JMIR Aging

Journal Impact Factor (JIF) (2023): 5.0
Volume 8 (2025) ISSN 2561-7605 Editor in Chief: Yun Jiang, PhD, MS, RN, FAMIA; Jinjiao Wang, PhD, RN, MPhil

Contents

Reviews

Effectiveness of Gamification on Enjoyment and Satisfaction in Older Adults: Systematic Review and Meta-Analysis (e72559)
Javier Bravo-Aparicio, Iria Trillo-Charlín, Juan Avendaño-Coy, Hector Beltran-Alacreu. 17

Effect of Physical Exercise on Telomere Length: Umbrella Review and Meta-Analysis (e64539)
Juan Sánchez-González, Juan Sánchez-Rodríguez, Rogelio González-Sarmiento, Víctor Navarro-López, Raúl Juárez-Vela, Jesús Pérez, Javier Martín-Vallejo. 157

Values of Stakeholders Involved in Applying Surveillance Technology for People With Dementia in Nursing Homes: Scoping Review (e64074)
Daniëlle van Gaans-Riteco, Annerieke Stoop, Eveline Wouters. 297

Evaluating the User Experience and Usability of Game-Based Cognitive Assessments for Older People: Systematic Review (e65252)
Rhys Mantell, Ye Hwang, Matthew Dark, Kylie Radford, Michael Kasumovic, Lauren Monds, Peter Schofield, Tony Butler, Adrienne Withall. . .
1 1 1 3

Patient-Related Barriers to Digital Technology Adoption in Alzheimer Disease: Systematic Review (e64324)
Andrea Panzavolta, Andrea Arighi, Emanuele Guido, Luigi Lavorgna, Francesco Di Lorenzo, Alessandra Dodich, Chiara Cerami. 1352

Ethics and Equity Challenges in Telerehabilitation for Older Adults: Rapid Review (e69660)
Mirella Veras, Louis-Pierre Auger, Jennifer Sigouin, Nahid Gheidari, Michelle Nelson, William Miller, Anne Hudon, Dahlia Kairy. 1377

Best Practices for Implementing Electronic Care Records in Adult Social Care: Rapid Scoping Review (e60107)
Martha Snow, Wagner Silva-Ribeiro, Mary Baginsky, Sonya Di Giorgio, Nicola Farrelly, Cath Larkins, Karen Poole, Nicole Steils, Joanne Westwood, Juliette Malley. 1392

The Impact of Digital Technology on the Physical Health of Older Workers: Scoping Review (e78406)
Jeroen Spijker, Hande Barlın, Melina Dritsaki, Yang Gu, Aija Klavina, Nilufer Korkmaz Yaylagul, Gunilla Kulla, Murat Mercan, Eda Orhun, Anna Sevcikova, Brigid Unim, Gunay Yildizer, Cristina Tofan. 1478

Technology-Assisted Physical Activity Interventions for Older People in Their Home-Based Environment: Scoping Review (e65746)
Rosemary Dubbeldam, Rafal Stemplewski, Iuliia Pavlova, Magdalena Cyma-Wejchenig, Sunwoo Lee, Patrick Esser, Ellen Bentlage, Veysel Alcan, Özge Çevik, Eleni Epiphaniou, Francesca Gallè, Antoine Langeard, Simone Gafner, Mona Ahmed, Niharika Bandaru, Arzu Erden Güner, Evrim Göz, Ilke Kara, Ay e Kabuk, Ilayda Türkoglu, Zada Pajalic, Jan Vindiš, Damjan Jaksic, U ur Verep, Ioanna Chouvarda, Vera Simovska, Yael Netz, Jana Pelclova. 2223

XSL•FO
RenderX

Using Artificial Intelligence–Based Technologies for the Early Detection of Behavioral and Psychological Symptoms of Dementia: Scoping Review (e76074)
Sofia Fernandes, Joëlle Rosselet Amoussou, Carla Gomes da Rocha, Elodie Perruchoud, Armin von Gunten, Cédric Mabire, Henk Verloo. . . 2 2 4 7

Effectiveness of mHealth-Based Gamified Interventions on Physical Activity in Older Adults: Systematic Review (e78686)
Lin Chen, Fenglin Jang, Min Li, Wei Zong, Huiqin Yu. . . . 2267

Examining Technology Perspectives of Older Adults With Mild Cognitive Impairment: Scoping Review (e78229)
Snezna Schmidt, Stephen Isbel, Blooma John, Ramanathan Subramanian, Nathan D'Cunha. . . . 2281

Patient and Carer-Related Facilitators and Barriers to the Adoption of Assistive Technologies for the Care of Older Adults: Systematic Review (e73917)
Stephen Malden, Kris McGill, Bruce Guthrie, Helen Frost, Susan Shenkin, Adanna Ezike, Bethany Bareham, Stewart Mercer, Caroline Pearce, Cara Wilson, Ian Underwood, John Vines, Sue Lewis, Amy O'Donnell. . . . 2306

Applications of Machine Learning for Cognitive Health in Older Individuals With HIV: Rapid Systematic Review (e80433)
Hwayoung Cho, Jiyoun Song, Hannah Cho, Lin Li, Renjie Liang, Railton Miranda, Qianqian Song, Jiang Bian. . . . 2326

Exploring Literature on Data Governance in the Health Care of Older Persons: Scoping Review (e73625)
Angely Garcia, Ayra Balingbing, Yves Palad. . . . 2337

Nonpharmacological Multimodal Interventions for Cognitive Functions in Older Adults With Mild Cognitive Impairment: Scoping Review (e70291)
Raffy Chan, Joson Zhou, Yuan Cao, Kenneth Lo, Peter Ng, David Shum, Arnold Wong. . . . 2354

Advancing Remote Monitoring for Patients With Alzheimer Disease and Related Dementias: Systematic Review (e69175)
Mohmmad Shaik, Fahim Anik, Md. Hasan, Sumit Chakravarty, Mary Ramos, Mohammad Rahman, Sheikh Ahamed, Nazmus Sakib. . . . 2380

Enhancing Enrollment and Adherence in Long-Term Wearable Research on Dementia: Qualitative Systematic Review and Meta-Synthesis (e63768)
Colleen Peterson, Renée St Louis, Carol Flannagan. . . . 2399

Barriers to and Facilitators of Digital Health Technology Adoption Among Older Adults With Chronic Diseases: Updated Systematic Review (e80000)
Jennifer Hepburn, Lynn Williams, Lisa McCann. . . . 2416

Comparative Diagnostic Accuracy of AI-Assisted Fluorine-18 Fluorodeoxyglucose Positron Emission Tomography Versus Structural Magnetic Resonance Imaging in Alzheimer Disease: Systematic Review and Meta-Analysis (e76981)
Bingbing Wang, Tailiang Zhao, Rongrong Ma, Xiaochuan Huo, Xiaoxiao Xiong, Minjie Wu, Yuran Wang, Liu Liu, Zhijiang Zhuang, Bin Wang, Jixin Shou. . . . 2458

Ethical and Quality of Care–Related Challenges of Digital Health Twins in Care Settings for Older Adults: Scoping Review (e73925)
Md Jabin, Ayesha Mirza, Adaobi Ilodibe, Tillal Eldabi, Emilia Vann Yaroson. . . . 2697

XSL•FO
RenderX

Original Papers

Frailty, Fitness, and Quality of Life Outcomes of a Healthy and Productive Aging Program (GrandMove) for Older Adults With Frailty or Prefrailty: Cluster Randomized Controlled Trial (e65636)
Jennifer Tang, Hao Luo, Michael Tse, Joseph Kwan, Angela Leung, Teresa Tsien Wong, Terry Lum, Gloria Wong. 30

Principal Component Analysis of Stair Negotiation and Floor Transition Kinematics in Older Adults With and Without Functional Disability: Cross-Sectional Study (e71530)
Juliana Moreira, Ivone Teles, Bruno Cunha, José Félix, Diana Guedes, Leonel Alves, Rubim Santos, Andreia Sousa. 45

Effectiveness of Iso-Inertial Resistance Training on Muscle Power in Middle-Older Adults: Randomized Controlled Trial (e66414)
Aïda Cadellans Arroniz, Daniel Romero Rodríguez, Víctor Zárate, Flora Dantony, Marc Madruga Parera, Silvia Ortega Cebrian, David Blanco. 6 2

Effectiveness of Walking Prescription Using Mobile Health Technology on the Changes in Daily Steps in Older Adults With Cognitive Impairment: Randomized Controlled Study (e63081)
Hee Kim, Yun Hwang, Jee Park, Dong Lee. 77

Association of a Healthy Lifestyle With All-Cause and Cause-Specific Mortality Among Individuals With Probable Sarcopenia: Population-Based Cohort Study (e65374)
Ning Wang, Yuqing Zhang, Junqing Xie, Na Lu, Aojie Zheng, Changjun Li, Jie Wei, Chao Zeng, Guanghua Lei, Yilun Wang. 90

A 12-Month Digital Peer-Supported App Intervention to Promote Physical Activity Among Community-Dwelling Older Adults: Follow-Up Study of a Nonrandomized Controlled Trial (e66610)
Kento Tabira, Yuko Oguma, Shota Yoshihara, Megumi Shibuya, Manabu Nakamura, Natsue Doihara, Akihiro Hirata, Tomoki Manabe, Takashi Yamashita. 106

Impact of a Light Volleyball Intervention Program on Improving Physical Attributes of Older Adults in Hong Kong: Preliminary Study of a Randomized Controlled Trial (e62886)
Ka Leung, Yuchen Shi. 123

Sexual Response Problems and Their Correlates Among Older Adults From the Sexual Well-Being (SWELL) Study in China: Multicenter Cross-Sectional Study (e66772)
Bingyu Liang, Chen Xu, Bingyi Wang, Xinyi Li, Xin Peng, Ying Wang, Hui Li, Yong Lu, Xiaopei Shen, Lin Ouyang, Guohui Wu, Maohe Yu, Jiewei Liu, Xiaojun Meng, Yong Cai, Huachun Zou. 133

Older Adults' Perspectives on Participating in a Synchronous Online Exercise Program: Qualitative Study (e66473)
Giulia Coletta, Kenneth Noguchi, Kayla Beaudoin, Angelica McQuarrie, Ada Tang, Rebecca Ganann, Stuart Phillips, Meridith Griffin. 146

Exploring the Feasibility of a 5-Week mHealth Intervention to Enhance Physical Activity and an Active, Healthy Lifestyle in Community-Dwelling Older Adults: Mixed Methods Study (e63348)
Kim Daniels, Sharona Vonck, Jolien Robijns, Kirsten Quadflieg, Jochen Bergs, Annemie Spooren, Dominique Hansen, Bruno Bonnechère. . . 1 7 5

Voice-Based Remote Care Program for Vulnerable Older Adults in a Rural Community: Single-Arm Pilot Clinical Study (e76653)
Geon Jang, Sunghwan Ji, Ji Baek, Eunju Lee, Seungryong Chong, Il-Young Jang. 203

Evaluating the Effectiveness of Digital Social Robots in Reducing Loneliness Among Community-Dwelling Older Adults in Japan: Randomized Controlled Trial and Qualitative Analysis (e74422)
Hiroshi Murayama, Mai Takase. 215

Adapting the Technology Acceptance Model to Examine the Use of Information Communication Technologies and Loneliness Among Low-Income, Older Asian Americans: Cross-Sectional Survey Analysis (e63856)
Pauline DeLange Martinez, Daniel Tancredi, Misha Pavel, Lorena Garcia, Heather Young. 227

XSL•FO
RenderX

Implementation of an Electronic Medication Management System in 41 Residential Care Homes in Hong Kong: Pre–Post Interventional Study (e79262)
Phillip Au-Doung, Ho Chau, Chit Wong, Kai Lau, Yi Chan, Ho Yip, Chiu Hung, Bik-Wai Tai, Teddy Lam, Chui Lee, Sau Chiang, Yin Cheung. . . 2 4 5

The Use of Technology to Deliver In-Home Aged Care Services: Mixed Methods Study of Australian Staff Perspectives (e76141)
Baldwin Kwan, Marissa Dickins, Sue Williams, Frances Batchelor, Kerry Hwang, Kate Fulford, Tony Walsh, Helena Jakupovic, Tanya Davison. 2 6 4

Identifying Food Preferences and Malnutrition in Older Adults in Care Homes: Co-Design Study of a Digital Nutrition Assessment Tool (e64661)
Jenni Connelly, Kevin Swingler, Nidia Rodriguez-Sanchez, Anna Whittaker. 284

Barriers to and Facilitators of Implementing Overnight Nursing Teleconsultation in Small, Rural Long-Term Care Facilities: Qualitative Interview Study (e71950)
Veronique Nabelsi, Véronique Plouffe, Marie Leclerc. 315

Intraoperative Hypotension and Major Adverse Cardiac Events Among Older Adult Patients Undergoing Noncardiac Surgery: Retrospective Cohort Study (e67177)
Kai Zhang, Chang Liu, Meng Wang, Ting Zhang, Bingbing Meng, Siyi Yao, Jingsheng Lou, Qiang Fu, Yanhong Liu, Jiangbei Cao, Lulong Bo, Weidong Mi, Hao Li. 337

Sleep Disturbance as a Catalyst in the Cyclical Link Between Depressive Symptoms and Disability in Instrumental Activities of Daily Living in Older Chinese Adults: Longitudinal Cohort Study (e76643)
Hao Wu, Lu Wang, Xin Liu, Yuyue Yang, Wei Gao. 349

Development and Validation of the Healthy Longevity Index for Personalized Healthy Aging in Primary Care: Cross-National Retrospective Analysis (e80034)
Hsi-Yu Lai, Shu Zhang, Rei Otsuka, Shih-Tsung Huang, Hidenori Arai, Fei-Yuan Hsiao, Liang-Kung Chen. 365

Real-World Insights Into Dementia Diagnosis Trajectory and Clinical Practice Patterns Unveiled by Natural Language Processing: Development and Usability Study (e65221)
Hunki Paek, Richard Fortinsky, Kyeryoung Lee, Liang-Chin Huang, Yazeed Maghaydah, George Kuchel, Xiaoyan Wang. 381

Development of a Longitudinal Model for Disability Prediction in Older Adults in China: Analysis of CHARLS Data (2015-2020) (e66723)
Jingjing Chu, Ying Li, Xinyi Wang, Qun Xu, Zherong Xu. 393

Factors Influencing the Severity of Urinary and Defecatory Dysfunction Among the Middle-Aged and Older Adult Chinese Population: Longitudinal Study of a 5-Wave Survey Cohort (e70541)
Haoran Zhu, Xiaoming Li, Liwei Jing, Jingli Kou, Lichao Gong, Rui Wang, Guangtian Liu, Chao Zhang, Jiayi Zhao, Jing Zhang. 411

The Role of Physical Activity and Physical Function in Predicting Physical Frailty Transitions in Chinese Older Adults: Longitudinal Observational Study From CHARLS (e75887)
Ziwei Zeng, Chun Hsu, Cindy Sit, Stephen Wong, Yijian Yang. 426

Predictive Validity of Hospital-Associated Complications of Older People Identified Using Diagnosis Procedure Combination Data From an Acute Care Hospital in Japan: Observational Study (e68267)
Seigo Mitsutake, Tatsuro Ishizaki, Shohei Yano, Takumi Hirata, Kae Ito, Ko Furuta, Yoshitomo Shimazaki, Hideki Ito, Alison Mudge, Kenji Toba. 442

XSL•FO
RenderX

Older Perpetrators of Domestic Violence: Mixed-Effects Logistic Regression Analysis of Police Records (e75993)
Sharon Reutens, Emaediong Akpanekpo, George Karystianis, Adrienne Withall, Tony Butler. 455

Associations Between Hearing Loss and Depressive Symptom Trajectories in Middle-Aged and Older People in China: Retrospective Analysis (e75545)
Yuchen Liu, Wenwen Liu, Jun Ma, Yangfan Chai, Guilan Kong. 471

Detecting Sleep/Wake Rhythm Disruption Related to Cognition in Older Adults With and Without Mild Cognitive Impairment Using the myRhythmWatch Platform: Feasibility and Correlation Study (e67294)
Caleb Jones, Rachel Wasilko, Gehui Zhang, Katie Stone, Swathi Gujral, Juleen Rodakowski, Stephen Smagula. 485

Identifying Optimal Wearable Devices for Monitoring Mobility in Hospitalized Older Adults: Feasibility, Acceptability, and Validity Study (e64372)
Paulo Nascimento, Renata Kirkwood, Lauren Griffith, Mylinh Duong, Cody Cooper, Yujiao Hao, Rong Zheng, Samir Raza, Marla Beauchamp.
4 9 4

Performance of a Digital Cognitive Assessment in Predicting Dementia Stages Delineated by the Dementia Severity Rating Scale: Retrospective Study (e65292)
Duong Huynh, Kevin Sun, Mary Patterson, Reza Hosseini Ghomi, Bin Huang. 511

Home-Based Exercise and Fall Prevention in Older Adults: Development, Validation and Usability of the Mais Equilíbrio Mobile App (e80724)
Mateus Leite, Alessandro Silva, Silvana Funghetto, Luciano de Lima, Samuel Mezavila Abdelmur, Hudson Pinheiro, Calliandra de Souza Silva, Maurílio Dutra, Marina Stival. 523

Risk Factors for Community-Dwelling Older Adults Dropping Out of Self-Guided, Remote, and Web-Based Longitudinal Research: Predictive Modeling of Data from the Web-LABrainS Platform (e68826)
Luke Braun, H Allen, Jeffrey Keller. 538

Multimodal Detection of Agitation in People With Dementia in Clinical Settings: Observational Pilot Study (e68156)
Abeer Badawi, Somayya Elmoghazy, Samira Choudhury, Sara Elgazzar, Khalid Elgazzar, Amer Burhan. 553

Usability and Impact of the Web-Based Dementia Foundations Educational Program in Personal Support Workers (PSWs), PSW Trainees, and Care Companions: Quasi-Experimental Study (e67889)
Anthony Levinson, Stephanie Ayers, Sandra Clark, Patricia Gerantonis, Amy Schneeberg, Richard Sztramko. 569

Implementation of New Technologies in an Aged Care Social Day Program: Mixed Methods Evaluation (e60297)
Dannielle Post, Kathleen Whitson, Gaynor Parfitt. 582

Unveiling the Frailty Spatial Patterns Among Chilean Older Persons by Exploring Sociodemographic and Urbanistic Influences Based on Geographic Information Systems: Cross-Sectional Study (e64254)
Yony Ormazábal, Diego Arauna, Juan Cantillana, Iván Palomo, Eduardo Fuentes, Carlos Mena. 593

A Comparison of Patient and Provider Perspectives on an Electronic Health Record–Based Discharge Communication Tool: Survey Study (e60506)
Dorothy Wang, Eliza Wong, Annie Cheung, Kam-Shing Tang, Eng-Kiong Yeoh. 606

Association of Subjective Cognitive Concerns With Performance on Mobile App–Based Cognitive Assessment in Cognitively Normal Older Adults: Observational Study (e64033)
Caroline Nester, Alyssa De Vito, Sarah Prieto, Zachary Kunicki, Jennifer Strenger, Karra Harrington, Nelson Roque, Martin Sliwinski, Laura Rabin, Louisa Thompson. 622

XSL•FO
RenderX

Effectiveness of a Dyadic Technology–Enhanced Home-Based Horticultural Therapy on Psychosocial Well-Being Among People With Dementia and Their Family Caregivers: Multimethods Pilot Study (e66017)
Patrick Kor, Justina Liu, Arkers Wong, Alex Tsang, Han Tan, Daphne Cheung, Humphrey Leung, Frances Wong. 638

Enhancing Older Adults' Lives Through Positive Aging Perception, Quality-of-Life Enhancement, and Social Support to Drive Acceptance and Readiness Toward Indoor Assistive Technology: Cross-Sectional Study (e59665)
Ka Wong, Pei-Lee Teh, Weng Lim, Shaun Lee. 653

A Self-Adaptive Serious Game to Improve Motor Learning Among Older Adults in Immersive Virtual Reality: Short-Term Longitudinal Pre-Post Study on Retention and Transfer (e64004)
Gauthier Everard, Louise Declerck, Thierry Lejeune, Martin Edwards, Justine Bogacki, Cléo Reiprich, Kelly Delvigne, Nicolas Legrain, Charles Batcho. 672

Socioculturally Appropriate Internet-Based Geriatric Care Model for Older Adults Living With HIV: Experience-Based Co-Design Approach (e67122)
Kristina Kokorelias, Marina Wasilewski, Dean Valentine, Andrew Eaton, Erica Dove, Paige Brown, Stuart McKinlay, Christine Sheppard, Esther Su, Hardeep Singh, Ashley Flanagan, Alice Zhabokritsky, Reham Abdelhalim, Rabea Parpia, Rahel Zewude, Laura Jamieson, Sharon Walmsley, Luxey Sirisegaram. 689

Longitudinal Remote Sleep and Cognitive Research in Older Adults With Mild Cognitive Impairment and Dementia: Prospective Feasibility Cohort Study (e72824)
Victoria Gabb, Jonathan Blackman, Hamish Morrison, Haoxuan Li, Adrian Kendrick, Nicholas Turner, Rosemary Greenwood, Bijetri Biswas, Amanda Heslegrave, Elizabeth Coulthard. 707

Staff Enablement of the Tovertafel for Enrichment in Residential Aged Care: Field Study (e67919)
Ryan Kelly, Asmita Manchha, Jenny Waycott, Rajna Ogrin, Judy Lowthian. 726

Detection of Vascular Mild Cognitive Impairment in Southeast Asia Using the Visual Cognitive Assessment Test: Machine Learning Analysis From the BIOCIS (Biomarkers and Cognition Study, Singapore) (e76847)
Pricilia Tanoto, Hannah Ye, Seyed Saffari, Yi Leow, Ashwati Vipin, Faith Lee, Smriti Ghildiyal, Shan Liew, Adnan Mohammed, Gurveen Sandhu, Kiirtaara Aravindhan, Gursimar Bhalla, Rasyiqah Shaik Mohamed Salim, Nagaendran Kandiah. 741

Using Smart Displays to Implement an eHealth System for Older Adults With Multiple Chronic Conditions: Randomized Controlled Trial (e75991)
Gina Landucci, David Gustafson Sr, Marie-Louise Mares, Klaren Pe-Romashko, John Curtin, Yaxin Hu, Adam Maus, Kasey Thompson, Sydney Saunders, Kaitlyn Brown, Judith Woodburn, Bilge Mutlu. 753

Engaging Older Adults and Staff in the Co-Design and Evaluation of Socially Assistive Robot and Virtual Reality Activities for Long-Term Care: User-Centered Study (e75288)
Ritam Ghosh, Nibraas Khan, Miroslava Migovich, Judith Tate, Cathy Maxwell, Paul Newhouse, Douglas Scharre, Alai Tan, Lorraine Mion, Nilanjan Sarkar. 773

A Hybrid Rule- and Large Language Model–Based Embodied Voice Assistant (GRACE) for Cognitive Stimulation in Older Adults: Usability Study Assessing Technical Feasibility, Technology Acceptance, and Working Alliance (e76489)
Rasita Vinay, Ekaterina Uetova, Nora Tommila, Nikola Biller-Andorno, Tobias Kowatsch. 792

An Accessible Multifunctional System to Support Safe and Independent Aging in Place: Iterative Development and Qualitative Analysis (e72579)
Céline Aldenhoven, Matthew Milton, Lisa D'Ambrosio, Chaiwoo Lee, Sophia Ashebir, Elisabeth André. 808

XSL•FO
RenderX

Exploring Technology Supporting Aging-in-Place Using an Equity Lens Through Focus Groups and World Café–Informed Research Agenda: Qualitative Study (e71093)
Marianne Saragosa, Ibukun-Oluwa Abejirinde, Evan MacEachern, Michelle Nelson, Kristina Kokorelias, Sidra Bharmal, Brina Ludwig Prout, Marian Mohamed . . . 825

The Caregiver Support Model for Informal Caregivers of Frail Older Adults: Randomized Controlled Trial (e71638)
Cyrus Leung, Kin-Kit Li, Dannii Yeung, Alice Chong, Marcus Chiu, Xin Guan, Tit Lo . . . 841

Supporting Dementia Caregiving With a Mobile Care Ecosystem: Development and Mixed Methods Study (e78759)
Chetna Malhotra, Yanzhen Yue, Chandrika Ramakrishnan, Shimoni Shah, Wenda Chen, Philip Yap, Chin Cheong, Irene Teo, Shiou-Liang Wee, Xiangming Lan, Yi Chen, Chee Chong, Xueying Huang, Ivy Chua . . . 858

Assessing the Time for Living and Caring (TLC) Study: Mixed-Methods Feasibility Study of a Web-Based Caregiver Intervention to Improve Respite (e71792)
Amber Thompson, Alexandra Terrill, Michael Caserta, Bob Wong, Eli Iacob, Catharine Sparks, Louisa Stark, Rebecca Utz . . . 872

Medicaid Education and Eligibility Planning for Caregivers: Website Usability and Validation Study (e77441)
Marguerite DeLiema, Siyu Gao, Justine Scattarelli, Kelly Moeller, Olu Olofinboba . . . 885

Culturally Adapted STAR-Caregivers Virtual Training and Follow-Up for Latino Caregivers of People Living With Dementia: Single-Arm Pre-Post Mixed Methods Study (e66053)
Miguel Mariscal, Celeste Garcia, Lily Zavala, Magaly Ramirez . . . 903

Identifying Unmet Needs of Informal Dementia Caregivers in Clinical Practice: User-Centered Development of a Digital Assessment Tool (e59942)
Olga Biernetzky, Jochen Thyrian, Melanie Boekholt, Matthias Berndt, Wolfgang Hoffmann, Stefan Teipel, Ingo Kilimann . . . 919

Uncovering Specific Navigation Patterns by Assessing User Engagement of People With Dementia and Family Caregivers With an Advance Care Planning Website: Quantitative Analysis of Web Log Data (e60652)
Charlèss Dupont, Tinne Smets, Courtney Potts, Fanny Monnet, Lara Pivodic, Aline De Vleminck, Chantal Van Audenhove, Maurice Mulvenna, Lieve Van den Block . . . 934

The PDC30 Chatbot—Development of a Psychoeducational Resource on Dementia Caregiving Among Family Caregivers: Mixed Methods Acceptability Study (e63715)
Sheung-Tak Cheng, Peter Ng . . . 946

Designing a Multimodal and Culturally Relevant Alzheimer Disease and Related Dementia Generative Artificial Intelligence Tool for Black American Informal Caregivers: Cognitive Walk-Through Usability Study (e60566)
Cristina Bosco, Ege Otenen, John Osorio Torres, Vivian Nguyen, Darshil Chheda, Xinran Peng, Nenette Jessup, Anna Himes, Bianca Cureton, Yvonne Lu, Carl Hill, Hugh Hendrie, Priscilla Barnes, Patrick Shih . . . 958

Assessment of Technology Readiness in Norwegian Older Adults With Long-Term Health Conditions Receiving Home Care Services: Cross-Sectional Questionnaire Study (e62936)
Sverre Bergh, J rat Benth, Lisbeth Høgset, Britt Rydjord, Lars Kayser . . . 972

Experiences of Older Mental Health Patients and Their Care Partners Using a Proxy Account to Access Open Notes: Qualitative Interview Study (e66690)
Eva Meier-Diedrich, Tobias Esch, Maria Hägglund, Martin Heinze, Stefan Hochwarter, Justin Speck, Marie Wagener, Volker Dahling, Julian Schwarz . . . 985

Employers' Perspectives of Caregiver-Friendly Workplace Policies for Caregiver-Employees Caring for Older Adults in Hong Kong: Thematic Analysis (e68061)
Maggie Lee, Eng-Kiong Yeoh, Eliza Wong . . . 1000

XSL•FO
RenderX

Creating User Personas to Represent the Needs of Dementia Caregivers Who Support Medication Management at Home: Persona Development and Qualitative Study (e63944)
Anna Jolliff, Priya Loganathar, Richard Holden, Anna Linden, Himalaya Patel, Jessica Lee, Aaron Ganci, Noll Campbell, Malaz Boustani, Nicole Werner. . . . 1017

Stressors Disclosed on Reddit by Caregivers of Older Adults and Social Support Received: Content Analysis (e71452)
Nova Huang, Bryan Timothy, Shirley Ho. . . . 1035

Development and Validation of a Rule-Based Natural Language Processing Algorithm to Identify Falls in Inpatient Records of Older Adults: Retrospective Analysis (e65195)
Xing Qian, Pui Chau, Daniel Fong, Mandy Ho, Jean Woo. . . . 1048

The Impact of Vision Impairment on Self-Reported Falls Among Older US Adults: Cross-Sectional and Longitudinal Study (e68771)
Kasem Seresirikachorn, Rachasak Somyanonthanakul, Matthew Johnson, Panisa Singhanetr, Jiraporn Gatedee, David Friedman, Nazlee Zebardast. . . . 1057

AI-Assisted Dynamic Postural Control Screening to Improve Functional Mobility in Older Adult Populations: Quasi-Experimental Study (e73290)
Kai-Chih Lin, Rong-Jong Wai, Hung-Yu Chang Chien. . . . 1072

Building Strong Foundations: Nonrandomized Interventional Study of a Novel, Digitally Delivered Fall Prevention Program for Older Adults (e68957)
David Wing, Jeanne Nichols, Hava Barkai, Olivia Culbert, Daniel Moreno, Michael Higgins, Anna O'Brien, Mariana Perez, Hope Davey, Ryan Moran. . . . 1083

Development of a Predictive Dashboard With Prescriptive Decision Support for Falls Prevention in Residential Aged Care: User-Centered Design Approach (e63609)
S Silva, Nasir Wabe, Amy Nguyen, Karla Seaman, Guogui Huang, Laura Dodds, Isabelle Meulenbroeks, Crisostomo Mercado, Johanna Westbrook. . . . 1097

Dynamic Interactions Among Sleep Duration, Cognitive Function, and Depressive Symptoms in Middle-Aged and Older Chinese Adults: Temporal Network Analysis From CHARLS (e76210)
Furong Chen, Jiaying Li, Junchen Guo, Ying Xiong, Zengjie Ye. . . . 1129

The Impact of Social Stress and Healthy Lifestyle on the Mortality of Chinese Older Adults: Prospective Cohort Study (e75942)
Jin Yang, Jilong Huang, Qingmei Huang, Jian Gao, Dan Liu, Zhihao Li, Yuebin Lv, Xiaoming Shi, Chen Mao. . . . 1142

Edentulousness and the Likelihood of Becoming a Centenarian: Longitudinal Observational Study (e68444)
Xindi Wei, Longfei Zhuang, Yuan Li, Junyu Shi, Yijie Yang, Hongchang Lai, Beilei Liu. . . . 1153

Association Between Age and Severity at Disability Onset and All-Cause Mortality: Longitudinal Observational Study From the Health and Retirement Study (e73254)
Anying Bai, Cuie Liu, Yu Jiang, Weihao Xu, Jian Cao. . . . 1162

Artificial Intelligence-Driven Biological Age Prediction Model Using Comprehensive Health Checkup Data: Development and Validation Study (e64473)
Chang-Uk Jeong, Jacob Leiby, Dokyoon Kim, Eun Choe. . . . 1174

Peer Volunteers' Journeys Through Training and Engagement in Older Adult Communities: Descriptive Qualitative Study (e71810)
Shefaly Shorey, Jeremy Seow, Rathi Mahendran, Geok Wee, Lee Goh, Ee Kua. . . . 1186

XSL•FO
RenderX

Knowledge, Attitudes, and Behaviors Related to Dementia Prevention and Caregiving Among Korean Americans (the KIMCHI Project): Pre- and Posttest Evaluation Study (e72147)
Van Ta Park, Bora Nam, Daren Huang, Stacy Yun, Nicole Phan, Eun Lee, Hye-Won Shin. 1199

Associations Between Dietary Habits and Accelerated Aging and the Establishment of an Accelerated Aging Interpretable Risk Prediction Model via Shapley Additive Explanations: Cross-Sectional Study From Two Representative Populations (e72020)
Zhengyang Wu, Ning Zhang, Haiwei Wang, Yang Yang, Houhao Shen, Qidi Zhang, Yunxia Cao, Yinan Du, Dongmei Ji. 1220

Digital Isolation and Depression Risk in Older Adults Using the National Health and Aging Trends Study Database: 8-Year Longitudinal Study (e75174)
Zhili He, Shijun Yang, Wei Tong, Wei Zhang. 1234

Technology Activities and Cognitive Trajectories Among Community-Dwelling Older Adults: National Health and Aging Trends Study (e77227)
Erh-Chi Hsu, Erin Spaulding, Eric Jutkowitz. 1248

Using Virtual Reality to Assess Spatial Navigation Ability in Individuals With Mild Cognitive Impairment and Older Adults: Cross-Sectional Study (e75952)
Ming-Chyi Pai, Yuh-Ting Lin, Chung-Yao Hsiao, Chia-Hung Lai, Chih-Jen Chen, Chaur-Jong Hu, Cheng-Yu Chen. 1259

The Critical Moderating Role of Cognitive Function in Digital Inclusion: Data Analysis Study on Depression Risk Among Older Adults (e80700)
Gang Xiao, Tingting Nie. 1274

Research Status and Trends in Virtual Reality Technology for Older Adults: Bibliometric and Visual Analysis (e76609)
Jing Xu, Wenjin Zhang, WenLi Liu. 1291

The Double-Edged Sword of Digital Engagement—How Digital Access and Internet Use Reshape Sleep Schedules and Underlying Mechanisms in Older Adults: Longitudinal Observational Study (e79731)
Chi Zhang, Longxuan Lin, Li Wang, Han Hu, Heyang Zhang. 1306

Aesthetic Experience in the Acceptance of Wearable Technology for People With Dementia: Critical Interpretive Synthesis (e72082)
Yixuan Wei, John Ratcliffe, Dag Aarsland, Wei Liu. 1322

Factors Associated With the Ability To Keep Up With Technology Developments: Findings From a National Multigenerational Cross-Sectional Survey in Sweden (e77930)
Jens Offerman, Sofi Fristedt, Steven Schmidt, Susanne Iwarsson. 1342

Frequency of Electronic Personal Health Record Use in US Older Adults: Cross-Sectional Study of a National Survey (e71460)
Lavlin Agrawal, Richelle DaSouza, Pavankumar Mulgund, Pankaj Chaudhary. 1364

Implementation of a Web-Based Program for Advance Care Planning and Evaluation of its Complexity With the Nonadoption, Abandonment, Scale-Up, Spread, And Sustainability (NASSS) Framework: Qualitative Evaluation Study (e49507)
Doris van der Smissen, Maud Schreijer, Lisette van Gemert-Pijnen, Rudolf Verdaasdonk, Agnes van der Heide, Ida Korfage, Judith Rietjens.
1 4 1 1

Determinants of Telehealth Adoption Among Older Adults: Cross-Sectional Survey Study (e60936)
Siow-Hooi Tan, Yee-Yann Yap, Siow-Kian Tan, Chee-Kuan Wong. 1426

XSL•FO
RenderX

Gender Differences in the Digital Divide, Digital Back-Feeding, and Health-Related Quality of Life Among Rural Older Adults: Cross-Sectional Study (e75925)
Xin Che, Shujun Chai, Dan Zhao, Shirong Chen, Chengchao Zhou. 1444

Digital Pathways to Reducing Depression Among Aging Populations Through the “Broadband China” Pilot Program: Quasi-Natural Experiment (e79260)
Zhiying Li. 1457

Association of Insulin Resistance, Sarcopenia, and Risk of Cardiovascular Disease: Findings From the China Health and Retirement Longitudinal Study (e80115)
Huilin Liu, Sijing Li, Ximin Zhang, Wenjie Long, Huili Liao, Lu Lu, Shihao Ni, Zhongqi Yang. 1492

Exploring Risk Factors Related to Low Calf Circumference in Older Adults With Multimorbidity: Cross-Sectional Latent Class Analysis (e68760)
Xilin Peng, Xudong Chen, Ruihao Zhou, Fanfan Shi, Tao Zhu, Guo Chen. 1509

Using Machine Learning to Predict Cognitive Decline in Older Adults From the Chinese Longitudinal Healthy Longevity Survey: Model Development and Validation Study (e67437)
Hao Ren, Yiying Zheng, Changjin Li, Fengshi Jing, Qiting Wang, Zeyu Luo, Dongxiao Li, Deyi Liang, Weiming Tang, Li Liu, Weibin Cheng. . .
1 5 2 2

The Relationship Between Sleep Disorders and Combination of Diabetes and Sarcopenia in Adults Aged 45 Years or Older: 10-Year Nationwide Prospective Cohort Study (e66372)
Shugang Li, Yimi Wang, Linzhi Li, Hao Wu. 1537

Sarcopenia and Risk of Cognitive Impairment: Cohort Study and Mendelian Randomization Analysis (e66031)
Tingting Sha, Yuqing Zhang, Jie Wei, Changjun Li, Chao Zeng, Guanghua Lei, Yilun Wang. 1551

Feasibility of a Cinematic–Virtual Reality Program Educating Health Professional Students About the Complexity of Geriatric Care: Pilot Pre-Post Study (e64633)
Elizabeth Beverly, Samuel Miller, Matthew Love, Carrie Love. 1565

Exploring Dance as a Therapeutic Approach for Parkinson Disease Through the Social Robotics for Active and Healthy Ageing (SI-Robotics): Results From a Technical Feasibility Study (e62930)
Roberta Bevilacqua, Elvira Maranesi, Marco Benadduci, Gabriella Cortellessa, Alessandro Umbrico, Francesca Fracasso, Giovanni Melone, Arianna Margaritini, Angela La Forgia, Pierpaolo Di Bitonto, Ada Potenza, Laura Fiorini, Carlo La Viola, Filippo Cavallo, Alessandro Leone, Andrea Caroppo, Gabriele Rescio, Mauro Marzorati, Amedeo Cesta, Giuseppe Pelliccioni, Giovanni Riccardi, Lorena Rossi. 1578

Development and Validation of a Predictive Model Based on Serum Silent Information Regulator 6 Levels in Chinese Older Adult Patients: Cross-Sectional Descriptive Study (e64374)
Yuzi You, Wei Liang, Yajie Zhao. 1592

Wearable Smartphone-Based Multisensory Feedback System for Torso Posture Correction: Iterative Design and Within-Subjects Study (e55455)
Amanda Pereira, Olibario Machado Neto, Valeria Elui, Maria Pimentel. 1611

Estimation of Machine Learning–Based Models to Predict Dementia Risk in Patients With Atherosclerotic Cardiovascular Diseases: UK Biobank Study (e64148)
Zhengsheng Gu, Shuang Liu, Huijuan Ma, Yifan Long, Xuehao Jiao, Xin Gao, Bingying Du, Xiaoying Bi, Xingjie Shi. 1633

XSL•FO
RenderX

Development and Validation of a Brain Aging Biomarker in Middle-Aged and Older Adults: Deep Learning Approach (e73004)
Zihan Li, Jun Li, Jiahui Li, Mengying Wang, Andi Xu, Yushu Huang, Qi Yu, Lingzhi Zhang, Yingjun Li, Zilin Li, Xifeng Wu, Jiajun Bu, Wenyuan Li. 1650

Changes in Sarcopenia Status and Subsequent Cardiovascular Outcomes: Prospective Cohort Study (e69860)
Yuanyue Zhu, Kan Wang, Zuolin Lu, Feika Li, Yu Xu, Linhui Shen, Yufang Bi, Weiguo Hu. 1673

The Dual Impact of Time and Content Exposure of Social Media on Diabetes Self-Management in Older Adults: Cross-Sectional Study (e67312)
Qingyuan Ye, Ruiyang Xu, Li Li, Meng Zhao, Shan Wang, Sijing Peng, Si Chen, Fatema Ahmed, Chen Wu, Kefang Wang. 1685

Impact of a Consumer e-Learning Course on Beliefs, Treatment Choices, and Outcomes Among People With Hip and Knee Osteoarthritis: Qualitative Interview Study (e80282)
Rachel Nelligan, Libby Spiers, Rana Hinman, Kim Bennell. 1698

Implementation of the German Clinical Practice Guideline for Multimorbidity Using a Digital Tool in Primary Care: Pilot Cluster Randomized Clinical Trial (e79767)
Julia Nothacker, Valentina Paucke, Susanne Lezius, Antonia Zapf, Dagmar Lühmann, Martin Scherer, Ingmar Schäfer. 1709

Associations Between Social Media Use and Anxiety and Depression Among Older Adults : Cross-Sectional Study (e71712)
Jiaoling Huang, Zhenxing Ge, Yijing Chu, Yuge Yan, Wei Zhang, Hong Liang, Yuqi Yang, Hui Wang. 1723

Online Communities as a Support System for Alzheimer Disease and Dementia Care: Large-Scale Exploratory Study (e68890)
Sidharth Kaliappan, Chunyu Liu, Yoshee Jain, Ravi Karkar, Koustuv Saha. 1736

Analyzing Disparity in Geographical Accessibility to Home Medical Care Using a Claims Database and Geographical Information System: Simulation Study (e70040)
Yasuhiro Morii, Yasuhiro Nakanishi, Yuichi Nishioka, Yukio Tsugihashi, Tatsuya Noda, Tomoya Myojin, Tomoaki Imamura, Manabu Akahane. . . 1751

Determinants of Having Online Health Consultations During the COVID-19 Pandemic Among Middle-Aged and Older Adults in Germany: Representative Longitudinal Survey Study (e60311)
Ariana Neumann, Hans-Helmut König, André Hajek. 1763

Revisits, Readmission, and Mortality From Emergency Department Admissions for Older Adults With Vague Presentations: Longitudinal Observational Study (e55929)
Sebastian Alvarez Avendano, Amy Cochran, Valerie Odeh Couvertier, Brian Patterson, Manish Shah, Gabriel Zayas-Caban. 1777

Passive Remote Monitoring Technologies' Influence on Home Care Clients' Ability to Stay Home: Multiprovincial Randomized Controlled Trial (e69107)
Lorie Donelle, Bradley Hiebert, Grace Warner, Michael Reid, Jennifer Reid, Salimah Shariff, Emily Richard, Sandra Regan, Lori Weeks, Kathleen Ledoux. 1788

Artificial Intelligence–Enhanced Multi-Algorithm R Shiny Application for Predictive Modeling and Analytics: Case Study of Alzheimer Disease Diagnostics (e70272)
Han Wenzheng, Edmund Agyemang, Sudesh Srivastav, Jeffrey Shaffer, Samuel Kakraba. 1809

Daily-Life Walking Characteristics of Older Adults in Relation to Age, Sex, and Physical Function: the HUNT4 Trondheim 70+ Observational Study (e75835)
Karoline Grønvik, Anisoara Paraschiv-Ionescu, Gro Tangen, Øyvind Salvesen, Jorunn Helbostad, Nina Skjæret-Maroni, Beatrix Vereijken. . . 1826

XSL•FO
RenderX

Implementing a Digital Physical Activity Intervention for Older Adults: Qualitative Study (e64953)
Cherish Boxall, Laura Dennison, Sascha Miller, Judith Joseph, Kate Morton, Jenny Corser, Joanna Kesten, Asgar Electicwala, Max Western, Stephen Lim, Chloe Grimmett, Lucy Yardley, Katherine Bradbury. . . . 1861

Exploring Smart Health Wearable Adoption Among Singaporean Older Adults Based on Self-Determination Theory: Web-Based Survey Study (e69008)
Hyunjin Kang, Tingting Yang, Nazira Banu, Sheryl Ng, Jeong Lee. . . . 1875

A Smartphone-Based Timed Up and Go Test Self-Assessment for Older Adults: Validity and Reliability Study (e67322)
Melissa Böttinger, Sabato Mellone, Jochen Klenk, Carl-Philipp Jansen, Marios Stefanakis, Elena Litz, Anastasia Bredenbrock, Jan-Philipp Fischer, Jürgen Bauer, Clemens Becker, Katharina Gordt-Oesterwind. . . . 1891

Perceptions of the Use of Mobile Apps to Assess Sleep-Dependent Memory in Older Adults With Subjective and Objective Cognitive Impairment: Focus Group Approach (e68147)
Aaron Lam, Simone Simonetti, Angela D'Rozario, David Ireland, DanaKai Bradford, Jurgen Fripp, Sharon Naismith. . . . 1904

Relationship Between Within-Session Digital Motor Skill Acquisition and Alzheimer Disease Risk Factors Among the MindCrowd Cohort: Cross-Sectional Descriptive Study (e67298)
Andrew Hooyman, Matt Huentelman, Matt De Both, Lee Ryan, Kevin Duff, Sydney Schaefer. . . . 1916

Implications of Public Disclosure of Personal Information in a Mobile Alert App for People Living With Dementia Who Go Missing: Qualitative Descriptive Study (e64847)
Adebusola Adekoya, Christine Daum, Noelannah Neubauer, Antonio Miguel-Cruz, Lili Liu. . . . 1928

Transcultural Adaptation, Validation, Psychometric Analysis, and Interpretation of the 22-Item Thai Senior Technology Acceptance Model for Mobile Health Apps: Cross-Sectional Study (e60156)
Nida Buawangpong, Penprapa Siviroj, Kanokporn Pinyopornpanish, Wachiranun Sirikul. . . . 1940

Developing a Dyadic Immersive Virtual Environment Technology Intervention for Persons Living With Dementia and Their Caregivers: Multiphasic User-Centered Design Study (e66212)
Elizabeth Rochon, Ayush Thacker, Mirelle Phillips, Christine Ritchie, Ana-Maria Vranceanu, Evan Plys. . . . 1958

Design of a Mobile App and a Clinical Trial Management System for Cognitive Health and Dementia Risk Reduction: User-Centered Design Approach (e66660)
Hannes Hilberger, Bianca Buchgraber-Schnalzer, Simone Huber, Theresa Weitlaner, Markus Bödenler, Alara Abaci, Jeroen Bruinsma, Ana Diaz, Anna Guazzarini, Jenni Lehtisalo, Seungjune Lee, Vasileios Loukas, Francesca Mangialasche, Patrizia Mecocci, Tiia Ngandu, Anna Rosenberg, Elisabeth Stögmann, Konsta Valkonen, Elena Uhlik, Helena Untersteiner, Laura Kneß, Helmut Ahammer, Sten Hanke. . . . 1975

A Digital Serious Game (Anticip'action) to Support Advance Care Planning Discussions in the General Population: Usability Study (e73378)
Dafne Campioni, Frederic Ehrler, Antoine Berger, Christine Clavien. . . . 2001

Evaluating the Feasibility of a Dyadic, Touch-Based Multimedia Tablet Intervention and Its Effects on the Caregiver-Patient Relationship Among Individuals With Mild Cognitive Impairment: Qualitative Triangulation Study (e75189)
Alexa von Bosse, Anna Thum, Sonja Mahler, Henriette Schulz, Paulina Weuthen, Christophe Kunze. . . . 2024

Design and Evaluation of a Mobile App for Intergenerational Communication: User-Centered Participatory Design and Experimental Mixed Methods Study (e75950)
Soondool Chung, Hannah Lee, Jeehye Jung. . . . 2039

XSL•FO
RenderX

An mHealth App (NeoMayor) to Promote Healthy Lifestyles and Brain Health in Older Adults: Design and Validation Study (e71936)
Myriam Gutiérrez, Victoria Cabello, Carol SanMartín, Jorge Mauro, Gada Musa, Melissa Torres, Maria San Martin, Carlos Márquez, Rodrigo Beltrán, Nicole Rogers, Felipe Salech, Daniela Ponce, Christian Ampuero, Pia Varas, Bastián Gamboa-Labbé, Víctor Cortés, Carlos Vega, Rocío Ruiz, Juan Velásquez, Rodrigo Vergara, Maria Behrens, Jamileth More, Carolina Delgado Derio. 2056

End User and Primary Care Physicians' Perspectives on Digital Innovations in Dementia Risk Detection: Focus on a Digital Sleep Biomarker (e74307)
Ríona Mc Ardle, Marie Poole, Sophie Horrocks, Josh King-Robson, Jonathan Schott, David Sharp, Matthew Harrison, Louise Robinson. 2072

Associations of Wearable-Measured Sleep and Physical Activity With Memory Performance in Older Adults: Cross-Sectional Study With Actigraphy and MRI (e80584)
Geng-Hao Liu, Yueh-Hsiang Huang, Tzu-Chiao Yuan, Yueh-Peng Chen, Jung-Lung Hsu, Shwu-Hua Lee, Chih-Ming Lin, Cheng-Hong Toh, Ji-Tseng Fang, Shih-Wei Lin, Li-Pang Chuang, Ning-Hung Chen. 2092

Web-Based Assessment of Cognition and Dementia Risk Factors in Over 3000 Norwegian Adults Aged 50 Years and Older: Cross-Sectional PROTECT Norge Study (e69867)
Ingelin Testad, Jon Aakre, Martha Gjestsen, Clive Ballard, Anne Corbett, Dag Aarsland, Ellie Pickering, Anastasia Ushakova. 2111

Network Analysis of Key Instrumental Activities of Daily Living and Cognitive Domains for Targeted Intervention in US Older Adults Without Dementia: Cross-Sectional Study (e67632)
Jiaying Li, Rendong He, Erh-Chi Hsu, Junxin Li. 2120

Online Group–Based Dual-Task Training to Improve Cognitive Function of Community-Dwelling Older Adults: Randomized Controlled Feasibility Study (e67267)
Pui Chau, Denise Cheung, Jojo Kwok, Wai Chan, Doris Yu. 2134

Predicting Postoperative Delirium in Older Patients Before Elective Surgery: Multicenter Retrospective Cohort Study (e67958)
Shun-Chin Wu, Nitin Sharma, Anne Bauch, Hao-Chun Yang, Jasmine Hect, Christine Thomas, Sören Wagner, Bernd Förstner, Christine von Arnim, Tobias Kaufmann, Gerhard Eschweiler, Thomas Wolfers, PAWEL Study Group. 2150

Cognitive Functioning Trajectories and Their Association With Mental Health in Older Adults With Hypertension: Secondary Analysis of the Chinese Longitudinal Healthy Longevity Survey Data (e74916)
Peiyun Wu, Cheng Lin, Jing Yang, Zhizhuo Wang. 2162

Comparison of 3 Aging Metrics in Dual Declines to Capture All-Cause Dementia and Mortality Risk: Cohort Study (e66104)
Anying Bai, Shan He, Yu Jiang, Weihao Xu, Zhanyi Lin. 2173

Building Consensus on the Relevant Criteria to Screen for Depressive Symptoms Among Near-Centenarians and Centenarians: Modified e-Delphi Study (e64352)
Carla Gomes da Rocha, Armin von Gunten, Pierre Vandel, Daniela Jopp, Olga Ribeiro, Henk Verloo. 2186

Cyberchondria in Older Adults and Its Relationship With Cognitive Fusion, Health-Related Quality of Life, and Mental Well-Being: Mediation Analysis (e70302)
Richard Xu, Vladan Starcevic. 2209

Evaluation of an AI-Based Clinical Decision Support System for Perioperative Care of Older Patients: Ethical Analysis of Focus Groups With Older Adults (e71568)
Nina Parchmann, Marcin Orzechowski, Simone Brefka, Florian Steger. 2448

Toward Data-Informed Care in Long-Term Care: Qualitative Analysis (e69423)
Suleyman Bouchmal, Katya Sion, Jan Hamers, Sil Aarts. 2478

XSL•FO
RenderX

Machine Learning for Predicting Postoperative Functional Disability and Mortality Among Older Patients With Cancer: Retrospective Cohort Study (e65898)
Yuki Hashimoto, Norihiko Inoue, Takuaki Tani, Shinobu Imai. . . . 2492

Unsupervised Deep Learning of Electronic Health Records to Characterize Heterogeneity Across Alzheimer Disease and Related Dementias: Cross-Sectional Study (e65178)
Matthew West, You Cheng, Yingnan He, Yu Leng, Colin Magdamo, Bradley Hyman, John Dickson, Alberto Serrano-Pozo, Deborah Blacker, Sudeshna Das. . . . 2509

Reconstruction of a Risk Analysis Tool That Uses Nursing Data for Frailty Assessment Among Older Adults: Derivation and Internal Evaluation Study (e67888)
Brian Celso, Sarah Ser, Urszula Snigurska, Laurence Solberg, Laurie Duckworth, Elizabeth Dunn, Mattia Prosperi, Robert Lucero, Ragnhildur Bjarnadottir. . . . 2536

Insights Into Associations Between Oral and General Health Outcomes of Nursing Home Residents Based on InterRAI Data: Cross-Sectional Study (e72308)
Emilie Schoebrechts, Johanna de Almeida Mello, Patricia Vandenbulcke, Hein van Hout, Jan De Lepeleire, Anja Declercq, Dominique Declerck, Joke Duyck. . . . 2551

Urban Quality and Biochemical, Hematological, and Nutritional Markers in Older Adults: Cross-Sectional Geospatial Study (e74313)
Carlos Mena, Yony Ormazabal, Nacim Molina, Eduardo Fuentes, Juan Cantillana, Victoria Villalobos, Moises Sandoval, Ivan Palomo, Diego Arauna. . . . 2569

The Dynamic Associations of Social and Intellectual Activity With Frailty Trajectory in Middle-Aged and Older Adults in China: Nationwide Longitudinal Study (e80152)
Fei Xu, Xinlei Miao, Shuang Liu, Yangxuan He, Meng Li, Jiayi Deng, Yilin Zhu, Jingshan Jiang, Song Leng. . . . 2584

Transitions in Sarcopenia Status and Cognitive Trajectories Among Middle-Aged and Older Adults in China: Longitudinal Cohort Study (e78277)
Chun Luo, Hao Wu, Xiaoying Shen, Shuang Han, Lv Lin, Bingyang Liu. . . . 2597

Machine Learning Models for Frailty Classification of Older Adults in Northern Thailand: Model Development and Validation Study (e62942)
Natthanaphop Isaradech, Wachiranun Sirikul, Nida Buawangpong, Penprapa Siviroj, Amornphat Kitro. . . . 2611

Association Between Sleep Duration and Cognitive Frailty in Older Chinese Adults: Prospective Cohort Study (e65183)
Ruixue Cai, Jianqian Chao, Chenlu Gao, Lei Gao, Kun Hu, Peng Li. . . . 2627

Identification of Target Body Composition Parameters by Dual-Energy X-Ray Absorptiometry, Bioelectrical Impedance, and Ultrasonography to Detect Older Adults With Frailty and Prefrailty Status Using a Mobile App in Primary Care Services: Descriptive Cross-Sectional Study (e67982)
Beatriz Ortiz-Navarro, José Losa-Reyna, Veronica Mihaiescu-Ion, Jerónimo Garcia-Romero, Margarita Carrillo de Albornoz-Gil, Alejandro Galán-Mercant. . . . 2640

Association Between the Frailty Index Based on Laboratory Tests and All-Cause Mortality in Hospitalized Older Adults: Retrospective Cohort Study (e70204)
Eyal Pasternak, Tamar Freud, Yan Press. . . . 2658

Handwriting in Mild Cognitive Impairment: Reliability Assessment and Machine Learning–Based Screening (e73074)
Simone Toffoli, Carlo Abbate, Francesca Lunardini, Edoardo Corno, Nicholas Diani, Alessia Gallucci, Emanuele Tomasini, Pietro Trimarchi, Simona Ferrante. . . . 2672

XSL•FO
RenderX

Using Wearable Sensors to Measure and Predict Personal Circadian Lighting Exposure in Nursing Home Residents: Model Development and Validation (e72338)
Shevvaa Beiglary, Yanxiao Feng, Nan Wang, Neda Ghaeili, Ying-Ling Jao, Yo-Jen Liao, Yuxin Li, Julian Wang. 2712

Using Deep Learning to Perform Automatic Quantitative Measurement of Masseter and Tongue Muscles in Persons With Dementia: Cross-Sectional Study (e63686)
Mahdi Imani, Miguel Borda, Sara Vogrin, Erik Meijering, Dag Aarsland, Gustavo Duque. 2734

Exploring Older Adults' Perspectives and Acceptance of AI-Driven Health Technologies: Qualitative Study (e66778)
Arkers Wong, Jessica Lee, Yue Zhao, Qi Lu, Shulan Yang, Vivian Hui. 2746

Development and Feasibility Study of HOPE Model for Prediction of Depression Among Older Adults Using Wi-Fi-based Motion Sensor Data: Machine Learning Study (e67715)
Shayan Nejadshamsi, Vania Karami, Negar Ghourchian, Narges Armanfard, Howard Bergman, Roland Grad, Machelle Wilchesky, Vladimir Khanassov, Isabelle Vedel, Samira Abbasgholizadeh Rahimi. 2758

Model-Based Feature Extraction and Classification for Parkinson Disease Screening Using Gait Analysis: Development and Validation Study (e65629)
Ming Lim, Tee Connie, Michael Goh, Nor Saedon. 2777

Identifying Deprescribing Opportunities With Large Language Models in Older Adults: Retrospective Cohort Study (e69504)
Vimig Socrates, Donald Wright, Thomas Huang, Soraya Fereydooni, Christine Dien, Ling Chi, Jesse Albano, Brian Patterson, Naga Sasidhar Kanaparthy, Catherine Wright, Andrew Loza, David Chartash, Mark Iscoe, Richard Taylor. 2806

Machine Learning Approach for Frailty Detection in Long-Term Care Using Accelerometer-Measured Gait and Daily Physical Activity: Model Development and Validation Study (e77140)
Xiaoping Zheng, Ziwei Zeng, Kimberley S van Schooten, Yijian Yang. 2821

Representations of Older Adults' Digital Literacy in Canadian News Media: Critical Discourse Analysis Using Unified Theory of Acceptance and Use of Technology 2 (e69373)
Alixe Ménard, Alexandra Maharaj, Sophie Harb, Sarah Fraser, Tracey O'Sullivan. 2834

Exploring Older Adult's Views of the Age-Inclusivity of Physical Activity Websites Using the Think Aloud Method: Qualitative Analysis (e68951)
Veda Clemson, Elisabeth Grey, Julie Barnett, Ella Burfitt, Fiona Gillison. 2848

Factors Influencing Older Adults' Perception of the Age-Friendliness of Their Environment and the Impact of Loneliness, Technology Use, and Mobility: Quantitative Analysis (e67242)
Eric Balki, Niall Hayes, Carol Holland. 2859

Buffering Effects of Internet Use on Caregiving-Related Health Impacts and Loneliness Among Older Informal Caregivers in California: Cross-Sectional Study (e74209)
Xiang Qi, Ruotong Liu, Eunjung Ko, Yaolin Pei, Bei Wu. 2872

Comparative Analysis of 16 Aging Concepts and Their Influence on Aging Narratives: Bibliometric and Content Analysis (e72011)
Na Xiao, Bo Xia, Laurie Buys, Connie Susilawati, Martin Larbi, Qing Chen. 2885

The Experience and Impact of Digital Technologies on Indigenous Populations in New Zealand During the COVID-19 Pandemic and Cyclone Gabrielle: The Kaupapa M ori Methodology (e73974)
Dianne Wepa, Shiji Thomas, Md Jabin. 2911

XSL•FO
RenderX

Digital Literacy and Its Association With Subjective Health Status and Healthy Lifestyle Behaviors Among Korean Older Adults: Cross-Sectional Study (e64974)
Soon Lee, Yejin Kim, Bomgyeol Kim, Sang Lee, Suk-Yong Jang, Tae Kim. . . . 2925

Selection, Optimization, and Compensation Strategies Used by Older Adults to Live Well With Technology: Qualitative Study (e75019)
Wei Koh, Kristiana Ludlow, Jacki Liddle, Nancy Pachana. . . . 2937

Free-Text Responses in a Nationally Representative Experimental Survey about End-of-Life Care Choices: ChatGPT-4o-Assisted Qualitative Analytical Study (e76335)
Elizabeth Goldberg, Mario Macis, Megan Bounds, Jonathan Picazo, Lauren Nicholas. . . . 2948

Viewpoints

Digital Measurement of Subjective Experiences in Alzheimer Disease and Related Dementias (AD/ADRD) (e71920)
Colin Depp, Jason Holden, Eric Granholm. . . . 2432

A Digitally Capable Aged Care Workforce: Demands and Directions for Workforce Education and Development (e54143)
Kathleen Gray, Kerryn Butler-Henderson, Karen Day. . . . 2437

Advancing Emergency Care With Digital Twins (e71777)
Haoran Li, Jingya Zhang, Ning Zhang, Bin Zhu. . . . 2442

Corrigenda and Addendas

Correction: Machine Learning Models for Frailty Classification of Older Adults in Northern Thailand: Model Development and Validation Study (e75690)
Natthanaphop Isaradech, Wachiranun Sirikul, Nida Buawangpong, Penprapa Siviroj, Amornphat Kitro. . . . 2528

Correction: Internet-Based Supportive Interventions for Family Caregivers of People With Dementia: Randomized Controlled Trial (e69493)
Yanhong Xie, Shanshan Shen, Caixia Liu, Hong Hong, Huilan Guan, Jingmei Zhang, Wanqi Yu. . . . 2530

Correction: Changes in Sarcopenia Status and Subsequent Cardiovascular Outcomes: Prospective Cohort Study (e84642)
Yuanyue Zhu, Kan Wang, Zuolin Lu, Feika Li, Yu Xu, Linhui Shen, Yufang Bi, Weiguo Hu. . . . 2532

Correction: Machine Learning Approach for Frailty Detection in Long-Term Care Using Accelerometer-Measured Gait and Daily Physical Activity: Model Development and Validation Study (e85173)
Xiaoping Zheng, Ziwei Zeng, Kimberley van Schooten, Yijian Yang. . . . 2534

XSL•FO
RenderX

Effectiveness of Gamification on Enjoyment and Satisfaction in Older Adults: Systematic Review and Meta-Analysis

Javier Bravo-Aparicio[1,2*], MSc; Iria Trillo-Charlín[1,2*], MSc; Juan Avendaño-Coy[1,2*], PhD; Hector Beltran-Alacreu[1,2*], PhD

[1]Toledo Physiotherapy Research Group, Faculty of Physiotherapy and Nursing of Toledo, Universidad de Castilla-La Mancha, Av Carlos III sn, Toledo, Spain

[2]Toledo Physiotherapy Research Group, Instituto de Investigación Sanitaria de Castilla-La Mancha, Castilla-La Mancha, Spain

[*]all authors contributed equally

Corresponding Author:
Javier Bravo-Aparicio, MSc
Toledo Physiotherapy Research Group, Faculty of Physiotherapy and Nursing of Toledo, Universidad de Castilla-La Mancha, Av Carlos III sn, Toledo, Spain

Abstract

Background: Sedentary behavior is highly prevalent among older adults, with adherence to exercise being a major challenge. Exercise offers substantial physical, psychological, and social benefits, but enjoyment is a key factor influencing adherence. Technology-based interventions have shown promise in enhancing motivation and participation, demonstrating higher adherence rates than conventional treatments, although challenges such as motivation loss and technological barriers persist. This review evaluates the effectiveness of active video game interventions on enjoyment and satisfaction in older adults.

Objective: This systematic review and meta-analysis aims to determine whether active video games are superior to other interventions in generating greater enjoyment or satisfaction in older adults.

Methods: PubMed, Cochrane, PEDro, SPORTDiscus, CINAHL, Web of Science, and Scopus databases were searched from inception to September 30, 2024, to identify randomized clinical trials or crossover studies. The primary outcome was enjoyment or satisfaction, assessed using various scales, including the Physical Activity Enjoyment Scale, Intrinsic Motivation Inventory, User Satisfaction Questionnaire, and Likert-type scoring scales. Secondary outcomes included adherence rates and adverse effects. Cochrane Risk of Bias 2 tool was used to evaluate the risk of bias.

Results: Five studies were included in the quantitative analysis. The results indicated a significant improvement in enjoyment or satisfaction compared to the control groups (standardized mean difference 0.34, 95% CI 0.05-0.64; P=.02; I^2=24%), although the effect size was small. Secondary outcomes could not be analyzed due to insufficient data in the selected studies.

Conclusions: Active video game interventions may improve enjoyment and satisfaction in older adults, but the evidence remains of low certainty.

(JMIR Aging 2025;8:e72559) doi:10.2196/72559

KEYWORDS

gamification; exergaming; enjoyment; satisfaction; older adults

Introduction

The global population of older adults is projected to surpass 994 million by 2030, and this trend is expected to continue in the following years [1]. A major concern in aging is the reduced expectancy of healthy life years. The 3 aspects of healthy or successful aging are cognitive or mental well-being, social fulfillment, and physical health, with exercise being crucial for physical health [2]. However, older adults are among the most sedentary population groups, as sedentary behavior seems to increase with age, with older adults spending 62% to 86% of the day in sedentary behavior [3-5].

In 2018, the World Health Organization (WHO) launched an intervention plan aimed at reducing physical inactivity among adults and adolescents by 15% by 2030 [6]. This initiative enhances individual and community health by creating active societies, environments, people, and systems. The intervention emphasizes the importance of exercise or physical activity as a primary treatment approach, given its numerous benefits and low risks of side effects. Regular physical activity improves mortality rates, life expectancy, and physical and functional health outcomes [7-12].

There are discrepancies in the literature regarding the risks and benefits of sedentary behavior and physical activity on physical, psychological, and social outcomes [13,14]. Consequently,

XSL•FO
RenderX

sedentary lifestyles remain prevalent among older populations, as barriers or facilitators to physical activity adherence arise from intrapersonal factors (physical and mental health and individual preferences), interpersonal influences, as well as physical, structural, and organizational environments [15]. Numerous barriers and facilitators affect older adults' engagement and adherence to exercise, shaped by individual experiences and preferences. Researchers have suggested that fun should be incorporated into the FITT (Frequency, Intensity, Time, Type) prescription model [16], as enjoyment may be a critical factor in exercise adherence. Studies have shown that patients perceive exercise differently—some view it as a pleasant activity, while others regard it as an obligation like taking medicine [15,17]. Addressing this barrier through an immediate reward system like enhancing enjoyment could positively transform patients' exercise experiences and potentially improve adherence to physical activity interventions [18,19]. This can be explained through various theoretical frameworks such as operant conditioning theory, self-determination theory, or Ekkekakis model, which link enjoyment to perceived exertion [20,21]. Technology-based interventions, particularly those that integrate engaging and interactive elements, offer a promising solution to enhance motivation, make exercise more enjoyable, and encourage sustained participation among older adults [19].

Gamification, which applies video game design elements such as points, badges, leaderboards, and avatars in nongame contexts, has become an increasingly popular tool in recent years for enhancing adherence to various interventions. Gamification may positively influence user behavior and experience, although its effectiveness may vary depending on the intervention, as inconsistent results have been reported across different age groups [22,23]. Video games or technology-based interventions have demonstrated adherence rates as high as 91%, and in some cases, rates up to 1.38 times higher than conventional exercise treatments or no intervention, which could suggest that greater adherence to physical activity might lead to enhanced health benefits [19,24,25]. Adherence rates in exercise programs for older adults range from 65% to 86% but tend to decline in unsupervised training programs or when the duration exceeded 12 weeks, suggesting that factors such as supervision, program length, and the engaging nature of the intervention play a crucial role in maintaining adherence [19,24,26]. Despite these benefits, such interventions also present challenges, including loss of motivation or interest, space limitations, technological barriers, and feelings of embarrassment when using video games [24,27].

This systematic review aims to primarily evaluate the satisfaction and enjoyment experienced by older adults through active video games. The secondary objective was to determine their adherence to treatment and the possible side effects of the intervention. These metrics are essential for understanding intervention efficacy and long-term adherence.

Methods

Study Registration

The protocol for this systematic review and meta-analysis was registered in PROSPERO (CRD42024593212). This analysis was conducted following the recommendations of the Cochrane Collaboration and the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines.

Search Strategy

A bibliographic search was completed between September 21 and 30, 2024, in the following health and sports science databases: PubMed, Cochrane, Web of Science, CINAHL, SPORTDiscus, PEDro, and Scopus. The search strategies used are available in Multimedia Appendix 1.

Inclusion and Exclusion Criteria

The inclusion criteria for this review were as follows: (1) randomized clinical or controlled trials and crossover studies; (2) patients older than 60 years; (3) exercise or physical activity intervention using gamification, including commercial apps, exergames, or serious games; (4) the comparison group had to perform some form of active exercise, follow their usual treatment, or do nothing; and (5) use of enjoyment or satisfaction scales.

Studies that met any of the following criteria were excluded: (1) case series, observational studies, and conference proceedings and (2) use of enjoyment or satisfaction scales in only one of the groups. No language exclusion criteria were applied.

Study Selection

Studies were selected based on a predefined PICOS (Population, Intervention, Comparison, Outcome) framework established at the outset of the review. The population included older adults, the intervention involved exercise delivered through video games, and the outcome focused on assessing the levels of enjoyment or satisfaction. After defining the search strategy, studies were entered into Rayyan (an app) [28] to exclude duplicate papers. Two researchers selected the studies according to the inclusion and exclusion criteria; in case of disagreement, a third researcher reviewed the study until a consensus was reached.

Data Extraction

First, 2 reviewers extracted informative data from the studies independently; in case of discrepancies, a third reviewer resolved this. The data to be extracted were first author and year, number of participants, design, groups, type of intervention in both groups, outcomes, number of sessions, session time, perceived effort, hardware and software used, and follow-up.

The second part consisted of extracting data values for the different outcomes—both primary and secondary. For the primary outcome of exercise enjoyment or satisfaction and for the secondary outcomes of adherence and adverse effects, mean and standard deviation values were extracted. When values were reported as change or as final values, the extraction of the final values for the analysis was determined as the preferred option.

XSL•FO
RenderX

If data were only available in graphs, the graph digitization software GraphGrabber version 2.0.2 [29] was used for extraction.

Risk of Bias

The methodological quality of the included studies was assessed by 2 independent reviewers (JBA and ITC) using the Cochrane Risk of Bias 2 tool, which evaluates the possible risk of bias in randomized trials for both parallel and crossover design studies [30]; in case of disagreement, the third reviewer broke the tie (HBA). This scale assesses bias based on 5 domains: process randomization, missing data on outcomes, outcome measurement, selection of reported outcomes, and deviations from intended interventions. An additional domain, bias arising from period and carryover effects, was assessed in crossover studies.

The GRADE (Grades of Recommendation Assessment Development and Evaluation) rating system was used to assess the quality of evidence. Publication bias was also assessed using the funnel plot and Egger test for publication asymmetry.

Main Outcomes

The primary outcome variable for this review was exercise enjoyment or satisfaction, which was assessed using the scales reported in the included studies. These scales were not predetermined but were identified during the review process based on the methodologies of the selected studies. Each scale was included because it was used by the respective studies to measure enjoyment or satisfaction, ensuring consistency with their reported outcomes. The following tools were identified.

1. Physical Activity Enjoyment Scale: This scale is a validated and reliable tool used to assess the level of enjoyment individuals experience during a physical activity. The studies reviewed utilized a modified 5-item version of the scale, with responses recorded on a Likert scale ranging from 1 to 7 [31].
2. Intrinsic Motivation Inventory: This is a multidimensional scale designed to assess intrinsic motivation, with various subscales, including interest or enjoyment [32]. Only the interest or enjoyment subscale was used as an outcome measure in the study reviewed.
3. User Satisfaction Questionnaire: This is a 15-item questionnaire divided into 2 parts, that is, the benefits and pitfalls of the intervention and self-perceived improvements in physical and cognitive outcomes.

Two studies [33,34] did not use specific enjoyment or satisfaction scales. Instead, the participants were directly asked about their levels of enjoyment or satisfaction, and their responses were measured using Likert-type scales.

Statistical Analysis

We assessed the overall effects of exercise through video games on enjoyment or satisfaction in older adults. As secondary outcomes, the effects of gamification compared to those of other interventions on adherence and adverse effects were analyzed. Subgroup analyses of the primary outcome (exercise enjoyment or satisfaction) were conducted to explore the key variables potentially influencing variations in enjoyment. These included session time (<10 min vs >10 min), target population (older adults without reported health conditions vs older adults with reported health conditions), immersion type (virtual reality vs augmented reality), type of control group (active vs passive), and the number of sessions (1 session vs >10 sessions).

The inverse variance method analyzed the primary variable (exercise enjoyment or satisfaction). Statistical heterogeneity was assessed using the chi-square test, and the I^2 value was calculated. Heterogeneity was established as low for I^2=25%, moderate for I^2=50%, and high for I^2=75%. The random effects analysis model was used when the heterogeneity was I^2≥50%, and the fixed effects analysis model was used when the heterogeneity was I^2<50%.

The standardized mean difference (SMD) was used for the overall effect on enjoyment or satisfaction, as different scales were implemented in the included studies. For all enjoyment or satisfaction scales, higher scores implied a better result on this outcome.

For all variables, a statistical significance level of P<.05 and 95% CIs were established. The effect size was determined as low when SMD was 0.2, moderate when SMD was 0.5, and high when SMD was 0.8, according to Cohen. Sensitivity analysis was performed individually per study to analyze their influence on the overall results and changes in heterogeneity according to study weight. RevMan software (version 5.4.1; The Cochrane Collaboration) was used for the quantitative analysis.

Deviations From the Protocol

Some analyses foreseen in the protocol registered in PROSPERO could not be performed in this review. Secondary outcomes were foreseen to meta-analyze the adherence to exercise through video games and the adverse effects that these interventions could have; however, this analysis could not be performed, as adherence and the appearance of adverse effects were measured in only 1 [33] and 2 [33,34] studies, respectively.

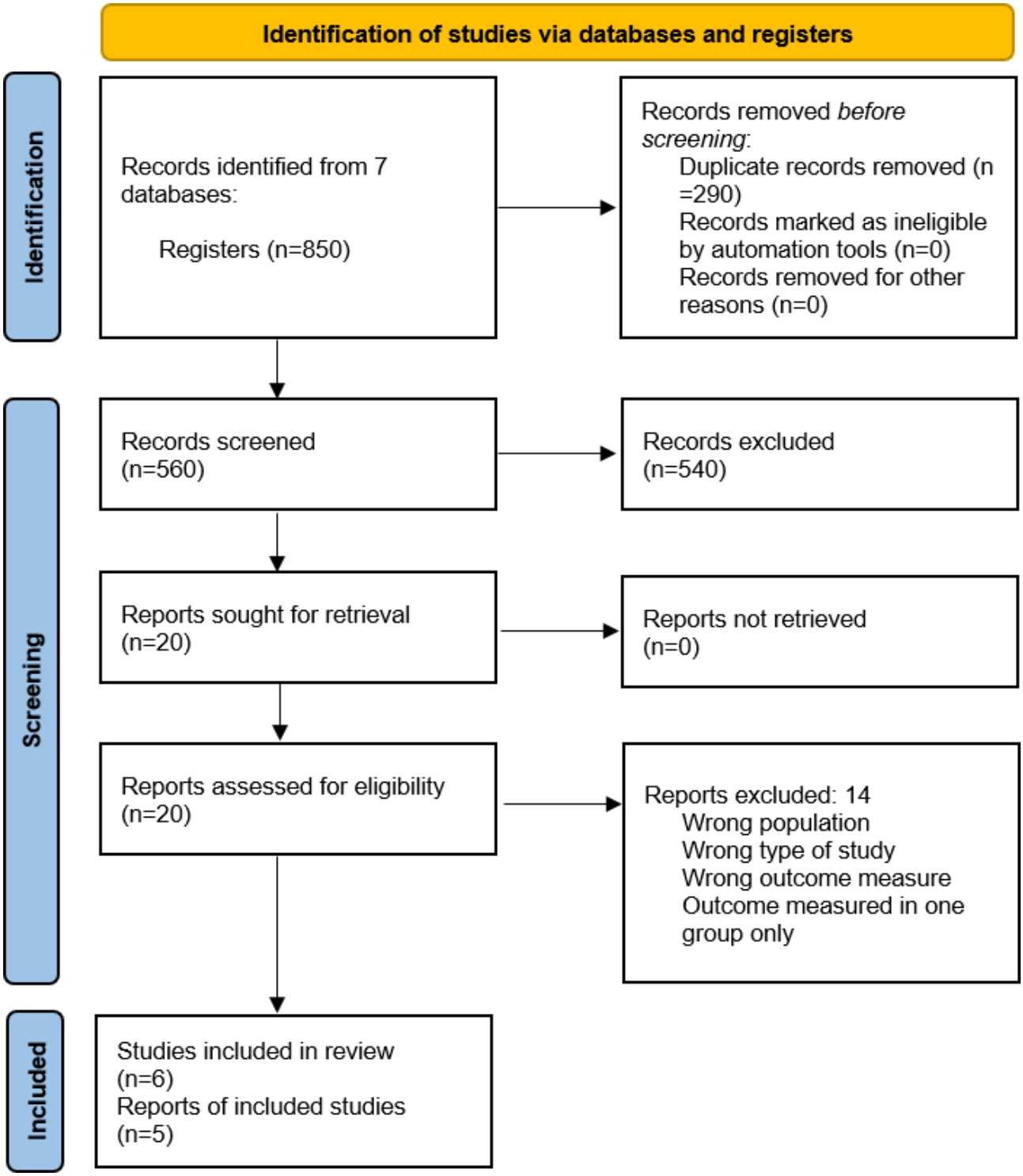
Results

Study Selection

A total of 850 studies were retrieved from the search strategy. After the elimination of 290 duplicated papers, 560 papers were screened, of which 540 were excluded after reading the title and abstract. The remaining 20 studies were included for full-text reading, of which only 6 [33-38] were included in this systematic review, as represented in Figure 1. In the quantitative analysis, 5 studies [33,35-38] encompassing a total of 419 participants were included, with a mean age of 74.72 (SD 6.4) years; 215 participants played an active video game, while 204 participants received other interventions. Takei et al [34] were contacted to obtain the unavailable data in their published paper, but no response was received. This selection process underscores the robustness of the study inclusion.

XSL•FO
RenderX

Figure 1. Flow diagram in this review.



Characteristics of the Included Studies

This systematic review was performed on 6 studies: 2 randomized controlled trials and 4 crossover studies, involving 419 participants [33-38]. Among these, 3 studies focused on older adults without reported health conditions [35,36,38], 2 on older adults undergoing rehabilitation [33,34], and 1 specifically on older adults with Parkinson disease and mild cognitive impairment [37].

Most studies implemented active video game interventions that incorporated full-body movements, combining various therapeutic approaches such as strength training, balance exercises, flexibility routines, yoga, and jogging [33-36], while 2 studies focused on specific exercises for upper limbs and gait, respectively [37,38].

The comparison groups varied across studies. Dockx et al [37], Sayar et al [36], and Takei et al [34] compared the effects of active video games to those of another active intervention, while Kruse et al [38] and Oesch et al [33] compared the effects of video games to those of videos or exercise leaflets. Ferreira et al's [35] trial contrasted the effects of active video games with those of watching television. In terms of immersion, 5 studies

XSL•FO
RenderX

[33-37] employed augmented reality, while Kruse et al's [38] study used virtual reality.

Regarding session duration and frequency, training times typically ranged from 30 minutes to 1 hour, except for Kruse et al's [38] study, which had sessions lasting for 7-10 minutes. Four studies conducted only 1 session of the active video games [34-36,38], while Oesch et al [33] and Dockx et al [37] implemented interventions 2-3 times per week, with overall duration varying between 10 days and 6 weeks.

Two studies [34,36] included exercise intensity parameters measured using the Rate of Perceived Exertion scale, allowing participants to self-regulate the intensity of their interventions. In both cases, participants adjusted their exercise levels based on their own perceptions of effort. Oesch et al [33] mentioned that their exercise intervention was self-regulated, but they did not provide specific data on how this was measured or its effects.

Secondary outcomes such as adherence were investigated only by Oesch et al [33] who reported an overall adherence rate of 85% for both groups. However, the control group showed higher adherence, as the exergame group had more dropouts due to dissatisfaction with the intervention [33]. Regarding adverse effects, Oesch et al [33] and Takei et al [34] reported no adverse effects in their studies. The remaining studies [35-38] did not specify whether any adverse effects occurred. The general characteristics of the included studies and the intervention characteristics in the included studies are shown in Tables 1 and 2, respectively [33-38].

Table . General study characteristics.

Study ID	Study design	Participants (n)	Age (years), mean (SD)	Pathology	Intervention frequency	Time (min)
Dockx et al [37], 2017	RCT[a]	281 (114M[b], 167W[c])	73.75 (6.66)	Older adults without reported health conditions, older adults with mCI[d], older adults with Parkinson disease	3 days a week for 6 weeks	45
Ferreira et al [35], 2022	Crossover study	32 (15M, 17W)	66.70 (4.98)	Older adults without reported health conditions	1 session	50
Kruse et al [38], 2021	Crossover study	25 (3M, 22W)	81.24 (4.97)	Older adults without reported health conditions	1 session	7 - 10
Oesch et al [33], 2017	RCT	54 (29M, 25W)	74.05 (9.25)	Older adults in rehabilitation	Twice a day for 10 days	30
Sayar et al [36], 2023	Crossover study	40 (17M, 23W)	69.60 (4.16)	Older adults without reported health conditions	1 session	30
Takei et al [34], 2023	Crossover study	16 (3M, 13W)	83 (7)	Older adults in rehabilitation	1 session	60

[a]RCT: randomized controlled trial.

[b]M: men.

[c]W: women.

[d]mCI: mild cognitive impairment.

XSL•FO
RenderX

Table . Intervention characteristics in the included studies in this review.

Study ID	Experimental group	Control group	Video game type	Hardware	Software	Movement required	Outcomes	Intensity
Dockx et al [37], 2017	Treadmill with augmented reality	Treadmill	AR[a]	Screen for projecting visual content	Screen simulating walking in the street	Gait	USQ[b]	N/A[c]
Ferreira et al [35], 2022	"Your Shape Fitness Evolved" video game	Watch television	AR	Xbox Kinect	"Your Shape Fitness Evolved" (Stack' em up, zen develop it, pump it, wall breaker, hurricane)	Full-body movement	PACES[d]	N/A
Kruse et al [38], 2021	VR[e] video game	Exercise video	VR	Valve Index VR headset	Maestro game VR	Upper limbs exercises	IMI[f]	N/A
Oesch et al [33], 2017	Windows Kinect video games from GameUp Project	Exercise leaflet	AR	Windows Kinect	Game up	Full-body movement	Enjoyment (Likert type scale) adherence and adverse effects	Self-regulated
Sayar et al [36], 2023	Xbox Kinect video game	Brisk walking	AR	Xbox Kinect	"Kinect Adventures!" and "Your Shape Fitness Evolved 2012"	Full-body movement	PACES	RPE[g] (1-10)
Takei et al [34], 2023	Nintendo switch video game	Physical therapy	AR	Nintendo switch, ring fit, and leg sensor	Nintendo switch video games	Full-body movement	Enjoyment (Likert type scale) and adverse effects	RPE (6-20)

[a]AR: augmented reality.

[b]USQ: User Satisfaction Questionnaire.

[c]N/A: not applicable.

[d]PACES: Physical Activity Enjoyment Scale.

[e]VR: virtual reality.

[f]IMI: Intrinsic Motivation Inventory.

[g]RPE: Rate of Perceived Exertion.

Risk of Bias in the Included Studies

The agreement rate achieved between the two authors who completed the risk of bias assessment was 80%; in case of disagreement (20%), the third reviewer resolved it. The risk of bias in the 6 studies is represented in Figure 2 [33-38].

The Egger regression–based test was conducted to evaluate the presence of publication bias. The results indicated that the intercept was not significantly different from 0 (intercept=2.985; P=.83), suggesting no evidence of small-study effects.

The funnel plot (Figure 3) visually supports these findings, showing a relatively symmetric distribution of the effect sizes around the estimated overall effect size. The absence of asymmetry further suggests that publication bias is unlikely to have significantly influenced the results of this meta-analysis.

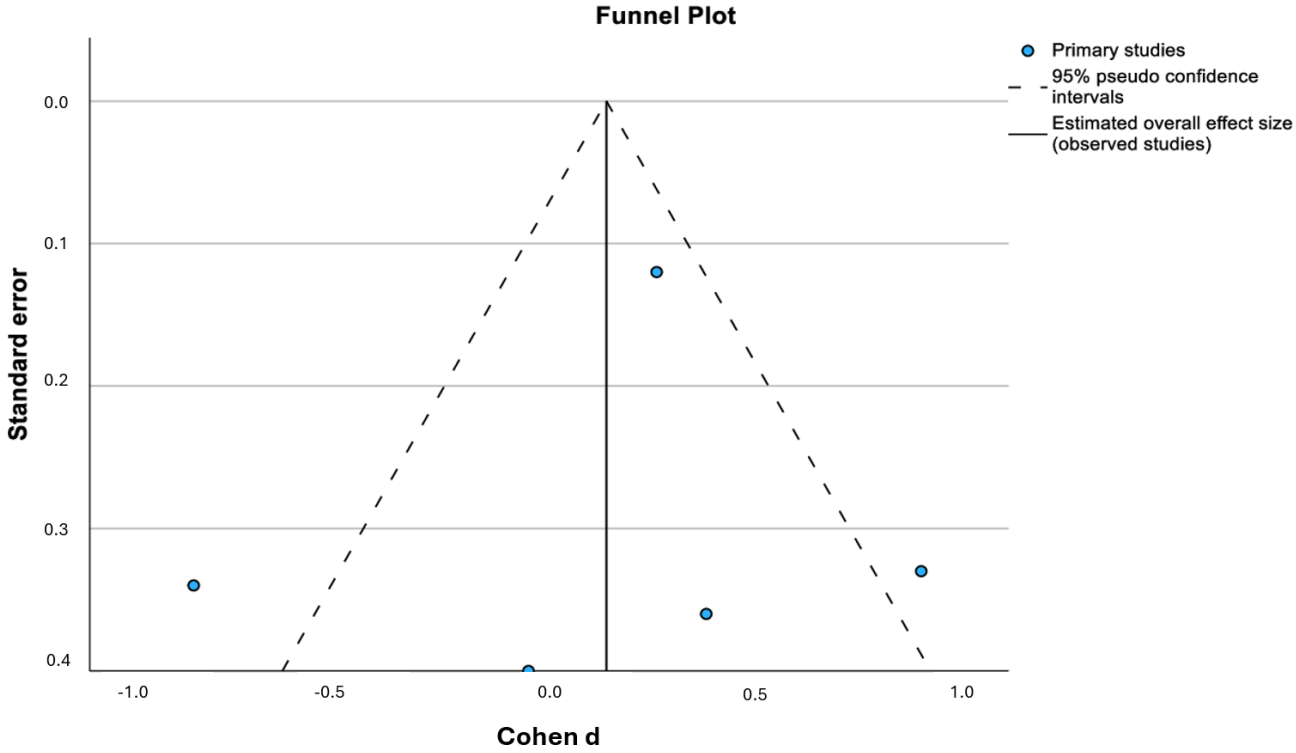
Based on these findings, there is no statistical evidence of publication bias in the included studies.

XSL•FO
RenderX

Figure 2. Risk of bias assessment. D1: randomization process. DS: bias arising from period and carryover effects. D2: deviation from intended interventions. D3: missing outcome data. D4: measurement of the outcome. D5: selection of the reported result.



Figure 3. Funnel plot of the included studies.

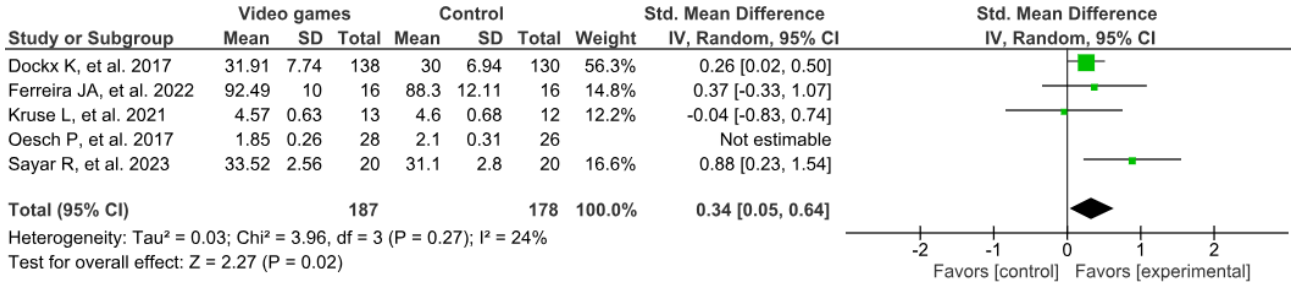


Quantitative Analysis

The active video game group showed an improvement in the overall enjoyment or satisfaction compared to the control group after the intervention period, as shown in Figure 4 [33,35-38], which shows a small effect size and low heterogeneity (SMD 0.34, 95% CI 0.05-0.64; P=.02; I^2=24%). The certainty of the evidence for overall enjoyment in active video games versus that in control interventions was rated as low according to the GRADE approach. This was based on data from 4 randomized trials involving 187 participants in the intervention groups and 178 in the control groups. The SMD for enjoyment was 0.34 SD (95% CI 0.05-0.64) higher in the active video game group. The evidence was downgraded due to very serious risk of bias, while inconsistency, indirectness, and imprecision were not considered serious. No other concerns were identified (Multimedia Appendix 2).

XSL•FO
RenderX

Figure 4. Forest plot of the overall enjoyment.



For the quantitative analysis, the study by Oesch et al [33] was removed after the sensitivity analysis because it had a small sample size and a large effect size (SMD −0.86, 95% CI −1.42 to −0.3), which increased heterogeneity by 55% for a study weight of 20.6%. Its inclusion significantly affected the overall effect result (SMD 0.12, 95% CI −0.41 to 0.64; P=.67; I^2=79%), as shown in Multimedia Appendix 3.

The effect of active video games on exercise enjoyment or satisfaction by subgroups is shown in Table 3. No significant differences were found based on session time, target population, immersion type, number of sessions, or control group type. However, some subgroup comparisons were close to reaching statistical significance. Notably, when the effects of active video games were compared to those of an active intervention, enjoyment was higher with active video games, although this difference was not statistically significant (P=.08). Additionally, 2 other subgroup analyses approached significance: older adults without reported health conditions appeared to enjoy active video games more (P=.12), and fewer sessions seemed to result in greater enjoyment or satisfaction (P=.12).

XSL•FO
RenderX

Table . Subgroup analysis.

Subgroup	Studies (n)	Participants (n)	Random effect		Heterogeneity (%), I^2	Subgroup difference	
			SMD[a] (95%CI)	*P* value		Chi-square *(df)*	*P* value
Session time (min)						0.1 (1)	0.71
<10	1	25	−0.04 (-−0.83 to 0.74)	0.91	N/A[b]		
>10	4	394	0.15 (−0.48 to 0.78)	0.67	79		
Target population						1.3 (1)	0.25
Older adults without reported health condition-ss	3	97	0.44 (−0.09 to 0.96)	0.1	39		
Older adults with reported health conditions	2	322	−0.27 (−1.37 to 0.83)	0.63	92		
Immersion type						0.1 (1)	0.71
Virtual reality	1	25	−0.04 (−0.83 to 0.74)	0.91	N/A		
Augmented reali-ty	4	394	0.15 (−0.48 to 0.78)	0.64	84		
Number of sessions						1.3 (1)	0.25
1	3	97	0.44 (−0.09 to 0.96)	0.1	39		
>10	2	322	−0.27 (−1.37 to 0.83)	0.63	92		
Control group type						1.9 (1)	0.16
Active interven-tion	2	308	0.49 (−0.1 to 1.09)	0.1	68		
Passive interven-tion	3	111	−0.21 (−0.98 to 0.57)	0.6	74		

[a]SMD: standardized mean difference.

[b]N/A: not applicable.

Discussion

Principal Findings

This systematic review with meta-analysis evaluates the specific effectiveness of active video games on enjoyment and satisfaction experienced by older adults—outcomes that are crucial for adherence to physical activity programs. Our findings indicate that exercise delivered through active video games could provide greater enjoyment or satisfaction than control interventions.

Enjoyment is a key determinant in long-term adherence to physical activity, as it enhances engagement and sustainability. Consequently, incorporating enjoyable components such as active video games into exercise regimens aligns with proposals to include fun within the FITT principles for a more holistic exercise prescription [16]. Studies have shown that enjoyment could serve as either a barrier or a facilitator in the adherence to exercise routines [15,26]. Therefore, incorporating enjoyable elements such as active video games into exercise routines or treatments could potentially help individuals stay engaged in a physical activity, as the overall effect on enjoyment and satisfaction in this review suggests that active video games may offer a modest advantage over control interventions, potentially enhancing the appeal of exercise for older participants.

In this systematic review, we found that only Oesch et al [33] examined adherence to active video game interventions, and they reported that the control group showed higher adherence, while the experimental group showed a higher dropout rate due to participants disliking the treatment. These findings contrast with those of Valenzuela et al [24] who demonstrated increased adherence to technology-based interventions. This divergence may stem from differences in the intervention design, participant characteristics, or contextual factors, warranting further exploration. However, the results from Oesch et al's [33] study may align with general adherence trends for physical activity, where nonadherence rates range from 47% to 96% within the first year in healthy populations and from 50% to 70% in

XSL•FO
RenderX

patients undergoing physical therapy, while adherence rates in older adults range from 65% to 86% [26,39-41].

Among the reviewed studies, only Oesch et al [33] and Takei et al [34] reported adverse effects, with neither identifying any incidents during their interventions. Although these findings suggest that active video game interventions are generally safe, the absence of reporting in other studies limits definitive conclusions regarding their safety profiles [33,34]. However, most of the studies [35-38] included did not report the occurrence of adverse effects.

Subgroup analyses revealed no statistically significant differences between active video game interventions and control interventions; however, comparisons between the active video game intervention group and the control group approached significance (P=.08). This trend suggests a potential for differential effects that may become apparent with larger sample sizes or more targeted studies. Sayar et al [36] used a crossover design, allowing participants to experience both interventions and compare them directly in terms of enjoyment. This design made it possible to observe which intervention generated a greater sense of enjoyment among participants. In contrast, Dockx et al [37] compared the effects of the usual treadmill walking intervention with those of an intervention that had a screen simulating standard treadmill walking, and they suggested that the added visual and auditory distractions may have contributed to participants' preference for the screen-enhanced intervention, as participants may have perceived less exertion [42,43].

Other comparisons approaching significance (P=.12) were found in older adults without reported health conditions versus older adults with reported health conditions and in 1 session versus >10 sessions, with the same studies [33,35-38] included in each subgroup for the number of sessions. One possible explanation for greater enjoyment in fewer sessions is that repeated exposure to the same intervention might lead to decreased motivation, as older adults could lose interest in the video game or view the technology as more of a barrier than a facilitator [19,24]. Similarly, this may also explain why some older adults with health issues did not favor this type of treatment. A study [44] on older adults experiencing chronic low back pain indicated that continued engagement in physical activity was often due to the enjoyable experience of the exercise itself. In contrast, in Oesch et al's [33] study, older adults who did not find the activity enjoyable frequently might have perceived it as a barrier, which in some cases contributed to their decision to drop out of the study [33].

Limitations

This review has several limitations. The generalizability of our findings is limited by the small number of the included studies and their high risk of bias, primarily due to issues in randomization and blinding of participants and assessors. Although the statistical heterogeneity was low after removing one study [33] that significantly increased the variability, differences in the intervention types and outcome measurement methods still contribute to some methodological inconsistencies. Additionally, only 2 studies [34,36] reported on the exercise intensity—a factor known to influence enjoyment and satisfaction through established models [21,45,46]. Future research should explore how exercise enjoyment affects adherence in rehabilitation programs and examine whether perceived exertion influences enjoyment or satisfaction. Addressing these gaps could strengthen the evidence base. Other possible research directions include a systematic review on the role of active video games in adherence or a qualitative study exploring factors that influence older adults' adherence to exercise programs.

Clinical Implications

From a clinical perspective, this review does not establish a definitive advantage of active video games over traditional interventions. However, tailoring exercise programs to individual preferences and integrating enjoyable elements may optimize patient adherence and satisfaction, aligning with patient-centered care principles.

Conclusion

Active video games could help improve enjoyment or satisfaction in older adults, with a low certainty of evidence. In this systematic review, active video games did not show a superior effect to conventional treatment on adherence. Future research should explore optimizing gamification techniques to maximize adherence and satisfaction.

Acknowledgments

This research received no external funding. Part of JBA's and ITC's salaries were financed by the European Social Fund Plus.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File, 15 KB - aging_v8i1e72559_app1.docx]

Multimedia Appendix 2
Assessment of evidence according to GRADE (Grades of Recommendation Assessment Development and Evaluation).
[DOCX File, 16 KB - aging_v8i1e72559_app2.docx]

XSL•FO
RenderX

Multimedia Appendix 3
Forest plot of the overall enjoyment, including Oesch et al's [33] study.
[PNG File, 91 KB - aging_v8i1e72559_app3.png]

Checklist 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[PDF File, 84 KB - aging_v8i1e72559_app4.pdf]

References

1. Gerland P, Hertog S, Wheldon MC, et al. World population prospects 2022 summary of results. United Nations. URL: https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/wpp2022_summary_of_results.pdf [accessed 2025-05-20]
2. Urtamo A, Jyväkorpi SK, Strandberg TE. Definitions of successful ageing: a brief review of a multidimensional concept. Acta Biomed 2019 May 23;90(2):359-363. [doi: 10.23750/abm.v90i2.8376] [Medline: 31125022]
3. Harvey JA, Chastin SFM, Skelton DA. How sedentary are older people? A systematic review of the amount of sedentary behavior. J Aging Phys Act 2015 Jul;23(3):471-487. [doi: 10.1123/japa.2014-0164] [Medline: 25387160]
4. Kim Y, Lee E. The association between elderly people's sedentary behaviors and their health-related quality of life: focusing on comparing the young-old and the old-old. Health Qual Life Outcomes 2019 Jul 26;17(1):131. [doi: 10.1186/s12955-019-1191-0] [Medline: 31349858]
5. Gorman E, Hanson HM, Yang PH, Khan KM, Liu-Ambrose T, Ashe MC. Accelerometry analysis of physical activity and sedentary behavior in older adults: a systematic review and data analysis. Eur Rev Aging Phys Act 2014;11(1):35-49. [doi: 10.1007/s11556-013-0132-x] [Medline: 24765212]
6. Global action plan on physical activity 2018–2030: more active people for a healthier world. World Health Organization. URL: https://iris.who.int/bitstream/handle/10665/272722/9789241514187-eng.pdf [accessed 2025-05-20]
7. Ekelund U, Steene-Johannessen J, Brown WJ, et al. Does physical activity attenuate, or even eliminate, the detrimental association of sitting time with mortality? A harmonised meta-analysis of data from more than 1 million men and women. Lancet 2016 Sep 24;388(10051):1302-1310. [doi: 10.1016/S0140-6736(16)30370-1] [Medline: 27475271]
8. Moore SC, Patel AV, Matthews CE, et al. Leisure time physical activity of moderate to vigorous intensity and mortality: a large pooled cohort analysis. PLoS Med 2012;9(11):e1001335. [doi: 10.1371/journal.pmed.1001335] [Medline: 23139642]
9. Hall AK, Chavarria E, Maneeratana V, Chaney BH, Bernhardt JM. Health benefits of digital videogames for older adults: a systematic review of the literature. Games Health J 2012 Dec;1(6):402-410. [doi: 10.1089/g4h.2012.0046] [Medline: 26192056]
10. Mazeas A, Duclos M, Pereira B, Chalabaev A. Evaluating the effectiveness of gamification on physical activity: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2022 Jan 4;24(1):e26779. [doi: 10.2196/26779] [Medline: 34982715]
11. Taylor LM, Kerse N, Frakking T, Maddison R. Active video games for improving physical performance measures in older people: a meta-analysis. J Geriatr Phys Ther 2018;41(2):108-123. [doi: 10.1519/JPT.0000000000000078] [Medline: 26974212]
12. Xu W, Liang HN, Baghaei N, Wu Berberich B, Yue Y. Health benefits of digital videogames for the aging population: a systematic review. Games Health J 2020;9(6):389-404. [doi: 10.1089/g4h.2019.0130] [Medline: 32589482]
13. McGowan LJ, Powell R, French DP. Older adults' construal of sedentary behaviour: implications for reducing sedentary behaviour in older adult populations. J Health Psychol 2021 Oct;26(12):2186-2199. [doi: 10.1177/1359105320909870] [Medline: 32114825]
14. Compernolle S, De Cocker K, Cardon G, De Bourdeaudhuij I, Van Dyck D. Older adults' perceptions of sedentary behavior: a systematic review and thematic synthesis of qualitative studies. Gerontologist 2020;60(8):572-582. [doi: 10.1093/geront/gnz127] [Medline: 31670766]
15. Bethancourt HJ, Rosenberg DE, Beatty T, Arterburn DE. Barriers to and facilitators of physical activity program use among older adults. Clin Med Res 2014 Sep;12(1-2):10-20. [doi: 10.3121/cmr.2013.1171] [Medline: 24415748]
16. Burnet K, Kelsch E, Zieff G, Moore JB, Stoner L. How fitting is F.I.T.T.?: a perspective on a transition from the sole use of frequency, intensity, time, and type in exercise prescription. Physiol Behav 2019 Feb 1;199:33-34. [doi: 10.1016/j.physbeh.2018.11.007] [Medline: 30414400]
17. Devereux-Fitzgerald A, Powell R, Dewhurst A, French DP. The acceptability of physical activity interventions to older adults: a systematic review and meta-synthesis. Soc Sci Med 2016;158(14–23):14-23. [doi: 10.1016/j.socscimed.2016.04.006] [Medline: 27104307]
18. Woolley K, Fishbach A. Immediate rewards predict adherence to long-term goals. Pers Soc Psychol Bull 2017;43(2):151-162. [doi: 10.1177/0146167216676480] [Medline: 27899467]

XSL•FO
RenderX

19. Collado-Mateo D, Lavín-Pérez AM, Peñacoba C, et al. Key factors associated with adherence to physical exercise in patients with chronic diseases and older adults: an umbrella review. Int J Environ Res Public Health 2021;18(4):2023. [doi: 10.3390/ijerph18042023] [Medline: 33669679]
20. Teixeira PJ, Carraça EV, Markland D, Silva MN, Ryan RM. Exercise, physical activity, and self-determination theory: a systematic review. Int J Behav Nutr Phys Act 2012 Jun 22;9(1):78. [doi: 10.1186/1479-5868-9-78] [Medline: 22726453]
21. Ekkekakis P. Pleasure and displeasure from the body: perspectives from exercise. Cogn Emot 2003 Mar;17(2):213-239. [doi: 10.1080/02699930302292] [Medline: 29715726]
22. Krath J, Schürmann L, von Korflesch HFO. Revealing the theoretical basis of gamification: a systematic review and analysis of theory in research on gamification, serious games and game-based learning. Comput Human Behav 2021 Dec;125:106963. [doi: 10.1016/j.chb.2021.106963]
23. Johnson D, Deterding S, Kuhn KA, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. Internet Interv 2016 Nov;6(6):89-106. [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]
24. Valenzuela T, Okubo Y, Woodbury A, Lord SR, Delbaere K. Adherence to technology-based exercise programs in older adults: a systematic review. J Geriatr Phys Ther 2018;41(1):49-61. [doi: 10.1519/JPT.0000000000000095] [Medline: 27362526]
25. Wen CP, Wai JPM, Tsai MK, et al. Minimum amount of physical activity for reduced mortality and extended life expectancy: a prospective cohort study. Lancet 2011 Oct 1;378(9798):1244-1253. [doi: 10.1016/S0140-6736(11)60749-6] [Medline: 21846575]
26. Rivera-Torres S, Fahey TD, Rivera MA. Adherence to exercise programs in older adults: informative report. Gerontol Geriatr Med 2019;5:2333721418823604. [doi: 10.1177/2333721418823604] [Medline: 30733977]
27. Ning H, Jiang D, Du Y, et al. Older adults' experiences of implementing exergaming programs: a systematic review and qualitative meta-synthesis. Age Ageing 2022 Dec 5;51(12):afac251. [doi: 10.1093/ageing/afac251] [Medline: 36571772]
28. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
29. Scientific software development and application services. Quintessa. URL: https://www.quintessa.org/software/ [accessed 2025-05-28]
30. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898. [doi: 10.1136/bmj.l4898] [Medline: 31462531]
31. Kendzierski D, DeCarlo KJ. Physical activity enjoyment scale: two validation studies. J Sport Exerc Psychol 1991;13(1):50-64. [doi: 10.1123/jsep.13.1.50]
32. Ryan RM, Mims V, Koestner R. Relation of reward contingency and interpersonal context to intrinsic motivation: a review and test using cognitive evaluation theory. J Pers Soc Psychol 1983;45(4):736-750. [doi: 10.1037//0022-3514.45.4.736]
33. Oesch P, Kool J, Fernandez-Luque L, et al. Exergames versus self-regulated exercises with instruction leaflets to improve adherence during geriatric rehabilitation: a randomized controlled trial. BMC Geriatr 2017 Mar 23;17(1):77. [doi: 10.1186/s12877-017-0467-7] [Medline: 28330455]
34. Takei K, Morita S, Watanabe Y, Suganami M, Inao M. Safety, feasibility, and acceptability of physiotherapy combined with strength training using active video games for older patients with musculoskeletal conditions. Disabil Rehabil Assist Technol 2024 Apr;19(3):641-647. [doi: 10.1080/17483107.2022.2115562] [Medline: 36001528]
35. Ferreira JA, Passos GS, Youngstedt SD, et al. Are there sex differences in energy expenditure and enjoyment in older adults playing active video games? J Bodyw Mov Ther 2022 Jul;31:72-76. [doi: 10.1016/j.jbmt.2022.04.008] [Medline: 35710225]
36. Sayar R, Sinaei E, Daryanoosh F, Koshki M, Sobhani S. Comparing the effects of 30 minutes exergaming versus brisk walking on physiological and psychological measurements of older adults. J Bodyw Mov Ther 2023 Oct;36:69-73. [doi: 10.1016/j.jbmt.2023.06.011] [Medline: 37949602]
37. Dockx K, Alcock L, Bekkers E, et al. Fall-prone older people's attitudes towards the use of virtual reality technology for fall prevention. Gerontology 2017;63(6):590-598. [doi: 10.1159/000479085] [Medline: 28817828]
38. Kruse L, Karaosmanoglu S, Rings S, Ellinger B, Steinicke F. Enabling immersive exercise activities for older adults: a comparison of virtual reality exergames and traditional video exercises. Societies (Basel) 2021 Sep;11(4):134. [doi: 10.3390/soc11040134]
39. Sperandei S, Vieira MC, Reis AC. Adherence to physical activity in an unsupervised setting: explanatory variables for high attrition rates among fitness center members. J Sci Med Sport 2016 Nov;19(11):916-920. [doi: 10.1016/j.jsams.2015.12.522] [Medline: 26874647]
40. Sluijs EM, Kok GJ, van der Zee J. Correlates of exercise compliance in physical therapy. Phys Ther 1993 Nov;73(11):771-782. [doi: 10.1093/ptj/73.11.771] [Medline: 8234458]
41. McLean SM, Klaber Moffett JA, Sharp DM, Gardiner E. A randomised controlled trial comparing graded exercise treatment and usual physiotherapy for patients with non-specific neck pain (the GET UP neck pain trial). Man Ther 2013 Jun;18(3):199-205. [doi: 10.1016/j.math.2012.09.005] [Medline: 23085116]
42. Stewart TH, Villaneuva K, Hahn A, et al. Actual vs. perceived exertion during active virtual reality game exercise. Front Rehabil Sci 2022;3:887740. [doi: 10.3389/fresc.2022.887740] [Medline: 36189005]

XSL•FO
RenderX

43. Neumann DL, Moffitt RL. Affective and attentional states when running in a virtual reality environment. Sports (Basel) 2018 Jul 27;6:71. [doi: 10.3390/sports6030071] [Medline: 30060451]
44. Hay ME, Connelly DM. Exploring the experience of exercise in older adults with chronic back pain. J Aging Phys Act 2020 Apr 24;28(2):294-305. [doi: 10.1123/japa.2019-0030] [Medline: 31722296]
45. Tenenbaum G, Hutchinson JC. A Social-Cognitive Perspective of Perceived and Sustained Effort Handbook of Sport Psychology: John Wiley and Sons; 2012. [doi: 10.1002/9781118270011]
46. Leventhal H, Everhart D. Emotion, pain, and physical illness. In: Emotions in Personality and Psychopathology: Springer; 1979:261-299. [doi: 10.1007/978-1-4613-2892-6_10]

Abbreviations

FITT: Frequency, Intensity, Time, and Type
GRADE: Grades of Recommendation Assessment Development and Evaluation
PICOS: Population, Intervention, Comparison, Outcome
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SMD: standardized mean difference
WHO: World Health Organization

Edited by C Lee; submitted 12.02.25; peer-reviewed by J Ogunsakin, S Hill; revised version received 25.03.25; accepted 09.04.25; published 12.06.25.

Please cite as:
Bravo-Aparicio J, Trillo-Charlín I, Avendaño-Coy J, Beltran-Alacreu H
Effectiveness of Gamification on Enjoyment and Satisfaction in Older Adults: Systematic Review and Meta-Analysis
JMIR Aging 2025;8:e72559
URL: https://aging.jmir.org/2025/1/e72559
doi:10.2196/72559

© Javier Bravo-Aparicio, Iria Trillo-Charlín, Juan Avendaño-Coy, Hector Beltran-Alacreu. Originally published in JMIR Aging (https://aging.jmir.org), 12.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Frailty, Fitness, and Quality of Life Outcomes of a Healthy and Productive Aging Program (GrandMove) for Older Adults With Frailty or Prefrailty: Cluster Randomized Controlled Trial

Jennifer Yee Man Tang[1], PhD; Hao Luo[2,3], PhD; Michael Tse[4], PhD; Joseph Kwan[5], MBChB, MD; Angela Yee Man Leung[6], PhD; Teresa Bik-kwan Tsien Wong[7], MSc; Terry Yat Sang Lum[3], PhD; Gloria Hoi Yan Wong[3,8], PhD

[1]Department of Educational Psychology, The Chinese University of Hong Kong, Shatin, China (Hong Kong)

[2]School of Public Health Sciences, University of Waterloo, Waterloo, ON, Canada

[3]Department of Social Work and Social Administration, The University of Hong Kong, Hong Kong, Pok Fu Lam, China (Hong Kong)

[4]Centre for Sports and Exercise, The University of Hong Kong, Hong Kong, Pok Fu Lam, China (Hong Kong)

[5]Department of Brain Sciences, Imperial College London, London, United Kingdom

[6]School of Nursing, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[7]Institute of Active Ageing, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[8]School of Psychology and Clinical Language Sciences, University of Reading, Earley Gate, Whiteknights Campus, Reading, United Kingdom

Corresponding Author:
Gloria Hoi Yan Wong, PhD
Department of Social Work and Social Administration, The University of Hong Kong, Hong Kong, Pok Fu Lam, China (Hong Kong)

Abstract

Background: Exercise interventions can reverse frailty. However, their scalability and sustainability are limited by manpower, which is reducing due to population aging. GrandMove is a program that combines healthy and productive aging strategies to (1) train and employ robust older adults as exercise coaches and (2) improve fitness and motivate the adoption of an exercise habit in older adults with frailty and prefrailty.

Objective: The aim of this study is to examine the effectiveness of GrandMove in improving frailty, fitness, and quality of life in older adults with frailty and prefrailty.

Methods: This cluster randomized controlled trial recruited older adults with frailty and prefrailty (N=390) living in the community. The 18-month exercise program consisted of three 6-month phases of lifestyle education (E), resistance exercise (R), and aerobic exercise (A). Each group of participants was randomized into 3 intervention sequence arms: the E-R-A group, the A-R-E group, and the R-A-E group.

Results: At 6, 12, and 18 months, 346, 305, and 264 participants completed the frailty assessment, respectively. At 6 months, 100 of 346 participants (28.9%) were robust. A-R-E and R-A-E were no better than E-R-A as the active control in addressing frailty over the first 6 months (A-R-E: interaction coefficient 0.07, 95% CI −0.35 to 0.49, *P*=.68; R-A-E: interaction coefficient −0.02, 95% CI −0.42 to 0.38, *P*=.90). Compared to lifestyle education, resistance training and aerobic training over the first 6 months were associated with greater improvement in fitness measures of grip strength for the left hand (A-R-E: interaction coefficient 2.99, 95% CI 0.76 to 5.23, *P*=.009; R-A-E: interaction coefficient 2.21, 95% CI 0.63 to 4.36, *P*=.04) and right hand (A-R-E: interaction coefficient 3.75, 95% CI 1.54 to 5.97, *P*=.001; R-A-E: interaction coefficient 2.29, 95% CI 0.16 to 4.42, *P*=.04) and arm curl test (A-R-E: interaction coefficient 1.42, 95% CI 0.39 to 2.46, *P*=.007; R-A-E: interaction coefficient 1.11, 95% CI 0.12 to 2.11, *P*=.03). The sequence of exercise interventions (R-A-E vs A-R-E) did not make a difference in primary outcomes at 12 months, but the R-A-E group showed better quality of life (interaction coefficient 4.50, 95% CI 0.12 to 8.88, *P*=.008). Improved frailty outcomes were maintained by the end of the study, but the change in overall physical activity level was limited.

Conclusions: Combining healthy and productive aging strategies is a scalable and sustainable way to improve frailty, fitness, and quality of life in older adults with frailty and prefrailty. Different combinations of lifestyle education and physical interventions improved frailty.

Trial Registration: HKU Clinical Trials Registry HKUCTR-1964; https://www.hkuctr.com/Study/Show/75c5d2e6825c4b5498f0c65c82714c4b

(JMIR Aging 2025;8:e65636) doi:10.2196/65636

XSL•FO
RenderX

KEYWORDS

physical activity; physical exercise; exercising; gerontology; geriatric; older adult; older person; older people; active aging; postretirement work; cluster randomized controlled trial; RCT

Introduction

Frailty, a common condition that increases with age, is a dynamic clinical state that can improve or worsen over time [1,2]. About 1 in 10 older adults are frail and 46% are prefrail, a state with higher risk of progression to frailty [3]. Frailty is known to predict disability, mortality, risk of fractures, and many other adverse health outcomes [4]. To date, evidence for the management of frailty is the strongest with physical activity [5,6]. In older adults who are very frail, even small gains in strength can result in important functioning benefits and promote independence [5]. Evidence suggested superior outcomes with training carried out 3 or more times per week, for 30-45 minutes per session, and lasting at least 3-5 months [7]. Adherence to exercise training is good even in older adults with frailty, with few adverse events or safety concerns [5].

Practice guidelines widely recommend multicomponent exercise programs for the management of frailty [8]. Many programs concurrently prescribe multimodal exercise components, even though sequential prescription may promote adherence and minimize attrition [8,9]. The optimal sequence is yet to be investigated. Theoretically, strength and aerobic training contribute in different ways by slowing or compensating for muscle wasting and loss of endurance, as well as preventing other diseases and resisting the cascade of disability [10]. These exercise modalities impact on the older person's quality of life by affecting their ability to lift load (eg, arising from a chair) and endurance in performing activities of daily living (eg, walking), allowing them to take control and lead a purposeful life. These may have implications for the design of exercise training protocols for frailty intervention.

Past development of exercise programs put a greater emphasis on maximizing physical benefits, while less focus was put on the scalability and sustainability of the program. It is foreseeable that the health care manpower gap as a result of population aging is increasing [11] and heavy reliance on professionally led formal health care services for addressing frailty is unviable and unsustainable. Our team developed a multicomponent exercise program, namely the GrandMove program, which has a group of trained coaches in their 50s to 60s deliver the interventions. Delivery of exercise interventions by peer coaches rather than health professionals is a key to the long-term scalability and sustainability of such programs in the context of population aging. Our program incorporated both healthy and productive aging strategies. Although the program had clear goals of improving frailty, fitness, and quality of life in older adults with frailty and prefrailty, our peer coaches took on their own path to successful aging by adopting an active lifestyle and engaging in postretirement paid work. A detailed description of the program design is provided in the Methods and Multimedia Appendix 1.

In this study, three intervention sequences were designed to (1) examine the effectiveness of a 6-month aerobic or resistance training program compared with lifestyle education as the active control intervention at 6 months, (2) investigate any order effect in resistance and aerobic training in older adults with frailty and prefrailty at 12 months, and (3) evaluate exercise habit formation over 18 months.

Methods

Study Design and Participants

This is an 18-month, multicenter, cluster randomized controlled trial with 3 intervention arms. A cluster design was used because of the logistical issues associated with the implementation of the interventions. The study sites included 14 community service centers for older adults and 15 public rental estates in Hong Kong. The staff from the participating sites referred potential participants based on their age and frailty status at screening.

Ethical Considerations

Designated research assistants, who were not involved in the intervention, obtained written informed consent and conducted screening interviews with the participants. The study protocol was approved by the institutional ethics committee (reference number EA1511048). The study was registered with the HKU Clinical Trials Registry (reference number HKUCTR-1964). All data used in this study were deidentified before analysis. Participants did not receive any compensation for their involvement.

Participants

Target study participants were older adults living in the community who did not have any contradictions to participation in a moderate level of physical activity. Between February 2016 and May 2017, older adults were invited to take part in this study if they were aged 65 years or older and were screened as either prefrail (scores of 1‐2 on the FRAIL [Fatigue, Resistance, Ambulation, Illness, and Loss of weight] scale) or frail (scores of 3‐5 on the FRAIL scale) [2]. To ensure the safety of the participants, older adults were excluded from participating if they reported having specific conditions. Specific exclusion criteria are described in Multimedia Appendix 2.

The sample size was calculated based on an estimated 0.5-point improvement in the FRAIL score (SD 1) between an exercise condition versus lifestyle education, using a 2-sided test at 1% significance level with 80% power. Assuming a 20% dropout rate, the minimum sample size is 120 per arm. It was estimated that about 1 in 10 older adults was frail. To allow subgroup analysis by frailty status, the ratio of participants with prefrailty to those with frailty was set at about 6:4 (Multimedia Appendix 2).

Randomization and Masking

As shown in Multimedia Appendix 3, each group of participants was randomly assigned into one of three parallel arms with different intervention sequences: (1) lifestyle education – resistance training – aerobic training (E-R-A); (2) aerobic

XSL•FO
RenderX

training – resistance training – lifestyle education (A-R-E); and (3) resistance training – aerobic training – lifestyle education (R-A-E). Each intervention component lasted for 6 months, totaling up to 18 months. The sequence of the intervention components was designed to address the three research aims: (1) to evaluate the effectiveness of aerobic and resistance training at 6 months compared to lifestyle education, (2) to investigate the order effect of aerobic and resistance training at 12 months, and (3) to evaluate exercise habit formation over 18 months.

Randomization (1:1:1) was done by a research assistant using computer-generated random numbers stratified by frailty status. The person who was responsible for generating the random allocation sequence was not involved in any other parts of the research. Each group of consecutive participants was assigned to one of the 3 arms. Group allocation could not be masked for persons delivering or receiving the interventions. Our program exercise physiologists, who were responsible for developing the exercise protocols, training exercise coaches, and conducting the fitness tests with participants, were blind to the group assignment. Research assistants designated to assess outcomes were blind to group assignment and had no involvement in the delivery of interventions.

Intervention Program: GrandMove

The GrandMove program is a structured exercise training program that contains 2 protocols, one focusing on aerobic exercise and the other on resistance exercise. Two protocols for aerobic training and resistance training were designed by our program exercise physiologists with certified training in strength and conditioning and tailored for older adults with prefrailty and frailty in Hong Kong. Robust older people were trained as exercise coaches to deliver the exercise protocols. The design of the program was guided by social learning theory [12] and behavioral principles [13] (see Multimedia Appendix 1 for program characteristics). Both protocols were designed to be workable in a small home or group, using only small training implements such as rubber bands, towels, and water bottles for exercising. Each protocol had 5 levels, which indicate different levels of intensity. Based on the initial assessment, participants would start at a level that was realistic and attainable. A participant who reached a standard of fitness and strength in that level would continue with the next level.

The lifestyle education condition was comprised of 12 group sessions of health talks and 36 telehealth sessions. All health talks that covered different topics were delivered in a small group format by a retired nurse. Telehealth sessions involved a research assistant reviewing the health topics with participants and consolidating their understanding over the phone.

Each exercise intervention component lasted for 6 months (Multimedia Appendix 4). The 6-month schedule was designed to provide active coaching in the first 3 months, followed by monitoring and supervision (4th and 5th month), and self-sustained practice (6th month). Group sessions were provided in a small group format (8 - 10 older adults per group led by 2 coaches). Lifestyle education was designed to match the frequency of contacts with the exercise interventions.

Measures

The participants were assessed at baseline, 6 months, 12 months, and 18 months. Each assessment included a battery of self-report instruments administered through a structured interview and physical tests. Basic demographic characteristics were obtained, including age, gender, education level, marital status, and living arrangement.

Primary Outcomes

There were 3 primary outcomes. First, the frailty score was calculated using the 5-item FRAIL scale [2]. The items cover areas of fatigue, resistance, aerobic fitness, illnesses, and loss of weight. FRAIL scores are classified into three categories: robust (score of 0), prefrail (score of 1-2), and frail (score of 3-5). Second, physical performance was measured using the Short Physical Performance Battery (SPPB) [14]. The SPPB score ranges from 0 - 12, with a higher score indicating better performance. Third, quality of life was measured using the World Health Organization Quality of Life – Older adults module (WHOQoL-OLD) scale validated in a Chinese population [15]. The total score ranges from 0 - 100, with a higher score indicating better quality of life.

Secondary Outcomes

There were 3 fitness performance measures including isometric handgrip strength as measured by a digital hand dynamometer (Jamar Plus+), the 30-second arm curl test [16], and the 2-minute step test [17]. The hand grip strength of each hand was the average score of 3 trials. Higher fitness scores indicate better performance.

Other secondary outcomes included instrumental activities of daily living (IADL) as measured using the Lawton IADL scale (score range 0 - 100, with higher scores indicating a lower level of disability) [18], level of physical activity as measured using the Physical Activity Scale for the Elderly (PASE; higher scores indicate a higher level of physical activity) [19], sleep quality as measured using the Pittsburgh Sleep Quality Index (PSQI; higher scores indicate worse overall sleep quality) [20], social functioning as measured using the Lubben Social Network Scale (LSNS; higher scores indicate a greater level of social support) [21], and depressive symptoms as measured using the Patient Health Questionnaire-9 (PHQ-9; score range 0 - 27, with higher scores indicating greater symptom severity) [22].

Statistical Analysis

We conducted all analyses on an intention-to-treat basis. We generated descriptive statistics for baseline characteristics. The proportions of participants who achieved a robust or an improved frailty status (from frail to prefrail/robust or from prefrail to robust) at each follow-up were described. Within-group changes in proportions of robust or improved frailty outcome were tested using mixed effect logistic regression models.

The difference in outcome measures between intervention groups was estimated at each follow-up time point using repeated measures mixed effect linear regression models, adjusting for the fixed effects of gender and age at baseline. The main analysis using mixed models in the whole sample will

XSL•FO
RenderX

further adjust for other baseline covariates that were found to be significantly different between groups. To examine the effectiveness of 6-month aerobic or resistance training, outcomes in the A-R-E group and R-A-E group were compared to the E-R-A group. To investigate the order effect in aerobic and resistance training, we compared 12-month outcomes between the R-A-E group and the A-R-E group (reference group). To evaluate the maintenance of exercise effect and habit formation, we compared the outcomes including the level of physical activity between groups at 18 months. As intervention components were switched at the 6th and 12th month, time was treated as a categorical variable instead of continuous to account for the possible nonlinear relationship between time and outcome variables. Treatment effect referred to the coefficient of the intervention group × time interaction. Subgroup analyses by baseline frailty status and gender were individually made to delineate the differential treatment effects. Since 12 outcomes were examined in this study, we set the significance level at 1% and reported the 99% CIs for the primary outcomes to account for multiple comparison bias. The significance level for the secondary outcomes was set at 5%. All analyses were conducted using Stata/MP (version 17.0; StataCorp LLC).

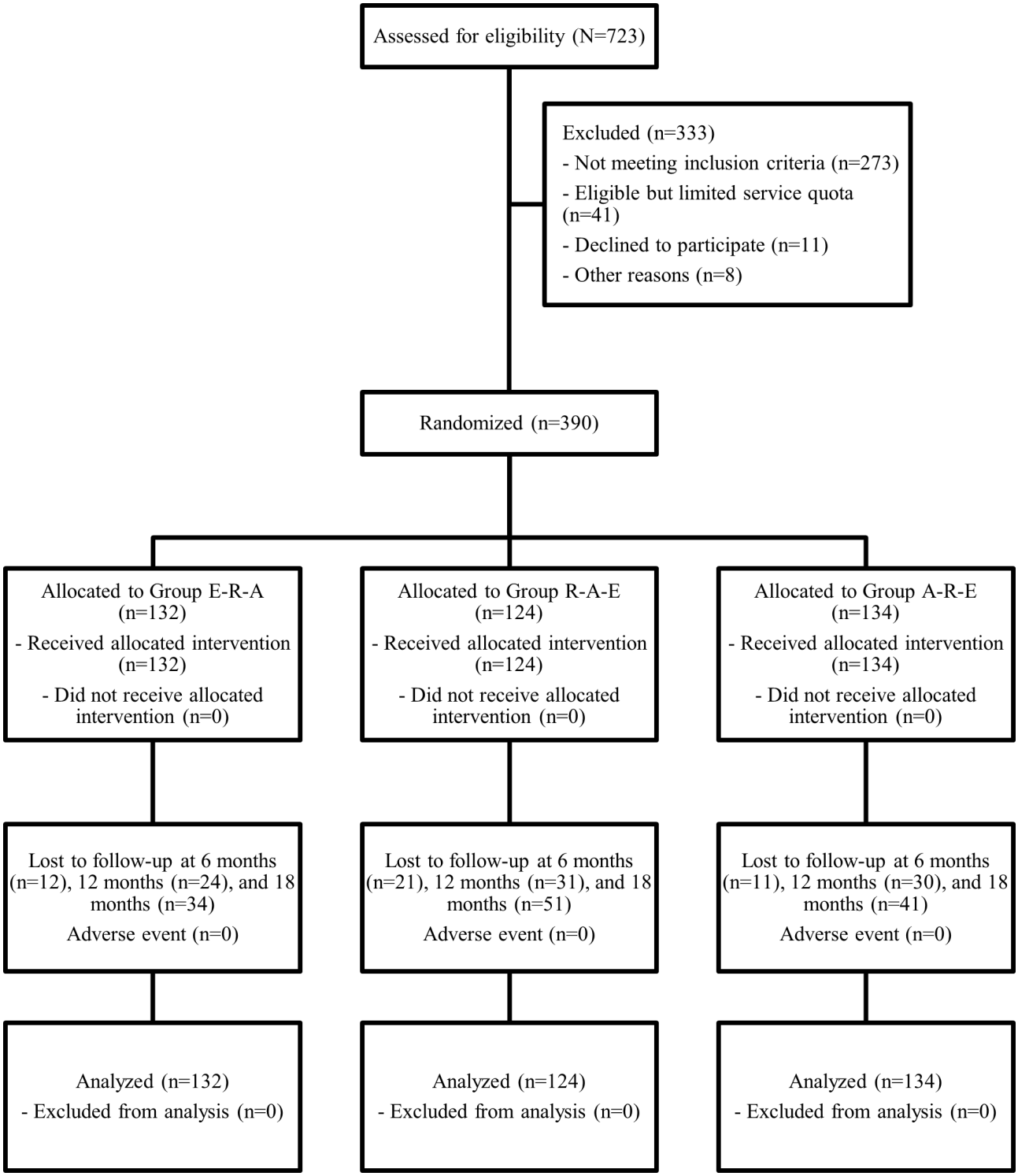
Results

Sample Characteristics

In total, 723 potential participants were assessed for eligibility, of whom 390 (53.9%) were randomized (Figure 1). Of the 390 participants included in the study, 132 (33.8%) were assigned to the E-R-A group, 124 (31.8%) to A-R-E, and 134 (34.4%) to R-A-E.

XSL•FO
RenderX

Figure 1. Study flow chart. All participants were analyzed in repeated measures mixed-effect regression models.



Baseline characteristics for participants by intervention group are shown in Table 1. The average FRAIL score, SPPB score, and overall quality of life at baseline were 2.1 (SD 0.9), 7.8 (SD 2.8), and 91.6 (SD 13.1), respectively. Characteristics of participants in the 3 groups were similar, except for the quality of life ($F_{2,385}$=3.26, P=.04) and arm curl test scores ($F_{2,382}$=5.47, P=.005). As suggested by the mean PASE scores, all 3 groups were considered as having low activity levels in general [23]. No reports of adverse events that led to death, hospitalization, or medical attendance were received by the end of the program.

XSL•FO
RenderX

Table . Baseline characteristics of the study population (n=390).

	E[a]-R[b]-A[c] group (n=132)	A-R-E group (n=124)	R-A-E group (n=134)	F/χ^2 (df)	*P* value
Age (years), mean (SD)	80.5 (7.5)	80.9 (7.1)	81.2 (7.4)	F=0.25 (2, 387)	.78
Gender, female, n (%)	106 (80.3)	97 (78.2)	105 (78.4)	χ^2=0.21 (2)	.90
Education, n (%)				χ^2=6.15 (6)	.41
No formal education	79 (59.9)	66 (53.2)	73 (54.5)		
Primary school	27 (20.5)	29 (23.4)	30 (22.4)		
Junior middle school	11 (8.3)	10 (8.1)	19 (14.2)		
High school or above	15 (11.4)	19 (15.3)	12 (9)		
Marital status, n (%)				χ^2=2.04 (4)	.73
Married	47 (35.6)	49 (40.2)	45 (33.8)		
Widowed	75 (56.8)	62 (50.8)	79 (59.4)		
Others	10 (7.6)	11 (9)	9 (6.8)		
Living alone	38 (28.8)	46 (37.4)	47 (35.3)	χ^2=2.34 (2)	.31
Frailty status, n (%)				χ^2=0.70 (2)	.70
Prefrail	80 (60.6)	80 (64.5)	80 (59.7)		
Frail	52 (39.4)	44 (35.5)	54 (40.3)		
5-item FRAIL[d] scale, mean (SD)	2.2 (0.9)	2.1 (0.9)	2.1 (1)	F=0.27 (2, 387)	.76
SPPB[e], mean (SD)	8 (2.7)	7.5 (2.9)	7.8 (2.9)	F=1.14 (2, 387)	.32
WHOQoL-OLD[f], mean (SD)	93.5 (12.3)	91.9 (12.4)	89.4 (14.1)	F=3.26 (2, 385)	.04
Grip strength (left hand), mean (SD)	34.5 (11.5)	33.6 (12.7)	33 (14.6)	F=0.46 (2, 381)	.63
Grip strength (right hand), mean (SD)	36.2 (12.6)	33.7 (13)	35.7 (14.9)	F=0.06 (2, 382)	.94
30-s arm curl test, mean (SD)	12.1 (4.3)	10.6 (4.3)	10.5 (4.1)	F=5.47 (2, 382)	.005
2-min step test, mean (SD)	69.2 (28.2)	65.3 (28.2)	66.6 (30.8)	F=0.60 (2, 382)	.55
IADL[g], mean (SD)	14.8 (2.7)	15.2 (3.1)	15.2 (2.7)	F=0.99 (2, 385)	.37
PASE[h], mean (SD)	70.6 (41.9)	73.1 (41.7)	72.3 (47.7)	F=0.11 (2, 382)	.90
LSNS[i], mean (SD)	21.1 (9.6)	21.8 (9.5)	21.4 (9.7)	F=0.14 (2, 385)	.87
PSQI[j], mean (SD)	8.3 (4.3)	8.1 (3.9)	8.3 (3.7)	F=0.10 (2, 356)	.91
PHQ-9[k], mean (SD)	4 (4.9)	3.8 (4.7)	3.3 (4.5)	F=0.86 (2, 384)	.42

[a]E: lifestyle education.

[b]R: resistance training.

[c]A: aerobic training.

[d]FRAIL: Fatigue, Resistance, Ambulation, Illness, and Loss of weight.

[e]SPPB: Short Physical Performance Battery.

[f]WHOQoL-OLD: Cantonese version of the World Health Organization Quality of Life – Older adults module.

[g]IADL: Lawton's Instrumental Activities of Daily Living scale.

[h]PASE: Physical Activity Scale for the Elderly.

[i]LSNS: Lubben Social Network Scale.

XSL•FO
RenderX

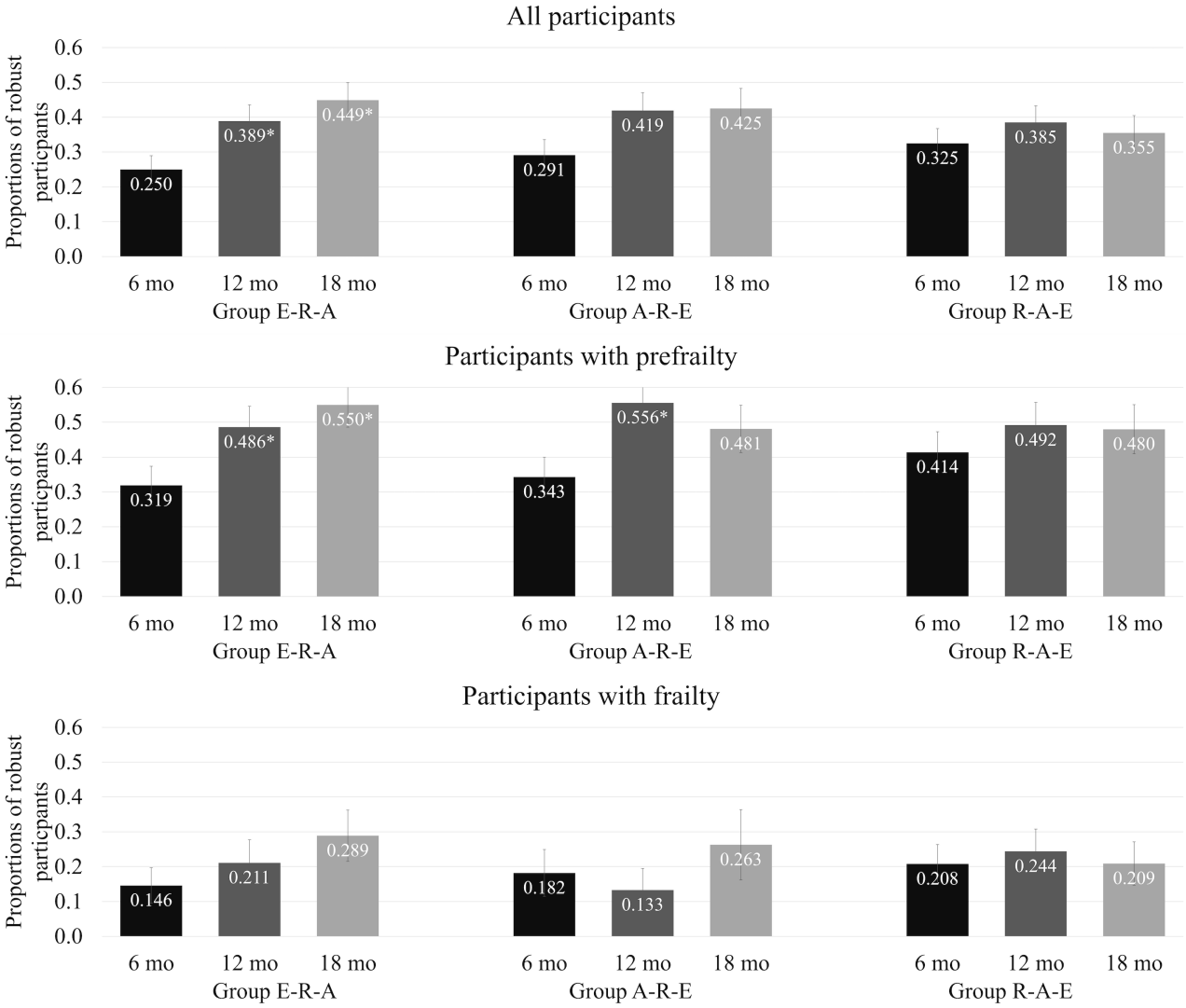
[j]PSQI: Pittsburgh Sleep Quality Index.

[k]PHQ-9: Patient Health Questionnaire-9.

Change in Frailty Status

We found a substantial overall improvement in frailty status (Figure 2 and Multimedia Appendices 5 and 6), both in terms of achieving a robust status and an improved status (from frail to prefrail/robust or from prefrail to robust).

Figure 2. Percentage of older participants obtaining robust status at the follow-up assessment time points. Absolute numbers are reported in Multimedia Appendix 5. The intervention only targeted older adults with frailty and prefrailty, and thus no robust participants were included at baseline. Asterisks (*) indicate a further significant increase (P<.05) in the proportion of robust status from 6 months. A: aerobic training; E: lifestyle education; R: resistance training.



In the prefrail sample, 35.8% (76/212) of all participants became robust after 6 months. The proportion of robust participants increased to 51.0% (98/192) at 12 months and remained stable (83/164, 50.6%) at 18 months. In the frail sample, 17.9% (24/134) of participants became robust at 6 months. The proportion increased to 20.4% (23/113) at 12 months and continually increased to 25% (25/100) at 18 months. In terms of within-group changes, participants from the E-R-A group and A-R-E group (prefrail sample only) continued to have a significant improvement in frailty beyond 6 months (Figure 2 and Multimedia Appendices 5 and 6). Although a smaller proportion of participants with frailty obtained a robust status within the first 6 months, the majority of them progressed to prefrailty.

Aerobic Versus Resistance Training: 6-Month Outcomes

As shown in Table 2, there was an overall improvement (having a significant time effect) in FRAIL score (coefficient −0.81, 95% CI −1.09 to –0.52; P<.001), IADL (coefficient 0.48, 95% CI 0.04 to 0.91; P=.03), PSQI (coefficient −0.74, 95% CI −1.34 to –0.14; P=.02), and PHQ-9 (coefficient −0.98, 95% CI −1.81 to −0.16; P=.02) in all groups over 6 months, suggesting that participants had improved frailty and functional and psychological outcomes.

XSL•FO
RenderX

Table . Summary of time and group × time interaction effects on primary and secondary outcomes (whole sample).

	Baseline to 6 months		Baseline to 12 months		Baseline to 18 months	
	Coefficient (95% or 99% CI)	*P* value	Coefficient (95% or 99% CI)	*P* value	Coefficient (95% or 99% CI)	*P* value
5-item FRAIL [a] scale						
Time effect	−0.81 (−1.09 to −0.52)	<.001	−1.10 (−1.40 to −0.80)	<.001	−1.06 (−1.36 to −0.75)	<.001
A-R-E[b] group × time	0.07 (−0.35 to 0.49)	.68	0.16 (−0.27 to 0.60)	.33	0.16 (−0.29 to 0.62)	.35
R-A-E group × time	−0.02 (−0.42 to 0.38)	.90	0.25 (−0.17 to 0.67)	.13	0.20 (−0.23 to 0.64)	.23
SPPB [c]						
Time effect	−0.22 (−0.66 to 0.31)	.20	−0.61 (−1.08 to −0.15)	.003	−0.64 (−1.11 to −0.16)	.002
A-R-E group × time	0.15 (−0.51 to 0.80)	.57	0.27 (−0.41 to 0.95)	.31	0.34 (−0.25 to 1.34)	.22
R-A-E group × time	0.03 (−0.60 to 0.65)	.91	0.04 (−0.62 to 0.70)	.88	−0.22 (−0.90 to 0.46)	.52
WHOQoL-OLD [d]						
Time effect	0.22 (−2.62 to 3.07)	.84	1.66 (−1.31 to 4.63)	.15	0.52 (−2.53 to 3.57)	.66
A-R-E group × time	3.34 (−0.85 to 7.52)	.04	−0.73 (−5.10 to 3.63)	.66	1.85 (−2.75 to 6.45)	.30
R-A-E group × time	5.23 (1.21 to 9.24)	.001	3.77 (−0.47 to 8.01)	.02	4.13 (−0.23 to 8.49)	.02
Grip strength (left)						
Time effect	−2.81 (−4.31 to −1.30)	<.001	−0.22 (−1.81 to 1.36)	.78	Not reported	
A-R-E group × time	2.99 (0.76 to 5.23)	.009	0.20 (−2.14 to 2.54)	.87		
R-A-E group × time	2.21 (0.63 to 4.36)	.04	−0.55 (−2.82 to 1.72)	.63		
Grip strength (right)						
Time effect	−3.12 (−4.62, to –1.63)	<.001	−1.12 (−1.12 to 0.45)	.16	Not reported	
A-R-E group × time	3.75 (1.54 to 5.97)	.001	1.06 (−1.25 to 3.38)	.37		
R-A-E group × time	2.29 (0.16 to 4.42)	.04	−0.50 (−2.75 to 1.75)	.66		
30-s arm curl						
Time effect	−0.77 (−0.78 to 0.63)	.83	0.12 (−0.62 to 0.85)	.76	0.12 (−0.63 to 0.88)	.75
A-R-E group × time	1.42 (0.39 to 2.46)	.007	1.46 (0.38 to 2.54)	.008	0.52 (−0.63 to 1.66)	.38
R-A-E group × time	1.11 (0.12 to 2.11)	.03	1.03 (−0.02 to 2.09)	.054	0.65 (−0.44 to 1.73)	.24
2-min step test						
Time effect	2.75 (−1.81 to 7.31)	.24	8.46 (3.69 to 13.24)	.001	7.99 (2.99 to 12.85)	.002
A-R-E group × time	5.51 (−1.23 to 12.25)	.11	2.49 (−4.57 to 9.54)	.49	1.10 (−6.32 to 8.52)	.77

XSL•FO
RenderX

	Baseline to 6 months		Baseline to 12 months		Baseline to 18 months	
	Coefficient (95% or 99% CI)	*P* value	Coefficient (95% or 99% CI)	*P* value	Coefficient (95% or 99% CI)	*P* value
R-A-E group × time	1.38 (−5.09 to 7.84)	.68	−3.17 (−9.99 to 3.66)	.36	−4.46 (−11.50 to 2.59)	.22
IADL [e]						
Time effect	0.48 (0.04 to 0.91)	.03	0.64 (0.19 to 1.09)	.006	0.74 (0.27 to 1.20)	.002
A-R-E group × time	−0.32 (−0.95 to 0.32)	.33	-0.77 (−1.43 to −0.11)	.02	-0.88 (−1.59 to −0.18)	.01
R-A-E group × time	−0.38 (−0.99 to 0.24)	.23	−1.09 (−1.74 to 0.45)	.001	−1.22 (−1.89 to −0.56)	<.001
PASE [h]						
Time effect	2.25 (−6.18 to 10.67)	.60	3.11 (−5.67 to 11.89)	.49	10.99 (2.06 to 19.90)	.02
A-R-E group × time	4.65 (−7.70 to 17.00)	.46	0.37 (−12.45 to 13.19)	.96	−3.29 (−16.72 to 10.14)	.63
R-A-E group × time	2.32 (−9.59 to 14.24)	.70	0.88 (−11.60 to 13.36)	.89	−3.00 (−15.76 to 9.77)	.64
LSNS [f]						
Time effect	0.89 (−0.66 to 2.44)	.26	2.43 (0.81 to 4.06)	.003	1.39 (−0.27 to 3.06)	.10
A-R-E group × time	−1.73 (−4.02 to 055)	.14	−1.82 (−4.21 to 0.56)	.13	−0.68 (−3.20 to 1.83)	.60
R-A-E group × time	−0.48 (−2.67 to 1.72)	.67	−1.28 (−3.60 to 1.04)	.28	0.23 (−2.15 to 2.62)	.85
PSQI [g]						
Time effect	−0.74 (−1.34 to −0.14)	.02	−0.96 (−1.59 to −0.33)	.003	0.17 (−0.48 to 0.81)	.61
A-R-E group × time	0.47 (−0.43 to 1.36)	.31	1.16 (0.23 to 2.09)	.02	−0.31 (−1.29 to 0.68)	.54
R-A-E group × time	0.58 (−0.28 to 1.44)	.19	0.94 (0.03 to 1.85)	.04	0.64 (−0.29 to 1.57)	.18
PHQ-9 [i]						
Time effect	−0.98 (−1.81 to −0.16)	.02	−1.14 (−2.00 to −0.27)	.01	−1.19 (−2.38 to −0.61)	.001
A-R-E group × time	0.43 (−0.79 to 1.65)	.49	0.49 (−0.78 to 1.76)	.45	0.69 (−0.65 to 2.03)	.31
R-A-E group × time	1.34 (0.17 to 2.52)	.02	1.51 (0.27 to 2.75)	.02	2.14 (0.87 to 3.41)	.001

[a]FRAIL: Fatigue, Resistance, Ambulation, Illness, and Loss of weight.

[b]A: aerobic training, R: resistance training, E: lifestyle education.

[c]SPPB: Short Physical Performance Battery.

[d]WHOQoL-OLD: Cantonese version of the World Health Organization Quality of Life - Older adults module.

[e]IADL: Lawton's Instrumental Activities of Daily Living scale.

[f]LSNS: Lubben Social Network Scale.

[g]PASE: Physical Activity Scale for the Elderly.

[h]PSQI: Pittsburgh Sleep Quality Index.

[i]PHQ-9: Patient Health Questionnaire-9.

[j]Measurements reporting 99% CI: 5-item FRAIL scale, SPPB, and WHOQoL-OLD. All other scales used 95% CI.

At 6 months, participants with resistance training (R-A-E group) compared with those with lifestyle education (E-R-A group) had greater improvement in quality of life (interaction coefficient 5.23, 99% CI 1.21 to 9.24; *P*<.001) but not in FRAIL

XSL•FO
RenderX

score and SPPB. Similar findings were observed in the prefrail subsample (Multimedia Appendix 7). Aerobic training (A-R-E group) was not different from lifestyle education in the primary outcomes.

For the secondary outcomes, both aerobic training (A-R-E group) and resistance training (R-A-E group) performed better in several fitness outcomes than the E-R-A group. The A-R-E group was associated with greater improvement in left hand grip strength (interaction coefficient 2.99, 95% CI 0.76 to 5.23; *P*=.009), right hand grip strength (interaction coefficient 3.75, 95% CI 1.54 to 5.97; *P*=.001), and the arm curl test (interaction coefficient 1.42, 95% CI 0.39 to 2.46; *P*=.007) than lifestyle education (E-R-A group). The R-A-E group was also associated with greater improvement in the 3 outcomes (Table 2). The A-R-E and R-A-E groups did not differ from the E-R-A group in all other nonfitness measures including the 2-minute step test, IADL, PASE, LSNS, and PSQI over 6 months. The R-A-E group had a significantly higher level of PHQ-9 scores (interaction coefficient 1.34, 95% CI 0.17 to 2.52; *P*=.02). Results of the subgroup analysis by frailty status were summarized in Multimedia Appendices 7 and 8, whereas results of the subgroup analysis by gender were summarized in Multimedia Appendices 9 and 10.

Aerobic and Resistance Training: The Order Effect (12-Month Outcomes)

To compare the order effect, we also compared the R-A-E group with the A-R-E group (reference group) on 12-month outcomes (Multimedia Appendix 11). The 2 groups did not differ in all outcomes except that the R-A-E group was associated with greater improvement in WHOQoL-OLD (interaction coefficient 4.50, 95% CI 0.12 to 8.88; *P*=.008). The result was similar when the analysis was applied in the prefrail subsample. Participants with frailty in the R-A-E group achieved significantly fewer steps in the 2-minute step test compared to their counterparts in the A-R-E group at 12 months (interaction coefficient −12.48, 95% CI −24.78 to −0.18; *P*=.047).

In sum, the findings suggested that the order of aerobic and resistance training did not have an impact on fitness outcome at 12 months, although some preliminary data suggested that undergoing resistance training before aerobic training might improve quality of life.

Maintenance of Exercise Effect and Habit Formation

The A-R-E and R-A-E groups did not receive a physical intervention between 12 and 18 months. If exercise effects were maintained, we expected to observe gains to be maintained by the end of the study. As shown in Table 2, the improvement in WHOQoL-OLD seemed to be maintained at 18 months in these 2 groups compared to the E-R-A group but the *P* value did not reach the .01 threshold (interaction coefficient 4.13, 95% CI −0.23 to 8.49; *P*=.02).

However, the gains in IADL appeared to reduce in the A-R-E (interaction coefficient −0.88, 95% CI −1.59 to −0.18; *P*=.013) and the R-A-E groups (interaction coefficient −1.22, 95% CI −1.89 to −0.56; *P*<.001) and also for PHQ-9 in the R-A-E group (interaction coefficient 2.14, 95% CI 0.87 to 3.41; *P*=.001).

Neither aerobic training nor resistance training improved PASE at 6 months compared to lifestyle education. We observed a significant time effect on PASE over 18 months in the whole sample (coefficient 10.99, 95% CI 2.06 to 19.90; *P*=.02) and in the frail subsample (coefficient 14.6, 95% CI 1.39 to 27.81; *P*=.03), indicating that combined exercise and lifestyle education, regardless of intervention sequence, might improve individuals' level of physical activity.

Discussion

Principal Findings

This study examined the effectiveness of both aerobic training and resistance training using a sequential cluster randomized controlled trial design, which enabled us to examine how these physical activity interventions and their sequence might influence frailty outcomes. Although lifestyle education was initially added as a comparator intervention, our findings have demonstrated that the 18-month intervention combining lifestyle education and physical interventions is a viable strategy to address frailty and is health-promoting. The frailty score improved across groups, providing further evidence that physical frailty is reversible. Our study also produced a few other key findings. Aerobic training or resistance training tended to improve fitness performance but was not superior to an intensive lifestyle education program in addressing frailty. It is possible that 6 months of resistance training followed by another 6 months of aerobic training might be better at improving quality of life. Although the overall improvement in frailty outcome was maintained over 18 months, IADL and PHQ-9 appeared to worsen over the post–physical intervention period. We observed a small increase in the level of physical activity over 18 months, but its relation to the formation of exercise habits and an active lifestyle is yet to be determined.

Despite physical activity intervention being the most widely studied and recommended approach for the management of frailty [24], we found that a single-mode exercise program (aerobic or resistance) for 6 months was no better than lifestyle education in addressing frailty. The findings coincided with a recent systematic review suggesting that physical activity intervention, when compared with an active control intervention, was not associated with a significant reduction in frailty [25]. We found that the administration of 2 physical interventions in sequence did not further improve frailty scores but participants were able to maintain the gain accrued over the first 6 months. The 3 study arms approximated a multifaceted intervention that incorporated physical, psychosocial, and educational components. There were a few multifaceted studies [26-28] but cross-study comparison was difficult due to their different designs and the lack of frailty as an outcome. The Hatoyama Cohort Study [27] found that resistance exercise in combination with nutritional education and psychosocial programs for 3 months successfully reduced the prevalence of frailty by 24%. Our study further added that the sequence of intervention components, if not delivered concurrently, might not make a significant difference. Nonetheless, older adults with frailty may find the program more acceptable if they have greater control over the sequence of interventions.

XSL•FO
RenderX

We added to the literature that psychosocial intervention might be complementary to physical activity intervention for the management of frailty. Following the lifestyle education program, the subsequent addition of physical intervention appeared to generate further gain. Participants from the E-R-A group were observed to have continued improvement in frailty outcomes after switching from lifestyle education to resistance training. There were only a few known studies to date exploring the effectiveness of psychosocial interventions on frailty. A Swedish study evaluated a 4-week senior group program similar to our lifestyle education program, which introduced and discussed various healthy lifestyle topics [29]. The senior group program had no impact on frailty outcomes, but it was reported to delay activities of daily living deterioration. Our lifestyle education program was more intensive in terms of frequency and duration than the senior group program. Further studies may explore whether shorter or longer combined interventions may make a difference in frailty outcomes.

Since frailty is suggested to result from cumulative declines in multiple physiologic systems [1], it is not surprising that interventions targeting different systems may yield similar positive results (ie, the equifinality principle). There might be common factors across the intervention approaches that mitigate frailty, such as social support from peers (exercise coaches and retired nurses). In our study, all 3 intervention components encouraged social engagement through the group sessions and peer coaching. Previous studies have suggested that social support might directly or indirectly contribute to higher levels of physical activity in older adults [30-32] and could therefore be health-promoting and lead to a healthier lifestyle. Further studies should include a comparison group without social influence and a measure of perceived social support to allow for an estimation of its impact on frailty and quality of life. Future studies may also include measures of other psychological mediators such as self-efficacy, enjoyment, and the use of behavioral and cognitive processes [33] for mediation analyses.

The subgroup analysis by frailty status showed that participants with frailty appeared to have a greater magnitude of reduction in FRAIL score, further demonstrating that frailty is a modifiable condition. The frail subgroup also had consistently reported increased social engagement and some reduction in depressive symptoms over 18 months. Some improvements in physical fitness were observed but may not have been sustained beyond the active intervention period. The results echoed our hypothesis that active physical or psychosocial interventions might enhance social support for older adults at high risk of vulnerability and social isolation [34]. Physical activity guidelines should recommend older adults with frailty to engage in moderate-intensity exercise regularly [35]. The design of physical activity interventions should incorporate social or group elements that enhance social learning and reinforcement. With a sample predominantly comprised of female older adults (~80%), the results for the female subsample were largely consistent with those in the full sample. There may be a lack of power to detect significant intervention effects in the male subsample. These results should be interpreted with caution.

We observed a significant increase in the level of physical activity over 18 months, probably due in part to the significant change within the frail subsample. No evidence showed that participants with prefrailty increased their level of physical activity after interventions. It might take as long as 18 months to observe an overall increase in the level of physical activity. The potential of exercise and lifestyle interventions to modify long-term exercise habits remains unclear. It is common to observe a relapse pattern in health behavior once an intervention has ended [36]. Program effectiveness and acceptability are equally important as pain and discomfort associated with physical activity could be an obstacle to engaging in further physical activity [37]. In addition, a physical intervention combined with smartphone-assisted e-reminders, an activity tracker, and e-coaching may help habit formation [38]. It is equally important to implement strategies to discourage unhealthy habits simultaneously. It is possible that our lifestyle education program helped increase participants' awareness about unhealthy habits and alternative healthy options in the environment. Further investigation to determine the optimal form of intervention(s) that best maintains exercise habit formation is warranted.

We consider the delivery of exercise interventions by peer coaches to be the key to long-term scalability and sustainability in the context of population aging. More research is needed to formally evaluate the sustainability and cost-effectiveness of the intervention model. Future research should consider the productive aging component for engaging retired older adults in productive activities, task shifting from health professionals to more available human resources (ie, peer coaches), and the implications of an increased number of robust older people in the community. Health care utilization associated with a positive change in frailty outcomes has not been explored, but it is worth further investigation.

Limitations

Due to the complexity of the study design, we lacked a care-as-usual comparator group. The active control group only abstained from physical intervention in the first 6 months. Therefore, we were unable to evaluate the effectiveness of physical activity interventions for 12 months compared with an active control intervention. Further multifaceted intervention studies may include both psychoeducational components in combination with physical interventions. An accelerometer-based measure of physical activity level may be better than a self-reported one in estimating the effect of a physical intervention on habit formation [39]. Similar to other physical intervention programs, the generalization of results might be limited by attrition bias and a predominantly female sample. This program was tested in a densely populated community. This setting may enhance the viability and cost-effectiveness of delivering both group and home sessions via peer coaches. Adaptations may be needed if the model is applied in other contexts, for example, in rural areas, which will require a separate evaluation of feasibility and effectiveness.

Conclusions

In contrast to most previous trials, we attempted to address frailty by examining physical interventions in comparison with a novel comparator, that is, lifestyle education. Both tested physical interventions and lifestyle education were effective in

XSL•FO
RenderX

improving frailty status. A simultaneous improvement in quality of life was observed. Participants with frailty appeared to benefit beyond a frailty outcome. Despite the positive findings, the impact of the interventions on exercise habit formation is yet to be investigated.

Acknowledgments

This work was supported by theSimon KY Lee Foundation, Hong Kong. The funding body had no role in the design, methods, subject recruitment, data collection, analysis, or preparation of the paper.

Authors' Contributions

All authors, led by GHYW and TYSL, developed the study concept and design. MT, TYSL, and GHYW were directly involved in activities related to the acquisition of data. JYMT, HL, and GHYW conducted data analysis, and all authors participated in the interpretation of the data. JYMT, HL, and GHYW drafted the manuscript. All authors critically revised the manuscript for important intellectual content.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Program design.
[DOCX File, 36 KB - aging_v8i1e65636_app1.docx]

Multimedia Appendix 2
Additional information about study methods and attendance.
[DOCX File, 32 KB - aging_v8i1e65636_app2.docx]

Multimedia Appendix 3
Components and sequence of the 3-arm trial.
[PNG File, 72 KB - aging_v8i1e65636_app3.png]

Multimedia Appendix 4
Schedule for each intervention component.
[DOCX File, 33 KB - aging_v8i1e65636_app4.docx]

Multimedia Appendix 5
Percentage of participants achieving robust status (from prefrail or frail to robust).
[DOCX File, 37 KB - aging_v8i1e65636_app5.docx]

Multimedia Appendix 6
Percentage of older participants obtaining an improved frailty status at the follow-up assessment time points. The intervention only targeted older adults with prefrailty and frailty, and thus no robust participants were included at baseline. Asterisks (*) indicate a further significant increase (P<.05) in the proportion of improved frailty status from 6 months.
[PNG File, 1647 KB - aging_v8i1e65636_app6.png]

Multimedia Appendix 7
Summary of time and group × time interaction effects on primary and secondary outcomes (participants with prefrailty only).
[DOCX File, 41 KB - aging_v8i1e65636_app7.docx]

Multimedia Appendix 8
Summary of time and group × time interaction effects on primary and secondary outcomes (participants with frailty only).
[DOCX File, 43 KB - aging_v8i1e65636_app8.docx]

Multimedia Appendix 9
Summary of time and group × time interaction effects on primary and secondary outcomes (female participants only).
[DOCX File, 44 KB - aging_v8i1e65636_app9.docx]

XSL•FO
RenderX

Multimedia Appendix 10
Summary of time and group × time interaction effects on primary and secondary outcomes (male participants only).
[DOCX File, 44 KB - aging_v8i1e65636_app10.docx]

Multimedia Appendix 11
Comparison of outcomes between the R-A-E group and A-R-E group (reference) at the 12-month follow-up.
[DOCX File, 36 KB - aging_v8i1e65636_app11.docx]

Checklist 1
CONSORT checklist. CONSORT: Consolidated Standards of Reporting Trials.
[PDF File, 100 KB - aging_v8i1e65636_app12.pdf]

References

1. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
2. Morley JE, Malmstrom TK, Miller DK. A simple frailty questionnaire (FRAIL) predicts outcomes in middle aged African Americans. J Nutr Health Aging 2012 Jul;16(7):601-608. [doi: 10.1007/s12603-012-0084-2] [Medline: 22836700]
3. O'Caoimh R, Sezgin D, O'Donovan MR, et al. Prevalence of frailty in 62 countries across the world: a systematic review and meta-analysis of population-level studies. Age Ageing 2021 Jan 8;50(1):96-104. [doi: 10.1093/ageing/afaa219] [Medline: 33068107]
4. Vermeiren S, Vella-Azzopardi R, Beckwée D, et al. Frailty and the prediction of negative health outcomes: a meta-analysis. J Am Med Dir Assoc 2016 Dec 1;17(12):1163. [doi: 10.1016/j.jamda.2016.09.010] [Medline: 27886869]
5. Merchant RA, Morley JE, Izquierdo M. Editorial: Exercise, aging and frailty: guidelines for increasing function. J Nutr Health Aging 2021;25(4):405-409. [doi: 10.1007/s12603-021-1590-x] [Medline: 33786554]
6. Travers J, Romero-Ortuno R, Bailey J, Cooney MT. Delaying and reversing frailty: a systematic review of primary care interventions. Br J Gen Pract 2019 Jan;69(678):e61-e69. [doi: 10.3399/bjgp18X700241] [Medline: 30510094]
7. 2018 Physical Activity Guidelines Advisory Committee scientific report. Office of Disease Prevention and Health Promotion. 2018. URL: https://health.gov/sites/default/files/2019-09/PAG_Advisory_Committee_Report.pdf [accessed 2025-04-30]
8. Izquierdo M, Merchant RA, Morley JE, et al. International exercise recommendations in older adults (ICFSR): expert consensus guidelines. J Nutr Health Aging 2021;25(7):824-853. [doi: 10.1007/s12603-021-1665-8] [Medline: 34409961]
9. de Labra C, Guimaraes-Pinheiro C, Maseda A, Lorenzo T, Millán-Calenti JC. Effects of physical exercise interventions in frail older adults: a systematic review of randomized controlled trials. BMC Geriatr 2015 Dec 2;15(154):154. [doi: 10.1186/s12877-015-0155-4] [Medline: 26626157]
10. Tseng BS, Marsh DR, Hamilton MT, Booth FW. Strength and aerobic training attenuate muscle wasting and improve resistance to the development of disability with aging. J Gerontol A Biol Sci Med Sci 1995 Nov;50 Spec No:113-119. [doi: 10.1093/gerona/50a.special_issue.113] [Medline: 7493203]
11. Report of the strategic review on healthcare manpower planning and professional development. Health Bureau. 2017. URL: https://www.healthbureau.gov.hk/en/press_and_publications/otherinfo/180500_sr/srreport.html [accessed 2025-04-30]
12. Bandura A. Social Learning Theory: Prentice-Hall; 1977.
13. Lally P, Gardner B. Promoting habit formation. Health Psychol Rev 2013 May;7(sup1):S137-S158. [doi: 10.1080/17437199.2011.603640]
14. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994 Mar;49(2):M85-M94. [doi: 10.1093/geronj/49.2.m85] [Medline: 8126356]
15. Liu R, Wu S, Hao Y, et al. The Chinese version of the World Health Organization Quality of Life instrument-older adults module (WHOQOL-OLD): psychometric evaluation. Health Qual Life Outcomes 2013 Sep 16;11(156):24034698. [doi: 10.1186/1477-7525-11-156] [Medline: 24034698]
16. James TW. The 30-second arm curl test as an indicator of upper body strength in older adults [Dissertation]. : California State University; 1999 URL: https://www.proquest.com/dissertations-theses/30-second-arm-curl-test-as-indicator-upper-body/docview/304580972/se-2 [accessed 2025-04-30]
17. Rikli RE, Jones CJ. Functional fitness normative scores for community-residing older adults, ages 60-94. J Aging Phys Act 1999;7(2):162-181. [doi: 10.1123/japa.7.2.162]
18. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179-186. [Medline: 5349366]
19. Washburn RA, McAuley E, Katula J, Mihalko SL, Boileau RA. The physical activity scale for the elderly (PASE): evidence for validity. J Clin Epidemiol 1999 Jul;52(7):643-651. [doi: 10.1016/s0895-4356(99)00049-9] [Medline: 10391658]

XSL•FO
RenderX

20. Buysse DJ, Reynolds CF 3rd, Monk TH, Hoch CC, Yeager AL, Kupfer DJ. Quantification of subjective sleep quality in healthy elderly men and women using the Pittsburgh Sleep Quality Index (PSQI). Sleep 1991 Aug;14(4):331-338. [doi: 10.1093/sleep/14.4.331] [Medline: 1947597]
21. Lubben J, Blozik E, Gillmann G, et al. Performance of an abbreviated version of the Lubben Social Network Scale among three European community-dwelling older adult populations. Gerontologist 2006 Aug;46(4):503-513. [doi: 10.1093/geront/46.4.503] [Medline: 16921004]
22. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
23. Laddu D, Parimi N, Cauley JA, et al. The association between trajectories of physical activity and all-cause and cause-specific mortality. J Gerontol A Biol Sci Med Sci 2018 Nov 10;73(12):1708-1713. [doi: 10.1093/gerona/gly037] [Medline: 29529273]
24. Fit for frailty: consensus best practice guidance for the care of older people living with frailty in community and outpatient settings. : British Geriatrics Society; 2014 URL: https://www.bgs.org.uk/sites/default/files/content/resources/files/2018-05-14/fff2_short_0.pdf [accessed 2025-04-30]
25. Negm AM, Kennedy CC, Thabane L, et al. Management of frailty: a systematic review and network meta-analysis of randomized controlled trials. J Am Med Dir Assoc 2019 Oct;20(10):1190-1198. [doi: 10.1016/j.jamda.2019.08.009] [Medline: 31564464]
26. Oh SL, Kim HJ, Woo S, et al. Effects of an integrated health education and elastic band resistance training program on physical function and muscle strength in community-dwelling elderly women: Healthy Aging and Happy Aging II study. Geriatr Gerontol Int 2017 May;17(5):825-833. [doi: 10.1111/ggi.12795] [Medline: 27138245]
27. Seino S, Nishi M, Murayama H, et al. Effects of a multifactorial intervention comprising resistance exercise, nutritional and psychosocial programs on frailty and functional health in community-dwelling older adults: a randomized, controlled, cross-over trial. Geriatr Gerontol Int 2017 Nov;17(11):2034-2045. [doi: 10.1111/ggi.13016] [Medline: 28393440]
28. Spoorenberg SLW, Wynia K, Uittenbroek RJ, Kremer HPH, Reijneveld SA. Effects of a population-based, person-centred and integrated care service on health, wellbeing and self-management of community-living older adults: a randomised controlled trial on Embrace. PLoS One 2018;13(1):e0190751. [doi: 10.1371/journal.pone.0190751] [Medline: 29351295]
29. Gustafsson S, Wilhelmson K, Eklund K, et al. Health-promoting interventions for persons aged 80 and older are successful in the short term--results from the randomized and three-armed Elderly Persons in the Risk Zone study. J Am Geriatr Soc 2012 Mar;60(3):447-454. [doi: 10.1111/j.1532-5415.2011.03861.x] [Medline: 22409735]
30. Lieber SB, Moxley J, Mandl LA, Reid MC, Czaja SJ. Social support and physical activity: does general health matter? Eur Rev Aging Phys Act 2024 Jun 20;21(1):16. [doi: 10.1186/s11556-024-00347-6] [Medline: 38902616]
31. Resnick B, Orwig D, Magaziner J, Wynne C. The effect of social support on exercise behavior in older adults. Clin Nurs Res 2002 Feb;11(1):52-70. [doi: 10.1177/105477380201100105] [Medline: 11845515]
32. Smith GSE, Moyle W, Burton NW. The relationship between social support for physical activity and physical activity across nine years in adults aged 60-65 years at baseline. Int J Environ Res Public Health 2023 Mar 3;20(5):36901538. [doi: 10.3390/ijerph20054531] [Medline: 36901538]
33. Lewis BA, Marcus BH, Pate RR, Dunn AL. Psychosocial mediators of physical activity behavior among adults and children. Am J Prev Med 2002 Aug;23(2 Suppl):26-35. [doi: 10.1016/s0749-3797(02)00471-3] [Medline: 12133735]
34. Lindsay Smith G, Banting L, Eime R, O'Sullivan G, van Uffelen JGZ. The association between social support and physical activity in older adults: a systematic review. Int J Behav Nutr Phys Act 2017 Apr 27;14(1):56. [doi: 10.1186/s12966-017-0509-8] [Medline: 28449673]
35. Izquierdo M, Duque G, Morley JE. Physical activity guidelines for older people: knowledge gaps and future directions. Lancet Healthy Longev 2021 Jun;2(6):e380-e383. [doi: 10.1016/S2666-7568(21)00079-9] [Medline: 36098146]
36. Wood W, Neal DT. Healthy through habit: interventions for initiating & maintaining health behavior change. Behavioral Science & Policy 2016 Apr;2(1):71-83. [doi: 10.1177/237946151600200109]
37. Wang Z, Fang Y, Chan PSF, Yu FY, Sun F. The changes in levels and barriers of physical activity among community-dwelling older adults during and after the fifth wave of COVID-19 outbreak in Hong Kong: repeated random telephone surveys. JMIR Aging 2023 Jan 23;6:e42223. [doi: 10.2196/42223] [Medline: 36599172]
38. Kwan RY, Lee D, Lee PH, et al. Effects of an mHealth brisk walking intervention on increasing physical activity in older people with cognitive frailty: pilot randomized controlled trial. JMIR mHealth uHealth 2020 Jul 31;8(7):e16596. [doi: 10.2196/16596] [Medline: 32735218]
39. Valenti G, Bonomi AG, Westerterp KR. Walking as a contributor to physical activity in healthy older adults: 2 week longitudinal study using accelerometry and the doubly labeled water method. JMIR mHealth uHealth 2016 Jun 7;4(2):e56. [doi: 10.2196/mhealth.5445] [Medline: 27268471]

Abbreviations

A-R-E: aerobic training - resistance training - lifestyle education
E-R-A: lifestyle education - resistance training - aerobic training
FRAIL: Fatigue, Resistance, Ambulation, Illnesses, and Loss of weight

XSL•FO
RenderX

IADL: Instrumental Activities of Daily Living
LSNS: Lubben Social Network Scale
PASE: Physical Activity Scale for the Elderly
PHQ-9: Patient Health Questionnaire-9
PSQI: Pittsburgh Sleep Quality Index
R-A-E: resistance training - aerobic training - lifestyle education
SPPB: Short Physical Performance Battery
WHOQoL-OLD: World Health Organization Quality of Life – Older adults module

Edited by M Gray; submitted 21.08.24; peer-reviewed by N Jiang, Y Zhou; revised version received 11.03.25; accepted 25.03.25; published 14.05.25.

Please cite as:
Tang JYM, Luo H, Tse M, Kwan J, Leung AYM, Tsien Wong TBK, Lum TYS, Wong GHY
Frailty, Fitness, and Quality of Life Outcomes of a Healthy and Productive Aging Program (GrandMove) for Older Adults With Frailty or Prefrailty: Cluster Randomized Controlled Trial
JMIR Aging 2025;8:e65636
URL: https://aging.jmir.org/2025/1/e65636
doi:10.2196/65636

© Jennifer Yee Man Tang, Hao Luo, Michael Tse, Joseph Kwan, Angela Yee Man Leung, Teresa Bik-kwan Tsien Wong, Terry Yat-sang Lum, Gloria Hoi-yan Wong. Originally published in JMIR Aging (https://aging.jmir.org), 14.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Principal Component Analysis of Stair Negotiation and Floor Transition Kinematics in Older Adults With and Without Functional Disability: Cross-Sectional Study

Juliana Moreira[1,2], PhD; Ivone da Silva Teles[1], MSc; Bruno Cunha[3], PhD; José Félix[1,4,5], MSc; Diana C Guedes[1], MSc; Leonel A T Alves[1], MSc; Rubim Santos[1], PhD; Andreia S P Sousa[1], PhD

[1]CIR, E2S, Polytechnic of Porto, Rua Dr. António Bernardino de Almeida nº 400, Porto, Portugal

[2]Research Center in Physical Activity, Health and Leisure, Faculty of Sports, University of Porto, Porto, Portugal

[3]CINTESIS@RISE, CINTESIS.UPT, Portucalense University, Porto, Portugal

[4]Department of Physiotherapy, Institute of Health of the North - Escola Superior de Saúde do Vale do Ave (ESSVA), Cooperativa de Ensino Superior Politécnico e Universitário (CESPU), Vila Nova de Famalicão, Portugal

[5]Department of Medical Sciences, University of Aveiro, Aveiro, Portugal

Corresponding Author:
Andreia S P Sousa, PhD
CIR, E2S, Polytechnic of Porto, Rua Dr. António Bernardino de Almeida nº 400, Porto, Portugal

Abstract

Background: Stair negotiation (ascending and descending) and transitions to level walking are complex motor tasks influenced by aging; yet the impact of functional disability on these changes remains underexplored.

Objective: This study aimed to evaluate the lower limb joint positions, velocities, and the center of mass (CoM) displacement and velocity during stair negotiation and transitions in older adults with and without functional disability.

Methods: Sixty community-dwelling adults, aged 60 years and older, were assessed for lower limb tridimensional joint positions and velocities during the instances of foot contact or leaving the step and foot contact or leaving the floor; the joint range of motion, angular velocity range, and the CoM displacement and angular velocity range were assessed during stair ascend and descend and transitions floor-to-stair and stair-to-floor through an optoelectronic system. Principal component analysis was used to assess 8 groups of variables to compute principal component models (I-VIII). Participants were classified as with or without disability based on functional disability indicators. Group differences were assessed using the Mann-Whitney *U* test.

Results: From 240 variables, 41 key parameters were identified, mainly related to hip and knee angular velocities in the sagittal plane. Significant differences between older adults with disability (n=25) and without disability (n=35) were found in 12 principal components.

Conclusions: Older adults with functional disability showed changes in the sagittal plane hip as well as in the knee angular velocity and mediolateral and vertical CoM displacement and velocity during stair negotiation and transitions. These findings can inform targeted strategies to improve mobility and stability in this population.

Trial Registration: ClinicalTrials.gov NCT05611723; https://clinicaltrials.gov/study/NCT05611723

(JMIR Aging 2025;8:e71530) doi:10.2196/71530

KEYWORDS

disability; biomechanics; stair negotiation; gerontology; kinematics; factor analysis

Introduction

Disability, as described by the International Classification of Functioning, Disability, and Health, embraces the impairments, activity limitations, and participation restrictions [1]. It is a condition that requires assistance to overcome reliance or difficulty with everyday tasks and independent living [2]. Aging is associated with systemic changes in the body, including deficits in cardiorespiratory function, slowness, strength loss, reduced flexibility, and weight loss, resulting in a decreased ability to perform daily activities [3,4]. These alterations increase the risk of mobility limitations [4] and physical disability [2] in older adults, making it crucial to reduce the length of handicap time and increase the well-being and independent time [5]. Research and development investment is fundamental for effective intervention strategies, especially in the rapidly aging global population, requiring policy makers to anticipate these changes [5]. Accordingly, to delay disability or prevent its occurrence, a multidimensional assessment needs to be performed [6]. One hallmark of disability and functional

XSL•FO
RenderX

decline in older adults is related to impairments in stairs negotiation, the act of ascending or descending stairs [7]. These tasks are associated with a high risk of mortality and serious injuries [8], with estimates that 20% of falls occur on stairs [9], making them 10 times more dangerous than walking [10]. The fact that this is one of the most hazardous daily activities is justified by its motor complexity, which, requiring precise coordination of joint movements, muscle forces, and balance control, poses a significant challenge for older adults [11], particularly those with functional disability [12].

The association between aging and biomechanical performance during stair negotiation has been reported in previous studies [9,10,13-15]. Specifically, older adults seem to present slower cadence and higher variability of hip joint angle curves in the sagittal plane during stair ascent and descent [13,16]. They also present higher maximum dorsiflexion angles [17] and increased time to descend stairs, which has been considered a predictor of functional decline [18]. A reduction of peak downwards and upwards center of mass (CoM) acceleration during the descent and landing phases, respectively [19], was also identified in older adults' performance.

While age-related biomechanical adaptations in stair negotiation are well-documented, the extent to which these are associated with disability, as opposed to healthy aging, remains underexplored. Moreover, stair negotiation assessment should also encompass the sequences of walking to ascend stairs and descending stairs to walk, to represent the complete range of this daily activity. Therefore, it is important to evaluate not only stair use (stair ascent and descent) but also floor transitions (floor-to-stair [F-S] and stair-to-floor [S-F]). During stair use and transitions, beyond the range of kinematic parameters during full cycles of stair ascent and descent as well as F-S and S-F, the assessment of key events could bring relevant data because at these transitional moments, older adults demonstrated more impairments in margin of stability [20], lower downwards CoM velocity [19], and lower peak moments during stair ascent and descent [17]. These changes lead to compensatory strategies prioritizing stability over range of motion (ROM), particularly when making or leaving contact with the step [17]. Additionally, anatomical changes related to aging such as sagittal malalignment, were associated with ascending and descending stairs with increased thorax flexion, increasing the fall risk [21], and previous research highlighted the roles of pelvis and head kinematics [22].

Addressing this research gap can provide valuable insights into the biomechanical contributors to stair negotiation impairments. Moreover, the lack of research is even more pronounced in studies comparing older adults with functional disability to those experiencing healthy aging. Such investigations can inform the development of targeted interventions to improve stair performance, reduce fall risk, and enhance functional independence in older populations.

Principal component analysis (PCA), a multivariate and nonparametric statistical technique, reduces the dimensionality of biomechanical data by generating a new set of uncorrelated, orthogonal variables known as principal components (PCs) [23,24]. It has been used to analyze biomechanical data from older adults during stair negotiation, primarily in comparisons with younger adults [14].

Following this rationale, this study examines the kinematic differences in lower limb joint position and velocity at the instant of contact and leaving the step, the joint ROM, the variation of angular velocity range (Δvelocity), and CoM displacement and velocity variation during stair ascent and descent cycles in older adults with and without disabilities. Additionally, it investigates the same kinematic variables at contact and leaving the floor moments, and F-S and S-F cycles.

Methods

Study Design

A cross-sectional study following the recommendations of the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) [25] guideline was designed and implemented as part of an observational registered study in the ClinicalTrials.gov database (NCT05611723). Between June 1, 2022, and March 31, 2023, a sample selection questionnaire was distributed to the partners of Center for Rehabilitation Research of the School of Health Polytechnic of Porto, where this study took place. The data collection continued until April 30, 2023.

Participants

Individuals were enrolled in this study if they were aged older than 60 years, community residents of the northern region of Portugal, and independently capable of completing the tasks (walking, sitting, and standing from a chair, and ascending and descending stairs). The following conditions precluded volunteers from participating in this study: institutionalization; had an established diagnosis of malignancy or terminal diseases with an anticipated survival of less than 1 year; had stroke history, cerebral hemorrhage, head trauma, or Parkinson disease; had cognitive impairment, specifically Mini-Mental State Examination (MMSE, Portuguese version) score of 22 for 0 to 2 years of literacy, 24 for 3 to 6 years of literacy, and 27 for more than 6 years of literacy [26]; had rheumatic pathologies that interfere with the tasks performance; had diabetic foot, lower limb fracture in the past 6 weeks, or other related conditions; had unstable cardiovascular disease, liver or renal function failure; had a BMI greater than 30 kg/m^2; had uncorrected vestibular or audition or vision impairments that could influence the performance of tasks; or had symptoms at the time of evaluation that could influence task performance.

Procedures and Instruments

Sample Selection and Characterization

A questionnaire was used to characterize the participants regarding demographic data, health conditions, and history of falls in the past 12 months. Additionally, participants were characterized regarding anthropometric parameters (height, weight, and BMI), cognitive function (MMSE), self-reported medication intake, and self-reported physical activity (International Physical Activity Questionnaire: Short Form—Portuguese version).

XSL•FO
RenderX

Anthropometric parameters, specifically body mass (kg), were obtained by bioimpedance assessment (Tanita Inner Scan BC-601), with a precision of 0.1 kg. Height was measured using the seca 222 stadiometer (seca—Medical Scales and Measuring Systems), with a 1 mm scale. The MMSE (Folstein et al, 1975 [27]) was applied to assess the cognitive function of the participants. The total score ranges from 0 to 30, with a score of 22 for literacy from 0 to 2 years, 24 for literacy from 3 to 6 years, and 27 for literacy of 7 years or more [26]. Medication intake was self-reported, and polypharmacy was defined as taking 5 or more prescription drugs [28]. Self-reported physical activity was assessed by the 7-item short version of International Physical Activity Questionnaire [29].

Disability Indicators Characterization and Assessment

Participants were divided into 2 groups: older adults without disabilities and those with 2 or more disability indicators of the previously identified in the literature [30] namely: the handgrip strength (hand-held dynamometer), balance (single leg balance test time), overall self-reported health status (SRH; self-reported question), basic activities of daily living (ADL; Barthel Index) and instrumental activities of daily living (IADL; Lawton and Brody).

The handgrip strength was measured using a handheld dynamometer Jamar Plus+ Digital (Performance Health Supply). For the assessment, participants were seated with the shoulder adducted, the elbow flexed at 90 degrees and unsupported, the forearm neutral, and the wrist at 30 degrees of extension, as advised by the American Society of Hand Therapists [31]. The total time spent standing was measured in seconds using the participants' favored leg and with eyes open. The test was completed when the participants could not maintain the position or reached the maximum duration of 60 seconds. This time was used to determine when the individual became imbalanced (eg, flexed leg contacting the ground) [32].

By responding to the query "In general, how do you rate your health today?," the overall SRH was evaluated as ("very bad," "bad," "fair," "good," and "very good." Good SRH (defined as "very good" or "good") and poor SRH (defined as "fair," "bad," or "very bad") were dichotomized [33]. The participant's level of independence in the basic ADL was assessed using the Barthel Index [34]. The scoring system, ranging from 0 (complete dependency) to 20 (functional independence), was used as this scoring corrects for an uneven impression of accuracy [35]. The Lawton and Brody [36] IADL Scale was used to assess the functional capacity of older adults. It assesses 7 instrumental activities, and the final score ranges from 7 to 23, corresponding to the maximum score of total independence [37].

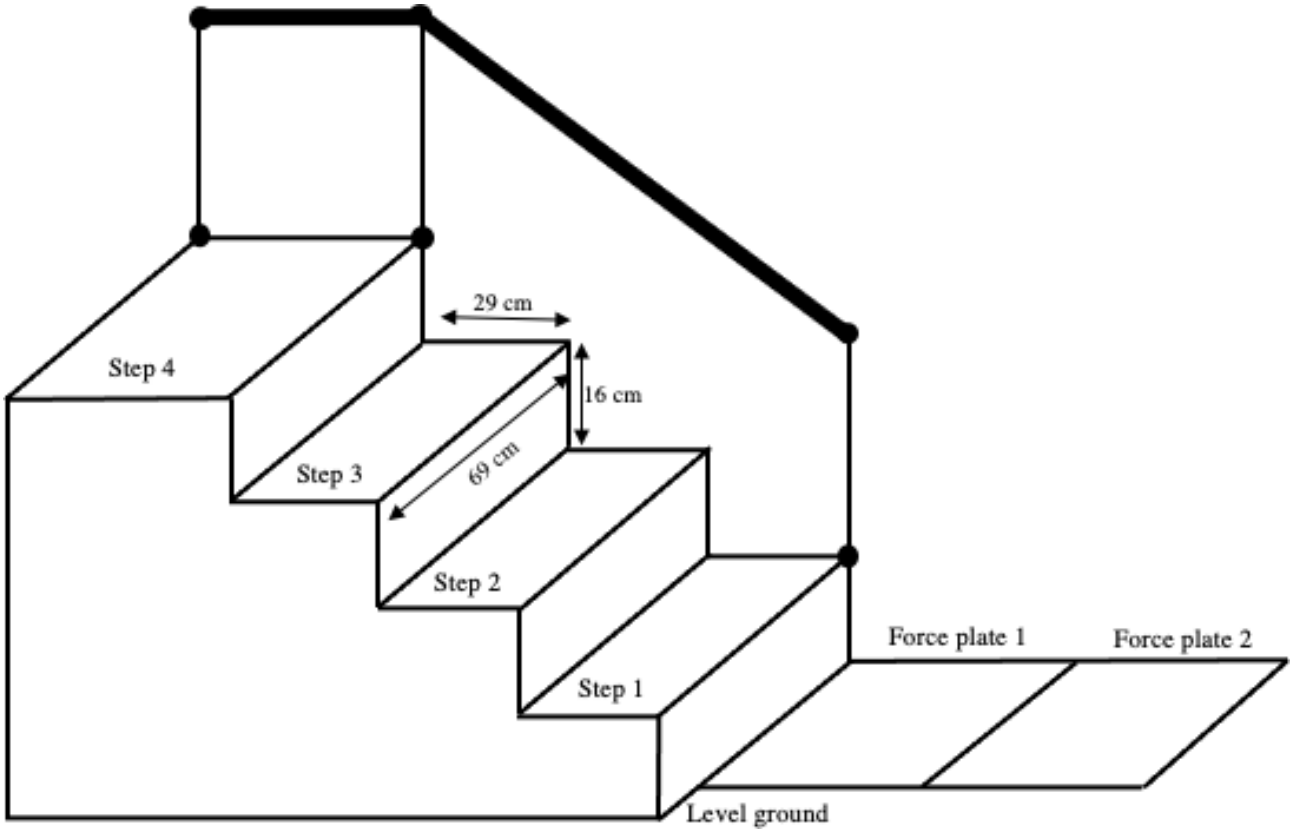
Kinematic Assessment

The stair negotiation was recorded using the Qualisys Track Manager (Qualisys AB), with 12 optoelectronic cameras (8 Oquos500 and 3 MiqusM3), 1 Miqus video camera, and 2 force plates (FP4060-08/10, Bertec). Kinematic data were collected at a 100 Hz frequency.

Stair negotiation was performed in a 4-step staircase (step height: 16 cm, depth: 29 cm, and width: 69 cm; Figure 1). The topmost step of the staircase was labeled "step 4," and the subsequent steps were labeled "step 3," "step 2," and "step 1." To successfully complete the stair ascent, participants had to stand in front of the staircase before the force plates, take the first step on level ground on the force plate 2, the second step on the force plate 1, and then ascend 3 stairs using a step-over-step sequence at their own leisure [13]. Each participant was informed on how to descend the staircase, which included starting at step 4 in an upright stance with their feet side by side and moving down on their own self-selected initiating limb and velocity [19,38]. Additionally, participants were told to keep walking over the force plates for at least 2 steps after they reached level ground [39]. When participants reached a standing position on level ground, the trial was considered over [19]. Throughout the testing, handrails were installed as a safety measure; however, participants were told not to use them unless absolutely essential. A total of 3 valid trials were recorded for each participant [10].

XSL•FO
RenderX

Figure 1. Scheme of stairs ascent, descent, and transitions with the floor setup.



Data Processing

Overview

Qualisys Track Manager (Qualisys AB; version 2020.3) was used to process marker trajectories of the 3 trials per participant. The data was then exported to Visual3D Professional (Has-Motion, Inc) [40].

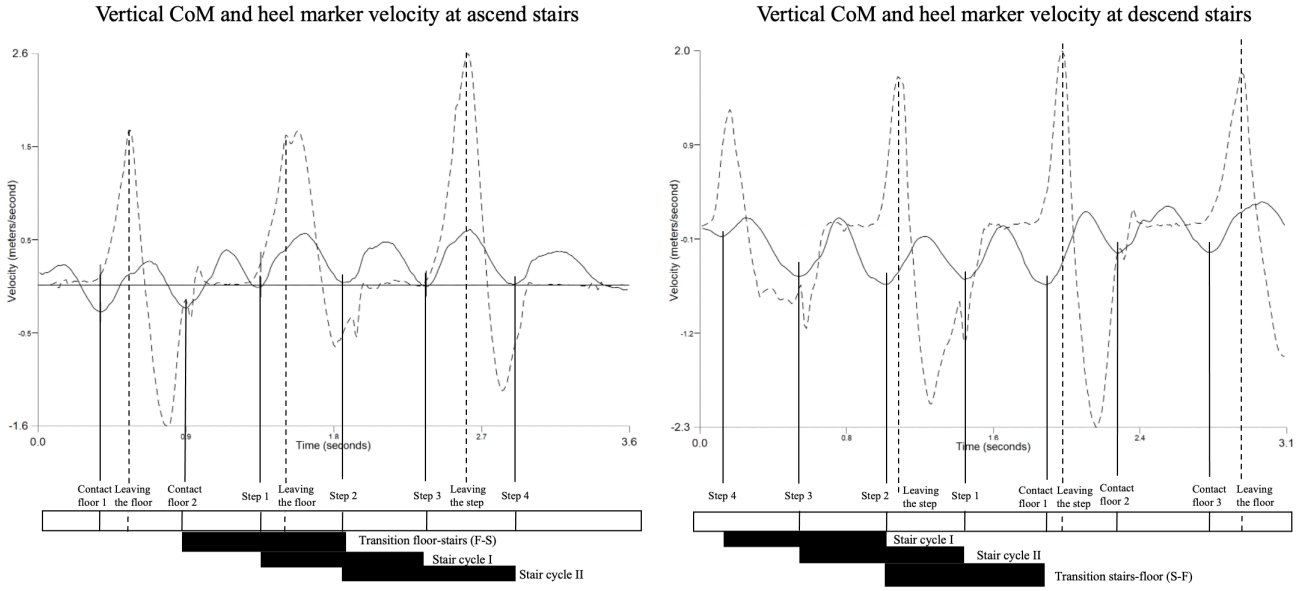
The joint angles were calculated using the Cardan sequence that assumes that the x-axis is in the medial-lateral direction, the y-axis is anterior-posterior, and the z-axis is in the up and down or axial direction.

To characterize the sequences of walking to ascending and descending stairs to walk, the values from cycle time (s), speed (m/s), and peak vertical velocity from the CoM (absolute value in meters per second) were computed.

The complete daily activity sequences of walking to ascend stairs and descending stairs to walk were divided into subtasks, specifically stair use (stair ascent and descent) and floor transitions (F-S and S-F). Thus, 2 distinct cycles were identified for stair ascent and descent. For the ascending stairs task, 1 stair cycle began in step 1 and ended in step 3, and another began in step 2 and ended in step 4. The descending task cycles were defined in the opposite direction (Figure 2). The F-S was considered as the instant the foot left the floor and ended when the foot left the step, with the opposite transition (S-F) following the reverse pattern (Figure 2). The moment of leaving the step or the ground was defined by the maximum heel marker's vertical velocity, and the moment of contact with the step or the ground was determined by the minimum downward CoM velocity [41].

XSL•FO
RenderX

Figure 2. Events definition for each task, stair ascent and stair descent, and respective transitions with the floor F-S and S-F. The regular line represents vertical CoM velocity, and the dashed line represents the vertical heel marker velocity. CoM: center of mass; F-S: floor-to-stair transition; S-F: stair-to-floor transition.



Processing Joint Position and Velocities at Events

During both stair cycles and the S-F and F-S, data on position and angular velocity of the hip, knee, and ankle joints were extracted from the ipsilateral limb performing the contact with the step or floor and the moment the foot left the step or floor. Accordingly, the variables selected were position and angular velocity of hip, knee, and ankle according to the 3 axes (X, Y, Z) at the contact with the step and leaving the step at ascent and descent stairs, and the contact with the floor and leaving the floor at the F-S and S-F.

Processing Joint Ranges of Motion and Velocities, CoM Displacement and Velocity Range at Cycles

The ROM was calculated as the difference between the maximum and minimum lower limb angles during the stair ascent and descent cycles and F-S and S-F cycles. The difference between the maximum and minimum angular velocities of the lower limb joints was defined as Δvelocity. The COM was computed by incorporating all body segments, with segment lengths determined by joint centers, and its position had as a resolution coordinate system the laboratory coordinate system, and the CoM velocity by taking the derivative of the CoM position over time [42]. The CoM displacement and CoM Δvelocity in the 3 planes (X, Y, Z) were also computed for both the stair cycles and the S-F and F-S. All the computations were performed in the software Visual3D Professional.

Statistical Analysis

The statistical software G*Power (version 3.1.9.7; Heinrich-Heine-Universität Düsseldorf) was used to determine a 0.948 power (1-β), introducing as independent variables the significance level (α=.05), effect size (*f*=0.6667), and sample size (n=60) [43].

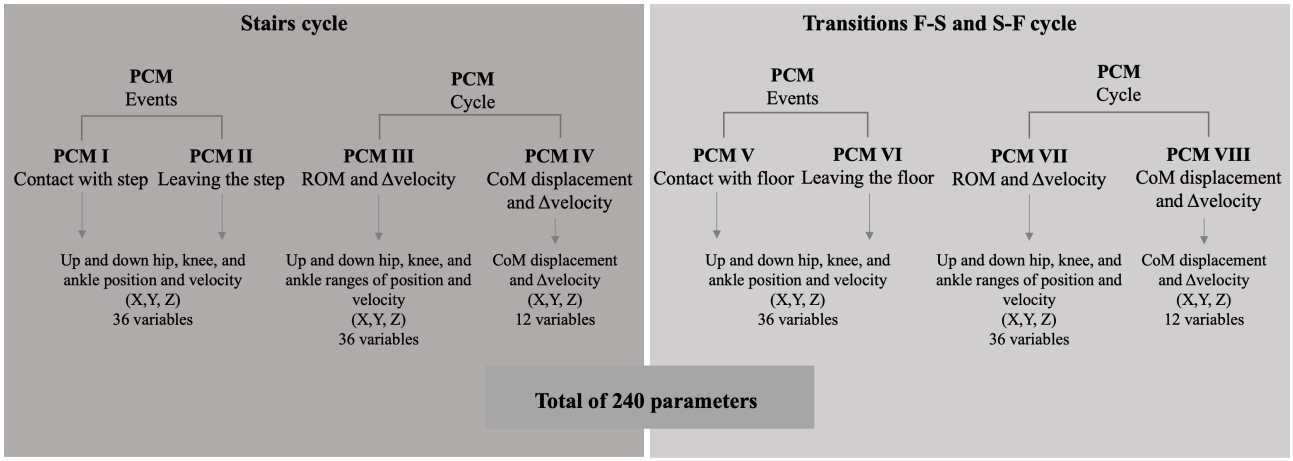
The IBM SPSS (version 28.0; SPSS Inc) statistical package was used to compile the data and for statistical analysis. The normality of the distribution was assessed by the Kolmogorov-Smirnov and Shapiro-Wilk tests for the group of older adults without and with disability, respectively. Between-group comparisons of demographic and clinical characteristics were conducted using Mann-Whitney tests or an independent sample *t* test (2-tailed). χ^2 tests were used to evaluate between-group association in the distribution of the categorical variables, namely sex, history of fall, and SRH.

Individual principal component models (PCMs) were made for different sets of variables. Particularly, PCM I included the position and velocity of the hip, knee, and ankle on the 3 planes (X, Y, Z) at the moment of contact with the step at stairs ascending and descending, and PCM II the same parameters at the moment the foot left the step during stair ascending and descending. The PCM V and VI included the same set of variables at contact with floor and leaving the floor moments, respectively, during the S-F and F-S. PCM III was computed for ROM and Δvelocity of hip, knee, and ankle on the 3 planes at the stairs ascent and descent, and PCM VII for the same set of variables at S-F and F-S. Additionally, PCMs IV and VII were conducted for COM displacement and Δvelocity on stairs and transition to floor cycles, respectively (Figure 3). PCA was carried out utilizing the correlation matrix to guarantee that every variable contributed equally to the analysis; accordingly, *z* scores were used to account for the varying measurement scales of the variables. The correlation matrix is based on standardized values, meaning that each variable was divided by its SD, effectively normalizing the variables to have a mean of 0 and an SD of 1. This step ensures that variables with different units and every variable contribute equally to the analysis. The suitability of the dataset for PCA was evaluated using Kaiser-Meyer-Olkin (KMO), as a measure of sampling adequacy for the overall data, and Bartlett Test of Sphericity as a measure of data suitability for reduction. A significant Bartlett test (*P*<.05) and KMO values larger than 0.5 were used to determine whether the data were suitable for PCA. Potential outliers were manually inspected in the data and verified not to unduly influence the PCA results. Using eigenvalues greater than 1 as

XSL•FO
RenderX

the criterion, the PCs that accounted for most of the variance were extracted, ensuring that only meaningful variance contributors were retained for biomechanical interpretation. To aid in the interpretability of the components, a Varimax rotation was applied to redistribute variance across PCs, making it easier to identify biomechanical patterns. The rotated component matrix was examined to identify clusters of variables that loaded onto the same components. The components were interpreted based on the variables that exhibited high loadings (loading >0.7, as the sample is n=60 [44]) on each component. The PC scores were tested for normality using the Kolmogorov-Smirnov and Shapiro-Wilk tests. As the data did not follow a normal distribution, the nonparametric Mann-Whitney *U* test was applied to assess differences between older adults with and without disability indicators, with the significance level set at 0.05.

Figure 3. Parameters included in each PCM, divided by the subtasks, stairs, and transitions with the floor. CoM: center of mass; F-S: floor-to-stair transition; PCM: principal component model; ROM: range of motion; S-F: stair-to-floor transition; Δvelocity: angular velocity range.



Ethical Considerations

This study adhered rigorously to the principles outlined in the Declaration of Helsinki. All participants received information about this study's purpose and methodology, and they completed and signed an informed consent form. This study was submitted to the Institutional Ethics Committee of E2S, Polytechnic of Porto on April 27, 2022, and obtained authorization on May 25, 2022 (CE0064C). Participants' identities and confidentiality were maintained throughout the research study. All participant data were deidentified and were stored on password-protected secure servers to prevent unauthorized access. There was no known risk or harm to participating in this study or publicizing its results or findings; participation was voluntary, and the participants did not receive compensation in this study.

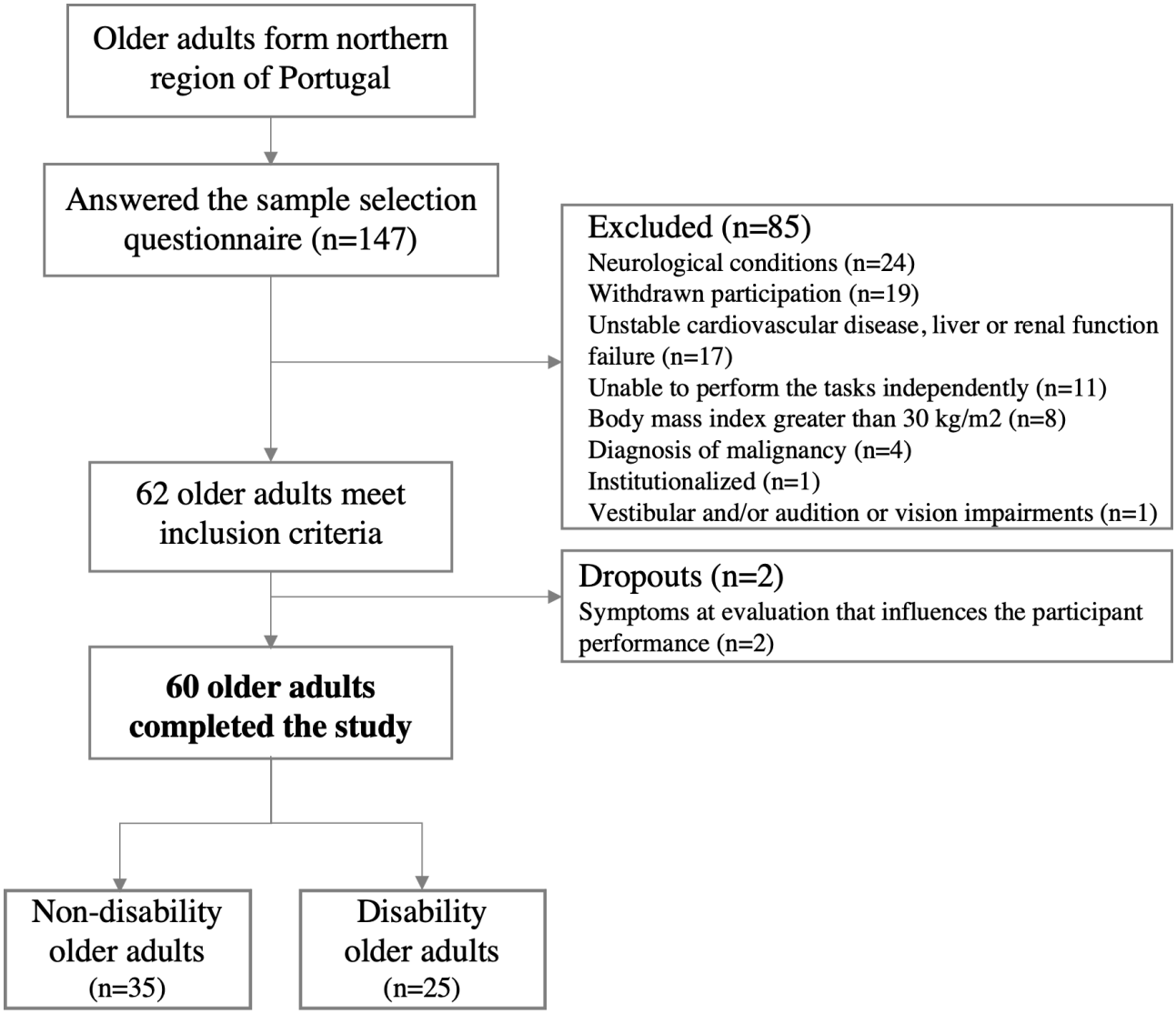
Results

Sample and Tasks Characterization

A total of 62 individuals, from 147 older adults, enrolled in the study, as shown in Figure 4. During the evaluation, 2 participants exhibited symptoms that affected the activities performance. As a result, 60 participants were included in the analysis of this study. Table 1 provides a summary of the descriptive statistics of their clinical and demographic features.

XSL•FO
RenderX

Figure 4. Flowchart of the sample selection process.



In examining the factors serving as markers for disability, substantial differences were observed between groups with and without disability (Table 1). Older individuals with disabilities exhibited poorer independence in ADL and IADL (P=.01 and P=.01, respectively), as well as reduced hand grip strength (P=.018) and diminished time spent in a 1-leg stand posture (P<.001). Furthermore, in comparison to the nondisabled group, older persons with disabilities rated their health as poorer (P<.001). Polypharmacy also showed significant disparity between the groups (P<.001).

Regarding stair negotiation, older adults with disabilities demonstrated a longer cycle time (P=.003), reduced speed (P<.001), and lower peak vertical velocity of the CoM (P=.003) while ascending stairs (Table 2). In the descending stairs task, older adults with disabilities exhibited lower speed (P<.009) and CoM peak velocity (P<.013).

XSL•FO
RenderX

Table . Demographic and clinical characterization of the sample. The *P* value reflects the comparison between ND[a] and D[b].

	ND (n=35)	D (n=25)	*P* value	Test value
Age (years)	66.34 (5.60)[c]	68.60 (6.77)[c]	.15	534[d]
Gender (female)	19 (54.29)[e]	19 (76)[e]	.09	2.9 (*df*=1)[f]
BMI (kg/m^2)	25.22 (3.08)[c]	26.02 (2.66)[c]	.30	−1.049[g]
History of falls in the previous 12 months (fallers)	11 (31.4)[e]	11 (44)[e]	.47	0.5 (*df*=1)[f]
Polypharmacy (polymedicated)	2 (5.71)[e]	11 (44)[e]	<.001	12.5 (*df*=1)[f]
Cognitive function (MMSE[h] score)	28.94 (1.31)[c]	28.68 (1.49)[c]	.50	394[d]
Self-reported physical activity (IPAQ[i] MET[j]-min/wk)	3186.46 (2964.91)[c]	3519.66 (2822.11)[c]	.51	393.5[d]
Self-reported health			<.001	21.8 (*df*=1)[f]
Poor	8 (22.86)[e]	21 (84)[e]		
Good	27 (77.14)[e]	4 (16)[e]		
Hand grip strength (kg)	36.59 (39.86)[c]	25.07 (7.54)[c]	.02	279.5[d]
One leg standing time (s)	38.83 (20.93)[c]	18.19 (20.72)[c]	<.001	192.5[d]
ADL[k] independence (Barthel Index score)	19.97 (0.17)[c]	19.76 (0.44)[c]	.01	345[d]
IADL[l] independence (Lawton & Brody score)	23 (0.00)[c]	21.96 (2.67)[c]	.01	332.5[d]

[a]ND: older adults without disability.

[b]D: older adults with disability.

[c]Mean (SD).

[d]Mann-Whitney *U* test

[e]n (%).

[f]χ^2 test

[g]Independent samples *t* test (2-tailed).

[h]MMSE: Mini-Mental State Examination.

[i]IPAQ: International Physical Activity Questionnaire.

[j]MET: metabolic equivalent of task.

[k]ADL: activities of daily living

[l]IADL: instrumental activities of daily living.

XSL•FO
RenderX

Table . Walk to ascending stairs and descend stairs to characterize walk tasks according to cycle time, double limb support time, speed, and CoM[a] peak velocity mean and SD. The *P* value reflects the comparison between ND[b] and D[c].

	Walking to ascend stairs				Descending stairs to walk			
	ND, mean (SD)	D, mean (SD)	*P* value	Test value	ND, mean (SD)	D, mean (SD)	*P* value (test value)	Test value
Cycle time (s)	1.15 (0.23)	1.31 (0.16)	.003	−2.876[d]	1.13 (0.21)	1.22 (0.16)	.06	562.0[e]
Speed (m/s)	0.68 (0.11)	0.60 (0.08)	<.001	3.278[d]	0.86 (0.18)	0.76 (0.14)	.009	2.450[d]
CoM peak vertical velocity (m/s)	0.20 (0.05)	0.17 (0.03)	.003	−2.853[d]	0.56 (0.07)	0.52 (0.08)	.013	−2.284[d]

[a]CoM: center of mass.

[b]ND: older adults without disability.

[c]D: older adults with disability.

[d]Independent samples *t* test (2-tailed).

[e]Mann-Whitney *U* test

PCA Models

PCMs I to VIII presented KMO values above 0.5, and Bartlett Test of Sphericity was significant (*P*<.001) for all PCMs, indicating that the data were suitable for dimensionality reduction. Table S1 in Multimedia Appendix 1 includes a detailed description of the PCs with eigenvalues greater than 1 for each of the PCMs.

This study observed statistically significant differences (*P*<.05) between older adults with and without disabilities in 12 PCs across PCM I to VIII. After applying Varimax rotation, the components were interpreted based on the variables with loadings exceeding 0.7, as outlined in Table 3. Analysis of the variables with higher loadings in the PCs revealed a prevalence of variables associated with angular velocity, distributed similarly during ascending and descending tasks, particularly in the sagittal plane. Notably, the contact of the foot with the step was found to be important in the ascending task regarding frontal hip angular velocity and sagittal knee velocity (PCM I).

XSL•FO
RenderX

Table . Variables with loadings higher than 0.7 of the PC[a] that presented significant differences between ND[b] and D[c], with respective mean and SD of each group. The tasks ascending stairs and descending stairs are designated up and down, respectively.

PCM[d]	PC	Variable	Load	ND, mean (SD)	D, mean (SD)
Stairs					
PCM I (contact with step)	4				
		Up hip velocity (Y)	0.814	18.7 (22.0)	10.3 (20.5)
		Up knee velocity (X)	0.888	−140.2 (62.2)	−89.1 (42.9)
PCM II (leaving the step)					
	2	Up hip angle (X)	0.750	34.7 (12.7)	30.4 (11.4)
	2	Up hip velocity (X)	0.784	199.7 (43.2)	168.7 (36.1)
	2	Up knee angle (X)	0.811	58.4 (11.6)	49.7 (12.7)
	2	Up knee velocity (X)	0.847	392.5 (91.4)	336.5 (63.9)
	2	Up ankle velocity (X)	0.744	128.0 (87.0)	84.9 (79.4)
	3	Down hip velocity (X)	0.804	153.0 (43.4)	113.5 (35.1)
	3	Down knee velocity (X)	0.877	283.2 (86.2)	197.0 (87.3)
	3	Down ankle velocity (X)	−0.853	−212.6 (50.8)	−160.0 (58.8)
PCM III (up and down cycles)					
	1	Up hip Δvelocity[e] (X)	0.856	316.5 (45.4)	286.3 (35.1)
	1	Up knee Δvelocity (X)	0.901	651.1 (112.3)	587.0 (83.8)
	1	Down hip ROM[f] (X)	0.803	27.8 (4.0)	27.0 (3.0)
	1	Down hip Δvelocity (X)	0.853	295.8 (85.4)	239.5 (45.4)
	1	Down knee Δvelocity (X)	0.807	690.6 (138.0)	570.7 (97.2)
	10	Down ankle ROM (Y)	0.867	12.4 (2.8)	14.4 (3.3)
PCM IV (up and down cycles)					
	1	Up CoM[g] displacement (X)	0.824	0.05 (0.01)	0.04 (0.01)
		Up CoM Δvelocity (X)	0.855	0.22 (0.05)	0.19 (0.04)
		Down CoM displacement (X)	0.820	0.05 (0.01)	0.04 (0.01)
		Down CoM Δvelocity (X)	0.801	0.31 (0.06)	0.30 (0.05)
F-S[h] and S-F[i]					
PCM V (contact with floor)	3				
		S-F hip velocity (Y)	0.852	7.2 (21.3)	−8.7 (15.0)
		S-F hip velocity (Z)	−782	−12.3 (23.2)	−5.4 (18.8)
		S-F knee velocity (Z)	0.841	10.3 (40.0)	−8.8 (28.9)
PCM VI (leaving the floor)					
	2	F-S ankle velocity (Z)	−0.746	4.7 (10.3)	−4.8 (16.6)

XSL•FO
RenderX

PCM[d]	PC	Variable	Load	ND, mean (SD)	D, mean (SD)
	2	S-F hip velocity (Z)	0.788	−14.6 (35.1)	0.7 (29.5)
	2	S-F ankle velocity (X)	0.718	−230.4 (61.8)	−168.0 (63.1)
	2	S-F ankle velocity (Z)	−0.816	4.2 (32.7)	−10.9 (30.1)
	3	F-S hip velocity (X)	−0.857	142.2 (34.0)	130.0 (24.2)
	3	F-S knee velocity (X)	−0.929	234.3 (116.5)	192.1 (97.6)
	9	F-S hip angle (Y)	0.784	1.8 (2.9)	3.7 (2.4)
	9	S-F hip angle (Y)	0.901	3.1 (3.4)	4.8 (3.4)
PCM VII (F-S and S-F cycles)	1				
		S-F hip ROM (X)	0.747	46.1 (7.0)	46.3 (6.0)
		S-F hip Δvelocity (X)	0.861	296.6 (63.6)	274.7 (50.1)
		S-F knee ROM (X)	0.789	75.5 (4.4)	73.4 (4.4)
		S-F knee Δvelocity (X)	0.888	750.6 (130.3)	652.7 (114.1)
		F-S hip Δvelocity (X)	0.865	329.4 (59.1)	297.7 (47.4)
		F-S knee ROM (X)	0.795	70.2 (6.6)	67.7 (5.7)
		F-S knee Δvelocity (X)	0.907	658.5 (122.5)	575.0 (91.6)
PCM VIII (F-S and S-F cycles)	2				
		Up CoM displacement Z	0.710	0.22 (0.01)	0.22 (0.01)
		Up CoM Δvelocity Z	0.833	0.60 (0.10)	0.65 (0.08)
		Down CoM Δvelocity Z	0.798	0.31 (0.06)	0.30 (0.05)

[a]PC: principal component.
[b]ND: older adults without disability.
[c]D: older adults with disability.
[d]PCM: principal component model.
[e]Δvelocity: angular velocity range.
[f]ROM: range of motion.
[g]CoM: center of mass.
[h]F-S: floor-to-stair transition.
[i]S-F: stair-to-floor transition.

Additionally, significant differences between groups were demonstrated in 2 PCs at the moment of leaving the step (PCM II), 1 defining the ascending (PC2) and the other the descending (PC3) tasks. Position and angular velocity on the sagittal plane of the hip, and knee and ankle angular velocity are present in PC 2, with higher loadings for the ones related to the knee joint. The PC3, defining the descending task, retained the hip, knee, and ankle angular velocities in the sagittal plane. The ROM and the range of angular velocities were divided into the sagittal and frontal planes, with PC1 retaining sagittal hip ROM and Δvelocity, and knee Δvelocity when descending stairs and hip and knee Δvelocity when ascending stairs. The PC 10 retained the frontal plane ROM of the ankle joint during descending (PCM III).

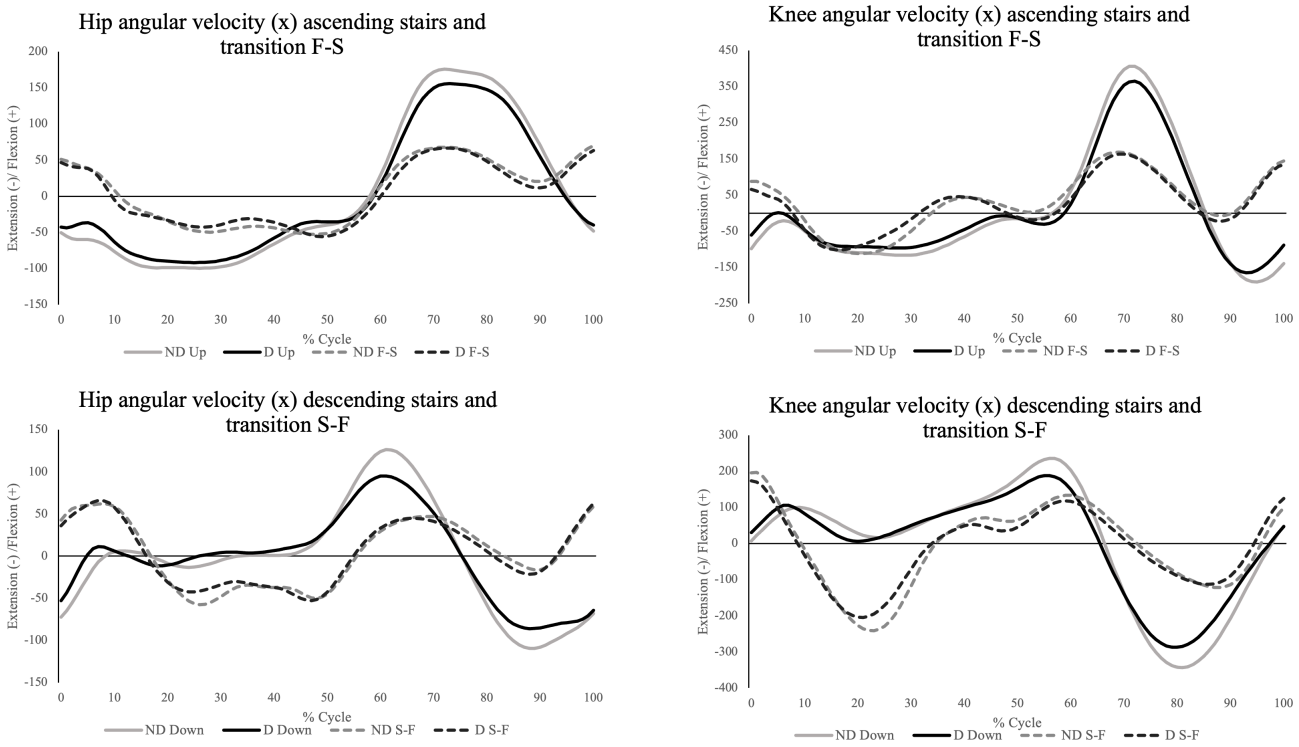
The PCM V regarding the contact moment retained only variables of the S-F, particularly, frontal and transverse hip angular velocity and transverse knee angular velocity. At the moment of leaving the floor (PCM VI), 3 PCs demonstrated differences between groups, gathering variables of S-F and F-S. PC2 presented higher loadings in angular velocities of the S-F of the transverse plane in the hip joint and sagittal and transverse planes in the ankle joint. This component also presented higher loading for the transverse ankle velocity in the F-S. PC3 retained F-S sagittal hip and knee velocities, as PC 9 retained frontal hip angles of both F-S and S-F. The PCM VII demonstrated significant differences between groups only in PC1, which explained 25% of the total variance. The variables with higher loadings were all of the sagittal planes (S-F hip and knee ROM and Δvelocity, and F-S hip Δvelocity and knee ROM and Δvelocity). The variables with higher loadings were all of the sagittal plane (S-F hip and knee ROM and Δvelocity, and F-S hip Δvelocity and knee ROM and Δvelocity). In analyzing both stair ascent and descent cycles as well as transitional movements, the parameters most consistently preserved were the angular velocities of the hip and knee in the sagittal plane.

XSL•FO
RenderX

Accordingly, the complete cycle for the groups without and with disability is plotted in Figure 5. Therefore, this figure highlights the main differences between the groups identified by the PCA.

The CoM displacement and Δvelocity PCM IV built with stairs cycle data demonstrated differences between older adults with and without disability in PC 1 explained approximately 33% of the variance and retained the displacement and Δvelocity of CoM in the sagittal plane, both in ascending and descending stairs. The PC2 of PCM VIII, computed with the CoM data of transition to floor cycles, presented differences between the groups, retaining the variables considering the transverse plane, except for CoM displacement in the S-F.

Figure 5. Sagittal hip and knee angular velocities on ascending stairs (up) and F-S and descending stairs (down) and S-F, within the cycle for ND and D. D: older adults with disability; F-S: floor-to-stair transition; ND: older adults without disability; S-F: stair-to-floor transition.



Discussion

Principal Results and Comparison With Prior Work

This study investigated the kinematic differences in joint angles, joint velocities, and CoM displacement during stair ascent, descent, and F-S and S-F among older adults with and without disability. This study is believed to be the first to delve into the distinguishing parameters of kinematics between older adults with and without disability when negotiating stairs and transitions to the floor. As hypothesized, older adults with disabilities demonstrated significant differences in characterization parameters during stair ascent to walk and stair descent to walk, particularly in cycle time, speed, and CoM vertical peak velocity when compared to their nondisabled counterparts.

In this study, the time taken to walk to ascend stairs was significantly longer in older adults with disability. This contrasts with previous research, such as the study by Oh-Park et al (2011) [18], which considered stair descent time, but not ascent time, as a predictor of functional decline. This difference may be attributed to the fact that the participants in the study by Oh-Park et al (2011) [18] commenced the task in front of the stairs, whereas the current study evaluated stair negotiation and the transition to the floor, beginning and ending with walking. This approach is more challenging and reflective of real-life activities, as highlighted in the study by Vallabhajosula et al (2012) [45]. In addition to a longer time to complete the ascending task, older adults with disabilities also demonstrated reduced speed and lower CoM peak velocity while walking to ascent stairs and descending stairs to walk. This is consistent with a more conservative strategy that could reduce the risk for a fall that was identified in older adults compared to younger adults [19], which seems to be augmented in older adults with disability.

In addition to the characterization parameters of the tasks, this study encompassed a comprehensive analysis of a wide array of variables, including those related to the position and velocity of lower limb joints, CoM displacement and velocity in 3 axes, and the description of 2 tasks: ascending and descending. These tasks were further subdivided into subtasks, particularly using stairs and transitioning to the floor, and were characterized at 2 distinct moments: contact with the step or floor and leaving the step or floor, along with their respective ranges. Given the complexity and multidimensionality of the dataset, a PCA approach was deemed appropriate. PCA was used to elucidate the shared variance of the parameters associated with stair negotiation, thereby identifying a simplified representation that consolidates the most significant parameters. This approach simplifies the analysis of stair negotiation in older adults with disability. While the mechanics of successful stair negotiation for healthy older adults have been extensively documented [8], studies focusing on older adults with disability have

XSL•FO
RenderX

predominantly centered on measuring the time taken to complete tasks [18]. The present study sought to assess the aforementioned set of parameters to gain a comprehensive understanding of the behavior of older adults with disability during stair negotiation. Moreover, it is fundamental to consider different axes, as previous studies have shown that factors in the sagittal plane are crucial [13,16,17], and limited studies have compared older adults to younger individuals in terms of the frontal and transverse planes [46,47]. Furthermore, the PCA analysis was instrumental in reducing and analyzing the data curves of knee angles, moments, and forces in a study by Reid et al (2010) [14]. That decision was informed by existing literature, which has highlighted the direct relationship between functional impairment during stair negotiation and the knee joints of older adults compared to their younger counterparts [14]. In line with this rationale, a conservative approach was adopted in the present study, entailing the analysis of discrete data collected at different events as an initial step in exploring the task of stair negotiation among older adults with disability. By focusing on discrete events rather than continuous data, our PCM approach provides a more granular analysis of stair negotiation mechanics.

Accordingly, the use of individual PCMs allowed us to target specific, biomechanically meaningful parameters in the stair negotiation cycle, such as angular velocity, especially within the sagittal plane. In the cycles of stair ascent and descent, both the instances of foot contact with the step and subsequent lift-off were marked by significant retention of hip and knee sagittal angular velocity in PCM I and II. This significance was further highlighted by the range analyses conducted in PCM III, which preserved the ranges of sagittal angular velocities of the hip and knee joints during both ascending and descending tasks. Notably, older adults with disabilities demonstrated lower mean sagittal angular velocities at the hip and knee joints compared to their nondisabled peers. This disparity is visually represented in Figure 5, which indicates that while older adults with disabilities exhibit movement patterns similar to those of nondisabled older adults, they tend to adopt more conservative movement strategies. These findings align with previous aging research, which demonstrated that knee joint angular velocity at the point of maximal knee joint moment demand during the stride cycle was significantly lower in older adults compared to younger adults, a pattern not observed for ankle joint angular velocity [17,48]. These studies also found that variation in knee extension-flexion angles and ankle plantarflexion-dorsiflexion angles during stair ascent [48] and descent [17] was comparable between young and older adults, consistent with our observations among older adults with and without disability. Prior research on stair descent also indicated reduced knee ROM in the sagittal plane and increased hip ROM in the frontal and transverse planes in older adults compared to younger adults [47]. It is hypothesized that older adults may rely more on ankle movements to maintain stability, especially in the frontal plane, as a balance-preserving strategy that can help manage potential mediolateral instability on descending stairs compared with younger adults [12,47]. In the present study, although lower limb joint positions did not appear to be predominant in stair descent, the PCA indicated that older adults with disabilities descended stairs with lower sagittal hip ROM and higher frontal ankle ROM, suggesting that older adults with disabilities further used this strategy compared with those without disability.

Within the F-S and S-F cycles, the hip and knee sagittal angular velocity variables were also retained. Additionally, hip angular velocity in the transverse plane was notably retained at the moments of floor contact and leaving the floor during these transitions. The emergence of variables associated with frontal and transverse planes in F-S and S-F aligns with previous findings indicating that older adults exhibit reduced independent control of adjacent joints during these transitions compared to younger adults [49]. Moreover, the hip angle in the frontal plane was higher in older adults with disability in F-S and S-F. The transition from level gait to stair and the opposite necessitates a shift in motor control between these tasks [15], which seems to be significantly more impaired in older adults with disability.

CoM parameters associated with mediolateral displacement and velocity were retained within the stair cycle. In contrast, during the F-S and S-F cycles, the retained parameters predominantly described vertical displacement and velocity, except for F-S velocity. Buckley et al (2013) [19] proposed that reduced peak downwards CoM velocity of older participants may result from stiffening the limb, prolonging limb flexion when lowering the CoM. This effect may be amplified in older adults with disability, according to our findings. Additionally, Nadeau et al (2003) [46] observed that mediolateral CoM control demands are greater during stair negotiation than during level walking, corroborating the critical role of CoM parameters in influencing stair negotiation dynamics.

Given that stair use in the home environment can help older adults without disability retain the ability to perform IADL [50], these insights carry important implications for the development of targeted interventions aimed at supporting the stability and mobility of older adults, especially those with disability. These findings thus provide clinicians and researchers with actionable information to better address the unique mobility needs of this population, facilitating improvements in both safety and quality of life for older adults navigating stair environments.

Limitations

Despite the original contributions of this study, certain limitations are noted. Older adults with disabilities take significantly more medication than their counterparts. As the analysis of the medication type was not the object of this study, future studies should address that relationship because stair negotiation is influenced by medication intake [51]. Additionally, due to the structure of our PCA-based analysis and the lack of raw group-wise distributions for each retained PC variable, we were unable to calculate reliable effect sizes post hoc. Nevertheless, this study's PCA marks initial steps in stair negotiation analysis; future research should explore full waveform data of high-loading parameters to gain more comprehensive insights.

The findings of this study align with established literature on stair negotiation biomechanics in older adults, highlighting compensatory behaviors such as reduced sagittal velocities and conservative movement strategies. The unique contribution of this research lies in the detailed retention of kinematic

XSL•FO
RenderX

parameters across various movement planes and phases, offering a more nuanced understanding of stability and movement adaptations in older adults with disability. By focusing on discrete events rather than continuous data, our PCM approach provides a more granular analysis of stair negotiation mechanics, identifying critical sagittal angular velocities of the hip and knee joints during stair ascent and descent, as well as frontal and transverse plane movements of these joints during S-F. Additionally, this approach highlights the potential application of PCMs to the data curves for these parameters, enhancing their utility in both research and clinical contexts.

Conclusions

Overall, this study examined kinematic differences in joint angles, joint velocities, and CoM displacement and velocity between older adults with and without disabilities during stair ascent and descent. Movements F-S and S-F were also analyzed. PCA of 240 kinematic variables led to the selection of 41 key parameters, mainly related to knee and hip joint angular velocity. In the cycles of stair ascent and descent, both the foot contact with the step and subsequent lift-off were marked by significant retention of hip and knee sagittal angular velocity as well as respective Δvelocity. These parameters were also retained within the F-S and S-F cycles, in addition to the hip angular velocity in the transverse plane at the moments of floor contact and leaving the floor. For the stair cycle, CoM parameters were largely linked to mediolateral displacement and velocity range, while during F-S and S-F, parameters describing vertical displacement and velocity were predominant, except for F-S velocity.

The findings from this study provide valuable insights that could inform disability prevention initiatives. This PCA approach represents a foundational step in understanding stair negotiation in older adults, and further research should focus on full waveform data of high-loading parameters to provide a more comprehensive understanding.

Acknowledgments

The authors would like to thank all included study participants for their invaluable cooperation and commitment, which made this research possible.

This work was supported by the Fundação para a Ciência e Tecnologia (FCT), Portuguese Ministry of Education, Science and Innovation, NORTE 2020, and European Social Fund of European Union (2020.05356.BD; DOI 10.54499/2020.05356.BD) and through R&D Units funding (UIDB/05210/2020), Fundação para a Ciência e Tecnologia (FCT), Portugal, and the European Union.

Authors' Contributions

JM and ASPS worked on this study by handling the conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, validation, visualization, and writing of the original draft and its review and editing. IdST, JF, DCG, LATA assisted with the validation, investigation, data curation, and writing of the original draft and its review of this study. BC aided with the formal analysis, review of the writing, and validation of this study. RS reviewed of the writing, validated and supervised this study, and administrated this project.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Table comparing the principal component scores between older adults without and with disability.
[DOCX File, 35 KB - aging_v8i1e71530_app1.docx]

References

1. International classification of functioning, disability and health: ICF. World Health Organization. 2001. URL: http://who.int/standards/classifications/international-classification-of-functioning-disability-and-health [accessed 2025-08-05]
2. Fried LP, Ferrucci L, Darer J, Williamson JD, Anderson G. Untangling the concepts of disability, frailty, and comorbidity: implications for improved targeting and care. J Gerontol A Biol Sci Med Sci 2004 Mar 1;59(3):M255-M263. [doi: 10.1093/gerona/59.3.M255]
3. Partridge L, Deelen J, Slagboom PE. Facing up to the global challenges of ageing. Nature New Biol 2018 Sep;561(7721):45-56. [doi: 10.1038/s41586-018-0457-8] [Medline: 30185958]
4. Khan SS, Singer BD, Vaughan DE. Molecular and physiological manifestations and measurement of aging in humans. Aging Cell 2017 Aug;16(4):624-633. [doi: 10.1111/acel.12601] [Medline: 28544158]
5. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2019;396(10258):1204-1222. [doi: 10.1016/S0140-6736(20)30925-9]

XSL•FO
RenderX

6. Han Y, Zhang L, Fang Y. Multidimensional disability evaluation and confirmatory analysis of older adults in a home-based community in China. Front Public Health 2022;10:899303. [doi: 10.3389/fpubh.2022.899303] [Medline: 35801251]
7. Tiedemann AC, Sherrington C, Lord SR. Physical and psychological factors associated with stair negotiation performance in older people. J Gerontol A Biol Sci Med Sci 2007 Nov;62(11):1259-1265. [doi: 10.1093/gerona/62.11.1259] [Medline: 18000146]
8. Jacobs JV. A review of stairway falls and stair negotiation: lessons learned and future needs to reduce injury. Gait Posture 2016 Sep;49:159-167. [doi: 10.1016/j.gaitpost.2016.06.030] [Medline: 27427833]
9. Ackermans TMA, Francksen NC, Casana-Eslava RV, et al. A novel multivariate approach for biomechanical profiling of stair negotiation. Exp Gerontol 2019 Sep;124:110646. [doi: 10.1016/j.exger.2019.110646] [Medline: 31269462]
10. Kováčiková Z, Sarvestan J, Zemková E. Age-related differences in stair descent balance control: are women more prone to falls than men? PLoS One 2021;16(1):e0244990. [doi: 10.1371/journal.pone.0244990] [Medline: 33411803]
11. Williamson JD, Fried LP. Characterization of older adults who attribute functional decrements to “old age”. J Am Geriatr Soc 1996 Dec;44(12):1429-1434. [doi: 10.1111/j.1532-5415.1996.tb04066.x] [Medline: 8951311]
12. Startzell JK, Owens DA, Mulfinger LM, Cavanagh PR. Stair negotiation in older people: a review. J Am Geriatr Soc 2000 May;48(5):567-580. [doi: 10.1111/j.1532-5415.2000.tb05006.x] [Medline: 10811553]
13. Novak AC, Brouwer B. Sagittal and frontal lower limb joint moments during stair ascent and descent in young and older adults. Gait Posture 2011 Jan;33(1):54-60. [doi: 10.1016/j.gaitpost.2010.09.024]
14. Reid SM, Graham RB, Costigan PA. Differentiation of young and older adult stair climbing gait using principal component analysis. Gait Posture 2010 Feb;31(2):197-203. [doi: 10.1016/j.gaitpost.2009.10.005]
15. Di Giulio I, Reeves ND, Roys M, et al. Stair gait in older adults worsens with smaller step treads and when transitioning between level and stair walking. Front Sports Act Living 2020;2:63. [doi: 10.3389/fspor.2020.00063] [Medline: 33345054]
16. Novak AC, Li Q, Yang S, Brouwer B. Mechanical energy transfers across lower limb segments during stair ascent and descent in young and healthy older adults. Gait Posture 2011 Jul;34(3):384-390. [doi: 10.1016/j.gaitpost.2011.06.007] [Medline: 21733694]
17. Reeves ND, Spanjaard M, Mohagheghi AA, Baltzopoulos V, Maganaris CN. The demands of stair descent relative to maximum capacities in elderly and young adults. J Electromyogr Kinesiol 2008 Apr;18(2):218-227. [doi: 10.1016/j.jelekin.2007.06.003] [Medline: 17822923]
18. Oh-Park M, Wang C, Verghese J. Stair negotiation time in community-dwelling older adults: normative values and association with functional decline. Arch Phys Med Rehabil 2011 Dec;92(12):2006-2011. [doi: 10.1016/j.apmr.2011.07.193] [Medline: 22133249]
19. Buckley JG, Cooper G, Maganaris CN, Reeves ND. Is stair descent in the elderly associated with periods of high centre of mass downward accelerations? Exp Gerontol 2013 Feb;48(2):283-289. [doi: 10.1016/j.exger.2012.11.003] [Medline: 23178302]
20. Bosse I, Oberländer KD, Savelberg HH, Meijer K, Brüggemann GP, Karamanidis K. Dynamic stability control in younger and older adults during stair descent. Hum Mov Sci 2012 Dec;31(6):1560-1570. [doi: 10.1016/j.humov.2012.05.003] [Medline: 22853941]
21. Fakhoury M, Rachkidi R, Semaan K, et al. Kinematic limitations during stair ascent and descent in patients with adult spinal deformity. Brain Spine 2025;5:104153. [doi: 10.1016/j.bas.2024.104153] [Medline: 39759966]
22. Ayoub E, Rteil A, Chaaya C, et al. Head and pelvis are the key segments recruited by adult spinal deformity patients during daily life activities. Sci Rep 2024 Sep 17;14(1):21716. [doi: 10.1038/s41598-024-70038-w] [Medline: 39289378]
23. Brandon SCE, Graham RB, Almosnino S, Sadler EM, Stevenson JM, Deluzio KJ. Interpreting principal components in biomechanics: representative extremes and single component reconstruction. J Electromyogr Kinesiol 2013 Dec;23(6):1304-1310. [doi: 10.1016/j.jelekin.2013.09.010] [Medline: 24209874]
24. Abdi H, Williams LJ. Principal component analysis. WIREs Computational Stats 2010 Jul;2(4):433-459. [doi: 10.1002/wics.101]
25. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008 Apr;61(4):344-349. [doi: 10.1016/j.jclinepi.2007.11.008]
26. Morgado J, Rocha CS, Maruta C, Guerreiro M, Martins IP. Novos valores normativos do Mini-Mental State Examination [Article in Portuguese]. Sinapse 2009;9(2):10-16 [FREE Full text]
27. Folstein MF, Folstein SE, McHugh PR. “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
28. Lee EA, Brettler JW, Kanter MH, et al. Refining the definition of polypharmacy and its link to disability in older adults: conceptualizing necessary polypharmacy, unnecessary polypharmacy, and polypharmacy of unclear benefit. Perm J 2020;24:18.212. [doi: 10.7812/TPP/18.212] [Medline: 31905333]
29. Craig CL, Marshall AL, Sjöström M, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]

XSL•FO
RenderX

30. Moreira JS, Melo A, Santos R, Sousa ASP. Indicators and Instruments to assess components of disability in community-dwelling older adults: a systematic review. Sensors (Basel) 2022 Oct 28;22(21):21. [doi: 10.3390/s22218270] [Medline: 36365968]
31. Jansen CWS, Niebuhr BR, Coussirat DJ, Hawthorne D, Moreno L, Phillip M. Hand force of men and women over 65 years of age as measured by maximum pinch and grip force. J Aging Phys Act 2008 Jan;16(1):24-41. [doi: 10.1123/japa.16.1.24] [Medline: 18212392]
32. Furuna T, Nagasaki H, Nishizawa S, et al. Longitudinal change in the physical performance of older adults in the community. J Jpn Phys Ther Assoc 1998;1(1):1-5. [doi: 10.1298/jjpta.1.1] [Medline: 25792874]
33. Crimmins EM. Trends in the health of the elderly. Annu Rev Public Health 2004;25:79-98. [doi: 10.1146/annurev.publhealth.25.102802.124401] [Medline: 15015913]
34. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. Md State Med J 1965 Feb;14:61-65. [Medline: 14258950]
35. Collin C, Wade DT, Davies S, Horne V. The Barthel ADL Index: a reliability study. Int Disabil Stud 1988;10(2):61-63. [doi: 10.3109/09638288809164103] [Medline: 3403500]
36. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179-186. [Medline: 5349366]
37. Araújo F, Pais-Ribeiro J, Oliveira A, et al. Validação da escala de lawton e brody numa amostra de idosos não institucionalizados [article in portuguese]. Presented at: Actas Do 7° Congresso Nacional de Psicologia Da Saúde; Jan 31 to Feb 2, 2008; Oporto and Lisbon p. 217-220 URL: https://sigarra.up.pt/fpceup/pt/pub_geral.pub_view?pi_pub_base_id=14792 [accessed 2022-05-28]
38. Singhal K, Kim J, Casebolt J, Lee S, Han KH, Kwon YH. Kinetic comparison of older men and women during walk-to-stair descent transition. Gait Posture 2014 Sep;40(4):600-604. [doi: 10.1016/j.gaitpost.2014.07.004] [Medline: 25082325]
39. Moniz-Pereira V, Kepple TM, Cabral S, João F, Veloso AP. Joint moments' contributions to vertically accelerate the center of mass during stair ambulation in the elderly: an induced acceleration approach. J Biomech 2018 Oct 5;79:105-111. [doi: 10.1016/j.jbiomech.2018.07.040] [Medline: 30104054]
40. Moreira J, Cunha B, Félix J, Santos R, Sousa ASP. Kinematic and kinetic gait principal component domains in older adults with and without functional disability: a cross-sectional study. J Funct Morphol Kinesiol 2025 Apr 23;10(2):140. [doi: 10.3390/jfmk10020140] [Medline: 40407424]
41. Foster RJ, Maganaris CN, Reeves ND, Buckley JG. Centre of mass control is reduced in older people when descending stairs at an increased riser height. Gait Posture 2019 Sep;73:305-314. [doi: 10.1016/j.gaitpost.2019.08.004] [Medline: 31415979]
42. List of tutorials has-motion software documentation. Has-motion. 2024. URL: https://wiki.has-motion.com/doku.php?id=visual3d:tutorials:list_of_tutorials [accessed 2025-08-05]
43. Kang H. Sample size determination and power analysis using the G*Power software. J Educ Eval Health Prof 2021;18:17. [doi: 10.3352/jeehp.2021.18.17] [Medline: 34325496]
44. Hair J, Black WC, Babin BJ, Anderson RE. Multivariate Data Analysis, 5th edition: Prentice-Hall; 1998.
45. Vallabhajosula S, Yentes JM, Stergiou N. Frontal joint dynamics when initiating stair ascent from a walk versus a stand. J Biomech 2012 Feb 2;45(3):609-613. [doi: 10.1016/j.jbiomech.2011.11.035] [Medline: 22172606]
46. Nadeau S, McFadyen BJ, Malouin F. Frontal and sagittal plane analyses of the stair climbing task in healthy adults aged over 40 years: what are the challenges compared to level walking? Clin Biomech (Bristol) 2003 Dec;18(10):950-959. [doi: 10.1016/s0268-0033(03)00179-7] [Medline: 14580839]
47. Mian OS, Thom JM, Narici MV, Baltzopoulos V. Kinematics of stair descent in young and older adults and the impact of exercise training. Gait Posture 2007 Jan;25(1):9-17. [doi: 10.1016/j.gaitpost.2005.12.014] [Medline: 16481170]
48. Reeves ND, Spanjaard M, Mohagheghi AA, Baltzopoulos V, Maganaris CN. Older adults employ alternative strategies to operate within their maximum capabilities when ascending stairs. J Electromyogr Kinesiol 2009 Apr;19(2):e57-e68. [doi: 10.1016/j.jelekin.2007.09.009] [Medline: 18053743]
49. Chiu SL, Chang CC, Dennerlein JT, Xu X. Age-related differences in inter-joint coordination during stair walking transitions. Gait Posture 2015 Jul;42(2):152-157. [doi: 10.1016/j.gaitpost.2015.05.003] [Medline: 26043669]
50. Tomioka K, Kurumatani N, Hosoi H. Association between stairs in the home and instrumental activities of daily living among community-dwelling older adults. BMC Geriatr 2018 Jun 4;18(1):132. [doi: 10.1186/s12877-018-0830-3] [Medline: 29898678]
51. Mustafaoğlu R, Unver B, Karatosun V. Evaluation of stair climbing in elderly people. J Back Musculoskelet Rehabil 2015;28(3):509-516. [doi: 10.3233/BMR-140549] [Medline: 25335569]

Abbreviations

ADL: activities of daily living
CoM: center of mass
F-S: floor-to-stair transition
IADL: instrumental activities of daily living

XSL•FO
RenderX

KMO: Kaiser-Meyer-Olkin
MMSE: Mini-Mental State Examination
PC: principal component
PCA: principal component analysis
PCM: principal component model
ROM: range of motion
S-F: stair-to-floor transition
SRH: self-reported health status
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
Δvelocity: angular velocity range

Edited by J Wang; submitted 22.01.25; peer-reviewed by A Assi, S Huang; revised version received 14.05.25; accepted 07.06.25; published 27.08.25.

Please cite as:
Moreira J, Teles IDS, Cunha B, Félix J, Guedes DC, Alves LAT, Santos R, Sousa ASP
Principal Component Analysis of Stair Negotiation and Floor Transition Kinematics in Older Adults With and Without Functional Disability: Cross-Sectional Study
JMIR Aging 2025;8:e71530
URL: https://aging.jmir.org/2025/1/e71530
doi:10.2196/71530

© Juliana Moreira, Ivone da Silva Teles, Bruno Cunha, José Félix, Diana C Guedes, Leonel A T Alves, Rubim Santos, Andreia S P Sousa. Originally published in JMIR Aging (https://aging.jmir.org), 27.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effectiveness of Iso-Inertial Resistance Training on Muscle Power in Middle-Older Adults: Randomized Controlled Trial

Aïda Cadellans Arroniz, PhD; Daniel Romero Rodríguez, PhD; Víctor Zárate, BS; Flora Dantony, MS; Marc Madruga Parera, PhD; Silvia Ortega Cebrian, PhD; David Blanco, PhD

Department of Physiotherapy, Universitat Internacional de Catalunya, Carrer de Josep Trueta, Sant Cugat del Vallès, Barcelona, Spain

Corresponding Author:
Daniel Romero Rodríguez, PhD
Department of Physiotherapy, Universitat Internacional de Catalunya, Carrer de Josep Trueta, Sant Cugat del Vallès, Barcelona, Spain

Abstract

Background: Resistance training is commonly used to prevent the decline in muscle power associated with aging.

Objective: This study aims to evaluate the effectiveness of iso-inertial (IN) training on power, physical performance, and variables associated with the risk of falls, compared to gravitational (GR) training, in physically active middle-older adults.

Methods: A parallel-group, randomized controlled trial was conducted at Espai Esport Wellness Center (Granollers, Spain). In total, 44 physically active adults (age >57) were randomized 1:1 to either the IN (n=21) or GR (n=23) training groups (using R software; R Core Team). Participants completed a 6-week training program (2 sessions/week) consisting of 3 exercises (forward lunge, side lunge, and forward lunge with row). The primary outcome includes power in the eccentric phase of each exercise, evaluated using both IN and GR devices. Secondary outcomes include concentric power, physical performance, and variables associated with the risk of falls. Only outcome evaluators were blinded. We used multivariate linear regression models for the analysis.

Results: In total, 27 participants completed the program (IN: n=15 IN; GR: n=12). IN training resulted in better eccentric power gains compared to GR training when assessed using the IN system, although the difference was only statistically significant for the side lunge. For forward lunge, between-group difference was 4.50 W (95% CI −2.94 to 11.94 W, *P*=.23); for side lunge, between-group difference was 9.24 W (95% CI 2.99-15.49 W; *P*<.01); and for forward lunge with row, between-group difference was 15.25 W (95% CI −0.63 to 31.13 W; *P*=.06). We observed no significant differences for the eccentric power using the GR system evaluation, concentric power, physical performance, and variables associated with the risk of falls. Both groups showed significant improvements from baseline across all outcomes.

Conclusions: Although IN training appeared to result in greater power gains during the eccentric phase when assessed with the IN system, statistically significant differences were observed only for the side lunge exercise. Both training systems seemed equally effective in improving eccentric power as evaluated with the GR system, concentric power, physical performance, and reducing variables associated with the risk of falls.

Trial Registration: ClinicalTrials.gov NCT06160089; https://clinicaltrials.gov/study/NCT06160089

(JMIR Aging 2025;8:e66414) doi:10.2196/66414

KEYWORDS

resistance training; power; middle-older adults; iso-inertial training; flywheel

Introduction

Different indicators show that the European Union population is increasingly aging. The median age has risen from 38 years in 2001 to 41 years in 2010 and 44 years in 2020. Additionally, the proportion of older adults in the population has grown, with 21% of the population being 65 years or older in 2020, compared to 16% in 2001, representing an increase of 5 percentage points [1]. In Spain, it is projected that the percentage of the population aged 65 and over will peak at 30.5% around 2055 [2]. This, combined with the population's progressive increase in life expectancy, has led to a growing interest in aging-related issues in recent years.

The World Health Organization recommends physical activity (PA) for the prevention of mortality, cardiovascular diseases, hypertension, specific cancers, type 2 diabetes, as well as for mental health (reducing symptoms of anxiety and depression), cognitive health, sleep quality, and potential improvements in measures of adiposity [3]. PA is also advised for managing health outcomes specific to older adults, such as fall prevention, fall-related injuries, physical function, frailty, and osteoporosis, which are consequences of the physiological decline in physical

XSL•FO
RenderX

capacity [3]. Recent evidence shows that PA can reduce the fall rate in older adults by up to 23%, significantly decreasing the risk of fall-related injuries, including serious falls resulting in bone fractures, head trauma, open wounds, soft tissue injuries, or other injuries requiring medical attention or hospitalization [4].

After reaching its peak at the beginning of adulthood, muscle and bone mass begin to decrease over the years, accompanied by losses in muscle strength and power [5-7]. Recent research has revealed a significant decrease in muscle strength of 1.5% - 5% each year from age 50 [8]. For this reason, the early implementation of preventive strategies, such as promoting PA, should focus not only on older adults but also on middle-aged individuals. In addition, older adults seem to preserve the ability to produce eccentric force better than concentric or isometric force [9]. The mechanisms underlying this phenomenon appear to be both mechanical and cellular [10]. The age-related accumulation of noncontractile material in the muscle-tendon unit increases passive stiffness, which may offer a mechanical advantage during eccentric contractions [11]. This makes eccentric strength training a powerful tool for exercising or restoring muscle strength in individuals who may have a limited capacity to train at high intensities, such as older adults [9]. Additionally, some authors have suggested that eccentric contractions are more effective because they provide intense muscular work with lower metabolic expense [9]. It has also been observed that, compared to concentric-based exercises, eccentric-based exercises lead to greater maximal strength with less muscle activation and a more significant increase in muscle mass [12]. To maintain an upright position or during certain activities of daily living, such as getting up and down from a chair or walking down stairs, the lower limb muscles function in a controlled manner against the force of gravity, which requires eccentric muscle contractions [13]. Indeed, it has been suggested that maximal knee and ankle eccentric strength may be critical for safe stair descent in older adults [14].

The study of muscle power in older adults has gained interest in recent years. Power, defined as the ability to exert force in a short time interval, appears to be more closely associated with certain mobility tasks than muscle strength [15-17]. It is considered a predictor of functional capacity, as it is also linked to the execution of activities of daily living such as climbing stairs, standing up from a chair, or walking [18,19]. Previous studies have reported that power diminishes more than strength over time [12].

Resistance training is one of the main strategies used to prevent decreased functional capacity [12]. The gravity-dependent resistance exercise method, in which resistance is opposed through gravity-dependent free weights or weight machines, has the limitation that the workload applied in the concentric phase of the lengthening-shortening cycle conditions the muscle capacity progression in the eccentric phase [20]. This limitation reduces the potential of this method for generating improvements associated with eccentric training [13]. In contrast, the iso-inertial (IN) training method generates resistance through an IN device, where the inertia of a rotating mass provides the workload. Unlike the gravitational (GR) system, the IN method can generate greater force during the eccentric phase of an exercise, a phenomenon known as eccentric overload. Due to the force-generating system of IN devices, high workloads can be applied in both phases of a given action, whereas eccentric overload in GR systems can only be achieved with external assistance [13,21]. Another benefit of the IN method is that, unlike GR training, it ensures accommodated resistance and optimal muscle loading [22].

Maroto-Izquierdo et al [20] conducted a systematic review of randomized controlled trials to evaluate the effects of IN training among athletes and healthy subjects, reporting increased muscle strength, power, and muscle size compared to GR training. Nevertheless, they highlighted some methodological limitations in the studies included, such as the difficulty in isolating the eccentric phase from the concentric phase and the fact that trial participants, therapists, and outcome assessors were typically not blinded—that is, they were aware of the intervention assignment. These factors may have introduced bias into the results. More recent studies have also focused on older adults, showing improvements in postural control, maximal isometric strength, isokinetic power, and other factors in participants who trained using IN systems, compared to those who trained with GR systems [23-25].

However, the interventions in these studies focused on analytical exercises (eg, leg curl and leg extension), which did not involve large muscle groups working synergistically and in a coordinated manner. Functional exercises focus on multijoint, complex, and dynamic movements, offering a promising approach for this population [12]. These exercises are designed to improve overall functionality and facilitate the performance of daily tasks, such as lifting objects, bending down, or walking. To date, no study has compared the impact of IN and GR training on power during functional exercises or on physical performance variables in middle-older adults.

The primary objective of this study was to evaluate the effectiveness of an IN resistance training program on eccentric power compared to the same program executed with GR resistance in physically active middle-older adults. The secondary objective was to compare these 2 programs in terms of concentric power, physical performance, and variables associated with the risk of falls.

Methods

Trial Design and Setting

This was a parallel-group, randomized controlled trial (allocation ratio 1:1) with a blinded outcome assessment, conducted at the Espai Esport Wellness Center gymnasium (Granollers, Spain). This study was approved by the Drug Research Ethics Committee of the Universitat Internacional de Catalunya (Code: FIS-2023 - 03)

We followed the CONSORT (Consolidated Standards of Reporting Trials) guidelines to produce this report [26] (the CONSORT-EHEALTH [Consolidated Standards of Reporting Trials of Electronic and Mobile Health Interventions and Online Telehealth] checklist is provided in Checklist 1).

XSL•FO
RenderX

Participants

We recruited physically active middle-older adults (aged 57 y or older) who engaged in at least 2 days of moderate-intensity aerobic PA or strength training per week [3]. In accordance with United Nations standards, we define middle adults as those aged between 40 and 60 years, and older adults as those aged 60 and older [27]. Participants younger than 60 years of age were considered for inclusion because muscle and bone mass loss, as well as reduced physical capacity, have been reported starting at around 50 years of age [8]. The initial phase of recruitment targeted individuals aged 60 and older. However, as the recruitment period neared its end, we lowered the age threshold to maximize participation before the deadline.

Participants with osteoarticular or acute musculoskeletal injuries, or those with systemic neurodegenerative diseases, were excluded. Recruitment was carried out in collaboration with the administrative staff of the gymnasium, who contacted all members over 57 years old and invited them to participate. Participants had to sign an informed consent form before participating in the study.

Interventions

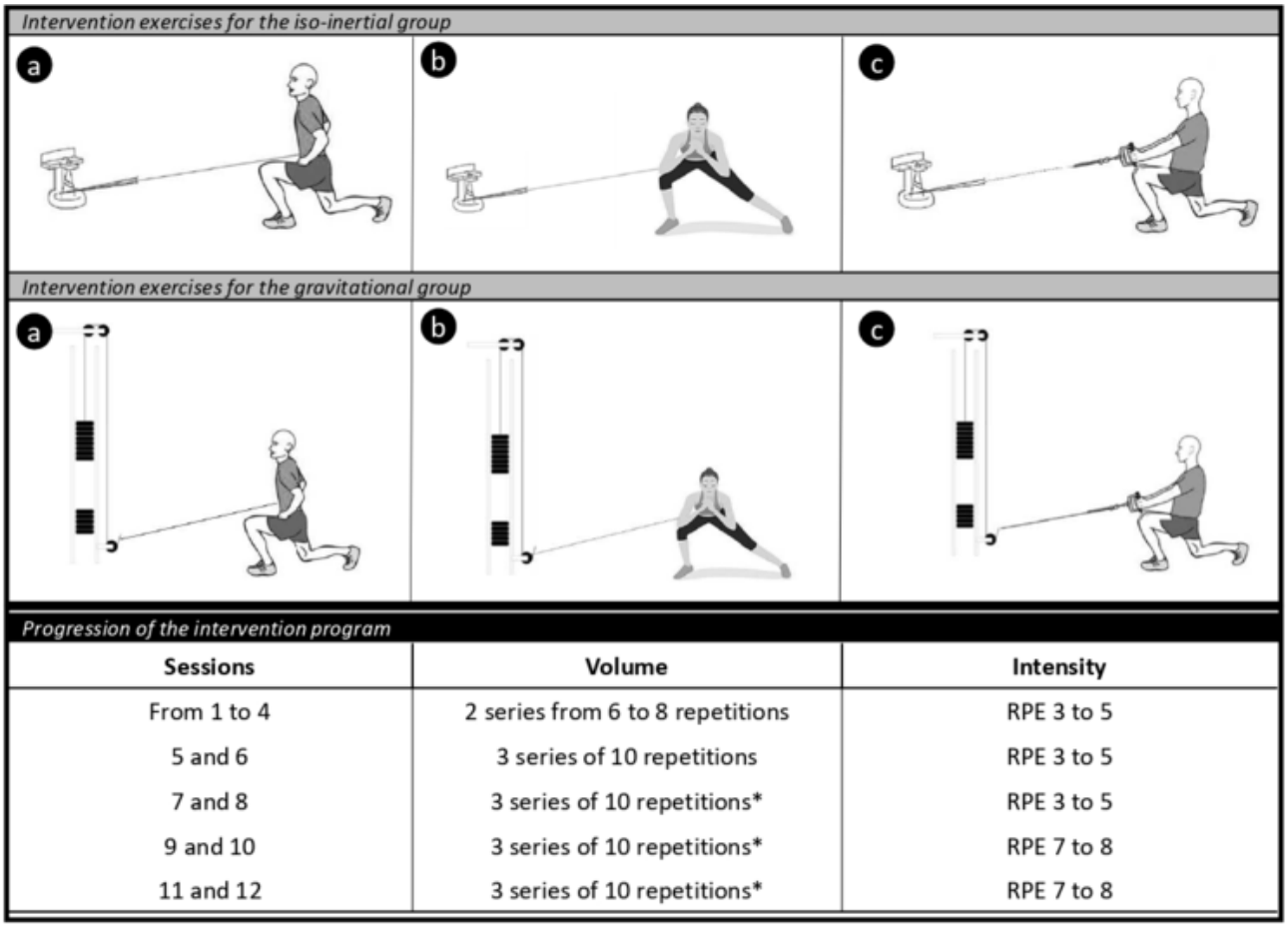
Participants completed a 6-week program (2 sessions per week separated by at least 48 h) using either an IN or a GR resistance device. The 6-week duration was chosen based on previous studies evaluating the effectiveness of IN resistance training in older adults [19,28-30].

Each session started with a warm-up consisting of (1) 4 minutes of moderate aerobic exercise (treadmill, elliptical bike, or stationary bike), (2) active stretching exercises for 6 seconds each (hip adductors, hamstrings, gastrocnemius, quadriceps, and gluteus), and (3) 6‐8 repetitions of each of the 3 intervention exercises without resistance to familiarize participants with the exercises.

Regardless of the assigned group, participants were required to perform the same 3 intervention exercises: forward lunge, side lunge, and forward lunge with row (Figure 1). Each exercise was performed individually. The IN group used the Nessinertial conical pulley (6 inertial loads), while the GR group used the MFT CSX-5000 device (weight: 3.75 kg). Each exercise was performed with both limbs. The starting limb for each exercise was randomized for each subject. Study investigators instructed participants to perform the exercises at a specific intensity based on the Borg Rating of Perceived Exertion (RPE) scale (0‐10, where 0 means no exertion at all and 10 means maximal exertion) [31]. In the first 2 weeks, participants were instructed to perform 2 sets of 6-8 repetitions at a 3‐5 RPE. In week 3, the program was increased to 3 sets of 10 repetitions at 3‐5 RPE. In week 4, volume and intensity remained the same as week 3, but participants performed lunges at 3 different distances marked on the floor (short, normal, and long), alternating between them. Weeks 5 and 6 followed a similar structure to weeks 3 and 4, but participants performed the exercises at 7‐8 RPE. A rest period of 30 seconds to 1 minute between sets was indicated. All this information is summarized in Figure 1. The load was identical for all participants, but each adjusted the execution speed of each exercise to achieve the desired RPE. More detailed information on the exercises and devices is provided in Multimedia Appendix 1. Participants were instructed not to perform any other lower limb resistance training during the intervention period, but they were free to engage in their usual aerobic exercise and upper limb resistance training. We did not monitor the training performed outside the study.

XSL•FO
RenderX

Figure 1. Intervention exercises and progression of the intervention program. (**A**) Forward lunge, (**B**) side lunge, and (**C**) forward lunge with row. RPE: rate of perceived exertion; *lunges performed at different distances.



Progression of the intervention program

Sessions	Volume	Intensity
From 1 to 4	2 series from 6 to 8 repetitions	RPE 3 to 5
5 and 6	3 series of 10 repetitions	RPE 3 to 5
7 and 8	3 series of 10 repetitions*	RPE 3 to 5
9 and 10	3 series of 10 repetitions*	RPE 7 to 8
11 and 12	3 series of 10 repetitions*	RPE 7 to 8

All participants were supervised by one of the study investigators or by the center's instructors until they were shown to be autonomous. From that point onward, the sessions were performed without supervision. Additionally, participants were provided with documentary information, including images and videos of the exercises. We monitored the completion of the training program by asking participants to contact us after each completed session via WhatsApp or phone call. This communication also allowed participants to report any potential adverse events. In response, we acknowledged their feedback by congratulating them on completing the session and motivating them by providing a countdown of the remaining sessions in the training program. If participants missed a session, they were instructed to perform an extra session during the following week, ensuring a 48-hour break between sessions. Adherence was assessed by dividing the number of completed sessions by the total number of sessions expected.

Outcomes and Instruments

Outcomes were assessed at the gymnasium facilities between 2 and 5 days before the intervention began (T0) and 2 to 5 days after its completion (T1). All the evaluation instruments described below are valid and reliable.

The primary outcome was the power in the eccentric phase (ie, eccentric power) of each exercise. This was measured using both IN and GR resistance devices, with a rotatory and a linear encoder (Chronojump Boscosystem; r=0.98, 90% CI 0.98 - 0.99) [32], respectively. We used the free software Chronojump, in conjunction with the open hardware Chronopic V.3, to record the data. For each exercise, Chronojump provided the average power for both the eccentric and concentric phases of each repetition, calculated as the average speed multiplied by the average force for each phase. The average speed was determined by dividing the distance covered by the time taken to perform each phase. For the statistical analysis, we calculated the eccentric power for each exercise by taking the mean of the average power during the eccentric phase of the 3 best consecutive repetitions. These 3 repetitions were selected based on having the highest combined eccentric-concentric average power. Importantly, the decision to compute the average power instead of the maximum power for each phase and repetition, as well as the mean of the 3 repetitions, was aimed at providing a stable measure that minimized the impact of power peaks and fluctuations within a repetition, thus preventing atypical results or computational noise from distorting the power values.

The secondary outcomes were (1) power in the concentric phase (ie, concentric power) of each exercise, measured with both the IN and GR devices, and calculated in a manner analogous to the procedure described earlier for the eccentric power; (2) physical performance, assessed using the Short Physical Performance Battery (SPPB; intraclass correlation coefficient=0.839 - 0.940) [33], which includes a balance test, a gait speed test, and a five-times sit-to-stand test; and (3) risk of falls, evaluated through the timed up and go test (r:

XSL•FO
RenderX

0.51 - 0.78) [34]. Participants made a single attempt for each of these tests, which are further explained in Multimedia Appendix 2. For the muscle power variables, participants were instructed to perform the exercises at their maximum possible speed, while for the functional tests, we followed the verbal instructions provided in the test protocols [33,34].

We also collected anthropometric and sociodemographic data, including date of birth, weight (kg), height (cm), sex (male or female), and work status (working or retired).

Sample Size

Given that no studies have compared eccentric power between IN and GR resistance training protocols in middle-older adults, we used a convenience sample of 44 participants, which was the maximum sample size we could achieve within the established recruitment period (November 2023 to February 2024).

Randomization (Sequence Generation, Allocation Concealment Mechanism, and Implementation) and Blinding

The team's statistician (DB) created the randomization sequence using R Statistical Software (R Core Team, version 4.5.0) in blocks of 4, stratified by sex and age (57 - 63, 64 - 70, and 71 years or older). This sequence was matched by the administrative staff of the center to the participants' membership numbers, according to their recruitment order. As a result, both participants and researchers were unaware of the allocation until the training program started. Neither participants nor investigators monitoring the training protocols could be blinded to the allocation. However, the outcome assessors were blinded.

Statistical Methods

For the statistical analysis, we used R Statistical Software. Initially, we presented descriptive data for all included subjects, those who completed the study, and those who were lost to follow-up. We then applied linear regression models, where the dependent variables were the change scores of each outcome (T1 minus T0), and the independent variables were the allocation group (IN or GR), baseline score, limb, and selected demographic variables (age and sex). We checked the model application assumptions (linearity, homoscedasticity, normality, and independence). Since the characteristics of the participants lost to follow-up were similar to those who completed the study, we performed a complete case analysis. We reported the adjusted between-group difference of means, along with its 95% CI and the *P* value of the comparison. We used an alpha level of .05 for statistical significance. Additionally, we calculated the intragroup improvement (both absolute value and percentage) for each study outcome. All R scripts are publicly available [23].

Ethical Considerations

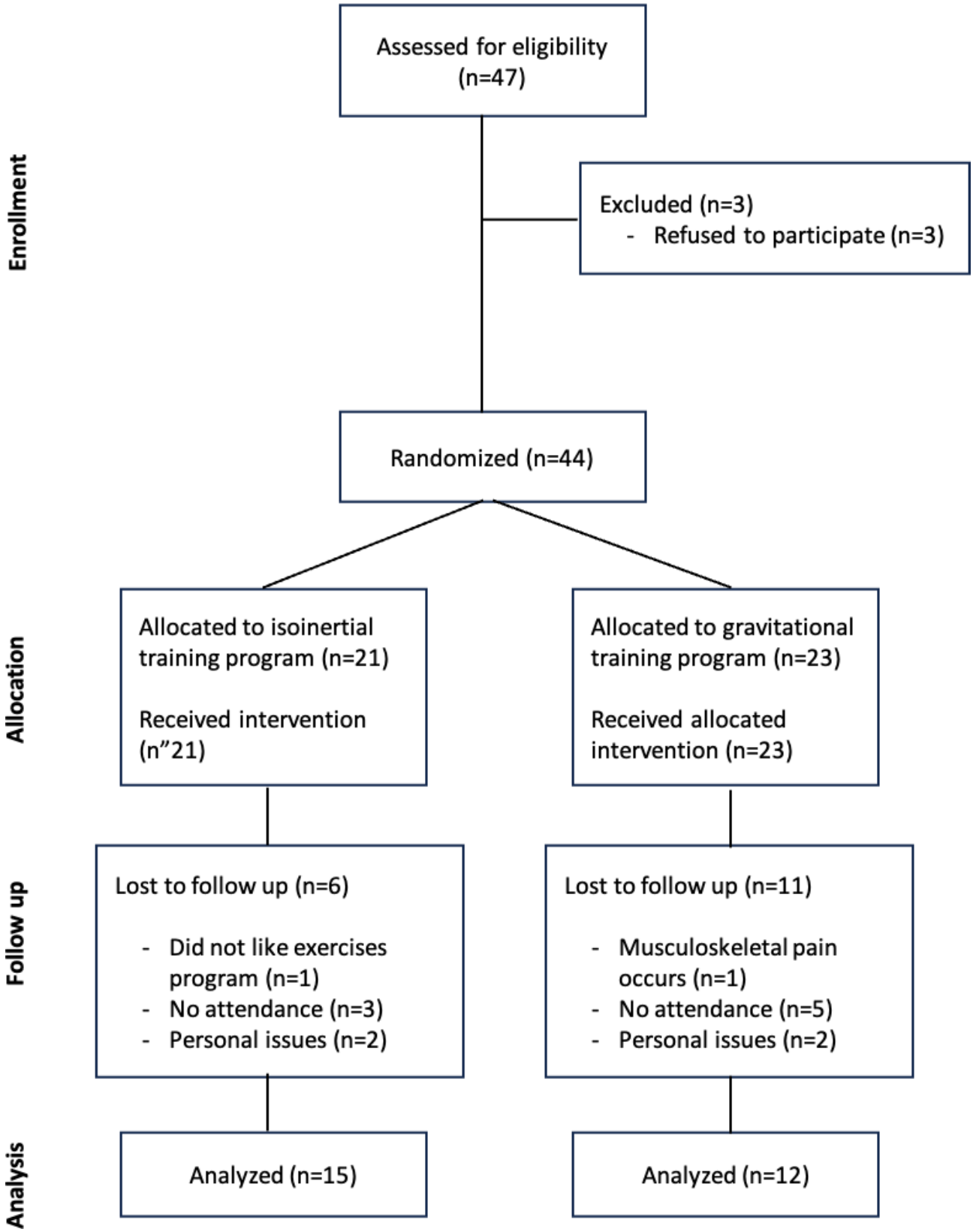
This study was approved by the Drug Research Ethics Committee of the Universitat Internacional de Catalunya (approval number: FIS-2023 - 03). The protocol can be found in Multimedia Appendix 3. Participants provided informed consent to participate in the study before taking part (Multimedia Appendix 4). No identifiable features of research participants or users have been shared in any images of the manuscript or supplementary materials. No compensation was provided to subjects for participation.

Results

Between November 10, 2023, and February 28, 2024, we recruited 47 participants. Of them, 3 refused to participate, leaving 44 participants to be randomized: 21 to the IN group and 23 to the GR group. In total, 17 patients were lost to follow-up (6 in the IN group, 11 in the GR group). The reasons for loss to follow-up were as follows: (1) failure to attend training sessions (n=8), (2) personal problems (n=4), (3) musculoskeletal problems (n=4), and (4) dissatisfaction with the training program (n=1). In the end, 27 participants completed the program (15 in the IN group and 12 in the GR group) (Figure 2). All of the completers had 100% adherence to the program.

XSL•FO
RenderX

Figure 2. Explanation of the short physical performance battery and timed up and go tests.



The mean (SD) age of participants who completed the program was 63.52 (6.76) years. Of these, 16 out of 27 participants (59.2%) were female, and 11 (40.7%) were male. The baseline characteristics of all included participants were similar to those of the participants who completed the program and to those lost to follow-up (Table 1).

XSL•FO
RenderX

Table . Baseline characteristics of the participants.

	Total participants included (n=44)		Participants with full program completed (n=27)		Participants lost to follow-up (n=17)	
	IN[a] group (n=21)	GR[b] group (n=23)	IN group (n=15)	GR group (n=12)	IN group (n=6)	GR group (n=11)
Age (years), mean, SD	65.00 (6.36)	62.74 (6.87)	64.53 (6.90)	62.25 (6.61)	67.80 (3.56)	62.83 (7.25)
Gender						
Female, n (%)	13 (61.90)	12 (52.17)	8 (53.33)	8 (66.67)	5 (83.33)	4 (36.36)
Male, n (%)	8 (38.10)	11 (47.83)	7 (46.67)	4 (33.33)	1 (16.66)	7 (63.63)
Weight (kg), mean (SD)	68.18 (10.28)	73.09 (11.18)	68.41 (9.35)	70.67 (0.59)	1.68 (0.06)	73.90 (16.51)
Height (m), mean (SD)	1.65 (0.08)	1.69 (0.08)	1.65 (0.06)	1.67 (0.10)	1.66 (0.15)	1.66 (0.15)
Status						
Active, n (%)	10 (47.62)	15 (65.22)	8 (53.33)	8 (66.67)	3 (50)	6 (54.54)
Retired, n (%)	11 (52.38)	8 (34.78)	7 (46.67)	4 (33.33)	3 (50)	5 (45.45)

[a]IN: iso-inertial.

[b]GR: gravitational.

Primary Outcome

Although improvements in eccentric phase power (evaluated using the IN system) were greater in the IN group across all exercises—ranging from 78.61% to 101.86%—compared to the GR group (55.43% to 86.10%), only the side lunge showed a statistically significant between-group difference ($P<.05$), as reported in Table 2. For the forward lunge, the mean eccentric power at T1 was 42.67 W in the IN group and 35.59 W in the GR group (between-group difference: 4.50 W, 95% CI −2.94 to 11.94 W; $P=.23$). For the side lunge, the mean eccentric power was 51.06 W in the IN group and 40.22 W in the GR group (between-group difference: 9.24 W, 95% CI 2.99 to 15.49 W; $P<.01$). For the forward lunge with row, the mean eccentric power was 128.39 W in the IN group and 102.41 in the GR group (between-group difference: 15.25 W, 95% CI −0.63 -to 31.13 W; $P=.06$).

XSL•FO
RenderX

Table . Intra- and intergroup analysis of power changes in the eccentric phase (in W).

Exercise	GR[a] group				IN[b] group				Differences between groups (95% CI)	*P* value
	T0, mean (SD)	T1, mean (SD)	Change T1-T0	% Change T1-T0	T0, mean (SD)	T1, mean (SD)	Change T1-T0	% Change T1-T0		
IN										
Forward lunge	19.12 (9.37)	35.59 (21.11)	16.47	86.10	21.14 (9.52)	42.67 (15.78)	21.53	101.86	4.50 (−2.94 to 11.94)[c]	.23
Side lunge	25.52 (12.21)	40.22 (15.25)	14.70	57.59	26.75 (12.04)	51.06 (19.37)	24.31	90.88	9.24 (2.99 to 15.49)[c]	<.001
Forward lunge with row	65.89 (45.58)	102.41 (62.81)	36.52	55.43	71.89 (40.08)	128.39 (51.49)	56.51	78.61	15.25 (−0.63 to 31.13)[c]	.06
GR										
Forward lunge	21.52 (6.36)	26.08 (8.03)	4.56	21.18	21.53 (6.50)	24.99 (5.41)	3.47	16.10	−0.51 (−2.59 to 1.58)	.63
Side lunge	19.11 (4.26)	21.75 (4.22)	2.64	13.80	19.39 (4.07)	23.03 (4.40)	3.64	18.77	0.69 (−1.20 to 2.58)[c]	.47
Forward lunge with row	42.37 (16.38)	60.08 (27.24)	17.71	41.81	42.41 (18.57)	55.58 (16.79)	13.16	31.03	−2.94 (−12.76 to 6.87)	.55

[a]GR: gravitational.

[b]IN: iso-inertial.

[c]Favors the IN group.

When using the GR system for evaluation, the IN group (ranging from 13.80% to 41.81% improvement) and the GR group (ranging from 16.10% to 31.03% improvement) showed similar eccentric power improvements (Table 2). For the forward lunge, the mean eccentric power was 24.99 W in the IN group and 26.08 W in the GR group (between-group difference: −0.51 W, 95% CI −2.59 to 1.58 W; *P*=.63). For the side lunge, the mean eccentric power was 23.03 W in the IN group and 21.75 W in the GR group (between-group difference: 0.69 W, 95% CI −1.20 to 2.58; *P*=.47). For the forward lunge with row, the mean eccentric power was 55.58 W in the IN group and 60.08 W in the GR group (between-group difference: −2.94 W, 95% CI −12.76 to 6.87; *P*=.55).

Secondary Outcomes

The power in the concentric phase improved similarly in both the IN group (with improvements ranging from 47.81% to 81.73% in the IN system and from 5.66% to 61.23% in the GR system) and the GR group (with improvements ranging from 57.81% to 67.61% in the IN system and from 10.09% to 55.97% in the GR system) (Table 3). Additionally, improvements were observed in both groups, but no significant between-group differences were found in physical performance or in variables associated with the risk of falls (Table 4).

XSL•FO
RenderX

Table . Intra- and intergroup analysis of power changes in the concentric phase (in W).

Exercise	GR[a] group				IN[b] group				Differences between groups (95% CI)	*P* value
	T0, mean (SD)	T1, mean (SD)	Change T1-T0	% Change T1-T0	T0, mean (SD)	T1, mean (SD)	Change T1-T0	% Change T1-T0		
IN										
Forward lunge	21.74 (9.18)	39.51 (24.24)	17.77	81.73	26.55 (10.74)	42.29 (14.45)	15.74	59.27	−3.13 (−11.09 to 4.82)	.43
Side lunge	25.49 (12.40)	41.09 (19.00)	15.61	61.23	28.65 (12.66)	48.02 (17.43)	19.37	67.61	4.27 (−3.03 to 11.56)[c]	.25
Forward lunge with row	72.19 (48.28)	106.71 (60.33)	34.52	47.81	79.91 (44.43)	126.10 (50.11)	46.19	57.81	8.26 (−6.08 to 22.59)[c]	.25
GR										
Forward lunge	19.14 (5.43)	21.74 (4.74)	2.61	13.61	19.91 (5.27)	21.92 (5.09)	2.01	10.09	−0.20 (−1.97 to 1.56)	.82
Side lunge	17.43 (3.61)	18.42 (3.46)	0.99	5.66	17.45 (3.04)	19.60 (2.81)	2.15	12.31	0.91 (−0.12 to 1.94)[c]	.08
Forward lunge with row	43.44 (20.31)	71.44 (48.66)	28.00	64.47	52.60 (27.81)	82.04 (37.79)	29.44	55.97	7.46 (−12.42 to 27.34)	.45

[a]GR: gravitational.

[b]IN: iso-inertial.

[c]Favors the IN group.

XSL•FO
RenderX

Table . Intergroup analysis of physical fitness and variables associated with risk of falls changes.

Physical fitness test and out-come	GR[a] group				IN[b] group				Differences between groups (95% CI)	*P* value
	T0, mean (SD)	T1, mean (SD)	Change T1-T0	% Change T1-T0	T0, mean (SD)	T1, mean (SD)	Change T1-T0	% Change T1-T0		
SPPB[c]										
Score (0 to 12)	10.92 (1.16)	11.67 (0.65)	0.16	4.34	10.85 (0.80)	10.92 (1.32)	−0.07	−1.92	0.77 (−0.13 to 1.66)	.09
Balance (sec)	3.83 (0.58)	4.00 (0.00)	−0.09	−2.82	4.00 (0.00)	3.92 (0.28)	−0.31	−9.98	0.09 (−0.09 to 0.27)[f]	.31
Walking speed (sec)	3.18 (0.59)	3.09 (0.46)	−1.41	−11.54	3.18 (0.62)	2.86 (0.46)	−1.35	−10.73	0.20 (−0.11 to 0.51)[f]	.20
FTSST[d] (sec)	12.25 (2.01)	10.84 (2.18)	−0.255	−3.07	12.66 (2.08)	11.30 (2.36)	−0.66	−7.98	−0.39 (−1.55 to 0.76)	.49
Variables associated with risk of falls										
TUG[e] (seconds)	8.29 (1.60)	8.04 (1.02)	0.75	6.87	8.33 (1.28)	7.67 (1.20)	0.07	0.70	0.30 (−0.53 to 1.13)[f]	.46

[a]GR: gravitational.

[b]IN: iso-inertial.

[c]SPPB: Short Physical Performance Battery.

[d]FTSST: five times sit-to-stand test.

[e]TUG: time up and go test.

[f]Favors the IN group.

Discussion

Principal Results

IN training appeared to result in greater power gains during the eccentric phase compared to GR training when power was assessed using the IN system. However, statistically significant differences were observed only for the side lunge exercise. In contrast, when evaluated with the GR system, both types of training performed similarly. Additionally, we observed no differences in the concentric power, physical performance, and variables associated with the risk of falls. This is the first study in middle-older adults to (1) compare the effectiveness of IN and GR resistance training on the power of a coordinative action, (2) separately analyze the power in the eccentric and concentric phases of each exercise, and (3) assess power using both IN and GR devices regardless of the participant's training group.

Limitations

This research has several limitations. First, 38% of participants were lost to follow-up, preventing us from reaching a larger sample size. However, the characteristics of those were similar to those of participants who completed the study. Also, we did not measure other physiological parameters involved in functional and structural adaptations. Future studies could include these parameters to explore the underlying mechanisms behind the effects of IN and GR training. Although participants were warned about the incompatibility of performing lower limb strength training during the study period, the lack of monitoring other types of training outside the study may have introduced bias. Finally, even though the TUG test is a widely used tool to assess fall risk in older adults, its predictive validity remains controversial in the current literature. Therefore, our results with this variable are not entirely conclusive. A long-term follow-up of falls could provide more clarity on the issue.

Comparison With Prior Work

Despite the growing number of studies focusing on the training and assessment of muscle power in older adults [15,16,35], only 2 studies to date have directly compared IN and GR training using this parameter. However, neither of these studies assessed power during the eccentric phase of the movement. Therefore, their results can only be compared to ours with regard to concentric power. First, Floreani et al [36] observed that IN training led to an increase in absolute maximum explosive power (MEP) during a leg press compared to GR training. Specifically, the absolute MEP improved by +10.8% for the IN group versus +0.31% for the GR group. Power evaluation was performed using a GR system. The results for the IN group are similar to those we obtained for the frontal lunge (+10.09% improvement) and side lunge (+12.31% improvement), but not for the frontal lunge with row (+55.97% improvement) when evaluating concentric power using the GR system. In contrast, the results

XSL•FO
RenderX

for the GR group differ from ours. While they found little to no improvements, we observed significant gains. This discrepancy may be explained by various factors, such as differences in the outcome measure (they used MEP) and the training program. Second, the study by Sañudo et al [37] reported a 63% increase in power during a concentric action for the IN group (no data were provided for the GR group). In this case, power was evaluated using an iso-inertial system. Their results are consistent with ours, as we observed an increase in concentric power ranging from 59.27% to 67.61% for the IN group when evaluated using the IN system.

Importantly, there are 2 key differences between the 2 studies mentioned earlier and ours. First, those studies proposed training programs based on a single analytical exercise (squat or leg extension). In contrast, our protocol is functional, as it involves trunk stabilization and targets the major upper and lower extremity muscle groups, combining resistance and high-velocity exercises. This approach aims to have a greater transfer to activities of daily living. Additionally, our protocol, which includes aerobic exercise and a stretching routine in the warm-up, aligns with current position statements and consensus guidelines for PA in older adults. These guidelines recommend a multimodal approach that incorporates aerobic, strength, balance, and flexibility exercises [3].

Second, none of the 2 studies differentiated between eccentric and concentric power. Studies indicate that the preservation of muscle tension and the increased stiffness of muscle fibers in aging muscles contribute to greater active stiffness, which may enhance performance during eccentric actions. In fact, a 21.6% higher functional reserve of eccentric strength compared to concentric strength has been reported, which could be particularly relevant when initiating resistance training and rehabilitation programs for individuals with low strength levels [9,11]. Furthermore, it has been suggested that the loss of muscle power in the lower extremities may be a key factor in the limitation of mobility-related activities of daily living and in the onset of falls in older adults [16,17,38]. Given the implications of eccentric actions in terms of functionality, training, and risk of falls prevention in middle-older adults, a separate analysis of eccentric power should be included when evaluating training programs for this population.

The IN group appeared to achieve greater power in the eccentric phase than the GR group when power was assessed with the IN system. These results align with the theoretical foundation of IN training, which facilitates maximal concentric and eccentric muscle actions with brief episodes of eccentric overload [22]. Previous studies have reported that the peak force generated during the eccentric phase of movement can exceed that of the preceding concentric phase by 15% - 30%, likely due to the elastic energy storage properties inherent to the IN system [29,39]. However, our results should be interpreted with caution, as the improvements were only statistically significant in one of the 3 exercises. It is also important to note that participants' familiarity with the IN system may have introduced bias into the results. Consequently, future studies could incorporate a longer familiarization period for all participants prior to conducting the assessments.

No differences between groups were observed in eccentric power for the GR system evaluation. This may be because of the limited power generation in the eccentric phase with this system. Previous studies in athletes have reported improvements in power favoring IN training compared to GR training; however, both groups were evaluated only with the iso-inertial system [20,40,41]. The use of a single assessment tool may have influenced the results, as some participants were already familiar with the IN system, while others were not. By assessing our participants with both systems, we were able to avoid this bias. Indeed, if the differences favoring the IR training group were due to familiarization, the GR group would have achieved more power in the GR system evaluation. This may confirm that the eccentric overload generated by IN training results in genuinely higher eccentric power values.

We found that all participants improved their physical performance and increased their score on the TUG test, with no significant between-group differences. The GR group completed the test 3.08% faster, while the IN group took 7.98% less time. The TUG test is widely recommended for detecting fall risk and is commonly used in clinical practice and geriatric research [42]. However, there is ongoing controversy regarding the determination of the optimal cutoff value for identifying fall risk [43]. For this reason, some studies consider the TUG a tool for assessing functional mobility [44]. Given the debate over cutoff points, we cannot confirm that training reduces fall risk based on the TUG test. However, we can confirm that there was an improvement in the time it took participants to complete tasks involving functional mobility.

These results align with a systematic review that reported improvements in functional performance for all participants, with a slight advantage for eccentric-based compared to concentric-based exercises [13]. Another review found improvements in unipedal balance in older adults who used an IN device compared to those using a GR one [20]. Notably, these studies are focused on analytical exercises. However, our protocol proposed a more functional approach. First, we used lunges because they involve complex functions such as decelerating a limb, force absorption, and controlling movement against an external force [45]. These functions are key components of an eccentric action [46]. In addition, all exercises emphasized trunk stabilization, which is an important factor to consider when developing interventions for middle-older adults. Some studies have reported that core stability training for older adults can improve balance and coordination and reduce the risk of falls [47].

Few studies have compared the effects of performing the same functional exercises with different types of resistance. Madruga-Parera et al [48] evaluated the effectiveness of functional exercises in handball that were biomechanically identical for the 2 study groups, with only the type of resistance changed. Our protocol was based on this approach. We suggest that future studies incorporate this methodology when evaluating the differences between training methods to minimize performance biases.

The scales used for evaluating physical performance and the variables associated with the risk of falls lacked sensitivity for

XSL•FO
RenderX

the type of participants included, as we encountered a ceiling effect. For example, the participants' mean baseline SPPB score was 10.88 out of 12 points, indicating excellent physical condition. A similar ceiling effect was noted in another study, which stated that the Berg scale could not predict the risk of falls in individuals with high levels of balance ability [49]. Therefore, we suggest that scales originally designed for middle-older adults be modified for physically active middle-older adults.

Despite the limitations mentioned earlier, there are several strengths to highlight in this study: (1) the randomized controlled trial design, (2) the pragmatic design, which is ideal for testing the effectiveness of interventions under real-life conditions, (3) the outcome evaluations using 2 different systems to minimize familiarization effects, (4) the functional training program, and (5) the inclusion of identical exercises for both groups.

Conclusions

Although IN training appeared to result in greater power gains during the eccentric phase compared to GR training when assessed using the IN system, statistically significant differences were observed only for the side lunge exercise. There were no differences between the 2 methods in terms of eccentric power assessed using the GR system evaluation, nor for concentric power, physical performance, or variables associated with the risk of falls. Regardless of the training system, the resistance training program led to significant improvements in all outcomes.

This study helps further understand the effects of IN and GR resistance training, which can be useful for clinicians to prescribe more effective training programs for middle-older adults. Using IN devices for resistance training in middle-older adults may be a promising way to improve power during the eccentric phase of an action. Additionally, we provide strong empirical data that support orienting resistance training toward exercises involving complex coordinative actions (rather than analytical exercises) with a transfer to daily activities such as walking or going up and down stairs. This approach may increase middle-older adults' autonomy and therefore promote healthier longevity. Finally, we warn about the ceiling effect of the current clinical evaluation tests and suggest avoiding their use in middle-older adults. Instead, we propose to use power assessments in clinical practice to monitor the improvements associated with resistance training.

Acknowledgments

The authors would like to thank the staff of Espai Esport Wellness (Granollers, Spain) for allowing us to conduct this project in their facilities, as well as for assisting us with participant recruitment and monitoring the intervention.

Data Availability

Data were stored and encoded in 3 Excel files to protect confidentiality before, during, and after the study. The first of these files contains the personal data of the participants (name and surnames, age, and sex), is accessible only to the administrative staff of the centre, and served only to manage the recruitment of the participants. The second and third files are managed entirely by the research team and contain the registry of all the study outcomes. In these files, participants were identified by a code. Only the research team can relate the data collected in the study to the identity of the participants, which will not be available to anyone except in the event of a medical emergency or legal requirement. These two datasets are publicly available in Zenodo repository [50].

Authors' Contributions

Conceptualization for this study was done by AC, DR, MMP, and DB. AC and DB contributed to writing – original draft. AC, DB, DR, MMP, VZ, FD, and SO contributed to writing – review & editing. Methodology, software, data curation, and formal analysis were done by DB and VZ. Resources were provided by DR and MMP. Investigation was done by AC, DB, DR, and FD. Supervision and project administration were done by AC and DB.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Information about the execution of the exercises.
[DOCX File, 8 KB - aging_v8i1e66414_app1.docx]

Multimedia Appendix 2
Explanation of the Short Physical Performance Battery and Timed Up and Go tests.
[DOCX File, 7 KB - aging_v8i1e66414_app2.docx]

Multimedia Appendix 3
Study protocol approved by the ethics committee.

XSL•FO
RenderX

[DOCX File, 3463 KB - aging_v8i1e66414_app3.docx]

Multimedia Appendix 4
Informed consent.
[PDF File, 85 KB - aging_v8i1e66414_app4.pdf]

Checklist 1
CONSORT-EHEALTH checklist.
[PDF File, 1051 KB - aging_v8i1e66414_app5.pdf]

References

1. An ageing population. 2025. URL: https://ec.europa.eu/eurostat/cache/digpub/demography_2021/bloc-1c.html
2. Nota de prensa: proyecciones de población. Años 2024-2074. 2025. URL: https://www.ine.es/dyngs/Prensa/PROP20242074.htm [accessed 2025-08-13]
3. RECOMENDACIONES. In: Directrices de La OMS Sobre Actividad Física y Comportamientos Sedentarios - NCBI Bookshelf: World Health Organization; 2025. URL: https://www.ncbi.nlm.nih.gov/books/NBK581973 [accessed 2025-08-13]
4. Sherrington C, Fairhall NJ, Wallbank GK, et al. Exercise for preventing falls in older people living in the community. Cochrane Database Syst Rev 2019 Jan 31;1(1):CD012424. [doi: 10.1002/14651858.CD012424.pub2] [Medline: 30703272]
5. Suchomel TJ, Nimphius S, Bellon CR, Stone MH. The importance of muscular strength: training considerations. Sports Med 2018 Apr;48(4):765-785. [doi: 10.1007/s40279-018-0862-z]
6. Chuang YF, Chen CC, Hsu MJ, et al. Age related changes of the motor excitabilities and central and peripheral muscle strength. J Electromyogr Kinesiol 2019 Feb;44(November 2018):132-138. [doi: 10.1016/j.jelekin.2018.12.007]
7. Foldvari M, Clark M, Laviolette LC, et al. Association of muscle power with functional status in community-dwelling elderly women. J Gerontol A Biol Sci Med Sci 2000 Apr;55(4):M192-M199. [doi: 10.1093/gerona/55.4.m192] [Medline: 10811148]
8. Keller K, Engelhardt M. Strength and muscle mass loss with aging process. Age and strength loss. Muscle Ligaments Tendons J 2013;03(4):346. [doi: 10.32098/mltj.04.2013.17]
9. Roig M, Macintyre DL, Eng JJ, Narici MV, Maganaris CN, Reid WD. Preservation of eccentric strength in older adults: evidence, mechanisms and implications for training and rehabilitation. Exp Gerontol 2010 Jun;45(6):400-409. [doi: 10.1016/j.exger.2010.03.008] [Medline: 20303404]
10. Ochala J, Dorer DJ, Frontera WR, Krivickas LS. Single skeletal muscle fiber behavior after a quick stretch in young and older men: a possible explanation of the relative preservation of eccentric force in old age. Pflugers Arch - Eur J Physiol 2006 Jul;452(4):464-470. [doi: 10.1007/s00424-006-0065-6]
11. Roig M, O'Brien K, Kirk G, et al. The effects of eccentric versus concentric resistance training on muscle strength and mass in healthy adults: a systematic review with meta-analysis. Br J Sports Med 2009 Aug;43(8):556-568. [doi: 10.1136/bjsm.2008.051417]
12. Fragala MS, Cadore EL, Dorgo S, et al. Resistance training for older adults: position statement from the National Strength and Conditioning Association. J Strength Cond Res 2019;33(8):2019-2052. [doi: 10.1519/JSC.0000000000003230]
13. Kulkarni D, Gregory S, Evans M. Effectiveness of eccentric-biased exercise interventions in reducing the incidence of falls and improving functional performance in older adults: a systematic review. Eur Geriatr Med 2022 Apr;13(2):367-380. [doi: 10.1007/s41999-021-00571-8] [Medline: 34633637]
14. Tinetti ME, Speechley M, Ginter SF. Risk factors for falls among elderly persons living in the community. N Engl J Med 1988 Dec 29;319(26):1701-1707. [doi: 10.1056/NEJM198812293192604]
15. Coelho-Júnior HJ, Uchida MC. Effects of low-speed and high-speed resistance training programs on frailty status, physical performance, cognitive function, and blood pressure in prefrail and frail older adults. Front Med 2021;8 [FREE Full text] [doi: 10.3389/fmed.2021.702436]
16. Marsh AP, Miller ME, Saikin AM, et al. Lower extremity strength and power are associated with 400-meter walk time in older adults: The InCHIANTI study. J Gerontol A Biol Sci Med Sci 2006 Nov;61(11):1186-1193. [doi: 10.1093/gerona/61.11.1186] [Medline: 17167161]
17. Bean JF, Leveille SG, Kiely DK, Bandinelli S, Guralnik JM, Ferrucci L. A comparison of leg power and leg strength within the InCHIANTI study: which influences mobility more? J Gerontol A Biol Sci Med Sci 2003 Aug;58(8):728-733. [doi: 10.1093/gerona/58.8.m728] [Medline: 12902531]
18. Balachandran A, Martins MM, De Faveri FG, Alan O, Cetinkaya F, Signorile JF. Functional strength training: seated machine vs standing cable training to improve physical function in elderly. Exp Gerontol 2016 Sep;82:131-138. [doi: 10.1016/j.exger.2016.06.012]
19. Naczk M, Marszalek S, Naczk A. Inertial training improves strength, balance, and gait speed in elderly nursing home residents. Clin Interv Aging 2020;15:177-184. [doi: 10.2147/CIA.S234299] [Medline: 32103920]

XSL•FO
RenderX

20. Maroto-Izquierdo S, García-López D, Fernandez-Gonzalo R, Moreira OC, González-Gallego J, de Paz JA. Skeletal muscle functional and structural adaptations after eccentric overload flywheel resistance training: a systematic review and meta-analysis. J Sci Med Sport 2017 Oct;20(10):943-951. [doi: 10.1016/j.jsams.2017.03.004]
21. Fiorilli G, Mariano I, Iuliano E, et al. Isoinertial eccentric-overload training in young soccer players: effects on strength, sprint, change of direction, agility and soccer shooting precision. J Sports Sci Med 2020 Mar;19(1):213-223. [Medline: 32132845]
22. Tesch PA, Fernandez-Gonzalo R, Lundberg TR. Clinical applications of iso-inertial, eccentric-overload (YoYoTM) resistance exercise. Front Physiol 2017:241. [doi: 10.3389/fphys.2017.00241] [Medline: 28496410]
23. Onambélé GL, Maganaris CN, Mian OS, et al. Neuromuscular and balance responses to flywheel inertial versus weight training in older persons. J Biomech 2008 Nov;41(15):3133-3138. [doi: 10.1016/j.jbiomech.2008.09.004]
24. Brzenczek-Owczarzak W, Naczk M, Arlet J, Forjasz J, Jedrzejczak T, Adach Z. Estimation of the efficacy of inertial training in older women. J Aging Phys Act 2013;21(4):433-443. [doi: 10.1123/japa.21.4.433]
25. Kowalchuk K, Butcher S. Eccentric overload flywheel training in older adults. J Funct Morphol Kinesiol 2019 Aug 22;4(3):61. [doi: 10.3390/jfmk4030061] [Medline: 33467376]
26. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010 Mar 23;340:c332 [FREE Full text] [doi: 10.1136/bmj.c332] [Medline: 20332509]
27. DellaPergola S. World jewish population, 2012. In: Dashefsky A, Sheskin I, editors. American Jewish Year Book 2012: Springer; 2012:109-112 URL: http://link.springer.com/chapter/10.1007/978-94-007-5204-7_6 [accessed 2025-08-13]
28. Sañudo B, González-Navarrete Á, Álvarez-Barbosa F, de Hoyo M, Del Pozo J, Rogers ME. Effect of flywheel resistance training on balance performance in older adults. A randomized controlled trial. J Sports Sci Med 2019 Jun;18(2):344-350. [Medline: 31191105]
29. Fernandez-Gonzalo R, Fernandez-Gonzalo S, Turon M, Prieto C, Tesch PA, García-Carreira MDC. Muscle, functional and cognitive adaptations after flywheel resistance training in stroke patients: a pilot randomized controlled trial. J Neuroeng Rehabil 2016 Apr 6;13(1):37. [doi: 10.1186/s12984-016-0144-7] [Medline: 27052303]
30. Naczk A, Huzarski T, Doś J, et al. Impact of inertial training on muscle strength and quality of life in breast cancer survivors. Int J Environ Res Public Health 2022 Mar 10;19(6):3278. [doi: 10.3390/ijerph19063278] [Medline: 35328964]
31. Bray NW, Smart RR, Jakobi JM, Jones GR. Exercise prescription to reverse frailty. Appl Physiol Nutr Metab 2016 Oct;41(10):1112-1116. [doi: 10.1139/apnm-2016-0226]
32. Illera-Domínguez V, Fernández-Valdés B, Gisbert-Orozco J, et al. Validity of a low-cost friction encoder for measuring velocity, force and power in flywheel exercise devices. Biol Sport 2023 Jul;40(3):805-811. [doi: 10.5114/biolsport.2023.119991] [Medline: 37398963]
33. Santamaría-Peláez M, González-Bernal JJ, Da Silva-González Á, et al. Validity and reliability of the short physical performance battery tool in institutionalized spanish older adults. Nurs Rep 2023 Sep 30;13(4):1354-1367. [doi: 10.3390/nursrep13040114] [Medline: 37873821]
34. Nightingale CJ, Mitchell SN, Butterfield SA. Validation of the timed up and go test for assessing balance variables in adults aged 65 and older. J Aging Phys Act 2019 Apr;27(2):230-233. [doi: 10.1123/japa.2018-0049]
35. Reid KF, Callahan DM, Carabello RJ, Phillips EM, Frontera WR, Fielding RA. Lower extremity power training in elderly subjects with mobility limitations: a randomized controlled trial. Aging Clin Exp Res 2008 Aug;20(4):337-343. [doi: 10.1007/BF03324865]
36. Floreani M, Rejc E, Gambin S, Vavassori L, Lazzer S. Effects of gravitational versus iso-inertial resistance training on leg muscle force and metabolic cost of walking in healthy older adults. J Sports Med Phys Fitness 2022 Jul 1;62(7):910-920. [doi: 10.23736/S0022-4707.21.12649-0]
37. Sañudo B, González-Navarrete Á, Álvarez-Barbosa F, Hoyo M, Del Pozo J, Rogers ME. Effect of flywheel resistance training on balance performance in older adults. a randomized controlled trial. J Sport Sci Med 2019;18(2):344-350.
38. Skelton DA. Explosive power and asymmetry in leg muscle function in frequent fallers and non-fallers aged over 65. Age Ageing 2002 Mar 1;31(2):119-125. [doi: 10.1093/ageing/31.2.119]
39. Fernandez-Gonzalo R, Lundberg TR, Alvarez-Alvarez L, de Paz JA. Muscle damage responses and adaptations to eccentric-overload resistance exercise in men and women. Eur J Appl Physiol 2014 May;114(5):1075-1084. [doi: 10.1007/s00421-014-2836-7]
40. Raya-González J, de Keijzer KL, Bishop C, Beato M. Effects of flywheel training on strength-related variables in female populations. A systematic review. Res Sports Med 2022;30(4):353-370. [doi: 10.1080/15438627.2020.1870977] [Medline: 33401963]
41. Pozzo M, Alkner B, Norrbrand L, Farina D, Tesch PA. Muscle‐fiber conduction velocity during concentric and eccentric actions on a flywheel exercise device. Muscle Nerve 2006 Aug;34(2):169-177. [doi: 10.1002/mus.20574]
42. Barry E, Galvin R, Keogh C, Horgan F, Fahey T. Is the timed up and go test a useful predictor of risk of falls in community dwelling older adults: A systematic review and meta-analysis. BMC Geriatr 2014 Feb 1;14:14. [doi: 10.1186/1471-2318-14-14] [Medline: 24484314]
43. Beauchet O, Fantino B, Allali G, Muir SW, Montero-Odasso M, Annweiler C. Timed up and go test and risk of falls in older adults: a systematic review. J Nutr Health Aging 2011 Dec;15(10):933-938. [doi: 10.1007/s12603-011-0062-0]

XSL•FO
RenderX

44. Podsiadlo D, Richardson S. The Timed “Up & Go”: A Test of Basic Functional Mobility for Frail Elderly Persons. J American Geriatrics Society 1991 Feb;39(2):142-148. [doi: 10.1111/j.1532-5415.1991.tb01616.x]
45. Jönhagen S, Halvorsen K, Benoit DL. Muscle activation and length changes during two lunge exercises: implications for rehabilitation. Scandinavian Med Sci Sports 2009 Aug;19(4):561-568. [doi: 10.1111/j.1600-0838.2007.00692.x]
46. Lindstedt SL, LaStayo PC, Reich TE. When active muscles lengthen: properties and consequences of eccentric contractions. Physiology (Bethesda) 2001 Dec;16(6):256-261. [doi: 10.1152/physiologyonline.2001.16.6.256]
47. Haruyama K, Kawakami M, Otsuka T. Effect of core stability training on trunk function, standing balance, and mobility in stroke patients. Neurorehabil Neural Repair 2017 Mar;31(3):240-249. [doi: 10.1177/1545968316675431]
48. Madruga-Parera M, Bishop C, Fort-Vanmeerhaeghe A, Beato M, Gonzalo-Skok O, Romero-Rodríguez D. Effects of 8 weeks of isoinertial vs. cable-resistance training on motor skills performance and interlimb asymmetries. J Strength Cond Res 2022;36(5):1200-1208. [doi: 10.1519/JSC.0000000000003594]
49. Lee DK, Kang MH, Lee TS, Oh JS. Relationships among the Y balance test, Berg Balance Scale, and lower limb strength in middle-aged and older females. Braz J Phys Ther 2015;19(3):227-234. [doi: 10.1590/bjpt-rbf.2014.0096] [Medline: 26039033]
50. Replication data for: “effectiveness of iso-inertial resistance training on eccentric and concentric power, physical performance, and risk of falls in physically active middle-older adults: a randomised controlled trial”. Zenodo. URL: https://doi.org/10.5281/zenodo.11125189 [accessed 2025-08-13]

Abbreviations:

CONSORT: Consolidated Standards of Reporting Trials
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Interventions and Online Telehealth
GR: gravitational
IN: iso-inertial
MEP: maximum explosive power
PA: physical activity
RPE: Borg Rating of Perceived Exertion
SPPB: Short Physical Performance Battery

Edited by M Antoniou; submitted 12.09.24; peer-reviewed by E luneau, L llurda; revised version received 03.06.25; accepted 03.06.25; published 21.08.25.

Please cite as:
Cadellans Arroniz A, Romero Rodríguez D, Zárate V, Dantony F, Madruga Parera M, Ortega Cebrian S, Blanco D
Effectiveness of Iso-Inertial Resistance Training on Muscle Power in Middle-Older Adults: Randomized Controlled Trial
JMIR Aging 2025;8:e66414
URL: https://aging.jmir.org/2025/1/e66414
doi:10.2196/66414

© Aïda Cadellans Arroniz, Daniel Romero Rodríguez, Víctor Zárate, Flora Dantony, Marc Madruga Parera, Silvia Ortega Cebrian, David Blanco. Originally published in JMIR Aging (https://aging.jmir.org), 21.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effectiveness of Walking Prescription Using Mobile Health Technology on the Changes in Daily Steps in Older Adults With Cognitive Impairment: Randomized Controlled Study

Hee Jung Kim[1], PhD; Yun Jung Hwang[2], PhD; Jee Eun Park[3,4,5], MD, PhD; Dong Young Lee[3,4,5,6], MD, PhD

[1]Biomedical Research Institute, Seoul National University Hospital, Seoul, Republic of Korea

[2]Department of Occupational Therapy, Yeoju University, Yeoju, Republic of Korea

[3]Department of Psychiatry, Seoul National University College of Medicine, 101, Daehak-ro, Jongno-gu, Seoul, Republic of Korea

[4]Department of Neuropsychiatry, Seoul National University Hospital, Seoul, Republic of Korea

[5]Medical Research Center, Institute of Human Behavioral Medicine, Seoul National University, Seoul, Republic of Korea

[6]Interdisiplinary Program in Cognitive Science, Seoul National University, Seoul, Republic of Korea

Corresponding Author:
Jee Eun Park, MD, PhD
Department of Psychiatry, Seoul National University College of Medicine, 101, Daehak-ro, Jongno-gu, Seoul, Republic of Korea

Abstract

Background: Walking is frequently recommended as a beneficial physical activity for older adults, as it can enhance both their physical and mental well-being and help prevent cognitive decline and dementia. While it is known that mobile health (mHealth) technology can help improve physical activity among older adults, there is limited research on its effectiveness for older individuals with cognitive impairment.

Objective: This study aimed to determine the effectiveness and feasibility of walking prescriptions using mHealth technology for older adults with cognitive impairment.

Methods: In total, 60 older adults (mean=76.1, SD 5.4) years; female, n=34) with mild cognitive impairment or mild dementia (n=28 and n=32, respectively; Mini-Mental State Examination [MMSE], mean=20.7, SD 4.0) who visited the memory clinic were enrolled. They were randomly assigned into three groups: (1) group A (n=20) was prescribed with a goal of daily steps based on their telemonitored activity using a smart band; (2) group B (n=19) only wore a smart band without a prescription; and (3) group C (n=21) took a monthly education to encourage their walking. All participants took monthly face-to-face sessions with a coach to check their performance and modify the goal of daily steps. Changes in daily steps (primary outcome), cognitive function, physical status, and depressive symptoms from baseline to post-intervention (12 weeks) and follow-up (24 weeks) were assessed by unblinded researchers. Linear mixed effect models with factors of group (reference: control), time (reference: baseline), and their interaction were used for data analysis. Post hoc analyses using paired *t* tests were also conducted.

Results: For group A, there was a significant group × time interaction effect on daily steps both at 12 and 24 weeks (β (SE)=2205.88 (672.34), *P*=.001; β (SE)=2194.63 (884.33), *P*=.015). Group B showed increased numbers of steps only at 12 weeks but not at 24 weeks. Group C showed a continuous decrease in daily steps during the study period. Regarding secondary outcomes, group C showed a significant decline in cognitive function measured by MMSE both at 12 and 24 weeks. However, groups A and B showed stationary MMSE scores during 24 weeks. The number of withdrawn participants did not differ among the 3 groups.

Conclusions: Our findings suggest that walking prescriptions using mHealth technology can effectively increase daily steps in older adults with cognitive impairment.

Trial Registration: Clinical Research Information Service (CRIS) of Republic of Korea; CRIS KCT0002610; https://cris.nih.go.kr/cris/search/detailSearch.do?seq=10195&search_page=L

(JMIR Aging 2025;8:e63081) doi:10.2196/63081

KEYWORDS

dementia; mild cognitive impairment; mobile health; physical activity; exercise

XSL•FO
RenderX

Introduction

Physical activity (PA) is protective against age-related cognitive decline and incidence of dementia in older adults [1,2]. It can reduce symptoms of depression and anxiety. It can also increase the psychological well-being of the aging population [3]. Even in older adults diagnosed with mild cognitive disorder (MCI) or dementia, moderate-to-vigorous intensity PA can effectively delay cognitive deterioration and prevent comorbid physical illnesses [4,5]. Therefore, clinical experts have advised older people to exercise regularly and continuously.

Unfortunately, the number of people doing exercise decreases as they get older due to their declined physical ability, lack of knowledge of age-appropriate exercise, false belief that exercise can cause injury, and psychosocial prejudice that older adults should behave carefully and calmly [6]. According to a national survey in Korea, only 37.6% of the population aged 65 years or older maintain PA at a recommended level of above moderate intensity for 150 minutes or more per week, and 46.3% do not exercise at all [7]. A US national survey has also found that only about 27% of individuals aged 65 years or older exercise for the recommended level for physical health [8]. According to a review paper summarizing studies from the United States, Europe, and Japan, 67% of older adults aged 60 years or older spend more than 8.5 hours sitting down [9].

It has been revealed that above moderate-intensity PA can protect or improve cognitive function, even in older adults with neurodegenerative brain diseases [10-13]. A recent study has analyzed the relationship between PA level, cognitive function, and brain pathology of older people using clinical data for 2 years before death and pathology data from brain autopsy and found that exercise and motor ability are associated with higher cognitive function independent of brain pathology, including Alzheimer's disease [14]. PA can have a counter-effect on degenerative brain diseases by improving brain plasticity and cognitive reserve. However, studies examining the effectiveness of PA programs for older adults with cognitive impairment have been mainly conducted on patients with advanced dementia in long-term care facilities due to limitations of clinic- or community-based programs. It is difficult to monitor PAs of older adults with cognitive impairment and lead them to plan and practice exercise continuously [11,13].

Recent advances in mobile health (mHealth) technology can be used to help individuals form exercise habits on their own or through expert assistance. Continuous monitoring and feedback of PA are possible because mHealth technology allows accurate activity measurements [15] and accumulation of activity information [16-18]. However, many studies on PA programs using mHealth technology have mainly been conducted on young people [19-22]. Although it has also been suggested that PA can be increased in the older population using mHealth technology [23], most studies have been conducted on older adults with physical diseases such as cardiovascular disease, diabetes, and obesity without cognitive dysfunction [23-25]. To the best of our knowledge, few studies have examined the effectiveness of PA programs using mHealth technology in older people with cognitive impairment.

Walking is the most commonly recommended PA. It is also suitable for older adults because it is easy to practice in daily life with few side effects. A relationship between maintaining recommended daily steps in older adults and low mortality has been reported [26]. While most previous studies have reported that moderate to vigorous PA is associated with improved cognitive function [10-13], a cohort study of 78,430 adults found that a higher number of steps (even if less than 10,000 steps per day) was associated with a lower risk of incident dementia [27]. Walking is not only directly influenced by physical functions such as aerobic capacity, balance ability, and limb strength but also by cognitive functions such as attention and executive function [28]. Regular walking can reflect an active lifestyle and self-directed health care. Mild cognitive impairment has been reported to be accompanied by a decrease in gait speed [29] and the decreased interest and motivation often seen in older adults with dementia leads to a decrease in PA [30]. Increasing steps in older adults with cognitive impairment may be a way to compensate for physical and mental frailty.

Clinical practice with prescribing and giving feedback to older patients to exercise at their physical level can lead them to set appropriate goals and practice PA [11,31-34]. While walking education alone has been reported to help increase step counts in older adults [32], it can be more helpful to directly prescribe a target step count based on assessing physical function for older patients with cognitive impairment. In addition, older patients with MCI or dementia may have difficulty remembering or reporting their daily steps; therefore, using mHealth technology to monitor daily step counts is useful to monitor their daily activity. However, simply monitoring daily activity can be insufficient to enhance daily steps; it requires setting appropriate goals, developing strategies to achieve them, evaluating outcomes, and refining the approach, which demands substantial cognitive resources. Consequently, human-led coaching, or personalized guidance, is essential for older adults with cognitive impairment [24]. In this study, mHealth technology is expected to function as a tool for clinicians to modify the daily activity of older patients with MCI or dementia. In other words, clinicians can improve their step counts more effectively if they use mHealth-based monitored data to make personalized prescriptions of steps.

Therefore, this study targeted older adults with cognitive impairment who visited a memory clinic to examine the effectiveness and feasibility of a program for changing daily steps by individualized walking prescription and monitoring using mHealth technology. A 3-arm design was used to compare the use of mHealth technology alone with the addition of human monitoring and feedback.

Methods

Participants

In total, 60 participants with mild cognitive impairment or mild dementia (clinical dementia rating [CDR] ≤1) were recruited from a memory clinic at a university hospital. A geriatric psychiatrist confirmed participants' cognitive diagnosis and evaluated their clinical status via CDR. Diagnoses of MCI and dementia were based on the comprehensive clinical assessments.

XSL•FO
RenderX

MCI was defined as individuals who met the core clinical criteria of MCI according to the recommendations of the National Institute on Aging and Alzheimer's Association guidelines [35]: (1) memory complaint corroborated by self, an informant, or clinician; (2) objective memory impairment for age, education, and gender; (3) largely intact functional activities; and (4) not demented. All MCI individuals had a global CDR score of 0.5. Participants diagnosed with mild dementia met the criteria for dementia in the Diagnostic and Statistical Manual of Mental Disorders 4th Edition (DSM-IV-TR [36]) and CDR score of 0.5 or 1.

Older patients performing no or low level of moderate PA (eg, fast walking for 30 min less than 3 times a week) were included, while those who had at least one of the following were excluded: (1) impossible to walk independently (cane users were not excluded); (2) diagnosed with an above-moderate degree of dementia (CDR ≥2); (3) having a physical illness that might affect their safety during the study period; (4) not being able to use a smart band by oneself or not having caregivers who could help them use it; and (5) participating in other exercise programs within the past year.

The sample size was based on a previous study reporting the effectiveness of the intervention using mHealth technology combined with SMS text messages to increase PA [20]. This study had a similar idea to ours, comparing 3 groups (mHealth-based monitoring with SMS text message vs mHealth-based monitoring without SMS text message vs control). The sample size was conservatively determined to yield a larger number of subjects after comparing 3 groups one by one. For achieving an 80% power ($1-\beta=.8$) at the 5% level of significance ($\alpha=.05$) with equal allocation, the sample sizes for intervention and active control are 10 and 10, respectively [37]. Then, we doubled the number of subjects to reflect the possibility of a small effect on steps from the intervention, given that the subjects in the reference study had twice as many baseline steps as our participants (9600 vs 4500 steps/day). The drop-out rate was assumed to be 40% since the participants of this study were older adults with cognitive impairment. Therefore, this resulted in 20 subjects in each arm, for a total of 60 subjects.

Ethical Considerations

This study was approved by the Institutional Review Board of Seoul National University Hospital (H-1708-118-879), and it was registered in a clinical trial registry (KCT0002610). All participants provided written informed consent. To protect privacy, the data used in this study were anonymized before analysis. Participants did not receive any compensation for their participation.

Study Design

Participants were randomly assigned into 2 intervention groups (group A, walking prescription and using a smart band; group B, only using a smart band) and a control group (group C) at a 1:1:1 ratio. For the randomization, an independent researcher who was not involved in patient evaluation or intervention created a random number table prior to registration of the subject. This study was not blinded even though 2 different researchers conducted baseline and follow-up assessments, respectively.

Intervention

Our 12-week protocol of intervention was designed for use in a hospital setting. Physicians, nurses, psychologists, physical or occupational therapists, and any other trained paramedic professionals can use this program. In this study, a clinical psychologist and occupational therapist alternatively checked each participant's performance and compliance with the program and provided educational sessions. Both conducted study procedures uniformly according to the protocol, which included detailed methods of assessment, consultation, and feedback. Prior to the start of the study, a mock coaching session was conducted to validate the interventionists' adherence to the protocol and feedback skills with the supervision of a physician (a geriatric psychiatrist). Walking education was provided to all participants, including the control group, during the study period. It was conducted in monthly face-to-face sessions, which involved individual guidance for correct posture, appropriate intensity, and a favorable environment for walking.

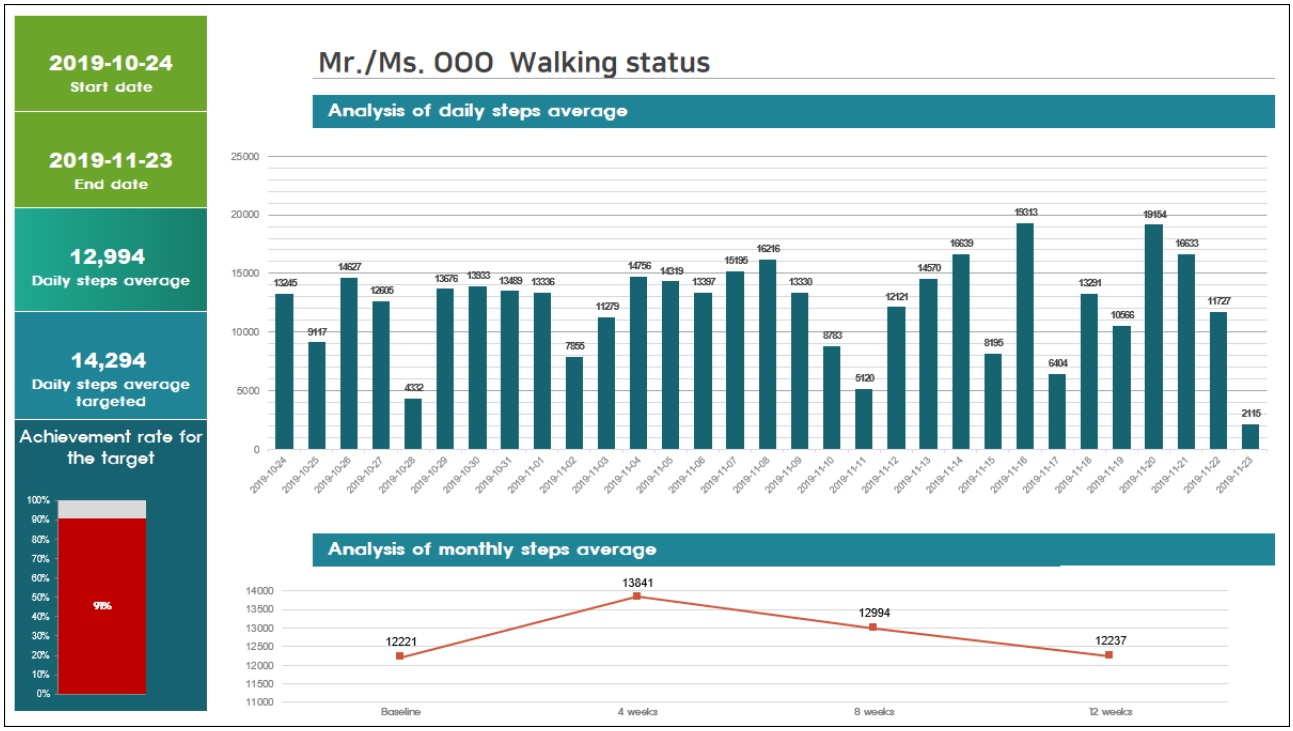
Participants in group A wore a smart band and received a personal prescription of daily steps based on their usual walking for the first week. Daily steps were monitored and timely feedback for encouragement was given via telephone or SMS text message at least once a week. If steps were not recorded for 5 consecutive days, the researcher made an emergency call, as this was the deadline to recover data in the event that there was a connection issue between devices. Monthly face-to-face feedback sessions were adjusted depending on individuals' daily step goals based on their accumulated step data. The target number of steps gradually increased if participants achieved the goal in the previous month, up to the recommended steps by age based on the guidelines for walking in older adults [38]. Participants also discussed with a researcher about their walking tasks, such as when, where, and how much they were walking. A solution was found together in case they had problems in performing tasks. A graph sheet based on telemonitored data was used for these sessions (Figure 1). Even after the 12-week intervention, they were encouraged to wear a smart band and keep walking.

Group B participants were also instructed to wear a smart band. However, they were not given a walking prescription. They were also encouraged to keep wearing the band after the intervention period. The control group (group C) received monthly walking education only and was provided with the same protocol as the intervention groups. A smart band was applied only at baseline, 12 weeks, and 24 weeks for assessment. App screenshots can be found in Multimedia Appendices 1-3.

Subjects who withdrew consent to participate in this study, failed the baseline assessment, missed more than 2 face-to-face visits, or did not obtain activity data for more than 2 weeks were withdrawn early.

XSL•FO
RenderX

Figure 1. A monthly walking sheet for group A subjects.



Outcome Assessment and Measures

At baseline, information on demographics, anthropometrics, and histories of physical and mental illnesses was collected. Primary and secondary outcomes were assessed at baseline, postintervention (12 weeks), and follow-up (24 weeks).

Primary Outcome

A change in daily steps was the primary outcome of this study. Smart bands (HR LS405-B6 & HR2.5 Gold edition; Seven Elec Co., LTD) were used to measure subjects' daily steps. Participants downloaded a mobile application that was linked to the wearable device. They could check their accumulated data on steps, which was transmitted automatically to the database. Average daily steps in a week were calculated at baseline, 12 weeks, and 24 weeks.

Secondary Outcomes

Secondary outcomes included three domains: physical function, cognitive function, and depressive symptoms. The Short Physical Performance Battery (SPPB) was used to evaluate physical function [39,40]. Cognitive function was measured using the Mini-Mental State Examination in the Korean version of the CERAD assessment packet (MMSE-KC) [41,42]. The Korean version of the Geriatric Depression Scale [43-45] was used for evaluating depressive symptoms in older patients.

Statistical Analyses

Analysis of variance (ANOVA) and χ^2 tests were used for continuous and categorical variables, respectively, to compare baseline characteristics among groups. A modified intent-to-treat approach was used to examine effectiveness of the intervention. The analysis included all randomly assigned participants with at least 1 post-baseline observation. Linear mixed-effect models with factors of group (reference: control), time (reference: baseline), and their interaction were used to examine group differences in changes of primary and secondary outcomes from baseline to 12- and 24-week follow-ups. To analyze the primary outcome, age, sex, education, cognitive function, and depressive symptoms at baseline were included in the model as covariates. In addition, we conducted a 2-way mixed ANOVA, where missing data of drop-out subjects were conservatively input as "no change" by carrying forward previous assessment values. For post hoc analyses, time-related changes within each group were examined using paired 2-tailed *t* test (within-group analyses). Whether different patterns were observed by participants' cognitive status (MCI and mild dementia) was also tested. All statistical analyses were performed with SPSS software (version 23.0; SPSS Inc.).

Results

In total, 60 older patients diagnosed with MCI or mild dementia were enrolled in this trial and randomly allocated into three groups. Among them, 13 participants dropped out throughout the study period. They did not attend post-baseline assessments. The number of withdrawn participants did not differ among the three groups. Drop-outs were caused by refusals, poor device operation, or health problems. Figure 2 shows the overall study flow, allocation, drop-out, and reasons for withdrawal. As one participant dropped out from group B soon after allocation, a person on the waitlist was allocated to group C with an independent randomization procedure. No adverse events were reported during the study period.

Table 1 shows baseline characteristics of participants according to group allocation. Participants had a mean (SD) age of 76.07 (5.43) years and 11.38 (5.10) years of education. Of them, 34 (57%) were females. Among the three groups, there were no significant differences in demographics or baseline measures

XSL•FO
RenderX

such as daily steps, physical status, cognitive function, or depressive symptoms.

A linear mixed-effect model showed a significant group-by-time interaction for the primary outcome measure with adjustment of age, sex, cognitive function, physical status, and depression (Table 2).

Figure 2. Flowchart for enrollment, allocation, and participation.

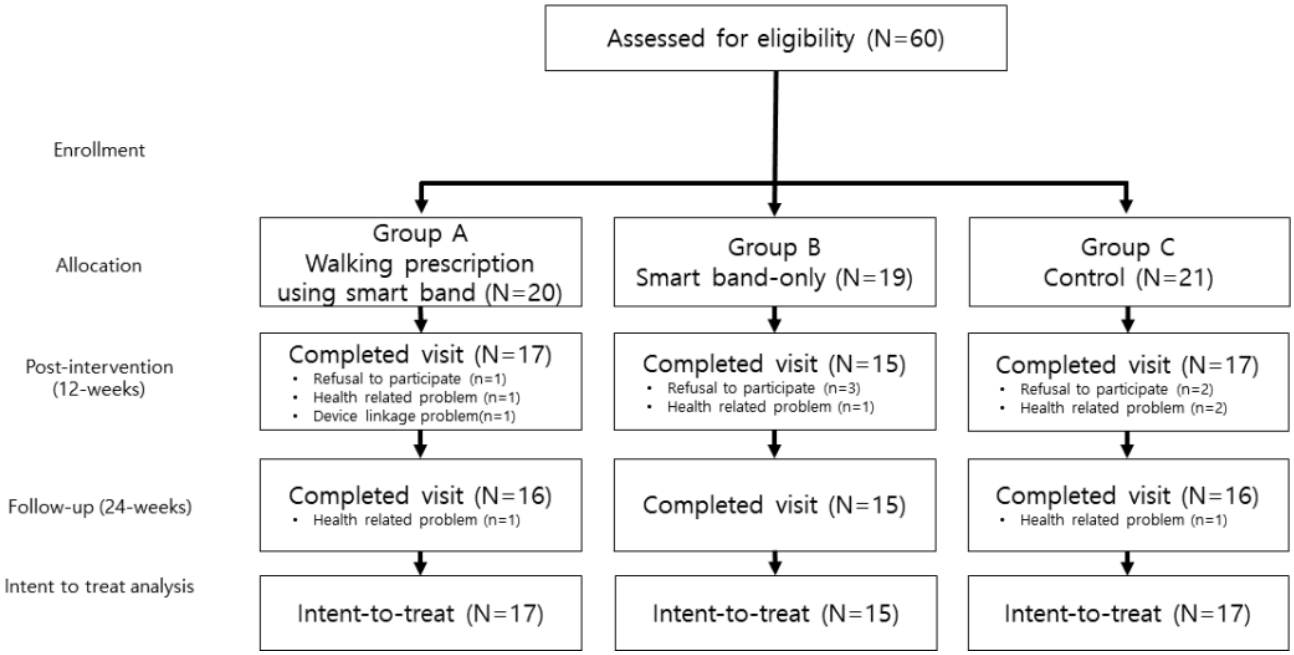


Table . Baseline characteristics of participants.

	Group A (n=20): walking prescription wearing smart band	Group B (n=19): wearing smart band only	Group C (n=21): control	*P* value
Demographics				
Age, years, mean (SD)	74.50 (7)	77.37 (4.15)	76.38 (4.57)	.31
Female, n (%)	12 (60)	11 (58)	11 (52)	.88
Education, years, mean (SD)	10.80 (4.83)	11.95 (5.75)	11.43 (4.95)	.79
Primary outcome				
Daily steps, n (SD)	4439.40 (3443.42)	4122.00 (1743.89)	5157.67 (2305.32)	.3
Secondary outcomes				
Physical status				
SPPB[a], scores (SD)	8.10 (1.80)	7.63 (1.92)	7.71 (1.95)	.93
Cognitive function				
MMSE-KC[b], scores (SD)	20.80 (3.44)	19.32(4.93)	21.71 (3.20)	.16
Depressive symptom				
GDS-K[c], scores (SD)	12.00 (6.00)	12.11 (7.59)	12.19 (5.86)	>.99

[a]SPPB, Short Physical Performance Battery.

[b]MMSE-KC, Mini-Mental State Examination in the Korean version of the CERAD assessment packet.

[c]GDS-K, Korean version of Geriatric Depression Scale.

XSL•FO
RenderX

Table . Linear mixed-effect models for differences in primary and secondary outcomes.

Intervention group[a]	Primary outcome	Secondary outcomes		
	Daily steps	Physical status	Cognitive function	Depressive symptom
	β (SE)[b]	SPPB[c]	MMSE-KC[d]	GDS-K[e]
			β (SE)[b]	
Constant	17,067.35 (4741.57)	13.29 (3.90)	3.70 (3.75)	1.53 (5.83)
Walking prescription with a smart band	−817.46 (853.10)	0.24 (0.66)	−0.30 (0.80)	−0.30 (1.35)
Smart band only	−97.67 (897.63)	−0.10 (0.70)	−0.35 (0.84)	0.02 (1.41)
Time[f]				
12 Weeks	−639.94 (475.42)	0.35 (0.29)	−1.76 (0.65)	−0.47 (1.20)
24 Weeks	−1541.50 (625.43)	0.16 (0.40)	−1.43 (0.76)	−2.29 (1.33)
Interaction of group by time				
12 Weeks—Smart band+Prescription	2205.88 (672.34)[g]	−0.12 (0.41)	1.29 (0.92)	0.29 (1.70)
24 Weeks—Smart band+Prescription	2194.63 (884.33)[h]	−0.15 (0.56)	0.88 (1.07)	1.79 (1.88)
12 Weeks—Smart band only	1715.21 (694.39)[h]	−0.02 (0.42)	2.36 (0.95)[h]	−0.06 (1.76)
24 Weeks—Smart band only	1088.30 (905.86)	−0.02 (0.57)	2.83 (1.10)[h]	3.96 (1.93)[h]

[a]Control as a reference group.

[b]Linear mixed model, including factors of group, time, and group x time interaction.

[c]SPPB, Short Physical Performance Battery.

[d]MMSE-KC, Mini-Mental State Examination in the Korean version of the CERAD assessment packet.

[e]GDS-K, Korean version of Geriatric Depression Scale.

[f]Baseline as a reference time.

[g]$P<.001$.

[h]$P<.05$.

Daily step numbers were significantly increased in group A (walking prescription wearing a smart band) at 12 weeks and 24 weeks, whereas group B (wearing a smart band only) showed an improvement only at 12 weeks, with the increase not sustained till the follow-up. Meanwhile, a continuous decrease in daily steps was observed in group C (control) during the study period. Differences of changes in daily steps by groups were confirmed with a 2-way mixed ANOVA, in which missing data were conservatively substituted with “no change” by carrying forward values. Analyses were adjusted for baseline daily steps (Figure 3; $F(3.38, 2{,}146{,}208.77)=3.66$, $P=.01$).

Post hoc within-group analyses repeatedly showed a differential pattern of daily step changes according to intervention (Table 3). Step numbers in group A showed significant increases at 12 and 24 weeks. Although fewer steps were observed at follow-up compared to post-intervention timing, daily steps after 24 weeks were still more than those at the baseline. Participants in group B tended to walk more after wearing a smart band at 12 weeks than at baseline, whereas their daily steps decreased at 24 weeks below their baseline numbers of steps. Finally, group C participants showed a continuous decrease in daily steps over 6 months.

XSL•FO
RenderX

Figure 3. Daily steps by groups using 2-way mixed analysis of variance adjusted with baseline steps. Two-way mixed model analyses of covariance, in which missing data of follow-up were conservatively substituted by the value measured immediately before, included factors of group, time, and group × time interaction, and adjusted for individuals' daily steps at baseline. Significant group-by-time interaction was shown in daily steps (F(3.38, 2,146,208.77)=3.66, *P*=.01). The raw data on daily steps (mean (SD)) without adjustment of individuals' daily steps at baseline are as follows: group A: 4553.18 (3598.33) at baseline, 6119.12 (4218.33) at 12 weeks and 5235.76 (3626.98) at 24 weeks; group B: 4449.07 (1708.51) at baseline, 5524.33 (2870.45) at 12 weeks, and 3995.87 (1629.55) at 24 weeks; group C: 5218.12 (2512.09) at baseline, 4578.18 (2425.97) at 12 weeks, and 3707.59 (1799.96) at 24 weeks.

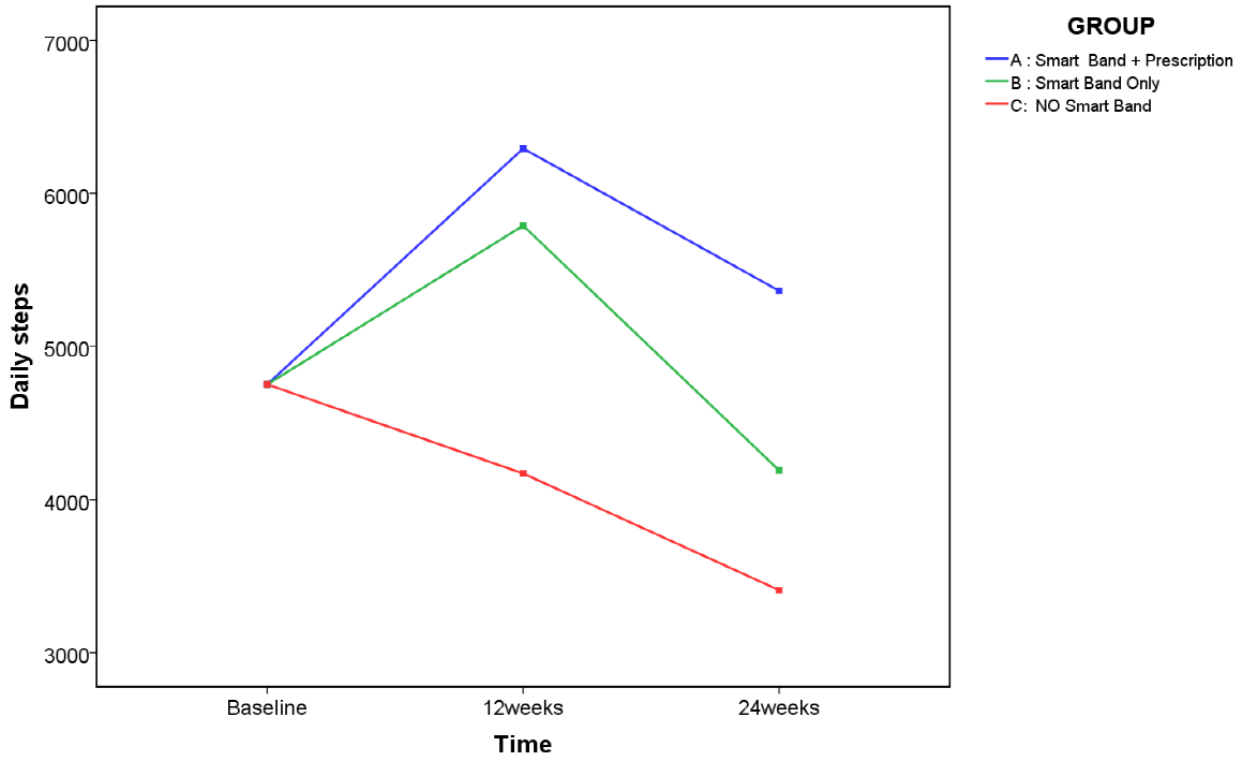


Table . Intra-group changes in daily steps and cognitive function from baseline to 12 weeks and 24 weeks follow-up.

	Group A (Smart band-based walking prescription, n=16)		Group B (Smart band only, n=15)		Group C (Control group, n=16)	
	Changes at 12 weeks Estimated mean difference (95% CI)[a]	Changes at 24 weeks Estimated mean difference (95% CI)[a]	Changes at 12 weeks Estimated mean difference (95% CI)[a]	Changes at 24 weeks Estimated mean difference (95% CI)[a]	Changes at 12 weeks Estimated mean difference (95% CI)[a]	Changes at 24 weeks Estimated mean difference (95% CI)[a]
Primary outcome						
Daily steps	1,549,000[a] (510.279 to 2587.721)	610.438 (−614.599 to 1835.474)	1075.267 (2400.059 to −249.525)	−453.200 (−1502.691 to 596.291)	−788.000 (−2089.440 to 513.444)	−1713.000[a] (−2761.966 to −664.034)
Secondary outcomes						
MMSE-KC[b]	−0.125 (−1.485 to 1.235)	−0.438 (−2.077 to 1.202)	0.600 (−1.136 to 2.336)	1.400 (−.627 to 3.427)	−1.875[a] (−2.991 to −.759)	−1.500[a] (−2.834 to 0.166)
GDS-K[c]	−.500 (−2.170 to 1.710)	−.750 (−2.841 to 1.341)	−.533 (−3.070 to 2.003)	1.667 (−1.123 to 4.457)	−.625 (−4.692 to 3.442)	−2.250 (−4.857 to 0.357)

[a]Significance at *P*<.008 with Bonferroni correction.

[b]MMSE-KC: Mini-Mental State Examination in the Korean version of the CERAD assessment packet.

[c]GDS-K: Geriatric Depression Scale-Korean.

XSL•FO
RenderX

Among secondary outcomes, cognitive function measured by MMSE-KC was changed differently by groups. A significant interaction between groups and time was shown in group B (Table 2). However, this result was not replicated in 2-way ANOVA in which missing data were conservatively substituted with their forward values ($F(4, 4.04)=1.73$, $P=.15$). In post hoc intra-group analyses, a significant decrease in MMSE-KC score was found in the control group at 12 weeks and 24 weeks. On the other hand, cognitive function in groups A and B was maintained at a level similar to the baseline (Table 3).

For depressive symptoms (Korean version of the Geriatric Depression Scale), a significant group-by-time interaction was revealed in group B at the follow-up point, although it was not confirmed in 2-way ANOVA ($F(4, 11.51)=1.62$, $P=.07$). In post hoc within-group analyses, a significant decrease in depressive symptoms was found in MCI individuals of group C at 24 weeks.

Discussion

Principal Findings and Comparison With Previous Works

This study examined the effectiveness and feasibility of mHealth-assisted walking prescriptions in older adults with cognitive impairment. It is the first study to demonstrate the effectiveness of a clinic-based PA program using mHealth technology, including people with dementia. Participants who were provided walking prescriptions and feedback based on data from smart bands (group A) showed increased steps in their daily lives both at postintervention (12 week) and follow-up (24 week) periods, whereas participants who were provided smart bands without personal prescriptions (group B) showed an increase in step numbers immediately after the intervention; however, such effect was not sustained at follow-up. Meanwhile, steps of the control group (group C) continued to decrease over the course of this study. As a secondary outcome measure, general cognitive function also showed a decreased pattern in the control group, while participants in groups A and B did not show decreased scores during 6 months, although they were diagnosed with MCI and mild dementia.

Previous studies have reported that walking can improve physical function, such as aerobic capacity [46], and reduce mortality in older adults with cognitive impairment [26]. However, a recent meta-analysis of randomized clinical trials reported that walking did not significantly improve cognitive function in older adults with MCI [46], whereas most previous studies demonstrated that moderate-to-vigorous PA improves cognitive function [10-13]. However, there are other benefits to maintaining adequate step counts for cognitively impaired older adults. Maintaining walking performance, such as pace, has been known to be associated with a lower risk of cognitive decline [47] and improved activities of daily living in older adults [48]. Regular walking can help maintain a robust rest-activity rhythm in cognitively impaired older adults who are vulnerable to circadian disruption, which is associated with the progression of dementia [49]. Since walking itself has cognitive components, including attention and executive function [28], it can stimulate cognitive function in MCI and dementia patients. Maintaining PA is challenging for older adults with cognitive impairment: guiding and monitoring high-intensity PA is nearly impossible in a memory clinic setting. Even though prescribing walking is not sufficient for direct effects on cognitive improvement, it can be a good alternative for maintaining PA for MCI and dementia patients.

A recent meta-analysis has demonstrated that mHealth intervention can increase PA, with such effects maintained for a long term [50]. However, most of the previous studies intervened with young adults. Older adults have many disadvantages when applying mHealth technology. They are less skilled at operating devices. They also have lower levels of digital literacy, making it difficult for them to adopt self-directed exercise programs using mHealth technology. Furthermore, it is almost impossible to have older adults with cognitive impairment, especially those diagnosed with dementia, stick with the exercise on their own. On the other hand, because they regularly visit clinics for treatment, it is easier for clinicians to implement programs in a direct manner. It is important to consider that older adults with cognitive impairment might have difficulty performing complex exercise tasks and accurately recalling their activities. Therefore, this study designed a program that could use mHealth technology to prescribe walking, a simple form of PA, and provide clinician-led monitoring and feedback.

To the best of our knowledge, only a few studies have tested the effectiveness of mHealth-assisted PA intervention in older adults with cognitive impairment. A previous study has confirmed the effects of a mHealth brisk walking intervention in increasing moderate-to-vigorous activity in older people with cognitive frailty [25]. Similar to our study, that study also prescribed and provided feedback on individualized goals. However, it primarily used automated monitoring and feedback generated by a smartphone application. Subjects in that study were recruited from the community. They had mild cognitive impairment with a baseline step count of >12,000 steps. Compared to subjects in our study, who were recruited from hospitals, including those with dementia and those who had a baseline step count of around 4000 steps, subjects in that study were cognitively and physically healthy. Therefore, a relatively self-directed intervention might show effectiveness. Group B subjects of our study were given a smart band and asked to perform self-directed walking. Similar to group A, group B also showed an increase in steps after 12 weeks. However, at the 24-week follow-up, steps tended to decrease back to baseline. That is, the effectiveness of mHealth-guided walking without constant monitoring and feedback was not sustained over a long term in older adults with cognitive impairment. As the previous study only tested the effect after 12 weeks, it was hard to know if the effect was sustained or not. However, the present study's results suggest that mHealth-supported PA is effective in increasing daily steps when it is accompanied by human-driven feedback, at least in older adults with dementia-level cognitive impairment.

Another previous study has also examined the feasibility and effectiveness of a memory clinic-based walking prescription for individuals with cognitive impairment [24]. In that study, participants did not show an increase in their step count after

XSL•FO
RenderX

applying the mHealth-supported walking prescription, similar to our approach. The main difference between that study and ours was the prescribing protocol. The previous study asked participants to double their steps in 6 weeks while our participants were asked to increase their steps by 500 biweekly. If such an increase was not achieved, they repeated the same goal. Our final steps were set at a realistic number based on age-appropriate guideline [38]. In addition, we checked participants' step counts daily and gave timely feedback at least once a week, whereas the previous study had a fixed biweekly feedback schedule. For older adults with cognitive impairment, the individualized coaching protocol might be more important than the use of mHealth technology itself (such as how to prescribe, monitor, and provide feedback on exercise). Realistic goals, frequent interventions, and flexible approaches might be essential to ensure the effectiveness of the mHealth walking program.

The study also suggests the potential for mHealth-supported walking prescriptions to help maintain cognitive function in older adults with cognitive impairment. Although this program did not improve participants' cognitive function, MMSE scores of intervention groups (groups A and B) were maintained during the study period, whereas the control group, MCI individuals in particular, showed a significant decrease in this global cognitive function score. A number of previous studies have reported that moderate-to-vigorous PA is an effective modality for maintaining cognitive function in patients with dementia and for improving cognitive plasticity in older adults with MCI [51-53]. Although the mHealth-assisted walking prescription in this study did not increase participants' cognitive scores, it has the possibility to prevent further deterioration of cognitive function in older adults diagnosed with MCI and early dementia, at least for some time. Given that subjects were already experiencing pathologic cognitive decline, a realistic goal of a clinic-based program might be to delay the rapid progression. At the same time, even though the sample size of this study was calculated to be more conservative than the previous study of similar design, it may have been insufficient to confirm the effectiveness on cognition. Future studies should include a larger number of subjects and different intensities of PA to test the program's effectiveness on cognitive function.

The mHealth-guided walking prescription in the present study was confirmed to be feasible to use in a clinic-based environment. There were no adverse events in our program. The number of withdrawn participants did not differ among the three study groups. Walking is the most accessible form of PA for older adults. It can be done anywhere and anytime. It is also easy to personalize and monitor individuals' goals [54]. Techniques we used in this program can be replaced by any application that includes an accelerator. This mHealth-supported walking may be one of the easiest and simplest ways to apply a PA program to older adults having a cognitive problem. Only one dropped out due to difficulty in using a mobile device, although all participants had cognitive impairment.

Strengths

We designed a 3-arm randomized controlled study. Therefore, we were able to analyze the effectiveness of the program from multiple perspectives: technology use alone versus technology use combined with personalized coaching. We also validated this mHealth-based program in a well-defined population with MCI and mild dementia through comprehensive clinical assessments. Lastly, our program and protocol were confirmed to be feasible in a clinical setting. Only 1 coaching person managed all three study groups' participants at a time. The coach monitored group A individuals' steps from the webpage, provided feedback via text or phone at least once a week, and conducted monthly in-person sessions with all participants. Training the coach on the structured protocol to deliver the program took about 4 hours, and the time required for the coach to manage the participants averaged 2-3 hours per day for 6 months. Assuming that there are 20-30 older adults with cognitive impairment over a 6-month period, the program could easily be replicated in other centers with 1 staff member.

Limitations

This study has several limitations. First, the numbers of participants were small. Although the effectiveness of the primary outcome (changes in step numbers) was confirmed, larger numbers of subjects and differentiated intensity of PA interventions might be needed to confirm the effect on cognitive function. Second, participants and researchers were not blinded to which group individuals were in. However, it was not possible to blind them as different protocols were applied by study groups. In addition, 2 different researchers conducted baseline and follow-up assessments, respectively. Third, the effectiveness of the program could have been different between individuals according to their baseline activity levels. Although we adjusted baseline steps in the analyses of group-level comparison, different effectiveness by individuals needs to be examined in the future. This approach will allow us to develop more personalized guides for PA in older adults regarding the intensity, duration, and interval of exercise. Fourth, the smart band used in this study had yet to be validated. Even though the accelerometer is a simple feature commonly found on smart bands, lack of validation is a limitation of this study. Fifth, apathy may affect motivation and adherence to a PA program. This study assessed depressive symptoms and found no difference in baseline depressive symptoms between participants who completed the program and those who dropped out. However, future research needs to measure apathy, which may directly affect motivation levels, and explore its impact on the effectiveness of the PA program. Lastly, impact of caregivers' support and assistance on the effectiveness of this program was not assessed. As some caregivers were essential in getting older adults to adhere to our program, it should be included as an important factor in future studies [55].

Conclusions

Our findings suggest that walking prescriptions using mHealth technology can effectively increase PAs and maintain cognitive health in older adults with cognitive impairment. It is also feasible to apply this mHealth-assisted program to older adults with MCI and mild dementia in the clinic setting. However, the effectiveness of our protocol needs to be confirmed with larger samples and more personalized methods in the future.

XSL•FO
RenderX

Acknowledgments

This study was supported by the Haesong Geriatric Psychiatry Research Fund in 2016.

Authors' Contributions

HJK was involved in data curation and writing – original draft preparation; YJH was involved in data curation; DYL was involved in conceptualization, methodology, writing – review & editing; JEP was involved in conceptualization, methodology, writing – original draft, review & editing, supervision, project administration, and funding acquisition.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Mobile app screenshot 1: app login page.
[PNG File, 299 KB - aging_v8i1e63081_app1.png]

Multimedia Appendix 2
Mobile app screenshot 2: daily activity including step counts.
[PNG File, 574 KB - aging_v8i1e63081_app2.png]

Multimedia Appendix 3
Mobile app screenshot 3: detailed page for daily steps.
[PNG File, 393 KB - aging_v8i1e63081_app3.png]

Checklist 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File, 8695 KB - aging_v8i1e63081_app4.pdf]

References

1. Di Marco LY, Marzo A, Muñoz-Ruiz M, et al. Modifiable lifestyle factors in dementia: a systematic review of longitudinal observational cohort studies. J Alzheimers Dis 2014;42(1):119-135. [doi: 10.3233/JAD-132225] [Medline: 24799342]
2. Laurin D, Verreault R, Lindsay J, MacPherson K, Rockwood K. Physical activity and risk of cognitive impairment and dementia in elderly persons. Arch Neurol 2001 Mar;58(3):498-504. [doi: 10.1001/archneur.58.3.498] [Medline: 11255456]
3. McAuley E, Rudolph D. Physical activity, aging, and psychological well-being. J Aging Phys Act 1995;3(1):67-96. [doi: 10.1123/japa.3.1.67]
4. Panza F, Lozupone M, Solfrizzi V, et al. Cognitive frailty: a potential target for secondary prevention of dementia. Expert Opin Drug Metab Toxicol 2017 Oct;13(10):1023-1027. [doi: 10.1080/17425255.2017.1372424] [Medline: 28849681]
5. Solfrizzi V, Scafato E, Seripa D, et al. Reversible cognitive frailty, dementia, and all-cause mortality. The Italian Longitudinal Study on Aging. J Am Med Dir Assoc 2017 Jan;18(1):89. [doi: 10.1016/j.jamda.2016.10.012] [Medline: 28012505]
6. Daley MJ, Spinks WL. Exercise, mobility and aging. Sports Med 2000 Jan;29(1):1-12. [doi: 10.2165/00007256-200029010-00001] [Medline: 10688279]
7. Korea Institute for Health and Social Affairs. The survey of Korean elderly [Report in Korean]. : Sejong-si: Korea Institute for Health and Social Affairs; 2020 URL: https://www.mohw.go.kr/board.es?mid=a10411010100&bid=0019&act=view&list_no=366496 [accessed 2025-06-03]
8. National Center for Health Statistics. Health, United States 2019 2021. [doi: 10.15620/cdc:100685]
9. Harvey JA, Chastin SFM, Skelton DA. Prevalence of sedentary behavior in older adults: a systematic review. Int J Environ Res Public Health 2013 Dec 2;10(12):6645-6661. [doi: 10.3390/ijerph10126645] [Medline: 24317382]
10. Heyn P, Abreu BC, Ottenbacher KJ. The effects of exercise training on elderly persons with cognitive impairment and dementia: a meta-analysis. Arch Phys Med Rehabil 2004 Oct;85(10):1694-1704. [doi: 10.1016/j.apmr.2004.03.019] [Medline: 15468033]
11. van Uffelen JGZ, Chinapaw MJM, Hopman-Rock M, van Mechelen W. Feasibility and effectiveness of a walking program for community-dwelling older adults with mild cognitive impairment. J Aging Phys Act 2009 Oct;17(4):398-415. [doi: 10.1123/japa.17.4.398] [Medline: 19940321]
12. Ströhle A, Schmidt DK, Schultz F, et al. Drug and exercise treatment of Alzheimer disease and mild cognitive impairment: a systematic review and meta-analysis of effects on cognition in randomized controlled trials. Am J Geriatr Psychiatry 2015 Dec;23(12):1234-1249. [doi: 10.1016/j.jagp.2015.07.007] [Medline: 26601726]

XSL•FO
RenderX

13. Liu IT, Lee WJ, Lin SY, Chang ST, Kao CL, Cheng YY. Therapeutic effects of exercise training on elderly patients with dementia: a randomized controlled trial. Arch Phys Med Rehabil 2020 May;101(5):762-769. [doi: 10.1016/j.apmr.2020.01.012] [Medline: 32084347]
14. Buchman AS, Yu L, Wilson RS, et al. Physical activity, common brain pathologies, and cognition in community-dwelling older adults. Neurology (ECronicon) 2019 Feb 19;92(8):e811-e822. [doi: 10.1212/WNL.0000000000006954] [Medline: 30651386]
15. Case MA, Burwick HA, Volpp KG, Patel MS. Accuracy of smartphone applications and wearable devices for tracking physical activity data. JAMA 2015 Feb 10;313(6):625-626. [doi: 10.1001/jama.2014.17841] [Medline: 25668268]
16. Steinhubl SR, Muse ED, Topol EJ. Can mobile health technologies transform health care? JAMA 2013 Dec 11;310(22):2395. [doi: 10.1001/jama.2013.281078]
17. Patel MS, Asch DA, Volpp KG. Wearable devices as facilitators, not drivers, of health behavior change. JAMA 2015 Feb 3;313(5):459-460. [doi: 10.1001/jama.2014.14781] [Medline: 25569175]
18. McCormack GR, Petersen J, Ghoneim D, Blackstaffe A, Naish C, Doyle-Baker PK. Effectiveness of an 8-week physical activity intervention involving wearable activity trackers and an ehealth app: mixed methods study. JMIR Form Res 2022 May 3;6(5):e37348. [doi: 10.2196/37348] [Medline: 35404832]
19. Burke LE, Ma J, Azar KMJ, et al. Current science on consumer use of mobile health for cardiovascular disease prevention: a scientific statement from the American Heart Association. Circulation 2015 Sep 22;132(12):1157-1213. [doi: 10.1161/CIR.0000000000000232] [Medline: 26271892]
20. Martin SS, Feldman DI, Blumenthal RS, et al. mActive: a randomized clinical trial of an automated mHealth intervention for physical activity promotion. J Am Heart Assoc 2015 Nov 9;4(11):26553211. [doi: 10.1161/JAHA.115.002239] [Medline: 26553211]
21. Walsh JC, Corbett T, Hogan M, Duggan J, McNamara A. An mHealth intervention using a smartphone app to increase walking behavior in young adults: a pilot study. JMIR Mhealth Uhealth 2016 Sep 22;4(3):e109. [doi: 10.2196/mhealth.5227] [Medline: 27658677]
22. Feter N, Dos Santos TS, Caputo EL, da Silva MC. What is the role of smartphones on physical activity promotion? A systematic review and meta-analysis. Int J Public Health 2019 Jun;64(5):679-690. [doi: 10.1007/s00038-019-01210-7] [Medline: 30758514]
23. Kwan RYC, Salihu D, Lee PH, et al. The effect of e-health interventions promoting physical activity in older people: a systematic review and meta-analysis. Eur Rev Aging Phys Act 2020;17(1):7. [doi: 10.1186/s11556-020-00239-5] [Medline: 32336996]
24. Vidoni ED, Watts AS, Burns JM, et al. Feasibility of a memory clinic-based physical activity prescription program. J Alzheimers Dis 2016 Apr 21;53(1):161-170. [doi: 10.3233/JAD-160158] [Medline: 27104905]
25. Kwan RY, Lee D, Lee PH, et al. Effects of an mHealth brisk walking intervention on increasing physical activity in older people with cognitive frailty: pilot randomized controlled trial. JMIR Mhealth Uhealth 2020 Jul 31;8(7):e16596. [doi: 10.2196/16596] [Medline: 32735218]
26. Lee IM, Shiroma EJ, Kamada M, Bassett DR, Matthews CE, Buring JE. Association of step volume and intensity with all-cause mortality in older women. JAMA Intern Med 2019 Aug 1;179(8):1105-1112. [doi: 10.1001/jamainternmed.2019.0899] [Medline: 31141585]
27. Del Pozo Cruz B, Ahmadi M, Naismith SL, Stamatakis E. Association of daily step count and intensity with incident dementia in 78 430 adults living in the UK. JAMA Neurol 2022 Oct 1;79(10):1059-1063. [doi: 10.1001/jamaneurol.2022.2672] [Medline: 36066874]
28. Snijders AH, van de Warrenburg BP, Giladi N, Bloem BR. Neurological gait disorders in elderly people: clinical approach and classification. Lancet Neurol 2007 Jan;6(1):63-74. [doi: 10.1016/S1474-4422(06)70678-0] [Medline: 17166803]
29. Peel NM, Alapatt LJ, Jones LV, Hubbard RE. The association between gait speed and cognitive status in community-dwelling older people: a systematic review and meta-analysis step count and intensity with incident dementia. J Gerontol A Biol Sci Med Sci 2019 May 16;74(6):943-948. [doi: 10.1093/gerona/gly140] [Medline: 29917045]
30. Stubbs B, Eggermont L, Soundy A, Probst M, Vandenbulcke M, Vancampfort D. What are the factors associated with physical activity (PA) participation in community dwelling adults with dementia? A systematic review of PA correlates. Arch Gerontol Geriatr 2014;59(2):195-203. [doi: 10.1016/j.archger.2014.06.006] [Medline: 25034708]
31. Petrella RJ, Koval JJ, Cunningham DA, Paterson DH. Can primary care doctors prescribe exercise to improve fitness? The STEP Test Exercise Prescription (STEP) project. Am J Prev Med 2003 May;24(4):316-322. [doi: 10.1016/s0749-3797(03)00022-9] [Medline: 12726869]
32. Kolt GS, Schofield GM, Kerse N, Garrett N, Ashton T, Patel A. Healthy steps trial: pedometer-based advice and physical activity for low-active older adults. Ann Fam Med 2012;10(3):206-212. [doi: 10.1370/afm.1345] [Medline: 22585884]
33. Northey JM, Cherbuin N, Pumpa KL, Smee DJ, Rattray B. Exercise interventions for cognitive function in adults older than 50: a systematic review with meta-analysis. Br J Sports Med 2018 Feb;52(3):154-160. [doi: 10.1136/bjsports-2016-096587] [Medline: 28438770]

XSL•FO
RenderX

34. Park JE, Jeon SY, Kim SA, et al. A multidomain intervention for modifying lifestyle habits reduces the dementia risk in community-dwelling older adults: a single-blinded randomized controlled pilot study. J Alzheimers Dis 2019;70(1):51-60. [doi: 10.3233/JAD-190016] [Medline: 31127782]
35. Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011 May;7(3):270-279. [doi: 10.1016/j.jalz.2011.03.008] [Medline: 21514249]
36. Association AP. Diagnostic and Statistical Manual of Mental Disorders:DSM-IV-TR: American Psychiatric Association; 2000. [doi: 10.1002/9780470479216.corpsy0271]
37. Chow SJ, Wang H, Lokhnygina Y. Size Calculations in Clinical Research, 3rd edition: New York: Chapman and Hall/CRC; 2017. [doi: 10.1201/9781315183084]
38. Korea Golden Age Forum. The walking guideline for the elderly people [Report in Korean]. : Seoul: Korea Golden Age Forum; 2010 URL: https://www.goldenageforum.org/ [accessed 2025-06-03]
39. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994 Mar;49(2):M85-M94. [doi: 10.1093/geronj/49.2.m85] [Medline: 8126356]
40. Han SH, Kim SM, Hong DJ, Choi BY, Kim MJ, Park SB. The short physical performance battery in some Korean rural community. J Korean Acad Rehabil Med 2009;33(2):177-185.
41. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
42. Lee JH, Lee KU, Lee DY, et al. Development of the Korean version of the Consortium to Establish a Registry for Alzheimer's Disease Assessment Packet (CERAD-K): clinical and neuropsychological assessment batteries. J Gerontol B Psychol Sci Soc Sci 2002 Jan;57(1):47-53. [doi: 10.1093/geronb/57.1.p47] [Medline: 11773223]
43. Yesavage JA, Brink TL, Rose TL, et al. Development and validation of a geriatric depression screening scale: a preliminary report. J Psychiatr Res 1982;17(1):37-49. [doi: 10.1016/0022-3956(82)90033-4] [Medline: 7183759]
44. Bae J. Diagnostic accuracy of geriatric depression scale (GDS) for DSM-III-R major depression [Master's thesis]. : Seoul National University; 1996.
45. Cho MJ, Bae JN, Suh GH, et al. Validation of geriatric depression scale, Korean version (GDS) in the assessment of DSM-III-R major depression. J Korean Neuropsychiatr Assoc 1999:48-63.
46. Lin JC, Chen IH, Cheng FY. Review articles (meta-analyses) effects of walking on cognitive function in individuals with mild cognitive impairment: a systematic review and meta-analysis. BMC Geriatr 2023 Aug 21;23(1):500. [doi: 10.1186/s12877-023-04235-z] [Medline: 37605156]
47. Quan M, Xun P, Chen C, et al. Walking pace and the risk of cognitive decline and dementia in elderly populations: a meta-analysis of prospective cohort studies. J Gerontol A Biol Sci Med Sci 2017 Feb;72(2):266-270. [doi: 10.1093/gerona/glw121] [Medline: 27927757]
48. Ungvari Z, Fazekas-Pongor V, Csiszar A, Kunutsor SK. The multifaceted benefits of walking for healthy aging: from Blue Zones to molecular mechanisms. Geroscience 2023 Dec;45(6):3211-3239. [doi: 10.1007/s11357-023-00873-8] [Medline: 37495893]
49. Lim ASP, Ellison BA, Wang JL, et al. Sleep is related to neuron numbers in the ventrolateral preoptic/intermediate nucleus in older adults with and without Alzheimer's disease. Brain (Bacau) 2014 Oct;137(Pt 10):2847-2861. [doi: 10.1093/brain/awu222] [Medline: 25142380]
50. Mönninghoff A, Kramer JN, Hess AJ, et al. Long-term effectiveness of mHealth physical activity interventions: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2021 Apr 30;23(4):e26699. [doi: 10.2196/26699] [Medline: 33811021]
51. Lautenschlager NT, Cox KL, Flicker L, et al. Effect of physical activity on cognitive function in older adults at risk for Alzheimer disease: a randomized trial. JAMA 2008 Sep 3;300(9):1027-1037. [doi: 10.1001/jama.300.9.1027] [Medline: 18768414]
52. Baker LD, Frank LL, Foster-Schubert K, et al. Effects of aerobic exercise on mild cognitive impairment: a controlled trial. Arch Neurol 2010 Jan;67(1):71-79. [doi: 10.1001/archneurol.2009.307] [Medline: 20065132]
53. Foster PP, Rosenblatt KP, Kuljiš RO. Exercise-induced cognitive plasticity, implications for mild cognitive impairment and Alzheimer's disease. Front Neurol 2011;2(28):28. [doi: 10.3389/fneur.2011.00028] [Medline: 21602910]
54. Bravata DM, Smith-Spangler C, Sundaram V, et al. Using pedometers to increase physical activity and improve health: a systematic review. JAMA 2007 Nov 21;298(19):2296-2304. [doi: 10.1001/jama.298.19.2296] [Medline: 18029834]
55. Bateman DR, Srinivas B, Emmett TW, et al. Categorizing health outcomes and efficacy of mHealth apps for persons with cognitive impairment: a systematic review. J Med Internet Res 2017 Aug 30;19(8):e301. [doi: 10.2196/jmir.7814] [Medline: 28855146]

Abbreviations

ANOVA: analysis of variance

XSL•FO
RenderX

CDR: clinical dementia rating
DSM-IV-TR: Diagnostic and Statistical Manual of Mental Disorders 4th Edition
MCI: mild cognitive disorder
mHealth: mobile health
MMSE-KC: Mini-Mental State Examination in the Korean version of the CERAD assessment packet
SPPB: Short Physical Performance Battery

Edited by R Yang; submitted 10.07.24; peer-reviewed by L Taylor, P Sanprakhon; revised version received 29.03.25; accepted 31.03.25; published 11.06.25.

Please cite as:
Kim HJ, Hwang YJ, Park JE, Lee DY
Effectiveness of Walking Prescription Using Mobile Health Technology on the Changes in Daily Steps in Older Adults With Cognitive Impairment: Randomized Controlled Study
JMIR Aging 2025;8:e63081
URL: https://aging.jmir.org/2025/1/e63081
doi:10.2196/63081

© Hee Jung Kim, Yun Jung Hwang, Jee Eun Park, Dong Young Lee. Originally published in JMIR Aging (https://aging.jmir.org), 11.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Association of a Healthy Lifestyle With All-Cause and Cause-Specific Mortality Among Individuals With Probable Sarcopenia: Population-Based Cohort Study

Ning Wang[1,2,3], MD, PhD; Yuqing Zhang[4,5], DSC; Junqing Xie[6], PhD; Na Lu[7], MPH; Aojie Zheng[1,2,3], MD; Changjun Li[2,8,9], MD, PhD; Jie Wei[1,2,3,8,10], PhD; Chao Zeng[1,2,3,8,10], MD, PhD; Guanghua Lei[1,2,3,8], MD, PhD; Yilun Wang[1,2,3], MD, PhD

[1]Department of Orthopaedics, Xiangya Hospital, Central South University, 87 Xiangya Road, Kaifu District, Changsha, China

[10]Department of Epidemiology and Health Statistics, Xiangya School of Public Health, Central South University, Changsha, China

[2]Key Laboratory of Aging-related Bone and Joint Diseases Prevention and Treatment, Ministry of Education, Xiangya Hospital, Central South University, Changsha, China

[3]Hunan Key Laboratory of Joint Degeneration and Injury, Xiangya Hospital, Central South University, Changsha, China

[4]Division of Rheumatology, Allergy, and Immunology, Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, United States

[5]The Mongan Institute, Massachusetts General Hospital, Harvard Medical School, Boston, MA, United States

[6]Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, United Kingdom

[7]Arthritis Research Canada, Richmond, BC, Canada

[8]National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, China

[9]Department of Endocrinology, Endocrinology Research Center, Xiangya Hospital of Central South University, Changsha, China

Corresponding Author:
Yilun Wang, MD, PhD
Department of Orthopaedics, Xiangya Hospital, Central South University, 87 Xiangya Road, Kaifu District, Changsha, China

Abstract

Background: Individuals with probable sarcopenia have shown excess mortality, yet no specific treatment regimen has been established. While lifestyle factors improve health and longevity in general populations, their role in probable patients with sarcopenia remains unclear due to differing lifestyle patterns. Clarifying this could inform strategies to address this unmet need.

Objective: We aim to quantify the impact of a healthy lifestyle on all-cause and cause-specific mortality in probable sarcopenic populations using a large-scale prospective cohort study.

Methods: Participants were selected from the UK Biobank, aged 40 - 69 years, during 2006 - 2010. Probable sarcopenia was identified according to EWGSOP2 (European Working Group on Sarcopenia in Older People 2) criteria, resulting in 20,654 participants being included in this study. Death dates and underlying causes were obtained from the National Health Service Information Center. Cox proportional hazard models and population-attributable risk were used to assess the associations between healthy lifestyle factors and premature mortality risk.

Results: A total of 20,654 individuals with probable sarcopenia were included in this study. The median age of the population was 62.0 (IQR 56.0-66.0) years, and 60.6% (n=12,528) were women. During a median follow-up duration of 11.5 (IQR 10.8-12.3) years, 2447 participants died. All healthy lifestyle factors, including nonsmoking (P<.001), moderate alcohol intake (P<.001), regular physical activity (P<.001), a healthy diet (P=.01), limited television-watching time (P<.001), adequate sleep duration (P=.001), and strong social connections (P<.001), were independently associated with lower mortality risk. To evaluate the cumulative associations between modifiable lifestyle factors and mortality outcomes (all-cause and cause-specific) among patients with probable sarcopenia, we developed a healthy lifestyle index. Participants were assigned one point per adherence to each optimal lifestyle factor. Compared with individuals with 0 - 2 healthy lifestyle scores, hazard ratios of all-cause mortality for those with 3 to 6 - 7 factors were 0.67 (95% CI 0.59 - 0.76), 0.51 (95% CI 0.45 - 0.57), 0.43 (95% CI 0.38 - 0.49), and 0.33 (95% CI 0.29 - 0.39), respectively (P for trend <.001). There was also a dose-response relationship between the number of healthy lifestyle factors and mortality from cancer, cardiovascular disease, respiratory disease, digestive disease, and other causes (all P for trend<.001). Population-attributable risk analysis indicated that 25.7% (95% CI 22% - 29%) of deaths were attributable to a poor lifestyle (scoring 0 - 5).

XSL•FO
RenderX

Conclusions: A healthy lifestyle is associated with a lower risk of all-cause mortality and mortality due to cancer, cardiovascular disease, respiratory disease, and digestive disease among individuals with probable sarcopenia. Adopting a healthy lifestyle (scoring 6‐7) could prevent 25.7% of deaths in this population.

(JMIR Aging 2025;8:e65374) doi:10.2196/65374

KEYWORDS

healthy lifestyle; sarcopenia; mortality; cohort study; muscle strength

Introduction

Sarcopenia, a progressive and generalized skeletal muscle disorder, affects approximately 10%‐27% of people older than 60 years of age worldwide [1]. The sarcopenia diagnosis follows a three-stage hierarchy per the 2019 EWGSOP2 (European Working Group on Sarcopenia in Older People 2) guidelines: probable sarcopenia is defined by low muscle strength (primary criterion, measured via handgrip dynamometry); confirmed sarcopenia requires concurrent low muscle mass; and severe sarcopenia is diagnosed with additional functional limitation [2]. Even patients in the mildest stage, probable sarcopenia, exhibit a nearly 1.6 times higher risk of mortality compared to the general population [3-6]. Given the high prevalence and significant adverse consequences of sarcopenia, it is crucial to identify potentially modifiable factors to lower the premature mortality risk among this population.

The EWGSOP2 guideline underscores the critical role of lifestyle management in sarcopenia [7], given that physical activity and dietary interventions have been shown to retard sarcopenia progression and reduce mortality risk in this population [8]. However, current evidence is limited to a narrow range of lifestyle factors, notably neglecting other critical determinants, especially emerging ones such as sleep quality, sedentary behavior, and social engagement [9-11]. These lifestyle factors have already been proven to promote health in the general population by reducing inflammation and comorbid burdens [12-14], which are also relevant to sarcopenia [15]. However, whether and to what extent individuals with sarcopenia can benefit from such multidimensional lifestyle factors remains unclear due to their distinct behavioral patterns; for example, individuals with sarcopenia usually reduce their regular physical activity and tend to be more socially isolated [16,17]. Additionally, there is still a lack of evidence on the combined effects of lifestyle factors among individuals with sarcopenia, despite their strong interrelation [18]. Addressing this evidence gap is critical to inform evidence-based guidelines and population-level strategies for behavior change.

To bridge the knowledge gap, we analyzed data from the UK Biobank study to examine the relation of 7 healthy lifestyle factors, that is, no current smoking, moderate alcohol consumption, healthy diet pattern, regular physical activity, adequate sleep duration, short television watching time, and appropriate social connection to all-cause mortality. We also estimated the proportion of deaths that can be prevented theoretically by simultaneously adopting several healthy lifestyle factors among individuals with sarcopenia.

Methods

Study Population

The UK Biobank study is a population-based cohort that recruited more than 500,000 participants aged 40‐69 years during 2006‐2010 [19]. Participants provided detailed health-related data and lifestyle information, had physical measurements taken, and provided biological samples at 22 assessment centers across England, Scotland, and Wales [20].

Handgrip strength was measured using a Jamar J00105 hydraulic dynamometer by a trained nurse in a standardized clinical setting [3]. Participants were seated upright with forearms supported on armrests, and bilateral measurements were obtained via a single 3-second maximal voluntary contraction of each hand. The average of right- and left-hand measurements, expressed in kilograms, was used in this study. Sarcopenia was diagnosed according to the EWGSOP2 criteria [2]. Due to the scarcity of confirmed sarcopenia cases in the UK Biobank [21], participants with probable sarcopenia, defined as handgrip strength <27 kg among men and <16 kg among women, were finally included in this study.

Ethical Considerations

All participants provided written informed consent before the data collection. The study was approved by the National Information Governance Board for Health and Social Care and the National Health Service North West Multicenter Research Ethics Committee (16/NW/0274; ethics approval for UK Biobank studies). This study was conducted with permission (UKB application 77,646) from the UK Biobank. The UK Biobank has made necessary efforts to safeguard participant privacy and provided appropriate compensation to the participanting subjects; all patients'data have been anonymized and deidentified.

Assessment of Lifestyle Factors

The impact of seven modifiable healthy lifestyles was evaluated in the current analysis. These included four well-established factors (ie, smoking status, alcohol intake, physical activity, and diet) [22] and three emerging factors (television watching time, sleep duration, and social connection) [12,18,23]. All information on lifestyle factors was self-reported and assessed using a touchscreen questionnaire at baseline. Detailed definitions of each lifestyle factor are shown in Multimedia Appendix 1.

Each lifestyle variable was coded 0 or 1, with 1 indicating the adoption of a particularly healthy lifestyle factor based on the following criteria: no current smoking (never or previous smoker), moderate alcohol consumption (women ≤1 drink per

XSL•FO
RenderX

day, and men ≤2 drinks per day regularly), healthy diet pattern (≥5 recommended food component groups), regular physical activity (≥150 min per week of moderate activity, or ≥75 min per week of vigorous activity), adequate sleep duration (7 - 8 h per day) [24], short television watching time (<4 h per day) [25], and appropriate social connection (ie, moderately active and active social connection status) [26]. An overall healthy lifestyle score was constructed as the sum of the scores of 7 lifestyle factors, with a higher score indicating higher adoption of an overall healthy lifestyle. For the avoidance of extreme groups with limited cases, we collapsed scores of 0, 1, and 2 into one category and 6 and 7 into another category. The categorization choices are arbitrary but justifiable to avoid floor and ceiling effects.

Ascertainment of Outcomes

The outcome variables consisted of all-cause mortality and cause-specific mortality, that is, death from cancer, cardiovascular disease (CVD), respiratory disease, neurodegenerative disease, digestive disease, and other causes. Information about the death date and underlying cause was obtained from the National Health Service Information Center (for England and Wales) and the National Health Service Central Register Scotland (for Scotland). The cause of death was defined based on the International Classification of Diseases, 10th Revision code. Detailed information about the linkage procedure is web-accessible [27].

Assessment of Covariates

Covariates included age (years); sex (women or men); ethnicity (White, Black, Asian, Mixed or other); education (higher, ie, college or university degree or other professional qualification; upper secondary, ie, second or final stage of secondary education; lower secondary, ie, first stage of secondary education; vocational, ie, work-related practical qualifications; or other); social deprivation (Townsend deprivation index); and employment (currently employed or not).

Statistical Analysis

Baseline characteristics are presented as the mean (SD) if normally distributed or median (IQR) if nonnormally distributed for continuous variables and frequency (%) for categorical variables. The distribution of quantitative variables was evaluated using both graphical assessments (ie, histograms and Q-Q plots) and formal normality tests (ie, skewness and kurtosis metrics).

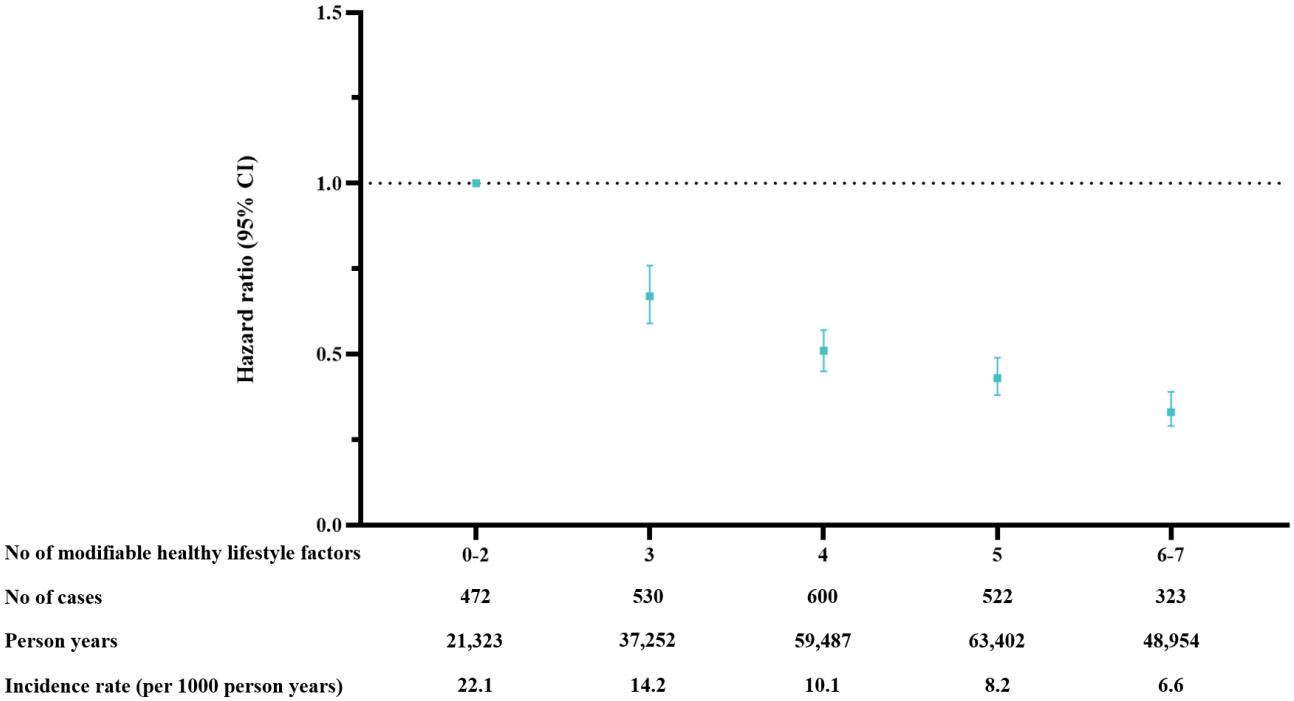
For each participant, we calculated person-years of follow-up from the date of recruitment (between 2006 and 2010) to the date of death or the end of follow-up (December 2020 for England and Wales and November 2020 for Scotland and elsewhere), whichever occurred first. Cox proportional hazard models were applied to examine the associations of each lifestyle factor and the overall lifestyle score with all-cause and cause-specific mortality risk. The results were reported as hazard ratios (HRs) with 95% CIs. We assigned a median value to each lifestyle score category to test the linear trend. Kaplan-Meier survival analysis and curve were also applied to reveal the association between lifestyle score and the cumulative incidence of death (Figure 1). We adjusted for age, sex, education, Townsend deprivation index, and employment to control potential confounders. Besides, lifestyle factors were mutually adjusted for when the relation of each lifestyle factor to all-cause mortality was examined. In addition, we performed sex-stratified analyses to evaluate the relation of healthy lifestyle scores to all-cause mortality in men and women. We took the same approach to examine the relation of healthy lifestyle scores to all-cause mortality. The proportional hazards assumption was tested using a Schoenfeld residuals plot. The population-attributable risk (PAR), an estimate of the proportion of all-cause deaths that would theoretically have been prevented if all individuals had a lifestyle score of at least 3, 4, 5, or >6, was calculated, with the 0 - 2 category as the reference group.

We conducted 2 sensitivity analyses to evaluate the robustness of our primary findings. First, we performed multiple imputations with chained equations with 5 datasets to deal with the missing values of exposure or covariates (Table S1 in Multimedia Appendix 1). Then, we excluded participants who died within 2 or 4 years of follow-up to reduce immortal bias (Table S2 in Multimedia Appendix 1).

P values were 2-sided with statistical significance set at less than .05. All analyses were performed using Stata SE (version 15; StataCorp).

XSL•FO
RenderX

Figure 1. Hazard ratios of all-cause mortality according to combined modifiable healthy lifestyle factors.



No of modifiable healthy lifestyle factors	0-2	3	4	5	6-7
No of cases	472	530	600	522	323
Person years	21,323	37,252	59,487	63,402	48,954
Incidence rate (per 1000 person years)	22.1	14.2	10.1	8.2	6.6

Results

Population Characteristics

The baseline characteristics of 20,654 individuals with probable sarcopenia are shown in Table 1. The mean age of the population was 59.7 (SD 7.20) years, and 39.8% (n=8126) were men. The proportion of healthy lifestyle scores of 0‐2, 3, 4, 5, and 6‐7 was 9.7% (n=2015), 16.4% (n=3382), 25.7% (n=5316), 27.2% (n=5610), and 21% (n=4331), respectively. Participants with a lower lifestyle score (0‐2) were younger (mean 58.85, SD 7.15 vs 60.43, SD 7.11 y), less educated (20.8% (n=419) vs 51.5% (n=2232) with higher education level), less employed (19.1% (n=385) vs 42.3% (1833) employment rate), and more socioeconomically deprived (Townsend score: 1.27 vs −1.43), compared to high scorers (6-7). Conversely, higher lifestyle scorers (6-7) exhibited a lower prevalence of CVD [8.8% (n=383) vs 19.6% (n=394])], hypertension [33.7% (n=1460) vs 49.4% (n=996)], diabetes [7.5% (n=325) vs 15.4% (n=311)], hyperlipidemia [24.5% (n=1,061) vs 37.3% (n=752)], and depression [5.7% (n=245) vs 18.1% (n=364); Table 1].

XSL•FO
RenderX

Table . Baseline characteristics of participants with probable sarcopenia by lifestyle score category[a].

Variables	Lifestyle score				
	0 - 2 (n=2015)	3 (n=3382)	4 (n=5316)	5 (n=5610)	6 - 7 (n=4331)
Age, mean (SD)	58.85 (7.15)	59.70 (7.20)	59.84 (7.17)	60.12 (7.10)	60.43 (7.11)
Male, n (%)	874 (43.4)	1281 (37.9)	2046 (38.5)	2123 (37.8)	1802 (41.6)
Not current smoking, n (%)	1122 (55.7)	2816 (83.3)	4866 (91.5)	5393 (96.1)	4298 (99.2)
Moderate alcohol consumption, n (%)	216 (10.7)	570 (16.9)	1439 (27.1)	2003 (35.7)	2729 (63)
Diet >=5 recommended components, n (%)	246 (12.2)	1064 (31.5)	2587 (48.7)	3748 (66.8)	3884 (89.7)
Regular physical activity, n (%)	196 (9.7)	700 (20.7)	2008 (37.8)	3427 (61.1)	3772 (87.1)
Adequate sleep, n (%)	353 (17.5)	1211 (35.8)	2808 (52.8)	4039 (72)	3903 (90.1)
Never or short television watching time, n (%)	309 (15.3)	1129 (33.4)	2835 (53.3)	4099 (73.1)	3955 (91.3)
Not isolated social connection, n (%)	1034 (51.3)	2656 (78.5)	4721 (88.8)	5341 (95.2)	4264 (98.5)
Ethnicity, n (%)					
Asian	74 (3.7)	194 (5.7)	358 (6.7)	352 (6.3)	270 (6.2)
Black	42 (2.1)	54 (1.6)	76 (1.4)	75 (1.3)	42 (1)
Mixed	16 (0.8)	30 (0.9)	23 (0.4)	34 (0.6)	17 (0.4)
Other	39 (1.9)	36 (1.1)	77 (1.4)	80 (1.4)	47 (1.1)
White	1844 (91.5)	3068 (90.7)	4782 (90)	5069 (90.4)	3955 (91.3)
Education category, n (%)					
Higher education	419 (20.8)	930 (27.5)	1876 (35.3)	2367 (42.2)	2232 (51.5)
Upper secondary	110 (5.5)	188 (5.6)	353 (6.6)	405 (7.2)	276 (6.4)
Lower secondary	385 (19.1)	749 (22.1)	1222 (23)	1247 (22.2)	892 (20.6)
Vocational	131 (6.5)	231 (6.8)	316 (5.9)	307 (5.5)	214 (4.9)
Other or prefer not to answer	970 (48.1)	1284 (38)	1549 (29.1)	1284 (22.9)	717 (16.6)
Employment, n (%)	385 (19.1)	926 (27.4)	1918 (36.1)	2253 (40.2)	1833 (42.3)
Townsend, mean (SD)	1.27 (3.68)	0.01 (3.46)	−0.57 (3.30)	−1.17 (3.08)	−1.43 (2.90)
Conditions, n (%)					
Cancer	255 (12.7)	400 (11.8)	639 (12)	622 (11.1)	486 (11.2)
Cardiovascular disease	394 (19.6)	568 (16.8)	692 (13)	601 (10.7)	383 (8.8)
Hypertension	996 (49.4)	1520 (44.9)	2228 (41.9)	2100 (37.4)	1460 (33.7)
Diabetes mellitus	311 (15.4)	439 (13)	632 (11.9)	527 (9.4)	325 (7.5)
Hyperlipidemia	752 (37.3)	1112 (32.9)	1604 (30.2)	1529 (27.3)	1061 (24.5)
Neurodegenerative disease	24 (1.2)	32 (0.9)	45 (0.8)	25 (0.4)	18 (0.4)
Respiratory disease	251 (12.5)	226 (6.7)	226 (4.3)	200 (3.6)	104 (2.4)
Digestive disease	54 (2.7)	54 (1.6)	55 (1)	41 (0.7)	24 (0.6)
Depression	364 (18.1)	423 (12.5)	490 (9.2)	386 (6.9)	245 (5.7)

[a]Data are mean (SD) for continuous variables or n (%) for categorical variables.

XSL•FO
RenderX

Association of Individual Lifestyle Factors With the Risk of All-Cause Mortality

During 230,418 person-years of follow-up, 2447 participants died. As shown in Table 2, each lifestyle factor was significantly associated with all-cause mortality. Compared with no current smoking participants, the adjusted HR of mortality was 1.98 (95% CI 1.78 - 2.20) for those current smokers. Abstaining from smoking appears to be the most effective in reducing the risk of mortality. There was a U-shaped association between alcohol consumption and all-cause mortality, with moderate alcohol consumption (male ≤16 g per day; female ≤8 g per day regularly) being associated with lower mortality; other alcohol consumption levels (never or excessive) had a higher mortality risk (HR 1.28, 95% CI 1.17 - 1.39). Engaging in irregular physical activity and following an unhealthy diet pattern were associated with a higher mortality rate; the corresponding HRs were 1.36 (95% CI 1.25 - 1.47) and 1.11 (95% CI 1.02 - 1.20). Adoption of 3 emerging healthy lifestyle factors was also associated with lower mortality risk. Short (≤6 h per day) or long duration (≥9 h per day) of sleep versus adequate sleeping duration (7 - 8 h per day) was associated with a higher risk of mortality (HR 1.14, 95% CI 1.05 - 1.24), especially the long sleep duration group had a significantly higher mortality rate (HR 1.39, 95% CI 1.25 - 1.55; Table S3 in Multimedia Appendix 1). Participants with longer television-watching time (≥4 h per day) had higher mortality than those with short television-watching participants (HR 1.20, 95% CI 1.10 - 1.31), and those who had an isolated social connection could also significantly increase the risk of all-cause mortality (HR 1.22, 95% CI 1.09 - 1.35; Table 2). Besides, no apparent effect modification was observed by sex (Tables S4 and S5 in Multimedia Appendix 1).

XSL•FO
RenderX

Table . HR[a] (95% CI) of all-cause mortality according to individual modifiable healthy lifestyle factors.

Modifiable healthy lifestyle factors[b]	Participants, n	Cases, n	Person-years	Follow-up time, mean (SD)	Incidence rate per 1000 person-year (95% CI)	HR (95% CI)[c]	*P* value
Smoking							<.001
No current	18,495	1966	207526.68	11.22 (1.91)	9.47 (9.06 - 9.90)	1.00 (reference)	
Current	2159	481	22890.80	10.60 (2.75)	21.01 (19.18 - 22.98)	1.98 (1.78 - 2.20)	
Alcohol consumption							<.001
Moderate	6957	785	77110.68	11.08 (1.92)	10.18 (9.48 - 10.92)	1.00 (reference)	
Never or excessive	13,697	1662	153306.80	11.19 (2.07)	10.84 (10.33 - 11.38)	1.28 (1.17 - 1.39)	
Diet							.01
>=5 recommended components	11,529	1205	129681.51	11.25 (1.83)	9.29 (8.77 - 9.83)	1.00 (reference)	
<5 recommended components	9125	1242	100735.97	11.04 (2.24)	12.33 (11.65 - 13.03)	1.11 (1.02 - 1.20)	
Physical activity							<.001
Regular	10,103	978	113630.74	11.25 (1.80)	8.61 (8.08 - 9.16)	1.00 (reference)	
Irregular	10,551	1469	116786.74	11.07 (2.21)	12.58 (11.94 - 13.24)	1.36 (1.25 - 1.47)	
Sleep							.001
Adequate	12,314	1307	138003.41	11.21 (1.92)	9.47 (8.96 - 10.00)	1.00 (reference)	
Short or long	8340	1140	92414.07	11.08 (2.17)	12.34 (11.63 - 13.07)	1.14 (1.05 - 1.24)	
Television watching time							<.001
Short	12,327	1161	138817.19	11.26 (1.81)	8.36 (7.89 - 8.86)	1.00 (reference)	
Long	8327	1286	91600.29	11.00 (2.29)	14.04 (13.28 - 14.83)	1.20 (1.10 - 1.31)	
Social connection							<.001
Appropriate	18,016	1994	201699.16	11.20 (1.96)	9.89 (9.46 - 10.33)	1.00 (reference)	
Isolated	2638	453	28718.32	10.89 (2.41)	15.77 (14.35 - 17.30)	1.22 (1.09 - 1.35)	

[a]HR: hazard ratio.

[b]Low-risk lifestyle factors: no current smoking, moderate alcohol consumption (must be drinking but no more than 1 drink per day for women and 2 drinks per day for men on a relatively regular frequency, no drinking is risk factor), healthy diet (adequate intake of at least one-half of 10 recommended food groups), regular physical activity (150 min per week of moderate activity or 75 min per week of vigorous activity, or an equivalent combination), adequate sleep duration (7-8 h per day), short television watching time (<4 h per day), and appropriate social connection (not isolated).

[c]Adjustment for age (years), sex (women or men), ethnicity (White, Black, Asian, Mixed, or other), education (higher, ie, college or university degree or other professional qualification; upper secondary, ie, second or final stage of secondary education; lower secondary, ie, first stage of secondary education; vocational, ie, work-related practical qualifications; or other), Townsend deprivation index, and employment (currently employed or not). Lifestyle factors were mutually adjusted for analyses on the association of each individual lifestyle factor with all-cause mortality risk.

Association of Lifestyle Score With Risk of All-Cause Mortality

The healthy lifestyle score was inversely associated with all-cause mortality (Table 3). For each one-point increase, the HR of all-cause mortality risk was 0.78 (95% CI 0.76 - 0.81; *P* for trend <.001). Compared with those scoring 0 - 2, the multivariable-adjusted HRs of mortality for participants with healthy lifestyle scores of 3, 4, 5, and 6 - 7 were 0.67 (95% CI 0.59 - 0.76), 0.51 (95% CI 0.45 - 0.57), 0.43 (95% CI

XSL•FO
RenderX

0.38 - 0.49), and 0.33 (95% CI 0.29 - 0.39), respectively (*P* for trend <.001; Figure 2). The PAR for all-cause mortality increased as the healthy lifestyle score increased, from 3.1% for a score of ≥2% to 9.7%, 18.3%, 25.7%, and 34.6% for scores of ≥3, ≥4, ≥5, and ≥6, respectively (Table 4).

XSL•FO
RenderX

Table . HR[a] (95% CI) of all-cause or cause-specific mortality according to combined modifiable healthy lifestyle factors.

Variables	Number of modifiable healthy lifestyle factors[b]					*P* value	HR of each point increase (95% CI)
	0 - 2 (n=2015)	3 (n=3382)	4 (n=5316)	5 (n=5610)	6 - 7 (n=4331)		
All-cause mortality						<.001	0.78 (0.76 - 0.81)
Number of cases	472	530	600	522	323		
Person-years	21,323	37,252	59,487	63,402	48,954		
HR (95% CI)	1.00 (reference)	0.68 (0.60 - 0.77)	0.51 (0.45 - 0.57)	0.43 (0.37 - 0.49)	0.33 (0.28 - 0.38)		
Cancer						<.001	0.82 (0.78 - 0.86)
Number of cases	172	218	245	207	151		
Person-years	21,323	37,252	59,487	63,402	48,954		
HR (95% CI)	1.00 (reference)	0.76 (0.62 - 0.94)	0.57 (0.47 - 0.70)	0.46 (0.37 - 0.57)	0.43 (0.34 - 0.55)		
Cardiovascular disease						<.001	0.83 (0.77 - 0.88)
Number of cases	100	111	142	114	66		
Person-years	21,323	37,252	59,487	63,402	48,954		
HR (95% CI)	1.00 (reference)	0.71 (0.54 - 0.94)	0.65 (0.50 - 0.85)	0.53 (0.40 - 0.71)	0.39 (0.28 - 0.55)		
Respiratory disease						<.001	0.73 (0.66 - 0.80)
Number of cases	74	62	48	47	31		
Person-years	21,323	37,252	59,487	63,402	48,954		
HR (95% CI)	1.00 (reference)	0.57 (0.41 - 0.81)	0.31 (0.21 - 0.45)	0.31 (0.21 - 0.45)	0.26 (0.17 - 0.41)		
Neurodegenerative disease						.08	0.89 (0.95 - 1.02)
Number of cases	18	25	37	44	32		
Person-years	21,323	37,252	59,487	63,402	48,954		
HR (95% CI)	1.00 (reference)	0.70 (0.38 - 1.29)	0.61 (0.35 - 1.08)	0.63 (0.36 - 1.10)	0.55 (0.30 - 0.98)		
Digestive disease						<.001	0.66 (0.58 - 0.75)
Number of cases	43	37	28	27	8		
Person-years	21,323	37,252	59,487	63,402	48,954		
HR (95% CI)	1.00 (reference)	0.60 (0.39 - 0.94)	0.33 (0.20 - 0.53)	0.32 (0.19 - 0.52)	0.12 (0.06 - 0.27)		
Other[c]						<.001	0.81 (0.75 - 0.88)
Number of cases	65	77	100	83	35		
Person-years	21,323	37,252	59,487	63,402	48,954		

XSL•FO
RenderX

Variables	Number of modifiable healthy lifestyle factors[b]					*P* value	HR of each point increase (95% CI)
	0 - 2 (n=2015)	3 (n=3382)	4 (n=5316)	5 (n=5610)	6 - 7 (n=4331)		
HR (95% CI)	1.00 (reference)	0.79 (0.56 - 1.10)	0.71 (0.51 - 0.98)	0.59 (0.41 - 0.84)	0.33 (0.21 - 0.51)		

[a]HR: hazard ratio.

[b]Adjustment for age (years), sex (women or men), ethnicity (White, Black, Asian, Mixed or other), education (higher, ie, college or university degree or other professional qualification; upper secondary, ie, second or final stage of secondary education; lower secondary, ie, first stage of secondary education; vocational, ie, work-related practical qualifications; or other), Townsend deprivation index, and employment (currently employed or not); adjusting for competing risk of death of other causes.

[c]Mortality from causes other than cancer, CVD, respiratory disease, neurodegenerativee disease, or digestive disease.

Figure 2. Kaplan-Meier survival curve of all-cause mortality according to combined modifiable healthy lifestyle factors.

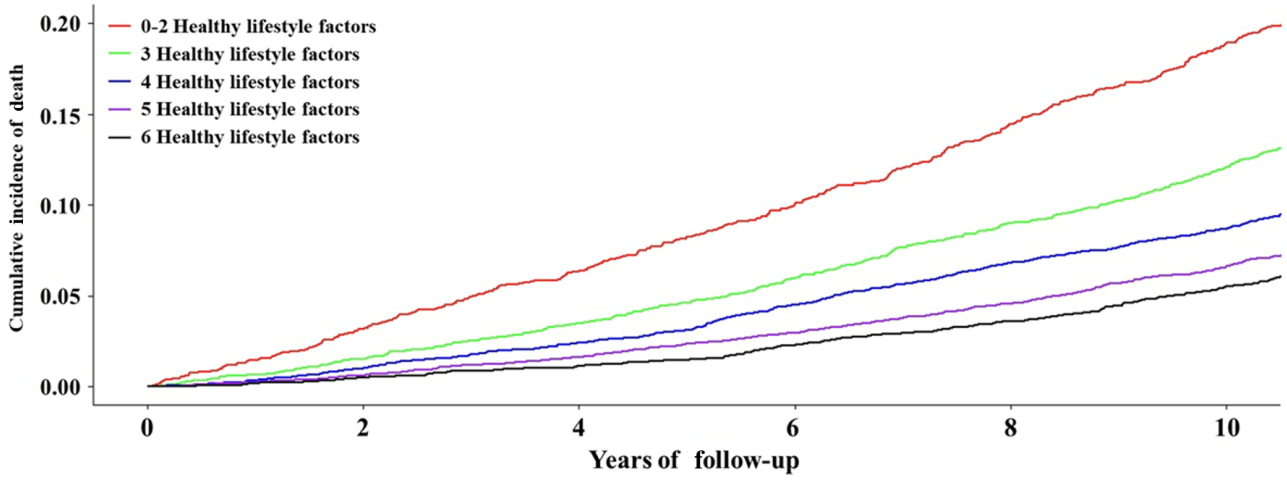


Table . Relative and population-attributable risks of all-cause mortality for groups defined by combinations of modifiable healthy lifestyle factors.

Modifiable healthy lifestyle factors[a]	Number of cases	% of total population	Hazard ratio (95% CI)[b]	PAR%[c] (95% CI)
At least 2 vs <2	2309	94.4	0.46 (0.38 - 0.55)	3.1 (2.6 - 3.5)
At least 3 vs <3	1975	80.7	0.50 (0.45 - 0.55)	9.7 (8.6 - 10.7)
At least 4 vs <4	1445	59.1	0.55 (0.51 - 0.60)	18.3 (16.3 - 20.2)
At least 5 vs <5	845	34.5	0.61 (0.56 - 0.66)	25.7 (22.0 - 29.0)
At least 6 vs <6	323	13.2	0.60 (0.53 - 0.68)	34.6 (27.9 - 40.5)

[a]Low-risk lifestyle factors: no current smoking, moderate alcohol consumption (no more than 1 drink per day for women and 2 drinks per day for men on a relatively regular frequency), healthy diet (adequate intake of at least one-half of 10 recommended food groups), regular physical activity (150 min per week of moderate activity or 75 min per week of vigorous activity, or an equivalent combination), adequate sleep duration (7-8 h per day), short television watching time (<4 h per day), and appropriate social connection (not isolated).

[b]Adjustment for age (years), sex (women or men), ethnicity (White, Black, Asian, Mixed, or other), education (higher, ie, college or university degree or other professional qualification; upper secondary, ie, second or final stage of secondary education; lower secondary, ie, first stage of secondary education; vocational, ie, work-related practical qualifications; or other), Townsend deprivation index, and employment (currently employed or not).

[c]PAR%: population attributable risk percent.

HRs and their 95% CIs were calculated in Cox proportional hazard models after adjusting for age (years), sex (women or men), ethnicity (White, Black, Asian, Mixed or other), education (higher, ie, college or university degree or other professional qualification; upper secondary, ie, second or final stage of secondary education; lower secondary, ie, first stage of secondary education; vocational, ie, work-related practical qualifications; or other), Townsend deprivation index, and employment (currently employed or not).

Association of Lifestyle Score With Risk of Cause-Specific Mortality

During the follow-up, 993 participants died from cancer, 533 participants died from CVD, 262 participants died from respiratory disease, 156 participants died from neurodegenerative disease, 143 participants died from digestive disease, and 360 participants died from other causes. There was a significantly inverse dose-response association between the health lifestyle score and the mortality from cancer, CVD, respiratory disease, digestive disease, and other causes (*P* for

XSL•FO
RenderX

trend <.001); however, no such pattern was observed for neurodegenerative disease (Table 3). Participants with higher overall lifestyle scores (6-7) had significantly lower mortality rates compared to those with scores of 0‐2. The multivariate-adjusted HRs of cause-specific mortality were 0.43 (95% CI 0.34‐0.55) for cancer, 0.39 (95% CI 0.28‐0.55) for CVD, 0.26 (95% CI 0.17‐0.41) for respiratory disease, 0.12 (95% CI 0.06‐0.27) for digestive disease, and 0.33 (95% CI 0.21‐0.51) for other causes.

Discussion

Statement of Principal Findings

In this cohort study conducted among individuals with probable sarcopenia, we found that adopting a healthy lifestyle was associated with significantly reduced all-cause mortality, notably deaths from cancer, CVD, respiratory disease, and digestive disease. We estimated that approximately 25.7% of deaths could be prevented if they could adopt six or more healthy lifestyle factors. These findings underscore the pivotal role of embracing a healthy lifestyle in alleviating mortality among individuals with probable sarcopenia.

Comparison With Previous Studies

Physical activity and nutritional interventions have already been recommended as primary strategies for managing sarcopenia, but previous evidence was inconclusive, with limitations such as small sample sizes and inconsistent adherence to interventions [28,29]. This study provides further evidence for the effectiveness of these two interventions among individuals with probable sarcopenia. Meanwhile, this study highlights divergent associations between lifestyle factors and mortality in probable sarcopenia populations versus the general population. Among individuals with probable sarcopenia, current smoking and excessive alcohol consumption exert the strongest mortality impacts (Table S3 in Multimedia Appendix 1). By contrast, in the general population, smoking cessation remains the most impactful lifestyle determinant on mortality [30], and Locquet et al [31] also discovered that smokers have a 2.36-fold higher likelihood of developing sarcopenia. Regular physical activity conferred the second-greatest survival benefit in the general population [30], which was consistent with prior reports by Li et al [32]; while dietary quality (HR 0.88, 95% CI 0.82‐0.94 in men; HR 0.97, 95% CI 0.90‐1.05 in women) and excessive alcohol intake (HR 0.95, 95% CI 0.89‐1.01 in men; HR 1.03, 95% CI 0.94‐1.12 in women) had weaker associations with mortality [30]. Notably, a meta-analysis also indicated that alcohol consumption is not associated with the presence of sarcopenia in the general population, with a pooled odds ratio of 0.964 [33]. Overall, these results indicate that it is meaningful to explore the impact of these healthy lifestyle factors on the mortality risk within the specific population of probable sarcopenia.

Three emerging healthy lifestyle factors also show a significant association with reduced risk of mortality. Social connection has garnered significant attention since the onset of the COVID-19 pandemic [34]. Previous studies have reported that individuals with sarcopenia are more prone to experiencing social isolation [17]. This study, for the first time, provides evidence for the beneficial effects of social connection on the probable sarcopenia population. Additionally, a recent meta-analysis showed that both short and long sleep durations are associated with an increased risk of all-cause mortality and cardiovascular events in the general population [35]. In this study, however, only long sleep duration was significantly associated with an increased mortality risk. There are two possible reasons: (1) individuals with probable sarcopenia who spend less time on sleep may allocate more time for physical activities, the interconnectedness among different lifestyle factors thereby potentially influencing the underlying association between lifestyle factors and mortality risk; and (2) sleep quality, such as chronotype, insomnia, snoring, and daytime sleepiness, can also affect the health outcomes [36].

To our knowledge, no study has examined the association between the combined healthy lifestyle score and mortality risk among individuals with sarcopenia. We have used Cox proportional hazard models and PAR to uncover whether and to what extent individuals with probable sarcopenia can benefit from a healthy lifestyle. These findings indicate that, beyond physical activity and diet, other healthy lifestyle factors also confer benefits to the probable sarcopenia population. This knowledge can contribute to the development of “core set outcomes” for lifestyle-based intervention trials in the field of sarcopenia. Additionally, a higher healthy lifestyle score is also associated with a reduced risk of mortality due to cancer, CVD, respiratory disease, and digestive disease, although not for neurodegenerative disease. Previous studies have demonstrated that adopting a healthy lifestyle can lower the risk of developing neurodegenerative diseases [37]. Nevertheless, since neurodegenerative diseases are typically not immediately life-threatening, their specific mortality risk might be attenuated.

Potential Mechanisms

Biological mechanisms linking various lifestyle factors to health outcomes have been postulated. Physical activity can impact systemic immune and metabolic response, stimulating adenosine monophosphate-activated protein kinase phosphorylation or other critical signaling pathways to promote multiple organ health and increase survival probability [38-40]. A balanced diet can provide sufficient protein, amino acids, and other essential nutrients [29]. Additionally, it can also help maintain intestinal microflora homeostasis and influence the aging gut, extending life span through metabolites produced by gut microbiota, such as short-chain fatty acids [41].

Cigarettes contain many toxic and carcinogenic components, which can impair mitochondrial function, increase oxidative stress [42], and cause epigenetic changes [43]. Therefore, avoiding tobacco exposure is an effective measure to lower disease and mortality risks. Individuals who frequently consume moderate amounts of alcohol show a lower mortality risk than those who drink alcohol rarely or excessively. A possible explanation for this phenomenon may lie in the fact that individuals who engage in regular moderate drinking tend to have better financial situations and social connections. Mechanistically, moderate alcohol consumption can lower the activity of a mental stress–related brain network [44] and reduce the risk of major cardiovascular events [45]. Besides, excessive

XSL•FO
RenderX

drinking leads to inflammation and immune-metabolic dysregulation, gut leak and dysbiosis, mitochondrial dysfunction, and epigenomic modifications [46-48]. These mechanisms synergistically interact to cause alcohol-mediated multiorgan injury and even death.

Sedentary behavior is directly associated with elevated levels of circulating inflammatory markers, such as interleukin-6 and C-reactive protein [49,50]. These markers have been proposed as senescent biomarkers due to their positive correlation with age and potential promotion of adverse health outcomes in older adults [51]. Growing evidence indicates that social connection influences various biological pathways, including blood pressure [52], oxidative stress [53], neuroendocrine dysregulation [54], chronic inflammation [55], and gut-microbiome interactions [56]. Similarly, social connections may indirectly influence health outcomes via stress perception and other behavioral factors, such as sleep quality and quantity, smoking, or even drug abuse [57,58]. These social stress-related behaviors are strongly associated with biological health.

Sarcopenia is associated with the aforementioned pathologic mechanisms that drive increased mortality risk and are influenced by lifestyle [59,60]. Individuals with sarcopenia often have multiple comorbidities (eg, diabetes, kidney disease, and chronic obstructive pulmonary disease) [61,62], and evidence indicates that healthy lifestyles significantly aid in managing these conditions [63,64], thereby reducing mortality risk.

Clinical and Research Implications

First, our findings provide incentive and support for the notion that lifestyle intervention strategies can have an impressive impact on mortality risk among individuals with probable sarcopenia, and this could help health professionals and policy makers to make preventive advice and policy recommendations for this population. Second, this study presented precise numerical values to which individuals with probable sarcopenia can benefit from a healthy lifestyle. Such knowledge might help motivate individuals with probable sarcopenia to change habits and adhere to a healthy lifestyle. Third, only 21% of the participants adhered to a healthy lifestyle (6‐7 healthy factors) in this study. In comparison, 51.9% adopted less than or equal to four healthy lifestyle factors, suggesting a large gap between current and ideal population lifestyles. Last but not least, this study found that certain lifestyle factors are more relevant than others; therefore, public health policies could focus on a few more potent risk factors (ie, smoking) rather than on costly strategies addressing multiple risk factors.

Strengths and Limitations

This study has several strengths. Foremost among these is the large sample size and comprehensive data resources, which encompassed detailed information on potential confounding variables. This robust dataset facilitated thorough analyses of mortality across various causes and allowed for stratification based on potential risk factors [22]. Additionally, this study pioneered the assessment of the combined impact of an overall healthy lifestyle on the probable sarcopenia population. This approach was particularly important due to the strong intercorrelations among various lifestyle factors [65]. Using an overall healthy lifestyle score enabled a comprehensive evaluation of the intricate relationships between lifestyle factors and mortality among individuals with probable sarcopenia.

However, several limitations should be acknowledged. First, the study's participant pool primarily consisted of Caucasians in the United Kingdom. Consequently, the generalizability of our findings to other ethnic groups may be limited. Second, due to the nature of the observational study, we cannot derive causality between lifestyle modification and mortality in this population, which warrants more well-conducted interventional studies to verify. Third, the temporal relationship between lifestyle factors and probable sarcopenia could not be clearly demarcated in this study. Thus, potential collider bias may arise when conditioning on probable sarcopenia status in analyses of lifestyle factors and mortality. This residual collider bias could attenuate associations toward the null [66], yet our identification of strong relationships strengthens confidence in these conclusions. Given this bias, estimating potentially preventable deaths via postprobable sarcopenia lifestyle changes would better clarify the importance of lifestyle management. However, UK Biobank's limited longitudinal data on postdiagnosis lifestyle modifications precluded formal analysis of how such changes affect mortality [19]. Fourth, further research can be conducted on confirmed patients with sarcopenia, although investigating lifestyles' impact on the risk of premature mortality among the probable sarcopenia population holds greater preventive implications. Fourth, measurement errors could occur in the self-reported lifestyle data. Fifth, those who died during the study period might have had serious diseases at baseline. Although the study excluded deaths within the first 2‐4 years of follow-up, the possibility of reverse causation and residual confounding remains. Finally, even though various covariates were adjusted for in our analyses, other confounders, such as BMI and comorbidities, may not have been included, which could result in residual confounding.

Conclusions

A healthy lifestyle was associated with a lower risk of all-cause mortality and mortality due to cancer, CVD, respiratory disease, and digestive disease among individuals with probable sarcopenia. A healthy lifestyle (scoring 6‐7) could prevent 25.7% of deaths in this population.

Acknowledgments

This work was supported by the National Key Research and Development Plan (2022YFC3601900 and 2022YFC2505500), the National Natural Science Foundation of China (81930071, 82072502, 82202769, 82302771, and U21A20352), the China Postdoctoral Science Foundation funded project (2023M733959), the Natural Science Foundation of Hunan Province (2023JJ40993 and 2022JJ40821), and the Fundamental Research Funds for the Central Universities of Central South University (2025ZZTS0867).

XSL•FO
RenderX

Authors' Contributions

YW designed the concept. NW drafted the manuscript. NL and JW did the statistical analysis. CZ, YW, NW, and GL obtained funding. NL, JW, and YZ gave administrative, technical, or material support. All authors read, gave critical feedback on intellectual content, interpreted data, revised the manuscript, and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplemental methods and results.
[DOCX File, 64 KB - aging_v8i1e65374_app1.docx]

References

1. Petermann-Rocha F, Balntzi V, Gray SR, et al. Global prevalence of sarcopenia and severe sarcopenia: a systematic review and meta-analysis. J Cachexia Sarcopenia Muscle 2022 Feb;13(1):86-99. [doi: 10.1002/jcsm.12783] [Medline: 34816624]
2. Cruz-Jentoft AJ, Bahat G, Bauer J, et al. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing 2019 Jan 1;48(1):16-31. [doi: 10.1093/ageing/afy169] [Medline: 30312372]
3. Celis-Morales CA, Welsh P, Lyall DM, et al. Associations of grip strength with cardiovascular, respiratory, and cancer outcomes and all cause mortality: prospective cohort study of half a million UK Biobank participants. BMJ 2018 May 8;361:k1651. [doi: 10.1136/bmj.k1651] [Medline: 29739772]
4. López-Bueno R, Andersen LL, Calatayud J, et al. Associations of handgrip strength with all-cause and cancer mortality in older adults: a prospective cohort study in 28 countries. Age Ageing 2022 May 1;51(5):afac117. [doi: 10.1093/ageing/afac117] [Medline: 35639798]
5. García-Hermoso A, Cavero-Redondo I, Ramírez-Vélez R, et al. Muscular strength as a predictor of all-cause mortality in an apparently healthy population: a systematic review and meta-analysis of data from approximately 2 million men and women. Arch Phys Med Rehabil 2018 Oct;99(10):2100-2113. [doi: 10.1016/j.apmr.2018.01.008] [Medline: 29425700]
6. Newman AB, Kupelian V, Visser M, et al. Strength, but not muscle mass, is associated with mortality in the health, aging and body composition study cohort. J Gerontol A Biol Sci Med Sci 2006 Jan;61(1):72-77. [doi: 10.1093/gerona/61.1.72] [Medline: 16456196]
7. Sayer AA, Cruz-Jentoft A. Sarcopenia definition, diagnosis and treatment: consensus is growing. Age Ageing 2022 Oct 6;51(10):afac220. [doi: 10.1093/ageing/afac220] [Medline: 36273495]
8. Beaudart C, Dawson A, Shaw SC, et al. Nutrition and physical activity in the prevention and treatment of sarcopenia: systematic review. Osteoporos Int 2017 Jun;28(6):1817-1833. [doi: 10.1007/s00198-017-3980-9] [Medline: 28251287]
9. McCarron HR, Zmora R, Gaugler JE. A web-based mobile app with a smartwatch to support social engagement in persons with memory loss: pilot randomized controlled trial. JMIR Aging 2019 Jun 18;2(1):e13378. [doi: 10.2196/13378] [Medline: 31518270]
10. Wu J, Fu Y, Chen D, et al. Sedentary behavior patterns and the risk of non-communicable diseases and all-cause mortality: a systematic review and meta-analysis. Int J Nurs Stud 2023 Oct;146:104563. [doi: 10.1016/j.ijnurstu.2023.104563] [Medline: 37523952]
11. Cai R, Chao J, Gao C, Gao L, Hu K, Li P. Association between sleep duration and cognitive frailty in older Chinese adults: prospective cohort study. JMIR Aging 2025 Apr 23;8:e65183. [doi: 10.2196/65183] [Medline: 40267503]
12. Sun JW, Zhao LG, Yang Y, Ma X, Wang YY, Xiang YB. Association between television viewing time and all-cause mortality: a meta-analysis of cohort studies. Am J Epidemiol 2015 Dec 1;182(11):908-916. [doi: 10.1093/aje/kwv164] [Medline: 26568572]
13. Itani O, Jike M, Watanabe N, Kaneita Y. Short sleep duration and health outcomes: a systematic review, meta-analysis, and meta-regression. Sleep Med 2017 Apr;32:246-256. [doi: 10.1016/j.sleep.2016.08.006] [Medline: 27743803]
14. Holt-Lunstad J. Why social relationships are important for physical health: a systems approach to understanding and modifying risk and protection. Annu Rev Psychol 2018 Jan 4;69:437-458. [doi: 10.1146/annurev-psych-122216-011902] [Medline: 29035688]
15. Petermann-Rocha F, Pell JP, Celis-Morales C, Ho FK. Frailty, sarcopenia, cachexia and malnutrition as comorbid conditions and their associations with mortality: a prospective study from UK Biobank. J Public Health (Oxf) 2022 Jun 27;44(2):e172-e180. [doi: 10.1093/pubmed/fdaa226] [Medline: 33423060]
16. Parry SW, Bamford C, Deary V, et al. Cognitive-behavioural therapy-based intervention to reduce fear of falling in older people: therapy development and randomised controlled trial—the Strategies for Increasing Independence, Confidence and Energy (STRIDE) study. Health Technol Assess 2016 Jul;20(56):1-206. [doi: 10.3310/hta20560] [Medline: 27480813]
17. Yang J, Huang J, Yang X, Li S, Wu X, Ma X. The association of living alone and social isolation with sarcopenia: a systematic review and meta-analysis. Ageing Res Rev 2023 Nov;91:102043. [doi: 10.1016/j.arr.2023.102043] [Medline: 37647996]

XSL•FO
RenderX

18. Han H, Cao Y, Feng C, et al. Association of a Healthy lifestyle with all-cause and cause-specific mortality among individuals with type 2 diabetes: a prospective study in UK Biobank. Diabetes Care 2022 Feb 1;45(2):319-329. [doi: 10.2337/dc21-1512] [Medline: 34857534]
19. Gu Z, Liu S, Ma H, et al. Estimation of machine learning-based models to predict dementia risk in patients with atherosclerotic cardiovascular diseases: UK Biobank study. JMIR Aging 2025 Feb 26;8:e64148. [doi: 10.2196/64148] [Medline: 40009844]
20. Sudlow C, Gallacher J, Allen N, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015 Mar;12(3):e1001779. [doi: 10.1371/journal.pmed.1001779] [Medline: 25826379]
21. Dodds RM, Granic A, Robinson SM, Sayer AA. Sarcopenia, long-term conditions, and multimorbidity: findings from UK Biobank participants. J Cachexia Sarcopenia Muscle 2020 Feb;11(1):62-68. [doi: 10.1002/jcsm.12503] [Medline: 31886632]
22. Zhang YB, Chen C, Pan XF, et al. Associations of healthy lifestyle and socioeconomic status with mortality and incident cardiovascular disease: two prospective cohort studies. BMJ 2021 Apr 14;373:n604. [doi: 10.1136/bmj.n604] [Medline: 33853828]
23. Jike M, Itani O, Watanabe N, Buysse DJ, Kaneita Y. Long sleep duration and health outcomes: a systematic review, meta-analysis and meta-regression. Sleep Med Rev 2018 Jun;39:25-36. [doi: 10.1016/j.smrv.2017.06.011] [Medline: 28890167]
24. Wang C, Bangdiwala SI, Rangarajan S, et al. Association of estimated sleep duration and naps with mortality and cardiovascular events: a study of 116 632 people from 21 countries. Eur Heart J 2019 May 21;40(20):1620-1629. [doi: 10.1093/eurheartj/ehy695] [Medline: 30517670]
25. van de Vegte YJ, Said MA, Rienstra M, van der Harst P, Verweij N. Genome-wide association studies and Mendelian randomization analyses for leisure sedentary behaviours. Nat Commun 2020 Apr 21;11(1):1770. [doi: 10.1038/s41467-020-15553-w] [Medline: 32317632]
26. Moore AR, Brolin TJ, Smith RA, Azar FM, Throckmorton TW. Outcomes after primary reverse total shoulder arthroplasty in patients with comorbid anxiety and depressive disorders: a retrospective cohort study. Curr Orthop Pract 2021;32(1):37-42. [doi: 10.1097/BCO.0000000000000940]
27. Services. National Health Service Information Centre. URL: http://content.digital.nhs.uk/services [accessed 2025-04-20]
28. Brown JC, Harhay MO, Harhay MN. Physical activity, diet quality, and mortality among sarcopenic older adults. Aging Clin Exp Res 2017 Apr;29(2):257-263. [doi: 10.1007/s40520-016-0559-9] [Medline: 27020695]
29. Dent E, Morley JE, Cruz-Jentoft AJ, et al. International clinical practice guidelines for sarcopenia (ICFSR): screening, diagnosis and management. J Nutr Health Aging 2018;22(10):1148-1161. [doi: 10.1007/s12603-018-1139-9] [Medline: 30498820]
30. Chudasama YV, Khunti K, Gillies CL, et al. Healthy lifestyle and life expectancy in people with multimorbidity in the UK Biobank: a longitudinal cohort study. PLoS Med 2020 Sep;17(9):e1003332. [doi: 10.1371/journal.pmed.1003332] [Medline: 32960883]
31. Locquet M, Bruyère O, Lengelé L, Reginster JY, Beaudart C. Relationship between smoking and the incidence of sarcopenia: the SarcoPhAge cohort. Public Health (Fairfax) 2021 Apr;193:101-108. [doi: 10.1016/j.puhe.2021.01.017] [Medline: 33773322]
32. Li Y, Pan A, Wang DD, et al. Impact of healthy lifestyle factors on life expectancies in the US population. Circulation 2018 Jul 24;138(4):345-355. [doi: 10.1161/CIRCULATIONAHA.117.032047] [Medline: 29712712]
33. Bu YL, Wang C, Zhao C, Lu X, Gao W. The association of alcohol consumption with the risk of sarcopenia: a dose-response meta-analysis. Am J Drug Alcohol Abuse 2024 May 3;50(3):305-320. [doi: 10.1080/00952990.2023.2300049] [Medline: 38232346]
34. Lengelé L, Locquet M, Moutschen M, et al. Frailty but not sarcopenia nor malnutrition increases the risk of developing COVID-19 in older community-dwelling adults. Aging Clin Exp Res 2022 Jan;34(1):223-234. [doi: 10.1007/s40520-021-01991-z] [Medline: 34689315]
35. Yin J, Jin X, Shan Z, et al. Relationship of sleep duration with all-cause mortality and cardiovascular events: a systematic review and dose-response meta-analysis of prospective cohort studies. J Am Heart Assoc 2017 Sep 9;6(9):e005947. [doi: 10.1161/JAHA.117.005947] [Medline: 28889101]
36. Huang BH, Duncan MJ, Cistulli PA, Nassar N, Hamer M, Stamatakis E. Sleep and physical activity in relation to all-cause, cardiovascular disease and cancer mortality risk. Br J Sports Med 2022 Jul;56(13):718-724. [doi: 10.1136/bjsports-2021-104046] [Medline: 34187783]
37. Lourida I, Hannon E, Littlejohns TJ, et al. Association of lifestyle and genetic risk with incidence of dementia. JAMA 2019 Aug 6;322(5):430-437. [doi: 10.1001/jama.2019.9879] [Medline: 31302669]
38. Sossdorf M, Fischer J, Meyer S, et al. Physical exercise induces specific adaptations resulting in reduced organ injury and mortality during severe polymicrobial sepsis. Crit Care Med 2013 Oct;41(10):e246-e255. [doi: 10.1097/CCM.0b013e31828a2ae3] [Medline: 23887230]
39. Ferri G, Serano M, Isopi E, et al. Resolvin D1 improves airway inflammation and exercise capacity in cystic fibrosis lung disease. FASEB J 2023 Nov;37(11):e23233. [doi: 10.1096/fj.202301495R] [Medline: 37823221]

XSL•FO
RenderX

40. Calderin EP, Zheng JJ, Boyd NL, et al. Exercise-induced specialized proresolving mediators stimulate AMPK phosphorylation to promote mitochondrial respiration in macrophages. Mol Metab 2022 Dec;66:101637. [doi: 10.1016/j.molmet.2022.101637] [Medline: 36400404]
41. Hou Q, Huang J, Zhao L, et al. Dietary genistein increases microbiota-derived short chain fatty acid levels, modulates homeostasis of the aging gut, and extends healthspan and lifespan. Pharmacol Res 2023 Feb;188:106676. [doi: 10.1016/j.phrs.2023.106676] [Medline: 36693599]
42. Badran M, Laher I. Waterpipe (shisha, hookah) smoking, oxidative stress and hidden disease potential. Redox Biol 2020 Jul;34:101455. [doi: 10.1016/j.redox.2020.101455] [Medline: 32086009]
43. Mao Y, Huang P, Wang Y, Wang M, Li MD, Yang Z. Genome-wide methylation and expression analyses reveal the epigenetic landscape of immune-related diseases for tobacco smoking. Clin Epigenetics 2021 Dec 9;13(1):215. [doi: 10.1186/s13148-021-01208-0] [Medline: 34886889]
44. Mezue K, Osborne MT, Abohashem S, et al. Reduced stress-related neural network activity mediates the effect of alcohol on cardiovascular risk. J Am Coll Cardiol 2023 Jun 20;81(24):2315-2325. [doi: 10.1016/j.jacc.2023.04.015] [Medline: 37316112]
45. Di Castelnuovo A, Costanzo S, Bonaccio M, et al. Moderate alcohol consumption is associated with lower risk for heart failure but not atrial fibrillation. JACC Heart Fail 2017 Nov;5(11):837-844. [doi: 10.1016/j.jchf.2017.08.017] [Medline: 29032141]
46. Simon L, Souza-Smith FM, Molina PE. Alcohol-associated tissue injury: current views on pathophysiological mechanisms. Annu Rev Physiol 2022 Feb 10;84:87-112. [doi: 10.1146/annurev-physiol-060821-014008] [Medline: 35143331]
47. Forsyth CB, Voigt RM, Keshavarzian A. Intestinal CYP2E1: a mediator of alcohol-induced gut leakiness. Redox Biol 2014;3:40-46. [doi: 10.1016/j.redox.2014.10.002] [Medline: 25462064]
48. Jung J, McCartney DL, Wagner J, et al. Additive effects of stress and alcohol exposure on accelerated epigenetic aging in alcohol use disorder. Biol Psychiatry 2023 Feb 15;93(4):331-341. [doi: 10.1016/j.biopsych.2022.06.036] [Medline: 36182531]
49. Healy GN, Matthews CE, Dunstan DW, Winkler EAH, Owen N. Sedentary time and cardio-metabolic biomarkers in US adults: NHANES 2003-06. Eur Heart J 2011 Mar;32(5):590-597. [doi: 10.1093/eurheartj/ehq451] [Medline: 21224291]
50. Patel SR, Zhu X, Storfer-Isser A, et al. Sleep duration and biomarkers of inflammation. Sleep 2009 Feb;32(2):200-204. [doi: 10.1093/sleep/32.2.200]
51. Picca A, Coelho-Junior HJ, Calvani R, Marzetti E, Vetrano DL. Biomarkers shared by frailty and sarcopenia in older adults: a systematic review and meta-analysis. Ageing Res Rev 2022 Jan;73:101530. [doi: 10.1016/j.arr.2021.101530] [Medline: 34839041]
52. Fortmann AL, Gallo LC. Social support and nocturnal blood pressure dipping: a systematic review. Am J Hypertens 2013 Mar;26(3):302-310. [doi: 10.1093/ajh/hps041] [Medline: 23382479]
53. Li H, Xia N. The role of oxidative stress in cardiovascular disease caused by social isolation and loneliness. Redox Biol 2020 Oct;37:101585. [doi: 10.1016/j.redox.2020.101585] [Medline: 32709420]
54. Cacioppo JT, Cacioppo S, Capitanio JP, Cole SW. The neuroendocrinology of social isolation. Annu Rev Psychol 2015 Jan 3;66:733-767. [doi: 10.1146/annurev-psych-010814-015240] [Medline: 25148851]
55. Uchino BN, Trettevik R, Kent de Grey RG, Cronan S, Hogan J, Baucom BRW. Social support, social integration, and inflammatory cytokines: a meta-analysis. Health Psychol 2018 May;37(5):462-471. [doi: 10.1037/hea0000594] [Medline: 29565600]
56. Sylvia KE, Demas GE. A gut feeling: microbiome-brain-immune interactions modulate social and affective behaviors. Horm Behav 2018 Mar;99:41-49. [doi: 10.1016/j.yhbeh.2018.02.001] [Medline: 29427583]
57. de Grey RGK, Uchino BN, Trettevik R, Cronan S, Hogan JN. Social support and sleep: a meta-analysis. Health Psychol 2018 Aug;37(8):787-798. [doi: 10.1037/hea0000628] [Medline: 29809022]
58. Shankar A, McMunn A, Banks J, Steptoe A. Loneliness, social isolation, and behavioral and biological health indicators in older adults. Health Psychol 2011 Jul;30(4):377-385. [doi: 10.1037/a0022826] [Medline: 21534675]
59. Narasimhulu CA, Singla DK. Amelioration of diabetes-induced inflammation mediated pyroptosis, sarcopenia, and adverse muscle remodelling by bone morphogenetic protein-7. J Cachexia Sarcopenia Muscle 2021 Apr;12(2):403-420. [doi: 10.1002/jcsm.12662] [Medline: 33463042]
60. Hashimoto Y, Takahashi F, Okamura T, Hamaguchi M, Fukui M. Diet, exercise, and pharmacotherapy for sarcopenia in people with diabetes. Metab Clin Exp 2023 Jul;144:155585. [doi: 10.1016/j.metabol.2023.155585] [Medline: 37156410]
61. Sepúlveda-Loyola W, Osadnik C, Phu S, Morita AA, Duque G, Probst VS. Diagnosis, prevalence, and clinical impact of sarcopenia in COPD: a systematic review and meta-analysis. J Cachexia Sarcopenia Muscle 2020 Oct;11(5):1164-1176. [doi: 10.1002/jcsm.12600] [Medline: 32862514]
62. Wang C, Guo X, Xu X, et al. Association between sarcopenia and frailty in elderly patients with chronic kidney disease. J Cachexia Sarcopenia Muscle 2023 Aug;14(4):1855-1864. [doi: 10.1002/jcsm.13275] [Medline: 37300354]
63. Au DH, Gleason E, Hunter-Merrill R, et al. Lifestyle intervention and excess weight in chronic obstructive pulmonary disease (COPD): INSIGHT COPD randomized clinical trial. Ann Am Thorac Soc 2023 Dec;20(12):1743-1751. [doi: 10.1513/AnnalsATS.202305-458OC] [Medline: 37769182]

XSL•FO
RenderX

64. Kalantar-Zadeh K, Bross R, Wang C, Moore LW. Community nutrition and lifestyle medicine for kidney health: awareness of the social determinants of health in renal nutrition. J Ren Nutr 2023 Jul;33(4):505-507. [doi: 10.1053/j.jrn.2023.05.003]
65. Johansson J, Morseth B, Scott D, Strand BH, Hopstock LA, Grimsgaard S. Moderate-to-vigorous physical activity modifies the relationship between sedentary time and sarcopenia: the Tromsø Study 2015-2016. J Cachexia Sarcopenia Muscle 2021 Aug;12(4):955-963. [doi: 10.1002/jcsm.12718] [Medline: 34060236]
66. Munafò MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection bias can substantially influence observed associations. Int J Epidemiol 2018 Feb 1;47(1):226-235. [doi: 10.1093/ije/dyx206] [Medline: 29040562]

Abbreviations

CVD: cardiovascular disease
EWGSOP2: European Working Group on Sarcopenia in Older People 2
HR: hazard ratio
PAR: population-attributable risk

Edited by J Wang; submitted 14.08.24; peer-reviewed by DE Sacar, J Zhang, M Locquet; revised version received 22.04.25; accepted 12.05.25; published 28.07.25.

Please cite as:
Wang N, Zhang Y, Xie J, Lu N, Zheng A, Li C, Wei J, Zeng C, Lei G, Wang Y
Association of a Healthy Lifestyle With All-Cause and Cause-Specific Mortality Among Individuals With Probable Sarcopenia: Population-Based Cohort Study
JMIR Aging 2025;8:e65374
URL: https://aging.jmir.org/2025/1/e65374
doi:10.2196/65374

© Ning Wang, Yuqing Zhang, Junqing Xie, Na Lu, Aojie Zheng, Changjun Li, Jie Wei, Chao Zeng, Guanghua Lei, Yilun Wang. Originally published in JMIR Aging (https://aging.jmir.org), 28.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A 12-Month Digital Peer-Supported App Intervention to Promote Physical Activity Among Community-Dwelling Older Adults: Follow-Up Study of a Nonrandomized Controlled Trial

Kento Tabira[1,2], PhD, PT; Yuko Oguma[1,2], MPH, MD, PhD; Shota Yoshihara[3,4], MSc, OT; Megumi Shibuya[4], MSc; Manabu Nakamura[1,2], MSc, PT; Natsue Doihara[1], PhD; Akihiro Hirata[1,5], PhD; Tomoki Manabe[1,2], MPH, RN, PHN; Takashi Yamashita[6], MA, MPH, PhD

[1]Sports Medicine Research Center, Keio University, 4-1-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa, Japan

[2]Graduate School of Health Management, Keio University, Fujisawa, Kanagawa, Japan

[3]Department of Rehabilitation Sciences, Graduate School of Medical Sciences, Kitasato University, Sagamihara, Kanagawa, Japan

[4]A10 Lab Inc, Chuo-ku, Tokyo, Japan

[5]Japan Society for the Promotion of Science, Chiyoda-ku, Tokyo, Japan

[6]Department of Sociology, Anthropology, and Public Health, Gerontology Doctoral Program, University of Maryland Baltimore County, Baltimore, MD, United States

Corresponding Author:
Yuko Oguma, MPH, MD, PhD
Sports Medicine Research Center, Keio University, 4-1-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa, Japan

Abstract

Background: Mobile apps and peer support are known to effectively promote physical activity in older adults, which, in turn, improves physical function. Previously, we investigated the feasibility and impact of using digital peer-supported apps (DPSAs) to increase physical activity among older adults over a 3-month period. However, the long-term feasibility and impact on sustainable behavior change remain unknown.

Objective: This study aims to evaluate the 12-month feasibility of the DPSA and to obtain preliminary estimates of its effects on physical activity and physical function among older Japanese adults.

Methods: This nonrandomized controlled trial recruited older adults aged 65 years or older from 2 physical activity programs. Participants chose either the intervention (app program + exercise instruction) group or the control (exercise instruction only) group. Only those participants who had completed the 3-month intervention and wished to continue in the 12-month follow-up intervention study were included. DPSA feasibility was assessed using retention and adherence rates. Physical activity was assessed using accelerometers, capturing daily step count, light-intensity activity, moderate to vigorous intensity activity, and sedentary behavior. Physical function was evaluated using grip strength and the 30-second chair stand test (CS-30). Accelerometer measurements were collected every 3 months over 12 months (5 time points, including baseline), whereas physical function was measured at baseline, 3 months, and 12 months.

Results: The follow-up study included 44 of 66 participants from the 3-month intervention study, with 26 participants in the intervention group and 18 participants in the control group. The 12-month retention rate for participants in the DPSA intervention group was 73% (19/26), whereas the retention rate among all 41 participants, including those who chose not to participate in the follow-up study, was 46% (19/41). The adherence rate was 85.9%. The average number of steps per day (95% CI) in the intervention group changed before and after DPSA use (P=.048). We observed an increase of 1736 (β=1736, 95% CI 232-3241) steps per day compared with baseline. No significant change was observed in the control group. There were significant within-group differences in CS-30 scores for both intervention (P<.001) and control (P=.03) groups over the 12-month period. Specifically, there was a significant change in CS-30 scores (95% CI) between the baseline and 12-month assessments for the intervention (β=6.5, 95% CI 3.8-9.1; P<.001) and control (β=3.8, 95% CI 0.6-7.1; P=.02) groups.

Conclusions: Participants with long-term DPSA use observed increases in average daily steps and CS-30 scores before and after DPSA use, although only a limited number of older adults had long-term access to the DPSA. Identifying ways to expand long-term DPSA use among older adults is necessary. Additionally, randomized controlled trials should be conducted to determine the long-term effects of DPSAs on physical activity and function in older adults.

Trial Registration: University Hospital Medical Information Network UMIN000050618; https://center6.umin.ac.jp/cgi-open-bin/ctr_e/ctr_view.cgi?recptno=R000057008

XSL•FO
RenderX

(JMIR Aging 2025;8:e66610) doi:10.2196/66610

KEYWORDS

physical activity; exercise; exercising; physical function; older adult; aging; eHealth; peer support; mobile phone; mHealth; mobile health; mobile app; app; application; smartphone; digital

Introduction

Background

The world's population is aging at an unprecedented rate [1]. The number of adults older than 65 years has tripled over the past 50 years, and by 2050, older adults are expected to account for a quarter of the global population [2-4]. Japan has a high proportion of older adults: 29.1% of the total population was aged 65 years or older in 2023 [5]. Healthy aging is a global health care challenge as population aging accelerates [6]. Regular physical activity aids in reducing the risk of noncommunicable diseases [7] and is associated with improved physical health [8] and increased life expectancy [9]. However, despite decades of public health interventions, the global physical activity level has remained stable or even declined, making it an important health policy challenge [10]. Given the world's aging population and the health benefits of physical activity, it is critical to promote regular physical activity among older adults. The Japanese guidelines for physical activity [11] recommend a minimum of 15 metabolic equivalent (MET) hours per week of physical activity with an intensity of at least 3 METs in older adults. Physical activity of 15 MET hours per week can be converted into steps, which is more than 6000 steps per day [11]. However, few older adults meet this recommendation: among men, 45% of those aged 65 - 74 years, 32% of those aged 75 - 84 years, and 11% of those aged 85 years or older meet this recommendation, and among women, 38% of those aged 65 - 74 years, 22% of those aged 75-84 years, and 5% of those aged 85 years or older meet this recommendation [11]. Regular physical activity improves physical function in older adults [12]. Declining physical function is linked with the loss of mobility and activities of daily life, which are core dimensions of physical disabilities [13,14], and thus, both physical activity and physical function need to be improved.

Recently, mobile apps have been successfully used to increase physical activity levels [15,16]. eHealth, or electronic health, encompasses health care services supported by information and communication technology, including computers, mobile phones, and satellite communications, for health services and information. Moreover, mHealth, or mobile health, refers to the use of smart or portable devices for providing health services and information [17]. These interventions for older adults have been shown to be effective in increasing the time spent in physical activity, energy expenditure in physical activity, and steps walked [18,19]. In a review comparing mHealth with face-to-face interventions, interventions that included mHealth were shown to have increased steps and total physical activity, but there was no observed difference in physical function [20].

In the systematic review by Duan et al [21], eHealth interventions for physical activity have shown that theory-based interventions are more effective than non–theory-based interventions. The transtheoretical model and social cognitive theory were the top 2 most frequently supporting theories, and the studies included in this systematic review with the largest effect sizes were based on these 2 theories [22]. The social cognitive theory proposed by Bandura [23] stipulates that behavior is learned by observing and imitating others. This process is called observational learning or modeling and has been extensively studied in the context of motor skill development and education [24-26]. Self-efficacy, an important aspect of social cognitive theory [23], is an crucial determinant of exercise persistence and outcomes; interventions based on self-efficacy can promote exercise participation [27].

The effectiveness of peer support interventions for physical activity is often explained by social cognitive theory [28]. Webel et al [29] defined peer support as "a method of teaching or facilitating health promotion that makes use of people sharing specific health messages with members of their own community." Our previous study using a digital peer-supported app (DPSA) framed by the social cognitive theory showed that the feasibility of the DPSA was adequate and that the number of daily steps and the level of moderate to vigorous intensity physical activity (MVPA) increased in older participants [30]. There are 2 main types of peer support [31]: the first includes methods related to education and information, such as peer tutoring and mentoring; the second is the emotional support provided by peers. Our research is based on peer support interventions that provide emotional support. Peer support is provided by comparable peers and promotes physical activity in ways that cannot be done by professionals or family members; Burton et al [32] reported that peer support increased adherence to an exercise program; Ginis et al [28] reported that peer support was as effective as professional intervention. Peer support may be cost-effective when considering the expense of paying professionals [33]. Peer support through the DPSA includes social support, which contributes to the success of eHealth and mHealth interventions for increasing physical activity among older adults [34]. In addition, DPSA interventions do not require in-person gatherings, thus reducing constraints owing to scheduling issues, meeting locations, and costs (eg, transportation) [35]. Thus, the DPSA may be effective in promoting physical activity among older adults. However, our previous study was a short-term intervention of 3 months, and the long-term feasibility and impact for sustainable behavior change remains unknown.

Three of 4 review studies concluded that mHealth or eHealth interventions are effective over short term (1 - 6 months) in promoting physical activity in adults aged 50 years or older [34]. All 3 reviews incorporated randomized controlled trials (RCTs) comparing interventions that were not eHealth or mHealth (eg, paper-based intervention, professional face-to-face intervention, and group face-to-face intervention), or no intervention. Despite the demonstrated long-term health benefits of physical activity [36], long-term empirical evidence of

XSL•FO
RenderX

mHealth and eHealth, beyond 6 - 12 months, remains scarce [37-39]. Furthermore, no study has continued the app intervention for 12 months and collected device-based physical activity measures in community-dwelling older adults older than 65 years [40]. Physical activity interventions for older adults often face challenges regarding long-term participation owing to age-related health decline, low self-efficacy, and poor geographic access to physical activity spaces [41,42]. There is a need to test the long-term effectiveness of the DPSA in promoting physical activity among older adults. However, before testing the long-term effectiveness of the DPSA on a large scale, a reasonable first step is to examine the feasibility and preliminary changes in physical activity, physical function, and self-efficacy in community settings. We hypothesized that 1 year of DPSA use would increase physical activity owing to increased self-efficacy for exercise. We also expect that the increase in physical activity will be accompanied by an increase in physical function.

Objectives

This study was a 12-month longitudinal study of participants in a 3-month DPSA intervention study who volunteered to participate in a follow-up study. The objectives of this study were twofold: (1) to evaluate the feasibility (retention and adherence rates) of using the DPSA to promote physical activity in older adults over a 12-month period, and (2) to measure preliminary estimates of the effects of physical activity, physical function, and self-efficacy for exercise through the use of the DPSA.

Methods

Study Design

This study is a nonrandomized controlled trial of 2 groups conducted over 12 months and is a follow-up study of a 3-month intervention trial [30]. This study was conducted in Fujisawa City, Kanagawa, Japan. Fujisawa City is in the southeastern part of Kanagawa and is an urban area close to Tokyo. As of April 2023, the city had a population of 445,291; of those, 24.5% (109,005) were aged 65 years or older. The percentage of older adults in the total population is increasing year by year [43]. This study was conducted as a collaboration between local governments, mobile app development companies, and universities. Industry-government-academia collaboration is important to further scientific research that is relevant to real-world community issues [44,45].

Ethical Considerations

This study was approved by the research ethics committee of Sports Medicine Research Center at Keio University (approval no. 2022 - 07). Informed consent for the follow-up study was obtained from all participants in the 3-month intervention study. The data obtained were anonymized. The study protocol was registered in the University Hospital Medical Information Network (UMIN000050618).

Participants

The study included Fujisawa City residents aged 65 years or older. In Japan, older adults are generally defined as persons aged 65 years or older [40]. We recruited participants for two 3-month programs designed to increase physical activity [30]. Participants from 2 different areas within the municipality of Fujisawa City were recruited through flyers, local newsletters, and calls to related organizations. Participants chose either intervention (app program and exercise instruction) or control (exercise instruction only) group. The 3-month intervention study [30] included 74 participants (intervention group: 41, control group: 33). The follow-up study was introduced to 66 participants who completed the 3-month intervention. Participants (n=8) who did not provide their informed consent were excluded from the follow-up study. The eligibility criteria were adults aged 65 years or older who could walk independently and perform daily activities without being advised by a physician to refrain from physical activity (self-reported criterion). Prospective participants were screened using a personal health status questionnaire based on the Physical Activity Readiness Questionnaire [46-48] to ascertain any potential health problems with study participation. Because the purpose of this study was to assess feasibility and obtain preliminary estimates, sample size was not calculated; the number of participants was limited because the study was conducted in collaboration with the local government.

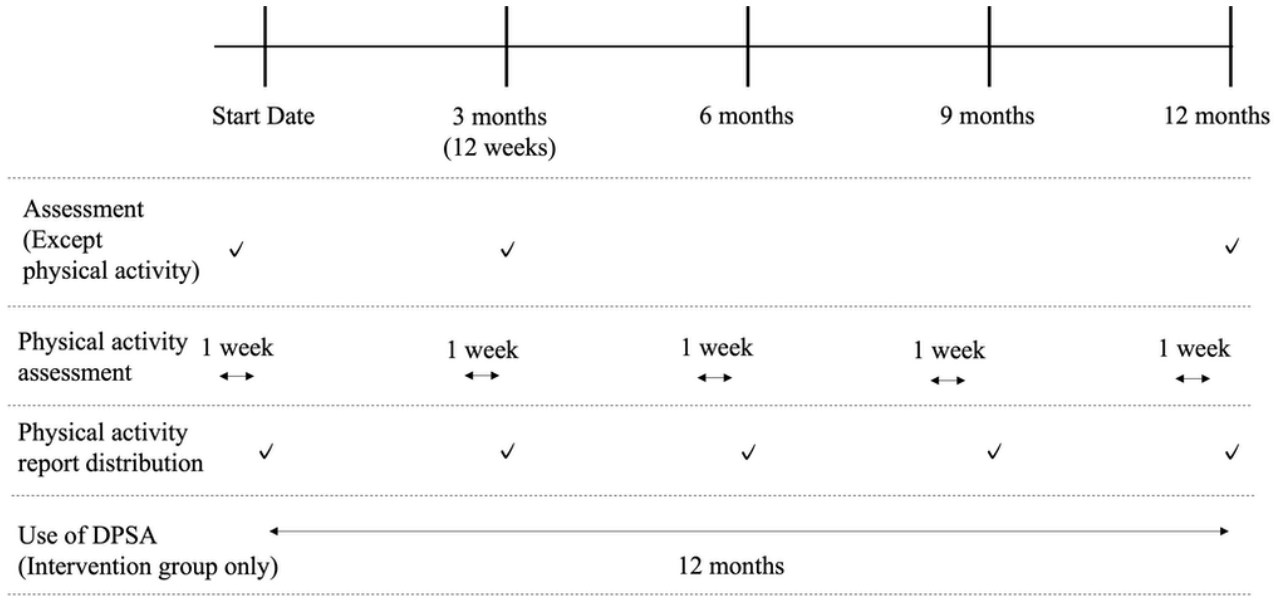
Intervention

Program

Regardless of program selection, all participants underwent face-to-face exercise instruction, program introduction, and baseline assessment by a physical therapist or health fitness instructor. Exercise instructions focused on aerobic, stretching, muscle strengthening, and balance exercises based on the original “Fujisawa + 10 exercise” program [49,50]. The timeline of the study procedure is shown in Figure 1. Both intervention and control groups were instructed to increase their daily physical activities. The participants completed surveys and physical function measurements at baseline (start date), 3 months, and 12 months postintervention. Additionally, physical activity levels were measured every 3 months, 5 times in total, using triaxial accelerometers. Individualized physical activity reports were generated from the collected data and provided as feedback to the participants. The intervention group began using the app 1 week after the baseline outcome assessment.

XSL•FO
RenderX

Figure 1. Timeline of 12-month intervention procedure. DPSA: digital peer-supported app.



Digital Peer-Supported App

This study used Minchalle, a commercially available DPSA [51]. This mobile app was developed in June 2015 by A10 Lab Inc, with an initial release in November 2015. Figure 2 shows a sample app screen. The DPSA creates a group chat for up to 5 people with a common goal, and participants anonymously interact with each other in the group. The common goal of the intervention group was to increase their physical activity by walking and exercising. Once a day, participants posted their step counts, photographs, and comments in a group chat box. The main functions of the DPSA used in this study were as follows: (1) posting photographs, step counts, and comments about the day, (2) reaction buttons from group members (Figure 2), (3) setting step count goals on a group basis, and (4) providing feedback on the group's total daily step count. Step counts were measured using a smartphone, and the DPSA reported the number of steps taken on that day at the time of posting. The participants were asked to carry smartphones throughout the day while they were awake. Participants had the option to post comments or photographs multiple times a day and engage with other members. The mobile app was provided to the participants free of charge. Details of the DPSA's functionality are summarized in Multimedia Appendix 1.

Figure 2. Examples of mobile app screens. (**A**) A group is selected. (**B**) Photographs, step count, and comments are posted on the group. A photograph taken that day is posted and comments are added on the day's events. (**C**) Contents of the posts are displayed in the group. The total number of steps for the group is displayed. (**D**) Responses to posts by group members.



Characteristics of Research Participants

In addition to general characteristics, such as age and sex, the survey enquired about smartphone ownership, frequency of app use (except DPSA), and exercise habits. Body weight (kg) was measured using a digital scale, and height (m) was measured with a stadiometer after removing shoes. BMI was calculated as body weight divided by the square of height. Exercise habits

XSL•FO
RenderX

were considered as “those who exercised at least twice a week for 30 minutes or longer each time for at least 1 year” [52].

The frequency of neighborhood interaction was assessed by asking the participants how many times they interact with people in the neighborhood within 1 week. Group exercise participation was defined as study participants who participate in a group of 3 or more people who meet voluntarily to exercise.

DPSA Feasibility

DPSA feasibility was assessed based on retention and adherence rates during the year of program implementation. The retention rate indicates how many of the participants continued to use the DPSA for 12 months. The adherence rate indicates how often participants used the DPSA during the intervention period. The DPSA used in this study excludes a person from a group if they have not posted a set of step counts, photographs, or comments for 15 consecutive days. Dropouts were defined as those excluded from the group during the 12 months of DPSA use by researchers. Retention rates were calculated using a population of 26 participants in the intervention group and a population of all 41 participants who decided not to participate in the follow-up study. The retention rate was considered good if it was ≥70% (≥29 retention out of 41) based on previous studies by Farrance et al [53] and Picorelli et al [54]. The adherence rate was calculated by dividing the number of sets of step counts, photographs, and comments posted during the intervention period by the duration of the intervention. Adherence was calculated as the percentage of both participants, including dropouts and not including dropouts. Considering that the adherence rate for participants in the 3-month program was 87.7% [30], an adherence rate of ≥80% was considered good. The adherence rates were also calculated by group (7 groups: A-G). The number of all chat posts per person by group was calculated to assess the degree to which the group was used. The observed negative physical conditions during the intervention were ascertained by interviewing participants 12 months later. The app developers and the municipality were available to support the participants for any privacy breaches and technical issues.

Outcome Measure

To assess physical activity, participants were asked to wear a triaxial accelerometer [55] (Active Style Pro HJA-750C activity meter; Omron Healthcare) at the waist level for 7 consecutive days for a total of 5 times every 3 months starting before the intervention. This accelerometer provides a relatively accurate measure of physical activity in healthy older adults [56]. Participants were instructed not to remove the device unless required for certain tasks, such as changing clothes and bathing. At the end of the measurement, all the data collected were transferred from the accelerometer to a personal computer. Following the suggested method [57] for estimating physical activity, an individual was required to record ≥10 hours of activity per day for 3 days to be included in the subsequent analyses. The data were collected in 60-second epochs for data analysis and used to estimate the intensity of the activity (METs). Outcome measurements of physical activity included the mean daily step count and time spent in sedentary behavior (SB: ≤1.5 METs), light-intensity physical activity (LPA: 1.6 - 2.9 METs), and MVPA (≥3 METs) per day. The number of steps reported to the group chat in the DPSA was measured by the smartphone but was not used as an outcome.

Physical function was assessed using grip strength and the 30-second chair stand test (CS-30). The grip strength was measured using a digital dynamometer (Grip D; TKK 5401; Takei Scientific Instruments). This digital dynamometer was reliable and was validated relative to the Jamar dynamometer, which is the most frequently cited instrument for assessing grip strength in adults aged older than 60 years [58]. Measurements were taken in the standing position with the elbow joint in extension and the wrist joint in midextension. The left and right hands were measured once, and the highest value was used for data analysis. For the CS-30 test [59], seated participants were instructed to stand up from the chair with their arms crossed at the chest level as many times as possible in 30 seconds. The CS-30 has been reported to be quite reliable and valid as an indicator of lower-limb function in older adults [59].

Self-efficacy for exercise consisted of 4 questions on self-confidence in exercising under the following conditions [60]: physical fatigue, mental stress, lack of time, and bad weather. In response to the question, “Do you have the confidence to exercise regularly under the following conditions?” participants were asked to select 1 of 5 answers ranging from “No, I don’t have any confidence at all (1 point)” to “Yes, I am quite confident (5 points).” The total score ranged from 4 to 20.

Statistical Analysis

This study used intention-to-treat analysis. Participant characteristics between groups were compared using independent sample *t*, chi-square, and Mann-Whitney *U* tests. Fixed-effects models were used because of the intensive repeated-measures design [61]. The advantage of this method is that it can handle nested observations, unbalanced numbers of observations, and missing values [62]. Although it would have been desirable to use a model that included random effects in this study, sample size limitations impeded the convergence of the mixed-effects model, and thus, we applied a model with fixed effects only. Yet, the fixed-effects model is still capable of capturing changes in the repeated measures in the outcomes. On a related note, linear mixed-effects models can be used with small sample sizes [63,64]. Between-group differences (intervention vs control) were analyzed using fixed-effects models adjusted for baseline age, sex, and app usage frequency (at baseline). The interaction between the groups and the time of the intervention was then analyzed.

Subsequently, the effects of the intervention for each group were analyzed using linear mixed-effects models, and significant differences compared with preintervention were evaluated using the Bonferroni method (accelerometer data were adjusted for wear time). Dependent variables, such as daily step count, SB, LPA, MVPA, grip strength, CS-30, and self-efficacy for exercise, were analyzed in separate models. Although the daily step count distribution did not precisely follow the normal distribution, the consistent results were obtained when applying a square root transformation, and, therefore, the results without the square root transformation are presented for interpretation.

XSL•FO
RenderX

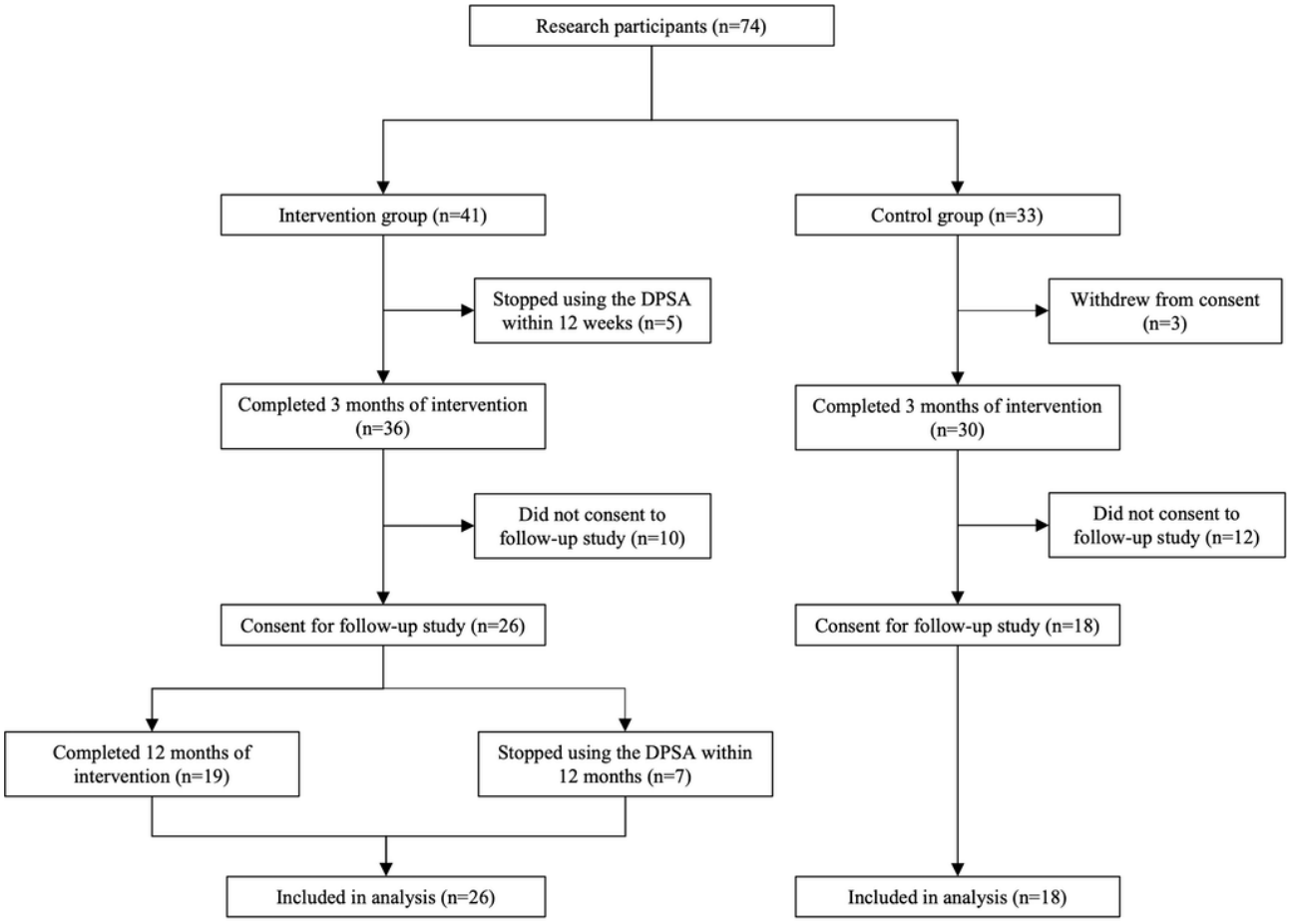
Data were analyzed using IBM SPSS Statistics 29 (IBM). The statistical significance level was set to 5%.

Results

Participants

The follow-up study included 44 of 66 participants from the 3-month intervention study. Of these, 26 were in the intervention group and 18 in the control group (Figure 3). The intervention group consisted of 7 groups of 3-4 people.

Figure 3. Flow diagram of participant enrollment and follow-up. DPSA: digital peer-supported app.



Participant Characteristics

Table 1 shows the baseline participant characteristics. The mean age (SD) of the participants in the intervention group was 75.1 (5.1) years, with 13/26 men (50%). In the control group, the mean age was 77.4 (SD 5.3) years, with 6/18 men (33%). Although the participants were relatively older (average age of 76.0 years), both groups consisted of active older adults who had regular exercise habits, engaged in active interactions with their neighbors, and had no health problems that would interfere with study participation. No statistically significant differences were observed in the baseline demographic characteristics between the intervention and control groups. Although statistically significant differences were not observed arguably owing to the small sample sizes, the intervention group had larger proportions of smartphone ownership and app use frequency.

XSL•FO
RenderX

Table . Participant characteristics.

Characteristics	Total sample (n=44)	Intervention group (n=26)	Control group (n=18)	*P* value
Age (years), mean (SD)	76.0 (5.2)	75.1 (5.1)	77.4 (5.3)	.15[a]
Sex, n (%)				.27[b]
Male	19 (43)	13 (50)	6 (33)	
Female	25 (57)	13 (50)	12 (66)	
BMI (kg/m^2), mean (SD)	22.9 (3.0)	23.3 (3.0)	22.2 (3.0)	.27[a]
Living alone, n (%)	11 (26)	8 (31)	3 (17)	.24[b]
Self-rated health, n (%)				.33[b]
Excellent, good, or normal	39 (89)	24 (92)	15 (83)	
Fair or poor	5 (11)	2 (8)	3 (17)	
Perceived household economic status, n (%)				.36[b]
Excellent, good, or normal	41 (93)	25 (96)	16 (89)	
Fair or poor	3 (7)	1 (3)	2 (7)	
Life satisfaction, n (%)				.52[b]
Excellent or good, or normal	38 (86)	22 (85)	16 (90)	
Fair or poor	7 (14)	4 (15)	3 (10)	
Working, n (%)	12 (28)	7 (27)	5 (28)	.61[b]
Smartphone owner, n (%)	41 (93)	26 (100.0)	15 (83)	.06[b]
Frequency of app use, n (%)				.09[b]
Usually or sometimes	34 (80)	24 (92)	11 (61)	
Rarely or never	9 (21)	2 (8)	7 (39)	
Exercise habits[c], n (%)	24 (55)	16 (62)	8 (44)	.26[b]
Frequency of neighborhood interaction, n (%)				.30[b]
≥3 times per week	20 (45)	13 (50)	7 (39)	
≤2 times per week	24 (55)	13 (50)	11 (61)	
Participation in group exercise, n (%)	19 (43)	10 (38)	9 (50)	.45[b]
History falls in the past year, n (%)	6 (14)	3 (12)	3 (17)	.48[b]
Effect of COVID-19 on decreased physical activity, n (%)				.62[b]
Greatly/slightly	31 (70)	19 (73)	12 (67)	
Not much/unchanged	13 (30)	7 (27)	6 (33)	
Self-reported decrease in walking speed, n (%)	31 (70)	19 (73)	12 (67)	.65[b]
Triaxial accelerometer				
Steps per day, median (IQR)	6849 (4187 - 8688)	7082 (4434 - 9866)	5276 (4062 - 7143)	.15[d]

XSL•FO
RenderX

Characteristics	Total sample (n=44)	Intervention group (n=26)	Control group (n=18)	*P* value
LPA[e] (minutes per day), mean (SD)	330.5 (89.5)	303 (72.1)	369.1 (98.8)	.03[a]
MVPA[f] (minutes per day), mean (SD)	51.4 (27.9)	57.7 (25.3)	42.7 (29.7)	.047[a]
SB[g] (minutes per day), mean (SD)	540.0 (113.4)	538.7 (85.8)	541.8 (146.2)	.82[a]
Triaxial accelerometer wearing time (minutes per day), mean (SD)	921.9 (115.6)	899.2 (66.6)	953.6 (157.8)	.18[a]
Physical function, mean (SD)				
Grip strength (kg)	26.4 (8.3)	28.1 (8.7)	24.2 (7.3)	.13[a]
CS-30[h]	20.1 (6.8)	20.4 (7.6)	19.6 (5.6)	.81[a]
Self-efficacy for exercise, mean (SD)	13.4 (3.4)	13.6 (3.2)	13.1 (3.2)	.47[a]

[a]Analysis was conducted using the independent samples *t* test (2-tailed).

[b]Analysis was conducted using the chi-squared test.

[c]Exercise habit was defined as exercising at least twice a week for ≥30 minutes each time for ≥1 year.

[d]Analysis was conducted using the Mann-Whitney *U* tests.

[e]LPA: light-intensity physical activity.

[f]MVPA: moderate to vigorous intensity physical activity.

[g]SB: sedentary behavior.

[h]CS-30: 30-second chair stand test.

Feasibility: Retention Rate, Number of Posts, and Negative Impact

The retention rate among the 26 participants in the intervention group was 96% (25/26) in the 6th month, 92% (24/26) in the 9th month, and 73% (19/26) in the 12th month. The retention rate, based on 41 participants, all of whom did not participate in the follow-up survey, was 61% (25/41) in the 6th month, 59% (24/41) in the 9th month, and 46% (19/41) in the 12th month; thus, this retention rate was <70% at the beginning of the follow-up study. The reasons for dropping out of DPSA were “contracted COVID-19 and stopped submitting” (n=1), “unknown cause” (n=2), and “after discussion in a group chat, everyone stopped using DPSA” (n=4).

The adherence rate and number of total posts per day among members of the intervention group are summarized in Table 2. The adherence for the DPSA was 85.9%. The total number of chats per person averaged 2.55 (SD 1.28) per day. Excluding dropouts, the adherence rate was 92.3%, with a total of 2.88 (SD 1.24) posts per day per person. Adherence rates were good among participants in the follow-up study. One group had all members drop out; all members of the group were male. Three cases of mild physical discomfort that did not interfere with daily life were reported, with 2 participants reporting knee pain and 1 reporting foot pain.

Table . Digital peer-supported app adherence rate and number of total posts per day among members of the intervention group.

	All participants (n=26)		Excluding dropout (n=19)	
Group	Adherence rate, n (%)	Total posts/person /day, mean (SD)	Adherence rate, n (%)	Total posts/person /day, mean (SD)
All	26 (85.9)	2.55 (1.28)	19 (92.3)	2.88 (1.24)
A	4 (86.3)	1.40 (0.33)	4 (86.3)	1.40 (0.33)
B	4 (95.4)	2.61 (0.77)	3 (99.1)	2.63 (0.94)
C	4 (75.0)	1.17 (0.41)	All dropouts	All dropouts
D	4 (98.6)	2.24 (0.58)	4 (98.6)	2.24 (0.58)
E	3 (84.7)	4.17 (0.22)	3 (84.7)	4.17 (0.22)
F	4 (85.4)	3.28 (0.63)	4 (85.4)	3.28 (0.63)
G	3 (75.9)	3.66 (2.10)	2 (99.6)	4.74 (1.38)

XSL•FO
RenderX

No privacy breaches were associated with app usage. There were 2 inquiries from participants, including account transfer after a smartphone model change (about 30 minutes) and uninstallation of the DPSA (about 20 minutes).

Changes in Physical Activity and Function and Self-Efficacy for Exercise

Table 3 shows the analysis results. In the group comparison of the linear mixed-effects model analyses of physical activity and function and self-efficacy for exercise, no differences were observed. However, a significant change was observed in step count over time only in the intervention group (P=.048), wherein we observed an increase of 1736 (β=1736, 95% CI 232-3241) steps per day compared with baseline. LPA and SB showed differences in the control group, but no significant difference was noted at any time point compared with baseline. Regarding the CS-30, there was a significant within-group difference in the increase in CS-30 scores for the intervention (P<.001) and control (P=.03) groups over the 12-month period. Additionally, the change in CS-30 scores between the baseline and 12-month assessments was 6.5 (β=6.5, 95% CI 3.8-9.1) times in the intervention group (P<.001) and 3.8 (β=3.8, 95% CI 0.6-7.1) times in the control group (P=.02). Regarding the self-efficacy for exercise, a significant change over time was observed only in the intervention group (P=.03), wherein an increase of 1.6 (β=1.6, 95% CI 0.2-3.1) points was observed after 12 months compared with baseline (P=.03).

XSL•FO
RenderX

Table . Included outcome measures at baseline and 3, 6, and 12 months with within-group and between-group comparisons (26 participants in the intervention group and 18 participants in the control group).

Outcome measures	Intervention			Control			Group × time[a]
	Comparison with baseline		Within-group changes	Comparison with baseline		Within-group changes	
	β (95% CI)	*P* value	*P* value	β (95% CI)	*P* value	*P* value	*P* value
Steps per day[b]			.048			.08	.25
Baseline	Reference						
3 months	960 (−505 to 2425)	.39		150 (−765 to 1065)	.99		
6 months	1213 (−231 to 2657)	.14		274 (−633 to 1181)	.99		
9 months	581 (−910 to 2072)	.99		−653 (−1566 to 260)	.28		
12 months	1736 (232 to 3241)	.02		147 (−803 to 1096)	.99		
LPA[b,c] (minutes per day)			.18			.044	.71
Baseline	Reference						
3 months	−6 (−35 to 23)	.99		−21 (−49 to 7)	.23		
6 months	16 (−13 to 44)	.67		3 (−24 to 31)	.99		
9 months	5 (−25 to 35)	.99		−16 (−44 to 12)	.57		
12 months	18 (−12 to 48)	.048		7 (−22 to 36)	.99		
MVPA[b,d] (minutes per day)			.07			.96	.28
Baseline	Reference						
3 months	14 (1 to 26)	.02		1 (−7 to 10)	.99		
6 months	8 (−4 to 21)	.32		0 (−8 to 8)	.99		
9 months	5 (−7 to 18)	.99		−1 (−9 to 7)	.99		
12 months	10 (−3 to 22)	.23		0 (−9 to 8)	.99		
SB[b,e] (minutes per day)			.16			.049	.62
Baseline	Reference						
3 months	−8 (−40 to 24)	.99		20 (−8 to 48)	.30		
6 months	−24 (−56 to 7)	.22		−4 (−31 to 24)	.99		
9 months	−10 (−43 to 22)	.99		17 (−11 to 45)	.47		
12 months	−28 (−61 to 5)	.13		−6 (−36 to 23)	.99		
Triaxial accelerometer wearing time (minutes per day)			.51			.20	.27
Baseline	Reference						
3 months	9 (−38 to 56)	.99		−26 (−70 to 19)	.59		
6 months	−8 (−54 to 39)	.99		−14 (−58 to 31)	.99		
9 months	−19 (−67 to 28)	.99		−23 (−68 to 21)	.73		
12 months	9 (−39 to 58)	.99		−42 (−87 to 4)	.09		

XSL•FO
RenderX

Outcome measures	Intervention			Control			Group × time[a]
	Comparison with baseline		Within-group changes	Comparison with baseline		Within-group changes	
	β (95% CI)	*P* value	*P* value	β (95% CI)	*P* value	*P* value	*P* value
Grip strength (kg)			.09			.15	.12
Baseline	Reference						
3 months	−0.8 (−2.1 to 0.5)	.31		−0.9 (−2.5 to 0.7)	.35		
12 months	−1.1 (−2.3 to 0.1)	.07		−6 (−36 to 23)	.95		
CS-30 (times)[b,f]			<.001			.03	.41
Baseline	Reference						
3 months	1.4 (−1.1 to 4.0)	.40		0.6 (−2.1 to 3.3)	.99		
12 months	6.5 (3.8 to 9.1)	<.001		3.8 (0.6 to 7.1)	.02		
Self-efficacy for exercise (points)			.03			.54	.53
Baseline	Reference						
3 months	1.1 (−0.2 to 2.5)	.12		0.3 (−0.8 to 1.5)	.99		
12 months	1.6 (0.2 to 3.1)	.02		0.7 (−0.8 to 2.1)	.58		

[a]Analyses were adjusted for age, sex, and frequency of app use (baseline).

[b]Triaxial accelerometer data were adjusted for wear time.

[c]LPA: light-intensity physical activity.

[d]MVPA: moderate to vigorous intensity physical activity.

[e]SB: sedentary behavior.

[f]CS-30: 30-second chair stand test.

Discussion

Principal Results

This study aimed to confirm the feasibility of a 12-month intervention using the DPSA to improve physical activity among older adults and to obtain preliminary estimates of its effects on physical activity, physical function, and self-efficacy for exercise. The retention rate in the intervention group (n=26) over the 12-month period was 73% (19/26). Considering the 41 participants who did not express interest in the follow-up surveys as the denominators, the retention rate was 46% (19/41). The adherence rate was 87.7%. This study obtained preliminary estimates of the effects of DPSA use on physical activity, physical function, and self-efficacy for exercise.

Comparison With Previous Studies

This is a rare study that examined the 12-month long-term feasibility and changes of an app intervention aimed at promoting physical activity in older adults. The 26 participants in the intervention group who used the DPSA had a 12-month retention rate of 73%. Including participants who did not indicate a desire for a follow-up survey, the 12-month retention rate was 46%. A previous study that used a smartphone app and smart band for weight loss, physical activity, and caloric intake in an overweight and obese population aged between 20 and 65 years for 12 months reported a 12-month retention rate of 68.4% (227/332) [65]. Moreover, a previous study of adults aged 30 - 60 years on physical activity and weight loss in Japan using a smartphone app focused on steps reported a 12-month retention rate of 95% (52/55) [66]. Compared with that reported by these previous studies, the 12-month retention rate was lower. The low retention rate might have been because the study included older adults who were less familiar with the app than younger adults [67], and daily posting may have been stressful for participants with the limited app experience. Only about half of the older adults were capable of long-term retention in the DPSA. One of the 7 groups had all its members drop out; this group was unique in that all members were males. Groups comprising a mix of male and female members may last longer. Group chat members with fewer posts, indicating lower engagement, tended to drop out. The following strategies can be adopted to prevent dropouts: providing opportunities for interactions among group members, encouraging people to make supportive posts to each other, and providing canned messages, such as greeting and appreciation messages, to allow group chat members communicate with each other through simple operations. Only 3 negative physical effects were reported; however, they were all minor and did not cause privacy issues.

In this study, significant changes in the number of steps taken and the self-efficacy for exercise score (Table 3) were observed

XSL•FO
RenderX

within the intervention group, but there was no significant difference between groups. Self-efficacy is an important aspect of social cognitive theory [23]. As hypothesized, peer support based on social cognitive theory improves self-efficacy for exercise, resulting in increased steps. In the intervention group, an increase of more than 1000 average daily steps was observed. Increasing the number of steps taken daily by ≥1000 reduces the risk of various diseases and mortality [68,69] . A systematic review of 17 prospective studies by Hall et al [68] showed that each 1000-step increase in the daily step count decreases the risk of death and heart disease, with a 6% - 36% decrease in all-cause mortality risk and a 5% - 21% decrease in heart disease risk. Furthermore, an increase of 1000 steps per day decreases a woman's risk of diabetes by 6% and an increase of 2000 steps per day decreases the risk of diabetes by 12% [69]. Although there was an increase in MVPA in the 3-month intervention, there were no significant differences within groups in this study. However, MVPA increased by 10 (95% CI −3 to 22) minutes per day at 12 months compared with baseline. This result may be an effect of the small sample size. Peer support can build trust and provide social support through interpersonal communication [70]. In peer-based intervention strategies aimed at promoting physical activity among older adults, social support is considered a key factor in facilitating behavior change [71]. In this study, social support through peer support may have influenced physical activity levels. However, the evaluation of social support provided by the DPSA was lacking and should be addressed in future research.

In this study, CS-30 scores changed from baseline to 12 months for both intervention and control groups (Table 3), but no significant differences were observed between groups. In the intervention group, long-term continuation of the DPSA may have increased self-efficacy for exercise and the number of steps taken, leading to improved lower limb function. The DPSA may be effective as a means to improve lower limb function. This is a meaningful result because improving lower limb function may lead to the prevention of falls [72,73], sarcopenia [9,74,75], frailty [74,76], and dementia [34,72,77], and may also lead to reductions in health care costs associated with these conditions [78]. The control group also improved lower extremity function with an increase in CS-30 scores. Older adults in the control group did not use the DPSA. They attended exercise instruction and continued regular physical activity monitoring with an accelerometer. The improvement in lower limb function may have been due to voluntary physical activity or strength training that could not be adequately measured with an activity meter. The DPSA is not for everyone, as it requires possession of a smartphone and an understanding of its usage. It may be important to select a menu of interventions that is tailored to the characteristics of the participants.

Limitations

This study has the following 4 limitations. First, the study design was less capable of demonstrating the effects of the DPSA, compared with an RCT. Participants were nonrandomly assigned to the intervention and control groups and free to choose the group in which they would participate. Older adults who did not own a smartphone were unable to participate in the intervention group, and those unfamiliar with the app's operation were less likely to join. Given that this was a non-RCT, fully accounting for possible confounding bias was challenging, making it difficult to accurately estimate the intervention's effect by comparing the 2 groups. Additionally, follow-up participation was voluntary. The use of the DPSA is applicable to only eligible older Fujisawa City residents who own smartphones and are interested in mobile apps.

Second, the sample size was small. The follow-up participation rate was 63% (26/41) in the intervention group and 55% (18/33) in the control group. This low participation rate reduces the study's validity and may have impacted the feasibility and estimates of the effects on physical activity, physical function, and self-efficacy for exercise. The small sample size might have resulted in insufficient statistical power to detect differences between groups, and the model parameter estimation might have been unstable. Furthermore, the small sample size did not allow the convergence of the mixed-effects model. The older adults in this study took more steps per day and were originally sufficiently physically active compared with the general healthy older adult population [79]. The mean baseline score for adults in Japan aged 60 years or older for CS-30 score was 17.3 times [80]. At baseline, the intervention group averaged 20.4 times and the control group averaged 19.6 times. The study participants originally had the adequate level of lower extremity function. Future studies may benefit an aging society by targeting many older adults who are less physically active and have poor lower extremity function.

Third, given that this study included only those who participated in the follow-up, survival bias may have been present. In the intervention group, participants who did not complete the follow-up study were older and engaged in less physical activity at baseline than those who did. Therefore, the feasibility findings and estimates of changes in physical activity, physical function, and self-efficacy for exercise may be overestimated.

Fourth, the generalizability of this study is limited owing to potential selection bias. Participants in the intervention group were not only motivated to increase physical activity but also familiar with using the app. The DPSA was not adaptable to all participants, as it required a certain level of information technology literacy and cognitive function. Social, cultural, and economic factors (eg, older age, privacy concerns, and low income) may influence preference and feasibility with smartphone apps [81-83]. Therefore, the use of DPSA may not be widely accepted.

Conclusions

This study assessed the 12-month feasibility of using the DPSA and measured preliminary estimates of its effects on physical activity, physical function, and self-efficacy for exercise. The 12-month retention rate for participants in the DPSA intervention group was 73% (19/26), and that for 41 participants including those who decided not to participate in the follow-up study was 46% (19/41). The DPSA adherence rate was 85.9%. Only a limited number of older adults had long-term access to the DPSA. Preliminary estimates suggest that DPSA use may improve step count, lower extremity function, and self-efficacy for exercise. However, various biases were introduced, preventing the demonstration of clear intervention effects. There

XSL•FO
RenderX

is a need to identify ways in which more older adults can use DPSAs for extended periods of time; RCTs should be conducted to ascertain the long-term effects of DPSAs on physical activity and function in older adults.

Acknowledgments

This study was supported by a Doctoral Student Grant-in-Aid Program from the Ushioda Memorial Fund. The Fujisawa City government, A10 Lab Inc., and Keio University independently funded their respective parts of the research. The authors thank the staff of the Elderly Support Division of the Fujisawa City government (Shinichi Sakae, Mayumi Taguchi, Mari Sakata, and Shizuka Kobayashi) for recruiting participants and managing the project, Ken Takahashi of the Fujisawa City Health and Medical Foundation for his cooperation as a health fitness instructor, and Oguma laboratory staff for their help in collecting the data.

Authors' Contributions

KT, YO, SY, and MS were responsible for designing this study. KT, YO, SY, MS, MN, ND, AH, and TM were responsible for the data collection. KT and YO were responsible for analyzing and interpreting the data. All authors have edited, reviewed, and approved the final manuscript.

Conflicts of Interest

SY and MS were employees of A10 Lab Inc at the time of the research.

Multimedia Appendix 1
Details of the digital peer-supported app's functionality.
[DOCX File, 70532 KB - aging_v8i1e66610_app1.docx]

References

1. Zhang D. Perceived neighborhood conditions, psychosocial factors, and sleep problems among urban and rural older adults in China. J Aging Health 2024 Jun;36(5-6):337-349. [doi: 10.1177/08982643231159709] [Medline: 37395560]
2. Lutz W, Sanderson W, Scherbov S. Doubling of world population unlikely. Nature New Biol 1997 Jun 19;387(6635):803-805. [doi: 10.1038/42935] [Medline: 9194559]
3. Lang PO, Govind S, Aspinall R. Reversing T cell immunosenescence: why, who, and how. Age (Dordr) 2013 Jun;35(3):609-620. [doi: 10.1007/s11357-012-9393-y] [Medline: 22367580]
4. Bouaziz W, Vogel T, Schmitt E, Kaltenbach G, Geny B, Lang PO. Health benefits of aerobic training programs in adults aged 70 and over: a systematic review. Arch Gerontol Geriatr 2017;69:110-127. [doi: 10.1016/j.archger.2016.10.012] [Medline: 27912156]
5. White paper on the aging society. Cabinet Office Japan. 2023. URL: https://www8.cao.go.jp/kourei/whitepaper/w-2024/zenbun/06pdf_index.html [accessed 2024-09-17]
6. Dementia. World Health Organization. 2022. URL: https://www.who.int/news-room/fact-sheets/detail/dementia#:~:text=Studies%20show%20that%20people%20can,cholesterol%20and%20blood%20sugar [accessed 2024-05-17]
7. Katzmarzyk PT, Friedenreich C, Shiroma EJ, Lee IM. Physical inactivity and non-communicable disease burden in low-income, middle-income and high-income countries. Br J Sports Med 2022 Jan;56(2):101-106. [doi: 10.1136/bjsports-2020-103640] [Medline: 33782046]
8. Manning KM, Hall KS, Sloane R, et al. Longitudinal analysis of physical function in older adults: the effects of physical inactivity and exercise training. Aging Cell 2024 Jan;23(1):e13987. [doi: 10.1111/acel.13987] [Medline: 37681737]
9. Wang H, Huang WY, Zhao Y. Efficacy of exercise on muscle function and physical performance in older adults with sarcopenia: an updated systematic review and meta-analysis. Int J Environ Res Public Health 2022 Jul 5;19(13):8212. [doi: 10.3390/ijerph19138212] [Medline: 35805870]
10. Lee IM, Shiroma EJ, Lobelo F, et al. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet 2012 Jul 21;380(9838):219-229. [doi: 10.1016/S0140-6736(12)61031-9] [Medline: 22818936]
11. Promotion of physical activity and exercise 2023. Ministry of Health, Labour and Welfare. 2024 Jan. URL: https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryou/kenkou/undou/index.html [accessed 2025-02-12]
12. Dipietro L, Campbell WW, Buchner DM, et al. Physical activity, injurious falls, and physical function in aging: an umbrella review. Med Sci Sports Exerc 2019 Jun;51(6):1303-1313. [doi: 10.1249/MSS.0000000000001942] [Medline: 31095087]
13. Katz S. Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living. J Am Geriatr Soc 1983 Dec;31(12):721-727. [doi: 10.1111/j.1532-5415.1983.tb03391.x] [Medline: 6418786]
14. Rejeski WJ, Ip EH, Marsh AP, Miller ME, Farmer DF. Measuring disability in older adults: the International Classification System of Functioning, Disability and Health (ICF) framework. Geriatr Gerontol Int 2008 Mar;8(1):48-54. [doi: 10.1111/j.1447-0594.2008.00446.x] [Medline: 18713189]

XSL•FO
RenderX

15. Fukuoka Y, Haskell W, Lin F, Vittinghoff E. Short- and long-term effects of a mobile phone app in conjunction with brief in-person counseling on physical activity among physically inactive women: the mPED randomized clinical trial. JAMA Netw Open 2019 May 3;2(5):e194281. [doi: 10.1001/jamanetworkopen.2019.4281] [Medline: 31125101]
16. Laranjo L, Ding D, Heleno B, et al. Do smartphone applications and activity trackers increase physical activity in adults? Systematic review, meta-analysis and metaregression. Br J Sports Med 2021 Apr;55(8):422-432. [doi: 10.1136/bjsports-2020-102892] [Medline: 33355160]
17. Moss RJ, Süle A, Kohl S. eHealth and mHealth. Eur J Hosp Pharm 2019 Jan;26(1):57-58. [doi: 10.1136/ejhpharm-2018-001819] [Medline: 31157099]
18. Elavsky S, Knapova L, Klocek A, Smahel D. Mobile health interventions for physical activity, sedentary behavior, and sleep in adults aged 50 years and older: a systematic literature review. J Aging Phys Act 2019 Aug 1;27(4):565-593. [doi: 10.1123/japa.2017-0410] [Medline: 30507266]
19. Kwan RYC, Salihu D, Lee PH, et al. The effect of e-health interventions promoting physical activity in older people: a systematic review and meta-analysis. Eur Rev Aging Phys Act 2020;17:7. [doi: 10.1186/s11556-020-00239-5] [Medline: 32336996]
20. D'Amore C, Reid JC, Chan M, et al. Interventions including smart technology compared with face-to-face physical activity interventions in older adults: systematic review and meta-analysis. J Med Internet Res 2022 Oct 31;24(10):e36134. [doi: 10.2196/36134] [Medline: 36315229]
21. Duan Y, Shang B, Liang W, Du G, Yang M, Rhodes RE. Effects of eHealth-based multiple health behavior change interventions on physical activity, healthy diet, and weight in people with noncommunicable diseases: systematic review and meta-analysis. J Med Internet Res 2021 Feb 22;23(2):e23786. [doi: 10.2196/23786] [Medline: 33616534]
22. Morey MC, Snyder DC, Sloane R, et al. Effects of home-based diet and exercise on functional outcomes among older, overweight long-term cancer survivors: RENEW: a randomized controlled trial. JAMA 2009 May 13;301(18):1883-1891. [doi: 10.1001/jama.2009.643] [Medline: 19436015]
23. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory: International Bibliography of the Social Sciences (IBSS); 1986.
24. McCullagh P, Weiss MR, Ross D. Modeling considerations in motor skill acquisition and performance: an integrated approach. Exerc Sport Sci Rev 1989;17:475-513. [Medline: 2676555]
25. Posavac EJ, Kattapong KR, Dew DE Jr. Peer-based interventions to influence health-related behaviors and attitudes: a meta-analysis. Psychol Rep 1999 Dec;85(3 Pt 2):1179-1194. [doi: 10.2466/pr0.1999.85.3f.1179] [Medline: 10710973]
26. Ashford D, Bennett SJ, Davids K. Observational modeling effects for movement dynamics and movement outcome measures across differing task constraints: a meta-analysis. J Mot Behav 2006 May;38(3):185-205. [doi: 10.3200/JMBR.38.3.185-205] [Medline: 16709559]
27. McAuley E, Blissmer B. Self-efficacy determinants and consequences of physical activity. Exerc Sport Sci Rev 2000 Apr;28(2):85-88. [Medline: 10902091]
28. Ginis KAM, Nigg CR, Smith AL. Peer-delivered physical activity interventions: an overlooked opportunity for physical activity promotion. Transl Behav Med 2013 Dec;3(4):434-443. [doi: 10.1007/s13142-013-0215-2] [Medline: 24294332]
29. Webel AR, Okonsky J, Trompeta J, Holzemer WL. A systematic review of the effectiveness of peer-based interventions on health-related behaviors in adults. Am J Public Health 2010 Feb;100(2):247-253. [doi: 10.2105/AJPH.2008.149419] [Medline: 20019321]
30. Tabira K, Oguma Y, Yoshihara S, et al. Digital peer-supported app intervention to promote physical activity among community-dwelling older adults: nonrandomized controlled trial. JMIR Aging 2024 May 30;7:e56184. [doi: 10.2196/56184] [Medline: 38814686]
31. van der Meulen K, Granizo L, Del Barrio C. Emotional peer support interventions for students with SEND: a systematic review. Front Psychol 2021;12:797913. [doi: 10.3389/fpsyg.2021.797913] [Medline: 35027905]
32. Burton E, Farrier K, Hill KD, Codde J, Airey P, Hill AM. Effectiveness of peers in delivering programs or motivating older people to increase their participation in physical activity: systematic review and meta-analysis. J Sports Sci 2018 Mar;36(6):666-678. [doi: 10.1080/02640414.2017.1329549] [Medline: 28535358]
33. Iliffe S, Kendrick D, Morris R, et al. Multi-centre cluster randomised trial comparing a community group exercise programme with home based exercise with usual care for people aged 65 and over in primary care: protocol of the ProAct 65+ trial. Trials 2010 Jan 18;11:6. [doi: 10.1186/1745-6215-11-6] [Medline: 20082696]
34. McGarrigle L, Todd C. Promotion of physical activity in older people using mHealth and eHealth technologies: rapid review of reviews. J Med Internet Res 2020 Dec 29;22(12):e22201. [doi: 10.2196/22201] [Medline: 33372894]
35. Royse LA, Baker BS, Warne-Griggs MD, et al. "It's not time for us to sit down yet": how group exercise programs can motivate physical activity and overcome barriers in inactive older adults. Int J Qual Stud Health Well-being 2023 Dec;18(1):2216034. [doi: 10.1080/17482631.2023.2216034] [Medline: 37220010]
36. Reiner M, Niermann C, Jekauc D, Woll A. Long-term health benefits of physical activity—a systematic review of longitudinal studies. BMC Public Health 2013 Sep 8;13:813. [doi: 10.1186/1471-2458-13-813] [Medline: 24010994]

XSL•FO
RenderX

37. Jonkman NH, van Schooten KS, Maier AB, Pijnappels M. eHealth interventions to promote objectively measured physical activity in community-dwelling older people. Maturitas 2018 Jul;113:32-39. [doi: 10.1016/j.maturitas.2018.04.010] [Medline: 29903646]
38. Yerrakalva D, Yerrakalva D, Hajna S, Griffin S. Effects of mobile health app interventions on sedentary time, physical activity, and fitness in older adults: systematic review and meta-analysis. J Med Internet Res 2019 Nov 28;21(11):e14343. [doi: 10.2196/14343] [Medline: 31778121]
39. Mönninghoff A, Kramer JN, Hess AJ, et al. Long-term effectiveness of mHealth physical activity interventions: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2021 Apr 30;23(4):e26699. [doi: 10.2196/26699] [Medline: 33811021]
40. Stecher C, Pfisterer B, Harden SM, et al. Assessing the pragmatic nature of mobile health interventions promoting physical activity: systematic review and meta-analysis. JMIR Mhealth Uhealth 2023 May 4;11:e43162. [doi: 10.2196/43162] [Medline: 37140972]
41. Nyman SR, Victor CR. Older people's participation in and engagement with falls prevention interventions in community settings: an augment to the Cochrane systematic review. Age Ageing 2012 Jan;41(1):16-23. [doi: 10.1093/ageing/afr103]
42. Olanrewaju O, Kelly S, Cowan A, Brayne C, Lafortune L. Physical activity in community dwelling older people: a systematic review of reviews of interventions and context. PLoS One 2016;11(12):e0168614. [doi: 10.1371/journal.pone.0168614] [Medline: 27997604]
43. Fujisawa City future population estimates. Fujisawa City. 2023 Apr. URL: https://www.city.fujisawa.kanagawa.jp/documents/30105/fujisawasi-syourai-jinnkou-suikei-2022.pdf [accessed 2025-02-12]
44. Landry R, Traore N, Godin B. An econometric analysis of the effect of collaboration on academic research productivity. High Educ 1996 Oct;32(3):283-301. [doi: 10.1007/BF00138868]
45. Dooley L, Kirk D. University - industry collaboration. Eur J Innov Manage 2007 Aug 7;10(3):316-332. [doi: 10.1108/14601060710776734]
46. Thomas S, Reading J, Shephard RJ. Revision of the Physical Activity Readiness Questionnaire (PAR-Q). Can J Sport Sci 1992 Dec;17(4):338-345. [Medline: 1330274]
47. Bredin SSD, Gledhill N, Jamnik VK, Warburton DER. PAR-Q+ and ePARmed-X+: new risk stratification and physical activity clearance strategy for physicians and patients alike. Can Fam Physician 2013 Mar;59(3):273-277. [Medline: 23486800]
48. Active guide: Japanese official physical activity guidelines for health promotion brochure in English. Ministry of Health, Labour and Welfare. 2013. URL: https://www.nibn.go.jp/eiken/programs/pdf/active2013-e.pdf [accessed 2025-12-02]
49. Fujisawa +10 exercise. Fujisawa City Health and Medical Center. 2014. URL: http://iryo.city.fujisawa.kanagawa.jp/health/health01_01.htm#02 [accessed 2025-01-17]
50. Osawa Y, Saito Y, Tsunekawa N, Manabe T, Oguma Y. Exercise workload of the "Fujisawa +10 exercise" program in older women. J Exer Physiol Online 2015;18:79-85.
51. Minchalle Web. A10Lab. Inc. 2023. URL: https://www.mhlw.go.jp/content/10904750/000999445.pdfhttps://minchalle.com/ [accessed 2025-02-12]
52. Ministry of Health, Labour and Welfare. National health and nutrition survey. 2017. URL: https://www.mhlw.go.jp/content/10904750/000351576.pdf [accessed 2025-02-12]
53. Farrance C, Tsofliou F, Clark C. Adherence to community based group exercise interventions for older people: a mixed-methods systematic review. Prev Med 2016 Jun;87:155-166. [doi: 10.1016/j.ypmed.2016.02.037] [Medline: 26921655]
54. Picorelli AMA, Pereira LSM, Pereira DS, Felício D, Sherrington C. Adherence to exercise programs for older people is influenced by program characteristics and personal factors: a systematic review. J Physiother 2014 Sep;60(3):151-156. [doi: 10.1016/j.jphys.2014.06.012] [Medline: 25092418]
55. Ohkawara K, Oshima Y, Hikihara Y, Ishikawa-Takata K, Tabata I, Tanaka S. Real-time estimation of daily physical activity intensity by a triaxial accelerometer and a gravity-removal classification algorithm. Br J Nutr 2011 Jun;105(11):1681-1691. [doi: 10.1017/S0007114510005441] [Medline: 21262061]
56. Park J, Ishikawa-Takata K, Tanaka S, Bessyo K, Tanaka S, Kimura T. Accuracy of estimating step counts and intensity using accelerometers in older people with or without assistive devices. J Aging Phys Act 2017 Jan;25(1):41-50. [doi: 10.1123/japa.2015-0201] [Medline: 27180730]
57. Jefferis BJ, Sartini C, Ash S, et al. Trajectories of objectively measured physical activity in free-living older men. Med Sci Sports Exerc 2015 Feb;47(2):343-349. [doi: 10.1249/MSS.0000000000000410] [Medline: 24988411]
58. Savas S, Kilavuz A, Kayhan Koçak F, Cavdar S. Comparison of grip strength measurements by widely used three dynamometers in outpatients aged 60 years and over. J Clin Med 2023 Jun 25;12(13):4260. [doi: 10.3390/jcm12134260] [Medline: 37445293]
59. Jones CJ, Rikli RE, Beam WC. A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. Res Q Exerc Sport 1999 Jun;70(2):113-119. [doi: 10.1080/02701367.1999.10608028] [Medline: 10380242]
60. Oka K. Stages of change for exercise behavior and self-efficacy for exercise among middle-aged adults. Nihon Koshu Eisei Zasshi 2003 Mar;50(3):208-215. [Medline: 12704833]

XSL•FO
RenderX

61. Li Y, Baron J. Linear mixed-effects models in analyzing repeated-measures data. In: Li Y, Baron J, editors. Behavioral Research Data Analysis With R: Springer; 2012:177-204. [doi: 10.1007/978-1-4614-1238-0_10]
62. Raudenbush SW, Xiao-Feng L. Effects of study duration, frequency of observation, and sample size on power in studies of group differences in polynomial change. Psychol Methods 2001 Dec;6(4):387-401. [doi: 10.1037/1082-989X.6.4.387] [Medline: 11778679]
63. Muth C, Bales KL, Hinde K, Maninger N, Mendoza SP, Ferrer E. Alternative models for small samples in psychological research: applying linear mixed effects models and generalized estimating equations to repeated measures data. Educ Psychol Meas 2016 Feb;76(1):64-87. [doi: 10.1177/0013164415580432] [Medline: 29795857]
64. Wiley RW, Rapp B. Statistical analysis in Small-N Designs: using linear mixed-effects modeling for evaluating intervention effectiveness. Aphasiology 2019;33(1):1-30. [doi: 10.1080/02687038.2018.1454884] [Medline: 33012945]
65. Lugones-Sanchez C, Recio-Rodriguez JI, Agudo-Conde C, et al. Long-term effectiveness of a smartphone app combined with a smart band on weight loss, physical activity, and caloric intake in a population with overweight and obesity (evident 3 study): randomized controlled trial. J Med Internet Res 2022 Feb 1;24(2):e30416. [doi: 10.2196/30416] [Medline: 35103609]
66. Yoshimura E, Tajiri E, Michiwaki R, Matsumoto N, Hatamoto Y, Tanaka S. Long-term effects of the use of a step count-specific smartphone app on physical activity and weight loss: randomized controlled clinical trial. JMIR Mhealth Uhealth 2022 Oct 24;10(10):e35628. [doi: 10.2196/35628] [Medline: 36279159]
67. Mubarak F, Suomi R. Elderly forgotten? Digital exclusion in the information age and the rising grey digital divide. Inquiry 2022;59. [doi: 10.1177/00469580221096272] [Medline: 35471138]
68. Hall KS, Hyde ET, Bassett DR, et al. Systematic review of the prospective association of daily step counts with risk of mortality, cardiovascular disease, and dysglycemia. Int J Behav Nutr Phys Act 2020 Jun 20;17(1):78. [doi: 10.1186/s12966-020-00978-9] [Medline: 32563261]
69. Garduno AC, LaCroix AZ, LaMonte MJ, et al. Associations of daily steps and step intensity with incident diabetes in a prospective cohort study of older women: the OPACH Study. Diabetes Care 2022 Feb 1;45(2):339-347. [doi: 10.2337/dc21-1202] [Medline: 35050362]
70. Matz-Costa C, Howard EP, Castaneda-Sceppa C, Diaz-Valdes Iriarte A, Lachman ME. Peer-based strategies to support physical activity interventions for older adults: a typology, conceptual framework, and practice guidelines. Gerontologist 2019 Nov 16;59(6):1007-1016. [doi: 10.1093/geront/gny092] [Medline: 30085074]
71. Mead S, Hilton D, Curtis L. Peer support: a theoretical perspective. Psychiatr Rehabil J 2001;25(2):134-141. [doi: 10.1037/h0095032] [Medline: 11769979]
72. Ikegami S, Takahashi J, Uehara M, et al. Physical performance reflects cognitive function, fall risk, and quality of life in community-dwelling older people. Sci Rep 2019 Aug 22;9(1):12242. [doi: 10.1038/s41598-019-48793-y] [Medline: 31439872]
73. Sanchez M, Vidal JS, Bichon A, et al. Impact of a public open-access community-based physical activity and fall prevention program on physical performance in older adults. Eur J Public Health 2023 Feb 3;33(1):132-138. [doi: 10.1093/eurpub/ckac186] [Medline: 36594685]
74. Buchman AS, Leurgans SE, Wang T, et al. Motor function is the primary driver of the associations of sarcopenia and physical frailty with adverse health outcomes in community-dwelling older adults. PLoS One 2021;16(2):e0245680. [doi: 10.1371/journal.pone.0245680] [Medline: 33529220]
75. Sawada S, Ozaki H, Natsume T, et al. The 30-s chair stand test can be a useful tool for screening sarcopenia in elderly Japanese participants. BMC Musculoskelet Disord 2021 Jul 24;22(1):639. [doi: 10.1186/s12891-021-04524-x] [Medline: 34303339]
76. Pilotto A, Custodero C, Maggi S, Polidori MC, Veronese N, Ferrucci L. A multidimensional approach to frailty in older people. Ageing Res Rev 2020 Jul;60:101047. [doi: 10.1016/j.arr.2020.101047] [Medline: 32171786]
77. Doi T, Tsutsumimoto K, Nakakubo S, et al. Physical performance predictors for incident dementia among Japanese community-dwelling older adults. Phys Ther 2019 Sep 1;99(9):1132-1140. [doi: 10.1093/ptj/pzz077] [Medline: 31162614]
78. Arai H, Ouchi Y, Toba K, et al. Japan as the front-runner of super-aged societies: perspectives from medicine and medical care in Japan. Geriatr Gerontol Int 2015 Jun;15(6):673-687. [doi: 10.1111/ggi.12450] [Medline: 25656311]
79. Tudor-Locke C, Craig CL, Aoyagi Y, et al. How many steps/day are enough? For older adults and special populations. Int J Behav Nutr Phys Act 2011 Jul 28;8(1):80. [doi: 10.1186/1479-5868-8-80] [Medline: 21798044]
80. Nakazono T, Kamide N, Ando M. The reference values for the chair stand test in healthy Japanese older people: determination by meta-analysis. J Phys Ther Sci 2014 Nov;26(11):1729-1731. [doi: 10.1589/jpts.26.1729] [Medline: 25435687]
81. Hoehle H, Zhang X, Venkatesh V. An espoused cultural perspective to understand continued intention to use mobile applications: a four-country study of mobile social media application usability. Eur J Inf Syst 2015 May;24(3):337-359. [doi: 10.1057/ejis.2014.43]
82. Gera R, Chadha P, Ahuja V. Mobile app usage and adoption: a literature review. Int J Electron Bus 2020;15(2):160. [doi: 10.1504/IJEB.2020.106546]
83. Wang C, Qi H. Influencing factors of acceptance and use behavior of mobile health application users: systematic review. Healthcare (Basel) 2021 Mar 22;9(3):357. [doi: 10.3390/healthcare9030357] [Medline: 33809828]

XSL•FO
RenderX

Abbreviations

BMI: body mass index
CS-30: 30-second chair stand test
DPSA: digital peer-supported app
IQR: interquartile range
LPA: light-intensity physical activity
MET: metabolic equivalent
MVPA: moderate to vigorous intensity physical activity
RCT: randomized controlled trial
SB: sedentary behavior

Edited by Q Fan; submitted 26.09.24; peer-reviewed by J Smith-Turchyn, W Liang; revised version received 13.04.25; accepted 14.04.25; published 28.05.25.

Please cite as:
Tabira K, Oguma Y, Yoshihara S, Shibuya M, Nakamura M, Doihara N, Hirata A, Manabe T, Yamashita T
A 12-Month Digital Peer-Supported App Intervention to Promote Physical Activity Among Community-Dwelling Older Adults: Follow-Up Study of a Nonrandomized Controlled Trial
JMIR Aging 2025;8:e66610
URL: https://aging.jmir.org/2025/1/e66610
doi:10.2196/66610

© Kento Tabira, Yuko Oguma, Shota Yoshihara, Megumi Shibuya, Manabu Nakamura, Natsue Doihara, Akihiro Hirata, Tomoki Manabe, Takashi Yamashita. Originally published in JMIR Aging (https://aging.jmir.org), 28.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Impact of a Light Volleyball Intervention Program on Improving Physical Attributes of Older Adults in Hong Kong: Preliminary Study of a Randomized Controlled Trial

Ka Man Leung[1], BSci, MSocSci, MSci, PhD; Yuchen Shi[2], PhD

[1]Department of Health and Physical Education, Faculty of Liberal Arts and Social Science, Education University of Hong Kong, Tai Po, China (Hong Kong)

[2]School of Liberal Arts, Macau University of Science and Technology, Avenida Wai Long, Taipa, Macau, China

Corresponding Author:
Yuchen Shi, PhD
School of Liberal Arts, Macau University of Science and Technology, Avenida Wai Long, Taipa, Macau, China

Abstract

Background: Physical inactivity, which increases the risk of chronic diseases in older adults, is prevalent among older adults in Hong Kong. To address this problem, the Hong Kong government has been proactively promoting active aging.

Objective: Following the World Health Organization's strategy to prevent chronic diseases in older adults and aligning with the global goal of active aging, this study evaluated the effects of a 16-week light volleyball (LVB) intervention program on the physical health of older adults in Hong Kong.

Methods: A total of 276 participants aged ≥60 years were recruited and randomly assigned to 1 of 3 groups: an LVB intervention group, a Taichi control group (ie, with light physical activity), and a control group. Tests on components of fitness were conducted before and after the intervention.

Results: Participants from the LVB intervention group exhibited significant improvements in lower body strength ($F_{2,272}$= 7.23, P=.001, η^2=.05), agility ($F_{2,272}$= 6.05, P=.003, η^2=.043), and dynamic balance ($F_{2,272}$= 9.41, P=.001, η^2=.065) when compared with those from the Taichi active control group and control group.

Conclusions: To promote active aging among older adults in Hong Kong, the findings of this preliminary study, along with forthcoming follow-up tests, will provide health specialists and practitioners with valuable insights regarding the health benefits of the LVB community program for older adults.

Trial Registration: Chinese Clinical Trial Register ChiCTR1900026657; https://www.chictr.org.cn/showprojEN.html?proj=44350

(JMIR Aging 2025;8:e62886) doi:10.2196/62886

KEYWORDS

adapted physical activity; older adults; gerontology; geriatrics; older; aging; randomized controlled trial; volleyball; light volleyball; intervention; sports; physical activity; exercise; physical attributes; RCT; controlled trial

Introduction

Given the low birth rate and increasing life expectancy in Hong Kong, the trend toward an aging population is expected to persist in the coming decades in this region. The number of individuals aged ≥65 years is projected to rise substantially, reaching 1.89 million by 2019 and 30.5 million by 2069 [1]. This continual demographic shift highlights several societal challenges, including a decline in the working population and an increased burden on Hong Kong's social welfare and health care systems. With an expanding share of older adults in the population, public health expenditure in Hong Kong is forecasted to increase by 394%, from US $4.9 billion in 2004 to US $24.4 billion in 2033, accounting for 5.5% of Hong Kong's total GDP in 2033 [2]. The simultaneous decrease in the working population and increase in social welfare expenditures for the aging population are expected to create severe financial strain for the Hong Kong government in the coming decades.

Although the benefits of physical activity have been well-documented [3], approximately 40% of older adults aged 60 years or older in Hong Kong engage in insufficient physical activity [4]. More than 50% of older adults in Hong Kong have overweight or obesity [5], and 75% have one or more chronic illnesses [6]. Additionally, older adults in Hong Kong spend an average of 7.5 hours sitting and 6.7 hours lying down daily [7]. Spiteri et al [8] identified various barriers to physical activity for older adults, including perceptions of negative consequences (ie, pain, risk of injury, and fear of falling), and a lack of knowledge and skills. Similar barriers to physical activity have been reported among older adults in Hong Kong. To address

XSL•FO
RenderX

these challenges, promoting active aging through physical activity is essential [9]. The World Health Organization (WHO) [10] introduced the "Active Ageing: A Policy Framework," warning that failure to address the aging population crisis could lead to the collapse of health care and social welfare systems. The Hong Kong government interprets active aging as achieving full physical, social, and mental well-being [6]. Since 2001, it has developed a healthy aging report outlining strategies to promote active aging, including physical activity initiatives. For instance, the Hong Kong government installed 150 new fitness equipment sets for older adults across Hong Kong between 2017 and 2018 to encourage active aging [11].

Meta-analyses have demonstrated that physical activity interventions have a significant effect on improving balance, reducing falls [12], and enhancing physical function [13] in older adults. Both supervised resistance and aerobic physical activity interventions have demonstrated positive effects on physical function outcomes. Future research directions include (1) increasing the number of physical activity interventions conducted in Asia, (2) aligning intervention programs with the WHO's recommendation of 150 minutes of physical activity per week, (3) using more rigorous randomized controlled trial designs, and (4) implementing supervised physical activity interventions with larger sample sizes in Asia [14,15]. Recent reviews have emphasized the need to examine the effects of sports interventions on older adults' health, particularly among Chinese older adults. Wong et al [16] reviewed 371 studies from the past 15 years on the effects of physical activity on older adults' physical health, and they recommended exploring the health benefits of newly emerging sports, such as light volleyball, in future research [16,17]. Franco et al [17] qualitatively reviewed the literature on older adults' perceptions of physical activity participation, suggesting that older adults prefer physical activity programs that are (1) professionally instructed, (2) group-based and conducive for peer interaction, and (3) highly accessible, with affordability being a key factor influencing the willingness to participate [17]. Similar findings were reported by Van Dyck et al [18], who highlighted older adults' preferences for innovative physical activity activities, such as aqua fitness or light volleyball, over traditional options, such as walking and cycling, in intervention programs.

Light volleyball (LVB) is a newly adapted physical activity derived from traditional volleyball. Compared with traditional volleyball, LVB uses balls with a larger circumference (80 - 83 cm vs 65 - 67 cm) and lighter weight (150g vs 250g), allowing the balls to travel at a slower speed. These features make LVB more accessible, particularly for individuals with reduced physical capacity due to age-related decline (eg, older adults). Additionally, the LVB court is smaller (similar in size to a badminton court), and the net is set lower (1.8 m), further reducing the physical demands on participants. Studies from China have reported physical health benefits associated with regular LVB practice; however, most of these studies did not have standardized fitness measurements and control groups. In 2020, the first author and her team conducted a quasi-experimental intervention to examine the effects of LVB on the physical and psychological health attributes of 78 older adults aged ≥60 years. The results indicated significant improvements in physical (ie, lower and upper body muscle strength, agility, balance, and aerobic endurance) and psychological (ie, physical activity enjoyment) attributes among participants in the LVB group when compared with the control group [19]. Furthermore, improvements in upper body muscle strength, aerobic endurance, and physical activity enjoyment were significantly more pronounced in the LVB group than in the active control group, which participated in Rouliqiu. That study highlighted the effectiveness of LVB intervention programs in enhancing both physical and psychological health attributes in older adults. Despite these promising findings, that study had notable limitations. First, its sample size was small, with only 62 participants completing the screening tests, pre- and postintervention functional tests, and data analysis. Second, participants were not randomly assigned to groups, which could have introduced selection bias.

Building on the positive results of the LVB pilot study and the prioritization of resource allocation for promoting active aging in Hong Kong, the first author and her team secured Research Impact funding amounting to US $0.95 million for further research in this area. This funding supported an investigation into the effectiveness of an LVB intervention on the physical and psychological health attributes of older adults in Hong Kong using both quantitative and qualitative methods and examining a larger sample size of approximately 300 participants. In this study, we present the preliminary results of this LVB intervention in our quantitative arm using the results from the pretest and posttest. In the current study, Taichi was selected as the intervention for the active control group because both Taichi and LVB are whole-body exercises originating from China and suitable for older adults [20,21]. Compared with the team-based LVB, Taichi is an individual exercise, which may lead to differing effects on older adults' health and quality of life [22]. Prior research suggests that older adults with greater social support are more likely to continue exercising regularly [23,24]. Community-based group physical activity interventions with increased social support have also been associated with greater beneficial effects and program adherence. Additionally, Taichi is a popular activity among Chinese older adults [25]. With LVB gaining traction in Hong Kong, Taichi serves as a relevant comparison, providing valuable insights into physical activity promotion among Hong Kong's older adult population. This study aimed to assess the effects of a 16-week LVB intervention compared with Taichi and a control group on two key outcomes in Chinese older adults aged ≥65 years: (1) functional fitness and (2) balance. The hypothesis was that both LVB and Taichi would result in significant and comparable improvements in these physical health attributes relative to the control group.

Methods

Study Design

A randomized controlled trial design was used to assess the effects of the LVB intervention on participants' physical health outcomes. Following the CONSORT (Consolidated Standards of Reporting Trials) guidelines (Checklist 1) [26], detailed information about the intervention can be found in our

XSL•FO
RenderX

previously published study protocol (trial registration number ChiCTR1900026657) [19]. The LVB group was compared against an active control group (ie, Taichi) and a control group. Participants were randomly assigned into LVB group, Taichi group, and control group in 1:1:1 ratio.

Study Intervention

The intervention program was conducted from mid-2021 to early-2023, with data collection performed at pretest (immediately before the intervention) and posttest (immediately after the intervention). The preliminary results presented here are based on pretest and posttest data only. Participants in the LVB and Taichi groups attended a 16-week training program consisting of two 90-minute sessions per week following the pretest. This intervention duration aligns with the guidelines of the Centers for Medicare and Medicaid Services of the US Department of Health and Human Services, which suggest that the health benefits of physical activity for older adults can be observed within 1 to 3 months after the start of a program [27]. In contrast, the control group was instructed to continue their usual daily activities while participating in monthly social gatherings (eg, health workshops) to control for psychosocial effects.

Participants

Participants were eligible for recruitment if they (1) were aged ≥60 years, (2) lived independently, (3) had no cognitive impairment, (4) had not participated in physical activity programs for 2 consecutive years prior to the program, and (5) had passing scores on the Abbreviated Mental Test (AMT) and Timed-up-and-go (TUG) test [28,29]. Specifically, individuals had to obtain a score of at least 6 out of 10 on the AMT to demonstrate sufficient cognitive capability and complete the TUG test within 20 seconds. Participants were ineligible for recruitment if they had steady hypertension (≥160/90 mmHg), arthritis, or neurological disorders.

Recruitment and Procedures

Participants were recruited through informational sessions conducted by the research team and advertisements placed in local neighborhood elderly centers. All participants were informed about the confidentiality of their personal data collected for the study and were assured they could withdraw from the program at any time. The research team sought consent from each participant before collecting data and proceeded immediately to conduct the AMT and TUG test to screen unqualified participants. After participants passed the screening tests, they completed the following steps: (1) questionnaires, (2) sociodemographic questions, (c) measurements of weight and body fat percentage using the Tanita machine (TBF-410GS), and (d) a functional fitness test based on the work of Leung et al [19,20]. After the pretest, an independent researcher used a computer-generated random number system to assign participants to the intervention conditions. Data collection and entry personnel were blinded to the group assignments of participants. The intervention program commenced the following week in sports complexes at community centers or community halls in Hong Kong and lasted for 16 weeks (approximately 3.5 months). The research team arranged the posttest for participants within 7 days after they completed the intervention program. All participants received a supermarket cash voucher worth US$ 12.84 as an incentive for their participation.

Ethical Considerations

The study was reviewed and approved by the Education University of Hong Kong's Research Ethics Committee (approval number E2022-2023-0013).

Measures

Functional Fitness

The research team used the Senior Fitness Test Manual to assess the physical attributes of older adults [30]. The test comprises 6 items: the chair stand test (lower body strength), arm curl test (upper body strength), chair sit-and-reach test (lower body flexibility), back-scratch test (upper body flexibility), 8-foot up-and-go test (agility and balance), and 2-minute step test (aerobic endurance). These have been demonstrated to be reliable, with intraclass correlation coefficients ranging between .80 and .98 in participant trials. Their validity has been supported through content, criterion-related analyses, and construct validation, including comparisons of senior fitness test scores with other established measures, such as treadmill VO2 testing [31]. Higher scores in the chair stand test (repetitions), arm curl test (repetitions), chair sit-and-reach test (cm), back-scratch test (cm), and 2-minute step test (repetitions) indicate higher levels of lower and upper body strength, flexibility, and aerobic endurance; the converse is the case for the 8-foot up-and-go test (seconds). For the flexibility assessments, negative scores reflect an inability to reach the toes during the chair sit-and-reach test or to make hand contact in the back-scratch test.

Balance Test

The Balance System SD (BBS-SD, 950 - 441 model) was used to measure the dynamic balance of older adults in the current study. Participants were asked to stand upright at the center of the platform while observing a screen situated 30 cm (approximately 11.1 in) in front of them. They completed three 20-second trials with 10-second breaks in between. The results of 3 trials were collected, and the mean values were recorded by the research team. The findings of previous studies support the reliability and validity of this balance test [32,33]. The measurement index and overall stability index generated by the Biodex balance system were included in the current study. These indices measured the participants' dynamic balance, specifically assessing their balance fluctuations across multiple axes. Higher values indicated greater deviations and poorer balance control.

Data Analysis

Data were analyzed using SPSS 27.0 software (IBM Corp). Descriptive statistics, specifically mean and SD values for continuous variables and frequencies and percentages for categorical variables, were used to describe the data. Preliminary checks were conducted to ensure that the assumptions of normality and homogeneity of variance were met. One-way ANOVA was performed to evaluate baseline differences between the groups. To examine the effect of the 16-week intervention program on physical attributes, a series of analysis of covariance tests were conducted, comparing the 3 groups

XSL•FO
RenderX

(LVB, Taichi, and control groups) at 2 time points (pre- and posttest). Body mass was used as a covariate because of its correlation with the outcome measures [34]. Partial eta squared (η^2), *P* values, and contrast *t*-statistics were calculated. Planned contrasts were used to analyze the differences between the LVB and control groups, and between the LVB and Taichi groups. Cohen *d* was calculated as a measure of effect size, with thresholds of 0.1 for a small effect, 0.3 for a medium effect, and 0.5 for a large effect [35]. Statistical significance was set at $P<.05$.

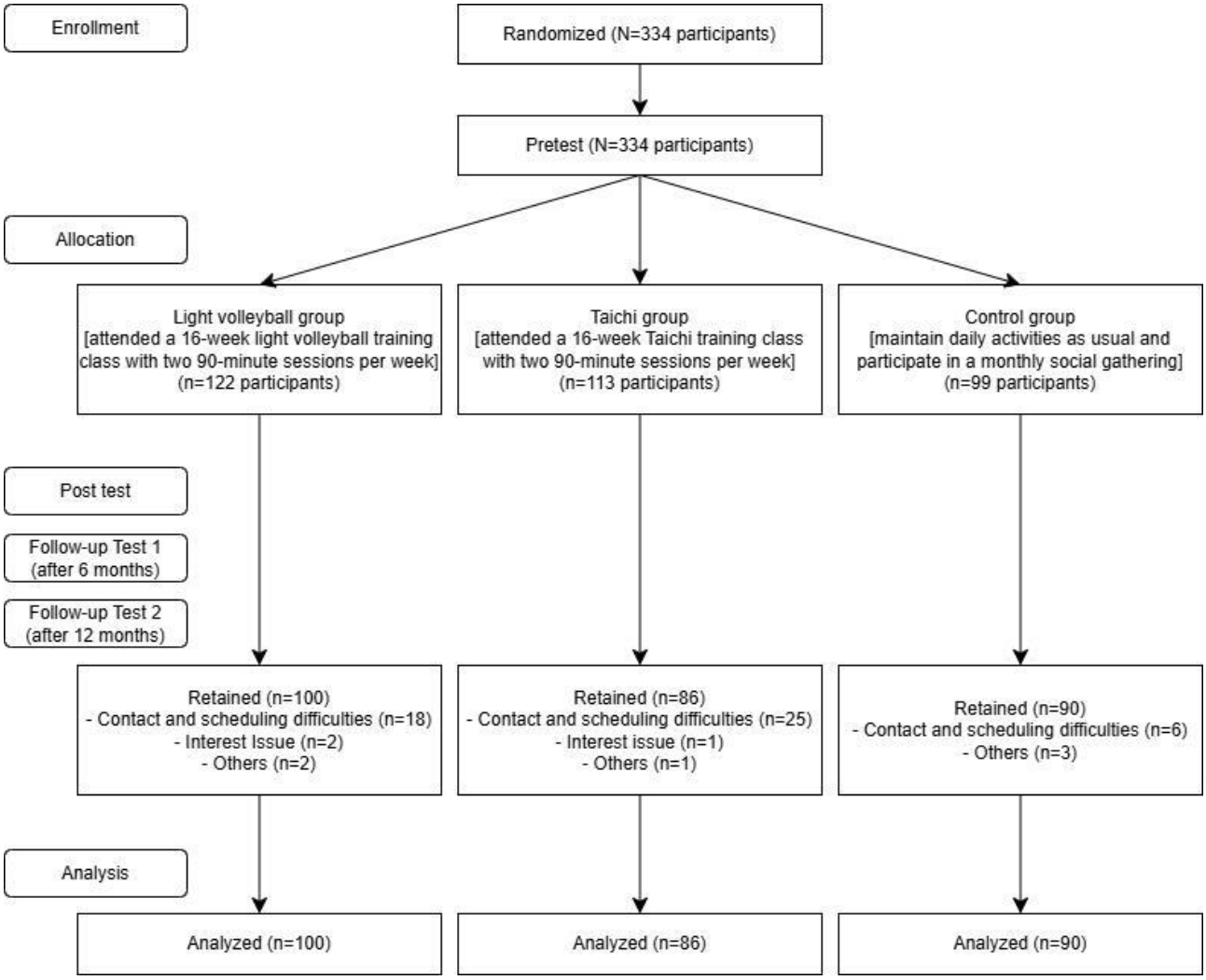
Results

Overview

Figure 1 presents the recruitment process for the current study. At the start of the intervention, 334 older adults from 7 elderly centers were enrolled in the program and completed the pretest. With simple randomization, there were 122 older adults in LVB group, 113 in Taichi group, and 99 in control group. A total of 3 participants (2 from the LVB group and 1 from the Taichi group) withdrew from the program because of personal reasons, and 7 others (2 from the LVB group, 1 from the Taichi group, and 3 from the control group) withdrew without providing any reason. Among the remaining 318 participants, 42 did not complete the posttest for various personal reasons (eg, illness, Lunar New Year gatherings, and concerns about social distancing during the late COVID-19 pandemic). This sample size met our calculated sample size for this intervention (with effect size of 0.5 [Cohen *d*]) in order to achieve a power of 80% at a significance level of 5% [19]. The retention rate was 82.6%, slightly higher than the expected 80%. Ultimately, 276 participants (LVB group, n=100; Taichi group, n=86; control group, n=90) who completed both the pre- and posttest were included in the final data analysis.

Table 1 presents the sociodemographic characteristics of the participants. Approximately 56% (156/276) of the participants were aged ≤70 years, and only 8% (22/276) were aged ≥80 years. The majority of the participants were female (229/276, 83%) and retired (217/276, 78.6%). Approximately 57.3% (158/276) had attained secondary education or higher. Approximately half (161/276, 58.3%) reported their financial status as average. The average BMI was 24 (SD 3.6) kg/m^2. The groups did not significantly differ with each other at the pretest with respect to all variables ($P>.05$).

Figure 1. Recruitment statistics.



XSL•FO
RenderX

Table . Sociodemographic characteristics of participants.

Characteristics	LVB[a] (n=100)	TC[b] (n=86)	Control[c] (n=90)
Age (years), n (%)			
60 - 64	30 (30)	14 (16.3)	10 (11.1)
65 - 69	35 (35)	34 (39.5)	33 (36.7)
70 - 74	24 (24)	26 (30.2)	17 (18.9)
75 - 79	8 (8)	8 (9.3)	15 (16.7)
80 or above	3 (3)	4 (4.7)	15 (16.7)
Sex, n (%)			
Male	17 (17)	18 (20.9)	12 (13.3)
Female	83 (83)	68 (79.1)	78 (86.7)
Occupation, n (%)			
Full-time job	0 (0)	0 (0)	1 (1.1)
Housewife	16 (16)	11 (12.8)	21 (23.3)
Retired	79 (79)	72 (83.7)	66 (73.3)
Part-time job or others	5 (5)	3 (3.5)	2 (2.2)
Education level, n (%)			
No education	5 (5)	4 (4.7)	21 (23.3)
Primary education	35 (35)	23 (26.7)	30 (33.3)
Secondary education	40 (40)	41 (47.7)	29 (32.2)
Tertiary education	20 (20%)	18 (20.9%)	10 (11.1%)
Perceived financial status, n (%)			
Low	19 (19)	18 (20.9)	18 (20)
Below average	21 (21)	17 (19.8)	10 (11.1)
Average	53 (53)	50 (58.1)	58 (64.4)
Above average	3 (3)	1 (1.2)	3 (3.3)
Higher	4 (4)	0 (0)	1 (1.1)
House nature, n (%)			
Bought	46 (46)	33 (38.4)	32 (35.6)
Rent	54 (54)	53 (61.6)	58 (64.4)
BMI (kg/m^2), mean (SD)	24.1 (3.7)	23.5 (3.7)	24.2 (3.4)

[a]LVB: light volleyball.

[b]TC: Taichi.

[c]CG: control group.

Improvement in Physical Fitness

Table 2 presents the mean and SD values of physical measures for the groups at pre- and posttest, along with the results of analysis of covariance with repeated measures, with BMI controlled for. After the intervention, the LVB group achieved statistically significant improvements in lower limb muscular endurance (chair stand test, $F_{2,272}$= 7.2, P=.001, η^2=.05), agility (up-and-go test, $F_{2,272}$= 6.1, P=.003, η^2=.04), and dynamic balance (Biodex balance test, $F_{2,272}$= 9.4, P=.001, η^2=.07) compared with the other groups.

Pairwise comparisons revealed that both the LVB and Taichi groups exhibited significantly improved adjusted performance in lower limb muscular endurance (chair stand test, LVB group, mean 17.6, SD 5.0; Taichi group, mean 15.3, SD 4.9, P=.001), agility (up-and-go test, LVB group, mean 5.6, SD 1.3; Taichi group, mean 6.2, P=.001), and dynamic balance (Biodex balance test, LVB group, mean 0.5, SD 0.5; Taichi group, mean 0.7, SD 0.6, P=.001) compared with the control group (chair stand test, mean 15.3, SD 4.9; up-and-go test, mean 6.1, SD 1.2; Biodex balance test, mean 0.7, SD 0.5). Notably, the LVB group performed significantly better on all three measures (chair stand test, up-and-go test, and Biodex balance test) at posttest compared with the Taichi group. No significant group

XSL•FO
RenderX

differences were identified for the arm curl test ($F_{2,270}$=2.8, P=.1, partial η^2=.02), chair sit-and-reach test ($F_{2,270}$=0.06, P=.5, partial η^2=.005), and back scratch test ($F_{2,270}$=1.3, P=.3, partial η^2=.01), and 2-min step test ($F_{2,270}$=1.1, P=.3, partial η^2=.008).

Table . Mean and SD values for measures in groups at pre- and posttest.

Measures	LVB[a] (n=100), mean (SD)		TC[b] (n=86), mean (SD)		CG[c] (n=90), mean (SD)		*F* test (*df*)	Mean difference (95% CI)	
	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest		LVB-TC	LVB-CG
Chair stand test (frequency)	16.4 (5.1)	17.6 (5.0)	15 (5.1)	15.3 (4.9)	15 (3.7)	15.3 (3.5)	7.2[d] (2, 272)	2.4[d] (0.8 to 3.9)	2.3[d] (0.8 to 3.9)
Arm curl (frequency)	15.3 (4.9)	16.0 (4.8)	14.4 (5.1)	14.6 (4.9)	14.1 (5.3)	14.2 (5.0)	2.8 (2, 270)	1.5 (−0.2 to 3.2)	1.9[e] (0.2 to 3.5)
Chair sit-and-reach test (cm)	7.3 (11.5)	7.8 (11.4)	5 (13.0)	5.2 (12.6)	5.1 (9.5)	4.8 (8.8)	0.7 (2, 272)	2.7 (−0.8 to 6.1)	3 (−0.4 to 6.4)
Back scratch (cm)	0.6 (9.3)	1.0 (9.5)	0.9 (8.9)	0.1 (8.7)	−2 (9.2)	−2.7 (9.4)	1.3 (2, 270)	0.9 (−1.9 to 3.6)	3.7[e] (0.1 to 6.3)
Up-and-go test (s)	6 (1.5)	5.6 (1.3)	6.2 (1.8)	6.2 (1.7)	6.5 (1.4)	6.1 (1.2)	6.1[d] (2, 272)	−.6[e] (−1.1 to −0.1)	−.7[d] (−1.2 to −0.2)
Step test (frequency)	92.1 (18.3)	94.6 (17.6)	89.4 (22.9)	90.0 (21.8)	85.3 (17.6)	85.9 (16.7)	1.3 (2, 272)	4.56 (−1.5 to 10.7)	8.7[e] (2.7 to 14.7)
Overall stability index (score)	0.7 (0.6)	0.5 (0.5)	0.7 (0.6)	0.7 (0.6)	0.8 (0.6)	0.7(0.5)	9.4[d] (2, 272)	−0.2[e] (−0.3 to −.01)	−0.2[d] (−0.4 to −0.04)

[a]LVB: light volleyball.

[b]TC: Taichi.

[c]CG: control group.

[d]P<.01.

[e]P<.05.

Discussion

Principal Findings

Given the continued aging of the population in Hong Kong and the limitations of our previous LVB pilot study, the current study aimed to investigate the effect of a 16-week LVB intervention program on physical health outcomes among older adults. Compared with the pilot study, the current study used an randomized controlled trial design and recruited 5.45 times more participants. The results indicated that the LVB intervention had a greater effect on improving lower body strength, agility, and dynamic balance in older adults when compared with both the Taichi intervention and the control group. Although significant improvements in lower body strength, agility, and dynamic balance were identified in the LVB intervention group, no significant changes in aerobic endurance and upper body strength were identified.

Improvement in Physical Health

The current study hypothesized that the LVB intervention would lead to greater improvements in lower body strength, agility, and dynamic balance relative to the control. This expectation is supported by our pilot study in 2018, which revealed significant improvements in these physical attributes for the LVB intervention group when compared with the control group [20].

We also hypothesized that the LVB intervention would improve lower body strength, agility, and dynamic balance relative to Taichi. First, LVB shares several similarities with traditional volleyball, such as involving a considerable amount of lower body movement, which helps enhance lower body strength [36]. Second, LVB was expected to result in more significant improvements in agility and dynamic balance compared with Taichi. This is because LVB requires players to use open skills, enabling them to react and adapt to dynamic, constantly changing environments. In contrast, Taichi primarily involves closed skills, which do not require such adaptive responses to environmental changes [37]. Although there may be differences, Sheppard and Young [38] defined agility as the ability to change speed or direction in response to a stimulus, such as an environmental change. A recent systematic review comparing the effects of open-skill physical activity and closed-skill physical activity on cognitive function reported that open-skill physical activity was more effective in enhancing cognitive

XSL•FO
RenderX

function than closed-skill exercises were [39]. Furthermore, Young et al [40] suggested that agility is strongly linked to cognitive functions such as decision-making ability and perception.

Additionally, a strong correlation has been observed between lower body strength and the ability to change direction (change-of-direction ability) [41]. Recent research has demonstrated a significant correlation between relative and absolute strength and agility, including change-of-direction ability and linear speed [42]. In the current study, we also identified a correlation between lower body strength and agility. These were the 2 physical attributes for which more pronounced improvements were identified in the LVB intervention group than in the Taichi group. In addition to lower body strength, balance plays a crucial role in maintaining good posture during acceleration, deceleration, and sudden changes in location or direction. Balance training has been found to be beneficial in improving the agility of volleyball players [43].

Although the participants in the LVB group exhibited greater improvements in lower body strength, agility, and dynamic balance, an unexpected finding was the minimal differences in upper body strength and aerobic endurance, 2 physical attributes that were initially hypothesized to exhibit significant improvement in the LVB group compared with the Taichi group. For upper body strength, LVB involves more frequent and vigorous arm and shoulder movements relative to Taichi, such as spiking and blocking, which were expected to result in greater upper body strength improvements compared with Taichi [44]. Similarly, the higher energy demand in volleyball was assumed to lead to greater improvements in aerobic endurance in the LVB group compared with Taichi [45]. However, despite these expectations, no significant differences in improvements were observed between the LVB and Taichi groups for these two physical attributes. Nonetheless, both the LVB and Taichi programs were demonstrated to improve upper body strength and aerobic endurance in older adults, although LVB did not significantly outperform Taichi as was hypothesized. Previous studies have indicated that both LVB and Taichi can be beneficial for improving upper body muscle strength in older adults [19,44]. Regarding aerobic capacity, although we assumed that volleyball would require higher aerobic endurance,Taichi's breathing techniques were found to be beneficial for improving aerobic capacity in older adults, providing benefits similar to those achieved through LVB [46].

Limitations

Although this study addressed the limitations of the previous pilot study, it still has several limitations of its own. First, these are only preliminary findings, and the scheduled follow-up evaluation in 6 or 12 months will help determine whether these advantages are seen or sustained over the longer term after the intervention. Second, the gender representation was disproportionate; the number of female participants was higher because recruitment was conducted through local elderly centers. These centers in Hong Kong have predominantly female memberships, with older women showing a greater tendency to participate in activities [47,48]. This imbalance in gender representation limits the generalizability of the findings. Furthermore, the study was conducted during the COVID-19 pandemic and the proposed intervention period was postponed from 2020 - 2022 to 2021 - 2023, which may have introduced variables that are not typically present in nonpandemic conditions. These factors could have affected the results of the interventions.

Conclusions

The current study examined the effects of a 16-week LVB intervention program on the physical health of older adults in Hong Kong. The results indicated that the participants in the LVB group experienced significant improvements in lower body strength, agility, and dynamic balance compared with both the Taichi active control group and the control group. This study builds on the previous pilot by adopting an randomized controlled trial design, incorporating dynamic balance as a fitness component, increasing the sample size, and collecting data from 7 local elderly centers instead of just one.

Future studies should address the following: first, studies should aim for more balanced gender representation and include follow-up tests to monitor the long-term maintenance of physical improvements. Numerous studies have highlighted the positive effects of physical activity on aspects such as cognition and psychology in older adults, whereas the current study focused solely on physical attributes. Future research could explore the effects of LVB on cognitive and psychological outcomes to provide a more comprehensive understanding of its benefits. Furthermore, future studies should consider qualitative approaches to understand participants' experiences and assess how LVB intervention influences various health dimensions. Given the potential limitations due to the COVID-19 pandemic, future studies may benefit from replication under more stable conditions, free from social distancing policies, to determine whether similar results can be observed. In conclusion, this large-scale study provides strong evidence supporting the physical health benefits of LVB for older adults. The results from our future qualitative studies and follow-up measures will further inform researchers and practitioners about the acceptability and appropriateness of LVB interventions for older adults.

Conflicts of Interest

None declared.

Checklist 1
CONSORT Checklist
[PDF File, 227 KB - aging_v8i1e62886_app1.pdf]

XSL•FO
RenderX

References

1. Persons with disabilities and chronic diseases. Census and Statistics Department, Hong Kong Special Administrative Region. 2020. URL: https://www.censtatd.gov.hk/en/data/stat_report/product/C0000055/att/B11301632021XXXXB0100.pdf [accessed 2025-04-08]
2. Synopsis of healthcare financing studies projection of Hong Kong's healthcare expenditure. Food and Health Bureau, Hong Kong Special Administrative Region Government. 2008. URL: https://www.healthbureau.gov.hk/beStrong/files/consultation/projecthealthexp_eng.pdf [accessed 2025-04-08]
3. Moreno-Agostino D, Daskalopoulou C, Wu YT, et al. The impact of physical activity on healthy ageing trajectories: evidence from eight cohort studies. Int J Behav Nutr Phys Act 2020 Jul 16;17(1):92. [doi: 10.1186/s12966-020-00995-8] [Medline: 32677960]
4. Territory-wide physical fitness survey for the community. Sports Commission of Hong Kong Special Administrative Region Government. 2024. URL: https://www.legco.gov.hk/yr2023/english/panels/ha/papers/ha20230612cb2-539-5-e.pdf [accessed 2025-04-08]
5. Healthy exercise for all campaign - fitness programmes for the elderly. Leisure and Cultural Services Department, Hong Kong Special Administrative Region Government. 2023. URL: https://www.lcsd.gov.hk/en/healthy/fitness/elderly.html [accessed 2025-04-08]
6. Elderly commission: report on healthy ageing executive summary. The Government of Hong Kong Special Administrative Region fo the People's Republic of China. 2022. URL: https://www.elderlycommission.gov.hk/en/library/Ex-sum.htm [accessed 2025-04-08]
7. Chief Secretary for Administration's Office. Population policy strategies and initiatives. 2016. URL: https://gia.info.gov.hk/general/201501/15/P201501150880_0880_140608.pdf [accessed 2025-04-08]
8. Spiteri K, Broom D, Bekhet AH, de Caro JX, Laventure B, Grafton K. Barriers and motivators of physical activity participation in middle-aged and older-adults - a systematic review. J Aging Phys Act 2019 Sep 1;27(4):929-944. [doi: 10.1123/japa.2018-0343] [Medline: 31141447]
9. World Health Organization. Active ageing: a policy framework (no. WHO/NMH/NPH/02.8). 2004. URL: https://extranet.who.int/agefriendlyworld/wp-content/uploads/2014/06/WHO-Active-Ageing-Framework.pdf [accessed 2025-04-08]
10. Leisure and Cultural Services Department, Hong Kong Special Administrative Region Government. Use of public open space. 2016. URL: https://www.legco.gov.hk/yr15-16/english/panels/ha/papers/ha20160614cb2-1680-3-e.pdf [accessed 2025-04-08]
11. Sherrington C, Michaleff ZA, Fairhall N, et al. Exercise to prevent falls in older adults: an updated systematic review and meta-analysis. Br J Sports Med 2017 Dec;51(24):1750-1758. [doi: 10.1136/bjsports-2016-096547] [Medline: 27707740]
12. Sun M, Min L, Xu N, Huang L, Li X. The effect of exercise intervention on reducing the fall risk in older adults: a meta-analysis of randomized controlled trials. Int J Environ Res Public Health 2021 Nov 29;18(23):12562. [doi: 10.3390/ijerph182312562] [Medline: 34886293]
13. Chase JAD, Phillips LJ, Brown M. Physical activity intervention effects on physical function among community-dwelling older adults: a systematic review and meta-analysis. J Aging Phys Act 2017 Jan;25(1):149-170. [doi: 10.1123/japa.2016-0040] [Medline: 27620705]
14. Taylor J, Walsh S, Kwok W, et al. A scoping review of physical activity interventions for older adults. Int J Behav Nutr Phys Act 2021 Jun 30;18(1):82. [doi: 10.1186/s12966-021-01140-9] [Medline: 34193157]
15. Pinheiro MB, Reis AHS, Baldwin JN, et al. Quantity and quality are increasing but there's room for improvement: A scoping review of physical activity intervention trials. Braz J Phys Ther 2024;28(2):101051. [doi: 10.1016/j.bjpt.2024.101051] [Medline: 38574557]
16. Wong MYC, Ou KL, Chung PK, Chui KYK, Zhang CQ. The relationship between physical activity, physical health, and mental health among older Chinese adults: A scoping review. Front Public Health 2022;10:914548. [doi: 10.3389/fpubh.2022.914548] [Medline: 36684983]
17. Franco MR, Tong A, Howard K, et al. Older people's perspectives on participation in physical activity: a systematic review and thematic synthesis of qualitative literature. Br J Sports Med 2015 Oct;49(19):1268-1276. [doi: 10.1136/bjsports-2014-094015] [Medline: 25586911]
18. Van Dyck D, Mertens L, Cardon G, De Cocker K, De Bourdeaudhuij I. Opinions toward physical activity, sedentary behavior, and interventions to stimulate active living during early retirement: a qualitative study in recently retired adults. J Aging Phys Act 2017 Apr;25(2):277-286. [doi: 10.1123/japa.2015-0295] [Medline: 27759483]
19. Leung KM, Chung PK, Chan AWK, et al. Promoting healthy ageing through light volleyball intervention in Hong Kong: study protocol for a randomised controlled trial. BMC Sports Sci Med Rehabil 2020;12(1):6. [doi: 10.1186/s13102-019-0151-7] [Medline: 32002186]
20. Leung KM, Chung PK, Hagger MS. The effects of light volleyball intervention programme in improving selected physical and psychological attributes of older adults in Hong Kong. Int J Sport Exerc Psychol 2020 Jan 2;18(1):1-12. [doi: 10.1080/1612197X.2018.1462231]

XSL•FO
RenderX

21. Chan AWK, Lee A, Lee DTF, et al. The sustaining effects of Tai chi Qigong on physiological health for COPD patients: a randomized controlled trial. Complement Ther Med 2013 Dec;21(6):585-594. [doi: 10.1016/j.ctim.2013.09.008] [Medline: 24280465]
22. Lindsay Smith G, Banting L, Eime R, O'Sullivan G, van Uffelen JGZ. The association between social support and physical activity in older adults: a systematic review. Int J Behav Nutr Phys Act 2017 Apr 27;14(1):56. [doi: 10.1186/s12966-017-0509-8] [Medline: 28449673]
23. Burke SM, Carron AV, Eys MA, Ntoumanis N, Estabrooks PA. Group versus individual approach? A meta-analysis of the effectiveness of interventions to promote physical activity. Int Rev Sport Exerc Psychol 2006 Feb;2(1):13-29. [doi: 10.53841/bpssepr.2006.2.1.13]
24. Healthy exercise for all campaign – physical fitness test for the community. Sports Commission of Hong Kong Special Administrative Region Government. 2012. URL: https://www.censtatd.gov.hk/en/data/stat_report/product/FA100265/att/B71302FA2013XXXXB0100.pdf [accessed 2025-04-08]
25. Farrance C, Tsofliou F, Clark C. Adherence to community based group exercise interventions for older people: a mixed-methods systematic review. Prev Med 2016 Jun;87:155-166. [doi: 10.1016/j.ypmed.2016.02.037] [Medline: 26921655]
26. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMC Med 2010 Mar 24;8(1):18. [doi: 10.1186/1741-7015-8-18] [Medline: 20334633]
27. Shekelle PG, Maglione MA, Mojica W, et al. Exercise programs for older adults: a systematic review and meta-analysis. RAND. 2003 URL: https://www.rand.org/pubs/reprints/RP1257.html [accessed 2025-05-08]
28. Chu LW, Pei CKW, Ho MH. Validation of the abbreviated mental test (Hong Kong version) in the elderly medical patient. Hong Kong Med J 1995:207-211.
29. Schoene D, Wu S, Mikolaizak AS, Wu S‐, et al. Discriminative ability and predictive validity of the timed up and go test in identifying older people who fall: systematic review and meta‐analysis. J American Geriatrics Society 2013 Feb;61(2):202-208 [FREE Full text] [doi: 10.1111/jgs.12106]
30. Rikli RE, Jones CJ. Senior fitness test manual. In: Human Kinetics 2013.
31. Rikli RE, Jones CJ. Development and validation of a functional fitness test for community-residing older adults. J Aging Phys Act 1999 Jan;7(2):129-161. [doi: 10.1123/japa.7.2.129]
32. Parraca JA, Olivares PR, Carbonell-Baeza A, Aparicio VA, Adsuar JC, Gusi N. Test-retest reliability of Biodex Balance SD on physically active old people. JHSE 2011 Jun 1;6(2):444-451 [FREE Full text] [doi: 10.4100/jhse.2011.62.25]
33. Finn JA, Alvarez MM, Jett RE, Axtell RS, Kemler DS. Stability performance assessment among subjects of disparate balancing abilities. Med Sci Sports Exerc 1999 May;31(Supplement):S252. [doi: 10.1097/00005768-199905001-01209]
34. Riebe D, Blissmer BJ, Greaney ML, Garber CE, Lees FD, Clark PG. The relationship between obesity, physical activity, and physical function in older adults. J Aging Health 2009 Dec;21(8):1159-1178. [doi: 10.1177/0898264309350076] [Medline: 19897781]
35. Cohen J. Statistical power analysis for the behavioral sciences. Comput Environ Urban Syst 1990 Jan;14(1):71. [doi: 10.1016/0198-9715(90)90050-4]
36. Tramel W, Lockie RG, Lindsay KG, Dawes JJ. Associations between absolute and relative lower body strength to measures of power and change of direction speed in division II female volleyball players. Sports (Basel) 2019 Jul 1;7(7):160. [doi: 10.3390/sports7070160] [Medline: 31266193]
37. Koch P, Krenn B. Executive functions in elite athletes – comparing open-skill and closed-skill sports and considering the role of athletes' past involvement in both sport categories. Psychol Sport Exerc 2021 Jul;55:101925. [doi: 10.1016/j.psychsport.2021.101925]
38. Sheppard JM, Young WB. Agility literature review: classifications, training and testing. J Sports Sci 2006 Sep;24(9):919-932. [doi: 10.1080/02640410500457109] [Medline: 16882626]
39. Gu Q, Zou L, Loprinzi PD, Quan M, Huang T. Effects of open versus closed skill exercise on cognitive function: a systematic review. Front Psychol 2019;10:1707. [doi: 10.3389/fpsyg.2019.01707] [Medline: 31507472]
40. Young WB, Dawson B, Henry GJ. Agility and change-of-direction speed are independent skills: implications for training for agility in invasion sports. Int J Sports Sci Coach 2015 Feb;10(1):159-169. [doi: 10.1260/1747-9541.10.1.159]
41. Sheppard JM, Dawes JJ, Jeffreys I, Spiteri T, Nimphius S. Broadening the view of agility: a scientific review of the literature. Journal of Australian Strength & Conditioning 2014;22(3):6-25 [FREE Full text]
42. Andersen E, Lockie RG, Dawes JJ. Relationship of absolute and relative lower-body strength to predictors of athletic performance in collegiate women soccer players. Sports (Basel) 2018 Sep 29;6(4):106. [doi: 10.3390/sports6040106] [Medline: 30274252]
43. Gadre HV, Lele DC, Deo M, Mathur C. Effect of dynamic balance training on agility in adolescent volleyball players. JCDR 2019 Jan 1. [doi: 10.7860/JCDR/2019/42282.13310]
44. Hedrick A. Training for high level performance in women's collegiate volleyball: part I training requirements. J Strength Cond 2007;29(6):50. [doi: 10.1519/1533-4295(2007)29[50:TFHLPI]2.0.CO;2]
45. Wehner C, Blank C, Arvandi M, Wehner C, Schobersberger W. Effect of tai chi on muscle strength, physical endurance, postural balance and flexibility: a systematic review and meta-analysis. BMJ Open Sport Exerc Med 2021;7(1):e000817. [doi: 10.1136/bmjsem-2020-000817] [Medline: 33614126]

XSL•FO
RenderX

46. Zou L, Loprinzi PD, Yu JJ, et al. Superior effects of modified Chen-style tai chi versus 24-style tai chi on cognitive function, fitness, and balance performance in adults over 55. Brain Sci 2019 May 4;9(5):102. [doi: 10.3390/brainsci9050102] [Medline: 31060221]
47. Caritas hong kong – servies for the elderly. Service Communications. 2007. URL: https://www.caritasse.org.hk/zh/wp-content/uploads/2020/07/%E5%AE%89%E8%80%81%E6%9C%8D%E5%8B%99%E9%80%9A%E8%A8%8A_%E7%AC%AC11%E6%9C%9F.pdf [accessed 2025-04-08]
48. St. James Settlement. The Voice Issue No 350 The Voice 2006 [Website in Chinese]. URL: https://www.thevoice.org.hk/v0350/00002.htm [accessed 2025-05-08]

Abbreviations

AMT : Abbreviated Mental Test
CONSORT: Consolidated Standards of Reporting Trials
LVB: light volleyball
TUG: Timed-Up-and-Go
WHO: World Health Organization

Edited by J Job; submitted 03.06.24; peer-reviewed by C Sherrington, R Strauss; revised version received 22.01.25; accepted 29.01.25; published 28.05.25.

Please cite as:
Leung KM, Shi Y
Impact of a Light Volleyball Intervention Program on Improving Physical Attributes of Older Adults in Hong Kong: Preliminary Study of a Randomized Controlled Trial
JMIR Aging 2025;8:e62886
URL: https://aging.jmir.org/2025/1/e62886
doi:10.2196/62886

© Ka Man Carman Leung, Yuchen Shi. Originally published in JMIR Aging (https://aging.jmir.org), 28.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Sexual Response Problems and Their Correlates Among Older Adults From the Sexual Well-Being (SWELL) Study in China: Multicenter Cross-Sectional Study

Bingyu Liang[1,2*], PhD; Chen Xu[3*], PhD; Bingyi Wang[4*], PhD; Xinyi Li[4*], BM; Xin Peng[4*], MD; Ying Wang[5], PhD; Hui Li[6], BM; Yong Lu[7], PhD; Xiaopei Shen[8], MD; Lin Ouyang[9], MD; Guohui Wu[9], MD; Maohe Yu[10], MD; Jiewei Liu[11], BM; Xiaojun Meng[12], PhD; Yong Cai[3*], PhD; Huachun Zou[4,13*], PhD

[1]School of Public Health, Guangxi Key Laboratory of AIDS Prevention and Treatment, Guangxi Medical University, Nanning, China

[10]Department of AIDS/STD Control and Prevention, Tianjin Center for Disease Control and Prevention, Tianjin, China

[11]Baiyun District Center for Disease Control and Prevention, Guanghzou, China

[12]Department of AIDS/STD Control and Prevention, Wuxi Municipal Center for Disease Control and Prevention, Wuxi, China

[13]School of Public Health, Fudan University, 130 Andong Road, Xuhui District, Shanghai, China

[2]Department of Clinical Research, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom

[3]Public Health Research Center, Hongqiao International Institute of Medicine, Tongren Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[4]School of Public Health (Shenzhen), Sun Yat-sen University, Shenzhen, China

[5]School of Public Health, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[6]Department of AIDS/STD Control and Prevention, Shizhong District Center for Disease Control and Prevention, Jinan, China

[7]School of Public Health, the Key Laboratory of Environmental Pollution Monitoring and Disease Control, Guizhou Medical University, Guizhou, China

[8]Department of AIDS/STD Control and Prevention, Affiliation Shanghai Municipal Center for Disease Control and Prevention, Shanghai, China

[9]Department of AIDS/STD Control and Prevention, Chongqing Municipal Center for Disease Control and Prevention, Chongqing, China

[*]these authors contributed equally

Corresponding Author:
Huachun Zou, PhD
School of Public Health (Shenzhen), Sun Yat-sen University, Shenzhen, China

Abstract

Background: Sexual response problems among older adults are not an inevitable consequence of aging but rather a response to sexual health. However, there is a lack of recent and multicenter data on this issue in China.

Objective: This study aims to assess the prevalence of sexual response problems and their correlates among older adults.

Methods: A multicenter cross-sectional study on sexual well-being was conducted among individuals aged more than 50 years in China between June 2020 and December 2022. Data on sociodemographics, physical health, psychological health, and sexual response problems were collected through face-to-face interviews. We included sexually active older adults who reported either vaginal, oral, or anal sex in the past 12 months for this study. Sexual response problems included a lack of interest or enjoyment in sex; feeling anxious, having pain, or no excitement during sex; no desire or orgasms; and the lack of lubrication in sex. The stepwise logistic regression models were used to examine the correlates of sexual response problems.

Results: A total of 1317 sexually active older adults (842 men, 475 women) were included. Older women reported a higher prevalence of sexual response problems than older men (52.0% [247/475] vs 43.1% [363/842]). Common factors associated with at least one of the sexual response problems included living in rural areas (men: adjusted odds ratio [aOR]=0.31, 95% CI 0.22 - 0.43; women: aOR=0.29, 95% CI 0.19 - 0.43) and abnormal BMI (aOR=men: 1.52, 95% CI1.11 - 2.07; women: aOR=2.19, 95% CI 1.47 - 3.28). Among older men, sleep quality (aOR=1.87, 95% CI 1.30 - 2.68), emotional connection with sex partners during sexual intercourse (aOR=0.69, 95% CI 0.50 - 0.96), frequently experienced fatigue (aOR=2.47, 95% CI 1.59 - 3.90), anxiety (aOR=4.26, 95% CI 1.12 - 21.27), and seeking professional help for sex life (aOR=1.58, 95% CI 1.14 - 2.21) were associated with sexual response problems. Among older women, sexual response problems were associated with a lack of physical exercise (aOR=1.69, 95% CI 1.13 - 2.54), poor sex-partner relationships (aOR=1.70, 95% CI 1.12 - 2.60), and depressive symptoms (aOR=3.18, 95% CI 1.18 - 10.24).

XSL•FO
RenderX

Conclusions: Sexual response problems are common among older adults. These problems were associated with adverse physical health, mental health, and poor sex-partner relationships. These findings highlight the importance for health care providers to take into account the physical and psychological health of older adults, as well as the quality of their relationships with sexual partners when diagnosing and addressing sexual response problems.

(JMIR Aging 2025;8:e66772) doi:10.2196/66772

KEYWORDS

dysfunction; sexual health; sexual well-being; sexually active; sexual activity; well-being; correlate; sex partner relationship; gerontology; geriatrics; older adults; elder; elderly; older person; aging; China; cross-sectional study

Introduction

Sexual response problems are characterized by diminished or absent sexual interest and disruptions in the physiological or psychosocial patterns associated with the sexual response cycle. These disruptions include a lack of interest, enjoyment, desire, orgasms, and lubrication in sex [1]. Sexual response problems have a profound impact on various aspects of life for older adults, affecting their quality of life, interpersonal relationships, dissatisfaction in marriage [2,3], work efficiency [4], self-esteem [5], and physical and mental health [2,3,6].

Recent studies have shown a high prevalence of sexual problems in both older men and older women. For instance, in Europe, a majority of men (73.7% - 79.8%) and women (23.5% - 50.2%) reported experiencing at least one sexual problem, with the most prevalent sexual problems being orgasmic difficulties and reaching orgasm more quickly than desired and failure to reach orgasm or taking too long to climax [7,8]. An early national survey in urban China indicated that 21% of men and 35% of women aged 20 - 64 years had at least one persistent sexual dysfunction, with lack of sex interest, erectile difficulties (men), and inability to reach orgasm being the most common problems [9]. There are substantial variations in the prevalence of sexual response problems across different global regions, with noticeable differences between men and women. As extensively demonstrated by a significant number of epidemiological studies worldwide, there are substantial differences in the prevalence of sexual response problems among men and women. A few researchers have suggested that the prevalence of sexual response problems was higher in men compared to women [10,11]. Conversely, more studies have reported a higher prevalence among females [12-15]. It is necessary to elucidate the specific causes of these differences and inform gender-specific prevention and intervention strategies.

Sociodemographic, psychological, sex partner relationships, chronic disease, and physical health were found to be important determinants of sexual response problems among older adults. The prevalence of sexual problems tends to increase with age [16], and there are notable differences between men and women in terms of prevalence and types of sexual response problems [17]. Psychosocial factors such as anxiety, depression, stress, and the quality of marital relationships play significant roles in sexual response problems [16,18,19]. Intimate partner violence is associated with a higher likelihood of experiencing sexual problems [20]. In addition, chronic diseases like cancer [17], diabetes [21], and coronary heart disease [22] are linked to an increased risk of sexual dysfunction. Moreover, sexual response problems in one partner may influence the sexual function of the other partner [23].

Despite extensive studies on sexual problems, most of the existing literature was from developed countries or early studies, and there is a lack of nationally representative, large sample and recent data on the prevalence and correlates of sexual response problems among older adults in China. Given China's rapidly aging population [24,25], the sexual health of older adults is a growing concern. A comprehensive understanding of older adults' sexual response problems may enhance sex education, research, policy, and clinical care for this growing population. This multicentre cross-sectional study, using data from the Sexual Well-being (SWELL) study in China, aims to fill the research gap by examining the prevalence of sexual response problems and their correlates among older adults. These epidemiological data are essential for andrologists, gynecologists, urologists, venereologists, and other health care providers involved in treating and caring for older adults. They may help them counsel their patients on the potential adverse effects of different treatment modalities. Moreover, the findings are expected to contribute valuable insights for developing targeted interventions to enhance sexual relationships, improve quality of life, and address the sexual health challenges faced by aging populations in China.

Methods

Study Participants and Procedures

Our study was based on the SWELL study, a multicenter cross-sectional survey conducted between June 2020 and December 2022. The survey spanned four different regions in China, including Shanghai (Eastern China), Jinan (Eastern China), Chongqing (Western China), Guangzhou (Southern China), and Tianjin (Northern China). Participants were recruited using a multistage sampling design, and more detailed sampling procedures are provided in our previous protocol [26]. We collected data on demographic characteristics, physical health characteristics, mental health characteristics, sex partner relationship characteristics, and sexual behavior characteristics through face-to-face interviews. All participants provide formal informed consent to participate in the study.

Participants were enlisted from subdistricts within each chosen city. Eligibility criteria for participants in this study included: (1) aged 50 years and older; (2) only heterosexual orientation; (3) having engaged in heterosexual activities (including oral or vaginal intercourse) in the preceding year; (4) being able to comprehend the survey instrument of the SWELL Study.

XSL•FO
RenderX

Ethical Consideration

The SWELL study was approved by the School of Public Health (Shenzhen), Sun Yat-sen University Research Ethics Committee (approval number SYSU-PHS [2019] 006) and was performed following the Helsinki Declaration. Written consent was obtained from all participants, who were informed of their right to withdraw from the study at any time. Participant information and responses remained confidential, with anonymized data stored in password-protected folders accessible exclusively to the research team and supervisors.

Study Variables

Sexual Response Problems (Outcomes Variables)

Respondents were asked to report if they had experienced any of the following sexual response problems for three months in the preceding year: (1) lacked desire for sex, (2) lacked enjoyment in sex, (3) anxiety during sex, (4) discomfort or pain in sex, (5) no sexual arousal or excitement during sex, (6) lack of or delayed orgasm despite arousal, and (7) reaching orgasm faster than you would like, (8) lubrication difficulties (women only) or erectile function difficulties (men only). These items captured the major sexual response problem domains in the classification of sexual dysfunction in the *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition [27]. The reliability of sexual response problems in this study was 0.75.

Sociodemographic Characteristics

Age, gender (men and women), living area (rural or urban), monthly income (Chinese Yuan [CNY]), and years of education (<6 years: primary and lower; 7 - 12 years: senior or high school; >12 years: college and above) were included in demographic characteristics. Age was categorized into 3 age groups (50 - 59 years, 60 - 69 years, and older than 70 years). Monthly income (CNY) was categorized into 2 groups (≥5000 CNY [US $700] and <5000 CNY [US $700]).

Lifestyle Characteristics

Physical Exercise

Physical exercise was assessed with 5 items that inquired about the frequency of participation (6 times a week, 3 - 5 times a week, 1 - 2 times a week, no more than once a week, hardly ever, or never). The participants who reported engaging in physical exercise more than 1 - 2 times a week were categorized as often engaging in physical exercise. The remaining responses were categorized as not often exercised [28].

Seeking Professional Help for Sex Life

Participants were asked whether they had sought help or advice about their sex life from a range of sources in the past year. These sources included 4 informal sources (family member or friend, information and support sites on the internet) and 6 professional sources (general practitioner or family doctor, sexual health clinic, genitourinary clinic, sexually transmitted infection clinic, or relationship counselor); more than one answer was allowed. If a participant responds that they have previously sought help for sex life from 6 professional sources, this is defined as “seeking professional help for sex life.”

Physical Health Characteristics

Frequently Experiencing Fatigue

As for frequently experiencing fatigue, the participants were asked about their fatigue levels using a question: “Do you frequently experience fatigue?” with response options of “Yes” or “No.”

Chronic Disease

Chronic disease is defined as one or more diseases that last for 3 months or more including cardiovascular diseases (including myocardial infarction, coronary heart disease, angina, other forms of heart disease, and hypertension), arthritis, diabetes, or hyperglycemia, cerebral hemorrhage or stroke, chronic lung disease (excluding asthma), Parkinson disease, epilepsy, hyperlipidemia, gout or hyperuric acid, chronic gastroenteritis and chronic pain. A separate section was directed at women participants, whether they have a history of one of the following, including bladder surgery, genital or gynecologic surgery, cesarean section, abdominal surgery, and hip pelvic fractures or hip replacement.

Body Mass Index and Sleep Quality

In the SWELL Study, BMI is considered abnormal when lower than 18.5 or higher than 25.

Sleep quality was evaluated using the validated single-item sleep quality scale (SQS), which ranges from 1 to 10 and has been proven to be divided into 2 categories for analysis: those indicating poor sleep quality (scores 1 - 6) and those indicating good sleep quality (scores 7 - 10) [29]. The SQS had an acceptable internal consistency (Cronbach α=0.85).

Mental Health Characteristics

Depressive Symptoms

Depressive symptoms were measured by the 9-item Patient Health Questionnaire (PHQ-9), which has been validated and proven [29,30]. The scale’s total score ranges from 0 to 27, with scores≥10 representing clinically significant depressive symptoms. This study defined a score greater than or equal to 10 as depressive symptoms.

Anxiety Symptoms

The anxiety symptoms were measured on the generalized anxiety disorder-7 (GAD-7) scale with 7-item. Mild or normal anxiety was defined as a GAD-7 score <10, while moderate-to-vigorous anxiety was defined as a GAD-7 score ≥10 [29]. In this study, a score greater or equal to 10 was defined as with anxiety symptoms.

Sexual Relationship Characteristics

Emotional Connection With a Sexual Partner During Sex

The participants were asked, “How often would you say you feel emotionally close to your partner when you have sex together? (options: always, most of the time, sometimes, not very often, hardly ever).” The participants who selected the options of “always” and “most of the time” were defined as having a “good emotional connection with a sexual partner

XSL•FO
RenderX

during sex;" otherwise, they were defined as having a "poor emotional connection with a sexual partner during sex."

Relationship With Sex Partners

Regarding relationships with sex partners, the participants were asked, "How do you evaluate the relationship with your recent sex partner? Please assign a rating to the quality of your partnership with them." The rating scale ranges from 1 to 7, with 1 indicating "very good" and 7 representing "very poor." We categorized the responses into 2 categories for analysis: good relationship with a sex partner (scores 1 - 4) and poor relationship with a sex partner (scores 5 - 7).

Sexual Satisfaction

Sexual satisfaction was measured on a 5-point Likert scale, with responses ranging from 1 to 5 (strongly agree, agree, medium, disagree, and strongly disagree). In this study, we reclassified scores of 1 - 3 as sexual satisfaction and scores of 4 - 5 as sexual dissatisfaction [30].

Statistical Analyses

Descriptive analyses were conducted to characterize the study sample, including presenting percentages, means, and SD. The χ^2 test was used to compare the proportions of characteristics between the sex groups.

For the multivariable logistic regression analysis, collinearity diagnostics were initially performed for all potential variables (Multimedia Appendices 1 and 2). Subsequently, multivariate logistic regression analysis was carried out for noncollinear variables. The multivariable logistic regression models selected variables using a stepwise method based on the Akaike Information Criterion (AIC). The stepwise regression method combines forward selection and backward elimination approaches, adding and removing predictors in the model-building process. This approach effectively minimizes the inclusion of covariates, thereby enhancing the robustness of the analysis. Finally, the model with the minimum AIC was adopted (men: 1005.929; women: 582.316). Adjusted odds ratios (aORs) and their corresponding 95% CIs were estimated.

All statistical analyses were performed using R software version 4.2.3 (R Project). The Stats package (version 4.2.2) was used to build the stepwise multivariable logistic regression models. In addition, the figures were generated using the ggplot2 package (version 3.4.3) and forestmodel package (version 0.6.2) from CRAN.

Results

Demographic and Health Characteristics of the Participants

As shown in Table 1, 1317 older adults were included in this analysis. The average age was 64 years (SD 8.4 years, ranging from 50 to 86 years). Over half of the participants resided in rural areas (men: 53.1% [447/842], women: 50.9% [242/475]), and the majority reported 7 - 12 years of education (junior or senior high school; men: 70.0% [589/842], women: 59.6% [283/475]). In addition, a significant proportion of participants reported infrequent engagement in physical exercise (men: 54.8% [461/842], women: 48.2% [229/475]). Regarding physical health, more than half of the participants did not frequently experience fatigue (men: 85% [716/842], women: 74.9% [356/475]) and did not have chronic diseases (men: 55.2% [465/842], women: 62.9% [299/475]). Regarding sexual relationship characteristics, the majority of male participants reported sexual satisfaction (479/842, 56.9%) and a good relationship with their sex partner (615/842, 73.0%). In comparison, women reported slightly lower rates of sexual satisfaction (228/475, 48.0%) and a good relationship with their sex partner (306/475, 64.4%). Furthermore, less than 5.1% (24/475) of both men and women reported symptoms of anxiety and depression.

XSL•FO
RenderX

Table . Demographic, lifestyle, health, and sexual relationship characteristics among older adults aged more than 50 years in China (stratified by sex).

Characteristics	Men				Women			
				P values				*P* values
At least one of the sexual response problems, n (%)	All	Yes	No		All	Yes	No	
	842 (100.0)	363 (43.1)	479 (56.9)	N/A[a]	475 (100.0)	247 (52.0)	228 (48.0)	N/A
Demographic characteristics								
Living area				<.001				<.001
Rural	447 (53.1)	132 (29.5)	315 (70.5)		242 (50.9)	89 (36.8)	153 (63.2)	
Urban	395 (46.9)	231 (58.5)	164 (41.5)		233 (49.1)	158 (67.8)	75 (32.2)	
Age (years)				.04				.13
50 - 59	440 (52.3)	175 (39.8)	265 (60.2)		272 (57.3)	137 (50.4)	135 (49.6)	
60 - 69	300 (35.6)	134 (44.7)	166 (55.3)		156 (32.8)	79 (50.6)	77 (49.4)	
70+	102 (12.1)	54 (52.9)	48 (47.1)		47 (9.9)	31 (66)	16 (34)	
Monthly income (RMB)				.37				.72
≥5000	229 (27.2)	105 (45.9)	124 (54.1)		398 (83.8)	205 (51.5)	193 (48.5)	
<5000	613 (72.8)	258 (42.1)	355 (57.9)		77 (16.2)	42 (54.5)	35 (45.5)	
Education level				.009				.02
≤6	160 (19.0)	75 (46.9)	85 (53.1)		121 (25.5)	51 (42.1)	70 (57.9)	
7 - 12	589 (70.0)	236 (40.1)	353 (59.9)		283 (59.6)	152 (53.7)	131 (46.3)	
>12	93 (11.0)	52 (55.9)	41 (44.1)		71 (14.9)	44 (62)	27 (38)	
Lifestyle characteristics								
Physical exercise				.81				.001
Often	381 (45.2)	162 (42.5)	219 (57.5)		246 (51.8)	109 (44.3)	137 (55.7)	.
Not Often	461 (54.8)	201 (43.6)	260 (56.4)		229 (48.2)	138 (60.3)	91 (39.7)	
Seeking professional help for sex life				.001				.96
Yes	266 (31.6)	138 (51.9)	128 (48.1)		138 (29.1)	71 (51.4)	67 (48.6)	
No	576 (68.4)	225 (39.1)	351 (60.9)		337 (70.9)	176 (52.2)	161 (47.8)	
Physical health characteristics								
Frequently experienced fatigue				<.001				.07
Often	126 (15.0)	86 (68.3)	40 (31.7)		119 (25.1)	71 (59.7)	48 (40.3)	.07
Not often	716 (85.0)	277 (38.7)	439 (61.3)		356 (74.9)	176 (49.4)	180 (50.6)	
BMI				.001				<.001
Normal	434 (51.5)	162 (37.3)	272 (62.7)		230 (48.4)	95 (41.3)	135 (58.7)	
Abnormal	408 (48.5)	201 (49.3)	207 (50.7)		245 (51.6)	152 (62)	93 (38)	
Sleep quality				<.001				.35

XSL•FO
RenderX

Characteristics	Men				Women			
				P values				*P* values
Good	630 (74.8)	232 (36.8)	398 (63.2)		314 (66.1)	158 (50.3)	156 (49.7)	
Poor	212 (25.2)	131 (61.8)	81 (38.2)		161 (33.9)	89 (55.3)	72 (44.7)	
Chronic disease				.10				.06
Yes	377 (44.8)	163(43.2)	214 (56.8)		176 (37.1)	81 (46)	95 (54)	
No	465 (55.2)	200 (43)	265 (57)		299 (62.9)	166 (55.5)	133 (44.5)	
Mental health characteristics								
Anxiety symptoms				.002				.34
Yes	17 (2.0)	14 (82.4)	3 (17.6)		20 (4.2)	13 (65)	7 (35)	
No	825 (98.0)	349 (42.3)	476 (57.7)		455 (95.8)	234 (51.4)	221 (48.6)	
Depressive symptoms				<.001				.012
Yes	37 (4.4)	30 (81.1)	7 (18.9)		24 (5.1)	19 (79.2)	5 (20.8)	
No	805 (95.6)	333 (41.4)	472 (58.6)		451 (94.9)	228 (50.6)	223 (49.4)	
Sexual relationship characteristics								
Emotional connection with a sexual partner during sex				<.001				.004
Yes	573 (68.1)	223 (38.9)	350 (61.1)		294 (61.9)	144 (49)	150 (51)	
No	269 (31.9)	140 (52.0)	129 (48.0)		181 (38.1)	103 (56.9)	78 (43.1)	
Relationship with a sex partner				.002				.005
Good	615 (73.0)	245 (39.8)	370 (60.2)		306 (64.4)	144 (47.1)	162 (52.9)	
Poor	227 (27.0)	118 (52)	109 (48)		169 (35.6)	103 (60.9)	66 (39.1)	
Sexual satisfaction[b]				<.001				<.001
Yes	479 (56.9)	0 (0)	479 (100)		228 (48.0)	0 (0)	228 (100)	
No	363 (43.1)	363 (100)	0 (0)		247 (52.0)	247 (100)	0 (0)	

[a]N/A: not applicable

[b]Sexual response problems include lacked interest in having sex, lacked enjoyment in sex, feeling anxiety during sex, feeling physical pain as a result of sex, feeling no excitement or arousal during sex, difficulty in reaching climax, reaching climax more quickly than desired, trouble getting or keeping an erection(men) or uncomfortable dry vagina(women).

Prevalence of Sexual Response Problems

The prevalence of at least one sexual response problem (including or excluding lack of interest in sex) is shown in Figure 1 and Table 2. In total, 610 out of 1317 participants had sexual response problems, with an overall prevalence of sexual response problems of 46.3% (610/1317). There was a significant difference in the prevalence of the reported at least one sexual response problem between women and men, with being significantly higher in women than in men (52.0% [247/475] vs 43.1% [363/842], χ^2_1=9.6, *P*=.002). The prevalence of at least one sexual response problem increased with age among men (The Cochran-Armitage Trend Test, Z=−2.476, *P* for trend=.01). However, this trend was only observed in the age group between the older than 70 years age group and the other 2 age groups among men. There was no significant difference in the prevalence of sexual response problems between the 60 - 69 years age group and the 50 - 59 age group among women.

XSL•FO
RenderX

Figure 1. Prevalence of sexual response problems among older adults aged 50+ years**.** (**A**) At least one of sexual response problems among men and women; (**B**) At least one of sexual response problems excluding lack of interest in sex among men and women.

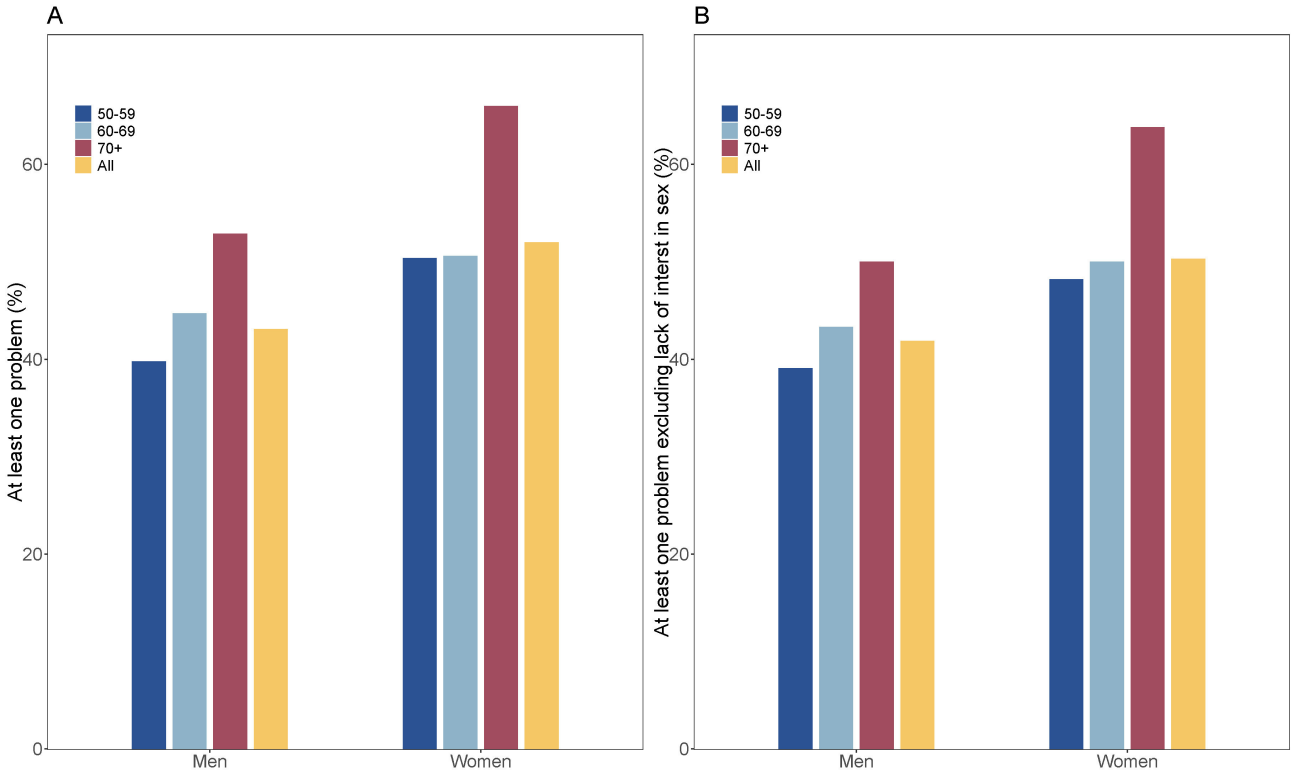


Table . The prevalence of sex response problems among men and women by age group.

Age group	Men	Women	Chi-square (df)	*P* values
At least one sex response problem				
50 - 59	175 (39.8)	137 (50.4)	7.24 (1)	.007
60 - 69	134 (44.7)	79 (50.6)	1.24 (1)	.27
Older than 70 years	54 (52.9)	31 (66.0)	1.73 (1)	.19
All age groups	363 (43.1)	247 (52.0)	9.65 (1)	.002
At least one sex response problem excluding lack of interest in sex				
50 - 59 years	172 (39.1)	131 (48.2)	5.29 (1)	.02
60 - 69 years	130 (43.3)	78 (50.0)	1.58 (1)	.21
Older than 70 years	51 (50)	30 (63.8)	1.95 (1)	.16
All age groups	353 (41.9)	239 (50.3)	8.31 (1)	.004

Correlates of Sexual Response Problems

The results of multivariable logistic regression analysis stratified by sex, which was presented in Figure 2, revealed several significant associations with reporting at least one sexual response problem.

XSL•FO
RenderX

Figure 2. Correlates of sexual response problems among older adults aged 50+ years**.** (**A**) Correlates of sexual response problems among men aged 50+ years; the multivariable logistic regression analysis was adjusted by living area, age, educational level, feeling of fatigue, body mass index, sleeping quality, seeking professional help for sex life, emotional connection with sexual partner during sex, anxiety, depressive symptoms. (**B**) Correlates of sexual response problems among women aged 50+ years; the logistic regression model was adjusted by age, living area, physical exercise, body-mass index, sleeping quality, sexual relationship with partner, emotional connection with sexual partner during sex, depressive symptoms. SRP: Sexual Response Problem; aOR: adjusted odds ratio; CI: confidence interval.

A

Variables	Level	N *(%)*	Prevalence of *SRP (%)*	*Adjusted odds ratio*	*aOR (95%CI)*	*p*
Living area	Urban	395 (46.91)	231 (58.48)		Reference	
	Rural	447 (53.09)	132 (29.53)		**0.31 (0.22, 0.43)**	**<0.001**
Age	50-59	440 (52.26)	175 (39.77)		Reference	
	60-69	300 (35.63)	134 (44.67)		1.28 (0.91, 1.79)	0.155
	70+	102 (12.11)	54 (52.94)		**1.72 (1.05, 2.81)**	**0.030**
Education level	Primary school and lower	93 (11.05)	52 (55.91)		Reference	
	Senior or high school	160 (19.00)	75 (46.88)		1.44 (0.78, 2.66)	0.241
	College and above	589 (69.95)	236 (40.07)		0.78 (0.47, 1.29)	0.336
Feeling of fatigue	No	716 (85.04)	277 (38.69)		Reference	
	Yes	126 (14.96)	86 (68.25)		**2.47 (1.59, 3.90)**	**<0.001**
Body-mass index	Normal	434 (51.54)	162 (37.33)		Reference	
	Abnormal	408 (48.46)	201 (49.26)		**1.52 (1.11, 2.07)**	**0.008**
Sleep quality	Good	630 (74.82)	232 (36.83)		Reference	
	Bad	212 (25.18)	131 (61.79)		**1.87 (1.30, 2.68)**	**<0.001**
Seeking professional help for sex life	No	576 (68.41)	225 (39.06)		Reference	
	Yes	266 (31.59)	138 (51.88)		**1.58 (1.14, 2.21)**	**0.007**
Emotional connection with sex partner during sex	No	269 (31.95)	140 (52.04)		Reference	
	Yes	573 (68.05)	223 (38.92)		**0.69 (0.50, 0.96)**	**0.029**
Anxiety (moderate to severe anxiety)	No	825 (97.98)	349 (42.30)		Reference	
	Yes	17 (2.02)	14 (82.35)		**4.26 (1.12, 21.27)**	**0.048**
Depressive symptoms	No	805 (95.61)	333 (41.37)		Reference	
	Yes	37 (4.39)	30 (81.08)		2.32 (0.94, 6.44)	0.083

B

Variables	Level	N *(%)*	Prevalence of *SRP (%)*	*Adjusted odds ratio*	*aOR (95%CI)*	*p*
Living area	Urban	233 (49.05)	158 (67.81)		Reference	
	Rural	242 (50.95)	89 (36.78)		**0.29 (0.19, 0.43)**	**<0.001**
Age	50-59	272 (57.26)	137 (50.37)		Reference	
	60-69	156 (32.84)	79 (50.64)		0.90 (0.58, 1.39)	0.62
	70+	47 (9.89)	31 (65.96)		1.93 (0.96, 3.99)	0.07
Physical exercise	Yes	246 (51.79)	109 (44.31)		Reference	
	No	229 (48.21)	138 (60.26)		**1.69 (1.13, 2.54)**	**0.01**
Body-mass index	Normal	230 (48.42)	95 (41.30)		Reference	
	Abnormal	245 (51.58)	152 (62.04)		**2.19 (1.47, 3.28)**	**<0.001**
Sleep quality	Good	314 (66.11)	158 (50.32)		Reference	
	Bad	161 (33.89)	89 (55.28)		1.42 (0.93, 2.19)	0.11
Relationship with sex partner	Good	306 (64.42)	144 (47.06)		Reference	
	Bad	169 (35.58)	103 (60.95)		**1.70 (1.12, 2.60)**	**0.01**
Emotional connection with sex partner during sex	No	181 (38.11)	103 (56.91)		Reference	
	Yes	294 (61.89)	144 (48.98)		0.88 (0.57, 1.34)	0.54
Depressive symptoms	No	451 (94.95)	228 (50.55)		Reference	
	Yes	24 (5.05)	19 (79.17)		**3.18 (1.18, 10.24)**	**0.03**

For older men, residing in rural areas (aOR 0.31, 95% CI 0.22 - 0.43) and maintaining an emotional connection with a sex partner during sexual intercourse (aOR 0.69, 95% CI 0.50 - 0.96) were associated with a reduced likelihood of reporting sexual response problems. In contrast, being 70 years or older (aOR 1.72, 95% CI 1.05 - 2.81), frequently experiencing fatigue (aOR 2.47, 95% CI 1.59 - 3.90), and poor sleeping quality (aOR 1.87, 95% CI 1.30 - 2.68) were associated with higher odds of reporting sexual response problems. In addition, moderate to severe anxiety symptoms (aOR 4.26, 95% CI 1.12 - 21.27), abnormal BMI (aOR 1.52, 95% CI 1.11 - 2.07), and seeking professional help for sex issues (aOR 1.58, 95% CI 1.14 - 2.21) were positively associated with reporting sexual response problems among older men.

Older women residing in rural areas (aOR 0.29, 95% CI 0.19 - 0.43) were less likely to report sexual response problems. Conversely, factors such as infrequent engagement in physical exercise (aOR 1.69, 95% CI 1.13 - 2.54), depressive symptoms (aOR 3.18, 95% CI 1.18 - 10.24), having an abnormal BMI (aOR 2.19, 95% CI 1.47 - 3.28), and being in a poor sex partner relationship (aOR 1.70, 95% CI 1.12 - 2.60) were associated with higher odds of reporting sexual response problems.

Discussion

Principal Results

Our study demonstrated a high prevalence of sexual response problems among older adults, with variations noted between men and women. In older men, sexual response problems correlate with advancing age. Our findings also linked sexual response problems with adverse physical health outcomes such as frequent experiencing fatigue, poor sleep quality, and abnormal BMI. In addition, we observed strong associations between sexual response problems and mental health issues, including anxiety and depressive symptoms. Moreover, poor sex partner relationships, sexual dissatisfaction, and lack of emotional connection during sex were also associated with sexual response problems. These findings significantly contribute to the existing literature, underscoring the importance of addressing sexual response problems within the domain of sexual health and enhancing our understanding of these issues among older adults.

XSL•FO
RenderX

Comparison With Previous Work

In this study, women reported a higher prevalence of sexual response problems compared to men, consistent with findings from other countries [12-14]. This discrepancy may be due to several factors, including physiological alterations, psychological elements, and sociocultural dynamics. With advancing age, women experience a significant decline in estrogen levels, especially after menopause [31]. Furthermore, older women are subject to various age-related psychological changes, such as concerns about body image, fears about aging, and self-consciousness about their sexual lives, which may lead to decreased sexual desire or arousal issues. In Chinese culture, the sexual lives of older adults are frequently neglected or deemed inappropriate, which may influence the perceptions and expectations of sex among older women, making them feel ashamed or uncomfortable in their sexual lives. This sex difference highlights the need for targeted interventions for women's sexual health. These interventions should address the multifaceted nature of sexual health in older women, combining physiological, psychological, and sociocultural interventions.

In this study, adverse physical health, such as abnormal BMI, frequent fatigue, and poor sleep quality, played a crucial role in sexual response problems. These findings were identified by evidence from high-income countries, which linked poor physical health to sexual response problems [32-34]. Obesity, characterized by an abnormal BMI, detrimentally impacts the reproductive system and sexual function [35-37]. Both men and women affected by obesity face a heightened risk of fertility challenges and sexual dysfunction [38]. Weight loss can reduce fatty tissue and diminish aromatase activity, leading to a relative increase in testosterone levels [39], which enhances sexual function in men. Besides, weight loss tends to improve sexual functioning for women and men [40]. Obesity significantly affects the hypothalamic-pituitary-gonadal axis in men, leading to diminished libido and erectile dysfunction [41]. The previous review highlighted that excess body weight negatively affected hormones contributing to sexual behavior, noting that adipose tissue facilitates the conversion of androgens to estrogens, further impacting sexual function [42]. Given the inverse association between body mass and sexual response problems, it is recommended that clinicians, both in general practice and in weight loss programs, should more fully address sexual response problems.

Poor sleep quality was associated with sexual response problems in previous studies [43,44], which was also demonstrated by this study. Poor sleep is closely linked to sexual dysfunction due to several physiological and psychological factors. Inadequate sleep can disrupt hormone production, notably reducing testosterone levels, which are essential for sexual desire and performance. Studies have shown that sleep deprivation can decrease testosterone production, leading to impaired sexual activity [45]. In addition, sleep disorders often contribute to stress, anxiety, and depression, all of which can diminish libido and sexual satisfaction [46]. Furthermore, chronic stress and anxiety can lead to sleep problems, which in turn may cause erectile dysfunction [47]. Sleep is fundamental to health, and its bidirectional relationship with sexual response difficulties necessitates that clinicians conduct thorough assessments to identify underlying causes of poor sleep, encourage patient-partner communication to alleviate psychological burdens, promote regular physical activity to enhance sleep quality, and recommend professional sleep therapy to improve physiological function when necessary.

Our study showed that lack of physical exercise was correlated with the occurrence of sex response problems among older women. Regular physical activity is a healthy practice that can mitigate the risk of sexual response problems [36,48]. These associations between sexual response problems and physical health underscore the importance of prioritizing sexual function within sexual health. Physical exercise and high-quality sleep are recommended, along with other lifestyle guidance, to improve sexual functioning in both men and women and to improve health across a range of domains.

Psychosocial factors, including symptoms of anxiety and depression, exhibited the strongest association with sexual response problems, as evidenced by findings from the Global Study of Sexual Attitudes and Behaviours (GSSAB) [49]. Other studies also linked the associations between mental health and sexual response problems, which highlighted that men with anxiety disorders are at a higher risk of developing erectile dysfunction [50,51]. Moreover, the physiological responses triggered by heightened anxiety levels may contribute to disruptions in sexual function [52]. This finding will capture the attention of practitioners exploring the causes and treatment of sexual problems in patients. In clinical practice, there should be heightened efforts to address sexual problems within integrated services of mental health and sexual health counseling, as well as in primary and secondary care.

Our findings underscored the importance of sex-partner relationships in the context of sexual response problems. Specifically, men who maintained an emotional connection with their sex partners during sex were less likely to report sex response problems, while women experiencing poor relationships with sex partners were more likely to encounter sexual response problems in older women. This underscores the role of brief emotional interactions during sexual activity in men's adaptation to sexual response problems. In contrast, long-term sexual relationships appear to be more influential in women's adjustment to sexual response problems. Existing studies have consistently underscored the significance of partner relationships for sexual response problems in women [53], emphasizing the impact of daily intimacies in relationships [9]. In contrast, emotional intimacy is presumed to play a significant role in maintaining sexual desire and partnered sexual activity for men [54,55]. In addition, emotional connection with a partner may foster open communication, trust, and mutual understanding, all essential to a satisfying sexual relationship. The variation in sexual response problems concerning sex partner relationships among older men and women underscores the importance of physicians considering sex differences when diagnosing and treating patients with sexual response problems. In future studies, the sex partner should be included and involved in the evaluation and management to achieve a better intimate relationship in an established couple and avoid sexual response problems.

XSL•FO
RenderX

Limitations

Our study has several limitations that should be considered. First, as a cross-sectional study, it cannot establish temporal order and causal direction. Second, the reliance on self-reporting in the questionnaires, especially for sensitive issues, introduced the possibility of recall biases. Third, the sexual response problems were reported by the participants experiencing them, and the differences in the sensitivity and understanding of the same sexual problem may exist among participants, potentially influenced by factors such as education, age, or living area. Finally, some analyses were based on small cell sizes, particularly for variables like anxiety and depressive symptoms, which may result in unstable estimates.

Conclusion

This study showed a substantial prevalence of sexual response problems among older adults, with women experiencing these problems more often than men. The study identified adverse physical health, poor mental health, and poor relationships with sex partners as factors contributing to increased sexual response problems among older adults. To address these concerns, health care professionals can implement interventions for older adults experiencing sexual response problems, such as enhancing physical health, supporting mental health, improving intimate relationships, and providing educational and cognitive-behavioral interventions. These insights drawn from the latest and representative SWELL study data enhance our understanding of sexual response problems among older adults and have the potential to promote overall health and well-being among the aging population in China.

Acknowledgments

We thank the investigation staff, all participants, and Ms Bowen Liang for her outstanding contribution to this study. This study was supported by the Natural Science Foundation of China International/Regional Research Collaboration Project (grant number 72061137001), and the Natural Science Foundation of China Excellent Young Scientists Fund (grant number 82022064). All funding parties had no role in the study design or data explanation.

Data Availability

The data collected in this study will not be publicly available. However, upon reasonable request, the corresponding author can be contacted for de-identified data.

Authors' Contributions

HZ conceived and designed the study in consultation with BL, CX, BW, XL, and XP. BW, XP, YW, HL, YL, XS, LO, GW, MY, JL, and XM contributed to data collection. BL, CX, BW, XM, WT, and JT contributed to data analysis and presentation. BL, CX, BW, XP, and XL drafted the manuscript, with all authors critically reviewing the paper. All authors approved the final report. BL, YC, and HZ contributed equally as co-corresponding authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Collinearity diagnostics of all potential variables for the multivariable logistic regression analysis (Men). Note that no value was more than 0.3, suggesting no collinearity between variables.
[PNG File, 345 KB - aging_v8i1e66772_app1.png]

Multimedia Appendix 2
Collinearity diagnostics of all potential variables for the multivariable logistic regression analysis (Women). Note that no value was more than 0.3, suggesting no collinearity between variables.
[PNG File, 341 KB - aging_v8i1e66772_app2.png]

References

1. Wang B, Peng X, Fu L, et al. Sexual function and correlates among adults aged 50+ years in China: findings from the sexual well-being (SWELL) study. J Am Geriatr Soc 2024 Apr;72(4):1122-1135. [doi: 10.1111/jgs.18789] [Medline: 38319006]
2. Ceyhan O, Ozen B, Simsek N, Dogan A. Sexuality and marital adjustment in women with hypertension in Turkey: how culture affects sex. J Hum Hypertens 2019 May;33(5):378-384. [doi: 10.1038/s41371-019-0181-3] [Medline: 30804462]
3. Galati MCR, Hollist CS, do Egito JHT, et al. Sexual dysfunction, depression, and marital dissatisfaction among Brazilian couples. J Sex Med 2023 Feb 27;20(3):260-268. [doi: 10.1093/jsxmed/qdac004] [Medline: 36763940]
4. Elterman DS, Bhattacharyya SK, Mafilios M, Woodward E, Nitschelm K, Burnett AL. The quality of life and economic burden of erectile dysfunction. Res Rep Urol 2021;13:79-86. [doi: 10.2147/RRU.S283097] [Medline: 33634039]

XSL•FO
RenderX

5. Lew-Starowicz Z, Czajkowska K. Prevalence of sexual dysfunctions and associated risk factors in Poland. Arch Med Sci 2022;18(4):1031-1040. [doi: 10.5114/aoms.2019.86794] [Medline: 35832714]
6. Atlantis E, Sullivan T. Bidirectional association between depression and sexual dysfunction: a systematic review and meta-analysis. J Sex Med 2012 Jun 1;9(6):1497-1507. [doi: 10.1111/j.1743-6109.2012.02709.x]
7. Hald GM, Graham C, Štulhofer A, Carvalheira A, Janssen E, Træen B. Prevalence of sexual problems and associated distress in aging men across 4 European countries. J Sex Med 2019 Aug;16(8):1212-1225. [doi: 10.1016/j.jsxm.2019.04.017] [Medline: 31255584]
8. Graham CA, Štulhofer A, Lange T, et al. Prevalence and predictors of sexual difficulties and associated distress among partnered, sexually active older women in Norway, Denmark, Belgium, and Portugal. Arch Sex Behav 2020 Nov;49(8):2951-2961. [doi: 10.1007/s10508-020-01742-7] [Medline: 32533517]
9. Parish WL, Laumann EO, Pan S, Hao Y. Sexual dysfunctions in urban china: a population-based national survey of men and women. J Sex Med 2007 Nov;4(6):1559-1574. [doi: 10.1111/j.1743-6109.2007.00596.x] [Medline: 17888072]
10. Yu Z, Niu J, Wang C. The prevalence and risk factors of sexual dysfunction in the elderly in Southern China. Int J Gen Med 2024;17:2355-2360. [doi: 10.2147/IJGM.S462124] [Medline: 38803552]
11. Spatz ES, Canavan ME, Desai MM, Krumholz HM, Lindau ST. Sexual activity and function among middle-aged and older men and women with hypertension. J Hypertens 2013 Jun;31(6):1096-1105. [doi: 10.1097/HJH.0b013e32835fdefa] [Medline: 23640604]
12. Mitchell KR, Mercer CH, Ploubidis GB, et al. Sexual function in Britain: findings from the third national survey of sexual attitudes and lifestyles (Natsal-3). Lancet 2013 Nov 30;382(9907):1817-1829. [doi: 10.1016/S0140-6736(13)62366-1] [Medline: 24286787]
13. Andresen JB, Graugaard C, Andersson M, Bahnsen MK, Frisch M. Sexual inactivity and dysfunction in Denmark: a project SEXUS study. Arch Sex Behav 2022 Nov;51(8):3669-3688. [doi: 10.1007/s10508-022-02348-x] [Medline: 36127559]
14. Quinta Gomes AL, Nobre PJ. Prevalence of sexual problems in Portugal: results of a population-based study using a stratified sample of men aged 18 to 70 years. J Sex Res 2014;51(1):13-21. [doi: 10.1080/00224499.2012.744953] [Medline: 23573897]
15. Ramírez-Santos J, Cristóbal-Cañadas D, Parron-Carreño T, Lozano-Paniagua D, Nievas-Soriano BJ. The problem of calculating the prevalence of sexual dysfunction: a meta-analysis attending gender. Sex Med Rev 2024 Mar 26;12(2):116-126. [doi: 10.1093/sxmrev/qead058] [Medline: 38336366]
16. Heidari M, Ghodusi M, Rezaei P, Kabirian Abyaneh S, Sureshjani EH, Sheikhi RA. Sexual function and factors affecting menopause: a systematic review. J Menopausal Med 2019 Apr;25(1):15-27. [doi: 10.6118/jmm.2019.25.1.15] [Medline: 31080785]
17. Hovén E, Fagerkvist K, Jahnukainen K, et al. Sexual dysfunction in young adult survivors of childhood cancer - a population-based study. Eur J Cancer 2021 Sep;154:147-156. [doi: 10.1016/j.ejca.2021.06.014] [Medline: 34273812]
18. Rahmani A, Afsharnia E, Fedotova J, et al. Sexual function and mood disorders among menopausal women: a systematic scoping review. J Sex Med 2022 Jul;19(7):1098-1115. [doi: 10.1016/j.jsxm.2022.03.614] [Medline: 35752457]
19. Calzo JP, Austin SB, Charlton BM, et al. Erectile dysfunction in a sample of sexually active young adult men from a U.S. cohort: demographic, metabolic and mental health correlates. J Urol 2021 Feb;205(2):539-544. [doi: 10.1097/JU.0000000000001367] [Medline: 32935616]
20. Bahrami Vazir E, Mohammad-Alizadeh-Charandabi S, Kamalifard M, Ghelichkhani F, Mohammadi A, Mirghafourvand M. The correlation between sexual dysfunction and intimate partner violence in young women during pregnancy. BMC Int Health Hum Rights 2020 Sep 14;20(1):24. [doi: 10.1186/s12914-020-00245-9] [Medline: 32928215]
21. Defeudis G, Mazzilli R, Tenuta M, et al. Erectile dysfunction and diabetes: a melting pot of circumstances and treatments. Diabetes Metab Res Rev 2022 Feb;38(2):e3494. [doi: 10.1002/dmrr.3494] [Medline: 34514697]
22. Jackson SE, Yang L, Koyanagi A, Stubbs B, Veronese N, Smith L. Declines in sexual activity and function predict incident health problems in older adults: prospective findings from the english longitudinal study of ageing. Arch Sex Behav 2020 Apr;49(3):929-940. [doi: 10.1007/s10508-019-1443-4] [Medline: 31432361]
23. Hamzehgardeshi Z, Sabetghadam S, Pourasghar M, Khani S, Moosazadeh M, Malary M. Prevalence and predictors of sexual distress in married reproductive-age women: A cross-sectional study from Iran. Health Sci Rep 2023 Sep;6(9):e1513. [doi: 10.1002/hsr2.1513] [Medline: 37655267]
24. Li L, Yu L. The influence of pension mode on the mental health of older adults—evidence from older adults in China. IJERPH 2021 Dec 23;19(1):119. [doi: 10.3390/ijerph19010119]
25. Fang EF, Xie C, Schenkel JA, et al. A research agenda for ageing in China in the 21st century (2nd edition): Focusing on basic and translational research, long-term care, policy and social networks. Ageing Res Rev 2020 Dec;64:101174. [doi: 10.1016/j.arr.2020.101174] [Medline: 32971255]
26. Wang B, Peng X, Liang B, et al. Sexual well-being among older adults in China (SWELL): protocol for a multicenter cross-sectional study. BMJ Open 2023 Jan 30;13(1):e067338. [doi: 10.1136/bmjopen-2022-067338] [Medline: 36717139]
27. Laumann EO, Gagnon JH, Michael RT, Michaels S. The Social Organization of Sexuality: Sexual Practices in the United States: University of Chicago press; 2000.
28. Dadvand P, Bartoll X, Basagaña X, et al. Green spaces and General Health: roles of mental health status, social support, and physical activity. Environ Int 2016 May;91:161-167. [doi: 10.1016/j.envint.2016.02.029]

XSL•FO
RenderX

29. Whittle HJ, Sheira LA, Wolfe WR, et al. Food insecurity is associated with anxiety, stress, and symptoms of posttraumatic stress disorder in a cohort of women with or at risk of HIV in the United States. J Nutr 2019 Aug 1;149(8):1393-1403. [doi: 10.1093/jn/nxz093] [Medline: 31127819]
30. Wang B, Peng X, Fu L, et al. Correlates of sexual lifestyles among older adults living with HIV in China: findings from the sexual well-being (SWELL) study. Infect Microb Dis 2023;5(3):145-151. [doi: 10.1097/IM9.0000000000000125]
31. Soysal P, Smith L. Sexual activity and successful aging. Eur Geriatr Med 2022 Dec;13(6):1235-1238. [doi: 10.1007/s41999-022-00694-6]
32. Lindau ST, Schumm LP, Laumann EO, Levinson W, O'Muircheartaigh CA, Waite LJ. A study of sexuality and health among older adults in the United States. N Engl J Med 2007 Aug 23;357(8):762-774. [doi: 10.1056/NEJMoa067423] [Medline: 17715410]
33. Field N, Mercer CH, Sonnenberg P, et al. Associations between health and sexual lifestyles in Britain: findings from the third national survey of sexual attitudes and lifestyles (Natsal-3). Lancet 2013 Nov 30;382(9907):1830-1844. [doi: 10.1016/S0140-6736(13)62222-9] [Medline: 24286788]
34. Lee DM, Nazroo J, O'Connor DB, Blake M, Pendleton N. Sexual health and well-being among older men and women in England: Findings from the English longitudinal study of ageing. Arch Sex Behav 2016 Jan;45(1):133-144. [doi: 10.1007/s10508-014-0465-1] [Medline: 25624001]
35. Kolotkin RL, Zunker C, Østbye T. Sexual functioning and obesity: a review. Obesity (Silver Spring) 2012 Dec;20(12):2325-2333. [doi: 10.1038/oby.2012.104] [Medline: 22522887]
36. Mollaioli D, Ciocca G, Limoncin E, et al. Lifestyles and sexuality in men and women: the gender perspective in sexual medicine. Reprod Biol Endocrinol 2020 Feb 17;18(1):10. [doi: 10.1186/s12958-019-0557-9] [Medline: 32066450]
37. Faubion SS, Fairbanks F, Kuhle CL, et al. Association between body mass index and female sexual dysfunction: a cross-sectional study from the data registry on experiences of aging, menopause, and sexuality. J Sex Med 2020 Oct;17(10):1971-1980. [doi: 10.1016/j.jsxm.2020.07.004] [Medline: 32771351]
38. Law S, Wan K, Yang W, Global Obesity C. Effects of bariatric surgery on sexual function and fertility: a narrative review. Obes Rev 2024 Aug;25(8):e13757. [doi: 10.1111/obr.13757] [Medline: 38689132]
39. Kumagai H, Yoshikawa T, Zempo-Miyaki A, et al. Vigorous physical activity is associated with regular aerobic exercise-induced increased serum testosterone levels in overweight/obese men. Horm Metab Res 2018 Jan;50(1):73-79. [doi: 10.1055/s-0043-117497] [Medline: 28934816]
40. McNabney SM, Gletsu-Miller N, Rowland DL. Sexual function and satisfaction in the context of obesity. Curr Diab Rep 2023 Nov;23(11):315-327. [doi: 10.1007/s11892-023-01523-1] [Medline: 37632680]
41. Santamaria A, Amighi A, Thomas M, Goradia R, Choy J, Hehemann MC. Effect of surgical, medical, and behavioral weight loss on hormonal and sexual function in men: a contemporary narrative review. Ther Adv Urol 2024;16:17562872241279648. [doi: 10.1177/17562872241279648] [Medline: 39285942]
42. Du Plessis SS, Cabler S, McAlister DA, Sabanegh E, Agarwal A. The effect of obesity on sperm disorders and male infertility. Nat Rev Urol 2010 Mar;7(3):153-161. [doi: 10.1038/nrurol.2010.6] [Medline: 20157305]
43. Smith L, Grabovac I, Veronese N, et al. Sleep quality, duration, and associated sexual function at older age: findings from the english longitudinal study of ageing. J Sex Med 2019 Mar;16(3):427-433. [doi: 10.1016/j.jsxm.2019.01.005] [Medline: 30773496]
44. Seftel AD. Re: Sleep quality, duration, and associated sexual function at older age: findings from the english longitudinal study of ageing. Journal of Urology 2019 Sep;202(3):443-443. [doi: 10.1097/01.JU.0000574448.61263.77]
45. Andersen ML, Tufik S. The association between sleep disturbances and erectile dysfunction during the COVID-19 pandemic. Sex Med Rev 2022 Apr;10(2):263-270. [doi: 10.1016/j.sxmr.2021.12.001] [Medline: 35042659]
46. Pigeon WR, Youngren W, Carr M, Bishop TM, Seehuus M. Relationship of insomnia to sexual function and sexual satisfaction: findings from the sleep and sex survey II. J Psychosom Res 2023 Dec;175:111534. [doi: 10.1016/j.jpsychores.2023.111534] [Medline: 37871507]
47. Zhang F, Xiong Y, Qin F, Yuan J. Short sleep duration and erectile dysfunction: a review of the literature. Nat Sci Sleep 2022;14:1945-1961. [doi: 10.2147/NSS.S375571] [Medline: 36325277]
48. Allen MS, Walter EE. Health-related lifestyle factors and sexual dysfunction: a meta-analysis of population-based research. J Sex Med 2018 Apr;15(4):458-475. [doi: 10.1016/j.jsxm.2018.02.008] [Medline: 29523476]
49. Laumann EO, Nicolosi A, Glasser DB, et al. Sexual problems among women and men aged 40-80 y: prevalence and correlates identified in the Global Study of Sexual Attitudes and Behaviors. Int J Impot Res 2005;17(1):39-57. [doi: 10.1038/sj.ijir.3901250] [Medline: 15215881]
50. Velurajah R, Brunckhorst O, Waqar M, McMullen I, Ahmed K. Erectile dysfunction in patients with anxiety disorders: a systematic review. Int J Impot Res 2022 Mar;34(2):177-186. [doi: 10.1038/s41443-020-00405-4] [Medline: 33603242]
51. Ciocca G, Di Lorenzo G, Comite G, et al. Sexual functioning in male patients suffering from depression and anxiety disorders. Eur psychiatr 2017 Apr;41(S1):s848-s849. [doi: 10.1016/j.eurpsy.2017.01.1681]
52. Yafi FA, Jenkins L, Albersen M, et al. Erectile dysfunction. Nat Rev Dis Primers 2016 Feb 4;2:16003. [doi: 10.1038/nrdp.2016.3] [Medline: 27188339]

XSL•FO
RenderX

53. Zhang H, Fan S, Yip PSF. Sexual dysfunction among reproductive-aged Chinese married women in Hong Kong: prevalence, risk factors, and associated consequences. J Sex Med 2015 Mar;12(3):738-745. [doi: 10.1111/jsm.12791] [Medline: 25470214]
54. van Lankveld J, Jacobs N, Thewissen V, Dewitte M, Verboon P. The associations of intimacy and sexuality in daily life: Temporal dynamics and gender effects within romantic relationships. J Soc Pers Relat 2018 May;35(4):557-576. [doi: 10.1177/0265407517743076] [Medline: 29899585]
55. Štulhofer A, Jurin T, Graham C, Janssen E, Træen B. Emotional intimacy and sexual well-being in aging European couples: a cross-cultural mediation analysis. Eur J Ageing 2020 Mar;17(1):43-54. [doi: 10.1007/s10433-019-00509-x] [Medline: 32158371]

Abbreviations

AIC: Akaike Information Criterion
aORs: adjusted odds ratios
GAD-7: generalized anxiety disorder-7
GSSAB: Global Study of Sexual Attitudes and Behaviours
PHQ-9: Patient Health Questionnaire
SQS: sleep quality scale
SRPs: Sexual response problems
SWELL: Sexual Well-being

Edited by G Seçkin; submitted 23.09.24; peer-reviewed by H Song, J Bian, T Tian; revised version received 17.01.25; accepted 23.02.25; published 01.05.25.

Please cite as:
Liang B, Xu C, Wang B, Li X, Peng X, Wang Y, Li H, Lu Y, Shen X, Ouyang L, Wu G, Yu M, Liu J, Meng X, Cai Y, Zou H
Sexual Response Problems and Their Correlates Among Older Adults From the Sexual Well-Being (SWELL) Study in China: Multicenter Cross-Sectional Study
JMIR Aging 2025;8:e66772
URL: https://aging.jmir.org/2025/1/e66772
doi:10.2196/66772

©Bingyu Liang, Chen Xu, Bingyi Wang, Xinyi Li, Xin Peng, Ying Wang, Hui Li, Yong Lu, Xiaopei Shen, Lin Ouyang, Guohui Wu, Maohe Yu, Jiewei Liu, Xiaojun Meng, Yong Cai, Huachun Zou. Originally published in JMIR Aging (https://aging.jmir.org), 1.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Older Adults' Perspectives on Participating in a Synchronous Online Exercise Program: Qualitative Study

Giulia Coletta[1], PhD; Kenneth S Noguchi[2], PhD; Kayla Beaudoin[1], BSc; Angelica McQuarrie[1], MSc; Ada Tang[2], PhD; Rebecca Ganann[3], PhD; Stuart M Phillips[1], PhD; Meridith Griffin[4,5], PhD

[1]Department of Kinesiology, McMaster University, Hamilton, ON, Canada

[2]School of Rehabilitation Science, McMaster University, Hamilton, ON, Canada

[3]School of Nursing, McMaster University, Hamilton, ON, Canada

[4]Department of Health, Aging and Society, McMaster University, Kenneth Taylor Hall, 241, 1280 Main Street West, Hamilton, ON, Canada

[5]Gilbrea Centre for Studies in Aging, McMaster University, Hamilton, ON, Canada

Corresponding Author:
Meridith Griffin, PhD
Department of Health, Aging and Society, McMaster University, Kenneth Taylor Hall, 241, 1280 Main Street West, Hamilton, ON, Canada

Abstract

Background: Older adults face several barriers to exercise participation, including transportation, lack of access, and poor weather conditions. Such barriers may influence whether older adults meet the Canadian 24-Hour Movement Guidelines. Recently, older adults have adopted technology for health care and are increasingly using digital health technologies to improve their access to care. Therefore, technology may be a valuable tool to reduce barriers to exercise and increase exercise participation rates within this population.

Objective: This study aimed to explore older adults' perceptions and experiences of exercise, in general, and specifically related to our synchronous online exercise program for community-dwelling older adults.

Methods: A total of 3 registered kinesiologists and 1 physiotherapist with experience working with older adults delivered an 8-week, thrice-weekly synchronous online group-based exercise program for older adults in 3 cohorts. The program focused on strength, balance, and aerobic activity. Following the program, a qualitative study with interpretive descriptive design was conducted to explore participants' perceptions and experiences. Participants were invited to take part in a 30-minute, one-on-one semistructured interview via Zoom with a research team member. Interview data were thematically analyzed to identify common themes.

Results: A total of 22 older adults (16 women, 6 men; mean age 70, SD 4 years) participated in interviews. Three themes were identified as follows: (1) health, exercise, and aging beliefs; (2) the pandemic interruption and impacts; and (3) synchronous online exercise programs attenuate barriers to exercise. Participants discussed their exercise beliefs and behaviors and their desire to safely and correctly participate in exercise. Older adults found that their physical activity was curtailed, routines disrupted, and access to in-person exercise programs revoked due to the pandemic. However, many suggested that our synchronous online exercise program was motivational and attenuated commonly reported environmental barriers to participation, such as transportation concerns (eg, time spent traveling, driving, and parking), accessibility and convenience by participating at a location of their choice, and removing travel-related concerns during poor weather conditions.

Conclusions: Given these reported experiences, we posit that synchronous online exercise programs may help motivate and maintain adherence to exercise programs for older adults. These findings may be leveraged to improve health outcomes in community-dwelling older adults.

Trial Registration: ClinicalTrials.gov NCT04627493; https://clinicaltrials.gov/study/NCT04627493

(JMIR Aging 2025;8:e66473) doi:10.2196/66473

KEYWORDS

exercise; older adults; qualitative study; qualitative; experience; attitude; opinion; perception; perspective; interview; internet; kinesiology; physiotherapy; synchronous; online; home-based; gerontology; geriatric; older; aging; physical activity

XSL•FO
RenderX

Introduction

Physical activity (PA) guidelines are not met by ≈85% of older adults in Canada [1,2]. The Canadian Society for Exercise Physiology recommends that older adults participate in 150 minutes of moderate-to-vigorous PA and 2 bouts of strengthening exercise each week [3], similar to the World Health Organization guidelines [4]. If PA guidelines are not met, older adults are at greater risk of adverse health outcomes, including falls, cardiovascular disease, mobility limitations, hospitalization, and mortality [4].

Older adults commonly report barriers to participating in PA and exercise. A systematic review by Spiteri et al [5] determined that environmental factors, resources, lack of assistance in managing change, and social influences are among the common barriers to PA participation for older adults aged 65 - 70 years. Baeret et al [6] reported similar barriers in their systematic review of 44 journal papers describing motivators and barriers to PA in older adults. Common environmental barriers for older adults include transportation, geographic accessibility, affordability, and weather [5-7]. Transportation barriers can be attributed to a lack of transportation to facilities, challenges navigating parking, and poor reliability of affordable transportation [5,6]. Lack of access to exercise facilities, the cost of exercise programming, and distance to programs contribute to environmental barriers [5,6]. Lastly, poor weather conditions reduce the likelihood of adults participating in recreational activities or attending recreation centers [7]. Among Canadians, 64% are less active in the winter compared with the summer [8]. Developing exercise programs to mitigate common environmental barriers may aid in improving older adults' abilities to participate in exercise and meet PA guidelines.

Recently, older adults have adopted technology for health care to improve access to care [9,10]. Technology may, therefore, be a valuable mode of delivery to reduce barriers to exercise and increase participation rates within this population [11]. Current evidence suggests that online exercise programs improve physical function and cognition in older adults, although these programs are exclusively asynchronous, web-based, or interactive exercise-based video games [12]. Older adults have reported not enjoying asynchronous exercise programs as they lack social connectivity, accountability, and adherence, making it difficult to stay motivated [12,13]. Recently, there has been a greater uptake of asynchronous and synchronous programming due to the COVID-19 pandemic [14-16]. A scoping review by Dagenais et al [16] described the nature and extent of the existing literature on online exercise programming among older adults and found 17 studies using asynchronous (9/17, 53%), synchronous (5/17, 29%), or both asynchronous and synchronous programming (3/17, 18%). For older adults, group-based synchronous programs have demonstrated improvements in muscle strength, balance, physical function, aerobic capacity, and quality of life [14]. Therefore, gaining insights into older adults' perceptions and experiences with synchronous programming is important.

We developed and delivered a synchronous online exercise program to help older adults engage in exercise while in-person programming was temporarily suspended [17]. The synchronous program was designed and delivered by registered kinesiologists and physiotherapists. The main aim of the parent trial was to investigate the preliminary effectiveness and feasibility of the synchronous online exercise program through a mixed method pilot randomized controlled trial (RCT) [17]. The aim of this paper was to explore older adult participants' perceptions and experiences of exercise, in general, and of our synchronous online community exercise program.

Methods

Design

This qualitative study was part of an 8-week community-based, pilot RCT [17]. An interpretive descriptive design explored community-dwelling older adults' experiences and perceptions of exercise, in general, and of the synchronous online exercise program [18]. Interpretive description offers a flexible approach to analyzing qualitative data within medical education research, as it can explore individuals' experiences while producing practical clinical outcomes [19]. The pilot RCT compared community-dwelling older adults aged 65 to 80 years participating in a synchronous online exercise program (ACTIVE) delivered on Zoom with a waitlist control group (CON). A total of 3 registered kinesiologists and 1 physiotherapist with experience working with older adults conducted the intervention online in 3 cohorts at the Physical Activity Centre of Excellence at McMaster University in Ontario, Canada. The intervention and all data collection, including the qualitative interviews, were conducted remotely via Zoom. Zoom meets McMaster University's privacy legislation. The study duration was from January 2021 to May 2022. The study was reported in accordance with the consolidated criteria for reporting qualitative research [20]. For the purpose of this paper, the qualitative findings alone will be reported.

Ethical Considerations

This study was approved by the Hamilton Integrated Research Ethics Board (#11429) and registered at ClinicalTrials.gov (NCT04627493). Participants were informed of the study content and procedures both verbally and in writing. Informed written and oral consent was obtained from all participants. All participants had the right to withdraw from the study at any time without any adverse consequences. All data were anonymized. Participants did not receive compensation for their participation.

Participants

ACTIVE and CON group participants enrolled in our 8-week synchronous online program were invited via email to participate in an individual qualitative interview after completing the exercise program. The eligibility criteria for the full trial have been previously reported [17]. Briefly, they were healthy older community-dwelling women and men between the ages of 65 and 80 years, had access to the internet via a personal smartphone, tablet, or computer, and participated in ≤150 minutes of moderate to vigorous PA per week. To be eligible for the qualitative portion of the study, participants must have

XSL•FO
RenderX

remained enrolled in the study and needed to participate in the online exercise program. Interested participants were contacted within 2 weeks of the end of the synchronous online exercise program to schedule an interview time. With a purposeful subsample of interested participants from the pilot RCT, we aimed to conduct between 20 and 24 interviews, as data saturation typically occurs within this range [21].

Intervention

The 8-week intervention consisted of thrice-weekly 60-minute synchronous online exercise classes (totaling 24) delivered online via Zoom by health professionals (ie, registered kinesiologists and physiotherapists). Briefly, the group-based classes began with a 5-minute warm-up, 50 minutes of progressive strength (25 min), aerobic (20 min), and balance training (5 min), and a 5-minute cool-down [17]. The physiotherapist (n=1) and registered kinesiologists (n=3) were trained to work with older adults and adjusted exercise intensity weekly to meet participants' abilities. The instructors and research team members supervised and moderated each class to ensure participant safety and assist with technological difficulties. All interactions between participants and the research team were conducted remotely via Zoom; no pre-established relationships existed prior to the study.

Qualitative Data Collection

Semistructured, open-ended one-on-one interviews were conducted via Zoom by a study investigator (GC, female, PhD Candidate) with experience in qualitative research [22]. A qualitative interview guide (Multimedia Appendix 1) was developed with a qualitative methodology expert (MG) to generate discussion around participants' experiences and perceptions of exercise, in general, and with our synchronous online exercise program [22]. Questions were not provided to participants before the interview. However, participants were aware of the interviewer's personal goals (ie, PhD research) for conducting the research. Transcripts were transcribed verbatim from Zoom audio recordings, checked for accuracy, and not returned to participants.

Data Analysis

The interviews were thematically analyzed using the computer software Dedoose (version 9.0.17; SocioCultural Research Consultants, LLC) [23]. Two authors, GC and KSN, independently coded the first 2 transcripts and discussed initial patterns and themes to inform the preliminary coding book [24], which was shared and refined with input from the qualitative expert. The coding book was used to review subsequent transcripts and iteratively revised to reflect new concepts. Themes and subthemes were inductively developed and refined with the research team. At least 2 interviewers (GC, KSN, and MG) discussed any disagreements in transcription or content to reach a consensus [23,25].

Results

Participants' Characteristics

Of the 32 who participated in the pilot RCT, which determined the preliminary effectiveness of the synchronous online exercise program [17], 22 older women and men participated in semistructured interviews. Ten participants declined or did not respond to the invitation to participate in this study. Table 1 presents the descriptive characteristics of the sample. There were no major differences between groups for any variable. The majority of participants were women (25/32, 78%) and had completed postsecondary education (29/32, 91%). All interviews were approximately 30 minutes in duration by design.

XSL•FO
RenderX

Table . Descriptive characteristics for the ACTIVE, control group (CON), and no-interview participants (N=32).

	ACTIVE[a] (n=13)	CON[b] (n=9)	No interview (n=10)
Women, n (%)	9 (69)	7 (78)	9 (90)
Age (years), mean (SD)	70 (4)	70 (4)	74 (4)
Race, n (%)			
Caucasian	12 (92)	9 (100)	9 (90)
Indigenous	1 (8)	0 (0)	0 (0)
Asian	0 (0)	0 (0)	1 (10)
Highest level of education earned, n (%)			
Some school or high school diploma	2 (15)	1 (11)	1 (10)
Some college, vocational, or training school	1 (8)	1 (11)	0 (0)
College graduate or bachelor's degree	5 (38)	5 (56)	7 (70)
Postgraduate training or degree	5 (38)	2 (22)	2 (20)
Partner status, n (%)			
Never married	0 (0)	2 (22)	0 (0)
Divorced or separated	4 (31)	1 (11)	1 (10)
Widowed	3 (23)	1 (11)	3 (30)
Presently married	6 (46)	5 (56)	6 (60)
Living arrangement, n (%)			
Wife, husband, or partner	6 (46)	5 (56)	6 (60)
Children	1 (8)	0 (0)	0 (0)
Friends	0 (0)	1 (11)	0 (0)
Living alone	5 (38)	3 (33)	4 (40)
Other	1 (8)	0 (0)	0 (0)
Physically active prior to the pandemic, n (%)	7 (54)	5 (56)	6 (60)

[a]ACTIVE: exercise group.

[b]CON: waitlist control group.

Qualitative Findings

Overview

Our thematic analysis identified 3 main themes with respect to participants' perceptions of exercise in general and with the synchronous online exercise program: health, exercise, and aging beliefs; the pandemic interruption and impacts; and synchronous online exercise programs attenuate barriers to exercise (Textbox 1). Participants described various subthemes within these 3 main themes, including their exercise beliefs and behaviors and their desire to safely and correctly participate in exercise as an older adult. They also reflected upon the impact of the pandemic on their PA, routine, and lack of access to exercise programs. Participants discussed how the online synchronous format contributed to their motivation to exercise, reduced concerns about transportation, enabled participation from their desired training location, and allowed them to avoid poor weather conditions. Our findings provide greater insight into older adults' perceptions and experiences of exercise, in general, through the themes of health, exercise, and aging beliefs and the impact of the pandemic on older adults' PA. We also provide insight into the delivery of synchronous online community exercise programs and how common barriers to exercise may be attenuated.

XSL•FO
RenderX

Textbox 1. Qualitative themes and subthemes.

Health, exercise, and aging beliefs

- Exercise beliefs and behaviors.
- Desire to safely and correctly participate in exercise as an older adult.

The pandemic interruption and impacts

- Physical activity curtailed.
- Routines disrupted.
- Access to exercise programs revoked.

Synchronous online exercise programs attenuate barriers to exercise

- Synchronous programming is motivating.
- Removed concerns about transportation.
- Ease of participation from desired training location.
- Removed poor weather condition concerns.

Health, Exercise, and Aging Beliefs

Participants believed in the benefits of exercise and sought ways to become more physically active. Many had participated in PA throughout their lives, and they valued the importance of the health benefits of exercise. Three participants described their beliefs about the benefits of exercise and its importance to health and longevity.

> *Yeah, it's like exercise should be long-term or lifetime commitment… it can really benefit you through your own lifetime.* [CON1]
>
> *I have the attitude exercise is important for everybody, but it's especially important as we get older.* [CON2]
>
> *I need physical exercise. My body feels better.* [CON3]

Two participants discussed the reasons for seeking out opportunities to increase their PA.

> *If I joined the program at least it would give me a reason to get up and get started in exercise.* [ACTIVE4]
>
> *Since I've retired, I'm trying to find some kind of routine that would keep me active, and [help with] losing weight and trying to get healthy for longevity.* [ACTIVE5]

Participants expressed concern about their declining fitness levels. Three participants described their desire to continue participating in meaningful activities and believed that becoming more active would help them maintain or improve their health.

> *I was getting too unfit basically… If I think in terms of basic fitness, it would be functional fitness. Can I bend down to get something on the bottom shelf? Can I reach the top shelf? Can I get in and out of a bathtub?… Activities of daily living, but all of them, like the reaching bending stretching [exercise would help with this].* [CON6]
>
> *Because I want to be active in my older age. I don't want to be, you know, stuck in a chair and very limited in what I can do. I still want to travel… You know I want to be mobile.* [ACTIVE7]
>
> *What does motivate me is as I'm aging, I want to get healthier… I want to maybe see my grandkids get married, you know, things like that, so as I age, I get more motivation.* [ACTIVE8]

One participant described their motivation to become more physically active for health and vanity.

> *It's for health, and to be honest with you, it partly is vanity, too.* [CON2]

The participants in our study had prior beliefs regarding exercise and believed in the long-term benefits of exercise. Their motivation stemmed from their desire to improve their health and continue to participate in activities that are meaningful to them. While participants expressed a desire to be active, many disclosed that they lacked knowledge of how to exercise. One participant described that as an older woman, strength training is not an easily accessible exercise modality. Aerobic is traditionally the commonly prescribed type of exercise, but older adults, including women, are interested in learning about strength training.

> *I think as an older person and a woman, I'm less likely to do [strength training] on my own. I can go out and walk and do aerobic stuff, but it's the resistance work that I would be interested in.* [ACTIVE5]

Three participants described concerns over incorrect form, potentially leading to injury. Older adults discuss a lack of confidence, fear of injury, and desire for more supervision when exercising.

> *I think as we get older, certainly doing it the proper way even becomes more important than when you were young.* [ACTIVE9]

XSL•FO
RenderX

> *It's not like I didn't want to try and join some of those [traditional] gyms sometimes, but I didn't feel as confident about it because of some of the medical part [injuries].* [CON3]

> *But there's not a lot of supervision happening there [traditional gyms]… the opportunity to do things incorrectly kind of opens you for doing something wrong and hurting yourself.* [ACTIVE7]

Participants in this study desired to become more physically active and believed in the benefits of exercise. Their motivations for exercising varied from wanting to continue with their activities of daily living to longevity and vanity. However, our participants suggested that traditional gym settings do not offer adequate supervision for older adults as they learn to participate in exercise safely.

The Pandemic Interruption and Impacts

The COVID-19 pandemic interrupted participants' daily routines and limited access to PA during this study. The first cohort of participants began the study during the second wave (Beta variant) of the COVID-19 pandemic and completed the study during the third wave (Gamma variant). The second cohort of participants began the study during the third wave of the COVID-19 pandemic and completed the study as the fourth wave (Delta variant) began. The third cohort of the study started and completed the study during the fifth wave (Omicron variant) of the COVID-19 pandemic. Participants voiced several concerns about their own inactivity levels and how these were affected by COVID-19. For example, 3 participants spoke of the secondary consequences of the pandemic resulting in unhealthy habits.

> *I had joined the gym, and then COVID happened. So, then I've put on another 30 pounds, so now I'm 70 pounds over my optimum weight…. I don't want to risk it right now because of the COVID stuff. Yeah, and I'm not doing massages and not doing physiotherapy, so you know that's kind of hindered things too.* [ACTIVE10]

> *I got out first thing in the morning; it was a part of a routine during the pandemic. I lost all of that; consequently, I put on weight… when the pandemic hit, I literally turned into a couch potato.* [CON2]

> *I was really surprised at how de-conditioned I became over the course of the pandemic; I mean, it wasn't great before, but it became a lot worse during the pandemic.* [ACTIVE11]

The pandemic's secondary consequences left participants more sedentary during this period and disrupted their routines. For example, 2 participants described the challenge to keep track of weekdays.

> *COVID really slowed us down, and I didn't know what day of the week it was.* [CON1]

> *Well, yeah, until the program that you gave us, I was getting lost on what day of the week it was. At least I knew when it was Monday, Wednesday, Friday.* [ACTIVE10]

In addition to disruptions to participants' daily routines, the pandemic revoked access to exercise programs. The public health-mandated lockdowns resulted in canceled in-person exercise programs, further limiting access to PA and exercise. Prior to the pandemic, participants had been engaging in exercise programs and were members of local gyms. COVID-19 interrupted their participation in programming, and many found it difficult to motivate themselves to exercise.

> *… once COVID hit, other classes [exercise programs] closed, it was very difficult… it's almost impossible to continue [exercising]. And I'm not very good at self-motivating.* [CON1]

> *Okay, I go to the gym quite often, like maybe four times or five times a week and, of course, the gyms are closed [COVID], and I found it hard to get motivated to do stuff on my own.* [ACTIVE4]

> *I just started at the Mac gym in the Senior's Program in February, so I was just sort of getting into the three-times-a-week schedule when, of course, we stopped in March because of the pandemic.* [ACTIVE11]

The pandemic interrupted many participants' routines, including canceled exercise programs and memberships and curtailing PA.

Synchronous Online Exercise Programs Attenuate Barriers to Exercise

Participants reported that our synchronous online exercise program removed several common environmental barriers to exercise and motivated them to become more physically active. For example, one participant discussed enjoying exercising from home while still connecting to people in real time.

> *The fact that we got to do it at home… that was cool. I think, maybe something like that would probably motivate me again that there's interaction, it's real people, and it's done in real-time… I actually liked having it [online].* [ACTIVE8]

Another participant spoke about how synchronous programming, in general, allows them to continue to exercise when they are not comfortable attending in-person programming.

> *Yeah, if you have sniffles, you might still want to do the exercises, but you don't want to spread the disease…* [ACTIVE10]

A third participant described how our synchronous online exercise program was advantageous compared with in-person programming.

> *So, it's just me moving around in the room, not, you know, bumping into other people like in a small gym… We're always overheated. Here I have a fan blowing on me, and I don't have somebody complaining that wind is blowing on them because they don't like wind.* [ACTIVE10]

Our synchronous online exercise program motivated participants to exercise as it allows individuals to remain connected with others while participating safely and easily from their own homes. Participants described enjoying the ease of participating

XSL•FO
RenderX

in our program from home since it removed their concerns about transportation and reduced time spent traveling to and from the exercise program, which allowed more time to engage in other activities. Participants discussed their concerns about driving, traffic, and parking when attending in-person exercise programs at the university or recreational facilities. For example, one participant referred to the difficulty of driving and finding their way around campus.

> *… I didn't like driving to McMaster, and then you have to figure out where you're going… that was just a part that I don't care about… So, I like local these days, and the fact that I could do that on Zoom is great. I think it's the best thing ever.* [CON12]

Another participant spoke about not having to drive (or otherwise find their way) to the session and how it streamlined their exercise experience.

> *No, that was actually really, really good. So, because it's COVID, you don't have to worry about being in contact. But yes, I don't like the concept of having to drive somewhere to exercise. It just defeats the whole purpose… I liked that.* [ACTIVE5]

A third participant discussed the inefficiency of driving to the gym.

> *I always used to say if people just walk to the gym or bike to the gym they wouldn't even have to go in the gym. So yes, you've cut out that timepiece of having to go somewhere to do something.* [ACTIVE7]

Our online exercise program eliminated transportation concerns and was described as motivational as it allowed individuals to participate from their desired training location, including their homes and while on holiday. For example, one participant described the ease of connecting to our program while away from home.

> *I was out of town, okay? I ended up at my nephew's place, and he works the night shift. So, I didn't want to disturb him inside, so I went outside, and they had a nice strong Wi-Fi connection, and it worked out really good.* [CON1]

Another participant enjoyed the flexibility of our program, allowing them to maintain their routine while away from home.

> *We'll sign up for that. We can do it at the lake as well… With this type of activity, you can do it in your living room, you can do it out on the deck, and you can do it wherever you want. I think it's great that there's more flexibility, that you can take the program to where you are, rather than you go[ing] there.* [ACTIVE9]

Two participants discussed how offering the exercise program through a synchronous videoconference platform enabled participation due to the ease of access.

> *It [Zoom] made me kind of want to go on… how can you argue about getting up five minutes before it starts? And, yeah, it's really, it's fantastic.* [ACTIVE13]

> *Having it be virtual means that you don't have to worry about traffic and parking and driving.* [CON6]

The synchronous program allowed flexibility in how and where participants engaged with our program, easing participants' concerns about traveling in poor weather conditions. Older adults described their fears and worries about navigating poor weather conditions (eg, extreme heat, rain, snow, and ice) on foot or by car to attend exercise programs or to be physically active. One participant described enjoying not needing to leave the house during the winter months to exercise.

> *Yeah, absolutely. In the winter, too, you didn't have to, you know, shiver and so on.* [CON14]

Two participants spoke of the inconvenience of navigating the weather conditions when exercising.

> *To go warm up my car, put my boots on, then go to the gym, take my boots off... It just doesn't make sense to me.* [ACTIVE8]

> *It was helpful that I didn't have to go running and try to park somewhere and find something [parking] and go through a snowbank.* [CON3]

Another participant discussed barriers when leaving the house when weather conditions are poor and how exercising from home helped to ensure they could still exercise.

> *… if it's like a torrential rain or snowstorm or ice, I don't go out on ice. Yeah, you know, then you would do it from home.* [6ACTIVE10]

Participants described poor weather conditions as negatively impacting their PA levels. Overall, they enjoyed the ease of the videoconference platform and participating in a synchronous online exercise program.

Discussion

Principal Findings

This study provides novel insights into synchronous online exercise programs for older adults. Specifically, participants identified themes including health, exercise, and aging beliefs; the pandemic interruption and impacts; and synchronous online exercise programs attenuating barriers to exercise. Our findings suggested that participants in this study believed in the benefits of exercise and wanted to learn how to exercise safely. The pandemic curtailed their PA, disrupted their routines, and revoked access to exercise programs, leaving them seeking new approaches to exercise. We found that our synchronous online exercise program was motivational and may reduce commonly reported environmental barriers, including removing concerns about transportation (eg, time spent traveling, driving, and parking), improving access through ease of participation from desired training locations, and removing poor weather condition concerns. Therefore, participants expressed a preference for synchronous online delivery of exercise programs, which may be a valuable option for motivating older adults to become physically active while reducing common environmental barriers.

Our findings suggested that participants believed in the health benefits of exercise. Participants described motivation to

XSL•FO
RenderX

exercise due to their desire to be healthy, mobile, and participate in meaningful activities as they aged. Notably, the majority of participants in our study were previously active and had prior experience with exercise, which may influence their perceptions and experiences with exercise. Our findings are similar to findings from a qualitative study by Harrison et al [26] focused on understanding the barriers and motivators of 58 older adults (49/58, 84% females; 60‐85 years) residing in an urban community in Washington, DC, which suggested that the main benefits and reasons for exercising are prolonged life, more energy, and a stronger body. Other participants in our study reported their appearance as a motivator for exercise. While participants in our study discussed their motivations for exercising, they lacked confidence in their abilities to exercise safely, feared injury, and desired supervision. Participants described how traditional gyms lack adequate supervision and support for older adults learning how to exercise. Access to exercise programs specific to older adults is critical to help motivate individuals to exercise and maintain PA levels [5,6].

Our study was conducted during the COVID-19 pandemic, which has been shown to limit walking, biking, PA, and mobility while increasing sedentary behavior across all age groups [27]. As a result, participants discussed the interruption of the pandemic to their lives and its impact in our interviews. While some evidence suggests that older adults became more active in the pandemic [28], this was not the case for our participants, who reported experiencing changes to their PA levels and daily routines and found it difficult to motivate themselves to be physically active independently. Similarly, in a 2021 qualitative study by Petersen et al [29], 12 healthy Canadian adults (6/12, 50% females; 20‐70 years), participants described disruptions to their daily routines and changes in PA as a result of the pandemic. Our study builds upon these findings by providing insight into how the closure of gyms and canceled exercise programs impacted participants in our study, as many were active members before the pandemic. Limited access to exercise programs may lead to difficulty for older adults in motivating themselves to be physically active.

Social connectivity is a key motivator for older adults to maintain adherence to exercise programs [6,30]. Our findings suggest that synchronous online exercise programs are motivational for older adults as they can connect in real time with instructors and other participants, which is unique to synchronous programming compared with asynchronous programs as older adults identify the lack of real-time objective feedback [31]. However, some older adults in our study had previously noted that socializing may not be important when exercising, and they prefer to focus on exercising alone [17]. The synchronous program made our participants feel comfortable exercising when they could not attend in-person programs, and it was advantageous compared with in-person exercise as they did not need to share their space. Synchronous programs may be a useful exercise approach as many older adults are shifting towards more digital connections with family and friends beyond the pandemic, as seen in a sedentary behavior reduction intervention with older adults [32]. However, this may not be a generalizable finding, particularly for older adults in Canada with lower socioeconomic status and lower educational attainment, as there may be barriers to technological access and literacy [33,34]. Careful consideration is needed to develop accessible future interventions and programs when leveraging technology to deliver synchronous PA and exercise programs more broadly. Future work should consider using in-person recruitment to ensure greater diversity in studies [35].

In addition to technology barriers, other commonly cited barriers to participation in PA and exercise included environmental factors, access to resources, assistance in managing change, and social influences for older community-dwelling adults [5,6]. Our findings reflect the environmental barriers reported in the literature, including transportation, access to exercise facilities, and poor weather conditions. Participants in our study expressed concerns about driving and parking at fitness centers, but these transportation concerns were alleviated through our online exercise program, which enabled participants to focus on exercising. Participants described how the “Internet” was a good way to “deal” with transportation concerns, as many described the extra mental load of in-person programming. Previous research has highlighted the importance of transportation assistance in promoting the uptake of programs and maintenance of PA for community-dwelling older adults who are socially disadvantaged or experience disability [36-38]. Reducing transportation concerns may improve the accessibility of exercise programs and the ability for older adults to participate from their desired training location [38].

Barriers to exercise accessibility due to lack of time (family and work) and lack of facilities are other commonly reported environmental barriers [5,6]. In focus groups with recently retired older adults, they described feeling like they are relied upon by children, grandchildren, parents, and friends, making it difficult to prioritize structured PA [39]. For some, online programs provide convenience in terms of not needing transportation, allowing older adults to spend more time on other priorities [40]. Our participants commented on the ease of getting up 5 minutes before class and our program’s convenience. The frequency of our program, 3 times per week, was also feasible for older adults [17]. Our synchronous online exercise program improved participants’ access to exercise specific for older adults and allowed the flexibility of attending class from home while still remaining socially connected. It additionally created a psychologically safe place for participants compared with traditional gyms, which are dominated by younger adults and lack supervised training specifically for older adults. Creating convenient, safe, and synchronous programs is important for ongoing adherence [39].

Poor weather conditions (eg, winter, rain, extreme heat) in Canada are another environmental barrier that contributes to increased sedentary time [41]. Our findings suggest that ice, snow, rain, and extreme heat discourage exercise, particularly since most older adults must travel to a facility to participate in an exercise program. Participants suggested that online delivery may effectively reduce this environmental barrier as participants do not need to be concerned with “cleaning their cars off” or “walking on ice” to attend a training facility to exercise. In older Canadians, there is a decrease in PA levels during precipitation [42] and winter months [8]. Synchronous delivery of exercise may be a feasible approach to deliver programs broadly,

XSL•FO
RenderX

particularly to populations who may not access in-person programs easily. Future work should explore hybrid exercise programming models, including in-person and synchronous delivery.

Strengths and Limitations

Our study suggests that synchronous exercise programs are motivational for older adults as they can connect in real-time with instructors while eliminating common environmental barriers to in-person exercise programming. Interpretive design allowed us to address a complex experiential question while producing a practical outcome [19]. A potential limitation of this study is the limited diversity among our sample. Our study population was predominately Caucasian, with high education attainment, women between 65 and 80, and motivated to exercise. Thus, the findings do not represent the perceptions and experiences of structurally marginalized populations, those experiencing lower socioeconomic status, men, older than 80 years, and those with low levels of education attainment or motivation to support behavioral change. We did not include nonparticipants in the program (n=1) in the interviews. Including nonparticipants may have allowed us to elucidate further barriers and health inequities for those unable to complete the program. Although the study lead provided technological support and assistance through Zoom or a phone call, including turning on cameras, muting and unmuting, and connecting speakers, all participants had good technology literacy and digital access. Future studies should consider including individuals in different populations with varying degrees of technological literacy. Additionally, conducting the synchronous online exercise program outside of the COVID-19 context would provide valuable insight into the long-term feasibility of this program. Trustworthiness may not have been achieved as the transcripts and codes were not checked and confirmed by the participants in this study [25]; however, triangulation across analysts strengthened this aspect of rigor.

Conclusion

The delivery of exercise programs using synchronous online classes may help older adults meet PA guidelines. Older adults report experiencing environmental barriers, including transportation, lack of access, and poor weather conditions when exercising. Our findings suggest that synchronous online exercise programs may serve as an approach to mitigate these environmental barriers and motivate older adults while keeping them socially connected. Considerations for designing exercise programs for older adults include delivery of the program by exercise and health professionals, synchronous programming to maintain social connectivity, and reducing environmental barriers such as transportation and weather concerns. Community programs may consider implementing synchronous online exercise programs as part of their recreational programming for older adults to increase engagement and reduce accessibility barriers. Future work should focus on leveraging synchronous exercise programs in community programs to engage older adults in PA and exercise and explore hybrid (synchronous and in-person) options.

Acknowledgments

We would like to acknowledge and thank the Physical Activity Centre of Excellence (PACE) registered kinesiologists and physiotherapists for their assistance in designing and delivering our online exercise program.

Conflicts of Interest

SMP reports grants or research contracts from the US National Dairy Council, Canadian Institutes for Health Research, Dairy Farmers of Canada, Roquette Freres, Ontario Centre of Innovation, Nestle Health Sciences, Myos, National Science and Engineering Research Council, and the US National Institutes of Health during the conduct of the study; personal fees from Nestle Health Sciences and nonfinancial support from Enhanced Recovery, outside the submitted work. SMP has patents licensed to Exerkine but reports no financial gains from patents or related work.

Multimedia Appendix 1
Semistructured interview questions.
[DOCX File, 18 KB - aging_v8i1e66473_app1.docx]

References

1. Bennie JA, De Cocker K, Teychenne MJ, Brown WJ, Biddle SJH. The epidemiology of aerobic physical activity and muscle-strengthening activity guideline adherence among 383,928 U.S. adults. Int J Behav Nutr Phys Act 2019 Apr 18;16(1):34. [doi: 10.1186/s12966-019-0797-2] [Medline: 30999896]
2. Colley RC, Garriguet D, Janssen I, Craig CL, Clarke J, Tremblay MS. Physical activity of Canadian adults: accelerometer results from the 2007 to 2009 Canadian Health Measures Survey. Health Rep 2011 Mar;22(1):7-14. [Medline: 21510585]
3. Ross R, Chaput JP, Giangregorio LM, et al. Canadian 24-Hour Movement Guidelines for adults aged 18-64 years and adults aged 65 years or older: an integration of physical activity, sedentary behaviour, and sleep. Appl Physiol Nutr Metab 2020 Oct;45(10 (Suppl. 2):S57-S102. [doi: 10.1139/apnm-2020-0467] [Medline: 33054332]
4. Bull FC, Al-Ansari SS, Biddle S, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med 2020 Dec;54(24):1451-1462. [doi: 10.1136/bjsports-2020-102955] [Medline: 33239350]

XSL•FO
RenderX

5. Spiteri K, Broom D, Bekhet AH, de Caro JX, Laventure B, Grafton K. Barriers and motivators of physical activity participation in middle-aged and older-adults - a systematic review. J Aging Phys Act 2019 Sep 1;27(4):929-944. [doi: 10.1123/japa.2018-0343] [Medline: 31141447]
6. Baert V, Gorus E, Mets T, Geerts C, Bautmans I. Motivators and barriers for physical activity in the oldest old: a systematic review. Ageing Res Rev 2011 Sep;10(4):464-474. [doi: 10.1016/j.arr.2011.04.001] [Medline: 21570493]
7. Chan CB, Ryan DA. Assessing the effects of weather conditions on physical activity participation using objective measures. Int J Environ Res Public Health 2009 Oct;6(10):2639-2654. [doi: 10.3390/ijerph6102639] [Medline: 20054460]
8. Merchant AT, Dehghan M, Akhtar-Danesh N. Seasonal variation in leisure-time physical activity among Canadians. Can J Public Health 2007;98(3):203-208. [doi: 10.1007/BF03403713] [Medline: 17626385]
9. Schroeder T, Dodds L, Georgiou A, Gewald H, Siette J. Older adults and new technology: mapping review of the factors associated with older adults' intention to adopt digital technologies. JMIR Aging 2023 May 16;6:e44564. [doi: 10.2196/44564] [Medline: 37191976]
10. Mace RA, Mattos MK, Vranceanu AM. Older adults can use technology: why healthcare professionals must overcome ageism in digital health. Transl Behav Med 2022 Dec 30;12(12):1102-1105. [doi: 10.1093/tbm/ibac070] [Medline: 36073770]
11. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556. [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
12. Raquel Costa-Brito A, Bovolini A, Rúa-Alonso M, et al. Home-based exercise interventions delivered by technology in older adults: a scoping review of technological tools usage. Int J Med Inform 2024 Jan;181:105287. [doi: 10.1016/j.ijmedinf.2023.105287] [Medline: 37972483]
13. Goethals L, Barth N, Guyot J, Hupin D, Celarier T, Bongue B. Impact of home quarantine on physical activity among older adults living at home during the COVID-19 pandemic: qualitative interview study. JMIR Aging 2020 May 7;3(1):e19007. [doi: 10.2196/19007] [Medline: 32356777]
14. Fuentes Diaz MF, Leadbetter B, Pitre V, Nowell S, Sénéchal M, Bouchard DR. Synchronous group-based online exercise programs for older adults living in the community: a scoping review. J Aging Phys Act 2024 Dec 1;32(6):703-717. [doi: 10.1123/japa.2023-0214] [Medline: 38823794]
15. Peng R, Chang J, Du Y, et al. Older adults' perceptions and experiences of engaging in web- and mobile-based physical activity interventions: a systematic review and qualitative meta-synthesis. Geriatr Nurs 2024;59(630-8):630-638. [doi: 10.1016/j.gerinurse.2024.08.025] [Medline: 39197354]
16. Dagenais M, Parker O, Galway S, Gammage K. Online exercise programming among older adults: a scoping review. J Aging Phys Act 2023 Apr 1;31(2):289-302. [doi: 10.1123/japa.2021-0417] [Medline: 36087933]
17. Coletta G, Noguchi KS, Beaudoin KD, et al. A live online exercise program for older adults improves depression and life-space mobility: a mixed-methods pilot randomized controlled trial. PLoS One 2024;19(11):e0312992. [doi: 10.1371/journal.pone.0312992] [Medline: 39527532]
18. Thorne S. Interpretive Description: Qualitative Research for Applied Practice: Taylor & Francis; 2016.
19. Thompson Burdine J, Thorne S, Sandhu G. Interpretive description: a flexible qualitative methodology for medical education research. Med Educ 2021 Mar;55(3):336-343. [doi: 10.1111/medu.14380] [Medline: 32967042]
20. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
21. Martínez-Mesa J, González-Chica DA, Duquia RP, Bonamigo RR, Bastos JL. Sampling: how to select participants in my research study? An Bras Dermatol 2016;91(3):326-330. [doi: 10.1590/abd1806-4841.20165254] [Medline: 27438200]
22. Adams WC. Conducting semi-structured interviews. In: Handbook of Practical Program Evaluation: John Wiley & Sons; 2015:492-505. [doi: 10.1002/9781119171386]
23. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
24. Creswell JW, Clark VLP. Designing and Conducting Mixed Methods Methods Research: SAGE Publications; 2017.
25. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. Int J Qual Methods 2017;16(1). [doi: 10.1177/1609406917733847]
26. Harrison EA, Lord LM, Asongwed E, et al. Perceptions, opinions, beliefs, and attitudes about physical activity and exercise in urban-community-residing older adults. J Prim Care Community Health 2020;11:2150132720924137. [doi: 10.1177/2150132720924137] [Medline: 32468912]
27. Park AH, Zhong S, Yang H, Jeong J, Lee C. Impact of COVID-19 on physical activity: a rapid review. J Glob Health 2022;12(5003):05003. [doi: 10.7189/jogh.12.05003] [Medline: 35493780]
28. Colley RC, Saunders TJ. The ongoing impact of the COVID-19 pandemic on physical activity and screen time among Canadian adults. Health Rep 2023 Oct 18;34(10):13-23. [doi: 10.25318/82-003-x202301000002-eng] [Medline: 37856572]
29. Petersen JA, Naish C, Ghoneim D, Cabaj JL, Doyle-Baker PK, McCormack GR. Impact of the COVID-19 pandemic on physical activity and sedentary behaviour: a qualitative study in a Canadian city. Int J Environ Res Public Health 2021 Apr 22;18(9):4441. [doi: 10.3390/ijerph18094441] [Medline: 33922094]

XSL•FO
RenderX

30. Killingback C, Tsofliou F, Clark C. Older people's adherence to community-based group exercise programmes: a multiple-case study. BMC Public Health 2017 Jan 25;17(1):115. [doi: 10.1186/s12889-017-4049-6] [Medline: 28122532]
31. Jansons P, Fyfe J, Via JD, Daly RM, Gvozdenko E, Scott D. Barriers and enablers for older adults participating in a home-based pragmatic exercise program delivered and monitored by Amazon Alexa: a qualitative study. BMC Geriatr 2022 Mar 25;22(1):248. [doi: 10.1186/s12877-022-02963-2] [Medline: 35337284]
32. Greenwood-Hickman MA, Dahlquist J, Cooper J, et al. "They're Going to Zoom It": a qualitative investigation of impacts and coping strategies during the COVID-19 pandemic among older adults. Front Public Health 2021;9(679976):679976. [doi: 10.3389/fpubh.2021.679976] [Medline: 34095079]
33. Arcury TA, Sandberg JC, Melius KP, et al. Older adult internet use and eHealth literacy. J Appl Gerontol 2020 Feb;39(2):141-150. [doi: 10.1177/0733464818807468] [Medline: 30353776]
34. Davidson J, Schimmele C. Evolving internet use among canadian seniors. Statistics Canada. 2019 Jul 7. URL: https://www150.statcan.gc.ca/n1/pub/11f0019m/11f0019m2019015-eng.htm [accessed 2025-03-27]
35. Carter CR, Maki J, Ackermann N, Waters EA. Inclusive recruitment strategies to maximize sociodemographic diversity among participants: a St. Louis case study. MDM Policy Pract 2023;8(1):23814683231183646. [doi: 10.1177/23814683231183646] [Medline: 37440792]
36. Nau T, Nolan G, Smith BJ. Enhancing engagement with socially disadvantaged older people in organized physical activity programs. Int Q Community Health Educ 2019 Jul;39(4):257-267. [doi: 10.1177/0272684X18821301] [Medline: 30638424]
37. Phoenix C, Griffin M, Smith B. Physical activity among older people with sight loss: a qualitative research study to inform policy and practice. Public Health 2015 Feb;129(2):124-130. [doi: 10.1016/j.puhe.2014.10.001] [Medline: 25687710]
38. Maula A, LaFond N, Orton E, et al. Use it or lose it: a qualitative study of the maintenance of physical activity in older adults. BMC Geriatr 2019 Dec 12;19(1):349. [doi: 10.1186/s12877-019-1366-x] [Medline: 31830900]
39. Van Dyck D, Mertens L, Cardon G, De Cocker K, De Bourdeaudhuij I. Opinions toward physical activity, sedentary behavior, and interventions to stimulate active living during early retirement: a qualitative study in recently retired adults. J Aging Phys Act 2017 Apr;25(2):277-286. [doi: 10.1123/japa.2015-0295] [Medline: 27759483]
40. Islam MS, Frazier MC, Harden SM, Lim S. Barriers and benefits of online group exercise programs for older adults. J Appl Gerontol 2024 Oct;43(10):1397-1407. [doi: 10.1177/07334648241240599] [Medline: 38536888]
41. Turrisi TB, Bittel KM, West AB, et al. Seasons, weather, and device-measured movement behaviors: a scoping review from 2006 to 2020. Int J Behav Nutr Phys Act 2021 Feb 4;18(1):24. [doi: 10.1186/s12966-021-01091-1] [Medline: 33541375]
42. Hoppmann CA, Lee JCM, Ziegelmann JP, Graf P, Khan KM, Ashe MC. Precipitation and physical activity in older adults: the moderating role of functional mobility and physical activity intentions. J Gerontol B Psychol Sci Soc Sci 2017 Sep 1;72(5):792-800. [doi: 10.1093/geronb/gbv107] [Medline: 26707498]

Abbreviations

CON: control group
PA: physical activity
RCT: randomized controlled trial

Edited by M O'Connell; submitted 14.09.24; peer-reviewed by B Leadbetter, IJ Liang, S Harden; revised version received 31.12.24; accepted 10.03.25; published 03.04.25.

Please cite as:
Coletta G, Noguchi KS, Beaudoin K, McQuarrie A, Tang A, Ganann R, Phillips SM, Griffin M
Older Adults' Perspectives on Participating in a Synchronous Online Exercise Program: Qualitative Study
JMIR Aging 2025;8:e66473
URL: https://aging.jmir.org/2025/1/e66473
doi:10.2196/66473

©Giulia Coletta, Kenneth S Noguchi, Kayla Beaudoin, Angelica McQuarrie, Ada Tang, Rebecca Ganann, Stuart M Phillips, Meridith Griffin. Originally published in JMIR Aging (https://aging.jmir.org), 3.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effect of Physical Exercise on Telomere Length: Umbrella Review and Meta-Analysis

Juan Luis Sánchez-González[1], PhD; Juan Luis Sánchez-Rodríguez[2], PhD; Rogelio González-Sarmiento[3], PhD; Víctor Navarro-López[4], PhD; Raúl Juárez-Vela[5], PhD; Jesús Pérez[6,7,8,9], PhD; Javier Martín-Vallejo[10], PhD

[1]Faculty of Nursing and Physiotherapy, Department of Physiotherapy, University of Salamanca, Salamanca, Spain

[10]Faculty of Medicine, Department of Statistics, University of Salamanca, Salamanca, Spain

[2]Faculty of Psychology, Department of Basic Psychology, Psychobiology and Methodology, University of Salamanca, Salamanca, Spain

[3]Faculty of Medicine, Department of Medicine, University of Salamanca, Salamanca, Spain

[4]Faculty of Health Sciences, Department of Physical Therapy, Occupational Therapy, Rehabilitation and Physical Medicine, Universidad Rey Juan Carlos, Madrid, Spain

[5]Faculty of Health Sciences, Department of Nursing, University of La Rioja, Logroño, Spain

[6]Faculty of Medicine, Department of Psychiatry, University of Salamanca, Avenida Donantes de Sangre s/n, Salamanca, Spain

[7]Department of Psychiatry, University of Cambridge, Cambridge, United Kingdom

[8]Cambridgeshire and Peterborough NHS Foundation Trust, Cambridge, United Kingdom

[9]Norwich Medical School, University of East Anglia, Norwich, United Kingdom

Corresponding Author:
Jesús Pérez, PhD
Faculty of Medicine, Department of Psychiatry, University of Salamanca, Avenida Donantes de Sangre s/n, Salamanca, Spain

Abstract

Background: Telomere length (TL) is a marker of cellular health and aging. Physical exercise has been associated with longer telomeres and, therefore, healthier aging. However, results supporting such effects vary across studies. Our aim was to synthesize existing evidence on the effect of different modalities and durations of physical exercise on TL.

Objective: The aim of this study was to explore the needs and expectations of individuals with physical disabilities and their interventionists for the use of a virtual reality physical activity platform in a community organization.

Methods: We performed an umbrella review and meta-analysis. Data sources included PubMed, Embase, Web of Science, Cochrane Library, and Scopus. We selected systematic reviews and meta-analyses of randomized and nonrandomized controlled clinical trials evaluating the effect of physical exercise on TL.

Results: Our literature search retrieved 12 eligible systematic reviews, 5 of which included meta-analyses. We identified 22 distinct primary studies to estimate the overall effect size of physical exercise on TL. The overall effect size was 0.28 (95% CI 0.118-0.439), with a heterogeneity test value Q of 43.08 (P=.003) and I^2 coefficient of 51%. The number of weeks of intervention explained part of this heterogeneity (Q_B=8.25; P=.004), with higher effect sizes found in studies with an intervention of less than 30 weeks. Exercise modality explained additional heterogeneity within this subgroup (Q_B=10.28, P=.02). The effect sizes were small for aerobic exercise and endurance training, and moderate for high-intensity interval training.

Conclusions: Our umbrella review and meta-analysis detected a small-moderate positive effect of physical exercise on TL, which seems to be influenced by the duration and type of physical exercise. High quality studies looking into the impact of standardized, evidence-based physical exercise programs on TL are still warranted.

Trial Registration: PROSPERO CRD42024500736; https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=500736

(JMIR Aging 2025;8:e64539) doi:10.2196/64539

KEYWORDS

aging; chromosome; exercise; meta-analysis; telomere; telomerase; genes; genome; DNA

Introduction

Life expectancy has increased worldwide over the last century. The United Nations World Population Prospects [1] estimates that the world's population aged 65 years or older will rise by 16% in 2050, doubling the number of children younger than 5 years. This prospect will lead to a significant increase in age-related illnesses, such as cancer, dementia, or cardiovascular diseases [2], and is prompting research into the primary

XSL•FO
RenderX

mechanisms of aging and, subsequently, strategies to promote a healthy late-life.

López-Otín et al [3] have recently postulated 12 basic biological mechanisms of aging: epigenetic alterations, loss of proteostasis, disabled macroautophagy, deregulated nutrient-sensing, mitochondrial dysfunction, cellular senescence, stem cell exhaustion, altered intercellular communication, chronic inflammation, dysbiosis, genomic instability, and telomere attrition. Indeed, one of the main theories that has attempted to explain aging relates to telomere length (TL) and the role of telomerase, an enzyme responsible for maintaining and elongating telomeres. Telomeres are essential nucleoprotein structures located at the termini of eukaryotic chromosomes that play pivotal roles in safeguarding genomic integrity and regulating processes such as tumor suppression and aging [4]. Typically composed of a repetitive guanine-rich sequence extending from the chromosome end in the 5' to 3' orientation, paired with a complementary cytidine-rich strand [5], telomeres exhibit variations in sequence among species, yet share a common repetitive pattern across organisms. Measuring around 15 kilobases at birth in human somatic cells, telomeres undergo gradual attrition, with approximately 25 to 200 bases lost from their ends during each cell division in the absence of telomerase [6]. Upon reaching a critical length, telomeres signal cell cycle arrest, leading to cellular senescence and eventual demise [7]. Telomeres serve to shield chromosome ends from degradation and fusion, thereby upholding genomic stability [8,9].

There are different demographic factors that influence TL, such as genetics that may explain the high heritability of TL [10,11]; sex, with longer telomeres found in adult females [12]; or ethnicity, with the White community usually having longer telomeres than Black or Hispanic communities [13]. Also, stress levels have been associated with shorter TL, which may be due to oxidative stress and reduced telomerase enzyme activity [14,15]. Obesity [16], alcohol consumption [17], and smoking [16,18] are also factors that negatively influence TL.

Notably, longer telomeres have been found in people who exercise regularly [19] or with higher levels of daily physical activity [20,21]. Indeed, one of the most studied interventions regarding TL is physical exercise. Recent systematic reviews and meta-analyses [22-26] have investigated the effect of physical exercise on the TL of clinical and nonclinical samples, showing some positive, but still inconclusive, results. Aerobic exercise, such as running or swimming, has been consistently associated with longer telomeres. Denham et al [27] conducted a study that found individuals who regularly engage in aerobic exercise show greater telomerase activity, the enzyme that helps maintain TL, compared with those who lead a sedentary lifestyle. In addition, Ludlow et al [28] demonstrated that resistance training, which includes exercises like weight lifting, may also have beneficial effects on TL, possibly through mechanisms involving the reduction of oxidative stress and inflammation. In addition, more recent research has indicated that high-intensity training, such as high-intensity interval training (HIIT), may be more effective in preserving TL compared with low- to moderate-intensity exercises [29]. This type of exercise can induce more robust adaptive responses at the cellular level, including increased expression of genes related to longevity and telomere protection [3].

The aim of this umbrella review was to synthesize existing evidence on the effect of different modalities and durations of physical exercise on TL to inform the implementation of evidence-based physical exercise programs or recommendations to add healthier years into longer lives.

Methods

Overview

This umbrella review was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [30], and it is registered in the International Prospective Register of Systematic Reviews (PROSPERO; registration number CRD42024500736).

Inclusion Criteria

The inclusion criteria and description of studies in this review followed the PICOS (Population, Intervention, Control, Outcomes, and Study design) framework for reviews [31].

Study Design

We selected systematic reviews, with or without meta-analyses, of randomized and nonrandomized controlled clinical trials, excluding nonexperimental designs, such as observational studies. No restrictions were applied on the basis of any particular language, following current international recommendations [32].

Population

Study participants were people with or without a clinical condition. The systematic reviews had to explicitly state that they included humans in their analyses. Therefore, studies involving experimental animals, such as rodents, were excluded.

Intervention and Control

We selected all systematic reviews and meta-analyses comparing the effects of physical exercise versus other intervention, or none, on TL. In addition, we further divided the results according to different types of physical exercise. If any reviews included primary studies combining different types of exercise, they were classified as "combined."

Outcome Measure

The outcome measure was TL-related calculations.

Search Strategy

We performed a literature search for scientific articles published until January 19, 2024, using the following databases: PubMed (Medline), PEDro, Embase, Web of Science, and Cochrane Library. Multimedia Appendix 1 provides the search strategies, tailored for each database. A total of 2 independent reviewers (JLS-G and JLS-R) conducted the search using the same methodology. Any discrepancies during this process were resolved through consensus including a third reviewer (VN-L). In addition, we manually examined the reference sections of the original studies, and, if needed, contacted the authors for additional information.

XSL•FO
RenderX

Selection Criteria and Data Extraction

A total of 2 independent reviewers (JLS-G and JLS-R) initiated a screening process to evaluate relevance of the systematic reviews and meta-analyses. The initial screening was based on the title, abstract, and keywords for each review. When there was no consensus, or if the abstracts provided insufficient information, the full text was examined. During the second screening phase, the full text was evaluated to confirm inclusion criteria. Data presented in the results section were extracted using a protocol to ensure retrieval of the most relevant information for each study. The sample sizes, type of exercise, duration, means, and SDs of the telomere size data were extracted for each of the studies included in this review. During the analysis, we checked whether the results were repeated in different reviews or meta-analyses; if the effect sizes coincided, only one of the effect sizes was chosen; if the effect sizes did not match, the primary study was selected.

Methodological Quality Assessment

A total of 2 independent reviewers (JLSG and VNL) evaluated the methodological quality of the systematic reviews and meta-analyses, using the modified quality assessment scale for systematic reviews (AMSTAR-2 [A Measurement Tool to Assess Systematic Reviews]). AMSTAR-2 [33] is a questionnaire comprised of 16 domains, with simple categorical options: "yes," when the result is positive; "no," when the standard is not met or there is not enough information to answer it; and "partial yes," when there was partial adherence to the standard. In addition, we calculated the kappa coefficient (κ) and percentage (%) agreement scores to assess reliability before reaching consensus. Interrater reliability was estimated using κ, where κ>0.7 indicates a high level of agreement between reviewers, κ of 0.5‐0.7 indicates a moderate level of agreement, and κ<0.5 suggests a low level of agreement [34].

Risk of Bias Assessment

We evaluated the risk of bias using the Risk Of Bias in Systematic reviews (ROBIS) assessment, which includes three domains: (1) relevance of assessment (optional); (2) identification of concerns with the review process through 4 domains related to study eligibility criteria, identification and selection of studies, data collection and study appraisal, and synthesis and findings; and (3) judgment on the risk of bias. The ROBIS tool includes signaling questions to assess specific domains and to guide the judgment of the systematic review's risk of bias, with responses categorized as "yes," "probably yes," "probably no," "no," or "no information." The risk of bias is then categorized as "low," "high," or "unclear" [35].

A total of 2 independent reviewers (JLS-G and JLS-R) assessed the risk of bias in the selected studies. In addition, we computed the kappa coefficient (κ) and percentage (%) agreement scores to evaluate reliability before reaching consensus.

Grading of Evidence

The Physical Activity Guidelines Advisory Committee Grading Criteria (PAGAC) were used to evaluate grading of evidence. The criteria used for assessing the quality of evidence included (1) applicability of the study sample, exposures, and outcomes to the research question; (2) generalizability to the population of interest; (3) risk of bias/study limitations; (4) quantity and consistency of findings across studies; and (5) magnitude and precision of the effect. Based on this information, final evidence grades and conclusion statements for each research question were formulated [36].

Overlap of Primary Studies

The overlap analysis of primary studies among the systematic reviews was performed with the Graphical Representation of Overlap for OVErviews (GROOVE) tool [37]. Using a matrix of evidence, GROOVE establishes the number of primary studies and systematic reviews, the absolute number of overlapped and nonoverlapped primary studies, and an overall corrected covered area (CCA) assessment. GROOVE also offers a detailed CCA assessment for each possible pair of systematic reviews (or "nodes"), with a graphical and easy-to-read representation of these results. In addition, it provides an optional function that incorporates the structural missingness within the matrix.

Data Synthesis and Analysis

Given the high variability of designs, patients and endpoints across studies, results were integrated following a random effects model. The Hedges unbiased standardized mean difference was used to determine effect sizes. Heterogeneity was measured using Q-tests, I^2 coefficients, and prediction intervals [38]. Metaregression and metapartition [39] were used to explain heterogeneity. These techniques involve partitioning sum of squares of the homogeneity test into 2 components, Q_B and Q_W, to account for quantitative or qualitative variables. If the Q_E value was large and statistically significant compared with the value of Q_W, it was considered that this variable explained part of the heterogeneity found in the integration of results across studies. If the variable was qualitative, subgroups of studies, which differed in the effect size and also showed less heterogeneity, were created. The procedure was repeated for subgroups that still presented large heterogeneity for other variables. Publication bias was assessed using the contour-enhanced funnel plot, Egger test, Doi plot, and LFK index procedures. 95% CIs were calculated for effect sizes. The significance level was set at 5%. The analyses were carried out with the *meta* v7.0, *metafor* v4.6 and *metasens* 1.5‐2 libraries of the R statistics software (version 4.4.0; R Foundation for Statistical Computing).

Ethical Considerations

As this is an umbrella review, it is not necessary to have the approval of the ethics committee.

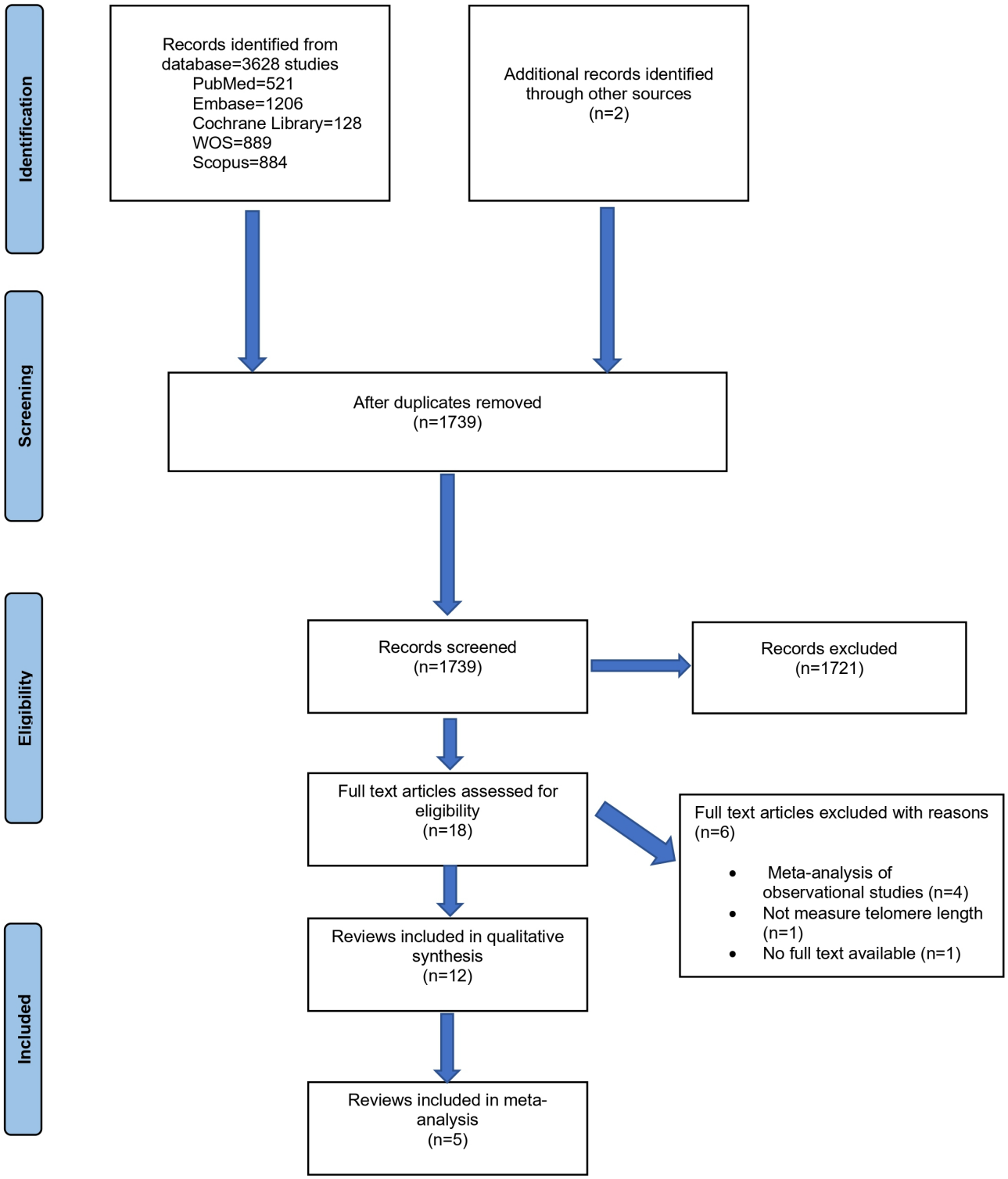
Results

Study Selection

The initial literature search revealed 3628 records. In addition, 2 more were retrieved manually from the references. A total of 12 systematic reviews were eligible for qualitative synthesis, and 5 of them allowed a meta-analysis. Figure 1 shows the study screening strategy.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart.



Characteristics of the Systematic Reviews

Table 1 lists the characteristics of the systematic reviews and meta-analyses (type of study, design, sample, intervention, lab technique to measure TL, risk of bias, evaluation of quality, and conclusions).

XSL•FO
RenderX

Table . Characteristics of the systematic reviews.

Study	Type of study	Study design, n		RCCTs[a] or CCTs[b] with physical exercise intervention					Risk of bias	Evaluation of quality	Conclusions
		Interventional	Observational	Sample		Intervention		Lab technique for TL[c]			
				n	Condition (number of studies, n)	Type (n)	Duration: weeks (n)				
Song et al [26]	Systematic review and meta-analysis	7 (RCCTs)	0	939	Healthy (3), breast cancer (2), polycystic ovary syndrome (1), and obese (1)	Aerobic exercise (4) and endurance training (3)	48 (2), 24 (3), and 16 (2)	qPCR[d]	Yes	No	The type and duration of exercise for positive improvement in TL is aerobic exercise for more than 6 months.
Buttet et al [22]	Systematic review and meta-analysis	13 (RCCTs) and 3 (CCTs)	5	908	Healthy (5), obese (1), and myocardial infarction (1)	Endurance training (7), strength (1), and HIIT[e] (1)	48 (2), 32 (1), and 24 (4)	qPCR	No	No	A lifestyle intervention with physical activity + diet can increase TL, independently of population characteristics or baseline TL.
Sánchez-González et al [19]	Systematic review and meta-analysis	8 (RCCTs) and 1 (CT)	0	1320	Healthy (9)	Endurance training (4), aerobic exercise (4), HIIT (3), and aerobic + endurance (1)	8 (1), 16 (1), 24 (4), and 52 (3)	qPCR	Yes	Yes	The findings suggest that HIIT seems to have a positive effect on TL compared with other types of exercise such as endurance training or aerobic exercise in healthy population.

XSL•FO
RenderX

Study	Type of study	Study design, n		RCCTs[a] or CCTs[b] with physical exercise intervention					Risk of bias	Evaluation of quality	Conclusions
		Interventional	Observational	Sample		Intervention		Lab technique for TL[c]			
				n	Condition (number of studies, n)	Type (n)	Duration: weeks (n)				
Valente et al [25]	Systematic review and meta-analysis	6 (RCCTs)	16 (case-control studies), 2 (prospective cohort), 3 (cross-sectional), and 3 (retrospective cohort)	612	Healthy (5) and obese (1)	Endurance training (5), strength (1), aerobic exercise (1), and HIIT (1)	56 (1), 48 (1), and 24 (4)	qPCR	Yes	No	There is very low certainty that physically active individuals have longer telomeres with a moderate effect, but this effect is probably overestimated.
Denham et al [23]	Systematic review and meta-analysis	12 (NR[f])	0	487	Healthy (6) and chronic fatigue (1)	HIIT (1), aerobic exercise (3), Pilates training (1), and endurance training (2)	52 (1), 24 (2), 22 (1), 8 (2), and 1 (1)	qPCR	Yes	No	Exercise training as an inexpensive lifestyle factor that increases telomerase reverse transcriptase expression and telomerase activity. Regular exercise training could attenuate telomere attrition through a telomerase-dependent mechanism and ultimately extend healthspan longevity.

XSL•FO
RenderX

Study	Type of study	Study design, n		RCCTs[a] or CCTs[b] with physical exercise intervention					Risk of bias	Evaluation of quality	Conclusions
		Interventional	Observational	Sample		Intervention		Lab technique for TL[c]			
				n	Condition (number of studies, n)	Type (n)	Duration: weeks (n)				
Barragán et al [40]	Systematic review	12 (NR)	53 (cross-sectional studies), 13 (case-control), and 9 (longitudinal)	1.056	Healthy (12)	Endurance training (5), combined training (1), aerobic exercise (6), NR (1), and HIIT (1)	56 (1), 52 (2), 24 (3), 8 (2), 1 (2), and NR (2)	qPCR	No	Yes	Although fewer sedentary activities, optimal sleep habits, and non- or ex-smoker status have been associated with less telomere shortening, several methodological issues were detected, including the need for more targeted interventions and standardized protocols to better understand how physical activity and sleep can impact TL and aging.
Schellnegger et al [41]	Systematic review	8 (RCCTs) and 7 (CCTs)	27 (observational studies)	1.700	Healthy (13) and obese (2)	Aerobic exercise (8), endurance training (8), and HIIT (1)	1 (6), 6 (1), 8 (1), 12 (1), 24 (5), and 52 (2)	qPCR	No	No	Physical activity with regular aerobic training of moderate to vigorous intensity appears to help preserve TL.

XSL•FO
RenderX

Study	Type of study	Study design, n		RCCTs[a] or CCTs[b] with physical exercise intervention					Risk of bias	Evaluation of quality	Conclusions
		Interventional	Observational	Sample		Intervention		Lab technique for TL[c]			
				n	Condition (number of studies, n)	Type (n)	Duration: weeks (n)				
Prathap et al [42]	Systematic review	2 (RCCTs): 1 in rodents	3 (literature review)	151	Breast cancer (1)	Aerobic exercise (1)	24 (1)	qPCR	No	No	Based on the evidence collected it can be suggested that chronic moderate intensity aerobic exercise in a lifelong practice shows beneficial effects in a dose-response manner in cancer prevention by modulating telomeres through epigenetic mechanism.
Quiao et al [43]	Systematic review	16 (RCCTs) and 14 (CCTs)	—[g]	562	Myocardial infarction (1), healthy adults (4), obese (1), and polycystic ovary syndrome (1)	Combined training (1), endurance training (3), and aerobic exercise (3)	8 (1), 12 (1), 16 (2), 20 (1), 24 (1), and 48 (1)	qPCR	No	Yes	Weight-loss and comprehensive lifestyle intervention strategies show encouraging impacts in delaying telomere shortening. More rigorous studies targeting populations at different age stages through life span are needed.

XSL•FO
RenderX

Study	Type of study	Study design, n		RCCTs[a] or CCTs[b] with physical exercise intervention					Risk of bias	Evaluation of quality	Conclusions
		Interventional	Observational	Sample		Intervention		Lab technique for TL[c]			
				n	Condition (number of studies, n)	Type (n)	Duration: weeks (n)				
Min et al [44]	Systematic review	2 (RCCTs)	3 (cross-sectional study)	247	2 (breast cancer survivor)	Aerobic exercise (2)	24 (1) and 48 (1)	qPCR	No	No	Three of the five studies reported that physical activity has a significant relationship in delaying TL shortening, but others observed no association between physical activity and TL in breast cancer survivors.

XSL•FO
RenderX

Study	Type of study	Study design, n		RCCTs[a] or CCTs[b] with physical exercise intervention					Risk of bias	Evaluation of quality	Conclusions
		Interventional	Observational	Sample		Intervention		Lab technique for TL[c]			
				n	Condition (number of studies, n)	Type (n)	Duration: weeks (n)				
Marques et al [45]	Systematic review	4 (RCCTs)	16 (cross-sectional)	647	Obese (1) and healthy (3)	Aerobic exercise (4), HIIT (1), and endurance training (1)	24 (3) and 48 (1)	qPCR	No	Yes	Better cardiorespiratory fitness or a large cardiorespiratory training load are associated with an increase in TL. Although, TL was related to regular moderate-to-vigorous aerobic exercise and cardiorespiratory fitness in older healthy humans, it was not related to cardiorespiratory fitness among young subjects.

XSL•FO
RenderX

Study	Type of study	Study design, n		RCCTs[a] or CCTs[b] with physical exercise intervention					Risk of bias	Evaluation of quality	Conclusions
		Interventional	Observational	Sample		Intervention		Lab technique for TL[c]			
				n	Condition (number of studies, n)	Type (n)	Duration: weeks (n)				
Himbert et al [46]	Systematic review	10 (RCCTs) and 11 (CCTs)	—	439	Obese (1)	Aerobic exercise (1)	48 (1)	qPCR	No	No	The inconsistent effects of weight loss on telomere length or DNA repair suggest the need for a re-assessment of intervention designs and assay methodology to definitively address this topic.

[a]RCCT: randomized controlled clinical trial.

[b]CCT: (nonrandomized) controlled clinical trials.

[c]TL: telomere length.

[d]qPCR: polymerase chain reaction quantitative.

[e]HIIT: high-intensity interval training.

[f]NR: not reported.

[g]Not applicable.

AMSTAR-2 Appraisal

The overall confidence ranged from critically low to moderate quality scores. A total of 2 studies (16.6%) obtained an overall confidence of moderate, and 10 studies (83.4%) had critically low quality (Multimedia Appendix 2). The items with the highest scores were “did the review authors use a comprehensive literature search strategy?”, “did the review authors describe the included studies in adequate detail?”, “did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?”, and “did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?” The lowest scoring items were “did the review authors account for risk of bias in individual studies when interpreting/discussing the results of the review?”, “did the review authors report on the sources of funding for the studies included in the review?”, and “if meta-analysis was performed, did the review authors assess the potential impact of risk of bias in individual studies on the results of the meta-analysis or other evidence synthesis?” The interrater reliability of the methodological quality assessment was high (κ≥0.8).

ROBIS Assessment

Multimedia Appendix 3 summarizes the risk of bias using ROBIS. 40% (5/12) of the studies had a low risk of bias, 20% (2/12) had an unclear risk of bias, and 40% (5/12) had a low risk of bias. The domains related to “data collection” and “synthesis and findings” had the highest risk of bias. On the other hand, the domain related to “eligibility criteria” had the lowest risk of bias. The interrater reliability of the risk of bias was high (k>0.8).

PAGAC Grades of Evidence

The grades of evidence were classified as not assignable, limited, moderate, or strong according to the PAGAC. The level of evidence was limited to moderate among the studies. Most reviews evaluated a heterogeneous population of individuals, which limited the applicability and generalizability of the results. The domain “quantity and consistency” was affected because

XSL•FO
RenderX

some studies had inconsistency in the direction of or the effect size itself (Multimedia Appendix 4).

GROOVE Analysis

A total of 39 primary studies were identified across all systematic reviews and meta-analyses, 24 of which were distinct studies. The overall overlap in the matrix of evidence was moderate (CCA=32.5%) and it remained moderate (CCA=15.63%) after adjusting for the chronological structural missingness. Multimedia Appendix 5 shows a graphical representation of the GROOVE results.

Telomere Length

The number of studies in the systemic reviews that were finally integrated in the quantitative (meta-analytic analysis) was 22. The estimated effect size was 0.28 (95% CI 0.118-0.439), with a heterogeneity test value (Q) of 43.08 (P=.003). The I^2 coefficient was 51% and the prediction interval was −0.26 to 0.81. The number of weeks of intervention explained part of this heterogeneity (Q_B=7.54; P=.006). This factor explained 20% of the overall heterogeneity. Figure 2 shows that studies where the number of weeks of intervention was greater than 30 had smaller effect sizes. In fact, if the intervention time was categorized into <30 and ≥30 weeks, Q_B was 11.64 (P<.001).

Figure 2. Scatterplot of primary studies including effect size and number of weeks of intervention.

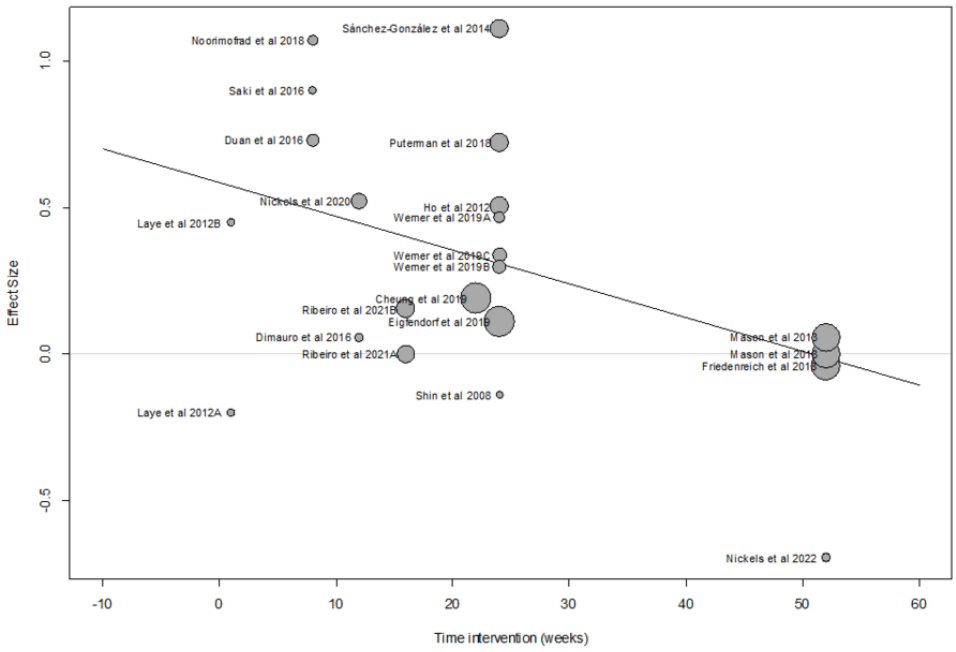
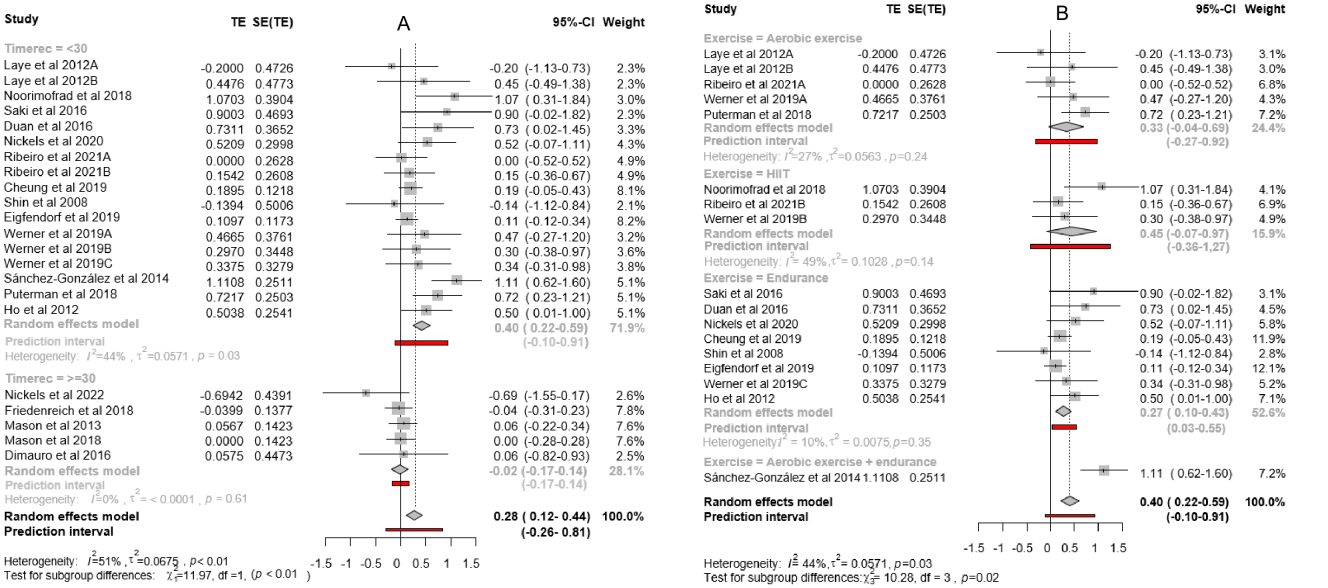


Figure 3A shows that the effect size was 0.39 for studies with an intervention length of less than 30 weeks, and −0.01 for those with interventions longer than 30 weeks. There was no significant difference in the average number of weekly sessions between these two groups (3.3 vs 3 sessions, respectively). Interestingly, the first group presented a statistically significant heterogeneity (I^2=41%), whilst the second, despite having a large range of outcomes, did not. Since heterogeneity was high in the group of studies with shorter interventions (<30 wk), it was analyzed with respect to other variables. The factor that better explained this heterogeneity was the type of physical exercise defined by the 4 categories (subgroups) as shown in Figure 3 (Q_B=11.15, P=.01). One subgroup was comprised of just 1 study that included a combination of aerobic exercise with endurance training and achieved a large effect size. The other 3 subgroups did not have a statistically significant heterogeneity between them, although the subgroup of studies defined by an intervention based on HIIT, which alternates short periods of intense anaerobic exercise with brief recovery periods, had a marked I^2 value (49%). The HIIT subgroup only included 3 studies, one of which had a very high effect size compared to the other two. When we compared the effect sizes of the 3 types of noncombined physical exercises, the value was small for aerobic exercise and endurance training, and moderate for HIIT.

XSL•FO
RenderX

Figure 3. (A) Forest plot of subgroups of primary studies defined by length of intervention. (B) Forest plot of subgroups of primary studies defined by type of exercise. TE: treatment effect.



The contour-enhanced funnel plot (Multimedia Appendix 6) did not identify publication bias. Many of the effect sizes were not significant, and at the top of the funnel, there was some asymmetry, which could be due to the heterogeneity of results. The nonpresence of publication bias was confirmed with the Egger method (*P*=.10). The Doi plot and LFK index also found a weak asymmetry (Multimedia Appendix 7).

Discussion

Principal Findings

This umbrella review studied the effects of physical exercise on TL. The estimated overall effect size indicated a small-moderate positive impact of exercise on TL. However, we found some discrepancies in the results across studies that may be explained by several methodological and contextual reasons, including different methodologies for measuring TL, such as quantitative polymerase chain reaction and terminal restriction fragment assays, that can contribute to inconsistencies in TL calculations. In this regard, Smith et al [47] pointed out that intraindividual TL can vary significantly depending on the lab technique used to measure it. Differences in study populations could also be a factor responsible for disparities in results. Demographic factors, such as age, sex, ethnicity, and health status, can influence the relationship between exercise and TL. For instance, LaRocca et al [48] and Puterman et al [21] suggest that the benefits of exercise on TL may be more pronounced in older adults and in those with higher stress levels. In addition, the duration and intensity of the intervention could be another confounding factor. Our findings suggest that interventions longer than 30 weeks tend to show a lesser positive impact on TL. This could be due to a ceiling effect where additional benefits of exercise do not translate into further increase of TL. This is consistent with Werner et al [29], who observed that the beneficial effects of exercise on TL are more evident in the early phases of physical exercise programs.

The high heterogeneity in study designs, including interventions, participant adherence, and control measures, could have also contributed to different outcomes. In fact, Denham et al [23] and Puterman et al [21] have already highlighted the importance of these methodological issues when interpreting TL results. Studies such as Cherkas et al [20] had found that vigorous aerobic physical activity was associated with longer telomeres in adults, which is in line with our finding of a positive effect of aerobic exercise, especially when combined with endurance training, on TL. However, a systematic review by Du et al [49] did not identify a significant association between aerobic exercise and TL in younger populations, suggesting that age may be an important moderating factor. Denham et al [23] found mixed results related to endurance physical training. Our review also found significant variation in the effects of noncombined endurance training, which suggests that factors such as the intensity and duration of this type of physical exercise may influence the magnitude of the effect on TL. Finally, HIIT had previously shown a significant positive effect on TL [29]. Our review results are consistent with this; nonetheless, the low number of HIIT studies and variety of designs limit generalizability.

Biological Understanding of the Effect of Exercise on TL

Reduction of Oxidative Stress and Inflammation

Oxidative stress and chronic inflammation are 2 major factors contributing to telomere erosion [50,51]. Studies have shown that regular exercise can mitigate these factors [52-54]. Acute and chronic physical activity have different effects on oxidative stress. While acute exercise triggers the production of reactive oxygen and nitrogen species, leading to oxidative stress, regular exercise training enhances the body's endogenous antioxidant system, offering protection against the harmful effects of oxidative damage [55]. This antioxidant effect protects cells from oxidative damage, which in turn preserves the integrity of telomeres. In addition, exercise reduces the levels of inflammatory markers such as C-reactive protein and tumor necrosis factor-alpha, which can protect telomeres from inflammation-induced damage [56].

XSL•FO
RenderX

Increase in Telomerase Activity

Telomerase is a crucial enzyme for maintaining TL, as it adds repetitive sequences to the end of chromosomes. Research suggests that exercise can increase telomerase activity. Ornish et al [57] demonstrated that an intensive lifestyle change program, which included regular exercise, significantly increased telomerase activity in men with prostate cancer. This finding suggests that exercise can directly influence telomere biology by activating telomerase, which helps maintain TL and protects cells from premature aging. In addition, a systematic review with meta-analysis published by Denham et al [23] concludes that regular exercise may attenuate telomere attrition through a telomerase-dependent mechanism and ultimately prolong lifespan and longevity. Furthermore, Ludlow et al [28] found that moderate physical activity levels are associated with increased telomerase activity in leukocytes, indicating that regular exercise not only improves overall health, but also acts at a molecular level to protect cells from premature aging. These studies highlight the importance of exercise continuity and intensity in maximizing the benefits related to telomerase activity and cellular longevity.

Improvement of Cardiovascular Capacity and Metabolic Health

Improved cardiovascular capacity and metabolic health are also associated with longer telomeres. Puterman et al [21] found that better aerobic capacity and higher insulin sensitivity are associated with longer telomeres. These cardiovascular and metabolic benefits reduce oxidative stress and inflammation, creating a healthier cellular environment that protects TL. Regular exercise enhances cardiovascular system efficiency, increases blood circulation, and facilitates the delivery of oxygen and nutrients to cells, which can contribute to the preservation of telomeres.

Strengths and Limitations

We present a comprehensive analysis of systematic reviews and meta-analyses of the effect of physical exercise on TL, providing a broad and robust synthesis of the available evidence. The methodology used to assess the quality of studies and the application of standardized tools for evaluating the risk of bias and heterogeneity strengthens the validity of our results. However, our work has some limitations. The methodological quality of the reviews included in this analysis varies, with many of them classified as critically low. Also, the heterogeneity in study designs, samples, and interventions makes difficult to generalize the results, even after the use of a random effects model to integrate outcomes. Furthermore, some meta-analyses reported results only using postintervention measures, whilst others used postintervention differences adjusted for preintervention measures; this may lead to variations in effect sizes.

Clinical Implications

The implementation of exercise programs into clinical practice could be an effective strategy to limit telomere shortening, promoting healthy aging and reducing the incidence of age-related diseases. HIIT and the combination of aerobic exercise with endurance training may be promising interventions in this regard but still require to be further tested in high-quality studies. In addition, physical exercise should be tailored to each individual's physical capability and health condition to maximize cellular health benefits. Implementing physical exercise programs will also require ongoing health education and monitoring to ensure adherence to and adjustment of interventions based on observed outcomes. Future studies should investigate sex and ethnic TL differences in response to various types of physical exercise. This could help personalize recommendations and optimize benefits on cellular health and aging.

Conclusions

Our umbrella review and meta-analysis identified a small-moderate positive effect of physical exercise on TL, which seems to be influenced by the duration and type of physical exercise. In health care systems, the implementation of evidence-based physical exercise, training programs, and/or recommendations tailored to each individual might be a valuable preventive strategy for a healthy aging. More studies, with larger sample sizes, testing the impact of standardized, evidence-based physical exercise interventions on TL are warranted.

Disclaimer

The lead author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained. Dissemination to participants and related patients and public communities: After publication, the findings of this review will be disseminated to appropriate audiences, such as academia, clinicians, policy makers, and the general public, through various channels including blogs, press releases, and social media.

Data Availability

The data sets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

JLS-G, JLS-R, RG-S, JP, and JM-V contributed to the design of this review. JLS-G, VN-L, JP, and JM-V conducted study selection, data extraction and analysis, and wrote the first draft of this manuscript. RJ-V provided a detailed review of the statistical

XSL•FO
RenderX

analysis. All authors reviewed and approved the final version of the manuscript. The corresponding author, JP, confirms that all listed authors meet authorship criteria and that no others meeting these criteria have been omitted.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Database formulas during literature search.
[DOCX File, 13 KB - aging_v8i1e64539_app1.docx]

Multimedia Appendix 2
Risk of bias.
[PNG File, 263 KB - aging_v8i1e64539_app2.png]

Multimedia Appendix 3
Graphical representation of the overlap of primary studies between reviews.
[PNG File, 100 KB - aging_v8i1e64539_app3.png]

Multimedia Appendix 4
Contour-enhanced funnel plot of the primary studies included in the meta-analysis.
[PNG File, 22 KB - aging_v8i1e64539_app4.png]

Multimedia Appendix 5
Doi plot and LFK index of the primary studies included in the meta-analysis.
[PNG File, 7 KB - aging_v8i1e64539_app5.png]

Multimedia Appendix 6
Quality assessment scores (A Measurement Tool to Assess Systematic Reviews [AMSTAR-2]).
[DOCX File, 18 KB - aging_v8i1e64539_app6.docx]

Multimedia Appendix 7
Summary of findings and quality of evidence (Physical Activity Guidelines Advisory Committee Grading Criteria [PAGAC]).
[DOCX File, 16 KB - aging_v8i1e64539_app7.docx]

Checklist 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[PDF File, 83 KB - aging_v8i1e64539_app8.pdf]

References

1. World Population Prospects 2022: Summary of results. : United Nations; 2022.
2. Melzer D, Pilling LC, Ferrucci L. The genetics of human ageing. Nat Rev Genet 2020 Feb;21(2):88-101. [doi: 10.1038/s41576-019-0183-6] [Medline: 31690828]
3. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. Hallmarks of aging: an expanding universe. Cell 2023 Jan 19;186(2):243-278. [doi: 10.1016/j.cell.2022.11.001] [Medline: 36599349]
4. Zakian VA. Telomeres: beginning to understand the end. Science 1995 Dec 8;270(5242):1601-1607. [doi: 10.1126/science.270.5242.1601] [Medline: 7502069]
5. Smith EM, Pendlebury DF, Nandakumar J. Structural biology of telomeres and telomerase. Cell Mol Life Sci 2020 Jan;77(1):61-79. [doi: 10.1007/s00018-019-03369-x] [Medline: 31728577]
6. Opresko PL, Shay JW. Telomere-associated aging disorders. Ageing Res Rev 2017 Jan;33:52-66. [doi: 10.1016/j.arr.2016.05.009] [Medline: 27215853]
7. Cascales-Angosto M, Álvarez-Gómez JA. Telómeros y telomerasa, sus implicaciones en el envejecimiento y el cáncer [Article in Spanish]. An Real Acad Dr Esp 2010;14:49-70.
8. Erdem HB, Bahsi T, Ergün MA. Function of telomere in aging and age related diseases. Environ Toxicol Pharmacol 2021 Jul;85:103641. [doi: 10.1016/j.etap.2021.103641] [Medline: 33774188]
9. Armanios M. The role of telomeres in human disease. Annu Rev Genomics Hum Genet 2022 Aug 31;23:363-381. [doi: 10.1146/annurev-genom-010422-091101] [Medline: 35609925]

XSL•FO
RenderX

10. Codd V, Nelson CP, Albrecht E, et al. Identification of seven loci affecting mean telomere length and their association with disease. Nat Genet 2013 Apr;45(4):422-427. [doi: 10.1038/ng.2528] [Medline: 23535734]
11. Mangino M, Hwang SJ, Spector TD, et al. Genome-wide meta-analysis points to CTC1 and ZNF676 as genes regulating telomere homeostasis in humans. Hum Mol Genet 2012 Dec 15;21(24):5385-5394. [doi: 10.1093/hmg/dds382] [Medline: 23001564]
12. Willeit P, Willeit J, Mayr A, et al. Telomere length and risk of incident cancer and cancer mortality. JAMA 2010 Jul 7;304(1):69-75. [doi: 10.1001/jama.2010.897] [Medline: 20606151]
13. Diez Roux AV, Ranjit N, Jenny NS, et al. Race/ethnicity and telomere length in the Multi-Ethnic Study of Atherosclerosis. Aging Cell 2009 Jun;8(3):251-257. [doi: 10.1111/j.1474-9726.2009.00470.x] [Medline: 19302371]
14. Epel ES. Psychological and metabolic stress: a recipe for accelerated cellular aging? HORMONES (Athens) 2009;8(1):7-22. [doi: 10.14310/horm.2002.1217] [Medline: 19269917]
15. Simon NM, Smoller JW, McNamara KL, et al. Telomere shortening and mood disorders: preliminary support for a chronic stress model of accelerated aging. Biol Psychiatry 2006 Sep 1;60(5):432-435. [doi: 10.1016/j.biopsych.2006.02.004] [Medline: 16581033]
16. Valdes A, Andrew T, Gardner J, et al. Obesity, cigarette smoking, and telomere length in women. Lancet 2005 Aug;366:662-664. [doi: 10.1016/S0140-6736(05)66630-5]
17. Pavanello S, Hoxha M, Dioni L, et al. Shortened telomeres in individuals with abuse in alcohol consumption. Int J Cancer 2011 Aug 15;129(4):983-992. [doi: 10.1002/ijc.25999] [Medline: 21351086]
18. McGrath M, Wong JYY, Michaud D, Hunter DJ, De Vivo I. Telomere length, cigarette smoking, and bladder cancer risk in men and women. Cancer Epidemiol Biomarkers Prev 2007 Apr;16(4):815-819. [doi: 10.1158/1055-9965.EPI-06-0961] [Medline: 17416776]
19. Sánchez-González JL, Sánchez-Rodríguez JL, Martín-Vallejo J, Martel-Martel A, González-Sarmiento R. Effects of physical exercise on cognition and telomere length in healthy older women. Brain Sci 2021 Oct 27;11(11):1417. [doi: 10.3390/brainsci11111417] [Medline: 34827416]
20. Cherkas LF, Hunkin JL, Kato BS, et al. The association between physical activity in leisure time and leukocyte telomere length. Arch Intern Med 2008 Jan 28;168(2):154-158. [doi: 10.1001/archinternmed.2007.39] [Medline: 18227361]
21. Puterman E, Lin J, Blackburn E, O'Donovan A, Adler N, Epel E. The power of exercise: buffering the effect of chronic stress on telomere length. PLoS One 2010 May 26;5(5):e10837. [doi: 10.1371/journal.pone.0010837] [Medline: 20520771]
22. Buttet M, Bagheri R, Ugbolue UC, et al. Effect of A lifestyle intervention on telomere length: a systematic review and meta-analysis. Mech Ageing Dev 2022 Sep;206:111694. [doi: 10.1016/j.mad.2022.111694] [Medline: 35760212]
23. Denham J, Sellami M. Exercise training increases telomerase reverse transcriptase gene expression and telomerase activity: a systematic review and meta-analysis. Ageing Res Rev 2021 Sep;70:101411. [doi: 10.1016/j.arr.2021.101411] [Medline: 34284150]
24. Mundstock E, Zatti H, Louzada FM, et al. Effects of physical activity in telomere length: systematic review and meta-analysis. Ageing Res Rev 2015 Jul;22:72-80. [doi: 10.1016/j.arr.2015.02.004] [Medline: 25956165]
25. Valente C, Andrade R, Alvarez L, Rebelo-Marques A, Stamatakis E, Espregueira-Mendes J. Effect of physical activity and exercise on telomere length: systematic review with meta-analysis. J Am Geriatr Soc 2021 Nov;69(11):3285-3300. [doi: 10.1111/jgs.17334] [Medline: 34161613]
26. Song S, Lee E, Kim H. Does exercise affect telomerelength? a systematic review and meta-analysis of randomized controlled trials. Medicina (Kaunas) 2022 Feb 5;58(2):242. [doi: 10.3390/medicina58020242] [Medline: 35208566]
27. Denham J, O'Brien BJ, Charchar FJ. Telomere length maintenance and cardio-metabolic disease prevention through exercise training. Sports Med 2016 Sep;46(9):1213-1237. [doi: 10.1007/s40279-016-0482-4] [Medline: 26914269]
28. Ludlow AT, Zimmerman JB, Witkowski S, Hearn JW, Hatfield BD, Roth SM. Relationship between physical activity level, telomere length, and telomerase activity. Med Sci Sports Exerc 2008;40(10):1764-1771. [doi: 10.1249/MSS.0b013e31817c92aa]
29. Werner CM, Hecksteden A, Morsch A, et al. Differential effects of endurance, interval, and resistance training on telomerase activity and telomere length in a randomized, controlled study. Eur Heart J 2019 Jan 1;40(1):34-46. [doi: 10.1093/eurheartj/ehy585] [Medline: 30496493]
30. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
31. Stone PW. Popping the (PICO) question in research and evidence-based practice. Appl Nurs Res 2002 Aug;15(3):197-198. [doi: 10.1053/apnr.2002.34181] [Medline: 12173172]
32. Moher D, Pham B, Jones A, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? Lancet 1998 Aug;352:609-613. [doi: 10.1016/S0140-6736(98)01085-X]
33. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017 Sep 21;358:j4008. [doi: 10.1136/bmj.j4008] [Medline: 28935701]
34. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282. [Medline: 23092060]

XSL•FO
RenderX

35. Whiting P, Savović J, Higgins JPT, et al. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. J Clin Epidemiol 2016 Jan;69:225-234. [doi: 10.1016/j.jclinepi.2015.06.005] [Medline: 26092286]
36. Torres A, Tennant B, Ribeiro-Lucas I, Vaux-Bjerke A, Piercy K, Bloodgood B. Umbrella and systematic review methodology to support the 2018 Physical Activity Guidelines Advisory Committee. J Phys Act Health 2018 Nov 1;15(11):805-810. [doi: 10.1123/jpah.2018-0372] [Medline: 30336718]
37. Bracchiglione J, Meza N, Bangdiwala S, et al. Graphical Representation of Overlap for OVErviews. : OSF; 2022. [doi: 10.17605/OSF.IO/U2MS4]
38. Skipka G. The inclusion of the estimated inter-study variation into forest plots for random effects meta-analysis - a suggestion for a graphical representation. In: Program and Abstract Book: Cochrane Colloquium; 2006:23-26.
39. Ortega Z, Martín-Vallejo J, Mencía A, Galindo-Villardón MP, Pérez-Mellado V. Introducing meta-partition, a useful methodology to explore factors that influence ecological effect sizes. PLoS One 2016;11(7):e0158624. [doi: 10.1371/journal.pone.0158624] [Medline: 27409084]
40. Barragán R, Ortega-Azorín C, Sorlí JV, et al. Effect of physical activity, smoking, and sleep on telomere length: a systematic review of observational and intervention studies. J Clin Med 2021 Dec 24;11(1):76. [doi: 10.3390/jcm11010076] [Medline: 35011817]
41. Schellnegger M, Lin AC, Hammer N, Kamolz LP. Physical activity on telomere length as a biomarker for aging: a systematic review. Sports Med Open 2022 Sep 4;8(1):111. [doi: 10.1186/s40798-022-00503-1] [Medline: 36057868]
42. Prathap L, Suganthirababu SP, Kumar KP, S P. Effect of physical exercise in remodeling telomere length and cancer prevention in an epigenetic prospect – a systematic review. Biomed Pharmacol J 2021 Jun 30;14(2):891-901. [doi: 10.13005/bpj/2191]
43. Qiao S, Jiang Y, Li X. The impact of health promotion interventions on telomere length: a systematic review. Am J Health Promot 2020 Jul;34(6):633-647. [doi: 10.1177/0890117120906958] [Medline: 32103672]
44. Min J, Kim JY, Choi JY, Kong ID. Association between physical activity and telomere length in women with breast cancer: a systematic review. J Clin Med 2022 Apr 30;11(9):2527. [doi: 10.3390/jcm11092527] [Medline: 35566652]
45. Marques A, Gouveira É, Peralta M, et al. Cardiorespiratory fitness and telomere length: a systematic review. J Sports Sci 2020 Jul;38(14):1690-1697. [doi: 10.1080/02640414.2020.1754739] [Medline: 32284029]
46. Himbert C, Thompson H, Ulrich CM. Effects of intentional weight loss on markers of oxidative stress, DNA repair and telomere length - a systematic review. Obes Facts 2017;10(6):648-665. [doi: 10.1159/000479972] [Medline: 29237161]
47. Smith DL, Wu C, Gregorich S, Dai G, Lin J. Impact of DNA extraction methods on quantitative PCR telomere length assay precision in human saliva samples. Int J Methodol 2022;1(1):44-57. [doi: 10.21467/ijm.1.1.5784]
48. LaRocca TJ, Seals DR, Pierce GL. Leukocyte telomere length is preserved with aging in endurance exercise-trained adults and related to maximal aerobic capacity. Mech Ageing Dev 2010 Feb;131(2):165-167. [doi: 10.1016/j.mad.2009.12.009] [Medline: 20064545]
49. Du M, Prescott J, Kraft P, et al. Physical activity, sedentary behavior, and leukocyte telomere length in women. Am J Epidemiol 2012 Mar 1;175(5):414-422. [doi: 10.1093/aje/kwr330] [Medline: 22302075]
50. Finkel T, Holbrook NJ. Oxidants, oxidative stress and the biology of ageing. Nat New Biol 2000 Nov 9;408(6809):239-247. [doi: 10.1038/35041687] [Medline: 11089981]
51. von Zglinicki T. Oxidative stress shortens telomeres. Trends Biochem Sci 2002 Jul;27(7):339-344. [doi: 10.1016/S0968-0004(02)02110-2]
52. Powers SK, Jackson MJ. Exercise-induced oxidative stress: cellular mechanisms and impact on muscle force production. Physiol Rev 2008 Oct;88(4):1243-1276. [doi: 10.1152/physrev.00031.2007] [Medline: 18923182]
53. Gomez-Cabrera MC, Domenech E, Viña J. Moderate exercise is an antioxidant: upregulation of antioxidant genes by training. Free Radic Biol Med 2008 Jan 15;44(2):126-131. [doi: 10.1016/j.freeradbiomed.2007.02.001] [Medline: 18191748]
54. Radak Z, Chung HY, Koltai E, Taylor AW, Goto S. Exercise, oxidative stress and hormesis. Ageing Res Rev 2008 Jan;7(1):34-42. [doi: 10.1016/j.arr.2007.04.004] [Medline: 17869589]
55. Belviranlı M, Gökbel H. Acute exercise induced oxidative stress and antioxidant changes. Electron J Gen Med 2006;3(3):126-131. [doi: 10.29333/ejgm/82392]
56. Fischer CP. Interleukin-6 in acute exercise and training: what is the biological relevance? Exerc Immunol Rev 2006;12:6-33. [Medline: 17201070]
57. Ornish D, Lin J, Daubenmier J, et al. Increased telomerase activity and comprehensive lifestyle changes: a pilot study. Lancet Oncol 2008 Nov;9(11):1048-1057. [doi: 10.1016/S1470-2045(08)70234-1] [Medline: 18799354]

Abbreviations

AMSTAR-2: A Measurement Tool to Assess Systematic Reviews
CCA: corrected covered area
GROOVE: Graphical Representation of Overlap for OVErviews
HIIT: high-intensity interval training
PAGAC: Physical Activity Guidelines Advisory Committee Grading Criteria

XSL•FO
RenderX

PICOS: Population, Intervention, Control, Outcomes, and Study design
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
ROBIS: Risk Of Bias in Systematic reviews
TL: telomere length

Edited by M O'Connell; submitted 19.07.24; peer-reviewed by MN Shalaby, R Sobral-Monteiro-Junior; revised version received 10.09.24; accepted 16.10.24; published 10.01.25.

Please cite as:
Sánchez-González JL, Sánchez-Rodríguez JL, González-Sarmiento R, Navarro-López V, Juárez-Vela R, Pérez J, Martín-Vallejo J
Effect of Physical Exercise on Telomere Length: Umbrella Review and Meta-Analysis
JMIR Aging 2025;8:e64539
URL: https://aging.jmir.org/2025/1/e64539
doi:10.2196/64539

© Juan Luis Sánchez-González, Juan Luis Sánchez-Rodríguez, Rogelio González-Sarmiento, Víctor Navarro-López, Raúl Juárez-Vela, Jesús Pérez, Javier Martín-Vallejo. Originally published in JMIR Aging (https://aging.jmir.org), 10.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Exploring the Feasibility of a 5-Week mHealth Intervention to Enhance Physical Activity and an Active, Healthy Lifestyle in Community-Dwelling Older Adults: Mixed Methods Study

Kim Daniels[1,2], MSc; Sharona Vonck[1], PhD; Jolien Robijns[1], PhD; Kirsten Quadflieg[1], PhD; Jochen Bergs[1,3], Prof Dr; Annemie Spooren[1,2], Prof Dr; Dominique Hansen[2,4], Prof Dr; Bruno Bonnechère[1,2,5], Prof Dr

[1]Centre of Expertise in Care Innovation, Department of PXL – Healthcare, PXL University of Applied Sciences and Arts, Hasselt, Belgium

[2]REVAL Rehabilitation Research Center, Faculty of Rehabilitation Sciences, Hasselt University, Diepenbeek, Belgium

[3]THINK3 Simulation & Innovation Lab, Faculty of Medicine and Life Sciences, Hasselt University, Diepenbeek, Belgium

[4]BIOMED, Faculty of Medicine and Life Sciences, Hasselt University, Diepenbeek, Belgium

[5]Technology-Supported and Data-Driven Rehabilitation, Data Sciences Institute, Hasselt University, Diepenbeek, Belgium

Corresponding Author:
Kim Daniels, MSc
Centre of Expertise in Care Innovation, Department of PXL – Healthcare
PXL University of Applied Sciences and Arts
Guffenslaan 39
Hasselt, 3500
Belgium
Phone: 32 485763451
Email: kim.daniels@pxl.be

Abstract

Background: Advancements in mobile technology have paved the way for innovative interventions aimed at promoting physical activity (PA).

Objective: The main objective of this feasibility study was to assess the feasibility, usability, and acceptability of the More In Action (MIA) app, designed to promote PA among older adults. MIA offers 7 features: personalized tips, PA literacy, guided peer workouts, a community calendar, a personal activity diary, a progression monitor, and a chatbot.

Methods: Our study used a mixed methods approach to evaluate the MIA app's acceptability, feasibility, and usability. First, a *think-aloud* method was used to provide immediate feedback during initial app use. Participants then integrated the app into their daily activities for 5 weeks. Behavioral patterns such as user session duration, feature use frequency, and navigation paths were analyzed, focusing on engagement metrics and user interactions. User satisfaction was assessed using the System Usability Scale, Net Promoter Score, and Customer Satisfaction Score. Qualitative data from focus groups conducted after the 5-week intervention helped gather insights into user experiences. Participants were recruited using a combination of web-based and offline strategies, including social media outreach, newspaper advertisements, and presentations at older adult organizations and local community services. Our target group consisted of native Dutch-speaking older adults aged >65 years who were not affected by severe illnesses. Initial assessments and focus groups were conducted in person, whereas the intervention itself was web based.

Results: The study involved 30 participants with an average age of 70.3 (SD 4.8) years, of whom 57% (17/30) were female. The app received positive ratings, with a System Usability Scale score of 77.4 and a Customer Satisfaction Score of 86.6%. Analysis showed general satisfaction with the app's workout videos, which were used in 585 sessions with a median duration of 14 (IQR 0-34) minutes per day. The Net Promoter Score was 33.34, indicating a good level of customer loyalty. Qualitative feedback highlighted the need for improvements in navigation, content relevance, and social engagement features, with suggestions for better calendar visibility, workout customization, and enhanced social features. Overall, the app demonstrated high usability and satisfaction, with near-daily engagement from participants.

Conclusions: The MIA app shows significant potential for promoting PA among older adults, evidenced by its high usability and satisfaction scores. Participants engaged with the app nearly daily, particularly appreciating the workout videos and educational content. Future enhancements should focus on better calendar visibility, workout customization, and integrating social networking features to foster community and support. In addition, incorporating wearable device integration and predictive analytics could

XSL•FO
RenderX

provide real-time health data, optimizing activity recommendations and health monitoring. These enhancements will ensure that the app remains user-friendly, relevant, and sustainable, promoting sustained PA and healthy behaviors among older adults.

Trial Registration: ClinicalTrials.gov NCT05650515; https://clinicaltrials.gov/study/NCT05650515

(JMIR Aging 2025;8:e63348) doi:10.2196/63348

KEYWORDS

mobile health; mHealth; feasibility; physical activity; older adults; health promotion; usability; mobile phone

Introduction

Background

Ensuring the well-being of the global aging population is a pressing concern, particularly as the World Health Organization predicts a significant rise in the number of individuals aged >65 years by 2050 [1]. To confront this impending public health challenge, the promotion of regular physical activity (PA) is paramount [2].

Research has consistently demonstrated the multifaceted advantages of regular PA on physical, cognitive, and mental health even in advanced age. In recognition of these benefits, the World Health Organization advocates for specific PA guidelines for individuals aged ≥65 years [3,4]. These guidelines encompass moderate-intensity aerobic activities, muscle-strengthening exercises, and balance training. However, despite the wealth of evidence supporting these recommendations, a substantial proportion of older adults worldwide fail to meet the prescribed PA levels [5,6]. Hence, adherence to recommended guidelines remains a challenging issue [7].

Addressing the gap in PA participation among older adults necessitates the implementation of innovative and sustainable interventions. Mobile technology has opened up avenues for innovative approaches to foster PA and cultivate healthier lifestyles even among older adults [8,9]. Mobile health (mHealth) apps have emerged as promising and cost-effective health intervention tools, especially in promoting PA. These apps harness technology to offer personalized guidance and track and encourage PA [10]. Capitalizing on the widespread use of smartphones and tablets, these apps possess the potential to revolutionize how we address and promote PA among aging populations [11]. They offer convenient access to PA advice and support, deliver an enjoyable user experience, and furnish feedback on progress over time [12-14].

Despite these positive trends, adoption of these technologies in real-life conditions is still relatively limited. Several challenges and hurdles must be overcome before such interventions can be successfully and sustainably implemented in the field. The challenge is not solely in developing mHealth apps but also in maintaining long-term engagement and motivation among older adults. In addition, studies often report high attrition rates and small intervention effects [14-16]. These problems must be addressed to use the full potential of mHealth interventions. First, high attrition is likely due to the intervention not matching the users' needs, goals, and expectations [17,18]. This may be avoided by involving potential users during the entire cycle of intervention development [19-22]. Second, research has revealed that theory-based interventions are more effective at modifying health behaviors than traditional interventions [23,24]. Thus, interventions should be grounded within and informed by theoretical models.

Objectives

In response to these challenges, a collaborative cocreation process led to the development and refinement of an mHealth app named More In Action (MIA) [19]. The app's content and design were crafted through cocreative workshops and based on the theoretical framework of the Behavior Change Wheel (BCW) [23]. Despite the active involvement of end users from the outset of development and throughout the iterative process, as well as preliminary results indicating high levels of enjoyment and ease of use of the app [19], the long-term retention rate and motivation levels of the participants remain unknown. Therefore, the primary objective of this study was to thoroughly explore the acceptability, feasibility, and usability of the MIA app in promoting PA and encouraging sustained, active, and healthy behaviors among its users.

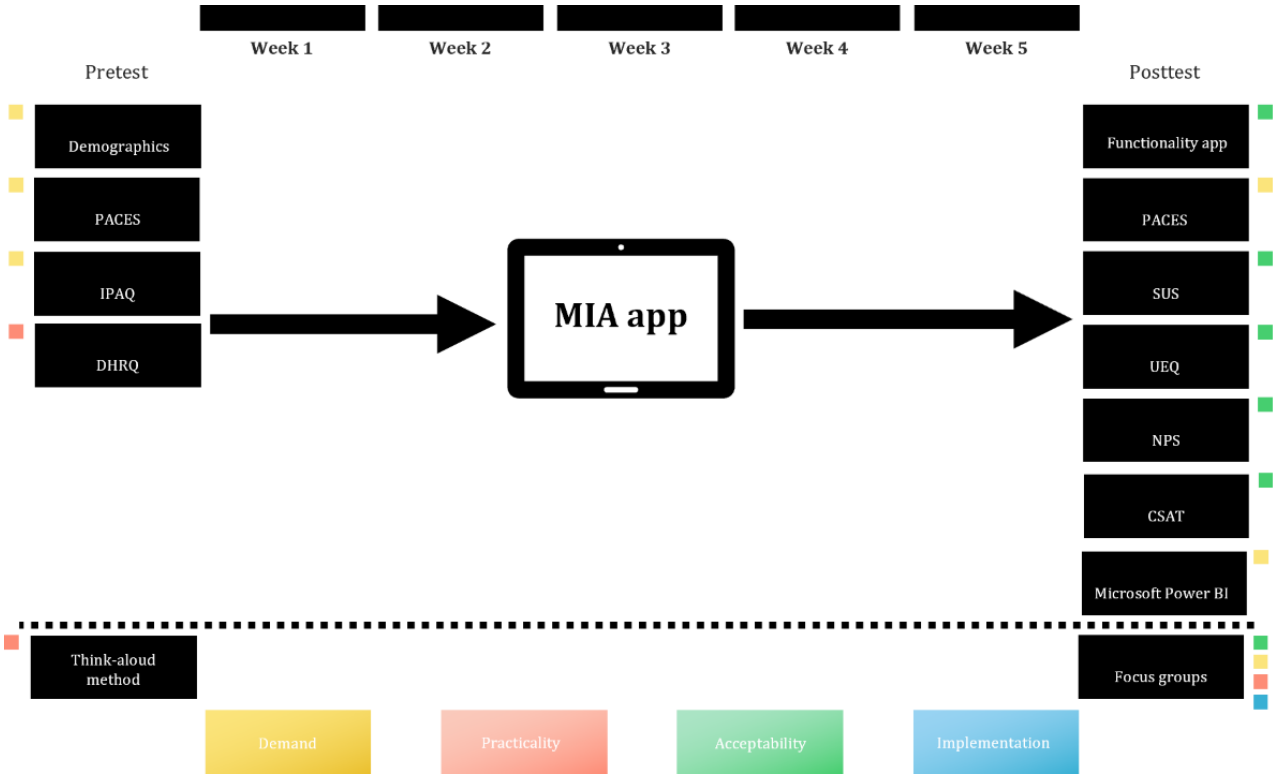
Methods

Study Design

We used a comprehensive mixed methods approach to assess the MIA app's acceptability, feasibility, and usability, as shown in Figure 1.

XSL•FO
RenderX

Figure 1. Study design. The colors represent the different evaluation focuses: yellow for demand (interest and need), orange for practicality (feasibility), green for acceptability (user satisfaction and acceptance), and blue for implementation (integration into the study). CSAT: Customer Satisfaction Score; DHRQ: Digital Health Readiness Questionnaire; IPAQ: International Physical Activity Questionnaire; MIA: More In Action; NPS: Net Promoter Score; PACES: Physical Activity Enjoyment Scale; SUS: System Usability Scale; UEQ: User Experience Questionnaire.



Our methodology adhered to the recommended framework for feasibility studies by Bowen et al [25]. This framework allowed us to explore critical aspects of the MIA app, such as *acceptability*, *demand*, *practicality*, and *implementation*. These interrelated concepts were systematically integrated into our study design and procedures, as detailed in Table 1.

Acceptability was defined as the perceived suitability of the app and is directly tied to user satisfaction and their intent to continue using the app, as shown in Table 1. Feasibility, on the other hand, addresses the ease of implementing the app among older adults [26], focusing on demand and implementation and examining whether it can be effectively implemented and sustained in users' daily routines. In Table 1, feasibility is broken down into 2 distinct areas. The first is demand, which addresses whether older adults are willing to adopt and regularly use the app. This area evaluates actual use, intention to use, and perceived demand. The second is practicality, which assesses how easily the MIA app can be integrated into the users' daily lives. This area was explored through the *think-aloud* method, focus groups, and questionnaires. Practicality measures the ease of integrating the app into daily routines, user willingness to pay, and any perceived positive or negative effects. Usability, a key aspect of our investigation, was defined as the ease of use and suitability of the system or product for a specific user group performing designated tasks in a particular environment, including practicality. In this context, *ease of use* directly impacted user performance and satisfaction, whereas *acceptability* determined the likelihood of the product being embraced and used [27]. In Table 1, usability directly impacts acceptability (satisfaction and intent to continue using the app) and feasibility (ease of integration and ability to perform the required intervention activities).

XSL•FO
RenderX

Table 1. Key areas of focus for the More In Action (MIA) feasibility study based on the framework by Bowen et al [25].

Area of focus and research question	Method	Outcome	Measures
Usability			
Acceptability			
To what extent is the MIA app as a means to promote PA[a] acceptable among community-dwelling adults aged ≥65 years?	• Questionnaire • Focus group	• Satisfaction • Intent to continue use	SUS[b], CSAT[c], NPS[d], and Likert-scale questions on how the app fits into the end users' daily-life activities
Feasibility			
Demand			
What is the level of adoption of the MIA app among community-dwelling older adults? What factors influence the intention to use and engage with the app?	• Questionnaire • Microsoft Power BI (Microsoft Corp) analytics • Plausible Analytics	• Actual use • Intention to use • Perceived demand	Questions regarding the influence of the app on modulating PA behavior (eg, PACES[e] and IPAQ[f]) and analytics to compare the frequency of use and patterns of use across the participants
Practicality			
To what extent can the MIA app be integrated into the daily lives of older adults aged ≥65 years residing within the community?	• Think-aloud method • Questionnaire • Focus group	• Positive or negative effects • Ability of participants to execute intervention activities • Willingness to pay	First impressions, Likert scale, and in-depth questions regarding integrating the app into their daily lives (eg, DHRQ[g])
Implementation			
How can the MIA app be optimally implemented to facilitate sustained engagement in PA among community-dwelling older adults?	• Focus group	• Amount and type of resources needed to implement • Factors affecting implementation	In-depth questions on how the MIA app can be deployed in the community context

[a]PA: physical activity.

[b]SUS: System Usability Scale.

[c]CSAT: Customer Satisfaction Score.

[d]NPS: Net Promoter Score.

[e]PACES: Physical Activity Enjoyment Scale.

[f]IPAQ: International Physical Activity Questionnaire.

[g]DHRQ: Digital Health Readiness Questionnaire.

Ethical Considerations

This study was registered at Clinical Trials.gov (NCT05650515) and was approved by the ethical committee of Hasselt University (B1152023000011). All participants provided informed consent. To ensure privacy, the data used in this study were deidentified before analysis. Participants did not receive any compensation for their involvement.

Participants

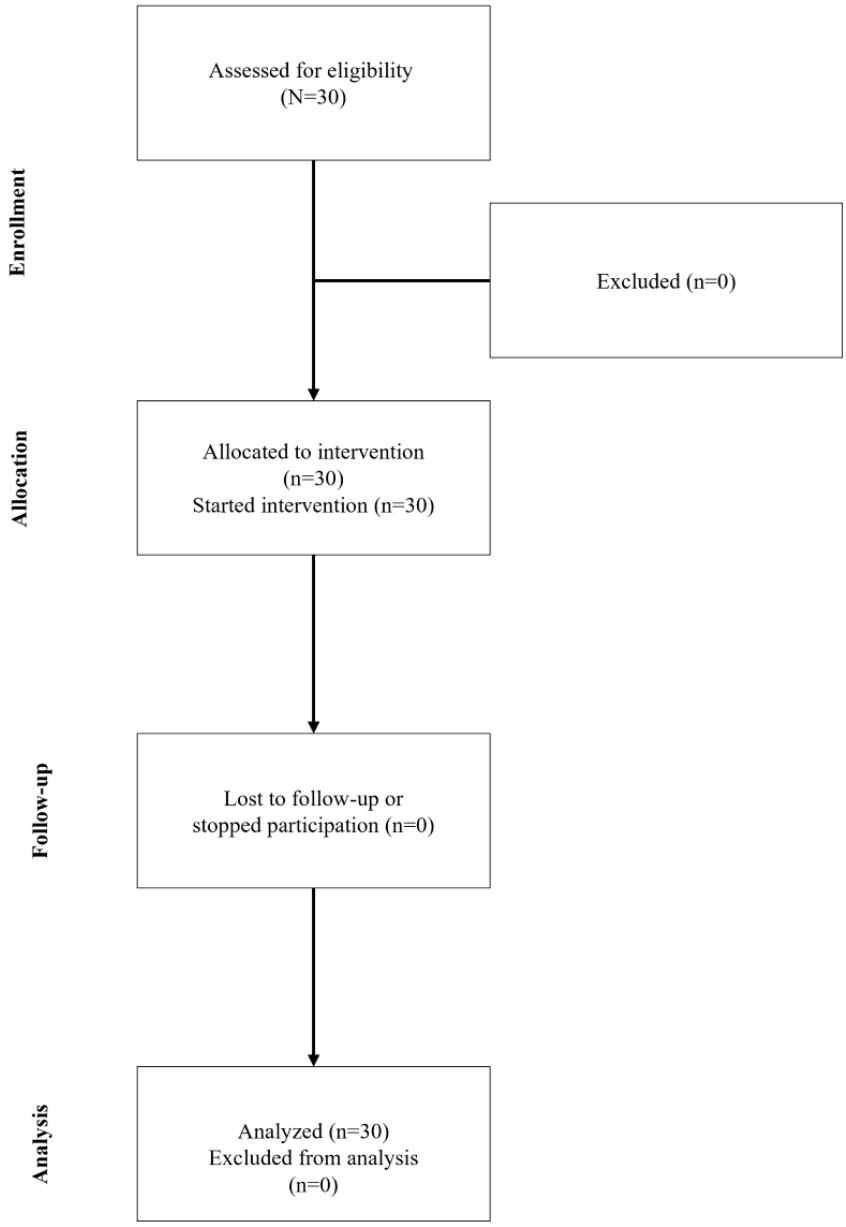
Older adults without severe illness were invited to participate in this study. They were recruited via social media outreach, newspaper advertisements, and pitches at several older adult organizations and through the local community services during the recruitment period from August 2023 to September 2023.

The inclusion criteria for participants encompassed that they had to be aged ≥65 years, competent to provide informed consent (Multimedia Appendix 1), able to actively participate in the study, community dwelling (living either independently at home or in a serviced apartment), without any severe illnesses, and native Dutch speakers. The exclusion criteria were the presence of current neurological, cardiovascular, respiratory, severe metabolic, or cognitive disorders. Participants were not excluded based on digital literacy, ensuring a diverse range of digital competence levels. The complete list of exclusion criteria is presented in Multimedia Appendix 2.

Figure 2 provides a CONSORT (Consolidated Standards of Reporting Trials) flowchart for this single-arm feasibility study. The methodology and results were reported following the CONSORT 2010 checklist [28].

XSL•FO
RenderX

Figure 2. Modified CONSORT (Consolidated Standards of Reporting Trials) flow diagram for a single-arm feasibility study of the 5-week More In Action intervention.



Intervention

General Description

The mHealth app MIA represents an innovative approach to promoting PA and fostering a lifestyle centered on health and activity among older adults. Developed through a collaborative cocreation process, MIA has been refined over several iterations to meet the specific needs and preferences of its target user group [19]. MIA is optimized for both smartphone and tablet use, thereby ensuring accessibility and user engagement across a broad spectrum of mobile devices.

The app's design and content draw upon insights gained from cocreative workshops and are firmly grounded in the intervention functions of the BCW theoretical framework [29] as well as principles from self-identification theory [30,31] as key elements to motivating individuals toward healthier behaviors. Textbox 1 presents a comprehensive summary of the behavior change techniques incorporated into MIA.

XSL•FO
RenderX

Textbox 1. Overview of behavior change techniques implemented in the More In Action (MIA) app.

Behavior Change Wheel (BCW) [32]

- The BCW framework is centered on 3 core components: capability, opportunity, and motivation (Capability, Opportunity, and Motivation–Behavior system), which can drive behavior change. The MIA app enhances users' *capability* through educational content and reminders, increases *opportunity* through social features such as the community calendar, and boosts *motivation* by offering personalized goals and feedback.

Self-determination theory (SDT) [33]

- SDT emphasizes 3 fundamental psychological needs—autonomy, competence, and relatedness—as essential for motivation toward healthy behavior changes:
- *Autonomy*: the MIA app can provide users with the ability to customize their exercise routines, choose their goals, and select the types of physical activity (PA) they prefer. The workouts are composed personally based on the tailored personal goals users enter after registering.
- *Competence*: to enhance users' feelings of competence, the MIA app incorporates a system of progressive challenges and feedback. The app tracks users' progress in real time, offering progress bars or celebratory messages for achieving milestones (eg, achieving the World Health Organization guideline for PA). In addition, educational content on the benefits of regular PA and other health benefits helps users feel more skilled and capable.
- *Relatedness*: the MIA app fosters a sense of relatedness by integrating a community calendar that allows users to connect with others.

Self-identification theory (SIT) [31]

- Implementing SIT within the MIA app involved creating features that allow older adults to integrate PA into their self-concept, making it a core part of their identity. Users start by *selecting goals* that resonate with their personal aspirations and lifestyle. These goals can range from improving health, gaining strength, and enhancing mobility to participating in community activities or playing with grandchildren. They choose freely. The key is for users to choose goals that reflect their values and how they perceive themselves.

Unique Attributes of MIA

The MIA app exhibits several unique attributes within the scope of gerontechnology that set it apart from other existing mHealth technologies. To compare MIA with existing solutions, we conducted a review to map the currently available mHealth technologies for promoting PA in older adults.

All technologies were scored according to 11 key components previously identified [19,20,34]. These findings are detailed in Table 2.

XSL•FO
RenderX

Table 2. Comparative analysis of mobile health app features for enhancing physical activity among older adults.

Existing apps	User-centered design[a]	Behavior change techniques[b]	Personalized intervention[c]	Interactivity[d]	Activities of daily living[e]	Integration with wearable devices[f]	Social cohesion[g]	Education and information[h]	Rewards and incentives[i]	Accessibility and inclusion[j]	Older adult–specific features[k]
Standing-Tall [35]			✓[l]	✓						✓	✓
Web + [36]		✓	✓			✓	✓	✓	✓	✓	✓
Ready Steady Go [37]		✓		✓			✓	✓	✓	✓	✓
HBex [24]	✓	✓	✓	✓	✓			✓	✓	✓	
Vivo [38]	✓			✓						✓	
App-based exercise program [39]			✓		✓			✓		✓	✓
Physitrack [40]			✓	✓						✓	
My plan 2.0 [41]	✓	✓	✓	✓				✓	✓	✓	✓
Bingocize [42]				✓			✓	✓	✓	✓	✓
Fit for All platform [43]			✓	✓					✓		
ActiveLifestyle [44]		✓	✓					✓		✓	✓
Vibrotactile app [45]	✓		✓	✓					✓		✓
PACE[m] app [46]			✓		✓	✓			✓	✓	
Gymcentral [47]	✓	✓	✓	✓			✓		✓	✓	✓
Telehealth intervention [48]			✓	✓			✓				
Nymbl [49]			✓	✓				✓	✓		✓
eLIFE [50]		✓	✓	✓				✓	✓	✓	✓
Exercise app [51]			✓	✓					✓	✓	✓
Make Movement Your Mission [52]		✓		✓	✓		✓	✓	✓	✓	✓
Evident [53]		✓	✓	✓		✓		✓	✓		✓
MIA[n] app [19]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

[a]This design prioritizes the end user's needs, preferences, and limitations throughout the development process.

[b]These are systematic strategies derived from behavioral science theories to influence and sustain behavior modification.

XSL•FO
RenderX

[c]Personalized interventions involve tailoring health-related strategies and communications to individual users based on specific data gathered about their behaviors, preferences, and environmental contexts.

[d]Refers to the dynamic capability of the app to engage users through direct and responsive interactions.

[e]Integrates physical activity into routine daily tasks to reduce perceived barriers and enhance the practicality of exercises.

[f]This component involves the app's capability to synchronize with wearable technology to gather continuous physiological data, which can be used for monitoring health conditions in real time.

[g]Encourages the formation of supportive social networks within the app, enhancing user engagement through community building.

[h]Delivers evidence-based health information and instructional content to improve knowledge and skills related to physical activity.

[i]Uses motivational elements such as web-based badges, achievement unlocking, and progress tracking to enhance motivation and encourage continual app engagement.

[j]Accessibility ensures that products and services are usable by people with various abilities, whereas inclusivity focuses on creating environments that accommodate and welcome diverse individuals across all backgrounds and needs.

[k]An older adult–specific feature is a design element in a product or service, particularly mobile apps, that addresses the physical, cognitive, and social needs of older adults, such as mobility, sensory impairments, chronic conditions, and isolation, to enhance usability, safety, and autonomy.

[l]Presence of component.

[m]PACE: Physical Activity Cardiorespiratory Exercise.

[n]MIA: More In Action.

Features

MIA incorporates 7 major features: tailor-made tips, literacy initiatives to enhance awareness about PA and a healthy lifestyle, guided exercise workouts with peers, a community calendar fostering social connections, a personal diary for manual uploads of non–app-based PAs, a progression monitor, and a chatbot named MIA. The chatbot serves as a platform where users can pose questions and receive insights, motivation, and support in return. The current chatbot function is managed manually, with researchers on the development team directly addressing user inquiries. This approach allows for personalized responses and ensures that user concerns are addressed with accuracy and expertise but also for the collection of data that will be used later on to develop an artificial intelligence (AI)–driven chatbot. This AI chatbot would maintain human oversight to ensure the accuracy and relevance of the responses, particularly for complex queries. The app also includes 195 guided workout videos, each featuring 30-second exercise bursts followed by breaks. These exercises cover strength, endurance, coordination, and balance and are demonstrated by peers. Workouts are tailored to 3 skill levels—beginner, advanced, and expert—and range from 10 to 20 minutes. MIA continuously adapts workouts based on user experiences; for example, if an exercise is too challenging, the algorithm adjusts future sessions to ensure that they remain enjoyable and achievable.

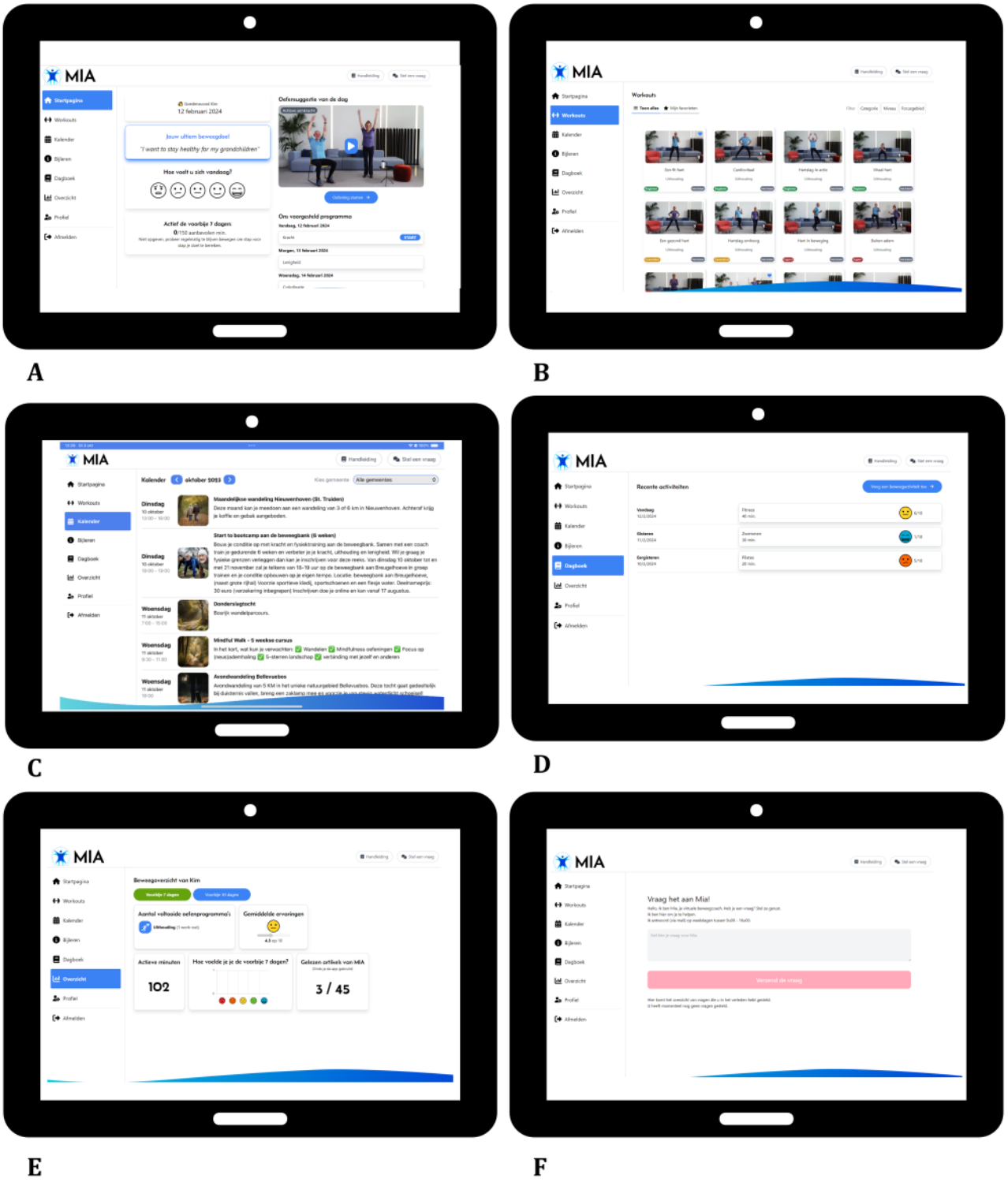
User Onboarding and Interaction

Upon initiating MIA, users undergo a preliminary assessment through a questionnaire that informs the customization of their exercise regimen. This personalization is central to MIA's approach, tailoring the app experience to individual needs and preferences.

As shown in Figure 3, the home page presents a personalized PA agenda that adapts daily based on user feedback to align with their physical and mental health status. This adaptive feature is critical for tailoring the experience to the user's day-to-day condition, emphasizing the interconnection between mental well-being and PA.

XSL•FO
RenderX

Figure 3. Screenshots of the More In Action interface: (A) home page displaying a physical activity (PA) agenda, suggested workout videos, and daily tips designed to engage users with personalized fitness guidance; (B) library of workout videos with filters for exercise type, fitness level, and target body areas supporting a tailored workout experience; (C) a community calendar for local PA events; (D) manual logging of nonapp PAs; (E) PA progress monitor; and (F) a web-based coach. MIA: More In Action.



Throughout the intervention period, participants were not required to use the app daily. Instead, following the principles of behavior change techniques, autonomy was emphasized, allowing users to set personalized goals and choose their own frequency of app engagement. This flexibility enabled participants to engage with the app at their preferred rate, from as little as once a week to daily over the course of the 5-week intervention.

XSL•FO
RenderX

Outcome Measures

Quantitative Measures

At Baseline

Demographics and PA levels were first assessed during the baseline measurement. The International Physical Activity Questionnaire–Short Form (IPAQ-SF) [54], the Physical Activity Enjoyment Scale (PACES) [55], and the Digital Health Readiness Questionnaire (DHRQ) [56] were used to measure digital health readiness via a web-based questionnaire using Qualtrics (Qualtrics International Inc). The IPAQ-SF showed moderate validity with accelerometry but had wide limits of agreement, indicating caution for longitudinal use [57,58]. Both the PACES [55,59] and DHRQ [56] demonstrated strong internal consistency and reliability, making them suitable for assessing enjoyment of PA and digital health readiness, respectively.

The 5-Week Postintervention Assessment

To evaluate user satisfaction and experience, 3 established instruments were used: the System Usability Scale (SUS) [60], Net Promoter Score (NPS) [61], and Customer Satisfaction Score (CSAT) [62]. The SUS provides a standardized assessment of usability through 10 items with response options ranging from *Strongly agree* to *Strongly disagree*. Scores of >68 indicate above-average usability. The SUS has demonstrated high internal consistency (Cronbach α=0.74) and test-retest reliability (Pearson correlation coefficient=0.75), supporting its use in evaluating digital health interventions [63]. The NPS measures user loyalty by asking participants to rate their likelihood of recommending the app on a scale from 0 to 10. This metric is calculated by asking older adults the following: *On a scale from 0 to 10, how likely are you to recommend MIA-app to a friend or peer?* Respondents are classified into 3 categories: promoters (those who provide a score of 9-10), passives (those who provide a score of 7-8), and detractors (those who provide a score of 0-6). The NPS, derived by subtracting the percentage of detractors from that of promoters, offers a straightforward approach to understanding user loyalty and referral potential [62].

The CSAT quantifies user satisfaction using a 10-point scale, capturing immediate feedback on the app experience. App-specific concerns (user-friendliness, layout, and utility) were also evaluated across multiple categories, with ratings ranging from 1 to 5 (1-2 indicating dissatisfaction, 3 indicating neutrality, and 4-5 indicating satisfaction).

Microsoft Power BI (Microsoft Corp) was used to analyze participant behavior during app use to provide insights into app use analytics and user engagement patterns [64]. Data on metrics such as workout frequency, session duration, and instances of premature workout termination were collected using integrated tracking tools such as Plausible Analytics. These data were imported into Microsoft Power BI for further analysis. Web-based dashboards and visualizations, including bar charts, were created to represent key engagement metrics and user navigation patterns.

Qualitative Measures

At Baseline

A *concurrent think-aloud* approach allowed participants to voice their thoughts and actions while using the app throughout installation and use [65,66]. Usability and enhancement suggestions were supplied to the researchers in real time [67-69]. Multimedia Appendix 3 details the think-aloud protocol, with an average procedure time of 20 (SD 6.2) minutes. These sessions were conducted in our laboratory at PXL University of Applied Sciences and Arts.

The 5-Week Postintervention Assessment

Qualitative data from focus groups were also collected to better understand app users' experiences. Following the 5-week trial of the app, the participants were invited to take part in focus groups in our laboratory at PXL University of Applied Sciences and Arts, each consisting of 6 older adults. Each focus group session lasted 2 hours and aimed to explore participants' perceptions of the app along with their reflections and suggestions for improvement. The sessions covered all app features. The focus groups were facilitated by experienced researchers (JR and KD) who are specialized in qualitative research. Multimedia Appendix 4 contains the interview guide.

Data Analysis

Quantitative Analysis

The quantitative analysis used mostly descriptive statistics to examine the data. Frequency distributions were also used to observe the occurrence of data values or categories, offering a detailed view of data distribution across outcomes. In addition, Microsoft Power BI analytics played a crucial role in visualizing and interpreting the data. Microsoft Power BI allowed for dynamic dashboards and visualizations, making it easier to explore trends, patterns, and relationships within the data. By integrating various data sources, it enabled the creation of charts, graphs, and tables that provided real-time insights into participant behavior and app use.

Qualitative Analysis

The think-aloud method and focus groups were audio recorded with participants' consent for the qualitative component. The transcriptions of the focus groups were carried out verbatim, followed by a rigorous content analysis using the NVivo software (QSR International) [70]. Themes were systematically coded based on the guidelines by Braun and Clarke [71]. The coding process was independently carried out by 2 experienced researchers (KD and JR), ensuring reliability and reducing the potential for bias in theme interpretation.

Integration of Qualitative and Quantitative Results

The Nonadoption, Abandonment, Scale-up, Spread, and Sustainability (NASSS) framework was used to integrate both the qualitative and quantitative results of the study, providing a structured analysis of the potential for sustainable implementation of the MIA app [72]. Recognized for its utility in evaluating complex technological health interventions, the NASSS framework allows for a systematic examination of key factors influencing the deployment and long-term sustainability

XSL•FO
RenderX

of digital health technologies. The analysis encompassed the framework's 7 domains: the condition being addressed (physical inactivity among older adults), the technological characteristics of the MIA app (including usability and technical robustness), the value proposition (perceived benefits to end users and stakeholders), the adopter system (end users and supporters), the organizational context (integration within health care systems), the wider sociopolitical environment (regulatory and cultural considerations), and the processes of embedding and adapting the technology over time.

Results

Demographic Characteristics

A total of 30 participants (mean age 70.3, SD 4.8 y; n=17, 57% female) were included in this study. Participants had different educational backgrounds; the largest proportion (15/30, 50%) had completed higher education. Their level of digital literacy was commendable, displaying a mean score of 59.6 (SD 8.8) out of 75 on the DHRQ. Notably, 80% (24/30) of the participants acknowledged the affirmative impact of digitalization on health, whereas 73% (22/30) actively engaged with social media, among whom 36% (8/22) used it daily, 23% (5/22) used it often, 23% (5/22) used it occasionally, and 14% (3/22) used it rarely. Furthermore, 67% (20/30) of the participants sought health-related information on the web, whereas all 30 participants reported internet use, with 27 (90%) of them accessing it daily, and 17 (57%) owned a smartwatch. When it came to monitoring PA through their smartphones' health apps, 73% (22/30) of the participants checked their step count daily, 13% (4/30) did so frequently, 3% (1/30) did so occasionally, 3% (1/30) did so rarely, and 10% (3/30) never did. Regarding health-related app use, 17% (5/30) of the participants actively used this type of apps.

The IPAQ-SF revealed variability in PA levels among individuals. The median activity levels indicated moderate engagement in walking (1386, IQR 284.8-2178 metabolic equivalent of task min/wk) and lighter involvement in vigorous activities (median 0, IQR 0-1830), suggesting a skew toward lower-intensity exercises. Moreover, it was found that 10% (3/30) of the participants were classified as inactive, 40% (12/30) fell into the category of minimal activity, and a significant 50% (15/30) were categorized as vigorously active. High SDs across all categories underscored a wide range of PA levels, from minimal to extremely active, reflecting the diverse nature of physical engagement in this sample.

The PACES results revealed that participants, on average, reported a moderate to high level of enjoyment during PA, with a mean score of 107 [55]. Considering that the PACES scale ranges from 18 to 126, where higher values signify greater enjoyment, the mean score of 107 suggests that participants' enjoyment levels were significantly skewed toward the upper end of the scale. This positioning implies a generally positive perception of PA among the participants. The complete sociodemographic characteristics of the participants are presented in Table 3.

XSL•FO
RenderX

Table 3. Participant characteristics (N=30).

Variable	Values
Age (y), mean (SD)	70.3 (4.8)
Sex (female), n (%)	17 (57)
Marital status, n (%)	
Single	2 (7)
Living together	1 (3)
Married	21 (70)
Divorced	4 (13)
Widowed	2 (7)
Educational level, n (%)	
Primary school	1 (3)
Middle school	6 (20)
University of applied sciences	15 (50)
University	8 (27)
Fall incidence (yes), n (%)	3 (10)
Digital literacy score (DHRQ [a] ; out of 75), mean (SD)	59.6 (8.8)
Use (out of 20)	16.3 (3.1)
Skills (out of 25)	21.3 (3.2)
Literacy (out of 15)	12.3 (2.2)
Health literacy (out of 15)	9.8 (3.2)
Learnability (out of 25)	20.7 (2.5)
Physical activity level (IPAQ - SF [b]), median (IQR)	
Total	3273 (1345.5-3873)
Walking	1386 (284.8-2178)
Moderate	1020 (310-2070)
Vigorous	0 (0-1830)

[a]DHRQ: Digital Health Readiness Questionnaire.

[b]IPAQ-SF: International Physical Activity Questionnaire–Short Form.

Quantitative Analysis

Acceptability

First, the general acceptability was accessed using the 3 main indicators, as shown in Figure 4.

XSL•FO
RenderX

Figure 4. Results of the System Usability Scale, Customer Satisfaction Score, and Net Promoter Score (NPS). MIA: More In Action.

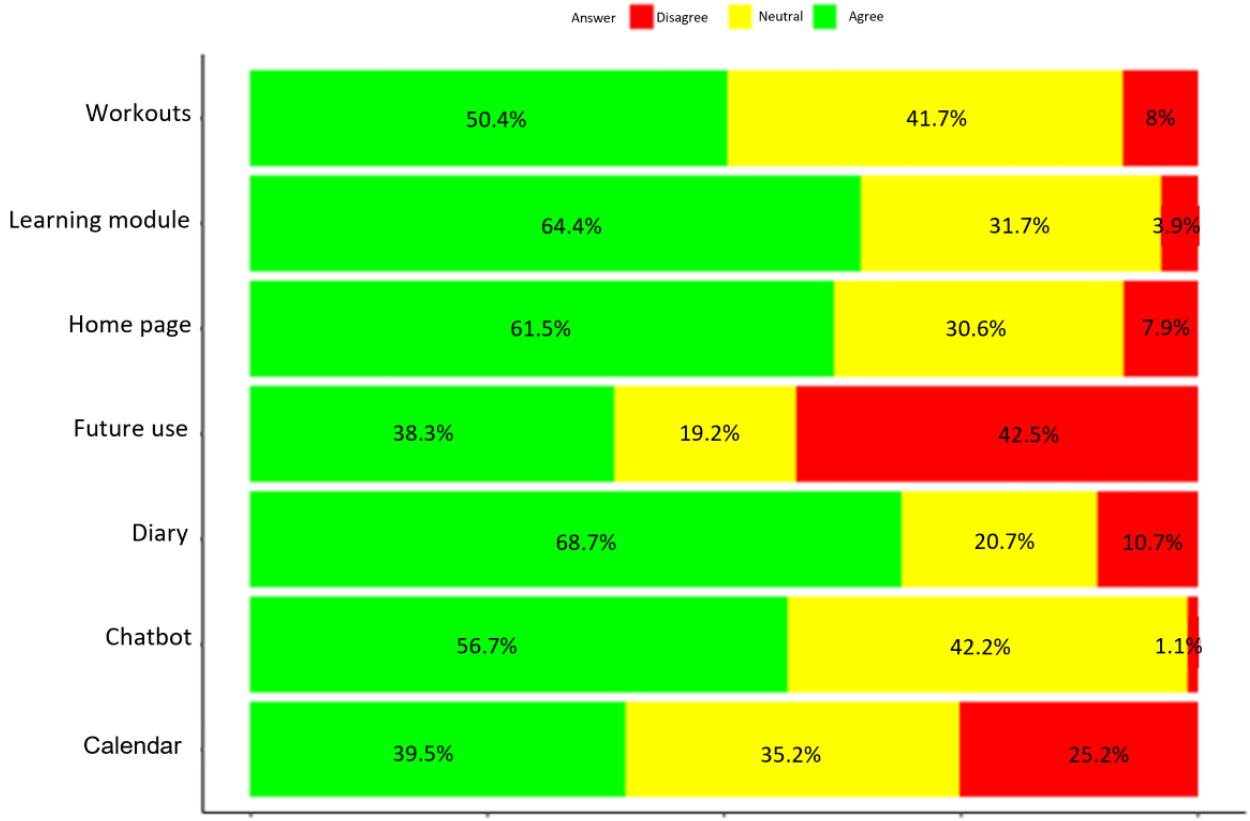


According to the SUS, the mHealth app was considered *acceptably good*, with a mean score of 77.4 (SD 14.3). The results of the CSAT revealed that 47% (14/30) of the participants indicated that they were *very satisfied*, whereas 40% (12/30) reported being *satisfied*. Another 10% (3/30) had a neutral response, and only 3% (1/30) were *dissatisfied*. None of the participants provided a *very dissatisfied* score. This distribution led to a CSAT score of 86.6%, indicating a high satisfaction level among older adults using the MIA app. The NPS was used to assess customer satisfaction with and loyalty and enthusiasm toward the app. A total of 50% (15/30) of the participants were *promoters*, indicating that they would recommend the MIA app to peers and family. In addition, 33% (10/30) were *passives*, reflecting moderate satisfaction without strong advocacy, whereas 17% (5/30) were *detractors*, indicating that they were unlikely to recommend the app. The resulting NPS is calculated by subtracting the percentage of detractors from the percentage of promoters, yielding a total NPS of 33.34.

Furthermore, the specific acceptability of the different features was assessed. The quantitative results also guided the qualitative part of the study, where focus groups were conducted to dive deeper into participants' reflections and gather additional insights. The aggregated results are presented in Figure 5. The complete results per category are presented in Multimedia Appendix 5.

XSL•FO
RenderX

Figure 5. User satisfaction and engagement—categorized ratings of app features and interface (1-2: dissatisfaction [red]; 3: neutrality [yellow]; 4-5: satisfaction [green]).



User Behavioral Patterns

To track user navigation and behavior within the app, Microsoft Power BI analytics were collected on the following features: workout videos, learning modules, community calendar, manual diary, progression monitor, and chatbot. Visual representations of the Microsoft Power BI analytics are available in Multimedia Appendix 6.

Workout Videos

A total of 585 workout videos were analyzed, highlighting general satisfaction with a median satisfaction score of 4 (IQR 3-4; scale of 1-5) and physical exertion rated at a median of 5 (IQR 5-7; scale of 0-10). Users engaged for a median duration of 14 (IQR 0-34) minutes per session, with use times ranging from 0 to 34 minutes. Strength workouts were most common, comprising 40% (234/585) of the sessions, followed by endurance, flexibility, and balance workouts. Most workouts (429/585, 73.3%) were aimed at beginners, with only 5% (29/585) targeting expert levels, reflecting the app's focus on older adults. Notably, 77% (23/30) of users expressed no preference for a specific type of exercise, suggesting a wide acceptance of the available workout options.

Learning Modules

The distribution of the learning modules comprised 255 read articles across 6 different topics: mental health, social well-being, physical well-being, nutrition, sleep, and risk of falling. Physical well-being was the most frequently engaged with topic, with 26.3% (67/255) of the total read articles. Nutrition followed with 17.3% (44/255) of the read articles, whereas mental health comprised 16.9% (43/255) of the read articles.

Community Calendar

Only 3 activities were added to the participants' community agenda, all of which were organized walks.

Manual Entries in Diary

The analysis of 545 external activities showed a median *general feeling after exercise* score of 4 (IQR 4-4) on a scale from 2 to 5, indicating generally positive responses. The range of perceived exertion scores was broader, with a median of 5 (IQR 5-7) indicating a balance between lower and higher levels of perceived effort. Walking emerged as the most popular activity (210/545, 38.5% of external activities), followed by cycling with electric assistance (78/545, 14.3%); cycling without electric assistance (66/545, 12.1%); and other activities such as padel, tennis, strength training, and running. The duration of these activities varied significantly, with a median time of 60 (IQR 40-90) minutes ranging from 5 to 400 minutes, indicating diverse engagement levels. The data revealed that a quarter of the participants (8/30, 27%) engaged in activities for <40 minutes, whereas a significant portion spent ≥90 minutes.

Chatbot

During the 5-week period, the chatbot was consulted 66 times. Most of the interactions involved feedback, accounting for 48% (32/66) of instances. Communication regarding exercise differentiation occurred 15% (10/66) of times, whereas

XSL•FO
RenderX

medical-related queries were raised 5% (3/66) of times. Technical issues led to 9% (6/66) of consultations, and motivational support prompted 17% (11/66) of interactions. The remaining 6% (4/66) of consultations fell into other categories.

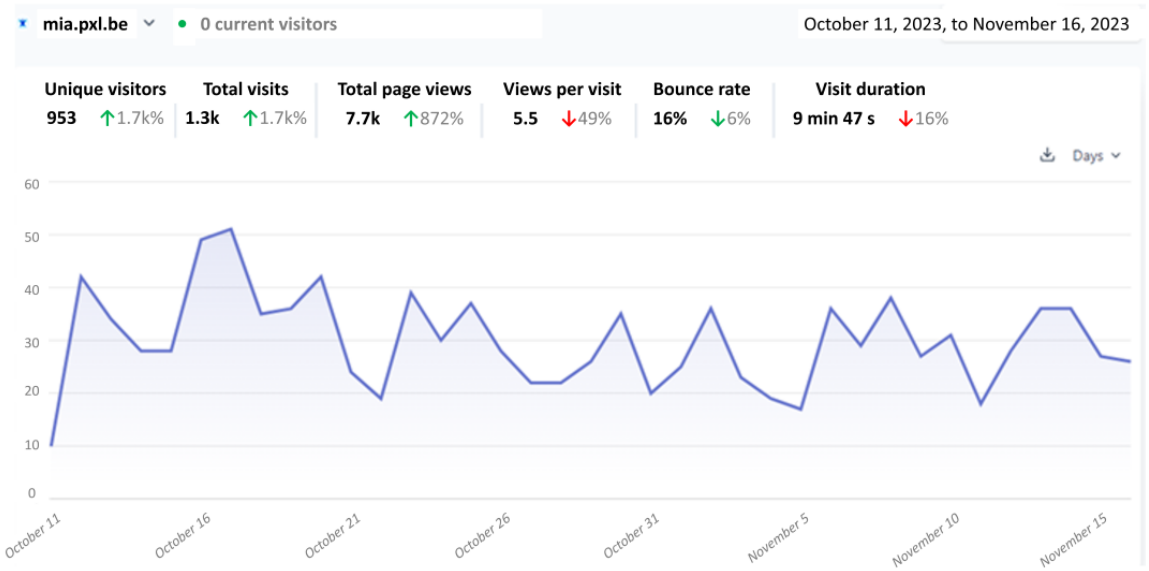
Adherence

During the 5-week intervention, app adherence was measured by tracking both the frequency of use and interaction patterns. Rather than requiring daily use, the study sought to assess whether participants engaged with the app according to the personalized goals and self-determined use frequencies they had set. The intervention, guided by self-determination theory [33], prioritized autonomy, enabling participants to select their own goals and decide how often to use the app. Engagement levels varied, with participants using the app between 3 and 7 times per week, with a median of 5 (IQR 4-6) times per week. Adherence rates averaged 135%, with individual rates ranging from 21% to 271%, reflecting the diverse levels of user interaction throughout the intervention.

On average, participants engaged with the app 32.88 (SD 14.06) times, ranging from 11 to 59 times used, which translates to an approximate daily use rate of 0.94 (SD 0.40) times. This indicates near-daily use of the app among the study cohort.

Despite the considerable variability in workout interactions within the app, with a median workout frequency of 17 sessions (IQR 11.2-28.5), motivation remained notably stable throughout the study, with some dips during the weekends. Participant engagement and interaction patterns throughout the 5-week intervention are visualized in Figure 6. Daily engagement varied between 0.66 and 3 sessions per day; however, there were no discernible fluctuations in adherence over time. Unlike the common pattern of an initial surge in activity followed by a dip and potential recovery, participants demonstrated consistent interaction with the app across the 5-week intervention.

Figure 6. Participant engagement and interaction patterns over the 5-week intervention. The y-axis shows the number of visits to the More In Action app.



Qualitative Analysis

Think-Aloud Method

The thematic analysis of the *think aloud* data resulted in the identification of 3 major key themes presented in Textbox 2. The other results and functionality are briefly described in the following sections.

Participants began interacting with the MIA app by opening it and completing their profiles. The welcome screen was inviting and easy to navigate, but some participants encountered difficulties with the profile questionnaire, highlighting a need for clearer guidance. When exploring the home screen, the interface was straightforward, with large icons and a simple layout. However, some features such as mood indicators were confusing, indicating a need for better explanations. The workout section was well received for its variety and suitability for different fitness levels, enhancing user experience. However, navigating the calendar feature proved challenging for some, with issues in adding or removing items pointing to a need for a more intuitive design. The learning module page was appreciated for its informative content, although some noted a lack of retirement-related topics, suggesting room for expansion. The diary functionality faced some hurdles as users found it difficult to add past activities, needing more flexible controls. Visually, the app was praised for its appealing design with bright colors and clear typography, making it esthetically pleasing and not overly complex. However, specific challenges such as difficulties with the profile questionnaire and mood indicators revealed that clarity in guidance is also crucial. Thus, while the app's visual appeal contributes to its overall usability, the complexity of certain features points to the need for balancing esthetics with intuitive guidance.

XSL•FO
RenderX

Textbox 2. Key themes of the think-aloud protocol.

Usability and accessibility

- The participants' experiences reveal that usability and accessibility are crucial for a successful app. Challenges with completing profiles, navigating calendars, and adding diary entries highlight areas in which the app could improve. Ensuring that the app is intuitive and easy to use can lead to a better user experience.

User interface and design

- The overall design and layout of the app play a significant role in user satisfaction. The positive feedback on the app's simplicity, large icons, and bright colors shows that a visually appealing and straightforward design enhances usability. This theme indicates that maintaining a clean and user-friendly interface is essential.

Feature functionality and clarity

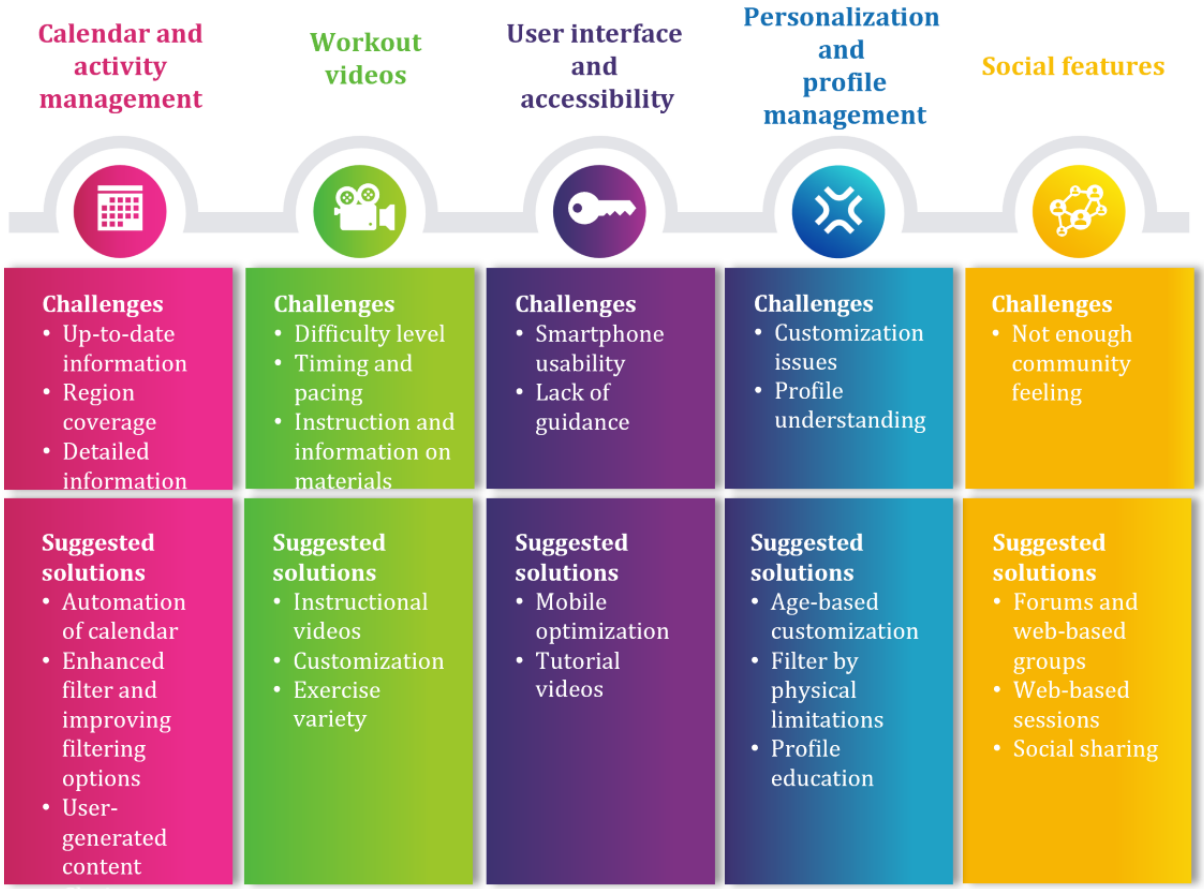
- The confusion regarding certain features, such as the mood indicators and the calendar, suggests that functionality and clarity are important. Ensuring that each feature is well explained and operates smoothly can lead to a more satisfying user experience. This theme points to the need for clear communication and improved design in specific areas of the app.

Focus Groups

Overview

The transcripts of these focus groups led to identifying 5 key themes, presented in Figure 7: calendar and activity management, workout videos, user interface and accessibility, personalization and profile management, and social features. Participants valued the community calendar for fostering social connections, but engagement in activities was low. Tailored workout videos were appreciated for their adaptability, although more flexibility in goal setting and detailed instructions were desired. Feedback highlighted the intuitive interface but suggested improvements in flexibility, readability, and tracking. Personalization features such as the chatbot were appreciated, although users wanted better clarity in activity plan customization. Social features, including a buddy system, were seen as ways to boost user interaction and engagement. Each of these themes is further detailed in the following subsections.

Figure 7. Key themes of the postassessment focus groups.



XSL•FO
RenderX

Calendar and Activity Management

Participants valued the community calendar for enhancing social connections but reported limited participation in the listed activities. A participant aged 68 years highlighted the following:

> *The agenda is there because of the social connection with others. I really liked it, although I didn't go to an organized activity. You should definitely keep it in there.*

Others suggested that participation might increase with better visibility and advertising of the activities, along with reminders and notifications to prompt more active involvement.

Workout Videos

The tailored workout plans and videos were popular among users. A participant aged 67 years wanted greater control over his workout goals, saying the following:

> *The target of 150 minutes per week is good, but I would like to choose my target minutes myself, for instance, 500 minutes.*

This comment indicated that some users prefer more flexibility in setting personal goals. Another participant aged 72 years commented the following:

> *The algorithm when a workout is too challenging works well; the next video suggestion is indeed tailored to my level. I tested it to see if it would work.*

This adaptability was valued, but users pointed out that the video instructions could sometimes be too quick to follow. Another participant aged 70 years who was already active and doing a lot of cardiovascular exercise mentioned the following—"The app and the exercises made me realize I need to work on my strength as well"—suggesting that the app can broaden users' fitness horizons.

Participants also suggested improvements such as being informed about required materials in advance. A participant aged 69 years noted the following:

> *It would be good to know which materials to use in advance.*

This feedback suggests that clearer preparation instructions could enhance user experience.

User Interface and Accessibility

Various aspects of the app's interface and accessibility were highlighted, noting strengths and areas for improvement. The calendar feature was appreciated for fostering social interactions, although it lacked sufficient engagement tools. Users found the manual diary and progression monitor helpful for tracking activities. Still, they criticized their limited functionality, particularly the inability to edit past entries and the short view range of 7 days. Suggestions included enhancing flexibility and extending tracking capabilities to improve usability.

The *Learning* section received mixed reviews; while the content was engaging, users suggested features such as checkboxes to track articles read, indicating a need for more interaction and relevance to older adults. Although the app's layout was praised for its esthetic and ease of navigation, calls for better readability and customization were prominent.

Feedback on the chatbot was mixed, with some users finding it motivating whereas others saw room for personalization. General usability issues related to smartphone optimization and intuitive navigation were raised but reportedly resolved quickly, demonstrating effective technical support. While the app was generally well received, users desired more robust features and personalization to enhance their experience.

Personalization and Profile Management

Users appreciated features that allowed for personalization, such as the chatbot and manual diary, recognizing their value despite some limitations. A user aged 68 years praised the diary feature, stating that "The diary was great. It is really good to track your activities," but also recommended enhancements, suggesting that "It would be nice if the diary could go back in the past, more than now." In addition, it became apparent that users were unaware that the initial registration questions were intended to personalize their plans as a user stated the following:

> *I didn't know we could adjust the information in the PA plan.*

This misunderstanding also led to them not expecting the system to adapt over time.

Nonetheless, several participants acknowledged the app's added value in promoting PA. A participant aged 66 years noted the following:

> *Getting the push notifications motivated me; it was like a digital motivator in being more active.*

Social Features

The social features such as the community calendar were highlighted as significant opportunities to enhance user engagement. A participant aged 66 years suggested a buddy system to increase interaction:

> *Perhaps a buddy system where I can see my friend and we can motivate each other.*

Suggestions for Improvement

Multimedia Appendix 7 summarizes the suggestions and iterations from the focus groups, offering insights into lessons learned during the development process.

Integration of Quantitative and Qualitative Analysis

The NASSS framework [72] provided a structured approach to integrating the quantitative and qualitative results, revealing barriers to and facilitators of adoption and the interactions influencing scalability and sustainability. It enabled comparisons among user experiences, usability scores (eg, SUS and CSAT), and app interaction data while incorporating qualitative insights from focus group feedback on usability challenges and user perceptions. This enabled us to identify alignment and discrepancies between users' perceived value of the MIA app and their actual app interactions, as well as broader contextual factors influencing sustained use. Figure 8 visualizes this analysis, emphasizing how the NASSS framework bridges the gap between numerical data and user narratives.

XSL•FO
RenderX

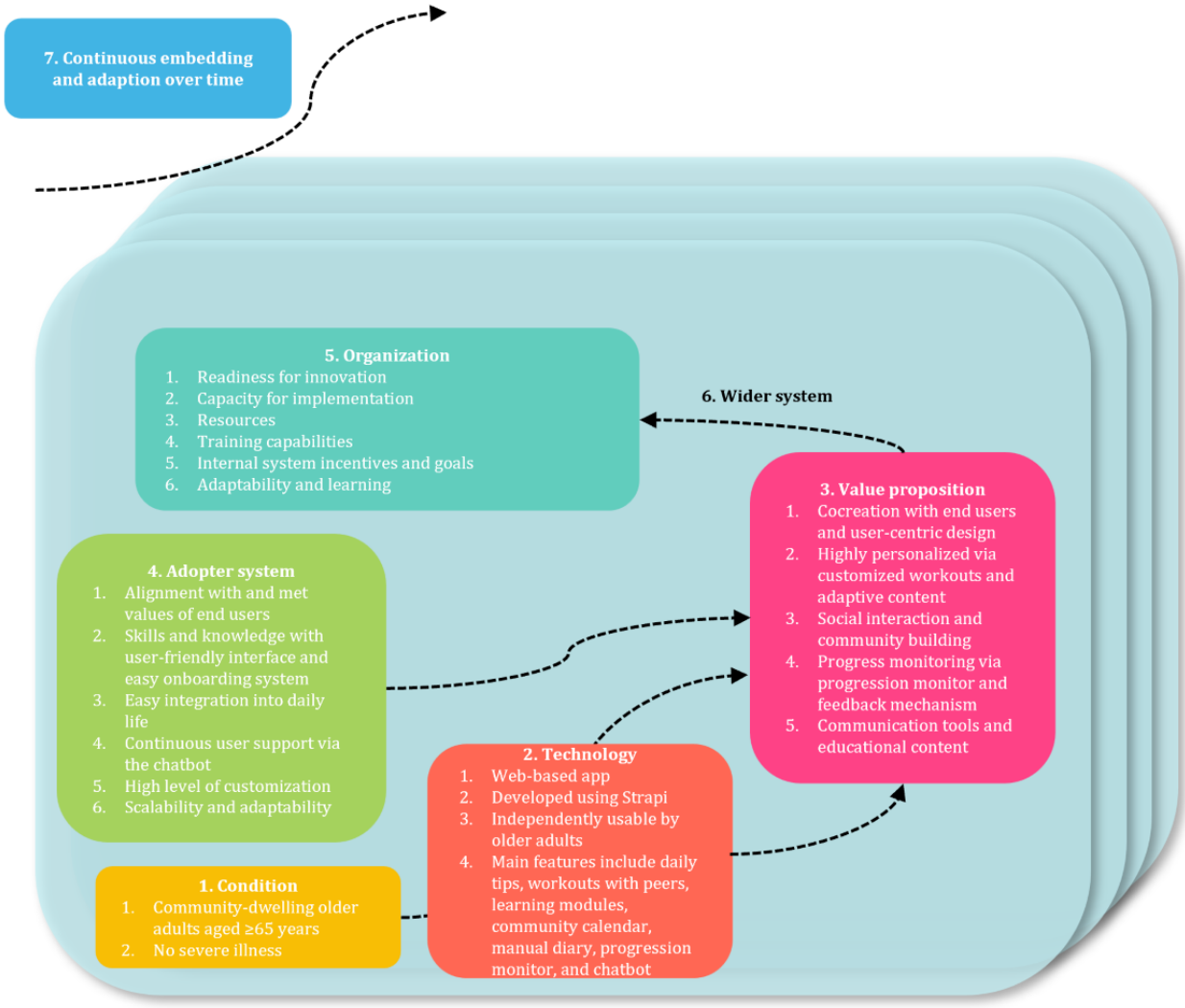
Addressing physical inactivity among older adults is challenging due to varying digital literacy, physical abilities, and health conditions. A key facilitator was the app's customization, allowing users to adjust exercise recommendations, leading to positive SUS and CSAT scores. The app's technical stability and responsive support further encouraged adoption among digitally literate users.

However, barriers such as navigation difficulties and insufficient guidance on features such as profile setup and calendar use limited engagement, with 30% (9/30) of the users classified as passive and 17% (5/30) classified as detractors in the NPS, indicating a risk of user churn. While users appreciated tailored workouts and real-time feedback, some sought greater control over goal setting, suggesting the need for enhanced personalization.

The NASSS analysis identified opportunities to boost engagement, including social features such as a buddy system to align with older adults' social preferences. Integration with health care systems could facilitate broader adoption if endorsed by providers. Alignment with health policies and partnerships may also expand the app's reach.

Despite these strengths, challenges such as technological intimidation, market competition, age-related declines, and data privacy concerns pose risks to long-term adoption. Adherence to ethical standards mitigated some concerns. Sustained engagement depended on the app's adaptability to user feedback, requiring continuous updates and technical support to maintain user interest. Detailed information is available in Multimedia Appendix 8.

Figure 8. Nonadoption, Abandonment, Scale-up, Spread, and Sustainability framework adapted for the feasibility study of the More In Action app.



Discussion

Principal Findings

In recent years, the intersection of aging and technology has gained recognition as a vital factor in enhancing quality of life for older adults [73]. Technology holds substantial potential to improve everyday living for this demographic, yet adoption barriers remain a significant concern as outcomes often rely on user-driven interactions, which can vary significantly [10,34,74,75]. Effective technology design for older users is best achieved through *user involvement and cocreation*, a

XSL•FO
RenderX

strategy that leverages the unique insights of users, who are experts in their own experiences [19,20,61,76,77]. This participatory approach not only results in products that better meet older adults' specific needs, desires, and challenges but also counters age-related stereotypes in gerontechnology [78]. MIA exemplifies this by having been developed through a collaborative cocreation process with older adults. However, to achieve acceptable technologies for older adults, an in-depth understanding of their acceptability, feasibility, and usability is necessary [79]. Against this backdrop, the main aim of this study was to thoroughly explore the acceptability, feasibility, and usability of the MIA app in promoting PA and encouraging sustained, active, and healthy behaviors among its users.

The app received positive feedback, achieving an SUS score of 77.4 and a high CSAT rate of 86.6%. The NPS stood at 33.34, indicating a good level of customer loyalty, with half (15/30, 50%) of the participants categorized as promoters. Feedback on the app was positive for its clarity and engagement. In addition, feasibility and adherence levels were comparable to those observed in other apps evaluated. For example, StandingTall showed acceptable usability, with exercise adherence improving over time [35], whereas Physitrack reported a high study retention rate of 95% and an adherence to prescribed exercises of 84% [40]. Similarly, Fit for All [43] recorded an adherence rate of 82%, supporting the app's potential in facilitating engagement and promoting PA among users. The workout videos on the MIA app received a commendation, although some participants suggested improvements. The learning module and diary functions were praised for their utility, yet some users noted challenges with navigation and desired the ability to edit past diary entries. The progression tracker and chatbot were also seen as valuable, although there was a call for more detailed tracking and enhanced personalization. Analysis showed strong engagement with the workout videos, especially strength-focused sessions, and high satisfaction levels with these workouts. Insights from the think-aloud protocols and focus groups highlighted the need for clearer guidance on profile management, workout customization, and calendar functionality. Participants also expressed a desire for more engaging social features, such as a buddy system. In general, the app was well received for its educational content and tracking features. However, there were recommendations for improving the user interface, enhancing social connectivity, and further personalizing the content.

There are 3 reasons why the MIA app offers more value compared to other alternatives (Table 2). First, the development and design process of the MIA app was based on the design thinking method and user-centered approach, guided by older adults at every stage. Their input was pivotal, ensuring that the app's features were tailored to their specific needs and preferences. Moreover, most existing apps in this domain predominantly cater to the younger segment of the older population, typically those aged 50 to 55 years [10,12,13,80-82]. This demographic focus leaves individuals aged ≥65 years underrepresented. The MIA app addresses this gap by concentrating research and design efforts on the age group of ≥65 years, thereby providing a more inclusive and representative technological solution for older adults.

Second, the MIA app incorporates behavior change theories such as the BCW [29], self-determination theory [33], and self-identification theory [30] to enhance user engagement.

Finally, the MIA app is engineered to dynamically adapt to the diverse capabilities and fluctuating motivations of its users [83,84]. It features an extensive repertoire of 195 distinct exercises encompassing various PA domains such as strength, coordination, balance, flexibility, and endurance, as well as incorporating activities pertinent to daily living. During the feasibility study, sending push notifications with personalized health messages was linked to increased user engagement. In addition, research indicates that users are more likely to interact with the app within 24 hours when push notifications are sent at midday on weekends [85-87]. Personalization is achieved by accounting for the inherently dynamic nature of motivation, which is known to vary not only from day to day but also within a single day [88]. To systematically capture these fluctuations, the app uses the technique of ecological momentary assessment [89]. This approach involves the intensive and recurrent collection of data on an individual's behavior and motivation in a real-time environment [90-92]. Moreover, the app continuously refines its exercise recommendations based on user interactions and feedback. If an exercise is deemed excessively challenging, the algorithm records this, adjusting future workouts to better match the user's current abilities and psychological state. This ensures that subsequent sessions are both enjoyable and attainable, fostering ongoing engagement and facilitating progressive improvement.

To ensure long-term retention and motivation, new videos and activities are regularly added to the app. Future updates to the MIA app will emphasize social features and gamification, including a buddy system for users to connect with exercise partners for mutual support, as well as earning badges, completing challenges, and participating in leaderboards for friendly competition. These elements, along with personalized workouts, progress tracking, content refreshers, and timely reminders, aim to enhance user engagement and sustained adherence.

Strengths and Limitations

The results of this study should be interpreted in light of some limitations. First, the participant profile may not fully represent the broader older adult population. The individuals involved in this feasibility study were notably well educated and digitally literate, with nearly 60% (17/30, 57%) owning smartwatches. This contrasts with broader trends identified in a recent study [93], which found that only 25% of those aged ≥65 years use smartwatches or health apps, a number that declines to 16% among those aged >75 years. In our study, most participants already owned smartphones and were more acquainted with mobile technology. Future studies could explore the perceptions and willingness to adopt this mHealth app among older adults less familiar with such technology.

In addition, the participants who volunteered for this study might have been inherently more motivated and interested in PA than the general older adult population. This could limit the generalizability of the findings as their feedback might not reflect the perspectives of a broader, less technologically adept

XSL•FO
RenderX

audience. Involving representatives of the older people demographic was very challenging because older adults are an extremely heterogeneous group with highly varied characteristics and needs who use, modify, and interact with technologies in rather diverse ways [94]. To mitigate the limitations of this study and enhance its generalizability, several strategies need to be implemented in the future. This involves targeting individuals who are less educated and less familiar with digital technology and those who do not own smart devices. Collaborating with general practitioners can help reach this more varied participant pool.

In addition, focus groups with older adults can sometimes result in conflicting opinions regarding an app's design and features, which proves that a one-size-fits-all approach does not work for this population. Fortunately, the opinions of the older adults were well studied and were considered as much as possible during the iteration process after the feasibility study. To address the issue of conflicting opinions in focus groups, it would be beneficial to implement a more personalized approach in future studies [95,96]. This could include individual interviews that can capture a wider range of perspectives and experiences. This would allow researchers to explore deeper insights and accommodate each participant's unique needs and preferences, moving beyond a one-size-fits-all solution.

To enhance both the reliability and validity of future studies, it is recommended to incorporate randomized controlled designs that include blinding procedures. This approach will help minimize bias by ensuring that assessors do not know which participants have received the intervention, thereby providing a more objective assessment of the outcomes [97,98]. In addition, conducting longitudinal assessments could yield a richer understanding of the intervention's long-term attrition rates, further strengthening the study's findings and their applicability across different settings and populations [99,100].

Finally, unrealistic expectations surrounding an mHealth app could pose a challenge. Users might expect immediate and significant improvements, failing to recognize that behavior change is typically a gradual process [24]. Therefore, it is crucial to underscore that the app is designed to be a supportive aid in the journey toward change, not a quick fix. To address this issue, managing expectations proactively is essential. Effective communication and educational initiatives are key to establishing a realistic understanding of what the app can and cannot do. By clearly outlining the app's functionalities and limitations, users can be better prepared for their experience. This approach helps maintain sustained engagement and maximizes the app's potential as a tool for positive transformation.

Despite these limitations, this study's primary strength lay in its adoption of a mixed methods approach, which facilitated a comprehensive understanding of the app's feasibility. The methodology combined surveys using validated tools, interviews, focus groups, and Microsoft Power BI analytics. This robust approach enabled us to construct a detailed profile of user behavior within the app. Through the systematic analysis of interaction data, we identified the features that garnered the most engagement and pinpointed areas for potential enhancement to improve user experience and retention. The insights derived from this extensive data analysis are fundamental in shaping the future development strategies of the app, ensuring that it aligns more closely with user needs and preferences. Furthermore, by using these diverse research techniques, we were able to gather rich data on participants' experiences concerning the feasibility of the MIA app. This collaborative methodology proved especially beneficial, fostering meaningful dialogues between participants and researchers. These discussions yielded valuable perspectives on potential improvements to the mHealth app.

Future Use and Implications in the Field

The MIA app is currently free and open source. However, to ensure its long-term viability, partnerships with health care insurers and other market players are being explored. A willingness-to-pay analysis [101], detailed in Multimedia Appendix 9, indicated that end users are willing to pay €4.50 to €7 (US $4.66 to US $7.25 at a conversion rate of €1=US $1.04) per month for access to the app. This willingness to pay suggests a perceived value of the app beyond its initial free access, which is further supported by user behavior. Specifically, after 6 months, 43% (13/30) of the participants continued using the app daily without being prompted.

Implementing MIA in real-world settings involves integrating it into health care systems, community programs, and support networks, positioning it as a preventive tool. Partnerships with health care providers and insurers can promote MIA's role while community centers leverage its social features to encourage group activities.

As digital health technologies continue to evolve, the MIA app needs to take several opportunities into account to expand its applicability and feasibility. This offers a range of possibilities for the field.

Tailored user experience remains a very important subject to keep motivating older adults by providing relevant and achievable goals that cater to their individual fitness levels and health conditions [87,102,103]. Enhanced just-in-time interventions enabled by the app's real-time data capabilities through ecological momentary assessment could revolutionize preventive health measures [104,105]. This merges with the potential for integration with wearable technologies, promising a more holistic approach to health monitoring that could improve predictive health interventions for older adults [106-108]. Moreover, integration with wearable devices can enhance the MIA app's functionality by providing real-time, accurate data on PA levels and health metrics. This integration can promise a more holistic view of a user's health and a more precise adjustment of their activity recommendations [50].

As social engagement appeared to be a critical component in maintaining motivation for PA among older adults, the MIA app could include more robust social networking features such as support groups, cooperative challenges, and shared fitness goals [30,109,110].

Another important evolution to consider is the evolution of advanced predictive analytics [111-113]. With advancements in AI and machine learning, the app could incorporate predictive

XSL•FO
RenderX

analytics to forecast potential health risks based on user activity and health data. This feature could alert users and health care providers to potential health issues before they become severe, facilitating timely intervention.

In addition, the MIA app's potential expansion to provide targeted support for informal caregivers, especially those managing care for partners with dementia, offers a pathway to significantly alleviate caregiver burden [114,115]. This feature could become increasingly vital during crisis situations such as pandemics, where the app's adaptability could ensure continuous support for physical and mental health under restrictive conditions [116-119]. Furthermore, by enabling older adults to exercise independently, the app could empower health care professionals by reducing the frequency of in-person checkups, thereby optimizing health care resources [120].

Finally, a broader health integration more closely with health care systems allowing for a smoother exchange of information between the app and health care providers could be a great opportunity [121]. This integration could enable the development of personalized health care plans based on the app's insights, enhancing the overall health care experience by keeping physicians informed and engaged in their patients' lifestyle changes.

Conclusions

This study highlights the potential in merging aging with technology to enhance quality of life and prevention through promotion of PA. Despite some limitations, the app received positive feedback for its usability and customer satisfaction. This underscores the value of involving users in the design process, adhering to a cocreation model that caters to their specific needs and counters age-related stereotypes in technology design. However, the insights gathered suggest a need for broader inclusivity in future studies, targeting less technologically savvy older adults to improve generalizability.

The analysis revealed strong engagement with specific app features (eg, workout videos) and highlighted areas for improvement, such as user interface and social connectivity enhancements. Longitudinal studies and ongoing iterations informed by user feedback will be crucial in refining these features. Managing expectations is also essential as technology adoption among older adults often requires recognizing that behavior change is gradual. Future directions include integrating the app into health care systems to tailor health plans more precisely to individual needs and expanding the app's functionality using predictive analytics to pre-empt health issues. Ultimately, by continuing to evolve and adapt to user feedback and technological advancements, technology such as the MIA app can significantly contribute to promoting sustained, active, and healthy behaviors among older adults, demonstrating the profound impact of well-designed gerontechnology.

Acknowledgments

This work was supported by PXL University of Applied Sciences and Arts (2/DWO/2021/HC/P133) and Flanders Innovation and Entrepreneurship (2/DWO/2022/HC/VL041).

Authors' Contributions

KD developed the initial concept for the project and was responsible for acquiring funding. KD undertook the data analysis and curated the necessary resources while also creating visual representations of the data. KD drafted the original manuscript and led the editing and revision process. JR and KQ both contributed to the investigation process and assisted in editing and reviewing the manuscript. SV was similarly involved in the investigation and provided critical review of the manuscript drafts. JB, AS, and DH provided feedback on the manuscript. BB, aside from assisting in manuscript revision, developed the methodological framework of the study and supervised the research process, ensuring adherence to academic standards and contributing significantly to the original draft. All authors have read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The full informed consent form used to secure participant consent for the study. The form is organized into the following sections: study overview, confidentiality and data privacy, participation and withdrawal details, contact information, consent declaration, and signature and date fields.
[DOCX File , 32 KB - aging_v8i1e63348_app1.docx]

Multimedia Appendix 2
Inclusion and exclusion criteria for trial participation.
[DOCX File , 28 KB - aging_v8i1e63348_app2.docx]

Multimedia Appendix 3

XSL•FO
RenderX

Think-aloud protocol.
[DOCX File , 30 KB - aging_v8i1e63348_app3.docx]

Multimedia Appendix 4
Interview guide for the focus groups.
[DOCX File , 32 KB - aging_v8i1e63348_app4.docx]

Multimedia Appendix 5
Complete acceptability results per category.
[DOCX File , 361 KB - aging_v8i1e63348_app5.docx]

Multimedia Appendix 6
Visual representations of the Microsoft Power BI (Microsoft Corp) analytics.
[DOCX File , 672 KB - aging_v8i1e63348_app6.docx]

Multimedia Appendix 7
Areas of improvement.
[DOCX File , 52 KB - aging_v8i1e63348_app7.docx]

Multimedia Appendix 8
Strengths, weaknesses, opportunities, and threats matrix of the More In Action app according to the Nonadoption, Abandonment, Scale-up, Spread, and Sustainability components.
[DOCX File , 31 KB - aging_v8i1e63348_app8.docx]

Multimedia Appendix 9
Willingness-to-pay analysis.
[DOCX File , 33 KB - aging_v8i1e63348_app9.docx]

References

1. Ageing. World Health Organization. URL: https://www.who.int/health-topics/ageing#tab=tab_1 [accessed 2023-12-01]
2. Bangsbo J, Blackwell J, Boraxbekk CJ, Caserotti P, Dela F, Evans AB, et al. Copenhagen Consensus statement 2019: physical activity and ageing. Br J Sports Med 2019 Jul 21;53(14):856-858 [FREE Full text] [doi: 10.1136/bjsports-2018-100451] [Medline: 30792257]
3. WHO guidelines on physical activity and sedentary behaviour. World Health Organization. 2020 Nov 25. URL: https://www.who.int/publications/i/item/9789240015128 [accessed 2023-12-01]
4. Ekelund U, Steene-Johannessen J, Brown WJ, Fagerland MW, Owen N, Powell KE, et al. Does physical activity attenuate, or even eliminate, the detrimental association of sitting time with mortality? A harmonised meta-analysis of data from more than 1 million men and women. The Lancet 2016 Sep;388(10051):1302-1310. [doi: 10.1016/s0140-6736(16)30370-1]
5. Guthold R, Stevens GA, Riley LM, Bull FC. Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. The Lancet Global Health 2018 Oct;6(10):e1077-e1086. [doi: 10.1016/s2214-109x(18)30357-7]
6. Harvey JA, Chastin SF, Skelton DA. How sedentary are older people? A systematic review of the amount of sedentary behavior. J Aging Phys Act 2015 Jul;23(3):471-487. [doi: 10.1123/japa.2014-0164] [Medline: 25387160]
7. Sherrington C, Fairhall NJ, Wallbank GK, Tiedemann A, Michaleff ZA, Howard K, et al. Exercise for preventing falls in older people living in the community. Cochrane Database Syst Rev 2019 Jan 31;1(1):CD012424 [FREE Full text] [doi: 10.1002/14651858.CD012424.pub2] [Medline: 30703272]
8. Bonnechère B, Kossi O, Mapinduzi J, Panda J, Rintala A, Guidetti S, et al. Mobile health solutions: an opportunity for rehabilitation in low- and middle income countries? Front Public Health 2022 Jan 24;10:1072322 [FREE Full text] [doi: 10.3389/fpubh.2022.1072322] [Medline: 36761328]
9. Nussbaum R, Kelly C, Quinby E, Mac A, Parmanto B, Dicianno BE. Systematic review of mobile health applications in rehabilitation. Arch Phys Med Rehabil 2019 Jan;100(1):115-127. [doi: 10.1016/j.apmr.2018.07.439] [Medline: 30171827]
10. Kampmeijer R, Pavlova M, Tambor M, Golinowska S, Groot W. The use of e-health and m-health tools in health promotion and primary prevention among older adults: a systematic literature review. BMC Health Serv Res 2016 Sep 05;16 Suppl 5(Suppl 5):290 [FREE Full text] [doi: 10.1186/s12913-016-1522-3] [Medline: 27608677]
11. Hsieh KL, Chen L, Sosnoff JJ. Mobile technology for falls prevention in older adults. J Gerontol A Biol Sci Med Sci 2023 May 11;78(5):861-868 [FREE Full text] [doi: 10.1093/gerona/glac116] [Medline: 35640254]

XSL•FO
RenderX

12. Yerrakalva D, Yerrakalva D, Hajna S, Griffin S. Effects of mobile health app interventions on sedentary time, physical activity, and fitness in older adults: systematic review and meta-analysis. J Med Internet Res 2019 Nov 28;21(11):e14343 [FREE Full text] [doi: 10.2196/14343] [Medline: 31778121]
13. Sohaib Aslam A, van Luenen S, Aslam S, van Bodegom D, Chavannes NH. A systematic review on the use of mHealth to increase physical activity in older people. Clin eHealth 2020;3:31-39. [doi: 10.1016/j.CEH.2020.04.002]
14. Mönninghoff A, Kramer JN, Hess AJ, Ismailova K, Teepe GW, Tudor Car L, et al. Long-term effectiveness of mHealth physical activity interventions: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2021 Apr 30;23(4):e26699 [FREE Full text] [doi: 10.2196/26699] [Medline: 33811021]
15. Free C, Phillips G, Felix L, Galli L, Patel V, Edwards P. The effectiveness of M-health technologies for improving health and health services: a systematic review protocol. BMC Res Notes 2010 Oct 06;3(1):250 [FREE Full text] [doi: 10.1186/1756-0500-3-250] [Medline: 20925916]
16. Meyerowitz-Katz G, Ravi S, Arnolda L, Feng X, Maberly G, Astell-Burt T. Rates of attrition and dropout in app-based interventions for chronic disease: systematic review and meta-analysis. J Med Internet Res 2020 Sep 29;22(9):e20283 [FREE Full text] [doi: 10.2196/20283] [Medline: 32990635]
17. Edney S, Ryan JC, Olds T, Monroe C, Fraysse F, Vandelanotte C, et al. User engagement and attrition in an app-based physical activity intervention: secondary analysis of a randomized controlled trial. J Med Internet Res 2019 Nov 27;21(11):e14645 [FREE Full text] [doi: 10.2196/14645] [Medline: 31774402]
18. Madeira RN, Germano H, Macedo P, Correia N. Personalising the user experience of a mobile health application towards patient engagement. Procedia Comput Sci 2018;141:428-433. [doi: 10.1016/j.procs.2018.10.173]
19. Daniels K, Lemmens R, Knippenberg E, Marinus N, Vonck S, Baerts J, et al. Promoting physical activity and a healthy active lifestyle in community-dwelling older adults: a design thinking approach for the development of a mobile health application. Front Public Health 2023 Nov 29;11:1280941 [FREE Full text] [doi: 10.3389/fpubh.2023.1280941] [Medline: 38106904]
20. Janols R, Sandlund M, Lindgren H, Pettersson B. Older adults as designers of behavior change strategies to increase physical activity-report of a participatory design process. Front Public Health 2022 Dec 21;10:988470 [FREE Full text] [doi: 10.3389/fpubh.2022.988470] [Medline: 36620266]
21. Nyman SR, Victor CR. Older people's participation in and engagement with falls prevention interventions in community settings: an augment to the Cochrane systematic review. Age Ageing 2012 Jan;41(1):16-23. [doi: 10.1093/ageing/afr103] [Medline: 21875865]
22. Lindsay S, Jackson D, Schofield G, Olivier P. Engaging older people using participatory design. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2012 Presented at: CHI '12; May 5-10, 2012; Austin, TX. [doi: 10.1145/2207676.2208570]
23. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6(1):42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
24. Mehra S, Visser B, Dadema T, van den Helder J, Engelbert RH, Weijs PJ, et al. Translating behavior change principles into a blended exercise intervention for older adults: design study. JMIR Res Protoc 2018 May 02;7(5):e117 [FREE Full text] [doi: 10.2196/resprot.9244] [Medline: 29720358]
25. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457 [FREE Full text] [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
26. Feeley N, Cossette S, Côté J, Héon M, Stremler R, Martorella G, et al. The importance of piloting an RCT intervention. Can J Nurs Res 2009 Jun;41(2):85-99. [Medline: 19650515]
27. Beyan N. The music methodology for usability measurement. In: Proceedings of the Posters and Short Talks of the 1992 SIGCHI Conference on Human Factors in Computing Systems. 1992 Presented at: CHI '92; May 3-7, 1992; Monterey, CA. [doi: 10.1145/1125021.1125114]
28. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
29. Michie S, Atkins L, West R. The Behaviour Change Wheel: A Guide to Designing Interventions. Surrey, UK: Silverback Publishing; 2014.
30. Boen F, Seghers J, van Uffelen J. Motiveren tot verandering. In: Cahn W, Deenik J, Vermeulen J, editors. Leefstijlpsychiatrie. Houten, The Netherlands: Bohn Stafleu van Loghum; 2022.
31. Teixeira PJ, Carraça EV, Markland D, Silva MN, Ryan RM. Exercise, physical activity, and self-determination theory: a systematic review. Int J Behav Nutr Phys Act 2012 Jun 22;9(1):78 [FREE Full text] [doi: 10.1186/1479-5868-9-78] [Medline: 22726453]
32. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug 20;46(1):81-95 [FREE Full text] [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]

XSL•FO
RenderX

33. Deci EL, Ryan RM. Motivation, personality, and development within embedded social contexts: an overview of self-determination theory. In: Ryan RM, editor. The Oxford Handbook of Human Motivation. Oxford, UK: Oxford University Press; 2012.
34. Aranha M, James K, Deasy C, Heavin C. Exploring the barriers and facilitators which influence mHealth adoption among older adults: a literature review. Gerontechnology 2021 Jan 01;20(2):1-16. [doi: 10.4017/gt.2021.20.2.424.06]
35. Delbaere K, Valenzuela T, Lord SR, Clemson L, Zijlstra GA, Close JC, et al. E-health StandingTall balance exercise for fall prevention in older people: results of a two year randomised controlled trial. BMJ 2021 Apr 06;373:n740 [FREE Full text] [doi: 10.1136/bmj.n740] [Medline: 33824131]
36. Muellmann S, Bragina I, Voelcker-Rehage C, Rost E, Lippke S, Meyer J, et al. Development and evaluation of two web-based interventions for the promotion of physical activity in older adults: study protocol for a community-based controlled intervention trial. BMC Public Health 2017 May 25;17(1):512 [FREE Full text] [doi: 10.1186/s12889-017-4446-x] [Medline: 28545506]
37. Brew-Sam N, Chib A, Torres AY, Ng JX, Wong YT, Sze-G Y. An integrated mHealth campaign to reduce the risk of falling for older adults. J Appl Gerontol 2022 May 27;41(5):1336-1347. [doi: 10.1177/07334648211062877] [Medline: 35085043]
38. Thompson C, Porter Starr KN, Kemp EC, Chan J, Jackson E, Phun J. Feasibility of virtually delivering functional fitness assessments and a fitness training program in community-dwelling older adults. Int J Environ Res Public Health 2023 May 30;20(11):5996 [FREE Full text] [doi: 10.3390/ijerph20115996] [Medline: 37297600]
39. Jungreitmayr S, Kranzinger C, Venek V, Ring-Dimitriou S. Effects of an app-based physical exercise program on selected parameters of physical fitness of females in retirement: a randomized controlled trial. Front Physiol 2022 Mar 4;13:821773 [FREE Full text] [doi: 10.3389/fphys.2022.821773] [Medline: 35317213]
40. Daly RM, Gianoudis J, Hall T, Mundell NL, Maddison R. Feasibility, usability, and enjoyment of a home-based exercise program delivered via an exercise app for musculoskeletal health in community-dwelling older adults: short-term prospective pilot study. JMIR Mhealth Uhealth 2021 Jan 13;9(1):e21094 [FREE Full text] [doi: 10.2196/21094] [Medline: 33439147]
41. van Dyck D, Herman K, Poppe L, Crombez G, de Bourdeaudhuij I, Gheysen F. Results of MyPlan 2.0 on physical activity in older Belgian adults: randomized controlled trial. J Med Internet Res 2019 Oct 07;21(10):e13219 [FREE Full text] [doi: 10.2196/13219] [Medline: 31593541]
42. Shake MC, Crandall KJ, Mathews RP, Falls DG, Dispennette AK. Efficacy of Bingocize: a game-centered mobile application to improve physical and cognitive performance in older adults. Games Health J 2018 Aug;7(4):253-261. [doi: 10.1089/g4h.2017.0139] [Medline: 30089018]
43. Konstantinidis EI, Billis AS, Mouzakidis CA, Zilidou VI, Antoniou PE, Bamidis PD. Design, implementation, and wide pilot deployment of FitForAll: an easy to use exergaming platform improving physical fitness and life quality of senior citizens. IEEE J Biomed Health Inform 2016 Jan;20(1):189-200. [doi: 10.1109/jbhi.2014.2378814]
44. van Het Reve E, Silveira P, Daniel F, Casati F, de Bruin ED. Tablet-based strength-balance training to motivate and improve adherence to exercise in independently living older people: part 2 of a phase II preclinical exploratory trial. J Med Internet Res 2014 Jun 25;16(6):e159 [FREE Full text] [doi: 10.2196/jmir.3055] [Medline: 24966165]
45. Bao T, Carender WJ, Kinnaird C, Barone VJ, Peethambaran G, Whitney SL, et al. Effects of long-term balance training with vibrotactile sensory augmentation among community-dwelling healthy older adults: a randomized preliminary study. J Neuroeng Rehabil 2018 Jan 18;15(1):5 [FREE Full text] [doi: 10.1186/s12984-017-0339-6] [Medline: 29347946]
46. Albergoni A, Hettinga FJ, Stut W, Sartor F. Factors influencing walking and exercise adherence in healthy older adults using monitoring and interfacing technology: preliminary evidence. Int J Environ Res Public Health 2020 Aug 24;17(17):6142 [FREE Full text] [doi: 10.3390/ijerph17176142] [Medline: 32846988]
47. Nikitina S, Didino D, Baez M, Casati F. Feasibility of virtual tablet-based group exercise among older adults in Siberia: findings from two pilot trials. JMIR Mhealth Uhealth 2018 Feb 27;6(2):e40. [doi: 10.2196/mhealth.7531] [Medline: 29487045]
48. Johnson N, Bradley A, Klawitter L, Johnson J, Johnson L, Tomkinson GR, et al. The impact of a telehealth intervention on activity profiles in older adults during the COVID-19 pandemic: a pilot study. Geriatrics (Basel) 2021 Jun 30;6(3):68 [FREE Full text] [doi: 10.3390/geriatrics6030068] [Medline: 34209416]
49. Papi E, Chiou SY, McGregor AH. Feasibility and acceptability study on the use of a smartphone application to facilitate balance training in the ageing population. BMJ Open 2020 Dec 02;10(12):e039054 [FREE Full text] [doi: 10.1136/bmjopen-2020-039054] [Medline: 33268409]
50. Taraldsen K, Mikolaizak AS, Maier AB, Mellone S, Boulton E, Aminian K, et al. Digital technology to deliver a lifestyle-integrated exercise intervention in young seniors-the PreventIT feasibility randomized controlled trial. Front Digit Health 2020 Jul 31;2:10 [FREE Full text] [doi: 10.3389/fdgth.2020.00010] [Medline: 34713023]
51. Lugade V, Torbitt M, O'Brien SR, Silsupadol P. Smartphone- and paper-based delivery of balance intervention for older adults are equally effective, enjoyable, and of high fidelity: a randomized controlled trial. Sensors (Basel) 2023 Aug 27;23(17):7451 [FREE Full text] [doi: 10.3390/s23177451] [Medline: 37687907]

XSL•FO
RenderX

52. Bosco A, McGarrigle L, Skelton DA, Laventure RM, Townley B, Todd C. Make movement your mission: evaluation of an online digital health initiative to increase physical activity in older people during the COVID-19 pandemic. Digit Health 2022;8:20552076221084468 [FREE Full text] [doi: 10.1177/20552076221084468] [Medline: 35295764]
53. Recio-Rodríguez JI, Gonzalez-Sanchez S, Tamayo-Morales O, Gómez-Marcos MA, Garcia-Ortiz L, Niño-Martín V, et al. Changes in lifestyles, cognitive impairment, quality of life and activity day living after combined use of smartphone and smartband technology: a randomized clinical trial (EVIDENT-Age study). BMC Geriatr 2022 Oct 06;22(1):782 [FREE Full text] [doi: 10.1186/s12877-022-03487-5] [Medline: 36203135]
54. Lee PH, Macfarlane DJ, Lam TH, Stewart SM. Validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF): a systematic review. Int J Behav Nutr Phys Act 2011 Oct 21;8(1):115 [FREE Full text] [doi: 10.1186/1479-5868-8-115] [Medline: 22018588]
55. Kendzierski D, DeCarlo KJ. Physical activity enjoyment scale: two validation studies. J Sport Exerc Psychol 1991;13(1):50-64. [doi: 10.1123/jsep.13.1.50]
56. Scherrenberg M, Falter M, Kaihara T, Xu L, van Leunen M, Kemps H, et al. Development and internal validation of the digital health readiness questionnaire: prospective single-center survey study. J Med Internet Res 2023 Mar 10;25:e41615 [FREE Full text] [doi: 10.2196/41615] [Medline: 36897627]
57. van Holle V, de Bourdeaudhuij I, Deforche B, van Cauwenberg J, van Dyck D. Assessment of physical activity in older Belgian adults: validity and reliability of an adapted interview version of the long International Physical Activity Questionnaire (IPAQ-L). BMC Public Health 2015 Apr 28;15(1):433 [FREE Full text] [doi: 10.1186/s12889-015-1785-3] [Medline: 25928561]
58. Meh K, Jurak G, Sorić M, Rocha P, Sember V. Validity and reliability of IPAQ-SF and GPAQ for assessing sedentary behaviour in adults in the European Union: a systematic review and meta-analysis. Int J Environ Res Public Health 2021 Apr 26;18(9):4602. [doi: 10.3390/ijerph18094602] [Medline: 33926123]
59. Chung PK, Leung KM. Psychometric properties of eight-item physical activity enjoyment scale in a Chinese population. J Aging Phys Act 2018 Oct 12;27(1):1-6. [doi: 10.1123/japa.2017-0212] [Medline: 29722637]
60. Lewis JR. The system usability scale: past, present, and future. Int J Hum Comput Interact 2018 Mar 30;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
61. Fisher NI, Kordupleski RE. Good and bad market research: a critical review of Net Promoter Score. Appl Stochastic Models Bus Ind 2018 Nov 22;35(1):138-151. [doi: 10.1002/asmb.2417]
62. Baquero A. Net promoter score (NPS) and customer satisfaction: relationship and efficient management. Sustainability 2022 Feb 10;14(4):2011. [doi: 10.3390/su14042011]
63. Ensink CJ, Keijsers NL, Groen BE. Translation and validation of the System Usability Scale to a Dutch version: D-SUS. Disabil Rehabil 2024 Jan 27;46(2):395-400. [doi: 10.1080/09638288.2022.2160837] [Medline: 36573399]
64. Town P, Thabtah F. Data analytics tools: a user perspective. J Inf Knowl Manag 2019 Mar 22;18(01):1950002. [doi: 10.1142/s0219649219500023]
65. Alhadreti O, Mayhew P. Rethinking thinking aloud: a comparison of three think-aloud protocols. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI '18; April 21-26, 2018; Montreal, QC. [doi: 10.1145/3173574.3173618]
66. Jääskeläinen R. Think-aloud protocol. In: Handbook of Translation Studies. Amsterdam, The Netherlands: John Benjamins Publishing Company; 2010.
67. Eccles DW, Arsal G. The think aloud method: what is it and how do I use it? Qual Res Sport Exerc Health 2017 May 26;9(4):514-531. [doi: 10.1080/2159676x.2017.1331501]
68. McDonald S, Edwards HM, Zhao T. Exploring think-alouds in usability testing: an international survey. IEEE Trans Prof Commun 2012 Mar;55(1):2-19. [doi: 10.1109/tpc.2011.2182569]
69. Ericsson KA. Valid and non-reactive verbalization of thoughts during performance of tasks: towards a solution to the central problems of introspection as a source of scientific data. J Conscious Stud 2003;10(9-10):1-18 [FREE Full text]
70. Dhakal K. NVivo. J Med Libr Assoc 2022 Apr 01;110(2):270-272 [FREE Full text] [doi: 10.5195/jmla.2022.1271] [Medline: 35440911]
71. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
72. Greenhalgh T, Abimbola S. The NASSS framework - a synthesis of multiple theories of technology implementation. Stud Health Technol Inf 2019 Jul 30;263:193-204. [doi: 10.3233/SHTI190123] [Medline: 31411163]
73. Pruchno R. Technology and aging: an evolving partnership. Gerontologist 2019 Jan 09;59(1):1-5. [doi: 10.1093/geront/gny153] [Medline: 30629258]
74. Lee CC, Czaja SJ, Moxley JH, Sharit J, Boot WR, Charness N, et al. Attitudes toward computers across adulthood from 1994 to 2013. Gerontologist 2019 Jan 09;59(1):22-33 [FREE Full text] [doi: 10.1093/geront/gny081] [Medline: 29982458]
75. Nebeker C, Zlatar ZZ. Learning from older adults to promote independent physical activity using mobile health (mHealth). Front Public Health 2021 Aug 12;9:703910 [FREE Full text] [doi: 10.3389/fpubh.2021.703910] [Medline: 34476229]

XSL•FO
RenderX

76. Matthew-Maich N, Harris L, Ploeg J, Markle-Reid M, Valaitis R, Ibrahim S, et al. Designing, implementing, and evaluating mobile health technologies for managing chronic conditions in older adults: a scoping review. JMIR Mhealth Uhealth 2016 Jun 09;4(2):e29 [FREE Full text] [doi: 10.2196/mhealth.5127] [Medline: 27282195]
77. Fischer B, Peine A, Östlund B. The importance of user involvement: a systematic review of involving older users in technology design. Gerontologist 2020 Sep 15;60(7):e513-e523 [FREE Full text] [doi: 10.1093/geront/gnz163] [Medline: 31773145]
78. Frennert S, Östlund B. Review: seven matters of concern of social robots and older people. Int J Soc Robot 2014 Jan 29;6(2):299-310. [doi: 10.1007/s12369-013-0225-8]
79. Degroote L, van Dyck D, de Bourdeaudhuij I, de Paepe A, Crombez G. Acceptability and feasibility of the mHealth intervention 'MyDayPlan' to increase physical activity in a general adult population. BMC Public Health 2020 Jun 29;20(1):1032 [FREE Full text] [doi: 10.1186/s12889-020-09148-9] [Medline: 32600352]
80. McGarrigle L, Todd C. Promotion of physical activity in older people using mHealth and eHealth technologies: rapid review of reviews. J Med Internet Res 2020 Dec 29;22(12):e22201 [FREE Full text] [doi: 10.2196/22201] [Medline: 33372894]
81. Muellmann S, Forberger S, Möllers T, Bröring E, Zeeb H, Pischke CR. EHealth interventions for the promotion of physical activity in older adults: a systematic review. In: Proceedings of the 9th European Public Health Conference. 2016 Presented at: EPH 2016; November 9-12, 2016; Vienna, Austria. [doi: 10.1093/eurpub/ckw165.023]
82. Vasquez BA, Betriana F, Nemenzo E, Inabangan AK, Tanioka R, Garcia L, et al. Effects of healthcare technologies on the promotion of physical activities in older persons: a systematic review. Inform Health Soc Care 2023 Apr 03;48(2):196-210. [doi: 10.1080/17538157.2022.2086874] [Medline: 35699246]
83. Maes I, Mertens L, Poppe L, Crombez G, Vetrovsky T, van Dyck D. The variability of emotions, physical complaints, intention, and self-efficacy: an ecological momentary assessment study in older adults. PeerJ 2022;10:e13234 [FREE Full text] [doi: 10.7717/peerj.13234] [Medline: 35611175]
84. Maes I, Mertens L, Poppe L, Vetrovsky T, Crombez G, de Backere F, et al. Within-person associations of accelerometer-assessed physical activity with time-varying determinants in older adults: time-based ecological momentary assessment study. JMIR Aging 2023 Nov 23;6:e44425 [FREE Full text] [doi: 10.2196/44425] [Medline: 37995131]
85. Hernández-Reyes A, Molina-Recio G, Molina-Luque R, Romero-Saldaña M, Cámara-Martos F, Moreno-Rojas R. Effectiveness of PUSH notifications from a mobile app for improving the body composition of overweight or obese women: a protocol of a three-armed randomized controlled trial. BMC Med Inform Decis Mak 2020 Feb 24;20(1):40 [FREE Full text] [doi: 10.1186/s12911-020-1058-7] [Medline: 32093701]
86. Bidargaddi N, Almirall D, Murphy S, Nahum-Shani I, Kovalcik M, Pituch T, et al. To prompt or not to prompt? A microrandomized trial of time-varying push notifications to increase proximal engagement with a mobile health app. JMIR Mhealth Uhealth 2018 Nov 29;6(11):e10123 [FREE Full text] [doi: 10.2196/10123] [Medline: 30497999]
87. Mair JL, Hayes LD, Campbell AK, Buchan DS, Easton C, Sculthorpe N. A personalized smartphone-delivered just-in-time adaptive intervention (JitaBug) to increase physical activity in older adults: mixed methods feasibility study. JMIR Form Res 2022 Apr 07;6(4):e34662 [FREE Full text] [doi: 10.2196/34662] [Medline: 35389348]
88. Rothman AJ, Sheeran P. The operating conditions framework: integrating mechanisms and moderators in health behavior interventions. Health Psychol 2021 Dec;40(12):845-857. [doi: 10.1037/hea0001026] [Medline: 32914997]
89. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008 Apr 01;4(1):1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
90. Pickering TA, Huh J, Intille S, Liao Y, Pentz MA, Dunton GF. Physical activity and variation in momentary behavioral cognitions: an ecological momentary assessment study. J Phys Act Health 2016 Mar;13(3):344-351. [doi: 10.1123/jpah.2014-0547] [Medline: 26284314]
91. Dunton GF. Ecological momentary assessment in physical activity research. Exerc Sport Sci Rev 2017 Jan;45(1):48-54 [FREE Full text] [doi: 10.1249/JES.0000000000000092] [Medline: 27741022]
92. Firth J, Torous J, Yung AR. Ecological momentary assessment and beyond: the rising interest in e-mental health research. J Psychiatr Res 2016 Sep;80:3-4 [FREE Full text] [doi: 10.1016/j.jpsychires.2016.05.002] [Medline: 27236099]
93. de Marez L, Sevenhant R, Denecker F, Georges A, Wuyts G, Schuurman D. imec.digimeter 2023 : digitale trends in Vlaanderen. Imec. URL: https://biblio.ugent.be/publication/01HJ69C9YQEE4C4KXKBC7AYS2J [accessed 2023-12-01]
94. Grates MG, Heming AC, Vukoman M, Schabsky P, Sorgalla J. New perspectives on user participation in technology design processes: an interdisciplinary approach. Gerontologist 2019 Jan 09;59(1):45-57. [doi: 10.1093/geront/gny112] [Medline: 30260407]
95. Agan J, Koch LC, Rumrill PDJ. The use of focus groups in rehabilitation research. Work 2008 Jan 01;31(2):259-269. [doi: 10.3233/wor-2008-00731]
96. O.Nyumba T, Wilson K, Derrick CJ, Mukherjee N. The use of focus group discussion methodology: insights from two decades of application in conservation. Methods Ecol Evol 2018 Jan 11;9(1):20-32. [doi: 10.1111/2041-210x.12860]
97. El-Kotob R, Giangregorio LM. Pilot and feasibility studies in exercise, physical activity, or rehabilitation research. Pilot Feasibility Stud 2018 Aug 14;4(1):137 [FREE Full text] [doi: 10.1186/s40814-018-0326-0] [Medline: 30123527]

XSL•FO
RenderX

98. Shanyinde M, Pickering RM, Weatherall M. Questions asked and answered in pilot and feasibility randomized controlled trials. BMC Med Res Methodol 2011 Aug 16;11(1):117 [FREE Full text] [doi: 10.1186/1471-2288-11-117] [Medline: 21846349]
99. Bonnechère B, Timmermans A, Michiels S. Current technology developments can improve the quality of research and level of evidence for rehabilitation interventions: a narrative review. Sensors (Basel) 2023 Jan 12;23(2):875 [FREE Full text] [doi: 10.3390/s23020875] [Medline: 36679672]
100. Beukenhorst AL, Howells K, Cook L, McBeth J, O'Neill TW, Parkes MJ, et al. Engagement and participant experiences with consumer smartwatches for health research: longitudinal, observational feasibility study. JMIR Mhealth Uhealth 2020 Jan 29;8(1):e14368 [FREE Full text] [doi: 10.2196/14368] [Medline: 32012078]
101. Ceylana HH, Koseb B, Aydin M. Value based pricing: a research on service sector using Van Westendorp Price Sensitivity Scale. Procedia Soc Behav Sci 2014 Aug;148:1-6. [doi: 10.1016/j.sbspro.2014.07.013]
102. Cao W, Milks MW, Liu X, Gregory ME, Addison D, Zhang P, et al. mHealth interventions for self-management of hypertension: framework and systematic review on engagement, interactivity, and tailoring. JMIR Mhealth Uhealth 2022 Mar 02;10(3):e29415 [FREE Full text] [doi: 10.2196/29415] [Medline: 35234655]
103. Nimmanterdwong Z, Boonviriya S, Tangkijvanich P. Human-centered design of mobile health apps for older adults: systematic review and narrative synthesis. JMIR Mhealth Uhealth 2022 Jan 14;10(1):e29512 [FREE Full text] [doi: 10.2196/29512] [Medline: 35029535]
104. Hardeman W, Houghton J, Lane K, Jones A, Naughton F. A systematic review of just-in-time adaptive interventions (JITAIs) to promote physical activity. Int J Behav Nutr Phys Act 2019 Apr 03;16(1):31 [FREE Full text] [doi: 10.1186/s12966-019-0792-7] [Medline: 30943983]
105. Schembre SM, Liao Y, Robertson MC, Dunton GF, Kerr J, Haffey ME, et al. Just-in-time feedback in diet and physical activity interventions: systematic review and practical design framework. J Med Internet Res 2018 Mar 22;20(3):e106 [FREE Full text] [doi: 10.2196/jmir.8701] [Medline: 29567638]
106. Wilmink G, Dupey K, Alkire S, Grote J, Zobel G, Fillit HM, et al. Artificial intelligence-powered digital health platform and wearable devices improve outcomes for older adults in assisted living communities: pilot intervention study. JMIR Aging 2020 Sep 10;3(2):e19554 [FREE Full text] [doi: 10.2196/19554] [Medline: 32723711]
107. Burnham JP, Lu C, Yaeger LH, Bailey TC, Kollef MH. Using wearable technology to predict health outcomes: a literature review. J Am Med Inform Assoc 2018 Sep 01;25(9):1221-1227 [FREE Full text] [doi: 10.1093/jamia/ocy082] [Medline: 29982520]
108. Baig MM, Afifi S, GholamHosseini H, Mirza F. A systematic review of wearable sensors and IoT-based monitoring applications for older adults - a focus on ageing population and independent living. J Med Syst 2019 Jun 15;43(8):233. [doi: 10.1007/s10916-019-1365-7] [Medline: 31203472]
109. Pelssers J, Delecluse C, Opdenacker J, Kennis E, van Roie E, Boen F. "Every step counts!": effects of a structured walking intervention in a community-based senior organization. J Aging Phys Act 2013 Apr;21(2):167-185. [doi: 10.1123/japa.21.2.167] [Medline: 22898451]
110. Boulton ER, Horne M, Todd C. Involving older adults in developing physical activity interventions to promote engagement: a literature review. Popul Ageing 2019 Jun 24;13(3):325-345. [doi: 10.1007/s12062-019-09247-5]
111. Bonnechère B. Unlocking the black box? A comprehensive exploration of large language models in rehabilitation. Am J Phys Med Rehabil 2024 Jan 12;103(6):532-537. [doi: 10.1097/phm.0000000000002440]
112. Golas SB, Nikolova-Simons M, Palacholla R, Op den Buijs J, Garberg G, Orenstein A, et al. Predictive analytics and tailored interventions improve clinical outcomes in older adults: a randomized controlled trial. NPJ Digit Med 2021 Jun 10;4(1):97 [FREE Full text] [doi: 10.1038/s41746-021-00463-y] [Medline: 34112921]
113. Nikolova-Simons M, Golas SB, den Buijs JO, Palacholla RS, Garberg G, Orenstein A, et al. A randomized trial examining the effect of predictive analytics and tailored interventions on the cost of care. NPJ Digit Med 2021 Jun 03;4(1):92 [FREE Full text] [doi: 10.1038/s41746-021-00449-w] [Medline: 34083743]
114. Chiao CY, Wu HS, Hsiao CY. Caregiver burden for informal caregivers of patients with dementia: a systematic review. Int Nurs Rev 2015 Sep 08;62(3):340-350. [doi: 10.1111/inr.12194] [Medline: 26058542]
115. Alltag S, Conrad I, Riedel-Heller SG. [Caregiver burden among older informal caregivers of patients with dementia and its influence on quality of life: a systematic literature review]. Z Gerontol Geriatr 2019 Aug;52(5):477-486. [doi: 10.1007/s00391-018-1424-6] [Medline: 29992357]
116. Yang Y, Koenigstorfer J. Determinants of physical activity maintenance during the Covid-19 pandemic: a focus on fitness apps. Transl Behav Med 2020 Oct 08;10(4):835-842 [FREE Full text] [doi: 10.1093/tbm/ibaa086] [Medline: 32926160]
117. Symons M, Meira Cunha C, Poels K, Vandebosch H, Dens N, Alida Cutello C. Physical activity during the first lockdown of the COVID-19 pandemic: investigating the reliance on digital technologies, perceived benefits, barriers and the impact of affect. Int J Environ Res Public Health 2021 May 22;18(11):5555 [FREE Full text] [doi: 10.3390/ijerph18115555] [Medline: 34067390]
118. Petersen JM, Kemps E, Lewis LK, Prichard I. Promoting physical activity during the COVID-19 lockdown in Australia: the roles of psychological predictors and commercial physical activity apps. Psychol Sport Exerc 2021 Sep;56:102002 [FREE Full text] [doi: 10.1016/j.psychsport.2021.102002] [Medline: 36567740]

XSL•FO
RenderX

119. Banskota S, Healy M, Goldberg E. 15 smartphone apps for older adults to use while in isolation during the COVID-19 pandemic. West J Emerg Med 2020 Apr 14;21(3):514-525 [FREE Full text] [doi: 10.5811/westjem.2020.4.47372] [Medline: 32302279]
120. Ghani Z, Jarl J, Sanmartin Berglund J, Andersson M, Anderberg P. The cost-effectiveness of mobile health (mHealth) interventions for older adults: systematic review. Int J Environ Res Public Health 2020 Jul 22;17(15):5290 [FREE Full text] [doi: 10.3390/ijerph17155290] [Medline: 32708016]
121. Clet E, Leblanc P, Alla F, Cohidon C. Factors for the integration of prevention in primary care: an overview of reviews. BJGP Open 2024 Apr 05;8(3):BJGPO.2023.0141. [doi: 10.3399/bjgpo.2023.0141]

Abbreviations

AI: artificial intelligence
BCW: Behavior Change Wheel
CONSORT: Consolidated Standards of Reporting Trials
CSAT: Customer Satisfaction Score
DHRQ: Digital Health Readiness Questionnaire
IPAQ-SF: International Physical Activity Questionnaire–Short Form
mHealth: mobile health
MIA: More In Action
NASSS: Nonadoption, Abandonment, Scale-up, Spread, and Sustainability
NPS: Net Promoter Score
PA: physical activity
PACES: Physical Activity Enjoyment Scale
SUS: System Usability Scale

Edited by J Wang; submitted 17.06.24; peer-reviewed by N Santos, T Sun, M Shi, I Hedderson; comments to author 02.10.24; revised version received 23.10.24; accepted 03.12.24; published 27.01.25.

Please cite as:
Daniels K, Vonck S, Robijns J, Quadflieg K, Bergs J, Spooren A, Hansen D, Bonnechère B
Exploring the Feasibility of a 5-Week mHealth Intervention to Enhance Physical Activity and an Active, Healthy Lifestyle in Community-Dwelling Older Adults: Mixed Methods Study
JMIR Aging 2025;8:e63348
URL: https://aging.jmir.org/2025/1/e63348
doi:10.2196/63348
PMID:

©Kim Daniels, Sharona Vonck, Jolien Robijns, Kirsten Quadflieg, Jochen Bergs, Annemie Spooren, Dominique Hansen, Bruno Bonnechère. Originally published in JMIR Aging (https://aging.jmir.org), 27.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Voice-Based Remote Care Program for Vulnerable Older Adults in a Rural Community: Single-Arm Pilot Clinical Study

Geon Young Jang[1*], MD; Sunghwan Ji[1,2*], MD; Ji Yeon Baek[1], MD, PhD; Eunju Lee[1], MD, PhD; Seungryong Chong[3], MS; Il-Young Jang[1,4], MD, PhD

[1]Division of Geriatrics, Department of Internal Medicine, Asan Medical Center, University of Ulsan College of Medicine, 88, Olympic-ro 43-gil, Songpa-gu, Seoul, Republic of Korea

[2]Department of Digital Health, Samsung Advanced Institute for Health Science & Technology, Sungkyunkwan University, Seoul, Republic of Korea

[3]AI Business Partnership Team, SK Telecom, Seoul, Republic of Korea

[4]PyeongChang Health Center and County Hospital, PyeongChang-gun, Gangwon-do, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Eunju Lee, MD, PhD
Division of Geriatrics, Department of Internal Medicine, Asan Medical Center, University of Ulsan College of Medicine, 88, Olympic-ro 43-gil, Songpa-gu, Seoul, Republic of Korea

Abstract

Background: Voice-based digital health technologies are highly feasible and acceptable tools for supporting older adults. However, their development has rarely focused on caregiving needs, and it is often poorly integrated with existing care services, thereby limiting their sustained effect.

Objective: This study aimed to evaluate the feasibility and effectiveness of a comprehensive voice-based remote care program developed in partnership with a local public health center.

Methods: A single-center, single-arm clinical study involving community-dwelling, socioeconomically vulnerable older adults was conducted using a Clinical Frailty Scale of 4 - 5. Participants received a 6-month voice-based care program comprising smart speaker daily check-ins, an emergency response system, and artificial intelligence–driven well-being check calls. These components were integrated with the public health center for continuous monitoring. The primary outcome was caregiver burden, assessed using the Korean version of the Zarit Burden Interview. Secondary outcomes include depression (Patient Health Questionnaire-9), anxiety (Generalized Anxiety Disorder-7), and quality of life (Korean version of the Control, Autonomy, Self-realization, and Pleasure scale).

Results: Among 100 enrolled participants, 96 (96%) completed the program. The caregiver burden slightly decreased from 17.1 - 16.2 points (mean difference −1, 95% CI −2.17 to 0.24; *P*=.12). However, caregivers reported a significant reduction in their perception of being the sole support provider (*P*=.003). Among older adults, significant improvements were observed in depression (Patient Health Questionnaire-9; *P*<.001), anxiety (Generalized Anxiety Disorder-7; *P*=.008), and quality of life (Korean version of the Control, Autonomy, Self-Realization, and Pleasure scale; *P* =.048).. Program adherence was high, with participants engaging for a median of 184 (IQR 154 - 203; 186/214, 87%) days.

Conclusions: Whereas the voice-based remote care program did not significantly reduce the overall caregiver burden, it significantly reduced the perception of the caregivers as being the sole support system. Furthermore, it influenced the psychological well-being of older adults by reducing depression and anxiety and enhancing their quality of life. High adherence and engagement enhance the feasibility and acceptability of scalable digital health interventions for vulnerable older adults in rural settings.

Trial Registration: Clinical Research Information Service (CRIS) KCT0009015; https://cris.nih.go.kr/cris/search/detailSearch.do?seq=27113

(JMIR Aging 2025;8:e76653) doi:10.2196/76653

KEYWORDS

caregiving; informal caregiver; remote care; smart speaker; telemedicine; digital health technology

Introduction

Frailty is defined as a state of increased vulnerability resulting from age-related functional decline across multiple physiological systems, leading to decreased resilience to stress [1]. With global population aging, the number of older adults with frailty is steadily increasing [2]. In South Korea, the prevalence of frailty among individuals aged 65 years and older is approximately

XSL•FO
RenderX

23.1%. Given the rapid pace of population aging in South Korea, the absolute number of older adults with frailty is expected to continue increasing [3]. Older adults experiencing frailty require ongoing, long-term care, creating significant caregiver burden [4]. Informal caregivers, particularly family members or close friends, provide approximately 80% of long-term care in various countries [5]. These caregivers often face mental health challenges, including anxiety, depression, and burnout [6-8]. In addition, they face substantial financial strain, averaging US $7242 annually in the United States, US $10,800 - US 32,400 in South Korea, and US $6000-US $12,000 across Organisation for Economic Co-operation and Development countries [5,9,10]. Despite community support programs, workforce shortages and limited resources hinder continuous care [6,7,11-15], particularly in rural areas with inadequate support services [16].

Digital health technologies that enhance remote monitoring and home-based intervention are essential tools for easing caregiver burden [17,18]. However, digital technologies, including wearable devices, have innate limitations in terms of long-term adherence, particularly among older adults. Voice-based technologies, including smart speakers, demonstrate high feasibility and acceptability among older adults [19-23]. However, most digital health interventions focus on caregiver education or alleviating psychological distress [24-26], primarily addressing mental health, cognitive function, and social engagement [27-30]. In addition, most interventions are device-centric and lack integration with ongoing care services, limiting their ability to respond to evolving care needs [18].

Therefore, this study aims to evaluate the effectiveness of the program on the primary outcomes of caregiver burden and secondary outcomes, including well-being, depression, anxiety, and quality of life among older adults. To bridge this gap, we developed a comprehensive voice-based remote care program in partnership with a local public health center. This intervention integrates an artificial intelligence (AI)-powered speaker and an AI-driven interactive voice response (IVR) call system to provide daily check-ins and support for older adults with pre- to mild frailty in a rural South Korean community.

Methods

Trial Design and Study Participants

A single-arm clinical trial was conducted at PyeongChang Health Center and County Hospital (PCH), the principal public health facility in Pyeongchang County, South Korea. Participants were recruited from the Aging Study of the Pyeongchang Rural Area (ASPRA), a prospective cohort study targeting older adults in the region. The specific protocol for the ASPRA cohort has previously been published [31]. Briefly, the ASPRA inclusion criteria required participants to be ≥65 years old, enrolled in the National Healthcare Service, and able to walk with or without aid. In addition, they needed to live within the community (noninstitutionalized) and be capable of providing informed consent personally or through a legal proxy. Individuals living in nursing homes, hospitalized, or receiving equivalent care at home were excluded. All study participants underwent an annual Comprehensive Geriatric Assessment.

Overall, 1955 older adults were screened from the ASPRA database. The selection was based primarily on frailty and socioeconomic vulnerability. Eligible participants were ≥65, residing within the community, with a Clinical Frailty Scale (CFS) of 4 - 5, and living alone or with low income. Individuals with significant cognitive impairments limiting the use of smart speakers, those with severe hearing loss unable to manage incoming calls, and those with <1-year life expectancy based on patient-reported terminal conditions (metastatic cancer or other end-stage diseases) were excluded.

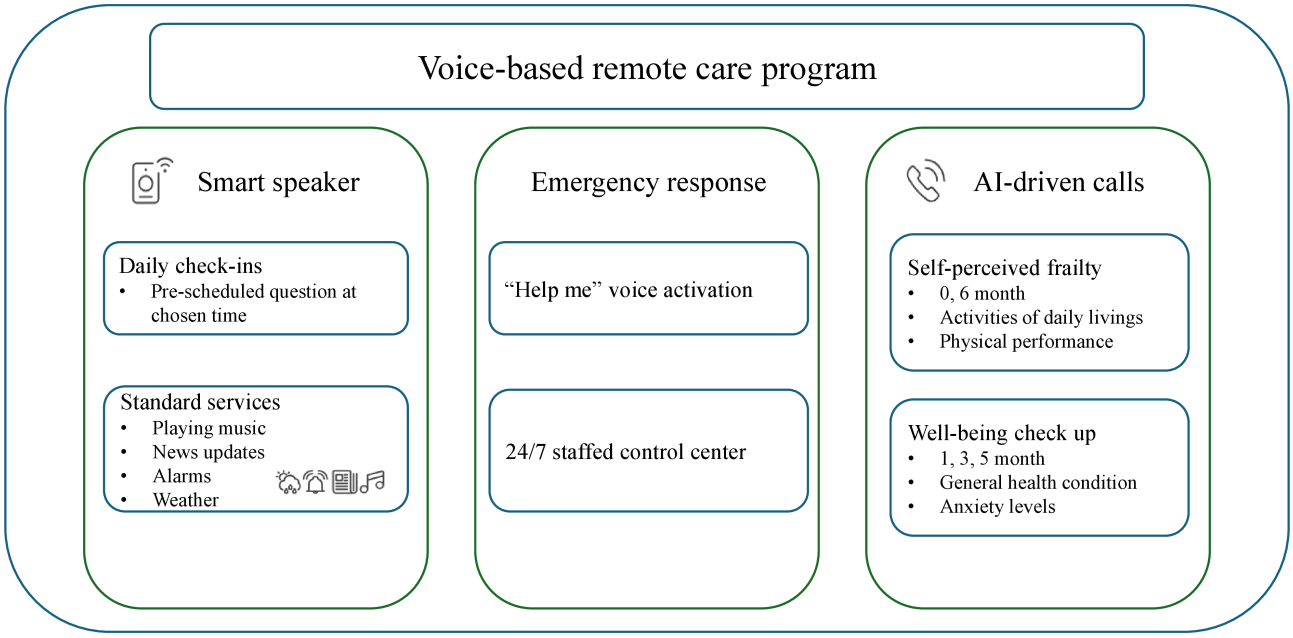
Since this study involved a novel intervention with limited prior research, formal sample size calculations based on statistical power were not applicable. A target sample size of 100 participants was determined through consultation with local health care stakeholders, considering feasibility and regional context.

Intervention

This voice-based remote care program is a structured intervention designed to provide continuous in-home support for older adults using integrated smart technologies. Participants were enrolled in a 6-month program operated by PCH, which delivered services through a smart speaker (SKT NUGU by SK Telecom, Seoul, Korea) and an IVR system. The intervention comprised 3 key components: smart speaker-based daily check-ins, an emergency response system, and AI-driven well-being check calls (care calls), Figure 1.

XSL•FO
RenderX

Figure 1. Conceptual framework of a voice-based remote care program featuring 3 key components: smart speaker, emergency response, and AI-driven calls. AI: artificial intelligence.



Smart speaker-based daily check-ins allowed participants to receive simple, prescheduled questions at their preferred time. These questions consisted of general wellness inquiries with health behavior recommendations (Multimedia Appendix 1). The smart speaker provided immediate responses tailored to inputs, including interactive health guidance and retrieval of relevant health information. When participants requested immediate assistance, the system facilitated direct phone calls to their caregivers or PCH staff. The system automatically logged the participant's response status (responded or did not respond). If no response was detected within a specified period (default: 3 days, adjustable based on participant preference), the system flagged the case for review by PCH staff during regular working days. Staff then conducted follow-up phone calls to participants or their caregivers to confirm their status. In addition, the smart speakers provided standard services, including music playback, news updates, alarm settings, and weather forecasts.

The emergency response system, also managed through a smart speaker, enabled participants to request immediate assistance during falls or other home emergencies. Participants could activate the smart speaker by saying "Help me," to connect directly to a 24/7 staffed control center managed by a private security company. Control center staff assessed the situation in real time via the smart speaker or through a follow-up phone call and, if necessary, promptly dispatched emergency services for immediate intervention. All activations, including details of communications between participants and control center staff, were recorded for monitoring and quality assurance purposes.

The AI-driven well-being check calls used an IVR system to conduct periodic well-being assessments throughout the 6-month intervention period. The IVR system incorporated AI technology (NUGU Interactive technology by SK Telecom), which includes speech recognition, natural language understanding, and dialog management. During this period, the system made 5 scheduled calls to each participant. At baseline and 6 months, comprehensive self-perceived frailty assessments were conducted using a branching questioning algorithm that included Activities of Daily Living (ADL), Instrumental Activities of Daily Living (IADL), and physical performance measures. At 1, 3, and 5 months, brief health status and anxiety assessments were administered. The AI system automatically interpreted participant responses, classified general health status as positive or negative, determined self-perceived frailty scores, and identified newly developed disabilities in ADL or IADL. PCH staff monitored these categorized responses monthly. When participants showed signs of deterioration, they were advised to visit PCH or other health care facilities.

At initial setup, PCH staff visited participant homes to install and configure smart speakers and Wi-Fi systems, providing user training. Technical support was available throughout the intervention period, with participants able to contact PCH for basic troubleshooting or specialized technical assistance for complex issues.

Measurements

Participants underwent a comprehensive geriatric assessment. Demographic data collected included age, sex, height, weight, medical aid status, living arrangement (living alone or not), caregiver relationship, and smartphone usage. Comorbidities and the number of regularly prescribed medications were also recorded. Physical performance was assessed using the short physical performance battery, which ranges from 0 - 12, along with a measurement of handgrip strength [32]. Disability status was defined as the presence of 1 or more impairments in any of the 7 ADL items or 10 IADL items [33]. Cognitive function was evaluated using the Mini-Cog test, with cognitive dysfunction defined as a score of ≤3 [34]. Frailty was assessed using the Korean version of the CFS and a 45-item Frailty Index (Multimedia Appendix 2) [35]. All assessments were conducted in person by trained nurses at PCH.

XSL•FO
RenderX

Outcomes

The primary outcome was caregiver burden, measured using the Korean version of the Zarit Burden Interview (ZBI-K) [36], administered to caregivers in person at baseline and the end of the 6-month intervention. The ZBI-K comprised 22 items, each rated on a 5-point Likert scale (0=never to 4=always), yielding an overall score range of 0 - 88 [37]. Higher ZBI-K scores indicated a greater level of caregiver burden and were categorized as follows: 0 - 20 (little or no burden), 21 - 40 (mild to moderate burden), 41 - 60 (moderate to severe burden), and 61 - 88 (severe burden) [38,39]. Multimedia Appendix 3 provides the complete ZBI-K questionnaire.

Secondary outcomes included changes in psychological well-being and quality of life among older adult participants, assessed in person at PCH at baseline and 6 months. Depressive symptoms were assessed using the Patient Health Questionnaire-9 (PHQ-9), a 9-item instrument with an overall score range of 0 - 27 [40,41]. Anxiety levels were measured using the Generalized Anxiety Disorder-7 (GAD-7), a 7-item scale ranging from 0 - 21 [42,43]. Quality of life was evaluated using the Korean version of the Control, Autonomy, Self-Realization, and Pleasure Scale (K-CASP-19), which has a score range of 0 - 57 [44,45] and the EuroQol-5 Dimension [46,47].

In addition, smart speaker usage patterns were monitored throughout the intervention period. Data collected included the number of days the smart speaker was used, total usage frequency, and average daily interactions. The activation of the emergency response system and engagement with the AI telephone service (care call response rate) were also recorded to assess feasibility and participant compliance.

Statistical Methods

Baseline characteristics were summarized using means (SD) values for continuous variables and frequencies with percentages for categorical variables. To evaluate changes in primary and secondary outcomes over time, a linear mixed-effect model was used, specifying time as a fixed effect and participant ID as a random effect. Estimated marginal means were calculated for each time point, and pairwise comparisons were performed to assess mean differences between baseline and follow-up assessments. CI values for mean changes were calculated, and effect sizes were estimated using Cohen *d*. For individual ZBI-K items, pre–post differences were evaluated using the Wilcoxon signed-rank test. All statistical analyses were performed using R version 4.4.1 (R Core Team), with a significance level set at α=.05.

Ethical Considerations

The trial was registered with the Clinical Research Information Service of the Republic of Korea (identifier: KCT0009015) and conducted in accordance with the Declaration of Helsinki. Ethical approval was obtained from the Institutional Review Board of Asan Medical Center (institutional review board approval number 2023-0587). All participants provided written informed consent prior to enrollment. Participant confidentiality was strictly maintained, and all data were anonymized before analysis. No monetary compensation was provided for participation.

Results

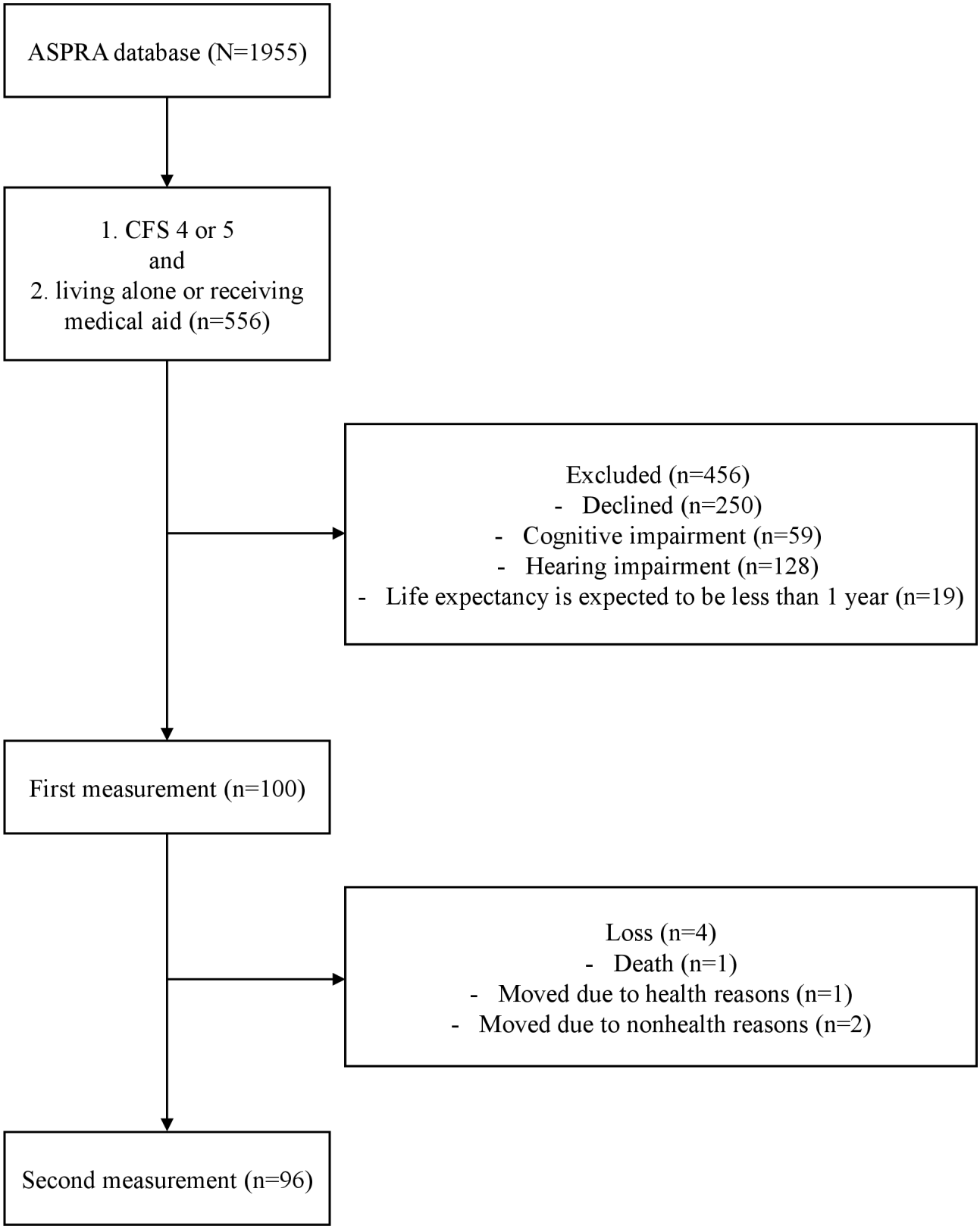
Participants

Between June 1 and December 31, 2023, 556 participants were identified from the ASPRA database based on the eligibility criteria: a CFS score of 4 or 5 and either living alone or receiving medical aid. However, 456 individuals were excluded owing to refusal to participate (n=250), cognitive impairment (n=59), hearing impairment (n=128), or a life expectancy of <1 year (n=19). Consequently, 100 participants were enrolled in the study (Figure 2). During the study period, 4 participants withdrew from the intervention: 1 owing to death, 2 owing to relocation for non–health-related reasons, and 1 relocated for health-related reasons. Consequently, 96 participants completed the second measurement. Table 1 shows the baseline characteristics of the participants. The majority were female (84/100, 84%), with a mean age of 76.5 (SD 6.1) years. A significant proportion of participants (66/100, 66%) lived alone and most were capable of using smartphones (80/100, 80%).

Regarding caregiving support, most participants received care from their immediate family members. Among male participants, the primary caregivers were most often their spouses (10/16, 62.5%). In contrast, female participants were predominantly cared for by their children (58/84, 70.2%). A smaller proportion of the participants received care from their daughters-in-law (5/84, 6%), siblings (3/84, 3.6%), or other caregivers (2/84, 2.4%).

XSL•FO
RenderX

Figure 2. Flowchart of participant selection and follow-up. ASPRA: Aging Study of the Pyeongchang Rural Area; CFS: Clinical Frailty Scale.



XSL•FO
RenderX

Table . Baseline characteristics of the study groups.

Characteristics	Male (n=16)	Female (n=84)
Age (years), mean (SD)	75.69 (6.83)	76.70 (5.88)
Medical aid, n (%)	3 (18.8)	9 (10.7)
Living alone, n (%)	6 (37.5)	59 (70.2)
Smartphone usage, n (%)	14 (87.5)	66 (78.6)
Caregiver, n (%)		
Spouses	10 (62.5)	16 (19)
Children	6 (37.5)	58 (70.2)
Daughters-in-law	0 (0)	5 (6)
Siblings	0 (0)	3 (3.6)
Others	0 (0)	2 (2.4)
BMI (kg/m^2), mean (SD)	25.66 (2.61)	26.02 (4.24)
Skeletal muscle index (kg/m^2), mean (SD)	7.46 (0.70)	6.12 (0.88)
Gait speed (m/s), mean (SD)	0.78 (0.18)	0.81 (0.28)
SPPB[a] (score), mean (SD)	8.80 (1.61)	8.70 (2.38)
Grip strength (kg), mean (SD)	30.49 (8.49)	21.13 (5.15)
Mini-Cog≤3, n (%)	9 (56.2)	50 (59.5)
No. of chronic conditions, mean (SD)	2.25 (1.48)	2.46 (1.43)
No. of medication, mean (SD)	4.25 (2.62)	3.56 (2.39)
Fall in the last year, n (%)	4 (25)	36 (42.9)
ADL[b] disability, n (%)	2 (12.5)	44 (52.4)
IADL[c] disability, n (%)	3 (18.8)	42 (50)
CFS[d], mean (SD)	3.44 (0.63)	3.57 (0.72)
Frailty index, mean (SD)	0.19 (0.09)	0.25 (0.12)

[a]SPPB: short physical performance battery.

[b]ADL: activities of daily living.

[c]IADL: instrumental activities of daily living.

[d]CFS: Clinical Frailty Scale.

Primary Outcome

Caregiver burden, assessed using the Korean version of the ZBI-K, showed a slight, nonsignificant reduction more than the 6-month intervention period. The mean score decreased from 17.1 at baseline to 16.2 at follow-up, with a mean difference of −1 (95% CI −2.17 to 0.24; *P*=.12). When individual ZBI-K items were analyzed, most showed minimal or nonsignificant changes. However, item 14—which assessed the perception of the caregivers as being the sole source of support—demonstrated a significant reduction from baseline to follow-up (mean change −0.43, Cohen *d*=−0.34; *P*=.003). Multimedia Appendix 4 presents details of other item-level analyses.

Secondary Outcomes

Regarding psychological function, anxiety levels, as measured by the GAD-7 scale, decreased significantly from a mean score of 2.97 to 2.18 (mean change −0.79, 95% CI −1.37 to −0.81; *P*=.008). Similarly, depression, assessed using the PHQ-9, showed a significant reduction from a mean score of 5.68 to 4.02 (mean change −1.66, 95% CI −2.41 to −0.90; *P*<.001).

Quality of life, evaluated via the K-CASP-19, significantly improved, with scores increasing from 27.3 - 28.8 (mean change 1.53, 95% CI 0.01 - 3.05; *P*=.048). However, the EuroQol-5 Dimension score exhibited only a minimal increase from 0.77 to 0.78, which was not statistically significant (mean change 0.01, 95% CI −0.01 to 0.03; *P*=.46). A summary of the primary and secondary outcomes is presented in Table 2.

XSL•FO
RenderX

Table . Primary and secondary outcomes.

Outcomes	Values, mean (SD)		Change		
	Baseline	Follow-up	Mean change (95% CI)	Cohen *d*	*P* value[a]
Primary outcome					
ZBI-K[b]	17.1 (7.9)	16.2 (8.9)	−1.0 (−2.17 to 0.24)	−0.15	.12
Secondary outcomes					
GAD-7[c]	2.97 (3.49)	2.18 (2.87)	−0.79 (−1.37 to −0.81)	−0.23	.008
K-CASP-19[d]	27.3 (7.6)	28.8 (8.0)	1.53 (0.01 to 3.05)	0.21	.048
EQ-5D[e]	0.77 (0.12)	0.78 (0.01)	0.01 (−0.01 to 0.03)	0.08	.46
PHQ-9[f]	5.68 (4.86)	4.02 (3.73)	−1.66 (−2.41 to −0.90)	−0.37	<.001

[a]Linear mixed-effects model.

[b]ZBI-K: Korean version of the Zarit Burden Interview.

[c]GAD-7: Generalized Anxiety Disorder-7.

[d]K-CASP-19: Korean version of the Control, Autonomy, Self-realization, and Pleasure scale–19 items.

[e]EQ-5D: EuroQol 5-dimension questionnaire.

[f]PHQ-9: Patient Health Questionnaire-9.

Technology Usage and Compliance

Overall, the participants exhibited high usage rates of the smart speaker device. Multimedia Appendix 5 illustrates that the participants used the smart speaker for a median of 184 (IQR 154 - 203.25) days out of the total 214 days, representing approximately 87% (186/214) of the intervention period. Total usage frequency varied across participants, with a median of 508 interactions (IQR 374.75 - 729.25). The lowest usage was 7, whereas the highest reached 2136 interactions, highlighting the diverse usage patterns. On average, daily usage was 2.38 interactions (IQR 1.75 - 3.41), with individual rates ranging from 0.03 - 9.98 interactions per day.

Multimedia Appendix 6 shows that the AI-driven well-being check calls demonstrated high response rates throughout the study. At baseline (0 months), all participants (100/100, 100%) responded to the care call. This response rate remained unchanged for 1 month (99/99, 100%) and 3 months (98/98, 100%). However, in the 5-month assessment, the response rate decreased to 91.8% (90/97) among the 97 remaining participants. Reasons for nonresponse at the 5-month mark included connection failures (n=5), inability to verify identity (n=3), and refusal to engage in conversation (n=1). By the final 6-month assessment, the response rate returned to 100% among the 96 participants who completed the study.

The emergency response system, specifically the “Help me” function, was activated 3 times during the study period. However, each activation was determined to be a false alarm caused by the misinterpretation of voice commands rather than actual emergency situations. No genuine emergency events requiring intervention were recorded throughout the study period.

Discussion

Principal Findings

This pilot study evaluated a voice-based remote care system, integrating an AI speaker and IVR call system to support vulnerable older adults in rural South Korea, focusing on caregiver burden, psychological well-being, and intervention feasibility. Although the primary outcome—caregiver burden, measured using the ZBI-K—did not show a statistically significant improvement, the total score showed slight improvement. In addition, item-level analysis revealed a significant reduction in the perception of caregivers as the sole provider of support. Older adults showed significant improvements in depression, anxiety, and quality of life. High engagement with the smart speaker, along with near-complete response rates to AI-driven check-in calls, supports the feasibility and acceptability of the intervention.

The mean ZBI-K score showed a slight decrease from 17.1 at baseline to 16.2 at follow-up (*P*=.12). This minimal change may be partly attributed to the low baseline burden among caregivers, as a score of 17.1 typically indicates low burden [38,39]. Since our participants were community-dwelling older adults, primarily socioeconomically vulnerable and ranging from prefrail to mildly frail, the caregiver burden was relatively low compared to those of caregivers of individuals with multiple chronic conditions (mean ZBI-K 38.27, SD 14.58) [48] or dementia (mean ZBI-K 51.09, SD 17.95) [49].

Nevertheless, ZBI-K item 14 (“Do you feel that your relative seems to expect you to take care of them as if you were the only one they could depend on?”) showed significant improvement with a mean reduction of 0.43 points (Cohen *d*=−0.34; *P*=.003). It specifically measures the caregiver’s perception of being the sole provider of support. Previous qualitative studies have similarly suggested that voice-based devices provide caregivers with psychological reassurance, emotional support, scheduling

XSL•FO
RenderX

assistance, and emergency response [50-52]. Direct comparisons with these studies are challenging due to differences in research settings, such as dementia status, caregiver type, and living arrangements. Despite these limitations, our findings provide quantitative support for the psychological reassurance and reduced caregiving responsibility indicated by earlier qualitative research.

This intervention had significant effects on secondary outcomes assessed among older adult participants. Notable improvements were observed in depression (PHQ-9 scores decreased from 5.68 to 4.02, mean change −1.66; P<.001), anxiety (GAD-7 scores decreased from 2.97 to 2.18, mean change −0.79; P=.008), and quality of life (K-CASP-19 scores increased from 27.3 to 28.8, mean change 1.53; P=.048). These findings align with previous qualitative studies, which suggested companionship, enhanced social connectedness, and personalization through user experiences as potential mechanisms underlying improvements in loneliness and depression [27-29]. While our study used standardized assessment tools and quantitatively demonstrated psychological improvements, it did not directly investigate these mechanisms. Future research could further explore these proposed hypotheses.

This intervention demonstrated 3 key strengths that contributed to overall effectiveness, high adaptability, and sustainability of this program. First, it is a voice-based system using an AI-powered speaker and IVR calls, providing a more user-friendly interface for older adults compared to smartphones or wearable devices, which often present usability challenges [53]. Second, the program integrates 3 core components—smart speaker-based daily check-ins, emergency response system, and AI-driven well-being check calls—into a single, seamless system that connects directly with the local public health infrastructure. Unlike previous interventions, which typically focused on a single component [20-22,27-30,50-52,54,55], this multi-layered design offers more comprehensive support and responsive care. Third, the intervention was designed to reduce dependence on human resources, addressing a key limitation of traditional telecare systems such as videoconferencing or telephone support, which usually require active involvement from health care professionals, thereby limiting scalability [17,56].

This study has some limitations. First, the absence of a control group limits our ability to establish causality, as improvements in secondary outcomes may have been influenced by external factors. However, to mitigate selection bias, we included all eligible participants from the ASPRA cohort database who met our inclusion criteria without additional selection processes. Second, the intervention primarily focused on care recipients rather than caregivers, which may explain the lack of significant improvement in overall caregiver burden. Third, the focus on rural older adults is a notable strength of this study, addressing a population typically underrepresented in digital health research; however, these findings may not be directly applicable to urban settings or different cultural settings. Fourth, our study was limited to informal caregivers. Formal caregivers differ from informal caregivers in their expectations toward voice-based technology, particularly regarding professional accuracy and efficiency versus emotional support.

Therefore, future research should implement randomized controlled trial designs to establish causality. In addition, incorporating more diverse populations and expanding the intervention to support caregivers alongside care recipients may offer a more comprehensive approach to reducing caregiver burden. A mixed-methods approach, combining quantitative assessments with qualitative feedback from caregivers, could also provide valuable insights into psychological improvements and practical impacts of the program.

Conclusions

Voice-based remote care program using smart speakers and AI-driven telephone services for vulnerable older adults in rural areas did not significantly reduce the overall caregiver burden. However, it reduced the perception of being the sole support provider. For participants, the intervention improves psychological well-being, significantly reducing depression and anxiety and enhancing quality of life. High engagement and adherence confirm the feasibility and acceptability of this intervention among older adults in rural settings. These findings highlight the potential of smart speaker-based remote care as a scalable supplement to traditional services in rural areas, improving psychological well-being and achieving high acceptance among vulnerable older adults.

Funding

This research was supported by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute, funded by the Ministry of Health and Welfare, Republic of Korea (grant number HR20C0026). This study was supported by a grant from the Asan Institute for Life Sciences, Asan Medical Center, Seoul, Korea (grant number 2024IF0033).

Conflicts of Interest

SC is an employee of SK Telecom but participated in this study in a personal capacity. SK Telecom had no role in the study design, data collection, analysis, interpretation, or article preparation.

Multimedia Appendix 1
Two-week examples of questions from “the smart speaker-based daily check-ins”.
[DOCX File, 18 KB - aging_v8i1e76653_app1.docx]

Multimedia Appendix 2

XSL•FO
RenderX

The 45 items of the frailty index.
[DOCX File, 18 KB - aging_v8i1e76653_app2.docx]

Multimedia Appendix 3
Zarit Burden interview.
[DOCX File, 20 KB - aging_v8i1e76653_app3.docx]

Multimedia Appendix 4
Changes in individual Korean version of the Zarit Burden Interview items from baseline to 6-month follow-up.
[DOCX File, 21 KB - aging_v8i1e76653_app4.docx]

Multimedia Appendix 5
Smart speaker usage patterns during the 6-month intervention period.
[DOCX File, 18 KB - aging_v8i1e76653_app5.docx]

Multimedia Appendix 6
Care call response rates throughout the intervention period.
[DOCX File, 17 KB - aging_v8i1e76653_app6.docx]

References

1. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet 2013 Mar 2;381(9868):752-762. [doi: 10.1016/S0140-6736(12)62167-9] [Medline: 23395245]
2. Hoogendijk EO, Afilalo J, Ensrud KE, Kowal P, Onder G, Fried LP. Frailty: implications for clinical practice and public health. Lancet 2019 Oct 12;394(10206):1365-1375. [doi: 10.1016/S0140-6736(19)31786-6] [Medline: 31609228]
3. Kang MG, Kim OS, Hoogendijk EO, Jung HW. Trends in frailty prevalence among older adults in Korea: a nationwide study from 2008 to 2020. J Korean Med Sci 2023 Jul 24;38(29):e157. [doi: 10.3346/jkms.2023.38.e157] [Medline: 37489714]
4. Ringer T, Hazzan AA, Agarwal A, Mutsaers A, Papaioannou A. Relationship between family caregiver burden and physical frailty in older adults without dementia: a systematic review. Syst Rev 2017 Mar 14;6(1):55. [doi: 10.1186/s13643-017-0447-1] [Medline: 28292313]
5. OECD Health Policy Studies. Who cares? Attracting and retaining care workers for the elderly. : OECD Publishing; 2020 May 26. [doi: 10.1787/92c0ef68-en]
6. Adelman RD, Tmanova LL, Delgado D, Dion S, Lachs MS. Caregiver burden: a clinical review. JAMA 2014 Mar 12;311(10):1052-1060. [doi: 10.1001/jama.2014.304] [Medline: 24618967]
7. Schulz R, Newsom J, Mittelmark M, Burton L, Hirsch C, Jackson S. Health effects of caregiving: the caregiver health effects study: an ancillary study of the cardiovascular health study. Ann Behav Med 1997;19(2):110-116. [doi: 10.1007/BF02883327] [Medline: 9603685]
8. Huis in ‘t Veld T, Korver D, Orhan R, Berenschot L. The economic value of informal long-term care. Eur J Public Health 2022 Oct 21;32(Supplement_3). [doi: 10.1093/eurpub/ckac130.121]
9. Caregiving out-of-pocket costs study. : AARP; 2021. [doi: 10.26419/res.00473.001]
10. Kim S, Lee SH, Namkung EH, et al. Changes in the lives and perceptions of the elderly in Korea: National survey of older Koreans. Research@KIHASA. 2021. URL: https://repository.kihasa.re.kr/en/handle/201002/39783 [accessed 2025-11-04]
11. Bremer P, Cabrera E, Leino-Kilpi H, et al. Informal dementia care: consequences for caregivers’ health and health care use in 8 European countries. Health Policy 2015 Nov;119(11):1459-1471. [doi: 10.1016/j.healthpol.2015.09.014] [Medline: 26518906]
12. Hussein S, Manthorpe J. An international review of the long-term care workforce: policies and shortages. J Aging Soc Policy 2005;17(4):75-94. [doi: 10.1300/J031v17n04_05] [Medline: 16380370]
13. Toseland RW, McCallion P, Gerber T, Banks S. Predictors of health and human services use by persons with dementia and their family caregivers. Soc Sci Med 2002 Oct;55(7):1255-1266. [doi: 10.1016/s0277-9536(01)00240-4] [Medline: 12365535]
14. Phillipson L, Jones SC, Magee C. A review of the factors associated with the non-use of respite services by carers of people with dementia: implications for policy and practice. Health Soc Care Community 2014 Jan;22(1):1-12. [doi: 10.1111/hsc.12036] [Medline: 23530738]
15. Greenwood N, Habibi R, Smith R, Manthorpe J. Barriers to access and minority ethnic carers’ satisfaction with social care services in the community: a systematic review of qualitative and quantitative literature. Health Soc Care Community 2015 Jan;23(1):64-78. [doi: 10.1111/hsc.12116] [Medline: 25135207]
16. Miller KEM, Ornstein KA, Coe NB. Rural disparities in use of family and formal caregiving for older adults with disabilities. J Am Geriatr Soc 2023 Sep;71(9):2865-2870. [doi: 10.1111/jgs.18376] [Medline: 37081828]
17. Chi NC, Demiris G. A systematic review of telehealth tools and interventions to support family caregivers. J Telemed Telecare 2015 Jan;21(1):37-44. [doi: 10.1177/1357633X14562734] [Medline: 25475220]

XSL•FO
RenderX

18. Lindeman DA, Kim KK, Gladstone C, Apesoa-Varano EC. Technology and caregiving: emerging interventions and directions for research. Gerontologist 2020 Feb 14;60(Suppl 1):S41-S49. [doi: 10.1093/geront/gnz178] [Medline: 32057082]
19. McCloud R, Perez C, Bekalu MA, Viswanath K. Using smart speaker technology for health and well-being in an older adult population: pre-post feasibility study. JMIR Aging 2022 May 9;5(2):e33498. [doi: 10.2196/33498] [Medline: 35532979]
20. Kim S, Choudhury A. Exploring older adults' perception and use of smart speaker-based voice assistants: a longitudinal study. Comput Human Behav 2021 Nov;124:106914. [doi: 10.1016/j.chb.2021.106914]
21. Nunez M, Patel P, Ulin L, et al. Feasibility and usage of a virtual assistant device in cognitively impaired homebound older adults. J Appl Gerontol 2025 Oct;44(10):1651-1660. [doi: 10.1177/07334648251314284] [Medline: 39792590]
22. Quinn K, Leiser Ransom S, O'Connell C, Muramatsu N, Marquez DX, Chin J. Assessing the feasibility and acceptability of smart speakers in behavioral intervention research with older adults: mixed methods study. J Med Internet Res 2024 Aug 30;26:e54800. [doi: 10.2196/54800] [Medline: 39213034]
23. Kokorelias KM, Grigorovich A, Harris MT, et al. Longitudinal coadaptation of older adults with wearables and voice-activated virtual assistants: scoping review. J Med Internet Res 2024 Aug 7;26:e57258. [doi: 10.2196/57258] [Medline: 39110963]
24. Zhai S, Chu F, Tan M, Chi NC, Ward T, Yuwen W. Digital health interventions to support family caregivers: an updated systematic review. Digit Health 2023;9:20552076231171967. [doi: 10.1177/20552076231171967] [Medline: 37223775]
25. Su Z, McDonnell D, Liang B, et al. Technology-based health solutions for cancer caregivers to better shoulder the impact of COVID-19: a systematic review protocol. Res Sq 2021 Jan 19. [doi: 10.21203/rs.3.rs-66218/v2] [Medline: 32908975]
26. Gentili A, Failla G, Melnyk A, et al. The cost-effectiveness of digital health interventions: a systematic review of the literature. Front Public Health 2022;10:787135. [doi: 10.3389/fpubh.2022.787135] [Medline: 36033812]
27. Marziali RA, Franceschetti C, Dinculescu A, et al. Reducing loneliness and social isolation of older adults through voice assistants: literature review and bibliometric analysis. J Med Internet Res 2024 Mar 18;26:e50534. [doi: 10.2196/50534] [Medline: 38498039]
28. Corbett CF, Wright PJ, Jones K, Parmer M. Voice-activated virtual home assistant use and social isolation and loneliness among older adults: mini review. Front Public Health 2021;9:742012. [doi: 10.3389/fpubh.2021.742012] [Medline: 34708017]
29. Park S, Kim B. The impact of everyday AI-based smart speaker use on the well-being of older adults living alone. Technol Soc 2022 Nov. [doi: 10.1016/j.techsoc.2022.102133]
30. Kim J, Shin E, Han K, et al. Efficacy of smart speaker-based metamemory training in older adults: case-control cohort study. J Med Internet Res 2021 Feb 16;23(2):e20177. [doi: 10.2196/20177] [Medline: 33591276]
31. Baek JY, Lee E, Oh G, et al. The aging study of pyeongchang rural area (ASPRA): findings and perspectives for human aging, frailty, and disability. Ann Geriatr Med Res 2021 Sep;25(3):160-169. [doi: 10.4235/agmr.21.0100] [Medline: 34610665]
32. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994 Mar;49(2):M85-M94. [doi: 10.1093/geronj/49.2.m85] [Medline: 8126356]
33. Won CW, Yang KY, Rho YG, et al. The development of Korean activities of daily living(K-ADL) and Korean instrumental activities of daily living(K-IADL) scale [Article in Korean]. J Korean Geriatr Soc 2002 Jun;6(2):107-120 [FREE Full text]
34. Wright AEH, Harrell HE. Physical examination in the evaluation of dementia. Med Clin North Am 2022 May;106(3):471-482. [doi: 10.1016/j.mcna.2021.12.009] [Medline: 35491067]
35. Ko RE, Moon SM, Kang D, et al. Translation and validation of the Korean version of the clinical frailty scale in older patients. BMC Geriatr 2021 Jan 13;21(1):47. [doi: 10.1186/s12877-021-02008-0] [Medline: 33441092]
36. Yoon E, Robinson M. Psychometric properties of the Korean version of the Zarit Burden Interview (K-ZBI): preliminary analyses. J Soc Work Res Eval 2005;6(1):75 [FREE Full text]
37. Zarit SH, Todd PA, Zarit JM. Subjective burden of husbands and wives as caregivers: a longitudinal study. Gerontologist 1986 Jun;26(3):260-266. [doi: 10.1093/geront/26.3.260] [Medline: 3721233]
38. Stagg B, Larner AJ. Zarit Burden Interview: pragmatic study in a dedicated cognitive function clinic. Prog Neurol Psychiatry 2015 Jul;19(4):23-27. [doi: 10.1002/pnp.390]
39. James K, Chin-Bailey C, Holder-Nevins D, Thompson C, Donaldson-Davis K, Eldemire-Shearer D. Zarit burden interview among caregivers of community-dwelling older adults in a caribbean setting (Jamaica): reliability and factor structure. Health Soc Care Community 2021 Sep;29(5):e79-e88. [doi: 10.1111/hsc.13244] [Medline: 33252838]
40. Park SJ, Choi HR, Choi JH, Kim KW, Hong JP. Reliability and validity of the Korean version of the Patient Health Questionnaire-9 (PHQ-9) [Article in Korean]. Anxiety and mood 2010;6(2):119-124 [FREE Full text]
41. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
42. Ahn JK, Kim Y, Choi KH. The psychometric properties and clinical utility of the Korean version of GAD-7 and GAD-2. Front Psychiatry 2019;10:127. [doi: 10.3389/fpsyt.2019.00127] [Medline: 30936840]
43. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]

XSL•FO
RenderX

44. Kim GR, Netuveli G, Blane D, et al. Psychometric properties and confirmatory factor analysis of the CASP-19, a measure of quality of life in early old age: the HAPIEE study. Aging Ment Health 2015 Jul;19(7):595-609. [doi: 10.1080/13607863.2014.938605] [Medline: 25059754]
45. Hyde M, Wiggins RD, Higgs P, Blane DB. A measure of quality of life in early old age: the theory, development and properties of a needs satisfaction model (CASP-19). Aging Ment Health 2003 May;7(3):186-194. [doi: 10.1080/1360786031000101157] [Medline: 12775399]
46. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [doi: 10.1016/0168-8510(90)90421-9] [Medline: 10109801]
47. Lee YK, Nam HS, Chuang LH, et al. South Korean time trade-off values for EQ-5D health states: modeling with observed values for 101 health states. Value Health 2009;12(8):1187-1193. [doi: 10.1111/j.1524-4733.2009.00579.x] [Medline: 19659703]
48. Lee ES, Kim MY. Effect of perceived chronic illness management support, health literacy, and social support on the care burden of families caring for older people with multiple chronic conditions at home: a cross-sectional study. J Korean Gerontol Nurs 2023 Feb;25(1):76-88. [doi: 10.17079/jkgn.2023.25.1.76]
49. Sung MR, Lee DY, Kim KA. Factors influencing burnout in primary family caregivers of elders with dementia [Article in Korean]. J Korean Gerontol Nurs 2012 Dec;14(3):200-208 [FREE Full text]
50. O'Brien K, Liggett A, Ramirez-Zohfeld V, Sunkara P, Lindquist LA. Voice-controlled intelligent personal assistants to support aging in place. J Am Geriatr Soc 2020 Jan;68(1):176-179. [doi: 10.1111/jgs.16217] [Medline: 31617581]
51. Bults M, van Leersum CM, Olthuis TJJ, et al. Acceptance of a digital assistant (Anne4Care) for older adult immigrants living with dementia: qualitative descriptive study. JMIR Aging 2024 Apr 19;7:e50219. [doi: 10.2196/50219] [Medline: 38639994]
52. Tennant R, Allana S, Mercer K, Burns CM. Caregiver expectations of interfacing with voice assistants to support complex home care: mixed methods study. JMIR Hum Factors 2022 Jun 30;9(2):e37688. [doi: 10.2196/37688] [Medline: 35771594]
53. Khamaj A, Ali AM. Examining the usability and accessibility challenges in mobile health applications for older adults. Alexandria Engineering Journal 2024 Sep;102:179-191. [doi: 10.1016/j.aej.2024.06.002]
54. Shah RR, Dixon CC, Fowler MJ, et al. Using voice assistant systems to improve dietary recall among older adults: perspectives of registered dietitians. J Nutr Gerontol Geriatr 2024;43(1):1-13. [doi: 10.1080/21551197.2024.2302619] [Medline: 38287658]
55. Chung S, Woo BK. Using consumer perceptions of a voice-activated speaker device as an educational tool. JMIR Med Educ 2020 Apr 24;6(1):e17336. [doi: 10.2196/17336] [Medline: 32329740]
56. Wu M, Li C, Hu T, et al. Effectiveness of telecare interventions on depression symptoms among older adults: systematic review and meta-analysis. JMIR Mhealth Uhealth 2024 Jan 17;12:e50787. [doi: 10.2196/50787] [Medline: 38231546]

Abbreviations

ADL: activities of daily living
AI: artificial intelligence
ASPRA: Aging Study of Pyeongchang Rural Area
CFS: Clinical Frailty Scale
GAD-7: Generalized Anxiety Disorder-7
IADL: instrumental activities of daily living
IVR: interactive voice response
PCH: Pyeongchang County Health Center and County Hospital
PHQ-9: Patient Health Questionnaire-9
ZBI-K: Korean version of Zarit Burden Interview

Edited by Y Jiang; submitted 28.04.25; peer-reviewed by JY Lee, LF Belo, S Sharma; revised version received 23.07.25; accepted 08.10.25; published 13.11.25.

Please cite as:
Jang GY, Ji S, Baek JY, Lee E, Chong S, Jang IY
Voice-Based Remote Care Program for Vulnerable Older Adults in a Rural Community: Single-Arm Pilot Clinical Study
JMIR Aging 2025;8:e76653
URL: https://aging.jmir.org/2025/1/e76653
doi:10.2196/76653

XSL•FO
RenderX

© Geon Young Jang, Sunghwan Ji, Ji Yeon Baek, Eunju Lee, Seungryong Chong, Il-Young Jang. Originally published in JMIR Aging (https://aging.jmir.org), 13.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluating the Effectiveness of Digital Social Robots in Reducing Loneliness Among Community-Dwelling Older Adults in Japan: Randomized Controlled Trial and Qualitative Analysis

Hiroshi Murayama, MPH, PhD; Mai Takase, PhD

Research Team for Social Participation and Healthy Aging, Tokyo Metropolitan Institute for Geriatrics and Gerontology, 35-2 Sakae-cho, Itabashi-ku, Tokyo, Japan

Corresponding Author:
Hiroshi Murayama, MPH, PhD
Research Team for Social Participation and Healthy Aging, Tokyo Metropolitan Institute for Geriatrics and Gerontology, 35-2 Sakae-cho, Itabashi-ku, Tokyo, Japan

Abstract

Background: Most studies on interventions using social robots to reduce loneliness have been conducted in facilities in Western countries.

Objective: This study evaluated the effectiveness of digital social robot interventions in reducing loneliness among community-dwelling older Japanese adults using a randomized controlled trial and qualitative analysis.

Methods: Individuals aged ≥65 years who lived alone in Tokyo and neighboring areas and experienced loneliness were recruited. In total, 73 eligible participants were randomly assigned to either an intervention or a control group. The 4-week intervention involved a humanoid social communication robot (BOCCO emo), which facilitated conversations with human operators and family members and reminded participants of daily tasks. The primary outcome was loneliness, with mental health (psychological well-being, depression, and self-rated health), the frequency of laughter in daily life, health competence, and interpersonal relationships (social network and generalized trust) as secondary outcomes. Participants were evaluated at baseline and follow-up using a self-administered questionnaire. In the follow-up survey, participants in the intervention group provided open-ended responses regarding their experiences using the social robot.

Results: In total, 68 participants completed both the baseline and follow-up surveys (34 in each group). The average age of the participants was 82.3 (SD 6.5) years, and 64 (N=68, 94%) participants were women. A linear mixed-effects model with a random intercept indicated that loneliness decreased more in the intervention group than in the control group (difference-in-difference −3.1, 95% CI −5.9 to −0.4). Psychological well-being also improved in the intervention group (difference-in-difference 1.9, 95% CI 0.1 to 3.7). We identified 4 categories through content analysis: *emotional support and psychological connection*, *lifestyle assistance*, *enrichment of social interaction*, and *cognitive and mental stimulation*.

Conclusions: Social robots can reduce loneliness among community-dwelling older adults in non-Western societies. Information and communication technology appears to be an effective approach for alleviating loneliness and enhancing well-being among older adults in community settings.

Trial Registration: UMIN Clinical Trials Registry UMIN000050644; https://center6.umin.ac.jp/cgi-open-bin/ctr_e/ctr_view.cgi?recptno=R000057702

(JMIR Aging 2025;8:e74422) doi:10.2196/74422

KEYWORDS

social robots; loneliness; psychological well-being; randomized controlled trial; qualitative analysis

Introduction

Loneliness is an increasing public health concern, particularly among older adults [1]. The prevalence of loneliness is expected to increase as the global population ages, placing greater strain on health care systems [2,3]. Subjectively, loneliness arises when actual and desired social connections vary, leading to negative emotions such as anxiety, low self-esteem, and social distress [4]. Meta-analyses associate loneliness with an increased risk of mortality [5,6], cardiovascular and cerebrovascular diseases [6,7], mental disorders [6], and dementia [8]. Psychosocial interventions have been effective in reducing loneliness among older adults [2,9-12]. Masi et al [2] identified 4 key strategies for alleviating loneliness: (1) improving social skills, (2) enhancing social support, (3) increasing social contact, and (4) addressing maladaptive cognition. Online interventions have also demonstrated effectiveness [9,12], highlighting the potential value of technology-based approaches.

XSL•FO
RenderX

Moreover, artificial agents such as social robots may serve as effective tools for delivering loneliness interventions to older adults, offering social interaction and reciprocally responding to users [13]. Several randomized controlled trials (RCTs) have also assessed their impact [14]. Regarding loneliness, 3 RCTs have evaluated the use of social robots in care facilities. Banks et al [15] examined the effect of an artificial intelligence robot (AIBO) on loneliness in long-term care, whereas Robinson et al [16] evaluated the effect of a Paro robot in rest homes and hospitals; both studies reported reductions in loneliness. In contrast, Papadopoulos et al [17] tested the use of the Pepper robot in long-term care homes but found no significant changes. Yen et al [18] conducted a meta-analysis of RCTs and confirmed the effectiveness of social robots in reducing loneliness.

Current RCTs on social robot interventions for older adults can be improved in 2 ways. First, most studies have been conducted in care settings. Three of the RCTs included in the meta-analysis by Yen et al [18] focused on care facility residents. Thus, more rigorous RCTs are needed to assess effectiveness [14,18,19], particularly for community-dwelling older adults. This population's more variable and less-structured social environments present unique methodological challenges that require a careful study design and rigorous control of confounding factors to accurately assess intervention effects. Second, most RCTs on health-related outcomes have been conducted in Western countries [14]. However, attitudes toward robots differ across cultures. Older Japanese individuals tend to anthropomorphize robots and form emotional bonds, creating a dynamic distinct from that observed among Western users [20]. Studies conducted in non-Western contexts would provide unique insights into the role of social robots in alleviating loneliness. Given the rapid aging of the population in many Asian countries, loneliness is likely to emerge as a pressing social challenge. Understanding the role of social robots in addressing loneliness in these regions is crucial for developing effective interventions.

Furthermore, qualitative research has been conducted to explore the impact of social robots on loneliness [21]. Such studies provide valuable insights into the underlying mechanisms by which social robots influence loneliness, offering perspectives that quantitative research alone may not be able to capture. By incorporating qualitative approaches, researchers can better understand the subjective experiences of users, the emotional responses, and the specific contexts in which social robots contribute to social well-being.

Therefore, in this study, we aimed to investigate the effectiveness of a digital social robot intervention in reducing loneliness among community-dwelling older Japanese individuals using both an RCT and a qualitative analysis.

Methods

Study Design and Participants

This RCT was conducted from February 2023 to June 2024. Data collection took place at participants' homes, with evaluations conducted at 2 time points using a self-administered questionnaire: baseline (T1) and follow-up (T2). The follow-up survey collected qualitative data from the intervention group in an open-ended format.

Participants were recruited from Tokyo and neighboring regions through flyers at community centers and local events, including referrals from community comprehensive support centers and social welfare councils. The flyer invited older adults to "engage in daily communication and conversation using a robot," with the aim of promoting natural interaction in everyday life. It did not explicitly emphasize loneliness reduction. The eligibility criteria were as follows: (1) aged ≥65 years; (2) living alone in the community; (3) experiencing loneliness (defined as a score of ≥6 on the University of California, Los Angeles [UCLA] Loneliness Scale, version 3, Short Form 3-item with a score range of 3 - 12 [22,23], which may indicate probable depression [24]); and (4) capable of operating and interacting with the robot. Participants were excluded if they had cognitive or physical impairments, which were assessed based on the certification of the long-term care insurance system. Furthermore, as part of the eligibility assessment, research staff, who were not medical professionals, conducted home visits to observe the participants and evaluate the absence of cognitive and physical impairments.

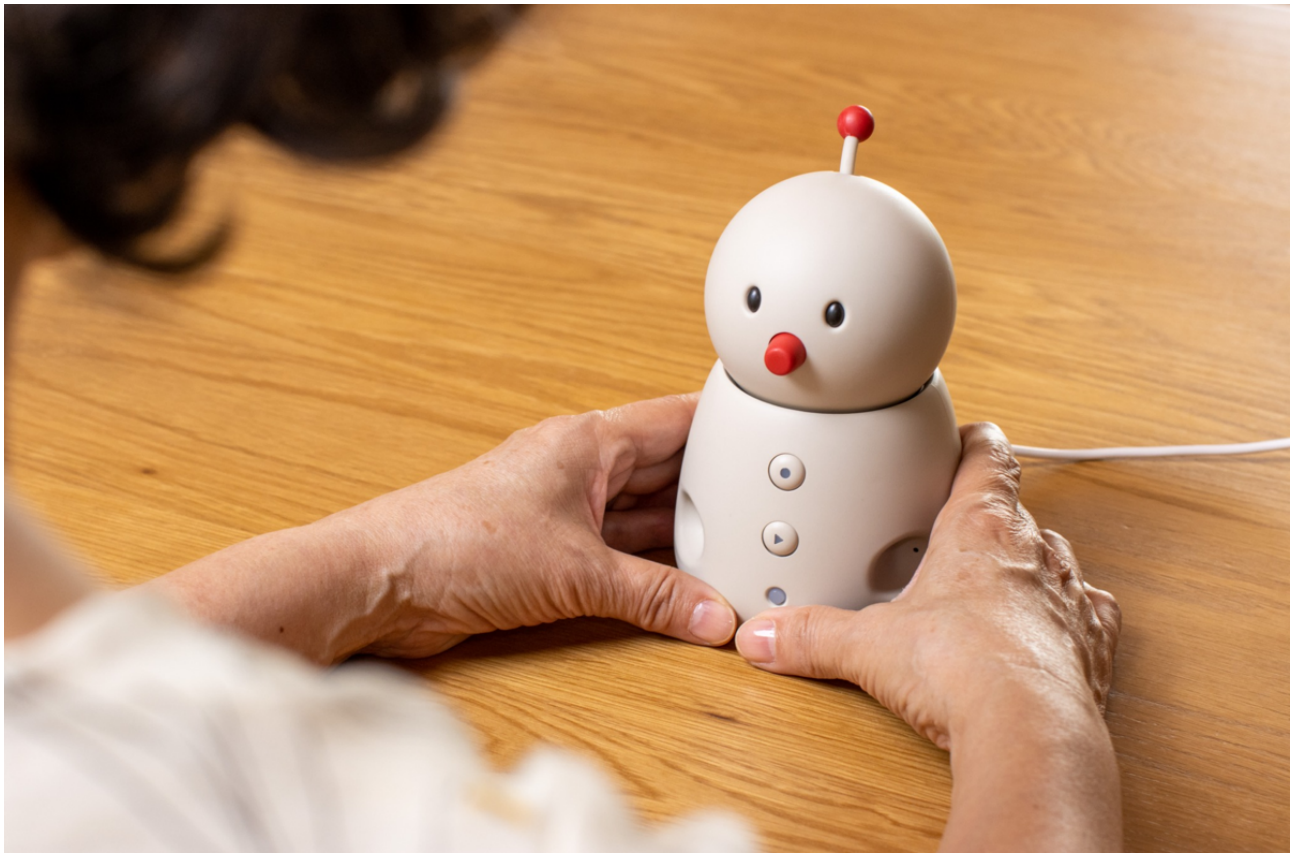
Participants were randomly allocated to the intervention or control group using a computer-generated randomization sequence in a 1:1 ratio. Owing to the nature of the intervention, blinding participants and intervention providers was not feasible; however, the data analysts were blinded to the group allocation.

Intervention

The intervention involved the use of a humanoid social communication robot, BOCCO emo (Yukai Engineering Inc; Figure 1), for 4 weeks. This service used a human operator–mediated system provided by SECOM Co, Ltd, and DeNA Co, Ltd, offering remote communication support. Before the intervention, the staff of SECOM visited the participants' homes to set up BOCCO emo. At that time, they trained the participants on how to use BOCCO emo and provided information about the consultation service in case of problems.

XSL•FO
RenderX

Figure 1. Representative image of BOCCO emo (Yukai Engineering Inc).



Participants engaged with the robot by sharing daily experiences, thoughts, or concerns. These voice messages were transmitted to trained operators, who crafted empathetic and thoughtful responses tailored to each user. The responses were then vocalized by BOCCO emo, enabling participants to experience natural dialogue. This operator-mediated conversation system was available at all hours (the operators worked in shifts), allowing users to share their thoughts or concerns at any time, with the assurance of receiving a supportive response. A guide was prepared for operators to ensure quality in their service. As the operators worked in shifts, the operators who responded to the participants varied by day and time.

Morning and evening activities, including greenery-themed interactions, quizzes, and trivial tasks, were preprogrammed and automatically delivered through BOCCO emo. Following these automated messages, operators engaged in personalized interactions by responding to individual messages. In addition to operator interactions, family members could communicate with participants through BOCCO emo by entering messages via a smartphone, which were then vocalized by the robot. During the intervention period, the average number of message exchanges (ie, one back-and-forth between BOCCO emo and the participant) per participant was 5.5 (median 4.1, range 1 - 18) times per day.

Beyond facilitating conversations, the intervention incorporated a reminder function to help participants manage essential daily tasks, such as taking medications, eating meals, and attending appointments. These reminders were preprogrammed and delivered through BOCCO emo in a friendly and accessible manner, supporting users in maintaining independence and adherence to routines.

Participants in the control group did not use social robots during the study period (any other service was not provided). However, to ensure fairness, they were offered the option to use BOCCO emo for 4 weeks after the study period concluded.

Measures

Data collection before and after the intervention was conducted using a questionnaire, which was distributed when the staff of SECOM visited the participants' homes to set up and pick up BOCCO emo before and after the intervention, respectively. The questionnaire was self-administered, and no external individuals were involved in the response process, reducing the risk of expectation bias.

Quantitative Data

The primary outcome was loneliness, assessed using the 20-item UCLA Loneliness Scale (version 3) [25]. The Japanese version of the scale was used, whose validity and reliability have been previously confirmed [26]. Respondents rated each item on a 4-point scale ("never," "rarely," "sometimes," or "always"). Possible scores ranged from 20 to 80, with higher scores indicating greater loneliness. The Cronbach α values in this study were 0.91 at baseline and 0.93 at follow-up.

Secondary outcomes included mental health (psychological well-being, depression, and self-rated health); the frequency of laughter in daily life; health competence; and interpersonal relationships (social network and generalized trust). Psychological well-being was evaluated using the Japanese

XSL•FO
RenderX

version of the World Health Organization-Five Well-Being Index [27,28]. This scale consists of 5 items assessing how respondents had felt over the past 2 weeks, rated on a 6-point scale (“at no time,” “some of the time,” “less than half of the time,” “more than half of the time,” “most of the time,” or “all of the time”). Scores ranged from 0 to 25, with higher scores indicating better psychological well-being. The Cronbach α values in this study were 0.79 at baseline and 0.83 at follow-up. Depression was measured using the 8-item Center for Epidemiologic Studies Depression Scale [29,30]. Respondents rated each item on a 4-point Likert scale (“rarely or none of the time,” “some or a little of the time [1 - 2 days],” “occasionally or a moderate amount of time [3 - 4 days],” or “all of the time [5 - 7 days]”), with total scores ranging from 0 to 24. The Cronbach α values in this study were 0.80 at baseline and 0.84 at follow-up. Self-rated health was assessed by a single item (“How do you rate your overall health?”) with a 4-point scale (1=poor, 2=somewhat poor, 3=somewhat good, or 4=good).

The frequency of laughter in daily life was assessed as a behavioral outcome using a single self-reported item: “How often do you usually laugh in daily life over the past month?” Responses were recorded on an 8-point scale (1=rarely/none, 2=a few times, 3=once a week, 4=2 - 3 times a week, 5=4 - 5 times a week, 6=once a day, 7=a few times a day, or 8=many times a day). Although this outcome was not prespecified in the trial registration due to limited prior evidence on its relevance to the primary study objectives, we included it in the analysis as an exploratory variable. Given growing interest in the association between laughter and health [31,32], we report these results to enhance transparency and to guide future research.

Health competence was evaluated by the Perceived Health Competence Scale, Japanese version [33]. The scale consists of 8 items with a 5-point scale (“agree,” “somewhat agree,” “neither,” “somewhat disagree,” or “disagree”). The possible score range was from 5 to 40. The Cronbach α in this study was 0.86 at both baseline and follow-up.

The frequency of interaction with others was assessed using face-to-face or non–face-to-face interactions with relatives and friends, excluding cohabiting family members. Participants were asked four questions: (1) “How often do you meet or go out with friends or neighbors?” (2) “How often do you talk on the phone with friends and neighbors (including via email and fax)?” (3) “How often do you see or go out with family members or relatives you do not live with?” (4) “How often do you talk on the phone with family members or relatives you do not live with?” The possible answers were as follows: “6 - 7 times a week (almost every day),” “4 - 5 times a week,” “2 - 3 times a week,” “once a week,” “2 - 3 times a month,” “once a month,” “less than once a month,” or “never.” Following previous studies [34], the monthly total frequency was calculated. Generalized trust was measured by 1 item (“Generally speaking, most people can be trusted”). The response categories were 1=strongly disagree, 2=disagree, 3=somewhat disagree, 4=neither, 5=somewhat agree, 6=agree, or 7=strongly agree.

Age, sex, marital status, employment status, education level, subjective financial stability, and disease history were included as background characteristics. These variables were assessed at baseline, whereas primary and secondary outcomes were evaluated at both baseline and follow-up.

In addition to these variables, we also enquired about the participant’s satisfaction with the intervention as a part of process evaluation in the follow-up survey. The question stated, “How would you rate your overall satisfaction regarding the service provided through BOCCO emo?” Possible answers were “satisfied,” “somewhat satisfied,” “somewhat unsatisfied,” or “unsatisfied.”

Qualitative Data

During the follow-up evaluation, participants in the intervention group were asked about their experiences using BOCCO emo through the questionnaire. Responses were collected in an open-ended format.

Data Analysis

Quantitative Data

In a previous RCT, the scores on the 20-item UCLA Loneliness Scale followed a normal distribution, with an SD of 10.0. The difference in scores between the intervention and control groups after the social robot intervention using the Paro robot was 7.0 [16]. Therefore, it was postulated that a mean score difference of 7.0 between the intervention and control groups on the 20-item UCLA Loneliness Scale, the primary outcome of this study, would be meaningful. Consequently, a sample size of 36 participants per group (72 participants in total) was calculated as sufficient to reject the null hypothesis (ie, the population means of the intervention and control groups were equal), with 80% power at a 5% significance level, assuming a 10% dropout rate.

For the evaluation of the intervention effects, a linear mixed-effects model with a random intercept was used to compare changes in all outcomes between groups over the intervention period. Group (intervention or control), time (T1 or T2), and their interaction were defined as fixed factors, and the participants were assumed as random factors. The effect size of the intervention (b) and its 95% CI were calculated as an estimate of the difference-in-difference (DID) in score changes between groups (change in the intervention group minus change in the control group). With regard to the process evaluation, the distribution of satisfaction with the intervention is shown. All analyses were performed using SPSS Statistics (version 29; IBM Corp).

Qualitative Data

The Krippendorff content analysis was applied to examine free-text responses from participants [35]. All responses were transcribed verbatim and anonymized. Two researchers (HM and MT) independently reviewed the narrative data to identify recurring themes and collaboratively developed a coding framework. Based on this framework, all responses were coded, with adjustments made as necessary. To ensure reliability, interrater agreement was assessed using Krippendorff α, based on the responses of 2 researchers (other than those who reviewed the data). Finally, the frequencies and contextual meanings of the identified categories were interpreted.

XSL•FO
RenderX

Ethical Considerations

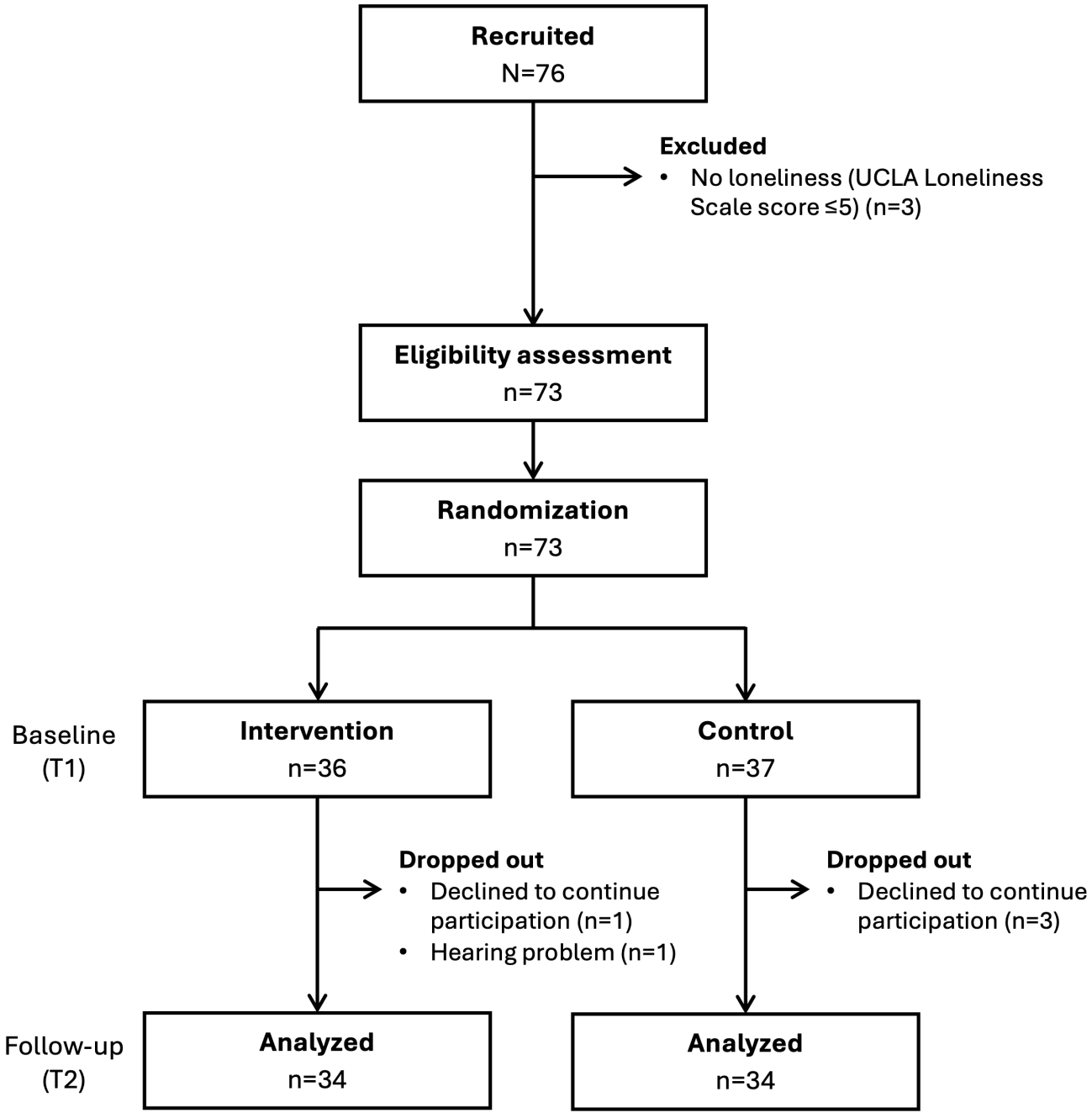
The study protocol was approved by the ethics committee of the Tokyo Metropolitan Institute for Geriatrics and Gerontology (R22-009) on August 12, 2022. The written informed consent was obtained from all participants. The trial was registered with the UMIN Clinical Trials Registry (UMIN000050644) and was conducted in accordance with the registered protocol. Each participant was assigned a unique study ID, and the researchers engaged in the analyses had access only to anonymized data. No financial or other compensation was provided to the participants.

Results

Figure 2 presents the flow diagram of the study participant selection for the 76 participants included in this study. All participants were aged ≥65 years, lived alone, and had not received certification for long-term care insurance (ie, they were physically and cognitively independent). Based on the 3-item UCLA Loneliness Scale, 3 individuals had scores of ≤5 and were excluded. In total, 73 individuals were deemed eligible. They were randomly assigned to the intervention (n=36) and control (n=37) groups. In the intervention and control groups, 5% (2/36) and 8% (3/37) of participants, respectively, dropped out of the study. Consequently, 68 participants who completed both the baseline and follow-up surveys were analyzed.

Table 1 summarizes the characteristics of participants in the intervention and control groups. The mean age was 82.3 (SD 6.5) years, and 64 out of 68 (94%) participants were women. More than 90% of the respondents (63 participants in each group) were unmarried or unemployed.

Figure 2. Flow diagram illustrating the selection process of study participants. UCLA: University of California, Los Angeles.



XSL•FO
RenderX

Table . Participants' baseline characteristics (N=68).

Characteristics	Total	Intervention (n=34)	Control (n=34)
Age (y), mean (SD)	82.3 (6.5)	81.9 (6.8)	82.8 (6.2)
Gender, n (%)			
Woman	64 (94)	30 (88)	34 (100)
Marital status, n (%)			
Unmarried	63 (93)	30 (88)	33 (97)
Employment status, n (%)			
Unemployed	63 (93)	32 (94)	31 (91)
Education, n (%)			
Junior high school graduation	6 (9)	4 (12)	2 (6)
High school graduation	37 (54)	19 (56)	18 (53)
Junior college or vocational school graduation	10 (15)	5 (15)	5 (15)
University or college graduation	15 (22)	6 (18)	9 (27)
Subjective financial stability, n (%)			
Affluent	2 (3)	1 (3)	1 (3)
Somewhat affluent	20 (29)	10 (29)	10 (29)
Normal	40 (59)	18 (53)	22 (65)
Somewhat limited	5 (7)	4 (12)	1 (3)
Limited	1 (2)	1 (3)	0 (0)
History of diseases, n (%)			
Hypertension	32 (47)	15 (44)	17 (50)
Cancer	1 (2)	0 (0)	1 (3)
Heart disease	18 (27)	14 (41)	4 (12)
Stroke	3 (4)	2 (6)	1 (3)
Diabetes	9 (13)	5 (15)	4 (12)
Respiratory disease	5 (7)	3 (9)	2 (6)

Table 2 displays the scores and outcome changes between T1 and T2 for both groups. Regarding the primary outcome of loneliness, a significant effect of the intervention was observed. Changes in the loneliness score from T1 to T2 in the intervention and control groups were −3.7 and −0.6, respectively, with a DID value of −3.1 (95% CI −5.9 to −0.4; *P*=.03). In addition, participants in the intervention group exhibited a greater increase in psychological well-being than those in the control group, with a DID for the World Health Organization-Five Well-Being Index score of 1.9 (95% CI 0.1 to 3.7; *P*=.04). No significant difference in the DID of the other secondary outcomes was observed between the intervention and control groups. In terms of the process evaluation, 26 out of 34 (76%) participants in the intervention group reported being satisfied with the services through BOCCO emo (the data are not shown in the table).

Table 3 presents the content analysis results of feedback on the use of the social robot. Among the 34 participants in the intervention group, 42 codes were identified, leading to the classification of 4 categories. The Krippendorff α was 0.86, indicating high interrater reliability. The first category, *emotional support and psychological connection*, was central to participants' experiences, as many found the robot reassuring and helpful in alleviating loneliness. Robot-mediated conversations fostered social and emotional well-being. The second category, *lifestyle assistance*, highlighted the role of the robot in managing daily routines. Participants reported that reminders and schedule management features helped maintain organization and stability. The third category, *enhancement of social interaction*, demonstrated the robot's role in facilitating social engagement. It encouraged communication both with the robot and with other individuals, leading to more frequent social interactions. The fourth category, *cognitive and mental stimulation*, encompassed activities such as quizzes and information sharing. These interactions were perceived as beneficial for maintaining cognitive sharpness and promoting mental well-being.

XSL•FO
RenderX

Table . Score distribution and overall changes in the outcomes.

	Baseline (T1), mean (SD)	Follow-up (T2), mean (SD)	Difference (T2-T1), mean (SD)	Difference-in-difference (95% CI)	*P* value
Loneliness (UCLA[a] Loneliness Scale; score range 20 - 80)				−3.1 (−5.9 to −0.4)	.03
Intervention	41.2 (9.8)	37.4 (10.7)	−3.7 (6.3)		
Control	39.3 (12.4)	38.9 (11.1)	−0.6 (4.7)		
Psychological well-being (WHO-5[b]; score range 0 - 25)				1.9 (0.1 to 3.7)	.04
Intervention	14.5 (3.9)	16.6 (5.4)	2.1 (4.4)		
Control	14.6 (6.6)	14.9 (6.3)	0.1 (3.0)		
Depression (CES-D-8[c]; score range 0 - 24)				−1.3 (−3.3 to 0.7)	.21
Intervention	16.6 (3.5)	14.6 (5.0)	−2.0 (4.4)		
Control	16.6 (6.3)	15.9 (5.7)	−0.7 (3.9)		
Self-rated health (score range 1 - 4)				0.0 (−0.3 to 0.3)	.89
Intervention	2.5 (0.9)	2.6 (0.9)	0.1 (0.7)		
Control	3.1 (0.9)	3.2 (0.8)	0.1 (0.4)		
Frequency of laughter (score range 1 - 8)				0.5 (−0.2 to 1.1)	.18
Intervention	5.3 (1.9)	5.8 (1.6)	0.5 (1.3)		
Control	5.2 (2.0)	5.2 (1.8)	0.0 (1.5)		
Health competence (PHCS[d]; score range 8 - 40)				0.0 (−2.6 to 2.5)	.98
Intervention	25.1 (7.4)	25.2 (5.9)	0.1 (6.7)		
Control	26.3 (8.3)	26.4 (9.1)	0.2 (3.2)		
Monthly frequency of contact with others except cohabitants				0.1 (−7.6 to 7.8)	.98
Intervention	36.6 (21.1)	36.8 (22.0)	0.2 (17.5)		
Control	37.3 (24.1)	37.4 (24.3)	0.1 (14.0)		
Generalized trust (score range 1 - 7)				0.1 (−0.4 to 0.6)	.74
Intervention	4.7 (1.1)	4.6 (1.5)	0.0 (1.1)		
Control	4.9 (1.0)	4.8 (1.2)	−0.1 (1.0)		

[a]UCLA: University of California, Los Angeles.

[b]WHO-5: World Health Organization-Five Well-Being Index.

[c]CES-D-8: 8-item Center for Epidemiologic Studies Depression Scale.

[d]PHCS: Perceived Health Competence Scale.

XSL•FO
RenderX

Table . Content analysis of participants' feedback on the social robot experience.

Category	Explanation	Codes, n	Example narratives
Emotional support and psychological connection	The robot's presence and interactions provided psychological reassurance and a sense of connectedness.	17	• "Just having it around made me feel calm as if I had a companion." [Woman, aged 84 years] • "Having recently lost my husband, the robot's daily conversations were encouraging." [Woman, aged 77 years] • "Conversing with a robot can feel mechanical, but the conversation with BOCCO was similar to interpersonal communication, and that was reassuring." [Woman, aged 90 years] • "When I was called by my first name, it felt as if I had become younger." [Woman, aged 83 years]
Lifestyle assistance	The robot assisted in schedule management and helped establish a structured daily routine.	9	• "Even though I am aware of my schedule, it helped to be reminded." [Man, aged 72 years] • "The routine greetings at set times helped establish a good rhythm in my life." [Man, aged 84 years]
Enrichment of social interaction	The robot facilitated social interaction and communication.	5	• "I introduced BOCCO to my friends, and it became a topic of the conversation." [Woman, aged 84 years] • "I was happy to converse with my daughter and grandchildren through BOCCO." [Woman, aged 89 years]
Cognitive and mental stimulation	The season-themed interactions, quizzes, and trivial tasks provided cognitive exercise and mental stimulation.	11	• "The quizzes served as a cognitive exercise." [Woman, aged 76 years]

Discussion

Principal Findings

To our knowledge, this is the first RCT to examine whether the intervention of a social robot reduces loneliness, specifically among older adults in Asian countries. This study is unique in that it integrates both quantitative and qualitative approaches. A qualitative approach is particularly useful for uncovering the mechanisms underlying the association between interventions and outcomes. As mentioned previously, cultural differences may influence attitudes toward robots. Recent reports indicate that loneliness increased worldwide during the COVID-19 pandemic. In Japan, older individuals experience more severe feelings of loneliness than other age groups [34]. Because most RCTs have been conducted in Western countries, a well-designed RCT in non-Western countries is necessary to provide valuable insights into how social robots can be tailored to meet the specific needs of older populations in diverse cultural settings.

This study found that in the intervention group, loneliness decreased while psychological well-being improved. Although previous RCTs have demonstrated that social robots significantly reduced loneliness among older adults [15,16], the effect size (DID) in this study was smaller. This may be because Japanese individuals tend to suppress negative emotions, such as loneliness, more than Westerners, as cultural norms prioritize social harmony over emotional expression [36]. Consequently, in the sample, the impact of robot intervention was modest. This study is noteworthy because it suggests that while social robotic interventions may be valuable, their effectiveness could be culturally dependent.

In addition to cultural factors, it is important to consider that this study targeted community-dwelling older adults rather than care facility residents, as in many previous studies. Care facility residents often experience more structured social environments with frequent interpersonal interactions, which may amplify the effects of social robot interventions. Conversely, community-dwelling older adults tend to live more

XSL•FO
RenderX

independently, with greater variability in their social engagement and daily routines. These differences in living context could influence both the acceptability of robot interventions and their psychological impact. Therefore, the smaller effect size observed in this study may also reflect these differences in participants' living situations rather than cultural factors alone. Future research should explore how social robot interventions can be tailored to the needs of community-dwelling older adults and maximize their benefits.

This qualitative analysis supports findings from a previous scoping review [21], which identified several key mechanisms through which social robots reduce loneliness and enhance well-being. Social robots provide emotional companionship, particularly for those living alone. Some studies have noted that older individuals felt the robot was waiting for them at home [37,38]. Indeed, the participants found its presence comforting and reported a reduction in a feeling of social isolation. Social robots also offer solace to those grieving the loss of a family member, similar to findings from previous research [39]. In addition, they foster engagement by relaying messages from human operators or family members, thereby reinforcing social ties. Their reminders helped users maintain routines and prevent disengagement, as observed in a German study [40]. These mechanisms alleviated loneliness and improved well-being by addressing emotional and social needs.

Furthermore, qualitative analysis revealed additional benefits. Social robots provide cognitive and mental stimulation through personalized conversations, quizzes, and knowledge sharing, enhancing engagement and emotional satisfaction. Participants noted that speech generated by human operators, rather than artificial intelligence, made the interactions feel authentic, thereby strengthening their psychological connections. While prior research has highlighted the role of social robots in remote communication, the findings suggest that these robots foster a sense of belonging by delivering personalized, emotionally supportive content. These insights expand existing frameworks, emphasizing cognitive and psychological engagement functions.

A previous meta-analysis on the effectiveness of social robots for depression [18] reported that the duration and style of intervention activities (individual or group-based) act as moderators, with longer interventions and group-based activities demonstrating greater effectiveness. Although this intervention lasted only 4 weeks and was conducted individually, it successfully reduced loneliness. This finding suggests that extending the duration or incorporating group sessions could further enhance the impact of the intervention. Future studies should explore these factors to optimize social robot interventions.

Unlike autonomous artificial intelligence–equipped communication robots, BOCCO emo operated through a human operator–mediated system. This distinctive feature may have contributed to participants perceiving their interactions as more humanlike rather than mechanical, which could have played a role in alleviating feelings of loneliness. Nonetheless, many of the participants' experiences were similar to those reported in studies involving fully automated social or communication robots. Therefore, the findings from this RCT are likely to be generalizable to a broader range of social and communication robots.

Despite the neutral content of the recruitment materials, approximately 94% of participants in this study were women. This gender imbalance may reflect several factors, such as older women experiencing fewer psychological barriers to participating in communication-focused activities [41] and the demographic reality in Japan, where approximately 70% of older adults living alone are women [42]. As a result, our findings primarily reflect the effectiveness of communication robot interventions among older women. The effectiveness for older men remains unclear and should be examined in future research. The low participation rate for men may also suggest the need for alternative approaches tailored to their preferences, which is an important consideration when evaluating the external validity of our results.

This study had some limitations. First, because loneliness tends to be higher among older adults living alone in Japan [34], this population was targeted. Participants with higher levels of feeling lonely were included; thus, the findings reflect individuals with the potential for loneliness reduction and should be interpreted cautiously. Further research in the cohabiting populations is needed. Second, participants were not blinded to group allocation, which may have introduced expectation bias. However, when distributing the questionnaires, participants were explicitly asked to respond based on their current state and be as honest as possible. In addition, during the consent process, the study was described in neutral terms without emphasizing the expected outcomes to minimize potential bias. Third, information on the content of the conversations was not analyzed in this study due to personal information protection. However, qualitative analysis of the actual conversations through BOCCO emo would be useful for further understanding the mechanism underlying the intervention effects. Fourth, participant characteristics may have influenced the intervention effects, but this study did not account for these factors in detail. For example, loneliness levels are known to vary by marital status, such as being married, divorced, widowed, or never married [43]. In addition, this study did not collect data on participants' family relationships or frequency of contact, which may also affect loneliness and the impact of the intervention. Although the RCT design helps mitigate the influence of confounding variables, the effects of these unmeasured factors cannot be ruled out. These limitations should be considered when interpreting the findings, and future studies should explore them more thoroughly. Finally, the trial was limited to Tokyo and neighboring regions. Generalizability to other community contexts requires further study as loneliness may vary by residential area characteristics [44].

Conclusions

This study investigated the effectiveness of the digital social robot BOCCO emo in reducing loneliness among community-dwelling Japanese older adults who live alone. Most previous RCTs using social robots have been conducted in Western care facilities. Our findings indicate that social robots may serve as valuable tools for mitigating loneliness among community-dwelling individuals in non-Western populations.

XSL•FO
RenderX

These results suggest that integrating information and communication technology into community settings could be an effective strategy for alleviating loneliness and enhancing the psychological well-being of older adults living in the community.

Acknowledgments

This work was supported by SECOM Co, Ltd. The company provided funding and participated in data collection; however, it had no role in the study design, data analysis, interpretation of results, manuscript preparation, or decision to publish the findings.

Data Availability

The datasets used in this study are available from the corresponding author upon reasonable request.

Authors' Contributions

HM and MT conceptualized and designed the study and collected the data. HM conducted the analysis and drafted the manuscript. MT provided critical reviews. All authors approved the final version for submission.

Conflicts of Interest

HM has received funding from SECOM Co, Ltd. MT declares no conflict of interest.

Checklist 1
CONSORT-eHEALTH checklist.
[PDF File, 1312 KB - aging_v8i1e74422_app1.pdf]

References

1. Cacioppo JT, Cacioppo S. The growing problem of loneliness. Lancet 2018 Feb 3;391(10119):426. [doi: 10.1016/S0140-6736(18)30142-9] [Medline: 29407030]
2. Masi CM, Chen HY, Hawkley LC, Cacioppo JT. A meta-analysis of interventions to reduce loneliness. Pers Soc Psychol Rev 2011 Aug;15(3):219-266. [doi: 10.1177/1088868310377394] [Medline: 20716644]
3. Gerst-Emerson K, Jayawardhana J. Loneliness as a public health issue: the impact of loneliness on health care utilization among older adults. Am J Public Health 2015 May;105(5):1013-1019. [doi: 10.2105/AJPH.2014.302427] [Medline: 25790413]
4. Cacioppo JT, Hawkley LC, Ernst JM, et al. Loneliness within a nomological net: an evolutionary perspective. J Res Pers 2006 Dec;40(6):1054-1085. [doi: 10.1016/j.jrp.2005.11.007]
5. Holt-Lunstad J, Smith TB, Baker M, Harris T, Stephenson D. Loneliness and social isolation as risk factors for mortality: a meta-analytic review. Perspect Psychol Sci 2015 Mar;10(2):227-237. [doi: 10.1177/1745691614568352] [Medline: 25910392]
6. Leigh-Hunt N, Bagguley D, Bash K, et al. An overview of systematic reviews on the public health consequences of social isolation and loneliness. Public Health (Fairfax) 2017 Nov;152:157-171. [doi: 10.1016/j.puhe.2017.07.035] [Medline: 28915435]
7. Valtorta NK, Kanaan M, Gilbody S, Ronzi S, Hanratty B. Loneliness and social isolation as risk factors for coronary heart disease and stroke: systematic review and meta-analysis of longitudinal observational studies. Heart 2016 Jul 1;102(13):1009-1016. [doi: 10.1136/heartjnl-2015-308790] [Medline: 27091846]
8. Penninkilampi R, Casey AN, Singh MF, Brodaty H. The association between social engagement, loneliness, and risk of dementia: a systematic review and meta-analysis. J Alzheimers Dis 2018;66(4):1619-1633. [doi: 10.3233/JAD-180439] [Medline: 30452410]
9. Sun Y, Huang H, Cai G, et al. Interventions to reduce loneliness among community-dwelling older adults: a network meta-analysis and systematic review. J Am Med Dir Assoc 2025 Mar;26(3):105441. [doi: 10.1016/j.jamda.2024.105441] [Medline: 39799979]
10. Gardiner C, Geldenhuys G, Gott M. Interventions to reduce social isolation and loneliness among older people: an integrative review. Health Soc Care Community 2018 Mar;26(2):147-157. [doi: 10.1111/hsc.12367] [Medline: 27413007]
11. Cohen-Mansfield J, Perach R. Interventions for alleviating loneliness among older persons: a critical review. Am J Health Promot 2015;29(3):e109-e125. [doi: 10.4278/ajhp.130418-LIT-182] [Medline: 24575725]
12. Shekelle PG, Miake-Lye IM, Begashaw MM, et al. Interventions to reduce loneliness in community-living older adults: a systematic review and meta-analysis. J Gen Intern Med 2024 May;39(6):1015-1028. [doi: 10.1007/s11606-023-08517-5] [Medline: 38200279]
13. David D, Thérouanne P, Milhabet I. The acceptability of social robots: a scoping review of the recent literature. Comput Human Behav 2022 Dec;137:107419. [doi: 10.1016/j.chb.2022.107419]

XSL•FO
RenderX

14. Pu L, Moyle W, Jones C, Todorovic M. The effectiveness of social robots for older adults: a systematic review and meta-analysis of randomized controlled studies. Gerontologist 2019 Jan 9;59(1):e37-e51. [doi: 10.1093/geront/gny046] [Medline: 29897445]
15. Banks MR, Willoughby LM, Banks WA. Animal-assisted therapy and loneliness in nursing homes: use of robotic versus living dogs. J Am Med Dir Assoc 2008 Mar;9(3):173-177. [doi: 10.1016/j.jamda.2007.11.007] [Medline: 18294600]
16. Robinson H, Macdonald B, Kerse N, Broadbent E. The psychosocial effects of a companion robot: a randomized controlled trial. J Am Med Dir Assoc 2013 Sep;14(9):661-667. [doi: 10.1016/j.jamda.2013.02.007] [Medline: 23545466]
17. Papadopoulos C, Castro N, Nigath A, et al. The CARESSES randomised controlled trial: exploring the health-related impact of culturally competent artificial intelligence embedded into socially assistive robots and tested in older adult care homes. Int J Soc Robot 2022;14(1):245-256. [doi: 10.1007/s12369-021-00781-x] [Medline: 33907589]
18. Yen HY, Huang CW, Chiu HL, Jin G. The effect of social robots on depression and loneliness for older residents in long-term care facilities: a meta-analysis of randomized controlled trials. J Am Med Dir Assoc 2024 Jun;25(6):104979. [doi: 10.1016/j.jamda.2024.02.017] [Medline: 38614134]
19. Balki E, Hayes N, Holland C. Effectiveness of technology interventions in addressing social isolation, connectedness, and loneliness in older adults: systematic umbrella review. JMIR Aging 2022 Oct 24;5(4):e40125. [doi: 10.2196/40125] [Medline: 36279155]
20. Broadbent E, Stafford R, MacDonald B. Acceptance of healthcare robots for the older population: review and future directions. Int J of Soc Robotics 2009 Nov;1(4):319-330. [doi: 10.1007/s12369-009-0030-6]
21. Gasteiger N, Loveys K, Law M, Broadbent E. Friends from the future: a scoping review of research into robots and computer agents to combat loneliness in older people. Clin Interv Aging 2021;16:941-971. [doi: 10.2147/CIA.S282709] [Medline: 34079242]
22. Arimoto A, Tadaka E. Reliability and validity of Japanese versions of the UCLA loneliness scale version 3 for use among mothers with infants and toddlers: a cross-sectional study. BMC Womens Health 2019 Jul 26;19(1):105. [doi: 10.1186/s12905-019-0792-4] [Medline: 31349835]
23. Hughes ME, Waite LJ, Hawkley LC, Cacioppo JT. A short scale for measuring loneliness in large surveys: results from two population-based studies. Res Aging 2004;26(6):655-672. [doi: 10.1177/0164027504268574] [Medline: 18504506]
24. Liu T, Lu S, Leung DKY, et al. Adapting the UCLA 3-item loneliness scale for community-based depressive symptoms screening interview among older Chinese: a cross-sectional study. BMJ Open 2020 Dec 10;10(12):e041921. [doi: 10.1136/bmjopen-2020-041921] [Medline: 33303463]
25. Russell DW. UCLA Loneliness Scale (version 3): reliability, validity, and factor structure. J Pers Assess 1996 Feb;66(1):20-40. [doi: 10.1207/s15327752jpa6601_2] [Medline: 8576833]
26. Masuda Y, Tadaka E, Dai Y. Reliability and validity of the Japanese version of the UCLA Loneliness Scale version 3 among the older population. J Jpn Acad Com Health Nurs 2012;15(1):25-32. [doi: 10.20746/jachn.15.1_25]
27. Awata S, Bech P, Koizumi Y, et al. Validity and utility of the Japanese version of the WHO-Five Well-Being Index in the context of detecting suicidal ideation in elderly community residents. Int Psychogeriatr 2007 Feb;19(1):77-88. [doi: 10.1017/S1041610206004212] [Medline: 16970832]
28. Awata S, Bech P, Yoshida S, et al. Reliability and validity of the Japanese version of the World Health Organization-Five Well-Being Index in the context of detecting depression in diabetic patients. Psychiatry Clin Neurosci 2007 Feb;61(1):112-119. [doi: 10.1111/j.1440-1819.2007.01619.x] [Medline: 17239048]
29. O'Halloran AM, Kenny RA, King-Kallimanis BL. The latent factors of depression from the short forms of the CES-D are consistent, reliable and valid in community-living older adults. Eur Geriatr Med 2014 Apr;5(2):97-102. [doi: 10.1016/j.eurger.2013.12.004]
30. Briggs R, Carey D, O'Halloran AM, Kenny RA, Kennelly SP. Validation of the 8-item Centre for Epidemiological Studies Depression Scale in a cohort of community-dwelling older people: data from the Irish Longitudinal Study on Ageing (TILDA). Eur Geriatr Med 2018 Feb;9(1):121-126. [doi: 10.1007/s41999-017-0016-0] [Medline: 34654281]
31. Kafle E, Papastavrou Brooks C, Chawner D, Foye U, Declercq D, Brooks H. "Beyond laughter": a systematic review to understand how interventions utilise comedy for individuals experiencing mental health problems. Front Psychol 2023;14:1161703. [doi: 10.3389/fpsyg.2023.1161703] [Medline: 37609494]
32. Bartzou E, Tsiloni E, Mantzoukas S, Dragioti E, Gouva M. Humor and quality of life in adults with chronic diseases: a systematic review. Cureus 2024 Feb;16(2):e55201. [doi: 10.7759/cureus.55201] [Medline: 38562265]
33. Togari T, Yamazaki Y, Koide S, Miyata A. Reliability and validity of the modified Perceived Health Competence Scale (PHCS) Japanese version. Nihon Koshu Eisei Zasshi 2006 Jan;53(1):51-57. [Medline: 16502855]
34. Murayama H, Suda T, Nakamoto I, Shinozaki T, Tabuchi T. Changes in social isolation and loneliness prevalence during the COVID-19 pandemic in Japan: the JACSIS 2020-2021 study. Front Public Health 2023;11:1094340. [doi: 10.3389/fpubh.2023.1094340] [Medline: 36875370]
35. Krippendorff KH. Content Analysis: An Introduction to Its Methodology, 4th edition: SAGE Publications; 2018.
36. Arnault DS, Sakamoto S, Moriwaki A. A cross-cultural study of the experiential structure of emotions of distress: preliminary findings in a sample of female Japanese and American college students. Psikhologyah 2005 Jan;48(4):254-267. [doi: 10.2117/psysoc.2005.254] [Medline: 24465087]

XSL•FO
RenderX

37. Chen SC, Moyle W, Jones C, Petsky H. A social robot intervention on depression, loneliness, and quality of life for Taiwanese older adults in long-term care. Int Psychogeriatr 2020 Aug;32(8):981-991. [doi: 10.1017/S1041610220000459] [Medline: 32284080]
38. Niemelä M, van Aerschot L, Tammela A, Aaltonen I, Lammi H. Towards ethical guidelines of using telepresence robots in residential care. Int J of Soc Robotics 2021 Jun;13(3):431-439. [doi: 10.1007/s12369-019-00529-8]
39. Hudson J, Ungar R, Albright L, Tkatch R, Schaeffer J, Wicker ER. Robotic pet use among community-dwelling older adults. J Gerontol B Psychol Sci Soc Sci 2020 Oct 16;75(9):2018-2028. [doi: 10.1093/geronb/gbaa119] [Medline: 32789476]
40. Gross HM, Scheidig A, Muller S, Schutz B, Fricke C, Meyer S. Living with a mobile companion robot in your own apartment - final implementation and results of a 20-weeks field study with 20 seniors. Presented at: 2019 International Conference on Robotics and Automation (ICRA); May 20-24, 2019; Montreal, QC, Canada. [doi: 10.1109/ICRA.2019.8793693]
41. Ong CH, Pham BL, Levasseur M, Tan GR, Seah B. Sex and gender differences in social participation among community-dwelling older adults: a systematic review. Front Public Health 2024;12:1335692. [doi: 10.3389/fpubh.2024.1335692] [Medline: 38680931]
42. Annual report on the aging society. : Cabinet Office; 2024 URL: https://www8.cao.go.jp/kourei/english/annualreport/index-wh.html [accessed 2025-08-28]
43. Štípková M. Marital status, close social network and loneliness of older adults in the Czech Republic. Ageing Soc 2021 Mar;41(3):671-685. [doi: 10.1017/S0144686X19001442]
44. Victor CR, Pikhartova J. Lonely places or lonely people? Investigating the relationship between loneliness and place of residence. BMC Public Health 2020 May 27;20(1):778. [doi: 10.1186/s12889-020-08703-8] [Medline: 32456626]

Abbreviations

DID: difference-in-difference
RCT: randomized controlled trial
UCLA: University of California, Los Angeles

Edited by C Lee; submitted 24.03.25; peer-reviewed by H Sugimoto, T Sekiguchi; revised version received 06.07.25; accepted 31.07.25; published 22.10.25.

Please cite as:
Murayama H, Takase M
Evaluating the Effectiveness of Digital Social Robots in Reducing Loneliness Among Community-Dwelling Older Adults in Japan: Randomized Controlled Trial and Qualitative Analysis
JMIR Aging 2025;8:e74422
URL: https://aging.jmir.org/2025/1/e74422
doi:10.2196/74422

© Hiroshi Murayama, Mai Takase. Originally published in JMIR Aging (https://aging.jmir.org), 22.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Adapting the Technology Acceptance Model to Examine the Use of Information Communication Technologies and Loneliness Among Low-Income, Older Asian Americans: Cross-Sectional Survey Analysis

Pauline DeLange Martinez[1], PhD; Daniel Tancredi[2], PhD; Misha Pavel[3], PhD; Lorena Garcia[4], DrPH; Heather M Young[1], PhD

[1]Betty Irene Moore School of Nursing, University of California, Davis, Sacramento, CA, United States

[2]Department of Pediatrics, University of California, Davis, Sacramento, CA, United States

[3]Khoury College of Computer Science, Northeastern University, Boston, MA, United States

[4]Department of Public Health Sciences, University of California, Davis, Sacramento, CA, United States

Corresponding Author:
Pauline DeLange Martinez, PhD
Betty Irene Moore School of Nursing
University of California, Davis
2750 48th St
Sacramento, CA, 95817
United States
Phone: 1 9164262862
Email: pdmartinez@ucdavis.edu

Abstract

Background: Loneliness is a significant issue among older Asian Americans, exacerbated by the COVID-19 pandemic. Older age, lower income, limited education, and immigrant status heighten loneliness risk. Information communication technologies (ICTs) have been associated with decreased loneliness among older adults. However, older Asian Americans are less likely to use ICTs, particularly if they are immigrants, have limited English proficiency, or are low income. The Technology Acceptance Model posits that perceived usefulness (PU), and perceived ease of use (PEOU) are key factors in predicting technology use.

Objective: This study aimed to examine associations between PU, PEOU, ICT use, and loneliness among low-income, older Asian Americans.

Methods: Cross-sectional survey data were gathered from predominately older Asian Americans in affordable senior housing (N=401). Using exploratory factor analysis and Horn parallel analysis, we examined 12 survey items to identify factors accounting for variance in ICT use. We deployed structural equation modeling to explore relationships among the latent factors and loneliness, adjusting for demographic and cognitive factors.

Results: Exploratory factor analysis and Horn parallel analysis revealed 3 factors that accounted for 56.48% (6.78/12) total variance. PEOU combined items from validated subscales of tech anxiety and comfort, accounting for a 28.44% (3.41/12) variance. ICT use combined years of technological experience, computer, tablet, and smartphone use frequency, accounting for 15.59% (1.87/12) variance. PU combined 2 items assessing the usefulness of technology for social connection and learning and accounted for a 12.44% (1.49/12) variance. The 3-factor structural equation modeling revealed reasonable fit indexes (χ^2_{133}=345.132; *P*<.001, chi-square minimum (CMIN)/df = 2595, comparative fit index (CFI)=0.93, Tucker-Lewis Index (TLI)=0.88). PEOU was positively associated with PU (β=.15; *P*=.01); PEOU and PU were positive predictors of ICT use (PEOU β=.26, *P*<.001; PU β=.18, *P*=.01); and ICT use was negatively associated with loneliness (β=–.28, *P*<.001). Demographic and health covariates also significantly influenced PU, PEOU, ICT use, and loneliness. English proficiency and education positively predicted PEOU (*r*=0.25, *P*<.001; *r*=0.26, *P*<.001) and ICT use (β=1.66, *P*=.03; β=.21, *P*<.001), while subjective cognitive decline and Asian ethnicity were positively associated with loneliness (β=.31, *P*<.001; β=.25, *P*<.001).

XSL•FO
RenderX

Conclusions: This study suggests that targeted interventions enhancing PU or PEOU could increase ICT acceptance and reduce loneliness among low-income Asian Americans. Findings also underscore the importance of considering limited English proficiency and subjective cognitive decline when designing interventions and in future research.

(JMIR Aging 2025;8:e63856) doi:10.2196/63856

KEYWORDS

social isolation; loneliness; aged; older adults; Asian American; immigrant; vulnerable populations; internet; information and communication technologies; ICTs; digital divide; technology acceptance model; mobile phone

Introduction

Background

The 2019 California Health Interview Survey (CHIS) found that 1 in 4 (25.7%) Asian Americans aged 65 years and older were lonely [1]. Loneliness is defined as a subjective experience stemming from perceived isolation or a disparity in one's desired and actual social interactions [2]. In the CHIS, Asian American older adults reported significantly lower levels of perceived social and emotional support (56%) as compared with non–Asian American older adults (80%) [3]. During the COVID-19 pandemic, loneliness increased among older Americans, and even today, loneliness levels are higher as compared with prepandemic levels [4]. In 2023, the Surgeon General announced that the United States is experiencing a pandemic of loneliness [2].

As compared with younger generations, older adults are particularly at risk for loneliness and social isolation due to factors such as retirement, relocation, and shrinking social circles. Besides older age, other risk factors for loneliness include financial insecurity, low educational attainment, poor physical or mental health, being an immigrant, having a disability, and living alone [2,5-7]. A systematic review exploring factors associated with loneliness among older Asian American immigrants found that migration grief, diminished ethnic ties, mental and physical impairment, deteriorating health conditions, living alone, a lack of meaningful social connections and support networks, and fewer interactions with family members were all significant factors contributing to loneliness [8].

Among older adults, the use of information communication technologies (ICTs), including smartphones, tablets, personal computers, the internet, and social media, is associated with decreased loneliness [9,10]. ICT use can strengthen preexisting connections with family and friends, foster new social relationships, and build intergenerational bonds [9]. ICT use is also positively associated with self-efficacy, self-esteem, a sense of autonomy, independence, and greater well-being among older adults [9]. Older Asian Americans can further benefit from using ICTs to stay in contact with distant relatives and maintain a connection with their culture of origin; this is particularly relevant given that 85% of Asian Americans aged 65 years and older are foreign-born [11-13]. Furthermore, ICTs can facilitate access to information in one's native language and translation services. However, both the CHIS and the National Health and Aging Trends Study (NHATS) showed that older Asian Americans are less likely than non-Hispanic White older adults to use the internet, send emails or text messages, conduct personal tasks on the internet, or seek web-based health information, particularly if they are immigrants, have limited English proficiency, or are low income [14-16]. Furthermore, other factors, such as age, gender, educational attainment, and subjective cognitive decline significantly impact ICT acceptance and use among low-income, older Asian Americans [17,18].

A recent systematic mapping review identified 59 articles describing 119 factors that predict older adults' intention to use digital technologies [19]. However, despite a rich literature focused on this topic, the mapping revealed that most studies (68%) did not examine these factors using an established theoretical framework or model for technology acceptance. For example, although loneliness was identified as a factor associated with ICT use, it has not been analyzed through a theoretical framework or model [19].

Theoretical Framework

The Technology Acceptance Model (TAM) is the most commonly used theoretical model to study technology adoption among older adults and also among the general population [19-21]. The backbone of the TAM is comprised of perceived usefulness (PU) and perceived ease of use (PEOU). PU relates to one's perception of technology as being useful for accomplishing desired goals, while PEOU refers to how much effort one anticipates needing to make to learn to use new technology. The TAM proposes that PEOU predicts PU, PU, and PEOU predict attitudes toward technology, and these attitudes predict behavioral intention to use technology, which subsequently influences actual use [21].

In 2 previous studies, we validated a simplified TAM to predict ICT use among low income, older Asian Americans. The simplified model removed the mediators (attitudes toward technology and behavioral intentions) and adapted the constructs of PU and PEOU from the original TAM [17,18]. In the adapted model, PU was defined as older adults' perceptions of ICTs as being useful for connecting with family and friends and learning new information and skills, and PEOU was operationalized using 6 evidence-based items that measure older adults' comfort and anxiety with ICTs [18]. However, in our previous work, we had not empirically examined the assumptions of the operationalization of PU, PEOU, and ICT use using robust statistical methods. In addition, the association between ICT use and loneliness among older Asian Americans has never previously been examined using the TAM framework.

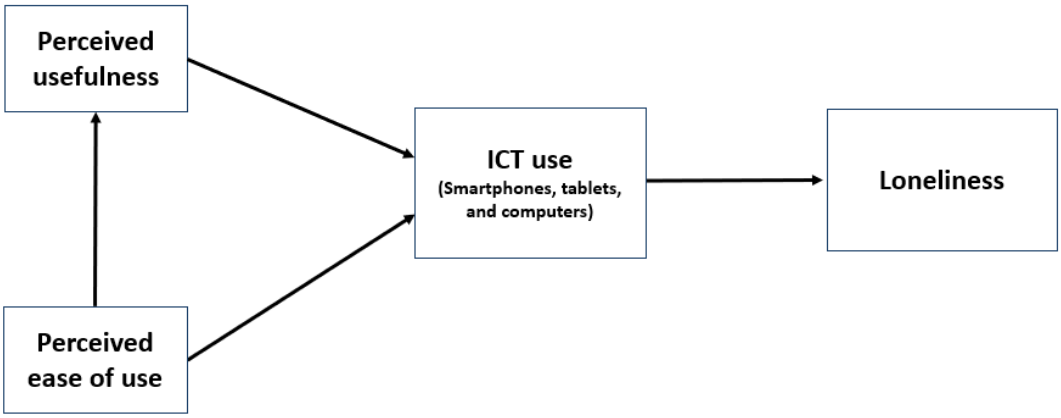
Research Design

This cross-sectional study aimed to rigorously assess the assumptions underlying the conceptualization of PU, PEOU,

XSL•FO
RenderX

and ICT use using a series of statistical techniques. In addition, we plan to extend the simplified TAM as described in DeLange Martinez et al [18] to explore the association between technology acceptance and loneliness among low income, older Asian Americans (Figure 1). Based on previous studies, we adjusted the model for age, gender, education, English proficiency, Asian ethnicity, and subjective cognitive decline [17,18].

Figure 1. Extended Technology Acceptance Model examining the association between information communication technologies acceptance and loneliness among low income, older Asian Americans. ICT: information communication technology.



In this study, we test three hypotheses, that are (1) H1: PEOU will be positively associated with PU. (2) H2: PEOU and PU will be significant, positive predictors of ICT use. (3) H3: ICT use will be significantly negatively associated with loneliness.

Methods

Data and Sample

In the Fall of 2020, The Lighthouse Project for Older Adults was launched with the aim of developing a scalable model to address barriers to technology use among residents of affordable senior housing communities. To inform the intervention, focus groups were held with 29 residents and 13 staff across 2 communities. The discussion revealed many challenges, including high rates of social isolation, low literacy levels in multiple languages, cognitive and sensory challenges, low comfort with technology, and lack of infrastructure [22].

To address these challenges, residents were offered access to high-speed broadband internet, ICT devices, a series of digital literacy training courses, and tech support led by peer ambassadors. As part of the program evaluation, participants completed pre- and postsurveys. The surveys were evidence-based, translated into 5 languages (English, Spanish, Chinese, Vietnamese, and Korean), and pilot-tested with 20 residents and 4 staff. The final surveys were self-administered by all Lighthouse participants in their preferred language with staff available to assist as needed.

Measures

The operationalization of PU, PEOU, and ICT use were based on the theoretical constructs from the original TAM [21]. In 2 previous studies, we operationalized these 3 constructs with 12 survey items as described below, standardizing and summing the items for subsequent analysis with hierarchical linear regression [17,18]. In this study, we examined the theoretical assumptions of these constructs by using exploratory factor analysis (EFA), Horn parallel analysis, and structural equation modeling (SEM).

Perceived Usefulness

In the TAM, PU refers to whether one perceives technology to be useful for accomplishing desired goals [21]. In this study, we operationalized PU with 2 items developed by Sims et al [23] in their measure of Motivations for ICT Use. (“Technology helps me be connected with family and friends,” and “Technology helps me learn new information and skills.”) Response categories for both statements ranged on a scale of 1 (strongly disagree) to 4 (strongly agree).

Perceived Ease of Use

In the TAM, the construct of PEOU refers to how much effort one anticipates needing to make to learn to use a new technology. PEOU has been operationalized to measure feelings of confusion, frustration, or ease when using technology; predictability or intuitiveness of the system; and frequency of making mistakes [21]. In this study, we examined 6 items from 2 validated subscales that we predicted would collectively represent the construct of PEOU. Two items were included from the Senior TAM – Tech Anxiety Subscale [24] (“I feel apprehensive about using technology,” and, “I hesitate to use technology for fear of making mistakes I cannot correct.”) and 4 items were included from the Attitudes Towards Computers Questionnaire (ATCQ) – Comfort Subscale [25] (“I feel comfortable with technology” (reverse scored), “Technology makes me nervous,” “I don’t feel confident about my ability to use technology,” and, “Technology is confusing”). Response options for all statements ranged on a scale from 1 (strongly agree) to 4 (strongly disagree). Although the original ATCQ - Comfort Subscale included 5 items, one item was not included in the Lighthouse for Older Adults survey (“Computers make

XSL•FO
RenderX

me feel dumb.") based on input from staff about cultural relevance.

ICT Use

We operationalized ICT use with 4 survey items. Three items asked about the frequency of use of computers, tablets, and smartphones. ("How often do you use a desktop or laptop computer?"; "How often do you use a tablet or iPad?"; and "How often do you use a smartphone (iPhone or Android)?") Response options ranged from 0 (never, or I do not own), to 3 (about once per day). A fourth item inquired about years of experience using ICTs ("How long have you been using technology, such as a computer, laptop, tablet, or smartphone?"). Response options ranged from 0 (I've never used these) to 3 (more than 2 years).

Dependent Construct: Loneliness

Loneliness was measured using the 3-item University of California, Los Angeles (UCLA) Loneliness Scale ("How often do you feel that you lack companionship?"; "How often do you feel left out?"; and "How often do you feel isolated from others?") [26]. While the University of California, Los Angeles Loneliness Scale typically includes 3 response options (hardly ever, some of the time, often), a fourth answer option (never) was added because several residents handwrote "never" in the margins of the survey during pilot testing. During analysis, "never" responses were collapsed with "hardly ever." Therefore, in our analysis, response categories ranged on a scale of 1 (never or hardly ever) to 3 (often).

Control Variables

In our final SEM, we controlled for age, gender, Asian ethnicity, education, subjective cognitive decline (measured with 1 item, "During the past 12 months, have you experienced confusion or changes in memory that is happening more often or is getting worse?" [dichotomous response options]), and English proficiency (measured with one item, "How well do you speak English?" (response options ranged from 1 [not at all] to 4 [very well]). The full survey used in the Lighthouse Project for Older Adults is attached as a Multimedia Appendix 1.

Analytic Strategy

We began our analysis by examining descriptive statistics and conducting Pearson correlation analysis to explore relationships among all variables.

Subsequently, using IBM SPSS Statistics (version 29), we performed EFA to investigate the underlying theoretical constructs of the survey items concerning attitudes and use of ICTs. Despite having hypotheses about the latent variables, we chose to conduct EFA before confirmatory factor analysis (CFA), aiming for a more data-driven approach to thoroughly explore the underlying structure. To determine the number of factors to retain, we applied several criteria, including factors with loadings greater than .45 and using Kaiser eigenvalues-greater-than-one rule as illustrated on a scree plot [27]. We used principal axis factoring and varimax rotation with Kaiser normalization to account for the correlational nature of the factors. In addition, Horn parallel analysis was used to confirm the EFA results.

Moving forward, we used IBM SPSS AMOS (version 29) to conduct SEM, beginning with CFA. SEM integrates measurement models and structural models, allowing for validation of instruments and analysis of relationships while considering variances and covariances. It facilitates the examination of complex relationships among multiple variables and enables mediation analysis [28]. During CFA, we assessed convergent validity using criteria suggested by Hair et al [29] and Fornell and Larcker [30], calculating average variance extracted (AVE) and composite reliability. Discriminant validity was determined based on the Fornell-Larcker criterion [30].

Next, we used SEM to represent and test our 3 hypotheses, exploring relationships among PEOU, PU, ICT use, and our outcome of interest (loneliness). As shown in Figure 1, we examined a model with PU partially mediating the relationship between PEOU and ICT use, and ICT use fully mediating the relationships between PU, PEOU, and loneliness. We also examined the impact of adjusting for demographic and cognitive factors based on previous findings [15-18]. Maximum likelihood estimation was used for factor structure verification, and missing data were imputed by estimating means and intercepts. Measures of fit, including chi-square statistic (χ^2), chi-square divided by degrees of freedom (χ^2/df), comparative fit index (CFI), Tucker-Lewis coefficient (TLI), and root-mean-square error of approximation (RMSEA), were reported. Regression weights and correlation estimates among latent factors, the outcome of interest, and control factors were also provided. For all analyses, the α-level for testing significance was set to .05.

Ethical Considerations

For this study, we analyzed presurveys from 5 Lighthouse communities, collected between July 2021 and July 2022, before receiving the intervention. All participants were aged 62 years and older, based on housing eligibility criteria. In total, 31 participants were excluded from the analysis due to missing at least 1 of the dependent variables. The final dataset included 401 participants. On the basis of the HRP-210 Determination Request, the University of California, Davis, institutional review board determined that this research is exempt as it did not directly involve human participants and used deidentified secondary data (ID: 1938286-1).

Results

Overview

Participant (n=401) demographics are described in Table 1. Participants ranged in age from 62 to 97 (mean 79.07, SD 7) years, most were female, had a high school degree or less, reported limited English proficiency, and were Asian. Over a quarter of participants reported subjective cognitive decline and over a third reported loneliness.

XSL•FO
RenderX

Table 1. Participant demographics (N=401).

Participant demographics	Frequency, n	Percentage, %
Gender		
Female	276	69.2
Male	123	30.8
Ethnicity		
Non-Hispanic White	48	12.0
Asian	320	79.8
Korean	262	65.3
Chinese	50	12.5
Japanese	2	0.5
Filipino	2	0.5
Vietnamese	1	0.2
Latinx	11	2.7
Black or African American	6	1.5
American Indian or Alaskan Native	2	0.5
More than one race or ethnicity	7	1.7
English proficiency		
Very well	58	14.8
Well	62	15.8
Not well	172	43.8
Not at all	101	25.7
Educational attainment		
Never attended school	21	5.5
Some high school	127	33.0
Completed high school or general educational development	72	18.7
Some college	75	19.5
College degree	65	16.9
Graduate degree	25	6.5
Subjective cognitive decline		
No	283	72.2
Yes	109	27.8
Years of experience using information communication technologies		
More than 2 years	239	59.6
1 to 2 years	39	9.7
Less than 1 year	32	8.0
I have never used these	91	22.7
Computer use		
About once per day	87	21.7
2 to 4 times per week	28	7.0
Once or less than once per week	25	6.2
Never	261	65.1
Tablet use		
About once per day	92	22.9

XSL•FO
RenderX

Participant demographics	Frequency, n	Percentage, %
2 to 4 times per week	29	7.2
Once or less than once per week	25	6.2
Never	255	63.6
Smartphone use		
About once per day	247	61.6
2 to 4 times per week	39	9.7
Once or less than once per week	16	4.0
Never	99	24.7

About 1 in 5 participants reported they have never used a smartphone, tablet, or computer, while 6 in 10 reported over 2 years of experience using these devices. When asked about the frequency of use of each type of device, two-thirds reported that they never use a computer, slightly less than two-thirds never use a tablet, and a quarter never use a smartphone. Despite this, 89.2% (340/381) somewhat or strongly agreed that technology is useful for connecting with family and friends, and 90.4% (341/377) agreed that technology is useful for learning new information and skills.

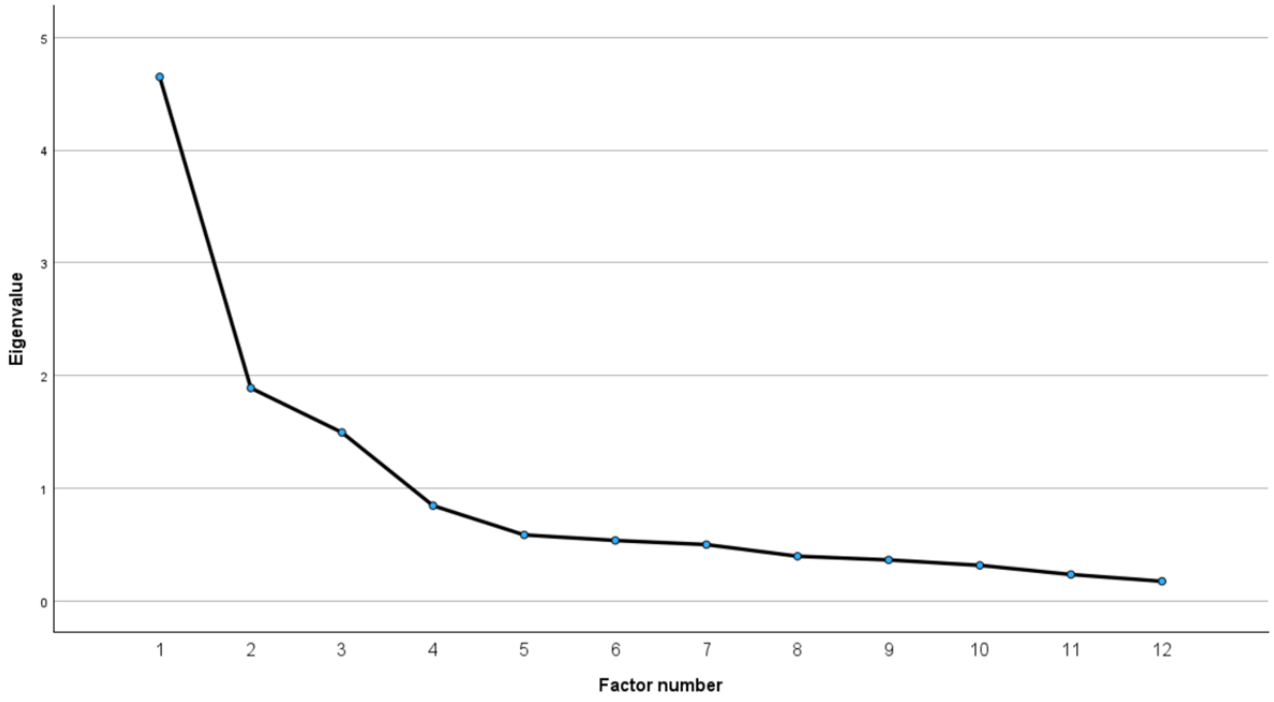
Next, we used correlation analysis to examine relationships among items measuring ICT use, PEOU, PU, loneliness, and our control variables (age, gender, Asian ethnicity, education, subjective cognitive decline, and English proficiency). The assumptions for factor analysis were met. We observed multiple correlations among the items, most ranging from 0.30 and above. Importantly, none of these correlations exceeded 0.9, indicating no issues with multicollinearity. Correlation results are included in Multimedia Appendix 2.

EFA and Horn Parallel Analysis

The Kaiser-Meyer-Olkin measure of sampling adequacy was .83 indicating sufficient correlation among the variables. The Bartlett test of Sphericity indicated $P<.001$, allowing us to reject the null hypothesis that the correlation matrix is an identity matrix and that it is reasonable to proceed with EFA.

Three factors had an initial eigenvalue greater than one (Factor 1=4.65, Factor 2=1.89, and Factor 3=1.50). This is illustrated with the scree plot (Figure 2). The rotated factor matrix is shown in Table 2.

Figure 2. Scree plot.



XSL•FO
RenderX

Table 2. Rotated factor matrix of independent variables.

Variables	Factor		
	PEOU[a]	ICT[b] use	PU[c]
Tech comfort	.40	.42	.35
Tech nervous	.72	.08	.01
Tech confidence	.82	.16	.10
Tech confusion	.88	.14	.08
Tech apprehension	.79	.20	.10
Tech fear	.75	.22	.10
Tech connection	.06	.07	.75
Tech learning	.09	.15	.85
Years of tech experience	.21	.60	.16
Computer use frequency	.19	.67	.03
Tablet use frequency	.03	.68	–.03
Smartphone use frequency	.11	.52	.17

[a]PEOU: perceived ease of use.

[b]ICT: Information communication technology.

[c]PU: perceived usefulness.

The first factor, PEOU, combined the 2 items from the Senior TAM – Tech Anxiety Subscale and 3 of the 4 items from the Attitudes Towards Computers Questionnaire – Comfort Subscale. One item, "I feel comfortable with technology," had a factor loading of .40. According to Tabachnick and Fidell [27], loading above 0.71 is excellent, 0.63 is very good, 0.55 is good, 0.45 fair, and 0.32 poor. Therefore, we dropped the tech comfort item from PEOU in further analyses.

The second factor, ICT use, combined 4 items: years of tech experience, computer use frequency, tablet use frequency, and smartphone use frequency.

The third factor, PU, combined 2 items: "Technology helps me be connected with family and friends," and "Technology helps me learn new information and skills."

After rotation, the 3 factors combined accounted for a total variance of 56.48% (6.78/12), with PEOU accounting for 28.44% (3.41/12) of the variance, ICT use for 15.59% (1.87/12), and PU for 12.44% (1.49/12). The 3 factors were confirmed when running a Horn parallel analysis. These factors also showed high internal reliability with Cronbach α scores of .90, .74, and .76 for PEOU, ICT use, and PU, respectively.

EFA was conducted separately for the dependent variables and generated 1 interpretable factor, loneliness, with an eigenvalue of 2.29 (Table 3). Loneliness had strong internal reliability with a Cronbach α score of .84.

Table 3. Component matrix for dependent variables using principal extraction method.

Variables	Factor (loneliness)
Lack companionship	.69
Feel left out	.93
Feel isolated	.79

CFA Findings

We ran CFA to further examine the relationships among the latent variables and to assess our conceptual model (Figure 3). As shown in Table 4, each of the items significantly loaded to form the 3 latent factors, confirming our EFA results.

Except for chi-square, all fit indices reached recommended level of fit: (χ^2_{41}=182.114; P<.001, chi-square minimum (CMIN)/df=4.44, CFI=0.92, TLI=0.87). The RMSEA of 0.09 was borderline. Since χ^2 is sensitive to large sample sizes, with a large sample of 401 participants, it was not unusual to get a significant value; for sample sizes greater than 250, a significant χ^2 value is acceptable [31].

XSL•FO
RenderX

Figure 3. Standardized results from confirmatory factor analysis to assess our conceptual model. ICT: information communication technology.

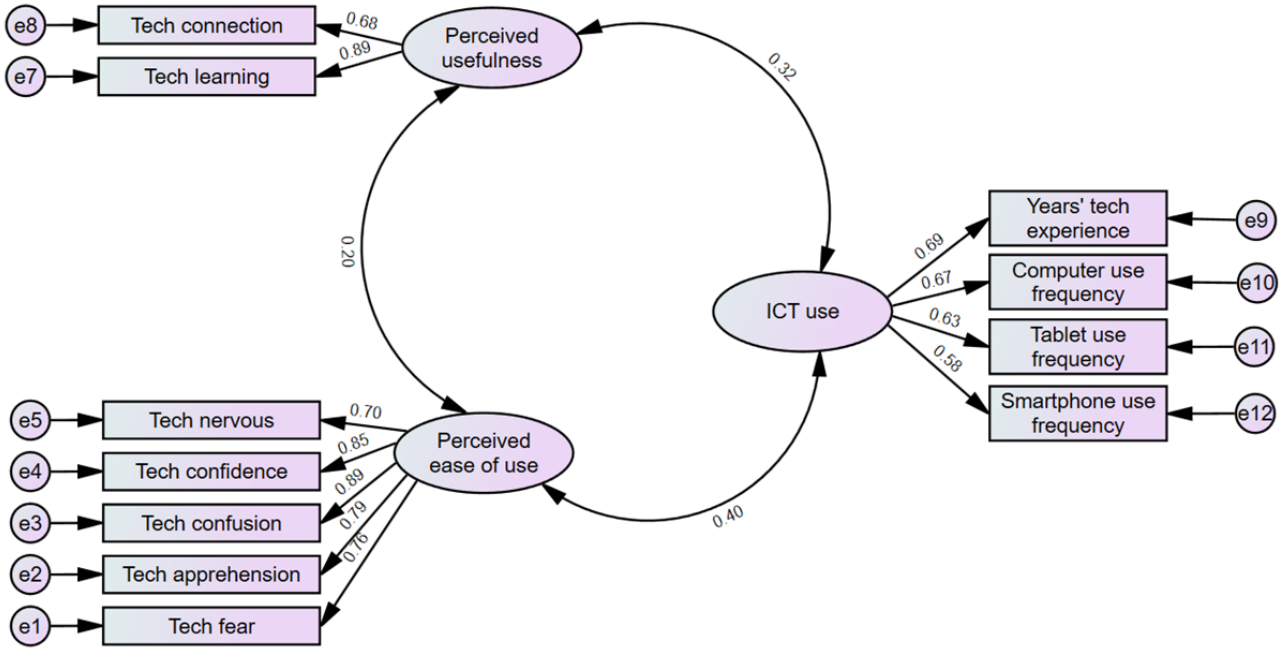


Table 4. Regression estimates of items loading into 3 factors.

Observed variables		Latent variables	Standardized regression estimate (factor loading)	Unstandardized regression estimate	SE	Composite reliability	*P* value
Tech fear	←	PEOU[a]	.76	1.00	—[b]	—	—
Tech apprehension	←	PEOU	.80	.99	.06	15.963	<.001
Tech confusion	←	PEOU	.89	1.09	.06	18.156	<.001
Tech confidence	←	PEOU	.85	1.05	.06	17.130	<.001
Tech nervous	←	PEOU	.70	.86	.06	13.797	<.001
Tech learning	←	PU[c]	.90	1.00	—	—	—
Tech connection	←	PU	.68	.77	.16	4.796	<.001
Years of tech experience	←	ICT[d] use	.69	1.00	—	—	—
Computer use frequency	←	ICT use	.67	.97	.10	10.053	<.001
Tablet use frequency	←	ICT use	.63	.93	.10	9.719	<.001
Smartphone use frequency	←	ICT use	.58	.86	.09	9.166	<.001

[a]PEOU: perceived ease of use.

[b]Not applicable.

[c]PU: perceived usefulness.

[d]ICT: Information communication technology.

There was evidence for convergent validity because all three of the conditions were fulfilled, that is, (1) composite reliability values are 0.7 or greater, (2) all standardized factor loadings are 0.5 or greater, and (3) AVE values are 0.5 or greater [29]. All 3 of these criteria were met for PEOU and PU, which had composite reliability values of .90 and .77, standardized factor loadings all greater or equal to .68, and AVE values of .64 and .63, respectively.

ICT use had an AVE of .42, slightly lower than ideal [29]. However, according to Fornell and Larcker [30], if the AVE is less than 0.5, but composite reliability is higher than 0.6, the convergent validity of the construct is acceptable [30]. ICT use had a composite reliability of .74, therefore we concluded that the latent variable had convergent validity.

Discriminant validity was met, with discriminant values of .80, .80, and .65 for PEOU, PU, and ICT use, respectively, while the correlation estimates all fell below 0.4 as shown in Table 5 [30].

Finally, as shown in Table 6, the relationships between the latent variables were all significant.

XSL•FO
RenderX

Table 5. Correlations and covariances of latent variables in confirmatory factor analysis.

Latent variables			Correlation estimate	Covariance estimate	SE	Composite reliability	*P* value
PEOU[a]	↔	PU	.200	.107	.033	3.273	.001
ICT[b] Use	↔	PU	.317	.190	.042	4.501	<.001
ICT Use	↔	PEOU	.399	.269	.048	5.643	<.001

[a]PEOU: perceived ease of use.

[b]ICT: Information communication technologies.

Table 6. Regression weights (unadjusted model).

Latent variables			Standardized regression estimate	Unstandardized regression estimate	SE	Composite reliability	*P* value
PU[a]	←	PEOU	.20	.18	.05	3.38	<.001
ICT[b] Use	←	PU	.24	.31	.10	3.19	.001
ICT Use	←	PEOU	.37	.41	.07	5.75	<.001
Loneliness	←	ICT Use	-.29	-.15	.03	-4.48	<.001

[a]PU: perceived usefulness.

[b]ICT: Information communication technologies.

Structural Equation Modeling

The 3-factor SEM revealed reasonable fit indexes (χ^2_{73}=231.835, CMIN/DF=3.18, CFI=0.93, TLI=0.90). An RMSEA of 0.07 was acceptable. Once again, the chi-square value was significant, which, as previously described, is acceptable for this sample size [31].

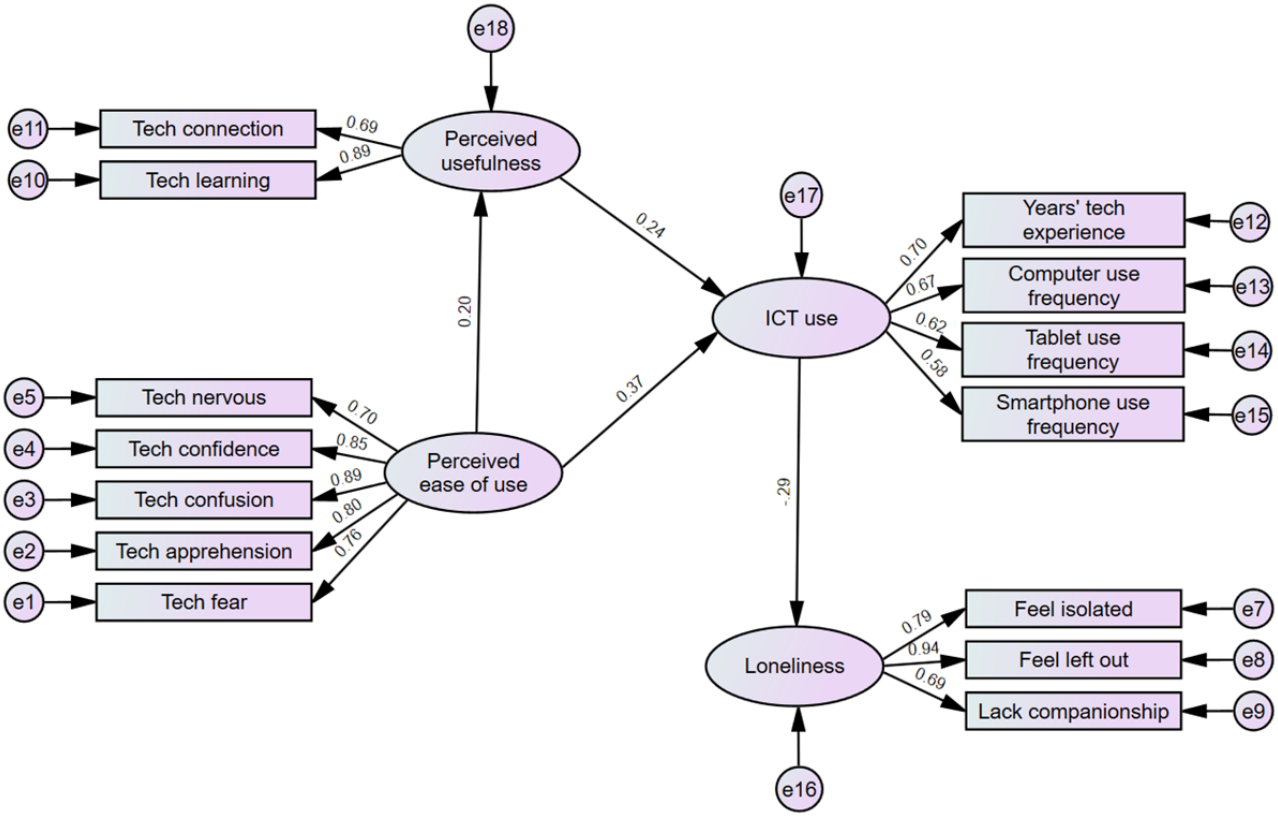
As shown in Figure 4 and Table 6, hypothesis 1 was supported; PEOU was significantly, and positively associated with PU (β=.20, *P*<.001).

Hypothesis 2 was also supported; both PEOU and PU were significant, positive predictors of ICT use (PEOU: β=.37; *P*<.001; PU: β=.24; *P*=.001).

Finally, hypothesis 3 was supported; ICT use was significantly, negatively associated with loneliness (β=–.29; *P*<.001).

XSL•FO
RenderX

Figure 4. Standardized results from structural equation modeling (unadjusted model).



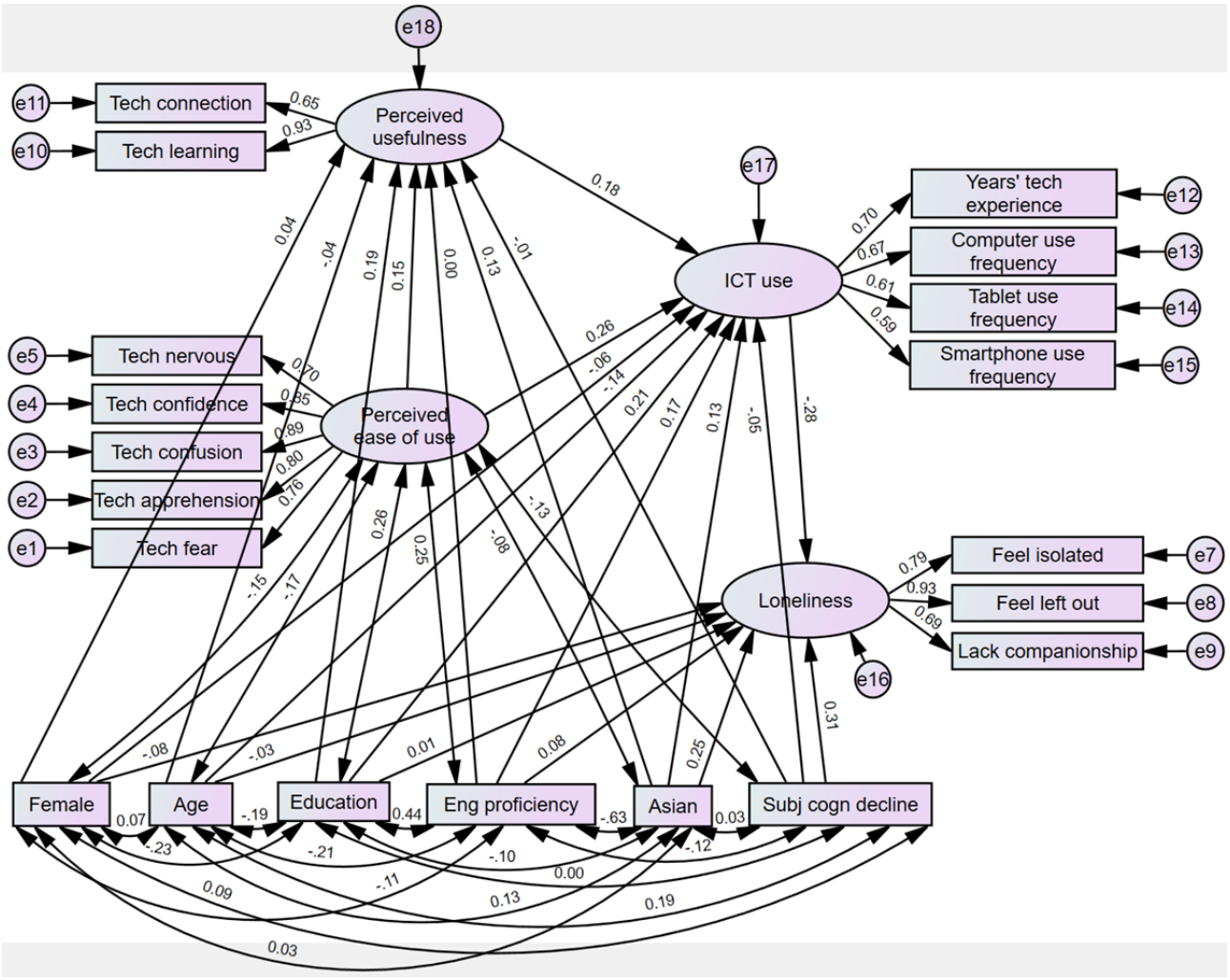
Structural Equation Modeling Adjusting for Control Factors

When we reran the model adjusting for age, gender, education, Asian ethnicity, English proficiency, and subjective cognitive decline, the model fit improved (Figure 5; χ^2_{133}=345.13, *P*<.001, CMIN/DF=2595, CFI=0.93, TLI=0.88). The RMSEA of 0.06 was acceptable.

As shown in Table 7, the adjusted results continued to support Hypotheses 1, 2, and 3. PEOU continues to be significantly positively associated with PU (β=.152, *P*=.01). PEOU and PU were significant, positive predictors of ICT use (PEOU: β=.26, *P*<.001; PU: β=.179, *P*=.01). And, ICT use was significantly negatively associated with loneliness (β=–.28, *P*<.001). In addition, some of the control variables were significant predictors of the endogenous variables, PU and ICT use, and the dependent variable, loneliness (Table 7). Education was significantly, positively associated with PU (β=.19; *P*=.003). English proficiency and education significantly, positively predicted ICT use (English proficiency: β=1.66; *P*=.03; Education: β=.21; *P*<.001), while age was negatively associated with ICT use (β=–1.36; *P*=.01). Finally, subjective cognitive decline and Asian ethnicity were each positively associated with loneliness (subjective cognitive decline: β=.31; *P*<.001; Asian ethnicity: β=.25; *P*<.001).

XSL•FO
RenderX

Figure 5. Adjusted, standardized results from structural equation modeling. ICT: information communication technology.



XSL•FO
RenderX

Table 7. Regression weights (adjusted model).

Variables	Standardized regression estimate	Unstandardized regression estimate	SE	Composite reliability	*P* value
PU[a]←PEOU[b]	.15[c]	.14	.06	2.55	.01
PU←English proficiency	–.003	–.002	.06	–.04	.97
PU←Asian	.13	.24	.13	1.83	.07
PU←subjective cognitive decline	–.01	–.02	.09	–.22	.83
PU←age	–.04	–.004	.01	–.78	.44
PU←education	.19[c]	.10	.03	2.98	.003
PU←female	.04	.07	.09	.80	.42
ICT[d] use←PU	.18[c]	.22	.08	2.70	.01
ICT use←female	–.06	–.12	.10	–1.17	.24
ICT use←age	–.14[c]	–.02	.01	–2.52	.01
ICT use←English proficiency	.17[c]	.15	.07	2.13	.03
ICT use←subjective cognitive decline	–.05	–.10	.10	–.92	.36
ICT use←PEOU	.26[c]	.29	.07	4.30	<.001
ICT use←education	.21[c]	.13	.04	3.30	<.001
ICT use←Asian	.13	.29	.15	1.93	.05
Loneliness←female	–.08	–.08	.05	–1.59	.11
Loneliness←age	–.03	–.002	.003	–.63	.53
Loneliness←English proficiency	.08	.04	.04	1.01	.31
Loneliness←ICT use	–.28[c]	–.14	.04	–3.93	<.001
Loneliness←subjective cognitive decline	.31[c]	.31	.05	5.76	<.001
Loneliness←education	.01	.004	.02	.21	.84
Loneliness←Asian	.25[c]	.28	.08	3.64	<.001

[a]PU: perceived usefulness.

[b]PEOU: perceived ease of use.

[c]$P \leq .05$.

[d]ICT: Information communication technology.

While PEOU was an exogenous factor, correlation estimates reveal significant associations with all the control variables except for Asian ethnicity (Table 8). English proficiency and education were each significantly positively associated with PEOU (English proficiency: r=0.25, P<.001; Education: r=0.26; P<.001). Subjective cognitive decline, age, and female gender were each significantly, and negatively associated with PEOU (Subjective cognitive decline: r=–0.130; P=.02; Age: r=–0.18; P=.001; Female: r=–0.15; P=.01).

Table 8. Correlations and covariances of control variables with perceived ease of use (adjusted model).

Control variables	Correlation estimate	Covariance estimate	SE	Composite reliability	*P* value
Asian	–.09	–.03	.02	–1.58	.11
Subjective cognitive decline	–.13	–.05	.02	–2.40	.02
English proficiency	.25	.19	.04	4.47	<.001
Education	.26	.28	.06	4.54	<.001
Age	–.18	–.95	.30	–3.18	.001
Female	–.15	–.05	.02	–2.73	.01

XSL•FO
RenderX

Discussion

Principal Findings

In this study, we examined the association of PU, PEOU, ICT use, and loneliness among low income, predominately Asian American, older adults. This research built upon 2 previous studies that simplified the TAM to examine demographic and health factors that impact technology acceptance among Asian Americans aged 62 years and older living in affordable senior housing communities [17,18]. In these previous analyses, the conceptualization of PU, PEOU, and ICT use were based on the existing literature and theory around how these constructs apply to older adults. The constructs were developed by summing and normalizing 12 survey items. In this study, the assumptions underlying the conceptualization of PU, PEOU, and ICT use and the TAM framework were rigorously assessed and confirmed using a combination of statistical techniques. Initially, EFA was conducted to determine the appropriate number of factors and explore the underlying structure of the constructs. Horn parallel analysis was subsequently used to validate the EFA results. Following this, SEM (beginning with CFA) was performed to validate and confirm the factor structure identified in EFA and examine the relationships among PU, PEOU, ICT use, and loneliness.

The original TAM focused on technology acceptance in the workplace. PU was measured with items such as, "Using [chart master] in my job would enable me to accomplish tasks more quickly," and, "Using [chart master] would improve my job performance" [21]. Our research findings support a modified construct of PU that accounts for older adults' perceptions of ICTs as being useful for social connection and for learning new information and skills. We found that the item, "Technology helps me learn new information and skills," had a slightly higher factor loading than the item, "Technology helps me be connected with family and friends." Mixed methods longitudinal data from the Lighthouse Project for Older Adults supported this finding; participants reported that they most frequently used ICTs to access YouTube (eg, to view videos related to nutrition, exercise, and cultural content), followed by accessing entertainment and checking the weather [32]. The Pew Research Center reported that, in 2021 among adults aged 65 years and older, the use of YouTube experienced the most growth as compared with any other app [33]. Notably, studies suggest that app- and web-based activities among older adults vary significantly by age group and gender [23,34-36].

When it comes to PU, it is essential to note that the COVID-19 pandemic spurred an exponential increase in the use of ICTs for social connection. American Association of Retired Persons (AARP) 2021 Tech Trends reported a notable surge in the use of various communication technologies among individuals aged 50 years and above to stay connected with others. A significant portion of this demographic reported an uptick in their usage of video chats (45%, 1022/2271), texting (37%, 840/2271), emailing (26%, 590/2271), and phone calls (29%; 659/2271) compared with prepandemic levels. In 2019, approximately half had never used video chat, whereas by 2020, this figure rose to 70%, with one out of 3 engaging in video chats on a weekly basis [37].

Our findings supported the operationalization of PEOU combining 2 items from the Senior TAM – Tech Anxiety Subscale [24] and 3 items from the ATCQ – Comfort Subscale [25]. Previous studies note an array of emotions that influence technology acceptance and use, including enjoyment, effort expectancy [38], control, efficacy [24,39], confidence [40,41], comfort [25,42,43], and anxiety [24,40]. Due to the variety of existing measures and constructs highlighted in the literature (some developed in the 1980s and with highly educated, White, middle-aged adults), it can be challenging for researchers to select a concise set of items that are specific to older adults and modern technology. We believe our findings can inform future studies with Asian American older adults, who require survey modification [44]. During exploratory analysis, one item, "I feel comfortable with technology," was dropped due to low factor loading. We believe this may have been due to the item being reverse scored, which may have been confusing due to participants' limited literacy levels. Our findings align with previous research which suggests that assessment scales containing reverse-scored items impose higher cognitive processing demands, potentially resulting in measurement challenges for older adult participants [45], particularly since our participants had self-reported limited English proficiency and low educational attainment (Table 1).

Our mediating factor, ICT use, was unique in that it combined measures assessing the frequency of use of smartphones, tablets, and computers, as well as years of experience. We believe this measure is valuable to better understand ICT use since these devices are often used interchangeably among older adults for multitasking [46]. Furthermore, we are not the first to identify years of experience as an important factor in understanding technology acceptance [38].

Using CFA, we found significant, positive relationships between PU, PEOU, and ICT use. This finding is aligned with the Senior Technology Exploration, Learning, and Acceptance model [47] and the Senior Technology Acceptance and Adoption Model (STAM) [48]. The concept of reinforcement of use is consistent with a model of technology acceptance or rejection from an ease-of-learning perspective [49]. When ICTs are perceived as more useful, their usage tends to increase, forming a reinforcing cycle. Increased usage and familiarity make it easier, encouraging individuals to master new skills and diversify their usage, such as progressing from social chats to watching YouTube videos, further reinforcing their mastery and use.

SEM confirmed our modified TAM, supporting our three hypotheses. PEOU significantly, and positively predicted PU and ICT use. Further, ICT use was significantly negatively associated with loneliness. These relationships remained significant, even when adjusting for gender, age, education, English proficiency, Asian ethnicity, and subjective cognitive decline. Interestingly, English proficiency was related to PEOU and ICT use, suggesting a cultural or linguistic bias in technology and application development favoring English-speaking users. Asian ethnicity was associated with loneliness, affirming observations of the vital role technology

XSL•FO
RenderX

can play for immigrants in maintaining social and cultural connections [11]. Subjective cognitive decline was also associated with loneliness, perhaps a reflection of lower social and technology engagement [50].

This study has implications for interventions, particularly among populations with lower literacy. Both PEOU and PU are modifiable factors and could be enhanced by offering digital literacy training and support to new learners [43,51-53] or tailoring user interfaces [54,55]. The relevance of technology for the individual can be increased by demonstrating various culturally relevant use cases for devices, such as connecting with distant relatives, accessing health information, or enjoying entertainment in one's native language. In the Lighthouse Project for Older Adults, the low-income housing providers made a commitment to enhance technology use and offered the technology, training, and support to encourage adoption. This not only provided tangible support in the form of equipment and training but also created a community of learners where individuals could benefit from the experience of peers [32]. This is an example of how service providers can play a role in potentially improving quality of life and reducing loneliness by encouraging the adoption of technology.

Study Limitations

This study was partially limited by the use of a convenience sample of participants who were interested in learning more about technology, not a random sample. This potentially biases the sample toward those with more positive attitudes toward technology. Even with this bias, we observed variability in perceptions about technology but did not capture the full array of attitudes likely present in the population. In addition, participants all lived in age-restricted affordable housing communities. Therefore, our findings are not representative of older adults living in multigenerational households. We removed participants whose data were missing the dependent variable, loneliness. We retained all other participants, yet 4%-7% of responses were missing for items related to PEOU and PU.

Methodological considerations included minimizing participant burden to maximize potential engagement with technology and the most complete dataset. This required modifications to measures, to accomplish parsimony and ease of administration, potentially compromising psychometric properties. However, our CFA indicated the adequacy of the measures for crucial constructs, despite having fewer than 3 observations for each latent variable in the case of PU.

EFA and SEM were limited by the inclusion of binary and ordinal observed variables (eg, gender, Asian ethnicity, and subjective cognitive decline). While this represents a methodological weakness, these variables were included because of their known importance to the constructs of interest. Furthermore, we used the same dataset for all multivariate analyses because of the early phase of understanding relationships among our variables of interest, recognizing that typically the confirmatory analysis should be conducted on an independent sample.

Finally, many other factors potentially impact technology acceptance among older adults, including social and health factors, access to Wi-Fi and devices, and digital literacy training.

Conclusion

Despite the limitations, this study affirms the usefulness of the TAM in understanding the dynamics of technology adoption among a low-income Asian American population. At baseline, there is considerable interest in technology, affirming its relevance in the lives of older adults. The role of English proficiency in ICT use warrants further exploration to identify ways to increase equity and access for those who have another primary language. The study further highlights the potential role that technology could play in alleviating loneliness through greater engagement with family and friends and the ability to maintain cultural ties. Future studies could explore the most effective ways to overcome resistance to technology, the most meaningful ways to support novice users to adopt a new device, and ways to increase the diversity of use once a user has become comfortable with basic functions. While technology cannot replace the human touch, it has the promise to improve engagement and social connection among isolated, low-income older adults.

Acknowledgments

This work was funded by the Healthy Aging in a Digital World Initiative at the University of California, Davis and the Center for Information Technology Research in the Interest of Society (CITRIS) and the Banatao Institute.

Other acknowledgments: The University of California, Davis, School of Medicine Office of Research, Evaluation Unit, including Rebeca Giacinto, Pauline DeLange Martinez, Stuart Henderson, and Julie Rainwater, designed the Lighthouse surveys and provided oversight for data collection and entry. We thank all Lighthouse participants for their time and effort in completing the evaluation surveys. We also thank the staff at each site for their support with participant recruitment, data collection, and entry.

Lighthouse project leadership included the Front Porch Center for Innovation and Wellbeing (Davis Park, Kari Olson, Jennifer Lee), Eskaton (Therese ten Brinke, Sheri Peifer), and CITRIS (David Lindeman, Gale Berkowitz).

Data Availability

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

XSL•FO
RenderX

Authors' Contributions

PDM led conceptual design, data analysis, and authorship. HMY made substantial conceptual and design contributions, providing invaluable input along all stages of manuscript preparation. MP and DJT provided oversight for the operationalization of key variables, data analysis, reporting results, and interpretation of the data. LG helped to critically revise the manuscript and agreed to be accountable for the accuracy of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Lighthouse survey.
[DOCX File , 53 KB - aging_v8i1e63856_app1.docx]

Multimedia Appendix 2
Correlation matrix.
[XLSX File (Microsoft Excel File), 20 KB - aging_v8i1e63856_app2.xlsx]

References

1. UCLA Center for Health Policy Research. AskCHIS. Loneliness Scale. Los Angeles. URL: https://healthpolicy.ucla.edu/our-work/askchis [accessed 2024-11-20]
2. New Surgeon General Advisory Raises Alarm about the Devastating Impact of the Epidemic of Loneliness and Isolation in the United States. 2023. URL: https://www.hhs.gov/about/news/2023/05/03/new-surgeon-general-advisory-raises-alarm-about-devastating-impact-epidemic-loneliness-isolation-united-states.html [accessed 2024-01-28]
3. Shimkhada R, Tse HW, Ponce NA. Life satisfaction and social and emotional support among asian American older adults. J Am Board Fam Med 2022;35(1):203-205 [FREE Full text] [doi: 10.3122/jabfm.2022.01.210232] [Medline: 35039430]
4. Malani P, Singer D, Kirch M, Solway E, Roberts S, Smith E. Trends in loneliness among older adults from 2018-2023. University of Michigan National Poll on Healthy Aging; 2023 updated March 2019;3(1):600-601 [FREE Full text] [doi: 10.1093/geroni/igz038.2233]
5. Pinquart M, Sörensen S. Risk Factors for Loneliness In Adulthood and Old Age--A Meta-Analysis. Hauppauge, New York: Nova Science Publishers; 2003.
6. Cohen-Mansfield J, Hazan H, Lerman Y, Shalom V. Correlates and predictors of loneliness in older-adults: a review of quantitative results informed by qualitative insights. Int. Psychogeriatr 2015;28(4):557-576. [doi: 10.1017/s1041610215001532]
7. Delaruelle K. Migration-related inequalities in loneliness across age groups: a cross-national comparative study in Europe. Eur J Ageing 2023;20(1):35 [FREE Full text] [doi: 10.1007/s10433-023-00782-x] [Medline: 37612470]
8. Wang Y, Woo B, Jiang N. Correlates of loneliness among older asian immigrants: a systematic review. Innovation in Aging 2019;3(1):629. [doi: 10.1093/geroni/igz038.2345]
9. Petersen B, Khalili-Mahani N, Murphy C, Sawchuk K, Phillips N, Li KZH, et al. The association between information and communication technologies, loneliness and social connectedness: a scoping review. Front Psychol 2023;14:1063146 [FREE Full text] [doi: 10.3389/fpsyg.2023.1063146] [Medline: 37034933]
10. Ibarra F, Baez M, Cernuzzi L, Casati F. A systematic review on technology-supported interventions to improve old-age social wellbeing: loneliness, social isolation, and connectedness. J Healthc Eng 2020;2020:2036842 [FREE Full text] [doi: 10.1155/2020/2036842] [Medline: 32765823]
11. Khvorostianov N, Elias N, Nimrod G. 'Without it i am nothing': the internet in the lives of older immigrants. New Media & Society 2011;14(4):583-599. [doi: 10.1177/1461444811421599]
12. Melenhorst AS, Rogers WA, Caylor EC. The use of communication technologies by older adults: exploring the benefits from the user's perspective. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2001;45(3):221-225. [doi: 10.1177/154193120104500305]
13. Blacher K. National asian pacific center on aging. Asian Americans and Pacific Islanders in the United States Aged 65 Years and Older: Population, Nativity, and Language. 2013. URL: https://www.napca.org/wp-content/uploads/2017/10/65-population-report-FINAL.pdf [accessed 2021-12-30]
14. Yoon H, Jang Y, Kim S, Speasmaker A, Nam I. Trends in internet use among older adults in the United States, 2011-2016. J Appl Gerontol 2021;40(5):466-470. [doi: 10.1177/0733464820908427] [Medline: 32131670]
15. Yoon H, Jang Y, Vaughan PW, Garcia M. Older adults' internet use for health information: digital divide by race/ethnicity and socioeconomic status. J Appl Gerontol 2020;39(1):105-110. [doi: 10.1177/0733464818770772] [Medline: 29661052]
16. Kim J, Lee HY. Older minority immigrants' digital technology use. Does English Proficiency Matter? The Gerontologist 2016;56(3):139. [doi: 10.1093/geront/gnw162.557]

XSL•FO
RenderX

17. DeLange Martinez P, Tancredi D, Pavel M, Garcia L, Heather MY. The role of health in the technology acceptance model among low-income, asian American older adults: a cross sectional survey analysis. JMIR Formative Research 2024:TBD (forthcoming) [FREE Full text] [doi: 10.2196/57009]
18. Delange Martinez P, Tancredi DJ, Pavel M, Garcia L, Young HM. JMIR. Technology Acceptance Among Low-In come, Asian American Older Adults: Cross-Sectional Survey Analysis 2024 Nov 22;26:e52498 [FREE Full text] [doi: 10.2196/52498] [Medline: 39576987]
19. Schroeder T, Dodds L, Georgiou A, Gewald H, Siette J. Older adults and new technology: mapping review of the factors associated with older adults' intention to adopt digital technologies. JMIR Aging 2023;6:e44564 [FREE Full text] [doi: 10.2196/44564] [Medline: 37191976]
20. Charness N, Boot WR. Chapter 20 - technology, gaming, and social networking. In: Schaie KW, Willis SL, editors. Handbook of the Psychology of Aging (Eighth Edition). Saon Dieg: Academic Press; 2016.
21. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989;13(3):319. [doi: 10.2307/249008]
22. CITRIS and the Banatao Institute. Brings technology-enabled health and well-being to low-income california seniors during COVID crisis. Lighthouse for Older Adults, a CITRIS and University of California Initiative. 2020. URL: https://citris-uc.org/lighthouse-for-older-adults/ [accessed 2020-09-16]
23. Sims T, Reed AE, Carr DC. Information and communication technology use is related to higher well-being among the oldest-old. J Gerontol B Psychol Sci Soc Sci 2017;72(5):761-770. [doi: 10.1093/geronb/gbw130] [Medline: 27702839]
24. Chen K, Chan AHS. Gerontechnology acceptance by elderly Hong Kong Chinese: a senior technology acceptance model (STAM). Ergonomics 2014;57(5):635-652. [doi: 10.1080/00140139.2014.895855] [Medline: 24655221]
25. Jay GM, Willis SL. Influence of direct computer experience on older adults' attitudes toward computers. J Gerontol 1992;47(4):250-257. [doi: 10.1093/geronj/47.4.p250] [Medline: 1624702]
26. Hughes ME, Waite LJ, Hawkley LC, Cacioppo JT. A short scale for measuring loneliness in large surveys: results from two population-based studies. Res Aging 2004;26(6):655-672 [FREE Full text] [doi: 10.1177/0164027504268574] [Medline: 18504506]
27. Tabachnick BG, Fidell LS. In: Boston MA, editor. Using Multivariate Statistics. USA: Pearson Education, Inc; 2018.
28. Kaplan D. Structural equation modeling. In: Smelser NJ, Baltes PB, editors. International Encyclopedia of the Social & Behavioral Sciences. Oxford: Pergamon; 2001:15215-15222.
29. Hair JFJ, Black WC, Babin BJ, Anderson RE, Tatham RL. Multivariate Data Analysis. New Jersey: Prentice-Hall; 2009.
30. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research 1981;18(1):39. [doi: 10.2307/3151312]
31. Hair JFJ, Black WC, Babin BJ, Anderson RE, Tatham RL. Multivariate Data Analysis. New Jersey: Prentice-Hall; 2018.
32. Giacinto RE, Martinez P, Neves S, Sullivan M, Kipke R, Henderson S. University of california, davis, school of medicine office of research evaluation unit,. Lighthouse for Older Adults Evaluation Report. URL: https://health.citris-uc.org/wp-content/uploads/2023/08/Lighthouse-UCD-SOMOR-Evaluation-Summary-Report-FINAL-PDF.pdf [accessed 2024-11-20]
33. Faverio M. Pew Research Center. Share of those 65 and older who are tech users has grown in the past decade. URL: https://www.pewresearch.org/short-reads/2022/01/13/share-of-those-65-and-older-who-are-tech-users-has-grown-in-the-past-decade/ [accessed 2022-01-13]
34. Anderson GO. Getting Connected Older Americans Embrace Technology to Enhance Their Lives. AARP Research. 2018. URL: https://www.aarp.org/research/topics/technology/info-2018/technology-use-attitudes.html [accessed 2018-02-23]
35. Kim J, Lee HY, Christensen MC, Merighi JR. Technology access and use, and their associations with social engagement among older adults: do women and men differ? J Gerontol B Psychol Sci Soc Sci 2017;72(5):836-845. [doi: 10.1093/geronb/gbw123] [Medline: 28073816]
36. Yu RP, Ellison NB, McCammon RJ, Langa KM. Mapping the two levels of digital divide: internet access and social network site adoption among older adults in the USA. Information, Communication & Society 2015;19(10):1445-1464. [doi: 10.1080/1369118x.2015.1109695]
37. Kakulla B. Personal Tech and the Pandemic.: Older Adults Are Upgrading for a Better Online Experience. AARP Research URL: https://www.aarp.org/research/topics/technology/info-2021/2021-technology-trends-older-americans.html-CMP=RDRCT-PRI-TECH-040721/?cmp=RDRCT-907b618d-20210416 [accessed 2024-02-05]
38. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Quarterly 2003;27(3):425. [doi: 10.2307/30036540]
39. Cassidy S, Eachus P. Developing the computer user self-efficacy (Cuse) scale: investigating the relationship between computer self-efficacy, gender and experience with computers. Journal of Educational Computing Research 2002;26(2):133-153. [doi: 10.2190/jgjr-0kvl-hrf7-gcnv]
40. Loyd BH, Gressard C. Reliability and factorial validity of computer attitude scales. Educational and Psychological Measurement 1984;44(2):501-505. [doi: 10.1177/0013164484442033]
41. Nickell GS, Pinto JN. The computer attitude scale. Computers in Human Behavior 1986:301-306 [FREE Full text] [doi: 10.1016/0747-5632(86)90010-5]

XSL•FO
RenderX

42. Laganà L, Oliver T, Ainsworth A, Edwards M. Enhancing computer self-efficacy and attitudes in multi-ethnic older adults: a randomised controlled study. Ageing Soc 2011;31(6):911-933 [FREE Full text] [doi: 10.1017/S0144686X10001340] [Medline: 25512679]
43. Arthanat S. Promoting information communication technology adoption and acceptance for aging-in-place: a randomized controlled trial. J Appl Gerontol 2021;40(5):471-480 [FREE Full text] [doi: 10.1177/0733464819891045] [Medline: 31782347]
44. Stewart AL, Thrasher AD, Goldberg J, Shea JA. A framework for understanding modifications to measures for diverse populations. J Aging Health 2012;24(6):992-1017 [FREE Full text] [doi: 10.1177/0898264312440321] [Medline: 22495768]
45. Carlson M, Wilcox R, Chou CP, Chang M, Yang F, Blanchard J, et al. Psychometric properties of reverse-scored items on the CES-D in a sample of ethnically diverse older adults. Psychol Assess 2011;23(2):558-562 [FREE Full text] [doi: 10.1037/a0022484] [Medline: 21319906]
46. Kononova A, Joshi P, Cotten S. Contrary to myth, older adults multitask with media and technologies, but studying their multitasking behaviors can be challenging. Innov Aging 2019;3(4):igz029 [FREE Full text] [doi: 10.1093/geroni/igz029] [Medline: 31637313]
47. Tsai HYS, Rikard RV, Cotten SR, Shillair R. Senior technology exploration, learning, and acceptance (STELA) model: from exploration to use – a longitudinal randomized controlled trial. Educational Gerontology 2019;45(12):728-743. [doi: 10.1080/03601277.2019.1690802]
48. Renaud K, Biljon JV. Predicting technology acceptance and adoption by the elderly: a qualitative study. 2008 Presented at: SAICSIT '08: Proceedings of the 2008 annual research conference of the South African Institute of Computer Scientists and Information Technologists on IT research in developing countries: riding the wave of technology; 2008 October 8; New York. [doi: 10.1145/1456659.1456684]
49. Barnard Y, Bradley MD, Hodgson F, Lloyd AD. Learning to use new technologies by older adults: perceived difficulties, experimentation behaviour and usability. Computers in Human Behavior 2013;29(4):1715-1724 [FREE Full text] [doi: 10.1016/j.chb.2013.02.006]
50. Li Y, Godai K, Kido M, Komori S, Shima R, Kamide K, et al. Cognitive decline and poor social relationship in older adults during COVID-19 pandemic: can information and communications technology (ICT) use helps? BMC Geriatr 2022;22(1):375 [FREE Full text] [doi: 10.1186/s12877-022-03061-z] [Medline: 35484494]
51. Lagana L, Garcia JJ. The mental health impact of computer and internet training on a multi-ethnic sample of community-dwelling older adults: results of a pilot randomised controlled trial. Int. J Biomed Sci 2013;9(3):135-147 [FREE Full text] [doi: 10.59566/ijbs.2013.9135]
52. González A, Ramírez MP, Viadel V. ICT learning by older adults and their attitudes toward computer use. Curr Gerontol Geriatr Res 2015;2015:849308 [FREE Full text] [doi: 10.1155/2015/849308] [Medline: 26346158]
53. Senteio CR. Investigating the enduring impact of a community-based health education program to promote African American elders' use of technology designed to support chronic disease self-management. Geriatrics (Basel) 2018;3(4):70 [FREE Full text] [doi: 10.3390/geriatrics3040070] [Medline: 31011105]
54. Czaja SJ, Boot WR, Charness N, Rogers WA, Sharit J. Improving social support for older adults through technology: findings from the PRISM randomized controlled trial. Gerontologist 2018;58(3):467-477 [FREE Full text] [doi: 10.1093/geront/gnw249] [Medline: 28201730]
55. Lor M, Bowers BJ. Feasibility of audio-computer-assisted self-interviewing with color-coding and helper assistance (ACASI-H) for hmong older adults. Res Nurs Health 2017;40(4):360-371 [FREE Full text] [doi: 10.1002/nur.21796] [Medline: 28431187]

Abbreviations

ATCQ: Attitudes Towards Computers Questionnaire
AVE: average variance extracted
CFA: confirmatory factor analysis
CFI: Comparative Fit Index
CHIS: California Health Interview Survey
CMIN: chi-square minimum
EFA: exploratory factor analysis
ICT: Information communication technology
NHATS: National Health and Aging Trends Study
PEOU: perceived ease of use
PU: perceived usefulness
RMSEA: root-mean-square error of approximation
SEM: structural equation modeling
STAM: Senior Technology Acceptance and Adoption Model
TAM: Technology Acceptance Model

XSL•FO
RenderX

TLI: Tucker-Lewis Index

Edited by PH Chau; submitted 01.07.24; peer-reviewed by R Shimkhada, Z Luo; comments to author 12.09.24; revised version received 22.10.24; accepted 06.11.24; published 08.01.25.

Please cite as:
DeLange Martinez P, Tancredi D, Pavel M, Garcia L, Young HM
Adapting the Technology Acceptance Model to Examine the Use of Information Communication Technologies and Loneliness Among Low-Income, Older Asian Americans: Cross-Sectional Survey Analysis
JMIR Aging 2025;8:e63856
URL: https://aging.jmir.org/2025/1/e63856
doi:10.2196/63856
PMID:

©Pauline DeLange Martinez, Daniel Tancredi, Misha Pavel, Lorena Garcia, Heather M Young. Originally published in JMIR Aging (https://aging.jmir.org), 08.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Implementation of an Electronic Medication Management System in 41 Residential Care Homes in Hong Kong: Pre–Post Interventional Study

Phillip Lung Wai Au-Doung[1], BPharm; Ho Cheung Chau[2], BPharm; Chit Hin Wong[1], BPharm; Kai Cheong Lau[2], BPharm; Yi Qi Chan[2], BPharm; Ho Lam Yip[2], BPharm; Chiu Ngai Hung[2], BPharm; Bik-Wai Bilvick Tai[1], PharmD; Teddy Tai-Ning Lam[1], PharmD, PhD; Chui Ping Lee[1], PharmD; Sau Chu Chiang[2*], BPharm, MHA; Yin Ting Cheung[1*], PhD

[1]School of Pharmacy, Faculty of Medicine, The Chinese University of Hong Kong, Lo Kwee-Seong Integrated Biomedical Sciences Building, Hong Kong, China (Hong Kong)

[2]Hong Kong Pharmaceutical Care Foundation Limited, Hong Kong, China (Hong Kong)

[*]these authors contributed equally

Corresponding Author:
Yin Ting Cheung, PhD
School of Pharmacy, Faculty of Medicine, The Chinese University of Hong Kong, Lo Kwee-Seong Integrated Biomedical Sciences Building, Hong Kong, China (Hong Kong)

Abstract

Background: The medication management process in resident care homes for the elderly (RCHEs) is complex and can be labor intensive. In 2019, a nongovernment organization led by pharmacists with special interest in informatics developed the SafeMed Medication Management System (SMMS), which is a digital web-based system that integrates electronic medical profiles and medication profiles to revamp the traditional manual medication management process in RCHEs in Hong Kong.

Objective: This study aimed to assess the effectiveness of the SMMS in improving RCHE staff's time efficiency and competencies in the medication management process and how this could potentially reduce human resource costs after its implementation.

Methods: This was a pre–post interventional study conducted from September 2022 to August 2024. Time efficiency was evaluated using time–motion analysis. The time spent on each process—preparing the medication doses, checking the prepared doses, and administering the doses to residents—was evaluated in 10-minute blocks. The mean numbers of doses prepared, checked, and administered were calculated for each block. A three-way ANOVA was used to compare the doses before and after the system implementation. Staff competencies and perceived acceptance of the system were evaluated using a structured survey adapted from the technology acceptance model. An exploratory analysis was conducted to estimate the potential financial savings attributable to SMMS implementation using RCHE staff salary data obtained from publicly available governmental sources. Results from the time–motion analysis were used to estimate the cost per dose prepared by RCHE staff before and after the system implementation.

Results: Forty-one RCHEs implemented the SMMS, serving a total of 3911 residents. The time–motion analysis (n=6 RCHEs) revealed that the mean (SD) number of doses significantly increased in 10-minute blocks after system implementation (medication preparation: 25, SD 14 to 49, SD 15 doses; medication checking: 21, SD 6 to 85, SD 33 doses; medication administration: 9, SD 1 to 16, SD 6 doses). The overall mean number of doses handled across all processes combined was significantly higher after implementation (18.9 vs 51.9 doses, P=.02). RCHE staff (n=392) reported significantly improved competencies in entering and accessing residents' records and preparing, checking, and administering medications after the system implementation (all P<.001). The estimated cost of managing one dose of medication dropped substantially from HKD 2.00 (US $0.25) before to HKD 0.74 (US $0.09) after system implementation. If fully implemented in all RCHEs across Hong Kong, the daily human resource cost associated with the medication management process could potentially be reduced from HKD 2,574,000 (US $330,000) to HKD 952,380 (US $122,100).

Conclusions: The time–motion analysis and quantitative survey findings suggest that digital technology combined with automation can improve staff's time efficiency and competencies and promote human resource cost-saving in the medication management process. Future work should evaluate the long-term impact of this system on medication safety and its cost-effectiveness in RCHEs.

(JMIR Aging 2025;8:e79262) doi:10.2196/79262

XSL•FO
RenderX

KEYWORDS

medication management; residential care homes; electronic medication records; time motion study; cost saving; older adults; nursing homes

Introduction

Hong Kong has a rapidly aging population. From 1997 to 2023, the number of people aged 65 years and above more than doubled to 1.68 million, representing 22.4% of the total population [1]. Currently, there are around 74,900 residential care places across the city, with occupancy of approximately 58,500 residents [2]. In addition to adopting a multipronged approach to increase the number of resident care homes for the elderly (RCHEs), improving the standard of health care in RCHEs has become a top priority in Hong Kong.

The working environment in RCHEs is often busy and characterized by a heavy workload [3,4]. The dynamic nature of the workplace necessitates that staff prioritize tasks and perform multiple activities simultaneously while performing medication management [3,4]. In Hong Kong, medication management in RCHEs is mostly handled manually, making the process highly inefficient and labor-intensive. The RCHE residents receive regular medical follow-ups from the public hospitals and clinics under the Hong Kong Hospital Authority (HA). However, the electronic prescription records of HA are not synced with or integrated into the medication administration records in RCHEs. Therefore, the medication management tasks in RCHEs often include manually transcribing, preparing, dispensing, and checking medications, as well as managing documentation on medication administration records [5,6]. According to the literature, various medication errors occur in RCHEs, including errors in preparing drugs, their quantities, and the timing of dispensing medications [7-10]. Such errors could cause serious medication-related problems to RCHE residents, who tend to be vulnerable because of functional decline, multimorbidity, and polypharmacy [11-13]. There is therefore a crucial need to improve the efficiency of medication management without compromising patient safety.

Human-related medication errors can be minimized by incorporating IT [14]. For example, the use of electronic medical profiles and electronic medication records has been shown to reduce errors [15,16] by identifying mismatches in medication records and alerting RCHE staff automatically [17]. The use of electronic medication administration records (eMARs) is associated with a significantly lower rate of medication administration errors than the use of paper-based records [16]. Furthermore, the use of technology can potentially improve the efficiency of the medication management workflow as administrative paperwork is replaced electronically [17,18]. Digitalization allows instant access to residents' medication profiles [15], enabling the immediate tracing and checking of records [18]. By leveraging the power of IT, RCHE staff can enhance the safety and efficiency of workflows, thereby minimizing the risk of medication errors and improving patient outcomes.

In 2019, a group of pharmacists specializing in clinical informatics was appointed by a local nongovernmental organization, the Hong Kong Pharmaceutical Care Foundation Ltd (HKPCF), to develop the SafeMed Medication Management system (SMMS) [19]. The SMMS is a digital system that integrates electronic medical profiles and eMARs. A pilot study showed that the SMMS improved time efficiency and reduced medication errors [5]. However, the pilot study only included two RCHEs and did not evaluate the system from the perspectives of RCHE staff. Hence, the findings of the pilot study might not be generalizable to other RCHEs, and the study lacked a crucial perspective on the system's usability and effectiveness from the standpoint of RCHE staff who would be using it daily.

The primary objectives of this study were to evaluate the system's effects in two key areas: (1) to assess the effectiveness of the SMMS in improving time efficiency in the medication management process and (2) to examine the RCHE staff's competencies and perceived acceptance of the SMMS. To provide a financial perspective on medication management, an exploratory analysis was also conducted to estimate the potential cost savings in human resources after the implementation of the SMMS.

Methods

Study Design

This was a pre–post interventional study conducted in the participating RCHEs from January 2022 to August 2024. An invitation was extended to all RCHEs across various geographical regions of Hong Kong to participate in the Integrated Old Age Home Medication Management Program initiated by the HKPCF. The eligible RCHEs had to meet the SMMS criteria, that is, they (1) had a Wi-Fi network; (2) were receiving continued operational funding from the government, the private sector, or philanthropic organizations; and (3) consented in writing to provide deidentified data for evaluation and research purposes. All participating RCHEs were included in the RCHE staff survey, with a subset being further recruited for the time–motion analysis.

Ethical Considerations

The study was performed in accordance with the Declaration of Helsinki and was approved by the Survey and Behavioral Research Ethics Committee of the Chinese University of Hong Kong (reference 033 - 22). The ethics board agreed to waive written informed consent from each RCHE resident individually as one of the terms and conditions in the service agreement. RCHEs receiving the service had agreed to the collection and use of anonymized medication records available in SMMS, prior to the implementation of the Integrated Old Age Home Medication Management Programme. Verbal consent (Table S1 in Multimedia Appendix 1) was further obtained from the RCHE's superintendents and participating RCHE staff prior to participating in the time–motion analysis. HKPCF pharmacists provided the RCHE staff with regular training and technical

XSL•FO
RenderX

assistance in using the SMMS. No compensation was given to the participants.

Intervention

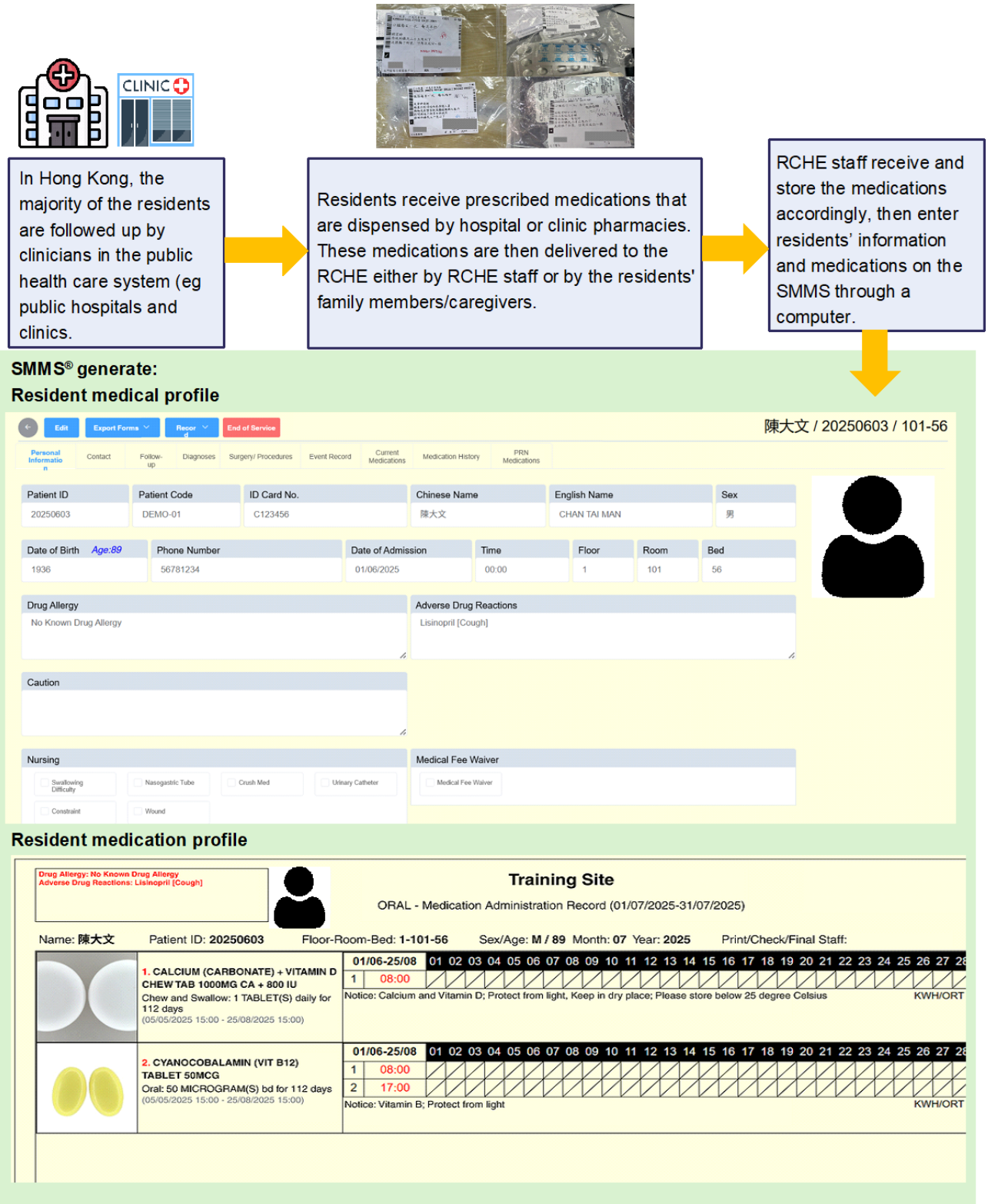
The SMMS is a digital web-based system that integrates electronic medical profiles and eMARs. In Hong Kong, RCHE residents primarily receive follow-up care from clinicians at public hospitals or general outpatient clinics (GOPCs). These hospitals and GOPCs are managed by the HA, providing over 80% of the medical services to the older people and those with chronic illnesses in Hong Kong. To a lesser extent, general practitioners and clinicians from private clinics or hospitals may also prescribe medications for patients. The pharmacies in HA or private institutions dispense the medications to the patients. Notably, the electronic medical and medication profiles in both the public health care system and private institutions are currently not integrated with the medication management system in the RCHEs.

Figures 1-4 summarize the medication management procedures with SMMS. After each visit, the medications were collected by RCHE staff, who also entered residents' information and medications into the database of the SMMS (Figure 1). A new profile was created for each resident in the SMMS. Their profiles were updated by RCHE staff after each follow-up at hospitals or GOPCs. During the medication preparation process, RCHE staff checked and verified the medications against the images shown on the electronic tablet (Figure 2). The medications are then administered to the residents after the RCHE staff verified the residents' identifiers and medications on the electronic medication records on the tablet (Figure 3). When administering injectable medications, RCHE staff would check the previous injection site and record the current injection site on the tablet, as well as select the appropriate site for the next injection (Figure 4).

The SMMS aims to enhance the efficiency of and reduce potential medication errors in the medication management process (Table 1). It includes all medications commonly prescribed by clinicians from the HA. The system's detailed drug database is regularly maintained by HKPCF pharmacists. The medication details available in the database include trade name, dosage form, strength, legal classification, and images of the medications. RCHE staff can verify the medications by comparing with the images provided by the system. Additionally, the system can generate administration schedules so that the medications can be delivered to residents in a real-time manner. It also supports electronic signatures for medication checking, preparation, and administration, eliminating the need for paper-based records.

XSL•FO
RenderX

Figure 1. Use of the SMMS in the medication management workflow: generating a medical profile and medication profile of resident, After receiving the medications, RCHE staff enter the information into SMMS. It is able to generate a medical and medication profile for residents with a standardized format. The medication profile is supported by a drug database with drug images by Hong Kong registration number. The database is maintained by pharmacists and covers the majority of medications from the Hong Kong Hospital Authority. RCHE: resident care homes for the elderly; SMMS: SafeMed Medication Management System.



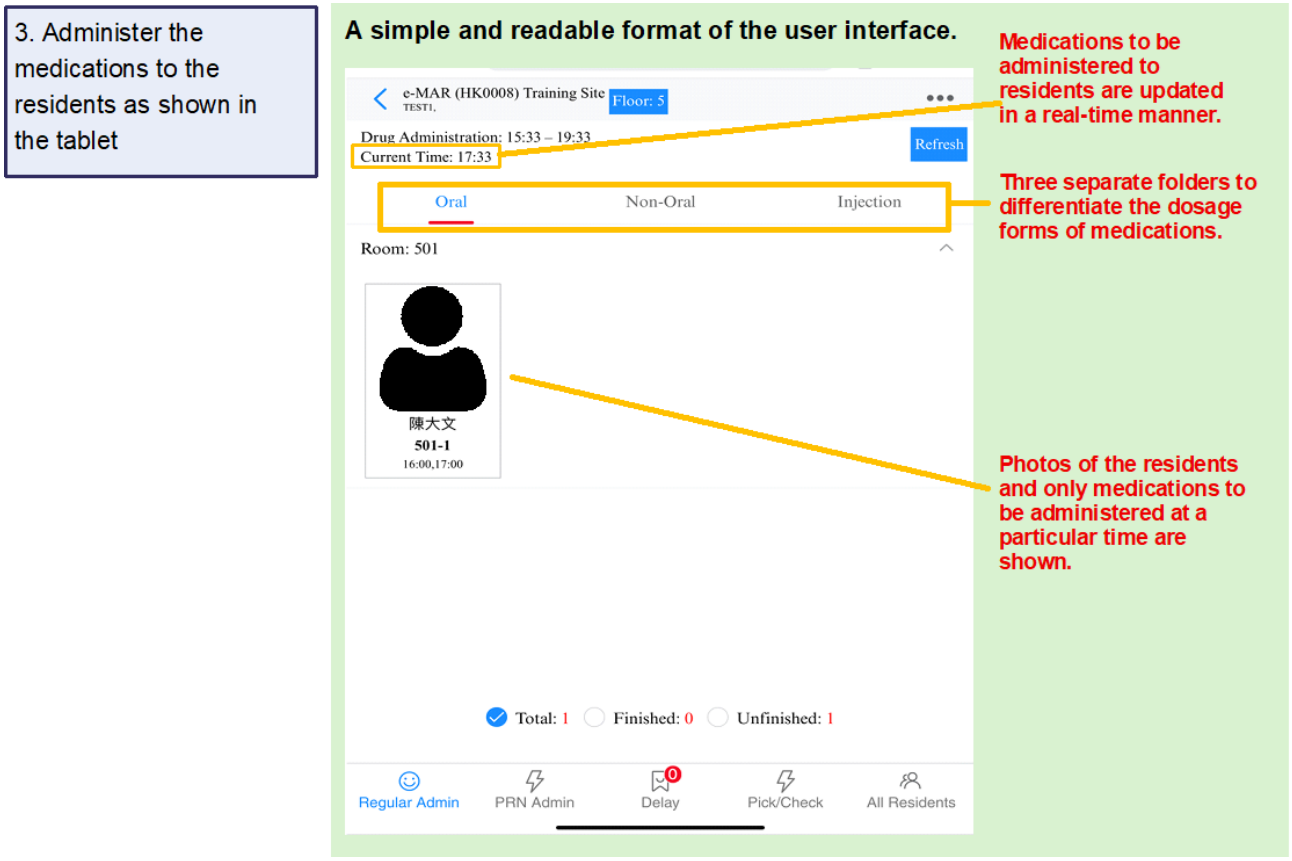
XSL•FO
RenderX

Figure 2. Use of the SMMS in the medication management workflow: conducting medication preparation and checking. RCHE staff prepare and check the medications by using a tablet computer. They can verify the medications by comparing them with the images provided by the system. An electronic signature will be generated automatically after clicking “complete.” eMAR: electronic medication administration record; SMMS: SafeMed Medication Management System.



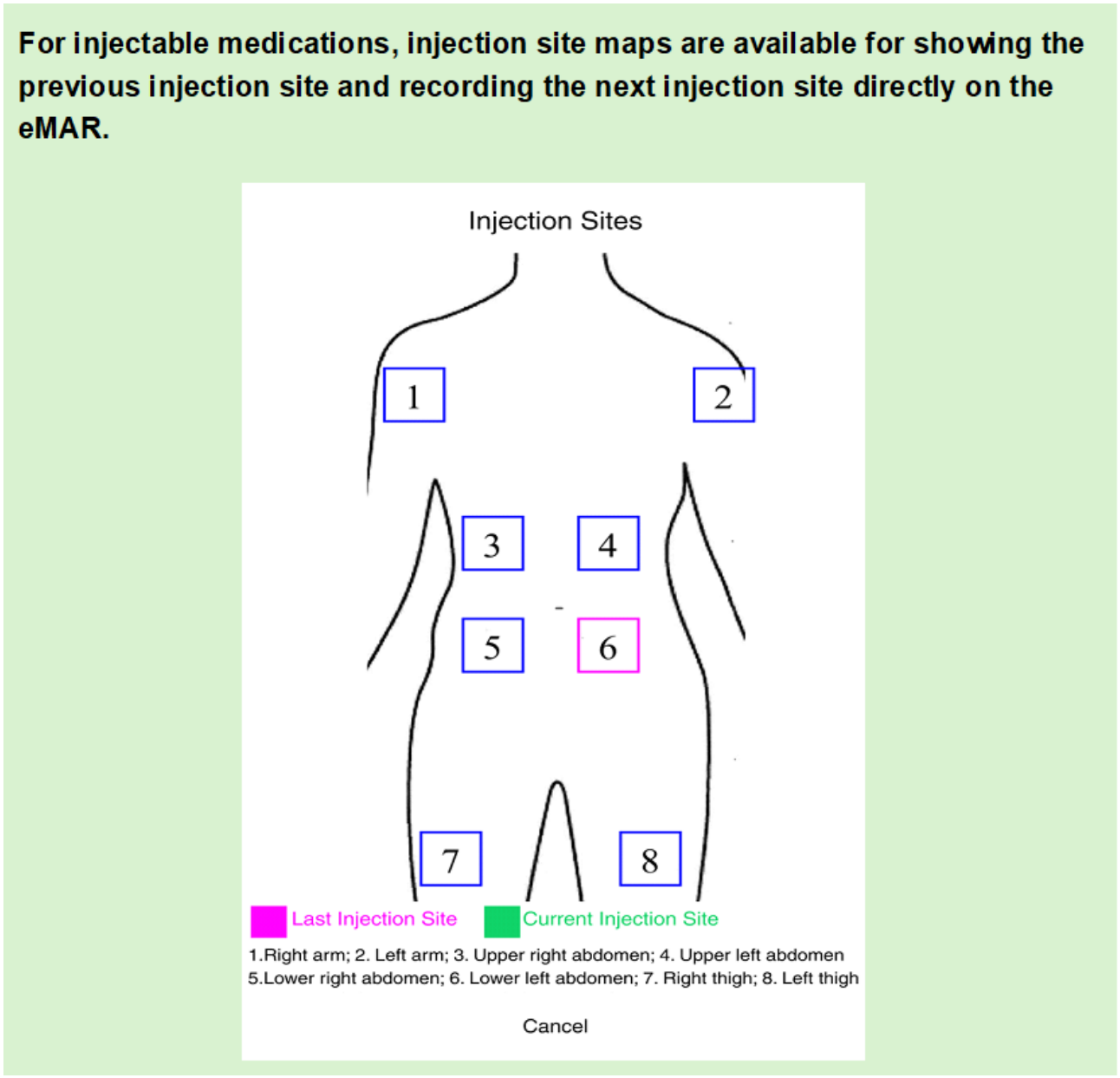
XSL•FO
RenderX

Figure 3. Use of the SMMS in the medication management workflow: conducting medication administration. RCHE staff use the tablet to verify the residents' details and medications before administering them to the residents. There are three separate folders (oral, nonoral, and injection) to differentiate the dosage forms of medications. RCHE staff can focus on each of the three routes, as each route may require different grades of RCHE staff to execute or supervise the administration. Photos of the resident and medications to be administered at that timeslot will be shown. eMAR: electronic medication administration record; RCHE: resident care homes for the elderly; SMMS: SafeMed Medication Management System.



XSL•FO
RenderX

Figure 4. Use of the SMMS in the medication management workflow: conducting medication administration of injectable medications. When administering injectable medications, RCHE staff record the injection site from the tablet to select the appropriate site for the next injection. eMAR: electronic medication administration record; RCHE: resident care homes for the elderly; SMMS: SafeMed Medication Management System.



XSL•FO
RenderX

Table . Features and advantages of the SMMS in the medication management workflow.

System features	Advantages
Data entry	
A comprehensive drug database with drug codes in the public health care system	• Minimizes errors resulting from the manual data entry of drug names and doses.
Able to enter complex regimens and calculate the dispensing quantity automatically (eg, Warfarin 2.5 mg daily on odd days and 3 mg daily on even days)	• Reduces errors and saves time by eliminating the need to manually calculate the dispensing quantity.
Automatically records the responsible personnel for each update in the resident profile and eMAR[a] in the system	• Enhances traceability and accountability.
Residents' records	
A simple and readable format	• Improves efficiency by reducing the time to search from paper MARs[b].
Captures current medications from different prescribing sources	• Quickly reviews resident's past medication history. • Improves efficiency and reduces errors by flagging up repeat prescriptions from residents (older adults with multimorbidity who consult multiple specialties from different specialty clinics [19]); thus, duplicate regimens may occur between different specialties even at the same health care institutions [19].
Real-time updating of the eMAR after any updates are made to the resident's record	• Improves the accuracy of checking and administration relative to the outdated paper MARs, which staff may need more time to check one by one or may even forget to update.
Medication preparation and checking	
A simple and readable format on tablets	• Improves efficiency (reduces the time to search from paper MARs) and potentially reduces errors.
Provides images of medications and administration quantity	• Improves efficiency and reduces errors as the RCHE[c] staff can check the appearance and view the dispensing quantity of the medication immediately.
Real-time electronic signature	• Improves accountability and saves time by eliminating the need to sign paper MARs one by one.
eMARs	
A simple and readable format on tablets	• Improves efficiency (saves time by eliminating the need to search from paper MARs) and potentially reduces errors.
Real-time updates of medications to be administered to residents	• Improves efficiency and reduces the chances of forgetting to update the paper MARs.
Shows only medications to be administered at a particular time during the time of drug administration	• Improves efficiency as traditionally RCHE staff needed to search the paper MARs one by one to determine who needed to take medications in a particular time period. • Helps to reduce errors in the timely administration of medications.
Real-time notification of late administration	• Reduces errors in the timely administration of medications.
Provides photos of the residents and medications	• Improves efficiency and reduces errors as the RCHE staff do not need to search the paper records one by one to identify the residents and medications.
Real-time electronic signature	• Improves accountability and saves time by eliminating the need to sign the paper MARs one by one.
Group medications of different dosage forms under separate folders, including oral, nonoral, and injectable medications	• Improves efficiency as each administration route may require different grades of nursing personnel to execute or supervise the administration.

XSL•FO
RenderX

System features	Advantages
Displays the last administration site and allows for selecting the current administration site for injectable medications and transdermal patches	• Improves safety by facilitating administration site selection and rotation.

[a]eMAR: electronic medication administration record.

[b]MAR: medication administration record.

[c]RCHE: residential care home for the elderly.

Outcomes

Time Efficiency

Time efficiency was evaluated using the time–motion analysis method, which has been used to analyze the efficiency of workflow in RCHEs [5,20]. A detailed description and sample size calculation for the time–motion analysis is presented in Table S1 in Multimedia Appendix 1.

To summarize, the investigators (HCC, PLWAD, and CHW) recorded videos of RCHE staff performing their usual duties in medication preparation, checking, and administration before and after the implementation of the SMMS. The recording protocol was developed based on recommendations from the literature [20] and had been published in our previous study [5]. A breakdown of the specific steps of the recording is presented in Table S2 in Multimedia Appendix 1.

Staff Competencies and Perceived Acceptance of the SMMS

The RCHE staff competencies in completing medication management procedures and their acceptance of the system were evaluated using a structured survey, which was developed based on the technology acceptance model, one of the most widely applied models to evaluate consumer acceptability of IT [21]. The model posits that perceived usefulness and perceived ease of use are important factors determining whether a newly introduced technology would be accepted by its potential users [21]. The questionnaire items were adapted from another local study [22] and reviewed by clinical pharmacists (SCC, HCC, and KCL) and survey methodologists (YTC and TTNL). A pilot study of the questionnaire was conducted with five RCHE staff, and the questionnaire was modified accordingly.

The competency part of the questionnaire comprised 4 domains: entering residents' information, accessing residents' records, preparing and checking medication, and administering medication. The RCHE staff rated their competencies in completing the tasks in the medication management process on a 10-point scale (−5=not able to complete at all, 0=same as traditional practice, and 5=able to complete). A higher score indicates a greater competency in completing the task. The competency survey was distributed to the RCHE staff a maximum of 12 weeks prior to the implementation of the program. RCHE staff was recruited through convenience sampling with the support of the RCHE management. The RCHE staff were asked to self-administer the survey. Each RCHE staff member was assigned a code to enable matching for the postintervention survey, while maintaining anonymity. Only the RCHE management has access to the identified matching list.

The acceptance part of the questionnaire comprised 4 domains: ease of use and technical support, perceived benefits to residents, perceived benefits to RCHE staff, and perceived support from RCHE management. The RCHE staff rated their acceptance on a 6-point scale (1=strongly disagree and 6=strongly agree). A higher score indicates a higher acceptance of the system. At the end, RCHE staff also provided an overall satisfaction score on a 10-point scale. The acceptance part of the questionnaire was only completed at the postintervention assessment.

The internal reliability of the scale was evaluated using Cronbach α and the item–total correlations based on 156 surveys conducted from September 2022 to February 2023. Cronbach α [23] and corrected item–total correlations [24] were calculated to measure the strength of the relationship between each item and the total score of the scale (Table S3 in Multimedia Appendix 1). The internal reliability for both surveys showed satisfactory psychometric properties.

Predictors of Staff Competencies and Perceived Acceptance

Data were collected on individual characteristics, namely, age, gender, working experiences, and qualifications [25], as well as organizational factors, namely, the financial model of operation and number of staff in each RCHE [16,26,27]. The organizational data were retrieved from the Hong Kong Government Social Welfare Department website [28].

Potential Cost Savings

An exploratory analysis was conducted to estimate the potential financial savings attributable to SMMS implementation using the official nursing staff salary data obtained from the Master Pay Scale of the Social Welfare Department's (MPS) Salary Scale of common posts in nongovernmental organizations effective from April 1, 2024 [29]. The average monthly salary was calculated to be HKD 39,922.50 (US $5118) based on the midpoint salaries for registered nurses (MPS point 20; HKD 44,765 [US $5740]) and enrolled nurses (MPS point 15; HKD 35,080 [US $4498]). Assuming a standard working schedule of 176 hours per month, equivalent to a 44-hour work week, the average hourly wage was calculated to be HKD 227 (US $29), that is, approximately HKD 38 (US $4.9) per 10 minutes. The time–motion analysis provided the mean number of doses processed per 10 minutes before and after SMMS implementation and was used with the calculated salary per 10-minute period (HKD 38 [US $4.9]) to estimate the cost per dose prepared by RCHE staff. The difference in cost per dose before and after the system implementation was used to estimate

XSL•FO
RenderX

potential labor cost savings associated with improved time efficiency in the medication management process.

Statistical Analysis

The characteristics of the participating RCHEs and RCHE staff were summarized using descriptive statistics. The numbers of doses prepared, checked, and administered before and after the system implementation were stratified by implementation status and analyzed using a 3-way ANOVA. One sample 2-tailed *t* test (reference value=0, indicating no change) was used to compare the changes in staff's competency before versus after the system implementation. The survey results were analyzed using the Mann-Whitney *U* test (categorical variables) and the Spearman correlation test (continuous variables) to identify individual and organizational factors associated with perceived competency and acceptance.

All analyses were performed using SPSS software, version 25. The results were considered statistically significant at *P* values of ≤.05.

Results

Baseline Characteristics of the RCHEs and RCHE Staff

Forty-one RCHEs participated in this study (Table 2). Nearly 90% (n=36) of the participating RCHEs were care and attention homes for older people. Twenty-five (61%) of them were subvented homes. Approximately 75% (n=31) of the participating RCHEs had more than 40 residents who were on at least one oral medication. The average RCHE nursing staff to resident ratio was 1:10. Regarding the baseline characteristics of the 3911 residents, 61.2% (n=2722) of them were female, their mean age was 83.8 years, and the median number of oral medications per resident was 11.

XSL•FO
RenderX

Table . Baseline characteristics of the participating RCHEs, residents, and RCHE staff.

Characteristics	Sample
RCHEs[a] (n=41)	
Care level[b], n (%)	
Care and attention home for the older people	36 (87.8)
Others	5 (12.2)
Mode of operation, n (%)	
Subvented home[c]	25 (61.0)
Self-financing home or private home	14 (34.1)
Contract home (with both subsidized and nonsubsidized places)	2 (4.9)
Number of residents in each RCHE with at least one oral medication, n (%)	
1 - 40 residents	10 (24.4)
41 - 80 residents	9 (22.0)
≥81 residents	22 (53.6)
Average staff-to-resident ratio	1:2
Average nursing staff-to-resident ratio	1:10
Residents (n=3,911)	
Age of residents (y), mean (SD)	83.8 (12.3)
Gender, n (%)	
Male	1519 (38.8)
Female	2722 (61.2)
Oral medications per residents, median (IQR)	11 (8-15)
RCHE staff (n=392)	
Gender, n (%)	
Male	45 (11.5)
Female	278 (70.9)
Prefer not to answer	69 (17.6)
Age (y), n (%)	
19 - 39 y	196 (50.0)
≥40 y	180 (45.9)
Prefer not to answer	16 (4.1)
Job nature, n (%)	
Full-time	329 (83.9)
Part-time	29 (7.4)
Mixed	34 (8.7)
RCHE staff working experience, n (%)	
<10 y	247 (63.0)
≥10 y	136 (34.7)
Missing	9 (2.3)
Working hours per day, median (IQR)	8 (5-10)
Qualification, n (%)	
Registered nurses or enrolled nurses	196 (50.0)
Health care workers or dispensers	175 (44.6)

XSL•FO
RenderX

Characteristics	Sample
Others	21 (5.4)
Duration between preimplementation assessment and SMMS implementation, median (IQR)	37 (4–57) days
Duration between SMMS implementation and postimplementation assessment, median (IQR)	253 (195–353) days

[a]RCHEs: residential care homes for the elderly.

[b]Care and attention homes for older people refer to RCHEs that provide personal care to older people with moderate impairment who are unable to live at home.

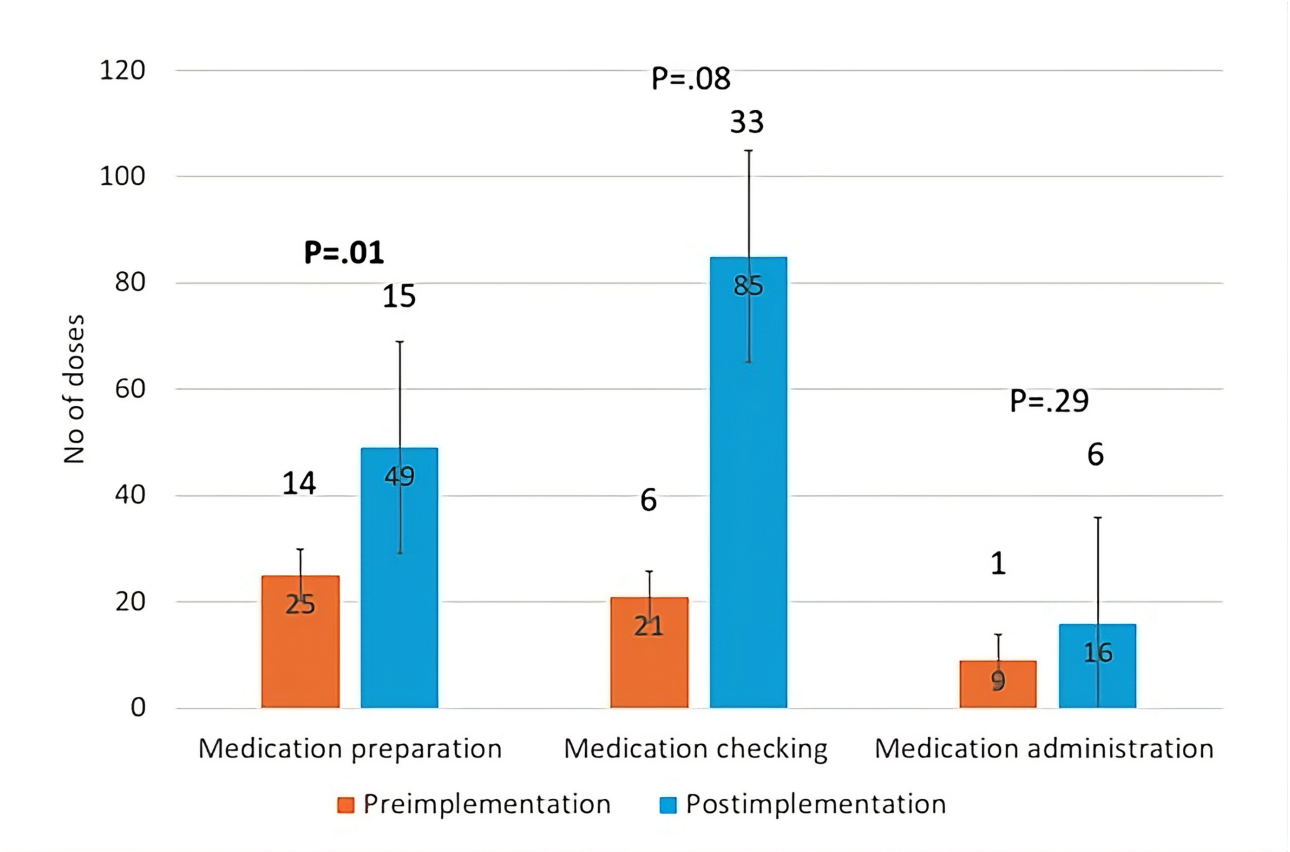
[c]Subvented homes refer to RCHEs that are operated by nongovernmental organizations and are subsidized by the government.

In the 41 participating RCHEs, 392 RCHE staff completed the pre–post surveys (response rate: 65%, loss-to-follow-up rate: 34%; Table 2). The median time between preintervention survey completion and system implementation was 37 days, while the median time for postintervention survey completion since system implementation was 253 days. Among the participants, 70.9% (n=278) were female. Half of the participants were between 19 and 39 years of age and were qualified as registered nurses or enrolled nurses. Additionally, 63% of the participants had less than 10 years of experience working in RCHEs.

Time Efficiency

Time–motion data from 770 minutes of recording were collected from six participating RCHEs. The average number of doses prepared, checked, and administered in 10-minute blocks all increased following the system implementation. Specifically, medication preparation increased significantly from 25 (SD 14) to 49 (SD 15) doses (*P*=.01; Figure 5). Medication checking increased marginally from 21 (SD 6) to 85 (SD 33) doses (*P*=.08), though the difference was not statistically significant. There was no significant change in the doses administered (9, SD 1 to 16, SD 6 doses; *P*=.29). The detailed breakdown of the time–motion analysis is presented in Table S4 in Multimedia Appendix 1.

Figure 5. The ANOVA results revealed that the overall mean number of doses processed across all processes combined was significantly higher after the system implementation (18.9 vs 51.9 doses, *P*=.02). Additionally, the estimated time required (per 1000 doses) reduced for all processes after the system implementation: the time required for medication preparation decreased from 824.0 to 432.1 min; for medication checking, from 750.0 to 250.8 min; and for medication administration, from 1277.2 to 905.6 min.



XSL•FO
RenderX

RCHE Staff Perceived Competency

Staff competencies in completing the medication management workflow improved significantly after the system implementation. This improvement was observed in the areas of entering and accessing residents' records and in the processes of medication preparation, checking, and administration (all *P*'s<.001; Table S5 in Multimedia Appendix 1).

Staff competencies were not associated with gender, age, qualifications, work experience, or mode of operations (Table S5 in Multimedia Appendix 1).

RCHE Staff Perceived Acceptance

The RCHE staff agreed that the system was easy to use and beneficial to the residents and them for their daily work (Table 3). The mean ratings were between 3.8 and 4.2 out of 6. The overall satisfaction score was 6.9 (SD 1.9) out of 10.

XSL•FO
RenderX

Table . RCHE staff perceived acceptance to the system. A Mann–Whitney test was conducted to compare staff-perceived acceptance by staff-level factors and institutional-level factors.

Domain	Ease of use and technical support[a]		Perceived benefits to residents[a]		Perceived benefits to RCHE[b] staff[a]		Perceived support from RCHE management[a]		Overall satisfaction towards the system[a]	
	Mean (SD)	*P* value	Mean (SD)	*P* value	Mean (SD)	*P* value	Mean (SD)	*P* value	Mean (SD)	*P* value
Overall cohort (points)	3.8 (1.0)		4.2 (1.1)		3.9 (1.1)		4.1 (1.0)		6.9 (1.9)	
Comparison by staff-level factors:										
Qualifications		.85		.86		.54		.77		.02
Registered or enrolled nurses	3.8 (1.0)		4.2 (1.1)		4.0 (1.2)		4.1 (1.0)		6.7 (2.0)	
Health care workers or dispensers	3.9 (1.0)		4.2 (1.1)		4.0 (1.0)		4.2 (1.0)		7.2 (1.8)	
Gender		.24		.18		.29		.34		.02
Male	4.0 (1.1)		4.4 (1.0)		4.1 (1.4)		4.3 (1.2)		7.2 (2.1)	
Female	3.8 (1.1)		4.2 (1.1)		3.9 (1.1)		4.1 (1.0)		6.7 (1.9)	
Age (y)		.09		.22		.13		.29		.44
19 - 39	3.9 (1.1)		4.3 (1.0)		4.1 (1.1)		4.1 (1.1)		6.8 (1.9)	
≥40	3.7 (0.9)		4.1 (1.2)		3.8 (1.1)		4.0 (1.0)		6.9 (2.0)	
RCHEs working experiences		.02		.69		.53		.23		.33
<10 y	3.9 (1.1)		4.2 (1.2)		4.0 (1.1)		4.2 (1.1)		6.9 (2.0)	
≥10 y	3.6 (0.9)		4.2 (1.0)		3.9 (1.1)		4.0 (1.0)		6.8 (1.8)	
Comparison by Institutional-level factors										
Mode of operation		.13		.44		<.001		.08		.25
Subvented home	3.9 (1.0)		4.2 (1.1)		4.1 (1.0)		4.2 (1.0)		6.8 (1.8)	
Self-financing home, private home, or contract home	3.6 (1.2)		4.1 (1.2)		3.4 (1.2)		3.8 (1.1)		7.0 (2.2)	
Number of residents who have at least one medication in each RCHE		.28		.63		.42		.36		.02
1 - 40	3.8 (1.1)		3.9 (1.2)		3.8 (1.3)		3.8 (1.0)		7.3 (2.0)	
41 - 80	3.7 (1.1)		4.2 (1.0)		3.8 (1.1)		4.1 (1.1)		6.5 (1.9)	

XSL•FO
RenderX

Domain	Ease of use and technical support[a]		Perceived benefits to residents[a]		Perceived benefits to RCHE[b] staff[a]		Perceived support from RCHE management[a]		Overall satisfaction towards the system[a]	
	Mean (SD)	*P* value	Mean (SD)	*P* value	Mean (SD)	*P* value	Mean (SD)	*P* value	Mean (SD)	*P* value
>80	3.9 (1.0)		4.2 (1.1)		4.0 (1.1)		4.1 (1.0)		6.9 (1.9)	

[a]The RCHE staff rated their acceptance on a 6-point scale (1=strongly disagree and 6=strongly agree) for all subscales, except for "overall satisfaction" on a 10-point scale. A higher score indicates a higher acceptance of the system.

[b]RCHE: residential care home for the elderly.

RCHE staff with ≥10 years of work experience gave a lower rating for the system's ease of use than those with shorter experiences (3.6 [SD 0.9] vs 3.9 [1.1], *P*=.02). The staff from government-funded RCHEs tended to agree that the system benefited their daily work more than staff from RCHEs with other funding models (private, self-financing, or contract homes; 4.1 [1.0] vs 3.4 [1.2], *P*<.001). Staff in RCHEs with a smaller number of residents (1-40) were more satisfied with the system than those in larger RCHEs with >40 residents (7.3 [2.0] vs 6.5 [1.9] vs 6.9 [1.9], *P*=.022).

Potential Cost Savings

The ANOVA results revealed significantly improved efficiency in the medication management process after SMMS implementation and that the mean doses processed per 10-minute interval increased from 18.9 (preimplementation) to 51.9 (postimplementation). Using a calculated labor cost of HKD 38 per 10-minute interval, the estimated cost per dose dropped substantially from HKD 2.00 (US $0.25) before to HKD 0.74 (US $0.09) after SMMS implementation.

The median number of oral medications per RCHE resident is 11. Assuming that each medication is taken, on average, twice daily, each resident requires 22 doses of medications prepared, checked, and administered daily. There are currently 58,500 residents in the RCHEs across Hong Kong [2]; hence, the system, if fully implemented in all RCHEs, could lead to a daily reduction in human resource costs associated with medication management from HKD 2,574,000 (US $330,000) to HKD 952,380 (US 122,100).

Discussion

Summary of Key Findings and Implications

This is the first and largest study in Hong Kong to evaluate the impact of a digital web-based system integrating electronic medical profiles and eMARs on time efficiency and potential human resource cost reduction from the perspectives of RCHE staff. Our results show that medications were more efficiently managed after system implementation across all processes combined, particularly during the medication preparation process, though improvement in medication checking and medication administration processes was marginal and statistically insignificant. The RCHE staff reported improved competency in completing the medication management process and supported the long-term use of the electronic system. The findings demonstrate that the system can address the current limitations in medication management processes in RCHEs.

The time–motion analysis findings show encouraging results as the number of packed and checked doses increased almost 3-fold from before to after the system implementation. This is a conservative estimate as this evaluation was conducted during the early implementation phase; the absolute efficiency would improve further once the RCHEs have had a longer run-in and usage time. Efficiency was improved because the paperless processes had replaced the paper-based processes to eliminate the need for physical signatures on each paper MAR during medication packing and checking. eMARs display only the medications that need preparation at a specific time, significantly reducing the time required to verify the medications (Table S4 in Multimedia Appendix 1). Furthermore, structural previews and automatic calculations for dispensing quantities contributed to increased time efficiency, particularly for complex medication regimens. Notably, the SMMS appeared to have a less significant impact on medication administration, as the considerable variation in this process may be influenced by environmental and patient factors. For example, RCHE staff might still prefer to perform manual and visual inspections of medication and patient identifiers as a safety measure, even though these processes are supported by the features of the SMMS. Nevertheless, the RCHE staff felt satisfied with the system. This demonstrates that the integration of IT into daily routines can enhance the efficiency of the medication management process, allowing staff to contribute more time to frontline nursing care.

The exploratory analysis showed that the professional labor costs associated with medication handling in RCHEs may potentially lead to a daily reduction of HKD 1,621,620 (US $207,900) in human resource costs associated with medication management in RCHEs. We acknowledge that this estimate should be interpreted cautiously, given the assumptions underlying this analysis. These assumptions include the scaling of the program with consistent workflow to all RCHEs in Hong Kong, as well as considerations related to the costs of purchasing, installing, and maintaining the system, along with training expenses, among other factors. However, this exploratory analysis suggests that such financial benefits could allow residential care facilities to redirect available resources toward other critical areas, including enhancing staff training, improving patient care quality, and addressing urgent staffing needs. This is especially encouraging as there is currently a shortage of health care personnel in the primary care setting of Hong Kong. This system is in line with the recently released Hong Kong Primary Healthcare Blueprint [30], which explicitly acknowledges the need for innovative technologies, including automation and artificial intelligence, to alleviate workforce burdens. The SMMS has the potential to directly reduce the workloads of nursing staff and spare their human resources for other duties.

XSL•FO
RenderX

Our results reveal that the RCHE staff reported significantly improved competencies in performing different aspects of the medication management process using the SMMS. Our findings align with recent studies showing that a considerable proportion of RCHE staff hold positive perceptions of the use of electronic systems [31], though this result should be interpreted cautiously, given that the competency assessment tool has not been extensively validated. The time between competency survey completion and system implementation varied widely among the RCHE staff, which may have resulted in inaccuracies and recall bias. However, the competency findings are still important in shaping future optimization of the SMMS as the tool examines staff competency in handling specific features of the intervention based on the widely used technology acceptance model. For example, the RCHE staff acknowledged that the system offers substantial benefits for residents by enhancing the safety of the medication management process through the automatic calculation of dispensing quantity and detection of duplicate prescriptions and late administration. Future work should conduct in-depth studies on the user experience using conceptual implementation science framework specific to technology adoption, such as the Unified Theory of Acceptance and Use of Technology and the Consolidated Framework for Implementation Research to assess the barriers and facilitators of the SMMS implementation success [32,33].

We also acknowledge that the current evaluation did not include safety and reduction in medication errors as an outcome due to the difficulty in capturing self-reported errors and "near-misses" in a large-scale and consistent manner during the preimplementation phase. However, our previous pilot study demonstrated a significant reduction in medication errors (preimplementation: 10/9504 doses, 0.1%; postimplementation: 0/5731 doses; P=.02) [5]. The study also found that the SMMS effectively identified one-third of the residents as taking potentially inappropriate medications based on the Beers criteria [34]. Future work should focus on quantifying the medication safety aspects of the SMMS, particularly in facilitating pharmacists-led medication reviews, establishing referral pathways to promote deprescribing and therapy modifications, and computerized decision-support tools to reduce potentially inappropriate prescriptions and the use of potentially inappropriate medications in this special population.

The literature presents varying findings regarding the factors associated with RCHE staff's acceptance of new technology. For example, some studies have shown that female or older RCHE staff and those with longer work experience might have lower computer literacy and a lower acceptance of new technology [35]. Our study did not identify any significant association between staff demographic characteristics and competency, suggesting that most of the RCHE staff considered the system user-friendly and manageable. However, we found that institution-level factors were associated with staff acceptance of and satisfaction with the system. Importantly, the RCHE staff from self-financing and private RCHEs were less likely than the RCHE staff from subvented homes to agree that the SMMS was beneficial for their daily work. This could be due to a higher expectation of quality and less subsidized support in private RCHEs than in government-funded RCHEs. Small-scale RCHEs (with ≤40 residents) had a higher satisfaction with the system than large-scale RCHEs, probably due to a less demanding workload. Our results indicate a strong association between management support and staff satisfaction with the system, suggesting that effective communication between the management team and frontline staff can facilitate this transition in the early implementation phase. More in-depth interviews with RCHE staff and stakeholders from different operation modes are needed to explore their expectations of and support for the system.

Limitations

This study has several limitations. First, only 6 RCHEs participated in the time–motion study. One reason may be the ongoing COVID-19 pandemic and the infection control measures implemented during the preimplementation phase, which affected the willingness of RCHEs to participate in the time–motion analysis. Nevertheless, this is still considered a reasonably adequate sample size as few studies in the literature have systematically evaluated the effect of automation on time efficiency at an institutional level. Furthermore, as the six participating RCHEs are all subsidized facilities, the findings may be generalizable to other government-subsidized homes, which represent the majority of funding models for RCHEs in Hong Kong. Second, the Hawthorne effect may have influenced the participants' behavior while under observation, potentially leading to the recorded number of doses not accurately reflecting the true situation [36]. Therefore, attempts were made to record more than one staff per nursing home, and the videos were limited to less than 30 minutes per staff to reduce their stress levels. The improvement in outcomes may be a "registration effect," suggesting that the initial improvements could diminish over time. However, considering that the average run-in period between implementation and evaluation was 6 to 12 months, the findings may reflect the program's effects in the intermediate term. A longer follow-up study is needed to assess the sustainability of the program's outcomes.

The potential for sampling bias cannot be ruled out as the participants who completed the survey may have had a greater awareness of the new system implementation than those who did not complete the survey. Although a proportion of the RCHE staff might have missed the survey due to conflicting schedules, we could not determine the reasons for the unwillingness to respond among RCHE staff who explicitly declined to participate in the survey. However, a moderate response rate (65%) is acceptable in most survey studies involving convenience samples. Lastly, unlike our pilot study [5], we did not capture the rate of medication errors, as this exercise likely led to underreporting and a reluctance among RCHE staff to report errors or near-misses. However, it is reasonable to postulate that the rate of medication errors could be effectively minimized due to the additional layers of safety nets outlined in Table 1.

Future Work

To further explore how the digital medication management system could reduce costs compared with traditional workflows, our next step involves investigating the actual cost savings achieved through reduced dispensing and preparation work in

XSL•FO
RenderX

the medication management process. This work also lays the foundation for the potential integration of electronic health and medication records in the public health care system with the SMMS in RCHEs. This integration could enable a seamless, closed-loop medication management system that encompasses prescribing, dispensing, preparing, and administering. Future work could extend the application of the system to other settings that involve complex medication management, such as rehabilitation centers and long-term care facilities for individuals with special needs.

In 2020, the Hong Kong government released "The Smart City Blueprint for Hong Kong 2.0" to expand IT applications to enhance the quality of life of older people [37]. A notable initiative highlighted within the Blueprint is the establishment of the HKD 1 billion Innovation and Technology Fund for Application in Elderly and Rehabilitation Care (I&T Fund), initiated in 2018 [38]. This fund aims to encourage older people care units to adopt recognized gerontechnology solutions, including digital tools that streamline medication management and electronic health records to reduce the burden and pressure on care staff. The SMMS was officially listed as a "Recognized Technology Application Product" in the Social Welfare Department's technology directory, under the category of computerized medication management systems [39]. This government-level recognition indicates government stakeholders' acknowledgment of the potential of the SMMS to enhance the quality of older people care through reducing medication errors and streamlining operational workflows. Our findings provide policymakers with insights to allocate additional funding for implementing IT in RCHE medication management. These results support expanding and integrating "smart living" technologies into various operation modes of RCHEs. Future efforts should focus on collaboration with policymakers to address potential barriers such as hardware or resource supports for the system implementation.

Conclusions

This is the first study in Hong Kong to evaluate the impact of using a digital web-based system integrating eMARs from the perspectives of RCHE staff. Our results show that the system was effective in improving time efficiency and significantly enhancing competencies in medication preparation, checking, and administration. This highlights the importance of implementing electronic systems in RCHEs, guiding relevant stakeholders and policymakers tasked with enhancing the safety and efficiency of medication management for the aging population.

Acknowledgments

We would like to acknowledge all participating nursing homes and staff for their contributions to this study.

Funding

The authors acknowledge the Vera Ruttonjee Desai Charitable Fund for providing financial support for the implementation of the medication management system. The funder played no role in investigating, analyzing, or interpreting the results of this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional information on methodology and results.
[DOCX File, 45 KB - aging_v8i1e79262_app1.docx]

References

1. Residential care services for the elderly in Hong Kong. Research Office Legislative Counsel Secretariat: Welfare Services. URL: https://app7.legco.gov.hk/rpdb/en/uploads/2024/ISSH/ISSH08_2024_20240517_en.pdf [accessed 2025-05-28]
2. Highlights of the year 2021-22: residential care services for the elderly. Social Welfare Department. The Government of the Hong Kong Special Administrative Region. URL: https://www.swd.gov.hk/en/pubsvc/elderly/elderly_info/elderly_ah_sps/elderlyah/rcse/rcse_2021-2022 [accessed 2025-05-28]
3. Barber ND, Alldred DP, Raynor DK, et al. Care homes' use of medicines study: prevalence, causes and potential harm of medication errors in care homes for older people. Qual Saf Health Care 2009 Oct;18(5):341-346. [doi: 10.1136/qshc.2009.034231] [Medline: 19812095]
4. Odberg KR, Hansen BS, Aase K, Wangensteen S. Medication administration and interruptions in nursing homes: a qualitative observational study. J Clin Nurs 2018 Mar;27(5-6):1113-1124. [doi: 10.1111/jocn.14138] [Medline: 29076582]
5. So KH, Ting CW, Lee CP, Lam TTN, Chiang SC, Cheung YT. Medication management service for old age homes in Hong Kong using information technology, automation technology, and the internet of things: pre-post interventional study. JMIR Med Inform 2021 Feb 10;9(2):e24280. [doi: 10.2196/24280] [Medline: 33565993]
6. Wong A, Chan P, Chan E, Lam M, Wong I. Pharmaceutical services of elderly patients in Hong Kong: scoping study to identify and analyse the nature of pharmaceutical services provided to the elderly in Hong Kong. 2013 Jan 1 URL: http:/

XSL•FO
RenderX

/www.ps.org.hk/uploads/files/Pharmaceutical%20Services%20of%20Elderly%20Patients%20in%20Hong%20Kong.pdf [accessed 2025-09-04]
7. Crespin DJ, Modi AV, Wei D, et al. Repeat medication errors in nursing homes: contributing factors and their association with patient harm. Am J Geriatr Pharmacother 2010 Jun;8(3):258-270. [doi: 10.1016/j.amjopharm.2010.05.005] [Medline: 20624615]
8. Parsons C, Lapane K, Kerse N, Hughes C. Prescribing for older people in nursing homes: a review of the key issues. Int J Older People Nurs 2011 Mar;6(1):45-54. [doi: 10.1111/j.1748-3743.2010.00264.x] [Medline: 21303465]
9. Esi Owusu Agyemang R, While A. Medication errors: types, causes and impact on nursing practice. Br J Nurs 2010;19(6):380-385. [doi: 10.12968/bjon.2010.19.6.47237] [Medline: 20335918]
10. Prasanna SMS, Cader TSB, Sabalingam S, Shanika LGT, Samaranayake NR. Are medications safely used by residents in elderly care homes? - A multi-centre observational study from Sri Lanka. PLoS ONE 2020;15(6):e0233486. [doi: 10.1371/journal.pone.0233486] [Medline: 32497110]
11. Bao Y, Shao H, Bishop TF, Schackman BR, Bruce ML. Inappropriate medication in a national sample of US elderly patients receiving home health care. J Gen Intern Med 2012 Mar;27(3):304-310. [doi: 10.1007/s11606-011-1905-4] [Medline: 21975822]
12. Ayani N, Oya N, Kitaoka R, et al. Epidemiology of adverse drug events and medication errors in four nursing homes in Japan: the Japan Adverse Drug Events (JADE) Study. BMJ Qual Saf 2022 Dec;31(12):878-887. [doi: 10.1136/bmjqs-2021-014280] [Medline: 35450935]
13. O Riordan D, Aubert CE, Walsh KA, et al. Prevalence of potentially inappropriate prescribing in a subpopulation of older European clinical trial participants: a cross-sectional study. BMJ Open 2018 Mar 22;8(3):e019003. [doi: 10.1136/bmjopen-2017-019003] [Medline: 29567842]
14. Führling C, Maas R. Extended pharmacist assessment of medication safety for nursing home residents - a cross-sectional and prospective study. J Clin Med 2022 Nov 7;11(21):21. [doi: 10.3390/jcm11216602] [Medline: 36362830]
15. Ausserhofer D, Favez L, Simon M, Zúñiga F. Electronic health record use in Swiss nursing homes and its association with implicit rationing of nursing care documentation: multicenter cross-sectional survey study. JMIR Med Inform 2021 Mar 2;9(3):e22974. [doi: 10.2196/22974] [Medline: 33650983]
16. Alenius M, Graf P. Use of electronic medication administration records to reduce perceived stress and risk of medication errors in nursing homes. Comput Inform Nurs 2016 Jul;34(7):297-302. [doi: 10.1097/CIN.0000000000000245] [Medline: 27270628]
17. Qian S, Yu P, Hailey DM. The impact of electronic medication administration records in a residential aged care home. Int J Med Inform 2015 Nov;84(11):966-973. [doi: 10.1016/j.ijmedinf.2015.08.002] [Medline: 26358850]
18. Karnehed S, Erlandsson LK, Norell Pejner M. Nurses' perspectives on an electronic medication administration record in home health care: qualitative interview study. JMIR Nurs 2022 Apr 22;5(1):e35363. [doi: 10.2196/35363] [Medline: 35452400]
19. SafeMed Medication Management System. Hong Kong Pharmaceutical Care Foundation. URL: http://www.pcfhk.org/en/smms [accessed 2025-11-21]
20. Lopetegui M, Yen PY, Lai A, Jeffries J, Embi P, Payne P. Time motion studies in healthcare: what are we talking about? J Biomed Inform 2014 Jun;49:292-299. [doi: 10.1016/j.jbi.2014.02.017] [Medline: 24607863]
21. Kim J, Park HA. Development of a health information technology acceptance model using consumers' health behavior intention. J Med Internet Res 2012 Oct 1;14(5):e133. [doi: 10.2196/jmir.2143] [Medline: 23026508]
22. Cheung YT, Lam PH, Lam TTN, Lam HHW, Li CK. Technology acceptance among patients with hemophilia in Hong Kong and their expectations of a mobile health app to promote self-management: survey study. JMIR Form Res 2021 Sep 9;5(9):e27985. [doi: 10.2196/27985] [Medline: 34499034]
23. Bland JM, Altman DG. Statistics notes: Cronbach's alpha. BMJ 1997 Feb 22;314(7080):572-572. [doi: 10.1136/bmj.314.7080.572]
24. Nunnally JC, Bernstein IH. Psychometric Theory, 3rd edition: McGraw-Hill; 1994. URL: https://2024.sci-hub.st/4140/22bb3ccd98afffa4e86a0997355d8049/thorndike1995.pdf [accessed 2025-11-27]
25. Aldosari B, Al-Mansour S, Aldosari H, Alanazi A. Assessment of factors influencing nurses acceptance of electronic medical record in a Saudi Arabia hospital. Inform Med Unlocked 2018;10:82-88. [doi: 10.1016/j.imu.2017.12.007]
26. Davis JA, Brannon D, Whitman MV. Organizational factors associated with the use of information systems in nursing homes. Health Care Manage Rev 2009;34(2):141-151. [doi: 10.1097/HMR.0b013e31819e912f] [Medline: 19322045]
27. Shiells K, Diaz Baquero AA, Štěpánková O, Holmerová I. Staff perspectives on the usability of electronic patient records for planning and delivering dementia care in nursing homes: a multiple case study. BMC Med Inform Decis Mak 2020 Jul 13;20(1):159. [doi: 10.1186/s12911-020-01160-8] [Medline: 32660474]
28. Search for residential care homes for the elderly. Social Welfare Department. 2023. URL: https://elderlyinfo.swd.gov.hk/en/ltc_search/rcs [accessed 2025-11-21]
29. Salary scale of common posts in the non-governmental organisations w.e.f 1.4.2024. Social Welfare Department. The Government of the Hong Kong Special Administrative Region. 2024 Apr 1. URL: https://www.swd.gov.hk/storage/asset/section/586/tc/Salary%20Scale%20of%20Common%20Posts%20w_e_f_%2001_04_2024.pdf [accessed 2025-11-21]

XSL•FO
RenderX

30. LCQ21: implementation situation of the primary healthcare blueprint. The Government of the Hong Kong Special Administrative Region. URL: https://www.info.gov.hk/gia/general/202504/02/P2025040200493.htm [accessed 2025-05-28]
31. Alenius M, Graf P. Use of electronic medication administration records to reduce perceived stress and risk of medication errors in nursing homes. Comput Inform Nurs 2016 Jul;34(7):329. [doi: 10.1097/01.NCN.0000489937.56735.df]
32. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009 Aug 7;4(1):50. [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
33. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view1. MIS Q 2003 Sep 1;27(3):425-478. [doi: 10.2307/30036540]
34. Chau HC, Zhang K, Tai BWB, et al. Analysis of medication management system data to determine potentially inappropriate medication use and hospitalization among older adults living in residential care homes for the elderly population. BMC Geriatr 2025 May 6;25(1):314. [doi: 10.1186/s12877-025-05989-4] [Medline: 40329170]
35. Reed K, Doty DH, May DR. The impact of aging on self-efficacy and computer skill acquisition. J Managerial Issues 2005;17(2):212-228 [FREE Full text]
36. Qian SY, Yu P, Zhang ZY, Hailey DM, Davy PJ, Nelson MI. The work pattern of personal care workers in two Australian nursing homes: a time-motion study. BMC Health Serv Res 2012 Sep 6;12(1):305. [doi: 10.1186/1472-6963-12-305] [Medline: 22953995]
37. Support for the elderly and persons with disabilities and their carers. HK Smart City. 2025. URL: https://www.smartcity.gov.hk/living.html#16 [accessed 2025-11-21]
38. Smart city blueprint for Hong Kong 2.0. Innovation and Technology Bureau, Hong Kong SAR Government. URL: https://www.smartcity.gov.hk [accessed 2025-05-28]
39. Past winners (silver award “safemed medication management matrix”). Hong Kong ICT Awards Digital Policy Office. The Government of the Hong Kong Special Administrative Region. URL: https://www.hkictawards.hk/past-winners [accessed 2025-05-28]

Abbreviations

eMAR: electronic medication administration record
GOPC: general outpatient clinics
HA: Hospital Authority
HKPCF: Hong Kong Pharmaceutical Care Foundation Ltd
MPS: Master Pay Scale of the Social Welfare Department
RCHE: resident care homes for the elderly
SMMS: SafeMed Medication Management System

Edited by A Garcia; submitted 18.06.25; peer-reviewed by R Lim, WV Harten; revised version received 23.10.25; accepted 07.11.25; published 04.12.25.

Please cite as:
Au-Doung PLW, Chau HC, Wong CH, Lau KC, Chan YQ, Yip HL, Hung CN, Tai BWB, Lam TTN, Lee CP, Chiang SC, Cheung YT
Implementation of an Electronic Medication Management System in 41 Residential Care Homes in Hong Kong: Pre–Post Interventional Study
JMIR Aging 2025;8:e79262
URL: https://aging.jmir.org/2025/1/e79262
doi:10.2196/79262

© Phillip Lung Wai Au-Doung, Ho Cheung Chau, Chit Hin Wong, Kai Cheong Lau, Yi Qi Chan, Ho Lam Yip, Chiu Ngai Hung, Bik-Wai Bilvick Tai, Teddy Tai-ning Lam, Chui Ping Lee, Sau Chu Chiang, Yin Ting Cheung. Originally published in JMIR Aging (https://aging.jmir.org), 4.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Use of Technology to Deliver In-Home Aged Care Services: Mixed Methods Study of Australian Staff Perspectives

Baldwin Pok Man Kwan[1], MPH; Marissa Dickins[1,2,3], PhD; Sue Williams[4,5], MPhysio; Frances Batchelor[4,5], PhD; Kerry Hwang[4,6], PhD; Kate Fulford[1], MPharm; Tony Walsh[4]; Helena Jakupovic[1], EMBA; Tanya E Davison[1,3], DPsych(Clin)

[1]Research and Innovation, Silverchain, 6 Sundercombe Street, Osborne Park, Western Australia, Australia

[2]School of Public Health and Preventive Medicine, Faculty of Medicine, Nursing and Health Sciences, Monash University, Clayton, Victoria, Australia

[3]Institute of Health and Wellbeing, Federation University, Ballarat, Victoria, Australia

[4]National Ageing Research Institute, Parkville, Victoria, Australia

[5]Department of Physiotherapy, The University of Melbourne, Parkville, Victoria, Australia

[6]School of Nursing and Midwifery, Monash University, Clayton, Victoria, Australia

Corresponding Author:
Baldwin Pok Man Kwan, MPH
Research and Innovation, Silverchain, 6 Sundercombe Street, Osborne Park, Western Australia, Australia

Abstract

Background: With a global aging population, technology has been proposed as a solution to address the growing demand for services in the in-home aged care sector. Despite the potential of technology, there are difficulties when implementing technology into routine care delivery. There is a lack of evidence regarding the specific factors affecting technology use in the in-home aged care setting from the perspective of the direct care workforce.

Objective: This study aimed to understand in-home aged care staff members' views of (1) the digital enablement potential of direct in-home care tasks, (2) benefits and drawbacks of technology use, and (3) enablers and barriers for technology use in Australian in-home aged care.

Methods: An explanatory sequential mixed methods research design was used, with a cross-sectional survey and semistructured staff interviews. Participants were recruited from in-home aged care staff members working at a national Australian in-home health and aged care organization.

Results: In total, 226 participants completed the survey, and 18 participants completed the interviews. Overall, participants felt that many care tasks within in-home aged care could be digitally enabled, with more than half (56%) of the common direct care tasks identified as being likely to be digitally enabled. Participants also discussed a range of quality of care-, staff-, and organization-related benefits and drawbacks in the use of technology. Finally, participants agreed that most of the researcher-proposed enablers and barriers were important, while suggesting additional enablers and barriers such as client preferences regarding technology use and poor data connectivity.

Conclusions: This study provides insight into staff members' views regarding the use of technology to deliver in-home aged care services. The results could help inform technology developers and in-home aged care providers, providing key information to guide technology implementation into care delivery. Further research is required to ensure that appropriate strategies are available to ensure successful implementation of technology into in-home aged care.

(JMIR Aging 2025;8:e76141) doi:10.2196/76141

KEYWORDS

aged; digitally enabled care; home care services; technology; workforce

Introduction

Globally, the older adult population is expected to continually grow over the next 25 years, with projections showing that the proportion of adults aged 65 years or older will increase to 16% in 2050 from around 10% in 2022 [1]. The World Health Organization has highlighted care at home as a key component in enabling older adults to age in a suitable environment and live a healthy and long life [2]. In Australia, this proportion is expected to be higher than the worldwide average, with 20% of people predicted to be aged 65 and older by 2031 [3]. This will drive a rapid rise in demand for in-home aged care services. In particular, the proportion of older-old Australians (aged 85 years and older), who may require additional services to remain safely at home, is expected to double from 2.1% in 2020 to up

XSL•FO
RenderX

to 4.4% in 2066 [4], which may continue to drive demand for in-home aged care services.

Australian aged care services are generally delivered within a residential aged care facility, for individuals who have moved into aged care homes or the older person's own home, where in-home aged care workers visit to provide care [5]. The number of Australian in-home aged care recipients, who receive a broad range of services, including domestic assistance, care coordination, personal care, social support, nursing, and allied health, to support their care needs, increased by 35% from 2017 to more than 1 million in 2023 [6]. Australia—like other countries—faces significant challenges related to the recruitment and retention of the aged care workforce, as well as funding of services [7]. Novel approaches to service delivery are required to meet the increasing demand for in-home aged care.

Technology has been suggested to have a vital role in improving the quality and transforming in-home aged care services to meet the rising needs of the older population [8]. A potential approach to use technology would be through digitally enabled care, which refers to using technology to augment new or existing care tasks within in-home aged care. A literature review has listed some common aspects of care in the in-home aged care sector that could be digitally enabled, such as falls prevention and detection, pain management, and care need assessments, using novel technologies including artificial intelligence, sensors, and interactive video games [9]. Research has shown that technology within in-home aged care can improve communication and care coordination, support the lone worker on the road and their geographically isolated clients, and enable social interactions for socially isolated clients [10]. For example, an Australian randomized controlled trial found that smart home technology improved social care–related quality of life in individuals receiving in-home aged care services [11]. With a significant proportion of Australian aged care recipients facing complex issues such as social isolation [12], and physical and mental health comorbidities [13,14], technology may serve as a crucial component of the solution to meet their needs.

Recent research has found that the majority of the current Australian in-home aged care workforce has sufficient digital readiness to use technology in the provision of care [15]. Despite this, and the potential benefits of technology within in-home aged care, the implementation and use of technology for aged care provision remain low [16]. Although the health and aged care sectors share many similarities, the aged care sector has lagged behind in the use of technological solutions [5,8]. There are suggestions that barriers such as ethical issues as well as individual, organizational, and system readiness impede the implementation of technology into aged care [9]. In addition, concerns have been raised about the potential impact of technology on the relational aspect of care, loneliness, inequities, and unintended consequences within in-home aged care services [10]. These concerns originated from apprehension surrounding the replacement of face-to-face contact with technology, ineffectiveness of technology to alleviate loneliness, unease around the accessibility of technology for older adults, and unforeseen consequences from lack of staff involvement when implementing technology [10].

A range of reviews have investigated the enablers and barriers of technology use in the health and aged care sectors. Enjoyment of technology, technology usability, personalization of the technology, familiarity with the specific technology, previous technology experience, support for the use of technology, and technology training acted as enablers for technology implementation into care from health or social care recipients' perspective [17,18]. In contrast, technical issues, negative perceptions regarding technology use in care for older adults, limitations of the technology design, privacy concerns, technology usability, cost of technology, and lack of digital literacy were seen as barriers for home or health care recipients [17-19]. Although significant research has been conducted in the broader health, home care, and aged care sectors, there is a lack of evidence regarding the specific enablers and barriers of technology use in the in-home aged care setting from the perspective of the aged care workforce. Given the increasing role of the in-home aged care workers in both using technologies themselves and in supporting their clients to use technologies, this gap in knowledge requires attention. Specific research in the in-home care setting is required given its distinct characteristics, where providers deliver services to a cohort of older adults with unique care needs in their own homes, typically working on their own and without on-hand technical support.

At a time when several countries, including Australia, are developing strategies to increase the use of technologies to support health and aged care services, it is crucial to understand the potential of technology for direct care and the enablers and barriers of technology in the in-home aged care setting. This knowledge is critical for appropriate development of policy and implementation strategies and may also inform the development of future technologies to improve the quality of care provided to older people and enable them to remain safely at home for longer. Hence, this study aimed to understand in-home aged care staff members' views of (1) the digital enablement potential of direct in-home care tasks, (2) benefits and drawbacks of technology use, and (3) enablers and barriers for technology use in Australian in-home aged care.

Methods

Design

An explanatory sequential mixed methods research design was selected to explore the digital enablement potential of direct care tasks and enablers and barriers for technology use within in-home aged care [20]. This study design was chosen to collect quantitative data first and then use qualitative methodology to explore the quantitative results in more depth [20].

Cross-Sectional Survey

Research Objective

The aim of the survey was to understand staff members' views on (1) the digital enablement potential of direct in-home care tasks, and (2) the importance of enablers and barriers for technology use within in-home aged care.

XSL•FO
RenderX

Sampling, Data Collection, and Measures

The sampling and recruitment procedure for the survey has been described in detail in a previous publication, which reported on a separate component of the survey [15]. In short, using convenience sampling, direct care in-home aged care workers from a large Australian in-home aged care provider were invited to participate in a staff survey between May and August 2023.

Specifically designed questions collected data on participants' demographic characteristics, the digital enablement potential of direct in-home care tasks, and the importance of specific enablers and barriers for technology use within in-home aged care. For this study, direct care tasks were defined as tasks completed by in-home aged care staff that may require visiting a client's home and support aged care recipients' goals and care needs. For digital enablement potential, common direct care tasks for different roles were listed, with separate tasks listed for the key roles found in the Australian in-home aged care sector: managers or team leaders, Home Care Package coordinators or care managers, domestic assistants, care aides or therapy assistants, nurses, and allied health professionals. The list of common care tasks, which included items such as support with personal care, linen services, and clinical assessments, was developed by researchers through mapping of organizational service delivery documents used by a national in-home aged care provider, as well as Australian in-home aged care service guidelines [21-24]. Participants were asked to rate how likely it was that each of the listed direct care tasks could be digitally enabled (with existing or future technology) using a Likert scale from 1 = extremely unlikely, 2 = unlikely, 3 = likely, and 4 = extremely likely. Higher scores indicated a higher digital enablement potential.

A list of potential enablers and barriers was drafted by the research team through a comprehensive literature search focusing on potential enablers and barriers of technology use in home care settings and the research team's clinician researchers' lived experience working within in-home aged care. For example, "having staff involvement in how technology is implemented" and "cost to client" were listed as an enabler and barrier of staff technology use within in-home aged care, respectively. The full list of enablers and barriers is available in the "Results" section. Participants rated the importance of each enabler and barrier in technology use on a Likert scale from 0 = don't know, 1 = not at all important, 2 = low importance, 3 = important, and 4 = very important. Higher scores indicated higher importance. In addition, open-ended questions were used to elicit responses from participants regarding additional enablers and barriers related to technology use within in-home aged care.

Data Analysis

Descriptive statistics were generated using frequencies and percentages for categorical variables and means and SDs for continuous variables in Stata (version 18.0 BE; StataCorp). The digital enablement potential of tasks was ranked from highest to lowest for each role based on the mean score of each task. Similarly, the importance of enablers and barriers for technology use was ranked from highest to lowest using the mean score. Participants who rated zero (don't know) for the enablers and barriers were removed from analysis for the specific enabler or barrier. Open-ended questions were coded inductively using applied thematic analysis in NVivo (version 12.0; Lumivero) [25]. Applied thematic analysis is a pragmatic, inductive qualitative data analysis technique [25]. The 5 elements of applied thematic analysis (text segmentation, codebook creation, structural coding, content coding, and theme development) were used during analysis [25].

Semistructured Staff Interviews

Research Objective

The staff interviews aimed to describe (1) the digital enablement potential of direct in-home care tasks, (2) benefits and drawbacks of technology use, and (3) enablers and barriers for technology use within in-home aged care.

Sampling

A purposive sampling approach was used for the semistructured interviews. Current in-home aged care staff from a single provider organization working in South Australia, Victoria, or Western Australia were eligible to participate and recruited in 2 ways. First, individuals who completed the staff survey and opted in for further research were invited to participate in a semistructured interview. Second, managers were contacted to nominate in-home aged care workers from different roles to be invited to participate in the interviews. This approach ensured that participants from a diverse range of roles and locations were recruited to participate in the study. In both cases, invitations were sent to potential participants via email, with the participant information and consent forms. Interested staff were asked to respond to the email, provide a signed consent form, and arrange a suitable time for the interview.

Data Collection

Web-based semistructured interviews were conducted using Microsoft Teams by SW and BK. The interviews included questions related to the digital enablement potential of direct care tasks, benefits and drawbacks of technology use, and enablers and barriers for technology use within in-home aged care. The interview schedule is available in Multimedia Appendix 1. As the research team began to conduct interviews during the survey period, the indicative aggregate survey results were used as interview materials to elicit feedback and discussion from the interviewees. Aggregate responses to open-ended questions in the survey regarding additional enablers and barriers were also presented for discussion. These interview materials are also available in Multimedia Appendix 1. The interviews were transcribed verbatim using Microsoft Teams.

Data Analysis

Data were entered into NVivo and analyzed using applied thematic analysis [25]. Analysis was conducted by BK, with MD conducting double coding of 20% of the transcripts to ensure rigor of coding. The research team familiarized themselves with the data by reading the transcripts. The codebook was initially constructed based on the answers to the survey questions and then further developed iteratively to develop and analyze the relationships between codes and themes.

XSL•FO
RenderX

Positionality Statement

BK is an in-home aged care physiotherapist with qualitative research training. MD and TD are researchers embedded within an in-home aged care provider, with TD having a clinical psychology background. SW, FB, and KH are researchers at a national aging research institute, with SW and FB having clinical experience as physiotherapists in health care and aged care settings. KF has experience working as an in-home aged care pharmacist and developing innovative models of care. TW and HJ are consumers with lived experience of in-home aged care.

Ethical Considerations

Ethics approval was granted by the Silverchain Low Risk Ethics Committee (RG-054). Informed consent was obtained from participants prior to study participation through the provision of a participant information and consent form. All collected data were deidentified during data analysis and project reporting. No financial compensation was given to the participants of the survey. Financial compensation was provided to the operational team to backfill staff participation in the interviews.

Results

Descriptive Statistics

For the survey, 226 participants consented and completed questions relevant to this study. Eighteen participants consented to participate in the interviews. Descriptive statistics of both groups of participants are shown in Table 1. In the survey, 69.5% (157/226) of participants were aged 45 years or older, and 89.8% (203/226) were female. The majority worked in a part-time or casual position (144/226, 63.7%), with a broad range of different roles reported. Across the survey and interviews, most participants worked in metropolitan areas (162/226, 71.7% and 11/18, 61.1%, respectively), with approximately half of them having 5 or fewer years of experience working within in-home aged care (110/226, 48.7% and 10/18, 55.6%, respectively).

XSL•FO
RenderX

Table . Detailed sociodemographic characteristics (frequency, %).

Characteristics	Staff survey (N=226), n (%)	Staff interviews (N=18), n (%)
Total participants	226 (100)	18 (100)
Age group (years)		N/A[a]
18 - 24	6 (2.7)	
25 - 34	21 (9.3)	
35 - 44	42 (18.6)	
45 - 54	71 (31.4)	
55 - 64	75 (33.2)	
65+	11 (4.9)	
Gender		N/A
Male	21 (9.3)	
Female	203 (89.8)	
Nonbinary/third gender	1 (0.4)	
Prefer not to say	1 (0.4)	
Employment type		N/A
Full-time	82 (36.3)	
Part-time	123 (54.4)	
Casual	21 (9.3)	
Location of work		
Metropolitan	162 (71.7)	11 (61.1)
Regional, rural, and remote[b]	N/A	7 (38.9)
Regional	29 (12.8)	N/A
Rural	25 (11.1)	N/A
Remote	10 (4.4)	N/A
Education		N/A
Some secondary schooling	10 (4.4)	
Senior secondary schooling (completed high school or equivalent)	48 (21.2)	
Trade/technical/vocational training or diploma	90 (39.8)	
Bachelor's degree (including Honours)	49 (21.7)	
Postgraduate degree (eg, master's degree, PhD, or professional degree)	16 (7.1)	
Missing	13 (5.8)	
Length of time working within in-home aged care (years)		
0 - 1	35 (15.5)	3 (16.7)
2 - 5	75 (33.2)	7 (38.9)
6 - 10	32 (14.2)	0 (0)
11 - 15	32 (14.2)	1 (5.6)
16 - 20	28 (12.4)	4 (22.2)
21 - 25	9 (4.0)	1 (5.6)
25+	15 (6.6)	2 (11.1)
Main role		
Allied health professional	28 (12.4)	3 (16.7)

XSL•FO
RenderX

Characteristics	Staff survey (N=226), n (%)	Staff interviews (N=18), n (%)
Care aides/therapy assistants	95 (42.0)	3 (16.7)
Enrolled nurses/registered nurses	5 (2.2)	1 (5.6)
Domestic assistants	51 (22.6)	1 (5.6)
Home care package coordinators/care managers	28 (12.4)	6 (33.3)
Team leader/manager	19 (8.4)	4 (22.2)

[a]N/A: not available.

[b]Regional, rural, and remote were categorized together for the purposes of the interviews.

Digital Enablement Potential of Direct In-Home Aged Care Tasks

The mean digital enablement potential of each direct care task is reported in Table 2, with illustrative quotes regarding representative tasks from the interviews presented alongside the quantitative results. Overall, 56% (14/25) of the listed direct care tasks were rated a mean of 3 or above, indicating that participants felt that many direct care tasks could be digitally enabled within in-home aged care. This was reinforced by the interviews, with most interview participants reporting that some direct care tasks could be digitally enabled, but this varied across roles. Notably, participants who worked as a care aide or a therapy assistant ranked all their listed tasks to a mean of below 3 (ranging between 2.22 and 2.82), and participants who worked as a domestic assistant ranked their listed tasks below or just slightly above a mean of 3 (ranging between 2.35 and 3.04), indicating these subgroups of in-home aged care staff felt that their tasks were less likely to be digitally enabled. For example, a participant working as a domestic assistant reported:

> *I see my work as hands on.* [Interview participant 17, Care aide/therapy assistant]

Survey participants in clinical or care coordination roles, such as allied health professionals and care coordinators, rated communication-related tasks in their roles relatively more likely to be digitally enabled than noncommunication tasks, with communication-related tasks with staff or external health care providers rated as the top or second most likely task to be digitally enabled. While managers and team leaders rated communication with team members as the task most likely to be digitally enabled, they also rated client communication as the least likely task to be digitally enabled. In contrast, interview participants reported that communication with clients is currently digitally enabled, using common technology such as text messages or emails.

> *I think a huge one for me [as] the [home care] package coordinator. There's actually emailing . and that's with clients.* [Interview participant 07, Home care package coordinator/care manager]

In addition, in the survey, allied health professionals and nurses rated clinical assessment, which was ranked first or second, more likely to be digitally enabled than clinical treatment, which was ranked last for both professions. However, interview participants voiced a range of different opinions regarding the digital enablement potential of clinical treatments:

> *Nursing treatments. Obviously, that's...face to face.* [Interview participant 09, Home care package coordinator/care manager]

> *In terms of social work...we could deliver a fair amount of education and treatment digitally.* [Interview participant 02, Allied health professional]

While the majority of tasks conducted by domestic assistants were considered unlikely to be digitally enabled, an exception was unaccompanied shopping services, which could be digitally enabled. For example, a participant (occupational therapist) suggested that unaccompanied shopping services could be digitally enabled to empower in-home aged care recipients.

> *The other programs have been helping people with shopping online, so people who are unable to leave home...but with support they can...I have heard good feedback about that...You know the two people are sitting there in front of the computer doing it together, but I think that's that is a good example of an enabling...task [because] you're helping that person to choose their own shopping...so I think it's important.* [Interview participant 11, Team leader/manager]

While many direct care tasks were seen as having potential to be digitally enabled, some participants expressed concern that the quality of the service delivered may be decreased or the service may not achieve its intended purpose:

> *[Virtual occupational therapy] assessments, while they can be done and I have done them, [they're] definitely not as easy as in person and you don't get the same depth and richness of information.* [Interview participant 11, Team leader/manager]

> *It's interesting that...by thinking that just by seeing a face online [during social support services] is going to solve the problem. But as you say, maybe it's more the actual going to the cafe and seeing the world and having that coffee as much as that person that they're with, but it's a combination of everything rather than just the face.* [Interview participant 08, Home care package coordinator/care manager]

XSL•FO
RenderX

Table . Likelihood of digitally enabling direct care task[a] (n=226; ranked from highest to lowest likelihood in each role).

Direct care task	Score, mean (SD)	Representative quotes
Managers/team leads (n=19)		Staff management: "It's not brilliant...just to have one team on a video call and then the others in the room. It would be better to have everyone in the room, but...it's a way that...we can still interact, so it's better than nothing" [Interview participant 09, Home care package coordinator/care manager].
Communicating with team members	3.21 (0.63)	
Staff management	3.16 (0.60)	
Onboarding/training team members	3.16 (0.77)	
Client communication	2.79 (0.63)	
Home care package coordinators/care managers (n=28)		Care coordination: "Technology is really beneficial in that being able to do things like email or...Teams meetings...you can...fit things into a busy schedule a lot easier. And a lot of my clients have kids who are still working but still really [want to] provide input and be collaborative on services. But you know, when I'm working full time and they're working full time, it's hard to kind of find for them to take time off to do whatever, so it can really facilitate carer and family involvement in services as well" [Interview participant 02, Allied health professional].
Liaising with other health professionals	3.50 (0.51)	
Staff management	3.21 (0.79)	
Care coordination	3.11 (0.83)	
Client assessment	3.04 (0.96)	
Domestic assistants (n=51)		General house cleaning: "You need to physically clean a house" [Interview participant 17, Care aide/therapy assistant].
Unaccompanied shopping (delivered to home)	3.04 (0.60)	
Linen services	2.45 (0.73)	
General house cleaning	2.35 (0.80)	
Care aides/therapy assistants (n=95)		Assistance with self-administration of medications: "Medication administration that is in a dispenser and the dispenser is...electronic and it dispenses the tablets at the right time and tells them when to get up. [The client will] know when to take them" [Interview participant 01, Home care package coordinator/care manager].
Assistance with self-administration of medications	2.82 (0.67)	
Support with mobilizing (eg, getting in and out of bed/moving around the house)	2.72 (0.75)	
Support with transport	2.68 (0.73)	
Assistance with allied health therapy	2.67 (0.57)	
Social support	2.54 (0.73)	
Support with personal care (eg, toileting, bathing, dressing, and grooming)	2.22 (0.84)	

XSL•FO
RenderX

Direct care task	Score, mean (SD)	Representative quotes
Enrolled nurses/registered nurses (n=5**)**		Nursing assessment: “We have attempted some remote clinical assessments via [Microsoft] Teams which are OK, but connection drops out sometimes” [Interview participant 09, Home care package coordinator/care manager].
Nursing assessment	3.60 (0.55)	
Liaising with other health professionals	3.40 (0.55)	
Client education	3.20 (0.84)	
Nursing treatment	2.60 (1.14)	
Allied health professionals (n=28)		Allied health assessment: “My allied health assessments, all of my assessments are done on my laptop. I don’t feel like I could go into an assessment, write on paper and then come back out and do that onto my laptop. I feel like that would be such a waste of time” [Interview participant 14, Allied health professional].
Liaising with other health professionals	3.50 (0.58)	
Allied health assessment	3.43 (0.63)	
Client education	3.25 (0.75)	
Allied health treatment	2.89 (0.79)	

[a]Extremely unlikely = 1, unlikely = 2, likely = 3, and extremely likely = 4.

Benefits and Drawbacks of Technology Within In-Home Aged Care

The benefits and drawbacks of technology within in-home aged care were explored in semistructured interviews, with findings summarized in Table 3. Findings were categorized into benefits and drawbacks relating to quality of care, staff, and organization.

XSL•FO
RenderX

Table . Benefits and drawbacks of technology within in-home aged care and representative quotes from semistructured interviews.

Benefits and drawbacks	Representative quotes
Quality of care–related benefits	
Improving availability of care	"Considering how big and remote [Western Australia] is...probably be an advantage just because for them to get to sort of someone else to see it might be next to impossible" [Interview participant 12, Enrolled nurse/registered nurse].
Increasing data available to inform care	"They can actually track to see whether they're actually having [the medication] or not" [Interview participant 16, Domestic assistant].
Meeting client preferences regarding technology-supported care	"Depending on the client and what they are wanting because it is [all] about delivering best care to each and every client and we want to make sure that they're happy and living the life that they want to live, so if it could change, maybe [having technology in their care] would be the way forward" [Interview participant 10, Team leader/manager].
Reducing cost of care for clients	"I think it's very valuable for administrative purposes for doing...reviews to help cut down on time and therefore cost to client" [Interview participant 11, Team leader/manager].
Improving flexibility of scheduling of care	"I guess we should always give the clients the option of whether or they want us like for example, if we're doing...a virtual...session rather than face to face...There will be some clients who, especially if they're time poor and...they might not get up early in the morning they want you to...see them at like 3:30 in the afternoon for example...there might be a bit more flexibility around times if we don't have to then get back to the office and finish work" [Interview participant 05, Allied health professional].
Empowering clients' independence and function	"If we could show them, I think the benefits of some of these things like the banking and things like that and that they can do online shopping...I think the [technology] would make them also feel a bit more independent as well, because I think they have to rely on technology a lot and they don't understand it" [Interview participant 01, Home care package coordinator/care manager].
Enabling clients and carers to be more involved in their care	"I have clients that don't speak English and the family member might [want to] be there to help out...We also have access to interpreters, but often they prefer to have a family member there. The family member might be they could be Interstate...they might be on the other side of town. [You] could even do it like as a 3-way conversation [using technology]" [Interview participant 05, Allied health professional].
Improving client safety and security	"Wearables have been amazing for our clients...We started off people having personal alarms around their neck that they had to press if they've had a fall. Now we've got wearables that will alert family members when clients have had a fall and let them know where they are, so I think technology is assisting with giving us fantastic care for our client group" [Interview participant 04, Home care package coordinator/care manager].
Staff and organization–related benefits	
Helping the organization continually improve	"You know it's going to be a stepping stone for what we're evolving into as a company, as an organization" [Interview participant 10, Team leader/manager].
Improving accessibility to up-to-date information and data	"The software that we use, you live by it in a way because it's got your schedule, it's got every detail that you need to know" [Interview participant 10, Team leader/manager].
Improving flexibility of staff rostering	"Because we've got the roster on our phones, which I find really much, much better...in the sense that we can kind of like change things a little bit without...being ridiculous about it but adjust things so that everything fits nicely on for the day" [Interview participant 18, Care aide/therapy assistant].
Improving staff efficiency	"It's really good to be able to do more regular follow up because...you don't have to be driving everywhere. You're not spread as thin. And again, good for staff shortages" [Interview participant 11, Team leader/manager].

XSL•FO
RenderX

Benefits and drawbacks	Representative quotes
Meeting the needs of on-the-road staff	"The laptop really does help, because we need to be portable, we need to be mobile. You know, our laptops are connected to the Internet, which is really helpful to have a SIM card built in" [Interview participant 11, Team leader/manager].
Quality of care–related drawbacks	
Impact rapport building due to having technology	"I take my laptop in and sometimes I feel that it hinders building that relationship with a client because I've got a screen in front of me. I work sometimes from both my laptop and my phone at the same time, so I feel like they are not giving me their full story because...I look like I'm playing on my phone or playing on my laptop rather than building that eye contact, building that relationship with the client" [Interview participant 14, Allied health professional].
Impact on client privacy with use of technology for safety	"I have seen a couple of client situations where family members have put cameras around that they can't be in the same house. They've put a lot of cameras with motion sensors and those sort of things for clients with dementia, and I find that quite intrusive as a way of caring for somebody. So that use of technology I find a bit jarring" [Interview participant 04, Home care package coordinator/care manager].
Reducing human interaction during care	"We can't have robots going in and seeing the clients for us, we still need...that might come eventually one day, but we still need the face-to-face human interaction...We need to keep that face-to-face. Definitely" [Interview participant 16, Domestic assistant].
Reducing time for care	"Sometimes it's the time spent on technology that might not be good...You lose time to spend on care" [Interview participant 17, Care aide/therapy assistant].
Technology issues may negatively impact quality of care	"[Electronic clinical records] can be very glitchy. It's already happened to me this morning where it had some sort of fatal error and booted me out. I had to log back in and start from scratch again and that all takes time. Uh, it can be very slow to download documents...and things like that. If you need to access documents like medication orders...that have been uploaded [but] haven't seen yet that are on [the electronic clinical records], that can be very slow to access" [Interview participant 12, Enrolled nurse/registered nurse].
Increasing clients' functional decline and care needs	"I think that implementing too much technology into home care. I think I would have to disagree with it...getting their...services like...social support...and replacing it with technology, I think that could be really a part of a client's decline, so they could start getting depressed or feel socially isolated" [Interview participant 08, Home care package coordinator/care manager].
Staff and organization–related drawbacks	
Increased physical burden due to technology use	"You don't [want to] be [staring] at a screen all day. Mainly so it does become quite hard" [Interview participant 08, Home care package coordinator/care manager].

Quality of Care–Related Benefits and Drawbacks

Several quality of care–related benefits were discussed by interview participants, including benefits for care delivery, as well as direct benefits for clients. Care delivery benefits included improvements to the availability of care, better access to data to inform care, and greater flexibility of scheduling of care (Table 3). In addition, participants suggested that technology use within in-home aged care can empower clients' independence and function, enabling them to be more involved in their care and allowing them to remain connected to the world.

> *I think that's a great idea because clients are always wanting to keep in the new age. They don't want to feel like they're going back to the old ways, as they say. So, helping them or supporting them to be able to access [technology] and be able to know what's happening in the world and then I think it's a really good benefit for them.* [Interview participant 10, Team leader/manager]

In contrast, participants reported that there were several drawbacks that could arise from technology use within in-home aged care, including the potential for technical issues with technology to negatively affect the quality of care:

> *What we do is often we'll obviously restart, try to re-establish a connection or...abandon the video call and then go through to the phone call. When we do abandon the video call, obviously we can't see the client, so we can't do necessary skin checks*

XSL•FO
RenderX

> *and...more visual type assessments that we can do via video call...That does impact our assessment.* [Interview participant 09, Home care package coordinator/care manager]

Many participants discussed concerns that technology use could result in less time to spend delivering in-person care, impacting on their ability to build a meaningful relationship with the client, and reducing interpersonal interaction during direct care delivery. This participant felt that many older people rely on aged care providers for regular human interaction:

> *The thing that you just miss is that real human connection of being in a room with somebody. And for a lot of our older adults, [the] only connection with [a] person on a regular basis is when someone comes in to help him with their shower or do their shopping or cook their meal. And I think if you switch now while you could never help someone with their shower virtually, there are lots of things you could cut out. And I just feel the person would lose that. That's where the quality goes. It's that human connection.* [Interview participant 11, Team leader/manager]

Staff and Organization–Related Benefits and Drawbacks

Similarly, participants suggested that technology could bring multiple benefits for staff and in-home aged care organizations (Table 3). Several participants proposed that technology use can improve the accessibility of up-to-date information and data, including access to care-relevant information, such as information on client preferences regarding their care.

> *If I was to go in...and do an assessment, and then I would put on [the electronic medical record] alerts, this is a danger I could put in a care plan. This is the way that we should probably [be] providing the care, so the next provider would have all of that information. I think that would enhance the care [provided to the client].* [Interview participant 14, Allied health professional]

> *There can also be alerts on what a client's preferences are, so you can already...be equipped with the tools to go in and provide the best care possible with...access to that information.* [Interview participant 14, Allied health professional]

Furthermore, many participants discussed the improvements to staff efficiency that could be provided through technology use, such as reduction of time spent on travel and administrative tasks.

> *I think it's very valuable for administrative purposes for doing...reviews to help cut down on time.* [Interview participant 11, Team leader/manager]

> *It's [going to] make us more efficient...Let's say the average day...we're expected to do X hours face to face [client] time...If some of that time [we] could still [be] in the office doing an assessment or a review via this kind of [technology] arrangement, then...we will be able to fit in more clients because we're not [wasting] time traveling.* [Interview participant 05, Allied health professional]

Conversely, interview participants proposed only 1 staff-related drawback of technology use within in-home aged care. Participants reported that technology use can result in increased physical burden related to sitting at a computer or using a device screen for extended periods:

> *It is unpleasant having to do so much with mobile. [My] neck gets very sore looking down all the [time].* [Survey participant 11]

Enablers and Barriers for Technology Use Within In-Home Aged Care

Enablers and barriers for technology use were explored in both the survey and the interviews. The quantitative findings from the survey are reported in Tables 4-6. Illustrative quotes from survey and interview participants regarding the enablers and barriers are reported in Table 7. Overall, the survey participants rated all the listed enablers (ranging between 3.12 and 3.81) and the majority of listed barriers (ranging between 2.92 and 3.74) as having an important impact on their use of technology within in-home aged care. In the interviews, participants also mostly agreed with enablers and barriers listed in the survey.

XSL•FO
RenderX

Table . Enablers of staff using technology (n=199, missing=27; ranked from highest to lowest importance)[a].

Enablers	Score, mean (SD)
Having someone to go to if or when the technology does not work	3.81 (0.41)
Having technology that is reliable	3.79 (0.43)
Having technology that is suitable and appropriate	3.69 (0.49)
Having good organizational and managerial support	3.62 (0.57)
Having staff involvement in how technology is implemented	3.57 (0.55)
Having enough time to get familiar with the technology	3.56 (0.58)
Having education and training on how to use technology for care tasks	3.51 (0.62)
Your confidence using technology for care tasks	3.34 (0.66)
Your previous experience in using technology for care tasks	3.12 (0.64)

[a]Don’t know = 0, not at all important = 1, low importance = 2, important = 3, and very important = 4.

Table . Barriers of staff using technology (staff-related factors) (n=199, missing = 27; ranked from highest to lowest importance)[a].

Barriers	Score, mean (SD)
Having unreliable technology	3.74 (0.59)
Having limited or no support when technology does not work	3.70 (0.60)
Having inappropriate or unsuitable technology	3.69 (0.64)
Having limited or no staff involvement in how technology is implemented	3.48 (0.69)
Having limited organizational and managerial support	3.46 (0.72)
Having limited time to get familiar with the technology	3.46 (0.75)
Limited education and training on how to use technology for care tasks	3.38 (0.76)
Cost to client	3.17 (0.90)
Concern that introducing technology will increase to workload	3.08 (0.79)
Your confidence using technology for care tasks	3.05 (0.90)
Client ability, familiarity, and confidence using technology	3.04 (0.83)
Client health status	2.96 (0.87)
Your previous experience in using technology for care tasks	2.92 (0.84)

[a]Don’t know = 0, not at all important = 1, low importance = 2, important = 3, and very important = 4.

XSL•FO
RenderX

Table . Barriers of staff using technology (client-related factors) (n=199, missing = 27; ranked from highest to lowest importance)[a].

Barriers	Score, mean (SD)
Cost to client	3.27 (0.88)
Client ability, familiarity, and confidence using technology	3.23 (0.88)
Client health status	3.13 (0.84)

[a]Don’t know = 0, not at all important = 1, low importance = 2, important = 3, and very important = 4.

XSL•FO
RenderX

Table . Enablers and barriers of staff using technology and representative quotes from semistructured interviews.

Enablers and barriers	Representative quotes
Having someone to go to if or when the technology does not work/Having limited or no support when technology does not work	"We have a really good IT department, so if it is...my computer is not turning on or my phones [is] not doing something, I will always send them to IT when it is IT specific" [Interview participant 11, Team leader/manager].
Having technology that is reliable/Having unreliable technology	"When you get to a client and it's not loaded in properly and you might have to restart your phone. Then you've just got to wait for it to have to do that, and sometimes it might cut into the client's time, or you're stood there waiting because you've [got to] wait for your phone to log in, so you can log in and start that client's visit" [Interview participant 10, Team leader/manager].
Having technology that is suitable and appropriate/Having inappropriate or unsuitable technology	"The program we have is pretty comprehensive. It's got all our schedule, and it's got all the maps and everything...I plug it into my car with Android Auto, but hit the map thing and it directs me to my next client, so it makes it easier to [find and] not having to go and hunt up addresses and do all that thing and type anything in to actually do that. That makes it really simple" [Interview participant 12, Enrolled nurse/registered nurse].
Having good organizational and managerial support/Having limited organizational and managerial support	"We do have quite good managerial support in terms of the technology" (Interview participant 02, Allied health professional)./"We did do a trial with the IT [department and they] were quite supportive" [Interview participant 11, Team leader/manager].
Having staff involvement in how technology is implemented/Having limited or no staff involvement in how technology is implemented	"The frontline staff are often the ones that are left out or it's only trialed with a small...area of the frontline staff. And I don't think that gives a broad understanding of the bigger picture" [Interview participant 13, Team leader/manager].
Having enough time to get familiar with the technology/Having limited time to get familiar with the technology	"When you're starting the job...getting used to the program is a big part. I've been training the staff member for a couple of weeks now and it even when I...started it probably took me a good couple of months to find all the little different bits and pieces and shortcuts and things like that" [Interview participant 12, Enrolled nurse/registered nurse].
Having education and training on how to use technology for care tasks/Limited education and training on how to use technology for care tasks	"Yeah, I think we definitely need to look at doing it better. Because I think with a lot of these programs, you know you spend so much time on it...building them and implementing them. And then it's only a short window of training to make sure everybody's on board and then off you go. You'll be right. And like I said, and then it's the staff on the ground that are left picking up the pieces like, yeah, I'll walk you through that. So that's probably a big gap is training on digital for sure" [Interview participant 07, Home care package coordinator/care manager].
Your confidence using technology for care tasks	"I think it's what we just spoke about like...making sure they've got competence and the domestic [worker] actually go out there and basically given the tools to start using that technology and not be scared to use [the] technology" [Interview participant 16, Domestic assistant].
Your previous experience in using technology for care tasks	"Some of them don't even have a smartphone that they use in their own life, so having to explain how to use a mobile phone is quite tricky as part of an onboarding process" [Interview participant 13, Team leader/manager].
Cost to client	"Cost to client...Yeah, there's not a lot. Well, I wouldn't say that there's a huge cost to client because you know not in our setting because they would just have a normal clinical assessment. They'd be billed for that. They wouldn't be [billed] for the technology aspect" [Interview participant 09, Home care package coordinator/care manager].
Concern that introducing technology will increase to workload	"We've attempted some remote clinical assessments via teams which are OK, but connection drops out sometimes. You know, if clients have dementia or they've got cognitive issues, they obviously can't control the technology themselves, so we have to rely on external or other people to assist us with that and that becomes an issue. That's sort of put...a stop to us doing it because it's taking more time for...staff to sit in the home with the clients while we do our work" [Interview participant 09, Home care package coordinator/care manager].
Client ability, familiarity, and confidence using technology	"Familiarity and confidence with using [technology]. That's a big one. I think people get scared of technology and especially the elderly" [Interview participant 09, Home care package coordinator/care manager].

XSL•FO
RenderX

Enablers and barriers	Representative quotes
Client health status	"If clients have dementia or they've got cognitive issues, they obviously can't control the technology themselves" [Interview participant 09, Home care package coordinator/care manager].
Additional enablers	
Client preference of technology	"People are asking for it...And especially our Parkinson's clients that are looking at, 'This is my future. I'm [going to] need a computer to be able to get my meals into and I don't know how to do that and to be able to communicate with my family and to be able to...Our speech [pathologist] at the moment, go out and do a little bit with clients'" [Interview participant 06, Home care package coordinator/care manager].
Paced introduction of technology into care	"I just think that it's definitely the way [of] the future and as long as it's introduced slowly into people's lives...I think as long as it's introduced slowly, people are much more open to it. And the more you kind of introduce, the more people will accept it" [Interview participant 14, Allied health professional].
Additional barriers	
Client preference of technology	"Some of them don't like it and send it back. It's not what they wanted" [Interview participant 01, Home care package coordinator/care manager].
Lack of technology infrastructure for clients	"They either don't have an Internet connection, they don't have smartphones, or they don't feel comfortable using it to do phone calls like that or telehealth. Most who do telehealth and even during COVID, they were supported by a family member...I don't [want to] say none, but the large majority just don't have the technology yet preexisting in their home" [Interview participant 11, Team leader/manager].
Lack of client trust of technology	"The most part of it is they don't trust technology, and that's what they say. Because with all the scams that goes around, they just don't have that belief that it's a safe way for communication" [Interview participant 10, Team leader/manager].
Digital literacy of staff	"I would say to understand the differences between different software and how to manage them, because it is quite common for staff members not to be very tech savvy with them" [Interview participant 10, Team leader/manager].
Staff preference regarding the technology use in care	"It should be something that's not mandatory...I believe that it should be something that individual clinicians should be able to do if they are confident to do so. If clinicians want to do that, but aren't confident, then there should be training offered because it's going to make us more efficient" [Interview participant 05, Allied health professional].
Complexities related to the aged care system	"The only issue I would see with that is when we get agency staffing or like agency nurses, we have a lot of brokerage staff. For example, if we don't have enough cleaners, we might go to another agency...so they don't wouldn't be privy to being able to see what's on [electronic care plans]" [Interview participant 06, Home care package coordinator/care manager].
Cost to the organization	"Cost is huge, and that's always going to be [a] huge cost to the organization and cost to client because you know it's always going to be the biggest barrier. I think everybody wants bang for their buck and it's not always [going to] be at the top of the list" [Interview participant 04, Home care package coordinator/care manager].
Having poor data connectivity	"This is...definitely a limiting thing in some areas, and downloading different stuff can take quite a long time or just doesn't work" [Interview participant 12, Enrolled nurse/registered nurse].

In the survey, the 3 highest ranked enablers for staff use of technology within in-home aged care were having someone to go to if or when the technology does not work (mean 3.81, SD 0.41), having technology that is reliable (mean 3.79, SD 0.43), and having technology that is suitable and appropriate (mean 3.69, SD 0.49) (Table 4). Similarly, albeit in a different order, the 3 highest ranked barriers for staff technology use were having unreliable technology (mean 3.74, SD 0.59), having limited or no support when technology does not work (mean 3.70, SD 0.60), and having inappropriate or unsuitable technology (mean 3.69, SD 0.64) (Table 5). Other factors, such as organizational support, staff involvement in the implementation of the technology, and opportunities for training and familiarization with technology, were also strongly endorsed as enablers (or barriers when not present).

XSL•FO
RenderX

Participants highlighted the importance of having support in how to use technology. Some participants also suggested that unofficial peer support from colleagues facilitated technology use, as well as more formal organizational support, for example, from their manager or IT department.

> *Having someone to go to when technology doesn't work is vitally [important]. I ring IT on a very regular basis.* [Interview participant 12, Enrolled nurse/registered nurse]

> *And I was constantly calling other team members...Just "How do you do this?," "How do you do that?" and having within my own team someone that I could call to get that advice was very helpful.* [Interview participant 12, Enrolled nurse/registered nurse]

The reliability of technology was highlighted by participants as important, with participants reporting additional time and work are required when the technology is unreliable.

> *A software, say like one of our scheduling software. Sometimes that does [crash] on you, and sometimes you have to load out of it and load back in, so it would be nice to have something that runs very smoothly and is reliable.* [Interview participant 10, Team leader/manager]

Different aspects of the suitability and appropriateness of technology were discussed by participants, such as the design of technology and the intersection of technology with the features of in-home aged care. The participants highlighted the importance of the suitability of technology design for the on-the-road nature of in-home aged care and the direct care tasks delivered to aged care clients. For example, one participant talked about the limitations of a smartphone for some tasks, whereas another participant discussed the suitability of a software for in-home aged care:

> *The phones are ridiculously hard to use if we're trying to seek information...because it's so little and small to try and either type on if you're out in the field or try to navigate through the system to try and get the information that you need. That's been a hindrance.* [Interview participant 06, Home care package coordinator/care manager]

> *I think it's good for being on the road because it gives you all the information that you need.* [Interview participant 10, Team leader/manager]

Participants also raised additional enablers and barriers to technology use by staff within in-home aged care. Client preferences regarding the use of technology emerged as both an enabler and a barrier. Participants reported that aged care clients have differing preferences for technology, with some clients requesting more technology use and others requesting technology be removed from their home. Other barriers to technology use discussed included lack of technology infrastructure for clients, lack of client trust, and staff digital literacy and preferences. Poor data connectivity was highlighted as a barrier by multiple participants:

> *Depending on where you are in Perth, unfortunately there are some black spots around and we have no coverage...This is...definitely a limiting thing in some areas, and downloading different stuff can take quite a long time or just doesn't work...There are some spots where we just have no coverage and there's nothing you can do in that situation.* [Interview participant 12, Enrolled nurse/registered nurse]

> *We have a lot of issues with coverage in certain areas.* [Interview participant 13, Team leader/manager]

Discussion

Principal Results

In this study, we used a mixed methods approach to investigate Australian in-home aged care workers' views on the digital enablement potential of direct care tasks, benefits and drawbacks of technology use, and the enablers and barriers to technology use within in-home aged care. Overall, participants felt that many direct care tasks within in-home aged care could be digitally enabled, with more than half of the common direct care tasks completed by staff identified as being likely to be digitally enabled. This is in line with research findings from Ireland and Spain, which show that health and social care professionals appear positive and hopeful regarding the role and use of technology in care delivery within health and social care [26].

However, our study found that staff members' views on the digital enablement potential of their tasks differed by professions, with care aides, therapy assistants, and domestic assistants rating their tasks less likely to be digitally enabled. This is similar to research conducted in the Finnish setting, which suggested that care workers' views on the digital enablement potential may be partially due to staff in these roles viewing their work to be more physical and hands-on in nature [10]. This is an interesting finding as new technology, such as robot vacuums (which are available for Australian in-home aged care clients to access as part of their services) [27], is becoming increasingly common and could complement direct care tasks, such as general house cleaning. Previous research has found that Australian in-home aged care workers' role was associated with their digital readiness, with staff with less clinical responsibilities having higher proportions of individuals with lower digital readiness [15]. This may have negatively influenced participants' views on the digital enablement potential of their direct care tasks. Therefore, some in-home aged care workers may require additional support to improve their digital readiness and, in turn, change their views on digitally enabling direct care tasks. This study showed that staff in clinical or care coordination roles felt that communication tasks with colleagues or external health providers had a high digital enablement potential. This concurs with the existing literature that technology in home care services could enhance communication and coordination and free up time for more care provision [10,26].

Despite the overall positive views from staff regarding the digital enablement potential of direct care tasks, concerns were raised regarding the potential impact of digital enablement on the

XSL•FO
RenderX

quality of care. In the interviews, staff discussed multiple quality of care-related benefits from the use of technology within in-home aged care, such as improved care availability, increased access to care information, client enablement, and improved client safety and security. On the other hand, staff were concerned that digital enablement could have drawbacks, including reduced human interactions during care, impact on client privacy, and the negative impact of technology issues on care quality. Similar to our results, research from Finland and Ireland showed that there were concerns from health and social care providers that the increasing use of technology would lead to the loss of the "human touch" and reduce human interactions during care [10].

Interestingly, government strategies supporting technology in aged care have suggested that technology would be able to reduce the time spent on noncare activities such as administration to free up time for direct care [28]. However, the mismatch between the government and the workers' views may stem from the wide range of technology that could be implemented within in-home aged care, which may have different impacts on human interactions during care. The potential loss of "human touch" that might occur with technology would need to be carefully considered to avoid negative impacts on the quality of in-home aged care. Participants also discussed benefits and drawbacks of digital enablement related to staff and the organization. One such benefit of technology is the improved accessibility of up-to-date information. This is consistent with previous research showing that digitalization in health and social care services could improve the care-related information availability for staff regarding care recipients [26].

In order to realize the potential benefits of technology, factors related to technology use need to be considered during the implementation of technology into in-home aged care. The majority of the study cohort agreed with the importance of most of the enablers and barriers generated by the research team and suggested additional enablers and barriers from their experience of working in Australian in-home aged care. The most important factor for staff technology use was the availability of support when technical issues occur. The different types of support that were important for staff use of technology were "official" support provided by the internal information technology department and "unofficial" support provided by their colleagues. As shown in an Australian telehealth trial, clinicians require support for a wide range of technical tasks, such as software and hardware issues, assistance with technology setup, and device management [29]. It is critical for technical support to be readily available for staff as and when required. Another barrier raised by multiple participants was poor data connectivity and its negative effect on staff technology use. Connectivity issues have been raised as a key barrier in previous research in rural Canadian home care services [30]. These study findings extend this concern about data connectivity issues to the Australian context—in both metropolitan and regional areas.

Furthermore, participants raised several additional important ethical issues related to technology use within in-home aged care, such as client technology preferences and trust of technology. While additional research is required on staff perspectives of ethical issues, the findings align with those from a scoping review of ethical issues related to technology use within aged care, which suggests that ethical concerns, such as lack of regulatory oversight, client autonomy, and nonmaleficence, are important topics for older Australians and their carers [31]. Research also highlights that ethical issues are particularly complex when the aged care recipient lacks the capacity to make an informed decision [9]. Although technology use can offer multiple potential benefits as identified by the study findings, it is paramount for in-home aged care service providers to thoroughly consider the ethical concerns of using technology with care recipients, and that staff are well versed in how to navigate these concerns with care recipients.

Limitations and Strengths

There were a few study limitations that should be considered when interpreting the findings. First, both the survey and the interviews were conducted using digital platforms. This may have potentially excluded in-home aged care staff members who may feel less comfortable using technology and thus affecting the generalizability and transferability of the findings. However, the digital platforms used for this study are commonly used for a range of operational and administrative purposes within the organization and most staff should be familiar with them. Second, while the demographic characteristics of the staff survey participants were similar to the overall Australian in-home aged care workforce [32], nursing staff were underrepresented. In addition, there was a 12% missing response rate in questions related to enablers and barriers for technology use. This may have limited the generalizability of the quantitative results. Furthermore, interview participants were recruited in 2 ways, with interested staff members self-nominating for the interviews or being nominated by their managers. This may have biased the results of the qualitative interviews. Moreover, the terminology used for direct care tasks may have been interpreted differently by participants. However, the list of direct care tasks was generated through the mapping of organizational service delivery documents and should be relatively familiar and easy to understand for participants. Finally, there were technical issues with the audio recordings of the staff interviews. Due to a storage error, the audio recordings were deleted prior to the completion of data analysis, meaning that the research team was required to rely on the initial transcripts of the interviews. Nonetheless, a strength of the study was the use of a sequential explanatory mixed methods approach, which allowed for an in-depth exploration of the quantitative results [20].

This is the first study specifically investigating Australian in-home aged care workers' perspectives of the digital enablement potential of direct care tasks, benefits and drawbacks of technology use, and enablers and barriers of technology use. The results of the study would be beneficial for both technology developers and in-home aged care providers, providing key information to guide the implementation of technology into the care provided to older people. As previous literature has pointed out, the importance of understanding end users' needs and requirements must be prioritized when developing technology for aged care [33]. Our findings will inform technology developers of the relevant tasks and benefits for novel

XSL•FO
RenderX

technology and barriers that should be overcome when designing technology for in-home aged care. Previous research has identified that technology implementation is complex and multifaceted in residential aged care, with staff members being a crucial part of the implementation process [34]. This study will allow in-home aged care providers to gain a better understanding of workers' views on technology use and inform their future technology implementation efforts. Larger-scale studies of multiple aged care providers are required to validate the study findings. Future research should focus on ascertaining evidence-based strategies to facilitate the enablers and overcome the barriers of technology use within Australian in-home aged care.

Conclusions

This study provides insight into staff members' views of the digital enablement potential of direct care tasks, benefits and drawbacks of technology use, and the enablers and barriers to technology use within Australian in-home aged care. The study findings would be beneficial for both technology developers and in-home aged care providers, providing key information to guide the implementation of technology into the care provided to older people. Further research is required to ensure that appropriate strategies are available to facilitate the enablers and overcome the barriers to ensure successful implementation of technology into in-home aged care.

Acknowledgments

The authors would like to acknowledge the staff members who took time to take part in the survey and interviews. The authors disclosed receipt of the following financial support for the research, authorship, and publication of this paper: this work was supported by an Aged Care Research and Industry Innovation Australia (ARIIA) grant. The funders did not take part in planning, conducting, or reporting of this research. Silverchain took part in the review and approval of the manuscript for publication.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to ethical and privacy restrictions.

Conflicts of Interest

BPMK, MD, KF, and TED were employees of Silverchain, a leading national in-home aged care provider, when this research was conducted. The remaining authors declare no conflicts of interest.

Multimedia Appendix 1
Interview guide and materials.
[DOCX File, 34 KB - aging_v8i1e76141_app1.docx]

References

1. United Nations Department of Economic and Social Affairs, Population Division. World Population Prospects 2022: Summary of Results: United Nations; 2022. URL: https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/wpp2022_summary_of_results.pdf [accessed 2025-09-09]
2. Global Strategy and Action Plan on Ageing and Health: World Health Organization; 2017. URL: https://iris.who.int/bitstream/handle/10665/329960/9789241513500-eng.pdf?sequence=1 [accessed 2025-09-09]
3. Duty of care: meeting the aged care workforce challenge. : Committee for Economic Development of Australia; 2021.
4. Older Australians. Australian Institute of Health and Welfare. 2024. URL: https://www.aihw.gov.au/reports/older-people/older-australians/contents/demographic-profile#Australia%E2%80%99s%20older%20population [accessed 2025-02-28]
5. Royal Commission Into Aged Care Quality and Safety. Final report: care, dignity and respect (summary and recommendations). : Royal Commissions; 2021 URL: https://www.royalcommission.gov.au/aged-care/final-report [accessed 2025-09-09]
6. People using aged care. Australian Institute of Health and Welfare. 2022. URL: https://www.gen-agedcaredata.gov.au/topics/people-using-aged-care#Aged%20care%20use%20in%20Australia [accessed 2023-05-10]
7. State of the sector: aged care 2024. : Aged & Community Care Providers Association; 2024.
8. RMIT-Cisco Health Transformation Lab. Transforming aged care: towards a future in which digitisation clasps hands with respect, and connection drives improvement. : Health Transformation Lab; 2022.
9. Barnett K, Livingstone A, Margelis G, Tomlinson G, Young R. Technology ageing and aged care: literature review. : Aged Care Industry IT Council; 2019 URL: https://www.researchgate.net/profile/Kate-Barnett-4/publication/338763552_TECHNOLOGY_AGEING_AND_AGED_CARE_LITERATURE_REVIEW/links/5e292b4ea6fdcc70a1436ea2/TECHNOLOGY-AGEING-AND-AGED-CARE-LITERATURE-REVIEW.pdf [accessed 2025-09-09]
10. Lolich L, Pirhonen J, Turja T, Timonen V. Technology in the home care of older people: views from Finland and Ireland. J Cross Cult Gerontol 2022 Jun;37(2):181-200. [doi: 10.1007/s10823-022-09449-z] [Medline: 35435600]

XSL•FO
RenderX

11. Lu W, Silvera-Tawil D, Yoon HJ, et al. Impact of the smarter safer homes solution on quality of life and health outcomes in older people living in their own homes: randomized controlled trial. J Med Internet Res 2025 Jan 22;27:e59921. [doi: 10.2196/59921] [Medline: 39608020]
12. McKittrick R, Orellana L, Manias E, Hensher M, Hutchinson AM. Characteristics, care and support needs of older Victorians requiring a government-funded Home Care Package: an observational study. Australas J Ageing 2025 Mar;44(1):e13400. [doi: 10.1111/ajag.13400] [Medline: 39799959]
13. Inacio MC, Lang C, Bray SCE, et al. Health status and healthcare trends of individuals accessing Australian aged care programmes over a decade: the Registry of Senior Australians historical cohort. Intern Med J 2021 May;51(5):712-724. [doi: 10.1111/imj.14871] [Medline: 32359019]
14. Dickins M, Kelly J, Paxton S, et al. The prevalence and correlates of depression and anxiety symptoms in older adults receiving in-home aged care: a cross-sectional survey. Australas J Ageing 2025 Mar;44(1):e13396. [doi: 10.1111/ajag.13396] [Medline: 39940120]
15. Kwan BPM, Dickins M, Batchelor F, et al. Digital readiness of the in-home aged care workforce: a latent profile analysis. Australas J Ageing 2025 Mar;44(1):e70010. [doi: 10.1111/ajag.70010] [Medline: 40044589]
16. Douglas HE, Georgiou A, Tariq A, et al. Implementing information and communication technology to support community aged care service integration: lessons from an Australian aged care provider. Int J Integr Care 2017 Apr 10;17(1):9. [doi: 10.5334/ijic.2437] [Medline: 29042851]
17. Papadopoulos I, Koulouglioti C, Lazzarino R, Ali S. Enablers and barriers to the implementation of socially assistive humanoid robots in health and social care: a systematic review. BMJ Open 2020 Jan 9;10(1):e033096. [doi: 10.1136/bmjopen-2019-033096] [Medline: 31924639]
18. Basile I, Consolo L, Colombo S, Rusconi D, Rampichini F, Lusignani M. Technology to support older adults in home palliative care: a scoping review. Am J Hosp Palliat Care 2024 Jun;41(6):673-690. [doi: 10.1177/10499091231189502] [Medline: 37473720]
19. Galavi Z, Montazeri M, Ahmadian L. Barriers and challenges of using health information technology in home care: a systematic review. Int J Health Plann Manage 2022 Sep;37(5):2542-2568. [doi: 10.1002/hpm.3492] [Medline: 35643906]
20. Ivankova NV, Creswell JW, Stick SL. Using mixed-methods sequential explanatory design: from theory to practice. Field methods 2006 Feb;18(1):3-20. [doi: 10.1177/1525822X05282260]
21. Commonwealth home support programme: program manual 2022-2023. : Australian Government Department of Health and Aged Care; 2022.
22. Transition care programme guidelines updated December 2022. : Australian Government Department of Health and Aged Care; 2022.
23. Short-term restorative care programme manual. Australian Government Department of Health. 2021. URL: https://www.health.gov.au/sites/default/files/documents/2021/04/short-term-restorative-care-programme-manual.pdf [accessed 2023-03-28]
24. Home care packages program manual for care recipients–version 14. Australian Government Department of Health and Aged Care. 2023. URL: https://www.myagedcare.gov.au/sites/default/files/2023-01/operational-manual-for-home-care-package-consumers.pdf [accessed 2023-03-03]
25. Guest G, MacQueen K, Namey E. Applied Thematic Analysis: SAGE Publications, Inc; 2012. [doi: 10.4135/9781483384436]
26. Lolich L, Riccò I, Deusdad B, Timonen V. Embracing technology? Health and social care professionals' attitudes to the deployment of e-Health initiatives in elder care services in Catalonia and Ireland. Technol Forecast Soc Change 2019 Oct;147:63-71. [doi: 10.1016/j.techfore.2019.06.012]
27. Home care packages (HCP) program inclusions and exclusions faqs for providers–version 2. Australian Government Department of Health and Aged Care. 2024. URL: https://www.health.gov.au/sites/default/files/2024-11/home-care-packages-program-inclusions-and-exclusions-faqs-for-providers-version-1.pdf [accessed 2024-08-05]
28. Aged care data and digital strategy 2024-2029. Australian Government Department of Health and Aged Care. 2024. URL: https://www.health.gov.au/sites/default/files/2024-07/aged-care-data-and-digital-strategy-2024-2029_0_0.pdf [accessed 2024-08-05]
29. Taylor A, Wade V, Morris G, et al. Technology support to a telehealth in the home service: qualitative observations. J Telemed Telecare 2016 Jul;22(5):296-303. [doi: 10.1177/1357633X15601523] [Medline: 26362562]
30. Elliott J, Gordon A, Tong CE, Stolee P. "We've got the home care data, what do we do with it?": understanding data use in decision making and quality improvement. BMC Health Serv Res 2020 Mar 25;20(1):251. [doi: 10.1186/s12913-020-5018-9] [Medline: 32213180]
31. Ethics and technology: technology in aged care evidence theme. ARIIA Knowledge & Implementation Hub. 2023. URL: https://www.ariia.org.au/sites/default/files/2023-02/ARIIA-technology-ethics-and-technology.pdf [accessed 2025-07-25]
32. 2020 aged care workforce census report. Australian Government Department of Health and Aged Care. 2021. URL: https://www.health.gov.au/resources/publications/2020-aged-care-workforce-census?language=en [accessed 2025-04-04]
33. Lion KM, Murfield J, Sriram D, et al. Technology in aged care: a qualitative survey of academic, research, and technology industry professionals. Contemp Nurse 2023;59(4-5):311-322. [doi: 10.1080/10376178.2023.2242978] [Medline: 37537756]

XSL•FO
RenderX

34. Cavenett W, Baker S, Waycott J, et al. Deploying new technology in residential aged care: staff members' perspectives. Presented at: Proceedings of the 30th Australian Conference on Computer-Human Interaction; Dec 4-7, 2018; Melbourne, Australia p. 200-204. [doi: 10.1145/3292147.3292214]

Edited by M Antoniou; submitted 24.04.25; peer-reviewed by A Hall, J Waycott; revised version received 07.08.25; accepted 24.08.25; published 14.10.25.

Please cite as:
Kwan BPM, Dickins M, Williams S, Batchelor F, Hwang K, Fulford K, Walsh T, Jakupovic H, Davison TE
The Use of Technology to Deliver In-Home Aged Care Services: Mixed Methods Study of Australian Staff Perspectives
JMIR Aging 2025;8:e76141
URL: https://aging.jmir.org/2025/1/e76141
doi:10.2196/76141

© Baldwin Pok Man Kwan, Marissa Dickins, Sue Williams, Frances Batchelor, Kerry Hwang, Kate Fulford, Tony Walsh, Helena Jakupovic, Tanya E Davison. Originally published in JMIR Aging (https://aging.jmir.org), 14.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Identifying Food Preferences and Malnutrition in Older Adults in Care Homes: Co-Design Study of a Digital Nutrition Assessment Tool

Jenni Connelly[1*], BA, MSc, PhD; Kevin Swingler[2], BSc, MSc, PhD; Nidia Rodriguez-Sanchez[1], BSc, PhD; Anna C Whittaker[1*], BSc, MSc, PhD

[1]Faculty of Health Sciences and Sport, University of Stirling, Stirling, United Kingdom

[2]Faculty of Natural Sciences, University of Stirling, Stirling, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Jenni Connelly, BA, MSc, PhD
Faculty of Health Sciences and Sport
University of Stirling
3a74a Cottrell Building
Stirling, FK9 4LA
United Kingdom
Phone: 44 1786 466399
Email: jenni.connelly1@stir.ac.uk

Abstract

Background: Malnutrition is a challenge among older adults and can result in serious health consequences. However, the dietary intake monitoring needed to identify malnutrition for early intervention is affected by issues such as difficulty remembering or needing a dietitian to interpret the results.

Objective: This study aims to co-design a tool using automated food classification to monitor dietary intake and food preferences, as well as food-related symptoms and mood and hunger ratings, for use in care homes.

Methods: Participants were 2 separate advisory groups and 2 separate sets of prototype testers. The testers for the first prototype were 10 community-dwelling older adults based in the Stirlingshire area in Scotland who noted their feedback on the tool over 2 weeks in a food diary. The second set of testers consisted of 14 individuals (staff: n=8, 57%; and residents: n=6, 43%) based in 4 care homes in Scotland who provided feedback via interview after testing the tool for a minimum of 3 days. In addition, 130 care home staff across the United Kingdom completed the web-based survey on the tool's needs and potential routes to pay for it; 2 care home managers took part in follow-up interviews. Data were collected through food diaries, a web-based survey, audio recordings and transcriptions of focus groups and interviews, and research notes. Systematic text condensation was used to describe themes across the different types of data.

Results: Key features identified included ratings of hunger, mood, and gastrointestinal symptoms that could be associated with eating specific foods, as well as a traffic light system to indicate risk. Issues included staff time, Wi-Fi connectivity, and the accurate recognition of pureed food and fortified meals. Different models for potential use and commercialization were identified, including peer support among residents to assist those considered less able, staff-only use of the tool, care home–personalized database menus for easy meal photo selection, and targeted monitoring of residents considered to be at the highest risk using the traffic light system.

Conclusions: The tool was deemed useful for monitoring dietary habits and associated symptoms, but necessary design improvements were identified. These should be incorporated before formal evaluation of the tool as an intervention in this setting. Co-design was vital to help make the tool fit for the intended setting and users.

(JMIR Aging 2025;8:e64661) doi:10.2196/64661

KEYWORDS

ageing; digital technology; dietary measurement; care homes; co-design; dietary intake; food diary

XSL•FO
RenderX

Introduction

Background

Diet is an important component of a healthy lifestyle, and it remains one of the most common challenges among older adults in the United Kingdom [1]. Malnutrition is also a prevalent issue in this population. The World Health Organization defines malnutrition as "deficiencies or excesses in nutrient intake, imbalance of essential nutrients or impaired nutrient utilization" [2]. Approximately 10% of those aged 65 years are malnourished or are at risk of malnutrition in the United Kingdom [3], a figure that rose to 25% of those aged 60 years during the COVID-19 pandemic [4-6]. The prevalence of severe malnutrition is reported to be as high as 45% in those living in residential care homes [7] but may be worse because approximately 70% of malnutrition in the United Kingdom is unrecognized [8,9]. Malnutrition is strongly associated with frailty, as well as functional and health decline [10-12]; thus, it is important to monitor dietary intake and identify malnutrition risk before symptoms are already evident. However, tracking and understanding the impact of dietary intake is not easy. It is possible for people to become malnourished if they do not eat enough nutritionally rich foods for 2 to 5 days [6,13].

Malnutrition in older adults is commonly identified via the Malnutrition Universal Screening Tool (MUST) [14], and when individuals are classed as medium to high risk, it is recommended to document food and drink intake for 3 days. However, there are difficulties with self-reported food diaries, which can be inaccurate due to recording error, misremembering, and socially desirable responses. Additional barriers include limited staff time in care homes and the need for analysis and interpretation by a dietitian [15].

Computer-based tailored interventions may offer a solution because there is evidence that they have positively impacted dietary monitoring in the general population [16]. Automated food classification is a rapidly advancing field of artificial intelligence research that uses a computer source to identify different foods from an image [17,18]. Creating new nutrition databases or using or adapting existing databases means that it is possible to infer detailed nutritional data from photos of food [19]. Convolutional neural network algorithms are used to identify food automatically from photos, and these are linked to a nutritional database to calculate the calories in the identified foods. Machine learning techniques can then be applied to correlate dietary intake with important individual differences such as food preferences, physical symptoms, and malnutrition risk [20]. Automated food classification can consider the portion size and regularity of food intake, which is crucial to monitor in older people because this can significantly influence malnutrition risk. Linking food intake to preferences and symptoms through technology and artificial intelligence means this information could be used to identify potential risks quickly and accurately. This would provide care staff with information to help monitor older people's diet and decrease malnutrition risk. Studies that have evaluated photo-based dietary assessment tools have found measurement errors comparable to traditional methods [21], and those that have found it a valid method have focused on a younger population, reporting positive feedback [22,23].

Objectives

Consequently, following the 6 steps for quality intervention development (6SQuID) [24], we sought to co-design a prototype tool with key stakeholders, which we refer to as a "digital nutritional assessment tool." We aimed to work with advisory groups (AGs) and participants to develop the key features of a prototype tool for monitoring dietary intake, map this onto food preferences and symptoms, and test the prototype with stakeholders. It is important to note that this study reports on developmental research to identify the tool's suitability, features needed, and usability, rather than its efficacy at capturing and estimating nutrient intake, which is part of the next stage of development and testing. Its key objectives were to complete the first two 6SQuID steps: (1) defining and understanding the problem and its causes and (2) identifying which causal or contextual factors are modifiable—which have the greatest scope for change and who would benefit the most—using a co-design and coproduction approach. This inclusive approach was taken due to the wish for the resulting tool to have real-life application and translation outside of the academic setting while acknowledging the importance of specifying up front the motivations of coproduction, what outcomes are required for whom, and how these might be achieved [25]. This qualitative approach to data collection is necessarily reflective and conscious and has attempted to follow recently produced resources as a guide to quality coproduction methods in health research [26,27]. In this way, it positions end users as essential to the research process; thus, it is collaborative and equitable [27]. In this study, stakeholders were identified as older adults, the tool software designers, dietitians, and ultimately care home staff and residents themselves.

Methods

Recruitment

Participants consisted of 2 AGs and 2 separate sets of prototype testers. Both AGs comprised different older adults and dietitians. Care home staff as well as app and tool designers were consistent across the 2 groups. The testers for the first prototype were community-dwelling older adults and care home employees recruited across Stirlingshire in Scotland. For care home testing, the aim was to recruit 3 care homes, 2 staff members, and 4 to 5 residents in each care home. Care home staff and managers were recruited via social media and Scottish Care newsletters to complete a survey and semistructured interviews, respectively.

Study Design

The study used a co-design and coproduction approach [27] to determine key components of the tool to be developed and used through multiple iterations of usability feedback from older adults, care staff, and dietitians. The tool was developed using the information provided by the AGs in a collaboration between the research team and the app developers, Game Doctor. The development process involved AGs with key stakeholders in the design process (phase 1); community testing and feedback (phase 2); and care setting feasibility and usability testing, as

XSL•FO
RenderX

well as consultation and data gathering with a wider group of care home staff and managers (phase 3). This was a continual process, with feedback resulting in a new build of the tool prototype, to 3 builds. A summary of the timeline for the study design and the different groups involved is shown below in Table 1.

Table 1. Timeline and summary of the study design.

	Phase 1			Phase 2							Phase 3				
	Month 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
AG[a] 1		✓		✓											
Testing				✓				✓				✓			
Survey									✓	✓				✓	
Interviews										✓			✓		
AG 2											✓				✓

[a]AG: advisory group.

Procedure

Phase 1: AG Recruitment and Co-Design of Tool

According to the 6SQuID framework [24], the first step of intervention development is "defining and understanding the problem." This was achieved through the creation of the first AG (AG 1) with key stakeholders to develop ideas and identify features that would be vital for a food recognition tool in the care home setting. The study and AG role were advertised through email to the Stirling 1000 Elders email list used by ACW to recruit 2 older adults. A snowball technique was then used to recruit key stakeholders working within nutrition and care settings. Through this process, AG 1 with 6 key stakeholders (older adults: n=2, 33%; care home staff: n=1, 17%; app developers: n=2, 33%; and dietitian: n=1, 17%) was formed. The first session was run face-to-face and was used to define and understand the problem (ie, the need for a food recognition model for dietary monitoring and how artificial intelligence could be used as a potential solution). In addition, AG 1 members were asked to brainstorm key features that the tool should include, such as the identification of symptoms related to eating a particular food and potential rating scales to assess mental and physical well-being after eating. AG 1 also identified potential issues and challenges that would be faced in using the tool with older adults in a health care setting generally.

Feedback from this first AG 1 meeting provided the app developers, Game Doctor, with information on which to base the initial build of the tool and the 2 subsequent builds based on later AG meetings. The LogMeal (AIGecko Technologies SL, Artificial Intelligence and Deep Learning Food Division) photo database used by the tool to identify the foods photographed was selected in collaboration between the research team and app developers based on a review of potential existing services, considering the range of foods recognized and financial costs. The first prototype was then developed by Game Doctor, linked to the photo database, and demonstrated in a second meeting of AG 1 for prototype testing and feedback before bug fixing and independent feasibility testing.

Phase 2: Feasibility Testing of the Tool and Food Diaries

This phase was based on 6SQuID step 2: "identify which causal or contextual factors are modifiable—which have the greatest scope for change and who would benefit the most." We recruited 10 community-dwelling older adults via email to Stirling 1000 Elders. Eligibility included owning an Android mobile phone or tablet capable of downloading the tool. The older adults were provided with a pseudonymized log-in ID for the tool and asked to use it to record each meal over 2 weeks by photographing their plate before and after eating. They were also asked to select the foods present that were identified by the tool or to correct any misidentified foods by entering what the foods circled in the tool were. They could then also rate on a 5-point emoji-based Likert scale their mood, level of hunger after eating, and the presence of up to 7 further symptoms. The symptoms included in the tool were based on feedback from AG 1 and included tiredness, diarrhea, cramping, swallowing difficulties, constipation or discomfort, and bloating. A written diary to complete alongside using the tool was provided to participants to indicate the contents of each meal and note down any problems with using the tool and provide general feedback about it. A third and final meeting of AG 1 was planned to discuss the findings from prototype testing and give the members the opportunity for further testing to provide feedback to the app developer so that a third version of the prototype could be developed. Due to difficulties in arranging a final meeting, AG 1 members were asked to use the tool for a week and then complete a short web-based survey. The survey asked about missing tool features and symptoms, strategies to support the tool's integration in care settings, and other populations that would benefit from using the tool. This was then used as the basis for discussions among members of the second AG (AG 2).

Phase 3

Overview

As the tool was intended for eventual use in care homes, phases 1 and 2 were repeated with AG 2, a newly recruited group of 4 key stakeholders (older adults: n=2, 50%; care home staff: n=1, 25%; and dietitian [also an older adult]: n=1, 25%). In addition,

XSL•FO
RenderX

a new cohort of tester participants, including staff and residents, was recruited through 4 care homes.

At the first AG 2 meeting, the members discussed version 3 of the prototype and the challenges of its application in care homes.

Care Home Testing and Focus Groups

For testing, 8 care homes were contacted through existing networks with information on the study and what was expected of staff and residents. The lead researcher (JC) prepared an introduction to the tool and led a separate discussion at each care home on key features that care home staff and residents wanted to see in this type of tool. Each care home then had access to version 3 of the prototype for 2 weeks, and staff and residents were asked to record food intake and food-related symptoms. The care homes were given an Android tablet preloaded with the tool that could be used to take photos of the food. During this period, residents (n=6) were asked to take photos of all foods they consumed, while staff members (n=8) recorded food intake for other residents, across a minimum of 3 days, ensuring that each type of meal (breakfast, lunch, dinner, and snacks) was recorded at least once. The photos were then added to a timeline, and participants had the option to indicate how hungry they felt after the meal using an emoji-based Likert scale and click on any symptoms they were experiencing from a list of common symptoms, which were then added to the timeline. Participants also entered any verbal preferences and indicated whether they needed support with eating.

After prototype testing, JC conducted semistructured interviews with 5 participants (care home staff: n=4, 80%; and resident: n=1, 20%) to obtain their feedback on the usability and acceptability of the tool. Questions covered topics such as what they liked about the tool, difficulties encountered with its use, missing features (eg, physical or mental symptoms that should be included), additional features that should be included, and advice on the usability of the tool in practice (eg, who would take the photos and whether there was enough time to use the tool during a typical care home day). A second and final AG 2 meeting was convened to talk through the feedback from care homes and suggest future directions.

Care Home Staff Perspectives Survey and Manager Interviews

Concurrently with the care home testing, a web-based survey was devised for care home staff that included 16 questions under three headings: (1) “About you” (including job title, organization, and time spent in the care sector), (2) “Nutrition” (including questions about monitoring diet and dietary outcomes in their care home organization), and (3) “Routes to integrate tool.” Of the 16 questions, 6 (38%) were open ended. These were included (under the appropriate headings listed previously) to explore any care home–specific strategies or features that would improve the integration of the tool. The survey was designed to take no longer than 20 minutes to complete. This included a 7-minute video on the web showcasing the tool and its current features, which was required to be watched before completing the survey.

A total of 130 participants (care home staff) were recruited across Scotland via social media (ie, Twitter [subsequently rebranded as X] and LinkedIn), emails to existing partner care homes, and through Scottish Care. The inclusion criteria included any worker from within a UK care home. The survey was available through the advertising link, and, when clicked, the individual was taken to the participant information sheet.

After the survey, semistructured interviews were conducted with care home managers to explore the survey feedback and assess how the tool would fit within their care homes. Care home managers could sign up for these interviews via the last question in the survey.

Data Analysis

This project used systematic text condensation [28] across the different types of qualitative data collected. This is a descriptive and explorative method for gathering themes across different types of qualitative data, including interviews, observations, and written text analysis. As such, it represents a pragmatic approach that can turn total impressions of the overall chaos of the data into themes with units of meaning that can then be condensed and synthesized into descriptions and concepts.

Ethics Approval

This study was conducted in accordance with the Declaration of Helsinki and approved by the General University Ethics Committee of the University of Stirling (8857; September 7, 2023) for studies involving human participants. As a token of thanks, the participants testing the app received £50 (US $63.32) at the end, once the device had been returned. The feedback provided on the tool was anonymized in terms of users but not per care home. This was carried out in this manner to ensure feedback could be related to the size and location of the care home. Those who took part in the advisory group received £50 (US $63.32) per meeting they attended. For both the survey and the follow-up interviews, participants implied consent by clicking the provided link to take part in the survey.

Results

Participants

Participants consisted of 2 separate AGs (AG 1 and AG 2) and 2 separate sets of prototype testers. The testers for the first prototype were 10 community-dwelling older adults and care home employees based in the Stirlingshire area. Four care homes responded about participating in the study. The care home testers consisted of 14 individuals (staff: n=8, 57%; and residents: n=6, 43%) in 4 care homes across Crieff, Stirling, and Fife, differing in size and type of organization (nursing care, older adults only, mixed needs, and dementia specialty or general care). Altogether, 130 individuals completed the web-based survey, with 2 managers volunteering to take part in follow-up interviews.

Phase 1: AG Recruitment and Co-Design of Tool

AG 1 members met with the research team for 2 hours on 2 occasions over the course of 1 year to design, test, and improve the first protype. They also tested the final prototype and gave feedback before testing in community-dwelling older adults.

XSL•FO
RenderX

Game Doctor staff took part in all AGs and produced an initial prototype in March 2022. Game Doctor. They suggested that the 3-sprint model would work well for the development of up to 3 builds of the initial prototype within the project timeline and budget. At the first meeting of AG 1, key features for the prototype were discussed, and priority was assessed. The need for a food recognition model and key features it should have, such as simple and quick food recognition, were noted. How it would work in the care home environment, issues of data security and consent, symptoms that needed to be included, potential for a mood scale, and overall usability of the tool were also considered. Priorities were summarized into a set of features to be incorporated (Table 2).

Table 2. Key features identified through the advisory group, along with the priority level determined in collaboration with Game Doctor.

Feature	ID	Description	Priority
Food tracking via image recognition model	F1V1	Food tracking using off the shelf image recognition model	High
Patient profiles	F2V1	A way to create profiles for each patient or resident using the application	High
Correct food button input	F3V1	Input button to confirm that the model has recognized food accurately	High
Before or after plate recognition	F4V1	Ability of model or application to recognize plates as before or after meal	Medium
Manual input form	F5V1	Manual input form to record symptoms and mood via a 5-point Likert scale	High
Analytics	F6V1	Analytics to track user use of prototype	Medium
Opt-in or consent form	F7V1	Consent form for users during testing of prototype	High
Resident data storage	F8V1	Centralized storage of resident data so multiple devices can access resident data	High
Push notifications	F9V1	Notifications or prompts such as push notifications to attract the user back into the application	Medium

Next, Game Doctor produced a short video to guide participants on how to download the tool and use its initial features (ie, taking photos of plates of food before and after meals, logging symptoms via tick-box selections, and recording mood using a face emoji scale). The basic initial prototype was presented at the second meeting of AG 1. Game Doctor provided a walk-through of the basic prototype and presented different food recognition models Game Doctor. LogMeal was chosen as the food recognition model provider due to the size of its database and the accuracy of food recognition (refer to Table 3 for the models compared). This meeting lasted twice as long as the first session to allow for Game Doctor to talk through the functions of the tool and food recognition model, for the group to try it and test with singular foods, and then test with mixed foods over a restaurant meal. The prototype was tested by taking photos of fruits and vegetables after Game Doctor's presentation. Food in different presentations and textures was provided on plates for testing with the tool. Basic feedback was given when the group tested the tool with whole meals during dinner, with participants providing feedback on its usability. Other features identified to enhance the tool included having an "after" photo to calculate food eaten; adding an element to note whether food had been fortified (eg, high fat or additional protein added); including additional relevant medical conditions; and keeping the tool simple by, for example, using a tick-box approach and having symptom recording options appear when the "after" photo is uploaded.

Table 3. Different food recognition models available that could be integrated into the tool.

Provider	Pricing
Clarifai	Free for academic projects
LogMeal	Depends on project
FoodAI	To be confirmed
BiteAI	To be confirmed
Calorie Mama	To be confirmed

Key design issues noted in feedback included the positioning of the "confirm" screen, which sometimes overlapped the photo taken; how to enter a food if the database does not recognize it and provide a correct option; whether the highlighting circles link to each food on the plate; how to cancel the recognition system if there is an error in recognizing a food; and how plate size would be distinguished to calculate nutrition from different portion sizes. Important features discussed as being needed in the next build, based on feedback from using the tool, included (1) the ability to select resident ID and room number; (2) "before" and "after" photos; (3) adding a food fortification button to indicate the addition of butter, cream, protein powder, and so on; (4) having tick boxes for various symptoms that can be personalized to residents and structuring the page to follow the "after" photo or providing an option to enter symptoms later; (5) plate size option; (6) the option to add medical conditions to the user profile; and (7) having the symptoms page appear

XSL•FO
RenderX

right after the second photo, including options to indicate fullness, mood, swallowing difficulties, and so on.

Phase 2: Feasibility Testing of the Tool and Food Diaries

The next phase of the project involved further testing and feedback by users in their home settings. We recruited 10 community-dwelling older adults to test the tool for 2 weeks. Of the 10 participants, 9 (90%) managed to use the tool, while 1 (10%) had difficulty downloading it and withdrew from the study. Of the 9 participants, 8 (89%) used the tool for the full 2 weeks, while 1 (11%) used it for 1 week. Each day, they documented their food intake and any difficulties encountered with the tool in food diaries. In addition, members of AG 1 tested the tool in free-living conditions over 1 week and submitted comments about their experience through a brief web-based survey. The findings from the diaries and survey covered 3 topics: symptoms, features, and populations. Regarding symptoms, it was noted that the tool monitored a good range of symptoms specific to the target group; however, there were suggestions to include additional symptoms, such as thirst, confusion, and anxiety. Regarding features, it was reported that the tool was easy to use when it connected well with the food recognition model. Most of the group said that they did not feel that any key features were missing, although they noted that testing in care home settings would help identify site-specific needs. Suggested features included an "empty plate" button and a directory of foods. As for other populations and settings that might benefit from the tool, the survey responses highlighted community care, patients with dementia, family carers, and hospitals.

Phase 3

Overview

At the first meeting of AG 2, members suggested that care homes would face unique challenges in using the tool, different from those within the community. These might include potential difficulties in recognizing pureed food, fortified meals, or culturally specific foods (eg, haggis), residents with cognitive impairment not being able to use the tool, and the possibility for staff or visitors to assist with tool use. It was decided that we would trial the tool across a range of care homes with residents of varying abilities to gain direct feedback on these issues.

Care Home Testing and Focus Groups

The prototype was tested in 4 care homes with 8 staff members (n=2, 25% in each home) and in 1 care home with 6 residents (older adults: n=4, 67%; and younger adults [aged 54 y and 58 y] with learning disabilities: n=2, 33%). Of the 4 care homes, 3 (75%) did not recruit older adults for the testing due to the cognitive functioning status of their residents who would not be able to use the tool; therefore, the staff members used it themselves. The 6 residents who were recruited all tested the tool themselves. The staff recruited had from 6 months to 20 years of experience of working in the care sector. The older adults recruited were aged between 78 and 84 years and had lived in the care home for at least 6 months. After prototype testing in the care homes, the semistructured interviews with staff (4/8, 50%) and residents (1/6, 17%) identified (1) the pros and cons of using the tool in their care home, (2) how it can be adapted to enhance its usability, and (3) common symptoms reported in the tool. The resident who tested the tool was 1 (50%) of 2 younger residents with learning disabilities (aged 58 y). The main outcomes from these interviews are displayed in Textbox 1.

XSL•FO
RenderX

Textbox 1. Themes emerging from semistructured interviews with care home staff and residents.

Key design issues

- Wi-Fi issues slowed down the uploading of data
- Tool struggled to recognize pureed food
- Camera picked up patterns on plates as food
- Design features are needed to speed up the process

Key features

- Staff would prefer photo profiles of residents instead of names
- International Dysphagia Diet Standardization Initiative framework [29] and Malnutrition Universal Screening Tool [14] malnutrition risk score should be built in
- Traffic light system approach to identifying risk
- Self-choose whom to measure option
- Empty plate button
- Monitor fluid intake too
- Pureed food and fortification recognition is essential
- A notification or prompt system as a reminder

Positive feedback

- Recognized food well
- Residents liked monitoring their own food intake
- Enjoyed playing around with the tool and learned how to use it quickly

Key Design Issues and Usability

All care homes reported difficulties in using the tool when the home Wi-Fi connection was poor:

> *We had connection issues and found the app to be very slow or would freeze-could there be an easy offline system to use throughout the home that would automatically upload when a strong wifi connection was made?* [Staff, care home 2]

All care homes reported slow uploads, with the tool freezing; therefore, participants had difficulty clicking on the foods on the screen. This then led to slower identification of foods and a slow search for foods that were not identified, highlighting the need for features that would speed up this process.

A resident who was using the tool revealed that they did not have these issues because they just saved the photo to upload later when their Wi-Fi connection was stronger:

> *I found the tool easy to use, wifi is terrible here so I just waited until the signal was stronger and uploaded the photo then.* [Resident, care home 2]

Staff members expressed concern that if they did not upload photos and ratings immediately, they would not be able to do so later due to a change of shift or being given another role or task within the home. Furthermore, choking was identified as a major concern in all participating care homes, and many of the residents were on a pureed diet. The tool struggled to recognize pureed foods, highlighting the need for better pureed food recognition or an alternative solution:

> *A number of our residents are on pureed diets or we add thickening agents (which contain calories) to some of their foods, the tool could not properly detect it, although funnily enough we use molds shaped like the foods and it could sometimes pick that up.* [Staff, care home 1]

Staff members also wanted to add a measure of swallowing difficulties to the tool, which would link to food intake and the International Dysphagia Diet Standardization Initiative (IDDSI) framework [25]:

> *Having a feature and notes in IDDSI, malnutrition risk and fortification is really important for us and our residents.* [Staff, care home 4]

The current version did not link a swallowing difficulties feature, but adding this would encourage care homes to change the texture of the food.

Finally, when photos were taken of plates, the food recognition function of the tool sometimes misidentified the design patterns on the plate as food:

> *The plate was empty but it was saying the patterns were food and giving it a value.* [Staff, care home 3]

However, the more the tool was used, the less this became a problem because users worked out how to delete the circle around the patterned part that was misidentified as food. In addition, continued use improved the tool's accuracy through machine learning, reducing misidentifications over time.

XSL•FO
RenderX

Key Features Missing From the Tool

Regarding features needed in the tool for the care home setting, all homes reported the need for a resident profile instead of the current name function, which takes users straight to the diary:

> *We need a profile with the resident's picture which we choose from to quickly find the resident—we have a few residents with the same name.* [Staff, care home 2]

Staff members wanted to be able to upload a photo of each resident to click on instead of selecting a name, which would take them directly to the resident's profile. This profile would incorporate relevant information about the resident such as weight, changes in weight, issues with previous foods such as stomach problems or swallowing difficulties, risky foods and allergies, IDDSI [29] and MUST [14] scores, medications that may affect diet, and a category for food monitoring (high risk, medium risk, or low risk). Having the malnutrition risk score assessed via the tool and keeping track of IDDSI scores was a high priority for staff. These features would support staff, particularly new staff who are less familiar with a resident, in ensuring that the correct food is provided:

> *Choosing the picture would help with our new staff or our bank staff to recognize the resident, we have quite a high turnover here so a photo would quicken the recognition process.* [Staff, care home 4]

The profile would be accessed, leading to the diary, and a warning signal would be produced when a picture is taken to highlight risky foods for that resident. This was suggested as an important future feature:

> *If the tool was able to actually warn or prompt staff of certain foods it would really support our staff-we have a lot of residents here who have allergies or are choking hazards so it would support staff unfamiliar with these residents when giving them their meals.* [Staff, care home 2]

A key concern for care home staff using the tool was the time it would require to take the photo and deliver a meal to the resident while it was still hot. Some homes had 80 residents, and staff would not be able to monitor everyone in addition to their other duties. Upon further discussion, a traffic light system was suggested:

> *Having something like a traffic light system for overall nutrition risk of the resident as well as risk of the food would be brilliant. We could then monitor our resident over time and give more detailed reports when the dietitian came to who was priority.* [Staff, care home 3]

This would be a system that would categorize individuals as *green* (does not need regular monitoring), *amber* (needs monitoring but is not of major concern), and *red* (needs monitoring and may need observation or support with eating due to key issues such as swallowing difficulties or choking concerns). Care staff also wanted to be able to choose which residents to monitor because this would help reduce the time spent using the tool by prioritizing residents who needed monitoring the most while eliminating those who did do not require it.

Given concerns over staff time, methods to speed up the tool's use were highlighted as essential components to integrate. One suggestion was allowing staff to self-select individuals they needed to monitor based on their own experience of who needed support the most:

> *We know our staff well, we know who has problems when it comes to eating, and we know who to highlight to our manager, if we could choose who to monitor and have the tool to support our statements it would help.* [Staff, care home 1]

Another suggestion was to add an "empty plate" button:

> *Being able to just even say they ate everything would really help.* [Staff, care home 4]

This would eliminate the need for a second photo. A "half-finished" button was also suggested; however, it was highlighted that this could lead to users not knowing what parts of the meal had been eaten, resulting in an inaccurate estimate of nutritional intake:

> *Being able to click that they ate half would quicken things up, but saying that, we have certain residents who only eat their meat and leave their veggies.* [Staff, care home 2]

Having a database with preexisting photos based on current care home menus was identified to speed up use of the tool, with the staff only needing to take an "after" photo. Another suggestion to expedite the process was to work with the care homes to integrate their menus into the tool. Of the 4 care homes, 3 (75%) reported a 3-week menu that could be incorporated, and staff would pick from preassigned photos of each meal rather than take "before" photos of all meals. They would then only need to capture "before" photos for individuals with smaller portions or meal modifications (eg, substitution or fortification). Afterward, staff would take "after" photos of unfinished plates, using an "empty plate" button for completely finished meals.

Care staff mentioned the importance of monitoring fluid intake as well as food consumption. Linking fluid intake to bathroom breaks and urinary tract infection diagnoses would motivate staff to provide better fluid care for residents:

> *If we could monitor what they are drinking and then when we take them to the toilet, particularly during the night, it would be really useful so our staff can monitor who hasn't drank in a while. It would be even better if the tool could tell us when to stop giving them liquids or if certain liquids made them need more, to prevent them needing at night time.* [Staff, care home 3]

Staff felt that it would be useful to build a feature into the tool that gave personalized fluid feedback and indicated when best to give out fluids, the best type of fluids to give, and whether there should be any time restrictions to ensure the best level of hydration for each resident.

Given that most residents were on restricted diets—some requiring pureed food due to swallowing difficulties and others

XSL•FO
RenderX

needing fortified meals—having these dietary modifications recognized by, or incorporated into, the tool was deemed essential. Staff identified taking multiple photos as the only feasible solution that would still provide accurate nutritional information. This process would involve taking (1) a photo before pureeing, (2) a photo of the pureed meal, and (3) a photo after consumption. At the second meeting of AG 2, it was confirmed that further discussions with app developers would be needed to determine the quickest and most accurate method for implementing this approach.

Residents who tested the tool suggested having a notification or prompt system to remind them to use it. They mentioned that they sometimes forgot to log their food intake and became upset when they missed the opportunity to do so:

> *I eat at different times to the other residents, and was worried I wasn't allowed to use it out with mealtimes so I sometimes forgot to take pictures. This was annoying as I really enjoyed using the tool. Will I still have access to it?* [Resident, care home 2]

Residents reported that being able to set their own prompts or notifications would be beneficial. However, staff did not share this preference because they felt that they themselves would remember to use the tool at mealtimes as part of their work routine. The only instance in which staff considered a prompt system useful was for alerts warning them if a resident's meal contained an allergen or had previously posed a choking risk.

Care Home Staff Survey and Interviews

Altogether, 130 individuals working in the care industry completed the survey. Most of the respondents were health care staff working in a care home (62/130, 47.7%), followed by care workers (25/130, 19.2%), support staff (19/130, 14.6%), and managers (9/130, 6.9%). Of the 9 managers, 2 volunteered to participate in the postsurvey interview. Of the 130 respondents, 60 (46.2%) had worked in the care industry for >5 years, while 30 (23.1%) had worked for <2 years. When asked how they monitored residents' diets, most (114/130, 87.7%) did not provide specific examples or measurable approaches used in their care homes. Many mentioned menu planning, staff training, and nutritional analysis but without detailing the methods or explaining how food intake was tracked for each resident. Among those who provided examples, methods included written reports (9/16, 56%), diet diaries (4/16, 25%), and computer platforms (2/16, 13%). When asked about additional symptoms that they felt were important to monitor in residents but were not available in the current version of the tool, the most commonly mentioned were hunger (45/130, 34.6%), mood (41/130, 31.5%), and swallowing difficulties (39/130, 30%). However, these symptoms were already included in the tool but may need to be more visually prominent and not limited to being reported only after meals. Other suggested symptoms included tiredness and pain. When asked which features they felt should be added that were not available in the current version, the most common responses were profiles that use photos to identify resident rather than just their name (42/130, 32.3%), a traffic light system (36/130, 27.7%), and integration of IDDSI and MUST scores (36/130, 27.7%). Additional suggestions aligned with feedback from individual care homes, including a regular weigh-in section and fluid intake monitoring. These suggestions were discussed at the second meeting with AG 2, focusing on developing a basic version of the tool that meets general needs across care homes, with the option for individual care homes to add personalized features if deemed essential, provided they were willing to cover the additional cost.

The survey data were supported by the interviews with care home managers. Both managers reported using multiple methods to monitor residents' diets, although these were not always consistent across staff members or health professionals. They noted that having an app capable of tracking diet, identifying malnutrition, and monitoring associated symptoms would help standardize care across residents and life situations. The managers agreed that such a tool would allow dietary monitoring beyond the care home, including when residents were out with family or hospitalized, ensuring that symptoms could be monitored during these periods. A traffic light system that was visible to staff but not residents was considered highly important to ensure close monitoring of only residents who needed it. The managers were enthused about the tool and using it as part of care but had some concerns. They were worried about the time required to take photos and monitor symptoms, but when prompted during the interview, they came up with strategies themselves to speed up the process. Both managers expressed concerns about the cost of the tool and its uptake. While they acknowledged the tool's value, they noted that traditional paper methods were more economical unless the tool provided clear value for money. When discussing ways to make this feasible, they suggested (1) making it free for National Health Service (NHS) users, (2) implementing a tiered fee structure for access to different tool features, (3) generating revenue from external users to help subsidize the costs for care homes, and (4) finding funding to support its implementation in care homes.

Discussion

Principal Findings

We co-designed a prototype tool that enabled users to take photos of food plates using Android phones or tablets and link foods to a variety of chosen symptoms and mood ratings. The prototype was then tested in care homes. While the ultimate version of the tool would also calculate nutritional intake from the photos, the aim of this study was to co-design the tool to maximize usability. The co-design involved older adults, care home staff, dietitians, app developers, and researchers. The repeated design and test approach ensured that the prototype was co-designed with AGs, with participants' feedback used to make incremental improvements before testing whether the suggested features were usable and integral to the overall objectives of the tool. This agile methodology implemented by Game Doctor allowed for rapid prototyping and testing of each of the 3 builds by end users. Testing in the community and later in care homes revealed what worked well and what needed updating as well as features that should be included to make the tool specifically suitable for use in care home environments. This collaboration with end users was vital to developing a tool suited for diet and symptom monitoring in care homes.

XSL•FO
RenderX

Care home staff and residents provided feedback on how to ensure that the tool could be implemented in their care home and were very positive about using it overall, although it should be acknowledged that this may have been influenced by the relative experience of the staff. However, the range of experience across the testers gives us confidence that we have captured a range of views. All participants emphasized the tool's value in supporting staff; facilitating quicker diagnosis of dietary-related symptoms; and enabling earlier detection of malnutrition, which is an important issue in care homes [7]. They reported that they would not need to monitor every resident or use every aspect of the tool but that it would be very valuable to use and could additionally form part of their marketing approach to attract residents.

Care home staff and residents suggested several important features and ways to increase usability. Priority features that will be integrated into future design are enhanced resident profiles with a traffic light system, integration of IDDSI [29] and MUST [14] scores, and recognition of pureed food and fortified meals. Traffic light systems have been found to be successful in dietary monitoring previously by encouraging users to choose lower-calorie alternatives [30]. A similar approach could be applied to identifying "safe foods" in the current targeted group. This would circumvent the need for diet diaries and bringing in a dietitian to interpret data [15], potentially accelerating the recognition of malnutrition with the addition of a nutrition calculator and other tool features to enable appropriate actions to be taken before hospitalization becomes necessary. Previous tools [22,23] that have used photo-based dietary assessment have found this to be a valid, feasible, and user-friendly approach to monitoring dietary intake. However, these tools have focused on younger populations, who may have different dietary requirements and greater familiarity with technology compared to the current target group.

Regarding wider use across care homes and potential methods of commercialization, the care staff survey and manager interviews indicated a preference for the tool to be free for NHS users. This would support its adoption across services and ensure that those who most needed it could fully benefit from its features. Finding funding from the local council or government or having the cost subsidized by nonresident users was brought up as another way to keep costs low or eliminate them for care homes. Managers suggested that even staff members interested in using the tool for their own dietary monitoring could help subsidize costs. When discussing payment for the tool, managers remained positive about using it, provided it was not too expensive, which would divert funding from other activities. Their preference was to pay an annual fee that would give them access to the tool, including updates and maintenance support. A tiered fee structure was also suggested, where costs would increase based on the number of features accessed. Although the managers would prefer the tool to be free, they were still keen to use it even if it was associated with a small fee. The global revenue for nutrition apps was projected to reach US $5.4 billion in 2024, with an expected annual growth rate of 11.2% [31]. This highlights the popularity of dietary monitoring and the importance of working with care home organizations to develop a financially feasible model for the tool.

Co-design between end users, the tool developers, and the research team was integral to this tool's production. As previous tools [22,23] and reviews of image-assisted dietary assessment methods [32] have focused on younger populations, it was essential to involve older adults in the codevelopment process. The feedback gathered played a crucial role in shaping the tool's functionality and features, ensuring that it was suited to the intended environment and users. This emphasizes the importance of co-design and coproduction in this type of research and innovation [27]. Applying the 3-sprint approach to design enabled a continual process of refinement, resulting in the production of a tool created by and for older adults and staff in care homes. Without this iterative process, the research team could have estimated which features might be important for users in care homes, but direct lived-experience evidence was essential to verifying whether these were the right features to incorporate and ensuring the tool's feasibility for the intended end users. However, despite this inclusive approach to ensuring that the tool incorporated the most important features for the setting, several technical improvements are still needed, along with the integration of new features; for example, because this study focused on co-design and feasibility testing, it is not yet clear how well the tool scores in terms of the accuracy of food recognition, and further testing, as well as comparison to similar tools, will be needed to ensure that this critical aspect performs well. Future versions would need to use a specially trained model that is linked to known menus; this should improve food recognition substantially. Other tools have reported mixed accuracy estimates, ranging from 9% to 63% [33]. One advantage of our tool is that it allows manual corrections or direct entry of specific foods and can be directly linked to a care home's menu plan to increase the accuracy of food recognition. Many similar tools have been criticized for their inability to accurately assess the quantity of food eaten [33]. Our tool addresses this limitation by incorporating "before" and "after" meal photos, enabling it to calculate the actual amount consumed rather than just identifying the food on the plate. Future research is needed to evaluate food quantity estimation as well as food recognition performance. However, the tool has the potential to simplify and speed up malnutrition recognition and dietary monitoring in later life among older adults, who are at higher risk than younger populations.

Limitations

This study is not without limitations. While the use of co-design and coproduction ensured that the tool included relevant features suited to its intended setting, our testing sample lacked diversity. All care homes recruited were private rather than council funded, and all AG members and tool testers were of White British origin. However, this reflects the typical ethnic composition of the Scottish population [34], and there was diversity in terms of gender, age, occupation, and specific care home site across participants. Future research would benefit from recruiting more diverse samples, including participants from both council-funded and private care homes, as well as recruiting across a range of geographic locations, including inner-city areas where communities are more ethnically diverse. A second limitation was the specific database used, which was developed in Spain and optimized for a more European-style diet. As a result, it did

XSL•FO
RenderX

not include many foods commonly consumed in Scotland and misidentified some foods as typical Spanish dishes (eg, paella). However, a future direction for this research is to create a bespoke database of food consumed in Scottish care homes as well as exploring the possibilities of working with individual homes' menu planners to create a tool specific to each home. This approach could be fundamental to the economic model for maintaining and continually updating the tool as a way of charging for a bespoke service rather than tool *use* in future, allowing us to provide the tool free in other settings, such as the NHS. One limitation of the co-design and coproduction methodology was the diversity of feedback regarding the tool and its usability and key features. While some care homes wanted enhanced features, such as malnutrition risk scores, a traffic light system, and fluid measurement with urinary continence tracking, others preferred to keep the tool's functions simple to enable ease of use. Further discussions clarified that the enhanced features could be optional, tailored to individual care homes' needs, and potentially cost associated, which eliminated the contradictions in the feedback.

Future Directions

Building on the outlined future directions, the next steps for this research project are to incorporate the care home feedback into the next build of the tool to ensure its suitability for care homes. To do this, we plan to apply steps 3 to 6 of the 6SQuID framework [24], test the tool in a more diverse sample of care homes, link it to a specially developed database of food photos reflecting the typical Scottish care home diet, and explore cost-effectiveness options and sustainability strategies to ensure that the tool remains up to date.

Conclusions

This journey through the co-design of a digital food recognition tool has revealed both its considerable potential and areas for further design improvement and feature enhancement. As we progress toward further prototype refinement and wider testing in diverse care home settings, the feedback and recommendations from this project will guide our approach, ensuring that the tool is practical, feasible, and robustly designed to comprehensively monitor diet and identify malnutrition risk and food-related problematic symptoms in older adults in residential care homes.

Acknowledgments

The authors acknowledge the artificial intelligence company specializing in food recognition and nutrition analysis LogMeal (AIGecko Technologies SL, Artificial Intelligence and Deep Learning Food Division) for the use of its database and thank all advisory group members and tester participants. This research was funded by a UK Research and Innovation Healthy Ageing Catalyst grant (ES/X006271/1), an Economic and Social Research Council–Stirling Social Science Impact Acceleration Account institutional block award to the University of Stirling, and a specific exploratory–Impact Acceleration Account and routes to commercialization grant to JC and ACW (STI-23/24-P0001). The article processing charge was funded by the University of Stirling. The funders had no role in the study's design; the collection, analyses, or interpretation of data; the writing of the manuscript; or the decision to publish the results.

Conflicts of Interest

None declared.

References

1. Stratton R, Cawood A, Anderson L, Burden S, Collery R, Holmes E, et al. Malnutrion and nutrional care survey in adults: UK malnutrion awareness week, October 2022. BAPEN. 2022. URL: https://www.bapen.org.uk/pdfs/reports/mag/national-survey-of-malnutrition-and-nutritional-care-2022.pdf [accessed 2025-02-06]
2. Malnutrition. World Health Organization. URL: https://www.who.int/health-topics/malnutrition#tab=tab_1 [accessed 2024-12-13]
3. Roe A. The importance of eating well in later life. Age UK. 2018 Sep 30. URL: https://www.ageuk.org.uk/discover/2018/importance-of-eating-well-in-later-life/ [accessed 2023-05-12]
4. Profile of the older population living in England and Wales in 2021 and changes since 2011. Office for National Statistics. 2023 Apr 3. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/ageing/articles/profileoftheolderpopulationlivinginenglandandwalesin2021andchangessince2011/2023-04-03#:~:text=2.-,Population%20ageing,from%2016.4%25%20to%2018.6%25 [accessed 2024-03-26]
5. Dent E, Wright OR, Woo J, Hoogendijk EO. Malnutrition in older adults. The Lancet 2023 Mar;401(10380):951-966. [doi: 10.1016/s0140-6736(22)02612-5]
6. Norman K, Haß U, Pirlich M. Malnutrition in older adults-recent advances and remaining challenges. Nutrients 2021 Aug 12;13(8):2764 [FREE Full text] [doi: 10.3390/nu13082764] [Medline: 34444924]
7. State of the nation 2021: older people and malnutrition in the UK today. The Malnutrition Task Force. 2021. URL: https://www.malnutritiontaskforce.org.uk/sites/default/files/2021-10/State%20of%20the%20Nation%202020%20F%20revise.pdf [accessed 2023-03-26]
8. Schenker S. Undernutrition in the UK. Nutr Bull 2003 Mar 17;28(1):87-120. [doi: 10.1046/j.1467-3010.2003.00303.x]

XSL•FO
RenderX

9. Stratton R, Smith T, Gabe S. Managing malnutrition to improve lives and save money. BAPEN. 2018 Oct. URL: https://www.bapen.org.uk/pdfs/reports/mag/managing-malnutrition.pdf [accessed 2025-02-06]
10. Fávaro-Moreira NC, Krausch-Hofmann S, Matthys C, Vereecken C, Vanhauwaert E, Declercq A, et al. Risk factors for malnutrition in older adults: a systematic review of the literature based on longitudinal data. Adv Nutr 2016 May;7(3):507-522 [FREE Full text] [doi: 10.3945/an.115.011254] [Medline: 27184278]
11. Wei K, Nyunt MS, Gao Q, Wee SL, Ng TP. Frailty and malnutrition: related and distinct syndrome prevalence and association among community-dwelling older adults: Singapore longitudinal ageing studies. J Am Med Dir Assoc 2017 Dec 01;18(12):1019-1028. [doi: 10.1016/j.jamda.2017.06.017] [Medline: 28804010]
12. Norman K, Pichard C, Lochs H, Pirlich M. Prognostic impact of disease-related malnutrition. Clin Nutr 2008 Feb;27(1):5-15. [doi: 10.1016/j.clnu.2007.10.007] [Medline: 18061312]
13. Dean M, Raats MM, Grunert KG, Lumbers M, Food in Later Life Team. Factors influencing eating a varied diet in old age. Public Health Nutr 2009 Dec;12(12):2421-2427. [doi: 10.1017/S1368980009005448] [Medline: 19344544]
14. Introducing 'MUST'. BAPEN. URL: https://www.bapen.org.uk/must-and-self-screening/introducing-must/ [accessed 2024-01-29]
15. Ortega RM, Pérez-Rodrigo C, López-Sobaler AM. Dietary assessment methods: dietary records. Nutr Hosp 2015 Feb 26;31 Suppl 3:38-45 [FREE Full text] [doi: 10.3305/nh.2015.31.sup3.8749] [Medline: 25719769]
16. Kroeze W, Werkman A, Brug J. A systematic review of randomized trials on the effectiveness of computer-tailored education on physical activity and dietary behaviors. Ann Behav Med 2006 Jun;31(3):205-223. [doi: 10.1207/s15324796abm3103_2] [Medline: 16700634]
17. Fakhrou A, Kunhoth J, Al Maadeed S. Smartphone-based food recognition system using multiple deep CNN models. Multimed Tools Appl 2021 Aug 12;80(21-23):33011-33032. [doi: 10.1007/s11042-021-11329-6]
18. Sahoo D, Hao W, Xiongwei W, Le H, Achananuparp P, Lim EP, et al. FoodAI: food image recognition via deep learning for smart food logging. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2019 Presented at: KDD '19; August 4-8, 2019; Anchorage, AK URL: https://arxiv.org/abs/1909.11946
19. Banerjee S, Mondal AC. Nutrient food prediction through deep learning. In: Proceedings of the 2021 Asian Conference on Innovation in Technology. 2021 Presented at: ASIANCON 2021; August 27-29, 2021; Pune, India. [doi: 10.1109/asiancon51346.2021.9545014]
20. Zeng J, Zhang J, Li Z, Li T, Li G. Prediction model of artificial neural network for the risk of hyperuricemia incorporating dietary risk factors in a Chinese adult study. Food Nutr Res 2020;64 [FREE Full text] [doi: 10.29219/fnr.v64.3712] [Medline: 32047420]
21. Ho DK, Tseng SH, Wu MC, Shih CK, Atika AP, Chen YC, et al. Validity of image-based dietary assessment methods: a systematic review and meta-analysis. Clin Nutr 2020 Oct;39(10):2945-2959. [doi: 10.1016/j.clnu.2020.08.002] [Medline: 32839035]
22. Prinz N, Bohn B, Kern A, Püngel D, Pollatos O, Holl RW. Feasibility and relative validity of a digital photo-based dietary assessment: results from the Nutris-Phone study. Public Health Nutr 2018 Mar 6;22(7):1160-1167. [doi: 10.1017/s1368980018000344]
23. Budiningsari D, Syahrian F. Validity of a digital photo-based dietary assessment tool: development and initial evaluation. Nutr Health 2024 Jul 23:2601060241239095. [doi: 10.1177/02601060241239095] [Medline: 39043223]
24. Wight D, Wimbush E, Jepson R, Doi L. Six steps in quality intervention development (6SQuID). J Epidemiol Community Health 2016 May;70(5):520-525 [FREE Full text] [doi: 10.1136/jech-2015-205952] [Medline: 26573236]
25. Oliver K, Kothari A, Mays N. The dark side of coproduction: do the costs outweigh the benefits for health research? Health Res Policy Syst 2019 Mar 28;17(1):33 [FREE Full text] [doi: 10.1186/s12961-019-0432-3] [Medline: 30922339]
26. NIHR guidance on co-producing a research project. National Institute for Health and Care Research. 2024. URL: https://www.learningforinvolvement.org.uk/content/resource/nihr-guidance-on-co-producing-a-research-project/? [accessed 2024-02-23]
27. Smith B, Williams O, Bone L, the Moving Social Work Co-production Collective. Co-production: a resource to guide co-producing research in the sport, exercise, and health sciences. Qual Res Sport Exerc Health 2022 Mar 19;15(2):159-187. [doi: 10.1080/2159676x.2022.2052946]
28. Malterud K. Systematic text condensation: a strategy for qualitative analysis. Scand J Public Health 2012 Dec 04;40(8):795-805. [doi: 10.1177/1403494812465030] [Medline: 23221918]
29. The IDDSI framework (the standard). International Dysphagia Diet Standardisation Initiative. URL: https://www.iddsi.org/standards/framework [accessed 2025-02-25]
30. Osman M, Thornton K. Traffic light labelling of meals to promote sustainable consumption and healthy eating. Appetite 2019 Jul 01;138:60-71. [doi: 10.1016/j.appet.2019.03.015] [Medline: 30880087]
31. Nutrition apps- worldwide. Statista. URL: https://www.statista.com/outlook/hmo/digital-health/digital-fitness-well-being/health-wellness-coaching/nutrition-apps/worldwide [accessed 2024-12-13]
32. Höchsmann C, Martin CK. Review of the validity and feasibility of image-assisted methods for dietary assessment. Int J Obes (Lond) 2020 Dec 08;44(12):2358-2371 [FREE Full text] [doi: 10.1038/s41366-020-00693-2] [Medline: 33033394]

XSL•FO
RenderX

33. Van Asbroeck S, Matthys C. Use of different food image recognition platforms in dietary assessment: comparison study. JMIR Form Res 2020 Dec 07;4(12):e15602 [FREE Full text] [doi: 10.2196/15602] [Medline: 33284118]
34. Ethnicity. Scotland's Census. URL: https://www.scotlandscensus.gov.uk/census-results/at-a-glance/ethnicity/ [accessed 2025-02-25]

Abbreviations

6SQuID: 6 steps for quality intervention development
IDDSI: International Dysphagia Diet Standardization Initiative
MUST: Malnutrition Universal Screening Tool
NHS: National Health Service

Edited by J Wang; submitted 23.07.24; peer-reviewed by A Coltman, O Ueland, DR Crabtree; comments to author 11.11.24; revised version received 18.12.24; accepted 30.01.25; published 03.03.25.

Please cite as:
Connelly J, Swingler K, Rodriguez-Sanchez N, Whittaker AC
Identifying Food Preferences and Malnutrition in Older Adults in Care Homes: Co-Design Study of a Digital Nutrition Assessment Tool
JMIR Aging 2025;8:e64661
URL: https://aging.jmir.org/2025/1/e64661
doi:10.2196/64661
PMID:40053797

©Jenni Connelly, Kevin Swingler, Nidia Rodriguez-Sanchez, Anna C Whittaker. Originally published in JMIR Aging (https://aging.jmir.org), 03.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Values of Stakeholders Involved in Applying Surveillance Technology for People With Dementia in Nursing Homes: Scoping Review

Daniëlle van Gaans-Riteco[1,2], MSc; Annerieke Stoop[1], PhD; Eveline Wouters[1,3], PhD, MD

[1]Academic Collaborative Center Care for Older Adults, Tranzo, Scientific Center for Care and Wellbeing, Tilburg School of Social and Behavioral Sciences, Tilburg University, Tilburg, The Netherlands

[2]Groenhuysen, Roosendaal, The Netherlands

[3]School for Allied Health Professions, Fontys University of Applied Science, Eindhoven, The Netherlands

Corresponding Author:
Daniëlle van Gaans-Riteco, MSc
Academic Collaborative Center Care for Older Adults
Tranzo, Scientific Center for Care and Wellbeing
Tilburg School of Social and Behavioral Sciences, Tilburg University
Professor Cobbenhagenlaan 125
Tilburg, 5037DB
The Netherlands
Phone: 31 134662969
Email: d.p.c.vangaans-riteco@tilburguniversity.edu

Abstract

Background: Due to the progressive nature of dementia, concerns about the safety of nursing home residents are frequently raised. Surveillance technology, enabling visual and auditory monitoring, is often seen as a solution for ensuring safe and efficient care. However, tailoring surveillance technology to individual needs is challenging due to the complex and dynamic care environment involving multiple formal and informal stakeholders, each with unique perspectives.

Objective: This study aims to explore the scientific literature on the perspectives and values of stakeholders involved in applying surveillance technology for people with dementia in nursing homes.

Methods: We conducted a scoping review and systematically searched 5 scientific databases. We identified 31 articles published between 2005 and 2024. Stakeholder characteristics were extracted and synthesized according to the theory of basic human values by Schwartz.

Results: In total, 12 stakeholder groups were identified, with nursing staff, residents, and informal caregivers being the most frequently mentioned. Among stakeholder groups close to residents, values related to benevolence, security, conformity, and tradition were most commonly addressed. Furthermore, values such as self-direction, power, and achievement seemed important to most stakeholder groups.

Conclusions: Several stakeholder groups emphasized the importance of being and feeling involved in the application of surveillance technologies. In addition, they acknowledged the necessity of paying attention to stakeholders' perspectives and values. Across these stakeholder groups, values related to benevolence, security, and self-direction were represented, although various stakeholders assigned different meanings to these values. Awareness of stakeholders' perspectives demands a willingness to acknowledge each other's values and bridge differences.

(JMIR Aging 2025;8:e64074) doi:10.2196/64074

KEYWORDS

surveillance technology; nursing home; stakeholders; values; dementia; safety

XSL•FO
RenderX

Introduction

Background

Globally, people are living longer. Every country in the world expects growth in the number and proportion of older persons [1]. As the population ages, diseases such as dementia are diagnosed more often because age is their strongest known risk factor [2,3]. Dementia is a major cause of disability and dependency, affecting cognitive abilities and behavior, leading to an inability to signal when help is needed, which is associated with safety concerns [3,4]. Compared with 2015, the number of people affected with dementia will triple by 2050 [3].

In several countries worldwide, the number of care professionals is insufficient to meet the growing care demands [5]. In addition, global shortages of skilled care professionals will increase due to the aging workforce [6,7]. To meet the increasing care demands, documents from the Netherlands government show that by 2024, 1 in 4 employees should work in the health care sector, a figure that should rise to 1 in 3 by 2060 [8,9]. One of the proposed solutions to this challenge is allowing people with dementia to live at home as long as possible. However, due to their increasing care demands, a need for long-term care (LTC) facilities providing high-quality intensive dementia care will continue. Studies in different countries worldwide indicate that approximately 30% to 40% of people with dementia will eventually need a care environment in nursing homes [10-12]. The Netherlands is known for its high percentage of residents receiving end-of-life care in nursing homes, which makes nursing homes the most frequent place of death [13]. Consequently, there has been a large increase in health care expenditures for the population with dementia, especially in LTC facilities [12].

One recommendation to address these societal challenges in home care and nursing homes is to foster investment in health technologies that contribute to sustainable and high-quality care for people with dementia, such as assistive and innovative care technologies [3,14,15]. Care technologies can delay or replace admission in a nursing home and reduce the workload of nursing staff and informal caregivers in community care and nursing homes [16,17]. There are different types of care technologies. One is surveillance technology, which allows visual and auditory monitoring and registration of events, including residents' activities. Surveillance technologies include tagging and tracking technology, sensors, and audio and video surveillance [18,19]. Surveillance technologies are increasingly focused on supporting autonomy and respecting privacy while enhancing safety and individualized care for people with dementia [7,20-23].

Surveillance technologies are often regarded as a solution for ensuring safe and efficient health care, including in nursing homes [14,19,20]. These technologies have the potential to provide high-quality care and relieve nursing staff as staff shortages increase [7]. Due to the potential benefits of using surveillance technologies for quality of life and care, general attitudes toward these technologies have become more positive [16]. Nonetheless, surveillance technologies can affect privacy, autonomy, and freedom of movement [16,20,21]. Therefore, the use of these technologies should comply with regulations governing privacy and involuntary care, including requirements for subsidiarity, proportionality, and expediency. In addition, the use of surveillance technologies has to be justified in the care plan [24-26].

The application of surveillance technologies for people with dementia living in nursing homes is complex in practice. Successfully implementing care technologies, including surveillance technologies in psychogeriatric nursing homes, appears to be challenging as it involves more than just implementing a technological application successfully used elsewhere [27]. One of the greatest challenges in implementing care technologies such as surveillance technology seems to be integrating technology into the care process. Surveillance technologies affect residents and other primary stakeholders, such as residents' representatives and formal caregivers [16,28]. The involvement of these primary stakeholders, also known as end users, and secondary stakeholders such as managers, information and communication technology (ICT) employees, developers, and vendors of surveillance technologies is necessary to increase stakeholder commitment [28].

Early involvement of relevant stakeholders increases the likelihood of successful implementation [29]. A prerequisite for their involvement is knowing and acknowledging stakeholders' cultures, perspectives, and interests [29]. Given the broad spectrum of stakeholders involved in applying surveillance technologies for people with dementia in nursing homes, there is a great diversity of backgrounds, resulting in differences in values and interests [27,30]. Values represent what is (most) important to people and direct their attitudes, behaviors, and actions [31,32]. Differences in stakeholder values and interests can complicate the creation of support among stakeholders [30]. In addition, dealing with different perceptions and values among a range of stakeholders is a major challenge, further exacerbated by a limited understanding of stakeholders' values [33]. Therefore, a knowledge of these values can help explain decision-making processes, attitudes, and behaviors of persons or groups in different contexts [31,34,35].

This situation necessitates exploring the stakeholder groups that are involved in implementing surveillance technologies, and their respective values. Nursing staff and informal caregivers' attitudes toward using surveillance technologies [18,36] and their ethical dilemmas when using surveillance technology in psychogeriatric nursing homes have been explored [23,36-38]. For example, Rostad and Stokke [39] noted the high complexity of the LTC setting, involving "wicked problems," such as many and changing stakeholders, competing interests, and disagreements regarding the nature of problems. However, it remains unclear what these competing interests consist of. In addition, little is known about the variation in the perspectives and values of the stakeholders.

This Study

To the best of our knowledge, no literature review has been conducted to explore the stakeholders involved, their perspectives, and values in the application of surveillance technologies for people with dementia living in nursing homes. Therefore, this scoping review aimed to explore which stakeholders are described in the scientific literature concerning

XSL•FO
RenderX

surveillance technologies for people with dementia in nursing homes. In addition, we seek to identify what is known about these stakeholders' values.

Methods

Overview

We conducted a scoping review to systematically explore, map, and synthesize the characteristics of stakeholders involved in applying surveillance technology for people with dementia in nursing homes and identify existing knowledge gaps. The PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) checklist was used as a guideline for this review [40] (Multimedia Appendix 1). The corresponding steps were followed [41]: (1) identifying research questions; (2) identifying relevant literature in databases; (3) selecting the literature; (4) charting the data; and (5) collating, summarizing, and reporting the results.

Identifying the Research Questions

The research questions formulated were as follows: (1) Which stakeholders are involved in applying surveillance technology for people with dementia residing in psychogeriatric wards in nursing homes? and (2) What is known about the values of these stakeholders?

Identifying Relevant Literature

We believed that articles of interest had been published in psychological, health care, medical, nursing, and technological journals. Therefore, we conducted a literature search using the following databases: MEDLINE, CINAHL, PsycINFO, ACM Digital Library, and IEEE Xplore. Search terms encompassed the LTC setting and the use of surveillance technology for people with dementia. A search string for each database was developed and programmed with the help of an information specialist (Multimedia Appendix 2). A search was first performed in April 2022 and fully updated in August 2023 and December 2024.

Studies were eligible for inclusion if surveillance technology was applied to people with dementia in nursing homes or a comparable 24×7 LTC setting. The use of surveillance technology was evaluated or monitored using qualitative, quantitative, or mixed methods designs. Furthermore, studies had to be peer reviewed, written in English or Dutch, and published between 2002 and 2023. The year 2002 was chosen because it was then that literature on the implementation of surveillance technologies became increasingly prevalent. In addition, studies had to mention the stakeholders who were involved in the process of applying surveillance technology. Studies mostly focusing on assistive technologies, such as automatic lights, or supportive technologies, for example, medication dispensers, health care apps for managing chronic diseases, etc were excluded. Studies conducted in an experimental or laboratory setting and nonoriginal research, such as scoping reviews and systematic reviews, were excluded.

Literature Selection

First, duplicate studies were removed. Title screening was performed by one of the authors (DvG-R), and in case of doubt, one of the other authors (AS) was consulted. Two authors (DvG-R and AS) independently screened abstracts using the literature review management tool Rayyan (Rayyan Systems Inc) [42]. The full text of articles considered eligible by both authors was reviewed for relevance. In all the selection steps, the results were compared and discussed until a consensus was reached. In case of doubt, the third author (EW) was consulted.

Charting the Data

A format for further data extraction was agreed upon and included the title, authors, year, country, aim, study design, method of data collection, study population, sample size, setting, technology type, an overview of the results per identified stakeholder, and limitations for this scoping review. Using this format, 2 authors (DvG-R and AS) independently reviewed 10 (32%) of the 31 included articles. When the reviews were compared, only minor differences were found. The remaining articles were reviewed by DvG-R, who consulted one of the two other authors (AS or EW) when appropriate. When no consensus about data extraction was reached, the other author (EW or AS) was consulted.

Collating, Summarizing, and Reporting the Results

We categorized the findings from each article per stakeholder group in data extraction forms. Subsequently, an overview of results per stakeholder group was compiled. Through an inductive process, we categorized our findings into frequently mentioned words, such as acceptance; privacy; safety; freedom of movement; person-centered care; quality of life; quality-of-care characteristics; technology characteristics; and resident characteristics, involvement, concerns, and their values. The findings were linked to human values to deepen an understanding of stakeholders' perspectives. Describing and defining values is considered complex, and analyzing them is an even greater challenge [43]. Therefore, we used the theory of basic human values by Schwartz et al [31], an empirically tested framework of values that is recognized across many cultures [44]. This theory is an important and well-known theory and is widely used to predict attitudes and behaviors in different contexts and situations [34,35]. These values are grounded in the 3 universal requirements of human existence: the needs of individuals as biological organisms, requirements of coordinated social interaction, and survival and welfare needs of groups [31]. This framework conceptualizes values ordered by importance relative to one another, and they form a system of priorities for groups, societies, and individuals [45]. The refined theory of basic human values has 19 values grouped into 4 higher-order categories as follows: openness to change, self-enhancement, conservation, and self-transcendence [31]. Table 1 presents the motivational goals of the Schwartz values based on the circular motivational continuum [31].

XSL•FO
RenderX

Table 1. The 19 values of Schwartz et al [31] explained in terms of their motivational goals based on the circular motivational continuum.

Higher order value and values	Conceptual definition in terms of motivational goals
Openness to changes	
Self-direction-thought	Freedom to cultivate one's own ideas and abilities
Self-direction-action	Freedom to determine one's own actions
Stimulation	Excitement, novelty, and change
Self-enhancement	
Hedonism	Pleasure and sensuous gratification
Achievement	Success according to social standards
Power-dominance	Power through exercising control over people
Power-resources	Power to control material and social resources
Conservation	
Face	Security and power through maintaining one's public image and avoiding humiliation
Security-personal	Safety in one's immediate environment
Security-societal	Safety and stability in the wider society
Tradition	Maintaining and preserving cultural, family, or religious traditions
Conformity-rules	Compliance with rules, laws, and formal obligations
Conformity-interpersonal	Avoidance of upsetting or harming other people
Self-transcendence	
Humility	Recognizing one's insignificance in the larger scheme of things
Benevolence-dependability	Being a reliable and trustworthy member of the ingroup
Benevolence-caring	Devotion to the welfare and well-being of ingroup members and being empathic
Universalism-concern	Commitment to equality, justice, and protection for all people
Universalism-nature	Preservation of the natural environment
Universalism-tolerance	Acceptance and understanding of those who are different from oneself

Results

General Findings

We identified 4116 unique studies. After screening titles, 298 (7.2%) abstracts were screened. After applying inclusion and exclusion criteria, we selected 91 (30.5%) articles for full-text screening. Eventually, we included 29 (32%) articles in this study. The search was updated in August 2023, and December 2024, resulting in 2 additional articles, bringing the total to 31 articles. The included studies were published between 2005 and 2023. Studies were conducted in 14 countries: The Netherlands (n=7, 23%), Finland (n=2, 6%), Denmark (n=1, 3%), Norway (n=2, 6%), Sweden (n=2, 6%), Spain (n=1, 3%), Germany (n=1, 3%), Switzerland (n=1, 3%), the United Kingdom (n=3, 10%), the United States (n=5, 16%), Canada (n=3, 10%), Mexico (n=1, 3%), Singapore (n=1, 3%), and China (n=1, 3%). A total of 10 (32%) articles had a quantitative research design, 18 (58%) articles had a qualitative design, and 3 (10%) articles had a mixed methods design. Figure 1 shows the flowchart of the selection process. The key characteristics of the included articles are available in Multimedia Appendix 3 [7,18-22,37,38,46-68].

In total, 12 stakeholder groups were identified in the publications. The most frequently identified stakeholder groups were nursing staff (23/31, 74%), residents (20/31, 65%), and informal caregivers (13/31, 42%). The stakeholder groups care managers (9/31, 29%), developers of surveillance technologies (7/31, 23%), physicians (5/31, 16%), LTC administrators (4/31, 13%), maintenance employees (2/31, 6%), ICT (2/31, 6%), vendors of surveillance technologies (1/31, 3%), project managers (1/31, 3%), and academics (1/31, 3%) were mentioned less frequently. A total of 9 (29%) of the 31 articles included 4 or more stakeholder groups in their research. Table 2 shows the stakeholders' frequency of occurrence in the included articles.

XSL•FO
RenderX

Figure 1. Flowchart of the selection process.

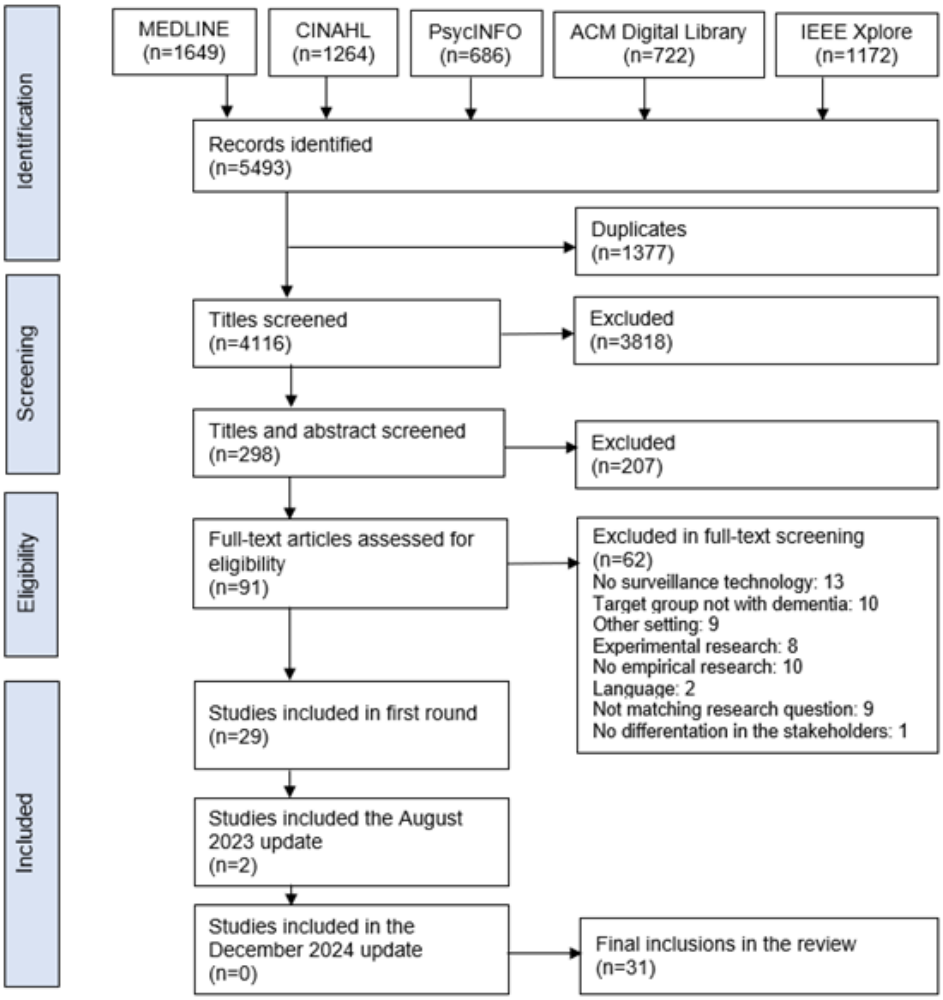


Table 2. Frequency of stakeholder occurrence in included articles and proxy perspectives[a].

Stakeholder	Frequency (N=31), n (%)	Proxy or partly proxy, n (%)[b]
Nursing staff	23 (74)	7 (30)
Residents	20 (65)	17 (85)
Informal caregivers	13 (42)	7 (53)
Managers	9 (29)	—[c]
Developers	7 (23)	—
Physicians	5 (16)	—
Long-term care administrators	4 (13)	—
Maintenance	2 (6)	—
ICT[d]	2 (6)	—
Vendors	1 (3)	—
Project managers	1 (3)	—
Academics	1 (3)	—

[a]Proxy perspective refers to the perspective of a stakeholder filled in by another stakeholder.

[b]Percentages in this column represent the proportion of articles in the corresponding cells in the "Frequency" column.

[c]Not applicable.

[d]ICT: information and communication technology.

XSL•FO
RenderX

Values of Stakeholders

Nursing Staff

Of all identified stakeholder groups, nursing staff (registered nurses, assistant nurses, and nurse aids) were mentioned in 23 (74%) out of 31 articles. Of these 23 articles, 7 (23%) articles were written from a proxy perspective of care managers (n=4, 17%), LTC administrators (n=2, 8%), and informal caregivers (n=1, 2%).

Several values of Schwartz were represented in the stakeholder group nursing staff. The values *benevolence* and *security* were strongly represented. The value *benevolence* is related to being helpful, supporting, and assisting those in need [31]. Staff's desire to be helpful and respond to the needs of residents also partly overlapped with the value *security*, which is characterized by striving to assure the safety of loved ones, taking precautions to prevent harm, and being warned in case of threats [45]. From the nursing staff perspective, surveillance technologies were most often applied to enhance the general safety of residents, mitigate the risk of falling, alert staff promptly, and contribute to efficient care [7,22,46]. Surveillance technologies can help nursing staff prioritize and direct their attention to where care and support are most needed [7,22,47]. Nursing staff expected monitoring technologies to increase the safety of residents [20], although they were aware of its limitations [21]. For example, they mentioned that surveillance technologies could not guarantee that residents would never fall [22]. Nursing staff rated residents' safety higher than they rated the experience of freedom [20], possibly due to the fear of being blamed for accidents or injuries to residents [22].

Because nursing staff felt responsible for their residents, the former expressed hesitation toward entirely relying on new surveillance technologies [48]. The values *conformity* and *tradition* were reflected in the nursing staff's conservativeness and preference to maintain their traditional routines. The nursing staff mentioned that altering their care routines was more difficult than expected. They tended to continue doing their rounds and checking on residents as usual despite the use of new technologies [49]. Managers' proxy perspectives also recognized the difficulties in altering care routines [48]. Although nursing staff tended to keep their traditions, they were generally supportive of new technologies that contributed to improvements in daily practice, particularly when the technologies functioned as intended [46]. However, new surveillance technologies never functioned properly from the outset, and implementing them often resulted in initial malfunctions [50]. These initial malfunctions, such as false alarms, poor Wi-Fi, and slow software had a negative influence on the level of acceptance, partly due to the nursing staff's great sense of responsibility for the residents [7,46,51]. These challenges limited the usability and accuracy of detecting unsafe situations and led to an even higher workload [7,22,46,48]. In addition, these challenges resulted in nurses' alarm fatigue and, as a result, nursing staff primarily relied on their previous experiences with residents' routines [21,48,51].

Moreover, nursing staff found new technologies challenging because these technologies required skills they did not previously need in their daily practice [52]. The extent to which procedures and instructions were tailored to the nursing staff's daily work, their shifts, and professional language influenced the nursing staff's openness to changes and the experienced extent of *self-direction* [7,20,51,53]. Moreover, they expected the technology vendors to help them, for instance, by providing support even outside regular business hours [7,51]. Nursing teams also felt supported when some of their colleagues took the lead in teaching them [7,47,51].

Although surveillance technologies contributed to providing optimal security and safety for residents, nursing staff expressed concerns about the impact of the use of monitoring technology on privacy, their competence in using the technologies, and the replacement of their roles [51,54]. This outcome relates to the value *self-direction.* Nursing staff expressed concerns that the monitoring technology was applied as a "big brother tool," indicating a lack of confidence [7,19,37,54]. Nursing staff preferred the application of codes of ethics and limited access by authorized professionals to protect their privacy and that of residents [7,53]. In addition, staff were concerned that monitoring had an impact on resident relationships [54] and that their role would be replaced by technology [51]. The latter was also recognized by care managers, as reflected in their proxy perspectives [51,52]. Particularly, older nursing staff expressed concerns about experiencing challenges while working with technology, resulting in a reserved attitude toward care technology [51]. The extent to which nursing staff felt involved influenced the freedom they experienced to determine their ideas and actions. This outcome impacted their openness to changes, represented by the values *self-direction in action and thought*. From a proxy perspective, care managers and administrators recognized these challenges faced by nursing staff, and they mentioned that nursing staff needed time to adapt to new care technologies before they could appreciate them [55,56].

Because surveillance technologies directly interfered with the nursing process, nursing staff expressed a desire to be involved from the beginning, for instance by being involved in discussions and decision-making processes [22], to be able to express their needs, opinions, and concerns [53]. The values *achievement* and *power* are related to the nursing staff's desire to be acknowledged and appreciated as important stakeholders.

In summary, the values *benevolence, security, conformity,tradition*, *self-direction in action and thought*, *achievement,* and *power* were represented among the stakeholder group of nursing staff.

Residents

Of the 31 included articles, 20 (65%) described residents as stakeholders, with 17 (85%) articles based on a (partial) proxy perspective of nursing staff (n=12, 47%), informal caregivers (n=4, 20%), LTC administrators (n=2, 10%), and developers of surveillance technologies (n=2, 10%).

Several residents who were interviewed, said they were aware of their dependency on care as a result of their cognitive decline. They adapted to life's circumstances, reflecting the value *humility*. Moreover, they indicated that care customized to their preferences and needs supported their independence and

XSL•FO
RenderX

contributed to their safety [52,55]. The value *(personal) security* is represented in the residents' feeling of being heard by care professionals. This means that professionals know where they are, respond to their alarms, and can care for them [22,53,57]. Consequently, surveillance technologies were perceived as part of the deal, improving their safety, and receiving individualized care [38,57]. This was recognized by nursing staff because they mentioned that surveillance technologies enhanced safety and contributed to the care, in line with residents' personal preferences, contributing to the latter's level of independence [52,53,55].

Residents knew that they partly gave up their privacy when they moved to a nursing home. Nevertheless, surveillance technologies could contribute to their feeling of *(personal) security* because the technologies protect their privacy. Coded doors, for example, could prevent other residents from unintentionally entering their rooms [19,38]. Although residents expressed a feeling of increased *(personal) security* due to surveillance technologies, they feared the consequences of these technologies as they expressed worries that these technologies would replace the valuable human contact with staff, because social contacts were an essential and highly valued part of life [37,38].

Furthermore, the values *conformity* and *tradition* seemed important to residents. These values include the avoidance of significant changes in their living environment, for instance, due to the use of surveillance technologies. Residents emphasized the importance of maintaining a feeling of homeliness [38,57]. In addition, surveillance technology should not jeopardize their feeling of homeliness or be (too) visible. Instead, surveillance technology should be aesthetically pleasing, easy to use, and not disrupt their daily routines [57,58]. Caregivers, as proxies, noticed that the extent of devices' intrusiveness to residents influenced the acceptability of devices [47,55,59]. Moreover, several residents expressed resistance to technology based on the usefulness they experienced [60]. The relevance of these factors was recognized in the proxy perspective of nursing staff and informal caregivers [47,59,61]. In addition, nursing staff experienced fewer nighttime disturbances for residents resulting in calmer nights [7,51]. Nursing staff noticed that the residents' openness to changes and the experienced *self-direction* decreased when there were more false alarms [57], which was recognized in the reluctance the latter expressed and is related to the experienced usefulness.

Residents' feelings of being stigmatized or being regarded as patients increased when their wishes regarding the visibility, appearance, and usability of surveillance technologies were not met [18]. This outcome aligns with the value *face*, which emphasizes maintaining one's public image and avoiding humiliation. Regarding this value, residents expressed greater concern about cameras than about other devices. These concerns were particularly about being recognizable in images while performing personal and hygienic activities, evoking feelings of intrusion and vulnerability [37,38,54]. This factor was similarly mentioned by managers when they expressed their concerns about the invasion of residents' privacy and dignity arising from surveillance cameras [54].

In summary, the values *humility, (personal) security, conformity, tradition, self-direction, and face* were most represented in the resident stakeholder group.

Informal Caregivers

Of the 31 articles, 13 (42%) mentioned informal caregivers (such as family caregivers and authorized representatives) as stakeholders. In total, 7 (53%) of these articles (partly) described a proxy perspective. Proxy perspectives were mostly represented by nursing staff (n=6, 46%), LTC administrators (n=2, 15%), and physicians (n=1, 8%).

Informal caregivers, mostly family members, expressed their concerns about the safety of their loved ones. Preventing (new) falls was often mentioned as a reason to use surveillance technology [19,22]. In their desire to contribute to the safety and well-being of their loved ones, deriving from the values *benevolence* and *security*, informal caregivers often valued residents' personal safety above possible threats to their privacy and freedom of movement [22,38]. This was also recognized by nursing staff and physicians, who observed informal caregivers' peace of mind when surveillance technologies were used [21,54]. In line with this value benevolence, informal caregivers felt responsible for the well-being of their loved ones. Informal caregivers feared that surveillance technologies would replace valuable human contact that arises from this value, and this fear may be reinforced by informal caregivers' awareness of the staff shortages in nursing homes [38,51]. This fear was recognized by nursing staff who indicated that they perceived it among informal caregivers [38,51]. Informal caregivers noted surveillance technologies should support nursing staff rather than replace them [38]. From a transcendent perspective, the value *universalism* seemed to be represented in informal caregivers' concerns about striving for equality and protection of people who were weak [31].

Informal caregivers mentioned they were willing to accept a wide range of surveillance technologies, including video surveillance, as long as they were convinced about the contribution these technologies made to the safety, quality of life, and well-being of their loved ones [22,37,38]. In addition, nursing staff and managers said that the level of informal caregivers' openness to changes was also determined by their perception of usefulness and their interest in technologies [53]. Informal caregivers' willingness to accept a broad range of surveillance technologies and their openness to changes reflected the value *self-direction*, embodying their ability to choose their goals and be involved in decision-making. The prerequisite of getting involved in (discussions about) applying surveillance technologies to their loved one [38,47,53,62] arises from the value *power*. Informal caregivers mentioned that these discussions should occur between all relevant stakeholders, such as residents, relatives, and nursing staff. In addition, informal caregivers wanted to be asked for formal consent as they were (authorized) representatives [38]. In practice, informal caregivers mentioned that they were not or not sufficiently informed about the available surveillance technologies [62].

In summary, the values *benevolence* and *security, universalism, self-direction, and power* were represented in the stakeholder group of informal caregivers.

XSL•FO
RenderX

Care Managers

Of the 31 articles, 9 (29%) mentioned care managers as stakeholders. Several care managers mentioned that their priority was to manage the 24×7 health care service and that surveillance technologies could contribute to achieving this aim [53,54]. The application of surveillance technologies increased the safety of residents, which is related to the values *benevolence* and *security*. Furthermore, data from these systems could be used to defend nursing homes against allegations of negligence leveled by families [19]. In line with this situation, care managers mentioned that data could help them monitor staff and hold them to account. However, care managers acknowledged that responding to an alarm was no guarantee that care was being provided [19]. The ethical objections against surveillance technologies that care managers mentioned were particularly aimed at the potential impact on residents' privacy, rather than the impact on nursing staff [19,22]. They acknowledged camera surveillance could contribute to a "big brother" effect and a culture of mistrust [19,54]. For care managers, surveillance technologies particularly seemed to represent values with a personal focus, namely, to have control over and manage the residential care facility. This outcome aligned with values such as *power* and *achievement.*

Care managers mentioned they were often insufficiently prepared for new ways of working and, subsequently, the different authority structures resulting from implementation strategies. In addition, they mentioned they were unable to make implementations a priority due to other organizational priorities [51]. The unpreparedness for changes they experienced could be a consequence of the changed ways of working, fear of the unknown, and clinging to the values *conformity and tradition*. Nevertheless, they were open to changes, although they also experienced unpreparedness for cocreation from several stakeholders [51]. This unpreparedness stems from an expectation of a tailored solution from vendors—not a realization that everyone's input including their own was a prerequisite for a joint implementation process [51]. Care managers taking the initiative in prioritizing reflections with other stakeholders to discuss dilemmas was mentioned as a facilitator for the successful use of surveillance technologies [51]. In this respect, care managers faced challenges in their *self-direction in action and thought* [31].

Furthermore, care managers felt that values *power* and *achievement* might be occasionally threatened. They indicated that they were concerned about increasing costs associated with new technologies, such as surveillance systems, while revenues remained stagnated [63]. In addition, surveillance technologies were not always as robust as they needed to be to withstand use in nursing home practice, leading to recurring costs due to damaged products [19]. Furthermore, care managers mentioned they felt restricted by rules and contractual obligations with vendors, which could hinder their access to technologies [19].

In summary, the values *benevolence* and *security* were somewhat represented in the stakeholder group of care managers. However, the values *power,achievement, conformity,tradition,* and *self-direction in action and thought* were more clearly represented.

Developers

A total of 7 (23%) of the 31 articles mentioned developers as stakeholders. Developers envisioned surveillance technologies would be used to enhance the safety of residents with dementia and improve the security of residents and nursing staff while respecting the privacy of both [19,55,58]. This outcome aligned with the value *security*.

In addition, developers mentioned they sought to use their technologies to enable nursing staff to support and assist residents with progressive diseases such as dementia. Due to the characteristics of the resident population, developers mentioned that surveillance technologies should be dynamic and scalable and be designed for failure and intensive use [55]. This outcome reflected the value *benevolence* as it underscored their determination to support and assist those in need.

Developers emphasized that by testing monitoring technologies new insights were created and improvements could be made. Testing in real life supported them to achieve success, in line with the values *achievement* and *power*. However, they acknowledged that high error rates in initial tests had an influence on the nursing staff's openness to change [52,58,60,61]. Conversely, many initial errors occurred due to unskillful use. A higher level of training for nursing staff could reduce these errors [58]. This outcome challenged developers' values *stimulation* and *self-direction* because they were generally excited about new technologies when they noticed the impact of their technologies on end users in nursing homes [60]. Furthermore, developers emphasized the importance of evaluating the effects of surveillance technologies in nursing homes [60]. Moreover, they underlined the importance of development in close collaboration with end users and specialists in dementia care to meet their needs and requirements [52,55]. Developers acknowledged that this close collaboration was an intensive process [52,55].

In summary, for the developer stakeholder group, the values *security, benevolence, achievement, power, stimulation,* and *self-direction* were most represented.

Physicians

Of the 31 articles, 5 (16%) included physicians as stakeholders in the application of surveillance technologies.

As with nursing staff and informal caregivers, the values *security* and *benevolence* were represented, as physicians agreed that providing safety was an important reason to apply these technologies. Moreover, they noted these technologies offered peace of mind to nurses and informal caregivers [21,37]. This outcome might explain physicians' high acceptance rate of surveillance technologies [55,58].

In summary, the values *security* and *benevolence* were clearly represented in the stakeholder group of physicians.

LTC Administrators

LTC administrators were mentioned in 4 (13%) of the 31 articles. LTC administrators expressed their vision of being at the forefront of implementing new technologies, especially in newly constructed nursing homes. Generally, new technologies

XSL•FO
RenderX

were considered an important solution for health care challenges [52]. In these administrators' perspectives, the values *self-direction in action and thought* and *stimulation* were reflected. However, LTC administrators acknowledged that implementing these technologies challenged institution's openness to changes because these new technologies demanded time and resources across various roles and professions [51]. LTC administrators noted that education for staff was an important issue; however, they sometimes questioned the abilities of nursing staff to master the new technologies despite offering education to the latter [37].

The values *power* and *achievement* emerged in the acknowledgment by LTC administrators that their residential care facilities were perceived as more attractive employers when they used modern surveillance technologies [37,52]. The financial investments required could force organizations to opt for a cheaper but more generic design, although this may be less suitable than preferred designs [19]. Consequently, LTC administrators could feel restricted in their control over resources. In line with care managers' perspective, LTC administrators mentioned that organizational contracts limited their scope and flexibility around product choices and ongoing maintenance [19].

In summary, the values *self-direction in action and thought, stimulation, power,* and *achievement* were represented in the stakeholder group of LTC administrators.

Values of the Other Stakeholders

The other stakeholder groups, that is, maintenance staff, ICT staff, vendors of surveillance technologies, project managers, and academics were mentioned only once or twice in the 31 articles included. Maintenance staff mentioned that although new surveillance technologies might seem to have a limited scope, only affecting night shift workers and residents, these technologies could also have an influence on janitors, cleaning staff, and substitute personnel [37,51]. In their view, maintenance staff members were also in need of information and education to accommodate the technologies in their (cleaning and maintenance) routines. When they were not informed or did not receive education, they reported difficulties, for example, in replacing sensors and reconnecting cables after cleaning [37,51]. Because a new system is only as good as the people who are responsible for operating it, the people operating the system could affect its reliability [37,51]. Maintenance staff expressed a desire to be acknowledged as important stakeholders and be involved in education, which relates to their openness to changes in the values *self-direction* and *stimulation*. In line with the apprehensions of nursing staff, maintenance staff expressed concerns about being observed [37]. For instance, maintenance staff sometimes felt threatened due to their openness to changes and *self-direction*, reducing their motivation to engage with new technologies.

ICT staff emphasized the importance of system and component interoperability for ensuring system reliability because new systems were often installed into existing systems and infrastructure [19,51]. However, they mentioned that this interoperability between systems often was not facilitated by manufacturers [19]. Furthermore, the staff highlighted the importance of their involvement from the outset in exploring this compatibility and interoperability to prevent compromises in residents' safety and security [51]. This outcome reflected the need for ICT staff to be involved and acknowledged as stakeholders, in line with the values *power* and *achievement*. However, in practice, it was seen that ICT staff only became seriously involved when systems were unstable or errors occurred [51].

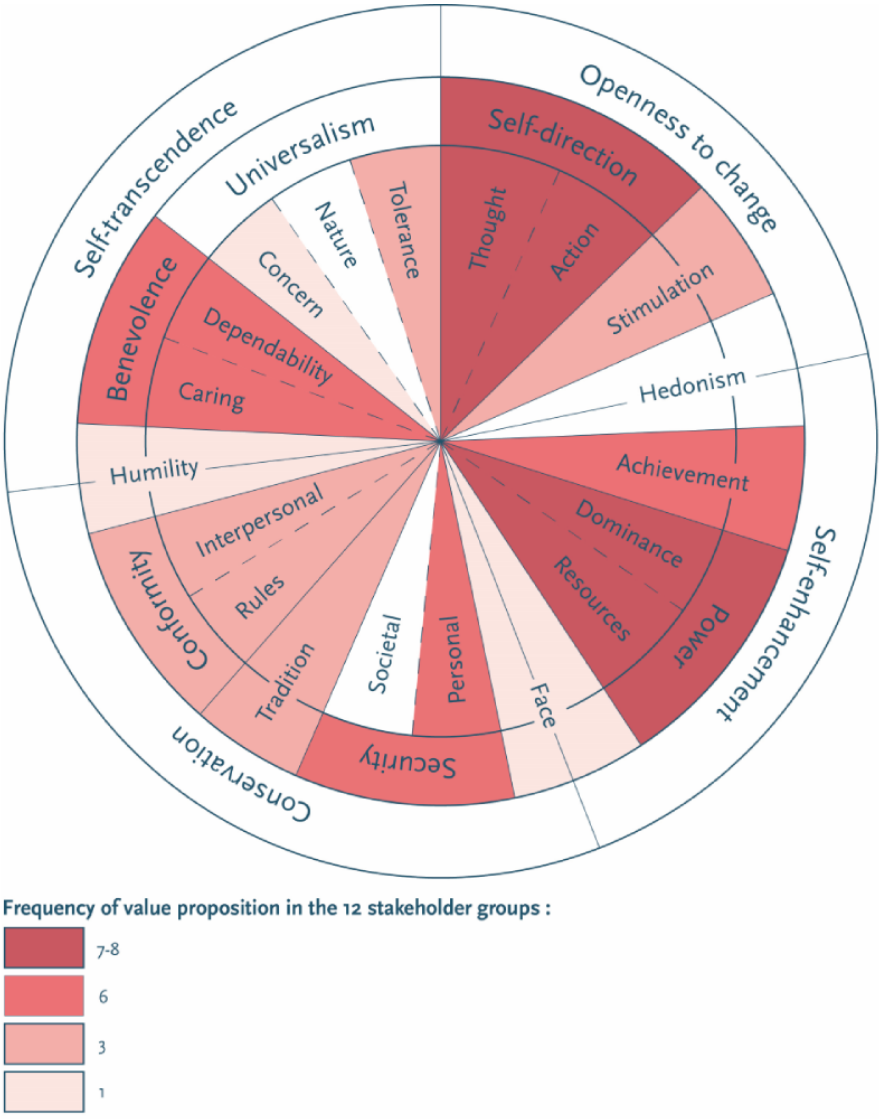
In the research by Dugstad et al [51], vendors of surveillance technologies and ICT were challenged to adopt a more socially focused approach to bridge differences between stakeholder groups. The vendors mentioned that despite their knowledge about their products, they needed the nursing staff's insights to ensure their technologies worked in specific care environments due to the great variety of care practices and infrastructure. Besides vendors striving for their own success, reflecting the values *power* and *achievement*, they were challenged to pursue a higher purpose, in line with the value *universalism*. Universalism relates to vendors' willingness to engage with other stakeholders to contribute to successful applications. In line with this value, ICT staff and vendors recognized being challenged to adopt a language that was more understandable for nursing staff who experienced differences in professional cultures and language (jargon) [51]. The ICT staff and vendors mentioned they were not used to adapting their language; therefore, misunderstandings were initially manifested due to a lack of knowledge and insight into each other's workflows. This situation challenged them to transcend their profession and jargon in order to contribute to a successful digital transformation [51]. Local project managers also underlined the importance of recognizing and bridging cultural differences as conditions for ensuring successful digital transformation. Hence, project managers identified their role as translators between different stakeholders [51], in line with the value *universalism*.

Finally, academics mentioned that the balance between freedom and security is important. They tended to value freedom of movement over security [20]. The socially focused value *benevolence* was reflected in being devoted to the welfare of others and protecting people with an increased care dependency. However, academics tend to attach great importance to residents' freedom to determine their own actions despite the dementia, which relates to the *value self-direction*.

Figure 2 presents the stakeholders' basic human values. The darker the color scheme, the more often this value is applied to various stakeholders.

XSL•FO
RenderX

Figure 2. Value palette for all stakeholders in the proposed circular structure of the 19 basic human values by Schwartz [32].



Discussion

Principal Findings

This scoping review aimed to identify the stakeholders involved in applying surveillance technologies for people with dementia in nursing homes and to describe the values of these stakeholders. Overall, 12 stakeholder groups were identified in 31 articles. The stakeholder groups of nursing staff, residents, and informal caregivers were most often mentioned in these articles. Several stakeholder groups, such as physicians, LTC administrators, maintenance staff, ICT staff, vendors, and project managers, were mentioned significantly less frequently [51,52,55].

Although many stakeholder groups emphasized the importance of being acknowledged and feeling involved as stakeholders [22,37,51,53,62], they said that their involvement should be improved [22,51,53,62]. Moreover, several stakeholders mentioned feeling dissatisfied when they did not feel sufficiently involved [53,62]. As dissatisfaction could detract from success, all relevant stakeholders should be involved [53]. Furthermore, in literature, the importance of involving both primary stakeholders, also known as end users, and secondary stakeholders (ie, more distantly involved stakeholders) is underlined as contributing to the successful implementation of care technologies, such as surveillance technologies [22,28,29].

To determine which primary and secondary stakeholders should be involved and engaged, it is important to identify these stakeholders [69]. Identifying stakeholders can be accomplished, for example, through a stakeholder analysis [69]. Several stakeholder matrices such as the power-interest matrix, the 3D matrix, and the responsible, accountable, consulted, and informed matrix offer tools to identify stakeholders and categorize their attributes such as their power, position, and level of interest [69,70]. In addition, these matrices support prioritizing who should be involved and to what extent. These matrices reflect the increased recognition of how the characteristics of stakeholders influence innovation and implementation processes [69,71,72]. Regardless of the distinction between stakeholders' level of involvement and their responsibilities, the resistance of any stakeholder can have an influence on the success of the implementation of care technologies.

XSL•FO
RenderX

Resistance by stakeholders is a challenge, necessitating the identification of alternative perspectives to a situation [73-75]. Therefore, besides identifying stakeholders when applying care technology, it is necessary to realize that care technologies strongly influence the care process and everyone involved. Therefore, cooperation between stakeholders who have previously rarely cooperated is essential. Hence, it is important to thoroughly know and acknowledge the different stakeholders and their differences and respect their differing interests [29].

Several implementation theories provide theoretical support to implementation processes, especially in complex contexts such as health care. One of the theoretical constructs for involving multiple stakeholders in implementation processes in health care is the normalization process theory (NPT) by May et al [76]. NPT provides tools to enhance an understanding of the social processes of thinking, enacting, and organizing work to implement and adopt interventions in care processes within health care organizations [77]. NPT distinguishes 4 constructs: coherence, cognitive participation, collective action, and reflexive monitoring. These 4 constructs emphasize the relevance of knowledge about the value that stakeholders assign to care technology and attention to stakeholders' willingness to participate and cooperate. NPT also focuses on keeping people engaged during the whole implementation process, including reflecting upon and appraising the effect of newly implemented technology [29,78].

A common thread across the NPT as well as other implementation theories is the need for cooperation and communication between stakeholders focusing on knowing and acknowledging each other's perspectives, values, and interests [29,79]. For instance, in the early 90s, Gregory and Keeney [72] wrote about the right of multiple stakeholders to be involved in policy decision processes and consequently balance between conflicting objectives. In addition, they mentioned that values, beliefs, cooperative potential, and stakeholders' concerns are known to influence the outcome of innovation processes [72]. Similarly, in business ethics, Dentoni et al [80] mentioned the complexity of multiple stakeholder involvement in dynamic settings because each stakeholder group has its own set of values, perceptions, and interests that may clash. This situation requires fundamentally different approaches that demand understanding differing values, complex settings, and acting upon uncertain knowledge [80].

The active participation of stakeholders challenges them to effectively collaborate with a critical and open stance toward their perspectives and values. Several stakeholders in this scoping review mentioned that this collaboration demands not only time and effort but also challenges them to adapt their jargon to interprofessional cultures and bridge their differences [51,52]. Although all stakeholders should be able to operate and communicate across boundaries between different practices with each other, collaborating with people is difficult and can lead to tensions and misunderstandings related to values and interests [81]. Such collaboration demands competence to perceive differences as learning opportunities and to cross boundaries between multiple stakeholders [81,82]. Collaboration between stakeholders is receiving increasing attention from organizations. In addition, organizations have shown a growing interest in creating value through participation and interaction with multiple stakeholders. However, until now, most of the attention has been given to creating value *for* stakeholders and not *with* them. In cases of increased awareness about stakeholders' input, values and interests may be well identified [83]. Thus, besides merely identifying and superficially involving stakeholders, it is important to pay attention to their perspectives, values, and interests.

Therefore, this scoping review also focused on what is already known about the stakeholder group values identified in the 31 articles reviewed. Several values in the theory of basic human values by Schwartz et al [31] were frequently represented among the stakeholder groups. The values *benevolence* and *security* were represented in 6 (50%) out of 12 stakeholder groups. This outcome is unsurprising given the progressive nature of dementia. Most stakeholder groups mentioned they experienced a feeling of being responsible for caring for residents with dementia and responding to their needs, in line with the value *benevolence* [21,22,37,38,55]. Concerning the value *security*, striving for safety for residents with dementia is often mentioned as a reason to apply surveillance technologies [7,22,46]. However, the meaning assigned to this value varies among the different stakeholder groups. The trade-off between safety and aspects such as privacy and freedom of movement differs among various stakeholders and appears to be related to how closely a stakeholder is involved with a resident [20]. For nursing staff, *security* is related to enhancing the safety of residents and mitigating their risk of falling [7,22,46]. This relates to the responsibility they feel for their residents because they feel accountable for accidents or injuries of residents [22,48]. For care managers, the value *security* is also related to their accountability for providing care, managing the residential care facility, assuring families that care is provided, and monitoring staff [19,54]. For residents, *security* is related to the feeling of being heard regarding their personal preferences and (care) needs and experiencing a feeling of homeliness [22,52,53,57]. Informal caregivers have concerns about their loved ones and they seem to experience more peace of mind when surveillance technologies are applied [21,54]. Consequently, informal caregivers often value safety above freedom of movement and the possible threats to privacy [19,22]. In contrast, academics, a group of stakeholders who do not have a close relationship with residents, tend to value freedom of movement above safety, giving high importance to the residents' experience of freedom and self-determination [20].

As is evident in the considerations regarding safety, surveillance technologies that are applied in practice affect the work and living environment of several stakeholder groups [37,51,53]. Accordingly, many stakeholders emphasized their desire to be and to feel involved in the application of surveillance technologies [22,53]. Arising from this desire, the values *self-direction in action and thought* were represented in 8 (67%) of 12 stakeholder groups. This outcome underlines the relevance of thoroughly involving stakeholders throughout the implementation and application process. In addition, the degree of stakeholders' openness to changes, their experienced *self-direction*, and their tendency to cling to traditional routines seem to be related to the extent to which they feel involved.

XSL•FO
RenderX

The more they feel involved, the more they are open to changes and willing to collaborate [22,38,47,53,62]. Being and feeling involved is especially an important issue for nursing staff because surveillance technologies directly interfere with the nursing process. Therefore, nursing staff were the most frequently cited stakeholders. They were mentioned in 23 (74%) of the 31 articles. Accordingly, they expected to be involved from the beginning; asked to express their needs, opinions, and concerns [53]; involved in decision-making; and acknowledged as important stakeholders [22]. Informal caregivers mentioned that they were willing to accept a broad range of technologies for their relatives as long as the former were informed about the surveillance technologies, were involved in deliberations about the technologies, and were asked for their consent regarding the application of the technologies [38,47,53,62].

Related to the value *self-direction* were the values *power* and *achievement*, representing stakeholders' desire to exert their influence on other people or use material resources and pursue success in competencies and performance. The values *power* and *achievement* were mentioned by 7 and 6 stakeholder groups, respectively. These values are a reflection that stakeholders such as nursing staff, informal caregivers, developers, ICT staff, and vendors want to feel recognized in their knowledge and experiences, and want to exert their influence in the application of surveillance technologies [22,38,51-53,61,62]. For care managers and LTC administrators, the values *power* and *achievement* are related to control over and the management of residential care facilities [53,54].

Although several values such as benevolence, security, self-direction, power, and achievement were represented by most of the stakeholders, the various stakeholder groups assigned different meanings to these values. Knowing and understanding diverse stakeholders' perspectives and attitudes, including the different meanings they attribute to values, is crucial because this knowledge and understanding influence the adoption and use of technologies [84]. Integrating multiple perspectives is valuable to fully understand the complexities of care practices [29] and dementia care technology [84]. In addition, it is challenging to distinguish whether stakeholders give meaning to a certain value based on their interests or whether they act from the resident's perspectives. This situation demands that designated persons in health care organizations have courage and take the lead in initiating meaningful and in-depth conversations where the diverse stakeholders will be challenged to communicate across their boundaries, looking beyond their own perspectives.

Methodological Considerations

A proxy perspective was often observed in the stakeholder groups of nursing staff, residents, and informal caregivers. This outcome is not unusual because proxy perspectives often originate from stakeholders with whom one works or lives [85]. In addition, proxy perspectives are often seen in stakeholder groups that collaborate and deal with matters that touch on values and interests. Hence, a proxy perspective was regularly seen among nursing staff and informal caregivers in this scoping review. Furthermore, when making assumptions, people are less understanding of others' actual motivations [85]. Accordingly, Kloos et al [86] researched the well-being of residents and found that nursing staff tended to overestimate the well-being of residents. Their study underlines the importance of combining proxy assessments with self-reports whenever possible. In addition, Kunicki et al [87] noted that the level of involvement of proxies and their sense of well-being could influence their perception of the resident's preferences. Kunicki et al [87] recommended that proxies find methods to better understand residents' preferences when residents were not able to express their preferences properly [87]. This scoping review raises the question of whether proxy perspectives of residents with dementia reveal resident's perspectives. Although it is challenging to interview residents with dementia due to the impairments related to their condition, it is possible to involve them. Therefore, in future research, residents should be made participants in research to understand their perspectives [88,89].

Strengths and Limitations

Previous studies primarily focused on facilitating and limiting factors and ethical dilemmas when applying surveillance technologies to people with dementia [18,23,36-38]. However, this scoping review is the first to explore which universally recognized human values are reflected in the experiences and opinions of stakeholders. Because values reflect human thinking and determine attitude and behavior, it is crucial to consider the underlying values that explain stakeholders' behavior and reactions [34,35,79]. Stakeholders' values were classified using the basic human values model by Schwartz. Although this model offers an empirically tested theory to predict attitudes and behaviors in different contexts and situations, it is possible that not all opinions and experiences of stakeholders could be categorized using the model. Furthermore, it is likely that the stakeholders' values do not provide a complete picture of the values that stakeholders hold in practice because the representation of stakeholders' values in this scoping review is based on information that did not primarily focus on stakeholders' values.

Recommendations for Research

Although this scoping review identified 12 stakeholder groups, most (26/31, 84%) of the articles included in the review described only 4 or fewer stakeholder groups. Therefore, various stakeholder groups were underrepresented. Future research should emphasize the involvement of all relevant stakeholders. In addition, our explorations of stakeholders' values revealed there is insufficient information about stakeholders' values. Hence, more research about the values that influence stakeholders' actions and decisions should be conducted. In exploring stakeholders' perspectives and values, a proxy perspective should be avoided where possible. In future research, inviting stakeholders to look beyond their perspectives and boundaries could be useful in mitigating language and knowledge boundaries between different stakeholders. This approach could facilitate constructive cooperation between stakeholders. In addition, efforts should be made to include residents with dementia in research to explore their perspective rather than assumptions about residents' perspectives being largely based on a proxy perspective. Listening to people with dementia can enhance the quality, relevance, and impact of

XSL•FO
RenderX

dementia research, which contributes to the enhancement of knowledge based on what we learn *from* them and their informal caregivers, in order to create knowledge *with* them [90]. Participatory research could apply to this kind of research about complex dynamic subjects; research should be conducted with rather than about stakeholders.

Conclusions

All stakeholders involved in applying surveillance technologies expressed a desire for their perspectives and values to be acknowledged. This desire stems from the human need to be acknowledged and appreciated. Moreover, all the stakeholders manifested a willingness to be engaged and participate. The broad acknowledgment and involvement of stakeholders and an understanding of their perspectives and values contribute to the successful implementation and application of surveillance technologies for people with dementia in nursing homes. Therefore, when applying surveillance technologies for people with dementia, residential care facilities are expected to intensively collaborate with an increasing number of stakeholders. Therefore, stakeholders' active engagement, with attention to everyone's perspectives and values, is more important than ever.

Acknowledgments

The authors thank information specialist C Niehot for the support in drawing up the search string for this scoping review. In addition, the authors thank Professor Dr M Verkerk for the conversations about values, interests, and ideals and Professor Dr S Schwartz, Drs D Arets, R van de Vorst, and Dr B Wernaart for the helpful feedback in using the theory of values by Schwartz et al [31] to organize the results of this scoping review according to human values.

Authors' Contributions

All authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA-ScR checklist.
[PDF File (Adobe PDF File), 167 KB - aging_v8i1e64074_app1.pdf]

Multimedia Appendix 2
Overview of the search string for the 5 databases.
[DOCX File , 22 KB - aging_v8i1e64074_app2.docx]

Multimedia Appendix 3
Mapping overview of the included articles.
[DOCX File , 73 KB - aging_v8i1e64074_app3.docx]

References

1. Ageing and health. World Health Organization. 2024 Oct 1. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2024-12-12]
2. Dementia in Europe Yearbook 2019: estimating the prevalence of dementia in Europe. Alzheimer Europe. 2019. URL: https://www.alzheimer-europe.org/sites/default/files/alzheimer_europe_dementia_in_europe_yearbook_2019.pdf [accessed 2023-05-04]
3. Global action plan on the public health response to dementia 2017 - 2025. World Health Organization. 2017 Dec 07. URL: https://www.who.int/publications/i/item/global-action-plan-on-the-public-health-response-to-dementia-2017---2025 [accessed 2023-05-04]
4. Huijsman R, Boomstra R, Veerbeek M, Dopp C. Zorgstandaard dementie. Dementiezorg voor Elkaar. 2020. URL: https://www.zorgstandaarddementie.nl/zorgstandaard [accessed 2021-09-21]
5. Strengthening the dementia care workforce: a public health priority. American Public Health Association. 2020 Oct 24. URL: https://www.apha.org/policies-and-advocacy/public-health-policy-statements/policy-database/2021/01/13/strengthening-the-dementia-care-workforce [accessed 2023-12-18]
6. Global strategy on human resources for health: workforce 2030. World Health Organization. 2020 Jul 07. URL: https://www.who.int/publications/i/item/9789241511131 [accessed 2023-05-04]
7. Emilsson M, Karlsson C, Svensson A. Experiences of using surveillance cameras as a monitoring solution at nursing homes: the eldercare personnel's perspectives. BMC Health Services Research 2023 Feb 10;23(1):144 [FREE Full text] [doi: 10.1186/s12913-023-09130-2] [Medline: 36765321]

XSL•FO
RenderX

8. Zorg voor de toekomst: over de toekomstbestendigheid van de zorg. Sociaal-Economische Raad. 2020 Jun. URL: https://www.ser.nl/-/media/ser/downloads/adviezen/2020/zorg-voor-de-toekomst.pdf [accessed 2023-05-04]
9. Bijleveld C, Boot A, Bovens M, Engbersen G, Hulscher S. Kiezen voor houdbare zorg. Wetenschappelijke Raad voor het Regeringsbeleid. 2021. URL: https://www.wrr.nl/publicaties/rapporten/2021/09/15/kiezen-voor-houdbare-zorg [accessed 2022-05-01]
10. Dementia statistics hub. Alzheimer's Research UK. URL: https://dementiastatistics.org/ [accessed 2023-04-05]
11. Mjørud M, Selbæk G, Bjertness E, Edwin TH, Engedal K, Knapskog AB, et al. Time from dementia diagnosis to nursing-home admission and death among persons with dementia: a multistate survival analysis. PLoS One 2020 Dec 4;15(12):e0243513 [FREE Full text] [doi: 10.1371/journal.pone.0243513] [Medline: 33275638]
12. Factsheet dementie. Vektis. 2022 Jan 24. URL: https://tinyurl.com/59xxymzn [accessed 2022-01-27]
13. Bauer AK, Fassmer AM, Zuidema SU, Janus SI, Hoffmann F. End-of-life care in German and Dutch nursing homes: a cross-sectional study on nursing home staff's perspective in 2022. Arch Public Health 2024 Jun 14;82(1):85 [FREE Full text] [doi: 10.1186/s13690-024-01316-2] [Medline: 38877598]
14. Nap HH, van der Weegen S, Cornelisse L, Lukkien D, van der Leeuw J. Zorgdomotica in de nachtzorg: onderzoek naar de mogelijkheden en onmogelijkheden van de nieuwe generatie zorgdomotica. Vilans. 2015. URL: https://www.researchgate.net/publication/287216465_Zorgdomotica_in_de_nachtzorg [accessed 2023-05-04]
15. Stroomversnelling voor krappe arbeidsmarkt in verpleeghuiszorg – uitgebreide rapportage. Ministerie van Volksgezondheid, Welzijn en Sport. 2020 Dec. URL: https://www.igj.nl/publicaties/rapporten/2020/12/10/stroomversnelling-voor-krappe-arbeidsmarkt-in-verpleeghuiszorg-%E2%80%93-uitgebreide-rapportage [accessed 2023-05-04]
16. van der Roest H, van der Kuil M, Overbeek A, Hartstra E. 414 - Supporting staff delivering person-centered care to people with dementia in Dutch nursing homes. Int Psychogeriatr 2021 Oct;33(Supplement 1):37-38. [doi: 10.1017/s1041610221001733]
17. Mohammadnejad F, Freeman S, Klassen-Ross T, Hemingway D, Banner D. Impacts of technology use on the workload of registered nurses: a scoping review. J Rehabil Assist Technol Eng 2023 Jun 17;10:20556683231180189 [FREE Full text] [doi: 10.1177/20556683231180189] [Medline: 37342268]
18. Niemeijer AR, Depla MF, Frederiks BJ, Hertogh CM. The experiences of people with dementia and intellectual disabilities with surveillance technologies in residential care. Nurs Ethics 2015 May;22(3):307-320. [doi: 10.1177/0969733014533237] [Medline: 24913545]
19. Hall A, Brown Wilson C, Stanmore E, Todd C. Moving beyond 'safety' versus 'autonomy': a qualitative exploration of the ethics of using monitoring technologies in long-term dementia care. BMC Geriatr 2019 May 24;19(1):145 [FREE Full text] [doi: 10.1186/s12877-019-1155-6] [Medline: 31126240]
20. Niemeijer AR, Frederiks BJ, Depla MF, Legemaate J, Eefsting JA, Hertogh CM. The ideal application of surveillance technology in residential care for people with dementia. J Med Ethics 2011 May;37(5):303-310. [doi: 10.1136/jme.2010.040774] [Medline: 21292695]
21. Zwijsen SA, Depla MF, Niemeijer AR, Francke AL, Hertogh CM. Surveillance technology: an alternative to physical restraints? A qualitative study among professionals working in nursing homes for people with dementia. Int J Nurs Stud 2012 Feb;49(2):212-219. [doi: 10.1016/j.ijnurstu.2011.09.002] [Medline: 21962335]
22. Hall A, Wilson CB, Stanmore E, Todd C. Implementing monitoring technologies in care homes for people with dementia: a qualitative exploration using Normalization Process Theory. Int J Nurs Stud 2017 Jul;72:60-70 [FREE Full text] [doi: 10.1016/j.ijnurstu.2017.04.008] [Medline: 28494333]
23. Niemeijer A, Depla M, Frederiks B, Hertogh C. Toezichthoudende domotica: een handreiking voor zorginstellingen. VUmc. 2012. URL: https://nvavg.nl/wp-content/uploads/2014/03/2012-Toezichthoudende-domotica-een-handreiking-voor-zorginstellingen.pdf [accessed 2022-04-12]
24. Wet zorg en dwang. Ministerie van Volksgezondheid, Welzijn en Sport. 2019. URL: https://www.dwangindezorg.nl/wzd [accessed 2021-05-28]
25. Handreiking Wet zorg en dwang voor specialisten ouderengeneeskunde & artsen verstandelijk gehandicapten in intramurale setting. Verenso. 2022 Jun. URL: https://www.verenso.nl/_asset/_public/Richtlijnen_kwaliteit/richtlijnen/Handreiking-Wzd-SO-en-AVG_versie-2-0-_DEFINITIEF.pdf [accessed 2025-02-18]
26. Recommendations on how to improve the legal rights and protection of adults with incapacity due to dementia. Alzheimer Europe. 2017. URL: https://tinyurl.com/3w9uby7f [accessed 2023-04-05]
27. Greenhalgh T. How to improve success of technology projects in health and social care. Public Health Res Pract 2018 Sep 27;28(3):2831815 [FREE Full text] [doi: 10.17061/phrp2831815] [Medline: 30406256]
28. Wernaart B. In: Wernaart B, editor. Moral Design and Technology. Wageningen, The Netherlands: Wageningen Academic Publishers; 2022.
29. Wouters EJ. Technologie: een zorg minder? Tilburg University. 2018. URL: https://www.tilburguniversity.edu/sites/default/files/download/Oratie_Wouters_TilburgU_5okt2018_1.pdf [accessed 2021-12-10]
30. Verkerk MJ, Holtkamp FC, Wouters EJ, van Hoof J. Professional practices and user practices: an explorative study in health care. Philos Reformata 2017 Dec;82(2):167-191. [doi: 10.1163/23528230-08202001]

XSL•FO
RenderX

31. Schwartz SH, Cieciuch J, Vecchione M, Davidov E, Fischer R, Beierlein C, et al. Refining the theory of basic individual values. J Pers Soc Psychol 2012 Oct;103(4):663-688. [doi: 10.1037/a0029393] [Medline: 22823292]
32. Holtkamp FC, Wouters EJ, Verkerk MJ. Understanding user practices when drawing up requirements-the case of designing assistive devices for mobility. Int J Environ Res Public Health 2019 Jan 24;16(3):318 [FREE Full text] [doi: 10.3390/ijerph16030318] [Medline: 30678372]
33. Keller EJ, Crowley-Matoka M, Collins JD, Chrisman HB, Milad MP, Vogelzang RL. Fostering better policy adoption and inter-disciplinary communication in healthcare: a qualitative analysis of practicing physicians' common interests. PLoS One 2017 Feb 24;12(2):e0172865 [FREE Full text] [doi: 10.1371/journal.pone.0172865] [Medline: 28235088]
34. Kleyweg J. Moraliteitontwikkeling en de 10 waardentheorie van Schwartz: exploratief onderzoek naar de ontwikkeling van moreel oordelen in de maatschappij anno 2006 en de relatie tot de waardentheorie van Schwartz. Universiteit Twente. 2006 Aug 31. URL: https://essay.utwente.nl/58804/2/scriptie_J_Kleyweg.pdf [accessed 2023-08-12]
35. Fujihara R, Montezano L, Alfinito S. Are we ready for new technologies? The relation between human values and technology readiness applied to m-commerce in Brazil. Organ Soc 2022 Jul;29(102):537-568. [doi: 10.1590/1984-92302022v29n0024EN]
36. Wangmo T, Lipps M, Kressig RW, Ienca M. Ethical concerns with the use of intelligent assistive technology: findings from a qualitative study with professional stakeholders. BMC Med Ethics 2019 Dec 19;20(1):98 [FREE Full text] [doi: 10.1186/s12910-019-0437-z] [Medline: 31856798]
37. Dorsten AM, Sifford KS, Bharucha A, Mecca LP, Wactlar H. Ethical perspectives on emerging assistive technologies: insights from focus groups with stakeholders in long-term care facilities. J Empir Res Hum Res Ethics 2009 Mar;4(1):25-36. [doi: 10.1525/jer.2009.4.1.25] [Medline: 19374478]
38. Sallinen M, Hentonen O, Teeri S. Ethical dilemmas related to the use of safety technology in service house environments. Scand J Caring Sci 2020 Mar;34(1):199-205. [doi: 10.1111/scs.12721] [Medline: 31250937]
39. Rostad HM, Stokke R. Integrating welfare technology in long-term care services: nationwide cross-sectional survey study. J Med Internet Res 2021 Aug 16;23(8):e22316 [FREE Full text] [doi: 10.2196/22316] [Medline: 34398791]
40. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
41. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005;8(1):19-32 [FREE Full text] [doi: 10.1080/1364557032000119616]
42. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 05;5(1):210 [FREE Full text] [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
43. Arets D, Ilievski D, Wernaart BF, Kamp JA, van Vliet M, Emini S. The right to Öffentlichkeit. In: Applied Human Rights. Wageningen, The Netherlands: Wageningen Academic; Jul 27, 2023.
44. Ponizovskiy V, Ardag M, Grigoryan L, Boyd R, Dobewall H, Holtz P. Development and validation of the personal values dictionary: a theory–driven tool for investigating references to basic human values in text. Eur J Pers 2020 Sep 01;34(5):885-902. [doi: 10.1002/per.2294]
45. Schwartz SH. An overview of the Schwartz theory of basic values. Online Read Psychol Cult 2012 Dec;2(1):1-20. [doi: 10.9707/2307-0919.1116]
46. Coahran M, Hillier LM, Van Bussel L, Black E, Churchyard R, Gutmanis I, et al. Automated fall detection technology in inpatient geriatric psychiatry: nurses' perceptions and lessons learned. Can J Aging 2018 Jul 03;37(3):245-260. [doi: 10.1017/s0714980818000181]
47. Nijhof N, van Gemert-Pijnen JE, de Jong GE, Ankoné JW, Seydel ER. How assistive technology can support dementia care: a study about the effects of the IST Vivago watch on patients' sleeping behavior and the care delivery process in a nursing home. Technol Disabil 2012 May 01;24(2):103-115. [doi: 10.3233/tad-2012-0339]
48. Niemeijer AR, Depla M, Frederiks B, Francke AL, Hertogh C. CE: original research: the use of surveillance technology in residential facilities for people with dementia or intellectual disabilities: a study among nurses and support staff. Am J Nurs 2014 Dec;114(12):28-37; quiz 38,48. [doi: 10.1097/01.NAJ.0000457408.38222.d0] [Medline: 25396755]
49. Holmes D, Teresi JA, Ramirez M, Ellis J, Eimicke J, Kong J, et al. An evaluation of a monitoring system intervention: falls, injuries, and affect in nursing homes. Clin Nurs Res 2007 Nov;16(4):317-335. [doi: 10.1177/1054773807307870] [Medline: 17991911]
50. Lauriks S, Meiland F, Osté JP, Hertogh C, Dröes RM. Effects of assistive home technology on quality of life and falls of people with dementia and job satisfaction of caregivers: results from a pilot randomized controlled trial. Assist Technol 2020 Sep 02;32(5):243-250 [FREE Full text] [doi: 10.1080/10400435.2018.1531952] [Medline: 30592439]
51. Dugstad J, Eide T, Nilsen ER, Eide H. Towards successful digital transformation through co-creation: a longitudinal study of a four-year implementation of digital monitoring technology in residential care for persons with dementia. BMC Health Serv Res 2019 Jun 10;19(1):366 [FREE Full text] [doi: 10.1186/s12913-019-4191-1] [Medline: 31182093]
52. Brodersen S, Lindegaard H. The smart floor: how a public-private partnership co-developed a heterogeneous healthcare technology system. Stud Health Technol Inform 2015;215:191-205. [Medline: 26249197]

XSL•FO
RenderX

53. Bourbonnais A, Rousseau J, Lalonde MH, Meunier J, Lapierre N, Gagnon MP. Conditions and ethical challenges that could influence the implementation of technologies in nursing homes: a qualitative study. Int J Older People Nurs 2019 Dec;14(4):e12266. [doi: 10.1111/opn.12266] [Medline: 31475466]
54. Berridge C, Halpern J, Levy K. Cameras on beds: the ethics of surveillance in nursing home rooms. AJOB Empir Bioeth 2019 Feb 22;10(1):55-62. [doi: 10.1080/23294515.2019.1568320] [Medline: 30794112]
55. Aloulou H, Mokhtari M, Tiberghien T, Biswas J, Phua C, Kenneth Lin JH, et al. Deployment of assistive living technology in a nursing home environment: methods and lessons learned. BMC Med Inform Decis Mak 2013 Apr 08;13(1):42 [FREE Full text] [doi: 10.1186/1472-6947-13-42] [Medline: 23565984]
56. Meng F, Song F, Guo M, Wang F, Feng X, Wang D, et al. Status and influential factors of intelligent healthcare in nursing homes in China. Comput Inform Nurs 2021 May 01;39(5):265-272 [FREE Full text] [doi: 10.1097/CIN.0000000000000685] [Medline: 33950900]
57. Sallinen M, Hentonen O, Kärki A. Technology and active agency of older adults living in service house environment. Disabil Rehabil Assist Technol 2015 Jan 17;10(1):27-31. [doi: 10.3109/17483107.2013.836685] [Medline: 24044649]
58. Fraile JA, Bajo J, Corchado JM, Abraham A. Applying wearable solutions in dependent environments. IEEE Trans Inform Technol Biomed 2010 Nov;14(6):1459-1467. [doi: 10.1109/titb.2010.2053849]
59. Bankole A, Anderson M, Knight A, Oh K, Smith-Jackson T, Hanson MA, et al. Continuous, non-invasive assessment of agitation in dementia using inertial body sensors. In: Proceedings of the 2nd Conference on Wireless Health. 2011 Presented at: WH '11; October 10-13, 2011; San Diego, CA. [doi: 10.1145/2077546.2077548]
60. Soto-Mendoza V, García-Macías JA, Chávez E, Martínez-García AI, Favela J, Serrano-Alvarado P, et al. Design of a predictive scheduling system to improve assisted living services for elders. ACM Trans Intell Syst Technol 2015 Jul 24;6(4):1-31 [FREE Full text] [doi: 10.1145/2736700]
61. Abbate S, Avvenuti M, Light J. Usability study of a wireless monitoring system among Alzheimer's disease elderly population. Int J Telemed Appl 2014;2014:617495-617498 [FREE Full text] [doi: 10.1155/2014/617495] [Medline: 24963289]
62. Anker-Hansen C, Grøndahl VA, Helgesen AK, Fagerli LB, Rummelhoff G, Bååth C, et al. Informal caregivers and assistive technology in Norwegian nursing homes. PLoS One 2022 Oct 4;17(10):e0275474 [FREE Full text] [doi: 10.1371/journal.pone.0275474] [Medline: 36194609]
63. Stark M, Tietz R, Gattinger H, Hantikainen V, Ott S. Effects of a mobility monitoring system on the cost of care in relation to reimbursement at Swiss nursing homes: learnings from a randomized controlled trial. Health Econ Rev 2017 Dec 01;7(1):43 [FREE Full text] [doi: 10.1186/s13561-017-0178-y] [Medline: 29192353]
64. Al‐Oraibi S, Fordham R, Lambert R. Impact and economic assessment of assistive technology in care homes in Norfolk, UK. J Assist Technol 2012;6(3):192-201. [doi: 10.1108/17549451211261317]
65. Engström M, Ljunggren B, Lindqvist R, Carlsson M. Staff perceptions of job satisfaction and life situation before and 6 and 12 months after increased information technology support in dementia care. J Telemed Telecare 2005;11(6):304-309. [doi: 10.1258/1357633054893292] [Medline: 16168167]
66. Offermann-van Heek J, Ziefle M. They don't care about us! Care personnel's perspectives on ambient assisted living technology usage: scenario-based survey study. JMIR Rehabil Assist Technol 2018 Sep 24;5(2):e10424 [FREE Full text] [doi: 10.2196/10424] [Medline: 30249592]
67. Te Boekhorst S, Depla MF, Francke AL, Twisk JW, Zwijsen SA, Hertogh CM. Quality of life of nursing-home residents with dementia subject to surveillance technology versus physical restraints: an explorative study. Int J Geriatr Psychiatry 2013 Apr;28(4):356-363. [doi: 10.1002/gps.3831] [Medline: 22644773]
68. Wigg JM. Liberating the wanderers: using technology to unlock doors for those living with dementia. Sociol Health Illn 2010 Feb 01;32(2):288-303 [FREE Full text] [doi: 10.1111/j.1467-9566.2009.01221.x] [Medline: 20422746]
69. Franco-Trigo L, Fernandez-Llimos F, Martínez-Martínez F, Benrimoj SI, Sabater-Hernández D. Stakeholder analysis in health innovation planning processes: a systematic scoping review. Health Policy 2020 Oct;124(10):1083-1099. [doi: 10.1016/j.healthpol.2020.06.012] [Medline: 32829927]
70. Riahi Y. Project stakeholders: analysis and management processes. Int J Econ Manag Stud 2017;4(3):39-45 [FREE Full text] [doi: 10.14445/23939125/IJEMS-V4I3P108]
71. Brugha R, Varvasovszky Z. Stakeholder analysis: a review. Health Policy Plan 2000 Sep;15(3):239-246. [doi: 10.1093/heapol/15.3.239] [Medline: 11012397]
72. Gregory R, Keeney RL. Creating policy alternatives using stakeholder values. Manag Sci 1994 Aug;40(8):1035-1048. [doi: 10.1287/mnsc.40.8.1035]
73. Braun D, Kramer J. Weerstand is bullshit. Boom Management. 2016 Oct 27. URL: https://boommanagement.nl/artikel/weerstand-is-bullshit/ [accessed 2024-01-14]
74. Lohss S. A new perspective on resistance to change. wibas. 2023 Feb 25. URL: https://www.wibas.com/blog/a-new-perspective-on-resistance-to-change/ [accessed 2024-01-14]
75. Willis R. What drives resistance to change? A leader's perspective. Assentire. 2012. URL: https://www.researchgate.net/publication/329514636_What_Drives_Resistance_to_Change_A_Leader's_Perspective [accessed 2024-01-12]

XSL•FO
RenderX

76. May C, Finch T, Mair F, Ballini L, Dowrick C, Eccles M, et al. Understanding the implementation of complex interventions in health care: the normalization process model. BMC Health Serv Res 2007 Sep 19;7:148 [FREE Full text] [doi: 10.1186/1472-6963-7-148] [Medline: 17880693]
77. Braspenning AM, Cranen K, Snaphaan LJ, Wouters EJ. A multiple stakeholder perspective on the drivers and barriers for the implementation of lifestyle monitoring using infrared sensors to record movements for vulnerable older adults living alone at home: a qualitative study. Int J Environ Res Public Health 2022 Jan 05;19(1):570 [FREE Full text] [doi: 10.3390/ijerph19010570] [Medline: 35010829]
78. May CR, Mair F, Finch T, MacFarlane A, Dowrick C, Treweek S, et al. Development of a theory of implementation and integration: Normalization Process Theory. Implement Sci 2009 May 21;4:29 [FREE Full text] [doi: 10.1186/1748-5908-4-29] [Medline: 19460163]
79. Verkerk M. A philosophy-based 'toolbox' for designing technology: the conceptual power of Dooyeweerdian philosophy. Koers- Bull Christ Scholarsh 2014 Aug 11;79(3). [doi: 10.4102/koers.v79i3.2164]
80. Dentoni D, Bitzer V, Schouten G. Harnessing wicked problems in multi-stakeholder partnerships. J Bus Ethics 2018 Apr 7;150:333-356. [doi: 10.1007/s10551-018-3858-6]
81. Oonk C, Gulikers J, den Brok P, Mulder M. Stimulating boundary crossing learning in a multi-stakeholder learning environment for sustainable development. Int J Sustain High Educ 2022 Jan 13;23(8):21-40. [doi: 10.1108/IJSHE-04-2021-0156]
82. Saidi T, Mork E, Aminoff S, Lobban F, Romm KL. Crossing boundaries in the delivery of healthcare - a qualitative study of an eHealth intervention in relation to boundary object theory. Digit Health 2023 Aug 30;9:20552076231196970 [FREE Full text] [doi: 10.1177/20552076231196970] [Medline: 37662674]
83. Santos Jhunior RO, Johnston NG, Boaventura JM, Barbero ER. Value co-creation within the stakeholder theory: taking stock and moving forward. Revista de Administração Unimep 2021 Jun;19(1) [FREE Full text]
84. Sorrentino M, Fiorilla C, Mercogliano M, Esposito F, Stilo I, Affinito G, et al. Technological interventions in European dementia care: a systematic review of acceptance and attitudes among people living with dementia, caregivers, and healthcare workers. Front Neurol 2024;15:1474336 [FREE Full text] [doi: 10.3389/fneur.2024.1474336] [Medline: 39416661]
85. Selderijk E. Aannames zijn killing voor je team. De Teamontwikkelaars. 2023 Mar 23. URL: https://deteamontwikkelaars.nl/aannames-zijn-killing-voor-je-team/ [accessed 2023-12-18]
86. Kloos N, Drossaert C, Bohlmeijer E, Westerhof GJ. How well do nursing staff assess the wellbeing of nursing home residents? An explorative study of using single-question scales. Ageing Soc 2020 Sep 22;42(4):812-828. [doi: 10.1017/S0144686X20001178]
87. Kunicki ZJ, Madrigal C, Quach LT, Riester MR, Jiang L, Duprey MS, et al. Comparing resident, proxy, and staff respondents for nursing home residents' preferences for everyday living. J Appl Gerontol 2023 Jan 26;42(1):28-36. [doi: 10.1177/07334648221123059] [Medline: 36029016]
88. Roelofs TS, Luijkx KG, Embregts PJ. A person-centered approach to study intimacy and sexuality in residential care facility (RCF) clients with dementia: methodological considerations and a guide to study design. NeuroRehabilitation 2017 Oct 14;41(2):359-373. [doi: 10.3233/NRE-172205] [Medline: 29060944]
89. van Boekel LC, Wouters EJ, Grimberg BM, van der Meer NJ, Luijkx KG. Perspectives of stakeholders on technology use in the care of community-living older adults with dementia: a systematic literature review. Healthcare (Basel) 2019 May 28;7(2):73 [FREE Full text] [doi: 10.3390/healthcare7020073] [Medline: 31141999]
90. Frank L, Jennings LA, Petersen RC, Majid T, Gilmore-Bykovskyi A, Schicker L, et al. Participation of persons with dementia and their caregivers in research. J Am Geriatr Soc 2021 Jul 10;69(7):1784-1792 [FREE Full text] [doi: 10.1111/jgs.17340] [Medline: 34245586]

Abbreviations

ICT: information and communication technology
LTC: long-term care
NPT: normalization process theory
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews

XSL•FO
RenderX

Edited by R Bjarnadottir; submitted 11.07.24; peer-reviewed by J McMurray, M Nakanishi; comments to author 27.11.24; revised version received 09.01.25; accepted 30.01.25; published 20.03.25.

Please cite as:
van Gaans-Riteco D, Stoop A, Wouters E
Values of Stakeholders Involved in Applying Surveillance Technology for People With Dementia in Nursing Homes: Scoping Review
JMIR Aging 2025;8:e64074
URL: https://aging.jmir.org/2025/1/e64074
doi:10.2196/64074
PMID:39899267

©Daniëlle van Gaans-Riteco, Annerieke Stoop, Eveline Wouters. Originally published in JMIR Aging (https://aging.jmir.org), 20.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Barriers to and Facilitators of Implementing Overnight Nursing Teleconsultation in Small, Rural Long-Term Care Facilities: Qualitative Interview Study

Veronique Nabelsi[1], BSc, MScA, PhD; Véronique Plouffe[2], BSc, MSc; Marie Chantal Leclerc[3], BSc, MSc, DBA

[1]Département des sciences administratives, Université du Québec, Québec, Canada

[2]Département des sciences comptables, Université du Québec en Outaouais, Gatineau, Canada

[3]Département des sciences infirmières, Université du Québec en Outaouais, Gatineau, Canada

Corresponding Author:
Veronique Nabelsi, BSc, MScA, PhD
Département des sciences administratives
Université du Québec
101, rue St-Jean-Bosco
Québec, J8X 3X7
Canada
Phone: 1 8195953900 ext 1915
Email: veronique.nabelsi@uqo.ca

Abstract

Background: Teleconsultation has expanded rapidly in recent years, especially during the COVID-19 pandemic, and has become standard practice among physicians. The benefits of teleconsultation, namely, improving access to care, ensuring continuity and quality of care, increasing patient satisfaction, and reducing costs and wait times, are well documented. However, its use in nursing practice, especially in long-term care settings, remains underresearched despite its significant transformative potential, particularly in resource-limited and rural settings, where it could address major challenges such as nursing shortages and access to care.

Objective: This study aimed to identify barriers to and facilitators of implementing overnight nursing teleconsultation in rural residential and long-term care centers in Quebec, Canada (*centres d'hébergement et de soins de longue durée* [CHSLDs]), with ≤50 beds.

Methods: A 6-month pilot project was rolled out sequentially in 3 rural CHSLDs in 2 administrative regions of Quebec between July 2022 and March 2023. A total of 38 semistructured interviews were conducted with managers (n=27, 71%), nursing staff members (n=9, 24%), and resident committee presidents (n=2, 5%) between February 2023 and July 2023.

Results: The study identified several barriers to the implementation of teleconsultation. The main barriers reported included union opposition (managers: 23/27, 85%), network instability (resident committee presidents: 2/2, 100%), limited technology skills (nursing staff members: 7/9, 78%), a perceived increase in workload (nursing staff members: 8/9, 89%; resident committee presidents: 2/2, 100%), and a low volume of teleconsultations (nursing staff members: 8/9, 89%). Despite the barriers, participants also identified key facilitators. These included the care setting (nursing staff members: 9/9, 100%; managers: 21/27, 78%), buy-in from senior management and managers (managers: 27/27, 100%; resident committee presidents: 2/2, 100%), collaboration between the departments (nursing staff members: 9/9, 100%), nursing staff motivation (nursing staff members: 9/9, 100%), and improvements in professional practices (nursing staff members: 8/9, 89%). Finally, the relative benefits of teleconsultation, such as enhanced mutual vision, faster assessment of clinical situations, improved resident care management quality, and greater flexibility and safety, were unanimously recognized (38/38, 100%) as contributing to its acceptability and potential for success.

Conclusions: This study provides an in-depth understanding of the barriers to and facilitators of implementing overnight nursing teleconsultation in small rural CHSLDs. This constitutes a sound basis for developing tailored strategies aimed at overcoming identified barriers and optimizing facilitators. The results also provide practical guidelines for decision makers, highlighting the need to adapt implementation approaches to the unique context of each facility. Furthermore, this study highlights the importance of further research to broaden our knowledge on the dissemination and scale-up of health care innovations. This includes the development of learning health systems capable of responding in an agile and effective way to the needs of rural and vulnerable populations both in Quebec and elsewhere.

XSL•FO
RenderX

(JMIR Aging 2025;8:e71950) doi:10.2196/71950

KEYWORDS

teleconsultation; long-term facilities; nursing; barriers and facilitators; rural; telehealth; qualitative; pilot study; Quebec

Introduction

Background

Teleconsultation has evolved considerably in recent years, especially during the COVID-19 pandemic, when it became standard practice for general practitioners and specialists alike [1-3]. Numerous studies have demonstrated its effectiveness in improving access to care [4-6], ensuring continuity [7,8] and quality of care [9,10], increasing patient satisfaction [11-15], and reducing costs [16-18] and wait times [2,19-22]. Moreover, internet-based clinical support initiatives between novice and expert professionals are being implemented in both urban [23-25] and rural [26,27] settings specifically to address challenges related to the shortage of qualified health care workers.

Despite the increasing adoption of teleconsultation, its use in nursing practice remains largely unexplored [28-31], even though its potential to transform care, including in long-term care centers, is well recognized [32,33]. This potential is even more significant in rural and resource-limited settings, where access to health care services remains a major challenge [34-37].

The lack of data on the use of teleconsultation in nursing practice is concerning as nurses play a key role in the continuum of care, especially in rural settings and long-term care facilities, where they are often patients' first point of contact [38-40]. Although recent studies have focused on the facilitators of and barriers to implementing geriatric teleconsultation in home nursing care [33,41], its impact on nurses' and nursing assistants' workflows [35], and the costs and cost savings associated with its use in residential and long-term care centers [34], these studies remain limited.

To date, studies have not yet provided a comprehensive picture of the effectiveness and impact of teleconsultation in nursing practice. Thus, in-depth exploration of this approach is essential to optimizing its benefits and supporting nursing professionals in the adoption and integration of this technological innovation.

Context

This research gap is of concern in residential and long-term care centers (*centres d'hébergement et de soins de longue durée* [CHSLDs]) in Quebec, Canada, where continuity and quality of care are essential to residents' well-being. Directives from the Ministry of Health and Social Services (Ministère de la Santé et des Services sociaux [MSSS]) and guidelines from the Order of Nurses of Quebec (Ordre des infirmières et infirmiers du Québec) stress the need to ensure the presence of nurses 24 hours a day in these facilities. However, smaller CHSLDs, namely, those with ≤50 beds, face major challenges in meeting this requirement due to nursing staff shortages, especially during night shifts. These frequent periods when no resources are available expose residents to increased risks to their safety and well-being [42,43]. This situation is even more critical in semiremote and remote areas, where difficulties associated with recruiting and retaining nursing staff exacerbate the challenges related to access and quality of care. A better understanding of the opportunities offered by teleconsultation could help alleviate these structural challenges and help build the CHSLDs' capacity in these settings.

As a response to these challenges, the National Directorate of Nursing Care and Services of the MSSS has initiated a pilot project to assess the impact of access to overnight nursing teleconsultation in rural CHSLDs with ≤50 beds. This initiative is based on the implementation of nursing teleconsultation, a promising solution to reinforce continuity of care and reduce regional disparities in access to health care services in Quebec.

Recent studies have shown that teleconsultation can mitigate limitations related to the lack of on-site nursing staff and offer real-time clinical support, thus reducing the risk of adverse events during periods of understaffing [35,44,45].

In line with this approach, this project explored innovative solutions to build resilience in long-term care systems and respond more effectively to the growing challenges associated with nursing shortages.

Objectives

The aim of this study was to identify barriers to and facilitators of implementing overnight nursing teleconsultation in rural Quebec CHSLDs with ≤50 beds. Specifically, this study aimed to gather the views of managers, nursing staff, and resident committee presidents. Exploring these viewpoints will fill a significant gap in the current literature and suggest critical avenues to support the successful integration of teleconsultation in long-term care settings.

Methods

Study Design and Setting

The 6-month pilot project was rolled out in 3 rural CHSLDs located in 2 administrative regions of Quebec. The regions were selected by the MSSS for their alignment with the project's outlined criteria, which include facilities located in semiremote and remote areas, those already experiencing nursing shortages during the night shift, and those reporting issues and risks related to these shortages. In addition, at least 30% of all CHSLD facilities in the territory have a capacity of ≤50 beds. The rollout was conducted sequentially from July 2022 to March 2023 at different sites. Given the innovative nature of the pilot project, an exploratory qualitative study was conducted to identify the barriers to and facilitators of implementing teleconsultation in overnight nursing care.

Data Collection

An interview guide was designed, tested, and validated by the research team. The guide, comprising 12 open-ended questions, aimed to identify the barriers to and facilitators of implementing

XSL•FO
RenderX

teleconsultation. This guide provided a better understanding of the context and experiences surrounding the pilot project's deployment. This study was guided by key factors influencing the implementation of health innovations as outlined in the framework proposed by Chaudoir et al [46]. The interview guide is included in Multimedia Appendix 1.

The framework developed by Chaudoir et al [46] was selected for several key reasons that ensure that it is an appropriate framework to assess the implementation of health innovations. In fact, this model is based on a systematic review of the literature. It also captures the complexity of implementing health innovations by considering various levels of influence, such as organizational structures, health care providers, patients, and the specific characteristics of the innovation.

This holistic approach enables a more comprehensive and nuanced assessment of the factors that can affect an innovation's success. Furthermore, by incorporating levels of analysis that reflect practical realities in the field—such as the nursing staff and residents—the model provides health care professionals with an applicable and relevant framework. It helps identify the barriers and facilitators that are specific to each setting, thus facilitating the design of customized implementation strategies. Finally, the model is flexible and can be tailored to different health innovations and settings. Whether we are examining the implementation of a new technology or a care protocol, the model proposes analysis categories that can be adjusted according to the features of the innovation and environment.

The model's components can be described on five levels: (1) the structural level, which includes factors related to the broader context in which the innovation is implemented, such as health care policies, funding structures, or available resources. These variables directly or indirectly influence an organization's ability to adopt and integrate new practices. (2) the organizational level, which focuses on the specific characteristics of the organizations themselves, such as organizational culture, leadership, support systems, and communication dynamics. It examines how these internal elements facilitate or hinder the implementation of innovations. (3) the health care provider level; at this level, the focus is on the individuals responsible for implementation, such as physicians, nurses, or other health care professionals. The model assesses health care providers' attitudes, knowledge, skills, and beliefs, all of which can influence how an innovation is adopted and applied. (4) the patient level, which analyzes patients' perceptions, attitudes, and behaviors, as well as their level of innovation engagement and buy-in. Patients' psychosocial factors, such as their understanding, beliefs, or preferences, are crucial to the successful implementation of health innovations. (5) the innovation level, which examines the specific characteristics of the innovation itself, such as complexity, compatibility with existing practices, cost, and flexibility. An innovation perceived as easy to use, relevant, and beneficial is more likely to be adopted.

This framework provides a solid foundation to analyze the perspectives of managers, nursing staff, and resident committee presidents, highlighting the key factors influencing the adoption and effectiveness of nursing teleconsultation in small CHSLDs.

Participants

Participant recruitment was conducted using nonprobability sampling [47], through which participants were identified by pilot project managers in each region. This selection method is designed to maximize participants' intrinsic motivation by giving them the opportunity to become involved on their own terms. By promoting this freedom of choice, this study aimed to attract participants who were especially motivated and engaged, thereby improving the quality of the collected data as well as enhancing the relevance and validity of the findings.

Participants were initially contacted through an email that included the interview guide and a detailed consent form. The consent form outlined the context of the study, the project's objectives, the procedures and expected duration of participation, the anticipated benefits, and assurances regarding anonymity and confidentiality. A total of 38 semistructured individual interviews were conducted with managers (n=27, 71%), nursing staff (n=9, 24%), and resident committee presidents (n=2, 5%) between February 2023 and July 2023.

These semistructured interviews were carried out in French and were conducted via videoconference (Zoom; Zoom Video Communications, Inc). The principal investigator (VN) conducted the interviews. Participants were given the option to review the transcripts of each interview, but none of the participants chose to receive the transcripts. No additional recruitment process was necessary as information redundancy indicated data saturation [48].

Data Analysis

On the basis of the framework by Chaudoir et al [46], we conducted a pattern analysis of interview transcripts to identify factors describing facilitators of and barriers to implementing overnight nursing teleconsultation in small CHSLDs. This approach emphasized hierarchical coding, enabling rigorous structuring of the textual data analysis while offering the flexibility required to meet the specific needs of the study.

The analysis prioritized participants' responses, highlighting their descriptions of barriers and facilitators. To do this, we immersed ourselves in the data by reading and rereading transcripts while taking handwritten notes on emerging factors and codes. This iterative process carried out by a single coder (the principal investigator, VN) is a valid method in qualitative analysis of thematic data, enabling researchers to understand how participants gave meaning to their experiences. A handwritten thematic map was created to group data extracts into broad categories of barriers and facilitators, which fostered a thorough review process and helped generate initial codes. These codes were then applied during a second analysis phase using the NVivo software (version 14; QSR International). The initial codes were used to identify overlaps and search for emerging themes.

To ensure a comprehensive and structured analysis, the coder (VN) applied the 5 levels of the framework by Chaudoir et al [46]—structural, organizational, health care provider, patient, and innovation—as primary coding categories. Each identified barrier or facilitator was initially coded under one of these levels. The coder developed subthemes within each level, enabling a

XSL•FO
RenderX

more nuanced examination of the data. This structured approach reinforced the consistency of our analysis, ensuring that all aspects of the implementation process were thoroughly explored while maintaining alignment with established theoretical constructs.

Finally, potential factors were examined and verified across the dataset by rereading the transcripts and checking themes against identified codes. This approach ensured the robustness and consistency of the findings.

Ethical Considerations

Ethics approval was obtained from the Research Ethics Committee of the Outaouais Integrated Health and Social Services Centre before the beginning of the study (reference 2022-353_195) in Quebec, Canada. All participants gave their consent electronically before beginning the interviews. Participation was anonymous and voluntary. Study participants did not receive monetary compensation. All interviews were audio recorded with the participants' permission and then transcribed in compliance with ethical and confidentiality standards. The deidentified recordings were transcribed verbatim by a third-party transcription service bound by a confidentiality agreement. The study's findings will be disseminated through presentations at conferences and publications in peer-reviewed journals using anonymized data. The findings will also be shared through presentations to various MSSS stakeholders and the nursing community.

Results

Overview of Barriers

This study aimed to identify the barriers to and facilitators of implementing overnight nursing teleconsultation in 3 of Quebec's rural CHSLDs with ≤50 beds. The results are presented in accordance with the 5 levels of the framework by Chaudoir et al [46]: structural, organizational, health care provider, patient, and innovation.

Table 1 presents a comprehensive overview of the framework detailing the barriers identified at each level and for each participant group.

XSL•FO
RenderX

Table 1. Level, factors, and number of participants who mentioned each factor.

Level and factor	Managers[a] (n=27), n (%)	Nursing staff members[b] (n=9), n (%)	Resident committee presidents[c] (n=2), n (%)
Structural			
Union opposition	23 (85)	—[d]	—
Network instability	11 (41)	4 (44)	2 (100)
Overburdened managers	16 (59)	—	—
Lack of support from project leaders	—	6 (67)	—
Organizational			
Lack of leadership from the site manager	—	4 (44)	—
Health care provider			
Resistance to change	12 (44)	6 (67)	—
Limited technology skills	12 (44)	7 (78)	—
Insecurity about using technology	12 (44)	4 (44)	—
Increased workload associated with the technology	10 (37)	8 (89)	2 (100)
Patient			
Concerns about quality of care	—	—	1 (50)
Innovation			
Low volume of teleconsultations	14 (52)	8 (89)	—
Complexity of the process compared to a phone call with a remote on-call nurse	16 (59)	—	—
Increased time to initiate care management	—	5 (56)	—
Insecurity about the quality of nursing assessments	—	5 (56)	—
Difficulty using the tablet	—	4 (44)	—

[a]Individuals in positions of authority. They oversee operations, manage resources, and supervise personnel in health care settings. Their role is to ensure the smooth functioning of the facility.

[b]This refers to all individuals involved in providing direct care to patients or residents, including both nurses and nursing assistants.

[c]These individuals lead the residents' committee. The committee represents residents in health care facilities such as long-term care homes. Its mandate is to protect residents' rights. The committee ensures that residents are treated with dignity and that their rights and freedoms are respected. It also serves as a key spokesperson for residents. It brings residents' concerns and needs to the attention of the institution's governing bodies.

[d]Not applicable.

Barriers: Structural-Level Factors

Union Opposition

Union opposition was mentioned by most managers (23/27, 85%) as a major factor before the launch of the pilot project. According to respondents' testimonials, for several months, the union waged a disinformation campaign conveying alarming messages to health care professionals and the community.

The union disseminated messages stating that the "government was planning to replace nurses with tablets," which it claimed would "diminish the quality of care and endanger the safety of residents" (M10). Another manager (M9) specified that "They ran a lot of ads saying nurses were being replaced by tablets." Some managers reported that this campaign had a disruptive effect, creating "a shock wave and a wave of fear in the community" (M22), and the pilot project was perceived "as devaluing this client group, as if they were receiving second-class care" (M25).

The union's intervention was not confined to the public sphere; it also manifested itself directly in the workplace. Managers reported that union representatives visited CHSLDs attempting to dissuade nurses from participating in the pilot project. One manager (M8) explained the following:

> *The union would come directly into the workplace, to frighten employees.*

The union told nursing assistants that their participation was "super dangerous, because they would be going beyond their scope of practice and their professional order would turn against them," added one manager (M9). Despite this initial pressure, none of the interviewed nursing staff members (9/9, 100%) reported any problems with their union after the beginning of the pilot project. Unlike the managers, nursing staff members did not perceive union opposition as a barrier.

Network Instability

Network instability was identified as a barrier by some managers (11/27, 41%) and nursing staff members (4/9, 44%), as well as

XSL•FO
RenderX

by the resident committee presidents (2/2, 100%), especially in rural and remote areas.

Interviewees' testimonials revealed that connectivity was not uniform across the facilities. One manager (M1) explained the following:

> *It didn't necessarily work everywhere in the CHSLD.*

In total, 44% (4/9) of nursing staff members added that there were sometimes 15- to 20-minute delays in logging in. This wasted time, although occasional, can have significant repercussions on the quality of care, as mentioned by one resident committee president (RC36):

> *From time to time, it won't work there...you have to take that into account, because it would be a huge waste of time to start an intervention, then you lose the network, you have to restart another way, by telephone, etc.*

Overburdened Managers

The extra workload associated with the pilot project was a barrier for 59% (16/27) of the managers, especially for project leaders, project coleaders, and site managers. They had to reconcile their usual tasks while ensuring rapid deployment of the project within the context of a nursing shortage. One manager (M4) explained the following:

> *We weren't optimal in our monitoring, which created an obstacle, because, basically, we weren't as present.*

Moreover, daily monitoring of nursing staff practices increased the burden on site managers, who feared that implementation of the project would not be feasible without additional resources. One manager (M9) raised the following question:

> *Is the workload going to be realistic for managers who are already highly solicited?*

Lack of Support From Project Leaders

Lack of support from project leaders was identified as a barrier by 67% (6/9) of nursing staff members. This limited support manifested as reduced availability as project leaders were often overwhelmed by their many responsibilities. One nursing staff member (NS34) indicated the following:

> *We'd have liked to have a little more time, but they're kind of busy with everything.*

This created a feeling of abandonment among nursing staff, with some expressing a lack of guidance. One staff member (NS29) reported that "The project leaders didn't always have the answers to our questions."

Barriers: Organizational-Level Factors (Lack of Leadership From the Site Manager)

In one administrative region, the site manager's lack of leadership was perceived as a barrier by some nursing staff members (4/9, 44%). This lack of leadership manifested itself as a lack of proximity between the site manager and the nursing staff, namely, infrequent travel to meet the team and gaps in communication. This distance hindered the exchange of information and the understanding of the pilot project's issues, leading the site manager to become disinterested, which generated frustration and reduced nursing staff buy-in and motivation at the outset of the pilot project. However, as the project progressed, the situation improved, and they were able to overcome these challenges.

Barriers: Health Care Provider–Level Factors

Resistance to Change

Resistance to change was identified as a barrier by 44% (12/27) of managers and 67% (6/9) of nursing staff members. This resistance took the form of a reluctance to participate in training and simulations, as well as a marked preference for using the telephone, which was perceived as quicker and more effective. For example, one nurse refused to use teleconsultation to assess a resident after a fall, preferring to travel to the CHSLD herself. One manager (M3) indicated the following:

> *The nursing staff found it cumbersome.*

Indeed, nursing staff members reported that their colleagues preferred traditional methods such as on-call nursing or traveling to visit the resident in person. Even when teleconsultation would have been more appropriate, using the telephone remained the preferred method.

Limited Technology Skills

Limited technology skills represented a barrier, as identified by 67% (18/27) of the managers. One manager (M20) pointed out the following:

> *One of the barriers we encountered very, very quickly was that people were not familiar with the technology and then were not able to use it.*

According to one manager (M23), this shortcoming can be explained by "a lack of simulation and comfort as well as by constraints such as nursing shortages, heavy workloads, and COVID-19 outbreaks." One manager (M8) said the following:

> *In addition to learning new tools, our nurses had to continue providing care. At one point, they were saying, "It's just not working".... It was a real challenge to implement this on a daily basis.*

According to 78% (7/9) of nursing staff members, this barrier is especially significant among older nurses, who are often less comfortable with new technology. One nursing staff member (NS33) described this generational challenge:

> *We weren't all born with a keyboard or tablet in our hands...we have a few who are in their fifties. Not all of them were comfortable with it either.*

One nursing staff member (NS29) added the following:

> *...although the younger generation showed an initial interest, this desire was curbed by older nurses' reluctance to embrace the technology, creating a barrier to the successful integration of digital tools into professional practices.*

Furthermore, the low volume of remote activities and insufficient monitoring limited the practice of teleconsultation, resulting in the loss of acquired skills. One nursing staff member (NS33) described it as follows:

XSL•FO
RenderX

> *We had practices, and I fell on the practice that was postponed.... It was never rescheduled. So, the first time, I'd never even practiced.*

Insecurity About Using Technology

Insecurity about using technology was a concern due to the novelty of teleconsultation and was mentioned by 44% (12/27) of managers. In addition, 44% (4/9) of care staff members reported that nursing assistants in particular felt vulnerable when they were alone on-site, fearing that they would not know how to use teleconsultation properly or solve technical problems. As one manager (M4) explained, "It's my nurse who takes charge" when a resident is not doing well. Another manager (M9) confirmed this feeling of insecurity:

> *This practice, in terms of the technology, well, it caused a little insecurity at first, because we didn't know that much about it.*

Similarly, one nursing staff member (NS32) noted that "a fear of computer technology" discouraged some of the more experienced nurses from becoming involved in the pilot project.

Increased Workload Associated With the Technology

The implementation of teleconsultation led to an increased workload for remote nurses, a challenge that was highlighted by 37% (10/27) of managers, 89% (8/9) of nursing staff members, and 100% (2/2) of resident committee presidents. One nursing staff member (NS31) reported the following:

> *We were the ones who had to adapt the most.*

Before every night shift, nurses had to make sure that they had the necessary tools for teleconsultation, which often meant traveling to the CHSLD even on their days off. In addition to their usual tasks, they had to manage interdepartmental reports, fill in specific follow-up forms, and immediately document each teleconsultation. One resident committee president (RC36) explained the following:

> *We also had to foresee working time to connect, use the equipment, get everything working and then provide electronic notes afterward, transfer them, etc.*

One manager (M23) pointed out the following:

> *...the people who are working remotely, the ones who are doing the teleconsultation, are the same people who are there in the evening, during the night...it creates a great deal of anxiety.*

Barriers: Patient-Level Factors (Concerns About Quality of Care)

In the context of teleconsultation, concerns about the quality of care were primarily raised by 50% (1/2) of the resident committee presidents. He feared that teleconsultation would undermine the personal nature of care, concerned that technology would compromise the human contact that is essential to in-person interactions. Initially opposed to the project, fearing the impersonality and disempowerment of nurses, he eventually recognized the benefits of teleconsultation as the pilot progressed. One of the presidents (RC36) expressed the following:

> *My initial opposition to the project was based on what I didn't want: That it would become impersonal, that it would prevent human contact...that it would change the on-call nurse's responsibility, relying on a screen, which is not the same as what you might experience during an in-person visit.*

Barriers: Innovation-Level Factors

Low Volume of Teleconsultations

According to 52% (14/27) of managers, the low volume of teleconsultations was a barrier and could be attributed to the small size of CHSLDs, but the low volume of activity also raised important questions, as one manager pointed out (M25):

> *Was the volume low because there wasn't a need? Was the volume low because practices were already good on both sides, and [the person carrying out an intervention] acted preventively?*

Another manager (M4) added the following:

> *Nursing assistants who said, "Oh no, look, it's 4 o'clock. The nurse is coming in two hours, we'll wait two hours." Did this harm the resident? Well, indirectly, for someone who is in pain, yes it did. But there was no report of an incident or accident that could have or did cause harm to the resident's health, safety, and well-being.*

According to 89% (8/9) of nursing staff members, the low volume of teleconsultations hindered their ability to maintain their skills. One of them (NS26) said the following:

> *It had been a month since we'd had one...I forgot to fill in the smartsheet.*

Moreover, a manager (M10) observed the following:

> *...when comparing data from the previous year to [the data related to] the implementation of teleconsultation, the number of telephone calls received is equal to the number of teleconsultations over the same period.*

Complexity of the Teleconsultation Process Compared to On-Call Nursing

The complexity of the teleconsultation process compared to on-call nursing was perceived as a barrier by 59% (16/27) of managers. Unlike on-call nursing, when the nursing assistant can contact the on-call nurse directly, teleconsultation involves a series of more complex actions, such as waking up the on-call nurse to initiate the consultation and, sometimes, the need to call back the CHSLD. This complexity prolongs the time it takes to obtain a nursing assessment, as explained by one manager (M18):

> *It's too slow...the time to turn on the laptop, to connect safely.*

Another manager (M10) added that the speed of the on-call process influenced perceptions of teleconsultation:

> *It influenced the teleconsultation project.*

XSL•FO
RenderX

In addition, the necessity of having the teleconsultation on-call travel case added another layer of difficulty, especially when staff forgot the equipment, as described by one manager (M12):

> *Ah OK, but now I don't have the equipment, I have to go to the CHSLD to get the equipment.*

Increased Time to Initiate Care Management

According to 56% (5/9) of nursing staff members, the use of teleconsultation led to an increase in the time taken to initiate the residents' care management. Unlike previous practices, the nurse had to assess the resident over a digital platform before intervening, adding an extra step that delayed the response to immediate needs. This delay was exacerbated in remote areas, where unstable internet connections complicated access to teleconsultation, potentially leading to a deterioration of the resident's condition. One nursing staff member (NS26) illustrated this problem by describing a situation in which the requirement to use a tablet for teleconsultation interfered with the prompt management of a resident's pain:

> *A lady was experiencing pain.... I found the computer-based support detrimental to immediate care.... Meanwhile, the lady was in pain. You know, we are managing pain at the same time as we manage the tablet.*

Insecurity About the Quality of Nursing Assessments

There were concerns about the quality of nursing assessments carried out via teleconsultation, including the fear that visual assessment cannot adequately replace a physical examination. In total, 56% (5/9) of nursing staff members shared these concerns. One of them (NS30) explained the following:

> *I had concerns about the physical assessment in the sense that shifting to a visual assessment instead of doing it in real life...that my assessment would not be complete.*

The absence of physical contact with the resident was perceived as a limitation as nonverbal communication plays an important part in a comprehensive assessment. Another nursing staff member (NS33) illustrated this difficulty:

> *Non-verbal and verbal [messages] contradict each other in residents...with the tablet, it wasn't easy because I couldn't look at my resident's face and leg movement at the same time.*

The use of technology such as the digital stethoscope also prompted reservations. One nurse expressed unease:

> *Listening over the phone, it's not like performing the auscultation myself.* [NS31]

Difficulty Using the Tablet

When using a tablet, nursing staff members experienced physical constraints in terms of mobility and effectiveness. These constraints made it difficult to carry out teleconsultations, as one nurse described:

> *The hindrance was caused by the darn arm they set up to hold that tablet.... It would swing around. You know, to be honest, it wasn't the best.* [NS33]

A total of 44% (4/9) of the nurses emphasized the need for a stand to hold the tablet, freeing up the nursing assistant's hands.

Overview of Facilitators

Table 2 presents a comprehensive overview of the framework detailing the facilitators identified at each level and for each participant group.

XSL•FO
RenderX

Table 2. Level, dimension, and number and percentage of participants who mentioned each factor.

Level and dimension	Managers[a] (n=27), n (%)	Nursing staff members[b] (n=9), n (%)	Resident committee presidents[c] (n=2), n (%)
Structural			
Care setting	21 (78)	9 (100)	—[d]
Legitimization of the practice of overnight on-call nursing	18 (67)	—	—
Culture of on-call nursing	—	—	2 (100)
Implementation monitoring	21 (78)	—	—
Organizational			
Buy-in from senior management and managers	27 (100)	—	2 (100)
Support from project leaders	19 (70)	—	—
Support from site managers	19 (70)	7 (78)	—
Team involvement, motivation, and stability	15 (56)	—	—
Collaboration between the nursing department and the Support Program for the Autonomy of Seniors	18 (67)	9 (100)	—
Transfer of knowledge and experience	21 (78)	—	—
Health care provider			
Nursing staff buy-in	13 (48)	—	—
Nursing staff motivation	—	9 (100)	—
Development of the nursing staff's skills	13 (48)	—	—
Ability to adapt and use technology	—	7 (78)	—
Openness to change	—	—	2 (100)
Patient			
Buy-in from residents, families, and resident committees	12 (44)	—	—
Communication	—	6 (67)	—
Innovation			
Relative benefits	27 (100)	9 (100)	2 (100)
Development of nursing staff's roles	19 (70)	—	—
Improved professional practices	—	8 (89)	—

[a]Individuals in positions of authority. They oversee operations, manage resources, and supervise personnel in health care settings. Their role is to ensure the smooth functioning of the facility.

[b]This refers to all individuals involved in providing direct care to patients or residents, including both nurses and nursing assistants.

[c]These individuals lead the residents' committee. The committee represents residents in health care facilities such as long-term care homes. Its mandate is to protect residents' rights. The committee ensures that residents are treated with dignity and that their rights and freedoms are respected. It also serves as a key spokesperson for residents. It brings residents' concerns and needs to the attention of the institution's governing bodies.

[d]Not applicable.

Facilitators: Structural-Level Factors

Care Setting

Testimonials from 78% (21/27) of the managers and all nursing staff members (9/9, 100%) highlighted the care setting's decisive role in the success of the pilot project. Faced with a nursing staff shortage, this initiative was viewed as a promising solution to optimize practices and the management of available resources while maintaining the quality and safety of resident care. One manager (M1) explained the following:

> *Our objective is to ensure that every nursing staff member is in the right place, playing their role to the full and that we are using our resources wisely.*

Legitimization of the Practice of Overnight On-Call Nursing

Legitimization of the practice of overnight on-call nursing factored positively in the project's success. According to 67% (18/27) of managers, the fact that this practice was framed within a specific, temporary context reassured stakeholders such as the Order of Nursing Assistants of Quebec (Ordre des infirmières

XSL•FO
RenderX

et infirmiers auxiliaires du Québec), the Quebec ombudsman, and user and resident committees. One manager (M1) explained the following:

> *This legitimization enabled nursing staff to feel they had greater authorization to use teleconsultation, mitigating fears related to professional compliance.*

Culture of On-Call Nursing

The existing culture of on-call nursing was a significant facilitator. All resident committee presidents (2/2, 100%) stated that this culture, which was already well established and accepted in CHSLDs, facilitated the implementation of teleconsultation. Considered "an innovative and adaptive solution, on-call nursing was seen as essential to maintaining optimal quality of care," as indicated by one of the presidents (RC36). He added the following:

> *...the pre-existing culture facilitated the transition to teleconsultation by normalizing the idea of a remote nurse and positioning it as a safe and effective approach.*

Implementation Monitoring

Close monitoring by the nursing department (ND) and the Support Program for the Autonomy of Seniors (SAPA) within the health care system in Quebec was a key facilitator according to 78% (21/27) of managers. Monitoring took place at three levels: (1) strategic level—committees and regular meetings with the MSSS promoted fluid communication on project advancement; (2) operational level—project leaders organized regular meetings with site managers, enabling actions to be adjusted quickly and providing immediate feedback (as one manager [M16] reported, "monitoring by the Ministry...was highly beneficial"); and (3) day-to-day level—ongoing monitoring of teleconsultation practice was implemented, including a review of ministry forms, hospital transfers, and incident reports, as well as audits in CHSLDs to ensure that nursing staff had the support they needed.

Facilitators: Organizational-Level Factors

Buy-In From Senior Management and Managers

Senior management buy-in was viewed as a facilitator by all managers (27/27, 100%), namely owing to the support of the chief executive officer of the Integrated Health and Social Services Centre of Abitibi-Témiscamingue and the Integrated University Health and Social Services Centre of Mauricie and Centre-du-Québec, as reported by one manager (M8):

> *He took it on, then he defended it.*

When senior management prioritizes a project, it motivates other managers to engage, facilitating rollout and the resolution of challenges such as acquiring equipment—"We received our equipment very quickly because it was a priority," according to one manager (M1). The resident committee presidents (2/2, 100%) also confirmed that this support was important to the project's success. One of the resident committee presidents (RC35) explained the following:

> *When senior management is supportive of the project, it sends a strong signal to the members of the organization about the strategic importance of teleconsultation. This approval from leadership can positively influence the levels of acceptance and engagement within the team.*

In addition, the managers' buy-in was unanimously recognized as a facilitator by participating managers (27/27, 100%). Their engagement made it possible to effectively navigate MSSS requirements and ensure the buy-in of the nursing staff members who were consulted, underlining the importance of creating an environment that is conducive to the implementation of teleconsultation.

Support From Project Leaders

Support from project leaders was a determining factor according to 70% (19/27) of managers. Project leaders played a key role in motivating nursing staff by clarifying objectives, allaying concerns, and fostering champions within the teams, creating a conducive environment for the adoption of teleconsultation. As one manager (M5) explained, "Having a dedicated person to answer questions and solve problems" was essential.

Support From Site Managers

According to 70% (19/27) of managers, site managers also fostered the implementation of teleconsultation. One manager (M12) reported the following:

> *Their knowledge of the environment and their existing bonds of trust played a key role in the human management of change and in nursing staff mobilization.*

The nursing staff (7/9, 78%) also appreciated this support, underlining the site managers' guidance and availability. One nursing staff member (NS32) explained the following:

> *I felt supported throughout the project. If I had any questions, I knew where to turn. I had a lot of support from my manager.*

Involvement, Motivation, and Stability of Nursing Staff

According to 56% (15/27) of managers, the involvement, motivation, and stability of nursing staff were key facilitators. Team cohesion facilitated flexibility and mutual support, as stated by one manager (M16):

> *Nursing staff demonstrated solidarity by swapping shifts during snowstorms to ensure staff availability.*

This solidarity enabled staff to respond effectively to residents' needs and maintain reasonable response times.

Collaboration Between the ND and SAPA

According to 67% (18/27) of managers, collaboration between the ND and SAPA was a key facilitator. A clear division of roles enabled the ND to manage external relationships with the MSSS and other agencies, whereas the SAPA dealt directly with the implementation of teleconsultation, ensuring effective coordination. One manager (M16) reported the following:

> *Having a single point of entry was helpful.*

This synergy promoted the cocreation of solutions to the project's challenges, namely in terms of training and monitoring, reinforcing the effectiveness and success of the initiative.

XSL•FO
RenderX

The training plan, including coaching, simulations, and tool adaptations, was tailored to meet regional needs and reinforce the safety of nursing practices. One manager (M10) underlined the following:

> *Training was customized...to reinforce safety.*

Practical guides and equipment such as headsets supported the practice of teleconsultation, and simulations boosted the nursing staff's confidence. Nursing staff members (9/9, 100%) unanimously appreciated the training, deeming it essential to the adoption of the technology and success of the pilot project.

Transfer of Knowledge and Experience

The transfer of knowledge and experience was a significant factor for most managers (21/27, 78%), facilitating collaboration and adaptation between regions and within participating CHSLDs.

Between regions, managers worked closely, sharing their experiences and adjusting approaches according to the specific needs of each region. Although support tools were not cocreated systematically, these exchanges enabled participants to adjust based on local context. Regarding collaboration and adaptation within CHSLDs, in one region, the level of sharing between 2 CHSLDs was particularly striking. Nursing staff from the first site shared their experiences with that of the second, fostering buy-in to the pilot project. For example, a manager's guide created from lessons learned was passed on to the other site, facilitating the implementation of teleconsultation in similar settings. One manager (M16) described it as follows:

> *The manager took notes on everything she had implemented...and brought it back to the ND....So, it ranged from the criteria we had to meet, to making sure we met them, to the tasks we had to carry out...all in one guide.*

Facilitators: Health Care Provider–Level Factors

Nursing Staff Buy-In

For 48% (13/27) of managers, nursing staff buy-in was a facilitator. Nursing staff members were not only favorable to the idea, they were also motivated to actively participate in the pilot project, demonstrating a willingness to move forward with teleconsultation. One manager (M5) noted the following:

> *The nursing staff were very open and aligned with the project. They wanted to move forward with the change.*

Despite initial stress, the nursing staff adapted quickly. As one manager (M24) pointed out, "Once the adaptation period was over...there was no more stress. It went well," underlining their ability to overcome resistance and make a successful transition to teleconsultation.

Nursing Staff Motivation

All nursing staff members (9/9, 100%) considered that the motivation of health care staff was a facilitator. The main sources of motivation included commitment to the team, interest in technological tools, and the desire to help resolve the nursing shortage. Nursing staff members appreciated the creation of overnight nursing assistant positions with remote support, helping address the shortage and improve the quality of care. As one staff member put it, "We're not as effective after 1 p.m" (NS28), highlighting the challenges of working long hours. Being motivated to use technological tools such as teleconsultation reflects a desire to explore innovative solutions to improve working conditions and better meet residents' needs.

Development of the Nursing Staff's Skills

Development of the nursing staff's skills was a key factor, as highlighted by 48% (13/27) of managers. Training tailored to the staff's needs fostered their preparedness and engagement, making them champions of the pilot project. Younger members showed greater mastery of computer and technological skills. In addition, the training improved nursing assistants' level of autonomy and clinical judgment, contributing to the development of clinical leadership and team management skills. One manager (M7) described it as follows:

> *When she's on her own, she also takes on a bit of a coordination role...something she didn't usually do because it was part of the nurse's role.*

Ability to Adapt and Use Technology

For 78% (7/9) of nursing staff members, the ability to use technology was directly related to computer literacy and adaptability. Despite a low level of initial fluency, some managed to overcome their difficulties. One staff member (NS33) put it as follows:

> *I'm not very tech-savvy. I managed. It went well.*

Thus, adaptability was a key factor, with nursing staff members demonstrating an ability to adjust to teleconsultation, including those who were not as comfortable with computers.

Openness to Change

Nursing staff's openness to change was essential according to both resident committee presidents (2/2, 100%). This positive attitude was fostered by the fact that the project was recognized as an opportunity for exploration. One of the resident committee presidents (RC36) said the following:

> *Given that it was a pilot project, everyone agreed to give it a try.*

RC35 added the following:

> *...the transfer of experience from nurses performing on-call nursing to teleconsultation is also a concrete example of this openness to change. The fact that the staff had prior experience in similar practices made it easier to adapt to new methods.*

Facilitators: Resident-Level Factors

Buy-In From Residents, Families, and Resident Committees

Buy-in from residents, families, and resident committees greatly facilitated the implementation of teleconsultation according to 44% (12/27) of managers. This support was reinforced by the creation of a relationship of trust through transparent

XSL•FO
RenderX

communication and regular meetings with the site manager. One manager (M17) pointed out the following:

> *The site manager kept us informed on a regular basis, establishing a climate of trust.*

Communication

Communication was a facilitator, especially in interactions with residents and their families. In total, 67% (6/9) of nursing staff members noted that proactive communication with residents and their families, namely, explaining the project, answering questions, and obtaining informed consent, facilitated resident buy-in. One nursing staff member (NS30) said the following:

> *Transparency, especially about concerns such as data leaks, fostered a positive reception to teleconsultation.*

Families generally welcomed the technology, recognizing the additional benefit to care delivery. Nursing staff members observed that there was no negative impact on residents, which could be attributed to the effective communication that reassured families about data confidentiality and security.

Facilitators: Innovation-Level Factors

Relative Benefits

Analysis of the relative benefits of teleconsultation showed that this modality represented a significant added value for the dyad composed of the remote nurse and the CHSLD nursing assistant according to all managers (27/27, 100%), nursing staff members (9/9, 100%), and resident committee presidents (2/2, 100%). Benefits included, first, improved mutual vision. All nursing staff members (9/9, 100%) agreed that teleconsultation enabled direct observation of the resident and of the nursing assistant's nonverbal cues, fostering a greater understanding of the situation than was possible through on-call nursing. One staff member (NS31) pointed out the following:

> *The way [the nursing assistants] report it to us over the phone and the way we see it through our own assessment, are two different things. Sometimes it's minimized, and sometimes it's exaggerated.*

This visual component enhanced the remote nurse's ability to carry out a more accurate and thorough assessment. One of the resident committee presidents (RC35) stated the following:

> *...the use of video in teleconsultation provides a clear advantage in terms of assessment quality compared to voice-only interactions. This ability to visualize the patient can lead to more informed decisions regarding necessary interventions.*

The second benefit was faster assessment. Teleconsultation reduced the wait time for nursing assessments, enabling more effective interventions according to all nursing staff members (9/9, 100%). One staff member (NS32) noted the following:

> *...an intervention that might have taken an hour could take only 15 minutes with teleconsultation.*

Moreover, the direct electronic transmission of nursing notes to the CHSLD promoted continuity of care. One of the resident committee presidents (RC36) noted the following:

> *...that despite the implementation of teleconsultation, it is not intended to completely replace in-person visits, but rather to complement care, thus offering flexibility and adaptability in the delivery of healthcare services.*

The third benefit was improved quality of care management. The visual component of teleconsultation helped identify signs or symptoms that were not described verbally, leading to more informed decisions according to all nursing staff members (9/9, 100%). In addition, one manager (M25) reported that teleconsultation "provides visual support that the telephone does not, enabling [the nurse] to validate the on-site nursing assistant's hypothesis or to support her in her contribution to the assessment."

The fourth benefit was flexibility and safety. According to the resident committee presidents (2/2, 100%), teleconsultation did not replace in-person visits but was complementary to them, providing flexibility in the delivery of health care services. One of the presidents (RC36) mentioned the following:

> *Nurses are not prevented from physically traveling if necessary.*

This innovation also contributed to resident safety according to the other president (RC35), noting that "residents were safe with this technological innovation, reinforcing the idea that teleconsultation does not entail any compromise in terms of patient safety."

All managers (27/27, 100%) reported that the introduction of teleconsultation was perceived as a major step forward, enhancing the quality and safety of care delivery. According to them, teleconsultation provided greater safety by enabling nurses to exercise their clinical judgment under improved conditions. One manager (M4) stated the following:

> *Teleconsultation gives the manager a sense of security because they know that the healthcare professional is in a better position to exercise their clinical judgment.*

Development of Nursing Staff's Roles

The pilot project contributed significantly to the development of nursing staff's roles according to 70% (19/27) of managers. Although overnight on-call nursing was already integrated into the organizational culture, the pilot project improved the organization and quality of care by strengthening nursing staff's roles.

Key improvements included enhancement of the remote nurse and nursing assistant dyad as the nursing assistant's practice was strengthened, including more effective nursing assessments and better care planning during the night shift, and expanded scope of practice for nursing assistants as changes to nursing care procedures enabled nursing assistants to fully exercise their skills. One manager (M1) pointed out the following:

> *Nursing assistants have activities...that we've allowed them to carry out.*

XSL•FO
RenderX

Another manager (M16) added that the pilot project enabled them to "apply their entire scope of practice, enhancing the value and recognition of their role."

Improved Professional Practices

The implementation of teleconsultation significantly improved professional nursing practices according to most nursing staff members (8/9, 89%). The main observed benefits were, first, the development of nursing assistants' skills. Teleconsultation enabled nursing assistants to develop their scope of practice, acquiring new skills and playing a more active role in assessing and delivering care. Structuring communication when transmitting information also helped strengthen their communication skills. The second main observed benefit was proactive information updates. Teleconsultation facilitated the updating of therapeutic nursing plans and medication administration records, enhancing nursing assistants' autonomy. One nursing staff member explained that teleconsultation eliminated the need to constantly contact the nurse for simple decisions such as administering medication, ensuring greater autonomy for nursing assistants. This staff member (NS29) underlined that "this project has really helped to make us more autonomous...we can manage almost everything on our own."

Discussion

Principal Findings

Overview

The aim of this study was to gain a better understanding of the experiences of managers, nursing staff, and resident committee presidents involved in the pilot project to identify the factors that may promote the implementation of teleconsultation. Our multilevel analysis revealed the presence of facilitators and barriers.

Structural-Level Factors

Among the identified structural barriers, union opposition, which was reported by most managers (23/27, 85%), initially represented a significant obstacle, but it subsided once the project was in place. The union recognized that teleconsultation complied with professional standards while at the same time complementing the work done by nurses in CHSLDs without seeking to replace it. This helped dispel any initial fears. In addition, managers' leadership played a decisive role by supporting the nursing staff, responding to their concerns, and facilitating a smooth and harmonious transition.

However, network instability in rural areas was a barrier for some managers (11/27, 41%) and nursing staff members (4/9, 44%). Connectivity issues coupled with the lack of overnight technical support led to delays and interruptions, compromising the effectiveness and reliability of teleconsultation. These difficulties highlighted the limitations of technological infrastructures in these environments and created situations in which interventions were delayed, compromising continuity of care. This problem is corroborated by studies that underline the importance of stable network connectivity to ensure the smooth operation of teleconsultation services. Indeed, network quality is a determining factor in ensuring the fluidity of exchanges and the effectiveness of remote consultations, as highlighted by several research studies [28,49-52]. In these settings, ensuring reliable network coverage and efficient technical support is imperative to avoid interruptions that could adversely affect the quality of care.

Overburdened managers, who had to juggle their usual responsibilities with the demands of the pilot project, led to suboptimal management according to 59% (16/27) of managers. This double workload, compounded by a nursing shortage, made it difficult to adequately monitor the project and led to management being less present and reactive. This finding is in line with observations found in the literature, which stress that work overload is a major challenge in project management, especially in settings with limited resources [53]. Thus, managers' inability to respond optimally to project requirements due to the pressure of their day-to-day responsibilities contributed to the suboptimal implementation of the initiative, underlining the need for greater support for managers during the implementation of complex projects.

Lack of support from project leaders was a barrier for 67% (6/9) of nursing staff members. Insufficient training and monitoring limited the nursing staff's ability to use teleconsultation effectively. This situation reflects the importance of organizational support, which studies have shown to be a key factor in the success of training and implementation programs [54]. Indeed, constant support and adequate coaching help build the staff's skills and ensure the successful adoption of new technology, such as teleconsultation, in care settings.

However, several facilitators contributed to the project's success. The context of the health care system, marked by a nursing shortage that was exacerbated by the COVID-19 pandemic, acted as a catalyst for the implementation of teleconsultation according to several managers (21/27, 78%) and all nursing staff members (9/9, 100%). Perceived as an effective solution to the shortage, teleconsultation enabled better management of human resources, which reduced excessively long shifts and improved the nursing staff's working conditions. This approach is supported by studies demonstrating that technology can optimize the use of human resources during a shortage [55,56].

Some managers (18/27, 67%) highlighted that legitimization of the practice of overnight on-call nursing played a key role in nursing staff's acceptance and adoption of teleconsultation. According to the literature, framing new clinical practices within a defined, transparent, and temporary framework is essential to reassuring stakeholders such as professional orders, user committees, and regulatory bodies [35,57,58]. Such a framework helps build an environment of trust, reducing reluctance to adopt innovative practices and ensuring compliance with professional standards.

The preexisting culture of on-call nursing facilitated the acceptance of the pilot project by making the transition to teleconsultation smoother according to the resident committee presidents (2/2, 100%). Familiarity with remote practices fostered acceptance of the new technology. This observation is supported by the literature, which indicates that preexisting practices and organizational cultures play a decisive role in the acceptance of health care innovations [35,59].

XSL•FO
RenderX

Finally, close monitoring at several levels was a key factor in the successful implementation of teleconsultation for a large proportion of managers (21/27, 78%). By integrating strategic, operational, and daily monitoring, the project benefited from rapid adjustments, ensuring compliance with project objectives and requirements. This systematic monitoring not only facilitated proactive management of challenges but also enhanced responsiveness to emerging issues, helping maintain coherent and fluid processes. According to the literature, structured and sustained monitoring is essential to optimizing the implementation of new practices, especially in a context of technological transformation, helping overcome barriers and ensure project sustainability [60,61].

Organizational-Level Factors

According to the nursing staff in 1 of the 2 administrative regions (4/9, 44%), one of the barriers at the organizational level was the site managers' lack of leadership. This lack of leadership, characterized by minimal presence and limited communication with nursing staff, was viewed as a barrier to the implementation of teleconsultation. The perceived distance between the site managers and care teams created a climate of frustration and disengagement. This situation is corroborated by the literature, which underlines that weak leadership can hinder the implementation of change initiatives by generating feelings of frustration and disengagement among health care teams [62-64].

As part of the implementation of teleconsultation, several organizational factors facilitated its rollout, highlighting the importance of a coordinated approach and sustained engagement at all levels of the organization. Buy-in and active support from senior management and managers were crucial to the project's success according to all managers (27/27, 100%) and resident committee presidents (2/2, 100%). The literature supports this observation, stating that the engagement of organizational leaders is a key factor in the success of change initiatives in health care facilities [65,66]. This engagement helped mobilize the necessary resources and promote a culture conducive to innovation.

The role of project leaders and site managers was considered essential by 70% (19/27) of managers. Their availability and expertise not only enabled the effective resolution of operational issues but also played a key role in maintaining the nursing staff's level of motivation.

Specifically, nursing staff members (7/9, 78%) stated that regular meetings and personalized support helped clarify objectives, allayed concerns, and ensured constant monitoring, which helped overcome challenges and optimize implementation processes. These results are in line with the literature, which emphasizes the importance of management practices and managers' commitment to the success of digital transformation initiatives [67,68].

According to some managers (15/27, 56%), the involvement, motivation, and stability of nursing staff members were also facilitators. The staff's ability to adapt and maintain a high level of service despite challenges was facilitated by increased motivation and cohesiveness, which is supported by the work by Nabelsi et al [35]. Their research shows that staff motivation and stability are essential elements to ensuring efficient processes in health care. The solidarity and cooperation observed within the teams helped maintain high levels of service even under difficult conditions.

Another facilitator was the collaboration between the ND and the SAPA according to all nursing staff members (9/9, 100%) and several managers (18/27, 67%). This collaboration clarified the division of roles, avoided duplication, and ensured smooth project management. A clear division of responsibilities and coordination between departments are essential [35]. This model of cross-directorate collaboration helped maximize efficiency and avoid overlaps that might slow down the implementation process.

Finally, most managers (21/27, 78%) said that the transfer of knowledge and experience between regions and within CHSLDs played a significant role in localizing the project. This approach, which centered on collaboration and the sharing of best practices, enhanced the effectiveness of the pilot project [35]. The study by Nabelsi et al [35] shows that knowledge sharing between teams and sites not only fosters the adoption of technology, it also enables better management of the challenges encountered in the field by adapting to local needs and constraints.

Health Care Provider–Level Factors

Resistance to change was a barrier to the implementation of teleconsultation, especially for certain nursing staff members (6/9, 67%). Their marked preference for the telephone, perceived as faster and more effective, highlighted their resistance to the adoption of teleconsultation. The literature on organizational change in health care indicates that this resistance may be fueled by the perception of being overloaded and by deeply ingrained habits, making it difficult to accept new technology [28,69,70].

According to nursing staff members (7/9, 78%), a lack of skill in using new technology was also a limiting factor, especially among older nursing staff members, who were not as comfortable with technological tools. This shortcoming is a recognized factor in the failure to implement technology in health care [44,71,72]. Work overload, exacerbated by crises such as the COVID-19 pandemic, added a further dimension to this challenge, creating a cognitive overload that complicated the effective integration of new technology.

Insecurity about using technology was another barrier for a small proportion of managers (12/27, 44%) and nursing staff members (4/9, 44%). Nursing assistants in particular felt vulnerable due to their lack of familiarity with teleconsultation and fear of dealing with technical problems. The literature on technology acceptance underlines that insecurity and fear of the unknown can reduce user motivation and performance [73,74].

The increased workload associated with technology was also a barrier. The implementation of teleconsultation led to increased management of reports, specific forms, and detailed documentation, an issue recognized by most nursing staff members (8/9, 89%) as well as all resident committee presidents (2/2, 100%). Studies show that increased workloads can cause

XSL•FO
RenderX

stress and reduce job satisfaction, negatively influencing the implementation of new practices [75,76].

Among the facilitators, nursing staff buy-in was important, although it was mentioned by a lower proportion of managers (13/27, 48%). The nursing staff's willingness to participate in the pilot project and their positive attitude toward teleconsultation were key to its success. The literature on organizational change in health care suggests that stakeholder buy-in is important to the success of change initiatives [77,78].

Nursing staff members' motivation also played a significant role in the pilot project's success, a factor unanimously recognized by all staff members (9/9, 100%). Commitment to the team, interest in technological tools, and the desire to help solve the nursing shortage were all motivating factors. Research shows that motivation is a facilitator of the acceptance and successful use of health technology [79,80].

Development of the nursing staff's skills was a facilitator, although it was mentioned by a lower proportion of managers (13/27, 48%). Adequate training strengthened nurses' readiness and commitment, enabling them to become project champions. The literature indicates that skill development is key to the successful adoption of new technology [66,81,82]. Customized training and support tools such as practical guides and simulations helped build the nursing staff's skills and confidence [35].

The ability to use the technology was also crucial to the success of the pilot project, as underlined by nursing staff members (7/9, 78%). Rapid adaptation to new technology is indeed a facilitator, as confirmed in research on technology acceptance [74,80].

Finally, openness to change facilitated the implementation of the project, a factor that was unanimously recognized by resident committee presidents (2/2, 100%). This openness to change led to smoother adoption of teleconsultation, which is supported by studies demonstrating that acceptance of change is essential to the success of transformation initiatives [83].

Resident-Level Factors

The only barrier was fear concerning the quality of care provided via teleconsultation, which was mentioned by 50% (1/2) of the resident committee presidents. One of the presidents expressed concerns about the reduction in human contact, which could lead to perceived depersonalization of care. However, these concerns were dispelled as the pilot progressed, and he eventually recognized the benefits of teleconsultation. This fear that technology would disempower nurses and create a sense of impersonality represented a barrier to the acceptance of teleconsultation. Studies indicate that concerns about quality of care and depersonalization can negatively influence acceptance of telehealth technology by residents and their families [50,72,84,85].

According to some managers (12/27, 44%), buy-in from residents, families, and resident committees was a facilitator of the implementation of teleconsultation. The managers observed that transparent communication and regular meetings with these stakeholders helped establish a climate of trust. This approach is in line with the literature, which stresses the importance of trust and effective communication when fostering acceptance of health technology by patients and their families [86-88].

For most nursing staff members (6/9, 67%), effective communication itself was a facilitator. Research shows that managing expectations and clarifying the benefits of new technology are important to their acceptance by residents and their families. The ability to clearly explain the benefits of teleconsultation and address residents' concerns contributed greatly to their buy-in [35].

Innovation-Level Factors

The low volume of teleconsultations was recognized as a barrier by more than half (14/27, 52%) of the managers. Several possible explanations were put forward. Some managers suggested that this low volume may reflect a lack of real need, the preexisting effectiveness of preventive practices, or some nursing staff members' reluctance to use teleconsultation. However, when comparing data from the previous year with the data related to implementation of teleconsultation, findings reveal that the number of telephone calls received was equal to the number of teleconsultations over the same period. In addition, the low volume of consultations impacted the nursing staff's ability to maintain their skills. Most nursing staff members (8/9, 89%) reported a decrease in their level of comfort with the technological tools due to sporadic use.

The complexity of the teleconsultation process was also a barrier for more than half (16/27, 59%) of the managers. Compared to traditional on-call nursing, teleconsultation involves more complex technological processes. Connolly et al [89] underline that this complexity can reduce the effectiveness of interventions and increase staff frustration, hindering the adoption and effectiveness of technology [51,90].

Another barrier was the increased time to initiate care management according to over half (5/9, 56%) of nursing staff members. Using digital platforms for assessments can increase response times, a problem that is exacerbated by technological limitations and connectivity issues. The research by Pilosof et al [91] shows that these delays can adversely impact the quality of care by affecting the responsiveness of interventions.

Concerns about the quality of assessments conducted via teleconsultation were noteworthy. Just over half (5/9, 56%) of nursing staff members expressed concern about the ability of visual assessment to effectively replace a physical assessment, highlighting the potential risk of compromising quality of care. This fear is corroborated by studies revealing that telemedicine can sometimes alter the quality of clinical assessments if not properly integrated into care practices [85-87,92-95].

Finally, the difficulty of using a tablet for teleconsultations represented a barrier for a small proportion of nursing staff members (4/9, 44%). Srinivasan et al [85] highlight that ergonomic issues and difficulties in handling technological equipment can reduce the effectiveness of interventions and user satisfaction, complicating the integration of teleconsultation [86,87,89].

Despite these challenges, several innovation-level factors facilitated the implementation of teleconsultation. The relative

XSL•FO
RenderX

benefits of this technology were viewed positively by all managers (27/27, 100%), nursing staff members (9/9, 100%), and resident committee presidents (2/2, 100%). They appreciated the improved mutual vision and faster assessment. Teleconsultation solutions offer significant benefits in terms of visual communication and speed of intervention [35]. This perception of the benefits fostered acceptability and support for the project.

The development of nursing staff's roles was another facilitator according to several managers (19/27, 70%). The project strengthened professional practices and broadened the nursing assistants' skills. The work by Nabelsi et al [35] indicates that teleconsultation technology can support the expansion of professional roles and improve quality of care.

Moreover, the implementation of teleconsultation led to improved professional practices for most nursing staff members (8/9, 89%). Research has shown that technology integration improves professional skills and care management, highlighting the potential of innovations to positively transform practices in care settings [35,91,96].

Limitations and Future Research

This study has a number of limitations that must be taken into account when interpreting the results. First, this study was conducted in only 2 Quebec regions, including a total of 3 CHSLDs. This limited scope may restrict the generalizability of the results to other geographical settings or types of facilities. In addition, the small number of sites included in the study may not enable researchers to capture the diversity of practices and challenges encountered in other regions or in facilities of different sizes. Second, this study focused exclusively on smaller facilities with ≤50 beds. While this is in line with the study's objective of targeting small CHSLDs, the results may not be directly applicable to larger facilities, which may have different organizational structures and needs. Finally, this study's evaluation period was short, making it impossible to observe the long-term impact of nursing teleconsultation, especially in terms of the continuous improvement of nursing practices, the sustainability of interventions, and the changes in stakeholder perceptions.

Furthermore, this study used a nonprobabilistic sampling method, which may have resulted in the inclusion of participants who were more inclined to view teleconsultation favorably or who had a particular interest in the topic. To mitigate this potential selection bias, the team researcher actively sought a diversity of perspectives during data collection, encouraging participants to share both supportive and critical viewpoints regarding the implementation of teleconsultation. Moreover, a rigorous qualitative analysis was conducted, with particular attention given to dissenting opinions and negative experiences, ensuring a comprehensive representation of the facilitators and barriers encountered. Despite these efforts, the inherent limitations of qualitative research, including the subjectivity of self-reported experiences, necessitate further investigation through complementary methodologies. A quantitative study conducted through a survey would strengthen the robustness of our findings and allow for a more generalizable assessment of the impact of teleconsultation

To broaden our understanding of the implementation of nursing teleconsultation in long-term care, it would be relevant to conduct studies in a larger number of CHSLDs, including facilities of various sizes located in different regions, to assess the transferability of this study's findings and their effectiveness on a larger scale. Longitudinal research would also be needed to assess the long-term effects of nursing teleconsultation on quality of care, resident satisfaction, and human resource management in CHSLDs, as well as to identify any adjustments needed to ensure the sustainability of these practices.

It would also be useful to study the impact of teleconsultation on nurses' well-being and workload, namely by examining how this technology can be optimized to effectively support their role without increasing their stress level or mental load. Finally, economic studies could help quantify the costs associated with implementing nursing teleconsultation and compare them with potential savings in terms of decreased hospitalizations, adverse events, and improved quality of care.

Conclusions

This study provides the first in-depth analysis of barriers and facilitators related to the implementation of overnight nursing teleconsultation in small long-term care facilities. The findings provide a better understanding of these barriers, which can be used to develop strategies to overcome them during implementation. These findings are also particularly relevant to decision makers who are responsible for designing health initiatives as their choices influence the implementation and scaling-up process.

Broadly, the results provide a comprehensive overview of the factors influencing the successful implementation of teleconsultation in long-term care. This can help identify key factors to consider when scaling up teleconsultation in CHSLDs. The framework developed by Chaudoir et al [46] highlights the concept of adaptability, emphasizing the importance of adjusting the deployment of an innovation to suit the specific context. When scaling up teleconsultation, it is important to consider the specific characteristics of each CHSLD and region and tailor the implementation of teleconsultation accordingly.

While resistance to change is often considered a major barrier to implementing new health care technologies, our findings challenge this assumption. In the rural CHSLDs studied, the preexisting on-call nursing culture not only facilitated the adoption of teleconsultation but also eased its integration into clinical practice. This contrasts with previous research suggesting that health care professionals may resist new technology due to concerns about workflow disruptions or unfamiliarity with remote care models. In this context, previous experience with remote support likely mitigated these challenges, highlighting the importance of accounting for contextual factors when implementing teleconsultation. Furthermore, the identification of a low volume of teleconsultations as a barrier contradicts the common assumption that a phased rollout is always beneficial. Instead, our results suggest that achieving a critical mass of teleconsultations is essential to maintaining staff engagement and competencies. These findings provide new insights into teleconsultation implementation by demonstrating how preexisting practices

XSL•FO
RenderX

and use patterns can significantly influence the adoption and sustainability of technological innovations in long-term care settings.

Efforts to implement overnight nursing teleconsultation in long-term care are more likely to succeed if they are based on an understanding of the forces driving the dissemination and scale-up of teleconsultation. Therefore, further research is needed to develop and strengthen the conceptual and applied foundations of the dissemination and scale-up of health care innovations, especially in the context of Quebec's emerging learning health care systems.

Acknowledgments

This study was funded and supported by the Ministry of Health and Social Services (Ministère de la Santé et des Services sociaux) in Quebec, Canada. The authors are most grateful to the health professionals and managers of the Integrated Health and Social Services Centre of Abitibi-Témiscamingue and the Integrated University Health and Social Services Centre of Mauricie and Centre-du-Québec. The authors wish to thank everyone who participated in interviews and discussions.

Authors' Contributions

VN conceptualized and designed the study, collected the data, conducted the interviews, conducted the analysis, and wrote the first draft of the manuscript. VN and VP have read and revised the manuscript. All authors (VN, VP, and MCL) approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File , 16 KB - aging_v8i1e71950_app1.docx]

References

1. Fernández Coves A, Yeung KH, van der Putten IM, Nelson EA. Teleconsultation adoption since COVID-19: comparison of barriers and facilitators in primary care settings in Hong Kong and the Netherlands. Health Policy 2022 Oct;126(10):933-944 [FREE Full text] [doi: 10.1016/j.healthpol.2022.07.012] [Medline: 36050194]
2. Smith AC, Thomas E, Snoswell CL, Haydon H, Mehrotra A, Clemensen J, et al. Telehealth for global emergencies: implications for coronavirus disease 2019 (COVID-19). J Telemed Telecare 2020 Jun;26(5):309-313 [FREE Full text] [doi: 10.1177/1357633X20916567] [Medline: 32196391]
3. Omboni S, Padwal RS, Alessa T, Benczúr B, Green BB, Hubbard I, et al. The worldwide impact of telemedicine during COVID-19: current evidence and recommendations for the future. Connect Health 2022 Jan 04;1:7-35 [FREE Full text] [doi: 10.20517/ch.2021.03] [Medline: 35233563]
4. Williams S, Barnard A, Collis P, Correia de Sousa J, Ghimire S, Habib M, et al. Remote consultations in primary care across low-, middle- and high-income countries: implications for policy and care delivery. J Health Serv Res Policy 2023 Jul;28(3):181-189 [FREE Full text] [doi: 10.1177/13558196221140318] [Medline: 36484225]
5. Wherton J, Greenhalgh T, Hughes G, Shaw SE. The role of information infrastructures in scaling up video consultations during COVID-19: mixed methods case study into opportunity, disruption, and exposure. J Med Internet Res 2022 Nov 10;24(11):e42431 [FREE Full text] [doi: 10.2196/42431] [Medline: 36282978]
6. Salisbury C, Murphy M, Duncan P. The impact of digital-first consultations on workload in general practice: modeling study. J Med Internet Res 2020 Jun 16;22(6):e18203 [FREE Full text] [doi: 10.2196/18203] [Medline: 32543441]
7. Norberg BL, Austad B, Kristiansen E, Zanaboni P, Getz LO. The impact and wider implications of remote consultations for general practice in Norway: qualitative study among Norwegian contract general practitioners. JMIR Form Res 2024 Dec 17;8:e63068 [FREE Full text] [doi: 10.2196/63068] [Medline: 39688890]
8. Xuan Y, Guo C, Lu W. The effects of information continuity and interpersonal continuity on physician services online: cross-sectional study. JMIR Med Inform 2022 Jul 21;10(7):e35830 [FREE Full text] [doi: 10.2196/35830] [Medline: 35862183]
9. Hawrysz L, Kludacz-Alessandri M, Walczak R. Predictive factors of physicians' satisfaction and quality of work under teleconsultation conditions: structural equation analysis. JMIR Hum Factors 2024 Jun 10;11:e47810 [FREE Full text] [doi: 10.2196/47810] [Medline: 38857081]
10. Campbell K, Greenfield G, Li E, O'Brien N, Hayhoe B, Beaney T, et al. The impact of virtual consultations on the quality of primary care: systematic review. J Med Internet Res 2023 Aug 30;25:e48920 [FREE Full text] [doi: 10.2196/48920] [Medline: 37647117]

XSL•FO
RenderX

11. Meng G, McAiney C, McKillop I, Perlman CM, Tsao S, Chen H. Factors that influence patient satisfaction with the service quality of home-based teleconsultation during the COVID-19 pandemic: cross-sectional survey study. JMIR Cardio 2024 Feb 16;8:e51439 [FREE Full text] [doi: 10.2196/51439] [Medline: 38363590]
12. Andrews E, Berghofer K, Long J, Prescott A, Caboral-Stevens M. Corrigendum to "satisfaction with use of telehealth during COVID-19: an integrative review" [international journal of nursing studies advanced 2 (2020). Article number 100008]. Int J Nurs Stud Adv 2022 Dec;4:100092 [FREE Full text] [doi: 10.1016/j.ijnsa.2022.100092] [Medline: 35989885]
13. Hong YR, Lawrence J, Williams Jr D, Mainous III A. Population-level interest and telehealth capacity of us hospitals in response to COVID-19: cross-sectional analysis of google search and national hospital survey data. JMIR Public Health Surveill 2020 Apr 07;6(2):e18961 [FREE Full text] [doi: 10.2196/18961] [Medline: 32250963]
14. Koonin LM, Hoots B, Tsang CA, Leroy Z, Farris K, Jolly T, et al. Trends in the use of telehealth during the emergence of the COVID-19 pandemic - United States, January-March 2020. MMWR Morb Mortal Wkly Rep 2020 Oct 30;69(43):1595-1599 [FREE Full text] [doi: 10.15585/mmwr.mm6943a3] [Medline: 33119561]
15. Tenforde AS, Hefner JE, Kodish-Wachs JE, Iaccarino MA, Paganoni S. Telehealth in physical medicine and rehabilitation: a narrative review. PM R 2017 May;9(5S):S51-S58. [doi: 10.1016/j.pmrj.2017.02.013] [Medline: 28527504]
16. Lin Y, Xu X, Liu Y, Alias H, Hu Z, Wong LP. Perception and acceptance of telemedicine use in health care among the general public in China: web-based cross-sectional survey. J Med Internet Res 2024 Jul 16;26:e53497 [FREE Full text] [doi: 10.2196/53497] [Medline: 39012687]
17. Huguet M, Sarazin M, Perrier L, Augusto V. How we can reap the full benefit of teleconsultations: economic evaluation combined with a performance evaluation through a discrete-event simulation. J Med Internet Res 2022 May 20;24(5):e32002 [FREE Full text] [doi: 10.2196/32002] [Medline: 35594065]
18. Snoswell CL, Taylor ML, Comans TA, Smith AC, Gray LC, Caffery LJ. Determining if telehealth can reduce health system costs: scoping review. J Med Internet Res 2020 Oct 19;22(10):e17298 [FREE Full text] [doi: 10.2196/17298] [Medline: 33074157]
19. Nabelsi V, Lévesque-Chouinard A. Successful electronic consultation service initiative in Quebec, Canada with primary care physicians' and specialists' experiences on acceptance and use of technological innovation: cross-sectional exploratory study. JMIR Form Res 2024 May 30;8:e52921 [FREE Full text] [doi: 10.2196/52921] [Medline: 38814689]
20. Almathami HK, Win KT, Vlahu-Gjorgievska E. Empirical evidence of internal and external factors influencing users' motivation toward teleconsultation use. Telemed J E Health 2024 Jan;30(1):141-156. [doi: 10.1089/tmj.2022.0527] [Medline: 37343179]
21. Kichloo A, Albosta M, Dettloff K, Wani F, El-Amir Z, Singh J, et al. Telemedicine, the current COVID-19 pandemic and the future: a narrative review and perspectives moving forward in the USA. Fam Med Community Health 2020 Aug;8(3):e000530 [FREE Full text] [doi: 10.1136/fmch-2020-000530] [Medline: 32816942]
22. Nabelsi V, Lévesque-Chouinard A, Liddy C, Dumas Pilon M. Improving the referral process, timeliness, effectiveness, and equity of access to specialist medical services through electronic consultation: pilot study. JMIR Med Inform 2019 Jul 10;7(3):e13354 [FREE Full text] [doi: 10.2196/13354] [Medline: 31293239]
23. Roberson AE, Carlson M, Kohler CM, Harris PA, Volkmann CL. Initiating virtual nursing in general inpatient care. Am J Nurs 2023 Jun 01;123(6):48-54. [doi: 10.1097/01.NAJ.0000938736.42266.5e] [Medline: 37233140]
24. Savitz ST, Frederick RK, Sangaralingham LR, Lampman MA, Anderson SS, Habermann EB, et al. Evaluation of safety and care outcomes after the introduction of a virtual registered nurse model. Health Serv Res 2023 Oct;58(5):999-1013. [doi: 10.1111/1475-6773.14208] [Medline: 37525521]
25. Steaban RL, Morrison J, Simmons S, Ivory C, France D, Leming-Lee S, et al. Strategies to evaluate new models of nursing care to meet hospital staffing and patient care needs. Nurse Lead 2024 Dec;22(6):718-724. [doi: 10.1016/j.mnl.2024.05.007]
26. Nott S, Wingfield G, Haigh A, Luscombe GM, Thompson AE, Saurman E, et al. The Virtual Rural Generalist Service: a hybrid virtual model of care designed to improve health access and outcomes in rural and remote communities. Med J Aust 2024 Dec 09;221 Suppl 11:S3-S7. [doi: 10.5694/mja2.52529] [Medline: 39647931]
27. Thompson AE, Saurman E, Nott S, Wilson A, Shaw T. Clinician experiences of a hybrid virtual medical service supporting rural and remote hospitals: a qualitative study. Med J Aust 2024 Dec 09;221 Suppl 11:S16-S21. [doi: 10.5694/mja2.52525] [Medline: 39647921]
28. Regragui S, Abou Malham S, Gaboury I, Bois C, Deville-Stoetzel N, Maillet L, et al. Nursing practice and teleconsultations in a pandemic context: a mixed-methods study. J Clin Nurs 2023 Sep;32(17-18):6339-6353. [doi: 10.1111/jocn.16756] [Medline: 37202866]
29. Vaughan C, Lukewich J, Mathews M, Hedden L, Poitras ME, Asghari S, et al. Nursing contributions to virtual models of care in primary care: a scoping review protocol. BMJ Open 2022 Sep 20;12(9):e065779 [FREE Full text] [doi: 10.1136/bmjopen-2022-065779] [Medline: 36127080]
30. Booth RG, Strudwick G, McBride S, O'Connor S, Solano López AL. How the nursing profession should adapt for a digital future. BMJ 2021 Jun 14;373:n1190. [doi: 10.1136/bmj.n1190]
31. Bashshur R, Doarn CR, Frenk JM, Kvedar JC, Woolliscroft JO. Telemedicine and the COVID-19 pandemic, lessons for the future. Telemed J E Health 2020 May;26(5):571-573. [doi: 10.1089/tmj.2020.29040.rb] [Medline: 32275485]

XSL•FO
RenderX

32. Dai Z. Telehealth in long-term care facilities during the COVID-19 pandemic - lessons learned from patients, physicians, nurses and healthcare workers. BMC Digit Health 2023;1(1):2 [FREE Full text] [doi: 10.1186/s44247-022-00003-y] [Medline: 38014371]
33. Tan AJ, Rusli KD, McKenna L, Tan LL, Liaw SY. Telemedicine experiences and perspectives of healthcare providers in long-term care: a scoping review. J Telemed Telecare 2024 Feb;30(2):230-249. [doi: 10.1177/1357633X211049206] [Medline: 34666535]
34. Nabelsi V, Plouffe V. Assessing cost and cost savings of teleconsultation in long-term care facilities: a time-driven activity-based costing analysis within a value-based healthcare framework. BMC Health Serv Res 2024 Sep 13;24(1):1064 [FREE Full text] [doi: 10.1186/s12913-024-11578-9] [Medline: 39272121]
35. Nabelsi V, Leclerc MC, Plouffe V. Nurses' and nursing assistants' experiences with teleconsultation in small rural long-term care facilities: semistructured interview pilot study. JMIR Aging 2024 Nov 27;7:e65111 [FREE Full text] [doi: 10.2196/65111] [Medline: 39602791]
36. Özkaytan Y, Schulz-Nieswandt F, Zank S. Acute health care provision in rural long-term care facilities: a scoping review of integrated care models. J Am Med Dir Assoc 2023 Oct;24(10):1447-57.e1. [doi: 10.1016/j.jamda.2023.06.013] [Medline: 37488029]
37. Qiao Y, Ran L, Li J. Optimization of teleconsultation using discrete-event simulation from a data-driven perspective. Telemed J E Health 2020 Jan 01;26(1):114-125. [doi: 10.1089/tmj.2018.0229] [Medline: 30994407]
38. Abdolkhani R, Petersen S, Walter R, Zhao L, Butler-Henderson K, Livesay K. The impact of digital health transformation driven by COVID-19 on nursing practice: systematic literature review. JMIR Nurs 2022 Aug 30;5(1):e40348 [FREE Full text] [doi: 10.2196/40348] [Medline: 35867838]
39. Santos MT, Halberstadt BM, Trindade CR, Lima MA, Aued GK. Continuity and coordination of care: conceptual interface and nurses' contributions. Rev Esc Enferm USP 2022;56:e20220100 [FREE Full text] [doi: 10.1590/1980-220X-REEUSP-2022-0100en] [Medline: 35951450]
40. Cacchione PZ. Innovative care models across settings: providing nursing care to older adults. Geriatr Nurs 2020;41(1):16-20. [doi: 10.1016/j.gerinurse.2020.01.011] [Medline: 32033809]
41. Qi Tan AJ, Chua WL, McKenna L, Chin Tan LL, Lim YJ, Liaw SY. Enablers and barriers to nurse-facilitated geriatric teleconsultations in nursing homes: a qualitative descriptive multi-site study. Age Ageing 2022 Dec 05;51(12):afac268. [doi: 10.1093/ageing/afac268] [Medline: 36580553]
42. Chez soi : le premier choix La politique de soutien à domicile. Ministère de la santé et des Services Sociaux. 2003. URL: https://publications.msss.gouv.qc.ca/msss/fichiers/2002/02-704-01.pdf [accessed 2024-08-03]
43. Documentation des soins infirmiers : Norme d'exercice. Ordre des infirmières et infirmiers du Québec. 2023. URL: https://www.oiiq.org/norme-exercice-documentation [accessed 2024-08-04]
44. Rutledge CM, O'Rourke J, Mason AM, Chike-Harris K, Behnke L, Melhado L, et al. Telehealth competencies for nursing education and practice: the four P's of telehealth. Nurse Educ 2021 Jan 19;46(5):300-305 [FREE Full text] [doi: 10.1097/NNE.0000000000000988] [Medline: 33481494]
45. Edelman LS, McConnell ES, Kennerly SM, Alderden J, Horn SD, Yap TL. Mitigating the effects of a pandemic: facilitating improved nursing home care delivery through technology. JMIR Aging 2020 May 26;3(1):e20110 [FREE Full text] [doi: 10.2196/20110] [Medline: 32412909]
46. Chaudoir SR, Dugan AG, Barr CH. Measuring factors affecting implementation of health innovations: a systematic review of structural, organizational, provider, patient, and innovation level measures. Implement Sci 2013 Feb 17;8(10):22-29 [FREE Full text] [doi: 10.1186/1748-5908-8-22] [Medline: 23414420]
47. Fortin MF. Foundations and stages of the research process: quantitative and qualitative methods. 2nd edition. Chenelière éducation, Montréal. URL: https://www.scirp.org/reference/referencespapers?referenceid=2631793 [accessed 2024-04-29]
48. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018 Mar;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
49. Sousa VL, Dourado Júnior FW, Anjos SD, Moreira AC. Nursing teleconsultation in primary health care: scoping review. Rev Lat Am Enfermagem 2024;32:e4329 [FREE Full text] [doi: 10.1590/1518-8345.7212.4329] [Medline: 39230178]
50. Hughes L, Petrella A, Phillips N, Taylor RM. Virtual care and the impact of COVID-19 on nursing: a single centre evaluation. J Adv Nurs 2022 Feb;78(2):498-509 [FREE Full text] [doi: 10.1111/jan.15050] [Medline: 34590738]
51. Mohammed HT, Hyseni L, Bui V, Gerritsen B, Fuller K, Sung J, et al. Exploring the use and challenges of implementing virtual visits during COVID-19 in primary care and lessons for sustained use. PLoS One 2021;16(6):e0253665 [FREE Full text] [doi: 10.1371/journal.pone.0253665] [Medline: 34166441]
52. Teasley SL, Sexton KA, Carroll CA, Cox KS, Riley M, Ferriell K. Improving work environment perceptions for nurses employed in a rural setting. J Rural Health 2007;23(2):179-182. [doi: 10.1111/j.1748-0361.2007.00087.x] [Medline: 17397376]
53. Li S, Jeffs L, Barwick M, Stevens B. Organizational contextual features that influence the implementation of evidence-based practices across healthcare settings: a systematic integrative review. Syst Rev 2018 May 05;7(1):72 [FREE Full text] [doi: 10.1186/s13643-018-0734-5] [Medline: 29729669]

XSL•FO
RenderX

54. Hollander JE, Carr BG. Virtually perfect? Telemedicine for COVID-19. N Engl J Med 2020 Apr 30;382(18):1679-1681. [doi: 10.1056/NEJMp2003539] [Medline: 32160451]
55. Zhang T, Mosier J, Subbian V. Identifying barriers to and opportunities for telehealth implementation amidst the COVID-19 pandemic by using a human factors approach: a leap into the future of health care delivery? JMIR Hum Factors 2021 Apr 09;8(2):e24860 [FREE Full text] [doi: 10.2196/24860] [Medline: 33779566]
56. Mohammed HM, El-sol AE. Tele-nursing: opportunities for nurses to shape their profession's future. Int J Nov Res Healthc Nurs 2020;7(3):660-676 [FREE Full text]
57. Nittari G, Khuman R, Baldoni S, Pallotta G, Battineni G, Sirignano A, et al. Telemedicine practice: review of the current ethical and legal challenges. Telemed J E Health 2020 Dec;26(12):1427-1437 [FREE Full text] [doi: 10.1089/tmj.2019.0158] [Medline: 32049608]
58. Gillespie SM, Handler SM, Bardakh A. Innovation through regulation: COVID-19 and the evolving utility of telemedicine. J Am Med Dir Assoc 2020 Aug;21(8):1007-1009 [FREE Full text] [doi: 10.1016/j.jamda.2020.06.054] [Medline: 32736843]
59. Jacob C, Sanchez-Vazquez A, Ivory C. Understanding clinicians' adoption of mobile health tools: a qualitative review of the most used frameworks. JMIR Mhealth Uhealth 2020 Jul 06;8(7):e18072 [FREE Full text] [doi: 10.2196/18072] [Medline: 32442132]
60. Cannavacciuolo L, Capaldo G, Ponsiglione C. Digital innovation and organizational changes in the healthcare sector: multiple case studies of telemedicine project implementation. Technovation 2023 Feb;120:102550. [doi: 10.1016/j.technovation.2022.102550]
61. Kumar S, Tian E. Critical factors in the development, implementation and evaluation of telemedicine. In: Rashid TA, Chakraborty C, Fraser K, editors. Advances in Telemedicine for Health Monitoring: Technologies, Design and Applications. London, UK: The Institution of Engineering and Technology; 2020:1-14.
62. Beauséjour W, Hagens S. Uncovering important drivers of the increase in the use of virtual care technologies in nursing care: quantitative analysis from the 2020 national survey of Canadian nurses. JMIR Nurs 2022 Mar 31;5(1):e33586 [FREE Full text] [doi: 10.2196/33586] [Medline: 35357326]
63. Consolidated telemedicine implementation guide. World Health Organization. URL: https://www.who.int/publications/i/item/9789240059184 [accessed 2024-04-29]
64. Kho J, Gillespie N, Martin-Khan M. A systematic scoping review of change management practices used for telemedicine service implementations. BMC Health Serv Res 2020 Sep 01;20(1):815 [FREE Full text] [doi: 10.1186/s12913-020-05657-w] [Medline: 32873295]
65. Fennelly O, Cunningham C, Grogan L, Cronin H, O'Shea C, Roche M, et al. Successfully implementing a national electronic health record: a rapid umbrella review. Int J Med Inform 2020 Dec;144:104281 [FREE Full text] [doi: 10.1016/j.ijmedinf.2020.104281] [Medline: 33017724]
66. Ifinedo P. Empirical study of nova scotia nurses' adoption of healthcare information systems: implications for management and policy-making. Int J Health Policy Manag 2018 Apr 01;7(4):317-327 [FREE Full text] [doi: 10.15171/ijhpm.2017.96] [Medline: 29626399]
67. Sriharan A, Sekercioglu N, Mitchell C, Senkaiahliyan S, Hertelendy A, Porter T, et al. Leadership for AI transformation in health care organization: scoping review. J Med Internet Res 2024 Aug 14;26:e54556 [FREE Full text] [doi: 10.2196/54556] [Medline: 39009038]
68. Pfitzer E, Bitomsky L, Nißen M, Kausch C, Kowatsch T. Success factors of growth-stage digital health companies: systematic literature review. J Med Internet Res 2024 Dec 11;26:e60473 [FREE Full text] [doi: 10.2196/60473] [Medline: 39661978]
69. Greenhalgh T, Ladds E, Hughes G, Moore L, Wherton J, Shaw SE, et al. Why do GPs rarely do video consultations? Qualitative study in UK general practice. Br J Gen Pract 2022 May;72(718):e351-e360 [FREE Full text] [doi: 10.3399/BJGP.2021.0658] [Medline: 35256385]
70. Lin CC, Dievler A, Robbins C, Sripipatana A, Quinn M, Nair S. Telehealth in health centers: key adoption factors, barriers, and opportunities. Health Aff (Millwood) 2018 Dec;37(12):1967-1974. [doi: 10.1377/hlthaff.2018.05125] [Medline: 30633683]
71. Garber K, Gustin T. Telehealth education: impact on provider experience and adoption. Nurse Educ 2022;47(2):75-80 [FREE Full text] [doi: 10.1097/NNE.0000000000001103] [Medline: 34628435]
72. James S, Ashley C, Williams A, Desborough J, Mcinnes S, Calma K, et al. Experiences of Australian primary healthcare nurses in using telehealth during COVID-19: a qualitative study. BMJ Open 2021 Aug 06;11(8):e049095 [FREE Full text] [doi: 10.1136/bmjopen-2021-049095] [Medline: 34362804]
73. Esmaeilzadeh P, Mirzaei T. Using electronic health records to mitigate workplace burnout among clinicians during the COVID-19 pandemic: field study in Iran. JMIR Med Inform 2021 Jun 03;9(6):e28497 [FREE Full text] [doi: 10.2196/28497] [Medline: 34033578]
74. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
75. Opoku D, Scott P, Quentin W. Healthcare professionals' perceptions of the benefits and challenges of a teleconsultation service in the Amansie-west district of Ghana. Telemed J E Health 2015 Sep;21(9):748-755. [doi: 10.1089/tmj.2014.0210] [Medline: 25942579]

XSL•FO
RenderX

76. Gagnon MP, Paré G, Pollender H, Duplantie J, Côté J, Fortin JP, et al. Supporting work practices through telehealth: impact on nurses in peripheral regions. BMC Health Serv Res 2011 Feb 04;11:27 [FREE Full text] [doi: 10.1186/1472-6963-11-27] [Medline: 21294882]
77. White M, Butterworth T, Wells JS. Reported implementation lessons from a national quality improvement initiative; productive ward: releasing time to Care™. A qualitative, ward-based team perspective. J Nurs Manag 2017 Oct;25(7):519-530. [doi: 10.1111/jonm.12489] [Medline: 28799269]
78. Rafferty AE, Jimmieson NL, Armenakis AA. Change readiness: a multilevel review. J Manag 2012 Sep 05;39(1):110-135. [doi: 10.1177/0149206312457417]
79. Meurs M, Keuper J, Sankatsing V, Batenburg R, van Tuyl L. "Get used to the fact that some of the care is really going to take place in a different way": general practitioners' experiences with e-health during the COVID-19 pandemic. Int J Environ Res Public Health 2022 Apr 22;19(9):5120 [FREE Full text] [doi: 10.3390/ijerph19095120] [Medline: 35564519]
80. Harst L, Lantzsch H, Scheibe M. Theories predicting end-user acceptance of telemedicine use: systematic review. J Med Internet Res 2019 May 21;21(5):e13117 [FREE Full text] [doi: 10.2196/13117] [Medline: 31115340]
81. Fronczek AE, Rouhana NA, Kitchin JM. Enhancing telehealth education in nursing: applying King's conceptual framework and theory of goal attainment. Nurs Sci Q 2017 Jul 21;30(3):209-213. [doi: 10.1177/0894318417708418] [Medline: 28899265]
82. van Houwelingen CT, Moerman AH, Ettema RG, Kort HS, Ten Cate O. Competencies required for nursing telehealth activities: a Delphi-study. Nurse Educ Today 2016 Apr;39:50-62 [FREE Full text] [doi: 10.1016/j.nedt.2015.12.025] [Medline: 27006033]
83. Nilsen P, Seing I, Ericsson C, Birken SA, Schildmeijer K. Characteristics of successful changes in health care organizations: an interview study with physicians, registered nurses and assistant nurses. BMC Health Serv Res 2020 Feb 27;20(1):147 [FREE Full text] [doi: 10.1186/s12913-020-4999-8] [Medline: 32106847]
84. Arem H, Moses J, Cisneros C, Blondeau B, Nekhlyudov L, Killackey M, et al. Cancer provider and survivor experiences with telehealth during the COVID-19 pandemic. JCO Oncol Pract 2022 Apr;18(4):e452-e461. [doi: 10.1200/OP.21.00401] [Medline: 34714706]
85. Srinivasan M, Asch S, Vilendrer S, Thomas SC, Bajra R, Barman L, et al. Qualitative assessment of rapid system transformation to primary care video visits at an academic medical center. Ann Intern Med 2020 Oct 06;173(7):527-535. [doi: 10.7326/m20-1814]
86. Belcher RH, Phillips J, Virgin F, Werkhaven J, Whigham A, Wilcox L, et al. Pediatric otolaryngology telehealth in response to COVID-19 pandemic: lessons learned and impact on the future management of pediatric patients. Ann Otol Rhinol Laryngol 2021 Jul;130(7):788-795 [FREE Full text] [doi: 10.1177/0003489420976163] [Medline: 33238714]
87. Franzosa E, Gorbenko K, Brody AA, Leff B, Ritchie CS, Kinosian B, et al. "There is something very personal about seeing someone's face": provider perceptions of video visits in home-based primary care during COVID-19. J Appl Gerontol 2021 Nov;40(11):1417-1424 [FREE Full text] [doi: 10.1177/07334648211028393] [Medline: 34210200]
88. Gilkey MB, Kong WY, Huang Q, Grabert BK, Thompson P, Brewer NT. Using telehealth to deliver primary care to adolescents during and after the COVID-19 pandemic: national survey study of US primary care professionals. J Med Internet Res 2021 Sep 10;23(9):e31240 [FREE Full text] [doi: 10.2196/31240] [Medline: 34406974]
89. Connolly SL, Gifford AL, Miller CJ, Bauer MS, Lehmann LS, Charness ME. Provider perceptions of virtual care during the coronavirus disease 2019 pandemic: a multispecialty survey study. Med Care 2021 Jul 01;59(7):646-652 [FREE Full text] [doi: 10.1097/MLR.0000000000001562] [Medline: 34009880]
90. Chen W, Flanagan A, Nippak PM, Nicin M, Sinha SK. Understanding the experience of geriatric care professionals in using telemedicine to care for older patients in response to the COVID-19 pandemic: mixed methods study. JMIR Aging 2022 Aug 10;5(3):e34952 [FREE Full text] [doi: 10.2196/34952] [Medline: 35830331]
91. Pilosof NP, Barrett M, Oborn E, Barkai G, Pessach IM, Zimlichman E. Telemedicine implementation in COVID-19 ICU: balancing physical and virtual forms of visibility. HERD 2021 Jul 02;14(3):34-48 [FREE Full text] [doi: 10.1177/19375867211009225] [Medline: 34075789]
92. Kendzerska T, Zhu DT, Gershon AS, Edwards JD, Peixoto C, Robillard R, et al. The effects of the health system response to the COVID-19 pandemic on chronic disease management: a narrative review. Risk Manag Healthc Policy 2021;14:575-584 [FREE Full text] [doi: 10.2147/RMHP.S293471] [Medline: 33623448]
93. Park HY, Kwon YM, Jun HR, Jung SE, Kwon SY. Satisfaction survey of patients and medical staff for telephone-based telemedicine during hospital closing due to COVID-19 transmission. Telemed J E Health 2021 Jul;27(7):724-732 [FREE Full text] [doi: 10.1089/tmj.2020.0369] [Medline: 33216710]
94. Hunter SB, Dopp AR, Ober AJ, Uscher-Pines L. Clinician perspectives on methadone service delivery and the use of telemedicine during the COVID-19 pandemic: a qualitative study. J Subst Abuse Treat 2021 May;124:108288 [FREE Full text] [doi: 10.1016/j.jsat.2021.108288] [Medline: 33771285]
95. Lehoux P, Sicotte C, Denis JL, Berg M, Lacroix A. The theory of use behind telemedicine: how compatible with physicians' clinical routines? Soc Sci Med 2002 Mar;54(6):889-904. [doi: 10.1016/s0277-9536(01)00063-6] [Medline: 11996023]

XSL•FO
RenderX

96. Safaeinili N, Vilendrer S, Williamson E, Zhao Z, Brown-Johnson C, Asch SM, et al. Inpatient telemedicine implementation as an infection control response to COVID-19: qualitative process evaluation study. JMIR Form Res 2021 Jun 16;5(6):e26452 [FREE Full text] [doi: 10.2196/26452] [Medline: 34033576]

Abbreviations

CHSLD: centre d'hébergement et de soins de longue durée
MSSS: Ministère de la Santé et des Services sociaux
ND: nursing department
SAPA: Support Program for the Autonomy of Seniors

Edited by AF Abdulai; submitted 30.01.25; peer-reviewed by R Tripathi, T David, C Udensi; comments to author 23.03.25; revised version received 26.03.25; accepted 12.04.25; published 07.05.25.

Please cite as:
Nabelsi V, Plouffe V, Leclerc MC
Barriers to and Facilitators of Implementing Overnight Nursing Teleconsultation in Small, Rural Long-Term Care Facilities: Qualitative Interview Study
JMIR Aging 2025;8:e71950
URL: https://aging.jmir.org/2025/1/e71950
doi:10.2196/71950
PMID:

©Veronique Nabelsi, Véronique Plouffe, Marie Chantal Leclerc. Originally published in JMIR Aging (https://aging.jmir.org), 07.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Intraoperative Hypotension and Major Adverse Cardiac Events Among Older Adult Patients Undergoing Noncardiac Surgery: Retrospective Cohort Study

Kai Zhang[1,2*], PhD; Chang Liu[1,3*], PhD; Meng Wang[4*], MD; Ting Zhang[1,2], MD; Bingbing Meng[1,2], MD; Siyi Yao[1,2], MD; Jingsheng Lou[1,2,5], MD, PhD; Qiang Fu[1,2,5], MD, PhD; Yanhong Liu[1,2,5], MD, PhD; Jiangbei Cao[1,2,5], MD, PhD; Lulong Bo[4*], MD, PhD; Weidong Mi[1,2,5*], MD, PhD; Hao Li[1,2,5*], MD, PhD

[1]Department of Anesthesiology, The First Medical Center, Chinese People's Liberation Army General Hospital, 28th Fuxing Road, Haidian District, Beijing, China

[2]Medical School of Chinese People's Liberation Army General Hospital, Beijing, China

[3]School of Medicine, Nankai University, Tianjin, China

[4]Department of Anesthesiology, Changhai Hospital, Naval Medical University, Shanghai, China

[5]National Clinical Research Center for Geriatric Diseases, Chinese People's Liberation Army General Hospital, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Hao Li, MD, PhD
Department of Anesthesiology, The First Medical Center, Chinese People's Liberation Army General Hospital, 28th Fuxing Road, Haidian District, Beijing, China

Abstract

Background: Intraoperative hypotension (IOH) is an important risk factor for major adverse cardiac events (MACE) in patients undergoing noncardiac surgery. However, the IOH threshold in older adult patients remains controversial.

Objective: This study aimed to explore an appropriate IOH threshold in older adult patients to decrease the risk of MACE.

Methods: This study involved older adult patients undergoing noncardiac surgery (age ≥65 y) from January 2012 to August 2019 in the Chinese People's Liberation Army General Hospital (PLAGH; 35,262 patients) and Shanghai Changhai Hospital from January 2024 to December 2024 (13,418 patients). Univariate moving-average plots and multivariate restricted cubic splines were used to determine the IOH thresholds associated with an increased risk of MACE. The relationship between the IOH threshold and MACE was assessed using univariate and multivariate logistic regression analyses by 3 different hypotension exposure forms (duration, area, and time-weighted average mean arterial pressure [MAP]).

Results: Out of 35,262 patients, 874 developed MACE in PLAGH, and 296 of 13,418 patients developed MACE in Changhai Hospital. In PLAGH, MAP below an absolute threshold of 70 mm Hg was associated with MACE. When the IOH absolute threshold was 70 mm Hg, the risk of MACE demonstrated a "dose-increasing" effect with changes in IOH exposure, and the risk of MACE was significantly increased when the duration lasted >15 minutes (odds ratio 1.51, 95% CI 1.22-1.88; *P*<.001). The stratified analysis showed that in patients younger than 80 years, when intraoperative MAP dropped below 70 mm Hg for more than 15 minutes, the odds ratio was 1.38 (95% CI 0.86 - 2.28), *P*<.01. In Changhai hospital, intraoperative MAP <70 mm Hg was also significantly associated with MACE. Furthermore, IOH lasting longer than 15 minutes substantially increased the risk of MACE.

Conclusions: For older adult patients undergoing noncardiac surgery, intraoperative MAP should be kept above 70 mm Hg to reduce the risk of postoperative MACE.

(JMIR Aging 2025;8:e67177) doi:10.2196/67177

KEYWORDS

intraoperative hypotension; major adverse cardiac events; geriatric patients; threshold; mean arterial pressure

Introduction

Each year, more than 300 million surgical procedures are conducted globally, with noncardiac surgeries accounting for 85% of the total [1]. With the global rise in the aging population, there has been a corresponding increase in the number of older adult surgical patients, leading to a higher incidence of postoperative complications [2]. Previous studies have suggested that major adverse cardiac events (MACE) occur in more than 2% of patients and are responsible for one-third of postoperative

XSL•FO
RenderX

deaths [3-5]. Because of the decreased physiologic reserve and multiple coexisting diseases in older adult patients [5,6], the risk of postoperative MACE in these patients is higher and needs more attention.

Intraoperative hypotension (IOH) is reportedly associated with postoperative cardiovascular complications and mortality in patients undergoing noncardiac surgery, and the association becomes magnified as the severity of hypotension increases [7-10]. With the continuous advancement of medical technology, recent years have witnessed the emergence of numerous intervention measures aimed at preventing IOH. One notable example is the intraoperative hypotension prediction index developed by Hatib et al [11], which uses machine learning algorithms for its construction. Clinical studies have substantiated the efficacy of this index in reducing the incidence of IOH [12-14]. However, further research is warranted to ascertain whether its application can also lead to a reduction in postoperative complications. Concurrently, the majority of studies using the hypotension prediction index have established a hypotension threshold of 65 mm Hg of mean arterial pressure (MAP) [12,13]. This threshold is largely derived from research conducted on adult patients. However, older adult patients exhibit reduced cardiopulmonary reserve capacity, diminished autonomic nervous system self-regulation, and a cardiovascular system that is often in a precarious state of equilibrium. Therefore, the applicability of this threshold to older adult patients remains a subject of debate. Establishing an accurate threshold for hypotension is crucial in optimizing older adult patient outcomes.

Therefore, this study was performed to explore the IOH threshold and its relationship with MACE in older adult patients undergoing noncardiac surgery.

Methods

Study Population

Perioperative data for patients required in this study were obtained from the First Medical Center of Chinese People's Liberation Army General Hospital (PLAGH) and Shanghai Changhai Hospital, respectively. Perioperative data from the First Medical Center (January 2012 to August 2019) involved older adult surgical patients and were primarily used for exploring IOH thresholds associated with increased risk of MACE in this population. Data from Shanghai Changhai Hospital (January 2024 to December 2024) were specifically used for validating IOH thresholds in older adult surgical patients.

Ethical Considerations

The Research Ethics Committee of the Chinese PLAGH approved this study (approval reference S2019-311-02). The requirement for informed consent was waived because this retrospective cohort study posed minimal risk to participants. The patients' names and ID numbers were anonymized during data collection.

Inclusion and Exclusion Criteria

Patients aged ≥65 years, who were scheduled to undergo noncardiac surgery under anesthesia, were included in this study. The exclusion criteria were nongeneral anesthesia, American Society of Anesthesiologists (ASA) physical status classification of >IV, hysteroscopic or body surface surgery [15], surgery duration of ≤30 minutes, incomplete clinical data, no intraoperative blood pressure values at 30-second intervals, and missing data on patient characteristics.

Primary Outcomes

The primary outcome was MACE, defined as the composite of acute myocardial infarction (MI), unstable angina, heart failure (HF), new-onset severe arrhythmia, nonfatal cardiac arrest, and cardiac death, during the operation and within 30 days after noncardiac surgery [16,17]. Patients with MACE were identified by examining their medical records. For data elements in the patients' medical records, we used structured query language, a standard language for storing, manipulating, and retrieving data from databases, to identify instances of data domains. We extracted the patients' postoperative biochemical test results, electrocardiograms (ECGs), coronary angiograms, postoperative course descriptions, and consultation records describing the specific MACE. A patient could have more than one MACE. The complications were independently evaluated by 3 experienced doctors. Any disputes were resolved through discussion and negotiation.

MI [18] was defined as an increase and gradual decrease in troponin levels accompanied by at least one of the following: ischemic symptoms, abnormal Q waves on ECG, ST-segment elevation or depression, or coronary artery intervention (ie, coronary angioplasty). HF [17] was defined as new in-hospital signs or symptoms of dyspnea or fatigue, orthopnea, paroxysmal nocturnal dyspnea, increased jugular venous pressure, pulmonary rales on physical examination, cardiomegaly, or pulmonary vascular engorgement. New-onset severe arrhythmia was defined as ECG evidence of atrial flutter, atrial fibrillation, ventricular tachycardia, ventricular fibrillation, and second- or third-degree atrioventricular conduction block. Nonfatal cardiac arrest was defined as the absence of a cardiac rhythm or the presence of a chaotic rhythm requiring any component of basic or advanced cardiac life support. Cardiac death was defined as any death unless an unequivocal noncardiovascular cause could be established.

Data Collection

The following preoperative and intraoperative variables related to MACE were collected [19]. First, demographic variables (age, sex, and BMI) and smoking and drinking status. Smoking (drinking) status included nonsmoker (nondrinker), current smoker (drinker), and former smoker (drinker). Patients who had stopped smoking (drinking) for >1 year were considered former smokers (drinkers). Second, medical history (cerebrovascular disease, coronary heart disease, valvular heart disease, HF, arrhythmia, hypertension, diabetes mellitus, conditions requiring cardiac interventions, or peripheral arterial disease) and cardiac interventions, including coronary artery bypass grafting and percutaneous coronary intervention. Third,

XSL•FO
RenderX

the most recent laboratory test results before surgery (hemoglobin, serum creatinine [SCr], and fibrinogen). Fourth, preoperative medicine (antiplatelet drugs, β-blockers, angiotensin-converting enzyme inhibitors [ACEI] and angiotensin receptor blockers [ARB], diuretic, and statin). Finally, surgery-related variables (ASA score, surgery type, surgery duration, emergency surgery, blood loss, urine volume, crystalloid infusion, and blood pressure data). The blood pressure data included noninvasive and invasive blood pressure, which were recorded at 30-second intervals from the Madison database. If no invasive blood pressure measurement was obtained during the operation, noninvasive blood pressure data were recorded. The baseline MAP was defined as the patients' mean preoperative MAP. Artifactual data were removed using previously published criteria; for instance, an invalid MAP value was defined as (1) systolic blood pressure of ≥300 or ≤20 mm Hg, (2) systolic blood pressure ≤ diastolic blood pressure +5 mm Hg, (3) diastolic blood pressure of ≤5 or ≥225 mm Hg, or (4) MAP of >155 or <25 mm Hg [20].

Statistical Analysis

The patients were categorized by whether they had developed MACE. Continuous variables are expressed as mean and SD if normally distributed, and as median and IQR if nonnormally distributed. Categorical data are presented as frequency (%). For comparison of differences between the 2 groups, the variance test or Kruskal–Wallis rank-sum test was used for continuous variables, and the chi-square test or Fisher exact test was used for categorical variables, as appropriate.

Moving-average smoothing plots with widths of 1, 3, and 5 minutes were used to determine the lowest MAP. Then, as previously reported [9], the linearity between each MAP exposure and outcome was modeled by a restricted cubic spline function with 3 knots located at the 10th, 50th, and 90th percentiles, and univariate moving-average plots and multivariate smoothed cubic spline curves were used to determine the IOH threshold associated with an increased risk of MACE. The confounding variables were determined from the variables with a standard mean difference of <0.2 in the baseline and with clinical meaning. These confounding variables included age, cerebrovascular disease, coronary heart disease, valvular heart disease, HF, arrhythmia, hypertension, peripheral arterial disease, renal insufficiency, hemoglobin, SCr, fibrinogen, antiplatelet drugs, β-blockers, ACEI, ARB, diuretic, ASA score, surgery type, surgery duration, emergency surgery, blood loss, urine volume, and crystalloid infusion.

Next, univariate and multivariate logistic regression analyses were performed to validate the practical value of IOH threshold by three different IOH exposure methods as described in the previous study: (1) time spent at the threshold associated with harmful outcomes, (2) total area under the threshold–time plot, and (3) time-weighted average mean arterial pressure (TWA-MAP) under the threshold. After that, stratification analysis was used to explore whether the association between the IOH threshold and the occurrence of postoperative MACE differed between the very old (≥80 y old) and older adults.

Except where noted, a 2-tailed *P* value of<.05 indicated a significant difference. All analyses were performed using R version 3.6.3 (The R Foundation for Statistical Computing).

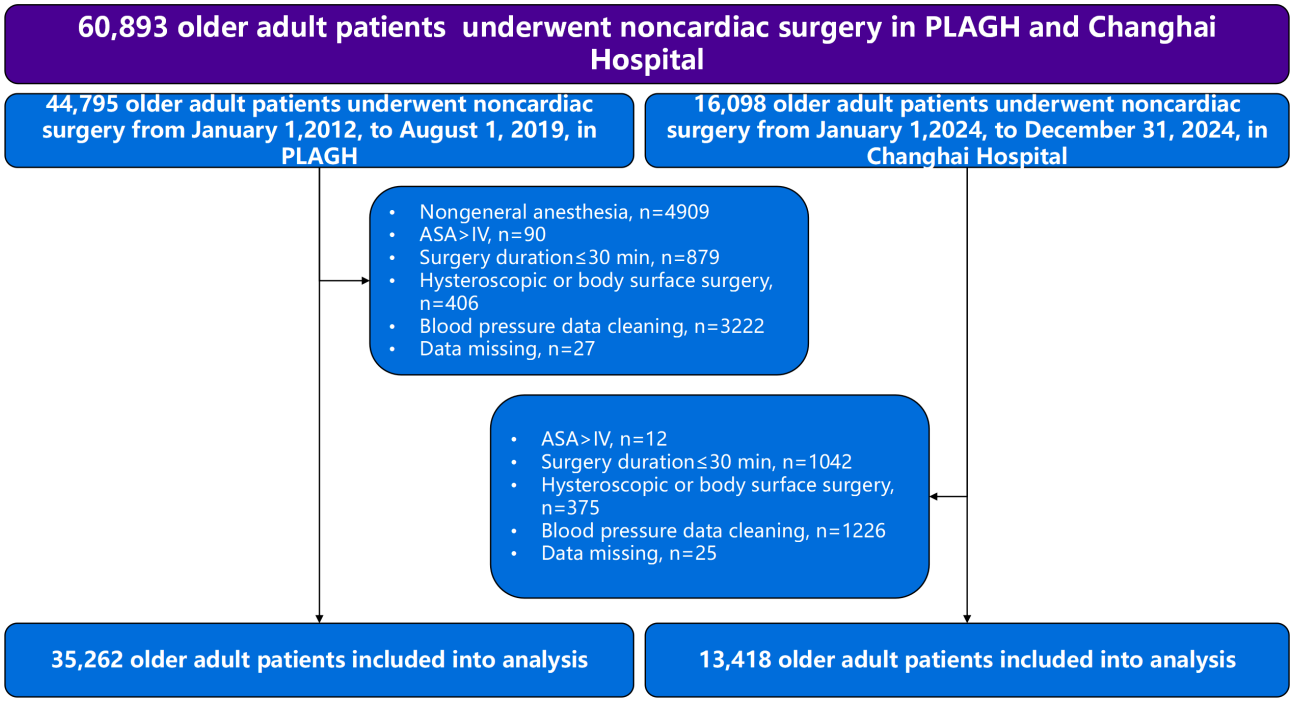
Results

Patient Characteristics and Confounding Variables in PLAGH

In total, 44,795 older adult patients (≥65 years old) undergoing noncardiac surgery were identified from January 1, 2012, to August 1, 2019. Of these patients, 35,262 were included in the study after application of the inclusion and exclusion criteria (Figure 1), and 874 (2.48%) patients developed MACE. Of these 874 patients, 101 (0.29%) had acute MI, 139 (0.39%) had HF, 391 (1.11%) had arrhythmia, 215 (0.61%) had angina, and 65 (0.184%) had nonfatal cardiac arrest or cardiac death. There were significant differences between patients with and without MACE in terms of the ASA grade, hypertension, arrhythmia, HF, coronary artery disease, valvular heart disease, peripheral artery disease, hemoglobin concentration, SCr concentration, fibrinogen concentration, statin therapy, antiplatelet drugs, diuretics, type of surgery, surgery duration, colloid use, blood loss, and urine volume. However, there was no difference in the baseline MAP between the 2 groups (Table 1).

XSL•FO
RenderX

Figure 1. Patient flowchart. ASA: American Society of Anesthesiologists; PLAGH: People's Liberation Army General Hospital.



XSL•FO
RenderX

Table . Patients' baseline and intraoperative characteristics in People's Liberation Army General Hospital.

Demographic characteristics, variable, and level	Non-MACE[a] (n=34,388)	MACE (n=874)	*P* value	SMD[b]
Sex, n (%)			.54	0.02
Female	16527 (48.1)	410 (46.9)		
Male	17861 (51.9)	464 (53.1)		
Age (y), mean (SD)	71 (5.11)	73 (6.11)	<.001	0.40
Age≥80 y, n (%)	2392 (7)	132 (15.1)	<.001	0.26
BMI[c] (kg/m^2), median (IQR[d])	24.4 (22.2-26.8)	24.3 (22.1-26.8)	.28	0.04
Smoking status, n (%)			.15	0.06
Nonsmoker	32249 (93.8)	806 (92.2)		
Current smoker	1079 (3.1)	36 (4.1)		
Former smoker	1060 (3.1)	32 (3.7)		
Drinking status, n (%)			.16	0.06
Nondrinker	27732 (80.6)	696 (79.6)		
Current drinker	4553 (13.2)	111 (12.7)		
Former drinker	2103 (6.1)	67 (7.7)		
Previous history, n (%)				
Hypertension	15243 (44.3)	472 (54)	<.001	0.20
Diabetes	7845 (22.8)	231 (26.4)	.01	0.08
Arrhythmia	3226 (9.4)	207 (23.7)	<.001	0.39
Congestive heart failure	66 (0.2)	44 (5)	<.001	0.31
Coronary heart disease	3802 (11.1)	219 (25.1)	<.001	0.37
Valvular heart disease	148 (0.4)	24 (2.8)	<.001	0.19
Peripheral artery disease	750 (2.2)	48 (5.5)	<.001	0.17
Cerebrovascular disease	2649 (7.7)	103 (11.8)	<.001	0.14
Renal insufficiency	423 (1.2)	43 (4.9)	<.001	0.22
Preoperative laboratory data				
Hgb[e] (g/L), median (IQR)	130 (119-141)	127 (111-139)	<.001	0.24
SCr[f] (umol/L), median (IQR)	70.5 (60.0-82.5)	72.4 (61.2-88.0)	<.001	0.28
FB[g] (g/L), median (IQR)	3.4 (2.9-4.0)	3.6 (3.0-4.5)	<.001	0.23
Preoperative medication, n (%)				
Statin	1862 (5.4)	89 (10.2)	<.001	0.18
Antiplatelet drugs	2886 (8.4)	108 (12.4)	<.001	0.13
β-blockers	2956 (8.6)	159 (18.2)	<.001	0.29
ACEI[h]	1306 (3.8)	51 (5.8)	.003	0.10
ARB[i]	3132 (9.1)	109 (12.5)	.001	0.10
Diuretic	2479 (7.2)	105 (12)	<.001	0.16
Surgery-related factors				
ASA[j], n (%)			<.001	0.49
I	579 (1.7)	7 (0.8)		

XSL•FO
RenderX

Demographic characteristics, variable, and level	Non-MACE[a] (n=34,388)	MACE (n=874)	*P* value	SMD[b]
II	26963 (78.4)	522 (59.7)		
III	6598 (19.2)	288 (33)		
IV	248 (0.7)	57 (6.5)		
Emergency surgery, n (%)	840 (2.4)	63 (7.2)	<.001	0.22
Type of surgery, n (%)			<.001	0.46
Ear and nose	1681 (4.9)	19 (2.2)		
Hepatobiliary	5885 (17.1)	139 (15.9)		
Orthopedics	9567 (27.8)	192 (22)		
Urology	3273 (9.5)	52 (6)		
General	6008 (17.5)	210 (24)		
Neurosurgery	1569 (4.6)	42 (4.8)		
Thoracic	2594 (7.5)	133 (15.2)		
Vascular	928 (2.7)	58 (6.6)		
Others	2883 (8.4)	29 (3.3)		
Surgery duration (min)			<.001	0.46
0-60	889 (2.6)	37 (4.2)		
60-180	17102 (49.7)	272 (31.1)		
180-240	7725 (22.5)	213 (24.4)		
240-infinity	8672 (25.2)	352 (40.3)		
Baseline MAP[k] (mm Hg), median (IQR)	96 (89-103)	95 (87-102)	.04	0.07
Crystalloid infusion rate (ml/kg·min), mean (SD)	0.05 (0.0)	0.051 (0.0)	<.001	0.16
Blood loss (mL), median (IQR)	100 (50-200)	200 (100-300)	<.001	0.21
Urine volume (mL), median (IQR)	300 (100-550)	400 (200-700)	<.001	0.25

[a]MACE: major adverse cardiac events.
[b]SMD: standard mean difference.
[c]BMI; body mass index
[d]IQR; interquartile range
[e]Hgb: hemoglobin.
[f]SCr: serum creatinine.
[g]FB: fibrinogen.
[h]ACEI: angiotensin-converting enzyme inhibitors.
[i]ARB: angiotensin receptor blockers.
[j]ASA: American Society of Anesthesiologists.
[k]MAP: mean arterial pressure.

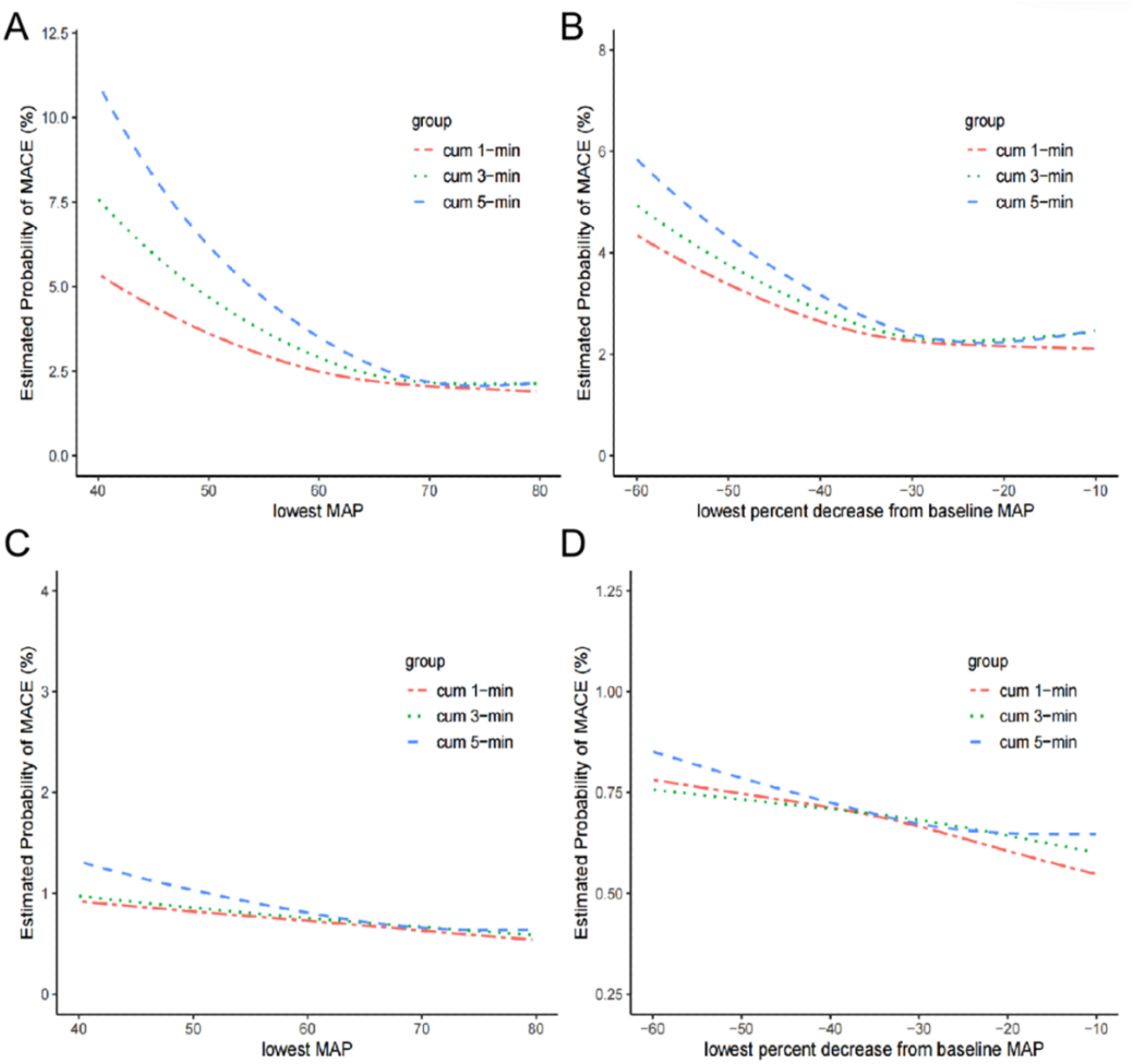
Determination of IOH Thresholds

As illustrated in Figure 2A and B, the analysis of univariate moving average smoothing curves indicates that the curves converge at a MAP of 70 mm Hg and a 30% reduction from baseline values across various cumulative time intervals. This finding suggests that an absolute MAP below 70 mm Hg or a relative decrease of 30% from baseline may represent a critical threshold for IOH in older adult patients. After adjusting for confounding factors—including age, cerebrovascular disease, coronary heart disease, valvular heart disease, HF, arrhythmia, hypertension, peripheral arterial disease, renal insufficiency, hemoglobin, SCr, fibrinogen, antiplatelet drugs, β-blockers, ACEI, ARB, diuretic, ASA score, surgery type, surgery duration, emergency surgery, blood loss, urine volume, crystalloid infusion—Figures 2C and 2D reveal that the curves maintain convergence at the 70 mm Hg mark but lose convergence at the 30% reduction point, thereby underscoring the robustness of

XSL•FO
RenderX

using an absolute MAP below 70 mm Hg as a threshold for IOH in older adult patients. To further validate this threshold's efficacy, the study used 3 widely adopted indicators of IOH exposure—time below the threshold, area under the curve (AUC) beneath the threshold, and TWA-MAP—subjecting them to multivariate logistic regression analysis. The adjusted model demonstrates a significant positive correlation between the duration of MAP below 70 mm Hg and the incidence of MACE. Specifically, when the duration of IOH exceeds 15 minutes, the risk of MACE substantially increases after adjusting for all confounding variables (adjusted odds ratio [OR] 1.51, 95% CI 1.22 - 1.88, P<.001). Notably, the AUC for the duration below the hypotensive threshold, measured in minutes×mm Hg, and the TWA-MAP also significantly affect MACE occurrence. As exposure intensity rises, the risk of MACE exhibits a graded increase, which further illustrates the dose-dependent effect of hypotension (Table 2).

Figure 2. Univariable and multivariable relationship between MACE and absolute and relative lowest MAP thresholds. (**A, B**) Estimated probability of MACE was determined from the univariable moving-average smoothing plots. (**C, D**) Estimated probability of MACE was determined from smoothed multivariable logistic regression. MACE: major adverse cardiac events; MAP: mean arterial pressure; cum: cumulative.



XSL•FO
RenderX

Table . Univariable and multivariable relationship between mean arterial pressure thresholds and outcomes.

Duration of intraoperative hypotension	OR[a] (95% CI)	Adjusted OR (95% CI)	*P* value
Time with MAP[b]<70 mm Hg (per min)	1.01 (1.00-1.01)	1.00 (1.00-1.01)	<.001
Time with MAP<70 mm Hg (min)			
0.5 - 5	1.23 (0.94-1.60)	1.26 (0.96-1.61)	.10
5 - 10	1.27 (0.97-1.65)	1.23 (0.94-1.62)	.13
10 - 15	1.10 (0.81-1.49)	1.06 (0.78-1.45)	.72
15	1.69 (1.38-2.07)	1.51 (1.22-1.88)	<.001
Area under curve<70 mm Hg (min×mm Hg)			
Q1	1.19 (0.93-1.52)	1.20 (0.93-1.54)	.16
Q2	1.28 (1.01-1.63)	1.22 (0.95-1.57)	.12
Q3	1.46 (1.15-1.85)	1.37 (1.07-1.75)	.01
Q4	2.09 (1.67-2.61)	1.82 (1.43-2.32)	<.001
TWA[c]-MAP<70 mm Hg (mm Hg)			
Q1	1.35 (1.06-1.71)	1.30 (1.01-1.66)	.04
Q2	1.36 (1.07-1.73)	1.29 (1.01-1.66)	.04
Q3	1.48 (1.17-1.87)	1.42 (1.11-1.82)	.005
Q4	1.83 (1.46-2.29)	1.56 (1.22-1.99)	<.001

[a]OR: odds ratio.

[b]MAP: mean arterial pressure.

[c]TWA: time-weighted average.

As illustrated in Table 3, among patients aged <80 years, compared with intraoperative MAP less than 70 mm Hg durations ≤5 minutes, the ORs of intraoperative MAP less than 70 mm Hg durations of 5 - 10 minutes, 10 - 15 minutes, and >15 minutes were 1.11 (95% CI 0.85 - 1.44), 0.86 (95% CI 0.62 - 1.17), and 1.31 (95% CI 1.09 - 1.59), respectively. Among patients aged ≥80 years, the ORs of intraoperative MAP less than 70 mm Hg durations of 5 - 10 minutes, 10 - 15 minutes, and >15 minutes were 1.27 (95% CI 0.66 - 2.54), 1.28 (95% CI 0.58 - 2.64), and 1.38 (95% CI 0.86 - 2.28), respectively.

Table . Results of stratification analysis.

Stratifications	Time with MAP[a]<70 mm Hg (min)	OR[b] (95% CI)	*P* value
Age <80	0.5 - 5	Reference	Reference
Age <80	5 - 10	1.11 (0.85 - 1.44)	.43
Age <80	10 - 15	0.86 (0.62 - 1.17)	.34
Age <80	15	1.31 (1.09 - 1.59)	<.01
Age ≥80	0.5 - 5	Reference	Reference
Age ≥80	5 - 10	1.27 (0.66 - 2.54)	.42
Age ≥80	10 - 15	1.28 (0.58 - 2.64)	.52
Age ≥80	15	1.38 (0.86 - 2.28)	.19

[a]MAP: mean arterial pressure.

[b]OR: odds ratio.

Validation of the IOH Threshold for Older Adult Patients in Shanghai Changhai Hospital

A total of 13,418 older adult surgical patients were included at Changhai Hospital. Among them, 296 patients (2.2%) experienced MACE (Figure 1). The baseline characteristics of the patients are presented in Table S1 in Multimedia Appendix 1. Univariate logistic regression analysis revealed a significant association between IOH and MACE (OR 1.01, 95% CI 1.00 - 1.01, *P*<.001). Following adjustment for confounding

XSL•FO
RenderX

factors, including age, sex, history of cerebrovascular disease, history of coronary heart disease, history of valvular heart disease, history of HF, history of arrhythmia, history of hypertension, history of peripheral arterial disease, history of renal insufficiency, SCr, preoperative hemoglobin, preoperative fibrinogen, ASA grades, surgery duration, emergency surgery status, intraoperative blood loss, and intraoperative crystalloid infusion volume, the association between IOH and MACE remained robust (adjusted OR 1.01, 95% CI 1.00 - 1.01, P<.001). Furthermore, we found that IOH lasting ≥15 minutes was associated with a significantly increased risk of MACE (adjusted OR 2.01, 95% CI 1.55 - 2.59, P<.001; Table S2 in Multimedia Appendix 1). These results are consistent with findings from the PLAGH study.

Discussion

Principal Findings

In this study, we investigated and validated the correlation between IOH and MACE in older adult patients undergoing noncardiac surgery from PLAGH and Shanghai Changhai Hospital. The findings indicate that the risk of MACE in older adult patients increases as MAP decreases during noncardiac surgery. Specifically, when the lowest intraoperative MAP falls below 70 mm Hg, there is a significant increase in the risk of MACE. Using an intraoperative MAP of 70 mm Hg as the reference threshold for IOH, the risk of MACE demonstrates a dose-dependent increase. In addition, if the duration of intraoperative MAP below 70 mm Hg exceeds 15 minutes, the risk of MACE increases above 50%. Consequently, using a MAP of 70 mm Hg as the reference threshold for IOH could offer more valuable guidance for older adult patients.

The relationship between IOH and postoperative complications has been extensively investigated by researchers [9,21-26]. Previous studies have consistently demonstrated an association between the severity of IOH and postoperative complications [8,27-30]. This severity encompasses not only the extent of the reduction but also the duration of hypotension and the area under the hypotension threshold. In line with previous studies, the results of our study demonstrate a strong correlation between the severity of IOH, as measured by time, AUC, or TWA-MAP, using a reference threshold of MAP 70 mm Hg, and the occurrence of MACE. Notably, our study also found that when MAP falls below 70 mm Hg for more than 15 minutes during surgery, the risk of MACE increases significantly. This finding may have important clinical implications. In liver resection surgery, where hepatic inflow occlusion is often necessary, the determination of an appropriate occlusion time is crucial. Our observation that MAP below 70 mm Hg for more than 15 minutes is associated with a 50% increased risk of MACE can serve as a valuable reference point in this context.

However, our study differs from previous research in terms of the currently accepted threshold for IOH. While the most used IOH threshold is a MAP of 65 mm Hg, our study suggests that a higher threshold of 70 mm Hg may be more suitable for older adult surgical patients. This discrepancy may be attributed to the fact that previous studies encompassed a general population, including patients aged 18 years and older or 45 years and older, whereas our study specifically focused on individuals aged 65 years and older. As mentioned earlier, the cardiovascular system of older adult patients is often delicately balanced, and even minor fluctuations in hemodynamics can have irreversible and severe consequences.

It is important to note the physiological differences in the very older adult population (aged 80 years and older). Within this cohort, blood pressure variability tends to be greater, necessitating more cautious monitoring and management during surgery. Although our study accounted for various confounding factors that may affect blood pressure fluctuations and the occurrence of postoperative MACE, the limitations inherent to retrospective studies prevented the adjustment for all potential confounders. As a result, our analyses did not yield sufficient statistical significance in this very old group. The specific factors contributing to the lack of significant findings could include the unique cardiovascular responses and decreased physiologic reserves that characterize this demographic. Our results underscore the need for future studies to refine the thresholds for IOH in older patients, as traditional metrics may not adequately reflect the risks faced by this vulnerable population.

This study has 2 main limitations. First, the occurrence of MACE was lower than that in other research, and this was primarily attributed to the limitations inherent in retrospective studies. The reliance solely on patients' medical records for the identification of postoperative MACE may have led to underrecognition of specific cases of acute MI because of the absence of postoperative troponin monitoring. Second, although this study showed that a 25% decrease from baseline MAP was associated with MACE, no significant association was found in the subsequent validation analyses. This was also probably due to the limitations inherent in retrospective studies.

Conclusion

For older adult patients undergoing noncardiac surgery, intraoperative MAP should be kept above 70 mm Hg to reduce the risk of postoperative MACE. However, for patients older than 80 years, further research is needed to explore the accurate threshold for IOH.

Acknowledgments

We thank Professor Lan Sun, Professor Longtao Wu, Professor Wei Wei, Professor Linlin Xu, and Lei Mo of Shanghai Lejiu Healthcare Technology for their technical assistance in the data extraction and statistical analyses. We also thank Angela Morben, DVM, ELS, from Liwen Bianji (Edanz) for editing the English text of a draft of this manuscript.

This research was supported by the National Key Research and Development Program of China (2018YFC2001900) and the Beijing Nova Program (Z211100002121171). The authors declare no competing financial interests.

XSL•FO
RenderX

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' Contributions

Conception and design of study: KZ
Conceptualization: HL
Data curation: MW, TZ, BM, JL, QF, SY, LB
Formal analysis: KZ
Funding acquisition: WM
Methodology: KZ
Project administration: WM
Resources: LB, WM, HL
Software: KZ
Supervision: YL, JC
Writing-original draft: KZ, CL
Writing-review & editing: WM, HL, LB

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary material about Changhai Hospital.
[DOCX File, 21 KB - aging_v8i1e67177_app1.docx]

References

1. Puelacher C, Bollen Pinto B, Mills NL, et al. Expert consensus on peri-operative myocardial injury screening in noncardiac surgery: a literature review. Eur J Anaesthesiol 2021 Jun 1;38(6):600-608. [doi: 10.1097/EJA.0000000000001486] [Medline: 33653981]
2. Wang H, Chen H. Aging in China: challenges and opportunities. China CDC Wkly 2022 Jul 8;4(27):601-602. [doi: 10.46234/ccdcw2022.130] [Medline: 35919296]
3. Botto F, Alonso-Coello P, Chan MTV, et al. Myocardial injury after noncardiac surgery: a large, international, prospective cohort study establishing diagnostic criteria, characteristics, predictors, and 30-day outcomes. Anesthesiology 2014 Mar;120(3):564-578. [doi: 10.1097/ALN.0000000000000113] [Medline: 24534856]
4. Bartels K, Karhausen J, Clambey ET, Grenz A, Eltzschig HK. Perioperative organ injury. Anesthesiology 2013 Dec;119(6):1474-1489. [doi: 10.1097/ALN.0000000000000022] [Medline: 24126264]
5. Kazaure HS, Roman SA, Rosenthal RA, Sosa JA. Cardiac arrest among surgical patients: an analysis of incidence, patient characteristics, and outcomes in ACS-NSQIP. JAMA Surg 2013 Jan;148(1):14-21. [doi: 10.1001/jamasurg.2013.671] [Medline: 23324834]
6. Abbott TEF, Pearse RM, Archbold RA, et al. A prospective international multicentre cohort study of intraoperative heart rate and systolic blood pressure and myocardial injury after noncardiac surgery: results of the VISION study. Anesth Analg 2018 Jun;126(6):1936-1945. [doi: 10.1213/ANE.0000000000002560] [Medline: 29077608]
7. Gregory A, Stapelfeldt WH, Khanna AK, et al. Intraoperative hypotension is associated with adverse clinical outcomes after noncardiac surgery. Anesth Analg 2021 Jun 1;132(6):1654-1665. [doi: 10.1213/ANE.0000000000005250] [Medline: 33177322]
8. Hallqvist L, Granath F, Fored M, Bell M. Intraoperative hypotension and myocardial infarction development among high-risk patients undergoing noncardiac surgery: a nested case-control study. Anesth Analg 2021 Jul 1;133(1):6-15. [doi: 10.1213/ANE.0000000000005391] [Medline: 33555690]
9. Salmasi V, Maheshwari K, Yang D, et al. Relationship between intraoperative hypotension, defined by either reduction from baseline or absolute thresholds, and acute kidney and myocardial injury after noncardiac surgery: a retrospective cohort analysis. Anesthesiology 2017 Jan;126(1):47-65. [doi: 10.1097/ALN.0000000000001432] [Medline: 27792044]
10. Wesselink EM, Wagemakers SH, van Waes JAR, Wanderer JP, van Klei WA, Kappen TH. Associations between intraoperative hypotension, duration of surgery and postoperative myocardial injury after noncardiac surgery: a retrospective single-centre cohort study. Br J Anaesth 2022 Oct;129(4):487-496. [doi: 10.1016/j.bja.2022.06.034]
11. Hatib F, Jian Z, Buddi S, et al. Machine-learning algorithm to predict hypotension based on high-fidelity arterial pressure waveform analysis. Anesthesiology 2018 Oct;129(4):663-674. [doi: 10.1097/ALN.0000000000002300] [Medline: 29894315]
12. Wijnberge M, Geerts BF, Hol L, et al. Effect of a machine learning-derived early warning system for intraoperative hypotension vs standard care on depth and duration of intraoperative hypotension during elective noncardiac surgery: the

XSL•FO
RenderX

HYPE randomized clinical trial. JAMA 2020 Mar 17;323(11):1052-1060. [doi: 10.1001/jama.2020.0592] [Medline: 32065827]
13. Runge J, Graw J, Grundmann CD, Komanek T, Wischermann JM, Frey UH. Hypotension prediction index and incidence of perioperative hypotension: a single-center propensity-score-matched analysis. J Clin Med 2023 Aug 23;12(17):5479. [doi: 10.3390/jcm12175479] [Medline: 37685546]
14. Yoshikawa Y, Maeda M, Kunigo T, et al. Effect of using hypotension prediction index versus conventional goal-directed haemodynamic management to reduce intraoperative hypotension in non-cardiac surgery: a randomised controlled trial. J Clin Anesth 2024 May;93:111348. [doi: 10.1016/j.jclinane.2023.111348] [Medline: 38039629]
15. Halvorsen S, Mehilli J, Cassese S, et al. 2022 ESC Guidelines on cardiovascular assessment and management of patients undergoing non-cardiac surgery. Eur Heart J 2022 Oct 14;43(39):3826-3924. [doi: 10.1093/eurheartj/ehac270] [Medline: 36017553]
16. Jammer I, Wickboldt N, Sander M, et al. Standards for definitions and use of outcome measures for clinical effectiveness research in perioperative medicine: European Perioperative Clinical Outcome (EPCO) definitions: a statement from the ESA-ESICM joint taskforce on perioperative outcome measures. Eur J Anaesthesiol 2015 Feb;32(2):88-105. [doi: 10.1097/EJA.0000000000000118] [Medline: 25058504]
17. Sabaté S, Mases A, Guilera N, et al. Incidence and predictors of major perioperative adverse cardiac and cerebrovascular events in non-cardiac surgery. Br J Anaesth 2011 Dec;107(6):879-890. [doi: 10.1093/bja/aer268] [Medline: 21890661]
18. Alpert JS, Thygesen K, Antman E, Bassand JP. Myocardial infarction redefined--a consensus document of The Joint European Society of Cardiology/American College of Cardiology Committee for the redefinition of myocardial infarction. J Am Coll Cardiol 2000 Sep;36(3):959-969. [doi: 10.1016/s0735-1097(00)00804-4] [Medline: 10987628]
19. Devereaux PJ, Sessler DI. Cardiac complications in patients undergoing major noncardiac surgery. N Engl J Med 2015 Dec 3;373(23):2258-2269. [doi: 10.1056/NEJMra1502824] [Medline: 26630144]
20. Wachtendorf LJ, Azimaraghi O, Santer P, et al. Association between intraoperative arterial hypotension and postoperative delirium after noncardiac surgery: a retrospective multicenter cohort study. Anesth Analg 2022 Apr 1;134(4):822-833. [doi: 10.1213/ANE.0000000000005739] [Medline: 34517389]
21. Sessler DI, Short TG. Intraoperative hypotension and complications. J Am Coll Cardiol 2021 Nov 2;78(18):1765-1767. [doi: 10.1016/j.jacc.2021.08.049] [Medline: 34711334]
22. Sessler DI, Khanna AK. Perioperative myocardial injury and the contribution of hypotension. Intensive Care Med 2018 Jun;44(6):811-822. [doi: 10.1007/s00134-018-5224-7]
23. Walsh M, Devereaux PJ, Garg AX, et al. Relationship between intraoperative mean arterial pressure and clinical outcomes after noncardiac surgery: toward an empirical definition of hypotension. Anesthesiology 2013 Sep;119(3):507-515. [doi: 10.1097/ALN.0b013e3182a10e26] [Medline: 23835589]
24. Mascha EJ, Yang D, Weiss S, Sessler DI. Intraoperative mean arterial pressure variability and 30-day mortality in patients having noncardiac surgery. Anesthesiology 2015 Jul;123(1):79-91. [doi: 10.1097/ALN.0000000000000686] [Medline: 25929547]
25. Müller-Wirtz LM, Ruetzler K, Rössler J. Intraoperative hypotension and delirium. J Clin Anesth 2023 Oct;89:111153. [doi: 10.1016/j.jclinane.2023.111153]
26. van Zuylen ML, Gribnau A, Admiraal M, et al. The role of intraoperative hypotension on the development of postoperative cognitive dysfunction: a systematic review. J Clin Anesth 2021 Sep;72:110310. [doi: 10.1016/j.jclinane.2021.110310]
27. Bijker JB, Persoon S, Peelen LM, et al. Intraoperative hypotension and perioperative ischemic stroke after general surgery: a nested case-control study. Anesthesiology 2012 Mar;116(3):658-664. [doi: 10.1097/ALN.0b013e3182472320] [Medline: 22277949]
28. Wang J, Lin F, Zeng M, et al. Intraoperative blood pressure and cardiac complications after aneurysmal subarachnoid hemorrhage: a retrospective cohort study. Int J Surg 2024;110(2):965-973. [doi: 10.1097/JS9.0000000000000928]
29. Sun LY, Chung AM, Farkouh ME, et al. Defining an intraoperative hypotension threshold in association with stroke in cardiac surgery. Anesthesiology 2018 Sep;129(3):440-447. [doi: 10.1097/ALN.0000000000002298] [Medline: 29889106]
30. Etemadi M, Hogue CW. Preventing intraoperative hypotension. Anesthesiology 2020 Dec 1;133(6):1170-1172. [doi: 10.1097/ALN.0000000000003561] [Medline: 34352077]

Abbreviations

ACEI: angiotensin-converting enzyme inhibitor
ARB: angiotensin receptor blocker
ASA: American Society of Anesthesiologists
AUC: area under the curve
ECG: electrocardiogram
HF: heart failure
IOH: intraoperative hypotension
MACE: major adverse cardiac event

XSL•FO
RenderX

MAP: mean arterial pressure
MI: myocardial infarction
OR: odds ratio
PLAGH: People's Liberation Army General Hospital
SCr: serum creatinine
TWA-MAP: time-weighted average mean arterial pressure

Edited by M Mardini; submitted 04.10.24; peer-reviewed by D Stokanovic, K Bhatt, PS Pillai; revised version received 24.06.25; accepted 22.07.25; published 16.10.25.

Please cite as:
Zhang K, Liu C, Wang M, Zhang T, Meng B, Yao S, Lou J, Fu Q, Liu Y, Cao J, Bo L, Mi W, Li H
Intraoperative Hypotension and Major Adverse Cardiac Events Among Older Adult Patients Undergoing Noncardiac Surgery: Retrospective Cohort Study
JMIR Aging 2025;8:e67177
URL: https://aging.jmir.org/2025/1/e67177
doi:10.2196/67177

© Kai Zhang, Chang Liu, Meng Wang, Ting Zhang, Bingbing Meng, Siyi Yao, Jingsheng Lou, Qiang Fu, Yanhong Liu, Jiangbei Cao, Lulong Bo, Weidong Mi, Hao Li. Originally published in JMIR Aging (https://aging.jmir.org), 16.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Sleep Disturbance as a Catalyst in the Cyclical Link Between Depressive Symptoms and Disability in Instrumental Activities of Daily Living in Older Chinese Adults: Longitudinal Cohort Study

Hao Wu[1,2], MPH; Lu Wang[1,2], MPH; Xin Liu[1,2], BMed; Yuyue Yang[1,2], BMed; Wei Gao[1,2], PhD

[1]Department of Epidemiology and Health Statistics, School of Public Health, Jiangxi Medical College, Nanchang University, No. 1299 Xuefu Avenue, Nanchang, China

[2]Jiangxi Provincial Key Laboratory of Disease Prevention and Public Health, Nanchang University, Nanchang, China

Corresponding Author:
Wei Gao, PhD
Department of Epidemiology and Health Statistics, School of Public Health, Jiangxi Medical College, Nanchang University, No. 1299 Xuefu Avenue, Nanchang, China

Abstract

Background: Depressive symptoms, sleep disturbances, and functional disability are interrelated. However, the bidirectional pathways between depression, sleep disturbances, and disability in instrumental activities of daily living (IADLs) remain underexplored in China.

Objective: We aimed to examine the bidirectional longitudinal relationships between depression and disability in IADLs among older Chinese adults, with a focus on elucidating the mediating role of sleep disturbances in this dynamic interplay.

Methods: The study encompassed 2677 older adults who provided complete data at T1 (2015), T2 (2018), and T3 (2020) for the China Health and Retirement Longitudinal Study (CHARLS). Depressive symptoms were assessed using the 10-item Center for Epidemiological Studies Depression (CESD-10) scale, and a 6-item scale was used to measure disability in IADLs. Sleep disturbances were self-reported. Temporal associations between depressive symptoms and disability in IADLs as well as the longitudinal mediating effect of sleep disturbances were examined using a cross-lagged panel model.

Results: Prior depression significantly predicted subsequent disability in IADLs at T2 (β=0.070, *P*<.001) and T3 (β=0.074, *P*<.001), and prior disability in IADL predicted subsequent depression at T2 (β=0.094, *P*<.001) and T3 (β=0.100, *P*<.001). Additionally, the indirect effect of prior disability in IADLs on subsequent depression via sleep disturbances was statistically significant (β=0.062, SE=0.010, *P*<.001), with the mediation effect accounting for 50.41% of the total effect. In contrast, after accounting for this mediation, the direct effect of prior depression on subsequent disability in IADLs was not significant (β=0.009, SE=0.018, *P*=.61). Consequently, the impact of depression on disability in IADLs was fully mediated through sleep disturbances in this cohort of older Chinese adults.

Conclusions: Depressive symptoms and disability in IADLs are bidirectionally linked, and sleep disturbances play a longitudinal mediating role in the bidirectional relationship among older Chinese adults. The potential longitudinal bidirectionality highlights the importance of sleep health for interventions on depression and functional disability in older adults.

(JMIR Aging 2025;8:e76643) doi:10.2196/76643

KEYWORDS

sleep disturbance; depression; disability in IADLs; cross-lagged panel model; longitudinal mediation; older adults

Introduction

According to the World Health Organization, the number of people aged 60 years and older outnumbered children younger than 5 years in 2020. Between 2015 and 2050, the proportion of the world's population older than 60 years will nearly double from 12% to 22% [1]. China's accelerated demographic transition since achieving aging society status 2 decades ago coincides with growing global health challenges, including cognitive impairment and physical frailty.

Aging, as an inescapable natural process, progressively influences our functional capacities, mental health, a growing risk of disease, and social interactions. Functional disability affects more than 1 billion people worldwide according to the 74th World Health Assembly, and this number is increasing as the aging population increases [2]. Activities of daily living (ADLs) and instrumental activities of daily living (IADLs) are key measures, with ADLs covering basic self-care and IADLs involving more complex tasks necessary for independent living and social participation [3,4]. Although there is some debate as to whether ADLs or IADLs reflect separate constructs, most

XSL•FO
RenderX

conceptualizations assume a hierarchical relation between the 2 constructs, with disability in IADLs preceding disability in ADLs [5]. Therefore, impairments of IADLs may indicate an early decline in physical function [6]. The mental health of older adults, particularly depression, demands urgent attention. With approximately 4.7% of the global population experiencing a depressive episode annually [7] and rates in Chinese people older than 45 years reaching 26.67% for men and 38.37% for women [8], the rising prevalence of depressive symptoms poses a significant public health challenge. Late-life depression, linked to physical illness, functional impairment, and mortality [9], underscores the need for identifying modifiable risk factors. Sleep, a potentially modifiable lifestyle behavior, plays a very important role in health and well-being. Nearly one-half of the world's population experience sleep disturbances and an estimated one-third of adults have experienced insomnia symptoms [10]. Further, older adults may under-report the severity of sleep disturbances compared with objective assessments by polysomnography [11]. Both sleep disturbances and depression are associated with a greater risk of adverse health-related outcomes, including poor health status [12], disability [13], and poorer physical functioning in older populations [14].

Currently, multiple cross-sectional and longitudinal studies have confirmed the pairwise bidirectional associations between sleep, depressive symptoms, and functional disability in older adults globally [15-17]. A study with older adults aged 60 years and older in 3 provinces in China showed that the effect of limitations in IADLs on depressive symptoms in older adults was greater than that of limitations in ADLs [18]. Additionally, a bidirectional association between sleep disturbances and depression was found in older adults [19,20], in which sleep disturbances might exacerbate depression, and older adults with depression often experience poor sleep quality [21,22]. Some studies found that short [14] or long [23] sleep duration was associated with disability in IADLs, whereas others showed that both sleep durations were related to disability in IADLs in older people [24]. Older adults with more limitations in ADLs and IADLs had a higher risk of experiencing declines in sleep quality and the transition from meeting to not meeting the recommended sleep duration over time [25].

Previous studies have indicated that depression can impair IADLs by reducing motivation and exacerbating fatigue, while declines in IADLs may in turn increase the risk of depression [15,16]. Therefore, we hypothesized a bidirectional relationship between depression and disability in IADLs. Furthermore, extensive studies have firmly established significant bidirectional relationships between sleep and depression as well as between sleep and functional disability [20-23,25]. Given that sleep disturbances are particularly prevalent among older adults and could act as both a consequence of depression and a precursor to functional disability, we further proposed that sleep plays a mediating role in the relationship between depression and functional disability. Based on substantial literature, we also suggested that gender differences might lead to distinct stress coping mechanisms and health behavior patterns [26], regional disparities reflect differences in health care accessibility and lifestyle [27], and empty nest status—as an important psychosocial factor—might influence physical and mental health outcomes among older adults [28]. Additionally, our research, using a conceptual framework, had found that sleep significantly mediated this interaction, acting as both a consequence and an active contributor [29]. However, despite these conceptual advances, a critical gap remains. Few longitudinal studies have specifically tested the mediating role of sleep in the bidirectional relationship between depression and functional disability among older adults in China. The Chinese context, characterized by its unique sociocultural dynamics and rapidly aging population, may shape these pathways in distinct ways. Furthermore, the changing morbidity patterns in China warrant a specific focus on IADL-related issues [30]. To address this gap, our study aimed to elucidate this complex interplay by using a cross-lagged panel model (CLPM) and testing 3 main hypotheses: (1) A bidirectional relationship exists between depression and disability in IADLs; (2) sleep disturbances longitudinally mediate this relationship; and (3) the strength of these associations and the mediating effect vary by gender, district, and empty nester status.

Methods

Data Collection and Study Participants

This study used data from the China Health and Retirement Longitudinal Study (CHARLS), a national longitudinal study of adults aged ≥45 years across 28 provinces in China. Initially conducted in 2011, it included 5 waves of data collection as of publication of this study. The CHARLS was approved by the Biomedical Ethics Review Committee of Peking University [31], and informed consent was obtained from all participants. Detailed information on study sampling, including the inclusion and exclusion criteria, can be found in earlier publications [31]. For this study, data from 2015, 2018, and 2020 were used, focusing on individuals aged ≥60 years without cognitive impairments (score<1.5 SD [32]) or missing data on sleep, depression, or disability in IADLs. Ultimately, a total of 2677 participants were included in this study (Multimedia Appendix 1). Reporting follows the Reporting of Studies Conducted Using Observational Routinely-Collected Health Data (RECORD) guidelines (Checklist 1).

Measurements

Depressive symptoms were assessed using the 10-item Center for Epidemiologic Studies Depression (CESD-10) scale, with scores ranging from 0 to 30. Participants were defined as having a higher sensitivity for clinically significant depressive symptoms if they scored 12 or higher on the CESD-10 scale [33]. IADLs, as assessed using the Lawton IADL Scale, represent the abilities required for community living that depend on advanced cognitive functioning [6]. The final score was a sum of responses to the 6 items and ranged from 0 to 18 points: the higher the score, the higher the level of disability in IADLs for the individual. Sleep disturbances were assessed using a single-item sleep quality scale and nighttime sleep duration, which were both commonly used in previous studies [34]. The total score for sleep disturbances was calculated as the sum of the scores for the single-item sleep quality scale and nighttime sleep duration, ranging from 0 to 6. It is important to note that

XSL•FO
RenderX

this composite measure captured subjective sleep disturbances and was not equivalent to a clinically diagnosed sleep disorder. Although this approach had not been formally validated against established instruments such as the Pittsburgh Sleep Quality Index or objective measures (actigraphy), similar methods were used in prior research using the CHARLS data [34]. Detailed information is in Multimedia Appendix 1.

Covariates

To minimize the potential confounding variables influencing the relationship between depressive symptoms, sleep disturbances, and disability in IADLs and to keep the model simplicity, limited covariates were adjusted. Covariates were selected based on the established theoretical and empirical associations with depressive symptoms, sleep disturbances, and functional disability in the existing literature on aging populations [6,15-17,22,23,25]. The covariates consisted of sociodemographic characteristics (age, gender, marital status, education, empty nester status, district, and working status), lifestyle behaviors (smoking and drinking status), cognitive function, and the number of noncommunicable diseases. Sociodemographic factors and lifestyle behaviors are fundamental determinants of health outcomes. The number of noncommunicable diseases was included as a proxy for overall physical health burden, a key potential confounder. Cognitive function was adjusted for due to its close linkages with both depression and the ability to perform IADLs [14,24]. Detailed information is in Multimedia Appendix 1.

Statistical Analysis

Demographic characteristics of the participants were analyzed and presented as frequencies and proportions for categorical variables and means and SDs for continuous variables. To study the causal relationships between depressive symptoms, disability in IADLs, and sleep disturbances over time, a CLPM with maximum likelihood estimation was conducted [35]. The CLPM with observed variables evaluated the association between depressive symptoms and disability in IADLs at all time points and the mediating role of sleep disturbances, with 4 models constructed with varying constraints on parameters. We tested 4 models: Model 1a-2a, an unconstrained CLPM adjusted for covariates; Model 1b-2b, with constrained cross-lagged parameters; Model 1c-2c, with constrained autoregressive parameters; and Model 1d-2d, with both autoregressive and cross-lagged parameters constrained. Model fit was evaluated using the root mean square error of approximation (RMSEA), comparative fit index (CFI), standardized root mean square residual (SRMR), and Tucker-Lewis index (TLI), with specific cutoffs for good fit (Multimedia Appendix 1). A potential conceptual overlap was noted between item 7 of the CESD-10 scale (“My sleep was restless”) and our measure of sleep disturbance. To assess whether this overlap inflated the mediation effects, a sensitivity analysis was performed. All CLPMs were reestimated using a modified CESD-10 score from which item 7 was excluded. The results of this sensitivity analysis are reported in Multimedia Appendices 2 and 3.

Descriptive statistic analysis was conducted using R software (Version 4.2.2, R Foundation for Statistical Computing). The CLPMs were constructed using Mplus software (Version 8.3, Muthén & Muthén). A significance level of $P<.05$ was used to indicate statistical significance of differences.

Ethical Considerations

Our study used publicly available anonymized data from the CHARLS and did not require further review. Ethical approval for all the CHARLS waves was granted from the Institutional Review Board at Peking University (IRB00001052-11015). The original ethical approval and participant consent include the use of data. Participants gave informed consent to participate in the study before taking part.

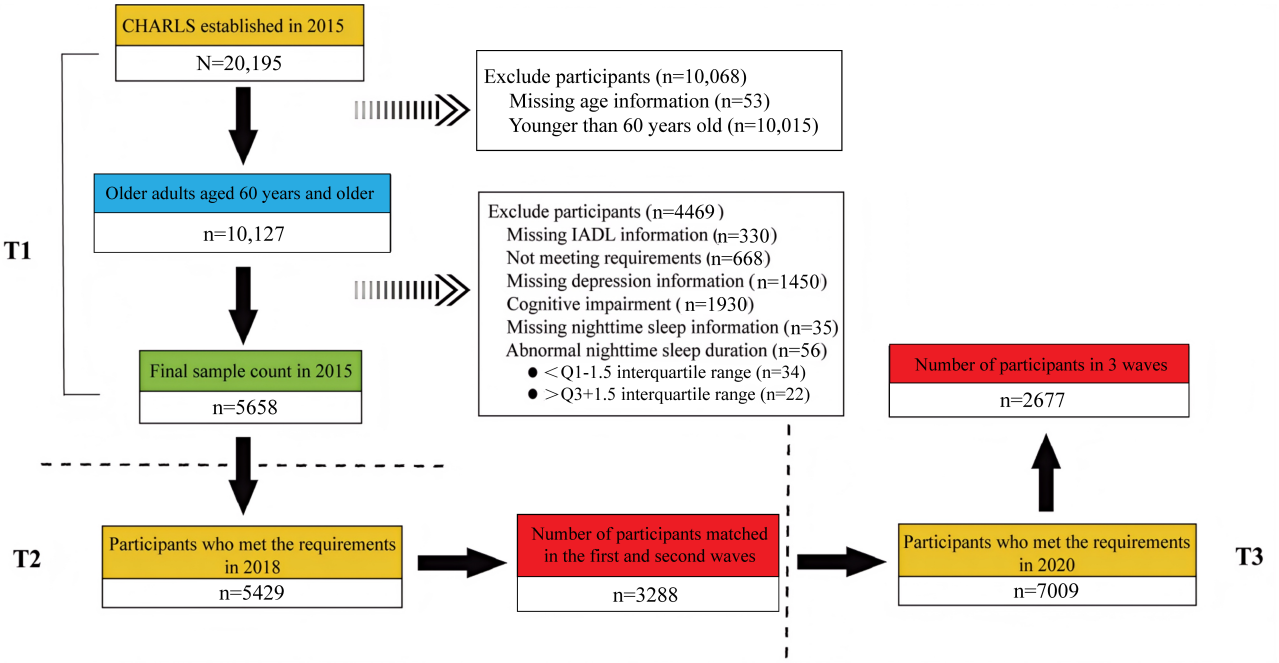
Results

Descriptive Statistics

A total of 2677 older adults aged 60 years and older participated in this study and completed data collection from 2015 to 2020 (Figure 1). The sample characteristics based on baseline data and the distribution of key variables by period are presented in Table 1. The majority of participants resided in villages (1639/2677, 61.23%) rather than in urban areas (1038/2677, 38.77%), and most were empty nesters (1690/2677, 63.13%). Among the participants, 61.75% (1653/2677) had completed at most a primary education. From T1 to T3, there was a slight decline in the proportion of married older adults, by 5.46%. The number of older adults who smoked was slightly higher than that of nonsmokers (T1: 1395/2677, 52.11% vs 1282/2677, 47.89%; T2: 1424/2677, 53.19% vs 1253/2677, 46.81%; T3: 1417/2677, 52.93% vs 1260/2677, 47.07%), and the majority were nondrinkers (T1: 1590/2677, 59.39%; T2: 1600/2677, 59.77%; T3: 1577/2677, 58.91%). As age increased, the prevalences of depressive symptoms, sleep disturbances, and disability in IADL among older adults also increased.

XSL•FO
RenderX

Figure 1. Flowchart of participant selection for the study. CHARLS: China Health and Retirement Longitudinal Study; IADL: instrumental activity of daily living; T1: 2015; T2: 2018; T3: 2020.



XSL•FO
RenderX

Table . Demographic characteristics and changes in depressive symptoms, sleep disturbance, and disability in instrumental activities of daily living (IADLs) from T1 to T3 (N=2677).

Characteristic	2015 (T1)	2018 (T2)	2020 (T3)
Age (years), mean (SD)	65.78 (4.84)	—[a]	—
Gender, n (%)			
Male	1630 (60.89)	—	—
Female	1047 (39.11)	—	—
Empty nester status, n (%)			
Empty nesters	1690 (63.13)	—	—
Not empty nesters	987 (36.87)	—	—
District, n (%)			
Village	1639 (61.23)	—	—
Urban	1038 (38.77)	—	—
Education, n (%)			
Illiteracy	781 (29.17)	—	—
Primary education or less	872 (32.57)	—	—
Junior high school	646 (24.13)	—	—
High school or more	378 (14.13)	—	—
Marital status, n (%)			
Married	2379 (88.87)	2297 (85.81)	2233 (83.41)
Not in marriage	298 (11.13)	380 (14.19)	444 (16.59)
Working status, n (%)			
Government/institutions/firm	147 (5.49)	128 (4.78)	144 (5.38)
Self-employed individual	1207 (45.09)	1179 (44.04)	208 (7.77)
Farmer	161 (6.01)	57 (2.13)	1020 (38.10)
Others	1162 (43.41)	1313 (49.05)	1305 (48.75)
Smoking status, n (%)			
Yes	1395 (52.11)	1424 (53.19)	1417 (52.93)
No	1282 (47.89)	1253 (46.81)	1260 (47.07)
Drinking status, n (%)			
More than once a month	842 (31.45)	838 (31.30)	821 (30.67)
Less than once a month	245 (9.16)	239 (8.93)	279 (10.42)
None of these	1590 (59.39)	1600 (59.77)	1577 (58.91)
Depressive symptoms (score), mean (SD)	8.22 (4.58)	8.85 (4.79)	9.25 (5.25)
Disability in IADLs (score), mean (SD)	0.55 (1.54)	0.68 (1.80)	0.63 (1.75)
Sleep disturbances (score), mean (SD)	1.95 (1.88)	2.13 (1.92)	2.19 (1.93)

[a]Not applicable.

The Reciprocal Relationship Between Depressive Symptoms and Disability in IADLs

Figure 2 depicts the CLPM of the bidirectional relationship between depressive symptoms and disability in IADLs (Model 1d). The fit indices for all 4 models are summarized in Multimedia Appendix 4. After controlling for covariates, Model 1d still had a good fit to the data (RMSEA=0.048, SRMR=0.029, CFI=0.903, TLI=0.900). As expected, depressive symptoms at each time point were positively related to depressive symptoms over time, as was disability in IADLs. The cross-lagged effects of prior depressive symptoms on subsequent disability in IADLs

XSL•FO
RenderX

at T2 (β=0.070, *P*<.001) and T3 (β=0.074, *P*<.001) as well as disability in IADLs on subsequent depressive symptoms at T2 (β=0.094, *P*<.001) and T3 (β=0.100, *P*<.001) were significant (Table 2). Greater prior depressive symptoms predicted greater subsequent disability in IADLs, and greater prior disability in IADLs predicted greater subsequent depressive symptoms. Additionally, the effect of disability in IADLs on depression was greater than the impact of depression on disability in IADLs.

Figure 2. Cross-lagged panel model of the bidirectional relationship between depressive symptoms and disability in instrumental activities of daily living (IADLs). T1: 2015; T2: 2018; T3: 2020. ***P<.001.

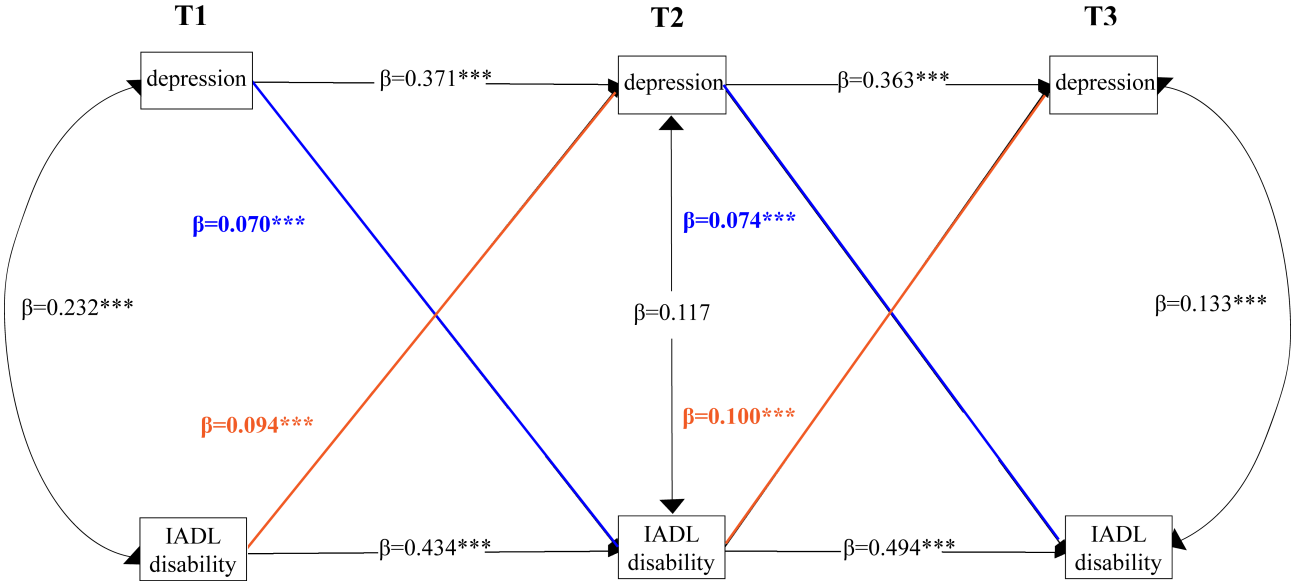


Table . Statistical results of the bidirectional relationship between depressive symptoms and disability in instrumental activities of daily living (IADLs).

Variable	Autoregression estimates				Cross-lagged estimates			
	Depressive symptoms→disability in IADLs		Disability in IADLs→depressive symptoms		Depressive symptoms→disability in IADLs		Disability in IADLs→depressive symptoms	
	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95%CI)	*P* value	β (95% CI)	*P* value
Model 1a[a]								
2015→2018	0.351 (0.317-0.384)	<.001	0.474 (0.444-0.504)	<.001	0.061 (0.028-0.095)	<.001	0.096 (0.060-0.132)	<.001
2018→2020	0.381 (0.349-0.414)	<.001	0.470 (0.440-0.500)	<.001	0.079 (0.045-0.113)	<.001	0.101 (0.066-0.136)	<.001
Model 1b[b]								
2015→2018	0.351 (0.318-0.384)	<.001	0.472 (0.443-0.502)	<.001	0.068 (0.045-0.091)	<.001	0.095 (0.071-0.119)	<.001
2018→2020	0.380 (0.348-0.412)	<.001	0.472 (0.442-0.501)	<.001	0.073 (0.048-0.098)	<.001	0.102 (0.076-0.127)	<.001
Model 1c[c]								
2015→2018	0.372 (0.347-0.397)	<.001	0.433 (0.410-0.456)	<.001	0.076 (0.043-0.110)	<.001	0.084 (0.049-0.119)	<.001
2018→2020	0.363 (0.335-0.390)	<.001	0.494 (0.469-0.520)	<.001	0.068 (0.035-0.101)	<.001	0.109 (0.075-0.143)	<.001
Model 1d[d]								
2015→2018	0.371 (0.346-0.396)	<.001	0.434 (0.411-0.457)	<.001	0.070 (0.047-0.093)	<.001	0.094 (0.070-0.119)	<.001
2018→2020	0.363 (0.336-0.391)	<.001	0.494 (0.468-0.519)	<.001	0.074 (0.049-0.098)	<.001	0.100 (0.075-0.126)	<.001

[a]Model 1a: unconstrained model.

[b]Model 1b: constrained cross-lagged paths.

[c]Model 1c: constrained autoregressive paths.

[d]Model 1d: constrained all paths.

The Longitudinal Mediating Role of Sleep Disturbances

As Figure 3 shows, after adding 2 indirect paths that shared sleep disturbances as a potential mediator and adjustment for control variables, Model 2d still fitted the data adequately (RMSEA=0.050, SRMR=0.028, CFI=0.921, TLI=0.909). The fit indices for all 4 models are summarized in Multimedia Appendix 4. Similarly, we found cross-lagged effects of significant paths from depressive symptoms to disability in

XSL•FO
RenderX

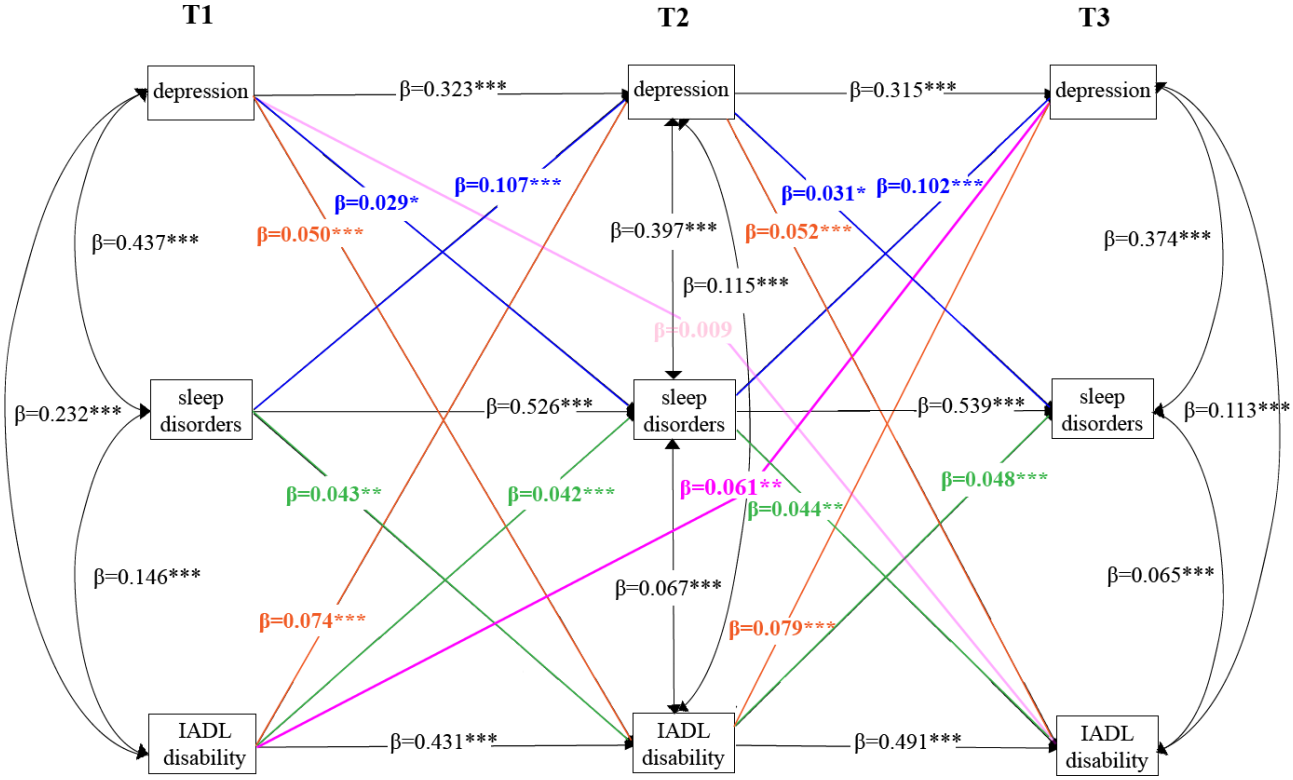
IADLs at T2 (β=0.050, *P*<.001) and T3 (β=0.052, *P*<.001) and from disability in IADLs to depressive symptoms at T2 (β=0.074, *P*<.001) and T3 (β=0.079, *P*<.001), which slightly reduced in size compared with those in Model 1d (Table 3). The cross-lagged effect of disability in IADLs on depression was still greater than the impact of depression on disability in IADLs.

Our analysis identified 2 indirect pathways from sleep disturbances. First, higher levels of prior depressive symptoms were found to be a greater risk of more severe sleep disturbances at T2 (β=0.029, *P*=.02) and T3 (β=0.031, *P*=.02), and following that, those with worse prior sleep disturbances were more likely to report higher amounts of subsequent disability in IADLs at T2 (β=.0043, *P*<.001) and T3 (β=0.044, *P*<.001). Second, greater prior disability in IADLs was a significant predictor of more severe subsequent sleep disturbances at T2 (β=0.042, *P*<.001) and T3 (β=0.048, *P*<.001); subsequently, severe prior sleep disturbances predicted greater depressive symptoms at T2 (β=0.107, *P*<.001) and T3 (β=0.102, *P*<.001).

β, standardized coefficient; CI: confidence interval.

To further test the longitudinal mediation, we computed the MacKinnon formula for calculating the mediated percentage, which is the indirect effect divided by the total effect (Multimedia Appendix 5). The indirect effects of prior disability in IADLs on depressive symptoms via sleep disturbances (β=0.062, SE=0.010, *P*<.001) and prior depressive symptoms on disability in IADLs via sleep disturbances (β=0.043, SE=0.011, *P*<.001) were significant across the 2 time intervals. Importantly, we observed that the direct effect of disability in IADLs on depression was significant (β=0.061, SE=0.018, *P*=.001) after considering the longitudinal mediating role of sleep disturbances, and the mediating effect of sleep disturbances reached 50.41%. Conversely, the path from depression to disability in IADLs was insignificant (β=0.009, SE=0.018, *P*=.61). It is worth noting that sleep disturbance was the fully mediating factor in the path from depression symptoms to disability in IADLs.

Figure 3. Longitudinal mediating effect of sleep disturbances in the bidirectional relationship between depressive symptoms and disability in instrumental activities of daily living (IADLs). T1: 2015; T2: 2018; T3: 2020. *P<.05, **P<.01, ***P<.001.



XSL•FO
RenderX

Table . Statistical results of the longitudinal mediating of sleep disturbances in the reciprocal pathway between disability in instrumental activities in daily living (IADLs) and depressive symptoms.

Variable	Autoregression estimates				Cross-lagged estimates			
	2015→2018		2018→2020		2015→2018		2018→2020	
	β (95%CI)	*P* value	β (95%CI)	*P* value	β (95%CI)	*P* value	β (95%CI)	*P* value
Disability in IADLs→sleep disturbances→depressive symptoms								
Model 2a[a]								
Disability in IADLs→sleep disturbances	0.472 (0.442-0.503)	<.001	0.466 (0.435-0.496)	<.001	0.057 (0.024-0.090)	.001	0.036 (0.004-0.068)	.03
Sleep disturbances→depressive symptoms	0.506 (0.474-0.538)	<.001	0.557 (0.526-0.588)	<.001	0.097 (0.058-0.135)	<.001	0.111 (0.073-0.150)	<.001
Disability in IADLs→depressive symptoms	0.309 (0.271-0.347)	<.001	0.327 (0.290-0.365)	<.001	0.092 (0.056-0.127)	<.001	0.067 (0.028-0.106)	.001
Model 2b[b]								
Disability in IADLs→sleep disturbances	0.470 (0.440-0.499)	<.001	0.468 (0.438-0.498)	<.001	0.042 (0.021-0.063)	<.001	0.049 (0.024-0.073)	<.001
Sleep disturbances→depressive symptoms	0.520 (0.492-0.549)	<.001	0.543 (0.514-0.572)	<.001	0.108 (0.080-0.136)	<.001	0.100 (0.074-0.127)	<.001
Disability in IADLs→depressive symptoms	0.300 (0.266-0.335)	<.001	0.335 (0.302-0.369)	<.001	0.077 (0.051-0.103)	<.001	0.082 (0.054-0.109)	<.001
Model 2c[c]								
Disability in IADLs→sleep disturbances	0.323 (0.295-0.351)	<.001	0.491 (0.466-0.517)	<.001	0.051 (0.019-0.083)	.002	0.040 (0.009-0.071)	.01
Sleep disturbances→depressive symptoms	0.523 (0.499-0.547)	<.001	0.545 (0.517-0.573)	<.001	0.100 (0.064-0.135)	<.001	0.111 (0.075-0.147)	<.001

XSL•FO
RenderX

Variable	Autoregression estimates				Cross-lagged estimates			
	2015→2018		2018→2020		2015→2018		2018→2020	
	β (95%CI)	*P* value	β (95%CI)	*P* value	β (95%CI)	*P* value	β (95%CI)	*P* value
Disability in IADLs→depressive symptoms	0.431 (0.408-0.454)	<.001	0.314 (0.285-0.343)	<.001	0.081 (0.046-0.117)	<.001	0.073 (0.035-0.112)	<.001
Model 2d[d]								
Disability in IADLs→sleep disturbances	0.431 (0.408-0.454)	<.001	0.491 (0.465-0.516)	<.001	0.042 (0.021-0.063)	<.001	0.048 (0.024-0.072)	<.001
Sleep disturbances→depressive symptoms	0.526 (0.502-0.550)	<.001	0.539 (0.512-0.566)	<.001	0.107 (0.079-0.135)	<.001	0.102 (0.075-0.129)	<.001
Disability in IADLs→depressive symptoms	0.323 (0.295-0.351)	<.001	0.315 (0.286-0.344)	<.001	0.074 (0.049-0.100)	<.001	0.079 (0.052-0.106)	<.001
Depressive symptoms→sleep disturbances→disability in IADLs								
Model 2a[a]								
Depressive symptoms→sleep disturbances	0.309 (0.271-0.347)	<.001	0.327 (0.290-0.365)	<.001	0.052 (0.016-0.088)	.005	0.048 (0.012-0.076)	.03
Sleep disturbances→disability in IADLs	0.506 (0.474-0.538)	<.001	0.557 (0.526-0.588)	<.001	0.036 (0.005-0.073)	.04	0.050 (0.012-0.088)	.009
Depressive symptoms→Disability in IADLs	0.472 (0.442-0.503)	<.001	0.466 (0.435-0.496)	<.001	0.046 (0.009-0.083)	.02	0.051 (0.012-0.090)	.01
Model 2b[b]								
Depressive symptoms→sleep disturbances	0.300 (0.266-0.335)	<.001	0.335 (0.302-0.369)	<.001	0.029 (0.004-0.054)	.02	0.030 (0.005-0.056)	.02
Sleep disturbances→disability in IADLs	0.520 (0.492-0.549)	<.001	0.543 (0.514-0.572)	<.001	0.042 (0.017-0.068)	.001	0.044 (0.017-0.071)	.001

XSL•FO
RenderX

Variable	Autoregression estimates				Cross-lagged estimates			
	2015→2018		2018→2020		2015→2018		2018→2020	
	β (95%CI)	*P* value	β (95%CI)	*P* value	β (95%CI)	*P* value	β (95%CI)	*P* value
Depressive symptoms→disability in IADLs	0.470 (0.440-0.499)	<.001	0.468 (0.438-0.498)	<.001	0.046 (0.020-0.072)	<.001	0.050 (0.022-0.078)	<.001
Model 2c[c]								
Depressive symptoms→sleep disturbances	0.323 (0.295-0.351)	<.001	0.314 (0.285-0.343)	<.001	0.049 (0.016-0.081)	.004	0.035 (0.008-0.070)	.03
Sleep disturbances→disability in IADLs	0.523 (0.499-0.547)	<.001	0.545 (0.517-0.573)	<.001	0.041 (0.004-0.078)	.03	0.047 (0.010-0.084)	.01
Depressive symptoms→disability in IADLs	0.431 (0.408-0.454)	<.001	0.491 (0.466-0.517)	<.001	0.058 (0.021-0.095)	.002	0.043 (0.005-0.082)	.03
Model 2d[d]								
Depressive symptoms→sleep disturbances	0.323 (0.295-0.351)	<.001	0.315 (0.286-0.344)	<.001	0.029 (0.004-0.054)	.02	0.031 (0.005-0.057)	.02
Sleep disturbances→disability in IADLs	0.526 (0.502-0.550)	<.001	0.539 (0.512-0.566)	<.001	0.043 (0.017-0.069)	.001	0.044 (0.018-0.071)	.001
Depressive symptoms→disability in IADLs	0.431 (0.408-0.454)	<.001	0.491 (0.465-0.516)	<.001	0.050 (0.023-0.076)	<.001	0.052 (0.024-0.080)	<.001

[a]Model 2a: unconstrained model.

[b]Model 2b: constrained cross-lagged paths.

[c]Model 2c: constrained autoregressive paths.

[d]Model 2d: constrained all paths.

Stratified Analysis

The findings regarding the bidirectional relationship between depressive symptoms and disability in IADLs by gender, district, and empty nester status are shown in Multimedia Appendix 6. After controlling for covariates, there was no significant difference in the bidirectional relationship between depressive symptoms and disability in IADLs in empty nesters and non-empty nesters. Furthermore, in the 2 time intervals, the cross-lagged effect of higher prior depressive symptoms on adverse subsequent disability in IADLs was greater for women than for men; however, the pathway from severe disability in IADLs to greater subsequent depressive symptoms was not different for men and women. Another finding was that the effect of prior depressive symptoms was found to predict greater subsequent disability in IADLs in urban areas than in village areas, whereas the effect of prior disability in IADLs on

XSL•FO
RenderX

subsequent depressive symptoms was not significant in different districts.

In general, sleep disturbance was a longitudinal mediator in the bidirectional association between depression and disability in IADLs in different empty-nester statuses, genders, and districts (Multimedia Appendices 7 and 8). We did not discover significant differences in the path from depressive symptoms to sleep disturbances at T2 and T3 across different empty nester statuses (*P*=.78). Nevertheless, we observed a significant difference in the effect of prior sleep disturbances on subsequent disability in IADLs (*P*=.04), which was higher in non-empty nesters (Multimedia Appendix 7). Similarly, a clear difference existed in the effect of prior depressive symptoms on subsequent sleep disturbances by gender: Women had higher cross-lagged effects (*P*=.008). In addition, we did not observe a significant difference in the path from sleep disturbances to disability in IADLs by gender (*P*=.93). Likewise, no differences were observed in the cross-lagged effects between depressive symptoms and sleep disturbances nor between sleep disturbances and disability in IADLs by district. Additionally, we found no differences in the longitudinal mediating effect of sleep disturbances in the reciprocal relationship between depressive symptoms and disability in IADLs across various empty-nester statuses, genders, and districts (Multimedia Appendix 8).

The mediating effect of sleep disturbances in the pathway from disability in IADLs to depressive symptoms was significantly stronger among empty nesters, accounting for 57.48%, compared with a lower proportion of 35.65% among non-empty nesters. We found that the mediated percentage of sleep disturbances in prior disability in IADLs on subsequent depressive symptoms among women was 31.51%. In addition, the mediating effect of sleep disturbance showed little difference across districts in older adults (Multimedia Appendix 9).

Sensitivity Analysis

Acknowledging the potential conceptual overlap between item 7 of the CESD-10 scale (“My sleep was restless”) and our core measure of sleep disturbance, a sensitivity analysis was performed to assess the robustness of the mediation effects. All CLPMs were re-estimated using a modified CESD-9 score that excluded item 7 (Multimedia Appendices 2 and 3). The sensitivity analysis confirmed that all key cross-lagged paths underlying the mediation pathways remained statistically significant after removing the overlapping item. Although there was conceptual overlap between the sleep item in the CESD-10 scale and the sleep disturbance mediator, it did not fundamentally alter the main findings of this study. The longitudinal mediating role of sleep disturbances in the bidirectional relationship between depression and disability in IADLs was robust.

Discussion

Principal Findings

In this dynamic cohort study based on longitudinal data from older Chinese adults, we examined the bidirectional relationship between depressive symptoms and disability in IADLs mediated by sleep disturbances within a 5-year period. Moreover, we found the effect of disability in IADLs on depression was greater than the impact of depression on disability in IADLs. Notably, sleep disturbances served as the primary mediator in the path from depressive symptoms to disability in IADLs, and the mediating effect of sleep disturbances accounted for 50.41% of the association between disability in IADLs and subsequent depressive symptoms in older adults.

Whereas cross-sectional studies have presented a static view of pairwise correlations between depression, sleep, and disability in IADLs, a longitudinal analysis was better able to reveal how progressions of sleep disturbances affected the depression-disability cycle over time. The observed bidirectional association between depressive symptoms and disability in IADL aligns with previous research in older populations [15,36]. There were several potential mechanisms underlying the relationship between depressive symptoms and disability in IADLs. First, the prolonged presence of certain somatic depressive symptoms, particularly fatigue and pain, may contribute to enhanced decline of physical functioning over time [36]. Older adults with depressive symptoms are more likely to experience amplified symptom burden and complications of chronic medical conditions, both of which may increase the risks for disability. Another potential explanation is that biological changes associated with depressive symptoms, such as elevated cortisol levels and insulin resistance, may increase disability risks [37]. The severity of depression was associated with disability in IADLs but not disability in ADLs, possibly due to its impact on discretionary activities [38]. IADLs, which decline before basic ADLs, are key indicators of independence loss in older adults. We observed that the effect of disability in IADLs on depression was greater than the impact of depression on disability in IADLs; however, there are currently no studies explaining this phenomenon. In our opinion, the possible reason is the decrease in physical activity, cognitive impairment, and changes in comorbidity patterns with advancing age, all of which could lead to physical disability [39,40]. These somatic reactions make it difficult for older adults to adapt to the changes in the early stages, increasing the risk of depression. Conversely, for individuals with depression or those at a high risk of developing depression, the onset of functional disability is characterized by a certain time delay [41], resulting in a smaller effect size in the early stages. This reciprocal relationship highlights the need for comprehensive management of both physical and psychological factors in older adults to mitigate the bidirectional, vicious cycle of depression and functional disability.

Our study provided new evidence on how sleep disturbances influence the reciprocal relationship between depression and disability in IADLs. As a result, the following conclusions could be drawn. Sleep disturbances served as the primary mediator in the path from depressive symptoms to disability in IADLs, indicating a critical link between mental health and physical functioning in older adults. This underscores the need for integrated interventions that target both mental and sleep health to promote overall well-being and maintain independence in the aging population. The mediating effect of sleep disturbances accounted for 50.41% of the association between disability in IADLs and subsequent depressive symptoms in older adults.

XSL•FO
RenderX

The remaining variance in the relationship may be influenced by other factors, such as cognitive decline [39], social isolation [42], or chronic health conditions [43]. Sleep is crucial for maintaining overall well-being and plays a significant role in disrupting the detrimental cycle between depression and disability in IADLs. Several mechanisms may contribute to this relationship. The bidirectional associations and co-occurrence between sleep disturbances and late-life depression demonstrated both concurrent and sequential comorbidity patterns. Therefore, sleep disturbance was not only a comorbidity of depression but also a prodromal symptom, which could predict the occurrence and outcome of depression [21]. It is important to highlight the potential impact of treatment of sleep disturbances before, during, and after depression and functional disability. Timely identification of and interventions for sleep disturbances are crucial to reducing the increasing risks of depression and functional disability in older adults.

In the subgroup analysis, our findings also highlighted the important role of sleep disturbances in the context of specific subgroups. The mediating effect of sleep disturbances in the reciprocal relationship between disability in IADLs and depressive symptoms was significantly stronger among empty nesters, who might have less awareness and ability to maintain health, which means they might not have a better quality of life [44]. Additionally, the mediating effect of sleep disturbances was more pronounced in women, which may be related to hormonal differences and greater susceptibility to sleep disturbances among women [45]. In particular, it is necessary to focus on empty nesters, provide necessary financial support and basic social security, and develop effective intervention strategies to improve the quality of life of empty nesters. We concluded that managing or treating sleep disturbances could play an important role in reducing levels of future depression and disability. Our results have implications for planning and maintaining healthy sleep environments. Specifically, the findings underscore the importance of addressing sleep disturbance as a critical component in the prevention and management of depression and functional disability among older adults. Intervention strategies should focus on creating sleep-friendly living spaces, promoting good sleep hygiene, and providing education on the impact of poor sleep on mental and physical health. Furthermore, there is an urgent need to use portable wearable devices to objectively measure sleep-related indicators, thereby identifying and conducting evidence-based clinical research to further confirm that sleep can break the detrimental cycle between depression and functional disability. By improving sleep quality, we may not only enhance the overall well-being of older individuals but also potentially reduce the burden on health care systems by preventing the exacerbation of depressive symptoms and a decline in functional abilities.

Limitations

Although some interesting findings are presented here, there were limitations to this study. First, due to the lack of objective measurement data, the self-reported nature of the data meant that results must be interpreted with caution. There may be a tendency to over- or underestimate symptoms. It is crucial to acknowledge the limitations inherent in our assessment of sleep disturbance. Although the composite score used in this study provided a measure for our large-scale epidemiological investigation, its exclusive reliance on self-reported sleep quality and duration, without validation against established instruments such as the Pittsburgh Sleep Quality Index or objective measures, meant that our findings should be interpreted as reflecting subjective sleep experiences rather than clinically diagnosed sleep disorders. Second, the impact on specific items of disability in IADLs was not explored nor was the severity of depression differentiated. Therefore, future subgroup analyses of such data would be more informative in clarifying the underlying mechanisms. Third, we were unable to include all potential confounding factors, such as stressful life events, medication use for depression, family history of psychiatric disorders, health behaviors and conditions, and retirement information, in the analysis or adjust for them [42]. Fourth, the CLPM used in our study did not distinguish within-person and between-person variance, which is considered a limitation by some researchers [35]. The method was not applicable when the goal was to examine the association of within-person variance. In addition, this study primarily focused on older adults in China; it is unclear whether study findings may be relevant to other countries or ethnic groups.

Conclusions

This longitudinal cohort study provided evidence for the bidirectional relationship between depressive symptoms and disability in IADLs in older adults and highlighted the mediating role of sleep disturbances in this relationship. Addressing sleep disturbances and promoting sleep health may be key components of effective interventions for depression and functional disability in older adults. Overall, public health policies could benefit from incorporating initiatives that support healthy sleep habits across the lifespan, recognizing that sleep is a fundamental pillar of health and well-being.

Acknowledgments

We would like to thank the staff and participants of the China Health and Retirement Longitudinal Study (CHARLS).

This work was supported by the Senior Talent Startup Fund of Nanchang University (28170120/9167) and the Postgraduate Innovation Special Fund of Jiangxi Province (YC2024-B052). The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

XSL•FO
RenderX

Data Availability

The data that support the findings of this study are available from the China Health and Retirement Longitudinal Study (CHARLS) [46].

Authors' Contributions

HW and WG conceptualized and designed the study. HW built the analytical datasets, conducted the data analysis, and drafted and revised the manuscript. LW, YY, and XL reviewed the analysis and contributed to the interpretation of data. HW drafted the manuscript, and all other authors critically reviewed and revised the manuscript. All authors accept responsibility for the decision to submit for publication. WG is responsible for the overall work as the guarantor.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional methodology information.
[DOCX File, 20 KB - aging_v8i1e76643_app1.docx]

Multimedia Appendix 2
Statistical results of the longitudinal mediating of sleep disturbances in the reciprocal pathway between disability in instrumental activities of daily living (IADLs) and depressive symptoms using a modified Center for Epidemiologic Studies Depression (CESD) scale score (excluding Item 7).
[DOCX File, 23 KB - aging_v8i1e76643_app2.docx]

Multimedia Appendix 3
Statistical results of the longitudinal mediating effect of sleep disturbances in the bidirectional relationship between depressive symptoms and disability in instrumental activities of daily living (IADLs) using a modified Center for Epidemiologic Studies Depression (CESD) scale score (excluding Item 7).
[DOCX File, 16 KB - aging_v8i1e76643_app3.docx]

Multimedia Appendix 4
Model fit indices and comparison of 4 models.
[DOCX File, 16 KB - aging_v8i1e76643_app4.docx]

Multimedia Appendix 5
Statistical results of the longitudinal mediating effect of sleep disturbances in the bidirectional relationship between depressive symptoms and disability in instrumental activities of daily living (IADLs).
[DOCX File, 16 KB - aging_v8i1e76643_app5.docx]

Multimedia Appendix 6
Statistical results of the bidirectional relationship between depressive symptoms and disability in instrumental activities of daily living (IADLs) by empty nester status, gender, and district.
[DOCX File, 20 KB - aging_v8i1e76643_app6.docx]

Multimedia Appendix 7
Statistical results of the longitudinal mediating effect of sleep disturbances in the pathway from depressive symptoms to disability in instrumental activities of daily living (IADLs) by empty nester status, gender, and district.
[DOCX File, 18 KB - aging_v8i1e76643_app7.docx]

Multimedia Appendix 8
Statistical results of the longitudinal mediating effect of sleep disturbances in the pathway from disability in instrumental activities of daily living (IADLs) to depressive symptoms by empty nester status, gender, and district.
[DOCX File, 18 KB - aging_v8i1e76643_app8.docx]

Multimedia Appendix 9

XSL•FO
RenderX

Statistical results of the longitudinal mediating effect of sleep disturbances in the bidirectional relationship between depressive symptoms and disability in instrumental activities of daily living (IADLs) by gender, empty nester status, and district.
[DOCX File, 18 KB - aging_v8i1e76643_app9.docx]

Checklist 1
RECORD statement checklist of items that should be reported in observational studies using routinely collected health data.
[DOCX File, 32 KB - aging_v8i1e76643_app10.docx]

References

1. Ageing and health. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2025-04-27]
2. Update from the Seventy-fourth World Health Assembly. World Health Organization. 2021 May 27. URL: https://www.who.int/news/item/27-05-2021-update-from-the-seventy-fourth-world-health-assembly-27-may-2021 [accessed 2024-05-18]
3. Guo L, An L, Luo F, Yu B. Social isolation, loneliness and functional disability in Chinese older women and men: a longitudinal study. Age Ageing 2021 Jun 28;50(4):1222-1228. [doi: 10.1093/ageing/afaa271] [Medline: 33352582]
4. Gao Y, Du L, Cai J, Hu T. Effects of functional limitations and activities of daily living on the mortality of the older people: a cohort study in China. Front Public Health 2022;10:1098794. [doi: 10.3389/fpubh.2022.1098794] [Medline: 36743188]
5. Mlinac ME, Feng MC. Assessment of activities of daily living, self-care, and independence. Arch Clin Neuropsychol 2016 Sep;31(6):506-516. [doi: 10.1093/arclin/acw049] [Medline: 27475282]
6. Wang W, Liu Y, Ji D, et al. The association between functional disability and depressive symptoms among older adults: findings from the China Health and Retirement Longitudinal Study (CHARLS). J Affect Disord 2024 Apr;351:518-526. [doi: 10.1016/j.jad.2024.01.256]
7. Herrman H, Patel V, Kieling C, et al. Time for united action on depression: a Lancet–World Psychiatric Association Commission. The Lancet 2022 Mar;399(10328):957-1022. [doi: 10.1016/S0140-6736(21)02141-3]
8. Luo H, Li J, Zhang Q, et al. Obesity and the onset of depressive symptoms among middle-aged and older adults in China: evidence from the CHARLS. BMC Public Health 2018 Jul 24;18(1):909. [doi: 10.1186/s12889-018-5834-6] [Medline: 30041632]
9. Alexopoulos GS. Depression in the elderly. Lancet 2005;365(9475):1961-1970. [doi: 10.1016/S0140-6736(05)66665-2] [Medline: 15936426]
10. First MB. Diagnostic and statistical manual of mental disorders, 5th edition, and clinical utility. J Nerv Ment Dis 2013 Sep;201(9):727-729. [doi: 10.1097/NMD.0b013e3182a2168a] [Medline: 23995026]
11. Vitiello MV, Larsen LH, Moe KE. Age-related sleep change: gender and estrogen effects on the subjective-objective sleep quality relationships of healthy, noncomplaining older men and women. J Psychosom Res 2004 May;56(5):503-510. [doi: 10.1016/S0022-3999(04)00023-6] [Medline: 15172206]
12. Ancoli-Israel S, Cooke JR. Prevalence and comorbidity of insomnia and effect on functioning in elderly populations. J Am Geriatr Soc 2005 Jul;53(7 Suppl):S264-S271. [doi: 10.1111/j.1532-5415.2005.53392.x] [Medline: 15982375]
13. Cole MG, Dendukuri N. Risk factors for depression among elderly community subjects: a systematic review and meta-analysis. Am J Psychiatry 2003 Jun;160(6):1147-1156. [doi: 10.1176/appi.ajp.160.6.1147] [Medline: 12777274]
14. Spira AP, Gonzalez CE, Venkatraman VK, et al. Sleep duration and subsequent cortical thinning in cognitively normal older adults. Sleep 2016 May 1;39(5):1121-1128. [doi: 10.5665/sleep.5768] [Medline: 26951390]
15. Wang J, Luo N, Sun Y, et al. Exploring the reciprocal relationship between activities of daily living disability and depressive symptoms among middle-aged and older Chinese people: a four-wave, cross-lagged model. BMC Public Health 2023;23(1). [doi: 10.1186/s12889-023-16100-0]
16. Bacon KL, Heeren T, Keysor JJ, Stuver SO, Cauley JA, Fredman L. Longitudinal and reciprocal relationships between depression and disability in older women caregivers and noncaregivers. Gerontologist 2016 Aug;56(4):723-732. [doi: 10.1093/geront/gnu157] [Medline: 26035874]
17. Gayman MD, Turner RJ, Cui M. Physical limitations and depressive symptoms: exploring the nature of the association. J Gerontol B Psychol Sci Soc Sci 2008 Jul;63(4):S219-S228. [doi: 10.1093/geronb/63.4.s219] [Medline: 18689771]
18. Wang D, Qi S, Wang BH, Hu QT, Wang ZK, Wang Z. Depression status and its influencing factors among the elderly aged 60 years and above in three provinces of China. Zhonghua Liu Xing Bing Xue Za Zhi 2022;43(12):1925-1931. [doi: 10.3760/cma.j.cn112338-20220418-00313] [Medline: 36572465]
19. Takano K, Tanno Y. Diurnal variation in rumination. Emotion 2011 Oct;11(5):1046-1058. [doi: 10.1037/a0022757] [Medline: 21517158]
20. Bao YP, Han Y, Ma J, et al. Cooccurrence and bidirectional prediction of sleep disturbances and depression in older adults: meta-analysis and systematic review. Neurosci Biobehav Rev 2017 Apr;75:257-273. [doi: 10.1016/j.neubiorev.2017.01.032] [Medline: 28179129]
21. Fang H, Tu S, Sheng J, Shao A. Depression in sleep disturbance: a review on a bidirectional relationship, mechanisms and treatment. J Cell Mol Med 2019 Apr;23(4):2324-2332. [doi: 10.1111/jcmm.14170] [Medline: 30734486]

XSL•FO
RenderX

22. Yasugaki S, Okamura H, Kaneko A, Hayashi Y. Bidirectional relationship between sleep and depression. Neurosci Res 2025 Feb;211:57-64. [doi: 10.1016/j.neures.2023.04.006] [Medline: 37116584]
23. Song Y, Dzierzewski JM, Fung CH, et al. Association between sleep and physical function in older veterans in an adult day healthcare program. J American Geriatrics Society 2015 Aug;63(8):1622-1627. [doi: 10.1111/jgs.13527]
24. Liu M, Du X, Sun Y, Zhou A, Sun S, Wu Y. The mediating role of cognition in the relationship between sleep duration and instrumental activities of daily living disability among middle-aged and older Chinese. Arch Gerontol Geriatr 2021;94:104369. [doi: 10.1016/j.archger.2021.104369] [Medline: 33556636]
25. Lee YH, Kong D, Lee YTH, Lin CH, Liu CT, Chang YC. Functional disabilities and changes in sleep quality and duration among older adults: results from a longitudinal study in China, 2005-2014. Eur Geriatr Med 2022 Aug;13(4):967-975. [doi: 10.1007/s41999-022-00619-3] [Medline: 35191012]
26. Hodes GE, Bangasser D, Sotiropoulos I, Kokras N, Dalla C. Sex differences in stress response: classical mechanisms and beyond. Curr Neuropharmacol 2024;22(3):475-494. [doi: 10.2174/1570159X22666231005090134] [Medline: 37855285]
27. Dong E, Liu S, Chen M, et al. Differences in regional distribution and inequality in health-resource allocation at hospital and primary health centre levels: a longitudinal study in Shanghai, China. BMJ Open 2020 Jul 19;10(7):e035635. [doi: 10.1136/bmjopen-2019-035635] [Medline: 32690509]
28. Xu S, Yang X, Liu J, et al. Health and wellbeing among the empty nest and non-empty nest elderly in China-results from a national cross-sectional study. PLoS One 2023;18(9):e0291231. [doi: 10.1371/journal.pone.0291231] [Medline: 37699029]
29. Wu H, Wang L, Chen H, Gao W. The longitudinal mediating role of sleep in the bidirectional relationship between depression and functional disability among older adults: a systematic review and conceptual framework. medRxiv. Preprint posted online on Apr 28, 2025. [doi: 10.1101/2025.02.27.25323031]
30. Lu B, Liu X, Lim J, Yang M. Changes in the morbidity prevalence and morbidity-free life expectancy of the elderly population in China from 2000 to 2010. The Journal of the Economics of Ageing 2019 May;13:113-121. [doi: 10.1016/j.jeoa.2018.05.002]
31. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). Int J Epidemiol 2014 Feb;43(1):61-68. [doi: 10.1093/ije/dys203] [Medline: 23243115]
32. Chertkow H, Nasreddine Z, Joanette Y, et al. Mild cognitive impairment and cognitive impairment, no dementia: part A, concept and diagnosis. Alzheimers Dement 2007 Oct;3(4):266-282. [doi: 10.1016/j.jalz.2007.07.013] [Medline: 19595948]
33. Cheng ST, Chan ACM. The Center for Epidemiologic Studies Depression Scale in older Chinese: thresholds for long and short forms. Int J Geriatr Psychiatry 2005 May;20(5):465-470. [doi: 10.1002/gps.1314] [Medline: 15852439]
34. Fan S, Wang Q, Zheng F, et al. Depression as a mediator and social participation as a moderator in the bidirectional relationship between sleep disorders and pain: dynamic cohort study. JMIR Public Health Surveill 2023 Jul 26;9:e48032. [doi: 10.2196/48032] [Medline: 37494109]
35. Hamaker EL, Kuiper RM, Grasman R. A critique of the cross-lagged panel model. Psychol Methods 2015 Mar;20(1):102-116. [doi: 10.1037/a0038889] [Medline: 25822208]
36. Lee G, Arieli R, Ryou YJ, Martin P. The bidirectional relationship between depressive symptoms and functional limitations among centenarian survivors in their 80s: testing bivariate latent change score models. Aging Ment Health 2023;27(9):1720-1728. [doi: 10.1080/13607863.2023.2177830] [Medline: 36786734]
37. Katon WJ. Epidemiology and treatment of depression in patients with chronic medical illness. Dialogues Clin Neurosci 2011;13(1):7-23. [doi: 10.31887/DCNS.2011.13.1/wkaton] [Medline: 21485743]
38. Reichardt LA, Nederveen FE, van Seben R, et al. The longitudinal association between depressive symptoms and functional abilities in older patients. J Psychosom Res 2020 Oct;137:110195. [doi: 10.1016/j.jpsychores.2020.110195] [Medline: 32731046]
39. Kim BJ, Liu L, Nakaoka S, Jang S, Browne C. Depression among older Japanese Americans: the impact of functional (ADL & IADL) and cognitive status. Soc Work Health Care 2018 Feb;57(2):109-125. [doi: 10.1080/00981389.2017.1397588] [Medline: 29236614]
40. Zhu X, Wang Y, Luo Y, Ding R, Shi Z, He P. Bidirectional, longitudinal associations between depressive symptoms and IADL/ADL disability in older adults in China: a national cohort study. BMC Geriatr 2024 Aug 6;24(1):659. [doi: 10.1186/s12877-024-05248-y] [Medline: 39107705]
41. Tian F, Yang H, Pan J. Association between functional disability and long-term trajectories of depressive symptoms: evidence from the China Health and Retirement Longitudinal Study. J Affect Disord 2022 Aug 1;310:10-16. [doi: 10.1016/j.jad.2022.05.001] [Medline: 35525506]
42. Wang Y, Liu M, Yang F, Chen H, Wang Y, Liu J. The associations of socioeconomic status, social activities, and loneliness with depressive symptoms in adults aged 50 years and older across 24 countries: findings from five prospective cohort studies. Lancet Healthy Longev 2024 Sep;5(9):100618. [doi: 10.1016/j.lanhl.2024.07.001] [Medline: 39208829]
43. Huang D, Wang J, Fang H, Fu Y, Lou J. Longitudinal association of chronic diseases with depressive symptoms in middle-aged and older adults in China: mediation by functional limitations, social interaction, and life satisfaction. J Glob Health 2023 Sep 29;13:04119. [doi: 10.7189/jogh.13.04119] [Medline: 37766663]

XSL•FO
RenderX

44. Zhang C, Cai Y, Xue Y, et al. Exploring the influencing factors of quality of life among the empty nesters in Shanxi, China: a structural equation model. Health Qual Life Outcomes 2021 May 28;19(1):156. [doi: 10.1186/s12955-021-01793-x] [Medline: 34049560]
45. Pengo MF, Won CH, Bourjeily G. Sleep in women across the life span. Chest 2018 Jul;154(1):196-206. [doi: 10.1016/j.chest.2018.04.005] [Medline: 29679598]
46. China Health and Retirement Longitudinal Study. URL: https://charls.pku.edu.cn/index.htm [accessed 2025-10-19]

Abbreviations

ADL: activity of daily living
CESD: Center for Epidemiologic Studies Depression
CFI: comparative fit index
CHARLS: China Health and Retirement Longitudinal Study
CLPM: cross-lagged panel model
IADL: instrumental activity of daily living
RECORD: Reporting of Studies Conducted Using Observational Routinely-Collected Health Data
RMSEA: root mean square error of approximation
SRMR: standardized root mean square residual
T1: 2015
T2: 2018
T3: 2020
TLI : Tucker-Lewis index

Edited by A Garcia; submitted 28.04.25; peer-reviewed by Y Liang, Y Duan; revised version received 18.09.25; accepted 02.10.25; published 06.11.25.

Please cite as:
Wu H, Wang L, Liu X, Yang Y, Gao W
Sleep Disturbance as a Catalyst in the Cyclical Link Between Depressive Symptoms and Disability in Instrumental Activities of Daily Living in Older Chinese Adults: Longitudinal Cohort Study
JMIR Aging 2025;8:e76643
URL: https://aging.jmir.org/2025/1/e76643
doi:10.2196/76643

© Hao Wu, Lu Wang, Xin Liu, Yuyue Yang, Wei Gao. Originally published in JMIR Aging (https://aging.jmir.org), 6.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Development and Validation of the Healthy Longevity Index for Personalized Healthy Aging in Primary Care: Cross-National Retrospective Analysis

Hsi-Yu Lai[1], MSc; Shu Zhang[2], PhD; Rei Otsuka[2], PhD; Shih-Tsung Huang[3,4], PhD; Hidenori Arai[2], MD, PhD; Fei-Yuan Hsiao[1,5,6*], PhD; Liang-Kung Chen[4,7,8*], MD, PhD

[1]Graduate Institute of Clinical Pharmacy, College of Medicine, National Taiwan University, Taipei, Taiwan

[2]Department of Epidemiology of Aging, National Center for Geriatrics and Gerontology, Obu, Japan

[3]Department of Pharmacy, National Yang Ming Chiao Tung University, Taipei, Taiwan

[4]Center for Healthy Longevity and Aging Sciences, National Yang Ming Chiao Tung University, Taipei, Taiwan

[5]Graduate Institute of Clinical Pharmacy, College of Medicine, National Taiwan University, Taipei, Taiwan

[6]Department of Pharmacy, College of Medicine, National Taiwan University Hospital, Taipei, Taiwan

[7]Center for Geriatrics and Gerontology, Taipei Veterans General Hospital, No. 201, Sec 2, Shih-Pai Road, Taipei, Taiwan

[8]Taipei Municipal Gan-Dau Hospital, Taipei, Taiwan

[*]these authors contributed equally

Corresponding Author:
Liang-Kung Chen, MD, PhD
Center for Healthy Longevity and Aging Sciences, National Yang Ming Chiao Tung University, Taipei, Taiwan

Abstract

Background: Measuring and promoting healthy aging at an individual level remains challenging as promoting healthy longevity requires real-time, personalized tools to assess risk and guide interventions in clinical practice.

Objective: This study aimed to develop and validate a novel Healthy Longevity Index (HLI) for use in primary care settings in older adults.

Methods: Using data from the Taiwan Longitudinal Study on Aging (TLSA; n=4470), we developed a nomogram-based HLI incorporating demographics, lifestyle factors, intrinsic capacity (IC) measures, and chronic conditions to predict 4-, 8-, and 12-year disability- and dementia-free survival (absence of physical disability, dementia, or mortality). The HLI was internally validated in a TLSA subset and externally validated in the Japanese National Institute for Longevity Sciences, Longitudinal Study of Aging (NILS-LSA) cohort (n=1090).

Results: The 12-year HLI nomogram demonstrated robust performance, with C-statistics of 0.79 (bootstrapped 95% CI 0.78 - 0.80) in the TLSA training cohort and 0.77 (bootstrapped 95% CI 0.75 - 0.79) in the TLSA validation cohort. External validation in the NILS-LSA yielded a C-statistic of 0.71 (bootstrapped 95% CI 0.66 - 0.76). The HLI effectively stratified participants into risk tertiles, with the highest-risk group showing only 27.8% probability of 12-year disability- and dementia-free survival compared to 87.8% in the lowest-risk group. Key predictors included age, sex, education, and, particularly, IC impairments in locomotion, visual acuity, and cognition—all assessable during routine primary care consultations.

Conclusions: The HLI provides a practical tool for real-time, personalized assessment of healthy longevity risk in primary care settings. Its design enables providers to deliver person-centered care through targeted interventions and individualized prevention strategies that promote healthy aging across populations, especially in older adults.

(JMIR Aging 2025;8:e80034) doi:10.2196/80034

KEYWORDS

healthy longevity; disability-free survival; intrinsic capacity; primary care; precision health

Introduction

The burgeoning global demographic of older adults presents complex challenges across physical, mental, and social spheres, necessitating comprehensive strategies to optimize later-life well-being, as conceptualized by the World Health Organization (WHO)'s definition of healthy aging, that is, the enhancement and preservation of functional ability over time [1]. Consequently, primary care systems worldwide face increasing pressure to address these challenges as the first point of contact for aging populations. The paramount goal and pragmatic approach is to extend health span by minimizing

XSL•FO
RenderX

disability-adjusted life years (DALYs) throughout the life course [2]. While DALYs constitute a robust metric for quantifying unhealthy life years at a macro level, their derivation is contingent upon extensive public health data. Moreover, DALYs are inherently lagging indicators, typically available with a 2-year temporal offset from the reference period. Besides, DALYs reflect population-level status, limiting their applicability to individuals and clinical decision-making in primary care settings. In 2016, the WHO's integrated care for older people (ICOPE) framework operationalized the measurement of intrinsic capacity (IC) as a composite biological construct to facilitate personalized care planning for healthy aging [3]. However, translating IC assessment into routine primary care practice remains challenging. Despite impairments in IC, overall or individual domain, being a substantial predisposing factor for disability and mortality, and predictive of diverse adverse outcomes [4-9], its correlation with DALYs and capacity for real-time estimation within the context of healthy aging remains ill-defined.

Technological advancements have facilitated the development of diverse real-time personal risk assessments, thereby revolutionizing health care delivery through enhanced precision, accuracy, and patient outcomes [10]. For instance, a smartwatch app effectively detected atrial fibrillation in a large population, offering potential for early intervention and improved outcomes [11]. Furthermore, a study of over 40,000 people with insulin-treated diabetes found that real-time continuous glycemic monitoring initiation led to a significant decrease in glycated hemoglobin (−0.40%), hypoglycemia (2.7% reduction), and emergency department visits for hypoglycemia, with no change in hyperglycemia or overall hospitalizations [12]. The presented examples unequivocally illustrate the beneficial effects of real-time, personalized health risk assessment, with the potential to significantly enhance health outcomes across a range of conditions. Similar technological advances could transform healthy aging assessment in primary care, where early detection and intervention are paramount for maintaining dignity, independence, and quality of life.

A key advantage of real-time health monitoring lies in its ability to track and potentially modify risk factors that influence health outcomes. In the context of healthy aging, several modifiable risk factors have been identified as crucial determinants of disability-free survival and cognitive health, including lifestyle behaviors (physical activity, smoking cessation, and alcohol consumption), nutritional status, management of chronic conditions (such as hypertension and diabetes), and maintenance of domains such as mobility and cognitive function [7,10,13-19]. However, current approaches to healthy aging assessment often fail to provide individuals with real-time feedback on how changes in these modifiable factors might influence their long-term outcomes, highlighting the need for dynamic, personalized assessment tools.

Unlike the readily quantifiable metrics used in specific disease management, such as diabetes and atrial fibrillation, IC and DALYs present challenges in real-time, personalized assessment and care planning within routine clinical practice. To prospectively estimate the potential burden of health conditions, disability-free survival has emerged as a promising outcome measure for specific disease states [20], treatments [21], and the broader context of healthy aging [22]. Although the exact correspondence between disability-free survival and DALYs warrants further elucidation, a rigorously formulated equation for the former is congruent with empirically determined DALYs [23]. On the other hand, while a previous study established age- and sex-specific reference centiles for IC as a potential indicator of aging trajectories, it solely captured the normative status of IC without quantifying the risk of adverse outcomes or encompassing other factors contributing to healthy longevity, not to mention the possibility of real-time measurements suitable for clinical implementation [5]. To facilitate proactive self-management and tailored interventions for healthy aging, this study sought to develop a novel Health Longevity Index (HLI)—a probabilistic estimate of disability- and dementia-free survival within a specified timeframe—and its corresponding nomogram using a nationally representative cohort sample. In addition, the study endeavored to substantiate the predictive capacity of the index through external validation using a Japanese population-based cohort.

Methods

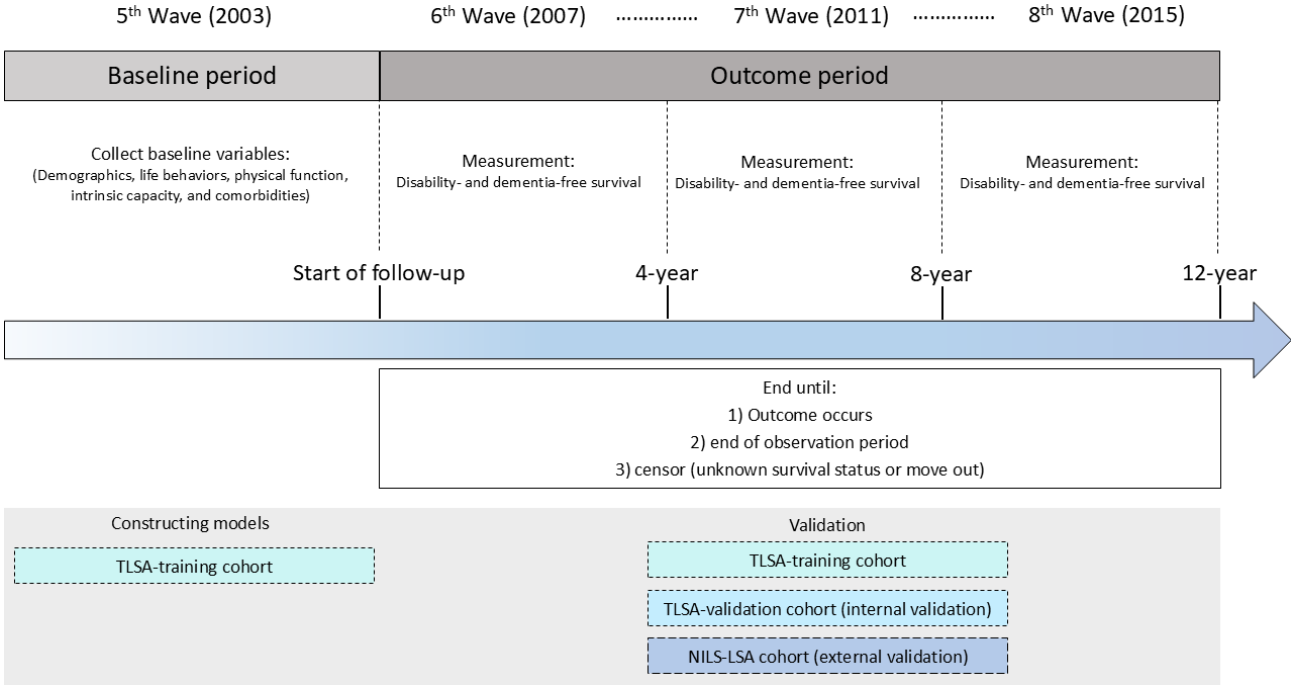
Healthy longevity index (probability of disability- and dementia-free survival) and its nomogram construction cohort: Taiwan Longitudinal Study on Aging (TLSA)

Our HLI was deliberately theory-driven based on the WHO's conceptualization of healthy aging (ie, IC) and fundamental health behavior. Based on the WHO's "ICOPE Handbook of Guidance on person-centered assessment and pathways in primary care" [24], our HLI aims to screen for loss in a range of domains of IC and assess health and social care needs to develop a personalized care plan. To effectively identify individuals at risk of care dependency, the ICOPE pathway begins with ICOPE Step 1, which involves screening for failing tasks across 5 domains of IC: locomotion, vitality, sensory (visual and hearing loss), cognitive function, and psychological well-being.

The data used to develop a novel HLI (probability of disability- and dementia-free survival) and its nomogram (Figure 1) in this study were obtained from the Taiwan Longitudinal Study on Aging (TLSA), a comprehensive national survey initiated by the Health Promotion Administration of the Ministry of Health and Welfare. The TLSA was aimed at deriving deep insights into the overall well-being of middle-aged as well as older adults in Taiwan. To ensure a representative sample of the community, a 3-stage random-sampling method was systematically used, which facilitated the recruitment of participants and achieved a response rate of up to 90%. The scope of the data collection included assessments of participants' physical activity, comorbidities, lifestyle choices, and mental health status. The details for participant recruitment and data collection have been documented in previous publications [7,10,16] and are readily available on the official Health Promotion Administration website [25]. This study was designed and reported following the TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) checklist.

XSL•FO
RenderX

Figure 1. Study design of the Healthy Longevity Index (HLI). NILS-LSA: Japanese National Institute for Longevity Sciences, Longitudinal Study of Aging; TLSA: Taiwan Longitudinal Study on Aging.



The study cohort for constructing the nomogram consisted of 5377 participants from the fifth wave of the TLSA survey (2003). Among them, we excluded individuals residing in nursing homes (n=110), those whose responses were provided by proxies (n=328), participants with unknown residence information (n=15), with unknown BMI (n=164), and with unknown cognitive function (n=35). Given that the missingness rate was less than 5%, a complete case approach was adopted for the analysis (Figure S1 in Multimedia Appendix 1). To minimize reverse causality (ie, undiagnosed, subclinical diseases leading to disability and mortality), we further excluded participants with physical disability (n=201), cognitive disability (n=40), and those who died within the first year of follow-up (n=14). Ultimately, the study cohort included 4470 participants. For the study, this final cohort was stratified into 2 distinct groups using simple random sampling: a training set comprising 70% of the cohort (TLSA-training cohort), which was used for model construction, and a validation set (TLSA-validation cohort) constituting the remaining 30% for further internal validation.

External Validation Cohort: Japan National Institute for Longevity Sciences, Longitudinal Study of Aging (NILS-LSA)

To generalize the applicability of the proposed prediction model, we used the National Institute for Longevity Sciences, Longitudinal Study of Aging (NILS-LSA) cohort from Japan to further validate the model derived from the TLSA cohort. The NILS-LSA is a Japanese population-based prospective cohort study that collects and analyzes variables relating to normal aging and age-related diseases. The data of this study constituted age- and sex-stratified participants recruited by random sampling from Obu and Higashiura, Aichi Prefecture, Japan. The details of the NILS-LSA were previously reported and have been well-studied in previous research [26-29]. The first-wave examination of the NILS-LSA was conducted from November 1997 to April 2000 and included 2267 participants (aged 40 - 79 years). These participants were followed up every 2 years and were replaced by new, age- and sex-matched randomly recruited participants when participants (aged ≤79 years) could not attend follow-up investigations. Participants aged 40 years were also newly recruited every year.

Participants for the external validation were derived from the fourth (June 2004 to July 2006) wave of the NILS-LSA. We chose this wave because this wave covers all the data needed for the analysis and has a relatively long follow-up duration. Among 1225 participants aged 60 years and older (the Mini-Mental State Examination [MMSE] was only conducted in participants aged ≥60 years), 1172 participants had not been certified as having a disability before or at the baseline survey. We further excluded 32 participants with incomplete covariate data at baseline and 50 with a MMSE score <24. Finally, a total of 1090 participants were included in the external validation analysis.

Feature Selection to Construct the HLI and its Nomogram

The study collected a broad spectrum of features, including demographics (including fundamental lifestyle behaviors), underlying medical conditions (self-reported chronic conditions), and, most importantly, IC, to construct the nomogram. Demographic data include age, gender, geographic region of residence, educational attainment, marital status, living condition, as well as their health behaviors (current smoker and alcohol drinker). Furthermore, we include self-reported chronic conditions, encompassing cardiovascular diseases (hypertension, heart disease, and stroke), diabetes mellitus, liver disease, and renal disease. In particular, the baseline status of IC was also assessed for each participant (Table 1).

XSL•FO
RenderX

Table . Description of variables from the Taiwan Longitudinal Study on Aging (TLSA).

Variable	Description	Data type
Demographics		
Age (years)	Respondent's age in 2003 (50-64, 65 - 74, or older than 75)	Categorical (3 levels)
Sex	Respondent's sex (male or female)	Categorical (2 levels)
Living region	Respondent's living region (rural or urban)	Categorical (2 levels)
Education	Respondent's completed education (illiterate, elementary, junior or senior high school, college or above)	Categorical (4 levels)
Marital status	Respondent's marital status in 2003 (marital, divorced or separated, or never married, or widowed)	Categorical (3 levels)
Live with spouse	Does your spouse often live with you?	Categorical (binary)
Life behaviors		
Current smoker	Do you currently smoke?	Categorical (binary)
Current alcohol drinker	Do you currently drink alcohol?	Categorical (binary)
Physical activity	Do you usually exercise?	Categorical (binary)
BMI	Respondent's BMI in 2003 (underweight [<18.5], healthy [18.5-24], overweight [24-28], obesity [≥28])	Categorical (ordered, 4 levels)
Functional status		
Difficulty in climbing stairs	Can you walk up to the second or third floor?	Categorical (binary)
Intrinsic capacity		
Locomotion impairment	Do you have difficulty walking 200-300 meters alone without any assistance or tools?	Categorical (binary)
Visual acuity impairment	How well can you perceive objects (even with glasses or contact lenses)?	Categorical (binary)
Hearing impairment	How well can you hear sounds (even with hearing aids)?	Categorical (binary)
Vitality impairment	In the past week, have you experienced a poor appetite or lack of desire to eat?	Categorical (binary)
Psychological impairment	CES-D[a] scale (I felt that everything I did required significant effort; I had difficulty initiating tasks)	Categorical (binary)
Cognition impairment	SPMSQ[b] (Spatial orientation: What is the name of this place?; Temporal orientation: What is the date today?)	Categorical (binary)
Self-reported chronic conditions		
Hypertension	In the past year, have you seen a doctor for this illness?	Categorical (binary)
Diabetes mellitus	In the past year, have you seen a doctor for this illness?	Categorical (binary)
Heart disease	In the past year, have you seen a doctor for this illness?	Categorical (binary)
Stroke	In the past year, have you seen a doctor for this illness?	Categorical (binary)
Liver disease	In the past year, have you seen a doctor for this illness?	Categorical (binary)
Renal disease	In the past year, have you seen a doctor for this illness?	Categorical (binary)

[a]CES-D: Center for Epidemiologic Studies Depression.

XSL•FO
RenderX

[b]SPMSQ: Short Portable Mental Status Questionnaire.

As our study participants were community-dwelling middle-aged and older adults, the IC was constructed in accordance with the WHO's ICOPE Step 1 (screening) framework [30], encompassing 5 domains (locomotion, sensory, vitality, psychological, and cognitive function). As the ICOPE step 1 (screening) is intended to use a feasible tool to screen those with impairment of IC and recommend them to ICOPE step 2 (assessment) for further evaluation, the screening questions are all very easy to read for community-dwelling adults. This screening approach is particularly suited for primary care settings as it requires minimal time and no specialized equipment.

In Table 1, we have specified the screening questions for each domain of IC in TLSA. Our approach closely followed the original ICOPE Step 1 definitions for most domains (Table S1 in Multimedia Appendix 2). However, due to data availability, we also adapted the operationalization of the locomotion domain through expert consensus and evidence from existing literature to ensure alignment with ICOPE screening criteria while maintaining methodological rigor. The evaluation of the locomotion domain followed the Nagi questionnaire [31]. Participants who faced challenges or were unable to cover a distance of 200‐300 meters were categorized as having limited mobility, indicating impairment in the locomotion domain [32,33]. The sensory domain was further divided into visual and hearing subareas. Visual acuity was assessed by asking participants, "How well can you perceive objects (even with glasses or contact lenses)?" Those who responded with "poor" or "very poor" were categorized as experiencing visual impairment [34,35]. Hearing capacity was evaluated through self-reported hearing impairment, with individuals reporting hearing difficulties in either ear (even with a hearing aid) being classified as having impaired hearing [34,36]. For the vitality domain, a decrease in appetite was used as an indicator of nutritional status. Participants who reported experiencing a diminished appetite in the preceding week were categorized as having an impairment in the vitality domain [33]. To assess the psychological domain, 2 questions from The Center for Epidemiologic Studies Depression (CES-D) questionnaire were used: "I felt that everything I did required significant effort," and "I had difficulty initiating tasks." Participants who responded with "often" or "most of the time" to either of these questions were identified as having an impairment in the psychological domain [33]. To evaluate the cognition domain, 2 questions were used from the Short Portable Mental Status Questionnaire (SPMSQ) [37] related to spatial and temporal orientation. Failing to correctly answer either of these items was indicative of an impairment in the cognition domain.

Disability- and Dementia-Free Survival

The primary endpoint of this study was disability- and dementia-free survival, a composite outcome defined by the absence of physical disability, dementia, or mortality. We further calculated survival probability and transformed it into an HLI by multiplying by 100, yielding a 0‐100 scale that reflects the likelihood of maintaining physical and cognitive function throughout the observation period. In TLSA, each participant was monitored at each wave from the baseline year of 2003 until the occurrence of incident disability and dementia, death, emigration, loss of follow-up, or the end of the study period in 2015, whichever occurred first. Survival status was ascertained from multiple sources, including the Household Registration and Conscription Information System of the Ministry of the Interior, the Death Certificate Report System of the Ministry of Health and Welfare, and survey data collected by interviewers. To evaluate physical disability, we used the validated activities of daily living scale for each participant. This scale encompasses 6 essential tasks: bathing, dressing or undressing, eating, transferring (getting out of bed standing up or sitting in a chair), ambulating within the home, and toileting [38]. Physical disability was defined as the inability to perform, having severe difficulty in performing, or requiring assistance to complete at least 1 of these 6 basic activities of daily living tasks. Dementia was assessed using the SPMSQ [37]. At the baseline wave in 2003, the 9-item version was used, while a 10-item version was used in subsequent follow-up waves, as the question "When is your birthday?" was not collected in 2003. For participants with 6 or more years of education, a cut-off of 4 or more errors was used. For those with fewer than 6 years of education, the cut-off was 6 or more errors. These cut-off points demonstrated balanced sensitivity and specificity for detecting dementia and were highly correlated with MMSE scores [39].

In NILS-LSA, incident disability and dementia were defined as a care need of level 1 or higher, as defined by the long-term care insurance (LTCI) system used in Japan [40]. We calculated the person-years of follow-up for each participant from the first participation date until the (1) date of incident disability (LTCI certification), (2) the date of death, (3) the date of moving out of Obu and Higashiura, or (4) the end of the study period (January 31, 2022), whichever occurred first. A dataset that included information on LTCI certification, death, or emigration was obtained from Obu and Higashiura. All data were transferred from the Obu and Higashiura Governments under an agreement related to the Epidemiologic Research and Privacy Protection Guideline.

Statistical Analysis

The descriptive analysis of demographic and clinical features was performed. For categorical variables, the chi-square test was used to determine the statistical significance, whereas the Kruskal-Wallis test was used for continuous variables to assess the difference between groups.

Phase 1: Features Selection, Model Development, and Nomogram Construction for HLI

In the first phase of this study, we used a univariate Cox proportional hazards regression model to identify distinct potential features for 4-, 8-, and 12-year disability- and dementia-free survival to build up the HLI. A backward elimination with a *P* value threshold of .05 was then used to determine features at each follow-up to be included in the multivariate models. The method offers the advantage of evaluating the collective predictive capacity of the features

XSL•FO
RenderX

involved by initially incorporating all features into the model, systematically eliminating the least significant variables at an early stage, and ensuring only the most essential variables remain in the final model [41].

Significant features were subsequently integrated into the model that develops a nomogram for predicting the probability of disability- and dementia-free survival to build up the HLI. To be noteworthy, the ultimate features also underwent manual evaluation by clinical experts to ensure that no clinically important features were omitted during the aforementioned process.

To identify optimal survival models, we compared the semiparametric Cox proportional hazards model and a range of parametric models, including Weibull, exponential, log-logistic, and log-normal models. The selection of the most suitable model was performed using the Akaike information criterion (AIC). Models demonstrating lower AIC values were considered to have superior fit characteristics. Based on these criteria, the Weibull regression model was selected. The coefficients derived from this model were subsequently used to compute scores that were allocated to variables on the nomogram. We further calculated survival probability and then transformed it into an HLI by multiplying by 100, yielding a 0 - 100 scale.

The performance of the model was assessed through the Harrell concordance index (C-index) to evaluate the discriminative capacity. The Harrell C-index is an extension of the receiver operating characteristic curve specifically adapted for survival analysis and was frequently used in the validation of survival models in the context of their time-dependent properties [42], with the index higher than 0.7 suggesting good model performance. The calibration plot was constructed using smooth restricted cubic splines with 4 knots modified by a complementary log-log transformation. This method is often used in time-to-event data structures and helps reduce the number of knots and increase the likelihood of a linear relationship between the outcome probability and the linear predictor [43]. The 95% CI for the C-index was obtained from bootstrapping with 1000 resamplings to reduce overfitting bias. In addition, the Brier score was calculated to assess the overall performance of the predictive models, with bootstrapped 95% CIs derived from 1000 resamplings. The Brier score is calculated by squaring the largest difference between a predicted probability (which lies between 0 and 1) and the actual outcome (which is either 0 or 1). As a result, the score falls within the range of 0 to 1, where a lower Brier score indicates better calibration of the predictions. In addition, we generated Kaplan-Meier plots stratified into 3 groups according to tertiles of predicted probability (Tertile 1, Tertile 2, and Tertile 3) to enhance the examination of calibration accuracy. A log-rank test was used to assess differences in outcomes among these groups.

Phase 2: Internal Validation in TLSA and External Validation in Japan NILS-LSA Cohort

To calculate the probability of disability- and dementia-free survival, we first extracted coefficients derived from the prediction model of the TLSA (TLSA training cohort) and further applied them in the validation set from the remaining 30% of the TLSA cohort (TLSA validation cohort). To generalize the applicability of the proposed prediction model, we further used the NILS-LSA cohort from Japan to externally validate the model derived from the TLSA cohort. The overall model performance was assessed by using the Brier score, while the calibration plot and the C-statistic were obtained to determine the model's accuracy and discriminability, respectively. The operational definition of variables incorporated into the nomogram measured in the NILS-LSA cohort (external validation) generally corresponded to those in the TLSA cohort (Table S2 in Multimedia Appendix 2). Both the interval validation sets of the TLSA and NILS-LSA cohorts were categorized into tertiles based on the predicted probabilities (Tertile 1, Tertile 2, and Tertile 3), respectively.

Sensitivity Analysis

To obtain a more robust and unbiased estimate of the initial model's performance metrics, including the mean C-statistic and mean Brier score, we conducted a sensitivity analysis by using 10-fold cross-validation using the TLSA cohort across 4-, 8-, and 12-year follow-up periods. All statistical analyses were performed using SAS software (version 9.4; SAS Institute) and the R (version 4.3.3) statistical package. We considered a 2-tailed $P \leq .05$ to indicate statistical significance.

Ethical Considerations

Ethical approval was granted by the Research Ethics Committee of National Taiwan University Hospital (202204025RINA). In addition, the Committee on the Ethics of Human Research of the National Center for Geriatrics and Gerontology (NCGG), Japan, approved the study protocol (1665 - 3). Written informed consent was obtained from all participants before their inclusion in the study. All potentially identifying data were encrypted to protect anonymity. Only investigators who received approval from the principal investigator of each cohort and signed a data access agreement could access the data.

Results

Features of Study Participants

A total of 4470 community-dwelling participants in the TLSA were included for HLI model construction, and their characteristics are summarized in Table S3 in Multimedia Appendix 2. The cohort predominantly comprised individuals aged 50 - 64 years (n=2451, 54.8%), followed by those aged 65 - 74 years (n=986, 22.1%), and 75 years or older (n=1033, 23.1%). The cohort comprised 53.0% (n=2367) males, with over half residing in rural areas (n=2503, 56.0%), 18.1% (n=808) reporting illiteracy, and 67.7% (n=3028) engaging in regular physical activity. IC impairments were most prevalent in the psychological domain (n=939, 21%), followed by vitality (n=593, 13.3%), cognition (n=534, 12%), locomotion (n=608, 11.4%), visual acuity (n=500, 11.2%), and hearing (n=311, 7%). Cardiometabolic conditions were common in the study sample, with hypertension affecting 29.4% (n=1314) of participants, heart disease 13.7% (n=614), diabetes mellitus 11.7% (n=524), and stroke 1.8% (n=82).

For the study, the cohort was stratified into 2 distinct groups using simple random sampling: a training set comprising 70%

XSL•FO
RenderX

(n=3129) of the cohort (TLSA-training cohort), which was used for model construction, and a validation set constituting the remaining 30% (n=1341) for internal validation (TLSA-validation cohort). No significant differences were observed between the TLSA training and validation cohorts regarding baseline characteristics.

During the 12-year observation period, 1131 (36.1%) participants in the TLSA-training cohort were documented as deceased or with incident physical disability or dementia, with 556 (13.6%) participants at 4 years and 821 (26.2%) participants by the 8-year observation period. Similarly, in the TLSA-validation cohort, 517 (38.5%) participants were recorded with the same outcomes over 12 years, with 230 (14.6%) participants at 4 years and 261 (26.2%) participants by the 8-year observation period. In the full dataset, the number of participants with unknown survival status or who had moved out was 305 (6.8%), 492 (11.0%), and 679 (15.2%) at the 4-, 8-, and 12-year follow-up, respectively.

Model Specification and Features Associated With HLI

Features selection in TLSA training cohort using both univariate and multivariate Cox hazard regression analyses at the 12-year follow-up are provided in Table S4 in Multimedia Appendix 2. After excluding nonsignificant features from the univariate analysis and applying backward elimination, the multivariate analysis was conducted. The results indicated that older adults (65 - 74 years: HR 2.85, 95% CI 2.40 - 3.37; older than 75 years: HR 6.32, 95% CI 5.37 - 7.42), female (HR 0.74, 95% CI 0.64 - 0.85), higher level of education (elementary: HR 0.83, 95% CI 0.71 - 0.98; junior or senior high school: HR 0.74, 95% CI 0.60 - 0.91; college and above: HR 0.54, 95% CI 0.40 - 0.71), current smoker (HR 1.39, 95% CI 1.20 - 1.62), current drinker (HR 0.81, 95% CI 0.70 - 0.94), having difficulty in climbing stairs (HR 1.24, 95% CI 1.04 - 1.47), and IC impairment (locomotion: HR 1.32, 95% CI 1.10 - 1.59; visual acuity: HR 1.21, 95% CI 1.03 - 1.42; cognition: HR 1.20, 95% CI 1.02 - 1.42) were significantly associated with 12-year disability- and dementia-free survival probability for building up HLI. Furthermore, cardiometabolic conditions were also associated with 12-year disability- and dementia-free survival, including hypertension (HR 1.14, 95% CI 1.01 - 1.29), diabetes (HR 1.54, 95% CI 1.32 - 1.81), and stroke (HR 1.57, 95% CI 1.14 - 2.14). Liver disease (HR 1.54, 95% CI 1.24 - 1.92) and renal disease (HR 1.46, 95% CI 1.17 - 1.82) were also indicative of the outcome. Details of the features selected and their association with 4-year and 8-year disability- and dementia-free survival for building up HLI are also provided in Table S5 and Table S6 in Multimedia Appendix 2.

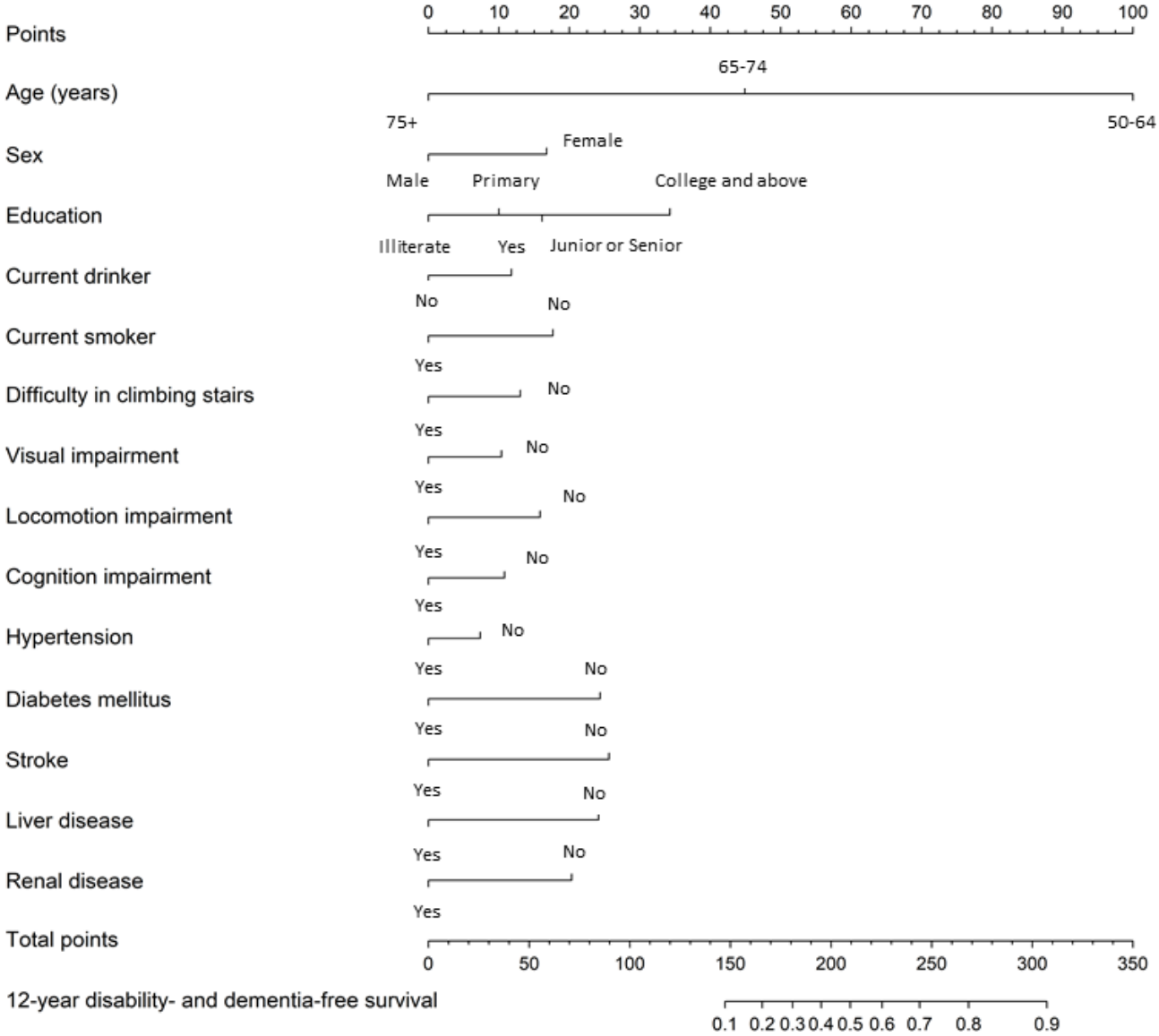
The goodness of fit was compared between the Cox regression model and 4 types of parametric survival models, including Weibull, exponential, log-logistic, and log-normal. The Weibull model had the lowest AIC across the 4-, 8-, and 12-year observation periods, denoting a better overall fit than the other models (Table S7 in Multimedia Appendix 2; Figures S2, S3, and S4 in Multimedia Appendix 1).

Nomogram Development and Validation

The features selected to construct the 12-year probability of disability- and dementia-free survival nomogram from the TLSA training cohort using the Weibull model were used to develop a nomogram. These features included demographics (age, sex, level of education, current smoking status, and current drinking status), functional status indicator (difficulty in climbing stairs), and IC (locomotion impairment, visual acuity, and cognition impairment). In addition, self-reported chronic diseases such as hypertension, diabetes mellitus, stroke, liver disease, and renal disease were incorporated (Figure 2). Nomograms for the 4-year and 8-year probabilities of disability- and dementia-free survival are presented in Figures S5 and S6 in Multimedia Appendix 1, respectively.

XSL•FO
RenderX

Figure 2. Nomogram for 12-year probability of disability- and dementia-free survival.



Consequently, the nomogram assigns points to each feature, which are used to calculate the probability of disability- and dementia-free survival, and then transformed into a HLI by multiplying by 100, yielding a 0 - 100 scale. For example, a 59-year-old female with primary school education, locomotion impairment, and visual impairment would accumulate a score of 270, with scores distributed as follows: 100 points for age (50 - 64 years), 17 points for being female, 10 points for having primary school education, 18 points for not smoking, 13 points for having no difficulty in climbing stairs, 11 points for having no cognitive impairment, 7 points for having no hypertension, 24 points for having no diabetes mellitus, 26 points for having no stroke, 24 points for having no liver disease, and 20 points for having no renal disease. Overall, this score predicts a 12-year HLI of 80.6. The formulas, model parameters, and the detailed weighting scores are provided in Table S8 in Multimedia Appendix 2.

The model demonstrated robust performance at the 12-year follow-up period, with consistent results observed between the TLSA training and validation cohort (Table 2). In the TLSA training cohort, the C-statistic was 0.79 (bootstrapped 95% CI 0.78 - 0.80) and the Brier score was 0.16 (bootstrapped 95% CI 0.15 - 0.17) at the 12-year follow-up period. Similarly, the TLSA validation cohort showed a C-statistic of 0.77 (bootstrapped 95% CI 0.75 - 0.79) and a Brier score of 0.17 (bootstrapped 95% CI 0.16 - 0.18), showing good discriminability. Calibration plots were used to visually assess the model's accuracy (Figures S9 in Multimedia Appendix 1). The graphs showed good agreement in both the TLSA training and TLSA validation cohorts, indicating a well-calibrated model. In addition, the graphical presentations also confirmed good accuracy for both the TLSA training and TLSA internal validation cohorts throughout the 4-year and 8-year observation period, as provided in Figures S7 and S8 in Multimedia Appendix 1.

XSL•FO
RenderX

Table . Model performance for disability- and dementia-free survival at 4-, 8-, and 12-follow-up periods in the Taiwan Longitudinal Study on Aging (TLSA) and the National Institute for Longevity Sciences, Longitudinal Study of Aging (NILS-LSA).

Follow-up	TLSA training cohort	TLSA validation cohort	NILS-LSA validation cohort
4-year			
C-statistic (95% CI)	0.80 (0.78-0.81)	0.78 (0.74 - 0.81)	0.71 (0.66-0.76)
Brier score (95% CI)	0.10 (0.09-0.11)	0.10 (0.09-0.12)	0.07 (0.06-0.08)
8-year			
C-statistic (95% CI)	0.79 (0.77-0.81)	0.77 (0.74-0.79)	0.70 (0.67-0.73)
Brier score (95% CI)	0.14 (0.13-0.15)	0.15 (0.14-0.17)	0.16 (0.15-0.17)
12-year			
C-statistic (95% CI)	0.79 (0.78-0.80)	0.77 (0.75-0.79)	0.71 (0.68-0.73)
Brier score (95% CI)	0.16 (0.15-0.17)	0.17 (0.16-0.18)	0.20 (0.19-0.21)

We further plotted Kaplan-Meier curves based on the tertiles of the 12-year predicted HLI, as generated by the nomogram. In the TLSA training cohort, the overall median predicted HLI was 74. “Tertile 3” exhibited the poorest actual survival outcomes, with a median 12-year HLI of 27.8, compared to 74.7 in “Tertile 2” and 87.8 in “Tertile 1.” Similar patterns were observed in the TLSA internal validation cohort, where the median 12-year HLI was 28.1 for “Tertile 3,” and 73.5 and 87.8 for “Tertile 2” and “Tertile 1,” respectively (Figures S10 in Multimedia Appendix 1).

External Validation (NILS-LSA Cohort)

The models constructed were further validated in the NILS-LSA cohort from Japan using Brier score, the C-statistic, and graphical calibration plots (Figures S11 in Multimedia Appendix 1). The comparison of features included in the models between TLSA and NILS-LSA is provided in Table 3. In the NILS-LSA cohort, the statistics demonstrated good model performance with a C-statistic of 0.71 (bootstrapped 95% CI 0.66 - 0.76) and a Brier score of 0.07 (bootstrapped 95% CI 0.06 - 0.08) at the 4-year follow-up period. Similar results were also observed for the 8- and 12-year disability- and dementia-free survival periods (Table 2).

XSL•FO
RenderX

Table . Comparison of variables included in final models between the Taiwan Longitudinal Study on Aging (TLSA) and the National Institute for Longevity Sciences, Longitudinal Study of Aging (NILS-LSA).

Variable	TLSA training cohort (n=3129), n (%)	TLSA validation cohort (n=1341), n (%)	NILS-LSA validation cohort (n=1090), n (%)
Demographics			
Age (years)			
50 - 64	1715 (54.8)	736 (54.9)	0 (0)
65 - 74	681 (21.8)	305 (22.7)	756 (69.4)
75+	733 (23.4)	300 (22.4)	334 (30.6)
Sex			
Female	1475 (47.1)	628 (46.8)	554 (50.8)
Education			
Illiterate	556 (17.8)	252 (18.8)	0 (0)
Elementary	1511 (48.3)	635 (47.4)	20 (1.8)
Junior or senior	751 (24.0)	320 (23.9)	799 (73.3)
College and above	311 (9.9)	134 (10.0)	271 (24.9)
Life behaviors			
Current smoker	722 (23.1)	306 (22.8)	137 (12.6)
Current alcohol drinker	1000 (32.0)	430 (32.1)	529 (48.5)
Functional status			
Difficulty in climbing stairs	484 (15.5)	209 (15.6)	123 (11.3)
Intrinsic capacity			
Locomotion impairment	350 (11.2)	158 (11.8)	104 (9.5)
Visual acuity impairment	341 (10.9)	159 (11.9)	169 (15.5)
Vitality impairment	401 (12.8)	192 (14.3)	188 (17.2)
Psychological impairment	643 (20.6)	296 (22.1)	94 (8.6)
Cognition impairment	373 (11.9)	161 (12.0)	109 (10.0)
Self-reported chronic conditions			
Hypertension	926 (29.6)	388 (28.9)	453 (41.6)
Diabetes mellitus	370 (11.8)	154 (11.5)	111 (10.2)
Stroke	59 (1.9)	23 (1.7)	64 (5.9)
Liver disease	182 (5.8)	93 (6.9)	65 (6.0)
Renal disease	160 (5.1)	78 (5.8)	60 (5.5)

We also plotted Kaplan-Meier curves based on the tertiles of the 12-year HLI in the NILS-LSA cohort (Figure S10 in Multimedia Appendix 1). The overall median predicted 12-year HLI was 59.5. “Tertile 3” exhibited a median 12-year HLI of 32.6, compared to 59.5 in “Tertile 2” and 71.5 in “Tertile 1.”

Developing the Real-Time, Personalized HLI Platform

A user-friendly app was developed to optimize accessibility and facilitate real-world implementation. The real-time model is presented in a dashboard format, as provided in Figure 3. This personalized application calculates the HLI using the prediction model of specific timeframes, stratifying individuals into 3 distinct risk groups: high risk (Tertile 1, red light), medium risk (Tertile 2, yellow light), and low risk (Tertile 3, green light). The electronic interface is designed for use by clinical practitioners and researchers, facilitating decision-making within clinical settings and enabling personalized lifestyle modifications in response to fluctuating HLI levels.

XSL•FO
RenderX

Figure 3. Real-time dashboard of personalized Healthy Longevity Index.



Sensitivity Analysis

We conducted 10-fold cross-validation in estimating the initial model's performance metrics using the entire TLSA cohort. The results aligned with our previous findings across the 4-, 8-, and 12-year follow-up periods. Specifically, at the 4-year follow-up, the C-statistic was 0.79 (95% CI 0.77 - 0.81), with a Brier score of 0.10 (95% CI 0.09 - 0.11). At the 8-year follow-up, the C-statistic was 0.78 (95% CI 0.77 - 0.79), and the Brier score was 0.15 (95% CI 0.14 - 0.15). Similarly, at the 12-year follow-up, the C-statistic was 0.78 (95% CI 0.77 - 0.79), while the Brier score increased to 0.16 (95% CI 0.16 - 0.17).

Discussion

Principal Findings

This study introduces a novel, holistic HLI designed to promote healthy aging through real-time assessments, lifestyle optimization, and the development of personalized care plans. The HLI predicts the probability of maintaining active physical and cognitive function within specific timeframes. Data from 2 large-scale national longitudinal cohorts, the TLSA and the NILS-LSA in Japan, were used for HLI development and validation (internal and external). The nomogram integrates a multifaceted array of predictors, including demographic characteristics, lifestyle behaviors, IC measures, and chronic conditions, to generate individualized risk estimates suitable for primary care implementation. Notably, the model demonstrated robust predictive performance, with C-indices of 0.79 (95% CI 0.78 - 0.80) in the TLSA training cohort and 0.71 (95% CI 0.66 - 0.76) in the NILS-LSA cohort (external validation), indicating good discrimination across diverse populations. Furthermore, the electronic HLI dashboard generates a real-time individual HLI score. Primary care providers and their patients can periodically monitor this score to assess their potential for healthy longevity—defined as the transformed probability of disability- and dementia-free survival within specified timeframes—and to evaluate the impact of lifestyle modifications or targeted intervention plans.

Comparison With Previous Work

Early detection and continuous monitoring through appropriate biomarkers have become fundamental to modern disease management, historically relying on periodic biomarker assessments to inform health care decision-making. Recent advancements in biomedical and digital technologies, coupled with the availability of population-scale datasets, have created the potential for real-time or continuous biomarker monitoring, promising a transformative impact on health management. While the efficacy of biosensors, wearable devices, and continuous glycemic monitoring in enhancing clinical outcomes has been well established, a comparable paradigm for the WHO's healthy aging framework remains elusive. Despite the WHO's intention to promote self-assessment, IC evaluation still heavily relies on health care professionals in community settings [44]. The HLI addresses this gap by providing primary care teams with a practical tool that can be integrated into routine consultations. Conversely, the HLI in this study was derived from questionnaire-based data (including IC assessment) and fundamental health behavior metrics, potentially serving as an effective marker for real-time monitoring to promote healthy aging via self-assessments. By incorporating a multitude of modifiable factors, including lifestyle behaviors and health conditions, the HLI enables individuals and health care providers to target interventions toward these modifiable factors, thereby enhancing the potential for healthy aging.

Regarding real-world implementation, growing evidence supports the feasibility and acceptance of digital health assessment tools among older adults in community settings. Studies have demonstrated that older adults are increasingly receptive to smartphone-based health apps and web-based questionnaire tools when they provide meaningful health information [45]. Furthermore, questionnaire-based screening tools similar to our HLI approach have been successfully implemented in primary care settings and community health programs, with evidence suggesting that older adults value receiving personalized health risk information when it is presented in an understandable format and coupled with guidance for improvement [46]. The integration of our HLI dashboard with existing health care workflows and community wellness programs could facilitate adoption while providing necessary support for interpreting results and developing action plans. In addition, research on health behavior change in older adults suggests that personalized risk feedback, when delivered appropriately, can serve as an effective motivator for lifestyle modifications rather than a source of distress, particularly when combined with self-efficacy building and social support [47].

XSL•FO
RenderX

However, while these studies demonstrate technical feasibility and user acceptance, further implementation research is needed to evaluate the real-world impact of tools such as the HLI on actual health outcomes, behavior change sustainability, and long-term well-being in diverse older adult populations.

Building on this demonstrated acceptance among older adults, an important clinical application of the HLI lies in its potential to guide referral decisions and care pathways as individuals' risk profiles change over time. Our tertile-based risk stratification system could serve as objective thresholds to support primary care clinicians in making timely referrals for specialist assessment [48]. For instance, individuals transitioning from low-risk to moderate- or high-risk categories may benefit from comprehensive geriatric assessment, referral to memory clinics for cognitive evaluation, or enrollment in structured fall prevention programs [49]. Furthermore, the granular assessment of IC domains within the HLI could help primary care providers identify which specific specialists (geriatricians, neurologists, or rehabilitation specialists) would be most appropriate for referral based on the individual's risk profile [50]. While we do not yet have a well-designed clinical protocol for HLI-based referrals, prioritizing interventions that target modifiable risk factors—such as lifestyle behaviors, physical activity, IC maintenance, and chronic disease management—should remain the primary focus for translating risk assessment into actionable clinical care. However, formal validation of HLI-based referral thresholds and their impact on clinical outcomes, health care usage, and cost-effectiveness will be essential before widespread clinical implementation [51].

The HLI may be implemented in primary health care settings and preventive screening among community-dwelling older adults, where the goal is early identification and prevention rather than management of established disease. This implementation context aligns well with our study population and addresses the growing need for proactive risk assessment in preventive care [51,52]. For clinical interpretation in primary care, physicians should use HLI scores as decision support for preventive interventions, health promotion activities, and determining appropriate monitoring intensity [53]. Integration with electronic health records in primary care settings will enable seamless incorporation into routine wellness visits and preventive care workflows [54,55]. While the current HLI is optimized for community-dwelling older adults, the methodological principles could be adapted to develop similar instruments for populations with specific diseases or conditions; however, such applications would require separate development and validation studies.

Among the modifiable risk factors encapsulated within the HLI, IC holds promise for addressing population health challenges related to aging within primary care systems. A recent systematic review and meta-analysis, encompassing 37 studies with a total of 206,693 participants, established a significant inverse relationship between IC and functional declines, as well as mortality risk, in older adults. These findings support the potential of IC as a biomarker of healthy aging; however, further research is necessary to refine its measurement and application within diverse populations and settings [56]. The utility of IC as a biomarker for healthy aging is limited by the inconclusive causal link between IC and DALYs, and the challenges associated with self-assessment and monitoring of IC. The INSPIRE Project has established age- and sex-specific reference centiles for IC using data from 975 adults aged 20 - 102 years to offer a framework for individuals to assess their healthy aging trajectory, but the limitations to developing real-time personalized assessment persist. The calculation of centiles and the comparative nature of individual scores within the cohort do not provide explicit probabilities of healthy aging. Furthermore, evaluating the efficacy of intervention programs using this approach presents methodological challenges [18,57].

Limitations and Future Directions

Despite the development and validation of the HLI as a tool for web-based, real-time, personalized assessment of healthy aging status, several limitations warrant consideration. First, the relatively healthy profiles of participants in the TLSA and NILS-LSA cohorts may have introduced a selection bias toward more favorable outcomes. However, it universally exists in all cohort studies. Second, while this study aligned with the WHO's conceptualization of healthy aging by focusing on physical and cognitive capacity, the exclusion of social and other mental well-being dimensions necessitates further research to establish a comprehensive, consensual operational definition of healthy aging. Third, while both the cognition predictor and dementia outcome derive from the SPMSQ, we used only 2 orientation questions for the predictor versus the full SPMSQ score with education-adjusted cutoffs for the outcome. Although this represents different levels of cognitive assessment, we acknowledge this methodological approach may have some impact on results for cognitive-related outcomes. Future studies using independent cognitive assessment tools would provide stronger validation of these findings. Fourth, the integration of novel digital biomarkers, such as activity patterns, voice, and facial expression recognition, into the HLI framework would further enhance real-time risk assessment and intervention monitoring. Fifth, future studies should evaluate the HLI's implementation in diverse primary care settings and assess its impact on clinical decision-making, patient outcomes, and health care resource usage.

Conclusion

This study successfully developed and validated the HLI using large-scale nationally representative cohort data from Taiwan and Japan, which demonstrates potential as a real-time and personalized biomarker for healthy aging in primary care practice. By providing a more dynamic monitoring approach, the index offers greater responsiveness to interventions and lifestyle modifications in real-world settings. The integration of readily accessible digital biomarkers from wearable devices is anticipated to further refine the index's accuracy and utility as a mobile tool for promoting healthy longevity within primary care systems.

XSL•FO
RenderX

Acknowledgments

We thank all members of the National Center for Geriatrics and Gerontology (NCGG), Japan, for their contributions and efforts in making the databases used in this study available. A list of members can be found on the official NCGG website. This research was funded by the Taiwan Ministry of Science and Technology (MOST 110-2634-F-010-001), the Taiwan National Science and Technology Council (NSTC 113-2321-B-A49-028- and NSTC 114-2321-B-A49-012-), and the Interdisciplinary Research Center for Healthy Longevity, National Yang Ming Chiao Tung University, under the Featured Areas Research Center Program within the Higher Education Sprout Project, Ministry of Education (MOE) Taiwan. The funding sources had no role in the conduct of this study, including study design, data collection and analysis, manuscript preparation and review, and the decision to submit the manuscript for publication. During the preparation of this work, the authors declare to have used Claude Sonnet 4 to refine and polish the writing.

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Authors' Contributions

All authors meet the criteria for authorship as stated in the Uniform Requirements for Manuscripts submitted to Biomedical Journals. HYL, FYH, and LKC designed the research. HYL and SZ analyzed the data. HYL, SZ, FYH, and LKC wrote the original draft. HYL, SZ, RO, HA, FYH, and LKC contributed to the conceptualization and critical revision of the manuscript. RO, STH, HA, FYH, and LKC provided clinical insights.

FYH is the co-corresponding author of this paper and can be reached at Graduate Institute of Clinical Pharmacy, College of Medicine, National Taiwan University; No. 33 Linsen S. Road, Taipei, 10050, Taiwan; +886 2-33668787; mailto:fyshsiao@ntu.edu.twfyshsiao@ntu.edu.tw.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional figures.
[PDF File, 782 KB - aging_v8i1e80034_app1.pdf]

Multimedia Appendix 2
Additional tables.
[PDF File, 689 KB - aging_v8i1e80034_app2.pdf]

References

1. Beard JR, Officer A, de Carvalho IA, et al. The world report on ageing and health: a policy framework for healthy ageing. The Lancet 2016 May;387(10033):2145-2154. [doi: 10.1016/S0140-6736(15)00516-4]
2. Ferrari AJ, Santomauro DF, Aali A, et al. Global incidence, prevalence, years lived with disability (YLDs), disability-adjusted life-years (DALYs), and healthy life expectancy (HALE) for 371 diseases and injuries in 204 countries and territories and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. The Lancet 2024 May;403(10440):2133-2161. [doi: 10.1016/S0140-6736(24)00757-8]
3. Integrated care for older people: guidelines on community-level interventions to manage declines in intrinsic capacity. : World Health Organization (WHO); 2017.
4. Lee WJ, Peng LN, Lin MH, Loh CH, Hsiao FY, Chen LK. Intrinsic capacity differs from functional ability in predicting 10-year mortality and biological features in healthy aging: results from the I-Lan longitudinal aging study. Aging (Milano) 2023 Feb 14;15(3):748-764. [doi: 10.18632/aging.204508]
5. Lu WH, Rolland Y, Guyonnet S, de Souto Barreto P, Vellas B. Reference centiles for intrinsic capacity throughout adulthood and their association with clinical outcomes: a cross-sectional analysis from the INSPIRE-T cohort. Nat Aging 2023 Dec;3(12):1521-1528. [doi: 10.1038/s43587-023-00522-x] [Medline: 37946044]
6. Lee WJ, Peng LN, Lin MH, Loh CH, Hsiao FY, Chen LK. Intrinsic capacity and multimorbidity predicting incident disability-Insights from the I-Lan Longitudinal Aging Study. Arch Gerontol Geriatr 2024 Jun;121:105357. [doi: 10.1016/j.archger.2024.105357] [Medline: 38340587]
7. Meng LC, Chuang HM, Lu WH, Lee WJ, Liang CK, Loh CH, et al. Multi-trajectories of intrinsic capacity decline and their impact on age-related outcomes: a 20-year national longitudinal cohort study. Aging Dis 2023 Nov 21;15(6):2697-2709. [doi: 10.14336/AD.2023.1115-1] [Medline: 38029399]

XSL•FO
RenderX

8. Meng LC, Hsiao FY, Huang ST, Lu WH, Peng LN, Chen LK. Intrinsic capacity impairment patterns and their associations with unfavorable medication utilization: a nationwide population-based study of 37,993 community-dwelling older adults. J Nutr Health Aging 2022;26(10):918-925. [doi: 10.1007/s12603-022-1847-z] [Medline: 36259580]
9. Meng LC, Huang ST, Peng LN, Chen LK, Hsiao FY. Biological features of the outcome-based intrinsic capacity composite scores from a population-based cohort study: pas de deux of biological and functional aging. Front Med (Lausanne) 2022;9:851882. [doi: 10.3389/fmed.2022.851882] [Medline: 35308493]
10. Guan ST, Lai HY, Chen LK, Hsiao FY. Advancing nutrition risk assessment in middle-aged and older individuals with diverse food cultures: a data-driven personalized approach to predict incident hypertension, diabetes and mortality. J Nutr Health Aging 2024 Feb;28(2):100025. [doi: 10.1016/j.jnha.2023.100025] [Medline: 38218677]
11. Perez MV, Mahaffey KW, Hedlin H, et al. Large-scale assessment of a smartwatch to identify atrial fibrillation. N Engl J Med 2019 Nov 14;381(20):1909-1917. [doi: 10.1056/NEJMoa1901183] [Medline: 31722151]
12. Karter AJ, Parker MM, Moffet HH, Gilliam LK, Dlott R. Association of real-time continuous glucose monitoring with glycemic control and acute metabolic events among patients with insulin-treated diabetes. JAMA 2021 Jun 8;325(22):2273-2284. [doi: 10.1001/jama.2021.6530] [Medline: 34077502]
13. Dominguez LJ, Veronese N, Barbagallo M. Dietary patterns and healthy or unhealthy aging. Gerontology 2024;70(1):15-36. [doi: 10.1159/000534679] [Medline: 37883935]
14. Chen X, Giles J, Yao Y, et al. The path to healthy ageing in China: a Peking University–Lancet Commission. The Lancet 2022 Dec;400(10367):1967-2006. [doi: 10.1016/S0140-6736(22)01546-X]
15. Yu YY, Lai HY, Loh CH, Chen LK, Hsiao FY. Multi-trajectories of physical activity and health in aging: a 20-year nationwide population-based longitudinal study. Arch Gerontol Geriatr 2025 Aug;135:105878. [doi: 10.1016/j.archger.2025.105878] [Medline: 40349593]
16. Hsiao FY, Peng LN, Lee WJ, Chen LK. Higher dietary diversity and better healthy aging: a 4-year study of community-dwelling middle-aged and older adults from the Taiwan Longitudinal Study of Aging. Exp Gerontol 2022 Oct 15;168:111929. [doi: 10.1016/j.exger.2022.111929] [Medline: 35977645]
17. Lee WJ, Peng LN, Lin CH, et al. Effects of incorporating multidomain interventions into integrated primary care on quality of life: a randomised controlled trial. Lancet Healthy Longev 2021 Nov;2(11):e712-e723. [doi: 10.1016/S2666-7568(21)00248-8] [Medline: 36098028]
18. Liang CK, Lee WJ, Chou MY, et al. Roles of baseline intrinsic capacity and its subdomains on the overall efficacy of multidomain intervention in promoting healthy aging among community-dwelling older adults: analysis from a nationwide cluster-randomized controlled trial. J Prev Alzheimers Dis 2024;11(2):356-365. [doi: 10.14283/jpad.2024.20] [Medline: 38374742]
19. Ahn SH, Lee JH, Lee JH. Body mass index changes and femur fracture risk in Parkinson's disease: national cohort study. J Cachexia Sarcopenia Muscle 2025 Jun;16(3):e13860. [doi: 10.1002/jcsm.13860] [Medline: 40464204]
20. Zhou Z, Curtis AJ, Owen A, et al. Prediabetes, diabetes and loss of disability-free survival in a community-based older cohort: a post-hoc analysis of the ASPirin in reducing events in the elderly trial. Age Ageing 2023 Apr 1;52(4):afad060. [doi: 10.1093/ageing/afad060] [Medline: 37097768]
21. McNeil JJ, Woods RL, Nelson MR, et al. Effect of aspirin on disability-free survival in the healthy elderly. N Engl J Med 2018 Oct 18;379(16):1499-1508. [doi: 10.1056/NEJMoa1800722] [Medline: 30221596]
22. Neumann JT, Thao LTP, Murray AM, et al. Prediction of disability-free survival in healthy older people. Geroscience 2022 Jun;44(3):1641-1655. [doi: 10.1007/s11357-022-00547-x]
23. Campbell BCV, Mitchell PJ, Churilov L, et al. Endovascular thrombectomy for ischemic stroke increases disability-free survival, quality of life, and life expectancy and reduces cost. Front Neurol 2017;8:657. [doi: 10.3389/fneur.2017.00657] [Medline: 29312109]
24. Integrated care for older people (ICOPE): guidance for person-centred assessment and pathways in primary care, 2nd ed. World Health Organization (WHO). URL: https://www.who.int/publications/i/item/9789240103726 [accessed 2025-11-01]
25. Taiwan Longitudinal Study on Aging (TLSA). Taiwan Health Promotion Administration. URL: https://www.hpa.gov.tw/EngPages/Detail.aspx?nodeid=1077&pid=6197 [accessed 2025-10-30]
26. Zhang S, Peng LN, Otsuka R, et al. Comparative analysis of intrinsic capacity impairments, determinants, and clinical consequences in older community-dwellers in Japan and Taiwan: longitudinal studies showing shared traits and distinct presentations. J Nutr Health Aging 2023;27(11):1038-1046. [doi: 10.1007/s12603-023-2020-z] [Medline: 37997726]
27. Otsuka R, Zhang S, Tange C, et al. Association of dietary intake with the transitions of frailty among Japanese community-dwelling older adults. J Frailty Aging 2022;11(1):26-32. [doi: 10.14283/jfa.2021.42] [Medline: 35122087]
28. Huang ST, Tange C, Otsuka R, Nishita Y, Peng LN, Hsiao FY, et al. Subtypes of physical frailty and their long-term outcomes: a longitudinal cohort study. J Cachexia Sarcopenia Muscle 2020;11(5):1223-1231. [doi: 10.1002/jcsm.12577]
29. Shimokata H, Ando F, Niino N. A new comprehensive study on aging--the National Institute for Longevity Sciences, Longitudinal Study of Aging (NILS-LSA). J Epidemiol 2000 Apr;10(1 Suppl):S1-S9. [doi: 10.2188/jea.10.1sup_1] [Medline: 10835822]
30. Integrated care for older people (ICOPE) implementation framework: guidance for systems and services. World Health Organization (WHO). URL: https://apps.who.int/iris/handle/10665/325669 [accessed 2025-10-30]

XSL•FO
RenderX

31. Nagi SZ, Marsh J. Disability, health status, and utilization of health services. Int J Health Serv 1980;10(4):657-676. [doi: 10.2190/81FB-J0X4-R9EY-06PV] [Medline: 7429695]
32. Hwang AC, Lee WJ, Huang N, et al. Longitudinal changes of frailty in 8 years: comparisons between physical frailty and frailty index. BMC Geriatr 2021 Dec 18;21(1):726. [doi: 10.1186/s12877-021-02665-1] [Medline: 34922488]
33. Chu WM, Ho HE, Yeh CJ, et al. Self-rated health trajectory and frailty among community-dwelling older adults: evidence from the Taiwan Longitudinal Study on Aging (TLSA). BMJ Open 2021 Aug 6;11(8):e049795. [doi: 10.1136/bmjopen-2021-049795] [Medline: 34362805]
34. Aliberti MJR, Bertola L, Szlejf C, et al. Validating intrinsic capacity to measure healthy aging in an upper middle-income country: findings from the ELSI-Brazil. Lancet Reg Health Am 2022 Aug;12:100284. [doi: 10.1016/j.lana.2022.100284] [Medline: 36776430]
35. Liljas AEM, Carvalho LA, Papachristou E, et al. Self-reported vision impairment and incident prefrailty and frailty in English community-dwelling older adults: findings from a 4-year follow-up study. J Epidemiol Community Health 2017 Nov;71(11):1053-1058. [doi: 10.1136/jech-2017-209207] [Medline: 28798152]
36. Liljas AEM, Carvalho LA, Papachristou E, et al. Self‐reported hearing impairment and incident frailty in English community‐dwelling older adults: a 4‐year follow‐up study. J American Geriatrics Society 2017 May;65(5):958-965 [FREE Full text] [doi: 10.1111/jgs.14687]
37. Pfeiffer E. A Short Portable Mental Status Questionnaire for the assessment of organic brain deficit in elderly patients†. J American Geriatrics Society 1975 Oct;23(10):433-441 [FREE Full text] [doi: 10.1111/j.1532-5415.1975.tb00927.x]
38. Edemekong PF, Bomgaars DL, Sukumaran S, Schoo C. Activities of Daily Living: StatPearls Publishing; 2024. URL: https://www.ncbi.nlm.nih.gov/books/NBK470404/ [accessed 2025-10-30]
39. Malhotra C, Chan A, Matchar D, Seow D, Chuo A, Do YK. Diagnostic performance of Short Portable Mental Status Questionnaire for screening dementia among patients attending cognitive assessment clinics in Singapore. Ann Acad Med Singap 2013 Jul;42(7):315-319. [doi: 10.47102/annals-acadmedsg.V42N7p315] [Medline: 23949260]
40. Standardized assessment manual for the levels of activity daily living deficiencies among elderly with dementia. : Ministry of Health; 1993.
41. Chowdhury MZI, Turin TC. Variable selection strategies and its importance in clinical prediction modelling. Fam Med Community Health 2020;8(1):e000262. [doi: 10.1136/fmch-2019-000262] [Medline: 32148735]
42. Pencina MJ, D'Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. Stat Med 2004 Jul 15;23(13):2109-2123. [doi: 10.1002/sim.1802]
43. Austin PC, Harrell FE Jr, van Klaveren D. Graphical calibration curves and the integrated calibration index (ICI) for survival models. Stat Med 2020 Sep 20;39(21):2714-2742. [doi: 10.1002/sim.8570]
44. de Souto Barreto P, Gonzalez-Bautista E, Bischoff-Ferrari HA, et al. Real-life intrinsic capacity screening data from the ICOPE-Care program. Nat Aging 2024 Sep;4(9):1279-1289. [doi: 10.1038/s43587-024-00684-2] [Medline: 39122839]
45. Mace RA, Mattos MK, Vranceanu AM. Older adults can use technology: why healthcare professionals must overcome ageism in digital health. Transl Behav Med 2022 Dec 30;12(12):1102-1105. [doi: 10.1093/tbm/ibac070] [Medline: 36073770]
46. Lau LK, Lun P, Gao J, Tan E, Ding YY. Application and implementation of brief geriatric assessment in primary care and community settings: a scoping review. BMC Geriatr 2025 Jan 2;25(1):2. [doi: 10.1186/s12877-024-05615-9] [Medline: 39748279]
47. Manuel DG, Abdulaziz KE, Perez R, Beach S, Bennett C. Personalized risk communication for personalized risk assessment: real world assessment of knowledge and motivation for six mortality risk measures from an online life expectancy calculator. Inform Health Soc Care 2018 Jan;43(1):42-55. [doi: 10.1080/17538157.2016.1255632] [Medline: 28068156]
48. Dom Dera J. Risk stratification: a two-step process for identifying your sickest patients. Fam Pract Manag 2019;26(3):21-26. [Medline: 31083868]
49. Tatum Iii PE, Talebreza S, Ross JS. Geriatric assessment: an office-based approach. Am Fam Physician 2018 Jun 15;97(12):776-784. [Medline: 30216018]
50. Bell SP, Saraf A. Risk stratification in very old adults: how to best gauge risk as the basis of management choices for patients aged over 80. Prog Cardiovasc Dis 2014;57(2):197-203. [doi: 10.1016/j.pcad.2014.08.001] [Medline: 25216619]
51. Field TS, Rochon P, Lee M, et al. Costs associated with developing and implementing a computerized clinical decision support system for medication dosing for patients with renal insufficiency in the long-term care setting. J Am Med Inform Assoc 2008;15(4):466-472. [doi: 10.1197/jamia.M2589] [Medline: 18436908]
52. Taylor DA, Wallis KA, Feki S, et al. Cardiovascular disease risk prediction in older people: a qualitative study. Br J Gen Pract 2021 Oct;71(711):e772-e779. [doi: 10.3399/BJGP.2020.1038] [Medline: 34019484]
53. Kong LN, Yang L, Lyu Q, Liu DX, Yang J. Risk prediction models for frailty in older adults: a systematic review and critical appraisal. Int J Nurs Stud 2025 Jul;167:105068. [doi: 10.1016/j.ijnurstu.2025.105068] [Medline: 40184783]
54. Tan A, Kuo YF, Goodwin JS. Predicting life expectancy for community-dwelling older adults from Medicare claims data. Am J Epidemiol 2013 Sep 15;178(6):974-983. [doi: 10.1093/aje/kwt054] [Medline: 23851579]
55. Peng LN, Hsiao FY, Lee WJ, Huang ST, Chen LK. Comparisons between hypothesis- and data-driven approaches for multimorbidity frailty index: a machine learning approach. J Med Internet Res 2020 Jun 11;22(6):e16213. [doi: 10.2196/16213] [Medline: 32525481]

XSL•FO
RenderX

56. Sánchez-Sánchez JL, Lu WH, Gallardo-Gómez D, et al. Association of intrinsic capacity with functional decline and mortality in older adults: a systematic review and meta-analysis of longitudinal studies. Lancet Healthy Longev 2024 Jul;5(7):e480-e492. [doi: 10.1016/S2666-7568(24)00092-8] [Medline: 38945130]
57. Lee WJ, Peng LN, Lin MH, Kim S, Hsiao FY, Chen LK. Enhancing intrinsic capacity and related biomarkers in community-dwelling multimorbid older adults through integrated multidomain interventions: ancillary findings from the Taiwan Integrated Geriatric (TIGER) trial. J Am Med Dir Assoc 2024 May;25(5):757-763. [doi: 10.1016/j.jamda.2023.10.006] [Medline: 37949432]

Abbreviations

AIC: Akaike information criterion
C-index: concordance index
CES-D: Center for Epidemiologic Studies Depression
DALY: disability-adjusted life year
HLI: Healthy Longevity Index
IC: intrinsic capacity
ICOPE: integrated care for older people
LTCI: long-term care insurance
MMSE: Mini-Mental State Examination
NILS-LSA: National Institute for Longevity Sciences, Longitudinal Study of Aging
SPMSQ: Short Portable Mental Status Questionnaire
TLSA: Taiwan Longitudinal Study on Aging
TRIPOD: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis
WHO: World Health Organization

Edited by H LaMonica; submitted 03.07.25; peer-reviewed by SC Chuang; revised version received 02.09.25; accepted 15.09.25; published 03.11.25.

Please cite as:
Lai HY, Zhang S, Otsuka R, Huang ST, Arai H, Hsiao FY, Chen LK
Development and Validation of the Healthy Longevity Index for Personalized Healthy Aging in Primary Care: Cross-National Retrospective Analysis
JMIR Aging 2025;8:e80034
URL: https://aging.jmir.org/2025/1/e80034
doi:10.2196/80034

© Hsi-Yu Lai, Shu Zhang, Rei Otsuka, Shih-Tsung Huang, Hidenori Arai, Fei-Yuan Hsiao, Liang-Kung Chen. Originally published in JMIR Aging (https://aging.jmir.org), 3.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Real-World Insights Into Dementia Diagnosis Trajectory and Clinical Practice Patterns Unveiled by Natural Language Processing: Development and Usability Study

Hunki Paek[1*], PhD; Richard H Fortinsky[2*], PhD; Kyeryoung Lee[1], PhD; Liang-Chin Huang[1], PhD; Yazeed S Maghaydah[2], MD; George A Kuchel[2], MD; Xiaoyan Wang[3,4], PhD

[1]IMO Health, Rosemont, IL, United States

[2]UConn Center on Aging, University of Connecticut School of Medicine, Farmington, CT, United States

[3]Center for Quantitative Medicine, University of Connecticut School of Medicine, 195 Farmington Ave, Farmington, CT, United States

[4]Department of Health Policy and Management, Tulane University, New Orleans, LA, United States

[*]these authors contributed equally

Corresponding Author:
Xiaoyan Wang, PhD
Center for Quantitative Medicine, University of Connecticut School of Medicine, 195 Farmington Ave, Farmington, CT, United States

Abstract

Background: Understanding the dementia disease trajectory and clinical practice patterns in outpatient settings is vital for effective management. Knowledge about the path from initial memory loss complaints to dementia diagnosis remains limited.

Objective: This study aims to (1) determine the time intervals between initial memory loss complaints and dementia diagnosis in outpatient care, (2) assess the proportion of patients receiving cognition-enhancing medication prior to dementia diagnosis, and (3) identify patient and provider characteristics that influence the time between memory complaints and diagnosis and the prescription of cognition-enhancing medication.

Methods: This retrospective cohort study used a large outpatient electronic health record (EHR) database from the University of Connecticut Health Center, covering 2010 - 2018, with a cohort of 581 outpatients. We used a customized deep learning–based natural language processing (NLP) pipeline to extract clinical information from EHR data, focusing on cognition-related symptoms, primary caregiver relation, and medication usage. We applied descriptive statistics, linear, and logistic regression for analysis.

Results: The NLP pipeline showed precision, recall, and F_1-scores of 0.97, 0.93, and 0.95, respectively. The median time from the first memory loss complaint to dementia diagnosis was 342 (IQR 200-675) days. Factors such as the location of initial complaints and diagnosis and primary caregiver relationships significantly affected this interval. Around 25.1% (146/581) of patients were prescribed cognition-enhancing medication before diagnosis, with the number of complaints influencing medication usage.

Conclusions: Our NLP-guided analysis provided insights into the clinical pathways from memory complaints to dementia diagnosis and medication practices, which can enhance patient care and decision-making in outpatient settings.

(JMIR Aging 2025;8:e65221) doi:10.2196/65221

KEYWORDS

dementia; memory loss; memory; cognitive; Alzheimer disease; natural language processing; NLP; deep learning; machine learning; real-world insights; electronic health records; EHR; cohort; diagnosis; diagnostic; trajectory; pattern; prognosis; geriatric; older adults; aging

Introduction

The rising prevalence of dementia, driven by an aging population, presents a profound concern for society [1-4], placing substantial burdens on individuals and imposing high financial costs [5-7]. Despite these challenges, no curative treatments are currently available [8,9], highlighting the critical need for early detection of prodromal symptoms of dementia, such as mild cognitive decline [10-13], and timely diagnosis. Early intervention can delay disease progression or alter the trajectory toward dementia [9,14,15]. However, dementia and its associated symptoms are frequently underreported and underdiagnosed in clinical practice [16,17]. As the condition progresses, patients commonly experience increased memory loss, deteriorating cognitive ability, heightened confusion, and changes in personality like agitation. The conversion from mild

XSL•FO
RenderX

cognitive impairment to Alzheimers disease has been explored using patient health data [18].

Electronic health records (EHRs) offer a valuable resource for enhancing the detection and management of disease by providing comprehensive data on patient health and history [19-23]. However, much of the nuanced patient information is embedded within unstructured clinical notes and is not accessible through structured data. Natural language processing (NLP), a subfield of artificial intelligence that enables computers to understand, interpret, and generate human language, holds promise for extracting meaningful information from vast and complex free-text EHRs [24-27]. NLP has been instrumental in automatically extracting clinical information in various medical domains, including geriatric care [28-35]. For instance, Kharrazi et al [36] showcased higher rates of geriatric syndrome extraction from unstructured EHR using NLP compared to relying solely on claim data or structured EHR data. Studies have successfully extracted cognitive status and measurement scores [37,38], as well as lifestyle exposures and discourse production for Alzheimers disease [39], from clinical documentation using NLP. Additionally, multiple studies have applied NLP methods to extract neuropsychiatric symptoms and cognitive or function impairment information [40-43]. State-of-the-art models, such as pretrained Bidirectional Encoder Representation from Transformer (BERT), have been applied in clinical settings for tasks like detecting inpatient falls [44] and classifying dementia risk [45] using clinical notes. While previous research has made significant strides in the earlier detection of cognitive decline using EHR, most studies have focused on extracting symptoms or cognitive measurement scores rather than other clinical features that affect disease progression, such as the relationship of the primary caregivers with patients.

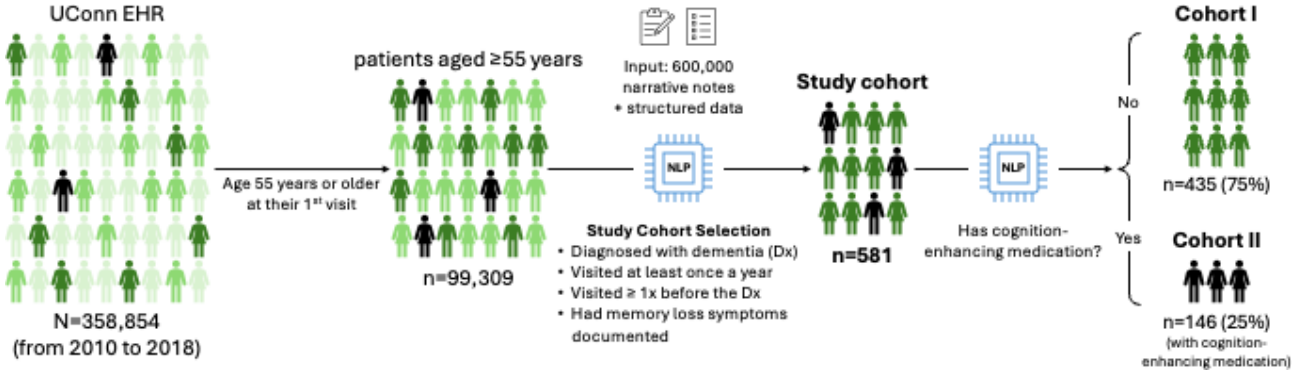
We aimed to investigate the time interval from initial memory loss complaints to dementia diagnosis and explore the association between various clinical features, including the family primary caregiver relationship, using real-world outpatient clinical notes. Additionally, we aimed to analyze the pattern of cognition-enhancing medication prescriptions before diagnosis. To achieve this, we developed a customized NLP pipeline using deep learning techniques, based on a prodromal dementia symptom ontology that we established.

Methods

Study Cohorts

This retrospective study used data from the UConn Health Center between 2010 and 2018. The use of longitudinal EHR data allowed us to track all patients' clinical information, including demographic characteristics, diagnoses, measurements, medications, and signs. The study cohort was defined as patients who met the following criteria: (1) received a dementia diagnosis, (2) had at least one outpatient visit per year, (3) had at least one visit before the dementia diagnosis, and (4) had documented memory loss–related symptoms (eg, memory loss, confusion, cognition impairment, trouble remembering, not recalling, forgetting, and blackout) in the EHR. Dementia was defined based on the presence of 3 or more *ICD* (*International Classification of Diseases*) codes used for dementia phenotyping in our study (as detailed in Multimedia Appendix 1) and dementia documentation in their clinical notes. We analyzed demographic and clinical characteristics from structured EHR data, including insurance details, the initial location (medical unit) of memory loss complaints, and the location of the first dementia diagnosis. These locations encompass various settings within this health care system, such as geriatric medicine, internal medicine, and neurology outpatient clinics. NLP was used to extract symptoms and primary caregiver (family supporter) relationship information from clinical notes. Both diagnosis and cognition-improving medication information were extracted from both structured and unstructured data. Figure 1 provides an overview of the selection process of the study cohort and the information extraction process.

Figure 1. The workflow for the study cohort selection and information extraction. EHR: electronic health record; NLP: natural language processing.



Building a Deep Learning–Based NLP Pipeline

A framework was developed to curate the prodromal dementia symptoms comprising four stages: (1) preprocessing and query expansion, (2) ontology construction and annotation, (3) NLP model development, and (4) system evaluation.

Preprocessing and Query Expansion

In this stage, a query was expanded to extract and identify a broad set of patient notes with documented memory loss symptoms. A list of seed terms was obtained through a manual survey of the literature and a clinical note review by a clinical researcher and domain expert. A bigram word2vec algorithm [46] was used to identify additional significant terms potentially related to memory loss symptoms to ensure the encapsulation

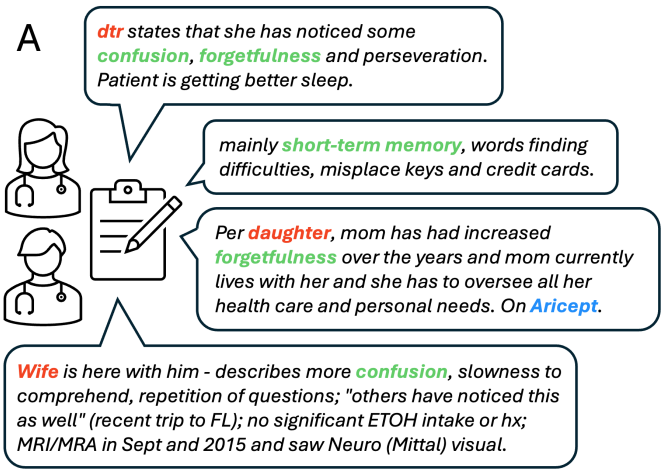
XSL•FO
RenderX

of an expansive cohort. The expanded, rule-based query terms provided in Multimedia Appendix 2 were subsequently applied to extract the relevant patient notes for NLP modeling.

Ontology Construction and Annotation

This stage involves the simulation of an expert's knowledge and understanding of the free text. A prodromal dementia symptom ontology was built based on the physician's opinion, a comprehensive literature review, and a clinical note review. The ontology includes 9 entities and 9 relations. The 9 entities are "memory loss symptom (Sx)," "dementia diagnosis (Dx)," "temporal," "duration," "status change: worse," "other symptoms," "cognitive test result," "caregiver relation," and "cognition enhancing medication (Rx)." The 9 relations are "has complaint date," "has diagnosis date," "has other symptom information," "has status change information," "has duration information," "has test information," "has caregiver information," "has treatment information," and "has effects" as depicted in Figure 2A. Two independent annotators manually annotated notes using Clinical Language Annotation, Modeling, and Processing (CLAMP) [47], an NLP toolkit, guided by the constructed ontology. Figure 2B shows an example note with entities and relations annotated.

Figure 2. The ontology of memory loss and the annotated sample note. (A) The ontology of prodromal dementia symptoms. In total, 9 entities and 9 semantic relations between entities were defined in the ontology. (B) A sample note that has been deidentified and annotated with prodromal dementia symptoms. In this note, the following entities have been annotated and related to each other: symptoms, relation, medication, duration, status change, test, and temporal. MRI/MRA: magnetic resonance imaging/magnetic resonance angiography, dtr:daughter, ETOH: ethanol, FL: florida, hx: history.



B

Entity	Standardized concept values	Original strings
Symptom	Memory loss	confusion
Symptom	Memory loss	forgetfulness
Symptom	Memory loss	short-term memory
Medication	Donepezil	Aricept
Relation	Daughter	drt
Relation	Daughter	daughter
Relation	Wife	Wife

NLP Model Development

The annotated notes (n=815) obtained in the previous stage were used for NLP model training. These notes were randomly split into a training set (80%, n=652, of annotated notes) and an independent validation set (the remaining 20%, n=163). The manual annotation and training processes were iteratively performed with additional manually annotated notes to enhance model performance until the model achieved an F_1-score of >0.8. For model training, a multilayer deep learning architecture was adopted, which involved transforming the text into sequential vectors of characterization through the embedding step. The vectors were then fed into a Bidirectional Long Short-Term Memory (BiLSTM), a text classification architecture based on artificial neural networks, for pattern recognition in both forward and backward directions. Finally, the patterns were sent to the next layer of a Conditional Random Field (CRF) model for prediction probability computation. BiLSTM-CRF architectures are widely used in clinical NLP tasks, demonstrating robustness in recognizing entities in sequential data like clinical notes [48] and effectively leveraging moderate-sized datasets with lower computational resource requirements.

NLP Pipeline Evaluation

The performance of the pipeline was evaluated in the validation set through precision (positive predictive value [PPV]), recall (sensitivity), and F_1-score (a balanced score between false positives [FPs] and false negatives [FNs]). Recall was calculated as the ratio of the number of entities that were identified by the pipeline over the total number of the corresponding entities in the manually annotated gold standard (ie, true positive [TP]/[TP+FN]). Precision was measured as the ratio of the number of distinct entities returned by the correct pipeline according to the gold standard divided by the total number of entities found by our pipeline (ie, TP/[TP+FP]). F_1-score was calculated as the harmonic mean of PPV and sensitivity (ie, 2 × PPV × sensitivity / [PPV + sensitivity]).

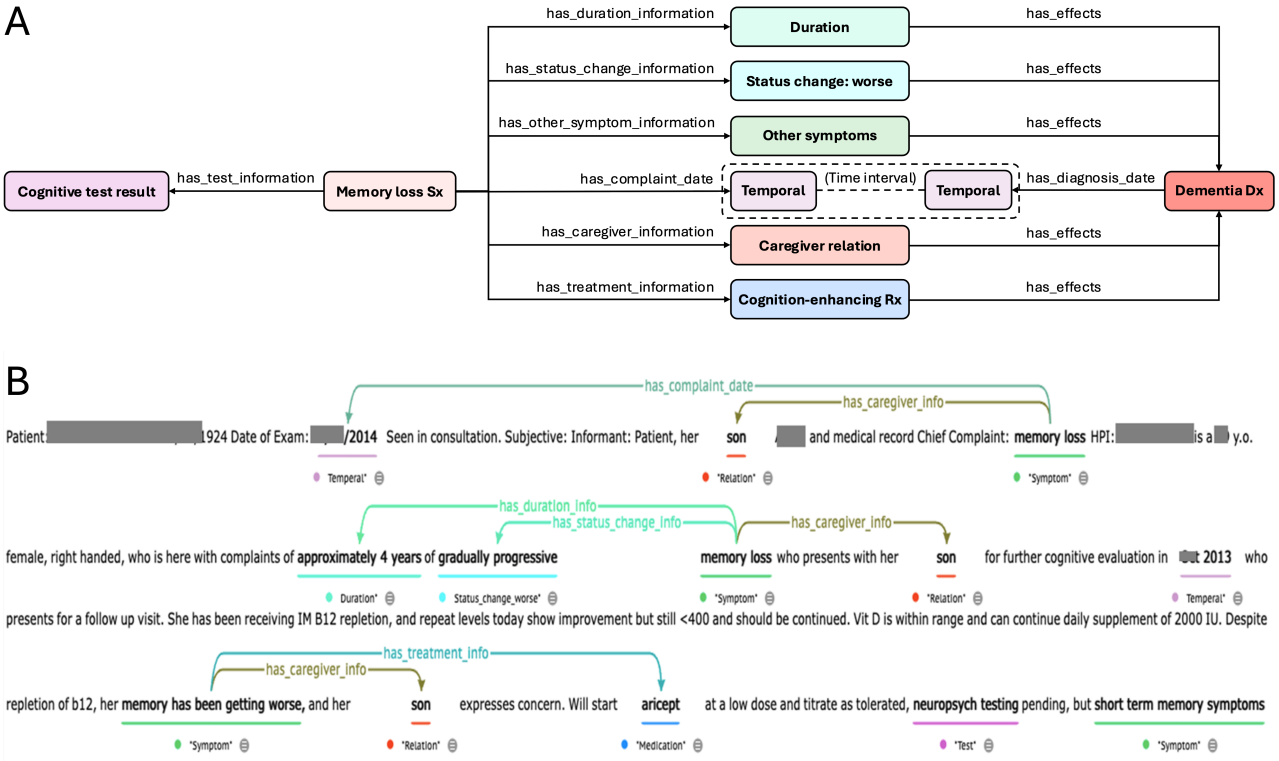
Standardization of Concept Values Leveraging NLP

Clinical notes contain abundant information but are often heterogeneous in form. To enable use case analysis, these heterogeneous entities needed to be standardized. Figure 3A illustrates the various forms in which cognition-related concepts like forgetfulness, memory loss, short-term memory, and confusion were documented in the clinical notes. Abbreviations (eg, dtr for daughter) and mixed use of brand names and generic names of the same drugs (eg, Aricept and donepezil) were commonly found. Figure 3B provides examples of standardized concept values obtained through the NLP process. The cognition-enhancing medications discussed include donepezil, memantine, rivastigmine, and galantamine. For the analysis, the son, son-in-law, daughter-in-law, grandson, and granddaughter were classified as other adult children, while the

XSL•FO
RenderX

nephew, niece, cousin, brother, and sister were classified as other family support.

Figure 3. Clinical note standardization via NLP. (A) The diverse expressions used in clinical notes depict concepts such as memory loss symptoms, primary caregiver relationships, and cognition-enhancing medications. (B) A sample of the standardized output generated by the NLP process, representing the standardized values of the various original expressions. Dx: diagnosis; HPI: history of present illness; NLP: natural language processing; Rx: medication; Sx: symptom.



Characterization of Dementia Cohorts and Longitudinal Analysis of Dementia Trajectory

The collection of clinical information of patients was accomplished through the extraction of both structured and unstructured EHR data. The statistical analysis was performed using R software (R Core Team). A logistic regression model was used to assess the relationship between medication prescription history and other factors, such as memory loss complaints and the primary caregiver relation information, by calculating the odds ratio (OR) and 95% CI. A total of 149 patients who had memory loss complaints and diagnoses recorded on the same day were excluded from this study.

Ethical Considerations

This study involved no interaction with patients as it is a retrospective cohort study that used a deidentified dataset. As such, it was deemed as exempt from institutional review board approval.

Results

Building a Memory Loss NLP Pipeline

The performance of our memory loss NLP pipeline was evaluated using precision, recall (sensitivity), and F_1-score metrics in the validation set, with detailed results presented in Multimedia Appendix 3. Our system achieved high scores across all semantic types, including “memory loss symptom,” “dementia diagnosis,” “duration,” “primary caregiver,” and “status change” (overall precision, recall [sensitivity], and F_1-scores of 0.97, 0.93, and 0.95, respectively). For example, the precision for “memory loss symptoms” was 0.97, indicating that 97% of “memory loss symptoms” were identified by our NLP system in patients’ clinical notes, as verified against a manually curated gold standard. The recall (sensitivity) of 0.93 implies that our system correctly identified 93% of actual memory loss cases, with only 7% missed. Multimedia Appendix 3 provides comprehensive details for various semantic types and relations, enabling a thorough understanding of the NLP pipeline’s performance.

Clinical Characterization of Study Cohorts

Cohort Identification

An average of 358,854 patients visited the UConn Health system for outpatient care between 2010 and 2018, with the number of visits increasing annually from 224,488 in 2010 to 1,024,349 in 2017. There were 8686 patient visits in 2018 until the time point we collected data. Out of these patients, 99,039 patients were aged 55 years or older at their first visit. From more than 600,000 narrative records and coded data of those patients, we identified 730 patients who had at least one outpatient visit per year, at least one visit before their dementia diagnosis, and documented memory loss symptoms in their clinical notes. Of these, 149 (20.4%) patients reported memory loss complaints and were diagnosed with dementia on the same day, while 581 (79.6%) patients had at least a 1-day gap between the complaint and diagnosis. For the following analysis, 581 patients with

XSL•FO
RenderX

memory loss symptoms documented at different days were included in study cohorts (Figure 1).

Cohort Demographics

The demographic characteristics, primary insurance, and medication information of study cohorts are summarized in Table 1. Most cohort members were non-Hispanic White individuals (509/581, 87.6%) and female (381/581, 65.6%), with an age distribution of over 85 years (315/581, 54.2%), 75 - 84 years (171/581, 29.4%), 65-74 years (63/581, 10.8%), and under 65 years (32/581, 5.5%).

Table . Demographic characteristics, primary insurance, and medication information of study cohorts (N=581).

Characteristics		Values, n (%)
Age (years)		
	<65	32 (5.5)
	65 - 74	63 (10.8)
	75 - 84	171 (29.4)
	85+	315 (54.2)
Race		
	White	509 (87.6)
	African American	33 (5.7)
	Others	39 (6.7)
Sex		
	Female	381 (65.6)
	Male	200 (34.4)
Primary caregiver (family supporter) relation (n=291)		
	Husband	26 (8.9)
	Wife	39 (13.4)
	Daughter	135 (46.4)
	Son	50 (17.2)
	Other family members	41 (14.1)
	Missing	290 (49.9)
Prior medication before the first diagnosis of dementia		146 (25.1)

Primary Caregiver and Medication Information

Out of 291 patients with primary caregiver relation information, adult children were the main caregivers (n=185, 63.6%, with n=135, 46.4%, being daughters and n=50, 17.2%, being sons), followed by spouses (n=65, 22.3%, with n=39, 13.4%, being wives and n=26, 8.9%, being husbands). Other family members (eg, nephew) made up 14.1% (41/291) of the cohort. Out of 581 patients, 146 (25.1%) patients had been prescribed cognition-improving medications prior to the dementia diagnosis.

Outpatient Care Locations

Next, we investigated the outpatient care locations where the first memory loss complaints were reported and where dementia was diagnosed. Geriatrics is the most frequent location for both the first memory loss complaints made (308/581, 53%) and the diagnosis of dementia (350/581, 60.2%), followed by primary care (185/581, 31.8%, and 163/581, 28.1%, respectively) and neurology (39/581, 6.7%, and 61/581, 10.5%, respectively). The majority of the cohort was covered by Medicare (354/581, 60.9%) or Medicaid (72/581, 12.4%) as primary insurance, while 26.2% (152/581) had commercial insurance. Only a small percentage of patients (2/581, 0.3%) had no insurance coverage (Multimedia Appendix 4).

Distribution of Time Intervals Between Cognitive Symptom Complaints and Dementia Diagnosis and the Number of Complaints

Time Interval

The median time interval between the first memory loss complaints and dementia diagnosis was 342 days, ranging from a minimum of 1 day to a maximum of 1458 days in our study cohort (n=581) (Multimedia Appendix 5).

Health Care Use

Additionally, the number of complaints made before being diagnosed was analyzed, with a median of 3 complaints, ranging from a minimum of 1 complaint to a maximum of 18 complaints (Multimedia Appendix 5).

Association Analysis for the Earlier Dementia Diagnosis

We aimed to identify the clinical features that are associated with earlier diagnosis of dementia from the first memory loss

XSL•FO
RenderX

complaints. Results indicated that the location of the first complaints made and the diagnosis, as well as the relation of the primary caregiver, were significantly associated with earlier diagnosis of dementia. Patients who made complaints in geriatrics (−141 days, $P<.001$, χ^2 test) or neurology (−158 days, $P=.02$) were diagnosed with dementia earlier compared to those who made complaints in primary care. Furthermore, patients diagnosed with dementia in geriatrics had a shorter interval of 152.9 days ($P<.001$) compared to those diagnosed in primary care. Additionally, having a wife or a daughter as a primary caregiver was associated with an earlier diagnosis of dementia, with a shorter interval of 249.6 days ($P=.01$) and 176.8 days ($P=.04$), respectively, compared to those who had a husband as a primary caregiver. However, factors such as age or insurance types were not found to have a significant impact on earlier diagnosis (Table 2).

Table . Statistical analysis of time intervals between the first complaints and dementia diagnosis.

Features		Estimate (95% CI)	*P* value (χ^2 test)
Age (years)			
	<65	109.8 (–26.7 to 246.3)	.12
	65 - 74	–61.6 (–163.2 to 40)	.24
	75 - 84	0.08 (–69.4 to 69.6)	≥.99
	85+	Ref[a]	—[b]
Primary insurance			
	No insurance	36.8 (–25 to 98.6)	.24
	Medicaid	5.8 (–255.5 to 267.1)	.97
	Medicare	36.8 (–11.9 to 85.5)	.14
	Commercial	Ref	—
The location of the first memory loss complaint			
	Geriatrics	–141 (–212 to –70)	<.001
	Neurology	–158 (–286.2 to –29.8)	.02
	Primary care	Ref	—
	Other	81 (–35.1 to 197.1)	.17
The location of the first diagnosis of dementia			
	Geriatrics	–152.9 (–244 to –61.8)	<.001
	Neurology	–82.2 (–191.4 to 27)	.14
	Primary care	Ref	—
	Other	42 (–225.4 to 309.4)	.76
Primary caregiver (family supporter) relation			
	Wife	–249.6 (–437.2 to –62)	.009
	Daughter	–176.8 (–346.1 to –7.5)	.04
	Other adult children	–127.4 (–256.9 to 2.1)	.05
	Other family support	–257.5 (–577.1 to 62.1)	.11
	Husband	Ref	—

[a]Reference.

[b]Not applicable.

Association Analysis for the Medication Usage

Medication was prescribed in 25.1% (146/581) of patients before dementia diagnosis. We next analyzed factors associated with the usage of cognition-enhancing medication before the diagnosis of dementia after the 1st complaints of memory loss. The only factor that was significantly associated with medication usage was the total number of memory loss complaints made; each additional memory complaint was associated with a 15% greater likelihood that cognition-enhancing medications were prescribed (OR 1.148, 95% CI 1.027 - 1.283; Table 3).

XSL•FO
RenderX

Table . An analysis of the factors associated with the usage of medication before the diagnosis of dementia.

Features		OR[a] (95% CI)
Age (years)		
	<65	3.827 (0.403 - 23.32)
	65 - 74	1.251 (0.507 - 3.727)
	75 - 84	1.127 (0.615 - 2.445)
	85+	Ref[b]
The location of the first memory loss complaint		
	Geriatrics	1.477 (0.551 - 3.959)
	Neurology	2.124 (0.449 - 10.05)
	Primary care	Ref
	Other	0.331 (0.103 - 1.058)
The location of the first diagnosis		
	Geriatrics	0.489 (0.172 - 1.39)
	Neurology	0.65 (0.182 - 2.319)
	Primary care	Ref
Family support		
	Wife	4.367 (0.85 - 22.447)
	Daughter	1.831 (0.263 - 12.74)
	Other adult children	1.609 (0.538 - 4.816)
	Other family support	1.033 (0.276 - 3.871)
	Husband	Ref
Total number of memory loss complaints before the diagnosis of dementia		1.148 (1.027 - 1.283)

[a]OR: odds ratio.

[b]Reference.

Discussion

Principal Findings

We developed a high-performance deep learning–based NLP algorithm on an EHR dataset of dementia patients to delve into the real-world trajectory of dementia, starting from initial memory loss complaints to dementia diagnosis. Our investigation focused on the time interval from the first memory loss complaints to dementia diagnosis, the proportion of prescribed cognition-enhancing medication before diagnosis during this trajectory, and the clinical characteristics associated with these features.

We found that 20.4% (149/730) of patients had same-day documentation of memory loss complaints and dementia diagnosis. Among the remaining 79.5% (580/730) of patients with at least a 1-day gap between complaints and diagnosis, over half of the patients received a dementia diagnosis within a year of their initial memory loss complaints, with a median time of 342 days. The location of the first complaint and diagnosis and the relationship with the primary caregiver emerged as influential factors in achieving an earlier diagnosis. Notably, patients who initiated complaints or were diagnosed in geriatrics or neurology received earlier diagnoses compared to those in primary care. This underscores the important role of the initial complaint's location and the dementia diagnosis's setting in the early detection and management of dementia. Our findings align with previous research indicating missed and delayed diagnoses in primary care [49]. Geriatricians and neurologists possess significantly more expertise and practical experience in diagnosing dementia and prescribing these meds than most primary care doctors. To enhance early detection in primary care, it is crucial to train primary care providers to recognize the nuances of dementia symptoms and to appreciate the importance of thorough assessments during initial consultations. This approach could significantly reduce delays in diagnosis. Furthermore, improving caregiver education regarding the signs and symptoms of dementia is essential, as it can lead to earlier recognition of concerns and encourage timely visits to health care professionals. Additionally, understanding the factors that lead patients to receive care in a geriatric or neurological department rather than primary care would be an important question for further investigation. Similar to the previous study that identified dementia severity and marital status as independent predictors of receiving a clinical cognitive evaluation [50], other factors such as more complex medical needs (eg, multiple chronic conditions and polypharmacy), severe function decline, and the primary caregiver's educational level or relationship with the patient

XSL•FO
RenderX

could be associated with visits to geriatric or neurologic departments.

Remarkably, we found that patients with a wife or daughter as their primary caregiver were diagnosed earlier and more frequently used cognitive-improving medication before the dementia diagnosis. This emphasizes the vital role of primary caregivers in the diagnosis and treatment of dementia patients. Mahmoudi et al [35] previously emphasized the importance of extracting caregiver information in dementia patient notes and developed the rule-based NLP algorithm to identify caregiver availability. In our work, we extended this by also extracting family-caregiver relationships with patients and analyzing their impact on the early diagnosis of dementia. Subsequent research should explore the underlying mechanisms and factors of caregivers in this context such as the association between the relationship of primary caregivers and visits to geriatric or neurologic departments. Contrary to expectations, our study revealed that age and insurance were not associated with earlier diagnoses.

Surprisingly, the total number of memory loss complaints emerged as the sole factor significantly linked to medication usage, with other factors showing no significant association. The correlation between increased medication prescriptions and additional memory loss complaints highlights health care providers' responsiveness to escalating symptoms and underscores the importance of proactive monitoring of cognitive symptoms. This finding aligns with the Alzheimer's Association guidelines for early identification and treatment of Alzheimers disease, particularly in its initial stages. By recognizing and documenting memory loss complaints, clinicians can better initiate appropriate therapeutic interventions, particularly those aimed at altering disease progression. Additionally, our findings emphasize the need for enhanced caregiver education on the importance of reporting memory loss and other cognitive changes.

We demonstrated that extracting cognitive symptom-related terms from longitudinally documented patient notes before dementia diagnosis could be an alternative approach to analyzing documented cognitive measurement scores during patient visits, potentially aiding in identifying dementia patients. Previous studies have highlighted a significant lack of such documentation in clinical notes [37,41,51]. For instance, Harding et al [51] found that cognitive measurement scores were rarely available in their cohorts when establishing the algorithm for identifying dementia patients in EHR. Similarly, Maserejian et al [37] demonstrated a low percentage of dementia (11%) or Alzheimers disease (24%) in patients with cognitive measurement scores such as Mini Mental State Examination (MMSE), a recall test, a clock drawing, Montreal Cognitive Assessment (MoCA), Mini-Cog, or Saint Louis University Mental Status (SLUMS) documented and suggested prompts of cognitive measurement. McCoy et al [41] attempted to extract cognitive symptom-related terms (eg, impulsive, forgetful, cognitive, and memory) and converted them into scores, given the issue of reliability and scalability of the cognitive measurement test. Consistently, the proportion of patient notes with cognitive test names, including MMSE, SLUMS, MoCA, Mini-Cog, clock drawing, trail making, Boston naming test, and Wisconsin card sorting test, was very low in our study, so these were not used in further analysis.

Our NLP approach in automatically identifying cognitive symptom-related terms and primary caregivers, and systematically analyzing these factors along with other structured data, enhanced our understanding of dementia progression and management. This approach provides a practical and scalable method for identifying cognitive impairment, especially when traditional cognitive measurement scores are lacking. Further exploration using this NLP method could significantly advance the field, providing deeper insights and more effective interventions for dementia care. Moreover, our study findings align with the Alzheimer's Association's recommendations for using simple practical tests like Mini-Cog or General Practitioner Assessment of Cognition (GPCOG), developed by a group of clinical dementia experts, during annual visits, particularly when symptoms are reported by patients or caregivers, thereby supporting current clinical practice in dementia care.

Strengths and Limitations

Our study has several strengths. First, our work presents a novel approach to understanding clinical practices in dementia by examining the time interval from the symptom complaints to the diagnosis of dementia using real-world data. Second, our study highlights how the relationship between primary caregivers and patients and the location (medical units: geriatrics, neurology, and primary care) of complaints made may influence the time to diagnosis. While existing ontologies provide comprehensive medical concepts, they often lack the specificity necessary to accurately capture the relationships between concepts. By defining the relationships between entities such as "has date" and "has caregiver information" in the ontology, our study could provide deeper insights into the clinical features that affect the time interval from the symptom presentation to dementia diagnosis in real-world contexts. Third, our study developed and validated a customized NLP model to be used to predict an outcome in a clinical setting using EHRs.

Several limitations of this study should be considered. First, the study relied on EHR data from a single health care system, which may limit the generalizability of the findings to other populations and health care settings. This system-specific reliance may introduce potential biases related to local clinical practices, documentation standards, and patient demographics, which could affect the study's findings if applied to broader or more diverse health care environments. Additionally, the patient population in our dataset was predominantly White, female, and older patients above 85 years with Medicare or Medicaid insurance, which may further limit the generalizability of the findings to other demographic groups, including younger patients, males, and individuals from diverse racial backgrounds. Future studies should aim to include a more diverse patient population in terms of age, race, and insurance type across multiple health care systems to validate and potentially broaden the applicability of the findings.

Another limitation is the potential for loss of follow-up within the EHR data, as patients with less frequent or inconsistent visits may have different clinical trajectories. We preselected patients

XSL•FO
RenderX

who had been diagnosed with dementia, had at least one documented memory loss complaint before diagnosis, and visited the health care system at least once per year to reduce the likelihood of significant follow-up loss. Nonetheless, variability in follow-up could still influence our results. Furthermore, potential biases within EHR documentation may impact the findings. For example, memory complaints may be underreported or inconsistently documented, depending on clinician practices and the completeness of note-taking. This variability could affect the accuracy of data extraction and the insights derived from the patient journey. Lastly, the study did not account for the potential impact of other medical conditions on the result. Incorporating objective measures of cognitive decline and other co-occurring neuropsychiatric symptoms could enhance the assessment of dementia.

Conclusions

Our study highlights the importance of the location of initial memory loss complaints, the location of the dementia diagnosis, and the role of the primary caregiver in the early diagnosis and treatment of dementia patients. By analyzing complex clinical dementia care practice patterns within a real-world setting on a large scale using NLP, our exploratory analysis demonstrates the potential of advanced analytical techniques in achieving earlier and more accurate diagnoses of dementia.

Acknowledgments

RHF and GAK acknowledge grant support from the National Institute on Aging (P30AG067988).

Data Availability

The data used in this study are not openly accessible due to patient privacy, security, and the Health Insurance Portability and Accountability Act of 1996 (HIPAA) requirement.

Authors' Contributions

Conceptualization: HP, GAK, KL, RHF, XW
Data curation: HP, KL, XW
Formal analysis: HP, KL, LCH, XW
Funding acquisition: GAK, RHF, XW
Methodology: HP, KL, LCH, XW
Writing – original draft: HP, GAK, KL, RHF, XW
Writing – review & editing: HP, GAK, KL, RHF, YSM, XW

Conflicts of Interest

HP, KL, LCH, and XW are current employees of IMO Health. All other authors declare no other competing financial or nonfinancial interests.

Multimedia Appendix 1
ICD (*International Classification of Diseases*) codes used for the phenotyping of dementia cohort.
[DOCX File, 15 KB - aging_v8i1e65221_app1.docx]

Multimedia Appendix 2
Query terms used for identifying memory loss–related symptoms.
[DOCX File, 15 KB - aging_v8i1e65221_app2.docx]

Multimedia Appendix 3
Evaluation of memory loss NLP pipeline. NLP: natural language processing.
[DOCX File, 15 KB - aging_v8i1e65221_app3.docx]

Multimedia Appendix 4
Descriptive statistics of providers and insurance information.
[DOCX File, 14 KB - aging_v8i1e65221_app4.docx]

Multimedia Appendix 5
The distribution time intervals and complaints. (A) Distribution of the time intervals between the first memory loss complaints and the diagnosis of dementia. (B) Distribution of the number of complaints made before the diagnosis of dementia.
[DOCX File, 82 KB - aging_v8i1e65221_app5.docx]

XSL•FO
RenderX

References

1. Hebert LE, Weuve J, Scherr PA, Evans DA. Alzheimer disease in the United States (2010-2050) estimated using the 2010 census. Neurology 2013 May 7;80(19):1778-1783. [doi: 10.1212/WNL.0b013e31828726f5] [Medline: 23390181]
2. Livingston G, Sommerlad A, Orgeta V, et al. Dementia prevention, intervention, and care. Lancet 2017 Dec;390(10113):2673-2734. [doi: 10.1016/S0140-6736(17)31363-6]
3. Matthews KA, Xu W, Gaglioti AH, et al. Racial and ethnic estimates of Alzheimer's disease and related dementias in the United States (2015-2060) in adults aged ≥65 years. Alzheimers Dement 2019 Jan;15(1):17-24. [doi: 10.1016/j.jalz.2018.06.3063] [Medline: 30243772]
4. Plassman BL, Langa KM, Fisher GG, et al. Prevalence of dementia in the United States: the aging, demographics, and memory study. Neuroepidemiology 2007;29(1-2):125-132. [doi: 10.1159/000109998] [Medline: 17975326]
5. Velandia PP, Miller-Petrie MK, Chen C, et al. Global and regional spending on dementia care from 2000-2019 and expected future health spending scenarios from 2020-2050: an economic modelling exercise. eClinicalMedicine 2022 Mar;45:101337. [doi: 10.1016/j.eclinm.2022.101337] [Medline: 35299657]
6. GBD 2016 Dementia Collaborators. Global, regional, and national burden of Alzheimer's disease and other dementias, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol 2019 Jan;18(1):88-106. [doi: 10.1016/S1474-4422(18)30403-4]
7. Frankish H, Horton R. Prevention and management of dementia: a priority for public health. Lancet 2017 Dec;390(10113):2614-2615. [doi: 10.1016/S0140-6736(17)31756-7]
8. Tisher A, Salardini A. A comprehensive update on treatment of dementia. Semin Neurol 2019 Apr;39(2):167-178. [doi: 10.1055/s-0039-1683408] [Medline: 30925610]
9. Long JM, Holtzman DM. Alzheimer disease: an update on pathobiology and treatment strategies. Cell 2019 Oct 3;179(2):312-339. [doi: 10.1016/j.cell.2019.09.001] [Medline: 31564456]
10. Gale SA, Acar D, Daffner KR. Dementia. Am J Med 2018 Oct;131(10):1161-1169. [doi: 10.1016/j.amjmed.2018.01.022] [Medline: 29425707]
11. 2021 Alzheimer's disease facts and figures. Alzheimers Dement 2021 Mar;17(3):327-406. [doi: 10.1002/alz.12328] [Medline: 33756057]
12. Ferencz B, Gerritsen L. Genetics and underlying pathology of dementia. Neuropsychol Rev 2015 Mar;25(1):113-124. [doi: 10.1007/s11065-014-9276-3] [Medline: 25567624]
13. Flicker C, Ferris SH, Reisberg B. Mild cognitive impairment in the elderly: predictors of dementia. Neurology 1991 Jul;41(7):1006-1009. [doi: 10.1212/wnl.41.7.1006] [Medline: 2067629]
14. Ossenkoppele R, Smith R, Mattsson-Carlgren N, et al. Accuracy of tau positron emission tomography as a prognostic marker in preclinical and prodromal Alzheimer disease: a head-to-head comparison against amyloid positron emission tomography and magnetic resonance imaging. JAMA Neurol 2021 Aug 1;78(8):961-971. [doi: 10.1001/jamaneurol.2021.1858] [Medline: 34180956]
15. Vermunt L, Sikkes SAM, van den Hout A, et al. Duration of preclinical, prodromal, and dementia stages of Alzheimer's disease in relation to age, sex, and APOE genotype. Alzheimers Dement 2019 Jul;15(7):888-898. [doi: 10.1016/j.jalz.2019.04.001] [Medline: 31164314]
16. Amjad H, Roth DL, Sheehan OC, Lyketsos CG, Wolff JL, Samus QM. Underdiagnosis of dementia: an observational study of patterns in diagnosis and awareness in US older adults. J Gen Intern Med 2018 Jul;33(7):1131-1138. [doi: 10.1007/s11606-018-4377-y] [Medline: 29508259]
17. Maclagan LC, Abdalla M, Harris DA, et al. Can patients with dementia be identified in primary care electronic medical records using natural language processing? J Healthc Inform Res 2023 Mar;7(1):42-58. [doi: 10.1007/s41666-023-00125-6] [Medline: 36910911]
18. Cui Y, Liu B, Luo S, et al. Identification of conversion from mild cognitive impairment to Alzheimer's disease using multivariate predictors. PLOS ONE 2011 Jul;6(7):e21896. [doi: 10.1371/journal.pone.0021896]
19. Yao L, Zhang Y, Li Y, Sanseau P, Agarwal P. Electronic health records: implications for drug discovery. Drug Discov Today 2011 Jul;16(13-14):594-599. [doi: 10.1016/j.drudis.2011.05.009] [Medline: 21624499]
20. Ellsworth MA, Dziadzko M, O'Horo JC, Farrell AM, Zhang J, Herasevich V. An appraisal of published usability evaluations of electronic health records via systematic review. J Am Med Inform Assoc 2017 Jan;24(1):218-226. [doi: 10.1093/jamia/ocw046] [Medline: 27107451]
21. Wang Y, Wang L, Rastegar-Mojarad M, et al. Clinical information extraction applications: a literature review. J Biomed Inform 2018 Jan;77:34-49. [doi: 10.1016/j.jbi.2017.11.011] [Medline: 29162496]
22. Bacigalupo I, Lombardo FL, Bargagli AM, et al. Identification of dementia and MCI cases in health information systems: an Italian validation study. Alzheimers Dement 2022;8(1):e12327. [doi: 10.1002/trc2.12327] [Medline: 36320346]
23. Tjandra D, Migrino RQ, Giordani B, Wiens J. Cohort discovery and risk stratification for Alzheimer's disease: an electronic health record-based approach. Alzheimers Dement 2020;6(1):e12035. [doi: 10.1002/trc2.12035] [Medline: 32548236]
24. Fu S, Chen D, He H, et al. Clinical concept extraction: a methodology review. J Biomed Inform 2020 Sep;109:103526. [doi: 10.1016/j.jbi.2020.103526] [Medline: 32768446]

XSL•FO
RenderX

25. Wang X, Chused A, Elhadad N, Friedman C, Markatou M. Automated knowledge acquisition from clinical narrative reports. AMIA Annu Symp Proc 2008 Nov 6;2008:783-787. [Medline: 18999156]
26. Wang X, Hripcsak G, Markatou M, Friedman C. Active computerized pharmacovigilance using natural language processing, statistics, and electronic health records: a feasibility study. J Am Med Inform Assoc 2009;16(3):328-337. [doi: 10.1197/jamia.M3028] [Medline: 19261932]
27. Koleck TA, Dreisbach C, Bourne PE, Bakken S. Natural language processing of symptoms documented in free-text narratives of electronic health records: a systematic review. J Am Med Inform Assoc 2019 Apr 1;26(4):364-379. [doi: 10.1093/jamia/ocy173] [Medline: 30726935]
28. Lee K, Liu Z, Chandran U, et al. Detecting ground glass opacity features in patients with lung cancer: automated extraction and longitudinal analysis via deep learning–based natural language processing. JMIR AI 2023;2:e44537. [doi: 10.2196/44537] [Medline: 38875565]
29. Weegar R. Applying natural language processing to electronic medical records for estimating healthy life expectancy. Lancet Reg Health West Pac 2021 Apr;9:100132. [doi: 10.1016/j.lanwpc.2021.100132] [Medline: 34327441]
30. Maclagan LC, Abdalla M, Harris DA, et al. Using natural language processing to identify signs and symptoms of dementia and cognitive impairment in primary care electronic medical records (EMR). Alzheimers Dement 2021 Dec;17(S7):e054091. [doi: 10.1002/alz.054091]
31. Al-Harrasi AM, Iqbal E, Tsamakis K, et al. Motor signs in Alzheimer's disease and vascular dementia: detection through natural language processing, co-morbid features and relationship to adverse outcomes. Exp Gerontol 2021 Apr;146:111223. [doi: 10.1016/j.exger.2020.111223] [Medline: 33450346]
32. Byrd RJ, Steinhubl SR, Sun J, Ebadollahi S, Stewart WF. Automatic identification of heart failure diagnostic criteria, using text analysis of clinical notes from electronic health records. Int J Med Inform 2014 Dec;83(12):983-992. [doi: 10.1016/j.ijmedinf.2012.12.005] [Medline: 23317809]
33. Divita G, Luo G, Tran LTT, Workman TE, Gundlapalli AV, Samore MH. General symptom extraction from VA electronic medical notes. Stud Health Technol Inform 2017;245:356-360. [Medline: 29295115]
34. Fu S, Lopes GS, Pagali SR, et al. Ascertainment of delirium status using natural language processing from electronic health records. J Gerontol A Biol Sci Med Sci 2022 Mar 3;77(3):524-530. [doi: 10.1093/gerona/glaa275] [Medline: 35239951]
35. Mahmoudi E, Wu W, Najarian C, Aikens J, Bynum J, Vydiswaran VGV. Identifying caregiver availability using medical notes with rule-based natural language processing: retrospective cohort study. JMIR Aging 2022 Sep 22;5(3):e40241. [doi: 10.2196/40241] [Medline: 35998328]
36. Kharrazi H, Anzaldi LJ, Hernandez L, et al. The value of unstructured electronic health record data in geriatric syndrome case identification. J Am Geriatr Soc 2018 Aug;66(8):1499-1507. [doi: 10.1111/jgs.15411] [Medline: 29972595]
37. Maserejian N, Krzywy H, Eaton S, Galvin JE. Cognitive measures lacking in EHR prior to dementia or Alzheimer's disease diagnosis. Alzheimers Dement 2021 Jul;17(7):1231-1243. [doi: 10.1002/alz.12280] [Medline: 33656251]
38. Noori A, Magdamo C, Liu X, et al. Development and evaluation of a natural language processing annotation tool to facilitate phenotyping of cognitive status in electronic health records: diagnostic study. J Med Internet Res 2022 Aug 30;24(8):e40384. [doi: 10.2196/40384] [Medline: 36040790]
39. Zhou X, Wang Y, Sohn S, Therneau TM, Liu H, Knopman DS. Automatic extraction and assessment of lifestyle exposures for Alzheimer's disease using natural language processing. Int J Med Inform 2019 Oct;130:103943. [doi: 10.1016/j.ijmedinf.2019.08.003] [Medline: 31476655]
40. Topaz M, Adams V, Wilson P, Woo K, Ryvicker M. Free-text documentation of dementia symptoms in home healthcare: a natural language processing study. Gerontol Geriatr Med 2020;6. [doi: 10.1177/2333721420959861] [Medline: 33029550]
41. McCoy TH, Han L, Pellegrini AM, Tanzi RE, Berretta S, Perlis RH. Stratifying risk for dementia onset using large-scale electronic health record data: a retrospective cohort study. Alzheimers Dement 2020 Mar;16(3):531-540. [doi: 10.1016/j.jalz.2019.09.084] [Medline: 31859230]
42. McCoy TH Jr, Yu S, Hart KL, et al. High throughput phenotyping for dimensional psychopathology in electronic health records. Biol Psychiatry 2018 Jun 15;83(12):997-1004. [doi: 10.1016/j.biopsych.2018.01.011] [Medline: 29496195]
43. Oh IY, Schindler SE, Ghoshal N, Lai AM, Payne PRO, Gupta A. Extraction of clinical phenotypes for Alzheimer's disease dementia from clinical notes using natural language processing. JAMIA Open 2023 Apr;6(1):ooad014. [doi: 10.1093/jamiaopen/ooad014] [Medline: 36844369]
44. Cheligeer C, Wu G, Lee S, et al. BERT-based neural network for inpatient fall detection from electronic medical records: retrospective cohort study. JMIR Med Inform 2024 Jan 30;12:e48995. [doi: 10.2196/48995] [Medline: 38289643]
45. Ben Miled Z, Dexter PR, Grout RW, Boustani M. Feature engineering from medical notes: a case study of dementia detection. Heliyon 2023 Mar;9(3):e14636. [doi: 10.1016/j.heliyon.2023.e14636] [Medline: 37020943]
46. Mikolov T, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space. arXiv. Preprint posted online on Jan 16, 2013. [doi: 10.48550/ARXIV.1301.3781]
47. Soysal E, Wang J, Jiang M, et al. CLAMP – a toolkit for efficiently building customized clinical natural language processing pipelines. J Am Med Inform Assoc 2018 Mar 1;25(3):331-336. [doi: 10.1093/jamia/ocx132] [Medline: 29186491]

XSL•FO
RenderX

48. Lee K, Liu Z, Chandran U, et al. Detecting ground glass opacity features in patients with lung cancer: automated extraction and longitudinal analysis via deep learning-based natural language processing. JMIR AI 2023 Jun 1;2:e44537. [doi: 10.2196/44537] [Medline: 38875565]
49. Bradford A, Kunik ME, Schulz P, Williams SP, Singh H. Missed and delayed diagnosis of dementia in primary care: prevalence and contributing factors. Alzheimer Dis Assoc Disord 2009;23(4):306-314. [doi: 10.1097/WAD.0b013e3181a6bebc] [Medline: 19568149]
50. Kotagal V, Langa KM, Plassman BL, et al. Factors associated with cognitive evaluations in the United States. Neurology 2015 Jan 6;84(1):64-71. [doi: 10.1212/WNL.0000000000001096] [Medline: 25428689]
51. Harding BN, Floyd JS, Scherrer JF, et al. Methods to identify dementia in the electronic health record: comparing cognitive test scores with dementia algorithms. Healthcare 2020 Jun;8(2):100430. [doi: 10.1016/j.hjdsi.2020.100430] [Medline: 32553526]

Abbreviations

BERT: Bidirectional Encoder Representation from Transformer
BiLSTM: Bidirectional Long Short-Term Memory
CLAMP: Clinical Language Annotation, Modeling, and Processing
CRF: Conditional Random Field
EHR: electronic health record
FN: false negative
FP: false positive
GPCOG: General Practitioner Assessment of Cognition
***ICD*:** *International Classification of Diseases*
MMSE: Mini Mental State Examination
MoCA: Montreal Cognitive Assessment
NLP: natural language processing
OR: odds ratio
PPV: positive predictive value
SLUMS: Saint Louis University Mental Status
TP: true positive

Edited by J Job; submitted 08.08.24; peer-reviewed by K Elechi, M Elbattah, P Han; revised version received 13.01.25; accepted 14.01.25; published 25.02.25.

Please cite as:
Paek H, Fortinsky RH, Lee K, Huang LC, Maghaydah YS, Kuchel GA, Wang X
Real-World Insights Into Dementia Diagnosis Trajectory and Clinical Practice Patterns Unveiled by Natural Language Processing: Development and Usability Study
JMIR Aging 2025;8:e65221
URL: https://aging.jmir.org/2025/1/e65221
doi:10.2196/65221

© Hunki Paek, Richard H Fortinsky, Kyeryoung Lee, Liang-Chin Huang, Yazeed S Maghaydah, George A Kuchel, Xiaoyan Wang. Originally published in JMIR Aging (https://aging.jmir.org), 25.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Development of a Longitudinal Model for Disability Prediction in Older Adults in China: Analysis of CHARLS Data (2015-2020)

Jingjing Chu[1], MSc; Ying Li[1], MSc; Xinyi Wang[1], BMed; Qun Xu[2], MSc; Zherong Xu[1], MD

[1]The First Affiliated Hospital, Zhejiang University School of Medicine, 79 Qingchun Road, Hangzhou, China

[2]Zhejiang University School of Medicine, Hangzhou, China

Corresponding Author:
Zherong Xu, MD
The First Affiliated Hospital, Zhejiang University School of Medicine, 79 Qingchun Road, Hangzhou, China

Abstract

Background: Disability profoundly affects older adults' quality of life and imposes considerable burdens on health care systems in China's aging society. Timely predictive models are essential for early intervention.

Objective: We aimed to build effective predictive models of disability for early intervention and management in older adults in China, integrating physical, cognitive, physiological, and psychological factors.

Methods: Data from the China Health and Retirement Longitudinal Study (CHARLS), spanning from 2015 to 2020 and involving 2450 older individuals initially in good health, were analyzed. The dataset was randomly divided into a training set with 70% data and a testing set with 30% data. LASSO regression with 10-fold cross-validation identified key predictors, which were then used to develop an Extreme Gradient Boosting (XGBoost) model. Model performance was evaluated using receiever operating characteristic curves, calibration curves, and clinical decision and impact curves. Variable contributions were interpreted using SHapley Additive exPlanations (SHAP) values.

Results: LASSO regression was used to evaluate 36 potential predictors, resulting in a model incorporating 9 key variables: age, hand grip strength, standing balance, the 5-repetition chair stand test (CS-5), pain, depression, cognition, respiratory function, and comorbidities. The XGBoost model demonstrated an area under the curve of 0.846 (95% CI 0.825 - 0.866) for the training set and 0.698 (95% CI 0.654 - 0.743) for the testing set. Calibration curves demonstrated reliable predictive accuracy, with mean absolute errors of 0.001 and 0.011 for the training and testing sets, respectively. Clinical decision and impact curves demonstrated significant utility across risk thresholds. SHAP analysis identified pain, respiratory function, and age as top predictors, highlighting their substantial roles in disability risk. Hand grip and the CS-5 also significantly influenced the model. A web-based application was developed for personalized risk assessment and decision-making.

Conclusion: A reliable predictive model for 5-year disability risk in Chinese older adults was developed and validated. This model enables the identification of high-risk individuals, supports early interventions, and optimizes resource allocation. Future efforts will focus on updating the model with new CHARLS data and validating it with external datasets.

(JMIR Aging 2025;8:e66723) doi:10.2196/66723

KEYWORDS

disability; prediction model; older adults; China Health and Retirement Longitudinal Study; CHARLS; medical resources allocation

Introduction

The aging of the population presents a significant global challenge, with profound implications for health care systems, economic stability, and social services [1,2]. In China, where the older population is rapidly increasing, the prevalence of disability among older adults has become a pressing concern. According to the Chinese Centers for Disease Control and Prevention, the number of older individuals with disabilities reached 52.71 million in 2020 and is projected to exceed 77.65 million by 2030. By 2030, disabled older adults are expected to account for over 57% of the total disabled population, potentially rising to more than 70% by 2050 if no preventive measures are implemented [3]. Disability in older adults encompasses various limitations or difficulties in performing daily activities independently, typically due to chronic degenerative changes in function. It is commonly assessed through activities of daily living (ADL) and instrumental activities of daily living (IADL). This growing burden of disability impacts quality of life and places a significant strain on families and public resources. Therefore, accurate prediction of disability is crucial for early intervention and effective management.

Despite significant research efforts to forecast disability in older adults, existing models often lack sufficient precision and fail to account for the complex, multifactorial nature of disability.

XSL•FO
RenderX

These models typically overlook the broader context of risk factors and offer limited utility for public health decision-making [4-6]. For instance, a study by Sun et al [5] identified depressive symptoms as a significant predictor across different types of disability. However, many existing models still fail to incorporate mental health factors alongside physical health indicators, limiting their real-world applicability. Furthermore, cognitive impairment has been found to be a strong predictor of disability in specific ADL and IADL tasks [6], highlighting the importance of integrating mental, cognitive, and physical health factors in predictive models. Disability, as a complex health issue, is influenced by multiple risk domains, including chronic diseases, polypharmacy, aging, mental health problems, unhealthy lifestyles, and the family social environment. A comprehensive predictive model should serve as a vital tool to improve early identification of at-risk individuals, inform public health strategies, and optimize resource allocation.

This study seeks to address the limitations in existing research by developing a disability prediction model specifically designed for the older Chinese population, using longitudinal data from the China Health and Retirement Longitudinal Study (CHARLS) collected between 2015 and 2020 [7]. While previous research has developed disability prediction models based on CHARLS data from 2015 to 2018 [8], the release of the 2020 survey data enables the extension of the analysis over a longer time frame. This study will leverage the validated Extreme Gradient Boosting (XGBoost) algorithm [8] to explore disability predictors in greater depth. In addition, by integrating variables such as sarcopenia and frailty-related indicators, which have previously been underexplored in predictive models for disability, we offer a more nuanced understanding of the physical, cognitive, physiological, and psychological factors contributing to disability risk. We aim to create a predictive model that not only offers high precision but also provides practical insights for health care professionals and policymakers.

Methods

Study Population

The data for this study were sourced from the CHARLS, initiated in 2011 by the National School of Development at Peking University. The CHARLS used a stratified, multistage Probability Proportional to Size random sampling method, covering 150 counties and 450 villages and urban communities across 28 provinces, involving 17,708 individuals from 10,257 households. Follow-up surveys were conducted in 2013, 2015, 2018, and 2020, with detailed methodology available in other publications [7].

Initially, 21,095 participants from the 2015 baseline survey were included. The final cohort consisted of 2450 individuals after applying the following exclusion criteria: (1) no information on biomarker or blood data; (2) younger than 60 years (for this study, older adults were defined as individuals aged 60 years or older, in accordance with the World Health Organization and the Chinese government’s standard for aging population classification); (3) missing ADL or IADL data in 2015; (4) missing follow-up ADL or IADL data in 2020; (5) having ADL and IADL limitations or any form of disability in 2015, including physical, intellectual, visual, auditory, or significant speech impairments; and (6) other relevant data missing.

Ethical Considerations

The Institutional Review Board of Peking University (IRB No. IRB00001052-11014) approved the research, and all respondents provided informed consent. CHARLS adheres to the Declaration of Helsinki and China's Personal Information Protection Law. The CHARLS database adheres to strict privacy protection and anonymization principles during data collection and processing to ensure the security of participants' personal information.

Assessment of Disability

Disability in this study was defined as impairment in performing ADL and IADL, which is commonly used in geriatric research to evaluate functional limitations in older adults. ADL assessed the ability to perform fundamental self-care tasks such as dressing, bathing, eating, getting out of bed, toileting, and managing urination and bowel movements. IADL measured more complex daily tasks, including household chores, cooking, shopping, phone use, financial management, and medication adherence. Responses were categorized into four levels: (1) no difficulty, (2) difficulty but can still do it, (3) difficulty and need help, and (4) cannot do it. To create a binary outcome, responses of (2) “Difficulty but can still do it,” (3) “Difficulty and need help,” and (4) “Cannot do it” were coded as 1 (indicating ADL and IADL disability), while the response “No difficulty” was coded as 0 (no disability). Participants were classified as having ADL and IADL disabilities if they reported any level of difficulty (levels 2 - 4) in at least one ADL or IADL item [9].

Predictive Variables

Clinical Factors

Laboratory assessments included a range of biomarkers: white blood cell count, hemoglobin, hematocrit, triglycerides, total cholesterol, glucose, uric acid, creatinine, blood urea nitrogen (BUN), high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, cystatin C, C-reactive protein, and glycated hemoglobin.

Depressive symptoms were evaluated using the 10-item Center for Epidemiologic Studies Depression Scale (CES-D) [10]. CES-D scores ranged from 0 to 30, with higher scores indicating more severe symptoms. The CES-D has been validated for Chinese middle-aged and older populations [11]. A score of 10 or above was used to indicate depression, while scores below 10 indicated no depressive symptoms [12]. Cognitive function was assessed using a modified version of the telephone interview for cognitive status (TICS) questionnaire [13]. The overall cognitive score was calculated by summing the scores from four domains: (1) orientation (5 points), (2) computation (5 points), (3) memory (20 points), and (4) drawing (1 point), with a total possible score of 31 points [14]. Higher scores indicate better cognitive performance.

Physical Performance

The physical examination included measurements of systolic and diastolic blood pressure, pulse, and respiratory function. Respiratory function was measured using a peak flow meter.

XSL•FO
RenderX

Participants were instructed to stand, take a deep breath, and blow as hard and fast as possible into the mouthpiece. The highest value from 3 attempts was recorded for analysis. Physical performance was assessed using gait speed, the 5-repetition chair stand test (CS-5), and standing balance. Gait speed was measured to evaluate lower limb function and mobility. A walking course of 2.5 meters was set up, and participants were instructed to walk the course twice at their usual pace. The average gait speed was calculated by dividing the distance by the time taken. The 5-repetition CS-5 was conducted to assess lower limb strength and endurance. Participants were asked to sit in a chair with their arms folded across their chest and, upon the examiner's command, to stand up and sit down 5 times consecutively at their fastest pace without using their arms for support. The total time required to complete the 5 repetitions was recorded, with a longer duration indicating poorer lower limb function. The standing balance assessment involved maintaining a standing position for 10 seconds in three distinct foot placements: (1) side-by-side, (2) semitandem, and (3) full tandem. Handgrip strength, the primary indicator of muscle strength, was measured for each participant using a Yuejian WL-1000 dynamometer (Nantong Yuejian Physical Measurement Instrument Co). Handgrip strength was measured in both the dominant and nondominant hands, with 2 measurements per hand. The higher value for each hand was recorded, and the average value for the 2 hands was taken to represent the handgrip strength. Together, these assessments provided a comprehensive evaluation of physical function and performance.

Appendicular skeletal muscle mass (ASM) was estimated using a formula specifically developed for the Chinese population, which closely corresponds with dual-energy X-ray absorptiometry measurements [15,16]. The formula accounts for weight, height, sex (1 for males, 2 for females), and age as follows:

$$ASM = 0.193 \times \text{weight (kg)} + 0.107 \times \text{height (cm)} - 4.157 \times \text{sex} - 0.037 \times \text{age} - 2.631$$

Potential Covariates

Covariates for our study were identified from previous literature and grouped into 2 main categories. The first included social and lifestyle factors: age, gender, BMI, marital status, residential area, daily sleep hours, and alcohol and tobacco use. BMI categories were defined as underweight (BMI <18.5 kg/m^2), normal weight (BMI =18.5 - 24 kg/m^2), and overweight (BMI ≥24 kg/m^2). The second category addressed pain, incidents of falling, and number of comorbidities (hypertension, dyslipidemia, diabetes, cancer, stroke, heart disease, lung disease, liver disease, kidney disease, digestive disease, mental health disorders, memory disorders, asthma, and arthritis) [14,17]. The comorbidity classification was based on the CHARLS questionnaire design. Neurological disorders, including Parkinson disease and Alzheimer disease, are included under memory-related diseases.

Statistical Analysis

The preprocessed dataset was split into a training subset with 70% data and a testing subset with 30% data. Continuous variables were described using medians and IQR, with comparisons using the Mann–Whitney *U* test. Count variables were expressed as frequencies and percentages and assessed using the χ^2 test. Model development and testing were performed using the training and testing sets, respectively.

Initial correlation analysis identified potential multicollinearity. Variable selection was conducted exclusively on the training set using LASSO regression with 10-fold cross-validation to prevent information leakage and ensure an unbiased evaluation of model performance. LASSO regression was chosen over other methods due to its ability to perform simultaneous variable selection and regularization, reduce overfitting, and enhance model interpretability. In addition, ablation experiments were conducted to evaluate the effect of removing specific features related to sarcopenia and frailty on model performance. The logloss metric was used to assess the performance of the models with and without these features. The selected variables informed the development of an XGBoost model, a machine learning algorithm that uses gradient boosting through decision trees to iteratively minimize prediction errors. For the optimization of the XGBoost model, hyperparameter tuning was performed using a grid search approach. Key hyperparameters were tuned, including the number of boosting rounds (nrounds), which determines the number of iterations for boosting, and the maximum tree depth, which controls the complexity of each individual tree. The learning rate (eta) was adjusted to control the weight of each update during training. In addition, the minimum loss reduction (gamma) for tree splitting, the feature subsampling ratio (colsample_bytree), and the minimum child weight were optimized to control the model's complexity and prevent overfitting. The subsample ratio (subsample) was also tuned to control the fraction of training data used in each boosting round. The optimal parameters were selected based on the lowest logloss value obtained during cross-validation.

Model performance was assessed using receiver operating characteristic (ROC) curves and area under the curve (AUC) values, with higher AUC indicating better discrimination. Calibration curves evaluated the agreement between predicted and observed outcomes. Decision curve analysis (DCA) and clinical impact curves (CIC) aided in determining optimal application and estimating the model's impact on patient management. SHapley Additive exPlanations (SHAP) values were used to interpret variable importance and model transparency. Four key SHAP plots were generated: (1) a summary plot, (2) dependence plot, (3) interaction plot, and (4) force plot. The model has been deployed on a web-based platform.

Analyses were conducted with R software (R Foundation for Statistical Computing), version 4.3.2. A *P* value <.05 was considered statistically significant.

Results

Baseline Characteristics

This study assessed a cohort of 2450 older adults initially in good health. Over a 5-year follow-up period, 610 participants developed disabilities, resulting in a disability incidence rate of

XSL•FO
RenderX

24.90%. The dataset was split 7:3, with the training set consisting of 1715 individuals (427 with disabilities) and the testing set comprising 735 participants (183 with disabilities). Baseline characteristics of both sets were detailed in Table 1 and Figure 1. Except for differences in sleep duration and white blood cell counts, no statistically significant differences were observed between the 2 groups ($P>.05$).

XSL•FO
RenderX

Table . Baseline characteristics in the training and testing subset.

Variable	Total	Training set	Testing set	*P* value
	N=2450	N=1715	N=735	
Dependent variable, n (%)				
Disability	610 (24.90)	427 (24.90)	183 (24.90)	—
Non-disability	1840 (75.10)	1288 (75.10)	552 (75.10)	—
ASM/Ht^2	6.89 [5.97-7.59]	6.87 [5.96-7.56]	6.92 [6-7.70]	.12
Hand grip (kg)	31 [25-38]	31 [25-38]	31 [24.60-38]	.78
Gait speed (m/s)	0.83 [0.69-0.97]	0.83 [0.69-0.97]	0.83 [0.70-0.99]	.17
CS-5 (s)	8.78 [7.22-10.60]	8.81 [7.30-10.70]	8.71 [7.08-10.50]	.11
Balance	3 [3-3]	3 [3-3]	3 [3-3]	.99
Age (year)	65 [62-70]	65 [62-70]	65 [62-70]	.70
Sex, n (%)				.50
Male	1393 (56.86)	967 (56.38)	426 (57.96)	
Female	1057 (43.14)	748 (43.62)	309 (42.04)	
BMI level, n (%)				.14
Underweight	142 (5.80)	104 (6.06)	38 (5.17)	
Normal weight	1299 (53.02)	926 (54.00)	373 (50.75)	
Overweight	1009 (41.18)	685 (39.94)	324 (44.08)	
Marital status, n (%)				.21
Unmarried	417 (17.02)	303 (17.67)	114 (15.51)	
Married	2033 (82.98)	1412 (82.33)	621 (84.49)	
Living area, n (%)				.59
Rural	1958 (79.92)	1376 (80.23)	582 (79.18)	
Urban	492 (20.08)	339 (19.77)	153 (20.82)	
Education level, n (%)				.24
Illiterate	811 (33.10)	577 (33.64)	234 (31.84)	
Primary school	1062 (43.35)	753 (43.91)	309 (42.04)	
Middle school	388 (15.84)	256 (14.93)	132 (17.96)	
High school and above	189 (7.71)	129 (7.52)	60 (8.16)	
Sleeping, n (%)				.04[a]
<6 h	699 (28.53)	511 (29.80)	188 (25.58)	
≥6 h	1751 (71.47)	1204 (70.20)	547 (74.42)	
Pain, n (%):				.17
No	1981 (80.86)	1374 (80.12)	607 (82.59)	
Yes	469 (19.14)	341 (19.88)	128 (17.41)	
Falldown, n (%)				.81
No	2102 (85.80)	1469 (85.66)	633 (86.12)	
Yes	348 (14.20)	246 (14.34)	102 (13.88)	
Smoking, n (%)				.35
No	1497 (61.10)	1037 (60.47)	460 (62.59)	
Yes	953 (38.90)	678 (39.53)	275 (37.41)	
Drinking, n (%)				.45

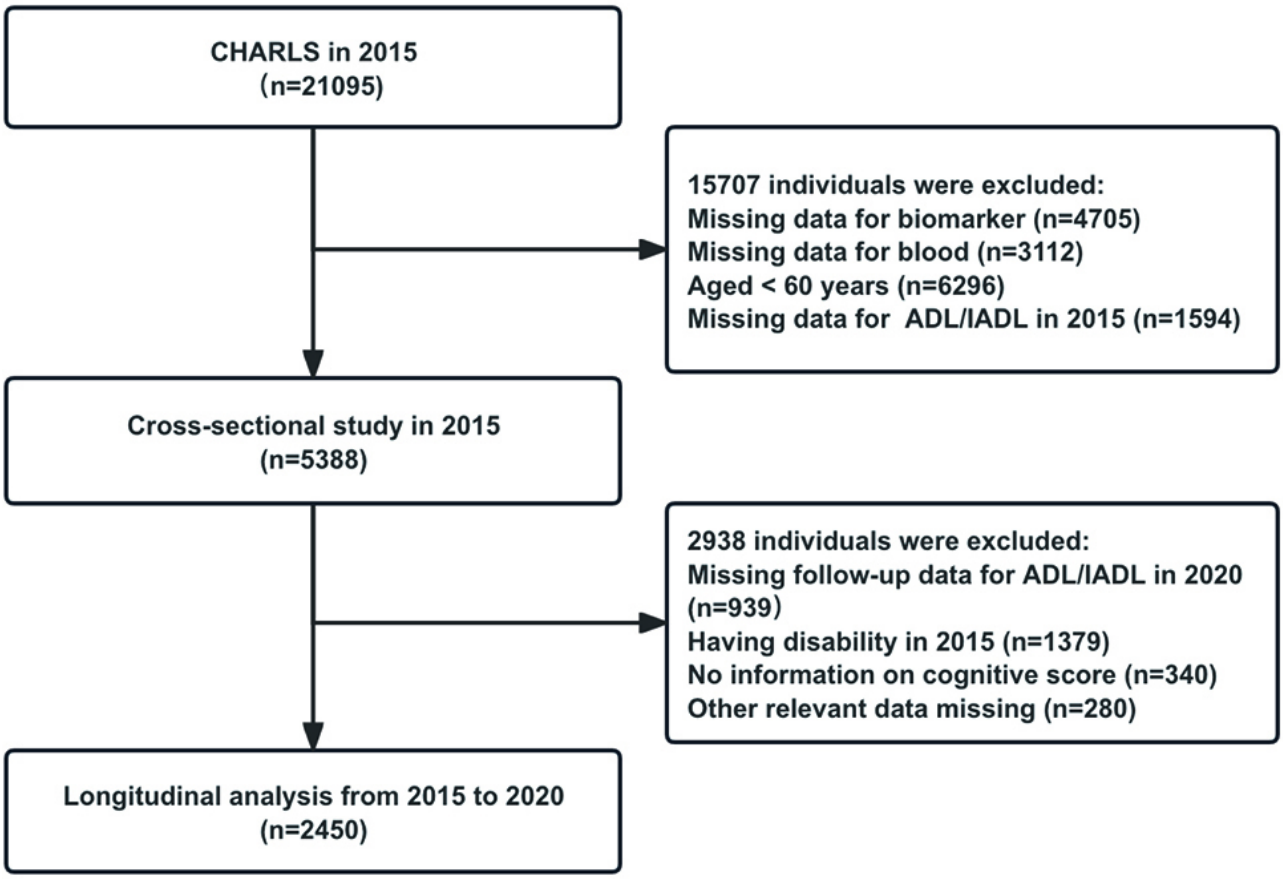
XSL•FO
RenderX

Variable	Total	Training set	Testing set	*P* value
	N=2450	N=1715	N=735	
No	1496 (61.06)	1056 (61.57)	440 (59.86)	
Yes	954 (38.94)	659 (38.43)	295 (40.14)	
Comorbidities, n (%)				.33
0	1705 (69.59)	1209 (70.50)	496 (67.48)	
1	538 (22.96)	365 (21.28)	173 (23.54)	
≥2	207 (8.45)	141 (8.22)	66 (8.98)	
Depression, n (%)				.37
No	1900 (77.55)	1339 (78.08)	561 (76.33)	
Yes	550 (22.45)	376 (21.92)	174 (23.67)	
Cognition	15.00 [12.00-18.00]	15.00 [11.50-18]	16 [12-19]	.46
Systolic BP[b]	129.67 [116.67-143.67]	129.67 [116.33-143.33]	130.33 [118-143.67]	.21
Diastolic BP	74.00 [67.33-81.33]	74 [67.33-81.33]	74.33 [67.33-81.83]	.71
Pulse	72.3 [66-79.3]	72.67 [66-79.67]	72 [66-79]	.53
Respiratory function	280 [203.33-360]	276.67[200-356.67]	280 [213-363.00]	.15
WBC[c] (1000)	5.70 [4.75-6.80]	5.67 [4.72-6.78]	5.90 [4.80-6.90]	.01[a]
HGB[d] (g/dl)	13.74 [12.66-14.80]	13.70 [12.60-14.80]	13.90 [12.70-14.90]	.13
HCT[e] (%)	41.60 [38.50-45.00]	41.40 [38.40-44.80]	41.90 [38.70-45.20]	.19
TG[f] (mg/dl)	110.62 [81.42-161.95]	109.73 [80.53-161.06]	114.16 [81.42-163.72]	.30
CHO[g] (mg/dl)	182.63 [161-205.79]	183.40 [161.39-206.56]	181.47 [159.85-203.86]	.24
GLU[h] (mg/dl)	97.30 [90.09-108.11]	95.50 [90.09-108.11]	97.30 [90.10-108.11]	.46
UA[i] (mg/dl)	5.00 [4.10-5.90]	5.00 [4.10-5.90]	4.90 [4.10-5.90]	.51
CRP[j] (mg/l)	1.40 [0.80-2.60]	1.40 [0.80-2.70]	1.40 [0.80-2.40]	.17
HbA1c[k] (%)	5.90 [5.60-6.20]	5.90 [5.60-6.20]	5.90 [5.60-6.20]	.81
CREA[l] (mg/dl)	0.81 [0.69-0.93]	0.80 [0.69-0.93]	0.81 [0.70-0.94]	.27
BUN[m] (mg/dl)	15.13 [12.89-18.49]	15.41 [12.89-18.49]	15.13 [12.61-18.49]	.37
HDL[n] (mg/dl)	50.19 [43.24-57.92]	50.19 [43.24-58.30]	50.19 [43.24-57.14]	.50
LDL[o] (mg/dl)	101.93 [84.56-120.85]	102.32 [84.56-121.24]	99.61 [84.17-120.08]	.15

[a]*P*< .05
[b]BP: blood pressure.
[c]WBC: white blood cell.
[d]HGB: hemoglobin.
[e]HCT: hematocrit.
[f]TG: triglycerides.
[g]CHO: total cholesterol.
[h]GLU: glucose.
[i]UA: uric acid.
[j]CRP: C-reactive protein.
[k]HbA1c: glycated hemoglobin.
[l]CREA: creatinine.
[m]BUN: blood urea nitrogen.
[n]HDL: high-density lipoprotein cholesterol.
[o]LDL: low-density lipoprotein cholesterol.

XSL•FO
RenderX

Figure 1. Flowchart of the data selection. CHARLS: China health and retirement longitudinal study, ADL/IADL: Activities of daily living and instrumental activities of daily living.



Predictor Selection

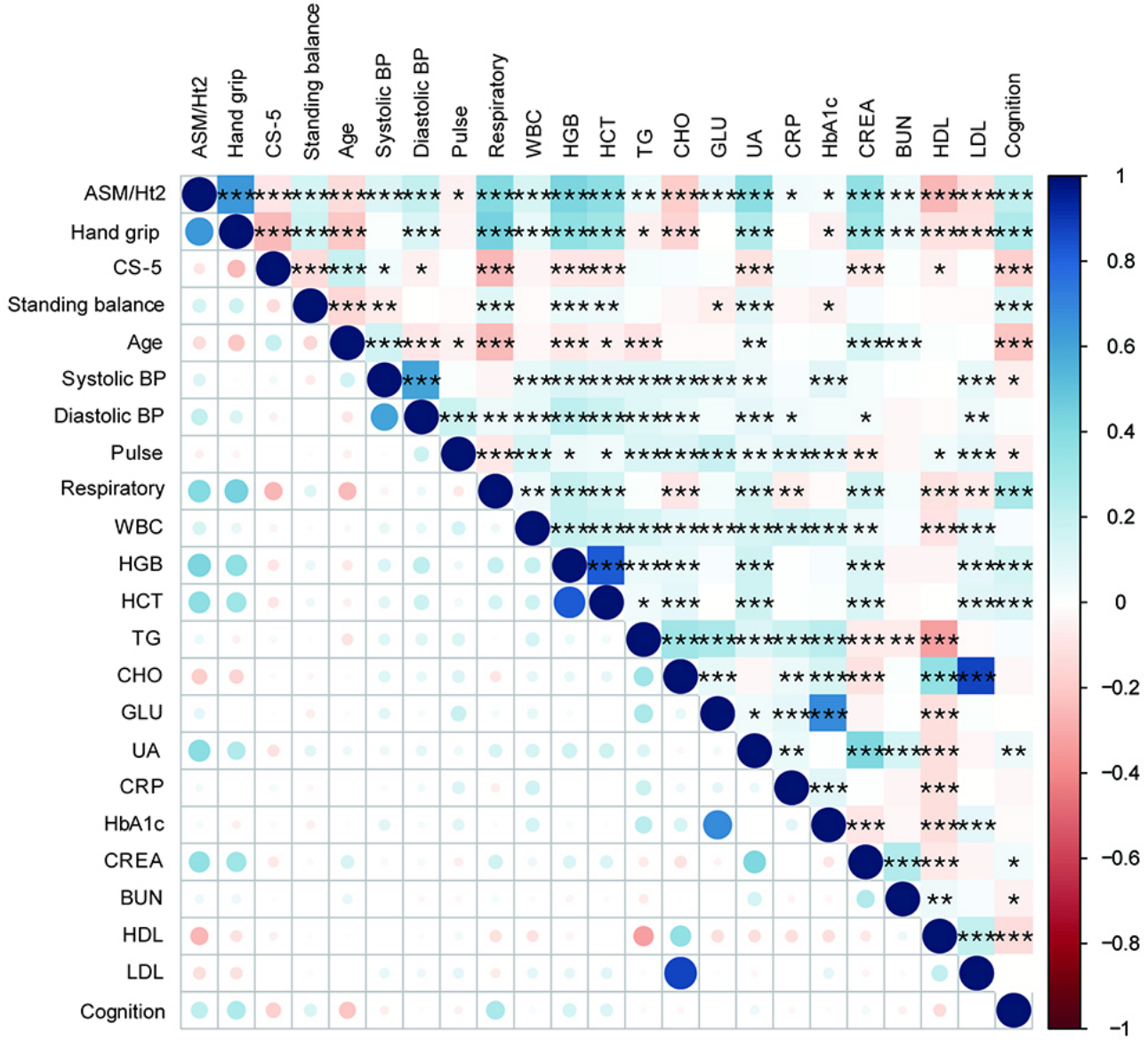
Figure 2 represented the interrelations among the continuous independent variables measured in the study. The matrix used varying shades of color and circle sizes to illustrate the magnitude and direction of correlation coefficients. The analysis revealed significant correlations, such as a negative association between hand grip and age, and a positive association between CHO and high-density lipoprotein (P<.001).

To identify the strongest predictors of disabilities, the training dataset was normalized to account for different measurement units across variables. With disability as the dependent variable, 36 potential predictors were evaluated using LASSO regression. The compressive variable coefficient was used to avoid overfitting and improve predictive accuracy. The parameter λ was selected based on the largest λ within 1 SD of the minimal binomial deviance to enforce stricter penalty constraints. The LASSO regression retained 9 predictors with non-zero coefficients (Figure 3): age, hand grip, standing balance, CS-5, pain, depression, cognition, respiratory function, and the count of comorbidities.

Following LASSO regression, Table 2 summarizes the results of the ablation experiments, which evaluates the impact of removing specific sarcopenia- and frailty-related features (hand grip, CS-5, and standing balance) on model performance. Logloss was used as the primary metric to evaluate the model's performance, with lower values indicating better predictive accuracy. The results showed that removing these features increased the logloss, with the most significant increase observed when all 3 features were removed simultaneously. These findings suggest that including sarcopenia-related parameters is crucial for maintaining the model's predictive accuracy.

XSL•FO
RenderX

Figure 2. Correlation matrix of continuous independent variables. *: *P*<.05; **: *P*<.01; ***: *P*<.001. ASM/Ht2: appendicular skeletal muscle mass and height 2, CS-5: five-repetition chair stand test, Systolic and Diastolic BP: systolic blood pressure, WBC: white blood cell, HGB: hemoglobin, HCT: hematocrit, TG: triglycerides, CHO: total cholesterol, GLU: glucose, UA: uric acid, CRP: C-reactive protein, HbA1c: glycated hemoglobin, CREA: creatinine, BUN: blood urea nitrogen, HDL: high-density lipoprotein cholesterol, LDL: low-density lipoprotein cholesterol. Positive correlations are represented by blue tones, and negative correlations by red tones, with the intensity of the color indicating the strength of the correlation. Circle size is proportional to the absolute value of the correlation coefficient.



XSL•FO
RenderX

Figure 3. Variable selection via LASSO regression model. (A)Optimal parameter selection in LASSO regression. This plot illustrates the choice of the optimal λ, displaying log(λ) on the horizontal axis and regression coefficients on the vertical axis. (B)LASSO regression parameter (λ) selection via binomial deviance plot. Each point represents the model's deviance at varying log (λ) values, with the vertical dotted line indicating the λ value that minimizes the binomial deviance.

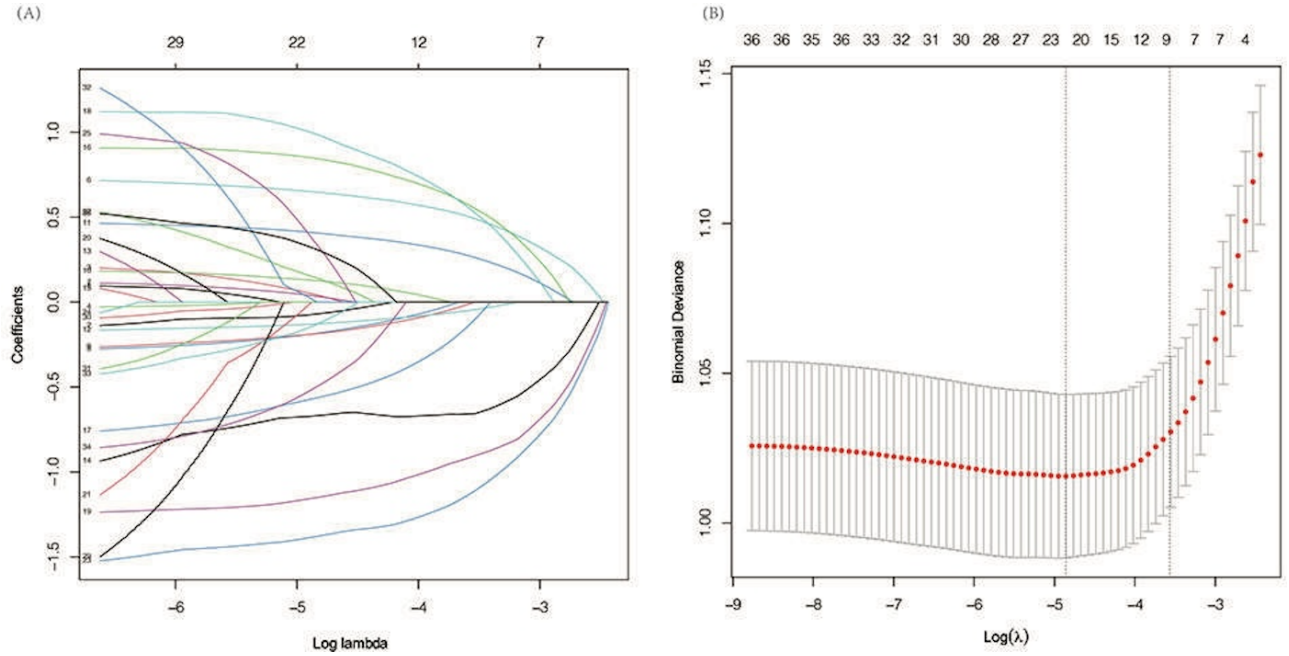


Table . Performance comparison of ablation experiments on model performance.

Model configuration	Logloss	Change	Number of features
All features (full model)	0.411	0	9
Removing hand grip	0.427	0.016	8
Removing CS-5[a]	0.421	0.010	8
Removing standing balance	0.413	0.002	8
Removing hand grip, CS-5, and standing balance	0.442	0.031	6

[a]CS-5: five-repetition chair stand test

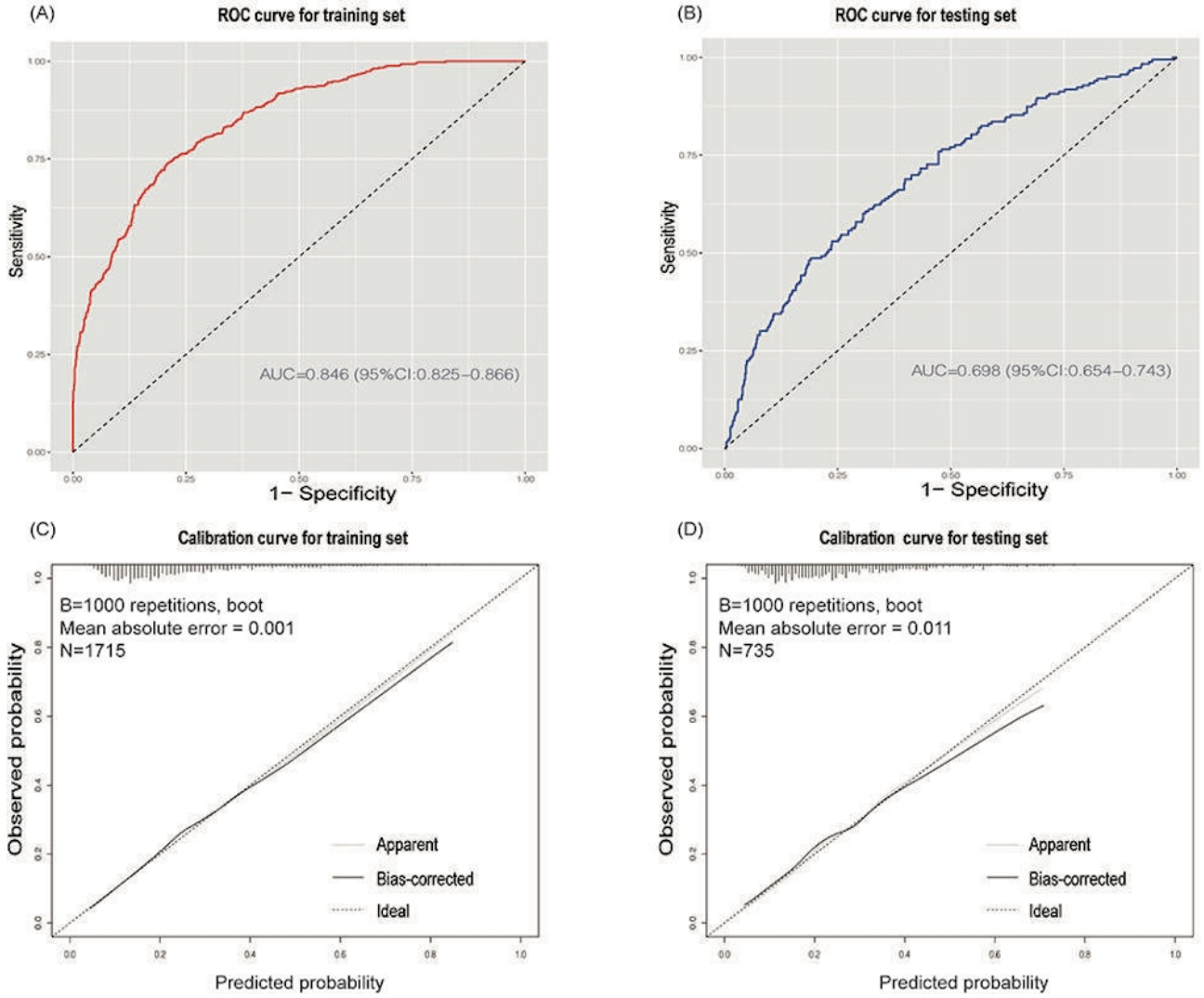
Construction and Assessment of the Predictive Model

Using disability outcomes from 2020 as the dependent variable and 9 predictors selected through LASSO regression, a predictive model was constructed using the XGBoost algorithm. The model's hyperparameters were optimized through grid search and cross-validation. The best parameters identified were: nrounds=100, max_depth=3, eta=0.1, gamma=0.1, colsample_bytree=0.8, min_child_weight=3, and subsample=0.8. These parameters were used to train the final XGBoost model, which was then evaluated on the testing dataset.

The performance of the XGBoost model was evaluated using ROC curves to assess its discrimination ability. In the training set, the model achieved an AUC of 0.846 (95% CI 0.825 - 0.866), indicating good discrimination (Figure 4A). In the testing set, the AUC was 0.698 (95% CI 0.654 - 0.743), reflecting moderate predictive accuracy (Figure 4B).

XSL•FO
RenderX

Figure 4. Receiver operating characteristic curves.



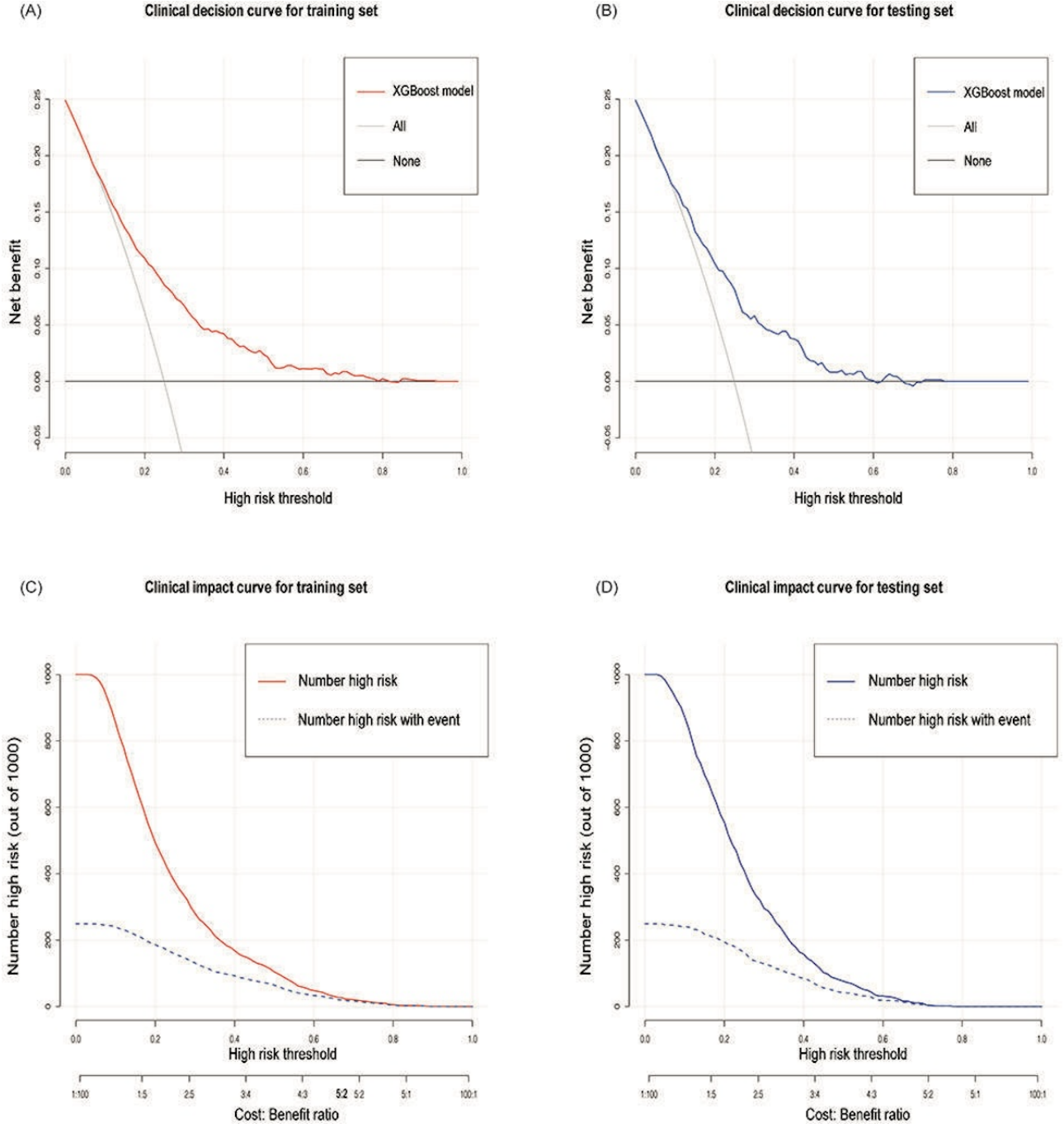
Calibration curves for the training and testing sets assessed the model's predictive accuracy. The training set showed a mean absolute error of 0.001 (Figure 4C), suggesting high precision, while the testing set had a mean absolute error of 0.011 (Figure 4D). Both curves closely approximated the ideal line, confirming the model's reliable prediction of disability risk.

Clinical decision analysis demonstrated the effectiveness of the XGBoost model in predicting disability across different risk thresholds. In the training set, the DCA showed that using the XGBoost model to identify high-risk patients provided a net benefit (Figure 5A). For example, at a chosen risk threshold of 0.30, applying the model's predictions would result in a better net benefit than treating all patients or treating none, highlighting the model's clinical utility in improving decision-making. The testing set also demonstrated a similar net benefit (Figure 5B), confirming the model's robustness and clinical applicability in an independent dataset.

In both datasets, the "Number high risk" line decreased steeply with increasing thresholds, indicating fewer individuals were classified as high-risk under stricter criteria (Figure 5C and D). In contrast, the "Number high risk with event" line, representing individuals who experienced disability, showed a more gradual decline. These trends highlight the model's ability to focus predictions on a targeted group as thresholds increase, demonstrating its utility in guiding clinical decision-making and optimizing interventions for those most likely to benefit. For a more comprehensive evaluation of the model's performance, the specificity, accuracy, positive predictive value (PPV), and negative predictive value at thresholds of 0.2 and 0.5 are provided in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 5. Clinical decision curves and impact curves for XGBoost model. (A)The red line represents the net benefit of the training set. (B)The blue line represents the net benefit of the testing set. The 'all' line indicates the benefit when all patients are treated, and the 'none' line when no patients are treated. (C-D) The solid lines depict the total number of individuals identified as high risk, and the dashed lines represent those at high risk who experienced the true event.

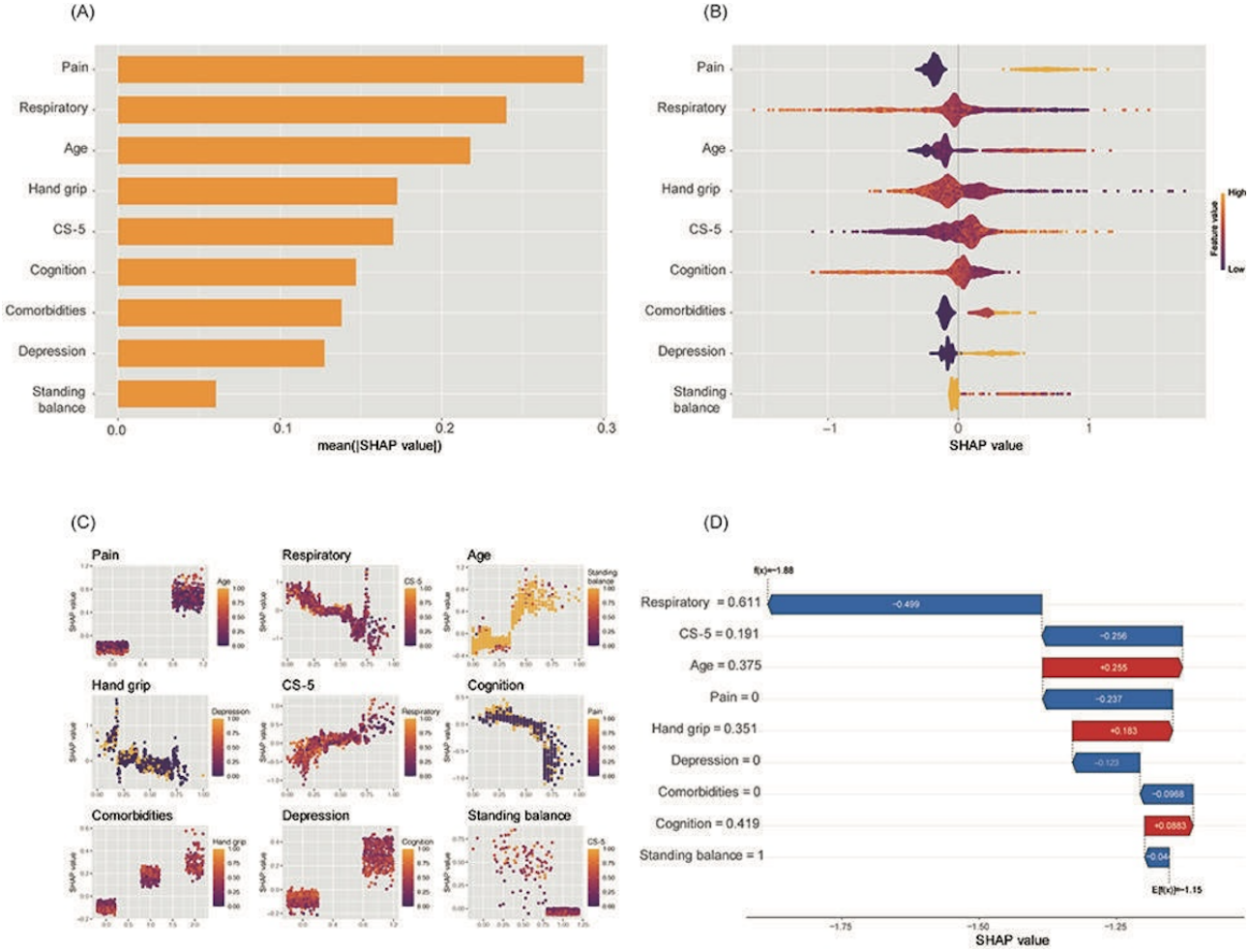


SHAP for Model Interpretation

We used SHAP values to assess the influence of each variable on the 5-year disability risk. Figure 6A ranked predictors by their mean SHAP values, reflecting their average contribution to the model's output. Pain had the highest mean SHAP value, followed by respiratory function and age, indicating their strong overall influence on disability prediction. Figure 6B showed the SHAP value distributions, where pain, respiratory function, and age exhibited the broadest ranges, suggesting their dynamic and individualized impact. Hand grip and CS-5 also significantly influenced the model, highlighting their importance in predicting physical function-related disability, while cognition, comorbidities, and depression showed more consistent contributions. Balance had the least impact.

XSL•FO
RenderX

Figure 6. SHAP: SHapley additive exPlanations value interpretation diagram for predicting disability.(A)Variable importance in the predictive model as measured by SHAP values.(B)SHAP value distribution for predictive model variables. SHAP values for model variables are shown as violins. The color represents feature value intensity, and the width indicates impact density. (C)Scatterplot matrix of SHAP values for model predictors. Each plot reveals the influence of a single variable on the model output, with color intensity indicating the magnitude of the feature value. (D)Individual sample SHAP value analysis for disability risk prediction.

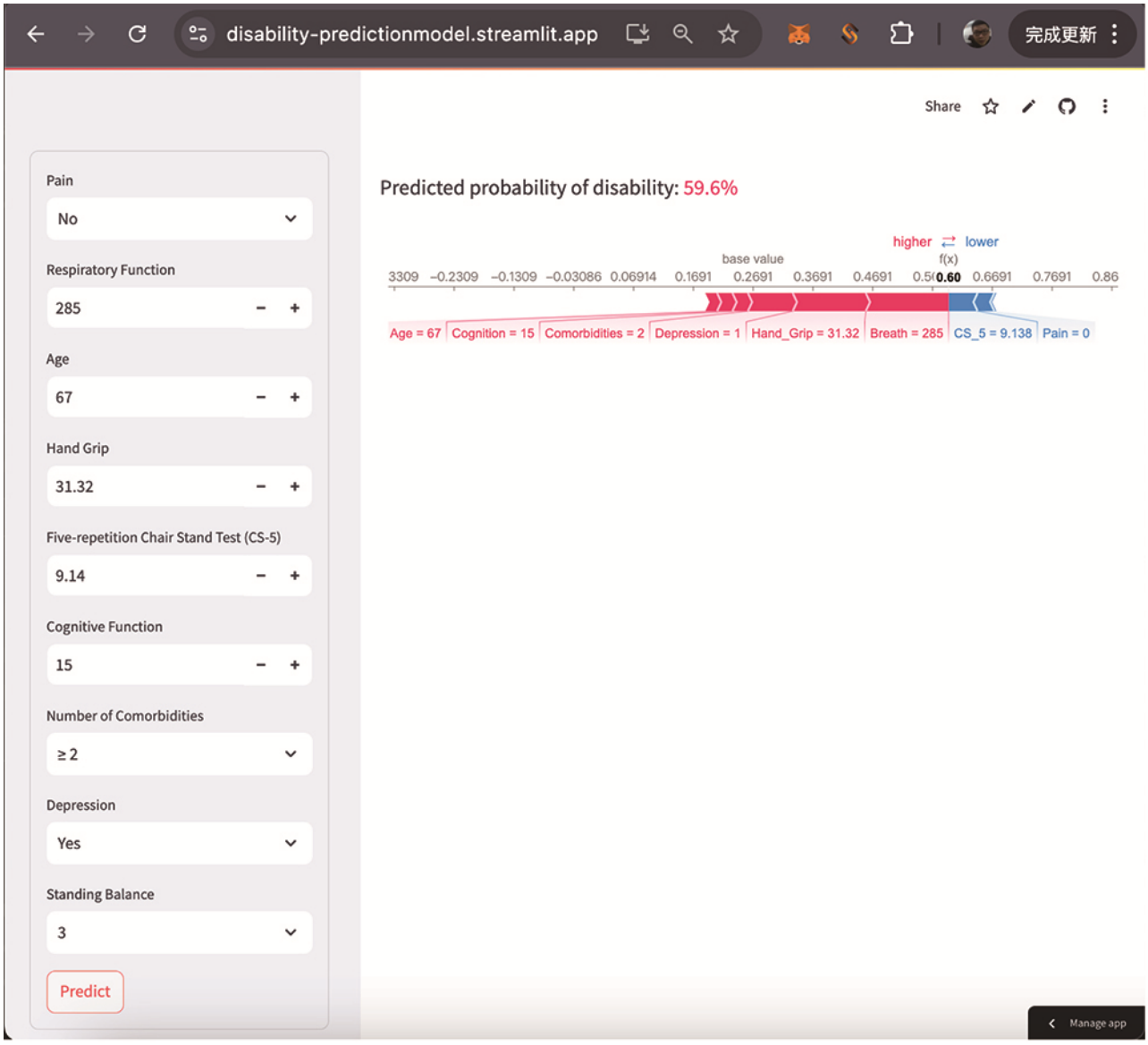


The SHAP summary plot (Figure 6C) provided an overview of the overall influence of each predictor on the model's output, revealing that pain, age, and respiratory function had the most substantial and wide-ranging influence on the predicted disability risk. The SHAP dependence plot (Figure 6D) visualized the individualized impact of these predictors on a single patient's disability risk profile, offering insights into the model's decision-making process at both macro and micro levels. Together, these plots provided a comprehensive understanding of the predictors' contributions to disability risk prediction.

A web-based calculator [18] enables clinicians to estimate the 5-year disability probability by entering patient-specific data [18], aiding in personalized clinical decisions (Figure 7).

XSL•FO
RenderX

Figure 7. Online calculator for individual disability prediction.



Discussion

Principal Findings and Comparison With Previous Works

This study constructed an effective 5-year disability prediction model using baseline data from CHARLS 2015 to forecast disability occurrence in 2020. The model identified 9 key predictive variables (pain, respiratory function, age, handgrip strength, CS-5, cognitive function, depression, comorbidities, and standing balance) that are closely associated with disability incidence in Chinese older adults. Calibration curves demonstrated the model's strong discrimination and consistency in both training and test sets, while DCA and CIC highlighted its positive clinical and social application value. The use of longitudinal data from CHARLS allowed for a more accurate, data-driven understanding of aging-related disability trends, leveraging demographic and health-related variables highly relevant to the Chinese context. In addition, the inclusion of sarcopenia and frailty-related diagnostic indicators as predictive variables represents a novel aspect of this research. These indicators enhance the model's sensitivity to physical, cognitive, physiological, and psychological changes associated with aging that contribute to disability. This model offers a practical tool for improving disability prevention and management in older adults.

One notable observation in this study is the difference in AUC values between the training set (0.846) and the testing set (0.698). This discrepancy suggests potential overfitting, where the model may have captured noise or random fluctuations in the training data that do not generalize to unseen data. The observed discrepancy may be due to the imbalance in the dataset with respect to the outcome variable (disability status). To address this, we attempted oversampling to balance the data. However, this approach increased model complexity by retaining additional statistically significant predictors, raising concerns about overfitting. As a result, we proceeded with the original dataset to prioritize model simplicity and reduce overfitting risks. Although the model's discriminative performance decreased in the test set, the overall trends remain robust, providing valuable insights into disability prediction for older populations. Moving forward, we plan to conduct external

XSL•FO
RenderX

validation using additional datasets or longitudinal CHARLS follow-up data to further assess the model's generalizability across diverse settings and time points. We selected the XGBoost algorithm for modeling based on its demonstrated ability to handle complex datasets and address class imbalance by assigning higher weights to the minority class [8]. This approach enhances the model's ability to predict the minority class accurately. XGBoost's advantages over other algorithms include its superior performance with structured and unstructured data, its regularization techniques to reduce overfitting, and its efficiency in training large datasets. In addition, its gradient boosting framework captures complex variable interactions, making it more suitable than traditional linear models. To optimize performance, we used advanced techniques such as cross-validation and hyperparameter tuning, including a comprehensive grid search over key parameters. This rigorous process ensured the model was robust and well-calibrated, improving its predictive performance. By extending the analysis to the latest 2020 CHARLS data, this study offers a more comprehensive 5-year prediction window compared with previous work focused on shorter time frames.

SHAP summary charts clarified the role and importance of each variable in predicting disability, providing transparency and interpretability to the model. According to the SHAP chart, pain, respiratory function, and age were the top 3 factors in importance, with the wide distribution of SHAP values. This indicates that changes in these variables significantly alter the risk of disability. Chronic pain, particularly lower back and neck pain, is a leading cause of disability globally, as highlighted by the Global Burden of Disease Study 2015. These types of pain are major causes of years lived with disability in many regions, including Latin America, the Caribbean, most regions of Asia, Oceania, and sub-Saharan Africa [19]. Chronic pain is closely linked to functional disability and poor physical performance in the older, as supported by various studies [20,21]. Respiratory function declines with age in older people, and respiratory impairment accounts for 20.7% of all types of disability [22]. Maximal inspiratory pressure and maximal expiratory pressure are correlated with hand-grip strength and skeletal muscle mass index [23,24]. Respiratory sarcopenia, characterized by a decrease in respiratory muscle strength alongside systemic skeletal muscle with aging [25], can lead to deterioration in respiratory force generation, adversely affecting activities of daily living [26]. Overall, respiratory impairment is prevalent among older individuals and is linked to physical inactivity and poor performance-based mobility [27]. Age is an independent risk factor for disability, with intrinsic capacity and functional ability declining with age. Disability levels are highest in the oldest patients [28], and age correlates with increased pain and respiratory impairment. Older adults are more likely to experience these issues, further increasing their disability risk [29]. In summary, higher pain scores, poorer respiratory function, and older age are associated with a greater risk of disability. Clinically, this suggests the need for emphasis on pain management and respiratory exercises in old people, particularly for those with chronic respiratory diseases.

The concentrated distribution of SHAP values for handgrip strength, CS-5, and cognitive function indicates these variables significantly influenced disability prediction. The ablation experiments further confirmed the impact of handgrip strength and CS-5 on the model. When these sarcopenia-related features were removed, the model's performance was notably affected, as indicated by a significant increase in logloss. This finding aligns with the SHAP analysis results, which showed a wide distribution of values for these variables. Weak handgrip strength is identified as a key component of sarcopenia, strongly associated with subsequent disability and mortality [30]. Reduced handgrip strength and lower extremity strength, as measured by the CS-5, are strong predictors of functional impairment, disability, and low health-related quality of life, significantly increasing the risk of severe disability, frailty, and other health limitations in older adults [31,32]. Cognitive function is another crucial risk factor for disability. Studies have shown that cognitive decline is associated with ADL disability [33], and longitudinal research indicates that cognitive impairment may precede ADL disability, serving as a predictor of intermediate and late-stage ADL loss [34]. Physical and cognitive functions are closely related, with physical activity enhancing neurogenesis in the adult brain. Dual-task training, which enhances both cognitive and physical functions, has shown positive effects on cognitive function and physical activity in older individuals [35]. In summary, declines in handgrip strength, lower extremity strength (CS-5), and cognitive function are positively correlated with an increased risk of disability. This underscores the importance of targeted interventions, such as early muscle strength training for the upper and lower limbs and cognitive function exercises, to help reduce the risk of disability.

Although the roles of depression, comorbidities, and standing balance are less significant in predicting disability, they still contribute to its progression in the elderly and remain non-negligible factors. Depression, in particular, is a common psychological disorder among older adults and continues to be one of the most prevalent and disabling biopsychosocial conditions in this population. A Chinese cross-sectional study provided evidence of the association between depressive symptoms and ADL disability [36]. There is also a strong association between depression and physical activity, with significant mental health benefits gained from being physically active, even at levels below public health recommendations [37]. This may explain why depression can affect disability progression through physical function measures such as handgrip strength and CS-5. Similarly, comorbidities and standing balance are associated with disability and are critical factors in the multifactorial process of disability [38,39]. Comorbidity, the coexistence of 2 or more chronic diseases in older adults is a well-documented risk factor for increased mortality, reduced quality of life, and functional decline, ultimately leading to disability [40]. As a consequence of managing multiple chronic conditions, polypharmacy, defined as the concurrent use of multiple medications, becomes increasingly common in older populations [41]. Polypharmacy has been associated with a higher risk of falls, frailty, cognitive impairment, and adverse drug interactions, further exacerbating health deterioration and disability [42]. However, in this study, medication use was categorized in the CHARLS questionnaire only as Chinese traditional medicine or Western modern medicine, without

XSL•FO
RenderX

detailed data on specific medications. This limitation prevented a comprehensive analysis of polypharmacy's impact. The decline in standing balance in older adults indicates decreased postural control and increased risk of falls, often seen in populations with sarcopenia and frailty, which eventually progresses to disability [43]. Overall, these findings suggest that we should consider the combined effects of mental health, management of multiple chronic diseases, and balance function when predicting disability and formulating prevention strategies.

In our study, we evaluated the treatment benefits of the model using DCA. Figure 6A shows the net benefit across different risk thresholds. At low thresholds (<0.2), the net benefit was high but gradually declined as the threshold increased, approaching zero around 0.6. The model outperformed both all-treatment and no-treatment strategies across most thresholds, demonstrating its clinical utility. In practice, selecting an appropriate threshold is critical for clinical decision-making and resource allocation. A low threshold is suitable for high-sensitivity scenarios, such as community screening for early intervention in older adults. A medium threshold balances sensitivity and specificity, making it ideal for resource-limited settings where the model can precisely identify high-risk individuals for targeted interventions. Many clinicians have used DCA to test various disease prediction models, such as those for 30-day mortality in MIMIC-III patients with sepsis-3, major adverse cardiovascular events in older patients, and hypertension risk in patients with type 2 diabetes mellitus [44-46].

- We evaluated the model's predictive efficacy at different risk thresholds using CIC. As shown in Figure 6C, at thresholds below 0.2, the model identifies over 500 high-risk patients, with approximately 200 actual events, resulting in a high false-positive rate and increased resource consumption. This range is suitable for early widespread screening when follow-up resources are available. At thresholds between 0.2 and 0.5, the number of high-risk patients identified aligns more closely with actual events, balancing sensitivity and specificity while improving cost-effectiveness. This range is ideal for resource-limited settings. At thresholds above 0.5, the number of high-risk patients decreases significantly, nearly matching actual events but potentially missing some high-risk cases. The CIC provides clinicians with a visual tool to balance sensitivity and specificity, optimizing disability prediction and intervention strategies in older adults. CIC is commonly used to evaluate the predictive accuracy and clinical value of clinical prediction model for various diseases [47-49]. However, they are rarely used to evaluate the clinical usefulness of disability prediction models in older adults.

The 9 variables selected through LASSO regression form a streamlined yet effective set of predictors that can be easily integrated into routine clinical practice. Specifically, the inclusion of sarcopenia and frailty-related features provides health care professionals with clear and actionable insights into the ability of older adults to live independently, enabling timely interventions to prevent disability. By focusing on these key variables, the model remains interpretable, reducing the risk of "black-box" complexity in clinical decision-making. To facilitate practical application, we developed a web-based application via the Streamlit platform that uses these 9 predictors to calculate the 5-year risk of disability for individual patients. This user-friendly tool allows clinicians to input patient-specific data and receive immediate risk assessments, integrating predictive analytics into clinical workflows and bridging complex data models with everyday decision-making. Future applications of this model can aid healthcare professionals in identifying individuals at high risk of disability and implementing early, targeted interventions. This approach has the potential to delay the onset of disability and improve the quality of life for older individuals.

Limitations

The limitations of this study include the selection of predictor variables. While the selected predictors are based on the best available evidence, other important variables, such as activity intensity, were not included due to high missing values. In addition, the model generalization and optimization is a limitation. The model performs well on the internal test set but lacks external validation due to the unavailability of a suitable external dataset. We plan to collect data from multicenter older care communities for external validation to further improve and optimize the model. As the CHARLS database updates, the model may need periodic updates to maintain accuracy and usefulness. Moreover, predicting long-term disability risk is challenging due to complex time interactions that may alter the risk trajectory.

Conclusions

Our research incorporates parameters aligned with the diagnostic criteria for sarcopenia and frailty. These physical function measures are combined with predictors from cognitive and psychological health dimensions, recognizing the complex interplay of physical capability, aging, and mental health in the development of disability. This approach enhances the model's precision and considers the need for the efficient identification of at-risk individuals and the optimization of medical resources in clinical practice. Consequently, the model provides a highly reliable disability prediction tool for older patients, health care workers, and policymakers. In the future, we will adjust the model based on updates to the CHARLS database to ensure its suitability for the older population in China. In addition, we will seek appropriate external databases for validation and promote the model's application across different ethnic groups.

Acknowledgments

We would like to thank the staff and participants of CHARLS. Publicly available datasets can be found at http://charls.pku.edu.cn/
.

XSL•FO
RenderX

This research was supported by a grant from the National Key R&D Program of China (No.: 2020YFC2005601).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Model performance metrics at thresholds of 0.2 and 0.5 for the training and testing sets, respectively.
[DOC File, 32 KB - aging_v8i1e66723_app1.doc]

References

1. Foreman KJ, Marquez N, Dolgert A, et al. Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: reference and alternative scenarios for 2016-40 for 195 countries and territories. Lancet 2018 Nov 10;392(10159):2052-2090. [doi: 10.1016/S0140-6736(18)31694-5] [Medline: 30340847]
2. Wang S. Spatial patterns and social-economic influential factors of population aging: A global assessment from 1990 to 2010. Soc Sci Med 2020 May;253:112963. [doi: 10.1016/j.socscimed.2020.112963] [Medline: 32289647]
3. Luo Y, Su B, Zheng X. Trends and challenges for population and health during population aging - China, 2015-2050. China CDC Wkly 2021 Jul 9;3(28):593-598. [doi: 10.46234/ccdcw2021.158] [Medline: 34594944]
4. Oppewal A, Hilgenkamp TIM, van Wijck R, Schoufour JD, Evenhuis HM. Physical fitness is predictive for a decline in the ability to perform instrumental activities of daily living in older adults with intellectual disabilities: Results of the HA-ID study. Res Dev Disabil 2015;41-42:76-85. [doi: 10.1016/j.ridd.2015.05.002] [Medline: 26079525]
5. Jung SW, Yoon JH, Lee W. Predictors for depressive symptoms by four types of disability. Sci Rep 2021 Sep 29;11(1):19371. [doi: 10.1038/s41598-021-98765-4] [Medline: 34588530]
6. Dodge HH, Kadowaki T, Hayakawa T, Yamakawa M, Sekikawa A, Ueshima H. Cognitive impairment as a strong predictor of incident disability in specific ADL-IADL tasks among community-dwelling elders: the azuchi study. Gerontologist 2005 Apr;45(2):222-230. [doi: 10.1093/geront/45.2.222] [Medline: 15799987]
7. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the china health and retirement longitudinal study (CHARLS). Int J Epidemiol 2014 Feb;43(1):61-68. [doi: 10.1093/ije/dys203] [Medline: 23243115]
8. Han Y, Wang S. Disability risk prediction model based on machine learning among Chinese healthy older adults: results from the China Health and Retirement Longitudinal Study. Front Public Health 2023;11(1271595):1271595. [doi: 10.3389/fpubh.2023.1271595] [Medline: 38026309]
9. Gill TM, Williams CS, Tinetti ME. Assessing risk for the onset of functional dependence among older adults: the role of physical performance. J American Geriatrics Society 1995 Jun;43(6):603-609 [FREE Full text] [doi: 10.1111/j.1532-5415.1995.tb07192.x]
10. Kohout FJ, Berkman LF, Evans DA, Cornoni-Huntley J. Two shorter forms of the CES-D (Center for Epidemiological Studies Depression) depression symptoms index. J Aging Health 1993 May;5(2):179-193. [doi: 10.1177/089826439300500202] [Medline: 10125443]
11. Cheng ST, Chan ACM, Fung HH. Factorial structure of a short version of the Center for Epidemiologic Studies Depression Scale. Int J Geriat Psychiatry 2006 Apr;21(4):333-336. [doi: 10.1002/gps.1467]
12. Hu H, Cao Q, Shi Z, Lin W, Jiang H, Hou Y. Social support and depressive symptom disparity between urban and rural older adults in China. J Affect Disord 2018 Sep;237:104-111. [doi: 10.1016/j.jad.2018.04.076]
13. Mo S, Wang Y, Peng M, et al. Sex disparity in cognitive aging related to later-life exposure to ambient air pollution. Sci Total Environ 2023 Aug 15;886(163980):163980. [doi: 10.1016/j.scitotenv.2023.163980] [Medline: 37150467]
14. Hu Y, Peng W, Ren R, Wang Y, Wang G. Sarcopenia and mild cognitive impairment among elderly adults: The first longitudinal evidence from CHARLS. J Cachexia Sarcopenia Muscle 2022 Dec;13(6):2944-2952. [doi: 10.1002/jcsm.13081] [Medline: 36058563]
15. Wen X, Wang M, Jiang CM, Zhang YM. Anthropometric equation for estimation of appendicular skeletal muscle mass in Chinese adults. Asia Pac J Clin Nutr 2011;20(4):551-556. [Medline: 22094840]
16. Yang M, Hu X, Wang H, Zhang L, Hao Q, Dong B. Sarcopenia predicts readmission and mortality in elderly patients in acute care wards: a prospective study. J Cachexia Sarcopenia Muscle 2017 Apr;8(2):251-258. [doi: 10.1002/jcsm.12163] [Medline: 27896949]
17. Chen J, Zhang F, Zhang Y, et al. Trajectories network analysis of chronic diseases among middle-aged and older adults: evidence from the China Health and Retirement Longitudinal Study (CHARLS). BMC Public Health 2024 Feb 22;24(1):559. [doi: 10.1186/s12889-024-17890-7] [Medline: 38389048]
18. Lu J. Predicted probability of disability. 2024. URL: https://disability-predictionmodel.streamlit.app [accessed 2025-01-15]
19. GMaCoD C. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016 Oct 8;388(10053):1459-1544. [doi: 10.1016/S0140-6736(16)31012-1] [Medline: 27733281]

XSL•FO
RenderX

20. van der Gaag WH, Chiarotto A, Heymans MW, et al. Developing clinical prediction models for nonrecovery in older patients seeking care for back pain: the back complaints in the elders prospective cohort study. Pain 2021 Jun 1;162(6):1632-1640. [doi: 10.1097/j.pain.0000000000002161] [Medline: 33394879]
21. Nicholas M, Vlaeyen JWS, Rief W, et al. The IASP classification of chronic pain for ICD-11: chronic primary pain. Pain 2019 Jan;160(1):28-37. [doi: 10.1097/j.pain.0000000000001390] [Medline: 30586068]
22. Alene KA, Wangdi K, Colquhoun S, et al. Tuberculosis related disability: a systematic review and meta-analysis. BMC Med 2021 Sep 9;19(1):203. [doi: 10.1186/s12916-021-02063-9] [Medline: 34496845]
23. Watsford ML, Murphy AJ, Pine MJ. The effects of ageing on respiratory muscle function and performance in older adults. J Sci Med Sport 2007 Feb;10(1):36-44. [doi: 10.1016/j.jsams.2006.05.002]
24. Shin HI, Kim DK, Seo KM, Kang SH, Lee SY, Son S. Relation between respiratory muscle strength and skeletal muscle mass and hand grip strength in the healthy elderly. Ann Rehabil Med 2017 Aug;41(4):686-692. [doi: 10.5535/arm.2017.41.4.686] [Medline: 28971054]
25. Kera T, Kawai H, Hirano H, et al. Definition of respiratory sarcopenia with peak expiratory flow rate. J Am Med Dir Assoc 2019 Aug;20(8):1021-1025. [doi: 10.1016/j.jamda.2018.12.013] [Medline: 30737167]
26. Ohara DG, Pegorari MS, Oliveira Dos Santos NL, et al. Respiratory muscle strength as a discriminator of sarcopenia in community-dwelling elderly: a cross-sectional study. J Nutr Health Aging 2018;22(8):952-958. [doi: 10.1007/s12603-018-1079-4] [Medline: 30272099]
27. Vaz Fragoso CA, Beavers DP, Hankinson JL, et al. Respiratory impairment and dyspnea and their associations with physical inactivity and mobility in sedentary community-dwelling older persons. J Am Geriatr Soc 2014 Apr;62(4):622-628. [doi: 10.1111/jgs.12738] [Medline: 24635756]
28. Zhou Y, Li Y, Zhu X, Ma L. Medical and old-age care integration model and implementation of the integrated care of older people (ICOPE) in China: opportunities and challenges. J Nutr Health Aging 2021;25(6):720-723. [doi: 10.1007/s12603-021-1595-5] [Medline: 34179923]
29. Guido D, Leonardi M, Mellor-Marsá B, et al. Pain rates in general population for the period 1991-2015 and 10-years prediction: results from a multi-continent age-period-cohort analysis. J Headache Pain 2020 May 13;21(1):52. [doi: 10.1186/s10194-020-01108-3] [Medline: 32404046]
30. Dodds RM, Syddall HE, Cooper R, Kuh D, Cooper C, Sayer AA. Global variation in grip strength: a systematic review and meta-analysis of normative data. Age Ageing 2016 Mar;45(2):209-216. [doi: 10.1093/ageing/afv192] [Medline: 26790455]
31. Brooks JM, Titus AJ, Bruce ML, et al. Depression and handgrip strength among US. adults aged 60 years and older from NHANES 2011-2014. J Nutr Health Aging 2018;22(8):938-943. [doi: 10.1007/s12603-018-1041-5] [Medline: 30272097]
32. Coelho-Júnior HJ, Calvani R, Álvarez-Bustos A, et al. Physical performance and negative events in very old adults: a longitudinal study examining the ilSIRENTE cohort. Aging Clin Exp Res 2024;36(1):38345698. [doi: 10.1007/s40520-024-02693-y]
33. Garrett SL, Kennedy RE, Sawyer P, Williams CP, Brown CJ, Allman RM. Association between executive dysfunction and instrumental activities of daily living: racial and ethnic differences among community-dwelling older Adults in the Southeastern US. J Natl Med Assoc 2019 Jun;111(3):320-327. [doi: 10.1016/j.jnma.2018.11.002] [Medline: 30527966]
34. Sun Q, Jiang N, Lu N, Lou VWQ. Bidirectional relationship between cognitive function and loss hierarchy of activities of daily living among older adults with disabilities in urban China: a cross-lagged analysis. BMJ Open 2022 Sep;12(9):e057211. [doi: 10.1136/bmjopen-2021-057211]
35. Ali N, Tian H, Thabane L, et al. The effects of dual-task training on cognitive and physical functions in older adults with cognitive impairment; a systematic review and meta-analysis. J Prev Alzheimers Dis 2022;9(2):359-370. [doi: 10.14283/jpad.2022.16] [Medline: 35543010]
36. Feng Z, Li Q, Zhou L, Chen Z, Yin W. The relationship between depressive symptoms and activity of daily living disability among the elderly: results from the China Health and Retirement Longitudinal Study (CHARLS). Public Health (Fairfax) 2021 Sep;198:75-81. [doi: 10.1016/j.puhe.2021.06.023]
37. Pearce M, Garcia L, Abbas A, et al. Association between physical activity and risk of depression: a systematic review and meta-analysis. JAMA Psychiatry 2022 Jun 1;79(6):550-559. [doi: 10.1001/jamapsychiatry.2022.0609] [Medline: 35416941]
38. Vargese SS, Jylhä M, Raitanen J, Enroth L, Halonen P, Aaltonen M. Dementia-related disability in the population aged 90 years and over: differences over time and the role of comorbidity in the vitality 90 + study. BMC Geriatr 2023;23(1):37149593. [doi: 10.1186/s12877-023-03980-5]
39. Boyle PA, Wilson RS, Buchman AS, et al. Lower extremity motor function and disability in mild cognitive impairment. Exp Aging Res 2007;33(3):355-371. [doi: 10.1080/03610730701319210] [Medline: 17497375]
40. Yarnall AJ, Sayer AA, Clegg A, Rockwood K, Parker S, Hindle JV. New horizons in multimorbidity in older adults. Age Ageing 2017 Nov 1;46(6):882-888. [doi: 10.1093/ageing/afx150] [Medline: 28064173]
41. Masnoon N, Shakib S, Kalisch-Ellett L, Caughey GE. What is polypharmacy? A systematic review of definitions. BMC Geriatr 2017 Oct 10;17(1):230. [doi: 10.1186/s12877-017-0621-2] [Medline: 29017448]

XSL•FO
RenderX

42. van Dam CS, Labuschagne HA, van Keulen K, et al. Polypharmacy, comorbidity and frailty: a complex interplay in older patients at the emergency department. Eur Geriatr Med 2022 Aug;13(4):849-857. [doi: 10.1007/s41999-022-00664-y] [Medline: 35723840]
43. Billot M, Calvani R, Urtamo A, et al. Preserving mobility in older adults with physical frailty and sarcopenia: opportunities, challenges, and recommendations for physical activity interventions. Clin Interv Aging 2020;15:1675-1690. [doi: 10.2147/CIA.S253535] [Medline: 32982201]
44. Hou N, Li M, He L, et al. Predicting 30-days mortality for MIMIC-III patients with sepsis-3: a machine learning approach using XGboost. J Transl Med 2020 Dec 7;18(1):462. [doi: 10.1186/s12967-020-02620-5] [Medline: 33287854]
45. Zhang K, Liu C, Sha X, et al. Development and validation of a prediction model to predict major adverse cardiovascular events in elderly patients undergoing noncardiac surgery: a retrospective cohort study. Atherosclerosis 2023 Jul;376:71-79. [doi: 10.1016/j.atherosclerosis.2023.06.008] [Medline: 37315395]
46. Yang J, Wang X, Jiang S. Development and validation of a nomogram model for individualized prediction of hypertension risk in patients with type 2 diabetes mellitus. Sci Rep 2023 Jan 23;13(1):1298. [doi: 10.1038/s41598-023-28059-4] [Medline: 36690699]
47. Shen J, Xiao Z, Wang X, Zhao Y. A nomogram clinical prediction model for predicting urinary infection stones: development and validation in A retrospective study. World J Urol 2024;42(1):38573354. [doi: 10.1007/s00345-024-04904-7]
48. He M, Li C, Kang Y, Zuo Y, Duo L, Tang W. Clinical predictive model for the 1-year remission probability of IgA vasculitis nephritis. Int Immunopharmacol 2021 Dec;101:108341. [doi: 10.1016/j.intimp.2021.108341]
49. Hu P, Xu Y, Liu Y, et al. An externally validated dynamic nomogram for predicting unfavorable prognosis in patients with aneurysmal subarachnoid hemorrhage. Front Neurol 2021;12:683051. [doi: 10.3389/fneur.2021.683051] [Medline: 34512505]

Abbreviations

ADL: activities of daily living
ASM: appendicular skeletal muscle mass
AUC: area under the curve
CES-D: Center for Epidemiologic Studies Depression Scale
CHARLS: China Health and Retirement Longitudinal Study
CIC: clinical impact curve
CS-5: chair stand test
DCA: decision curve analysis
IADL: instrumental activities of daily living
LDL: low-density lipoprotein
ROC: receiver operating characteristic
SHAP: SHapley Additive exPlanations
TICS: telephone interview for cognitive status

Edited by P Jansons; submitted 21.09.24; peer-reviewed by J Hu, J Yue, L Ma, Y Qiu; revised version received 26.02.25; accepted 13.03.25; published 17.04.25.

Please cite as:
Chu J, Li Y, Wang X, Xu Q, Xu Z
Development of a Longitudinal Model for Disability Prediction in Older Adults in China: Analysis of CHARLS Data (2015-2020)
JMIR Aging 2025;8:e66723
URL: https://aging.jmir.org/2025/1/e66723
doi:10.2196/66723

© Jingjing Chu, Ying Li, Xinyi Wang, Qun Xu, Zherong Xu. Originally published in JMIR Aging (https://aging.jmir.org), 17.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Factors Influencing the Severity of Urinary and Defecatory Dysfunction Among the Middle-Aged and Older Adult Chinese Population: Longitudinal Study of a 5-Wave Survey Cohort

Haoran Zhu[1], PhD; Xiaoming Li[2], MPH; Liwei Jing[1], PhD; Jingli Kou[3], BSN; Lichao Gong[3], BSN; Rui Wang[1], PhD; Guangtian Liu[4], MSPT; Chao Zhang[5], PhD; Jiayi Zhao[6], MSN; Jing Zhang[7], PhD

[1]School of Nursing, Capital Medical University, 10 Xitoutiao, You'anmenwai, Fengtai District, Beijing, China

[2]School of Public Health, North China University of Science and Technology, Tangshan, China

[3]Department of Geriatrics, Xuanwu Hospital of Capital Medical University, Beijing, China

[4]College of Nursing and Rehabilitation, North China University of Science and Technology, Tangshan, China

[5]Department of Gastroenterology, Peking University Third Hospital, Beijing, China

[6]St Michael College, University of Toronto, Toronto, ON, Canada

[7]School of Public Health, Harbin Medical University, Harbin, China

Corresponding Author:
Liwei Jing, PhD
School of Nursing, Capital Medical University, 10 Xitoutiao, You'anmenwai, Fengtai District, Beijing, China

Abstract

Background: Urinary and defecatory dysfunction (UDD) is a significant concern among the aging population in China. However, there is a lack of longitudinal research exploring the risk factors of UDD severity in Chinese older adults.

Objective: This study uses data from the China Health and Retirement Longitudinal Study spanning 2011 to 2020 to explore UDD risk factors in the middle-aged and older adult Chinese population, focusing on epidemiological characteristics and potential influences on severity.

Methods: A longitudinal cohort of over 10,000 participants from the China Health and Retirement Longitudinal Study was analyzed across 5 waves using Bayesian logistic regression. This analysis examined associations between UDD severity and factors including demographic, lifestyle, and health-related factors, including comorbidities, BMI, and handgrip strength.

Results: Higher UDD prevalence was observed among female population, older adults, those with low education levels, and rural residents. Depression, arthritis, and low handgrip strength emerged as critical predictors of severe UDD. Additionally, abnormal BMI, both underweight (odds ratio [OR] 3.019, 95% CI 1.484 - 5.951; *P*=0.002) and obesity (OR 2.697, 95% CI 1.338 - 5.217; *P*=0.005), was strongly linked to increased severity and persistence of UDD. Participants aged 66 years and older exhibited the highest UDD prevalence, with both underweight and obese individuals facing the greatest risk of persistent and worsening symptoms.

Conclusions: This study is the first to longitudinally examine the risk factors of UDD severity in China's middle-aging and aging population. The findings underscore the need for targeted interventions focusing on muscle strength rehabilitation and comorbidity management to mitigate UDD progression, contributing to improved quality of life for older individuals.

(JMIR Aging 2025;8:e70541) doi:10.2196/70541

KEYWORDS

China; characteristic; risk factors; urination and defecatory dysfunction; longitudinal cohort study

Introduction

Population aging has become a public health issue, imposing a major global challenge across countries in recent decades [1]. The population aged 60 years and older in China is projected to exceed 300 million by 2025, accounting for more than 20% of the total population [2]. Due to the decline in muscle function and disorders of the nervous system among older people, particularly hormonal changes in female population that lead to muscle changes, and urethral sphincter atrophy in male population, millions experience urinary and defecatory dysfunction (UDD) [3-7].

The World Health Organization's integrated care for older people model emphasizes the significance of UDD, particularly urinary incontinence, highlighting its impact on health systems worldwide. Additionally, the care burden associated with urinary retention, fecal incontinence, and constipation is substantial, underscoring the need for comprehensive strategies to address

XSL•FO
RenderX

these pervasive health issues [8]. UDD is defined as a disorder that affects the storage or elimination of urine and feces [9]. Recent studies on UDD reveal varying prevalence rates. Research shows 37.1% global prevalence for female urinary incontinence [10], 18.9% for constipation in older adults [11], and 8% for fecal incontinence [12]. However, these studies, mainly from China, are limited by small, regional samples and a lack of comprehensive data on influencing factors, highlighting the need for broader research to fully understand UDD's impact. This common disorder can have a negative impact on the quality of life of patients and exacerbate their economic burden, particularly in long-term or severe cases [13-16].

Numerous studies have demonstrated that UDD may be associated with chronic conditions (such as hypertension, diabetes, stroke, depression, and so on), abnormal BMI (overweight, obesity, or underweight), handgrip strength, and unhealthy lifestyle factors such as smoking and drinking [10,17-22]. Due to limitations in existing cross-sectional studies, which cannot confirm causal relationships, we established a cohort of over 10,000 participants, with an innovative focus on UDD severity, to study the factors influencing its severity.

The study aimed to use data from a 5-wave survey conducted by the China Health and Retirement Longitudinal Study (CHARLS) [23] to (1) describe the characteristics of UDD from 2011 to 2020 and (2) to identify the risk factors impacting the severity of UDD in China.

Methods

Study Design and Participants

The CHARLS, a nationally representative survey targeting individuals aged 45 years and older from 450 villages or communities in 28 provinces and 150 counties or districts across Mainland China, was conducted by the National School of Development at Peking University from 2011 to 2020 [23]. A total of 150 counties were randomly selected, with stratification by region and urban or rural classification. Administrative villages and neighborhoods were designated as primary sampling units. Households with at least 1 member aged 45 years or older were selected, and 1 individual, along with their spouse, was interviewed. The CHARLS database encompasses a broad array of variables related to demographic statistics, socioeconomic conditions, and health status. Participants in CHARLS are interviewed face-to-face every 2 years using a computer-assisted personal interview technique.

To map the characteristics and temporal trends in UDD prevalence, all participants with UDD were identified in each wave of the cross-sectional survey from 2011 to 2020. For identification of risk factors, participants with UDD who enrolled in 2011 and were continuously followed up in the subsequent 4 cross-sectional surveys were included. Detailed information about the sampling design used in the survey has been described in earlier publications [24].

Ethical Considerations

The CHARLS was approved by the Biomedical Ethics Review Committee of Peking University (IRB0000105211015) [23]. All participants provided signed informed consent forms and received modest financial compensation. Data were anonymized before release and are only available to approved researchers.

Data Collection

In each county or district, trained staff collected data at participants' homes and local community health centers as well as County Centers for Disease Prevention and Control in accordance with the reported protocol [23]. Deidentified information was collected for analysis, consisting of data on participants' demographic characteristics, health-related behaviors and outcomes, childhood conditions, community environment, cognitive and physical health, current economic status, social and family support, health insurance, health care use, comorbidities (chronic diseases such as hypertension and diabetes), and their urinary and defecation control abilities.

Referring to the World Health Organization standard, participants were divided into the following age groups: 45 - 59 years (middle-aged), 60 - 74 years (young-old), 75 - 89 years (old), and ≥90 years (very old) [25]. Residences of participants included rural and urban areas. Education levels were categorized as illiterate, primary school graduate, secondary school graduate, and college graduate or above. Regions of China were classified as Southwest, Southern, Eastern, Northwest, Northern, and Northeast.

The CHARLS did not include data from Ningxia Hui Autonomous Region, Xizang Autonomous Region, Hainan province, Hong Kong, Macau, and Taiwan province. Additionally, the 2020 follow-up survey in the CHARLS did not include participants from the Xinjiang Uygur Autonomous Region due to the impact of the COVID-19 pandemic.

Case Identification and Study Outcomes

According to the International Classification of Functioning, Disability, and Health, urinary function refers to the ability to discharge urine from the bladder, while bowel function refers to the ability to expel waste materials and undigested food in the form of feces, along with the associated physiological processes [26]. Therefore, impairment in either of these functions is classified as UDD. Based on participants' responses to the question "Do you have any difficulties with controlling urination and defecation?" in CHARLS, all participants reporting "I have difficulty" or "I cannot do it" were identified as having UDD.

The second aim of the study was to explore the risk factors for UDD. To this end, participants with UDD were divided into the 4 subgroups based on clinical expert experience and the actual disease status of the participants, each comprising, individuals with increasingly severe UDD, from group A to group D, in order, as follows:

- Group A: Participants who had never experienced UDD.
- Group B: Participants who identified as having UDD in 1 or 2 consecutive surveys, without recurrence in the later follow-ups.
- Group C: Participants who reported having experienced UDD in 1 or 2 consecutive surveys and had recovered at the time of the subsequent 1 or 2 follow-ups but eventually experienced recurrence.

XSL•FO
RenderX

- Group D: Participants who identified as having UDD in consecutive 3 or more surveys.

Statistical Analysis

To achieve the first study aim, a descriptive analysis was conducted to examine the characteristic of prevalence of UDD by sex, residence, marital status, age group, education level, and geographic region over 5 waves from 2011 to 2020. The calculation of the prevalence rate is the cumulative number of cases divided by the total number of participants followed up per year.

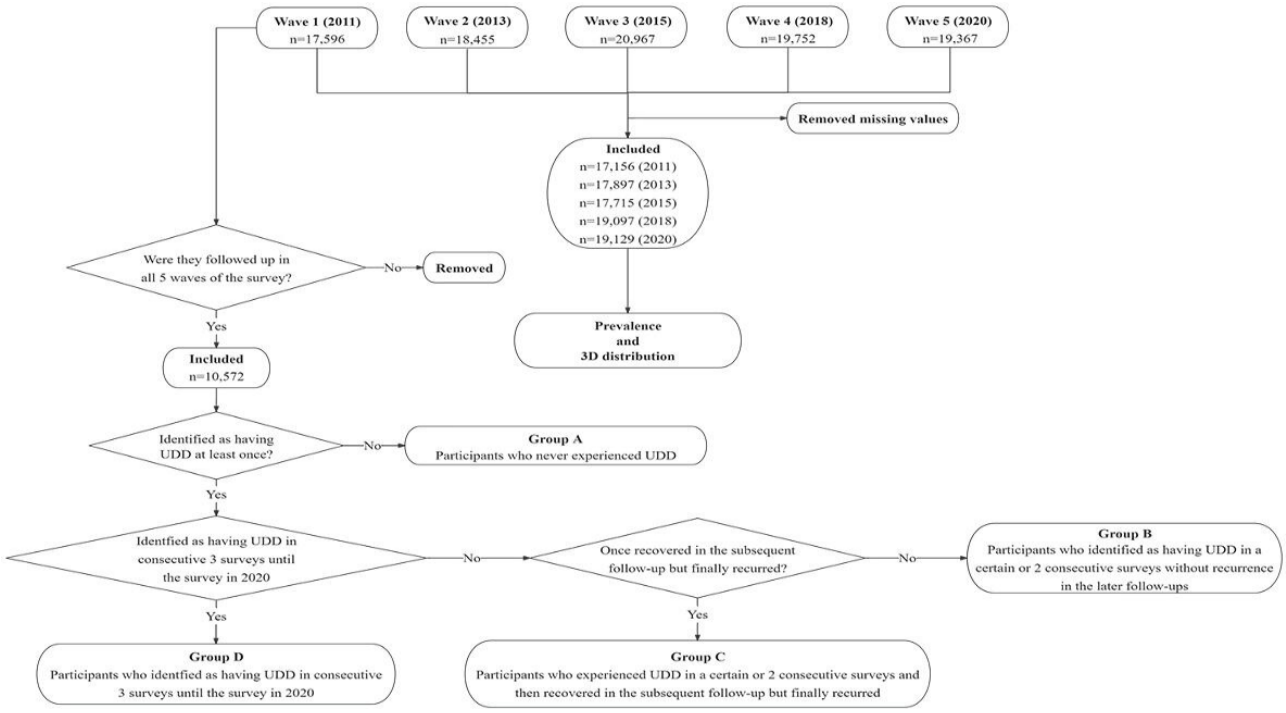
To achieve the second study aim, a chi-square test was initially performed to identify potential risk factors among the 4 groups. These factors were then further examined through Bayesian logistic regression analysis. Prior to regression analysis, chi-square tests were conducted to assess univariate associations. A weakly informative prior was used to ensure robust parameter estimation. Parameter estimation was conducted using Markov Chain Monte Carlo sampling, and model results are presented as odds ratios (ORs) with 95% CIs. To enhance interpretability, exponentiation was applied to the estimated coefficients, and 2-sided *P* values were calculated. R (version 64 4.4.3; R Foundation for Statistical Computing) was used for descriptive analysis and logistic regression. *P* values were 2-sided with a significance level of .05.

Results

Characteristics of Study Participants in the 5 Waves

The participants in the 5 waves comprised 17,156, 17,897, 17,715, 19,097, and 19,129 individuals, respectively. Although there were no significant differences between sexes, significant differences were found among age groups, places of residence, education levels, marital status, and regions across the 5 waves. Figure 1 shows the screening process of this study. Table 1 presents the characteristics of patients with UDD across various groups.

Figure 1. Flowchart of participant inclusion or exclusion and grouping diagram. UDD: urinary and defecatory dysfunction.



XSL•FO
RenderX

Table . Demographic characteristics of participants with urinary and defecatory dysfunction in the 5-wave survey from 2011 to 2020.

	Baseline wave in 2011 (n=17,156), n (%)	Second wave in 2013 (n=17,897), n (%)	Third wave in 2015 (n=17,715), n (%)	Fourth wave in 2018 (n=19,097), n (%)	Fifth wave in 2020 (n=19,129), n (%)	*P* value
Sex						.98
Female	388 (4.41)	349 (3.77)	429 (4.66)	491 (4.89)	513 (5.09)	
Male	365 (4.63)	372 (4.31)	401 (4.71)	411 (4.54)	447 (4.94)	
Age group (years)						<.001
45 - 50	57 (1.49)	54 (1.46)	47 (1.62)	31 (1.30)	32 (1.27)	
51 - 55	59 (2.11)	58 (2.18)	76 (2.43)	54 (1.47)	73 (1.97)	
56 - 60	130 (3.72)	95 (2.75)	97 (3.34)	67 (2.48)	90 (3.21)	
61 - 65	132 (4.91)	133 (4.31)	136 (4.19)	132 (3.87)	161 (4.69)	
66 - 70	106 (5.92)	104 (5.15)	138 (5.99)	185 (6.24)	199 (6.66)	
71 - 75	92 (7.55)	110 (7.71)	129 (8.40)	134 (7.49)	154 (8.75)	
76 - 80	67 (8.54)	80 (8.71)	111 (11.28)	144 (12.04)	122 (11.37)	
≥81	110 (20.04)	87 (13.30)	96 (13.64)	155 (15.78)	129 (15.64)	
Education						<.001
Illiterate[a]	476 (6.15)	414 (5.09)	482 (5.99)	545 (6.48)	537 (6.51)	
Primary[b]	146 (3.98)	147 (3.88)	177 (4.59)	185 (4.42)	205 (4.90)	
Second[c]	122 (2.29)	150 (2.70)	162 (3)	162 (2.64)	182 (3.13)	
College[d]	9 (2.14)	10 (2.43)	9 (2.17)	10 (2.63)	36 (4.17)	
Marital status						<.001
Married[e]	593 (3.96)	560 (3.61)	625 (4.09)	635 (3.92)	695 (4.34)	
Other status[f]	160 (7.30)	161 (6.78)	205 (8.37)	267 (9.19)	265 (8.51)	
Residence						<.001
Rural	627 (4.72)	585 (4.27)	637 (4.70)	738 (4.87)	728 (5.11)	
City	122 (3.24)	131 (3.27)	180 (4.64)	163 (4.17)	143 (4.99)	
Region						<.001
Northeast	66 (5.18)	52 (3.75)	59 (4.52)	68 (5.16)	64 (5.41)	
East	222 (4.27)	159 (2.93)	201 (3.63)	213 (3.62)	263 (4.38)	
North	87 (3.72)	124 (5.13)	119 (5.19)	146 (5.75)	136 (5.64)	
Central	94 (3.49)	119 (4.26)	126 (4.61)	150 (5.11)	162 (5.32)	
South	42 (2.76)	49 (3.25)	64 (4.37)	58 (3.66)	78 (4.75)	
Southwest	199 (6.84)	163 (5.25)	217 (7.04)	204 (6.06)	185 (5.39)	
Northwest	43 (3.51)	55 (4.40)	44 (3.39)	63 (4.29)	72 (5.18)	

[a]Without formal education, did not finish primary school, was homeschooled.

[b]With primary school education.

[c]With middle school or high school education.

[d]With college education and above.

[e]Married with spouse present or married but not living with spouse temporarily for reasons such as work.

[f]Separated, divorced, widowed, or never married.

XSL•FO
RenderX

Trends and Distributions of UDD Prevalence From 2011 to 2020

The prevalence of UDD from 2011 to 2020 was 4.39% (n=753), 4.03% (n=721), 4.69% (n=830), 4.72% (n=902), and 5.02% (n=960), respectively. Figure 2 demonstrates that the prevalence of UDD by sex and residence rose slowly over the 10 years, with female participants or participants living in rural areas having higher prevalence than their counterparts. The prevalence of UDD by education level and marital status remained essentially stable across the decade in China, although illiterate or married participants had higher prevalence than other categories.

Figure 2. Temporal trends of prevalence of urinary and defecatory dysfunction by sex, education level, marital status, and residence among middle-aged and older people in China from 2011 to 2020.

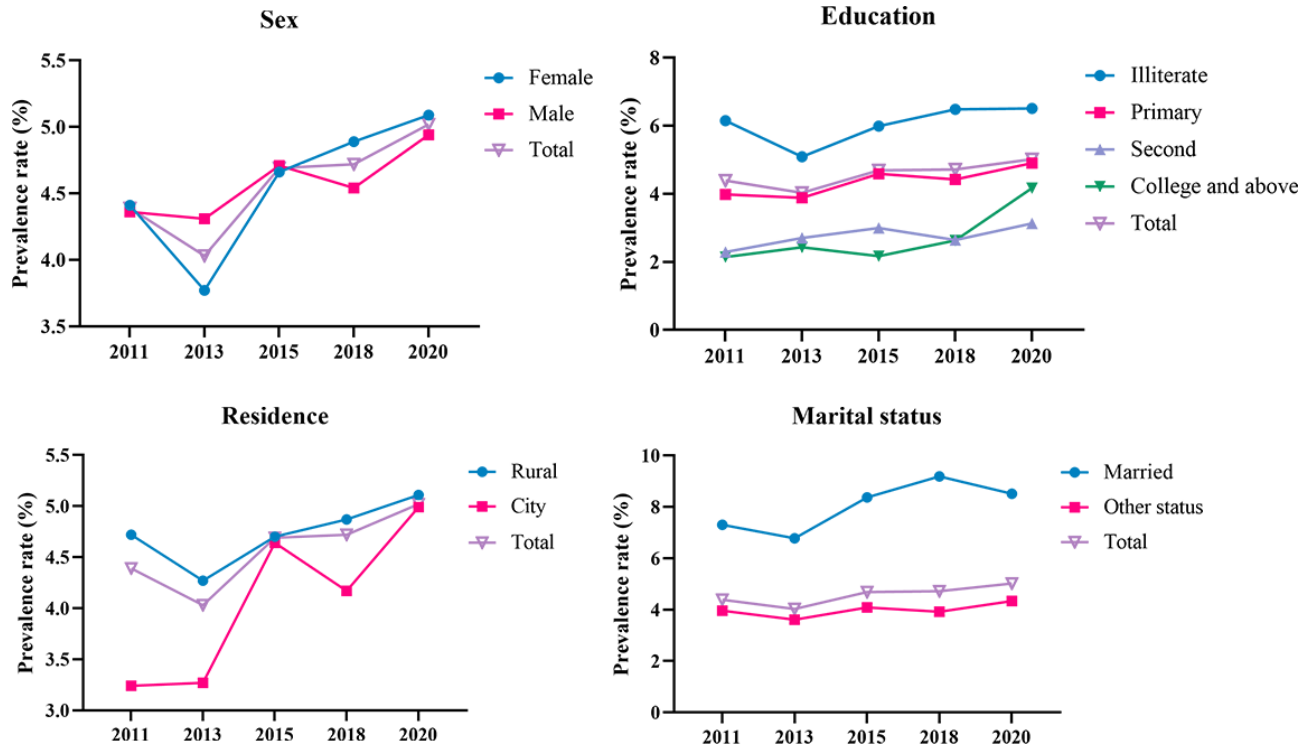
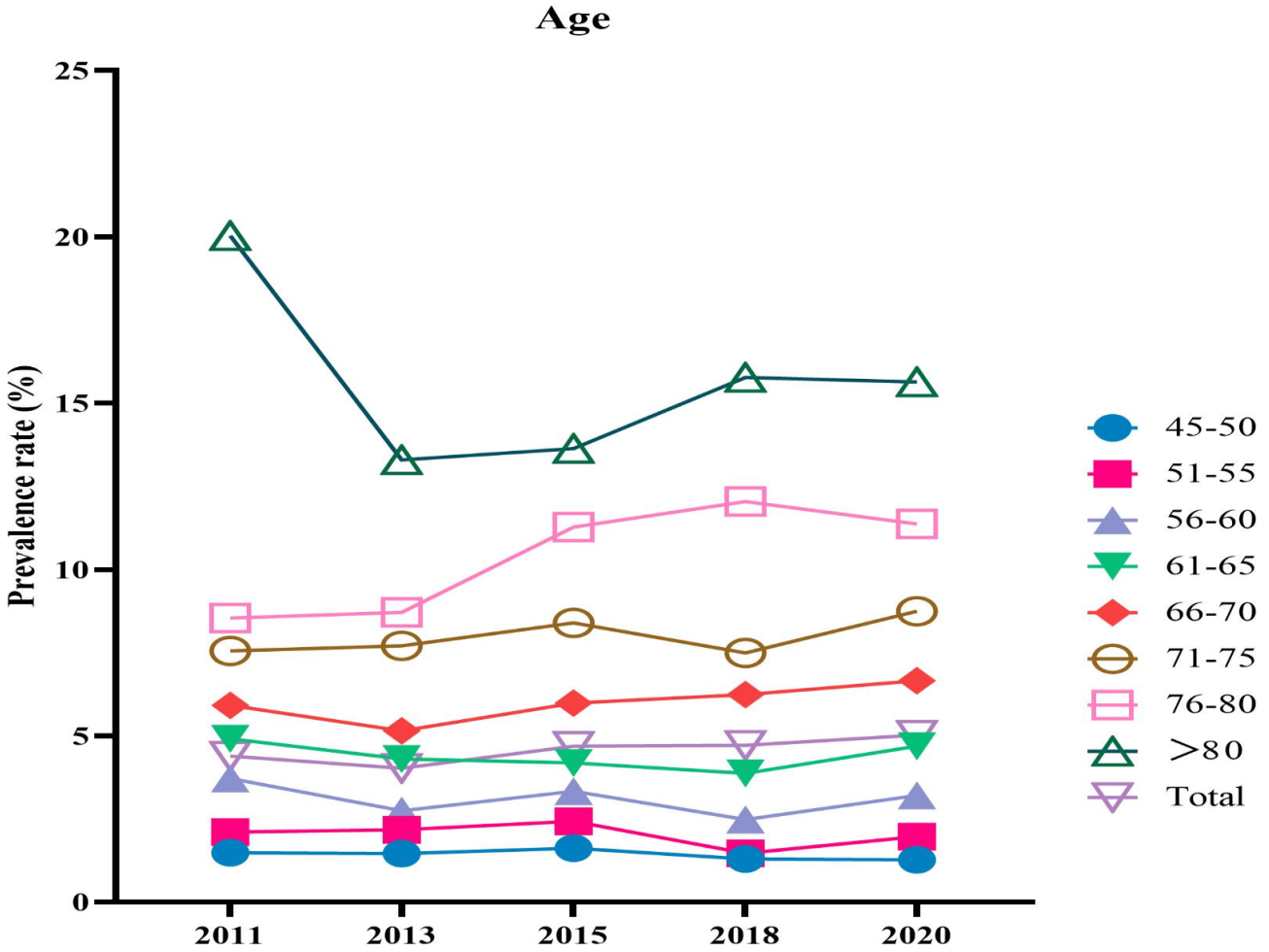


Figure 3 illustrates that the prevalence of UDD gradually increased with age over the decade, particularly among those aged 66 years and older who had a higher prevalence than the overall prevalence. The prevalence of the 4 groups aged 66 years and older slightly increased, while the corresponding figures for the 4 younger groups younger than 65 years were almost stable during the 10 years of surveys.

XSL•FO
RenderX

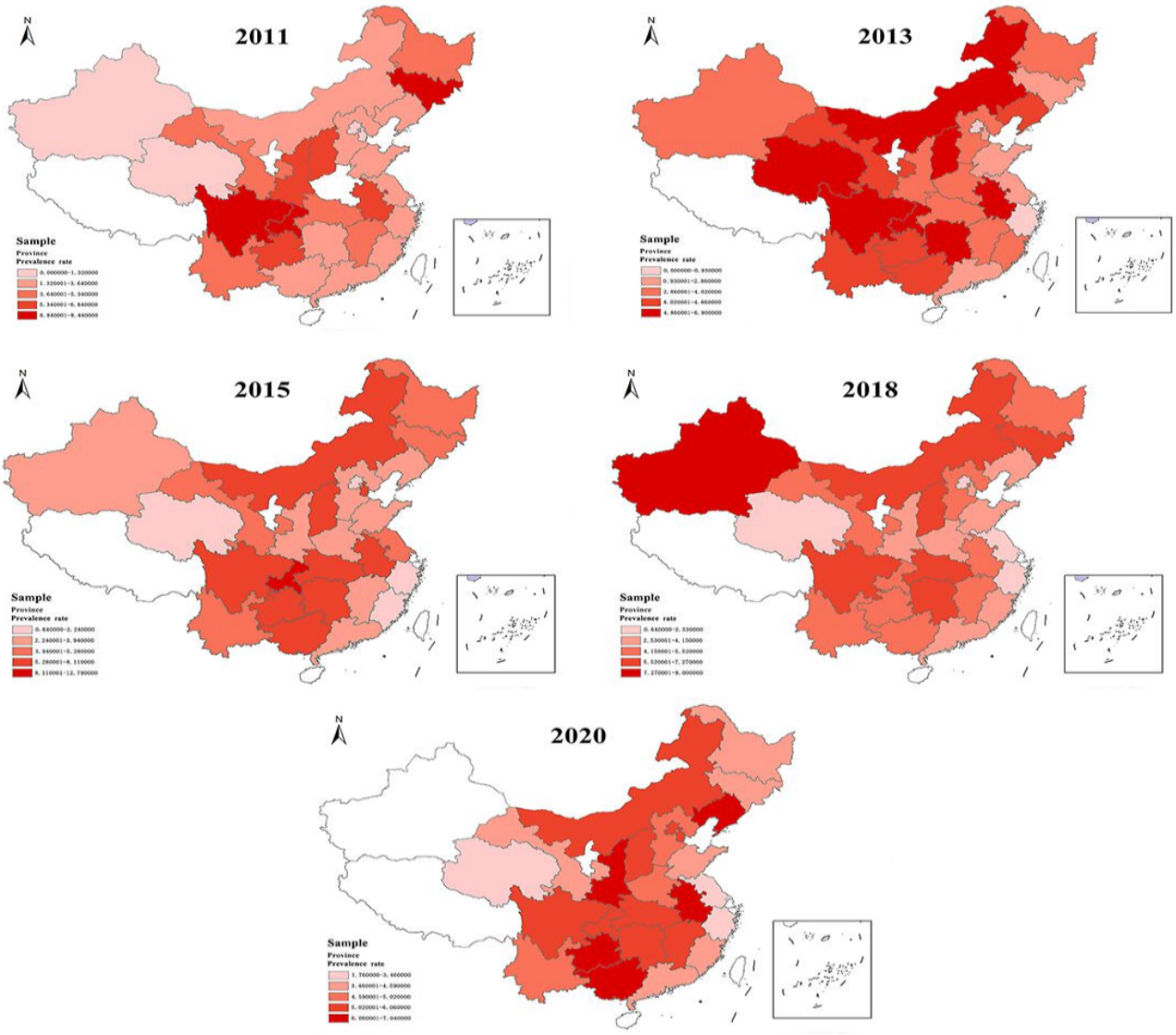
Figure 3. Temporal trends of age-specific prevalence (in years) of urinary and defecatory dysfunction among middle-aged and older people in China from 2011 to 2020.



The prevalence of UDD varied significantly across different regions over the years, being consistently higher in the Southwest region than in other regions. Both the Southwest and Northwest regions saw an increasing prevalence of UDD in the past 10 years (Figure 4).

XSL•FO
RenderX

Figure 4. Regional distribution of urinary and defecatory dysfunction prevalence among middle-aged and older people in China from 2011 to 2020 (because of the impact of the pandemic, information from Xinjiang participants in 2020 was not included in the China Health and Retirement Longitudinal Study database).



Identification of Risk Factors for Different UDD Conditions in a Cohort Study

There were 10,572 participants enrolled in the cohort study; among them, 8966 participants had never experienced UDD over the 10-year period, while 1275 had recovered, 245 reported recurrence, and 86 reported persistence at the last follow-up. There were significant differences for factors other than sex (P=.19), cancer (P=.12), and smoking (P=.39). Among all groups with various severity of UDD, the proportion of female participants was slightly higher than that of male participants. Higher proportions of participants were aged 60 - 75 years, illiterate, married, living in rural areas, living in the Southwest and East regions, with normal BMI, nonsmokers, and nondrinkers. The comorbidities of participants with UDD are shown in Table 2.

XSL•FO
RenderX

Table . Univariable analysis of 4 groups of urinary and defecatory dysfunction (UDD).

	Without UDD[a] (group A; n=8966), n (%)	Recovered from UDD[b] (group B; n=1275), n (%)	Reported recurrent[c] UDD (group C; n=245), n (%)	Reported persistent UDD[d] (group D; n=86), n (%)	*P* value[e]
Sex					
Female	4751 (52.99)	717 (56.24)	132 (53.88)	47 (54.65）	.19
Age group (years)					<.001
45 - 59	5604 (62.50)	555 (43.53)	85 (34.69)	30 (34.88)	
60 - 75	3015 (33.63)	608 (47.69)	137 (55.92)	47 (54.65)	
76 - 90	343 (3.83)	109 (8.55)	22 (8.98)	8 (9.30)	
＞90	4 (0.04)	3 (0.24)	1 (0.41)	1 (1.16)	
Education					<.001
Illiterate	3935 (43.89)	715 (56.08)	147 (60)	47 (54.65)	
Primary	1967 (21.94)	289 (22.67)	46 (18.78)	26 (30.23)	
Second	2918 (32.55)	260 (20.39)	50 (20.41)	12 (13.95)	
College and above	146 (1.63)	11 (0.86)	2 (0.82)	1 (1.16)	
Marital status					
Married	8169 (91.11)	1088 (85.33)	210 (85.71)	74 (86.05)	<.001
Residence					
Rural	7400 (82.53)	1115 (87.45)	213 (86.94)	74 (86.05)	.002
Region					<.001
Northeast	557 (6.21)	75 (5.88)	15 (6.12)	6 (6.98)	
East	2891 (32.24)	351 (27.53)	63 (25.71)	22 (25.58)	
North	1149 (12.82)	183 (14.35)	30 (12.24)	11 (12.79)	
Central	1482 (16.53)	215 (16.86)	40 (16.33)	14 (16.28)	
South	735 (8.20)	84 (6.59)	13 (5.31)	4 (4.65)	
Southwest	1521 (16.96)	279 (21.88)	68 (27.76)	26 (30.23)	
Northwest	631 (7.04)	88 (6.90)	16 (6.53)	3 (3.49)	
BMI[e]					<.001
Underweight	466 (5.41)	83 (6.83)	24 (10.57)	14 (17.07)	
Normal weight	4486 (52.09)	636 (52.35)	113 (49.78)	29 (35.37)	
Overweight	2615 (30.36)	330 (27.16)	64 (28.19)	20 (24.39)	
Obese	1018 (11.82)	157 (12.92)	25 (11.01)	17 (20.73)	
Smoking[e]	3434 (38.32)	483 (37.97)	107 (43.67)	33 (38.37)	.39
Drinking[e]	3065 (34.21)	342 (26.89)	83 (33.88)	25 (29.07)	<.001
Comorbidities					
Hypertension[e]	1882 (21.09)	387 (30.57)	71 (29.22)	32 (37.21)	<.001
Dyslipidemia[e]	772 (8.64)	144 (11.31)	30 (12.35)	11 (12.79)	.002
Diabetes[e]	424 (4.73)	83 (6.52)	22 (9.05)	10 (11.63)	<.001
Cancer	76 (0.85)	19 (1.49)	1 (0.41)	1 (1.16)	.12
Chronic lung disease[e]	701 (7.84)	184 (14.51)	40 (16.39)	20 (23.26)	<.001
Liver disease[e]	320 (3.57)	59 (4.63)	16 (6.53)	3 (3.49)	.03

XSL•FO
RenderX

	Without UDD[a] (group A; n=8966), n (%)	Recovered from UDD[b] (group B; n=1275), n (%)	Reported recurrent[c] UDD (group C; n=245), n (%)	Reported persistent UDD[d] (group D; n=86), n (%)	*P* value[e]
Heart problems[e]	854 (9.53)	210 (16.51)	45 (18.44)	18 (21.18)	<.001
Stroke[e]	125 (1.40)	41 (3.24)	12 (4.92)	6 (7.06)	<.001
Kidney disease[e]	472 (5.27)	110 (8.63)	35 (14.29)	14 (16.28)	<.001
Digestive disease[e]	1954 (21.86)	370 (29.20)	96 (39.18)	37 (43.53)	<.001
Memory-related disease[e]	66 (0.74)	26 (2.05)	12 (4.90)	6 (6.98)	<.001
Arthritis or rheumatism[e]	2839 (31.74)	578 (45.62)	141 (57.55)	50 (58.14)	<.001
Asthma[e]	239 (2.68)	53 (4.18)	24 (9.84)	6 (6.98)	<.001
Depression[e]	3334 (37.18)	748 (58.67)	174 (71.02)	65 (75.58)	<.001
Handgrip strength					
Low handgrip strength	1005 (11.84)	289 (24.41)	70 (31.39)	30 (37.04)	<.001

[a]Participants who had never experienced UDD.

[b]Participants who identified as having UDD in 1 or 2 consecutive surveys without recurrence in the later follow-ups.

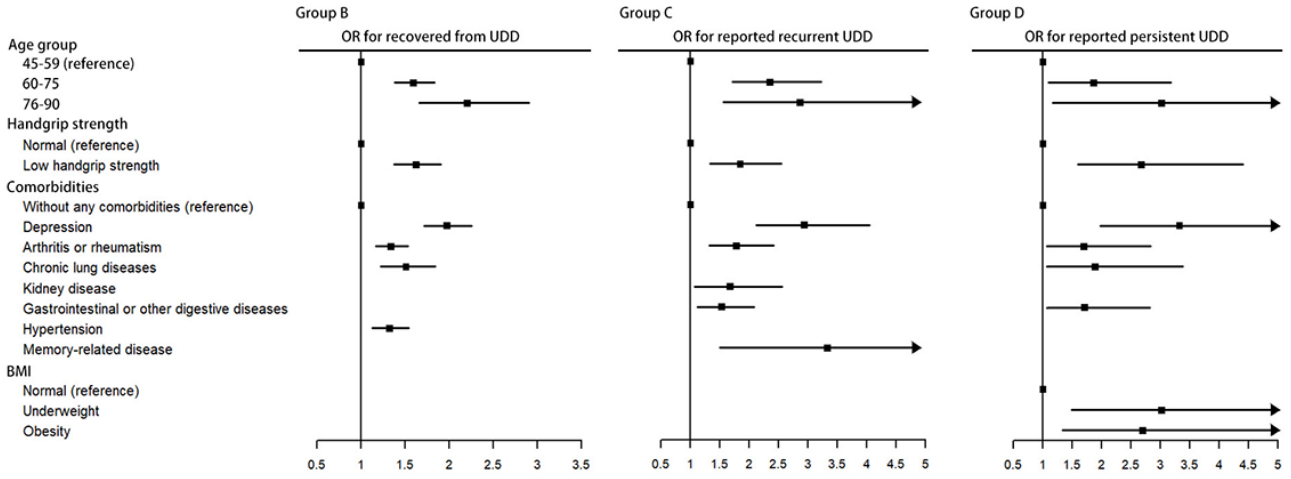
[c]Participants who reported experiencing UDD in 1 or 2 consecutive surveys and having recovered in the subsequent 1 or 2 follow-ups but eventually experienced recurrence.

[d]Participants who identified as having UDD in 3 or more consecutive surveys.

[e]Missing data: 436 for BMI, 7 for smoking, 10 for drinking, 53 for hypertension, 30 for dyslipidemia, 14 for diabetes, 27 for chronic lung disease, 4 for liver disease, 13 for heart problem, 36 for stroke, 5 for kidney disease, 37 for digestive disease, 42 for memory-related disease, 29 for arthritis or rheumatism, 52 for asthma, and 593 for handgrip strength.

Figure 5 shows the ORs and 95% CIs of the risk factors for UDD in patients with different conditions. Four common risk factors were identified across the 3 groups, namely, older age (60 - 89 years), depression, arthritis (or rheumatism), and handgrip strength. For group C and group D, who had more severe UDD than group B, participants with gastrointestinal diseases had a higher risk of experiencing recurrent or persistent UDD. Memory-related diseases (OR 3.328, 95% CI 1.505 - 6.836; *P*=.002) may cause UDD to recur. Additionally, underweight (OR 3.019, 95% CI 1.484 - 5.951; *P*=.002) and obesity (OR 2.697, 95% CI 1.338 - 5.217; *P*=.005) were identified as potential independent risk factors for the persistent nature of UDD among middle-aged and older people in China. The detailed results of the Bayesian logistic regression model are provided in Multimedia Appendix 1.

Figure 5. OR and 95% CI of risk factors for UDD in different population groups with UDD. Forest plots show ORs and 95% CIs for group (B) recovered from UDD, (C) reported UDD recurrence, and (D) UDD persisted adjusted for sex, age, education, marital status, residence, region, BMI, handgrip strength, smoking status, drinking status, and comorbidities. In the forest plot, only the significant results within each group (*P*<.05) are shown. The reference group for groups B, C, and D consists of participants who had never experienced UDD. OR: odds ratio; UDD: urinary and defecatory dysfunction.



XSL•FO
RenderX

Discussion

Principal Findings

This is the first study to use a representative sample to describe the temporal, spatial, and interpersonal distributions of UDD among middle-aged and older people in China. The study identified that over the past decade, the prevalence of UDD has remained stable across different education levels, kinds of marital status, and age groups. However, it significantly increased among both sexes, urban residents, and in the Northwest and Southwest regions. Participants who were female, illiterate, or married and those aged 66 years and older exhibited a higher prevalence of UDD. The incidence, recurrence, and persistence of UDD were more likely to be influenced by aging and comorbidities such as depression, arthritis (or rheumatic diseases), and handgrip strength in China. More importantly, being underweight or obese might contribute to the long-term persistence of UDD.

The Prevalence and Time Trends of UDD

Due to its increased prevalence, the disease burden associated with UDD has shown a gradually escalating trend in China over the past 10 years. However, despite a slight increase, the prevalence of UDD in China (5.02% in 2020) has remained lower than the global prevalence, which ranges from 7% to 55%. This lower prevalence might be attributed to lower self-report due to participants' potential embarrassment or their belief that UDD is a natural aging-related phenomenon among the Chinese population [10,21,27,28]. Approximately, 50% to 67% of patients were unwilling to report their UDD conditions to health care providers [29,30]. The slowly increasing trend in prevalence observed in this study may reflect improved awareness and diagnosis of UDD over time, leading to more UDD cases being identified.

Over the past decade, the prevalence of UDD in female individuals has consistently remained higher than that in male individuals. This finding underscores the necessity of addressing the unique physiological and reproductive health challenges that female individuals face, as the relaxation of pelvic floor muscles during childbirth significantly contributes to this increased prevalence [31]. This suggests that we should place greater emphasis on female individual's pelvic health. By implementing early interventions and training, such as pelvic floor muscle exercises, we can significantly improve the quality of life for female individuals in their later years [32].

It was notable that the Northwest and Southwest provinces—2 resource-limited regions in China—experienced an obviously higher prevalence of UDD during the 5-wave survey. The evidence aligned with other studies demonstrating that economic status may influence the prevalence of UDD, while UDD places a significant economic burden on patients too [14,15,33]. The higher prevalence of UDD among participants from resource-limited areas may be related to the insufficient health insurance or poor access to medical services, such as postpartum pelvic floor rehabilitation [34,35]. Therefore, economic development may serve as a solution to improve the quality of life of patients with UDD in these regions along with enhancing the accessibility of medical resources.

Given the impact of UDD, it is necessary to establish a surveillance system for UDD to provide more evidence for the identification of the long-term impact of UDD and the development of prevention and treatment strategies at the population level.

Population Aging, Comorbidities, and Low Handgrip Strength May Lead to All Severity of UDD Among the Chinese Population as Marked Risk Factors

Population aging is an important risk factor influencing the occurrence, recurrence, and persistence of UDD. This study suggests that the physiological senescence associated with aging may contribute to the development and exacerbation of UDD, particularly among participants aged 66-90 years. It is widely recognized that aging can lead to the occurrence or aggravation of UDD, likely due to the frailty and functional impairment associated with aging [13,36-39]. However, the potential mechanism underlying the association between frailty and UDD remains unknown. Finite element models can make it possible to explore the mechanism underpinning the relationship between frailty and UDD [40]. By comparing defecation outcomes under different parameter settings, the association between training methods and urinary and fecal control ability can be quantified, providing accurate guidance for rehabilitation.

This study demonstrated that comorbidities are the second important risk factor for UDD. For instance, depression was found to significantly impact the incidence, recurrence, and persistence of UDD, as it was usually associated with reduced serotonin function, which leads to urgency urinary incontinence [41]. Some studies argue that the medicines patients take potentially result in gastrointestinal side effects such as constipation [42]. In addition, participants with urinary and fecal incontinence and constipation exhibit significantly higher levels of depression and stress [43-49], possibly due to the production of certain gut microbiota that lead to depressive symptoms and lower levels of *g_Pseudoramibacter-Eubacterium* and *g_Candidatus-Solibacter*. [50] Functional constipation in middle-aged and older individuals may lead to a decrease in the abundance of these microbiota [51], ultimately exacerbating depression. This interaction creates a vicious cycle between depression and UDD.

In addition to depression, arthritis or rheumatism likely also contributes to the increased risk of incidence, recurrence, and persistence of UDD, although its impact on UDD might have been underestimated over the past 20 years. The most recent studies on the association between arthritis and UDD were conducted in the 1990s. These studies implied that arthritis may increase susceptibility to urinary tract infections, and atlantoaxial subluxation in the late stage of arthritis may lead to neurogenic bladder [52], which in turn increases the incidence of UDD. Medications such as misoprostol and cyclophosphamide, used to treat arthritis, may also cause UDD [52,53]. Moreover, opioids, commonly taken by patients with arthritis and rheumatic diseases, are likely to cause constipation. Additionally, the limited mobility of patients with osteoarthritis is perhaps associated with urinary incontinence [54,55].

XSL•FO
RenderX

Other comorbidities, such as hypertension and memory-related diseases, were also identified as potential contributors to the occurrence or recurrence of UDD. For instance, studies indicate a correlation between hypertension and constipation, potentially linked by shared physiological mechanisms and lifestyle factors [56]. Patients with Alzheimer disease often lose the ability to send bladder signals to the brain's urination center, resulting in an inability to urinate normally [57,58].

This study, therefore, suggested potential causal relationships between comorbidities and UDD, indicating that health care providers should pay more attention to the prevention and treatment of these comorbidities. Optimizing pharmaceutical therapy may effectively alleviate the incidence and progression of UDD in middle-aged and older individuals. Additionally, increasing physical therapy or dietary and nutritional treatments for mental health, arthritis or rheumatic diseases, and other chronic conditions could be beneficial. Dissemination of knowledge about these chronic diseases to enhance public awareness at the community level might help address the occurrence, recurrence, and persistence of UDD.

Our study also reveals a significant correlation between grip strength and the severity of UDD. Specifically, individuals with reduced grip strength exhibited a higher incidence of severe UDD manifestations, including recurrent and persistent conditions. This relationship may suggest that grip strength serves as an indirect biomarker for overall muscle integrity and functional capacity, which could influence the severity of UDD symptoms [59,60]. These results align with existing literature emphasizing the importance of muscle function in maintaining urogenital health and also underscore the need for further research into preventive strategies focusing on strength training in vulnerable populations.

Being Underweight or Obese Can Lead to Long-Term Persistence of UDD

This study found that almost half of the older Chinese population in the cohort had abnormal weight, with underweight or obesity considered independent risk factors for persistent UDD. This finding aligned with evidence from other studies showing that obesity is closely related to the occurrence of UDD [10,61-63]. The mechanism underlying this association is that metabolic changes and increased abdominal pressure caused by obesity may lead to the development of UDD. Additionally, underweight older adults might experience reduced muscle strength, particularly in the pelvic floor, possibly triggering the incidence of UDD. Some studies argue that population aging–related muscle atrophy and decreased strength are more likely to cause persistence of UDD [64,65], particularly among those who are underweight. Consequently, the improvement of nutritional intake is vital for older people in both communities and health care facilities. The awareness of healthy diet and appropriate physical exercises be enhanced among the older population, and health care providers and community health workers should develop individual tailored interventions for patients with UDD who are underweight or obese.

Strengths and Limitations

The study population was selected from 28 provinces in China, ensuring sufficient representativeness to reflect the characteristics and risk factors of UDD in the middle-aged and older Chinese populations. To achieve the second study objective, a cohort study was designed to establish clear causal relationships between the identified risk factors and the incidence, progression, and persistence of UDD in the Chinese population. Additionally, the population with UDD was classified based on the severity of UDD, facilitating the exploration of risk factors and the development of personalized interventions.

There are some limitations to this study. The CHARLS does not provide detailed information on urinary and fecal incontinence or constipation, making it difficult to distinguish risk factors for diverse UDD conditions. The CHARLS also lacks information on factors that may affect UDD, such as bladder outlet obstruction, urinary tract infections, benign prostatic hyperplasia, use of analgesics, and history of abuse, and factors such as hospital capacity, network scale, and accessibility of medical services over the past decade could not be analyzed due to the same limitations. Additionally, due to the statistical methods used in the CHARLS database, the impact of incontinence related to childbirth in female population may have been underestimated. Consequently, this study could not assess the impact of these factors on UDD. Furthermore, those who had UDD in both 2018 and 2020 could not have their UDD prognosis determined after the CHARLS survey: they may have recovered and not be included in the most-severe group, which may have resulted in misclassification bias. In addition, the emergence of recovered participants may also be related to respondents' misjudgment of their own UDD conditions. In the Results section, middle school and high school education emerged as a protective factor for the recovered population, though the underlying reasons require further investigation.

Conclusions

In summary, the prevalence of UDD increased with age and was found to be higher among illiterate individuals, married people, and those living in the Southwest and Northwest regions. Depression, arthritis and rheumatic diseases, and other chronic comorbidities contributed to the occurrence, recurrence, and persistence of UDD; further, being underweight or obese independently affected the persistence of UDD among the middle-aged and older Chinese population. Enhancing the treatment of psychological and chronic diseases and improving BMI may alleviate the occurrence, recurrence, and persistence of UDD.

Acknowledgments

This work was supported by the National Key R&D Program of China under the project "Development of Enhanced Training and Nursing Care Robot Systems for Bedridden Elderly with Urinary and Defecatory Dysfunction" (2022YFB4703300). The

XSL•FO
RenderX

authors would like to thank Fan Xuanxuan as well as Xu Ruoyu, Yang Muqi, and Yuan Qi from the School of Nursing, Capital Medical University, for their assistance in data processing and analysis (listed in no particular order).

Data Availability

The data used in this study are publicly available from the China Health and Retirement Longitudinal Study.

Authors' Contributions

HZ designed the study and drafted the manuscript. LJ supervised the study and approved the final version. J Zhang interpreted the data and reviewed the manuscript. GL and XL helped analyze the data. JK contributed to the conceptual framework and provided resources. LG administered the project and funding. RW contributed to methodological expertise and data curation. J Zhao and CZ validated the results. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The detailed results of the Bayesian logistic regression model.
[DOCX File, 31 KB - aging_v8i1e70541_app1.docx]

References

1. Ferrucci L, Gonzalez-Freire M, Fabbri E, et al. Measuring biological aging in humans: a quest. Aging Cell 2020 Feb;19(2):e13080. [doi: 10.1111/acel.13080] [Medline: 31833194]
2. Wang H. Prioritizing the health of the elderly population in China to promote the national strategy for healthy aging. Chin J Epidemiol 2019;40:259-265.
3. Ness W. Faecal incontinence: causes, assessment and management. Nurs Stand 2012;26(42):52-54. [doi: 10.7748/ns2012.06.26.42.52.c9162] [Medline: 22908765]
4. Coyne KS, Sexton CC, Bell JA, et al. The prevalence of lower urinary tract symptoms (LUTS) and overactive bladder (OAB) by racial/ethnic group and age: results from OAB-POLL. Neurourol Urodyn 2013 Mar;32(3):230-237. [doi: 10.1002/nau.22295] [Medline: 22847394]
5. Goepel M, Hoffmann JA, Piro M, Rübben H, Michel MC. Prevalence and physician awareness of symptoms of urinary bladder dysfunction. Eur Urol 2002 Mar;41(3):234-239. [doi: 10.1016/s0302-2838(02)00014-3] [Medline: 12180221]
6. Markland AD, Goode PS, Redden DT, Borrud LG, Burgio KL. Prevalence of urinary incontinence in men: results from the national health and nutrition examination survey. J Urol 2010 Sep;184(3):1022-1027. [doi: 10.1016/j.juro.2010.05.025] [Medline: 20643440]
7. Whitehead WE, Borrud L, Goode PS, et al. Fecal incontinence in US adults: epidemiology and risk factors. Gastroenterology 2009 Aug;137(2):512-517. [doi: 10.1053/j.gastro.2009.04.054] [Medline: 19410574]
8. Integrated care for older people: guidelines on community-level interventions to manage declines in intrinsic capacity. World Health Organization. 2017. URL: https://www.who.int/publications/i/item/9789241550109 [accessed 2025-05-15]
9. Elder JS, Diaz M. Vesicoureteral reflux—the role of bladder and bowel dysfunction. Nat Rev Urol 2013 Nov;10(11):640-648. [doi: 10.1038/nrurol.2013.221] [Medline: 24126731]
10. Batmani S, Jalali R, Mohammadi M, Bokaee S. Prevalence and factors related to urinary incontinence in older adults women worldwide: a comprehensive systematic review and meta-analysis of observational studies. BMC Geriatr 2021 Mar 29;21(1):212. [doi: 10.1186/s12877-021-02135-8] [Medline: 33781236]
11. Salari N, Ghasemianrad M, Ammari-Allahyari M, Rasoulpoor S, Shohaimi S, Mohammadi M. Global prevalence of constipation in older adults: a systematic review and meta-analysis. Wien Klin Wochenschr 2023 Aug;135(15-16):389-398. [doi: 10.1007/s00508-023-02156-w] [Medline: 36826591]
12. Mack I, Hahn H, Gödel C, Enck P, Bharucha AE. Global prevalence of fecal incontinence in community-dwelling adults: a systematic review and meta-analysis. Clin Gastroenterol Hepatol 2024 Apr;22(4):712-731. [doi: 10.1016/j.cgh.2023.09.004] [Medline: 37734583]
13. Vaughan CP, Markland AD. Urinary incontinence in women. Ann Intern Med 2020 Feb 4;172(3):ITC17-ITC32. [doi: 10.7326/AITC202002040] [Medline: 32016335]
14. Milsom I, Coyne KS, Nicholson S, Kvasz M, Chen CI, Wein AJ. Global prevalence and economic burden of urgency urinary incontinence: a systematic review. Eur Urol 2014 Jan;65(1):79-95. [doi: 10.1016/j.eururo.2013.08.031] [Medline: 24007713]
15. Miner PB Jr. Economic and personal impact of fecal and urinary incontinence. Gastroenterology 2004 Jan;126(1 Suppl 1):S8-13. [doi: 10.1053/j.gastro.2003.10.056] [Medline: 14978633]

XSL•FO
RenderX

16. Frigerio M, Barba M, Cola A, et al. Quality of life, psychological wellbeing, and sexuality in women with urinary incontinence—where are we now: a narrative review. Med Bogota Colomb 2022;58(4):525. [doi: 10.3390/medicina58040525]
17. Paquette IM, Varma MG, Kaiser AM, Steele SR, Rafferty JF. The American Society of Colon and Rectal Surgeons' clinical practice guideline for the treatment of fecal incontinence. Dis Colon Rectum 2015;58(7):623-636. [doi: 10.1097/DCR.0000000000000397]
18. Drake MJ, Williams J, Bijos DA. Voiding dysfunction due to detrusor underactivity: an overview. Nat Rev Urol 2014 Aug;11(8):454-464. [doi: 10.1038/nrurol.2014.156] [Medline: 25002201]
19. Chaichanavichkij P, Hartmann M, Scott SM, Fenton N, Knowles CH, Disorders of Defaecation Delphi Group. Evaluating the risk factors for the development of benign disorders of defaecation: a surgical perspective. Tech Coloproctol 2023 Oct;27(10):847-857. [doi: 10.1007/s10151-023-02843-w] [Medline: 37498418]
20. Heitmann PT, Vollebregt PF, Knowles CH, Lunniss PJ, Dinning PG, Scott SM. Understanding the physiology of human defaecation and disorders of continence and evacuation. Nat Rev Gastroenterol Hepatol 2021 Nov;18(11):751-769. [doi: 10.1038/s41575-021-00487-5] [Medline: 34373626]
21. Bag Soytas R, Soytas M, Danacioglu YO, et al. Relationship between the types of urinary incontinence, handgrip strength, and pelvic floor muscle strength in adult women. Neurourol Urodyn 2021 Aug;40(6):1532-1538. [doi: 10.1002/nau.24699] [Medline: 34004030]
22. Barberio B, Judge C, Savarino EV, Ford AC. Global prevalence of functional constipation according to the Rome criteria: a systematic review and meta-analysis. Lancet Gastroenterol Hepatol 2021 Aug;6(8):638-648. [doi: 10.1016/S2468-1253(21)00111-4] [Medline: 34090581]
23. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). Int J Epidemiol 2014 Feb;43(1):61-68. [doi: 10.1093/ije/dys203] [Medline: 23243115]
24. Zhao Y, Strauss J, Yang G, et al. 2011–2012 national baseline users' guide. China Health and Retirement Longitudinal Study. 2013. URL: https://charls.pku.edu.cn/wd/a2011nqgjxdc.htm [accessed 2025-05-15]
25. Ageing and health. World Health Organization. 2022. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2024-06-18]
26. International Classification of Functioning, Disability and Health: ICF. World Health Organization. 2001. URL: https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health [accessed 2024-06-24]
27. Mahfouz IA, Blanker MH, Asali F, et al. Seeking consultation for urinary incontinence: behaviours and barriers among Jordanian women. Neurourol Urodyn 2023 Aug;42(6):1299-1310. [doi: 10.1002/nau.25189] [Medline: 37186396]
28. Jarbøl DE, Haastrup PF, Rasmussen S, Søndergaaard J, Balasubramaniam K. Women's barriers for contacting their general practitioner when bothered by urinary incontinence: a population-based cross-sectional study. BMC Urol 2021 Jul 12;21(1):99. [doi: 10.1186/s12894-021-00864-x] [Medline: 34247613]
29. Walden AL, Salsbury SA, Reed WR, Lawrence DJ. Bladder and bowel symptoms among adults presenting with low back pain to an academic chiropractic clinic: results of a preliminary study. J Chiropr Med 2014 Sep;13(3):178-187. [doi: 10.1016/j.jcm.2014.07.006] [Medline: 25225466]
30. Yan F, Xiao LD, Zhou K, Li Z, Tang S. Perceptions and help-seeking behaviours among community-dwelling older people with urinary incontinence: a systematic integrative review. J Adv Nurs 2022 Jun;78(6):1574-1587. [doi: 10.1111/jan.15183] [Medline: 35150161]
31. Sheng Y, Carpenter JS, Ashton-Miller JA, Miller JM. Mechanisms of pelvic floor muscle training for managing urinary incontinence in women: a scoping review. BMC Womens Health 2022 May 13;22(1):161. [doi: 10.1186/s12905-022-01742-w] [Medline: 35562699]
32. Curillo-Aguirre CA, Gea-Izquierdo E. Effectiveness of pelvic floor muscle training on quality of life in women with urinary incontinence: a systematic review and meta-analysis. Medicina (Kaunas) 2023 May 23;59(6):1004. [doi: 10.3390/medicina59061004] [Medline: 37374208]
33. Harris LA, Chang CH. Burden of constipation: looking beyond bowel movements. Am J Gastroenterol 2022 Apr 1;117(4S):S2-S5. [doi: 10.14309/ajg.0000000000001708] [Medline: 35354769]
34. Barakat C, Konstantinidis T. A review of the relationship between socioeconomic status change and health. Int J Environ Res Public Health 2023 Jun 29;20(13):6249. [doi: 10.3390/ijerph20136249] [Medline: 37444097]
35. McMaughan DJ, Oloruntoba O, Smith ML. Socioeconomic status and access to healthcare: interrelated drivers for healthy aging. Front Public Health 2020;8:231. [doi: 10.3389/fpubh.2020.00231] [Medline: 32626678]
36. Tamanini JTN, Franceschi Júnior O, Santos JLF, et al. Fecal incontinence: incidence and risk factors from the SABE (Health, Wellbeing and Aging) study. Int Urogynecol J 2022 Nov;33(11):2993-3004. [doi: 10.1007/s00192-021-04914-8] [Medline: 35015091]
37. Bauer SR, McCulloch CE, Cawthon PM, et al. Longitudinal associations between concurrent changes in phenotypic frailty and lower urinary tract symptoms among older men. The Journal of Frailty & Aging 2023 Apr;12(2):117-125. [doi: 10.14283/jfa.2022.33]
38. Liu X, Zhao D, Zhao S, Li Z, Wang Y, Qin X. Deciphering the correlations between aging and constipation by metabolomics and network pharmacology. Aging (Albany NY) 2021 Jan 10;13(3):3798-3818. [doi: 10.18632/aging.202340] [Medline: 33428599]

XSL•FO
RenderX

39. Veronese N, Soysal P, Stubbs B, et al. Association between urinary incontinence and frailty: a systematic review and meta-analysis. Eur Geriatr Med 2018 Oct;9(5):571-578. [doi: 10.1007/s41999-018-0102-y] [Medline: 34654231]
40. Wang R, Liu G, Jing L, Zhang J, Ye Y, Zhu H. Quantifying the effects of five rehabilitation training methods on the ability of elderly men to control bowel movements: a finite element analysis study. Front Bioeng Biotechnol 2024;12:1392448. [doi: 10.3389/fbioe.2024.1392448] [Medline: 38988865]
41. Zorn BH, Montgomery H, Pieper K, Gray M, Steers WD. Urinary incontinence and depression. J Urol 1999 Jul;162(1):82-84. [doi: 10.1097/00005392-199907000-00020] [Medline: 10379745]
42. Oliva V, Lippi M, Paci R, et al. Gastrointestinal side effects associated with antidepressant treatments in patients with major depressive disorder: a systematic review and meta-analysis. Prog Neuropsychopharmacol Biol Psychiatry 2021 Jul 13;109:110266. [doi: 10.1016/j.pnpbp.2021.110266] [Medline: 33549697]
43. Lee HY, Rhee Y, Choi KS. Urinary incontinence and the association with depression, stress, and self-esteem in older Korean Women. Sci Rep 2021 Apr 27;11(1):9054. [doi: 10.1038/s41598-021-88740-4] [Medline: 33907278]
44. Wu S, Wu F. Association of urinary incontinence with depression among men: a cross-sectional study. BMC Public Health 2023;23(1):944. [doi: 10.1186/s12889-023-15961-9]
45. Wang Y, Li N, Zhou Q, Wang P. Fecal incontinence was associated with depression of any severity: insights from a large cross-sectional study. Int J Colorectal Dis 2023 Nov 22;38(1):271. [doi: 10.1007/s00384-023-04563-x] [Medline: 37991579]
46. Gevelinger MM, Jaworski EM, Speranza JR, Sanderson DJ. Improvements in self-reported depression following treatment of fecal incontinence with sacral neuromodulation. Neuromodulation 2020 Dec;23(8):1158-1163. [doi: 10.1111/ner.13249]
47. Whiting D, Shane AI, Pope R, Payne S, Venn S. Female urinary incontinence in sub-Saharan Africa. BJU Int 2022 Nov;130(5):543-549. [doi: 10.1111/bju.15903] [Medline: 36161452]
48. Ballou S, Katon J, Singh P, et al. Chronic diarrhea and constipation are more common in depressed individuals. Clin Gastroenterol Hepatol 2019 Dec;17(13):2696-2703. [doi: 10.1016/j.cgh.2019.03.046] [Medline: 30954714]
49. Wang Y, Shen X, Wang P. Constipation is associated with depression of any severity, but not with suicidal ideation: insights from a large cross-sectional study. Int J Colorectal Dis 2023;38(1):231. [doi: 10.1007/s00384-023-04520-8]
50. Liang J, Zhao Y, Xi Y, et al. Association between depression, anxiety symptoms and gut microbiota in Chinese elderly with functional constipation. Nutrients 2022 Nov 25;14(23):5013. [doi: 10.3390/nu14235013] [Medline: 36501044]
51. Costa LM, Mendes MM, Oliveira AC, et al. Dietary inflammatory index and its relationship with gut microbiota in individuals with intestinal constipation: a cross-sectional study. Eur J Nutr 2022 Feb;61(1):341-355. [doi: 10.1007/s00394-021-02649-2] [Medline: 34351455]
52. Turner-Stokes L, Frank AO. Urinary incontinence among patients with arthritis—a neglected disability. J R Soc Med 1992 Jul;85(7):389-393. [Medline: 1629846]
53. Fossaluzza V, Di Benedetto P, Zampa A, De Vita S. Misoprostol-induced urinary incontinence. J Intern Med 1991 Nov;230(5):463-464. [doi: 10.1111/j.1365-2796.1991.tb00474.x] [Medline: 1940783]
54. Mat S, Jaafar MH, Razack AHA, et al. Gender differences in the associations between knee pain and urinary incontinence in older adults: cross-sectional analysis from the Malaysian Elders Longitudinal Research Study (MELoR). Neurourol Urodyn 2023 Mar;42(3):641-649. [doi: 10.1002/nau.25136] [Medline: 36728321]
55. de Melo G, Fernandes VDO, Costa HS, et al. Urinary incontinence in middle-aged and older women with hip and knee osteoarthritis: an outpatient study of frequency and risk factors. Neurourol Urodyn 2024 Mar;43(3):672-679. [doi: 10.1002/nau.25395] [Medline: 38247352]
56. Judkins CP, Wang Y, Jelinic M, et al. Association of constipation with increased risk of hypertension and cardiovascular events in elderly Australian patients. Sci Rep 2023;13(1):10943. [doi: 10.1038/s41598-023-38068-y]
57. Davis NJ, Wyman JF, Gubitosa S, Pretty L. Urinary incontinence in older adults. Am J Nurs 2020;120(1):57-62. [doi: 10.1097/01.NAJ.0000652124.58511.24]
58. Camilleri M. Gastrointestinal motility disorders in neurologic disease. J Clin Invest 2021 Feb 15;131(4):e143771. [doi: 10.1172/JCI143771] [Medline: 33586685]
59. Erdogan T, Bahat G, Kilic C, et al. The relationship between sarcopenia and urinary incontinence. Eur Geriatr Med 2019 Dec;10(6):923-929. [doi: 10.1007/s41999-019-00232-x]
60. Li L, Chen F, Li X, Gao Y, Wang N, Xu T. Association between low handgrip strength and incontinence among Chinese middle-aged and older people: a cross-sectional study. Aging Med (Milton) 2024 Jun;7(3):360-367. [doi: 10.1002/agm2.12309] [Medline: 38975299]
61. Doumouchtsis SK, Loganathan J, Pergialiotis V. The role of obesity on urinary incontinence and anal incontinence in women: a review. BJOG 2022 Jan;129(1):162-170. [doi: 10.1111/1471-0528.16848] [Medline: 34312978]
62. Subak LL, Richter HE, Hunskaar S. Obesity and urinary incontinence: epidemiology and clinical research update. J Urol 2009 Dec;182(6 Suppl):S2-S7. [doi: 10.1016/j.juro.2009.08.071] [Medline: 19846133]
63. Zhu F, Chen M, Xiao Y, Huang X, Chen L, Hong L. Synergistic interaction between hyperlipidemia and obesity as a risk factor for stress urinary incontinence in Americans. Sci Rep 2024 Mar 27;14(1):7312. [doi: 10.1038/s41598-024-56744-5] [Medline: 38538650]

XSL•FO
RenderX

64. Suskind AM, Cawthon PM, Nakagawa S, et al. Urinary incontinence in older women: the role of body composition and muscle strength: from the health, aging, and body composition study. J Am Geriatr Soc 2017 Jan;65(1):42-50. [doi: 10.1111/jgs.14545] [Medline: 27918084]
65. Bauer SR, Grimes B, Suskind AM, Cawthon PM, Cummings S, Huang AJ. Urinary incontinence and nocturia in older men: associations with body mass, composition and strength in the health ABC study. Journal of Urology 2019 Nov;202(5):1015-1021. [doi: 10.1097/JU.0000000000000378]

Abbreviations

CHARLS: China Health and Retirement Longitudinal Study
OR: odds ratio
UDD: urinary and defecatory dysfunction

Edited by A Garcia; submitted 24.12.24; peer-reviewed by D Shaker, KH Goh; revised version received 30.03.25; accepted 17.04.25; published 26.05.25.

Please cite as:
Zhu H, Li X, Jing L, Kou J, Gong L, Wang R, Liu G, Zhang C, Zhao J, Zhang J
Factors Influencing the Severity of Urinary and Defecatory Dysfunction Among the Middle-Aged and Older Adult Chinese Population: Longitudinal Study of a 5-Wave Survey Cohort
JMIR Aging 2025;8:e70541
URL: https://aging.jmir.org/2025/1/e70541
doi:10.2196/70541

© Haoran Zhu, Xiaoming Li, Liwei Jing, Jingli Kou, Lichao Gong, Rui Wang, Guangtian Liu, Chao Zhang, Jiayi Zhao, Jing Zhang. Originally published in JMIR Aging (https://aging.jmir.org), 26.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Role of Physical Activity and Physical Function in Predicting Physical Frailty Transitions in Chinese Older Adults: Longitudinal Observational Study From CHARLS

Ziwei Zeng[1], MSc; Chun Liang Hsu[2], PhD; Cindy Hui-ping Sit[1], PhD; Stephen Heung-sang Wong[1], PhD; Yijian Yang[1,3], MD, PhD

[1]Department of Sports Science and Physical Education, The Chinese University of Hong Kong, Shatin, Hong Kong, China (Hong Kong)

[2]Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[3]The Jockey Club Institute of Ageing, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

Corresponding Author:
Yijian Yang, MD, PhD
Department of Sports Science and Physical Education, The Chinese University of Hong Kong, Shatin, Hong Kong, China (Hong Kong)

Abstract

Background: Frailty is a dynamic geriatric syndrome associated with adverse health outcomes, yet its progression can be mitigated through targeted interventions.

Objective: This study aimed to investigate predictors of frailty transitions in Chinese older adults, focusing on physical activity (PA) and physical function.

Methods: Using data from the China Health and Retirement Longitudinal Study (CHARLS), we examined transitions between frailty states (robust, prefrail, and frail) from 2011 (baseline) to 2013 (follow-up) among 1014 participants aged 65 years and older. The following outcomes were assessed, including frailty using the physical frailty phenotype, PA using a modified International Physical Activity Questionnaire, and physical function using the Short Physical Performance Battery (SPPB) and handgrip strength. Ordinal logistic regression models were used to examine the relationship between PA, physical function, and frailty transitions.

Results: Results showed that higher PA levels and better physical function reduced the likelihood of worsening frailty or increased the probability of transitioning to robustness. Key findings from the subgroup include: among robust individuals, greater handgrip strength predicted maintained robustness (average marginal effects [AME]=1.12%; *P*=.02); in prefrail individuals, higher vigorous PA (AME=21.76%; *P*=.04) and handgrip strength (AME=0.64%; *P*=.003) increased transitions to robustness; for frail individuals, increased low-intensity PA (AME =22.48%; *P*=.04) and higher SPPB walking subscores (AME=27.73%; *P*=.02) promoted improvement to nonfrailty.

Conclusions: These findings highlight the importance of tailored interventions based on baseline frailty status. Promoting PA and improving physical function, particularly muscle strength and mobility function, may help delay or reverse frailty progression.

(JMIR Aging 2025;8:e75887) doi:10.2196/75887

KEYWORDS

elderly; frailty trajectories; physical performance; strength; mobility

Introduction

The global population is aging at an unprecedented rate, presenting significant challenges to public health and social care systems worldwide [1]. By 2050, the proportion of adults aged 65 years and older in China is projected to reach 26%, reflecting a broader demographic shift that underscores the urgency of addressing age-related health issues such as physical frailty [2]. Frailty, a geriatric syndrome characterized by diminished physiological reserves and increased vulnerability to stressors, is associated with adverse outcomes including falls, disability, hospitalization, and mortality [3,4]. Defined by Fried et al [3] as the presence of 3 or more criteria: weakness, slow gait speed, exhaustion, low activity, and unintentional weight loss, physical frailty is a dynamic condition that is not an inevitable consequence of aging but rather a potentially reversible state [3,5]. Rates of physical frailty among Chinese older adults range from 6% to 17% [6], highlighting the importance of early detection and intervention to mitigate its progression.

Frailty is typically classified into 3 stages: robust, prefrail, and frail, with individuals transitioning between these stages over

XSL•FO
RenderX

time due to a combination of physical, psychological, and environmental factors [7]. Understanding these transitions is crucial for developing targeted interventions to prevent or delay frailty progression. Physical function measures, such as handgrip strength and the Short Physical Performance Battery (SPPB), are increasingly recognized as critical indicators of frailty [8-11]. Handgrip strength, for example, provides insight into muscular strength, functional limitations, and the onset of age-related adverse health conditions, making it a more useful marker of frailty than chronological age alone [8,9]. Similarly, the SPPB, which assesses gait speed, sit-to-stand (STS) performance, and balance, has been widely used to evaluate physical function and predict frailty progression [10-12]. These measures offer a comprehensive understanding of the factors that influence frailty dynamics and complement the role of physical activity (PA) in frailty transitions [13].

PA has long been recognized as a cornerstone of healthy aging, encompassing any bodily movement produced by skeletal muscles that increases energy expenditure [14]. Beyond its general health benefits, PA has emerged as a key modifiable factor in frailty prevention and management [15]. Regular PA has been shown to improve muscle strength, cardiovascular health, and overall physical function, all of which are critical for reducing the risk of frailty [16-22]. However, the specific types and intensities of PA that influence frailty transitions remain unclear, highlighting the need for further investigation.

In China, the rapid aging population presents challenges in addressing the burden of frailty and functional decline [23]. Despite the growing recognition of frailty as a public health issue, few studies have examined the specific factors that predict frailty transitions among Chinese older adults. Existing research has primarily focused on Western populations [24,25], and there is limited evidence on how PA and physical function influence frailty transitions in the Chinese context. This gap underscores the need for population-specific research, as contextual factors (eg, lifestyle and health care infrastructure [26]) in China may differ from Western settings. Furthermore, longitudinal follow-up is critical for understanding frailty as a dynamic process rather than a static condition. By observing the same individuals over time, researchers can quantify both progression and recovery across frailty states and relate these transitions to changes in PA and physical function. This design overcomes the limitations of cross-sectional studies where reverse causation and unmeasured baseline differences may obscure causal pathways and provides more robust evidence to inform stage-specific interventions. Understanding these relationships is essential for developing tailored strategies to mitigate frailty in Chinese older adults.

This study aimed to investigate the predictors of physical frailty transitions in Chinese older adults, with a particular focus on the role of PA participation and physical function. We hypothesized that higher amounts of light, moderate, and vigorous PA, along with greater muscle strength and mobility, would increase the likelihood of transitioning to a more robust status. Conversely, lower amounts of PA across intensity categories and poor physical function would increase the probability of worsening frailty. By examining these relationships, this study provided valuable insights into the role of PA and physical function in physical frailty transitions, with the ultimate goal of informing intervention strategies to delay or prevent frailty progression.

Methods

Study Design

This was a 2-wave longitudinal observational study using data from the China Health and Retirement Longitudinal Study (CHARLS). We leveraged the baseline survey (Wave 1, 2011) and first follow-up (Wave 2, 2013) to examine within-person transitions among 3 frailty states (robust, prefrail, and frail), and to evaluate how baseline PA and physical function predicted these transitions.

Ethical Considerations

Data were obtained from CHARLS, a nationally representative survey designed to examine the aging process among Chinese households and individuals aged 45 years and older [27]. The sampling design used a 4-stage, stratified, cluster method, covering 150 county-level and 450 village-level units across China [28]. Ethical approval was granted by the ethical review committee of Peking University (approval IRB00001052-11015), and all participants provided written informed consent. CHARLS adheres to the declaration of Helsinki and China's Personal Information Protection Law. The CHARLS database ensures strict privacy protection and anonymization during data collection and processing to safeguard participants' personal information.

Study Participants

The CHARLS survey collected comprehensive data on demographics, anthropometrics, health status, functioning, and physical measurements [27]. This study used data from the first 2 waves of CHARLS (2011 and 2013), which share identical sampling procedures and therefore draw from the same population base. Although frailty is typically a slow-progressing condition, previous longitudinal research suggests that meaningful changes, both improvements and declines, can be captured within a 2-year period among older adults [29,30]. Thus, the Wave 1 (2011) and Wave 2 (2013) data are appropriate for examining frailty transitions in this context.

While the Law of the People's Republic of China on the Protection of the Rights and Interests of the Elderly (1996) defines older adults as individuals aged 60 years and older, we adopted the United Nations' recommendation of ages 65 years and older. This definition aligns with international standards in frailty research and better targets the population at higher risk for physical frailty in China [31].

For this study, we first restricted the sample to respondents who completed both Wave 1 and Wave 2 data collection. From this sample, we included participants who met the following criteria: (1) participants aged 65 years and older at baseline (2011); (2) they provided sufficient data on physical frailty (at least 3 components), PA, and physical function (at least 1 component of the SPPB) across both waves; and (3) had consistent and valid data for gender, body height, and body weight across both

XSL•FO
RenderX

waves (eg, mismatched gender data between waves were excluded).

Measures

Physical Frailty

Physical frailty was assessed using the Physical Frailty Phenotype (PFP) scale, developed from the Cardiovascular Health Study (CHS) [3]. The scale comprises five criteria: (1) weakness, (2) slowness, (3) exhaustion, (4) low activity, and (5) shrinking. The PFP scale was adapted and validated for use in the CHARLS to examine physical frailty in older Chinese adults [32-35]. Each criterion was scored as 1 point if met, with a total score ranging from 0 to 5. Participants were categorized as "Robust" (0 criteria met), "Pre-frail" (1 - 2 criteria met), or "Frail" (3 - 5 criteria met) [36]. Detailed definitions and scoring for each criterion are provided in Table 1.

Table . Operationalization of the Physical Frailty Phenotype in the China Health and Retirement Longitudinal Study (CHARLS) compared to the Cardiovascular Health Study (CHS).

PFP[b] component	CHS[c] definition	CHARLS[a] adaptation		
		Definition	Values	
			Women	Men
Weakness	Handgrip strength: lowest 20% (by gender and BMI).	Handgrip strength	• BMI ≤20.6kg/m^2: ≤18.0kg • 20.6<BMI≤23.1kg/m^2: ≤18.2kg • 23.1<BMI ≤25.7kg/m^2: ≤20.0kg • BMI >25.7kg/m^2: ≤20.0kg	• BMI ≤20.0kg/m^2: ≤27.0kg • 20.0<BMI≤22.0kg/m^2: ≤28.5kg • 22.0<BMI ≤24.4kg/m^2: ≤30.0kg • BMI >24.4kg/m^2: ≤31.5kg
Slowness	Walking time/15 feet: slowest 20% (by gender and height).	Walking speed (2.5 meters)	• Body height ≤151cm: ≤0.41m/s • Body height >151cm: ≤0.45m/s	• Body height ≤162cm: ≤0.47m/s • Body height >162cm: ≤0.50m/s
Exhaustion	"Exhaustion" (self-report).	Self-reported	• Felt "I felt everything I did was an effort" or "I could not get going" about 3 - 4 days or 5 - 7 days in a week.	
Low activity	Kcals/week: lowest 20%: men:<383 Kcals/week; women:<270 Kcals/week.	Self-reported	• Answered "no" to: "During a usual week, did you do any vigorous activities/moderate physical effort/walking for at least ten minutes continuously."	
Shrinking	Baseline: >10 lbs lost unintentionally in the previous year.	Self-reported	• Lost 5 or more kilograms in the last year in Wave 1 (2011). • Weight declines 5 or more kilograms between Wave 1 (2011) and Wave 2 (2013).	

[a]CHARLS China Health and Retirement Longitudinal Study.

[b]PFP Physical Frailty Phenotype.

[c]CHS Cardiovascular Health Study.

Initially, mortality was considered a key transition outcome, but upon further review, it was found that participants who had died by Wave 2 did not provide sufficient baseline data for inclusion. Therefore, mortality was not considered as a transition outcome in this study. Instead, we focused on the 3 states of physical frailty (robust, prefrail, and frail). Frailty transitions were defined as movements between these three states over time, categorized as follows: remain robust, robust to prefrail, robust to frail, prefrail to robust, remain prefrail, prefrail to frail, frail to robust, frail to prefrail, and remain frail.

Physical Activity

PA was assessed using a modified version of the International Physical Activity Questionnaire (IPAQ), a widely used tool for PA assessment [27]. Differences between CHARLS and the original IPAQ included: (1) PA was assessed over a "usual week" rather than the "last 7 days," (2) sedentary behavior was not recorded, and (3) PA was reported as discrete time categories rather than continuous data [37]. PA data included intensity (vigorous, moderate, and low-intensity), duration (<30 min, 30 min–2 h, 2 - 4 h, and ≥4 h), and frequency (1 - 7 d/wk). PA volume was calculated using metabolic equivalents (METs): vigorous-intensity=8.0 METs, moderate-intensity=4.0 METs, and low-intensity (walking)=3.3 METs [37,38]. Vigorous PA MET-minutes/week=8.0×vigorous-intensity activity minutes×vigorous-intensity days; Moderate PA MET-minutes/week=4.0×moderate-intensity activity minutes × moderate-intensity days; and low-intensity PA MET-minutes/week=3.3×walking minutes×walking days [37,38]. Total PA was categorized into four types based on

XSL•FO
RenderX

MET-minutes/week for each intensity level: (1) null type=0, (2) type 1: <600, (3) type 2: 600 - 3000, and (4) type 3: >3000, following the IPAQ scoring protocol and previously applied in similar Chinese cohorts [37,39].

Physical Function

Physical function was assessed using the SPPB, which includes tasks measuring gait speed, STS test time, and standing balance. Implementation details and scoring for the Chinese population were described by Zhong et al [40]. Each test is scored from 0 to 4, with higher scores indicating better performance. The total SPPB score (range 0 - 12) was classified into three groups: 0 - 6 (poor performers), 7 - 9 (fair performers), and 10 - 12 (good performers) [41].

Handgrip strength was assessed using a handheld dynamometer (Yuejian WL-1000; Nantong Yuejian Physical Measurement Instrument Co Ltd). Participants were instructed to stand upright, maintain a 90° elbow flexion, and perform maximal grip exertion for 3 - 5 seconds. A total of 2 trials were conducted on each hand, and the highest value (kg) across all 4 trials was used for analysis [27].

Covariates

Covariates were selected based on their relevance to physical frailty, PA, and physical function. These included age, gender (man=0 and woman=1), BMI, marital status (unmarried=0 and married=1), education level (illiterate=1, no formal education=2, elementary school=3, and middle school or above=4), self-rated health (poor or fair=0 and good or very good or excellent=1), fall history (no=0 and yes=1), activity of daily living (ADL) limitation (without=0 and with=1), instrumental activity of daily living (IADL) limitation (without=0 and with=1), number of chronic diseases (0=0, 1=1, and ≥2=2), pain status (no=0 and yes=1), current smoking status (no=0 and yes=1), and current drinking status (no=0 and yes=1).

Statistical Analysis

Data analysis was conducted using Stata 15.0 (StataCorp LP), with statistical significance set at $P<.05$. Descriptive statistics (mean [SD] for continuous variables and n [%] for categorical variables) were used to summarize baseline characteristics. Multiple imputation by chained equations addressed missing data, with results combined across 20 imputations [42]. Differences across frailty groups at baseline were examined using the 1-way ANOVA test (normally distributed data) or the Kruskal-Wallis H test (non-normally distributed data) for continuous variables, and χ^2 tests for categorical variables. Post hoc pairwise comparisons were performed with Bonferroni corrections as appropriate.

Ordinal logistic regression was used to model the relationship between predictors (PA and physical function) and the ordered outcome of frailty transitions, adjusting for demographic and health-related covariates [43]. Average marginal effects (AMEs) were calculated to assess the impact of a 1-unit increase in a predictor on the probability of frailty transition state, when holding other variables constant. To assess potential circularity from including handgrip strength (one frailty criterion) as a predictor, we conducted a sensitivity analysis omitting handgrip strength from the model. Robust standard errors were used to account for potential heteroskedasticity. Predictive accuracy was assessed using a classification table, and discriminate ability (remain frail or worsen to frailty vs others) was evaluated using receiver operating characteristic (ROC) curves and the area under the curve (AUC) was calculated along with corresponding 95% CIs and *P* values to assess statistical significance [44]. Separate ROC curves were generated for each baseline frailty status.

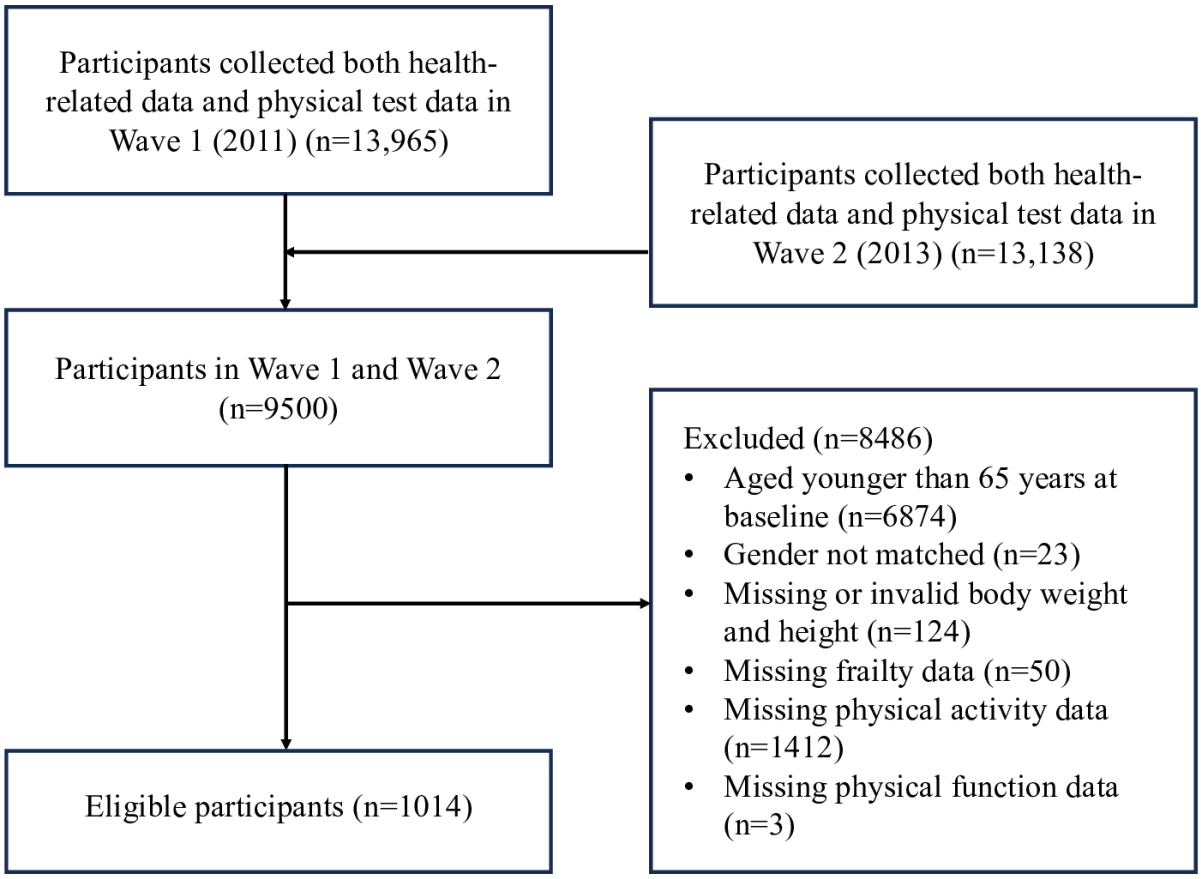
Results

Sample Characteristics and Comparisons

A total of 1014 participants were included (see Figure 1), with 371 categorized as robust, 561 as prefrail, and 82 as frail at baseline. Figure 2 shows the transition patterns from baseline to follow-up. Among robust participants, 39% (145/371) transitioned to prefrailty, and 1% (4/371) transitioned to frailty. Among prefrail participants, 7% (42/561) became frail, while 30% (169/561) reverted to a robust state. Among frail participants, 15% (12/82) regained robustness, and 57% (47/82) became prefrail.

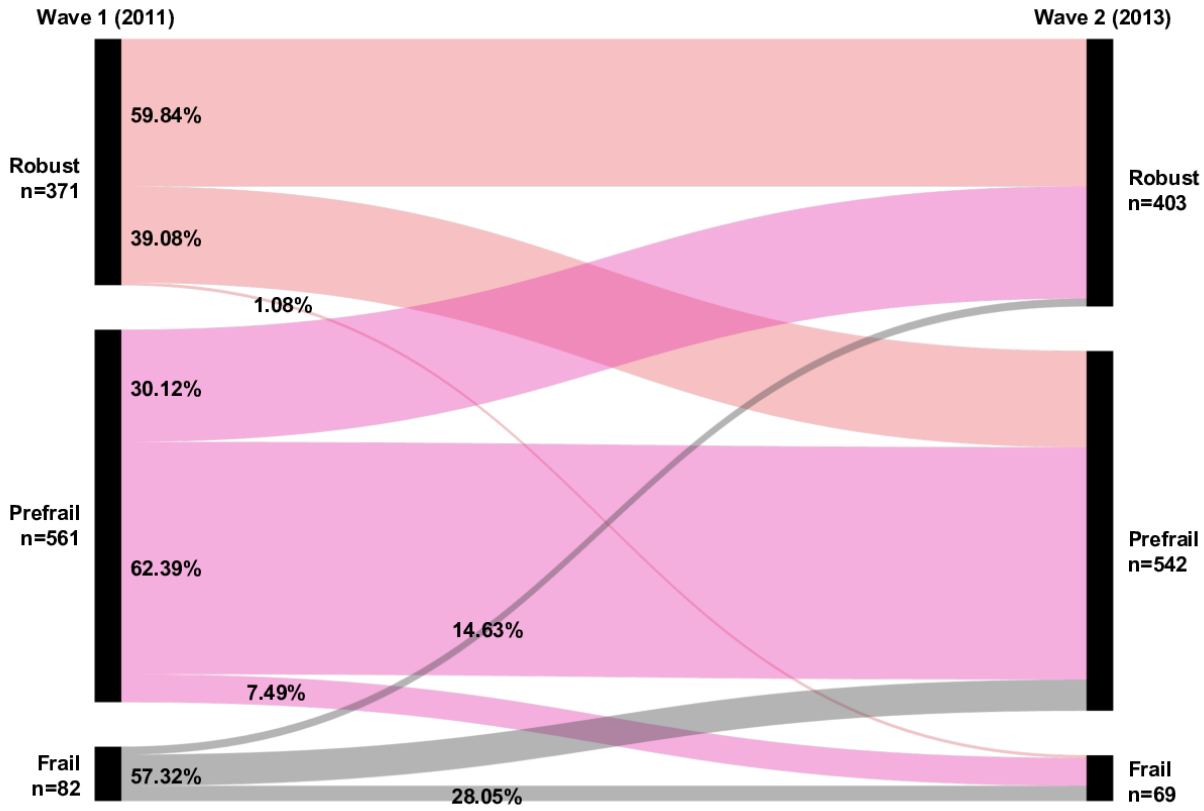
XSL•FO
RenderX

Figure 1. Flowchart of participant selection.



XSL•FO
RenderX

Figure 2. Sankey diagram of frailty status transition from Wave 1 (2011) to Wave 2 (2013).



The mean age of the study population was 71.1 years (SD 5.0, range 65 - 90 y), with significant variations across frailty groups (robust: 70.2 y [SD 4.4], prefrail: 71.4 y [SD 5.1], and frail: 73.0 y [SD 5.8]; $P<.001$). Gender distribution was balanced (48% women). Frailer participants were significantly older, had lower educational attachment, poorer self-rated health, higher fall history, and greater limitations in ADL and IADL, as well as more pain and chronic diseases compared to their healthier counterparts ($P\leq.01$). Detailed baseline characteristics are presented in Table 2.

XSL•FO
RenderX

Table . Demographic characteristics of participants by baseline frailty status.

Variables	Total sample (N=1014)	Robust (n=371)	Prefrail (n=561)	Frail (n=82)	*P* value
Age (years), n (%)					<.001[a]
65 - 74	775 (76.43)	312 (84.10)	414 (73.80)	49 (59.76)	
75 - 84	229 (22.58)	58 (15.63)	141 (25.13)	30 (36.59)	
≥85	10 (0.99)	1 (0.27)	6 (1.07)	3 (3.66)	
Gender, n (%)					.02[a]
Women	486 (47.93)	156 (42.05)	286 (50.98)	44 (53.66)	
Men	528 (52.07)	215 (57.95)	275 (49.02)	38 (46.34)	
BMI (kg/m^2), mean (SD)	22.37 (3.57)	22.65 (3.64)	22.14 (3.47)	22.66 (3.89)	.09
Marital status, n (%)					.91
Unmarried	253 (24.95)	91 (24.53)	140 (25.96)	22 (26.83)	
Married	761 (75.05)	280 (75.47)	421 (75.04)	60 (73.17)	
Education level, n (%)					.004[a]
Illiterate	392 (38.66)	123 (33.15)	226 (40.29)	43 (52.44)	
Primary and lower	201 (19.82)	69 (18.60)	119 (21.21)	13 (15.85)	
Secondary education	237 (23.37)	93 (25.07)	130 (23.17)	14 (17.07)	
College and above	184 (18.15)	86 (23.18)	86 (15.33)	12 (14.63)	
Self-rated health, n (%)					<.001[a]
Poor or fair	825 (81.36)	268 (72.24)	481 (85.74)	76 (92.68)	
Good or very good or excellent	189 (18.64)	103 (27.76)	80 (14.26)	6 (7.32)	
Fall history, n (%)					.01[a]
Yes	176 (17.36)	47 (12.67)	111 (19.79)	18 (21.95)	
No	838 (82.64)	324 (87.33)	450 (80.21)	64 (78.05)	
ADL[b] limitation, n (%)					<.001[a]
With	261 (25.74)	58 (15.63)	160 (28.52)	43 (52.44)	
Without	753 (74.26)	313 (84.37)	401 (71.48)	39 (47.56)	
IADL[c] limitation, n (%)					<.001[a]
With	291 (28.70)	60 (16.17)	178 (31.73)	53 (64.63)	
Without	723 (71.30)	311 (83.83)	383 (68.27)	29 (35.37)	
Number of chronic diseases, n (%)					.007[a]
0	270 (26.63)	118 (31.81)	141 (25.13)	11 (13.41)	
1	285 (28.11)	102 (27.49)	159 (28.34)	24 (29.27)	
≥2	459 (45.27)	151 (40.70)	261 (46.52)	47 (57.32)	
Pain, n (%)					<.001[a]
Yes	358 (35.31)	80 (21.56)	232 (41.35)	46 (56.10)	
No	656 (64.69)	291 (78.44)	329 (58.65)	36 (43.90)	
Current smoking, n (%)					.51
Yes	423 (41.72)	162 (43.67)	225 (40.11)	36 (43.90)	
No	591 (58.28)	209 (56.33)	336 (59.89)	46 (56.10)	

XSL•FO
RenderX

Variables	Total sample (N=1014)	Robust (n=371)	Prefrail (n=561)	Frail (n=82)	*P* value
Current drinking, n (%)					.02[a]
Yes	306 (30.18)	129 (34.77)	160 (28.52)	17 (20.73)	
No	708 (69.18)	242 (65.23)	401 (71.48)	65 (79.27)	
SPPB[d] total score (points), mean (SD)	8.92 (2.11)	9.84 (1.68)	8.66 (2.06)	6.55 (1.89)	<.001[a]
SPPB-Walk (points), mean (SD)	2.47 (0.99)	2.82 (0.85)	2.38 (1.00)	1.46 (0.72)	<.001[a]
SPPB-STS[e] (points), mean (SD)	2.98 (1.16)	3.37 (0.93)	2.86 (1.18)	2.01 (1.20)	<.001[a]
SPPB-Balance (points), mean (SD)	3.48 (0.81)	3.64 (0.68)	3.42 (0.84)	3.07 (0.91)	<.001[a]

[a]Indicates statistical significance (*P*<.05).

[b]ADL: activity of daily living.

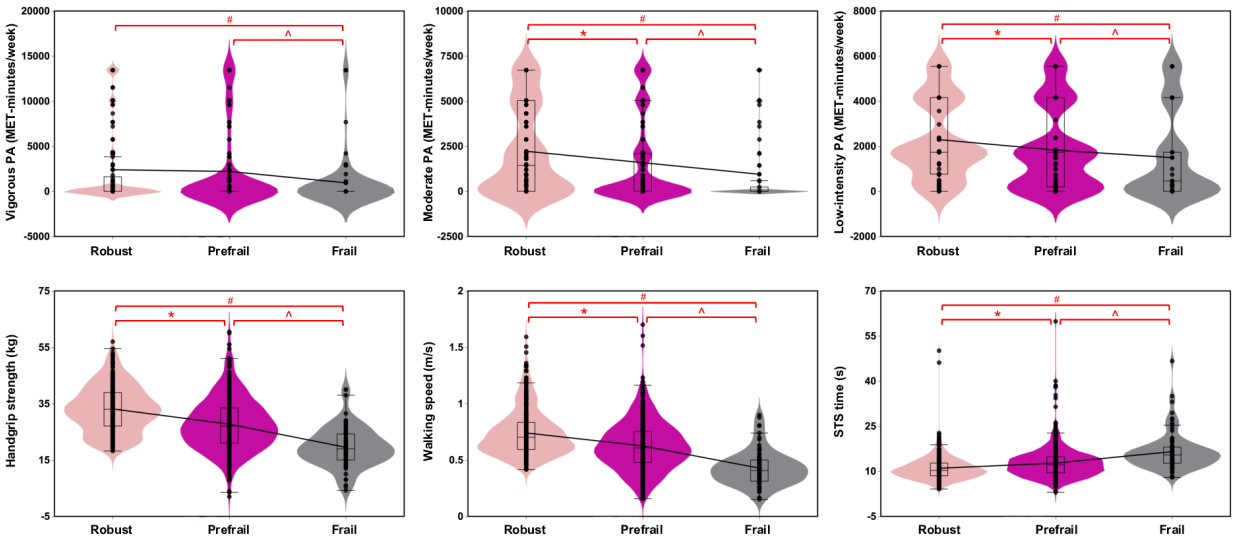
[c]IADL: instrumental activity of daily living.

[d]SPPB: Short Physical Performance Battery.

[e]STS: sit-to-stand.

Table 2 and Figure 3 show that robust participants engaged in significantly more moderate and low-intensity PA compared to pre-frail and frail individuals (*P*<.001), with vigorous PA significantly higher in robust individuals compared to frail (*P*=.003). Prefrail participants also demonstrated higher PA levels compared to frail individuals (*P*≤.02). Robust participants exhibited greater handgrip strength, faster walking speeds, shorter STS time, and higher SPPB scores compared to prefrail participants, and prefrail participants performed better than frail participants (*P*<.001)

Figure 3. Physical activity participation and physical function across physical frailty groups at baseline. The "*" sign denotes significant differences between robust and pre-frail; "^" denotes significant differences between prefrail and frail; "#" denotes significant differences between robust and frail. The shape of the violin plot represents a kernel density estimate, illustrating the distribution of data across different values. The wider section of the violin indicates a denser data distribution. The thick line in the center of the violin marks the median, while the box plot shows the IQR and outliers. MET: metabolic equivalent; PA: physical activity; STS: sit-to-stand.

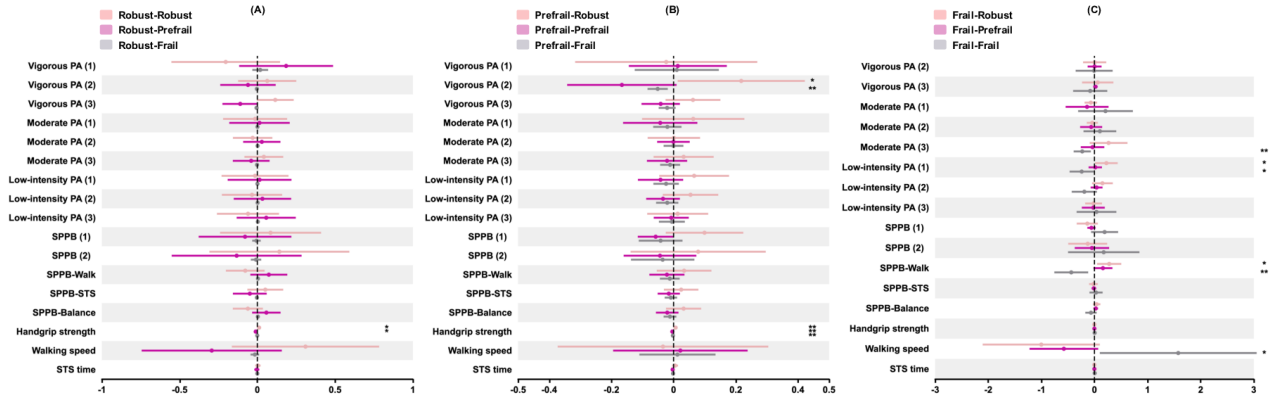


Predictors of Frailty Transitions

For participants robust at baseline, increased handgrip strength (AME=1.12%, 95% CI 0.21-2.03; *P*=.02) was associated with a higher probability of remaining robust, while decreased handgrip strength (AME=−1.06%, 95% CI −1.93 to −0.19; *P*=.02) was linked to a higher probability of transitioning to prefrailty (see Figure 4A).

XSL•FO
RenderX

Figure 4. Average marginal effect (AME) of each predictor on the probability of frailty transition in older adults. (**A**) "Robust" at baseline; (**B**) "Pre-frail" at baseline; (**C**) "Frail" at baseline. Notes: AME reflects the change in the probability of each frailty transition at follow-up based on a 1-unit change in a predictor. The "*" sign denotes *P*<.05 and "**" denotes *P*<.01. PA: physical activity; SPPB: Short Physical Performance Battery; STS: sit-to-stand.



Among prefrail participants, increased participation in vigorous PA (type 2; AME=21.76%, 95% CI 1.31%-42.20%; *P*=.04) and better handgrip strength (AME=0.64%, 95% CI 0.21%-1.07%; *P*=.003) were associated with a greater probability of transitioning to robust state. Conversely, decreased vigorous PA (type 2) participation (AME=−5.12%, 95% CI −8.38% to −1.87%; *P*=.002) and handgrip strength (AME=−0.23%, 95% CI −0.40% to −0.07%; *P*=.005) were linked to a higher likelihood of worsening to frailty. Furthermore, decreased handgrip strength (AME=−0.41%, 95% CI −0.70% to −0.13%; *P*=.005) also increased the probability of remaining prefrail (see Figure 4B).

For frail participants, increased low-intensity PA (type 1) participation (AME=22.48%, 95% CI 1.24-43.72; *P*=.04) and higher SPPB walking subscore (AME=27.73%, 95% CI 4.95%-50.51%; *P*=.02) were associated with an increased probability of improving to robust state. Decreased moderate (type 3; AME=−22.98%, 95% CI −39.17% to −6.79%; *P*=.005) and low-intensity (type 1; AME=−24.02%, 95% CI −46.76% to −1.27%; *P*=.04) PA participation and SPPB walking subscore (AME=−43.68%, 95% CI −75.56% to −11.80%; *P*=.007), along with increased walking speed (AME =157.77%, 95% CI 10.06%-305.49%; *P*=.04), were linked to a higher likelihood of remaining frail (see Figure 4C).

In addition to PA and physical function, several covariates were significantly associated with frailty transitions. Detailed regression results are presented in Tables S1-S3 in Multimedia Appendix 1.

Sensitivity Analysis

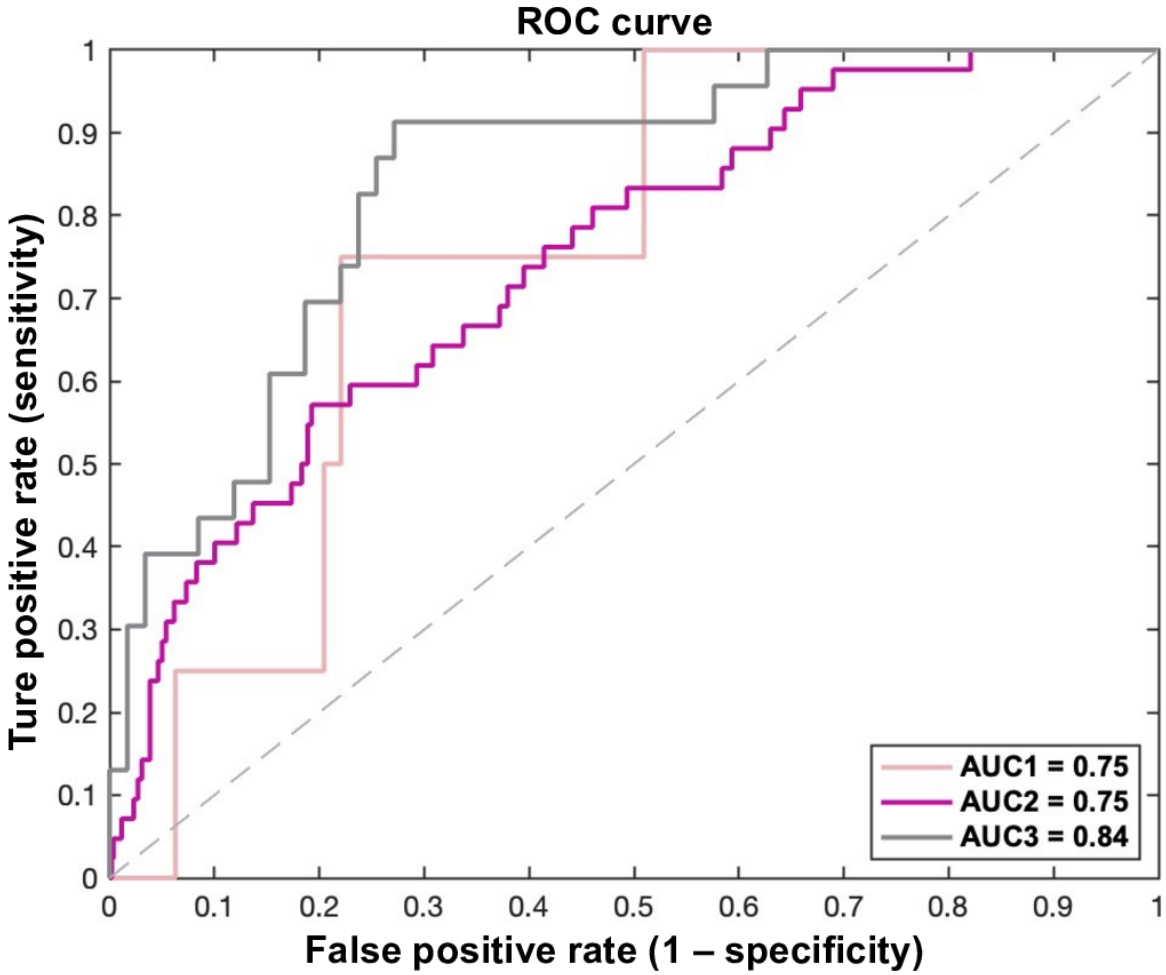
When handgrip strength was removed from the predictor list, the direction and magnitude of AMEs for PA and SPPB subcomponents were unchanged across all baseline frailty states. However, there is one exception: among prefrail participants, the AME for vigorous PA (type 2) on transitioning to robust status attenuated to 20.68% (95% CI −0.23 to 41.60) with *P*=.05, compared to *P*=.04 in the full model. This suggests that the previously observed significance for vigorous PA in prefrail individuals is moderately dependent on concurrent strength measures (see Tables S4–S6 in Multimedia Appendix 1).

Discriminative Ability of Physical Activity and Physical Function Variables for Frailty Transition

Figure 5 presents the discriminative performance of PA and physical function variables for frailty transitions across baseline frailty states. Among robust participants, predictors distinguished those who transitioned to frailty from those who remained non-frail with fair accuracy (AUC=0.75, 95% CI 0.53 - 0.92; *P*=.04). In the prefrail group, the model similarly discriminated individuals who worsened to frailty from those who remained or improved (AUC=0.75, 95% CI 0.67 - 0.82, *P*<.001). Finally, among frail participants, predictors showed good accuracy in identifying those who persisted in frailty versus those who recovered (AUC=0.84, 95% CI 0.75 - 0.93; *P*<.001).

XSL•FO
RenderX

Figure 5. Receiver operating characteristic (ROC) curves for discriminative ability of physical activity and physical function variables. (1) ROC curve 1 (AUC1) for discriminating transition to frailty versus nonfrailty among participants with robust baseline status; (2) ROC curve 2 (AUC2) for discriminating transition to frailty versus nonfrailty among participants with prefrail baseline status. (3) ROC curve 3 (AUC3) for discriminating frailty persistence versus transition to nonfrailty among participants with frail baseline status. AUC: area under the curve.



Discussion

Principal Findings and Comparison With Previous Works

The findings in this study provide valuable insights into how levels of PA participation and physical function measures influence frailty dynamics and highlight how these factors may differ across baseline frailty stages in Chinese older adults. Our results are consistent with the hypothesis that greater amounts of PA and better physical function are associated with a decreased likelihood of worsening frailty or a greater chance of transitioning to a more robust status. This is consistent with previous studies suggesting that PA and physical function are crucial in mitigating frailty progression [13,15]. However, the effect of these factors was not consistent across different baseline frailty states, suggesting that interventions may need to be tailored according to an individual's frailty stage.

Among participants who were robust at baseline, increased handgrip strength was associated with a higher probability of remaining robust. Conversely, decreased handgrip strength was linked to a higher probability of transitioning to the prefrail stage. Handgrip strength is widely recognized as a reliable indicator of overall muscular strength because it reflects the integrity of the neuromuscular system and correlates strongly with strength in other muscle groups [9,45]. As a simple, noninvasive measure, it provides valuable insights into an individual's functional capacity and overall health status, making it a practical tool for assessing frailty risk [9,46]. Our finding highlights the importance of maintaining muscular strength in healthy older adults, as it not only helps preserve physical function but also reduces the risk of transitioning to a more frail state [8,9]. In addition, our results suggest that PA level alone did not significantly predict the maintenance or decline in frailty state in this population. It is possible that PA in this group may not have been sufficiently differentiated or intense enough to cause a clear shift in frailty status, as the robust state may represent a relatively stable and resilient group. These findings support that strength-based indicators like handgrip strength may play a more critical role in maintaining or improving frailty status [8,9], indicating that muscle strength could be a more direct determinant of frailty transitions in this group. Strengthening interventions, such as resistance training, could be particularly effective for maintaining robustness in this group and preventing the onset of frailty [47].

XSL•FO
RenderX

For participants who were prefrail at baseline, the findings underscore the importance of vigorous PA and handgrip strength in promoting robustness among prefrail individuals. The amount of vigorous PA per week, which exceeds the recommended PA levels for older adults [48], has been shown to enhance overall health status, and thus for reversing prefrailty [49]. The positive association between vigorous PA and improved frailty status suggests that higher-intensity activities may stimulate physiological adaptations that counteract the decline in physical function associated with aging [50,51]. Conversely, the observed relationships between reduced PA, decreased handgrip strength, and worsening frailty highlight the risks of physical inactivity and muscle loss in this population [52,53]. These results emphasize the need for early and targeted interventions that combine vigorous PA and strength training to mitigate frailty progression for individuals in the prefrail stage. By addressing these modifiable risk factors early in the prefrail stage, it may be possible to delay or even prevent the onset of frailty, ultimately improving the quality of life of older adults and reducing the burden on healthcare systems [54].

In frail participants, increased participation in low-intensity PA and a higher SPPB walking subscore were associated with an increased probability of improving to a robust status. This finding is encouraging, as it suggests that even low-intensity PA can be effective in improving function in frail individuals, potentially reversing frailty and promoting recovery to a more robust state. This aligns with previous studies demonstrating the benefits of low-intensity exercises, such as walking and flexibility training, for frail older adults [55]. Furthermore, decreased participation in moderate and low-intensity PA, along with a lower SPPB walking subscore, was linked to a higher likelihood of remaining frail. This highlights the bidirectional relationship between PA and frailty: while reduced PA increases the risk of frailty, frailty itself can limit PA engagement due to physical limitations. This self-reinforcing cycle of decline, wherein frailty impedes PA participation, has been well-documented in the literature [15,47,56-64]. Our findings further emphasize this dynamic interaction, underscoring the need for early intervention to disrupt this cycle.

Unexpectedly, increased walking speed was associated with a higher likelihood of remaining frail in this study. While counterintuitive, this finding may reflect methodological and physiological nuances specific to the frail cohort. First, the wide CI (10.06% - 305.49%) indicates substantial variability in individual responses, suggesting that small speed improvements in some participants did not translate to holistic functional recovery. Second, frail older adults often experience fluctuations in mobility due to underlying health conditions, which may temporarily elevate walking speed without altering overall frailty status [65-67]. Thus, while improved speed is generally desirable, its isolated increase may be insufficient to reverse multifactorial frailty in this population. Future research should incorporate other mobility measures, such as walking endurance or dual-task gait performance, to better capture functional changes in frail individuals.

Interestingly, only 28% of participants (23/82) classified as frail at baseline remained frail after 2 years, with a recovery rate notably higher than the roughly 50% reported in other longitudinal studies [24]. Several methodological and cohort-specific factors likely contribute to this finding. For instance, selective attrition (ie, exclusion of participants who died or were unable to complete Wave 2) removes the most vulnerable individuals, biasing recovery rates to be upward. In addition, the relatively shorter follow-up period (2 y) and younger age of participants in our study may have contributed to these findings, as improvements in frailty are often more pronounced in younger populations over shorter timescales [24,25]. Finally, reliance on self-reported and occasionally incomplete frailty component data can introduce measurement error, further exaggerating transition rates. Taken together, these factors suggest that the high observed recovery rate may overstate true physiological reversal of frailty and should be interpreted with caution. Future studies with longer follow-up, objective frailty measures, and comprehensive attrition tracking are needed to distinguish genuine recovery from methodological artifacts.

The findings of this study underscore the importance of tailoring interventions based on baseline frailty status. For robust individuals, PA interventions should focus on maintaining strength through resistance training and preventing the decline of muscle function. In prefrail individuals, the emphasis should be on increasing the intensity of PA, especially vigorous activities, and improving handgrip strength to delay or prevent further frailty progression. Our sensitivity analysis confirms that most associations between PA, SPPB subcomponents, and frailty transitions are robust to the exclusion of handgrip strength. The borderline attenuation of the vigorous PA effect among prefrail participants (P=.05) indicates a possible interplay between high-intensity activity and muscle strength, warranting further investigation into tailoring PA prescriptions according to baseline strength levels. For frail individuals, interventions should prioritize low-intensity PA and the improvement of basic functional abilities to prevent further deterioration. In addition, factors such as age and gender differences may influence responses to PA interventions, warranting further investigation. Future research with larger sample sizes is needed to explore how personalized interventions can better address the needs of frail individuals by setting achievable goals that gradually build functional capacity without overexertion [68].

This study has several strengths, including its focus on a large cohort of Chinese older adults and the inclusion of both PA and physical function measures. The AUC of discriminative ability for frailty transition capacity is good for participants with a frail state at baseline, suggesting that the model performs well. For individuals with nonfrail status at baseline, the AUC is 0.75, indicating fair discriminative capacity [44]. This is considered meaningful in the context of frailty transitions, which are influenced by a wide range of biological, behavioral, and environmental factors [4].

However, the study also has several limitations. First, the use of self-reported PA data may introduce recall bias, although the IPAQ has been widely used and validated in older populations [37]. Second, the exclusion of participants with missing data may have introduced selection bias. While the CHARLS manual outlined potential reasons for missing data (eg, feeling unsafe or incapable to perform the tests), these reasons were not clearly

XSL•FO
RenderX

indicated in the dataset. As a result, it is challenging to accurately determine the cause of missing data for most participants. To avoid speculation, we only proceeded with the available data. This limitation may have contributed to the selection bias as participants appeared healthier at follow-up compared to baseline. Furthermore, our analytic sample included 52% (528/1014) men and 48% (486/1014) women, whereas many older-adult cohort studies report a female majority due to women's greater longevity [69]. This near-equal sex distribution likely reflects our requirement that participants survive and complete both CHARLS waves (2011 and 2013). Frailer individuals, disproportionately women at advanced ages, were more likely to die or be lost to follow-up, thus attenuating the expected women predominance. We acknowledge this potential selection bias and its implications for generalizability. Third, despite the longitudinal design, this study is observational and cannot establish causal relationships between PA amounts, physical function, and frailty transitions. Longitudinal studies with longer follow-up periods and larger sample sizes could provide further insights into the dynamic nature of frailty and its predictors. Fourth, our findings are based on a specific cohort of Chinese older adults, and the generalizability of these results to other populations remains uncertain. Further studies are needed to explore how these predictors of frailty transitions may vary across different cultural, social, and economic contexts. Finally, our study focused on physical frailty and did not consider other dimensions of frailty, such as cognitive or social frailty, which may also influence frailty transitions. Future research should incorporate a broader range of determinants to provide a more holistic understanding of frailty dynamics.

Conclusions

In conclusion, this study provides important insights into the role of PA participation and physical function in frailty transitions among older Chinese adults. Our findings suggest that both PA and physical function measures significantly influence frailty dynamics, with varying effects depending on baseline frailty status. Tailored interventions that address the specific needs of robust, prefrail, and frail individuals are essential for preventing or delaying frailty progression. Future research should explore more personalized approaches to PA interventions, especially for frail individuals, to ensure that activities align with their functional capacity and promote meaningful improvement in physical function.

Acknowledgments

This study used data from the two waves of the China Health and Retirement Longitudinal Study (CHARLS), collected from 2011 to 2013. We would like to express our sincere gratitude to the CHARLS research team and all the participants for their valuable contributions to this study.

Authors' Contributions

ZZ and YY conceptualized and designed the research. ZZ performed the data analysis, interpreted the results, and drafted the initial manuscript. CLH, CH-PS, SH-SW, and YY provided critical review and edits to the manuscript. All authors have read and approved the final version of the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Ordinal logistic regression analysis of predictors and covariates associated with frailty transitions.
[DOCX File, 56 KB - aging_v8i1e75887_app1.docx]

References

1. Chen X, Giles J, Yao Y, et al. The path to healthy ageing in China: a Peking University–Lancet Commission. The Lancet 2022 Dec;400(10367):1967-2006. [doi: 10.1016/S0140-6736(22)01546-X]
2. Mather M, Kilduff L. Aging and health in China: what can we learn from the world's largest population of older people. : Population Reference Bureau; 2020 URL: https://www.prb.org/resources/china-aging-worlds-largest-population-older-people/ [accessed 2025-08-25]
3. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
4. Bauer JM, Sieber CC. Sarcopenia and frailty: a clinician's controversial point of view. Exp Gerontol 2008 Jul;43(7):674-678. [doi: 10.1016/j.exger.2008.03.007] [Medline: 18440743]
5. Kojima G, Taniguchi Y, Iliffe S, Jivraj S, Walters K. Transitions between frailty states among community-dwelling older people: A systematic review and meta-analysis. Ageing Res Rev 2019 Mar;50:81-88. [doi: 10.1016/j.arr.2019.01.010] [Medline: 30659942]
6. He B, Ma Y, Wang C, et al. Prevalence and risk factors for frailty among community-dwelling older people in China: a systematic review and meta-analysis. J Nutr Health Aging 2019;23(5):442-450. [doi: 10.1007/s12603-019-1179-9] [Medline: 31021361]

XSL•FO
RenderX

7. Op het Veld LPM, van Rossum E, Kempen G, de Vet HCW, Hajema K, Beurskens A. Fried phenotype of frailty: cross-sectional comparison of three frailty stages on various health domains. BMC Geriatr 2015 Jul 9;15:77. [doi: 10.1186/s12877-015-0078-0] [Medline: 26155837]
8. Chung CJ, Wu C, Jones M, et al. Reduced handgrip strength as a marker of frailty predicts clinical outcomes in patients with heart failure undergoing ventricular assist device placement. J Card Fail 2014 May;20(5):310-315. [doi: 10.1016/j.cardfail.2014.02.008] [Medline: 24569037]
9. Syddall H, Cooper C, Martin F, Briggs R, Aihie Sayer A. Is grip strength a useful single marker of frailty? Age Ageing 2003 Nov;32(6):650-656. [doi: 10.1093/ageing/afg111] [Medline: 14600007]
10. Cruz-Jentoft AJ, Baeyens JP, Bauer JM, et al. Sarcopenia: European consensus on definition and diagnosis: Report of the European Working Group on Sarcopenia in Older People. Age Ageing 2010 Jul;39(4):412-423. [doi: 10.1093/ageing/afq034] [Medline: 20392703]
11. Verghese J, Xue X. Identifying frailty in high functioning older adults with normal mobility. Age Ageing 2010 May;39(3):382-385. [doi: 10.1093/ageing/afp226] [Medline: 20051607]
12. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994 Mar;49(2):M85-M94. [doi: 10.1093/geronj/49.2.m85] [Medline: 8126356]
13. Lang PO, Michel JP, Zekry D. Frailty syndrome: a transitional state in a dynamic process. Gerontology 2009;55(5):539-549. [doi: 10.1159/000211949] [Medline: 19346741]
14. Physical activity. World Health Organization. 2018. URL: https://www.who.int/news-room/fact-sheets/detail/physical-activity [accessed 2025-08-21]
15. O'Connell ML, Coppinger T, McCarthy AL. The role of nutrition and physical activity in frailty: A review. Clin Nutr ESPEN 2020 Feb;35:1-11. [doi: 10.1016/j.clnesp.2019.11.003] [Medline: 31987100]
16. Pate RR, Pratt M, Blair SN, et al. Physical activity and public health. A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. JAMA 1995 Feb 1;273(5):402-407. [doi: 10.1001/jama.273.5.402] [Medline: 7823386]
17. Sattelmair J, Pertman J, Ding EL, Kohl HW III, Haskell W, Lee IM. Dose response between physical activity and risk of coronary heart disease. Circulation 2011 Aug 16;124(7):789-795. [doi: 10.1161/CIRCULATIONAHA.110.010710] [Medline: 21810663]
18. Aune D, Norat T, Leitzmann M, Tonstad S, Vatten LJ. Physical activity and the risk of type 2 diabetes: a systematic review and dose–response meta-analysis. Eur J Epidemiol 2015 Jul;30(7):529-542. [doi: 10.1007/s10654-015-0056-z]
19. Giovannucci E, Ascherio A, Rimm EB, Colditz GA, Stampfer MJ, Willett WC. Physical activity, obesity, and risk for colon cancer and adenoma in men. Ann Intern Med 1995 Mar 1;122(5):327-334. [doi: 10.7326/0003-4819-122-5-199503010-00002] [Medline: 7847643]
20. Sherrington C, Michaleff ZA, Fairhall N, et al. Exercise to prevent falls in older adults: an updated systematic review and meta-analysis. Br J Sports Med 2017 Dec;51(24):1750-1758. [doi: 10.1136/bjsports-2016-096547]
21. Vagetti GC, Barbosa Filho VC, Moreira NB, Oliveira VD, Mazzardo O, Campos WD. Association between physical activity and quality of life in the elderly: a systematic review, 2000-2012. Braz J Psychiatry 2014;36(1):76-88. [doi: 10.1590/1516-4446-2012-0895] [Medline: 24554274]
22. McPhee JS, French DP, Jackson D, Nazroo J, Pendleton N, Degens H. Physical activity in older age: perspectives for healthy ageing and frailty. Biogerontology 2016 Jun;17(3):567-580. [doi: 10.1007/s10522-016-9641-0]
23. Jia L, Du Y, Chu L, et al. Prevalence, risk factors, and management of dementia and mild cognitive impairment in adults aged 60 years or older in China: a cross-sectional study. Lancet Public Health 2020 Dec;5(12):e661-e671. [doi: 10.1016/S2468-2667(20)30185-7] [Medline: 33271079]
24. Rivera-Almaraz A, Salinas-Rodríguez A, Gutiérrez-Peña E, Manrique-Espinoza BS. Predictors of frailty transitions in Mexican older adults. J Gerontol A Biol Sci Med Sci 2024 Apr 1;79(4):glae068. [doi: 10.1093/gerona/glae068] [Medline: 38401153]
25. Rodríguez-Laso Á, García-García FJ, Rodríguez-Mañas L. Transitions Between Frailty States and Its Predictors in a Cohort of Community-Dwelling Spaniards. J Am Med Dir Assoc 2022 Mar;23(3):524. [doi: 10.1016/j.jamda.2021.07.008] [Medline: 34389334]
26. Majid Z, Welch C, Davies J, Jackson T. Global frailty: The role of ethnicity, migration and socioeconomic factors. Maturitas 2020 Sep;139:33-41. [doi: 10.1016/j.maturitas.2020.05.010] [Medline: 32747038]
27. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). Int J Epidemiol 2014 Feb;43(1):61-68. [doi: 10.1093/ije/dys203] [Medline: 23243115]
28. Wang R, Chen Z, Zhou Y, Shen L, Zhang Z, Wu X. Melancholy or mahjong? Diversity, frequency, type, and rural-urban divide of social participation and depression in middle- and old-aged Chinese: A fixed-effects analysis. Soc Sci Med 2019 Oct;238:112518. [doi: 10.1016/j.socscimed.2019.112518] [Medline: 31473574]
29. Lee JSW, Auyeung TW, Leung J, Kwok T, Woo J. Transitions in frailty states among community-living older adults and their associated factors. J Am Med Dir Assoc 2014 Apr;15(4):281-286. [doi: 10.1016/j.jamda.2013.12.002] [Medline: 24534517]

XSL•FO
RenderX

30. McHugh JE, Dowling M, Butler A, Lawlor BA. Psychological distress and frailty transitions over time in community-dwelling older adults. Ir J Psychol Med 2016 Jun;33(2):111-119. [doi: 10.1017/ipm.2015.67] [Medline: 30115144]
31. Wang H, Chen H. Aging in China: challenges and opportunities. China CDC Wkly 2022 Jul 8;4(27):601-602. [doi: 10.46234/ccdcw2022.130] [Medline: 35919296]
32. Xu W, Li YX, Hu Y, Wu C. Association of frailty with recovery from disability among community-dwelling Chinese older adults: China health and retirement longitudinal study. BMC Geriatr 2020 Mar 30;20(1):119. [doi: 10.1186/s12877-020-01519-6] [Medline: 32228463]
33. Li Y, Xue QL, Odden MC, Chen X, Wu C. Linking early life risk factors to frailty in old age: evidence from the China Health and Retirement Longitudinal Study. Age Ageing 2020 Feb 27;49(2):208-217. [doi: 10.1093/ageing/afz160] [Medline: 31957780]
34. Wu C, Smit E, Xue QL, Odden MC. Prevalence and correlates of frailty among community-dwelling Chinese older adults: the China health and retirement longitudinal study. The Journals of Gerontology 2018 Jan 1;73(1):102-108. [doi: 10.1093/gerona/glx098]
35. Xu W, Li YX, Wu C. Incidence of frailty among community-dwelling older adults: a nationally representative profile in China. BMC Geriatr 2019 Dec 30;19(1):378. [doi: 10.1186/s12877-019-1393-7] [Medline: 31888498]
36. Sha S, Chan SHW, Chen L, Xu Y, Pan Y. The association between trajectories of loneliness and physical frailty in Chinese older adults: does age matter? Int J Environ Res Public Health 2022 Apr 22;19(9):5105. [doi: 10.3390/ijerph19095105] [Medline: 35564500]
37. Li X, Zhang W, Zhang W, et al. Level of physical activity among middle-aged and older Chinese people: evidence from the China health and retirement longitudinal study. BMC Public Health 2020 Dec;20(1):1-13. [doi: 10.1186/s12889-020-09671-9]
38. Tian Y, Shi Z. Effects of physical activity on daily physical function in chinese middle-aged and older adults: a longitudinal study from CHARLS. J Clin Med 2022 Nov 2;11(21):6514. [doi: 10.3390/jcm11216514] [Medline: 36362742]
39. Li S, Zhang J, Yang Y. Correlation between the physical activity volume and cognitive and mental capacity among older adult people in China: a cross-sectional study based on the 2020 CHARLS database. Front Public Health 2024;12:1462570. [doi: 10.3389/fpubh.2024.1462570]
40. Zhong BX, Zhong HL, Zhou GQ, Xu WQ, Lu Y, Zhao Q. Physical performance and risk of hip fracture in community-dwelling elderly people in China: A 4-year longitudinal cohort study. Maturitas 2021 Apr;146:26-33. [doi: 10.1016/j.maturitas.2021.01.003] [Medline: 33722361]
41. Guralnik JM, Ferrucci L, Simonsick EM, Salive ME, Wallace RB. Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. N Engl J Med 1995 Mar 2;332(9):556-561. [doi: 10.1056/NEJM199503023320902] [Medline: 7838189]
42. Rubin DB. Multiple Imputation for Survey Nonresponse: Wiley; 1987. [doi: 10.1002/9780470316696]
43. Schmoor C, Schumacher M, Finke J, Beyersmann J. Competing risks and multistate models. Clin Cancer Res 2013 Jan 1;19(1):12-21. [doi: 10.1158/1078-0432.CCR-12-1619] [Medline: 23172882]
44. Swets JA. Measuring the accuracy of diagnostic systems. Science 1988 Jun 3;240(4857):1285-1293. [doi: 10.1126/science.3287615] [Medline: 3287615]
45. Shaughnessy KA, Hackney KJ, Clark BC, et al. A narrative review of handgrip strength and cognitive functioning: bringing a new characteristic to muscle memory. Journal of Alzheimer's Disease 2020 Feb 18;73(4):1265-1278. [doi: 10.3233/JAD-190856]
46. Vaishya R, Misra A, Vaish A, Ursino N, D'Ambrosi R. Hand grip strength as a proposed new vital sign of health: a narrative review of evidences. J Health Popul Nutr 2024 Jan 9;43(1):7. [doi: 10.1186/s41043-024-00500-y] [Medline: 38195493]
47. Peterson MJ, Giuliani C, Morey MC, et al. Physical activity as a preventative factor for frailty: the health, aging, and body composition study. J Gerontol A Biol Sci Med Sci 2009 Jan;64(1):61-68. [doi: 10.1093/gerona/gln001] [Medline: 19164276]
48. Piercy KL, Troiano RP, Ballard RM, et al. The physical activity guidelines for Americans. JAMA 2018 Nov 20;320(19):2020-2028. [doi: 10.1001/jama.2018.14854] [Medline: 30418471]
49. Takamura M, Sone T, Kawamura T, Suzuki R, Moriyama N, Yasumura S. A cross-sectional study on the characteristics of physical activity in pre-frail older adults. Int J Environ Res Public Health 2021 Nov 24;18(23):23. [doi: 10.3390/ijerph182312328] [Medline: 34886059]
50. Kendall KL, Fairman CM. Women and exercise in aging. J Sport Health Sci 2014 Sep;3(3):170-178. [doi: 10.1016/j.jshs.2014.02.001]
51. Stamatakis E, Huang BH, Maher C, et al. Untapping the health enhancing potential of vigorous intermittent lifestyle physical activity (VILPA): rationale, scoping review, and a 4-pillar research framework. Sports Med 2021 Jan;51(1):1-10. [doi: 10.1007/s40279-020-01368-8] [Medline: 33108651]
52. Hengeveld LM, Wijnhoven HAH, Olthof MR, et al. Prospective associations of diet quality with incident frailty in older adults: the health, aging, and body composition study. J Am Geriatr Soc 2019 Sep;67(9):1835-1842. [doi: 10.1111/jgs.16011] [Medline: 31267522]
53. Doi T, Makizako H, Tsutsumimoto K, et al. Transitional status and modifiable risk of frailty in Japanese older adults: A prospective cohort study. Geriatrics Gerontology Int 2018 Nov;18(11):1562-1566. [doi: 10.1111/ggi.13525]

XSL•FO
RenderX

54. Zeng XZ, Meng LB, Li YY, et al. Prevalence and factors associated with frailty and pre-frailty in the older adults in China: a national cross-sectional study. Front Public Health 2023;11:1110648. [doi: 10.3389/fpubh.2023.1110648]
55. Tse ACY, Wong TWL, Lee PH. Effect of low-intensity exercise on physical and cognitive health in older adults: a systematic review. Sports Med Open 2015;1(1):37. [doi: 10.1186/s40798-015-0034-8] [Medline: 26512340]
56. Kehler DS, Theou O. The impact of physical activity and sedentary behaviors on frailty levels. Mech Ageing Dev 2019 Jun;180:29-41. [doi: 10.1016/j.mad.2019.03.004] [Medline: 30926562]
57. Apóstolo J, Cooke R, Bobrowicz-Campos E, et al. Effectiveness of interventions to prevent pre-frailty and frailty progression in older adults: a systematic review. JBI Database System Rev Implement Rep 2018 Jan;16(1):140-232. [doi: 10.11124/JBISRIR-2017-003382] [Medline: 29324562]
58. Giné-Garriga M, Roqué-Fíguls M, Coll-Planas L, Sitjà-Rabert M, Salvà A. Physical exercise interventions for improving performance-based measures of physical function in community-dwelling, frail older adults: a systematic review and meta-analysis. Arch Phys Med Rehabil 2014 Apr;95(4):753-769. [doi: 10.1016/j.apmr.2013.11.007] [Medline: 24291597]
59. Angulo J, El Assar M, Álvarez-Bustos A, Rodríguez-Mañas L. Physical activity and exercise: Strategies to manage frailty. Redox Biol 2020 Aug;35:101513. [doi: 10.1016/j.redox.2020.101513] [Medline: 32234291]
60. Savela SL, Koistinen P, Stenholm S, et al. Leisure-time physical activity in midlife is related to old age frailty. J Gerontol A Biol Sci Med Sci 2013 Nov;68(11):1433-1438. [doi: 10.1093/gerona/glt029] [Medline: 23525478]
61. Borda MG, Pérez-Zepeda MU, Samper-Ternent R, Gómez RC, Avila-Funes JA, Cano-Gutierrez CA. The influence of lifestyle behaviors on the incidence of frailty. J Frailty Aging 2020;9(3):144-149. [doi: 10.14283/jfa.2019.37] [Medline: 32588028]
62. Soares WJS, Lima CA, Bilton TL, Ferrioli E, Dias RC, Perracini MR. Association among measures of mobility-related disability and self-perceived fatigue among older people: a population-based study. Braz J Phys Ther 2015;19(3):194-200. [doi: 10.1590/bjpt-rbf.2014.0091] [Medline: 26039035]
63. Blodgett J, Theou O, Kirkland S, Andreou P, Rockwood K. The association between sedentary behaviour, moderate-vigorous physical activity and frailty in NHANES cohorts. Maturitas 2015 Feb;80(2):187-191. [doi: 10.1016/j.maturitas.2014.11.010] [Medline: 25542406]
64. Stojanovic M, Babulal GM, Head D. Determinants of physical activity engagement in older adults. J Behav Med 2023 Oct;46(5):757-769. [doi: 10.1007/s10865-023-00404-y] [Medline: 36920727]
65. Fried LP, Ferrucci L, Darer J, Williamson JD, Anderson G. Untangling the concepts of disability, frailty, and comorbidity: implications for improved targeting and care. J Gerontol A Biol Sci Med Sci 2004 Mar;59(3):255-263. [doi: 10.1093/gerona/59.3.m255] [Medline: 15031310]
66. Avila-Funes JA, Helmer C, Amieva H, et al. Frailty among community-dwelling elderly people in France: the three-city study. J Gerontol A Biol Sci Med Sci 2008 Oct;63(10):1089-1096. [doi: 10.1093/gerona/63.10.1089] [Medline: 18948560]
67. Cowper PA, Peterson MJ, Pieper CF, et al. Economic analysis of primary care-based physical activity counseling in older men: the VA-LIFE Trial. J Am Geriatr Soc 2017 Mar;65(3):533-539. [doi: 10.1111/jgs.14567] [Medline: 28152170]
68. Allison R 2nd, Assadzandi S, Adelman M. Frailty: evaluation and management. Am Fam Physician 2021 Feb 15;103(4):219-226. [Medline: 33587574]
69. Zeidan RS, McElroy T, Rathor L, Martenson MS, Lin Y, Mankowski RT. Sex differences in frailty among older adults. Exp Gerontol 2023 Dec;184:112333. [doi: 10.1016/j.exger.2023.112333] [Medline: 37993077]

Abbreviations

ADL: activity of daily living
AME: average marginal effect
AUC: area under the curve
CHARLS: China Health and Retirement Longitudinal Study
CHS: Cardiovascular Health Study
IADL: instrumental activity of daily living
IPAQ: International Physical Activity Questionnaire
MTE: metabolic equivalent
PA: physical activity
PFP: physical frailty phenotype
ROC: receiver operating characteristic
SPPB: Short Physical Performance Battery
STS: sit-to-stand

XSL•FO
RenderX

Edited by M Gray; submitted 13.04.25; peer-reviewed by AC Cavalcanti Maciel, C Maxwell, S Mizutani; revised version received 28.07.25; accepted 30.07.25; published 15.09.25.

Please cite as:
Zeng Z, Hsu CL, Sit CHP, Wong SHS, Yang Y
The Role of Physical Activity and Physical Function in Predicting Physical Frailty Transitions in Chinese Older Adults: Longitudinal Observational Study From CHARLS
JMIR Aging 2025;8:e75887
URL: https://aging.jmir.org/2025/1/e75887
doi:10.2196/75887

© Ziwei Zeng, Chun Liang Hsu, Cindy Hui-ping Sit, Stephen Heung-sang Wong, Yijian Yang. Originally published in JMIR Aging (https://aging.jmir.org), 15.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Predictive Validity of Hospital-Associated Complications of Older People Identified Using Diagnosis Procedure Combination Data From an Acute Care Hospital in Japan: Observational Study

Seigo Mitsutake[1,2], PhD, PT; Tatsuro Ishizaki[1], MD, PhD, MPH; Shohei Yano[1,3], PT, PhD; Takumi Hirata[1], MD, PhD, MPH; Kae Ito[1], MD, PhD; Ko Furuta[4], MD; Yoshitomo Shimazaki[5], RPh, PhD; Hideki Ito[6], MD, PhD; Alison Mudge[7,8], MD, PhD; Kenji Toba[6], MD, PhD

[1]Human Care Research Team, Tokyo Metropolitan Institute for Geriatrics and Gerontology, Tokyo, Japan

[2]Australian Institute of Health Innovation, Macquarie University, Sydney, Australia

[3]The Salvation Army Booth Memorial Hospital, Tokyo, Japan

[4]Department of Psychiatry, Tokyo Metropolitan Institute for Geriatrics and Gerontology, Tokyo, Japan

[5]Department of Pharmacy, Tokyo Metropolitan Institute for Geriatrics and Gerontology, Tokyo, Japan

[6]Tokyo Metropolitan Institute for Geriatrics and Gerontology, Tokyo, Japan

[7]Department of Internal Medicine and Aged Care, Royal Brisbane and Women's Hospital, Brisbane, Australia

[8]Centre of Health Services Research, The University of Queensland, Brisbane, Australia

Corresponding Author:
Seigo Mitsutake, PhD, PT
Human Care Research Team
Tokyo Metropolitan Institute for Geriatrics and Gerontology
35-2 Sakae-cho
Itabashi-ku
Tokyo, 173 0015
Japan
Phone: 81 3 3964 3241 ext 4229
Email: mitsu@tmig.or.jp

Abstract

Background: A composite outcome of hospital-associated complications of older people (HAC-OP; comprising functional decline, delirium, incontinence, falls, and pressure injuries) has been proposed as an outcome measure reflecting quality of acute hospital care. Estimating HAC-OP from routinely collected administrative data could facilitate the rapid and standardized evaluation of interventions in the clinical setting, thereby supporting the development, improvement, and wider implementation of effective interventions.

Objective: This study aimed to create a Diagnosis Procedure Combination (DPC) data version of the HAC-OP measure (HAC-OP-DPC) and demonstrate its predictive validity by assessing its associations with hospital length of stay (LOS) and discharge destination.

Methods: This retrospective cohort study acquired DPC data (routinely collected administrative data) from a general acute care hospital in Tokyo, Japan. We included data from index hospitalizations for patients aged ≥65 years hospitalized for ≥3 days and discharged between July 2016 and March 2021. HAC-OP-DPC were identified using diagnostic codes for functional decline, incontinence, delirium, pressure injury, and falls occurring during the index hospitalization. Generalized linear regression models were used to examine the associations between HAC-OP-DPC and LOS, and logistic regression models were used to examine the associations between HAC-OP-DPC and discharge to other hospitals and long-term care facilities (LTCFs).

Results: Among 15,278 patients, 3610 (23.6%) patients had coding evidence of one or more HAC-OP-DPC (1: 18.8% and ≥2: 4.8%). Using "no HAC-OP-DPC" as the reference category, the analysis showed a significant and graded association with longer LOS (adjusted risk ratio for patients with one HAC-OP-DPC 1.29, 95% CI 1.25-1.33; adjusted risk ratio for ≥2 HAC-OP-DPC 1.97, 95% CI 1.87-2.08), discharge to another hospital (adjusted odds ratio [AOR] for one HAC-OP-DPC 2.36, 95% CI 2.10-2.65; AOR for ≥2 HAC-OP-DPC 6.96, 95% CI 5.81-8.35), and discharge to LTCFs (AOR for one HAC-OP-DPC 1.35, 95% CI

XSL•FO
RenderX

1.09-1.67; AOR for ≥2 HAC-OP-DPC 1.68, 95% CI 1.18-2.39). Each individual HAC-OP was also significantly associated with longer LOS and discharge to another hospital, but only delirium was associated with discharge to LTCF.

Conclusions: This study demonstrated the predictive validity of the HAC-OP-DPC measure for longer LOS and discharge to other hospitals and LTCFs. To attain a more robust understanding of these relationships, additional studies are needed to verify our findings in other hospitals and regions. The clinical implementation of HAC-OP-DPC, which is identified using routinely collected administrative data, could support the evaluation of integrated interventions aimed at optimizing inpatient care for older adults.

(JMIR Aging 2025;8:e68267) doi:10.2196/68267

KEYWORDS

delirium; functional decline; Japan; older adult; routinely collected health data; elder; hospital complication; HAC-OP; incontinence; pressure injury; inpatient care; diagnosis procedure combination; predictive validity; hospital length of stay; administrative data; acute care; index hospitalization; diagnostic code; linear regression; logistic regression; long-term care; retrospective cohort; observational study; patient care; gerontology; hospital care; patient complication

Introduction

Hospitalization can impose a heavy physical and psychological burden on older adults, leading to in-hospital complications such as functional decline (the loss of independence in activities of daily living) and delirium [1]. A recent meta-analysis reported that the prevalence of functional decline following acute hospitalization was 30% in older adults aged ≥65 years [2]. Such functional decline is associated with various adverse outcomes, including prolonged hospital length of stay (LOS) and increased risks of hospital readmission and mortality within 30 days after discharge [3,4]. Similarly, delirium is estimated to affect around one quarter of older hospitalized adults [5-7], and is associated with longer stays, higher mortality, and higher health care costs [8,9].

Previous studies have developed and demonstrated the effectiveness of interventions to prevent such complications in hospitalized older adults [10,11]. While these interventions have generally focused on single complications as outcomes, they may also be effective in preventing other complications (eg, delirium prevention programs can also reduce falls) [1,12-16]. This is because in-hospital complications common in older people (often termed “geriatric syndromes”) tend to have overlapping risk factors that can be collectively targeted by complex and multifaceted interventions [1,12-16]. Mudge et al [1] proposed a multicomponent measure of “hospital-associated complications of older people” (HAC-OP) comprising common complications (functional decline, hospital-associated incontinence, hospital-associated delirium, pressure injury, or fall) among older adults admitted to acute care hospitals. In a prospective study using regular structured patient assessments as well as document review, the authors demonstrated a significant graded association between the HAC-OP measure and hospital LOS, facility discharge, and mortality within 6 months after admission. However, collecting this research measure was resource-intensive [1,17], which makes it difficult to replicate in routine practice. Routinely collected administrative data have been examined to assist clinicians, patients, and policy makers in making informed decisions [18,19]. Estimating HAC-OP from routinely collected administrative data could facilitate the rapid and standardized evaluation of interventions in the clinical setting, thereby supporting the development, improvement, and wider implementation of effective interventions.

In Japan, the majority of acute care hospitals use the Diagnosis Procedure Combination (DPC) case-mix patient classification system, which is linked to a lump-sum payment system for inpatients [20]. DPC-compliant hospitals must generate and submit DPC data to the government. These data include administrative claims and discharge abstracts containing patient-level information on diagnoses (recorded using *ICD-10* [*International Statistical Classification of Diseases, Tenth Revision*] codes), treatments, and prescribed drugs. Unlike many other countries, the required data also include activities of daily living (ADL) assessment scores at admission and discharge [20]. Prior studies have used DPC data to assess ADL scores in patients to examine the effects of rehabilitation services and to identify patients at high risk of early readmission in acute care settings [21,22]. Hospital outcomes, such as LOS and readmission within 28 days after discharge from hospital, are published annually for each DPC hospital [23]. However, in-hospital complications are not currently reported or benchmarked as they are in some other countries [24,25]. The systematic and accurate identification of HAC-OP from DPC data would support the evaluation of inpatient care and interventions that target these complications in Japan’s acute care hospitals.

Although the validity of chronic disease diagnoses in DPC data has been reported to be generally high [26], no study has tested the validity of in-hospital complications in DPC data, and complications are often underestimated in administrative datasets [27]. Understanding the current reporting of in-hospital complications and the association with important outcomes such as LOS and discharge destination could inform their use as an efficient and standardized evaluation of system-level interventions. Therefore, this study was conducted to develop a DPC data version of the HAC-OP measure (HAC-OP-DPC), describe the incidence of HAC-OP-DPC in a cohort of older acute care inpatients, and evaluate the predictive validity of this composite measure and its components by assessing associations with LOS and discharge destination.

XSL•FO
RenderX

Methods

Study Design and Patients

This retrospective cohort study was conducted using an anonymized DPC database obtained from a large general public acute care hospital (550 beds: 520 beds on general wards and 30 beds on psychiatric wards) in Tokyo, Japan. The DPC data comprised patient-level demographic characteristics, *ICD-10* codes, treatments, and prescribed drugs during all insurance-covered clinical encounters. The study used data from July 2016 to March 2021.

Patients who had been admitted to the study hospital from home or a long-term care facility (LTCF) and discharged during the study period were eligible for inclusion; the first hospitalization episode during the study period was designated the index hospitalization and included in the analysis. We excluded patients aged <65 years, patients discharged within 2 days of admission, patients who died during the index hospitalization, and patients with missing data in the study variables.

Ethical Considerations

The study protocol was approved by the Ethics Committee of the Tokyo Metropolitan Geriatric Hospital and Institute of Gerontology (approval number R18-20). All procedures followed the ethical guidelines of the Medical and Biological Research Involving Human Subjects established by the Japanese government. Opt-out consent was used because all data were anonymized before being received by the authors.

Measures

HAC-OP-DPC

Based on the original HAC-OP measure [1], we identified the following 5 conditions as HAC-OP-DPC: hospital-associated functional decline, incontinence, delirium, pressure injury, and fall. Each condition was defined using DPC data.

First, hospital-associated functional decline was defined as a decrease in ADL scores for the Barthel Index (BI) components of bathing, dressing, toileting, transfers, mobility, and feeding from hospital admission to discharge. In the BI, each component is given a score of 0, 5, 10, or 15 points (maximum scores vary among the components), with higher scores indicating greater independence in that activity [28]. BI was assessed at admission and hospital discharge by bedside nurses. Second, hospital-associated incontinence was defined as a decrease in scores for the BI components of bladder function and bowel function from hospital admission to discharge. Third, hospital-associated delirium was identified based on a recorded diagnosis of delirium (Multimedia Appendix 1) as a postadmission complication and recorded prescriptions of drugs used to manage agitation in delirium (injections of haloperidol or other antipsychotic drugs identified using prescription codes that remained constant throughout the study period). Fourth, hospital-associated pressure injury was identified based on a recorded diagnosis of pressure injury as a postadmission complication and discharge abstract records indicating pressure injury treatment during hospitalization without any similar treatment at admission. Fifth, a hospital-associated fall was identified based on a recorded diagnosis of fall as a postadmission complication. The overall multicomponent HAC-OP-DPC measure was categorized into none, 1, and 2 or more complications based on the count of conditions occurring in each patient.

Outcome Measures

The study outcome measures were hospital LOS during the index hospitalization, discharge to other hospitals, and discharge to LTCFs. LOS was calculated as the number of days between the dates of admission and discharge. Discharges to other hospitals, such as rehabilitation hospitals and LTCFs (including special nursing homes, private paid care facilities, and social welfare institutions), were identified using the relevant DPC codes indicating discharge destination.

Covariates

Using the subject hospital's DPC data, we extracted demographic variables of patient sex, age group (65-74, 75-84, and ≥85 years), and annual household income (<¥3.7 million, ≥¥3.7 million, and unknown; ¥1=US $0.0092 in 2016) at the index hospitalization [29]. Income was estimated from the available data about insurance copayments. Insurance copayment rates are the designated rates that patients pay at the point of care in Japan. For patients who have an annual household income below ¥3.7 million (approximately US $34,040; ¥1=US $0.0092), the copayment rates are 10% and 20% for patients aged ≥75 years and 70-74 years, respectively [29]. For patients aged ≥70 years who have an annual household income of ¥3.7 million or higher (¥1=US $0.0092), the copayment rate is 30%. The DPC data did not indicate the copayment rates for patients who received public medical assistance and patients aged 65-69 years. Therefore, income for these cases was categorized as "unknown."

We calculated variables for disease category, comorbidity, and frailty using *ICD-10* codes. Using previously described methods [30,31], we grouped patients into 12 disease categories based on their recorded primary diagnosis for admission. Next, we determined each patient's score in the Charlson Comorbidity Index (CCI), which is a weighted index of specific comorbidities that were identified using *ICD-10* codes [32]. CCI scores were divided into 3 categories (0, 1-2, and ≥3). Similarly, we calculated each patient's Hospital Frailty Risk Score (HFRS), which was developed to identify older adults experiencing frailty with a higher risk of adverse outcomes [33]. The total HFRS ranges from 0 to 99 and was divided into 2 categories (<5 and ≥5). To determine baseline functional dependence and incontinence levels in patients at admission, we analyzed the following 2 variables: dependence in ≥1 ADL items (BI components of bathing, dressing, toileting, transfers, mobility, and feeding) at admission and urinary and fecal incontinence (BI components of bladder function and bowel function) at admission. We dichotomized each of these 2 variables into independent (ie, scoring the maximum score on all components) or dependent (all other patients). We also determined each patient's location before admission (home or LTCF) and the surgical treatment received during the index hospitalization.

XSL•FO
RenderX

Statistical Analysis

The chi-square test was used to compare the differences in patient characteristics among the 3 HAC-OP-DPC categories. We generated cross-tabulations to examine the co-occurrences of each complication. Pearson correlation coefficients were calculated to measure the associations between each complication.

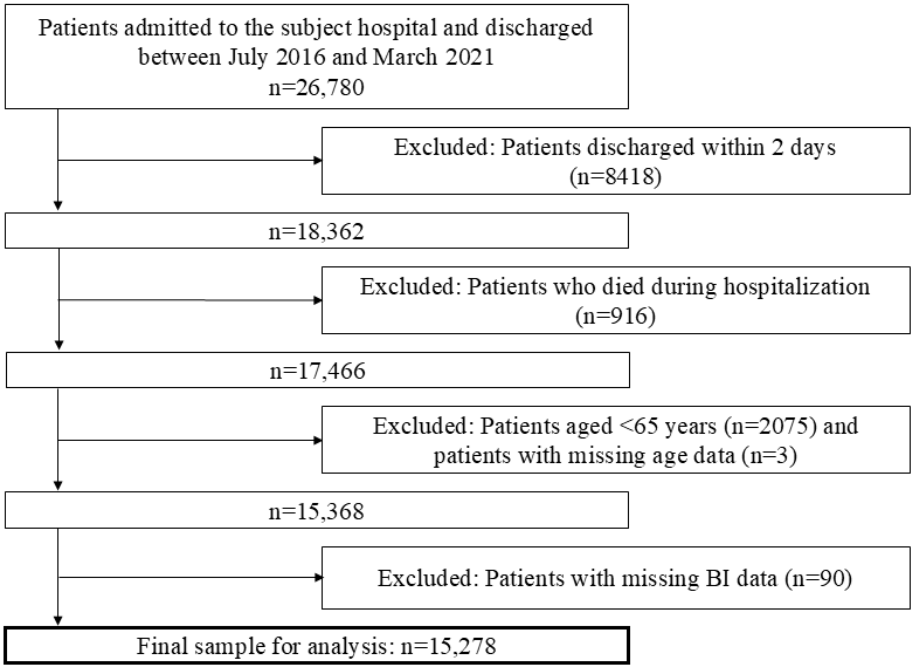
The associations between HAC-OP-DPC and LOS were examined using multivariable generalized linear regression models for gamma-distributed data with a log-link function that adjusted for all covariates. Effect sizes for 1 HAC-OP-DPC and ≥2 HAC-OP-DPC were quantified using adjusted risk ratios (ARRs) and their 95% CIs, which indicated the likelihood of having a longer LOS. Next, the associations of HAC-OP-DPC with discharge to other hospitals and discharge to LTCFs were examined using multivariable logistic regression models that adjusted for all covariates. Effect sizes for 1 HAC-OP-DPC and ≥2 HAC-OP-DPC were quantified using adjusted odds ratios (AORs) and their 95% CIs, which indicated the odds of being discharged to another hospital or LTCF. In addition to analyzing the associations between the number of HAC-OP-DPC and the 3 outcomes, we constructed models to examine the associations between individual HAC-OP-DPC components and the outcomes. For the analysis of hospital-associated functional decline, we excluded patients who were already dependent in all the BI components at admission because they could not experience any further functional decline. Similarly, for the analysis of hospital-associated incontinence, we excluded patients who were already dependent in both bladder function and bowel function at admission because they could not experience new-onset incontinence during hospitalization. We also conducted several sensitivity analyses. First, we re-ran the primary analysis for each of the 3 outcomes, excluding patients who could not experience functional decline (fully dependent on admission) or hospital-associated incontinence (incontinent on admission). Second, we recalculated the HAC-OP-DPC composite outcome excluding cases of delirium that were identified only by drug prescribing as we recognized that antipsychotic drugs might be prescribed for other indications (eg, behavioral and psychological symptoms of dementia, psychosis). All analyses were conducted using SPSS (version 28.0; IBM Corp), and the cross-tabulations were created using R (version 4.3.2; R project for Statistical Computing). *P* values (2-tailed) below .05 were considered statistically significant.

Results

Study Participant Selection

We first identified 26,780 candidate patients who were admitted to the subject hospital and discharged to home, another hospital, or an LTCF during the study period (Figure 1). After applying the exclusion criteria, the final sample for analysis consisted of 15,278 patients aged ≥65 years. Multimedia Appendix 2 shows the characteristics among 833 patients ≥65 years who died during hospitalization and without missing data at admission. These excluded participants were aged ≥85 years, had ≥2 CCI, were more likely to have ADL dependency and incontinence, and were more likely to have cancer, pneumonia, or heart failure.

Figure 1. Flowchart of patient selection. BI: Barthel Index.



Characteristics of Study Patients

Their mean age was 81.2 (SD 7.9) years, and women accounted for 55.4% of all patients (Table 1). There were 2877 (18.8%) patients who experienced 1 HAC-OP-DPC and 733 (4.8%) patients who experienced ≥2 HAC-OP-DPC during the index hospitalization.

XSL•FO
RenderX

Table 1. Patient characteristics according to the number of HAC-OP-DPC[a].

Characteristics	Total (N=15,278)		HAC-OP-DPC						*P* value
			None (n=11,668)		1% (n=2877)		≥2% (n=733)		
	N	%	n	%	n	%	n	%	
Sex									.16
Men	6815	44.6	5183	44.4	1280	44.5	352	48.0	
Women	8463	55.4	6485	55.6	1597	55.5	381	52.0	
Age (years)									<.001
65-74	3353	21.9	2771	23.7	508	17.7	74	10.1	
75-84	6396	41.9	5009	42.9	1117	38.8	270	36.8	
≥85	5529	36.2	3888	33.3	1252	43.5	389	53.1	
Annual household income (¥; ¥ 1=US $0.0092)									<.001
<3.7 million	11,600	75.9	8759	75.1	2230	77.5	611	83.4	
≥3.7 million	1570	10.3	1195	10.2	308	10.7	67	9.1	
Unknown	2108	13.8	1714	14.7	339	11.8	55	7.5	
Primary diagnosis for admission									<.001
Musculoskeletal diseases	1127	7.4	890	7.6	203	7.1	34	4.6	
Coronary heart disease	338	2.2	288	2.5	47	1.6	3	0.4	
Congestive heart failure	689	4.5	505	4.3	133	4.6	51	7.0	
Cerebrovascular disease	1039	6.8	874	7.5	133	4.6	32	4.4	
Pneumonia or acute bronchitis	1323	8.7	974	8.3	274	9.5	75	10.2	
Fracture	576	3.8	407	3.5	137	4.8	32	4.4	
Metabolic diseases	972	6.4	735	6.3	186	6.5	51	7.0	
Renal diseases	879	5.8	646	5.5	190	6.6	43	5.9	
Neurological diseases	631	4.1	510	4.4	100	3.5	21	2.9	
Gastrointestinal diseases	1694	11.1	1272	10.9	352	12.2	70	9.5	
Cancer	1867	12.2	1344	11.5	400	13.9	123	16.8	
Other	4143	27.1	3223	27.6	722	25.1	198	27.0	
CCI[b]									<.001
0	9041	59.2	7116	61.0	1563	54.3	362	49.4	
1-2	4852	31.8	3578	30.7	1000	34.8	274	37.4	
≥3	1385	9.1	974	8.3	314	10.9	97	13.2	
HFRS[c]									<.001
<5	13,278	86.9	10,324	88.5	2371	82.4	583	79.5	
≥5	2000	13.1	1344	11.5	506	17.6	150	20.5	
Dependence in ≥1 ADL[d] items at admission	12,658	82.9	9764	83.7	2239	77.8	655	89.4	<.001
Urinary and fecal incontinence at admission	5158	33.8	3932	33.7	1038	36.1	188	25.6	<.001
Location before admission									<.001
Home	14,186	92.9	10,875	93.2	2642	91.8	669	91.3	
LTCF[e]	1092	7.1	793	6.8	235	8.2	64	8.7	
Surgical treatment	3337	21.8	2444	20.9	732	25.4	161	22.0	<.001

[a]HAC-OP-DPC: hospital-associated complications of older people-Diagnosis Procedure Combination data version.

XSL•FO
RenderX

[b]CCI: Charlson Comorbidity Index.
[c]HFRS: Hospital Frailty Risk Score.
[d]ADL: activities of daily living.
[e]LTCF: long-term care facility.

HAC-OP-DPC Among This Study's Patients

The most common complication was functional decline (n=2103, 13.8%), followed by delirium (n=1345, 8.8%: recorded diagnosis of delirium (n=59); recorded prescriptions of drugs for delirium, n=1286), new incontinence (n=860, 5.6%), pressure injury (n=104, 0.7%), and fall (n=59, 0.4%). Figure 2 demonstrates the patterns of co-occurrence amongst HAC-OP-DPC. For example, among patients with functional decline, 20.6% also experienced incontinence during hospitalization, while 50.5% of patients with incontinence, 33.9% of patients with falls, and 22.5% of patients with delirium also experienced functional decline during hospitalization. While there were statistically significant correlations between most complications, the strength of these correlations was weak (maximum correlation coefficient of 0.26; Table 2). As shown in Table 1, older age, annual household income, multimorbidity, frailty, baseline functional impairment and incontinence, and living in LTCF were all significantly associated with HAC-OP-DPC. Patients with ≥2 HAC-OP-DPC had a significantly greater proportion of patients with congestive heart failure, pneumonia or acute bronchitis, metabolic diseases, and cancer as the principal diagnosis.

Figure 2. Co-occurrence of each HAC-OP-DPC among older patients aged ≥65 years (N=15,278). Values show the percentages of patients in each row who also had the column condition during hospitalization. HAC-OP-DPC: hospital-associated complications of older people-Diagnosis Procedure Combination data version.



Table 2. Within-patient correlations between each HAC-OP-DPC[a] among older patients aged ≥65 years (N=15,278). Values show Pearson correlation coefficients (r) between each HAC-OP-DPC.

	Hospital-associated functional decline	Hospital-associated incontinence	Hospital-associated delirium	Hospital-associated pressure injury	Hospital-associated fall
Hospital-associated functional decline	1.000	0.260[b]	0.079[b]	–0.008	0.036[b]
Hospital-associated incontinence	—	1.000	0.099[b]	0.021[b]	0.040[b]
Hospital-associated delirium	—	—	1.000	0.016[c]	0.025[b]
Hospital-associated pressure injury	—	—	—	1.000	0.008
Hospital-associated fall	—	—	—	—	1.000

[a]HAC-OP-DPC: hospital-associated complications of older people-Diagnosis Procedure Combination data version.
[b]*P*<.01.
[c]*P*<.05.

XSL•FO
RenderX

Association of HAC-OP-DPC With Outcome Measures

Table 3 demonstrates the significant and graded association observed between HAC-OP-DPC and outcomes. Adjusting for all covariates, those with 1 and ≥2 HAC-OP-DPC were significantly more likely to have a longer LOS during the index hospitalization (ARR 1.29, 95% CI 1.25-1.33, and ARR 1.97, 95% CI 1.87-2.08, respectively) compared to those with no HAC-OP-DPC. They were significantly more likely to be discharged to other hospitals (AOR 2.36, 95% CI 2.10-2.65, and AOR 6.96, 95% CI 5.81-8.35, respectively), and be discharged to LTCFs (AOR 1.35, 95% CI 1.09-1.67, and AOR 1.68, 95% CI 1.18-2.39, respectively). The analyses of individual HAC-OP-DPC showed that patients who experienced hospital-associated functional decline, incontinence, delirium, pressure injury, and falls were significantly more likely to have a longer LOS and be discharged to other hospitals than patients without these complications. However, only delirium was significantly associated in this individual complication analysis with discharge to LCTFs. A sensitivity analysis (n=11,075) that excluded 4203 patients who could neither experience hospital-associated functional decline nor incontinence found that HAC-OP-DPC still had a significant and graded association with longer LOS and discharge destination (Multimedia Appendix 3), similar to the results of the main analysis. Moreover, the associations of the number of HAC-OP-DPC using delirium identified from a recorded diagnosis alone with these outcomes were similar to the results of the main analysis (Multimedia Appendix 4).

XSL•FO
RenderX

Table 3. Association of HAC-OP-DPCa with outcome measures in 15,278 participants.

	Participants, n	LOS[b]				Discharge to other hospitals			Discharge to LTCFs[c]		
		Median	IQR	RR[d] (95% CI)[e]	ARR[f] (95% CI)[g]	%	OR[h] (95% CI)[i]	AOR[j] (95% CI)[k]	%	OR (95% CI)[i]	AOR (95% CI)[k]
HAC-OP-DPC											
No (Ref[l])	11,482	14	(7-25)	1.00	1.00	14.4	1.00	1.00	7.3	1.00	1.00
1	2,838	18	(9.5-32)	1.28 (1.25-1.32)	1.29 (1.25-1.33)	24.6	1.94 (1.76-2.14)	2.36 (2.10-2.65)	9.7	1.42 (1.24-1.64)	1.35 (1.09-1.67)
≥2	724	29	(14-44)	1.88 (1.78-1.99)	1.97 (1.87-2.08)	42.6	4.40 (3.76-5.14)	6.96 (5.81-8.35)	11.1	1.76 (1.38-2.23)	1.68 (1.18-2.39)
Hospital-associated functional decline[m]											
No (Ref)	10,124	13	(7-24)	1.00	1.00	10.4	1.00	1.00	4.3	1.00	1.00
Yes	2,103	16	(8-31)	1.35 (1.30-1.40)	1.35 (1.30-1.40)	23.6	2.67 (2.37-3.00)	3.56 (3.11-4.09)	7.3	1.77 (1.46-2.14)	1.09 (0.81-1.47)
Hospital-associated incontinence[n]											
No (Ref)	10,615	12	(7-22)	1.00	1.00	8.6	1.00	1.00	3.4	1.00	1.00
Yes	860	27	(14-42)	1.93 (1.83-2.04)	1.80 (1.71-1.9)	43.3	8.12 (6.98-9.44)	7.33 (6.16-8.73)	13.3	4.39 (3.51-5.49)	1.24 (0.84-1.83)
Hospital-associated delirium											
No (Ref)	13,933	14	(8-26)	1.00	1.00	16.1	1.00	1.00	7.6	1.00	1.00
Yes	1,345	25	(14-40.5)	1.61 (1.54-1.68)	1.53 (1.47-1.59)	34.1	2.70 (2.39-3.05)	2.23 (1.94-2.56)	12.3	1.71 (1.44-2.04)	1.61 (1.24-2.10)
Hospital-associated pressure injury											
No (Ref)	15,174	15	(8-27)	1.00	1.00	17.5	1.00	1.00	7.9	1.00	1.00
Yes	104	28	(16.25-42.75)	1.68 (1.45-1.95)	1.43 (1.24-1.64)	48.1	4.37 (2.97-6.43)	2.78 (1.83-4.22)	17.3	2.44 (1.46-4.07)	0.69 (0.31-1.56)
Hospital-associated fall											
No (Ref)	15,219	15	(8-27)	1.00	1.00	17.6	1.00	1.00	8.0	1.00	1.00
Yes	59	38	(22-50)	2.05 (1.68-2.49)	2.02 (1.68-2.43)	49.2	4.53 (2.72-7.56)	4.65 (2.63-8.23)	5.1	0.62 (0.19-1.98)	0.68 (0.15-3.19)

[a]HAC-OP-DPC: hospital-associated complications of older people-Diagnosis Procedure Combination data version.

[b]LOS: length of stay.

[c]LTCF: long-term care facility.

[d]RR: risk ratio.

[e]Generalized linear regression analysis.

[f]ARR: adjusted risk ratio.

[g]Generalized linear regression analysis that adjusted for all covariates (sex, age group, annual household income, primary diagnosis for admission, Charlson Comorbidity Index score, Hospital Frailty Risk Score, dependence in ≥1 activities of daily living items at admission, urinary and fecal incontinence at admission, location before admission, and surgical treatment).

[h]OR: odds ratio.

[i]Logistic regression analysis.

[j]AOR: adjusted odds ratio.

[k]Logistic regression analysis that adjusted for all covariates (sex, age group, annual household income, primary diagnosis for admission, Charlson Comorbidity Index score, Hospital Frailty Risk Score, dependence in ≥1 activities of daily living items at admission, urinary and fecal incontinence at

XSL•FO
RenderX

admission, location before admission, and surgical treatment).

[l]Ref: reference.

[m]Models for functional decline excluded participants with pre-exiting full dependence (model n=12,227).

[n]Models for hospital-associated incontinence excluded those with pre-existing incontinence (model n=11,475).

Discussion

Principal Findings

This retrospective cohort study is the first to develop and apply a tool to assess HAC-OP from routinely collected administrative data and evaluate its predictive validity for hospital outcomes. Our analysis showed that almost one quarter of older inpatients with multiday stays of more than 2 days had a coded HAC-OP, and that having one or more HAC-OP-DPC was associated with longer LOS and discharge to other hospitals and LTCFs. The clinical implementation of the HAC-OP-DPC measure could support comparative analyses of clinical and policy interventions aimed at preventing these complications, thereby contributing to the optimization of acute care for older adults.

Almost 1 in 4 older patients had coding documentation of any HAC-OP-DPC in the present study, which was approximately half that of the incidence of any HAC-OP described in the Australian study [1]. Nonetheless, the incidences of each HAC-OP-DPC were ranked in similar order, with functional decline and delirium being most common. A large contribution to the disparity in incidence is likely to be reliance on DPC data to identify individual complications as HAC-OP-DPC, while the HAC-OP study used repeated patient and clinical record assessments by trained research assistants [1,17]. Although we made efforts to minimize the underestimation of delirium, falls, and pressure injuries by including additional codes and data sources in their identification criteria, it is very likely that the coding data underestimated the incidences of all complications. For example, there were no coded falls without fracture, and only 59 cases of direct recording of a delirium diagnosis. There are recognized gaps in clinician recognition and documentation of hospital complications as well as translation into coding [27]. Understanding and improving the accuracy and usability of HAC-OP-DPC may require correlation with clinical data and comparisons between sites. There are other reasons that our incidence estimates may have been lower than expected. The denominator in our main analysis included 3051 patients who could not experience further functional decline and 3803 patients who could not experience new-onset incontinence, which may have led to an underestimation of these HAC-OP-DPC. Our analyses also excluded patients who died during their inpatient stay, who may have had a higher rate of HAC-OP.

Nonetheless, our study found that the HAC-OP-DPC measure was associated with longer LOS and discharge destination, thereby demonstrating its predictive validity for outcomes in an acute care setting. Furthermore, our analysis found nonoverlapping risk estimates for LOS and discharge to other hospitals between patients with 1 and ≥2 HAC-OP-DPC, showing a significant exposure-outcome effect of a graded nature. These findings were consistent with the results of the original HAC-OP study [1] and suggest that prevention of further HAC-OP is important in those who have already acquired 1 complication. Our observations that HAC-OP-DPC was significantly associated with older age and baseline function are consistent with previous reports [1]. Importantly, our study is the first to demonstrate the strong associations of HAC-OP with higher CCI scores and HFRS, consistent with existing knowledge that comorbidities and frailty are risk factors for the individual complications included in the composite measure [13,16,34-37].

Our analysis also confirmed that the individual HAC-OP-DPC were associated with longer LOS and discharge to other hospitals, which was congruent with previous studies [6,14,35-39]. All HAC-OP-DPC have been individually recognized as important outcomes in older patients [1,6,14,17,36-42]. Although our study showed that there were patterns of co-occurrence and significant correlations among the individual HAC-OP-DPC, these correlations were weak. This suggests that these complications represent relatively distinct conditions and that they can be treated individually [1]. Moreover, a systematic review of composite outcomes in clinical trials proposed that studies should list results for all components of a composite outcome to avoid confusion and bias [43], and an outcome study using the original multicomponent HAC-OP measure demonstrated significant and clinically important improvements in individual outcomes but not in the composite measure [17]. We recommend that future studies should report the effects of interventions not only on the HAC-OP-DPC measure as a composite outcome but also on its individual complications.

Strengths and Limitations

Strengths of this study include a large, representative dataset with high levels of item completeness (including functional variables), adjustment for important covariates, and use of sensitivity analyses to explore data assumptions. We also recognize several limitations. First, we included additional information to reduce these anticipated underestimates that may have reduced precision; for example, by including drug prescribing of antipsychotics in the delirium diagnosis, we may have included some patients with other indications such as behavioral and psychological symptoms of dementia. However, our sensitivity analysis using a more stringent definition of delirium suggests that this had minimal impact on our overall findings. Second, our study was conducted in a single acute care hospital in Japan, and its findings may not be generalizable to other hospitals, regions, or countries. Nevertheless, the majority of acute care hospitals in Japan have adopted the DPC system [20], and future studies could compare the incidence of HAC-OP-DPC in each acute hospital throughout Japan. Also, although a previous study identified the validity of chronic disease diagnoses in DPC data as being generally high [26], we recognize that the performance of comorbidity scoring methods based on administrative data may vary between health systems [44]. Third, we could not infer a causal relationship between HAC-OP-DPC and LOS because we were unable to identify the date when each complication occurred, and we cannot

XSL•FO
RenderX

exclude reverse causality between HAC-OP-DPC and LOS [1]. Fourth, our study did not examine the association between HAC-OP-DPC and mortality, which had been analyzed in the original HAC-OP study [1], because there are no BI scores assigned on discharge for in-hospital deaths to calculate functional decline and incontinence of HAC-OP-DPC, and DPC data lack information on death after discharge. Future studies could link DPC data with mortality data to examine the associations between HAC-OP-DPC and mortality. Finally, our database lacked information on potentially important confounders, such as residential status (eg, living alone or with others), the presence of caregivers, and disease severity that could be included in future studies.

Conclusions

This study showed that almost one quarter of older acute care inpatients in a Japanese hospital have coding indicating a HAC-OP and demonstrated the predictive validity of the HAC-OP-DPC measure for longer LOS and discharge to other hospitals and LTCFs. To attain a more robust understanding of these relationships, additional studies are needed to verify our findings in other hospitals and regions. The clinical implementation of HAC-OP-DPC, which are identified using routinely collected administrative data, could support the efficient evaluation of integrated interventions aimed at optimizing inpatient care for older adults.

Acknowledgments

This study was supported by KAKENHI Grants-in-Aid for Scientific Research (grant numbers 20H03924 and 24K13306) from the Japan Society for the Promotion of Science.

Conflicts of Interest

None declared.

Multimedia Appendix 1
ICD-10 (International Statistical Classification of Diseases, Tenth Revision) codes for identifying hospital-associated complications of older people-Diagnosis Procedure Combination data version.
[DOCX File , 34 KB - aging_v8i1e68267_app1.docx]

Multimedia Appendix 2
Characteristics among patients who died during hospitalization (n=833).
[DOCX File , 36 KB - aging_v8i1e68267_app2.docx]

Multimedia Appendix 3
Association of HAC-OP-DPC (hospital-associated complications of older people-Diagnosis Procedure Combination data version) with the outcome measures after excluding patients who could neither experience hospital-associated functional decline nor incontinence (n=11,075).
[DOCX File , 36 KB - aging_v8i1e68267_app3.docx]

Multimedia Appendix 4
Association of HAC-OP-DPC (hospital-associated complications of older people-Diagnosis Procedure Combination data version) with the outcome measures after excluding patients who could neither experience hospital-associated functional decline nor incontinence (N=15,278).
[DOCX File , 36 KB - aging_v8i1e68267_app4.docx]

References

1. Mudge AM, McRae P, Hubbard RE, Peel NM, Lim WK, Barnett AG, et al. Hospital-associated complications of older people: a proposed multicomponent outcome for acute care. J Am Geriatr Soc 2019 Feb;67(2):352-356 [FREE Full text] [doi: 10.1111/jgs.15662] [Medline: 30423197]
2. Loyd C, Markland AD, Zhang Y, Fowler M, Harper S, Wright NC, et al. Prevalence of hospital-associated disability in older adults: a meta-analysis. J Am Med Dir Assoc 2020;21(4):455-461.e5 [FREE Full text] [doi: 10.1016/j.jamda.2019.09.015] [Medline: 31734122]
3. Boyd CM, Ricks M, Fried LP, Guralnik JM, Xue Q, Xia J, et al. Functional decline and recovery of activities of daily living in hospitalized, disabled older women: the women's health and aging study I. J Am Geriatr Soc 2009;57(10):1757-1766 [FREE Full text] [doi: 10.1111/j.1532-5415.2009.02455.x] [Medline: 19694869]
4. Berian JR, Mohanty S, Ko CY, Rosenthal RA, Robinson TN. Association of loss of independence with readmission and death after discharge in older patients after surgical procedures. JAMA Surg 2016;151(9):e161689. [doi: 10.1001/jamasurg.2016.1689] [Medline: 27409710]

XSL•FO
RenderX

5. Gibb K, Seeley A, Quinn T, Siddiqi N, Shenkin S, Rockwood K, et al. The consistent burden in published estimates of delirium occurrence in medical inpatients over four decades: a systematic review and meta-analysis study. Age Ageing 2020;49(3):352-360 [FREE Full text] [doi: 10.1093/ageing/afaa040] [Medline: 32239173]
6. Geriatric Medicine Research Collaborative. Delirium is prevalent in older hospital inpatients and associated with adverse outcomes: results of a prospective multi-centre study on world delirium awareness day. BMC Med 2019;17(1):229 [FREE Full text] [doi: 10.1186/s12916-019-1458-7] [Medline: 31837711]
7. Siddiqi N, House A, Holmes J. Occurrence and outcome of delirium in medical in-patients: a systematic literature review. Age Ageing 2006;35(4):350-364. [doi: 10.1093/ageing/afl005] [Medline: 16648149]
8. Milbrandt EB, Deppen S, Harrison PL, Shintani AK, Speroff T, Stiles RA, et al. Costs associated with delirium in mechanically ventilated patients. Crit Care Med 2004;32(4):955-962. [doi: 10.1097/01.ccm.0000119429.16055.92] [Medline: 15071384]
9. Leslie DL, Marcantonio ER, Zhang Y, Leo-Summers L, Inouye SK. One-year health care costs associated with delirium in the elderly population. Arch Intern Med 2008;168(1):27-32 [FREE Full text] [doi: 10.1001/archinternmed.2007.4] [Medline: 18195192]
10. Cohen Y, Zisberg A, Chayat Y, Gur-Yaish N, Gil E, Levin C, et al. Walking for better outcomes and recovery: the effect of WALK-FOR in preventing hospital-associated functional decline among older adults. J Gerontol A Biol Sci Med Sci 2019;74(10):1664-1670. [doi: 10.1093/gerona/glz025] [Medline: 30726886]
11. Clinical guidelines, No. 103. Delirium: prevention, diagnosis and management in hospital and long-term care. National Institute for Health and Care Excellence (NICE). London: (NICE) URL: https://www.nice.org.uk/guidance/cg103 [accessed 2024-03-18]
12. Hshieh TT, Yue J, Oh E, Puelle M, Dowal S, Travison T, et al. Effectiveness of multicomponent nonpharmacological delirium interventions: a meta-analysis. JAMA Intern Med 2015;175(4):512-520 [FREE Full text] [doi: 10.1001/jamainternmed.2014.7779] [Medline: 25643002]
13. Geyskens L, Jeuris A, Deschodt M, Van Grootven B, Gielen E, Flamaing J. Patient-related risk factors for in-hospital functional decline in older adults: a systematic review and meta-analysis. Age Ageing 2022;51(2):afac007. [doi: 10.1093/ageing/afac007] [Medline: 35165688]
14. Ahmed S, Leurent B, Sampson EL. Risk factors for incident delirium among older people in acute hospital medical units: a systematic review and meta-analysis. Age Ageing 2014;43(3):326-333 [FREE Full text] [doi: 10.1093/ageing/afu022] [Medline: 24610863]
15. Deandrea S, Bravi F, Turati F, Lucenteforte E, La Vecchia C, Negri E. Risk factors for falls in older people in nursing homes and hospitals. A systematic review and meta-analysis. Arch Gerontol Geriatr 2013;56(3):407-415. [doi: 10.1016/j.archger.2012.12.006] [Medline: 23294998]
16. Palmer MH, Baumgarten M, Langenberg P, Carson JL. Risk factors for hospital-acquired incontinence in elderly female hip fracture patients. J Gerontol A Biol Sci Med Sci 2002;57(10):M672-M677. [doi: 10.1093/gerona/57.10.m672] [Medline: 12242323]
17. Mudge AM, McRae P, Banks M, Blackberry I, Barrimore S, Endacott J, et al. Effect of a ward-based program on hospital-associated complications and length of stay for older inpatients: the cluster randomized CHERISH trial. JAMA Intern Med 2022;182(3):274-282 [FREE Full text] [doi: 10.1001/jamainternmed.2021.7556] [Medline: 35006265]
18. Daniels H, Hollinghurst J, Fry R, Clegg A, Hillcoat-Nallétamby S, Nikolova S, et al. The value of routinely collected data in evaluating home assessment and modification interventions to prevent falls in older people: systematic literature review. JMIR Aging 2021;4(2):e24728 [FREE Full text] [doi: 10.2196/24728] [Medline: 33890864]
19. Han E, Kharrazi H, Shi L. Identifying predictors of nursing home admission by using electronic health records and administrative data: scoping review. JMIR Aging 2023;6:e42437 [FREE Full text] [doi: 10.2196/42437] [Medline: 37990815]
20. Yasunaga H. Real world data in Japan: chapter II the diagnosis procedure combination database. ACE 2019;1(3):76-79. [doi: 10.37737/ace.1.3_76]
21. Mitsutake S, Ishizaki T, Yano S, Tsuchiya-Ito R, Uda K, Toba K, et al. All-cause readmission or potentially avoidable readmission: which is more predictable using frailty, comorbidities, and ADL? Innov Aging 2023;7(5):igad043 [FREE Full text] [doi: 10.1093/geroni/igad043] [Medline: 37342490]
22. Uda K, Matsui H, Fushimi K, Yasunaga H. Intensive in-hospital rehabilitation after hip fracture surgery and activities of daily living in patients with dementia: retrospective analysis of a nationwide inpatient database. Arch Phys Med Rehabil 2019;100(12):2301-2307. [doi: 10.1016/j.apmr.2019.06.019] [Medline: 31421098]
23. Health survey on impact assessment of DPC introduction: summary of statistics. Ministry of Health, Labour and Welfare. URL: https://www.mhlw.go.jp/stf/shingi2/0000196043_00006.html; [accessed 2024-10-02]
24. Hospital-acquired complications (HACs). Australian Commission on Safety and Quality in Health Care. URL: https://www.safetyandquality.gov.au/our-work/indicators-measurement-and-reporting/hospital-acquired-complications-hacs [accessed 2024-10-31]
25. Department of health and human services. Patient safety indicators (PSI) benchmark data tables. Agency for Healthcare Research and Quality, U.S. URL: https://qualityindicators.ahrq.gov/Downloads/Modules/PSI/V2024/Version_2024_Benchmark_Tables_PSI.pdf [accessed 2024-10-31]

XSL•FO
RenderX

26. Yamana H, Moriwaki M, Horiguchi H, Kodan M, Fushimi K, Yasunaga H. Validity of diagnoses, procedures, and laboratory data in Japanese administrative data. J Epidemiol 2017;27(10):476-482 [FREE Full text] [doi: 10.1016/j.je.2016.09.009] [Medline: 28142051]
27. Ibitoye T, So S, Shenkin S, Anand A, Reed MJ, Vardy ERLC, et al. Delirium is under-reported in discharge summaries and in hospital administrative systems: a systematic review. Delirium (Bielef) 2023;2023:74541. [doi: 10.56392/001c.74541] [Medline: 39654697]
28. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. Md State Med J 1965;14:61-65. [doi: 10.1037/t02366-000]
29. Exchange rates. Organization for economic co-operation and development. URL: https://www.oecd.org/en/data/indicators/exchange-rates.html [accessed 2024-03-18]
30. Mitsutake S, Ishizaki T, Tsuchiya-Ito R, Furuta K, Hatakeyama A, Sugiyama M, et al. Association of cognitive impairment severity with potentially avoidable readmissions: a retrospective cohort study of 8897 older patients. Alzheimers Dement (Amst) 2021;13(1):e12147 [FREE Full text] [doi: 10.1002/dad2.12147] [Medline: 33816752]
31. Mitsutake S, Ishizaki T, Tsuchiya-Ito R, Furuta K, Hatakeyama A, Sugiyama M, et al. Association of cognitive impairment severity with potentially avoidable readmissions: a retrospective cohort study of 8897 older patients. Alzheimers Dement (Amst) 2021;13(1):e12147 [FREE Full text] [doi: 10.1002/dad2.12147] [Medline: 33816752]
32. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi J, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care 2005;43(11):1130-1139. [doi: 10.1097/01.mlr.0000182534.19832.83] [Medline: 16224307]
33. Gilbert T, Neuburger J, Kraindler J, Keeble E, Smith P, Ariti C, et al. Development and validation of a hospital frailty risk score focusing on older people in acute care settings using electronic hospital records: an observational study. Lancet 2018;391(10132):1775-1782 [FREE Full text] [doi: 10.1016/S0140-6736(18)30668-8] [Medline: 29706364]
34. Tran NN, Hoang TPN, Ho TKT. Diagnosis and risk factors for delirium in elderly patients in the emergency rooms and intensive care unit of the national geriatric hospital emergency department: a cross-sectional observational study. Int J Gen Med 2021;14:6505-6515 [FREE Full text] [doi: 10.2147/IJGM.S325365] [Medline: 34675618]
35. Raats JW, Steunenberg SL, de Lange DC, van der Laan L. Risk factors of post-operative delirium after elective vascular surgery in the elderly: a systematic review. Int J Surg 2016;35:1-6 [FREE Full text] [doi: 10.1016/j.ijsu.2016.09.001] [Medline: 27613124]
36. Hori K, Usuba K, Sakuyama A, Adachi Y, Hirakawa K, Nakayama A, et al. Hospitalization-associated disability after cardiac surgery in elderly patients - exploring the risk factors using machine learning algorithms. Circ Rep 2021;3(8):423-430 [FREE Full text] [doi: 10.1253/circrep.CR-21-0057] [Medline: 34414331]
37. Bates DW, Pruess K, Souney P, Platt R. Serious falls in hospitalized patients: correlates and resource utilization. Am J Med 1995;99(2):137-143. [doi: 10.1016/s0002-9343(99)80133-8] [Medline: 7625418]
38. Moghadamyeghaneh Z, Stamos MJ, Stewart L. Patient co-morbidity and functional status influence the occurrence of hospital acquired conditions more strongly than hospital factors. J Gastrointest Surg 2019;23(1):163-172. [doi: 10.1007/s11605-018-3957-9] [Medline: 30225796]
39. Campbell J, Hubbard R, Ostaszkiewicz J, Green T, Coyer F, Mudge A. Incontinence during and following hospitalisation: a prospective study of prevalence, incidence and association with clinical outcomes. Age Ageing 2023;52(9):afad181 [FREE Full text] [doi: 10.1093/ageing/afad181] [Medline: 37738169]
40. Montero-Marco J, Charlo-Bernardos M, Subirón-Valera AB, Erickson H, Herrero-Cortina B, Altarribas-Bolsa E. The role of nursing care continuity report in predicting length of hospital stay in older people: a retrospective cohort study. J Clin Nurs 2024;33(5):1830-1838. [doi: 10.1111/jocn.16953] [Medline: 38178555]
41. Brand CA, Sundararajan V. A 10-year cohort study of the burden and risk of in-hospital falls and fractures using routinely collected hospital data. Qual Saf Health Care 2010;19(6):e51. [doi: 10.1136/qshc.2009.038273] [Medline: 20558479]
42. Rubin EB, Buehler AE, Halpern SD. States worse than death among hospitalized patients with serious illnesses. JAMA Intern Med 2016;176(10):1557-1559 [FREE Full text] [doi: 10.1001/jamainternmed.2016.4362] [Medline: 27479808]
43. Cordoba G, Schwartz L, Woloshin S, Bae H, Gøtzsche PC. Definition, reporting, and interpretation of composite outcomes in clinical trials: systematic review. Br Med J 2010;341:c3920 [FREE Full text] [doi: 10.1136/bmj.c3920] [Medline: 20719825]
44. Sundararajan V, Quan H, Halfon P, Fushimi K, Luthi JC, Burnand B, International Methodology Consortium for Coded Health Information (IMECCHI). Cross-national comparative performance of three versions of the ICD-10 Charlson index. Med Care 2007;45(12):1210-1215. [doi: 10.1097/MLR.0b013e3181484347] [Medline: 18007172]

Abbreviations

ADL: activities of daily living
AOR: adjusted odds ratio
ARR: adjusted risk ratio
BI: Barthel Index
CCI: Charlson Comorbidity Index

XSL•FO
RenderX

DPC: Diagnosis Procedure Combination
ICD-10: International Statistical Classification of Diseases, Tenth Revision
HAC-OP: hospital-associated complications of older people
HAC-OP-DPC: hospital-associated complications of older people-Diagnosis Procedure Combination data version
HFRS: Hospital Frailty Risk Score
LOS: length of stay
LTCF: long-term care facility

Edited by M Gray; submitted 01.11.24; peer-reviewed by T Karen; comments to author 01.12.24; revised version received 14.12.24; accepted 02.01.25; published 06.02.25.

Please cite as:
Mitsutake S, Ishizaki T, Yano S, Hirata T, Ito K, Furuta K, Shimazaki Y, Ito H, Mudge A, Toba K
Predictive Validity of Hospital-Associated Complications of Older People Identified Using Diagnosis Procedure Combination Data From an Acute Care Hospital in Japan: Observational Study
JMIR Aging 2025;8:e68267
URL: https://aging.jmir.org/2025/1/e68267
doi:10.2196/68267
PMID:

©Seigo Mitsutake, Tatsuro Ishizaki, Shohei Yano, Takumi Hirata, Kae Ito, Ko Furuta, Yoshitomo Shimazaki, Hideki Ito, Alison Mudge, Kenji Toba. Originally published in JMIR Aging (https://aging.jmir.org), 06.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Older Perpetrators of Domestic Violence: Mixed-Effects Logistic Regression Analysis of Police Records

Sharon Reutens[1], MBBS, PhD; Emaediong Akpanekpo[2], MBBS, MPH, PhD; George Karystianis[2], BSc, MSc, PhD; Adrienne Withall[2], BSc, PhD; Tony Butler[2], BSc, MSc, PhD

[1]Faculty of Medicine, UNSW Sydney, Sydney, Australia

[2]UNSW Sydney, Sydney, Australia

Corresponding Author:
Sharon Reutens, MBBS, PhD
Faculty of Medicine
UNSW Sydney
223 Anzac Parade Kensington
Sydney, 2033
Australia
Phone: 61 93858765
Email: sreutens@unsw.edu.au

Abstract

Background: Domestic violence (DV) among older adults is an understudied area, often overlapping with abuse of older people, intimate partner violence, and behavioral and psychological symptoms of dementia.

Objective: This study aimed to examine the characteristics of older persons of interest—individuals suspected or charged with a DV offence—and survivors involved in police-attended DV events in New South Wales, Australia, and assess associations with physical and nonphysical abuse.

Methods: Police records of 10,708 DV events involving 8247 adults aged ≥55 years who were identified as persons of interest from 2005 to 2016 were analyzed using text mining. A 3-level mixed-effects logistic regression model was used to identify predictors of physical and nonphysical abuse.

Results: Physical abuse formed a greater proportion of all abuse committed by female persons of interest aged >65 years compared to female persons of interest aged between 55 and 64 years and male persons of interest; however, after stratified analysis, female persons of interest had similarly elevated odds of physical abuse perpetration to male persons of interest. Other factors associated with increased odds of perpetrating physical abuse were persons of interest with dementia and alcohol-related events. Dementia increased the odds of combined physical and nonphysical abuse. Substance use disorders increased the odds of events with combined physical and nonphysical abuse.

Conclusions: The findings of this study suggest that DV, including physical violence, is an important issue in later life. Alcohol is a situational factor, and dementia is associated with perpetration and exposure to violence. The study highlights the need for clinicians to evaluate the risk of violence and exposure to violence in patients with dementia and for policy interventions targeting alcohol and substance use in older adults.

(JMIR Aging 2025;8:e75993) doi:10.2196/75993

KEYWORDS

dementia; domestic violence; abuse of older people; abuse; police; text mining; older

Introduction

Domestic violence (DV) is a pervasive problem experienced by individuals across different ages, sexes, cultural backgrounds, and socioeconomic status. In Australia, the New South Wales Crimes (Domestic and Personal Violence) Act 2007 defines DV as any physical, sexual, financial, emotional or psychological abuse perpetrated within a current or past intimate partner and family relationship and includes DV against household residents and kin. DV is estimated to affect 20% of Australian people aged 18 years or older [1]. Survivors experience a wide range of adverse physical and mental health outcomes (eg, posttraumatic stress disorder, suicidal behavior, substance abuse, and injuries) or death as well as socioeconomic

XSL•FO
RenderX

impacts from controlling behaviors about finances and the risk of homelessness [2-4].

There is a paucity of literature on DV in older survivors and perpetrators that is further complicated because violence in this age group can fall into several categories such as abuse of older people, intimate partner violence (IPV), or, if a person has dementia, perceived as behavioral and psychological symptoms of dementia (BPSD). The World Health Organization defines IPV as "any behaviour within an intimate relationship that causes physical, psychological or sexual harm to those in the relationship" [5], while abuse of older people is seen as "a single or repeated act or lack of appropriate action, occurring within any relationship where there is an expectation of trust, which causes harm or distress to an older person" [4]. BPSD is "an umbrella term for a heterogenous group of non-cognitive symptoms that are almost ubiquitous in dementia" [6]. While the categories are useful for examining factors related to the relationship between perpetrator and survivor (abuse of older people and IPV) and problem behaviors in people with dementia (BPSD), these divisions have hampered the understanding of DV committed by older people.

Studies of DV with a focus on older people suggest that while physical abuse declines with age, emotional abuse does not [7-9]. However, if research excludes perpetrators and survivors living with dementia, the overall picture of violence perpetrated by older people will be incomplete. For instance, it is known that people with dementia are more likely to be abused if their care needs are greater [10,11], and aggression by people with dementia is common, with physical aggression occurring in 31% to 42% of cases [12].

Factors associated with DV perpetration by older people include male sex, depression, alcohol use, low income, functional impairment, and previous violence exposure [8,13]. DV survivor status is associated with several factors including female sex, low social support, and cognitive and physical impairments [14]. A history of trauma has been linked to both being abused and perpetration of violence [14,15]. These risk factors are similar to those studied by researchers of older person abuse. Associations with being a survivor of older person abuse include disability, poor physical and mental health, cognitive impairment, and female sex [10,16,17]. Perpetrators of older person abuse are most often offspring or spouses [16-18], have mental health problems including depression and substance misuse, and are dependent on those they abuse [17,19].

IPV is often viewed as gendered [18], but this might not be the case for older person abuse. While research suggests that most survivors of IPV and older person abuse are female [7,10], some studies suggest that older male individuals are at least equally, if not more likely to experience abuse [11,20,21], and one study found that female perpetration of violence increases with age [22]. Underreporting and underdetection of the abuse of older male individuals have been noted [21,23] and occur against a broader background of underreporting of older people's mistreatment, with estimates that 1 in 6 older people are mistreated, but only 1 in 24 cases are reported [10].

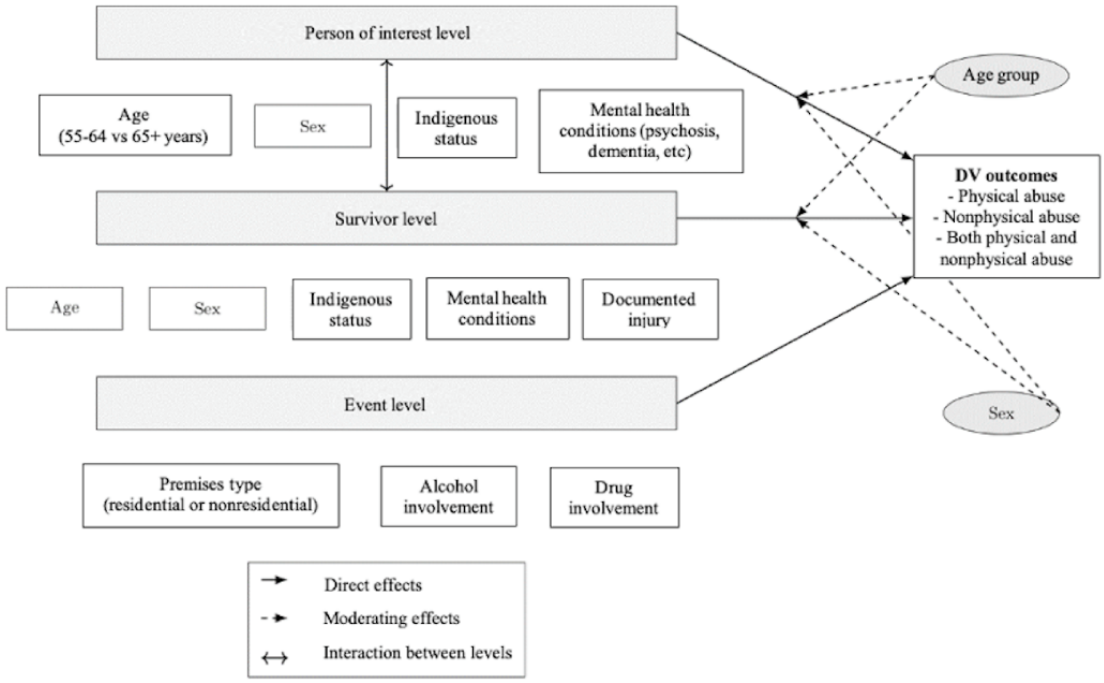
However, most of the data used for research in this area are obtained from survivor reports [9,13,24-26], which can be prone to recall bias. Thus, there is a need to explore other sources of potential information. Police records of attended DV events can provide an overview of DV by older people with the advantage of offering a relatively contemporaneous third-party perspective.

This study examined patterns of physical and nonphysical abuse among older adults using data from police-attended DV events in New South Wales. Our primary objectives were to describe the characteristics of DV events involving older persons of interest—individuals accused of or charged for perpetrating DV—and their survivors and identify potential risk factors in this understudied population. To guide our study, we developed a multilevel conceptual model (Figure 1) integrating elements from life course theory [27] and the ecological model of violence [28]. At the individual level, factors such as age-related cognitive changes affect impulse control and emotional regulation, which can lead to aggression [29]. Caregiver stress plays a role in perpetration when coping mechanisms are overwhelmed, for instance, managing behavioral symptoms of dementia [30]. The interplay between survivor and perpetrator is considered at the mesosystem level: the nature of the caregiver-recipient relationship can increase the risk of violence while feelings of threat and vulnerability can be expressed aggressively. Power dynamics can shift with aging, and expressions of control may change, for instance, from financial control to violence or from physical to verbal aggression [31]. Broader societal factors such as isolation and lack of social support for caregivers can increase violence [30], while macrosystem considerations include how violence by older people and those with cognitive impairment is construed by society, including the terminology used to describe it.

We structured the conceptual model as a 3-level hierarchical system including person of interest, survivor, and event levels. The person of interest level included characteristics of older perpetrators, including sociodemographic factors, mental health conditions (eg, dementia and psychosis), and developmental factors (eg, childhood onset disorders). The survivor level captured the dyadic context of DV, including survivor characteristics and vulnerability factors. The event level represented the immediate context of DV incidents, such as premises type and substance use involvement. We hypothesized that factors at each level exert both direct and indirect effects on DV outcomes (physical and nonphysical abuse). We hypothesized further that the perpetrator's age group and sex may function as a cross-level moderator, which may modify the relationships between predictors and outcomes across all levels.

XSL•FO
RenderX

Figure 1. Multilevel conceptual framework for domestic violence perpetration in older adults. DV: domestic violence.



Methods

Ethical Considerations

Ethics approval for this study was obtained from the Human Research Ethics Committee of the University of New South Wales (reference: HC16558). Permission to access deidentified domestic violence events was granted by the New South Wales Police. Participants were not recruited for this study.

Data

For the purpose of this study, an older person was defined as being aged 55 years and older based on previous DV research, which used this age as the lower age boundary [14,25,32]. The study population was taken from 492,393 police events recorded in the Computerized Operational Policing System (COPS)—a database maintained by the NSW Police Force that contains records of criminal events, including those specifically related to DV—from January 2005 to December 2016. Details regarding 64,587 events that involved either multiple persons of interest or survivors were excluded from the analysis, as it was not possible to deduce to whom the extracted mental health, survivor injury, and abuse type mention referred to (ie, if an event had 2 persons of interest and 1 extracted mental health mention for persons of interest, it is not known which person of interest the mention referred to), leaving 416,441 events for analysis, from which 221,995 events that documented the age of persons of interest were extracted. Of these, 10,708 events involved persons of interest aged 55 years and older, yielding 8247 individual persons of interest and 8629 survivors.

Each DV event in COPS has 2 primary formats: structured fields and unstructured event narratives. The structured fields contained numeric or categorical data relating to the DV event, whereas unstructured event narratives are free-text accounts that detail the circumstances surrounding the DV event. DV events were identified in COPS using specific structured fields: the offence classification being flagged as "domestic," associated factors set as "DV-related," and a documented domestic relationship between person of interest and survivor (including categories such as "spouse or partner," "ex-spouse or ex-partner," "boyfriend or girlfriend [including ex-boyfriend or ex-girlfriend, because relationship status was sometimes disputed]," "parent or guardian [including step or foster]," "child [including step or foster]," "sibling," "other member of family [including kin]," and "caregiver").

A text-mining approach was used to extract key information from the free-text narratives. This approach has been extensively described in a previous study [33,34]. Briefly, a set of rules based on common syntactical patterns was observed in a sample of 200 DV event narratives and identified mentions of mental illnesses, abuse types, and survivor injuries. Additionally, this involved manually constructing dictionaries containing relevant terms related to the DV event (eg, mental illness and anatomical parts). This approach identified 44 different abuse types, which included both physical and nonphysical abuse [33,34] (Multimedia Appendix 1).

The method was evaluated in a random sample of 100 narratives for the mental illness mentions with an average 92% (97.5% for persons of interest and 87.1% for survivors) precision (ie, the percentage of correctly identified mental illness mentions against the total number of identified mental illness mentions, a denominator that includes both true positives and false positives) and 90.2% and 85% precision for the abuse types and survivor injuries, respectively [33,34].

Outcome

The primary outcome of interest in this study was to describe DV in older people and any associations with the perpetration of physical and nonphysical abuse during a DV event. Physical abuse was broadly defined to include violent actions inflicted by a person of interest on a survivor (Multimedia Appendices

XSL•FO
RenderX

1 and 2). These included life-threatening actions (eg, strangulation, stabbing, and attempting to set fire to premises), actions causing significant injury (eg, hitting with an object or weapon, kicking, and punching), actions using objects as weapons, actions involving forceful restriction (eg, blocking, throwing, and physical restraint), actions causing pain and discomfort (eg, limb twisting, scratching, and chasing), and sexual assault. Nonphysical abuse was defined as emotional (eg, threats of self-harm), social (eg, social restriction), and verbal abuse (eg, yelling profanities). Breaches of Apprehended Domestic Violence Orders (ADVO)—an order made for the protection of a person against another person with whom they have or have had a domestic relationship—were included as a separate form of abuse. "Miscellaneous or other" forms of abuse include malicious damage and unspecified (in text) abuses (eg, "assault"). Physical and nonphysical abuses that occurred in the same event were treated as separate outcomes. Injury types were grouped into 10 categories: soft tissue injuries, nonspecific wounds, soreness, fracture, periorbital bruising, stab wound, bite mark, burn mark, broken tooth, and miscellaneous (Multimedia Appendix 3).

Explanatory Variables

Explanatory variables were classified into 3 levels based on the hierarchical nature of the data: event level (level 1), survivor level (level 2), and person of interest level (level 3). Each observation represented a unique DV event, characterized by a specific combination of persons of interest, survivor, and event date. A single survivor could have multiple events if abused by the same person of interest or by different persons of interest on different dates. A single person of interest could also be associated with multiple survivors on different dates. Individual abuse events (level 1) were nested within survivors (level 2), which were further nested within persons of interest (level 3).

Event-level variables included the following fixed field variables: premises type (residential or nonresidential), alcohol involvement in the event as assessed by police (yes or no), and drug involvement in the event as assessed by police (yes or no). Survivor-level variables encompassed age, sex (female or male), Indigenous status (non-Indigenous or Indigenous), documented injury (yes or no), and the presence or absence of psychosis, mood or anxiety disorder, personality disorder, and dementia. Person of interest–level variables included age (grouped into those aged 55-64 years and 65 years and older), sex (female or male), Indigenous status (non-Indigenous or Indigenous), and the presence or absence of psychosis, mood or anxiety disorder, personality disorder, dementia, and childhood-onset disorders, which included intellectual disability, tic disorders, and attention-deficit/hyperactivity disorder. Survivor-level and person of interest–level variables were obtained from the structured field data and the text-mined information.

Statistical Analysis

Characteristics of persons of interest and survivors were described using summary statistics. Categorical variables were reported as frequencies and percentages, while continuous variables were summarized using means, SDs, medians, and IQRs. Differences between persons of interest and survivors were assessed using Welch *t* test for age (due to unequal variances) and chi-square tests for categorical variables. The annual proportion of DV events from 2005 to 2016 was estimated, stratified by sex and age groups. For our analysis, DV events were classified into three mutually exclusive categories: (1) physical abuse only, (2) nonphysical abuse only, and (3) both physical and nonphysical abuse occurring in the same event. This classification was implemented to accurately estimate the prevalence and predictors of each distinct type of abuse and ensure that the analysis was not confounded by the co-occurrence of multiple forms of abuse within a single event.

A 3-level mixed-effects logistic regression model was used to identify predictors of physical and nonphysical abuse while accounting for the hierarchical structure of the data. The model accounted for the clustering of individual abuse events (level 1) within survivors (level 2) and survivors within persons of interest (level 3). Variables were selected for inclusion in the regression analyses based on a literature review linking them to DV [14,15,25,35]. These included age, sex, residential status, alcohol-related event, substance use disorders, cognitive impairments, mood or anxiety disorders, psychosis, and personality factors.

The primary focus of the analysis was on estimating the fixed effects, which represented the associations between event-level, survivor-level, and person of interest–level predictors and the odds of physical and nonphysical abuse. The mixed-effects logistic regression models were systematically built. First, an empty model with only random intercepts for survivors and persons of interest was fitted to assess the baseline variability in abuse outcomes attributable to differences across persons of interest and survivors. Second, event-level, survivor-level, and perpetrator-level factors were included in the multivariable model (full model) to determine adjusted estimates. The initial multivariable model included all variables of interest regardless of statistical significance from the univariable analyses. Variables with a *P* value of >.20 from the univariable analyses were individually dropped, and the Akaike information criterion was used to compare the initial multivariable model and the model without these variables. The model resulting in a smaller Akaike information criterion was considered as the better-fitting model. Adjusted odds ratios (aORs) with 95% CIs were reported as measures of association. Statistical significance was assessed using a 2-tailed *P* value threshold of .05. We examined whether the associations between explanatory variables and outcomes of interest differed significantly between person of interest age groups (55-64 years and 65 years and older) by conducting a stratified analysis by person of interest age groups.

Random effects parameters were estimated to quantify the variability in abuse outcomes attributable to the hierarchical structure of the data. The intraclass correlation coefficient was calculated to determine the proportion of the total variance in abuse events that was due to differences across persons of interest and survivors. The proportional change in variance was computed to assess the reduction in unexplained variance after adding predictors to the models. Multiple imputation by chained equations was used to address the missing data, which was determined to be missing at random. In total, 100 datasets were imputed. All analyses were conducted with Stata statistical software (version 18; StataCorp LLC).

XSL•FO
RenderX

Results

Characteristics of Police-Attended DV Events

Our analysis included 8247 persons of interest and 8629 survivors involved in 10,708 DV events (Multimedia Appendices 4 and 5). Male individuals constituted 81.4% (n=6710) of persons of interest, and female individuals comprised 73.9% (n=6375) of survivors (Table 1). This sex difference was consistent across age groups. Among persons of interest aged 55-64 years, 80.9% (n=4752) were male individuals, and 19% (n=1172) were female individuals. In the 65+-year age group, 82.5% (n=1958) of persons of interest were male individuals, and 17.5% (n=415) were female individuals. Persons of interest were typically older than survivors, with a median age of 60 (IQR 57-65) years for persons of interest compared to 50 (IQR 37-60) years for survivors.

Table 1. Characteristics of persons of interest and survivors of domestic violence events.

	Persons of interest (n=8247)	Survivors (n=8629)	*P* value[a]
Age (years)			<.001[b]
Mean (SD)	62 (7.0)	49 (16.9)	
Median (IQR)	60 (57-65)	50 (37-60)	
Sex, n (%)			<.001
Female	1532 (18.6)	6375 (73.9)	
Male	6710 (81.4)	2243 (26)	
Indigenous status, n (%)			.19
Indigenous	303 (3.7)	285 (3.3)	
Non-Indigenous	7944 (96.3)	8344 (96.7)	
Dementia, n (%)			
Present	191 (2.3)	104 (1.2)	<.001
Psychosis, n (%)			
Present	142 (1.7)	87 (1)	<.001
Mood or anxiety disorder			
Present	56 (0.7)	47 (0.5)	.13
Personality disorder, n (%)			
Present	28 (0.3)	13 (0.2)	.02
Childhood-onset disorder, n (%)			
Present	97 (1.2)	92 (1.1)	.51
Alcohol and substance use disorder, n (%)			
Present	200 (2.4)	65 (0.8)	<.001

[a]*P* values for categorical variables calculated using the chi-square test of independence.

[b]*P* value for age calculated using the Welch *t* test (*t* test for unequal variances) due to the large difference in SDs between the 2 groups.

Overall, persons of interest had higher proportions of mental health conditions compared with survivors. Dementia was observed in 2.3% (n=191) of persons of interest compared to 1.2% (n=104) of survivors (difference: n=87, 1.1%; *P*<.001). Psychosis showed a similar pattern, with a prevalence of 1.7% (n=142) in persons of interest versus 1% (n=87) in survivors (difference: n=55, 0.7%; *P*<.001). The largest difference was observed in alcohol and substance use disorders: 2.4% (n=200) of persons of interest had these disorders compared to 0.8% (n=65) of survivors (difference: n=135, 1.6%; *P*<.001). Personality disorders were slightly more common in persons of interest (n=28, 0.3%) than in survivors (n=13, 0.2%), with a small difference (n=15, 0.1%; *P*=.01).

Most (n=7441, 64.5%) events involved male persons of interest and female survivors, and the majority (n=9399, 87.8%) occurred in residential settings (Table 2). Persons of interest aged 55-64 years had a higher proportion of survivors who were ex-spouses or ex-partners (n=890, 11.5% vs n=240, 8.1%), and fewer survivors who were spouses (n=2191, 28.3% vs n=1126, 37.8%) compared with those persons of interest aged 65 years and older. Most of the charges were related to physical violence, predominantly assaults. Homicide was uncommon and more prevalent in people aged 65 years and older, as was firearm use (n=11, 0.4%). In most events (n=6636, 62%), no weapon was recorded.

XSL•FO
RenderX

Table 2. Characteristics of police-attended domestic violence (DV) events in New South Wales, stratified by age.

Characteristics	All DV events (N=10,708), n (%)	DV events from persons of interest aged 55-64 years (n=7732), n (%)	DV events from persons of interest aged 65+ years (n=2976), n (%)
Paired sex (person of interest-survivor)			
Male-female	7441 (64.5)	5347 (69.2)	2094 (70.4)
Female-female	752 (7)	539 (7)	213 (7.2)
Male-male	1486 (13.9)	1077 (13.9)	409 (13.7)
Female-male	1013 (9.5)	757 (9.8)	256 (8.6)
Person of interest relationship with survivor			
Boyfriend or girlfriend (including ex)	704 (6.6)	597 (7.7)	107 (3.6)
Spouse or partner	3317 (31)	2191 (28.3)	1126 (37.8)
Ex-spouse or partner	1130 (10.6)	890 (11.5)	240 (8.1)
Parent or guardian of survivor	951 (8.9)	679 (8.8)	272 (9.1)
Child (including step or foster child) of survivor	273 (2.6)	213 (2.8)	60 (2)
Sibling	227 (2.1)	170 (2.2)	57 (1.9)
Other member of family	1017 (9.5)	695 (9)	322 (10.8)
Unknown or unrecorded	3089 (28.9)	2297 (29.7)	792 (26.6)
Event premises			
Residential	9399 (87.8)	6762 (87.5)	2637 (88.6)
Nonresidential	1309 (12.2)	970 (12.6)	339 (11.4)
Events associated with drugs			
Present	30 (0.3)	27 (0.3)	3 (0.1)
Events associated with alcohol			
Present	3631 (33.9)	2845 (36.8)	786 (26.4)
Alleged offence associated with event			
Homicide	28 (0.3)	16 (0.2)	12 (0.4)
Assault	7213 (67.4)	5085 (65.8)	2128 (71.5)
ADVO[a] breach	2203 (20.6)	1711 (22.1)	492 (16.5)
Other offense against person	1779 (16.6)	1273 (16.5)	506 (17)
Malicious damage	734 (6.9)	593 (7.7)	141 (4.7)
Survivor injury documented[b]			
Soft tissue injury	2304 (21.5)	1660 (21.5)	644 (21.6)
Wound (nonspecific)	1468 (13.7)	1089 (14.1)	379 (12.7)
Soreness	641 (6)	473 (6.1)	168 (5.7)
Fracture	318 (3)	228 (3)	90 (3)
Periorbital bruising	48 (0.4)	22 (0.3)	26 (0.9)
Stab wound	34 (0.3)	25 (0.3)	9 (0.3)
Bite or burn mark	27 (0.3)	21 (0.3)	6 (0.2)
Broken tooth	14 (0.1)	12 (0.2)	2 (0.1)
Other	416 (3.9)	303 (3.9)	113 (3.8)
No injuries reported	6,931 (64.7)	4982 (64.4)	1949 (65.5)
Weapon use			
Sharp instrument	308 (2.9)	218 (2.8)	90 (3)

XSL•FO
RenderX

Characteristics	All DV events (N=10,708), n (%)	DV events from persons of interest aged 55-64 years (n=7732), n (%)	DV events from persons of interest aged 65+ years (n=2976), n (%)
Blunt instrument	98 (0.9)	68 (0.9)	30 (1)
Firearm-related	23 (0.2)	12 (0.2)	11 (0.4)
Weapon unrecorded	6636 (62)	4657 (60.2)	1979 (66.5)
Abuse type			
Physical abuse only	3651 (34.1)	2610 (33.8)	1041 (35)
Nonphysical abuse only	1286 (12)	918 (11.9)	368 (12.4)
Physical and nonphysical abuse	2452 (22.9)	1771 (22.9)	681 (22.9)
No abuse event recorded	3319 (31)	2433 (31.5)	886 (29.8)

[a]ADVO: Apprehended Domestic Violence Order.

[b]Subcategories are not mutually exclusive. Multiple injuries and alleged offences could be recorded in each event; hence, subcategories exceed the total number of events.

At least 1 type of abuse was recorded in 69% (n=7389) of event narratives. Physical abuse events were the most common type of abuse (n=3651, 34.1%), followed by combined physical and nonphysical abuse (n=2452, 22.9%) and nonphysical abuse events (n=1286, 12%).

Trends in Police-Attended DV Events From 2005 to 2016

Over the period 2005-2016, the proportions of physical and nonphysical abuse events involving persons of interest aged 55-64 years (“younger” group) and those aged 65 years and older (“older” group) remained relatively stable over time for both sexes. However, female persons of interest aged 65 years and older showed higher and more variable proportions of physical abuse events compared to their male counterparts (Multimedia Appendix 6).

Female persons of interest in the older group had higher proportions of physical abuse and nonphysical abuse compared to those in the younger group (Multimedia Appendix 7). They also had smaller proportions of abuse events related to combined physical and nonphysical abuse events, breach of ADVOs, and other events. The proportion of abuse types was similar in male counterparts aged 55-64 years and 65 years and older, except for proportionally fewer breaches of ADVOs in the older group.

Multilevel Associations of Physical and Nonphysical Abuse

The multilevel analysis revealed several factors associated with physical abuse, nonphysical abuse, and both physical and nonphysical abuse (Tables 3 and 4). Alcohol-related events (aOR 1.04, 95% CI 1.02-1.06), survivor dementia (aOR 1.10, 95% CI 1.01-1.21), and persons of interest aged 65 years and older (aOR 1.02, 95% CI 1.00-1.05) were associated with higher odds of physical abuse. Conversely, persons of interest with substance use disorders had a lower odds of physical abuse (aOR 0.94, 95% CI 0.89-0.99). The inclusion of predictors in the model for the physical abuse outcome explained 69.40% of the person of interest–level variance compared to the null model.

For nonphysical abuse, male persons of interest had higher odds of perpetration (aOR 1.02, 95% CI 1.01-1.04), as did Indigenous persons of interest (aOR 1.04, 95% CI 1.00-1.08). Survivors with dementia had lower odds of experiencing nonphysical abuse (aOR 0.93, 95% CI 0.87-0.99). Indigenous survivors showed decreased odds of nonphysical abuse (aOR 0.96, 95% CI 0.92-1.00), albeit at the threshold of statistical significance. The predictors in the model for nonphysical abuse explained 96.50% of the survivor-level variance compared to the null model.

Events at residential premises were associated with higher odds of both (or co-occurring) physical and nonphysical abuse (aOR 1.05, 95% CI 1.03-1.08), as were alcohol-related events (aOR 1.03, 95% CI 1.02-1.05). Male survivors had lower odds of experiencing both physical and nonphysical abuse (aOR 0.97, 95% CI 0.95-0.99). Indigenous persons of interest had lower odds of perpetrating both physical and nonphysical abuse (aOR 0.94, 95% CI 0.90-0.99), while persons of interest with dementia (aOR 1.07, 95% CI 1.01-1.13) or a substance use disorder (aOR 1.13, 95% CI 1.08-1.19) had higher odds. The predictors in the model for both physical and nonphysical abuse explained 33.35% of the survivor-level variance compared to the null model.

In examining the consistency of findings across the 3 models (physical abuse only, nonphysical abuse only, and both physical and nonphysical abuse), several patterns emerged. Alcohol-related events were consistently associated with higher odds of physical abuse and both physical and nonphysical abuse. Person of interest with substance use disorder showed opposite effects: lower odds of physical abuse but higher odds of both physical and nonphysical abuse. Dementia in survivors was associated with higher odds of physical abuse but lower odds of nonphysical abuse, while person of interest with dementia was associated with higher odds of both physical and nonphysical abuse. Indigenous status showed varying associations: Indigenous survivors had lower odds of nonphysical abuse, while Indigenous persons of interest had higher odds of nonphysical abuse but lower odds of both physical and nonphysical abuse.

XSL•FO
RenderX

Table 3. Multilevel predictors of physical and nonphysical abuse.

Characteristics	Full model for physical abuse only		Full model for nonphysical abuse only		Full model for both physical and nonphysical abuse	
	aOR[a] (95% CI)	*P* value	aOR (95% CI)	*P* value	aOR (95% CI)	*P* value
Event level						
Premises						
Nonresidential	1.0 (reference[b])	—[c]	1.0 (reference)	—	1.0 (reference)	—
Residential	1.01 (0.99-1.04)	.34	1.01 (0.99-1.02)	.57	1.05 (1.03-1.08)	<.001
Alcohol-related event						
No	1.0 (reference)	—	1.0 (reference)	—	1.0 (reference)	—
Yes	1.04 (1.02-1.06)	<.001	1.01 (0.99-1.02)	.36	1.03 (1.02-1.05)	<.001
Drug-related-event						
No	1.0 (reference)	—	1.0 (reference)	—	1.0 (reference)	—
Yes	1.01 (0.85-1.19)	.92	0.92 (0.82-1.03)	.15	0.96 (0.83-1.11)	.58
Survivor level						
Survivor injury documented						
No	1.0 (reference)	—	1.0 (reference)	—	1.0 (reference)	—
Yes	1.16 (1.14-1.18)	<.001	0.93 (0.92-0.94)	<.001	1.08 (1.07-1.10)	<.001
Survivor sex						
Female	1.0 (reference)	—	1.0 (reference)	—	1.0 (reference)	—
Male	1.01 (0.99-1.04)	.29	1.00 (0.98-1.02)	.98	0.97 (0.95-0.99)	.003
Survivor age (years)						
Per-year increase in age	0.99 (0.99 – 0.99)	.04	1.00 (1.00-1.00)	.03	—	—
Indigenous status (survivor)						
Non-Indigenous	1.0 (reference)	—	1.0 (reference)	—	1.0 (reference)	—
Indigenous	1.05 (0.99-1.12)	.11	0.96 (0.92-1.00)	.05	1.03 (0.98-1.09)	.28
Psychosis (survivor)						
Not present	1.0 (reference)	—	1.0 (reference)	—	1.0 (reference)	—
Present	1.00 (0.91-1.10)	.95	1.02 (0.96-1.09)	.47	1.00 (0.92-1.09)	.96
Dementia (survivor)						
Not present	1.0 (reference)	—	1.0 (reference)	—	1.0 (reference)	—
Present	1.10 (1.01-1.21)	.03	0.93 (0.87-0.99)	.02	0.95 (0.88-1.03)	.23
Childhood-onset disorder (survivor)						
Not present	1.0 (reference)	—	1.0 (reference)	—	1.0 (reference)	—
Present	0.97 (0.89-1.07)	.54	1.00 (0.94-1.07)	.93	0.96 (0.89-1.05)	.39
Self-harm (survivor)						
Not present	1.0 (reference)	—	1.0 (reference)	—	1.0 (reference)	—
Present	1.04 (0.89-1.21)	.62	1.02 (0.91-1.13)	.76	0.93 (0.81-1.06)	.29
Substance use disorder (survivor)						
Not present	1.0 (reference)	—	1.0 (reference)	—	1.0 (reference)	—
Present	0.94 (0.84-1.04)	.24	1.06 (0.99-1.14)	.10	1.00 (0.91-1.10)	.98
Personality disorder (survivor)						
Not present	1.0 (reference)	—	1.0 (reference)	—	1.0 (reference)	—
Present	1.02 (0.80-1.30)	.85	1.05 (0.89-1.25)	.53	1.00 (0.80-1.23)	.97

XSL•FO
RenderX

Characteristics	Full model for physical abuse only		Full model for nonphysical abuse only		Full model for both physical and nonphysical abuse	
	aOR[a] (95% CI)	*P* value	aOR (95% CI)	*P* value	aOR (95% CI)	*P* value
Person of interest level						
Person of interest sex						
Female	1.0 (reference)	—	1.0 (reference)	—	1.0 (reference)	—
Male	0.98 (0.95-1.00)	.07	1.02 (1.01-1.04)	.01	0.98 (0.96-1.01)	.14
Person of interest age						
55-64	1.0 (reference)	—	1.0 (reference)	—	1.0 (reference)	—
65+	1.02 (1.00-1.05)	.03	1.00 (0.98-1.01)	.91	1.00 (0.98-1.02)	.92
Indigenous status (person of interest)						
Non-Indigenous	1.0 (reference)	—	1.0 (reference)	—	1.0 (reference)	—
Indigenous	1.01 (0.95-1.07)	.74	1.04 (1.00-1.08)	.04	0.94 (0.90-0.99)	.02
Psychosis (person of interest)						
Not present	1.0 (reference)	—	1.0 (reference)	—	1.0 (reference)	—
Present	1.00 (0.93-1.08)	>.99	1.02 (0.97-1.08)	.36	1.06 (0.99-1.13)	.09
Dementia (person of interest)						
Not present	1.0 (reference)	—	1.0 (reference)	—	1.0 (reference)	—
Present	0.98 (0.92-1.04)	.52	1.02 (0.97-1.06)	.50	1.07 (1.01-1.13)	.02
Childhood-onset disorder (person of interest)						
Not present	1.0 (reference)	—	1.0 (reference)	—	1.0 (reference)	—
Present	1.00 (0.92-1.09)	.96	1.03 (0.97-1.10)	.26	1.06 (0.98-1.15)	.13
Self-harm (person of interest)						
Not present	1.0 (reference)	—	1.0 (reference)	—	1.0 (reference)	—
Present	1.10 (0.97-1.24)	.14	0.98 (0.90-1.06)	.57	0.98 (0.88-1.10)	.77
Substance use disorder (person of interest)						
Not present	1.0 (reference)	—	1.0 (reference)	—	1.0 (reference)	—
Present	0.94 (0.89-0.99)	.02	1.00 (0.96-1.04)	.92	1.13 (1.08-1.19)	<.001
Personality disorder (person of interest)						
Not present	1.0 (reference)	—	1.0 (reference)	—	1.0 (reference)	—
Present	0.89 (0.74-1.06)	.19	1.00 (0.88-1.13)	.99	1.02 (0.87-1.19)	.81

[a]aOR: adjusted odds ratio.

[b]Reference category indicates the baseline group against which the other category or categories are compared for the specific abuse category.

[c]Not applicable.

XSL•FO
RenderX

Table 4. Random effects parameters.

Parameters	Full model for physical abuse only	Full model for nonphysical abuse only	Full model for both physical and nonphysical abuse
Persons of interest			
Variance	0.003	0.01	0.06
ICC[a] (%)	1.45	7.51	3.32
PCV[b] (%)	69.40	–17.27	6.23
Survivors nested within persons of interest			
Variance	0.02	Negligible	0.05
ICC (%)	9.24	0	3.16
PCV (%)	4.45	96.50	33.35

[a]ICC: intraclass correlation coefficient.

[b]PCV: proportional change in variance.

Stratified Multilevel Analysis by Person of Interest Age Groups

Factors With the Same Direction of Effect Among Both Age Groups

In the stratified analysis by person of interest age (Multimedia Appendices 8-10), several factors showed consistent directions of effect for both person of interest age groups across all 3 abuse outcomes. For physical abuse, residential premises increased odds in both age groups (55-64 years: aOR 1.38, 95% CI 1.17-1.63 and 65 years and older: aOR 1.33, 95% CI 1.02-1.74). Person of interest sex (female) showed an increase in odds of physical abuse for those aged 55-64 years (aOR 1.22, 95% CI 1.04-1.44) but lower odds of nonphysical abuse (aOR 0.72, 95% CI 0.57-0.90). Among female persons of interest, female survivors had higher odds of experiencing physical abuse compared to male survivors (aOR 1.39, 95% CI 1.09-1.78); however, among male persons of interest, female survivors had significantly higher odds of experiencing both physical and nonphysical abuse (aOR 1.27, 95% CI 1.09-1.48). Male-perpetrated physical abuse was more likely in residential premises (aOR 1.42, 95% CI 1.21-1.66); however, no such difference was observed for female-perpetrated abuse. There were also no differences between survivor age and abuse type among male or female persons of interest.

Documented survivor injury was associated with physical abuse in both groups (55-64 years: aOR 3.28, 95% CI 2.88-3.75 and 65 years and older: aOR 2.86, 95% CI 2.35-3.48). A similar trend was observed for both physical and nonphysical abuse; however, substance use disorder in persons of interest was an additional risk factor in both groups (55-64 years: aOR 1.87, 95% CI 1.36-2.56 and 65 years and older: aOR 1.98, 95% CI 1.13-3.45). In the case of nonphysical abuse, documented survivor injury was associated with lower odds in both age groups (55-64 years: aOR 0.44, 95% CI 0.37-0.53 and 65 years and older: aOR 0.47, 95% CI 0.35-0.62).

Factors With Opposite Direction of Effect Among Both Age Groups

Some factors showed opposite directions of effect between the 2 person of interest age groups. For physical abuse, the Indigenous status of person of interest decreased odds for the 55- to 64-year group (aOR 0.68, 95% CI 0.49-0.95) but increased odds for the 65+-year group (aOR 1.60, 95% CI 0.71-3.58), although the latter was not statistically significant. A similar pattern was observed for both physical and nonphysical abuse regarding the Indigenous status of person of interest (55-64 years: aOR 0.62, 95% CI 0.44-0.88 and 65 years and older: aOR 1.34, 95% CI 0.58-3.08).

For nonphysical abuse, alcohol-related events were associated with decreased odds for the 55- to 64-year group (aOR 0.91, 95% CI 0.78-1.08) but with significantly higher odds for the 65+-year group (aOR 1.68, 95% CI 1.28-2.20). Substance use disorder in survivors was associated with increased odds for the 55- to 64-year group (aOR 2.15, 95% CI 1.03-4.48) but decreased odds for the 65+-year group (aOR 0.82, 95% CI 0.16-4.24), though this was not statistically significant for the older group.

Dementia in persons of interest was associated with significantly higher odds of both physical and nonphysical abuse in the 65+-year group (aOR 1.73, 95% CI 1.19-2.53), but not in the 55- to 64-year group (aOR 0.81, 95% CI 0.31-2.14).

Discussion

Main Findings

This study examined 10,708 police-attended DV events involving 8247 persons of interest aged 55 years and older and 8629 survivors in New South Wales, Australia, from 2005 to 2016. Consistent with the literature on DV across the age spectrum [4] including older people [7,14], most persons of interest were male, and most of the survivors were female. Initially, descriptive analyses suggested a higher and more variable proportion of physical abuse events among female persons of interest aged 65 years and older. However, after adjusting for confounders and accounting for data clustering, female persons of interest aged 55-64 years and 65 years and older showed similarly elevated odds of physical abuse compared to male persons of interest. This insight aligns with existing research that indicates that female-perpetrated abuse increases with increasing survivor age [22] and that older male individuals are more likely to report exposure to violence [13].

XSL•FO
RenderX

Female perpetration of DV is still underresearched, and this age-group association might reflect the stress of caregiver responsibilities or shifts in power relationships due to survivor dependency.

The association between abuse type and location differed by perpetrator sex, with male-perpetrated physical abuse more likely in residential settings, a pattern not observed for female perpetrators. Furthermore, while the overall odds of physical abuse decreased with survivor age and nonphysical abuse increased, these age-related patterns in exposure to violence were no longer evident after stratification by perpetrator sex. Notably, substance use disorder was consistently associated with higher odds of abuse across both person of interest age groups (55-64 and 65+ years). However, dementia and mental disorders typically diagnosed in childhood were associated with higher odds of both physical and nonphysical abuse only in the 65+-year age group. These findings suggest age-specific risk factors for DV perpetration in later life, with cognitive impairment and long-standing mental health issues emerging as significant contributors with increasing age. However, as these data are cross-sectional and observational in nature, an interplay of factors or unobserved factors might also account for these findings, and further studies are required.

At the person of interest level, there were higher odds of physical abuse in the group aged 65 years and older compared with the younger group, which might reflect the inclusion of physical abuse perpetrated by people with dementia in this study or abuse perpetrated by aging caregivers. Physical abuse and combined physical and nonphysical abuse events were more common than nonphysical abuse events. This might initially appear to contradict the typical picture of DV in older people, which is declining physical abuse with age [14,36], and could be because these data only cover events where police were called, which is more likely to occur if a person is concerned about physical safety. The decision to call the police in situations of DV is complex [32] and made more complex in intimate relationships where the perpetrator has a mental health condition that the survivor believes contributes to their behavior, with survivors reporting a sense of obligation to stay with the perpetrator [25]. The obligation to stay with the perpetrator can also occur in other relationships, such as adult children who care for their parents. One study found that police notifications for DV decreased with survivor age [29], possibly because the authors limited their data to events where survivors had contacted the police, whereas these data included all police callouts.

Compared with survivors, mental health conditions were more common in persons of interest who had a greater proportion of dementia (n=191, 2.3% vs n=104, 1.2%), psychosis (n=142, 1.7% vs n=87, 1%), and alcohol and substance use disorder (n=200, 2.4% vs n=65, 0.8%), consistent with other research [9,14]. Alcohol may have played a situational role in events involving physical abuse compared to nonphysical abuse, while persons of interest with substance use disorders (which include alcohol use disorders) increased the odds of DV perpetration involving both physical and nonphysical abuse when compared with physical-only and nonphysical-only abuse. Alcohol use disorders have been consistently associated with DV in later age [7,9,14,15], and the role of acute alcohol consumption has also been noted in IPV research [8,25]. Alcohol tolerance decreases with age, leading to higher and more prolonged alcohol levels compared with younger people [34]. Acute intoxication impairs most cognitive functions [35], including judgment and impulse control, and may be pertinent for new-onset DV in older people who consume alcohol and are unaware of their reduced tolerance. Substance use disorders in older people are associated with cognitive impairment, particularly visuoperceptual and executive dysfunction [35], which could explain the association with the commission of violence.

The effects of dementia were exerted at both a survivor and perpetrator level, increasing the odds of solely physical and solely nonphysical exposure to violence compared with the combined abuse group. The finding that dementia was associated with increased exposure to violence is consistent with other studies of DV survivors with dementia [36,37]. Exposure to violence of people with dementia has been viewed through the prism of caregiver stress [15,38], and this study adds to the literature, finding that the odds of such exposure to violence are increased for physical abuse but decreased for nonphysical abuse types. The role of caregiver stress in the perpetration of abuse is not linear, and there are likely mediating factors. The results of this study also found that dementia increased the odds of being a person of interest in events where combined physical and nonphysical abuse occurred, consistent with the findings of high levels of aggression in people with dementia [12,39]. This complex relationship between dementia and violent behavior suggests the need for further examination of specific characteristics that are associated with exposure to violence and perpetration of violence in this group.

A greater proportion of homicide (n=12, 0.4% vs n=16, 0.2%) and assault charges (n=2128, 71.5% vs n=5085, 65.8%) were laid against persons of interest in the older group. The older group also had a greater proportion of firearm use (n=11, 0.4% vs n=12, 0.2%), a finding that requires further examination. Weapons might be used by older people to compensate for decreased physical strength but have a greater propensity for resulting in injury and death, particularly if wielded against a spouse who is similarly aged.

Future Research and Policy Directions

This study has demonstrated shared and separate factors involved in the perpetration of physical and nonphysical abuse in older age. Future research is required into examining this association and determining the salience of these factors for DV that emerges in later life, as opposed to younger onset DV. For instance, Policastro et al [9] have suggested that new onset DV in older people may be more likely to be associated with mental illness and dementia, whereas DV “grown old” usually presents with emotional abuse.

Further, traditional theoretical frameworks underpinning the commission of violence in younger people may require adaptation for their application to older people. DV and, more recently, abuse of older people are viewed through a lens of power and control [7,19], while abuse of older people also acknowledges the role of caregiver stress [15]. These

XSL•FO
RenderX

frameworks will need to accommodate the role of biological aging on alcohol tolerance and cognition as well as changes in cognition associated with the onset of neurological disorders. Conversely, rather than viewing aggression in people with dementia solely as a symptom of the condition, a consideration of premorbid relationship dynamics and scrutiny of premorbid behaviors is required to determine if the dementia is an exculpatory factor. The role of premorbid personality was demonstrated in a study where caregivers of 99 people with dementia who displayed aggressive behavior were interviewed. In 3% of cases, the behaviors were similar to their premorbid state; in 43% of cases, it was an exaggeration of premorbid behavior; and in 53% of people, the aggression was different from their premorbid behavior [39].

Qualitative studies are recommended to understand the dynamics of relationships between the perpetrator and survivor, which may be influenced by caring responsibilities, frailty, and disability, in addition to earning constraints and social and cultural factors. The data also point to the need for further research to understand older female individuals who perpetrate violence, a particularly neglected group.

From a policy perspective, this research underlines the need for a multipronged approach to target substance use in older people, including education concerning safe levels of alcohol consumption in older age provided at an individual level through health providers and at a population level through government advertising campaigns with focus group evaluation of their success. This is particularly pertinent as the generation of adults born in the “baby boomer” generation because this generation has comparatively high use of alcohol and drugs compared to previous older generations [40].

The findings of this study have implications for clinicians who treat people with dementia, as dementia increases the odds of being both a survivor and a perpetrator of abuse. Health professionals who treat older people in primary and secondary care require education to recognize signs of nonaccidental injury in older people with dementia and to specifically enquire about the patient’s alcohol use, signs of aggression, agitation, and psychosis, particularly if they reside with another older person who might be at risk of serious injury or death from violence. This could be undertaken by professional specialty medical and nursing societies for emergency medicine, geriatrics, and old age psychiatry.

Strengths and Limitations

A strength of this study is that it did not rely on survivor recall of events, which could be contaminated by recall bias, although qualitative data would be useful for a greater understanding of the reasons why police are called. This is one of the largest datasets of older DV covering both male and female persons of interest and survivors. Yon et al [41] combined 2 cross-sectional datasets to obtain 6971 male and female survivors aged 60 years and older, Zink et al [36] looked at 995 female survivors aged 55 years and older, while Yan and Chan [13] examined 937 cases who were aged 60 years and older for IPV, including 540 male individuals.

This dataset is unable to cover the array of interconnected factors that influence DV, including cultural and environmental factors, such as resourcing, relationship factors, and individual factors, such as education and income level. These issues are multifaceted and, because they are strongly affected by cultural norms [42], these results may not be generalizable to all cultures and should be interpreted in conjunction with culture-specific qualitative data.

In addition, it is possible that older adults might be hesitant to report DV events to the police for several reasons, and as such, the number of police DV events involving older perpetrators and survivors could be underreported. Older people might not recognize abusive behaviors often focusing on intentional physical abuse while overlooking other forms, such as neglect, social, or financial abuse [43,44]. For example, Acierno et al [44] mentioned that financial abuse was less likely to be reported to authorities when perpetrated by a family member compared to similar acts by strangers. Furthermore, relationships between elder survivors and perpetrators such as caregivers or adult children may discourage reporting due to feelings of guilt and the desire to protect family members from legal consequences [45,46]. Finally, elder survivors may be unaware of the appropriate authority or available services to disclose their abuse to [47].

The use of police data required a notification, and deciding to call the police in cases of DV can be complex, depending on the relationship between the survivor and the person of interest. Further, concerns of police bias against people due to their race, socioeconomic status, sexuality, or past encounters can dissuade people from seeking police assistance [48]. Cases where the person experiencing abuse was reluctant to tell the police about the abuse (eg, due to resolution of the disagreement, fear of the person of interest, concerns about police bias, or stigma) and where there was no obvious sign of injury or disturbance may have been relegated to the “no abuse event recorded” category, such that underreporting of abuse types could occur, which is particularly salient for older people, who are more reluctant to disclose abuse if they are dependent on the perpetrator [32] and because of generational norms [49]. It is also likely that police would record only observable or reported survivor injuries, leading to potential underreporting in people with dementia who were unable to communicate how and where they were injured if it was not obviously visible, and the occurrence of nonphysical abuse types.

Police officers are not experts in diagnosing mental illnesses and rely on the testimony of witnesses at the DV event and on the self-reports of persons of interest and survivors. Therefore, the accuracy of mental illness mentions within police narratives remains unclear, considering that the purpose of collecting such data is not for research, and any interpretation of our findings should be taken with caution [50]. Any misclassification of mental health status for persons of interest (or survivors) in the narratives could potentially produce wrong conclusions for specific conditions (eg, dementia) in the DV setting. Underreporting of dementia in persons of interest might underestimate its association with physical and nonphysical abuse, affecting the respective aORs in Table 3 (eg, aOR 1.07, 95% CI 1.01-1.13 for dementia in persons of interest for

XSL•FO
RenderX

combined physical and nonphysical abuse). On the other hand, overreporting based on subjective observations in the police narratives could inflate these associations for conditions like dementia, which require clinical assessment, affecting the precision of our aORs, potentially leading to wider CIs or effect sizes. To address this, future research could focus on linking police records with clinical datasets to validate mental health mentions, enabling the application of correction methods to account for diagnostic uncertainty [51].

Conclusions

This study suggests that physical and nonphysical DV persists in later life. Alcohol and dementia increase the likelihood of perpetration, and screening should target these factors. Instead of reactively treating aggression that arises in the context of dementia, clinicians should regularly evaluate the risk of violence and exposure to violence in patients with dementia and educate caregivers about potential violent behavior early in the course of the condition. This would enable the provision of information on services that can be used to assist in the management of these behaviors. Additionally, clinicians require education to be alert for nonaccidental injury in people with dementia. Pediatric services have guidelines and are aware of red flags for suspected child abuse and neglect [52], which can inform the development of similar protocols for geriatric medical services.

Data Availability

Due to the sensitive nature of police documents, the data and materials are not available for replication purposes.

Authors' Contributions

SR conceived and planned the study, analyzed the data, interpreted the statistical results, and wrote the initial manuscript draft. EA planned the study, conducted the statistical analysis and statistical result interpretation, and revised the manuscript. GK provided the deidentified data, undertook data analysis, revised the manuscript, and provided supervision. AW interpreted the results, revised the manuscript, and provided supervision. TB planned the study, revised the manuscript, interpreted the results, and provided supervision. All authors reviewed the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Abuse categories and abuse types.
[DOCX File , 18 KB - aging_v8i1e75993_app1.docx]

Multimedia Appendix 2
Description of the extracted abuse types.
[DOCX File , 21 KB - aging_v8i1e75993_app2.docx]

Multimedia Appendix 3
Categories used to group survivor injury types derived from text mining of police narratives.
[DOCX File , 18 KB - aging_v8i1e75993_app3.docx]

Multimedia Appendix 4
Yearly breakdown of total police-reported events by sex (female and male) for persons of interest aged 55-64 years (2005-2016).
[DOCX File , 20 KB - aging_v8i1e75993_app4.docx]

Multimedia Appendix 5
Yearly breakdown of total police-reported events by sex (female and male) for persons of interest aged 65+ years (2005-2016).
[DOCX File , 20 KB - aging_v8i1e75993_app5.docx]

Multimedia Appendix 6
Proportion of physical and nonphysical abuse incidents involving female and male persons of interest aged 55-64 and 65+ years, annually from 2005 to 2016.
[DOCX File , 168 KB - aging_v8i1e75993_app6.docx]

Multimedia Appendix 7
Abuse types police-attended domestic violence events in New South Wales from 2005 to 2016, by sex and age group.
[PNG File , 16 KB - aging_v8i1e75993_app7.png]

XSL•FO
RenderX

Multimedia Appendix 8
Predictors of physical abuse, stratified by person of interest age groups.
[DOCX File , 20 KB - aging_v8i1e75993_app8.docx]

Multimedia Appendix 9
Predictors of both physical and nonphysical abuse, stratified by person of interest age groups.
[DOCX File , 22 KB - aging_v8i1e75993_app9.docx]

Multimedia Appendix 10
Predictors of nonphysical abuse, stratified by person of interest age groups.
[DOCX File , 20 KB - aging_v8i1e75993_app10.docx]

References

1. Personal safety, Australia. Australian Bureau of Statistics. 2021. URL: https://www.abs.gov.au/statistics/people/crime-and-justice/personal-safety-australia/latest-release [accessed 2025-08-30]
2. Bair-Merritt MH, Blackstone M, Feudtner C. Physical health outcomes of childhood exposure to intimate partner violence: a systematic review. Pediatrics 2006;117(2):e278-e290. [doi: 10.1542/peds.2005-1473] [Medline: 16452335]
3. Campbell JC. Health consequences of intimate partner violence. Lancet 2002;359(9314):1331-1336. [doi: 10.1016/s0140-6736(02)08336-8]
4. Family, domestic and sexual violence in Australia: continuing the national story. Australian Institute of Health and Welfare. 2019. URL: https://www.aihw.gov.au/reports/family-domestic-and-sexual-violence/family-domestic-sexual-violence-australia-2019/contents/summary [accessed 2025-08-30]
5. Understanding and addressing violence against women. World Health Organization. 2012. URL: https://iris.who.int/bitstream/handle/10665/77433/WHO_RHR_12.35_eng.pdf [accessed 2025-08-30]
6. Brodaty H, Draper BM, Low L. Behavioural and psychological symptoms of dementia: a seven-tiered model of service delivery. Med J Aust 2003;178(5):231-234. [doi: 10.5694/j.1326-5377.2003.tb05169.x] [Medline: 12603188]
7. Roberto KA, McPherson MC, Brossoie N. Intimate partner violence in late life: a review of the empirical literature. Violence Against Women 2013;19(12):1538-1558 [FREE Full text] [doi: 10.1177/1077801213517564] [Medline: 24476758]
8. Warmling D, Lindner S, Coelho E. Intimate partner violence prevalence in the elderly and associated factors: systematic review. Cien Saude Colet 2017;22(9):3111-3125 [FREE Full text] [doi: 10.1590/1413-81232017229.12312017] [Medline: 28954161]
9. Policastro C, Finn MA. Coercive control and physical violence in older Adults: analysis using data from the National Elder Mistreatment Study. J Interpers Violence 2017;32(3):311-330. [doi: 10.1177/0886260515585545] [Medline: 25976315]
10. Rogers MM, Storey JE, Galloway S. Elder mistreatment and dementia: a comparison of people with and without dementia across the prevalence of abuse. J Appl Gerontol 2023;42(5):909-918 [FREE Full text] [doi: 10.1177/07334648221145844] [Medline: 36564912]
11. Oh J, Kim HS, Martins D, Kim H. A study of elder abuse in Korea. Int J Nurs Stud 2006;43(2):203-214. [doi: 10.1016/j.ijnurstu.2005.03.005] [Medline: 15913631]
12. Brodaty H, Low LF. Aggression in the elderly. J Clin Psychiatry 2003;64(Suppl 4):36-43. [Medline: 12672263]
13. Yan E, Chan KL. Prevalence and correlates of intimate partner violence among older Chinese couples in Hong Kong. Int Psychogeriatr 2012;24(9):1437-1446 [FREE Full text] [doi: 10.1017/S1041610212000294] [Medline: 22452872]
14. Gerino E, Caldarera AM, Curti L, Brustia P, Rollè L. Intimate partner violence in the golden age: Systematic review of risk and protective factors. Front Psychol 2018;9:1595 [FREE Full text] [doi: 10.3389/fpsyg.2018.01595] [Medline: 30233454]
15. Yan E, Chan KL, Tiwari A. A systematic review of prevalence and risk factors for elder abuse in Asia. Trauma Violence Abuse 2015;16(2):199-219. [doi: 10.1177/1524838014555033] [Medline: 25380662]
16. DePrince AP, Olomi JM, Lee MS, Hasche LK, Casparian AM, Pettit LL. Exploring an emerging measurement issue in older adult maltreatment: exploitation of older adults' residences. J Trauma Dissociation 2020;21(2):158-171. [doi: 10.1080/15299732.2020.1692406] [Medline: 31762391]
17. Pillemer K, Burnes D, Riffin C, Lachs MS. Elder abuse: global situation, risk factors, and prevention strategies. Gerontologist 2016;56(Suppl 2):S194-S205 [FREE Full text] [doi: 10.1093/geront/gnw004] [Medline: 26994260]
18. De Donder L, Lang G, Luoma ML, Penhale B, Ferreira Alves J, Tamutiene I, et al. Perpetrators of abuse against older women: a multi‐national study in Europe. J Adult Prot 2011;13(6):302-314. [doi: 10.1108/14668201111194212]
19. Penhale B. Older women, domestic violence, and elder abuse: a review of commonalities, differences, and shared approaches. J Elder Abuse Neglect 2003;15(3-4):163-183. [doi: 10.1300/j084v15n03_10]
20. Khurana B, Loder RT. Injury patterns and associated demographics of intimate partner violence in older adults presenting to U.S. emergency departments. J Interpers Violence 2022;37(17-18):NP16107-NP16129. [doi: 10.1177/08862605211022060] [Medline: 34121494]

XSL•FO
RenderX

21. Zhang Kudon H, Herbst JH, Richardson LC, Smith SG, Demissie Z, Siordia C. Prevalence estimates and factors associated with violence among older adults: National Intimate Partner and Sexual Violence (NISVS) Survey, 2016/2017. J Elder Abuse Negl 2024;36(1):67-83. [doi: 10.1080/08946566.2023.2297227] [Medline: 38129823]
22. Ajdukovic M, Ogresta J, Rusac S. Family violence and health among elderly in Croatia. J Aggress Maltreatment Trauma 2009;18(3):261-279. [doi: 10.1080/10926770902835873]
23. Huntley AL, Potter L, Williamson E, Malpass A, Szilassy E, Feder G. Help-seeking by male victims of domestic violence and abuse (DVA): a systematic review and qualitative evidence synthesis. BMJ Open 2019;9(6):e021960 [FREE Full text] [doi: 10.1136/bmjopen-2018-021960] [Medline: 31186243]
24. Policastro C, Finn MA. Coercive control in intimate relationships: differences across age and sex. J Interpers Violence 2021;36(3-4):1520-1543. [doi: 10.1177/0886260517743548] [Medline: 29294994]
25. Zink T, Jacobson CJ, Regan S, Fisher B, Pabst S. Older women's descriptions and understandings of their abusers. Violence Against Women 2006;12(9):851-865. [doi: 10.1177/1077801206292680] [Medline: 16905677]
26. Zink T, Fisher BS. The prevalence and incidence of intimate partner and interpersonal mistreatment in older women in primary care offices. J Elder Abuse Negl 2006;18(1):83-105. [doi: 10.1300/j084v18n01_04] [Medline: 17439876]
27. Elder GJ. Life course theory. In: Kazdin EA, editor. Encyclopedia of Psychology. England: Oxford University Press; 2000.
28. Heise LL. Violence against women: an integrated, ecological framework. Violence Against Women 1998;4(3):262-290. [doi: 10.1177/1077801298004003002] [Medline: 12296014]
29. Gilmore MC, Stebbins L, Argüelles-Borge S, Trinidad B, Golden CJ. Development and treatment of aggression in individuals with dementia. Aggress Violent Behav 2020;54:101415. [doi: 10.1016/j.avb.2020.101415]
30. Schreiber E, Salivar EG. Using a vulnerability-stress-adaptation framework to model intimate partner violence risk factors in late life: a systematic review. Aggress Violent Behav 2021;57:101493. [doi: 10.1016/j.avb.2020.101493]
31. Herrenkohl TI, Fedina L, Roberto KA, Raquet KL, Hu RX, Rousson AN, et al. Child maltreatment, youth violence, intimate partner violence, and elder mistreatment: a review and theoretical analysis of research on violence across the life course. Trauma Violence Abuse 2022;23(1):314-328 [FREE Full text] [doi: 10.1177/1524838020939119] [Medline: 32723166]
32. Kang JH, Lynch JP. Calling the police in instances of family violence: effects of victim-offender relationship and life stages. Crime Delinq 2010;60(1):34-59. [doi: 10.1177/0011128709359655]
33. Karystianis G, Adily A, Schofield PW, Greenberg D, Jorm L, Nenadic G, et al. Automated analysis of domestic violence police reports to explore abuse types and victim injuries: text mining study. J Med Internet Res 2019;21(3):e13067 [FREE Full text] [doi: 10.2196/13067] [Medline: 30860490]
34. Karystianis G, Adily A, Schofield P, Knight L, Galdon C, Greenberg D, et al. Automatic extraction of mental health disorders from domestic violence police narratives: text mining study. J Med Internet Res 2018;20(9):e11548 [FREE Full text] [doi: 10.2196/11548] [Medline: 30213778]
35. Wilson M, Spike E, Karystianis G, Butler T. Nonfatal strangulation during domestic violence events in New South Wales: prevalence and characteristics using text mining study of police narratives. Violence Against Women 2022;28(10):2259-2285. [doi: 10.1177/10778012211025993] [Medline: 34581646]
36. Zink T, Fisher BS, Regan S, Pabst S. The prevalence and incidence of intimate partner violence in older women in primary care practices. J Gen Intern Med 2005;20(10):884-888 [FREE Full text] [doi: 10.1111/j.1525-1497.2005.0191.x] [Medline: 16191132]
37. Shen Y, Sun F, Feng Y, Lichtenberg PA, Meng H. Prevalence of elder abuse and neglect of persons with dementia in community settings: a systematic review and meta-analysis. Gerontology 2025;71(5):400-416. [doi: 10.1159/000543804] [Medline: 40552876]
38. Flannery RB. Domestic violence and elderly dementia sufferers. Am J Alzheimers Dis Other Demen 2003;18(1):21-23 [FREE Full text] [doi: 10.1177/153331750301800107] [Medline: 12613130]
39. Keene J, Hope T, Fairburn CG, Jacoby R, Gedling K, Ware CJ. Natural history of aggressive behaviour in dementia. Int J Geriatr Psychiatry 1999;14(7):541-548. [doi: 10.1002/(sici)1099-1166(199907)14:7<541::aid-gps961>3.0.co;2-p] [Medline: 10440974]
40. Barry KL, Blow FC. Drinking over the lifespan: focus on older adults. Alcohol Res 2016;38(1):115-120 [FREE Full text] [Medline: 27159818]
41. Yon Y, Wister AV, Mitchell B, Gutman G. A national comparison of spousal abuse in mid- and old age. J Elder Abuse Negl 2014;26(1):80-105. [doi: 10.1080/08946566.2013.784085] [Medline: 24313799]
42. Shorey S, Chua CMS, Chan V, Chee CYI. Women living with domestic violence: ecological framework-guided qualitative systematic review. Aggress Violent Behav 2023;71:101835. [doi: 10.1016/j.avb.2023.101835]
43. Mysyuk Y, Westendorp RGJ, Lindenberg J. Older persons' definitions and explanations of elder abuse in the Netherlands. J Elder Abuse Negl 2016;28(2):95-113. [doi: 10.1080/08946566.2015.1136580] [Medline: 26731491]
44. Acierno R, Steedley M, Hernandez-Tejada MA, Frook G, Watkins J, Muzzy W. Relevance of perpetrator identity to reporting elder financial and emotional mistreatment. J Appl Gerontol 2020;39(2):221-225. [doi: 10.1177/0733464818771208] [Medline: 29703128]
45. Jackson SL, Hafemeister TL. The impact of relationship dynamics on the detection and reporting of elder abuse occurring in domestic settings. J Elder Abuse Negl 2015;27(2):121-145. [doi: 10.1080/08946566.2015.1008085] [Medline: 25611176]

XSL•FO
RenderX

46. Kaspiew R, Carson R, Rhoades H. Elder abuse in Australia. Family Matters 2016;1(98):64-73 [FREE Full text]
47. Clare M, Clare B, Blundell B, Clare J. Conceptualising elder abuse: does this label fit? Commun Children Fam Australia 2014;8(1):37-48 [FREE Full text]
48. Wolf ME, Ly U, Hobart MA, Kernic MA. Barriers to seeking police help for intimate partner violence. J Family Violence 2003;18(2):121-129. [doi: 10.1023/a:1022893231951]
49. Straka SM, Montminy L. Responding to the needs of older women experiencing domestic violence. Violence Against Women 2006;12(3):251-267. [doi: 10.1177/1077801206286221] [Medline: 16456151]
50. Karystianis G, Cabral RC, Adily A, Lukmanjaya W, Schofield P, Buchan I, et al. Mental illness concordance between hospital clinical records and mentions in domestic violence police narratives: data linkage study. JMIR Form Res 2022;6(10):e39373 [FREE Full text] [doi: 10.2196/39373] [Medline: 36264613]
51. Fox MP, Lash TL, Greenland S. A method to automate probabilistic sensitivity analyses of misclassified binary variables. Int J Epidemiol 2005;34(6):1370-1376. [doi: 10.1093/ije/dyi184] [Medline: 16172102]
52. Appenteng R, Williams H, Fain E, Frazier S, Daly C, Phillips A, et al. Improving evidence-based non-accidental trauma screening in a pediatric emergency department. Pediatrics 2021;147:153. [doi: 10.1542/peds.147.3ma2.153a]

Abbreviations

ADVO: Apprehended Domestic Violence Order
aOR: adjusted odds ratio
BPSD: behavioral and psychological symptoms of dementia
COPS: Computerized Operational Policing System
DV: domestic violence
IPV: intimate partner violence

Edited by R Yang; submitted 14.04.25; peer-reviewed by D Saleh, MW Huang; comments to author 09.06.25; revised version received 11.07.25; accepted 28.08.25; published 29.09.25.

Please cite as:
Reutens S, Akpanekpo E, Karystianis G, Withall A, Butler T
Older Perpetrators of Domestic Violence: Mixed-Effects Logistic Regression Analysis of Police Records
JMIR Aging 2025;8:e75993
URL: https://aging.jmir.org/2025/1/e75993
doi:10.2196/75993
PMID:

©Sharon Reutens, Emaediong Akpanekpo, George Karystianis, Adrienne Withall, Tony Butler. Originally published in JMIR Aging (https://aging.jmir.org), 29.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Associations Between Hearing Loss and Depressive Symptom Trajectories in Middle-Aged and Older People in China: Retrospective Analysis

Yuchen Liu[1,2*], MD; Wenwen Liu[1*], MPH; Jun Ma[1,2], MD; Yangfan Chai[1,3], MPH; Guilan Kong[1,4], PhD

[1]National Institute of Health Data Science, Peking University, Beijing, China

[2]Peking University Third Hospital, Beijing, China

[3]Peking University Chongqing Research Institute of Big Data, Chongqing, China

[4]Advanced Institute of Information Technology, Peking University, Hangzhou, Zhejiang, China

[*]these authors contributed equally

Corresponding Author:
Guilan Kong, PhD
National Institute of Health Data Science
Peking University
Number 38 Xueyuan Road, Haidian District
Beijing, 100191
China
Phone: 86 10 82806542
Email: guilan.kong@hsc.pku.edu.cn

Abstract

Background: Hearing loss and depression are important health issues among the middle-aged and older population.

Objective: This study aimed to investigate the associations between hearing loss and depressive symptom trajectories in the Chinese middle-aged and older adult population.

Methods: The survey data of 2011, 2013, 2015, and 2018 waves collected in the China Health and Retirement Longitudinal Study were used for analysis. The latent growth mixture modeling approach was used to explore the trajectories of depressive symptoms. Hearing loss was identified through self-reporting, and depressive symptoms were assessed using the 10-item Center for Epidemiologic Studies Depression scale. The associations between hearing loss and depressive symptom trajectories were examined using logistic regression models.

Results: A total of 4768 participants without depressive symptoms at baseline were included for analysis. Among them, 4 depressive symptom trajectories, including “stable low symptoms” (n=3656, 76.68%), “slowly progressing symptoms” (n=503, 10.55%), “relieved symptoms after progression” (n=467, 9.79%), and “rapidly progressing symptoms” (n=142, 2.98%) were identified. Hearing loss was found to be significantly associated with the trajectory of “rapidly progressing symptoms.”

Conclusions: The trajectories of depressive symptoms in middle-aged and older people have 4 types with distinct patterns. Hearing loss is associated with the progression of depressive symptoms, and its impact is more significant for males, affecting not only symptom severity but also progression speed. These findings indicate that the mental health status of middle-aged and older people with hearing loss requires careful consideration, and timely interventions should be provided.

(JMIR Aging 2025;8:e75545) doi:10.2196/75545

KEYWORDS

depression symptom; latent growth mixture model; trajectory analysis; China Health and Retirement Longitudinal Study; CHARLS

Introduction

Depression is a common mental illness that may seriously damage human health, and it is usually caused by the complex interactions between social, psychological, and biological factors. Various life difficulties, such as childhood trauma, loss of loved ones, unemployment, and chronic diseases such as arthritis, asthma, cardiovascular disease, and cancer, can lead to or exacerbate depressive symptoms. The World Health Organization [1] reported that approximately 5% of adults

XSL•FO
RenderX

worldwide experience depressive disorders. The consequences of depression may affect every aspect of life, cause significant mental distress, and, in some cases, even lead to suicide. According to the 2017 Global Burden of Diseases study [2], depression now ranks as the third leading cause of years lived with disability. An epidemiological study [3] reported that the lifetime prevalence of depression in Chinese adults was 8% for female individuals and 5.7% for male individuals, and the 12-month depression prevalence was 4.2% for female individuals and 3% for male individuals. Among individuals aged 50 years and above, lifetime depression prevalence ranged from 7.3% to 7.8%, and 12-month depression prevalence ranged from 3.8% to 4.1%.

Most existing studies on depression [4-6] used the total score on an assessment scale as a quantitative measure to determine the severity level of depression or depressive symptoms of an individual, and this approach treats the population as a homogeneous group with a common pattern of depressive symptom progression. However, some studies have found heterogeneity in depressive symptoms among different groups, which necessitates individual-centered research methods to tailor interventions. Among the existing individual-centered studies on depression, the majority are cross-sectional and use the severity of depressive symptoms or specific symptoms as features to categorize patients [7-9]. In practice, the longitudinal patterns of depressive symptoms can reflect the heterogeneity of depression more comprehensively and objectively than the patterns identified using cross-sectional data, and the trajectories of depressive symptoms have gradually attracted research attention in recent years [10-12].

To explore the trajectories of or changes in depressive symptoms over time, latent growth modeling approaches have been used to identify unobserved latent groups within different study populations [12-14]. Latent growth modeling approaches consider individuals to be part of a heterogeneous population consisting of unobserved groups with similar development patterns [13,15]. One of the primary methods for latent growth modeling is latent growth mixture modeling (LGMM) [16], which enables the modeling of within-class heterogeneity [14], thereby facilitating a more accurate identification of the underlying trajectories of disease progression among different individuals. For example, Maccallum et al [16] used the LGMM method to portray depression trajectories following spousal and child bereavement using the Health and Retirement Study, resulting in the identification of 4 distinct types of depression: resilience (little or no depression; 68.2%), chronic grief (an onset of depression following loss; 13.2%), depressed-improved (high preloss depression that decreased following loss; 11.2%), and preexisting chronic depression (high depression at all assessments; 7.4%). Hong et al [17] used LGMM to analyze perinatal depression in Chinese perinatal women based on data from the First Affiliated Hospital of Wenzhou Medical University and identified 2 types of perinatal depressive symptom trajectory, including “decreasing” (95.3%) and “increasing” (4.7%). However, studies about the trajectories of depressive symptoms among middle-aged and older individuals in China are limited.

Hearing, a vital sensory system in humans, is crucial in daily life. Hearing loss profoundly impacts both physical and mental well-being, leading to communication challenges, cognitive decline, limited social interactions, social isolation, financial stress, and a significant reduction in quality of life [18-20]. China’s second national disability survey, conducted in 2006 [21], reported a prevalence of hearing impairment of 2.11%. Hearing loss increases with age, reaching 11.04% among those aged 60 years and above. In the stress process framework, hearing loss is considered a chronic strain, resulting in ongoing, long-term difficulties that negatively impact health and well-being [20]. It may give rise to the accumulation of additional stressors, including stigma and discrimination; poor mental health; and low levels of educational attainment, income, and employment [22].

The associations between hearing loss and depressive symptoms have been explored in some studies. Lee and Hong [23] found that hearing loss significantly contributed to higher rates of depression among solitary females in South Korea. Blay et al [24] investigated the factors of depression among community-dwelling individuals aged 60 years and above in Brazil and observed a significant correlation between hearing loss and depression. A multicenter study of adults aged 50 years or more in the United States also found that among those with hearing loss, the 10-item Center for Epidemiologic Studies Depression (CES-D10) scale score increased by 0.62 points (95% CI 0.23-1.01) for every 10 dB decrease in hearing [25]. Brewster et al [10] also found that age-related hearing loss is associated with increased depressive symptoms in older adults. Another study explored the associations between hearing loss and depressive symptoms in married couples and found that for women, their hearing loss was associated with increased depressive symptoms [26]. Some studies have reported different findings. For example, a study involving 2890 Norwegians aged 60 years and above found a significant association between hearing loss and depression in the baseline cross-sectional data; however, the longitudinal association was not significant over the 6-year follow-up period [27].

Hearing loss may lead to communication difficulties, which can subsequently result in emotional isolation and reduced social participation [28]. In fact, diminished social engagement is a risk factor for depression [29], and recent research shows that about 9% of the association between poor hearing and depressive symptoms could be explained by social isolation. Therefore, social isolation may play a mediating role in the relationship between hearing loss and depression [30]. Some other studies also suggest that there may be neuropsychological mechanisms underlying the association between hearing loss and depressive symptoms [31,32].

As the studies on the longitudinal relationship between hearing loss and depressive symptoms are limited in the literature, we aimed to investigate the depressive symptom trajectories in the middle-aged and older Chinese population first, with a nationally representative large-scale dataset in this study. Furthermore, we explored the relationship between hearing loss and the depressive symptom trajectories. Additionally, we investigated the potential mediating role of social engagement in the

XSL•FO
RenderX

relationship between hearing loss and the trajectories of depressive symptoms.

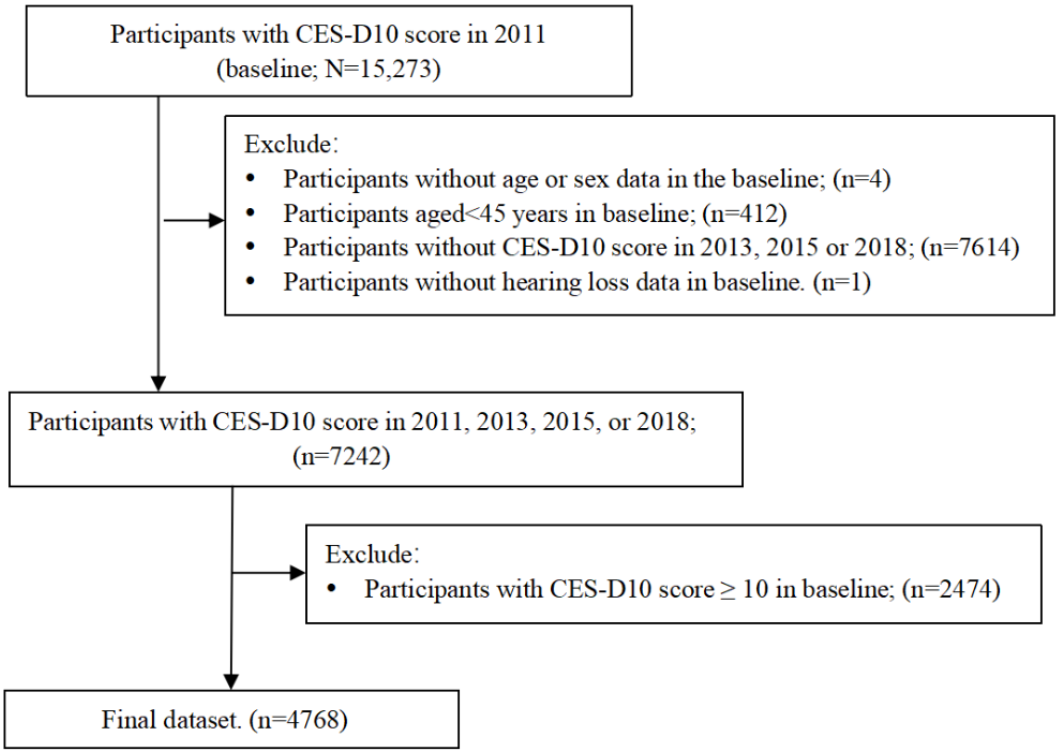
Methods

Data Source and Study Population

The China Health and Retirement Longitudinal Study (CHARLS) [33], a nationally representative longitudinal survey of the middle-aged and older population in China, served as the data source. In the CHARLS, information was collected from participants through face-to-face interviews on demographics, physical and mental health status, health-related behaviors and lifestyles, household economic situation, and health insurance. To ensure the representativeness of the samples, the CHARLS used a multistage probability sampling technique to select eligible samples, considering regional and socioeconomic differences. Detailed descriptions of the sampling methods can be found in previous publications [33,34].

In total, 15,273 participants involved in the 2011 survey have completed the CES-D10 scale. We then excluded 4 participants without gender or age data; 412 participants with a baseline age below 45 years; 7614 participants who did not participate in the follow-up survey in 2013, 2015, or 2018; and 1 person who had no data on hearing at baseline, resulting in 7242 participants as the potential study population. Furthermore, 2474 participants with depressive symptoms (a CES-D10 score of ≥10) in 2011 were excluded, and ultimately, 4768 participants with complete information were included in the analysis. The participant selection process is shown in Figure 1.

Figure 1. Flowchart of the participant selection. CES-D10: 10-item Center for Epidemiologic Studies Depression.



Measures

Depressive Symptoms

The CES-D, developed by Radloff [35], is one of the most commonly used tools for assessing the severity levels of depressive symptoms. It consists of 20 questions, but the response rate is often low due to the large number of questions and the long time needed to answer them. Andresen et al [36] simplified the CES-D and proposed a 10-question version, known as CES-D10, which has been demonstrated to have good reliability and validity in the Chinese population [37].

The depressive symptoms of participants in the CHARLS were measured using the CES-D10, and the detailed questions in the scale are shown in Multimedia Appendix 1. Participants were asked about how they felt and behaved during the last week before the interview, and the response for each question can be selected from the following four choices: (1) very little or not at all (<1 day); (2) not too much (1-2 days); (3) sometimes, or about half the time (3-4 days); and (4) most of the time (5-7 days).

In this study, participants' answers to the negative questions ranged from "seldom or not at all" to "most of the time" on a scale of 0 to 3. For the positive questions, responses ranging from "rarely or not at all" to "most of the time" were scored from 3 to 0. The total score was calculated by summing the scores of all questions, which ranged from 0 to 30, with higher scores indicating more severe depressive symptoms. According to previous studies [38], we defined a total CES-D10 score of 10 or greater as indicative of depressive symptoms in this study.

Hearing Loss

In this study, hearing loss was defined through participants' self-reports. The CHARLS questionnaire related to hearing is presented in Multimedia Appendix 2, and hearing loss was

XSL•FO
RenderX

defined as self-reported hearing loss or partial deafness, typically characterized by the use of a hearing aid or a poor hearing status.

Social Engagement

In the CHARLS, information on social engagement was primarily collected through the question "Have you participated in the following social activities in the past month?" with 12 response options: interacted with a friend; played ma-jong, chess, or cards; went to a community club; helped family, friends, or neighbors who do not live with you and did not pay you for the help; went to a sport, social, or other kind of club; took part in a community-related organization; participated in voluntary or charity work; cared for a sick or disabled adult who does not live with you and who did not pay you for the help; attended an educational or training course; invested in stocks; used the internet; other; and none of these. The first 10 activities were selected for analysis. Each activity was coded as 1 (participated) or 0 (did not participate), and the variable of "social engagement" was the total of these 10 items [38].

Confounding Variables

Confounding variables in this study included age, gender, education level, area of residence, marital status, smoking status, drinking status, household income, BMI, and comorbidities, including hypertension, diabetes, and cardiovascular disease (CVD). These confounding variables were selected by referring to the factors associated with depressive symptoms reported in previous studies [23,24]. Among these variables, demographics (age, gender, education level, area of residence, marital status, and household income) and health-related behaviors (smoking status and drinking status) were obtained from the questionnaire. BMI was defined as weight (kg) divided by the square of height (m^2), while weight and height were acquired through physical examination. It was categorized into 4 groups: underweight (<18.5 kg/m^2), normal (18.5-23.9 kg/m^2), overweight (24.0-27.9 kg/m^2), and obese (≥28.0 kg/m^2) according to the criteria set by the Working Group on Obesity in China [39]. Hypertension was defined as a mean systolic blood pressure of ≥140 mm Hg, a mean diastolic blood pressure of ≥90 mm Hg or more, or self-reported hypertension. Diabetes was defined as a fasting plasma glucose level of ≥126 mg/dL, HbA_{1c} concentration of ≥6.5%, or self-reported having physician-diagnosed diabetes. CVD was defined by self-reported stroke, heart attack, coronary heart disease, angina pectoris, congestive heart failure, or other physician-diagnosed heart problems.

Statistical Analysis

The baseline characteristics were analyzed using descriptive statistics, with frequencies and percentages used for categorical variables, the χ^2 test for between-group comparisons, and means (SDs) for continuous variables. Variables included for baseline analysis included age (45-54, 55-64, and ≥65 years); gender (male or female); education level (below high school, high school, or above high school); area of residence (urban or rural); marital status (married and living with a spouse, married and not living with a spouse, separated or divorced, widowed, or unmarried); household income (quintiles 1-5, bottom 20%, 21% to 40%, 41% to 60%, 61% to 80%, and 81% to 100% of income from low to high); smoking status (never smoked, currently smoking, or quit smoking); drinking status (never drank, currently drinking, or quit drinking); BMI (underweight, normal, overweight, or obese); high blood pressure (yes or no); diabetes (yes or no); and CVD (yes or no).

We used the LGMM approach to explore the depressive symptom trajectories of the study population. The model fitting process began with the initial assumption that only one trajectory type existed in the studied population. Subsequently, the number of trajectory types was set to increase gradually until the best-fitting model was found. The performance metrics, including Akaike information criterion, Bayesian information criterion, and sample size–adjusted Bayesian information criterion, were used to identify the best-fitting trajectory model. The smaller the metric values, the better the model is fitted [40]. Some other criteria commonly used for model selection, including a minimum sample size of at least 1% of the total sample for each trajectory category, higher entropy (<0.8), and statistically significant Lo-Mendell-Rubin Likelihood Ratio Test and Bootstrap Likelihood Ratio Test results (P<.001), were also considered [41]. Furthermore, the best-fitting model determined the number of trajectory types.

The associations between hearing loss and depressive symptom trajectories were analyzed using logistic regression, with the abovementioned confounding variables as covariates. In logistic regression modeling, we used any pair among the 4 trajectory patterns to form the binary dependent variable, and thus, 6 regression models were constructed for the association analysis. The odds ratios and 95% CIs generated in the regression analysis were reported. We also analyzed the association between hearing loss and social engagement using logistic regression. Furthermore, mediation analysis was conducted to assess the proportion of the association that was mediated through social engagement.

All the data preprocessing, descriptive statistical analysis, and logistic regression analysis in this study were done using STATA MP17. The "sgmediation" package in STATA was used to perform mediation analysis, and the LGMM analysis was done using Mplus 8.7.

Ethical Considerations

Our study used publicly available anonymized data from the CHARLS and did not require further review. The CHARLS was approved by the Biomedical Ethics Review Committee of Peking University (IRB00001052-11015). The original ethical approval and participant consent included the use of data, and all participants signed an informed consent form before participation.

Results

A total of 4768 participants, with a mean age of 56.42 (SD 7.87) years, were included in the final analysis. The baseline characteristics of the included participants are shown in Table 1.

XSL•FO
RenderX

Table 1. Baseline characteristics of the participants with different depressive symptom trajectories (N=4768).

Variables	Overall	Depressive symptom trajectories				*P* value
		Type 1	Type 2	Type 3	Type 4	
Total, n (%)	4768 (100)	3656 (76.68)	503 (10.55)	467 (9.79)	142 (2.98)	—[a]
Age (y), mean (SD)	56.42 (7.87)	57.11 (7.96)	56.41 (7.92)	56.25 (7.96)	56.48 (7.36)	—
Age groups (y), n (%)						.48
45-54	2069 (43.38)	1595 (43.63)	225 (44.73)	194 (41.54)	54 (38.03)	
55-64	1938 (40.64)	1477 (40.40)	194 (38.57)	206 (44.11)	61 (42.96)	
≥65	762 (15.98)	584 (15.97)	84 (16.7)	67 (14.35)	27 (19.01)	
Sex, n (%)						<.001
Male	2574 (53.97)	2098 (57.39)	220 (43.7)	211 (45.18)	45 (31.69)	
Female	2195 (46.03)	1558 (42.61)	283 (56.2)	256 (54.82)	97 (68.31)	
Education, n (%)						<.001
Below high school	3992 (83.71)	2982 (81.56)	464 (92.25)	412 (88.22)	134 (94.37)	
High school or above	776 (16.27)	674 (18.44)	39 (7.75)	55 (11.78)	7 (4.93)	
Area of residence, n (%)						<.001
Urban	1087 (22.79)	922 (25.22)	66 (13.12)	79 (16.92)	19 (13.38)	
Rural	3680 (77.17)	2732 (74.73)	437 (86.88)	388 (83.08)	123 (86.62)	
Marital status, n (%)						<.001
Married with spouse present	4279 (89.73)	3296 (90.15)	450 (89.76)	419 (89.72)	113 (79.58)	
Married with spouse away	172 (3.61)	129 (3.53)	22 (4.37)	16 (3.43)	5 (3.52)	
Separated, divorced, or widowed	297 (6.23)	223 (6.1)	26 (5.17)	28 (6)	20 (14.08)	
Never married	21 (0.44)	8 (0.22)	5 (0.99)	4 (0.86)	4 (2.82)	
Household income groups, n (%)						.006
Quintile 1 (lowest)	948 (19.88)	684 (18.71)	117 (23.26)	109 (23.34)	38 (26.76)	
Quintile 2	948 (19.89)	714 (19.53)	111 (22.07)	93 (19.91)	30 (21.13)	
Quintile 3	946 (19.84)	720 (19.69)	96 (19.06)	102 (21.84)	28 (19.72)	
Quintile 4	944 (19.82)	744 (20.35)	86 (17.1)	89 (19.06)	25 (17.61)	
Quintile 5 (highest)	946 (19.84)	770 (21.06)	87 (17.3)	69 (14.78)	20 (14.08)	
Smoking, n (%)						.008
Never	2776 (58.21)	2063 (56.43)	327 (65.01)	291 (62.31)	94 (66.20)	
Current	1598 (33.51)	1275 (34.87)	140 (27.83)	142 (30.41)	41 (28.87)	
Past	394 (8.26)	317 (8.67)	36 (7.16)	34 (7.28)	7 (4.93)	
Drinking, n (%)						<.001
Never	2659 (55.76)	1956 (53.50)	317 (63.02)	288 (61.67)	97 (68.31)	
Current	1787 (37.47)	1454 (39.77)	153 (30.42)	147 (31.48)	33 (23.24)	
Past	323 (6.77)	2446 (6.73)	33 (6.56)	32 (6.85)	12 (8.45)	
BMI groups (kg/m^2), n (%)						.47
<18.5	163 (3.42)	121 (3.31)	15 (2.98)	18 (3.85)	8 (5.63)	
18.5-23.9	2079 (43.59)	1586 (43.38)	226 (44.93)	212 (45.40)	55 (38.73)	
24-27.9	1350 (28.31)	1032 (28.23)	141 (28.03)	130 (27.84)	47 (33.1)	
≥28.0	619 (12.98)	467 (12.77)	74 (14.71)	59 (12.63)	19 (13.38)	
Missing	558 (11.7)	450 (12.31)	47 (9.34)	48 (10.28)	13 (9.15)	

XSL•FO
RenderX

Variables	Overall	Depressive symptom trajectories				*P* value
		Type 1	Type 2	Type 3	Type 4	
Hypertension, n (%)						.02
Yes	1063 (22.29)	801 (21.91)	118 (23.46)	110 (23.55)	34 (23.94)	
No	3345 (70.14)	2599 (71.09)	340 (67.59)	317 (67.88)	88 (61.97)	
Diabetes, n (%)						.33
Yes	243 (5.10)	177 (4.84)	27 (5.37)	33 (7.07)	6 (4.23)	
No	4483 (94.00)	3448 (94.31)	472 (93.84)	430 (92.08)	133 (93.66)	
Cardiovascular disease, n (%)						.04
Yes	469 (9.83)	373 (10.2)	37 (7.36)	40 (8.57)	19 (13.38)	
No	4276 (89.66)	3269 (89.41)	460 (91.45)	424 (90.76)	123 (86.62)	
Hearing loss, n (%)						<.001
Yes	465 (9.75)	333 (9.11)	59 (11.73)	45 (9.64)	28 (19.72)	
No	4303 (90.25)	3323 (90.89)	444 (88.27)	422 (90.36)	114 (80.28)	

[a]Not available.

Depressive Symptom Trajectories

In trajectory analysis, we set the number of potential trajectory types to range from 1 to 6 for LGMM modeling, and the performance metrics illustrating the model fitting status of the 6 trajectory models are presented in Table 2. With the increase in the number of trajectory types in LGMM models, the values of the Akaike information criterion, Bayesian information criterion, and sample size–adjusted Bayesian information criterion decreased gradually. Considering the simplicity, accuracy, and practical significance of the developed LGMM models, the model with 4 trajectory types was finally selected as the best fit.

The depressive symptom trajectories of the 4 types (types 1, 2, 3, and 4) are visually displayed in Figure 2. According to the characteristics of each depressive symptom trajectory, we named the 4 trajectories as type 1: “stable low symptoms” (having a steady low CES-D10 score, 74.5%); type 2: “slowly progressing symptoms” (gradually changing from a low to high CES-D10 score, 11.3%); type 3: “relieved symptoms after progression” (with a CES-D10 score growing first and then decreasing, 10.6%); and type 4: “rapidly progressing symptoms” (having a continuously and rapidly increasing CES-D10 score, 3.2%).

The average posterior probability of the most likely latent class membership for participants in each trajectory was calculated for the 4 trajectories, and the results are shown in Table 2. We observed that the average posterior probabilities (diagonal values) were all reasonably high (ranging from 0.779 to 0.939), indicating that the selected trajectory model was acceptable.

XSL•FO
RenderX

Table 2. Model fitting statistics of trajectory analysis.

Model fitting statistics	Number of trajectory types					
	1	2	3	4	5	6
AIC[a]	107,515.41	106,781.01	106,390.71	106,168.70	106,063.23	105,926.99
BIC[b]	107,599.52	106,891.00	106,526.58	106,330.45	106,250.86	106,140.49
Entropy	__[c]	0.82	0.85	0.82	0.77	0.78
LMR-LRT[d], *P* value	—	<.001	<.001	<.001	.26	.33
BLRT[e], *P* value	—	<.001	<.001	<.001	<.001	<.001
Class proportion (%)						
Type 1	100	85.66	82.48	74.51	48.71	45.84
Type 2	—	14.34	8.77	11.60	5.96	4.36
Type 3	—	—	8.75	10.74	35.10	3.62
Type 4	—	—	—	3.16	3.25	35.74
Type 5	—	—	—	—	7.00	5.87
Type 6	—	—	—	—	—	4.57
Average posterior probabilities						
Type 1	1	0.961	0.802	0.789	0.887	0.890
Type 2	—	0.865	0.821	0.779	0.752	0.743
Type 3	—	—	0.957	0.843	0.843	0.724
Type 4	—	—	—	0.938	0.789	0.838
Type 5	—	—	—	—	0.780	0.770
Type 6	—	—	—	—	—	0.753

[a]AIC: Akaike information criterion.

[b]BIC: Bayesian information criterion.

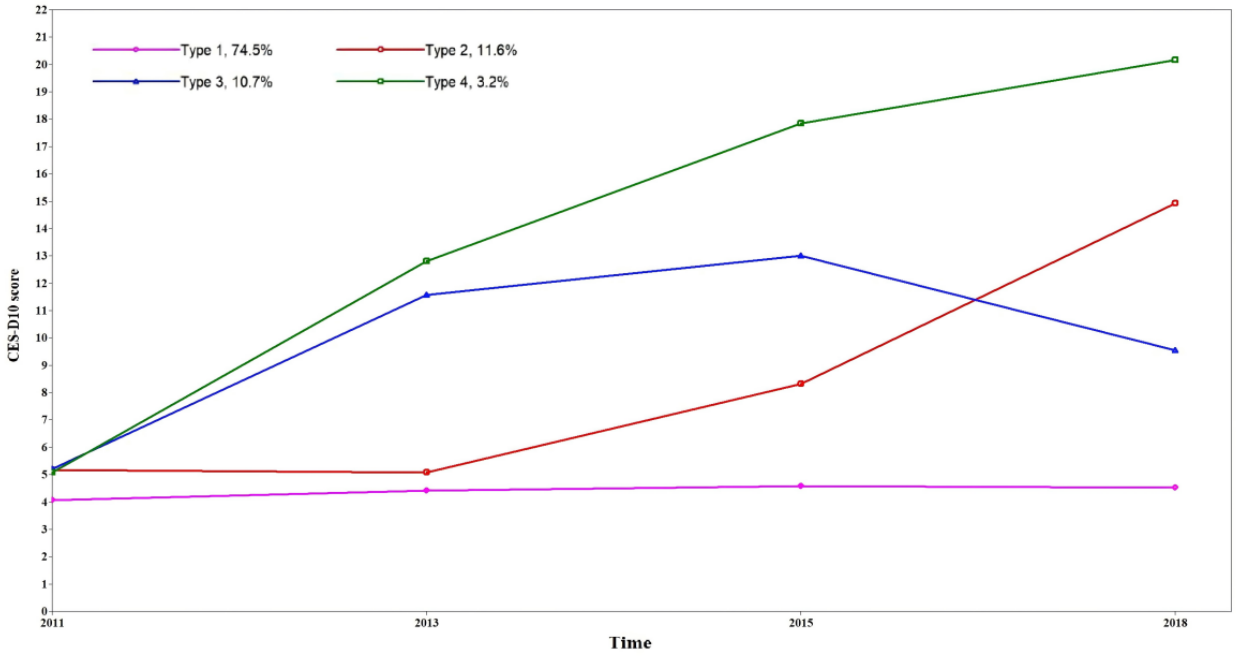
[c]Not available.

[d]LMR-LRT: Lo-Mendell-Rubin Likelihood Ratio Test.

[e]BLRT: Bootstrap Likelihood Ratio Test.

XSL•FO
RenderX

Figure 2. Depressive symptom trajectories for each latent category. CES-D10: 10-item Center for Epidemiologic Studies Depression.



Associations Between Hearing Loss and Depressive Symptom Trajectories

The results of logistic regression analyses for the associations between hearing loss and depressive symptom trajectories are presented in Table 3. Participants with hearing loss had a higher risk of having a "rapidly progressing symptoms" trajectory than the risk of the other 3 trajectories. Then, we conducted a subgroup analysis stratified by gender. Results showed that in female participants with hearing loss, they had a higher risk of having a "rapidly progressing symptoms" trajectory than "stable low symptoms" or "relieved symptoms after progression" trajectories, but the risks of having "rapidly progressing symptoms" and "slowly progressing symptoms" showed no difference. In male participants with hearing loss, the risk of having a "rapidly progressing symptoms" trajectory was still higher with statistical significance (P=.02) than the other 3 depressive symptom trajectories, which suggested that hearing loss may have a more substantial impact on depressive symptoms in male individuals.

XSL•FO
RenderX

Table 3. Odds ratios and 95% CIs for different depressive symptom trajectories associated with baseline hearing loss.

Predictors and depressive symptom trajectories						
Type 1	Type 2		Type 3		Type 4	
	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Hearing loss in men and women (N=4768)[a]						
Ref[b]	1.33 (0.99-1.80)	.06	1.07 (0.76-1.49)	.70	2.41 (1.54-3.76)	<.001
—[c]	Ref	—	0.80 (0.52-1.21)	.29	1.92 (1.14-3.23)	.01
—	—	—	Ref	—	2.49 (1.45-4.27)	<.001
Hearing loss in men (n=2574)[d]						
Ref	1.20 (0.79-1.82)	.40	1.12 (0.72-1.75)	.60	2.67 (1.29-5.50)	.008
—	Ref	—	0.93 (0.52-1.68)	.82	2.86 (1.17-6.98)	.02
—	—	—	Ref	—	2.76 (1.16-6.56)	.02
Hearing loss in women (n=2194)[d]						
Ref	1.51 (0.98-2.34)	.06	0.98 (0.59-1.64)	.94	2.38 (1.34-4.23)	.003
—	Ref	—	0.71 (0.38-1.30)	.27	1.90 (0.97-3.73)	.06
—	—	—	Ref	—	2.58 (1.25-5.34)	.01

[a]Adjusted for age, sex, education level, area of residence, marital status, household income, smoking status, drinking status, BMI, hypertension, diabetes, and cardiovascular disease.

[b]Ref: reference.

[c]Not available.

[d]Adjusted for age, education level, area of residence, marital status, household income, smoking status, drinking status, BMI, hypertension, diabetes, and cardiovascular disease.

The Role of Social Engagement

Logistic regression analyses revealed a significant association between hearing loss and reduced social engagement after adjusting for confounding variables (OR 0.79, 95% CI 0.62-0.99). Concurrently, the risk of exhibiting a rapidly progressing symptoms trajectory (vs stable low symptoms) after adjustment for social engagement (OR 2.19, 95% CI 1.30-3.70) was lower than before (OR 2.49, 95% CI 1.45-4.27). However, social engagement did not demonstrate a significant mediating effect in the relationship between hearing loss and depressive symptom trajectories. This suggests that the influence of hearing loss on depressive symptom progression may involve other factors, such as neurophysiological mechanisms. The conclusion remained consistent in gender-stratified analysis.

Discussion

Principal Findings

On the basis of a nationally representative dataset of middle-aged and older Chinese adults without depressive symptoms at baseline (2011), we used the LGMM approach to perform a depressive symptom trajectory analysis in this study, and identified 4 trajectory types: type 1, “stable low symptoms”; type 2, “slowly progressing symptoms”; type 3, “relieved symptoms after progression”; and type 4, “rapidly progressing symptoms.” Furthermore, we examined the impact of hearing loss on the depressive symptom trajectories. We found that hearing loss was significantly associated with the “rapidly progressing symptoms” trajectory, and the impact was more substantial in males. The association between hearing loss and rapidly progressing depressive symptoms was attenuated after adjustment for social engagement, but no significant evidence was found to support a mediating role of social isolation in the association between hearing loss and depressive symptom trajectories.

Comparison With Prior Work

The findings regarding the significant association between hearing loss and the rapid progression of depressive symptoms in this study are consistent with the conclusions of some existing studies. For instance, Simning et al [42] examined the relationship between hearing loss and depression in a study encompassing a US population aged 65 years and older, and found that hearing loss was significantly linked not only to the onset of depressive symptoms within 1 year but also to the risk of persistent depressive symptoms. Kiely et al [43] found that individuals with hearing loss exhibited more severe depressive symptoms compared to those without sensory impairments in their examination of 16 years of longitudinal data gathered by the Australian Longitudinal Study of Aging. Nevertheless, a few other studies have different findings. For example, Cosh et al [27] found a significant association between hearing loss and depression in cross-sectional data at baseline in a study involving 2890 Norwegians aged 60 years and older but did not observe a longitudinal association in the 6-year follow-up. This inconsistency in research findings may be attributed to variations in the criteria used to define hearing loss and depressive symptoms. In addition, we observed that there was a gender

XSL•FO
RenderX

difference in the impact of hearing loss on depressive symptom trajectories. While the risk of having a "rapidly progressing symptoms" trajectory was significantly higher than the risk of the 3 other depressive symptom trajectories in male participants with hearing loss, the risk of having a "rapidly progressing symptoms" or "slowly progressing symptoms" trajectory did not differ in female participants with hearing loss. This means that hearing loss may impact both the incidence and progression speed of depressive symptoms in males, but may not affect the depressive symptom progression speed in female individuals. Previous studies also found that the association between hearing loss and the risk of depressive symptoms was more significant in male individuals. These studies suggest that communication barriers resulting from hearing impairment, along with consequent social isolation, may exert a more pronounced effect in male individuals by exacerbating their already restricted social interactions [44-46].

It is generally believed that the mechanism behind this association operates through social isolation. Hearing problems lead to communication difficulties, making social interactions challenging, causing people to withdraw from social activities, and consequently generating feelings of loneliness and depression [47]. Studies indicate that after adjusting for social contact, the association between hearing loss and depressive symptoms weakened and became nonsignificant [43]. Research by Huang et al [30] showed that in mediation effect analysis, the proportion of the association mediated by social isolation was 9%.

However, in this study, after adjusting for social engagement, the association between hearing loss and depressive symptom trajectories remained significant, but no significant mediating effect was found. A study by Keidser and Seeto [47] demonstrated an independent association between hearing loss and depressive episodes, with social isolation having minimal influence on this association, which aligns with our findings. This outcome may be due to the fact that the social participation variable in this study relates to the quantity rather than the satisfaction of social interactions, and hearing-impaired individuals may not have withdrawn from social activities but rather become unable to engage in effective communication [48]. Additionally, our study population primarily consisted of middle-aged and older adults who may have adapted to changes in hearing conditions, with their communication skills having been changed or improved. Hawthorne [49] also found that the proportion of social isolation was higher among young people (ages 15-30 years) than among older adults (>60 years). It may be inferred that factors, such as social engagement, are more likely to mediate the association between hearing problems and depression in young and middle-aged adults compared to older adults. Furthermore, this study did not exclude hearing aid users. Older adults may use hearing aids to improve their hearing conditions, potentially mitigating the negative impacts of hearing loss [30].

Social engagement cannot fully explain the association between hearing loss and depressive symptoms, and potential neuropsychological mechanisms may exist. Neuroimaging studies have revealed similar alterations in patients with hearing loss and individuals with depression [50,51]. Additionally, research indicates that responses to emotional sounds in the amygdala and parahippocampal gyrus are diminished among hearing-impaired individuals [31]. However, the precise neuropsychological mechanisms remain unclear and warrant further investigation.

Our study demonstrates that hearing loss constitutes a risk factor for the rapid progression of depressive symptoms, warranting increased attention from clinicians and psychologists. It is imperative to enhance monitoring of mental health status among middle-aged and older adults with hearing impairment while proactively implementing auditory rehabilitation interventions, such as hearing aids, cochlear implants, or virtual reality cognitive–based interventions to mitigate the impact of hearing loss on patient well-being [52].

Strengths and Limitations

A strength of this study was a nationally representative, large-scale dataset with a high response rate and a rigorous quality control process, ensuring the reliability of the findings generated from the dataset. Furthermore, the 8-year follow-up period (2011-2018) in this study was sufficient to identify distinct and clinically meaningful trajectories of depressive symptoms, as many studies on depression trajectories have used similar follow-up periods [53-55]. The trajectory analysis showed 4 distinct categories, indicating that an 8-year follow-up period can provide a certain degree of generalization of the development trajectory of patients' depressive symptoms.

However, this study still has its limitations. First, the definition of hearing loss in this study was based solely on participants' self-reports, rather than results derived from clinical measurements [54]. The clinical hearing test was conducted without assistance, whereas self-reported hearing may reflect the hearing experience with the help of a hearing aid in daily life. This study did not exclude hearing aid users, and older adults may use hearing aids to improve their hearing conditions, potentially mitigating the negative impacts of hearing loss. Nevertheless, the objective measurement of hearing correlates well with the subjective sensory assessment [56]. Moreover, it has been suggested that self-reported hearing loss may better reflect the extent to which patients' daily lives are affected and may be more valuable when exploring the relationship between hearing loss and mental health status [20,57,58]. Second, some factors, such as psychiatric and medication history, were not available in the CHARLS dataset, and this may impact the analysis results. Third, due to the use of a preexisting national database, this study was limited by the inability to incorporate qualitative data, such as in-depth interviews, and insights into potential mechanisms were then supplemented through a literature review and mediation analysis of social engagement measures. Fourth, considering individuals who already have depression symptoms at baseline are likely to have received interventions or developed adaptive mechanisms, which could substantially alter the natural progression of depressive symptoms and confound the estimated effect of hearing loss, we excluded participants with baseline depression symptoms. However, this exclusion may limit the generalization of our findings to all depression trajectory types, and future studies incorporating participants across all baseline depression statuses

XSL•FO
RenderX

would be valuable to capture a fuller spectrum of symptom trajectories.

Conclusions

To summarize, this study was the first to analyze the trajectories of depressive symptoms in Chinese middle-aged and older adults without depressive symptoms at baseline, and further explored the association between hearing loss and depressive symptom trajectories. Findings of this study showed that the depressive symptom trajectories in Chinese middle-aged and older adults could be categorized into 4 types, including "stable low symptoms," "slowly progressing symptoms," "relieved symptoms after progression," and "rapidly progressing symptoms." Hearing loss was significantly associated with the "rapidly progressing symptoms" trajectory among the study population, and the impact was stronger for males. It suggests that more attention should be paid to the mental status of Chinese middle-aged and older adults with hearing loss, and if possible, appropriate interventions should be taken to prevent the onset of depressive symptoms.

Acknowledgments

This work was supported by grants from the Humanities and Social Science Project of the Ministry of Education of China (grant 22YJA630036), the National Natural Science Foundation of China (grant 82372095), the Zhejiang Provincial Natural Science Foundation of China (grant LZ22F020014), and the Beijing Natural Science Foundation (grant 7212201). The funders had no role in the study design, participant recruitment, data collection, data analysis, or paper preparation.

The authors thank the China Health and Retirement Longitudinal Study team for providing the data.

Data Availability

The data that support the findings of this study are available from the China Health and Retirement Longitudinal Study [59].

Conflicts of Interest

None declared.

Multimedia Appendix 1
The 10-item Center for Epidemiologic Studies Depression (CES-D10) scale.
[DOCX File , 16 KB - aging_v8i1e75545_app1.docx]

Multimedia Appendix 2
Survey questions related to hearing in the China Health and Retirement Longitudinal Study.
[DOCX File , 16 KB - aging_v8i1e75545_app2.docx]

References

1. Depression. World Health Organization. URL: https://www.who.int/health-topics/depression#tab=tab_1 [accessed 2023-12-13]
2. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1789-1858 [FREE Full text] [doi: 10.1016/S0140-6736(18)32279-7] [Medline: 30496104]
3. Lu J, Xu X, Huang Y, Li T, Ma C, Xu G, et al. Prevalence of depressive disorders and treatment in China: a cross-sectional epidemiological study. Lancet Psychiatry 2021 Nov;8(11):981-990. [doi: 10.1016/S2215-0366(21)00251-0] [Medline: 34559991]
4. Butcher E, Cortina-Borja M, Dezateux C, Knowles R. The association between childhood hearing loss and self-reported peer victimisation, depressive symptoms, and self-harm: longitudinal analyses of a prospective, nationally representative cohort study. BMC Public Health 2022 May 25;22(1):1045 [FREE Full text] [doi: 10.1186/s12889-022-13457-6] [Medline: 35614427]
5. Lisan Q, van Sloten TT, Lemogne C, Offredo L, Climie RE, Boutouyrie P, et al. Association of hearing impairment with incident depressive symptoms: a community-based prospective study. Am J Med 2019 Dec;132(12):1441-9.e4. [doi: 10.1016/j.amjmed.2019.05.039] [Medline: 31247178]
6. Wu C. Bidirectional association between depression and hearing loss: evidence from the China health and retirement longitudinal study. J Appl Gerontol 2022 Apr 06;41(4):971-981. [doi: 10.1177/07334648211042370] [Medline: 34486422]
7. Ferdinand RF, de Nijs PF, van Lier P, Verhulst FC. Latent class analysis of anxiety and depressive symptoms in referred adolescents. J Affect Disord 2005 Nov;88(3):299-306. [doi: 10.1016/j.jad.2005.08.004] [Medline: 16182373]
8. Putnam KT, Wilcox M, Robertson-Blackmore E, Sharkey K, Bergink V, Munk-Olsen T, et al. Clinical phenotypes of perinatal depression and time of symptom onset: analysis of data from an international consortium. Lancet Psychiatry 2017 Jun;4(6):477-485 [FREE Full text] [doi: 10.1016/S2215-0366(17)30136-0] [Medline: 28476427]

XSL•FO
RenderX

9. Zhang J, Liu D, Ding L, Du G. Prevalence of depression in junior and senior adolescents. Front Psychiatry 2023 Dec 11;14:1182024 [FREE Full text] [doi: 10.3389/fpsyt.2023.1182024] [Medline: 38152357]
10. Brewster KK, Ciarleglio A, Brown PJ, Chen C, Kim HO, Roose SP, et al. Age-related hearing loss and its association with depression in later life. Am J Geriatr Psychiatry 2018 Jul;26(7):788-796 [FREE Full text] [doi: 10.1016/j.jagp.2018.04.003] [Medline: 29752060]
11. Huang W, Zhu W, Chen H, Li F, Huang J, Zhou Y, et al. Longitudinal association between depressive symptoms and cognitive decline among middle-aged and elderly population. J Affect Disord 2022 Apr 15;303:18-23. [doi: 10.1016/j.jad.2022.01.107] [Medline: 35108603]
12. Chen X, Lin S, Zheng Y, He L, Fang Y. Long-term trajectories of depressive symptoms and machine learning techniques for fall prediction in older adults: evidence from the China Health and Retirement Longitudinal Study (CHARLS). Arch Gerontol Geriatr 2023 Aug;111:105012. [doi: 10.1016/j.archger.2023.105012] [Medline: 37030148]
13. Jung T, Wickrama KA. An introduction to latent class growth analysis and growth mixture modeling. Soc Personal Psychol Compass 2008 Jan;2(1):302-317. [doi: 10.1111/j.1751-9004.2007.00054.x]
14. Nguena Nguefack HL, Pagé MG, Katz J, Choinière M, Vanasse A, Dorais M, et al. Trajectory modelling techniques useful to epidemiological research: a comparative narrative review of approaches. Clin Epidemiol 2020;12:1205-1222 [FREE Full text] [doi: 10.2147/CLEP.S265287] [Medline: 33154677]
15. Nagin DS, Odgers CL. Group-based trajectory modeling in clinical research. Annu Rev Clin Psychol 2010;6:109-138. [doi: 10.1146/annurev.clinpsy.121208.131413] [Medline: 20192788]
16. Maccallum F, Galatzer-Levy IR, Bonanno GA. Trajectories of depression following spousal and child bereavement: a comparison of the heterogeneity in outcomes. J Psychiatr Res 2015 Oct;69:72-79. [doi: 10.1016/j.jpsychires.2015.07.017] [Medline: 26343597]
17. Hong L, Le T, Lu Y, Shi X, Xiang L, Liu M, et al. Distinct trajectories of perinatal depression in Chinese women: application of latent growth mixture modelling. BMC Pregnancy Childbirth 2022 Jan 10;22(1):24 [FREE Full text] [doi: 10.1186/s12884-021-04316-0] [Medline: 35012496]
18. Guthrie DM, Davidson JG, Williams N, Campos J, Hunter K, Mick P, et al. Combined impairments in vision, hearing and cognition are associated with greater levels of functional and communication difficulties than cognitive impairment alone: analysis of interRAI data for home care and long-term care recipients in Ontario. PLoS One 2018 Feb 15;13(2):e0192971 [FREE Full text] [doi: 10.1371/journal.pone.0192971] [Medline: 29447253]
19. Mick P, Kawachi I, Lin FR. The association between hearing loss and social isolation in older adults. Otolaryngol Head Neck Surg 2014 Mar 02;150(3):378-384. [doi: 10.1177/0194599813518021] [Medline: 24384545]
20. Pearlin LI. The sociological study of stress. J Health Soc Behav 1989 Sep;30(3):241-256. [Medline: 2674272]
21. Sun XB, Wei ZY, Yu LM, Wang Q, Liang W. Prevalence and etiology of people with hearing impairment in China. Zhonghua Liu Xing Bing Xue Za Zhi 2008 Jul;29(7):643-646. [Medline: 19031751]
22. West JS. Hearing impairment, social support, and depressive symptoms among U.S. adults: a test of the stress process paradigm. Soc Sci Med 2017 Nov;192:94-101. [doi: 10.1016/j.socscimed.2017.09.031] [Medline: 28963989]
23. Lee S, Hong GR. Predictors of depression among community-dwelling older women living alone in Korea. Arch Psychiatr Nurs 2016 Oct;30(5):513-520. [doi: 10.1016/j.apnu.2016.05.002] [Medline: 27654230]
24. Blay SL, Andreoli SB, Fillenbaum GG, Gastal FL. Depression morbidity in later life: prevalence and correlates in a developing country. Am J Geriatr Psychiatry 2007 Sep;15(9):790-799. [doi: 10.1097/JGP.0b013e3180654179] [Medline: 17698602]
25. Golub JS, Brewster KK, Brickman AM, Ciarleglio AJ, Kim AH, Luchsinger JA, et al. Subclinical hearing loss is associated with depressive symptoms. Am J Geriatr Psychiatry 2020 May;28(5):545-556 [FREE Full text] [doi: 10.1016/j.jagp.2019.12.008] [Medline: 31980375]
26. West JS, Smith SL, Dupre ME. The impact of hearing loss on trajectories of depressive symptoms in married couples. Soc Sci Med 2023 Mar;321:115780 [FREE Full text] [doi: 10.1016/j.socscimed.2023.115780] [Medline: 36801754]
27. Cosh S, von Hanno T, Helmer C, Bertelsen G, Delcourt C, Schirmer H. The association amongst visual, hearing, and dual sensory loss with depression and anxiety over 6 years: the Tromsø Study. Int J Geriatr Psychiatry 2018 Apr 28;33(4):598-605 [FREE Full text] [doi: 10.1002/gps.4827] [Medline: 29193338]
28. Dawes P, Emsley R, Cruickshanks KJ, Moore DR, Fortnum H, Edmondson-Jones M, et al. Hearing loss and cognition: the role of hearing AIDS, social isolation and depression. PLoS One 2015 Mar 11;10(3):e0119616 [FREE Full text] [doi: 10.1371/journal.pone.0119616] [Medline: 25760329]
29. Elmer T, Stadtfeld C. Depressive symptoms are associated with social isolation in face-to-face interaction networks. Sci Rep 2020 Jan 29;10(1):1444 [FREE Full text] [doi: 10.1038/s41598-020-58297-9] [Medline: 31996728]
30. Huang H, Wang J, Jiang CQ, Zhu F, Jin YL, Zhu T, et al. Hearing loss and depressive symptoms in older Chinese: whether social isolation plays a role. BMC Geriatr 2022 Jul 26;22(1):620 [FREE Full text] [doi: 10.1186/s12877-022-03311-0] [Medline: 35883170]
31. Husain FT, Carpenter-Thompson JR, Schmidt SA. The effect of mild-to-moderate hearing loss on auditory and emotion processing networks. Front Syst Neurosci 2014;8:10 [FREE Full text] [doi: 10.3389/fnsys.2014.00010] [Medline: 24550791]

XSL•FO
RenderX

32. Brewster KK, Golub JS, Rutherford BR. Neural circuits and behavioral pathways linking hearing loss to affective dysregulation in older adults. Nat Aging 2021 May 14;1(5):422-429 [FREE Full text] [doi: 10.1038/s43587-021-00065-z] [Medline: 37118018]
33. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). Int J Epidemiol 2014 Feb 12;43(1):61-68 [FREE Full text] [doi: 10.1093/ije/dys203] [Medline: 23243115]
34. Feng XL, Pang M, Beard J. Health system strengthening and hypertension awareness, treatment and control: data from the China Health and Retirement Longitudinal Study. Bull World Health Organ 2014;92:29-41 [FREE Full text] [doi: 10.2471/blt.13.124495]
35. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977;1(3):385-401. [doi: 10.1177/014662167700100306]
36. Andresen EM, Malmgren JA, Carter WB, Patrick DL. Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). Am J Prev Med 1994;10(2):77-84. [Medline: 8037935]
37. Chang SH, Chen MC, Chien NH, Lin TY, Chang YY. Gender differences in the relationship of lifestyle and depressive symptoms among Taiwanese older people. Collegian 2016 Sep;23(3):249-255 [FREE Full text] [doi: 10.1016/j.colegn.2015.04.005]
38. Liu J, Rozelle S, Xu Q, Yu N, Zhou T. Social engagement and elderly health in China: evidence from the China Health and Retirement Longitudinal Survey (CHARLS). Int J Environ Res Public Health 2019 Jan 18;16(2):278 [FREE Full text] [doi: 10.3390/ijerph16020278] [Medline: 30669415]
39. Zhou BF, Cooperative Meta-Analysis Group of the Working Group on Obesity in China. Predictive values of body mass index and waist circumference for risk factors of certain related diseases in Chinese adults--study on optimal cut-off points of body mass index and waist circumference in Chinese adults. Biomed Environ Sci 2002 Mar;15(1):83-96. [Medline: 12046553]
40. Mara CA, Carle AC. Understanding variation in longitudinal data using latent growth mixture modeling. J Pediatr Psychol 2021 Feb 19;46(2):179-188. [doi: 10.1093/jpepsy/jsab010] [Medline: 33609037]
41. Wang F, Mei X. Association of blood glucose change with postoperative delirium after coronary artery bypass grafting in patients with diabetes mellitus: a study of the MIMIC-IV database. Front Endocrinol (Lausanne) 2024 Jun 20;15:1400207 [FREE Full text] [doi: 10.3389/fendo.2024.1400207] [Medline: 38966222]
42. Simning A, Fox ML, Barnett SL, Sorensen S, Conwell Y. Depressive and anxiety symptoms in older adults with auditory, vision, and dual sensory impairment. J Aging Health 2019 Sep 13;31(8):1353-1375 [FREE Full text] [doi: 10.1177/0898264318781123] [Medline: 29896982]
43. Kiely KM, Anstey KJ, Luszcz MA. Dual sensory loss and depressive symptoms: the importance of hearing, daily functioning, and activity engagement. Front Hum Neurosci 2013;7:837 [FREE Full text] [doi: 10.3389/fnhum.2013.00837] [Medline: 24379769]
44. Harada S, Nishiwaki Y, Michikawa T, Kikuchi Y, Iwasawa S, Nakano M, et al. Gender difference in the relationships between vision and hearing impairments and negative well-being. Prev Med 2008 Oct;47(4):433-437. [doi: 10.1016/j.ypmed.2008.06.011] [Medline: 18619483]
45. Lyu J, Kim HY. Gender-specific associations of sensory impairments with depression and cognitive impairment in later life. Psychiatry Investig 2018 Sep;15(10):926-934 [FREE Full text] [doi: 10.30773/pi.2018.06.28.2] [Medline: 30205671]
46. Liu W, Yang C, Liu L, Kong G, Zhang L. Bidirectional associations of vision loss, hearing loss, and dual sensory loss with depressive symptoms among the middle-aged and older adults in China. J Affect Disord 2022 Mar 15;301:225-232. [doi: 10.1016/j.jad.2022.01.066] [Medline: 35038482]
47. Keidser G, Seeto M. The influence of social interaction and physical health on the association between hearing and depression with age and gender. Trends Hear 2017 Jul 28;21:2331216517706395 [FREE Full text] [doi: 10.1177/2331216517706395] [Medline: 28752806]
48. Routasalo PE, Savikko N, Tilvis RS, Strandberg TE, Pitkälä KH. Social contacts and their relationship to loneliness among aged people - a population-based study. Gerontology 2006;52(3):181-187. [doi: 10.1159/000091828] [Medline: 16645299]
49. Hawthorne G. Perceived social isolation in a community sample: its prevalence and correlates with aspects of peoples' lives. Soc Psychiatry Psychiatr Epidemiol 2008 Feb 9;43(2):140-150. [doi: 10.1007/s00127-007-0279-8] [Medline: 17994175]
50. Husain FT, Medina RE, Davis CW, Szymko-Bennett Y, Simonyan K, Pajor NM, et al. Neuroanatomical changes due to hearing loss and chronic tinnitus: a combined VBM and DTI study. Brain Res 2011 Jan 19;1369:74-88 [FREE Full text] [doi: 10.1016/j.brainres.2010.10.095] [Medline: 21047501]
51. Hickie I, Naismith S, Ward PB, Turner K, Scott E, Mitchell P, et al. Reduced hippocampal volumes and memory loss in patients with early- and late-onset depression. Br J Psychiatry 2005 Mar 02;186(3):197-202. [doi: 10.1192/bjp.186.3.197] [Medline: 15738499]
52. Makmee P, Wongupparaj P. VR cognitive-based intervention for enhancing cognitive functions and well-being in older adults with mild cognitive impairment: behavioral and EEG evidence. Psychosoc Interv 2025 Jan;34(1):37-51. [doi: 10.5093/pi2025a4] [Medline: 39781014]

XSL•FO
RenderX

53. Liu Y, Li W, Chen Z, He M, Zhang W, Wei Y, et al. Relationship between sarcopenia and the trajectories of depressive symptoms among Chinese older adults: the mediating effect of social participation. Int J Behav Med 2025 Apr 30. [doi: 10.1007/s12529-025-10366-x] [Medline: 40307534]
54. Liu L, Li X, Marshall IJ, Bhalla A, Wang Y, O'Connell MD. Trajectories of depressive symptoms 10 years after stroke and associated risk factors: a prospective cohort study. Lancet 2023 Nov;402 Suppl 1:S64. [doi: 10.1016/S0140-6736(23)02111-6] [Medline: 37997108]
55. Cheng Y, Thorpe L, Kabir R, Lim HJ. Latent class growth modeling of depression and anxiety in older adults: an 8-year follow-up of a population-based study. BMC Geriatr 2021 Oct 13;21(1):550 [FREE Full text] [doi: 10.1186/s12877-021-02501-6] [Medline: 34645416]
56. Cavazzana A, Röhrborn A, Garthus-Niegel S, Larsson M, Hummel T, Croy I. Sensory-specific impairment among older people. An investigation using both sensory thresholds and subjective measures across the five senses. PLoS One 2018 Aug 27;13(8):e0202969 [FREE Full text] [doi: 10.1371/journal.pone.0202969] [Medline: 30148857]
57. Dalton DS, Cruickshanks KJ, Klein BE, Klein R, Wiley TL, Nondahl DM. The impact of hearing loss on quality of life in older adults. Gerontologist 2003 Oct;43(5):661-668. [doi: 10.1093/geront/43.5.661] [Medline: 14570962]
58. Pope SK, Sowers M. Functional status and hearing impairments in women at midlife. J Gerontol B Psychol Sci Soc Sci 2000 May 01;55(3):S190-S194. [doi: 10.1093/geronb/55.3.s190] [Medline: 11833986]
59. China Health and Retirement Longitudinal Study. URL: http://charls.pku.edu.cn./ [accessed 2025-11-06]

Abbreviations

CES-D10: 10-item Center for Epidemiologic Studies Depression
CHARLS: China Health and Retirement Longitudinal Study
CVD: cardiovascular disease
LGMM: latent growth mixture modeling

Edited by H LaMonica; submitted 07.04.25; peer-reviewed by P Makmee, G Song; comments to author 15.06.25; revised version received 01.09.25; accepted 30.09.25; published 24.11.25.

Please cite as:
Liu Y, Liu W, Ma J, Chai Y, Kong G
Associations Between Hearing Loss and Depressive Symptom Trajectories in Middle-Aged and Older People in China: Retrospective Analysis
JMIR Aging 2025;8:e75545
URL: https://aging.jmir.org/2025/1/e75545
doi:10.2196/75545
PMID:

©Yuchen Liu, Wenwen Liu, Jun Ma, Yangfan Chai, Guilan Kong. Originally published in JMIR Aging (https://aging.jmir.org), 24.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Detecting Sleep/Wake Rhythm Disruption Related to Cognition in Older Adults With and Without Mild Cognitive Impairment Using the myRhythmWatch Platform: Feasibility and Correlation Study

Caleb D Jones[1], MS; Rachel Wasilko[2], MPH; Gehui Zhang[3], PhD; Katie L Stone[4,5], MA, PhD; Swathi Gujral[2], PhD; Juleen Rodakowski[6], MS, OTD, OTR/L; Stephen F Smagula[2], PhD

[1]Primary Care Accelerated Track, School of Medicine, University of Pittsburgh, Pittsburgh, PA, United States

[2]Department of Psychiatry, School of Medicine, University of Pittsburgh, 3811 O'Hara Street, Pittsburgh, PA, United States

[3]School of Sciences, Southwest Petroleum University, Chengdu, Sichuan, China

[4]California Pacific Medical Center Research Institute, San Francisco, CA, United States

[5]Department of Epidemiology and Biostatistics, University of California, San Francisco, CA, United States

[6]Department of Occupational Therapy, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, PA, United States

Corresponding Author:
Stephen F Smagula, PhD
Department of Psychiatry, School of Medicine, University of Pittsburgh, 3811 O'Hara Street, Pittsburgh, PA, United States

Abstract

Background: Consumer wearable devices could, in theory, provide sufficient accelerometer data for measuring the 24-hour sleep/wake risk factors for dementia that have been identified in prior research. To our knowledge, no prior study in older adults has demonstrated the feasibility and acceptability of accessing sufficient consumer wearable accelerometer data to compute 24-hour sleep/wake rhythm measures.

Objective: We aimed to establish the feasibility of characterizing 24-hour sleep/wake rhythm measures using accelerometer data gathered from the Apple Watch in older adults with and without mild cognitive impairment (MCI), and to examine correlations of these sleep/wake rhythm measures with neuropsychological test performance.

Methods: Of the 40 adults enrolled (mean [SD] age 67.2 [8.4] years; 72.5% female), 19 had MCI and 21 had no cognitive disorder (NCD). Participants were provided devices, oriented to the study software (myRhythmWatch or myRW), and asked to use the system for a week. The primary feasibility outcome was whether participants collected enough data to assess 24-hour sleep/wake rhythm measures (ie, ≥3 valid continuous days). We extracted standard nonparametric and extended-cosine based sleep/wake rhythm metrics. Neuropsychological tests gauged immediate and delayed memory (Hopkins Verbal Learning Test) as well as processing speed and set-shifting (Oral Trails Parts A and B).

Results: All participants meet the primary feasibility outcome of providing sufficient data (≥3 valid days) for sleep/wake rhythm measures. The mean (SD) recording length was somewhat shorter in the MCI group at 6.6 (1.2) days compared with the NCD group at 7.2 (0.6) days. Later activity onset times were associated with worse delayed memory performance (β=−.28). More fragmented rhythms were associated with worse processing speed (β=.40).

Conclusions: Using the Apple Watch-based myRW system to gather raw accelerometer data is feasible in older adults with and without MCI. Sleep/wake rhythms variables generated from this system correlated with cognitive function, suggesting future studies can use this approach to evaluate novel, scalable, risk factor characterization and targeted therapy approaches.

(JMIR Aging 2025;8:e67294) doi:10.2196/67294

KEYWORDS

sleep; sleep/wake; circadian; activity pattern; dementia; cognition; mobile sensing; actigraphy; accelerometer

Introduction

Twenty-four-hour sleep/wake characteristics, which are objectively measurable using accelerometer-containing devices, are related to both dementia biomarkers and dementia risk. Prior studies have shown that sleep/wake rhythm disruption, including fragmentation of 24-hour sleep/wake rhythms, temporally precedes the incidence of mild cognitive impairment (MCI) and dementia [1,2]. Even among adults with normal cognition, rhythm fragmentation correlates with greater brain amyloid deposition [3,4]. Over time, 24-hour rhythm fragmentation has been associated with increased risk and faster rates of cognitive decline in people with MCI and mild-to-moderate dementia [1,5] .

XSL•FO
RenderX

The above-mentioned studies linking sleep/wake rhythm disruption with cognitive impairment and neurodegenerative processes in aging raise the potential that accelerometer-based sleep/wake monitoring may have useful clinical applications. For example, wearable accelerometers could potentially be used to identify individuals who have established sleep/wake risk factors for dementia, assign targeted interventions, and track sleep/wake patterns throughout clinical trials. Compared with researcher-focused accelerometer devices, consumer wearables (which also contain triaxial accelerometers) could yield more widely scalable, and clinician and user-friendly, systems for developing and testing potential applications. Our prior pilot study demonstrated that it is possible to collect 24-hour accelerometer data from the Apple Watch and generate standard 24-hour sleep/wake rhythm measures [6]. However, this prior study was limited to a convenience sample of young adults.

Prior to studies evaluating the use of consumer wearable-based sleep/wake monitoring for dementia risk stratification and targeting prevention approaches, we sought to establish the feasibility of using a consumer wearable device to assess 24-hour sleep/wake patterns in older adults (including those with elevated dementia risk by virtue of having a diagnosis of MCI). Our first aim was therefore to evaluate the feasibility of using the Apple Watch and a software platform called myRhythmWatch (myRW) to obtain 24-hour accelerometer data assessing sleep/wake rhythms in older adults with and without MCI. Second, we sought to validate that the sleep/wake rhythm data collected from this system is relevant to cognition. To do so, we evaluated if sleep/wake rhythm variables extracted from this system correlated with cognitive function similar to prior published studies that used researcher-focused accelerometer devices.

Methods

Participants and Study Protocol

Participants were identified by referral from either local studies that adjudicated MCI diagnoses or a local research recruitment registry that is led by the University of Pittsburgh Clinical and Translational Science Institute (Pitt+Me). The inclusion criteria were (1) being 50 years of age or older; (2) passing the San Diego Brief Assessment of Capacity to Consent [7] with scores ≥14.5; (3) have a score of >27 on the Telephone Interview for Cognitive Status (TICS) [8]; (4) having a TICS score of either ≤34 (high-risk group) or ≥39 (no cognitive disorder or NCD); and (5) have a prior adjudicated diagnosis of MCI (high-risk group) or reporting no concerns regarding cognitive decline (ie, NCD). The exclusion criteria included (1) self-reporting active behavioral health treatments for insomnia or depression; (2) self-reporting of performing the prescribed exercises; and (3) self-reported use of sleep medications every night or nearly every night. After explaining the study purpose and procedures, we obtained verbal consent to screen interested individuals for eligibility. Of 82 potentially eligible participants who completed the verbal eligibility screening, 35 were ineligible, 6 refused, and 41 were enrolled. One participant withdrew after enrollment, resulting in an analytic sample of 40 individuals (19 high-risk (ie, having a low TICS score and prior MCI diagnosis) and 21 low-risk (ie, having a high TICS score and no prior MCI diagnosis), as shown in Figure S1 in Multimedia Appendix 1 (ie, a flow diagram illustrating how we arrived at our analytic sample).

After completing web-based written informed consent forms, participants completed baseline procedures and assessments via Health Insurance Portability and Accountability Act (HIPPA)-compliant video-conferencing. We sent participants the following study devices: an Apple Watch 8, iPhone SE 2nd generation with active data plans, and nonstock 40-Watt charger. After the participants received the devices, we conducted an additional study visit to instruct them on logging in and using the myRW application. We specifically instructed participants that about 20 - 30 minutes of charging a day is sufficient, and asked them to wear the watch whenever it was not charging for a week.

Ethical Considerations

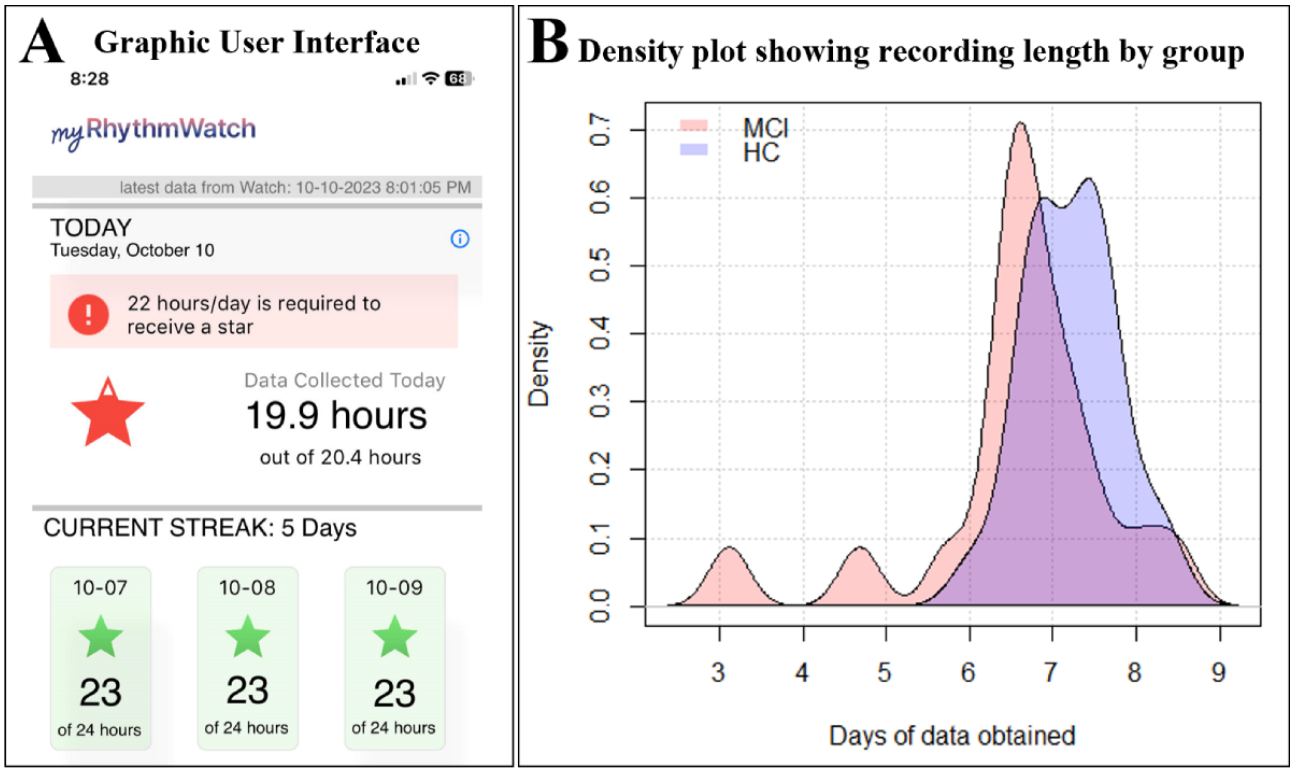
All study procedures were approved by the University of Pittsburgh Institutional Review Board (identifier: STUDY22080033). All the participants completed web-based written informed consent forms

Study Application

The myRW software system, developed by the corresponding author, extracts triaxial accelerometer data that are recorded at 50 hertz on the Apple Watch. After bandpass filtering, we aggregate across axes and into 30-second counts, and transmit data for storage and further computations on the Amazon Web Services platform. To encourage data collection, we depict the amount of data collected per day graphically by proportionally filling and color coding a star (Figure 1). We explicitly told participants that sufficient data were required to “get a star,” that the goal was to get 7 stars in a row, and we also displayed information on how many days in a row participants obtained stars (to encourage data collection “streaks”). Participants received reminders three times daily to check in to the application. Other features, including personal data graphs and sleep/wake metrics, were disabled for this study to prevent the feedback from potentially altering the users’ behavior/typical sleep/wake patterns.

XSL•FO
RenderX

Figure 1. Illustrations of the graphic user interface (left) and the amount of accelerometer data collected in the sample (right). The difference in the mean (SD) recording length between the no cognitive disorder (NCD; 7.2 [0.6]) and mild cognitive impairment (MCI; 6.6 [1.2]) groups is shown visually in the density plot (*P*=.04).



Sleep/Wake Measures

Consistent with prior publications using researcher-focused accelerometers [9-11], we required participants to have at least 3 valid continuous days of accelerometer data to consider the recording adequate for processing. Valid days were defined, consistent with National Sleep Research Resource data processing standards [12], as those with no more than 4 hours of nonwear time or any nonwear time during the main sleep period. Invalid days and remaining data missing within days were censored (not imputed). Although we applied this aforementioned 20 hours per day criterion, to maximize data collection, note that the app was programmed to reward participants with a green star only if they collected at least 22 hours of data per day.

Following previously published technical definitions, we calculated extended cosine-based [13] and nonparametric [14-17] sleep/wake measures using the R package 'RAR' and custom code. From the extended-cosine models, we extracted measures of 24-hour robustness (pseudo-F statistic, indicating how well the observed data fits the 24-hour curve); activity onset time (up-mesor, the time which the modeled activity level passes the middle modeled rhythm height prior to the peak); and activity offset time (down-mesor or the time which the modeled activity level passes the middle modeled rhythm height prior to the nadir). From the nonparametric method, we calculated the cross-daily stability (inter-daily stability, measuring the consistency of circadian sleep/wake activity rhythms across days); rhythm strength (relative amplitude, measuring the standardized peak-trough difference of 24-hour activity rhythms); and sleep/wake rhythm fragmentation (intradaily variability or IV, measuring the frequency and extent of transitions in activity levels). Although sleep/wake rhythm fragmentation is typically measured using differences in activity levels every hour, it can also be computed on a range of timescales [16,17]. For all nonparametric metrics, we used all-time series data and did not subsample to hourly activity levels, as described previously [14]. To explore the relevance of fragmentation timescale, we computed IV values using numerator timescales ranging from 5 to 60 minutes, but holding the denominator constant as done previously [16].

Neuropsychological Measures

Cognitive performance was measured using the Hopkins Verbal Learning Test (HVLT; PAR, Lutz, FL) and the Oral Trail Making Test (O-TMT).

The HVLT is a verbal list learning measure with three trials of 12 words from multiple representative semantic categories, and includes delayed recall and delayed recognition memory components [18,19]. For the delayed recall, participants were asked to repeat as many terms as they could recall from the word listed presented to them thrice, 30 minutes earlier. For delayed memory recognition, participants were asked to recognize those same terms appearing on a broader list containing semantically related terms, ie, "horse" must be differentiated from "dog," which was not part of the original term set. The HVLT assessed encoding, storage, and retrieval of noncontextual verbal information.

Processing speed was measured as the time taken on Part A of the O-TMT, which requires the subject to verbally count upwards from 1 to 25 as quickly as they are able [20]. Set-shifting, an aspect of cognitive function, was measured as

XSL•FO
RenderX

the time taken using Part B of the O-TMT, which requires the subject to alternate between listing numbers and letters in sequence (“1, A, 2, B,” etc). Note that longer times on both parts of the O-TMT indicates worse performance [21].

Statistical Analyses

To compare recording lengths between the MCI and NCD groups, we used an independent sample *t* test and density plots illustrating recording length distributions by groups. To examine relationships of sleep/wake rhythm measures with cognitive function, we first adjusted for key confounders by taking the residuals from linear regression models (one for each cognitive outcome variable) that had the confounders age, sex, education (college degree vs less than college degree), and accelerometer recording length as predictor variables. We used the residualized values as age, sex, education, and recording length-adjusted cognitive outcome variables in a series of linear regressions (one per sleep/wake predictor variable). Since there were many conceptually similar or highly inter-correlated intradaily variability metrics, for related analyses, we only report Benjamini-Hochberg corrected false discovery rates (which are here referred to as *q* values instead of *P* values) [22] to account for the 12 statistical tests relating intradaily variability metrics within each cognitive outcome.

Results

Sample Characteristics

The sample included older adults, with a mean (SD) age of 67.2 (8.4) years; the majority of the participants were females with college degrees (Table 1).

Table . Sample characteristics.

Characteristics	Value
Age, years, mean (SD)	67.2 (8.4)
Female sex, n (%)	29 (72.5)
College degree or greater, n (%)	29 (72.5)
Prior diagnosis of MCI[a], n (%)	19 (47.5)
Accelerometer recording length, days, mean (SD)	6.9 (0.9)
Immediate memory, mean (SD)	26.7 (5.2)
Delayed memory, mean (SD)	8.7 (2.8)
Processing speed, seconds, mean (SD)	10.3 (2.8)
Set shifting executive function, seconds, mean (SD)	40.1 (31.6)

[a]MCI: mild cognitive impairment

Accelerometer Recording Lengths in People With and Without a Diagnosis of MCI

In both the groups, all participants achieved the minimum data requirement for computing 24-hour sleep/wake rhythm assessments. There was, however, a statistically significant difference in the mean (SD) recording lengths of about a half a day when comparing the NCD (7.2 [0.6] days) and MCI groups (6.6 [1.2] days), with the Satterthwaite test assuming unequal group variances (*df* 25.033; *t* value −2.15; *P*=.04). This difference was due to 3 individuals in the MCI group who collected 3 - 6 days of data each (Figure 1).

Associations of Sleep/Wake Characteristics With Cognitive Performance

As shown in Table 2, there were small effect size associations of later activity onset time with lower delayed memory performance (β=−.28, 95% CI −0.55 to - 0.02; *t*=−2.17, *df*=38, *P*=.04) and more stable rhythms with better processing speed and performance (β=−0.27, 95% CI: −0.54 to 0.00; *t*=−2.00, *df*=38, *P*=.05). None of the other sleep/wake measures listed in Table 2 were associated with the cognitive outcomes. Regarding sleep/wake rhythm fragmentation, greater fragmentation levels in the 40 - 60-minute timeframe was significantly associated with worse processing speed and performance (Table 3; β point estimate range: 0.36, 0.40; *q*=.022). When repeating analyses in the subgroup with at least 6 days of data, results were not substantively altered (n=37; see Tables S1 and S2 in Multimedia Appendix 1).

To illustrate the accelerometer data collected in this study and visualize 24-hour sleep/wake fragmentation, Figure 2 shows data from example participants with lower and higher degrees of 24-hour sleep/wake rhythm fragmentation.

XSL•FO
RenderX

Table . Associations between 24-hour sleep/wake rhythm variables and cognitive performance.

Variable	Immediate memory[a]		Delayed memory[a]		Psychomotor speed/attention[b]		Set shifting executive function[b]	
	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value
24 h robustness	.06 (−.22 to .34)	.67	.17 (−.11 to 0.44)	.22	−.09 (−.38 to .20)	.53	−.18 (−.48 to .11)	.22
Cross-daily stability	.11 (−.17 to .39)	.44	.13 (−.14 to .41)	.34	−.27 (−.54 to .00)	.05	−.11 (−.42 to .19)	.45
Rhythm strength	.08 (−.21 to .36)	.59	.04 (−.23 to .32)	.75	−.04 (−.33 to .24)	.76	−.12 (−.42 to .18)	.43
Activity onset time	−.13 (−.41 to .15)	.36	−.28 (−.55 to -.02)	.04	.04 (−.24 to .33)	.76	−.11 (−.41 to .19)	.46
Activity offset time	−.03 (−.32 to .25)	.81	−.07 (−.35 to .21)	.60	−.12 (−.41 to .16)	.40	.17 (−.13 to .47)	.26

[a]Higher scores on the memory tests indicates better performance (as the outcome is number of items recalled)

[b]For psychomotor speed/attention and set-shifting, higher scores indicate worse performance (as the outcome is the duration of time to complete Oral Trails A and B, respectively)

Table . Associations between 24-hour sleep/wake rhythm fragmentation on various timescales with cognitive performance.

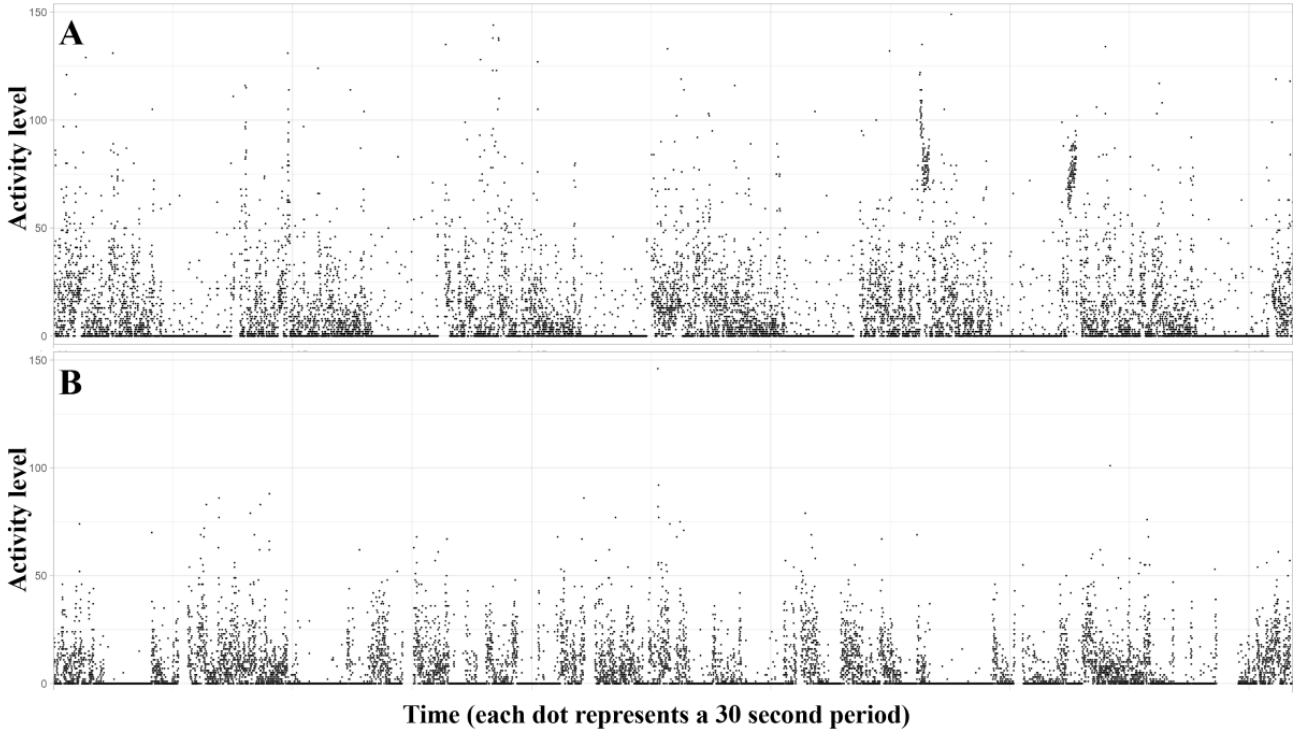
Time scale	Immediate memory[a]		Delayed memory[a]		Psychomotor speed/attention[b]		Set shifting executive function[b]	
	β (95% CI)	*q* value	β (95% CI)	*q* value	β (95% CI)	*q* value	β (95% CI)	*q* value
5 minutes	-.10 (−.38 to .18)	0.938	−.01 (−.29 to .27)	0.938	.14 (−.14 to .43)	0.314	.28 (−.01 to .58)	0.092
10 minutes	−.10 (−.38 to .18)	0.938	−.02 (−.3 to .26)	0.938	.14 (−.14 to .43)	0.314	.26 (−.03 to .55)	0.102
15 minutes	−.10 (−.39 to .18)	0.841	−.05 (−.33 to .23)	0.841	.16 (−.12 to .45)	0.298	.30 (.02 to .59)	0.092
20 minutes	−.12 (−.41 to .16)	0.628	−.11 (−.38 to .17)	0.628	.21 (−.07 to .49)	0.203	.34 (.06 to .62)	0.092
25 minutes	−.11 (−.40 to .17)	0.628	-.10 (-.38 to .18)	0.628	.20 (−.08 to .48)	0.203	.34 (.06 to .63)	0.092
30 minutes	−.11 (−.39 to .17)	0.628	−.11 (−.39 to .16)	0.628	.23 (−.05 to .51)	0.177	.32 (.03 to .61)	0.092
35 minutes	−.09 (−.37 to .19)	0.628	−.13 (−.41 to .14)	0.628	.28 (.01 to .55)	0.085	.28 (−.01 to .57)	0.093
40 minutes	−.08 (−.36 to .20)	0.628	−.15 (−.42 to .13)	0.628	.37 (.10 to .63)	0.022	.32 (.03 to .60)	0.092
45 minutes	-.09 (-.38 to .19)	0.628	−.16 (−.44 to .11)	0.628	.37 (.11 to .63)	0.022	.29 (.00 to .58)	0.092
50 minutes	−.11 (−.39 to .17)	0.628	−.18 (−.45 to .09)	0.628	.40 (.14 to .65)	0.022	.25 (−.05 to .54)	0.118
55 minutes	−.08 (−.36 to .21)	0.628	−.18 (−.45 to .1)	0.628	.36 (.09 to .62)	0.022	.17 (−.13 to .47)	0.254
60 minutes	−.12 (−.40 to .16)	0.628	−.23 (−.50 to .04)	0.628	.36 (.09 to .62)	0.022	.18 (−.12 to .48)	0.250

[a]Higher scores on the memory tests indicates better performance (as the outcome is number of items recalled).

[b]For psychomotor speed/attention and set-shifting, higher scores indicate worse performance (as the outcome is duration of time to complete Oral Trails A and B, respectively).

XSL•FO
RenderX

Figure 2. Accelerometer data from two participants. Top: The participant was in the low risk (no cognitive disease) group and had relatively lower sleep/wake rhythm fragmentation. Bottom: The participant is from the high-risk (mild cognitive impairment) group and had relatively higher sleep/wake rhythm fragmentation.



Discussion

These results demonstrate that it is feasible to collect raw accelerometer data and characterize 24-hour sleep/wake rhythms in older adults with and without MCI using the Apple Watch and myRW system. Each individual in our sample met the minimum data requirement to derive sleep/wake rhythms measures. Notably, however, several individuals in the MCI group generated shorter recordings. Thus, while supporting the feasibility of using the Apple Watch and myRW system to assess sleep/wake rhythms in people with MCI who are at risk for dementia, our findings also suggest that some individuals with MCI may require additional support (eg, additional human support or programmed reminders) and data imputation [23] to monitor sleep/wake rhythms over longer periods with this system.

With regard to our second aim, we found that sleep/wake rhythm variables extracted from the Apple Watch accelerometer data were correlated with cognitive function. We found small-to-medium effect size correlations between sleep/wake measures and cognitive performance. Specifically, we found that later activity onset times were related to worse delayed memory performances. In addition, less stable and more fragmented rhythms were correlated with worse processing speed. Detecting these signs of disrupted sleep/wake rhythms early on could help target prevention strategies, given prior research demonstrating that sleep/wake rhythm fragmentation [1], worse memory [24,25], and slower processing speeds [26] are all associated with the risk of developing dementia.

The study had several limitations. This was a cross-sectional observational study limited to accelerometer and neuropsychological measures, so there is no way to determine causality or the mechanisms underlying associations between the variables examined. The sample size was relatively small; therefore, there is a risk of false negative associations and effect size estimates (relating sleep/wake measures with cognition) that should be deemed less reliable than those from large epidemiologic studies. We failed to detect some associations that were expected based on prior literature examining dementia risk (eg, previous findings linking low rhythm strength with dementia risk [27]). Larger studies examining the relationships between these sleep/wake factors and specific domains of cognitive function, earlier in the disease processes, will be needed to verify our findings. We made efforts to minimize data loss as described above, but did not use data imputation, which could be applied in future studies. Additionally, given the small sample, future studies will be required to confirm results of using this system in samples that are more broadly representative, eg, samples including more ethnically diverse population subgroups. Finally, we note that determining the mechanism underlying these relationships between sleep/wake rhythm disruption, cognition, and dementia risk is outside the scope of this work. Recent literature has notably suggested that sleep/wake rhythm fragmentation relates to neurodegeneration of the locus coeruleus [28], which could hasten dementia pathology and cognitive decline [29,30].

In summary, we have demonstrated that it is feasible to use the Apple Watch and myRW system to gather accelerometer data and characterize sleep/wake risk factors for dementia in older adults including adults with MCI. A strength of this study is the use of the Apple Watch, which is already voluntarily being used by millions of people, as it may provide increased scalability for applications of sleep/wake risk factor monitoring into the general public and general practice settings. One implication

XSL•FO
RenderX

of this research is that it is feasible for future research to be conducted for evaluating if monitoring sleep/wake disruption using consumer wearable-based systems improves upon existing dementia risk factor detection and management approaches. Future studies will also be needed to examine if tailoring interventions using information on sleep/wake patterns derived from this system improves outcomes among older adults who are at risk for dementia.

Acknowledgments

CJ was supported by T32MH019986 in completing this work. The work presented herein was supported by a grant from the NIA to SFS (R21AG074094) and to JUR (R01AG056351).

Conflicts of Interest

SFS is owner and chief executive of a start-up, that was a University Licensed start-up, which worked on developing the myRW system with support from an National Institute on Aging STTR award (R41AG069596). KS reports grant funding from Eli Lilly and is supported by grant funding from NIH, PCORI, and DOD.

Multimedia Appendix 1
Supplementary Tables 1 and 2 and Figure 1.
[DOCX File, 36 KB - aging_v8i1e67294_app1.docx]

References

1. Li P, Gao L, Gaba A, et al. Circadian disturbances in Alzheimer's disease progression: a prospective observational cohort study of community-based older adults. Lancet Healthy Longev 2020 Dec;1(3):e96-e105. [doi: 10.1016/s2666-7568(20)30015-5] [Medline: 34179863]
2. Haghayegh S, Gao C, Sugg E, et al. Association of rest-activity rhythm and risk of developing dementia or mild cognitive impairment in the middle-aged and older population: prospective cohort study. JMIR Public Health Surveill 2024 May 7;10:e55211. [doi: 10.2196/55211] [Medline: 38713911]
3. Musiek ES, Bhimasani M, Zangrilli MA, Morris JC, Holtzman DM, Ju YES. Circadian rest-activity pattern changes in aging and preclinical Alzheimer disease. JAMA Neurol 2018 May 1;75(5):582-590. [doi: 10.1001/jamaneurol.2017.4719] [Medline: 29379963]
4. Nguyen Ho PT, Hoepel SJW, Rodriguez-Ayllon M, Luik AI, Vernooij MW, Neitzel J. Sleep, 24-hour activity rhythms, and subsequent amyloid-β pathology. JAMA Neurol 2024 Aug 1;81(8):824-834. [doi: 10.1001/jamaneurol.2024.1755] [Medline: 38913396]
5. Targa ADS, Benítez ID, Dakterzada F, et al. The circadian rest-activity pattern predicts cognitive decline among mild-moderate Alzheimer's disease patients. Alzheimers Res Ther 2021 Sep 25;13(1):161. [doi: 10.1186/s13195-021-00903-7] [Medline: 34563258]
6. Smagula SF, Stahl ST, Krafty RT, Buysse DJ. Initial proof of concept that a consumer wearable can be used for real-time rest-activity rhythm monitoring. Sleep 2022 Mar 14;45(3):zsab288. [doi: 10.1093/sleep/zsab288] [Medline: 34931683]
7. Jeste DV, Palmer BW, Appelbaum PS, et al. A new brief instrument for assessing decisional capacity for clinical research. Arch Gen Psychiatry 2007 Aug;64(8):966-974. [doi: 10.1001/archpsyc.64.8.966] [Medline: 17679641]
8. Ferrucci L, Del Lungo I, Guralnik JM, et al. Is the telephone interview for cognitive status a valid alternative in persons who cannot be evaluated by the mini mental state examination? Aging Clin Exp Res 1998 Aug;10(4):332-338. [doi: 10.1007/BF03339796]
9. Smagula SF, Gasperetti CE, Buysse DJ, et al. Efficacy of the transdiagnostic intervention for sleep and circadian dysfunction for depression symptoms and sleep-wake disruption in older and younger adults: secondary age-stratified analysis of a randomized controlled trial. Am J Geriatr Psychiatry 2024 Apr;32(4):478-488. [doi: 10.1016/j.jagp.2023.11.003] [Medline: 38040569]
10. Smagula SF, Zhang G, Krafty RT, et al. Sleep-wake behaviors associated with cognitive performance in middle-aged participants of the Hispanic Community Health Study/Study of Latinos. Sleep Health 2024 Aug;10(4):500-507. [doi: 10.1016/j.sleh.2024.02.002] [Medline: 38693044]
11. Lau SCL, Zhang G, Rueschman M, et al. Sleep-wake behavioral characteristics associated with depression symptoms: findings from the multi-ethnic study of atherosclerosis. Sleep 2024 Apr 12;47(4):zsae045. [doi: 10.1093/sleep/zsae045] [Medline: 38394355]
12. Dean DA 2nd, Goldberger AL, Mueller R, et al. Scaling up scientific discovery in sleep medicine: the national sleep research resource. Sleep 2016 May 1;39(5):1151-1164. [doi: 10.5665/sleep.5774] [Medline: 27070134]
13. Marler MR, Gehrman P, Martin JL, Ancoli-Israel S. The sigmoidally transformed cosine curve: a mathematical model for circadian rhythms with symmetric non-sinusoidal shapes. Stat Med 2006 Nov 30;25(22):3893-3904. [doi: 10.1002/sim.2466] [Medline: 16381069]

XSL•FO
RenderX

14. Suibkitwanchai K, Sykulski AM, Perez Algorta G, Waller D, Walshe C. Nonparametric time series summary statistics for high-frequency accelerometry data from individuals with advanced dementia. PLoS ONE 2020;15(9):e0239368. [doi: 10.1371/journal.pone.0239368] [Medline: 32976498]
15. Gonçalves BSB, Adamowicz T, Louzada FM, Moreno CR, Araujo JF. A fresh look at the use of nonparametric analysis in actimetry. Sleep Med Rev 2015 Apr;20(84-91):84-91. [doi: 10.1016/j.smrv.2014.06.002] [Medline: 25065908]
16. Fossion R, Rivera AL, Toledo-Roy JC, Ellis J, Angelova M. Multiscale adaptive analysis of circadian rhythms and intradaily variability: application to actigraphy time series in acute insomnia subjects. PLoS ONE 2017;12(7):e0181762. [doi: 10.1371/journal.pone.0181762] [Medline: 28753669]
17. Gonçalves BSB, Cavalcanti PRA, Tavares GR, Campos TF, Araujo JF. Nonparametric methods in actigraphy: an update. Sleep Sci 2014 Sep;7(3):158-164. [doi: 10.1016/j.slsci.2014.09.013] [Medline: 26483921]
18. Vanderploeg RD, Schinka JA, Jones T, Small BJ, Graves AB, Mortimer JA. Elderly norms for the Hopkins verbal learning test-revised. Clin Neuropsychol 2000 Aug;14(3):318-324. [doi: 10.1076/1385-4046(200008)14:3;1-P;FT318] [Medline: 11262707]
19. Tri Budi Rahardjo XX. The Hopkins verbal learning test and detection of MCI and mild dementia: a literature review. J Alzheimers Dis Parkinsonism 2014 Oct;04(6):166. [doi: 10.4172/2161-0460.1000166]
20. Mrazik M, Millis S, Drane DL. The oral trail making test: effects of age and concurrent validity. Arch Clin Neuropsychol 2010 May;25(3):236-243. [doi: 10.1093/arclin/acq006] [Medline: 20197294]
21. Ashendorf L, Jefferson AL, O'Connor MK, Chaisson C, Green RC, Stern RA. Trail making test errors in normal aging, mild cognitive impairment, and dementia. Arch Clin Neuropsychol 2008 Mar;23(2):129-137. [doi: 10.1016/j.acn.2007.11.005] [Medline: 18178372]
22. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B 1995 Jan 1;57(1):289-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
23. Weed L, Lok R, Chawra D, Zeitzer J. The impact of missing data and imputation methods on the analysis of 24-hour activity patterns. Clocks Sleep 2022 Sep 27;4(4):497-507. [doi: 10.3390/clockssleep4040039] [Medline: 36278532]
24. Hayat SA, Luben R, Khaw KT, Brayne C. The relationship between cognitive performance using tests assessing a range of cognitive domains and future dementia diagnosis in a British cohort: a ten-year prospective study. J Alzheimers Dis 2021;81(1):123-135. [doi: 10.3233/JAD-210030] [Medline: 33867360]
25. Rajan KB, Wilson RS, Weuve J, Barnes LL, Evans DA. Cognitive impairment 18 years before clinical diagnosis of Alzheimer disease dementia. Neurology (ECronicon) 2015 Sep 8;85(10):898-904. [doi: 10.1212/WNL.0000000000001774] [Medline: 26109713]
26. Amieva H, Meillon C, Proust-Lima C, Dartigues JF. Is low psychomotor speed a marker of brain vulnerability in late life? Digit symbol substitution test in the prediction of Alzheimer, Parkinson, stroke, disability, and depression. Dement Geriatr Cogn Disord 2019;47(4-6):297-305. [doi: 10.1159/000500597] [Medline: 31466055]
27. Winer JR, Lok R, Weed L, et al. Impaired 24-h activity patterns are associated with an increased risk of Alzheimer's disease, Parkinson's disease, and cognitive decline. Alzheimers Res Ther 2024 Feb 14;16(1):35. [doi: 10.1186/s13195-024-01411-0] [Medline: 38355598]
28. Van Egroo M, van Someren EJW, Grinberg LT, Bennett DA, Jacobs HIL. Associations of 24-hour rest-activity rhythm fragmentation, cognitive decline, and postmortem locus coeruleus hypopigmentation in Alzheimer's disease. Ann Neurol 2024 Apr;95(4):653-664. [doi: 10.1002/ana.26880] [Medline: 38407546]
29. Chen Y, Chen T, Hou R. Locus coeruleus in the pathogenesis of Alzheimer's disease: a systematic review. Alzheimers Dement (N Y) 2022;8(1):e12257. [doi: 10.1002/trc2.12257] [Medline: 35282658]
30. Bueichekú E, Diez I, Kim CM, et al. Spatiotemporal patterns of locus coeruleus integrity predict cortical tau and cognition. Nat Aging 2024 May;4(5):625-637. [doi: 10.1038/s43587-024-00626-y] [Medline: 38664576]

Abbreviations

HVLT: Hopkins Verbal Learning Test
MCI: mild cognitive impairment
NCD: no cognitive disorder
O-TMT: Oral Trail Making Test

XSL•FO
RenderX

Edited by AF Abdulai; submitted 21.10.24; peer-reviewed by R Lok, S Miyata; revised version received 04.03.25; accepted 07.03.25; published 07.04.25.

Please cite as:
Jones CD, Wasilko R, Zhang G, Stone KL, Gujral S, Rodakowski J, Smagula SF
Detecting Sleep/Wake Rhythm Disruption Related to Cognition in Older Adults With and Without Mild Cognitive Impairment Using the myRhythmWatch Platform: Feasibility and Correlation Study
JMIR Aging 2025;8:e67294
URL: https://aging.jmir.org/2025/1/e67294
doi:10.2196/67294

© Caleb D Jones, Rachel Wasilko, Gehui Zhang, Katie L Stone, Swathi Gujral, Juleen Rodakowski, Stephen F Smagula. Originally published in JMIR Aging (https://aging.jmir.org), 7.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Identifying Optimal Wearable Devices for Monitoring Mobility in Hospitalized Older Adults: Feasibility, Acceptability, and Validity Study

Paulo Nascimento[1], PhD; Renata Kirkwood[1], PhD; Lauren E Griffith[2], PhD; Mylinh Duong[3,4], PhD; Cody Cooper[1], BEng; Yujiao Hao[5], PhD; Rong Zheng[5], PhD; Samir Raza[4], MD; Marla Beauchamp[1], PhD

[1]School of Rehabilitation Sciences, McMaster University, 1280 Main Street West, Hamilton, ON, Canada

[2]Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada

[3]Population Health Research Institute, McMaster University, Hamilton, ON, Canada

[4]Department of Medicine, McMaster University, Hamilton, ON, Canada

[5]Department of Computing and Software, McMaster University, Hamilton, ON, Canada

Corresponding Author:
Marla Beauchamp, PhD
School of Rehabilitation Sciences, McMaster University, 1280 Main Street West, Hamilton, ON, Canada

Abstract

Background: Hospitalized, frail older adults have an increased risk of developing hospital-acquired disability associated with hospital practices of restricted physical activity and immobilization. The use of activity tracking wearable devices may allow identification and prevention of mobility decline, reducing hospital-acquired disability.

Objective: This study aimed to identify the optimal wearable device and wear location for monitoring mobility in older hospitalized patients. Specific objectives included (1) comparison of the feasibility and acceptability of ActiGraph wGT3X-BT (ActiGraph LLC), MOX1 (Maastricht Instruments), MetaMotionC (mBientLab), and Fitbit Versa (Google) for continuous mobility monitoring and (2) determination of the concurrent validity of the selected device for detecting body posture and step count.

Methods: Participants were recruited for this observational study in the acute medical care unit of an academic hospital in Hamilton, Ontario, Canada. Eligible patients were aged 60 years and older, able to undertake the mobility protocol, and had an anticipated length of stay greater than 4 days. The study was divided into 2 experiments. Experiment 1 evaluated the feasibility of 4 wearable devices and validated the derived data for body posture and step count. Experiment 2 involved a mobility assessment session and a 24-hour monitoring and feasibility period with the selected device from experiment 1.

Results: The ActiGraph wGT3X-BT emerged as the most feasible device, demonstrating superior usability, data acquisition, and management. The thigh-worn ActiGraph accurately detected sedentary behavior, while the ankle-worn device provided detailed information on step counts and body postures. Bland-Altman plots and intraclass correlation coefficients indicated that the ankle-worn ActiGraph showed excellent reliability for step counting, with minimal bias and narrow limits of agreement. Patients expressed a high willingness to wear a continuous mobility tracking device at the hospital and at home.

Conclusions: Thigh- and ankle-worn ActiGraph are optimal for assessing and monitoring mobility in older hospitalized patients. Challenges such as discomfort and device removal observed during the 24-hour monitoring period highlight areas for future studies. Overall, our findings support the integration of wearable technology in hospital settings to enhance mobility monitoring and early intervention strategies. Further research is warranted to evaluate the long-term use of wearable data for predicting health outcomes post hospitalization and informing clinical decision-making to promote early mobility.

(JMIR Aging 2025;8:e64372) doi:10.2196/64372

KEYWORDS

older adults; gerontology; geriatric; aging; feasibility; acceptability; mobility; wearable; inpatient; hospital-acquired disability; physical performance; mHealth; mobile health; hospital; physical activity; exercise; Fitbit; posture; walk

Introduction

Research indicates that a significant proportion of the adverse functional and health outcomes experienced by older adults during hospitalization may not be directly linked to their underlying health conditions or the reasons for hospitalization [1-3]. Instead, they may arise from certain hospital practices, such as restricted physical activity and immobilization, which may be harmful for older patients, especially for those who are frail [1]. Studies have consistently shown that older patients

XSL•FO
RenderX

spend a minimal amount of time standing or walking during their hospital stay, typically just 3% of the time [4-7]. This limited mobility can have significant consequences for function, with each day spent in bed associated with a 1% - 5% loss in muscle strength [3]. On the other hand, there is evidence showing that older adults who regain their prehospitalization level of function after discharge have lower mortality rates and maintain their functional levels 1 year post discharge. Therefore, early detection and prevention of mobility decline during hospitalization is critical to improving patient outcomes and reducing health care utilization [8].

Wearable technology provides a direct means of assessing and monitoring mobility, by gathering continuous information on patients' physical activities and mobility patterns. This capability would allow practitioners to create and monitor tailored mobility care for each patient, which may improve functional outcomes. Several wearable monitors have been validated in healthy individuals including among older adults living in the community [9,10]. However, experiment protocols, including population type, settings, sensor type, activity, and wear location (eg, wrist vs thigh) can influence the validity of these devices [11-14]. Despite the increased use of wearables in clinical settings, the feasibility, validity, and reliability of these devices have not been fully established in older hospitalized patients [15,16]. Collecting data accurately with wearables in hospitalized patients presents significant challenges, as they tend to be more sedentary and walk slower [17]. In addition, hospitalized patients often have medical devices attached to them, such as intravenous lines or heart monitors, which can impede mobility and make it challenging to select the optimal wear location for data collection.

The overall objective of this study was to identify the optimal wearable device and wear location for assessing and monitoring mobility in older hospitalized patients. Specifically, we aimed to (1) compare the feasibility and acceptability of different wearable devices to assess mobility for long-term continuous monitoring during inpatient hospital stays among older hospitalized patients, and (2) for the selected device from aim 1, determine its concurrent validity for detecting body posture and step count. To fulfill the study aims we performed 2 experiments with 2 independent samples. In the first experiment, we used a standardized protocol to test the feasibility of data collection with the ActiGraph wGT3X-BT, MOX 1, MetaMotionC (MMC), and Fitbit Versa over a short period of time, while validating the data collected for detecting body posture and step count using a standardized protocol. The second experiment was divided in 2 parts where patients wore the selected wearable from experiment 1 in 3 different body locations and performed a standardized activity protocol, followed by a 24-hour free-living protocol. This investigation was crucial for understanding the practicality of wearable devices for mobility monitoring in older adults in a hospital setting, as well as identifying potential barriers or limitations that may impact their utility.

Methods

Participants

Patients were recruited from the acute medical care unit of the Juravinski Hospital, Hamilton, Ontario, Canada. Inpatients were initially identified by the admitting physician, who sought approval before study personnel approached them. Patients were only included if, based on the judgment of the admitting physician, they were deemed capable of providing informed consent. A study coordinator, not directly involved in patient care, then contacted the eligible patients to obtain informed consent. During data collection, a hospital physical therapist, who accompanied all participants and assisted with data collection, ensured safety and facilitated communication when necessary.

Eligible patients were aged 60 years and older, able to undertake the mobility protocol with or without assistance, had an anticipated length of stay of more than 4 days, and able to provide written informed consent. Experiment 1 was conducted from November to December of 2019 and experiment 2 from September to December 2020.

Ethical Considerations

This study was approved by the Hamilton Integrated Research Ethics Board (HiREB #7145) and all participants provided written informed consent before participation. At the time of consent, we recorded information regarding participants' demographics, preadmission functional performance, and health status. All participants had the right to withdraw from the study at any time without any adverse consequences. All data were anonymized. Participants did not receive compensation for their participation.

Measurement Instruments

For the feasibility and device performance of experiment 1, 4 devices were compared, that is, the MetaMotionC (MMC), Fitbit Versa 1, MOX1, and the medical-grade monitor ActiGraph wGT3X-BT. The MMC by mBientLab was chosen for its open platform enabling on-board programming, in addition to its lower cost. The Fitbit Versa 1 (Google) was widely recognized in the community at the study's time, despite lacking direct access to raw sensor data extraction. For both devices, a Python algorithm was developed to extract the raw data. The MOX1 device (Maastricht Instruments , Netherlands) presents a triaxial accelerometer sensor and is equipped with proprietary software. During the study period, the proprietary software of the MOX1 did not offer direct access to the raw data and step count data. To address this limitation, the company provided us with a MATLAB function for extracting the raw data. Finally, the medical-grade ActiGraph wGT3X-BT (ActiGraph), often viewed as the gold standard in accelerometry movement analysis for research, features a built-in triaxial accelerometer that captures high-resolution raw acceleration data. The ActiLife software (ActiGraph LLC, version 6.11.4) was employed to initialize, process, and download data, extracting step counts and body posture measures (time spent lying down, sitting, and standing). In experiment 2, we used the ActiGraph wGT3X-BT as it was the selected device from experiment 1.

XSL•FO
RenderX

Protocols

Experiment 1

Patients from the first experiment were engaged in the activity data procedures to test the feasibility of 4 different wearable devices (ActiGraph wGT3X-BT, MOX1, MMC, and Fitbit Versa), as well as the concurrent validity of the ActiGraph in detecting body posture and step counts. Each patient wore all 4 devices on the waist, 3 devices on the thigh, and 3 on the ankle simultaneously and interchangeably, that is, up to 10 devices per patient. An elastic band equipped with Velcro on both ends was used to attach the devices securely to the body. The waistband featured 4 pockets, while the thigh and ankle bands had 3 pockets each, providing the flexibility to interchange the devices as needed during the study. The waistband was positioned at the level of the anterior superior iliac spine, the thigh band above the kneecap, and the ankle band above the malleolus. After positioning the elastic bands, the wearable devices were randomly assigned to each pocket. Once wearable devices were placed, patients performed the mobility protocol that included body posture tasks (standing, sitting, and lying) and the timed up and go (TUG) [18] mobility test.

During the body posture tasks, patients were asked to lie down, sit on the edge of the bed, and then stand with or without support for 5 minutes. The time spent in each body posture was observed and recorded by a physiotherapist. Following, the patients performed the TUG which is a reliable and valid test to assess mobility and balance in older adults [18]. It measures, in seconds, the time taken by an individual to stand up from a standard armchair, walk 3 meters, turn, walk back to the chair, and sit down. The participants wore their customary walking aid (none, cane, or walker), and no physical assistance was given. The recording initiated as the patient raised from the chair and stopped as the patient sat on the chair again. The time and the number of steps taken to complete the test were also counted and recorded. Following the activity procedures, patients were invited to complete the acceptability questionnaire (Textbox 1). The protocol took approximately 45 - 60 minutes to complete.

Textbox 1. Feasibility questionnaires for experiments 1 and 2.

Experiment 1

1. Have you ever used a device to measure physical activity in the past? (Responses: yes, no)
2. Would you be willing to wear the device for a longer period, 5 to 7 days, as part of a research study? (Responses: very likely, somewhat likely, not likely)
3. What part of the body would you prefer to wear the device? (Responses: waist, thigh, ankle)
4. Which of these devices would you likely use? (Responses: ActiGraph, MOX1, Versa, MetaMotionC)
5. How easy would it be for you to remember to use the device every day? (Responses: very easy, easy, very difficult, difficult)
6. Do you think this device would interfere with your daily routine? (Responses: no effect, minor effect, major effect)
7. Would you feel more motivated to move when wearing the device? (Responses: yes, no, no answer)

Experiment 2

1. Have you ever used a device to measure physical activity in the past? (Responses: yes, no)
   a. What devices did you use?
   b. What activity (or activities) did you track with the devices?
   c. How long ago was it that you used the devices?
   d. If you stopped using the devices, why did you stop?
2. Would you be willing to wear a device to measure physical activity for 5 to 7 days while in the hospital? (Responses: likely, uncertain, unlikely, very likely, very unlikely)
3. Would you be willing to wear a device daily to measure physical activity once you return home from the hospital, for a period of up to say 3 months? (Responses: yes, no)
4. What part of the body would you most prefer to wear the device? (Responses: wrist, waist, thigh, ankle, other)
5. Would you feel more motivated to move when wearing a device to measure physical activity? (Responses: yes, no)
6. Prior to being in the hospital, how many days per week on average did you engage in 30 minutes or more of physical activity, which was enough to raise your breathing rate? (Responses: 1, 2, 3, 4, 5, 6, 7, none)

Experiment 2

We conducted 1 mobility assessment session where patients wore the chosen device from experiment 1 (ActiGraph) in 3 different body locations (wrist, thigh, and ankle) and performed a standardized mobility protocol under the supervision of a trained physiotherapist. Patients wore the ActiGraph on the wrist using a wristband. At the thigh, the ActiGraph was attached to the anterior aspect of either the left or right thigh just above the kneecap, and at the ankle, the device was attached just above the malleoli. The Hypafix Stretch Non-Woven Adhesive (BSN Medical) was used to affix the devices. The

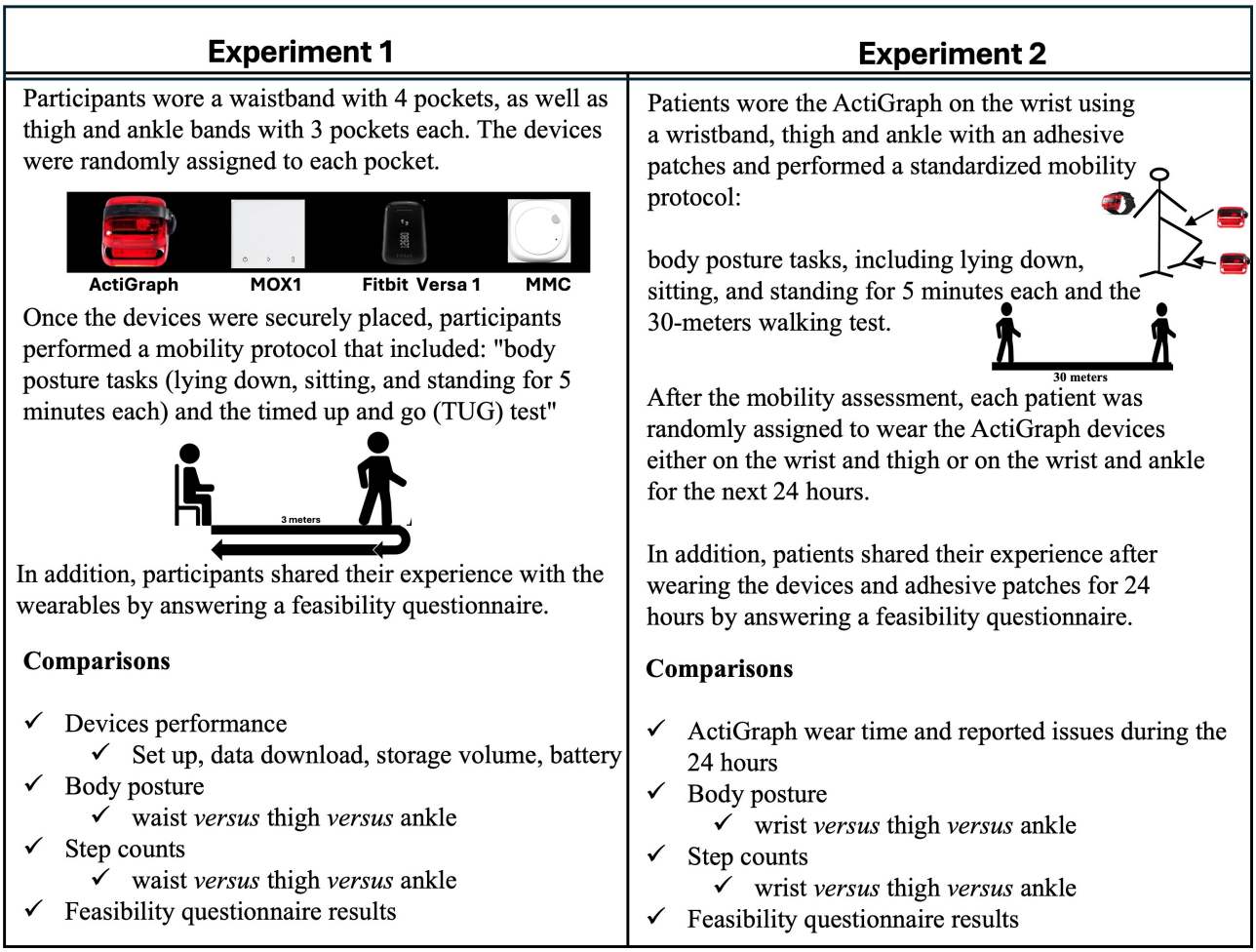
XSL•FO
RenderX

mobility protocol included the same body posture protocol as the first experiment, followed by a step count task recorded during a 30-meter walking test (30MWT) [19,20]. Patients were asked to stand up from a chair, walk 30 meter, turn around, and walk back at a comfortable pace using their usual walking aid if necessary. The time and the number of steps taken to complete the test was counted and recorded. The mobility protocol in experiment 2 lasted approximately 30 - 45 minutes.

After the mobility assessment, each patient was randomly assigned to wear the ActiGraph devices on either the wrist and thigh or wrist and ankle for the next 24 hours. We assessed the acceptability and feasibility of the wearable devices by location, as well as their acceptability of

wearing the adhesive patches over a 24-hour duration in the hospital. We also documented any interruptions or issues encountered during the 24-hour monitoring period. If patients removed or stopped using the device, the research assistant recorded feedback from both the patient and the attending nurses regarding the reasons for discontinuation. Figure 1 illustrated the variations within each experiment.

Figure 1. Summary of procedures for experiment 1 and 2.



Data Reduction

Experiment 1

Data were collected at 100 Hz for MOX1, and at 50 Hz for the other devices. In the second experiment, the ActiGraph data were collected at 30 Hz. Custom Python algorithms were created for the MMC and Fitbit devices to enable continuous saving and downloading of raw accelerometer data, necessitating the use of a companion device, a tablet. The MOX1 device was initialized using its proprietary software and data were downloaded using the provided MATLAB function. Raw data from these devices were not subjected to further processing; our focus was solely on recording downloading time, as well as the quantity and quality of the saved data.

In both experiments, the ActiLife software (version 6.11.4) was used to initialize, download, and process the ActiGraph data. Data were aggregated into 60 seconds time-stamped epochs. The following measures were obtained from the ActiLife: wear time, counts per minutes, activity intensity, step counts, and body posture (time spent in lying down, sitting, and standing). The following measures were obtained from the ActiLife algorithms: wear time, counts per minute, activity intensity, step counts, and body posture (time spent lying down, sitting, and standing). For step counts, since the ActiLife algorithms measures strides when attached to the thigh, we doubled the stride numbers to determine the actual number of steps taken. This adjustment is necessary because each stride recorded by the device corresponds to two steps—one for each leg—so the raw stride count is multiplied by 2 to obtain an accurate step count.

XSL•FO
RenderX

In addition, we applied the ActiGraph manufacturer's step algorithm, the low-frequency extension filter, to the step count data. This filter is designed to detect lower-amplitude movements, enhancing the accuracy of step detection [21]. Body posture classification in both experiments was obtained using the thigh-worn algorithm from the ActiGraph that relies on movement and the thigh angle to accurately classify lying and sitting versus standing positions [22]. The first and last 45 seconds of data from each activity were discarded to avoid potential participant error in recording time and to exclude the transition times.

Experiment 2

For the 24-hour protocol, the accelerometer data from the ActiGraph was screened for wear time using the method described by Choi et al [23]. Based on the activity counts determined by the ActiLife algorithms, and using an epoch length of 60 seconds, nonwear time was defined as 90 consecutive minutes of zero counts, with an allowance of 2 minutes of nonzero counts, provided there were 30-minute consecutive zero counts before and after that allowance. Based on the wear time, we determined the average amount of time the patients wore the devices at the wrist, thigh, and ankle during the 24-hour protocol.

Statistical Analysis

Descriptive data were analyzed using measures of central tendency and dispersion. Absolute percentage errors were calculated between the observed body posture time and step count against the values obtained using the ActiGraph algorithm (absolute percentage errors=[(observed data–ActiGraph data)÷observed data]×100). Intraclass correlation coefficients ($ICC_{2,1}$) [24] were used to examine criterion validity between step count taken from the ActiGraph compared with those observed during the TUG and 30MWT. By convention, an ICC ≥0.75 was considered excellent, 0.60 - 0.74 good, 0.40 - 0.59 fair, and <0.40 poor [25].

Bland-Altman plots [26,27] were constructed to show the variability of the ActiGraph in recording step count compared with the observed data. With this technique, the mean error score and the 95% prediction intervals can be examined in graphical form. Comparisons that are in closer agreement will have a mean bias close to zero and tighter 95% prediction intervals. Statistical analyses were conducted with SPSS version 26 (IBM Corp) with the α level set at .05.

Results

Overview

A total of 25 older adults (n=17, 65% women) with a mean age of 79.6 (SD 8.1) years participated in experiment 1, and 30 participants (n=24, 80% women), with a mean age of 81.4 (SD 8.8) years, participated in experiment 2 (Table 1). The patients were hospitalized due to the following conditions: cardiac conditions (n=6) including hypertension, congestive heart failure, atrial fibrillation, and coronary artery disease; urinary tract infection (n=5); deconditioning (n=5); falls (n=3); acute kidney injury (n=3); exacerbation of COPD (n=2); shortness of breath (n=2); sepsis (n=2); pneumonia (n=1); syncope (n=1); leg pain (n=1); vertigo (n=1); gastroenteritis (n=1); and back pain (n=1). Table 1 describes the patients' demographics and clinical characteristics. In general, patients were more independently mobile before admission.

XSL•FO
RenderX

Table . Demographics, comorbidities, and mobility characteristics of the patients.

Variables			Experiment 1	Experiment 2
Demographics, n			25	30
	Age (y), mean (SD)		79.6 (8.1)	81.4 (8.8)
	Sex (female), n (%)		17 (65)	24 (80)
	BMI (kg/m^2), mean (SD)		25.6 (6.2)	25.2 (7.7)
Gait aids, n (%)			25 (100)	29 (100)
	None		12 (48)	7 (24)
	Walker rollator		12 (48)	16 (55)
	Cane		1 (4)	2 (6)
	Cane and walker		0	4 (14)
Most common comorbidities, n			30	—[c]
	Cardiac condition (hypertension, congestive heart failure, atrial fibrillation, and coronary artery disease)		25	—
	Diabetes		9	—
	Osteoarthritis		8	—
	Cancer		7	—
	COPD[b] and Asthma		7	—
	Osteoporosis		5	—
	Kidney disease		5	—
	Early diagnosis of Alzheimer		2	—
	Parkinson		2	—
Mobility before admission, n (%)			—	30 (100)
	Independent		—	8 (27)
	Independent with cane		—	3 (10)
	Independent with walker		—	15 (50)
	1 Person assist		—	1 (3)
	1 Person assist with walker		—	2 (7)
	Independent with cane inside and walker outside		—	1 (3)
Mobility after admission			—	N=30
	Independent		—	1 (3)
	Independent with cane		—	1 (3)
	Independent with walker		—	11 (37)
	1 Person assist		—	3 (10)
	1 Person assist with walker		—	14 (47)
Mobility tests, n			23	—
	TUG[a] (s)			
		Mean (SD)	28.5 (13.7)	—
		Median (IQR)	23.3 (7.14-61.0)	—
	Average speed (m/s)			
		Mean (SD)	0.27 (0.16)	—
		Median (IQR)	0.22 (0.09-0.84)	—
	30MWT[d] (m-s) test 1			

XSL•FO
RenderX

Variables		Experiment 1	Experiment 2
	Mean (SD) (s)	—	50 (31)
	Median (IQR) (s)	—	35 (30-120)
	Percentile 25 (s)	—	32
	Percentile 75 (s)	—	67
30MWT (m-s) retest			
	Mean (SD) (s)	—	58 (17)
	Median (IQR) (s)	—	20 (30-50)
	Percentile 25 (s)	—	50
	Percentile 75 (s)	—	70

[a]TUG: timed up and go.

[b]COPD: chronic obstructive pulmonary disease.

[c]Not applicable.

[d]30MWT: 30-meter walking test.

Experiment 1

The acceptability questionnaire showed that out of 25 patients, 24 patients (96%) reported no previous experience with wearables. However, 20 patients (80%) expressed their willingness to wear a wearable for a period of 5-7 days. In terms of the preferred body location, the ankle was selected by 14 patients (58%), the waist was selected by 7 patients (29%), and the thigh by 4 patients (17%). Furthermore, 1 patient did not answer this question. Among the wearable devices, the Fitbit Versa was selected by 12 patients (48%), the ActiGraph was chosen by 6 patients (24%), and the MOX1 and MMC were selected by 3 (12%) and 2 (8%) patients, respectively.

We observed that both the ActiGraph and MOX1 devices were easy to set up and enabled faster data downloads compared with Fitbit and MMC devices. Of the 4 devices, only the ActiGraph retrieved 100% of the collected data. The MATLAB function from the MOX1 retrieved 72% of the files, with 1 file containing missing data, and 7 files having a lower frequency than the specified 100 Hz setup. In addition, the MATLAB function provided by the manufacturer did not include time stamps. The algorithms on the Fitbit Versa were able to retrieve 79% of the data, with the primary issue being missing data. Similarly, the MMC algorithms retrieved 81% of the data, although some data were missing due to a time stamp error that was not identified until later in the data collection process. In addition to its high performance, the ActiGraph device also had the longest battery life and storage volume compared to all other devices. The proprietary software offered by ActiGraph allowed us to process, quickly view and extract the data using a comprehensive selection of independently developed and validated algorithms. Given the ease of data collection using the ActiGraph in the above experiment, it was selected as the most user-friendly device for further evaluation.

For the body posture assessments, the waist- and thigh-worn ActiGraph identified the lying down position correctly 73.6% and 78.2% of the time, respectively. For the standing posture, both the thigh- and ankle-worn ActiGraph achieved high identification rates of 83.8% and 82.3%, respectively. However, all devices exhibited poor performance in identifying the sitting position, ranging from 25.4% to 49.6% (Table 2).

Table . Percentage of times the ActiGraph correctly detected body posture compared with the physiotherapist recordings during experiment 1 (n=25).

Patient body posture	Experiment 1 devices' attachment								
	Waist, %			Thigh, %			Ankle, %		
	Lie[a]	Sit	Stand	Lie	Sit	Stand	Lie	Sit	Stand
Lie	74[b]	0	0	78[b]	54	0	54[b]	0	0
Sit	11	41[b]	38	0	25[b]	15	0	50[b]	15
Stand	0	57	61[b]	1	1	84[b]	0.1	48	84[b]
Off[c]	16	1	1	21	20	1	45	3	3

[a]Lie: laying down.

[b]Percentage of time the ActiGraph device correctly identified the body postures.

[c]Off: percentage of time that the device detected it was off.

XSL•FO
RenderX

Table 3 displays the results of the ICC analysis, comparing the step counts during the TUG tests recorded by the ActiGraph devices worn on the waist, thigh, and ankle compared with direct observation by the physiotherapist. Data from 2 participants were excluded due to errors in registering the start and end of the tests. The ankle-worn device demonstrated the highest agreement with the physiotherapist ($ICC_{2,1}$=0.94, 95% CI 0.85-0.97), the lowest bias (average of the mean difference=0.9 steps), and a lower percentage of error counting steps (12.8%). The waist-worn device also shows excellent agreement with direct observation ($ICC_{2,1}$=0.85, 95% CI 0.65-0.94) but higher bias (1.4 steps) and a higher percentage of error counting steps (21%). The thigh-worn device has the lowest agreement ($ICC_{2,1}$=0.75, 95% CI −0.21 to 0.93), the highest bias (overcounted on average 7.3 steps), and the highest percentage of error counting steps (28.8%).

Table . Intraclass correlations and percentage of error in counting steps between the physiotherapist (observer) step count and ActiGraph during the timed up and go test in experiment 1 (n=23).

Body location	Bias	ICC[a]	95% CI	*P* value	Percentage error counting steps
Waist	1.4	0.85	0.65 to 0.94	.001	21
Thigh	7.3	0.75	−0.21 to 0.93	.001	28.8
Ankle	0.9	0.94	0.85 to 0.97	.001	12.8

[a]ICC: intraclass correlation coefficient.

Experiment 2

The acceptability questionnaire was completed by 25 patients. Their responses indicated that 22 (88%) had never used a wearable before, 19 (76%) would wear a device for 5 to 7 days while in hospital, and 16 (64%) were willing to wear the device daily at home, up to 3 months, as part of a research study. Furthermore, 12 (48%) patients felt motivated to move when wearing a device, and 8 (32%) would prefer to wear the device on the wrist. Intercurrences during the 24-hour protocol, include the following challenges with the devices: 1 participant had the thigh device removed and reapplied by nurses, while another participant removed both the ankle and wrist devices due to discomfort and itching. In 1 case, the wrist device was taken off because of itchiness, and another participant forgot the purpose of the devices, leading to the removal of both the wrist and thigh devices. In addition, the wrist device was loosened because of swelling, 1 participant had to remove the ankle device for 24 hours due to a major infection on the lower leg, and another was unable to wear the thigh device because the tape caused irritation.

Of the 30 patients, the ActiGraph devices were positioned on the wrist and ankle in 11 (37%) patients and on the wrist and thigh in 19 (63%) patients for the 24-hour protocol. One thigh-worn device recorded only 1:36 of data during the 24-hour protocol so the data were not included in the analysis. On average, the patients wore the device on the wrist for 22:45 (range 20:58-25:51), on the thigh for 24:36 (range 23:33-26:51), and on the ankle for 20:11 (range 8:03-26:16). Intercurrences reported while wearing the devices included itchiness at the wrist and thigh, patients removing devices due to forgetting their purpose, and patients loosening the device due to joint swelling.

The ActiGraph algorithms for the thigh-worn devices have recently changed, combining lying down and sitting as sedentary posture and including stepping detection (Table 4). The thigh-worn devices identified 100% of sedentary posture while the patient was lying down, 98% while they were sitting, and 91% while they were standing. The ankle-worn devices best identified lying (89%) and standing (84%) postures, and poorly identified the sitting posture (43.2%). In addition, the ankle-worn devices classified the position as sedentary on average 93% of the time when lying, but only 49% while the patients were seated. The wrist-worn devices performed poorly compared with the thigh- and ankle-worn devices, identifying lying down between 49% and 52% of the time and sitting and standing around 15% - 25% of the time. The wrist-worn devices were able to identify 80% of the sedentary posture while the patient was lying down, and 71% while sitting (Table 4).

XSL•FO
RenderX

Table . Percentage of times the ActiGraph correctly detected body posture compared with the physiotherapist recordings during experiment 2 (N=30).

Patient's body posture	Experiment 2 devices' attachment										
	Wrist, %			Thigh, %				Ankle, %			
	Sed[a]	Lie[b]	Sit	Stand	Sed	Stand	Step	Sed	Lie	Sit	Stand
Lie	84	50[c]	34	8	100[c]	0	0	92	89[c]	3	1.3
Sit	72	46	26[c]	23	98[c]	1	1	47	4	43[c]	49
Stand	76	10	66	24[c]	8	91[c]	0	13	0	13	84[c]
Off[d]	—	8	5	0	—	—	—	—	6	4	2

[a]Sed: sedentary (lying down + sitting).

[b]Lie: laying down.

[c]Percentage of time the ActiGraph device correctly identified the body postures.

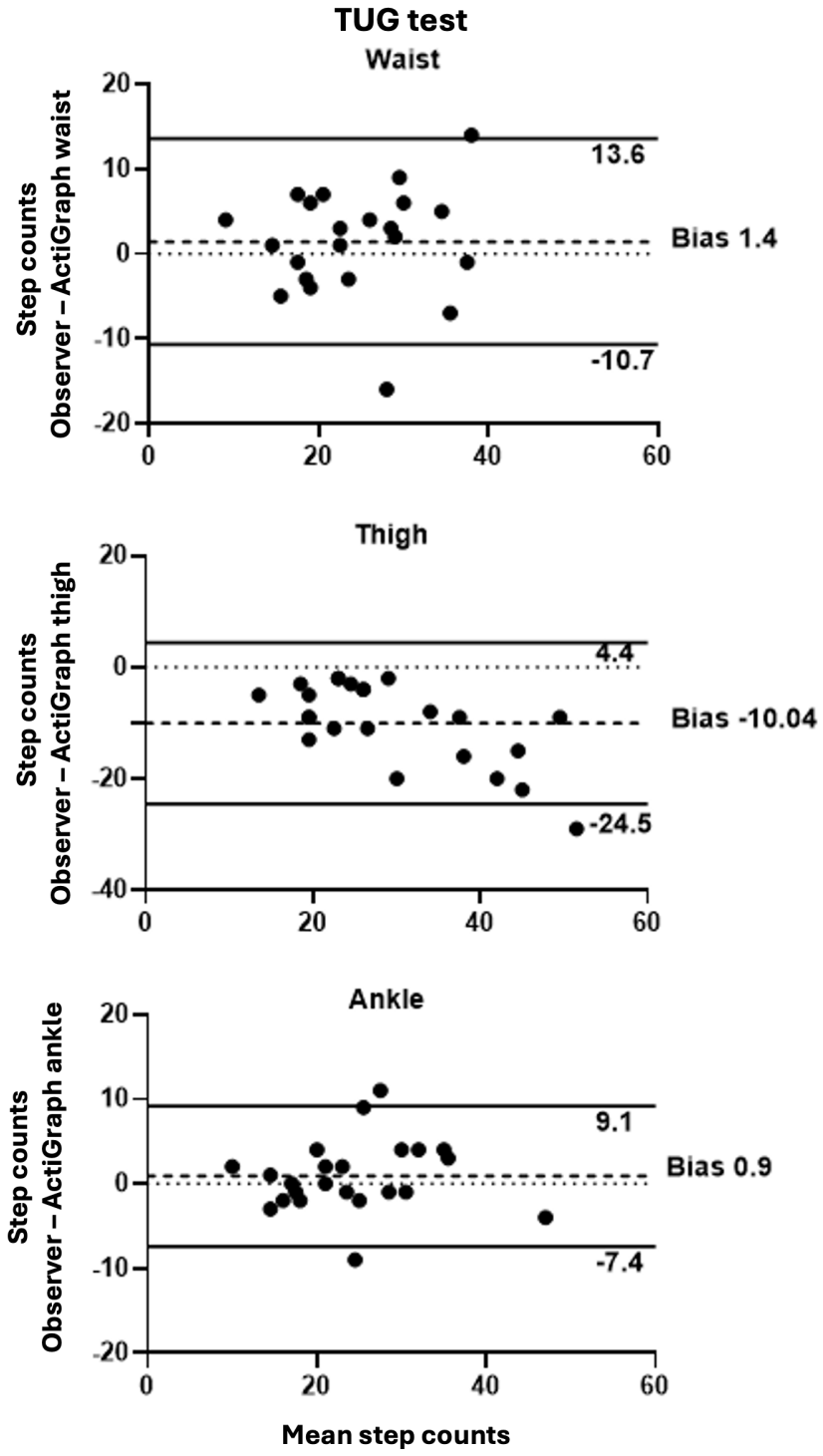
[d]Off: percentage time that the device detected it was off.

[e]Not applicable

Figure 2 shows the Bland-Altman plots of the difference between the observed step count and the ActiGraph placed on the waist, thigh, and ankle during the TUG test during experiment 1. Visual interpretation of the plots shows less bias (0.9), narrower limits of agreement (−7.4 to 9.1), and more observations closer to zero for the ankle-worn ActiGraph. In contrast, the thigh-worn device showed a larger limit of agreement, suggesting poor agreement between the 2 measurements, and a tendency to overcount steps.

XSL•FO
RenderX

Figure 2. Bland-Altman plots comparing observer measurements with ActiGraph data from devices worn on the waist, thigh, and ankle during the TUG test in experiment 1. The dashed lines denote the 95% limits of agreement (SD 1.96) of the mean difference, with darker dashed lines highlighting the mean difference or bias. TUG: timed up and go.



During the 30MWT in experiment 2, the wrist- and thigh-worn devices showed poor ICC values and a high percentage of miscounted steps. The ankle-worn devices showed excellent reliability and on average overcounted the number of steps by 1.9% (Table 5).

XSL•FO
RenderX

Table . Intraclass correlation coefficients and absolute percentage error of steps between the physiotherapist step count and the ActiGraph during the 30-Minute Walk Test in experiment 2 (n=30).

Body location	Bias	ICC[a]	95% CI	*P* value	Percentage error counting steps
Wrist	–42.1	–0.032	–0.217 to 0.218	.62	49.4
Thigh	–36.0	0.331	–0.149 to 0.706	<.001	56.8
Ankle	1.6	0.959	0.915 to 0.981	<.001	1.9

[a]ICC: intraclass correlation coefficient.

The Bland-Altman plots in Figure 3 illustrate the agreement between observer-rated step counts and ActiGraph step counts at the wrist, thigh, and ankle locations during the 30MWT. Among the 3 locations, the ankle-worn device demonstrated the best agreement with observer ratings, exhibiting a small bias and most observations falling within the 95% limits of agreement. In contrast, the wrist-worn device exhibited poor agreement, with the largest bias and widest limits of agreement.

XSL•FO
RenderX

Figure 3. Bland-Altman plots comparing observer measurements with ActiGraph data from devices worn on the wrist, thigh, and ankle during the 30 MWT in experiment 2. The dashed lines denote the 95% limits of agreement (SD 1.96) of the mean difference), with darker dashed lines highlighting the mean difference or bias. 30 MWT: 30-meter walk test.

XSL•FO
RenderX

Discussion

This study aimed to identify the optimal wearable device and wear location to assess and monitor mobility among older patients during hospitalization. We observed a high level of acceptability and feasibility regarding the usability and accuracy of wearable devices for detecting and monitoring activity in older patients hospitalized for acute medical illness. Although most patients had limited experience with activity monitoring devices, they were willing to wear one while hospitalized. Patients showed a preference for the Fitbit and ActiGraph devices, worn at the ankle or waist. Both the ActiGraph and Mox1 devices were easy to set up, but only the ActiGraph allowed for the retrieval of all data collected. Overall, the ActiGraph wGT3X-BT emerged as the preferred device with superior usability, data acquisition, and management.

The ActiGraph wGT3X-BT is widely recognized as the gold standard research-grade device for wearable mobility tracking, which is consistent with the findings of this study in our population. Its previous use in hospital settings underscores its utility for monitoring physical activity and posture in clinical contexts [28-30]. For example, other studies have employed the ActiGraph to predict hospital-acquired disability [28] and demonstrated its accuracy in quantifying postures and activity levels among hospitalized adults [29]. This study supports the device's reliability and validity among older hospitalized patients with unique challenges such as slower movement patterns and increased sedentary behavior.

However, when planning a cohort study in a hospital setting, it is crucial to consider multiple factors beyond just device accuracy, including accessibility to raw data and patient acceptability. During our study, the Fitbit Versa 1 was a popular device choice but lacked the option to download raw data directly. Instead, the download process relied on a companion device, a tablet, and any miscommunication between the 2 devices resulted in data lost. The MMC was selected for its affordability and open-source platform. However, we encountered challenges with the downloading process, which was time-consuming, and we also noted timestamp errors. The MATLAB function provided by the MOX 1 also resulted in missing data and lack of time stamps. Despite the challenges encountered with the MMC, MOX 1, and Fitbit devices in our study, it is plausible that these issues have since been addressed and resolved, underscoring the potential advancements made by these companies in their software and device functionalities. Our study highlights the importance of considering broader factors, such as data accessibility and device functionality, when evaluating the feasibility of wearable technology for hospital-based cohort studies.

Studies have suggested the lack of physical activity and immobilization during hospitalization may be more related to aspects of hospital care rather than to the patient's diagnosis [1,31]. While performance-based mobility tests can predict functional decline and hospital discharge, they are not commonly integrated into hospital measures [32-34]. Conducting a gait speed test, for example, is time-consuming and might not be feasible on a day-to-day basis in the hospital setting. Therefore, the use of wearable technology could be an attractive option, requiring minimal time investment for both patients and health professionals. There is a substantial body of literature on the utilization of wearables in hospitalized patients [16,28,34,35]. However, the research on patient feedback regarding their experiences and perceptions of wearing the devices while hospitalized is limited [17], which is crucial for optimizing their use in health care settings. In addition, we were particularly concerned about the potential interference of these devices with patients' medical conditions. Prolonged sitting and fluid intake, common in hospitalized patients, can lead to swelling, particularly in the ankles and wrists. We observed instances of wrist and ankle swelling in patients that resulted in the devices being removed, underscoring the importance of considering these issues when implementing wearable technology in hospital settings. Furthermore, we also considered the possibility of the devices interfering with medical equipment commonly used during hospitalization, including colostomy bags, urinary catheters, and wound dressings, as any disruption to these essential devices could compromise patient care and safety. For instance, in our first experiment, we applied an elastic band with pockets to the patient's waist, thigh, and ankle. We noticed that the waistband was problematic because it would interfere with heart sensor wires and colostomy bags. In addition, the thigh elastic band would easily fall from the participant's leg. Thus, in the second experiment, we disregarded the waistband and used an adhesive patch, which could also bring discomfort due to skin itchiness, as observed in our study. Considering these factors before launching a larger cohort is crucial and might save time and effort for both patients and researchers.

Our findings highlight the importance of considering both practical and contextual factors when selecting wear locations for mobility monitoring in hospital settings. While the literature supports the thigh for measuring postural behaviors and the waist or ankle for step counts, these locations may not always be feasible in clinical environments. Wrist-worn devices, although convenient and widely accepted, presented challenges such as patient discomfort due to swelling and skin irritation. This study serves as a preliminary exploration of acceptable and practical wear locations in a hospital setting, emphasizing the need to balance feasibility with the specific mobility metrics of interest.

It is well-established that sedentary behavior in hospitalized older adults is associated with a heightened risk of hospital-acquired disability (HAD) and functional decline [28,29]. Therefore, accurately measuring sedentary behavior in this population is essential for timely intervention and management strategies within the hospital setting. Body posture poses a challenge in accurate measurement, ideally requiring the use of at least 2 devices. For instance, thigh-worn devices have been reported as optimal for placing accelerometers to determine sedentary behavior (ie, lying and sitting). In this regard, the thigh-worn inclinometer algorithm provided by ActiGraph uses a movement threshold and thigh angular orientation to distinguish lying and sitting from standing and stepping [36]. Compared with other algorithms such as the activPAL, the ActiGraph's thigh angular parameter improves the classification of sitting posture, even when the participant

XSL•FO
RenderX

has their legs crossed or stretched, in both laboratory and free-living conditions [22,37]. Waist-worn devices are effective for differentiating lying from sitting but not from sitting and standing, while ankle-worn devices distinguish lying and standing. In our study, we tested the ActiGraph on the waist, thigh, and ankle, and in experiment 1, all wear locations performed well for lying and standing but poorly for sitting. We believe that the thigh elastic band likely changed position when participants transitioned from lying to sitting, changing the orientation of the ActiGraph, thus affecting the algorithms' ability to detect sitting posture. In experiment 2, we incorporated a wrist-worn device due to its user-friendly nature. To mitigate the issue encountered in experiment 1 with the elastic band, we adopted adhesive patches to secure the devices on the thigh and ankle. The thigh device exhibited the highest performance across various postures, followed by the ankle device, which showed adequate accuracy in detecting lying and standing postures. However, the wrist device, despite its ease of use, performed poorly in detecting all body postures. Our findings are supported by the literature [17,29], and we recommend using an accelerometer on the thigh and ankle to capture more detailed information on sedentary behavior.

Our investigation on step counts showed that the ankle is the most accurate body position to capture this metric. This finding aligns with previous research examining the validity of ankle-worn ActiGraph devices in conjunction with the lower frequency extension filter for step counting among hospitalized older adults [16,38]. In their study, Webber and St John [38] demonstrated that the ActiGraph positioned on the ankle was comparable with direct observation (ICC=0.94, median absolute error=2.5%) for monitoring step counts during the 10-meter walk test in hospitalized older adults. In addition, Anderson et al [16] reported that the ActiGraph positioned on the ankle (mean difference=−0.85 steps, ICC=0.99) accurately records step counts in hospitalized adults during free self-selected walking. By leveraging the ankle and thigh as the placement site for capturing mobility, researchers and health care practitioners can enhance the accuracy of monitoring and promote more effective interventions aimed at improving mobility and overall health outcomes in hospitalized adults and similar populations.

A potential limitation of this study is the bias introduced by patients wearing multiple devices simultaneously in experiment 1. The discomfort or inconvenience of wearing several devices may have influenced their feedback, as they were likely focused on the overall experience rather than evaluating each device individually. However, the primary goal of experiment 1 was to assess the feasibility of using multiple devices in a hospital setting, with a focus on understanding the practical and real-world challenges of device wearability, data retrieval, and integration in a clinical environment. While this may introduce some bias in the acceptability results, it does not diminish the value of these insights into how devices function together in practice.

Building on these findings, our study has several strengths, including identifying optimal devices and placement sites for wearable data collection, as well as validation through comparison with a gold-standard observer. However, it is important to note that our free-living protocol in experiment 2 was limited to a 24-hour duration. Despite this constraint, we were able to gather valuable insights, particularly regarding the feasibility and integration of wearable technology in hospital settings.

In conclusion, our study found that the ActiGraph wGT3X-BT was the most feasible device for assessing and monitoring mobility among older hospitalized patients. The ActiGraph's thigh-worn algorithm accurately detects sedentary behavior under supervised conditions and, when paired with a device at the ankle, provides detailed information on lying and sitting postures. In addition, our findings indicate that step counts can be accurately detected using the low-frequency extension with devices on the ankle. Therefore, we recommend the use of two devices, at the thigh and ankle, to accurately measure sedentary behavior and step count among older people in a hospital setting. Longer-term studies are warranted to evaluate the use of wearable data for predicting health outcomes after hospitalization and for informing clinical decision-making and efforts to promote early mobility among older hospitalized patients.

Acknowledgments

This project was supported by grants from the McMaster Institute for Research on Aging, McMaster University. MB holds a Tier2 Canada Research Chair in Mobility, Aging and Chronic Disease (CRC-2020-00043) and an Early Researcher Award from the Government of Ontario. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' Contributions

LG, MD, and MB contributed to the conception and design of the work. RK, CC, and PN contributed to the acquisition and analysis of data. PN and RK contributed to the interpretation of data. YH, RZ, and SR contributed to the development of the software for data analysis of the MetaMotionC wearable. All authors contributed to drafting the work (or substantial revisions). All authors have read and approved the final manuscript and agree to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Krumholz HM. Post-hospital syndrome--an acquired, transient condition of generalized risk. N Engl J Med 2013 Jan 10;368(2):100-102. [doi: 10.1056/NEJMp1212324] [Medline: 23301730]
2. Brown CJ, Foley KT, Lowman JD Jr, et al. Comparison of posthospitalization function and community mobility in hospital mobility program and usual care patients: a randomized clinical trial. JAMA Intern Med 2016 Jul 1;176(7):921-927. [doi: 10.1001/jamainternmed.2016.1870] [Medline: 27243899]
3. Creditor MC. Hazards of hospitalization of the elderly. Ann Intern Med 1993 Feb 1;118(3):219-223. [doi: 10.7326/0003-4819-118-3-199302010-00011] [Medline: 8417639]
4. Brown CJ, Redden DT, Flood KL, Allman RM. The underrecognized epidemic of low mobility during hospitalization of older adults. J Am Geriatr Soc 2009 Sep;57(9):1660-1665. [doi: 10.1111/j.1532-5415.2009.02393.x] [Medline: 19682121]
5. Jawad BN, Petersen J, Andersen O, Pedersen MM. Variations in physical activity and sedentary behavior during and after hospitalization in acutely admitted older medical patients: a longitudinal study. BMC Geriatr 2022 Mar 15;22(1):209. [doi: 10.1186/s12877-022-02917-8] [Medline: 35291952]
6. Rojer AGM, Ramsey KA, Trappenburg MC, et al. Patterns of objectively measured sedentary behavior and physical activity and their association with changes in physical and functional performance in geriatric rehabilitation inpatients. J Am Med Dir Assoc 2023 May;24(5):629-637. [doi: 10.1016/j.jamda.2023.01.011] [Medline: 36841261]
7. Callen BL, Mahoney JE, Grieves CB, Wells TJ, Enloe M. Frequency of hallway ambulation by hospitalized older adults on medical units of an academic hospital. Geriatr Nurs (Lond) 2004 Jul;25(4):212-217. [doi: 10.1016/j.gerinurse.2004.06.016]
8. Reynolds CD, Brazier KV, Burgess EAA, et al. Effects of unstructured mobility programs in older hospitalized general medicine patients: a systematic review and meta-analysis. J Am Med Dir Assoc 2021 Oct;22(10):2063-2073. [doi: 10.1016/j.jamda.2020.12.008] [Medline: 33434569]
9. Trost SG, O'Neil M. Clinical use of objective measures of physical activity. Br J Sports Med 2014 Feb;48(3):178-181. [doi: 10.1136/bjsports-2013-093173] [Medline: 24311601]
10. Ngueleu AM, Barthod C, Best KL, Routhier F, Otis M, Batcho CS. Criterion validity of ActiGraph monitoring devices for step counting and distance measurement in adults and older adults: a systematic review. J Neuroeng Rehabil 2022 Oct 17;19(1):112. [doi: 10.1186/s12984-022-01085-5] [Medline: 36253787]
11. Berendsen BAJ, Hendriks MRC, Meijer K, Plasqui G, Schaper NC, Savelberg HHCM. Which activity monitor to use? Validity, reproducibility and user friendliness of three activity monitors. BMC Public Health 2014 Jul 24;14:749. [doi: 10.1186/1471-2458-14-749] [Medline: 25059233]
12. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. Int J Behav Nutr Phys Act 2015 Dec 18;12:159. [doi: 10.1186/s12966-015-0314-1] [Medline: 26684758]
13. Van Remoortel H, Giavedoni S, Raste Y, et al. Validity of activity monitors in health and chronic disease: a systematic review. Int J Behav Nutr Phys Act 2012 Jul 9;9(1):84. [doi: 10.1186/1479-5868-9-84] [Medline: 22776399]
14. Karaca A, Demirci N, Yılmaz V, Hazır Aytar S, Can S, Ünver E. Validation of the ActiGraph WGT3X-BT accelerometer for step counts at five different body locations in laboratory settings. Meas Phys Educ Exerc Sci 2022 Jan 2;26(1):63-72. [doi: 10.1080/1091367X.2021.1948414]
15. Baldwin C, van Kessel G, Phillips A, Johnston K. Accelerometry shows inpatients with acute medical or surgical conditions spend little time upright and are highly sedentary: systematic review. Phys Ther 2017 Nov 1;97(11):1044-1065. [doi: 10.1093/ptj/pzx076] [Medline: 29077906]
16. Anderson JL, Yoward LS, Green AJ. A study investigating the validity of an accelerometer in quantification of step count in adult hospital inpatients recovering from critical illness. Clin Rehabil 2019 May;33(5):936-942. [doi: 10.1177/0269215519829893] [Medline: 30764647]
17. Hartley P, Keevil VL, Westgate K, et al. Using accelerometers to measure physical activity in older patients admitted to hospital. Curr Gerontol Geriatr Res 2018;2018:3280240. [doi: 10.1155/2018/3280240] [Medline: 30420882]
18. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991 Feb;39(2):142-148. [doi: 10.1111/j.1532-5415.1991.tb01616.x] [Medline: 1991946]
19. Andersson M, Moberg L, Svantesson U, Sundbom A, Johansson H, Emtner M. Measuring walking speed in COPD: test-retest reliability of the 30-metre walk test and comparison with the 6-minute walk test. Prim Care Respir J 2011 Dec;20(4):434-440. [doi: 10.4104/pcrj.2011.00082] [Medline: 21938352]
20. Bohm PE, Fehlings MG, Kopjar B, et al. Psychometric properties of the 30-m walking test in patients with degenerative cervical myelopathy: results from two prospective multicenter cohort studies. Spine J 2017 Feb;17(2):211-217. [doi: 10.1016/j.spinee.2016.08.033] [Medline: 27592193]
21. Feito Y, Hornbuckle LM, Reid LA, Crouter SE. Effect of ActiGraph's low frequency extension for estimating steps and physical activity intensity. PLoS One 2017;12(11):e0188242. [doi: 10.1371/journal.pone.0188242] [Medline: 29155845]
22. Steeves JA, Bowles HR, McClain JJ, et al. Ability of thigh-worn ActiGraph and activPAL monitors to classify posture and motion. Med Sci Sports Exerc 2015 May;47(5):952-959. [doi: 10.1249/MSS.0000000000000497] [Medline: 25202847]
23. Choi L, Liu Z, Matthews CE, Buchowski MS. Validation of accelerometer wear and nonwear time classification algorithm. Med Sci Sports Exerc 2011 Feb;43(2):357-364. [doi: 10.1249/MSS.0b013e3181ed61a3] [Medline: 20581716]

XSL•FO
RenderX

24. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979 Mar;86(2):420-428. [doi: 10.1037//0033-2909.86.2.420] [Medline: 18839484]
25. Cicchetti DV, Sparrow SA. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. Am J Ment Defic 1981 Sep;86(2):127-137. [Medline: 7315877]
26. Giavarina D. Understanding Bland Altman analysis. Biochem Med (Zagreb) 2015;25(2):141-151. [doi: 10.11613/BM.2015.015] [Medline: 26110027]
27. Martin Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986 Feb;327(8476):307-310. [doi: 10.1016/S0140-6736(86)90837-8] [Medline: 2868172]
28. Pavon JM, Sloane RJ, Pieper CF, et al. Accelerometer-measured hospital physical activity and hospital-acquired disability in older adults. J Am Geriatr Soc 2020 Feb;68(2):261-265. [doi: 10.1111/jgs.16231] [Medline: 31747050]
29. Anderson JL, Green AJ, Yoward LS, Hall HK. Validity and reliability of accelerometry in identification of lying, sitting, standing or purposeful activity in adult hospital inpatients recovering from acute or critical illness: a systematic review. Clin Rehabil 2018 Feb;32(2):233-242. [doi: 10.1177/0269215517724850] [Medline: 28805075]
30. Kirk AG, Kimmel LA, Behm KJ, Peiris CL, Ekegren CL. Validity of the activPAL and ActiGraph for measuring sitting time and steps in hospitalised orthopaedic patients with altered weight bearing. Disabil Rehabil 2024 Jan;46(2):378-386. [doi: 10.1080/09638288.2022.2157896] [Medline: 36541196]
31. Zisberg A, Syn-Hershko A. Factors related to the mobility of hospitalized older adults: a prospective cohort study. Geriatr Nurs 2016;37(2):96-100. [doi: 10.1016/j.gerinurse.2015.10.012] [Medline: 26597674]
32. Ibrahim K, Owen C, Patel HP, et al. Can routine clinical data identify older patients at risk of poor healthcare outcomes on admission to hospital? BMC Res Notes 2017 Aug 10;10(1):384. [doi: 10.1186/s13104-017-2705-7] [Medline: 28797300]
33. Kumar A, Karmarkar AM, Graham JE, et al. Comorbidity indices versus function as potential predictors of 30-day readmission in older patients following postacute rehabilitation. J Gerontol A Biol Sci Med Sci 2017 Feb;72(2):223-228. [doi: 10.1093/gerona/glw148] [Medline: 27492451]
34. Fisher SR, Graham JE, Ottenbacher KJ, Deer R, Ostir GV. Inpatient walking activity to predict readmission in older adults. Arch Phys Med Rehabil 2016 Sep;97(9 Suppl):S226-S231. [doi: 10.1016/j.apmr.2015.09.029] [Medline: 27264549]
35. van Roekel EH, Winkler EAH, Bours MJL, et al. Associations of sedentary time and patterns of sedentary time accumulation with health-related quality of life in colorectal cancer survivors. Prev Med Rep 2016 Dec;4:262-269. [doi: 10.1016/j.pmedr.2016.06.022] [Medline: 27419042]
36. How is inclination determined (for thigh wear location)? ActiGraph. URL: https://actigraphcorp.my.site.com/support/s/article/How-is-Inclination-Determined-for-Thigh-Wear-Location [accessed 2023-03-12]
37. Edwardson CL, Rowlands AV, Bunnewell S, et al. Accuracy of posture allocation algorithms for thigh- and waist-worn accelerometers. Med Sci Sports Exerc 2016 Jun;48(6):1085-1090. [doi: 10.1249/MSS.0000000000000865] [Medline: 26741122]
38. Webber SC, St John PD. Comparison of ActiGraph GT3X+ and StepWatch step count accuracy in geriatric rehabilitation patients. J Aging Phys Act 2016 Jul;24(3):451-458. [doi: 10.1123/japa.2015-0234] [Medline: 26751505]

Abbreviations

30MWT: 30-meter walking test
ICC: intraclass correlation coefficient
MMC: MetaMotionC
TUG: timed up and go

Edited by J Wang; submitted 21.07.24; peer-reviewed by A Habib, E Cenko; revised version received 13.12.24; accepted 30.01.25; published 12.05.25.

Please cite as:
Nascimento P, Kirkwood R, Griffith LE, Duong M, Cooper C, Hao Y, Zheng R, Raza S, Beauchamp M
Identifying Optimal Wearable Devices for Monitoring Mobility in Hospitalized Older Adults: Feasibility, Acceptability, and Validity Study
JMIR Aging 2025;8:e64372
URL: https://aging.jmir.org/2025/1/e64372
doi:10.2196/64372

© Paulo Nascimento, Renata Kirkwood, Lauren E Griffith, Mylinh Duong, Cody Cooper, Yujiao Hao, Rong Zheng, Samir Raza, Marla Beauchamp. Originally published in JMIR Aging (https://aging.jmir.org), 12.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits

XSL•FO
RenderX

unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Performance of a Digital Cognitive Assessment in Predicting Dementia Stages Delineated by the Dementia Severity Rating Scale: Retrospective Study

Duong Huynh[1], PhD; Kevin Sun[1], BS; Mary Patterson[1], BA, MS; Reza Hosseini Ghomi[1,2,3], MD, MSE; Bin Huang[1], PhD

[1]BrainCheck, Inc, Austin, TX, United States

[2]Frontier Psychiatry, PLLC, Billings, MT, United States

[3]Department of Neurology, Institute for Neuroengineering, & eScience Institute, University of Washington, Seattle, WA, United States

Corresponding Author:
Bin Huang, PhD
BrainCheck, Inc, Austin, TX, United States

Abstract

Background: Dementia is characterized by impairments in an individual's cognitive and functional abilities. Digital cognitive assessments have been shown to be effective in detecting mild cognitive impairment and dementia, but whether they can stage the disease remains to be studied.

Objective: In this study, we examined (1) the correlation between scores obtained from BrainCheck standard battery of cognitive assessments (BC-Assess), a digital cognitive assessment, and scores obtained from the Dementia Severity Rating Scale (DSRS), and (2) the accuracy of using the BC-Assess score to predict dementia stage delineated by the DSRS score. We also explored whether BC-Assess can be combined with information from the Katz Index of Independence in activities of daily living (ADL) to obtain enhanced accuracy.

Methods: Retrospective analysis was performed on a BrainCheck dataset containing 1751 patients with dementia with different cognitive and functional assessments completed for cognitive care planning, including the DSRS, the ADL, and the BC-Assess. The patients were staged according to their DSRS total score (DSRS-TS): 982 mild (DSRS-TS 10 - 18), 656 moderate (DSRS-TS 19-26), and 113 severe (DSRS-TS 37-54) patients. Pearson correlation was used to assess the associations between BC-Assess overall score (BC-OS), ADL total score (ADL-TS), and DSRS-TS. Logistic regression was used to evaluate the possibility of using patients' BC-OS and ADL-TS to predict their stage.

Results: We found moderate Pearson correlations between DSRS-TS and BC-OS (r=−0.53), between DSRS-TS and ADL-TS (r=−0.55), and a weak correlation between BC-OS and ADL-TS (r=0.37). Both BC-OS and ADL-TS significantly decreased with increasing severity. BC-OS demonstrated to be a good predictor of dementia stages, with an area under the receiver operating characteristic curve (ROC-AUC) of classification using logistic regression ranging from .733 to .917. When BC-Assess was combined with ADL, higher prediction accuracies were achieved, with an ROC-AUC ranging from 0.786 to 0.961.

Conclusions: Our results suggest that BC-Assess could serve as an effective alternative tool to DSRS for grading dementia severity, particularly in cases where DSRS, or other global assessments, may be challenging to obtain due to logistical and time constraints.

(JMIR Aging 2025;8:e65292) doi:10.2196/65292

KEYWORDS

stage; severity; progression; correlation; association; cognitive impairment; functional activities; cognitive assessment; BrainCheck; dementia; Alzheimer disease; gerontology; geriatric; old; elderly; aging; retrospective analysis; digital assessment; patient assessment; digital cognitive assessment; digital health; neurodegeneration; memory loss; memory function; risk factors

Introduction

Dementia is characterized by impairments in an individual's cognitive and everyday functional abilities. Although the pattern and advancement of these impairments vary among patients, the disease is usually considered as having 3 main stages: mild (early-stage), moderate (middle-stage), and severe (late-stage) [1]. The distinction between these categories lies in the extent to which a patient's physical, cognitive, and psychosocial well-being deteriorates. The rate of deterioration widely varies and progresses through subtle changes in daily functioning to complete loss of independence and the need for a caregiver [1]. Staging of dementia has important implications for practical

XSL•FO
RenderX

decision-making and research [2]. From the practical standpoint, knowledge of disease severity is helpful for selection of appropriate intervention options, for prognosis and communication with patients and their family about expectations, care needs, as well as early planning for the future [3-5]. From the research standpoint, staging patients is needed to determine their eligibility for participation, to achieve better clinical efficacy, and to obtain homogeneous sampling in research studies [4], particularly clinical trials.

As noted earlier, dementia is not a monolithic disease; therefore, both cognitive and functional abilities need to be measured to accurately assess its severity and progression. Standardized cognitive tests can provide objective measures of cognitive functioning in different domains such as memory and executive function. Brief cognitive screening tests such as the Montreal Cognitive Assessment and the Mini-Mental State Examination are widely used in clinical practices, but they are rarely used to grade dementia severity [6,7]. Formal neuropsychological tests provide a more comprehensive evaluation of cognitive functioning to support differential diagnoses [8,9]. However, these tests typically take hours and require administration by specialists. Digital cognitive assessments are emerging as an efficient solution due to their self-administration capability, remote accessibility, and automated scoring. Although these types of assessments have been shown to be effective in detecting mild cognitive impairment and dementia [10-12], their ability to stage dementia is not yet clear. Functional assessments quantify the ability to perform activities of daily living through questionnaires such as the Katz Index of Independence in activities of daily living (ADL) [13,14]. Functional assessments are valuable for helping to evaluate dementia severity and also for providing proper guidance to patients and their caregivers.

In clinical practice, both cognitive and functional deficits can be measured using global staging scales. These scales typically come in the form of a questionnaire or interview, relying on subjective judgments and reports from patients or their knowledgeable informants. Compared with other scales, such as the Global Deterioration Scale [15] or the Functional Assessment Staging [16], the Clinical Dementia Rating (CDR) [17-19] scale appears to be the most well-studied and best-evidenced for dementia staging [20,21]. However, the use of this instrument is often impractical in many situations where time and cost are concerns, due to its semi-structured nature, long administration time (30 - 60 min), and requirement of clinical judgment from a trained professional during administration and scoring [18,22]. In response to this, brief instruments have been developed to mirror the CDR, including the Dementia Severity Rating Scale (DSRS) [23]. The DSRS uses a multiple-choice format in which the caregiver rates the patient's cognitive and functional ability in 12 categories [23,24]. The DSRS has been shown to be effective in staging and determining the progression of dementia [23,25,26].

While cognitive tests have been shown to correlate with the above dementia staging tools [4,27,28], previous research primarily focused on traditional paper-based cognitive tests. The increasing adoption of digital solutions and tools in health care calls for the re-evaluation of dementia staging tools and digital cognitive assessments. The first goal of this study was to examine the correlation between scores obtained from the BrainCheck standard battery of cognitive assessments (BC-Assess), a digital cognitive assessment, and scores obtained from the DSRS, a global staging scale. The second goal of this study is to evaluate the effectiveness of using the BC-Assess score to predict dementia stage delineated by the DSRS score. We also explored whether BC-Assess can be combined with information from ADL to obtain enhanced predictive capability.

Methods

Data Source

This retrospective study analyzed a real-world dataset of patients who received cognitive care planning services from their providers through BrainCheck Plan. These patients and their caregivers had completed a series of assessments, including DSRS, ADL, and BC-Assess, and received a comprehensive and personalized cognitive care plan. Inclusion criteria for this study were (1) patients 60 years of age or older; (2) assessments of cognitive care planning completed in English on an iPad; and (3) Dementia Severity Rating Scale total score (DSRS-TS) ≥10. The criterion for DSRS-TS was to only include patients that were rated by DSRS to have mild, moderate, and severe dementia [29].

Patients that showed evidence of moderate to severe depression, defined by a Geriatric Depression Scale score of 9 or above [30], were excluded. This is to avoid including reversible causes of dementia, which may have poorly correlated impacts on cognitive and functional measures. Given that depression is common among patients with dementia [31-33], patients with mild depression were not excluded to avoid overfiltering of the data. For patients having multiple care plans, only the latest record was considered. To reduce the impact of practice effect, only data from providers who had completed ≥20 cognitive care planning services for their patients were included. In total, data from 1751 patients with their cognitive care plans completed between February 2022 and April 2024 across 48 providers were included for this analysis.

Measurements

The DSRS is a brief informant-based questionnaire made up of 12 items that measure functional abilities, including memory, orientation to time, orientation to place, decision-making, social interaction, home activities, personal care, eating, toileting, mobility, recognition, and speech and language. DSRS-TS is calculated by adding up scores across 12 functional areas, ranging from zero (no impairment) to 54 (extreme impairment) [25]. The patients could be categorized into 3 groups of different dementia severity levels based on their DSRS-TS. The majority 56% (n=982) were mild-stage patients (DSRS-TS 10 - 18), moderate-stage patients (DSRS-TS 19 - 26) accounted for 37.5% (n=656), and the remaining 6.5% (n=113) were severe-stage patients (DSRS-TS 37 - 54). The 3 severity levels serve as class labels in a logistic regression analysis in this study to predict patients' conditions based on their BC-Assess and ADL data.

The BC-Assess, completed by the patients, consisted of 6 individual cognitive assessments: Immediate Recognition and

XSL•FO
RenderX

Delayed Recognition (memory), Digit Symbol Substitution (processing speed), Stroop (executive function), Trails Making Test A, and Trails Making Test B (attention or mental flexibility). Detailed descriptions of these tests can be found in previous studies [12,34]. The raw score for each assessment is calculated using assessment-specific measurements based on accuracy or reaction time (Table S1 in Multimedia Appendix 1). The BC-Assess raw overall score is the average of raw scores from all completed assessments after each score has been transformed from its natural range into a common range [0,100] using the formula in Table S1 in Multimedia Appendix 1. A *z*-score is then calculated for each assessment score and for the overall score using the corresponding age- and device-specific mean and standard deviation values from the BrainCheck normative database. Assessment standardized scores and BC-Assess overall standardized score (BC-OS) are reported by rescaling the *z*-scores to have a mean of 100 and a standard deviation of 15.

The ADL is an informant-based 6-item survey that measures an individual's ability to independently perform basic activities of daily living, including bathing, dressing, going to the toilet, transferring, continence, and feeding. It is calculated by adding up scores from the 6 categories, each of which takes a value of 1 for independent and 0 for dependent, resulting in an ADL total score (ADL-TS) ranging from 0 to 6. An ADL-TS of 2 or less indicates severe functional impairment, 3 - 4 indicates moderate impairment, and 5 - 6 indicates full function [13,35].

Statistical Analysis

Data analyses were performed using Python (version 3.8.5). Descriptives were presented for demographics and each score. The χ^2 test was used to examine whether the distribution of gender was similar in patients from the 3 groups. The Kruskal-Wallis test was used to compare the mean age of patients across groups.

A 3-way multivariate analysis of variance (MANOVA) was used to examine the joint variation of ADL-TS and BC-OS as a function of dementia stage, age group, and gender. Post-hoc analysis employed 1-way MANOVAs to compare these 2 scores across dementia stages for each individual age and gender group. Tests for statistical significance of the mean differences across stages, age groups, and genders were also performed separately for each score using 3-way ANOVAs. For these statistical comparisons, age is treated as a factor with 3 levels: 60 - 69, 70 - 79, and 80+.

Logistic regression was used to investigate the effectiveness of using the patients' BC-OS and ADL-TS to predict their dementia stage, where age and gender served as covariates. In this analysis, age is treated as a continuous variable, and gender is treated as a binary variable: 1 for male and 0 for female. Although the 3 dementia stages form ordinal classes, separate binary logistic regressions were used to classify mild versus moderate; moderate versus severe; and mild versus severe, because the proportional odds assumption was not satisfied for both BC-OS (P=.01) and ADL-TS (P<.001) from Brant's Wald test [36], suggesting ordered logistic regression was not appropriate.

The binary logistic regression model for predicting a patient's condition is:

$$(1)\text{logit}(p)=\beta0+\beta1\cdot\text{stBC}+\beta2\cdot\text{stADL}+\beta3\cdot\text{stAge}+\beta4\cdot\text{stGender}$$

where p is the probability of predicting the patient as having a pre-specified positive class. In each of the above binary classifications, we chose the more severe stage to be the positive class. stBC,stADL,stAge,stGender are standardized values of predictor variables BC-OS, ADL-TS, age, and gender. The coefficients βi`s (i=1 - 4) represent the effects of the 4 predictors, and β0 is the intercept.

Model fitting was based on weighted loss functions to take into consideration class imbalance across the 3 dementia groups. Model performance was evaluated using 5-fold cross-validation with stratification, repeated 20 times (100 iterations), such that on each iteration, all training and testing subsets had roughly the same proportion of patients from each group as in the original dataset. We compared four different models:

(1) full model that included all 4 predictors as in equation (1)

(2) reduced model 1 that included BC-OS and ADL-TS:

$$(2)\text{logit}(p)=\beta0+\beta1\cdot\text{stBC}+\beta2\cdot\text{stADL}$$

(3) reduced model 2 that included BC-OS, age, and gender:

$$(3)\text{logit}(p)=\beta0+\beta1\cdot\text{stBC}+\beta3\cdot\text{stAge}+\beta4\cdot\text{stGender}$$

(4) reduced model 3 that included only BC-OS:

$$(4)\text{logit}(p)=\beta0+\beta1\cdot\text{stBC}$$

An receiver operating characteristic (ROC) curve was generated for each model on each cross-validation iteration. Paired-samples *t*-tests were used to compare areas under the receiver operating characteristic curves (ROC-AUC) across models.

Ethical Considerations

This study was conducted using existing deidentified data collected through BrainCheck. The dataset contained no personal identifiers, and no attempt was made to reidentify the individuals. As such, the research does not meet the definition of "human subjects research" as outlined by HHS 45 CFR 46.102. Therefore, this study did not require ethics review or approval by an institutional review board.

Results

Demographics and Assessment Performance

Table 1 summarizes the demographic characteristics of the patients in this study. Group sample size decreased with increasing severity for both genders. The distribution of gender was similar across the 3 groups (P=.84). Although the range of age was similar, mean age significantly increased with severity (P<.001; pairwise comparisons: P<.001 for mild vs moderate and mild vs severe, P=.02 for moderate vs severe). Statistical comparisons were not performed for education level and race due to a lot of missing information.

XSL•FO
RenderX

Table . Demographics and summary statistics of scores across dementia severity groups and the total sample.

Demographic characteristics	Mild group (n=982)	Moderate group (n=656)	Severe group (n=113)	Total (N=1751)
Gender, n (%)[a]				
Female	589 (60)	394 (60.1)	71 (62.8)	1054 (60.2)
Male	393 (40)	262 (39.9)	42 (37.2)	697 (39.8)
Age, years[b]				
Mean (SD)	78 (8.3)	80.3 (8.4)	80.9 (9.0)	79.0 (8.6)
Range	60 - 101	60 - 102	61 - 101	60 - 102
Age bucket (years), n (%)[c]				
60 - 69	161 (16.4)	74 (11.3)	15 (13.3)	250 (14.3)
70 - 79	393 (40)	232 (35.4)	34 (30.1)	659 (37.6)
≥80	428 (43.6)	350 (53.3)	64 (56.6)	842 (48.1)
Education (years), n (%)				
>12	72 (7.3)	49 (7.5)	1 (0.9)	122 (7)
≤12	48 (4.9)	41 (6.3)	9 (8)	98 (5.6)
Not reported	862 (87.8)	566 (86.3)	103 (91.2)	1531 (87.4)
Race, n (%)				
White	108 (11)	72 (11)	6 (5.3)	186 (10.6)
Black	23 (2.4)	12 (1.8)	2 (1.8)	37 (2.1)
Others	19 (1.9)	17 (2.6)	1 (0.8)	37 (2.1)
Not reported	832 (84.7)	555 (84.6)	104 (92.0)	1491 (85.2)
DSRS-TS[d], mean (SD)	13.4 (2.5)	25.4 (4.8)	43.3 (5.0)	19.8 (9.2)
BC-OS[e], mean (SD)	62.2 (35.1)	30.1 (38.9)	–5.0 (26.8)	45.8 (41.4)
ADL-TS[f], mean (SD)	4.8 (1.7)	3.2 (2.1)	1.2 (1.5)	4.0 (2.1)

[a]P=.84 (χ^2 test): distribution of gender was not significantly different across groups.

[b]P<.001 (Kruskal-Wallis test): mean age of patients was significantly different across groups.

[c]P<.001 (χ^2 test): distribution of age bucket was significantly different across groups.

[d]DSRS-TS: Dementia Severity Rating Scale total score.

[e]BC-OS: BrainCheck standard battery of cognitive assessments overall score.

[f]ADL-TS: activities of daily living total score.

The means and standard deviations of DSRS-TS, ADL-TS, and BC-OS are provided in Table 1. Overall, the BC-OS and ADL-TS decreased with increasing severity delineated by the DSRS-TS. Figure 1 shows the distributions of BC-OS and ADL-TS across patients within each group. For both scores, the distribution was systematically skewed toward the high values for the mild group, toward the low values for the severe group, and more evenly distributed for the moderate group. Table 2 shows the number and percentage of cases in each group that fall in different BC-Assess and ADL-TS score intervals.

XSL•FO
RenderX

Figure 1. Box plots of the distributions of BrainCheck standard battery of cognitive assessments overall score (BC-OS; left) and activities of daily living total score (ADL-TS; right) for each patient group: green=mild, yellow=moderate, and red=severe. Normally distributed random noise was used to add displacements along the x-axis for patients within each group (left and right) and along the y-axis at each ADL-TS value (right).

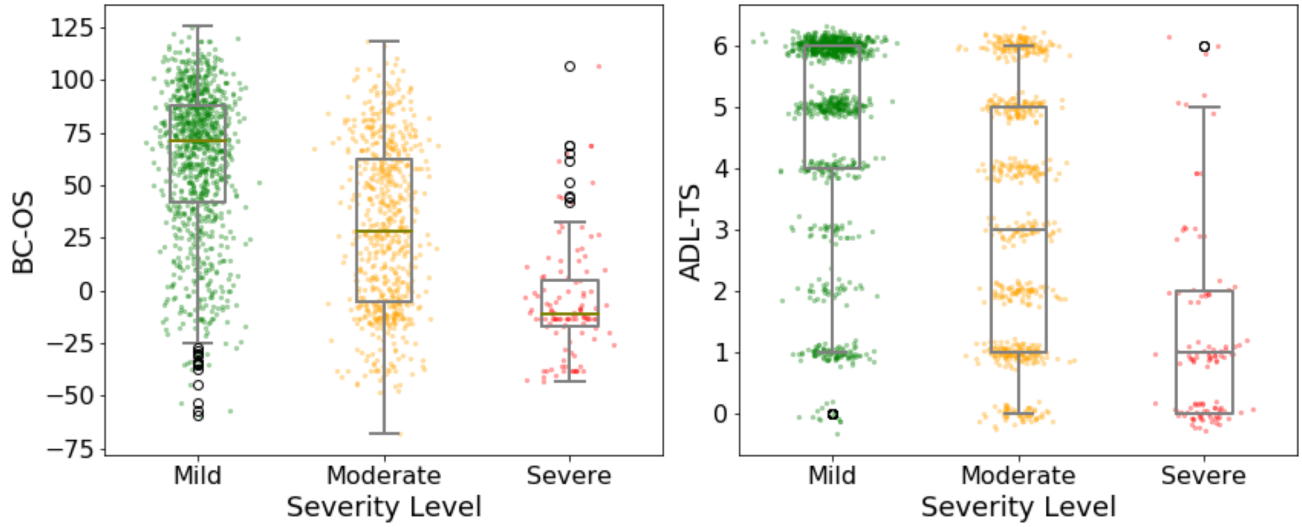


Table . ADL-TS[a] and BC-OS[b] distributions across dementia severity groups.

Score	Mild group (n=982)	Moderate group (n=656)	Severe group (n=113)
BC-OS, n (%)			
Beyond 2 SD of normative mean[c]	480 (48.9)	532 (81.1)	112 (99.1)
Within 2 SD of normative mean	502 (51.1)	124 (18.9)	1 (0.9)
ADL-TS, n (%)			
0 - 2 (severe)	137 (14)	275 (41.9)	96 (85)
3 - 4 (moderate)	119 (12.1)	142 (21.7)	10 (8.8)
5 - 6 (Full function)	726 (73.9)	239 (36.4)	7 (6.2)

[a]ADL-TS: activities of daily living total score.

[b]BC-OS: BrainCheck standard battery of cognitive assessments overall score.

[c]Based on a BC-OS normative mean of 100 and a standard deviation of 15.

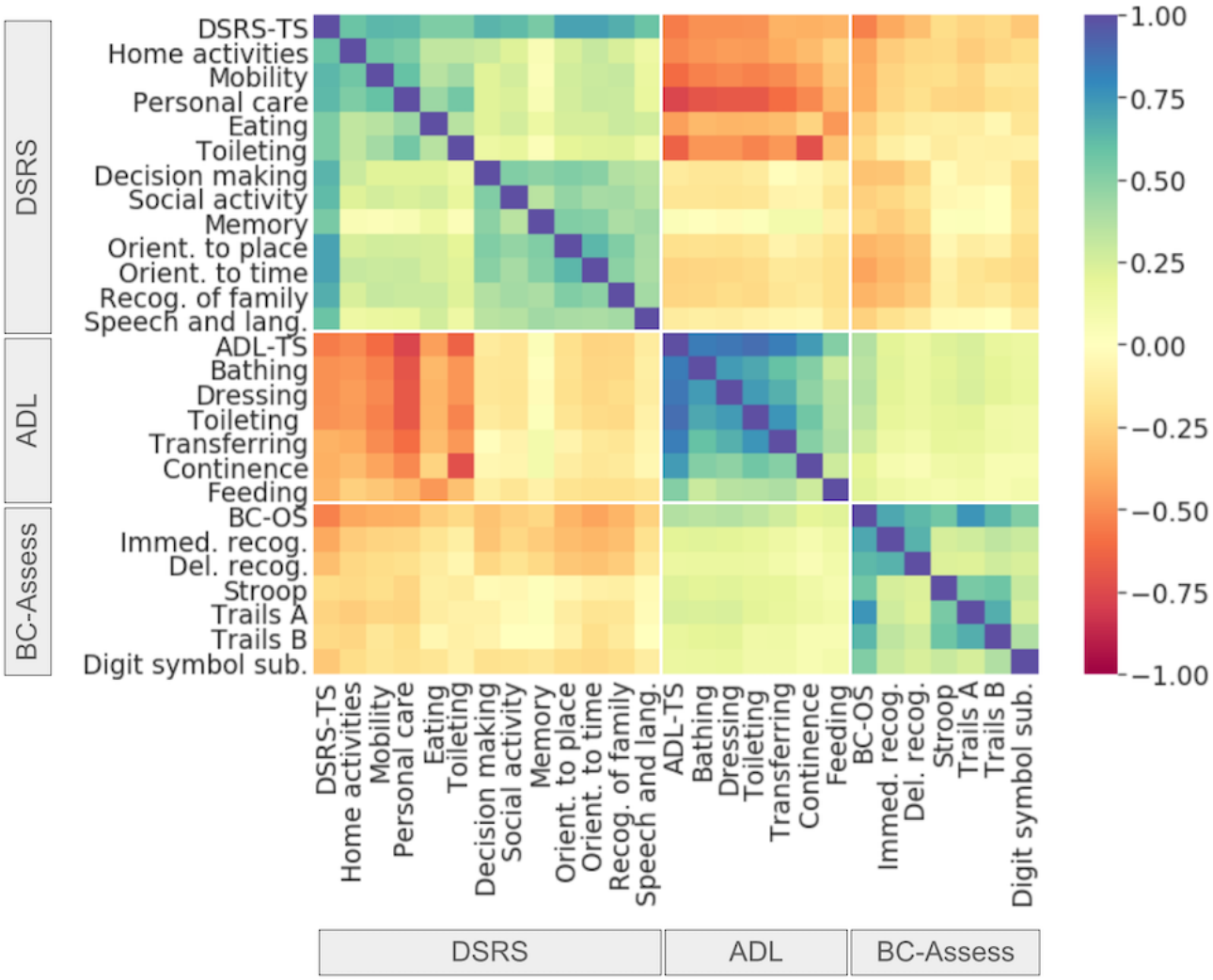
Correlations Between Assessments

We found moderate Pearson correlations between DSRS-TS and BC-OS (r=−0.53; P<.001), between DSRS-TS and ADL-TS (r=−0.55; P<.001), and a weak correlation between BC-OS and ADL-TS (r=0.37; P<.001). Since DSRS covers both cognitive and functional performance of a patient, moderate associations between DSRS-TS with BC-OS and ADL-TS were as expected. The weak correlation between ADL-TS and BC-OS suggests that cognitive and functional abilities are associated with each other, but certain discrepancies exist between the two.

The heatmap in Figure 2 plots Pearson correlations across DSRS, BC-Assess, and ADL subscores. Compared with BC-Assess subscores, BC-OS showed stronger correlations with DSRS and ADL subscores. However, these correlations were weak. Among DSRS subscores, BC-OS was most associated with home-activities (r=−0.41; P<.001), mobility (r=−0.39; P<.001), personal-care (r=−0.39; P<.001), orientation-to-time or orientation-to-place (r=−0.43 and −0.37; P<.001), and recognition-of-family (r=−0.37; P<.001). Among ADL subscores, BC-OS was most associated with dressing (r=0.36; P<.001), bathing (r=0.35; P<.001), and toileting (r=0.33; P<.001). Although weaker, a clear trend can be observed from Figure 2 for BC-Assess subscores. With regard to the DSRS, assessments of memory (Immediate or Delayed Recognition) were most associated with memory-demanding activities such as memory, orientation-to-time, orientation-to-place, recognition-of-family, and decision-making, whereas assessments of executive function, attention, or mental flexibility (Stroop, Trails Making A/B) were most associated with home-activities, mobility, and personal-care. With regard to the ADL, BC-Assess subscores were more associated with bathing, dressing, and toileting than with categories that are more essential, such as feeding, continence, and transferring. Between ADL and DSRS subscores, correlations mainly occurred within a subset of DSRS activities that are of the same types as those rated by the ADL such as eating, home-activities, mobility, personal-care, and toileting. Of these, the strongest correlations were found between DSRS toileting and ADL continence (r=−0.72; P<.001), and between DSRS personal-care and ADL bathing (r=−0.70; P<.001), dressing (r=−0.68; P<.001), and toileting (r=−0.68; P<.001).

XSL•FO
RenderX

Figure 2. Correlations between BC-Assess (BrainCheck standard battery of cognitive assessments), DSRS (Dementia Severity Rating Scale), and ADL (activities of daily living) subscores. DSRS-TS: DSRS total score; ADL-TS: ADL total score; BC-OS: BC-Assess overall score.



Comparison of BC-OS and ADL-TS Across Dementia Stages, Age Groups, and Genders

Given the correlation between BC-OS and ADL-TS, we analyzed the differences in these two scores across dementia stages, age groups, and genders by running a 3-way MANOVA (*BC-OS + ADL-TS~ Dementia Stage * Age Group * Gender*). Results based on the Pillais Trace method showed a significant effect of Dementia Stage (Pillais Trace =0.046 , $F_{43,466}$=20.5; *P*<.001) whereas the effects of age and gender and all interaction terms were not significant. Post hoc 1-way MANOVAs (*BC-OS + ADL-TS~ Dementia Stage*) were run to compare BC-OS and ADL-TS combined between mild versus moderate and between moderate versus severe separately for each age and gender group. Except for the 60 - 69 and female group (n=136) showing a nonsignificant difference between moderate versus severe, significant differences were observed for all other cases. We further performed 3-way ANOVAs where BC-OS (*BC-OS~ Dementia Stage * Age Group * Gender*) and ADL-TS (*ADL-TS~ Dementia Stage * Age Group * Gender*) were considered separately. For BC-OS, we found a significant main effect of Dementia Stage only ($F_{21,733}$=270.31; *P*<.001). The insignificant differences in BC-OS across age groups were as expected as this score had been standardized to adjust for age differences. For ADL-TS, we found significant effects of Dementia Stage ($F_{21,733}$=278.87; *P*<.001), Age Group ($F_{21,733}$=4.77; *P*=.009), and Gender ($F_{11,733}$=7.82; *P*=.005). For both scores, no interaction terms were significant.

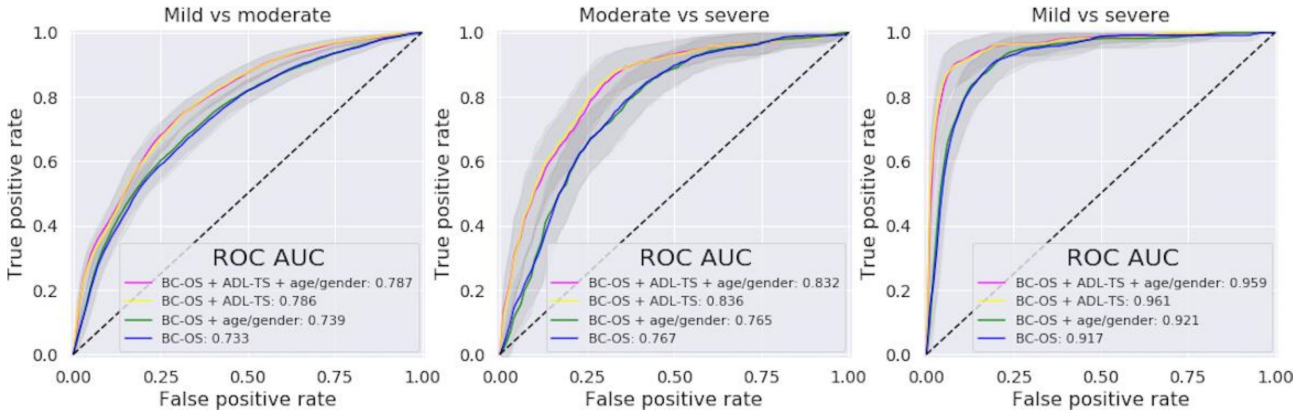
Logistic Regression to Examine the Roles of BC-OS, ADL-TS, Age, and Gender in Predicting a Patient's Condition

ROC analysis (Figure 3) showed a comparable performance between the full model (*BC-OS + ADL-TS + age or gender*: ROC-AUC=0.787 for mild vs moderate; 0.832 for moderate vs severe; and 0.959 for mild vs severe) and reduced model 1 (*BC-OS + ADL-TS*: ROC-AUC=0.786 for mild vs moderate; 0.836 for moderate vs severe; and 0.961 for mild vs severe), and between reduced model 2 (*BC-OS + age or gender*: ROC-AUC=0.739 for mild vs moderate; 0.765 for moderate vs severe; and 0.921 for mild vs severe) and reduced model 3 (*BC-OS* only: ROC-AUC=0.733 for mild vs moderate; 0.767 for moderate vs severe; and 0.917 for mild vs severe). The small differences in ROC-AUC generated by the omission of age and gender suggest that they are not important predictors. Moreover, including these demographic factors appears to have led to some degree of overfitting where reduced model 1 performed slightly but significantly better than the full model (mild vs severe:

XSL•FO
RenderX

P=.002; moderate vs severe: *P*<.001) and reduced model 3 performed slightly but significantly better than reduced model 2 (mild vs moderate: *P*<.001; moderate vs severe: *P*=.04; mild vs severe: *P*<.001).

Figure 3. Model comparisons for the classification between mild and moderate (left) and between moderate and severe (right): receiver operating characteristic curves and mean area under the receiver operating characteristic curve (ROC AUC) for the full model (BrainCheck standard battery of cognitive assessments overall score [BC-OS] + activities of daily living total score [ADL-TS] + age and gender; magenta), reduced model 1 (BC-OS + ADL-TS; yellow), reduced model 2 (BC-OS + age and gender; green), and reduced model 3 (BC-OS only; blue). The shaded area along each curve represents the corresponding standard deviation across cross validation iterations. In each classification, the more severe condition was chosen to be the positive class.



The 2 models that include ADL-TS (full model and reduced model 1) performed significantly better than the 2 models without ADL-TS (reduced models 2 and 3) (*P*<.001 for each pairwise comparison), suggesting the important role of ADL. For all binary classifications, although BC-OS alone can serve as a fairly effective predictor with an ROC-AUC of at least 0.733, including ADL-TS in the model significantly improved prediction accuracy.

As age and gender were not important factors, we excluded them from further analysis. We examined the diagnostic performance of reduced models 1 (*BC-OS + ADL-TS*) and 3 (*BC-OS* only) at the optimal cut-points from their ROC curves. These are points on the ROC curves that maximize true positive rate and minimize false positive rate. Table 3 shows sensitivity (true positive rate) and specificity (1–false positive rate) at the optimal cut-point for each model and classification. When included, ADL-TS improved the sensitivity and specificity by 3% - 5%. For reduced model 3, which included only BC-OS, we found that the optimal cut-point corresponded to a BC-OS of 52 (roughly 3 standard deviations below the normative mean) for the classification between mild versus moderate, and a BC-OS of 0 for the classification between moderate versus severe.

Table . Sensitivity (Se) and specificity (Sp) by model and classification: mean (SD) across cross-validation iterations. In each classification, the more severe condition was chosen to be the positive class.

Classification	BC-OS[a] + ADL-TS[b]		BC-OS only	
	Se, mean (SD)	Sp, mean (SD)	Se, mean (SD)	Sp, mean (SD)
Mild versus moderate	0.74 (0.04)	0.72 (0.04)	0.69 (0.05)	0.68 (0.05)
Moderate versus severe	0.84 (0.06)	0.76 (0.05)	0.77 (0.07)	0.72 (0.07)
Mild versus severe	0.92 (0.04)	0.93 (0.03)	0.89 (0.04)	0.86 (0.04)

[a]BC-OS: BrainCheck standard battery of cognitive assessments overall score.

[b]ADL-TS: activities of daily living total score.

The fitted model coefficients are provided for reduced models 1 and 3 in Table 4. The 1-sided *P* value obtained from bootstrapping for each coefficient is shown in the parentheses.

XSL•FO
RenderX

Table . Coefficients (*P* values) of the fitted reduced models 1 (*BC-OS*[a] + *ADL-TS*[b]) and model 3 (*BC-OS* only).

	Regression coefficients, β (*P* value)				
	Reduced model 1			Reduced model 3	
Classification	β0 (Intercept)	β1 (stBC)	β2(stADL)	β0 (Intercept)	β1 (stBC)
Mild versus moderate	−0.071 (.01)	−0.757 (<.001)	−0.717 (<.001)	−0.053 (.002)	−0.890 (<.001)
Moderate versus severe	−0.819 (<.001)	−0.989 (.03)	−1.059 (<.001)	−0.453 (<.001)	−1.215 (<.001)
Mild versus severe	−2.443 (<.001)	−1.611 (<.001)	−1.517 (<.001)	−1.557 (<.001)	−2.076 (<.001)

[a]BC-OS: BrainCheck standard battery of cognitive assessments overall score.

[b]ADL-TS: activities of daily living total score.

Discussion

By conducting a retrospective analysis of patient data in real-world clinical settings, this study sought to investigate the relationship between cognitive performance in a battery of digital cognitive assessments, BC-Assess, and dementia severity measured by the DSRS. We found a statistically significant moderate correlation between the BC-OS and the DSRS-TS. This correlation is comparable with that between the ADL-TS and the DSRS-TS. Both BC-OS and ADL-TS significantly decrease with increasing severity. BC-Assess demonstrated to be a good predictor of dementia severity, with ROC-AUC of classification using logistic regression ranging from 0.733 to 0.917. When BC-Assess was combined with ADL, higher prediction accuracies were achieved, with ROC-AUC ranging from 0.786 to 0.961.

Our results suggest that BC-Assess could serve as an alternative tool to DSRS for grading dementia severity, particularly in cases where DSRS, or other global assessments, may be challenging to obtain due to logistical and time constraints. Unlike DSRS, BC-Assess, as a brief digital cognitive assessment, offers the advantage of flexible choice of self-administration or administration by clinical support staff, runs on common consumer technology, and does not require availability of an informant. The significant improvement of prediction accuracies when BC-Assess is combined with ADL indicates the synergetic relationship between cognitive and functional measures in grading dementia severity. Previous studies have shown that patients' loss of independence to manage activities of daily living is nonlinearly related with their cognitive decline [37], and their correlation weakens as the disease progresses [38]. This is consistent with the finding of relatively low correlation (r=0.37) between the 2 measures in this study and elsewhere [39,40]. Together, these findings suggest that ADL carries additional information of functional abilities that is partially independent of cognitive abilities measured by BC-Assess. When combined, the 2 measures provide a more comprehensive understanding of a patient's condition.

While the BC-OS and the ADL-TS from the mild and severe groups separate well from each other, scores from the moderate group vary widely among patients and largely overlap with both the mild and severe groups. This is reflected in high sensitivity and specificity (.86 or higher) for the classification between the mild and severe groups, and moderate sensitivity and specificity (0.83 or lower) for the classifications of the moderate group. Overall, however, a gradual change in the distribution of each score across stages can be observed. In line with current knowledge about the progression and stages of dementia [1,41], this pattern of results suggests that cognitive and functional declines do not happen in the same way across patients with dementia, and that there might only be subtle differences in cognitive or functional performance, or both, across patients within 2 successive stages.

Implicit in this study is the assumption that the staging of dementia severity based on the DSRS-TS had accurately identified each patient's underlying degree of severity. Previous studies demonstrated that the DSRS has high test-retest and inter-rater reliability [25] and good concurrent validity with high correlations with the CDR, the Mini-Mental Status Examination, and the Consortium to Establish a Registry for Alzheimer's Disease [23,25]. Other studies showed that the DSRS-TS can effectively differentiate between individuals with dementia, MCI, and healthy controls [42], and that it changes at a constant linear rate throughout the entire clinical course of dementia [25]. However, the psychometric properties of the DSRS in distinguishing between patients with mild, moderate, and severe dementia have not been studied. The DSRS-TS cut-offs used for staging of dementia severity have been suggested based on the mapping of the DSRS-TS onto CDR stages where a CDR global score of 0, 0.5, 1, 2, and 3 represent no, questionable, mild, moderate, and severe dementia. However, it has been shown that there is a large variability in the degree of severity among patients placed in a particular CDR stage, and patients with the same degree of severity can be placed in different CDR stages [2,43,44]. Furthermore, the precision of severity grading depends on the scoring approach to the CDR, ie, the item-response-theory approach is more precise than the sum-of-the-boxes approach, which is more precise than the global score approach [43]. On top of that, the mapping of DSRS-TS to CDR global score was based on a limited sample of patients with dementia that might not be representative of patients at different severity stages in general [23]. These limitations are possible contributing factors to the widespread distributions and overlaps of BC-OS and ADL-TS across patient groups delineated by the DSRS-TS in this study. To allow for more systematic investigations into the effectiveness of using these scores in dementia staging, future research needs to address these limitations and establish more fine-grained and well-defined staging criteria as well as optimize

XSL•FO
RenderX

the scoring methods for the DSRS, the CDR, and other assessments of dementia severity.

Suboptimal and inconsistent data quality in a dataset acquired outside typical clinical research settings is another factor that potentially causes large variabilities of scores observed in this study [45]. While real-world data may better represent diverse clinical environments, which is desired to obtain a generalized relationship between assessments, limited control over the data collection process and differences in clinical practices may result in reduced accuracy and consistency of the data. Inconsistencies exist across clinical sites and across units within each site due to differences in internal policies, staff training, workflow, and expertise. Inconsistencies also come from the different patient or caregiver populations each site serves. Patients may also have different dementia pathologies, leading to significant differences in the pattern of scores.

The ADL and BC-Assess also have their own limitations that one should take into consideration when interpreting the current results. The ADL measures 6 basic activities of daily living and employs dichotomous scoring, which allows only 2 possible scores for each functional category, ie, 1 for *independent* and 0 for *dependent*. Therefore, it lacks the resolution to capture intermediate levels of dependence. Furthermore, as it is subjective ratings from informants, for cases with small changes in functional activities, patients may end up receiving substantially different scores depending on their caregivers' judgments, resulting in low interrater variability [46,47]. In this study, with only patients with dementia included (based on their DSRS-TS), we found high variability in the ADL-TS across patients within each group, especially for the moderate group. As for the BC-Assess, besides measurement errors that exist in any assessment and technical difficulties older adults may have when using smart devices for the assessment, it might have limited utility in severity staging because patients with extremely severe conditions might not be able to complete it [48], and it could also suffer from a floor effect, a common limitation of psychometric tests [49].

Our data show a high imbalance in the number of patients across the 3 groups, with mild, moderate, and severe dementia accounting for 56%, 37.5%, and 6.5%, respectively. Although the trend is similar, higher proportions of patients estimated to be in the moderate (31%) and severe (21%) stages of Alzheimer's Disease were reported in a previous study [50]. As patients included in this study were those that received cognitive care planning services from their providers through the BrainCheck Plan platform within a 2-year period, our data do not necessarily reflect the prevalence of each stage. The fact that we only included patients with DSRS, ADL, and BC-Assess data also limited the number of patients in later stages who might be too impaired to take the BC-Assess. Furthermore, given its main goals are to help patients and their families better understand the patients' condition and needs, to offer strategies to improve their overall quality of life, and to plan for the future when their condition gets worse, cognitive care planning is more meaningful for patients in early stages. Patients in the severe stage of dementia are completely dependent on their family or caregivers, and many of them require specialized care and attention in facilities. These institutionalized patients tend to have been diagnosed and given care plans tailored to their specific needs by the institution.

In addition to the findings of this study, the growing field of ecological digital assessment tools offers valuable insights into monitoring and predicting dementia progression using digital biomarkers in real-world settings. For example, Buegler et al [51] show how these tools provide individualized, context-sensitive data to better understand cognitive performance. Integrating ecological tools with measures like BC-Assess could enhance its utility by capturing real-time data and offering a more comprehensive view of a patient's condition. These tools could also address challenges in tracking cognitive and functional abilities over time, particularly when in-person assessments or informants are unavailable. Further research into combining ecological tools with BC-Assess could refine dementia severity assessments and improve patient outcomes across stages of the disease.

Despite the limitations, this study shows that BC-Assess could be a promising solution for measuring dementia severity. The use of BC-Assess for this purpose will be particularly useful in primary care settings, where DSRS or other comprehensive global assessments may pose implementation challenges. Due to its flexibility, efficiency, and ease of use, BC-Assess can help streamline the assessment process, supporting timely diagnosis and management of dementia. This, in turn, can improve patient outcomes and ease the burden on caregivers.

Acknowledgments

Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health (award R44AG078006). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Authors' Contributions

DH and BH were responsible for the conceptualization and methodology of the study. DH was responsible for the formal analysis and wrote the original draft of the manuscript. All authors reviewed and edited the manuscript.

Conflicts of Interest

All authors report receiving salaries and stock options from BrainCheck.

XSL•FO
RenderX

Multimedia Appendix 1
Raw score (RS) metric and transformed score (TS) calculation for each assessment.
[DOCX File, 7 KB - aging_v8i1e65292_app1.docx]

References

1. Stages of Alzheimer's. Alzheimer's Association. URL: https://alz.org/alzheimers-dementia/stages [accessed 2025-02-18]
2. O'Bryant SE, Waring SC, Cullum CM, et al. Staging dementia using Clinical Dementia Rating Scale Sum of Boxes scores: a Texas Alzheimer's research consortium study. Arch Neurol 2008 Aug;65(8):1091-1095. [doi: 10.1001/archneur.65.8.1091] [Medline: 18695059]
3. Horowitz FD. Developmental theory, prediction, and the developmental equation in follow-up research. In: Developmental Follow-Up: Elsevier; 1994:27-44. [doi: 10.1016/B978-0-12-267855-4.50006-4]
4. Perneczky R, Wagenpfeil S, Komossa K, Grimmer T, Diehl J, Kurz A. Mapping scores onto stages: mini-mental state examination and clinical dementia rating. Am J Geriatr Psychiatry 2006 Feb;14(2):139-144. [doi: 10.1097/01.JGP.0000192478.82189.a8] [Medline: 16473978]
5. Doody RS, Stevens JC, Beck C. Practice parameter: management of dementia (an evidence-based review). Report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology 2001 May 8;56(9):1154-1166. [doi: 10.1212/WNL.56.9.1154] [Medline: 11342679]
6. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
7. Vertesi A, Lever JA, Molloy DW, et al. Standardized Mini-Mental State Examination. Use and interpretation. Can Fam Physician 2001 Oct;47:2018-2023. [Medline: 11723596]
8. Lezak MD, Howieson DB, Bigler ED, Tranel D. Neuropsychological Assessment, 5th edition: Oxford University Press; 2012:1161.
9. Crawford JR. A Handbook of Neuropsychological Assessment: Erlbaum; 1994:454.
10. Wild K, Howieson D, Webbe F, Seelye A, Kaye J. Status of computerized cognitive testing in aging: a systematic review. Alzheimers Dement 2008 Nov;4(6):428-437. [doi: 10.1016/j.jalz.2008.07.003] [Medline: 19012868]
11. Chan JYC, Yau STY, Kwok TCY, Tsoi KKF. Diagnostic performance of digital cognitive tests for the identification of MCI and dementia: a systematic review. Ageing Res Rev 2021 Dec;72:101506. [doi: 10.1016/j.arr.2021.101506] [Medline: 34744026]
12. Ye S, Sun K, Huynh D, et al. A computerized cognitive test battery for detection of dementia and mild cognitive impairment: instrument validation study. JMIR Aging 2022 Apr 15;5(2):e36825. [doi: 10.2196/36825] [Medline: 35436212]
13. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged. The index of ADL: a standardized measure of biological and psychosocial function. J Am Med Assoc 1963 Sep 21;185:914-919. [doi: 10.1001/jama.1963.03060120024016] [Medline: 14044222]
14. Pashmdarfard M, Azad A. Assessment tools to evaluate Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL) in older adults: a systematic review. Med J Islam Repub Iran 2020;34:33. [doi: 10.34171/mjiri.34.33] [Medline: 32617272]
15. Reisberg B, Ferris SH, de Leon MJ, Crook T. The Global Deterioration Scale for assessment of primary degenerative dementia. Am J Psychiatry 1982 Sep;139(9):1136-1139. [doi: 10.1176/ajp.139.9.1136] [Medline: 7114305]
16. Sclan SG, Reisberg B. Functional assessment staging (FAST) in Alzheimer's disease: reliability, validity, and ordinality. Int Psychogeriatr 1992;4 Suppl 1(3):55-69. [doi: 10.1017/s1041610292001157] [Medline: 1504288]
17. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology (ECronicon) 1993 Nov;43(11):2412-2414. [doi: 10.1212/wnl.43.11.2412-a] [Medline: 8232972]
18. Morris JC. Clinical dementia rating: a reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. Int Psychogeriatr 1997;9 Suppl 1(S1):173-176. [doi: 10.1017/s1041610297004870] [Medline: 9447441]
19. Fillenbaum GG, Peterson B, Morris JC. Estimating the validity of the clinical Dementia Rating Scale: the CERAD experience. Consortium to establish a Registry for Alzheimer's Disease. Aging (Milano) 1996 Dec;8(6):379-385. [doi: 10.1007/BF03339599] [Medline: 9061124]
20. Hughes CP, Berg L, Danziger WL, Coben LA, Martin RL. A new clinical scale for the staging of dementia. Br J Psychiatry 1982 Jun;140(6):566-572. [doi: 10.1192/bjp.140.6.566]
21. Rikkert MGMO, Tona KD, Janssen L, et al. Validity, reliability, and feasibility of clinical staging scales in dementia: a systematic review. Am J Alzheimers Dis Other Demen 2011 Aug;26(5):357-365. [doi: 10.1177/1533317511418954] [Medline: 21914671]
22. Duara R, Loewenstein DA, Greig-Custo MT, et al. Diagnosis and staging of mild cognitive impairment, using a modification of the clinical dementia rating scale: the mCDR. Int J Geriatr Psychiatry 2010 Mar;25(3):282-289. [doi: 10.1002/gps.2334] [Medline: 19565573]
23. Clark CM, Ewbank DC. Performance of the dementia severity rating scale: a caregiver questionnaire for rating severity in Alzheimer disease. Alzheimer Dis Assoc Disord 1996;10(1):31-39. [Medline: 8919494]

XSL•FO
RenderX

24. Cognitive impairment care planning toolkit. Alzheimer's Association. URL: https://www.alz.org/getmedia/d548ca27-309c-43b5-a789-5de837401452/cognitive-impairment-care-planning-toolkit_012623-(1)-(1).pdf [accessed 2025-02-18]
25. Xie SX, Ewbank DC, Chittams J, Karlawish JHT, Arnold SE, Clark CM. Rate of decline in Alzheimer disease measured by a Dementia Severity Rating Scale. Alzheimer Dis Assoc Disord 2009;23(3):268-274. [doi: 10.1097/WAD.0b013e318194a324] [Medline: 19812470]
26. Weinstein JD, Gonzalez ER, Egleton RD, Hunt DA. A paradigm shift for evaluating pharmacotherapy for Alzheimer's disease: the 10-Patient Screening Protocol. Consult Pharm 2013 Jul 1;28(7):443-454. [doi: 10.4140/TCP.n.2013.443]
27. McCulloch K, Collins RL, Maestas KL, LeMaire AW, Pacheco V. Validation and diagnostic utility of the dementia rating scale in a mixed dementia population. Alzheimer Dis Assoc Disord 2014;28(2):168-174. [doi: 10.1097/WAD.0000000000000014] [Medline: 24231456]
28. Reisberg B, Ferris SH, Torossian C, Kluger A, Monteiro I. Pharmacologic treatment of Alzheimer's disease: a methodologic critique based upon current knowledge of symptomatology and relevance for drug trials. Int Psychogeriatr 1992;4 Suppl 1:9-42. [Medline: 1504291]
29. Roalf DR, Moberg PJ, Xie SX, Wolk DA, Moelter ST, Arnold SE. Comparative accuracies of two common screening instruments for classification of Alzheimer's disease, mild cognitive impairment, and healthy aging. Alzheimers Dement 2013 Sep;9(5):529-537. [doi: 10.1016/j.jalz.2012.10.001] [Medline: 23260866]
30. de Craen AJM, Heeren TJ, Gussekloo J. Accuracy of the 15-item geriatric depression scale (GDS-15) in a community sample of the oldest old. Int J Geriatr Psychiatry 2003 Jan;18(1):63-66. [doi: 10.1002/gps.773] [Medline: 12497557]
31. Steffens DC, Otey E, Alexopoulos GS, et al. Perspectives on depression, mild cognitive impairment, and cognitive decline. Arch Gen Psychiatry 2006 Feb;63(2):130-138. [doi: 10.1001/archpsyc.63.2.130] [Medline: 16461855]
32. Sultzer DL, Levin HS, Mahler ME, High WM, Cummings JL. A comparison of psychiatric symptoms in vascular dementia and Alzheimer's disease. Am J Psychiatry 1993 Dec;150(12):1806-1812. [doi: 10.1176/ajp.150.12.1806] [Medline: 8238634]
33. Forsell Y, Winblad B. Major depression in a population of demented and nondemented older people: prevalence and correlates. J Am Geriatr Soc 1998 Jan;46(1):27-30. [doi: 10.1111/j.1532-5415.1998.tb01009.x] [Medline: 9434662]
34. Yang S, Flores B, Magal R, et al. Diagnostic accuracy of tablet-based software for the detection of concussion. Plos One 2017 Jul 7;12(7):e0179352. [doi: 10.1371/journal.pone.0179352]
35. Wallace M, Shelkey M, Hartford Institute for Geriatric Nursing. Katz Index of Independence in Activities of Daily Living (ADL). Urol Nurs 2007 Feb;27(1):93-94. [Medline: 17390935]
36. Brant R. Assessing proportionality in the proportional odds model for ordinal logistic regression. Biometrics 1990 Dec;46(4):1171-1178. [Medline: 2085632]
37. Zidan M, Arcoverde C, Araújo NBD, et al. Motor and functional changes in different stages of Alzheimer's disease. Arch Clin Psychiatry São Paulo 2012;39(5):161-165. [doi: 10.1590/S0101-60832012000500003]
38. Monaci L, Morris RG. Neuropsychological screening performance and the association with activities of daily living and instrumental activities of daily living in dementia: baseline and 18- to 24-month follow-up. Int J Geriatr Psychiatry 2012 Feb;27(2):197-204. [doi: 10.1002/gps.2709] [Medline: 22223145]
39. Louis ED. Functional correlates of lower cognitive test scores in essential tremor. Mov Disord 2010 Mar 15;25(4):481-485. [doi: 10.1002/mds.22920] [Medline: 20108374]
40. Lee G. Impaired cognitive function is associated with motor function and activities of daily living in mild to moderate Alzheimer's dementia. Curr Alzheimer Res 2020;17(7):680-686. [doi: 10.2174/1567205017666200818193916] [Medline: 32811414]
41. The progression, signs and stages of dementia. Alzheimer's Society. 2021. URL: https://www.alzheimers.org.uk/about-dementia/symptoms-and-diagnosis/how-dementia-progresses/progression-stages-dementia [accessed 2025-02-18]
42. Mitchell JC, Dick MB, Wood AE, Tapp AM, Ziegler R. The utility of the Dementia Severity Rating Scale in differentiating mild cognitive impairment and Alzheimer disease from controls. Alzheimer Dis Assoc Disord 2015;29(3):222-228. [doi: 10.1097/WAD.0000000000000057] [Medline: 25187220]
43. Balsis S, Miller TM, Benge JF, Doody RS. Dementia staging across three different methods. Dement Geriatr Cogn Disord 2011;31(5):328-333. [doi: 10.1159/000328116] [Medline: 21555887]
44. Chang YL, Bondi MW, McEvoy LK, et al. Global clinical dementia rating of 0.5 in MCI masks variability related to level of function. Neurology (ECronicon) 2011 Feb 15;76(7):652-659. [doi: 10.1212/WNL.0b013e31820ce6a5] [Medline: 21321338]
45. Bhatt A. Data quality - The foundation of real-world studies. Perspect Clin Res 2023;14(2):92-94. [doi: 10.4103/picr.picr_12_23] [Medline: 37325580]
46. Brorsson B, Asberg KH. Katz index of independence in ADL. Reliability and validity in short-term care. Scand J Rehabil Med 1984;16(3):125-132. [Medline: 6494836]
47. Arik G, Varan HD, Yavuz BB, et al. Validation of Katz index of independence in activities of daily living in Turkish older adults. Arch Gerontol Geriatr 2015;61(3):344-350. [doi: 10.1016/j.archger.2015.08.019] [Medline: 26328478]

XSL•FO
RenderX

48. Kiselica AM, Johnson E, Benge JF. How impaired is too impaired? Exploring futile neuropsychological test patterns as a function of dementia severity and cognitive screening scores. J Neuropsychol 2021 Sep;15(3):410-427. [doi: 10.1111/jnp.12243] [Medline: 33655681]
49. Auer SR, Sclan SG, Yaffee RA, Reisberg B. The neglected half of Alzheimer disease: cognitive and functional concomitants of severe dementia. J Am Geriatr Soc 1994 Dec;42(12):1266-1272. [doi: 10.1111/j.1532-5415.1994.tb06509.x]
50. Hebert LE, Scherr PA, Bienias JL, Bennett DA, Evans DA. Alzheimer disease in the US population: prevalence estimates using the 2000 census. Arch Neurol 2003 Aug;60(8):1119-1122. [doi: 10.1001/archneur.60.8.1119] [Medline: 12925369]
51. Buegler M, Harms R, Balasa M, et al. Digital biomarker-based individualized prognosis for people at risk of dementia. Alzheimers Dement (Amst) 2020;12(1):e12073. [doi: 10.1002/dad2.12073] [Medline: 32832589]

Abbreviations

ADL: activities of daily living
ADL-TS: activities of daily living total score
BC-Assess: BrainCheck standard battery of cognitive assessments
BC-OS: BrainCheck standard battery of cognitive assessments overall score
CDR: Clinical Dementia Rating
DSRS: Dementia Severity Rating Scale
DSRS-TS: Dementia Severity Rating Scale total score
MANOVA: multivariate analysis of variance
ROC: receiver operating characteristic
ROC-AUC: area under the receiver operating characteristic curve

Edited by PH Chau; submitted 13.08.24; peer-reviewed by A Vella, M Colombo; revised version received 19.12.24; accepted 25.12.24; published 26.02.25.

Please cite as:
Huynh D, Sun K, Patterson M, Hosseini Ghomi R, Huang B
Performance of a Digital Cognitive Assessment in Predicting Dementia Stages Delineated by the Dementia Severity Rating Scale: Retrospective Study
JMIR Aging 2025;8:e65292
URL: https://aging.jmir.org/2025/1/e65292
doi:10.2196/65292

© Duong Huynh, Kevin Sun, Mary Patterson, Reza Hosseini Ghomi, Bin Huang. Originally published in JMIR Aging (https://aging.jmir.org), 26.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Home-Based Exercise and Fall Prevention in Older Adults: Development, Validation and Usability of the Mais Equilíbrio Mobile App

Mateus Medeiros Leite[1,2*], MSc; Alessandro de Oliveira Silva[1*], PhD; Silvana Schwerz Funghetto[2,3*], PhD; Luciano Ramos de Lima[3*], PhD; Samuel Barbosa Mezavila Abdelmur[2*], Spec; Hudson Azevedo Pinheiro[4*], PhD; Calliandra Maria de Souza Silva[2*], PhD; Maurílio Tiradentes Dutra[5*], PhD; Marina Morato Stival[2,3*], PhD

[1]Physical Education Department, University Center of Brasilia - UniCEUB, University Center of Brasília, SEPN 707/907 - Asa Norte, Brasília - DF, 70790-075, Brasília, Brazil

[2]Graduate Program in Health Sciences and Technologies, Faculty of Ceilândia, Universidade de Brasília, Ceilândia Sul Campus Universitário - Centro Metropolitano, Brasília, Brazil

[3]Nursing Course, Faculty of Ceilândia, Universidade de Brasília, Brasília, Brazil

[4]Health Department of the Federal District, Governo do Distrito Federal, Brasília, Brazil

[5]Federal Institute of Education, Science and Technology of Brasília, Instituto Federal de Educação, Ciência e Tecnologia de Brasília, Brasília, Brazil

[*]all authors contributed equally

Corresponding Author:
Mateus Medeiros Leite, MSc
Physical Education Department, University Center of Brasilia - UniCEUB, University Center of Brasília, SEPN 707/907 - Asa Norte, Brasília - DF, 70790-075, Brasília, Brazil

Abstract

Background: The global aging population and the high incidence of falls among this population highlight the need for effective preventive strategies. Home-based exercise programs, such as the Otago protocol, have demonstrated efficacy in reducing fall risk but often face barriers related to user adherence. Mobile health (mHealth) apps offer promising tools to support health promotion and enhance autonomy in older adults.

Objective: This study aims to develop and validate a prototype mobile app, *Mais Equilíbrio* (More Balance), designed to guide older adults in performing home-based physical exercises adapted from the Otago protocol.

Methods: This methodological study was conducted in two phases: (1) content validation by 22 experts in physical education and physiotherapy using the Suitability Assessment of Materials (SAM) scale, and (2) usability testing with 24 older adults (aged 60 to 80 y), using the System Usability Scale (SUS). An overall score above 70% on the SAM and above 85 on the SUS were considered indicators of high quality and excellent usability, respectively.

Results: The *Mais Equilíbrio* (More Balance) app was developed based on the Otago protocol and tailored for independent home use. A Content Validity Index above 0.95 was observed for all items. An overall average score of 81.20 (SD 15.78) on the SAM scale was found, classifying the material as “superior.” Usability tests with older adults showed an average score of 95.98 (SD 5.58) on the SUS, indicating excellent usability. The highest scores were observed in “ease of use” and “user confidence.”

Conclusions: The *Mais Equilíbrio* (More Balance) app, distinct for digitally adapting the Otago protocol to the Brazilian context and for its dual validation process with experts and older adults, has proven to be a valid and highly usable tool for guiding home-based physical exercise in older adults, with potential to promote fall prevention and autonomy.

(JMIR Aging 2025;8:e80724) doi:10.2196/80724

KEYWORDS

mobile apps; elderly; physical exercise; fall prevention; mHealth interventions; older adults

Introduction

Population aging has accelerated in recent decades, driven by social, health, and technological advances that have increased life expectancy while reducing mortality and fertility rates [1]. Estimates indicate that this trend will persist, with significant growth in the older adult population expected over the coming decades, especially in high-income countries [2,3].

This scenario is accompanied by a higher prevalence of chronic diseases and aging-related conditions, such as falls. Approximately 28% - 35% of older adults experience at least one fall per year, and 15% report recurrent episodes. Falls

XSL•FO
RenderX

contribute to nearly 300,000 deaths worldwide and are associated with severe consequences [4-6], including fractures, functional decline, fear of falling, and postfall syndrome, which often lead to loss of independence, reduced life expectancy, and increased morbidity and mortality [7-11].

Falls among older adults are multifactorial events influenced by intrinsic, extrinsic, and situational factors, including biological, behavioral, environmental, socioeconomic, psychosocial, and demographic aspects [7,12,13]. Specific risk factors such as advanced age, female sex, polypharmacy, history of falls, low cognitive performance [14], reduced physical function [15], decreased muscle strength (MS), particularly in the lower limbs [16], fear of falling, balance impairments, and low physical activity levels [17-21] further increase the likelihood of falls [22].

Low physical activity levels negatively affect older adults' ability to perform activities of daily living and are associated with adverse health outcomes and higher mortality [23,24]. Consequently, interventions targeting MS and balance are recommended for fall prevention. Programs such as the Otago Exercise Program (OEP) and the Home-based Older People's Exercise have demonstrated effectiveness, including in home-based implementations [25-27].

In this sense, mobile health (mHealth) technologies, an important component of gerontechnology, offer promising tools to support self-care, guide physical exercise practices, and assist in the monitoring of chronic health conditions [28-30]. In older adults, previous studies have demonstrated the use of mHealth solutions, especially with interventions that address health promotion, disease prevention, and long-term maintenance of regular exercise. Thus, these technologies stand out for their ability to be used independently and within home environments. However, the main barriers associated with this type of technology for older adults are declining motor and cognitive functioning, as well as low confidence in digital competence [31-34].

The development of the *Mais Equilíbrio* (More Balance) app was based on the Systematic Instructional Design model, which provides a structured framework for designing health-related apps [35]. To address the specific needs of the older population, the intervention design incorporated principles of universal usability and inclusive design. The choice of the OEP protocol as the scientific basis was motivated by its proven effectiveness in preventing falls, with risk reductions in older adults [4,36]. The central hypothesis is that technological mediation can overcome traditional barriers to adherence to home-based exercises, such as forgetfulness, low motivation, and lack of professional supervision.

Although several apps for fall prevention or risk assessment are already available, such as *FallSA*, designed as a self-assessment tool for community-dwelling older adults [37], the *Mais Equilíbrio* (More Balance) is distinctive in digitally integrating the OEP, adapted to the Brazilian context, with a strong focus on adherence and usability among older adults. Importantly, it also underwent a dual validation process, engaging both health experts and end users, which strengthens its contextual adequacy and reliability for future interventions. By converting the OEP into a mobile interface, *Mais Equilíbrio* (More Balance) combines the well-documented benefits of OEP, improvements in strength, balance, and fall reduction with the added advantages of digital health solutions, such as accessibility, remote monitoring, and scalability, benefits also highlighted in recent reviews of digitally delivered OEP interventions [38].

Despite the growing availability of mHealth solutions, there is a gap in the literature regarding the use of apps designed to promote physical exercise in older adults for improving balance, MS, and fall prevention. Most existing apps offer only educational content in text or images, with little interactivity, no self-monitoring, and no involvement of health professionals in the creation process. They also do not declare the scientific source of the content, which highlights the need for randomized controlled trials to assess their effectiveness. Given technological advancements, it is important to offer reliable material that uses appropriate language and attractive and dynamic features, and is easily accessible, facilitating understanding and encouraging home-based exercise routines [39].

Thus, this study aimed to develop and validate a prototype mobile app to promote home-based physical exercise in older adults, featuring instructional videos adapted from the Otago protocol.

Methods

Ethical Considerations

This study was conducted in accordance with national and international ethical guidelines and was approved by the Research Ethics Committee of the Faculty of Ceilândia, University of Brasília (Research Ethics Committee – CEP/FCE-UnB; approval number 5.530.239, approved on July 1, 2022). All procedures followed the guidelines of the National Research Council in accordance with current Brazilian legislation (Resolution No. 466/2012) and the principles of the Declaration of Helsinki. All participants provided informed consent prior to participation. The Free and Informed Consent Form (FICF) was obtained electronically through an online platform before data collection began. Participants were informed about the voluntary nature of the study and their right to withdraw at any time without penalty or loss of benefits. No financial or material compensation was provided to participants, in accordance with Brazilian ethical regulations for research involving human subjects. This research is part of the project entitled "Development and evaluation of home exercise technology for older adults".

Study Design and Period

This is a methodological development study with a quantitative and descriptive approach, conducted from March 2023 to September 2024. The study followed the guidelines of the Revised Standards for Quality Improvement Reporting Excellence (SQUIRE 2.0) instrument.

Implementation Scenario and Context

The development and validation of the *Mais Equilíbrio* (More Balance) app were carried out in the Brazilian health care

XSL•FO
RenderX

system, which, despite a robust primary health care (PHC) model with extensive coverage, has limitations and still faces challenges in terms of individualized care to older adults. A strategic partnership with a basic health unit facilitated participant recruitment, optimizing access to the community. In this context, the following factors were considered crucial for the implementation of digital technologies: (1) availability of home internet infrastructure, (2) family support in the use of technology, (3) the socioeconomic heterogeneity of the population, and (4) public policies aimed at promoting the health of older adults.

Technological Production: Prototype Development

To design the prototype of the *Mais Equilíbrio* (More Balance) mobile app, a structured workflow was established, with regular meetings between the development team and researchers to define the app's interfaces and screens. The purpose of the app is to present video demonstrations of the adapted Otago exercise protocol, along with its progressive stages, as detailed in Table 1.

Table . Otago exercise protocol adapted for the *Mais Equilíbrio* (More Balance) app.[a]

	Warm-up	Strength	Balance	Walk
Activities	1. Neck movements 2. Trunk rotation 3. Trunk extension 4. Ankle movements 5. March	1. Knee extension 2. Knee flexion 3. Hip abduction 4. Plantar flexion/Dorsi-flexion	1. Squat 2. Sitting down and standing up 3. Tandem balance 4. Balance 1 leg 5. Walking underfoot 6. Walking backwards 7. Walk to the side 8. Tandem walk 9. Walking and turning 10. Stair walking 11. Walking on tiptoes	Encouraged
Level				
1 (2x week)	1× 10 repetitions	2× 10 to 12 repetitions	2× 10 to 12 repetitions	2x week 30 min
2 (3x week)	1× 10 repetitions	3× 10 to 12 repetitions	3× 10 to 12 repetitions	2x week 30 min
3 (3x week)	1× 10 repetitions	4× 10 to 12 repetitions	4× 10 to 12 repetitions	2x week 30 min

[a]Adapted from Campbell and Robertson [40].

The app was developed using a cross-platform framework to ensure compatibility across devices, with a backend designed to provide a secure and scalable architecture for data management. An administration panel was also implemented to allow efficient monitoring and management of the app. This combination of technologies enabled the creation of a responsive, user-friendly, and maintainable system suitable for long-term use [41].

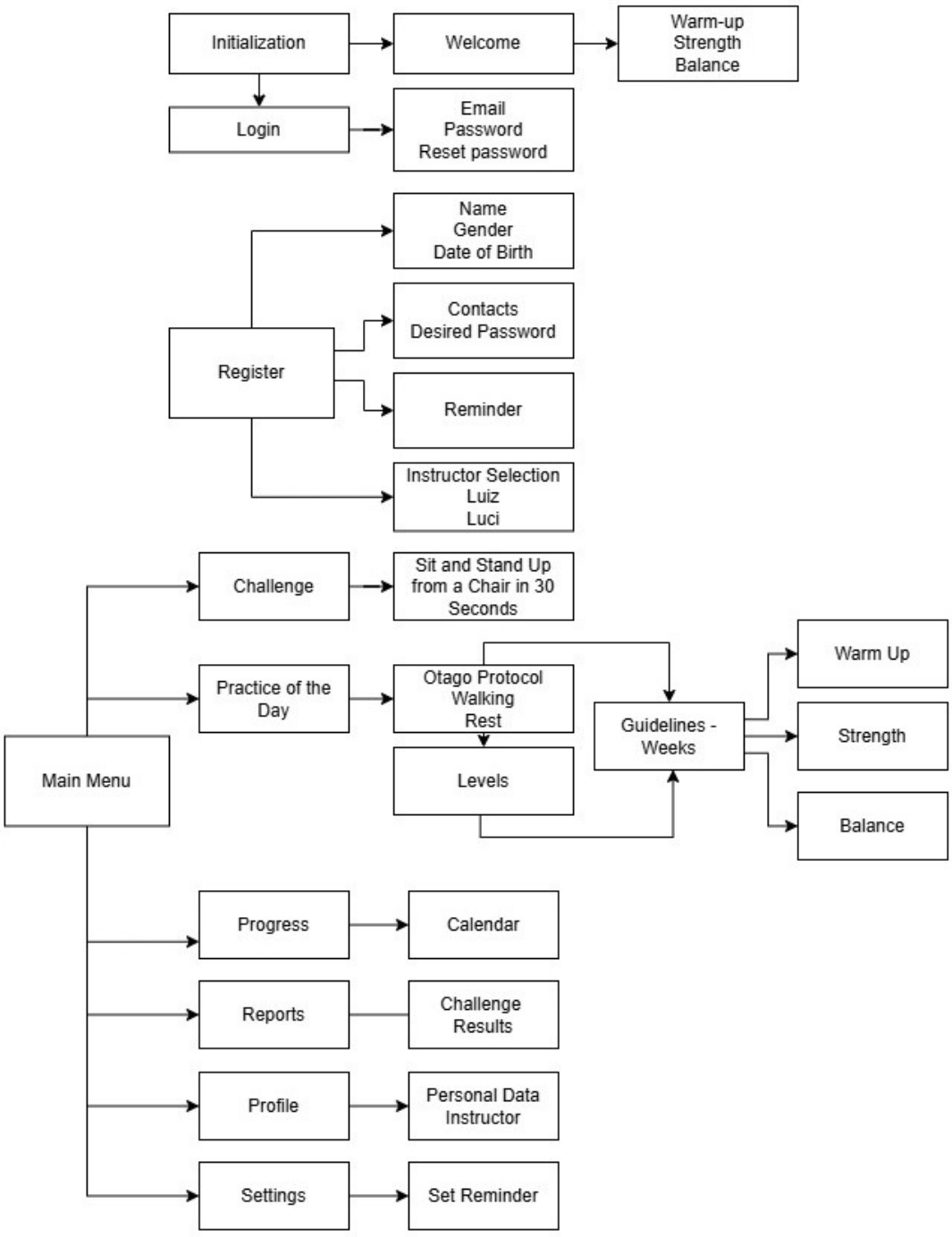
The software was developed following the steps of Systematic Design of Instruction, a structured approach that systematizes the creation of mobile apps in the health care field. This model includes four phases: analysis, design or development, implementation, and evaluation [35]. In this stage, the analysis and design or development phases were carried out.

During the analysis phase, the software objectives were established by professionals with expertise in physical exercise for older adults, together with a digital interface designer and a programmer. The instructional content was developed using videos adapted according to the OEP protocol [26].

During the design or development phase, the researchers created videos tailored to language to enhance understanding among older adult users. The app's layout composition was defined, ie, the visual organization and interaction models. For this activity, a digital interface designer and a programmer were consulted, both funded by the project. They developed the prototype for both Android and iOS operating systems, enabling future availability on the Play Store and Apple Store, thereby increasing the app's reach and accessibility. The system framework was built in subdivisions of sections and subsections, as shown in Figure 1.

XSL•FO
RenderX

Figure 1. Framework of the *Mais Equilíbrio* (More Balance) app.



Validation With Experts

A total of 310 experts were invited through different strategies, including dissemination on the Brazilian Lattes Platform of the National Council for Scientific and Technological Development, on institutional websites of major Brazilian universities, and through referrals from other experts who had already participated. The invitation letter was sent by email, along with the research objectives and a link to download the app, as well as access to the informed consent form and the app validation tool in Google Forms format. This instrument was structured into two sections: (1) characterization by experts and (2) Suitability Assessment of Materials (SAM) scale. The experts were selected using the method proposed by Fehring [42], with adaptations. A period of 30 days was granted for the completion of the validation, and, if it was not completed, a new email was sent every 7 days to reinforce the invitation.

XSL•FO
RenderX

The inclusion criteria for this group were as follows: being a physical education or physiotherapy professional; achieving a score ≥5 in Fehring's adapted criteria (master's degree: 4 points; master's degree with dissertation in the area of interest: 1 point; research in the area of interest: 2 points; article published in the area of interest: 2 points; doctorate degree with thesis in the area of interest: 2 points; professional experience in the area of interest: 2 points; and specialization in the area of interest: 2 points); and agreeing to participate by fully completing all study-related forms. Experts were excluded from the study if they did not submit their assessment on time or if they only partially participated in the study. Ultimately, 22 experts agreed to participate in the research, meeting the previous recommendations [43].

For content validation, the SAM scale was organized on an ordinal scale that considered the following items: 0, inadequate; 1, adequate; and 2, excellent [44]. When selecting option 0, experts were asked to provide justification or suggestions for improvement. To calculate the overall SAM score, the arithmetic mean of the sums for each expert (Σ) was calculated, divided by the number of experts and multiplied by 100. This yielded the following percentages: 70% - 100% (superior material), 40% - 69% (adequate material), or 0% - 39% (inadequate material). A minimum score of 70% was required for the app to be considered approved [45].

Validation With Target Audience

In this stage, older people aged 60 years or older (recognized as elderly in Brazil) who had access to smartphones with Android or iOS systems participated over the course of a week. The invitation to participate was made at a basic health unit providing PHC. Initially, 50 older adults were recruited. Of these, 24 were included, according to the inclusion criteria: age between 60 and 80 years; both sexes; literate; owning a smartphone that supports the app with internet access; walk without the aid of devices, as verified by self-report during screening. The usability of the app was evaluated in person using the System Usability Scale (SUS) [46,47].

To calculate usability by SUS, 1 was subtracted from the score of odd items, and for even items, the value of the 5-point scale was subtracted. The sum of the results was multiplied by 2.5, generating a final score between 0 and 100. Scores between 50 and 67 are considered borderline; 68 - 84 indicate good usability; and above 85, excellent acceptance [48-50]. The assessment was carried out independently by two researchers, with joint review in case of discrepancies until consensus was reached. In addition to the overall score, the SUS items allow for the evaluation of aspects such as memorization (item 2), learning (items 3, 4, 7, and 10), efficiency (items 5, 6, and 8), satisfaction (items 1, 4, and 9), and inconsistencies (item 6) [51,52].

To ensure completeness and accuracy of the data collected, several control measures were employed. The evaluation forms (SAM, Content Validity Index [CVI], and SUS) were reviewed immediately after completion to ensure that all questions had been answered. For the usability validation with older adults, qualitative feedback was recorded through detailed notes to complement quantitative data and to ensure the reliability of the information. Quantitative data were transcribed and subjected to a double-checking stage by an independent researcher to identify and correct possible typing errors or inconsistencies. In addition, all instruments were applied in standardized environments, and researchers were previously trained to ensure uniformity in data collection procedures.

Statistical Analysis

Data obtained were extracted from Google Forms to a Microsoft Excel (version 240) spreadsheet. IBM SPSS (Statistical Product and Service Solutions) software, version 25.0, was used for statistical analysis. The characteristics of the sample were categorized and presented by frequency and percentages. For content and appearance validation, the CVI was calculated to measure the experts' agreement on the validity of the educational technology content [53-55]. To determine the level of agreement among experts, CVI scores were calculated using the number of experts who chose response options 1 or 2 for an item divided by the total number of experts who evaluated the item. The average CVI values for all messages and the content validity of individual items were calculated. A CVI greater than or equal to 0.78 was considered to indicate content validity [56,57].

The exact binomial distribution test indicated for small samples was performed, with a statistical significance level of P value <.05 and a concordance ratio of 0.95, to estimate the statistical reliability of the CVI. The internal consistency of the usability instrument (SUS) was verified by Cronbach α. The analysis of qualitative data, which included comments and suggestions from both experts in Stage 1 and participants in Stage 2, was performed using thematic content analysis. This approach allowed for an in-depth understanding of the aspects perceived as positive and the areas that required adjustments in the app, providing valuable insights into content and usability. Due to the cross-sectional design of this study, the evaluation of temporal effects on the variables was not applicable.

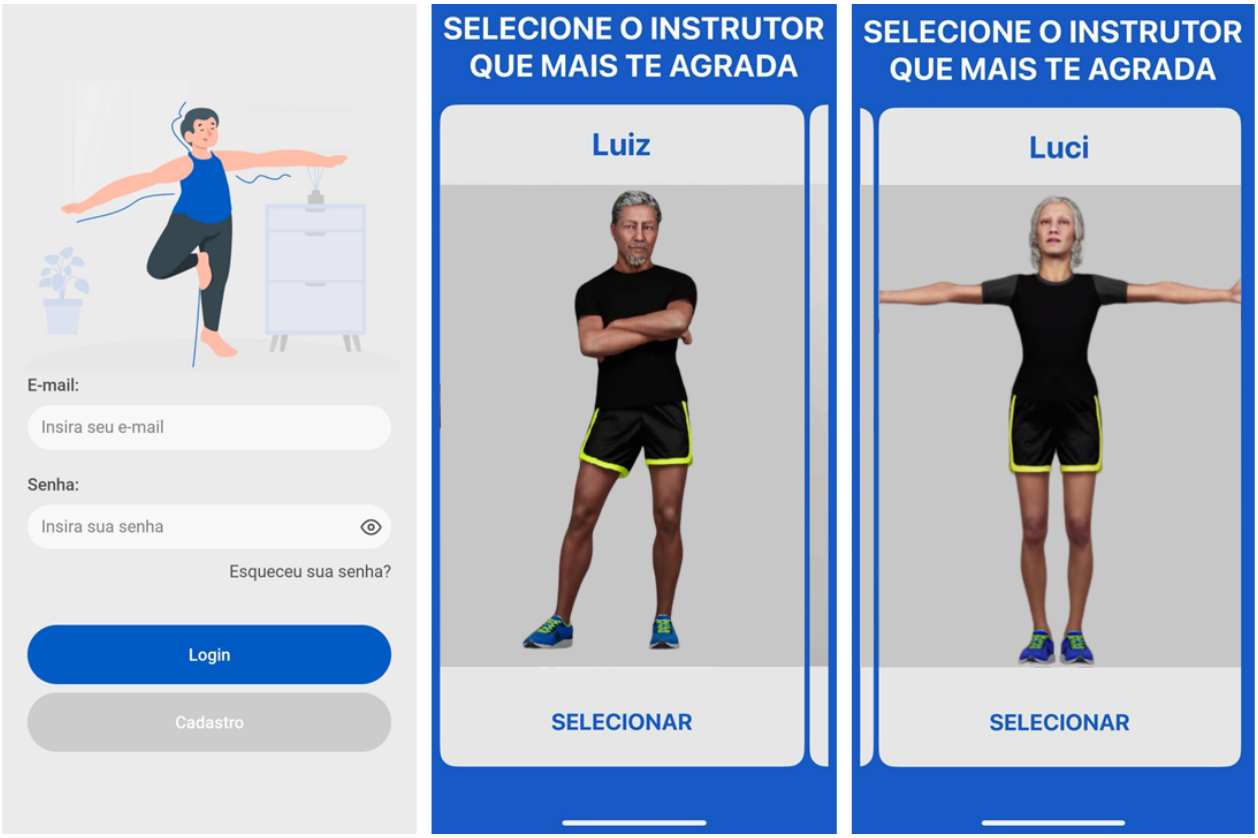
Results

The development of the *Mais Equilíbrio* (More Balance) app was an iterative process, led by a multidisciplinary research team composed of specialists and professionals in care for older people, aligned with Systematic Instructional Design. This process resulted in several modifications, such as adjustments to the interface, simplification of commands, and inclusion of audio reminders, based on internal feedback from the team.

Figure 2 presents an overview of the app's interface and features through a collection of main screens. This visual representation includes everything from the authentication process to the monitoring and customization features, providing a complete overview of the user experience. To ensure fidelity to the study validation, the screens are presented in their original language, Portuguese, with specific details for each one provided in the figure caption.

XSL•FO
RenderX

Figure 2. Tabs of the *Mais Equilíbrio* (More Balance) app. The figure shows three screens from the "*Mais Equilíbrio* (More Balance)" app. The first screen on the left displays a login interface, with fields labeled "E-mail" ("Enter your e-mail"), "Senha" ("password"), and buttons for "Login" and "*Cadastro*" ("Register"). The two screens on the right present the user with the option to select an instructor, with the title "SELECIONE O INSTRUCTOR QUE MAIS TE AGRADA" ("SELECT THE INSTRUCTOR YOU LIKE THE MOST)." The two available instructors are Luiz and Luci, each depicted with their name above and a "*SELECIONAR*" ("SELECT") button below.



Among the 22 experts who participated in the validation process, most were female (n=12, 54.55%), aged between 40 and 49 years (n=11, 50%), and physical education professionals (n=15, 68.18%). Regarding their qualifications, all held a master's degree (n=22, 100%), and most also held a doctorate (n=18, 81.82%). In terms of scientific production related to digital technologies, fall prevention, balance, or MS in older adults, 54.55% (n=12) had a master's dissertation on one of these topics, 45.45% (n=10) had a doctoral thesis in the area, and 86.36% (n=19) had published articles and ongoing research on the subject.

The results of the experts' evaluation demonstrated the adequacy of the educational material, considering the criteria of content, language, presentation, methodology, and cultural relevance. All 22 items assessed obtained a CVI equal to or greater than 0.95. Twenty items achieved a CVI of 1.00, with total agreement among the evaluators. Three items obtained 95.45% agreement (CVI 0.96), remaining within the validation parameters (Table 2).

XSL•FO
RenderX

Table . Expert agreement on material suitability assessment items, Brasília, Federal District, Brazil, 2025 (n=22).

	Values	CVI[a,b]	*P* value[c]
Objective, n (%)		1.00	
The objective is clear, covers the proposed topic, and facilitates understanding of the material	22 (100)	1.00	.32
Content covers relevant information to meet the objective	22 (100)	1.00	.32
The material proposal is limited to the objectives, clarifying doubts about the addressed topic	22 (100)	1.00	.32
Content highlights key points and encourages behavioral change	22 (100)	1.00	.32
Language, n (%)		0.99	
Reading level is appropriate for targeted audience comprehension	22 (100)	1.00	.32
Language style facilitates understanding and allows for active engagement of the targeted audience	22 (100)	1.00	.32
Appropriate language using everyday words	21 (95.5)	0.96	.70
Context comes before new information	22 (100)	1.00	.32
Logical sequence of ideas and the learning process facilitated by items	22 (100)	1.00	.32
Graphic Illustrations, n (%)		1.00	
Design and graphic resources express the objective	22 (100)	1.00	.32
Types of illustrations and graphic resources	22 (100)	1.00	.32
Relevant graphic resources	22 (100)	1.00	.32
The graphic resources used are self-explanatory	22 (100)	1.00	.32
Resources have explanatory captions	22 (100)	1.00	.32
Layout and presentation, n (%)		0.97	
Layout appearance	22 (100)	1.00	.32
Dimension and quality of spelling	21 (95.5)	0.96	.70
Use of subtitles	21 (95.5)	0.96	.70
Motivation, n (%)		1.00	
There is interaction between the text or graphic resources and the targeted audience, leading them to make choices or demonstrate skills	22 (100)	1.00	.32
Desired behavior patterns that are well modeled or well demonstrated/exemplified	22 (100)	1.00	.32
There is motivation for self-efficacy, that is, people are motivated to learn because they believe that tasks and behaviors are feasible	22 (100)	1.00	.32
Cultural suitability		1.00	

XSL•FO
RenderX

	Values	CVI[a,b]	*P* value[c]
The material is culturally appropriate to the logic, language, and experience of the targeted audience, n (%)	22 (100)	1.00	.32
Features culturally appropriate images and examples, n (%)	22 (100)	1.00	.32
SAM[d] global score, mean (SD)	81.20 (15.78)	—[e]	—

[a]Item content validation index.

[b]CVI: Content Validity Index.

[c]Binomial test.

[d]SAM: Suitability Assessment of Materials.

[e]Not applicable.

The average SAM global score was 81.20 (SD 15.78) points, classifying the material as "superior" according to the instrument's criteria. Thus, the results indicate that the content of the app was considered adequate in terms of clarity, relevance, accessible language, visual organization, and adaptation to the targeted audience (Table 2).

The target audience (older people) participated in the usability validation stage but were not involved in the development interactions. During this validation stage with the older people, it was observed that contextual elements, such as family support and internet access, interacted with the user experience. No direct adverse consequences were identified. However, based on the reports of the validation participants, unexpected benefits emerged, such as increased self-confidence in the use of technology and improvement in exercise routines. Rigorous data collection ensured that no relevant data were missing.

In the usability validation by 24 older adults, the majority were female (n=16, 66.67%); aged between 60 and 69 (n=18, 75.00%); had completed high school, undergraduate, or postgraduate degrees (n=18, 75.00%); and exercised (n=18, 83.33%). Table 3 shows means and SDs of the SUS domain scores by the target audience. The overall mean scale score was 95.98 (SD 5.58) points. Among the evaluated domains, the domain *Ease of memorization* obtained an average of 100.00 points, followed by the domains *System efficiency* and *Inconsistencies*, both with an average of 96.88 (SD 3.13 and SD 15.31 points, respectively). The domain *Ease of system knowledge* had an average of 94.79 (SD 2.55). Finally, the domain *User satisfaction* had an average of 94.44 (SD 3.18) points. The individual questions on the scale had averages ranging from 91.67 to 100.00 points.

XSL•FO
RenderX

Table . Distribution of average scores for the domains of the System Usability Scale for the *Mais Equilíbrio* (More Balance) app (n=24).

Usability feature	Scores, mean (SD)	Meaning
Ease of use	94.79 (2.55)	Easy-to-use system when used for the first time
Q3[a]	97.92 (7.06)	
Q4	91.67 (19.03)	
Q7	94.79 (12.72)	
Q10	94.79 (20.82)	
System efficiency	96.88 (3.13)	Speed in executing established tasks
Q5	93.75 (21.17)	
Q6	96.88 (15.31)	
Q8	100.00 (0.00)	
Inconsistencies	96.88 (15.31)	Absence of errors
Q6	96.88 (15.31)	
Ease of memorization	100.00 (0.00)	Easy to use even after a long period of not using it
Q2	100.00 (0.00)	
User satisfaction	94.44 (3.18)	Pleasant design
Q1	93.75 (21.17)	
Q4	91.67 (19.03)	
Q9	97.92 (7.06)	
SUS[c] Global Score	95.98 (5.58)	—[b]

[a]Q: System Usability Scale question.

[b]Not applicable.

[c]SUS: System Usability Scale.

Discussion

Main Results

The *Mais Equilíbrio* (More Balance) app was developed to encourage and educate older adults to adopt self-care and fall prevention practices, thereby promoting health and quality of life. To this end, this study aimed to develop and validate this prototype mobile app for promoting home-based physical exercise, containing videos that demonstrate adapted exercises based on the Otago protocol.

Managing health in old age can be challenging. Older people are often less likely to trust their own ability to access health information and advice, especially online. This reluctance can diminish their interest in using digital health information and advice modalities [58]. To reverse this situation, it is important to optimize information processing and minimize cognitive load for effective health education. This means promoting access to the digital era and stimulating interest in online health information, especially through reliable sources and with health professionals' participation. These strategies can contribute to autonomy, self-care, and health management during aging [59,60].

In this context, mobile technologies combined with internet access have great potential to contribute to health promotion and disease prevention, enabling access to knowledge and the development of skills and responsibilities necessary for the adoption of healthier behaviors [33,61]. The *Mais Equilíbrio* (More Balance) app can provide accessibility and usability for older people, focusing on simplicity, clarity, and the use of objective, audience-adapted language.

To facilitate use, buttons have been enlarged and strategically positioned. Content has been organized into blocks with intuitive navigation and voice assistance for the actions to be performed. To make the app more inclusive, visual elements considered common age-related visual limitations, employing large fonts, high-contrast colors, simple and recognizable icons, and clear, uniform illustrations. The content was built based on recommendations from guidelines and technical materials on elderly health, ensuring relevance and quality, prioritizing comprehensibility and usefulness [55,62-65].

It is believed that the user-centered approach was able to establish joint participation between the target audience and the design/research team, which is essential from the initial stages of conception [66]. In addition, innovative operational features were integrated to promote regular exercise adherence, such as customizable alarm reminders, which users can adjust according to their routines.

Unlike many digital health apps that do not undergo rigorous evaluation—leading to gaps in quality, accessibility, and coverage of user needs—this study included a structured

XSL•FO
RenderX

validation process with expert participation [67,68]. The results obtained show high levels of agreement among the evaluators, which reflected in the satisfactory CVI values, confirming the validity and reliability of the app. Such validation is essential to ensure safety, efficacy, and accessibility of digital health tools for older adults [68-70].

Validation by the target audience allowed us to verify the usability and content of the app, according to principles of comprehensibility and suitability to the educational and cultural profile of the participants, highlighting ease of use, simplicity of the interface, and memorization of functions. In this research, we opted to use the SUS Scale, and studies show that an SUS score above 68 indicates an acceptable degree of usability, while scores above 85 would be associated with excellent acceptance of software or an app [46,71]. In the present study, an overall score of >90 points was obtained, confirming high usability, with no inconsistencies in the responses. This suggests strong potential for adherence, satisfaction, and effectiveness of the app as a tool to support physical activity among older adults.

In this perspective, other studies focusing on the use of technologies to track and prevent fall episodes have been developed [37]. The researchers developed FallSA, a mobile fall screening app, and determined its acceptance, concurrent validity, test–retest reliability, discriminatory ability, and predictive validity as a self-screening tool to identify fall risk among older adults in Malaysia. The results of that study indicated a high level of acceptance, with 80% of older adults agreeing with its suitability as a fall self-screening tool, with moderate to strong discriminatory ability in classifying older adults as fallers and non-fallers [37].

In recent years, additional studies have expanded the evidence base beyond FallSA. In a randomized controlled trial combining wearable activity trackers with a home-based multicomponent exercise program in older adults with a history of falls, significant improvements in fear of falling and physical function were observed after 12 weeks [72]. Other app-based interventions, which prescribed exercise using reinforcement learning, demonstrated greater user satisfaction and higher exercise intensity under adaptive conditions [73]. These findings highlight the importance of systematic design guidelines for mobile apps targeting older adults, emphasizing usability features such as simplified navigation and visual clarity [69]. Such principles align with the design choices implemented in the *Mais Equilíbrio* (More Balance) app, supporting the notion that its dual focus, adaptation of the OEP protocol, and usability validation are consistent with current best practices in digital fall-prevention interventions.

Unlike previous mHealth solutions for fall prevention, which often focus exclusively on self-assessment tools [37], or generic exercise guidance without a validated scientific protocol [69,72,73], the *Mais Equilíbrio* (More Balance) app provides a unique contribution by digitally adapting and delivering the evidence-based OEP in the Brazilian context. Its dual validation process with both health professionals and older adults ensures not only scientific rigor but also usability and acceptability in the target population, addressing key barriers to adherence in home-based exercise interventions.

As such, it is believed that this kind of validation, based primarily on updated technical guidelines and protocols, ensures safety, quality, and effectiveness in health care. This is in line with the fundamental principles of usability assessment, especially when used on mobile devices by older adults. These results corroborate other studies that highlighted the need for design adaptations and combined alternatives for solutions, aiming to improve older adults' engagement with mHealth interventions [74]. Thus, usability testing is necessary before making the app available to the end user, as discussed in other studies [47,70,75].

Acceptance of usability by the target audience is directly associated with adherence to and continued use of health technologies [76]. In this sense, the validation carried out in this study revealed high scores in the SUS domains, highlighting the ease of use of the system from the first use, the agility in performing the proposed tasks, the low occurrence of errors, the ease of reuse even after a long period without contact with the system, in addition to a pleasant design. These results are consistent with findings from previous studies that identify key usability factors influencing user adherence, such as ease of use, efficiency through simple navigation, intuitive tasks, and accessible design [77-79]; high satisfaction and acceptance driven by a positive perception of the tool's usefulness [80,81]; and facilitated learning, characterized by fewer barriers and increased confidence among older adults in using technology [82].

Limitations

This study has some limitations. First, the validation was conducted over a limited time period, making it impossible to evaluate the prolonged use of the app or its effects on user behavior over time. Another limitation is related to the absence of clinical or functional assessment of participants regarding falls, which could deepen understanding of the practical impact of the tool. In addition, the absence of formal economic analyses, such as cost-effectiveness and budgetary impact, represents an important gap in assessing the sustainability and feasibility of large-scale implementation in the public health system. To this end, we aim to soon evaluate the prolonged use of the app or its effects on user behavior, adding clinical and functional assessments to measure the app's effectiveness in reducing fall risk, improving balance and MS, as well as incorporating economic analyses to understand costs and benefits. Furthermore, cognitive status, an important factor in studies involving older adults, was not formally assessed, which should be considered when interpreting the findings. Most older adult participants were women (n=16; 66.7%), which may limit generalizability and highlight potential gender differences in usability perception. Finally, as this study focused only on usability validation, there is no evidence of clinical outcomes, such as improved balance or reduced falls, and this limitation should be acknowledged. Future studies with larger, more diverse samples, including cognitive, clinical, and functional assessments, are warranted to provide more robust evidence of the app's effectiveness and sustainability.

XSL•FO
RenderX

Clinical and Research Implications

The app has proven to be a promising tool for promoting the health of older adults, with the potential to support fall prevention through accessible educational content and guided exercises. Among its strengths are its validation by physical education and physical therapy experts, its intuitive interface, and its high level of usability. Future studies may test the app's effectiveness with older adults in continuous use, evaluate health outcomes, and explore its effectiveness in reducing fear of falling, risk of falls, and increasing MS and balance through the exercises contained in the app.

In practical terms, integrating *Mais Equilíbrio* (More Balance) into Brazil's PHC system could follow community-based mHealth models used in other low- and middle-income countries (LMIC). Mobile interventions have improved strength, balance, and self-efficacy in preventing falls among older adults [83], while reviews emphasize the need for user-friendly design, trained health professionals, infrastructure, and alignment with public health policies [84].

In Brazil, the involvement of community health workers and family health teams may enhance the adoption and dissemination of the app, while its integration into routine primary care activities could support scalability and long-term sustainability in fall-prevention initiatives. Beyond local implementation, *Mais Equilíbrio* (More Balance) also aligns with broader policy agendas. Both the Sustainable Development Goals and the World Health Organization Global Digital Health Strategy emphasize the development of innovation ecosystems to strengthen PHC and improve population health [85]. Consistently, the Brazilian *Digital Health Strategy 2020 - 2028* establishes priorities for building an interconnected ecosystem that brings together the *Sistema Único de Saúde* (Unified Health System), technology companies, startups, universities, and research centers. A key action is to promote innovation environments that leverage digital connectivity to meet users' needs [86], underscoring the potential of *Mais Equilíbrio* (More Balance) to be scaled within the SUS and adapted to other LMIC contexts.

Conclusions

This study developed and validated the mobile app *Mais Equilíbrio* (More Balance), a digital tool designed to guide home-based physical exercises adapted for older adults. Content validation by experts ensured quality and instructional relevance, while high usability and acceptance confirmed by the older adults themselves attest to the tool's applicability and comprehensibility. The results confirmed *Mais Equilíbrio* (More Balance) as an understandable and applicable tool ready for use.

Therefore, the findings reinforce the potential of mHealth technologies to overcome barriers to health access, especially in the context of PHC, where challenges such as limited access to professionals and resources persist in LMICs. The high acceptance of the app by the target audience reaffirms its relevance. To ensure the sustainability of the intervention and expand its dissemination, we propose integrating the app into PHC health promotion strategies, training community agents to provide technical support, and establishing partnerships with health departments and active aging programs. In addition, this study offers valuable contributions to the development and validation of health communication strategies through mobile apps, highlighting the importance of training older adults in self-management of health, especially in the context of physical activity and fall prevention.

Although *Mais Equilíbrio* (More Balance) represents a valuable contribution to increasing research on digital health and the validation of interactive tools, future studies are needed to assess its long-term clinical effectiveness and impact on objective health outcomes in the older adults population, especially in fall prevention. It is also essential to explore the adaptation of the app to different cultural and linguistic contexts and to investigate the economic feasibility and financial sustainability of its large-scale implementation.

Acknowledgments

This study was funded by the Federal District Research Support Foundation, public notice 09/2022, with support from the University of Brasília. The funds covered technological development, infrastructure, and technical team costs. The funding institution had no influence on the design, analysis, or publication of the study. We would like to thank everyone who agreed to participate in this study. We also thank the National Council for Scientific and Technological Development, the Coordination for the Improvement of Higher Education Personnel, and the University of Brasília, which also supported this research.

Conflicts of Interest

None declared.

References

1. Dzau VJ, Inouye SK, Rowe JW, Finkelman E, Yamada T. Enabling healthful aging for All - the National Academy of Medicine grand challenge in healthy longevity. N Engl J Med 2019 Oct 31;381(18):1699-1701. [doi: 10.1056/NEJMp1912298] [Medline: 31633895]
2. Decade of healthy ageing: the global strategy and action plan on ageing and health 2016–2020: towards a world in which everyone can live a long and healthy life: report by the director-general. World Health Organization. 2020. URL: https://apps.who.int/iris/handle/10665/355618 [accessed 2025-10-24]

XSL•FO
RenderX

3. World population prospects 2019. United Nations Department of Economic and Social Affairs. 2020. URL: https://population.un.org/wpp/ [accessed 2025-10-24]
4. Yang Y, Wang K, Liu H, et al. The impact of Otago exercise programme on the prevention of falls in older adult: a systematic review. Front Public Health 2022;10:953593. [doi: 10.3389/fpubh.2022.953593]
5. Gawrońska K, Lorkowski J. Falls, aging and public health - a literature review. Ortop Traumatol Rehabil 2020 Dec 31;22(6):397-408. [doi: 10.5604/01.3001.0014.6044] [Medline: 33506804]
6. Otaka Y, Morita M, Mimura T, Uzawa M, Liu M. Establishment of an appropriate fall prevention program: a community-based study. Geriatr Gerontol Int 2017 Jul;17(7):1081-1089. [doi: 10.1111/ggi.12831] [Medline: 27492888]
7. WHO global report on falls prevention in older age. World Health Organization. 2008. URL: https://extranet.who.int/agefriendlyworld/wp-content/uploads/2014/06/WHo-Global-report-on-falls-prevention-in-older-age.pdf [accessed 2025-10-29]
8. Falls - key facts. World Health Organization. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/falls [accessed 2025-10-24]
9. Ehrhardt A, Hostettler P, Widmer L, et al. Fall-related functional impairments in patients with neurological gait disorder. Sci Rep 2020 Dec 3;10(1):21120. [doi: 10.1038/s41598-020-77973-4] [Medline: 33273488]
10. Florence CS, Bergen G, Atherly A, Burns E, Stevens J, Drake C. Medical costs of fatal and nonfatal falls in older adults. J Am Geriatr Soc 2018 Apr;66(4):693-698. [doi: 10.1111/jgs.15304] [Medline: 29512120]
11. Luebbert S, Christensen W, Finkel C, Worsowicz G. Falls in senior adults: demographics, cost, risk stratification, and evaluation. Mo Med 2022;119(2):158-163. [Medline: 36036034]
12. Kim J, Lee W, Lee SH. A systematic review of the guidelines and Delphi study for the multifactorial fall risk assessment of community-dwelling elderly. IJERPH 2020;17(17):6097. [doi: 10.3390/ijerph17176097]
13. Vaishya R, Vaish A. Falls in older adults are serious. Indian J Orthop 2020 Feb;54(1):69-74. [doi: 10.1007/s43465-019-00037-x] [Medline: 32257019]
14. Smith ADA, Silva AO, Rodrigues RAP, Moreira M, Nogueira JDA, Tura LFR. Assessment of risk of falls in elderly living at home. Rev Latino-Am Enfermagem 2017;25. [doi: 10.1590/1518-8345.0671.2754]
15. Drummond FMM, Lourenço RA, Lopes CDS. Incidence, persistence and risk factors of fear of falling in older adults: cohort study (2008-2013) in Rio de Janeiro, Brazil. Rev Saude Publica 2020;54:56. [doi: 10.11606/S1518-8787.2020054001939] [Medline: 32556022]
16. Alonso AC, Ribeiro SM, Luna NMS, et al. Association between handgrip strength, balance, and knee flexion/extension strength in older adults. PLoS ONE 2018;13(6):e0198185. [doi: 10.1371/journal.pone.0198185] [Medline: 29856802]
17. Cuevas-Trisan R. Balance problems and fall risks in the elderly. Phys Med Rehabil Clin N Am 2017 Nov;28(4):727-737. [doi: 10.1016/j.pmr.2017.06.006] [Medline: 29031339]
18. Makino K, Makizako H, Doi T, et al. Fear of falling and gait parameters in older adults with and without fall history. Geriatr Gerontol Int 2017 Dec;17(12):2455-2459. [doi: 10.1111/ggi.13102] [Medline: 28656737]
19. Pua YH, Ong PH, Clark RA, Matcher DB, Lim ECW. Falls efficacy, postural balance, and risk for falls in older adults with falls-related emergency department visits: prospective cohort study. BMC Geriatr 2017 Dec 21;17(1):291. [doi: 10.1186/s12877-017-0682-2] [Medline: 29268720]
20. Jiang Y, Wang M, Liu S, Ya X, Duan G, Wang Z. The association between sedentary behavior and falls in older adults: a systematic review and meta-analysis. Front Public Health 2022;10:1019551. [doi: 10.3389/fpubh.2022.1019551]
21. Cunningham C, O' Sullivan R, Caserotti P, Tully MA. Consequences of physical inactivity in older adults: a systematic review of reviews and meta‐analyses. Scand J Med Sci Sports 2020 May;30(5):816-827. [doi: 10.1111/sms.13616] [Medline: 32020713]
22. Qian XX, Chau PH, Kwan CW, et al. Investigating risk factors for falls among community-dwelling older adults according to WHO's risk factor model for falls. J Nutr Health Aging 2021;25(4):425-432. [doi: 10.1007/s12603-020-1539-5] [Medline: 33786558]
23. Batsis JA. Obesity in the older adult: special issue. J Nutr Gerontol Geriatr 2019;38(1):1-5. [doi: 10.1080/21551197.2018.1564197] [Medline: 30806583]
24. de Paiva FTF, Stival MM, de Lima LR, et al. Predictive factors for reduced functional mobility in elderly diabetics and non-diabetics. Int J Diabetes Dev Ctries 2021 Jun;41(2):314-321. [doi: 10.1007/s13410-020-00873-w]
25. Clegg A, Barber S, Young J, Iliffe S, Forster A. The Home-based Older People's Exercise (HOPE) trial: a pilot randomised controlled trial of a home-based exercise intervention for older people with frailty. Age Ageing 2014 Sep;43(5):687-695. [doi: 10.1093/ageing/afu033] [Medline: 24742587]
26. Jahanpeyma P, Kayhan Koçak F, Yıldırım Y, Şahin S, Şenuzun Aykar F. Effects of the Otago exercise program on falls, balance, and physical performance in older nursing home residents with high fall risk: a randomized controlled trial. Eur Geriatr Med 2021 Feb;12(1):107-115. [doi: 10.1007/s41999-020-00403-1] [Medline: 33237565]
27. Blain H, Miot S. How can we prevent falls? In: Orthogeriatrics: The Management of Older Patients with Fragility Fractures: Springer International Publishing; 2021:273-290. [doi: 10.1007/978-3-030-48126-1_16]
28. Gama LN. Desenvolvimento e avaliação de aplicativo móvel na prevenção de riscos osteomusculares no trabalho de enfermagem. Texto Context 2019. [doi: 10.1590/1980-265x-tce-2018-0214]

XSL•FO
RenderX

29. Debon R, Coleone JD, Bellei EA, De Marchi ACB. Mobile health applications for chronic diseases: a systematic review of features for lifestyle improvement. Diabetes Metab Syndr 2019;13(4):2507-2512. [doi: 10.1016/j.dsx.2019.07.016] [Medline: 31405669]
30. Gabbiadini A, Greitemeyer T. Fitness mobile apps positively affect attitudes, perceived behavioral control and physical activities. J Sports Med Phys Fitness 2019 Mar;59(3):407-414. [doi: 10.23736/S0022-4707.18.08260-9] [Medline: 29619794]
31. Mansson L, Lundin-Olsson L, Skelton DA, et al. Older adults' preferences for, adherence to and experiences of two self-management falls prevention home exercise programmes: a comparison between a digital programme and a paper booklet. BMC Geriatr 2020 Jun 15;20(1):209. [doi: 10.1186/s12877-020-01592-x] [Medline: 32539711]
32. Graham SA, Stein N, Shemaj F, Branch OH, Paruthi J, Kanick SC. Older adults engage with personalized digital coaching programs at rates that exceed those of younger adults. Front Digit Health 2021;3:642818. [doi: 10.3389/fdgth.2021.642818] [Medline: 34713112]
33. De Santis KK, Mergenthal L, Christianson L, Zeeb H. Digital technologies for health promotion and disease prevention in older people: protocol for a scoping review. JMIR Res Protoc 2022 Jul 21;11(7):e37729. [doi: 10.2196/37729] [Medline: 35862187]
34. Paiva JOV, Andrade RMC, de Oliveira PAM, et al. Mobile applications for elderly healthcare: a systematic mapping. PLoS ONE 2020;15(7):e0236091. [doi: 10.1371/journal.pone.0236091] [Medline: 32730266]
35. Dick W, Carey L, Carey JO. A Model for the Systematic Design of Instruction: Routledge; 2013, Vol. 1.
36. Sherrington C, Fairhall N, Wallbank G, et al. Exercise for preventing falls in older people living in the community: an abridged Cochrane systematic review. Br J Sports Med 2020 Aug;54(15):885-891. [doi: 10.1136/bjsports-2019-101512] [Medline: 31792067]
37. Singh DKA, Goh JW, Shaharudin MI, Shahar S. A mobile app (FallSA) to identify fall risk among Malaysian community-dwelling older persons: development and validation study. JMIR Mhealth Uhealth 2021 Oct 12;9(10):e23663. [doi: 10.2196/23663] [Medline: 34636740]
38. He Z, Wu H, Zhao G, et al. The effectiveness of digital technology-based Otago Exercise Program on balance ability, muscle strength and fall efficacy in the elderly: a systematic review and meta-analysis. BMC Public Health 2025 Jan 7;25(1):71. [doi: 10.1186/s12889-024-21251-9] [Medline: 39773392]
39. Barra DCC, Paim SMS, Sasso G, Colla GW. Métodos para desenvolvimento de aplicativos móveis em saúde: revisão integrativa da literatura. Texto Contexto Enferm 2018 Jan 8;26(4). [doi: 10.1590/0104-07072017002260017]
40. Campbell AJ, Robertson MC. Otago Exercise Programme to prevent falls in older adults. Live Stronger for Longer. 2003. URL: https://www.livestronger.org.nz/assets/Uploads/acc1162-otago-exercise-manual.pdf [accessed 2025-10-29]
41. Oliveira AKS, Lima ECS, de Sena Caridade ER. O uso dos framework flutter e nest.js para desenvolvimento de um front-end e um back-end de uma aplicação de help desk. Revista Ibero-Americana De Humanidades Ciências E Educação 2022;8(11):1766-1786. [doi: 10.51891/rease.v8i11.7771]
42. Fehring RJ. The Fering Model. In: Carrol-Johnson RM, Pasquete M, editors. 1994 Presented at: Conference on the Classification of Nursing Diagnoses; Philadelphia.
43. Bujang MA, Omar ED, Foo DHP, Hon YK. Sample size determination for conducting a pilot study to assess reliability of a questionnaire. Restor Dent Endod 2024 Feb;49(1):e3. [doi: 10.5395/rde.2024.49.e3] [Medline: 38449496]
44. Sousa CS, Turrini RNT, Poveda VB. Tradução e adaptação do instrumento "suitability assessment of materials" (sam) para o português. Rev enferm UFPE Online 2015;9(5):7854-7861. [doi: 10.5205/1981-8963-v9i5a10534p7854-7861-2015]
45. Presotto M, Rieder CDM, Olchik MR. Validação de conteúdo e confiabilidade do Protocolo de Avaliação dos Distúrbios Adquiridos de Fala em Indivíduos com Doença de Parkinson (PADAF). Codas 2019;31(5). [doi: 10.1590/2317-1782/20192018230]
46. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123 [FREE Full text]
47. Correia GDS, Silva AD, Pinto LF, Góes FGB, Goulart MDCEL, Pereira-Ávila FMV. Usabilidade do aplicativo móvel ROBOVID para educação em saúde acerca da COVID-19. Rev Latino-Am Enfermagem 2024;32. [doi: 10.1590/1518-8345.6924.4191]
48. Cestari VRF, Florêncio RS, Garces TS, Pessoa VDP, Moreira TMM. Benchmarking of mobile apps on heart failure. Rev Bras Enferm 2022;75(1). [doi: 10.1590/0034-7167-2020-1093]
49. Lourenço DF, Valentim EC, Lopes MDM. Translation and cross-cultural adaptation of the system usability scale to Brazilian Portuguese. Aquichan 2022 May 13;22(2):1-16. [doi: 10.5294/aqui.2022.22.2.8]
50. Hyzy M, Bond R, Mulvenna M, et al. System usability scale benchmarking for digital health apps: meta-analysis. JMIR Mhealth Uhealth 2022 Aug 18;10(8):e37290. [doi: 10.2196/37290] [Medline: 35980732]
51. Cavalcanti HDO, Bushatsky M, Barros M, Melo CDS, Delgado Filho AJF. Evaluation of the usability of a mobile application in early detection of pediatric cancer. Rev Gaucha Enferm 2021;42:e20190384. [doi: 10.1590/1983-1447.2021.20190384] [Medline: 34287592]
52. Lopes ROP, Chagas SR, Gomes EDS, Barbosa JDA, Silva Í, Brandão MAG. Benchmarking de aplicativos móveis voltados à saúde de pessoas com Diabetes Mellitus. Rev Latino-Am Enfermagem 2024;32:e4221. [doi: 10.1590/1518-8345.7182.4222]

XSL•FO
RenderX

53. Luz PD, Galindo NM Neto, Machado RDS, Marques MDC, Santos AD, Andrade E. Construção e validação de tecnologia educacional para adolescentes sobre reanimação cardíaca. Acta Paulista de Enfermagem 2023 Jun 7;36. [doi: 10.37689/acta-ape/2023AO016932]
54. Mattos S, Moreira T, Florêncio R, Cestari V. Elaboração e validação de um instrumento para mensurar Autopercepção de Saúde em adultos. Saúde debate 2021;45(129):366-377. [doi: 10.1590/0103-1104202112909]
55. Costa MD, Zandonadi RP, Ginani VC, et al. Connecting health and technology: validation of instant messaging for use as diabetes mellitus control strategy in older Brazilian adults. Int J Environ Res Public Health 2025 Feb 14;22(2):282. [doi: 10.3390/ijerph22020282] [Medline: 40003508]
56. Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Res Nurs Health 2007 Aug;30(4):459-467. [doi: 10.1002/nur.20199] [Medline: 17654487]
57. Oliveira L, Poínhos R, Afonso C, Vaz Almeida MD. Information sources on healthy eating among community living older adults. Int Q Community Health Educ 2021 Jan;41(2):153-158. [doi: 10.1177/0272684X20915362] [Medline: 32249704]
58. Gordon NP, Hornbrook MC. Older adults' readiness to engage with eHealth patient education and self-care resources: a cross-sectional survey. BMC Health Serv Res 2018 Mar 27;18(1):220. [doi: 10.1186/s12913-018-2986-0] [Medline: 29587721]
59. Ebaid D, Crewther SG. Visual information processing in young and older adults. Front Aging Neurosci 2019;11:116. [doi: 10.3389/fnagi.2019.00116] [Medline: 31156422]
60. Battineni G, Baldoni S, Chintalapudi N, et al. Factors affecting the quality and reliability of online health information. Digit Health 2020;6:2055207620948996. [doi: 10.1177/2055207620948996] [Medline: 32944269]
61. Turner A, Flood VM, LaMonica HM. Older adults' needs and preferences for a nutrition education digital health solution: a participatory design study. Health Expect 2024 Feb;27(1):e13923. [doi: 10.1111/hex.13923] [Medline: 38014900]
62. Goodman C, Lambert K. Scoping review of the preferences of older adults for patient education materials. Patient Educ Couns 2023 Mar;108:107591. [doi: 10.1016/j.pec.2022.107591] [Medline: 36584555]
63. Mansoor H, Gerych W, Alajaji A, et al. INPHOVIS: interactive visual analytics for smartphone-based digital phenotyping. Visual Informatics 2023 Jun;7(2):13-29. [doi: 10.1016/j.visinf.2023.01.002]
64. Montero-Odasso M, van der Velde N, Martin FC, et al. World guidelines for falls prevention and management for older adults: a global initiative. Age Ageing 2022 Sep 2;51(9):afac205. [doi: 10.1093/ageing/afac205] [Medline: 36178003]
65. Rodrigues VES, Oliveira FGL, Machado ALG, Beleza CMF, Pereira FGF. Construção e validação de gerontecnologias cuidativo-educacionais: revisão integrativa. Rev Bras Geriatr Gerontol 2021;24(4). [doi: 10.1590/1981-22562021024.210144.pt]
66. Marques ADB, Moreira TMM, Jorge TV, Rabelo SMS, Carvalho RD, Felipe GF. Usability of a mobile application on diabetic foot self-care. Rev Bras Enferm 2020;73(4):e20180862. [doi: 10.1590/0034-7167-2018-0862] [Medline: 32520095]
67. Nambisan P, Kapoor A, Kreps G, Dangeti RS. Framework-based evaluation of mobile health apps for older adults: implications for technology-enabled self-care. Innovation Aging 2024 Dec 31;8(Supplement_1):898-898. [doi: 10.1093/geroni/igae098.2900]
68. Park GE, Park YH, Kim KG, Park JY, Hwang M, Lee S. Mobile application for digital health coaching in the self-management of older adults with multiple chronic conditions: a development and usability study. Healthc Inform Res 2024 Oct;30(4):344-354. [doi: 10.4258/hir.2024.30.4.344] [Medline: 39551921]
69. Gomez-Hernandez M, Ferre X, Moral C, Villalba-Mora E. Design guidelines of mobile apps for older adults: systematic review and thematic analysis. JMIR Mhealth Uhealth 2023 Sep 21;11:e43186. [doi: 10.2196/43186] [Medline: 37733401]
70. Czuber NK, Garabedian PM, Rice H, Tejeda CJ, Dykes PC, Latham NK. Human-centered design and development of a fall prevention exercise app for older adults in primary care settings. Appl Clin Inform 2024 May;15(3):544-555. [doi: 10.1055/a-2267-1727] [Medline: 38350643]
71. Lima CD, Coelho RA, Medeiros MS, Kubrusly M, Marçal E, Peixoto Júnior AA. Development and validation of a mobile application for the teaching of electrocardiogram. Rev Bras Educ Med 2019;43(1 suppl 1):157-165. [doi: 10.1590/1981-5271v43suplemento1-20190164.ing]
72. Kim Y, Park KH, Noh HM. Effects of integrating wearable activity trackers with a home-based multicomponent exercise intervention on fall-related parameters and physical function in older adults: randomized controlled trial. JMIR Mhealth Uhealth 2025 May 8;13:e64458. [doi: 10.2196/64458] [Medline: 40340847]
73. Doherty C, Lambe R, O'Grady B, et al. An evaluation of the effect of app-based exercise prescription using reinforcement learning on satisfaction and exercise intensity: randomized crossover trial. JMIR Mhealth Uhealth 2024 Nov 26;12:e49443. [doi: 10.2196/49443] [Medline: 39622712]
74. van Acker J, Maenhout L, Compernolle S. Older adults' user engagement with mobile health: a systematic review of qualitative and mixed-methods studies. Innov Aging 2023;7(2):igad007. [doi: 10.1093/geroni/igad007] [Medline: 37007638]
75. Paulissen JMJ, Zegers CML, Nijsten IR, et al. Performance and usability evaluation of a mobile health data capture application in clinical cancer trials follow-up. Tech Innov Patient Support Radiat Oncol 2022 Dec;24:107-112. [doi: 10.1016/j.tipsro.2022.10.005] [Medline: 36387779]
76. Khamaj A, Ali AM. Examining the usability and accessibility challenges in mobile health applications for older adults. Alexandria Eng J 2024 Sep;102:179-191. [doi: 10.1016/j.aej.2024.06.002]

XSL•FO
RenderX

77. Takano E, Maruyama H, Takahashi T, et al. User experience of older people while using digital health technologies: a systematic review. Appl Sci (Basel) 2023;13(23):12815. [doi: 10.3390/app132312815]
78. Zainal A, Aziz NFA, Ahmad NA, et al. Usability measures used to enhance user experience in using digital health technology among elderly: a systematic review. Bulletin EEI 2023;12(3):1825-1832. [doi: 10.11591/eei.v12i3.4773]
79. McCarthy B, Sabharwal JK, Chawla S. Old age or cognitive decline? Examining the usability of a mobile health app for older Australians. Inform Health Soc Care 2024 Jan 2;49(1):83-97. [doi: 10.1080/17538157.2024.2332691] [Medline: 38529731]
80. Candelaria D, Serafica R, Reyes AT, et al. Does perceived usability of a multi-component mobile device intervention improve health-related quality of life? Eur Heart J 2024 Oct 28;45(Supplement_1). [doi: 10.1093/eurheartj/ehae666.3389]
81. Ahn J, Yang Y, Park G. Advancing elderly diabetes care: exploring the usability and acceptance of continuous glucose monitoring (CGM). Geriatr Nurs (Lond) 2024 Sep;59:15-25. [doi: 10.1016/j.gerinurse.2024.06.041]
82. Soffer T, Raban Y, Warshawski S, Barnoy S. The impact of emerging technologies on healthcare needs of older people. Health Policy Technol 2024 Nov;13(5):100935. [doi: 10.1016/j.hlpt.2024.100935]
83. Brew-Sam N, Chib A, Torres AYF, Ng JXJ, Wong YTJ, Sze-G Y. An integrated mHealth campaign to reduce the risk of falling for older adults. J Appl Gerontol 2022 May;41(5):1336-1347. [doi: 10.1177/07334648211062877] [Medline: 35085043]
84. Greenall-Ota J, Yapa HM, Fox GJ, Negin J. Qualitative evaluation of mHealth implementation for infectious disease care in low- and middle-income countries: narrative review. JMIR Mhealth Uhealth 2024 Dec 13;12:e55189. [doi: 10.2196/55189] [Medline: 39670953]
85. Global strategy on digital health 2020–2025. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/9789240020924 [accessed 2025-10-24]
86. Estratégia de saúde digital para o Brasil 2020–2028. Biblioteca Virtual em Saúde: Ministério da Saúde. 2020. URL: http://bvsms.saude.gov.br/bvs/publicacoes/estrategia_saude_digital_Brasil.pdf [accessed 2025-10-29]

Abbreviations

CVI: Content Validity Index
LMIC: low- and middle-income countries
mHealth: mobile health
MS: muscle strength
OEP: Otago Exercise Program
PHC: primary health care
Q: Question (related to System Usability Scale)
SAM: Suitability Assessment of Materials
SQUIRE 2.0: Revised Standards for Quality Improvement Reporting Excellence
SUS: System Usability Scale

Edited by P Jansons; submitted 15.07.25; peer-reviewed by D Parmak, P Adedigba; revised version received 06.10.25; accepted 13.10.25; published 04.11.25.

Please cite as:
Leite MM, Silva ADO, Funghetto SS, de Lima LR, Mezavila Abdelmur SB, Pinheiro HA, de Souza Silva CM, Dutra MT, Stival MM Home-Based Exercise and Fall Prevention in Older Adults: Development, Validation and Usability of the Mais Equilíbrio Mobile App
JMIR Aging 2025;8:e80724
URL: https://aging.jmir.org/2025/1/e80724
doi:10.2196/80724

© Mateus Medeiros Leite, Alessandro de Oliveira Silva, Silvana Schwerz Funghetto, Luciano Ramos de Lima, Samuel Barbosa Mezavila Abdelmur, Hudson Azevedo Pinheiro, Calliandra Maria de Souza Silva, Maurílio Tiradentes Dutra, Marina Morato Stival. Originally published in JMIR Aging (https://aging.jmir.org), 4.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Risk Factors for Community-Dwelling Older Adults Dropping Out of Self-Guided, Remote, and Web-Based Longitudinal Research: Predictive Modeling of Data from the Web-LABrainS Platform

Luke Daniel Braun[1], BS; H Raymond Allen[2], BS, PhD; Jeffrey N Keller[1], BS, PhD

[1]Pennington Biomedical Research Center, Institute for Dementia Research and Prevention, 6400 Perkins Rd, Baton Rouge, LA, United States

[2]Pennington Biomedical Research Center, Computing Services, Baton Rouge, LA, United States

Corresponding Author:
Jeffrey N Keller, BS, PhD
Pennington Biomedical Research Center, Institute for Dementia Research and Prevention, 6400 Perkins Rd, Baton Rouge, LA, United States

Abstract

Background: Little is currently known regarding the feasibility of using a self-guided, remote, web-based platform as the basis for a longitudinal study of aging in community-dwelling older adults (OAs). This study describes the feasibility and risk factors for participant dropout found when using this approach as part of the web-based Louisiana Aging Brain Study (web-LABrainS).

Objective: This study used data from 402 participants in the web-LABrainS effort to determine the feasibility of using a self-guided, remote, and web-based platform as the basis for conducting longitudinal research in community-dwelling older adults. Additionally, we sought to determine the risk factors associated with participant dropout over a 12-month period in web-LABrainS and determine whether the same risk factors associated with dropout from in-clinic longitudinal studies were observed in web-LABrainS dropouts.

Methods: Participants were enrolled in web-LABrainS on a rolling basis using word-of-mouth promotional efforts. Participants used the web-LABrainS platform to provide electronic consent, demographic and health information, answer questionnaires, and complete assessments as part of a self-guided and web-based effort off-site of the clinic (remote). Following completion of the baseline battery, participants were contacted by email every 6 months to complete another round of the web-LABrainS battery. The data in this study were collected from 402 participants, 217 (54.0%) of whom completed baseline, 6-month, and 12-month assessments (adherent participants) and 185 (46%) of whom participated in only the baseline and no subsequent web-LABrainS batteries (dropout participants).

Results: Our study indicates that even with limited participant outreach and retention efforts, it is feasible to conduct longitudinal clinical research studies in community-dwelling OAs using a self-guided, remote, and web-based approach. In contrast to traditional in-clinic longitudinal studies, dropouts from web-LABrainS were not observed to be significantly different with respect to age, education, gender, marital status, or living alone (P=.67, .16, .29, .051, .31). Similar to traditional longitudinal studies, dropouts from web-LABrainS had significantly higher use of depression medication, decreased self-reported mobility, and decreased delayed recall performance (P=.007, .007, .004). Interestingly, no differences in technology use, comfort with technology, time of assessment, or consent to be contacted about future research were observed between adherents and dropouts (P=.17, .36, .47, .40). Predictive binary logistic regression yielded a moderately accurate model and further supported a negative association between cognitive ability and dropout (OR 0.77, 95% CI 0.61-0.96).

Conclusions: Our study provides some of the first clinical evidence for the feasibility of conducting longitudinal human research using a self-guided, remote, and web-based approach. Additionally, these data highlight the similarities and differences in key factors associated with participant dropout using this type of approach compared to traditional longitudinal study formats. The findings from this study may help guide the design and deployment of future longitudinal studies of older adults focused on self-guided, remote, or web-based approaches.

(JMIR Aging 2025;8:e68826) doi:10.2196/68826

KEYWORDS

community-dwelling; longitudinal study; older adult; web-based assessment; study attrition

XSL•FO
RenderX

Introduction

Longitudinal studies of aging have proven to be invaluable in helping to identify the underlying clinical factors involved in the onset and progression of a diverse array of chronic conditions including dementia, frailty, and cardiovascular disease [1-4]. The vast majority of human longitudinal studies conducted to date have focused on the use of in-clinic study measures and the administration of in-person pencil-and-paper assessments and questionnaires [5]. There are multiple advantages to focusing on traditional onsite administration of assessments and measures for longitudinal studies. For example, traditional onsite evaluations significantly increase the control of the study environment, increase the rigor of assessments, and promote the centralization of clinical expertise and resources. There are also multiple disadvantages to focusing on the use of onsite visits for longitudinal human studies. Disadvantages include the high cost of conducting studies in a clinical setting and the limited availability of study visit slots within the typical daily clinic workflow. Additionally, the reliance on in-clinic visits selects for those participants who have the resources and flexibility to travel to the clinic (socioeconomic and geographical burden) and increases the overall entropy for participants to participate in clinical research.

The dropout rate for in-clinic longitudinal studies involving OAs has been reported to range from 15% to 60% [6-23]. Previous studies have identified a diverse array of risk factors associated with participant dropout from in-clinic longitudinal studies [6-23]. Factors reported to drive participant dropout in traditional longitudinal studies include advanced age, lower education, gender, and the presence of chronic health conditions [6-23].

In the last decade, there has been a dramatic increase in the development, validation, and use of self-guided, remote, and web-based assessments as part of clinical research [24-33]. While these efforts preceded COVID-19, the COVID-19 pandemic introduced an urgency and awareness for the need to quickly develop viable alternatives for conducting large-scale clinical research efforts independent of traditional research settings [34-38]. Together, these novel approaches have focused on the development of self-guided or web-based assessments for the collection of demographics, health information, and delivery of questionnaires and assessments outside of the clinic. As a consequence, today there is a growing scientific literature outlining self-guided, remote, or web-based assessments for cognitive, mobility, and mental health studies in older adults (OAs).

Very little is currently known in terms of the feasibility of using a self-guided, remote, and web-based platform for conducting longitudinal research in OAs. Previously, we validated a self-guided, remote, and web-based platform for the collection of demographics, health information, and delivery of questionnaires and clinical assessments [39]. We then initiated efforts to determine the feasibility of using this platform to conduct a self-guided, remote, and web-based longitudinal aging study in community-dwelling OAs. This study was defined as the web-based Louisiana Aging Brain Study (web-LABrainS).

In this research effort, we report on the feasibility of deploying web-LABrainS as a potential format for conducting longitudinal research in community-dwelling OAs. Feasibility was evaluated using extremely limited outreach (word-of-mouth) and limited retention efforts (2 email reminders to encourage participation) in order to define performance using minimal support conditions. We observed that even in this restricted paradigm, the web-LABrainS platform was a feasible option for conducting an initial 12-month period research follow-up. We also observed similarities and differences between the factors associated with study dropout in web-LABrainS, as compared to previous studies using traditional approaches for longitudinal research. Additionally, we describe the correlations between technology use, comfort with technology, time of day for assessments, and other novel findings toward adherence in web-LABrainS. Taken together, these efforts allowed us to achieve our goal of determining the feasibility of using a largely automated, self-guided, and web-based approach for onboarding, assessing, reporting, and conducting follow-up assessments as part of longitudinal studies involving community-dwelling OAs.

Methods

Study Participants

All data in the current study were obtained from participants enrolled in web-LABrainS. The web-LABrainS study is a self-guided, web-based battery of questionnaires (demographics, mobility, medication use, and health history) and assessments [39]. Web-LABrainS has been continually enrolling participants since 2020. Participants who heard about web-LABrainS via word-of-mouth (no promotions or outreach efforts for web-LABrainS have been conducted) contacted the Pennington Biomedical Research Center (PBRC) to be enrolled. Participants were then sent an email containing a hyperlink that took them to the web-LABrainS site to consent to and complete the web-LABrainS battery. Individuals were sent an additional email and hyperlink every 6 months to repeat the web-LABrainS battery.

Because subjects enroll in and complete the web-LABrainS battery of assessments on a rolling basis, for the purposes of this analysis, we pulled participant data from our servers on October 17, 2024. Data from participants who enrolled in the study any time after this date, as well as assessment data from participants enrolled in the study, were not included in this analysis.

Dropout and Adherent Classification

Subjects were classified as either dropouts or adherents of web-LABrainS. A participant was classified as a dropout if they completed their baseline assessment and had no subsequent assessments. A minimum of 9 months had to have passed since their baseline assessment for a participant to be considered a dropout. If a participant completed a second assessment—even if it was more than 9 months after their baseline assessment—they were not classified as a dropout. Participants were classified as adherents if they completed assessments at a minimum of 3 time points (baseline, 6-month, and 12-month assessments). Participants were recruited to the study via minimal outreach efforts that largely relied on word-of-mouth

XSL•FO
RenderX

and grassroots promotion by the participants in web-LABrainS. Participants were retained in the study by equally minimal retention efforts that were based on 2 email prompts for those individuals who did not open the email link to complete their web-LABrainS assessment at baseline, 6, or 12 months postbaseline. This approach was undertaken to evaluate the feasibility in a low recruitment and retention environment.

Components of Web-LABrainS

The components of the web-LABrainS assessment have been described previously in a study of the feasibility of this self-guided, web-based paradigm [39]. We will briefly summarize the components of the web-LABrainS assessments used in this investigation. First, participants were presented with demographic questions about their date of birth, gender identification, racial and ethnic background, marital status, and education. Participants were then asked about their living situation (including whether they live alone and their type of residence), the types of technology they use, their self-rated comfort with technology, as well as their driving status and frequency. Participants are asked about how they rate their own mobility and are given a short questionnaire adapted from the Life-Space Assessment [40]. Participants proceed to provide their level of concern for their memory and whether they rate it worse than others their age. Participants are asked about the number and type of prescription medications they use, and their ability to read a medication label. They are asked to provide their personal medical history including identifying the presence of current chronic health conditions. To conclude the survey portion of the web-LABrainS battery, participants are asked whether they wish to be contacted for research portions in the future.

In the next section of the web-LABrainS battery, participants take multiple validated cognitive assessments, including an 8-item orientation assessment as well as 4-item immediate and delayed recall assessments. Finally, they take an assessment of acute symptoms of depression and anxiety.

Statistical Analysis

We used 2 types of statistical analysis in this study. We first implemented a between-group comparison of dropouts and adherents. This involved 1-way chi-square tests of independence to assess the significance of the differences in the observed and expected frequencies of all categorical dependent variables that were examined. It also involved Mann-Whitney *U* tests to assess the significance of the differences in the distributions of all numerical-dependent variables that were examined. Exact *P* values were reported; however, a significance threshold (α) of .05 was used to report statistically significant results.

Our second form of statistical analysis was a binary logistic regression model, a predictive model trained on our data from which we calculated odds ratios (ORs) and CIs to determine the relative strength of our examined variables as predictors of dropout status. We trained and cross-validated 4 different models. We used different sets of predictors in each of these models and iteratively tuned their hyperparameters for optimal performance. Depending on the model, this involved adjusting the optimization algorithm, regularization method, regularization strength, size trim tolerance for predictor coefficients, convergence tolerance, or number of iterations over multiple runs of the model and selecting that which resulted in the strongest classification performance. We assessed our model's classification ability using k-fold cross-validation. We first shuffled the dataset and split it into 5 separate folds. Four of the folds were used to train the model, while the last fold was withheld to be used as the test set. This process was repeated with each fold being used as the test set once. The trained model was used to predict the outcomes of the unseen subjects in each set. For each iteration of the model, we generated a classification report that included the model's precision, recall, and F_1-scores to evaluate model performance for each class (dropout and adherent). The weighted and unweighted averages of these scores across the 2 classes were also provided along with the total model accuracy. The support values for these averages were also included. The 5 classification reports produced for each iteration were then averaged.

We then selected the best-performing model from the cross-validation phase. Of the 4 models tested, our chosen model was tied for the highest overall accuracy but exhibited more consistent performance across validation folds and better classification of the minority class. This model included the variables that significantly differed between groups from our first analysis (self-rated mobility score, delayed recall performance, total number of prescription medications taken, use of depression medication, and quality-of-life self-rating) along with demographic variables (age, race, ethnicity, gender, and years of education). With that subset of data selected, we preprocessed our data in the same way we did during our cross-validation stage. This involved applying 1-hot encoding to all categorical variables, converting them into a binary format that the model was able to use. It also involved converting all continuous variables to a standard scale with a mean of 0 and SD of 1 to ensure equal contribution of these variables to our model. Finally, we added a constant term to the dataset to be used as a reference by the model.

The model itself was based on a logit function. The coefficients of this function were estimated using a convex optimization algorithm with L1 or Lasso (Least Absolute Shrinkage and Selection Operator) regularization. We applied a regularization strength of 0.01 to the coefficient estimation. We also set a trim tolerance of 0.05 for coefficients with a small absolute value. After the model was fit to the full dataset, we retrieved the coefficients of each predictor and calculated the corresponding ORs. The ORs show how the probability of the outcome—in this case, dropping out— changes with an increase of 1 SD to a numerical variable or the presence of a binary variable.

All statistical analysis was performed in Python [41] using functions from the *numpy*, *scipy*, *sklearn*, and *statsmodels* packages [42-45]. Outputs from the Web-LABrainS tool were stored in data structures and manipulated using objects and functions from the *pandas* Python package [46]. Data visualization was accomplished using the *matplotlib* and *seaborn* packages [47,48].

XSL•FO
RenderX

Ethical Considerations

All study procedures were approved by the PBRC Institutional Review Board (IRB). Informed consent was provided for all participants prior to the initiation of study procedures. Participants were informed that participation was voluntary and that they could opt out at any time. Informed consent included the ability of PBRC researchers to conduct secondary analysis without additional consent. All data were stored in a deidentified manner to maintain participant confidentiality. No compensation was provided for study participation as outlined in the PBRC IRB-approved study protocol. All data used for this study were obtained from web-LABrainS. IRB approval for the web-LABrainS study was obtained from the PBRC IRB (FWA # 00006218) prior to the initiation of web-LABrainS research efforts. The PBRC IRB approval number is 2020-044-PBRC Web-LABrainS. The study complied with ethical standards outlined in the Belmont Report and Declaration of Helsinki.

Results

In this study, a total of 402 OA participants in web-LABrainS were examined. The demographics of all 402 OA participants at their baseline assessment are provided in Table 1. The baseline age of the entire group was 65.3 years (SD 11.6 years). In terms of representation of different races, genders, and levels of education, our overall sample was overwhelmingly White (355/402, 88.3%), female (298/402, 74.1%), and highly educated (mean 16.8, SD 2.5 years of education). Of the 402 participants in this study, a total of 217 participants (54.0%) completed the web-LABrainS battery at baseline, 6 months, and 12 months. These participants were classified as adherent participants. The remaining 185 (46.0%) completed only the baseline web-LABrainS assessment. These participants were referred to as the dropouts.

XSL•FO
RenderX

Table . Demographics and between-group comparisons for dropouts and adherents.

Measure	Total, N=402	Dropout, n=185	Adherent, n=217	*P* value
Year of baseline assessment, n (%)				.27
2021	100 (24.9)	44 (23.8)	56 (25.8)	
2022	68 (16.9)	33 (17.8)	35 (16.1)	
2023	231 (57.5)	105 (56.8)	126 (58.1)	
2024	3 (0.7)	3 (1.6)	0 (0.0)	
Time of baseline assessment, mean (SD) (HH:MM)	14:00 (4:00)	14:12 (04:00)	13:48 (04:00)	.47
Age (years), mean (SD)	65.3 (11.6)	65.5 (12.0)	65.1 (11.2)	.67
Age quintiles, n (%)				.52
<50	54 (13.4)	27 (14.6)	27 (12.4)	
50 - 60	64 (15.9)	28 (15.1)	36 (16.6)	
60 - 70	127 (31.6)	54 (29.2)	73 (33.6)	
70 - 80	117 (29.1)	53 (28.6)	64 (29.5)	
>80	40 (10.0)	23 (12.4)	17 (7.8)	
Gender, n (%)				.29
Men	103 (25.6)	42 (22.7)	61 (28.1)	
Women	298 (74.1)	143 (77.3)	155 (71.4)	
Prefer to self-describe	1 (0.2)	0 (0.0)	1 (0.5)	
Race, n (%)				.27
White	355 (88.3)	157 (84.9)	198 (91.2)	
Black or African American	30 (7.5)	19 (10.3)	11 (5.1)	
Asian	6 (1.5)	4 (2.2)	2 (0.9)	
American Indian or Alaska Native	0 (0.0)	0 (0.0)	0 (0.0)	
Other	6 (1.5)	3 (1.6)	3 (1.4)	
Ethnicity, n (%)				.16
Hispanic or Latino	5 (1.2)	2 (1.1)	3 (1.4)	
Non-Hispanic	344 (85.6)	165 (89.2)	179 (82.5)	
Other	53 (13.2)	18 (9.7)	35 (16.1)	
Marital status, n (%)				.05
Married	257 (63.9)	106 (57.3)	151 (69.6)	
Never married	28 (7.0)	11 (5.9)	17 (7.8)	
Common-law partner	5 (1.2)	4 (2.2)	1 (0.5)	
Divorced	70 (17.4)	39 (21.1)	31 (14.3)	
Widowed	39 (9.7)	23 (12.4)	16 (7.4)	
Years of education, mean (SD)	16.8 (2.5)	16.6 (2.6)	16.9 (2.5)	.16
Highest level of education, n (%)				.35
High school or GED	15 (3.7)	10 (5.4)	5 (2.3)	
Some college	68 (16.9)	35 (18.9)	33 (15.2)	
Associate's degree	20 (5.0)	7 (3.8)	13 (6.0%	
Bachelor's degree	126 (31.3)	60 (32.4)	66 (30.4)	

XSL•FO
RenderX

Measure	Total, N=402	Dropout, n=185	Adherent, n=217	*P* value
Master's degree	125 (31.1)	52 (28.1)	73 (33.6)	
Doctorate degree	48 (11.9)	21 (11.4)	27 (12.4)	
Living situation, n (%)				.31
Living alone	96 (23.9)	49 (26.5)	47 (21.7)	
Not living alone	306 (76.1)	136 (73.5)	170 (78.3)	
Housing, n (%)				.74
Single residence house	369 (91.8)	168 (90.8)	201 (92.6)	
Assisted living	1 (0.2)	1 (0.5)	0 (0.0)	
Apartment complex	13 (3.2)	7 (3.8)	6 (2.8)	
Stand- alone apartment	5 (1.2)	3 (1.6)	2 (0.9)	
Other	14 (3.5)	6 (3.2)	8 (3.7)	
Average technologies used, mean (SD)	4.0 (1.0)	3.9 (1.1)	4.1 (0.9)	.17
Individual technologies used, n (%)				
Smartphone use	390 (97.0)	179 (96.8)	211 (97.2)	≥.99
Tablet use	307 (76.4)	135 (73.0)	172 (79.3)	.17
Laptop use	342 (85.1)	156 (84.3)	186 (85.7)	.80
Desktop use	330 (82.1)	148 (80.0)	182 (83.9)	.38
Wearable use	252 (62.7)	109 (58.9)	143 (65.9)	.18
Comfort with computers score, mean (SD)				
Comfort with computers score, 1 - 5	3.8 (1.5)	3.8 (1.5)	3.9 (1.5)	.36
Comfort with computers, n (%)				.58
Very comfortable	212 (52.7)	92 (49.7)	120 (55.3)	
Slightly comfortable	54 (13.4)	28 (15.1)	26 (12.0)	
I'm okay	63 (15.7)	29 (15.7)	34 (15.7)	
Slightly uncomfortable	9 (2.2)	6 (3.2)	3 (1.4)	
Very uncomfortable	64 (15.9)	30 (16.2)	34 (15.7)	
Mobility, mean (SD)				
Mobility self-score, 1 - 5	3.8 (1.0)	3.7 (0.9)	3.9 (0.9)	.007
Life-space mobility score, 0 - 6	3.7 (1.2)	3.7 (1.1)	3.7 (1.2)	.96
Falls in last year, n (%)				.86
Fall	80 (19.9)	38 (20.5)	42 (19.4)	
No Fall	322 (80.1)	147 (79.5)	175 (80.6)	
Driving status, n (%)				.12
Regularly drive	383 (95.3)	173 (93.5)	210 (96.8)	
Occasionally drive	16 (4.0)	9 (4.9)	7 (3.2)	
Rarely drive	0 (0.0)	0 (0.0)	0 (0.0)	
Do not drive	3 (0.7)	3 (1.6)	0 (0.0)	
Driving frequency, mean (SD)				
Driving frequency self-score	4.4 (0.7)	4.4 (0.8)	4.5 (0.6)	.28
Cognition, mean (SD)				

XSL•FO
RenderX

Measure	Total, N=402	Dropout, n=185	Adherent, n=217	*P* value
Orientation score, 0 - 4	7.8 (0.4)	7.8 (0.4)	7.8 (0.4)	.42
Immediate recall score, 0 - 4	4.0 (0.1)	4.0 (0.1)	4.0 (0.1)	.53
Delayed recall score, 0 - 4	3.8 (0.6)	3.7 (0.7)	3.9 (0.5)	.004
Memory compared to others their age, n (%)				.30
Worse	77 (19.2)	40 (21.6)	37 (17.1)	
Not worse	325 (80.8)	145 (78.4)	180 (82.9)	
Memory concern score, mean (SD)				
Memory concern self-score	1.8 (0.6)	1.8 (0.6)	1.8 (0.6)	.42
Memory concern, n (%)				.22
Extremely concerned	4 (1.0)	1 (0.5)	3 (1.4)	
Very concerned	32 (8.0)	20 (10.8)	12 (5.5)	
Some concern	253 (62.9)	113 (61.1)	140 (64.5)	
No concern	113 (28.1)	51 (27.6)	62 (28.6)	
Depression, mean (SD)				
Depression score, 0 - 64	12.7 (9.8)	13.0 (9.6)	12.5 (10.0)	.39
Quality of life, mean (SD)				
Quality of life self-score	79.0 (16.5)	76.5 (17.2)	81.1 (15.5)	.003
Future contact for research, n (%)				.40
Consented	387 (96.3)	176 (95.1)	211 (97.2)	.
Did not consent	15 (3.7)	9 (4.9)	6 (2.8)	
Total prescription medications, mean (SD)	3.4 (2.7)	3.9 (2.9)	2.9 (2.4)	.001
Prescription medication types, n (%)				
Acid suppression	92 (22.9)	49 (26.5)	43 (19.8)	.14
Cholesterol	181 (45.0)	86 (46.5)	95 (43.8)	.66
Diabetes	39 (9.7)	22 (11.9)	17 (7.8)	.23
Sleep aids	83 (20.6)	46 (24.9)	37 (17.1)	.07
Depression	78 (19.4)	47 (25.4)	31 (14.3)	.007
Anxiety	90 (22.4)	47 (25.4)	43 (19.8)	.22
Medical conditions, n (%)				
Diabetes	31 (7.7)	16 (8.6)	15 (6.9)	.64
High blood pressure	185 (46.0)	93 (50.3)	92 (42.4)	.14
High cholesterol	197 (49.0)	92 (49.7)	105 (48.4)	.87
Thyroid deficiency	80 (19.9)	42 (22.7)	38 (17.5)	.24
Cancer	65 (16.2)	28 (15.1)	37 (17.1)	.70
Alcohol abuse	19 (4.7)	11 (5.9)	8 (3.7)	.41
Anxiety	102 (25.4)	49 (26.5)	53 (24.4)	.72
Stroke	8 (2.0)	2 (1.1)	6 (2.8)	.40
B12 deficiency	27 (6.7)	13 (7.0)	14 (6.5)	.98
Sleep apnea	66 (16.4)	28 (15.1)	38 (17.5)	.61
Depression	121 (30.1)	63 (34.1)	58 (26.7)	.14

XSL•FO
RenderX

Measure	Total, N=402	Dropout, n=185	Adherent, n=217	*P* value
Concussion or TBI	10 (2.5)	6 (3.2)	4 (1.8)	.56
TIA	7 (1.7)	3 (1.6)	4 (1.8)	≥.99
Atrial fibrillation	21 (5.2)	10 (5.4)	11 (5.1)	≥.99
Neurological disease	13 (3.2)	8 (4.3)	5 (2.3)	.39
Heart attack	9 (2.2)	6 (3.2)	3 (1.4)	.36
Drug abuse	4 (1.0)	1 (0.5)	3 (1.4)	.73
Parkinson disease	0 (0.0)	0 (0.0)	0 (0.0%	≥.99

Our between-group comparison identified that there were many metrics in which there were no significant differences between dropout and adherent participants. These included basic demographic variables such as average age, distribution of participants within age quintiles, gender, race, ethnicity, average years of education, and highest level of education (*P*=.67, .52, .29, .27, .16, .16, and .35). Differences in marital statuses neared significance (*P*=.051) but were not significant. Dropouts and adherents had statistically similar frequency of living alone, housing situations, driving statuses, and self-rated driving frequency (*P*=.31, .74, .12, .28). Results from the Life Space mobility scale did not significantly differ between groups (*P*=.96). Next, we sought to compare the 2 groups in terms of technology use and comfort with technology. Self-rated comfort with computers and average comfort with computers score (*P*=.58 and *P*=.36), average number of technologies used (*P*=.17) as well as the frequency of smartphone, tablet, laptop, desktop, and wearable technology use (*P*=*P*≥.99, *P*=.17, *P*=.80, *P*=.38, *P*=.18) were not significantly different between the 2 groups. Cognitively, dropouts and adherents did not display significantly different performance on our orientation or immediate recall assessments (*P*=.42, .53). There was also not a significant difference in depressive symptoms between groups (*P*=.39). There were no significant between-group differences in whether participants rated their memory as worse than others their age and in their self-rated memory concern and average memory concern score (*P*=.30, .22, .42) Use of medications for acid suppression, cholesterol, diabetes, sleep aid, and anxiety did not differ significantly between groups (*P*=.14, .66, .23, .07, .22). Medical history of diabetes, high blood pressure, high cholesterol, thyroid deficiency, cancer, alcohol abuse, anxiety, stroke, B12 deficiency, sleep apnea, depression, concussion or traumatic brain injury, transient ischemic attack, atrial fibrillation, neurological disease, heart attack, drug abuse, and Parkinson disease did not differ significantly between groups (*P*=.64, .14, .87, .24, .70, .41, .72, .40, .98, .61, .14, .56, ≥.99, .≥99, .39, .36, .73, ≥.99). Interestingly, consent to be contacted for future research did not differ significantly between groups (*P*=.40).

The metrics in which we observed differences between dropout and adherent groups ranged across multiple domains of health. The dropout group had significantly lower self-rated mobility scores (*P*=.007), significantly lower delayed recall scores (*P*=.004), higher total prescription medications taken (*P*=.001), higher self-reported use of depression medication (*P*=.007), and lower self-rated quality of life (*P*=.003) as compared to the adherent group.

With initial findings from our between-group comparisons, we proceeded to the modeling stage of our analysis. We constructed 4 binary logistic regression models consisting of different sets of predictors chosen based on the results of the between-group tests. The footnote of Table 2 defines the predictor set of each model. Hyperparameters were tuned for maximum classification strength according to the protocols described in the Methods section above. Ultimately, Model 3, a model with a predictor set consisting of all significant between-group variables and basic demographic variables, was tied for the highest overall accuracy but exhibited more consistent performance across validation folds and better classification of the minority class. The classification metrics from the full model comparison are shown in Table 2.

XSL•FO
RenderX

Table . Binary logistic regression model performance in the cross-validation stage.

	Model 1[a]	Model 2[b]	Model 3[c]	Model 4[d]
Adherent classification metrics, mean (SD)				
Precision	0.59 (0.06)	0.61 (0.05)	0.61 (0.04)	0.57 (0.03)
Recall	0.64 (0.05)	0.76 (0.09)	0.73 (0.07)	0.7 (0.08)
F_1-score	0.61 (0.05)	0.68 (0.06)	0.67 (0.05)	0.63 (0.04)
Support	54.25 (0.5)	54.25 (0.5)	54.25 (0.5)	54.25 (0.5)
Dropout classification metrics, mean (SD)				
Precision	0.53 (0.09)	0.61 (0.1)	0.6 (0.07)	0.53 (0.06)
Recall	0.48 (0.11)	0.44 (0.07)	0.46 (0.05)	0.39 (0.06)
F_1-score	0.5 (0.1)	0.51 (0.07)	0.52 (0.06)	0.45 (0.04)
Support	46.25 (0.5)	46.25 (0.5)	46.25 (0.5)	46.25 (0.5)
Classification metrics across classes, mean (SD)				
Unweighted average precision	0.56 (0.07)	0.61 (0.07)	0.6 (0.06)	0.55 (0.04)
Unweighted average recall	0.56 (0.07)	0.6 (0.06)	0.6 (0.05)	0.55 (0.03)
Unweighted average F_1-score	0.56 (0.08)	0.59 (0.07)	0.59 (0.05)	0.54 (0.03)
Unweighted average support	100.5 (0.58)	100.5 (0.58)	100.5 (0.58)	100.5 (0.58)
Weighted average precision	0.56 (0.07)	0.61 (0.07)	0.61 (0.05)	0.55 (0.04)
Weighted average recall	0.56 (0.07)	0.61 (0.07)	0.61 (0.05)	0.56 (0.04)
Weighted average F_1-score	0.56 (0.07)	0.6 (0.06)	0.6 (0.05)	0.55 (0.03)
Weighted average support	100.5 (0.58)	100.5 (0.58)	100.5 (0.58)	100.5 (0.58)
Overall model accuracy, mean (SD)				
Accuracy	0.56 (0.07)	0.61 (0.07)	0.61 (0.05)	0.56 (0.04)

[a]Model with predictor set consisting of all nonredundant metrics collected, L1 regularization with convex optimization algorithm, regularization strength of 0.01, size trim tolerance of 0.1 for coefficients, convergence tolerance of 1×10^{-10}, 100 iterations.

[b]Model with predictor set consisting of only variables significant in between-group comparison, Newton-Raphson root-finding algorithm for optimization, convergence tolerance of 1×10^{-10}, 100 iterations.

[c]Model with predictor set consisting of variables significant in between-group comparison and basic demographic variables, L1 regularization with convex optimization algorithm, regularization strength of 0.01, size trim tolerance of 0.05 for coefficients, convergence tolerance of 1×10^{-10}, 100 iterations.

[d]Model with predictor set consisting of variables significant in between-group comparison and basic demographic variables as well as all second-degree interactions, L1 regularization with convex optimization algorithm, regularization strength of 0.01, size trim tolerance of 0.1 for coefficients, convergence tolerance of 1×10^{-10}, 100 iterations.

Our logistic regression model, based primarily upon these significantly different between-group findings, showed several predictors with varying influence on the odds of the dropout (Table 3). Older age, identifying as female, having a Black or African American or Asian racial background, and the use of depression medication were all found to be associated with slightly increased odds, but the CIs for each encompass 1, indicating limited significance. Having a higher delayed recall was the most notable of the significant associations we found (OR 0.77, 95% CI 0.61-0.96), suggesting that higher delayed recall ability might reduce the odds of dropping out. Additionally, taking more prescription medications shows a significant association with an increase in the odds of dropping out (OR 1.3, 95% CI 1.01-1.68). Identifying as an ethnicity other than Hispanic or Latino or non-Hispanic was also shown to reduce dropout odds (OR 0.78, 95% CI 0.62-0.97). Overall, while some predictors hint at associations, few reach strong statistical significance.

XSL•FO
RenderX

Table . Binary logistic regression model results.

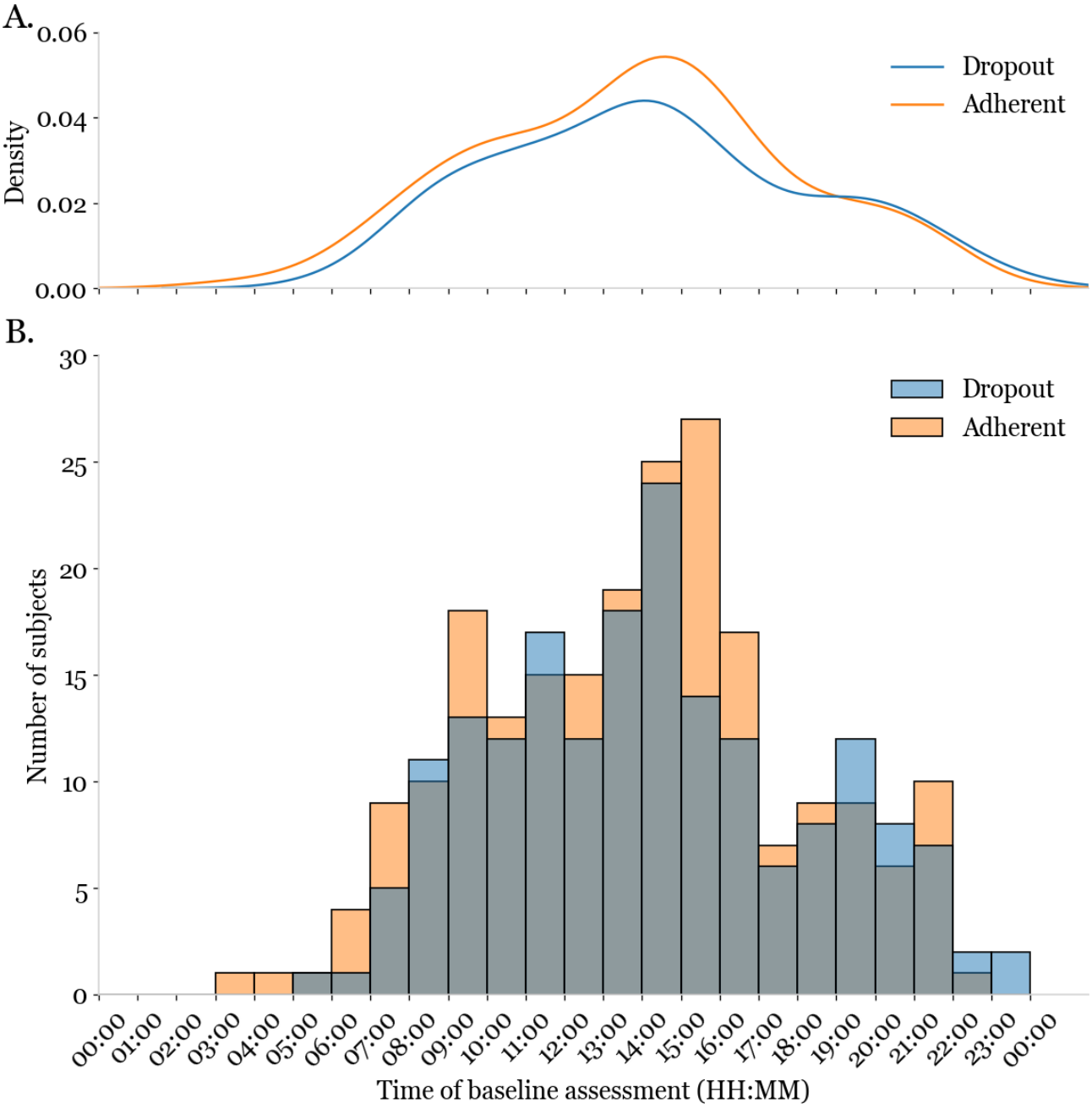
Predictor	Coefficient	Odds ratio (95% CI)
Constant term	−0.18	0.83 (0.62-1.12)
Age	0.13	1.14 (0.9-1.44)
Gender		
Women	0.13	1.14 (0.91-1.43)
Prefer to self-describe	−0.34	0.71 (0.01-51.46)
Race		
Black	0.17	1.18 (0.95-1.47)
Asian	0.2	1.22 (0.98-1.52)
Other	0.06	1.06 (0.84-1.34)
Two or more races	Trimmed	N/A[a]
Ethnicity		
Hispanic or Latino	−0.05	0.95 (0.77-1.17)
Other	−0.25	0.78 (0.62-0.97)
Years of education	−0.07	0.93 (0.75-1.15)
Self-rated mobility score	−0.14	0.87 (0.69-1.11)
Delayed recall score	−0.27	0.77 (0.61-0.96)
Total prescription medications	0.26	1.3 (1.01-1.68)
Use of depression medication	0.17	1.19 (0.95-1.49)
Quality of life	−0.2	0.82 (0.65-1.03)

[a]Not available.

As part of our feasibility efforts, we next sought to determine whether there were significant differences between the time of the day that adherents and dropouts completed their baseline web-LABrainS battery. We observed that dropouts and adherents largely completed their respective assessment batteries at the same times of day (Figure 1). The mean time of day at which all participants completed their assessments was 2:00 PM (SD 4 hours) and the distributions of the times of day at which dropouts and adherents completed their assessments did not differ significantly between groups (Table 1), an indication of the convenience of an online assessment tool that is always available to the user. Interestingly, none of the 402 participants took their assessment battery between midnight and 3:00 AM of the participant's local time, while all other times of day had some amount of representation.

XSL•FO
RenderX

Figure 1. Kernel density estimation function and histogram for web-based Louisiana Aging Brain Study assessment start times. (A) Density plots showing the temporal distribution of baseline assessments for dropout (blue) and adherent (orange) participants. (B) The number of participants who completed their baseline assessment within each hour-long bin in a 24-hour day is depicted in histogram form. Distributions for both dropouts (blue) and adherents (orange) were displayed together at partial opacity, with overlapping parts of the distributions shown in gray. All times were reported relative to each participant's time zone.



Discussion

Principal Findings

Our results in this study support our initial hypothesis that the use of a self-guided, remote, and web-based battery can be used to conduct longitudinal research in community-dwelling OAs. We identified that demographic characteristics such as age, education, gender, marital status, and living alone were not significantly correlated with participant dropout. Similarly, we observed that the reported comfort level with using technology, the amount of technology use, interest in research, and the time of assessment were also not correlated with participant dropout. Conversely, we identified a diverse array of medical-related issues that were significantly correlated with study dropout, including use of depression medication, decreased mobility, and impairments in delayed recall performance, all of which were significantly correlated with study dropout. The importance of each of these findings is discussed below.

While the level of dropouts in this study was high, it is in line with what has been reported by in-clinic longitudinal studies. It is important to point out that recruitment of participants to web-LABrainS in this feasibility study did not involve efforts to target specific populations or efforts to do large-scale recruitment, being that recruitment was largely restricted to word-of-mouth among participants and a byproduct of correspondence with our institute. Additionally, it should be

XSL•FO
RenderX

pointed out that participants received 2 reminder emails to conduct their 6- or 12-month assessment batteries and received limited or no additional follow-up (or forms of follow-up) to maintain participant retention in this feasibility study. Given the extremely restricted outreach and retention efforts in this study, we believe that even though our feasibility effort went out to only 12 months, it shows that this approach is a feasible effort to have community-dwelling OAs maintain participation in a longitudinal research effort that involves repeated assessments over time.

Many in-clinic longitudinal studies have investigated which of the variables collected in their demographic surveys, health questionnaires, and physical or cognitive assessments are significant predictors of participant attrition. One of the most commonly identified variables reported in the literature is age, although some studies claim older age makes a subject more prone to attrition [12,14,18,19,21] while others suggest younger subjects have a greater risk of study dropout [7-9,23]. Past research also shows variability in other predictors of study dropout including gender and education. Several studies identify being male as a significant predictor of study dropout [8,12,21], while others have shown that being female increases the risk for study dropout [16,19]. Likewise, past research has found both low levels of education [7,8,10,15,16,20,22,23] and high levels of education [21] to be significant predictors of study dropout. Taken together, data from this study and existing literature suggest that the retention of participants in longitudinal studies may be impacted by some factors (age, gender, and education) that are study-specific and therefore challenging to control for or successfully address.

Other predictors of participant attrition did not vary in directionality. These variables associated with participant dropout included having lower socioeconomic or job status, lower cognitive function, and lower self-reported or general health [7,10,12,15,19,21,22]. Medical conditions such as chronic stress, cardiovascular disease, coronary artery disease, heart attack, respiratory dysfunction, diabetes, depression, and anxiety were commonly identified predictors of study dropout when surveyed for or directly assessed [9-11,19,22]. Finally, lifestyle factors such as smoking, low physical activity, and low social and community involvement were also commonly associated with study dropout [7,10,12,21]. Living situations played an interesting role in the risk of attrition in previous studies. The presence of familial conflict or young children was identified as a significant predictor of study dropout [9]. Living alone or having fewer potential caretakers also significantly predicted attrition [12,18,19]. Based on the consistency of these findings, future studies should set out to fully address the potential impact of each of these aspects on participant retention. For example, future studies (including web-LABrainS 2.0) should consider incorporating strategies that minimize the negative impacts of lower socioeconomic status (providing participant stipends and maximizing flexibility for when study visits occur), impaired cognition (consider incorporating study partner and simplifying assessment screens), and poor health (allow study staff to assist with evaluation if requested, provide alternative types of assessments for individuals with physical limitations).

Our studies identified several factors associated with dropout in this web-LABrainS feasibility effort and the findings from studies using traditional longitudinal study approaches. For example, we observed that dropout in the web-LABrainS effort was not correlated with age, education, gender, marital status, or living alone. In contrast, dropouts in web-LABrainS had significantly higher levels of depression, decreased self-reported mobility, and decreased memory performance (delayed recall). In addition, we identified 2 novel factors—decreased quality of life and increased number of prescribed medications—that were associated with study dropout. Interestingly, in this web-LABrainS feasibility research effort, no differences in technology use, comfort with technology, or time of assessment were observed between adherent participants and dropouts. Another surprising finding in this study was the finding that there was an extremely high percentage (greater than 95%) of both adherent participants and dropout participants in web-LABrainS who confirmed a willingness to be contacted about future research studies. These data suggest that alternative factors, including the perceived value of participation in the specific study, play a role in participant retention. For example, it is likely that some participants continue in a longitudinal study solely based on personal reasons that have not been discussed above (family member impacted by the disease or dedication to a university, clinician, or researcher). Alternatively, participants may differentially stay in a longitudinal study based on their perceived value of the study results they receive. It should be noted that for web-LABrainS, we routinely receive participant correspondence saying they value the opportunity to participate because of a family member who had dementia or because they highly value the study summaries they receive at the conclusion of each assessment.

Limitations

There are multiple limitations in this feasibility study. For example, our web-LABrainS participants are not a representative sample of the community, and in future research efforts, it will be critical to optimize outreach such that the recruitment of a representative community sample is reliably obtained. With a more representative sample and larger sample size, we could also expect predictive modeling for risk of dropout to improve in accuracy from the binary logistic regression model shared in this study. Additionally, there was an expected elevated dropout in this web-LABrainS feasibility study due to the extremely limited nature of participant retention efforts (2 email reminders). In future studies, it will be key to identify and implement features of our platform to maximize the retention of study participants. We believe, based on the incidental feedback from web-LABrainS participants over the course of this feasibility study, that the major reason for study participation is the fact that each participant gets a detailed and automated report on their cognitive function at the conclusion of the web-LABrainS assessment. Continuing to optimize feedback and related collaterals to study participants will likely be critical to ensuring the long-term retention of study participants for years or even decades in longitudinal research efforts.

XSL•FO
RenderX

Conclusions

Taken together, the findings in this study provide a framework for the design and implementation of longitudinal research studies that focus on the use of a self-guided, remote, and web-based approach. Additionally, the data in this study provide important findings on how different demographic, health, and technology use aspects correlate with adherence or dropout in the web-LABrainS approach. In addition to being feasible, we believe using this new paradigm for longitudinal research will significantly decrease participant burden, dramatically lower research costs, and allow for significantly greater recruitment of participants across a wide geographic area. In future studies, it will be critical to develop methodologies and approaches that facilitate participation by a more racially and socioeconomically diverse array of study participants. These efforts will likely include outreach on social media platforms and collaborations with stakeholders in the community that align with racially and socioeconomically diverse populations. It will likely be important in these future efforts to make the outreach and onboarding of participants as frictionless as possible for participants by optimizing the user interface and user experience. In the interest of improving the experience of the community stakeholders involved in the future research effort, optimizing the application programming interface of the research platform will be a high priority. The impact of these efforts will have a significantly greater clinical impact if they can begin to be incorporated into clinical care as screening, patient engagement, and caregiver support tools. Our future efforts will attempt to incorporate each of these aspects as well as develop pathways for web-LABrainS participants to be connected quickly with study staff and health care professionals when mild cognitive impairment or dementia is identified as part of their longitudinal assessment.

Acknowledgments

The authors would like to thank the participants of web-LABrainS for their participation in this feasibility study. We thank John Ruth and Aimee Stewart for retrieving the time-of-day data from the Web-LABrainS assessment tool. Finally, we appreciate Robbie Beyl for his statistics consultation.

Funding

We would like to thank the Keller-Lamar Health Foundation for providing the funding to validate the web-LABrainS platform as part of a previous research effort [39].

Conflicts of Interest

None declared.

References

1. Newman AB. An overview of the design, implementation, and analyses of longitudinal studies on aging. J Am Geriatr Soc 2010 Oct;58 Suppl 2(Suppl 2):S287-S291. [doi: 10.1111/j.1532-5415.2010.02916.x] [Medline: 21029055]
2. Marengoni A, Angleman S, Melis R, et al. Aging with multimorbidity: a systematic review of the literature. Ageing Res Rev 2011 Sep;10(4):430-439. [doi: 10.1016/j.arr.2011.03.003] [Medline: 21402176]
3. Bektas A, Schurman SH, Sen R, Ferrucci L. Aging, inflammation and the environment. Exp Gerontol 2018 May;105:10-18. [doi: 10.1016/j.exger.2017.12.015] [Medline: 29275161]
4. Dawber TR, Kannel WB. The Framingham study. An epidemiological approach to coronary heart disease. Circulation 1966 Oct;34(4):553-555. [doi: 10.1161/01.cir.34.4.553] [Medline: 5921755]
5. Wild K, Howieson D, Webbe F, Seelye A, Kaye J. Status of computerized cognitive testing in aging: a systematic review. Alzheimers Dement 2008 Nov;4(6):428-437. [doi: 10.1016/j.jalz.2008.07.003] [Medline: 19012868]
6. Deeg DJH, van Tilburg T, Smit JH, de Leeuw ED. Attrition in the Longitudinal Aging Study Amsterdam. The effect of differential inclusion in side studies. J Clin Epidemiol 2002 Apr;55(4):319-328. [doi: 10.1016/s0895-4356(01)00475-9] [Medline: 11927198]
7. Young AF, Powers JR, Bell SL. Attrition in longitudinal studies: who do you lose? Aust N Z J Public Health 2006 Aug;30(4):353-361. [doi: 10.1111/j.1467-842x.2006.tb00849.x] [Medline: 16956166]
8. Nguyen T, Thomas AJ, Kerr P, et al. Recruiting and retaining community-based participants in a COVID-19 longitudinal cohort and social networks study: lessons from Victoria, Australia. BMC Med Res Methodol 2023 Feb 27;23(1):54. [doi: 10.1186/s12874-023-01874-z] [Medline: 36849927]
9. Ng SK, Scott R, Scuffham PA. Contactable non-responders show different characteristics compared to lost to follow-up participants: insights from an Australian Longitudinal Birth Cohort Study. Matern Child Health J 2016 Jul;20(7):1472-1484. [doi: 10.1007/s10995-016-1946-8] [Medline: 26976281]
10. Katsuno N, Li PZ, Bourbeau J, et al. Factors associated with attrition in a longitudinal cohort of older adults in the community. Chronic Obstr Pulm Dis 2023 Apr 27;10(2):178-189. [doi: 10.15326/jcopdf.2022.0380] [Medline: 37099700]
11. Huguet N, Kaufmann J, O'Malley J, et al. Using electronic health records in longitudinal studies: estimating patient attrition. Med Care 2020 Jun;58 Suppl 6 Suppl 1(Suppl 6 1):S46-S52. [doi: 10.1097/MLR.0000000000001298] [Medline: 32412953]

XSL•FO
RenderX

12. Jacobsen E, Ran X, Liu A, Chang CCH, Ganguli M. Predictors of attrition in a longitudinal population-based study of aging. Int Psychogeriatr 2021 Aug;33(8):767-778. [doi: 10.1017/S1041610220000447] [Medline: 32301414]
13. Davies K, Kingston A, Robinson L, et al. Improving retention of very old participants in longitudinal research: experiences from the Newcastle 85+ study. PLoS One 2014;9(10):e108370. [doi: 10.1371/journal.pone.0108370] [Medline: 25302500]
14. Chatfield MD, Brayne CE, Matthews FE. A systematic literature review of attrition between waves in longitudinal studies in the elderly shows a consistent pattern of dropout between differing studies. J Clin Epidemiol 2005 Jan;58(1):13-19. [doi: 10.1016/j.jclinepi.2004.05.006] [Medline: 15649666]
15. Brilleman SL, Pachana NA, Dobson AJ. The impact of attrition on the representativeness of cohort studies of older people. BMC Med Res Methodol 2010 Aug 5;10:71. [doi: 10.1186/1471-2288-10-71] [Medline: 20687909]
16. Van Beijsterveldt CEM, van Boxtel MPJ, Bosma H, Houx PJ, Buntinx F, Jolles J. Predictors of attrition in a longitudinal cognitive aging study: the Maastricht Aging Study (MAAS). J Clin Epidemiol 2002 Mar;55(3):216-223. [doi: 10.1016/s0895-4356(01)00473-5] [Medline: 11864790]
17. Helliwell B, Aylesworth R, McDowell I, Baumgarten M, Sykes E. Correlates of nonparticipation in the Canadian Study of Health and Aging. Int Psychogeriatr 2001;13 Supp 1(S1):49-56. [doi: 10.1017/s1041610202007986] [Medline: 11892974]
18. Zunzunegui MV, Béland F, Gutiérrez-Cuadra P. Loss to follow-up in a longitudinal study on aging in Spain. J Clin Epidemiol 2001 May;54(5):501-510. [doi: 10.1016/s0895-4356(00)00325-5] [Medline: 11337214]
19. Dapp U, Anders J, von Renteln-Kruse W, Golgert S, Meier-Baumgartner HP, Minder CE. The Longitudinal Urban Cohort Ageing Study (LUCAS): study protocol and participation in the first decade. BMC Geriatr 2012 Jul 9;12(1):35. [doi: 10.1186/1471-2318-12-35] [Medline: 22776679]
20. Matthews FE, Chatfield M, Freeman C, McCracken C, Brayne C, CFAS M. Attrition and bias in the MRC cognitive function and ageing study: an epidemiological investigation. BMC Public Health 2004 Apr 27;4(1):12. [doi: 10.1186/1471-2458-4-12] [Medline: 15113437]
21. Mein G, Johal S, Grant RL, Seale C, Ashcroft R, Tinker A. Predictors of two forms of attrition in a longitudinal health study involving ageing participants: an analysis based on the Whitehall II study. BMC Med Res Methodol 2012 Oct 29;12(1):164. [doi: 10.1186/1471-2288-12-164] [Medline: 23106792]
22. Jacomb PA, Jorm AF, Korten AE, Christensen H, Henderson AS. Predictors of refusal to participate: a longitudinal health survey of the elderly in Australia. BMC Public Health 2002;2(1):4. [doi: 10.1186/1471-2458-2-4] [Medline: 11914148]
23. Hübner S, Haijen E, Kaelen M, Carhart-Harris RL, Kettner H. Turn on, tune in, and drop out: predictors of attrition in a prospective observational cohort study on psychedelic use. J Med Internet Res 2021 Jul 28;23(7):e25973. [doi: 10.2196/25973] [Medline: 34319246]
24. Goetz ME, Hanfelt JJ, John SE, et al. Rationale and design of the Emory Healthy Aging and Emory Healthy Brain Studies. Neuroepidemiology 2019;53(3-4):187-200. [doi: 10.1159/000501856] [Medline: 31454799]
25. Fitzgerald D, Hockey R, Jones M, Mishra G, Waller M, Dobson A. Use of online or paper surveys by Australian women: longitudinal study of users, devices, and cohort retention. J Med Internet Res 2019 Mar 14;21(3):e10672. [doi: 10.2196/10672] [Medline: 30869647]
26. Ashford MT, Jin C, Neuhaus J, et al. Participant completion of longitudinal assessments in an online cognitive aging registry: The role of medical conditions. Alzheimers Dement (N Y) 2024 Jan;10(1):e12438. [doi: 10.1002/trc2.12438] [Medline: 38188606]
27. Weiner MW, Nosheny R, Camacho M, et al. The Brain Health Registry: an internet-based platform for recruitment, assessment, and longitudinal monitoring of participants for neuroscience studies. Alzheimers Dement 2018 Aug;14(8):1063-1076. [doi: 10.1016/j.jalz.2018.02.021] [Medline: 29754989]
28. Laidra K, Reile R, Havik M, et al. Estonian National Mental Health Study: design and methods for a registry-linked longitudinal survey. Brain Behav 2023 Aug;13(8):e3106. [doi: 10.1002/brb3.3106] [Medline: 37278143]
29. Ruano L, Sousa A, Severo M, et al. Development of a self-administered web-based test for longitudinal cognitive assessment. Sci Rep 2016 Jan 8;6:19114. [doi: 10.1038/srep19114] [Medline: 26743329]
30. Staguhn ED, Kirkhart T, Allen L, Campbell CM, Wegener ST, Castillo RC. Predictors of participation in online self-management programs: a longitudinal observational study. Rehabil Psychol 2024 May;69(2):102-109. [doi: 10.1037/rep0000521] [Medline: 37956087]
31. Chen Y, Ji H, Shen Y, Liu D. Chronic disease and multimorbidity in the Chinese older adults' population and their impact on daily living ability: a cross-sectional study of the Chinese Longitudinal Healthy Longevity Survey (CLHLS). Arch Public Health 2024 Feb 1;82(1):17. [doi: 10.1186/s13690-024-01243-2] [Medline: 38303089]
32. Testad I, Aakre JA, Gjestsen MT, et al. Web-based assessment of cognition and dementia risk factors in over 3000 Norwegian adults aged 50 years and older: cross-sectional PROTECT Norge Study. JMIR Aging 2025 Aug 25;8:e69867. [doi: 10.2196/69867] [Medline: 40854096]
33. Erb MK, Calcagno N, Brown R, et al. Longitudinal comparison of the self-administered ALSFRS-RSE and ALSFRS-R as functional outcome measures in ALS. Amyotroph Lateral Scler Frontotemporal Degener 2024 Aug;25(5-6):570-580. [doi: 10.1080/21678421.2024.2322549] [Medline: 38501453]

XSL•FO
RenderX

34. Taquet M, Skorniewska Z, De Deyn T, et al. Cognitive and psychiatric symptom trajectories 2-3 years after hospital admission for COVID-19: a longitudinal, prospective cohort study in the UK. Lancet Psychiatry 2024 Sep;11(9):696-708. [doi: 10.1016/S2215-0366(24)00214-1] [Medline: 39096931]
35. Yu T, Chen J, Gu NY, Hay JW, Gong CL. Predicting panel attrition in longitudinal HRQoL surveys during the COVID-19 pandemic in the US. Health Qual Life Outcomes 2022 Jul 6;20(1):104. [doi: 10.1186/s12955-022-02015-8] [Medline: 35794553]
36. Hansen T, Nilsen TS, Yu B, et al. Locked and lonely? A longitudinal assessment of loneliness before and during the COVID-19 pandemic in Norway. Scand J Public Health 2021 Nov;49(7):766-773. [doi: 10.1177/1403494821993711] [Medline: 33645336]
37. Ramiz L, Contrand B, Rojas Castro MY, et al. A longitudinal study of mental health before and during COVID-19 lockdown in the French population. Global Health 2021 Mar 22;17(1):29. [doi: 10.1186/s12992-021-00682-8] [Medline: 33752717]
38. Yu E, Hagens S. Socioeconomic disparities in the demand for and use of virtual visits among senior adults during the COVID-19 pandemic: cross-sectional study. JMIR Aging 2022 Mar 22;5(1):e35221. [doi: 10.2196/35221] [Medline: 35134746]
39. Calamia M, Weitzner DS, De Vito AN, Bernstein JPK, Allen R, Keller JN. Feasibility and validation of a web-based platform for the self-administered patient collection of demographics, health status, anxiety, depression, and cognition in community dwelling elderly. PLoS One 2021;16(1):e0244962. [doi: 10.1371/journal.pone.0244962] [Medline: 33465108]
40. Peel C, Sawyer Baker P, Roth DL, Brown CJ, Brodner EV, Allman RM. Assessing mobility in older adults: the UAB Study of Aging Life-Space Assessment. Phys Ther 2005 Oct;85(10):1008-1119. [doi: 10.1093/ptj/85.10.1008] [Medline: 16180950]
41. Rossum G, Drake FL. Python 3 Reference Manual: CreateSpace; 2009.
42. Harris CR, Millman KJ, van der Walt SJ, et al. Array programming with NumPy. Nature New Biol 2020 Sep;585(7825):357-362. [doi: 10.1038/s41586-020-2649-2] [Medline: 32939066]
43. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods 2020 Mar;17(3):261-272. [doi: 10.1038/s41592-019-0686-2] [Medline: 32015543]
44. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. arXiV. Preprint posted online on Jun 5, 2018. [doi: 10.48550/arXiv.1201.0490]
45. Seabold S, Perktold J. Statsmodels: econometric and statistical modeling with Python. Presented at: 9th Python in Science Conference; Jun 28 to Jul 3, 2010; Austin, Texas, USA p. 57-61. [doi: 10.25080/Majora-92bf1922-011]
46. McKinney W. Data structures for statistical computing in Python. Presented at: Proceedings of the 9th Python in Science Conference; Jun 28 to Jul 3, 2010; Austin, Texas, USA p. 56-61. [doi: 10.25080/Majora-92bf1922-00a]
47. Hunter JD. Matplotlib: a 2D graphics environment. Comput Sci Eng 2007;9(3):90-95. [doi: 10.1109/MCSE.2007.55]
48. Waskom ML. seaborn: statistical data visualization. J Open Source Softw 2021;6(60):3021. [doi: 10.21105/joss.03021]

Abbreviations

IRB: institutional review board
Lasso: Least Absolute Shrinkage and Selection Operator
OR: odds ratio
PBRC: Pennington Biomedical Research Center
web-LABrainS: web-based Louisiana Aging Brain Study

Edited by D Liu; submitted 14.11.24; peer-reviewed by L Bovsh, M Moore, WT Au-Yeung; revised version received 21.09.25; accepted 20.10.25; published 13.11.25.

Please cite as:
Braun LD, Allen HR, Keller JN
Risk Factors for Community-Dwelling Older Adults Dropping Out of Self-Guided, Remote, and Web-Based Longitudinal Research: Predictive Modeling of Data from the Web-LABrainS Platform
JMIR Aging 2025;8:e68826
URL: https://aging.jmir.org/2025/1/e68826
doi:10.2196/68826

© Luke Daniel Braun, H Raymond Allen, Jeffrey N Keller. Originally published in JMIR Aging (https://aging.jmir.org), 13.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Multimodal Detection of Agitation in People With Dementia in Clinical Settings: Observational Pilot Study

Abeer Badawi[1*], PhD; Somayya Elmoghazy[1*], MSc; Samira Choudhury[2,3], MSc, MD; Sara Elgazzar[4], BSc; Khalid Elgazzar[1,5], PhD; Amer M Burhan[2,6], MBChB, MSc

[1]Internet of Things (IoT) Research Lab, Department of Electrical and Computer Engineering, Ontario Tech University, 2000 Simcoe Street North, Oshawa, ON, Canada

[2]Ontario Shores Centre for Mental Health Sciences, Whitby, ON, Canada

[3]Department of Psychiatry, Western University, London, ON, Canada

[4]Faculty of Science, Ontario Tech University, Oshawa, ON, Canada

[5]Canadian University of Dubai, Dubai, United Arab Emirates

[6]Temerty Faculty of Medicine, University of Toronto, Toronto, Canada

[*]these authors contributed equally

Corresponding Author:
Somayya Elmoghazy, MSc
Internet of Things (IoT) Research Lab, Department of Electrical and Computer Engineering, Ontario Tech University, 2000 Simcoe Street North, Oshawa, ON, Canada

Abstract

Background: Dementia is a progressive neurodegenerative condition that affects millions worldwide, often accompanied by agitation and aggression (AA), which contribute to patient distress and increased health care burden. Existing assessment methods for AA rely heavily on caregiver reporting, introducing subjectivity and inconsistency.

Objective: This study proposes a novel, multimodal system for predicting AA episodes in individuals with severe dementia, integrating wearable sensor data and privacy-preserving video analytics.

Methods: A pilot study involving 10 participants was conducted at Ontario Shores Mental Health Institute. The system combines digital biomarkers collected from the EmbracePlus (Empatica Inc) wristband with video-based behavioral monitoring. Facial features in video frames were anonymized using a masking tool, and a deep learning model was used for AA detection. To determine optimal performance, various machine learning and deep learning models were evaluated for both wearable and video data streams.

Results: The Extra Trees model achieved up to 99% accuracy for personalized wristband data, while the multilayer perceptron model performed best in general models with 98% accuracy. For video analysis, the gated recurrent unit model achieved 95% accuracy and 99% area under the curve, and the long short-term memory model demonstrated superior response time for real-time use. Importantly, the system predicted AA episodes at least 6 minutes in advance in all participants based on wearable data.

Conclusions: The findings demonstrate the system's potential to autonomously and accurately detect and predict AA events in real-time. This approach represents a significant advancement in the proactive management of behavioral symptoms in dementia care.

(JMIR Aging 2025;8:e68156) doi:10.2196/68156

KEYWORDS

dementia; agitation; artificial intelligence; wearable sensors; video detection; multimodal sensing

Introduction

Background

Dementia is a neurodegenerative condition that leads to a progressive decline in cognition and is one of the leading causes of death, disability, and hospitalization in Canada and worldwide. Currently, dementia is the seventh cause of death worldwide [1]. Worldwide, over 55 million individuals are living with dementia; as the ratio of older people increases, this number will grow to 78 million by 2030 and 139 million by 2050, making dementia a major global health crisis [1]. In addition to cognitive and functional decline, people living with dementia also experience noncognitive neuropsychiatric symptoms (NPS) during their illness [2]. NPS commonly includes agitation, aggression, apathy, symptoms of psychosis, delusions, hallucinations, and disturbances of sleep and appetite. Among NPS, agitation and aggression (AA) occur frequently

XSL•FO
RenderX

in severe cases and are a common source of distress for patients and caregivers [3]. They commonly occur during care and are believed to be manifestations of perceived or real unmet needs [3]. Behaviors of AA include pacing, rocking, gesturing, restlessness, shouting, throwing objects, and destroying property [4]. These symptoms are the leading cause of hospitalizations, extended length of stay of inpatients, and increased demand for placement in long-term care facilities [5]. AA enormously burdens people living with dementia, their families, caregivers, and health care systems.

In current practices, AA is commonly assessed through caregiver reports. Many observational methods have been developed, including the Neuropsychiatric Inventory [6] and the Cohen-Mansfield Agitation Inventory [7]. These assessments are based on manual observations, which are subject to potential bias depending on the caregiver's memory or emotional state. It is possible to address these limitations by using artificial intelligence (AI) and predictive algorithms to predict episodes of AA in people living with dementia. In 2025, AI technologies are expected to be worth an estimated US $36 billion [8]. There is growing evidence that combining AI and sensory technologies to develop a solution for NPS detection will guide the provision of personalized interventions for people living with dementia [9-11]. The timely detection of critical events in people living with dementia using digital technologies is gaining wide acceptance [12,13]. Predicting and managing AA in people living with dementia requires innovative approaches that integrate multiple data sources. Wearable sensors and video-based monitoring systems provide unique opportunities for the real-time detection of AA and preagitation behaviors, but challenges such as privacy concerns and scalability have limited their adoption in clinical settings. This study addresses these challenges by combining biometric data from wearable sensors with AI-driven video analysis, enabling real-time detection and prediction of AA. This integrated approach aims to facilitate timely interventions, reduce care costs, and improve outcomes for both patients and caregivers. We conducted a pilot study in the Geriatric Dementia Unit (GDU) and the Geriatric Transitional Unit at the Ontario Shores Center for Mental Health Sciences [14]. We combine biometric data from the EmbracePlus wristband [15] and video data from CCTV cameras installed in common areas in both units. Machine learning and deep learning techniques, including Extra Trees, Gradient Boosting, Random Forest, multilayer perceptron (MLP), and recurrent neural networks (RNNs), are used to analyze biometric and skeletal data extracted from both the wristband and the video cameras. We achieved high accuracy in detecting AA from the EmbracePlus wristband through comprehensive data preprocessing, feature extraction, the Extra Trees classification algorithm for personalized models, and the MLP algorithm for the general model. Additionally, AA detection was enhanced by analyzing real-time video feeds with skeletal key points and using RNN-based neural networks, particularly long short-term memory (LSTM) and gated recurrent unit (GRU) [16]. These networks, optimized for real-time processing, facilitate timely interventions. The pilot study demonstrated the system's effectiveness through both the wristband and video detection.

Related Work

The growing number of people living with dementia causes significant challenges for health care systems and caregivers. One of these challenges is to deal with symptoms of AA that increase with the severity of dementia. Recent advancements in multimodal sensing technologies, including cameras and wearable wristbands, have shown promise in monitoring and managing AA in people living with dementia [10,17,18]. Wearable devices, capable of capturing physiological signals such as acceleration, heart rate, and skin conductance, have demonstrated potential for real-time AA detection [19-21]. Recent advancements in wearable sensor technology and AI have shown promise in addressing the early detection of AA behavior when focusing on signal processing and machine learning to extract features and classify AA events [22].

However, real-time video-based monitoring systems to monitor AA behavior in dementia patients are an area of interest for researchers today [10,17,23]. These systems use cameras positioned in patients' rooms or common areas to consistently record and monitor patients' behavioral data. Some research [24,25] used video cameras to detect AA from previously recorded videos at the Specialized Dementia Unit, Toronto Rehabilitation Institute, Canada. Their system focused on offline AA detection. Another work collected the training dataset from healthy people who imitated agitated hand movements [26]. Researchers have highlighted the importance of real-time feedback, which allows for timely interventions and reduces the severity and duration of AA events. This can help improve the quality of life and reduce the stress caused by agitated behaviors in both patients and caregivers. Researchers have also considered privacy factors and concerns while ensuring the accuracy of these systems. The work done in studies by Mishra et al [24] and Marshall [26] has proposed different masking methods for the people in the video frames that allow for AA detection while keeping the patients' identities and features hidden.

Multimodal sensing, which combines data from multiple sensors, types, and sources, has emerged as a promising approach to understanding AA in people living with dementia [27,28]. It enhances the detection of early signs of AA and the identification of relevant triggers [23,29]. Clinical trials evaluating the use of multimodal sensing in the context of dementia and AA behavior are crucial in this type of research. These trials assess the feasibility and efficacy of monitoring systems that combine various data sources to inform clinical decision-making. The results of these trials will provide valuable insights into the practical implications of multimodal sensing in real-world health care settings. Most existing studies on video systems and wearable sensors for AA detection in real-time have been conducted in controlled laboratory settings or residential care facilities with limited datasets [18,30]. Research is needed to use more diverse and general datasets gathered and applied in real hospital settings. This is essential since the data from hospitals would be more representative of people with severe dementia who are more prone to AA. Research in such a real-world setting with data gathered from real patients is vital for developing realistic and applicable solutions and ensuring

XSL•FO
RenderX

effective treatment and care for people living with dementia [18,30].

Moreover, it is challenging to identify the actual start and end times of AA episodes. This is because the AA events are extracted from nurse notes, which are not accurate and prone to human or individual error. Many AA episodes may also be overlooked and mislabeled as nonagitation. In addition to the datasets, there is a high demand for accurate and reliable end-to-end real-time monitoring solutions to actively predict and respond to AA events. There is also still a need for an in-depth investigation of AI and its features in addition to an in-depth investigation of preagitation patterns and signs that could trigger AA in people living with dementia. These investigations are necessary to indicate the usefulness of preagitation signs and patterns in early prediction. Digital biomarkers can help detect and predict AA early in real time [31]. The investigation by Alam et al [32] shows the correlation between motion biomarkers collected from the accelerometer and the early AA signs, which is particularly useful for personalized AA detection models.

This paper presents a unique system that combines video analysis and wearable sensor data to predict AA in people living with dementia. The video feeds are crucial for identifying the precise start and end of each AA episode. This precise timing enables us to incorporate data analysis from both the wristband and video footage into our research. We focus on AI and advanced feature engineering to improve the detection accuracy of AA in people living with dementia. We significantly enhance our chances of detecting AA by using 2 distinct yet cooperative models (one based on physiological data and the other on video analysis). This integrated approach also opens the doors to incorporating additional predictive methods, such as audio-based AA detection. The combined use of video-based systems and wearable wristbands offers deeper insights into AA management. While notable advancements have been achieved, further enhancements are needed, especially in real-time accuracy and system applicability. We are confident that our current research explores a vital area and will contribute new knowledge to the field and lay a solid foundation for future advancements.

Methods

Ethical Considerations

The research commenced in 2019. This study was approved by the Joint Research Ethics Board (JREB) at Ontario Shores Centre for Mental Health Sciences and Ontario Tech University (JREB Number: 21-011-B). Informed consent was obtained from all participants' substitute decision makers, as the participants had advanced dementia and were unable to provide consent themselves. Data collected by the cameras were masked, with any identifiable objects in the video frames blurred (a video demonstration was provided to the research ethics board [REB] chair and privacy officer). All information obtained from participants was kept confidential. Computer-based data were stored in password-protected databases, and paper-based files were kept in locked cabinets. Access to all data were restricted to authorized study personnel, who followed the confidentiality regulations of the JREB. Participants were not compensated for their participation.

This study was approved by the JREB at Ontario Shores Centre for Mental Health Sciences and Ontario Tech University (JREB Number: 21-011-B). Informed consent was obtained from all participants' substitute decision-makers, as the participants had advanced dementia and were unable to provide consent themselves. Data collected by the cameras were masked, with any identifiable objects in the video frames blurred (a video demonstration was provided to the REB chair and privacy officer). All information obtained from participants was kept confidential. Computer-based data were stored in password-protected databases, and paper-based files were kept in locked cabinets. Access to all data was restricted to authorized study personnel, who followed the confidentiality regulations of the JREB. Participants were not compensated for their participation.

Study Design

A significant challenge was presented using video cameras to document the behaviors and activities of patients, staff, and visitors in the public spaces of hospital inpatient units. Consent was secured for these patients through substitute decision makers considering their advanced dementia condition. At admission, capacity assessments were performed, and substitute decision makers were approached for permission to involve patients in potential research. The recording of video cameras was restricted to the common areas where patients usually gather during the day, and audio capture was turned off throughout the data collection period.

This study collected participant data using an EmbracePlus wristband [15] and video cameras installed in the GDU and the Geriatric Transitional Unit at the Ontario Shores Center for Mental Health Sciences. A total of 10 participants were recruited based on inclusion criteria, including being aged 60 years or older, a diagnosis of moderate to severe major neurocognitive disorder as determined by the Mini-Mental State Examination [33], and the ability to ambulate independently with or without a walking aid. Additionally, participants had to meet the agitation criteria defined by the Agitation Definition Working Group of the International Psychogeriatric Association [34], with a Functional Assessment Staging Tool scale score between 6a and 6e [35]. Moreover, each unit was equipped with a single AXIS M3077-PLVE Network Camera [36], and one AXIS P3225-VE Mk II camera [37] was installed in the hallway of the GDU to capture relevant footage. Participants wore the EmbracePlus wristband for 24 to 72 hours on 3 separate occasions within a 6-week study period. The wristband collected physiological parameters such as heart rate, electrodermal activity, and skin temperature.

During data collection (Multimedia Appendix 1), clinical staff, who are part of the research team of this study, monitored participants for episodes of AA, noting the start and end times of each event. CCTV cameras recorded footage to provide precise timestamps and additional context, as there can be a delay between observed behavior and recorded notes. To maintain privacy, faces and identifiable features in the video footage were blurred. Once an AA event is identified by the

XSL•FO
RenderX

staff in the video, the skeletal points of the recruited participants are extracted and added to the dataset, and all other skeletal points are discarded. The skeletal key points were analyzed using deep learning models to classify episodes of AA. Data from wearable sensors and clinical notes were integrated and analyzed to identify physiological and behavioral patterns preceding episodes of AA. The collected data formed the basis for developing and evaluating our multimodal system for real-time AA detection. The results demonstrate the system's feasibility and effectiveness in detecting and analyzing AA episodes.

Event Classification

The proposed system collects the biomarkers using the EmbracePlus wristband, which is considered the state-of-the-art wearable device for continuous health monitoring in the market today [15]. The device combines digital biomarkers, robust sensors, and a user-friendly design to continuously monitor participants with various health conditions. It collects electrodermal data of detected slight changes in skin conductance from the skin surface, and the data of a photoplethysmogram that calculates the pulse rate and pulse rate variability measurements, skin temperature, and raw accelerometry data for motion detection. The collected signals are sent to the cloud-based EmbracePlus Care platform [15] through a Bluetooth-connected gateway (eg, a smartphone).

The first type of data we deal with is the raw data from the accelerometer, heart rate, temperature, and electrodermal signals. We follow several preprocessing steps to clean, filter, and apply 1-minute window segmentation for the raw signals [21,38]. We then extract features from the signals, as shown in our previous work, from the statistical, time domain, frequency domain, and time-frequency domain with around 150 features [21,38]. Lastly, we evaluate multiple classification techniques, namely Random Forest, Extra Trees, and Gradient Boosting, to classify AA events. The performance of each model is evaluated using standard classification metrics such as accuracy, precision, recall, area under the curve (AUC), and F_1-score. Furthermore, we collect the digital biomarkers that are preprocessed data derived from Empatica's algorithms and calculate them minute-by-minute. The second type is the digital biomarkers, which include pulse rate variability, respiratory rate, movement intensity, accelerometer magnitude SD, steps, skin conductance level, wearing detection, temperature, and sleep detection as shown in Table 1. Digital biomarkers can effectively and accurately oversee human health from a distance, consistently, and without causing disruption. This applies across a spectrum of health conditions [15]. Figure 1 shows the classification workflow from the EmbracePlus wristband using raw data and digital biomarkers.

Table . The digital biomarker data description from the EmbracePlus wristband.

Digital biomarkers	Definition
PR[a]	The algorithm uses the photoplethysmogram and accelerometer data for PR monitoring with estimates on 10-second windows.
PRV[b]	The algorithm analyzes the photoplethysmogram for intermittent PRV, using accelerometer signals with nonoverlapping windows.
RR[c]	The algorithm processes the photoplethysmogram and accelerometer data to calculate RR values expressed in breaths per minute.
Movement intensity	The algorithm calculates activity count, steps, and accelerometer SD from the accelerometer sensor.
SCL[d]	SCL estimation from EmbracePlus electrodermal activity signal, output every 1 minute with nonoverlapping windows.
Wearing detection	The algorithm correlates device status with the photoplethysmogram patterns, indicating wearing time.
Temperature	The algorithm analyzes EmbracePlus data for continuous peripheral temperature estimation.

[a]PR: pulse rate.

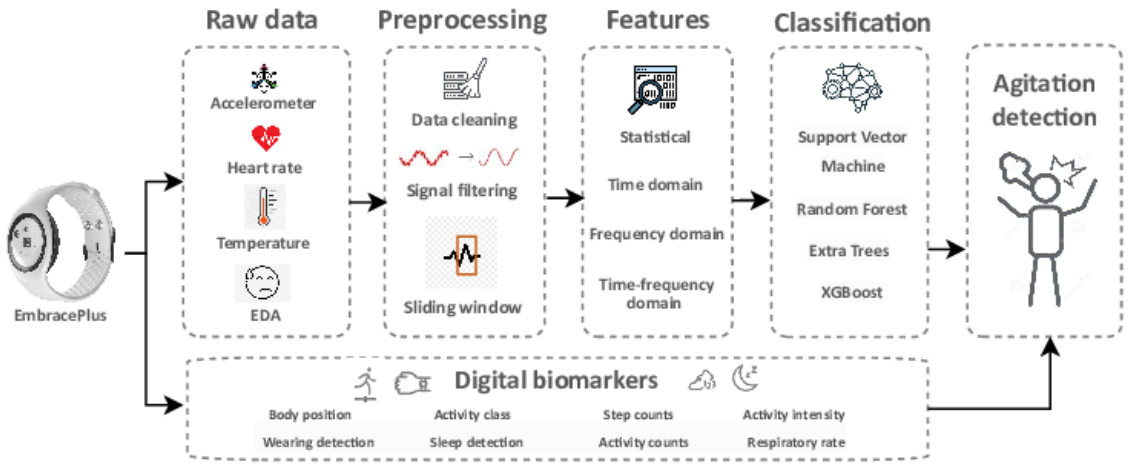
[b]PRV: pulse rate variability.

[c]RR: respiratory rate.

[d]SCL: skin conductance level.

XSL•FO
RenderX

Figure 1. The workflow of the proposed classification system using the EmbracePlus wristband. EDA: electrodermal activity.



The proposed work focuses on investigating the data from people living with dementia using machine learning, 2 of which were thoroughly investigated in our previous work [21,38]. The results concluded the most important features for this problem after performing feature engineering and proved that personalized models on individual patients outperform generic models. In this work, we report the results of the personalized model on 10 different participants from the Ontario Shores Mental Health Institute. We test our system in real-time once we determine the optimal classification system to predict AA events. In the real-time detection phase, real-time raw data is transmitted from the wristband. Following this, features are extracted from each 1-minute window, and these specific features are fed into the customized model to classify whether the data is considered normal or indicative of AA. The outcomes are subsequently transmitted to the backend system, and the health care provider is notified if the patient is agitated.

Video-Based Analysis

In addition to collecting data on physiological biomarkers, this study incorporates video analysis data for AA prediction. This approach uses an extra cooperative model, which improves our overall AA detection system and allows us to get the precise duration, including start and end times, of the collected AA episodes. Moreover, once an AA episode is detected, the cameras record a previously set preagitation, making it easier to observe any visual preagitation signs. We aim to provide real-time alerts to health care providers for timely intervention. The setup includes 3 CCTV cameras installed and a PC in the attending psychiatric office with access to this footage. Our system operates in 2 phases: the offline phase for manual labeling and model training, and the real-time stage for running the model. To protect the privacy of the participants and the staff present, we blur all faces and run OpenPose (Carnegie Mellon University) to capture the movement data of the participants [39-42]. OpenPose is an advanced real-time system for multiperson 2D pose estimation. It also helps to anonymize individuals in video frames. OpenPose uses convolutional neural networks to detect human body parts and map their skeletal structure onto the image or video frame. This allows for a detailed representation of movement data present in the collected frames. The model is trained on features extracted from skeletal point coordinates instead of the raw video frames. This approach has been recently used by researchers for AA detection in people living with dementia and has proven to be as successful in detecting AA while preserving the privacy of the people present [24,26].

After this, we use a preprocessing phase to enhance the generalizability of the model across various environments and datasets. This phase involves the elimination of extraneous noise that could otherwise impede model performance. The model considers the variations in camera angles and participant positioning within the frame, which can significantly influence the coordinate data. We calculate Euclidean distances and angle measurements between specific skeletal coordinates to determine movements. For example, the measured distance between the torso and feet is useful to identify potential kicking actions, which may indicate AA in certain contexts. Table 2 summarizes all 47 features extracted from these distance and angle measurements. Before training, a feature analysis step is introduced to remove highly correlated features and reduce the dimensionality of the model. This process reduced the features to 39 features. We tested the system on the same 10 participants whose wearable sensor data was used earlier.

XSL•FO
RenderX

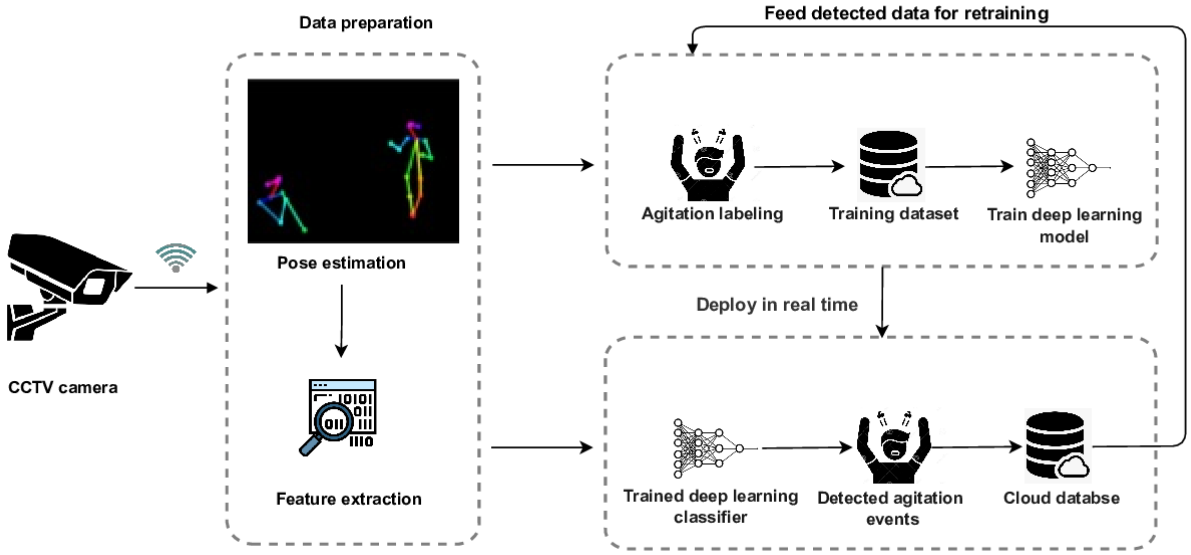
Table . Description of the extracted features from the skeletal data.

Feature	Description
eu_1-eu_14	Euclidean distances between different key point pairs; eu 1 represents the Euclidean distance between key point 1 and the previous position of key point 1
eu_1_3-eu_1_14	Euclidean distances between key point 1 and various other key points
por_2_1-por_14_1	POR[a] values between key points 2 - 14 and key point 1
ang_1_2-ang_1_14	Angles between key point 1 and key points 2 - 14

[a]POR: point of reference.

The offline and real-time (online) stages of the system are detailed in Figure 2. In the offline stage, we preprocess and extract features from skeletal data, which are then labeled using nurse notes from patient medical records. Access to the collected videos is restricted to the computer in the psychiatric office, and they are retained only until the AA episodes are accurately labeled with their start and end times. Once the dataset is finalized, all videos are securely discarded. Using this dataset, we train a deep learning model to differentiate between AA and nonagitation events. Our focus is on capturing a range of behaviors, from violent or aggressive actions to repetitive motions such as pacing or chair rocking. Hence, we use models that consider sequences of actions, such as RNN-based neural networks, to effectively recognize these sequences of actions. We specifically use and compare the results of the GRU and LSTM models. Both are adept at analyzing sequences of actions. The LSTM model is designed to capture long-term dependencies within sequences. The sophisticated cell structure of LSTM cells makes it highly effective in maintaining context over long intervals. However, the GRU model uses a simpler architecture that aims to achieve results comparable to LSTM models but with lower computational costs. Both models used in our research are composed of a single LSTM or GRU cell, followed by a fully connected sigmoid layer for the binary classification of AA episodes.

Figure 2. The system architecture of the video-based detection system.



In the real-time stage, we deploy our offline-trained model to classify AA in real-time. The model first processes real-time video data from hospital cameras. Video frames are analyzed using OpenPose, extracting features similar to those in the offline stage. We use a fixed-size window to input frame sequences into the classifier. Upon detecting AA, the system records the event, including a 5-minute buffer before and after the incident to capture the entire context. This approach helps identify potential triggers and patterns that lead to AA. The psychiatrist reviews these detected events for accuracy and confirms AA events, which are then added to our training dataset with appropriate labels. The model is then retrained in the offline stage with the added data. The primary goal of retraining the model is to continuously adapt and improve the model with newly detected data.

Results

Overview

A pilot study was conducted to validate the effectiveness and feasibility of the proposed system at Ontario Shores Mental Health Hospital. This initial investigation aimed to provide valuable insights into the system's functionality, usability, and overall potential before the implementation on a larger population. Details of the enrolled participants and the data

XSL•FO
RenderX

collected can be found in Table 3. Upon enrollment, the participants wore the EmbracePlus wristband for 24 to 72 hours. We turned on the cameras installed in the unit during the data collection days to record the participants' activities. Lastly, we assigned a nurse to observe the participant and provide a detailed report of behavior, AA events, and any abnormal behavior. During the 3 days, we collected 6 AA events, ranging from 2 to 23 minutes per AA event, with a total of 20 - 32 minutes of AA labels and 560 - 581 minutes of normal labels for each participant. The following sections will present the results from the EmbracePlus wristband and video cameras in detail.

Table . Overview of the demographic of the participants and the total collected data.

Participant	Gender	Age	Recruitment date	Total collected data (hours)
1	Female	83	August 2023	54.21
2	Female	63	January 2024	48.13
3	Female	77	March 2024	45.3
4	Female	78	March 2024	96.4
5	Female	79	March 2024	99.4
6	Male	79	May 2024	46.7
7	Male	67	August 2024	36
8	Male	68	August 2024	35.5
9	Male	85	September 2024	95.5
10	Male	86	September 2024	54.6

Performance Evaluation

The EmbracePlus Wristband Raw Data

This study uses raw data obtained from 4 wristband signals and personalized models, which achieved superior accuracy in AA detection from people living with dementia in previous research [21,38]. Subsequently, we conducted a comparative analysis of multiple machine learning and deep learning algorithms for AA detection, including Random Forest, Extra Trees, Gradient Boosting, and MLP. We trained and tested a personalized model for every participant and reported the evaluation results in Table 4. The dataset for each participant was randomly split into 70% for training and 30% for testing.

Table . Comparative performance metrics using raw data.

Participant	Extra Trees			Gradient Boosting			Random Forest			MLP[a]		
	Acc[b]	AUC[c]	Recall	Acc	AUC	Recall	Acc	AUC	Recall	Acc	AUC	Recall
1	0.98[d]	0.99[d]	0.99[d]	0.97	0.98	0.98	0.98[d]	0.99[d]	0.99[d]	0.97	0.98	0.98
2	0.90[d]	0.98[d]	0.96	0.88	0.96	0.93	0.88	0.96	0.94	0.88	0.96	0.97[d]
3	0.99[d]	0.99[d]	0.99[d]	0.98	0.99[d]	0.99	0.98	0.99[d]	0.98	0.99[d]	0.99[d]	0.99[d]
4	0.91[d]	0.99[d]	0.97[d]	0.89	0.97	0.96	0.88	0.97	0.99	0.89	0.96	0.98
5	0.99[d]	0.99[d]	0.99[d]	0.98	0.99[d]	0.98	0.98	0.99[d]	0.98	0.95	0.90	0.96
6	0.99[d]	0.99[d]	0.99[d]	0.98	0.99[d]	0.98	0.98	0.99[d]	0.98	0.99[d]	0.99[d]	0.99[d]
7	0.92[d]	0.98[d]	0.97	0.89	0.96	0.96	0.89	0.97	0.96	0.90	0.96	0.99[d]
8	0.91[d]	0.98[d]	0.97	0.91	0.96	0.96	0.88	0.97	0.95	0.89	0.96	0.98
9	0.99[d]	0.99[d]	0.99[d]	0.98	0.99[d]	0.98	0.98	0.99[d]	0.98	0.99[d]	0.99[d]	0.99[d]
10	0.99[d]	0.99[d]	0.99[d]	0.98	0.99[d]	0.99	0.98	0.99[d]	0.98	0.99[d]	0.99[d]	0.99[d]
All	0.85	0.94	0.93	0.98	0.96	0.90	0.93	0.95	0.65	0.99[d]	0.98[d]	0.98[d]

[a]MLP: multilayer perceptron.

[b]Acc: accuracy.

[c]AUC: area under the curve.

[d]Bold text indicates the best results for each participant in each metric.

XSL•FO
RenderX

The Extra Trees model outperformed the rest of the models for most of the participants, followed by the MLP model, which achieved similar results for 4 of the 10 participants. For example, the Extra Trees model achieved an accuracy of 98.67%, an AUC of 99.1%, and a recall of 99.76% for participant #1. It achieved the highest accuracy and AUC for participant #2 of 90% and 98%, respectively. For participants, #3, #6, #9, and #10, both the Extra Trees and MLP models achieved 99% across all evaluation metrics. These results underscore the efficacy of the chosen features, preprocessing methodologies, and up-sampling techniques. When tested on all 10 participants together, the Extra Trees model achieved a lower accuracy of 85%. The Gradient Boosting and Random Forest achieved higher accuracies in comparison with 98% and 93%, respectively. The Random Forest model, however, performed very poorly in other evaluation matrices, with a recall of 67% for the general model. The MLP model achieved a slightly higher accuracy than Gradient Boosting of 99% and even a higher AUC and recall of 98%, performing very similarly to the personalized models. This highlights the potential for a general model when enough data are collected.

Table 5 presents a summary of the statistical analysis of model performance metrics across all models. As expected, Extra Trees is the top-performing model for participant-level predictions, with the highest mean accuracy of 0.95 and a 95% CI of 0.928-0.986. This indicates consistent performance across the evaluation set. Pairwise statistical tests further confirmed the superior accuracy of Extra Trees compared to Gradient Boosting (P=.001), Random Forest (P=.006), and MLP (P=.002). Although MLP demonstrated a comparable mean accuracy of 0.94 and even outperformed Extra Trees in the general model setting, it exhibited a broader CI (0.909-0.979) and a similar SD (0.04). In terms of other metrics, Extra Trees also achieved the highest mean AUC of 0.98 and a narrow CI of 0.970-0.990. MLP and Gradient Boosting delivered competitive results in AUC 0.97 and recall 0.96. These findings suggest that while MLP performed well in the general model, Extra Trees consistently outperformed other models in the personalized model evaluations. This consistency identifies Extra Trees as the most reliable model for this multimodal system. Additionally, the observed trends highlight the importance of focusing on personalized models to consider individual variations.

Table . Statistical analysis of model performance metrics across all models.

Metric	Extra Trees	Gradient Boosting	Random Forest	MLP[a]
Mean accuracy	0.95[b]	0.94	0.94	0.94
SD (Acc[c])	0.04	0.04	0.05	0.04
95% CI (Acc)	0.928-0.986	0.912-0.976	0.905-0.977	0.909-0.979
Paired P value (vs Extra Trees)	—[d]	0.001	0.006	0.002
Mean AUC[e]	0.98[b]	0.97	0.96	0.97
SD (AUC)	0.01	0.01	0.01	0.01
95% CI (AUC)	0.970-0.990	0.960-0.980	0.950-0.970	0.960-0.980

[a]MLP: multilayer perceptron.

[b]Bold text indicates the best result for each metric.

[c]Acc: accuracy.

[d]Not applicable.

[e]AUC: area under the curve.

Furthermore, we studied the importance of different features for every participant (Multimedia Appendix 2). For participant #1, the top 10 features contributing to accurate AA classification using the Extra Trees model revealed that 5 were from electrodermal activity, 3 from the accelerometer, 1 from heart rate, and 1 from temperature. Notably, the electrodermal tonic mean was the most critical feature, suggesting a strong link between AA episodes and fluctuations in electrodermal activity, which is commonly associated with emotional arousal. For participant #2, the most important features were primarily the accelerometer and temperature signals. Features related to energy, root mean square, and variability in acceleration played a key role in AA classification. Additionally, temperature fluctuations were also significant, suggesting that both movement patterns and body temperature changes could indicate agitation onset in this participant.

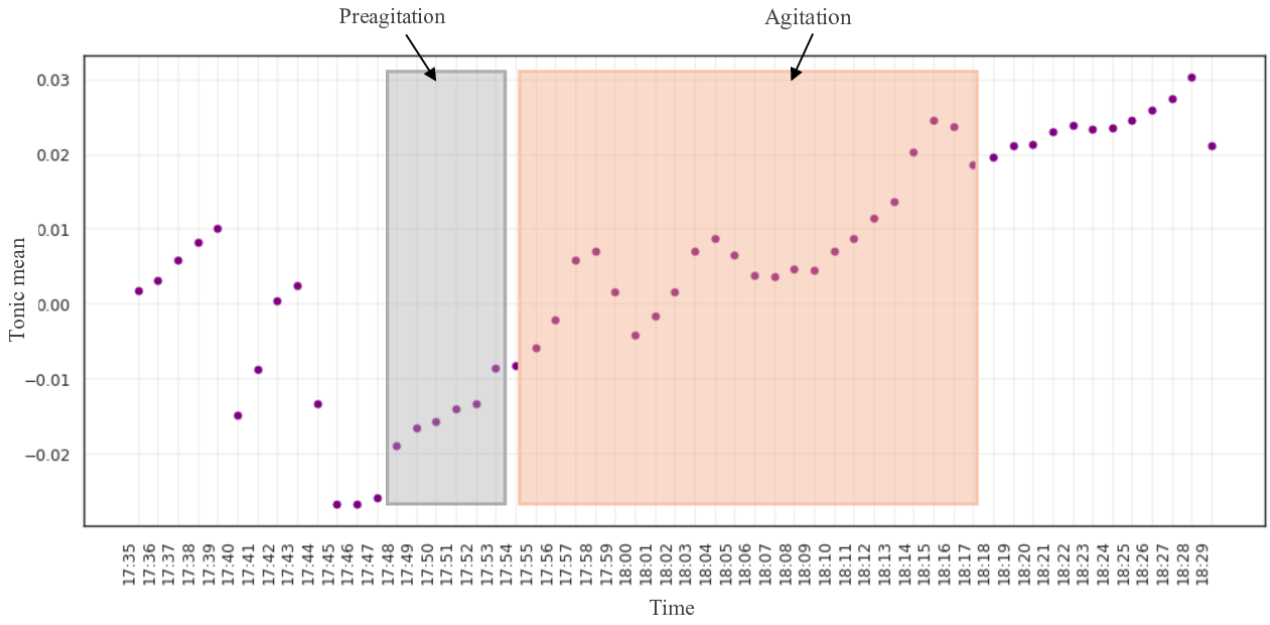
For participant #3, the temperature-related features were the most dominant in identifying AA episodes. The maximum temperature value, temperature root mean square, and energy were among the top contributors, highlighting a strong correlation between changes in body temperature and agitation. These findings emphasize the individual variability in AA predictors, reinforcing the necessity of personalized models for accurate classification. While some participants exhibit agitation-related physiological changes in electrodermal activity, others may show significant patterns in movement or temperature variations, underscoring the importance of a multimodal feature selection approach.

As the electrodermal tonic mean was the top feature to classify AA for participant #1, we investigated the AA labels. Figure 3 shows the tonic mean values of participant #1 from the electrodermal signal during labeled AA events from the camera

XSL•FO
RenderX

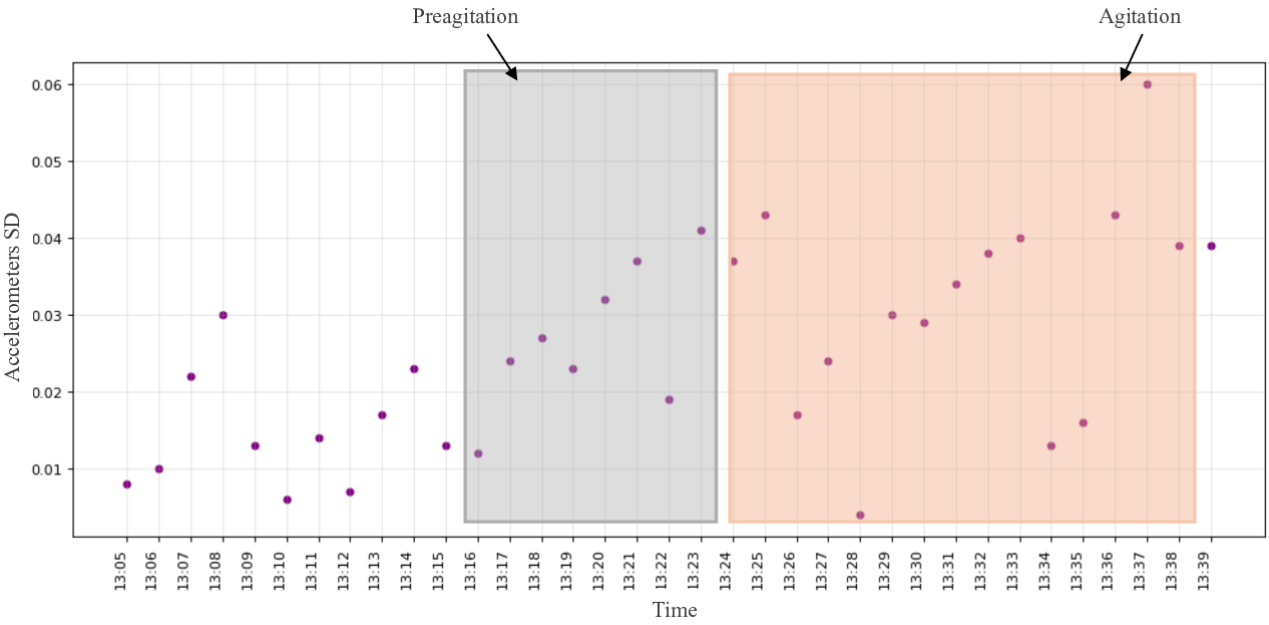
and nurse notes (highlighted in red). This event occurred during the second day and lasted for 23 minutes from 5:55 PM to 6:17 PM. This observation suggests that the patient's AA was related to the electrodermal signal connected to the emotions. We also observed an apparent change to the data before the actual AA occurred, which we manually marked as preagitation labels (highlighted in blue).

Figure 3. The EmbracePlus wristband raw data with agitation and preagitation annotation: the tonic mean plot for participant #1.



For participant #2, the acceleration features were the top features in identifying the AA event. Figure 4 shows one of the AA events for this participant occurring between 1:24 PM and 1:38 PM using the accelerometer data. Moreover, a change in the pattern of the signal 8 minutes before the event is manually labeled as preagitation.

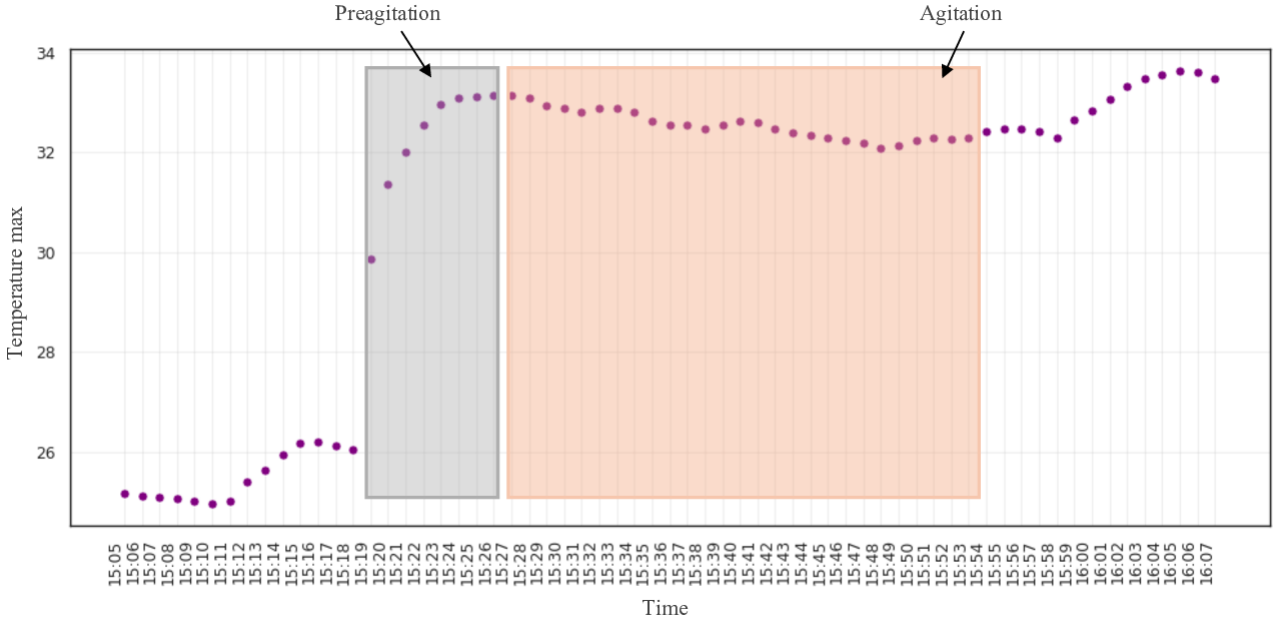
Figure 4. The EmbracePlus wristband raw data with agitation and preagitation annotation: the accelerometer plot for participant #2.



For participant #3, the most dominant features were the temperature features, so an example of the temperature readings for an AA event is shown in Figure 5. Just as before, a change in the pattern before the observed AA event is manually labeled as preagitation. These observations suggest the potential for detecting preagitation patterns from raw data, enabling the prediction of AA before it occurs.

XSL•FO
RenderX

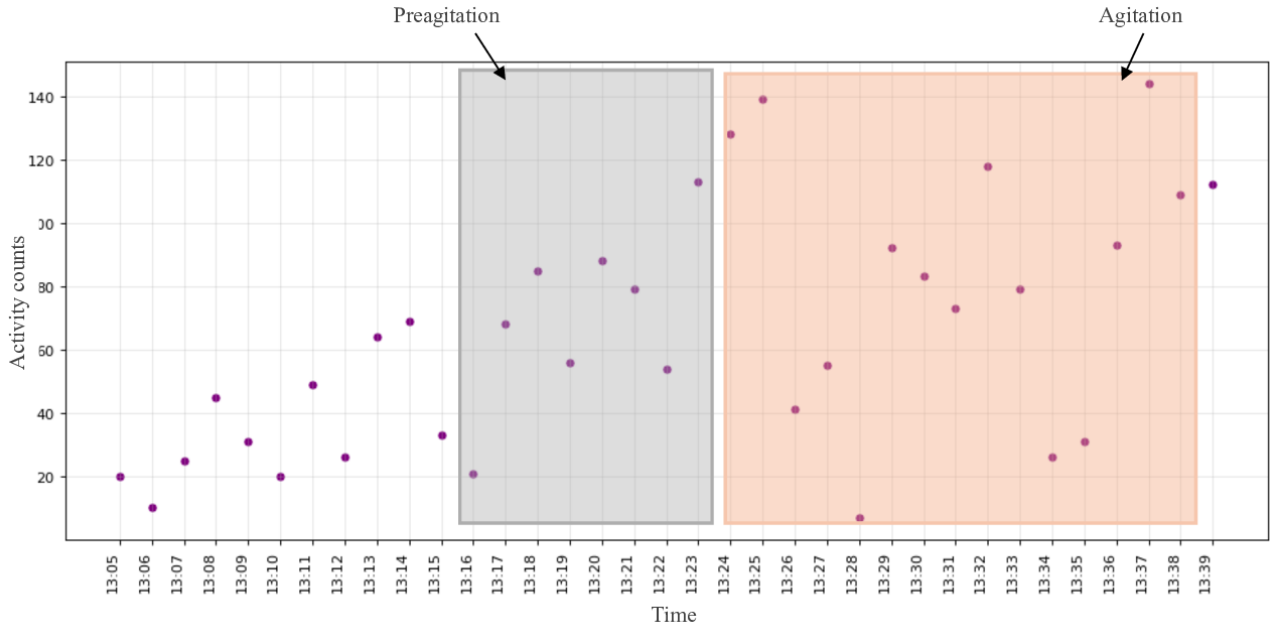
Figure 5. The EmbracePlus wristband raw data with agitation and preagitation annotation: the temperature plot for participant #3. max: maximum.



The Wristband EmbracePlus Digital Biomarkers

We explored all the digital biomarkers offered by EmbracePlus. Taking the first 3 participants as an example, we observed that pulse rate, activity counts, and activity class were the leading indicators for AA detection for the first 3 participants. In Figures 6-8 , the same AA events discussed in the previous subsection for the 3 participants are illustrated, and the values during labeled AA events from the camera and nurse notes are highlighted in red. Additionally, we observed a noticeable change in the data before the onset of AA, manually designated as preagitation labels and highlighted in gray. The manual labeling of the preagitation was done after reviewing all the signals for the participants and noting the same change across multiple patterns.

Figure 6. The activity counts from the digital biomarkers data for participant #2.



XSL•FO
RenderX

Figure 7. The pulse rate from the digital biomarkers data for participant #3.

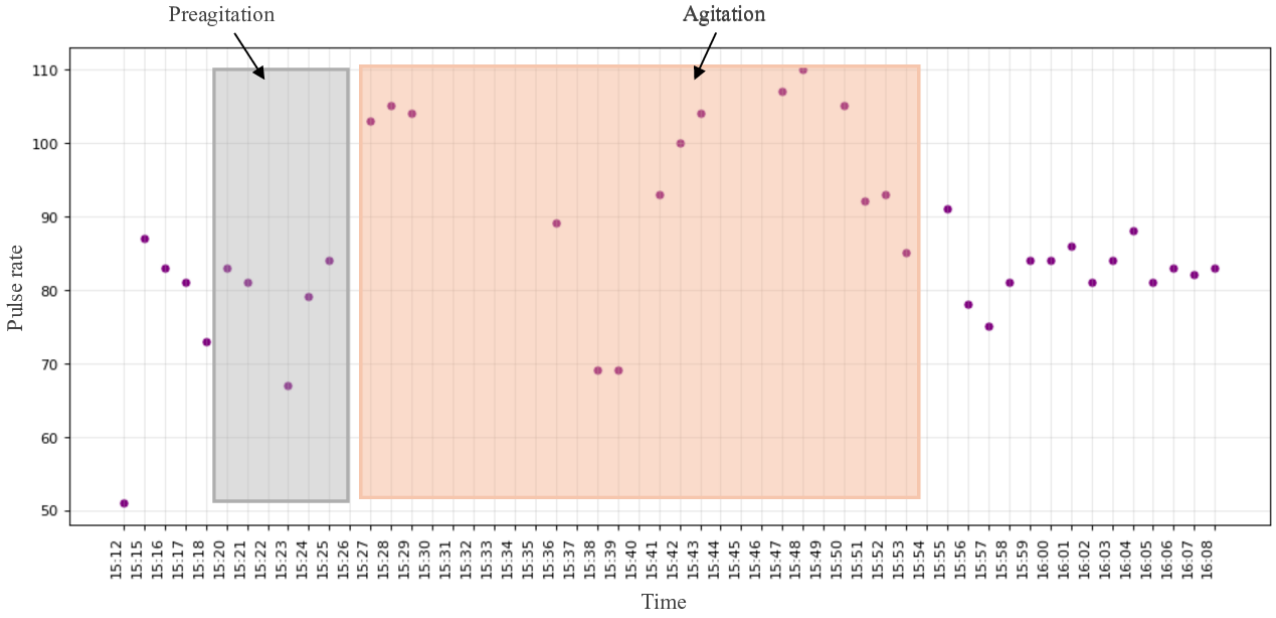


Figure 8. The activity class from the digital biomarkers data for participant #1.

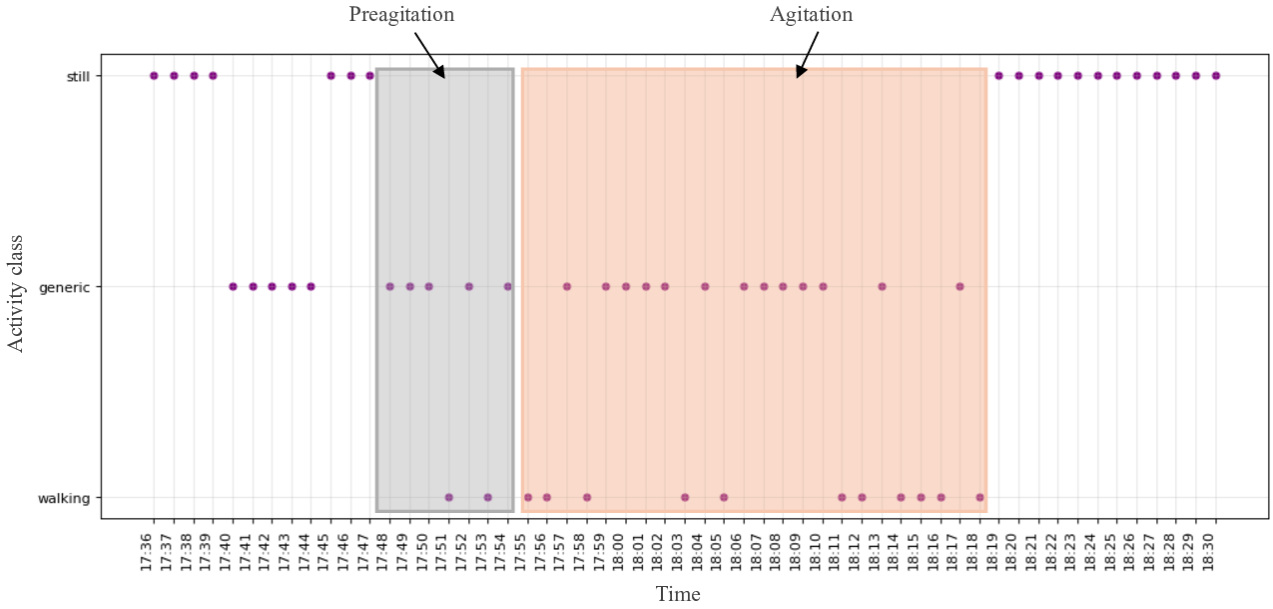


Figure 1 illustrates the activity class for participant #1, extracted from the accelerometer signal, revealing that the participant was in motion rather than stationary during AA and preagitation episodes, indicating body movement during these events. Figure 6 displays the total activity counts for participant #2 from the accelerometer signal. While the normal activity count for the participant ranged between 0 - 100 during AA and preagitation events, it surged to 50 - 140, signifying heightened activity levels during AA. Finally, Figure 7 presents the pulse rate derived from the heart rate signal for participant #3. Although the participant's average pulse rate ranged from 55 - 80 (SD 12.4) bpm, it increased to 90 - 110 (SD 11.6) bpm during AA and preagitation events. Across the raw and digital biomarkers data, we observed that the preagitation occurred from 3:20 PM to 3:27 PM. This indicates that signs of AA behavior occurred approximately 7 minutes before the actual event, suggesting the potential to predict and prevent AA events.

Performance of Video-Based Detection

We preprocessed our videos using OpenPose and performed feature extraction as described in the methodology section. As AA behaviors are repetitive in nature, we selected RNN models to capture the sequential patterns of these events. Features were extracted from 30-second windows. The window moves 1 second at a time to capture different variations of AA behaviors from the skeletal points. For the classification task, we tested 2 different network structures, which are LSTM and GRU.

In the feature analysis step, we focused on reducing the model's dimensionality without compromising its performance. Initially, 47 features were extracted based on skeletal movements,

XSL•FO
RenderX

including Euclidean distances and angles between key body parts. It is evident from the figures that more features are highly correlated in agitation events than in nonagitation events. This had a strong effect on the feature reduction as only the highly correlated features in both datasets were removed. We investigated the correlation of the features more deeply and tested the model on fewer features based on a correlation threshold. Setting the correlation threshold to 0.8 reduced the number of features to 39 features. The number of features based on the correlation threshold did not change even when the threshold was set to as low as 0.5 due to the huge difference in correlation between both types of events. Both datasets are randomly split into 70% for training and 30% for testing. The training was conducted on a Lambda server equipped with an RTX A6000 GPU, and each model was trained for 100 epochs with a batch size of 256. We used the Adam optimizer to efficiently handle sparse gradients and used a sigmoid activation function for binary classification (AA vs nonagitation). We report the results of all the tests on the testing set in Table 6.

Table . Video results comparison.

Model	Number of features	Accuracy	AUC[a]	F_1-score	Recall	Time (s)
LSTM[b]	47	0.94	0.98	0.97	0.96	16.1
LSTM	39	0.94	0.98	0.98	0.97	**14.7**[c]
GRU[d]	47	0.95[e]	0.98[e]	0.98[e]	0.97	30.2
GRU	39	0.95[e]	0.99[e]	0.98[e]	0.98[e]	29.8

[a]AUC: area under the curve.

[b]LSTM: long short-term memory.

[c]Bold text indicates the shortest inference time.

[d]GRU: gated recurrent unit.

[e]Bold text indicates the best result for each metric.

We compared metrics such as accuracy, AUC, and F_1-score. We also compared the response time, an essential factor in real-time applications, of all models. As observed in the table, the reduction in the number of features did not affect the performance of the models. The LSTM model, in both cases, achieved an accuracy of 94% and an AUC of 98%. The GRU model reached 95% accuracy and 99% AUC. Although the performance of both models is similar, the response time of the GRU model is double that of the LSTM. The response time for the LSTM was 15.9 seconds when all the features were used in training and was almost one second faster with fewer features. The GRU models, however, had a response time of 30.1 seconds for 47 features and 29.7 seconds for 39 features. The results show that the GRU model is superior across all evaluation metrics, albeit for the response time, where it lags behind the LSTM model by a huge margin. As we aim to detect AA as early as possible, the swifter response time can allow for timely interventions by health care providers in case of any AA event. The high AUC values of both models signify a strong ability to minimize the rate of false positives. This is crucial to ensure the reliability of the model in detecting real AA with lower false alarms, causing less overhead for health care providers. The faster response time posits the LSTM model as possibly the more advantageous model for real-time deployment.

Discussion

Principal Findings

The successful implementation of the system within the hospital setting, considering privacy and the positive feedback from patients and health care professionals, highlights the system's viability in a real-world clinical environment. The system used in this study integrated physiological data from the EmbracePlus wristband and video footage from CCTV cameras, allowing for a comprehensive and multimodal approach to AA detection. The EmbracePlus wristband system demonstrates promising results in detecting and classifying AA and preagitation events in individuals with severe dementia. The AA detection results are reflected in the video detection system, and the preagitation labels can be added to the system from EmbracePlus. The following discussion highlights key findings and their implications, followed by suggestions for future work.

The EmbracePlus wristband, which leverages both raw data and digital biomarkers, demonstrated its efficacy in discerning patterns associated with AA and preagitation. The personalized Extra Trees model emerged as the top-performing algorithm for the raw data, achieving high performance. Furthermore, features such as electrodermal tonic mean, accelerometer activity class, and pulse rate highlighted the significance of identifying AA. Furthermore, the outcomes of our analysis are promising and demonstrate the potential of predicting AA in dementia care settings in real time. We were able to predict preagitation events from all participants at least 6 minutes before the actual AA event. The identification of preagitation patterns in the data suggests that physiological changes precede observable AA behavior. Being the first to explore these patterns in individuals with severe dementia from the EmbracePlus wristband, this study lays the groundwork for a deeper understanding of the dynamics and physiological signatures of AA behaviors. The newfound ability to identify preagitation patterns offers a potential window for early intervention and preventive measures.

Moreover, the video detection system, incorporating CCTV footage and advanced pose estimation techniques, was used along with the EmbracePlus wristband for AA detection. The privacy preservation technique, which follows the REB

XSL•FO
RenderX

protocols in the hospital, does not exploit the patient's personal features or body image without affecting the performance of the proposed model. The LSTM neural network and the GRU networks exhibited robust performances. The high AUC of both models is particularly crucial in the context of health care to minimize the risk of false negatives and ensure that true AA events are accurately identified. During the real-time deployment stage, the model adapts and continuously improves based on the collected data. The labeled preagitation labels collected from the wristband can be fed into the training set of the video detection system to provide insight into detecting preagitation from the video footage. The outcomes of our analysis are promising and demonstrate the potential of both LSTM and GRU neural networks in detecting AA in dementia care settings in real time.

Future Work

For future work, we will focus on expanding our dataset by recruiting more participants from the hospital. We aim to validate the system over the long term, assessing its stability, generalizability, and adaptability to health care and home care environments. This step is crucial for building a database with a substantial number of AA and preagitation events, which is essential for developing a high-performance detection system using machine learning. Once our detection system is established, we plan to automate the real-time system that is capable of predicting AA. This involves receiving data in real time, classifying the data, and sending notifications to the health care providers if an AA event occurs. Additionally, we plan to improve the real-time AA detection of the video system. We also plan to use the preagitation labels from the EmbracePlus wristband to help our video detection model predict AA before they happen. During the initial stages of this study, health care providers will review and confirm all collected AA events. Their feedback is crucial in refining and enhancing the model's performance and should aid in identifying any limitations or challenges. Once the model is reliable, it will automatically detect and predict AA with no human intervention.

Conclusions

This study represents a notable step forward in developing an AA and preagitation detection system for individuals with severe dementia. It uses a comprehensive approach by integrating psychological biomarker sensing and video detection systems. The results demonstrate the feasibility and efficacy of monitoring systems that combine various data sources for AA detection. This study recruited 10 participants from the Ontario Shores Center for Mental Health Sciences Institute. We used the EmbracePlus wristband for continuous health monitoring and video footage from CCTV cameras for real-time observation of AA events. In the preliminary data analysis, the features extracted from the raw data of the EmbracePlus wristband demonstrated exceptional performance in detecting AA events, with the Extra Trees model emerging as the top-performing algorithm for all the personalized models and MLP outperforming the rest of the models for the general model and achieving an accuracy of 98%. Exploring the digital biomarkers further strengthened the system's classification of AA, preagitation, and normal events. Pulse rate, activity class, and activity counts have emerged as critical indicators for detecting AA. This study revealed the potential for detecting preagitation patterns, showcasing a 6-minute lead time before actual AA events. This early detection capability holds promise for timely intervention and preventive measures. In addition to the EmbracePlus wristband, the video-based detection demonstrated promising results in detecting AA using GRU, achieving a 95% accuracy rate and a robust AUC of 98%. The data analysis' promising results highlight the potential of the multimodal approach to enhance patient care and safety by predicting AA events. This research will provide new directions for researchers interested in technologies for dementia care and provide challenging propositions in detecting and monitoring, modeling, and evaluating patient-specific interventions for people living with dementia demonstrating NPS.

Acknowledgments

This research was undertaken, in part, thanks to funding from the Canada Research Chairs Program.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Workflow of the data collection process.
[DOCX File, 280 KB - aging_v8i1e68156_app1.docx]

Multimedia Appendix 2
Feature importance for participants #1, #2, and #3.
[DOCX File, 104 KB - aging_v8i1e68156_app2.docx]

References

1. Dementia. World Health Organization. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2025-06-10]

XSL•FO
RenderX

2. Lyketsos CG, Lopez O, Jones B, Fitzpatrick AL, Breitner J, DeKosky S. Prevalence of neuropsychiatric symptoms in dementia and mild cognitive impairment: results from the cardiovascular health study. JAMA 2002 Sep 25;288(12):1475-1483. [doi: 10.1001/jama.288.12.1475] [Medline: 12243634]
3. Ballard C, Corbett A. Agitation and aggression in people with Alzheimer's disease. Curr Opin Psychiatry 2013 May;26(3):252-259. [doi: 10.1097/YCO.0b013e32835f414b] [Medline: 23528917]
4. Seitz D, Purandare N, Conn D. Prevalence of psychiatric disorders among older adults in long-term care homes: a systematic review. Int Psychogeriatr 2010 Nov;22(7):1025-1039. [doi: 10.1017/S1041610210000608] [Medline: 20522279]
5. Knuff A, Leung RH, Seitz DP, Pallaveshi L, Burhan AM. Use of actigraphy to measure symptoms of agitation in dementia. Am J Geriatr Psychiatry 2019 Aug;27(8):865-869. [doi: 10.1016/j.jagp.2019.02.013] [Medline: 30952608]
6. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. Neurology (ECronicon) 1994 Dec;44(12):2308-2314. [doi: 10.1212/wnl.44.12.2308] [Medline: 7991117]
7. Cohen-Mansfield J. Cohen-mansfield agitation inventory. Int J Geriatr Psychiatry 1986.
8. Artificial Intelligence in Healthcare Market Size & Trends Analysis Report. Research and Markets. 2022. URL: https://www.researchandmarkets.com/report/healthcare-ai?srsltid=AfmBOopaBJv8BVQKRVRLr1YHW1TDHwxLSyce4o5UHr0dB0DX9g43dh3N [accessed 2025-6-19]
9. Bharucha AJ, Anand V, Forlizzi J, et al. Intelligent assistive technology applications to dementia care: current capabilities, limitations, and future challenges. Am J Geriatr Psychiatry 2009 Feb;17(2):88-104. [doi: 10.1097/JGP.0b013e318187dde5] [Medline: 18849532]
10. Khan SS, Ye B, Taati B, Mihailidis A. Detecting agitation and aggression in people with dementia using sensors—a systematic review. Alzheimer's Dementia 2018 Jun;14(6):824-832. [doi: 10.1016/j.jalz.2018.02.004]
11. Fabrizio C, Termine A, Caltagirone C, Sancesario G. Artificial intelligence for Alzheimer's disease: promise or challenge? Diagnostics (Basel) 2021 Aug 14;11(8):1473. [doi: 10.3390/diagnostics11081473] [Medline: 34441407]
12. Goerss D, Köhler S, Rong E, et al. Smartwatch-based interventions for people with dementia: user-centered design approach. JMIR Aging 2024 Jun 7;7:e50107. [doi: 10.2196/50107] [Medline: 38848116]
13. Anderson MS, Bankole A, Homdee N, Mitchell BA, Byfield GE, Lach J. Dementia caregiver experiences and recommendations for using the behavioral and environmental sensing and intervention system at home: usability and acceptability study. JMIR Aging 2021 Dec 6;4(4):e30353. [doi: 10.2196/30353] [Medline: 34874886]
14. Ontario Shores Centre for Mental Health Sciences. 2022. URL: https://www.ontarioshores.ca [accessed 2025-06-10]
15. EmbracePlus. URL: https://www.empatica.com [accessed 2025-06-10]
16. Carrarini C, Russo M, Dono F, et al. Agitation and dementia: prevention and treatment strategies in acute and chronic conditions. Front Neurol 2021;12:644317. [doi: 10.3389/fneur.2021.644317] [Medline: 33935943]
17. Husebo BS, Heintz HL, Berge LI, Owoyemi P, Rahman AT, Vahia IV. Sensing technology to monitor behavioral and psychological symptoms and to assess treatment response in people with dementia. a systematic review. Front Pharmacol 2019;10:1699. [doi: 10.3389/fphar.2019.01699] [Medline: 32116687]
18. Yang P, Bi G, Qi J, Wang X, Yang Y, Xu L. Multimodal wearable intelligence for dementia care in healthcare 4.0: a survey. Inf Syst Front 2025 Feb;27(1):197-214. [doi: 10.1007/s10796-021-10163-3]
19. Valembois L, Oasi C, Pariel S, Jarzebowski W, Lafuente-Lafuente C, Belmin J. Wrist actigraphy: a simple way to record motor activity in elderly patients with dementia and apathy or aberrant motor behavior. J Nutr Health Aging 2015 Aug;19(7):759-764. [doi: 10.1007/s12603-015-0530-z] [Medline: 26193860]
20. Etcher L, Whall A, Kumar R, Devanand D, Yeragani V. Nonlinear indices of circadian changes in individuals with dementia and aggression. Psychiatry Res 2012 Aug 30;199(1):77-78. [doi: 10.1016/j.psychres.2012.03.042] [Medline: 22521234]
21. Badawi A, Elgazzar K, Ye B, et al. Investigating multimodal sensor features importance to detect agitation in people with dementia. : IEEE; 2023 Presented at: 2023 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE); Sep 24-27, 2023; Regina, SK, Canada p. 77-82. [doi: 10.1109/CCECE58730.2023.10288764]
22. Cheung JCW, So BPH, Ho KHM, Wong DWC, Lam AHF, Cheung DSK. Wrist accelerometry for monitoring dementia agitation behaviour in clinical settings: a scoping review. Front Psychiatry 2022;13:913213. [doi: 10.3389/fpsyt.2022.913213] [Medline: 36186887]
23. Chikhaoui B, Ye B, Mihailidis A. Ensemble learning-based algorithms for aggressive and agitated behavior recognition. : Springer Presented at: 10th International Conference, UCAmI 2016, Part II; Nov 29 to Dec 2, 2016; Canary Islands, Spain p. 9-20. [doi: 10.1007/978-3-319-48799-1_2]
24. Mishra PK, Iaboni A, Ye B, Newman K, Mihailidis A, Khan SS. Privacy-protecting behaviours of risk detection in people with dementia using videos. Biomed Eng Online 2023 Jan 21;22(1):4. [doi: 10.1186/s12938-023-01065-3] [Medline: 36681841]
25. Khan SS, Mishra PK, Javed N, et al. Unsupervised deep learning to detect agitation from videos in people with dementia. IEEE Access 2022;10:10349-10358. [doi: 10.1109/ACCESS.2022.3143990] [Medline: 35582495]
26. Marshall F. Video-based automatic detection of agitated behaviours Ph.D dissertation. : Ulster University; 2022.
27. Rose K, Specht J, Forch W. Correlates among nocturnal agitation, sleep, and urinary incontinence in dementia. Am J Alzheimers Dis Other Demen 2015 Feb;30(1):78-84. [doi: 10.1177/1533317514524814] [Medline: 24670931]

XSL•FO
RenderX

28. Gong J, Rose KM, Emi IA, et al. Home wireless sensing system for monitoring nighttime agitation and incontinence in patients with Alzheimer's disease. : ACM; 2015 Oct 14 Presented at: Proceedings of the Conference on Wireless Health, ser WH '15; Oct 14-16, 2015; Bethesda, MD p. 5. [doi: 10.1145/2811780.2822324]
29. Khan SS, Zhu T, Ye B, et al. DAAD: a framework for detecting agitation and aggression in people living with dementia using a novel multi-modal sensor network. : IEEE; 2017 Presented at: 2017 IEEE International Conference on Data Mining Workshops (ICDMW); Nov 18-21, 2017; New Orleans, LA p. 703-710. [doi: 10.1109/ICDMW.2017.98]
30. Ye B, Khan SS, Chikhaoui B, et al. Challenges in collecting big data in a clinical environment with vulnerable population: lessons learned from a study using a multi-modal sensors platform. Sci Eng Ethics 2019 Oct;25(5):1447-1466. [doi: 10.1007/s11948-018-0072-y] [Medline: 30357559]
31. Zhang K, Yuan H, Fang Y. Evaluation of smart agitation prediction and management for dementia care and novel universal village oriented solution for integration, resilience, inclusiveness and sustainability. : IEEE; 2022 Presented at: 2022 6th International Conference on Universal Village (UV); Oct 22-25, 2022; Boston, MA p. 1-34. [doi: 10.1109/UV56588.2022.10185497]
32. Alam R, Gong J, Hanson M, et al. Motion biomarkers for early detection of dementia-related agitation. 2017 Jun 23 Presented at: MobiSys'17; Jun 23, 2017; Niagara Falls, NY p. 15-20 URL: https://dl.acm.org/doi/proceedings/10.1145/3089341 [doi: 10.1145/3089341.3089344]
33. AP Association. Diagnostic and Statistical Manual of Mental Disorders, 5th edition: American Psychiatric Publishing, Inc; 2013.
34. Cummings J, Mintzer J, Brodaty H, et al. Agitation in cognitive disorders: International Psychogeriatric Association provisional consensus clinical and research definition. Int Psychogeriatr 2015 Jan;27(1):7-17. [doi: 10.1017/S1041610214001963] [Medline: 25311499]
35. Sclan SG, Reisberg B. Functional Assessment Staging (FAST) in Alzheimer's disease: reliability, validity, and ordinality. Int Psychogeriatr 1992;4 Suppl 1(3):55-69. [doi: 10.1017/s1041610292001157] [Medline: 1504288]
36. AXIS M3077-PLVE network camera. Axis. 2024. URL: https://www.axis.com/products/axis-m3077-plve#compatible-discontinued [accessed 2025-06-10]
37. AXIS P3225-VE Mk II Network Camera - product support. Axis Communications. 2024. URL: https://www.axis.com/products/axis-p3225-ve-mk-ii/support [accessed 2025-06-10]
38. Badawi A, Choudhury S, Elgazzar K, Burhan AM. Artificial intelligence and features investigating to detect neuropsychiatric symptoms in patients with dementia: a pilot study. 2023 Presented at: 2023 IEEE Symposium Series on Computational Intelligence (SSCI); Dec 5-8, 2023; Mexico City, Mexico. [doi: 10.1109/SSCI52147.2023.10371835]
39. Cao Z, Martinez GH, Simon T, Wei S, Sheikh YA. OpenPose: Realtime multi- person 2d pose estimation using part affinity fields. IEEE Trans Pattern Anal Mach Intell 2019;43(1):172-186. [doi: 10.1109/TPAMI.2019.2929257]
40. Simon T, Joo H, Matthews I, Sheikh Y. Hand keypoint detection in single images using multiview bootstrapping. 2017 Presented at: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Jul 21-26, 2017; Honolulu, HI. [doi: 10.1109/CVPR.2017.494]
41. Cao Z, Simon T, Wei SE, Sheikh Y. Realtime multi-person 2d pose estimation using part affinity fields. 2017 Presented at: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Jul 21-26, 2017; Honolulu, HI. [doi: 10.1109/CVPR.2017.143]
42. Wei SE, Ramakrishna V, Kanade T, Sheikh Y. Convolutional pose machines. 2016 Presented at: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Jun 27-30, 2016; Las Vegas, NV. [doi: 10.1109/CVPR.2016.511]

Abbreviations

AA: agitation and aggression
AI: artificial intelligence
AUC: area under the curve
GDU: Geriatric Dementia Unit
GRU: gated recurrent unit
JREB: Joint Research Ethics Board
LSTM: long short-term memory
MLP: multilayer perceptron
NPS: neuropsychiatric symptom
REB: research ethics board
RNN: recurrent neural network

XSL•FO
RenderX

Edited by R Yang; submitted 30.10.24; peer-reviewed by C Wu, S Kumar; revised version received 31.03.25; accepted 28.04.25; published 15.07.25.

Please cite as:
Badawi A, Elmoghazy S, Choudhury S, Elgazzar S, Elgazzar K, Burhan AM
Multimodal Detection of Agitation in People With Dementia in Clinical Settings: Observational Pilot Study
JMIR Aging 2025;8:e68156
URL: https://aging.jmir.org/2025/1/e68156
doi:10.2196/68156

© Abeer Badawi, Somayya Elmoghazy, Samira Choudhury, Sara Elgazzar, Khalid Elgazzar, Amer M Burhan. Originally published in JMIR Aging (https://aging.jmir.org), 15.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Usability and Impact of the Web-Based Dementia Foundations Educational Program in Personal Support Workers (PSWs), PSW Trainees, and Care Companions: Quasi-Experimental Study

Anthony J Levinson[1*], MA, MSc, MD; Stephanie Ayers[1*], BA, BEd; Sandra Clark[1], MSc; Patricia Gerantonis[1], BSc, MPH; Amy Schneeberg[2], PhD; Richard Sztramko[3], BSc, MD

[1]Division of e-Learning Innovation, Faculty of Health Sciences, McMaster University, 1280 Main Street West, Hamilton, ON, Canada

[2]Freelance Statistical Consultant, Vancouver, BC, Canada

[3]Department of Medicine, McMaster University, Hamilton, ON, Canada

[*]these authors contributed equally

Corresponding Author:
Anthony J Levinson, MA, MSc, MD
Division of e-Learning Innovation, Faculty of Health Sciences, McMaster University, 1280 Main Street West, Hamilton, ON, Canada

Abstract

Background: Personal support workers (PSWs) are often expected to provide ongoing support for complex conditions and have identified an increased need for training in several areas, including dementia and mental health. Web-based interventions may be helpful complements to traditional in-person continuing education and training, but their effectiveness must be explored further.

Objective: This study's objective was to evaluate the usability, usefulness, satisfaction with, and effectiveness of the web-based Dementia Foundations Program among unregulated care providers who provide care to persons living with dementia or are in training.

Methods: A cohort of 50 PSWs, PSW trainees, and paid care companions from 3 recruitment sites were invited to access the Dementia Foundations Program, a 4-hour self-paced web-based program composed of 4 courses, for up to 6 weeks. Usability, usefulness, and satisfaction were assessed using surveys after each course and following the program. Dementia knowledge and attitudes were measured using the Dementia Knowledge Assessment Scale and the Dementia Attitudes Scale, with differences between baseline and postprogram scores analyzed using repeated measures ANOVA.

Results: Participants reported high levels of satisfaction with the program. Of the 50 participants, 46 (92%) agreed that the web-based training met their expectations, 47 (94%) agreed that the training covered a broad range of topics and was not missing any important content, and 49 (98%) agreed that the web-based training would benefit them. There was a significant postprogram improvement in dementia knowledge as measured by the Dementia Knowledge Assessment Scale, with an average 30% improvement across all cohorts. Dementia Attitudes Scale scores were also significantly improved postprogram across all cohorts.

Conclusions: This pilot study in PSWs, PSW trainees, and unregulated care companions demonstrated high satisfaction levels with the web-based Dementia Foundations Program. There were substantial improvements in knowledge and small improvements in attitudes for participants, and it was perceived as a useful tool that complemented their existing education and training. The Dementia Foundations Program is a user-friendly and effective e-learning program, which can be conveniently scaled and spread to enhance unregulated care provider dementia education.

(JMIR Aging 2025;8:e67889) doi:10.2196/67889

KEYWORDS

dementia; caregivers; e-learning; education; training; mobile phone

Introduction

Background

The number of older adults aged 60+ years is expected to double worldwide by 2050 [1]. In Canada specifically, the older seniors population aged 75+ years is expected to grow even sooner, possibly tripling by 2046 [2]. This has the potential to increase the prevalence of age-related health conditions such as dementia, as well as the need for adequate human health resources [3].

The Need for Unregulated Care Providers

There are estimated to be about 600,000 people in Canada living with dementia currently, with that number projected to grow to

XSL•FO
RenderX

almost a million in 2030 [4]. Unregulated care professionals who provide care to older adults play an important role in the effort to meet the growing demand for care from Canada's aging population. Several occupational titles can be used to describe these unregulated care providers, including personal support workers (PSWs), health care aides, and home care workers [5]. For the purposes of this paper, we will predominantly use the term PSWs, with the intent that this information applies to most professionals working in similar unregulated care roles.

It is estimated that there are over 1 million PSWs in Canada [6] providing support, including assisting with daily living activities, light housekeeping, meal preparation, socialization, and companionship [7], in several settings, including home, community, and clinical. In Canada, PSWs are not currently required to hold accredited qualifications and are not formally regulated by an organization or governing body. This lack of regulation of training can lead to gaps in knowledge and caregiving abilities [6]. PSWs have indicated that more training is needed, specifically in the field of dementia and mental health [8]. With over 69% of residents in long-term care (LTC) homes living with dementia and 87% with some form of cognitive impairment [9], more specific education and training about dementia are needed for PSWs.

Dementia Education Approaches

There are currently a variety of training approaches for dementia for unregulated care providers. Some PSW training programs or colleges might offer traditional didactic instruction on topics related to dementia; some of these programs have transitioned to webinar-based instruction during the pandemic. For PSWs or other unregulated care providers already in the workplace, the most common approach would be occasional on-the-job training, such as periodic "in-service" or "lunch-and-learn" sessions. These approaches tend to be synchronous, nonstandardized, and not necessarily educationally designed. Very few offer foundational knowledge content about dementia. Some approaches may also involve face-to-face instruction, which did prove to be challenging during the pandemic, concerning public health guidance as well as participant hesitance, even with the loosening of public health guidance [10]. Many of those solutions would be both challenging to scale and likely inconvenient for these care providers, who may not have flexible schedules that can accommodate synchronous instruction, especially if expected to be done on unpaid (personal) time.

For these reasons, the Dementia Foundations Program was developed in 2021 at McMaster University, aligning with several performance objectives and standards of the PSW Program Standard for Ontario Colleges from the Ministry of Training, Colleges and Universities [11]. During program development, project team members and target audience stakeholders identified the need for flexible, asynchronous solutions to best facilitate delivery to their PSW employees and students.

Program Description

The Dementia Foundations Program is a web-based, asynchronous (self-paced) series of 4 courses that takes approximately 4 hours total to complete. Each of the 4 courses includes modules on various topics related to dementia and dementia care and is described in Table 1.

Table . Dementia foundations program course descriptions.

Course number and course name	Course content
1. Foundations of Dementia	• What cognition is and how it is impacted by dementia • Difference between normal aging and mild cognitive impairment • Different types and stages of dementia • Types of treatments available
2. Responsive Behaviors and Mental Health	• Psychiatric issues such as apathy, depression, and anxiety that may affect people with dementia • Behavioral, emotional, and psychiatric symptoms of dementia • Strategies to help manage responsive behaviors
3. Home Supports and Safety	• Identifying safety risks • Strategies to reduce harm • Home and community-based supports and services available for patients and families
4. Promoting Brain Health and Caregiver Wellness	• Importance of brain health • How to make positive brain health choices • Why caregiver wellness is important and what can be done to support it

The program is intended as a complementary dementia education resource that may benefit care providers' training for several reasons. First, the modules within the courses use high-quality, evidence-based, multimedia instructional design [12,13], which has been shown to improve learning outcomes. Second, it can be delivered asynchronously for self-paced learning at any time and from any place, unlike other training programs that may use real-time training formats such as instructor-led classroom or webinar training. Third, the program has been developed using responsive web design and is compatible with any internet-connected device (smartphone, tablet, or desktop). Fourth, it is more efficient and cost-effective to scale than standardized digital or in-person training, as it does not require the presence of instructors. Lastly, the Dementia Foundations

XSL•FO
RenderX

Program offers a certificate of completion for each course and the complete program that learners can share with future employers detailing the time they have invested in learning about caring for those living with dementia.

Study Purpose

The objective of this study was to evaluate the usability, usefulness, satisfaction, and effectiveness of the Dementia Foundations Program among unregulated care providers who currently provide care to, or could provide care to, persons living with dementia.

Methods

Study Design

A pilot pre-post design was used with measures taken at baseline, after each of the 4 courses, and after program completion. Participants were able to complete the program at their own pace over a maximum of 6 weeks.

Participants and Recruitment Cohorts

A total of 50 participants were recruited through collaborative partner networks via an email campaign, keeping in line with typical pilot study sample size guidance in the literature [14,15]. Participants were organized into 3 recruitment cohorts: those working in the city of Ottawa's LTC homes, those currently a part of Durham College's PSW training program, and those recruited from the uCarenet homecare digital marketplace. Participants who fulfilled the following inclusion criteria were eligible to participate: (1) had a good command of the English language, (2) were sufficiently computer literate to use the Dementia Foundations Program, and (3) were currently an unregulated health care provider or trainee in Ontario, Canada. Interested participants were directed to a private web-based sign-up link where they were provided further study information and could submit informed consent. Once informed consent was received, a member of the research team manually added participants to the Dementia Foundations Program within the learning management system. Participants were grouped according to their recruitment cohort to allow the research team to identify and analyze patterns and trends among these 3 different settings.

The Dementia Foundations Program

Participants had access to the Dementia Foundations Program from June 1 to July 9, 2021. The program could be accessed from any internet-enabled device at any time or location. In addition to the 4 courses, participants in this pilot study had access to a "Getting to Know You" course at the beginning (which included the baseline assessments), and a "Review and Final Quiz" course at the end, which included the postprogram assessments.

Program Usability, Usefulness, and Satisfaction

To evaluate perceived usability, usefulness, and satisfaction with the training, 2 questionnaires were developed based on the Information Assessment Method for all questionnaire—a content-validated questionnaire designed to collect feedback from health information consumers based on 4 domains: situational relevance, cognitive impact, information use, and health benefits [16]. The first was administered after each course (postcourse assessment), and the second was delivered upon program completion (postprogram assessment). All questions were required, and each survey could only be submitted once; participants were able to review their responses and were asked to complete any questions they had missed before submitting.

The postcourse assessment consisted of 7 questions in total: 4 questions that could be answered on a 5-point Likert scale ranging from "strongly agree" to "strongly disagree" ("I thought the course content was very important to my professional/learning needs" [usefulness], "I understood the content in this course" [usability], "I was able to complete the course in a reasonable amount of time" [usability], and "I would recommend this course to a colleague" [usefulness and satisfaction]), 2 questions that could be answered with multiple-select answers ("What do you think about this course?" [usefulness] and "Which benefit(s) are you expecting after taking this course?" [usefulness]), and 1 open text question ("Any other comments about this course?" [usability, usefulness, or satisfaction]).

The postprogram survey consisted of 8 questions total: 4 questions that could be answered on a 5-point Likert scale ranging from "strongly agree" to "strongly disagree" ("This online training met my expectations" [satisfaction], "I felt that this training covered a broad range of topics and was not missing any important content or topics" [usefulness and satisfaction], "This online training will benefit me" [usefulness and satisfaction], and "I feel my certificate/credential related to this program will be very valuable to me with current or future employers" [usefulness and satisfaction], 1 question that could be answered with a 3-point Likert scale ranging from "too low" to "too high" (thoughts on proposed pricing structure), 1 question that could be answered with multiple-select options ("Which of the following courses would you consider paying for?"), and 2 open text questions ("How much would you be willing to pay for this program?", and "Is there anything else you would like to say about this program?" (usability, usefulness, or satisfaction).

Dementia Knowledge

To evaluate the effectiveness of the Dementia Foundations Program, change in knowledge and attitudes related to dementia was assessed. Knowledge was measured at baseline and after program completion with the Dementia Knowledge Assessment Scale (DKAS; reliability α=.85; ω_h=0.87; overall scale) [17]. The DKAS consists of 25 items on different aspects of dementia that could be answered with "true," "probably true," "false," "probably false," or "I don't know." The total scores achievable for this scale range from 0 to 50, with greater dementia-related knowledge reflected by a higher score.

Dementia Attitudes

Attitudes about dementia were measured at baseline and after program completion with the Dementia Attitudes Scale (DAS) [18]. The DAS consists of 20 items on a 7-point Likert scale that reflect the affective, behavioral, and cognitive components of the attitudes toward individuals with Alzheimer disease and

XSL•FO
RenderX

related dementias. The total scores achievable for this scale range from 20 to 140, with a more positive attitude reflected by a higher score.

Learning Management System Setup

All assessments were offered digitally within a learning management system. The baseline "Getting to Know You" course included a demographic questionnaire (age, sex, race or ethnicity, highest level of education completed, occupation, and most common workplace setting), the DKAS, and the DAS. The postcourse assessment was built directly into the end of each of the 4 courses. The postprogram "Review and Final Quiz" course included the DKAS, DAS, and postprogram assessment.

Data Analysis

Descriptive analyses were used to summarize the baseline characteristics of the study population. Differences between baseline and postintervention DKAS and DAS scores for all groups were analyzed using repeated measures ANOVA. All analyses were performed using R (version 4.4.1; R Foundation; June 14, 2024).

Ethical Considerations

The Hamilton Integrated Research Ethics Board reviewed the study protocol and granted exemption from full review on March 18, 2021, as this was considered a quality improvement initiative. Participants were required to provide informed consent. All participants were informed of the length of time of the e-learning and surveys, as well as details surrounding data collection, storage, and investigator identities. Participants' identities and confidentiality were maintained throughout the research study. All participant data were deidentified and were stored on password-protected secure servers to prevent unauthorized access. There was no known risk or harm to participating in this study or publicizing its results or findings. Upon program completion, participants were provided with a CAD $80 (US $58.22) gift card and a certificate of completion for their professional portfolio.

Results

Participant Characteristics

A total of 50 unregulated care providers enrolled and completed this study. Of these, 24 (48%) participants identified as PSWs, 16 (32%) identified as a PSW trainee or student, 1 (2%) identified as a health care aide, 1 (2%) identified as a personal care assistant, 4 (8%) identified as other front line health care workers, 1 (2%) identified as unregulated care provider manager, and 3 (6%) identified as other (while with no further details on what their role was, care settings identified were retirement home or LTC [recruited from the Durham College cohort], group home [recruited from the uCarenet cohort], and possibly for a family member [recruited from the uCarenet cohort] for these 3 participants). Participant demographics broken down by the 3 recruitment cohorts are outlined in Table 2.

XSL•FO
RenderX

Table . Participant demographics in each recruitment cohort.

Characteristics	Number of responses per recruitment cohort		
	City of Ottawa LTC[a] (n=20), n (%)	Durham College (n=17), n (%)	uCarenet (n=13), n (%)
Age (years)			
18 to 24	1 (5)	2 (12)	0 (0)
25 to 34	2 (10)	9 (53)	5 (38)
35 to 44	9 (45)	3 (18)	4 (31)
45 to 54	5 (25)	3 (18)	2 (15)
55 to 64	3 (15)	0 (0)	2 (15)
Highest level of education completed			
High school	0 (0)	2 (12)	0 (0)
College or university	17 (85)	15 (88)	13 (100)
Graduate school	3 (15)	0 (0)	0 (0)
Role			
PSW[b]	19 (95)	0 (0)	5 (38)
PSW trainee or student	0 (0)	16 (94)	0 (0)
Health care aide	1 (5)	0 (0)	0 (0)
Personal care assistant	0 (0)	0 (0)	1 (8)
Other front-line health care worker	0 (0)	0 (0)	4 (31)
Unregulated care provider manager or administrator	0 (0)	0 (0)	1 (8)
Other	0 (0)	1 (6)	2 (15)

[a]LTC: long-term care.

[b]PSW: personal support worker.

Program Usability, Usefulness, and Satisfaction

Overall, the courses were positively rated, with most participants agreeing that the course was “very important” to their professional and learning needs (Table 3). Participants agreed that the training benefited them by “improving the health and well-being of the people that they care for.” Additionally, participants identified that the training allowed them to “feel more confident,” “prevent a problem,” “handle a problem or the worsening of a problem,” and “decide something with someone else.” The participants were also asked to provide any additional comments about the course; many open-text responses were very positive. Participants identified the importance and applicability of the information to their professional work.

In general, participants reported high levels of satisfaction with the program. A total of 46 of 50 (92%) of participants agreed or strongly agreed that the web-based training met their expectations, 47 (94%) agreed or strongly agreed that the training covered a broad range of topics and was not missing any important content, 49 (98%) agreed or strongly agreed that the web-based training would benefit them, and 46 (92%) agreed or strongly agreed that the program would be very valuable with current or future employers. No support inquiries were received from participants, further confirming the high usability of the program.

XSL•FO
RenderX

Table . Summary of postcourse feedback.

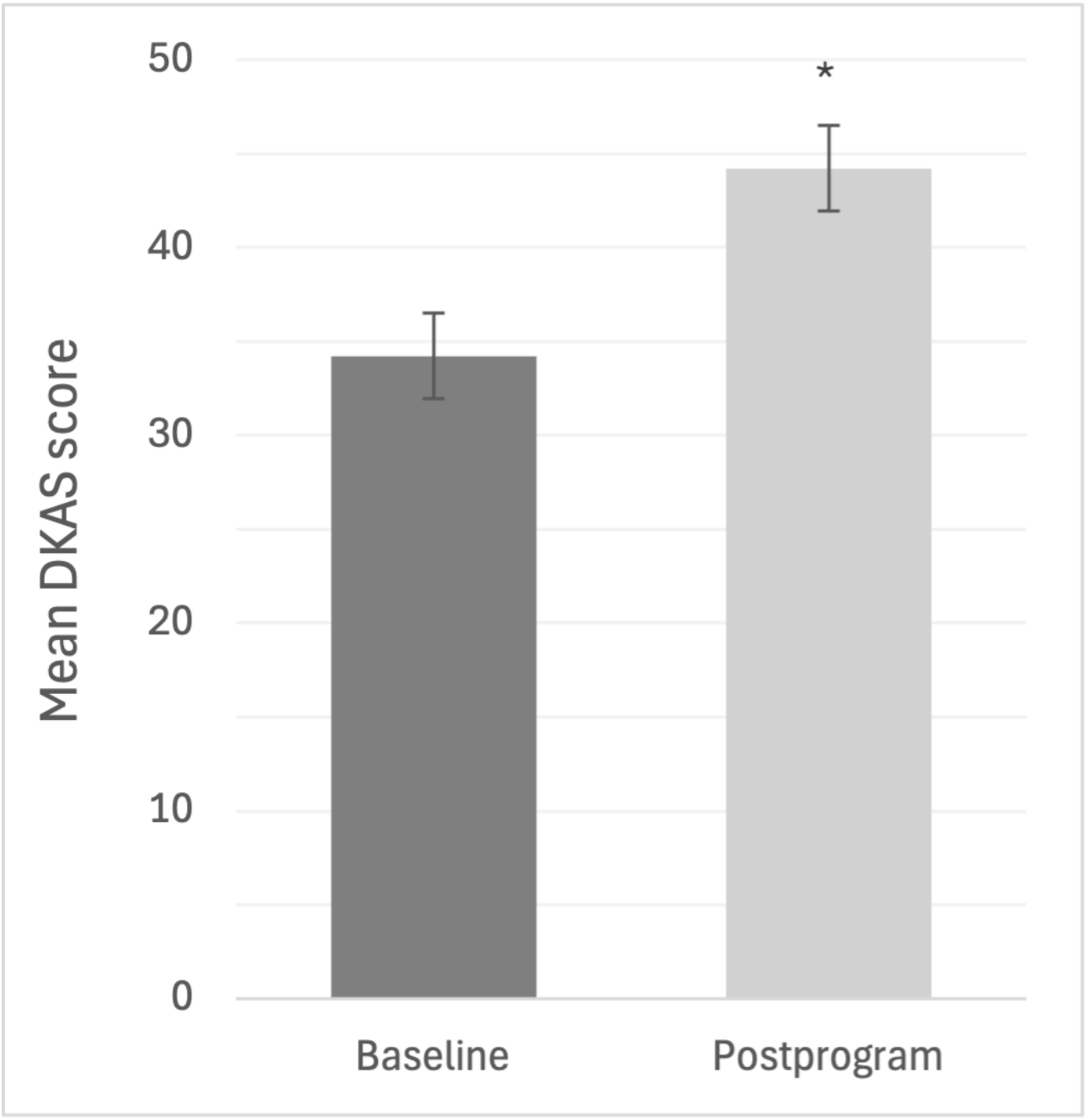
Question and course	Response breakdown, n (%)				
	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I thought the course content was very important to my profession/learning needs					
Course 1 (Foundations of Dementia)	4 (8)	0 (5)	1 (2)	18 (36)	27 (54)
Course 2 (Responsive Behaviours and Mental Health)	0 (0)	0 (0)	0 (0)	20 (40)	30 (60)
Course 3 (Home Supports and Safety)	1 (2)	1 (2)	1 (2)	19 (38)	28 (56)
Course 4 (Promoting Brain Health and Caregiver Wellness)	0 (0)	1 (2)	1 (2)	28 (56)	20 (40)
I was able to complete the course in a reasonable amount of time.					
Course 1 (Foundations of Dementia)	0 (0)	0 (0)	7 (14)	24 (48)	19 (38)
Course 2 (Responsive Behaviours and Mental Health)	0 (0)	0 (0)	2 (4)	29 (58)	19 (38)
Course 3 (Home Supports and Safety)	0 (0)	1 (2)	5 (10)	22 (44)	22 (44)
Course 4 (Promoting Brain Health and Caregiver Wellness)	1 (2)	0 (0)	3 (6)	28 (56)	18 (36)
I would recommend this course to a colleague.					
Course 1 (Foundations of Dementia)	0 (0)	0 (0)	2 (4)	13 (25)	35 (70)
Course 2 (Responsive Behaviours and Mental Health)	1 (2)	0 (0)	2 (4)	19 (38)	28 (56)
Course 3 (Home Supports and Safety)	0 (0)	0 (0)	0 (0)	19 (38)	31 (62)
Course 4 (Promoting Brain Health and Caregiver Wellness)	0 (0)	0 (0)	1 (2)	20 (40)	29 (58)

Change in Dementia Knowledge

All 50 participants completed the baseline and postprogram DKAS. One participant from the City of Ottawa LTC recruitment cohort was missing data for 15/25 items on the baseline DKAS due to a technical glitch. The most conservative approach was taken where it was assumed the participant would have received 2 points for each missing answer (ie, the highest score possible). The average baseline DKAS score was 34.2 of 50 (68%, SD 10.5) and the average postprogram score was 44.2 of 50 (88.4%, SD 4.7). Scores for all participants are shown in Figure 1, with scores for each recruitment cohort shown in Figure 2.

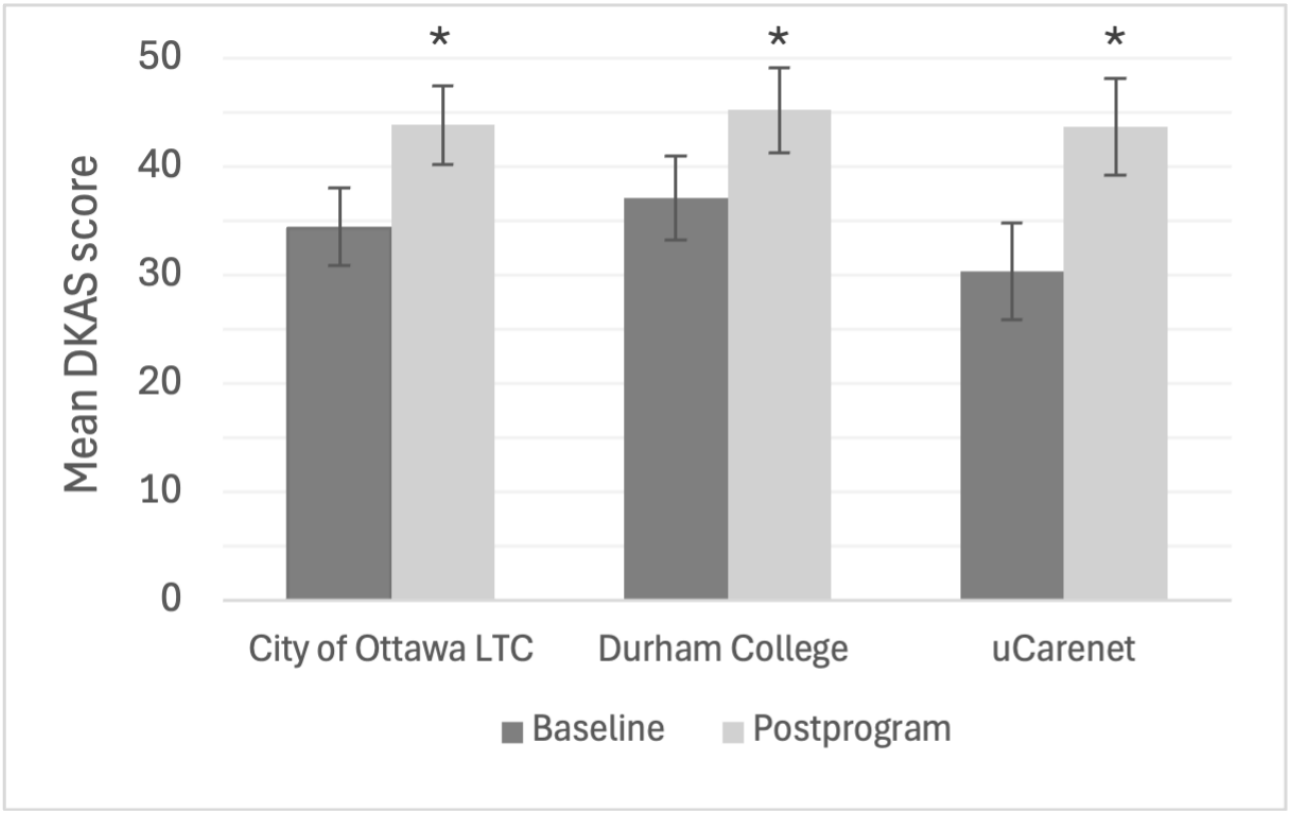
XSL•FO
RenderX

Figure 1. Baseline and postprogram DKAS scores for all participants. Results of repeated measures ANOVA (mean with 95% CI) with time as the only predictor of DKAS, including a random effect for participant ID. *Represents statistical significance (*P*<.001). DKAS: Dementia Knowledge Assessment Scale.



XSL•FO
RenderX

Figure 2. Baseline and postprogram DKAS by cohort and time. Results of repeated measures DKAS model (mean with 95% CI) with cohort time interaction, including a random effect for participant ID. *Represents statistical significance (P<.05). DKAS: Dementia Knowledge Assessment Scale; LTC: long-term care.



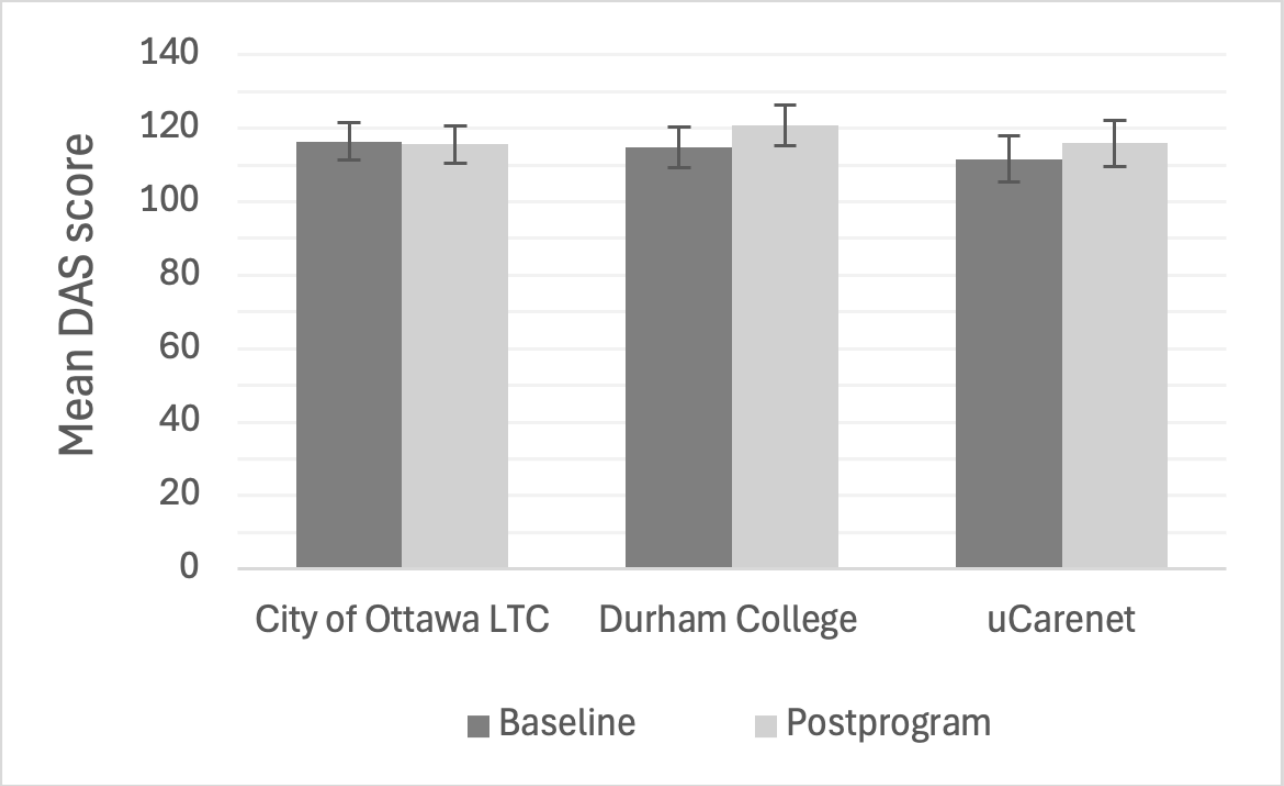
Repeated measures ANOVA showed a statistically significant effect on DKAS scores ($F_{1,49}$=58.533, P<.001), with scores increasing from baseline to postprogram across all 3 recruitment cohorts (η2=0.276). The model, including only time as a predictor and a random effect for participant ID, demonstrated that relative to baseline, there was an average 9.98-point increase on the DKAS score postintervention (95% CI 7.88-12.08; SE 1.065; P<.001). Post hoc testing did not reveal any significant differences in DKAS scores between recruitment cohorts; the model including an interaction with time and cohort, and a model with just time (not including cohort at all), were not statistically significantly different (P=.24).

Change in Dementia Attitudes

All 50 participants completed the baseline and postprogram DAS. On average across all recruitment cohorts, participants scored 114.7 of 140 (82%; SD 11.6) at baseline and 117.5 of 140 (84%; SD 11.2) postprogram. Scores for each recruitment cohort are shown in Figure 3.

XSL•FO
RenderX

Figure 3. Baseline and postprogram DAS scores among participants by cohort. Results from repeated measures DAS model (mean and 95% CI) including interaction for cohort and time. DAS: Dementia Attitudes Scale; LTC: long-term care.



Repeated measures ANOVA showed a statistically significant effect on DAS scores ($F_{1,49}$=4.383, P=.04; η2= 0.015). The model including only time as a predictor demonstrated that relative to baseline, there was an average 2.8 point increase on the DAS score postprogram (P=.04). Post hoc testing did not reveal any significant differences in DAS scores between recruitment cohorts; the model including an interaction with time and cohort, and a model with just time (not including cohort at all) were not statistically significantly different (P=.15).

Discussion

Principal Findings

The purpose of this pilot study was to evaluate the usability, usefulness, satisfaction, and effectiveness of the Dementia Foundations Program, a series of 4 web-based courses about various aspects of dementia developed for unregulated care providers such as PSWs. Three sites were recruited from slightly different populations (PSW trainees, PSWs working in LTC, and PSWs and other unregulated care providers working in community settings) who were all willing to take part in recruitment efforts due to the perceived benefit of additional dementia care training for their students, employees, and members. Overall, the program was positively valued for its usability, and participants reported a high level of satisfaction with the program. Participant knowledge, as measured by the DKAS, significantly increased after completing the program, with an average 30% improvement across all recruitment cohorts. A statistically significant increase was also seen for dementia attitudes as measured by the DAS, with a more modest average improvement of 2.5%. While scores for both the DKAS and DAS did not significantly differ between cohorts, there are a few trends worth highlighting.

PSW trainees from the Durham College cohort had higher baseline DKAS scores and slightly lower DAS scores compared to the PSWs from both the City of Ottawa LTC and uCarenet cohorts. The differences in knowledge scores could be explained by how trainees might have higher fact-based knowledge (from more recent studying) and may be more used to taking knowledge assessments, whereas practicing PSWs may have more everyday clinical knowledge and may not have taken a formal course in some time. When looking at dementia attitudes, the overall significant increase can be attributed to the Durham College and uCarenet cohorts, rather than the Ottawa LTC cohort. This conceptually makes sense, as practicing PSWs may already possess more developed attitudes toward dementia that would not be expected to change drastically after only 4 hours of knowledge-based education. Conversely, PSW trainees and PSWs not working in LTC homes who have had less exposure to working with people with dementia may be more prone to shifting their attitude in a significant way after 4 hours of education.

Comparison to Prior Work

Dementia Education

There is currently a lack of dementia-specific care training offered to unregulated care providers, leading to gaps in care [8,19]. This highlights the need for increased and enhanced training to improve education with the purpose of improving care for individuals with dementia. Research has shown that dementia education and training for health care workers improves knowledge, fosters positive attitudes, increases

XSL•FO
RenderX

confidence, and ultimately produces better outcomes for individuals living with dementia [20,21]. An extensive review by Surr et al [22] evaluated various features of dementia education programs for health care providers. Among the findings, the authors concluded that effective dementia education programs should be relevant to the workers' role and experience, include practice-based learning with theoretical content, be delivered by an experienced facilitator, and involve in-person facilitation. Although there are some existing on-the-job dementia training programs for unregulated care providers, they vary in availability, are more expensive to deliver, and are often less convenient for care providers who may not have flexible schedules to accommodate this type of instruction. The COVID-19 pandemic also imposed limitations on access to many in-person education programs. Thus, a blended knowledge approach that combines in-person training with web-based learning can deliver feasible and effective dementia education with the potential to be scaled and spread to a large population.

Furthermore, the recommendation by Surr et al [22] for in-person delivery may be influenced by several factors. For example, it is likely the case that most education and training programs about dementia have historically been delivered in person. Second, those programs delivered digitally (whether synchronously through webinar platforms or asynchronously) may not have incorporated best practices concerning the instructional design of multimedia e-learning. High-quality instructional design has been shown to improve learning transfer. The Dementia Foundations Program was created using evidence-based principles for multimedia learning outlined by Mayer [12] and Clark and Mayer [13]. The findings from the current study indicate that this entirely asynchronous web-based program is a useful and effective training modality for improving dementia-related knowledge among unregulated care providers. The high satisfaction with the program is evident in the course open-text comments, where most participants praised the program for its informativeness, user-friendliness, and usefulness to their profession as unregulated care providers.

About DKAS

Significant improvements in dementia knowledge were seen after the intervention for both PSWs and PSW trainees. As unregulated care providers in Canada have noted a lack of dementia-specific care training and education, this indicates that the Dementia Foundations Program is an effective way to impact dementia knowledge for this group. Other web-based dementia education programs have also been shown to be effective in increasing dementia knowledge. Eccleston et al [23] evaluated the efficacy of the 9-week-long Understanding Dementia Massive Open Online Course (UDMOOC) for people with varying professional and educational backgrounds. Similar to the Dementia Foundations Program, the UDMOOC aims to improve dementia knowledge by teaching a broad community audience basic neurobiology, dementia pathophysiology, medical management, and person-centered care. The median DKAS score at baseline for their study was 34.5 out of 50, which is comparable to our current pilot study's baseline median DKAS score of 35.5 out of 50, implying both groups started with similar levels of dementia knowledge before embarking on the training. The median postprogram scores for both studies were 45 of 50, both demonstrating a significant increase in dementia knowledge. These results are consistent with the current study's findings, indicating that web-based dementia education modalities are effective at increasing dementia knowledge. However, it is worth noting that the Dementia Foundations Program can be completed in a single day with only 4 hours' worth of content, whereas the UDMOOC estimates 21 hours over a 7- to 9-week period. Comparable baseline and postprogram median DKAS scores for both programs suggest the Dementia Foundations Program is more efficient for busy professionals, with similar knowledge gain in a much shorter amount of time.

When considering the mode of delivery of instruction, 1 study conducted by Parveen et al [24] also found that the most impactful dementia training programs involved a blend of both e-learning and in-person delivery, although another study conducted by Vollmar et al [25] found no significant differences between learning through in-person compared with web-based learning. With the significant increase in dementia knowledge shown by both the current study and Eccleston et al using only a web-based approach, it is reasonable to consider digital-only approaches to some aspects of dementia education, especially when considering the increased flexibility this offers to an unregulated care worker audience with potentially variable and less flexible schedules. A blended approach has been shown to be effective in other studies, and one can speculate that an efficient and effective approach might involve high-quality asynchronous web-based approaches for baseline knowledge education, complemented by in-person skills training.

About DAS

Dementia attitudes significantly increased from baseline to postprogram, with scores moving from an average of 114.7 to 117.5. While this increase is modest at 2.5%, baseline scores on the DAS started quite high compared to other studies looking at dementia attitudes in nursing students in Malta (mean DAS score of 107.9) [26] and health care professionals in China (mean DAS score of 91.3) [27]. The highest possible score on the DAS is 140, corresponding to the highest possible positive attitude toward dementia. The DAS scores from the current study indicate that this sample of participants already possesses relatively positive attitudes toward dementia compared to previous studies, therefore leaving only modest room for improvement on this measure.

Baseline DAS scores averaged 114.7 across all 3 recruitment cohorts, with uCarenet starting with the lowest score of 111.6 and the City of Ottawa LTC starting with the highest score at 116.4. After the 4-hour intervention, DAS scores were significantly higher at 117.5 across all 3 recruitment cohorts, with the Durham College PSW trainees showing the largest increase among cohorts from 114.9 to 120.9. Logically, this cohort showed the biggest increase in positive attitudes toward dementia, as trainees have very likely spent less time working with older adults and those with dementia. The experienced PSWs and other paid caregivers that comprised the Ottawa LTC and uCarenet cohorts have likely been working with people with dementia for many years and have had ample time to develop personal attitudes toward dementia, as previous research

XSL•FO
RenderX

has shown that working with and having ongoing and meaningful interactions with individuals living with dementia are key to fostering a change in attitudes [26,28]. It can therefore be reasonably assumed that 4 hours' worth of independent learning is not a sufficient intervention length to cause an attitude change in experienced workers, but remains sufficient exposure to improve attitudes toward dementia among trainees. This contrasts with a study that explored the impact of a single dementia awareness session on changing dementia attitudes among adolescents, with no significant improvements in either the intervention or control groups [29]. However, the difference may be that the Durham College PSW trainees were more accepting of a potential attitude change toward dementia given their desired career choice as a PSW.

Limitations

We acknowledge that there are limitations to this study. First, the relatively small sample size of 50 PSWs, PSW trainees, and other unregulated care professionals from Ontario who took part in this study is not necessarily representative of the diverse and large population of unregulated care providers in Canada or internationally. However, as the sample size was based on the typical large effect sizes for e-learning interventions, the results of this study remain valuable for an initial assessment of a new e-learning program for this population. This was further confirmed by the detection of statistically significant changes in both dementia knowledge and attitudes as measured by the DKAS and DAS. Second, the lack of a control group due to the pre-post study design introduces an unknown risk of bias; however, the pre-post design does alleviate intraparticipant variability and is consistent with the majority of studies in a 2019 systematic review of studies assessing technology-delivered dementia education to health care providers [30]. Lastly, not all participants had active experience working with people living with dementia, and as a result, may not have found the learning content to be relevant. However, trainees were intentionally included to be able to assess the program's suitability to incorporate into a training program, rather than continuing education once employed.

Conclusions and Future Directions

The Dementia Foundations Program was positively evaluated by a group of 50 unregulated care providers, predominantly PSWs and PSW trainees, with participants showing a significant improvement in both dementia knowledge and attitudes. Participants found the training very useful in supplementing their existing education and training, with particular relevance to the PSW trainee cohort. The findings of this study have implications for unregulated care providers who support people living with dementia, as well as trainees who may not yet interact with those living with dementia. The Dementia Foundations Program is an effective and user-friendly e-learning program that can be conveniently incorporated within existing dementia training and educational programs as part of a blended teaching and learning strategy to enhance dementia knowledge in unregulated care providers and trainees. Widespread implementation has the potential to increase training capacity and fill gaps in existing workforce training and improve dementia knowledge, attitudes, and awareness among care providers as dementia prevalence increases in Canada.

Since this pilot study was conducted, the Dementia Foundations Program has reached 1012 users from December 2021 to September 2024, with approximately 55% completing all 4 courses and receiving a certificate. Postprogram evaluations show high satisfaction with the courses from individuals who purchased the program digitally and employees and students from organizations who purchased the program for professional development.

Future work will focus on the spread and scale of this training program through advertising and partnerships with relevant organizations and institutions, and future research should use implementation science to assess organizational adoption and effectiveness. Continuous quality improvement evaluations will be conducted to determine the best approach for widespread implementation.

Acknowledgments

The authors disclose receipt of the following financial support for the research, authorship, or publication of this paper: This project received funding from the Centre for Aging & Brain Health Innovation (CABHI). AJL receives support through the John R Evans Chair in Health Sciences Educational Research and Instructional Development, Faculty of Health Sciences, McMaster University. The authors wish to thank Sheila Bauer, BScN, MHA, and the City of Ottawa LTC, Deborah Schuh, BScN, and the Durham College PSW Training Program, and Nectari Charitakis, BA, and uCarenet for their collaboration and assistance with recruitment. The authors also wish to thank Dr David Streiner for his guidance and mentorship throughout the statistical analysis.

Data Availability

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

Authors' Contributions

AJL and SA designed this study and led recruitment and data collection. SA, AJL, and AS performed data analyses. SA, SC, and PG drafted this paper. AS and SC created the figures. All authors read and approved this final paper.

XSL•FO
RenderX

Conflicts of Interest

AJL and RS are co-owners of the Dementia Foundations Program. There is a fee to access the Dementia Foundations Program.

References

1. Ageing and health. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2025-06-17]
2. A portrait of canada's growing population aged 85 and older from the 2021 census. Statistics Canada. URL: https://www12.statcan.gc.ca/census-recensement/2021/as-sa/98-200-X/2021004/98-200-X2021004-eng.cfm [accessed 2025-06-17]
3. Nuernberger K, Atkinson S, MacDonald G. Seniors in transition: exploring pathways across the care continuum. Healthc Q 2018 Apr;21(1):10-12. [doi: 10.12927/hcq.2018.25529] [Medline: 30051809]
4. Dementia numbers in Canada. Alzheimer Society. URL: https://alzheimer.ca/en/about-dementia/what-dementia/dementia-numbers-canada [accessed 2025-06-17]
5. PSW positions and job titles. Caring Support. URL: https://www.caringsupport.com/blog/consider-these-10-alternative-psw-positions-and-job-titles [accessed 2025-06-17]
6. Zagrodney K, Saks M. Personal support workers in Canada: the new precariat? Healthc Policy 2017 Nov;13(2):31-39. [doi: 10.12927/hcpol.2017.25324] [Medline: 29274225]
7. PSW roles & responsibilites. Ontario Personal Support Workers Association. URL: https://ontariopswassociation.com/psw-roles-and-responsibilities [accessed 2025-06-17]
8. Lum J, Sladek J, Ying A. Ontario personal support workers in home and community care: CRNCC/PSNO survey results. : CRNCC/RCRSC; 2009 URL: https://www.torontomu.ca/content/dam/crncc/knowledge/infocus/factsheets/InFocus-Ontario%20PSWs%20in%20Home%20and%20Community%20Care.pdf [accessed 2025-06-17]
9. Dementia in Canada: summary. Canadian Institute for Health Information. 2016. URL: https://www.cihi.ca/en/dementia-in-canada/dementia-in-canada-summary [accessed 2025-07-03]
10. Hapsari AP, Ho JW, Meaney C, et al. The working conditions for personal support workers in the Greater Toronto Area during the COVID-19 pandemic: a mixed-methods study. Can J Public Health 2022 Dec;113(6):817-833. [doi: 10.17269/s41997-022-00643-7] [Medline: 35616873]
11. Personal support worker standard. : Ontario Ministry of Colleges and Universities; 2022 URL: https://files.ontario.ca/mcu-personal-support-worker-jan2022-41469-en-2022-03-31.pdf [accessed 2025-06-17]
12. Mayer RE. Multimedia learning. In: Psychology of learning and motivation: Elsevier; 2002, Vol. 41:85-139.
13. Clark RC, Mayer RE. E-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning: John Wiley & Sons; 2016.
14. Lewis M, Bromley K, Sutton CJ, McCray G, Myers HL, Lancaster GA. Determining sample size for progression criteria for pragmatic pilot RCTs: the hypothesis test strikes back!. Pilot Feasibility Stud 2021 Feb 3;7(1):40. [doi: 10.1186/s40814-021-00770-x] [Medline: 33536076]
15. In J. Introduction of a pilot study. Korean J Anesthesiol 2017 Dec;70(6):601-605. [doi: 10.4097/kjae.2017.70.6.601] [Medline: 29225742]
16. Pluye P, El Sherif R, Granikov V, et al. Health outcomes of online consumer health information: a systematic mixed studies review with framework synthesis. J Assoc Inf Sci Technol 2019 Jul;70(7):643-659. [doi: 10.1002/asi.24178] [Medline: 31423458]
17. Annear MJ, Toye C, Elliott KEJ, McInerney F, Eccleston C, Robinson A. Dementia Knowledge Assessment Scale (DKAS): confirmatory factor analysis and comparative subscale scores among an international cohort. BMC Geriatr 2017 Jul 31;17(1):168. [doi: 10.1186/s12877-017-0552-y] [Medline: 28760154]
18. O'Connor ML, McFadden SH. Development and psychometric validation of the Dementia Attitudes Scale. Int J Alzheimers Dis 2010;2010:1-10. [doi: 10.4061/2010/454218]
19. Janes N, Sidani S, Cott C, Rappolt S. Figuring it out in the moment: a theory of unregulated care providers' knowledge utilization in dementia care settings. Worldviews Evid Based Nurs 2008;5(1):13-24. [doi: 10.1111/j.1741-6787.2008.00114.x] [Medline: 18333923]
20. Liu JYW, Lai C, Dai D, Ting S, Choi K. Attitudes in the management of patients with dementia: comparison in doctors with and without special training. East Asian Arch Psychiatry 2013 Mar;23(1):13-20. [Medline: 23535628]
21. Jennings A, McLoughlin K, Boyle S, et al. Development and evaluation of a primary care interprofessional education intervention to support people with dementia. J Interprof Care 2019;33(5):579-582. [doi: 10.1080/13561820.2018.1541876] [Medline: 30422731]
22. Surr CA, Gates C, Irving D, et al. Effective dementia education and training for the health and social care workforce: a systematic review of the literature. Rev Educ Res 2017 Oct;87(5):966-1002. [doi: 10.3102/0034654317723305] [Medline: 28989194]
23. Eccleston C, Doherty K, Bindoff A, Robinson A, Vickers J, McInerney F. Building dementia knowledge globally through the understanding dementia Massive Open Online Course (MOOC). NPJ Sci Learn 2019;4(1):3. [doi: 10.1038/s41539-019-0042-4] [Medline: 30993003]

XSL•FO
RenderX

24. Parveen S, Smith SJ, Sass C, et al. Impact of dementia education and training on health and social care staff knowledge, attitudes and confidence: a cross-sectional study. BMJ Open 2021 Jan 19;11(1):e039939. [doi: 10.1136/bmjopen-2020-039939] [Medline: 33468498]
25. Vollmar HC, Mayer H, Ostermann T, et al. Knowledge transfer for the management of dementia: a cluster randomised trial of blended learning in general practice. Implement Sci 2010 Jan 4;5(1):1. [doi: 10.1186/1748-5908-5-1] [Medline: 20047652]
26. Scerri A, Scerri C. Nursing students' knowledge and attitudes towards dementia - a questionnaire survey. Nurse Educ Today 2013 Sep;33(9):962-968. [doi: 10.1016/j.nedt.2012.11.001] [Medline: 23182308]
27. Zhao W, Moyle W, Wu MLW, Petsky H. Hospital healthcare professionals' knowledge of dementia and attitudes towards dementia care: a cross-sectional study. J Clin Nurs 2022 Jul;31(13-14):1786-1799. [doi: 10.1111/jocn.15590] [Medline: 33295010]
28. Tan WJ, Yeo D, Koh HJ, Wong SC, Lee T. Changing perceptions towards dementia: how does involvement in the arts alongside persons with dementia promote positive attitudes? Dementia (London) 2021 Jul;20(5):1729-1744. [doi: 10.1177/1471301220967805] [Medline: 33100055]
29. Farina N, Hughes LJ, Jones E, et al. The effect of a dementia awareness class on changing dementia attitudes in adolescents. BMC Geriatr 2020 Jun 2;20(1):188. [doi: 10.1186/s12877-020-01589-6] [Medline: 32487024]
30. Scerbe A, O'Connell ME, Astell A, Morgan D, Kosteniuk J, DesRoches A. Digital tools for delivery of dementia education for health-care providers: a systematic review. Educ Gerontol 2019 Nov 2;45(11):681-699. [doi: 10.1080/03601277.2019.1687149]

Abbreviations

DAS: Dementia Attitudes Scale
DKAS: Dementia Knowledge Assessment Scale
LTC: long-term care
PSW: personal support worker
UDMOOC: Understanding Dementia Massive Open Online Course

Edited by A Garcia; submitted 24.10.24; peer-reviewed by B Davis, M Sengupta; revised version received 22.05.25; accepted 26.05.25; published 21.07.25.

Please cite as:
Levinson AJ, Ayers S, Clark S, Gerantonis P, Schneeberg A, Sztramko R
Usability and Impact of the Web-Based Dementia Foundations Educational Program in Personal Support Workers (PSWs), PSW Trainees, and Care Companions: Quasi-Experimental Study
JMIR Aging 2025;8:e67889
URL: https://aging.jmir.org/2025/1/e67889
doi:10.2196/67889

© Anthony J Levinson, Stephanie Ayers, Sandra Clark, Patricia Gerantonis, Amy Schneeberg, Richard Sztramko. Originally published in JMIR Aging (https://aging.jmir.org), 21.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Implementation of New Technologies in an Aged Care Social Day Program: Mixed Methods Evaluation

Dannielle Post, PhD; Kathleen Whitson, BPsychSc (Hons); Gaynor Parfitt, Prof, PhD

Alliance for Research in Exercise, Nutrition, and Activity, Allied Health and Human Performance, University of South Australia, City East Campus, Frome Road, GPO Box 2471, Adelaide, Australia

Corresponding Author:
Gaynor Parfitt, Prof, PhD
Alliance for Research in Exercise, Nutrition, and Activity, Allied Health and Human Performance, University of South Australia, City East Campus, Frome Road, GPO Box 2471, Adelaide, Australia

Abstract

Background: Australia's aging population is looking to age in place, accessing care alternatives external to the traditional model of residential aged care facilities. This evaluation is situated in a Social Day Program, delivered by an aged care organization. It is designed to cater for people living with dementia, located in an environment equipped with new technologies including age-specific interactive computer gaming, social robots, sensory stimulation, and virtual reality. The technologies are designed to support older adults, enabling them to stay connected and maintain physical and cognitive functioning, independence, and quality of life.

Objective: This project aimed to undertake a multifaceted evaluation of the implementation of the new technologies, including an exploration of the barriers and enablers to uptake. The key issue is how to enhance the potential for optimizing the use of these technologies in the Social Day Program environment, to help inform decision-making regarding the implementation of these technologies at the organization's other sites, and future investment in such technologies by aged care organizations generally.

Methods: Observation of technology use within the organization was conducted over a 16-week period. Surveys and semistructured interviews were used to collect information from staff related to their experiences with the technology. Thematic analysis was used to analyze the interviews. Data were triangulated across the sample.

Results: Forty-eight observation periods were completed, totaling 126.5 observation hours. Technology use by clients was observed on 24 occasions, for 22 (17.4% of the observation time) hours. Nineteen staff completed surveys. Nearly three-quarters (n=14) of the staff perceived there to be barriers to the clients' use of technology, and 18 (95%) staff reported that they assisted clients to use the technology. Ten (53%) staff reported receiving training to use the technology and feeling confident in their knowledge of the technology to assist clients in using it. Twelve staff members participated in an interview. Key themes identified from the interview data were: technology has potential but is not for everyone, incorporating the subtheme technology as a placation tool, staff knowledge and confidence, and technology functionality and support.

Conclusions: This evaluation identified that technology was not being used for the purposes of enrichment or experience enhancement, nor extensively. Multiple barriers to the implementation and sustained use of the technology items were identified. Recommendations to improve implementation and promote sustained use of technology, based on the findings of this evaluation and evidence from the literature, may apply to other organizations seeking to implement these technologies in similar programs.

(JMIR Aging 2025;8:e60297) doi:10.2196/60297

KEYWORDS

aged care; older adults; interactive robots; social engagement; evaluation; geriatric; robot; day program; perception

Introduction

Background

With a shift to consumer-directed care, and Australia's aging population looking to age in place and access a range of supported care alternatives outside the traditional model of residential aged care facilities, there are a lot of new offerings for consumers. Often funded through the Australian government–supported Aged Care Community Packages, these offerings can include in-home care and home support services or colocated aging-specific services in centralized hubs. Services include Dietetics, Allied Health Group Programs, Exercise Physiology, Gym and Physical Activity, Massage, Music Therapy, Occupational Therapy, Physiotherapy, Podiatry, Speech Pathology, Gerontology, and Social Day Programs.

ECH Inc (hereinafter referred to as the organization) is a South Australian retirement living and in-home care services provider established in 1964 and based in the state of South Australia. The "profit for purpose" organization's mission is "helping

XSL•FO
RenderX

people to live confidently and independently, and to get more out of life" [1]. This project was situated in one of the organization's Social Day Programs, which occur at 4 centers across metropolitan South Australia. The Social Day Program involved in this evaluation is situated within a proof-of-concept Care Hotel that was purpose-designed to provide care givers respite options with 8 short-stay suites for people living with dementia or who may be recovering from surgery. These Social Day Programs, provided on weekdays between 9 AM and 3 PM, are designed to cater for community-dwelling people living with dementia or memory loss and are located in environments where the clients have space to engage socially with others or have supported, quiet alone time. Clients can be dropped off at and picked up from the Social Day Program by family members or transported to and from the Social Day Program by a bus provided by the organization. Program staff are trained to ensure that each client's interests and abilities are met, with the provision of activities that are enjoyable and meaningful. Participation in the Social Day Program is also available to Care Hotel guests, as is access to the space outside the Social Day Program hours (evenings and weekends). This space is equipped with new technologies that are designed to support older adults as they age, enabling them to stay connected and maintain physical and cognitive functioning, independence, and quality of life [2-5]. These technologies are varied and includes an age-specific interactive computer gaming console, "Obie" that projects interactive games onto a surface (table or floor) to encourage active play; a relaxation chair in a sensory stimulation room; virtual reality (VR) headsets (the Odysee system, with virtual experiences including travel, museums, and hot air ballooning); white noise bubble tubes (tubes with color changing lights, filled with bubbles, that are intended to provide visual distraction and calm); 2 robotic dolls; and a robotic sensory cat.

Systematic reviews support the effectiveness of these technologies, with interactive computer gaming improving physical and psychological functioning [6]; interactions with social robots positively affecting agitation, anxiety and loneliness, medication consumption, and quality of life [5]; and VR technology improving physical health outcomes [7] and cognition, memory, and depression [8] for older adults. The implementation of such technologies in the aged care environment is not without its challenges. Other studies have identified organizational factors, such as funding and staff engagement or knowledge, client-specific issues such as frailty, dementia, or limited prior exposure to technology, inappropriate or unsuitable technologies, or resistance from family members [9]. Similar barriers were identified in a systematic review of VR technology, with the addition of the potential for cybersickness during use also perceived to be a barrier [10]. Research focused on social and assistive robots has identified concerns around the potential for infantilization or loss of dignity for older adults who use them [11]. The items chosen by the organization were selected due to the evidence supporting their effectiveness and with the intention of supporting engagement and enrichment for their clients. Providing a variety of technologies, rather than a single type of technology, was intended to broaden the appeal of technology to clients.

This Study

The organization requested an extensive review of the use of these technologies in this unique setting. This research aimed to undertake a multifaceted evaluation of the implementation of technologies at the Social Day Program, including an exploration of the barriers and enablers to the use of technology within the Social Day Program. The key issue, and the focus of this paper, is identifying how to enhance the potential for optimizing the use of these technologies at the organization, making recommendations to help inform decision-making regarding the implementation of these technologies at other sites, and future investment in new technologies for the organization's clients.

The evaluation was intended to answer the following research questions:

1. What is the utility of these technologies?
2. What are the barriers and enablers to the uptake of these technologies?
3. What is the impact (physical, psychosocial, and quality of life) of technology use experienced by the clients, as perceived by the staff?

Methods

Study Design

Observation of technology use within the Social Day Program was conducted over 16 weeks by a research assistant (RA), with periodic support from undergraduate physiotherapy students completing a health promotion placement with the evaluation team. Surveys and semistructured interviews were used to collect information from staff related to their experiences with the technology. Data were triangulated across the sample, with observation, survey, and interview data integrated and compared to establish alignment and contrast among the findings of each method.

Participants

Participants were staff working in the organization's Care Hotel or Social Day Program. Staff provided written informed consent to participate in the evaluation. For the observation component, participants included clients of the Social Day Program and clients of the Care Hotel who were accessing the Social Day Care space.

Observation of Technology Use

Observation of the use of technology by staff, Social Day Program clients, and Care Hotel guests using the Social Day Program space was undertaken between July and November 2022. Observation periods of 4 hours each, either morning and overlapping early afternoon or late morning until the end of the day, occurred during the Social Day Program sessions across the week, but also after hours and on weekends to ensure capture of use of technology outside of typical program activity periods. An observation checklist was used to capture data related to technology-user type (ie, client, family member or carer, staff, or other), technology item, duration of use, concurrent use of an item by multiple participants, support or assistance provided, emotional impacts, communication, physical impacts, and other

XSL•FO
RenderX

notable information. This checklist was developed by the evaluation project team in consultation with the Stakeholder Advisory Group and piloted by the RA and placement student observers, with checklist items modified as required throughout piloting.

Surveys

Consenting staff completed an anonymous electronic survey (Qualtrics). Survey items were initially drafted by the evaluation team and refined using a co-design approach in collaboration with the Stakeholder Advisory Group. The Stakeholder Advisory Group was comprised of members of the organization and evaluation project team and a representative from each of the 3 stakeholder-participant groups (clients, family members or carers, and staff of the Social Day Program) and was formed to ensure consumer awareness and the suitability and feasibility of evaluation methods. The staff survey was intended to identify the perceived implications for staff involved in implementing the technology and supporting clients to use the technology.

Survey questions collected information about participant demographics, staff assistance with technology items, length of time staff assisted clients with the technology, and most- and least-assisted technology items. Perception of impact were collected across 6 domains relative to each of the items of technology: social engagement, cognitive awareness, communication, mood state, activities of daily living, and physical mobility. Responses were on a 4-point Likert scale, with responses being: don‘t know, no improvement, some improvement, and a lot of improvement. Additionally, there were questions relating to whether staff perceived that clients disliked any technology items; barriers to clients using the technology; staff training; staff confidence when supporting clients; and barriers to staff supporting clients with the technology items.

Interviews

Semistructured, once-off, one-on-one interviews were conducted by 1 of 2 University of South Australia RAs, with consenting staff. Interview questions were initially drafted by members of the evaluation project team and refined and co-designed in collaboration with the Stakeholder Advisory Group (Multimedia Appendix 1). Interviews were conducted either in person at the Social Day Program or over the phone. All interviews were recorded and then transcribed verbatim.

Data Analysis

Observation and survey data are reported as counts or counts and percentage responses as relevant. Content analysis of open-ended survey questions was intended; however, an insufficient number of responses was provided, and a descriptive approach to reporting the responses was instead used. Reflexive Thematic Analysis [12,13] was used to analyze interview data, which were coded independently by a member of the evaluation team. Semantic codes and candidate themes were reviewed and discussed among the team across multiple iterations before the final themes were decided. We applied a constructionist epistemology, with an experiential lens on the data, with the intention of highlighting the experiences and perspectives of staff related to the use of technology in the Social Day Program environment. While our approach was predominantly inductive, in that we sought to represent the meaning of the information as it was communicated by the staff, elements of a deductive approach were applied to ensure that the established themes aligned with the research questions, for example, what the staff perceived to be barriers to and enablers of technology use in this setting.

Ethical Considerations

This evaluation research was approved by University of South Australia's Human Research Ethics Committee (204457). A waiver of consent was approved for the purpose of observing technology use in the Social Day Program. This was on the proviso that before commencing each observation period, staff and the evaluation project team announced to clients during Circle Time—an activity during which information is relayed to clients and their family members—that an observation period was about to begin. This enabled clients to leave the room if they did not wish to participate in this part of the evaluation project. Participants remained anonymous, in that information that might identify them was not recorded per their use of technology, and they were not asked to modify their behavior for the evaluation. Participants provided written informed consent for use of their survey data. Written informed consent was also provided by interview participants. Each interview participant received an honorarium (AUS $30 gift card; US $19.43) to acknowledge their contribution to the project. Data are anonymized and no information is provided that would identify participants.

Results

Participants

Thirty-one staff were approached to participate in the evaluation project; 19 based in the Care Hotel, and 12 staff based in the Social Day Program. Nineteen (61.3%) staff completed surveys (9 Care Hotel and 10 Social Day Program). Ten (32.3%) of the 31 staff approached (3 Care Hotel and 7 Social Day Program) provided written informed consent to participate in the interview; all 10 interviews were completed.

Observation of Technology Use

Forty-eight observation periods were completed, totaling 126.5 observation hours. Technology use by clients was observed on 24 occasions, for 22 (17.4% of the observation time) hours. The relaxation chair and the robotic cat were the most frequently used technology items. The relaxation chair was used 7 times for a total of 9 hours and 36 minutes, and the robotic cat was used 7 times for a total of 3 hours and 14 minutes. Conversely, the Obie interactive games table and VR headsets were the least frequently used technology items (once for 64 min and once for 39 min, respectively). When technology was used, clients were mostly observed socially engaging and communicating with staff and other clients, with some clients also observed speaking to the robotic cat and dolls.

Surveys

Of the 19 staff members who consented and responded to the staff survey, 10 (53%) worked in the Social Day Program and

XSL•FO
RenderX

9 (47%) were Care Hotel staff. Ninety-five percent (n=18) of staff reported that they helped or supported clients in using the technology items. Social Day Program staff assisted clients more than Care Hotel staff across all technology items except for the relaxation chair. Furthermore, all 9 Social Day Program staff indicated they assisted clients with the Obie interactive games table and robotic cat. Seventeen of the nineteen (89%) staff members nominated the VR headsets as the technology item they engaged with the least. Staff reported that on average, clients engaged with the technology items from 10-minutes to more than 60-minutes. This finding is reflected in the observation data, with the average time clients spent using technology items ranging from 32-minutes to 96-minutes.

Improvements in social engagement, cognitive awareness, communication, mood state, activities of daily living, and physical mobility were seen to some extent across each of the technologies, although not for all items (Table 1).

XSL•FO
RenderX

Table . Perceived improvements associated with technology use.

Technology item and domain		Don't know, n (%)	No improvement, n (%)	Some or a lot of improvement, n (%)
Obie ICG[a] table				
	Social engagement	1 (5)	3 (16)	15 (79)
	Cognitive awareness	3 (16)	4 (21)	12 (64)
	Communication	1 (5)	2 (11)	16 (84)
	Mood state	1 (5)	0 (0)	18 (95)
	Activities of daily living	2 (11)	9 (47)	8 (42)
	Physical mobility	0 (0)	8 (42)	11 (58)
VR[b] headsets				
	Social engagement	10 (53)	4 (21)	5 (26)
	Cognitive awareness	12 (63)	5 (26)	2 (11)
	Communication	12 (63)	4 (21)	3 (16)
	Mood state	10 (53)	4 (21)	5 (26)
	Activities of daily living	11 (58)	7 (37)	1 (5)
	Physical mobility	12 (63)	6 (32)	1 (5)
Robotic cats				
	Social engagement	2 (11)	1 (5)	16 (84)
	Cognitive awareness	2 (11)	5 (26)	12 (63)
	Communication	2 (11)	2 (11)	15 (79)
	Mood state	2 (11)	0 (0)	17 (90)
	Activities of daily living	3 (16)	11 (58)	5 (26)
	Physical mobility	3 (16)	8 (42)	8 (42)
Robotic dolls				
	Social engagement	1 (5)	1 (5)	17 (90)
	Cognitive awareness	1 (5)	4 (21)	14 (74)
	Communication	3 (16)	1 (5)	15 (79)
	Mood state	2 (11)	0 (0)	17 (89)
	Activities of daily living	5 (26)	8 (42)	6 (32)
	Physical mobility	4 (21)	6 (32)	9 (47)
Relaxation chair				
	Social engagement	2 (11)	7 (37)	10 (52)
	Cognitive awareness	4 (21)	5 (26)	10 (53)
	Communication	3 (16)	5 (26)	11 (58)
	Mood state	1 (5)	0 (0)	18 (95)
	Activities of daily living	3 (16)	6 (32)	10 (53)
	Physical mobility	2 (11)	6 (32)	11 (58)

[a]ICG: interactive computer gaming.

[b]VR: virtual reality.

Nearly three-quarters (n=14) of the responding staff perceived there to be barriers to the clients' use of technology, with equal numbers responding "yes" when compared by staff type. Ten (53%) of the responding staff had received training. Ten (53%) staff indicated they felt confident when assisting clients to use the technology; however, ten (53%) staff also felt that barriers existed that inhibited them from assisting the clients to use the technology (5 each from the Social Day Program and Care Hotel).

XSL•FO
RenderX

Key concepts raised by staff in their responses to the open-ended questions about their perceptions of technology use within the Social Day Program related to why clients may dislike technology, the barriers to technology use by clients, factors related to staff confidence in supporting clients to use technology, and the barriers to staff supporting clients to use technology.

Staff perceived that clients may dislike the technology because they have health or cognition issues that limit their ability to engage with technology; due to negative perceptions or stigmatization by other clients of people who use the robotic doll or cat; that clients have no interest in technology or use of technology; and that clients did not like particular items, finding them to be noisy or creepy (ie, the robotic doll). These aspects were also considered to be barriers to clients' use of technology, with the addition of staff knowledge and confidence in the use of technology, and the time impact on staff to assist clients in using the technology.

Interviews

Interviews were conducted with 12 staff members to assess the use of technology within the Social Day Program, and the enablers of and barriers to technology use. Interviews ranged between 18 and 44 minutes. The three themes developed during the analysis of the interview data: (1) technology has potential but is not for everyone, with the subtheme technology as a placation tool; (2) staff knowledge and confidence in using technology; and (3) technology functionality and support, are based on our interpretation of how staff experienced technology use in the Social Day Program and the factors that are likely to be enablers or barriers that influence the adoption of technology and sustainability of technology use in this environment. These themes address the research questions that form the basis for the evaluation. Quotes from participants are used as illustrative examples of the themes established.

Technology Has Potential but it Is not for Everyone

This theme reflects the value and appropriateness that the staff perceive in technology use for their clients, assuming who might be likely to benefit from technology use and, conversely, for whom technology is not for. This theme incorporates the safety and personal considerations that were voiced as being needed when using technology. There were conflicting perspectives related to the use of technology within the Social Day Program. The potential benefits of the technology were raised often, including increased social engagement through the use of interactive games or being able to placate agitated clients with the robotic cat or time in the relaxation chair (a subtheme). While there was some concern about the cognitive benefits of the technologies, that the clients had to sit for prolonged periods to use some of them and therefore were being sedentary just as they would be at home (staff 14), and that some items had the potential to cause "sensory overload" (staff 2), others thought that the technology promoted social interaction, could "enrich the time clients spent at the Social Day Program" (staff 24), and could benefit future clients, in particular those who were tech-savvy. For one staff member, their initial perception that technology would "not be a great idea" (staff 1) changed positively when they saw the impact of technology use for some clients, for example, the connections made with the robotic cat. Similar observations of such connections are reflected by the following comment from a staff member.

> *I wasn't sure about the babies and the cat. I thought maybe that was a bit babyish...It's funny to watch how they [clients] perceive that and how quick they are to defend the cat, whereas sometimes you wouldn't even get a couple of words out of them, but they've got to make sure the cat is okay.* [Staff 7]

Generational factors, in that the client demographic needed assistance as they were not tech-savvy, or not familiar with, or not interested in technology, situated some clients in the "technology is not for everyone" group, in the eyes of the staff. These factors were often raised as a possible barrier to greater use of the technology in the Social Day Program. Future generations were viewed to be more likely to be receptive to technology, as evidenced by the following comments:

> *I think it's a double-edged sword and it has the potential to be brilliant, however unfortunately with that particular generation, because they were not particularly technologically minded or advanced, they are somewhat overwhelmed by it perhaps.* [Staff 6]

> *Look, we do have to help them most times just because they are technologies and it's not their era. Today's kids will be great in the future for these things, but these people were not brought up with these technologies.* [Staff 4]

While there was acknowledgment of the potential benefits of technology use, it was clear that staff were wary of the possible downfalls associated with it, demonstrating concern around physical and psychological safety for clients with frailty or cognition issues, or for whom memories of past traumas were triggered by using technology. Examples of this included clients who had worked in child abuse situations and had experienced miscarriages or the death of infants being exposed to the robotic dolls.

> *Some clients are really not suitable for the dolls and some aren't suitable for the animals. It does depend, because they've got either histories of – we've got some clients that have worked in child abuse situations. They're not good with the dolls, because it really upsets them. Knowing that background as well...we do the assessments and we ask the questions. But it's good knowing those things.* [Staff 4]

The juxtaposition between the potential benefit of technology use and its appropriateness was evident in the commentary around the use of the VR equipment. There was a common thread that the VR equipment could be the most beneficial for clients of all of the technologies introduced to the Social Day Program, allowing clients to reminisce about places they may have previously lived in or visited and activities previously performed, such as flying planes. For others, the VR was a source of concern; for example, clients with frailty issues may be scared to use it due to a fear of falling, and those with cognitive issues may be confused by what they are seeing.

XSL•FO
RenderX

> *I think that the VR would definitely be useful for some clients who may enjoy it, a lot of more cognitive ones I think, because people who have low cognition may get very frightened of some of the things they see.* [Staff 1]

Ensuring the safety of clients while they were using technology by having measures in place to assess physical capacity to use the VR, for example, or having background information that enabled identification of potential trauma triggers to avoid clients being exposed to them, was raised, with some staff indicating processes were in place within the organization to accommodate this.

> *We've got a safety test before we just chuck anybody on it [VR], because some people do have the same thing as me with the nausea and dizziness and stuff. There is a possibility of falling down.* [Staff 4]

Combined, not being tech-savvy, being averse to technology, or perceived to be at risk of physical or psychological injury due to technology use cemented the concept that technology is not for everyone involved in the Social Day Program.

Technology as a Placation Tool

Incorporated within this theme is the subtheme of technology as a placation tool. We initially wrestled with whether the use of technology in this way was situated as a potential benefit of technology use or a standalone theme; however, throughout the data, the utility of technology as a method of placation for clients, and therefore, it being perceived as a potential benefit, was evident. It seemed that while this was not an intended application of technology within the Social Day Program, placating agitated clients with items such as the robotic cat or time in the relaxation chair was rationalized by staff to lead to enriched experiences for those clients, as well as other clients in the program, whose experience was no longer being impinged upon by an agitated client. As a by-product, the benefit was extended to family members of the agitated clients who did not have to pick clients up from the Social Day Program earlier than intended, as described in the following comment:

> *It makes them extend their stay here in the day program. Rather than having to be picked up at lunchtime because their day is done and they've had enough stimulation for the day, they can be in there [the sensory room] for half an hour to an hour, come back out and enjoy the rest of the day 'til 3:30 [PM].* [Staff 4]

Knowledge and Confidence in Using Technology

This theme encapsulates concepts related to staff knowledge of the capabilities of the various pieces of technology and how to use them, as well as their confidence in assisting clients to use the technology, as a key driver in the use of technology in the Social Day Program. Further, these aspects are likely to influence whether technology continues to be used in the program. It was apparent across the data that training was important to staff; however, most reported that they had received little to no training in the use and maintenance of the technology. Further, there were perceptions that some of the technology required little training, such as the robotic doll and cat, whereas other types of technology, such as the VR or Obie, required more training of staff. This may have contributed to lower usage of these technologies as evidenced through observation, survey, and interview data, as may have confidence in using the technology. Perceptions of the relationship between training and confidence to use technology are contrasted. In some cases, staff reported that more training in the use of the available technology would increase their confidence to use the technology, for example:

> *I would like more training and capacity. I can do basic things, but I'm not confident using things...Yeah, I think I don't feel really confident with it. I'd like to be. I'd like to learn to do it.* [Staff 9]

Others felt confident to use the technology, irrespective of the level of training they had received. It was suggested that increasing confidence in the use of technology may increase the enthusiasm staff have for technology use, and subsequently, increase the regularity of technology use in the programs at the organization and assist in better tailoring technology use to individual clients.

Technology Functionality and Support

This theme reflects the importance of having available technology that is functional, as well as accessible support for staff in cases when the technology was not working. This would seem obvious in any environment where technology such as that implemented was being used; however, it was frequently raised by interviewees that technical and maintenance issues with the technology were common. These are clear barriers to the use of technology in the Social Day Program and are likely to influence the willingness of staff to use the technology or assist clients to use it. An example of this related to the VR technology, which staff reported had rarely been in working condition since its introduction. While some staff reported helping other staff to use the technology, obtaining higher-level support to repair the equipment had proved problematic, as illustrated by this comment from a staff member.

> *[If] I had a problem with something over the weekend let's say, I would have absolutely no idea who to call to get that support and I don't even think that the support necessarily is available. Therefore, from that perspective those clients would simply miss out until someone was able to be contacted to fix it. It would be a too bad, so sad kind of concept, which is not cool really.* [Staff 6]

Discussion

Principal Findings

This project aimed to undertake a multifaceted evaluation of the implementation of technologies at a Social Day Program to help inform decision-making regarding their implementation at the organization's other sites. The evaluation included (1) observation of use, (2) surveys of staff members' perceptions of enablers and barriers to use of the technologies, and (3) interviews with staff to further explore their experiences in use of technology and enablers and barriers to implementation and use of the technologies. The observation, staff survey, and

XSL•FO
RenderX

interviews provide a rich dataset to enable the evaluation. Observation over an extensive period identified limited use of technology, further verified by the survey data, with potential reasons for this, including barriers to and enablers of technology use identified in survey data and thematic analysis of interview data.

When technology was used it was predominantly used during the Social Day Program and by Social Day Program clients rather than after hours by the Care Hotel guests. One reason for this may be that guests in the Care Hotel might be recuperating from medical procedures, rather than solely attending for respite purposes. As such, these guests may not be able to engage, nor be interested in engaging, in technology use at this time. Alternatively, the technology located in the guests' Care Hotel rooms, such as an interactive television and tablet, which were not the focus of this evaluation, may have negated the need to access the technology located in the Social Day Program. For example, guests could relax and listen to music in their hotel rooms, and therefore would not need to access the sensory room and relaxation chair.

While items such as the robotic cat have been shown to be popular, a large portion of the use appears to be as a means of placating clients who were agitated or were disturbing other clients. This is not to say that the potential benefits of technology use at the organization were not recognized by some staff; however, encouragement of use may not be aligned with the original intention behind the implementation of the technology. Staff highlighted occasions when the use of the robotic cat by clients enhanced interaction between those clients, other clients, and staff. While mixed findings have been reported concerning the use of robotics in other aged care environments, some studies have shown that robotic animals such as cats and dogs can be conversation starters and contribute to enhanced engagement for older adults and people living with dementia [14,15].

In other cases, staff spoke of the opportunity that the sensory room provided for clients in being able to support an extended period of attendance at the Social Day Program, and in their eyes, held this as a potential benefit of technology use. Clients who may normally need to be picked up early, as they found the activities of the Social Day Program tiring or overwhelming, could instead remove themselves from the activities and take some time-out to relax in the sensory room. Being able to do this meant that these clients could spend longer days at the Social Day Program, rather than having to go home early. This has flow on benefits for their family carers, who, as a result, have a longer period on that day for respite. Research supports a reduction in care-related stressors for the carer during respite, and that the longer the respite duration, the greater the benefit to the carer [16]. Each of these factors can be highlighted to staff as potential benefits as a means of supporting implementation and greater use of the technology.

Harnessing the staff members' own perceptions of the emotional, social, cognitive, and physical benefits of technology use that were identified through this evaluation may be one approach to facilitating successful implementation. Observational, survey, and interview data indicated that technologies were perceived to positively impact the emotional state of clients. This was not only the calming effect reported and that is acknowledged in the literature [17], but also included observed happiness and enjoyment, as well as the social engagement that occurred among groups of Social Day Program clients as they used the Obie and robotic cats, for example.

Safety—both physical and psychological—was a prominent consideration for staff. Caution was raised about the risk of falls with some of the technology, in particular VR. The use of VR has been evaluated in other aged care environments and has been shown to have both positive and negative impacts. While it has the potential to engage and enrich the experience of the organization's clients by providing them access to activities, experiences, and environments that may no longer be physically accessible to them, VR may not be acceptable to, nor appropriate for, all people living with dementia [18]. Raised in the survey and interview data, these safety concerns extended to the possibility of psychological trauma for clients exposed to technology such as the robotic cats, categorizing them as part of the "technology is not for everyone" group. Concerns were raised about the stigma associated with people who use technologies such as the robotic cat or doll, because they are "not real." "Infantilization" of older adults living with dementia and their loss of dignity through the use of robotic toys, which they perceive to be real, has been explored in other research [17], and in some cases, it is the family members of the older adult holding this perception [19]. Strategies suggested to minimize this and maintain dignity when items such as robotic toys are used to engage people include creating an obvious environment of "play" [20].

Barriers to, and Enablers of, Technology Uptake and Use

As evidenced through this evaluation, several factors act as barriers to and enablers of technology use within the Social Day Program. The key factors, reflected by the developed themes, relate to staff training in, and knowledge of, technology operation and features; technology functionality and the availability of real-time technology support for staff; and perceptions of a generational influence on technology use. All of the barriers—and subsequently the enablers—to technology use identified in this evaluation were common to another recent Australian evaluation of technology use in aged care [9], suggesting they are not specific to this particular organization, but more reflective of the aged care environment generally.

Staff awareness and understanding of the underlying purpose of technology use for this population, beyond it being used as a behavior management tool, may also act as an enabler of greater use of technology within the organization. This includes opportunities for technology to promote "meaningful engagement" and social interaction [19], particularly, the technology should facilitate social interaction among those present and not be a replacement for it [9]. However, technologies that require staff to invest more of their already limited time to support clients in using them may also be less likely to be encouraged within the Social Day Program.

"Visibility" of the technology was considered by some staff to be an enabler of technology use; however, it was not raised sufficiently to warrant inclusion as a theme. Staff suggested

XSL•FO
RenderX

that technology items that are visible and easily accessible by staff and clients are more likely to be used regularly. Further, staff members who regularly use the technology and encourage clients to do so have the potential to encourage other staff members to do the same. As such, increased visibility of the technology has the potential to encourage more use within the organization.

Other potential barriers to the use of technology relate to perceptions that the current generation of clients at the Social Day Program are not "tech-savvy," do not like technology, or just have no interest in using it. Interviews identified that the clients were perceived not to have used much technology across their lifetime and that technology is likely to be more acceptable to future generations of clients who have grown up with and are familiar with technology. Staff reported observing positive impact related to social engagement, communication, and mood state, suggesting that some clients did enjoy using the technology. Enjoyment and acceptability of technology by older adults would seem to be obvious in whether technology is used and have been shown to be perceived by older adults as relevant in the successful implementation of technology to support ageing in place [20].

Sustainability

Targeting the enablers of technology use identified in this evaluation will contribute to ongoing technology use in the organization's Social Day Program. Taking a person-centered approach to technology use, by modifying organizational approaches to understanding the interests of clients, the physical and cognitive capabilities of clients, knowledge of potential triggers for negative or trauma responses associated with technology use; as well as technology that is maintained in good working order will also be a driver of sustainability.

Limitations

A limitation of this evaluation was having little data from clients and their family members, and therefore not being able to make their voices prominent in this evaluation. To an extent, this is reflective of other evaluations that have been undertaken in the aged care environment. In some cases, family members may not be involved in the lives of their guest or client family member. In others, the only respite family members may have is when the guest or client is at the Social Day Program. This may see family members not wanting to use that time for activities such as this evaluation.

In cases where guests or clients are living with cognitive decline or dementia, it was difficult to engage them in the interview process. However, it is important to ensure that people living with cognitive decline or dementia are included in projects about them, in particular, in the consultation and co-design process [21], which we were able to achieve through having clients and family members in the stakeholder group. While there were several instances in which staff provided negative feedback related to the technology not being operational, with limited support provided, it is also possible that other staff were reluctant to provide negative perspectives, which may introduce some bias to the findings.

Recommendations

While the use of technology may benefit this population, the potential benefits have not been fully realized at the organization due to several implementation barriers: there is a lack of staff knowledge, confidence in use, and training relevant to the operation of the different technologies, the technology has not always been operational and ongoing, accessible support is not available. There is a perception that technology has benefits, although the use of technology may not always be as intended by the organization, however, this is contrasted by the expressed view that technology is not for everyone and there are physical and psychological risks involved in use of technology in the Social Day Program population, many of whom live with cognitive decline. By addressing the barriers identified in this evaluation and strengthening the enablers per the following recommendations, the potential for clients to benefit across the domains could be enhanced.

Based on the results from the evaluation and considering the evidence from the literature, the following recommendations may enhance the implementation of technology in programs similar to the one described here:

- Increase staff knowledge of technology and its potential benefits for older adults living with cognitive decline, through the provision of more structured and regular training, a summary information sheet of project findings, quick refresher sessions, instruction sheets, and time to "play" and practice with the technology. This needs to include not only how to operate the technology and trouble shoot common problems or issues, but also provide an understanding of all the features and options (to enable tailoring to client needs or capabilities) and how to use the technology to best benefit the clients.
- Develop an information technology maintenance and support process that is resourced to ensure the equipment is functional at all times and enables staff to access real-time information technology support to address operational or technical issues. This should also include the development of comprehensive but user-friendly instructions to assist with understanding the features of the technology, operating the technology, and troubleshooting.
- Ensure that appropriate measures are in place to establish client physical and psychological safety, so that they are not exposed to technology that may be inappropriate for their circumstances.
- Increase the visibility and accessibility of the technology, so that it is at the forefront of staff's, clients' and their family members' or carers' thoughts as an activity. This may be through pictural posters placed around the room as reminders that the technology is there, with simple step by step instructions placed near technology or by leaving the technology in easy-to-access locations.
- Consider the placement of technology within the space.

Conclusions

This evaluation of technology use in a Social Day Program has identified that technology is not being used extensively, nor is it being used for the purposes of enrichment or experience enhancement. Multiple barriers to implementation and sustained

XSL•FO
RenderX

use of the technology items have been identified, spanning perceptions of clients' preference or low ability, to a lack of staff training and knowledge to support adequate use of the technology. Recommendations to improve implementation and promote sustained use of technology, based on the findings of this evaluation and evidence from the literature, may apply to other organizations seeking to implement these technologies in similar programs.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Semistructured interview guiding questions: family member or carer.
[DOCX File, 22 KB - aging_v8i1e60297_app1.docx]

References

1. About us. ECH. 2024. URL: https://www.ech.asn.au/about-us [accessed 2025-05-02]
2. Abdi S, de Witte L, Hawley M. Emerging technologies with potential care and support applications for older people: review of gray literature. JMIR Aging 2020 Aug 11;3(2):e17286. [doi: 10.2196/17286] [Medline: 32780020]
3. Khosravi P, Ghapanchi AH. Investigating the effectiveness of technologies applied to assist seniors: a systematic literature review. Int J Med Inform 2016 Jan;85(1):17-26. [doi: 10.1016/j.ijmedinf.2015.05.014] [Medline: 26216463]
4. Khosravi P, Rezvani A, Wiewiora A. The impact of technology on older adults' social isolation. Comput Human Behav 2016 Oct;63:594-603. [doi: 10.1016/j.chb.2016.05.092]
5. Pu L, Moyle W, Jones C, Todorovic M. The effectiveness of social robots for older adults: a systematic review and meta-analysis of randomized controlled studies. Gerontologist 2019 Jan 9;59(1):e37-e51. [doi: 10.1093/geront/gny046] [Medline: 29897445]
6. Bleakley CM, Charles D, Porter-Armstrong A, McNeill MDJ, McDonough SM, McCormack B. Gaming for health: a systematic review of the physical and cognitive effects of interactive computer games in older adults. J Appl Gerontol 2015 Apr;34(3):NP166-NP189. [doi: 10.1177/0733464812470747] [Medline: 24652863]
7. Dermody G, Whitehead L, Wilson G, Glass C. The role of virtual reality in improving health outcomes for community-dwelling older adults: systematic review. J Med Internet Res 2020 Jun 1;22(6):e17331. [doi: 10.2196/17331] [Medline: 32478662]
8. Yen HY, Chiu HL. Virtual reality exergames for improving older adults' cognition and depression: a systematic review and meta-analysis of randomized control trials. J Am Med Dir Assoc 2021 May;22(5):995-1002. [doi: 10.1016/j.jamda.2021.03.009] [Medline: 33812843]
9. Waycott J, Zhao W, Kelly RM, Robertson E. Technology-mediated enrichment in aged care: survey and interview study. JMIR Aging 2022 Apr 12;5(2):e31162. [doi: 10.2196/31162] [Medline: 34975014]
10. Kouijzer M, Kip H, Bouman YHA, Kelders SM. Implementation of virtual reality in healthcare: a scoping review on the implementation process of virtual reality in various healthcare settings. Implement Sci Commun 2023 Jun 16;4(1):67. [doi: 10.1186/s43058-023-00442-2] [Medline: 37328858]
11. Rigaud AS, Dacunha S, Harzo C, et al. Implementation of socially assistive robots in geriatric care institutions: healthcare professionals' perspectives and identification of facilitating factors and barriers. J Rehabil Assist Technol Eng 2024;11:20556683241284765. [doi: 10.1177/20556683241284765] [Medline: 39324049]
12. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
13. Braun V, Clarke V. Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. Couns Psychother Res 2021;21(1):37-47. [doi: 10.1002/capr.12360]
14. Campbell A. Dementia care: could animal robots benefit residents? Nurs Resid Care 2011 Dec;13(12):602-604. [doi: 10.12968/nrec.2011.13.12.602]
15. Naganuma M, Ohkubo E, Kato N. Use of robotic pets in providing stimulation for nursing home residents with dementia. Stud Health Technol Inform 2015;217(1009-12):1009-1012. [doi: 10.3233/978-1-61499-566-1-1009] [Medline: 26294602]
16. Zarit SH, Kim K, Femia EE, Almeida DM, Savla J, Molenaar PCM. Effects of adult day care on daily stress of caregivers: a within-person approach. J Gerontol B Psychol Sci Soc Sci 2011 Sep;66(5):538-546. [doi: 10.1093/geronb/gbr030] [Medline: 21642593]
17. Koh WQ, Ang FXH, Casey D. Impacts of low-cost robotic pets for older adults and people with dementia: scoping review. JMIR Rehabil Assist Technol 2021 Feb 12;8(1):e25340. [doi: 10.2196/25340] [Medline: 33497349]
18. Baker S, Waycott J, Robertson E, et al. Evaluating the use of interactive virtual reality technology with older adults living in residential aged care. Inf Process Manag 2020 May;57(3):102105. [doi: 10.1016/j.ipm.2019.102105]

XSL•FO
RenderX

19. Neal I, du Toit SHJ, Lovarini M. The use of technology to promote meaningful engagement for adults with dementia in residential aged care: a scoping review. Int Psychogeriatr 2020 Aug;32(8):913-935. [doi: 10.1017/S1041610219001388] [Medline: 31547900]
20. Peek STM, Wouters EJM, Luijkx KG, Vrijhoef HJM. What it takes to successfully implement technology for aging in place: focus groups with stakeholders. J Med Internet Res 2016 May 3;18(5):e98. [doi: 10.2196/jmir.5253] [Medline: 27143097]
21. Goeman DP, Corlis M, Swaffer K, et al. Partnering with people with dementia and their care partners, aged care service experts, policymakers and academics: a co-design process. Australas J Ageing 2019 Sep;38 Suppl 2(S2):53-58. [doi: 10.1111/ajag.12635] [Medline: 31496067]

Abbreviations

RA: research assistant
VR: virtual reality

Edited by R Yang; submitted 07.05.24; peer-reviewed by A Haynes, J Waycott, R Kelly; revised version received 25.03.25; accepted 29.03.25; published 12.05.25.

Please cite as:
Post D, Whitson K, Parfitt G
Implementation of New Technologies in an Aged Care Social Day Program: Mixed Methods Evaluation
JMIR Aging 2025;8:e60297
URL: https://aging.jmir.org/2025/1/e60297
doi:10.2196/60297

© Dannielle Post, Kathleen Whitson, Gaynor Parfitt. Originally published in JMIR Aging (https://aging.jmir.org), 12.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Unveiling the Frailty Spatial Patterns Among Chilean Older Persons by Exploring Sociodemographic and Urbanistic Influences Based on Geographic Information Systems: Cross-Sectional Study

Yony Ormazábal[1*], MSc; Diego Arauna[2*], PhD; Juan Carlos Cantillana[3], MSci; Iván Palomo[2], PhD; Eduardo Fuentes[2], PhD; Carlos Mena[1], PhD

[1]Longevity Center VITALIS, Faculty of Economics and Business, University of Talca, Universidad de Talca, Avenida Lircay S/N, Talca, Chile

[2]Thrombosis Research Center, Department of Clinical Biochemistry and Immunohematology, Faculty of Health Sciences, Interuniversity Center of Healthy Aging (CIES), Longevity Center VITALIS, University of Talca, Talca, Chile

[3]Facultad de Administración y Economía, Universidad Tecnológica Metropolitana, Santiago, Chile

[*]these authors contributed equally

Corresponding Author:
Carlos Mena, PhD
Longevity Center VITALIS, Faculty of Economics and Business, University of Talca, Universidad de Talca, Avenida Lircay S/N, Talca, Chile

Abstract

Background: Frailty syndrome increases the vulnerability of older adults. The growing proportion of older adults highlights the need to better understand the factors contributing to the prevalence of frailty. Current evidence suggests that geomatic tools integrating geolocation can provide valuable information for implementing preventive measures by enhancing the urban physical environment.

Objective: The aim of this study was to analyze the relationship between various elements of the urban physical environment and the level of frailty syndrome in older Chilean people.

Methods: A cohort of 251 adults aged 65 years or older from Talca City, Chile, underwent comprehensive medical assessments and were geographically mapped within a Geographic Information Systems database. Frailty was determined using the Fried frailty criteria. The spatial analysis of the frailty was conducted in conjunction with layers depicting urban physical facilities within the city, including vegetables and fruit shops, senior centers or communities, pharmacies, emergency health centers, main squares and parks, family or community health centers, and sports facilities such as stadiums.

Results: The studied cohort was composed of 187 women and 64 men, with no significant differences in age and BMI between genders. Frailty prevalence varied significantly across clusters, with Cluster 3 showing the highest prevalence (14/47, *P*=.01). Frail individuals resided significantly closer to emergency health centers (960 [SE 904] m vs 1352 [SE 936] m, *P*=.04), main squares/parks (1550 [SE 130] m vs. 2048 [SE 105] m, *P*=.03), and sports fields (3040 [SE 236] m vs 4457 [SE 322]m, *P*=.04) compared with nonfrail individuals. There were no significant differences in urban quality index across frailty groups, but frail individuals lived in areas with higher population density (0.013 [SE 0.001] vs 0.01 [SE 0.0007], *P*=.03).

Conclusions: Frail individuals exhibit geospatial patterns suggesting intentional proximity to health facilities, sports venues, and urban facilities, revealing associations with adaptive responses to frailty and socioeconomic factors. This highlights the crucial intersection of urban environments and frailty, which is important for geriatric medicine and public health initiatives.

(JMIR Aging 2025;8:e64254) doi:10.2196/64254

KEYWORDS

aging; frailty; geospatial clustering; urban factors; neighborhood conditions.

Introduction

Understanding the aging process and the sociodemographic determinants related to enhancing the quality of life has emerged as a very relevant research area in light of the rapid aging of the global population [1-3]. Currently, 12% of the world's population is aged ≥60 years, and projections suggest that this proportion may rise to 21.5% by the mid century [4]. Similarly, the ≥80 years age group is expected to increase from 1.7% to 4.5% [4]. In this context, Chile is experiencing a pronounced aging phenomenon [4,5]. Projections indicate that the Chilean population aged ≥60 years is set to surge from 15.7% to 32.9% by 2050, with the proportion of individuals aged ≥80 years potentially reaching 10.3% [4].

XSL•FO
RenderX

According to the World Health Organization, the frailty syndrome is a crucial determinant regarding the state of dependency, the presence of chronic diseases, polypharmacy, and the quality of life in older people [6,7]. The frailty syndrome is defined as a preventable and reversible clinical state in which the capacity of older people to cope with everyday stressors is compromised by an increase in vulnerability and the physiological deterioration associated with aging [8]. Recent results show a prevalence of frailty in Chile slightly higher than 20% [9,10]. Frail persons have higher risks of mortality, cognitive impairment, fractures, and hospitalization, among other adverse health events, which, considering the increase in the population of older people, represents a challenge for public health and social welfare systems [11,12].

The built environment refers to spaces altered or created by human activities, encompassing a spectrum from homes and schools to workplaces, highways, urban expanses, accessibility to amenities, recreational areas, and pollution [13]. This environment can be delineated into 2 primary components: the microenvironment, encapsulating neighborhood and street-level attributes, and the macroenvironment, which includes the degree of urbanization and patterns of land use [14]. Enhancing our understanding of how the urban physical environment impacts older adults can significantly aid in formulating effective plans and interventions to prevent the progression and onset of frailty while promoting the well-being of this population [15]. Previous research conducted by our group has demonstrated that leveraging geomatic tools, which integrate geolocation as an additional dimension of analysis, can provide valuable insights for studying frailty as a syndrome and supporting the implementation of preventive measures [16,17].

According to reports from the World Health Organization on aging and friendly cities, enhancing the environment through improvements in physical structures and community support is an effective approach to maintaining the health of older people [18]. Recent evidence underscores the impact of neighborhood characteristics on frailty among older people. Those residing in neighborhoods with abundant green spaces exhibit a lower incidence of frailty, whereas individuals perceiving precarious conditions in their surroundings, houses, and environment face a higher risk of frailty [19-21]. A comprehensive multilevel (individual and community) cross-sectional analysis highlighted that older adults living in aesthetically pleasing and walkable neighborhoods tend to exhibit lower levels of frailty. In contrast, areas with high-traffic roads, for example, were associated with a higher prevalence of frailty [22,23]. These findings emphasize the critical role of physical environmental factors in shaping the health and well-being of older populations, highlighting the importance of designing age-friendly communities that promote active and healthy aging.

In this context, this study aims to analyze the relationship between various elements of the urban physical environment and the level of frailty syndrome in older Chilean people.

Methods

Participants and Study Design

The research adopted a cross-sectional case-control design, with a representative sample of older persons (aged ≥65 years old, both men and women) randomly selected from various Family Health Centers and community groups of older people in Talca City, Chile (n=251). All medical centers that serve older adults in the city were considered, ensuring geographical representation. The inclusion criterion was adults aged 65 years and older. Participants with self-reported or medically documented cancer, Parkinson disease, or vascular events were excluded, as were older individuals unable to walk or speak [9]. The calculation of the sample size (aged ≥65 years old, both men and women) considered a prevalence of frailty in older adults of 24.6% [9], with a 95% CI, statistical power of 80%, and a loss percentage of 20%. The proportions of women and men in the sample were determined by the relative distribution of the adult population over 65 years using data from the National Socioeconomic Characterization Survey [24]. No additional stratification was applied.

Frailty Diagnosis

The Fried frailty phenotype criteria were used as the diagnostic tool for assessing frailty [9,25]. This method evaluates the presence or absence of the following components: slowness, weakness, weight loss, exhaustion, and low physical activity. These parameters were defined based on the criteria described previously by Palomo et al [9], which include: (1) slowness: walking velocity below a cut-off of 0.8 m/s average 3-meter walk at a usual pace, adjusted for sex and height according to the standards of the Short Physical Performance Battery, (2) weakness: handgrip strength measured using an Electronic Handgrip Dynamometer (Camry), with sex-specific cut-off (male <27 kg, female <15 kg), (3) weight loss: defined as loss of at least 5 kg in the previous 6 months, (4) exhaustion: a positive response to any of the following two questions from the Center for Epidemiological Studies Depression Scale: “I felt that anything I did was a big effort” and “I felt that I could not keep on doing things” at least 3 to 4 days a week,” (5) low physical activity: difficulty walking, assessed by the questions “Do you have difficulty walking a block?” or "Do you have difficulty climbing several flights of stairs without resting? Participants meeting 3 or more of these components were categorized as frail, those with no 1 or 2 components were considered prefrail, and individuals lacking all components were classified as non-frail or robust [9].

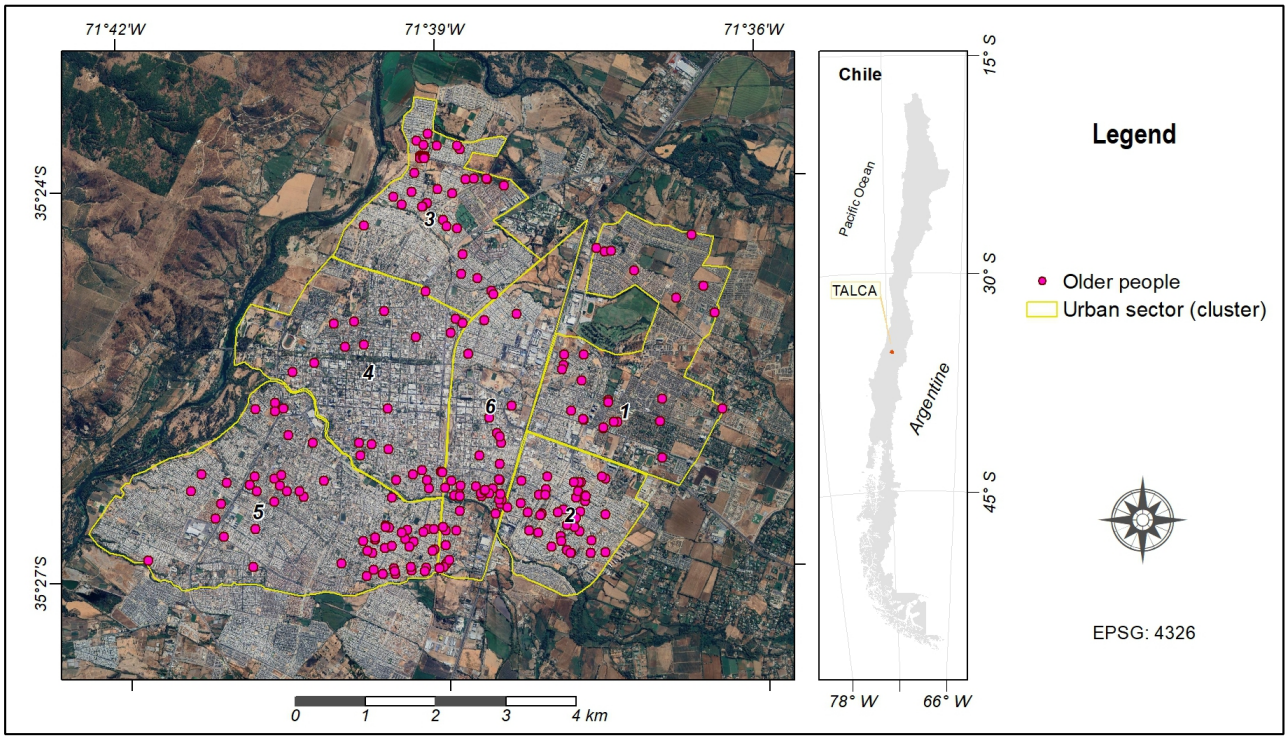
Geospatial Clusters

Each participant was geographically located according to the city address informed and represented as a point object based on the residence information provided during the medical evaluation. All data were organized into a point feature layer accompanied by its corresponding thematic table. This layer was integrated into a geodatabase for subsequent analysis within its geographical context, along with pertinent factors related to the urban physical environment. Figure 1 displays the individual residency positions of each older adult participating in the study within their respective geographical cluster. These geographical

XSL•FO
RenderX

areas were delimited in the city of Talca, Maule Region, Chile, based on geospatial location and sociodemographic characteristics. Cluster 1 corresponds to the northeastern sector characterized by high socioeconomic status. Cluster 2 encompasses the southeastern sector with low socioeconomic status. Cluster 3 covers the northern sector with a lower-middle socioeconomic class, and Cluster 4 represents the "historic center" area. The southern sector of medium-high socioeconomic level is covered by Cluster 5, and Cluster 6 corresponds to the "industrial center" area.

Figure 1. Location of older individuals within the 6 urban sectors.

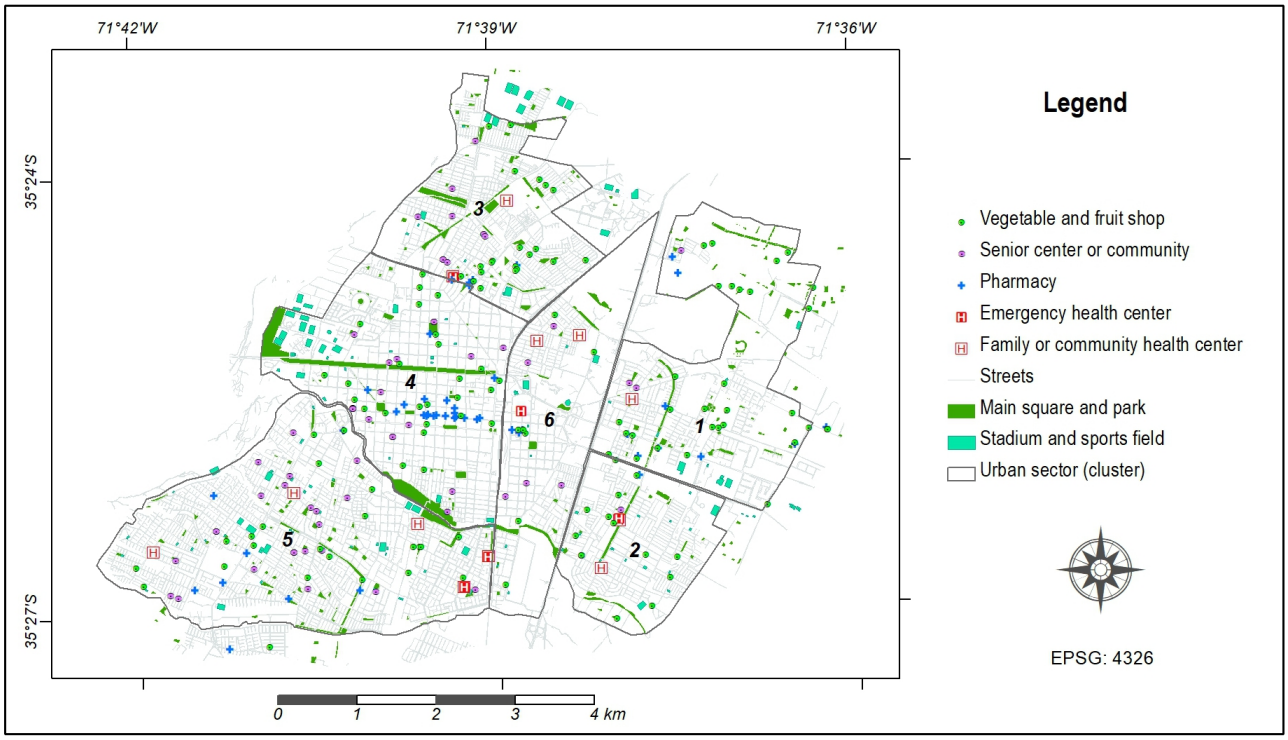


Urban Quality Level

A georeferenced database was constructed using Geographic Information Systems (GIS) technology to represent pertinent geographical information concerning urban physical facilities within the city, encompassing: (1) vegetables and fruits shops, (2) senior centers or communities, (3) pharmacies, (4) emergency health centers, (5) main squares and parks, (6) family or community health centers, and (7) stadiums and sports fields. Each component within the study area was depicted as a GIS layer, either in point or polygon form, at the neighborhood scale. This representation (Figure 2) was derived from data sourced from OpenStreetMap [26], Google Maps [27] and Infraestructura de Datos Geospaciales de Chile (IDE Chile) [28]. Subsequently, each GIS layer underwent analysis using the Euclidean distance method, providing insights into the proximity of every location within the city to the considered infrastructure. The resulting distance layers were subsequently classified to delineate 3 distinct zones encircling the urban facilities, categorizing their proximity as either close, medium, or distant. Next, each proximity class for every layer was assessed on a scale ranging from 1 to 3, wherein the closest proximity received a score of 3, and the more distant areas were assigned a score of 1. The distance ranges, and the corresponding values assigned to each urban facility were defined according to local context and are presented in Table 1. A general criterion for evaluation was that the closer the facility, the higher the value assigned. A raster calculator was used to aggregate all layers, yielding a summary index where a higher numerical value signifies enhanced urban quality in the depicted area. Afterward, the values derived from the distance analyses and the corresponding summary index for each participant were integrated into the point feature layer. Management, processing, and analyzing data were performed using ArcGIS software, version 10 (ESRI, Redlands, USA).

XSL•FO
RenderX

Figure 2. Geographical distribution of urban physical facilities within Talca City.



XSL•FO
RenderX

Table . Distance ranges and values for each urban physical facility.

Facility and distance ranges	Value
Vegetables and fruit shops (m)	
<300	3
300 - 600	2
>600	1
Senior centers or communities (m)	
<500	3
500 - 1000	2
>1000	1
Pharmacies (m)	
<500	3
500 - 1000	2
>1000	1
Emergency health centers (m)	
<1000	3
1000 - 2000	2
>2000	1
Main squares and parks (m)	
<200	3
200 - 400	2
>400	1
Family or community health centers (m)	
<700	3
700 - 1400	2
>1400	1
Stadiums and sports fields (m)	
<400	3
400 - 600	2
>600	1

Statistical Analysis

Statistical analyses were conducted using GraphPad Prism 9. Continuous variables were expressed as mean (SD) or median (95% CI). Categorical variables were expressed as percentages with a 95% CI. In the evaluation of differences between groups, the chi-square test with Yate correction was used to assess proportions, while ANOVA or the Kruskall-Wallis test, as appropriate, was applied to assess differences in means or medians. Statistical significance was considered at *P* values below .05.

Ethical Considerations

The institutional board review approval for this study was obtained from the Comité de Ética Científica (CEC) of Universidad de Talca (reference number 06 - 2021). All procedures followed adhered to the ethical standards of the CEC and the World Medical Association's Declaration of Helsinki. All study participants provided written informed consent.

Results

Sociodemographic Characteristics and Cluster Distribution

Table 2 presents the sociodemographic characteristics of the analyzed cohort of older people. The sample comprised 74.5% women and 25.2% men, with no significant differences observed in age and BMI between the 2 genders. In addition, Table 2 illustrates the distribution of the geospatial clusters established during the study. Analysis indicated no significant difference in the distribution of the 6 designated clusters between men and women.

XSL•FO
RenderX

Table . Sociodemographic description and geospatial distribution of the studied sample of older people.

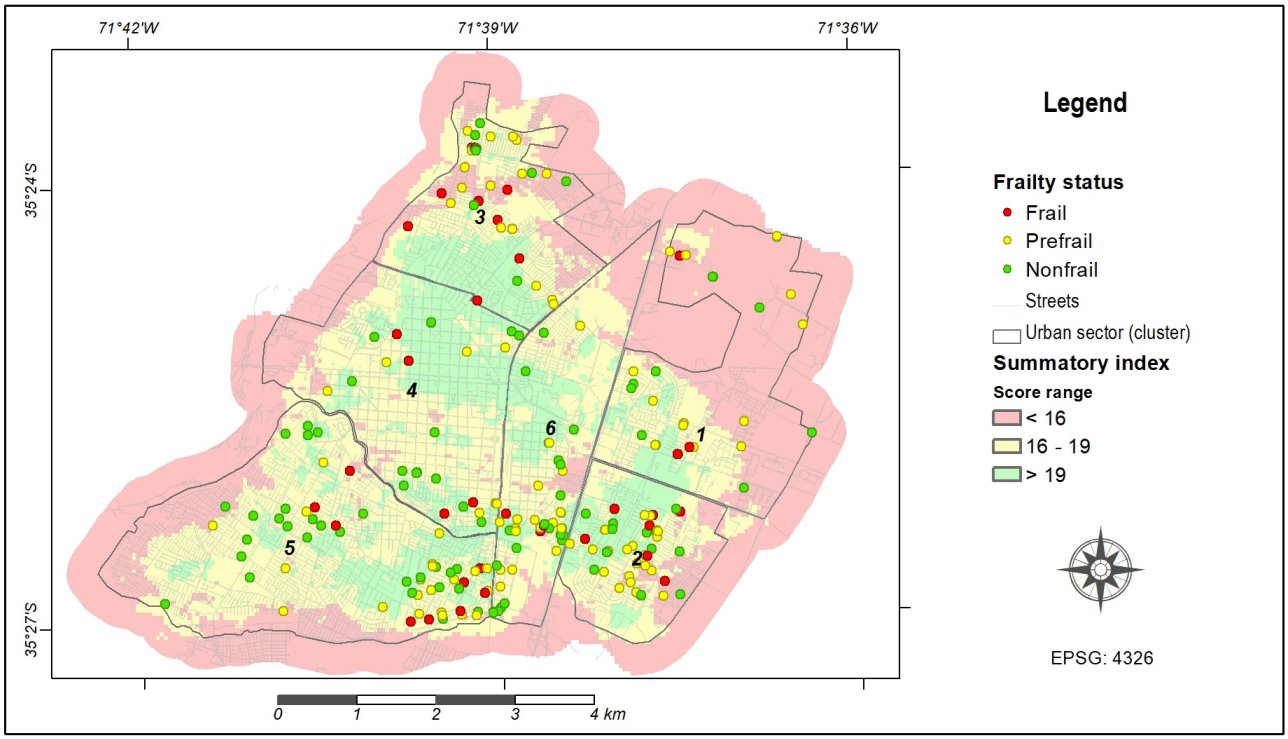
Variable	Women (n=187)	Men (n=64)	*P* value
Gender, % (95% CI)	74.5 (68.7-79.4)	25.5 (20.5-31.2)	—[a]
Age (years), mean (SD)	73.8 (5.9)	75 (5)	.152
BMI (kg/m^2), mean (SD)	33.6 (33.5)	28.4 (6.5)	.243
Spatial cluster showing percentage of prevalence, % (95% CI)			
Cluster 1	10.2 (6.6-15.3)	15.6 (8.7-26.4)	.260
Cluster 2	18.7 (13.7-24.9)	14.1 (7.6-24.6)	.452
Cluster 3	16.6 (11.9-22.5)	25 (16-36.8)	.142
Cluster 4	11.2 (7.5-16.6)	9.4 (4.4-18.9)	.817
Cluster 5	26.2 (20.4-32.9)	17.2 (9.8-28.2)	.176
Cluster 6	17.1 (12.4-23.2)	18.8 (11.1-29.9)	.845

[a] —: not applicable.

Analysis of Urban Quality

Figure 3 shows the different levels of urban quality [3] obtained from the cumulative assessment of physical environmental elements considered in this study. These elements mainly encompass basic urban services and infrastructures essential for the local population, with particular importance for the cohort of older people under study. The quality level is closely related to the accessibility of the various facilities from each location within the city. In addition, Figure 3 also shows the individual distribution of older adults, each denoted by their frailty status, which is subsequently analyzed in Table 3. This table displays the distribution of the frailty status through the different geospatial clusters. The prevalence of frailty varies between 7.3% and 34.1% among these clusters. Notably, cluster 3 exhibits a significantly high prevalence of frail people (34.1%, *P*=.006), followed by cluster 5 (21.9%, *P*=.417); however, this last prevalence is not significantly high.

Figure 3. Frailty status of older individuals over summary index within the geospatial clusters defined for Talca city.



XSL•FO
RenderX

Table . Distribution of older people by frailty status and geospatial clustering.

Spatial cluster	Frailty status considering the percentage of prevalence			*P* value
	Nonfrail (%, 95% CI)	Prefrail (%, 95% CI)	Frail (%, 95% CI)	
Cluster 1	14.3 (8.5 - 22.9)	11.7 (6.9 - 19)	7.3 (2.5 - 19.4)	.620
Cluster 2	14.3 (8.5 - 22.9)	21.6 (14.9 - 30.2)	17.7 (8.5 - 31.3)	.270
Cluster 3	12.1 (6.8 - 20.4)	19.8 (13.5 - 28.2)	34.1 (21.5 - 49.5)	.006
Cluster 4	14.3 (8.5 - 22.9)	8.1 (4.3 - 14.6)	12.2 (5.3 - 25.5)	.477
Cluster 5	30.7 (22.2 - 40.9)	20.7 (14.2 - 29.2)	21.9 (12 - 36.7)	.417
Cluster 6	14.3 (8.5 - 22.9)	18 (11.9 - 26.2)	7.3 (2.5 - 19.4)	.225
Total	100	100	100	—[a]

[a] —: not applicable.

Analysis of Distances to Urban Facilities, Summary Index, and Population Density

Figure 4 illustrates the variation in average distance between old persons (categorized by their frailty status) and relevant urban facilities. For the facilities of vegetable and fruit shops (A), senior centers or communities (B), pharmacies (C), and family and community health centers (F), there were no significant differences in the average distance across different frailty status groups. However, a clear linear trend is observed between the groups, where frail people tend to reside further from vegetable and fruit shops (frail: 335.6 [SE 31.2] vs nonfrail: 275.9 [SE 16.5]) and closer to the senior centers or communities (frail: 368.2 [SE 38.6] vs nonfrail: 435.9 [SE 35.9]) than the nonfrail people. On the other hand, the facilities of emergency health centers (D), main squares and parks (E), and stadiums and sports fields (G) present significant differences in the average distance across different frailty status groups, where frail people resided significantly closer to emergency health centers (frail: 960.4 [SE 90.4] vs nonfrail: 1352 [SE 93.6], *P*=.04), main squares and parks (frail: 155 [SE 13] vs prefrail: 204.8 [SE 10.5], *P*=.03), and stadiums and sports fields (frail: 304 [SE 23.6] vs prefrail: 445.7 [SE 32.2], *P*=.04), than both nonfrail and prefrail people, respectively. Finally, Figure 5 shows the variations in the summary index and population density between the frailty status groups. While the summary index (A) shows no significant differences between groups, there is a significant difference in population density (B) between the frail and nonfrail status groups, being higher for frail people (frail: 0.013 [SE 0.001] vs nonfrail: 0.01 [SE 0.0007], *P*=.03).

XSL•FO
RenderX

Figure 4. Comparison of the mean distance for different frailty status diagnosed using the Fried phenotype criteria with respect to relevant urban facilities of (A) vegetables and fruits shops; (B) senior centers or communities; (C) pharmacies; (D) emergency health centers; (E) main squares and parks; (F) family or community health centers; (G) stadiums and sports fields. The data presented are the mean (SE). Statistical analysis was performed using the ANOVA test with the Tukey pos-hoc test. *P<.05.

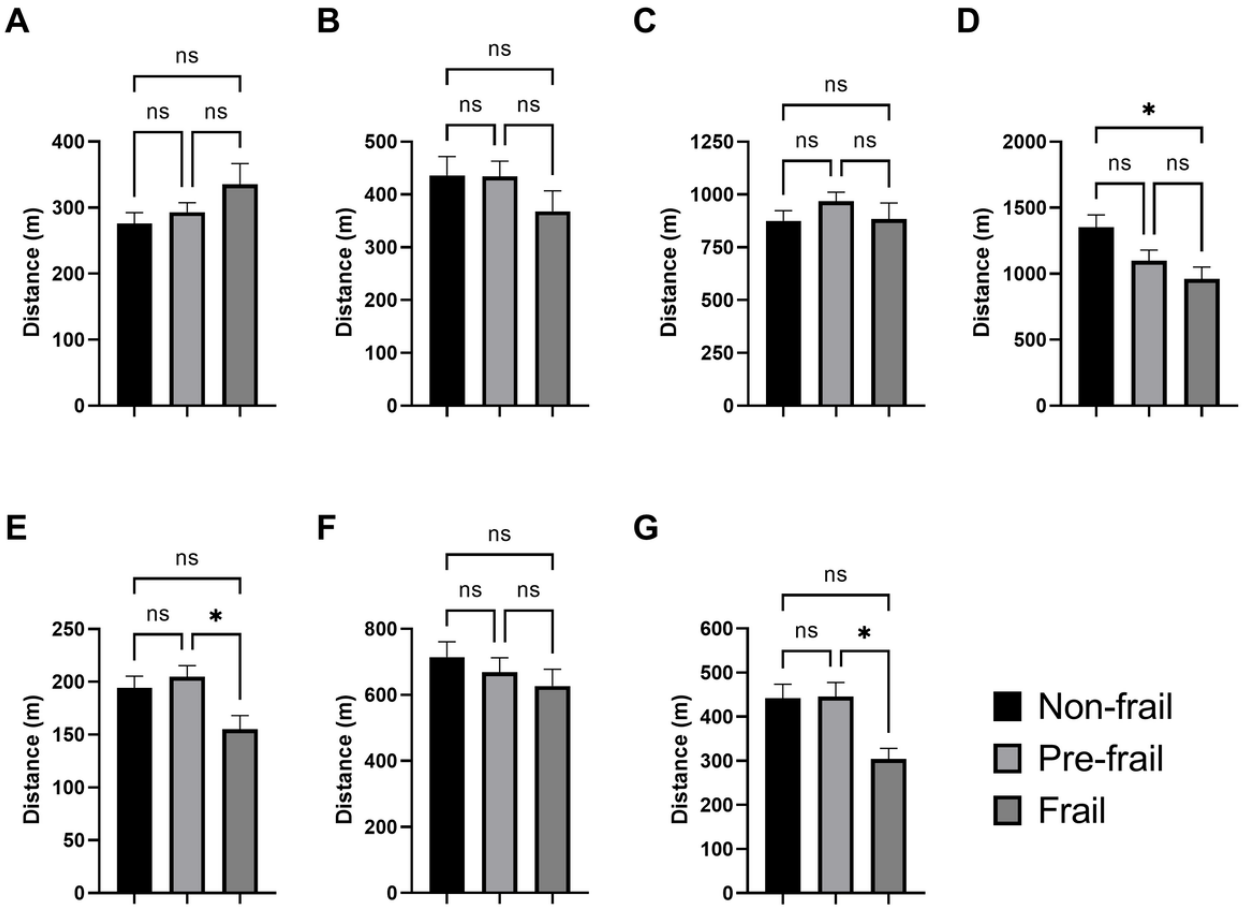
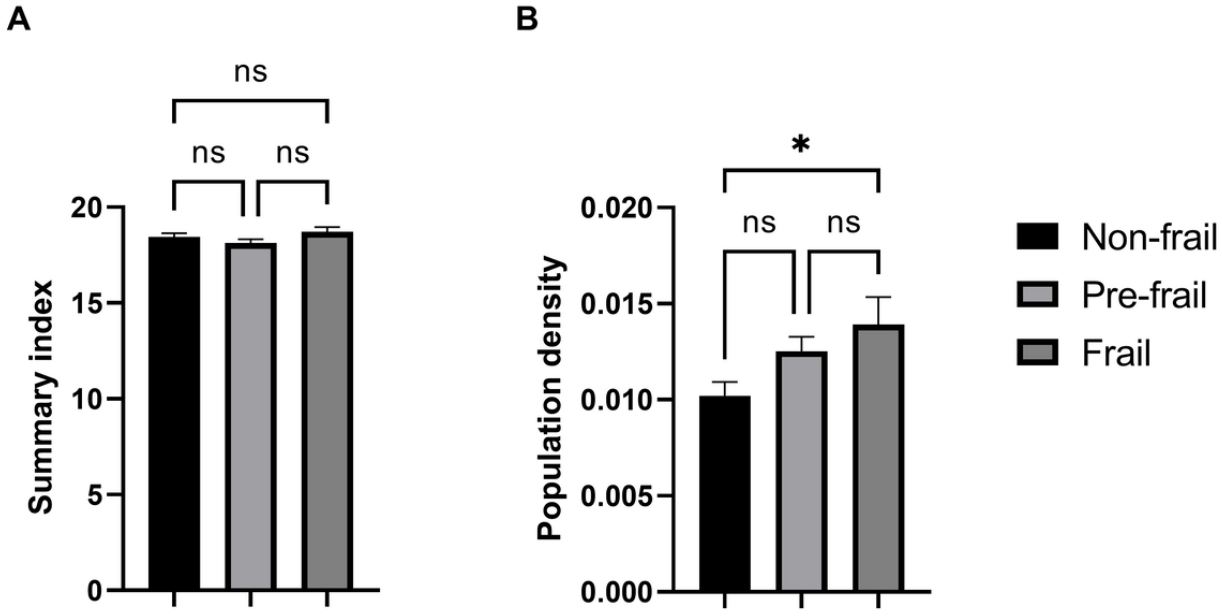


Figure 5. Comparison of mean values for different frailty status diagnosed according to the Fried phenotype criteria for the (A) summary index, (B) population density. The data presented are mean (SE). Statistical analysis was performed using the ANOVA test with the Tukey post-hoc test. *P<.05.



XSL•FO
RenderX

Discussion

Principal Findings

This study explores the relationship between urban physical environment factors and frailty syndrome, with a particular focus on its diagnostic criteria. The results mentioned above indicate that frail individuals tend to live closer to emergency health centers, stadiums, sports fields, and senior centers, suggesting efforts to improve their health and social engagement. In addition, frail individuals primarily reside in densely populated areas, which are associated with limited physical activity and higher mortality rates. Given the relevance of the Fried frailty phenotype as a standard frailty assessment tool in various studies involving older adults, it was selected for frailty characterization [11,29,30]. Investigating the link between frailty and urban environment is innovative and can provide invaluable insights for government programs aimed at enhancing the well-being of older people [31,32]. However, this emerging topic remains underdeveloped, limiting opportunities for comparisons specific to frailty. The incorporation of geomatics tools enables the use of geolocation as an additional analytical dimension, offering valuable insights into the study of frailty as a syndrome. This is especially important when the primary focus is on understanding the impact of the urban physical environment on the frail condition of older individuals [16,17]. Currently, the use of information and communication technologies among older persons is increasing, and different benefits have been developed about frailty with respect to predicting the risk, assisting in identifying changes in frailty parameters, enhancing adherence to a healthy diet, and distinguishing older rehabilitation patients who need to be readmitted to the hospital from those who can remain in the community which could facilitate this key process [33-37].

The data from Table 2 indicate homogeneity within the cohort, with no significant influences in the study parameters when stratified by gender. Participants had an average age range of 73 - 75 years old and exhibited a high BMI. This trend of elevated BMI is consistent with findings from other studies focused on older Chilean people, wherein a significant prevalence of obesity (BMI >25 kg/m^2) has been documented [8,9,38]. Gender did not appear to influence the distribution prevalence within the identified clusters. However, it is necessary to account for gender in geospatial clustering analyses involving older people, given its potential impact on various health outcomes, as recently evidenced in COVID-19 studies [39].

The findings presented in Table 3 suggest that geospatial distribution significantly impacts the prevalence of frailty syndrome. Notably, clusters 3 and 5, which are associated with middle to low socioeconomic classes, concentrate more than half of the frail people. This evidence is consistent with studies where low socioeconomic groups have been associated with a higher risk of developing frailty [31]. Furthermore, previous research indicates that frailty condition among older people tends to exhibit spatial clustering, wherein certain areas within and outside the city display localized concentrations of both high and low prevalence of frail people [16,40]. These clustering patterns can be attributed to the diverse urban infrastructures and socioeconomic disparities observed across different sectors within the city [16].

In this study, the outcomes depicted in Figure 3 reveal a geospatial clustering concentrated in the southern part of the city, primarily associated with clusters 3 and 5. However, when considering the general condition of the urban physical environment, characterized by the summary index, the group of frail people does not show significant differences compared with the other groups (Figure 5A). This fact suggests that, despite various groups of adults residing in environments characterized by comparable habitability conditions, distinctions in certain factors, as evidenced in Figure 4, can exert relevant positive or negative influences. Thus, the urban physical environment is likely connected with concurrent factors associated with the frailty syndrome, serving either as causal factors or as responses to such influences.

Frailty is acknowledged as a critical factor influencing the health and well-being of older people [11-18]. It stands out as the primary risk factor and indicator for the initiation of dependency, as well as the occurrence of chronic diseases, hospitalizations, falls, fractures, and mortality. The extensive array of health-related challenges observed in frail people may be associated with their proximity to emergency health centers, as well as family or community health centers [15-20]. Likewise, frailty has exhibited strong associations with obesity, sarcopenia, and low physical activity. Consequently, the proximity identified between the frail group and stadiums and sports fields may suggest a purposeful endeavor to improve body composition and address underlying clinical conditions [41,42].

Simultaneously, the frail group exhibited greater proximity to main squares and parks and to senior centers or communities renowned for fostering social activities among older people [43]. Given the well-established associations between frailty and social isolation, depression, and loneliness [44], the observed significant trend might signify a proactive response to enhance community engagement. This trend could also be indicative of the city government's concerted efforts to provide these facilities to a population in need. By strategically situating frail individuals near these spaces, urban planners and policymakers may be aiming to reduce the adverse effects of isolation and encourage greater social participation. Engaging in recreational activities and interacting with peers in accessible public spaces can help mitigate the mental health challenges often faced by frail older adults. Furthermore, these spaces serve as venues for both physical and mental stimulation, which are critical to maintaining functional independence and improving overall well-being.

Conversely, the robust association between frailty and obesity might elucidate the trend observed trend of increased distance from the frail group to fruit and vegetable shops. Nevertheless, validation of this hypothesis necessitates examination within a more extensive cohort. While elevated intake of fruits and vegetables has been linked to a diminished frailty risk, the impact of the proximity of these food supply points on frailty remains ambiguous [45]. Furthermore, greater distances to pharmacies may indicate challenges in accessing essential drugs

XSL•FO
RenderX

and medications, a factor that should be taken into account in health programs tailored to support the frailty group.

The findings depicted in Figure 5B indicate that the frail group tends to inhabit regions characterized by elevated population density. Older adults tend to live in densely populated marginalized areas due to economic, social, and structural factors. Financial limitations, the lack of adequate housing options, access to family support networks, and mobility barriers are some of the key factors influencing their stay in these areas. This circumstance poses an increased risk to this group, as current scientific understanding suggests a positive correlation between population density in a given neighborhood and increased susceptibility of middle-aged and older adults to overweight conditions [46]. This association can be ascribed to the proclivity of individuals residing in densely populated areas to adopt a sedentary lifestyle, marked by limited physical activity and diminished energy expenditure [20,46]. Notably, low physical activity and sedentarism are integral components of the frailty phenotype observed across diverse cohorts of older adults [11,12,26]. Furthermore, elevated population density has been lined to increased mortality rates across all causes in older people [11,12,47,48].

Initiatives aimed at preventing frailty underscore the imperative to advocate for physical activity, nutrition, and social engagement as primary and efficacious interventions. These interventions can be effectively implemented through a health education program tailored to inspire and engage older people [49,50]. Our findings underscore the importance of taking urban factors into account when examining the frailty condition of older people. This evidence aligns with previous research emphasizing the importance of these factors in the well-being and social engagement of older adults, as well as their association with frailty [16,51]. These results emphasize the imperative to investigate further and enhance our comprehension of the role played by urban factors in shaping frailty among older people. Regardless, it is imperative to consider the complex interrelationship between the urban physical environment and frailty condition when devising structural preventive measures aimed at improving the well-being of older people [16,47].

Conclusions

Contemporary evidence underscores the relevance of urban factors in influencing the onset of frailty and the diverse health factors linked to this syndrome. Frailty stands as a highly prevalent geriatric syndrome in the elderly, elevating the susceptibility to a range of adverse health and social outcomes. Addressing this challenge necessitates the development of age-friendly cities tailored to the needs of older populations. Our findings suggest that individuals classified as frail tend to reside in closer proximity to emergency health centers, as well as family or community health centers, which may be indicative of adaptive responses to the features associated with frailty, such as elevated mortality risk and diminished levels of physical activity. Likewise, frail people tend to reside closer to stadiums and sports fields, which may imply a deliberate endeavor to enhance body composition and address underlying clinical conditions. On its part, the closer proximity of frail people to urban infrastructures such as main squares and parks, and senior centers or communities may be indicative of the concerted effort by the municipal government to provide these facilities to a population in need. However, it is important to highlight that the frail group predominantly inhabits areas characterized by elevated population density, aligning with earlier research associating higher mortality rates with increased population density among older adults. Likewise, the geospatial clustering highlights the relevance of both socioeconomic status and geographic location in relation to frailty prevalence, unveiling an elevated occurrence of this syndrome in sectors characterized by a lower-middle socioeconomic class. This finding aligns with and reinforces previously established evidence. Ultimately, the existing body of evidence, coupled with our study findings, underscores the significance of investigating frailty and its associations with the urban environment and related factors. This emerging field of research holds groundbreaking potential, offering substantial implications for geriatric medicine. Furthermore, it provides invaluable insights that can inform the development of governmental initiatives aimed at promoting healthy aging and proactively preventing frailty.

Acknowledgments

This work has been funded by Interuniversity Center for Healthy Aging, (Code RED211993). Also, the authors would like to thank ANID-FONDECYT 1211136 and ANID-FONDECYT 3230783.

Authors' Contributions

DA, YO, and JCC contributed to conceptualization. DA and YO handled data curation and formal analysis. CM, IP, and EF managed funding acquisition. DA, YO, and JCC conducted investigation. DA, YO, and JCC managed methodology. CM, IP, and EF handled project administration. CM and IP managed resources. YO handled software. CM and IP supervised. CM and IP managed validation. DA and YO conducted writing-original draft and writing-review & editing.

Conflicts of Interest

None declared.

References

1. He W, Goodkind D, Kowal PR. An aging world: 2015. : United States Census Bureau Washington, DC; 2016.

XSL•FO
RenderX

2. Nguyen TN, Cumming RG, Hilmer SN. A review of frailty in developing countries. J Nutr Health Aging 2015 Nov;19(9):941-946. [doi: 10.1007/s12603-015-0503-2] [Medline: 26482697]
3. Bloom DE, Boersch-Supan A, McGee P, Seike A. Population aging: facts, challenges, and responses. Benefits and compensation international 2011;41(1):22.
4. Thumala D, Kennedy BK, Calvo E, et al. Aging and health policies in Chile: new agendas for research. Health Syst Reform 2017 Oct 2;3(4):253-260. [doi: 10.1080/23288604.2017.1353844] [Medline: 30359182]
5. Rudnicka E, Napierała P, Podfigurna A, Męczekalski B, Smolarczyk R, Grymowicz M. The World Health Organization (WHO) approach to healthy ageing. Maturitas 2020 Sep;139(6-11):6-11. [doi: 10.1016/j.maturitas.2020.05.018] [Medline: 32747042]
6. WHO guidelines on integrated care for older people (ICOPE). World Health Organization. 2018. URL: https://www.who.int/ageing/publications/guidelines-icope/en [accessed 2025-04-14]
7. Arauna D, Cerda A, García-García JF, et al. Polypharmacy Is Associated with Frailty, Nutritional Risk and Chronic Disease in Chilean Older Adults: Remarks from PIEI-ES Study. Clin Interv Aging 2020;15:1013-1022. [doi: 10.2147/CIA.S247444] [Medline: 32636616]
8. WHO. WHO clinical consortium on healthy ageing: topic focus: frailty and intrinsic capacity. 2017 Presented at: Report of Consortium Meeting; Dec 1-2, 2016; Geneva, Switzerland URL: https://www.who.int/publications/i/item/WHO-FWC-ALC-17.2 [accessed 2024-12-10]
9. Palomo I, Giacaman RA, León S, et al. Analysis of the characteristics and components for the frailty syndrome in older adults from central Chile. The PIEI-ES study. Arch Gerontol Geriatr 2019;80(70-5):70-75. [doi: 10.1016/j.archger.2018.10.004] [Medline: 30390427]
10. Palomo I, García F, Albala C, et al. Characterization by gender of frailty syndrome in elderly people according to FrailTrait Scale and Fried Frailty Phenotype. J Pers Med 2022 Apr 29;12(5):712. [doi: 10.3390/jpm12050712] [Medline: 35629135]
11. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
12. García-García FJ, Carcaillon L, Fernandez-Tresguerres J, et al. A new operational definition of frailty: the Frailty Trait Scale. J Am Med Dir Assoc 2014 May;15(5):371. [doi: 10.1016/j.jamda.2014.01.004] [Medline: 24598478]
13. Pasala SK, Rao AA, Sridhar GR. Built environment and diabetes. Int J Diabetes Dev Ctries 2010 Apr;30(2):63-68. [doi: 10.4103/0973-3930.62594] [Medline: 20535308]
14. Swinburn B, Egger G, Raza F. Dissecting obesogenic environments: the development and application of a framework for identifying and prioritizing environmental interventions for obesity. Prev Med 1999 Dec;29(6 Pt 1):563-570. [doi: 10.1006/pmed.1999.0585] [Medline: 10600438]
15. Di Ciaula A, Portincasa P. The environment as a determinant of successful aging or frailty. Mech Ageing Dev 2020 Jun;188(111244):111244. [doi: 10.1016/j.mad.2020.111244] [Medline: 32335099]
16. Mena C, Ormazabal Y, Fuentes E, Palomo I. Impacts of physical environment perception on the frailty condition in older people. Geospat Health 2020 Jun 18;15(1):32575969. [doi: 10.4081/gh.2020.888] [Medline: 32575969]
17. Mena C, Fuentes E, Ormazábal Y, Palomo-Vélez G, Palomo I. Effect of straight-line and road network distances to parks and markets on anthropometric measurements, biochemical markers, and a healthy lifestyle in adult people. Sport Sci Health 2016 Apr;12(1):55-61. [doi: 10.1007/s11332-015-0253-x]
18. World Health Organization. Decade of healthy ageing: baseline report. 2020 URL: https://www.who.int/publications/i/item/9789240017900 [accessed 2024-12-10]
19. Gobbens RJJ. Cross-sectional and longitudinal associations of environmental factors with frailty and disability in older people. Arch Gerontol Geriatr 2019;85:103901. [doi: 10.1016/j.archger.2019.103901] [Medline: 31352186]
20. Yu R, Wang D, Leung J, Lau K, Kwok T, Woo J. Is neighborhood green space associated with less frailty? Evidence from the Mr. and Ms. Os (Hong Kong) study. J Am Med Dir Assoc 2018 Jun;19(6):528-534. [doi: 10.1016/j.jamda.2017.12.015] [Medline: 29402649]
21. Duppen D, Van der Elst MCJ, Dury S, Lambotte D, De Donder L, D-SCOPE. The social environment's relationship with frailty: evidence from existing studies. J Appl Gerontol 2019 Jan;38(1):3-26. [doi: 10.1177/0733464816688310] [Medline: 28380715]
22. Abe T, Carver A, Sugiyama T. Associations of neighborhood built and social environments with frailty among mid-to-older aged Australian adults. Geriatr Gerontol Int 2021 Oct;21(10):893-899. [doi: 10.1111/ggi.14253] [Medline: 34355479]
23. Ye B, Gao J, Fu H. Associations between lifestyle, physical and social environments and frailty among Chinese older people: a multilevel analysis. BMC Geriatr 2018 Dec 14;18(1):314. [doi: 10.1186/s12877-018-0982-1] [Medline: 30547760]
24. CASEN 2017. Base de datos Casen 2017 SPSS (Base de datos) [Website in Spanish]. URL: https://observatorio.ministeriodesarrollosocial.gob.cl/encuesta-casen-2017 [accessed 2025-04-14]
25. Albala C, Lera L, Sanchez H, et al. Frequency of frailty and its association with cognitive status and survival in older Chileans. Clin Interv Aging 2017;12(995-1001):995-1001. [doi: 10.2147/CIA.S136906] [Medline: 28721027]
26. Open street map. OSM Foundation. URL: https://www.openstreetmap.org [accessed 2024-11-01]
27. Google Maps Google LLC. URL: https://www.google.com/maps [accessed 2024-11-01]
28. Geospatial infrastructure chile. IDE Chile. URL: https://www.ide.cl [accessed 2024-11-01]

XSL•FO
RenderX

29. Martin FC, O'Halloran AM. Tools for assessing frailty in older people: general concepts. Adv Exp Med Biol 2020;1216(9-19):9-19. [doi: 10.1007/978-3-030-33330-0_2] [Medline: 31894542]
30. Morley JE, Vellas B, van Kan GA, et al. Frailty consensus: a call to action. J Am Med Dir Assoc 2013 Jun;14(6):392-397. [doi: 10.1016/j.jamda.2013.03.022] [Medline: 23764209]
31. Hoogendijk EO, Afilalo J, Ensrud KE, Kowal P, Onder G, Fried LP. Frailty: implications for clinical practice and public health. Lancet 2019 Oct 12;394(10206):1365-1375. [doi: 10.1016/S0140-6736(19)31786-6] [Medline: 31609228]
32. Dent E, Martin FC, Bergman H, Woo J, Romero-Ortuno R, Walston JD. Management of frailty: opportunities, challenges, and future directions. Lancet 2019 Oct 12;394(10206):1376-1386. [doi: 10.1016/S0140-6736(19)31785-4] [Medline: 31609229]
33. Kim B, McKay SM, Lee J. Consumer-grade wearable device for predicting frailty in Canadian home care service clients: prospective observational proof-of-concept study. J Med Internet Res 2020 Sep 3;22(9):e19732. [doi: 10.2196/19732] [Medline: 32880582]
34. Ramezani R, Zhang W, Xie Z, et al. A combination of indoor localization and wearable sensor-based physical activity recognition to assess older patients undergoing subacute rehabilitation: baseline study results. JMIR Mhealth Uhealth 2019 Jul 10;7(7):e14090. [doi: 10.2196/14090] [Medline: 31293244]
35. Valdés-Aragonés M, Pérez-Rodríguez R, Carnicero JA, et al. Effects of monitoring frailty through a mobile/web-based application and a sensor kit to prevent functional decline in frail and prefrail older adults: FACET (frailty care and well function) pilot randomized controlled trial. J Med Internet Res 2024 Oct 22;26:e58312. [doi: 10.2196/58312] [Medline: 39436684]
36. Su Y, Wu KC, Chien SY, Naik A, Zaslavsky O. A mobile intervention designed specifically for older adults with frailty to support healthy eating: pilot randomized controlled trial. JMIR Form Res 2023 Nov 15;7:e50870. [doi: 10.2196/50870] [Medline: 37966877]
37. Keränen NS, Kangas M, Immonen M, et al. Use of information and communication technologies among older people with and without frailty: a population-based survey. J Med Internet Res 2017 Feb 14;19(2):e29. [doi: 10.2196/jmir.5507] [Medline: 28196791]
38. Bravo C, Godoy J, Sánchez Y, Riveros MJ. Asociación entre el nivel de autovalencia y el estado nutricional en adultos mayores chilenos [Article in Spanish]. Rev chil nutr 2021;48(5):741-747. [doi: 10.4067/S0717-75182021000500741]
39. Ramírez-Aldana R, Gomez-Verjan JC, Bello-Chavolla OY, García-Peña C. Spatial epidemiological study of the distribution, clustering, and risk factors associated with early COVID-19 mortality in Mexico. PLoS One 2021;16(7):e0254884. [doi: 10.1371/journal.pone.0254884] [Medline: 34288952]
40. Taylor D, Barrie H, Lange J, Thompson MQ, Theou O, Visvanathan R. Geospatial modelling of the prevalence and changing distribution of frailty in Australia - 2011 to 2027. Exp Gerontol 2019 Aug;123:57-65. [doi: 10.1016/j.exger.2019.05.010] [Medline: 31129145]
41. Crow RS, Lohman MC, Titus AJ, et al. Association of obesity and frailty in older adults: NHANES 1999-2004. J Nutr Health Aging 2019;23(2):138-144. [doi: 10.1007/s12603-018-1138-x] [Medline: 30697622]
42. Yuan L, Chang M, Wang J. Abdominal obesity, body mass index and the risk of frailty in community-dwelling older adults: A systematic review and meta-analysis. Age Ageing 2021 Jun 28;50(4):1118-1128. [doi: 10.1093/ageing/afab039] [Medline: 33693472]
43. Noon RB, Ayalon L. Older adults in public open spaces: age and gender segregation. Gerontologist 2018 Jan 18;58(1):149-158. [doi: 10.1093/geront/gnx047] [Medline: 28535273]
44. Davies K, Maharani A, Chandola T, Todd C, Pendleton N. The longitudinal relationship between loneliness, social isolation, and frailty in older adults in England: a prospective analysis. Lancet Healthy Longev 2021 Feb;2(2):e70-e77. [doi: 10.1016/S2666-7568(20)30038-6] [Medline: 36098160]
45. Kojima G, Avgerinou C, Iliffe S, Jivraj S, Sekiguchi K, Walters K. Fruit and vegetable consumption and frailty: a systematic review. J Nutr Health Aging 2018;22(8):1010-1017. [doi: 10.1007/s12603-018-1069-6] [Medline: 30272107]
46. Wang R, Feng Z, Xue D, Liu Y, Wu R. Exploring the links between population density, lifestyle, and being overweight: secondary data analyses of middle-aged and older Chinese adults. Health Qual Life Outcomes 2019 Jun 11;17(1):100. [doi: 10.1186/s12955-019-1172-3] [Medline: 31186044]
47. Meijer M, Kejs AM, Stock C, Bloomfield K, Ejstrud B, Schlattmann P. Population density, socioeconomic environment and all-cause mortality: a multilevel survival analysis of 2.7 million individuals in Denmark. Health Place 2012 Mar;18(2):391-399. [doi: 10.1016/j.healthplace.2011.12.001] [Medline: 22209409]
48. Verghese J, Ayers E, Sathyan S, et al. Trajectories of frailty in aging: prospective cohort study. PLOS ONE 2021;16(7):e0253976. [doi: 10.1371/journal.pone.0253976] [Medline: 34252094]
49. Shinkai S, Yoshida H, Taniguchi Y, et al. Public health approach to preventing frailty in the community and its effect on healthy aging in Japan. Geriatr Gerontol Int 2016 Mar;16 Suppl 1(87-97):87-97. [doi: 10.1111/ggi.12726] [Medline: 27018287]
50. Cesari M, Prince M, Thiyagarajan JA, et al. Frailty: an emerging public health priority. J Am Med Dir Assoc 2016 Mar 1;17(3):188-192. [doi: 10.1016/j.jamda.2015.12.016] [Medline: 26805753]

XSL•FO
RenderX

51. Dawson-Townsend K. Social participation patterns and their associations with health and well-being for older adults. SSM Popul Health 2019 Aug;8(100424):100424. [doi: 10.1016/j.ssmph.2019.100424] [Medline: 31312713]

Abbreviations

ANOVA: analysis of variance
CEC: Comité de Ética Científica
GIS: geographic information systems
IDE: Infraestructura de Datos Geospaciales

Edited by A Garcia; submitted 23.07.24; peer-reviewed by J Valenzuela, TTT Nguyen; revised version received 02.12.24; accepted 15.12.24; published 17.04.25.

Please cite as:
Ormazábal Y, Arauna D, Cantillana JC, Palomo I, Fuentes E, Mena C
Unveiling the Frailty Spatial Patterns Among Chilean Older Persons by Exploring Sociodemographic and Urbanistic Influences Based on Geographic Information Systems: Cross-Sectional Study
JMIR Aging 2025;8:e64254
URL: https://aging.jmir.org/2025/1/e64254
doi:10.2196/64254

© Yony Ormazábal, Diego Arauna, Juan Carlos Cantillana, Iván Palomo, Eduardo Fuentes, Carlos Mena. Originally published in JMIR Aging (https://aging.jmir.org), 17.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Comparison of Patient and Provider Perspectives on an Electronic Health Record–Based Discharge Communication Tool: Survey Study

Dorothy Yingxuan Wang[1,2], PhD; Eliza Lai-Yi Wong[1,2], PhD; Annie Wai-Ling Cheung[1,2], MPhil; Kam-Shing Tang[3], MD; Eng-Kiong Yeoh[1,2], MBBS

[1]JC School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

[2]Centre for Health Systems & Policy Research, JC School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

[3]Kwong Wah Hospital, Hospital Authority, Hong Kong, China (Hong Kong)

Corresponding Author:
Eliza Lai-Yi Wong, PhD
JC School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

Abstract

Background: Hospital discharge for older adult patients carries risks. Effective patient-provider communication is crucial for postacute care. Technology-based communication tools are promising in improving patient experience and outcomes. However, there is limited evidence comparing patient and provider user experiences on a large-scale basis, hindering the exploration of true patient-provider shared understanding.

Objective: This study aimed to evaluate an electronic health record–based discharge communication tool by examining and comparing patient and provider perspectives.

Methods: This study comprised a cross-sectional self-administered staff survey and a pre-post cross-sectional patient survey. Physicians, nurses, and older adult patients aged 65 years and older discharged from 4 public hospitals were included. Patient-provider comparison items focused on 3 aspects of the design quality of the tool (information clarity, adequacy, and usefulness) and overall satisfaction with the tool. In addition, patients' experience of discharge information and their medication-taking behaviors before and after the program implementation were compared based on a validated local patient experience survey instrument. Providers' perceived usefulness of this tool to their work and implementation intentions were measured based on the technology acceptance model to enhance understanding of their experiences by conducting structural equation modeling analysis.

Results: A total of 1375 and 2353 valid responses were received from providers and patients, respectively. Patients' overall satisfaction with this communication tool is significantly higher than providers', and patients rated the information clarity and usefulness presented by this tool higher as well (P<.001). However, patients rated information adequacy significantly lower than providers (P<.001). Meanwhile, patients reported a significant improvement in their experience of discharge medication information, and fewer patients reported side effects encounters after the program implementation (126/1083, 11.6% vs 111/1235, 9%; P=.04). However, providers showed inconsistent implementation fidelity. Providers' perceived quality of the tool design (β coefficient=0.24, 95% CI 0.08-0.40) and perceived usefulness to their work (β coefficient=0.57, 95% CI 0.43-0.71) significantly impacted their satisfaction. Satisfaction can significantly impact implementation intentions (β coefficient=0.40, 95% CI 0.17-0.64), which further impacts implementation behaviors (β coefficient=0.16, 95% CI 0.10-0.23).

Conclusions: A notable disparity exists between patients and health care providers. This may hinder the achievement of the tool's benefits. Future research should aim for a comprehensive overview of implementation barriers and corresponding strategies to enhance staff performance and facilitate patient-provider shared understanding.

(JMIR Aging 2025;8:e60506) doi:10.2196/60506

KEYWORDS

older adult; gerontology; geriatric; old; older; elderly; aging; aged; post-acute care; communication; satisfaction; medication information; patient-provider comparison; technology-based intervention; technology acceptance model; discharge; EHR; record; portal; cross-sectional; survey; questionnaire; experience; attitude; opinion; perception; perspective; acceptance; adoption; design; user experience

XSL•FO
RenderX

Introduction

At hospital discharge, health care providers play a crucial role in delivering comprehensive medication information, including side effects and warnings, to ensure safe medication therapy during patients' postacute care [1]. However, previous literature has reported that older adult patients often lack awareness or understanding of their medication regimen after being discharged home [2-4]. Insufficient knowledge is associated with their suboptimal adherence to treatment [5,6], an elevated likelihood of adverse events [7], increased readmissions and emergency department visits [8], and burden on the health care system [9].

A wide array of communication strategies has been documented in the literature with the aim of facilitating information provision by health care providers and enhancing patient awareness and understanding of health-related information [10]. Their effectiveness in reducing readmissions and enhancing patient satisfaction was supported by a recent meta-analysis [1]. Notably, information technology–based communication practices have emerged as a prominent and preferred mode for delivering discharge information, as highlighted in literature reviews [11]. In addition, a systematic review concluded that computer-enabled discharge communication interventions improve both patient and provider satisfaction and reduce perceived adverse events [12]. A Cochrane review further indicated that computer-generated reminders presented on paper can enhance the quality of care [13]. However, there is a scarcity of research comparing the perspectives of older adult patients and health care providers with concordance measures for a large-scale technology-based discharge communication tool. Measuring and comparing the alignment between patient and provider perspectives enables the unveiling of true shared understanding in terms of discharge education [14].

In Hong Kong, the provision of discharge medication information, particularly regarding side effects and warnings, was found to be suboptimal, according to a regular patient experience survey [15]. In 2017, the Hospital Authority developed a computer-generated written medication reminder called the postdischarge information summary (PDIS) to address this issue [16]. The key components of the PDIS were co-designed by a multidisciplinary program team consisting of government officials, clinicians, quality and safety representatives, and technology experts. The first component includes a salient medication reminder, a computer-based drug database encompassing 58 prescribed medications for local older adult patients, and 235 most pertinent side effects and warning items. This database underwent validation through 3 rounds of Delphi expert consensus meetings [17]. The second component comprises a list of follow-up appointments across all Hong Kong public hospitals. The PDIS system generates personalized information by integrating into the electronic health record (EHR). During discharge, physicians or nurses are required to print the written summary through the PDIS system and distribute it to discharged patients or their caregivers, along with a detailed explanation of its contents. No teach-back was required at the moment of program introduction. A comparison of the discharge communication workflow between usual practice and PDIS-incorporated practice is shown in Multimedia Appendix 1.

The objective of this study is to evaluate this EHR-based discharge communication tool by examining and contrasting the perspectives of both older adult patients and health care providers.

Methods

Study Design

This study comprises a self-administered cross-sectional staff survey and a pre-post cross-sectional patient survey. The pre-post patient surveys were conducted among 2 different patient groups. The STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guideline [18] was used to strengthen the reporting process (Checklist 1).

Setting and Sampling

The study involved 4 piloting public tertiary hospitals representing 3 out of 7 geographical clusters of Hong Kong. The PDIS was introduced in a phased manner within the geriatric and medicine department in January 2018 [16]. For the staff survey, all physicians and nurses involved in the PDIS implementation were invited to participate in the study. The surveys were conducted at least 6 months after the PDIS implementation, which spanned from August 2018 to June 2019. Paper-based promotion leaflets, invitation letters, and questionnaires were distributed through designated coordinators in each hospital. This survey was conducted anonymously and on a voluntary basis.

For the patient survey, the sample size was determined based on the inpatient discharge statistics for patients aged 65 years or older in 2015, as provided by the Hospital Authority. In order to achieve a precision level of ±4% with a 95% CI, a minimum of 1450 respondents was required for pre-post rounds. Assuming a 50% response rate, at least 2900 patients were randomly selected from the discharge records for each round. Responses from caregivers acting as surrogates were accepted if patients were unable to respond independently. Readmitted cases and day patients were excluded. Within 14 days of their discharge, patients were contacted by telephone. The pre-post survey was conducted from June to December 2017 and May to December 2018, accordingly.

The staff survey used English and the patient version used Chinese. We used English in staff survey because English is their working language and they are proficient in English. Furthermore, staff surveys are typically conducted in English in Hong Kong, so it was assumed that participants would feel comfortable with this language. To ensure language did not pose a barrier to participation or affect the responses collected, we decided to use Chinese for the patient survey. We do not anticipate any language concordance issues because the researchers who were responsible for collecting patient responses were well-trained before the survey to ensure they can correctly convey the meaning of the questions.

XSL•FO
RenderX

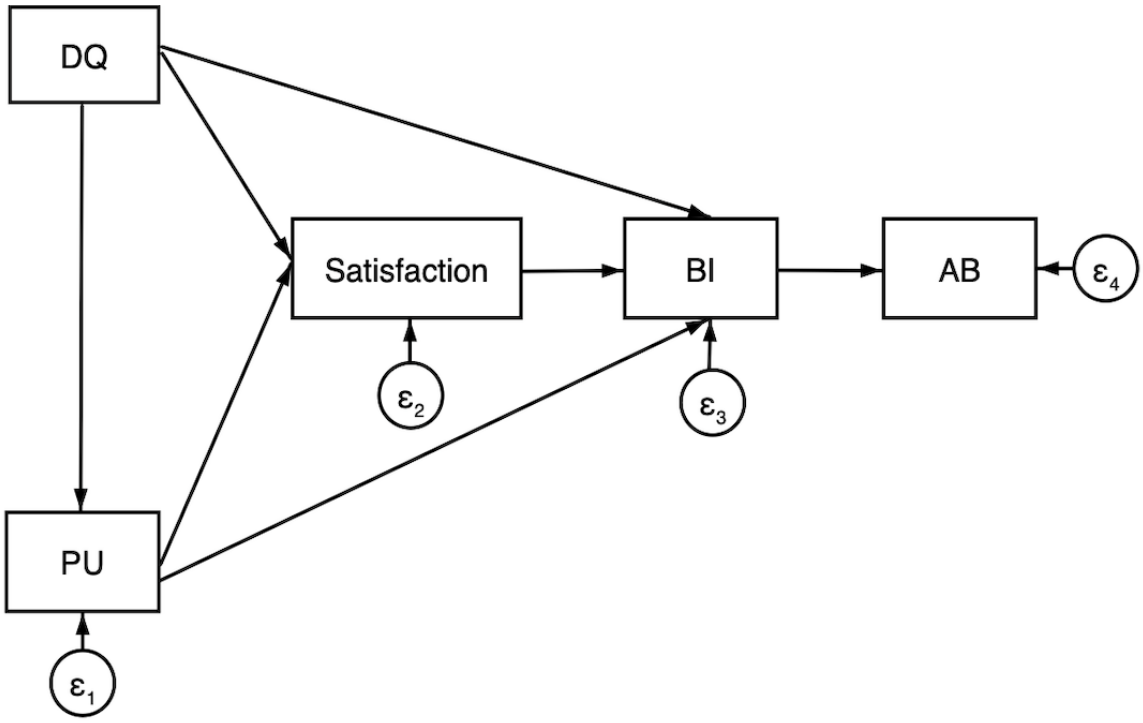
Theoretical Framework

The staff survey collected information on providers' practicing behaviors and user experience, adapted by the technology acceptance model (TAM) [19]. TAM has been designed to investigate why individuals adopt a specific technology and has been widely used in different settings [20].

According to the TAM, the perceived usefulness of the technology can impact users' behavior intention, which can be a determinant of users' actual behaviors. We added another domain named design quality to capture providers' perceived information quality of the tool. We also measured the overall satisfaction of the PDIS experience. Existing literature suggests that design quality can significantly influence staff adoption and implementation of the technologies [21]. The traditional TAM component of "Perceived Ease of Use" was not fully applicable in this context, as our primary aim was to compare the perceptions of patients and staff regarding their user experiences and perceptions of the PDIS. However, "Perceived Ease of Use" tends to be more relevant to staff, given that staff interact with the information system while patients interact with a paper-based version rather than a technology. Consequently, we adopted "Design Quality" as a domain to facilitate a meaningful comparison of perspectives between these 2 groups. In addition, we sought to streamline our survey to enhance response rates, particularly among staff who may have limited time to participate. Thus, we decided to use design quality as a domain in place of perceived ease of use.

We hypothesized that (1) design quality would have an impact on perceived usefulness, satisfaction, and behavior intention; (2) perceived usefulness would impact satisfaction and behavior intention; (3) satisfaction would impact behavior intention; and (4) behavior intention would impact actual behavior (Figure 1).

Figure 1. Conceptual framework of the factors impacting implementation fidelity. AB: actual behavior; BI: behavior intention; DQ: design quality; PU: perceived usefulness; "ε1", "ε2", "ε3", and "ε4" were the residual errors.



Measurements

The Staff Survey

Perceived Usefulness

In total, 4 items were used to measure perceived usefulness: (1) "PDIS supports my medication education," (2) "PDIS enhances my communication with patients or caregivers," (3) "PDIS enhances my job efficiency," and (4) "I find PDIS useful in my job." The Cronbach α for this domain was 0.97, indicating the good reliability. A Likert scale ranging from 0 ("strongly disagree") to 10 ("strongly agree") was used for these 4 items.

Design Quality

Design quality was measured using 3 items developed by the research team: (1) "The information provided by the PDIS is clear," (2) "The information provided by the PDIS is sufficient," and (3) "The information provided by the PDIS is useful to your patients or caregivers." The Cronbach α for this domain was 0.92, indicating good reliability. A 10-point Likert scale was used to measure these three items ranging from 0 ("strongly disagree") to 10 ("strongly agree").

Behavior Intention

One item, which was investigator team constructed, was adopted to measure behavior intention: "I am willing to use PDIS to deliver patient discharge information in the future." The 10-point

XSL•FO
RenderX

Likert scale was used (0: strongly disagree to 10: strongly agree) for this item.

Satisfaction

Satisfaction was measured using a single item: "How would you rate your experience of delivering patient discharge information using PDIS?" A 10-point Likert scale ranging from 0 ("very poor") to 10 ("very good") was used for this question.

Actual Behaviors

The actual behavior was defined as the implementation fidelity and measured by the frequency of distributing, reading, and explaining the PDIS to patients and caregivers with a 3-level ordinal scale (never, sometimes, and always). A composite score for actual behavior was created by converting the ordinal scale for each behavior into to a 0, 1, and 2 scale and summing it up across the 3 practicing behaviors (distribution, read, and explanation) to generate a total score.

Qualitative Comments

Free-text fields were provided to solicit provides' comments on their PDIS experiences. Respondents were prompted with the following question: "Is there anything you would like to tell us about the PDIS (eg, things that were particularly good, areas for improvement, or any other comments)?"

Demographic Information

Demographic information was collected in the final section of the survey. For example, we asked their self-reported gender, age, working experience in years, and professional role (eg, interns, residents, specialists, associate consultants, and consultants for physicians; enrolled nurses, registered nurses, advanced practice nurses, and ward managers for nurses).

The Patient Survey

The pre-post patient surveys consist of the following sections.

The Patient Experience

The items were drawn from the validated local assessment tool, the Short-form Hong Kong Inpatient Experience Questionnaire (SF-HKIEQ) [22], soliciting patient agreement on the clarity, adequacy, and usefulness of discharge medication information, including side effects and warnings, using an ordinal scale (yes, to some extent, or no) or Likert scale of 0 - 10 from strongly disagree to strongly agree; the Cronbach α is 0.87, indicating good reliability. The overall satisfaction of the PDIS experience was also solicited. A 0 - 10 Likert scale of strongly disagree/very bad to strongly agree/very good was applied to measure the items. In the patient survey, consistent with the previous analysis approach for the SF-HKIEQ, questions with an ordinal scale were converted to a 0, 5, and 10 scale and aggregated to calculate the mean and SD.

Self-Reported Medication-Taking Behavior

This section was developed based on the 4-item Morisky, Green, Levine (MGL) scale [23] and relevant studies [24]. Furthermore, 1 item was adopted from the MGL scale: "Ever forget to take medicines?" In addition, 2 items were investigator-constructed: "Whether you were compliant with their medication regimen" and "Whether you have ever experienced medication side effects." These 3 items were measured by binary responses (yes or no).

Design Quality

The patient postsurvey includes the same question of the design quality domain of the staff survey, allowing comparative analysis: (1) "The PDIS information is clear to me or patients," (2) "The PDIS information is enough to me or patients," and (3) "The PDIS information is useful to self-care or care for patients." A 10-point Likert scale was used (0: "strongly disagree" to 10: "strongly agree"). The Cronbach α was 0.94, indicating a high reliability.

Free-Text Field

A free-text section was added to solicit patients' comments on the PDIS experiences. Similarly with staff survey, we asked patients "what information do you need that was not provided by the PDIS, and any other comments?"

Patient Characteristics

Patient characteristics, such as their age, self-reported gender, education level (primary, secondary, college, and above), whether received the government subsidy, the comprehensive social security assistance designed for people whose income is not sufficient to meet basic needs), living arrangements (living alone or living with others), chronic conditions (including heart disease, hypertension, type 2 diabetes, and cancer), and self-reported quality of life using visual analog scale of the EQ-5D-5L Hong Kong [25], were also asked.

Data Analysis

The statistical analysis was performed utilizing R version 4.0.5 (R Project for Statistical Computing) and Stata version 18 for Mac (StataCorp). Provider and patient demographic information were summarized as means and percentages using descriptive statistical analysis. The staff survey and patient survey were analyzed separately. For the staff survey, subgroup analysis was conducted to evaluate the differences among physicians and nurses. The practicing behavior frequency and PDIS experiences were compared using the Pearson chi-square test and Mann-Whitney *U* test. The Mann-Whitney *U* test was used due to the outcome variables were skewed. In order to understand the relationship between the determinants of the behaviors like design quality, perceived usefulness, behavior intention, satisfaction of the PDIS, and the actual behavior, as well as the relationship between different determinants, the covariance-based structural equation modeling (CB-SEM) was applied. The goodness of fit index, comparative fit index (CFI), root mean square error approximation (RMSEA), Tucker-Lewis index (TLI), and standard root mean squared residual (SRMSR) were used to evaluate the goodness-of-fit of the model. The direct effects, indirect effects, and total effects between the core variables were assessed using a bootstrapping approach (n=2000). Standardized path coefficients (β) were used to estimate the path relationships. Statistical significance was set at P=.05.

For the patient survey, changes in patient experience regarding medication information before and after PDIS implementation were compared using the Mann-Whitney *U* test due to the

XSL•FO
RenderX

outcome variables were skewed. The difference between self-reported side effects encounters and compliance was assessed using the Pearson chi-square test.

To identify the disparity between staff and patient. the shared questions related to PDIS experience (eg, information clarity, adequacy, and usefulness, and the overall experience of the PDIS) in the staff and patient postsurveys were compared using the Mann-Whitney *U* test. For the qualitative comments from both staff and patient, thematic synthesis [26] was applied to identify and compare the common themes in free-text comments for the PDIS program between patients and providers. Furthermore, 2 coders independently coded all the utterances of patients and staffs to generate the initial codes to achieve consensus on coding. The disagreement were formally resolved through several discussions with a senior researcher experienced in qualitative study. After that, the inductive analysis was performed by one coder to generalize initial codes into overarching themes. The initial summary themes were developed and discussed with the senior research to reach consensus. The final list of themes was reviewed and consented by the research team. The theme list were identical for patient and staff and the frequency of each theme were counted in order to compare the pattern for patients and staffs.

Ethical Considerations

Ethical approval has been provided by the Joint Chinese University of Hong Kong – New Territories East Cluster Clinical Research Ethics Committee in compliance with the Declaration of Helsinki (CREC 2019.436). For the staff survey, implied consent was applied as participating staff members returned the completed questionnaires to the research team. For the patient survey, verbal consent was obtained from patients over the phone before the survey. All data were deidentified to protect participants' privacy and confidentiality.

Results

Comparative Analysis

A total of 1375 providers completed the survey with a 76% response rate, comprising 966 (72%) female participants, 650 (50%) participants aged 18 - 29 years old, 595 (50%) participants with 0 - 5 years of working experience, and 1216 (88%) nurses (Table 1).

Table . Demographic information of health care providers.

Characteristics		Total (N=1375)	Doctors (n=159)	Nurses (n=1216)
Sex, n (%) [a]				
	Female	966 (72)	55 (35)	911 (77)
	Male	382 (28)	103 (65)	279 (23)
Age (years), n (%) [b]				
	18-29	650 (50)	52 (34)	598 (52)
	30-39	349 (27)	50 (32)	299 (26)
	40-49	206 (16)	32 (21)	174 (15)
	50-59	86 (6.6)	21 (14)	65 (5.7)
	>59	5 (0.4)	0 (0)	5 (0.4)
Working experience (years), n (%) [c]				
	0-5	595 (50)	46 (34)	549 (53)
	6-10	278 (24)	25 (18)	253 (24)
	11-15	94 (8.0)	23 (17)	71 (6.8)
	16-20	125 (11)	18 (13)	107 (10)
	20-25	42 (3.6)	13 (9.6)	29 (2.8)
	26-30	37 (3.1)	10 (7.4)	27 (2.6)
	>30	9 (0.8)	1 (0.7)	8 (0.8)

[a]31 participants excluded from analysis due to missing information.

[b]82 participants excluded from analysis due to missing information.

[c]185 participants excluded from analysis due to missing information.

From the patient side, we received 2353 valid responses, including 1109 (47%) and 1244 (53%) responses collected through the pre- and postsurveys, respectively. The response rate was 55.5% for the presurvey and 59.4% for the postsurvey. The demographic composition was similar between the pre- and postsurvey groups, except that 6.4% (presurvey group: 853/1106; postsurvey group: 1023/1226) more participants were receiving the government subsidy in the postsurvey group (P<.001) (Table 2).

XSL•FO
RenderX

Table . Demographic information of older adult patients.

Characteristics		Total (N=2353)	Presurvey (n=1109)	Postsurvey (n=1244)
Age (years), mean (SD)		77.48 (7.98)	77.54 (8.00)	77.65 (7.93)
Sex, n (%)				
	Female	1070 (45.5)	517 (46.6)	553 (44.5)
	Male	1283 (54.5)	592 (53.4)	691 (55.5)
Education, n (%) [a]				
	≤Primary	1461 (62.8)	695 (63.4)	946 (62.3)
	Secondary	702 (30.2)	341 (31.1)	361 (29.4)
	≥College	162 (7)	60 (5.5)	102 (8.3)
Living status, n (%) [b]				
	Living alone	292 (12.4)	138 (12.4)	154 (12.4)
	Living with others	2056 (87.6)	971 (87.6)	1085 (87.6)
Government subsidy, n (%) [c]				
	Yes	1880 (80.5)	853 (77.1)	1023 (83.5)
	No	456 (19.5)	253 (22.9)	203 (16.5)
Heart diseases, n (%) [d]				
	Yes	844 (36.1)	398 (35.9)	446 (36.3)
	No	1492 (63.9)	710 (64.1)	782 (63.7)
Hypertension, n (%) [e]				
	Yes	1380 (59.1)	671 (60.6)	709 (57.7)
	No	956 (40.9)	437 (39.4)	519 (42.3)
Type 2 diabetes, n (%) [f]				
	Yes	754 (32.3)	358 (32.3)	396 (32.2)
	No	1582 (67.7)	750 (67.7)	832 (67.8)
Cancer, n (%) [g]				
	Yes	156 (6.7)	85 (7.7)	71 (5.8)
	No	2181 (93.3)	1023 (92.3)	1158 (94.2)
Length of stay (day), n (%) [h]				
	0-3	1256 (53.7)	592 (53.6)	664 (53.7)
	4-7	639 (27.3)	296 (26.8)	343 (27.7)
	>7	446 (19.1)	216 (19.6)	230 (18.6)
EQ-5D-VAS, mean (SD)[i]		66.61 (18.51)	68.37 (17.39)	65.04 (19.33)
Discharge day, n (%)				
	Weekday	1916 (81.5)	909 (82)	1007 (80.9)
	Weekend	437 (18.5)	200 (18)	237 (19.1)

[a]29 participants excluded from analysis due to missing information.
[b]5 participants excluded from analysis due to missing information.
[c]17 participants excluded from analysis due to missing information.
[d]17 participants excluded from analysis due to missing information.
[e]17 participants excluded from analysis due to missing information.
[f]17 participants excluded from analysis due to missing information.
[g]16 participants excluded from analysis due to missing information.

XSL•FO
RenderX

[h]12 participants excluded from analysis due to missing information.

[i]109 participants excluded from analysis due to missing information.

The comparative evaluation showed that patients consistently provided significantly higher ratings for their overall PDIS satisfaction compared with providers (mean 8.28, SD 1.60 vs mean 6.29, SD 1.88, respectively; *P*<.001) (Figure 2). Specifically, patients reported higher ratings for information clarity (mean 8.58, SD 2.50 vs mean 6.54, SD 1.86, respectively; *P*<.001) and usefulness of the PDIS (mean 8.14, SD 2.46 vs mean 6.50, SD 2.02, respectively; *P*<.001). On the contrary, patients were inclined to receive more information through the PDIS (mean 5.33, SD 1.35 vs mean 6.14, SD 2.08, respectively; *P*<.001).

Figure 2. Comparison of patient and health care providers' postdischarge information summary experiences. *P* value was obtained from the Mann-Whitney *U* test. **P*<.001.

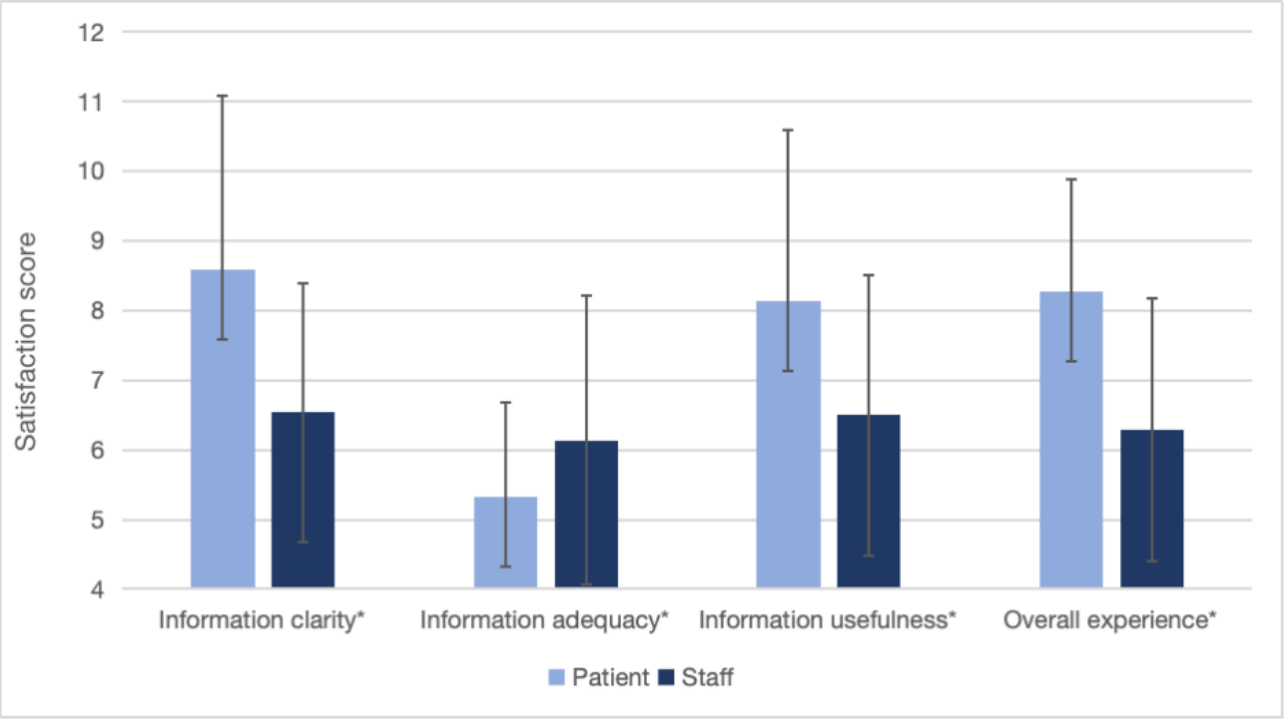
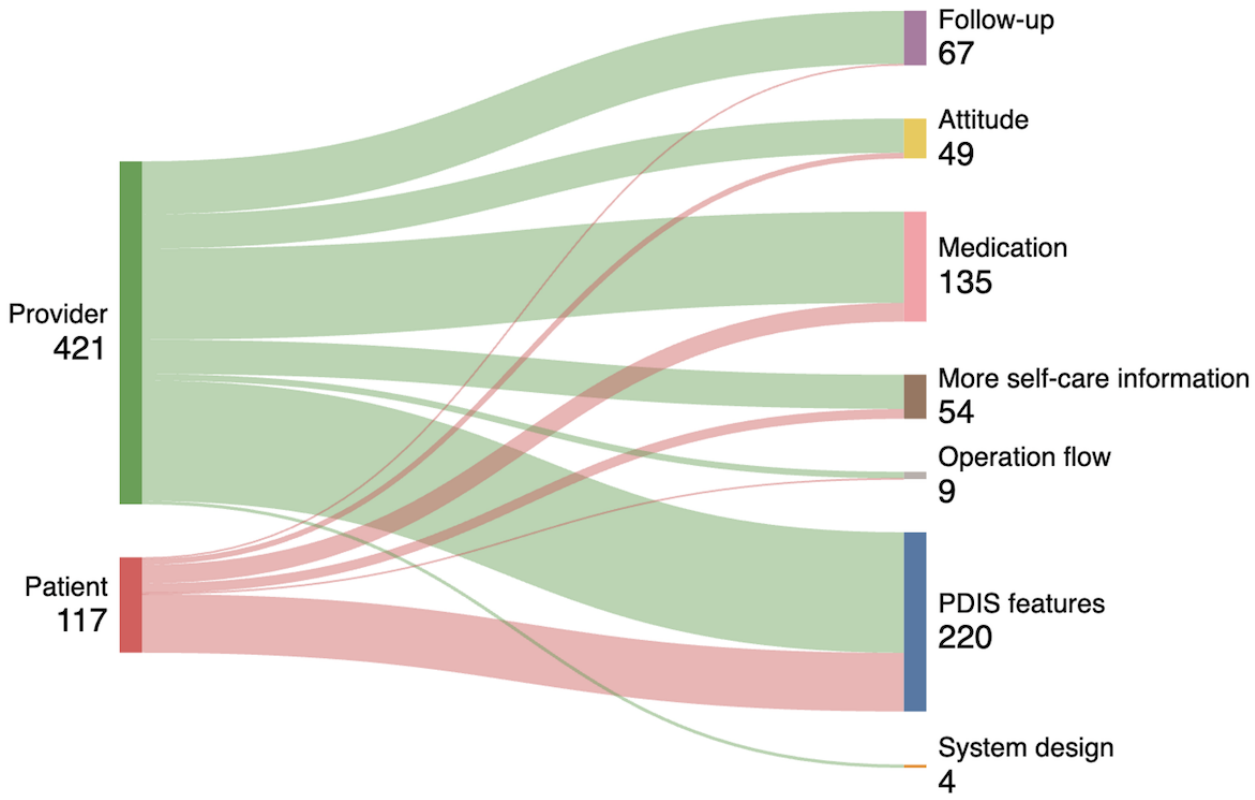


Figure 3 displays the similarities and differences between patients and providers regarding their comments on PDIS experiences. A total of 538 comments were received (421 were from providers and 117 were from patients). The most frequently commented aspect was PDIS features for both providers (147/421, 35%) and patients (73/117, 62%). However, the specific area of the feature was different. Providers emphasized the need for broader coverage of the drug databases (60/147, 41%) and the lack of multiple language versions (59/147, 40%), while patients' concerns revolve around the inconvenience of medication names and follow-up information being in English (45/73, 61%), the discomfort with medical jargon (9/73, 13%), and font size (7/73, 10%). In addition, providers frequently commented on the content of the PDIS form. For example, 27% (114/421) of the comments were related to the medication listed on the PDIS and emphasized the need for additional medication details such as medication changes, indications, and instructions.

XSL•FO
RenderX

Figure 3. Comparison of the qualitative comments on the postdischarge information summary between patients and health care providers. PDIS: postdischarge information summary.



Staff Survey

Analysis of the practicing behaviors (Table 3) revealed that 72.1% (n=922) of the providers reported being able to consistently distribute the PDIS form to patients or caregivers, whereas 56% (n=667) stated they could consistently explain its content. Subgroup analysis demonstrated significant variations across different roles. Regarding distribution, 78% (n=915) of nurses reported always doing so, compared with 6.4% (n=7) of doctors (P<.001). Similarly, 57% (n=666) of nurses reported always explaining the content, compared with 4% (n=1) of doctors (P<.001).

XSL•FO
RenderX

Table . Health care providers' practicing behaviors and experience of the postdischarge information summary.

Items		Total (N=1375)	Doctors (n=159)	Nurses (n=1216)	*P* value[a]
PDIS[b] implementation behavior, n (%)					
Distribute		1234 (89.7)	109 (68.6)	1170 (96.2)	<.001
	Always	922 (72)	7 (6.4)	915 (78)	
	Sometimes	261 (20)	17 (16)	244 (21)	
	Never	96 (7.5)	85 (78)	11 (0.9)	
Read		1190 (86.5)	26 (16.4)	1154 (94.9)	<.001
	Always	730 (61)	5 (19)	725 (62)	
	Sometimes	425 (36)	16 (62)	409 (35)	
	Never	35 (2.9)	5 (19)	30 (2.6)	
Explain		1190 (86.5)	26 (16.4)	1164 (95.7)	<.001
	Always	667 (56)	1 (3.8)	666 (57)	
	Sometimes	475 (40)	11 (42)	464 (40)	
	Never	48 (4)	14 (54)	34 (2.9)	
Perceptions of PDIS experiences, mean (SD)					
Perceived design quality					
	PDIS information is clear to your patients or careers	6.54 (1.86)	6.50 (1.61)	6.54 (1.86)	.67
	PDIS information is adequate to your patients or careers	6.14 (2.07)	6.24 (1.61)	6.14 (2.08)	.93
	PDIS information is useful for patients or careers	6.50 (2.02)	6.69 (1.81)	6.50 (2.02)	.70
Perceived usefulness					
	PDIS supports my medication education to patients or careers	6.35 (2.07)	6.08 (1.81)	6.36 (2.07)	.32
	Patient-provider communication becomes more effective with PDIS	6.21 (2.05)	6.00 (1.55)	6.21 (2.06)	.34
	PDIS enhances my job efficiency	5.97 (2.20)	6.08 (1.72)	5.97 (2.21)	.83
	PDIS is useful in my job	6.03 (2.16)	6.12 (1.68)	6.02 (2.17)	.97
Behavior intention					
	I am willing to use PDIS	6.09 (2.18)	6.12 (1.63)	6.09 (2.19)	.72
Overall satisfaction					
	Overall rating of PDIS user experiences	6.29 (1.88)	6.56 (1.51)	6.28 (1.88)	.75

[a]*P* value was obtained from the chi-square test and Mann-Whitney *U* test.

[b]PDIS: postdischarge information summary.

Physicians and nurses indicated moderate satisfaction with the design quality and perceived usefulness of the PDIS to their work, as reflected by mean agreement scores ranging between 5.96 and 6.54 (Table 3). The subgroup analysis did not identify

XSL•FO
RenderX

any differences between professional roles regarding user experiences. The CB-SEM analysis (Figure 4) showed that design quality significantly impacted their perceived usefulness (β coefficient=0.96, 95% CI 0.90-1.01) and behavior intention (β coefficient=0.14, 95% CI 0.06-0.21). In addition, perceived usefulness had a significant impact on behavior intention (β coefficient=0.48, 95% CI 0.26-0.70). Furthermore, behavior intention had a significant impact on the actual behavior (β coefficient=0.16, 95% CI 0.10-0.23). In addition, satisfaction can be significantly impacted by the design quality (β coefficient=0.24, 95% CI 0.08-0.40) and perceived usefulness (β coefficient=0.57, 95% CI 0.43-0.71). The structural equation modeling (SEM) model presents a good fit overall, with all indicators exceeding the recommended thresholds (Table 4). The results of the indirect effects and total effects can be found in Multimedia Appendices 2 and 3. In total, 3 mediating pathways were identified: (1) an indirect pathway from design quality through satisfaction to behavior intention (β coefficient=0.770, 95% CI 0.700 - 0.841, proportion mediated=15.1%); (2) an indirect pathway from perceived usefulness through satisfaction to behavior intention (β coefficient=0.228, 95% CI 0.185 - 0.272, proportion mediated=32.5%); and (3) an indirect pathway from design quality through perceived usefulness to satisfaction (β coefficient=0.544, 95% CI 0.482 - 0.605, proportion mediated=30.8%).

Figure 4. Structural equation modeling for factors impacting the providers' implementation fidelity. AB: actual behavior; BI: behavior intention; DQ: design quality; PU: perceived usefulness; "ε1", "ε2", "ε3", and "ε4" were the residual errors.

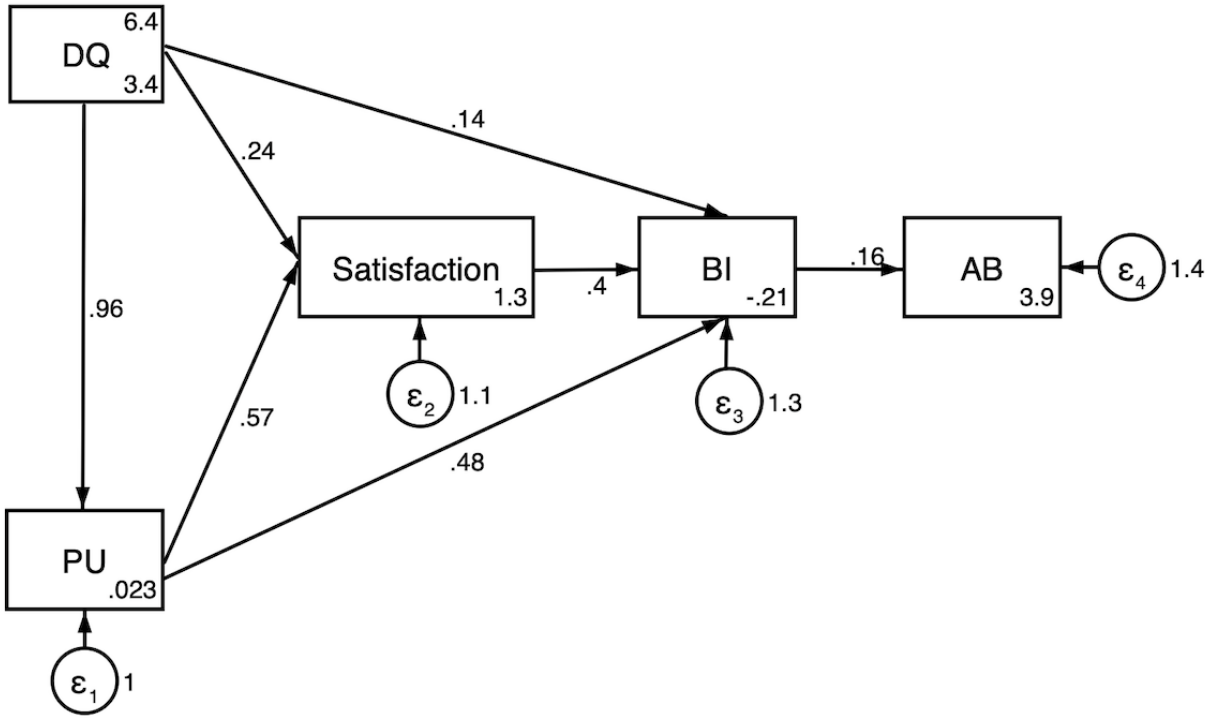


Table . Direct effects for the model[a].

Relationship	Standardized estimates (95% CI)	Remarks
Design quality → Perceived usefulness	0.955 (0.923-0.987)	Supported
Design quality → Satisfaction	0.242 (0.151 - 0.333)	Supported
Perceived usefulness → Satisfaction	0.569 (0.487-0.651)	Supported
Design quality → Behavior intention	0.137 (0.031-0.242)	Supported
Perceived usefulness → Behavior intention	0.476 (0.371-0.582)	Supported
Satisfaction → Behavior intention	0.402 (0.288-0.515)	Supported
Behavior intention → Actual behavior	0.164 (0.131-0.197)	Supported

[a]Model fit: χ_{441}^2=3088.344, *P*<.001; root mean square error approximation=0.052; comparative fit index=0.997; Tucker-Lewis index=0.990; standard root mean squared residual=0.021.

XSL•FO
RenderX

Patient Survey

Table 5 showed significant improvements in patient experience related to overall discharge information clarity (mean 8.18, SD 1.69 vs mean 7.93, SD 1.84, respectively; *P*=.002), adequacy (mean 8.15, SD 1.76 vs mean 7.92, SD 1.93, respectively; *P*=.01), and usefulness (mean 8.26, SD 1.70 vs mean 8.06, SD 1.80, respectively; *P*=.02). In addition, a significant increase was found in information adequacy for both side effects (mean 9.6, SD 2.0 vs mean 8.6, SD 3.4, respectively; *P*<.001) and warnings (mean 9.7, SD 1.8 vs mean 9.2, SD 2.7, respectively; *P*=.004). Only warning information reached statistically significant improvement in clarity (mean 8.77, SD 2.32 vs mean 8.45, SD 2.45, respectively; *P*=.03) and usefulness (mean 8.7, SD 2.45 vs mean 8.44, SD 2.39, respectively; *P*=.03).

Table . Older adult patients' or caregivers' perceptions of discharge medication information and medication-taking behavior between pre- and postsurvey groups.

Items		Total (N=2353)	Presurvey (n=1109)	Postsurvey (n=1244)	*P* value[a]
Perspectives of the discharge medication information					
Clarity, mean (SD)					
	Side effects	8.47 (2.50)	8.31 (2.63)	8.60 (2.39)	.10
	Warning signs	8.61 (2.39)	8.45 (2.45)	8.77 (2.32)	.03
	Overall	8.06 (1.77)	7.93 (1.84)	8.18 (1.69)	.003
Adequacy, mean (SD)					
	Side effects	9.2 (2.8)	8.6 (3.4)	9.6 (2.0)	<.001
	Warning signs	9.4 (2.3)	9.2 (2.7)	9.7 (1.8)	.004
	Overall	8.04 (1.85)	7.92 (1.93)	8.15 (1.76)	.02
Usefulness, mean (SD)					
	Side effects	8.47 (2.46)	8.34 (2.51)	8.57 (2.42)	.12
	Warning signs	8.57 (2.42)	8.44 (2.39)	8.70 (2.45)	.03
	Overall	8.17 (1.75)	8.06 (1.80)	8.26 (1.70)	.01
Medication-taking behaviors					
Self-reported side effects encounter, n (%)		2318 (86.1)	1083 (85.6)	1235 (86.6)	.04
	Yes	237 (10.2)	126 (11.6)	111 (9)	
	No	2081 (89.8)	957 (88.4)	1124 (91)	
Self-reported medication compliance, n (%)		2325 (86.4)	1085 (85.8)	1240 (87)	.39
	Yes	2226 (95.7)	1043 (96.1)	1183 (95.4)	
	No	99 (4.3)	42 (3.9)	57 (4.6)	

[a]*P* value was obtained from the Mann-Whitney *U* test and chi-square test.

No statistically significant difference was found between pre- and postsurvey groups in the percentage of self-reported medication compliance. Notably, the postsurvey group had a significantly lower percentage of self-reported side effects encounters (126/1083, 11.6% vs 111/1235, 9%, respectively; *P*=.04). Among participants who reported encountering side effects, the majority (96/111, 86.4%) still followed the medication instructions as prescribed.

Discussion

Principal Results

This is the first study to compare the perceptions of older adult patients and health care providers regarding the use of large-scale EHR-based discharge communication tools with concordance measures. There was a noticeable difference in ratings between patients and providers, with patients giving higher ratings in terms of design quality and overall experience of this tool. Qualitative comments indicated that patients and providers have different areas of concern regarding this communication tool. Furthermore, from the health care providers' side, inconsistent practicing behaviors were found, which were significantly influenced by the implementation intentions (represented as behavioral intentions in the SEM model), overall satisfaction, design quality, and perceived usefulness of the program. However, from the patients' side, older adults who received the written summary reported improved experiences with discharge information, including information clarity, adequacy, and usefulness.

XSL•FO
RenderX

Comparison With Previous Work

Providers assigned significantly lower scores to the perceived information clarity of the PDIS to their patients than patients themselves. This difference may be due to the providers' concern about the challenges associated with older adults' health literacy [27] and the potential negative consequences of sharing information on side effects, such as patients' anxiety and non-compliance [28-30]. However, the improved patient-reported ratings of information clarity on medication warnings and overall medication information, significantly fewer side effects encounters, and no evidence of patient noncompliance found by our study and others [31] suggested that taking action is no worse than inaction but yields better outcomes, contrary to the biased perception held by staff members [32]. In order to address staff concerns, facilitate their implementation, and fulfill patients' needs, rephrasing risk information by using lay language, shorter sentences, supplementing verbal descriptions with visual aids, and presenting medication benefits along with side effects can be considered [33-35].

Providers rated information usefulness for patients or caregivers lower than patients in this study, further impacting their implementation. The lower beliefs on the value of the communication tool for patient care from the provider side may be due to the beliefs that patients may not effectively follow the advice due to a lack of skills or inability to recall, despite clinicians appropriately delivering the instructions [36]. Therefore, it is suggested that using cognitive aid strategies such as teach-back techniques, repetition, demonstration, and reducing the complexity of the information to enhance patients' capacity to perform self-care tasks and recall of information [10]. The discrepancy can also be attributed to providers' lack of awareness regarding patient needs, which was also reported by previous studies [37]. Our study results, as well as other research [38], suggest that patients view information about medication side effects and warnings as crucial when making decisions about seeking professional assistance. Patient-provider information gap may not only lead to patient dissatisfaction but also levy stress on providers [37]. Therefore, it is important to leverage patient voices as credible sources and build long-term patient-provider relationships to address this gap.

It is important to note that older adults tend to rely more on health care providers and perceive them as trustworthy sources of information, as reflected in higher satisfaction with received medication information among older adults than younger individuals in a previous study [39]. Therefore, it is crucial to address provider-reported barriers in PDIS implementation to ensure a higher level of program satisfaction and implementation fidelity. Our study found that the design quality and perceived usefulness to providers' work can hinder their implementation, which is in accordance with previous studies [40,41]. This can be attributed to the low compatibility of the service with providers' existing workflow [21]. Our study found that providers expressed a preference for enriching the PDIS content with additional medication elements, such as medication changes, indications, and instructions, indicating their priorities for discharge education are not fully met. This suggested that involving frontline implementers as program designers and developers is crucial for program fit, staff self-efficacy, capacity, and commitment [21]. Other than the perceived usefulness and design quality of the program reported in our study, a comprehensive understanding of the complex elements involved in implementation, including context, stakeholders, and organizational factors, is needed [42]. This knowledge facilitates the creation of customized strategies and policies that have a higher likelihood of achieving success [43]. Therefore, conducting implementation research is essential in identifying the barriers and facilitators linked to the implementation of the PDIS to ensure providers' user experiences are optimized, leading to improved patient access to high-quality care and maintaining a high level of patient experiences among the older adult population.

Limitations

This study has several limitations. First, self-reported data may be subject to inaccuracies due to social desirability bias. However, including participants from diverse backgrounds may mitigate this limitation to some extent. Second, patient outcomes were not measured in this study, as our focus was on exploring and comparing patient and provider experiences with technology-based communication tools. Given the positive experiences reported by patients, future research could investigate clinical outcomes to further enhance the evidence base. Third, the pre-post survey design for a patient survey without a control group limits our ability to determine whether the observed changes are directly attributable to the PDIS. The cross-sectional design for the staff survey can limit our ability to establish causal inferences regarding factors influencing providers' inconsistent implementation. Subsequent studies using longitudinal or experimental designs are warranted to understand the causal mechanisms and develop effective strategies to enhance staff performance. Finally, as we did not impose a strict designation of specific professional roles (doctor or nurse) for the tasks involved in PDIS implementation, there may be an underestimation of the effects of the determinants on behavioral intention and actual behaviors.

Conclusions

EHR-based discharge communication tools have the potential to improve the patient experience with discharge information. However, there is a notable difference in user perceptions between patients and providers. This difference may hinder the full benefits of the program for patients. These findings have implications for future research, particularly in implementation research, where barriers and strategies to enhance staff performance can be investigated. In addition, the study provides valuable insights for organizations seeking to improve patient-provider shared understanding of postacute care plans among older adult patients during hospitalization, particularly through technology-based interventions.

XSL•FO
RenderX

Acknowledgments

We are grateful to the steering committee members for their professional advice regarding the implementation of the postdischarge information summary. We acknowledge the assistance of Ms Tang PF. We would also like to thank the HA Head Office and the cluster coordinators for their kind assistance in the logistics for conducting the study.

Authors' Contributions

ELYW, EKY, and KST were involved in the conception and design of the study. DYXW is the lead author and prepared the article. AWLC was involved in data collection. DYXW, ELYW, and AWLC were involved in data analysis. All the authors were involved in the critical revision and the final approval of the article.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Pre- and postdischarge information summary discharge communication workflow.
[DOCX File, 18 KB - aging_v8i1e60506_app1.docx]

Multimedia Appendix 2
Indirect effect for the structural equation modeling model.
[DOCX File, 15 KB - aging_v8i1e60506_app2.docx]

Multimedia Appendix 3
Total effect of the structural equation modeling model.
[DOCX File, 15 KB - aging_v8i1e60506_app3.docx]

Checklist 1
STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement checklist.
[PDF File, 135 KB - aging_v8i1e60506_app4.pdf]

References

1. Becker C, Zumbrunn S, Beck K, et al. Interventions to improve communication at hospital discharge and rates of readmission: a systematic review and meta-analysis. JAMA Netw Open 2021 Aug 2;4(8):e2119346. [doi: 10.1001/jamanetworkopen.2021.19346] [Medline: 34448868]
2. Davis TC, Wolf MS, Bass PF 3rd, et al. Literacy and misunderstanding prescription drug labels. Ann Intern Med 2006 Dec 19;145(12):887-894. [doi: 10.7326/0003-4819-145-12-200612190-00144] [Medline: 17135578]
3. Sanderson BK, Thompson J, Brown TM, Tucker MJ, Bittner V. Assessing patient recall of discharge instructions for acute myocardial infarction. J Healthc Qual 2009;31(6):25-33. [doi: 10.1111/j.1945-1474.2009.00052.x] [Medline: 19957461]
4. Wolf MS, Davis TC, Shrank W, et al. To err is human: patient misinterpretations of prescription drug label instructions. Patient Educ Couns 2007 Aug;67(3):293-300. [doi: 10.1016/j.pec.2007.03.024] [Medline: 17587533]
5. Sweileh WM, Zyoud SH, Abu Nab'a RJ, et al. Influence of patients' disease knowledge and beliefs about medicines on medication adherence: findings from a cross-sectional survey among patients with type 2 diabetes mellitus in Palestine. BMC Public Health 2014 Jan 30;14:94. [doi: 10.1186/1471-2458-14-94] [Medline: 24479638]
6. Al-Qazaz HK, Sulaiman SA, Hassali MA, et al. Diabetes knowledge, medication adherence and glycemic control among patients with type 2 diabetes. Int J Clin Pharm 2011 Dec;33(6):1028-1035. [doi: 10.1007/s11096-011-9582-2] [Medline: 22083724]
7. Husson O, Mols F, van de Poll-Franse LV. The relation between information provision and health-related quality of life, anxiety and depression among cancer survivors: a systematic review. Ann Oncol 2011 Apr;22(4):761-772. [doi: 10.1093/annonc/mdq413] [Medline: 20870912]
8. Koehler BE, Richter KM, Youngblood L, et al. Reduction of 30-day postdischarge hospital readmission or emergency department (ED) visit rates in high-risk elderly medical patients through delivery of a targeted care bundle. J Hosp Med 2009 Apr;4(4):211-218. [doi: 10.1002/jhm.427] [Medline: 19388074]
9. Jencks SF, Williams MV, Coleman EA. Rehospitalizations among patients in the Medicare fee-for-service program. N Engl J Med 2009 Apr 2;360(14):1418-1428. [doi: 10.1056/NEJMsa0803563] [Medline: 19339721]
10. Menichetti J, Lie HC, Mellblom AV, et al. Tested communication strategies for providing information to patients in medical consultations: A scoping review and quality assessment of the literature. Patient Educ Couns 2021 Aug;104(8):1891-1903. [doi: 10.1016/j.pec.2021.01.019] [Medline: 33516591]

XSL•FO
RenderX

11. Newnham H, Barker A, Ritchie E, Hitchcock K, Gibbs H, Holton S. Discharge communication practices and healthcare provider and patient preferences, satisfaction and comprehension: a systematic review. Int J Qual Health Care 2017 Oct 1;29(6):752-768. [doi: 10.1093/intqhc/mzx121] [Medline: 29025093]
12. Motamedi SM, Posadas-Calleja J, Straus S, et al. The efficacy of computer-enabled discharge communication interventions: a systematic review. BMJ Qual Saf 2011 May;20(5):403-415. [doi: 10.1136/bmjqs.2009.034587] [Medline: 21262793]
13. Arditi C, Rège-Walther M, Wyatt JC, Durieux P, Burnand B. Computer-generated reminders delivered on paper to healthcare professionals; effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2012 Dec 12;12:CD001175. [doi: 10.1002/14651858.CD001175.pub3] [Medline: 23235578]
14. Street RL. Information giving, managing, and understanding in clinical encounters. Patient Educ Couns 2021 Aug;104(8):1831-1833. [doi: 10.1016/j.pec.2021.05.027] [Medline: 34148629]
15. Wong ELY, Coulter A, Cheung AWL, Yam CHK, Yeoh EK, Griffiths SM. Patient experiences with public hospital care: first benchmark survey in Hong Kong. Hong Kong Med J 2012 Oct;18(5):371-380. [Medline: 23018064]
16. Wong EL, Tang KS, Dong D, et al. Evaluation of the implementation of information system for postdischarge with the theoretical domains framework by healthcare professionals: a multistage design with qualitative inquiry and Delphi expert discussion protocol. BMJ Open 2021 Dec 10;11(12):e046081. [doi: 10.1136/bmjopen-2020-046081] [Medline: 34893482]
17. Wong ELY, Tang KS, Cheung AWL, et al. Development of salient medication reminders to facilitate information transfer during transition from inpatient to primary care: the Delphi process. BMJ Open 2021 May 18;11(5):e041336. [doi: 10.1136/bmjopen-2020-041336] [Medline: 34006537]
18. von Elm E, Altman DG, Egger M, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ 2007 Oct 20;335(7624):806-808. [doi: 10.1136/bmj.39335.541782.AD] [Medline: 17947786]
19. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
20. Dehghani H, Mashhadi A. Exploring Iranian English as a foreign language teachers' acceptance of ChatGPT in English language teaching: extending the technology acceptance model. Educ Inf Technol 2024 Oct;29(15):19813-19834. [doi: 10.1007/s10639-024-12660-9]
21. Borges do Nascimento IJ, Abdulazeem H, Vasanthan LT, et al. Barriers and facilitators to utilizing digital health technologies by healthcare professionals. NPJ Digit Med 2023 Sep 18;6(1):161. [doi: 10.1038/s41746-023-00899-4] [Medline: 37723240]
22. Wong ELY, Coulter A, Hewitson P, et al. Patient experience and satisfaction with inpatient service: development of short form survey instrument measuring the core aspect of inpatient experience. PLoS One 2015;10(4):e0122299. [doi: 10.1371/journal.pone.0122299] [Medline: 25860775]
23. Morisky DE, Green LW, Levine DM. Concurrent and predictive validity of a self-reported measure of medication adherence. Med Care 1986 Jan;24(1):67-74. [doi: 10.1097/00005650-198601000-00007] [Medline: 3945130]
24. McConachie SM, Berri D, Konja J, Giuliano CA. Evaluation of adverse drug reaction formatting in drug information mobile phone applications. J Med Libr Assoc 2022 Jan 1;110(1):81-86. [doi: 10.5195/jmla.2022.1251] [Medline: 35210966]
25. Wong ELY, Cheung AWL, Wong AYK, Xu RH, Ramos-Goñi JM, Rivero-Arias O. Normative profile of health-related quality of life for Hong Kong general population using preference-based instrument EQ-5D-5L. V H 2019 Aug;22(8):916-924. [doi: 10.1016/j.jval.2019.02.014] [Medline: 31426933]
26. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
27. Shin DW, Roter DL, Cho J, et al. Attitudes toward disclosure of medication side effects: a nationwide survey of Korean patients, caregivers, and oncologists. Psychooncology 2015 Nov;24(11):1398-1406. [doi: 10.1002/pon.3807] [Medline: 25970704]
28. Curtis LM, Mullen RJ, Russell A, et al. An efficacy trial of an electronic health record-based strategy to inform patients on safe medication use: The role of written and spoken communication. Patient Educ Couns 2016 Sep;99(9):1489-1495. [doi: 10.1016/j.pec.2016.07.004] [Medline: 27444235]
29. McGrath JM. Physicians' perspectives on communicating prescription drug information. Qual Health Res 1999 Nov;9(6):731-745. [doi: 10.1177/104973299129122243] [Medline: 10662256]
30. Tarn DM, Wenger A, Good JS, Hoffing M, Scherger JE, Wenger NS. Do physicians communicate the adverse effects of medications that older patients want to hear? Drugs Ther Perspect 2015 Feb;31(2):68-76. [doi: 10.1007/s40267-014-0176-7] [Medline: 25750513]
31. Howland JS, Baker MG, Poe T. Does patient education cause side effects? A controlled trial. J Fam Pract 1990 Jul;31(1):62-64. [Medline: 2193996]
32. Spranca M, Minsk E, Baron J. Omission and commission in judgment and choice. J Exp Soc Psychol 1991 Jan;27(1):76-105. [doi: 10.1016/0022-1031(91)90011-T]
33. Längst G, Seidling HM, Stützle M, et al. Factors associated with medication information in diabetes care: differences in perceptions between patients and health care professionals. Patient Prefer Adherence 2015;9:1431-1441. [doi: 10.2147/PPA.S88357] [Medline: 26508840]

XSL•FO
RenderX

34. Hamrosi KK, Raynor DK, Aslani P. Pharmacist, general practitioner and consumer use of written medicine information in Australia: are they on the same page? Res Social Adm Pharm 2014;10(4):656-668. [doi: 10.1016/j.sapharm.2013.10.002] [Medline: 24239213]
35. McCormack L, Craig Lefebvre R, Bann C, Taylor O, Rausch P. Consumer understanding, preferences, and responses to different versions of drug safety messages in the United States: a randomized controlled trial. Drug Saf 2016 Feb;39(2):171-184. [doi: 10.1007/s40264-015-0358-9] [Medline: 26547718]
36. Peimani M, Nasli-Esfahani E, Sadeghi R. Patients' perceptions of patient-provider communication and diabetes care: a systematic review of quantitative and qualitative studies. Chron Illn 2020 Mar;16(1):3-22. [doi: 10.1177/1742395318782378] [Medline: 29895167]
37. Loewenthal JV, Beltran CP, Atalay A, Schwartz AW, Ramani S. "What's Going to Happen?": Internal medicine resident experiences of uncertainty in the care of older adults. J Gen Intern Med 2024 Mar 14. [doi: 10.1007/s11606-024-08720-y] [Medline: 38485878]
38. Grime J, Blenkinsopp A, Raynor DK, Pollock K, Knapp P. The role and value of written information for patients about individual medicines: a systematic review. Health Expect 2007 Sep;10(3):286-298. [doi: 10.1111/j.1369-7625.2007.00454.x] [Medline: 17678517]
39. van Geffen ECG, Philbert D, van Boheemen C, van Dijk L, Bos MB, Bouvy ML. Patients' satisfaction with information and experiences with counseling on cardiovascular medication received at the pharmacy. Patient Educ Couns 2011 Jun;83(3):303-309. [doi: 10.1016/j.pec.2011.04.004] [Medline: 21550196]
40. Ferdousi R, Arab-Zozani M, Tahamtan I, Rezaei-Hachesu P, Dehghani M. Attitudes of nurses towards clinical information systems: a systematic review and meta-analysis. Int Nurs Rev 2021 Mar;68(1):59-66. [doi: 10.1111/inr.12603] [Medline: 32608032]
41. Jimma BL, Enyew DB. Barriers to the acceptance of electronic medical records from the perspective of physicians and nurses:A scoping review. Inform Med Unlocked 2022;31:100991. [doi: 10.1016/j.imu.2022.100991]
42. Bauer MS, Damschroder L, Hagedorn H, Smith J, Kilbourne AM. An introduction to implementation science for the non-specialist. BMC Psychol 2015 Sep 16;3(1):32. [doi: 10.1186/s40359-015-0089-9] [Medline: 26376626]
43. Nilsen P. Making sense of implementation theories, models and frameworks. Implement Sci 2015 Apr 21;10:53. [doi: 10.1186/s13012-015-0242-0] [Medline: 25895742]

Abbreviations

CB-SEM: covariance-based structural equation modeling
CFI: comparative fit index
EHR: electronic health record
MGL: 4-item Morisky, Green, Levine
PDIS: postdischarge information summary
RMSEA: root mean square error approximation
SEM: structural equation modeling
SF-HKIEQ: Short-form Hong Kong Inpatient Experience Questionnaire
SRMSR: standard root mean squared residual
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
TAM: technology acceptance model
TLI: Tucker-Lewis index

Edited by X Xiang; submitted 14.05.24; peer-reviewed by T Tang; revised version received 25.11.24; accepted 27.11.24; published 29.01.25.

Please cite as:
Wang DY, Wong ELY, Cheung AWL, Tang KS, Yeoh EK
A Comparison of Patient and Provider Perspectives on an Electronic Health Record–Based Discharge Communication Tool: Survey Study
JMIR Aging 2025;8:e60506
URL: https://aging.jmir.org/2025/1/e60506
doi:10.2196/60506

© Dorothy Yingxuan Wang, Eliza Lai-Yi Wong, Annie Wai-Ling Cheung, Kam-Shing Tang, Eng-Kiong Yeoh. Originally published in JMIR Aging (https://aging.jmir.org), 29.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution,

XSL•FO
RenderX

and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Association of Subjective Cognitive Concerns With Performance on Mobile App–Based Cognitive Assessment in Cognitively Normal Older Adults: Observational Study

Caroline O Nester[1], PhD; Alyssa N De Vito[1], PhD; Sarah Prieto[1], PhD; Zachary J Kunicki[1], PhD; Jennifer Strenger[1,2], MA; Karra D Harrington[3], PhD; Nelson Roque[4], PhD; Martin J Sliwinski[3], PhD; Laura A Rabin[5], PhD; Louisa I Thompson[1], PhD

[1]Department of Psychiatry and Human Behavior, The Warren Alpert Medical School of Brown University, 222 Richmond St, Providence, RI, United States

[2]Memory and Aging Program, Butler Hospital, Providence, RI, United States

[3]Center for Healthy Aging, Penn State University, University Park, PA, United States

[4]Department of Psychology, University of Central Florida, Orlando, FL, United States

[5]Department of Psychology, Brooklyn College CUNY, Brooklyn, NY, United States

Corresponding Author:
Caroline O Nester, PhD
Department of Psychiatry and Human Behavior, The Warren Alpert Medical School of Brown University, 222 Richmond St, Providence, RI, United States

Abstract

Background: Subjective cognitive concerns (SCCs) may be among the earliest clinical symptoms of dementia. There is growing interest in applying a mobile app–based cognitive assessment to remotely screen for cognitive status in preclinical dementia, but the relationship between SCC and relevant mobile assessment metrics is uncertain.

Objective: This study aimed to characterize the relationship between SCC and adherence, satisfaction, and performance on mobile app assessments in cognitively unimpaired older adults.

Methods: Participants (N=122; $Mean_{age}$=68.85 [SD 4.93] years; $Mean_{education}$=16.85 [SD 2.39] years; female: n=82, 66.7%; White:n=106, 86.2%) completed 8 assessment days using Mobile Monitoring of Cognitive Change (M2C2), an app-based testing platform, with brief daily sessions within morning, afternoon, and evening time windows (24 total testing sessions). M2C2 includes digital working memory, processing speed, and episodic memory tasks. Participants provided feedback about their satisfaction and motivation related to M2C2 upon study completion. SCC was assessed using the Cognitive Function Instrument. Regression analyses evaluated the association between SCC and adherence, satisfaction, and performance on M2C2, controlling for age, sex, depression, and loneliness. Linear-mixed effects models evaluated whether SCC predicted M2C2 subtest performance over the 8-day testing period, controlling for covariates.

Results: SCC was not associated with app satisfaction or protocol motivation, but it was significantly associated with lower rates of protocol adherence (ß=−.20, *P*=.37, 95% CI −.65 to −.02). Higher SCC endorsement significantly predicted worse overall episodic memory performance (ß=−.20, *P*=.02, 95% CI −.02 to −.01), but not working memory or processing speed. There was a main effect of SCC on working memory performance at day 1 (estimate=−1.05, SE=0.47, *P*=.03) and a significant interaction between SCC and working memory over the 8-day period (estimate=0.05, SE=0.02, *P*=.03), such that SCC was associated with initially worse, then progressively better working memory performance.

Conclusions: SCCs are associated with worse overall memory performance on mobile app assessments, patterns of cognitive inefficiency (variable working memory), and mildly diminished adherence across an 8-day assessment period. Findings suggest that mobile app assessments may be sensitive to subtle cognitive changes, with important implications for early detection and treatment for individuals at risk for dementia.

(JMIR Aging 2025;8:e64033) doi:10.2196/64033

KEYWORDS

subjective cognitive concerns; subjective cognitive decline; digital cognitive assessment; mobile app; app-based; preclinical Alzheimer disease; mild cognitive impairment; MCI; preclinical dementia; mobile monitoring of cognitive change; Cognitive Function Instrument; mHealth; mobile health; applications; cognition; assessment; remote; geriatrics; gerontology; aging; memory; older adult; elderly; digital health; mobile phone

XSL•FO
RenderX

Introduction

Subjective cognitive decline (SCD) refers to self-perceived changes in cognitive functioning in the setting of normative performance on objective cognitive measures [1]. Older adults with SCD are more likely to demonstrate cognitive decline over time compared with their peers who do not endorse significant cognitive concerns [1-6]. This association has led many to posit that SCD may be one of the earliest clinical manifestations of Alzheimer disease and related dementias (ADRD) [1,3]. SCD has been linked with critical ADRD neural substrates and biomarkers, including structural [7-9] and functional [4,10] alterations, white matter dysfunction [11-13], and the presence of amyloid (Aβ) and tau [14-17]. However, it is important to acknowledge that most individuals notice changes in cognition as they age (up to 80% of people over the age of 70 years), and that many individuals with SCD do not ultimately progress to dementia [2]. As such, investigating novel markers within SCD may enhance early risk detection in this preclinical population.

Disentangling the relationship between SCD and performance on traditional neuropsychological measures is inherently complex. By definition, individuals with SCD perform in the normal range on objective cognitive tasks [3]. However, some studies have demonstrated that SCD is not without mild cognitive deficits. For example, there is evidence of mild reductions in processing speed, executive functioning, language, and memory in SCD [18-22], and such minor deficits have been correlated with concurrent Alzheimer disease (AD) biomarkers and future clinical progression [23]. There is presently a multitude of approaches in the field and widespread debate about how to best assess SCD, and measure selection may directly impact the association of SCD with concurrent cognition and risk for future cognitive decline [24-27]. It is also essential to recognize that SCD intrinsically reflects a longitudinal change over time [28]; however, traditional neuropsychological evaluations only capture a snapshot of in-the-moment cognitive status. Furthemore, for individuals with a high cognitive baseline, normative performance on objective testing may actually represent a decline in cognitive performance [1]. Novel and highly sensitive cognitive tools are needed to characterize more subtle deficits experienced by individuals with SCD [28].

There is growing interest in using smartphone-based, digital technology to remotely screen and track cognitive functioning in older adults at risk for dementia or those along the AD continuum. Smartphone usage among older Americans is highly prevalent (62% - 81% of people over the age of 60 use a smartphone) [29], and a number of smartphone-based cognitive assessment apps have been developed [30,31]. Such approaches are not only more accessible, convenient, time-effective, and scalable [31-33], but also may be more sensitive to subtle cognitive changes not captured by traditional paper and pencil cognitive assessments [31,32]. Smartphone-based cognitive assessments may also demonstrate superior ecological validity when completed in the individual's lived context with a familiar piece of technology [30,31,34]. Smartphone technology also facilitates high-frequency, repeated cognitive assessment, which allows for the quantification of subtle markers of decline in older adults, such as intraindividual variability and patterns in cognition over time [30-32,35,36]. However, the self-administered, unsupervised nature of smartphone-based cognitive assessments raises concerns related to task adherence, engagement, and potentially confounding factors that may occur within an uncontrolled environment [32].

Particularly relevant to aging research is the limited understanding of how subjective cognitive concerns (SCCs), the core feature of SCD, may be associated with engagement, satisfaction, and performance on such smartphone-based cognitive assessments. Preliminary evidence suggests that smartphone use may be associated with fewer SCCs in older adulthood and possibly constitute a protective factor against cognitive decline [37]; hence, smartphones may represent a promising tool for novel assessments and potential interventions for at-risk populations with self-perceived cognitive decline. Emerging literature suggests that smartphone-based cognitive testing generally has adequate adherence rates (>70%) [32], but little is understood about factors (such as SCC) which may impact engagement or motivation on unsupervised digital tests [38]. These are especially important relationships to disentangle, given the ubiquity of SCC in aging populations [2], high rates of smartphone usage [29], and growing interest in remote assessment to enhance accessibility to cognitive screening. In addition, the relationships between digital cognitive testing and AD biomarkers [39-41], as well as SCC and AD biomarkers [15,17,42], is supported by emerging research, but no studies to date have directly investigated the relationship between SCC, AD biomarkers, and digital smartphone cognitive tasks.

In light of the need to better understand the role of SCC in smartphone-based cognitive assessments in aging populations at risk for AD, the present analysis sought to characterize the relationship between SCC and adherence (aim 1), satisfaction and motivation (aim 2), and performance (aim 3) on a mobile app–based digital cognitive assessment in a sample of cognitively unimpaired older adults. Given the established association between SCC, cerebral Aβ positivity (eg, the hallmark biomarker of AD), and risk for cognitive decline [14-16,23], a secondary aim sought to understand if Aβ positivity moderated the relationship between SCC and performance on the mobile app–based cognitive assessment. Results represent the first step toward quantifying the impact of SCC on smartphone-based cognitive testing in older adult populations, with crucial implications for early detection in those at risk for cognitive decline and dementia.

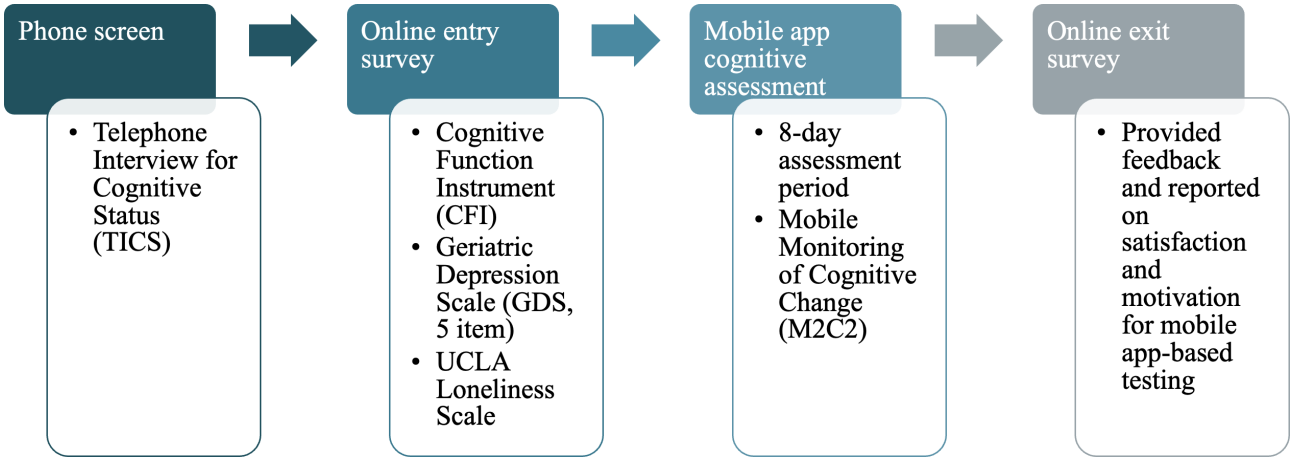
Methods

Participants, Recruitment, and Procedures

Participants consisted of 122 cognitively unimpaired older adults, between the ages of 60 and 80 years of age, recruited from the Butler Alzheimer's Prevention Registry, a local database of older adults interested in AD research at the Butler Hospital Memory and Aging Program (MAP) [43]. All study procedures were carried out remotely, and detailed recruitment and procedural information has been described previously [38,41,44]. Study procedures are presented in Figure 1.

XSL•FO
RenderX

Figure 1. Study procedures. UCLA: University of California, Los Angeles.



We used a targeted recruitment to enroll individuals with previous amyloid positron emission tomography (PET) data (elevated [Aß+] or nonelevated [Aß-] determined by clinical read by a radiologist) (final n=73) [41] as well as those without PET data. PET results were available if recorded in our Registry as part of previous research participation at the MAP. Individuals (n=256) were invited to participate in the study through an email or phone call. Individuals who consented (n=146) completed an online survey before the initiation of the remote testing protocol, which included various screening measures, including the Cognitive Functioning Index (CFI) [45] (subjective cognitive concern screen), Geriatric Depression Scale 5-item version (GDS-5) [46] (depressive symptoms screen; score range 0 - 5, with higher scores indicating worse depressive symptoms), and the University of California, Los Angeles (UCLA) 3-Item Loneliness Scale [47] (loneliness screen; score range 0 - 6, with higher scores indicating worse loneliness). The modified Telephone Interview for Cognitive Status (TICSm) [48] (objective cognition screen; score range 0 - 50, with lower scores indicating worse cognitive functioning) was also completed before the remote testing protocol by phone, administered by a research assistant. During screening, 23 participants were excluded, and 1 participant was lost to follow up. After the completion of the 8 day remote study protocol, participants (final n=122) were invited to provide feedback (including study satisfaction and motivation) on their experience through an online survey, and a US $20 gift card was provided for compensation.

Inclusion criteria involved unimpaired cognition, daily use of a smartphone, and familiarity with smartphone features. Unimpaired cognition was defined as a TICSm cutoff score of ≥34 [48]. Exclusion criteria included self-reported history of cognitive impairment or dementia, history of neurologic disease or severe mental illness, and physical inability to complete smartphone-based testing.

Subjective Cognitive Concerns

The CFI is a 14-item self-report measure which probes changes in cognitive and functional domains over the past year. The CFI was selected for inclusion in this study as it is a widely used, briefly administered measure of SCC that has been well-validated for used in cognitively healthy and preclinical dementia populations [42,45,49,50]. Items reflect memory concerns, increased use of compensatory strategies, changes in driving, or difficulty managing instrumental activities of daily living. Participants respond with yes (1 point), no (0 points), or maybe (0.5 points), with total scores ranging from 0 to 14 (continuous variable). Higher scores reflect greater SCC. As a sensitivity analysis, we also created a SCC variable that included CFI items that solely assessed cognition (CFI-Cog), by removing the functioning related items to create a refined examination of SCC. CFI-Cog included CFI items 1, 2, 3, 4, 5, 6, 8, 11, and 13. This approach is consistent with a recent study which identified that cognitively focused CFI items, rather than functionally related CFI items, were endorsed at a higher rate and more likely to be associated with Aß+ in cognitively unimpaired older adults [42].

Mobile App–Based Cognitive Assessment

Overview

Remote cognitive tasks were completed using the Mobile Monitoring of Cognitive Change (M2C2) app [51] (more details in Figure 2), a cognitive testing platform developed as part of the National Institute of Aging Mobile Toolbox initiative, with strong previous theoretical, empirical, and psychometric support, including evidence of sensitivity to age and age-related neuropathology [34,38,44]. Android smartphones preloaded with the cognitive assessment app were mailed to participants along with a detailed use guide. Phone functions were locked to prevent the use of features such as web browsing and the camera. Participants completed 3, brief (ie, 3 - 4 minutes) M2C2 sessions each day during morning, afternoon, and evening time windows for 8 consecutive days. Additional sessions could be completed on day 9 as optional or make-up sessions. Staff provided support through phone or email as needed, as described previously [41]. During each M2C2 session, participants completed 3 previously characterized cognitive measures assessing visual working memory (Color Shapes), processing speed (Symbol Match), and episodic memory (Prices). Each task took approximately 60 seconds to complete.

XSL•FO
RenderX

Figure 2. Mobile Monitoring of Cognitive Change (M2C2) app subtests.

XSL•FO
RenderX

Color Shapes

The Color Shapes visual working memory task is a visual array change detection test measuring intraitem feature binding. Participants determine if shapes change color across 2 sequential presentations in which the shape locations change. Performance was measured by discriminability (d-prime) performance calculated from the proportion of correct identifications and proportion of misidentified stimuli [51].

Symbol Match

The Symbol Match processing speed task is a speeded continuous performance test of conjunctive feature search. Participants are asked to identify matching symbol pairs. Performance was measured by the median reaction time to complete the task across all trials (in milliseconds) [51].

Prices

The Prices episodic memory task uses an immediate delayed forced-choice recognition paradigm. Participants incidentally encode 10 grocery item-price pairs for later recall while judging whether or not the item's price is "good." Recall trials begin immediately after the learning trials. Performance was measured by the proportion of correct responses on the 10 recall trials [51].

Study Adherence, Satisfaction, and Motivation

Adherence was quantified as the number of testing sessions completed across the 8-day ambulatory protocol (out of 24 sessions).

Participants' feedback covered various aspects, including their satisfaction with the protocols and their level of motivation to engage with the study tasks. Feedback related to satisfaction was measured with 2 items: "Completing the brain games was fun" and "I became bored with the brain games." Participants rated their responses on a 6-point Likert scale (1=strongly disagree to 6=strongly agree). For the purposes of this study, responses were dichotomized by those who agreed (Likert scale responses 4, 5, and 6) or those who disagreed (Likert scale responses 1, 2, and 3) on these items due to a largely positive response bias (most individuals found the tasks to be fun and not boring).

Feedback regarding motivation was measured on the item: "It was hard to motivate myself to complete the games each time." Participants rated their response on a 5-point Likert scale (1=Agree to 5=Disagree). As above, responses were dichotomized into those who agreed or were neutral (Likert scale responses 1, 2, and 3) and those who disagreed (Likert scale responses 4 and 5) for the purpose of this analysis due to a largely positive response bias (most individuals found themselves to be easily motivated to complete the tasks).

Statistical Analysis

Descriptive statistics were conducted to characterize the sample in terms of demographics (self-reported biological sex, race, ethnicity, and years of education), self-reported symptoms of depression (by the Geriatric Depression Scale 5-item version [46]), loneliness (by the UCLA Loneliness Scale [47]), and SCC (CFI total and CFI-Cog), relevant app-based metrics (ie, adherence, satisfaction and motivation), and objective cognitive performance (ie, visual working memory, processing speed, and episodic memory). Pearson correlation coefficient (for continuous variables) and point biserial correlation (for binary variables) were implemented as appropriate to assess the relationship between SCC and demographic variables, depression, loneliness, and outcomes of interest (eg, adherence, satisfaction, motivation, and performance on app-based cognitive assessment). To address aims 1 - 3 and secondary analyses, SCC (continuous variable) was the main predictor of interest and we adjusted for covariates, including age, sex, depression, and loneliness, in all models. Covariates were selected based on significant association with independent and dependent variables of interest. A priori significance threshold (α) was set at $P \leq .05$. For our moderation analyses, we examined α levels up to $P \leq .10$.

To test aim 1, linear regression models were used to evaluate the association between SCC and adherence (continuous variable) over the 8-day protocol. For the purposes of this analysis, the CFI total score was the main SCC predictor of interest. We present results for CFI-Cog only when they differ from the CFI total score. To test aim 2, logistic regression models were conducted to evaluate the association between SCC and satisfaction (dichotomized self-report items related to level of fun and boredom), and SCC and motivation (dichotomized self-report items related to motivation). Individual regression models were implemented for each motivation and satisfaction outcome of interest. To test aim 3, separate linear regression models were constructed to evaluate the association between SCC and performance on each M2C2 based cognitive assessment (Prices, Color Shapes, and Symbol Search, all continuous variables). Linear-mixed effects models were used to evaluate whether SCC predicted M2C2 subtest performance over time while controlling for age, sex, and loneliness. Finally, as a secondary analysis in a subsample of participants with Aβ PET data, we fit linear regression models with SCC and the interaction of SCC and Aβ PET positivity as predictors, to test the moderation effect of Aβ PET positivity on the association between SCC and study adherence and performance on M2C2 tasks, adjusting for covariates. All descriptive and regression analyses were conducted using SPSS version 28.0.1.1 (IBM Corp). Linear mixed-effects models were conducted using the *nlme* R package.

Ethical Considerations

The project received approval from the Butler Hospital Institutional Review Board (#1882523), and all the participants provided informed consent. All data was de-identified. Participants were compensated for their time with a $20 gift card.

Results

Participant Characteristics and Correlates of Subjective Cognitive Concerns

Overall sample demographics and SCC, adherence, satisfaction, motivation, and performance variables are presented in Table 1. The sample (N=122) was 68.85 (SD 4.93) years old

XSL•FO
RenderX

(range=60 - 81 y) with 16.52 (SD 2.39) years of formal education. The majority of the sample identified as female (n=82, 66.7%), White (n=106, 86.2%), and non-Hispanic (n=109, 88.6%). In the subset of individuals with Aβ PET scans, there were 25 Aβ positive participants and 48 Aβ negative participants. Subjective cognitive concerns, as measured on the CFI, were fairly minimal in this sample. On average, the sample scored 1.8 (SD 1.7) out of 14 possible points on the overall CFI ; however, the most common score on the CFI was 0 (mode=0), with 18% of the sample expressing no concerns about their cognition on the CFI.

Table . Sample demographic, subjective cognitive concerns, and mobile app protocol metrics.

	Total sample (N=122)	Subsample with Aβ[a] PET[b] (n=73)
Demographic characteristics		
Age (years), mean (SD)	68.85 (4.93)	69.25 (4.48)
Education (years), mean (SD)	16.52 (2.39)	16.45 (2.60)
Female, n (%)	82 (66.7)	52 (71.2)
White, n (%)	106 (86.2)	65 (89)
Non-Hispanic, n (%)	109 (88.6)	64 (87.7)
Depression (GDS[c]), mean (SD)	0.27 (0.54)	0.24 (0.46)
Loneliness (UCLA[d]), mean (SD)	0.70 (1.15)	0.56 (0.98)
Cognitive status (TICSm[e]), mean (SD)	39.32 (3.33)	39.01 (3.03)
Subjective cognitive concerns		
CFI[f] total score, mean (SD)	1.86 (1.70)	1.96 (1.80)
CFI-Cog[g], mean (SD)	1.39 (1.35)	1.96 (1.80)
Protocol adherence, mean (SD)	22.45 (2.58)	21.86 (2.30)
Protocol satisfaction		
Who had fun, n (%)	102 (85)	59 (83.1)
Who became bored, n (%)	30 (24.8)	17 (23.6)
Protocol motivation		
Who were motivated, n (%)	93 (77.5)	53 (73.6)
M2C2[h] task performance (average over 8-day protocol)		
Prices, mean (SD)	0.75 (0.09)	0.75 (0.09)
Color Shapes, mean (SD)	2.45 (0.76)	2.46 (0.80)
Symbol Search, mean (SD)	2224.78 (447.76)	2215.08 (410.92)

[a]Aβ: amyloid.

[b]PET: positron emission tomography.

[c]GDS: Geriatric Depression Scale.

[d]UCLA: University of California, Los Angeles.

[e]TICSm: modified Telephone Interview for Cognitive Status.

[f]CFI: Cognitive Function Instrument.

[g]CFI-Cog: Cognitive Function Instrument cognitive items.

[h]M2C2: Mobile Monitoring of Cognitive Change.

Bivariate associations are presented in Table 2. The CFI did show a small to moderate correlation with worse cognitive status (TICSm, r=−0.202, P=.03), greater loneliness (UCLA Loneliness Scale, r=0.248, P=.006), and worse adherence (r=−0.207, P=.02). In terms of overall M2C2 performances, the CFI showed a small correlation with episodic memory (ie, Prices) performance (r=−0.196, P=.03).

XSL•FO
RenderX

Table . Bivariate associations among demographic variables, subjective cognitive concerns, mobile app protocol metrics, and amyloid (Aβ) positron emission tomography (PET)[a].

		Age	Sex	Education	Depression	Loneliness	TICSm[b]	CFI[c]	CFI-Cog[d]	Adherence	Prices	Color Shape	Symbol Search	Aβ positive
Age														
	r	—[e]	−0.143	−0.005	−0.154	−0.024	−0.262	0.001	−0.020	−0.060	−0.369	−0.308	0.418	0.258
	P value	—	.12	.95	.09	.79	.004	.99	.83	.51	<.001	<.001	<.001	.03
Sex														
	r	−0.143	—	−0.066	−0.038	0.146	0.184	0.001	−0.050	0.031	0.221	0.271	−0.093	0.012
	P value	.12	—	.47	.68	.11	.04	.99	.60	.73	.01	.003	.31	.92
Education														
	r	−0.005	−0.066	—	0.011	−0.085	0.218	0.074	0.083	0.036	−0.028	0.020	−0.140	−0.015
	P value	.95	.47	—	.91	.36	.02	.42	.37	.70	.76	.83	.13	.90
Depression														
	r	−0.154	−0.038	0.011	—	0.170	0.063	0.040	0.064	0.054	0.202	−0.133	−0.028	0.052
	P value	.09	.68	.91	—	.06	.50	.67	.49	.56	.03	.15	.76	.66
Loneliness														
	r	−0.024	0.146	−0.085	0.170	—	−0.056	0.248	0.104	−0.094	0.001	0.033	−0.069	−0.031
	P value	.79	.11	.36	.06	—	.55	.006	.26	.31	.99	.72	.45	.80
TICSm														
	r	−0.262	0.184	0.218	0.063	−0.056	—	−0.202	−0.183	0.008	0.350	0.225	−0.216	−0.175
	P value	.004	.04	.02	.50	.55	—	.03	.05	.94	<.001	.01	.02	.14
CFI														
	r	0.001	0.001	0.074	0.040	0.248	−0.202	—	0.941	−0.176	−0.196	−0.028	−0.023	−0.002
	P value	.99	.99	.42	.67	.006	.03	—	<.001	.05	.03	.76	.80	.98
CFI-Cog														
	r	−0.020	−0.050	0.083	0.064	0.104	−0.183	0.941	—	−0.207	−0.171	−0.017	−0.038	−0.041
	P value	.83	.60	.37	.49	.26	.05	<.001	—	.02	.06	.86	.68	.73
Adherence														
	r	−0.060	0.031	0.036	0.054	−0.094	0.008	−0.176	−0.207	—	0.221	0.043	−0.075	0.210
	P value	.51	.73	.70	.56	.31	.94	.05	.02	—	.02	.64	.42	.08
Prices														
	r	−0.369	0.221	−0.028	0.202	0.001	0.350	−0.196	−0.171	0.221	—	0.309	−0.139	−0.325
	P value	<.001	.01	.76	.03	.99	<.001	.03	.06	.02	—	<.001	.13	.005
Color Shapes														
	r	−0.308	0.271	0.020	−0.133	0.033	0.225	−0.028	−0.017	0.043	0.309	—	−0.287	−0.141
	P value	<.001	.003	.83	.15	.72	.01	.76	.86	.64	<.001	—	.001	.24
Symbol Search														
	r	0.418	−0.093	−0.140	−0.028	−0.069	−0.216	−0.023	−0.038	−0.075	−0.139	−0.287	—	0.011
	P value	<.001	.31	.13	.76	.45	.02	.80	.68	.42	.13	.001	—	.93
Aβ positive														
	r	0.258	0.012	−0.015	0.052	−0.031	−0.175	−0.002	−0.041	0.210	−0.325	−0.141	0.011	—
	P value	.03	.92	.90	.66	.80	.14	.98	.73	.08	.005	.24	.93	—

[a]The Aβ PET data was only available for a subsample (n=73).

XSL•FO
RenderX

[b]TICSm: modified Telephone Interview for Cognitive Status

[c]CFI: Cognitive Function Instrument

[d]CFI-Cog: Cognitive Function Instrument cognitive items

[e]Not applicable

Aim 1: Subjective Cognitive Concerns and Adherence to Mobile App–Based Cognitive Assessment Protocol

Overall, remote assessment protocol adherence was high in the sample, with participants completing on average 93.5% of the 24 testing sessions over the 8-day period. CFI-Cog endorsement significantly predicted overall remote testing adherence over 8 days. This suggests that higher levels of SCC were associated with worse adherence (a diminished number of test sessions completed across the protocol) (ß=−.197, *P*=.04, 95% CI −.647 to −.021). However, this association was not observed when using the CFI total score as the predictor (Table 3).

XSL•FO
RenderX

Table . Association between subjective cognitive concerns and mobile app metrics.

	R^2	ß/aOR[a] (95% CI)	*P* value
Aim 1: Subjective cognitive concerns and adherence to app-based cognitive assessment protocol, ß			
Adherence	0.042		
SCC[b] (CFI[c])		−.157 (−.459 to .042)	.10
Age		−.046 (−.106 to .064)	.63
Sex		.051 (−.657 to 1.150)	.59
Depression (GDS[d])		.072 (−.479 to 1.068)	.45
Loneliness (UCLA[e])		−.071 (−.5.20 to .239)	.47
Aim 2: Subjective cognitive concerns, satisfaction, and motivation on app-based cognitive assessment, aOR			
Protocol satisfaction			
Level of fun	0.060		
SCC (CFI)		1.016 (0.738 to 1.398)	.92
Age		0.938 (0.844 to 1.042)	.23
Sex		0.328 (0.112 to 0.963)	.04
Depression (GDS)		0.679 (0.28 to 1.606)	.38
Loneliness (UCLA)		0.883 (0.554 to 1.408)	.60
Boredom	0.026		
SCC (CFI)		1.013 (0.783 to 1.310)	.92
Age		1.056 (0.966 to 1.155)	.23
Sex		0.575 (0.213 to 1.552)	.27
Depression (GDS)		1.060 (0.473 to 2.375)	.89
Loneliness (UCLA)		1.130 (0.785 to 1.625)	.12
Protocol motivation			
Motivation	0.028		
SCC (CFI)		0.891 (0.691 to 1.148)	.37
Age		0.962 (0.877 to 1.054)	.41
Sex		1.689 (0.613 to 4.603)	.31
Depression (GDS)		0.685 (0.312 to 1.503)	.35
Loneliness (UCLA)		1.286 (0.831 to 1.991)	.26
Aim 3: Subjective cognition and performance on app based cognitive assessment, ß			
Prices	0.234		
SCC (CFI)		−.200 (−.020 to −.002)	.02
Age		−.318 (−.009 to −.003)	<.001
Sex		.193 (.005 to .070)	.02
Depression (GDS)		.174 (.001 to .056)	.04
Loneliness (UCLA)		−.013 (−.015 to .013)	.88
Color Shapes	0.170		
SCC (CFI)		−.016 (−.085 to .071)	.86

XSL•FO
RenderX

	R^2	ß/aOR[a] (95% CI)	*P* value
Age		−.301 (−.073 to −.020)	<.001
Sex		.205 (.051 to .612)	.02
Depression (GDS)		−.181 (−.489 to −.009)	.04
Loneliness (UCLA)		.030 (−.098 to .138)	.74
Symbol Search	0.178		
SCC (CFI)		−.007 (−48.34 to 44.72)	.94
Age		.416 (22.302 to 53.977)	<.001
Sex		−.027 (−192.599 to 140.950)	.76
Depression (GDS)		.041 (−109.610 to 176.575)	.64
Loneliness (UCLA)		−.045 −92.465 to 55.043	.62

[a]aOR: adjusted odds ratio.

[b]SCC: subjective cognitive concern.

[c]CFI: Cognitive Function Instrument.

[d]GDS: Geriatric Depression Scale.

[e]UCLA: University of California, Los Angeles.

Aim 2: Subjective Cognitive Concerns, Satisfaction, and Motivation on Mobile App–Based Cognitive Assessment

In terms of overall protocol satisfaction, participants endorsed that completing the assessments was fun (n=102, 85%) and not boring (n=92, 75.2%). Participants also reported that they did not encounter difficulties in motivating themselves to give their best performance (n=93, 77.5%). CFI total score was not associated with metrics of study satisfaction in the adjusted models, including how fun or boring participants found the protocol to be. In terms of motivation, CFI total score was not predictive of self-reported effort across the protocol period (more details in Table 3).

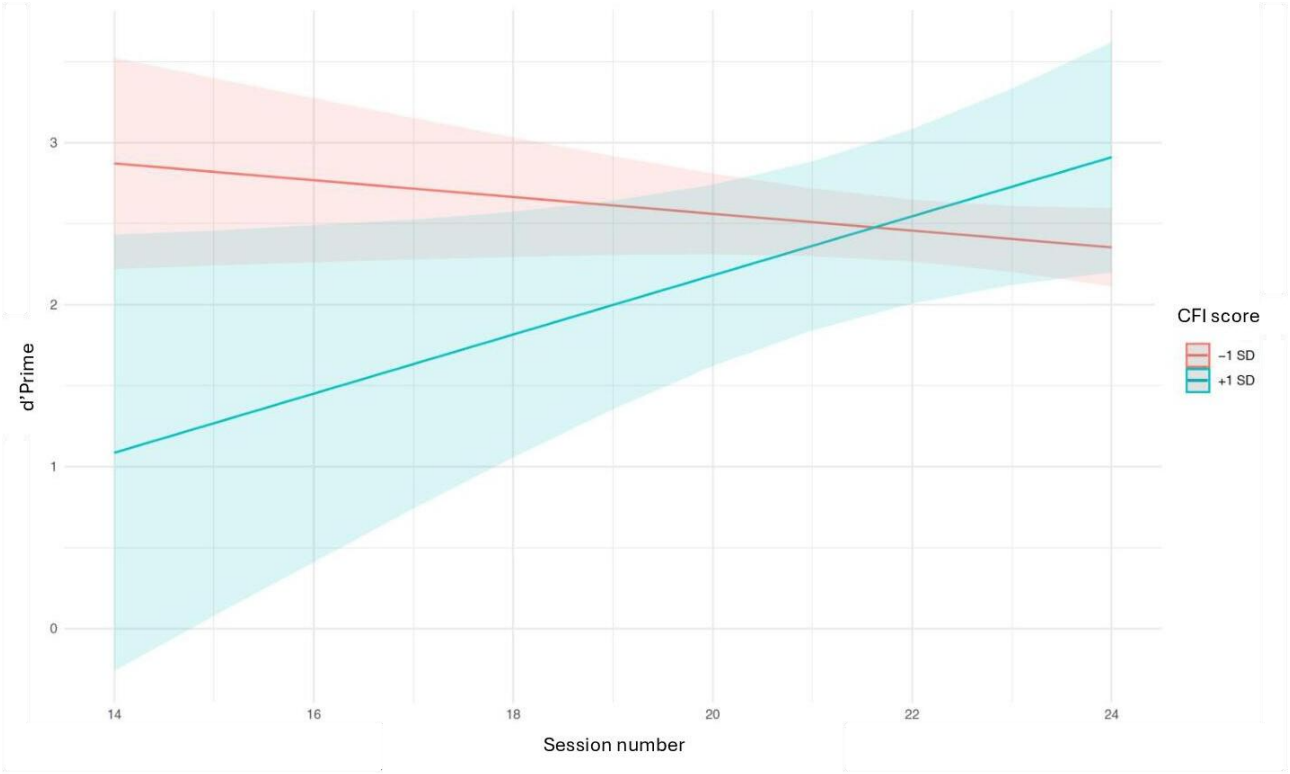
Aim 3: Subjective Cognitive Concerns and Performance on Mobile App–Based Cognitive Assessment

SCCs and their association with overall objective cognitive performances on digital cognitive tests was also investigated (more details in Table 3). High SCC endorsement on the CFI total score was significantly predictive of poorer overall performance on the Prices task (ß=−.200, *P*=.02, 95% CI −.020, to −.002). There were no significant associations between the CFI and performance metrics for Color Shapes (*P*=.86) or Symbol Search (*P*=.94).

We also investigated the association of SCC and longitudinal trends of cognitive performance over the 8-day testing period (M2C2 app performance descriptives for each study day have been previously reported) [38]. There was a main effect of the CFI total score suggesting that higher levels of SCC were associated with worse performance on Color Shapes (working memory) at day 1 (unstandardized estimate [b]=−1.047, SE=0.47, *P*=.03). Furthermore, there was a significant interaction between SCCs and Color Shapes performance (b=0.048, SE=0.02, *P*=.03), such that higher CFI total scores were associated with improved performance on this subtest over time. In other words, although high SCC was associated with worse working memory performance initially, individuals seemed to benefit from repeated practice and ultimately performed better (more details in Figure 3). There were no significant associations between SCC and intraindividual variability in Prices or Symbol Match performances over the protocol period.

XSL•FO
RenderX

Figure 3. Subjective cognitive concerns association with working memory longitudinally over the 8-day mobile-appmobile app protocoll. CFI: Cognitive Function Instrument.



Secondary Aim: Moderating Effect of Amyloid Status on the Relationship Between Subjective Cognitive Concerns and Mobile App–Based Cognitive Assessment

The moderation of Aβ positivity on the relationship between CFI and app-based metrics, as well as CFI and cognitive performance (ie, Prices, Color Shapes, and Symbol Search) was investigated in a subset of individuals with Aβ PET scans (n=73). The presence of Aβ did not moderate the relationship between CFI total score and adherence. There was evidence for moderation at the $P<.10$ level of Aβ positivity on the relationship between CFI total score and Prices was observed, but did not fall within the $\alpha=.05$ threshold (ß=.179, $P=.09$, 95% CI −.003 to .036). Aβ status did not moderate the association between SCC and Color Shapes or Symbol Search.

Discussion

Principal Findings

This study sought to characterize the relationship between SCC and mobile app–based protocol engagement (adherence, satisfaction, and motivation) and performance on cognitive tasks in a cohort of cognitively healthy older adults. While protocol satisfaction and motivation were not impacted by SCCs, overall adherence was associated, such that higher levels of SCCs predicted lower rates of adherence. Overall, we show that SCCs were associated with worse objective performance on mobile app–based assessments. Our findings suggest that mobile app–based assessments hold promise as sensitive tools to detect subtle neurodegenerative changes in individuals at risk for dementia, with potential application for early detection, diagnosis, and treatment.

We observed an association between SCCs, as measured on the Cognitive Function Index (CFI), and overall worse performance on a test of episodic memory (Prices). A significant association was not observed between SCC and overall performance on tasks of processing speed or working memory. Findings in the broader SCD literature are quite mixed as to whether SCCs are or are not associated with concurrent cognitive functioning abilities [52]; as such, our results may indicate that the use of sensitive digital cognitive tools could enhance the association between SCC and objective cognitive metrics. Our result reflecting a unique association with episodic memory is perhaps clinically meaningful, as memory concerns specifically are frequently considered to be the hallmark presenting symptom of AD [2,53], particularly in individuals over the age of 65 [53]. This observation is underscored by our finding of a moderating effect of cerebral Aβ positivity (at the $P<.10$ level) on the relationship between SCC and episodic memory (Prices) in a subsample of participants with biomarkers available for analysis (secondary aim), such that the association of SCC with episodic memory was stronger in the presence of Aβ. This result adds to the growing evidence base for an association between SCC and higher Aβ [52]. Replication of this analysis in a larger sample may clarify the interplay between SCC, AD biomarkers, and subtle memory deficits observed on mobile app–based assessment in otherwise cognitively and clinically normal older adults. That being said, it is worth considering that the unique association between SCC and episodic memory in our sample may be at least somewhat secondary to the way the CFI is constructed, as this instrument mostly focuses on episodic memory-related cognitive concerns [45]. Previous studies have

XSL•FO
RenderX

demonstrated the importance of querying beyond just memory concerns to most sensitively detect dementia risk [54,55]. Indeed, SCC can be sampled by a multitude of techniques and approaches (eg, memory only vs multiple cognitive domains; traditional paper and pencil vs digital assessments; self vs informant report; capturing current ability vs change in ability, etc) [24,25], all of which may impact the sensitivity of the SCC measures to detect current and risk for future cognitive impairment. Future research should carefully weigh SCC assessment approaches, and consider incorporating a more comprehensive SCC screener which queries across a broader range of cognitive domains, as this may be valuable for detecting more meaningful associations between SCC and mobile app–based cognitive performances.

We also evaluated patterns of cognitive performance across the 8-day assessment period. Analyses revealed that more SCC was associated with initially worse working memory (Color Shapes) performance on day one. However, individuals with higher SCC showed progressively improving working memory performance over the assessment period, and by day 8 were performing in a similar range with those with lower SCC. This pattern was not observed on the episodic memory or processing speed subtests. This finding is consistent with recent results from Aschenbrenner et al [56], which showed variability over high frequency app-based assessments in attention and working memory-based skills, but not in episodic memory, in older adults who were at increased genetic risk for AD [56]. Our finding of an initial relative weakness in working memory skills could reflect the presence of mild cognitive inefficiencies in older adults with SCC, which may be compensated for with repeated exposure to the task over time. This may be due to early dysfunction in working memory processes, which have been characterized in preclinical AD [57,58]. Subjectively, this may be experienced by these individuals as thinking which requires increased concentration, that is punctuated by occasional lapses, and is associated perhaps with a sense of being cognitively overwhelmed, all of which may be driving the report SCC in daily life, but may be subtle enough that it may not be detected on traditional neuropsychological tests. Hence, this type of unique marker of cognitive inefficiency would only be afforded through high frequency, mobile app–based cognitive assessment in this population. The plausible role of performance anxiety may also explain variability in cognitive performance overtime. Although beyond the scope of the present analysis, research which comprehensively measures the role of in-the-moment psychological and physiological states by ecological momentary assessment techniques in relation to in the moment SCC and objective cognition will be critical next steps.

Our findings related to SCC and metrics of protocol engagement and satisfaction with unsupervised, high frequency digital cognitive assessments have potentially broad implications for aging research, especially in light of increasing interest in using these techniques to remotely assess cognitive functioning in this population [31,32]. Indeed, whether or not a study is directly examining SCC or includes SCD as a clinical group of interest, aging studies must contend with the fact that SCCs are widespread and may be clinically meaningful feature among the older adults in their sample [2]. We show that SCC was not associated with protocol satisfaction (including how fun or how boring the experience was) nor did SCC effect how motivated a participant was to engage and put forth their best effort. However, higher levels of SCC did impact protocol adherence in our sample, with mildly worse compliance those with higher SCC observed. For context, protocol adherence was overall quite high in our population (93.5%) and similarly high compliance rates have been described in several other recent smartphone-based studies [32], yet no previous studies have directly explored the impact of SCCs on protocol adherence. It may be that high frequency, mobile assessments are able to detect mild patterns of forgetfulness (manifested in slightly worse adherence rates) in older adults with SCC, so future research should be attuned to this possibility. Replication of this association between SCC and adherence to remote protocols in larger, more demographically diverse samples is warranted.

Limitations

Some limitations of this study, as well as potential future directions of this research, warrant acknowledgment. Our sample was largely White, disproportionally female, and highly educated. This limits the generalizability of the study findings to broader demographic groups. Our sample was from a relatively small single cohort and only a subset had AD biomarkers available for analysis, restricting study power and may have impeded our ability to obtain significant findings. A measure of anxiety symptomology was not available in the present sample, limiting our ability to understand the plausible role of anxiety, SCC, and performance on digital assessments [32,59]. The main predictor of interest in this study, the CFI, shows adequate evidence for validity and reliability [42], but is somewhat limited in the domains of cognitive concerns it measures (eg, memory and daily functioning). Research has shown that comprehensive assessment of concerns across a broad range of cognitive domains may be most sensitive to risk in aging populations [60], so future smartphone-based research may consider including a broader range of cognitive concern items to better understand this relationship. We were unable to assess informant-report of SCC in this sample, which has been demonstrated to be optimized to predict risk above and beyond self-report [25], and this should be addressed in future smartphone studies. Although such metrics were not available in this study, there is also a small but growing literature which has examined the use of smartphones to assess SCC in a digital remote format in tandem with cognitive tests and relevant clinical outcomes in older adult samples [61-63]. Future research on SCC and smartphone-based digital cognitive assessments should investigate the relative value of traditional in person reported SCC (eg, the CFI) versus smartphone based digital SCC assessments. Finally, this cross-sectional study was not able to investigate the relationship between SCC and longitudinal decline on smartphone-based cognitive performance in individuals with positive AD biomarkers. It will be essential for future studies to follow participants over time to understand if SCC at baseline may predict incident decline on these novel cognitive metrics. Due to the limited nature of this pilot study, we cannot provide specific recommendations about optimal SCC assessment approach for use in future digital cognitive studies.

XSL•FO
RenderX

Conclusions

SCCs are frequently one of the earliest clinical symptoms of dementia [2]. However, clinical interpretations of SCCs are complicated by the ubiquitous endorsement of these concerns [2], their potential for non-neurodegenerative etiology (eg, psychiatric, medical, or sociodemographic factors) [2,3], and frequent lack of association with traditional objective cognitive testing [1]. Smartphone-based, digital cognitive assessments, which are increasingly used in aging research, may offer potential for improved ecological validity and sensitivity to subtle markers of cognitive decline [30-32,35]. Hence, such smartphone-based digital tools may be able to capture quite mild deficits in individuals with SCC who perform within normal limits on traditional neuropsychological tests. Results from the current study showed that SCCs are associated with worse overall memory performance on mobile app assessment, and patterns of cognitive inefficiency (variable working memory) and mild forgetfulness (diminished adherence) across an 8-day assessment period in cognitively intact older adults. Findings indicate that mobile app assessments may be uniquely sensitive to very subtle neurodegenerative changes in at risk older adults, with critical implications for early detection and timely intervention.

Acknowledgments

This work is supported by Alzheimer's Association grant AACSF-20-685786 (Thompson, PI) and by NIA grants T32 AG049676 and U2CAG060408 to Pennsylvania State University.

Authors' Contributions

CON took the lead on study conceptualization, formal analysis, original draft writing, and visualization. ANDV and ZJK contributed to formal analyses, visualization, and draft writing and editing. SP played an instrumental role in draft writing and editing. JS, KDH, NR, MJS, and LAR provided valuable feedback on study conceptualization and draft review and editing. LIT was responsible for funding acquisition, project administration, study conceptualization, and draft writing, review, and editing.

Conflicts of Interest

KDH is an employee of Cogstate Ltd. All remaining authors declare no conflicts of interest.

References

1. Jessen F, Amariglio RE, van Boxtel M, et al. A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. Alz Dement 2014 Nov;10(6):844-852. [doi: 10.1016/j.jalz.2014.01.001] [Medline: 24798886]
2. Jessen F, Amariglio RE, Buckley RF, et al. The characterisation of subjective cognitive decline. Lancet Neurol 2020 Mar;19(3):271-278. [doi: 10.1016/S1474-4422(19)30368-0] [Medline: 31958406]
3. Rabin LA, Smart CM, Amariglio RE. Subjective cognitive decline in preclinical Alzheimer's disease. Annu Rev Clin Psychol 2017 May 8;13:369-396. [doi: 10.1146/annurev-clinpsy-032816-045136] [Medline: 28482688]
4. Wang XT, Wang ZT, Hu HY, et al. Association of subjective cognitive decline with risk of cognitive impairment and dementia: a systematic review and meta-analysis of prospective longitudinal studies. J Prev Alzheimers Dis 2021;8(3):277-285. [doi: 10.14283/jpad.2021.27] [Medline: 34101784]
5. Parfenov VA, Zakharov VV, Kabaeva AR, Vakhnina NV. Subjective cognitive decline as a predictor of future cognitive decline: a systematic review. Dem Neuropsychol 2020;14(3):248-257. [doi: 10.1590/1980-57642020dn14-030007] [Medline: 32973979]
6. Pike KE, Cavuoto MG, Li L, Wright BJ, Kinsella GJ. Subjective cognitive decline: level of risk for future dementia and mild cognitive impairment, a meta-analysis of longitudinal studies. Neuropsychol Rev 2022 Dec;32(4):703-735. [doi: 10.1007/s11065-021-09522-3] [Medline: 34748154]
7. Valech N, Sánchez-Benavides G, Tort-Merino A, et al. Associations between the subjective cognitive decline-questionnaire's scores, gray matter volume, and amyloid-β levels. J Alzheimers Dis 2019;72(4):1287-1302. [doi: 10.3233/JAD-190624]
8. Ryu SY, Lim EY, Na S, et al. Hippocampal and entorhinal structures in subjective memory impairment: a combined MRI volumetric and DTI study. Int Psychogeriatr 2017 May;29(5):785-792. [doi: 10.1017/S1041610216002349] [Medline: 28067183]
9. Yue L, Wang T, Wang J, et al. Asymmetry of hippocampus and amygdala defect in subjective cognitive decline among the community dwelling Chinese. Front Psychiatry 2018;9:226. [doi: 10.3389/fpsyt.2018.00226] [Medline: 29942265]
10. Yan T, Wang W, Yang L, Chen K, Chen R, Han Y. Rich club disturbances of the human connectome from subjective cognitive decline to Alzheimer's disease. Theranostics 2018;8(12):3237-3255. [doi: 10.7150/thno.23772] [Medline: 29930726]
11. van Rooden S, van den Berg-Huysmans AA, Croll PH, et al. Subjective cognitive decline is associated with greater white matter hyperintensity volume. J Alzheimers Dis 2018;66(3):1283-1294. [doi: 10.3233/JAD-180285] [Medline: 30412485]
12. Sacuiu S, Eckerström M, Johansson L, et al. Increased risk of dementia in subjective cognitive decline if CT brain changes are present. J Alzheimers Dis 2018;66(2):483-495. [doi: 10.3233/JAD-180073]

XSL•FO
RenderX

13. Ohlhauser L, Parker AF, Smart CM, Gawryluk JR. White matter and its relationship with cognition in subjective cognitive decline. Alz Dem Diag Ass Dis Mo 2019 Dec;11(1):28-35. [doi: 10.1016/j.dadm.2018.10.008]
14. Chimthanawala NMA, Haria A, Sathaye S. Non-invasive biomarkers for early detection of Alzheimer's disease: a new-age perspective. Mol Neurobiol 2024 Jan;61(1):212-223. [doi: 10.1007/s12035-023-03578-3] [Medline: 37596437]
15. Rostamzadeh A, Bohr L, Wagner M, Baethge C, Jessen F. Progression of subjective cognitive decline to MCI or Dementia in relation to biomarkers for Alzheimer disease. Neurology (ECronicon) 2022 Oct 25;99(17):e1866-e1874. [doi: 10.1212/WNL.0000000000201072]
16. Wolfsgruber S, Molinuevo JL, Wagner M, et al. Prevalence of abnormal Alzheimer's disease biomarkers in patients with subjective cognitive decline: cross-sectional comparison of three European memory clinic samples. Alzheimers Res Ther 2019 Jan 17;11(1):8. [doi: 10.1186/s13195-018-0463-y] [Medline: 30654834]
17. Jadick MF, Robinson T, Farrell ME, et al. Associations between self and study partner report of cognitive decline with regional tau in a multicohort study. Neurology (ECronicon) 2024 Jun 25;102(12):e209447. [doi: 10.1212/WNL.0000000000209447] [Medline: 38810211]
18. Hao L, Sun Y, Li Y, et al. Demographic characteristics and neuropsychological assessments of subjective cognitive decline (SCD) (plus). Ann Clin Transl Neurol 2020 Jun;7(6):1002-1012. [doi: 10.1002/acn3.51068] [Medline: 32588989]
19. Jenkins A, Tree JJ, Thornton IM, Tales A. Subjective cognitive impairment in 55-65-year-old adults is associated with negative affective symptoms, neuroticism, and poor quality of life. J Alzheimer's Dis 2019;67(4):1367-1378. [doi: 10.3233/JAD-180810]
20. Kielb S, Rogalski E, Weintraub S, Rademaker A. Objective features of subjective cognitive decline in a United States national database. Alzheimers Dement 2017 Dec;13(12):1337-1344. [doi: 10.1016/j.jalz.2017.04.008]
21. Wolfsgruber S, Kleineidam L, Guski J, et al. Minor neuropsychological deficits in patients with subjective cognitive decline. Neurology (ECronicon) 2020 Sep 1;95(9):e1134-e1143. [doi: 10.1212/WNL.0000000000010142] [Medline: 32636322]
22. Webster-Cordero F, Giménez-Llort L. The challenge of subjective cognitive complaints and executive functions in middle-aged adults as a preclinical stage of dementia: a systematic review. Geriatrics (Basel) 2022 Mar 8;7(2):30. [doi: 10.3390/geriatrics7020030] [Medline: 35314602]
23. Stark M, Wolfsgruber S, Kleineidam L, et al. Minor neuropsychological deficits in SCD: association with AD biomarkers and cognitive decline. Alz Dement 2023 Dec;19(S18):e073332. [doi: 10.1002/alz.073332]
24. Rabin LA, Smart CM, Crane PK, et al. Subjective cognitive decline in older adults: an overview of self-report measures used across 19 international research studies. J Alzheimers Dis 2015 Sep 24;48:S63-S86. [doi: 10.3233/JAD-150154] [Medline: 26402085]
25. Chang KH, Wang C, Nester CO, et al. Examining the role of participant and study partner report in widely-used classification approaches of mild cognitive impairment in demographically-diverse community dwelling individuals: results from the Einstein Aging Study. Front Aging Neurosci 2023;15:1221768. [doi: 10.3389/fnagi.2023.1221768] [Medline: 38076542]
26. Butterbrod E, Rabin L, Tommet D, et al. Toward optimizing measurement of self‐perceived cognitive function‐a global expert survey on behalf of the SCD‐I Working Group. Alz Dement 2023 Dec;19(S18):e078216. [doi: 10.1002/alz.078216]
27. Rabin LA, Sikkes SAM, Tommet D, et al. Linking self-perceived cognitive functioning questionnaires using item response theory: the subjective cognitive decline initiative. Neuropsychology 2023 May;37(4):463-499. [doi: 10.1037/neu0000888] [Medline: 37276136]
28. Jessen F. Subjective and objective cognitive decline at the pre-dementia stage of Alzheimer's disease. Eur Arch Psychiatry Clin Neurosci 2014 Nov;264 Suppl 1:S3-S7. [doi: 10.1007/s00406-014-0539-z] [Medline: 25238934]
29. Gelles-Watnick R. Americans' use of mobile technology and home broadband. Pew Research Center. 2024. URL: https://www.pewresearch.org/internet/2024/01/31/americans-use-of-mobile-technology-and-home-broadband [accessed 2025-01-15]
30. Koo BM, Vizer LM. Mobile technology for cognitive assessment of older adults: a scoping review. Innov Aging 2019 Jan;3(1):igy038. [doi: 10.1093/geroni/igy038] [Medline: 30619948]
31. Öhman F, Hassenstab J, Berron D, Schöll M, Papp KV. Current advances in digital cognitive assessment for preclinical Alzheimer's disease. Alz Dem Diag Ass Dis Mo 2021 Jan;13(1):e12217. [doi: 10.1002/dad2.12217]
32. Harris C, Tang Y, Birnbaum E, Cherian C, Mendhe D, Chen MH. Digital neuropsychology beyond computerized cognitive assessment: applications of novel digital technologies. Arch Clin Neuropsychol 2024 Apr 24;39(3):290-304. [doi: 10.1093/arclin/acae016] [Medline: 38520381]
33. Leroy V, Gana W, Aïdoud A, et al. Digital health technologies and Alzheimer's disease clinical trials: might decentralized clinical trials increase participation by people with cognitive impairment? Alz Res Therapy 2023;15(1):87. [doi: 10.1186/s13195-023-01227-4]
34. Sliwinski MJ, Mogle JA, Hyun J, Munoz E, Smyth JM, Lipton RB. Reliability and validity of ambulatory cognitive assessments. Assessment 2018 Jan;25(1):14-30. [doi: 10.1177/1073191116643164] [Medline: 27084835]
35. Weizenbaum EL, Soberanes D, Hsieh S, et al. Capturing learning curves with the multiday Boston Remote Assessment of Neurocognitive Health (BRANCH): feasibility, reliability, and validity. Neuropsychology 2024 Feb;38(2):198-210. [doi: 10.1037/neu0000933] [Medline: 37971862]

XSL•FO
RenderX

36. Cerino ES, Katz MJ, Wang C, et al. Variability in cognitive performance on mobile devices is sensitive to mild cognitive impairment: results from the Einstein Aging Study. Front Digit Health 2021;3:758031. [doi: 10.3389/fdgth.2021.758031] [Medline: 34927132]
37. Benge JF, Kiselica AM, Aguirre A, et al. Technology use and subjective cognitive concerns in older adults. Arch Gerontol Geriatr 2023 Mar;106:104877. [doi: 10.1016/j.archger.2022.104877] [Medline: 36459914]
38. Thompson LI, Vito AN, Kunicki Z, et al. The feasibility and psychometric rigor of using high-frequency, smartphone-based assessments for cognitive screening in older adults. PsychArchives 2024. [doi: 10.1017/S1355617724000328]
39. Thompson LI, Cummings M, Emrani S, et al. Digital clock drawing as an Alzheimer's disease susceptibility biomarker: associations with genetic risk score and APOE in older adults. J Prev Alzheimers Dis 2024;11(1):79-87. [doi: 10.14283/jpad.2023.48] [Medline: 38230720]
40. Papp KV, Jutten RJ, Soberanes D, et al. Early detection of amyloid-related changes in memory among cognitively unimpaired older adults with daily digital testing. Ann Neurol 2024 Mar;95(3):507-517. [doi: 10.1002/ana.26833] [Medline: 37991080]
41. Thompson LI, Kunicki ZJ, Emrani S, et al. Remote and in‐clinic digital cognitive screening tools outperform the MoCA to distinguish cerebral amyloid status among cognitively healthy older adults. Alz Dem Diag Ass Dis Mo 2023 Oct;15(4):e12500. [doi: 10.1002/dad2.12500]
42. Amariglio RE, Sikkes SAM, Marshall GA, et al. Item-level investigation of participant and study partner report on the Cognitive Function Index from the A4 study screening data. J Prev Alzheimers Dis 2021;8(3):257-262. [doi: 10.14283/jpad.2021.8] [Medline: 34101781]
43. Alber J, Lee AKW, Menard W, Monast D, Salloway SP. Recruitment of at-risk participants for clinical trials: a major paradigm shift for Alzheimer's disease prevention. J Prev Alzheimers Dis 2017;4(4):213-214. [doi: 10.14283/jpad.2017.32] [Medline: 29181484]
44. Thompson LI, Harrington KD, Roque N, et al. A highly feasible, reliable, and fully remote protocol for mobile app‐based cognitive assessment in cognitively healthy older adults. Alz Dem Diag Ass Dis Mo 2022 Jan;14(1):e12283. [doi: 10.1002/dad2.12283]
45. Li C, Neugroschl J, Luo X, et al. The utility of the Cognitive Function Instrument (CFI) to detect cognitive decline in non-demented older adults. J Alzheimer's Dis 2017;60(2):427-437. [doi: 10.3233/JAD-161294]
46. Rinaldi P, Mecocci P, Benedetti C, et al. Validation of the five-item geriatric depression scale in elderly subjects in three different settings. J Am Geriatr Soc 2003 May;51(5):694-698. [doi: 10.1034/j.1600-0579.2003.00216.x] [Medline: 12752847]
47. Russell DW. UCLA Loneliness Scale (Version 3): reliability, validity, and factor structure. J Pers Assess 1996 Feb;66(1):20-40. [doi: 10.1207/s15327752jpa6601_2] [Medline: 8576833]
48. Cook SE, Marsiske M, McCoy KJM. The use of the Modified Telephone Interview for Cognitive Status (TICS-M) in the detection of amnestic mild cognitive impairment. J Geriatr Psychiatry Neurol 2009 Jun;22(2):103-109. [doi: 10.1177/0891988708328214] [Medline: 19417219]
49. Robinson T, Klinger H, Buckley R, et al. Subjective cognitive decline across ethnoracial groups in the A4 study. Alzheimers Dement 2023 Sep;19(9):4084-4093. [doi: 10.1002/alz.13138]
50. Li C, Hong Y, Yang X, et al. The use of subjective cognitive complaints for detecting mild cognitive impairment in older adults across cultural and linguistic groups: a comparison of the Cognitive Function Instrument to the Montreal Cognitive Assessment. Alz Dement 2023 May;19(5):1764-1774. [doi: 10.1002/alz.12804]
51. Hakun J, Elbich DB, Roque NA, Yabiku ST, Sliwinski M. Mobile monitoring of cognitive change (M2C2): high-frequency assessments and protocol reporting guidelines. PsyArXiv. Preprint posted online on Aug 22, 2024. [doi: 10.31234/osf.io/34ux5]
52. Munro CE, Boyle R, Chen X, et al. Recent contributions to the field of subjective cognitive decline in aging: a literature review. Alz Dem Diag Ass Dis Mo 2023 Oct;15(4):e12475. [doi: 10.1002/dad2.12475]
53. Barnes J, Dickerson BC, Frost C, Jiskoot LC, Wolk D, van der Flier WM. Alzheimer's disease first symptoms are age dependent: evidence from the NACC dataset. Alzheimers Dement 2015 Nov;11(11):1349-1357. [doi: 10.1016/j.jalz.2014.12.007]
54. La Joie R, Perrotin A, Egret S, et al. Qualitative and quantitative assessment of self‐reported cognitive difficulties in nondemented elders: association with medical help seeking, cognitive deficits, and β‐amyloid imaging. Alz Dem Diag Ass Dis Mo 2016 Jan;5(1):23-34. [doi: 10.1016/j.dadm.2016.12.005]
55. Nester C, Ayers E, Rabin L, Verghese J. Non-memory subjective cognitive concerns predict incident motoric cognitive risk syndrome. Eur J Neurol 2020 Jul;27(7):1146-1154. [doi: 10.1111/ene.14271] [Medline: 32319724]
56. Aschenbrenner AJ, Hassenstab J, Morris JC, Cruchaga C, Jackson JJ. Relationships between hourly cognitive variability and risk of Alzheimer's disease revealed with mixed-effects location scale models. Neuropsychology 2024 Jan;38(1):69-80. [doi: 10.1037/neu0000905] [Medline: 37079810]
57. Huntley JD, Howard RJ. Working memory in early Alzheimer's disease: a neuropsychological review. Int J Geriat Psychiatry 2010 Feb;25(2):121-132. [doi: 10.1002/gps.2314]
58. van Geldorp B, Heringa SM, van den Berg E, Olde Rikkert MGM, Biessels GJ, Kessels RPC. Working memory binding and episodic memory formation in aging, mild cognitive impairment, and Alzheimer's dementia. J Clin Exp Neuropsychol 2015;37(5):538-548. [doi: 10.1080/13803395.2015.1037722] [Medline: 26011711]

XSL•FO
RenderX

59. Liew TM. Subjective cognitive decline, anxiety symptoms, and the risk of mild cognitive impairment and dementia. Alzheimers Res Ther 2020 Sep 11;12(1):107. [doi: 10.1186/s13195-020-00673-8] [Medline: 32917264]
60. Nester CO, Gao Q, Katz MJ, et al. Does the cognitive change index predict future cognitive and clinical decline? Longitudinal analysis in a demographically diverse cohort. J Alzheimers Dis 2024;98(1):319-332. [doi: 10.3233/JAD-230752] [Medline: 38393900]
61. Rabin L, Mogle J, Katz MJ, Sliwinski MJ. Do repeated ambulatory cognitive assessments enhance the relation between objective and subjective cognition in non‐demented community dwelling older adults? Results from the Einstein Aging Study. Alzheimers Dement 2020 Dec;16(S6):e045918. [doi: 10.1002/alz.045918]
62. Mogle J, Turner JR, Bhargava S, Stawski RS, Almeida DM, Hill NL. Individual differences in frequency and impact of daily memory lapses: results from a national lifespan sample. BMC Geriatr 2023 Oct 17;23(1):670. [doi: 10.1186/s12877-023-04363-6] [Medline: 37848825]
63. Rabin L, Mogle J, Wang C, et al. Cognitive lapses reported by daily digital diary, and not by recall‐based questionnaires, predict future MCI in a demographically diverse cohort of community‐dwelling older adults. Alzheimers Dement 2023 Dec;19(S22):e077994. [doi: 10.1002/alz.077994]

Abbreviations

AD: Alzheimer disease
ADRD: Alzheimer disease and related dementias
Aβ: amyloid
CFI: Cognitive Function Instrument
CFI-Cog: Cognitive Function Instrument cognitive items
GDS: Geriatric Depression Scale
M2C2: Mobile Monitoring of Cognitive Change
MAP: Memory and Aging Program
PET: positron emission tomography
SCC: subjective cognitive concern
SCD: subjective cognitive decline
TICSm: modified Telephone Interview for Cognitive Status
UCLA: University of California, Los Angeles

Edited by J Job; submitted 08.07.24; peer-reviewed by I Liampas, S Young; revised version received 24.10.24; accepted 19.11.24; published 04.02.25.

Please cite as:
Nester CO, De Vito AN, Prieto S, Kunicki ZJ, Strenger J, Harrington KD, Roque N, Sliwinski MJ, Rabin LA, Thompson LI
Association of Subjective Cognitive Concerns With Performance on Mobile App–Based Cognitive Assessment in Cognitively Normal Older Adults: Observational Study
JMIR Aging 2025;8:e64033
URL: https://aging.jmir.org/2025/1/e64033
doi:10.2196/64033

© Caroline O Nester, Alyssa N De Vito, Sarah Prieto, Zachary J Kunicki, Jennifer Strenger, Karra D Harrington, Nelson Roque, Martin J Sliwinski, Laura A Rabin, Louisa I Thompson. Originally published in JMIR Aging (https://aging.jmir.org), 4.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Effectiveness of a Dyadic Technology–Enhanced Home-Based Horticultural Therapy on Psychosocial Well-Being Among People With Dementia and Their Family Caregivers: Multimethods Pilot Study

Patrick Pui Kin Kor[1], RN, PhD; Justina Yat Wa Liu[1,2], RN, PhD; Arkers Kwan Ching Wong[1], RN, PhD; Alex Pak Lik Tsang[1], CPsychol, PhD; Han Zhi Tan[1], MSc; Daphne Sze Ki Cheung[1,3,4], RN, PhD; Humphrey Kwong Wai Leung[5], DBA, FHKEng; Frances Kam Yuet Wong[1], RN, PhD

[1]School of Nursing, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[2]Research Institute for Smart Ageing, PolyU Academy for Interdisciplinary Research, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[3]School of Nursing and Midwifery, Faculty of Health, Deakin University, Melbourne, Australia

[4]Centre for Quality and Patient Safety Research/Alfred Health Partnership, Institute for Health Transformation, Deakin University, Melbourne, Australia

[5]University Fellows Association, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

Corresponding Author:
Patrick Pui Kin Kor, RN, PhD
School of Nursing
The Hong Kong Polytechnic University
11 Yuk Choi Road
Hung Hom, Kowloon
Hong Kong
China (Hong Kong)
Phone: 852 27665622
Email: patrick.kor@polyu.edu.hk

Abstract

Background: Horticultural therapy (HT) has been proposed to be an effective intervention for improving the psychosocial well-being of people with dementia and their caregivers. However, constraints such as limited land space in high-density cities, unstable weather, and lack of gardening experience may hamper the delivery of HT to people with dementia and their caregivers.

Objective: This pilot study aimed to examine the feasibility and preliminary effects of a technology-enhanced home-based HT for people with dementia and their caregivers using a hydroponic indoor growing system.

Methods: A single-group pre-post design was adopted. A total of 37 dyads of people with dementia and their caregivers participated in 3 weekly face-to-face sessions, followed by 8 weeks of home-based horticultural activities. Outcomes were measured at baseline and postintervention (at week 11), including feasibility outcomes, cognitive function, neuropsychiatric symptoms, and happiness levels of people with dementia. Caregivers' outcomes included positive aspects of caregiving, perceived stress levels, depressive symptoms, caregiver distress, and happiness levels. Semistructured focus group interviews were conducted with the caregivers to further explore their horticultural experience.

Results: Intervention feasibility was established with a completion rate of 83.78% and an attrition rate of 2.63% (n=1). Significant improvements were detected in caregiver distress (P<.05) and the happiness level of people with dementia (P<.01). The qualitative findings indicated that HT improved the psychological well-being of both people with dementia and caregivers, enhanced the relationships between caregivers and people with dementia, expanded the caregivers' social networks, and enhanced the autobiographical memory of people with dementia.

Conclusions: This pilot study provides evidence on the feasibility of using a hydroponic indoor grower to conduct home-based HT for people with dementia and their caregivers. The findings suggest positive effects on the psychological well-being of both people with dementia and their caregivers. Caregivers reported potential positive effects of HT on the autobiographical memory retrieval of people with dementia. Due to the pilot nature of this study, a control group was not employed. Therefore, large-scale randomized controlled trials are encouraged to further confirm the effectiveness of the intervention.

XSL•FO
RenderX

Trial Registration: ClinicalTrials.gov NCT05577975; https://clinicaltrials.gov/study/NCT05577975

(JMIR Aging 2025;8:e66017) doi:10.2196/66017

KEYWORDS

horticultural activity; dementia; caregivers; dyadic intervention; technology–enhanced intervention

Introduction

With the global aging population, dementia has become a significant public health concern. The World Health Organization estimates that the number of people with dementia worldwide is approximately 55 million, with this figure expected to rise to about 78 million by 2030 and 139 million by 2050 [1]. The global cost of dementia saw a substantial increase of 62% between 2010 and 2019, with the economic burden projected to reach an estimated US $1313.4 billion in 2019. Notably, informal care significantly contributes to roughly half of the total global costs [2]. People with dementia often exhibit behavioral and psychological symptoms of dementia (BPSD), including agitation, anxiety, irritability, depression, apathy, disinhibition, and changes in sleep or appetite [3]. More than 80% of people with dementia were affected by aggression, apathy, agitation, and depression, while around 30% experienced hallucinations and delusions [3,4]. Conceivably, BPSD adversely impacts the capacity to perform activities of daily living (ADL) and the maintenance of interpersonal relationships. This in turn results in significant stress and negative emotions for their family caregivers. For instance, caring for people with dementia has been associated with elevated levels of psychological distress and stress, as well as lower levels of self-efficacy, subjective well-being, and physical health [5]. Due to the significant level of psychological distress and stress involved in caregiving activities, approximately 50% of caregivers of people with dementia are susceptible to major depressive disorder [6]. A high level of caregiver stress is a risk factor for unhealthy behaviors, such as a sedentary lifestyle, inadequate dietary habits, and substance abuse, leading to an increased vulnerability to chronic illnesses among caregivers [7,8]. Caregiver burden is also linked to an increased risk of engaging in abusive behavior towards the care recipient [9]. It is therefore important to develop psychosocial interventions to promote the psychological well-being of both caregivers and people with dementia, as well as to alleviate the BPSD experienced by people with dementia.

Horticultural therapy (HT) has been proposed to be an effective intervention for improving the psychosocial well-being of people with dementia and their caregivers. HT is grounded in the Biophilia Hypothesis, which emphasizes the healing effects of human-plant interactions on physical, mental, and emotional domains [10]. According to the Attention Restoration Theory (ART), engaging with natural environments can restore cognitive fatigue by automatically capturing attention, thereby allowing the replenishment of directed attention capacity [11]. The Stress Recovery Theory (SRT) also proposes that interacting with nature can trigger physiological and psychological stress reduction and elicit positive affect, which in turn facilitates the replenishment of depleted psychological resources after stressful events [12,13]. For instance, HT establishes a sense of control, empowerment, and cooperation, which in turn can facilitate emotional stability when under a state of cognitive overload [14]. The physiological effects of HT on stress reduction in older adults are characterized by lower cortisol, heart rate, and blood pressure [15,16]. It has also been found that increased high α and low α wave activities after a single treatment [17]. Increasing α and theta activity in the brain's occipital lobes helped decrease anxiety and improve functioning in those with generalized anxiety disorder [18]. In addition, the social interactive nature of HT can facilitate social support for vulnerable groups [19].

More recently, HT has been extended in its application to people with dementia and demonstrated positive effects on BPSD, as well as enhancing physical and cognitive abilities, including memory and orientation, among people with dementia [20]. In the context of caregiving for people with dementia, HT can foster meaningful contact and social interaction between people with dementia and their caregivers [21,22]. A randomized controlled trial (RCT) demonstrated that HT can effectively reduce caregiving burden and promote the quality of life among caregivers of people with dementia [23]. Another RCT reported evidence of the protective effects of HT on the immune system and cognitive functioning, as well as its positive effects on reducing anxiety and promoting social interactions among older adults [24]. Several systematic reviews and meta-analyses have shown reliable effects of HT in improving cognitive functioning in people with dementia [25,26]. For instance, an improved capacity for autobiographical memory retrieval is often observed through engaging in horticultural activities [27]. In addition, HT also demonstrated benefits in alleviating BPSD in people with dementia, as suggested in multiple systematic reviews and meta-analyses [25,27-29].

Traditionally, HT is led by a registered therapist to achieve tailored rehabilitative or vocational goals. HT commonly includes participatory tasks such as cultivating, pruning, weeding, and planting flowers. Ornamental HT may also involve garden tours and nature viewing [16,22]. However, in densely populated areas, HT may not have the luxury of spaces to engage in home-based horticultural activities. The extreme climate in certain areas may create additional difficulties in engaging in outdoor horticultural activities. In addition, for caregivers who may be occupied with caregiving tasks, horticultural activities that require a significant amount of skills and time, such as monitoring the condition of various plants, may pose additional challenges for them [30].

While HT is promising in managing BPSD in people with dementia and improving the quality of life in both people with dementia and their caregivers, it is important to consider the environmental and spatial conditions in which HT is implemented, as these factors may inadvertently impact its

XSL•FO
RenderX

effectiveness. In this study, we proposed a novel approach to take advantage of technology to overcome the above limitations and to make horticultural activities more accessible for people with dementia and their caregivers, especially in densely populated areas. In recent years, there has been a growing body of research on the application of technology-assisted HT, such as using virtual reality to deliver HT (eg, VR Garden) [31]. However, there are certain limitations. One of the main drawbacks is the lack of real-world sensory experiences, such as smells and tactile sensations, which are inherent to traditional HT. Additionally, the use of virtual reality may lead to cybersickness, a condition characterized by symptoms similar to motion sickness, which can negatively impact the user experience and therapeutic outcomes [32]. To address these limitations, a hydroponic indoor grower can create a controllable, optimized environment for engaging in horticultural activities in a home-based HT for people with dementia and their caregivers. The hydroponic indoor grower also allows mobile app connectivity to track the progress of plant growth, which can guide people with dementia and their caregivers to engage in horticultural activities. Therefore, this pilot study aimed to address the following research objectives:

1. To examine the feasibility and acceptability of conducting a technology-enhanced home-based HT among people with dementia and their caregivers.
2. To investigate the preliminary effects of the intervention in (1) improving cognitive function and BPSD of people with dementia, (2) promoting positive caregiving experiences in family caregivers, and (3) reducing the caregivers' level of stress and depressive symptoms.

Methods

Study Design

To examine the feasibility and preliminary effects of the technology-enhanced home-based HT on improving the psychological well-being of people with dementia and their caregivers, we used an explanatory sequential multimethods approach. A single-group pre-post design is used for our quantitative assessment. The entire intervention lasted for 11 weeks, with outcome measures taken at baseline (T0) and post intervention (T1). A qualitative approach was used to explore the participants' experiences in-depth in the HT. Semistructured focus group interviews were conducted with participants after the intervention, with the aim of identifying the strengths, limitations, and difficulties experienced during the program.

Participants

Participants were recruited from a convenience sample of people with dementia and their caregivers in collaboration with 4 community centers operated by the nongovernmental organization (NGO) that provide gerontology services. Promotional posters were used to advertise the program in NGOs. The program allowed interested participants to enroll independently, while NGO staff reached out to interested individuals to provide further information and assistance. The Hong Kong version of the Montreal Cognitive Assessment-5-Minute (MoCA-5-min) [33] was administered by a research assistant in screening for stages of dementia in potential participants. Eligible participants were then enrolled based on the following inclusion and exclusion criteria. The inclusion criteria for people with dementia were as follows: (1) 65 years or above, (2) diagnosed with any type of dementia at the early to moderate stage, and (3) community-dwelling and living with family caregivers (ie, noninstitutionalized). The criteria for caregivers included the following: (1) at least 18 years old; (2) blood or by-marriage relatives (eg, spouses, siblings, children, and grandchildren) of a person who has been clinically diagnosed with dementia, regardless of its type; and (3) take up caring responsibilities ranging from physical aids (eg, transportation, financial assistance, personal hygiene, and decision-making) to emotional support, and provided most of the daily care and support for the people with dementia (daily contact of at least 4 hours). Participants were deemed ineligible if they had a current medical diagnosis with any acute mental illnesses that would otherwise prevent them from engaging in HT.

An a priori power analysis was performed using G*Power 3.1 [34], based on effect sizes in literature related to the effects of HT on psychological outcomes. A meta-analysis reported an effect size of 0.55 [35], the power analysis was formulated using a 1-tailed *t* test to detect a medium effect size with a power of 0.80 at a *P* value of .05, which resulted in a target sample size of 23. Considering an attrition rate of approximately 15% [36], it was necessary to include more than 27 dyads in the study.

Ethics Approval

Ethical approval for the study was obtained from the Institutional Review Board of The Hong Kong Polytechnic University (approval number HSEARS20220801002), and the trial was registered in ClinicalTrials.gov (NCT05577975). Prior to participation, all participants (both people with dementia and their caregivers) were required to provide informed consent. The process of obtaining informed consent from people with dementia adhered to the guidelines outlined by the Alzheimer's Society. A demonstration session was set up to explain the study's purpose, procedures, risks, benefits, and voluntary nature clearly to both people with dementia and their caregivers. During this session, the research assistant explained the consent form, demonstrated the equipment, and provided an overview of the participation process. People with dementia and their caregivers were allowed to try out some parts of the HT before signing the consent form. Only those people with dementia who were still willing to participate after this experience were allowed to sign the consent form. Ongoing assent was also obtained from participants with dementia throughout the study, and their willingness to participate was regularly monitored and respected.

Intervention

In the technology-enhanced home-based HT, the activities were also designed based on the Biophilia Hypothesis, with an emphasis on human-plant interactions. We further incorporated interactive components between people with dementia and their caregivers, in which people with dementia participated in the planting process alongside their caregivers. The caregivers assumed responsibility for decision-making and progress monitoring for the planting process, while people with dementia followed directions from caregivers and provided assistance

XSL•FO
RenderX

throughout the process. Examples of horticultural tasks performed by caregivers included formulating planting plans, adjusting optimal light conditions, strategizing fruit harvesting schedules, and devising culinary techniques for the harvest. Meanwhile, people with dementia executed tasks such as observation, flower watering, active participation in planting light regulation, and assisting caregivers with harvesting. According to the SRT and ART, these activities are posited to have cognitive and stress restoration effects. Based on the Sensory Integration Theory, these activities can mobilize the 5 senses through multisensory stimulation, tactile input, and olfactory and gustatory experiences [37,38], which can further enhance the cognitive functioning of people with dementia. Furthermore, the interactive components may enhance the dyadic relationship between people with dementia and their caregivers.

The technology-enhanced home-based HT was facilitated by a hydroponic indoor smart grower, which is a smart device that provides a greenhouse environment with adjustable LED grow lights and smart sensors (Figure 1). Features of the device include automatic watering and a removable water reservoir, as well as built-in smart sensors to detect nutrients, air temperature, humidity, light intensity, water temperature, and water level. The mobile app connectivity via Wi-Fi allows participants to monitor the growth conditions with notifications when water and nutrients need to be replenished, minimizing the probability of plant growing failure. The app can also customize planting plans for each type of plant and automatically adjust to suitable light intensity, resulting in faster plant growth and harvest. Additionally, the app can track the progress of plant growth and guide users to conduct horticultural activities.

Figure 1. Hydroponic indoor smart grower.



Upon the project's commencement, the hydroponic indoor smart grower was delivered to the enrolled participants. With reference to the optimal therapeutic dose between 100 and 500 minutes of HT for stress reduction [39], the intervention consisted of 3 weekly face-to-face (F-T-F) training sessions, each lasting 90 minutes. This was followed by 8 weeks of home-based horticultural activities (Table 1), with a total weekly time commitment of 1 hour. The program activities are presented in Table 2. A biweekly telephone call schedule was established to follow up on progress and program adherence. Additionally, a social media platform group was created for the participants to ask questions and share their horticultural experiences. The social media platform provided additional insight into participants' adherence to home-based horticultural activities.

XSL•FO
RenderX

Table 1. Intervention schedule of the technology-enhanced home-based horticultural therapy.

	F-T-F[a] sessions			Home-based horticultural activities with biweekly telephone follow-up							
	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11
Duration (minutes)	90	90	90	60	60	60	60	60	60	60	60

[a]F-T-F: face-to-face.

Table 2. Program content of the technology-enhanced home-based horticultural therapy.

Week	Theme	Objectives	Reality orientation contents
4	Vision; Hearing; Smell	Stimulating the senses of vision, hearing, and smell; focusing on the appearance of the smart grower; encouraging people with dementia-caregiver interaction.	Focusing on LED[a] light differences according to certain modes of the smart grower; focusing on the water sound when the smart grower performs auto-watering system; smelling the seeds to see if different seeds have different smells; and sharing their sensory information to each other.
5	Smell; Touch; Motor	Stimulating the senses of smell and touch; enhancing visual-spatial skills and executive; promoting motor skills.	Smelling the leaves at the beginning stage of the plants; touching the plants (To see the texture, eg, hard/soft/smooth/rough/solid/fluid); trimming plants; taking photos to record plant growth.
6	Vision; Touch	Stimulating the senses of vision and touch; enhancing visual-spatial skills and executive functioning; enhancing cognitive ability through comparison.	Examining the pattern, shape, size, or height of the vegetables (eg, bok choy vs tomato) in day and night; touching different types of vegetables to see if the textures are different; sharing their sensory information to each other; taking photos to record plant growth.
7	Smell; Vision; Motor; Memory	Stimulating the senses of smell and touch; promoting motor functioning; enhancing visual-spatial skills and executive; enhancing cognitive functioning (memory) through comparison; encouraging people with dementia-caregiver interaction.	Comparing the shape, size, and height of different vegetables to see the differences in the growing process; smelling the leaves or stem of vegetables; pruning the vegetables; sharing and discussing what are the differences from the beginning; taking photos to record plant growth.
8	Motor; Touch; Decision making	Stimulating the senses of touch; promoting motor functioning; stimulating cognitive functioning through decision-making process; enhancing visual-spatial skills and executive.	Touching the leaves or stem of the plants or vegetables to see if the textures are different from the beginning to mid-late stage of planting; planning and discussing which plants to cut; taking photos to record plant growth.
9	Vision; Motor; Decision making	Stimulating the senses of vision; promoting motor functioning; stimulating cognitive functioning through decision-making process; enhancing visual-spatial skills and executive.	Counting the number of vegetables (eg, how many bak choy/tomato?); deciding which vegetables to cut; cutting suitable vegetables; taking photos to record plant growth.
10	Smell; Vision; Motor; Memory	Stimulating the senses of smell and touch; promoting motor functioning; enhancing visual-spatial skills and executive; enhancing cognitive functioning (memory) through comparison; encouraging people with dementia-caregiver interaction.	Comparing the shape, size, and height of different vegetables to see the differences in the growing process; smelling the leaves or stem of vegetables; pruning the vegetables; sharing and discussing what are the differences from the beginning; taking photos to record plant growth.
11	Motor; Vision; Decision-making	Stimulating the senses of vision; promoting motor functioning; stimulating cognitive functioning through decision-making process.	Deciding what type of dishes to cook (salads or pasta); cooking the decided dishes based on the vegetable; recording the odor and color of the dishes; taking photos to record plant growth.

[a]LED: light-emitting diode.

Measures

Overview

Measurements were taken at T0 and T1. The feasibility of the intervention was evaluated by means of attrition rate, attendance rate, and completion rate. The outcome measures for people with dementia included self-care abilities, cognitive function, behavioral symptoms, and happiness levels. Caregivers' outcomes consisted of positive caregiving experiences, perceived stress levels, depressive symptomatology, caregiver distress, and happiness levels.

People With Dementia Outcomes

The Barthel Index was used to measure the capacity to perform ADLs. Barthel Index comprised 10 ADLs, including feeding, bathing, grooming, dressing, using the toilet, transferring (moving from the bed to the chair and back), mobility (on level surfaces), climbing stairs, and controlling bowel and bladder functions [40]. Items are rated based on whether individuals can perform activities independently, with some assistance, or

XSL•FO
RenderX

are dependent (scored as 10, 5, or 0). The index yields a total score out of 100, with higher scores indicating greater levels of functional independence [40]. We categorized the Barthel Index scores as follows: 0-20 (total dependence), 21-60 (severe dependence), 61-90 (moderate dependence), 91-99 (mild dependence), and 100 (total independence) [41]. The Barthel Index had excellent test-retest reliability (intraclass correlation coefficient [ICC]≥0.962) [42].

The MoCA-5-min was used to measure the cognitive functions of people with dementia over the telephone, administered by social workers [33]. It comprises 4 domains: attention (immediate recall of 5 words), executive function/language (1-minute verbal fluency), orientation (6-item date and geographic orientation), and memory (delayed recall and recognition of 5 words learned in item 1). The total scores of the MoCA-5-min range from 0 to 30, with higher scores representing a higher level of cognitive function. The MoCA-5-min had an excellent test-retest reliability (ICC=0.89) [33].

BPSD of people with dementia was measured using the Neuropsychiatric Inventory Questionnaire (NPI-Q) [43]. The NPI-Q consisted of 12 neuropsychiatric symptoms, including delusions, hallucinations, agitation/aggression, irritability/depression, anxiety, euphoria/exuberance, apathy/indifference, disinhibition, irritability/instability, abnormal motor behavior, nocturnal behavioral disorders, and appetite/eating disorders. Each symptom is evaluated based on frequency (1=rarely to 4=very often, once or more per day) and severity (1=mild [causing little distress in the patient] to 3=severe [very disturbing to the patient and challenging to redirect]). A total score for each domain is calculated by multiplying the frequency and severity ratings. The overall NPI score is derived by summing the scores across all domains, with a higher score indicating more severe neuropsychiatric symptoms. The NPI-Q has demonstrated satisfactory psychometric properties [43].

The Visual Analogue Scale of Happiness was used to measure the happiness level of people with dementia [44]. The Visual Analogue Scale of Happiness is an analog scale containing 6 faces expressing emotions ranging from great happiness to deep sadness, and people with dementia need to identify which face best represents their mood. Subjects that chose the faces “Little Unhappy,” “Unhappy,” or “Very Unhappy” were prone to exhibit depressive symptoms [45]. The reliability and validity of the scale have been validated [46].

Caregivers Outcomes

The Positive Aspects of Caregiving Scale was used to measure caregivers’ positive caregiving experience [47]. It consists of 11 items with 2 subscales: enriching life and affirming self. Participants answered on a 5-point scale ranging from 0 (strongly disagree) to 4 (strongly agree). Scores range from 0 to 44, with higher scores indicating more positive self-perceptions of caregiving. The Positive Aspects of Caregiving Scale demonstrates high levels of internal consistency, with a Cronbach α of 0.89 among the family caregivers of people with dementia in Hong Kong [47].

The Chinese version of the Perceived Stress Scale was used to measure the stress level in caregivers [48]. It is a 10-item measurement rated on a 5-point Likert-type scale, ranging from 0 (never) to 4 (very often). The total score ranges from 0 to 40, with higher scores indicating higher levels of perceived stress. The Perceived Stress Scale showed acceptable levels of psychometric properties, which included an internal consistency reliability of Cronbach α of 0.85 and a test-retest reliability coefficient of 0.85 [49].

The Center for Epidemiological Studies Depression Scale was used to measure depressive symptoms in caregivers over a 1-week recall period [50]. It consists of 20 items and is rated on a 4-point Likert scale, ranging from 0 (rarely or none of the time) to 3 (most or almost all the time). Scores range from 0 to 60, with higher scores indicating greater depressive symptoms. The Center for Epidemiological Studies Depression Scale has demonstrated acceptable psychometric properties, including a test-retest reliability of 0.91 and an internal consistency of 0.86 [51].

The Neuropsychiatric Inventory-Distress Scale (NPI-D) was used to measure caregiving distress [52]. It assesses the psychological distress levels caused by neuropsychiatric symptoms, as reported in the NPI-Q, to the caregivers [43]. Participants rated their distress levels according to the severity of symptoms in people with dementia using a 6-point Likert scale, which ranges from 0 (not at all distressing) to 5 (very severely or extremely distressing). The Cronbach α coefficient for the Chinese version of the NPI-D was 0.72 [53].

Perceived happiness was measured using the Subjective Happiness Scale (SHS) [54]. The SHS consists of 4 items; 2 of these assess the current state of happiness, including absolute happiness ratings and relative happiness ratings compared with peers. The remaining 2 items measure trait happiness by asking participants to indicate the extent to which they identify with descriptions of both happy and unhappy individuals. The SHS is rated on a 7-point Likert-type scale, ranging from negative to positive. Responses to the 4 items are averaged, yielding a range from 1 to 7. A higher score indicates greater levels of happiness. The Cronbach α for the Chinese version of the SHS was 0.82, and the test-retest reliability was 0.70 [55].

Feasibility Outcomes

Intervention feasibility was assessed using the completion rate, attendance rate, and attrition rate. The attrition rate was calculated by dividing the number of dropouts by the total number of enrolled subjects. The attendance rate was determined by averaging the individual attendance rates, which were calculated by dividing the number of F-T-F sessions attended by the total number of F-T-F sessions. The completion rate was defined as the percentage of participants who attended more than 80% of the sessions, which included home-based horticultural activities and totaled more than 600 minutes.

Quantitative Analysis

Data analysis was performed using IBM SPSS (version 23). We first explored patterns of sociodemographic characteristics of the participants. Then, paired-sample *t* tests were used to compare the outcomes before and after the intervention. Unless

XSL•FO
RenderX

otherwise specified, the statistical analyses performed did not exceed a 2-tailed α level of 0.05.

Qualitative Analysis

Given that caregivers of people with dementia experience varying levels of distress and may respond differently to the intervention, we adopted stratified purposive sampling based on this criterion to enhance the representativeness of the data. The caregiver's level of distress was assessed using NPI-D, with mild distress categorized as 0-13 points (77.78%), moderate distress as 14-26 points (16.67%), and severe distress as >27 points (5.56%) [52]. Participants for the qualitative study were selected through stratified sampling based on these different levels of distress. A total of 4 semistructured focus groups were conducted with 18 participants. A trained research assistant led the focus groups, asking participants to share their experiences using the smart growers, as well as their interactions with people with dementia during the program. A probing question—"What is your experience related to the use of the device in interacting with your relative?"—was used to facilitate the discussions. The focus group interviews were recorded and transcribed with the participants' consent.

Thematic analysis was used to analyze the transcribed verbatim; the authors read the texts, generated initial codes, and developed potential themes [56]. The process concluded with data saturation, when no new findings were discovered. Emergent themes were discussed and agreed upon by the researchers. The comprehensiveness of the content analysis was ensured through digital recordings and verbatim transcriptions, which were independently completed and then discussed by 2 researchers. The emergent themes, validated through researcher consensus after thorough discussion, exemplify the methodological rigor underlying the study's credibility.

Results

Quantitative Findings

A total of 50 dyads of people with dementia and their family caregivers were invited to participate in the study, and 38 dyads met the selection criteria and joined the study. The attrition rate was 2.63%, with 1 participant dropping out due to behavioral issues after the first F-T-F training session. The attendance rate for F-T-F sessions was 94.64%. A total of 36 out of 38 dyads attended the first session, while 35 out of 37 dyads attended both the second and third sessions. The completion rate was 83.78%, as 31 out of 37 participants achieved a participation rate of 80% or higher (>600 minutes).

Sample characteristics of people with dementia are presented in Table 3. The majority of participants were male (21/38, 55.3%), with a mean age of 81.8 (SD 8.4) years. Regarding the self-care abilities of the participants, only 21.1% (n=8) exhibited no dependence, while the majority displayed mild to moderate dependence.

Table 3. Sociodemographic characteristics of people with dementia (N=38).

Variables	Values
Age (in years), mean (SD)	81.8 (8.4)
Age (in years), range	60-102
Sex , n (%)	
Male	21 (55.3)
Female	17 (44.7)
Marital status , n (%)	
Married	23 (60.5)
Divorced/Widowed	15 (39.5)
Self-care abilities , n (%)	
Total independence	8 (21.1)
Mild dependence	25 (65.8)
Moderate dependence	5 (13.2)
Severe dependence	0 (0.0)
Total dependence	0 (0.0)

Demographic characteristics of caregivers are summarized in Table 4. Female caregivers accounted for 73.7%, with a mean age of 58.4 (SD 8.8) years. The average duration of care provided to people with dementia was 4.7 (SD 0.3) hours per day.

XSL•FO
RenderX

Table 4. Sociodemographic characteristics of caregivers (N=38).

Variables	Values
Age (in years), mean (SD)	58.4 (8.8)
Age (in years), range	34-76
Sex, n (%)	
Male	10 (26.3)
Female	28 (73.7)
Marital status, n (%)	
Single	2 (5.3)
Married	30 (78.9)
Divorced/widowed	6 (15.8)
Educational level, n (%)	
Primary	10 (26.3)
Secondary	20 (52.6)
College or above	8 (21.1)
Employment status, n (%)	
Unemployed/retired	20 (52.6)
Employed	18 (47.4)
Monthly household income, n (%)	
Less than 6000 HKD[a]	11 (28.9)
6000-9999 HKD	8 (21.1)
10,000-14,999 HKD	10 (26.3)
15,000-19,999 HKD	7 (18.4)
20,000 HKD or above	2 (5.3)
Relationship with people with dementia, n (%)	
Spouse	20 (52.6)
Children/in-laws	18 (47.4)
Living arrangement with people with dementia, n (%)	
Same household	23 (60.5)
Different household	15 (39.5)
Duration of care provided to people with dementia, in hours/day (hours), mean (SD)	4.7 (0.3)

[a]HKD: Hong Kong Dollar (1 HKD=US $ 0.13).

For people with dementia, there were significant pre-post improvements in cognitive functions and behavioral and BPSD, although the results did not reach statistical significance. Furthermore, we observed a significant improvement in the happiness level of people with dementia from T0 (mean 7.67, SD 1.40) to T1 (mean 8.79, SD 1.02; P<.01; Table 5).

Table 5. Outcome measures of the technology-enhanced home-based horticultural therapy for people with dementia (N=37, one participant dropped out).

Variables	T0[a], mean (SD)	T1[b], mean (SD)	*P* value	95% CI
Happiness	7.67 (1.40)	8.79 (1.02)	<.01	–1.75 to –0.46
Cognitive functions	14.07 (6.58)	14.36 (6.55)	.61	–1.41 to 0.84
BPSD[c]	10.03 (8.26)	7.72 (5.22)	.20	–1.29 to 5.91

[a]T0: preintervention.

[b]T1: postintervention.

[c]BPSD: behavioral and psychological symptoms of dementia.

XSL•FO
RenderX

Regarding caregivers, at a descriptive level, we observed pre-post improvements in perceived stress, depressive symptoms, and happiness, although the differences did not reach a statistically significant level. Specifically, we observed significant improvements in caregiver distress from T0 (mean 12.93, SD 9.40) to T1 (mean 7.83, SD 8.01; $P<.05$; Table 6).

Table 6. Outcome measures of the technology-enhanced home-based horticultural therapy for caregivers (N=37, one participant dropped out).

Variables	T0[a], mean (SD)	T1[b], mean (SD)	*P* value	95% CI
Stress	25.72 (6.17)	25.66 (6.00)	.96	–2.89 to 3.03
Depressive symptoms	16.79 (10.70)	17.03 (9.73)	.92	–5.28 to 4.80
Positive aspects of caregiving	31.34 (8.23)	31.00 (10.29)	.89	–4.86 to 5.55
Caregiver distress	12.93 (9.40)[c]	7.83 (8.01)[c]	<.05[c]	0.16 to 10.05[c]
Happiness	6.74 (3.93)	7.08 (4.37)	.41	–1.42 to 3.35

[a]T0: preintervention.

[b]T1: postintervention.

[c]Significant effects.

Qualitative Findings

Overview

A total of 18 caregivers joined the focus group. The majority of them were either children (n=12) or spouses (n=6) of the care recipients. Overall, the qualitative findings from the focus interviews can be summarized under the themes of (1) perceived benefits and (2) perceived challenges of the technology-enhanced home-based HT.

Perceived Benefits

Under the theme of perceived benefits, we further identified four subthemes: (1) improved emotional well-being, (2) improved relationship with people with dementia, (3) increased social network among caregivers, and (4) reinforced autobiographical memory retrieval and storytelling abilities.

Improved Emotional Well-Being

Caregivers positively recalled how they had enjoyed engaging in horticultural activities at home using the smart grower. They shared their excitement at witnessing the growing process of the vegetables. One participant shared that growing vegetables alleviated her stress, as it had become a meaningful goal in her life. Additionally, she appreciated the colorful lighting of the machine, which they found to be a decorative element in their homes.

> *Seeing the growth of vegetables is like looking after a baby. The first thing I do every day is to take photos of the vegetables… I also feel happy when I see the attractive symbols on the machine, such as the stars and mountain symbols.* [P3]

The caregivers were pleased that they could enjoy the fruits of their labor. The vegetables they grew, such as carrots and lettuce, were used to create delicious dishes. They felt a sense of pride and satisfaction after receiving compliments from friends and family members who saw photos of their harvest.

Similar perspectives were also evident in people with dementia. They felt happiness and excitement about the growing process, especially during the harvest. As one caregiver shared,

> *My father was also involved in the growing of lettuce…He looked at the smart grower and watered the plant every day. He felt happy when he saw the harvest in the smart grower…He could feel the nature at home.* [P10]

Another caregiver noted that while her husband could not actively participate in the growing process, he enjoyed watching it, particularly the colorful lighting on the smart growing machine.

Improved Relationship With People With Dementia

Caregivers found that their relationship with people with dementia improved through home-based horticultural activities. As one caregiver explained,

> *I put the smart grower near the window, and my husband (people with dementia) will help to water the plant…We are taking care of the plant together…It is just like a toy for us.* [P4]

These activities offered a way for them to collaborate and grow the vegetables together.

Increase Social Network Among Caregivers

Caregivers also reported increased interaction with other caregivers through the home horticultural activities. They were able to ask questions about the smart grower and share photos of their harvest in WhatsApp groups. This facilitated the formation of new friendships, as they had a common topic to discuss.

I uploaded photos to the WhatsApp group and others posted pictures on the group. So, I see how other people's plants are growing and talk to people.(P1) We have set a WhatsApp groups. When I don't understand sometimes, I will ask in the group. The student assistants will help me in the group and other caregivers will also answer and we will discuss together in the group. [P9]

The increased social interactions were corroborated by our WhatsApp group records. Participants actively shared recipes for their harvests with each other. Some caregivers displayed their stir-fried Chinese cabbage, while others shared recipes for garlic Chinese cabbage and methods for cooking baby turnip

XSL•FO
RenderX

with chieh-qua. More noteworthy was how caregivers enthusiastically shared their planting progress and complimented each other's growing vegetables. These positive interactions led to the formation of new friendships among participants.

Reinforced Autobiographical Memory Retrieval and Storytelling Abilities

One caregiver noted that the smart grower offered an opportunity for people with dementia to reminisce about the past. The people with dementia had experience growing crops in their homeland during childhood, and the home horticultural activities revived those childhood memories. This opportunity facilitated the sharing of historical anecdotes and the exercise of cognitive abilities, such as autobiographical memory retrieval.

> *Dad is interested in gardening, as growing vegetables with this machine reminded him of his country gardening days, and every now and then he would tell us about his previous experiences.* [P13]

Difficulties

While the horticultural activities generally had positive impacts on people with dementia and their caregivers, there were also some perceived challenges. These challenges can be summarized into two subthemes: (1) operational difficulties in using the smart grower, and (2) light disturbance from the smart grower.

Operational Difficulties in Using the Smart Grower

Some caregivers found the smart grower not easy to operate due to its sophisticated design featuring many high-tech components. Issues such as Wi-Fi connection problems were encountered during use. For people with dementia with mild to moderate dependence, the complexity of the operations seemed to diminish their patience and motivation, presenting a barrier to their active participation in the entire growth process.

> *Sometimes, the elder have no caregivers to help them. So, they can't do it by themselves because they don't understand English very well. For example, the word 'reset' they can't understand. Also, they would not know what to do if the Wi-Fi disconnected.* [P13]

Light Disturbance from the Smart Grower. Additionally, other caregivers expressed concerns about the lighting emitted by the smart grower. They reported using covers or towels to minimize light disturbance at night, indicating a need for better light management features in the device.

> *The machine emits light and is very disruptive to sleep. Sleeping in a very small house is really affected. When sleeping in the living room, the light turns red and is a little scary. It may have some effect on plants, but not sure. The next day it gives off normal light again.* [P14]

Discussion

Principal Findings

This pilot study aimed to investigate the feasibility of technology-enhanced home-based HT in improving the psychological well-being of people with dementia and their caregivers. Overall, our quantitative results suggested that technology-enhanced home-based HT is a feasible dyadic intervention for both people with dementia and their caregivers. The overall attrition rate was low (2.63%), with a satisfactory completion rate (83.78%). These findings indicate that caregivers and people with dementia are receptive to using technology for engaging in horticultural activities. Further qualitative findings revealed that both people with dementia and their caregivers were excited about using the smart growing machine, as it provided a novel experience for them. They also welcomed the interactive horticultural activities and found that their relationship with each other improved as a result.

The use of technology-enhanced HT has progressively become an increasingly popular approach to delivering HT. For example, virtual reality–based HT has recently been demonstrated as a viable and effective approach in promoting positive psychological outcomes in older adults, such as enhancing self-esteem and mastery [57]. Technology-enhanced HT can mitigate limitations posed by factors such as venue and weather conditions, offering a more flexible and convenient therapeutic approach. In contrast to a virtual reality–based HT, our technology-enhanced home-based HT not only exhibits the benefits of using technologies but also allows participants to touch and feel plants in reality, which can be more relatable to the feeling of nurturing life. According to the Biophilic Hypothesis, actual human-plant interactions play a vital role in exerting the therapeutic effects of HT [58].

Our quantitative results suggested a significant reduction in caregiver distress, aligning with previous studies that used a traditional delivery mode of HT. Previous research has demonstrated that horticultural activities can alleviate the caregiving burden [23]. Our qualitative findings further suggested that home-based horticultural activities may simultaneously enable people with dementia to engage in meaningful and socially interactive experiences with their caregivers, potentially improving their relationships and alleviating caregiver distress. Moreover, our findings indicated a significant increase in happiness levels among people with dementia. These findings also corroborated previous studies using traditional HT that reported a positive emotional impact on people with dementia [59]. Our qualitative findings provided further insights that the heightened levels of subjective well-being could be attributed to participants' enthusiasm for observing plant growth and acquiring a sense of satisfaction through harvesting activities.

Overall, our findings are in line with the ART, in that perceiving plants in their natural environment has the potential to enhance attention restoration and increase persuasive effects [11,60]. Our findings are also in line with SRT, suggesting that HT is associated with positive affect, which in turn facilitates stress recovery [12,61]. Although we did not observe a significant decrease in perceived stress levels, we found an overall improvement in emotional well-being among caregivers. For example, an increase in emotional well-being is inversely correlated with elevated levels of stress and anger [62]. Consistently, HTs can reduce cortisol levels through engagement in horticultural activities, such as self-harvesting processes, and the consumption of harvested fruits and vegetables [15]. The use of social media groups also facilitated social interactions

XSL•FO
RenderX

among caregivers and provided a source of support. The mutual encouragement and compliments about each other's vegetables, along with sharing recipes for their harvest created positive interactions that extended beyond the initial horticultural activities, offering an additional benefit of alleviating caregiving stress [63-65].

Our qualitative findings suggested that some participants reminisced about their past experiences in farming during the activities. Reminiscing about past events requires the recall of autobiographical memory [66,67], and actively externalizing autobiographical memory for later retrieval (ie, reminiscing about past events) can enhance memory performance and afford greater control over memory encoding processes [68]. While previous studies have suggested that engagement in horticultural activities can enhance the cognitive abilities of people with dementia [21,22], we did not detect a significant pre-post improvement in the cognitive functions of people with dementia. It is plausible that because the MoCA is a global measure of cognition, it may partially explain why we did not detect a significant pre-post improvement in MoCA scores, despite an increasing trend descriptively, as people with dementia primarily improved in memory abilities. However, this interpretation should be made cautiously due to our small sample size. Future studies should incorporate a larger-scale RCT to further investigate the cognitive benefits of the technology-enhanced home-based HT.

Limitations

This study had several limitations. Due to the pilot nature of the research to establish the feasibility of a technology-enhanced home-based HT, a control group was not employed. Therefore, a full-scale RCT is encouraged to provide further confirmatory evidence about the effectiveness of this intervention. Furthermore, there were only 3 F-T-F sessions provided to offer training and support to the participants. This dosage may have been insufficient. In the qualitative findings, caregivers reported having technical difficulties in using the smart grower at home, such as Wi-Fi connection issues. Previous studies have found similar observations that the older population may have a lower level of readiness for home technology [69]. Therefore, the complexity of the device use may be a significant barrier to the engagement of the older population, particularly for people with dementia [70]. Consequently, it is conceivable that the program outcomes could be further enhanced with more F-T-F sessions and better integrated technical support for people with dementia and their caregivers. A more comprehensive training and support system may help address the operational difficulties encountered with the smart grower technology.

Conclusion

The findings of this pilot study support the feasibility and offer preliminary evidence of using a hydroponic indoor grower to conduct technology-enhanced home-based HT to improve the psychological well-being of people with dementia and their caregivers. There were positive effects on the emotional well-being of people with dementia and a decrease in caregiver distress. The qualitative findings further suggested improved relationships between people with dementia and caregivers. Additionally, there were potential benefits of this intervention in improving the autobiographical memory retrieval capacity of people with dementia, indicating the cognitive benefits of this intervention. Future studies are encouraged to incorporate a larger and more diverse sample to further evaluate the sustainability of the positive psychological benefits of the technology-enhanced home-based HT for people with dementia and their caregivers.

Acknowledgments

This study received funding from Bright Future Charitable Foundation, Hong Kong, China. This publication reflects only the authors' views, and the Bright Future Charitable Foundation is not liable for any use that may be made of the information contained therein.

Authors' Contributions

PK, JL, AW, DC, HL, and FW conceptualized the research design. AT and HZT conducted the interview and assisted in data collection. AT and HZT conducted the qualitative analyses. PK, AT, and HZT performed statistical analyses and interpreted the data. PK, PK, AT, and HZT drafted the manuscript. All authors read and approved the final manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Shin JH. Dementia epidemiology fact sheet 2022. Ann Rehabil Med 2022;46(2):53-59 [FREE Full text] [doi: 10.5535/arm.22027] [Medline: 35508924]
2. Wimo A, Seeher K, Cataldi R, Cyhlarova E, Dielemann JL, Frisell O, et al. The worldwide costs of dementia in 2019. Alzheimers Dement 2023;19(7):2865-2873 [FREE Full text] [doi: 10.1002/alz.12901] [Medline: 36617519]
3. Scales K, Zimmerman S, Miller SJ. Evidence-based nonpharmacological practices to address behavioral and psychological symptoms of dementia. Gerontologist 2018;58(suppl_1):S88-S102 [FREE Full text] [doi: 10.1093/geront/gnx167] [Medline: 29361069]

XSL•FO
RenderX

4. Holroyd S, Currie L, Wooten GF. Prospective study of hallucinations and delusions in parkinson's disease. J Neurol Neurosurg Psychiatry 2001 Jun;70(6):734-738 [FREE Full text] [doi: 10.1136/jnnp.70.6.734] [Medline: 11385005]
5. Ballard C, Corbett A, Chitramohan R, Aarsland D. Management of agitation and aggression associated with alzheimer's disease: controversies and possible solutions. Curr Opin Psychiatry 2009;22(6):532-540. [doi: 10.1097/YCO.0b013e32833111f9] [Medline: 19696673]
6. Huang SS. Depression among caregivers of patients with dementia: associative factors and management approaches. World J Psychiatry 2022;12(1):59-76 [FREE Full text] [doi: 10.5498/wjp.v12.i1.59] [Medline: 35111579]
7. Hoffman D, Zucker H. A call to preventive action by health care providers and policy makers to support caregivers. Prev Chronic Dis 2016;13:E96 [FREE Full text] [doi: 10.5888/pcd13.160233] [Medline: 27442996]
8. Culberson JW, Kopel J, Sehar U, Reddy PH. Urgent needs of caregiving in ageing populations with alzheimer's disease and other chronic conditions: support our loved ones. Ageing Res Rev 2023;90:102001. [doi: 10.1016/j.arr.2023.102001] [Medline: 37414157]
9. Wang M, Sun H, Zhang J, Ruan J. Prevalence and associated factors of elder abuse in family caregivers of older people with dementia in central China cross-sectional study. Int J Geriatr Psychiatry 2019;34(2):299-307. [doi: 10.1002/gps.5020] [Medline: 30370657]
10. Niklasson J. Horticultural therapy for homeless people. SLU University Library. 2007. URL: https://stud.epsilon.slu.se/12506/ [accessed 2017-10-26]
11. Kaplan S. The restorative benefits of nature: toward an integrative framework. J Environ Psychol 1995;15(3):169-182. [doi: 10.1016/0272-4944(95)90001-2]
12. Ulrich RS, Simons RF, Losito BD, Fiorito E, Miles MA, Zelson M. Stress recovery during exposure to natural and urban environments. J Environ Psychol 1991;11(3):201-230. [doi: 10.1016/s0272-4944(05)80184-7]
13. Vacher P, Filaire E, Mourot L, Nicolas M. Stress and recovery in sports: effects on heart rate variability, cortisol, and subjective experience. Int J Psychophysiol 2019;143:25-35. [doi: 10.1016/j.ijpsycho.2019.06.011] [Medline: 31255740]
14. Kenmochi T, Kenmochi A, Hoshiyama M. Effects of horticultural therapy on symptoms and future perspective of patients with schizophrenia in the chronic stage. Journal of Therapeutic Horticulture 2019;29(1):1-10 [FREE Full text]
15. Yun J, Yao W, Meng T, Mu Z. Effects of horticultural therapy on health in the elderly: a review and meta-analysis. Z Gesundh Wiss 2023;32(10):1-27 [FREE Full text] [doi: 10.1007/s10389-023-01938-w] [Medline: 37361284]
16. Nicholas SO, Giang AT, Yap PLK. The effectiveness of horticultural therapy on older adults: a systematic review. J Am Med Dir Assoc 2019;20(10):1351.e1-1351.e11. [doi: 10.1016/j.jamda.2019.06.021] [Medline: 31402135]
17. Hassan A, Qibing C, Tao J. Physiological and psychological effects of gardening activity in older adults. Geriatr Gerontol Int 2018;18(8):1147-1152. [doi: 10.1111/ggi.13327] [Medline: 29626378]
18. Dadashi M, Birashk B, Taremian F, Asgarnejad AA, Momtazi S. Basic Clin Neurosci 2015;6(1):14-20 [FREE Full text] [Medline: 27504152]
19. Sempik J, Rickhuss C, Beeston A. The effects of social and therapeutic horticulture on aspects of social behaviour. Br J Occup Ther 2014;77(6):313-319. [doi: 10.4276/030802214x14018723138110]
20. Lin Y, Lin R, Liu W, Wu W. Effectiveness of horticultural therapy on physical functioning and psychological health outcomes for older adults: a systematic review and meta-analysis. J Clin Nurs 2022;31(15-16):2087-2099. [doi: 10.1111/jocn.16095] [Medline: 34694042]
21. Jarrott SE, Kwack HR, Relf D. An observational assessment of a dementia-specific horticultural therapy program. HortTechnology 2002;12(3):403-410. [doi: 10.21273/horttech.12.3.403]
22. Ferrini F. Horticultural therapy and its effect on people's health. Adv Hortic Sci 2003;17(2):77-87 [FREE Full text]
23. Kim YH, Park CS, Bae H, Lim EJ, Kang KH, Lee ES, et al. Horticultural therapy programs enhancing quality of life and reducing depression and burden for caregivers of elderly with dementia. J People Plants Environ 2020;23(3):305-320. [doi: 10.11628/ksppe.2020.23.3.305]
24. Ng KST, Sia A, Ng MKW, Tan CTY, Chan HY, Tan CH, et al. Effects of horticultural therapy on asian older adults: a randomized controlled trial. Int J Environ Res Public Health 2018;15(8):1705 [FREE Full text] [doi: 10.3390/ijerph15081705] [Medline: 30096932]
25. Zhao Y, Liu Y, Wang Z. Effectiveness of horticultural therapy in people with dementia: a quantitative systematic review. J Clin Nurs 2022;31(13-14):1983-1997. [doi: 10.1111/jocn.15204] [Medline: 32017241]
26. Tu HM, Chiu PY. Meta-analysis of controlled trials testing horticultural therapy for the improvement of cognitive function. Sci Rep 2020;10(1):14637 [FREE Full text] [doi: 10.1038/s41598-020-71621-7] [Medline: 32884098]
27. Lu LC, Lan SH, Hsieh YP, Yen YY, Chen JC, Lan SJ. Horticultural therapy in patients with dementia: a systematic review and meta-analysis. Am J Alzheimers Dis Other Demen 2020;35:1533317519883498 [FREE Full text] [doi: 10.1177/1533317519883498] [Medline: 31690084]
28. Whear R, Coon JT, Bethel A, Abbott R, Stein K, Garside R. What is the impact of using outdoor spaces such as gardens on the physical and mental well-being of those with dementia? a systematic review of quantitative and qualitative evidence. J Am Med Dir Assoc 2014;15(10):697-705 [FREE Full text] [doi: 10.1016/j.jamda.2014.05.013] [Medline: 25037168]

XSL•FO
RenderX

29. Xu M, Lu S, Liu J, Xu F. Effectiveness of horticultural therapy in aged people with depression: a systematic review and meta-analysis. Front Public Health 2023;11:1142456 [FREE Full text] [doi: 10.3389/fpubh.2023.1142456] [Medline: 36969640]
30. Kamioka H, Tsutani K, Yamada M, Park H, Okuizumi H, Honda T, et al. Effectiveness of horticultural therapy: a systematic review of randomized controlled trials. Complement Ther Med 2014;22(5):930-943. [doi: 10.1016/j.ctim.2014.08.009] [Medline: 25440385]
31. Xu T, Liao D, Peng X, Liang G, Li Y, Zheng X. Design of virtual reality scenes on horticultural therapy treatment. 2012 Presented at: International Conference on Human-Computer Interaction; 2022 June 26; Heidelberg. [doi: 10.15695/vurj.v8i0.3644]
32. Rebenitsch L, Owen C. Review on cybersickness in applications and visual displays. Virtual Reality 2016;20(2):101-125. [doi: 10.1007/s10055-016-0285-9]
33. Wong A, Nyenhuis D, Black SE, Law LS, Lo ES, Kwan PW, et al. Montreal cognitive assessment 5-minute protocol is a brief, valid, reliable, and feasible cognitive screen for telephone administration. Stroke 2015;46(4):1059-1064. [doi: 10.1161/strokeaha.114.007253]
34. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods 2009;41(4):1149-1160. [doi: 10.3758/brm.41.4.1149]
35. Pan iru I, Ronaldson A, Sima N, Dregan A, Sima R. The impact of gardening on well-being, mental health, and quality of life: an umbrella review and meta-analysis. Syst Rev 2024;13(1):45 [FREE Full text] [doi: 10.1186/s13643-024-02457-9] [Medline: 38287430]
36. Siu AMH, Kam M, Mok I. Horticultural therapy program for people with mental illness: a mixed-method evaluation. Int J Environ Res Public Health 2020;17(3) [FREE Full text] [doi: 10.3390/ijerph17030711] [Medline: 31979057]
37. Smith Roley S, Mailloux Z, Miller-Kuhaneck H, Glennon TJ. Understanding Ayres' sensory integration. Sacred Heart University 2007;12(17):CE1-CE8 [FREE Full text]
38. Yun HS, Yun SY, Choi BJ. Effects of horticultural activities designed to stimulate five senses on the sensory development of children. J People Plants Environ 2018;21(5):369-378. [doi: 10.11628/ksppe.2018.21.5.369]
39. Lu S, Liu J, Xu M, Xu F. Horticultural therapy for stress reduction: a systematic review and meta-analysis. Front Psychol 2023;14:1086121 [FREE Full text] [doi: 10.3389/fpsyg.2023.1086121] [Medline: 37564307]
40. Mahoney FI, Barthel DW. Functional evaluation: the Barthel index. Md State Med J 1965;14:61-65. [Medline: 14258950]
41. Filipska-Blejder K, Ślusarz R. Characteristics of the barthel scale in terms of neurological care. PNN 2022;11(4):180-184. [doi: 10.15225/pnn.2022.11.4.6]
42. Dos Santos Barros V, Bassi-Dibai D, Guedes CLR, Morais DN, Coutinho SM, de Oliveira Simões G, et al. Barthel Index is a valid and reliable tool to measure the functional independence of cancer patients in palliative care. BMC Palliat Care 2022;21(1):124 [FREE Full text] [doi: 10.1186/s12904-022-01017-z] [Medline: 35820921]
43. Kaufer DI, Cummings JL, Ketchel P, Smith V, MacMillan A, Shelley T, et al. Validation of the NPI-Q, a brief clinical form of the neuropsychiatric inventory. J Neuropsychiatry Clin Neurosci 2000;12(2):233-239. [doi: 10.1176/jnp.12.2.233] [Medline: 11001602]
44. Matsubayashi K, Kimura S, Iwasaki T, Okumiya K, Hamada T, Fujisawa M, et al. [Evaluation of subjective happiness in the elderly using a visual analogue scale of happiness to analyze the effect of life style and neurobehavioral function on subjective happiness]. Nihon Ronen Igakkai Zasshi 1992;29(11):817-822 [FREE Full text] [doi: 10.3143/geriatrics.29.817] [Medline: 1491476]
45. César KG, Brucki SMD, Takada LT, Nascimento LFC, Gomes CMS, Almeida MCS, et al. Performance of the visual analogue scale of happiness and of the cornell scale for depression in dementia in the tremembé epidemiological study, Brazil. Dement Neuropsychol 2014;8(4):389-393 [FREE Full text] [doi: 10.1590/S1980-57642014DN84000014] [Medline: 29213931]
46. Matsubayashi K, Kimura S, Iwasaki T, Okumiya S, Hamada T, Fujisawa M, et al. [Evaluation of subjective happiness in the elderly using a visual analogue scale of happiness in correlation with depression scale]. Nihon Ronen Igakkai Zasshi 1992;29(11):811-816 [FREE Full text] [doi: 10.3143/geriatrics.29.811] [Medline: 1491475]
47. Lou VW, Lau BH, Cheung KS. Positive aspects of caregiving (PAC): scale validation among Chinese dementia caregivers (CG). Arch Gerontol Geriatr 2015;60(2):299-306. [doi: 10.1016/j.archger.2014.10.019] [Medline: 25488014]
48. Cohen S, Kamarck T, Mermelstein R. Perceived stress scale. APA PsycTests 1983:10. [doi: 10.1037/t02889-000]
49. Leung DY, Lam TH, Chan SS. Three versions of perceived stress scale: validation in a sample of Chinese cardiac patients who smoke. BMC Public Health 2010;10:513 [FREE Full text] [doi: 10.1186/1471-2458-10-513] [Medline: 20735860]
50. Chin WY, Choi EP, Chan KT, Wong CK. The psychometric properties of the center for epidemiologic studies depression scale in Chinese primary care patients: factor structure, construct validity, reliability, sensitivity and responsiveness. PLoS One 2015;10(8):e0135131 [FREE Full text] [doi: 10.1371/journal.pone.0135131] [Medline: 26252739]
51. Miller WC, Anton HA, Townson AF. Measurement properties of the CESD scale among individuals with spinal cord injury. Spinal Cord 2008;46(4):287-292. [doi: 10.1038/sj.sc.3102127] [Medline: 17909558]

XSL•FO
RenderX

52. Kaufer DI, Cummings JL, Christine D, Bray T, Castellon S, Masterman D, et al. Assessing the impact of neuropsychiatric symptoms in alzheimer's disease: the neuropsychiatric inventory caregiver distress scale. J Am Geriatr Soc 1998;46(2):210-215. [doi: 10.1111/j.1532-5415.1998.tb02542.x] [Medline: 9475452]
53. Wang T, Xiao S, Li X, Wang H, Liu Y, Su N, et al. Reliability and validity of the Chinese version of the neuropsychiatric inventory in mainland China. Int J Geriatr Psychiatry 2012;27(5):539-544. [doi: 10.1002/gps.2757] [Medline: 21769938]
54. Lyubomirsky S, Lepper HS. A measure of subjective happiness: preliminary reliability and construct validation. Soc Indic Res 1999;46(2):137-155. [doi: 10.1023/A:1006824100041]
55. Nan H, Ni MY, Lee PH, Tam WW, Lam TH, Leung GM, et al. Psychometric evaluation of the Chinese version of the subjective happiness scale: evidence from the Hong Kong FAMILY cohort. Int J Behav Med 2014;21(4):646-652 [FREE Full text] [doi: 10.1007/s12529-014-9389-3] [Medline: 24515396]
56. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
57. Fan CC, Choy CS, Huang CM, Chih PS, Lee CC, Lin FH, et al. The effects of a combination of 3D virtual reality and hands-on horticultural activities on mastery, achievement motives, self-esteem, isolation and depression: a quasi-experimental study. BMC Geriatr 2022;22(1):744 [FREE Full text] [doi: 10.1186/s12877-022-03431-7] [Medline: 36096746]
58. Kaplan R. Some psycholog Ical benefits of gardening. Environ Behav 1973;5(2):145-162. [doi: 10.1177/001391657300500202]
59. Gigliotti CM, Jarrott SE, Yorgason J. Harvesting health: effects of three types of horticultural therapy activities for persons with dementia. Dementia 2004;3(2):161-180. [doi: 10.1177/1471301204042335]
60. Ulrich RS, Simons RF, Losito BD, Fiorito E, Miles MA, Zelson M. Stress recovery during exposure to natural and urban environments. J Environ Psychol 1991;11(3):201-230. [doi: 10.1016/s0272-4944(05)80184-7]
61. Zuckerman M. Development of a situation-specific trait-state test for the prediction and measurement of affective responses. J Consult Clin Psychol 1977;45(4):513-523. [doi: 10.1037//0022-006x.45.4.513] [Medline: 886036]
62. Steptoe A, Deaton A, Stone AA. Subjective wellbeing, health, and ageing. Lancet 2015;385(9968):640-648 [FREE Full text] [doi: 10.1016/S0140-6736(13)61489-0] [Medline: 25468152]
63. Hui Lin H, Mu Chuan L, Wan Lan C. Benefits of horticultural therapy on emotional and perceived stress of family caregivers: a pilot study. Research in Applied Psychology 2017;67:5-53.
64. Jueng RN, Lin CY, Huang YH. Systematic review on the positive mental health impact of older adults participation in horticultural activities in long term care facilities. Horticulturae 2023;9(10):1076. [doi: 10.3390/horticulturae9101076]
65. Houben M, Van Engen V, Kenning G, Brankaert R. Smile: capturing and sharing personal photos to stimulate social relations and support self-identity in dementia. New York: Springer; 2021.
66. Lu X, Kelly MO, Risko EF. Offloading information to an external store increases false recall. Cognition 2020;205:104428. [doi: 10.1016/j.cognition.2020.104428] [Medline: 32863020]
67. Janssen SMJ. Introduction to the cognitive abilities account for the reminiscence bump in the temporal distribution of autobiographical memory. Psychol Rep 2020;123(1):12-42. [doi: 10.1177/0033294119843221] [Medline: 31039675]
68. Kelly MO, Risko EF. Offloading memory: serial position effects. Psychon Bull Rev 2019;26(4):1347-1353. [doi: 10.3758/s13423-019-01615-8] [Medline: 31161530]
69. Liu L, Stroulia E, Nikolaidis I, Miguel-Cruz A, Rios Rincon AR. Smart homes and home health monitoring technologies for older adults: a systematic review. Int J Med Inform 2016;91:44-59. [doi: 10.1016/j.ijmedinf.2016.04.007] [Medline: 27185508]
70. Cheung DSK, Ho LYW, Chan LCK, Kwok RKH, Lai CKY. A home-based dyadic music-with-movement intervention for people with dementia and caregivers: a hybrid type 2 cluster-randomized effectiveness-implementation design. Clin Interv Aging 2022;17:1199-1216 [FREE Full text] [doi: 10.2147/CIA.S370661] [Medline: 35978943]

Abbreviations

ADL: activities of daily living
ART: Attention Restoration Theory
BPSD: behavioral and psychological symptoms of dementia
F-T-F: face-to-face
HT: horticultural therapy
ICC: intraclass correlation coefficient
MoCA-5-min: Montreal Cognitive Assessment-5-Minute
NGO: nongovernmental organization
NPI-D: Neuropsychiatric Inventory-Distress Scale
NPI-Q: Neuropsychiatric Inventory Questionnaire
RCT: randomized controlled trial
SHS: Subjective Happiness Scale
SRT: Stress Recovery Theory

XSL•FO
RenderX

Edited by X Xiang; submitted 02.09.24; peer-reviewed by A O'Donnell, LC Lu; comments to author 27.11.24; revised version received 05.12.24; accepted 18.12.24; published 05.02.25.

Please cite as:
Kor PPK, Liu JYW, Wong AKC, Tsang APL, Tan HZ, Cheung DSK, Leung HKW, Wong FKY
Effectiveness of a Dyadic Technology–Enhanced Home-Based Horticultural Therapy on Psychosocial Well-Being Among People With Dementia and Their Family Caregivers: Multimethods Pilot Study
JMIR Aging 2025;8:e66017
URL: https://aging.jmir.org/2025/1/e66017
doi:10.2196/66017
PMID:

©Patrick Pui Kin Kor, Justina Yat Wa Liu, Arkers Kwan Ching Wong, Alex Pak Lik Tsang, Han Zhi Tan, Daphne Sze Ki Cheung, Humphrey Kwong Wai Leung, Frances Kam Yuet Wong. Originally published in JMIR Aging (https://aging.jmir.org), 05.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Enhancing Older Adults' Lives Through Positive Aging Perception, Quality-of-Life Enhancement, and Social Support to Drive Acceptance and Readiness Toward Indoor Assistive Technology: Cross-Sectional Study

Ka Po Wong[1], PhD; Pei-Lee Teh[2], PhD; Weng Marc Lim[3], PhD; Shaun Wen Huey Lee[4], PhD

[1]Department of Applied Social Sciences, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[2]Gerontechnology Laboratory and School of Business, Monash University Malaysia, Sunway City, Malaysia

[3]Sunway Business School, Sunway University, Sunway City, Malaysia

[4]School of Pharmacy, Monash University Malaysia, Sunway City, Malaysia

Corresponding Author:
Pei-Lee Teh, PhD
Gerontechnology Laboratory and School of Business
Monash University Malaysia
Jalan Lagoon Selatan
Sunway City, 47500
Malaysia
Phone: 60 355144971
Email: teh.pei.lee@monash.edu

Abstract

Background: The growing aging population faces increasing mobility limitations, highlighting the need for assistive technologies as potential solutions. These technologies support the independence and well-being of older adults and individuals with mobility challenges. Indoor mobility is essential for daily activities and significantly impacts their lives. Limited indoor mobility can reduce quality of life and heighten the risk of falls.

Objective: This study explores how positive aging perceptions, quality-of-life enhancements, and social support influence the acceptance and readiness of indoor assistive technologies among older adults.

Methods: A cross-sectional study was conducted at a gerontechnology laboratory, requiring participants to visit the facility in person. Each 60-minute session included demonstrations of various indoor assistive technologies and the completion of a questionnaire. The assistive technologies showcased encompassed a wide range of devices. Participants' positive aging perceptions, quality-of-life enhancements, social support, technology acceptance, and readiness were measured using validated scales. Data were analyzed with AMOS (version 28; IBM Corp) and SPSS (version 28; IBM Corp), using structural equation modeling and multivariate analysis of covariance to assess the effects of predictors while controlling for demographic factors.

Results: A total of 104 older adults aged 60 years and older participated, with a mean age of 67.92 (SD 5.68) years. Structural equation modeling indicated that positive aging perception has a significant influence on older adults' control beliefs (P=.095), comfort (P=.047), and confidence (P<.001) in gerontechnology. Multivariate analysis revealed significant combined effects of quality-of-life enhancement (P=.01) and social support (P=.03) on technology acceptance and readiness, wherein quality-of-life enhancement (P=.001) and social support (P=.008) negatively impacted security perception. Among demographic variables, educational level significantly impacted gerontechnology confidence (P=.004) while ethnicity influenced optimism (P=.003).

Conclusions: This study sheds light on key factors affecting older adults' acceptance and readiness to adopt indoor assistive technologies. Findings highlight the importance of fostering positive aging perceptions through these technologies. Addressing issues related to control beliefs, comfort, and confidence in gerontechnology is essential to enhance technology acceptance and readiness among older adults. Future research should investigate the underlying mechanisms and create targeted interventions to support successful technology adoption in this population.

(JMIR Aging 2025;8:e59665) doi:10.2196/59665

XSL•FO
RenderX

KEYWORDS

indoor assistive technology; positive aging perceptions; quality of life; social support; technology acceptance; technology readiness

Introduction

Background

Recent years have seen an increased focus on addressing the needs of older adults and individuals with mobility challenges [1]. As the global population ages, the prevalence of mobility challenges has become a significant societal concern [2]. The World Health Organization (WHO) projects that by 2050, the number of people aged 60 years and older will reach 2 billion, with a substantial portion experiencing mobility limitations [3]. This demographic shift has intensified efforts to develop innovative solutions that support the independence and well-being of older adults and those with mobility challenges [4].

Indoor mobility, which includes activities such as sitting, standing, and walking, is essential for maintaining independence and overall well-being, especially among older adults and individuals with mobility difficulties [5,6]. Limited indoor mobility can lead to serious consequences, including an increased risk of falls and a diminished quality of life [5]. Therefore, enhancing indoor mobility is crucial for enabling these individuals to participate actively in daily activities and sustain their independence [7]. Assistive technologies, such as commodes, home care beds, and reclining wheelchairs, have emerged as solutions to address the specific needs of this population. These technologies help overcome mobility limitations and improve functional abilities [8], potentially transforming how older adults and individuals with mobility challenges interact with their living environments. Successful integration of assistive technologies requires careful selection and personalization to match the unique capabilities, needs, and preferences of each user [9].

Indoor assistive technologies encompass a wide range of tools, devices, and equipment designed to enhance and support the independence and mobility of older adults and individuals with functional limitations within indoor settings [10]. These technologies include mobility aids (eg, scooters, walkers, and wheelchairs), transfer and positioning devices (eg, grab bars, patient lifts, and reclining chairs), smart home systems (eg, automated lighting and voice-controlled appliances), and various adaptive equipment (eg, bed rails, shower chairs, and toilet risers). Tailored to address the specific challenges faced by older adults and those with mobility challenges, these technologies facilitate daily activities and help maintain independence at home and in other indoor environments [11]. Through targeted support to increase functional capacity, indoor assistive technologies empower older adults and individuals with mobility challenges, enhancing their daily lives [12] and improving overall quality of life and well-being [13,14].

Gerontechnology, the intersection of gerontology and advanced technology, aims to enhance the health, independence, and quality of life of older adults [15]. Incorporating gerontechnological advancements into indoor assistive technologies addresses the unique challenges faced by this population, fostering greater independence and well-being [15]. Examples of gerontechnology applications include advanced mobility aids such as smart walkers and wheelchairs, cognitive support tools such as memory aids and smart home systems, and social engagement platforms such as web-based communication tools [16]. These technologies assist older adults in maintaining physical movement, performing daily activities, and reducing social isolation, thereby creating an environment where they can live independently and with dignity. Acceptance and readiness to adopt these technologies are influenced by positive aging perceptions, enhanced quality of life, and social support systems. This study examines how these factors drive technology acceptance and readiness among older adults, ultimately improving their lives through indoor assistive technologies.

Theoretical Background

This study is grounded in the Theory of Planned Behavior (TPB) [17]. The TPB posits that individuals' behavioral intentions are influenced by 3 primary constructs: attitudes, subjective norms, and perceived behavioral control. In the context of this study, attitudes represent older adults' overall evaluation and perception of indoor assistive technologies. This study investigates whether older adults who hold a positive aging perception and believe that assistive technologies enhance their quality of life are more likely to adopt and use these technologies [18,19].

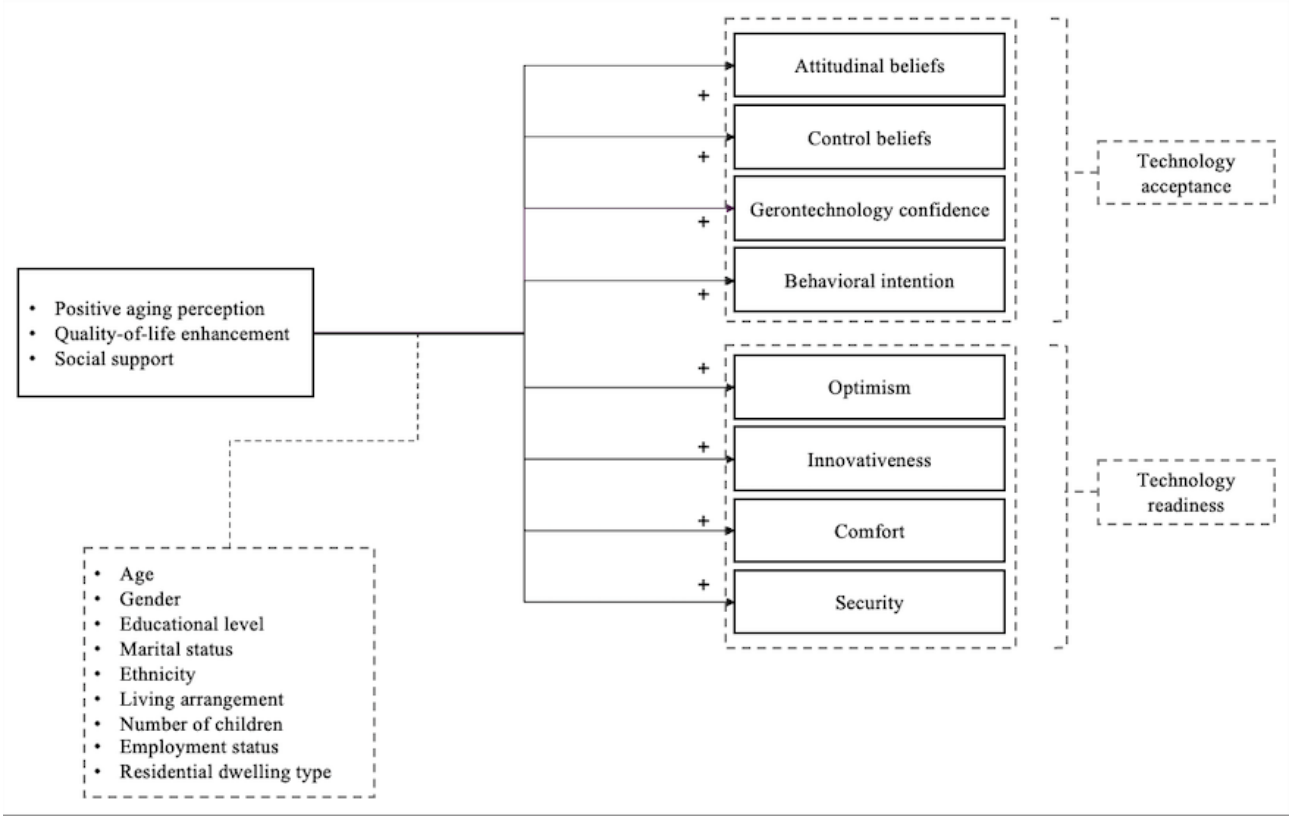
Subjective norms within the TPB refer to the social influences and societal expectations that shape an individual's behavior [20]. In this study, subjective norms are assessed through the role of social support—specifically, the assistance provided by family and friends—in influencing the acceptance and readiness to adopt indoor assistive technologies among older adults.

Perceived behavioral control is another critical construct of the TPB, encompassing individuals' beliefs and perceptions about their ability to perform a specific behavior [21,22]. This study measures perceived behavioral control by evaluating older adults' perceptions of their confidence and readiness to adopt and use indoor assistive technologies.

Applying the TPB enables this study to identify the factors that drive various aspects of older adults' acceptance and readiness to embrace indoor assistive technologies. Such insights are valuable for developing interventions and strategies aimed at promoting engagement, well-being, and overall quality of life among older adults [23]. The study's conceptual model is illustrated in Figure 1.

XSL•FO
RenderX

Figure 1. Study conceptual model.



Recognizing the needs of older adults with diminished mobility highlights the importance of creating innovative strategies to enhance their independence and physical and psychological well-being [1]. Although existing studies have examined the perceptions of technologies among older adults with and without disabilities or impairments, there remains a need for more evidence-based research focused on the specific factors that influence the acceptance and readiness to adopt indoor assistive technologies. Understanding these factors can provide promising pathways for improving indoor mobility and enhancing the quality of life for this population [13,22].

Previous research has explored various dimensions of indoor assistive technologies. For example, Gitlin et al [24] discovered that assistive devices and home modifications significantly improved functional abilities and reduced the risk of falls among older adults, indicating positive attitudes toward these technologies. Similarly, several studies have shown that the use of assistive devices in home environments leads to increased independence and reduced caregiver burden [15,25], underscoring the role of social support in technology acceptance. In addition, Demiris et al [26] and Liu et al [27] examined older adults' perceptions of their ability to use smart home technologies, aligning with the concept of perceived behavioral control in the TPB [17]. More recent studies, such as Peek et al [19], have identified key factors influencing the acceptance of technology for aging in place, including perceived ease of use, usefulness, and social influence, while Mitzner et al [28] highlighted the importance of user experience in technology acceptance among older adults. These studies provide empirical evidence supporting the benefits of indoor assistive technologies and emphasize the necessity of understanding user acceptance and readiness.

The global demographic shift toward an aging population and the rising prevalence of mobility limitations present urgent challenges that must be addressed. With the projected increase in the number of older adults, developing interventions and strategies that effectively cater to their unique needs and preferences becomes crucial [4]. To achieve a comprehensive understanding of older adults' needs regarding assistive technologies, targeted strategies must investigate their perspectives, acceptance, and readiness of aging in relation to indoor assistive technologies. Conducting such research can bridge existing knowledge gaps and inform the creation of strategies tailored to the specific needs of older adults, ultimately enhancing their overall independence and well-being [5].

This study aims to explore the influences of positive aging perception, quality-of-life enhancement through assistive technologies, and social support on technology acceptance and readiness among older adults in an indoor setting. Specifically, this study seeks to understand how these factors interact and contribute to older adults' attitudes and readiness to adopt and use indoor assistive technologies. The hypothesis posits that a positive aging perception, improved quality of life, and supportive social relationships positively affect older adults' technology acceptance, including attitudinal beliefs, control beliefs, confidence in gerontechnology, and behavioral intention. It is further hypothesized that these factors also positively impact older adults' readiness, which encompasses comfort, innovativeness, optimism, and security.

XSL•FO
RenderX

Methods

Study Setting

This study used a cross-sectional experiential design conducted at a gerontechnology laboratory located at an Australian university's international branch campus in Malaysia. Participants were required to visit the laboratory in person to take part in the study. Each session lasted approximately 60 minutes and included active participation in demonstrations of various indoor assistive technologies, followed by the completion of a questionnaire. The assistive technology demonstrations encompassed a wide range of devices tailored to support mobility.

Participants Recruitment

Recruitment efforts targeted individuals through advertisements placed across multiple social media platforms and in universities. Interested participants registered their interest by completing an online registration form through Google Forms or by contacting the provided phone number. Eligibility criteria were established to ensure suitable participation, including being aged 60 years or older; not having severe mobility challenges that would prevent full engagement in the experiential session; the ability to attend the demonstration session; and the capacity to provide written informed consent. These criteria ensured that participants could actively engage in the study activities and provide meaningful feedback on the assistive technologies demonstrated.

Study Design

Participants engaged in an experiential session featuring various indoor assistive technologies (given in Figure 2A-J). Each demonstrated device is specifically designed to support older adults and individuals with mobility challenges within their living environments [29]. The session included the following technologies: The SafeFree handling side guard and home care bed (Figure 2A) provide safe and comfortable support for resting, sleeping, and transferring activities within the home. The transfer roller kit board (Figure 2B) assists with safe transfers from beds, chairs, or the floor, enhancing indoor mobility. The tilt and reclining wheelchair (Figure 2C) allows users to adjust their position and posture to perform different activities at home. The shower commode chair (Figure 2D) enables individuals to bathe and use the toilet safely within their living spaces. The commode with a mobile pan (Figure 2E) offers discreet and convenient toileting support indoors. The recliner and stand-up sofa (Figure 2F) incorporate features that aid in sitting, standing, and repositioning for those with mobility challenges. The antislip seating mat (Figure 2G) enhances stability and safety on various seating surfaces within the home. The carbon fiber quad cane (Figure 2H) provides additional support and stability for individuals with limited mobility. The 3-in-1 stand assist walker (Figure 2I) supports users in performing various daily activities safely indoors. Finally, the nonintrusive activity-based sensor monitoring system (Figure 2J) offers remote monitoring and support for individuals in their living spaces.

XSL•FO
RenderX

Figure 2. Assistive technologies demonstrated in the experiential session: (A) handling side guard and home care bed; (B) transfer roller kit board; (C) tilt and reclining wheelchair; (D) shower commode chair; (E) commode with mobile pan; (F) recliner and stand-up sofa; (G) antislip seating mat; (H) carbon fiber quad cane; (I) 3-in-1 stand assist walker; and (J) nonintrusive activity-based sensor monitoring system.



A trained research assistant conducted the demonstrations to ensure consistency and quality throughout the session. Each participant received personalized attention from a single research assistant to maintain fidelity. The research team reviewed and approved all demonstration materials to ensure accuracy and effectiveness. Before the study, the research assistant conducted a trial demonstration for the research team, evaluating the clarity and conveyance of the content. In addition, a research team member conducted random visits to demonstration sessions to verify that the content remained consistent and adhered to the predetermined standards.

Outcome Measures

Participants were assessed after the demonstration session using a questionnaire divided into 5 main sections: positive aging

XSL•FO
RenderX

perception, quality-of-life enhancement, social support, technology acceptance, and technology readiness.

Positive Aging Perception

The Awareness of Age-Related Change (AARC) questionnaire [30] was used to assess participants' positive aging perception. A total of 10 items were selected for this study, using a 5-point Likert scale ranging from 1 (not at all) to 5 (very much). The internal consistency reliability demonstrated strong results for both AARC-Gains (α=.82) and AARC-Losses (α=.78) [31].

Quality-of-Life Enhancement

Quality-of-life enhancement was measured based on Moxley et al [32]. This section comprised 5 items evaluated using a 7-point Likert scale, from 1 (Not at all) to 7 (A lot). The items specifically addressed the extent to which assistive technologies can improve the quality of life for older adults. This scale has been widely adopted in gerontechnology research including studies by Dale et al [33].

Social Support

Social support was assessed using a 5-item scale adapted from Moxley et al [32]. This scale evaluated the level of assistance required from family and friends in becoming proficient with the presented technologies. Responses were measured on a 7-point Likert scale, ranging from 1 (Not at all) to 7 (A lot). The scale's widespread use in gerontechnology research, as demonstrated by its application in Dale et al [33], underscores its validity and reliability.

Technology Acceptance

Technology acceptance was evaluated using the Senior Technology Acceptance Model questionnaire, adapted from Chen and Lou [34] and Venkatesh et al [35]. This instrument included 4 variables: attitudinal beliefs (3 items), control beliefs (4 items), gerontechnology confidence (gerontechnology anxiety; 2 items), and behavioral intention (3 items). Participants responded on a 7-point Likert scale, from 1 (Strongly disagree) to 7 (Strongly agree). The questionnaire demonstrated strong internal consistency, with α values ranging from 0.96 to 0.97 [36]. Its widespread adoption across various fields highlights its effectiveness in assessing users' perceptions and attitudes toward technology adoption and utilization [37].

Technology Readiness

The Technology Readiness Questionnaire (TRQ) developed by Parasuraman and Colby [38] was used to measure participants' readiness and willingness to embrace technology. The TRQ includes 4 dimensions: comfort, innovativeness, optimism, and security, with 4 items assigned to each dimension. Participants rated each item on a 5-point Likert scale, where higher scores indicated a greater inclination to adopt and use technology. The TRQ's extensive application in aging-related research emphasizes its relevance in evaluating individuals' attitudes and preparedness for technology adoption [39].

Analytical Techniques

The measurement model was validated using confirmatory factor analysis (CFA) with SPSS (version 28.0; IBM Corp) and AMOS (version 28; IBM Corp). CFA evaluated the validity of the measurement model by examining the relationships between observed variables and their underlying latent constructs. A total of 8 key goodness-of-fit indices, recommended by Hu and Bentler [40], were used to assess the model fit. These indices included Cronbach α, chi-square and its respective degrees of freedom (χ^2/df), goodness-of-fit index (GFI), incremental fit index (IFI), comparative fit index (CFI), Tucker-Lewis index (TLI), root mean square error of approximation (RMSEA), and parsimony goodness-of-fit index (PGFI).

Descriptive statistics summarized the demographic variables, providing the mean and SD for each measure to offer an overview of the sample characteristics. Correlation analysis was performed to provide a preliminary assessment of the relationships between variables, while structural equation modeling explored the main relationships between predictor and outcome variables. Multivariate analysis of covariance (MANCOVA) was used as a post hoc analysis to scrutinize the effects of predictor variables on outcome variables while controlling for demographic characteristics such as age, gender, education level, and ethnicity. Wilks lambda (λ) was used to assess multivariate effects, followed by univariate analyses to identify specific between-subjects effects. Effect sizes were reported using partial eta squared (η^2). Statistical significance was set at $P<.05$, with marginal significance noted at $P<.10$.

Ethical Considerations

The research protocol was developed by Monash University Malaysia, and study approval was granted by the Monash University Human Research Ethics Committee (project ID: 39857; review reference: 2023-39857-98651) in September 2023. All participants provided written informed consent and received a token of appreciation of RM 50 (approximately US $11.2) for their participation. To ensure participant confidentiality, all collected data will be anonymized. In instances where full anonymization is not possible, strict protective measures will be implemented, including secure data storage and restricted access to authorized personnel only, to safeguard participant information.

Results

Participant Characteristics

A total of 104 older adults participated in the study, with an average mean age of 67.92 (SD 5.68) years. The sample comprised 58.7% (61/104) women and 41.3% (43/104) men. Educational levels varied, including 2.9% (3/104) with primary education and below, 24% (25/104) with secondary education, 27.9% (29/104) holding a diploma or preuniversity qualification, 29.8% (31/104) possessing a degree or professional certification, and 15.4% (16/104) with postgraduate degrees. Most participants were married (73/104, 70.2%), followed by those who were never married (5.8%, 6/104), divorced (9/104, 8.7%), widowed or widowers (10/104, 9.6%), separated (5/104, 4.8%), and others (1/104, 1%). Ethnically, the majority identified as Chinese (86/104, 82.7%), with smaller representations of Indian (16/104, 15.4%), Malay (1/104, 1%), and other ethnicities (1/104, 1%). Regarding living arrangements, 78.8% (82/104) resided with household members, 17.3% (18/104) lived alone, and 3.8% (4/104) had other living situations. Participants reported having

XSL•FO
RenderX

varying numbers of children, with 14.4% (15/104) having none, 14.4% (15/104) having 1 child, 32.7% (34/104) having 2 children, 32.7% (34/104) having 3 children, 1.9% (2/104) having 4 children, and 2.9% (3/104) having 5 children. In terms of employment status, 57.7% (60/104) were private retirees, 11.5% (12/104) were government retirees (including pensioners), 8.7% (9/104) were self-employed, 5.8% (6/104) were private sector employees, 7.7% (8/104) were homemakers, 5.8% (6/104) were unemployed, and 2.9% (3/104) were in other categories. The majority of participants lived in terraced, linked houses, or semidetached homes (60/104, 57.7%), followed by flats, apartments, condominiums, or townhouses (34/104, 32.7%); detached houses (9/104, 8.7%); and other dwelling types (1/104, 1%). Statistical analysis revealed that living arrangements significantly varied within the sample (P=.002), while other demographic variables did not show significant differences (Table 1).

XSL•FO
RenderX

Table 1. Demographic information (N=104).

Demographics	Values	*P* value
Age (years), mean (SD)	67.92 (5.68)	.22
Gender, n (%)		.08
Men	43 (41.3)	
Women	61 (58.7)	
Educational level, n (%)		.97
Primary school and below	3 (2.9)	
Secondary school	25 (24)	
Diploma or preuniversity	29 (27.9)	
Degree or professional	31 (29.8)	
Postgraduate	16 (15.4)	
Marital status, n (%)		.16
Never married	6 (5.8)	
Married	73 (70.2)	
Separated	5 (4.8)	
Divorcee	9 (8.7)	
Widow or widower	10 (9.6)	
Others	1 (1)	
Ethnicity, n (%)		.18
Malay	1(1)	
Chinese	86 (82.7)	
Indian	16 (15.4)	
Others	1 (1)	
Living arrangement, n (%)		.002
Living alone	18 (17.3)	
Living with household members	82 (78.8)	
Others	4 (3.8)	
Number of children, n (%)		.39
0	15 (14.4)	
1	15 (14.4)	
2	34 (32.7)	
3	34 (32.7)	
4	2 (1.9)	
5	3 (2.9)	
Employment status, n (%)		.41
Private sector employee	6 (5.8)	
Self-employed	9 (8.7)	
Government retirees (including pensioners)	12 (11.5)	
Private retiree	60 (57.7)	
Homemaker	8 (7.7)	
Unemployed	6 (5.8)	
Others	3 (2.9)	
Residential dwelling type, n (%)		.18

XSL•FO
RenderX

Demographics	Values	*P* value
Flat, apartment, condominium, or townhouse	34 (32.7)	
Detached house (bungalow or traditional house)	9 (8.7)	
Terrace, link house, or semidetached	60 (57.7)	
Others	1 (1)	

Correlations

The correlation analysis in Table 2 identified several significant relationships among the study variables. Age positively correlated with marital status (r= 0.278; P=.004), indicating that older participants were more likely to be married. Gender exhibited a significant negative correlation with living arrangements (r=–0.309; P=.001), suggesting that women were more likely to live with household members compared with men. Educational level was negatively associated with employment status (r=–0.233; P=.02). Residential dwelling type showed a positive correlation with living arrangements (r=0.397; P<.001) and the number of children (r=0.207; P=.04), indicating that certain housing types were more common among those living with more children. Quality-of-life enhancement was negatively related to age (r=–0.220; P=.02) and educational level (r=–0.201; P=.04). Social support positively correlated with quality-of-life enhancement (r=0.223; P=.02) and negatively correlated with gerontechnology confidence (r=–0.317; P=.001). Behavioral intention was positively associated with quality-of-life enhancement (r=0.246; P=.01), attitudinal beliefs (r=0.275; P=.005), control beliefs (r=0.396; P<.001), gerontechnology confidence (r=0.206; P=.04), optimism (r=0.428; P<.001), and innovativeness (r=0.204; P=.04). These correlations offer preliminary insights on how demographic factors and positive perceptions of aging, quality of life, and social support influence technology acceptance and readiness among older adults.

XSL•FO
RenderX

Table 2. Correlation matrix.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Age (mean 67.92, SD 5.68)																				
r	1	–.063	–.131	.278	.112	–.135	.199	–.014	.019	–.020	–.220	.01	–.129	.195	.042	–.068	.063	.132	.089	–.034
P value	—[a]	.52	.18	.004	.26	.17	.04	.89	.84	.84	.02	.92	.19	.047	.67	.49	.52	.18	.37	.73
2. Gender[b] (mean 1.59, SD 0.495)																				
r	–.063	1	.003	.14	.131	–.309	–.085	.08	–.133	.085	–.143	.096	.025	.035	–.045	–.010	–.050	–.189	–.099	–.150
P value	.52	—	.97	.16	.18	.001	.39	.42	.18	.39	.15	.33	.80	.72	.65	.92	.62	.06	.32	.13
3. Educational level[c] (mean 3.31, SD 1.089)																				
r	–.131	.003	1	–.162	–.140	–.045	–.137	–.233	.059	.178	–.201	–.289	.125	.114	.406	.088	.011	.025	.051	.095
P value	.18	.97	—	.101	.16	.65	.17	.02	.55	.07	.04	.003	.21	.25	<.001	.38	.91	.80	.61	.34
4. Marital status[d] (mean 2.49, SD 1.115)																				
r	.278	.14	–.162	1	.042	–.089	.107	.117	.007	.159	–.034	.025	.103	.169	.025	.008	–.045	.277	–.002	.048
P value	.004	.16	.101	—	.67	.37	.28	.24	.95	.11	.73	.80	.30	.09	.80	.94	.65	.004	.98	.63
5. Ethnicity[e] (mean 2.19, SD 0.609)																				
r	.112	.131	–.140	.042	1	.125	.046	.019	.128	–.043	.269	.260	.054	.146	–.030	.214	.268	.043	.098	–.092
P value	.26	.18	.16	.67	—	.21	.64	.85	.19	.66	.006	.008	.58	.14	.76	.03	.006	.66	.32	.36
6. Living arrangement[f] (mean 1.9, SD 0.566)																				
r	–.135	–.309	–.045	–.089	.125	1	.129	–.002	.397	–.060	.079	–.078	.065	.02	.182	.07	.16	.024	.195	.244
P value	.17	.001	.65	.37	.21	—	.19	.99	<.001	.55	.42	.43	.51	.84	.06	.48	.11	.81	.047	.01
7. Number of children (mean 2.02, SD 1.188)																				
r	.199	–.085	–.137	.107	.046	.129	1	–.111	.207	.043	–.001	.026	–.135	.052	.004	.011	.085	.031	–.077	–.055
P value	.04	.39	.17	.28	.64	.19	—	.26	.04	.66	.99	.80	.17	.60	.97	.91	.39	.75	.44	.58
8. Employment status[g] (mean 5.67, SD 1.517)																				
r	–.014	.08	–.233	.117	.019	–.002	–.111	1	.01	–.017	.049	.002	–.105	–.108	.01	–.062	–.075	.152	.067	.031
P value	.89	.42	.02	.24	.85	.99	.26	—	.92	.86	.62	.98	.29	.28	.92	.53	.45	.12	.50	.75
9. Residential dwelling type[h] (mean 2.28, SD 0.96)																				
r	.019	–.133	.059	.007	.128	.397	.207	.01	1	–.016	–.010	–.162	–.064	–.088	.093	.036	.105	.034	.104	.179
P value	.84	.18	.55	.95	.19	<.001	.04	.92	—	.87	.92	.101	.52	.38	.35	.71	.29	.73	.29	.07
10. Positive ageing perception (mean 3.959, SD 0.404)																				
r	–.020	.085	.178	.159	–.043	–.060	.043	–.017	–.016	1	–.077	–.090	.192	.053	.202	–.082	–.008	.073	.094	–.008
P value	.84	.39	.07	.11	.66	.55	.66	.86	.87	—	.44	.36	.051	.59	.04	.41	.93	.46	.34	.94
11. Quality-of-life enhancement (mean 4.687, SD 1.356)																				
r	–.220	–.143	–.201	–.034	.269	.079	–.001	.049	–.010	–.077	1	.223	.082	.175	–.061	.246	.201	.161	.072	.19
P value	.02	.15	.04	.73	.006	.42	.99	.62	.92	.44	—	.02	.41	.07	.54	.01	.04	.10	.47	.054
12. Social support (mean 3.483, SD 1.295)																				

XSL•FO
RenderX

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
r	.01	.096	–.289	.025	.260	–.078	.026	.002	–.162	–.090	.223	1	.16	.041	–.317	–.093	–.123	–.109	–.150	–.232
P value	.92	.33	.003	.80	.008	.43	.80	.98	.101	.36	.02	—	.104	.68	.001	.35	.21	.27	.13	.02
13. Attitudinal beliefs (mean 5.401, SD 1.293)																				
r	–.129	.025	.125	.103	.054	.065	–.135	–.105	–.064	.192	.082	.16	1	.276	.177	.275	.281	.129	.102	.152
P value	.19	.80	.21	.30	.58	.51	.17	.29	.52	.051	.41	.104	—	.005	.07	.005	.004	.19	.30	.12
14. Control beliefs (mean 5.337, SD 0.931)																				
r	.195	.035	.114	.169	.146	.02	.052	–.108	–.088	.053	.175	.041	.276	1	.201	.396	.177	.341	.022	.1
P value	.047	.72	.25	.09	.14	.84	.60	.28	.38	.59	.08	.68	.005	—	.04	<.001	.07	<.001	.82	.31
15. Gerontechnology confidence (mean 4.88, SD 1.471)																				
r	.042	–.045	.406	.025	–.080	.182	.004	.01	.093	.202	–.061	–.317	.177	.201	1	.206	.157	.003	.408	.335
P value	.67	.65	<.001	.80	.76	.06	.97	.92	.35	.04	.54	.001	.07	.04	—	.04	.11	.98	<.001	.001
16. Behavioral intention (mean 5.804, SD 0.786)																				
r	–.068	–.010	.088	.008	.214	.07	.011	–.062	.036	–.082	.246	–.093	.275	.396	.206	1	.428	.204	.076	.07
P value	.49	.92	.38	.94	.03	.48	.91	.53	.71	.41	.01	.35	.005	<.001	.04	—	<.001	.04	.44	.48
17. Optimism (mean 4.276, SD 0.515)																				
r	.063	–.050	.011	–.045	.268	.16	.085	–.075	.105	–.008	.201	–.123	.281	.177	.157	.428	1	.269	.145	.103
P value	.52	.62	.91	.65	.006	.11	.39	.45	.29	.93	.04	.21	.004	.07	.11	<.001	—	.006	.14	.30
18. Innovativeness (mean 3.26, SD 0.637)																				
r	.132	–.189	.025	.277	.043	.024	.031	.152	.034	.073	.161	–.109	.129	.341	.003	.204	.269	1	.173	.053
P value	.18	.06	.80	.004	.66	.81	.75	.12	.73	.46	.102	.27	.19	<.001	.97	.04	.006	—	.08	.59
19. Comfort (mean 3.202, SD 0.626)																				
r	.089	–.099	.051	–.002	.098	.195	–.077	.067	.104	.094	.072	–.150	.102	.022	.408	.076	.145	.173	1	.404
P value	.37	.32	.61	.98	.32	.047	.44	.50	.29	.34	.47	.13	.30	.82	<.001	.44	.14	.08	—	<.001
20. Security (mean 2.752, SD 0.746)																				
r	–.034	–.150	.095	.048	–.092	.244	–.065	.031	.179	–.008	.19	–.232	.152	.1	.335	.07	.103	.053	.404	1
P value	.73	.13	.34	.63	.36	.01	.58	.75	.07	.94	.054	.02	.12	.31	.001	.48	.30	.59	<.001	—

[a]Not applicable.

[b]Gender: 1=men and 2=women.

[c]Educational level: 1=primary school and below, 2=secondary school, 3=diploma or preuniversity, 4=degree or professional, and 5=postgraduate.

[d]Marital status: 1=never married, 2=married, 3=separated, 4=divorcee, 5=widow or widower, and 6=others.

[e]Ethnicity: 1=Malay, 2=Chinese, 3=Indian, 4=others.

[f]Living arrangement: 1=living alone, 2=living with household members, and 3=others.

[g]Employment status: 1=private sector employee, 2=self-employed, 3=government retiree (includes pensioners), 4=private retiree, 5=homemaker, 6=unemployed, and 7=others.

[h]Residential dwelling type: 1=flat, apartment, condominium, or townhouse; 2=detached house (bungalow or traditional house); 3=terrace, link house, or semidetached; 4=others.

Measurement Model

The measurement model was validated via CFA. The Cronbach α values for positive aging perception, quality-of-life enhancement, social support, technology acceptance, and technology readiness were 0.736, 0.855, 0.815, 0.740, and 0.731, respectively, demonstrating acceptable internal consistency.

XSL•FO
RenderX

Model fit indices were as follows: χ^2_{752}=1069.140, *P*<.001; GFI=0.703; IFI=0.858; CFI=0.850; TLI=0.829; RMSEA=0.064; and PGFI=0.585. These values met the recommended thresholds, indicating a moderately acceptable model fit [40].

Structural Model

The structural model was tested via structural equation modeling (Table 3), illustrating how positive aging perception, quality-of-life enhancement, and social support relate to various aspects of technology acceptance and readiness among older adults. Positive aging perception significantly enhanced gerontechnology confidence (β=.462; *P*<.001) and comfort (β=.323; *P*=.047) and exhibited a marginally significant positive relationship with control beliefs (β=.228; *P*=.095). In contrast, quality-of-life enhancement and social support did not show significant effects on any of the technology acceptance or readiness components. These findings suggest that positive aging perception plays a crucial role in fostering confidence and comfort with gerontechnology, whereas quality-of-life enhancement and social support have limited impacts on technology acceptance and readiness within this population.

Table 3. Structural equation modeling results.

Path	Standardized coefficient (β)	Unstandardized coefficient (B)	SE	*t* value	*P* value
Positive aging perception					
Attitudinal beliefs	.143	0.173	0.145	1.196	.23
Control beliefs	.228	0.242	0.145	1.670	.095
Gerontechnology confidence	.462	1.220	0.321	3.806	<.001
Behavioral intention	.190	0.239	0.146	1.635	.102
Optimism	.041	0.027	0.068	0.400	.69
Innovativeness	.061	0.042	0.094	0.453	.65
Comfort	.323	0.227	0.114	1.984	.047
Security	.171	0.082	0.067	1.219	.22
Quality-of-life enhancement					
Attitudinal beliefs	.097	0.932	1.687	0.552	.58
Control beliefs	.031	0.262	1.012	0.259	.80
Gerontechnology confidence	–.104	–2.177	3.615	–0.602	.55
Behavioral intention	.266	2.652	3.973	0.667	.50
Optimism	.165	0.878	1.421	0.618	.54
Innovativeness	.149	0.829	1.382	0.600	.55
Comfort	.160	0.896	1.475	0.607	.54
Security	.150	0.572	0.953	0.600	.55
Social support					
Attitudinal beliefs	.186	0.484	0.398	1.217	.22
Control beliefs	.142	0.323	0.323	1.002	.32
Gerontechnology confidence	–.215	–1.220	0.858	–1.422	.16
Behavioral intention	–.124	–0.333	0.336	–.993	.32
Optimism	–.192	–0.277	0.245	–1.130	.26
Innovativeness	–.176	–0.264	0.242	–1.090	.28
Comfort	–.289	–0.436	0.329	–1.323	.19
Security	–.357	–0.367	0.279	–1.314	.19

Post hoc Observations

The MANCOVA assessed the combined effects of independent variables and demographic factors on technology acceptance and readiness among older adults (Tables 4 and 5). The analysis revealed that positive aging perception did not significantly influence the various aspects of technology acceptance and readiness when controlling for demographic characteristics (Wilks λ=0.897, $F_{8,61}$=0.877, *P*=.54, partial η^2=0.103). This consistency reinforces the structural model findings, indicating that the relationships between positive aging perception and factors such as control beliefs, comfort, and gerontechnology confidence remain stable across different demographic groups. Conversely, quality-of-life enhancement (Wilks λ=0.737, $F_{8,61}$=2.717, *P*=.01, partial η^2=0.263) and social support (Wilks

XSL•FO
RenderX

λ=0.762, $F_{8,61}$=2.387, *P*=.03, partial η²=0.238) significantly impacted technology acceptance and readiness. A closer examination showed that both quality-of-life enhancement (*P*=.001) and social support (*P*=.008) negatively influenced security perceptions, suggesting that improvements in these areas are associated with reduced concerns about the security of assistive technologies. In addition, demographic factors played a notable role: educational level significantly predicted gerontechnology confidence (β=29.548, *P*=.004, partial η²=0.201) and ethnicity was a significant predictor of optimism (β=3.373, *P*=.003, partial η²=0.187). These findings indicate that while positive aging perceptions consistently affect certain aspects of technology acceptance, quality-of-life enhancements and social support are crucial in shaping specific perceptions such as security. Furthermore, educational attainment and ethnic background influence confidence and optimism regarding technology use. These granular insights highlight the importance of addressing both demographic and psychological factors to foster effective adoption and readiness for indoor assistive technologies among older adults. Further explanation is provided in the Discussion section.

Table 4. Multivariate analysis of covariance on the effects of predictor variables while controlling for demographic characteristics.

Variables	Wilks λ	*F* value (*df*)	*P* value	Partial η²
Predictor variables				
Positive aging perception	0.897	0.877 (8, 61)	.54	0.103
Quality-of-life enhancement	0.737	2.717 (8, 61)	.01	0.263
Social support	0.762	2.387 (8, 61)	.03	0.238
Demographic characteristics				
Age	0.770	1.066 (8, 61)	.39	0.123
Gender	0.920	0.663 (8, 61)	.72	0.080
Educational level	0.498	1.471 (8, 61)	.06	0.160
Marital status	0.584	0.883 (8, 61)	.67	0.102
Ethnicity	0.571	1.574 (8, 61)	.051	0.170
Living arrangement	0.791	0.947 (8, 61)	.52	0.110
Number of children	0.642	0.718 (8, 61)	.90	0.085
Employment status	0.519	0.904 (8, 61)	.66	0.104
Residential dwelling type	0.753	0.759 (8, 61)	.78	0.090
Significant between-subjects path effects (*F* >3.00, *P* <.05)				
Quality-of-life enhancement: security	6.404	12.913 (1, 68)	.001	0.160
Social support: security	3.674	7.408 (1, 68)	.008	0.098
Educational level: gerontechnology confidence	29.548	4.266 (4, 68)	.004	0.201
Ethnicity: optimism	3.373	5.198 (3, 68)	.003	0.187

Table 5. Multiple regression analyses examining the effects of predictor variables and demographic characteristics on technology acceptance and technology readiness outcomes.

Outcome variables	R^2	Adjusted R^2	*F* value (*df*)	*P* value
Overall model effects (predictor variables+demographic characteristics=outcomes)				
Attitudinal beliefs	0.439	0.158	1.562 (7, 95)	.06
Control beliefs	0.398	0.097	1.321 (7, 95)	.16
Gerontechnology confidence	0.471	0.207	1.783 (7, 95)	.02
Behavioral intention	0.303	–0.045	0.870 (7, 95)	.67
Optimism	0.441	0.161	1.576 (7, 95)	.06
Innovativeness	0.348	0.021	1.066 (7, 95)	.40
Comfort	0.337	0.005	1.015 (7, 95)	.47
Security	0.401	0.101	1.339 (7, 95)	.15

XSL•FO
RenderX

Discussion

Principal Findings

This study explored the relationships between positive aging perception, quality-of-life enhancement, social support, and various factors related to technology acceptance and readiness among older adults.

The preliminary analysis using correlation analysis revealed significant associations between demographic factors and key variables, such as age being positively related to marital status and negatively associated with quality-of-life enhancement and gerontechnology confidence; educational level being negatively correlated with living arrangements and gerontechnology confidence; and social support being positively correlated with quality-of-life enhancement. These insights suggest that the older adults who participated in this study predominantly consisted of married individuals with varying educational backgrounds, most of whom lived with household members and received substantial social support, which in turn influenced their perceptions and confidence regarding assistive technologies.

The primary analysis using structural equation modeling indicated that a positive perception of aging significantly enhances gerontechnology confidence and comfort, with a marginal positive relationship with control beliefs. These findings suggest that older adults who maintain a positive outlook on aging are more confident and comfortable with using gerontechnology, which aligns with existing literature emphasizing the role of positive aging attitudes in technology adoption [41,42]. However, quality-of-life enhancement and social support did not exhibit significant direct effects on most components of technology acceptance and readiness within the structural model. This contrasts with some previous studies that highlighted the importance of these factors in technology adoption [43,44], indicating that their influence may be more complex or mediated by other variables not captured in this model.

The post hoc analysis using MANCOVA further revealed that quality-of-life enhancement and social support significantly impact technology acceptance and readiness when controlling for demographic characteristics. Specifically, both quality-of-life enhancement and social support were found to negatively influence security perceptions, suggesting that as these factors increase, concerns about the security of assistive technologies decrease. This negative influence on security perceptions could be explained by demographic factors such as educational level and ethnicity. Educational level significantly predicted gerontechnology confidence, indicating that individuals with higher education may feel more confident in using technology, thereby reducing their security concerns. Similarly, ethnicity emerged as a significant predictor of optimism, suggesting that cultural or social backgrounds influence positive expectations toward technology, which can also alleviate security apprehensions. These findings highlight that while quality-of-life enhancements and social support do not directly influence the larger set of technology acceptance factors, they play a crucial role in alleviating security-related concerns, which are vital for the overall readiness to adopt assistive technologies.

Interestingly, social support showed negative associations with gerontechnology confidence and security perceptions in the structural model, although these were not statistically significant. This counterintuitive finding, if significant, may indicate that excessive reliance on social support could undermine individuals' confidence in using technology independently or heighten concerns about privacy and security. Such dynamics, which contrast past observations [45-48], warrant further investigation to understand the complex role of social support in technology adoption among older adults. One possible explanation is the influence of social norms, where older adults may experience expectations or pressures from their social networks that discourage full technological engagement [49]. An overreliance on assistance from others might also reduce the perceived need for technology, leading to decreased confidence and increased security concerns. Privacy and trust issues related to technology could also contribute to these negative relationships, as older adults may fear privacy breaches or distrust the reliability of assistive technologies [50,51].

The significant influence of demographic characteristics, particularly educational level and ethnicity, emphasizes the importance of considering sociodemographic factors in technology acceptance models. Higher educational levels were associated with greater confidence in using gerontechnology, likely due to better technological literacy and problem-solving skills acquired through education. Ethnic background influencing optimism suggests that cultural factors play a role in shaping positive attitudes toward technology adoption. These findings extend previous research [43,52] by quantifying the impact of these demographic factors and highlighting their specific effects on different aspects of technology acceptance.

Contrary to initial expectations, this study did not find significant associations between positive aging perception and several factors, including attitudinal beliefs, behavioral intention, innovativeness, optimism, and security, whereas quality-of-life enhancement and social support did not produce significant relationships. This may suggest that older adults' attitudes toward indoor assistive technologies may be more complex or mediated by other factors such as perceived need, technology complexity, or previous experience with similar devices that were not considered in this study [19,49]. While the lack of significant associations may seem surprising, it is important to acknowledge that the adoption and acceptance of assistive technologies among older adults are influenced by multifaceted factors [53]. The variables examined herein might be influenced by additional variables or interact with each other in ways that were not captured in this particular study. Future studies could explore other variables or contextual factors, such as cognitive abilities, physical health, technological literacy, or specific characteristics of the assistive technologies themselves, to develop more comprehensive models of technology acceptance among older adults [54].

Implications and Recommendations

This study underscores the importance of fostering positive aging perceptions to enhance gerontechnology confidence and

XSL•FO
RenderX

comfort among older adults. To achieve this, stakeholders should develop and implement educational programs and workshops that clearly demonstrate the benefits and practical applications of assistive technologies. These initiatives can help older adults build a more optimistic outlook on aging and increase their confidence in using these devices effectively.

Integrating quality-of-life enhancements and social support into the design and implementation of assistive technologies is crucial for addressing security concerns. Assistive devices should incorporate features that directly improve users' daily living experiences, such as intuitive user interfaces and robust safety mechanisms. These enhancements can help reduce apprehensions about the reliability and security of the technologies, making them more acceptable and trustworthy for older adults.

Demographic factors, including educational level and ethnicity, significantly influence technology confidence and optimism. To accommodate varying educational backgrounds, assistive technology solutions should offer clear instructions, accessible interfaces, and comprehensive training materials. Technologies should also be culturally tailored to respect and integrate the diverse cultural preferences of different ethnic groups. This cultural sensitivity can promote greater acceptance and positive attitudes toward the use of assistive technologies among ethnically diverse populations.

The negative associations observed between social support and gerontechnology confidence and security perceptions indicate that excessive reliance on social networks might undermine independent technology use and heighten security concerns. To address this, support systems should be designed to empower older adults to use assistive technologies independently. Training programs for caregivers and family members should focus on encouraging autonomous use of technology rather than providing constant assistance. Addressing privacy and trust issues through transparent data handling practices and robust security features should further help alleviate security concerns, thereby enhancing older adults' readiness to adopt assistive technologies.

To this end, enhancing technology acceptance and readiness among older adults requires a comprehensive approach that includes promoting positive aging attitudes, improving quality of life through targeted technology features, tailoring solutions to meet diverse educational and cultural backgrounds, and structuring social support to foster independence. Implementing these strategies should facilitate the effective adoption of indoor assistive technologies, ultimately improving the independence and well-being of older adults.

Limitations and Future Directions

This study presents several limitations that should be considered when interpreting the findings. First, the sample comprised older adults aged 60 years and older who participated in an indoor setting, predominantly Chinese and married, with most living with household members. While this demographic reflects typical patterns in urban Malaysian environments, it may limit the generalizability of the results to wider populations, including those in rural areas or living in assisted living facilities. Future research should use stratified sampling techniques to include a more diverse range of settings and demographics, enhancing the applicability of the findings across different contexts.

Second, the cross-sectional design of the study restricts the ability to establish causal relationships between variables. The associations identified provide a snapshot at a specific point in time but do not account for changes and developments over time. Longitudinal studies or experimental designs would be beneficial in examining the causal effects and temporal dynamics between positive aging perception, quality-of-life enhancement, social support, and technology acceptance and readiness.

Third, reliance on self-report measures to assess constructs such as positive aging perception, quality-of-life enhancement, social support, and technology acceptance introduces potential biases, including social desirability and recall bias. These biases may affect the accuracy of the data collected. Future studies could incorporate objective measures or combine self-report instruments with other assessment methods to achieve a more comprehensive understanding of these constructs.

Fourth, this study focused on specific variables—positive aging perception, quality-of-life enhancement, and social support—to explore their influence on technology acceptance and readiness. Other relevant factors, such as cognitive abilities, physical health, technological literacy, and specific characteristics of assistive technologies, were not included in the analysis. Incorporating a wider range of variables in future research would provide a more comprehensive view of the factors influencing older adults' technology acceptance and readiness. Using a multidimensional framework could capture the complex interactions among various factors more effectively.

Finally, the sample primarily consisted of Malaysian participants living in urban areas, which limits the diversity and generalizability of the findings to other cultural or demographic groups. Future studies should endeavor to include a more varied sample to examine potential cultural or contextual differences in the relationships between the variables. Using cross-cultural research designs or multisite studies can help identify and account for cultural and contextual variations, thereby enhancing the relevance and applicability of the findings across different populations.

Conclusion

Understanding the factors that influence older adults' acceptance and readiness to adopt indoor assistive technologies is essential for enhancing their independence and well-being. This study demonstrates that a positive perception of aging significantly increases gerontechnology confidence and comfort among older adults, while quality-of-life enhancements and social support play crucial roles in reducing security concerns related to technology use. Demographic factors, particularly educational level and ethnicity, also significantly influence confidence and optimism toward technology adoption. These insights highlight the need for targeted strategies to effectively promote the adoption of assistive technologies. Future research should explore the underlying mechanisms that drive these relationships and develop customized interventions to support successful technology integration among older adults.

XSL•FO
RenderX

Acknowledgments

This study was supported by the AGELESS research program funded by the Malaysian Ministry of Higher Education Long Term Research Grant Scheme (LRGS/1/2019/UM/01/1/1, LR005-2019, GOV-000047). We also extend our gratitude to Asian Integrated Medical Sdn Bhd for the loan of assistive technologies, which were used solely for demonstration purposes within the scope of this study.

Conflicts of Interest

None declared.

References

1. Freiberger E, Sieber CC, Kob R. Mobility in older community-dwelling persons: a narrative review. Front Physiol 2020;11:881 [FREE Full text] [doi: 10.3389/fphys.2020.00881] [Medline: 33041836]
2. Newman AB. The epidemiology and societal impact of aging-related functional limitations: a looming public health crisis. J Gerontol A Biol Sci Med Sci 2023;78(Suppl 1):4-7 [FREE Full text] [doi: 10.1093/gerona/glad021] [Medline: 37325965]
3. Leeson GW. The growth, ageing and urbanisation of our world. J Popul Ageing 2018;11:107-115. [doi: 10.1007/s12062-018-9225-7]
4. Moxham LJ, Stutchbury TK, Spinks G, de Vet E, Ikutegbe V, Traynor V, et al. Understanding the assistive technology needs of people over 55: the future of mobility aids. Australas J Ageing 2019;38(4):e127-e134. [doi: 10.1111/ajag.12648] [Medline: 30950196]
5. Brustio PR, Mulasso A, D'Emanuele S, Zia G, Feletti L, Del Signore S, et al. Indoor mobility, frailty, and disability in community-dwelling older adults: a mediation model. Int J Environ Res Public Health 2022;19(18):11386 [FREE Full text] [doi: 10.3390/ijerph191811386] [Medline: 36141659]
6. Kelsey JL, Berry SD, Procter-Gray E, Quach L, Nguyen USDT, Li W, et al. Indoor and outdoor falls in older adults are different: the maintenance of balance, independent living, intellect, and zest in the elderly of boston study. J Am Geriatr Soc 2010;58(11):2135-2141 [FREE Full text] [doi: 10.1111/j.1532-5415.2010.03062.x] [Medline: 20831726]
7. Hashidate H, Shimada H, Shiomi T, Shibata M, Sawada K, Sasamoto N. Measuring indoor life-space mobility at home in older adults with difficulty to perform outdoor activities. J Geriatr Phys Ther 2013;36(3):109-114. [doi: 10.1519/JPT.0b013e31826e7d33] [Medline: 22976813]
8. Carver J, Ganus A, Ivey JM, Plummer T, Eubank A. The impact of mobility assistive technology devices on participation for individuals with disabilities. Disabil Rehabil Assist Technol 2016;11(6):468-477. [doi: 10.3109/17483107.2015.1027295] [Medline: 25815679]
9. Kabacińska K, Vu K, Tam M, Edwards O, Miller WC, Robillard JM. "Functioning better is doing better": older adults' priorities for the evaluation of assistive technology. Assist Technol 2023;35(4):367-373 [FREE Full text] [doi: 10.1080/10400435.2022.2113180] [Medline: 35972791]
10. Cook AM, Polgar JM. Assistive Technologies: Principles and Practice. London, UK: Elsevier Health Sciences; 2019.
11. Fiorini L, de Mul M, Fabbricotti I, Limosani R, Vitanza A, D'Onofrio G, et al. Assistive robots to improve the independent living of older persons: results from a needs study. Disabil Rehabil Assist Technol 2021;16(1):92-102. [doi: 10.1080/17483107.2019.1642392] [Medline: 31329000]
12. Löfqvist C, Nygren C, Brandt ?, Oswald F, Iwarsson S. Use of mobility devices and changes over 12 months among very old people in five European countries. Aging Clin Exp Res 2007;19(6):497-505. [doi: 10.1007/BF03324737] [Medline: 18172373]
13. Agree EM, Freedman VA. A quality-of-life scale for assistive technology: results of a pilot study of aging and technology. Phys Ther 2011;91(12):1780-1788 [FREE Full text] [doi: 10.2522/ptj.20100375] [Medline: 22003159]
14. Matlabi H, Parker SG, McKee K. The contribution of home-based technology to older people's quality of life in extra care housing. BMC Geriatr 2011;11:68 [FREE Full text] [doi: 10.1186/1471-2318-11-68] [Medline: 22040111]
15. Marasinghe KM. Assistive technologies in reducing caregiver burden among informal caregivers of older adults: a systematic review. Disabil Rehabil Assist Technol 2016;11(5):353-360. [doi: 10.3109/17483107.2015.1087061] [Medline: 26371519]
16. Damant J, Knapp M, Watters S, Freddolino P, Ellis M, King D. The impact of ICT services on perceptions of the quality of life of older people. J Assist Technol 2013;7(1):21. [doi: 10.1108/17549451311313183]
17. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
18. Chen K, Chan AHS. Gerontechnology acceptance by elderly Hong Kong Chinese: a senior technology acceptance model (STAM). Ergonomics 2014;57(5):635-652. [doi: 10.1080/00140139.2014.895855] [Medline: 24655221]
19. Peek STM, Wouters EJM, van Hoof J, Luijkx KG, Boeije HR, Vrijhoef HJM. Factors influencing acceptance of technology for aging in place: a systematic review. Int J Med Inform 2014;83(4):235-248 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.01.004] [Medline: 24529817]

XSL•FO
RenderX

20. Wu J, Song S. Older adults' online shopping continuance intentions: applying the technology acceptance model and the theory of planned behavior. International Journal of Human–Computer Interaction 2020;37(10):938-948. [doi: 10.1080/10447318.2020.1861419]
21. Orellano-Colón EM, Harrison-Cruz S, López-Lugo E, Ramos-Peraza S, Meléndez-Ortiz A, Ortiz-Torres J, et al. Assistive technology self-management intervention for older Hispanics: a feasibility study. Disabil Rehabil Assist Technol 2020;15(8):862-870 [FREE Full text] [doi: 10.1080/17483107.2019.1621954] [Medline: 31172826]
22. Wang J, Fu Y, Lou V, Tan SY, Chui E. A systematic review of factors influencing attitudes towards and intention to use the long-distance caregiving technologies for older adults. Int J Med Inform 2021;153:104536. [doi: 10.1016/j.ijmedinf.2021.104536] [Medline: 34325206]
23. Beyer AK, Wolff JK, Freiberger E, Wurm S. Are self-perceptions of ageing modifiable? Examination of an exercise programme with vs. without a self-perceptions of ageing-intervention for older adults. Psychol Health 2019;34(6):661-676. [doi: 10.1080/08870446.2018.1556273] [Medline: 30628483]
24. Gitlin LN, Winter L, Dennis MP, Corcoran M, Schinfeld S, Hauck WW. A randomized trial of a multicomponent home intervention to reduce functional difficulties in older adults. J Am Geriatr Soc 2006;54(5):809-816 [FREE Full text] [doi: 10.1111/j.1532-5415.2006.00703.x] [Medline: 16696748]
25. Beyer S, Perry J. Promoting independence through the use of assistive technology. Tizard Learn Disabil Rev 2013;18:179-185. [doi: 10.1108/tldr-03-2013-0028]
26. Demiris G, Hensel BK, Skubic M, Rantz M. Senior residents' perceived need of and preferences for "smart home" sensor technologies. Int J Technol Assess Health Care 2008;24(1):120-124. [doi: 10.1017/S0266462307080154] [Medline: 18218177]
27. Liu L, Stroulia E, Nikolaidis I, Miguel-Cruz A, Rios Rincon A. Smart homes and home health monitoring technologies for older adults: a systematic review. Int J Med Inform 2016;91:44-59. [doi: 10.1016/j.ijmedinf.2016.04.007] [Medline: 27185508]
28. Mitzner TL, Boron JB, Fausset CB, Adams AE, Charness N, Czaja SJ, et al. Older adults talk technology: technology usage and attitudes. Comput Human Behav 2010;26(6):1710-1721 [FREE Full text] [doi: 10.1016/j.chb.2010.06.020] [Medline: 20967133]
29. Mann WC, Helal S. Smart Technology for Aging, Disability, and Independence: The State of the Science. Hoboken, NJ: Wiley; 2006.
30. Kaspar R, Gabrian M, Brothers A, Wahl HW, Diehl M. Measuring awareness of age-related change: development of a 10-item short form for use in large-scale surveys. Gerontologist 2019;59(3):e130-e140 [FREE Full text] [doi: 10.1093/geront/gnx213] [Medline: 29401221]
31. Nehrkorn-Bailey AM, Rodriguez D, Forsyth G, Braun B, Burke K, Diehl M. Change in views of aging, physical activity, and physical health over 8 weeks: results from a randomized study. J Aging Phys Act 2023;31(4):666-678 [FREE Full text] [doi: 10.1123/japa.2022-0133] [Medline: 36708712]
32. Moxley J, Sharit J, Czaja SJ. The factors influencing older adults' decisions surrounding adoption of technology: quantitative experimental study. JMIR Aging 2022;5(4):e39890 [FREE Full text] [doi: 10.2196/39890] [Medline: 36416885]
33. Dale J, Nanton V, Day T, Apenteng P, Bernstein CJ, Grason Smith G, et al. Uptake and use of care companion, a web-based information resource for supporting informal carers of older people: mixed methods study. JMIR Aging 2023;6:e41185 [FREE Full text] [doi: 10.2196/41185] [Medline: 37733406]
34. Chen K, Lou VWQ. Measuring senior technology acceptance: development of a brief, 14-item scale. Innov Aging 2020;4(3):igaa016 [FREE Full text] [doi: 10.1093/geroni/igaa016] [Medline: 32617418]
35. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478. [doi: 10.2307/30036540]
36. Goh TT, Dai X, Yang Y. Benchmarking ChatGPT for prototyping theories: experimental studies using the technology acceptance model. BenchCouncil Trans Benchmarks Stand Eval 2023;3(4):100153. [doi: 10.1016/j.tbench.2024.100153]
37. Kwan RYC, Yeung JWY, Lee JLC, Lou VWQ. The association of technology acceptance and physical activity on frailty in older adults during the COVID-19 pandemic period. Eur Rev Aging Phys Act 2023;20(1):24 [FREE Full text] [doi: 10.1186/s11556-023-00334-3] [Medline: 38114901]
38. Parasuraman A, Colby CL. An updated and streamlined technology readiness index: TRI 2.0. J Serv Res 2014;18(1):59-74. [doi: 10.1177/1094670514539730]
39. Kim S, Chow BC, Park S, Liu H. The usage of digital health technology among older adults in Hong Kong and the role of technology readiness and eHealth literacy: path analysis. J Med Internet Res 2023;25:e41915 [FREE Full text] [doi: 10.2196/41915] [Medline: 37043274]
40. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999;6(1):1-55 [FREE Full text] [doi: 10.1080/10705519909540118]
41. Lê Q, Nguyen HB, Barnett T. Smart homes for older people: positive aging in a digital world. Future Internet 2012;4(2):607-617. [doi: 10.3390/fi4020607]
42. Offermann J, Wilkowska W, Ziefle M. Interplay of perceptions of aging, care, and technology acceptance in older age. Int J Hum Comput 2022;39(5):1003-1015 [FREE Full text] [doi: 10.1080/10447318.2022.2050542]

XSL•FO
RenderX

43. Shin HR, Um SR, Yoon HJ, Choi EY, Shin WC, Lee HY, et al. Comprehensive senior technology acceptance model of daily living assistive technology for older adults with frailty: cross-sectional study. J Med Internet Res 2023;25:e41935 [FREE Full text] [doi: 10.2196/41935] [Medline: 37036760]
44. Delbreil E, Zvobgo G. Wireless sensor technology in dementia care: caregiver perceptions, technology take ‐ up and business model innovation. EuroMed J Bus 2013;8(1):79-97. [doi: 10.1108/emjb-05-2013-0019]
45. Huang G, Oteng SA. Gerontechnology for better elderly care and life quality: a systematic literature review. Eur J Ageing 2023;20(1):27 [FREE Full text] [doi: 10.1007/s10433-023-00776-9] [Medline: 37347277]
46. Aggar C, Sorwar G, Seton C, Penman O, Ward A. Smart home technology to support older people's quality of life: a longitudinal pilot study. Int J Older People Nurs 2023;18(1):e12489 [FREE Full text] [doi: 10.1111/opn.12489] [Medline: 35785517]
47. Gitlin LN. Assistive technology in the home and community for older people: psychological and social considerations. In: Scherer MJ, editor. Assistive Technology: Matching Device and Consumer for Successful Rehabilitation. Washington, DC: American Psychological Association; 2002.
48. Wei W, Gong X, Li J, Tian K, Xing K. A study on community older people's willingness to use smart home-an extended technology acceptance model with intergenerational relationships. Front Public Health 2023;11:1139667 [FREE Full text] [doi: 10.3389/fpubh.2023.1139667] [Medline: 37346111]
49. Tsertsidis A, Kolkowska E, Hedström K. Factors influencing seniors' acceptance of technology for ageing in place in the post-implementation stage: a literature review. Int J Med Inform 2019;129:324-333 [FREE Full text] [doi: 10.1016/j.ijmedinf.2019.06.027] [Medline: 31445274]
50. Bradford DK, Kasteren YV, Zhang Q, Karunanithi M. Watching over me: positive, negative and neutral perceptions of in-home monitoring held by independent-living older residents in an Australian pilot study. Ageing Soc 2017;38(7):1377-1398. [doi: 10.1017/s0144686x1700006x]
51. Schomakers EM, Biermann H, Ziefle M. Understanding privacy and trust in smart home environments. Cham: Springer; 2020 Presented at: HCI for Cybersecurity, Privacy and Trust. HCII 2020; July 19-24, 2020; Copenhagen, Denmark. [doi: 10.1007/978-3-030-50309-3_34]
52. Schroeder T, Dodds L, Georgiou A, Gewald H, Siette J. Older adults and new technology: mapping review of the factors associated with older adults' intention to adopt digital technologies. JMIR Aging 2023;6:e44564 [FREE Full text] [doi: 10.2196/44564] [Medline: 37191976]
53. Chen K. Why do older people love and hate assistive technology? an emotional experience perspective. Ergonomics 2020;63(12):1463-1474. [doi: 10.1080/00140139.2020.1808714] [Medline: 32780683]
54. Oh SS, Kim KA, Kim M, Oh J, Chu SH, Choi J. Measurement of digital literacy among older adults: systematic review. J Med Internet Res 2021;23(2):e26145 [FREE Full text] [doi: 10.2196/26145] [Medline: 33533727]

Abbreviations

AARC: Awareness of Age-Related Change
CFA: confirmatory factor analysis
CFI: comparative fit index
GFI: goodness-of-fit index
IFI: incremental fit index
MANCOVA: multivariate analysis of covariance
PGFI: parsimony goodness-of-fit index
RMSEA: root mean square error of approximation
TLI: Tucker-Lewis index
TPB: Theory of Planned Behavior
TRQ: Technology Readiness Questionnaire
WHO: World Health Organization

Edited by J Wang; submitted 18.04.24; peer-reviewed by H Shin, M Khamzina; comments to author 18.07.24; revised version received 01.08.24; accepted 14.01.25; published 05.02.25.

Please cite as:
Wong KP, Teh PL, Lim WM, Lee SWH
Enhancing Older Adults' Lives Through Positive Aging Perception, Quality-of-Life Enhancement, and Social Support to Drive Acceptance and Readiness Toward Indoor Assistive Technology: Cross-Sectional Study
JMIR Aging 2025;8:e59665
URL: https://aging.jmir.org/2025/1/e59665
doi:10.2196/59665
PMID:

XSL•FO
RenderX

©Ka Po Wong, Pei-Lee Teh, Weng Marc Lim, Shaun Wen Huey Lee. Originally published in JMIR Aging (https://aging.jmir.org), 05.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Self-Adaptive Serious Game to Improve Motor Learning Among Older Adults in Immersive Virtual Reality: Short-Term Longitudinal Pre-Post Study on Retention and Transfer

Gauthier Everard[1,2,3], PT, MSc, PhD; Louise Declerck[3], PT, MSc, PhD; Thierry Lejeune[3,4,5], MD, PhD; Martin Gareth Edwards[5,6], PsyD, PhD; Justine Bogacki[3], PT; Cléo Reiprich[3], PT; Kelly Delvigne[3], PT; Nicolas Legrain[3], PT; Charles Sebiyo Batcho[1,2], PT, MSc, PhD

[1]School of Rehabilitation Sciences, Faculty of Medicine, Laval University, Québec, QC, Canada

[2]Centre interdisciplinaire de recherche en réadaptation et intégration sociale, Université Laval, Québec, QC, Canada

[3]Neuro Musculo Skeletal Lab (NMSK), Institut de Recherche Expérimentale et Clinique, UCLouvain, Brussels, Belgium

[4]Services de Médecine Physique et Réadaptation, Cliniques Universitaires Saint-Luc, Brussels, Belgium

[5]Louvain Bionics, UCLouvain, Louvain-la-neuve, Belgium

[6]Psychological Sciences Research Institute (IPSY), UCLouvain, Louvain-la-neuve, Belgium

Corresponding Author:
Gauthier Everard, PT, MSc, PhD
School of Rehabilitation Sciences
Faculty of Medicine
Laval University
2325 Rue de l'Université
Québec, QC, G1V0A6
Canada
Phone: 1 4185299141 ext 46914
Email: gauthier.everard@uclouvain.be

Abstract

Background: Despite their potential, the use of serious games within immersive virtual reality (iVR) for enhancing motor skills in older adults remains relatively unexplored. In this study, we developed a self-adaptive serious game in iVR called REAsmash-iVR. This game involves swiftly locating and striking a digital mole presented with various distractors.

Objective: This short-term longitudinal pre-post study aims to evaluate REAsmash-iVR's efficacy in promoting motor learning in older adults. Specifically, we seek to determine the transfer and retention of motor learning achieved through REAsmash-iVR to other iVR tasks.

Methods: A total of 20 older adults participated in the study, engaging with REAsmash-iVR over 7 consecutive days. The evaluation included iVR tests such as KinematicsVR and a VR adaptation of the Box and Block Test (BBT-VR). KinematicsVR tasks included drawing straight lines and circles as fast and as accurately as possible, while BBT-VR required participants to move digital cubes as quickly as possible within 60 seconds. Assessments were conducted before and after the intervention, with a follow-up at 1 week post intervention. The primary outcome focused on evaluating the impact of REAsmash-iVR on speed-accuracy trade-off during KinematicsVR tasks. Secondary outcomes included analyzing movement smoothness, measured by spectral arc length, and BBT-VR scores.

Results: Results revealed significant improvements in speed-accuracy trade-off post intervention compared to that before the intervention, with notable retention of skills for straight lines (t_{19}=5.46; P<.001; Cohen d=1.13) and circle drawing (t_{19}=3.84; P=.001; Cohen d=0.787). Likewise, there was a significant enhancement in spectral arc length, particularly for circle drawing (χ^2_2=11.2; P=.004; ε^2=0.23), but not for straight-line drawing (χ^2_2=2.1; P=.35; ε^2=0.003). Additionally, participants demonstrated transfer with significant improvement (q=5.26; P<.001; Cohen r=0.678) and retention (q=6.82; P<.001; Cohen r=0.880) in BBT-VR skills.

Conclusions: These findings provide perspectives for the use of iVR to improve motor learning in older adults through delivering self-adaptive serious games targeting motor and cognitive functions.

XSL•FO
RenderX

Trial Registration: ClinicalTrials.gov NCT04694833; https://clinicaltrials.gov/study/NCT04694833

(JMIR Aging 2025;8:e64004) doi:10.2196/64004

KEYWORDS

virtual reality; aged; learning; upper extremity; video games; kinematics

Introduction

By 2050, the proportion of older adults (older than 65 years) worldwide will nearly reach 22% [1,2]. Within the population, some older adults experience progressive functional decline, encompassing both motor and cognitive aspects [3]. This decline exacerbates issues of inactivity and sedentariness [4] contributing to increased prevalence of age-related diseases [5,6]. As an illustration, it is projected that by the year 2050, the number of individuals affected by major neurocognitive disorders will surge from 50 to 152 million worldwide, marking a 3-fold increase in cases [7].

Older adults' functional decline is typically associated with lower motor functions [8] and reduced quality of life [9,10]. Consequently, older adults tend to adopt compensatory behaviors that mitigate the impact of these reduced functions on their daily living activities [11]. These behavioral compensations include making slower, less accurate, less linear, and less smooth movements [12,13].

Motor learning refers to any experience-dependent improvement of a skill and typically involves both motor and cognitive processes [14]. Once learned, motor skills can be retained for an extended duration, resulting in sustained enhancements in performance [14]. A skill is not deemed fully acquired until the ability to retain or apply it in different contexts is demonstrated [15]. Research suggests that both healthy older adults and those with neurocognitive disorders can improve their motor performance through motor learning, exhibiting enhancements in movement speed, smoothness, coordination, and accuracy [16,17]. However, current motor learning programs mainly focus on gait and balance and therefore, demand time and availability from caregivers [18]. In response to these demands, new portable devices with cost-saving potential such as virtual reality (VR) might be of interest to promote motor learning in older adults.

VR can be defined as a computerized technological system that allows users to interact with a simulated multisensorial environment while providing real-time performance feedback [19]. Two main types of VR experience exist. Nonimmersive VR (niVR) is where users maintain awareness of their physical surroundings and receive visual feedback via a 2D display. Immersive VR (iVR) facilitates total immersion in the digital environment (using a head-mounted display or a large, curved screen with a panoramic view), with a comprehensive panoramic perspective [20]. Recent research suggests that VR programs may enhance participants' level of physical activity [21] and cognitive skills such as reaction time [22].

In rehabilitation, VR devices are often combined with serious games. Serious games refer to any game-based initiative that primarily focuses on learning objectives (such as education or rehabilitation) rather than simple entertainment [23]. Serious games have the capability to fulfill motor learning principles [24-26], as they motivate participants to make numerous practice repetitions through the use of multisensory feedback, personalized challenges, and through use of compelling and enriched environments [27,28]. After a certain period of familiarization, VR devices may allow participants to follow self-directed interventions and complete remote assessment of objective motor and cognitive performance (eg, analysis of kinematics and reaction time) during interventions [29]. However, despite the potential, the use of serious games in VR to promote motor learning in older adults remains underexplored [30]. In addition, the generalization of skills acquired in iVR to other skills in iVR remains debated.

A recent review has proposed intriguing methods to enhance the comprehension of motor learning in VR, including tracking participants' kinematics, manipulating sensorial feedback and difficulty parameters, and precisely simulating VR physics [27]. Prior work showed that kinematic indexes acquired in iVR (eg, movement linearity) were reliable and could possibly differentiate hand movements between healthy older adults and those with major neurocognitive disorders [31]. Additionally, several studies have provided evidence to support the idea that the provisioning of haptic and visual feedback in VR positively influenced movement smoothness, accuracy, and rapidity, thereby contributing to the improvement of motor learning in VR [32-35]. Recent evidence also highlights the efficacy of incorporating bimanual tasks in VR for promoting unimanual motor learning, aligning with the notion that the acquisition of motor skills in a digital environment can be optimized when the tasks closely mimic the complexities of everyday activities [36-38]. In line with the Yerkes-Dodson law, there is also evidence indicating that to maintain participants' motivation, the level of difficulty should be optimally balanced, neither too hard nor too easy [39]. To achieve this optimal balance, research suggests that game difficulty should be adjustable and tailored to individual participants' motor and cognitive performance [29,40,41]. Research supports the idea that self-adaptive training, where the difficulty is adjusted based on real-time performance, can optimize learning [42]. For instance, 1 study has highlighted that individualized VR training for driving led to more effective learning and retention of performance compared to traditional VR, video, and manual training [43]. The literature indicates success rates ranging from 60% to 80% as ideal for effectively enhancing participants' motivation [40,44].

Regarding older adults, several studies have demonstrated that motor learning in VR can be effective, resulting in improvements in motor performance along with the retention of learned skills over time. For example, a longitudinal study found that both healthy older adults and those with Parkinson disease were able to achieve learning and retention of skills

XSL•FO
RenderX

across 10 different niVR games, with effective transfer of these skills to similar untrained tasks [45]. While these results are promising for niVR, the retention of iVR skills and transfer to other tasks remains underexplored. A recent multicentric large parallel randomized controlled trial (n=293) has nevertheless produced encouraging results, demonstrating that progressive cognitive-motor training in iVR was effective in older adults, with greater improvements in global cognition and physical frailty compared to traditional interventions [46].

Following current recommendations aiming at improving motor learning in older adults, we developed a self-adaptive serious game in iVR (REAsmash-iVR) [47]. This game consists of finding and hitting a digital mole as fast as possible when presented with different types of distractors. In this version, we use a regulator to continuously adapt exercise difficulty according to participants' performance. As this version has not yet been tested among older adults, this work aims to test the feasibility and effectiveness of REAsmash-iVR in promoting motor learning within this population. Specifically, we sought to determine the transfer and retention of motor learning achieved through REAsmash-iVR to other iVR tasks. We hypothesized that REAsmash-iVR would significantly improve participants' unimanual reaching velocity and accuracy in other iVR tasks. We also aimed to assess the effect of REAsmash-iVR on participants' simple reaction time.

Methods

Study Design

This study used a short-term longitudinal pre-post design, with data collected at 3 time points: baseline (T0), immediate postintervention (T1), and 1-week follow-up (T2). This design allowed us to assess both immediate and retained motor learning effects, aligning with established motor learning literature for intermediate-term retention and transfer [39].

Ethical Considerations

The study was conducted in adherence to the principles of the Helsinki Declaration and received approval from the Hospital-Faculty Ethics Committee of Saint-Luc-UCLouvain in Belgium (B403201524184) and the Recherche Sectorielle en Réadaptation et Intégration Sociale Ethics Committee in Canada (#2020-1909). Prior to commencing the trial, all participants provided written informed consent. The study adhered to Transparent Reporting of Evaluations with Nonrandomized Designs (TREND) guidelines. Participation was voluntary and uncompensated, and all data was collected, stored, and analyzed in a manner that ensured participant anonymity.

Participants

The study recruited participants from the Belgian and Canadian populations between October 2022 and December 2023 using convenient sampling, leveraging word of mouth, and community outreach strategies. Inclusion criteria were individuals aged 65 years and older, possessing corrected-to-normal vision, and demonstrating the ability to comprehend simple instructions. Older adults with orthopedic or neurological disorders that might have impacted their capacity to handle a controller or that could alter upper extremity movements were excluded from the study. Participants' cognition was screened using the Montreal Cognitive Assessment [48].

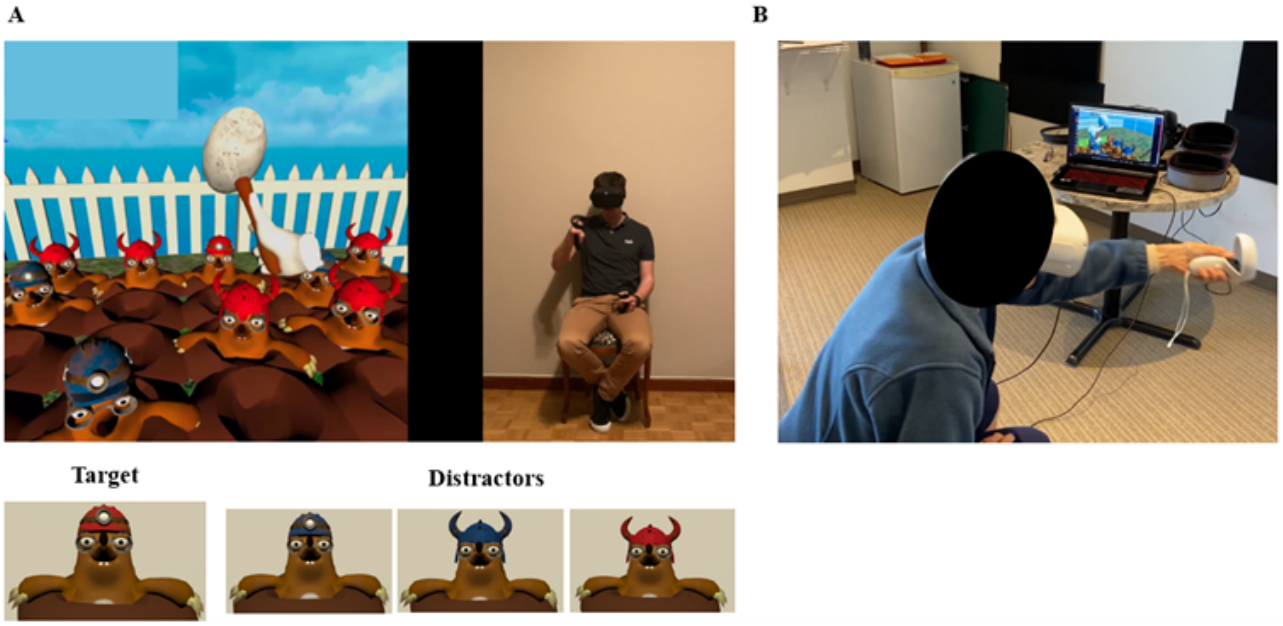
A flowchart diagram illustrating the participant flow through each stage of the study is presented in the Results section, as recommended by reporting guidelines such as TREND and CONSORT (Consolidated Standards of Reporting Trials) to ensure transparency and clarity in reporting [49,50].

Materials

The self-adaptive serious game REAsmash-iVR was developed using Unity 2019.3.15 software on the Oculus Quest 2 (Meta). This headset provides a high-resolution display (1832×1920 pixels per eye) and up to 90 Hz refresh rate, which ensures smooth and immersive interaction. The system includes 6 degrees of freedom tracking through integrated sensors, which allows participants to move freely within the digital environment. The device is equipped with 2 handheld controllers with motion tracking, which participants use to interact with the game, particularly for striking the target with a digital hammer. The controllers are equipped with motion sensors and buttons for precise input. As presented in Figure 1, Sidequest software was used to facilitate synchronization and video sharing from the headset to a laptop during the experiment.

XSL•FO
RenderX

Figure 1. Illustration of the REAsmash-iVR. (A) The upper part of the panel simultaneously depicts the REAsmash-iVR environment as seen through the VR headset (left) and the corresponding movements performed by a participant while interacting with the system (right). The lower part of the panel shows the target (a mole wearing a red miner helmet) and the distractors (moles wearing blue miner helmets, blue-horned helmets, and red-horned helmets). (B) Experimental setup: This panel illustrates a participant playing REAsmash-iVR using an immersive VR headset connected to a computer, which streams the application. iVR: immersive virtual reality; VR: virtual reality.



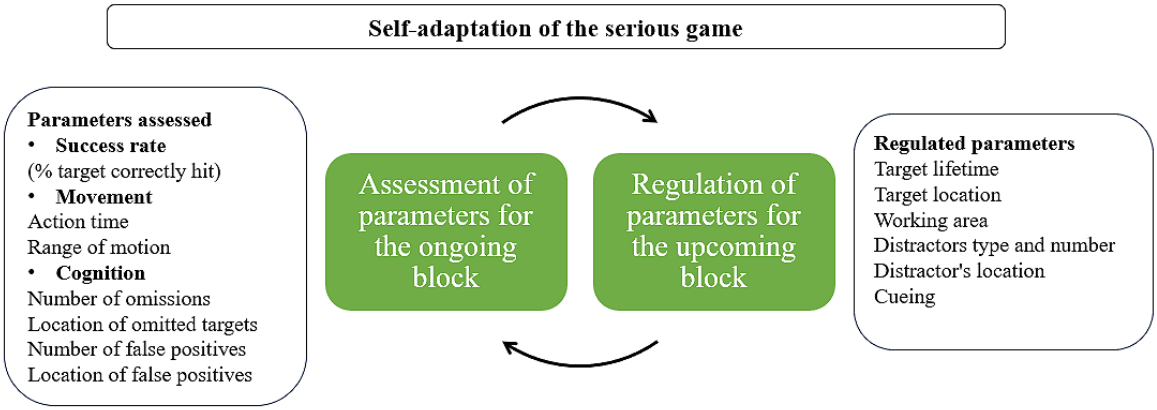
In the REAsmash-iVR serious game, the participant is asked to locate a target as quickly as possible. The target consists of a mole wearing a red miner's helmet and is presented among distractors (moles wearing helmets of different shapes and colors: a blue miner's helmet, a horned blue helmet, and a horned red helmet). These distractors manipulate cognitive difficulty by competing for attention during the task, requiring participants to focus and filter out irrelevant stimuli while searching for the target. To enhance cognitive engagement, distractors were designed to vary in salience, with some mimicking the target more closely in shape and color. Throughout the game, participants were instructed to only hit the target mole with the red miner's helmet, a task that requires both attention and precise motor action. To this end, participants used digital hammers, operated by the game controllers [47]. Motor function difficulty was manipulated through iterative practice of upper extremity reaching motions in different directions and with different levels of velocity.

In the version used here, the REAsmash-iVR used a regulator to adjust difficulty automatically and progressively based on the participants' motor and cognitive performance. The regulator of REAsmash-iVR difficulty aims to have the participant an average 75% successful performance. As the user improved and learned, the game progressively became more difficult, thereby maintaining the 75% optimal success rate. From a motor learning perspective, as the game escalated in difficulty, participants were compelled to execute a greater number of reaching movements toward the moles, spanning further distances, and within shorter timeframes, necessitating more efficient and precise upper extremity actions.

To ensure a continuous and progressive adaptation of the game difficulty, we used a dynamic regulation, with an infinite number of trial blocks. Each block involved finding and hitting a total of 1 to 24 target mole (trials), depending on the participant's level of performance. Between each block, the algorithm moderates the difficulty of the game based on the overall success rate (the ratio between the number of target moles accurately hit and the number of trials) of the prior block. If the success rate was superior to 75%, the game was considered too easy by the algorithm. If the success rate ranged between 50% and 75%, the game was considered difficult. If the success rate was below 50%, the game was considered excessively difficult. Depending on the success rate of the prior block (>75% vs 50%-75% vs <50%), the parameters that were considered responsible for the observed success rate were adjusted by the algorithm. More specifically, as presented in Figure 2, during each block, in addition to the overall success rate, the algorithm evaluated the proportion of omissions (instances when the target mole is not hit vs hit), the location of omitted moles, the number of false positives (distractor moles that were hit), and the location of distractor moles hit. These outcomes were used as indicators by the algorithm to see which of the following parameters had to be adjusted for the next block: the timing and location of the target appearance, the quantity (number) and types of distractors (high vs low salience contrast), the working area (where target and distractors appeared) and the delivery of cues that helped the participant to find the target (no cues, spatial auditive cue, and visual cues).

XSL•FO
RenderX

Figure 2. Self-adaptation of REAsmash-iVR. The system assesses performance parameters (success rate, movement, and cognition) after each block and regulates the task difficulty for the subsequent block by adjusting target and distractor characteristics, working area, and cueing to maintain an optimal challenge level for the participant. iVR: immersive virtual reality.



Experimental Protocol

Participants were tasked with engaging in REAsmash-iVR for 7 consecutive days for at least 15 minutes per day at their home. They were required to maintain a seated position throughout the gameplay and were asked to use both hands during the game. The feasibility of using REAsmash-iVR in older adults encompassed the documentation of adverse events, intervention duration (measured by the total minutes of active interaction with REAsmash-iVR, excluding pauses and time spent on menus and settings), and the evaluation of the game regulation efficacy.

To investigate whether participants' motor learning in REAsmash-iVR transferred to other iVR tasks and demonstrate retention of this transfer, the following measures were performed before the intervention (T0), immediately after (T1), and 7 days later (T2). We assessed motor learning transfer using 2 distinct tools: the KinematicsVR and the iVR version of the Box and Block Test (BBT-VR). These tests were administered by assessors with a background in physiotherapy and experience in using VR. Motor learning transfer refers to the ability to apply skills and improvements acquired during training to new, untrained tasks or contexts. Effective transfer indicates that participants have not only improved performance in the trained task (eg, REAsmash-iVR) but also developed adaptable motor strategies that can be used in other scenarios (eg, KinematicsVR and BBT-VR).

The 1-week duration was chosen as it is commonly used in motor learning literature to assess intermediate-term retention and transfer [39,51]. This timeframe is long enough to observe whether improvements persist beyond the immediate training environment but avoids confounding factors associated with longer intervals, such as unrelated learning or natural recovery. Prior studies have similarly used 1-week retention tests to evaluate skill stabilization and generalization in both laboratory and applied contexts [39,51].

We used the KinematicsVR to assess motor learning transfer in terms of unilateral reaching performance in a novel visuo-motor task [31]. The KinematicsVR requires using one of the VR headset controllers to swiftly and precisely draw 3D shapes (straight lines and a circle) visually presented in an iVR environment. The 3D positions of the controller were registered as an export file (.csv) in the hardware of the headset and analyzed offline. This test was deemed reliable and usable to assess upper extremity kinematics (especially during the drawing of straight lines and circles) in older adults with and without major neurocognitive disorder [31]. For this protocol, participants were asked to perform the movements with their dominant hand. All participants underwent familiarization trials before the assessment. To ensure that the differences between T0 and T1 primarily resulted from the REAsmash-iVR intervention rather than a general learning effect across trials, we contrasted the changes in our sample with those observed in a previous study [31] where an equal sample of older adults underwent the test twice consecutively (with no intervention in between). Transfer in this context indicates that the REAsmash-iVR intervention contributed to general improvements in unilateral reaching performance in a novel task (KinematicsVR). The performance in Kinematics-VR was analyzed based on metrics such as movement velocity, accuracy, and smoothness, allowing us to determine the extent to which the trained skills carried over to this analytic task.

We also used the BBT-VR to evaluate motor learning transfer in terms of gross unilateral manual dexterity [52,53]. This test involved moving digital cubes one at a time from 1 side of a box to the other within a 60-second timeframe. During the test, participants were required to grasp the cubes using their thumb, index, and middle fingers while pressing corresponding buttons on the controller. The BBT-VR was found to be valid, reliable, and usable to assess manual dexterity in healthy adults and individuals with stroke [52,53]. Transfer in this context indicates that the REAsmash-iVR intervention contributed to general improvements in fine motor control and hand-eye coordination. The performance in BBT-VR was analyzed based on metrics such as the number of blocks transferred, allowing us to determine the extent to which the trained skills carried over to this dexterity-focused task.

We also evaluated the participants' motor learning transfer to simple reaction time. The task involved detecting, as quickly as possible, a stimulus presented on a computer screen. Participants were instructed to click as quickly as possible on the touchpad when the stimuli were presented [54].

XSL•FO
RenderX

Kinematic Analyses

By analyzing the kinematic features of movements in KinematicsVR, we could assess whether the movement strategies developed during REAsmash-iVR translated into improved performance in a new visuomotor task.

The 3D positions of the controller obtained during the Kinematics-VR test were extracted from export files (.csv) at a sampling rate of 60 Hz. The analysis of kinematic data was then conducted using a program internally developed in Python (Python Software Foundation). For each participant, a preliminary visual analysis of the data was performed (to ensure that the data were correctly acquired) and signal smoothing was applied using a Butterworth filter (sampling frequency=60 Hz; cutoff frequency=10 Hz). The following kinematic indexes were calculated: the speed-accuracy trade-off (SAT) and the spectral arc length (SPARC).

The SAT, measured in arbitrary units, is a fundamental motor learning metric that quantifies the balance between movement speed and precision, often reflecting the extent to which participants prioritize speed over accuracy or vice versa during task execution. It is widely used to assess training-induced improvements in motor performance, as optimized performance behavior tends to achieve a more favorable balance between these competing demands [55]. In this study, we computed SAT as a ratio between speed and error, using the following equation.



Velocity refers to the first derivative of controller position. Error is measured based on movement linearity, which involves comparing the displacement of the controller with the ideal path. A higher SAT thus reflects a more efficient balance of speed and accuracy, indicating potential motor learning gains.

SPARC is a measure of movement smoothness, which is a key aspect of movement quality and skillful performance. Smoothness, as computed with SPARC, provides critical insights into whether a movement is natural and healthy or involves compensatory strategies. Natural and healthy movements are generally smoother, reflecting efficient neuromotor control, while compensatory movements tend to be less smooth and more erratic. SPARC is computed as the arc length of the instantaneous speed spectrum (ie, the length of the curve depicting the normalized amplitude of the "speed" signal  as a function of its frequency [ω]) [56]. A smoother movement involves less intermittency (alternance of acceleration and deceleration) typically resulting in a more compact and less erratic speed spectrum, leading to a small arc length. A negative sign is added to the computed arc length such a more negative SPARC value corresponds to a less smooth movement. This convention ensures that higher (less negative) SPARC values indicate smoother and more skillful motor performance.

Data Analysis

We performed statistical analyses using Sigmaplot (version; 13.0, Systat Software Inc) with α=.05. For each analysis, we explicitly tested the normality of the data using the Shapiro-Wilk test, and the results informed the selection of appropriate statistical methods (parametric or nonparametric). As this study is the first to test the feasibility and effectiveness of REAsmash-iVR in promoting motor learning, a convenience sample of 20 participants was determined.

To evaluate the efficacy of REAsmash-iVR self-regulation, we reported the percentage of instances where participants achieved a median success rate falling within the range of 60%-80% (a range deemed acceptable for enhancing motivation in a gamified learning context [40,44]). This evaluation was performed across the initial 55 blocks (which represented the minimum number of blocks observed in all participants).

Our primary outcome was to assess participants' motor learning transfer (and retention) to unilateral reaching performance in iVR. To analyze this, we used separate 1-way repeated measure ANOVA (or Friedman test for nonnormal data) for each shape used in the KinematicsVR assessment, comparing participants' SAT across 3 time points: before the intervention (T0), immediately after (T1), and at follow-up (T2). Post hoc pairwise comparisons were performed to detect changes between T0, T1, and T2. We used Bonferroni or Tukey adjustments (depending on the normality of the data) to control for the increased risk of type I error due to multiple comparisons. To ensure that the observed changes in KinematicsVR metrics between T0 and T1 were predominantly attributable to the REAsmash-iVR intervention and not merely a general learning effect over trials, we compared these changes with data from a previous study [31] where older adult participants underwent the test twice consecutively without any intervention in between [31]. To compare the T1-T0 changes between this study and our prior one, we used either the Mann-Whitney rank sum test (for nonnormal data) or the 2-tailed *t* test (for normally distributed data), depending on the distribution of the data.

Secondary outcomes included the assessment of REAsmash-iVR motor learning transfer to movement smoothness, manual dexterity performance in iVR, and simple reaction time. These outcomes were analyzed using 1-way repeated measures ANOVAs (or Friedman tests), with post-hoc pairwise comparisons. We also used Bonferroni or Tukey adjustments, depending on the normality of the data.

The effect size was computed using η^2 for ANOVAs, ε^2 for the Friedman test, and adjusted Cohen *d* for parametric post-hoc pairwise comparisons and 2-tailed *t* tests. For nonparametric tests, Cohen *r* was used. η^2 for ANOVAs and ε^2 for the Friedman test provide an estimate of the effect's magnitude relative to the total variance, with values of 0.01, 0.06, and 0.14 indicating small, medium, and large effects, respectively. Cohen *d* was used to quantify the standardized mean difference between conditions, with values of 0.2, 0.5, and 0.8 representing small, medium, and large effect sizes. For pairwise comparisons following the Friedman test and between-group comparisons with the Mann-Whitney *U* test, Cohen *r* was interpreted as small ($r\approx0.1$), medium ($r\approx0.3$), and large ($r\approx0.5$).

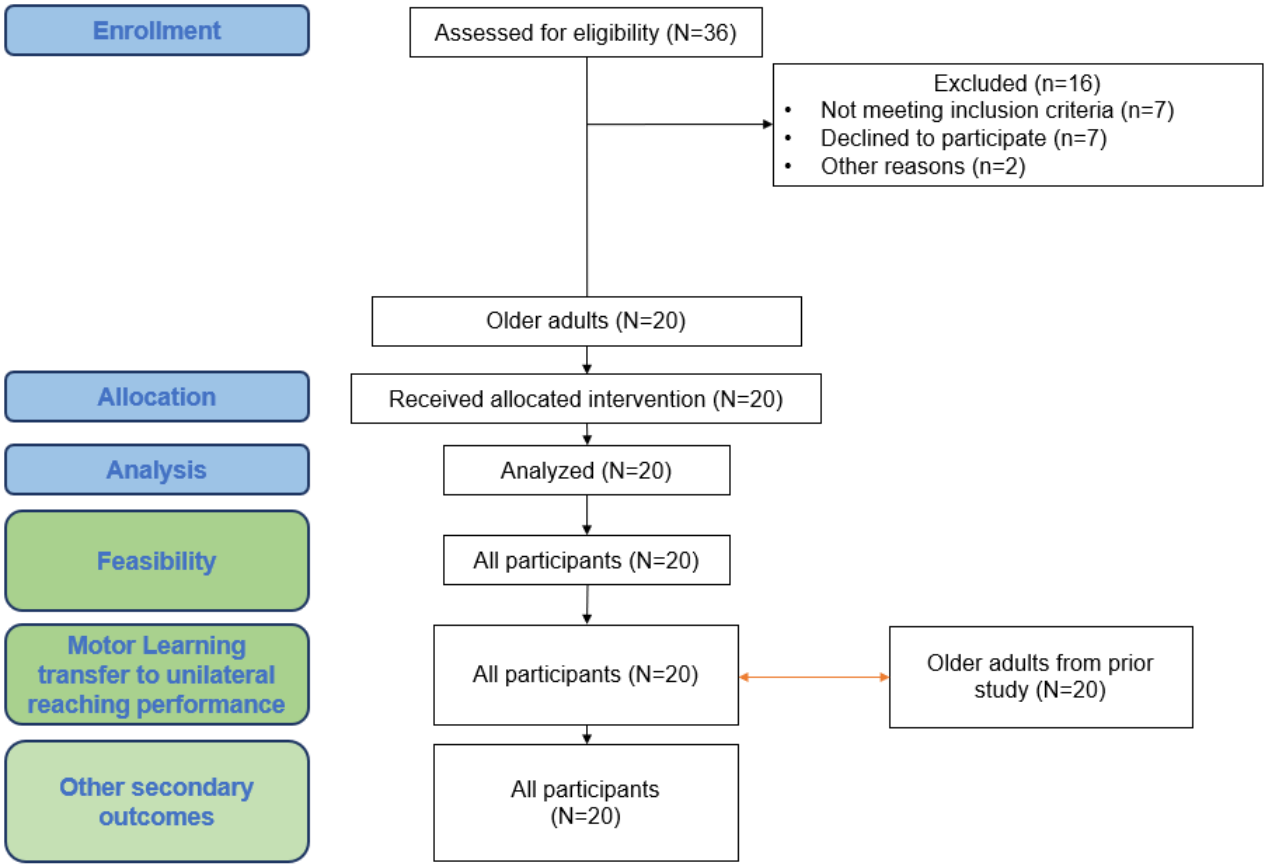
XSL•FO
RenderX

Results

Overview

A total of 20 older adults (of which 9 were women) with a mean age of 77.4 (SD 6.51) years participated in the study. Most of them were right-handed (n=17; 85%). A flowchart diagram is presented in Figure 3. Complementary information on participants' characteristics is provided in Table 1.

Figure 3. The flowchart diagram outlines the study process, including the assessment of 36 older adults for eligibility. Of these, 16 were excluded (7 did not meet the criteria, 7 declined, and 2 for other reasons). Twenty participants received the intervention, completed it, and were analyzed for feasibility, motor learning transfer, and other outcomes. Data were also compared with 20 older adult participants of a previous study.



XSL•FO
RenderX

Table 1. Participants' demographics.

Characteristic	Value
Age (years), median (IQR)	77.4 (69 to 89)
Sex, n (%)	
Female	9 (45)
Male	11 (55)
Dominant hand, n (%)	
Right	17 (85)
Left	3 (15)
Height (cm), mean (SD)	169.4 (7.73)
Weight (kg), mean (SD)	77.3 (14.40)
MoCA[a], median (IQR)	25 (6 to 30)
KinematicsVR: Straight lines	
Baseline SAT[b], mean (SD)	2.6 (1.09)
Baseline SPARC[c], median (IQR)	–1.97 (–2.027 to –1.896)
KinematicsVR: Circles	
Baseline SAT, mean (SD)	3.2 (1.51)
Baseline SPARC, mean (SD)	–4.45 (2.227)
Baseline BBT-VR[d], mean (SD)	29 (13.2)

[a]MoCA: Montreal Cognitive Assessment.

[b]SAT: speed-accuracy trade-off.

[c]SPARC: spectral arc length.

[d]BBT-VR: immersive virtual reality version of the Box and Block Test.

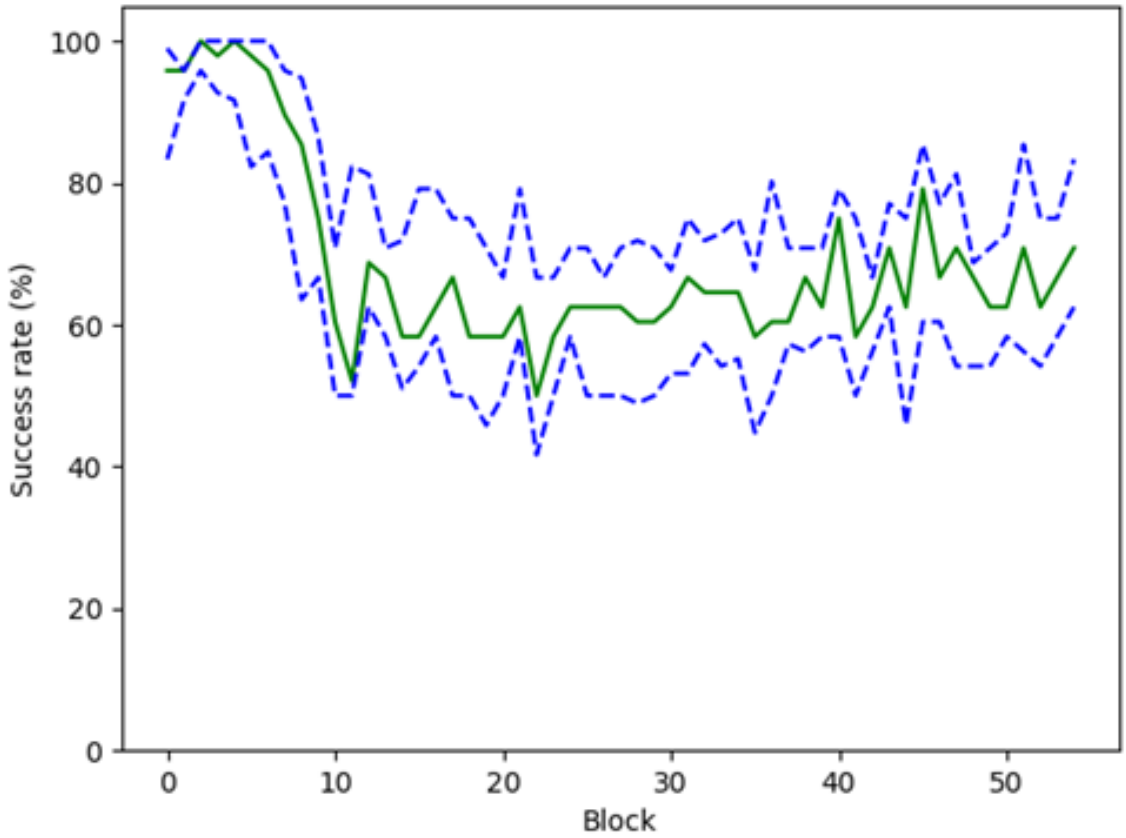
Feasibility

All participants finalized the study, and no adverse event occurred during the intervention. Participants actively played with REAsmash-iVR for 7 consecutive days for a median duration of 15.8 (IQR 9.73-15.12) minutes per day.

As illustrated in Figure 4, participants' median success rate reached a satisfactory value (60%-80%) after completing 11 blocks. Between the 11th and the 55th blocks, participants maintained a median success rate within the 60 to 80% range for 74% of the time (33 out of 44 blocks). Notably, during the final 11 blocks, participants consistently upheld a median success rate between 60% to 80%, reaching 100% coverage.

XSL•FO
RenderX

Figure 4. Evolution of success rate over the blocks. The x-axis represents the block number of the REAsmash-iVR intervention. The y-axis represents the participants' motor success rate. The green line represents the median and the blue lines the 1st and 3rd quartiles issued from all participants. iVR: immersive virtual reality; VR: virtual reality.



Primary Outcome: REAsmash-iVR Motor Learning Transfer to iVR Unilateral Speed-Accuracy Performance

As presented in Table 2 and Figure 5, separate repeated measures ANOVA ($F_{2,38}$=21.9; P<.001; η^2=0.535) and pairwise comparison revealed that directly after the REAsmash-iVR intervention, participants significantly improved their SAT in drawing straight lines (t_{19}=5.97; P<.001; Cohen d=1.20) with retention at T2 (T0 vs T2: t_{19}=5.46; P<.001; Cohen d=1.13). Regarding the drawing of circles, participants showed significant SAT improvements ($F_{2,38}$=7.7; P=.002; η^2=0.290) between T0 and T1 (T0 vs T1: t_{19}=2.64; P=.036; Cohen d=0.613) with retention at T2 (T0 vs T2: t_{19}=3.84; P=.001; Cohen d=0.787; Table 2 and Figure 5).

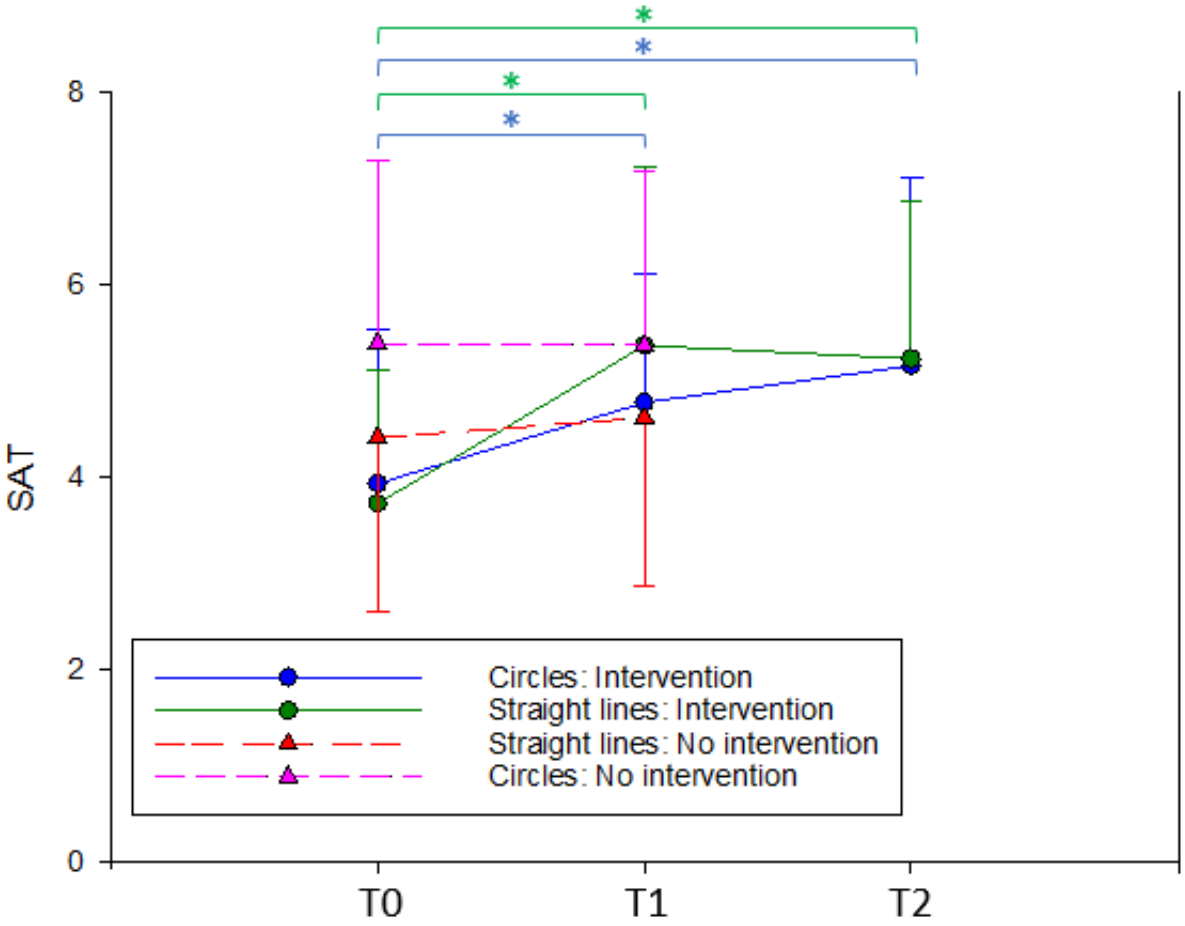
Table 2. Speed-accuracy trade-off changes over time.

Separate ANOVAs	T0	T1	T2	*F* test (*df*)	RM[a] ANOVA *P* value	Post hoc T0 versus T1 *P* value	Post hoc T0 versus T2 *P* value
Straight lines, mean (SD)	3.7 (1.38)	5.4 (1.85)	5.2 (1.63)	21.9 (2,38)	<.001	<.001	<.001
Circles, mean (SD)	3.9 (1.60)	4.8 (1.34)	5.2 (1.95)	7.7 (2,38)	.002	.036	.001

[a]RM: repeated measures.

XSL•FO
RenderX

Figure 5. Evolution of SAT over time. The * indicates the statistical significance of the ANOVA pairwise comparisons; Results are presented as mean and SDs. SAT: speed-accuracy trade-off.



When drawing straight lines, participants' SAT improvements between T0 and T1 were found to be significantly greater than those observed in participants who did not follow any intervention between the tests (t_{19}=3.0; *P*=.005; Cohen *d*=0.487). Contrastingly, when drawing circles, participants' SAT improvements between T0 and T1 were not significantly greater than those observed in participants who did not follow any intervention between the tests (t_{19}=1.96; *P*=.060; Cohen *d*=0.318).

Secondary Outcomes: REAsmash Motor Learning Transfer to iVR Movement Smoothness

As presented in Table 3, the Friedman test and pairwise comparisons showed that for the drawing of straight lines, participants did not observe significant SPARC changes between T0, T1, and T2 (χ^2_2=2.1; *P*=.35; ε^2=0.003; Table 3).

Table 3. Movement smoothness changes over time.

Separate ANOVAs	T0	T1	T2	Chi-square (*df*)	Friedman test *P* value	Post hoc T0 versus T1 *P* value	Post hoc T0 versus T2 *P* value
Straight lines, median (IQR)	–1.97 (–2.027 to –1.896)	–1.92 (–1.978 to –1.783)	–1.91 (–2.016 to –1.811)	2.1 (2)	.35	N/A[a]	N/A
Circles, median (IQR)	–3.77 (–5.876 to –2.517)	–2.67 (–4.000 to –2.359)	–2.44 (–3.287 to –2.273)	11.2 (2)	.004	.03	.004

[a]N/A: not applicable.

Contrastingly, for the drawing of circles, participants showed significant SPARC improvements (χ^2_2=11.20; *P*=.004; ε^2=0.23) between T0 and T1 (*q*=3.58; *P*=.03; Cohen *r*=0.8), with retention at T2 (T0 vs T2: *q*=4.47; *P*=.004; Cohen *r*=1.0; Table 3). Moreover, improvements between T0 and T1 were significantly greater than those observed in participants who did not follow any intervention between the tests (t_{19}=2.41; *P*=0.010; Cohen *d*=0.381).

Secondary Outcomes: REAsmash Motor Learning Transfer to Manual Dexterity Performance

As presented in Table 4, the Friedman test (χ^2_2=25.9; df=2; *P*<.001; ε^2=0.532) and pairwise comparison revealed that participants significantly improved their overall BBT-VR score

XSL•FO
RenderX

between T0 and T1 (q=5.26; P<.001; Cohen r=0.678) with significant retention at T2 (T0 vs T2: q=6.82; P<.001; Cohen r=0.880).

Table 4. BBT-VR[a] scores and reaction times change over time.

	T0	T1	T2	Statistics	RM[b] ANOVA or Friedman test P value	Post hoc T0 versus T1 P value	Post hoc T0 versus T2 P value
BBT-VR score (blocks), median (IQR)	27 (20.1 to 36.4)	36 (29.8 to 45.1)	38 (29.7 to 50.5)	χ^2_2=25.9	<.001	<.001	<.001
Simple reaction time (ms), mean (SD)	398.1 (116.89)	399.9 (94.12)	379.6 (80.31)	$F_{2,38}$=0.5	.64	N/A[c]	N/A

[a]BBT-VR: immersive virtual reality version of the Box and Block Test.

[b]RM: repeated measures.

[c]N/A: not applicable.

Secondary Outcomes: REAsmash Effect on Simple Reaction Time

Participants did not show significant simple reaction time changes ($F_{2,38}$=0.5; P=.64; η^2=0.024) between T0, T1, and T2 (Table 4).

Discussion

Principal Findings

This short-term longitudinal pre-post study aimed to evaluate the feasibility and effectiveness of using a self-adaptive serious game, REAsmash-iVR, to enhance motor learning and simple reaction time in older adults. Our results suggest that, on average, participants required 10 to 40 blocks to benefit from an acceptable-to-optimal level of success rate when starting from the easiest level of difficulty (block 0). Moreover, outcomes indicated that a 7-day intervention with REAsmash-iVR resulted in improved performance in other iVR tasks, as evidenced by enhanced SAT metrics in drawing straight lines and circles, and increased score in displacing digital blocks postintervention. Notably, significant retention of speed/accuracy improvement was observed at the 1-week follow-up for the drawing of straight lines and circles, and for the displacement of digital blocks. Similarly, secondary analyses revealed that the REAsmash-iVR intervention led to improved movement smoothness in drawing circles but not straight lines. Finally, we did not observe any significant effect of REAsmash-iVR on simple reaction time.

REAsmash-iVR Impact on Motor Learning

Our findings seem to indicate that older adults effectively achieved motor learning. After undergoing the REAsmash-iVR intervention, participants exhibited the ability to apply their skills in various contexts, as evidenced by enhanced speed and accuracy in KinematicsVR and improved performance in BBT-VR. Notably, certain improvements persisted even 1 week after the intervention, suggesting intermediate-term retention of learned skills and their potential transfer to other tasks. These findings may highlight a meaningful step in motor learning within this time frame.

On the one hand, our results [31] may indicate that the notable enhancements in speed-accuracy when drawing straight lines postintervention could be credited to the REAsmash-iVR intervention, rather than being merely a result of general learning or increased familiarity with the device and setup. In fact, in our prior study [31], where no intervention occurred between assessments, we observed no significant changes in SAT in KinematicsVR tasks, further supporting the conclusion that the improvements in this study are specifically attributable to the REAsmash-iVR intervention. The self-adaptive nature of REAsmash-iVR, which allowed for tailored and optimized adjustments to accommodate each participant's unique needs and capacities, likely played a role in optimizing engagement and facilitating skill acquisition, ultimately leading to the observed enhancements in motor performance. Previous research has demonstrated that older adults achieved significant motor learning and skill acquisition when provided with appropriate optimized interventions tailored to their specific needs and abilities. For instance, a study comparing a group who practiced a square-stepping task with enhanced feedback, autonomy-supportive choices, and optimized instructions to a control group practicing without these elements, found that the experimental group exhibited faster movement times during both practice and retention phases [57]. Similarly, feedback from participants in our study overwhelmingly attested to their enjoyment of the game. Many remarked on the engaging nature of the REAsmash-iVR intervention, highlighting its immersive qualities and the satisfaction derived from mastering new skills within the digital environment. Furthermore, studies have emphasized the importance of leveraging technological advancements to develop innovative interventions that cater to the unique challenges and preferences of older adults [58,59]. In a prior study, researchers showed that an adaptive video game training intervention led to generalized positive effects on cognitive control abilities in older adults [59]. Our findings also align with [45] a longitudinal, controlled clinical study investigating motor learning, retention, and transfer in older adults using VR-based training, specifically focusing on individuals with Parkinson disease [45]. Their results demonstrated that older adults with Parkinson disease could effectively learn new motor skills through VR-based interventions. Importantly, the study observed improvements not only in motor performance but also in the retention and transfer of learned skills to real-world tasks.

The greater performance improvement here, relative to our previous study [31] could be attributed to the participants

XSL•FO
RenderX

exhibiting lower baseline performance levels in this study, potentially allowing for a greater margin of improvement. This difference becomes particularly apparent when considering circle drawing tasks, where participants had similar baseline performance levels, and improvements in speed and accuracy did not exceed those seen in our prior study [31]. However, this divergence in outcomes could also stem from the inherent differences between drawing straight lines and circles (straight vs cyclic movements). Notably, it could be hypothesized that due to the similarity in movement nature (discrete reaching movements) between drawing straight lines and hitting digital moles, REAsmash-iVR likely played a significant role in augmenting speed and accuracy outcomes in straight-line drawing tasks.

Disparities in Transfer to Movement Smoothness and Manual Dexterity in iVR

Secondary analyses revealed that participants demonstrated enhanced movement smoothness in drawing circles but not straight lines. In a prior study, researchers have recently indicated that the intermittency of movement, as evidenced by the number of velocity peaks, is influenced by the specific task being performed in older adults [60]. In our study, the lack of significant changes observed in the drawing of straight lines could potentially be attributed to the 2D and discrete nature of these movements within the KinematicsVR application. This limitation may have restricted the scope for enhancement. Conversely, when considering the drawing of circles in KinematicsVR, the continuous and 3D nature of the movements offers potential for improvement in an additional dimension (compared to the drawing of straight lines where only 2 dimensions are considered in the KinematicsVR assessment). These hypotheses are supported by the baseline results, wherein participants demonstrated a median SPARC score of –1.97 (IQR –2.027 to –1.896) for straight lines, compared to a median SPARC score of –3.77 (IQR –5.876 to –2.517) for circles. Due to the tendency for smoother movements to be indicated by SPARC values closer to 0, these results underscore a heightened potential for enhancement in circle drawing tasks.

REAsmash-iVR Effect on Simple Reaction Time

The study did not observe any significant effect of REAsmash-iVR on simple reaction time. Although these findings could have been expected, they may reflect the specific design limitations of the intervention and assessment or the need for longer intervention durations to detect changes in these outcomes. The task in iVR involved locating and responding to digital moles as quickly as possible, suggesting the potential for improvements in simple reaction times. Especially since our sample had a mean reaction time (mean 398.1, SD 116.89 ms) more than the normal in equally aged standards [54]. In comparison, in a prior study [54], researchers observed that the average mean reaction time of individuals aged between 61 and 80 years for the same task was 296.1 (SD 63.9) milliseconds. However, REAsmash-iVR engages spatial attention and distractor inhibition, aspects that cannot be adequately assessed solely through simple reaction time measurements, as the latter primarily evaluates alertness levels. Moreover, it is important to consider that reaction time improvements tend to be modest in older adults due to age-related declines in neurological processing speed [61]. Several studies have documented a decline in reaction times with increasing age, reflecting changes in neurological networks and cognitive processing abilities [62-64]. A recent study assessing 861 participants aged 70-90 years also observed that an increase in intraindividual variability of reaction time, considered as a cognitive marker of neurobiological disturbance, was associated with dementia and mortality [65]. Therefore, longer intervention periods and more comprehensive assessments may be necessary to capture subtle improvements in reaction times among older adults participating in VR-based interventions.

Limitations

We acknowledge the following limitations. First, the study's sample size (n=20) and design (pre-post), while appropriate for initial exploration, pose limitations to the generalizability and robustness of the findings. Although a retrospective comparison was used to evaluate the effect of REAsmash-iVR on motor learning, using a randomized controlled trial design would provide stronger evidence for drawing conclusions regarding the intervention's efficacy.

Second, the length of the intervention period may have been too short to observe significant improvements in certain outcomes (eg, reaction time), particularly among healthy participants. A longer study duration would be beneficial for capturing more substantial changes. Future investigations involving healthy older adults and individuals with major neurocognitive conditions could provide valuable insights into the effects of REAsmash-iVR on a broader range of participants.

Third, for the BBT-VR, retrospective comparisons between participants who received the REAsmash-iVR intervention and those who did not were possible. As a result, it remains challenging to definitively attribute observed improvements solely to the intervention itself, as opposed to potential learning effects or mere familiarization with the VR device and testing procedures. Especially since the minimal detectable change of the BBT-VR in healthy adults is relatively high (14.06 for the dominant hand and 18.23 for the nondominant hand) [52]. Therefore, future studies incorporating appropriate control groups are essential to establish the causal relationship between the REAsmash-iVR intervention and the observed transfer of performance.

Implications

The results of this study carry significant clinical and research implications. Clinically, they underscore the potential of VR technology as a novel and engaging approach to promote motor learning and rehabilitation in older adults. Previous research has demonstrated that VR-based interventions can improve motor function and engagement in rehabilitation through gamified experiences, particularly in older populations with age-related declines [66-69]. Health care practitioners working in rehabilitation settings may consider integrating VR-based interventions into their programs targeting motor impairments and age-related declines in physical function. More importantly, the use of self-adaptive serious games such as REAsmash-iVR may offer participants the opportunity for self-rehabilitation

XSL•FO
RenderX

through tailored interventions. These interventions can be specifically designed to target individual motor and cognitive deficits, providing a personalized approach to rehabilitation. Personalization in VR rehabilitation has been shown to improve outcomes by adapting difficulty levels based on real-time performance data [43]. Furthermore, the diverse range of applications available within the VR headset may enable participants to receive real-time feedback on their performance, allowing for continuous monitoring of progress and identifying areas for improvement. Several studies have highlighted the potential of iVR to assess relevant quantitative outcomes, such as hand kinematics, gaze tracking, and reaction time, in a valid and reliable manner [47,70-72]. While traditional assessments may suffer from ceiling or floor effects, limiting their sensitivity to subtle changes in performance, these quantitative metrics provide precise, objective insights into how participants behave during the task [72,73]. This allows clinicians and researchers to track nuanced motor and cognitive responses, facilitating more tailored rehabilitation strategies. Such feedback mechanisms could enhance motivation and engagement, facilitating more effective rehabilitation outcomes. VR may boost motivation by providing immersive environments that promote goal-oriented tasks, immediate feedback, and a sense of achievement through gamified elements [74]. These features not only increase adherence to rehabilitation programs but also foster a positive emotional response, which is critical for sustaining long-term engagement and improving functional recovery [75,76]. Additionally, with advancements in VR and mixed-reality headset technologies, iVR devices hold promise for promoting social interaction and connectivity among older adults [77,78]. Research has shown that digital environments can foster social engagement, reducing isolation and improving mental well-being [79,80]. Digital conferences, collaborative gaming experiences, and social environments can be facilitated through VR platforms, fostering social engagement and reducing feelings of isolation, which are particularly relevant in the context of aging populations and social distancing measures. Incorporating these social aspects into VR-based interventions could not only enhance the overall user experience but would also contribute to the holistic well-being of older adults.

From a research perspective, the study highlights the importance of exploring optimal parameters and mechanisms underlying VR interventions to maximize their therapeutic benefits, eventually using regulators of difficulty. Larger controlled studies are needed to elucidate the long-term effects, optimal dosage, and generalizability of VR interventions across different populations and settings.

Conclusions

This study provides valuable insights into the feasibility and potential effectiveness of using a self-adaptive serious game to enhance motor learning in older adults. While the intervention demonstrated promising results in improving reaching accuracy and velocity balance, and movement smoothness, future research is warranted to elucidate its broader impact on physical and cognitive function in aging populations.

Acknowledgments

The authors thank Stéphane Grade, Camille Ganci, and Anthony Garcia for their technical support in the development of the game. This work was supported by postdoctoral grants of the Centre Interdisciplinaire de recherche en réadaptation et intégration sociale (Cirris), the Réseau provincial de recherche en adaptation-réadaptation, and Wallonie-Bruxelles International (all awarded to GE). The authors also thank the Région Wallonne, the SPW-Economie-Emploi-Recherche, and the Win2Wal Program (convention 1810108) for their support.

Data Availability

The data that support the findings of this study are available on request from the corresponding author.

Authors' Contributions

GE initiated the project. TL, MGE, and GE worked on the development of the application. CSB, LD, and GE designed the protocol. JB, CR, NL, and KD conducted the experiment. GE performed the statistical analyses. GE wrote the manuscript. GE realized the kinematic analyses. TL, MGE, CSB, and LD substantially contributed to the interpretation of the results. TL, LD, MGE, and CSB were major contributors to manuscript editing. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Wasay M, Grisold W, Carroll W, Shakir R. World Brain Day 2016: celebrating brain health in an ageing population. Lancet Neurol 2016;15(10):1008. [doi: 10.1016/S1474-4422(16)30171-5] [Medline: 27452915]
2. Ageing and health. World Health Organization. 2024. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health#:~:text=Between%202015%20and%202050%2C%20the,from%2012%25%20to%2022%25 [accessed 2024-02-13]
3. Colón-Emeric CS, Whitson HE, Pavon J, Hoenig H. Functional decline in older adults. Am Fam Physician 2013;88(6):388-394 [FREE Full text] [Medline: 24134046]

XSL•FO
RenderX

4. Botö S, Buvarp DJ, Hansson P, Sunnerhagen KS, Persson CU. Physical inactivity after stroke: incidence and early predictors based on 190 individuals in a 1-year follow-up of the fall study of Gothenburg. J Rehabil Med 2021;53(9):jrm00224 [FREE Full text] [doi: 10.2340/16501977-2852] [Medline: 34121128]
5. Marioni RE, Proust-Lima C, Amieva H, Brayne C, Matthews FE, Dartigues J, et al. Social activity, cognitive decline and dementia risk: a 20-year prospective cohort study. BMC Public Health 2015;15:1089 [FREE Full text] [doi: 10.1186/s12889-015-2426-6] [Medline: 26499254]
6. The Lancet Regional Health-Europe. Securing the future of Europe's ageing population by 2050. Lancet Reg Health Eur 2023;35:100807 [FREE Full text] [doi: 10.1016/j.lanepe.2023.100807] [Medline: 38115962]
7. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet 2020;396(10248):413-446 [FREE Full text] [doi: 10.1016/S0140-6736(20)30367-6] [Medline: 32738937]
8. Wang Z, Wang J, Guo J, Dove A, Arfanakis K, Qi X, et al. Association of motor function with cognitive trajectories and structural brain differences: a community-based cohort study. Neurology 2023;101(17):e1718-e1728 [FREE Full text] [doi: 10.1212/WNL.0000000000207745] [Medline: 37657942]
9. Pusswald G, Tropper E, Kryspin-Exner I, Moser D, Klug S, Auff E, et al. Health-related quality of life in patients with subjective cognitive decline and mild cognitive impairment and its relation to activities of daily living. J Alzheimers Dis 2015;47(2):479-486. [doi: 10.3233/JAD-150284] [Medline: 26401569]
10. Palacios N, Gao X, Schwarzschild M, Ascherio A. Declining quality of life in Parkinson disease before and after diagnosis. J Parkinsons Dis 2012;2(2):153-160 [FREE Full text] [doi: 10.3233/JPD-2012-12083] [Medline: 23939439]
11. Tomaszewski Farias S, Schmitter-Edgecombe M, Weakley A, Harvey D, Denny KG, Barba C, et al. Compensation strategies in older adults: association with cognition and everyday function. Am J Alzheimers Dis Other Demen 2018;33(3):184-191 [FREE Full text] [doi: 10.1177/1533317517753361] [Medline: 29357670]
12. Yan JH, Rountree S, Massman P, Doody RS, Li H. Alzheimer's disease and mild cognitive impairment deteriorate fine movement control. J Psychiatr Res 2008;42(14):1203. [doi: 10.1016/j.jpsychires.2008.01.006] [Medline: 18280503]
13. Yan JH. Effects of aging on linear and curvilinear aiming arm movements. Exp Aging Res 2000;26(4):393-407. [doi: 10.1080/036107300750015778] [Medline: 11091944]
14. Krakauer JW, Hadjiosif AM, Xu J, Wong AL, Haith AM. Motor learning. Compr Physiol 2019;9(2):613-663. [doi: 10.1002/cphy.c170043] [Medline: 30873583]
15. Muratori LM, Lamberg EM, Quinn L, Duff SV. Applying principles of motor learning and control to upper extremity rehabilitation. J Hand Ther 2013;26(2):94-103 [FREE Full text] [doi: 10.1016/j.jht.2012.12.007] [Medline: 23598082]
16. Yan JH, Dick MB. Practice effects on motor control in healthy seniors and patients with mild cognitive impairment and alzheimer's disease. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2006;13(3-4):385-410. [doi: 10.1080/138255890969609] [Medline: 16887780]
17. Onushko T, Kim C, Christou EA. Reducing task difficulty during practice improves motor learning in older adults. Exp Gerontol 2014;57:168-174. [doi: 10.1016/j.exger.2014.06.006] [Medline: 24909352]
18. Shafizadeh M, Farina N, Parvonpour S. Effects of motor learning interventions on walking performance and physical function in older adults with cognitive impairment and dementia: a systematic review and meta-analysis. J Aging Phys Act 2023;31(2):352-363. [doi: 10.1123/japa.2021-0442] [Medline: 35981711]
19. Saposnik G, Levin M, Outcome Research Canada (SORCan) Working Group. Virtual reality in stroke rehabilitation: a meta-analysis and implications for clinicians. Stroke 2011;42(5):1380-1386. [doi: 10.1161/STROKEAHA.110.605451] [Medline: 21474804]
20. Huygelier H, Mattheus E, Abeele VV, van Ee R, Gillebert CR. The use of the term virtual reality in post-stroke rehabilitation: a scoping review and commentary. Psychol Belg 2021;61(1):145-162 [FREE Full text] [doi: 10.5334/pb.1033] [Medline: 34131490]
21. Ng YL, Ma F, Ho FK, Ip P, Fu KW. Effectiveness of virtual and augmented reality-enhanced exercise on physical activity, psychological outcomes, and physical performance: a systematic review and meta-analysis of randomized controlled trials. Comput Hum Behav 2019;99:278-291. [doi: 10.1016/j.chb.2019.05.026]
22. Rutkowski S, Adamczyk M, Pastuła A, Gos E, Luque-Moreno C, Rutkowska A. Training using a commercial immersive virtual reality system on hand-eye coordination and reaction time in young musicians: a pilot study. Int J Environ Res Public Health 2021;18(3):1297 [FREE Full text] [doi: 10.3390/ijerph18031297] [Medline: 33535539]
23. Mubin O, Alnajjar F, Al Mahmud A, Jishtu N, Alsinglawi B. Exploring serious games for stroke rehabilitation: a scoping review. Disabil Rehabil Assist Technol 2022;17(2):159-165. [doi: 10.1080/17483107.2020.1768309] [Medline: 32508187]
24. Doumas I, Everard G, Dehem S, Lejeune T. Serious games for upper limb rehabilitation after stroke: a meta-analysis. J Neuroeng Rehabil 2021;18(1):100 [FREE Full text] [doi: 10.1186/s12984-021-00889-1] [Medline: 34130713]
25. Maier M, Rubio Ballester B, Duff A, Duarte Oller E, Verschure PFMJ. Effect of specific over nonspecific VR-based rehabilitation on poststroke motor recovery: a systematic meta-analysis. Neurorehabil Neural Repair 2019;33(2):112-129 [FREE Full text] [doi: 10.1177/1545968318820169] [Medline: 30700224]

XSL•FO
RenderX

26. Demers M, Fung K, Subramanian SK, Lemay M, Robert MT. Integration of motor learning principles into virtual reality interventions for individuals with cerebral palsy: systematic review. JMIR Serious Games 2021;9(2):e23822 [FREE Full text] [doi: 10.2196/23822] [Medline: 33825690]
27. Levac DE, Huber ME, Sternad D. Learning and transfer of complex motor skills in virtual reality: a perspective review. J Neuroeng Rehabil 2019;16(1):121 [FREE Full text] [doi: 10.1186/s12984-019-0587-8] [Medline: 31627755]
28. Saragih ID, Everard G, Lee B. A systematic review and meta-analysis of randomized controlled trials on the effect of serious games on people with dementia. Ageing Res Rev 2022;82:101740. [doi: 10.1016/j.arr.2022.101740] [Medline: 36191740]
29. Levin MF, Weiss PL, Keshner EA. Emergence of virtual reality as a tool for upper limb rehabilitation: incorporation of motor control and motor learning principles. Phys Ther 2015;95(3):415-425 [FREE Full text] [doi: 10.2522/ptj.20130579] [Medline: 25212522]
30. Abdi S, de Witte L, Hawley M. Emerging technologies with potential care and support applications for older people: review of gray literature. JMIR Aging 2020;3(2):e17286 [FREE Full text] [doi: 10.2196/17286] [Medline: 32780020]
31. Everard G, Boivin S, Boulay G, Duchemin R, Batcho CS. Immersive virtual reality to assess arm kinematics among older adults with and without major neurocognitive disorder—an exploratory cross-sectional study. Neuroscience 2024;537:47-57. [doi: 10.1016/j.neuroscience.2023.10.024] [Medline: 38006964]
32. Krogmeier C, Mousas C, Whittinghill D. Human-virtual character interaction: toward understanding the influence of haptic feedback. Comput Animat Virtual 2019;30(3-4):e1883. [doi: 10.1002/cav.1883]
33. Lontschar S, Deegan D, Humer I, Pietroszek K, Eckhardt C. Analysis of haptic feedback and its influences in virtual reality learning environments. 2020 Presented at: 6th International Conference of the Immersive Learning Research Network (iLRN); June 21-25, 2020; San Luis Obispo, CA. [doi: 10.23919/ilrn47897.2020.9155087]
34. Bugdadi A, Sawaya R, Bajunaid K, Olwi D, Winkler-Schwartz A, Ledwos N, et al. Is virtual reality surgical performance influenced by force feedback device utilized? J Surg Educ 2019;76(1):262-273. [doi: 10.1016/j.jsurg.2018.06.012] [Medline: 30072262]
35. Everard GJ, Lejeune TM, Batcho CS. Visual feedback and age affect upper limb reaching accuracy and kinematics in immersive virtual reality among healthy adults. Int J Rehabil Res 2023;46(3):221-229. [doi: 10.1097/MRR.0000000000000588] [Medline: 37334800]
36. Hoare B, Greaves S. Unimanual versus bimanual therapy in children with unilateral cerebral palsy: same, same, but different. J Pediatr Rehabil Med 2017;10(1):47-59. [doi: 10.3233/PRM-170410] [Medline: 28339410]
37. Kantak S, Jax S, Wittenberg G. Bimanual coordination: A missing piece of arm rehabilitation after stroke. Restor Neurol Neurosci 2017;35(4):347-364. [doi: 10.3233/RNN-170737] [Medline: 28697575]
38. Bhakuni R, Mutha PK. Learning of bimanual motor sequences in normal aging. Front Aging Neurosci 2015;7:76 [FREE Full text] [doi: 10.3389/fnagi.2015.00076] [Medline: 26005417]
39. Schmidt RA, Lee TD, Winstein C, Wulf G, Zelaznik HN. Motor Control and Learning: A Behavioral Emphasis. 6th edition. Champaign, IL: Human Kinetics; 2018.
40. Cameirão MS, Badia SBI, Oller ED, Verschure PFMJ. Neurorehabilitation using the virtual reality based rehabilitation gaming system: methodology, design, psychometrics, usability and validation. J Neuroeng Rehabil 2010;7:48 [FREE Full text] [doi: 10.1186/1743-0003-7-48] [Medline: 20860808]
41. Heins S, Dehem S, Montedoro V, Dehez B, Edwards M, Stoquart G. Robotic-assisted serious game for motor and cognitive post-stroke rehabilitation. 2017 Presented at: IEEE 5th International Conference on Serious Games and Applications for Health (SeGAH); April 2-4, 2017; Perth, WA, Australia.
42. Zahabi M, Abdul Razak AM. Adaptive virtual reality-based training: a systematic literature review and framework. Virtual Reality 2020;24(4):725-752. [doi: 10.1007/s10055-020-00434-w]
43. Lang Y, Wei L, Xu F, Zhao Y. Synthesizing personalized training programs for improving driving habits via virtual reality. 2018 Presented at: IEEE Conference on Virtual Reality and 3D User Interfaces (VR); March 18-22, 2018; Tuebingen/Reutlingen, Germany.
44. van der Kooij K, In 't Veld L, Hennink T. Motivation as a function of success frequency. Motiv Emot 2021;45(6):759-768 [FREE Full text] [doi: 10.1007/s11031-021-09904-3] [Medline: 34608344]
45. dos Santos Mendes FA, Pompeu JE, Modenesi Lobo A, Guedes da Silva K, Oliveira TDP, Peterson Zomignani A, et al. Motor learning, retention and transfer after virtual-reality–based training in Parkinson's disease—effect of motor and cognitive demands of games: a longitudinal, controlled clinical study. Physiotherapy 2012;98(3):217-223. [doi: 10.1016/j.physio.2012.06.001] [Medline: 22898578]
46. Kwan RYC, Liu J, Sin OSK, Fong KNK, Qin J, Wong JCY, et al. Effects of virtual reality motor-cognitive training for older people with cognitive frailty: multicentered randomized controlled trial. J Med Internet Res 2024;26:e57809 [FREE Full text] [doi: 10.2196/57809] [Medline: 39259959]
47. Ajana K, Everard G, Lejeune T, Edwards MG. A feature and conjunction visual search immersive virtual reality serious game for measuring spatial and distractor inhibition attention using response time and action kinematics. J Clin Exp Neuropsychol 2023;45(3):292-303. [doi: 10.1080/13803395.2023.2218571] [Medline: 37260369]

XSL•FO
RenderX

48. Hobson J. The Montreal Cognitive Assessment (MoCA). Occup Med 2015;65(9):764-765. [doi: 10.1093/occmed/kqv078] [Medline: 26644445]
49. Haynes AB, Haukoos JS, Dimick JB. TREND reporting guidelines for nonrandomized/quasi-experimental study designs. JAMA Surg 2021;156(9):879-880. [doi: 10.1001/jamasurg.2021.0552] [Medline: 33825826]
50. Hopewell S, Boutron I, Chan A, Collins GS, de Beyer JA, Hróbjartsson A, et al. An update to SPIRIT and CONSORT reporting guidelines to enhance transparency in randomized trials. Nat Med 2022;28(9):1740-1743. [doi: 10.1038/s41591-022-01989-8] [Medline: 36109642]
51. Walker MP, Stickgold R. Sleep-dependent learning and memory consolidation. Neuron 2004;44(1):121-133 [FREE Full text] [doi: 10.1016/j.neuron.2004.08.031] [Medline: 15450165]
52. Everard G, Otmane-Tolba Y, Rosselli Z, Pellissier T, Ajana K, Dehem S, et al. Concurrent validity of an immersive virtual reality version of the Box and Block Test to assess manual dexterity among patients with stroke. J Neuroeng Rehabil 2022;19(1):7 [FREE Full text] [doi: 10.1186/s12984-022-00981-0] [Medline: 35065678]
53. Everard G, Burton Q, Van de Sype V, Bibentyo TN, Auvinet E, Edwards MG, et al. Extended reality to assess post-stroke manual dexterity: contrasts between the classic box and block test, immersive virtual reality with controllers, with hand-tracking, and mixed-reality tests. J Neuroeng Rehabil 2024;21(1):36 [FREE Full text] [doi: 10.1186/s12984-024-01332-x] [Medline: 38491540]
54. Deary IJ, Liewald D, Nissan J. A free, easy-to-use, computer-based simple and four-choice reaction time programme: the Deary-Liewald reaction time task. Behav Res Methods 2011;43(1):258-268. [doi: 10.3758/s13428-010-0024-1] [Medline: 21287123]
55. FITTS PM. The information capacity of the human motor system in controlling the amplitude of movement. J Exp Psychol 1954;47(6):381-391. [Medline: 13174710]
56. Balasubramanian S, Melendez-Calderon A, Roby-Brami A, Burdet E. On the analysis of movement smoothness. J Neuroeng Rehabil 2015;12:112 [FREE Full text] [doi: 10.1186/s12984-015-0090-9] [Medline: 26651329]
57. Khalaji Z, Nezakat Alhosseini M, Safavi Hamami S, Iwatsuki T, Wulf G. Optimizing Motor Learning in Older Adults. J Gerontol B Psychol Sci Soc Sci 2024;79(1):gbad120. [doi: 10.1093/geronb/gbad120] [Medline: 37591686]
58. Fischl C, Blusi M, Lindgren H, Nilsson I. Tailoring to support digital technology-mediated occupational engagement for older adults - a multiple case study. Scand J Occup Ther 2020;27(8):577-590 [FREE Full text] [doi: 10.1080/11038128.2020.1760347] [Medline: 32396419]
59. Anguera JA, Boccanfuso J, Rintoul JL, Al-Hashimi O, Faraji F, Janowich J, et al. Video game training enhances cognitive control in older adults. Nature 2013;501(7465):97-101 [FREE Full text] [doi: 10.1038/nature12486] [Medline: 24005416]
60. Gulde P, Schmidle S, Aumüller A, Hermsdörfer J. The effects of speed of execution on upper-limb kinematics in activities of daily living with respect to age. Exp Brain Res 2019;237(6):1383-1395. [doi: 10.1007/s00221-019-05507-0] [Medline: 30887078]
61. Kempermann G, Gast D, Gage FH. Neuroplasticity in old age: sustained fivefold induction of hippocampal neurogenesis by long-term environmental enrichment. Ann Neurol 2002;52(2):135-143. [doi: 10.1002/ana.10262] [Medline: 12210782]
62. Der G, Deary IJ. Age and sex differences in reaction time in adulthood: results from the United Kingdom Health and Lifestyle Survey. Psychol Aging 2006;21(1):62-73. [doi: 10.1037/0882-7974.21.1.62] [Medline: 16594792]
63. Williams RS, Biel AL, Wegier P, Lapp LK, Dyson BJ, Spaniol J. Age differences in the Attention Network Test: Evidence from behavior and event-related potentials. Brain Cogn 2016;102:65-79. [doi: 10.1016/j.bandc.2015.12.007] [Medline: 26760449]
64. Novikov NI, Brazhnik ES, Kitchigina VF. Pathological correlates of cognitive decline in Parkinson's disease: from molecules to neural networks. Biochemistry 2023;88(11):1890-1904. [doi: 10.1134/S0006297923110172] [Medline: 38105206]
65. Kochan NA, Bunce D, Pont S, Crawford JD, Brodaty H, Sachdev PS. Is intraindividual reaction time variability an independent cognitive predictor of mortality in old age? Findings from the Sydney Memory and Ageing Study. PLoS One 2017;12(8):e0181719 [FREE Full text] [doi: 10.1371/journal.pone.0181719] [Medline: 28792946]
66. Zhu S, Sui Y, Shen Y, Zhu Y, Ali N, Guo C, et al. Effects of virtual reality intervention on cognition and motor function in older adults with mild cognitive impairment or dementia: a systematic review and meta-analysis. Front Aging Neurosci 2021;13:586999 [FREE Full text] [doi: 10.3389/fnagi.2021.586999] [Medline: 34025384]
67. Bixter M, Blocker K, Rogers W. Enhancing social engagement of older adults through technology. In: Pak R, McLaughlin AC, editors. Aging, Technology and Health. London, UK: Academic Press; 2018:179-214.
68. Kalantari S, Bill Xu T, Mostafavi A, Lee A, Barankevich R, Boot WR, et al. Using a nature-based virtual reality environment for improving mood states and cognitive engagement in older adults: a mixed-method feasibility study. Innov Aging 2022;6(3):igac015 [FREE Full text] [doi: 10.1093/geroni/igac015] [Medline: 35592668]
69. Arlati S, Di Santo SG, Franchini F, Mondellini M, Filiputti B, Luchi M, et al. Acceptance and usability of immersive virtual reality in older adults with objective and subjective cognitive decline. J Alzheimers Dis 2021;80(3):1025-1038. [doi: 10.3233/JAD-201431] [Medline: 33646164]
70. Llanes-Jurado J, Marín-Morales J, Guixeres J, Alcañiz M. Development and calibration of an eye-tracking fixation identification algorithm for immersive virtual reality. Sensors 2020;20(17):4956 [FREE Full text] [doi: 10.3390/s20174956] [Medline: 32883026]

XSL•FO
RenderX

71. Campo-Prieto P, Cancela-Carral JM, Rodríguez-Fuentes G. Immersive virtual reality reaction time test and relationship with the risk of falling in Parkinson's disease. Sensors 2023;23(9):4529 [FREE Full text] [doi: 10.3390/s23094529] [Medline: 37177733]
72. Negu A, Matu S, Sava FA, David D. Virtual reality measures in neuropsychological assessment: a meta-analytic review. Clin Neuropsychol 2016;30(2):165-184. [doi: 10.1080/13854046.2016.1144793] [Medline: 26923937]
73. Hao J, Li Y, Remis A, He Z, Yao Z, Pu Y. Performance-based outcome measures of upper extremity in virtual reality and telerehabilitation: a systematic review. Neurol Sci 2024;45(3):977-986. [doi: 10.1007/s10072-023-07158-0] [Medline: 37904015]
74. Aderinto N, Olatunji G, Abdulbasit MO, Edun M, Aboderin G, Egbunu E. Exploring the efficacy of virtual reality-based rehabilitation in stroke: a narrative review of current evidence. Ann Med 2023;55(2):2285907 [FREE Full text] [doi: 10.1080/07853890.2023.2285907] [Medline: 38010358]
75. Buyle M, Jung Y, Pavlou M, Gonzalez SC, Bamiou D. The role of motivation factors in exergame interventions for fall prevention in older adults: a systematic review and meta-analysis. Front Neurol 2022;13:903673 [FREE Full text] [doi: 10.3389/fneur.2022.903673] [Medline: 35989930]
76. Resnick B. Efficacy beliefs in geriatric rehabilitation. J Gerontol Nurs 1998;24(7):34-44. [doi: 10.3928/0098-9134-19980701-08] [Medline: 9801529]
77. Hughes S, Warren-Norton K, Spadafora P, Tsotsos L. Supporting optimal aging through the innovative use of virtual reality technology. Multimodal Technol Interact 2017;1(4):23. [doi: 10.3390/mti1040023]
78. Healy D, Flynn A, Conlan O, McSharry J, Walsh J. Older adults' experiences and perceptions of immersive virtual reality: systematic review and thematic synthesis. JMIR Serious Games 2022;10(4):e35802 [FREE Full text] [doi: 10.2196/35802] [Medline: 36472894]
79. Lin CX, Lee C, Lally D, Coughlin JF. Impact of virtual reality (VR) experience on older adults' well-being. In: Zhou J, Salvendy G, editors. Human Aspects of IT for the Aged Population. Applications in Health, Assistance, and Entertainment. Cham, Switzerland: Springer; 2018.
80. Matsangidou M, Solomou T, Frangoudes F, Papayianni E, Pattichis CS. Offering outworld experiences to in-patients with dementia through virtual reality: mixed methods study. JMIR Aging 2023;6:e45799 [FREE Full text] [doi: 10.2196/45799] [Medline: 37656031]

Abbreviations

BBT-VR: immersive virtual reality version of the Box and Block Test
CONSORT: Consolidated Standards of Reporting Trials
iVR: immersive virtual reality
niVR: nonimmersive virtual reality
SAT: speed-accuracy trade-off
SPARC: spectral arc length
TREND: Transparent Reporting of Evaluations with Nonrandomized Designs
VR: virtual reality

Edited by J Wang; submitted 05.07.24; peer-reviewed by S Balula, L Wang, S Pardini; comments to author 11.11.24; revised version received 13.12.24; accepted 30.01.25; published 03.03.25.

Please cite as:
Everard G, Declerck L, Lejeune T, Edwards MG, Bogacki J, Reiprich C, Delvigne K, Legrain N, Batcho CS
A Self-Adaptive Serious Game to Improve Motor Learning Among Older Adults in Immersive Virtual Reality: Short-Term Longitudinal Pre-Post Study on Retention and Transfer
JMIR Aging 2025;8:e64004
URL: https://aging.jmir.org/2025/1/e64004
doi:10.2196/64004
PMID:40053708

©Gauthier Everard, Louise Declerck, Thierry Lejeune, Martin Gareth Edwards, Justine Bogacki, Cléo Reiprich, Kelly Delvigne, Nicolas Legrain, Charles Sebiyo Batcho. Originally published in JMIR Aging (https://aging.jmir.org), 03.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Socioculturally Appropriate Internet-Based Geriatric Care Model for Older Adults Living With HIV: Experience-Based Co-Design Approach

Kristina M Kokorelias[1,2,3,4], PhD; Marina B Wasilewski[2,3,5], PhD; Dean Valentine[1]; Andrew D Eaton[6,7], MSW, PhD; Erica Dove[3], MSc; Paige Brown[8], BSc; Stuart McKinlay[8], BHSc; Christine L Sheppard[7], MSW, PhD; Esther Su[1]; Hardeep K Singh[2,3,9], PhD; Ashley Flanagan[4], PhD; Alice Zhabokritsky[10,11,12], MD; Reham Abdelhalim[13], MD, PhD; Rabea Parpia[14], MD; Rahel Zewude[11], MD; Laura Jamieson[15], MSc; Sharon Walmsley[10,11,12], MSc, MD; Luxey Sirisegaram[1], MD

[1]Section of Geriatric Medicine, Department of Medicine, Sinai Health System and University Health Network, Toronto, ON, Canada
[2]Department of Occupational Science and Occupational Therapy, University of Toronto, Toronto, ON, Canada
[3]Rehabilitation Sciences Institute, University of Toronto, Toronto, ON, Canada
[4]National Institute on Ageing, Toronto Metropolitan University, Toronto, ON, Canada
[5]St John's Rehab Research Program, Sunnybrook Health Sciences Centre, Toronto, ON, Canada
[6]Faculty of Social Work, Saskatoon Campus, University of Regina, Saskatoon, SK, Canada
[7]Factor-Inwentash School of Social Work, University of Toronto, Toronto, ON, Canada
[8]Undergraduate Medical Education, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada
[9]KITE, Toronto Rehabilitation Institute, University Health Network, Toronto, ON, Canada
[10]Infectious Diseases, Department of Medicine, University Health Network, Toronto, ON, Canada
[11]Division of Infectious Diseases, Department of Medicine, University of Toronto, Toronto, ON, Canada
[12]CIHR Canadian HIV Trials Network, Vancouver, BC, Canada
[13]Burlington Ontario Health Team, Joseph Brant Memorial Hospital, Burlington, ON, Canada
[14]St Michael's Hospital, Toronto, ON, Canada
[15]Ontario Federation of Indigenous Friendship Centres, Toronto, ON, Canada

Corresponding Author:
Kristina M Kokorelias, PhD
Section of Geriatric Medicine
Department of Medicine
Sinai Health System and University Health Network
Suite 18-208
600 University Avenue
Toronto, ON, M5G 1X5
Canada
Phone: 1 416 586 4800 ext 4374
Email: Kristina.kokorelias@sinaihealth.ca

Abstract

Background: Older adults living with HIV face challenges accessing regular geriatric care, and while virtual care services could offer a solution, they may come with limitations.

Objective: This study aimed to co-design a culturally appropriate virtual care model tailored to older adults' needs using the experience-based co-design methodology.

Methods: We used a qualitative, experience-based co-design approach with 19 older adults living with HIV. The process involved 3 phases: identifying needs through interviews and questionnaires, codeveloping a care model prototype through focus groups and a workshop, and refining the model using feedback from a world café format. Data were analyzed using thematic content analysis.

XSL•FO
RenderX

Results: The co-design process led to a virtual care model prototype that directly addressed participants' key needs. These included personalized communication methods, simplified technology interfaces for easier access, and culturally responsive care practices. Participants emphasized the importance of privacy in virtual consultations, flexible scheduling to accommodate health fluctuations, and ongoing support for managing both HIV and aging-related conditions. Their feedback shaped a model designed to bridge service gaps, offering a more inclusive, accessible, and patient-centered approach to virtual geriatric care.

Conclusions: This study co-designed a potential virtual geriatric care model grounded in the experiences of older adults living with HIV. By integrating participants' insights throughout the design process, the model offers a promising approach to improving care for this vulnerable population. Future directions for research to test this model are proposed.

(JMIR Aging 2025;8:e67122) doi:10.2196/67122

KEYWORDS

older adults; HIV; virtual care; co-design; qualitative research

Introduction

Background

The global population is aging, with the number of people aged ≥50 years projected to double by 2050, reaching approximately 2 billion individuals worldwide [1]. The aging global population also includes a growing number of older adults living with HIV [2]. Due to advancements in combination antiretroviral therapy, the life expectancy of people living with HIV has significantly increased [3], leading to the first generation of people living with HIV now reaching geriatric age [4]. This demographic shift brings unique challenges as older adults living with HIV experience a complex interplay of age-related comorbidities and HIV-related health issues [5-7].

Older adults living with HIV encounter a myriad of challenges that significantly impact their geriatric care. Medically, these individuals are more likely to experience comorbid conditions, including cardiovascular disease [6,8,9], diabetes [6,10], and osteoporosis [11-13], which can complicate HIV management and treatment [14]. These comorbidities complicate treatment regimens and require coordinated care, which may not always be available [15,16]. The specialization of health care often leads to fragmented care, where health care providers focus solely on HIV management without considering the broader geriatric care needs of the patient, such as mobility, nutrition, and mental health [17,18]. Furthermore, the physiological changes associated with aging, combined with the long-term effects of HIV and its treatment, can exacerbate these comorbidities, leading to a higher burden of illness and increased health care use [19]. Psychologically, older adults living with HIV may experience higher rates of depression [20-22], anxiety [23,24], and cognitive impairment [25-27] than their younger counterparts living with HIV. The stigma associated with both aging and HIV can combine and contribute to social isolation, discrimination in health care settings, and declines in mental health and overall well-being [28-30]. In addition, older adults may encounter ageism within health care settings, which can negatively affect the quality of care they receive and their willingness to seek further treatment [31-33]. Thus, the increase in older people living with HIV underscores the increasing need for comprehensive geriatric care that addresses the complex health needs of older adults [34,35].

HIV is highly stigmatized, and this can exacerbate other experiences of stigma and discrimination [32,36,37]. Notably, geriatric care (ie, care delivered specifically for older adults) is most effective when it meets the sociocultural needs of its patients [38,39]. Socioculturally appropriate care refers to health care that respects and responds to the cultural and social needs of patients, integrating their backgrounds, beliefs, and values into their care plans [40-43]. This approach is crucial for improving patient outcomes, enhancing satisfaction with care, and reducing health disparities [44-46]. However, delivering socioculturally appropriate care, particularly in the intersection of geriatrics and HIV, presents unique challenges [43]. These include a lack of cultural competency training [47,48], insufficient resources [49-51], language barriers [52,53], and implicit biases among health care providers [54-56], which can lead to misunderstandings and suboptimal care [57]. Moreover, individuals aging with HIV may come from diverse countries and have varying sexual identities and unique histories of trauma, discrimination, and stigma that differ from those of the general population [36,37,58]. Cultural mismatches such as failing to address the stigma associated with HIV in certain cultures can result in reduced patient engagement, lower treatment adherence, and worse health outcomes [36,59]. Therefore, addressing these challenges is essential to provide effective geriatric care to diverse people living with HIV.

Virtual care holds significant potential to address some of the challenges faced by older people living with HIV [60-63]. For instance, virtual care can help facilitate regular check-ins and monitoring without the need for frequent in-person clinic visits, which is particularly beneficial for those with mobility issues or transportation obstacles or for those living in remote areas [64-66]. As one example, virtual interventions for rural older veterans living with HIV have been shown to improve access to care and patient health outcomes, including quality of life, medication adherence, depressive symptoms, internalized stigma, and loneliness [67]. Virtual platforms can also provide a discrete way for patients to access care, potentially reducing stigma-related barriers to seeking health care in person [68-70].

While virtual care has revolutionized health care delivery, it also presents several challenges that need to be addressed for effective implementation [71,72]. One major issue is the digital divide, where limited access to technology and reliable internet connectivity, especially in rural and underserved areas, hampers the ability of patients to engage in virtual care [73-75]. In

XSL•FO
RenderX

addition, digital literacy poses a significant barrier as many patients, particularly older adults, may lack the knowledge and skills to navigate digital health platforms effectively [17,73,76]. There is a need to develop specialized socioculturally appropriate virtual geriatric care models tailored to the unique needs of diverse older people, including those living with HIV. Geriatric care refers to the specialized medical and supportive care tailored to meet the complex health and social needs of older adults, encompassing chronic disease management, functional support, and holistic well-being [50].

Delivering geriatric care virtually for people living with HIV presents both opportunities and significant challenges. While virtual care can enhance access to health care for older adults living with HIV, particularly in underserved or rural areas, it also requires careful consideration of various factors. One of the primary challenges is the complexity of managing multimorbidity, a hallmark of aging in people living with HIV, which often involves the treatment of multiple chronic conditions, including cardiovascular disease, diabetes, and mental health disorders [77,78]. Virtual care platforms may struggle to capture the full range of physical and cognitive assessments needed to effectively manage these conditions. In addition, older adults with HIV may face difficulties with digital literacy, which can hinder their ability to use virtual care tools effectively [79]. This is particularly true for older adults who may already experience cognitive decline or sensory impairments. The quality of the therapeutic relationship can also be affected in virtual settings as the absence of in-person interactions may reduce the ability of health care providers to observe subtle nonverbal cues and create a trusting environment [79]. Moreover, people living with HIV may have unique psychosocial challenges, such as stigma and social isolation, which can be exacerbated by virtual care [79]. Addressing these challenges requires the development of specialized, culturally competent virtual care models that incorporate both technological solutions and social support structures to ensure holistic, patient-centered care for older adults living with HIV [43].

Co-design is a valuable approach in developing and refining health interventions and services as it actively involves key players in the design process to ensure that the resulting solutions are highly relevant and effective for end users [80]. By engaging those who will ultimately benefit from these interventions, co-design fosters a deeper understanding of their needs, preferences, and sociocultural contexts, leading to more tailored and impactful outcomes [80,81]. By involving older adults in the design process, the resulting solutions will be more relevant and effective in addressing their unique health challenges and sociocultural contexts. *Co-design* promotes cultural sensitivity by incorporating the perspectives and experiences of diverse patient populations into the design and delivery of health services [82]. This is particularly important for older people living with HIV, who may face additional stigma and barriers underpinned by the intersection of sociocultural considerations and demographic factors such as age [83]. Addressing the gaps in the literature will ensure that the unique health challenges and sociocultural contexts of older people living with HIV are considered, leading to more effective and inclusive geriatric-HIV health care interventions [43].

Objectives

This paper reports on the co-design of a culturally appropriate virtual care model for older adults living with HIV using an experience-based co-design (EBCD) approach [84]. By involving older people living with HIV, this study sought to create a person-centered care model that enhances their health care experience and outcomes [85].

Methods

Study Design

In this community-based participatory research (CBPR), we used a qualitative EBCD methodology [84] that integrated participants as co-designers (details can be found in a published protocol [43]). CBPR is an approach that emphasizes collaboration between researchers and community members throughout the research process [86,87]. It aims to address community-identified needs and promote social change by integrating knowledge and action for social justice [86,88]. Co-design involves meaningful engagement of end users in the research design process, encompassing participation at every stage of the research and varying in intensity from passive involvement to highly active participation [84,89-91]. To achieve our research objective of co-designing a geriatric-HIV virtual care model, we adapted and used a variety of participatory EBCD methods, including interviews, focus groups, scenario design, and world cafés [84,92]. We briefly describe our methods in the following sections.

Ethical Considerations

Ethical considerations in CBPR include ensuring informed consent, where community members fully understand the research process and its implications [93]. This study was approved by the Sinai Health Research Ethics Board (approval 23-0106-E). All participants provided written informed consent to take part in the study, be audio recorded, and have anonymous quotations published. The participants were given the opportunity to have access to a translator and ask to clarify any concerns before signing the consent form. Participants in the focus groups and workshop were provided with refreshments and lunch and were reimbursed for their travel expenses. Participants in all phases also received an honorarium of $43 USD for their time, compensated at an hourly rate. Our study also emphasized community ownership of research findings, allowing the advisory committee to make decisions about how results are disseminated. All data shared with the advisory committee were anonymous.

Researcher Positionality

The research team (n=18) consisted of Canadian experts in co-design methodologies with diverse backgrounds and expertise in HIV research, equity-informed health care, health service research, and implementation science. This interdisciplinary team included PhD-trained researchers and clinicians specializing in geriatrics, family medicine, and infectious diseases, as well as occupational therapists, social workers,

XSL•FO
RenderX

health service administrators, and peer researchers with lived experience.

A community-based advisory team (n=10) was established before protocol development to ensure ongoing engagement throughout the project. This advisory team comprised knowledge users such as administrators from nonprofit organizations serving older persons living with HIV, including shelter staff, health care providers, charity organizations, and community centers. In addition, the advisory team included older individuals living with HIV and clinicians specializing in geriatric care. Regular individual meetings with principal investigators were conducted throughout the research process.

Recruitment and Participants

Our study sample was distinct from our advisory committee and included older individuals who self-identified as HIV positive; were aged ≥50 years; and resided in Ontario, Canada. We purposively sought diversity across various dimensions: (1) sex and gender, (2) age, (3) ethnicity and race, (4) socioeconomic status, (5) previous use of virtual geriatric care (ie, yes or no), (6) geographical location (rural vs urban), (7) time living with HIV, (8) non-English first language, and (9) level of educational attainment.

We used quota and purposive sampling to ensure a representative sample [94,95]. Our advisory team engaged in community outreach at relevant meetings and events to assist with recruitment, leveraging their networks and organizations. We also posted study flyers at community-based organizations, religious institutions, and culturally oriented events. Finally, our advisory team's websites and social media platforms were used to promote the study. All participants contacted the research coordinator to learn more about the study, and eligibility was confirmed based on the demographic characteristics to ensure alignment with our purposeful sampling technique.

Data Collection

Data collection occurred in 3 phases, which are outlined in this section and published in a previous protocol [43]. No deviations were made, aside from the fact that no focus groups occurred in phase 1. Phase 1 focused on understanding participants' needs and perspectives through semistructured interviews. The interviews were conducted by a trained research assistant with support from the primary author. The interviews focused on participants' experiences throughout their illness trajectory, current and anticipated needs for geriatric care, and priorities for virtual care. To ensure effective purposive quota sampling, demographic information was collected from participants at the time of recruitment, allowing us to determine whether they met the specific criteria for inclusion in the study. This demographic questionnaire captured diverse lived experiences and socioeconomic data relevant to this study's aims. Once participants were recruited, all interviews were audio recorded, professionally transcribed, and reviewed for accuracy by a research assistant. The narrative summaries of the interviews were created by the first author based on both the interviewer's notes and recollections and the transcripts of the interviews. These summaries were developed to provide a comprehensive overview of each interview, highlighting key themes, insights, and participant perspectives. Following their creation, the narrative summaries were shared with the research team and the advisory team via email and team meetings. The narrative summaries were included in the dataset for analysis and were considered part of the field and meeting notes, providing a synthesized account of key themes and insights from the interviews to support the interpretation and development of findings.

Before initiating the co-design process, we prioritized trust building to create a comfortable and inclusive environment for participants. To minimize power imbalances, we limited the number of researchers present, ensuring that the space felt less formal and more participant driven. In addition, we asked participants whom they wanted involved in the process, allowing them to shape the composition of the sessions. These efforts fostered a sense of ownership, encouraged open dialogue, and reinforced the collaborative nature of the co-design approach.

Phase 2 involved 2 focus groups and a workshop, all drawing on the same participants to apply an EBCD approach to develop a prototype for a culturally appropriate HIV virtual care model. The focus groups primarily aimed to gather in-depth qualitative data through structured discussions that allowed participants to share their experiences and insights about virtual care. In contrast, the workshop was more interactive and aimed at synthesizing the findings from the focus groups, where participants worked collaboratively to co-design and refine the elements of the HIV virtual care model. In preparation for phase 2, participants in the focus groups were provided with a detailed presentation summarizing the main needs identified in phase 1 before the workshop formally began. This presentation, based on comprehensive data collected during the initial phase, highlighted key insights into the unique challenges faced by older adults living with HIV. By sharing this information with the participants, we informed them about critical areas to address in the design process, ensuring that their feedback in the focus groups was grounded in the context of the identified needs. The focus groups and workshop were jointly led by a trained neutral facilitator and peer researcher who was trained through a structured program designed to build both facilitation and research skills. The training included an overview of the research objectives, ethical considerations, and techniques for creating a supportive and nonjudgmental environment for participants. The peer researcher was also trained in active listening, group dynamics, and how to manage sensitive topics, particularly those related to the experiences of older adults living with HIV. The workshop, guided by a modified 4D cycle of Appreciative Inquiry [96], aimed to identify strengths in existing care models and design principles for improved virtual care, accommodating both in-person and virtual attendance [97]. The modified 4D cycle of Appreciative Inquiry involves the phases of *Discovery*, *Dream*, *Design*, and *Destiny*, focusing on identifying strengths, envisioning an ideal future, cocreating actionable plans, and implementing sustained changes [96]. The *Definition* phase was also included for greater adaptability to specific contexts and goals. Discussions used visual tools such as concept mapping and word clouds [98] that were kept for analysis. All verbal focus group data were audio recorded and professionally

XSL•FO
RenderX

transcribed. Postworkshop reflections with the research and advisory teams guided phase 3.

In phase 3, there were 10 returning participants from phase 2 and 2 participants new to the study. The design principles for the culturally appropriate HIV virtual care model were developed using a world café format [99]. Small group discussions fostered dynamic exchanges, integrating feedback on proposed solutions and outlining recommendations for implementation and uptake. Facilitated by a peer researcher and supported by research team members, these sessions aimed to ensure participant engagement and generate robust data for ongoing project development [99]. Memos and reflexive notes were taken by the research team throughout the study. Transcripts and notes were anonymized by removing details that could lead to participant identification. No names of participants were used in the analysis or reporting of the results.

Data Analysis

The dataset comprised audio recordings, transcripts, field and meeting notes, reflexive notes, and physical artifacts designed by participants. The analysis team, consisting of research team members and select advisory committee members, conducted thematic content analysis [100,101]. The field and meeting notes were reviewed as part of the data analysis process, contributing to the identification of key themes. In phase 1, data analysis occurred concurrently with data collection to determine thematic saturation [102]. Thematic saturation was established when no new codes or concepts emerged from the data, indicating that additional interviews were unlikely to yield novel insights. Transcripts were coded using the NVivo software (version 14; QSR International) by the principal investigator and the research assistant. A codebook developed through inductive and deductive processes informed by the knowledge-to-action framework was used and refined iteratively through team meetings [103,104]. Physical artifacts created by participants during the workshop, including Post-it notes, drawings, and summary sheets, were also systematically analyzed and integrated into the overall data analysis process to enrich and deepen the understanding of participants' perspectives. Once scanned, the images were categorized based on themes and topics that emerged from the discussions. This allowed the research team to examine not only the content of the physical artifacts but also the ways in which participants visually expressed their ideas and emotions. The drawings and Post-it notes, for example, provided additional insights into participants' cognitive and emotional responses, offering a layer of qualitative data that complemented the verbal responses collected during the sessions. The images were analyzed using a thematic approach, identifying recurring patterns, symbols, and motifs that reflected key aspects of participants' experiences. Peer researchers were sent the preliminary findings 1 week ahead of the focus groups and workshop to enable thorough review and reflection before the subsequent phases. Themes were identified through comparative analysis and refined through team discussions, with a thematic map aiding in refining final themes to inform phase 2. Phase 2 and phase 3 continued with reflexive thematic content analysis to consolidate findings from the co-design workshop and refine themes for subsequent phases [101], including the world café discussions. The analysis refined the themes that emerged during these sessions and ensured that the feedback from the world café format was directly incorporated into the final design recommendations for the virtual care model. All researchers reviewed and endorsed the final version of the manuscript as coauthors to ensure the accuracy and completeness of the findings presented. A lay version of the manuscript was also sent to all members of the advisory team and participants of the study. Participants were invited to provide their feedback via email up to the point of manuscript submission.

Validation and Reliability Strategies

Rigor and trustworthiness in this study were ensured through several methodological strategies. First, we used a qualitative EBCD methodology, integrating participants as co-designers throughout the research process, which facilitated meaningful end-user engagement and ensured that participant perspectives were central to the study's design and outcomes. This approach was consistent with established principles of CBPR [86,105]. Second, we maintained methodological rigor by adhering strictly to our published protocol [43], which detailed the study procedures and ensured consistency in data collection and analysis methods. Our adherence to the protocol provided a structured framework for data collection and interpretation while remaining adaptable to participant and advisory committee input and emerging insights. This balance ensured that, while following a predefined approach, we also responded to and incorporated collaborators' feedback, thus enhancing the relevance and responsiveness of the methodology. Third, the study design included diverse methods that enhanced the richness and depth of the data collected and contributed to the triangulation of data sources [106,107]. The analysis team, consisting of research team members and selected advisory committee members, engaged in iterative coding and theme development supported by regular team meetings and peer review processes. This collaborative approach enhanced the reliability and credibility of the findings by ensuring that interpretations were grounded in the data and reflective of diverse perspectives [108]. Finally, transparency and reflexivity were maintained throughout the research process. That is, reflexive notes, memos, and detailed documentation of methodological decisions were kept, providing transparency in data interpretation and analysis. Regular interactions with the advisory team and participants, including sharing preliminary findings and seeking feedback, further strengthened the trustworthiness of this study by ensuring that the findings resonated with participant experiences and priorities [108].

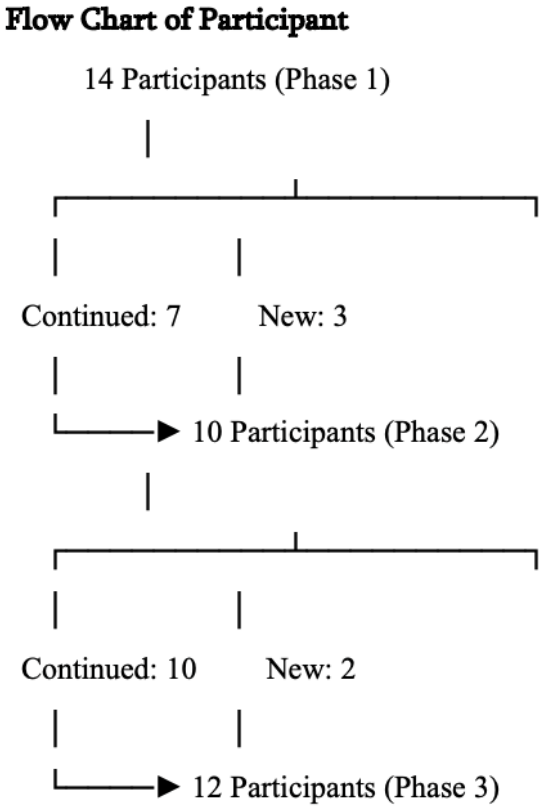
Results

Overview

The research team conducted the co-design process between June 2023 and June 2024. In total, 19 unique participants took part in this study, with 14 (74%) participants in phase 1, a total of 10 (53%) participants in phase 2 (n=7, 70% from phase 1), and 12 (63%) participants in phase 3 (n=10, 83% from phases 1 and 2). Multimedia Appendix 1 outlines the participant characteristics, and Figure 1 shows a flowchart of the participants.

XSL•FO
RenderX

Figure 1. Flowchart of participants.



Participants co-designed a culturally appropriate and flexible HIV virtual care model that was rooted in empathy and addressed both the clinical, psychological, and social elements and the emotional aspects of aging with HIV. Phase 1 aimed to understand participants' needs regarding virtual care, revealing four key findings: (1) the desire for accessible and integrated health care; (2) the complications caused by sociocultural and demographic factors; (3) the benefits of peer support without replacing formal care; and (4) the potential of virtual care to address some barriers, although with caveats related to access and technology. Participants expressed a strong need for easily accessible, coordinated care, particularly for older adults living with HIV who face complex, overlapping health and social challenges. They highlighted the difficulties of navigating the health care system, noting that virtual care could reduce barriers but only if it accounts for digital literacy and technology access. In addition, while peer support was valued, participants cautioned against overreliance on it, emphasizing that virtual care should supplement, not replace, formal services. They also pointed out that virtual care could improve access, especially for individuals facing logistical barriers, but it must be designed with flexibility and continuity in mind. In phase 2, the focus shifted to ideating solutions for the virtual geriatric-HIV care model. Participants suggested a multimodal care approach encompassing various forms of care delivery (eg, in-person and virtual visits and digital health records). They also stressed the importance of education and clear communication regarding technology use, ensuring that older adults are informed and confident in using virtual platforms. Privacy and data security were significant concerns, with participants requesting transparency about who accesses their information and how it is used. In addition, participants emphasized the need for reminder systems to help maintain continuity of care and reduce missed appointments. These solutions were aimed at improving both the practicality and trustworthiness of virtual care for older adults living with HIV. Phase 3 focused on co-designing actionable plans, with participants emphasizing the importance of linguistic and cultural equity in virtual care. They stressed that care models should be inclusive of diverse populations, particularly those who face language barriers. Participants highlighted the need for culturally appropriate services to ensure equitable care, pointing out that older adults from different cultural backgrounds often experience compounded difficulties when accessing health care services. The plans developed during this phase aimed to refine and prioritize these solutions to create a more inclusive and effective virtual care model for older adults living with HIV. Throughout the following sections, we illustrate the findings using quotes labeled with participant ID, race or ethnicity, gender, and age.

Phase 1: Understanding Participants' Needs Regarding Virtual Care

We identified the following key and interrelated insights: (1) older adults living with HIV want accessible and integrated care, (2) intersecting sociocultural and demographic factors complicate receipt of care for older adults living with HIV, (3) peer support is beneficial but should not replace formal health care services and supports, and (4) virtual care can address barriers to care with access caveats.

XSL•FO
RenderX

Older Adults Living With HIV Want Accessible and Integrated Care

Participants emphasized the need for health care systems to prioritize services that are easily accessible and tailored to the unique health challenges faced by older individuals living with HIV. Participants noted that they often had complex, overlapping medical and social needs that required coordinated care across multiple disciplines. While most participants described positive experiences with previous care, they also discussed challenges obtaining information and reaching the appropriate care provider. Several participants specifically recounted situations in which accessing services and support felt very difficult. One participant shared the following:

> *You can go to a family doc[tor] if you have one and that is so great—again, if you have one. But the issue becomes when the doc refers you to one specialist who refers you to another then back to your family doc and you repeat because no one knows how to care for you.* [Interview participant 11; man; White individual; aged 52 years]

Therefore, participants described desiring a health care system with the ability to easily access services from care providers.

Intersecting Sociocultural and Demographic Factors Complicate Receipt of Care for Older Adults Living With HIV

Participants highlighted that they often faced systemic barriers that hindered their ability to access health care services. These challenges were exacerbated by age-related biases within the system, which was not designed to fully address or accommodate their unique needs. Participants highlighted that those with lower incomes may experience additional obstacles, including lower access to resources; heightened stigma; and greater socioeconomic disadvantages that impact their health, such as the inability to purchase both food and medication. Participants agreed that virtual care has the potential to alleviate some of these systemic barriers by providing more accessible and flexible health care options. However, for virtual care to be an effective solution, it must address the challenges associated with digital literacy and technology access. One participant pointed out the following:

> *The level that we're to access the services at, is a bit crazy and some might have challenges. Like this morning, I tried to go on the touch screen, and I can't get it on to use my phone. And then I'm seeing that there is no Internet service.* [Interview participant 5; man; Black individual; aged 56 years]

Moreover, to ensure that virtual care is equitable, it is crucial that health care providers are aware of and address the specific needs and preferences of older adults. For instance, participants noted that they preferred a larger screen for virtual consultations, such as a tablet, over smaller devices such as phones.

Peer Support Is Beneficial but Should Not Replace Formal Health Care Services and Supports

Support systems among peers were noted as critical in the management of chronic conditions among older people living with HIV. Peer support groups offered participants emotional and practical guidance from individuals who shared similar experiences, fostering a sense of community and reducing feelings of isolation. Peer interactions were also thought to provide valuable insights into managing HIV and navigating the health care system as an older person as they offered lived experiences and practical advice that complemented professional care. One participant noted the positive impact of peer support, stating the following:

> *I found that connecting with others who are dealing with the same condition really helped me cope better and stay informed about my health.* [Interview participant 3; man; South Asian individual; aged 82 years]

In the context of peer support, health management and the pursuit of independence emerged as critical themes in the experiences of older adults living with HIV, particularly within a virtual care framework. Participants relied heavily on peer support to navigate the health care system, highlighting the need for virtual care models that offer tailored resources and programs to address their unique needs. One participant emphasized the value of community-based programs, noting the following:

> *First of all, it's a safe place for LGBTQ members. And it's a positive place for me because it's 50 plus. Basically, it focuses on age and being gay and what more can we want. Canada is great—it won't discriminate by colour in medical settings, but sometimes we can't find what we need.* [Interview participant 11; man; White individual; aged 52 years]

Despite the availability of such programs, challenges remained, particularly in accessing mental health services. The participants wished for more targeted online resources for aging, stress, or depression beyond medication. Furthermore, virtual care could enhance independence in health management by facilitating regular interactions with health care providers through digital platforms. Participants expressed a need for frequent virtual check-ins, such as weekly phone calls, to better monitor and manage their health conditions and maintain a sense of connection and support.

Participants noted that, while peer support was invaluable, they were cautious that their reliance on individuals with lived experience was due to systemic gaps in health care provision and the lack of formal support structures for addressing the complex interplay of health and social circumstances affecting older adults living with HIV. One participant remarked the following:

> *If I get sick, then I lose my job, and then the chance of losing my housing rise exponentially and so basically I have to have friends who have a house or know someone who is hiring to pay for medicine.* [Interview participant 1; woman; White individual; aged 57 years]

The same participant shared the following:

> *If you have money, you'll get the healthcare in the time that you need it. And if you don't, well, then good luck.*

XSL•FO
RenderX

Another participant shared the following:

> *When you're first diagnosed, all you're focused on is survival. You aren't trying to find services; you meet people who tell you what services exist. But the healthcare system, the focus isn't on us, who've been around for a long time, and now we need different supports in place than we did 20 years ago. So, we're stuck with the same services.* [Interview participant 2; woman; First Nations individual; aged 57 years]

Virtual Care Can Address Barriers to Care With Access Caveats

Participants noted that the integration of virtual care has the potential to alleviate some of these challenges by offering more flexible and accessible health care options, thereby significantly reducing the logistical challenges associated with in-person visits, as noted by a participant who appreciated the time savings:

> *I had a bone density scan...when I had it, he wanted to make an appointment, and I had the choice between virtual or going down to the office. I saved like 4 hours.* [Interview participant 12; man; First Nations individual; aged 51 years]

Another participant shared the following:

> *When the place is very cold and you got to catch two bus, three busses to meet your doctor, and pay for it each time, we can do virtual. Then when it's nice outside whether we don't mind going and spending the time and go out and see.* [Interview participant 10; woman; mixed heritage; aged 54 years]

However, one participant wisely pointed out the following:

> *But it has to be only when minor consultation is required...it should [not] replace the face-to-face consultation.* [Interview participant 9; man; White individual; aged 59 years]

Participants emphasized the importance of effective communication and continuity of care within the virtual geriatric-HIV care model. They expressed a strong desire for seamless interactions among care providers to prevent disruptions in their treatment and communication. This continuity was described as crucial for maintaining the stability of their HIV management, including regular blood work and medication adjustments. Although participants generally reported positive interactions with health care providers, challenges arose in distinguishing between symptoms related to HIV and those associated with normal aging. One participant remarked the following:

> *Sometimes it's hard to tell whether a symptom is due to the HIV or just getting older. Clear communication and understanding from my healthcare provider are essential to figuring out what's really going on.* [Interview participant 9; man; White individual; aged 59 years]

To address these complexities, participants recommended that novel care models integrate mechanisms for clear, ongoing communication and coordination among health care providers, ensuring that specialized care is tailored to meet the intricate needs of older adults living with HIV.

Phase 2: Ideating Solutions

Participants suggested improvements to current care models for older adults living with HIV and proposed their wishes regarding virtual care through a series of activities.

Importance of Flexible Care Delivery

Participants highlighted the importance of a multimodal care approach to enhance coordinated care and access to health information for older adults living with HIV. One participant explained the following:

> *...we need it [virtual care] to be offered in a lot of ways, so not just phones because not everyone has one, not everyone has a cell phone, maybe computer, so offer a lot of ways like a phone, cellphone, tablet, computer.* [Phase-2 participant 1; man; White individual; aged 56 years]

Participants also noted that the approach should integrate various modes of care delivery, such as in-person consultations, virtual visits, and digital health records, to create a more seamless, continuous, and comprehensive health care experience. One participant shared the following:

> *...you can't just have it be one approach, sometimes you need to come in, sometimes I might need a virtual visit, sometimes I might want to send information, we need options.* [Phase-2 participant 4; woman; Black individual; aged 71 years]

Education and Communication Regarding Technology Use and Privacy

The need for clear communication regarding new technologies was a recurring theme among participants. They emphasized that educational demonstrations are essential in helping patients understand and effectively use new technological tools. One participant said the following:

> *...the worst thing someone can do is if you assume we know what we are doing with technology, so it doesn't have to be a doctor but some instructions. I sometimes use Zoom but then they say go online and upload, I have no idea how to do that.* [Phase-2 participant 6; man; White individual; aged 62 years]

Transparency regarding technology use in virtual care emerged as a crucial concern for participants. Participants emphasized the need for clear communication about who accesses patient information, how it is used, and the procedures for withdrawing consent. In the context of virtual care, the digital environment can heighten concerns about privacy and confidentiality, making it imperative to establish transparent practices in data privacy and security. Participants indicated that, in virtual settings, where information is often shared and stored electronically, patients may feel more vulnerable and uncertain about their data. Therefore, ensuring that patients are well informed about the handling of their health information is essential for building trust and fostering a sense of comfort with using virtual care services. One participant explained the following:

XSL•FO
RenderX

> *...sometimes they think we are old, so we don't understand, but I know about technology and so I want to know who is going to look at my stuff in order to trust you. If I don't know, I won't trust this [virtual care].* [Phase-2 participant 8; woman; Indian-Caribbean individual; aged 54 years]

Participants also underscored the need for transparency regarding the environment in which virtual care appointments are conducted. They expressed concerns about ensuring that these appointments occur in private and secure settings to protect patient confidentiality. Specifically, participants highlighted the importance of knowing whether providers are conducting virtual appointments from locations that guarantee privacy and maintain the integrity of sensitive health information:

> *I don't want the doctor to be like talking about HIV with their friends in the background or at Starbucks.* [Phase-2 participant 8; woman; Southeast Asian individual; aged 70 years]

Reminder Systems

Participants suggested the implementation of reminder systems for appointments and follow-ups to ensure continuity of care and reduce missed appointments. Such systems can play a crucial role in maintaining regular contact between patients and health care providers, thereby improving adherence to treatment plans and enhancing health outcomes. By leveraging technology to provide timely reminders, health care systems can support patients in managing their care more effectively and reduce the risk of lapses in treatment. One participant shared the following:

> *...maybe it's age, maybe it's life, but anything to make it easier is good. Remind me when to log in, when to book an appointment, what to do, who to see, even the name, just anything to make life easier when it comes to doctors because sometimes, I see so many and now maybe a geriatrician.* [Phase-2 participant 4; woman; Black individual; aged 71 years]

Phase 3: Designing the Action Plan

Overview

Phase 3 entailed participants co-designing and planning the necessary actions and initiatives to achieve the desired future for virtual geriatric-HIV care as outlined in phase 2. Participants concluded phase 3 by refining the ideas and virtual geriatric care models co-designed in phase 2 and phase 3. This involved prioritizing linguistic and cultural equity.

Linguistic and Cultural Equity

Participants highlighted the need for the virtual geriatric care model to incorporate linguistic and cultural considerations to ensure equitable care for the population of older adults living with HIV. Participants shared stories exemplifying the diversity of the geriatric-HIV community (inclusive of several underserved populations) and emphasized the importance of increasing services to these populations. One participant stated the following:

> *...not everyone speaks English, especially when they come here and so we need to give back and help them. It's like if you dropped someone in their countries, it's like worse than a maze.* [Phase-3 participant 11; man; White individual; aged 69 years]

Participants noted that there needs to be health- and technology-related information in multiple languages and that service providers should make translation services known to their patients. Participants also suggested that health care providers offer education that addresses many of the culturally related myths and concerns of individuals about aging with HIV. One participant said the following:

> *...doctors and nurses need to look at who they are serving that is different than them. Not everyone understands HIV is not a life sentence so tell certain people that. Or maybe someone eats food that is different, not everyone understands quinoa. Maybe recommend cultural food.* [Phase-3 participant 1; woman; Black individual; aged 51 years]

Consider Affordability

Participants were broadly aware of the fact that the cost of technology could be a barrier to many older adults living with HIV. Participants highly recommended that geriatric-HIV virtual care models be made more accessible by offering subsidized internet access or technological devices. Participants also shared that some patients may have never used technology before virtual care due to high costs and, therefore, may need training on how to best use the technology for care. Participants recognized that health care providers may not have the capacity to teach their clients, and as such, someone else in the health care organization should. For instance, participants suggested that perhaps calling a patient on a landline or inviting the patient to in-person training could help support their use of virtual modalities for health care in the future. One participant described the following:

> *...maybe invite someone to a place in person just even once or call them and walk them through something like how to turn their video on or how to send a text to a doctor.* [Phase-3 participant 17; woman; Black individual; aged 54 years]

Participants emphasized the importance of early education on technology, as previously underscored in the other phases of the study.

Participants expressed significant concern about the high cost of HIV-related medications, particularly for those aged <65 years, who are ineligible for government-funded drug coverage in Ontario. Some participants felt that health care providers were not advocating enough for the use of lower-cost alternatives such as generic medications, which could be just as effective as brand-name drugs. For instance, one patient mentioned the following:

> *It [medication] can be the choice between food, rent or medications at times, so I want someone who maybe says a cheaper supplement or maybe what medication I can maybe take a different type of. Anything to keep the cost low.* [Phase-3 participant 9; woman; Asian individual; aged 70 years]

XSL•FO
RenderX

Participants indicated that a more proactive stance from providers in promoting cost-effective treatment options, both for HIV and age-related conditions, would improve their overall care experience.

Person-Centered Care for Age-Related Health Issues

Participants highlighted the importance of having health care providers who understood the dual challenges of aging and living with HIV. They emphasized that a key feature of an effective geriatric virtual care program would be the integration of expertise in both areas. This holistic approach was seen as essential for managing the complexities of age-related conditions alongside HIV, with many noting that current segmented specialist care often leads to fragmented treatment, polypharmacy, and inconsistent care coordination. An integrated model was viewed as a potential solution to improve care continuity and reduce these challenges.

Participants emphasized the importance of incorporating cognitive health screening and related support services into the virtual care model. This integration would address current gaps in HIV care by ensuring that the program not only focuses on medical treatment but also considers the emotional aspects of aging with HIV. Current gaps within HIV care were re-emphasized, and participants noted that, for uptake of the geriatric-HIV virtual care model to occur, participants needed to "feel as if the doctor or care professionals care of us as a person, not just a billable code or something, but to really understand how much [older adults living with HIV] struggle" (phase-3 participant 8; man; White individual; aged 67 years).

Patient Resilience and Support

Participants emphasized a virtual geriatric-HIV model of care that incorporates support networks to enhance patient resilience to living with HIV in older age. The involvement of trusted individuals such as existing care providers, friends, or service agencies was deemed essential to care by participants. This holistic support system helps patients navigate the challenges of their health journey, promotes a sense of connection and belonging, and strengthens their overall well-being. By embedding these peer support networks within the virtual care model, the program can provide a more robust and empathetic approach to patient care. One participant described the following:

> *I'd trust anything that [Community Agency] recommended, so if you bring your services to them so that they encourage them, I'll trust that it would help me.* [Phase-3 participant 4; woman; White individual; aged 58 years]

Moreover, participants emphasized that a virtual geriatric-HIV model of care should include features such as virtual support groups, peer mentoring programs, and online forums where individuals with similar experiences can connect and share insights. In addition, incorporating the peer mentorship system could allow patients to receive guidance and encouragement from those who have successfully managed similar health conditions.

Discussion

Principal Findings

This study was conducted to co-design a geriatric-HIV virtual care model for older adults living with HIV. The lack of tailored interventions that consider this population's specific needs and preferences necessitated a co-design approach. Using EBCD methodology [84,85,92], we engaged older adults living with HIV to collaboratively codevelop a virtual care model. Through interviews, focus groups, and a co-design workshop, participants identified key needs such as robust support systems, health management resources, and effective communication strategies. They cocreated and refined solutions including multimodal care options, technology transparency recommendations, and reminder systems.

Using co-design methodologies ensured that the voices and experiences of older adults living with HIV were central to the development and evaluation of the virtual care models. This inclusive, user-centered approach fosters solutions that are more likely to be accepted and effective as they are directly informed by the end users' needs and preferences. Engaging participants in the research process empowers them and enhances the relevance and applicability of the findings. A strength of this study was the inclusion of many English-as-a-second-language individuals and immigrants to Canada, who are typically underrepresented in research [109]. This diverse participant pool enhanced the transferability of the findings, ensuring that the developed virtual care models were relevant and accessible to a broader range of older adults living with HIV. By actively including these often-excluded groups, this study addresses potential disparities in health care access and provides insights into the unique needs and challenges faced by these populations. Thus, this research provides practical insights into the real-world application and potential barriers to adoption of virtual care models. This dual focus ensures that this study not only explores the theoretical potential of the interventions but also addresses the pragmatic challenges that might arise during implementation. As a result, the findings are highly relevant for informing policy and practice, ultimately contributing to the advancement of care for older adults living with HIV.

The findings presented in this paper have been echoed in various other studies with other populations. Research among rural and urban adults has shown high satisfaction with virtual care [110]. In addition, studies on virtual care for substance use indicate its ability to bridge service gaps, particularly in rural areas, reinforcing the importance of multimodal care options and effective communication [111]. Canadian case studies further demonstrate the necessity of involving users in the development of virtual care tools to ensure their effectiveness and transparency [112].

Our findings underscore the importance of integrating cultural and linguistic equity into virtual care models, which has substantial implications for enhancing health care accessibility and outcomes for older adults living with HIV. Research has shown that sociocultural stigma can severely impact health care engagement and access [36,113], highlighting the importance of designing solutions that are sensitive to these issues.

XSL•FO
RenderX

Involving older adults living with HIV in co-designing care solutions is essential to ensure that their needs are met [90]. Engaging older adults living with HIV in the development process ensures that the solutions are tailored to their specific needs and the challenges they face, thus enhancing their effectiveness and acceptance [90]. This study's findings have important implications for health policy, particularly regarding the integration of virtual care models for older persons living with HIV. Existing reviews have found that virtual geriatric-HIV programs are limited [17] despite their proposed benefits. Policy makers should consider the advantages of virtual care in addressing the unique needs of this population, especially in contexts in which traditional health care access is limited. Our study highlights that participants perceive virtual care to be feasible and acceptable for meeting their geriatric-HIV care needs, suggesting that policies should support the expansion of telehealth services and funding for technological infrastructure. In addition, policies should focus on inclusivity by ensuring that virtual care models are accessible to non-English speakers and diverse cultural groups who may have differing views on HIV care [50,114]. This could involve implementing language support services and cultural competency training for health care providers to enhance the effectiveness of virtual interventions. This could also involve funding geriatric health care settings such that they can hire multilingual staff, partner with translation services, and provide localized materials.

In practice, this study underscores the importance of adopting virtual care models as a viable alternative or complement to traditional in-person geriatric consultations for older persons living with HIV [67]. Health care providers should be encouraged to incorporate these models into their service offerings, with attention to user feedback and the co-design process highlighted in this study. Research emphasizes the need for health care professionals to collaborate with patients in co-design methodologies [115] to enhance patient participation in health care services. Projects involving creative workshops have shown that health care professionals can gain valuable insights and training through active participation in co-design processes [116]. In addition, discussions on EBCD underscore the necessity of equipping health care professionals with the tools and knowledge to effectively engage in patient-centered design and improvement initiatives [117]. Overall, training health care providers in these collaborative approaches is essential for fostering meaningful partnerships with patients in health care settings. For example, health care settings could employ staff trained in the use of virtual platforms and equipped with resources to support patients in navigating these technologies. Optimal methods for training older adults to use technology involve a combination of hands-on practice, clear instructions, and personalized support tailored to individual needs and cognitive abilities [118]. Effective approaches often include guided action and attention training, addressing fears, building confidence, and providing ongoing support, with a focus on practical, relevant applications that motivate learners [118,119]. Moreover, the integration of language support and culturally sensitive practices is crucial to ensure that virtual care is accessible to diverse populations. Other studies of virtual care have found that features such as multilingual support, screen readers, and voice commands can make virtual care more accessible to diverse patient populations [120]. Moreover, integrating peers into virtual care models can be achieved through comprehensive support, training, and collaboration across various providers [121].

To test the virtual geriatric care model, future research should use randomized controlled trials comparing virtual geriatric-HIV care with traditional care in terms of patient satisfaction, health outcomes, and cost-effectiveness. Specific technological platforms and features should also be evaluated to determine their effectiveness in addressing the unique needs of culturally diverse older adults living with HIV. In addition, implementation studies should examine barriers and facilitators in various care settings to optimize and scale the model effectively. Future research should build on the findings of this study by exploring the long-term impacts of virtual care models on health outcomes and quality of life for older persons living with HIV. Longitudinal studies could provide insights into the sustainability and efficacy of these interventions over time. In addition, research should investigate the effectiveness of different technological platforms and features in meeting the needs of culturally diverse older persons living with HIV. Existing comparative studies have not assessed virtual care against traditional models in the context of HIV. In addition, issues of equity and access must be addressed to avoid further disadvantaging vulnerable individuals who may lack access to necessary technology or resources. Furthermore, exploring the barriers to and facilitators of implementation in various settings, such as acute and community care, can offer valuable information for refining and scaling virtual geriatric-HIV care models. Finally, future research should continue to address the needs of non-English speakers and other underserved groups, ensuring that virtual care remains inclusive and equitable.

Limitations

This study, while innovative in its co-design approach to developing a virtual care model for older adults living with HIV, has noted limitations. First, the sample size was relatively small and geographically limited to a dense urban center, which may impact the transferability of the findings. Most participants were already engaged with technology, which may not accurately represent the broader population of older adults living with HIV who have varying levels of technological proficiency. This existing familiarity could have positively biased the feedback on virtual care models as these participants, who were actively seeking health care services, might be more comfortable and adept with digital tools. This may not be representative of others who avoid health care and may struggle more with virtual care platforms. Consequently, the findings may overestimate the acceptability of such interventions for the entire demographic, potentially overlooking significant barriers experienced by those less familiar or comfortable with technology. Future research should aim to include a more diverse sample in terms of technological experience to ensure that the developed care models are inclusive and accessible to all older adults living with HIV. Finally, geriatric-HIV care providers and community agencies were not engaged as formal participants in the co-design phases. Their exclusion may have limited insights into the clinical and service delivery perspectives necessary for developing a more holistic and integrated care model. Engaging

XSL•FO
RenderX

these stakeholders could have addressed gaps related to continuity of care and support services. Future iterations of this model would benefit from the involvement of health care providers and community organizations to ensure that the virtual care system is aligned with existing services and adequately meets the needs of older adults living with HIV.

Conclusions

This study highlights the crucial need for socioculturally appropriate virtual care models for older persons living with HIV. Using an EBCD approach, we engaged older persons living with HIV in co-designing a virtual care model that addresses their unique needs and challenges. The findings reveal several key areas for improvement in virtual geriatric-HIV care, including the necessity for robust support systems, effective communication strategies, and tailored health management resources. Participants identified critical needs such as the integration of multimodal care options, transparency in technology use, and the implementation of reminder systems to enhance patient engagement and continuity of care. They also emphasized the importance of linguistic and cultural equity, recommending that virtual care models incorporate diverse language options and cultural considerations to ensure accessibility and relevance. In addition, affordability was highlighted as a significant barrier, with participants advocating for subsidized technology and training to overcome economic challenges. Policy makers should consider the advantages of virtual care for older persons living with HIV and support the expansion of telehealth services and technological infrastructure. Future research should explore the long-term impacts of virtual care on health outcomes and quality of life on older persons living with HIV as well as evaluating different technological platforms and features. By embracing the diverse voices of older persons living with HIV, this study paves the way for a more inclusive, innovative, and equitable virtual future of care.

Conflicts of Interest

KMK is an Associate Editor for JMIR Rehabilitation and Assistive Technologies.

Multimedia Appendix 1
Demographic information for participants in phases 1, 2, and 3 (N=19).
[DOCX File , 41 KB - aging_v8i1e67122_app1.docx]

References

1. He W, Kinsella K. Global aging in the new millennium. In: Sokolovsky J, editor. The Cultural Context of Aging: Worldwide Perspectives. Westport, CT: Praeger/ABC-CLIO; 2020:27-46.
2. Autenrieth CS, Beck EJ, Stelzle D, Mallouris C, Mahy M, Ghys P. Global and regional trends of people living with HIV aged 50 and over: estimates and projections for 2000-2020. PLoS One 2018 Nov 29;13(11):e0207005 [FREE Full text] [doi: 10.1371/journal.pone.0207005] [Medline: 30496302]
3. Sapuła M, Suchacz M, Załęski A, Wiercińska-Drapało A. Impact of combined antiretroviral therapy on metabolic syndrome components in adult people living with HIV: a literature review. Viruses 2022 Jan 11;14(1):122 [FREE Full text] [doi: 10.3390/v14010122] [Medline: 35062326]
4. Grosso TM, Hernández-Sánchez D, Dragovic G, Vasylyev M, Saumoy M, Blanco JR, et al. Identifying the needs of older people living with HIV (≥ 50 years old) from multiple centres over the world: a descriptive analysis. AIDS Res Ther 2023 Feb 12;20(1):10 [FREE Full text] [doi: 10.1186/s12981-022-00488-7] [Medline: 36782210]
5. Guaraldi G, Milic J, Mussini C. Aging with HIV. Curr HIV/AIDS Rep 2019 Dec 10;16(6):475-481. [doi: 10.1007/s11904-019-00464-3] [Medline: 31823125]
6. Wing EJ. HIV and aging. Int J Infect Dis 2016 Dec;53:61-68 [FREE Full text] [doi: 10.1016/j.ijid.2016.10.004] [Medline: 27756678]
7. Onen NF, Overton ET, Seyfried W, Stumm ER, Snell M, Mondy K, et al. Aging and HIV infection: a comparison between older HIV-infected persons and the general population. HIV Clin Trials 2010 Jan 06;11(2):100-109. [doi: 10.1310/hct1102-100] [Medline: 20542846]
8. Aberg JA. Aging and HIV infection: focus on cardiovascular disease risk. Top Antivir Med 2020 Jan;27(4):102-105 [FREE Full text] [Medline: 32224501]
9. Martin-Iguacel R, Llibre JM, Friis-Moller N. Risk of cardiovascular disease in an aging HIV population: where are we now? Curr HIV/AIDS Rep 2015 Dec 30;12(4):375-387. [doi: 10.1007/s11904-015-0284-6] [Medline: 26423407]
10. Medapalli RK, Parikh CR, Gordon K, Brown ST, Butt AA, Gibert CL, et al. Comorbid diabetes and the risk of progressive chronic kidney disease in HIV-infected adults: data from the Veterans Aging Cohort Study. J Acquir Immune Defic Syndr 2012 Aug 01;60(4):393-399 [FREE Full text] [doi: 10.1097/QAI.0b013e31825b70d9] [Medline: 22592587]
11. Biver E. Osteoporosis and HIV Infection. Calcif Tissue Int 2022 May 30;110(5):624-640 [FREE Full text] [doi: 10.1007/s00223-022-00946-4] [Medline: 35098324]
12. Brown TT. HIV: an underrecognized secondary cause of osteoporosis? J Bone Miner Res 2013 Jun;28(6):1256-1258 [FREE Full text] [doi: 10.1002/jbmr.1967] [Medline: 23630095]

XSL•FO
RenderX

13. Cotter AG, Mallon PW. HIV infection and bone disease: implications for an aging population. Sex Health 2011;8(4):493. [doi: 10.1071/sh11014]
14. Frain J. Considering the impact of aging when caring for and treating adults with HIV. J Gerontol Nurs 2020 Apr 01;46(4):31-40. [doi: 10.3928/00989134-20200304-02] [Medline: 32219455]
15. Cherry ST, Tyagi K, Bailey J, Reagan B, Jacuinde V, Soto J, et al. More to come: perspectives and lived experiences of adults aging with HIV. J Gay Lesbian Soc Serv 2021 Jul 01;34(2):177-206. [doi: 10.1080/10538720.2021.1946669]
16. Cox LE, Brennan-Ing M. Medical, social and supportive services for older adults with HIV. Interdiscip Top Gerontol Geriatr 2017;42:204-221. [doi: 10.1159/000448563] [Medline: 27875834]
17. Kokorelias KM, Grosse A, Zhabokritsky A, Sirisegaram L. Understanding geriatric models of care for older adults living with HIV: a scoping review and qualitative analysis. BMC Geriatr 2023 Jul 08;23(1):417 [FREE Full text] [doi: 10.1186/s12877-023-04114-7] [Medline: 37422631]
18. Kokorelias KM, Grosse A, Zhabokritsky A, Walmsley SL, Sirisegaram L. Exploring implementation considerations for geriatric-HIV clinics: a secondary analysis from a scoping review on HIV models of geriatric care. HIV Med 2024 Jan 04;25(1):38-59. [doi: 10.1111/hiv.13549] [Medline: 37794607]
19. Brothers TD, Rockwood K. Frailty: a new vulnerability indicator in people aging with HIV. Eur Geriatr Med 2019 Apr 10;10(2):219-226. [doi: 10.1007/s41999-018-0143-2] [Medline: 34652747]
20. Kalichman SC, Heckman T, Kochman A, Sikkema K, Bergholte J. Depression and thoughts of suicide among middle-aged and older persons living with HIV-AIDS. Psychiatr Serv 2000 Jul;51(7):903-907. [doi: 10.1176/appi.ps.51.7.903] [Medline: 10875956]
21. Halkitis PN, Krause KD, Vieira DL. Mental health, psychosocial challenges and resilience in older adults living with HIV. Interdiscip Top Gerontol Geriatr 2017;42:187-203. [doi: 10.1159/000448564] [Medline: 27875833]
22. Mavandadi S, Zanjani F, Ten Have TR, Oslin DW. Psychological well-being among individuals aging with HIV: the value of social relationships. J Acquir Immune Defic Syndr 2009 May 01;51(1):91-98 [FREE Full text] [doi: 10.1097/QAI.0b013e318199069b] [Medline: 19282781]
23. Brandt CP, Sheppard DP, Zvolensky MJ, Morgan EE, Atkinson JH, Woods SP. Does age influence the frequency of anxiety symptoms and disorders in HIV disease? J HIV/AIDS Soc Serv 2016 Jun 28;15(4):380-403. [doi: 10.1080/15381501.2016.1189865]
24. Basu S, Chwastiak LA, Bruce RD. Clinical management of depression and anxiety in HIV-infected adults. AIDS 2005 Dec 02;19(18):2057-2067. [doi: 10.1097/01.aids.0000182518.84407.32] [Medline: 16284454]
25. Deng L, Zhang X, Gao Y, Turner D, Qian F, Lu H, et al. Association of HIV infection and cognitive impairment in older adults: a meta-analysis. Ageing Res Rev 2021 Jul;68:101310 [FREE Full text] [doi: 10.1016/j.arr.2021.101310] [Medline: 33640473]
26. Alford K, Vera JH. Cognitive impairment in people living with HIV in the ART era: a review. Br Med Bull 2018 Sep 01;127(1):55-68. [doi: 10.1093/bmb/ldy019] [Medline: 29868901]
27. Hosaka KR, Greene M, Premeaux TA, Javandel S, Allen IE, Ndhlovu LC, et al. Geriatric syndromes in older adults living with HIV and cognitive impairment. J Am Geriatr Soc 2019 Sep 26;67(9):1913-1916 [FREE Full text] [doi: 10.1111/jgs.16034] [Medline: 31241764]
28. Grov C, Golub SA, Parsons JT, Brennan M, Karpiak SE. Loneliness and HIV-related stigma explain depression among older HIV-positive adults. AIDS Care 2010 May 16;22(5):630-639. [doi: 10.1080/09540120903280901] [Medline: 20401765]
29. Brown MJ, Adeagbo O. HIV and aging: double stigma. Curr Epidemiol Rep 2021 Mar 12;8(2):72-78 [FREE Full text] [doi: 10.1007/s40471-021-00265-6] [Medline: 33728256]
30. Brennan-Ing M. Diversity, stigma, and social integration among older adults with HIV. Eur Geriatr Med 2019 Apr 07;10(2):239-246. [doi: 10.1007/s41999-018-0142-3] [Medline: 34652745]
31. DeMarco RF, Brennan-Ing M, Sprague C, Brown SM. Ageism, aging and HIV: community responses to prevention, treatment, care and support. Interdiscip Top Gerontol Geriatr 2017;42:234-239. [doi: 10.1159/000448567] [Medline: 27875836]
32. Emlet CA. "You're awfully old to have this disease": experiences of stigma and ageism in adults 50 years and older living with HIV/AIDS. Gerontologist 2006 Dec;46(6):781-790. [doi: 10.1093/geront/46.6.781] [Medline: 17169933]
33. Wallach I, Brotman S. Gaps in health and social care services to older adults living with HIV: a qualitative study on the perspectives of older adults and service providers. Nurs Care Open Access J 2019 Jan 21;6(1). [doi: 10.15406/ncoaj.2019.06.00179]
34. O'Neill G, Barry PP. Training physicians in geriatric care: responding to critical care. Public Policy Aging Rep 2003;13(2):17-21 [FREE Full text] [doi: 10.1093/ppar/13.2.17]
35. Harris J. Geriatric trends facing nursing with the growing aging. Crit Care Nurs Clin North Am 2019 Jun;31(2):211-224. [doi: 10.1016/j.cnc.2019.02.007] [Medline: 31047094]
36. Airhihenbuwa CO, Ford CL, Iwelunmor JI. Why culture matters in health interventions: lessons from HIV/AIDS stigma and NCDs. Health Educ Behav 2014 Feb 17;41(1):78-84. [doi: 10.1177/1090198113487199] [Medline: 23685666]

XSL•FO
RenderX

37. Arnold EA, Rebchook GM, Kegeles SM. 'Triply cursed': racism, homophobia and HIV-related stigma are barriers to regular HIV testing, treatment adherence and disclosure among young Black gay men. Cult Health Sex 2014 Jun 02;16(6):710-722 [FREE Full text] [doi: 10.1080/13691058.2014.905706] [Medline: 24784224]
38. Faller J, Marcon SS. Praticas socioculturais e de cuidado a saude de idosos em diferentes etnias. Esc Anna Nery 2013 Sep;17(3):512-519. [doi: 10.1590/S1414-81452013000300015]
39. Weßel M. Feminist approach to geriatric care: comprehensive geriatric assessment, diversity and intersectionality. Med Health Care Philos 2022 Mar 16;25(1):87-97 [FREE Full text] [doi: 10.1007/s11019-021-10052-1] [Medline: 34529218]
40. Holden K, McGregor B, Thandi P, Fresh E, Sheats K, Belton A, et al. Toward culturally centered integrative care for addressing mental health disparities among ethnic minorities. Psychol Serv 2014 Nov;11(4):357-368 [FREE Full text] [doi: 10.1037/a0038122] [Medline: 25383991]
41. Vu GT, Tran BX, Hoang CL, Hall BJ, Phan HT, Ha GH, et al. Global research on quality of life of patients with HIV/AIDS: is it socio-culturally addressed? (GAPRESEARCH). Int J Environ Res Public Health 2020 Mar 23;17(6):2127 [FREE Full text] [doi: 10.3390/ijerph17062127] [Medline: 32210042]
42. Ostiguy P. Populism: a socio-cultural approach. In: Kaltwasser CR, Taggart P, Espejo PO, Ostiguy P, editors. The Oxford Handbook of Populism. Oxford, UK: Oxford University Press; 2017.
43. Kokorelias KM, Wasilewski MB, Flanagan A, Zhabokritsky A, Singh H, Dove E, et al. Co-creating socio-culturally-appropriate virtual geriatric care for older adults living with HIV: a community-based participatory, intersectional protocol. Int J Qual Methods 2023 Oct 06;22 [FREE Full text] [doi: 10.1177/16094069231205189]
44. Alegria M, Pescosolido BA, Williams S, A G. Culture, race/ethnicity and disparities: fleshing out the socio-cultural framework for health services disparities. In: Pescosolido B, Martin J, McLeod J, Rogers A, editors. Handbook of the Sociology of Health, Illness, and Healing. New York, NY: Springer; 2011.
45. Goldstein MZ, Griswold K. Cultural sensitivity and aging. Psychiatr Serv 1998 Jun;49(6):769-771. [doi: 10.1176/ps.49.6.769] [Medline: 9634155]
46. Xakellis G, Brangman SA, Hinton WL, Jones VY, Masterman D, Pan CX, et al. Curricular framework: core competencies in multicultural geriatric care. J Am Geriatr Soc 2004 Jan 24;52(1):137-142. [doi: 10.1111/j.1532-5415.2004.52024.x] [Medline: 14687329]
47. Periyakoil VS. Building a culturally competent workforce to care for diverse older adults: scope of the problem and potential solutions. J Am Geriatr Soc 2019 May 10;67(S2):S423-S432. [doi: 10.1111/jgs.15939] [Medline: 31074850]
48. Iheneche E. Cultural diversity and cultural competency: new issues for elderly care and services. Arcada University of Applied Science. 2010. URL: https://www.theseus.fi/bitstream/handle/10024/7063/Iheneche_Ernest.pdf?sequence=1 [accessed 2025-05-01]
49. Freshman B, O'Lawrence H, Montion E. The workforce caring for an aging population: addressing the crisis of inadequate geriatric training among healthcare professionals. Workforce Educ Forum 2014;37(1):61-86 [FREE Full text]
50. Sangarlangkarn A, Avihingsanon A, Appelbaum JS. Application of geriatric principles and care models in HIV and aging. Interdiscip Top Gerontol Geriatr 2017;42:119-133. [doi: 10.1159/000448549] [Medline: 27875828]
51. Singh HK, Del Carmen T, Freeman R, Glesby MJ, Siegler EL. From one syndrome to many: incorporating geriatric consultation into HIV care. Clin Infect Dis 2017 Aug 01;65(3):501-506. [doi: 10.1093/cid/cix311] [Medline: 28387803]
52. Montoya ID, Whitsett DD. New frontiers and challenges in HIV research among older minority populations. J Acquir Immune Defic Syndr 2003 Jun 01;33 Suppl 2(Sup 2):S218-S221. [doi: 10.1097/00126334-200306012-00019] [Medline: 12853871]
53. Liboro RM, Bell S, Ranuschio B, Barnes L, Despres J, Sedere A, et al. Barriers and facilitators to promoting resilience to HIV/AIDS: a qualitative study on the lived experiences of HIV-positive, racial and ethnic minority, middle-aged and older men who have sex with men from Ontario, Canada. Int J Environ Res Public Health 2021 Jul 30;18(15):8084 [FREE Full text] [doi: 10.3390/ijerph18158084] [Medline: 34360376]
54. Dubov A, Krakower DS, Rockwood N, Montgomery S, Shoptaw S. Provider implicit bias in prescribing HIV pre-exposure prophylaxis (PrEP) to people who inject drugs. J Gen Intern Med 2023 Oct 24;38(13):2928-2935 [FREE Full text] [doi: 10.1007/s11606-023-08040-7] [Medline: 36964426]
55. Geter A, Herron AR, Sutton MY. HIV-related stigma by healthcare providers in the United States: a systematic review. AIDS Patient Care STDS 2018 Oct;32(10):418-424 [FREE Full text] [doi: 10.1089/apc.2018.0114] [Medline: 30277814]
56. Singleton MC, Green DC, Enguidanos SM. Identifying healthcare stereotype threat in older gay men living with HIV. J Appl Gerontol 2023 Sep 11;42(9):1965-1973. [doi: 10.1177/07334648231167944] [Medline: 37040260]
57. FitzGerald C, Hurst S. Implicit bias in healthcare professionals: a systematic review. BMC Med Ethics 2017 Mar 01;18(1):19 [FREE Full text] [doi: 10.1186/s12910-017-0179-8] [Medline: 28249596]
58. Montoya A. From “the people” to “the human”: HIV/AIDS, neoliberalism, and the economy of virtue in contemporary Vietnam. positions 2012;20(2):561-591. [doi: 10.1215/10679847-1538515]
59. Sankar A, Nevedal A, Neufeld S, Berry R, Luborsky M. What do we know about older adults and HIV? A review of social and behavioral literature. AIDS Care 2011 Oct 14;23(10):1187-1207 [FREE Full text] [doi: 10.1080/09540121.2011.564115] [Medline: 21939401]

XSL•FO
RenderX

60. Baim-Lance A, Angulo M, Chiasson MA, Lekas HM, Schenkel R, Villarreal J, et al. Challenges and opportunities of telehealth digital equity to manage HIV and comorbidities for older persons living with HIV in New York State. BMC Health Serv Res 2022 May 06;22(1):609 [FREE Full text] [doi: 10.1186/s12913-022-08010-5] [Medline: 35524251]
61. Lesko CR, Bengtson AM. HIV and COVID-19: intersecting epidemics with many unknowns. Am J Epidemiol 2021 Jan 04;190(1):10-16 [FREE Full text] [doi: 10.1093/aje/kwaa158] [Medline: 32696057]
62. Flickinger TE, DeBolt C, Xie A, Kosmacki A, Grabowski M, Waldman AL, et al. Addressing stigma through a virtual community for people living with HIV: a mixed methods study of the PositiveLinks mobile health intervention. AIDS Behav 2018 Oct 7;22(10):3395-3406 [FREE Full text] [doi: 10.1007/s10461-018-2174-6] [Medline: 29882048]
63. Rao D, Frey S, Ramaiya M. eHealth for stigma reduction efforts designed to improve engagement in care for people living with HIV. Curr HIV/AIDS Rep 2018 Dec 19;15(6):397-402 [FREE Full text] [doi: 10.1007/s11904-018-0414-z] [Medline: 30232579]
64. Lewis MA, Harshbarger C, Bann C, Burrus O, Peinado S, Garner BR, et al. Positive Health Check evaluation: a type 1 hybrid design randomized trial to decrease HIV viral loads in patients seen in HIV primary care. Contemp Clin Trials 2020 Sep;96:106097. [doi: 10.1016/j.cct.2020.106097] [Medline: 32738408]
65. Armstrong WS, Agwu AL, Barrette EP, Ignacio RB, Chang JJ, Colasanti JA, et al. Innovations in human immunodeficiency virus (HIV) care delivery during the coronavirus disease 2019 (COVID-19) pandemic: policies to strengthen the ending the epidemic initiative-a policy paper of the Infectious Diseases Society of America and the HIV Medicine Association. Clin Infect Dis 2021 Jan 23;72(1):9-14 [FREE Full text] [doi: 10.1093/cid/ciaa1532] [Medline: 33035296]
66. Cohn WF, Canan CE, Knight S, Waldman AL, Dillingham R, Ingersoll K, et al. An implementation strategy to expand mobile health use in HIV care settings: rapid evaluation study using the consolidated framework for implementation research. JMIR Mhealth Uhealth 2021 Apr 28;9(4):e19163 [FREE Full text] [doi: 10.2196/19163] [Medline: 33908893]
67. Walsh JL, Quinn KG, Hirshfield S, John SA, Algiers O, Al-Shalby K, et al. Acceptability, feasibility, and preliminary impact of 4 remotely-delivered interventions for rural older adults living with HIV. AIDS Behav 2024 Apr 03;28(4):1401-1414. [doi: 10.1007/s10461-023-04227-5] [Medline: 38170275]
68. Marent B, Henwood F, Darking M, EmERGE Consortium. Ambivalence in digital health: co-designing an mHealth platform for HIV care. Soc Sci Med 2018 Oct;215:133-141. [doi: 10.1016/j.socscimed.2018.09.003] [Medline: 30232053]
69. Liang C, Hutson JA, Keyes O. Surveillance, stigma and sociotechnical design for HIV. First Monday 2020 Oct 5;25(10). [doi: 10.5210/fm.v25i10.10274]
70. Massaroni V, Delle Donne V, Ciccarelli N, Ciccullo A, Borghetti A, Faliero D, et al. Use of telehealth for HIV care in Italy: are doctors and patients on the same page? A cross-sectional study. Int J Med Inform 2021 Dec;156:104616. [doi: 10.1016/j.ijmedinf.2021.104616] [Medline: 34695728]
71. Kim JY, Barth SK, Monroe AK, Ahsan S, Kovacic J, Senn S, et al. The impact of COVID-19 on the HIV continuum of care: challenges, innovations, and opportunities. Expert Rev Anti Infect Ther 2023 Jul 27;21(8):831-846. [doi: 10.1080/14787210.2023.2239503] [Medline: 37470436]
72. Phan JM, Kim S, Linh Đ, Cosimi LA, Pollack TM. Telehealth interventions for HIV in low- and middle-income countries. Curr HIV/AIDS Rep 2022 Dec 26;19(6):600-609 [FREE Full text] [doi: 10.1007/s11904-022-00630-0] [Medline: 36156183]
73. Shaw J, Brewer LC, Veinot T. Recommendations for health equity and virtual care arising from the COVID-19 pandemic: narrative review. JMIR Form Res 2021 Apr 05;5(4):e23233 [FREE Full text] [doi: 10.2196/23233] [Medline: 33739931]
74. Eyrich NW, Andino JJ, Fessell DP. Bridging the digital divide to avoid leaving the most vulnerable behind. JAMA Surg 2021 Aug 01;156(8):703-704. [doi: 10.1001/jamasurg.2021.1143] [Medline: 33769441]
75. Luque AE, Corales R, Fowler RJ, DiMarco J, van Keken A, Winters P, et al. Bridging the digital divide in HIV care: a pilot study of an iPod personal health record. J Int Assoc Provid AIDS Care 2013 Sep 10;12(2):117-121 [FREE Full text] [doi: 10.1177/1545109712457712] [Medline: 22965693]
76. Choukou MA, Sanchez-Ramirez DC, Pol M, Uddin M, Monnin C, Syed-Abdul S. COVID-19 infodemic and digital health literacy in vulnerable populations: a scoping review. Digit Health 2022 Feb 10;8:20552076221076927 [FREE Full text] [doi: 10.1177/20552076221076927] [Medline: 35223076]
77. Nixon E, Cooper V, Donetto S, Youssef E, Robert G. Co-designing health services for people living with HIV who have multimorbidity: a feasibility study. Br J Nurs 2023 Jun 08;32(11):S15-S21. [doi: 10.12968/bjon.2023.32.11.s15]
78. Yang C, Teh YE, Chua NG, Lee KL, Ng RQ. An overview of multimorbidity and polypharmacy in older people living with HIV. Geriatr Gerontol Int 2024 Mar 08;24 Suppl 1(S1):49-59. [doi: 10.1111/ggi.14717] [Medline: 37940135]
79. Kokorelias KM, Valentine D, Dove EM, Brown P, McKinlay S, Sheppard CL, et al. Exploring the perspectives of older adults living with HIV on virtual care: qualitative study. JMIR Aging 2024 Dec 04;7:e65730 [FREE Full text] [doi: 10.2196/65730] [Medline: 39630418]
80. Zamenopoulos T, Alexiou K. Co-Design as Collaborative Research. Bristol, England: University of Bristol and the AHRC Connected Communities Programme; Jan 09, 2019:1-53.
81. Castro EM, Malfait S, van Regenmortel T, van Hecke A, Sermeus W, Vanhaecht K. Co-design for implementing patient participation in hospital services: a discussion paper. Patient Educ Couns 2018 Jul;101(7):1302-1305. [doi: 10.1016/j.pec.2018.03.019] [Medline: 29602511]

XSL•FO
RenderX

82. Maye L, Claisse C. Co-design within and between communities in cultural heritage: current and open questions. Multimodal Technol Interact 2022 Dec 20;7(1):1. [doi: 10.3390/mti7010001]
83. Armoon B, Higgs P, Fleury MJ, Bayat AH, Moghaddam LF, Bayani A, et al. Socio-demographic, clinical and service use determinants associated with HIV related stigma among people living with HIV/AIDS: a systematic review and meta-analysis. BMC Health Serv Res 2021 Sep 22;21(1):1004 [FREE Full text] [doi: 10.1186/s12913-021-06980-6] [Medline: 34551772]
84. Donetto S, Tsianakas V, Robert G. Using experience-based co-design (EBCD) to improve the quality of healthcare: mapping where we are now and establishing future directions. King's College London. 2014. URL: https://www.scirp.org/reference/referencespapers?referenceid=3581170 [accessed 2025-05-01]
85. Raynor DK, Ismail H, Blenkinsopp A, Fylan B, Armitage G, Silcock J. Experience-based co-design-Adapting the method for a researcher-initiated study in a multi-site setting. Health Expect 2020 Jun 11;23(3):562-570 [FREE Full text] [doi: 10.1111/hex.13028] [Medline: 32045087]
86. Howrey BT, Thompson BL, Borkan J, Kennedy LB, Hughes LS, Johnson BH, et al. Partnering with patients, families, and communities. Fam Med 2015 Sep;47(8):604-611 [FREE Full text] [Medline: 26382118]
87. Lindamer LA, Lebowitz B, Hough RL, Garcia P, Aguirre A, Halpain MC, et al. Establishing an implementation network: lessons learned from community-based participatory research. Implement Sci 2009 Mar 31;4(1):17 [FREE Full text] [doi: 10.1186/1748-5908-4-17] [Medline: 19335915]
88. Eaton AD. Filmed simulation to train peer researchers in community-based participatory research. Soc Work Res 2019 Sep;43(3):195-199. [doi: 10.1093/swr/svz011]
89. Slattery P, Saeri AK, Bragge P. Research co-design in health: a rapid overview of reviews. Health Res Policy Syst 2020 Feb 11;18(1):17 [FREE Full text] [doi: 10.1186/s12961-020-0528-9] [Medline: 32046728]
90. Brown P, Singh H, Su E, Sirisegaram L, Munce SE, Eaton AD, et al. Understanding the use of co-design methods for research involving older adults living with HIV: a scoping review protocol. PLoS One 2024 May 30;19(5):e0303580 [FREE Full text] [doi: 10.1371/journal.pone.0303580] [Medline: 38814951]
91. Robert G. Participatory action research: using experience-based co-design to improve the quality of healthcare services. In: Ziebland S, Calabrase J, Coulter A, Locock L, editors. Understanding and Using Experiences of Health and Illness. Oxford, UK: Oxford University Press; 2013:138-149.
92. Pearce V, Baraitser P, Smith G, Greenhalgh T. Experience-based co-design. In: Greenhalgh T, Humphrey C, Woodard F, editors. User Involvement in Health Care. Hoboken, NJ: Blackwell Publishing Ltd; 2010.
93. Marshall PA, Rotimi C. Ethical challenges in community-based research. Am J Med Sci 2001 Nov;322(5):241-245. [doi: 10.1097/00000441-200111000-00001]
94. Moser CA. Quota sampling. J R Stat Soc Ser A 1952;115(3):411-423. [doi: 10.2307/2980740]
95. Suri H. Purposeful sampling in qualitative research synthesis. Qual Res J 2011;11(2):63-75. [doi: 10.3316/qrj1102063]
96. Calabrese R, Cohen E, Miller D. Creating a healthy workplace culture using an appreciative inquiry 4-D cycle. Organ Manag J 2013 Sep;10(3):196-207. [doi: 10.1080/15416518.2013.831706]
97. Faure M, Rosenzweig J, van Tiem D. Appreciative inquiry. In: Watkins R, Leigh D, editors. Handbook of Improving Performance in the Workplace: Selecting and Implementing Performance Interventions. Hoboken, NJ: John Wiley & Sons; 2009.
98. Kane M, Trochim WM. Concept Mapping for Planning and Evaluation. Thousand Oaks, CA: SAGE Publications; 2007.
99. Löhr K, Weinhardt M, Sieber S. The "world café" as a participatory method for collecting qualitative data. Int J Qual Methods 2020 Apr 24;19. [doi: 10.1177/1609406920916976]
100. Anderson R. Thematic content analysis (TCA): descriptive presentation of qualitative data. rosemarieanderson.com. 2007. URL: https://rosemarieanderson.com/wp-content/uploads/2014/08/ThematicContentAnalysis.pdf [accessed 2025-05-01]
101. Nicmanis M. Reflexive content analysis: an approach to qualitative data analysis, reduction, and description. Int J Qual Methods 2024;23. [doi: 10.1177/16094069241236603]
102. Francis JJ, Johnston M, Robertson C, Glidewell L, Entwistle V, Eccles MP, et al. What is an adequate sample size? Operationalising data saturation for theory-based interview studies. Psychol Health 2010 Dec;25(10):1229-1245. [doi: 10.1080/08870440903194015] [Medline: 20204937]
103. DeCuir-Gunby JT, Marshall PL, McCulloch AW. Developing and using a codebook for the analysis of interview data: an example from a professional development research project. Field Methods 2010 Dec 27;23(2):136-155. [doi: 10.1177/1525822x10388468]
104. Ando H, Cousins R, Young C. Achieving saturation in thematic analysis: development and refinement of a codebook. Comp Psychol 2014 Feb 20;3. [doi: 10.2466/03.cp.3.4]
105. Viswanathan M, Ammerman A, Eng E, Garlehner G, Lohr KN, Griffith D, et al. Community-based participatory research: assessing the evidence. Evid Rep Technol Assess (Summ) 2004 Aug(99):1-8. [Medline: 15460504]
106. Natow RS. The use of triangulation in qualitative studies employing elite interviews. Qual Res 2019 Feb 19;20(2):160-173. [doi: 10.1177/1468794119830077]
107. Carter N, Bryant-Lukosius D, DiCenso A, Blythe J, Neville AJ. The use of triangulation in qualitative research. Oncol Nurs Forum 2014 Aug 26;41(5):545-547. [doi: 10.1188/14.onf.545-547]

XSL•FO
RenderX

108. Cutcliffe JR, McKenna HP. Establishing the credibility of qualitative research findings: the plot thickens. J Adv Nurs 1999 Aug 25;30(2):374-380. [doi: 10.1046/j.1365-2648.1999.01090.x] [Medline: 10457239]
109. Passmore SR, Kisicki A, Gilmore-Bykovskyi A, Green-Harris G, Edwards DF. "There's not much we can do…" researcher-level barriers to the inclusion of underrepresented participants in translational research. J Clin Transl Sci 2022 Nov 11;6(1):e4 [FREE Full text] [doi: 10.1017/cts.2021.876] [Medline: 35154814]
110. Rush KL, Seaton CL, Corman K, Hawe N, Li EP, Dow-Fleisner SJ, et al. Virtual care prior to and during COVID-19: cross-sectional survey of rural and urban adults. JMIR Form Res 2022 Aug 22;6(8):e37059 [FREE Full text] [doi: 10.2196/37059] [Medline: 35849794]
111. Perri M, Schmidt RA, Guta A, Kaminski N, Rudzinski K, Strike C. COVID-19 and the opportunity for gender-responsive virtual and remote substance use treatment and harm reduction services. Int J Drug Policy 2022 Oct;108:103815 [FREE Full text] [doi: 10.1016/j.drugpo.2022.103815] [Medline: 35969920]
112. Griffin B. Virtual care in Canada: snapshots of innovative virtual care. Digital Health Canada. 2019 Nov. URL: https://divisionsbc.ca/sites/default/files/Divisions/Kootenay%20Boundary/Misc%20-%20events-imags/Digital%20Health%20Canada%20Virtual%20Care%20in%20Canada%20vFinal%20DEC2-2019%5B1%5D.pdf [accessed 2025-05-01]
113. Gari S, Doig-Acuña C, Smail T, Malungo JR, Martin-Hilber A, Merten S. Access to HIV/AIDS care: a systematic review of socio-cultural determinants in low and high income countries. BMC Health Serv Res 2013 May 28;13(1):198 [FREE Full text] [doi: 10.1186/1472-6963-13-198] [Medline: 23714167]
114. Rose MA. Planning HIV education programs for older adults: cultural implications. J Gerontol Nurs 2004 Mar;30(3):34-39. [doi: 10.3928/0098-9134-20040301-11] [Medline: 15061452]
115. Green T, Bonner A, Teleni L, Bradford N, Purtell L, Douglas C, et al. Use and reporting of experience-based codesign studies in the healthcare setting: a systematic review. BMJ Qual Saf 2020 Jan 23;29(1):64-76. [doi: 10.1136/bmjqs-2019-009570] [Medline: 31548278]
116. Webber R, Partridge R, Grindell C. The creative co-design of low back pain education resources. Evid Policy 2022;18(2):436-453 [FREE Full text] [doi: 10.1332/174426421x16437342906266]
117. Donetto S, Pierri P, Tsianakas V, Robert G. Experience-based co-design and healthcare improvement: realizing participatory design in the public sector. Des Journal 2015 May 07;18(2):227-248. [doi: 10.2752/175630615X14212498964312]
118. Mitzner TL, Fausset CB, Boron JB, Adams AE, Dijkstra K, Lee CC, et al. Older adults' training preferences for learning to use technology. Proc Hum Factors Ergon Soc Annu Meet 2008 Sep 01;52(26):2047-2051 [FREE Full text] [doi: 10.1177/154193120805202603] [Medline: 25309139]
119. Hickman J, Rogers WA, Fisk AD. Training older adults to use new technology. J Gerontol B Psychol Sci Soc Sci 2007 Jun;62 Spec No 1(Special_Issue_1):77-84. [doi: 10.1093/geronb/62.special_issue_1.77] [Medline: 17565168]
120. Patterson PB, Roddick J, Pollack CA, Dutton DJ. Virtual care and the influence of a pandemic: necessary policy shifts to drive digital innovation in healthcare. Healthc Manage Forum 2022 Sep 30;35(5):272-278 [FREE Full text] [doi: 10.1177/08404704221110084] [Medline: 35775144]
121. Panaite AC, Desroches OA, Warren É, Rouly G, Castonguay G, Boivin A. Engaging with peers to integrate community care: knowledge synthesis and conceptual map. Health Expect 2024 Apr 03;27(2):e14034 [FREE Full text] [doi: 10.1111/hex.14034] [Medline: 38567865]

Abbreviations

CBPR: community-based participatory research
EBCD: experience-based co-design

Edited by R Bjarnadottir; submitted 02.10.24; peer-reviewed by AL Nguyen, A Peeples; comments to author 10.03.25; revised version received 17.03.25; accepted 21.04.25; published 27.05.25.

Please cite as:
Kokorelias KM, Wasilewski MB, Valentine D, Eaton AD, Dove E, Brown P, McKinlay S, Sheppard CL, Su E, Singh HK, Flanagan A, Zhabokritsky A, Abdelhalim R, Parpia R, Zewude R, Jamieson L, Walmsley S, Sirisegaram L
Socioculturally Appropriate Internet-Based Geriatric Care Model for Older Adults Living With HIV: Experience-Based Co-Design Approach
JMIR Aging 2025;8:e67122
URL: https://aging.jmir.org/2025/1/e67122
doi:10.2196/67122
PMID:

XSL•FO
RenderX

©Kristina M Kokorelias, Marina B Wasilewski, Dean Valentine, Andrew D Eaton, Erica Dove, Paige Brown, Stuart McKinlay, Christine L Sheppard, Esther Su, Hardeep K Singh, Ashley Flanagan, Alice Zhabokritsky, Reham Abdelhalim, Rabea Parpia, Rahel Zewude, Laura Jamieson, Sharon Walmsley, Luxey Sirisegaram. Originally published in JMIR Aging (https://aging.jmir.org), 27.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Longitudinal Remote Sleep and Cognitive Research in Older Adults With Mild Cognitive Impairment and Dementia: Prospective Feasibility Cohort Study

Victoria Grace Gabb[1,2,3], BSc, MSc; Jonathan Blackman[1,2,3], BSc, BMBS, MRCP, MRCPsych, PhD; Hamish Morrison[1,2], BMedSci, MBChB, MRCP; Haoxuan Li[1,2,4,5,6], BSc, MBBS; Adrian Kendrick[7], BA, PgD, PhD; Nicholas Turner[8], BSc, MSc, PhD; Rosemary Greenwood[9,10], BSc, MSc; Bijetri Biswas[11], BSc, MSc; Amanda Heslegrave[12,13], BSc, PhD; Elizabeth Coulthard[1,2,3], BA, MBBS, FRCP, PhD

[1]Bristol Medical School, Translational Health Sciences, University of Bristol, Bristol, United Kingdom
[2]Neurology Department, Southmead Hospital, North Bristol NHS Trust, Bristol, United Kingdom
[3]Mental Health, NIHR Bristol Biomedical Research Centre, Bristol, United Kingdom
[4]Bristol Royal Infirmary, University Hospitals Bristol and Weston NHS Trust, Bristol, United Kingdom
[5]King's College London, London, United Kingdom
[6]King's College Hospital, London, United Kingdom
[7]University of the West of England, Bristol, United Kingdom
[8]Bristol Trials Centre, Bristol Medical School, University of Bristol, Bristol, United Kingdom
[9]Research & Innovation, University Hospitals Bristol and Weston NHS Foundation Trust, Bristol, United Kingdom
[10]York Trials Unit, Department of Health Sciences, University of York, York, United Kingdom
[11]School of Electrical, Electronic and Mechanical Engineering, University of Bristol, Bristol, United Kingdom
[12]Institute of Neurology, University College London, London, United Kingdom
[13]UK Dementia Research Institute, University College London, London, United Kingdom

Corresponding Author:
Victoria Grace Gabb, BSc, MSc
Bristol Medical School
Translational Health Sciences
University of Bristol
Bristol Brain Centre
Elgar House
Bristol, BS10 5NB
United Kingdom
Phone: 44 01174560700
Email: victoria.gabb@bristol.ac.uk

Abstract

Background: Sleep holds promise as a modifiable risk factor for neurodegenerative diseases and dementia. Clinical trials to modify sleep in people at risk of or in the early stages of dementia are needed. Monitoring natural sleep from home could support pragmatic and decentralized large-scale clinical trials. However, whether longitudinal sleep research can be successfully delivered remotely in this population has not been established yet.

Objective: We investigated the feasibility of remote longitudinal research using wearable devices, web-based cognitive tasks, and a smartphone app to record sleep and cognition in older adults with mild cognitive impairment (MCI) or dementia.

Methods: Older adults with MCI or dementia due to Alzheimer disease or Lewy body disease and cognitively healthy participants completed at-home sleep and circadian monitoring (digital sleep diaries, actigraphy, wearable sleep electroencephalography, and saliva samples) and digital cognitive assessments for 8 weeks. Feasibility outcomes included recruitment, retention, and data completeness.

Results: In total, 41 participants consented (n=10, 24% participants with Alzheimer disease; n=11, 27% participants with Lewy body disease; and n=20, 49% controls). There were predominantly male and White British participants, with a mean age of 70.9

XSL•FO
RenderX

(SD 5.9) years. Retention was very high, with 40 (98%) participants completing 8 weeks of remote monitoring. Data completeness for sleep electroencephalography was 91% and ranged from 79% to 97% for all remote tasks and was overall high across all participant subgroups. In total, 30% (12/40) of participants reported receiving external support with completing study tasks.

Conclusions: High rates of retention, data completeness, and data quality suggested that longitudinal multimodal sleep and cognitive profiling using novel and remote monitoring technology is feasible in older adults with MCI and dementia and healthy older adults, even without study partner support. Remote monitoring should be considered for mechanistic and interventional trials. Careful consideration should be given to how to ensure remote monitoring technologies reduce burden and enhance inclusivity, particularly in communities traditionally underserved by research and those with lower digital literacy.

International Registered Report Identifier (IRRID): RR2-10.2196/52652

(JMIR Aging 2025;8:e72824) doi:10.2196/72824

KEYWORDS

feasibility; sleep; remote study design; mild cognitive impairment; dementia; electroencephalography; actigraphy; saliva; digital biomarkers

Introduction

Background

Sleep disturbances such as insomnia, fragmented sleep, daytime sleepiness, and sleep-disordered breathing are common features of Alzheimer disease (AD) and Lewy body disease (LBD) and often appear early in the disease course and before clinical diagnosis [1,2]. Short and fragmented sleep are associated with processes linked to neurodegeneration, including reduced glymphatic clearance of waste [3], increased amyloid beta (Aβ) and tau burden [4], increased neuroinflammation [5], and impaired cardiovascular health [6]. Several small studies of sleep apnea treatment have demonstrated improvement in cognitive outcomes and blood biomarkers of Aβ and tau, suggesting that sleep interventions could improve prognosis in mild cognitive impairment (MCI) and dementia [7-9]. Poor quality or insufficient sleep in midlife increases the risk of all-cause dementia and MCI [10,11], indicating that improving sleep may also protect against dementia. Large-scale clinical studies are needed to enhance our mechanistic understanding of sleep, confirm the most promising therapeutic targets, and monitor the effectiveness and safety profile of sleep interventions [12,13].

However, selecting sleep assessment tools is challenging, particularly in populations with cognitive impairment. Self-report is convenient, inexpensive, and scalable, and has often been used to examine sleep in people with dementia and MCI [14]; however, it is potentially prone to recall bias due to memory deficits or anosognosia. Self-report also correlates poorly with objective sleep measures, particularly in participants with MCI or dementia [15,16] and those with subjectively poor sleep [17], and cannot inform on key components of sleep such as sleep staging or microarchitecture. Polysomnography (PSG) is typically considered the gold standard for sleep measurement, as it provides rich objective sleep data. However, PSG typically requires expert setup, analysis, and a controlled clinic environment, meaning it is expensive, not easily scalable, and therefore typically used for 1 or a few nights [18,19]. PSG setup, especially in an artificial environment, may also not reflect natural sleep [19] as it disrupts usual routines. Longitudinal data collection in sleep research would be beneficial for monitoring clinical trials and disease progression and could account for variation in sleep from external factors such as acute illness or stress, as well as natural intraindividual sleep variation [19].

Wearable devices, smartphone apps, and telemedicine, apart from actigraphy devices, have rarely been used in MCI or dementia research [20], but offer an opportunity to collect objective and subjective data longitudinally in a natural setting [21], often at relatively low cost. Detailed and accurate sleep analysis can now be achieved through wireless technology, including electroencephalography (EEG) headbands and overnight pulse oximetry [21,22]. While actigraphy has been used in dementia research, EEG headbands and pulse oximetry are less well tested, especially in the earlier stages of impairment [23-25]. There is also increasing interest in remote cognitive testing and digital biomarkers for diagnosis and monitoring progression [26]. Despite this, the adoption of digital health technologies into neurology clinical trials, particularly in older people, has been slow [27].

Remote sleep and cognitive data collection has the potential to decentralize clinical trials, making research more convenient for participants and reducing the costs, participant burden, and carbon footprint associated with study visits, while enabling real-time longitudinal data collection of treatment effects. Improved digital access and literacy among older adults [28] and the use of technology among patients living with MCI or dementia to support independent living and for recreation indicate increasing acceptance of technology [29]. However, not all older adults are comfortable using technology, and changes in cognition, sensory processing, and communication might negatively impact the usability and acceptability of novel devices for research purposes in people with cognitive impairment [29,30]. The few studies that have tested wearable devices and digital health technologies for sleep and dementia research in the home have predominantly collected feasibility data for a single device across only a few nights [31] and required support from a study partner or care home staff [32-34]. Before trials invest in digital health technologies and remote study designs, it is important to know whether participants can and are willing to engage and provide data in such studies over a longer period and how much support might be required.

XSL•FO
RenderX

Objectives

This study aimed to establish the feasibility of predominantly technology-based longitudinal sleep and cognitive assessments from home in older adults with and without MCI or dementia.

Methods

Overview

The Remote Evaluation of Sleep to Enhance Understanding of Early Dementia (RESTED) study was a prospective longitudinal observational cohort study involving remote sleep and cognitive monitoring of participants with MCI or dementia and age-matched cognitively healthy controls. The full study protocol has been previously published [35]. The study has been reported in line with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines [36] (Multimedia Appendix 1).

Ethical Considerations

This study was approved by the Health Research Authority (Yorkshire and the Humber-Bradford Leeds Research Ethics Committee, reference 21/YH/0177). This study was conducted in accordance with the principles of the Declaration of Helsinki. All participants provided written informed consent before engaging in any study activities and were reminded of their right to withdrawal without giving a reason and that withdrawal would not affect their health care. Following consent, data were pseudonymized with participants allocated a participant identifier code to maintain privacy. Participants were not compensated for their involvement in the study; however, travel expenses for research visits outside of usual clinic visits were reimbursed. All researchers engaging with participants were trained in good clinical practice.

Study Population

The RESTED study recruited community-dwelling adults aged ≥50 years with internet access at home. Participants were recruited to 1 of 3 participant subgroups according to clinical diagnosis meeting standardized diagnostic criteria [37-40]. Participants with MCI or mild dementia due to probable AD were recruited to the AD group, participants with MCI or mild dementia due to probable LBD or Parkinson disease (PD) were recruited to the LBD group, and sex- and age-matched individuals with no known neurodegenerative conditions or cognitive impairment were recruited as controls. Exclusion criteria included advanced dementia, acute or terminal illness, and significant unrelated comorbidities that might interfere with sleep, except for obstructive sleep apnea (OSA) if participants were already undergoing treatment.

Participants were recruited from the city of Bristol, United Kingdom, and the surrounding areas via cognitive and movement disorders clinics and research volunteer databases, including Join Dementia Research. Our original target sample size was 75 participants [41]. Due to the COVID-19 pandemic, the study was delayed, and the budget was partially reallocated to studies that required no patient contact [12,13,42]. Therefore, we revised the target sample size to 40 participants and recruited from a single site (North Bristol NHS Trust).

We did not recruit or require participants to have a study partner. Participants were welcome to invite someone to support them, and study support was recorded.

Study Procedures

Screening

Participants were prescreened for eligibility over the telephone and invited to complete consent and screening in person. Participants who scored less than 11 out of 30 on the Montreal Cognitive Assessment (MoCA) [43] at screening were considered too clinically impaired to participate and were withdrawn.

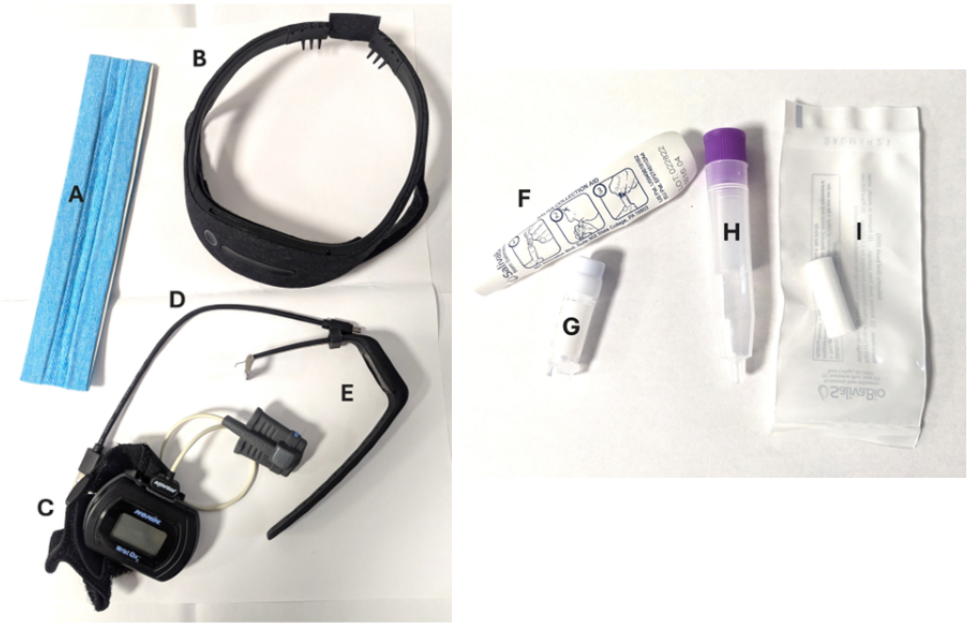
Baseline Assessments

Eligible participants completed baseline questionnaires with a researcher to assess sleep quality (Pittsburgh Sleep Quality Index [PSQI]) [44]; daytime sleepiness (Epworth Sleepiness Scale) [45]; OSA risk (STOP-Bang) [46]; symptoms of depression (Geriatric Depression Scale-15 item) [47]; anxiety (Generalized Anxiety Disorder-7 item) [48]; and apathy (Apathy Evaluation Scale-Self) [49]. Demographic information and medical histories were also recorded.

Participants were provided with a study kit (Figure 1) consisting of an Axivity AX3 actigraphy watch (Axivity Ltd); Dreem 2 wireless sleep EEG headband (Dreem) [23]; a USB charger; a home saliva collection kit for passive drool (for a dim-light melatonin onset assay); oral swabs (for a cortisol awakening response assay); and an overnight pulse oximeter. Participants who requested a study device were provided with a tablet. Details on devices and the study kit are provided in Multimedia Appendix 2. All participants were provided with a printed participant guide with instructions for each remote study task and research team contact details.

XSL•FO
RenderX

Figure 1. The study kit provided to the Remote Evaluation of Sleep to Enhance Understanding of Early Dementia (RESTED) participants. Participants were additionally required to use a smartphone or tablet to complete digital assessments. A: Sweatband to ensure tight fitting of Dreem 2; B: Dreem 2 electroencephalography (EEG) headband; C: Nonin 3150 WristOx 2 oximeter; D: USB-C charging cable (for Dreem 2 and AX3); E: Axivity AX3 actigraph and wrist strap; F: Saliva collection aid for passive drool; G: Collection tube for passive drool; H: Storage tube for oral swab; I: Oral swab.



At baseline, participants were provided with support to download and use the MyDignio (Dignio AS) app, a telemedicine cloud-based software tool that was used to deliver digital sleep diaries, questionnaires, and reminders to complete study tasks. Participants were also familiarized with the study schedule and tasks and offered support and training ad hoc throughout the study.

Main Study Period

Participants completed 56 days of continuous sleep and regular cognitive monitoring in their own home using an Axivity AX3 actigraphy watch and daily sleep diaries [50] delivered via the MyDignio app. Participants were scheduled to complete a set of 3 web-based cognitive tests (choice reaction time, forward digit span, and self-ordered search) twice weekly on a bespoke study version of the web-based assessment platform Cognitron [51].

For 7 days during the 8-week period, participants also completed an "intensive week," consisting of daily cognitive tasks on Cognitron and nighttime sleep recordings using Dreem 2, a wireless sleep EEG headband. The Dreem sleep staging algorithm has comparable accuracy to manual sleep expert scoring of PSG [23], including in cognitively healthy older adults as well as patients with AD [52] and PD [53]. Sleep recordings were initiated by the participant at their natural bedtime each night and terminated after natural awakening each morning. Data were uploaded via Bluetooth and Wi-Fi to a server accessible to the research team. Additional study tasks during the intensive week included 4 verbal memory assessments involving immediate and delayed free recall and a target-distractor recognition task with a researcher via videoconferencing software and serial saliva samples across 1 evening to assess dim-light melatonin onset and 1 morning to assess cortisol awakening response.

Participants were also invited to undergo 2 nights of pulse oximetry for sleep apnea screening and a blood test for plasma biomarker analysis of Aβ42:40, phosphorylated-tau 181 and 217, neurofilament light chain, and glial fibrillary acidic protein. A protocol amendment approved partway through the study introduced bespoke questionnaires probing study expectations, reasons for participation, experience with technology, and how acceptable they found the intensive week study tasks. Participants were also invited to attend a remote end-of-study interview to share their experiences and asked to return for a 6-month follow-up to complete an MoCA and any outstanding study tasks (eg, missed blood test). Participant feedback from qualitative interviews and questionnaires as well as sleep characteristics for the cohort will be analyzed and presented separately.

Feasibility and Acceptability Outcomes

Recruitment and Retention

Consent, recruitment, and retention rates were recorded and are summarized in a flowchart, alongside reasons for nonparticipation at each stage of the recruitment process.

Data Quality and Completeness

For each remote study task, data completeness was assessed by the average number and percentage of completed tasks or nights' use per participant per participant group (data completeness rate). The number and percentage of individuals who completed the maximum number of data points for a given study task (eg, completed all requested 7 nights of EEG) is also provided. For

XSL•FO
RenderX

EEG, signal quality is reported based on Dreem's automated algorithm. Optimal signal quality is considered ≥85% and good quality is considered ≥70%.

Associations Between Participant Characteristics and Adherence

To explore potential associations between key clinical and demographic variables and adherence, we examined the correlation between 2 continuous measures of adherence (number of sleep diaries completed and mean record quality of Dreem EEG data) and 4 continuous variables which might impact digital literacy or engagement (age at consent, apathy, subjective sleep quality as assessed by the PSQI total score, and baseline cognitive impairment as assessed by MoCA total score).

Study Support and Resource Use

The number of participants who attended study visits with someone to support them, reported having a partner or relative who could support study tasks, and reported receiving support on at least 1 study task were recorded and are summarized descriptively.

Data Analysis

Unless otherwise stated, descriptive statistics are provided as mean (SD) for continuous variables, and frequency and percentage for categorical variables, and provided for the full cohort as well as by participant subgroup (AD vs LBD vs controls). Data analysis was performed using R (version 4.3.1; R Foundation for Statistical Computing) and R Studio software (version 2023.6.0.421; R Foundation for Statistical Computing). Actigraphy data were processed using the open-source AX3/AX6 OMGUI application and analyzed in R using the *GGIR* package [54]. Where available, sleep timing was adjusted using participants' sleep diaries; otherwise, *GGIR* uses the heuristic algorithm looking at distribution of change in z-angle [55] to estimate sleep timing. All files were also visually inspected to check for the accuracy of sleep timing. The quality of the Dreem 2 recordings was assessed by inspecting Dreem's automated record quality index in the sleep report for each night, which indicates the percentage of the recording that is of scorable quality for sleep analysis. Associations between participant characteristics and adherence were calculated using Spearman rank correlation with α set to $P<.05$. Reasons for nonparticipation and missing data are provided where known and mapped to the Capability, Opportunity, Motivation model of behavior change [56]. Where additional data were collected on any outcome (eg, participants completed an additional night of EEG than instructed), additional data points were removed before analysis to avoid biasing feasibility metrics.

Patient and Public Involvement

Patient and public involvement (PPI) was sought from individuals with lived experience of MCI and dementia before and throughout the study. PPI contributors reviewed and improved study documents and advised on the acceptability of adding blood biomarker testing. The PPI group strongly endorsed our additional recruitment materials (including an advertisement poster and a plain English participant information sheet) that we introduced partway through the study, after we received feedback from a prospective participant that standard participant information sheets are too long and use inaccessible language for people with cognitive impairment. We also introduced a reduced study protocol, which involved use of a paper (rather than digital) sleep diary, actigraphy, and EEG, to encourage recruitment of participants who may be more comfortable with less frequent use of technology; however, nobody chose the reduced protocol, they either declined altogether or participated in the full study.

Results

Participant Characteristics and Enrollment

Recruitment

Participant flow through the study is shown in Figure 2. Recruitment was open for 17 months from February 2022 to July 2023, with an average recruitment rate of 3 participants per month. Of 129 individuals identified as potentially eligible before screening, 44 individuals consented to take part, giving a consent rate of 34%. Participants described different motivations for taking part, including wanting to support dementia research, perceiving the study as helpful for themselves or others, and because the study sounded interesting or novel. Reasons for declining to take part in the study are described in Figure 3 and mapped to the Capability, Opportunity, Motivation model of behavior change [56].

XSL•FO
RenderX

Figure 2. Participant flow through the Remote Evaluation of Sleep to Enhance Understanding of Early Dementia (RESTED) study. MCI: mild cognitive impairment; MoCA: Montreal Cognitive Assessment.

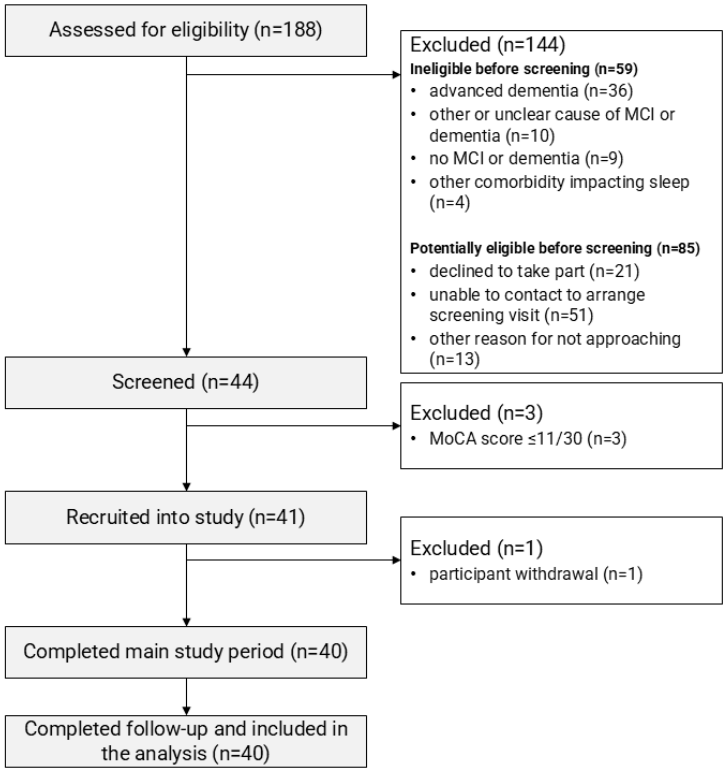
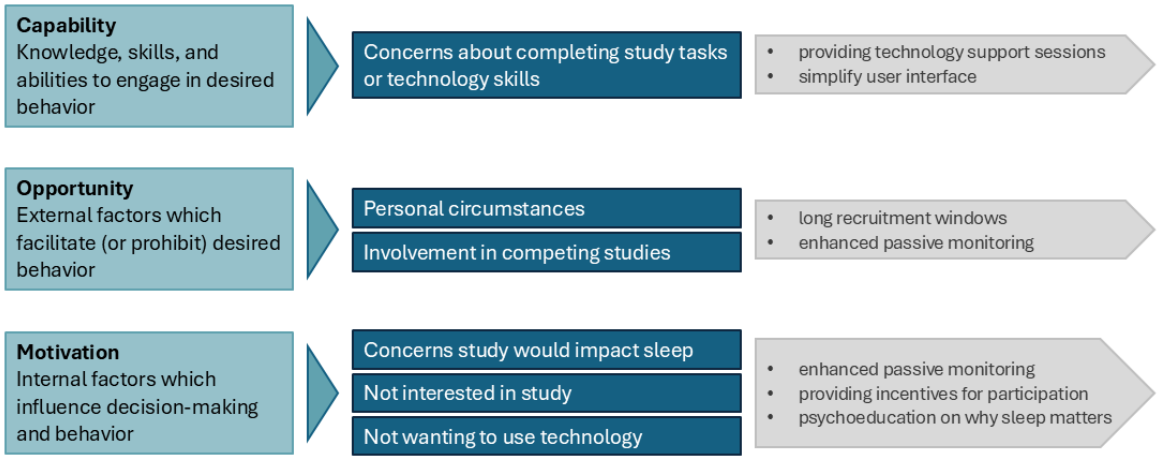


Figure 3. Reasons for declining to participate in the study could be categorized according to the capability, opportunity motivation, behavior (COM-B) model of behavior change in blue. Opportunities to increase capability, opportunity, and motivation to take part in the research are outlined in gray.



Retention

Of 44 individuals who completed screening, 3 scored less than 11 on the MoCA and were withdrawn. 41 participants were eligible and completed all baseline assessments and training in the at-home study tasks. Of the eligible participants, 1 withdrew before starting the at-home study tasks due to personal circumstances and perceived study burden. All remaining 40 participants completed the main 8-week study period and 6-month follow-up and were included in the analysis, giving a retention rate of 98%.

Participant Characteristics

Participant demographics and baseline variables are presented in Table 1. The sample was predominantly made up of male participants with a mean age at consent of 70.9 (SD 5.9, range 57-81) years. Most participants (33/40, 83%) were retired and all identified as White British. In total, 63% (25/40) of participants completed a questionnaire on experiences of digital technology and expectations of the study, which was introduced via an amendment partway through the recruitment period. Of these, most participants reported frequent use of smartphones or other smart technology (18/25, 72%), however only 12% (3/25) of participants had previously used a wearable to monitor their sleep.

XSL•FO
RenderX

Table 1. Participant characteristics for the study cohort.

Characteristics	AD[a] group (n=10)	LBD[b] group (n=10)	Control group (n=20)
Diagnosis at consent, n (%)			
Mild cognitive impairment	4 (40)	6 (60)	0 (0)
Dementia	6 (60)	4 (40)	0 (0)
Blood biomarkers of neurodegeneration or AD pathology (pg/mL), mean (SD)			
Aβ[c]42:40	0.064 (0.006)	0.062 (0.006)	0.072 (0.007)
Phosphorylated-tau 181	26.2 (7.1)	24.9 (12.7)	24.9 (12.7)
Phosphorylated-tau 217	0.7 (0.4)	0.8 (0.4)	0.6 (0.4)
GFAP[d]	102.2 (28.1)	140.9 (56.4)	103.5 (48.1)
NFL[e]	23.9 (18.0)	18.7 (6.5)	13.4 (5.1)
Age at consent (y), mean (SD)	69.2 (7.9)	73.9 (2.8)	70.3 (5.7)
Female participants, n (%)	2 (20)	2 (20)	5 (25)
Employment status, n (%)			
Full-time employment	2 (20)	0 (0)	2 (10)
Part-time employment	0 (0)	0 (0)	3 (15)
Retired	8 (80)	10 (100)	15 (75)
Education, n (%)			
Secondary (≤12 y)	4 (40)	5 (50)	4 (20)
Further or higher (>12 y)	6 (60)	5 (50)	16 (80)
Comorbid diagnoses, n (%)			
Obstructive sleep apnea[f]	0 (0)	1 (10)	2 (10)
Musculoskeletal or pain	3 (30)	3 (30)	8 (40)
Urinary or prostate	3 (30)	2 (20)	3 (15)
Depression or anxiety	1 (10)	0 (0)	1 (5)
Medications, n (%)			
Cognitive enhancer	6 (60)	6 (60)	0 (0)
Dopaminergic	0 (0)	8 (80)	0 (0)
Melatonin	0 (0)	3 (30)	0 (0)
Hypnotic or sedative	0 (0)	2 (20)	1 (5)
Antidepressant	4 (40)	4 (40)	4 (20)
Baseline cognition: MoCA[g]; mean (SD)	23.0 (4.9)	22.0 (4.3)	27.1 (1.4)
Subjective sleep quality: PSQI[h]; mean (SD)	4.3 (2)	7.8 (3)	6.0 (3.9)
Self-reported anxiety: GAD-7[i]; mean (SD)	5.6 (5.2)	3.5 (2.1)	3.6 (4.3)
Self-reported depression: GDS-15[j]; mean (SD)	4.5 (3.3)	4.1 (2.3)	2.2 (2.2)
Self-reported apathy: AES-S[k,l]; mean (SD)	62.2 (6.0)	53.8 (9.9)	62.8 (8.6)

[a]AD: Alzheimer disease.

[b]LBD: Lewy body disease.

[c]Aβ: amyloid beta.

[d]GFAP: glial fibrillary acidic protein.

[e]NFL: neurofilament light chain.

[f]Participants with obstructive sleep apnea were included only if they were being treated with continuous positive airway pressure.

[g]MoCA: Montreal Cognitive Assessment. Higher scores indicate better overall cognition.

XSL•FO
RenderX

[h]PSQI: Pittsburgh Sleep Quality Index. Higher scores indicate worse sleep quality.

[i]GAD-7: Generalized Anxiety Disorder-7. Higher scores indicate more symptoms of anxiety.

[j]GDS-15: Geriatric Depression Scale-15 item scale. Higher scores indicate more symptoms of depression.

[k]AES-S: Apathy Evaluation Scale-Self. Higher scores indicate more apathy.

[l]One participant had a missing value for an item, so this score was imputed using simple imputation.

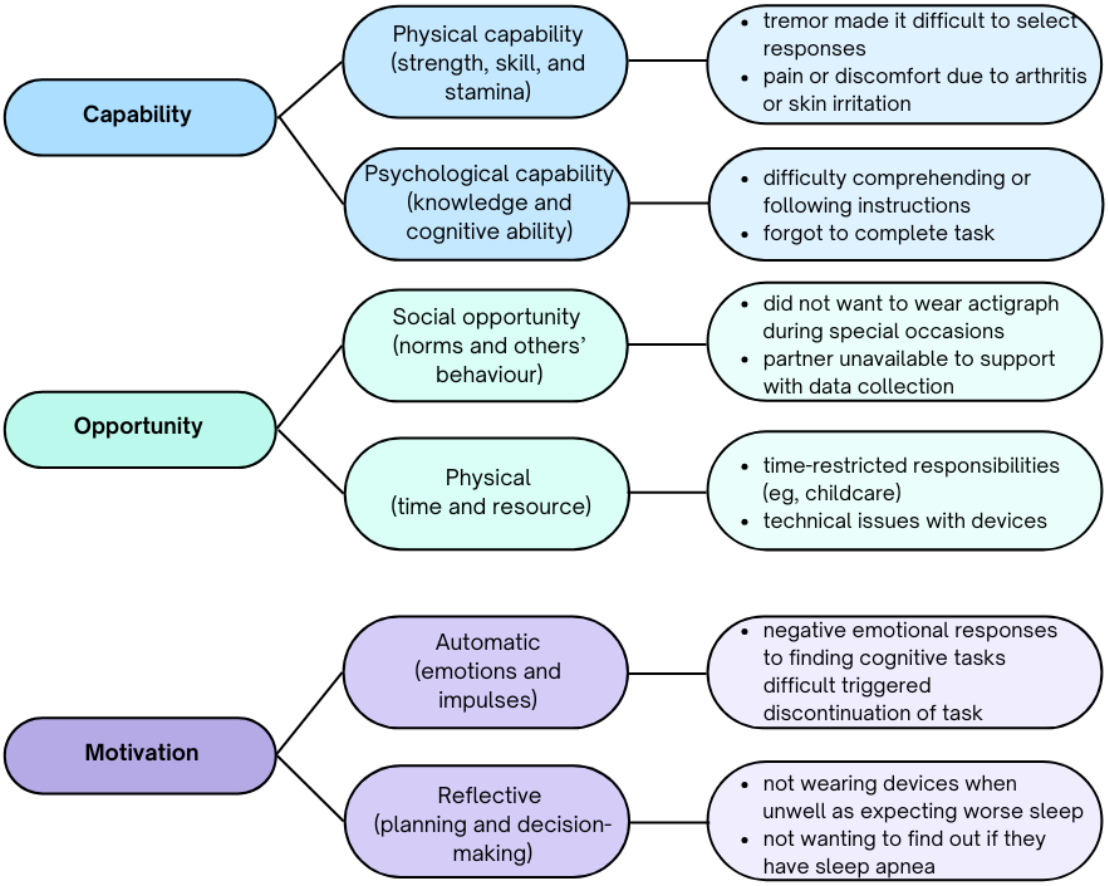
Data Quality and Completeness (Adherence)

Overview

Adherence to the sleep and cognitive tasks is shown in Table 2. Adherence was very high across all study groups and tasks.

Reasons for missing data are summarized in Figure 4 with additional detail provided in Multimedia Appendix 3.

Figure 4. Reasons for missing data are mapped according to the capability, opportunity motivation, behavior (COM-B) model of behavior change.



XSL•FO
RenderX

Table 2. A summary of feasibility outcomes relating to data quality and adherence for wearable devices and remote or web-based cognitive assessments.

Study task (device or software); description; and adherence or data quality metric	AD[a] (n=10)	LBD[b] (n=10)	Controls (n=20)	Total (n=40)
Wearable devices				
Sleep EEG[c] (Dreem 2): participants were asked to complete a week of consecutive overnight EEG recordings to record nighttime sleep (7 nights)				
Nights, mean (SD)	5.9 (2)	6.5 (0.7)	6.6 (0.9)	6.4 (1.2)
Full dataset provided (7 nights), n (%)	7 (70)	6 (60)	16 (80)	29 (73)
Record quality[d], mean (SD)	83.5 (12.1)	84.1 (15.0)	85 (18.4)	84.4 (15.8)
Actigraphy (Axivity AX3): participants were asked to complete 24-h wrist actigraphy continuously for 8 weeks (56 nights)				
Nights, mean (SD)	50.3 (13.1)	52.2 (6.8)	54.4 (3.2)	52.8 (7.6)
Full dataset provided (56 nights), n (%)	5 (50)	5 (50)	14 (70)	24 (60)
Pulse oximetry (Nonin WristOx2 3150): participants were asked to complete 2 consecutive nights of overnight pulse oximetry (2 nights)				
Completed ≥1 recording, n (%)	9 (90)	6 (75)	18 (100)	33 (89)
Full dataset provided (2 nights), n (%)[e]	6 (60)	5 (63)	15 (83)	26 (70)
Remote web-based assessments				
Digital sleep diary (MyDignio): participants were prompted to complete daily sleep diaries for 8 weeks (56 nights)				
Diaries completed, mean (SD)	48.2 (13.4)	51.1 (9.8)	53.8 (3.1)	51.7 (8.6)
Full dataset provided (56 nights), n (%)	1 (10)	4 (40)	8 (40)	12 (30)
Web-based cognitive tests (Cognitron): participants were prompted to complete a digit span, choice reaction time (CRT), and self-ordered search (SOS) tasks on twice weekly for 7 weeks and daily for 1 week (21 tasks)				
Digit spans completed, mean (SD)	17.6 (4.3)	17.6 (4.7)	16.8 (4.7)	17.2 (4.5)
CRT[f] completed, mean (SD)	16.9 (4.2)	13.9 (7.5)	16.7 (4.8)	16.1 (5.4)
SOS[g] completed, mean (SD)	17.5 (4.3)	17.5 (4.6)	16.6 (4.7)	17.0 (4.5)
Completed >1 web-based cognitive task, n (%)	10 (100)	8 (80)	20 (100)	38 (95)
Full dataset provided (all 3 cognitive tasks completed across 21 days), n (%)	2 (20)	2 (20)	2 (10)	6 (15)
Verbal memory tasks (videoconferencing): participants were asked to complete 2 evening "learn" tasks and 2 morning recall or recognition tasks scheduled with researchers (4 tasks)				
Memory tasks completed, mean (SD)	3.2 (1.7)	3.4 (1.3)	3.9 (0.4)	3.6 (1.1)
Full dataset provided (4 tasks completed), n (%)	8 (80)	8 (80)	19 (95)	35 (88)

[a]AD: Alzheimer disease.

[b]LBD: Lewy body dementia.

[c]EEG: electroencephalography.

[d]Mean first calculated for each participant across all available recordings.

[e]Three participants who already had a diagnosis of sleep apnea were not invited to complete the pulse oximetry.

[f]CRT: Choice Reaction Time.

[g]SOS: Self-ordered search.

Sleep EEG

In total, 257 recordings were made using Dreem 2. One participant accidentally completed an eighth recording, which was removed from subsequent analyses to reduce bias.

All participants successfully recorded at least 1 night of sleep, with an average of 6.4 (1.2) nights across the cohort or a 91.4% data completeness rate (Table 2). In total, 73% (29/40) of participants provided data for all 7 nights. Average record quality was also high at 84.4% (15.8), indicating nearly optimal record quality on average across the cohort. One participant wore the Dreem 2 but did not successfully initiate the recording during the intensive week but completed 7 successful nights of recording on a second attempt later in the study. Record quality for individual EEG channels is provided in Multimedia Appendix 4.

Actigraphy

In total, 2332 nights of data across the full cohort were collected. A total of 189 of these nights contained no data, either at the

XSL•FO
RenderX

beginning or the end of the actigraphy file and were removed from sleep analyses. Following visual inspection, a further 9 nights were removed due to sustained nonwear. Therefore, 2134 nights were of sufficient quality for sleep analysis.

Across the cohort, participants provided an average of 52.8 (7.6) out of 56 analyzable nights, giving a data completeness rate of 94.3%. In total, 60% (24/40) of participants provided 56 analyzable nights. All participants recorded at least 14 days of actigraphy. Unexpected battery failure outside of participants' control (due to long-term storage without use during the COVID-19 pandemic) affected 2 participants' recordings and was identified during data check at the midpoint study visit. No participants refused to wear or discontinued use of the actigraphy device, although several reported forgetting or choosing not to wear it on certain days or taking brief breaks due to minor skin irritation.

In total, 37 diary entries were manually entered based on visual analysis of the actigraphy data due to missing sleep diaries and poor heuristic algorithm looking at distribution of change in z-angle algorithm detection of the sleep period. Finally, the sleep analysis software was instructed to solely rely on sleep diary information for 166 nights due to a clear misclassification of sleep period.

Overnight Pulse Oximetry

As 3 (8%) out of 40 participants were already being treated for established OSA, 37 (93%) participants were offered overnight pulse oximetry. In total, 3 (8%) participants declined, and 1 (3%) participant was unable to tolerate wearing the device. The remaining 33 (89%) participants completed at least 1 successful overnight recording, with 26 (70%) recording successful oximetry traces on both nights. The data completeness rate was 79.7% for those eligible and asked to complete overnight pulse oximetry. Referrals to a sleep clinic for either sleep apnea or incidental findings (such as abnormal pulse rise index) were indicated for 57% (21/37) of participants. In total, 4 (11%) participants declined referral due to not wanting a formal diagnosis, not wanting to be put on sleep apnea treatment, and the inconvenience of traveling to the clinic. In total, 17 (46%) participants agreed to a referral.

Digital Sleep Diaries

Although few participants (12/40, 30%) completed all 56 sleep diaries, overall adherence was high. Participants completed an average of 51.7 (SD 8.6) sleep diaries, and only 2 (17%) participants completed fewer than 75% of their sleep diaries. The data completeness rate was 92.3%.

Only 1 (3%) participant completed sleep diaries on paper while her partner, who facilitated entries into MyDignio, was away. In total, 3 (8%) participants had technical issues with data entry into the MyDignio app lasting several days due to a software update and provided some of their sleep diaries via email which were then entered by the study team.

At the time of study setup, Dignio did not offer data format validation, and free-text response boxes were used for several questions. Participants often did not input data in the format requested, and inconsistency in reporting prohibited reliable automated recoding; however, most were manually interpretable by the research team. Where a range of values were provided, a mean was calculated and the value was rounded to the nearest whole minute (eg, 5-10 min was recoded to 8 min). One diary entry was manually corrected by the research team to align with other diary entries and the actigraphy file. It was not possible to recode some responses due to ambiguity (eg, responses of "not long," "several minutes," or "unsure"), leaving 43 completed sleep diaries incomplete on at least one sleep variable. Questions most likely to have a missing value were questions on sleep onset latency and length of nocturnal awakenings.

Video-Based Verbal Memory Tasks

In total, 92% (37/40) of participants completed at least 1 verbal memory task with a researcher. Data were missing for 1 (3%) participant as the task was considered inappropriate due to their specific language difficulties, 1 (3%) participant due to technical issues with video calling, and 1 (3%) participant due to distress during the first learning trial. Most participants (35/40, 88%) completed all 4 verbal memory tasks, with 2 (5%) participants completing only 2 of the tasks due to work or childcare commitments. Overall, the data completeness rate was 90%. In total, 3 (8%) participants reported finding the task difficult, 2 (5%) reported feeling distracted during the encoding tasks, 2 (5%) reported finding the encoding task distracting rather than helpful, and 2 (5%) reported finding that usual memory techniques such as the story technique or rehearsing words in an auditory loop were difficult to do in this task.

Web-Based Cognitive Tests

Few participants completed all 21 cognitive tasks: 6 (15%) out of 40 participants completed all 3 cognitive tasks on all 21 occasions. In total, 7 (18%) participants provided complete data for choice reaction time tasks, 8 (20%) for the self-ordered search, and 9 (23%) for the digit span task. However, on average data completeness was good, with 79.8% of allocated web-based cognitive tasks completed. Only 1 (3%) participant with PD attempted the tasks but discontinued due to frustration at not being able to do the tasks, as they required speed and accuracy. Several participants anecdotally reported misunderstanding the choice reaction time task and clicking outside of the response window but these instances were not formally recorded.

Saliva Samples

Overview

The outcomes for saliva and blood samples are summarized in Table 3.

XSL•FO
RenderX

Table 3. The outcomes relating to data quality and adherence for biomarker analysis: blood biomarkers, salivary dim-light melatonin, and salivary cortisol awakening response.

Study task (sample type) and description; adherence or data quality metric	AD[a] (n=10)	LBD[b] (n=10)	Controls (n=20)	Total (n=40)
Neurodegenerative and AD biomarker samples (blood): participants were asked to have blood drawn for biomarker testing				
Participants who provided blood for biomarker testing, n (%)	9 (90)	10 (100)	18 (90)	37 (92)
Samples analyzed for all biomarkers, n (%)	8 (80)	8 (80)	14 (70)	30 (75)
Samples analyzed for ≥1 biomarker, n (%)	9 (90)	10 (100)	18 (90)	37 (92)
Melatonin samples (passive drool): participants were asked to complete 7 hourly samples across 1 evening, starting from 5 h before bed, and record the sample time				
Analyzable samples provided per participant, mean (SD)	6.3 (2.2)	6.9 (0.3)	7.0 (0)	6.8 (1.1)
Participants who completed all 7 samples, n (%)	9 (90)	9 (90)	20 (100)	38 (95)
Average discrepancy between scheduled time and reported time of completion (min)[c], mean (SD)	5.6 (53.4)	−13.9 (32.6)	−9.2 (37.1)	−7.0 (41)
Cortisol samples (oral swab): participants were asked to complete 3 samples across 1 morning: upon waking up, after 30 min, and after 60 min, and record the sample time				
Analyzable samples provided per participant, mean (SD)	2.7 (0.9)	3.0 (0)	3.0 (0)	2.9 (0.5)
Participants who completed all 3 samples, n (%)	9 (90)	10 (100)	20 (100)	39(98)
Average discrepancy between scheduled time and time completed (min)[c], mean (SD)	22.5 (27.8)	24.3 (17)	24.6 (23.8)	24.0 (23.2)

[a]AD: Alzheimer disease.

[b]LBD: Lewy body dementia.

[c]Positive numbers here refer to a sample being taken later than scheduled, and negative numbers refer to a sample being taken earlier than scheduled.

Dim-Light Melatonin Assay: Passive Drool Samples

Most saliva samples (272/280, 97.1%) were collected as requested. Of 40 participants, 1 (3%) participant provided no saliva samples, and 1 (3%) participant provided 9 out of 10 samples. Data completeness rate for analyzable samples was 92.5% (259/280) across the cohort. Participant-reported sample times were available for 86.7% (236/272) of the total sample. In total, 50.8% (120/236) of the total sample were reported to have been completed within 15 minutes of the scheduled time, as calculated by average bedtimes reported in the sleep diary. On average, participants completed their samples around 7 minutes early (Table 3). Previous studies have identified highly variable salivary melatonin secretion, with peaks from 2 pg/mL to 84 pg/mL [57]. Across the cohort, values ranged from beneath detectable levels (<1.37 pg/mL) to 92.4 pg/mL.

Cortisol Awakening Response: Oral Saliva Swabs

Participants were instructed to provide saliva samples upon waking up, after 30 minutes, and after 60 minutes on one morning during the intensive week. Data completeness was 97.5%, with missing data from only 1 participant who did not provide any saliva samples. Data completeness rate for analyzable samples was 91.7% across the cohort. We compared reported saliva sample timings to their scheduled timings, based on final awakening time estimated by EEG. Participant-reported sample times and EEG awakening time were available for 72.6% (85/117) of cortisol samples. From the 85 samples with timings available, suggested wake times were later than the first sample recording time for 6 samples (7%) and were removed from analysis due to suspected error in recorded saliva timing or date. Participants recorded their first cortisol saliva swab an average of 23 minutes after awakening and overall cortisol samples were generally 24 minutes later than scheduled across the 3 time points (Table 3). Previous studies have identified variable salivary cortisol awakening values from 3 to 19 µg/L in healthy adults [58]. Across the cohort, values ranged from 0.9 to 13.4 µg/L.

Associations Between Participant Characteristics and Adherence

There was some evidence of a weak correlation between MoCA score at baseline and EEG record quality, where those with greater cognitive impairment had lower average EEG record quality (r=0.31; *P*=.05). Spearman rank correlations revealed no evidence of correlations between EEG record quality and age (r=−0.10; *P*=.54); apathy (r=0.16; *P*=.31); or PSQI score (r=0.19; *P*=.24). There was also no evidence of correlation between the number of sleep diaries completed and age (r=−0.15; *P*=.35); MoCA score (r=0.11; *P*=.48); apathy (r=0.07; *P*=.69); or PSQI score (r=0.12; *P*=.47).

Resource Use

Support From Partners and Relatives

In total, 36% (4/11) of individuals with AD, 85% (11/13) of individuals with PD or LBD, and 10% (2/20) of controls attended their consent visit with a partner or relative. In total, 80% (32/40) of participants reported that there would be someone external to the research team (eg, partner or relative) who could support them with study tasks if needed at baseline. However, only 30% (12/40) of participants reported receiving support with completing study tasks (n=3, 30% AD; n=8, 80% LBD; and n=1, 5% control). Support from outside of the study

XSL•FO
RenderX

team was predominantly from partners and included reminders to complete tasks, setting up devices, and troubleshooting technical problems. Of 40 participants, 3 (8%) participants who required technical support from a partner or relative, or used their devices to complete study tasks, reported missing study tasks due to their partner or relative being unavailable during the study period. The participant who withdrew following baseline reported having someone who could support them at home.

Support and Contact With the Research Team

Participants received in person and remote training and support from a researcher to complete study tasks at baseline (including downloading and setting up apps and turning on recording devices) and were offered a refresher training session before the intensive week. Participants were also provided with a written instruction manual, scheduled in-app and email reminders, and ad hoc support as requested by participants or where the team noticed ≥3 consecutive days of missing sleep diary data. All participants had at least 1 in-person visit from a researcher (eg, to download midpoint actigraphy data or retrieve study equipment) during the study period. Participants were advised to contact the research team if they had questions during the study by their preferred method (email, phone, video call, or instant messaging on the MyDignio app). Support from the research team during the study was most often provided by email and involved responding to participant queries on initial setup (eg, downloading the MyDignio app); reminders (eg, tasks to complete or passwords); and technical issues (eg, links not arriving for Cognitron). Some participants also requested telephone-based, video-based, or home-based support (eg, to refresh training on the intensive week study tasks). Most participants did not receive regular reminders from the study team to complete tasks.

Device Use

One participant was provided with a study tablet upon request, as they wanted to keep study activities and apps separate to their personal devices. All other participants used their own smart devices for study activities (ie, smartphones, tablets, and personal computers). Most participants used >1 device to complete web-based study tasks due to personal preference or convenience. No devices provided by the study team were lost or damaged during the study, although some participants cut the actigraphy watch strap for comfort.

Discussion

Principal Findings

In this study, we show that it is feasible to remotely measure sleep and cognition longitudinally in community-dwelling older adults with MCI and dementia due to AD and LBD and healthy older adults. Eligible participants were interested in, enrolled in, and remained in the study, despite being asked to complete a high volume of remote and novel study tasks across an 8-week period while continuing their usual routines. Only 1 participant withdrew from the study, and this was before remote data collection started. Most participants were receptive to multimodal home-based research using technology and alternatives to in-laboratory sleep assessment. Across the cohort, data completeness rate was high and ranged from 79.8% to 97.1%. Our findings support the use of remote, technology-supported research methods to study natural sleep in future trials and indicate some areas where further improvements and refinements would be helpful. With just under two-thirds of our sample (25/40, 63%) having sleep apnea, and only 3 (8%) being aware of this before joining the study, our results also highlight the importance of sleep apnea screening in older adults, as a risk factor and potentially reversible contributor to cognitive impairment and poor sleep quality [59].

Many older adults with MCI and dementia routinely leverage technology for cognitive stimulation, performing activities of daily living (such as shopping and banking), entertainment, and socializing, and use assistive-technology solutions, such as reminders and navigation to support independence [60]. However, cognitive impairment might impact ability to understand or remember to complete remote or technology-based study tasks, resulting in missing data, and may trigger anxiety if patients feel unable to perform tasks correctly [61]. Although a few previous studies had examined the use of sleep wearables in adults with AD, this was typically done over a brief period, examined feasibility of using only 1 digital health technology, or required input from a carer or study partner [31-34,62]. Our study demonstrated that older adults with MCI and dementia can successfully complete novel multimodal sleep and cognitive assessments, including longitudinal concurrent use of wrist actigraphy, web-based cognitive tests, digital sleep diaries, and wireless EEG headbands. However, more effort will be required to recruit samples which are representative of the older adult population and those with MCI or dementia, as our study predominantly recruited White male individuals who were familiar with smart technology, albeit not the devices used in RESTED.

Data completeness rates, reflecting both the ability to complete study tasks and produce data of analyzable quality, were high across all study tasks and participants. Remote data collection comes with a risk of missing data or nonadherence; however, we were able to mitigate this risk by using devices which upload data to servers at regular intervals (eg, Dreem and Dignio) and, less conveniently, by downloading data manually from devices to check compliance and data quality and offer support where needed (eg, actigraphy). Passive monitoring devices and devices that upload to servers automatically can simultaneously reduce the burden on participants and researchers while minimizing data loss by allowing researchers to monitor and respond quickly to any user or technical issues.

Reasons for missing data were usually known and typically related to infrequent but intentional decisions to remove a device (eg, due to a social event or illness) or, in most cases, technical issues with devices or software beyond the participants' control, which would not introduce bias. We observed a weak correlation between EEG record quality and baseline cognitive impairment, which could be explained by cognitive impairment impacting ability to correctly wear the EEG headband for maximal impedance or reflect that patients were moving more during the night. Even subtle movements can shift the headband and reduce

XSL•FO
RenderX

signal quality and getting optimal signal quality across entire or several recordings is a recognized challenge for dry EEG [52]. More comprehensive cognitive assessment than the MoCA, particularly a deeper assessment of executive dysfunction and long-term memory, might reveal a stronger relationship between cognitive impairment and adherence. However, recent studies using remote monitoring technologies found that, although patients with more advanced neurodegenerative disease or cognitive impairment reported more problems and were less compliant, the study remained feasible [63,64]. Larger studies may also want to consider regression analyses to identify key predictors of adherence to identify where support might be most indicated by the research team or caregiver support. However, importantly, overall record quality remained high across the cohort. Having MCI or mild dementia did not preclude study participation, adherence to remote supervised and unsupervised study tasks, or obtaining good-quality data.

Use of remote sleep and cognitive monitoring can help to partly or fully decentralize research, as participants can use their own devices or be provided with technology via post where needed [65], which is more scalable for clinical trials [66]. Frequent or lengthy clinic visits, particularly those that require overnight stays for sleep analysis, can be burdensome for participants, and may reduce inclusivity by requiring participants to live near a study site or travel long distances. More frequent but briefer remote assessments may also help to mitigate against the risk of large amounts of missing data resulting from a missed clinic appointment, can act as “digital biomarkers” that may detect changes more sensitively than an annual follow-up [67], and are more convenient for participants [68]. Although not the focus of our work, research could also explore the use of digital tools to assess fluctuations in cognition over the day in LBD (where fluctuation in cognition is a key diagnostic criteria) and AD (where a sundowning effect of heightened distress is often observed during the evening hours), both of which are difficult to assess in clinical settings [69,70].

While high adherence has been previously demonstrated in feasibility studies of remote monitoring technologies in participants with cognitive impairment or neurodegenerative conditions, most studies have required intensive study partner support throughout the study period as an eligibility criterion [63,71]. It has been argued that study partners are essential for participant safety and well-being in dementia research, even at the preclinical stage where cognitively normal individuals are enrolled [72]. For some participants, partners or relatives provided essential support for study participation. However, many participants provided good-quality data with minimal to no input from others. Crucially, several participants would not have been able to participate if there had been a study partner requirement. Strict eligibility criteria for trials are a recognized barrier to research participation [73]. Requiring a study partner for all studies may unnecessarily increase the burden for relatives already providing care or support, reducing the participant pool, and may disproportionately affect different groups in society who are underserved by research [74,75]. Our findings encourage researchers to consider inclusive and flexible research designs, including options to formally recruit and collect data from study partners without excluding participants who do not have someone available to support them in research studies.

There are an increasing number of options when considering how to measure sleep and cognition from home, including both research-grade and consumer sleep trackers [76,77]. We opted to use existing technologies that were noninvasive, affordable, required minimal training or supervision, were commercially available research-grade or consumer-grade devices that met necessary data privacy regulations, and provided data in a format that could be analyzed using open-source software. Existing technologies can be implemented immediately, reducing time and cost to setup research studies, have a more mature user interface, and are less error-prone than a newly developed solution, and crucially, may be easier to compare among studies for meta-analysis. Several study tasks required ongoing technical support or services from the manufacturers or developers. As a customer rather than a collaborator, we could not always identify or respond to technical problems, and with some providers, we were not informed in advance of several significant changes, which impacted the study or its participants resulting in data loss. Collaborating with industry, ideally at an early stage of the research process, might ensure longevity and continued support throughout the life course of the study and increased flexibility to adapt technology to better fit research (eg, data validation for digital patient-reported outcomes).

Limitations

Women, adults aged >80 years, and minoritized ethnic groups were underrepresented in our study, which is often observed in dementia research [78]. Although some barriers (eg, mistrust and accessibility of research) and motivators (eg, altruism) have been identified, further work is needed to identify how to improve research access, inclusion, and participation in dementia research [79]. There may also have been a self-selection bias, whereby individuals who were likely to be more competent and enthusiastic about using technology volunteered to participate in the study. Most of our participants were regular smartphone users, and several prospective participants declined to participate due to concerns regarding their confidence or interest in using the technologies. Although the proportion of older adults in the United Kingdom using the internet is increasing, a significant proportion of older adults do not regularly use the internet or lack fundamental skills such as being able to turn on a device and enter login details [28]. Older adults with MCI and dementia who regularly use devices report that smartphones and tablets can be useful to support activities of daily living, as well as for communication, entertainment, and recreation, but also list concerns, including cybersecurity and vulnerability to fraud [80]. Education around sleep and brain health, and basic digital skills training on how devices can support individuals living with cognitive impairment, may improve perceptions of capability and motivation to participate in similar studies [56,80]. Exploring the feasibility of more passive monitoring technologies, such as mattress sensors or smartphone-based passive sensing, may help to increase participation and confidence from those with less experience or interest in technology. Although these techniques may invite additional concerns around data privacy, passive monitoring technologies are already being used to support older adults aging in place

XSL•FO
RenderX

[81,82]. Increasing the pool of prospective participants through recruiting via multiple sites and meaningful engagement with community groups, and considering digital inclusivity is likely to enhance recruitment in future studies. Finally, we did not set a priori feasibility cutoffs for study tasks. Despite these limitations, the high completion and retention rate from the RESTED study suggests that remote sleep and cognitive monitoring could offer detailed sleep profiling suitable for tracking change in sleep over disease progression or for monitoring change in sleep clinical trials in older adults with or at risk of dementia.

Conclusions

Practical, detailed, and scalable assessment of sleep is essential for understanding how sleep disturbances affect neurodegeneration over time and for developing and evaluating effective interventions. Our results suggest that older adults with MCI and dementia, as well as healthy older adults, can and do engage in multimodal remote sleep and cognitive research, including using wireless EEG, actigraphy, and mobile apps or web applications. Remote monitoring technologies and research designs offer the opportunity to study natural sleep and its relationship with cognition and dementia over extended periods, are scalable, and should be considered when designing future clinical trials in sleep and dementia in these populations.

Acknowledgments

The authors sincerely thank the Remote Evaluation of Sleep to Enhance Understanding of Early Dementia (RESTED) study participants and their relatives and partners who supported them to take part. They would also like to thank the ReMemBr group lived experience experts for their contributions to improving the study design and supporting them to make the study as accessible as possible. They would also like to thank the North Bristol NHS Trust Respiratory Physiology department and the Bristol Bioresource Laboratories, Anglia Ruskin University Biomarker Lab, and UK DRI Fluid Biomarker Laboratory for providing their services and support for our saliva and blood biomarker analyses.

Finally, the authors would like to thank the teams at Dreem Research, Cognitron, Join Dementia Research, and Ewa G Truchanowicz and Phil Reay, and their teams at Dignio, United Kingdom.

The RESTED study was funded by charitable organizations (BRACE, Alzheimer's Research UK, the Bristol and Weston Hospitals Charity [previously Above and Beyond], and the David Telling Charitable Trust), the NIHR Bristol Biomedical Research Centre, and philanthropic donations from S Scobie and A Graham. JB also received funding from Alzheimer's Research United Kingdom (ARUK) supported by the Margaret Jost Fellowship and the Don Thoburn Memorial Scholarship. The funders had no role in the trial design, patient recruitment, data collection, analysis, interpretation, writing of the manuscript, or the decision to submit it for publication.

Data Availability

Data are available from the author upon reasonable request.

Authors' Contributions

VGG contributed to conceptualization, methodology, software, project administration, resources, investigation, data curation, formal analysis, visualization, writing—original draft preparation, and writing—review and editing. JB contributed to conceptualization, methodology, software, formal analysis, investigation, writing—review and editing, and funding acquisition. HM contributed to conceptualization, methodology, investigation, writing—review and editing, and funding acquisition. HL contributed to investigation, project administration, and formal analysis. AK contributed to conceptualization, methodology, and formal analysis. NT contributed to methodology, funding acquisition, and writing—review and editing. RG contributed to conceptualization, funding acquisition. BB contributed to methodology. EC contributed to conceptualization, methodology, resources, writing—review and editing, supervision, and funding acquisition.

Conflicts of Interest

EC has received funding from Biogen, Eisai, and Lilly for consultancy and providing educational resources. AH has received funding from Quanterix Corp for consultancy. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) Statement—checklist of items that should be included in reports of observational studies.
[DOCX File , 19 KB - aging_v8i1e72824_app1.docx]

Multimedia Appendix 2
Further information is provided on the devices used to measure sleep and circadian rhythms.

XSL•FO
RenderX

[DOCX File , 13 KB - aging_v8i1e72824_app2.docx]

Multimedia Appendix 3
Dreem electroencephalography individual channel metrics.
[DOCX File , 14 KB - aging_v8i1e72824_app3.docx]

Multimedia Appendix 4
Additional information on reasons for missing data and the affected study tasks is provided.
[DOCX File , 15 KB - aging_v8i1e72824_app4.docx]

References

1. Koren T, Fisher E, Webster L, Livingston G, Rapaport P. Prevalence of sleep disturbances in people with dementia living in the community: a systematic review and meta-analysis. Ageing Res Rev 2023 Jan;83:101782 [FREE Full text] [doi: 10.1016/j.arr.2022.101782] [Medline: 36356799]
2. Postuma RB, Iranzo A, Hu M, Högl B, Boeve BF, Manni R, et al. Risk and predictors of dementia and parkinsonism in idiopathic REM sleep behaviour disorder: a multicentre study. Brain 2019 Mar 01;142(3):744-759 [FREE Full text] [doi: 10.1093/brain/awz030] [Medline: 30789229]
3. Harenbrock J, Holling H, Reid G, Koychev I. A meta-analysis of the relationship between sleep and β-Amyloid biomarkers in Alzheimer's disease. Biomark Neuropsychiatry 2023 Dec;9:100068. [doi: 10.1016/j.bionps.2023.100068]
4. Zamore Z, Veasey SC. Neural consequences of chronic sleep disruption. Trends Neurosci 2022 Sep;45(9):678-691 [FREE Full text] [doi: 10.1016/j.tins.2022.05.007] [Medline: 35691776]
5. Zielinski MR, Gibbons AJ. Neuroinflammation, sleep, and circadian rhythms. Front Cell Infect Microbiol 2022 Mar 22;12:853096 [FREE Full text] [doi: 10.3389/fcimb.2022.853096] [Medline: 35392608]
6. Eshera YM, Gavrilova L, Hughes JW. Sleep is essential for cardiovascular health: an analytic review of the relationship between sleep and cardiovascular mortality. Am J Lifestyle Med 2024 Oct 31;18(3):340-350. [doi: 10.1177/15598276231211846] [Medline: 38737888]
7. Seda G, Matwiyoff G, Parrish JS. Effects of obstructive sleep apnea and CPAP on cognitive function. Curr Neurol Neurosci Rep 2021 May 06;21(7):32. [doi: 10.1007/s11910-021-01123-0] [Medline: 33956247]
8. Costa YS, Lim AS, Thorpe KE, Colelli DR, Mitchell S, Masellis M, et al. Investigating changes in cognition associated with the use of CPAP in cognitive impairment and dementia: a retrospective study. Sleep Med 2023 Jan;101:437-444. [doi: 10.1016/j.sleep.2022.11.037] [Medline: 36516600]
9. Liu WT, Huang HT, Hung HY, Lin SY, Hsu WH, Lee FY, et al. Continuous positive airway pressure reduces plasma neurochemical levels in patients with OSA: a pilot study. Life (Basel) 2023 Feb 22;13(3):613 [FREE Full text] [doi: 10.3390/life13030613] [Medline: 36983769]
10. Sabia S, Fayosse A, Dumurgier J, van Hees VT, Paquet C, Sommerlad A, et al. Association of sleep duration in middle and old age with incidence of dementia. Nat Commun 2021 Apr 20;12(1):2289 [FREE Full text] [doi: 10.1038/s41467-021-22354-2] [Medline: 33879784]
11. Mayer G, Frohnhofen H, Jokisch M, Hermann DM, Gronewold J. Associations of sleep disorders with all-cause MCI/dementia and different types of dementia - clinical evidence, potential pathomechanisms and treatment options: a narrative review. Front Neurosci 2024 Mar 22;18:1372326 [FREE Full text] [doi: 10.3389/fnins.2024.1372326] [Medline: 38586191]
12. Blackman J, Swirski M, Clynes J, Harding S, Leng Y, Coulthard E. Pharmacological and non-pharmacological interventions to enhance sleep in mild cognitive impairment and mild Alzheimer's disease: a systematic review. J Sleep Res 2021 Aug;30(4):e13229 [FREE Full text] [doi: 10.1111/jsr.13229] [Medline: 33289311]
13. Blackman J, Morrison HD, Lloyd K, Gimson A, Banerjee LV, Green S, et al. The past, present, and future of sleep measurement in mild cognitive impairment and early dementia-towards a core outcome set: a scoping review. Sleep 2022 Jul 11;45(7):3-11 [FREE Full text] [doi: 10.1093/sleep/zsac077] [Medline: 35373837]
14. Kainec KA, Caccavaro J, Barnes M, Hoff C, Berlin A, Spencer RM. Evaluating accuracy in five commercial sleep-tracking devices compared to research-grade actigraphy and polysomnography. Sensors (Basel) 2024 Jan 19;24(2):635 [FREE Full text] [doi: 10.3390/s24020635] [Medline: 38276327]
15. Tadokoro K, Ohta Y, Hishikawa N, Nomura E, Wakutani Y, Takao Y, et al. Discrepancy of subjective and objective sleep problems in Alzheimer's disease and mild cognitive impairment detected by a home-based sleep analysis. J Clin Neurosci 2020 Apr;74:76-80. [doi: 10.1016/j.jocn.2020.01.085] [Medline: 32007378]
16. Casagrande M, Forte G, Favieri F, Corbo I. Sleep quality and aging: a systematic review on healthy older people, mild cognitive impairment and Alzheimer's disease. Int J Environ Res Public Health 2022 Jul 11;19(14):8457 [FREE Full text] [doi: 10.3390/ijerph19148457] [Medline: 35886309]
17. Williams JM, Kay DB, Rowe M, McCrae CS. Sleep discrepancy, sleep complaint, and poor sleep among older adults. J Gerontol B Psychol Sci Soc Sci 2013 Sep 26;68(5):712-720 [FREE Full text] [doi: 10.1093/geronb/gbt030] [Medline: 23804432]

XSL•FO
RenderX

18. Markun LC, Sampat A. Clinician-focused overview and developments in polysomnography. Curr Sleep Med Rep 2020 Nov 23;6(4):309-321 [FREE Full text] [doi: 10.1007/s40675-020-00197-5] [Medline: 33251088]
19. Goparaju B, de Palma G, Bianchi MT. Naturalistic sleep tracking in a longitudinal cohort: how long is long enough? medRxiv Preprint posted online October 21, 2024 [FREE Full text] [doi: 10.1101/2024.10.19.24315818]
20. Blackman J, Morrison H, Lloyd K, Gimson A, Banerjee L, Green S, et al. Sleep measurement heterogeneity in mild cognitive impairment and early dementia - towards a core outcome set: a scoping review. Sleep 2022 Dec;100:S148-S149. [doi: 10.1016/j.sleep.2022.05.403]
21. Green SF, Frame T, Banerjee LV, Gimson A, Blackman J, Morrison H, et al. A systematic review of the validity of non-invasive sleep-measuring devices in mid-to-late life adults: future utility for Alzheimer's disease research. Sleep Med Rev 2022 Oct;65:101665 [FREE Full text] [doi: 10.1016/j.smrv.2022.101665] [Medline: 36208588]
22. Kourtis LC, Regele OB, Wright JM, Jones GB. Digital biomarkers for Alzheimer's disease: the mobile/ wearable devices opportunity. NPJ Digit Med 2019 Feb 21;2(1):9 [FREE Full text] [doi: 10.1038/s41746-019-0084-2] [Medline: 31119198]
23. Arnal PJ, Thorey V, Debellemaniere E, Ballard ME, Bou Hernandez A, Guillot A, et al. The Dreem Headband compared to polysomnography for electroencephalographic signal acquisition and sleep staging. Sleep 2020 Nov 12;43(11):zsaa097 [FREE Full text] [doi: 10.1093/sleep/zsaa097] [Medline: 32433768]
24. Miller DJ, Sargent C, Roach GD. A validation of six wearable devices for estimating sleep, heart rate and heart rate variability in healthy adults. Sensors (Basel) 2022 Aug 22;22(16):6317 [FREE Full text] [doi: 10.3390/s22166317] [Medline: 36016077]
25. Cay G, Ravichandran V, Sadhu S, Zisk AH, Salisbury AL, Solanki D, et al. Recent advancement in sleep technologies: a literature review on clinical standards, sensors, apps, and AI methods. IEEE Access 2022;10:104737-104756. [doi: 10.1109/access.2022.3210518]
26. Ding Z, Lee T, Chan AS. Digital cognitive biomarker for mild cognitive impairments and dementia: a systematic review. J Clin Med 2022 Jul 19;11(14):4191 [FREE Full text] [doi: 10.3390/jcm11144191] [Medline: 35887956]
27. Masanneck L, Gieseler P, Gordon WJ, Meuth SG, Stern AD. Evidence from ClinicalTrials.gov on the growth of digital health technologies in neurology trials. NPJ Digit Med 2023 Feb 10;6(1):23 [FREE Full text] [doi: 10.1038/s41746-023-00767-1] [Medline: 36765123]
28. Briefing: facts and figures about digital inclusion and older people. Age UK. URL: https://www.ageuk.org.uk/siteassets/documents/reports-and-publications/reports-and-briefings/active-communities/internet-use-statistics-june-2024.pdf [accessed 2024-05-29]
29. Holthe T, Halvorsrud L, Karterud D, Hoel K, Lund A. Usability and acceptability of technology for community-dwelling older adults with mild cognitive impairment and dementia: a systematic literature review. Clin Interv Aging 2018;13:863-886 [FREE Full text] [doi: 10.2147/CIA.S154717] [Medline: 29765211]
30. Chien SY, Zaslavsky O, Berridge C. Technology usability for people living with dementia: concept analysis. JMIR Aging 2024 Jul 03;7:e51987 [FREE Full text] [doi: 10.2196/51987] [Medline: 38959053]
31. Kent BA, Casciola AA, Carlucci SK, Chen M, Stager S, Mirian MS, et al. Home EEG sleep assessment shows reduced slow-wave sleep in mild-moderate Alzheimer's disease. Alzheimers Dement (N Y) 2022 Aug 16;8(1):e12347 [FREE Full text] [doi: 10.1002/trc2.12347] [Medline: 35992215]
32. Pavlickova H, Russell AE, Lightman S, McCabe R. Feasibility of salivary cortisol collection in patients and companions attending dementia diagnostic meetings in memory clinics. BMC Res Notes 2021 Jan 21;14(1):30 [FREE Full text] [doi: 10.1186/s13104-021-05446-6] [Medline: 33478564]
33. Jones C, Moyle W. A feasibility study of Dreampad™ on sleep, wandering and agitated behaviors in people living with dementia. Geriatr Nurs 2020 Nov;41(6):782-789. [doi: 10.1016/j.gerinurse.2020.04.014] [Medline: 32522427]
34. Guu TW, Brem A, Albertyn CP, Kandangwa P, Aarsland D, Ffytche D. Wrist-worn actigraphy in agitated late-stage dementia patients: a feasibility study on digital inclusion. Alzheimers Dement 2024 May 18;20(5):3211-3218 [FREE Full text] [doi: 10.1002/alz.13772] [Medline: 38497216]
35. Gabb VG, Blackman J, Morrison HD, Biswas B, Li H, Turner N, et al. Remote evaluation of sleep and circadian rhythms in older adults with mild cognitive impairment and dementia: protocol for a feasibility and acceptability mixed methods study. JMIR Res Protoc 2024 Mar 22;13:e52652 [FREE Full text] [doi: 10.2196/52652] [Medline: 38517469]
36. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007 Oct;370(9596):1453-1457. [doi: 10.1016/s0140-6736(07)61602-x]
37. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011 May 22;7(3):263-269 [FREE Full text] [doi: 10.1016/j.jalz.2011.03.005] [Medline: 21514250]
38. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011 May;7(3):270-279 [FREE Full text] [doi: 10.1016/j.jalz.2011.03.008] [Medline: 21514249]

XSL•FO
RenderX

39. Yamada M, Komatsu J, Nakamura K, Sakai K, Samuraki-Yokohama M, Nakajima K, et al. Diagnostic criteria for dementia with Lewy bodies: updates and future directions. J Mov Disord 2020 Jan;13(1):1-10 [FREE Full text] [doi: 10.14802/jmd.19052] [Medline: 31694357]
40. Litvan I, Goldman JG, Tröster AI, Schmand BA, Weintraub D, Petersen RC, et al. Diagnostic criteria for mild cognitive impairment in Parkinson's disease: movement disorder society task force guidelines. Mov Disord 2012 Mar 24;27(3):349-356 [FREE Full text] [doi: 10.1002/mds.24893] [Medline: 22275317]
41. Blackman J, Morrison HD, Gabb V, Biswas B, Li H, Turner N, et al. Remote evaluation of sleep to enhance understanding of early dementia due to Alzheimer's Disease (RESTED-AD): an observational cohort study protocol. BMC Geriatr 2023 Sep 23;23(1):590 [FREE Full text] [doi: 10.1186/s12877-023-04288-0] [Medline: 37742001]
42. Blackman J, Gabb VG, Carrigan N, Wearn A, Meky S, Selwood J, et al. Sleep quality during and after severe acute respiratory syndrome coronavirus 2 (COVID-19) lockdowns in the UK: results from the SleepQuest study. J Sleep Res 2024 Dec 23;33(6):e14205. [doi: 10.1111/jsr.14205] [Medline: 38650540]
43. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
44. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
45. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep 1991 Dec;14(6):540-545. [doi: 10.1093/sleep/14.6.540] [Medline: 1798888]
46. Chung F, Abdullah HR, Liao P. STOP-bang questionnaire: a practical approach to screen for obstructive sleep apnea. Chest 2016 Mar;149(3):631-638 [FREE Full text] [doi: 10.1378/chest.15-0903] [Medline: 26378880]
47. Yesavage JA, Sheikh JI. Geriatric Depression Scale (GDS): recent findings and development of a shorter version. In: Brink TL, editor. Clinical Gerontology: A Guide to Assessment and Intervention. Oxfordshire, UK: Routledge; Oct 25, 2008:165-173.
48. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
49. Marin RS, Biedrzycki RC, Firinciogullari S. Reliability and validity of the apathy evaluation scale. Psychiatry Res 1991 Aug;38(2):143-162. [doi: 10.1016/0165-1781(91)90040-v] [Medline: 1754629]
50. Carney CE, Buysse DJ, Ancoli-Israel S, Edinger JD, Krystal AD, Lichstein KL, et al. The consensus sleep diary: standardizing prospective sleep self-monitoring. Sleep 2012 Feb 01;35(2):287-302 [FREE Full text] [doi: 10.5665/sleep.1642] [Medline: 22294820]
51. Cognitron. URL: https://www.cognitron.co.uk/ [accessed 2025-05-29]
52. Ravindran KKG, Della Monica C, Atzori G, Nilforooshan R, Hassanin H, Revell V, et al. Evaluation of Dreem headband for sleep staging and EEG spectral analysis in people living with Alzheimer's and older adults. Sleep 2025 May 04:9240. [doi: 10.1093/sleep/zsaf122] [Medline: 40319389]
53. González DA, Wang D, Pollet E, Velarde A, Horn S, Coss P, et al. Performance of the Dreem 2 EEG headband, relative to polysomnography, for assessing sleep in Parkinson's disease. Sleep Health 2024 Feb;10(1):24-30. [doi: 10.1016/j.sleh.2023.11.012] [Medline: 38151377]
54. Migueles JH, Rowlands AV, Huber F, Sabia S, van Hees VT. GGIR: a research community–driven open source R package for generating physical activity and sleep outcomes from multi-day raw accelerometer data. J Meas Phys Behav 2022 Jul 01;31(5):629 [FREE Full text] [doi: 10.1123/jmpb.2018-0063]
55. van Hees VT, Sabia S, Jones SE, Wood AR, Anderson KN, Kivimäki M, et al. Estimating sleep parameters using an accelerometer without sleep diary. Sci Rep 2018 Aug 28;8(1):12975 [FREE Full text] [doi: 10.1038/s41598-018-31266-z] [Medline: 30154500]
56. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6(1):42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
57. Burgess HJ, Fogg LF. Individual differences in the amount and timing of salivary melatonin secretion. PLoS One 2008 Aug 26;3(8):e3055 [FREE Full text] [doi: 10.1371/journal.pone.0003055] [Medline: 18725972]
58. Trilck M, Flitsch J, Lüdecke DK, Jung R, Petersenn S. Salivary cortisol measurement--a reliable method for the diagnosis of Cushing's syndrome. Exp Clin Endocrinol Diabetes 2005 Apr 13;113(4):225-230. [doi: 10.1055/s-2005-837667] [Medline: 15891959]
59. Naismith S, Winter V, Gotsopoulos H, Hickie I, Cistulli P. Neurobehavioral functioning in obstructive sleep apnea: differential effects of sleep quality, hypoxemia and subjective sleepiness. J Clin Exp Neuropsychol 2004 Feb 09;26(1):43-54. [doi: 10.1076/jcen.26.1.43.23929] [Medline: 14972693]
60. Holthe T, Halvorsrud L, Lund A. Digital assistive technology to support everyday living in community-dwelling older adults with mild cognitive impairment and dementia. Clin Interv Aging 2022;17:519-544 [FREE Full text] [doi: 10.2147/CIA.S357860] [Medline: 35464157]

XSL•FO
RenderX

61. Ahuja M, Siddhpuria S, Reppas-Rindlisbacher C, Wong E, Gormley J, Lee J, et al. Sleep monitoring challenges in patients with neurocognitive disorders: a cross-sectional analysis of missing data from activity trackers. Health Sci Rep 2022 May 26;5(3):e608 [FREE Full text] [doi: 10.1002/hsr2.608] [Medline: 35509396]
62. Van Den Berg JF, Van Rooij FJ, Vos H, Tulen JH, Hofman A, Miedema HM, et al. Disagreement between subjective and actigraphic measures of sleep duration in a population-based study of elderly persons. J Sleep Res 2008 Sep 21;17(3):295-302. [doi: 10.1111/j.1365-2869.2008.00638.x] [Medline: 18321246]
63. Muurling M, de Boer C, Hinds C, Atreya A, Doherty A, Alepopoulos V, RADAR-AD consortium. Feasibility and usability of remote monitoring in Alzheimer's disease. Digit Health 2024 Apr 09;10:20552076241238133 [FREE Full text] [doi: 10.1177/20552076241238133] [Medline: 38601188]
64. Maas BR, Speelberg DH, de Vries G, Valenti G, Ejupi A, Bloem BR, et al. Patient experience and feasibility of a remote monitoring system in Parkinson's disease. Mov Disord Clin Pract 2024 Oct 26;11(10):1223-1231. [doi: 10.1002/mdc3.14169] [Medline: 39056543]
65. Aiyegbusi OL, Davies EH, Myles P, Williams T, Frost C, Haroon S, et al. Digitally enabled decentralised research: opportunities to improve the efficiency of clinical trials and observational studies. BMJ Evid Based Med 2023 Oct 21;28(5):328-331 [FREE Full text] [doi: 10.1136/bmjebm-2023-112253] [Medline: 36810190]
66. Inan OT, Tenaerts P, Prindiville SA, Reynolds HR, Dizon DS, Cooper-Arnold K, et al. Digitizing clinical trials. NPJ Digit Med 2020 Jul 31;3(1):101. [doi: 10.1038/S41746-020-0302-Y]
67. Dodge HH, Zhu J, Mattek NC, Austin D, Kornfeld J, Kaye JA. Use of high-frequency in-home monitoring data may reduce sample sizes needed in clinical trials. PLoS One 2015 Sep 17;10(9):e0138095 [FREE Full text] [doi: 10.1371/journal.pone.0138095] [Medline: 26379170]
68. Black BS, Taylor HA, Rabins PV, Karlawish J. Study partners perform essential tasks in dementia research and can experience burdens and benefits in this role. Dementia (London) 2016 May 13;17(4):494-514 [FREE Full text] [doi: 10.1177/1471301216648796] [Medline: 27179001]
69. O'Dowd S, Schumacher J, Burn DJ, Bonanni L, Onofrj M, Thomas A, et al. Fluctuating cognition in the Lewy body dementias. Brain 2019 Nov 01;142(11):3338-3350. [doi: 10.1093/brain/awz235] [Medline: 31411317]
70. Canevelli M, Valletta M, Trebbastoni A, Sarli G, D'Antonio F, Tariciotti L, et al. Sundowning in dementia: clinical relevance, pathophysiological determinants, and therapeutic approaches. Front Med (Lausanne) 2016 Dec 27;3:73 [FREE Full text] [doi: 10.3389/fmed.2016.00073] [Medline: 28083535]
71. Godkin FE, Turner E, Demnati Y, Vert A, Roberts A, Swartz RH, et al. Feasibility of a continuous, multi-sensor remote health monitoring approach in persons living with neurodegenerative disease. J Neurol 2022 May 27;269(5):2673-2686 [FREE Full text] [doi: 10.1007/s00415-021-10831-z] [Medline: 34705114]
72. Grill JD, Karlawish J. Study partners should be required in preclinical Alzheimer's disease trials. Alzheimers Res Ther 2017 Dec 06;9(1):93 [FREE Full text] [doi: 10.1186/s13195-017-0327-x] [Medline: 29212555]
73. Mitchell AK, Ehrenkranz R, Franzen S, Han SH, Shakur M, McGowan M, et al. Analysis of eligibility criteria in Alzheimer's and related dementias clinical trials. Sci Rep 2024 Jul 01;14(1):15036 [FREE Full text] [doi: 10.1038/s41598-024-65767-x] [Medline: 38951633]
74. Aranda MP, Marquez DX, Gallagher-Thompson D, Pérez A, Rojas JC, Hill CV, et al. A call to address structural barriers to Hispanic/Latino representation in clinical trials on Alzheimer's disease and related dementias: a micro-meso-macro perspective. Alzheimers Dement (N Y) 2023;9(2):e12389 [FREE Full text] [doi: 10.1002/trc2.12389] [Medline: 37287471]
75. Wise-Brown A, Brangman SA, Henderson JN, Willis-Parker M, Monroe S, Mintzer JE, et al. Promoting diversity in clinical trials: insights from planning the ALUMNI AD study in historically underrepresented US populations with early symptomatic Alzheimer's disease. EClinicalMedicine 2024 Jul;73:102693 [FREE Full text] [doi: 10.1016/j.eclinm.2024.102693] [Medline: 39429811]
76. de Zambotti M, Goldstein C, Cook J, Menghini L, Altini M, Cheng P, et al. State of the science and recommendations for using wearable technology in sleep and circadian research. Sleep 2024 Apr 12;47(4):zsad325. [doi: 10.1093/sleep/zsad325] [Medline: 38149978]
77. Lee T, Cho Y, Cha KS, Jung J, Cho J, Kim H, et al. Accuracy of 11 wearable, nearable, and airable consumer sleep trackers: prospective multicenter validation study. JMIR Mhealth Uhealth 2023 Nov 02;11:e50983 [FREE Full text] [doi: 10.2196/50983] [Medline: 37917155]
78. Banzi R, Camaioni P, Tettamanti M, Bertele' V, Lucca U. Older patients are still under-represented in clinical trials of Alzheimer's disease. Alzheimers Res Ther 2016 Aug 12;8(1):32 [FREE Full text] [doi: 10.1186/s13195-016-0201-2] [Medline: 27520290]
79. Gilmore-Bykovskyi AL, Jin Y, Gleason C, Flowers-Benton S, Block LM, Dilworth-Anderson P, et al. Recruitment and retention of underrepresented populations in Alzheimer's disease research: a systematic review. Alzheimers Dement (N Y) 2019 Jan;5(1):751-770 [FREE Full text] [doi: 10.1016/j.trci.2019.09.018] [Medline: 31921966]
80. Wilson SA, Byrne P, Rodgers SE. 'I'd be lost without my smartphone': a qualitative analysis of the use of smartphones and tablets by people living with dementia, mild cognitive impairment, and their caregivers. Aging Ment Health 2024 Apr 04;28(4):595-603 [FREE Full text] [doi: 10.1080/13607863.2023.2205585] [Medline: 37139944]

XSL•FO
RenderX

81. Read E, Woolsey C, Donelle L, Weeks L, Chinho N. Passive remote monitoring and aging in place: a scoping review. Can J Aging 2023 Mar 01;42(1):20-32. [doi: 10.1017/S0714980822000198] [Medline: 35912590]
82. Cornet VP, Holden RJ. Systematic review of smartphone-based passive sensing for health and wellbeing. J Biomed Inform 2018 Jan;77:120-132 [FREE Full text] [doi: 10.1016/j.jbi.2017.12.008] [Medline: 29248628]

Abbreviations

AD: Alzheimer disease
EEG: electroencephalography
LBD: Lewy body disease
MCI: mild cognitive impairment
MoCA: Montreal Cognitive Assessment
PD: Parkinson disease
PPI: patient and public involvement
PSG: polysomnography
PSQI: Pittsburgh Sleep Quality Index
RESTED: Remote Evaluation of Sleep to Enhance Understanding of Early Dementia
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by M O'Connell; submitted 19.02.25; peer-reviewed by N Misra, R Singh, R Tripathi; comments to author 11.03.25; revised version received 28.03.25; accepted 28.04.25; published 28.05.25.

Please cite as:
Gabb VG, Blackman J, Morrison H, Li H, Kendrick A, Turner N, Greenwood R, Biswas B, Heslegrave A, Coulthard E
Longitudinal Remote Sleep and Cognitive Research in Older Adults With Mild Cognitive Impairment and Dementia: Prospective Feasibility Cohort Study
JMIR Aging 2025;8:e72824
URL: https://aging.jmir.org/2025/1/e72824
doi:10.2196/72824
PMID:

©Victoria Grace Gabb, Jonathan Blackman, Hamish Morrison, Haoxuan Li, Adrian Kendrick, Nicholas Turner, Rosemary Greenwood, Bijetri Biswas, Amanda Heslegrave, Elizabeth Coulthard. Originally published in JMIR Aging (https://aging.jmir.org), 28.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Staff Enablement of the Tovertafel for Enrichment in Residential Aged Care: Field Study

Ryan M Kelly[1,2], BSc(Hons), MSc, PhD; Asmita Manchha[3], BCom, BBusMan(Hons), PhD; Jenny Waycott[2], BA(Hons), PhD; Rajna Ogrin[4], BSc, BPod(Hons), PhD; Judy A Lowthian[4,5,6], BAppSc(SpPath), MPH, PhD

[1]School of Computing Technologies, RMIT University, Melbourne, Australia

[2]School of Computing and Information Systems, The University of Melbourne, Melbourne, Australia

[3]Silverchain, Melbourne, Australia

[4]Bolton Clarke, Melbourne, Australia

[5]School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia

[6]Faculty of Health and Behavioural Sciences, The University of Queensland, St Lucia, Australia

Corresponding Author:
Ryan M Kelly, BSc(Hons), MSc, PhD
School of Computing Technologies
RMIT University
124 La Trobe St
Melbourne, 3000
Australia
Phone: 61 (03) 9925 2000
Email: ryan.kelly@rmit.edu.au

Abstract

Background: Enrichment activities are essential for enhancing the psychosocial well-being of older adults living in residential aged care homes. There has been increasing interest in using digital technology for enrichment, but the implementation of technology requires careful support and enablement from staff to ensure that residents experience the intended benefits.

Objective: This study aimed to understand how care staff facilitate aged care residents' use of the Tovertafel ("magic table" in Dutch), a technology that projects images onto a tabletop to enable groups of people to play games. The study further aimed to understand the benefits arising from the Tovertafel when facilitated by staff.

Methods: We conducted a field study in 1 residential aged care home in Queensland, Australia. The methods included semistructured interviews with the staff and residents about their experiences with the Tovertafel, observations of 4 sessions in which the residents and staff played Tovertafel games, and a diary completed by the staff after Tovertafel sessions. Data were analyzed through reflexive thematic analysis.

Results: We developed 3 themes through our analysis. Theme 1 highlights the need for the staff to overcome physical and personal barriers before Tovertafel sessions could take place. These included a lack of a dedicated space for playing Tovertafel games and the residents' reluctance to attend Tovertafel sessions. Theme 2 highlights how the staff used creative strategies to make Tovertafel sessions successful. These included helping the residents learn how to interact with the games; adapting the activity to suit the capabilities of the residents; sustaining engagement by choosing appropriate games; and using prompts, questions, and storytelling to make the games more engaging. Theme 3 describes the benefits and outcomes that arose from staff-supported enablement of the Tovertafel, including participation in an enjoyable physical activity, socialization, and reminiscence.

Conclusions: This study suggests that the Tovertafel provides opportunities for aged care staff to engage in creative play and personalization catering to residents with different capabilities. However, the benefits arising from the Tovertafel are unlikely to be achieved without substantial facilitation from the staff, who play a key role in enabling the participation of the residents. Sustaining the engagement of the residents is important during Tovertafel activities and can lead to beneficial outcomes.

(JMIR Aging 2025;8:e67919) doi:10.2196/67919

KEYWORDS

aged care; care staff; creative care; older adults; play; residential aged care; social enrichment; technology; Tovertafel

XSL•FO
RenderX

Introduction

Background

In 2023, a total of 185,000 Australians aged ≥65 years lived in residential aged care homes [1]. Older individuals usually move to residential care when they require high levels of support to enable appropriate clinical care alongside holistic well-being [2,3]. Holistic well-being refers to addressing residents' physical, psychological, and social needs [4]. To promote physical well-being, staff members from lifestyle teams support residents with activities of daily living, such as bathing, dressing, and eating, and enable them to participate in physical activities through exercise programs.

To promote psychological and social well-being, staff members provide an array of enrichment activities, such as art, gardening, music, games, and excursions. These activities are important for encouraging social interaction and preventing boredom [5], enhancing residents' quality of life [6], and promoting a sense of continuity, particularly when residents engage in activities that interest them [7]. However, activities in residential care are often group based and may not cater to the needs of all residents [6], particularly when there is diversity among residents regarding individual interests [7], cultural background [8], or functional and cognitive capabilities [9,10].

Recent years have seen an increasing uptake of digital technologies for enrichment in residential care homes [11]. Technology-based activities can both expand the range of available options and enable new experiences of joy and social engagement [12]. For example, immersive virtual reality (VR) offers the possibility of virtually "traveling" to distant locations, revisiting places from the past, and remotely attending concerts [13,14]. When used appropriately and in a way that demonstrates person-centered care, technology-based activities can address the individual needs and goals of residents, including those with cognitive impairment [11].

However, research has identified that appropriately trained staff are often required to enable the effective use of technology in aged care [15]. Residential care staff are known to be dealing with high workloads [16], and there is a risk that facilitating technology may add to this load. While technology facilitation can be enriching when it involves meaningful care work [17], significant work in setting up and manipulating the technology can be an unwanted burden that limits the use of the technology [18,19]. Therefore, evaluating the demands placed on care staff, along with the potential benefits of using technology with residents, is an important consideration when justifying the use of technology as part of a care home's lifestyle and social calendar.

The Tovertafel

This study focuses on staff involvement in enabling the Tovertafel ("magic table" in Dutch), a commercial system that enables groups of people to play games by tapping on moving images projected onto a tabletop [20]. The games vary in difficulty, from simple activities such as sweeping leaves or popping bubbles to more advanced puzzles and team-based games that require cognitive and social engagement. While the Tovertafel is aimed at providing stimulation for care home residents living with dementia, it can be used by residents with various levels of cognitive impairment. Table 1 provides examples of Tovertafel games.

Table 1. Examples of Tovertafel games.

Game name	Description
Leaves	Players use arm movements to sweep piles of leaves from the tabletop.
Football	Players try to score goals at either end of the table by tapping a soccer ball.
Music Box	Players tap on musical notes rotating around a music box in the center of the table. The notes play piano sounds when tapped. At the end of the game, the music box reveals a spinning ballerina and plays a song.
Rhymes	Players tap on feathers to reveal the words of a nursery rhyme, which can then be read aloud.
Sayings	Players tap on moving bubbles to reveal a phrase or saying, for example, "Never look a gift horse in the mouth."
Silverware	Players wave their arms over images of cutlery to simulate polishing a dinner set. Once all pieces are polished, plates and food appear on the tabletop, simulating the table being laid at dinner time.
The Veggie Patch	The game simulates gardening by requiring players to plant and water vegetables. Seed packets move over the table, causing vegetables to be planted when tapped. There are also moving clouds and sunbeams, which cause the plants to grow when tapped.

A small number of studies have examined the use of the Tovertafel in residential aged care. These studies have primarily focused on identifying benefits and drawbacks of the technology. Good et al [21] interviewed care staff about the perceived benefits of the Tovertafel for residents living with moderate to advanced dementia. Staff felt that using the Tovertafel had several benefits, including improved communication between staff and residents and short-term improvements in residents' mood. However, the Tovertafel was rarely used by residents without initiation from carers or loved ones, and residents with dementia could only use the Tovertafel for a short time because of impaired concentration. A study by Talman and Gustafsson [22] identified benefits of the Tovertafel for people living with intellectual disabilities. However, these findings may not be applicable to aged care. The Tovertafel's official website lists various studies that have been conducted to validate the system, but most are unpublished, originate from master's theses, or require readers to contact the company to obtain further information [23]. Other than the studies mentioned earlier

XSL•FO
RenderX

[21,22], the only publicly available study evaluating the Tovertafel is in Dutch [24].

While the impact of the Tovertafel may be positive, there has been limited examination of the care strategies required to facilitate the participation of residents in the activity. There is also limited understanding of how staff work to overcome issues that may impact the engagement of aged care residents with the Tovertafel. Research is needed to better understand how staff facilitate the Tovertafel activity and to inform decisions about using this technology, given the limited resources available in many residential aged care homes.

Objectives

This study aimed to address the following research questions:

1. How do care staff enable residents' use of the Tovertafel in residential aged care homes?
2. What benefits arise from staff-supported enablement of the Tovertafel?

Methods

Ethical Considerations

All procedures in this study received approval from the University of Melbourne's Human Research Ethics Committee (ID 12900) and the Bolton Clarke Human Research Ethics Committee (approval 230001). Our study involved participants who were either staff or residents from one aged care home in Australia. The staff participants signed a participant information sheet and consent form (PICF) and received an Aus $20 (US $13.37) gift voucher as a token of appreciation.

The care home residents who participated in the interviews were provided with a PICF that included pictures of Tovertafel games, along with a textual description of the technology. Some of these residents were shown a YouTube video of the Tovertafel if they were unable to understand its functionality from the PICF alone. The residents who agreed to be interviewed provided informed consent by signing the PICF and received baked goods as a token of gratitude. Some residents living in the memory support unit of the care home did not have the capacity to consent, so interviews were not sought about their experiences with the Tovertafel. However, we observed the sessions in which these residents participated and collected observational notes about their use of the Tovertafel, without capturing personal or identifying information, in line with the ethics approvals for the project.

All data from the study was storied securely in password-protected filing cabinets or secure cloud-based storage to protect the confidentiality of the data. All transcripts were anonymized, and names were replaced with pseudonyms to protect participants' privacy.

Study Design

We conducted a field study to investigate how the Tovertafel was used in 1 Australian residential aged care home. The study involved semistructured interviews, observations, and a participant diary. We chose these methods to gain in-depth insight into staff members' experiences of facilitating the Tovertafel and to understand how this facilitation elicits benefits for residents. All procedures were developed by the research team, drawing on our previous experience with the chosen methods [25,26] and our experience of conducting fieldwork in aged care [8,27,28].

First, the staff who were familiar with using the Tovertafel were interviewed about their experiences of using the system, the strategies that they had used to make the activity successful, and the perceived benefits. We also conducted interviews with the residents who had used the Tovertafel, provided they had the capacity to give informed consent and provide feedback. These interviews sought to understand what the residents liked about the activity and what they found challenging. The interview schedule for the residents contained simple questions to avoid overburdening them.

Second, we observed 4 activity sessions, each lasting 45 minutes, in which small groups of staff and residents played a variety of Tovertafel games. These sessions were scheduled by the lifestyle team of the care home as part of the usual activity schedule. The observations enabled us to understand how the Tovertafel was used in practice, what the staff did to make each session successful, and how the residents responded to the activity.

Third, the staff leading the Tovertafel activity were asked to complete a diary after each session. The diaries invited further reflection on the use of the Tovertafel, including the views of the staff about the impact of the Tovertafel on the mood of residents, what worked well, and what the staff found challenging about the session (refer to Multimedia Appendix 1 for the diary design). The staff were asked to complete diary entries after each Tovertafel session to capture their experiences as close to the time of delivery as possible.

Table 2 lists the observed Tovertafel sessions, the number of people involved, and the timing of diary entries completed by the staff members.

XSL•FO
RenderX

Table 2. List of the Tovertafel sessions examined.

ID	Date and time	Location	Number of people involved	Observed by researchers	Diary entry completed by staff
1	June 6, 2023, Tuesday, 9:15 AM-10 AM	The memory support unit	2 residents, 2 lifestyle staff, and 1 nursing staff	✓	
2	June 7, 2023, Wednesday, 9:15 AM-10 AM	The memory support unit	4 residents and 2 lifestyle staff	✓	✓
3	June 12, 2023, Monday, 9:15 AM-10 AM	The memory support unit	7 residents, 2 lifestyle staff, and 1 nursing staff		✓
4	June 21, 2023, Wednesday, 1:30 PM-2:15 PM	The nursing home	4 residents and 1 lifestyle staff		✓
5	June 27, 2023, Tuesday, 9:15 AM-10 AM	The nursing home	3 residents and 2 lifestyle staff	✓	✓
6	June 28, 2023, Wednesday, 9:15 AM-10 AM	The nursing home	6 residents and 2 lifestyle staff	✓	✓

Study Site

Our study took place at 1 residential aged care home of a not-for-profit aged care provider in Queensland, Australia, in 2023. The home provides care for approximately 80 residents divided across 3 residential wings: a memory support unit for residents with severe cognitive impairment, a nursing home area for other residents with high care needs, and a hostel area for residents with lower care needs. The home had 112 staff at the time of our study, including, for example, administrative staff, nurses, managers, and housekeeping staff. The home had 2 staff members who oversaw lifestyle activities: 1 diversional therapist and 1 lifestyle activities coordinator. These staff members were supported to run activity sessions by nursing and personal care staff, depending on availability and the number of people required.

The residents at the home predominantly follow a daily routine with fixed times for meals and personal care. Enrichment activities are scheduled around this routine. The lifestyle activities coordinator prepares a monthly activities calendar, which is distributed to residents on paper. Each day has an activity in the morning and another in the afternoon. Activities last between 45 and 60 minutes, and the residents are free to attend as they please. All activities are announced by staff over a public address system transmitted across the 3 residential wings.

The care home has a large room for providing lifestyle activities. At the time of our study, this room was undergoing renovation. Because of this, the equipment for the Tovertafel had been installed in the dining areas of the memory support unit and the nursing home.

During this study, the Tovertafel activity was scheduled in the activities calendar, allowing the residents from the memory support unit and the nursing home to participate (Table 2). The residents from the hostel area were also welcome to participate. The staff had been using the Tovertafel for 2 years, having received the device as a gift from a benefactor in 2021, but it had not been included in the activities schedule for >6 months before this research. The staff had not received formal training on how to use the Tovertafel. Instead, they learned on the job. After the Tovertafel was first installed in the home, the diversional therapist downloaded a PDF of the system's user manual to understand how the device should be operated.

Participant Recruitment

The staff members were invited to participate in the study if they were familiar with the Tovertafel or if they were scheduled to help facilitate the sessions during the study period. All 6 staff members who met either of these criteria agreed to participate. Volunteers and family members were also eligible to participate in the study, but none were present when data were collected.

Procedure

All procedures were carried out by 2 researchers (RMK and AM). Two fieldwork visits, each lasting 4 days, were conducted in June 2023. Both researchers spent several days on-site before and after the Tovertafel sessions, providing opportunities to learn about the social routines of the care home [29]. The researchers took photographs of the setting, without capturing images of the residents or the staff, and tested the Tovertafel games to understand what they involved. The dates and times of the observed sessions are provided in Table 2. The researchers created private, written fieldnotes after each day of the study to record their experiences and personal reflections about the care home.

During the first fieldwork visit, the researchers interviewed the staff members about their experiences with the Tovertafel and observed 2 Tovertafel sessions in the memory support unit. The researchers conducted nonintrusive observations of each session, observing the staffs' and the residents' interactions around the Tovertafel and with each other, from 2 to 3 meters away. The researchers took notes about which games were played in each session, what the staff did to facilitate the activity, and how the residents responded to the Tovertafel, including emotional responses and social interactions. All observations of emotional responses were noted in a free-text form, rather than using a predefined behavioral grid. The staff member in charge of each session was requested to fill out the research diary after they had facilitated the activity. All diary entries (5 in total) were completed by the coordinator of the lifestyle activities in the

XSL•FO
RenderX

care home, who attended all the Tovertafel sessions examined during the study.

The first fieldwork visit was followed by a 2-week gap during which the researchers were off-site. Two Tovertafel sessions were scheduled by the staff during this time (Table 2). These sessions were not observed by the researchers, but the lifestyle activities coordinator completed the diary after each one.

During the second fieldwork visit, the researchers conducted observations of 2 Tovertafel sessions in the nursing home and invited the residents to provide feedback through interviews after each session. Follow-up interviews were conducted with the staff who had facilitated the sessions.

Analysis

Data consisted of interview transcripts, observational and fieldwork notes, photographs, and diary entries. Interviews were transcribed using a professional transcription service. Data were analyzed using the 6 phases of the approach to reflexive thematic analysis by Braun and Clarke [30]. Reflexive thematic analysis is interpretivist and acknowledges that the subjectivity of the researcher can be a valued analytical resource, while also recognizing that the analysis is influenced by the positionality of the researcher [31]. This analysis was led by RMK, whose position is that technology has the potential to enrich the lives of people living in residential care but risks becoming an unnecessary burden, particularly if the technology is not well designed for this setting [17,26]. This position informed the analysis by highlighting both the burdens involved in enabling the Tovertafel and the enriching aspects of the work, particularly when it provided the staff with opportunities to enhance the activity for the residents.

Following the 6-phase approach by Braun and Clarke [30], RMK first read through the data before inductively creating semantic and latent codes at the sentence level [31]. RMK collated the codes into potential themes that captured interpretations of the problems the staff had faced, the ways that they facilitated sessions, and the perceived benefits of the Tovertafel. The analysis was discussed several times with AM, who iterated the themes based on her involvement in the data collection. RMK then drafted the themes into an initial version of the paper, which was discussed and reviewed by all authors [32]. This team input supported a critical and balanced perspective on the Tovertafel activity. After receiving feedback from peer reviewers, the themes were revised and reorganized to strengthen the interpretations.

In the following section, we present 3 themes. These are organized chronologically, as the themes we developed align with a before-during-after framing of the activity and reflect the temporal ordering of actions conducted by the staff when facilitating the Tovertafel sessions. The first theme highlights the need for the staff to overcome physical and individual barriers *before* initiating the Tovertafel activity. The second theme reveals the importance of creative care from the staff *during* Tovertafel sessions. The third theme details the benefits and outcomes *arising from* staff-supported facilitation of the Tovertafel. Verbatim quotes are used to illustrate findings, with pseudonyms and roles at the care home used to identify participants. Descriptions of the residents' emotions or reactions to the activity are based on observations by the researchers or entries made in the staff diaries.

Results

Participant Characteristics

All 6 staff members interviewed were women. The interviews lasted between 15 and 38 (mean 30, SD 8.7) minutes. Table 3 lists participants' pseudonyms, roles, and time at the workplace for the staff participants. Each Tovertafel session was led by Corina, the lifestyle activities coordinator, along with at least 1 other staff member.

Table 3. List of interviewees.

Pseudonym	Gender	Role	Time at the workplace	Time in current role	Interviewed during the first fieldwork visit	Interviewed during the second fieldwork visit
Anne	Woman	Care home manager	20 y	4 y	✓	
Brenda	Woman	Diversional therapist	11 y	4 y	✓	✓
Corina	Woman	Lifestyle activities coordinator	4 mo	4 mo	✓	✓
Diane	Woman	Special care unit nurse	16 y	16 y	✓	
Ellen	Woman	Personal care worker	11 y	11 y	✓	
Fiona	Woman	Enrolled nurse	8 y	3 y	✓	
George	Man	Nursing home resident	N/A[a]	N/A		✓
Herbert	Man	Nursing home resident	N/A	N/A		✓
Ingrid, Julie, and Karen	Women	Hostel residents (interviewed together)	N/A	N/A		✓

[a]Not applicable.

XSL•FO
RenderX

In addition, 5 residents (2 men and 3 women) were interviewed after the Tovertafel sessions. These interviews lasted between 10 and 19 (mean 14.3, SD 4.5) minutes. Other residents were approached for the interview but either declined to provide feedback, were unable to participate because of severe communication impairments, or were occupied with personal care activities.

Theme 1: Initiating Tovertafel Sessions Required Attending to Physical and Personal Barriers

Summary

Our analysis showed that the staff needed to overcome multiple barriers before running the Tovertafel activity. These included the lack of a dedicated space for play and the need to address residents' uncertainty about joining the activity. These barriers reflect the physical and personal challenges associated with using group-based digital technology, such as the Tovertafel, within residential care homes. We provide further explanations of physical space and perceptions of technology as subthemes in the following sections.

Physical Space: The Care Home Lacked a Dedicated Space for Play

At the time of our study, the care home lacked a dedicated activities space because of an ongoing renovation in the activities room. All group activities had to take place in other communal areas.

Equipment for the Tovertafel system, including an electrical connection and a ceiling bracket, had been installed in the dining areas of the memory support unit and the nursing home. These areas were typically set up with tables and chairs in preparation for mealtimes. To make each space suitable for the Tovertafel activity, the staff had to reconfigure the furniture to create a larger table suitable for the residents to sit around.

The staff did not feel that this practical work was especially burdensome, although it did require additional effort. As part of the setup, the staff also placed a white cloth over the table for use as a playing surface, after noting that the games were hard for some residents to see without using the tablecloth. This was because the wooden tabletops had not been designed for use with the Tovertafel and tended to absorb the projected images, making a lighter surface necessary. In addition, the dining areas had large windows that let in considerable sunlight, making it harder for residents to see the projections if the tablecloth was not used:

> *The setup doesn't take that long...but one thing I found was that the graphics were a bit too light. So maybe it needs to be a certain colour that it projects onto. Because I noticed that with a few of the games, [the residents] did struggle to see what it was.* [Diane, care nurse]

The residents noted that they were hesitant about attending Tovertafel sessions because of the lack of a dedicated space for play. As noted earlier, the activities room at the care home was undergoing renovation. The residents from the hostel area where the residents had lower care needs—were invited to join the activities that took place in the nursing home and memory care unit. However, they expressed reluctance to enter these areas and felt that that it would be better if the Tovertafel was in a shared activities room:

> *I was wondering if it's possible to have that in the activity room. More people might go then...a lot of people down here don't like going into the nursing home.* [Julie, hostel resident]

In addition to needing to overcome physical space barriers to initiate the Tovertafel sessions, the staff also had to ensure that they were able to reset the space at the end of each session. We observed that there was a risk of conflict if the chairs and tables were not rearranged in the same order as before. This was because specific residents could not be seated together, and the staff needed to place the tables carefully to minimize extra work for the care and culinary staff during mealtimes. This situation also meant that the Tovertafel could not be left available for the residents to use at their convenience. As highlighted in previous work [21], the staff thought that the system might be used more often if residents could use it with visiting family members, without the need for staff assistance.

Perceptions About Technology: Residents Were Often Reluctant to Join Tovertafel Sessions

A second barrier that the staff had to overcome was the reluctance of the residents to join the sessions. The staff found it challenging to encourage the residents to attend the Tovertafel sessions because many were not interested in technology-based activities:

> *I find they're not really interested in technology as a lot of the residents have farming backgrounds. They were used to milking the cows and doing all the farm work.* [Ellen, personal care worker]

> *People tend to go "under cover" when you mention the word technology.* [Corina, lifestyle activities coordinator]

In addition, the residents were often uncertain about what the Tovertafel involved if they had not seen it before. Most of the residents we encountered spoke English as their first language, but the word "Tovertafel" lacks obvious meaning for those who do not speak Dutch. The staff found that describing the Tovertafel as "the table game" or "something you play with your hands" was more effective, though still challenging without demonstrating the activity in person.

Because of this, the staff sought to increase attendance by going to the rooms of the residents before sessions took place to explain what the activity would involve and to assure the residents that it would be enjoyable. The staff also approached the residents who were in the nearby common areas when the activity began and asked them if they would like to give the activity a try. The staff felt this was different from other frequently facilitated activities (eg, choir and concerts), which the residents did not need encouragement to attend.

The issue of attendance was important because the staff felt that the Tovertafel worked best with groups of 4 to 6 players. Having 4 players was seen as too few to make the games enjoyable. This meant that achieving enough players was important for

XSL•FO
RenderX

the activity to be successful. The staff told us that the Tovertafel had not been used for some time before the study because attendance had been poor:

> *We have used it before, but we reached a stage where we weren't getting a lot of interest. So, I've taken it off the activity calendar for a time. We just weren't having the numbers turn up to the activity.* [Brenda, diversional therapist]

The staff were also conscious that running the activity too often with the same people risked creating disinterest. This view was echoed by the residents:

> *I think if we played it every day, it might get boring. Once a week, that's no problem.* [George, nursing home resident]

To initiate the Tovertafel activity, the staff needed to be aware of which residents were available to join, how many were willing to participate, and had to encourage those who were available to attend. Addressing these social dynamics was challenging and time consuming but was essential to ensure that sessions had sufficient participants to make the games worth playing.

Theme 2: The Staff Used Creative Care Strategies to Guide the Tovertafel Sessions

Summary

This theme highlights how the staff creatively enabled the use of the Tovertafel *during* the activity; that is, once the staff had created the conditions for a Tovertafel session by arranging the play area and encouraging the residents to attend, additional work was needed to make the sessions successful. We observed that the staff used creative approaches to help residents learn the activity, enable the continued participation of the residents, and sustain their engagement. Each of these activities required differing forms of creativity, which are described in the subsequent sections.

Selecting Gentle Games to Enable Initial Learning

At the start of each session, we observed that the residents were typically unsure what the Tovertafel was and what it involved. In 1 diary entry, Corina, the activities coordinator, similarly noted that the residents were hesitant at the beginning of a Tovertafel session but became more confident once they understood what to do:

> *The mood was a bit daunting and [residents were] unsure what was going on but once we got started the mood changed and [they] became interested.* [Corina, lifestyle activities coordinator, diary excerpt, session 2]

This highlighted the need for the staff to help the residents learn how to use the Tovertafel. The staff had identified a creative way to achieve this learning. At the start of each session, they chose to leave a simple game running on the tabletop while helping the residents join the activity. The staff typically chose the game Leaves for this situation. This game involves sweeping leaves from the tabletop using arm movements. We observed that playing with the leaves helped the residents to understand what would happen if they touched the images. Changes in the leaves were both noticeable and easy to trigger, helping the residents understand how the Tovertafel reacts to their movements.

In this way, starting each session with a gentle game provided opportunities for social learning, whereby the residents learned about the activity through watching others. In session 2, we observed a resident who repeatedly exclaimed that she could not join in because she "didn't know what it was all about." She eventually began playing after watching the interactions of other residents and after receiving encouragement from staff.

These instances highlight that creative work was needed to make the residents feel comfortable at the outset of each session. They also reveal that the staff were not simply supervising the activity. Rather, they were central to making the residents feel at ease and actively encouraged the participation of the residents during the sessions.

Developing Creative Solutions to Ensure Safe Participation

Creativity was evident in the way that the staff developed solutions to other problems that hindered the participation of the residents. For example, some residents found it challenging to trigger a response from the Tovertafel because of mobility impairment. Many could not stand up without assistance and therefore could not reach across the table, while others had difficulty moving their hands or arms.

Recognizing this, the staff had cut up a "pool noodle" (ie, a cylindrical flotation device made of foam) into pieces, which the residents then held to extend their reach and trigger the Tovertafel more easily:

> *Some [residents] have problems with their arms, or they've hurt their elbow in some way. So having the pool noodle, they don't have to reach. And a lot of residents tire very quickly when they're using the arms.* [Brenda, diversional therapist]

The staff were conscious that some residents might experience motion sickness when using the Tovertafel. In session 2, the staff paid close attention to a resident who was known to experience vertigo in response to moving images. The staff described the importance of attending to visual cues, such as body language and facial expression, while running the game to ensure that the residents remained safe. In this way, the staff exhibited creative care by monitoring for signs of distress during the games and changing to a simpler game if required:

> *You need to go by the visual cues and that's probably the advantage of Tovertafel. It's a visual reaction. And it can depend day to day on that resident, with how they're going to interact in an activity or with the Tovertafel.* [Anne, care home manager]

Sustaining Engagement by Choosing Appropriate Games

The importance of choosing games was mentioned repeatedly by the staff during the interviews. We observed that the staff always took responsibility for selecting games, using a remote control to pick games from the Tovertafel's menu system.

XSL•FO
RenderX

The staff explained that choosing appropriate games was essential to sustain the engagement of the residents. One reason was that the staff felt that there needed to be a good match between the games on offer and the physical or cognitive ability of the residents in attendance:

> *[It] depends on who you've got present in the session, because they've all got varying attention spans.* [Anne, care home manager]

In addition, the games needed to be relevant to the interests of the residents; otherwise, the sessions risked becoming boring. The staff explained how this risk could be mitigated by choosing games with features that appealed to the residents. Some Tovertafel games replicate tasks such as gardening and cooking. These games were perceived as engaging because they are modeled on familiar leisure activities. For instance, the staff often selected the Veggie Patch game because many residents were interested in gardening:

> *The gardening game is one where many of the ladies enjoy gardening, so we will do that and ask what they used to do in their garden. There's one person who got very engaged by that. She loves gardening and that's one way to try and get her to engage in the game because gardening was her hobby.* [Brenda, diversional therapist]

Sometimes the staff chose games that were known to be exciting and avoided games that they thought were boring. This was established through trial and error, as well as through verbal and nonverbal feedback from the residents. The staff also switched between games based on the current state of the residents. For example, if residents appeared to be losing interest in a gentle game, staff would switch to one involving physical activity to help engage them:

> *I find if they play it for too long, they get a bit tired of it and they get a little bit bored. So that's why you've got to change to a different game, just to change the mood. It seems like they like something a bit quicker as well, more so than something that's a little bit slow and repeated.* [Corina, lifestyle activities coordinator]

Finally, the staff had identified that staying with the same game for too long risked creating boredom among the players. Changing Tovertafel games during a session was essential to counter this risk, and the staff typically spent no longer than 3 to 5 minutes on each game. Tovertafel games have various difficulty levels, and some were seen as too simple to be played for more than a few minutes. The staff tried to introduce variety by including different types of games within a session, either by manually choosing games or using the Tovertafel's built-in shuffle function, which moves through games in a randomized order:

> *I don't stay on one game for too long, like maybe five minutes. Or I gauge by watching the people around the table and if they're disengaging or not finding it interesting enough, I'm like, okay, let's switch another game now. You know, to try and give them a bit of variety. And to keep their focus.* [Brenda, diversional therapist]

Using Prompts, Questions, and Storytelling to Make the Games More Engaging

A final creative strategy was the use of different spoken techniques to keep the residents engaged. Sometimes this involved gently prompting the residents. For example, a resident in session 4 was observed to be falling asleep during the game. The staff offered words of encouragement—"Come on George," "Wake up George"—to keep him interested. In session 5, a resident felt that she "wasn't doing it right" and stopped interacting with the game. A staff member provided gentle reassurance, saying "No, you're doing a good job," which encouraged her to become involved again.

Beyond this gentle prompting and encouragement, the staff tried to keep the residents' attention on the activity by creatively responding to elements of the games with questions that encouraged interaction. For example, during the Rhymes game, the staff chose to read each rhyme aloud and then asked the residents questions about them. These questions were not part of the game but were added by the staff to make the activity more engaging. In 1 instance, after seeing a rhyme from the song "Rock-a-bye baby," the staff asked the residents if they knew the rhyme and encouraged them to sing along. During the interviews, the staff explained that this was another technique to maintain the residents' focus:

> *To make it work, I keep communicating with the residents and try not to get them off track. Like, you know, just make them stay involved. Otherwise, they're going to drift off and get sidetracked with the game.* [Corina, lifestyle activities coordinator]

Other games involved solving puzzles to reveal pictures of places or animals. The staff used these as opportunities to invite conversation about the past, asking the residents questions such as, "Have you ever been to New Zealand?" or "Did you used to have chickens in your garden?" These questions prompted the residents to talk about their lives. Staff members noted that specific games provided inspiration for these conversations, leading to positive outcomes:

> *What I try to do is to get them to share their stories, so for example when we play the music box game, I might ask them if they had a music box when they were younger. The game with silver polishing is something they would have done when they were younger. You can ask them about that, ask who they had [over] for dinner and what they used to eat.* [Brenda, diversional therapist]

These behaviors were creative techniques that had been instigated by staff members. They were not a standard part of the Tovertafel games, nor were they part of any training that the staff had received. Rather, they were strategies that staff had found to work well and represented the attempts of staff members to build on the basic elements of the Tovertafel games. This appeared to make the activity more active and engaging, as opposed to one in which the residents passively interacted with the system without understanding what they were doing.

XSL•FO
RenderX

Theme 3: Staff Creativity Fostered Benefits and Outcomes

This theme highlights 3 benefits arising from the staff-supported enablement of the Tovertafel. These included moments of joy and fun from participating in a physical activity, opportunities for socialization with other residents, and reminiscence prompted by elements of the games.

Participating in a Fun Physical Activity

The staff universally viewed the Tovertafel as a positive activity for the residents, and in their view, the benefits outweighed the work involved in setting up and running the technology.

The staff described how keeping the residents engaged was important for enjoyment, otherwise attendees would receive limited value from the activity. Through supporting the residents to use the Tovertafel, we observed that the staff were able to elicit moments of joy and laughter, particularly when playing games that involved considerable physical activity. For example, in 1 diary entry, Brenda noted a positive outcome from the session:

> *Happy, smiling during the Tovertafel, lots of laughter. There was a change, they were feeling unsure but once participating their mood became more alert and wanted to give it a go.* [Brenda, diversional therapist, diary excerpt, session 3]

A related benefit was that the Tovertafel encouraged physical activity through the use of the hands and arms. The staff viewed this as a rare opportunity for the residents to engage in gentle exercise, as illustrated by the following comment:

> *There is a whack-a-mole game, and I like that because we've got people with mobility issues and so it's good exercise for them.* [Fiona, enrolled nurse]

Encouraging Social Interaction Between Residents

The Tovertafel provided an opportunity for the residents to socialize with others in the care home. We observed that, outside of mealtimes or scheduled activities, many of the residents would either be watching television or spending time alone in their room. The residents commented on how they enjoyed the social aspects of the activity:

> *I think it's more of a social occasion, when you play games or whatever you play, you can get to know people that way and everybody has different characters, you know what I mean?* [George, care home resident]

The staff felt that the Tovertafel enabled socialization in a unique way by encouraging residents to interact with people they might not usually engage with:

> *It gets them together, you know, like having a chat. So, there's just that bit more interaction. I think it's way better than the TV. And compared to our trivia games, not everyone gets involved in the trivia. Whereas with [Tovertafel] you can involve most of them. And even if they don't want to play it, at least they might just come and watch. They're still interacting in some way, you know?* [Diane, special care unit nurse]

As Diane noted, some residents did not always want to play with the Tovertafel, preferring to sit and watch. Others attended the sessions but were unable to take part because of physical impairment. The staff believed that the Tovertafel still offered a positive opportunity for these residents to socialize, either by watching the activity or by taking on a supporting role in the games. One staff member explained as follows:

> *The ones who honestly can't play, I think it's a good idea, even if they are there to join in and give them something to do, just remember a score or help out because then they are participating and joining in as well. So that's a good thing. You're not inviting them to sit there and fall asleep. The whole idea of it is to keep them inspired, just to let them know that we're here to have a good time and they're in a social environment.* [Corina, lifestyle activities coordinator]

Fostering Shared Reminiscence

Finally, the Tovertafel games provided opportunities for shared reminiscence. However, this reminiscence only occurred when the staff prompted the residents to share their stories in response to the games. By asking questions based on the game elements, the staff actively tried to prompt memories about the past. This included topics such as whether the residents used to keep animals in their garden, what they used to make for dinner, and what they typically did to celebrate birthdays—all of which were embodied within different games. This was seen as a particularly positive outcome of using the Tovertafel, as explained by 1 staff member:

> *Going back in time is good for them to feel good within themselves because it brings back memories. Like this morning, there were some songs that the residents remembered from their younger years. To me that's a good thing, reminiscing.* [Corina, lifestyle activities coordinator]

Prompting reminiscence enabled the residents to share their stories and learn about one another. In session 2, we observed that this reminiscence led to a protracted conversation in which a resident shared a story about a country property she helped renovate. The resident later retrieved a photo album from her room and began showing pictures of the renovation works to other residents. Interactions like these enabled the staff to engage in "biography work" [33], which involves learning about residents so that staff can address their needs using more culturally responsive practices. In this way, reminiscence appeared to be beneficial for the residents and the staff members attending each session.

Discussion

Principal Findings

Overview

This study aimed to understand how the care staff enable the Tovertafel in residential aged care and how this enablement contributes to benefits for residents. Our findings suggest that

XSL•FO
RenderX

staff facilitation was crucial for the effective use of the Tovertafel. This began with overcoming barriers to facilitating sessions, including the lack of a dedicated space for play and the need to encourage the residents to become involved. The staff then played a central role in supporting the residents through the activity, using creative techniques to guide the residents through the games and sustain engagement, while paying close attention to the safety of the residents. These actions helped residents gain meaningful benefits from the Tovertafel, which included fun physical activity, socializing with others, and opportunities for shared reminiscence.

Staff Enablement of the Tovertafel

Our findings highlight the central role the staff played in enabling the Tovertafel and the importance of creative care strategies in sustaining the participation of the residents. Previous work has argued that creativity can be required when providing enriching technology-mediated activities in care homes [34]. Our findings support this claim by showing how the care staff engaged in creative work during Tovertafel sessions. The Tovertafel provides a range of games, many of which are very simple. Although it may be hoped that the residents can play these games without staff oversight, our findings show that facilitation by the staff was needed to ensure that the residents understood the purpose of the activity, how to play the games, and how to interact with the system. In addition, the staff undertook a variety of creative actions to enhance the engagement of the residents *during* the sessions. The staff were conscious that the residents might choose to disengage from the Tovertafel, given the simplicity and repetitive nature of the games and because some of the residents had limited attention span. Creative actions around the selection and duration of the games helped to mitigate these risks and kept the residents' focus on the activity. These creative actions, in turn, were likely essential to fostering the benefits and outcomes observed in this study.

However, the involvement of the care staff was also needed to instigate the sessions. This work is arguably less desirable than tasks involving creative care. We observed how staff time was needed to set up the physical space for play—partly a result of where the Tovertafel was placed in the care home. The staff and the residents felt that the activity might be more popular if it were available to use in a dedicated space, such as an activities room. However, previous work has shown that when the Tovertafel was set up in such a space, it was still rarely used by the residents without prompting from caregivers [21]. This implies that facilitation may still be required *during* the use of the Tovertafel, irrespective of where it is located. We also found that staff time was needed to encourage the residents to attend the sessions, given residents' lack of experience with technology and uncertainty about what to do. These efforts can be viewed as a burden, but they also provided an opportunity for the staff to engage directly with the residents and understand their reasons for nonparticipation. The staff were then able to factor this into the running of the sessions, such as by starting with a simple game to help residents to become comfortable with the technology.

Finally, the care staff were involved in making sure that the residents remained safe during the activity. In residential aged care, many residents are frail or are living with cognitive impairment. This was true for many of the people who participated in the sessions we observed. The care staff ensured that these residents remained safe while using the Tovertafel. The staff believed that the Tovertafel involved some risks for the residents, particularly those who experience dizziness, and so they were vigilant about resident safety. Previous studies have highlighted the need for staff to ensure the safety of residents when using "risky" technologies such as exergames [35] or immersive VR [15,27]. Our study shows that the involvement of the staff was similarly essential to ensure that the residents remained safe when interacting with the Tovertafel.

Benefits of Using the Tovertafel

Our findings shed light on benefits arising from the staff-supported enablement of the Tovertafel. First, Tovertafel encouraged physical activity and provoked moments of fun and joy for the residents. It can be difficult to provide stimulating physical activities in residential aged care because of high rates of impairment among the residents [36]. Our observations suggest that the Tovertafel promotes gentle upper-body exercise but can require creative problem-solving and support from the staff to make the activities accessible. We observed this when the staff triggered interactions on behalf of the residents and through their efforts to make the game easier to play by using pieces of foam to extend the reach of the residents.

A second benefit was that the Tovertafel provided an opportunity for the residents to socialize with each other and with the care staff. Although care homes such as the one we studied frequently provide a range of lifestyle activities, opportunities for social interaction can remain scarce [37], contributing to experiences of loneliness [38]. To address these issues, previous work has explored the benefits of technologies that encourage "ludic engagement" through playful interactions between the residents [39,40]. The staff in our study felt that the Tovertafel was similarly beneficial because it brought residents together for playful, face-to-face interaction. This may be hard to achieve with technologies that are single-user focused, such as VR [15]. In addition, the staff noted that the Tovertafel demands cognitive engagement and encourages conversation. This was perceived to be more beneficial than passive activities, such as watching television. It was also perceived as beneficial for the residents who participated as "active audience members" [41], observing while others engaged with the Tovertafel.

A third benefit was that the Tovertafel enabled the staff to prompt reminiscence among the residents. When used effectively, reminiscence can be a positive activity for older people living in residential care [42] and can be supported using technology such as VR [43,44] and digital mapping systems [45]. However, these efforts are also typically single-user focused, and it has been found that while care home residents often engage in reminiscence with family and friends, they do so less frequently with other residents and the care staff [42]. This highlights a role for the Tovertafel in enhancing care: by drawing on elements of the games in a social setting, the staff were able to prompt reminiscence by encouraging residents to

XSL•FO
RenderX

share their life stories. Previous research has shown positive outcomes when care staff are involved in reminiscence because they can develop closer relationships with the residents and enhance their care work [46]. We saw this in our study when the staff used the Tovertafel to engage in "biography work," enabling them to learn new things about the residents [33]. Overall, these findings highlight how the Tovertafel can gently prompt reminiscence in ways that benefit both the residents and care staff. They also point to a positive feature of the Tovertafel games, whereby the open-ended design of most games allows for creative facilitation by the staff [34].

However, it is important to note that these benefits were contingent on the significant involvement of the care staff to enable the participation of the residents. On the basis of our findings, it is unlikely that similar outcomes will be achieved if the residents are not supported to use the Tovertafel. This may demand additional staff time in cases where the staff are not already rostered to support enrichment activities. The need for enablement by the staff, coupled with the high cost of the technology itself, should be a key consideration when deciding whether to invest in the Tovertafel. This study suggests that the Tovertafel can be an enriching small-group activity that can meet the needs of individual residents while affording opportunities for creative care work. However, as the staff had discovered at the home we studied, the Tovertafel should be adopted as part of a diverse activities program that includes large- and small-group activities; those that use digital technology; and those that encourage investment in physical infrastructure, such as gardening and interest-based hobbies [10,13].

Organizational Considerations for the Use of the Tovertafel

Our findings highlight several considerations that may inform decisions to implement the Tovertafel within residential aged care.

First, the Tovertafel is a high-cost investment that requires a significant amount of ongoing support to maintain use of the resource. The level of involvement from the staff required to effectively facilitate the Tovertafel is a potential barrier, especially in residential care homes where staff shortage is high, and priority is often given to meeting the personal care needs of the residents [2]. Aligning with previous work [21], our findings emphasize the competencies required by the staff to encourage the residents to use the Tovertafel, especially when those residents have cognitive impairment. This requires a team effort as the Tovertafel appears to work best with support from at least 1 staff member who can facilitate the engagement of the residents and maintain their attention.

Although training staff about the benefits and features of the Tovertafel is important [47], our findings demonstrate that training the staff *how* to creatively facilitate sessions may be even more valuable for maximizing its impact. This type of training is unlikely to be attained via reading manuals or resources and is better acquired through hands-on experience—as was the case with our participants. Training staff on good practices related to the Tovertafel could help with "onboarding" new staff to deliver the activity, given that staff turnover is a known barrier to the successful implementation of technology [16]. Alternatively, training could focus on developing the key competencies of the staff required for facilitation (particularly communication skills), allowing them to facilitate a range of activities [48].

In this study, we observed that facilitation skills were modeled to new staff by experienced staff and mentors. For instance, training typically involved a staff member demonstrating effective strategies while explaining the reasons for each approach aloud. However, this approach is likely to be an obstacle in residential aged care because of high staff turnover and limited staff-family training. Therefore, we recommend that staff receive general communication partner training, which can equip people with the skills required to respond effectively to the needs of the residents and to facilitate communication strategies [49]. Such training focuses on building interpersonal skills and fostering effective interactions, enabling staff to better engage with residents and enhance their facilitation capabilities.

Our study highlighted the need for an appropriate space to enable use of the Tovertafel. In the home we studied, the Tovertafel was situated in 2 specialized care areas: one for people with physical care needs and another for residents with severe cognitive impairment. However, in both locations, the Tovertafel was used in the main dining spaces, where the residents could not always see the projected images clearly. The need to set up the Tovertafel and pack it away meant it was not readily available for spontaneous activity or facilitation by family members and visitors. The residents and the staff felt that the system might be used more often if it were set up in a dedicated space, such as an activities room. This would enable the residents to access the resource with the staff as well as other available facilitators [21].

Third, appropriate scheduling of the Tovertafel needs to be considered to prevent disinterest among residents. According to the staff we interviewed, the Tovertafel had been dormant for some time before this study. This was because the residents had become uninterested in using it, leading to low attendance at the sessions. At the time of this study, staff were reintroducing the Tovertafel as a new activity to elicit interest. However, we observed that the Tovertafel was not well known among the residents compared to other activities. The Tovertafel was challenging to describe and required the staff to promote the benefits of the activity, relying on the residents' prior knowledge and interest in technology. Thus, the staff found it difficult to encourage residents to participate. This is challenging because the Tovertafel is an expensive technology and may be best deployed with larger numbers of residents so that it provides a return on investment.

Studies have reported that facilitators for promoting participation in technology-based activities can include communicating the benefits of use, building knowledge and confidence in using technology, and offering support during its use. In line with these findings, we offer some practical strategies for promoting interest in the Tovertafel: (1) communicating benefits; organizing informative sessions, where staff demonstrate the technology and its benefits, may increase residents' familiarity and engagement; (2) building knowledge and confidence in

XSL•FO
RenderX

using technology; sharing success stories and positive outcomes from previous sessions can serve as motivation for residents to participate; and (3) offering support; a buddy system could be established, pairing residents who are comfortable with technology with those who may be hesitant, fostering a supportive environment for participation. These strategies should be co-designed with residents and staff to ensure they are suitable for this context.

Limitations and Future Work

One limitation of this study is that we were only able to capture data about a handful of Tovertafel sessions, which occurred during a period of 4 weeks with a small number of residents. This may have constrained our observations on the breadth of strategies used by the staff. This limitation was mitigated through the use of diaries that were completed by the staff in additional sessions. Future studies can explore the impact of varying session lengths on the engagement of residents and strategies used by the staff to gain additional insights into effective facilitation techniques. Future research could also monitor the use of Tovertafel over the longer term to ascertain whether sustained engagement by both the staff and the residents occurs. Given the high cost of the device, ongoing use would be required for it to be cost-effective; otherwise, financial resources might be better spent on other beneficial activities and programs with higher engagement over the longer term.

Second, we focused on a single care home in Australia. However, gaining access to residential aged care for research is challenging, and studies are often based on a single site [8,10]. In addition, the Tovertafel is still not in widespread use, making it challenging to identify appropriate study sites to compare findings. Our findings have value in expanding the knowledge base about the Tovertafel and contribute to an improved understanding of the role that staff play in facilitating technologies that are hard for residents to use independently. For example, the benefits of enjoyment and physical activity from the Tovertafel align with a previous study conducted in the United Kingdom [21]. Nevertheless, further research should be conducted to identify context-dependent benefits and evaluate the transferability of our findings.

This study offers limited insight into the views of the residents. Many of the residents observed in our study did not have the cognitive capacity to consent to interviews. The residents who were interviewed had little to say about the Tovertafel, possibly because they were not fully aware of its potential features or because of their brief interaction with it. Future work should aim to provide a deeper understanding of the experiences of the residents. This information could then be used to inform best practices for care staff.

Conclusions

This study investigated the role of the staff in facilitating the effective use of the Tovertafel in 1 Australian residential aged care home. We found that creative strategies were needed to encourage the residents to use the Tovertafel and to support the participation of the residents in the activity. Beneficial outcomes included experiencing fun and joy, social interaction, and reminiscence. The staff proactively mitigated challenges in engagement, which included barriers to access and the adverse reactions that the residents may have to the experience. Aligning with previous research on the use of technology for enrichment in aged care [11,15,17], this study highlights the pivotal role of care staff in the adoption and use of digital technology, meaning that aged care providers need to account for staff involvement when deciding whether to adopt the Tovertafel.

Acknowledgments

This study was funded by the Australian government through a grant from the Australian Research Council awarded to JW (FT170100420) and by the University of Melbourne's Creativity and Wellbeing Research Institute. The authors would like to thank the participants for their willingness to support this study.

Authors' Contributions

JW conceived this study. JW, RMK, and AM designed the study procedures. RMK and AM collected the data. RMK led the data analysis and drafted the manuscript. RMK and AM revised the manuscript in response to reviewers' comments. RO, JAL, and JW edited each version of the manuscript and provided additional feedback.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Design of the feedback diary completed by the staff participants.
[DOCX File , 30 KB - aging_v8i1e67919_app1.docx]

References

1. Older Australians. Australian Institute of Health and Welfare. URL: https://www.aihw.gov.au/reports/older-people/older-australians/contents/aged-care [accessed 2025-10-11]

XSL•FO
RenderX

2. Monro C, Mackenzie L, O'Loughlin K, Low L, Du Toit SH. 'I could no longer cope at home': experiences of clients and families in residential aged care within the context of Australia's aged care reforms. Australas J Ageing 2023 Jun 28;42(2):344-354. [doi: 10.1111/ajag.13159] [Medline: 36437807]
3. Scott S, Raynor A, Dare J, Grieve J, Costello L. Improving the transition of older adults into residential aged care: a scoping review. Clin Gerontol 2024 Nov 06;47(5):746-759 [FREE Full text] [doi: 10.1080/07317115.2023.2274042] [Medline: 37929882]
4. Health promotion glossary of terms 2021. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/9789240038349 [accessed 2025-05-29]
5. Siette J, Dodds L, Surian D, Prgomet M, Dunn A, Westbrook J. Social interactions and quality of life of residents in aged care facilities: a multi-methods study. PLoS One 2022 Aug 29;17(8):e0273412 [FREE Full text] [doi: 10.1371/journal.pone.0273412] [Medline: 36037181]
6. Miller E. Beyond bingo: a phenomenographic exploration of leisure in aged care. J Leis Res 2017 Dec 13;48(1):35-49. [doi: 10.18666/jlr-2016-v48-i1-6254]
7. Tak SH, Kedia S, Tongumpun TM, Hong SH. Activity engagement: perspectives from nursing home residents with dementia. Educ Gerontol 2015 Mar 21;41(3):182-192 [FREE Full text] [doi: 10.1080/03601277.2014.937217] [Medline: 25489122]
8. Thach KS, Lederman R, Waycott J. Key considerations for the design of technology for enrichment in residential aged care: an ethnographic study. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. 2023 Presented at: CHI '23; April 23-28, 2023; Hamburg, Germany p. 1-16 URL: https://dl.acm.org/doi/10.1145/3544548.3581176 [doi: 10.1145/3544548.3581176]
9. Tierney L, Beattie E. Enjoyable, engaging and individualised: a concept analysis of meaningful activity for older adults with dementia. Int J Older People Nurs 2020 Jun 05;15(2):e12306. [doi: 10.1111/opn.12306] [Medline: 32022414]
10. Thomas JE, O'Connell B, Gaskin CJ. Residents' perceptions and experiences of social interaction and participation in leisure activities in residential aged care. Contemp Nurse 2013 Oct 17;45(2):244-254. [doi: 10.5172/conu.2013.45.2.244] [Medline: 24299253]
11. Waycott J, Zhao W, Kelly RM, Robertson E. Technology-mediated enrichment in aged care: survey and interview study. JMIR Aging 2022 Apr 12;5(2):e31162 [FREE Full text] [doi: 10.2196/31162] [Medline: 34975014]
12. Muñoz D, Pedell S, Sterling L. Evaluating engagement in technology-supported social interaction by people living with dementia in residential care. ACM Trans Comput Hum Interact 2022 Oct 20;29(5):1-31. [doi: 10.1145/3514497]
13. Wilding R, Barbosa Neves B, Waycott J, Miller E, Porter T, Johnston J, et al. Introducing virtual reality to older adults: a qualitative analysis of a co-design innovation with care staff. Arch Gerontol Geriatr 2024 Oct;125:105505 [FREE Full text] [doi: 10.1016/j.archger.2024.105505] [Medline: 38851090]
14. Hodge J, Balaam M, Hastings S, Morrissey K. Exploring the design of tailored virtual reality experiences for people with dementia. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI '18; April 21-26, 2018; Montreal, QC p. 1-13 URL: https://dl.acm.org/doi/10.1145/3173574.3174088 [doi: 10.1145/3173574.3174088]
15. Waycott J, Kelly RM, Baker S, Neves BB, Thach KS, Lederman R. The role of staff in facilitating immersive virtual reality for enrichment in aged care: an ethic of care perspective. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 2022 Presented at: CHI '22; April 29-May 5, 2022; New Orleans, LA p. 1-17 URL: https://dl.acm.org/doi/10.1145/3491102.3501956 [doi: 10.1145/3491102.3501956]
16. Hegney DG, Rees CS, Osseiran-Moisson R, Breen L, Eley R, Windsor C, et al. Perceptions of nursing workloads and contributing factors, and their impact on implicit care rationing: A Queensland, Australia study. J Nurs Manag 2019 Mar 17;27(2):371-380. [doi: 10.1111/jonm.12693] [Medline: 30221436]
17. Kelly RM, Xing Y, Baker S, Waycott J. "More than just holding the iPad": family members' perspectives on the work of video calling in aged care. Proc ACM Hum Comput Interact 2024 Nov 08;8(CSCW2):1-26. [doi: 10.1145/3687007]
18. Cavenett W, Baker S, Waycott J, Carrasco R, Robertson E, Vetere F, et al. Deploying new technology in residential aged care: staff members' perspectives. In: Proceedings of the 30th Australian Conference on Computer-Human Interaction. 2018 Presented at: OzCHI '18; December 4-7, 2018; Melbourne, Australia p. 200-204 URL: https://dl.acm.org/doi/10.1145/3292147.3292214 [doi: 10.1145/3292147.3292214]
19. Zhao W, Steven B, Jenny W. Challenges of deploying VR in aged care: a two-phase exploration study. In: Proceedings of the 32nd Australian Conference on Human-Computer Interaction. 2020 Presented at: OzCHI '20; December 2-4, 2020; Sydney, Australia p. 87-98 URL: https://dl.acm.org/doi/abs/10.1145/3441000.3441018 [doi: 10.1145/3441000.3441018]
20. The power of purposeful play. Tover. URL: https://www.tover.care/au/ [accessed 2024-10-22]
21. Good A, Omisade O, Ancient C, Andrikopoulou E. The use of interactive tables in promoting wellbeing in specific user groups. In: Proceedings of the 5th International Conference on Human Aspects of IT for the Aged Population, Held as Part of the 21st HCI International Conference. 2019 Presented at: HCII '19; July 26-31, 2019; Orlando, FL p. 506-519 URL: https://link.springer.com/chapter/10.1007/978-3-030-22015-0_39 [doi: 10.1007/978-3-030-22015-0_39]
22. Talman L, Gustafsson C. Evaluation of Tovertafel UP. Academy of Health, Care and Social Welfare. 2020. URL: https://www.diva-portal.org/smash/get/diva2:1415974/FULLTEXT01.pdf [accessed 2025-05-29]
23. Why is research so important? Tover. URL: https://www.tover.care/au/research [accessed 2024-10-22]

XSL•FO
RenderX

24. Bruil L, Adriaansen MJ, Groothuis JW, Bossema ER. [Quality of life of nursing home residents with dementia before, during and after playing with a magic table]. Tijdschr Gerontol Geriatr 2018 Apr 12;49(2):72-80. [doi: 10.1007/s12439-017-0243-3] [Medline: 29235073]
25. Kelly RM, Cheng Y, McKay D, Wadley G, Buchanan G. “It’s about missing much more than the people”: how students use digital technologies to alleviate homesickness. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. 2021 Presented at: CHI '21; May 8-13, 2021; Yokohama, Japan p. 1-17 URL: https://dl.acm.org/doi/10.1145/3411764.3445362 [doi: 10.1145/3411764.3445362]
26. Kelly RM, Xing Y, Baker S, Waycott J. Video calls as a replacement for family visits during lockdowns in aged care: interview study with family members. JMIR Aging 2023 Jun 12;6:e40953 [FREE Full text] [doi: 10.2196/40953] [Medline: 37191951]
27. Baker S, Waycott J, Robertson E, Carrasco R, Neves BB, Hampson R, et al. Evaluating the use of interactive virtual reality technology with older adults living in residential aged care. Inf Process Manag 2020 May;57(3):102105. [doi: 10.1016/j.ipm.2019.102105]
28. Meyer C, Ogrin R, Golenko X, Cyarto E, Paine K, Walsh W, et al. A codesigned fit-for-purpose implementation framework for aged care. J Eval Clin Pract 2022 Jun 06;28(3):421-435 [FREE Full text] [doi: 10.1111/jep.13660] [Medline: 35129259]
29. Spilsbury K, Haunch K, Hodge G, Pearce S. The opportunities and challenges of conducting observational research in care homes: what a researcher brings, does, and leaves behind? J Long Term Care 2024;63(2):360-370 [FREE Full text] [doi: 10.31389/jltc.280]
30. Braun V, Clarke V. Thematic Analysis: A Practical Guide. Thousand Oaks, CA: Sage Publication; 2022.
31. Terry G, Hayfield N. Essentials of Thematic Analysis. Washington, DC: American Psychological Association; 2021.
32. Braun V, Clarke V. One size fits all? What counts as quality practice in (reflexive) thematic analysis? Qual Res Psychol 2020 Aug 12;18(3):328-352. [doi: 10.1080/14780887.2020.1769238]
33. Müller C, Neufeldt C, Randall D, Wulf V. ICT-development in residential care settings: sensitizing design to the life circumstances of the residents of a care home. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2012 Presented at: CHI '12; May 5-10, 2012; Austin, TX p. 2639-2648 URL: https://dl.acm.org/doi/10.1145/2207676.2208655 [doi: 10.1145/2207676.2208655]
34. Waycott J, Davidson J, Baker F, Yuan S. Supporting creativity in aged care: lessons from group singing, music therapy, and immersive virtual reality programs. In: Proceedings of the 34th Australian Conference on Human-Computer Interaction. 2022 Presented at: OzCHI '22; November 29-December 2, 2022; Canberra, Australia p. 272-282 URL: https://dl.acm.org/doi/abs/10.1145/3572921.3572940 [doi: 10.1145/3572921.3572940]
35. Gerling KM, Mandryk RL, Linehan C. Long-term use of motion-based video games in care home settings. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 2015 Presented at: CHI '15; April 18-23, 2015; Seoul, Republic of Korea p. 1573-1582 URL: https://dl.acm.org/doi/10.1145/2702123.2702125 [doi: 10.1145/2702123.2702125]
36. Dahl Y, Löfström E. Supporting social interaction in care environments: exploring stakeholder perspectives on the potential of interactive technology. Int J Hum Comput Interact 2018 Jan 24;35(1):53-64. [doi: 10.1080/10447318.2018.1426897]
37. Barbosa Neves B, Sanders A, Kokanović R. "It's the worst bloody feeling in the world": experiences of loneliness and social isolation among older people living in care homes. J Aging Stud 2019 Jun;49:74-84. [doi: 10.1016/j.jaging.2019.100785] [Medline: 31229221]
38. Blythe M, Wright P, Bowers J, Boucher A, Jarvis N, Reynolds P, et al. Age and experience: ludic engagement in a residential care setting. In: Proceedings of the 8th ACM Conference on Designing Interactive Systems. 2010 Presented at: DIS '10; August 16-20, 2010; Aarhus, Denmark p. 161-170 URL: https://dl.acm.org/doi/10.1145/1858171.1858200 [doi: 10.1145/1858171.1858200]
39. Gaver W, Boucher A, Bowers J, Blythe M, Jarvis N, Cameron DW, et al. The photostroller: supporting diverse care home residents in engaging with the world. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 Presented at: CHI '11; May 7-12, 2011; Vancouver, BC p. 1757-1766 URL: https://dl.acm.org/doi/10.1145/1978942.1979198 [doi: 10.1145/1978942.1979198]
40. Downs J, Frank V, Wally S. Differentiated participation in social videogaming. In: Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction. 2015 Presented at: OzCHI '15; December 7-10, 2015; Parkville, Australia p. 92-100 URL: https://dl.acm.org/doi/10.1145/2838739.2838777 [doi: 10.1145/2838739.2838777]
41. Henkel LA, Kris A, Birney S, Krauss K. The functions and value of reminiscence for older adults in long-term residential care facilities. Memory 2017 Mar 10;25(3):425-435. [doi: 10.1080/09658211.2016.1182554] [Medline: 27161334]
42. Baker S, Kelly R, Waycott J, Carrasco R, Bell R, Joukhadar Z, et al. School's back: scaffolding reminiscence in social virtual reality with older adults. Proc ACM Hum Comput Interact 2021 Jan 05;4(CSCW3):1-25. [doi: 10.1145/3434176]
43. Porock D, Cain P, Young C, Hilbers J, Bomm M, Ghosh M, et al. The careful project: evaluating the acceptability of local virtual reality experiences as a leisure activity for residents in aged care. Australas J Ageing 2025 Mar 07;44(1):e13379. [doi: 10.1111/ajag.13379] [Medline: 39370985]
44. Webber S, Baker S, Waycott J. Virtual visits: reminiscence in residential aged care with digital mapping technologies. Australas J Ageing 2021 Sep 04;40(3):293-300. [doi: 10.1111/ajag.12902] [Medline: 33538405]

XSL•FO
RenderX

45. Cooney A, O'Shea E, Casey D, Murphy K, Dempsey L, Smyth S, et al. Developing a structured education reminiscence-based programme for staff in long-stay care facilities in Ireland. J Clin Nurs 2013 Jul 14;22(13-14):1977-1987. [doi: 10.1111/j.1365-2702.2012.04342.x] [Medline: 23241022]
46. Perion JJ. Tovertafel: evaluating the benefits of a novel multi-sensory intervention for nursing home residents with dementia [thesis]. The University of Toledo. 2021. URL: http://rave.ohiolink.edu/etdc/view?acc_num=toledo1620142242101885 [accessed 2025-08-04]
47. Bennett M, von Treuer K, McCabe MP, Beattie E, Karantzas G, Mellor D, et al. Resident perceptions of opportunity for communication and contribution to care planning in residential aged care. Int J Older People Nurs 2020 Mar 02;15(1):e12276. [doi: 10.1111/opn.12276] [Medline: 31578823]
48. Burton B, Shrubsole K, Manchha A, King M, Wallace SJ. Communication partner training for aged-care workers: a scoping review. Int J Lang Commun Disord 2025 Feb 20;60(2):e70016. [doi: 10.1111/1460-6984.70016] [Medline: 39977832]
49. Benoit-Dubé L, Jean EK, Aguilar MA, Zuniga A, Bier N, Couture M, et al. What facilitates the acceptance of technology to promote social participation in later life? A systematic review. Disabil Rehabil Assist Technol 2023 Apr 06;18(3):274-284. [doi: 10.1080/17483107.2020.1844320] [Medline: 33156714]

Abbreviations

PICF: participant information sheet and consent form
VR: virtual reality

Edited by J Wang; submitted 25.10.24; peer-reviewed by V Provencher, D Healy; comments to author 14.01.25; revised version received 04.06.25; accepted 03.07.25; published 12.08.25.

Please cite as:
Kelly RM, Manchha A, Waycott J, Ogrin R, Lowthian JA
Staff Enablement of the Tovertafel for Enrichment in Residential Aged Care: Field Study
JMIR Aging 2025;8:e67919
URL: https://aging.jmir.org/2025/1/e67919
doi:10.2196/67919
PMID:40794418

©Ryan M Kelly, Asmita Manchha, Jenny Waycott, Rajna Ogrin, Judy A Lowthian. Originally published in JMIR Aging (https://aging.jmir.org), 12.08.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Detection of Vascular Mild Cognitive Impairment in Southeast Asia Using the Visual Cognitive Assessment Test: Machine Learning Analysis From the BIOCIS (Biomarkers and Cognition Study, Singapore)

Pricilia Tanoto[1*], MPsych; Hannah En Ye[2*], BSc; Seyed Ehsan Saffari[3], PhD; Yi Jin Leow[1], PGDip; Ashwati Vipin[1], PhD; Faith Phemie Hui En Lee[1], BA; Smriti Ghildiyal[1], BA; Shan Yao Liew[1], BSc; Adnan Azam Mohammed[1], BE; Gurveen Kaur Sandhu[1], PhD; Kiirtaara Aravindhan[1], PhD; Gursimar Bhalla[1], BSc; Rasyiqah Binte Shaik Mohamed Salim[1], BEng; Nagaendran Kandiah[1,4,5], MBBS

[1]Dementia Research Centre Singapore, Lee Kong Chian School of Medicine, Singapore, Singapore

[2]School of Psychology, University of Wollongong, Wollongong, Australia

[3]AI+ Neurology Division, Duke-NUS Medical School, Singapore, Singapore

[4]Neuroscience and Mental Health Programme, Lee Kong Chian School of Medicine, Singapore, Singapore

[5]National Healthcare Group, Singapore, Singapore

[*]these authors contributed equally

Corresponding Author:
Nagaendran Kandiah, MBBS
Dementia Research Centre Singapore
Lee Kong Chian School of Medicine
11 Mandalay Road
Singapore, 308232
Singapore
Phone: 65 6513 2164
Email: nagaendran_kandiah@ntu.edu.sg

Abstract

Background: Vascular mild cognitive impairment (VMCI) is a significant global health concern, particularly in Asia. The visual cognitive assessment test (VCAT) has shown promise as a language-neutral screening tool for cognitive impairment.

Objective: This study aims to assess the effectiveness of the VCAT in detecting VMCI and compare its diagnostic performance with the widely used and validated Montreal Cognitive Assessment (MoCA).

Methods: Cross-sectional data from 524 community-dwelling participants were analyzed from the BIOCIS (Biomarkers and Cognition Study, Singapore) and classified into cognitively unimpaired, non-VMCI, and VMCI groups. The participants underwent neuropsychological assessments and 3-T magnetic resonance imaging. The random forest technique and multivariable logistic regression were applied to assess the discriminative properties of the tests.

Results: Participants with VMCI exhibited significantly lower performance across various neuropsychological tests (P<.001) and higher rates of vascular risk factors (P<.001). At a cutoff of 27, the VCAT achieved near-perfect accuracy in discriminating the VMCI group from the cognitively unimpaired group (area under the receiver operating characteristic curve=1; sensitivity=1; specificity=0.991). For differentiating the VMCI group from the non-VMCI group, both the VCAT and the MoCA showed optimal performance at a cutoff of 25 (area under the receiver operating characteristic curve=1.00; sensitivity=1.00; specificity=1.00).

Conclusions: The VCAT could be a valuable tool for detecting VMCI, particularly in diverse, multilingual populations. Its comparable or even superior performance to the MoCA, combined with its language-neutral design, positions the VCAT as a strong addition to cognitive assessment toolkits for VMCI. However, the complex nature of cognitive processing in VMCI suggests that a multifaceted approach that integrates both visual and verbal assessments may ultimately offer the most comprehensive evaluation.

International Registered Report Identifier (IRRID): RR2-10.14283/jpad.2024.89

XSL•FO
RenderX

(JMIR Aging 2025;8:e76847) doi:10.2196/76847

KEYWORDS

vascular mild cognitive impairment; cognition; neuropsychological test; Southeast Asia, machine learning

Introduction

Background

Vascular dementia is the second most prevalent form of dementia globally, including in Singapore and across Asia [1-3]. However, pure vascular dementia is relatively rare and often presents alongside other comorbidities that may worsen cognitive symptoms and accelerate decline [4]. To address this complexity, the term *vascular cognitive impairment* has emerged as an umbrella concept. This term encompasses a wide range of cognitive deficits, attributed partially or wholly to vascular causes, including vascular risk factors and cerebrovascular disease [5]. Within this spectrum, vascular mild cognitive impairment (VMCI) specifically denotes the early, mild stage of cognitive decline due to vascular causes [6]. Early recognition of VMCI is crucial for timely intervention and management.

A key contributor to VMCI is cerebral small vessel disease (CSVD), which frequently coexists with neurodegenerative disorders, such as Alzheimer disease and Parkinson disease. This overlap intensifies cognitive deficits, physical disabilities, and other symptoms of neurodegeneration [7]. Notably, individuals with moderate to severe CSVD tend to experience faster cognitive decline compared to those without such burdens [8]. Furthermore, patients who have experienced a stroke are at heightened risk of developing dementia within months to a year, primarily due to cumulative cerebral vascular damage from recurrent strokes [9].

To date, the gold standard for diagnosing VMCI involves combining neuroimaging techniques to detect CSVD [10,11]. However, the clinical utility of neuroimaging is often constrained by higher costs and limited accessibility, particularly in resource-constrained settings [10,11]. Existing cognitive screening tools, such as the Mini-Mental State Examination and the Montreal Cognitive Assessment (MoCA), are widely validated but were only designed to detect general cognitive changes rather than the specific symptoms of VMCI [12]. In addition, these tests pose challenges in multilingual settings, such as the Asia-Pacific region. Moreover, their language-specific designs usually require multiple translations and validations, potentially introducing cultural and linguistic biases [13,14].

Therefore, the visual cognitive assessment test (VCAT) is emerging as a promising, language-neutral alternative for global cognitive screening. As a visually based global tool, the VCAT eliminates the need for translations, thereby minimizing linguistic biases [15,16]. Although it is designed to be a global cognition test, it potentially offers a comprehensive method for detecting VMCI by assessing key cognitive domains, such as processing speed and executive function. These domains are closely linked to visual skills, as visual information is processed in the brain to guide accurate actions and judgments [17]. In addition, the VCAT also covers a broad spectrum of cognitive abilities, including attention, memory, language, and visuospatial abilities, and social cognition [15]. Previous research indicates that patients with VMCI, compared to healthy controls, show pronounced impairment in processing speed and executive functioning, further highlighting the potential utility of VCAT as a screening tool for VMCI [18]. While previous studies have assessed the validity of VCAT in detecting early dementia [19,20], its specific application and validation in detecting mild cognitive impairment (MCI) related to VMCI have not been extensively explored. This gap in knowledge presents an opportunity to evaluate the effectiveness of the VCAT in a critical area of cognitive health.

Objectives

The objectives of this study are twofold: first, to evaluate the utility of the VCAT in detecting VMCI and second, to determine the optimal cutoff points compared to other global cognitive screening tools for the identification of VMCI.

We hypothesize that the VCAT would be sensitive in detecting cognitive impairment in the setting of VMCI and that it is a reliable screening tool for VMCI. If validated, the VCAT could potentially provide clinicians with a more accessible and culturally flexible tool for early VMCI detection, facilitating timely interventions and improved patient outcomes across diverse populations.

Methods

Study Design and Participants

Participants were recruited from the local community as part of the BIOCIS (Biomarkers and Cognition Study, Singapore) conducted from February 2022 to February 2024. All participants were literate in English or Mandarin and had not experienced any major neurological, systemic, or psychiatric disorders within the past 2 years. Demographic data were collected through self-report questionnaires. Further details on the protocol can be found in the study conducted by Leow et al [21].

Ethical Considerations

This study was approved by the Nanyang Technological University Institutional Review Board (2021-1036). Informed consent was obtained from all participants according to the Declaration of Helsinki and local clinical research regulations, and the procedures used in this study were in accordance with ethical guidelines.

Neuropsychological Tests

Participants underwent a comprehensive set of neuropsychological tests, assessing global cognition (the MoCA and the VCAT), episodic memory (the Wechsler Memory Scale, fourth edition logical story delayed), executive function (the Trail Making Test Part B [TMT-B] and the Color Trails Test 2 [CT2]), visuospatial ability (the Wechsler Adult Intelligence

XSL•FO
RenderX

Scale, fourth edition block design and the Rey-Osterrieth complex figure test copy), processing speed (the symbol digit modalities test), and language ability (semantic fluency [animal] from the MoCA test).

Neuroimaging

Magnetic resonance imaging (MRI) scans were performed using a 3-T Prisma fit scanner (Siemens). MRI visual ratings were conducted on T1-weighted magnetization-prepared rapid gradient echo and T2-weighted fluid-attenuated inversion recovery sequences to quantify the Staals score (0-4). White matter hyperintensities (WMHs) volume was calculated using the *Computational Anatomy Toolbox 12*, an extension of Statistical Parametric Mapping within the MATLAB (MathWorks) environment.

Classification of Participants

Cognitive classification was based on the criteria described by Petersen [22] and the National Institute on Aging–Alzheimer's Association criteria [23]. Participants without subjective cognitive concerns and without impairment in cognitive tests (defined as scores of 1.5 SDs below the mean) were classified as cognitively unimpaired. Participants with subjective cognitive concerns and impaired cognitive test scores but without functional impairment were classified as having MCI. CSVD was determined using the Staals total small vessel disease score of 1 or more [24]. Consequently, participants were grouped into the cognitively unimpaired (CU), non-VMCI (NVMCI), and VMCI groups.

Statistical Analysis

Continuous variables were reported as means and SDs, while categorical variables were presented as frequencies and percentages. Group comparisons of demographic and cognitive scores were conducted using the chi-square test (or the Fisher exact test, where appropriate) and ANOVA (or the Kruskal-Wallis test, depending on the normality assumption) for categorical and continuous variables, respectively. The discriminative properties of the VCAT and the MoCA were assessed using the random forest (RF) technique as part of the machine learning methodology. A total of 100 trees were run to train the RF model, where at most 4 variables were randomly selected as candidates at each split. A 10-fold cross validation (CV) was performed to avoid overfitting and increase the generalization accuracy of the model. The area under the receiver operating characteristic (ROC) curve (AUC) was calculated, along with 95% CI computed using 2000 stratified bootstrap replicates. Sensitivity and specificity metrics were calculated for various threshold points, and the exact binomial distribution approach was used to calculate the corresponding 95% CI. Prespecified pairwise group comparisons were performed and adjusted for clinical confounders, including age, sex, and years of education. *P* values were corrected using the Tukey honestly significant difference post hoc test. Vascular risk factors (eg, hypertension and hyperlipidemia) and their corresponding medications were not adjusted for, as group classification was based on MRI evidence of cerebrovascular disease rather than the presence or absence of such clinical risk factors. This rationale aligns with our imaging-based definition of the VMCI group. All statistical analyses were performed using R (R Foundation for Statistical Computing) [25].

Results

Participant and Cognitive Test Characteristics

A total of 524 participants were enrolled in this study. The cohort had a mean age of 59.05 (SD 11.33) years, with 199 (37.9%) being male, and a mean education duration of 14.45 (SD 3.61) years. Among the participants, 211 (40.3%) were classified as the CU group, 92 (17.6%) as the NVMCI group, and 221 (42.2%) as the VMCI group as shown in Table 1. Table 1 illustrates the demographics and vascular risks of the participants. The VMCI group exhibited significantly higher rates of hypertension (106/221, 47.9%) and diabetes mellitus (39/221, 17.6%) diagnoses compared to the CU and NVMCI groups. Furthermore, the VCMI group demonstrated a larger volume of WMHs. However, there were no discernible differences in BMI status or the incidence of hyperlipidemia across the groups.

XSL•FO
RenderX

Table 1. Demographic and vascular risk characteristics (N=524). Italics indicate significant values.

Characteristic	Overall	CU[a] (n=211)	NVMCI[b] (n=92)	VMCI[c] (n=221)	*P* value[d]
Age (y), mean (SD)	59.05 (11.33)	52.18 (10.33)	58.12 (9.83)	65.91 (8.43)	*<.001*
Education (y), mean (SD)	14.45 (3.61)	15.41 (3.03)	14.60 (3.79)	13.48 (3.81)	*<.001*
Gender: male, n (%)	199 (38)	72 (34.1)	34 (37)	93 (42.1)	.30
BMI, mean (SD)	23.50 (3.34)	23.30 (3.32)	23.57 (3.43)	23.67 (3.33)	.30
Hypertension, n (%)	159 (30.3)	33 (15.6)	20 (21.7)	106 (48)	*<.001*
Diabetes mellitus, n (%)	53 (10.1)	8 (3.8)	6 (6.5)	39 (17.6)	*<.001*
Hyperlipidemia, n (%)	331 (63.2)	125 (59.2)	58 (63)	148 (67)	.20
White matter hyperintensities volume, mean (SD)	2.14 (3.13)	1.04 (0.87)	1.14 (0.66)	3.69 (4.36)	*<.001*

[a]CU: cognitively unimpaired.

[b]NVMCI: nonvascular mild cognitive impairment.

[c]VMCI: vascular mild cognitive impairment.

[d]*P* values were derived from ANOVA or the Kruskal-Wallis test for continuous variables and from the chi-square test or Fisher exact test for categorical variables.

ROC Analysis of VCAT in Detecting VMCI

To evaluate the performance of the VCAT and the MoCA in detecting VMCI, we conducted ROC analyses using RF classifiers with 10-fold CV. As summarized in Table 2 both tools showed comparable diagnostic accuracy at the same cutoff score of 27 or less: MoCA (AUC=0.98; sensitivity=0.97; specificity=0.92) and VCAT (AUC=0.98; sensitivity=0.93; specificity=0.93). While this analysis was not the main focus of this study, it provided supportive evidence for the discriminative utility of both instruments.

ROC analysis using RF and 10-fold CV showed that the VCAT outperformed the MoCA in distinguishing VMCI from CU, as presented in Multimedia Appendix 1. At cutoff points of 25 and 26, the sensitivity of VCAT (0.94, 95% CI 0.90-0.96 and 0.92, 95% CI 0.88-0.95, respectively) was comparable to that of MoCA (0.95, 95% CI 0.91-0.97 for both), with overlapping CIs. Notably, at a cutoff of 27, the VCAT achieved perfect accuracy in identifying VMCI (AUC=1; sensitivity=1; specificity=0.99), surpassing the performance of the MoCA (AUC=0.99; sensitivity=0.96; and specificity=0.95) as shown in Table 2. The multivariable logistic regression yielded a comparable trend, further validating our findings.

We performed an ROC analysis to distinguish between NVMCI and VMCI (Table 3). The RF results showed that the VCAT outperformed the MoCA at all 3 cutoff points in detecting VMCI from the NVMCI. This was further supported by a 10-fold CV, where the VCAT (AUC=0.99) slightly exceeded the performance of the MoCA (AUC=0.99) at cutoff points of 26 and 27, respectively. At a cutoff point of 25, both the VCAT and the MoCA achieved perfect accuracy (AUC=1.00; sensitivity=1.00; specificity=1.00) in distinguishing NVMCI from VMCI, indicating that 25 was the optimal cutoff point for differentiating between the 2 groups as shown in Table 3.

XSL•FO
RenderX

Table 2. Discriminative properties of Montreal Cognitive Assessment (MoCA) and visual cognitive assessment test (VCAT) in detecting vascular mild cognitive impairment (VMCI) from cognitively unimpaired (CU).

Test and cutoff point	Analysis	Training			10-fold CV[a]		
		AUC[b] (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
MoCA							
<25							
	RF[c]	0.90 (0.87-0.93)	0.86 (0.81-0.91)	0.77 (0.71-0.83)	0.99 (0.99-0.99)	0.95 (0.91-0.97)	0.97 (0.93-0.98)
	Multivariable logistic regression	0.87 (0.83-0.90)	0.84 (0.78-0.88)	0.75 (0.69-0.81)	0.87 (0.83-0.90)	0.84 (0.78-0.88)	0.75 (0.69-0.81)
<26							
	RF	0.90 (0.88-0.93)	0.83 (0.77-0.87)	0.80 (0.74-0.85)	0.99 (0.98-0.99)	0.95 (0.91-0.97)	0.95 (0.91-0.97)
	Multivariable logistic regression	0.86 (0.83-0.89)	0.84 (0.792-0.89)	0.74 (0.68-0.80)	0.86 (0.83-0.89)	0.84 (0.79-0.89)	0.74 (0.68-0.80)
<27							
	RF	0.90 (0.88-0.93)	0.85 (0.79-0.89)	0.79 (0.73-0.84)	0.99 (0.99-0.99)	0.96 (0.93-0.98)	0.95 (0.92-0.98)
	Multivariable logistic regression	0.86 (0.83-0.90)	0.76 (0.70-0.81)	0.84 (0.78-0.89)	0.86 (0.83-0.90)	0.76 (0.70-0.81)	0.84 (0.78-0.89)
VCAT							
<25							
	RF	0.91 (0.88-0.93)	0.86 (0.81-0.90)	0.79 (0.73-0.84)	0.99 (0.98-0.99)	0.94 (0.90-0.96)	0.95 (0.92-0.98)
	Multivariable logistic regression	0.86 (0.83-0.89)	0.82 (0.76-0.87)	0.80 (0.74-0.85)	0.86 (0.83-0.89)	0.82 (0.76-0.87)	0.80 (0.74-0.85)
<26							
	RF	0.90 (0.88-0.93)	0.87 (0.82-0.91)	0.78 (0.72-0.84)	0.99 (0.98-0.99)	0.92 (0.88-0.95)	0.98 (0.95-0.99)
	Multivariable logistic regression	0.86 (0.83-0.89)	0.84 (0.79-0.89)	0.77 (0.71-0.83)	0.86 (0.83-0.89)	0.84 (0.79-0.89)	0.84 (0.79-0.89)
<27							
	RF	0.90 (0.87-0.93)	0.83 (0.78-0.88)	0.80 (0.74-0.85)	1.00 (1.00-1)	1.00 (0.98-1.00)	0.99 (0.96-0.99)
	Multivariable logistic regression	0.86 (0.83-0.89)	0.86 (0.80-0.90)	0.75 (0.69-0.81)	0.86 (0.83-0.89)	0.86 (0.80-0.90)	0.75 (0.69-0.81)

[a]CV: cross validation.

[b]AUC: area under the curve.

[c]RF: random forest.

XSL•FO
RenderX

Table 3. Discriminative properties of Montreal Cognitive Assessment (MoCA) and visual cognitive assessment test (VCAT) in detecting vascular mild cognitive impairment (VMCI) from nonvascular mild cognitive impairment (NVMCI).

Test and cutoff point	Analysis	Training			10-fold CV[a]		
		AUC[b] (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
MoCA							
>25							
	RF[c]	0.85 (0.80-0.89)	0.85 (0.79-0.89)	0.69 (0.58-0.78)	1.00 (1.00-1.00)	1.00 (0.98-1.00)	1.00 (0.95-1.00)
	Multivariable logistic regression	0.74 (0.68-0.80)	0.79 (0.72-0.84)	0.60 (0.42-0.70)	0.74 (0.68-0.80)	0.70 (0.72-0.84)	0.60 (0.49-0.70)
>26							
	RF	0.84 (0.79-0.88)	0.75 (0.68-0.80)	0.76 (0.65-0.84)	0.99 (0.99-0.99)	0.94 (0.90-0.94)	0.97 (0.92-0.99)
	Multivariable logistic regression	0.73 (0.67-0.80)	0.82 (0.76-0.87)	0.56 (0.45-0.67)	0.73 (0.67-0.80)	0.82 (0.76-0.87)	0.56 (0.48-0.67)
>27							
	RF	0.85 (0.81-0.90)	0.72 (0.6-0.78)	0.81 (0.72-0.89)	0.99 (0.98-0.99)	0.96 (0.92-0.98)	0.96 (0.90-0.93)
	Multivariable logistic regression	0.74 (0.68-0.80)	0.57 (0.51-0.64)	0.79 (0.69-0.87)	0.74 (0.68-0.80)	0.57 (0.51-0.64)	0.79 (0.69-0.87)
VCAT							
>25							
	RF	0.85 (0.80-0.89)	0.85 (0.80-0.89)	0.70 (0.59-0.79)	1.00 (1.00-1.00)	1.00 (0.98-1.00)	1.00 (0.95-1.00)
	Multivariable logistic regression	0.73 (0.67-0.80)	0.82 (0.77-0.87)	0.56 (0.45-0.67)	0.73 (0.67-0.80)	0.82 (0.77-0.87)	0.56 (0.45-0.67)
>26							
	RF	0.85 (0.80-0.90)	0.72 (0.65-0.78)	0.79 (0.69-0.87	0.99 (0.99-1.00)	0.96 (0.93-0.98)	0.98 (0.93-1.00)
	Multivariable logistic regression	0.73 (0.67-0.79)	0.83 (0.77-0.87)	0.54 (0.46-0.65)	0.73 (0.67-0.79)	0.83 (0.77-0.87)	0.54 (0.43-0.65)
>27							
	RF	0.82 (0.80-0.89)	0.78 (0.72-0.88)	0.72 (0.62-0.81)	0.99 (0.99-1.00)	0.97 (0.94-0.92)	0.98 (0.93-1.00)
	Multivariable logistic regression	0.73 (0.66-0.79)	0.79 (0.72-0.84)	0.60 (0.49-0.70)	0.73 (0.66-0.79)	0.70 (0.72-0.82)	0.60 (0.42-0.70)

[a]CV: cross validation.

[b]AUC: area under the curve.

[c]RF: random forest.

Cognitive Performance

The analysis of covariance (ANCOVA revealed significant mean differences among the 3 groups based on neuropsychological tests, with the VMCI group performing lower than the CU and NVMCI groups after controlling for age, years of education, and sex (Table 4). However, this pattern was not consistent across all tests, particularly for the Wechsler Adult Intelligence Scale, fourth edition block design and the symbol digit modalities test assessments as in Table 4.

Further examination through pairwise comparisons using the Tukey honestly significant difference post hoc test indicated significant differences between NVMCI and VMCI groups in the VCAT, the MoCA, the CT2, the TMT-B, and semantic fluency (animals; $P<.05$). This suggested that the VCAT may be particularly effective in distinguishing between NVMCI and VMCI when compared to other global cognition tests, such as the MoCA, and domain-specific assessments, such as the CT2, the TMT-B, and semantic fluency.

XSL•FO
RenderX

Table 4. Mean differences of neuropsychological tests across 3 groups (N=524). Italics indicate significant values (*P*<.05).

Characteristic	Overall, mean (SD)	CU[a] (n=211), mean (SD)	NVMCI[b] (n=92), mean (SD)	VMCI[c] (n=221), mean (SD)	*P* value[d]	NVMCI vs VMCI[d]	CU vs VMCI[d]
VCAT[e]	26.79 (2.67)	27.83 (2.05)	26.98 (2.33)	25.71 (2.91)	*<.001*	*<.001*	*<.001*
MoCA[g]	25.73 (2.72)	26.74 (2.46)	25.59 (2.51)	24.82 (2.72)	*<.001*	*.04*	*<.001*
Logical story delayed	18.70 (3.97)	20.59 (2.66)	17.65 (4.66)	17.31 (3.99)	*<.001*	.74	*<.001*
Color Trails Test 2	94.31 (38.40)	78.86 (28.08)	93.59 (30.14)	109.41 (43.83)	*<.001*	*.001*	*<.001*
Trail Making Test Part B	75.76 (46.69)	54.31 (16.53)	77.57 (34.29)	96.26 (60.24)	*<.001*	*.001*	*<.001*
Block design	40.93 (11.18)	44.73 (11.39)	40.39 (10.75)	37.33 (9.89)	.06	.06	*<.001*
Rey-Osterrieth complex figure test	32.21 (3.64)	33.84 (1.79)	30.72 (4.12)	31.22 (4.15)	*<.001*	.46	*<.001*
Semantic fluency	18.53 (4.94)	20.52 (4.69)	18.32 (5.11)	16.72 (4.37)	*<.001*	*.02*	*<.001*
Symbol digit modalities test	70.97 (31.80)	77.52 (19.24)	72.33 (14.97)	64.13 (43.24)	.14	.08	*<.001*

[a]CU: cognitively unimpaired.

[b]NVMCI: nonvascular mild cognitive impairment.

[c]VMCI: vascular mild cognitive impairment.

[d]Adjusted for age, sex, and years of education.

[e]VCAT: visual cognitive assessment test.

[g]MoCA: Montreal Cognitive Assessment.

Discussion

Principal Findings

This study aimed to evaluate the utility of the VCAT in detecting VMCI, benchmarking its performance against the widely used MoCA. By comparing VMCI not only with healthy controls (CU group) but also with another form of MCI (ie, NVMCI), this study provides robust evidence of the diagnostic utility of the VCAT. The results demonstrate the capability of the VCAT to accurately distinguish VMCI from both healthy individuals and other MCI subtypes, reinforcing its specificity and clinical relevance. These findings support our hypothesis and underscore the potential of the VCAT as a targeted diagnostic tool for VMCI, offering significant implications for clinical practice, particularly in guiding early interventions and paving the way for future research in cognitive impairment diagnostics.

VCAT Performance and Clinical Implications

Our results demonstrate that VCAT performed comparatively better than the MoCA in detecting VMCI, particularly at higher cutoff points.

Although not the primary focus of this study, our supplementary analysis provides additional insight into the performance of the VCAT and the MoCA in detecting NVMCI. Both tools demonstrated high diagnostic accuracy at a cutoff score of 27 or less, with identical AUC values and comparable sensitivity and specificity. These findings suggest that the VCAT performs on par with the well-established MoCA, consistent with previous validation studies [26,27]. This further supports the diagnostic utility of the VCAT as a viable alternative cognitive screening tool, particularly in contexts where language or cultural factors may limit the effectiveness of the MoCA.

With an optimal cutoff of 27, the VCAT showed excellent discriminative properties in distinguishing individuals with VMCI from individuals without vascular cognitive impairment (ie, the CU group). The high sensitivity and specificity found in our study suggest that the VCAT could be a valuable tool for the early detection of VMCI, potentially allowing for earlier intervention and management strategies. Notably, previous studies have identified similar cutoff scores (approximately 27) for detecting MCI in general older Southeast Asian populations [19]. Our findings extend this evidence by demonstrating that the same cutoff may also be effective in detecting vascular cognitive impairment within a comparable cohort. This supports the utility of the VCAT in vascular-specific contexts and underscores the need for continued refinement of its thresholds across different MCI subtypes.

Furthermore, the VCAT demonstrated promise in differentiating between the vascular etiologies in populations with MCI groups (ie, NVMCI and VMCI). With a cutoff score of 25, the VCAT showed comparable performance to the MoCA in this differentiation. While neither the MoCA nor the VCAT was specifically designed to detect VMCI, these findings suggest that both, especially the VCAT, can capture broader cognitive impairments that may be present regardless of etiology. This potential to distinguish between vascular and nonvascular cognitive impairment in participants in our cohort may help guide clinicians in tailoring interventions and follow-up care more effectively.

In our study analysis, we used RF, as it offers several key benefits that enhance its effectiveness. This method is a useful tool in clinical research [28,29], as it aggregates the predictions of multiple decision trees and provides high accuracy and robustness, making it particularly valuable for distinguishing

XSL•FO
RenderX

between different cognitive states, such as VMCI, CU, and NVMCI. In addition, RF is less prone to overfitting compared to single decision trees, and it does not require extensive feature scaling, simplifying the data preprocessing steps [30]. Its robustness to noise and ability to handle missing data further contribute to reliable and insightful analysis, making RF a powerful tool for assessing and validating the discriminative properties of the VCAT.

The visual nature of the VCAT offers both advantages and potential limitations. Its language-independent approach makes it particularly valuable in diverse, multilingual populations where verbal tests may introduce linguistic biases. However, visual and verbal testing may use distinct pathways for processing semantic memory and implicit learning, emphasizing the importance of assessing both visual and verbal memory in cognitive evaluations, as highlighted in previous research [31].

Despite these limitations, the VCAT remains a valuable screening tool for VMCI. Its language-independent, visual approach addresses a key limitation of many existing cognitive assessment tools, which often require translation and may lose validity across different linguistic and cultural contexts [32].

Cognitive and Vascular Profile of VMCI

Our findings reinforce the distinct cognitive profile associated with VMCI. The VMCI group demonstrated significantly lower performance across various neuropsychological tests compared to the CU and NVMCI groups. This underscores the importance of comprehensive cognitive assessment in suspected VMCI cases, as global cognition tests alone may not fully capture the specific impairments due to different underlying pathophysiological mechanisms.

In our study, the VMCI group showed lower MoCA scores compared to both the CU and NVMCI groups. This may suggest that individuals with VMCI are at a more advanced stage of cognitive impairment. However, it is also possible that the lower scores reflect the specific effects of cerebrovascular disease, which can particularly impact attention and executive function [33]—domains heavily weighted in the MoCA. Therefore, the reduced performance in the VMCI group may be due to both greater overall cognitive decline and the influence of vascular pathology on specific cognitive areas.

Notably, our study revealed significant differences in processing speed and executive functioning between the VMCI and NVMCI groups, consistent with previous studies [34,35]. Furthermore, we observed significant differences between the NVMCI and VMCI groups across other cognitive domains, such as episodic memory, visuospatial ability, and language, as reflected in logical story delayed, block design, and fluency tests. This supports the growing body of evidence suggesting that VMCI is not necessarily confined to a single domain and may present with more subtle or multifaceted cognitive changes [36-38]. These findings highlight the complexity of cognitive deficits in VMCI and point to the potential influence of comorbid pathology.

The higher prevalence of hypertension and diabetes mellitus in the VMCI group, along with larger volumes of WMHs, highlights the strong association between vascular risk factors and WMHs with cognitive impairment. These findings support our hypothesis and align with previous research [39,40]. They also highlight the potential for preventive strategies targeting vascular health in mitigating cognitive decline. This underscores the importance of considering vascular risk factors in the assessment and management of cognitive impairment.

Clinical Implications

This study presents several important clinical implications. First, the VCAT demonstrates strong utility as a screening tool for VMCI, likely capturing broader cognitive impairments—particularly in processing speed and executive function—that are commonly affected in this population. Second, we established specific VCAT cutoff scores: 27 to distinguish VMCI from CU and 25 to differentiate VMCI from NVMCI. These cutoff points offer practical guidance for clinicians, enhancing accuracy and consistency in interpreting results. Third, the distinct cognitive profile observed in patients with VMCI, marked by deficits in processing speed and executive function, highlights that these key domains could be prime targets for early intervention. Addressing these specific deficits may improve patient outcomes and delay disease progression. Finally, the strong association between vascular risk factors and VMCI underscores the importance of aggressive management of vascular risk factors, both in reducing the incidence of VMCI and mitigating further cognitive decline.

Limitations and Future Directions

While our study provides valuable insights, several limitations should be noted. While the VCAT demonstrates comparable or slightly superior performance to the MoCA, particularly at higher cutoff points, its reliance on visual processing may not provide a complete picture of cognitive function.

First, although we applied a comprehensive set of exclusion criteria, we cannot entirely rule out the presence of undiagnosed neurodegenerative conditions, such as Parkinsonism or other cognitive disorders, especially given our community-based recruitment. Subclinical or prodromal cases may not be detectable through standard structural MRI sequences used in this study, which lack the sensitivity to identify early-stage Parkinsonian pathology. Second, the VCAT may have potential bias toward visual processing at the expense of verbal abilities. As highlighted earlier, while this characteristic makes the VCAT valuable in diverse, multilingual populations by minimizing linguistic biases, it may not capture the full spectrum of cognitive deficits in individuals with stronger verbal abilities.

Therefore, future studies should aim to delineate the specific pathways involved in visual and verbal cognitive processing and investigate their relative contributions to VMCI. In addition, while the VCAT has proven to be an effective screening tool, the complex nature of cognitive processing in VMCI suggests that a multifaceted approach, incorporating both visual and verbal assessments, may ultimately offer the most comprehensive evaluation. As our understanding of these cognitive pathways evolves, this may lead to refined screening protocols that leverage the strengths of both visual and verbal testing modalities. The cross-sectional design limits our understanding of the VCAT's predictive value for future

XSL•FO
RenderX

cognitive decline. The sample size, particularly in the NVMCI group, was relatively small, which may have affected the statistical power of some comparisons, particularly when using the RF technique. The imbalanced nature of the datasets could lead to biased predictions favoring the majority, potentially affecting the model's ability to accurately reflect the true predictive power. While we controlled for age, education, and sex, other potential confounding factors such as socioeconomic status or lifestyle factors were not accounted for. Furthermore, this study was conducted in a single geographic location, which may limit the generalizability of our findings. Therefore, future research should be conducted with longitudinal, larger, and more balanced samples that could provide more robust results.

Conclusions

In conclusion, our study demonstrates the potential of the VCAT as a valuable tool for detecting VMCI, particularly in diverse, multilingual populations. Its performance, which is comparable with or superior to established tools such as the MoCA, combined with its language-neutral design, positions the VCAT as a significant addition to the cognitive assessment toolkit. However, the complexity of cognitive processing in VMCI suggests that a multifaceted approach, incorporating both visual and verbal assessments, may offer the most comprehensive evaluation.

Given the growing global burden of vascular cognitive impairment, tools such as the VCAT could be pivotal in early detection and intervention, potentially improving outcomes for individuals affected by these conditions. Future research should prioritize longitudinal validation, cross-cultural applicability, and integration into clinical practice guidelines for cognitive impairment assessment.

Acknowledgments

The authors extend their appreciation to all individuals currently participating as well as those who will participate in research at the Dementia Research Centre Singapore. This project was funded by the Strategic Academic Initiative grant (-) from the Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore; the Singapore Ministry of Education Academic Research Fund tier 3 award ; and the National Medical Research Council, Singapore, under its clinician scientist award ().

Conflicts of Interest

None declared.

Multimedia Appendix 1
Discriminative properties of Montreal Cognitive Assessment and visual cognitive assessment testing in differentiating vascular mild cognitive impairment from cognitive unimpairment.
[DOCX File , 18 KB - aging_v8i1e76847_app1.docx]

References

1. Islam B, Li T, Xu M, Yang D, Lv H, Gassara G, et al. Emerging trends in cognitive impairment and dementia among older populations in Asia: a systematic review. J Glob Health 2024 Nov 08;14:04233 [FREE Full text] [doi: 10.7189/jogh.14.04233] [Medline: 39513289]
2. Mok VC, Cai Y, Markus HS. Vascular cognitive impairment and dementia: mechanisms, treatment, and future directions. Int J Stroke 2024 Oct 16;19(8):838-856 [FREE Full text] [doi: 10.1177/17474930241279888] [Medline: 39283037]
3. Vipin A, Satish V, Saffari SE, Koh W, Lim L, Silva E, et al. Dementia in Southeast Asia: influence of onset-type, education, and cerebrovascular disease. Alzheimers Res Ther 2021 Nov 30;13(1):195. [doi: 10.1186/s13195-021-00936-y] [Medline: 34847922]
4. Xiao X, Xiang S, Xu Q, Li J, Xiao J, Si Y. Comorbidity among inpatients with dementia: a preliminary cross-sectional study in West China. Aging Clin Exp Res 2023 Mar 08;35(3):659-667 [FREE Full text] [doi: 10.1007/s40520-023-02349-3] [Medline: 36754914]
5. Rundek T, Tolea M, Ariko T, Fagerli EA, Camargo CJ. Vascular Cognitive Impairment (VCI). Neurotherapeutics 2022 Jan;19(1):68-88 [FREE Full text] [doi: 10.1007/s13311-021-01170-y] [Medline: 34939171]
6. Meguro K, Dodge HH. Vascular mild cognitive impairment: identifying disease in community-dwelling older adults, reducing risk factors, and providing support. The Osaki-Tajiri and Kurihara projects. J Alzheimers Dis 2019 Mar 18;70(s1):S293-S302. [doi: 10.3233/jad-180899]
7. Arndt P, Pfister M, Perosa V, Mattern H, Bernal J, John A, et al. Risk factors and clinical significance of neurodegenerative co-pathologies in symptomatic cerebral small vessel disease. J Neurol 2025 Apr 18;272(5):349. [doi: 10.1007/s00415-025-13087-z] [Medline: 40251424]
8. Salvadori E, Brambilla M, Maestri G, Nicotra A, Cova I, Pomati S, et al. The clinical profile of cerebral small vessel disease: toward an evidence-based identification of cognitive markers. Alzheimers Dement 2023 Jan 31;19(1):244-260 [FREE Full text] [doi: 10.1002/alz.12650] [Medline: 35362229]

XSL•FO
RenderX

9. Rojas-Saunero LP, Zhou Y, Hayes-Larson E, Wu Y, Mobley T, Nianogo RA, et al. Effect of incident stroke on dementia risk over 10 years in a cohort of Asian American and non-Latino white older adults in California. Neurology 2025 Apr 08;104(7):e213488. [doi: 10.1212/WNL.0000000000213488] [Medline: 40048680]
10. Chávez-Fumagalli MA, Shrivastava P, Aguilar-Pineda JA, Nieto-Montesinos R, Del-Carpio GD, Peralta-Mestas A, et al. Diagnosis of Alzheimer's disease in developed and developing countries: systematic review and meta-analysis of diagnostic test accuracy. J Alzheimers Dis Rep 2020 Dec 29;5(1):15-30. [doi: 10.3233/adr-200263]
11. Sabbagh MN, Boada M, Borson S, Chilukuri M, Dubois B, Ingram J, et al. Early detection of mild cognitive impairment (MCI) in primary care. J Prev Alzheimers Dis 2020 Jun;7(3):165-170. [doi: 10.14283/jpad.2020.21] [Medline: 32463069]
12. Khaw J, Subramaniam P, Abd Aziz NA, Ali Raymond A, Wan Zaidi WA, Ghazali SE. Current update on the clinical utility of MMSE and MoCA for stroke patients in Asia: a systematic review. Int J Environ Res Public Health 2021 Aug 25;18(17):8962 [FREE Full text] [doi: 10.3390/ijerph18178962] [Medline: 34501552]
13. Cova I, Nicotra A, Maestri G, Canevelli M, Pantoni L, Pomati S. Translations and cultural adaptations of the Montreal cognitive assessment: a systematic and qualitative review. Neurol Sci 2022 Jan 09;43(1):113-124. [doi: 10.1007/s10072-021-05716-y] [Medline: 34750686]
14. Khan G, Mirza N, Waheed W. Developing guidelines for the translation and cultural adaptation of the Montreal Cognitive Assessment: scoping review and qualitative synthesis. BJPsych Open 2022 Jan 07;8(1):e21. [doi: 10.1192/bjo.2021.1067] [Medline: 34991771]
15. Kandiah N, Zhang A, Bautista DC, Silva E, Ting SK, Ng A, et al. Early detection of dementia in multilingual populations: visual cognitive assessment test (VCAT). J Neurol Neurosurg Psychiatry 2016 Feb 17;87(2):156-160. [doi: 10.1136/jnnp-2014-309647] [Medline: 25691617]
16. Ng KP, Chiew HJ, Lim L, Rosa-Neto P, Kandiah N, Gauthier S. The influence of language and culture on cognitive assessment tools in the diagnosis of early cognitive impairment and dementia. Expert Rev Neurother 2018 Nov 10;18(11):859-869. [doi: 10.1080/14737175.2018.1532792] [Medline: 30286681]
17. Pergantis P, Bamicha V, Chaidi I, Drigas A. Driving under cognitive control: the impact of executive functions in driving. World Electr Veh J 2024 Oct 16;15(10):474. [doi: 10.3390/wevj15100474]
18. Jokinen H, Laakso HM, Arola A, Paajanen TI, Virkkala J, Särkämö T, et al. Executive functions and processing speed in covert cerebral small vessel disease. Eur J Neurol 2025 Jan 30;32(1):e16533. [doi: 10.1111/ene.16533] [Medline: 39475227]
19. Soo SA, Kumar D, Leow YJ, Koh CL, Saffari SE, Kandiah N. Usefulness of the visual cognitive assessment test in detecting mild cognitive impairment in the community. J Alzheimers Dis 2023 Apr 15;93(2):755-763. [doi: 10.3233/jad-221301]
20. Ng LY, Chin CJ, Danial M, Albart SA, Suppiah PD, Ganasegeran K, et al. Validation of the Visual Cognitive Assessment Test (VCAT) for the early diagnosis of cognitive impairment in multilingual population in Malaysia. Psych 2022 Jan 01;4(1):38-48. [doi: 10.3390/psych4010003]
21. Leow YJ, Wang J, Vipin A, Sandhu GK, Soo S, Kumar D, et al. Biomarkers and Cognition Study, Singapore (BIOCIS): protocol, study design, and preliminary findings. J Prev Alzheimers Dis 2024 Aug;11(4):1093-1105. [doi: 10.14283/jpad.2024.89] [Medline: 39044522]
22. Petersen RC. Mild cognitive impairment as a diagnostic entity. J Intern Med 2004 Sep 20;256(3):183-194 [FREE Full text] [doi: 10.1111/j.1365-2796.2004.01388.x] [Medline: 15324362]
23. Petersen RC, Wiste HJ, Weigand SD, Fields JA, Geda YE, Graff-Radford J, et al. NIA-AA Alzheimer's disease framework: clinical characterization of stages. Ann Neurol 2021 Jun 06;89(6):1145-1156 [FREE Full text] [doi: 10.1002/ana.26071] [Medline: 33772866]
24. Staals J, Makin SD, Doubal FN, Dennis MS, Wardlaw JM. Stroke subtype, vascular risk factors, and total MRI brain small-vessel disease burden. Neurology 2014 Sep 30;83(14):1228-1234 [FREE Full text] [doi: 10.1212/WNL.0000000000000837] [Medline: 25165388]
25. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. URL: https://www.gbif.org/tool/81287/r-a-language-and-environment-for-statistical-computing [accessed 2025-05-29]
26. Liu Y, Xie B, Li Q, Xiao S, Li J, Kandiah N. Visual cognitive assessment test: utility of the brief cognitive battery for early screening of cognitive impairment in Chongqing, China. Brain Behav 2024;14(2):e3413. [doi: 10.1002/brb3.3413/v2/response1]
27. Ng KP, Ong GC, Saffari SE, Bernardes MM, de Carvalho AB, Schilling LP, et al. Visual Cognitive Assessment Test (VCAT): a language‐neutral test to detect mild cognitive impairment and dementia in multinational cohorts. Alzheimers Dement 2025 Jan 03;20(S3):e090328. [doi: 10.1002/alz.090328]
28. Liang L, Liu T, Ollier W, Peng Y, Lu Y, Che C. Identifying new risk associations between chronic physical illness and mental health disorders in China: machine learning approach to a retrospective population analysis. JMIR AI 2025 Jun 30;4:e72599 [FREE Full text] [doi: 10.2196/72599] [Medline: 40611696]
29. Dahu BM, Khan S, Toubal IE, Alshehri M, Martinez-Villar CI, Ogundele OB, et al. Geospatial modeling of deep neural visual features for predicting obesity prevalence in Missouri: quantitative study. JMIR AI 2024 Dec 17;3:e64362 [FREE Full text] [doi: 10.2196/64362] [Medline: 39688897]
30. Chen RC, Dewi C, Huang SW, Caraka RE. Selecting critical features for data classification based on machine learning methods. J Big Data 2020 Jul 23;7(1):52. [doi: 10.1186/S40537-020-00327-4]

XSL•FO
RenderX

31. Oltra-Cucarella J, Sánchez-SanSegundo M, Lipnicki DM, Crawford JD, Lipton RB, Katz MJ, et al. Visual memory tests enhance the identification of amnestic MCI cases at greater risk of Alzheimer's disease. Int Psychogeriatr 2019 Jul;31(7):997-1006. [doi: 10.1017/s104161021800145x]
32. Cruchinho P, López-Franco MD, Capelas ML, Almeida S, Bennett PM, Miranda da Silva M, et al. Translation, cross-cultural adaptation, and validation of measurement instruments: a practical guideline for novice researchers. J Multidiscip Healthc 2024 May;Volume 17:2701-2728. [doi: 10.2147/jmdh.s419714]
33. Siddiquee AT, Hwang YH, Kim S, Shin SJ, Lee JS, Kang J, et al. Middle-age cerebral small vessel disease and cognitive function in later life: a population-based prospective cohort study. Lancet Reg Health West Pac 2025 Feb;55:101284 [FREE Full text] [doi: 10.1016/j.lanwpc.2024.101284] [Medline: 39896231]
34. Lowry E, Puthusseryppady V, Johnen AK, Renoult L, Hornberger M. Cognitive and neuroimaging markers for preclinical vascular cognitive impairment. Cereb Circ Cogn Behav 2021;2:100029. [doi: 10.1016/j.cccb.2021.100029] [Medline: 36324708]
35. Oveisgharan S, Dawe RJ, Yu L, Kapasi A, Arfanakis K, Hachinski V, et al. Frequency and underlying pathology of pure vascular cognitive impairment. JAMA Neurol 2022 Dec 01;79(12):1277-1286. [doi: 10.1001/jamaneurol.2022.3472] [Medline: 36279115]
36. Macoir J. Language impairment in vascular dementia: a clinical review. J Geriatr Psychiatry Neurol 2024 Mar 08;37(2):87-95 [FREE Full text] [doi: 10.1177/08919887231195225] [Medline: 37551643]
37. Seki M, Yoshizawa H, Hosoya M, Kitagawa K. Neuropsychological profile of early cognitive impairment in cerebral small vessel disease. Cerebrovasc Dis 2022 Apr 4;51(5):600-607. [doi: 10.1159/000522438] [Medline: 35378532]
38. Cai LN, Yue J, Cao D, Wang P, Zhang Q, Li A, et al. Structural and functional activities of brain in patients with vascular cognitive impairment: a case-controlled magnetic resonance imaging study. Medicine (Baltimore) 2023 Apr 14;102(15):e33534. [doi: 10.1097/MD.0000000000033534] [Medline: 37058059]
39. Guo W, Shi J. White matter hyperintensities volume and cognition: a meta-analysis. Front Aging Neurosci 2022 Sep 1;14:949763 [FREE Full text] [doi: 10.3389/fnagi.2022.949763] [Medline: 36118701]
40. DeCarli C, Rajan KB, Jin L, Hinman J, Johnson DK, Harvey D, et al. WMH contributions to cognitive impairment: rationale and design of the diverse VCID study. Stroke 2025 Mar;56(3):758-776. [doi: 10.1161/strokeaha.124.045903]

Abbreviations

AUC: area under the receiver operating characteristic curve
BIOCIS: Biomarkers and Cognition Study, Singapore
CSVD: cerebral small vessel disease
CT2: Color Trails Test 2
CV: cross validation
MCI: mild cognitive impairment
MoCA: Montreal Cognitive Assessment
MRI: magnetic resonance imaging
CU: cognitively unimpaired
NVMCI: nonvascular mild cognitive impairment
RF: random forest
ROC: receiver operating characteristic
TMT-B: Trail Making Test Part B
VCAT: visual cognitive assessment test
VMCI: vascular mild cognitive impairment
WMH: white matter hyperintensity

Edited by H LaMonica; submitted 02.05.25; peer-reviewed by SA Soo, KP Ng; comments to author 23.06.25; revised version received 16.07.25; accepted 30.07.25; published 08.10.25.

Please cite as:
Tanoto P, Ye HE, Saffari SE, Leow YJ, Vipin A, Lee FPHE, Ghildiyal S, Liew SY, Mohammed AA, Sandhu GK, Aravindhan K, Bhalla G, Shaik Mohamed Salim RB, Kandiah N
Detection of Vascular Mild Cognitive Impairment in Southeast Asia Using the Visual Cognitive Assessment Test: Machine Learning Analysis From the BIOCIS (Biomarkers and Cognition Study, Singapore)
JMIR Aging 2025;8:e76847
URL: https://aging.jmir.org/2025/1/e76847
doi:10.2196/76847
PMID:

XSL•FO
RenderX

©Pricilia Tanoto, Hannah En Ye, Seyed Ehsan Saffari, Yi Jin Leow, Ashwati Vipin, Faith Phemie Hui En Lee, Smriti Ghildiyal, Shan Yao Liew, Adnan Azam Mohammed, Gurveen Kaur Sandhu, Kiirtaara Aravindhan, Gursimar Bhalla, Rasyiqah Binte Shaik Mohamed Salim, Nagaendran Kandiah. Originally published in JMIR Aging (https://aging.jmir.org), 08.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Using Smart Displays to Implement an eHealth System for Older Adults With Multiple Chronic Conditions: Randomized Controlled Trial

Gina Landucci[1], BS; David H Gustafson Sr[1], PhD; Marie-Louise Mares[2], PhD; Klaren Pe-Romashko[1], MS; John J Curtin[3], PhD; Yaxin Hu[4], MS; Adam Maus[1], MS; Kasey Thompson[1], BS; Sydney Saunders[1], BS; Kaitlyn Brown[1], MS; Judith Woodburn[1], BS; Bilge Mutlu[4], PhD

[1]Center for Health Enhancement Systems Studies, University of Wisconsin–Madison, Madison, WI, United States

[2]Department of Communication Arts, University of Wisconsin–Madison, Madison, WI, United States

[3]Department of Psychology, University of Wisconsin–Madison, Madison, WI, United States

[4]Department of Computer Sciences, University of Wisconsin–Madison, Madison, WI, United States

Corresponding Author:
Gina Landucci, BS
Center for Health Enhancement Systems Studies, University of Wisconsin–Madison
1513 University Ave
Madison, WI, 53706
United States
Phone: 1 608 890 1440
Email: gina.landucci@wisc.edu

Abstract

Background: Smart displays and speakers offer voice interaction, which may be more accessible and appealing to older adults with chronic pain and other multimorbid conditions. Previous trials found stronger socioemotional benefits of ElderTree (vs control) among those with high primary care use and multiple chronic conditions.

Objective: This study aims to test whether older adults with chronic pain and multiple other chronic conditions use and benefit more from ElderTree, an eHealth intervention targeting pain and quality of life, when delivered on a smart display.

Methods: We recruited 269 participants from the University of Wisconsin-Madison health system and community organizations and randomly assigned 1:1:1 to (1) smart display with internet and ElderTree, plus usual care; (2) touchscreen laptop with internet and ElderTree, plus usual care; or (3) usual care alone. Participants were aged ≥60 years, had a chronic pain diagnosis or reported chronic pain, and at least 3 common chronic conditions. Primary outcomes were pain interference and psychosocial quality of life. Data sources were baseline, 4-month, and 8-month surveys and continuous ElderTree usage data.

Results: No significant differences were found between the laptop versus smart display groups for pain interference (b=–0.11, 95% CI –1.07 to 0.85; *P*=.82) or psychosocial quality (b=–0.21, 95% CI –0.96 to 0.55; *P*=.56), nor between the combined laptop+smart display group versus control group for either outcome (pain interference: b=–0.41, 95% CI –1.23 to 0.41; *P*=.33; psychosocial quality of life: b=0.04, 95% CI –0.61 to 0.69; *P*=.90). Mediation was not tested because effects on primary outcomes were nonsignificant. Gender did not moderate the effect of laptop versus smart display groups in pain interference (b=–1.56, 95% CI –3.56 to 0.44; *P*=.13). Gender did moderate the effect of the combined laptop+smart display group versus control group (b=1.91, 95% CI 0.11 to 3.71; *P*=.04). Women showed a significant decrease in pain interference (b=–0.69, 95% CI –1.29 to –0.10; *P*=.02), whereas women in the control group showed no significant change (b=0.25, 95% CI –0.53 to 1.04; *P*=.53). Men in the combined group showed a nonsignificant decrease (b=–0.67, 95% CI –1.47 to 0.14; *P*=.10), whereas men in the control group showed a significant decrease (b=–1.61, 95% CI –2.88 to –0.35; *P*=.01). Participants assigned to the laptop versus smart display used ElderTree more frequently and had more favorable perceptions. Analyses of secondary and exploratory outcomes showed no significant differences between groups.

Conclusions: We found no significant differences between the combined ElderTree group and the control group for changes over time in any primary, secondary, or exploratory outcomes. Moderation analyses indicated that only gender moderated study arm effects, and only for the laptop+smart display versus control group on changes over time in the two primary outcomes.

Trial Registration: ClinicalTrials.gov NCT04798196; https://clinicaltrials.gov/ct2/show/NCT04798196

XSL•FO
RenderX

International Registered Report Identifier (IRRID): RR2-10.2196/37522

(JMIR Aging 2025;8:e75991) doi:10.2196/75991

KEYWORDS

aged; chronic pain; eHealth; geriatrics; health expenditures; multiple chronic conditions; primary care; quality of life; smart displays; smart speakers

Introduction

Background

This study reports the outcomes of a randomized controlled trial examining the uses and effects of an 8-month eHealth intervention for older adults, delivered through two platforms—laptop computer versus smart display—compared with a no-exposure control group. To our knowledge, this is one of the first large trials to compare these platforms for delivering a health intervention. As discussed below, studies of smart device interventions for health outcomes have typically not involved random assignment, comparisons with other platforms, or large sample sizes.

We focused on older adults experiencing chronic pain with multiple other chronic conditions. The intervention was designed to help them understand and manage their pain in the context of complex health challenges. As such, this was an intervention for individuals expected to desire relief but who might face challenges in achieving it. Chronic pain not only affects mobility [1,2], but is also associated with reductions in attention, working memory capacity [3], and executive function [4]. Easy, engaging methods of interacting with an intervention may be particularly important under these circumstances. Given this expectation, the core hypotheses were that the intervention would be more effective when delivered on a smart display than on a laptop, and that both platforms would outperform the control group on the primary outcomes of pain interference and quality of life.

Voice Interactivity and Smart Devices

Like laptops, smart displays are compact, internet-connected computers. A key difference, as reviewed in our protocol paper [5], is the supposed ease of voice interactivity. Users of smart displays can ask questions and give commands by speaking rather than typing and can hear rather than read the device's response. Laptops, on the other hand, are primarily designed for typing and reading. Moreover, chatbot functions, now rapidly evolving with progress in artificial intelligence (AI) and voice recognition, allow the audio interactions of smart displays to, to varying degrees, approximate human conversations. Smart speakers offer the same voice interactivity as smart displays, but the latter have the added advantage of a screen that supports touch navigation of visual menus and video viewing.

Various researchers have proposed that the voice interactivity of smart speakers and smart displays may be particularly accessible and appealing for older adults [6-11], and these claims have been echoed in mass media discussions of voice-interactive devices [12]. Proposed advantages of voice interactivity include compensating for mobility issues that make it difficult to walk across a room to a computer, hand tremors or arthritis that make typing or controlling a cursor challenging, and vision loss that makes reading text or deciphering small images difficult. An additional possibility is that voice interactivity may enhance feelings of companionship, particularly as natural language processing improves the quality of smart devices' responses. Relatedly, voice interactivity may offer easier, more intuitive access to content than computers or mobile phones, which require knowledge of where to tap or click, how to scroll, and how to bookmark content. Because smart speakers and smart displays are used with short vocal prompts, they have been suggested as better suited than computers or even smartphones to support older adults' health and well-being [6,7,13,14].

Researchers have also noted potential difficulties with voice interactivity for older adults [15-17]. Age-related physical changes may reduce the volume, pace, and clarity of older adults' speech, which in turn reduce the accuracy of voice-activated devices in interpreting their commands or questions. Hearing loss may impair the ability to understand the device's responses. Smart systems that require memorizing specific commands or "skills" may not be easy to master. Older adults, like others, may also have concerns about privacy and surveillance, including the possibility of accidentally activating the speech-controlled device.

Current evidence on the benefits and challenges of voice-activated smart systems for older adults is somewhat limited. A 2024 scoping review of studies published between 2010 and 2020 on community-dwelling older adults and "personal voice assistants" identified 22 studies with a total of 284 participants [18]. All 22 studies were exploratory examinations of older adults' use and acceptance of these technologies and none examined effects on quality of life or health-related outcomes. The overall conclusion was that although many older adults reported finding the devices convenient and easier to use than a computer for basic tasks, such as reminders or checking the weather, they sometimes struggled to make themselves understood or to remember the necessary commands. The authors noted the lack of research on effects of use on quality of life, well-being, or functional capacity.

Since then, at least 2 studies have examined socioemotional outcomes of smart device use, although neither compared these devices with text-based devices, such as computers. In a small, randomized trial, 34 older adults were assigned to use either a smart speaker (audio only) or a smart display (audio plus screen) for at least 30 minutes a day for specified activities (eg, listening to music, calling a friend, and playing games) over 12 weeks [19]. Both groups showed significant reductions in loneliness and increases in social support over 12 weeks, with greater changes in the smart display group. In a second, larger but nonrandomized study, 291 older adults living alone were given a smart speaker and training on its use [20]. Participants who

XSL•FO
RenderX

were more (vs less) lonely at baseline used it more often over 2 months. Frequent users showed significant decreases in loneliness but not depressive symptoms, whereas intermittent users showed no significant changes in either outcome. In sum, both studies found that frequent use of smart systems was associated with some positive socioemotional changes, although similar changes may have occurred with laptop use or in a control group.

In addition, at least 2 studies have examined health-related outcomes. In the smaller study, 34 adults with type 2 diabetes (mean age 55 years) were assigned to two smart speaker interventions: (1) "standard of care" daily prompts to maintain a log of insulin and blood glucose, or (2) "conversational AI," in which patients reported their insulin use and blood glucose levels to the smart speaker and received AI-generated dosing instructions based on the data they had just reported [21]. Participants in the conversational AI group achieved optimal dosing in fewer days, were more likely to achieve glycemic improvement and control, and reported greater reductions in diabetes-related emotional distress. The study suggested benefits of AI-generated feedback but did not examine the effectiveness of voice interactivity compared with a text-based system for delivering that feedback.

A larger quasi-experiment [22] examined whether the effects of a smartphone intervention targeting older adults' depression and health behaviors would be greater if participants also received a smart speaker with additional health-related audio functions (eg, medication reminders, exercise programs, and quizzes). All 170 participants were assigned to use the health app and wearable devices to monitor their health data over 6 months; those who agreed to also receive the smart speaker were compared with those who declined, matched on baseline depression scores. There were no differences between the two nonrandomized groups in changes over time in depression or in 5 of the 6 health behaviors assessed. Only dietary diversity increased more among participants who agreed to receive the smart speaker compared with those who declined.

In sum, despite claims that voice interactivity facilitates accessibility and companionship for older adults, few studies have examined the effects of smart displays or smart speakers on health-related outcomes. Those studies have generally not been large, randomized controlled trials and have seldom compared these devices with text-based platforms. The one quasi-experiment in which older adults chose whether to add a smart speaker to a mobile phone health intervention found minimal evidence of health benefits. As noted earlier, our project is one of the first major trials comparing a smart display with a laptop for delivering an intervention to older adults. It focuses on helping participants manage chronic pain in the context of comorbid health conditions.

Chronic Pain and Comorbidity

National data from 2023 indicated that 36% of adults aged 65 years and older in the United States experience chronic pain [23]. Chronic pain (ie, lasting more than 3 months) is often accompanied by other chronic conditions, including anxiety and depression [24-26], particularly in older adults [27]. In one study, over 67% of community-dwelling older adults with 2 or more chronic conditions also had chronic pain [28]. Such multimorbidity combined with chronic pain is associated with greater odds of opioid therapy, excessive polypharmacy and medication-related problems, and higher health care expenditures [29-31]. Chronic pain has complex, bidirectional relationships with sleep, exercise, diet, and depression [32-35], all of which affect the management of other chronic conditions. Given the significance of chronic pain for psychosocial quality of life and for management of comorbid conditions [24], we focused on pain interference and psychosocial quality of life as the primary outcomes of a health intervention targeting older adults with multiple chronic conditions.

Various digital therapeutics have been developed to help individuals manage chronic pain. A 2022 meta-analysis of 36 online cognitive and behavioral interventions for adults with chronic pain found small effects on pain intensity, catastrophizing, and interference (g=0.27-0.31), with stronger effects observed for interventions that included feedback or guidance from a clinician rather than solely self-paced progress through content [36,37]. Although some of these digital interventions included content targeting comorbid depression and anxiety [38,39], most focused only on pain, often specific to particular conditions such as fibromyalgia, arthritis, or spinal cord injury [40-42]. As such, they were not designed to address the broad range of experiences of older adults with chronic pain and other multimorbid chronic conditions.

In contrast, a recent trial specifically targeted adults with chronic pain and multimorbid chronic conditions, testing the effectiveness of usual care compared with a 6-week program on pain science, self-management, and tailored exercise, involving 2 weekly in-person meetings [43]. Adherence was high: the majority of participants attended more than 75% of the sessions, completed workbook assignments, and reported using the self-management strategies (eg, exercises, breathing, relaxation, and sleep plans). At 12 weeks postintervention, the intervention group showed greater improvements in physical function and greater reductions in various pain-related outcomes compared with the control group. A commentary on the study [44] noted the innovation of focusing on comorbid pain and other chronic health conditions, as well as the possibility of delivering such an intervention remotely, rather than in-person, for longer periods of time.

This Study

Overview

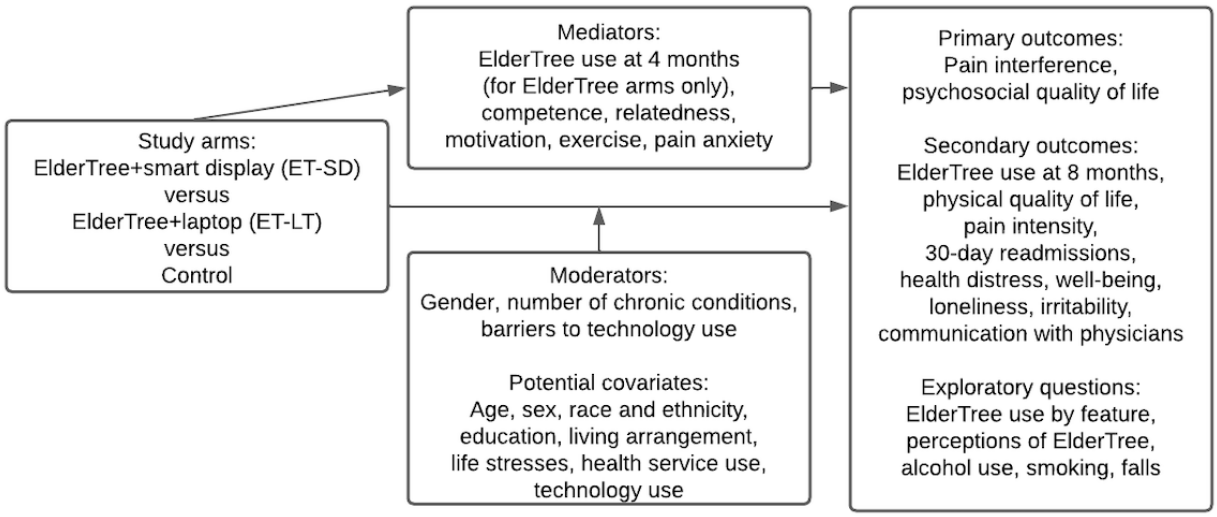
The broad goal was to assess whether delivering a digital eHealth intervention for older adults on a voice-based platform compared with a text-based platform would alter its effectiveness. More specifically, the central question was whether older adults with chronic pain, in the context of multiple comorbid chronic conditions, would show greater use and benefits of an intervention targeting pain and quality of life if they were received it on a smart display rather than a laptop, and whether both the smart display and laptop groups would show benefits compared with a control group that received no device or access to the intervention.

XSL•FO
RenderX

Proposed Main Effects

We hypothesized the greatest improvements over the 8-month intervention period in the primary outcomes in the smart display arm, followed by the laptop arm, and then the control group (Figure 1). We also examined whether there would be differences between study arms on the exploratory outcomes. In this study, we focus on the effects of study arm on engagement and the 2 primary outcomes, although we briefly report findings for the other outcomes and provide details in Multimedia Appendix 1.

Figure 1. Study logic.



Proposed Mediators

The intervention, called ElderTree, is one of a suite of eHealth systems known as Comprehensive Health Enhancement Support Systems [45,46]. These systems, including ElderTree, are based on principles of Self-Determination Theory, as it is intended to meet 3 core psychological needs related to a particular health situation: competence (self-efficacy to achieve health goals), relatedness (feelings of connection with others dealing with the same health challenges), and intrinsic motivation (ie, feelings of autonomy rather than coercion in working toward health goals) [47]. As shown in the logic diagram, we hypothesized that 4-month levels of these 3 constructs would mediate the effects of study arm on the primary outcomes. Given the focus on pain interference, we also examined rates of exercise and levels of pain anxiety as mediators. For comparisons of the 2 ElderTree arms, we proposed that higher use of ElderTree at 4 months in the smart display arm would mediate stronger effects of the smart display compared with the laptop arm on the primary outcomes.

Proposed Moderators

Randomized controlled trials of previous ElderTree platforms, which targeted older adults' quality of life and management of chronic conditions (but not pain specifically), found moderated effects. In the first study, socioemotional benefits of ElderTree compared with control were stronger for older adults with high rates of primary care use [48]. This pattern was interpreted as indicating stronger effects for those with more complex health conditions, such as multiple chronic conditions. In the second randomized trial, targeting older adults with multiple chronic conditions, socioemotional benefits of ElderTree compared with control were stronger for women than for men [49]. In this study, we examined whether effects in either of the 2 ElderTree arms, compared with control, would again be stronger for women than for men and for those with higher numbers of chronic conditions. Additionally, given the focus on platform, we considered whether the hypothesized benefits of a smart display compared with a laptop would vary by participants' reports of their barriers to using text-based versus voice-based devices.

The Historical Context

There are 3 aspects of the timing of this study that merit presentation up front to contextualize the methods and results. First, we selected and began rapid cycle testing on the laptop and smart display devices in 2020. From the options available at the time, we chose the Google Nest Hub Max. Although smart displays have since evolved, the device we used had some challenges with Wi-Fi connectivity, voice recognition, and the dictation program within the ElderTree app. More details are provided in the Results section. Second, the study period overlapped with the COVID-19 pandemic, as discussed in the section on recruitment and training. Third, Google discontinued "Conversational Actions" as of June 2023. This was a key functionality required to access ElderTree. We stopped recruitment 4 months early and ended the intervention early for 8 smart display participants.

Methods

Trial Design

We used a nonblinded, randomized controlled trial design with 1:1:1 allocation to (1) a smart display with internet access and ElderTree, along with usual care; (2) a touchscreen laptop with internet access and ElderTree, along with usual care; or (3) usual

XSL•FO
RenderX

care alone. Assessments were conducted at baseline, 4 months, and 8 months.

Participants

Older adult patients with at least 3 high-risk chronic conditions were recruited from the University of Wisconsin (UW)—Madison Department of Family Medicine and General Internal Medicine system as well as community organizations in the Madison, Milwaukee, and Beloit, Wisconsin, areas. Eligible participants were aged 60 years or older and had chronic pain—indicated by having received a chronic pain diagnosis, by reporting pain on some or most days lasting at least 12 weeks, by reporting pain in the past 3 months on some, most, or all days, and by reporting pain intensity of at least 3 on a scale of 0 (no pain) to 10 (worst imaginable pain) in the past 7 days. They also had at least 3 of the most prevalent chronic conditions among older adults as reported by the Centers for Medicare and Medicaid Services [50]—chronic obstructive pulmonary disease, asthma, diabetes, hyperlipidemia, hypertension, ischemic heart disease, atrial fibrillation, heart failure, stroke, cancer, chronic kidney disease, depression, osteoporosis, and arthritis. We modified the list by adding obesity (BMI≥30) and dizziness/falls/loss of vestibular function. In addition, patients agreed to allow reports about their health tracking to be sent to their primary care provider. Patients were excluded if they required an interpreter or had a medical diagnosis of Alzheimer disease, dementia, schizophrenia or other psychotic disorder, autism spectrum disorder, a known terminal illness with less than 6 months to live, or an acute medical problem requiring immediate hospitalization**.**

Intervention Groups

Patients in all 3 conditions continued with their usual care provided by primary care and internal medicine clinics in the UW–Madison system and were randomized to one of the following arms.

Control

Participants continued with usual care and did not receive a study device, ElderTree, or internet service. Participants were compensated US $30 for each survey completed at baseline, 4 months, and 8 months.

ElderTree Laptop

Participants in the laptop arm were given a Google Chromebook with a touchscreen, access to the ElderTree website, and internet access using a hotspot if needed, for 8 months. ElderTree staff conducted a home visit or phone call (if preferred by participant) during which the participant was trained using the laptop, navigating the ElderTree website, and reminded of study expectations: log into ElderTree regularly and explore its features, complete the weekly surveys, and fill out and return the 2 mailed surveys at 4 and 8 months. Participants were compensated US $10 for each mailed outcome survey completed at baseline, 4 months, and 8 months.

ElderTree Smart Display

Participants in the smart display arm received a Google Nest Hub Max smart device with a 10-inch touchscreen, access to the ElderTree app, and internet access using a hotspot if needed, for 8 months. Training on using the smart display and ElderTree was provided through a home visit or phone call. Study expectations and compensations were the same as those described for the laptop arm.

Laptop

As in earlier iterations of ElderTree, the laptop platform was a members-only website free of advertisements, with design features based on previous users' feedback and best practices for older populations, such as larger fonts, fewer options, and uncluttered screens for better comprehension, navigation, and usability [51,52]. The site included the following: "discussion groups," chatroom-like discussions with other participants; "private messages," an email-like feature for sending and receiving one-to-one messages; "wellness activities," including journal writing, breathing exercises, guided meditation, and exercise videos designed specifically for older adults; "health library," reliable online information on topics such as medication, mental and physical health, and talking with a doctor; "weekly survey," a short questionnaire that asked about general health variables relevant to all chronic conditions, such as sleep, relationship quality, falls, and missed medications; and "living well with chronic pain," a self-paced course presenting the latest science on chronic pain and strategies for managing it.

Smart Display

The smart display platform of ElderTree, developed as a third-party app on the Google Nest Hub Max, was intended to offer the same services as on the laptop platform, although participants navigated by voice and touch. Participants needed to "wake" the device with their voice and say "Hey Google, ElderTree" for the app to open. The device then gave prompts to take the weekly survey or to check out features such as the chronic pain lessons or health library, which included the same video content used on the laptop platform but not the text-based items. These videos were organized as supplemental resources related to each chronic pain lesson. Participants could use voice commands or tap the screen to select features. Because participants sometimes struggled to remember commands, "cheat sheets" with instructions were provided, and a small sticker with the phrase "Hey Google, ElderTree" was placed on the front of the device.

It became clear during development and the course of the study that the smart display dictation function was error-prone, could process only one sentence at a time, and was difficult to correct. Given this, the text-based discussion boards and journaling functions remained available to participants but were hard to use. As described below, a new feature was added to both study arms.

Adding the Weekly Meetup

The importance of maintaining social relatedness services was reinforced by the fact that the study took place during the later phases of the COVID-19 pandemic. To address this, we created a synchronous weekly virtual meetup. To maintain the integrity of the research design, meetups were made available to both laptop and smart display participants; they began in August 2022 and continued through the end of the intervention in

XSL•FO
RenderX

August 2023. Meetups focused on a new health-centered topic each week (eg, managing chronic pain, balance, memory, and eye health), were facilitated by research staff who provided a summary of the literature on the topic, and were followed by an open discussion. Time was also allocated for participants to share personal joys and sorrows. Each session ended with a group movement exercise and was documented and summarized in an email within a week, along with additional related information.

Measures

Participants were assessed for all study variables at baseline, 4 months, and 8 months.

Primary Outcomes

We assessed the 2 primary outcomes using subscales of the PROMIS-29 (Patient-Reported Outcomes Measurement Information System 29-item profile) Global Health measure [53,54], which has been found to be highly sensitive to changes in pain interference and quality of life [55].

The pain interference subscale included 4 items about the past 7 days (eg, “How much did pain interfere with your day-to-day activities?”). Scoring of individual items varies, and is determined by the scale developers, with higher values indicating greater pain interference.

Psychosocial quality of life was measured using the 12-item subscale from the PROMIS-29, which includes three subcategories: (1) ability to participate in social roles and activities (eg, “I have trouble doing all of my regular leisure activities with others”), (2) anxiety (eg, “I felt fearful”), and (3) depression (eg, “I felt worthless”). Items asked participants to report on the past 7 days. Scoring of individual items varies and is determined by the scale developers, with higher values indicating better quality of life.

Mediators

We measured 6 possible mediators. ElderTree use was measured by the number of days of use in the 4-month period before each person’s 4- and 8-month survey completion. Competence was assessed with the 6-item Self-Efficacy for Managing Chronic Disease scale, referring to the past 4 months. Relatedness was assessed with the 5-item McTavish Bonding Scale [56], referring to the past 4 months. Motivation was assessed with 4 items from the Treatment Self-Regulation Questionnaire, referring to the past 4 months [57]. Exercise was measured with a 4-item scale created for this study, referring to the past 4 months. Pain anxiety was assessed with the 13-item Pain Catastrophizing Scale [58], referring to the past 4 months.

Moderators

We investigated whether the effects of study arm on change from baseline to endpoint in primary outcomes were moderated by gender, number of chronic conditions, barriers to technology use, and ElderTree use (for laptop vs smart display only).

Gender was reported by participants at baseline (male, female, and self-describe); all participants identified as either male or female.

The number of chronic conditions was reported by participants, who checked all that applied from a list of the most prevalent chronic conditions among older adults, as reported by the Centers for Medicare and Medicaid Services.

Barriers to assigned technology use were reported by participants. Participants in both study arms indicated whether any of the following factors made it difficult for them to use a computer, iPad, or smart speaker (eg, Alexa or Google Home) on a scale from 0 (never used that technology) to 5 (extremely difficult), with 1 representing “not at all”: vision (even with glasses), hearing (even with a hearing aid), voice issues (difficulty speaking clearly), lack of knowledge on how to use the device, memory, and other (please specify).

Potential Covariates

Potential covariates included age, gender, race and ethnicity, education, living arrangement, life stressors, unplanned health care use (urgent care or emergency department visits), and comfort with technology use. Participants reported sociodemographic variables including gender, race and ethnicity, education, and whether they lived alone or with others. To assess life stressors, they reported on 14 possible stressors from the Social Readjustment Rating Scale (eg, “Death of a very close friend or family member,” “Change in financial status”) [59], checking all that applied (possible range=0-14). Participants reported unplanned health care use in the preceding 4 months to urgent care or the emergency department. Participants rated their comfort with 6 communication technologies—computer, smartphone, tablet, smart speaker, email, and Facebook—on a 6-point scale (total possible range=0-30), with higher scores indicating greater comfort.

Use and Perceptions of ElderTree

To assess system use, time-stamped usage data from ElderTree participants were continuously captured in our database, including specific services used, the date and time the system was accessed, and text entered in discussions and comment threads. Usage data should be interpreted with caution, as quantifying eHealth intervention use is complex and no established method has yet been validated [60].

Use was defined as visiting any service or feature, including the ElderTree home screen. Days of use were calculated as the number of days in which participants accessed ElderTree at least once in a 24-hour period, excluding the training period. Means for this variable included participants with no days of use (scored 0). We also attempted to assess hours of use of ElderTree. Because smart displays did not track time spent on each page, this calculation required estimation and cut-off rules. Calculations and findings for hours of use are presented in Multimedia Appendix 1 to contribute to the emerging discussion on measuring use [60].

Attendance and Participation at Meetups

We counted the number of meetup sessions attended by each participant and, for each session, recorded whether they contributed beyond an initial greeting (ie, raised hand and spoke or typed content into the chat).

XSL•FO
RenderX

Perceptions of ElderTree were assessed in the 8-month survey. Participants rated 9 items on a 1-4 scale (1=not at all to 4=very much so) addressing health benefits (eg, "it helped me learn more about my health issues," "it helped motivate me to take care of my health," Cronbach α=.80), socioemotional benefits (eg, "I enjoyed the other people on ElderTree," "it made me realize I am not alone," Cronbach α=.84), and ease of use (eg, "I found it easy to use," "it was easy to find what I was looking for," Cronbach α=.54). Staff also recorded any inquiries or issues reported by participants regarding ElderTree or either device throughout the trial.

Qualitative Insights

On the 4- and 8-month surveys, participants were asked 3 open-ended questions: "What stopped you or hindered you from using ElderTree or the computer/smart system in general? If you got a computer or smart system from us, what did you like best about getting the computer or smart system in this study? Any other comments or suggestions for us?" We also logged comments about ElderTree on the discussion boards or during phone calls to staff for help with the devices. Comments were unitized (broken into constituent parts) and inductively coded into core themes by the second author. An independent coder categorized a random sample of 25% of unitized comments into the themes (Krippendorff α=.90).

Sample Size Determination and Power

A 3-point difference (d=0.30) in pain interference is considered clinically meaningful [61]. Given this, the study was powered to detect d=0.30 between the laptop and smart display groups, and between the laptop and control groups. We ran 10,000 linear mixed model simulations, which indicated that, for a power of >80% to detect a study arm×time interaction of this magnitude, a postattrition sample of 255 participants would be required. Given relatively low rates of attrition and missing data in earlier trials of ElderTree, we recruited a total of 269 participants.

Recruitment

The UW-Madison Clinical and Health Informatics Institute queried UW Health clinic electronic health records to identify patients who met eligibility criteria. Potential participants received an opt-in letter and consent form from the UW-Madison's Clinical Trials Institute. The letter described the study and included a postage-paid return invitation for further contact from the study team.

To supplement UW Health recruitment and diversify the patient population, we worked with Access Community Health Centers in Madison, Wisconsin, and community collaborators with strong ties to local African American health care and patient communities. Staff at Access Community Health Centers identified potentially eligible patients and distributed the recruitment flyer, consent form, and return invitation. Community collaborators promoted the study opportunity within Black churches, community centers, fraternity and sorority health services programs, and other organizations in the Madison and Beloit, Wisconsin, areas.

The recruitment period (July 2021-December 2022) overlapped with the COVID-19 pandemic. Because the primary mode of recruitment was opt-in letters and participants could choose no-contact phone training rather than in person training, accrual proceeded on schedule. Supplemental community recruitment efforts focused mainly on flyer distribution, email newsletters, and other contact-free strategies.

Randomization and Training

The project manager used a computer-generated allocation sequence to randomize patients on a 1:1:1 ratio to ElderTree Smart Display (ET-SD), ElderTree Laptop (ET-LT), or control group, stratified by gender (male, female), site (UW Health, Access Community Health Centers, community), and number of chronic conditions (≤5 or ≥6). Block size was 12. When baseline assessment and consent were complete, research staff opened the sealed opaque envelope revealing group assignment and conducted equipment setup and training for the ET-SD and ET-LT arms, including all services and instructions for use). Training was conducted at the participants' homes or a public location (eg, public library), unless participants preferred no-contact training due to COVID-19 or other reasons, in which case equipment was mailed and training occurred by phone. No training was provided for the control arm. Once assigned, participants could not be blinded to condition, as those in the ET-SD and ET-LT arms were asked to use ElderTree while control participants were not. The training researcher also could not be blinded to condition after assignment was revealed.

Statistical Methods

Following randomization, the 3 groups were assessed for comparability across all prespecified stratification variables. No statistically significant differences were found between arms for gender, site, or number of chronic conditions.

We selected covariates by examining whether they were moderately correlated (r≥0.30, averaged across the 3 time points) with at least one of the primary outcomes [62]. Race and ethnicity were dichotomized (non-Hispanic White vs all other groups) and were unrelated to either outcome. Only life stressors met the correlation threshold (pain interference r=0.35; psychosocial PROMIS-29 r=−0.46) and was retained as a covariate.

Normality, linearity, and homogeneity of variance for outcome data were assessed using descriptive statistics and graphical representations. Outcomes were analyzed with mixed-effects models using the "glmer()" function from the *lme4* package in R (R Foundation for Statistical Computing. These models account for correlated measurements within participants, incorporated all available data (allowing for intention-to-treat rather than complete-case analysis), and provide unbiased estimates when data were missing at random. Each model included a random effect for participant and fixed effects for survey time point, study arm, and arm-by–time point interaction. Survey time point was entered as a continuous variable with 3 values (baseline, 4 months, and 8 months).

Ethical Considerations

This study was approved by the UW Health Sciences and Minimal Risk Research Institutional Review Board (reference number 2020-0868) and registered at ClinicalTrials.gov

XSL•FO
RenderX

(NCT04798196). Verbal informed consent was obtained by research staff during the initial recruitment call. A unique "study ID" was assigned at enrollment and associated with all data for the duration of the study to protect participant privacy and confidentiality. Participants received compensation (US $30) for completing surveys at baseline, 4 months, and 8 months.

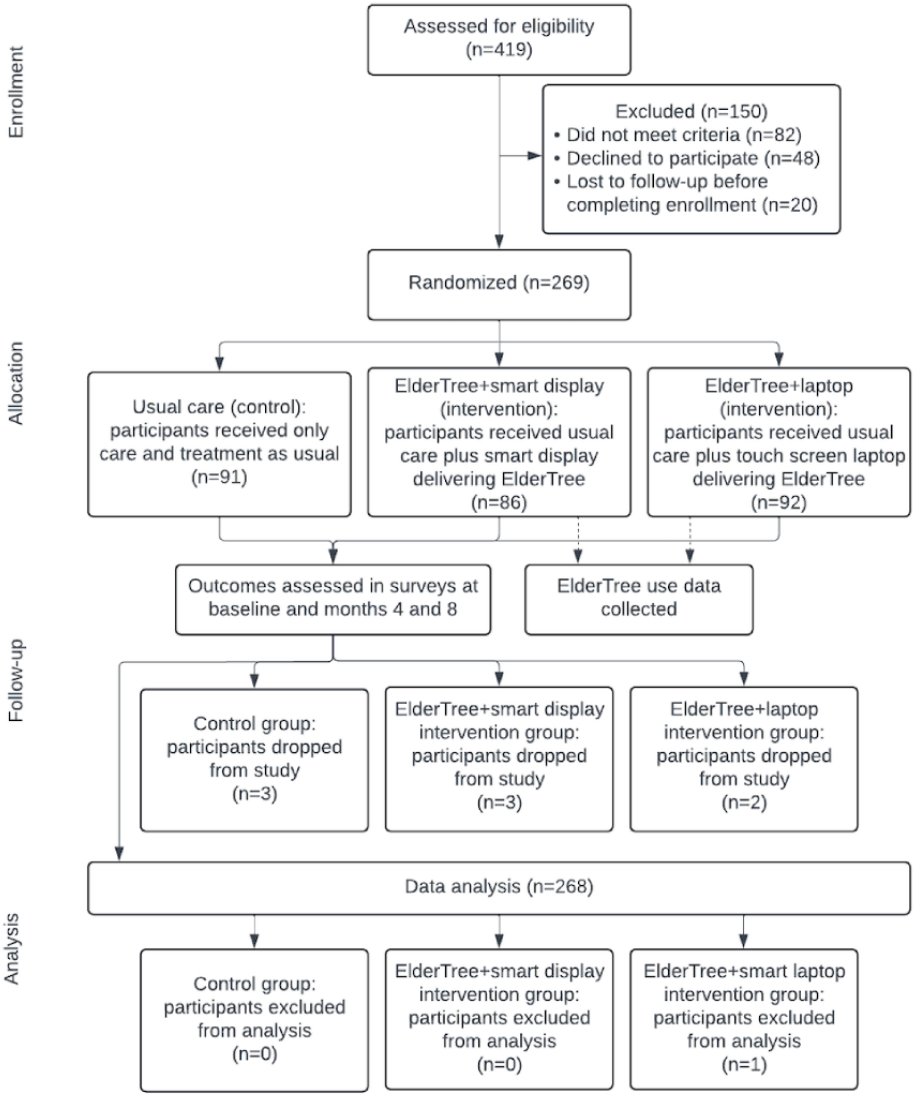
Results

Participants

A total of 269 participants were randomized (91 to control, 92 to ET-LT, and 86 to ET-SD); one participant was excluded because they did not complete the baseline survey, leaving 268 participants for analysis (Figure 2). Recruitment began in July 2021 and ended in December 2022; the intervention period continued through August 2023, which coincided with the historical context described above. Eight ET-SD participants were affected by Google's decision to discontinue the functionality that enabled access to ElderTree on the smart display: their intervention periods were 2 weeks to 2 months shorter, depending on enrollment date.

Participant baseline characteristics are reported by study arm in Table 1. Across all participants, 177 (66%) identified as female, and 211 (78.7%) identified as White only. The mean age was 69.8 (SD 7.34) years, and number of chronic conditions was 6.54 (SD 2.46).

Figure 2. Consolidated Standards of Reporting Trials (CONSORT) flow diagram. Data were analyzed using mixed-effects models, which incorporate all available data rather than only complete cases.



XSL•FO
RenderX

Table 1. Participant characteristics by study arm at baseline.

Characteristic	Control (N=91)	ET-LT[a] (N=91)	ET-SD[b] (N=86)
Gender, n (%)			
Female	63 (69)	59 (65)	55 (64)
Male	28 (31)	32 (35)	31 (36)
Race, n (%)			
Asian only	0 (0)	1 (1)	0 (0)
Black or African American and American Indian or Alaskan Native	2 (2)	1 (1)	0 (0)
Black, Euro-African, or African American	20 (22)	16 (18)	14 (16)
White only	68 (75)	72 (79)	71 (83)
White and Hispanic or Latino	1 (1)	1 (1)	0 (0)
White, Black or African American, American Indian or Alaskan Native	0 (0)	0 (0)	1 (1)
Highest level of education, n (%)			
Elementary school	3 (3)	0 (0)	0 (0)
High school	12 (13)	17 (19)	12 (14)
Some college	19 (21)	18 (20)	19 (22)
Vocational or technical school	16 (18)	13 (14)	11 (13)
College graduate	24 (26)	25 (27)	25 (29)
Postgraduate or professional	16 (18)	18 (20)	19 (22)
Living situation, n (%)			
Had significant other	50 (55)	49 (54)	52 (60)
Live alone	52 (57)	55 (60)	53 (62)
Had unplanned health service use, n (%)	18 (20)	21 (23)	20 (23)
Age (years), mean (SD)	69.49 (7.77)	69.76 (7.39)	70.13 (6.86)
Number of chronic conditions, mean (SD)	6.52 (2.31)	6.54 (2.54)	6.57 (2.55)
Life stressors (range 0-14), mean (SD)	2.81 (2.29)	2.69 (1.83)	2.65 (2.27)

[a]ET-LT: ElderTree laptop.

[b]ET-SD: ElderTree smart display.

Effects of Study Arm on Changes Over Time in Primary Outcomes

The central questions of this study were (1) whether delivering the intervention on different platforms affected changes over time in the primary outcomes of pain interference and psychosocial quality of life, and (2) whether both intervention arms would outperform the control group.

Pain Interference

Visualizations of the data indicated that, descriptively, the 3 study arms differed in baseline levels of pain interference (despite randomization), with different patterns by gender. Before examining the effects of study arm on changes over time, we first examined whether these baseline differences were significant. The results of 2-way ANOVAs predicting baseline pain interference found no significant effects of study arm, gender, or the gender×study arm interaction, comparing all 3 study arms (Table S1 in Multimedia Appendix 1). The analyses described below focus on changes over time in pain interference (ie, controlling for baseline scores), so initial differences would have been incorporated. Nonetheless, it is useful to indicate that the groups were statistically comparable at baseline despite descriptively different levels.

We did not find a significant difference between ET-LT and ET-SD for changes in pain interference over time (arm×time interaction b=–0.11, 95% CI –1.07 to 0.85; *P*=.82). There was also a nonsignificant difference between the combined ET-LT+ET-SD groups and control for change in pain interference over time (arm×time interaction b=–0.41, 95% CI –1.23 to 0.41; *P*=.33).

Psychosocial Quality of Life

Visualizations of the data indicated that, descriptively, the 3 study arms differed in baseline levels of psychosocial quality of life (despite randomization), with different patterns by gender. Again, we began by examining whether these differences were significant: they were not. Two-way ANOVAs predicting

XSL•FO
RenderX

baseline psychosocial quality of life found no significant differences by study arm (comparing all 3 arms or ET-LT+ET-SD vs control), nor significant interactions between study arm and gender. However, women scored significantly higher than men (Table S1 in Multimedia Appendix 1).

There was no significant difference between ET-LT and ET-SD for change in psychosocial quality of life over time (arm×time interaction b=–0.21, 95% CI –0.96 to 0.55; *P*=.56). There was also a nonsignificant difference between the combined ET-LT+ET-SD groups and control in change in psychosocial quality of life over time (arm×time interaction b=0.04, 95% CI –0.61 to 0.69; *P*=.90).

Mediation of Effects of Study Arm

No mediation analyses were conducted on the primary outcomes because the effects of study arm on these were nonsignificant.

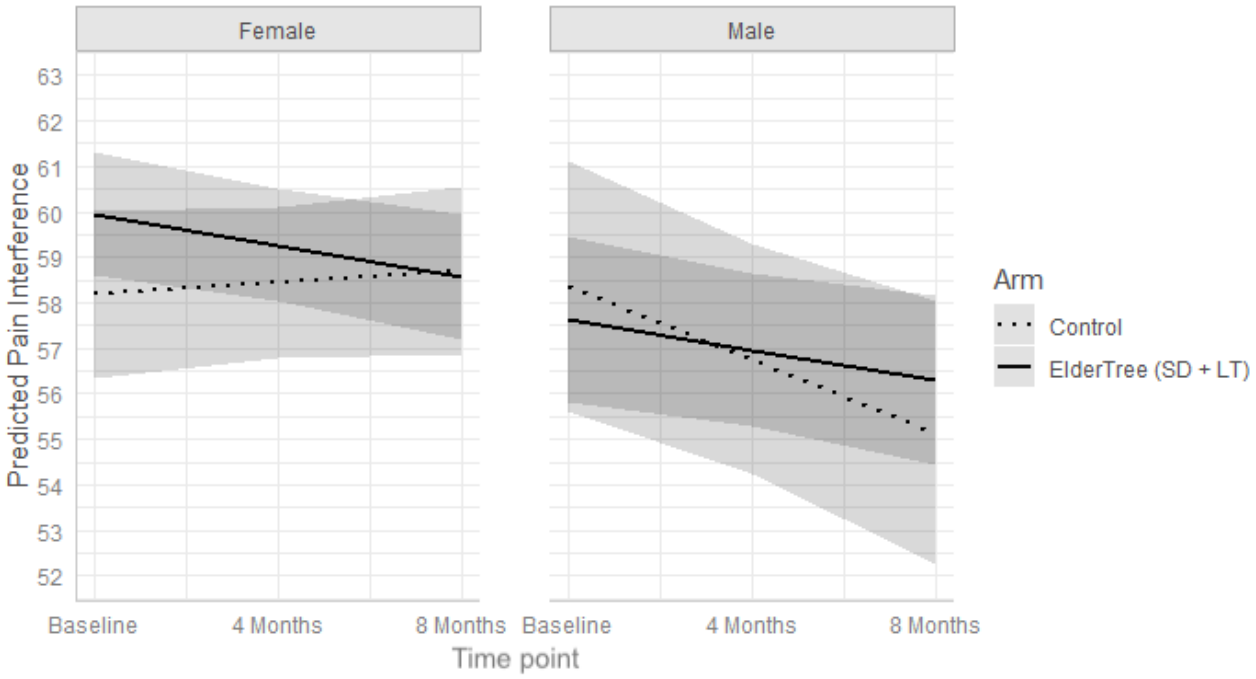
Moderation of Effects of Study Arm

We examined whether the effects of study arm on changes over time in the primary outcomes were moderated by gender, number of chronic conditions, and barriers to technology use.

Gender and Changes in Pain Interference

Gender did not moderate the effect of ET-LT versus ET-SD on changes over time in pain interference (b=–1.56, 95% CI –3.56 to 0.44; *P*=.13). However, as illustrated in Figure 3, gender did moderate the effect of the combined ET-LT+ET-SD groups versus control (b=1.91, 95% CI 0.11-3.71; *P*=.04). For women (n=177), those in the combined ET-LT+ET-SD group showed a significant decrease in pain interference over time (b=–0.69, 95% CI –1.29 to –0.10; *P*=.02) but those in the control group showed no significant change (b=0.25, 95% CI –0.53 to 1.04; *P*=.53). For men (n=91), those in the combined ET-LT+ET-SD group showed a nonsignificant decrease (b=–0.67, 95% CI –1.47 to 0.14; *P*=.10); those in the control group showed a significant decrease (b=–1.61, 95% CI –2.88 to –0.35; *P*=.01). However, even the significant decreases did not meet the benchmark for meaningful reductions. At 8 months, (controlling for baseline scores) there were no significant differences between study arms in pain interference for women (*P*=.10) or men (*P*=.27).

Figure 3. Predicted mean pain interference scores over time, stratified by gender.

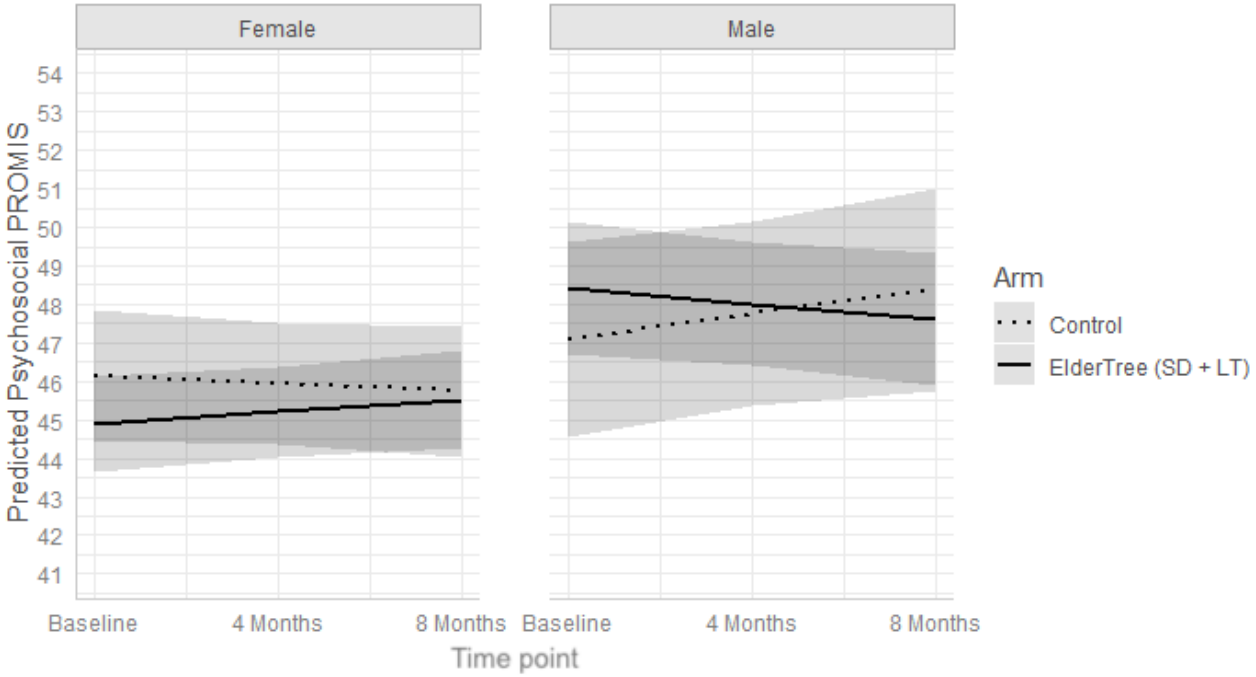


Gender and Changes in Psychosocial Quality of Life

Gender did not moderate the effect of ET-LT versus ET-SD on changes over time in psychosocial quality of life (b=1.27, 95% CI –0.31 to 2.85; *P*=.12). However, as illustrated in Figure 4, gender did moderate the effect of the combined ET-LT+ET-SD group versus control (b=–1.59, 95% CI –3.01 to –0.17; *P*=.03). For women, neither group showed significant changes over time: ET-LT+ET-SD (b=0.32, 95% CI –0.15 to 0.79; *P*=.18) and control (b=–0.21, 95% CI –0.36 to 1.64; *P*=.21). For men, the ET-LT+ET-SD group showed a significant decrease in psychosocial quality of life over time (b=–0.69, 95% CI –1.29 to –0.10; *P*=.02) whereas the control group remained stable (b=0.64, 95% CI –0.36 to 1.64; *P*=.21). None of the changes reached the benchmark for meaningful differences. At 8 months (controlling for baseline scores), there were no significant differences between study arms (control vs ET-LT vs ET-SD) for women (*P*=.25) or men (*P*=.13).

XSL•FO
RenderX

Figure 4. Predicted mean psychosocial quality of life scores over time, stratified by gender.



Other Moderators

To preview briefly, no additional moderation effects were observed. Number of chronic conditions did not moderate the effect of ET-LT versus ET-SD on changes over time in pain interference (b=–0.16, 95% CI –0.63 to 0.31; *P*=.51) or psychosocial quality of life (b=–0.32, 95% CI –0.70 to 0.06; *P*=.10). Nor did it moderate the effect of the combined ET-LT+ET-SD group versus control on changes over time in pain interference (b=–0.01, 95% CI –0.39 to 0.36; *P*=.94) or psychosocial quality of life (b=0.24, 95% CI –0.06 to 0.54; *P*=.12).

Barriers to technology did not moderate the effect of ET-LT versus ET-SD on changes over time in pain interference (b=0.97, 95% CI –1.20 to 3.15; *P*=.38) or psychosocial quality of life. Nor did they moderate the effect of the combined ET-LT+ET-SD group versus control on changes over time in pain interference (b=–0.55, 95% CI –2.50 to 1.40; *P*=.58) or psychosocial quality of life (b=0.28, 95% CI –1.27 to 1.84; *P*=.72).

Summary of Effects of Study Arm on Secondary and Exploratory Outcomes

As shown in Table S2 in Multimedia Appendix 1, there were no significant differences between ET-LT and ET-SD, or between the combined ET-LT+ET-SD groups versus control, in changes over time for any other secondary outcomes: physical quality of life, pain intensity, 30-day hospital readmissions, health distress, well-being, loneliness, irritability, or communication with physicians.

As shown in Table S3 in Multimedia Appendix 1, there were also no differences for either set of comparisons in changes over time for the exploratory outcomes of alcohol use, cigarette use, and falls.

ElderTree Usage and Ratings Overall and by Platform

Usage

Overall, the weekly survey was the most used service for both ET-LT and ET-SD. Weekly reminders by email and text prompted participants to take the survey; it took only a few minutes to complete and was easy to navigate on both platforms.

Other use patterns varied by platform. As shown in Table 2, participants assigned to ET-LT (vs ET-SD) used ElderTree on more days (Figure S1 in Multimedia Appendix 1 provides estimates of hours of use). In months 1-4, ET-LT participants used Elder Tree a mean of 28.47 (SD 19.28) days, compared with ET-SD participants, who used it for a mean of 19.02 (SD 16.09) days. In months 5-8, ET-LT participants used it a mean of 17.48 (SD 14.34) days, compared with ET-SD participants, who used it a mean of 10.08 (SD 9.85) days. These differences (approximately 1 week of use in each 4-month period) were significant (incidence rate ratio [IRR] 0.60, 95% CI 0.48-0.75; *P*<.001) and did not change in magnitude over time (IRR 0.85, 95% CI 0.71-1.02; *P*=.08).

XSL•FO
RenderX

Table 2. Laptop versus smart display in use and perceptions of ElderTree.

Outcome	Laptop (n=91), mean (SD)	Smart display (n=86), mean (SD)	Main effect of arm			Arm×time interaction		
			Estimate	95% CI	*P* value	Estimate	95% CI	*P* value
Days of ElderTree use								
Pain modules	4.95 (5.26)	4.92 (8.00)	0.97[a]	0.68 to 1.38	.86	1.25[a]	0.84 to 1.86	.27
Discussion group	12.36 (14.4)	2.39 (3.37)	0.18[a]	0.13 to 0.26	<.001	0.58[a]	0.38 to 0.87	.009
Health library	4.20 (5.50)	0.48 (0.89)	0.07[a]	0.04 to 0.12	<.001	0.29[a]	0.12 to 0.74	.01
weekly Survey	14.57 (6.89)	10.18 (6.06)	–4.32[b]	–6.12 to –2.52	<.001	0.56[b]	–1.12 to 2.23	.51
Perception of ElderTree								
ElderTree overall	3.02 0.58	2.59 0.65	–0.46[b]	–0.66 to –0.25	<.001	0.05[b]	–0.10 to 0.19	.52
Health benefits	2.72 0.74	2.44 0.78	–0.33[b]	–0.59 to –0.08	.01	0.08[b]	–0.12 to 0.27	.44
Socioemotional benefits	2.91 0.79	2.48 0.93	–0.51[b]	–0.80 to –0.21	<.001	0.13[b]	–0.09 to 0.35	.25
Ease of use	3.42 0.53	2.86 0.70	–0.55[b]	–0.77 to –0.34	<.001	–0.04[b]	–0.24 to 0.15	.66

[a]IRR: incidence rate ratio (for count outcomes).

[b]Regression coefficient (for continuous outcomes).

In terms of participation in the weekly meetups, among participants who remained in the study after meetups were offered, 39% (28/74) of ET-LT users and 36% (20/64) of ET-SD users attended at least once. In total, 48 participants (out of 138 possible) attended a meetup at least once (34.8%), and many attended multiple times resulting in 639 meetup admissions. Of those who participated each week, an average 70% (441/639) were ET-LT users and 30% (198/639) were ET-SD users. Only 6 (out of 28) ET-LT users and 4 (out of 20) ET-SD users attended but never contributed to the discussion during the meetup beyond an initial greeting. Participants were encouraged to participate, but it was not required if they felt more comfortable just listening.

There was no difference between ET-LT and ET-SD in the number of chronic pain modules started (overall mean 3.16 out of 8, SD 2.76), but more modules were completed in the ET-LT group (mean 2.63, SD 3.19) than in the ET-SD group (mean 1.74, SD 2.68; *P*=.04). There were no differences between ET-LT and ET-SD groups in the likelihood of attending one of the meetups (odds ratio [OR] 0.87, 95% CI 0.43-1.75; *P*=.71) or contributing verbally at a meetup (OR 0.92, 95% CI 0.43-1.95; *P*=.83).

Ratings of ElderTree

As shown in Table 2, mean ratings for benefits of ElderTree overall, health benefits, socioemotional benefits, and ease of use were generally above the mid-point (1=not at all to 4=very much). As Table 2 also shows, those in the ET-LT (vs ET-SD) group gave significantly higher ratings (averaged across the 4- and 8-month surveys) for all 4 measures. Significant arm×time effects are illustrated in Figures S2 and S3 in Multimedia Appendix 1.

Effects of ElderTree Usage and Perceptions on Primary Outcomes

Given that we did not find significant overall study arm effects on our primary outcomes, we examined whether ElderTree usage and perceptions predicted change in our primary outcomes. We calculated change scores for our primary outcomes by subtracting baseline scores for each outcome from the mean of that outcome across the 4- and 8-month assessments. We regressed these change scores for primary outcomes on the total number of days of ElderTree use and mean perceptions in separate models for each predictor and outcome. We also estimated separate models using either a combined sample with both ElderTree platforms (ET-LT+ET-SD) or separate samples for ET-LT and ET-SD. We did not apply any correction to *P* values to address α inflation due to multiple testing. Neither ElderTree usage nor any of the ElderTree perception scales significantly predicted change in primary outcomes across these 30 exploratory analyses (ie, 2 outcomes×5 predictors×3 samples; *P*=.10-.96).

Probing Possible Gender Differences in Uses

Although gender only moderated the effect of the combined ET-LT+ET-SD versus control, we explored whether there were gender differences in uses of ElderTree on the laptop versus smart display. We averaged across the entire study period for each person to simplify the models because models with time included produced convergence errors. Gender did not alter the differences between ET-LT and ET-SD for days of use (b=–0.21,

XSL•FO
RenderX

95% CI 0.23-0.94; *P*=.34), hours of use (b=–0.42, 95% CI 0.29-1.43; *P*=.15), likelihood of attending a meetup (arm×gender OR 0.90, 95% CI 0.18-4.27; *P*=.89), or talking at a meetup (arm×gender OR 0.81, 95% CI 0.14-4.36; *P*=.81).

Qualitative Insights

Textbox 1 lists core themes and example comments about perceived benefits of ElderTree. Health-related comments referenced the value of the breathing and exercise videos, increased awareness and knowledge of their conditions, and learning from the pain modules. Participants also commented on social connections, including the benefits of connecting with others facing similar challenges. Several participants wrote positively about platform affordances, with smart display users discussing the benefits of voice access and laptop users expressing a preference or interest in typing.

Textbox 1. Participant comments about perceived benefits: themes and examples. Codes are anonymized identifiers. Comments could reflect multiple themes and were categorized accordingly. LT: laptop user; SD: smart display user.

Health-related uses and benefits of ElderTree

I liked the fact that it answered a lot of my questions easily and I have access to information and the breathing techniques came in real handy. [SD AC53]

I learned so much on how to control and manage my pain! [SD AC65]

I now feel more in control of my physical problems and I'm more willing to try things I've learned in this study. [SD CO14]

I like the Weekly Survey because it makes me record my blood pressure readings daily and keep myself well hydrated daily! Plus keep a handle on my weight! [LT AC70]

I liked the 8 lessons I learned a lot and picked up some breathing exercises that helped my breathing.... lost 1/2 lung to cancer 10 years ago. [LT EC67]

Even when I dont open the computer I do the breathing exercises, lift lite weights, and stand on one leg. I plan to expand on this. I feel better and lost 6 lbs. What can be better than that. This program helped me achieve that…. feel so much better and healthier. [LT YA10]

Social connections

I really enjoy the Tuesday meet ups. I don't do much of interest so listen but it helps me feel involved. [SD EC86]

I brought my report to my doctor and she said "you seem so much calmer…why do you think that may be?" I think it's because I'm in this study where I'm learning and seeing that I'm not alone… there are others experiencing similar issues related to aging. [LT EC67]

I appreciate being able to talk to "real people" without needing to be the "perfect" senior. [LT HB11]

Other features and services of ElderTree

The games were a welcome diplatform sometimes. [SD CO14]

I liked… recipes, and playing some of the games. I also enjoyed doing some journaling. [LT EC15]

Platform benefits

What I like best about smart system is that I got to talk to it and it talk back. I was excited that it knew my name. [SD YA06]

I liked to voice command because the worst of all my chronic pain is in my hands! [SD EC44]

Helps me learn how to type [LT AC66]

Other benefits of device

Could access weather, news, music, podcasts, etc. [SD AC14]

Beautiful photos and music [SD AC16]

I can look up a lot of info to help myself get motivated. That I can keep in touch with doctors. [LT AC66]

The laptop is so convenient to take on our trips…i use it everywhere. in the condo, motels and again at home. [LT YA10]

Textbox 2 lists the core challenges and barriers referenced by participants. Chief among them were technical difficulties, including Wi-Fi connectivity issues and glitches, as well as limitations and slowness of the dictation function on the smart display. Relatedly, several participants wrote that they preferred their other devices (eg, phone or laptop) or felt too unfamiliar with their assigned device to use it often or effectively. Specific to ElderTree, some participants reported finding the material too easy or feeling that they were too young and healthy for the intervention and could not relate to other participants. Additional barriers included poor health and lack of time or motivation.

XSL•FO
RenderX

Textbox 2. Participant comments about challenges: themes and examples. Codes are anonymized identifiers. Comments could reflect multiple themes and were categorized accordingly. LT: laptop user; SD: smart display user.

Technical issues: Wi-Fi connectivity and ElderTree glitches

Every time (yes, every time) I used the ElderTree smart system, it couldn't stay connected to our Spectrum internet. Same problem when it was connected to ElderTree/Google's internet. Had to constantly reboot and/or just give up for that day. I was never able to connect to meetings with other participants. [SD CO14]

Lessons seemed difficult to complete in one session. Would be interrupted by "glitch" ie 2 pain lessons had to be completed in 3 or 4 attempts often resulting in loss of interest. [SD EC25]

Was very disappointed in the Journaling feature – too slow – made it useless for me. [SD YA03]

Google smart device – the dictation system was useless. [SD EC175]

For some reason, computer runs very slow. Going back to read next, notice it seems to lock up. Sometimes I need to log out and back in to get it to work. Then I just don't use it [LT YA46]

Unsure how to use device or navigate ElderTrees

I simply am not able to manage the device. Family members would respond with "Wow- you've got a...(thing)... began trying it out - I could not follow their enthusiastic directions- and any possible lesson was gone in a flash. I would call for help from the Tech Support number often. (I loved the pictures and time+temp). [SD EC84]

I am not very computer savvy. Took me awhile to navigate the various offerings (classes, library, etc). [LT EC67]

Did not like device (prefer other devices)

I am used to using a computer a lot. If I accessed ElderTree from a laptop rather than from the Google device I would have used it a lot more because I am often on the computer. But, I would need to make a special effort to go on to the Google device just for ElderTree [SD YA68]

Have the smart speaker. Don't like that I can't turn off the camera and audio separately, so I keep it turned off. [SD EC33]

Content pitched at wrong level, community not ideal fit

Sessions were very slow. Exercises very basic. Would have been better to offer levels of difficulty. [LT EC105]

Consider grouping participants. It's hard for me to relate to folks socially who live in care communities. [LT EC151]

Health issues

I have trouble with tremors, recently saw a neurologist and was told I am showing signs of Parkinsonisms. [LT EC194]

I had bronchitis and was too sick to use it for several weeks [SD HBO5]

Too busy with other things, under-motivated

I had other things, like reading books for 3 book groups I am in. [LT EC61]

I was often busy (taking a class, traveling, having guests, working on non-health life issues) and didn't spend consistent time on ElderTree. This interfered or slowed down my ability to make habit changes that ElderTree helped me to identify that were needed. Still working on these. [LT EC114]

Discussion

Overview

The goal of this large-sample, 8-month randomized controlled trial was to test the hypothesis that a smart display would be even better than a laptop for delivering a digital health intervention to older adults with chronic pain and multiple other chronic conditions. By "even better," we meant that participants randomized to use the intervention on the smart display, rather than the laptop, would use it more and show greater improvements over time in health-related outcomes. "Even better" also involved a comparison with the control group: older adults in both intervention arms were expected to show greater improvements over time relative to those receiving treatment as usual. As noted earlier, this appears to be one of the first large, randomized controlled trials to assess these comparisons, rather than focusing solely on feasibility and use. In discussing the findings, we first address the health-related outcomes before examining the use data and lessons learned.

First, in the sample as a whole, there were no significant differences between the two ElderTree platforms (laptop vs smart display) with respect to changes over time in any of the primary, secondary, or exploratory health-related outcomes. Second, there were no significant differences between the combined ElderTree group (laptop+smart display) and the control group on changes over time in any of the primary, secondary, or exploratory outcomes. Moreover, exploratory analyses did not detect any relationship between days of use or perceptions of ElderTree (whether combined across platforms or separately) and change in either primary outcomes. In other words, participants who used ElderTree more frequently or reported more positive perceptions of either platform did not show greater improvement in pain interference or psychosocial quality of life. However, our coarse measure of total days of

XSL•FO
RenderX

use may have been too coarse to detect meaningful differences based on usage.

The moderation analyses found no differences between ET-LT and ET-SD on any health-related outcomes for any of the hypothesized subgroups, including participants who reported few barriers to using their assigned platform. Although it was intuitive to expect that laptops would be superior for participants who found it easiest to type and read, and that smart displays would be superior for those who found it easiest to dictate and listen, we did not observe these patterns, even though a few participants mentioned appreciating their particular platform for precisely these reasons (Textbox 1).

For the comparison of the combined ElderTree group (ET-LT+ET-SD) versus control, only gender—not number of chronic conditions or reported barriers to use—was a significant moderator. Although the moderation results were somewhat complex, the overall takeaway is straightforward. Women showed the sole positive effect of ElderTree in the combined group, with a significant reduction in pain interference over time. Men showed the sole negative effect, with a significant decrease in psychosocial quality of life.

We found a similar gender difference in a previous trial of an earlier ElderTree version (focused on multiple chronic conditions), in which women—but not men—showed significant benefits of ElderTree versus an attention control on psychosocial quality of life and psychological well-being [49]. Given meta-analytic evidence that women tend to give and receive more social support online than men [63] we considered whether women in both studies engaged more with the discussion boards (and, in this study, with the meetups). In fact, in both studies, there were no differences between men and women in the amount or type of ElderTree use. Another possibility is that women benefited from the fact that there were more female than male participants in both studies (reflecting the demographics of this age group), so they likely experienced more same-sex interactions on ElderTree than men. Notably, however, when we offered same-sex discussion groups in the first iteration of ElderTree, participants made little use of them, instead converging on the shared discussion groups. Ultimately, we are reluctant to overinterpret the current gender differences, given that neither of the significant changes reached benchmarks for meaningful effects and that there were no significant differences between men and women by the end of the trial. We interpret the current findings as indicating that neither men nor women showed a meaningful effect of ElderTree, but future interventions should engage more men (as well as women) during development phase and in qualitative interviews to gain deeper insights into possible differences.

Why was our pain-focused intervention largely ineffective? As described in the Introduction, previous online interventions for chronic pain have typically found small effects unless participants' progress was guided by a clinician [38,39]. Our intervention was not clinician-guided, given the goal of creating a readily scalable program, and unlike previous online interventions for chronic pain, we targeted older adults concurrently managing multiple other chronic conditions. We hypothesized that ElderTree's combination of educational modules, discussion board and meetups, and other resources (eg, meditation and exercise videos, health library, and journaling) would yield positive results, particularly if offered on smart displays that could facilitate voice-activated engagement. We can identify at least 3 reasons for the overall lack of effects, in addition to the absence of clinician guidance and the complex health challenges of our older adult participants.

First, although we tried to make the modules engaging (eg, by adding inclusive images and short quizzes to highlight key points), the static slides with audio format and length (13-22 minutes), combined with some participants' difficulties with Wi-Fi connectivity and device glitches (Textbox 2), may have contributed to low use of that content. Second, use and effects of ElderTree may have been diminished by the context of the COVID-19 pandemic, not only because of the amount and urgency of other health-related information at the time, but also because, as one participant noted, there was a hunger for in-person rather than online interactions. Third, although the PROMIS-29 and its subscales have been validated, it is possible that the measures of pain interference (4 items) and psychosocial quality of life (12 items) were not subtle enough to assess change in these outcomes. Indeed, the reliance on self-report measures rather than clinical assessments is an important limitation of this study, though, given our primary focus on pain interference and quality of life, this seemed appropriate. Future studies may benefit from other measures, such as mobility and physical function.

Why were days of use and ratings of ElderTree higher in the laptop arm rather than in the smart display arm? We selected the Google Nest Hub Max because its Google Assistant–powered speech recognition, robust developer tools (Google Action SDK), and built-in camera seemed ideal for a voice‐first, multimedia intervention. In practice, however, core ElderTree features did not translate well. Threaded discussion, a core service of ElderTree that thrived on the laptop, became cumbersome on the smart display. Voice dictation forced users into short entries with repeated rerecording, so discussion‐group participation plummeted. Likewise, long‐form readings could not be scrolled or linked through URLs because the device was not designed for text‐heavy content.

The platform of smart display deployed in this study relied on pre-LLM (large language model) dialogue technology. These earlier systems had limited natural language capabilities, required a wake word, and were difficult to navigate, especially for users unfamiliar with command-based input. Additionally, some participants were more familiar with laptops or other devices like iPads or their phones than with smart displays.

We are currently concluding another randomized controlled trial of a subsequent iteration of ElderTree (ET-Move, focused on facilitating mobility in older adults with multiple chronic conditions), also comparing laptops and smart displays as platforms [64]. Based on the lessons of this study, we made several changes. First, to boost engagement and leverage the capabilities of the smart display, we shifted from text-based content to videos. Both the smart display and laptop platforms

XSL•FO
RenderX

offer many informational and movement videos. Most are under 5 minutes in length, and, except for the required safety videos, can be accessed in any order preferred by the user. Second, we shifted the emphasis away from pain to the gains achievable from safe exercise. Third, as suggested by some users in this study, we introduced different levels of content; ET-Move offers different types of exercise based on assessments of participants' physical capabilities. Fourth, we incorporated meetups throughout the study, to serve in part as a proxy for guidance. During these meetups, participants and moderators perform a simple set of exercises together, either before or after the social and educational aspects of the session.

We can see additional avenues for improving future interventions offered on smart displays. Most notably, interventions should incorporate LLM-based voice agents with natural, context-aware, and flexible dialogue. Such systems could support open-ended interaction and reduce the cognitive demands on users by eliminating the need for memorized commands. As AI models continue to evolve, so does the evidence base for their use in health interventions. A systematic review of 21 randomized controlled trials of AI-driven chatbot health interventions (thus far text-based rather than voice-driven on smart displays) found varied impacts on physical activity, functional abilities, adoption of healthier lifestyle habits, understanding of their medical conditions, and improvements in mental and psychosocial well-being [65]. A meta-analysis of 32 studies involving AI conversational agent interventions (also primarily text-based) showed statistically significant short-term improvements across a range of mental health outcomes, although long-term effects were generally nonsignificant [66]. Empathic responsiveness, extended interaction duration, and personal tailoring emerged as key factors enhancing efficacy.

Indeed, a core lesson of this study was that future systems should be tailored to user characteristics such as digital literacy and comfort with technology. Ultimately, while a randomized controlled trial examining platform or modality effects requires participants to use the assigned device, the real-world use of health interventions is likely enhanced by allowing access on the devices most familiar and useful to the participants. Moreover, for older adults who struggle to use the necessary technologies, extended onboarding periods and repeated, multimodal support are needed.

Overall, the results point to the challenges of using an online system to address chronic pain in older adults with multiple chronic conditions, particularly amid the challenges and stresses of a global pandemic. Despite the current findings, it is premature to declare that smart displays offer no benefit over laptops, especially given rapid advances in AI-based capabilities.

Conclusion

We found little evidence of meaningful effects of our online pain intervention for older adults with multiple chronic conditions, whether delivered on a voice-interactive smart display or on a laptop.

Acknowledgments

This study was funded by the Agency for Healthcare Research and Quality, United States Department of Health and Human Services (1R18HS026853-01A1), and supported by the Clinical and Translational Science Award program through the National Institutes of Health National Center for Advancing Translational Sciences (grant UL1TR002373). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors thank the research and technology staff at Comprehensive Health Enhancement Support Systems for their care and dedication to all aspects of this project: Dave Gustafson Jr, Hannah Lehr, Sam Nowak, Olivia Vjorn, Katherine Wright, and Matthew Wright.

Data Availability

Following the completion of the final analysis, the datasets generated or analyzed during this study will be available from the corresponding author upon reasonable request.

Conflicts of Interest

DHGS has a small shareholder interest in Comprehensive Health Enhancement Support Systems Health, a corporation that develops health care technology for patients and family members struggling with addiction; this relationship is managed by DHGS and the University of Wisconsin–Madison's Conflict of Interest Committee. The authors have no other conflicts to report.

Multimedia Appendix 1
Additional information for outcome analyses.
[DOCX File , 50 KB - aging_v8i1e75991_app1.docx]

Multimedia Appendix 2
CONSORT eHEALTH checklist V (1.6.1).
[PDF File (Adobe PDF File), 1072 KB - aging_v8i1e75991_app2.pdf]

References

XSL•FO
RenderX

1. Stubbs B, Binnekade TT, Soundy A, Schofield P, Huijnen IPJ, Eggermont LHP. Are older adults with chronic musculoskeletal pain less active than older adults without pain? A systematic review and meta-analysis. Pain Med 2013;14(9):1316-1331. [doi: 10.1111/pme.12154] [Medline: 23742160]
2. Wackström N, Koponen AM, Suominen S, Tarkka IM, Simonsen N. Does chronic pain hinder physical activity among older adults with type 2 diabetes? Health Psychol Behav Med 2020;8(1):362-382 [FREE Full text] [doi: 10.1080/21642850.2020.1807350] [Medline: 34040877]
3. Berryman C, Stanton TR, Jane Bowering K, Tabor A, McFarlane A, Lorimer Moseley G. Evidence for working memory deficits in chronic pain: a systematic review and meta-analysis. Pain 2013;154(8):1181-1196. [doi: 10.1016/j.pain.2013.03.002] [Medline: 23707355]
4. van der Leeuw G, Eggermont LHP, Shi L, Milberg WP, Gross AL, Hausdorff JM, et al. Pain and cognitive function among older adults living in the community. J Gerontol A Biol Sci Med Sci 2016;71(3):398-405 [FREE Full text] [doi: 10.1093/gerona/glv166] [Medline: 26433218]
5. Gustafson DH, Mares M, Johnston DC, Landucci G, Pe-Romashko K, Vjorn OJ, et al. Using smart displays to implement an ehealth system for older adults with multiple chronic conditions: protocol for a randomized controlled trial. JMIR Res Protoc 2022;11(5):e37522 [FREE Full text] [doi: 10.2196/37522] [Medline: 35511229]
6. McCloud R, Perez C, Bekalu MA, Viswanath K. Using smart speaker technology for health and well-being in an older adult population: pre-post feasibility study. JMIR Aging 2022;5(2):e33498 [FREE Full text] [doi: 10.2196/33498] [Medline: 35532979]
7. Edwards KJ, Jones RB, Shenton D, Page T, Maramba I, Warren A, et al. The use of smart speakers in care home residents: implementation study. J Med Internet Res 2021;23(12):e26767 [FREE Full text] [doi: 10.2196/26767] [Medline: 34932010]
8. Davitt J, Brown J. Using voice and touchscreen controlled smart speakers to protect vulnerable clients in long-term care facilities. Innov Aging 2022;6(4):igac024 [FREE Full text] [doi: 10.1093/geroni/igac024] [Medline: 35712325]
9. Zhong R, Ma M, Zhou Y, Lin Q, Li L, Zhang N. User acceptance of smart home voice assistant: a comparison among younger, middle-aged, and older adults. Univers Access Inf Soc 2022:1-18 [FREE Full text] [doi: 10.1007/s10209-022-00936-1] [Medline: 36338377]
10. Werner L, Huang G, Pitts BJ. Smart speech systems: a focus group study on older adult user and non-user perceptions of speech interfaces. Int J Hum Comput Interact 2022;39(5):1149-1161. [doi: 10.1080/10447318.2022.2050541]
11. Desai S, Chin J. OK Google, let's learn: using voice user interfaces for informal self-regulated learning of health topics among younger and older adults. 2023 Presented at: CHI '23: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems; April 19, 2023; Hamburg p. 1-21.
12. Sin J, Munteanu C, Ramanand N, Rong TY. VUI influencers: how the media portrays voice user interfaces for older adults. 2021 Presented at: CUI '21: Proceedings of the 3rd Conference on Conversational User Interfaces; July 27, 2021; Bilbao p. 1-13 URL: https://doi.org/10.1145/3469595.3469603 [doi: 10.1145/3469595.3469603]
13. Edwards KJ, Jones RB, Shenton D, Page T, Maramba I, Warren A, et al. The use of smart speakers in care home residents: implementation study. J Med Internet Res 2021;23(12):e26767 [FREE Full text] [doi: 10.2196/26767] [Medline: 34932010]
14. McCloud R, Perez C, Bekalu MA, Viswanath K. Using smart speaker technology for health and well-being in an older adult population: Pre-post feasibility study. JMIR Aging 2022;5(2):e33498 [FREE Full text] [doi: 10.2196/33498] [Medline: 35532979]
15. Kurokawa S, Urata M, Endo M, Yasuda T, Inoue A. Development of a smart speaker application to promote continuous exercise in the elderly. 2023 Presented at: 2023 IEEE 12th Global Conference on Consumer Electronics (GCCE); October 10-13, 2023; Nara p. 527-531 URL: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10315332
16. Zhong R, Ma M, Zhou Y, Lin Q, Li L, Zhang N. User acceptance of smart home voice assistant: a comparison among younger, middle-aged, and older adults. Univers Access Inf Soc 2024;23:275-292. [doi: 10.1007/s10209-022-00936-1]
17. Werner L, Huang G, Pitts BJ. Automated speech recognition systems and older adults: a literature review and synthesis. Proc Hum Factors Ergon Soc Annu Meet 2019;63(1):42-46 [FREE Full text] [doi: 10.1177/1071181319631121]
18. Werner L, Huang G, Pitts BJ. Smart speech systems: a focus group study on older adult user and non-user perceptions of speech interfaces. Int J Hum Comput Interact 2022;39(5):1149-1161 [FREE Full text] [doi: 10.1080/10447318.2022.2050541]
19. Pradhan A, Lazar A, Findlater L. Use of intelligent voice assistants by older adults with low technology use. ACM Trans Comput-Hum Interact 2020;27(4):1-27 [FREE Full text] [doi: 10.1145/3373759]
20. Arnold A, Kolody S, Comeau A, Miguel Cruz A. What does the literature say about the use of personal voice assistants in older adults? A scoping review. Disabil Rehabil Assist Technol 2024;19(1):100-111. [doi: 10.1080/17483107.2022.2065369] [Medline: 35459429]
21. Jones VK, Yan C, Shade MY, Boron JB, Yan Z, Heselton HJ, et al. Reducing loneliness and improving social support among older adults through different modalities of personal voice assistants. Geriatrics (Basel) 2024;9(2):22 [FREE Full text] [doi: 10.3390/geriatrics9020022] [Medline: 38525739]
22. Park S, Kim B. The impact of everyday AI-based smart speaker use on the well-being of older adults living alone. Technology in Society 2022;71:102133 [FREE Full text] [doi: 10.1016/j.techsoc.2022.102133]

XSL•FO
RenderX

23. Nayak A, Vakili S, Nayak K, Nikolov M, Chiu M, Sosseinheimer P, et al. Use of voice-based conversational artificial intelligence for basal insulin prescription management among patients with type 2 diabetes: a randomized clinical trial. JAMA Netw Open 2023;6(12):e2340232 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.40232] [Medline: 38039007]
24. Kim D. Can healthcare apps and smart speakers improve the health behavior and depression of older adults? A quasi-experimental study. Front Digit Health 2023;5:1117280 [FREE Full text] [doi: 10.3389/fdgth.2023.1117280] [Medline: 36910571]
25. Lucas JW, Sohi I. Chronic pain and high-impact chronic pain in U.S. adults, 2023. NCHS Data Brief 2024(518):CS355235. [doi: 10.15620/cdc/169630] [Medline: 39751180]
26. De La Rosa JS, Brady BR, Ibrahim MM, Herder KE, Wallace JS, Padilla AR, et al. Co-occurrence of chronic pain and anxiety/depression symptoms in U.S. adults: prevalence, functional impacts, and opportunities. Pain 2024;165(3):666-673 [FREE Full text] [doi: 10.1097/j.pain.0000000000003056] [Medline: 37733475]
27. Foley HE, Knight JC, Ploughman M, Asghari S, Audas R. Association of chronic pain with comorbidities and health care utilization: a retrospective cohort study using health administrative data. Pain 2021;162(11):2737-2749. [doi: 10.1097/j.pain.0000000000002264] [Medline: 33902092]
28. McQueenie R, Jani BD, Siebert S, McLoone P, McCowan C, Macdonald S, et al. Prevalence of chronic pain in LTCs and multimorbidity: a cross-sectional study using UK Biobank. J Multimorb Comorb 2021;11:26335565211005870 [FREE Full text] [doi: 10.1177/26335565211005870] [Medline: 35004337]
29. Nakad L, Booker S, Gilbertson-White S, Shaw C, Chi N, Herr K. Pain and multimorbidity in late life. Curr Epidemiol Rep 2020;7(1):1-8. [doi: 10.1007/s40471-020-00225-6]
30. Eckerblad J, Theander K, Ekdahl A, Unosson M, Wirehn A, Milberg A, et al. Symptom burden in community-dwelling older people with multimorbidity: a cross-sectional study. BMC Geriatr 2015;15:1 [FREE Full text] [doi: 10.1186/1471-2318-15-1] [Medline: 25559550]
31. Neba RA, Wang H, Kolala M, Sambamoorthi U. Multimorbidity and chronic pain management with opioids and other therapies among adults in the United States: a cross-sectional study. J Multimorb Comorb 2024;14:26335565241237889 [FREE Full text] [doi: 10.1177/26335565241237889] [Medline: 38454920]
32. Schneider J, Algharably EAE, Budnick A, Wenzel A, Dräger D, Kreutz R. High prevalence of multimorbidity and polypharmacy in elderly patients with chronic pain receiving home care are associated with multiple medication-related problems. Front Pharmacol 2021;12:686990 [FREE Full text] [doi: 10.3389/fphar.2021.686990] [Medline: 34168565]
33. Marupuru S, Axon DR. Association of multimorbidity on healthcare expenditures among older united states adults with pain. J Aging Health 2021;33(9):741-750. [doi: 10.1177/08982643211011841] [Medline: 33881371]
34. Runge N, Ahmed I, Saueressig T, Perea J, Labie C, Mairesse O, et al. The bidirectional relationship between sleep problems and chronic musculoskeletal pain: a systematic review with meta-analysis. Pain 2024;165(11):2455-2467. [doi: 10.1097/j.pain.0000000000003279] [Medline: 38809241]
35. Gao S, Zhou H, Luo S, Cai X, Ye F, He Q, et al. Investigating the causal relationship between physical activity and chronic back pain: a bidirectional two-sample mendelian randomization study. Front Genet 2021;12:758639 [FREE Full text] [doi: 10.3389/fgene.2021.758639] [Medline: 34987546]
36. Brain K, Burrows TL, Bruggink L, Malfliet A, Hayes C, Hodson FJ, et al. Diet and chronic non-cancer pain: the state of the art and future directions. J Clin Med 2021;10(21):5203 [FREE Full text] [doi: 10.3390/jcm10215203] [Medline: 34768723]
37. Werneck AO, Stubbs B. Bidirectional relationship between chronic pain and depressive symptoms in middle-aged and older adults. Gen Hosp Psychiatry 2024;89:49-54. [doi: 10.1016/j.genhosppsych.2024.05.007] [Medline: 38761582]
38. Gandy M, Pang ST, Scott AJ, Heriseanu AI, Bisby MA, Dudeney J, et al. Internet-delivered cognitive and behavioural based interventions for adults with chronic pain: a systematic review and meta-analysis of randomized controlled trials. Pain 2022;163(10):e1041-e1053. [doi: 10.1097/j.pain.0000000000002606] [Medline: 35121696]
39. Terpstra JA, van der Vaart R, van Beugen S, van Eersel RA, Gkika I, Erdős D, et al. Guided internet-based cognitive-behavioral therapy for patients with chronic pain: a meta-analytic review. Internet Interv 2022;30:100587 [FREE Full text] [doi: 10.1016/j.invent.2022.100587] [Medline: 36406977]
40. Baumeister H, Paganini S, Sander L, Lin J, Schlicker S, Terhorst Y, et al. Effectiveness of a guided internet- and mobile-based intervention for patients with chronic back pain and depression (WARD-BP): a multicenter, pragmatic randomized controlled trial. Psychother Psychosom 2021;90(4):255-268. [doi: 10.1159/000511881] [Medline: 33321501]
41. Buhrman M, Syk M, Burvall O, Hartig T, Gordh T, Andersson G. Individualized guided internet-delivered cognitive-behavior therapy for chronic pain patients with comorbid depression and anxiety: a randomized controlled trial. Clin J Pain 2015;31(6):504-516. [doi: 10.1097/AJP.0000000000000176] [Medline: 25380222]
42. Hedman-Lagerlöf M, Hedman-Lagerlöf E, Axelsson E, Ljótsson B, Engelbrektsson J, Hultkrantz S, et al. Internet-delivered exposure therapy for fibromyalgia: a randomized controlled trial. Clin J Pain 2018;34(6):532-542. [doi: 10.1097/AJP.0000000000000566] [Medline: 29077623]
43. Lorig KR, Ritter PL, Laurent DD, Plant K. The internet-based arthritis self-management program: a one-year randomized trial for patients with arthritis or fibromyalgia. Arthritis Rheum 2008;59(7):1009-1017. [doi: 10.1002/art.23817] [Medline: 18576310]

XSL•FO
RenderX

44. Burke D, Lennon O, Blake C, Nolan M, Barry S, Smith E, et al. An internet-delivered cognitive behavioural therapy pain management programme for spinal cord injury pain: a randomized controlled trial. Eur J Pain 2019;23(7):1264-1282. [doi: 10.1002/ejp.1402] [Medline: 31002442]
45. Miller J, MacDermid JC, Walton DM, Richardson J. Chronic pain self-management support with pain science education and exercise (COMMENCE) for people with chronic pain and multiple comorbidities: a randomized controlled trial. Arch Phys Med Rehabil 2020;101(5):750-761. [doi: 10.1016/j.apmr.2019.12.016] [Medline: 32004517]
46. Yih ET, Silver JK. Letter to the editor on "Chronic pain self-management support with pain science education and exercise (COMMENCE) for people with chronic pain and multiple comorbidities: a randomized controlled trial". Arch Phys Med Rehabil 2021;102(1):160-161 [FREE Full text] [doi: 10.1016/j.apmr.2020.06.032] [Medline: 32962799]
47. Gustafson DH, Mares M, Johnston DC, Mahoney JE, Brown RT, Landucci G, et al. A web-based ehealth intervention to improve the quality of life of older adults with multiple chronic conditions: protocol for a randomized controlled trial. JMIR Res Protoc 2021;10(2):e25175 [FREE Full text] [doi: 10.2196/25175] [Medline: 33605887]
48. Gustafson DH, Maus A, Judkins J, Dinauer S, Isham A, Johnson R, et al. Using the NIATx model to implement user-centered design of technology for older adults. JMIR Hum Factors 2016;3(1):e2 [FREE Full text] [doi: 10.2196/humanfactors.4853] [Medline: 27025985]
49. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000;55(1):68-78. [doi: 10.1037//0003-066x.55.1.68] [Medline: 11392867]
50. Gustafson DH, Kornfield R, Mares M, Johnston DC, Cody OJ, Yang EF, et al. Effect of an eHealth intervention on older adults' quality of life and health-related outcomes: a randomized clinical trial. J Gen Intern Med 2022;37(3):521-530 [FREE Full text] [doi: 10.1007/s11606-021-06888-1] [Medline: 34100234]
51. Gustafson DH, Mares M, Johnston D, Vjorn OJ, Curtin JJ, Landucci G, et al. An eHealth intervention to improve quality of life, socioemotional, and health-related measures among older adults with multiple chronic conditions: randomized controlled trial. JMIR Aging 2024;7:e59588 [FREE Full text] [doi: 10.2196/59588] [Medline: 39642938]
52. Chronic conditions among Medicare beneficiaries: chartbook 2012 edition. U.S. Centers for Medicare & Medicaid Services. 2012. URL: https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/chronic-conditions/downloads/2012chartbook.pdf [accessed 2024-07-14]
53. Gustafson DH, Maus A, Judkins J, Dinauer S, Isham A, Johnson R, et al. Using the NIATx model to implement user-centered design of technology for older adults. JMIR Hum Factors 2016;3(1):e2 [FREE Full text] [doi: 10.2196/humanfactors.4853] [Medline: 27025985]
54. Gustafson DH, Mares M, Johnston DC, Mahoney JE, Brown RT, Landucci G, et al. A web-based eHealth intervention to improve the quality of life of older adults with multiple chronic conditions: protocol for a randomized controlled trial. JMIR Res Protoc 2021;10(2):e25175 [FREE Full text] [doi: 10.2196/25175] [Medline: 33605887]
55. Hays RD, Bjorner JB, Revicki DA, Spritzer KL, Cella D. Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. Qual Life Res 2009;18(7):873-880. [doi: 10.1007/s11136-009-9496-9] [Medline: 19543809]
56. Global Health Scoring Manual: A Brief Guide to Scoring the PROMIS Global Health Instruments. 2021. URL: https://www.healthmeasures.net/images/PROMIS/manuals/Scoring_Manuals_/PROMIS_Global_Health_Scoring_Manual.pdf [accessed 2024-01-12]
57. Homco J, Rodriguez K, Bardach DR, Hahn EA, Morton S, Anderson D, et al. Variation and change over time in PROMIS-29 survey results among primary care patients with type 2 diabetes. J Patient Cent Res Rev 2019;6(2):135-147 [FREE Full text] [doi: 10.17294/2330-0698.1694] [Medline: 31414025]
58. Namkoong K, DuBenske LL, Shaw BR, Gustafson DH, Hawkins RP, Shah DV, et al. Creating a bond between caregivers online: effect on caregivers' coping strategies. J Health Commun 2012;17(2):125-140 [FREE Full text] [doi: 10.1080/10810730.2011.585687] [Medline: 22004055]
59. Levesque CS, Williams GC, Elliot D, Pickering MA, Bodenhamer B, Finley PJ. Validating the theoretical structure of the Treatment Self-Regulation Questionnaire (TSRQ) across three different health behaviors. Health Educ Res 2007;22(5):691-702. [doi: 10.1093/her/cyl148] [Medline: 17138613]
60. Sullivan MJL. The Pain Catastrophizing Scale: User Manual. Montreal: McGill University; 2009. URL: https://aspecthealth.ca/wpcontent/uploads/2017/03/PCSManual_English1.pdf [accessed 2025-03-17]
61. Wallace D, Cooper NR, Sel A, Russo R. The social readjustment rating scale: updated and modernised. PLoS One 2023;18(12):e0295943 [FREE Full text] [doi: 10.1371/journal.pone.0295943] [Medline: 38109368]
62. Goldberg SB, Kendall AD, Hirshberg MJ, Dahl CJ, Nahum-Shani I, Davidson RJ, et al. Is dosage of a meditation app associated with changes in psychological distress? It depends on how you ask. Clin Psychol Sci 2025;13(2):332-349. [doi: 10.1177/21677026241266567] [Medline: 40162042]
63. Meaningful Change for PROMIS® :Estimates for the Magnitude of Change that Corresponds to “Important” Change Depend on if You Are Comparing Groups or Monitoring an Individual. URL: https://www.healthmeasures.net/score-and-interpret/interpret-scores/promis/meaningful-change [accessed 2022-02-15]
64. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd Edition. Hillside: Lawrence Erlbaum Associates; 1988.

XSL•FO
RenderX

65. Tifferet S. Gender differences in social support on social network sites: a meta-analysis. Cyberpsychol Behav Soc Netw 2020;23(4):199-209. [doi: 10.1089/cyber.2019.0516] [Medline: 32096662]
66. Gustafson D, Mares M, Johnston D, Curtin J, Pe-Romashko K, Landucci G. Comparison of smart display versus laptop platforms for an eHealth intervention to improve functional health for older adults with multiple chronic conditions: protocol for a randomized clinical trial. JMIR Res Protoc 2025;14:e64449 [FREE Full text] [doi: 10.2196/64449] [Medline: 40080672]

Abbreviations

AI: artificial intelligence
ET-LT: ElderTree laptop
ET-SD: ElderTree smart display
IRR: incidence rate ratio
LLM: large language model
OR: odds ratio
PROMIS-29: Patient-Reported Outcomes Measurement Information System 29-item profile
UW: University of Wisconsin

Edited by P Jansons; submitted 01.05.25; peer-reviewed by N Lukkahatai, F Yeh; comments to author 10.06.25; revised version received 12.09.25; accepted 16.09.25; published 18.11.25.

Please cite as:
Landucci G, Gustafson Sr DH, Mares ML, Pe-Romashko K, Curtin JJ, Hu Y, Maus A, Thompson K, Saunders S, Brown K, Woodburn J, Mutlu B
Using Smart Displays to Implement an eHealth System for Older Adults With Multiple Chronic Conditions: Randomized Controlled Trial
JMIR Aging 2025;8:e75991
URL: https://aging.jmir.org/2025/1/e75991
doi:10.2196/75991
PMID:40961249

©Gina Landucci, David H Gustafson Sr, Marie-Louise Mares, Klaren Pe-Romashko, John J Curtin, Yaxin Hu, Adam Maus, Kasey Thompson, Sydney Saunders, Kaitlyn Brown, Judith Woodburn, Bilge Mutlu. Originally published in JMIR Aging (https://aging.jmir.org), 18.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Engaging Older Adults and Staff in the Co-Design and Evaluation of Socially Assistive Robot and Virtual Reality Activities for Long-Term Care: User-Centered Study

Ritam Ghosh[1], MSEE; Nibraas Khan[2], MSCS; Miroslava Migovich[3], PhD; Judith A Tate[4], PhD; Cathy A Maxwell[5], PhD; Paul A Newhouse[6], MD; Douglas W Scharre[7], MD; Alai Tan[4], PhD, MD; Lorraine C Mion[4], PhD; Nilanjan Sarkar[3,1,2], PhD

[1]Department of Electrical and Computer Engineering, School of Engineering, Vanderbilt University, Nashville, TN, United States

[2]Department of Computer Science, School of Engineering, Vanderbilt University, Nashville, TN, United States

[3]Department of Mechanical Engineering, School of Engineering, Vanderbilt University, Nashville, TN, United States

[4]College of Nursing, The Ohio State University, Columbus, OH, United States

[5]College of Nursing, University of Utah, Salt Lake City, UT, United States

[6]Center for Cognitive Medicine, Department of Psychiatry and Behavioral Sciences, Vanderbilt University Medical Center, Nashville, TN, United States

[7]Department of Neurology, The Ohio State University Wexner Medical Center, Columbus, OH, United States

Corresponding Author:
Ritam Ghosh, MSEE
Department of Electrical and Computer Engineering
School of Engineering
Vanderbilt University
2301 Vanderbilt Place
Nashville, TN, 37235
United States
Phone: 1 615 322 7311
Email: ritam.ghosh@vanderbilt.edu

Abstract

Background: Apathy is common among older adults residing in long-term care (LTC) and impairs quality of life for both older adults and care providers. Few pharmacological remedies exist, and nonpharmacologic approaches that engage those with apathy require extensive personnel time. Thus, technological approaches have been encouraged, including virtual reality (VR) and socially assistive robots (SAR). Despite a growing interest in their use, input from older adults and staff is often absent in their design. Involving older adults in the development of interactive health technologies is necessary to enhance the functionality, usability, and likelihood of promoting the intended health outcomes.

Objective: We aimed to design and evaluate SAR and nonimmersive VR (SAR-VR) activities for pairs of older adults that would encourage human-to-human interaction, an essential activity to mitigate apathy.

Methods: We implemented a multistep, user-centered design. A humanoid and dog SAR were used in combination with nonimmersive VR activities for pairs of older adults. An interdisciplinary team of engineers, nurses, and physicians collaborated with older adults and staff to create 4 activity prototypes, 3 with the humanoid robot and 1 with the dog robot. A total of 14 older adults at 2 sites participated in the design and evaluation of the different components of the system throughout all stages. Site 1 participants were instrumental in the development, and Site 2 participants validated the prototype activities. Data were collected at each session via observations, interviews, and a 6-item questionnaire that rated their degree of comfort and confidence in (1) using the wands, (2) interacting with the robot, and (3) interacting with the nonimmersive VR environment using a 5-point Likert response. Additionally, 5 staff from Site 2 were recruited to evaluate the ease of setting up and running the system at 2 different sessions. After each session, the system setup and interface were refined based on their feedback.

Results: A total of 4 of 6 older adults (mean age 85, SD 9.3 years; 2 male) at Site 1 completed field testing development, and 8 residents (mean age 80, SD 4.7 years; 2 male) at Site 2 completed field testing validation. Participant comfort and confidence increased significantly over successive iterations of the system across most categories (Site 1: Wilcoxon signed rank test P=.03;

XSL•FO
RenderX

Site 2: Wilcoxon signed rank test $P<.001$). Additionally, 5 LTC staff members successfully set up the system with minimal cueing from the researchers, demonstrating the usability of the system for caregivers. Iterative design changes incorporated hardware, software, and activity domains.

Conclusions: These initial results demonstrate that LTC older adults and staff are capable and critical to the development and implementation of SAR-VR activities. Future studies are needed to evaluate the feasibility of implementation and effectiveness in reducing apathy.

Trial Registration: ClinicalTrials.gov NCT05178992; https://clinicaltrials.gov/study/NCT05178992

(JMIR Aging 2025;8:e75288) doi:10.2196/75288

KEYWORDS

apathy; long-term care settings; nonimmersive virtual reality; older adults; socially assistive robots

Introduction

In the United States, 1.3 million older adults reside in nursing homes, and another 918,000 reside in assisted living facilities [1], representing 3.8% of the older adult population. However, the lifetime probability of requiring long-term care (LTC) is 70% for those who survive to the age of 65 years, and 15% will spend 2 or more years in a nursing home [2]. The presence of apathy complicates the care delivery to these vulnerable older adults. Apathy is a condition that results in a lack of initiative, loss of interest in daily activities, and reluctance toward social interactions [3]. Apathy is common among LTC residents, occurring in up to 82% of those with dementia and up to a third of those with depression [4-6]. Older adults experiencing apathy often remain disengaged with their surroundings and may require significant prompting and encouragement to perform basic activities, such as speaking, eating, bathing, or walking. Apathy has significant deleterious consequences, including accelerated cognitive decline, further functional deficits, diminished quality of life, and increased rates of mortality [7]. Not only does apathy negatively impact the older adult, but it also causes significant burden and stress on family and staff [3,4].

Despite the scope and significance of apathy, few pharmacologic options are available [8]; thus, managing apathy requires effective nonpharmacologic strategies. Activities that combine motor-based and cognitive aspects with social aspects, that is, human-to-human interaction (HHI), appear to be the most effective for enhancing engagement and reducing apathy among older adults [9-13]. Unfortunately, designing and delivering these multimodal activities require significant personnel resources, and many LTCs have inadequate staffing, either in labor quantity or skill [14,15]. Thus, technology has been suggested to complement existing staff and facilitate the delivery of care to older adults [16-18]. Although most technological interventions in LTCs have focused on the electronic health record, telehealth, and wearables, there is a growing interest in the use of nonimmersive or immersive virtual reality (VR) [19-21] or socially assistive robots (SARs) [22,23] to enhance function or quality of life. Our work focuses on the combination of SAR with nonimmersive VR, hereafter referred to as SAR-VR.

Both SARs and VR have shown promising results for older adults [19,22,24], and each has its strengths as well as limitations that a combined approach can address. SARs have the distinct advantage of physical presence, the ability to emulate body language, and use of nonverbal cues that foster trust and engagement beyond what on-screen avatars typically achieve [25,26]. However, the readily available commercial SARs have limited capability to manipulate physical objects, while custom SARs are expensive and often designed for a specific application, limiting their activities. Conversely, nonimmersive VRs allow for the emulation of various physical scenarios displayed on a standard computer monitor that can provide a greater range of activities compared to those provided by robots alone. VR activities can also be easily adjusted to accommodate an individual's physical capabilities (eg, impaired mobility, reduced strength, diminished range of motion). However, nonimmersive VR limits the interaction to the use of a mouse, joystick, or remote control, which may be difficult for older adults with reduced motor dexterity and proprioception in hands and wrists [27-29].

An SAR-VR system leverages the strengths of both modalities while compensating for their weaknesses. The SAR provides embodied, socially engaging interaction that motivates participation, while the VR component expands the scope and adaptability of activities beyond the robot's physical affordances. Together, the two create more holistic activities that are simultaneously socially engaging, physically accessible, and adaptable to the needs of aging populations. Recent work has demonstrated the successful combination of assistive robotics with nonimmersive VR for other populations, including stroke rehabilitation [30,31], rehabilitation for children with motor skill impairment [32], and conducting video lectures [33]. In our earlier pilot work [34], we successfully combined an SAR with a nonimmersive VR-based activity using gesture control for engaging pairs of older adults residing in the community and subsequently in LTCs. The activity was designed based on input from geriatric experts and refined based on older adults' usability testing.

Our long-term goal is to use SAR-VR systems to address apathy among older LTC residents. As an initial step, we needed to create SAR-VR activities that combined cognitive, physical, and social aspects and encouraged HHI during the activities. Given the range of physical and cognitive impairments in this population, we deemed it essential to involve older LTC adults in the development of these activities to enhance the functionality, usability, and likelihood of promoting their use [35,36]. However, older LTC adults are often absent during the

XSL•FO
RenderX

design of interactive health technologies [37]. To enhance the likelihood that the designed SAR-VR activities could effectively engage pairs of older LTC adults with varying levels of physical and cognitive function, we used a user-centered design (UCD) approach.

UCD, variously termed as human-centered design or participatory design, is a fundamental engineering philosophy in the design and development of products. It is a structured, iterative method that focuses on the end user throughout the entire design and development [38], but little information is available about the practical implementation of UCD methods with older adults with cognitive impairment [39]. Challenges to UCD implementation with this population include not only the characteristics of the older adults but also organizational and staff characteristics.

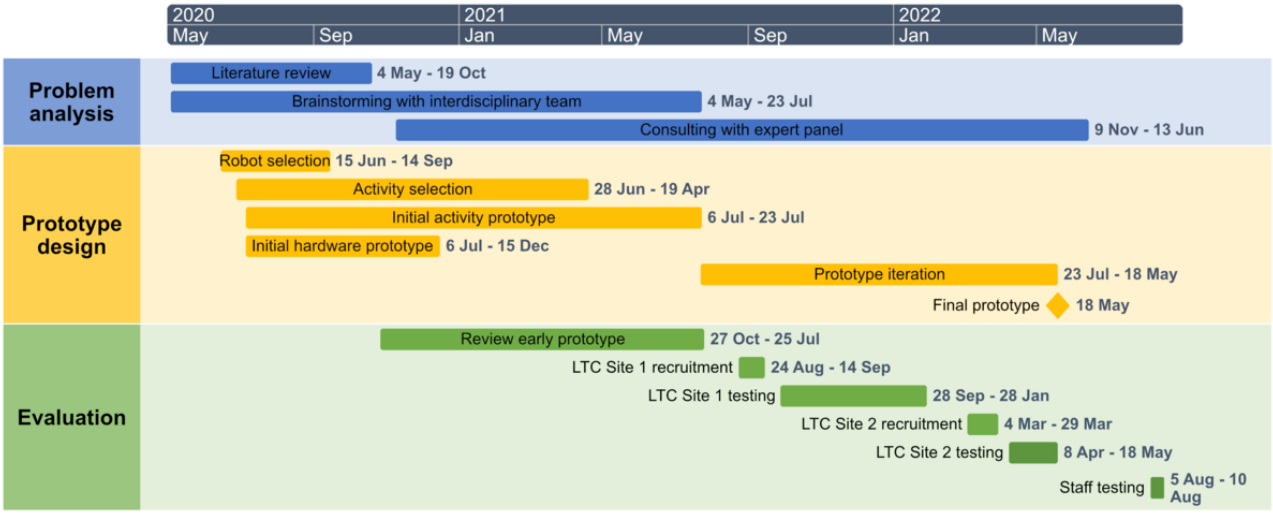
Our aim was to design and evaluate SAR-VR activities for pairs of older adults that would encourage HHI, an essential activity to mitigate apathy. To address gaps in knowledge and facilitate future work involving SAR-VR with this vulnerable population, we present our experiences in conducting UCD with older adults with cognitive impairment and characterize the challenges encountered in developing the SAR-VR system and paired activities.

Methods

User-Centered Study Design of the SAR-VR System

We implemented a multistep UCD process to design and evaluate SAR-VR activities that would encourage HHI mediated through human-computer interaction (HCI; nonimmersive VR) and human-robot interaction (HRI). Three overarching UCD principles guided the process throughout: (1) focus on users and tasks, (2) measure usability empirically, and (3) design and test usability iteratively [40,41]. Figure 1 displays the UCD timeline. There were distinct phases and activities to the UCD study process that are detailed below. These included problem analysis, prototype design development and evaluation, and field testing and validation.

Figure 1. Timeline for user-centered study process. LTC: long-term care.



Participant Recruitment

Participants consisted of older adults residing in an assisted living facility and staff. Eligibility criteria for older adults were age 65 years and older, at least 3 months of residence at the facility, the ability to hear, speak, and understand English, and the ability to sit comfortably and move both arms. Staff criteria included involvement directly or indirectly with the activities offered to residents within the setting and willingness to work with the researchers on the SAR-VR system. Investigators contacted several assisted living facilities by phone and email. Facilities needed to be within an hour's drive of the Vanderbilt University School of Engineering. Using institutional review board (IRB)–approved script, we explained the study purpose and our needs to conduct the study, that is, physical space adequate for the robot and nonimmersive VR setup, physical space conducive to the activity and evaluation (eg, adequate lighting and minimal noise or interruptions), and willingness to allow staff to interact with the research staff. Upon receiving permission from site administrators, Vanderbilt University engineers spoke to potentially interested residents at a resident meeting and to staff at a staff meeting, demonstrated the robot, and provided information regarding the study needs. Interested residents and staff contacted the research team directly, and a meeting was arranged to determine eligibility and conduct the informed consent process.

Phase 1: Problem Analysis

The first step was problem analysis, that is, to specify the context of use of SAR-VR activities with pairs of frail older adults. First, our interdisciplinary team brought expertise to the area of inquiry, including engineers from various disciplines (ie, mechanical engineering, electrical engineering, and computer science) as well as nurses and physicians who specialized in geriatrics and LTC. Based on our areas of expertise, we searched PubMed to (1) identify aspects of activities that effectively engage older LTC adults, especially those who have apathy, (2) SAR characteristics acceptable to older adults, and (3) nonimmersive VR factors suitable to older adults.

Results from our literature review were discussed with an interdisciplinary advisory panel (n=8), consisting of LTC

XSL•FO
RenderX

activity directors, nurses, managers, and a doctoral-prepared occupational therapist who specialized in geriatrics. Based on the literature review and conversations with the advisory panel and older adults, inspiration for activities was based on real-life activities the older adults performed at their LTCs and hobbies they enjoyed. Criteria for activities included (1) flexibility to accommodate older adults with varying levels of physical and cognitive abilities, (2) varying levels of collaboration required between 2 older adults, and (3) sufficient variation of activities to minimize monotony. Lastly, the system must have a user-friendly graphical user interface so that LTC staff can easily operate it and older adults can manipulate it. We conducted multiple design iterations and internal testing of prototype activities in the Vanderbilt University Engineering Laboratory. We presented ongoing work over several meetings with the advisory panel. The 2 LTC residents reviewed and approved the initial prototype activities for field testing via teleconferencing (due to COVID-19 isolation protocols).

Phase 2: Early Prototype Design

This phase focused on the choice of robots, software, and hardware development, and the choice of activities for each robot type.

Choice of Robots

In our earlier work [34], we found that older adults varied in their responsiveness and enjoyment to 2 types of SARs, humanoid and animal. For this study, we used NAO from SoftBank Robotics as our humanoid robot [42]. NAO's plastic body is child-sized at 58 cm and has 25 degrees of freedom with sensors ranging from tactile to audio microphone arrays. It has been used in numerous clinical and research settings, given its ease of use, sufficient expressive power for tasks, and its open architecture, allowing custom intelligent software development. We chose NAO because its limitations are not detrimental to our goals, its open architecture is capable of our specific modifications toward closed-loop intelligent interaction, and it is not prohibitively expensive. In our prior work, we augmented NAO's vision with a Kinect for Windows sensor (Microsoft Corp) for gesture recognition and gaze estimation using our prior developed algorithms [34,43]. We also designed a rule-based finite state machine method to recognize physical gestures (Multimedia Appendix 1).

We chose Aibo from SONY [44] as our dog robot. Aibo has been in the field of AI robotics for over 20 years. Aibo is designed to learn and develop a unique personality based on interactions with its owners and environment, has an open architecture allowing for custom software development, and has an array of sensors to detect colors, sound, touch, and movement. It has multiple degrees of freedom to allow for realistic and complex movements. Similar to NAO, Aibo has been used in research testing HRIs as well as for medical and therapeutic uses.

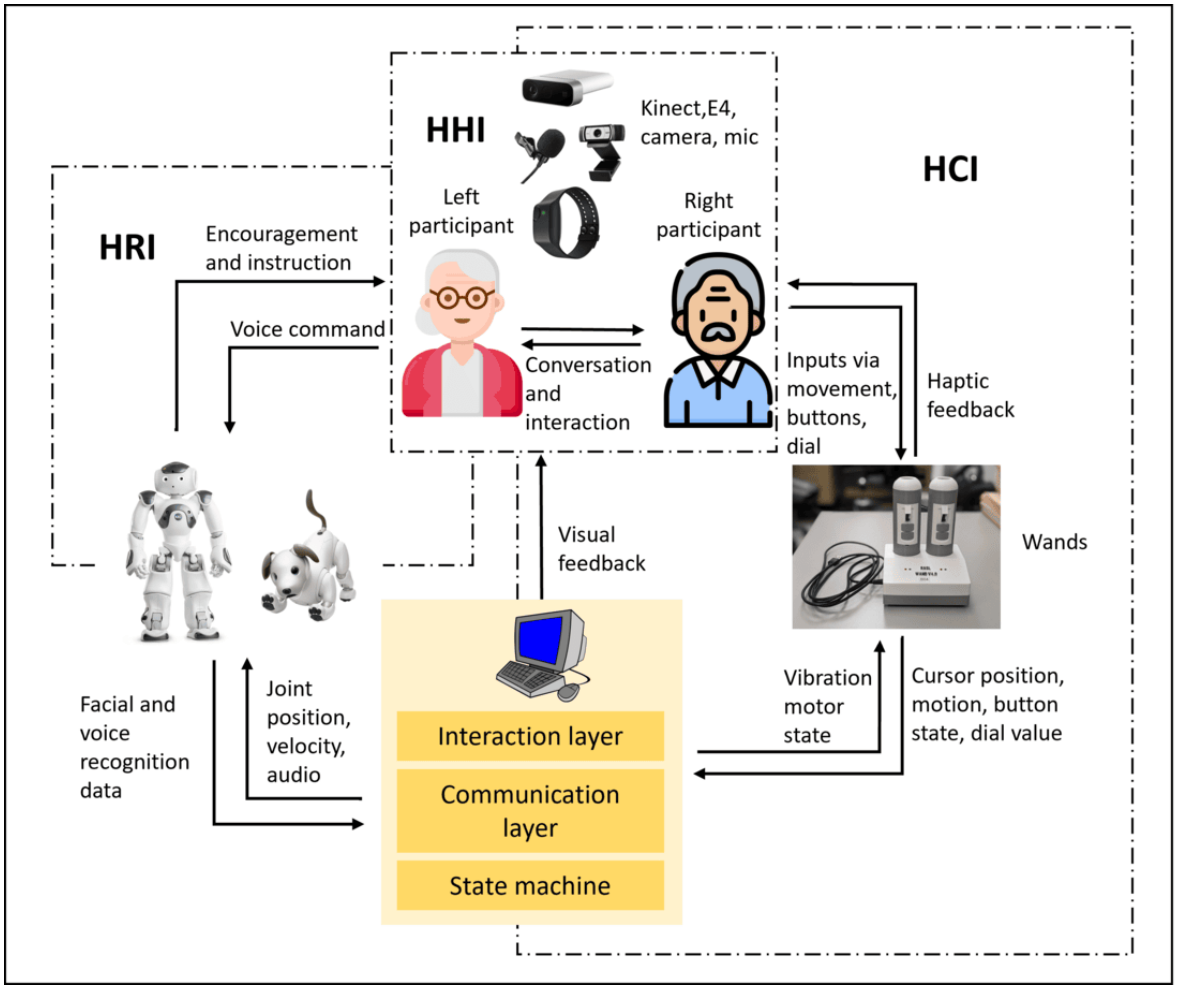
Software and Hardware Development

The long-term goal of the SAR-VR activities is to increase social engagement, defined as a person's activities performed within their social environment [45], to mitigate the effects of apathy in older adults. For this user-centered design study, we were interested in 3 types of interactions that could enhance social engagement. First, HHI referred to the collaborative behaviors needed between the 2 older adults to successfully complete the activity. Next, HCI, which encompasses the interactional aspects between the human user and computer [46], focused on the visual environment of the VR, the sounds, and tactile (haptic) wand feedback systems to maintain the older adults' interest and engagement. Lastly, HRI refers to the technical aspects of the older adults and robots communicating and collaborating, as well as the psychological and social considerations of human responses and perceptions to robots [47].

Activities were designed specifically for use with a humanoid SAR and a dog SAR. The humanoid SAR acted as a coach and cheerleader, while a computer-based avatar acted as the coach and cheerleader for activities involving the dog SAR. The dog SAR then functioned as an entertainment reward. The SAR-VR engineering architecture, that is, the organized design and integration of the software and hardware system components, consisted of the 3 interactive modules—an HCI, an HRI, and an HHI module, as shown in Figure 2.

XSL•FO
RenderX

Figure 2. Architecture of socially assistive robot and nonimmersive virtual reality. HCI: human-computer interaction; HHI: human-to-human interaction; HRI: human-robot interaction.



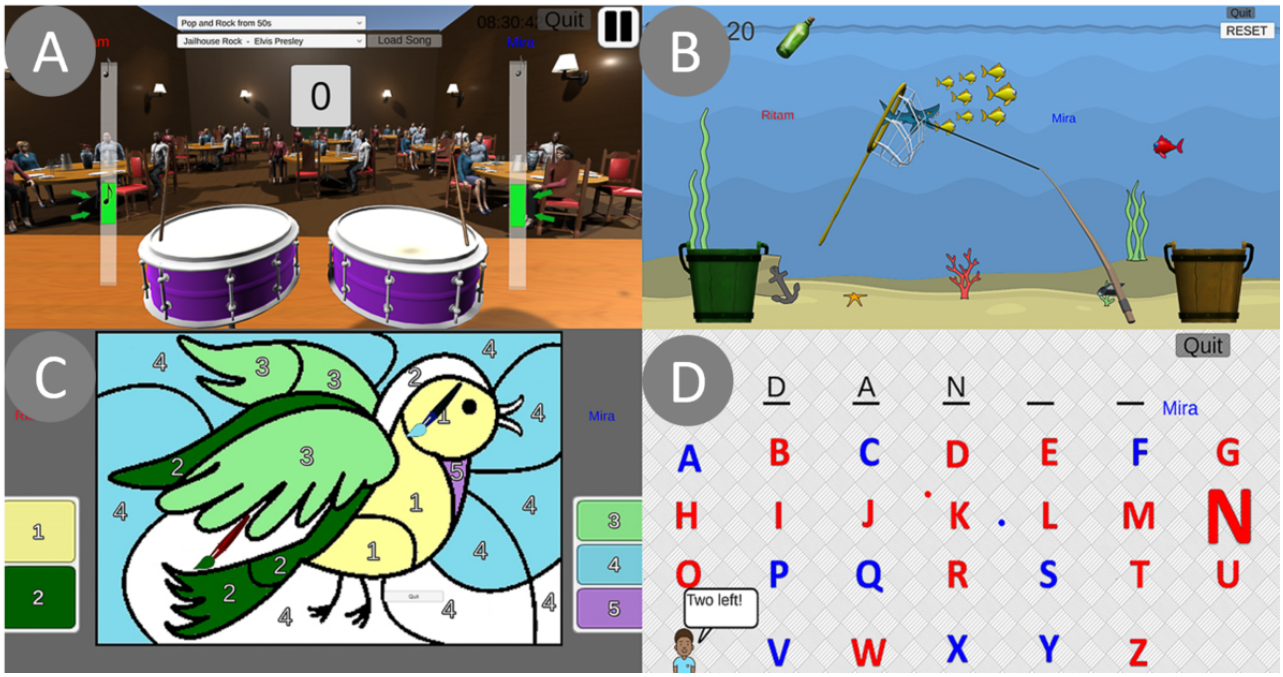
In brief, the HCI module was responsible for providing the participants with a virtual environment where they could perform the activities. The HRI module was responsible for using SARs to demonstrate the virtual activities to the participants, provide corrective feedback when necessary, encourage and motivate participants to keep them engaged, and provide rewards in the form of celebratory dance, tricks, etc, upon successful completion of the activities. The HHI module was deliberately designed to require the participants to communicate and collaborate to complete the activities since HHI is particularly effective in reducing apathy [9,11,12]. (Multimedia Appendix 1).

Initial Activity Choice and Rationale

Based on repeated brainstorming sessions and advisory panel meetings, inspiration for activities was taken from real-life activities that the older adults performed at the LTCs and hobbies they enjoyed. The interdisciplinary team determined potential virtual activities that (1) combined cognitive, physical, and social components; (2) could involve 2 older adults and encourage them to communicate and interact; and (3) were suitable for older adults with varying levels of cognitive and physical capabilities. We began development of 4 activities, 3 with Nao (music, fishing, and painting) and 1 with Aibo (spelling), as shown in Figure 3.

XSL•FO
RenderX

Figure 3. Screenshots of (A) music activity, (B) fishing activity, (C) painting activity, and (D) spelling activity.



Music can serve as an effective tool for reminiscence and autobiographical recall [48], improve cognitive function, and reduce depression [49]. The fishing and painting activities were chosen because many older adults enjoyed these activities in their youth. Research has also shown that creative activities, such as art and music, can lead to improved mental well-being, increased focus, and a sense of calmness [50]. The virtual spelling activity was chosen since word games, such as crossword puzzles, are often enjoyed by older adults in LTCs and have shown promising results in the preservation of cognitive function [51,52].

Activity Design Overview

Activities were structured to incorporate physical, cognitive, and collaborative aspects (Table 1). Each activity started with an interactive tutorial to familiarize participants with the objectives of the activity, provide the opportunity to practice using the wand, and get familiarized with interacting with NAO. The wand was the custom-made human interface device, similar to a Nintendo Wii remote controller, that was designed specifically with the older adult's physical function in mind [41,53] (for details on the wand, see Multimedia Appendix 1). NAO's verbal and physical gestures were informed by our earlier work that enhanced older adults' interaction with the robot. For example, at the start of each session, the engineer input the participants' names. NAO would wave and greet each participant by name and introduce itself as "Vandy from Vanderbilt." Ongoing positive reinforcement and feedback were directed to the individual. If one individual was having difficulty, NAO would ask the other participant to help them.

Table 1. Initial prototype activities, skills required, and socially assistive robot (SAR) input.

	Music	Fishing	Painting	Spelling
Description	Drumming motion with wand to play virtual drums in sync with displayed musical notes	Catch fish with virtual fishing rods using wands. Deposit captured fish into buckets	Paint-by-numbers format. Each participant has been assigned numbers with corresponding colors on palettes	A word (represents a dog command) was provided to participants. Each participant is assigned a color (red or blue) and must choose the correct letters to complete the spelling
Cognitive skills	Perception and attentional control	Problem-solving, attention, and working memory	Recognition and attention	Recall
Physical skills	Gross motor movements	Gross and fine motor movements	Fine motor skills	Fine motor skills
Social skills	Joint decision-making and participation	Cooperation and coordination	Participation	Participation and cooperation
SAR Input	NAO provides tutorials, encouragement, and reminders	NAO provides tutorials, encouragement, and reminders	NAO provides tutorials, encouragement, and reminders	NAO provides tutorials and feedback. Aibo performs the word command

Each activity had levels of increasing difficulty to accommodate varying levels of cognitive abilities among the participants (Multimedia Appendix 2). All activities underwent numerous modifications as the field testing progressed based on (1) enhancement of visual and audio characteristics and (2) meeting

XSL•FO
RenderX

our goal of encouraging HHI as determined through objective and subjective feedback (Multimedia Appendix 2).

Phase 3: Evaluation and Field Testing

Usability

Usability requirements of the SAR-VR activities and system were based on the intended end users—LTC older adults and staff. Data sources included observation, video recordings, and participant interviews during and after each session. Field notes, using a predetermined format, were generated after each testing session and reviewed by the research team. Videos were reviewed with the full interdisciplinary team to allow for behavioral analytics and task analysis.

We evaluated usability heuristically in 3 general areas. Contextual issues related to the environmental factors that could impact usability (lighting, sound, space issues, and furniture). Technology issues encompassed software and hardware components that resulted in errors or delays and prevented the older adults from effectively completing the tasks, potentially resulting in frustration. Activity issues were those germane to the actual activity; open-ended questions were used to obtain their thoughts on the activity, including their enjoyment, understanding, and suggestions for changes.

Confidence and Comfort

We examined older adults' confidence and comfort with the system. We adapted items from the System Usability Scale (SUS) [54] since in our earlier studies, we found older adults with cognitive impairment had difficulty completing the original scale. We asked 6 structured questions to elicit their degree of confidence and degree of comfort in using the wand (2 items), interacting with the VR environment (2 items), and interacting with the robots (2 items). We modified the 5-point Likert response as a horizontal scale depicting 5 facial expressions in combination with text descriptions rather than the use of numbers, a valid self-report method for older adults who have mild to moderate cognitive impairment [55,56]. The modified scale ranged from 6 to 30, with higher numbers indicating greater comfort and confidence. We compared the first and last ratings using the paired Wilcoxon signed rank test.

Interaction Indices

Lastly, we examined interaction indices for HHI and HRI. Since apathy is best treated by HHI activities [11,12], we were interested in knowing the degree to which the designed SAR-VR activities encouraged HHI. At each session, we measured the rate of interactions as the number of discrete events per minute. Rates were examined over time at each session.

Iterative Testing and Evaluation of the SAR-VR Activities

We conducted 2 field tests in separate LTC sites using the intended deployment setup to enhance discovery of potential challenges. Activity development and testing were conducted at the first site, and validation and final refinement occurred at the second site. During each round of field testing, multiple sessions were conducted with testing and analysis that required looping back to earlier stages of the prototype development and defined requirements for cyclical refinement and improvement of both the activity and the system. The design iterations continued until users deemed that the activities and SAR-VR system met the requirements.

Older Adult Study Procedure

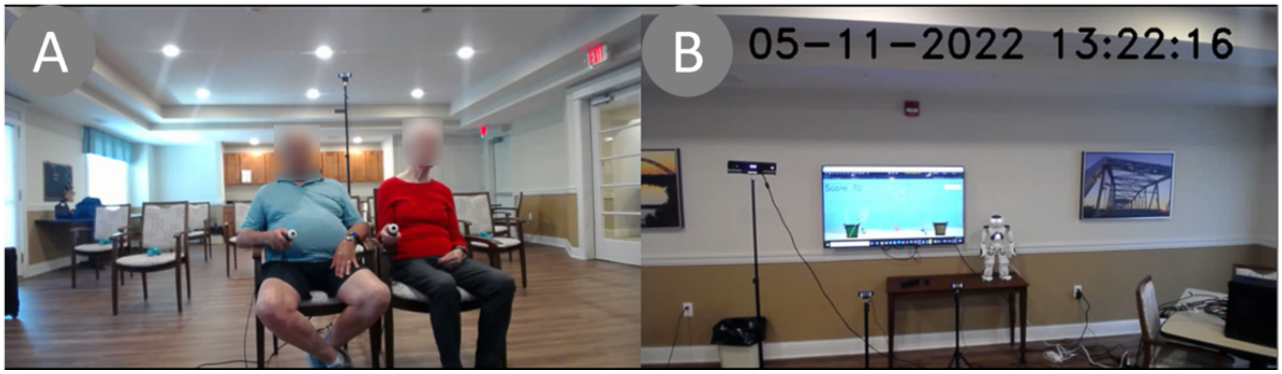
Prior to the first session, baseline information was gathered from each older adult, consisting of demographic information (age, sex, education level), frequency of use of technology (smartphones, computers, pads), the Apathy Evaluation Scale-Self Rated (AES-S), and the Self-Administered Gerocognitive Exam (SAGE). The AES-S [57] consists of 18 items that measure behavioral, cognitive, and emotional domains of apathy. Each item is rated on a 4-point scale ranging from 1 (not at all) to 4 (a lot), with a total score ranging from 18 (low apathy) to 72 (extreme apathy). Although there are no standard cutoff scores to indicate clinical apathy, a score of 40-42 on the AES-S generally indicates mild apathy in older populations. The AES-S has good internal consistency (Cronbach α=.86-.94), retest reliability (r=0.89-0.94), and convergent and divergent validity.

The SAGE [58] assesses cognition in 5 domains—language, reasoning and computation, visuospatial abilities, executive memory, and orientation. The SAGE has good interrater reliability (intraclass correlation coefficient=0.96), specificity (88%-95%), and sensitivity (62%-95%). A score of 17-22 (maximum) suggests normal cognition, 15-16 suggests mild cognitive impairment, and <15 suggests dementia.

The same study procedure was followed at all sessions at both sites. Figure 4 shows a setup, materials, and a session in progress. Each session was attended by 3-4 researchers. Participants sat in front of and facing the system. E4 physiological sensors (Empatica Inc) were placed on both participants' nondominant wrists. Calibration data and a 2-minute resting physiological signal were recorded for each participant. Each participant wore a clip-on microphone to record their speech. NAO was positioned in front and to the side of a large television screen. The Kinect was placed facing the 2 participants. An administrator operated the study workstation placed to the side of the participants. A total of 2 webcams were used to record the participants' behaviors and interactions, as well as the SAR and VR environment (Multimedia Appendix 3).

XSL•FO
RenderX

Figure 4. (A) Two participants performing the fishing activity (presented with permission from the participants) and (B) a virtual reality screen and NAO in front of the participants.



Each participant participated in 5 separate sessions lasting 30-45 minutes. Each session encompassed 2 activities with a 5- to 10-minute break between activities. At the start of each activity, participants completed an orientation where they were informed of the activity and objective and how to use the wands to map their movements onto the VR environment, and provided the opportunity to practice the basic motions needed for the activity. Most sessions were completed with an older adult pair; in the event of scheduling conflicts with older adults, some sessions were completed with one older adult and a research partner. Between each session, engineers modified activities and the SAR-VR system.

LTC Staff Study Procedure

Once the field testing was completed with the older adults, the final SAR-VR system was tested with LTC staff to determine whether they could administer the system. A total of 5 staff from Site 2 (mean age 34.2, SD 15.8 years) were recruited and consented to the study. Components of the system were set out; all components were labeled and color-coded. Staff were provided a written manual with a step-by-step guide in which components were illustrated (Multimedia Appendix 3). In addition, every section had a video guide. Staff participants were asked to set up and run the system based on the instructions. After each session, the system setup and interface were refined based on objective and subjective feedback.

Ethical Considerations

Approval for this study was obtained from the IRBs of the research team at The Ohio State University (IRB #2020B0342) and Vanderbilt University (IRB #210386). The procedures used in this study adhere to the tenets of the Declaration of Helsinki [59] and regulations of the IRBs. The 2 types of participants were recruited for this study—older adults residing in LTC and employed staff. Safeguards were in place for both groups to ensure participation was voluntary. Older adults were informed that their decision whether to participate would have no bearing on the care received at the LTC facility, and staff were informed that their decision regarding participation had no bearing on their employment. Written informed consent was obtained from all individuals participating in the study prior to data collection and study procedures. To ensure older adults understood the study purpose, risks, and benefits, we used the University of San Diego Brief Protection of Human Subjects Capacity to Consent instrument [60]. If a participant could not demonstrate understanding, they were ineligible. Research assistants underwent training to recognize signs of frustration, anxiety, or stress displayed by participants; any sign of discomfort resulted in termination of the session. Research assistants were trained in procedures to protect privacy and confidentiality. Participants were assigned unique study identification numbers, and all deidentified data were stored in a secure database and password-protected. Participants received financial compensation for study activities, including the interactive sessions and data collection procedures, to acknowledge their time and reduce attrition. An external safety officer, designated by the National Institute on Aging (NIA), reviewed and approved the study protocol and informed consent documents prior to the start of the study. Biannual reports were provided to both the safety officer and the NIA program officer to review progress and any untoward events.

Results

Older Adult Profile

A total of 12 older adults participated (Table 2). In the first field study, 6 older adults initially consented to participate. A 3-week delay occurred between consent and implementation due to a COVID-19 outbreak in the facility. Subsequently, 2 participants chose not to continue with the study, having lost interest over time; both these participants had significant cognitive impairment. The 4 remaining Site 1 participants (mean age 85, SD 9.3 years; mean AES-S score 29.3, SD 5; mean SAGE score 11.5, SD 3.6) completed all sessions. In Field Testing 2, eight residents (mean age 80, SD 4.7 years; mean AES-S score 26.35, SD 2.7; mean SAGE score 14.3, SD 4) were eligible, consented, and participated. Of the total sample, 4 participants had no experience with technology. Self-reported apathy scores were below the 40-cutoff point. Cognitive screening revealed that 3 (75%) of the 4 older adults at Site 1 had scores indicating dementia, whereas only 3 (38%) of 8 older adults at Site 2 had scores indicating dementia.

XSL•FO
RenderX

Table 2. Older adult characteristics for Field Testing 1 and 2.

Participant ID	Age (years)	Sex	Education	Technology use	AES-S[a] score	SAGE[b] score	Completed the study
Field Testing 1							
A1001	89	Female	High school	Everyday	25	11	Yes
A1002	90	Female	High school	Never use	32	17	Yes
A1003	92	Male	Graduate degree	Everyday	26	8	No
A1004	92	Male	Bachelor's degree	Everyday	24	11	Yes
A1005	69	Male	Graduate degree	Never use	36	7	Yes
A1006	89	Female	High school	Never use	25	5	No
Field Testing 2							
A1007	82	Male	Bachelor's degree	Everyday	26	18	Yes
A1008	81	Female	High school	Rarely	25	18	Yes
A1009	75	Female	Graduate degree	Everyday	28	18	Yes
A1010	80	Male	Bachelor's degree	2-3 times a week	31	11	Yes
A1011	80	Female	Graduate degree	Never use	29	6	Yes
A1012	72	Female	Trade school	Everyday	25	12	Yes
A1013	81	Female	Bachelor's degree	Everyday	22	16	Yes
A1014	89	Female	Graduate degree	Everyday	24	15	Yes

[a]AES-S: Apathy Evaluation Scale-Self Rated. Higher scores indicate greater apathy.

[b]SAGE: Self-Administered Gerocognitive Exam. A score of 17-22 (maximum) suggests normal cognition, 15-16 suggests mild cognitive impairment, and <15 suggests dementia.

Usability: Contextual, Technological, and Activity Issues

Attention to the environment was necessary. At the first site, older adults had difficulty seeing the monitor because of glare; subsequently, blinds needed to be drawn at all sessions. The voice of NAO was difficult for older adults to hear and understand clearly. After several trials of altering its pitch, style, volume, and speed, we found a setting that was the most clearly understood (Multimedia Appendix 1). Early on, we observed that older adults became fatigued with holding and manipulating the wands. They suggested having chairs with armrests and sitting at a table would provide them with the opportunity to rest their arms. The VR environment underwent a number of changes to enhance the visualization, including adjustments to the size, color, and placement of components.

Technology issues occurred in both the hardware and software components. Originally, NAO was programmed to do a short dance to celebrate the successful completion of an activity. After one of the NAO fell off the table and broke its hip, we elected to forgo that component despite the older adults' obvious enjoyment. NAO would give feedback based on the queuing of each observation of the older adults' performance, either positive or corrective. This resulted in a long list of feedback statements, ultimately causing a delayed response that did not correspond to the older adult's current performance. The mismatch of feedback with current behavior caused confusion and, at times, frustration. NAO was reprogrammed to provide feedback only after an older adult performed an errant behavior 3 times in a row. The customized wands were remodeled several times to enhance grip and controls; wands were also color-coded to match the person's cursor color in the VR environment.

Throughout the study, older adults provided suggestions to enhance the activities to make them more enjoyable. Table 3 highlights major changes that occurred during Field Testing 1 and 2.

XSL•FO
RenderX

Table 3. Changes to activities to enhance older adults' engagement.

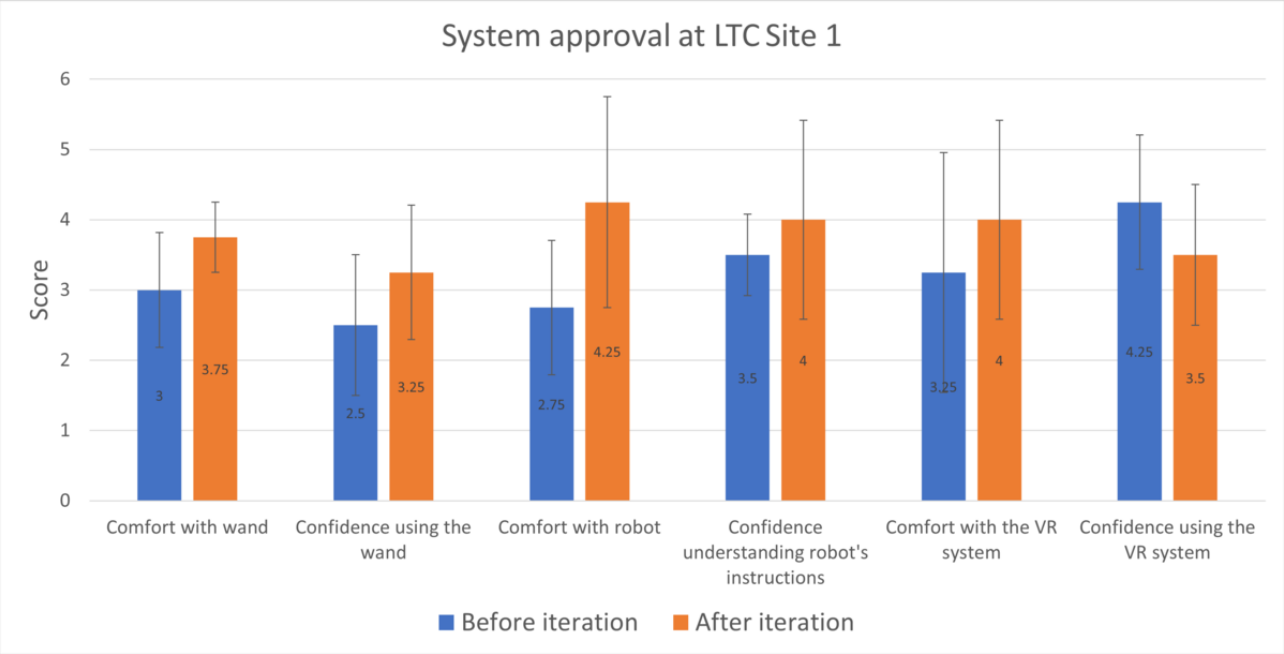
Component	Challenges	Changes
Field Testing 1		
Music activity	• Sometimes, only one participant got consecutive notes while the other had to wait for extended periods. • Sometimes NAO's consecutive feedback was too frequent and overwhelming.	• Note generation changed to "alternate" instead of "random." • Increased the length of the time between NAO's feedback statements.
Fishing activity	• Minimal collaboration was required between the 2 participants, who both had fishing rods.	• The initial activity of 2 fishing rods was replaced by 1 fishing rod and 1 net. The participant with the rod was responsible for catching a fish and passing it to the participant with the net. The participant with the net had to help guide the net to the fishing rod for the fish to be released, and then deposited the fish into the bucket. This increased collaboration and communication.
Painting activity	• Participants had difficulty seeing the color of the cursor "dot" and knowing which color was selected.	• The cursor changed from a "dot" to a "paintbrush," with the tip changing color to reflect the color that was selected.
Spelling activity	• Some of the participants did not realize it was their turn to pick the letters and would keep waiting. • Sometimes, the participants forgot what the target word was. • Some participants had difficulty locating the correct letters among the rows of letters.	• A reminder was added for NAO to cue participants when it was their turn to pick a letter. • A hint was added to the top of the screen to remind both participants of the word they are supposed to spell. • A slider was added that controls the number of excess letters visible on the screen. This allowed the administrator of the system to reduce visual stimuli overload and make it easier to locate the correct letter.
Field Testing 2		
Music activity	• Some participants were able to predict when it would be their turn to hit the drum and complained that it was boring.	• Alternate assigned notes were replaced with a probability approach to minimize predictability, yet ensuring each participant had sufficient notes to play.
Fishing activity	• Both buckets on the screen could accept fish, which minimized the amount of attention and movement by the participant with the net.	• Only 1 bucket at a time was active in the highest difficulty level. This required the participants to attend to which bucket was "active" during their play and provided the participants with the net more opportunities to play.
Painting activity	• Some participants had difficulty choosing a color from the palette, designed with 2 columns of colors.	• The palette layout changed to 1 column.
Avatar for dog activities	• Participants' attention was divided between the 2 robots.	• A screen avatar replaced NAO to provide instructions and feedback.

Comfort and Confidence Ratings

After each session, participants at both sites rated their degree of comfort and confidence in (1) using the wands, (2) interacting with the robot, and (3) interacting with the VR environment. Between the first and final sessions, participants at Site 1 increased their comfort and confidence in 5 of the 6 categories. The average increase across all 6 categories was 0.58 (95% CI 0.04-1.12; Cohen d=0.45; Wilcoxon signed rank test P=.03). The category-wise mean improvements and standard deviation for each item are displayed graphically in Figure 5 (see Multimedia Appendix 4 for numerical values).

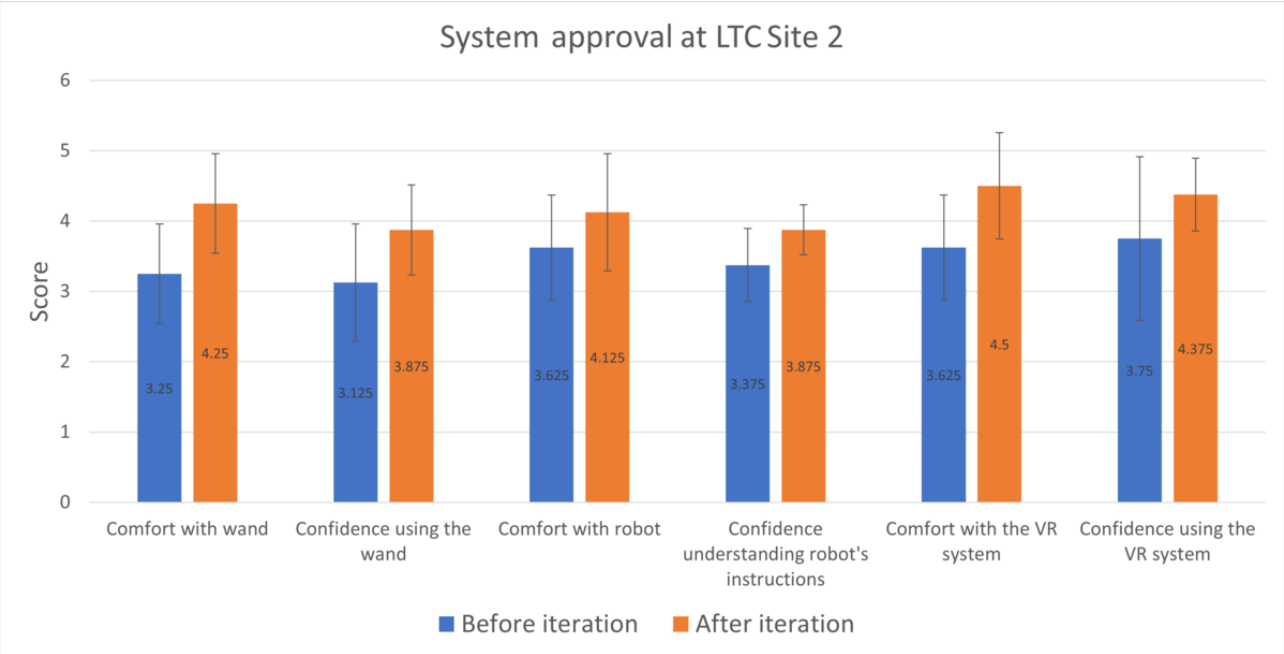
XSL•FO
RenderX

Figure 5. Site 1 older adults' rating of confidence and comfort with the socially assistive robot and nonimmersive virtual reality (VR) system elements from the first to last session. LTC: long-term care.



Participants at Site 2 showed enhanced comfort and confidence with an average increase of 0.71 (95% CI 0.43-0.99; Cohen *d*=0.73; Wilcoxon signed rank test *P*<.001) between the first and final sessions. The category-wise mean improvements and standard deviations are shown in Figure 6 (see Multimedia Appendix 4 for numerical values).

Figure 6. Site 1 older adults' rating of confidence and comfort with the socially assistive robot and nonimmersive virtual reality (VR) system elements from the first to last session. LTC: long-term care.



Interaction Indices: Participant-Participant, Participant-Robot, and Participant-Researcher

Analysis of the recorded videos revealed that the average interaction of older adults with their partner increased steadily across sessions, with iterations from 0.16 interactions per minute during the first set of sessions to 1.14 interactions per minute in the final set of sessions. The average number of robot interventions decreased from 1.10 per minute to 0.73 per minute. The average number of researcher-assisted interventions decreased from 0.34 per minute to 0.10 per minute. These results suggest an increased engagement and interaction among older adults and an increasing effectiveness of the system with each iteration.

Responses to Open-Ended Questions

Participants were asked open-ended questions (Multimedia Appendix 4) about their experience with the activities.

XSL•FO
RenderX

Participants reported they enjoyed the activities and expressed interest in playing them again. They described the games as fun and interesting, noting that once they became familiar with the mechanics (eg, casting the fishing rod), the experience was enjoyable and engaging. They also indicated that the robot's instructions were clear and sufficient, requiring no additional guidance.

Staff Setup of the SAR-VR System

A total of 5 staff members evaluated the system from the perspective of ease of use and convenience. They set up and ran the system following written instructions prepared by the researchers. The first 2 participants who worked with the system had difficulty following the steps in the correct order and required prompting from the research staff. To remedy this, the instruction manual was divided into 3 parts, that is, presession, during-session, and postsession. The presession consisted of initialization and calibration of components; during-session procedures involved running the system, activity selection, and changing the activity difficulty levels; and postsession procedures consisted of securely closing all applications and sensors, saving the data, and turning off the system. This division helped avoid confusion and led to a smoother setup by other participants. All the staff participants commented that the first-time setup had a learning curve, but they would feel more confident if they were to set up the system again. Overall, they found the integration of the different components of the system to be satisfactory and user-friendly.

Challenges in the UCD Process

Accessing Older Adult Participants

A major challenge to the UCD process was the ongoing COVID-19 pandemic. The study began in May 2020, at the start of COVID-19. Universities and LTC sites were on lockdown for months. The health care system was severely strained, delaying our access to experts from LTC sites. We were unable to conduct initial prototype testing in the Vanderbilt University Engineering Lab as initially planned, since older adults were kept in social isolation. The ongoing UCD was interrupted several times as COVID-19 waves continued to occur from 2020 to 2022, restricting access to LTC sites and older adults.

Working With Older Adults in LTC Sites

Most older adults had cognitive deficits that at times impeded their ability to interact with the system; by offering various levels of difficulty, we found that all participants could complete each activity. For some older adults, physical impairments, such as arthritis, made it difficult to handle the input device, and fatigue occurred easily. Subsequent changes to the wand facilitated grip and controls. Keeping activities to less than 20 minutes and the use of furniture supports eased the onset of fatigue. Visual impairments impeded their ability to see the VR environment. Ongoing adjustments to the color, size, and spacing of the elements resolved this issue. Hearing impairments impeded their ability to hear NAO's voice, and multiple changes were made until we were able to find a satisfactory pitch and volume. (see Multimedia Appendices 1 and 2 for refinements based on aging changes).

Working Within the LTC and With Staff

Logistics of the UCD proved challenging at times. Because we needed to take the equipment to the LTC, versus having the older adults come to the engineering laboratory, we needed to coordinate our schedules. We found that scheduling the sessions was complicated by competing priorities within the LTC, staff illnesses and shortages, and competing use of the activity room. The physical setup took additional time due to internet issues (that had not been an issue at the engineering laboratory).

Physiology and Motion Data for Future Research

Multimodal physiological and behavioral data were collected during the sessions to inform future research and system development. Physiological signals, including blood volume pulse, electrodermal activity, skin temperature, and accelerometer data, were recorded using the E4 sensor. These data can be analyzed to infer participants' affective states (eg, stressed vs relaxed, engaged vs disengaged) by examining the fluctuations in their autonomic responses during interaction with the system. Such insights can guide the design of adaptive features that adjust task difficulty, robot behavior, or feedback strategies in real time to better match the user's emotional and cognitive state.

Joint position and head orientation data were captured using the Kinect sensor. These data enable the quantification of motor performance (eg, range of motion, smoothness of movement) and the assessment of social engagement (eg, frequency of head turns toward the partner or robot, use of gestures, changes in posture). This allows for objective evaluation of both physical rehabilitation potential and the quality of social interaction.

Finally, time-stamped performance logs provide a record of task success, errors, and completion times, which can be correlated with physiological and kinematic measures. Together, these multimodal datasets will allow for richer models of user engagement, support longitudinal monitoring of progress, and inform further refinement of SAR-VR interventions for older adults.

Discussion

Overview

In this user-centered design study, we used the literature, our clinical expertise and prior research, and an advisory panel that included older adults to identify activities that could be created for older adults residing in LTC settings using a combined SAR and nonimmersive VR system (SAR-VR). The SAR-VR system was designed to engage pairs of older adults with the long-term goal of using this technology in LTC settings for those with apathy. The underlying premise is that those with apathy are best treated by activities that are performed with others and that combine cognitive and physical aspects [11,12]. We conducted numerous sessions with LTC older adults, who had varying levels of cognitive and physical impairments, to achieve the development and validation of 4 SAR-VR activities for this specific population. We are currently testing the effectiveness of the SAR-VR system to mitigate apathy in a multisite randomized controlled trial.

XSL•FO
RenderX

Although others [61,62] have highlighted the importance of UCD with older adults and adapting processes for this population, few studies have actually involved older adults with cognitive impairment in the design of technology [34,39]. Second, it is important to involve staff in UCD such that they feel comfortable operating new systems and thus ensure system scalability in LTC settings. Thus, we focused on the 2 key stakeholders, LTC older adults and LTC staff, in our UCD of the SAR-VR system and activities. Our findings demonstrated the effectiveness of an iterative UCD approach for designing SAR-VR activities and systems for older adults in LTC settings.

The study provided valuable lessons about conducting UCD of SAR-VR systems in LTC settings. First, the timeline for development and testing was prolonged due to multiple waves of COVID-19, virtually shutting down access to the older adults. Second, as is often the case in the design of technology, our initial design decisions were based on our expertise and experience, the literature, and experts in the LTC field. Some of these decisions were discovered to be incorrect during field testing. For example, we designed the graphics to be seen easily by the older adults (contrasting colors, font size) and lowered the pitch of NAO's voice to accommodate age-related vision and hearing changes [63]. However, we found our graphics required further modifications, and a series of voices had to be tested before older adults agreed to the pitch, tone, and speed. The initial creation of the handheld wand controller also required several changes to accommodate older adults' grips, maneuverability, and vision. Third, accessing older adults in the LTC sites proved more challenging than anticipated. It was not uncommon for staff to cancel and reschedule our sessions based on their availability, competing demands, or illness among the older adults—all of which further prolonged the process.

Despite these challenges, we found the UCD process successful in developing the SAR-VR system to deliver multimodal activities for older adults with cognitive and physical impairments. With the increasing population of older adults with dementia and insufficient skilled care staff, technological interventions are a promising strategy in the delivery of care processes. We succeeded in developing an innovative SAR-VR system that could deliver a variety of activities that combined physical, cognitive, and social domains for pairs of older adults residing in LTCs. Such multidomain stimuli, along with HHI, have been shown to be effective in reducing apathy [9,11]. Each activity required the participants to cooperate and make joint decisions, thus encouraging interaction and conversations. While other work has attempted to reduce apathy using immersive VR technology [64,65], research on developing technology to foster and mediate natural HHI among older adults with minimal caregiver intervention is scarce. Out of the total 14 participants recruited for the study, 12 stayed for the full duration of the study. Participants' confidence and comfort with the SAR-VR improved over time for both those who conducted the early development work (Field Testing 1) and the validation work (Field Testing 2); overall, the participants enjoyed doing the activities.

Since the system is designed to be deployed at LTCs, we tested the setup procedure of the system with LTC staff to ensure ease of use. All enrolled staff could set up and run the system with the instructions provided without any special training, thus proving the practicality and deployable nature of the system in an LTC. When asked if they would be more comfortable if they were to set up the system again, most replied in the affirmative and mentioned that the most challenging part was identifying components that they were otherwise unfamiliar with. The iterative design process allowed us to make modifications to the various components of the system, including the setup guide, until the system was suitable and enjoyable to both target populations, the older adults and the LTC staff.

While UCD has been used to develop immersive VR [39], eHealth applications [66,67], exercise games [68-70], and websites [71,72] for older adults, to the best of our knowledge, this is the first study to take an iterative UCD approach to developing a SAR-VR system and activities for pairs of older adults. The resultant modular architecture and activities set the foundation for investigating their effectiveness in mitigating apathy among older adults with dementia. We designed the system to easily be adapted to other robots, activities, and interaction devices. The intensity of the physical, cognitive, and social components of the activities can be individually tuned, as desired, to study the effect of each component.

Physiological, motion, and performance data collected from the study will provide insights for further analysis and refinement. Physiological data can provide insights into the level of stress and cognitive load of the participants during performance of the activities, which can later be used to design an affective state–based closed-loop system to automatically regulate the difficulty of the activities. The joint position data from the Kinect sensor can later help automate the detection of nonverbal communication by tracking head pose and gestures.

Limitations

This work, although promising, is not without limitations. First, it is limited by its small sample size and purposive sampling, as well as the short duration of the study procedures, that is, 5 sessions. Small sample sizes are typical in initial UCD studies given the need to facilitate rapid iterative testing, as well as keep personnel and financial resources within reasonable limits [38]. A larger study of longer duration is required to gauge user perception accurately and mitigate the novelty factor. Our target population is older adults who reside in LTC, with apathy, and with cognitive or physical impairments; our sample did not have evidence of clinically significant apathy, and several were screened with normal cognition. The user perception of the system depends on several factors, such as cognitive abilities, previous exposure, and familiarity with similar technology, as well as personal preferences toward particular activities. A larger sample size is required to generalize the preferences, as data from smaller samples are easily skewed by outlier individuals. A larger sample size would also allow for the identification of personal characteristics associated with successful and ongoing involvement with the SAR-VR system. Second, we used commercially available robots, thus limited by the physical attributes of the robots. For example, we found that we needed multiple adjustments to the humanoid robot's voice that would facilitate older adults' ability to hear and understand its language. We had to position the humanoid robot onto a table

XSL•FO
RenderX

to facilitate the visual field of the older adult who was focusing on the virtual environment displayed on a television or monitor. Nevertheless, we chose robots that have been available for many years to researchers because of their ease and versatility for customized programming. Lastly, we created only 4 activities during this study. We have since designed several more activities based on our experiences with this study that will allow for more variety in future studies.

Conclusions

Building upon our previous work on SARs and VR, this work details the use of user-centered iterative design techniques and processes to develop an innovative SAR-VR system that can provide activities that consist of physical, cognitive, and social aspects and are specifically designed to promote HHI through activities that require cooperation between participants. We created 4 activities—3 humanoid robot–based activities, where the humanoid robot NAO acted as the coach to provide hints and corrective feedback and mediate the interaction between older adults, and 1 animal robot–based activity, where an avatar in the VR environment provided the feedback and the puppy robot Aibo performed tricks as rewards to incentivize the older adults to complete the activity. A finite state machine was developed to provide appropriate real-time feedback in response to participants' performance in the activities. The system also collected video, physiology, joint position, and performance data that will be used for future investigations into how stress and engagement patterns in older adults varied with activity difficulty level. We incorporated various tunable difficulty levels in the activities to accommodate different physical and cognitive abilities. A variety of activities using both the humanoid and animal robots were developed to enhance engagement and reduce monotony. The system has a simple user interface and an easy-to-follow setup guide for the LTC staff. Currently, we are conducting a multisite randomized controlled trial [73] to compare the effect of usual care to the SAR-VR plus usual care on reducing apathy (primary outcome) among 188 older adults with cognitive impairment and enhancing executive cognitive function (secondary outcome). Eligibility of state-licensed LTCs includes an hour's drive from the principal investigators' offices, certification by the US Centers for Medicare and Medicaid, and physical space to store the research equipment. Eligible older adults are those aged 65 years or older, with 3 months or more of residence at the LTC, evidence of mild to moderate cognitive impairment, and symptoms of apathy. The intervention takes place twice weekly for 8 weeks; 4 weeks participating in NAO activities and 4 weeks participating in Aibo activities. Data are collected at baseline, each session, and post experiment. Findings will aid in determining the effectiveness of the SAR-VR system and identifying characteristics of the system, activities, and older adults that are associated with effectiveness. We are also gathering information to identify barriers and facilitators to implementation; this knowledge is needed to address the scalability of the intervention, including the training needs of staff.

Acknowledgments

The authors wish to thank Kelley Colopietro and Emily Latshaw for assisting with research activities and the 2 long-term care settings that graciously allowed us access to their environments.

Funding

The study was supported by the National Institute on Aging of the National Institutes of Health under award number R01AG062685. All authors are supported by National Institutes of Health grant funding. MM and NK received financial support from the National Science Foundation.

Authors' Contributions

Conceptualization: NS, LCM, JAT, PAN, DWS
Data curation: AT, RG, NK, MM
Formal analysis: RG, NK, MM, JAT, CAM, PAN, DWS, AT, LCM, NS
Funding acquisition: NS, LCM
Investigation: RG, NK, MM, LCM, JAT, NS, CAM
Methodology: LCM, JAT, NS, CAM
Project administration: NS, JAT
Resources: NS
Software: RG, MM, NK, NS
Supervision: NS
Visualization: RG, MM, NK
Writing—original draft: RG, MM, NK
Writing—review & editing: RG, MM, NK, NS, LCM, CAM, DWS, AT, PAN

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Details of the socially assistive robot with nonimmersive virtual reality system architecture.
[DOCX File , 2413 KB - aging_v8i1e75288_app1.docx]

Multimedia Appendix 2
Description of socially assistive robot with nonimmersive virtual reality activities.
[DOCX File , 765 KB - aging_v8i1e75288_app2.docx]

Multimedia Appendix 3
Study setup guide.
[DOCX File , 612 KB - aging_v8i1e75288_app3.docx]

Multimedia Appendix 4
Questionnaires and results.
[DOCX File , 130 KB - aging_v8i1e75288_app4.docx]

References

1. Sengupta M, Lendon JP, Caffrey C, Melekin A, Singh P. Post-acute and long-term care providers and services users in the United States, 2017–2018. CDC Stacks. 2022. URL: https://stacks.cdc.gov/view/cdc/115346 [accessed 2025-11-23]
2. Johnson RW. What is the lifetime risk of needing and receiving long-term services and supports? US Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation. 2019. URL: https://aspe.hhs.gov/reports/what-lifetime-risk-needing-receiving-long-term-services-supports-0 [accessed 2025-10-31]
3. Lanctôt KL, Ismail Z, Bawa KK, Cummings JL, Husain M, Mortby ME, et al. Distinguishing apathy from depression: a review differentiating the behavioral, neuroanatomic, and treatment-related aspects of apathy from depression in neurocognitive disorders. Int J Geriatr Psychiatry 2023;38(2):e5882 [FREE Full text] [doi: 10.1002/gps.5882] [Medline: 36739588]
4. 2024 Alzheimer's disease facts and figures. Alzheimer’s Association. 2024. URL: https://www.alz.org/getmedia/76e51bb6-c003-4d84-8019-e0779d8c4e8d/alzheimers-facts-and-figures.pdf [accessed 2025-10-31]
5. Lanctôt KL, Agüera-Ortiz L, Brodaty H, Francis PT, Geda YE, Ismail Z, et al. Apathy associated with neurocognitive disorders: recent progress and future directions. Alzheimers Dement 2017;13(1):84-100. [doi: 10.1016/j.jalz.2016.05.008] [Medline: 27362291]
6. Grossman HT, Sano M, Aloysi A, Elder GA, Neugroschl J, Schimming C, et al. Prevalent, persistent, and impairing: longitudinal course and impact of apathy in Alzheimer's disease. Alzheimers Dement (Amst) 2021;13(1):e12169 [FREE Full text] [doi: 10.1002/dad2.12169] [Medline: 35141391]
7. Tierney SM, Woods SP, Weinborn M, Bucks RS. Real-world implications of apathy among older adults: independent associations with activities of daily living and quality of life. J Clin Exp Neuropsychol 2018;40(9):895-903 [FREE Full text] [doi: 10.1080/13803395.2018.1444736] [Medline: 29532707]
8. Ruthirakuhan MT, Herrmann N, Abraham EH, Chan S, Lanctôt KL. Pharmacological interventions for apathy in Alzheimer's disease. Cochrane Database Syst Rev 2018;5(5):CD012197 [FREE Full text] [doi: 10.1002/14651858.CD012197.pub2] [Medline: 29727467]
9. Cai Y, Li L, Xu C, Wang Z. The effectiveness of non-pharmacological interventions on apathy in patients with dementia: a systematic review of systematic reviews. Worldviews Evid Based Nurs 2020;17(4):311-318. [doi: 10.1111/wvn.12459] [Medline: 32767834]
10. Li H, Wang H, Lu C, Chen T, Lin Y, Lee I. The effect of music therapy on reducing depression in people with dementia: a systematic review and meta-analysis. Geriatr Nurs 2019;40(5):510-516. [doi: 10.1016/j.gerinurse.2019.03.017] [Medline: 31056209]
11. Brodaty H, Burns K. Nonpharmacological management of apathy in dementia: a systematic review. Am J Geriatr Psychiatry 2012;20(7):549-564. [doi: 10.1097/JGP.0b013e31822be242] [Medline: 21860324]
12. Cohen-Mansfield J, Marx MS, Dakheel-Ali M, Thein K. The use and utility of specific nonpharmacological interventions for behavioral symptoms in dementia: an exploratory study. Am J Geriatr Psychiatry 2015;23(2):160-170 [FREE Full text] [doi: 10.1016/j.jagp.2014.06.006] [Medline: 25081819]
13. Ellis JM, Doyle CJ, Selvarajah S. The relationship between apathy and participation in therapeutic activities in nursing home residents with dementia: evidence for an association and directions for further research. Dementia (London) 2016;15(4):494-509. [doi: 10.1177/1471301214527300] [Medline: 24670286]
14. Historic staffing shortages continue to force nursing homes to limit new admissions, creating bottlenecks at hospitals and reducing access to care for seniors. American Health Care Association. 2022. URL: https://www.ahcancal.org/News-and-Communications/Press-Releases/Pages/

XSL•FO
RenderX

Historic-Staffing-Shortages-Continue-To-Force-Nursing-Homes-To-Limit-New-Admissions,-Creating-Bottlenecks-at-Hospitals-and-.aspx [accessed 2025-10-31]
15. Gibson DM, Greene J. State actions and shortages of personal protective equipment and staff in US nursing homes. J Am Geriatr Soc 2020;68(12):2721-2726 [FREE Full text] [doi: 10.1111/jgs.16883] [Medline: 33022757]
16. Keyworth C, Hart J, Armitage CJ, Tully MP. What maximizes the effectiveness and implementation of technology-based interventions to support healthcare professional practice? A systematic literature review. BMC Med Inform Decis Mak 2018;18(1):93 [FREE Full text] [doi: 10.1186/s12911-018-0661-3] [Medline: 30404638]
17. Meiland F, Innes A, Mountain G, Robinson L, van der Roest H, García-Casal JA, et al. Technologies to support community-dwelling persons with dementia: a position paper on issues regarding development, usability, effectiveness and cost-effectiveness, deployment, and ethics. JMIR Rehabil Assist Technol 2017;4(1):e1 [FREE Full text] [doi: 10.2196/rehab.6376] [Medline: 28582262]
18. Moyle W. The promise of technology in the future of dementia care. Nat Rev Neurol 2019;15(6):353-359. [doi: 10.1038/s41582-019-0188-y] [Medline: 31073242]
19. Appel L, Ali S, Narag T, Mozeson K, Pasat Z, Orchanian-Cheff A, et al. Virtual reality to promote wellbeing in persons with dementia: a scoping review. J Rehabil Assist Technol Eng 2021;8:20556683211053952 [FREE Full text] [doi: 10.1177/20556683211053952] [Medline: 35024166]
20. Molina KI, Ricci NA, de Moraes SA, Perracini MR. Virtual reality using games for improving physical functioning in older adults: a systematic review. J Neuroeng Rehabil 2014;11:156 [FREE Full text] [doi: 10.1186/1743-0003-11-156] [Medline: 25399408]
21. Lee L, Kim MJ, Hwang WJ. Potential of augmented reality and virtual reality technologies to promote wellbeing in older adults. Appl Sci 2019;9(17):3556 [FREE Full text] [doi: 10.3390/app9173556]
22. Yu C, Sommerlad A, Sakure L, Livingston G. Socially assistive robots for people with dementia: systematic review and meta-analysis of feasibility, acceptability and the effect on cognition, neuropsychiatric symptoms and quality of life. Ageing Res Rev 2022;78:101633 [FREE Full text] [doi: 10.1016/j.arr.2022.101633] [Medline: 35462001]
23. McColl D, Louie W, Nejat G. Brian 2.1: a socially assistive robot for the elderly and cognitively impaired. IEEE Robot Automat Mag 2013;20(1):74-83 [FREE Full text] [doi: 10.1109/mra.2012.2229939]
24. Appel L, Appel E, Bogler O, Wiseman M, Cohen L, Ein N, et al. Older adults with cognitive and/or physical impairments can benefit from immersive virtual reality experiences: a feasibility study. Front Med (Lausanne) 2019;6:329 [FREE Full text] [doi: 10.3389/fmed.2019.00329] [Medline: 32010701]
25. Pan Y, Steed A. A comparison of avatar-, video-, and robot-mediated interaction on users' trust in expertise. Front Robot AI 2016;3 [FREE Full text] [doi: 10.3389/frobt.2016.00012]
26. Nishio T, Yoshikawa Y, Sakai K, Iio T, Chiba M, Asami T, et al. The effects of physically embodied multiple conversation robots on the elderly. Front Robot AI 2021;8:633045 [FREE Full text] [doi: 10.3389/frobt.2021.633045] [Medline: 33829042]
27. Koyama S, Tatemoto T, Kumazawa N, Tanabe S, Nakagawa Y, Otaka Y. The effect of differences in powered wheelchair joystick shape on subjective and objective operability. Appl Ergon 2023;107:103920 [FREE Full text] [doi: 10.1016/j.apergo.2022.103920] [Medline: 36306702]
28. Carment L, Abdellatif A, Lafuente-Lafuente C, Pariel S, Maier MA, Belmin J, et al. Manual dexterity and aging: a pilot study disentangling sensorimotor from cognitive decline. Front Neurol 2018;9:910 [FREE Full text] [doi: 10.3389/fneur.2018.00910] [Medline: 30420830]
29. Goble DJ, Coxon JP, Wenderoth N, Van Impe A, Swinnen SP. Proprioceptive sensibility in the elderly: degeneration, functional consequences and plastic-adaptive processes. Neurosci Biobehav Rev 2009;33(3):271-278. [doi: 10.1016/j.neubiorev.2008.08.012] [Medline: 18793668]
30. Mirelman A, Bonato P, Deutsch JE. Effects of training with a robot-virtual reality system compared with a robot alone on the gait of individuals after stroke. Stroke 2009;40(1):169-174. [doi: 10.1161/STROKEAHA.108.516328] [Medline: 18988916]
31. Kayabinar B, Alemdaroğlu-Gürbüz İ, Yilmaz Ö. The effects of virtual reality augmented robot-assisted gait training on dual-task performance and functional measures in chronic stroke: a randomized controlled single-blind trial. Eur J Phys Rehabil Med 2021;57(2):227-237 [FREE Full text] [doi: 10.23736/S1973-9087.21.06441-8] [Medline: 33541040]
32. Garcia-Vergara S, Brown L, Park H, Howard A. Engaging children in play therapy: the coupling of virtual reality games with social robotics. Stud Comput Intell 2014;536:139-163 [FREE Full text]
33. Li J, Kizilcec R, Bailenson J, Ju W. Social robots and virtual agents as lecturers for video instruction. Comput Hum Behav 2016;55(Part B):1222-1230 [FREE Full text] [doi: 10.1016/j.chb.2015.04.005]
34. Fan J, Mion LC, Beuscher L, Ullal A, Newhouse PA, Sarkar N. SAR-Connect: a socially assistive robotic system to support activity and social engagement of older adults. IEEE Trans Robot 2022;38(2):1250-1269 [FREE Full text] [doi: 10.1109/tro.2021.3092162] [Medline: 36204285]
35. Fischer B, Peine A, Östlund B. The importance of user involvement: a systematic review of involving older users in technology design. Gerontologist 2020;60(7):e513-e523 [FREE Full text] [doi: 10.1093/geront/gnz163] [Medline: 31773145]

XSL•FO
RenderX

36. De Vito Dabbs A, Myers BA, Mc Curry KR, Dunbar-Jacob J, Hawkins RP, Begey A, et al. User-centered design and interactive health technologies for patients. Comput Inform Nurs 2009;27(3):175-183 [FREE Full text] [doi: 10.1097/NCN.0b013e31819f7c7c] [Medline: 19411947]
37. Peute LW, Wildenbos G, Engelsma T, Lesselroth BJ, Lichtner V, Monkman H, et al. Overcoming challenges to inclusive user-based testing of health information technology with vulnerable older adults: recommendations from a human factors engineering expert inquiry. Yearb Med Inform 2022;31(1):74-81 [FREE Full text] [doi: 10.1055/s-0042-1742499] [Medline: 35654432]
38. Holtzblatt K, Wendell JB, Wood S. Rapid Contextual Design: A How-to Guide to Key Techniques for User-Centered Design. Amsterdam, Netherlands: Elsevier; 2004.
39. Rochon EA, Thacker A, Phillips M, Ritchie C, Vranceanu A, Plys E. Developing a dyadic immersive virtual environment technology intervention for persons living with dementia and their caregivers: multiphasic user-centered design study. JMIR Aging 2025;8:e66212 [FREE Full text] [doi: 10.2196/66212] [Medline: 40397932]
40. Gulliksen J, Göransson B, Boivie I, Blomkvist S, Persson J, Cajander Å. Key principles for user-centred systems design. Behav Inf Technol 2003;22(6):397-409 [FREE Full text] [doi: 10.1080/01449290310001624329]
41. Ghosh R, Khan N, Migovich M, Wilson D, Latshaw E, Tate J. Iterative user centered design of robot-mediated paired activities for older adults with mild cognitive impairment (MCI). In: Gao Q, Zhou J, editors. Human Aspects of IT for the Aged Population. Technology in Everyday Living. HCII 2022. Lecture Notes in Computer Science, Vol 13331. Cham, Switzerland: Springer, Cham; 2022.
42. NAO. Softbank Robotics. URL: https://us.softbankrobotics.com/nao [accessed 2025-10-31]
43. Fan J, Ullal A, Beuscher L, Mion LC, Newhouse P, Sarkar N. Field testing of Ro-Tri, a robot-mediated triadic interaction for older adults. Int J Soc Robot 2021;13(7):1711-1727 [FREE Full text] [doi: 10.1007/s12369-021-00760-2] [Medline: 33643494]
44. Aibo website. Sony Group Corportation. URL: https://us.aibo.com [accessed 2025-10-31]
45. Holt-Lunstad J, Robles TF, Sbarra DA. Advancing social connection as a public health priority in the United States. Am Psychol 2017;72(6):517-530 [FREE Full text] [doi: 10.1037/amp0000103] [Medline: 28880099]
46. Rapp A. Human–computer interaction. Oxford Research Encyclopedia of Psychology. 2023. URL: https://oxfordre.com/psychology/view/10.1093/acrefore/9780190236557.001.0001/acrefore-9780190236557-e-47 [accessed 2025-09-17]
47. Singh KJ, Kapoor DS, Sohi BS. All about human-robot interaction. In: Mittal M, Shar RR, Roy S, editors. Cognitive Computing for Human-Robot Interaction. Cambridge, Massachusetts: Academic Press; 2021:199-229.
48. Basaglia-Pappas S, Laterza M, Borg C, Richard-Mornas A, Favre E, Thomas-Antérion C. Exploration of verbal and non-verbal semantic knowledge and autobiographical memories starting from popular songs in Alzheimer's disease. Int Psychogeriatr 2013;25(5):785-795 [FREE Full text] [doi: 10.1017/S1041610212002359] [Medline: 23388499]
49. Moreno-Morales C, Calero R, Moreno-Morales P, Pintado C. Music therapy in the treatment of dementia: a systematic review and meta-analysis. Front Med (Lausanne) 2020;7:160 [FREE Full text] [doi: 10.3389/fmed.2020.00160] [Medline: 32509790]
50. Cantu AG, Fleuriet KJ. "Making the ordinary more extraordinary": exploring creativity as a health promotion practice among older adults in a community-based professionally taught arts program. J Holist Nurs 2018;36(2):123-133. [doi: 10.1177/0898010117697863] [Medline: 29172944]
51. Murphy M, O'Sullivan K, Kelleher KG. Daily crosswords improve verbal fluency: a brief intervention study. Int J Geriatr Psychiatry 2014;29(9):915-919. [doi: 10.1002/gps.4079] [Medline: 24493633]
52. Ching-Teng Y. Effect of board game activities on cognitive function improvement among older adults in adult day care centers. Soc Work Health Care 2019;58(9):825-838. [doi: 10.1080/00981389.2019.1656143] [Medline: 31432758]
53. Migovich M, Ghosh R, Khan N, Tate J, Mion L, Sarkar N. System architecture and user interface design for a human-machine interaction system for dementia intervention. 2021 Presented at: Human Aspects of IT for the Aged Population. Supporting Everyday Life Activities: 7th International Conference; 2021 July 24-29; Berlin, Heidelberg p. 277-292.
54. Brooke J. SUS: a quick and dirty usability scale. In: Usability Evaluation in Industry. Boca Raton, FL: CRC Press; 1996.
55. Herr KA, Mobily PR, Kohout FJ, Wagenaar D. Evaluation of the faces pain scale for use with the elderly. Clin J Pain 1998;14(1):29-38. [doi: 10.1097/00002508-199803000-00005] [Medline: 9535311]
56. Chuang LL, Wu CY, Lin KC, Hsieh CJ. Relative and absolute reliability of a vertical numerical pain rating scale supplemented with a faces pain scale after stroke. Phys Ther 2014;94(1):129-138. [doi: 10.2522/ptj.20120422] [Medline: 24029301]
57. Marin RS, Biedrzycki RC, Firinciogullari S. Reliability and validity of the apathy evaluation scale. Psychiatry Res 1991;38(2):143-162. [doi: 10.1016/0165-1781(91)90040-v] [Medline: 1754629]
58. Scharre DW, Chang S, Murden RA, Lamb J, Beversdorf DQ, Kataki M, et al. Self-Administered Gerocognitive Examination (SAGE): a brief cognitive assessment Instrument for mild cognitive impairment (MCI) and early dementia. Alzheimer Dis Assoc Disord 2010;24(1):64-71. [doi: 10.1097/WAD.0b013e3181b03277] [Medline: 20220323]
59. WMA Declaration of Helsinki–ethical principles for medical research involving human subjects. World Medical Association. URL: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/ [accessed 2025-10-31]

XSL•FO
RenderX

60. Jeste DV, Palmer BW, Appelbaum PS, Golshan S, Glorioso D, Dunn LB, et al. A new brief instrument for assessing decisional capacity for clinical research. Arch Gen Psychiatry 2007;64(8):966-974. [doi: 10.1001/archpsyc.64.8.966] [Medline: 17679641]
61. Bradwell HL, Edwards KJ, Winnington R, Thill S, Jones RB. Companion robots for older people: importance of user-centred design demonstrated through observations and focus groups comparing preferences of older people and roboticists in South West England. BMJ Open 2019;9(9):e032468 [FREE Full text] [doi: 10.1136/bmjopen-2019-032468] [Medline: 31558461]
62. Duque E, Fonseca G, Vieira H, Gontijo G, Ishitani L. A systematic literature review on user centered design and participatory design with older people. 2019 Presented at: IHC '19: Proceedings of the 18th Brazilian Symposium on Human Factors in Computing Systems; 2019 Oct 22; Espírito Santo, Vitória, Brazil p. 1-11 URL: https://api.semanticscholar.org/CorpusID:203781081
63. Johnson J, Finn K. Designing User Interfaces for an Aging Population: Towards Universal Design. Burlington, Massachusetts: Morgan Kaufmann Publisher; 2017.
64. Brimelow RE, Dawe B, Dissanayaka N. Preliminary research: virtual reality in residential aged care to reduce apathy and improve mood. Cyberpsychol Behav Soc Netw 2020;23(3):165-170. [doi: 10.1089/cyber.2019.0286] [Medline: 31829729]
65. Saredakis D, Keage HA, Corlis M, Loetscher T. Using virtual reality to improve apathy in residential aged care: mixed methods study. J Med Internet Res 2020;22(6):e17632 [FREE Full text] [doi: 10.2196/17632] [Medline: 32469314]
66. Cornet VP, Toscos T, Bolchini D, Rohani Ghahari R, Ahmed R, Daley C, et al. Untold stories in user-centered design of mobile health: practical challenges and strategies learned from the design and evaluation of an app for older adults with heart failure. JMIR Mhealth Uhealth 2020;8(7):e17703 [FREE Full text] [doi: 10.2196/17703] [Medline: 32706745]
67. Ekstedt M, Kirsebom M, Lindqvist G, Kneck Å, Frykholm O, Flink M, et al. Design and development of an eHealth service for collaborative self-management among older adults with chronic diseases: a theory-driven user-centered approach. Int J Environ Res Public Health 2021;19(1):391 [FREE Full text] [doi: 10.3390/ijerph19010391] [Medline: 35010652]
68. Brox E, Konstantinidis ST, Evertsen G. User-centered design of serious games for older adults following 3 years of experience with exergames for seniors: a study design. JMIR Serious Games 2017;5(1):e2 [FREE Full text] [doi: 10.2196/games.6254] [Medline: 28077348]
69. Proffitt R, Sevick M, Chang C, Lange B. User-centered design of a controller-free game for hand rehabilitation. Games Health J 2015;4(4):259-264 [FREE Full text] [doi: 10.1089/g4h.2014.0122] [Medline: 26182212]
70. Chu CH, Biss RK, Cooper L, Quan AML, Matulis H. Exergaming platform for older adults residing in long-term care homes: user-centered design, development, and usability study. JMIR Serious Games 2021;9(1):e22370 [FREE Full text] [doi: 10.2196/22370] [Medline: 33687337]
71. Hoffman AS, Bateman DR, Ganoe C, Punjasthitkul S, Das AK, Hoffman DB, et al. Development and field testing of a long-term care decision aid website for older adults: engaging patients and caregivers in user-centered design. Gerontologist 2020;60(5):935-946 [FREE Full text] [doi: 10.1093/geront/gnz141] [Medline: 31773140]
72. Balakrishnan S, Salim SSB, Hong JL. User centered design approach for elderly people in using website. 2012 Presented at: International Conference on Advanced Computer Science Applications and Technologies (ACSAT); 2012 November 26-28; Kuala Lumpur, Malaysia p. 382-387. [doi: 10.1109/ACSAT.2012.11]
73. Tate J, Mion L, Migovich M, Ghosh R, Khan N, Kilpatrick A, et al. A multi-site randomized clinical trial of socially assistive robots on engaging older adults with cognitive impairment residing in long-term care settings: a protocol paper. Contemp Clin Trials 2025;155:107980 [FREE Full text] [doi: 10.1016/j.cct.2025.107980] [Medline: 40482961]

Abbreviations

AES-S: Apathy Evaluation Scale-Self Rated
HCI: human-computer interaction
HHI: human-to-human interaction
HRI: human-robot interaction
IRB: institutional review board
LTC: long-term care
NIA: National Institute on Aging
SAGE: Self-Administered Gerocognitive Examination
SAR: socially assistive robot
SAR-VR: socially assistive robot and nonimmersive virtual reality
SUS: System Usability Scale
UCD: user-centered design
VR: virtual reality

XSL•FO
RenderX

Edited by D Liu; submitted 01.04.25; peer-reviewed by M Hosseini, X Zhao; comments to author 26.07.25; revised version received 20.09.25; accepted 20.10.25; published 02.12.25.

Please cite as:
Ghosh R, Khan N, Migovich M, Tate JA, Maxwell CA, Newhouse PA, Scharre DW, Tan A, Mion LC, Sarkar N
Engaging Older Adults and Staff in the Co-Design and Evaluation of Socially Assistive Robot and Virtual Reality Activities for Long-Term Care: User-Centered Study
JMIR Aging 2025;8:e75288
URL: https://aging.jmir.org/2025/1/e75288
doi:10.2196/75288
PMID:

©Ritam Ghosh, Nibraas Khan, Miroslava Migovich, Judith A Tate, Cathy A Maxwell, Paul A Newhouse, Douglas W Scharre, Alai Tan, Lorraine C Mion, Nilanjan Sarkar. Originally published in JMIR Aging (https://aging.jmir.org), 02.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Hybrid Rule- and Large Language Model–Based Embodied Voice Assistant (GRACE) for Cognitive Stimulation in Older Adults: Usability Study Assessing Technical Feasibility, Technology Acceptance, and Working Alliance

Rasita Vinay[1,2], PhD; Ekaterina Uetova[3], MA; Nora Camilla Tommila[2], MSc; Nikola Biller-Andorno[1], Prof Dr; Tobias Kowatsch[2,4,5], Prof Dr

[1]Institute of Biomedical Ethics and History of Medicine, Faculty of Medicine, University of Zurich, Zurich, Switzerland

[2]Department of Management, Technology, and Economics, ETH Zurich, Zurich, Switzerland

[3]School of Computer Science, Technological University Dublin, Dublin, Ireland

[4]Institute for Implementation Science in Health Care, Faculty of Medicine, University of Zurich, Zurich, Switzerland

[5]School of Medicine, University of St. Gallen, St Gallen, Switzerland

Corresponding Author:
Rasita Vinay, PhD
Department of Management, Technology, and Economics
ETH Zurich
Weinbergstrasse 56/58
Zurich, 8092
Switzerland
Phone: 41 44 632 69 52
Email: rvinay@ethz.ch

Abstract

Background: The health and economic burden of dementia has led the World Health Organization to recognize it as a public health priority. Although there currently does not exist a cure for dementia, there are multiple interventions aimed at preventing the risk of dementia and improving the quality of life of people with dementia. Voice assistants (VAs), particularly those using large language models (LLMs), have emerged as promising tools to deliver these interventions to older adults due to their accessible and natural interface.

Objective: This pilot study aimed to evaluate the technical feasibility (ie, functional performance and usability) and user acceptance of the embodied rule-based and LLM VA GRACE, as well as the perceived strength of the collaborative relationship or working alliance, between GRACE and healthy older adults during the delivery of cognitive stimulation interventions.

Methods: A pilot study was conducted with 21 healthy German-speaking adults aged 60 years and older. Participants interacted with GRACE in a laboratory setting for 10-15 minutes. The interaction involved a structured cognitive stimulation session using rule-based and LLM components. Data were collected using pre- and postinteraction questionnaires and semistructured interviews. Quantitative analysis included descriptive statistics and Wilcoxon signed rank tests. Qualitative data were analyzed thematically.

Results: Participants rated GRACE positively, with statistically significant scores above neutral (P<.001 for perceived ease of use, usefulness, enjoyment, and working alliance; P=.009 for perceived control; and P=.009 for intention to continue interacting). Thematic analysis revealed that GRACE was perceived as easy to understand and unambiguous, friendly, and supportive, with intervention components viewed as enjoyable and appropriately challenging. Areas for improvement included personalization, response delays, and voice quality.

Conclusions: The results suggest that embodied rule-based and LLM VAs like GRACE are feasible and well-received tools for delivering cognitive interventions to older adults. Future iterations will incorporate feedback and extend testing to individuals at risk for dementia.

(JMIR Aging 2025;8:e76489) doi:10.2196/76489

XSL•FO
RenderX

KEYWORDS

voice assistant; dementia; older adults; human-computer interaction; embodied conversational agent; cognitive intervention; digital health interventions; large language model

Introduction

Background

Dementia is a collection of symptoms severely affecting memory, thinking, and other cognitive skills in an individual's daily life [1]. The most common form of dementia is Alzheimer disease [2], which was found to be the eighth leading cause of death worldwide [3]. In 2019, the estimated global societal cost of dementia was US $1.3 trillion, of which 50% of the costs were associated with informal caregiving [4]. The World Health Organization recognizes dementia as a public health priority through its global action plan [5], which hopes to minimize its health and economic burden worldwide. While this plan includes various action areas, such as dementia awareness, risk reduction, and treatment, digital assistive technologies should also be explored as prevention and management strategies.

Digital assistive technologies have emerged as valuable tools in supporting individuals living with dementia by enhancing independence, safety, and quality of life [6]. Designed to address challenges such as memory loss, confusion, and communication difficulties, these technologies support the autonomy of people with dementia while also alleviating the burden on caregivers [7]. Among the various forms of digital assistive technologies, voice-based digital assistive technologies, or voice assistants (VAs), offer distinct advantages due to an intuitive and hands-free conversational interface and low digital literacy requirement [8]. Unlike many other digital tools that require physical interaction or complex navigation, VAs rely on natural language processing, allowing users to engage through simple verbal commands. VAs can seamlessly integrate into daily routines, providing real-time reminders and scheduling medication [9-12] and even offering cognitive stimulation [5,13,14].

When VAs were generally assessed, they reported high user satisfaction for lifestyle interventions [15]. Prior research has also shown that the embodiment of these conversational VAs can further improve user experience [16].

Despite significant advances, a notable research gap remains in the development of embodied VAs that move beyond task-based support to deliver structured, evidence-based therapeutic interventions specifically tailored to people with early-stage dementia. While existing embodied VAs have demonstrated potential in offering reminders, companionship, and daily routine support, few have integrated clinically validated therapies [17]. Agents that do provide such therapeutic content, such as Eva [13], are often highly scripted, relying on predefined dialogue flows and rule-based interactions that limit personalization, adaptability, and natural dynamic conversation. To address this research gap, GRACE integrates cognitive stimulation therapy (CST) [18] with a hybrid architecture that combines rule-based logic and large language models (LLMs) to support more flexible, personalized, and empathetic conversations.

In a previous study, we used an embodied VA prototype, GRACE, to test user acceptance and working alliance between the user and the technology in healthy younger adults (18-60 years) [19,20]. User acceptance was assessed with constructs of perceived ease of use (PEU), perceived usefulness (PU), and perceived enjoyment (PEN) [21-23]. Working alliance, which can be understood as the perceived strength of the collaborative relationship between GRACE and the user, similar to a doctor-patient relationship, was measured through the 6-item Session Alliance Inventory [24]. The user acceptance constructs were selected based on the research aims of this study and consistent with prior work on conversational agents in digital health [25-27], rather than to test a specific theoretical framework (eg, the technology acceptance model). Results of this first study were used to optimize GRACE (version 1), including its visual features, interaction script, language (English to German), and voice quality.

This paper presents the results of our second pilot study, which evaluated GRACE (version 2) with healthy older adults (60 years and older). To extend our earlier work, we added the constructs of perceived control (PC) and intention to continue interacting (ITI) [23,28], which have also been applied in related studies of conversational agents [25-27]. Technical feasibility, which we defined as the usability and functional performance of GRACE when used by healthy older adults, was evaluated through descriptive analysis of user acceptance and working alliance constructs, as well as qualitative feedback gathered from a postinteraction interview. This interview also evaluated the ethical dimensions observed within the technology by users, using values presented in the European Union's Assessment List for Trustworthy Artificial Intelligence [29]. However, the qualitative findings from the interviews fall outside the scope of this paper and will be published in a separate paper that targets the bioethics research community.

Related Work

Older Adults and VAs

A growing area of interest among researchers involves leveraging VAs to support the health and well-being of older adults [17,30,31]. VAs have the potential to support older adults' independence by facilitating daily activities through smooth, speech-based communication, offering a user-friendly interface that requires minimal technological expertise and can accommodate varying levels of digital literacy [17,32-34]. Despite stereotypes about older adults' reluctance toward technology, studies have shown that older adults often have positive attitudes and are open to using these devices in their daily lives [17,32,35-38]. It was found that older adults felt VAs might be useful for alerts of serious illness, health information management, and search, facilitating collaboration between themselves and caregivers [31,39,40]. At the same time, older adults were concerned about confidentiality risks and receiving trusted information [31,39] and skeptical about personal privacy and relational trust [40].

XSL•FO
RenderX

Numerous VA applications have been developed to support older adults, generally falling into 2 main categories: conversational social robots, which have a physical presence and use multiple affective communication modalities such as facial expressions, gestures, or body movements; and virtual artificial intelligence (AI) assistants or smart speakers, which primarily rely on voice interaction [17]. GRACE falls into the first category—a conversational social robot that engages users through verbal and nonverbal cues (facial expressions). A recent systematic review highlighted the perception of conversational AI among older adults, where they emphasized the need for empathetic, personalized, and context-aware interactions [41]. Current technologies in aging and early dementia care primarily aim to support individuals by promoting independence, enhancing social engagement, improving home safety, and monitoring health and well-being [42]. These applications can serve various roles; some are designed to act as companions for older adults experiencing loneliness [43]. For example, a "digital radio" system integrated voice-controlled features to enhance older adults' connection with their family and address their needs for voice companionship, emotional interaction, and mental and physical well-being [44]. A companion robot "E-Bot" with integrated Amazon Alexa was designed to support older people living alone by enabling continuous interactions with caregivers, family members, and physicians while also monitoring physiological signals via wearable devices and an AI-enabled camera connected through Amazon AWS [45]. The NAO Robot provides single-session interactions, during which it reads newspapers to older adults, asks personal questions about their past, and listens to them, nodding occasionally and maintaining eye contact [46]. Studies have shown that robotic pets, which respond to user speech and petting with sounds and eye or body movements, can enhance communication, interaction skills, and activity participation of older adults [47].

VAs can also serve as home assistants, helping individuals manage their daily schedules. For instance, the Google Home Mini smart speaker can help cue individuals to take scheduled medications [9]. A companion bot named "Hello Steve" can send emails, open YouTube for entertainment, and track medication schedules to provide timely reminders for older users [10]. Another VA called "Mary" can assist older adults with daily tasks such as reminders, guidance with household activities, and locating objects [11]. Co-design approaches have also been explored with older adults to support health self-management, where the study found success in medication adherence and health-tracking features [12].

VAs have also been developed to provide therapeutic interventions for older adults. For example, VAs can be used for the prevention and management of chronic and mental health conditions [8], such as depression [48], encouraging physical activity [49], providing cognitive support [13,14], including VAs using LLMs with prompt engineering [50], and assisting with palliative care [51].

Despite this progress, most existing systems, including NAO, Hello Steve, and Mary, are limited in several key areas. First, these systems often rely on scripted interactions, which restrict their ability to provide personalized and context-aware engagement. Second, while some systems offer companionship or support daily living, few have integrated clinically validated interventions such as CST. Furthermore, the majority of VA applications are not co-designed with people living with dementia and their caregivers, resulting in limited adaptability to users' changing needs and preferences over time.

To address these gaps, GRACE was designed as an embodied VA that moves beyond basic reminders and social dialogue to deliver structured, evidence-based cognitive interventions. By combining CST with hybrid rule-based and LLM-powered dialogue and by grounding development in stakeholder-driven design, GRACE aims to provide sustained, meaningful, and personalized support for individuals with early-stage dementia, advancing the field toward truly person-centered digital health technologies.

Rule-Based Versus LLM VAs

Most current VAs operate as passive, rule-based systems that respond only to wake words, requiring users to adapt their communication style to achieve desired outcomes [52,53]. While this simplicity can make interactions straightforward, it often leads to frustration when the system cannot respond to user requests [52,54]. LLMs present a significant advancement, enabling more adaptive interactions by better "understanding" human requests and overcoming the limitations of traditional, rule-based voice interfaces [52]. Nevertheless, both systems have their strengths and limitations.

Rule-based VAs offer advantages such as transparency, reproducibility, and safety, as their functionality relies on predefined rules, allowing developers to maintain full oversight due to their predictable responses [52]. However, rule-based systems are limited by their lack of flexibility and adaptability, as they cannot handle out-of-scope queries or adjust their personality and the challenges of scaling and maintaining complex multiturn interactions [52].

LLM VAs, on the other hand, overcome many of the challenges faced by rule-based VAs [52]. LLMs can process complex queries, retain context in multiturn dynamic conversations where user interactions are less predictable, and generate nuanced, more human-like responses [52,55]. However, these advantages come with drawbacks. LLMs' reliance on large-scale computational resources raises challenges related to latency, particularly in real-time applications [56], hallucinations, that is, LLMs can generate nonsensical, incorrect text [57], unpredictable and unexpected responses [52], and bias [58,59].

Human Interactions With LLM VAs

LLM VAs have been shown to enhance individuals' interactions to be more natural and effective. For example, FunAudioLLM models improve multilingual speech recognition and generation, enabling applications like emotional voice chat and speech-to-speech translation [60]. Additionally, Sesame AI showed the ability to reproduce natural pauses and intonation during conversations [61]. A study on VAs demonstrated that the integration of LLMs enhances conversational adaptability, enabling more nuanced and context-aware responses, while effectively addressing intent recognition failures and conversational breakdowns [62]. A study on LLM-powered robots showed that they excel in fostering connections and

XSL•FO
RenderX

deliberation, but they struggle with logical communication and can sometimes cause anxiety [63]. However, more recent studies on LLM social-assistive robots found that older adults could successfully engage and improve on cognitive health-promoting tasks [64]. Another study introduced a framework for integrating LLMs as speech interfaces for assistive robots, demonstrated with the Obi feeding robot [65]. This highlights the importance of social capabilities that are often overlooked in physically assistive robots to enhance user engagement through social interaction. Talk2Care, an LLM asynchronous communication system designed to connect older adults at home with health care providers, demonstrated its potential to address communication challenges, enrich health information shared by older adults, and reduce providers' workload [66].

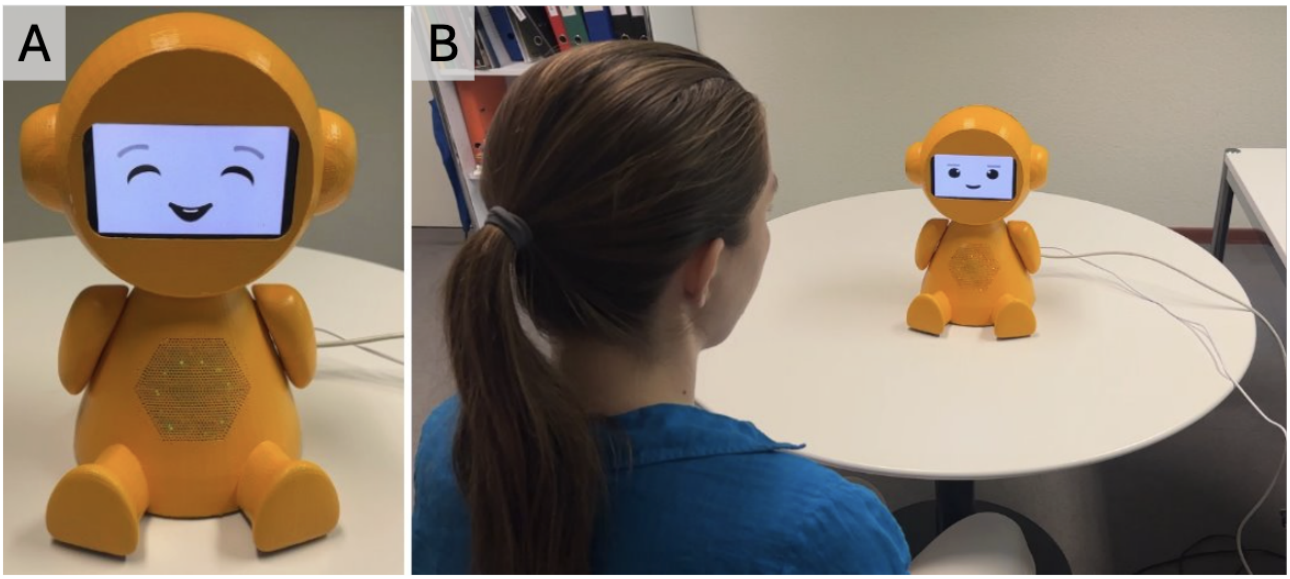
In this and previous work [20], we address the limitations of traditional rule-based VAs while leveraging the strengths of LLMs by extending the rule-based language [67] with OpenAI's application programming interface [68], enabling an embodied VA to combine the transparency and predictability of rule-based systems with the adaptability of LLM-generated responses within predefined, transparent boundaries.

Methods

Developing the Embodied GRACE Robot

GRACE was 3D-printed with polyethylene terephthalate (a type of polyester plastic) using designs from an open-source social robotics platform [69], with its color changed from gray to bright yellow based on feedback from our previous study [20]. GRACE has a reSpeaker 4-Mic Array microphone, an LED ring, a 5.5-inch Waveshare active-matrix organic LED screen, an 8 GB Raspberry Pi 4 Model B, and a Xiaomi Bluetooth speaker. GRACE needs to be connected to power and a wired internet. Communication between the Raspberry Pi and an external computer uses the Message Queuing Telemetry Transport protocol. The external computer runs software [67] that executes interaction scripts or simulates them for development without GRACE's body. GRACE's modular Python-based software (Python Software Foundation) includes LED, audio, speech-to-text, text-to-speech, and display components [69]. The LED lights, green for listening to the individual and red for when GRACE is speaking or processing information, can be seen from the mesh in the body of GRACE (Figure 1). Audio output uses a Bluetooth speaker placed in GRACE's head. The speech-to-text module uses Google Speech Recognition during <listen/> commands in the interaction script, while <talk> commands trigger the ElevenLabs service (Eleven Multilingual version 2) [70] to generate natural-sounding German speech ("Leonie" voice), cached for reuse. The screen shows animated facial expressions like blinking and smiling.

Figure 1. GRACE (version 2) prototype used for this pilot study: (A) front-facing image and (B) with a test user.



Design of Conversational Turns

The interaction script, adapted from a pilot study [20], was first drafted in plain text, and then, converted to XML with commands like <talk> and <listen/> (the full interaction script can be found in Multimedia Appendix 1) [67]. GRACE responds with predefined answers or dynamic replies via OpenAI's application programming interface [68]. This hybrid approach allows us to maintain consistency and control of rule-based interactions while enabling more nuanced and context-aware conversations, addressing the challenges faced by traditional systems and LLM VAs.

The script is based on CST and includes 3 intervention components and 2 additional activities (onboarding and conclusion). CST is an evidence-based group therapy program for people with dementia, encompassing a variety of themed activities (ie, word games, reminiscence, and creative games) to improve well-being and cognition [71,72]. A CST manual was used to guide development of the GRACE sessions as well as to identify different activities that could be included as part of the intervention components [73]. The script starts with an

XSL•FO
RenderX

onboarding activity (supported by LLM), where GRACE introduces itself to the participant and engages in conversation to get to know the participants (ie, by asking questions such as "Do you have siblings?", "Do you have a favorite hobby?", and "What is your hometown?"). In this interaction, the responses from GRACE are LLM-generated, where we prompted the LLM to provide a supportive statement or fun fact based on the participants' answer (see Multimedia Appendix 2 for the LLM prompts used in the complete interaction). After this, GRACE provides participants with an overview of the interventions it will offer.

The first intervention is a warm-up activity in the form of a guessing game, where different animal sounds are played, and the participant must guess the name of the animal that makes that sound. This interaction is rule-based with stored audio files playing selected animal sounds (ie, a cow, cat, dog, rooster, and horse), and GRACE responding with the name of the correct animal. Next, the participant can choose which intervention they want to complete first, a memory or breathing exercise. By allowing participants to select the order of the intervention, our goal is to promote autonomy and motivation, in line with the self-determination theory [74]. In the memory exercise, GRACE reads a short shopping list (1-5 items, an LLM is used to help determine the individual's answers and correct item recall), and the participant is asked to repeat the list in any order. If the participant cannot list all the items, GRACE repeats the missed items. The breathing exercise is rule-based and consists of a short, guided breathing training based on exhale-focused cyclic sighing [75]. GRACE allows participants to select from 3 acoustic relaxing environments (stored audio files of the forest, the sea, and a cabin with a fireplace), which are used to immerse them in the training environment through background noises and a description of the environment by GRACE. There is no participant communication beyond the environment selection; the stored audio files guide users to follow inhale or exhale techniques for 2 minutes. In the rule-based concluding activity, GRACE thanks the participant for doing the exercises, asks them about their favorite activity, provides health education on the importance of completing cognitive interventions, and farewells with a reminder to stay active.

Participants

Overall, 21 participants meeting the inclusion criteria took part in the study during August and November 2024. Study recruitment took place via email and posters, highlighting the study's aims and the inclusion and exclusion criteria.

The inclusion criteria were as follows: aged 60 years and older, ability to speak and understand German (at a B1 level), and no cognitive or hearing impairments. The exclusion criteria were as follows: having severe psychiatric disorders or a history of substance abuse.

The study was conducted in a laboratory setting in a single sitting, with participation lasting approximately 45 minutes.

Data Collection

The data collection consisted of a pre- and postinteraction questionnaire (Multimedia Appendices 3 and 4) and a short semistructured interview (Multimedia Appendix 5). The preinteraction questionnaire collected the participants' age, gender demographics, and general information regarding experience with VAs.

The postinteraction questionnaire consisted of 12 items used to assess technology acceptance [21]. The 7-point Likert scales were used for each item and anchored from 1=strongly disagree to 7=strongly agree. Consistent with prior work on conversational agents [27], we used self-report instruments for PEU, PU, PEN, PC of GRACE, and the ITI with GRACE [21,28,76]. Working alliance, which is an important relationship quality and robustly linked to treatment outcomes [77-79], was measured with the 6-item Session Alliance Inventory [80].

A semistructured interview with 10 questions was conducted to collect qualitative feedback on the interaction with GRACE. Questions 1-7 were developed by authors RV and NB-A, based on the Assessment List for Trustworthy Artificial Intelligence checklist [29], on the principles of human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity, nondiscrimination and fairness, societal and environmental well-being, and accountability. The results of interview questions 1-7 will be reported in a separate publication targeting the bioethics community. Questions 8-10 relate to the overall positive and negative aspects of interaction with GRACE and any further comments or suggestions that are reported in this work.

Procedure

Participants completed the preinteraction questionnaire (approximately 1 minute) after providing informed consent. Participants were then given instructions for the interaction with GRACE (ie, what the different LED-light colors mean and speaking clearly, etc). Once the participants were ready, the interaction was started by a researcher, after which GRACE led the entire interaction independently, with supervision from the researcher. The interaction with GRACE lasted 10-15 minutes and followed the interaction script as mentioned earlier. Following the interaction, participants completed a postinteraction questionnaire (5 minutes) and were then interviewed by a researcher (15 minutes).

Data Analysis

To summarize participants' responses on the postinteraction questionnaire, means and SDs were computed for each quantitative self-reported item. As the Shapiro-Wilk tests revealed deviations from normality in the data distribution, nonparametric analysis was warranted. Therefore, Wilcoxon signed rank tests were applied to examine whether the median scores significantly diverged from the neutral midpoint value of 4 on the 7-point Likert scale. Additionally, a composite score for the 6-item Session Alliance Inventory was created by averaging responses across all items, supported by high internal consistency (Cronbach α=0.92) [81]. Consistent with prior work on embodied VAs [20,82], thematic maps were used to present the results of interview questions 8-10. Authors RV and NCT conducted 3 rounds of thematic analysis until a consensus was reached.

XSL•FO
RenderX

Ethical Considerations

The study was exempted from formal review by the ethics committee of the University of St. Gallen on May 17, 2024. During the onboarding, participants were informed about the study's aims, procedure, and rights as participants. They were then provided with a written consent form for participation. All participants were assigned a participant code to maintain anonymity of responses (indicated as P#, where # was a randomly assigned number from 01 to 21). No compensation was provided to the participants.

Results

In total, 21 individuals (8 male and 13 female) participated in the study, where the average participant age was 69.3 (SD 6.2) years. From the preinteraction questionnaire, 8 participants indicated that they had used a VA previously, on average, a few times a month. Participants also indicated that they would use a VA if it provided them with activities and exercises to improve their quality of life (mean 5.38, SD 1.07, "somewhat agree" on the 7-point Likert scale).

Table 1 presents a summary of the quantitative evaluation (ie, technology acceptance and working alliance) of GRACE, showing that all mean scores exceeded the neutral midpoint of 4 on the scale. The postinteraction questionnaire revealed strong indicators of technological acceptance and a robust sense of working alliance with GRACE (mean 5.49, SD 0.81; Table 1). Participants rated the system as highly easy to use (mean 6.19, SD 0.81), suggesting that GRACE's conversational interface and structure were well understood and accessible. Moreover, PEU, PU, PEN, and overall assessment of the working alliance had a significance of $P<.001$. PC had a significance of $P=.009$, ITI with GRACE had a significance of $P=.009$, and item 5 from the Session Alliance Inventory had a significance of $P=.01$. Therefore, it can be concluded that participants responded positively to GRACE.

Table 1. Postinteraction quantitative evaluation of 21 participants on their perceived ease of use (PEU), perceived usefulness (PU), perceived enjoyment (PEN), perceived control (PC), intention to continue interacting (ITI), and perceived session alliance (PSA)[a].

	Item	Mean (SD)	V[b]
PEU	I thought the response from GRACE was easy to understand.	6.19 (0.81)	231.0[c]
PU1	I found it useful to work on the intervention with GRACE.	5.76 (1.14)	205.5[c]
PU2	I found GRACE motivated me to perform the exercises.	5.48 (1.21)	178.5[c]
PEN	I had fun using GRACE.	5.95 (0.86)	231.0[c]
PC	I was able to control the interaction with GRACE.	5.19 (1.69)	174.0[d]
ITI	I would continue interacting with GRACE.	5.14 (1.53)	105.0[d]
PSA1	I felt that GRACE and I respected each other.	5.81 (1.25)	204.0[c]
PSA2	I felt that GRACE appreciated me.	5.57 (1.21)	198.0[c]
PSA3	I felt that GRACE cared about me even when I did things it did not approve of.	5.33 (1.06)	120.0[c]
PSA4	I felt that GRACE and I are working toward mutually agreed upon goals.	5.67 (1.15)	222.0[c]
PSA5	I felt that GRACE and I agree on what is important for me to work on.	5.10 (1.73)	121.0[e]
PSA6	I believe the way GRACE and I are working with my problem is correct.	5.48 (1.03)	136.0[c]
PSA	Overall	5.49 (1.06)	222.0[c]

[a]The 7-point Likert scales were used from 1=strongly disagree to 7=strongly agree, and a 6-point scale for PSA items.

[b]V is the test statistic of the Wilcoxon signed rank with test value 4.

[c]$P<.001$.

[d]$P=.009$.

[e]$P=.01$.

Following their interaction with GRACE, a semistructured interview was conducted with all participants. Thematic analysis of the responses revealed key insights into what participants appreciated (Figure 2), areas for improvement (Figure 3), and further comments or suggestions (Figure 4). Direct participant quotes for the 3 interview questions are included in Multimedia Appendix 6.

XSL•FO
RenderX

Figure 2. Thematic map representing answers of study participants regarding the question: "What did you like about the interaction with GRACE?". Boxes in black indicate identified categories, and boxes in white indicate participant responses.

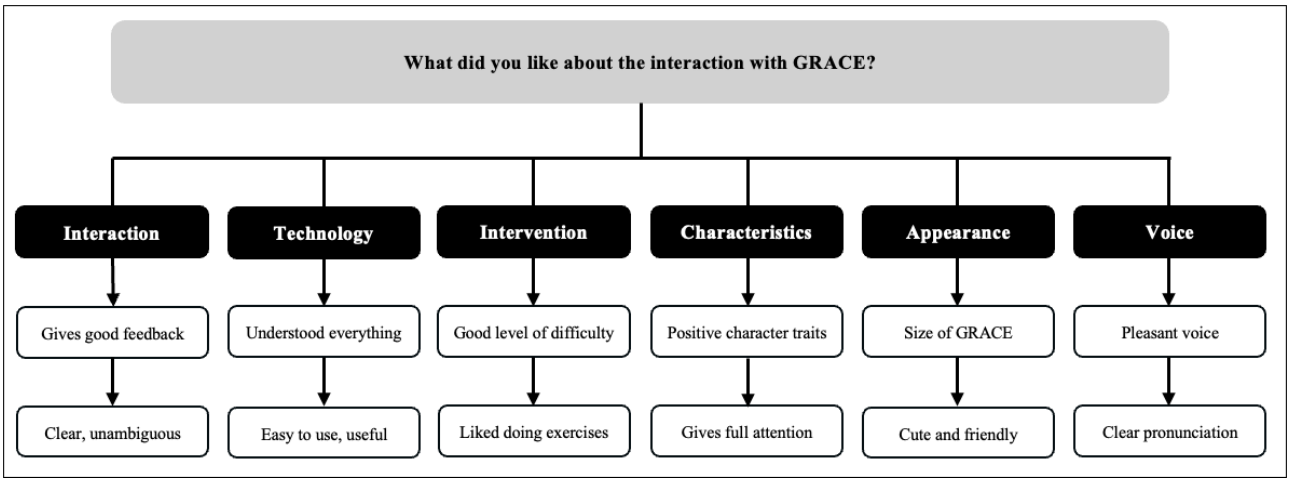


Figure 3. A thematic map representing answers of study participants regarding the question: "What needs to be improved about GRACE?". Boxes in black indicate identified categories, and boxes in white indicate participant responses.

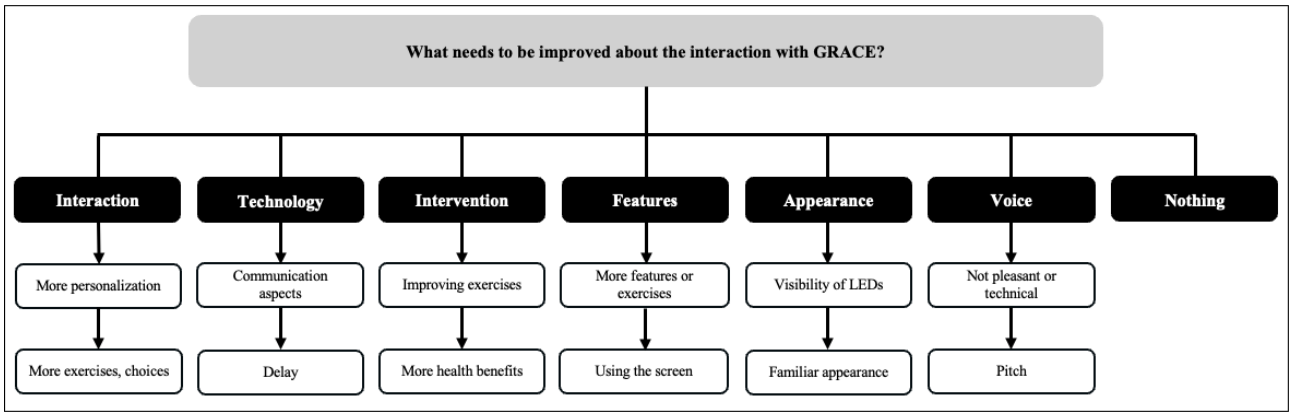
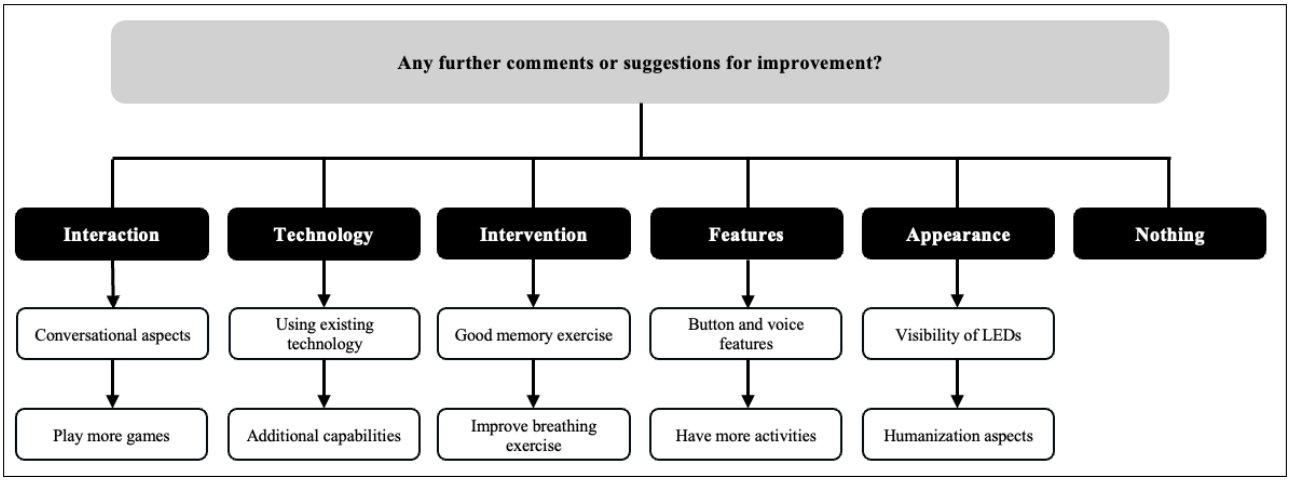


Figure 4. A thematic map representing answers of study participants regarding the question: "Any further comments or suggestions for improvement?". Boxes in black indicate identified categories, and boxes in white indicate participant responses.



Participants expressed a wide range of positive impressions regarding their interaction with GRACE. This included the ease of understanding and using the system, the friendly and supportive nature of GRACE, and the usefulness of the exercises. GRACE was described as "clear and unambiguous" (P01), "pleasant and friendly" (P08), and "able to give feedback" (P05 and P11). Many appreciated the ability to interact at their own pace and felt that GRACE responded accurately and attentively, even to mumbled speech (P13 and P14). The intervention itself, such as the memory and breathing exercises, was seen as engaging and appropriately challenging (P05, P11, and P19) and not too stressful (P04, P08, and P15). Participants also responded positively to GRACE's appearance and voice, describing her as "cute" and with "clear pronunciation" and a "pleasant voice" (P03, P08, P16, P19, and P21). Several participants also appreciated GRACE's personality traits,

XSL•FO
RenderX

perceiving her as “sympathetic” (P02), “polite” (P04), and able to offer “full attention” during the interaction (P06 and P12).

Despite overall satisfaction, participants identified several areas where GRACE could be improved, focusing on technology, interaction design, intervention content, and user experience. Technological aspects such as communication delays and accuracy were frequently mentioned: “Her reaction is a bit slow” (P20), “It didn’t always understand what I said” (P07), and “the communication is not great yet” (P21). The voice was another concern, described by some as “too technical” (P01), “not so pleasant” (P07), or “disturbing” due to pitch (P05). Participants desired more personalization and variety in interaction: “the answers were very stereotypical” (P08), and that “the exercises could be personalized” (P11 and P13). Some also found the content demanding or insufficiently explain: “the tasks and exercises were quite demanding” (P02) and “maybe it could explain the effects more, why should I do these exercises?” (P03). Many also felt the system to be “mechanical” (P20) and “impersonal” (P11), which prevented its ability to establish a more human connection (P12).

A range of additional reflections and suggestions to enhance GRACE’s functionality and relevance were also offered by participants in the last interview question. Some proposed expanding the system’s interactive capabilities, such as incorporating games (P06), adding a physical button (P08), or enabling more open-ended conversations (P12). Several comments emphasized the value of personalization, including the ability to choose a familiar image or voice: “Human face would be good or to be able to choose a photo, like of my grandson” (P08) and “maybe give also a choice for a nice, male voice” (P13). Others suggested improvements to specific features, such as enhancing the lighting (P07), adding more background sounds to the breathing exercise (P04), or clarifying the system’s nonhuman nature if a more lifelike appearance is used (P12). A few comments reflected on the context and potential of GRACE, including its suitability for people with dementia: “My mother responded to ‘garden’, ‘dog’, ‘favorite food’, ‘music’...” (P20) and how using personally meaningful cues could enhance engagement with GRACE. Others appreciated the potential for GRACE to support those who are lonely or in need of a companion-like presence: “It could be helpful to have an entity that asks questions and listens, even though it’s not a human interaction” (P12). The overarching tone of responses conveyed both interest and constructive engagement with GRACE’s future development.

Discussion

Principal Findings

This study aimed to assess technical feasibility, technology acceptance, and working alliance with GRACE in healthy older adults (60 years and older). The sample included a balanced group in terms of gender, with an average age of approximately 69 (SD 6.2) years, reflecting a relevant demographic for future health-related applications of VAs. While only a portion of participants reported prior experience with such technologies, there was a clear openness to engaging with GRACE, particularly when the system offered meaningful intervention components aimed at improving quality of life. Postinteraction feedback further supported this readiness: participants not only found GRACE highly usable and easy to navigate but also indicated a strong sense of working alliance with the system. This aligns with prior research, showing that older adults are receptive to voice interfaces when designed with clarity and intuitiveness in mind [83,84].

PU was also rated positively, with participants expressing that working with GRACE was beneficial (PU1: mean 5.76, SD 1.14) and motivating (PU2: mean 5.48, SD 1.21). Notably, PEN (mean 5.95, SD 0.86) was among the highest-rated constructs, indicating that beyond usability, participants genuinely enjoyed the interaction. These ratings support the potential for voice-based interventions to engage users in an emotionally positive way, which is particularly valuable for health-promoting interventions in older populations [8].

Participants also reported a relatively high sense of PC over the interaction (mean 5.19, SD 1.69). However, this dimension showed more variability, potentially reflecting differences in individual familiarity with conversational technologies or preferences for interaction pace. Intention to interact (mean 5.14, SD 1.53) suggests a moderate to strong willingness to engage with GRACE in the future, reinforcing the system’s potential for repeated use and longer-term integration into daily routines.

The assessment of perceived session alliance (PSA) further underscores the strength of the interaction. All 6 session alliance items scored well above the midpoint on a 6-point scale, with particularly high ratings for mutual respect (mean 5.81, SD 1.25) and feeling appreciated (mean 5.57, SD 1.21). These findings suggest that users not only felt heard and acknowledged, but also that GRACE was perceived as working collaboratively toward shared goals (eg, PSA-4: mean 5.67, SD 1.15). Although item PSA-5 (“I felt that GRACE and I agree on what is important for me to work on.”) received the lowest PSA rating (mean 5.10, SD 1.73), it remained significantly above the neutral test value, indicating room for improvement in tailoring goals and content to individual users.

In line with the quantitative findings, qualitative feedback highlighted positive perceptions of the interaction with GRACE. Participants described the interaction as “clear and unambiguous,” “well explained,” and “stress-free,” emphasizing the ease of use and the feeling of being understood. Many found GRACE’s voice and appearance to be “pleasant,” “cute,” or “friendly,” with others noting she gave “good feedback” and created a sense of receiving full attention, even if it was “pseudo-interest.” Several participants appreciated the exercises, especially memory and breathing exercises, citing them as engaging and motivating. GRACE’s character was also described as “sympathetic,” “positive,” and “polite,” contributing to the experience of a supportive and approachable interaction partner.

At the same time, participants offered constructive suggestions for improvement, pointing to areas that could strengthen engagement and usability. Most notably, concerns were raised around the technical quality of the voice (eg, “too technical,” “not so pleasant,” and “mechanical”), slow system responses,

XSL•FO
RenderX

and occasional misunderstandings in speech recognition. Some participants desired greater personalization in the interaction, such as tailoring the pace of exercises or receiving feedback that felt less generic. Suggestions also included expanding the scope of features, such as adding a button, screen use, or more varied exercises, and incorporating a more human or familiar appearance, while carefully clarifying that GRACE is not a real person. A few noted that the interaction felt "short" or "superficial," suggesting potential for deeper engagement in relational continuity in future iterations.

Additional comments reiterated the need for more flexibility, both in terms of content and delivery. Participants proposed more background noises for the breathing exercise, more open conversational exchanges, and the option to personalize GRACE's appearance or voice (eg, offering a male voice or selecting a familiar face). While some questioned the added value over existing VAs like Siri, others particularly emphasized GRACE's potential in supporting people experiencing loneliness or cognitive decline. A particularly insightful remark highlighted that "GRACE's productivity is not just about reducing loneliness" but also about sparking memory, interest, and emotional connection through personalized keywords or emotionally salient cues, which may be relevant for future use in dementia care.

Altogether, these results provide compelling early evidence that GRACE can foster both functional and relational engagement with older adults, meeting core criteria of usability, enjoyment, and alliance formation. When interpreted alongside the qualitative feedback, it becomes clear that the perceived success of GRACE hinges not only on task performance but also on emotional, aesthetic, and interpersonal dimensions that contribute to a meaningful user experience.

Comparison With Own Prior Work

This section briefly compares the significant improvement of outcomes in this study with a previous pilot study that was conducted with an earlier version of GRACE in healthy younger adults [19,20]. Here, several key differences can be highlighted.

Design Enhancement

In response to prior feedback, GRACE's color was adjusted to create a warmer and more inviting visual appearance (from gray to yellow/orange). Additionally, an LED light was also inserted to indicate to users when GRACE was waiting for a response and when it was not. These design shifts may have positively influenced user perception and emotional engagement with the system. Furthermore, the current version of GRACE delivered an additional intervention—an animal guessing game, which may help users become more familiar with its interaction style earlier on. The idea behind this additional intervention was to help users get familiar with interacting with GRACE in an exercise and add more cognitive interventions as part of GRACE's service. Therefore, adding this simple intervention could potentially reduce initial resistance and confusion to interaction and improve receptivity to the system's guidance.

Better Alignment With the Target Population

Participants in this pilot study were more closely matched with the intended user group (people with early dementia) regarding demographics and needs. The closer alignment likely contributed to increased relevance, engagement, and PU of the system. The idea behind slowly moving toward the target population was to ensure a safe and ethical approach to reaching vulnerable population groups, where the technology should first be tested among healthy adults, reiterated, tested among healthy older adults, reiterated, and then finally presented to other relevant stakeholders involved in dementia care. Once more, feedback is incorporated from this group of stakeholders; only then would GRACE be tested among people with early dementia.

Use of Participants' Native Language

Unlike the first pilot study, this study conducted interactions in the participants' native language—German. Language is a crucial component for trust building and user comprehension. Conducting the interaction in high German (not Swiss German) likely reduced the participants' cognitive load, minimized misunderstandings, and even made the system feel more inclusive and accessible. In multilingual settings, such as those in Switzerland where the study was conducted, even subtle linguistic mismatches can impact the perception of relevance and credibility. By aligning linguistic communication with participants' cultural context, this study likely enabled more natural interactions and a deeper sense of being understood. Moreover, native language may have helped convey empathy and nuance, which are particularly important in health and well-being domains, where the tone and phrasing matter greatly.

Comparison With Related Work

Our findings align with previous research, indicating that older adults are receptive to VAs, particularly when these are designed to be intuitive, emotionally engaging, and relevant to users' needs [17,32,35-38]. GRACE's high usability and enjoyment ratings echo studies emphasizing the value of speech-based, low-barrier interfaces for older populations [33,34]. Similar to social robots like NAO or robotic pets [46,47], GRACE fostered a sense of being heard and respected, supporting past evidence that effective engagement enhances users' experience.

A recent study has also started exploring the use of VAs to deliver individual CST to people with dementia, initially tested with a very small sample of family members and informal caregivers [85]. Their preliminary findings hinted at acceptance and technical feasibility by the participants. Based on this pilot study, we believe that GRACE provides a more holistic approach to CST by embedding engaging socially interactive intervention components to enhance emotional connection and provide cognitive support and user engagement.

GRACE's hybrid rule-based and LLM design may explain the balance between PC and conversational flexibility. This reflects recent work, suggesting that combining transparency with adaptability can improve trust and usability in health-focused VAs [52,62,65]. However, some participants still perceive GRACE's responses as "mechanical," consistent with concerns about emotional nuance and speech naturalness in LLM systems [57,63]. This echoes broader critiques of LLMs in the health

XSL•FO
RenderX

care context, particularly regarding their ability to generate responses that are grammatically correct but factually inaccurate, emotionally misaligned, or culturally insensitive, which poses particular risks when interacting with older adults or individuals with cognitive impairments [86]. Moreover, the opacity of LLM-generated dialogue poses ethical challenges around accountability and safety, especially when users interpret the VA's output as clinically authoritative [87]. Additionally, the opaque nature of LLM decision-making makes it difficult for users to anticipate or control responses, which may undermine trust [88]. Requests for personalization (eg, voice, pace, and feedback content) resonate with studies calling for more tailored and culturally sensitive VA designs in older adult care [10,12,39,66]. Notably, some users experienced GRACE as more than just a tool, highlighting memory, meaning, and emotional connection, which extend prior findings on the potential of VAs to support not only functionality but also relational and emotional well-being [40,44,51].

Our findings underscore the importance of control and personalization in the design of VAs for older adults and dementia care. Prior studies have shown that older adults emphasize the importance of thoughtful language, appropriate tone, and personalization when designing conversational agents for individuals affected by dementia [89]. Moreover, older adults value the ability to control the interaction and engage with systems that reflect cultural and contextual relevance [90,91]. Cultural appropriateness, language use, and visual or conversational representation play a key role in making intelligent systems feel relatable and comfortable to use, particularly among racially and culturally diverse older adults [92]. Furthermore, users' perceptions of empathy are shaped by how well the systems acknowledge and respond to aspects of users' identity, such as race, disability, and lived health care experiences [93].

Limitations and Future Work

In future work, we plan to reiterate GRACE based on this study's feedback and conduct interviews with other stakeholders such as caregivers, family members, and health care professionals to obtain their perspectives on GRACE. Such insights could be beneficial in identifying other interventions that may be useful for people with dementia or ways in which communication between GRACE and the user can be enhanced. It may also be beneficial to speak to health insurers and service providers to identify scalable solutions for implementing GRACE within the dementia community, so that they may be considered for future research directions that focus more on the implementation of GRACE within the community.

Although this study shows preliminary results that delivering interventions to older adults with embodied VAs is positively perceived, the effects of the interventions and the relationship between them remain unknown. A next step would be to test such interventions with individuals at risk of developing dementia, that is, those with mild cognitive impairments or subjective cognitive decline, to get closer to the target population. Additionally, integration of interventions, such as reminiscence therapy, which allows older adults to engage in conversations that evoke meaningful memories and foster emotional well-being [94], or cognitive storytelling [95,96], has shown promise in improving mood, reducing stress, and enhancing cognitive engagement in older adults [94-96]. These intervention examples may be used in future versions of GRACE, arising from existing research.

A potential limitation of this study is the possibility of social desirability bias, whereby participants may have provided more positive evaluations than they felt [97]. This bias can be particularly relevant in face-to-face settings, where participants may wish to please the research team or provide favorable evaluations of a new technology. While we cannot fully exclude this bias in participants' responses, we actively mitigated the risks by using anonymized questionnaires and having the interviews conducted by student researchers outside the study team. These measures reduce the likelihood that participants felt pressured in their responses, and that the consistently high ratings likely reflect the system's design refinements and closer alignment with the target population. Nonetheless, subtle forms of social desirability bias may still have influenced the findings, as is common in studies relying on self-reported perceptions.

Another potential limitation that must be mentioned is that the postinteraction technology acceptance questionnaire was constructed ad hoc from items adapted from prior work rather than using a standardized, validated instrument. While this approach ensured alignment with the specific aims of our study and comparability with earlier research on conversational agents, it may limit the generalizability of our findings and complicate direct comparison with studies that apply standardized scales, such as the technology acceptance model. Future work should consider validating these adapted items or adopting established measures to strengthen comparability and robustness.

More work will also be needed to adapt VAs, especially speech-to-text engines, to older adults' speech dynamics. Furthermore, speech-to-text and text-to-speech engines cannot efficiently work with dialects, such as those found in Swiss German. Swiss researchers have created speech corpora, which may be able to be used in future iterations of GRACE, to provide native language communication for Swiss German users [98].

Conclusions

This work developed and evaluated a reiteration of GRACE that provided cognitive interventions to healthy older adults. Technical feasibility, technology acceptance, and working alliance were evaluated with overall positive assessments. Participants notably enjoyed the unambiguous interactions with GRACE, while the interventions offered were deemed appropriately challenging. Suggestions for future improvement include expanding the use of physical features, such as the screen and a button, and implementing more exercises and open-conversational abilities. This work demonstrates that embodied VAs can be positively used to deliver cognitive interventions to older adults. Results from this work will inform future iterations of GRACE, where additional interventions will be identified and tested among people with early dementia.

XSL•FO
RenderX

Acknowledgments

The authors would like to thank Marcelo Marques da Rocha and his team for their kind and generous assistance throughout the development of GRACE. The authors would also like to thank Lola Ackermann, Anja Bischof, and Sophie Stephan for assisting with conducting interviews.

Funding

RV's research was supported by the Käthe Zingg-Schwichtenberg Fund (KZS) from the Swiss Academy of Medical Sciences (grant KZS 13/24).

Conflicts of Interest

RV and TK are affiliated with the Centre for Digital Health Interventions (CDHI), a joint initiative of the Institute for Implementation Science in Health Care, University of Zurich; the Department of Management, Technology, and Economics at ETH Zurich; and the Institute of Technology Management and School of Medicine at the University of St. Gallen. CDHI is funded in part by the Swiss health insurer CSS, the Austrian health care provider (and corporate start-up of UNIQA) Mavie Next, and the Swiss investor MTIP. TK was also a cofounder of Pathmate Technologies, a university spin-off company that creates and delivers digital clinical pathways. However, neither CSS, Mavie Next (UNIQA), nor MTIP was involved in this study. Furthermore, TK has neither shares of Pathmate Technologies nor any formal role in the company. All other authors declare no conflict of interest related to this study.

Multimedia Appendix 1
GRACE interaction script.
[PDF File (Adobe PDF File), 44 KB - aging_v8i1e76489_app1.pdf]

Multimedia Appendix 2
Large language model prompts used for GRACE interaction.
[TXT File , 2 KB - aging_v8i1e76489_app2.txt]

Multimedia Appendix 3
Preinteraction questionnaire (in German).
[PDF File (Adobe PDF File), 88 KB - aging_v8i1e76489_app3.pdf]

Multimedia Appendix 4
Postinteraction questionnaire (in German).
[PDF File (Adobe PDF File), 70 KB - aging_v8i1e76489_app4.pdf]

Multimedia Appendix 5
Interview questions (in German).
[PDF File (Adobe PDF File), 55 KB - aging_v8i1e76489_app5.pdf]

Multimedia Appendix 6
Direct participant quotes for questions 8-10 (translated to English).
[PDF File (Adobe PDF File), 96 KB - aging_v8i1e76489_app6.pdf]

References

1. What is dementia? Alzheimer's Association. 2024. URL: https://www.alz.org/alzheimers-dementia/what-is-dementia [accessed 2025-11-21]
2. Dementia. World Health Organization. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2022-05-25]
3. Global Burden of Disease 2021: findings from the GBD 2021 study. Institute for Health Metrics and Evaluation. 2024. URL: https://www.healthdata.org/research-analysis/library/global-burden-disease-2021-findings-gbd-2021-study [accessed 2025-11-24]
4. Wimo A, Seeher K, Cataldi R, Cyhlarova E, Dielemann JL, Frisell O, et al. The worldwide costs of dementia in 2019. Alzheimers Dement 2023;19(7):2865-2873 [FREE Full text] [doi: 10.1002/alz.12901] [Medline: 36617519]
5. Global action plan on the public health response to dementia. World Health Organization. 2017. URL: https://www.who.int/publications/i/item/global-action-plan-on-the-public-health-response-to-dementia-2017---2025 [accessed 2025-11-24]

XSL•FO
RenderX

6. Schneider C, Nißen M, Kowatsch T, Vinay R. Impact of digital assistive technologies on the quality of life for people with dementia: a scoping review. BMJ Open 2024;14(2):e080545 [FREE Full text] [doi: 10.1136/bmjopen-2023-080545] [Medline: 38341210]
7. Topo P. Technology studies to meet the needs of people with dementia and their caregivers. J Appl Gerontol 2008;28(1):5-37. [doi: 10.1177/0733464808324019]
8. Bérubé C, Schachner T, Keller R, Fleisch E, V Wangenheim F, Barata F, et al. Voice-based conversational agents for the prevention and management of chronic and mental health conditions: systematic literature review. J Med Internet Res 2021;23(3):e25933 [FREE Full text] [doi: 10.2196/25933] [Medline: 33658174]
9. Shade M, Rector K, Kupzyk K. Voice assistant reminders and the latency of scheduled medication use in older adults with pain: descriptive feasibility study. JMIR Form Res 2021;5(9):e26361 [FREE Full text] [doi: 10.2196/26361] [Medline: 34581677]
10. Lappromrattana T, Sooraksa P. Quick prototyping of companion bots for elderly people. Sens Mater 2023;35(4):1487-1495. [doi: 10.18494/SAM4268]
11. Tsiourti C, Moussa MB, Quintas J, Loke B, Jochem I, Lopes JA, et al. A virtual assistive companion for older adults: design implications for a real-world application. 2016 Presented at: Proceedings of SAI Intelligent Systems Conference (IntelliSys) 2016. IntelliSys 2016. Lecture Notes in Networks and Systems, Vol 15; September 21-22, 2016; London, United Kingdom URL: https://link.springer.com/chapter/10.1007/978-3-319-56994-9_69 [doi: 10.1007/978-3-319-56994-9_69]
12. Mahmood A, Cao S, Stiber M, Antony VN, Huang CM. Voice assistants for health self-management: designing for and with older adults. 2025 Presented at: CHI '25: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems; April 26-May 1, 2025; Yokohama, Japan p. 1-22. [doi: 10.1145/3706598.3713839]
13. Cruz-Sandova D, Morales-Tellez A, Sandoval EB, Favela J. A social robot as therapy facilitator in interventions to deal with dementia-related behavioral symptoms. 2020 Presented at: 15th ACM/IEEE International Conference on Human-Robot Interaction (HRI); March 23-26, 2020; Cambridge, United Kingdom p. 161-169 URL: https://ieeexplore.ieee.org/document/9484212
14. Morales-de-Jesús V, Gómez-Adorno H, Somodevilla-García M, Vilariño D. Conversational system as assistant tool in reminiscence therapy for people with early-stage of Alzheimer's. Healthcare (Basel) 2021;9(8):1036 [FREE Full text] [doi: 10.3390/healthcare9081036] [Medline: 34442173]
15. Sezgin E, Militello L, Huang Y, Lin S. A scoping review of patient-facing, behavioral health interventions with voice assistant technology targeting self-management and healthy lifestyle behaviors. Transl Behav Med 2020;10(3):606-628. [doi: 10.1093/tbm/ibz141] [Medline: 32766865]
16. Bickmore T, Cassell J. Social dialongue with embodied conversational agents. In: van Kuppevelt JCJ, Dybkjær L, Bernsen NO, editors. Advances in Natural Multimodal Dialogue Systems. Text, Speech and Language Technology, Vol 30. Dordrecht: Springer; 2005:23-54.
17. Lima MR, Wairagkar M, Gupta M, Rodriguez y Baena F, Barnaghi P, Sharp DJ, et al. Conversational affective social robots for ageing and dementia support. IEEE Trans Cogn Dev Syst 2022;14(4):1378-1397. [doi: 10.1109/tcds.2021.3115228]
18. Saragih ID, Tonapa S, Saragih I, Lee BO. Effects of cognitive stimulation therapy for people with dementia: a systematic review and meta-analysis of randomized controlled studies. Int J Nurs Stud 2022;128:104181. [doi: 10.1016/j.ijnurstu.2022.104181] [Medline: 35149325]
19. Vinay R, Tommila NC, Schlögl M, Klöppel S, Biller-Andorno N, Kowatsch T. Demonstrating GRACE: our embodied voice assistant providing cognitive interventions. 2024 Presented at: Mensch und Computer 2024—Workshopband; September 1-4, 2024; Karlsruhe, Germany. [doi: 10.18420/muc2024-mci-demo-378]
20. Vinay R, Tommila NC, Schlögl M, Klöppel S, Biller-Andorno N, Kowatsch T. GRACE: Towards an embodied voice assistant for improving quality of life by leveraging elements of cognitive stimulation therapy. 2024 Presented at: MuC '24: Proceedings of Mensch und Computer; September 1-4, 2024; Karlsruhe, Germany. [doi: 10.1145/3670653.3677521]
21. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manag Sci 1989;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
22. Davis FD, Davis F. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989;13(3):319-339. [doi: 10.2307/249008]
23. Venkatesh V. Determinants of perceived ease of use: integrating control, intrinsic motivation, and emotion into the technology acceptance model. Inf Syst Res 2000;11(4):342-365. [doi: 10.1287/isre.11.4.342.11872]
24. Falkenström F, Hatcher RL, Skjulsvik T, Larsson MH, Holmqvist R. Development and validation of a 6-item working alliance questionnaire for repeated administrations during psychotherapy. Psychol Assess 2015;27(1):169-183 [FREE Full text] [doi: 10.1037/pas0000038] [Medline: 25346997]
25. Ollier J, Suryapalli P, Fleisch E, von Wangenheim F, Mair JL, Salamanca-Sanabria A, et al. Can digital health researchers make a difference during the pandemic? Results of the single-arm, chatbot-led Elena+: care for COVID-19 interventional study. Front Public Health 2023;11:1185702 [FREE Full text] [doi: 10.3389/fpubh.2023.1185702] [Medline: 37693712]
26. Brill E, Vinay R, Nißen M, Joshi P, Klöppel S, Kowatsch T. Toward a smartphone-based and conversational agent-delivered just-in-time adaptive holistic lifestyle intervention for older adults affected by cognitive decline: two-week proof-of-concept study. JMIR Form Res 2025;9:e66885 [FREE Full text] [doi: 10.2196/66885] [Medline: 40720890]

XSL•FO
RenderX

27. Kowatsch T, Schachner T, Harperink S, Barata F, Dittler U, Xiao G, et al. Conversational agents as mediating social actors in chronic disease management involving health care professionals, patients, and family members: multisite single-arm feasibility study. J Med Internet Res 2021;23(2):e25060 [FREE Full text] [doi: 10.2196/25060] [Medline: 33484114]
28. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 2012;36(1):157-178 [FREE Full text] [doi: 10.2307/41410412]
29. Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment. European Commission. 2020 Jul 17. URL: https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment [accessed 2025-11-26]
30. Harrington CN, Garg R, Woodward A, Williams D. "It's kind of like code-switching": Black older adults' experiences with a voice assistant for health information seeking. 2022 Apr Presented at: CHI '22: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems; April 29-May 5, 2022; New Orleans, LA, United States p. 604. [doi: 10.1145/3491102.3501995]
31. Martin-Hammond A, Vemireddy S, Rao K. Exploring older adults' beliefs about the use of intelligent assistants for consumer health information management: a participatory design study. JMIR Aging 2019;2(2):e15381 [FREE Full text] [doi: 10.2196/15381] [Medline: 31825322]
32. Blocker KA, Kadylak T, Koon LM, Kovac CE, Rogers WA. Digital home assistants and aging: initial perspectives from novice older adult users. Proc Hum Factors Ergon Soc Annu Meet 2021;64(1):1367-1371. [doi: 10.1177/1071181320641327]
33. Koon LM, McGlynn SA, Blocker KA, Rogers WA. Perceptions of digital assistants from early adopters aged 55+. Ergon Des 2019;28(1):16-23. [doi: 10.1177/1064804619842501]
34. Yan Z, Dube V, Heselton J, Johnson K, Yan C, Jones V, et al. Understanding older people's voice interactions with smart voice assistants: a new modified rule-based natural language processing model with human input. Front Digit Health 2024;6:1329910 [FREE Full text] [doi: 10.3389/fdgth.2024.1329910] [Medline: 38812806]
35. Mitzner TL, Boron JB, Fausset CB, Adams AE, Charness N, Czaja SJ, et al. Older adults talk technology: technology usage and attitudes. Comput Human Behav 2010;26(6):1710-1721 [FREE Full text] [doi: 10.1016/j.chb.2010.06.020] [Medline: 20967133]
36. Razavi SZ, Schubert LK, van Orden K, Ali MR, Kane B, Hoque E. Discourse behavior of older adults interacting with a dialogue agent competent in multiple topics. ACM Trans Interact Intell Syst 2022;12(2):1-21. [doi: 10.1145/3484510]
37. Wolters M, Georgila K, Moore JD, MacPherson SE. Being old doesn't mean acting old: how older users interact with spoken dialog systems. ACM Trans Access Comput 2009;2(1):1-39. [doi: 10.1145/1525840.1525842]
38. Schlögl S, Chollet G, Garschall M, Tscheligi M, Legouverneur G. Exploring voice user interfaces for seniors. 2013 Presented at: PETRA '13: Proceedings of the 6th International Conference on PErvasive Technologies Related to Assistive Environments; May 29-31, 2013; Rhodes, Greece. [doi: 10.1145/2504335.2504391]
39. Nallam P, Bhandari S, Sanders J, Martin-Hammond A. A question of access: exploring the perceived benefits and barriers of intelligent voice assistants for improving access to consumer health resources among low-income older adults. Gerontol Geriatr Med 2020;6:2333721420985975 [FREE Full text] [doi: 10.1177/2333721420985975] [Medline: 33457459]
40. Joshi KR, Ulabhaje AA, Nataraj R, Martin-Hammond A. Skeptical yet curious: attitudes towards conversational assistants for supporting social engagement among older adults living alone. 2025 Presented at: CHI EA '25: Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems; April 26-May 1, 2025; Yokohama, Japan. [doi: 10.1145/3706599.3719917]
41. Huang Y, Zhou Q, Piper AM. Designing conversational AI for aging: a systematic review of older adults' perceptions and needs. 2025 Presented at: CHI '25: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems; April 26-May 1, 2025; Yokohama, Japan p. 1-20. [doi: 10.1145/3706598.3713578]
42. Moyle W. The promise of technology in the future of dementia care. Nat Rev Neurol 2019;15(6):353-359. [doi: 10.1038/s41582-019-0188-y] [Medline: 31073242]
43. Marziali RA, Franceschetti C, Dinculescu A, Nistorescu A, Kristály DM, Mo oi AA, et al. Reducing loneliness and social isolation of older adults through voice assistants: literature review and bibliometric analysis. J Med Internet Res 2024;26:e50534 [FREE Full text] [doi: 10.2196/50534] [Medline: 38498039]
44. Chen Y, Zeng J, Liu Y, Zhang J, Wu X, Li H. Digital radio: a one-touch interactive system for elderly companionship. In: Ying F, Jain LC, Wan R, Wu Q, Shi F, editors. Frontiers in Artificial Intelligence and Applications. Amsterdam: IOS Press; 2024:904-913.
45. Rincon JA, Marco-Detchart C, Julian V, Carrascosa C, Novais P. Towards a low-cost companion robot for helping elderly well-being. 2022 Presented at: Intelligent Data Engineering and Automated Learning—IDEAL 2022: 23rd International Conference, IDEAL 2022, Manchester, UK; November 24-26, 2022; Manchester, United Kingdom. [doi: 10.1007/978-3-031-21753-1_18]
46. Miehle J, Bagci I, Minker W, Ultes S. A social companion and conversational partner for the elderly. In: Eskenazi M, Devillers L, Mariani J, editors. Advanced Social Interaction with Agents. Lecture Notes in Electrical Engineering, Vol 510. Cham: Springer; 2019.

XSL•FO
RenderX

47. Sung H, Chang S, Chin M, Lee W. Robot-assisted therapy for improving social interactions and activity participation among institutionalized older adults: a pilot study. Asia Pac Psychiatry 2015;7(1):1-6. [doi: 10.1111/appy.12131] [Medline: 24692085]
48. Striegl J, Gotthardt M, Loitsch C, Weber G. Investigating the usability of voice assistant-based CBT for age-related depression. 2022 Presented at: Computers Helping People with Special Needs: 18th International Conference, ICCHP-AAATE 2022, Proceedings, Part I; July 11-15, 2022; Lecco, Italy p. 432-441. [doi: 10.1007/978-3-031-08648-9_50]
49. Wiratunga N, Cooper K, Wijekoon A, Palihawadana C, Mendham V, Reiter E, et al. FitChat: conversational artificial intelligence interventions for encouraging physical activity in older adults. ArXiv Preprint posted online on April 29, 2020 [FREE Full text] [doi: 10.48550/arXiv.2004.14067]
50. Favela J, Cruz-Sandoval D, Parra MO. Conversational agents for dementia using large language models. 2023 Presented at: 2023 Mexican International Conference on Computer Science (ENC); September 11-13, 2023; Guanajuato, Mexico URL: https://ieeexplore.ieee.org/document/10508610/
51. Utami D, Bickmore T, Nikolopoulou A, Paasche-Orlow M. Talk about death: end of life planning with a virtual agent. 2017 Presented at: Intelligent Virtual Agents. IVA 2017. Lecture Notes in Computer Science, Vol 10498; August 27-30, 2017; Stockholm, Sweden. [doi: 10.1007/978-3-319-67401-8_55]
52. Leusmann J, Wang C, Wang G. Comparing Rule-Based and LLM-Based Methods to Enable Active Robot Assistant Conversations. URL: https://cui.acm.org/workshops/CHI2024/wp-content/uploads/2024/04/Comparing-Rule-based-and-LLM-based-Methods-to-Enable-Active-Robot-Assistant-Conversations.pdf [accessed 2025-11-21]
53. Porcheron M, Fischer JE, Reeves S, Sharples S. Voice interfaces in everyday life. 2018 Presented at: CHI '18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems; April 21-26, 2018; Montreal, Quebec, Canada. [doi: 10.1145/3173574.3174214]
54. Myers C, Furqan A, Nebolsky J, Caro K, Zhu J. Patterns for how users overcome obstacles in voice user interfaces. 2018 Presented at: CHI '18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems; April 21-26, 2018; Montreal, Quebec, Canada. [doi: 10.1145/3173574.3173580]
55. Kadam AV. Using knowledge graphs and LLMs to enhance natural language understanding on voice assistants. Int J Comput Appl 2023;185(27):20-24. [doi: 10.5120/ijca2023923019]
56. Sheng Y, Zheng L, Yuan B, Li Z, Ryabinin M, Chen B, et al. FlexGen: high-throughput generative inference of large language models with a single GPU. 2023 Presented at: ICML'23: Proceedings of the 40th International Conference on Machine Learning; July 23-29, 2023; Honolulu, HI, United States p. 31094-31116 URL: https://dl.acm.org/doi/abs/10.5555/3618408.3619696
57. Zhang Y, Li Y, Cui L, Cai D, Liu L, Fu T, et al. Siren's song in the AI ocean: a survey on hallucination in large language models. ArXiv Preprint posted online on September 14, 2025 [FREE Full text] [doi: 10.48550/arXiv.2309.01219]
58. Bender EM, Gebru T, McMillan-Major A, Shmitchell S. On the dangers of stochastic parrots: can language models be too big? 2021 Presented at: FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency; March 3-10, 2021; Virtual Event p. 610-623. [doi: 10.1145/3442188.3445922]
59. Hastings J. Preventing harm from non-conscious bias in medical generative AI. Lancet Digit Health 2024;6(1):e2-e3 [FREE Full text] [doi: 10.1016/S2589-7500(23)00246-7] [Medline: 38123253]
60. An K, Chen Q, Deng C, Du Z, Gao C, Gao Z, et al. FunAudioLLM: voice understanding and generation foundation models for natural interaction between humans and LLMs. ArXiv Preprint posted online on July 11, 2024 [FREE Full text] [doi: 10.48550/arXiv.2407.04051]
61. Schalkwyk J, Kumar A, Lyth D. SesameAILabs: a conversational speech generation model. GitHub. URL: https://github.com/SesameAILabs/csm [accessed 2025-05-07]
62. Mahmood A, Wang J, Yao B, Wang D, Huang CM. User interaction patterns and breakdowns in conversing with LLM-powered voice assistants. Int J Hum Comput Stud 2025 Jan;195(103406):1-21. [doi: 10.1016/j.ijhcs.2024.103406]
63. Kim CY, Lee CP, Mutlu B. Understanding large-language model (LLM)-powered human-robot interaction. 2024 Mar 11 Presented at: HRI '24: Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction; March 11-14, 2024; Boulder, CO, United States p. 371-380 URL: https://dl.acm.org/doi/10.1145/3610977.3634966 [doi: 10.1145/3610977.3634966]
64. Lima MR, O'Connell A, Zhou F, Nagahara, A, Hulyalkar A, Deshpande A, et al. Promoting cognitive health in elder care with large language model-powered socially assistive robots. 2025 Presented at: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25); April 26-May 1, 2025; Yokohama, Japan p. 1-22. [doi: 10.1145/3706598.3713582]
65. Padmanabha A, Yuan J, Gupta J, Karachiwalla Z, Majidi C, Admoni H, et al. VoicePilot: harnessing LLMs as speech interfaces for physically assistive robots. 2024 Presented at: UIST '24: Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology; October 13-16, 2024; Pittsburgh, PA, United States. [doi: 10.1145/3654777.3676401]
66. Yang Z, Xu X, Yao B, Rogers E, Zhang S, Intille S, et al. Talk2Care: an LLM-based voice assistant for communication between healthcare providers and older adults. Proc ACM Interact Mob Wearable Ubiquitous Technol 2024;8(2):1-35. [doi: 10.1145/3659625]

XSL•FO
RenderX

67. Marques da Rocha M, Cruz-Sandoval D, Favela J, Muchaluat-Saade DC. Design and usability evaluation of the EvaSIM simulator for a socially assistive robot. Multimed Tools Appl 2024;83:69897-69922. [doi: 10.1007/s11042-024-18338-1]
68. OpenAI Platform. 2024. URL: https://platform.openai.com/ [accessed 2025-11-21]
69. EVA Social Robot. Github. 2024. URL: https://github.com/eva-social-robot/eva-social-robot.github.io [accessed 2025-11-21]
70. ElevenLabs: AI Voice Generator. 2024. URL: https://elevenlabs.io/ [accessed 2025-11-21]
71. Spector A, Orrell M, Davies S, Woods B. Can reality orientation be rehabilitated? Development and piloting of an evidence-based programme of cognition-based therapies for people with dementia. Neuropsychol Rehabil 2010;11(3-4):377-397. [doi: 10.1080/09602010143000068]
72. Spector A, Thorgrimsen L, Woods B, Royan L, Davies S, Butterworth M, et al. Efficacy of an evidence-based cognitive stimulation therapy programme for people with dementia: randomised controlled trial. Br J Psychiatry 2003;183:248-254. [doi: 10.1192/bjp.183.3.248] [Medline: 12948999]
73. Yates L, Orrell M, Leung P, Spector A, Woods B, Orgeta V. Making a Difference 3: Individualised CST? A Manual for Carers. Greater London: Hawker Publications; 2014.
74. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000;55(1):68-78. [doi: 10.1037//0003-066x.55.1.68] [Medline: 11392867]
75. Balban MY, Neri E, Kogon MM, Weed L, Nouriani B, Jo B, et al. Brief structured respiration practices enhance mood and reduce physiological arousal. Cell Rep Med 2023;4(1):100895 [FREE Full text] [doi: 10.1016/j.xcrm.2022.100895] [Medline: 36630953]
76. van der Heijden H. User acceptance of hedonic information systems. MIS Q 2004;28(4):695-704. [doi: 10.2307/25148660]
77. Del Re AC, Flückiger C, Horvath AO, Wampold BE. Examining therapist effects in the alliance-outcome relationship: a multilevel meta-analysis. J Consult Clin Psychol 2021;89(5):371-378. [doi: 10.1037/ccp0000637] [Medline: 33829817]
78. Flückiger C, Del Re AC, Wampold BE, Symonds D, Horvath AO. How central is the alliance in psychotherapy? A multilevel longitudinal meta-analysis. J Couns Psychol 2012;59(1):10-17. [doi: 10.1037/a0025749] [Medline: 21988681]
79. Flückiger C, Del Re AC, Wampold BE, Horvath AO. The alliance in adult psychotherapy: a meta-analytic synthesis. Psychotherapy (Chic) 2018;55(4):316-340 [FREE Full text] [doi: 10.1037/pst0000172] [Medline: 29792475]
80. Horvath AO, Greenberg LS. Development and validation of the Working Alliance Inventory. J Couns Psychol 1989;36(2):223-233. [doi: 10.1037/0022-0167.36.2.223]
81. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 2025;16:297-334. [doi: 10.1007/BF02310555]
82. Kowatsch T, Lohse K, Erb V, Schittenhelm L, Galliker H, Lehner R, et al. Hybrid ubiquitous coaching with a novel combination of mobile and holographic conversational agents targeting adherence to home exercises: four design and evaluation studies. J Med Internet Res 2021;23(2):e23612 [FREE Full text] [doi: 10.2196/23612] [Medline: 33461957]
83. Smith AL, Chaparro BS. Smartphone text input method performance, usability, and preference with younger and older adults. Hum Factors 2015;57(6):1015-1028. [doi: 10.1177/0018720815575644] [Medline: 25850116]
84. Kowalski J, Jaskulska A, Skorupska K, Abramczuk K, Biele C, Kopeć W, et al. Older adults and voice interaction: a pilot study with Google home. 2019 Presented at: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems (CHI EA '19); May 4-9, 2019; Glasgow, Scotland, United Kingdom URL: https://dl.acm.org/doi/10.1145/3290607.3312973 [doi: 10.1145/3290607.3312973]
85. Qiu L, Saragih ID, Fick DM, Sundar SS, Abdullah S. Voice assistants to deliver cognitive stimulation therapy for persons living with dementia. 2025 Presented at: Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25); April 26-May 1, 2025; Yokohama, Japan. [doi: 10.1145/3706599.3719707]
86. Manzini A, Keeling G, Alberts L, Vallor S, Morris MR, Gabriel I. The code that binds us: navigating the appropriateness of human-AI assistant relationships. Proc AAAI/ACM Conf AI Ethics Soc 2024;7:943-957. [doi: 10.1609/aies.v7i1.31694]
87. Chen G. The role of AI driven personal assistants in geriatric care: opportunities, challenges, and future directions. Ann Gerontol Geriatr Res 2025;1(1):1003. [doi: 10.52768/3068-8035/1003]
88. Martell MJ, Baweja JA, Dreslin BD. Mitigative strategies for recovering from large language model trust violations. J Cogn Eng Decis Mak 2024;19(1):76-95. [doi: 10.1177/15553434241303577]
89. Lima MR, Horrocks S, Daniels S, Lamptey M, Harrison M, Vaidyanathan R. The role of conversational AI in ageing and dementia care at home: a participatory study. 2023 Presented at: 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN); August 28-31, 2023; Busan, Korea, Republic of URL: https://ieeexplore.ieee.org/document/10309459
90. Bosco C, Otenen E, Osorio Torres J, Nguyen V, Chheda D, Peng X, et al. Designing a multimodal and culturally relevant Alzheimer disease and related dementia generative artificial intelligence tool for Black American informal caregivers: cognitive walk-through usability study. JMIR Aging 2025;8:e60566 [FREE Full text] [doi: 10.2196/60566] [Medline: 39778201]
91. Bosco C, Shojaei F, Theisz AA, Osorio Torres J, Cureton B, Himes AK, et al. Testing 3 modalities (voice assistant, chatbot, and mobile app) to assist older African American and Black adults in seeking information on Alzheimer disease and related dementias: Wizard of Oz usability study. JMIR Form Res 2024;8:e60650 [FREE Full text] [doi: 10.2196/60650] [Medline: 39653372]

XSL•FO
RenderX

92. Harrington CN, Egede L. Trust, comfort and relatability: understanding Black older adults' perceptions of chatbot design for health information seeking. 2023 Presented at: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23); April 23-28, 2023; Hamburg, Germany p. 1-18. [doi: 10.1145/3544548.3580719]
93. Moharana S, Washington G, Carrington P, Harrington CN. Towards contextualizing embodiment of intelligent systems. In: Arai K, editor. Intelligent Systems and Applications. IntelliSys 2024. Lecture Notes in Networks and Systems, vol 1065. Cham: Springer; 2024:109-116.
94. Douglas S, James IA, Ballard C. Non-pharmacological interventions in dementia. Adv Psychiatr Treat 2018 Jan 02;10(3):171-177. [doi: 10.1192/apt.10.3.171]
95. Phillips LJ, Reid-Arndt S, Pak Y. Effects of a creative expression intervention on emotions, communication, and quality of life in persons with dementia. Nurs Res 2010;59(6):417-425 [FREE Full text] [doi: 10.1097/NNR.0b013e3181faff52] [Medline: 21048483]
96. Ma J, Wang Q, Lang Y, Lv S, Xu Y, Wei B. Effectiveness of creative story therapy for dementia: a systematic review and meta-analysis. Eur J Med Res 2023;28(1):342 [FREE Full text] [doi: 10.1186/s40001-023-01337-7] [Medline: 37705077]
97. Grimm P. Social desirability bias. In: Sheth J, Malhotra N, editors. Wiley International Encyclopedia of Marketing. Hoboken, NJ: John Wiley & Sons; 2010.
98. Plüss M, Deriu J, Schraner Y, Paonessa C, Hartmann J, Schmidt L, et al. A speech corpus for all Swiss German dialect regions. 2023 Presented at: Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers); July 1, 2023; Toronto, Ontario, Canada p. 1763-1772 URL: https://aclanthology.org/2023.acl-short.150/

Abbreviations

AI: artificial intelligence
CST: cognitive stimulation therapy
ITI: intention to continue interacting
LLM: large language model
PC: perceived control
PEN: perceived enjoyment
PEU: perceived ease of use
PSA: perceived session alliance
PU: perceived usefulness
VA: voice assistant

Edited by M Mardini; submitted 09.05.25; peer-reviewed by C Bosco, A Carrillo; comments to author 22.07.25; revised version received 19.09.25; accepted 10.11.25; published 18.12.25.

Please cite as:
Vinay R, Uetova E, Tommila NC, Biller-Andorno N, Kowatsch T
A Hybrid Rule- and Large Language Model–Based Embodied Voice Assistant (GRACE) for Cognitive Stimulation in Older Adults: Usability Study Assessing Technical Feasibility, Technology Acceptance, and Working Alliance
JMIR Aging 2025;8:e76489
URL: https://aging.jmir.org/2025/1/e76489
doi:10.2196/76489
PMID:

©Rasita Vinay, Ekaterina Uetova, Nora Camilla Tommila, Nikola Biller-Andorno, Tobias Kowatsch. Originally published in JMIR Aging (https://aging.jmir.org), 18.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

An Accessible Multifunctional System to Support Safe and Independent Aging in Place: Iterative Development and Qualitative Analysis

Céline Madeleine Aldenhoven[1,2], BSc, MSc; Matthew Milton[2,3], BA&Sc, MS; Lisa D'Ambrosio[2], AB, PhD; Chaiwoo Lee[2], BS, MS, PhD; Sophia Ashebir[2], BA; Elisabeth André[4], Dr rer nat

[1]Software Engineering (Elite Graduate Program), Technical University of Munich, Ludwig Maximilian University, University of Augsburg, Munich and Augsburg, Germany

[2]AgeLab, Massachusetts Institute of Technology, 77 Massachusetts Ave, E40-275, Cambridge, MA, United States

[3]Department of Mechanical Engineering, Tufts University, Medford, MA, United States

[4]Chair for Human Centered Artificial Intelligence, Faculty of Applied Computer Science, University of Augsburg, Augsburg, Germany

Corresponding Author:
Céline Madeleine Aldenhoven, BSc, MSc
Software Engineering (Elite Graduate Program), Technical University of Munich, Ludwig Maximilian University, University of Augsburg, Munich and Augsburg, Germany

Abstract

Background: The global population of older adults is rapidly increasing, while the number of relative caregivers is declining. This creates a critical need for solutions that support caregiving and enable older adults to age in place while maintaining their independence. Many existing caregiving technologies focus on easing caregivers' burdens through surveillance-based systems, which often prioritize caregivers' needs over those of older adults. Such designs can unintentionally disempower older adults by devaluing their autonomy and decision-making capabilities.

Objective: This study explored older adults' and designers' reactions to a technology-enabled system called LifeTomorrow that was built by centering older adults as the primary users. The system aims to support their autonomy to make informed choices and their desire for independent living, while balancing their social and functional needs with those of informal caregivers.

Methods: A total of 37 participants, including older adults, caregivers, and designers, engaged in 2 iterative rounds of user studies to explore daily caregiving needs and technology usage. The system's design and features were refined based on these insights. A qualitative thematic analysis was conducted using the framework of the self-determination theory to evaluate how the system fulfills the basic psychological needs of competence, autonomy, and relatedness.

Results: The analysis underscored the system's value to older adults and their caregivers, and its fulfillment of basic human psychological needs (competence, autonomy, and relatedness), consistent with the goals of supporting a high quality of life for older users and caregivers. We found that relatedness is fostered through features enabling remote connection and communication, such as chat functions and shared health data. Autonomy is supported by empowering older adults to manage their health information, make informed choices about data sharing, and benefit from safety features like fall detection and emergency calls. Competence is enhanced through accessible design elements, including intuitive navigation, high-contrast visuals, and multigenerational usability. These features allow older adults and caregivers to confidently engage with the system and are targeted at improving their overall quality of life.

Conclusions: Through evaluation of the LifeTomorrow system, this study suggests possibilities for using a holistic, inclusive solution to support safe and independent aging in place and prioritizing the autonomy and empowerment of older adults while addressing caregivers' needs for support and connection. By centering older adults as active participants rather than passive recipients of care, the system exemplifies a shift toward equitable, user-centered technology in caregiving. Future research should investigate the long-term impacts of such systems on aging-in-place outcomes.

(JMIR Aging 2025;8:e72579) doi:10.2196/72579

KEYWORDS

gerontology; aging in place; older adults; informal caregivers; mobile apps; self-determination theory; iterative design; qualitative text analysis

XSL•FO
RenderX

Introduction

Background and Motivation

In the United States, 77% of adults aged 50 years or older desire to age in place, living independently in their homes for as long as possible [1]. However, an estimated 70% of people in the United States aged 65 years or older will require ongoing care and support at some point [2], often in their homes. This care is most often provided by informal or family caregivers [3]. Caregivers may assist their care recipients with a range of different tasks, from helping with instrumental activities of daily living (eg, transportation and managing finances) to providing direct health care support [4]. As a result of their caregiving, caregivers often experience negative impacts on their own physical and emotional well-being (eg, [5-7]). As the global population of older adults and their associated care needs rapidly increase, the demand for caregiving is also projected to grow [8-10]. Between the increasing demand for caregiving and the burden of caregiving on caregivers, there is an urgent need for solutions that support caregiving while respecting older adults' desires for autonomy and independence [11].

Technology offers significant potential to address these challenges, yet many existing systems are designed primarily with caregivers' needs in mind. These systems often prioritize reducing caregiving burden through surveillance and monitoring technologies, implicitly disempowering older adults by framing them as passive recipients of care [12-15]. Moreover, such solutions are frequently intrusive and expensive, and fail to center older adults as autonomous individuals capable of making decisions about their own care.

This study takes an alternative approach, examining user responses to a technology-enabled system designed with older adults' needs at its core. We build upon previous knowledge that many older adults are open to adopting mobile or wearable health technologies, especially to share data with their health care providers [16]. The LifeTomorrow system aims to empower healthy older adults to live safely and independently in their homes while supporting informal caregivers. By fostering communication, enabling selective sharing of health information, and providing caregivers with access to resources, the system was built to try to address the needs of both older adults and caregivers without relying on direct surveillance or intrusive monitoring.

Through 2 rounds of user studies, we explored the reactions of older adults, informal caregivers, and designers to the LifeTomorrow system. Using self-determination theory [17,18] as an analytical framework, we examined the extent to which the LifeTomorrow system fosters competence, autonomy, and relatedness, which are fundamental psychological needs critical to maintaining a high quality of life for older adults and their informal caregivers.

Prior Work

Technologies for Aging in Place

Aging in place refers to living within one's community with some level of independence for as long as possible, rather than entering residential care in later life [19]. To do so safely requires affordable, unobtrusive, and easy-to-use solutions for older adults. Health-related innovations in human-computer interaction (HCI) research have great potential to enhance older adults' access to health care resources while supporting long-term independence at home [20]. By developing intuitive, personalized, and accessible technologies, HCI can bridge gaps in health care delivery, promote self-management, and facilitate aging in place. For instance, voice assistants tailored for older adults can assist with health self-management tasks, such as debriefing doctor's visit notes and generating medication reminders, enhancing adherence and understanding. Mobile health apps designed with input from older adults and their support networks can facilitate communication with caregivers and provide timely health information, promoting engagement and autonomy [21].

Despite the common stereotype of older adults not using technology, as of 2021, 61% of older adults in the United States had smartphones [22], and 70% of older adults who are online use technology daily [23]. Older adults' intentions to use smart home technologies, which offer the potential to support greater in-home care and foster independence, are influenced by factors such as ease of use, accessibility, privacy, security, and cost [24]. Despite older adults' openness to health-monitoring technologies, barriers like steep learning curves and usability concerns remain significant challenges [25-28]. Designing technologies that emphasize simplicity, intuitive interfaces, and comfort is essential to promote adoption. Additionally, systems should accommodate support needs, such as setup assistance and ongoing technical help [29].

Several smart home care systems use cameras or voice recognition [30-36], but these raise privacy and security concerns. For example, Trajkova et al [37] found that older adults often limit or abandon voice assistants due to challenges with finding valuable uses, beliefs, and shared spaces.

A number of different systems also depend on more complicated architectures and physical components. For example, numerous studies have explored the use of a large number of sensors to support aging in place [38-41] and monitor daily activity phases [42]. While these systems may provide caregivers with valuable insights, they typically emphasize the need for caregivers to know where their care recipient is and what they are doing, overlooking older adults' autonomy. Moreover, many solutions rely on large sensor networks or machine learning systems that require substantial infrastructure or assume routine behavior patterns, which may not suit more active older adults [43,44].

Technologies to Support Caregivers

Remote caregiving technologies primarily aim to reduce caregivers' burdens by providing tools to monitor health, safety, and well-being from a distance [45]. Wearable devices and in-home monitoring systems can share vital health and safety information with caregivers and health care providers [46,47]. Additionally, internet-based interventions, such as virtual support groups and educational resources, have been shown to improve caregiver well-being by helping to manage caregivers' depression, reduce stress and anxiety, improve self-efficacy, and increase confidence [47-50].

XSL•FO
RenderX

Despite the benefits they may offer, much of the existing work on caregiving technologies focuses on narrower, piecemeal solutions addressing specific issues such as fall detection [51], activity monitoring [52], or communication and social isolation [49,53]. Systems often lack the capacity to address broader caregiving challenges, such as caregivers' mental health or access to resources. One exception is the SOS caregiver service by Graffigna et al [54], which offers a promising model for reducing caregiver isolation and providing access to knowledge, but most solutions fail to provide the same level of holistic support [55].

Supported Needs and Limitations

Because many current systems are designed around caregivers' needs and desires for information about the status of the care recipient, older adults may feel surveilled or disempowered and like their autonomy has been undermined. Privacy, stigma, and usability challenges further complicate adoption [26,56,57].

While collecting and sharing personal health data from older adults with caregivers may enable more customized care [45,58], concerns remain about reduced face-to-face interaction. It is vital to enhance, not replace, personal contact to prevent loneliness and social isolation [19,58].

For technological systems to achieve widespread adoption, they must also be cost-effective. Projects like the smart home system by Raad et al [39] and the SIMBAD project [59] are efforts to address this by using low-cost remote monitoring systems and fall detection sensors, respectively. Similarly, the survey by Visutsak et al [60] highlights the preference for wearable devices like smartwatches to support independent living within constrained budgets. While these examples provide critical functionality at a lower cost, they still largely focus on reactive or deficit-oriented designs, addressing specific concerns such as falls or emergencies.

In this research, we evaluated user responses to a system that follows a holistic design centering on the needs of older adults [47]. The approach to the system design emphasized empowering older adults as autonomous individuals while supporting informal caregivers with practical tools and resources. Unlike many existing systems, this system aims to avoid intrusive surveillance, respect privacy, and be cost-effective in order to improve accessibility and adoption. The research here explores users' and designers' reactions to 2 iterations of the prototype system designed to center and support human agency.

Analysis Framework: Self-Determination Theory

Self-determination theory was proposed by Deci and Ryan [61] in 1985, with a focus on an individual's ability to self-assure, make decisions, and think independently. It suggests that individuals need to meet their basic psychological needs, which include maintaining social connections (relatedness), feeling competent (competence), and self-regulating feelings and behaviors (autonomy) [17]. Satisfying these needs leads to a positive emotion and a sense of subjective well-being in a person's growth and development [17,61].

Self-determination theory is crucial in understanding HCI and technology use. Studies on gaming often use it to understand players' demand fulfillment, intrinsic drive, and addiction-like sensations [62], but in recent years, it has also been applied to people's interactions with whole-body control interfaces [63], artificial intelligence systems [64,65], and robots [66,67]. For example, Schwaninger et al [68] identified various types of relatedness, such as interpersonal and institutional, in people's usage of remote health monitoring technology. Zhao et al [18] applied self-determination theory to explore how using digital technologies for meaningful activities can support or undermine older adults' needs for autonomy, competence, and relatedness. Yang et al [65] leveraged self-determination theory to explore how users' competence, autonomy, and relatedness needs could be supported or undermined in experiences with conversational agents.

Competence, relatedness, and autonomy are also predictive and reliable mediators of motivation, engagement, and well-being [17,69]. The design of interactive applications has been effectively guided by self-determination theory to produce positive experiences [18,65,70,71]. In this work, we used self-determination theory to analyze how our proposed system contributes to fulfilling 3 basic human psychological needs (competence, autonomy, and relatedness) [17], consistent with the goals of supporting agency for older users and a high quality of life for them and their informal caregivers.

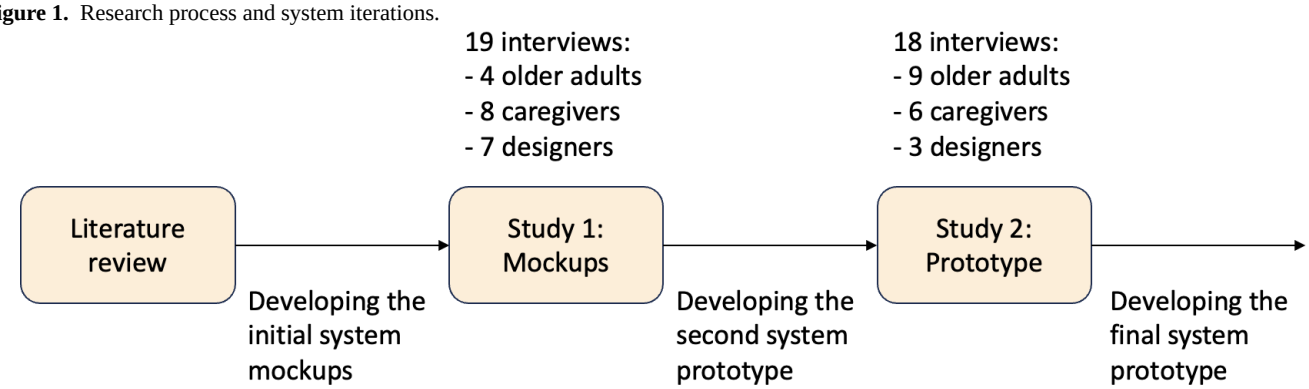
Methods

Study Procedure

Two rounds of user studies were conducted iteratively to improve the system design. Semistructured interviews facilitated participants' sharing of personal experiences and allowed tailored follow-up questions [72]. We built on an approach for analyzing think-aloud usability test sessions [73], which allowed people to engage with and react to the system design and prompts naturally and to respond using their own thoughts and words. The research process is outlined in Figure 1, and example questions from the interviews can be found in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 1. Research process and system iterations.



The first round of user studies for mockup evaluation was conducted online via Zoom (Zoom Communications, Inc), enabling broad participant inclusion regardless of location. After providing informed consent and demographic information, participants received an overview of the LifeTomorrow system. Older adults and informal caregivers reviewed mockups specific to their roles, while designers evaluated the full system. Participants provided feedback on the design and usability of each screen, discussed privacy concerns, and shared overall impressions. Concluding questions explored their likelihood of using or recommending the system.

The second round of user studies for prototype evaluation was conducted in person at the Massachusetts Institute of Technology (MIT) campus, enabling hands-on interaction with the implemented prototype. Participants provided written informed consent and demographic information on large-print forms. Older adults and informal caregivers tested their respective prototype apps on devices (iPad and Apple Watch for older adults; iPhone for informal caregivers) by completing tasks in the system with only a few sentences of introduction to their purpose. Tasks included finding specific information, sharing data, and navigating features, with task order randomized. Designers reviewed the complete system. Participants provided feedback on individual screens, privacy concerns, and system updates since round 1. Final questions addressed usability and willingness to recommend the system.

Study Stimulus: Overview of the LifeTomorrow System

The LifeTomorrow system is a multifunctional app designed to support independent living for older adults and their informal caregivers [74]. The system prioritizes support for older adults' independence [1]. Moving away from "elderly surveillance systems" [13], it allows older adults to manage their health data independently, monitor their personal data sharing, and grant or revoke access to their data at any time [12-14].

System functionalities include fall detection and communication, monitoring and display of vital health signs, emergency communication, and environmental safety (fire, smoke, and carbon monoxide detection), with an overall design flexible enough to enable the addition of future sensing and monitoring devices.

For older adults, an Apple Watch with automatic fall detection and emergency call capabilities is used, removing the need for specialized emergency call systems. The smartwatch's cellular connectivity and health tracking ability (eg, step counts, heart rate, and blood oxygen monitoring) eliminate the need for a smartphone or separate tracking device. The system includes a network of sensors in the older adult's home, including smart carbon monoxide detectors and a stove sensor. Collectively, the system provides a safer home environment, with capabilities to detect and notify events such as falls and fire hazards [75,76], which disproportionately have detrimental impacts on older adults. The system is further designed with flexibility to add other smart sensors, such as a door contact sensor, using the Matter framework [77,78], which allows interconnectivity between smart sensors from different manufacturers for different platforms. Overall, the system leverages products that are widely available and used, in order to make the system more accessible for economically less secure older adults [74,79] compared to large sensor networks [43,44].

The data from the Apple Watch can be displayed on an iPad using Apple's secure health data application programming interface, HealthKit [80], on iPad OS 17 Developer Beta, along with the data from the sensors. Inspired by previous research [81-83], we focused on an accessible design for older adults, including easy and intuitive navigation through the app. The app's interface is designed with an accessible color palette and high contrast, ensuring a better user experience. The 2 rounds of user studies were conducted to improve and redesign user interface (UI) visualizations and interactions as needed (Figure 1). Example screens after each round of user studies illustrating the system's evolution can be found in Figures 2 and 3.

XSL•FO
RenderX

Figure 2. Initial system used for round 1: (A) Caregiver's LifeTomorrow app; (B) Older adult's LifeTomorrow app.

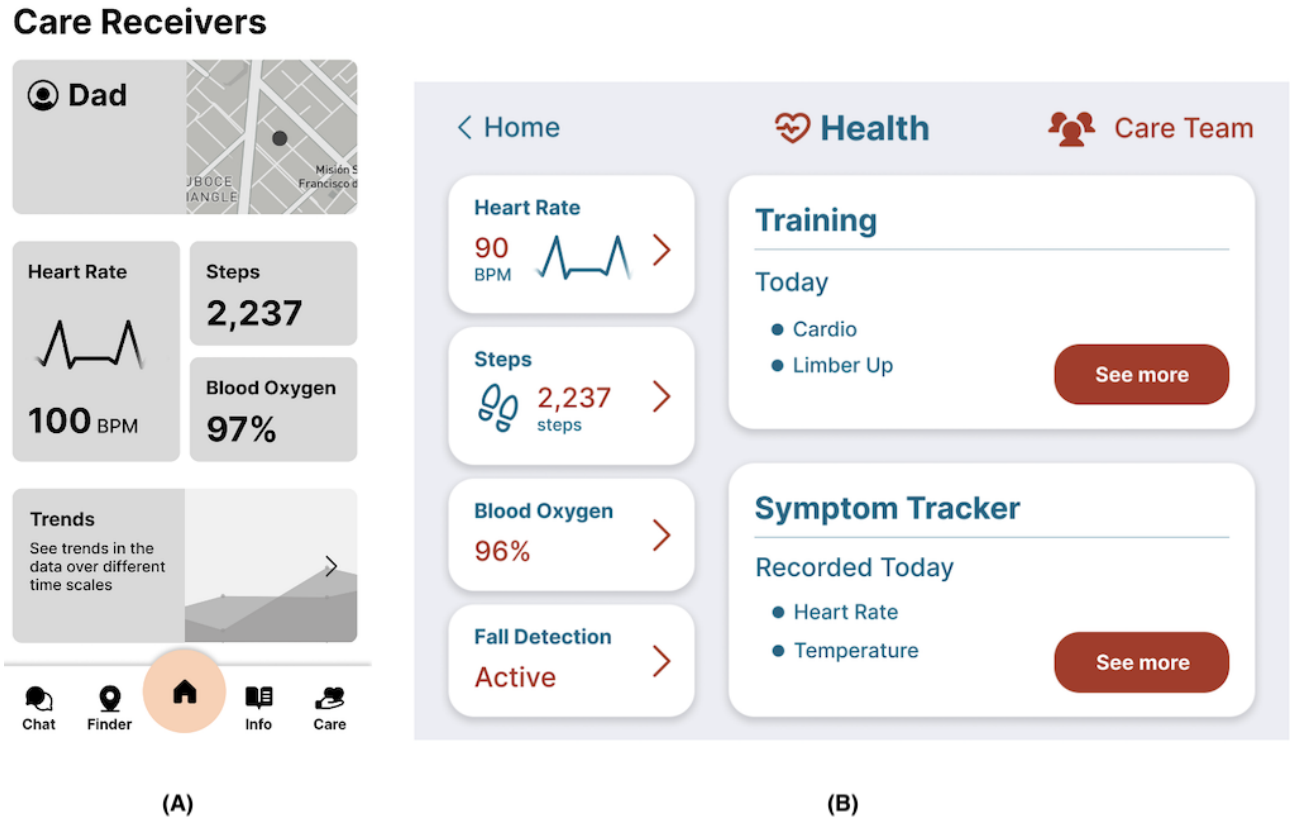
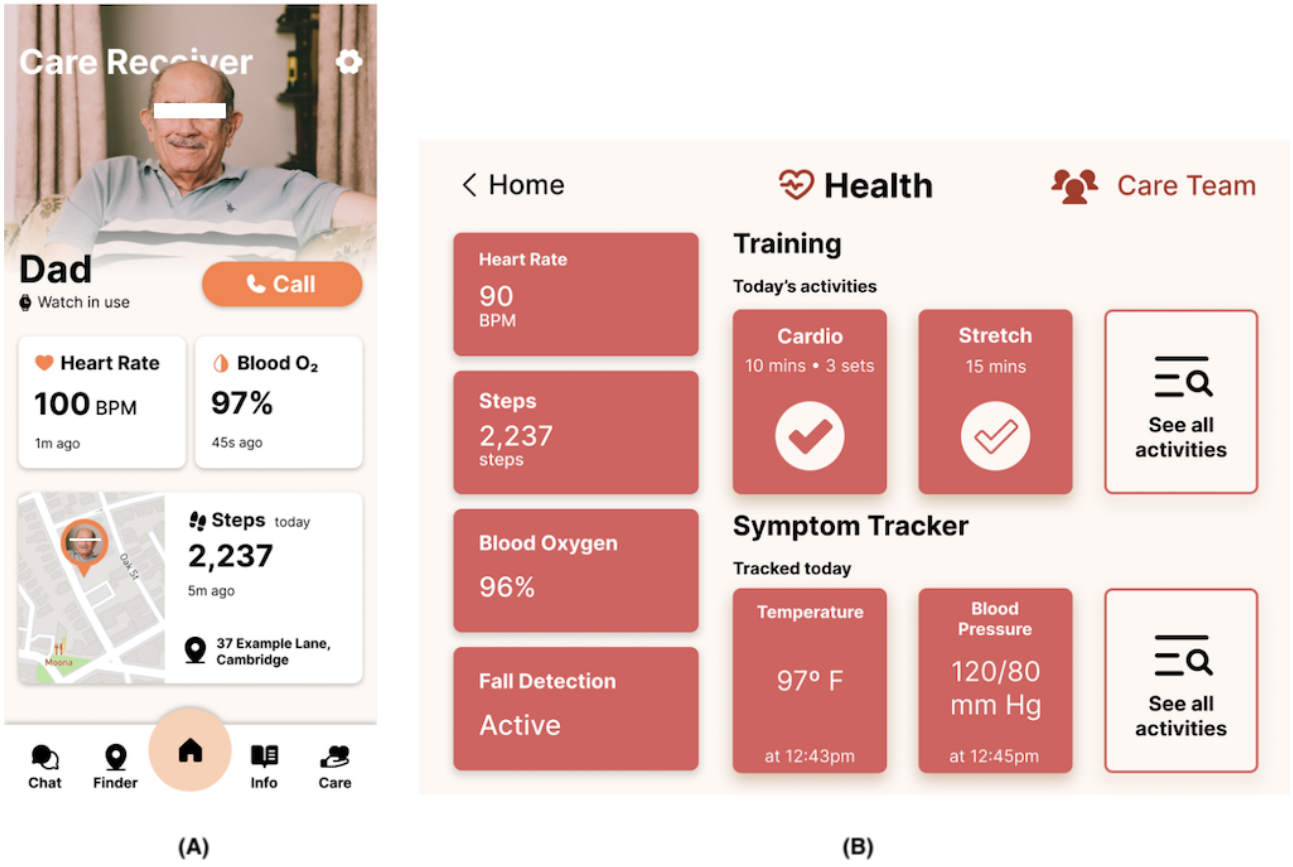


Figure 3. Redesigned system used for round 2: (A) Caregiver's LifeTomorrow app; (B) Older adult's LifeTomorrow app.



For informal caregivers, the app displays information shared by the care recipient. Informal caregivers can view vital data from their care recipients only if the care recipient chooses to share it (Figure 3). The app also includes features targeting

XSL•FO
RenderX

challenges informal caregivers face, such as isolation and information scarcity [77,84]. It offers a chat feature for private conversations or a forum for public conversations, local groups for finding informal caregivers nearby, and links to resources such as blog articles and videos.

Participants and Recruitment

Across both rounds of user studies, a total of 37 interviews were conducted with older adults, informal caregivers, and designers in the United States (see Multimedia Appendix 2). Recruitment used purposive sampling and targeted diverse roles, including informal caregivers, older adults, and designers. Older adults (aged ≥65 years [1,2]) were recruited from the MIT AgeLab Volunteer database. All older adults were in good health and were able to travel independently to the lab for study participation. Informal caregivers (nonprofessionals taking care of an older adult) came from the MIT AgeLab CareHive Panel, an open membership panel of informal caregivers who either currently provide care to an adult family member or had provided care to an adult family member in the past. Designers had at least 1 year of professional experience or relevant academic training and were recruited via snowball sampling. For the first round of user studies conducted online, participants were recruited from anywhere in the United States, whereas the second round of user studies focused on people in the Boston metropolitan area, as they were asked to come to MIT AgeLab to test the prototype in person.

In the first round of user studies, where visual mockups of the system were evaluated, 19 participants took part. There were 4 older adults (2 females and 2 males; aged 80 - 92 years), 8 informal caregivers (4 females and 4 males), and 7 designers (5 females and 2 males; aged 23 - 35 years).

Eighteen participants joined the second round of user studies, where an interactive prototype of the system was used. There were 9 older adults (8 females and 1 male; aged 68 - 92 years), 6 informal caregivers (5 females and 1 male; aged 24 - 60 years), and 3 designers (2 females and 1 male; aged 25 - 34 years). Five participants from round 1 also participated in round 2. All interviews were audio and visually recorded, and the audio was transcribed for coding. Transcripts were generated automatically via Zoom or Otter.ai and reviewed by a member of the research team.

Analysis

We conducted a qualitative thematic analysis [85] through the lens of self-determination theory [18]. Our analysis combined the interviews from both rounds of user studies to highlight themes that were dominant. The 6-phase approach by Braun and Clarke [86] guided the analysis. It involved the following steps:

1. Transcripts were reviewed, and preliminary notes were made (phase 1).
2. Inductive coding identified data relevant to the system's impact on users' well-being (phase 2).
3. Codes were clustered into themes, with a second researcher recoding 10% of the data from both rounds of user studies for reliability (phase 3).
4. Themes were deductively mapped to self-determination theory categories through multiple rounds of discussion, reflection, and interpretation of the content in the preliminary themes (phase 4).
5. Themes and labels were refined iteratively, resulting in 10 themes that mapped to autonomy, relatedness, or competence (phases 5 and 6), with refinement continuing throughout the writing process.

We aimed for richer interpretations of meaning rather than consensus, as guided by reflexive thematic analysis principles [87]. We contend that although many older individuals may use digital technology cautiously, they offer insightful perspectives on how technology affects society based on real-life experiences from all stages of their lives [88]. In keeping with the report by Durick et al [89], we considered older users as "specialist users," with their specialist status deriving from their experiences in the art of living [88].

Ethical Considerations

Participants completed a consent form prior to completing each interview and received US $75 as a gift card or via check for participation. Both rounds of user studies received exempt approval from the MIT Committee on the Use of Humans as Experimental Subjects (COUHES) with protocol numbers E-5008 and E-5079.

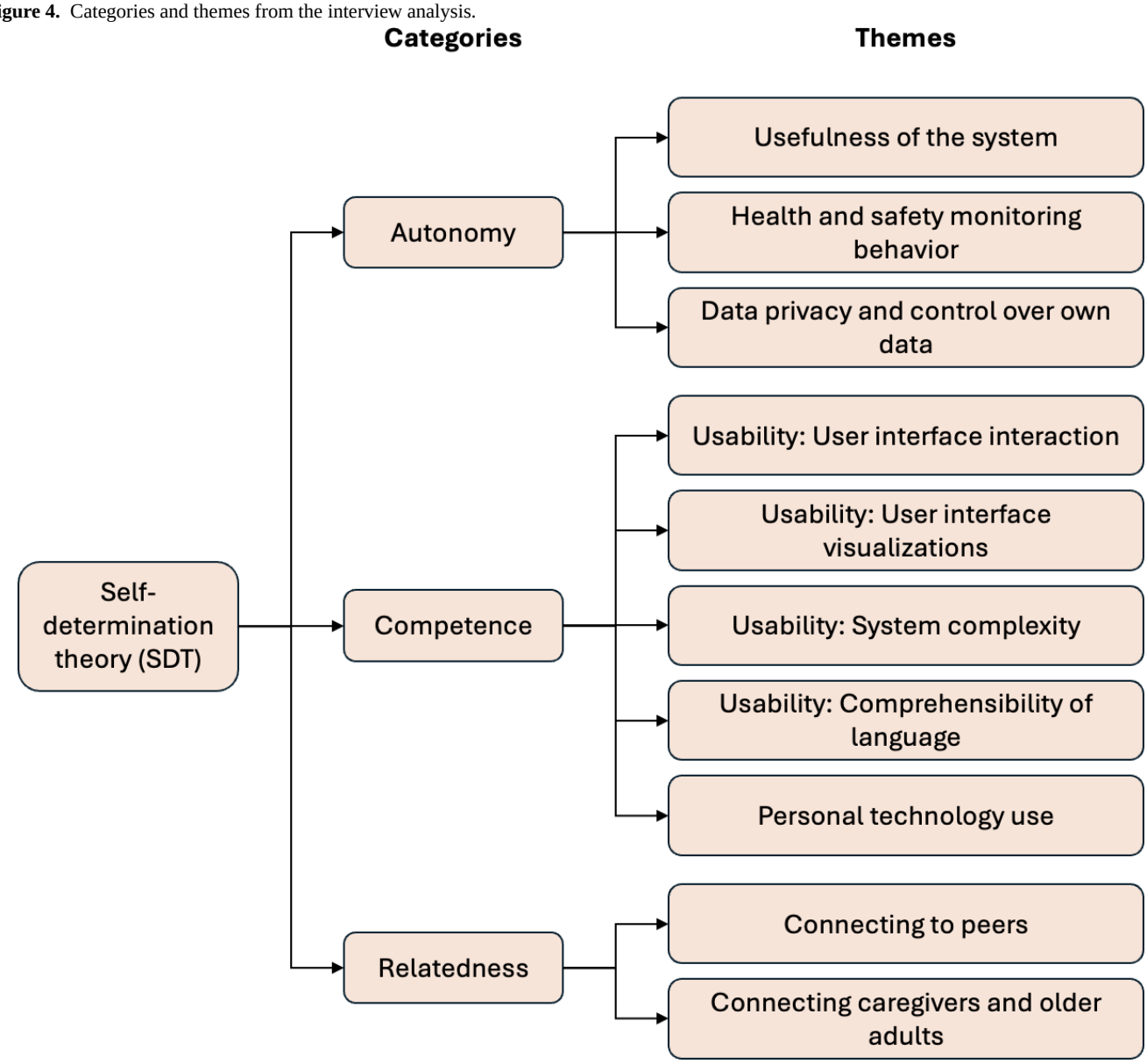
Results

Qualitative Analysis Results

Figure 4 displays the key themes that emerged from the qualitative analysis of the interviews and how they mapped to 3 categories related to self-determination.

XSL•FO
RenderX

Figure 4. Categories and themes from the interview analysis.



Autonomy

Autonomy refers to the ability to self-regulate and take ownership of one's needs and experiences [12]. Three themes from the interviews emerged around autonomy: the usefulness or benefits expected from using the system; the value of the health and safety monitoring features to support independent living; and, for older adults in particular, the ability to exert control over sharing their data and confidence in data privacy.

Usefulness of the System

In the first round of user studies, most older adults reported that they found the system helpful and indicated they would use it themselves. However, 1 female individual (older adult #4) chose not to use the LifeTomorrow system due to her existing use of a paper calendar and Apple Watch. She, however, appreciated the accessibility of the system, as it allowed her to track more data and create diagrams, which could not be done with her paper calendar system.

The system's functionalities, including health tracking, fall detection, safety sensors, and emergency alerts, were widely valued. People also appreciated having access to a range of content:

> *Well, I think it is wonderful to have all the resources presented in one place. It is like an encyclopedia for an older person.* [Older adult #1]

However, preferences for particular functions varied. Most older adults found either health tracking or safety features useful, while a minority expressed that they would derive use from both. For instance, older adult #2 desired expanded smart home sensors for added safety. Most older adults appreciated the system for fostering autonomy and managing their health data independently.

The integration of the Apple Watch into the system, rather than relying on specialized emergency devices, helped reduce stigma and increase the sense of autonomy for older adults [76,90]. Older adult #4 noted that she disliked wearing "an old lady

XSL•FO
RenderX

medallion," highlighting the watch's appeal as a device commonly used by multiple generations.

Revisions to the system for informal caregivers after the first round of user studies, including added local groups, resource summaries, and notifications for critical vitals, were positively received in the second round of user studies. However, the divide in what features participants found useful between health tracking and smart home features remained.

Among informal caregivers, 2 groups emerged. Socially well-connected informal caregivers valued features like remote health monitoring, but they were less interested in the potential for using community and resource-sharing functionalities. In contrast, informal caregivers who felt more isolated appreciated the app's ability to alleviate feelings of isolation through community features. One caregiver noted:

> *Part of caregiving is the isolation.* [Caregiver #5]

Regardless, informal caregivers reported that they found the app helpful overall, citing its comprehensive functionality as a significant advantage. Some concerns emerged about the validity of chat group information, but the mental health support provided by the community features was highly valued. From the feedback informal caregivers provided, tailoring resources to informal caregivers' specific experiences and care recipients' conditions would be essential for improving and further increasing the app's perceived value.

Health and Safety Monitoring Behavior

Questions about older adults' health and safety monitoring behaviors revealed a current reliance on simple tools like paper calendars or Apple Watches. One older adult highlighted the advantage of LifeTomorrow's centralized data storage, making the following statement:

> *I already do some of this, but it is in separate locations [paper calendar, smartwatch].* [Older adult #4]

Safety monitoring preferences among older adults varied. The older adults in our study frequently monitored steps, heart rate, and exercise, with only older adult #8 and older adult #11 having safety monitoring besides standard smoke detectors, through a security system installed in their homes. Some informal caregivers decided to use AirTags to increase the safety of their care recipients.

Older adults appreciated the system's affordability to provide health tracking, particularly its lack of a subscription model (older adult #4), which some participants contrasted to the service and business model of well-known Lifeline medical alert devices [91].

Data Privacy and Control Over One's Own Data

Control over personal data was a central concern, emphasizing the system's role in empowering users to decide how their data are shared. Older adults appreciated the app's transparency and ease of managing data access. One older adult commented:

> *It's surveilling me. That's what it's built for, but I'm choosing to use it. [...] I have complete control over who I share [my data] with.* [Older adult #7]

Informal caregivers valued that all shared data required care recipient consent, though some requested clearer explanations of data storage and sharing processes. The informal caregivers acknowledged the sensitivity of the data stored and shared via the app, but most older adults agreed with the informal caregivers that "there's more value than risk" (caregiver #6). Designers noted that providing older adults with insights into how their data appear to informal caregivers can enhance trust.

Competence

Competence is the sense that a user possesses the skills and knowledge to use the system independently and appropriately (or could acquire these relatively easily). Assessments of self-capability have been shown to be crucial for efficient work and completing moderately difficult tasks [17], and competence has been found to have a positive influence on older adults' intentions to use smart home technologies [24].

Themes around competence emerged primarily through comments around different dimensions of ease of usability of the system, notably the interactive features and behaviors of the UI, design of its visual components, system complexity, and comprehensibility of language, and by participants' reports of their existing experiences with technology.

For informal caregivers, accessible resources like the ones in the LifeTomorrow app addressed information gaps by providing features such as resource finders, blogs, and videos. These tools can empower informal caregivers, enhancing confidence in their abilities and reducing reliance on external support systems.

Usability: Interactions and UI Behaviors

Older adults were generally comfortable with touchscreens, finding the app intuitive and easy to navigate, particularly when clear descriptions (eg, "click to see history") were added to buttons. Their confidence and competence increased as whatever they clicked "you can always go back" (older adult #9). Interacting with the iPad touchscreen was perceived as accessible and easy to learn. For instance, older adult #9 commented:

> *Once I learned that you just could touch stuff and go where you want it to go. We're off to the races.*

Informal caregivers also found the interface intuitive and suggested features like customizable home screens. Designers proposed UI improvements, such as tab bars and clearer chat functionalities, which were implemented in the second prototype, with the goals of enhancing usability and confidence for all users.

Usability: Visual Design

Visual accessibility and well-structured visuals were positively received. Older adult #1 emphasized this by saying:

> *The presentations individually are lovely, perfect, have to be blind not to understand them.*

Moreover, older adult #5 commented:

> *I love the size because I can see. […] The dividers [to organize information] are wonderful, so it is not*

XSL•FO
RenderX

> *too much that older adults do not see anything anymore.*

Older adults appreciated the block organization, high contrast, and color-coded categories, saying that these made navigation easier. Large, clear health data graphs, adapted from iOS Health, ensured quick information retrieval. Feedback led to replacing unclear icons and refining the color scheme for improved accessibility.

Informal caregivers valued the visual consistency with other iOS apps and requested indicators for data updates, which were added.

> *I think that it was very clear. I feel like it feels clean. And there's not too many things going on.* [Caregiver #13]

Designers highlighted the importance of high-contrast colors and simple layouts. We received positive feedback on the accessibility of both apps, particularly after the second round of user studies, and it was summarized by designer #4 as "bright colors, high contrast, big text."

Usability: System Complexity

The perceived complexity of an app impacts users' confidence in using it, as apps seen as too complex can be overwhelming and hinder the psychological need for competence. LifeTomorrow's adaptability allowed users to focus only on desired functionalities, enabling lower complexity despite its multifunctionality.

> *I think I like the fact that it covers various areas. It's not one thing onlz, it's covering the health, your living quarters, emergencies, and so on. So, I like it. And it's user-friendly.* [Older adult #3]

In the second round of user studies, we found that older adults perceived the app to be intuitive and easy to navigate, as tasks, such as finding the highest heart rate of the week, were completed quickly. This low complexity fostered a high sense of competence among the older adults in our study. Older adult #4, for instance, stated how surprised she was that she could navigate around the app so intuitively. Older adult #8 described the app as "easy to look at," referring to the amount of information displayed on every page.

> *It's very smart. It's very clean, pleasing to the eye. It makes perfect sense. And it's, like, maneuverability to navigate so far has been critical.* [Older adult #9]

Informal caregivers appreciated the app's simplicity ("it feels clean. And there's not too many things going on" [caregiver #13]) and consistency with Apple interfaces. They emphasized the need for streamlined chat functionalities, which were redesigned for better navigation. Designers praised the system's organization and reduced complexity in the final version.

Usability: Comprehensibility of Language

Clear language and descriptions improved confidence among older adults, who previously struggled with terms like "fall detection" and "care team." Changes like replacing "hashtags" with "group chats" and "care team" with "community" reduced confusion and stigma. The second round of user studies showed significant improvements in understanding due to added explanations and user-friendly wording (eg, "training" instead of "exercise"). Overall, older adults were less likely to explore the functionality themselves if they did not know what different buttons or functions meant or led to, as older adult #6 described:

> *[My] Fitbit is very dense with information I don't understand, you know, blood glucose, […], I don't know what that means the don't offer explanations. So I sort of just ignore the functions. And in your system I understand what all these things are, […] zou can just glance at it and see what's useful.*

Relatedness

Relatedness concerns people's connections to one another, including their social ties [12]. From the interviews, data clustered around 2 themes under relatedness: connecting to peers (among informal caregivers) and connecting informal caregivers and older adults. Technologies that foster relatedness offer the potential for reducing isolation and loneliness among informal caregivers and older adults, and in turn supporting people's health and well-being [19,58,92].

Connecting to Peers

LifeTomorrow appeared to facilitate people's relatedness to peers via health dialog and community connections, offering mutual assistance. Older adult #9, for instance, highlighted that conversations about blood oxygen levels and exercise can improve older adults' sense of relatedness, increasing the value that LifeTomorrow provides. Informal caregivers, especially those who reported lacking a support network, valued the app's chat feature. Caregiver #5 highlighted its importance:

> *I like the idea of a community built in because part of caregiving is the isolation. I don't necessarily want to burden people in my day-to-day life with questions. But if it's a space that I'm opting into, the other folks are opting into. It would be really-I don't know-It would be pleasant.*

Feeling socially connected is a fundamental human need [92], yet informal caregivers often experience isolation [93]. By incorporating features, such as a chat function, the system aimed to provide informal caregivers with opportunities for peer support, sharing both emotional and informational resources. For instance, caregiver #3 noted:

> *It has just been invaluable, knowing other people who are taking care of their parents.*

Informal caregivers appreciated the ability to connect with peers virtually but also desired opportunities for in-person connections.

Older adults also expressed benefits from health-focused social interactions. Through shared knowledge, they could empower themselves and foster a sense of connection. Older adult #9 described:

> *I think seniors will connect about their health. So what's your heart rate? How many steps did you take yesterday? So the ways in which to foster positive kind of competition that I think is really a good thing*

XSL•FO
RenderX

that seniors get into, because they're going to call [a friend].

Connecting Informal Caregivers and Older Adults

Additionally, the system has the potential to facilitate informal caregiver–care recipient relationships by enabling older adults to control their health data and selectively share the data. Features, such as fall detection and emergency calls, may reduce informal caregivers' anxiety about their loved ones, creating opportunities for more meaningful, nonhealth-related interactions during phone calls or visits.

Older adults valued control over sharing health data, preferring LifeTomorrow's selective sharing approach over constant surveillance, which informal caregivers often favored. When in control of their own data, several older adults mentioned things like "I would love to data share with my primary care physicians" (older adult #8) or "I would certainly like it to tell somebody if I fallen" (older adult #2), indicating that they crave and would seek connection with their informal caregivers with regard to their health via LifeTomorrow.

While informal caregivers frequently visited or contacted their care recipients, they expressed constant worry about their safety. LifeTomorrow's features, such as heart rate and GPS tracking, fall detection, and emergency calls, addressed these concerns, aiming to give informal caregivers more peace of mind and strengthening remote connections. For instance, when talking about the LifeTomorrow function where the care recipient can choose to share some data, caregiver #12 commented:

What am I really going to do first thing every day, I'm really gonna go here just to check this, and I will hit this the most every single day.

This improved sense of relatedness was pivotal for both informal caregivers and older adults, fostering trust and support in caregiving relationships.

Discussion

Overview

The results of this research underscore the extent to which people's reactions to the system reflect the values they place on autonomy, competence, and relatedness, consistent with the core concepts in self-determination theory.

Core Concepts in Self-Determination Theory Are Implicated in How People Respond to New Technologies

Technologies to support care and independent living are becoming increasingly important for an aging population as the number of older adults grows and the relative availability of informal caregivers declines. The LifeTomorrow system was designed with older users at its center, with the goal of making it easy for older adults to use it and to be in control of their data. In this research, potential users (older adults and informal caregivers) interacted with a system prototype to evaluate its design and to describe their reactions (positive and negative) to it. The responses could be grouped into 10 themes that, in turn, mapped to 3 foundational constructs of self-determination theory (see Figure 4). The results of this research underscore the extent to which people's reactions to new technological systems reflect the values they place on autonomy, competence, and relatedness, consistent with the core concepts in self-determination theory.

From the interviews, it was apparent that older adults often strived to maintain their autonomy, avoiding safety monitoring systems that could stigmatize them as dependent or frail (eg, resistance to wearing an "old lady medallion"). The study also found that older adults are privacy-aware and prefer to have control of their own data. This aligns with prior research showing that older adults make rational, calculated decisions about information disclosures, challenging the perception that they are less privacy-conscious than younger individuals [94].

Competence, a key component of self-determination, was closely tied to the themes of usability and accessibility in the interview data. Many older adults, however, face barriers to technology proficiency, which in turn contribute to health inequities [11]. The intuitive interface design of the system prioritized simplicity, clear navigation, and accommodation for physical changes [95], and as such, the system reduced the likelihood of failed interactions, fostering user confidence. Past work has demonstrated that older adults can effectively use technology when it is designed to align with their needs and when the systems and devices are easier for people to use [13,24,88,96].

A technology-enabled solution can also foster self-determination by addressing the psychological need for relatedness in older adults and their informal caregivers. Informal caregivers described the value of knowing they were not alone, and older adults could envision how the system could serve as a point of conversation and connection with other older adults and their informal caregivers [97].

Self-Determination Theory and a Holistic Approach to Understanding Technology Adoption

To reap the benefits that technology potentially offers to older adults and informal caregivers, older adults need to actually adopt and use new technologies. Typically, when new technologies, products, or services are designed for older adults, there is a significant focus on the characteristics of direct, easily observable interactions between the individual and the product (eg, usability, ease of use, simplicity, effectiveness, and error prevention), which have been traditionally emphasized for products where the surrounding social context is less important [98]. While the 2 main factors of perceived usefulness and ease of use in the technology acceptance model [99] have been found repeatedly to be key to user adoption of new technologies and systems, the research also points to the need for a more comprehensive consideration of the whole user experience, including the broader environment and context of use, facilitating conditions and delivery channels, related social dynamics, emotional aspects of use, and behavioral characteristics (prior experiences, beliefs, and values) [98,100,101].

While Lee and Coughlin [98] and the UTAUT2 (Unified Theory of Acceptance and Use of Technology 2) [101] specifically developed frameworks that explain how technologies may be

XSL•FO
RenderX

integrated into and operated as part of a larger system, at the individual level, self-determination theory may offer a structure that can encompass many of these factors that are key for understanding technology adoption. Moreover, because of the implication of the self in self-determination theory, technology adoption and continued use may be more likely and "stickier" in cases where devices or systems arouse all 3 of the core values. For example, positive emotional responses to technologies and a perceived lack of stigma were consistent with the perspectives of Lee and Coughlin and aligned with people's sense of autonomy. Competence, which rests on the usability and ease of use of the device or system and the ability to access support, can similarly capture some of the key findings of both Lee and Coughlin [98] and Davis [99]. Finally, relatedness can capture what Lee and Coughlin [98] describe as social support for use. Assessing the extent to which autonomy, competence, and relatedness specifically are fostered or aroused by a device or system may be key to more consistent successful adoption and prolonged use of new technologies.

Thus, self-determination theory may offer a slightly different window into understanding technology adoption and use, particularly for older users who may be less confident or comfortable using new technologies, devices, or systems. Self-determination theory provides a means to evaluate new devices or systems not simply through a checklist of accessibility needs that focus solely on system function and feel. It points beyond these to capture how people's values, feelings, and experiences connect with the technology, which in turn may influence its use. This presents an opportunity to explore how specifically a self-determination metric might perform around gathering user evaluations or reactions to new technologies, and it opens up fruitful veins to examine whether and which different variables affect the weight or value users place on the 3 self-determination constructs (autonomy, competence, and relatedness) regarding adoption and how the weights on these values might also be affected by the purpose or function of the device.

Future Research Possibilities

Self-determination theory offers a perspective on how core values of autonomy, competence, and relatedness shape people's reactions to technologies. While this work demonstrates that these values are expressed through people's evaluations of a prototype system, future work could explore more systematically how different system features relate to different core values in self-determination theory. Further, future research could explore whether different users weigh core values differently (eg, "Do older adults place more weight on the value of autonomy that a technology or system provides compared to informal caregivers?" "Do informal caregivers value competence more highly?") and examine the extent to which these factors have an impact on user adoption of such technologies.

Limitations

There were limitations owing to the nature of the sample: a convenience, self-selected sample for older adults and informal caregivers, and a snowball sample for designers. Limited demographic and psychographic data were collected about the user participants in this study, who are likely not representative of the population as a whole. Given their residential location and willingness to participate in the study, they may have higher levels of education and income. This user sample is also likely to be more comfortable with technologies, as many consented and participated in interviews via Zoom, and might be more likely to be among those more interested and willing to adopt and use such a system. Regardless, participants could only test the system in the interviews. Use of the system in their daily lives (in a real-world setting) might lead to different opinions and results. As our sample was relatively small with regard to older adults and informal caregivers in the system, a larger trial is strongly encouraged to obtain further insights on this topic.

Conclusion

Technology offers the potential to support more older adults in terms of living at home independently for longer. However, to be successful, such technologies must involve systems that both older adults and informal caregivers feel confident and comfortable adopting and using. In a prototype system designed with an older adult agency at its heart, user evaluations of the system revealed that the 3 core values of self-determination theory (autonomy, competence, and relatedness) were useful for categorizing and understanding people's reactions. Further research should explore how different kinds of system features map specifically to different core values; whether users place different weights on different values; and how autonomy, competence, and relatedness affect adoption and continued use of technologies that provide these to users.

Acknowledgments

This work was supported in part by a fellowship of the German Academic Exchange Service (program number: 57515252) and by the Massachusetts Institute of Technology AgeLab.

Data Availability

Our responsibility is to protect the anonymity of the participants. We will discuss if and how the data could be shared to protect the participants' anonymity safely. In case safe access can be provided, the data can be shared upon reasonable request to the authors.

Authors' Contributions

Conceptualization: CMA (lead), CL (support), LD (support)

XSL•FO
RenderX

Data curation: CMA
Funding acquisition: CMA (equal), CL (equal), LD (equal)
Methodology: CMA (lead), CL (support), LD (support)
Resources: CMA
Software: CMA
Supervision: CMA (co-lead), CL (co-lead), EA (support)
Validation: SA
Visualization: MM (lead), CMA (support)
Writing: CMA (lead), CL (support), LD (support)
Writing – review & editing: CL (lead), LD (lead), MM (support), SA (support), EA (support)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Example study questions.
[DOCX File, 67 KB - aging_v8i1e72579_app1.docx]

Multimedia Appendix 2
Study participant demograpics.
[DOCX File, 70 KB - aging_v8i1e72579_app2.docx]

References

1. Despite pandemic, percentage of older adults who want to age in place stays steady. AARP: The American Association of Retired Persons. 2021. URL: https://www.aarp.org/home-family/your-home/info-2021/home-and-community-preferences-survey.html [accessed 2023-09-05]
2. How much care will you need? Administration for Community Living. 2020. URL: https://acl.gov/ltc/basic-needs/how-much-care-will-you-need [accessed 2023-09-05]
3. Celebrating national family caregivers month with BLS data. US Bureau of Labor Statistics. 2023. URL: https://www.bls.gov/blog/2023/celebrating-national-family-caregivers-month-with-bls-data.htm [accessed 2025-05-23]
4. Brady S, Ashebir S, D'Ambrosio L, Balmuth A, Felts A, Lee C. Mapping the caregiver experience: predicting dimensions of caregiver strain through task-based profiles. Res Aging 2025 Apr 14:1640275251334408. [doi: 10.1177/01640275251334408] [Medline: 40228530]
5. Keramat SA, Sarker PM, Comans T, Brooks D, Dissanayaka NN. Deterioration of health-related quality of life: the hidden health burden of informal caregiving. Eur J Health Econ 2025 Apr 17. [doi: 10.1007/s10198-025-01776-5] [Medline: 40240675]
6. Lambert SD, Bowe SJ, Livingston PM, et al. Impact of informal caregiving on older adults' physical and mental health in low-income and middle-income countries: a cross-sectional, secondary analysis based on the WHO's Study on global AGEing and adult health (SAGE). BMJ Open 2017 Nov 15;7(11):e017236. [doi: 10.1136/bmjopen-2017-017236] [Medline: 29146639]
7. Zueras P, Grundy E. Assessing the impact of caregiving for older parents on caregivers' health: initial health status and trajectories of physical and mental health among midlife caregivers for parents and parents-in-law in Britain. Soc Sci Med 2024 Feb;342:116537. [doi: 10.1016/j.socscimed.2023.116537] [Medline: 38181720]
8. Cattaneo A, Vitali A, Regazzoni D, Rizzi C. The burden of informal family caregiving in Europe, 2000-2050: a microsimulation modelling study. Lancet Reg Health Eur 2025 Jun;53:101295. [doi: 10.1016/j.lanepe.2025.101295] [Medline: 40255934]
9. The aging of the baby boom and the growing care gap: a look at future declines in the availability of family caregivers. AARP Public Policy Institute. 2023. URL: https://www.caregiver.org/uploads/2023/02/baby-boom-and-the-growing-care-gap-insight-AARP-ppi-ltc.pdf [accessed 2025-05-23]
10. Ribeiro O, Araújo L, Figueiredo D, Paúl C, Teixeira L. The caregiver support ratio in Europe: estimating the future of potentially (un)available caregivers. Healthcare (Basel) 2022;10(1):11. [doi: 10.3390/healthcare10010011]
11. Ageing. United Nations. URL: https://www.un.org/en/global-issues/ageing [accessed 2023-09-05]
12. García‐Rodríguez C, Martínez‐Tomás R, Cuadra‐Troncoso JM, Rincón M, Fernández‐Caballero A. A simulation tool for monitoring elderly who suffer from disorientation in a smart home. Expert Systems 2015 Dec;32(6):676-687 [FREE Full text] [doi: 10.1111/exsy.12107]
13. Kang KJ, Ka B, Kim SJ. A service scenario generation scheme based on association rule mining for elderly surveillance system in a smart home environment. Eng Appl Artif Intell 2012 Oct;25(7):1355-1364. [doi: 10.1016/j.engappai.2012.02.003]

XSL•FO
RenderX

14. Survadevara NK, Mukhopadhyay SC, Rayudu RK. Applying SARIMA time series to forecast sleeping activity for wellness model of elderly monitoring in smart home. Presented at: Sixth International Conference on Sensing Technology (ICST); Dec 18-21, 2012; Kolkata, India. [doi: 10.1109/ICSensT.2012.6461661]
15. Czech E, Soubutts E, Eardley R, O'Kane AA. Independence for whom? a critical discourse analysis of onboarding a home health monitoring system for older adult care. Presented at: 2023 CHI Conference on Human Factors in Computing Systems; Apr 23-28, 2023; Hamburg, Germany. [doi: 10.1145/3544548.3580733]
16. Wang J, Du Y, Coleman D, et al. Mobile and connected health technology needs for older adults aging in place: cross-sectional survey study. JMIR Aging 2019 May 15;2(1):e13864. [doi: 10.2196/13864] [Medline: 31518283]
17. Ryan RM, Deci EL. Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness: Guilford Press; 2017. [doi: 10.1521/978.14625/28806]
18. Zhao W, Kelly RM, Rogerson MJ, Waycott J. Older adults using technology for meaningful activities during COVID-19: an analysis through the lens of self-determination theory. Presented at: 2023 CHI Conference on Human Factors in Computing Systems; Apr 23-28, 2023; Hamburg, Germany. [doi: 10.1145/3544548.3580839]
19. Vines J, Lindsay S, Pritchard GW, et al. Making family care work: dependence, privacy and remote home monitoring telecare systems. Presented at: 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing; Sep 8-12, 2013; Zurich, Switzerland. [doi: 10.1145/2493432.2493469]
20. Harrington C, Martin-Hammond A, Bray KE. Examining identity as a variable of health technology research for older adults: a systematic review. Presented at: 2022 CHI Conference on Human Factors in Computing Systems; Apr 29 to May 5, 2022; New Orleans, LA, USA. [doi: 10.1145/3491102.3517621]
21. Nimmanterdwong Z, Boonviriya S, Tangkijvanich P. Human-centered design of mobile health apps for older adults: systematic review and narrative synthesis. JMIR Mhealth Uhealth 2022 Jan 14;10(1):e29512. [doi: 10.2196/29512] [Medline: 35029535]
22. Share of those 65 and older who are tech users has grown in the past decade. Pew Research Center. 2022. URL: https://www.pewresearch.org/short-reads/2022/01/13/share-of-those-65-and-older-who-are-tech-users-has-grown-in-the-past-decade [accessed 2023-09-05]
23. Zickuhr K, Madden M. Older adults and internet use. Pew Research Center. 2012. URL: https://www.pewresearch.org/internet/2012/06/06/older-adults-and-internet-use/ [accessed 2025-10-06]
24. Pal D, Funilkul S, Vanijja V, Papasratorn B. Analyzing the elderly users' adoption of smart-home services. IEEE Access 2018;6:51238-51252. [doi: 10.1109/ACCESS.2018.2869599]
25. Chernbumroong S, Atkins A, Yu H. Perception of smart home technologies to assist elderly people. Presented at: 4th International Conference on Software, Knowledge, Information Management and Applications; Aug 25-27, 2010; Paro, Bhutan p. 90-97 URL: https://eprints.staffs.ac.uk/1145/1/Proceeeding%20of%20SKIMA%202010.pdf [accessed 2025-10-06]
26. Demiris G, Rantz MJ, Skubic M, Aud MA, Tyrer HW Jr. Home-based assistive technologies for elderly: attitudes and perceptions. AMIA Annu Symp Proc 2005;2005:935. [Medline: 16779222]
27. Pal D, Triyason T, Funikul S. Smart homes and quality of life for the elderly: a systematic review. Presented at: 2017 IEEE International Symposium on Multimedia (ISM); Dec 11-13, 2017; Taichung, Taiwan. [doi: 10.1109/ISM.2017.83]
28. Zhu J, Shi K, Yang C, et al. Ethical issues of smart home‐based elderly care: a scoping review. J Nursing Management 2022 Nov;30(8):3686-3699. [doi: 10.1111/jonm.13521]
29. Schak M, Blum R, Bomsdorf B. Smart home for the elderly - a survey of desires, needs, and problems. Presented at: Human Aspects of IT for the Aged Population Technology in Everyday Living: 8th International Conference, ITAP 2022, Held as Part of the 24th HCI International Conference, HCII 2022; Jun 26 to Jul 1, 2022; Virtual. [doi: 10.1007/978-3-031-05654-3_7]
30. Kon B, Lam A, Chan J. Evolution of smart homes for the elderly. Presented at: 26th International Conference on World Wide Web Companion; Apr 3-7, 2017; Perth, Australia. [doi: 10.1145/3041021.3054928]
31. Portet F, Vacher M, Golanski C, Roux C, Meillon B. Design and evaluation of a smart home voice interface for the elderly: acceptability and objection aspects. Pers Ubiquit Comput 2013 Jan;17(1):127-144. [doi: 10.1007/s00779-011-0470-5]
32. Yu M, Rhuma A, Naqvi SM, Wang L, Chambers J. A posture recognition-based fall detection system for monitoring an elderly person in a smart home environment. IEEE Trans Inform Technol Biomed 2012;16(6):1274-1286. [doi: 10.1109/TITB.2012.2214786]
33. Hossain MS, Rahman MA, Muhammad G. Cyber–physical cloud-oriented multi-sensory smart home framework for elderly people: an energy efficiency perspective. J Parallel Distrib Comput 2017 May;103:11-21. [doi: 10.1016/j.jpdc.2016.10.005]
34. Jalal A, Kamal S, Kim D. A depth video sensor-based life-logging human activity recognition system for elderly care in smart indoor environments. Sensors (Basel) 2014 Jul 2;14(7):11735-11759. [doi: 10.3390/s140711735] [Medline: 24991942]
35. Kim K, Jalal A, Mahmood M. Vision-based human activity recognition system using depth silhouettes: a smart home system for monitoring the residents. J Electr Eng Technol 2019 Nov;14(6):2567-2573. [doi: 10.1007/s42835-019-00278-8]
36. McCloud R, Perez C, Bekalu MA, Viswanath K. Using smart speaker technology for health and well-being in an older adult population: pre-post feasibility study. JMIR Aging 2022 May 9;5(2):e33498. [doi: 10.2196/33498] [Medline: 35532979]
37. Trajkova M, Martin-Hammond A. "Alexa is a toy": exploring older adults' reasons for using, limiting, and abandoning echo. Presented at: 2020 CHI Conference on Human Factors in Computing Systems; Apr 25-30, 2020; Honolulu, HI, USA. [doi: 10.1145/3313831.3376760]

XSL•FO
RenderX

38. Pigini L, Bovi G, Panzarino C, et al. Pilot test of a new personal health system integrating environmental and wearable sensors for telemonitoring and care of elderly people at home (SMARTA Project). Gerontology 2017;63(3):281-286. [doi: 10.1159/000455168] [Medline: 28099965]
39. Raad MW, Yang LT. A ubiquitous smart home for elderly. Inf Syst Front 2009 Nov;11(5):529-536. [doi: 10.1007/s10796-008-9119-y]
40. Ransing RS, Rajput M. Smart home for elderly care, based on wireless sensor network. Presented at: 2015 International Conference on Nascent Technologies in the Engineering Field (ICNTE); Jan 9-10, 2015; Navi Mumbai, India. [doi: 10.1109/ICNTE.2015.7029932]
41. Suryadevara NK, Mukhopadhyay SC, Wang R, Rayudu RK, Huang YM. Reliable measurement of wireless sensor network data for forecasting wellness of elderly at smart home. Presented at: 2013 IEEE International Instrumentation and Measurement Technology Conference (I2MTC); May 6-9, 2013; Minneapolis, MN, USA. [doi: 10.1109/I2MTC.2013.6555372]
42. Demongeot J, Virone G, Duchêne F, et al. Multi-sensors acquisition, data fusion, knowledge mining and alarm triggering in health smart homes for elderly people. C R Biol 2002 Jun;325(6):673-682. [doi: 10.1016/s1631-0691(02)01480-4] [Medline: 12360854]
43. Lotfi A, Langensiepen C, Mahmoud SM, Akhlaghinia MJ. Smart homes for the elderly dementia sufferers: identification and prediction of abnormal behaviour. J Ambient Intell Human Comput 2012 Sep;3(3):205-218. [doi: 10.1007/s12652-010-0043-x]
44. Woznowski P, Burrows A, Diethe T. SPHERE: a sensor platform for healthcare in a residential environment. In: Angelakis V, Tragos E, Pöhls H, Kapovits A, Bassi A, editors. Designing, Developing, and Facilitating Smart Cities: Springer; 2017:315-333. [doi: 10.1007/978-3-319-44924-1_14]
45. Wang J, Fu Y, Lou V, Tan SY, Chui E. A systematic review of factors influencing attitudes towards and intention to use the long-distance caregiving technologies for older adults. Int J Med Inform 2021 Sep;153:104536. [doi: 10.1016/j.ijmedinf.2021.104536] [Medline: 34325206]
46. Guay C, Auger C, Demers L, et al. Components and outcomes of internet-based interventions for caregivers of older adults: systematic review. J Med Internet Res 2017 Sep 19;19(9):e313. [doi: 10.2196/jmir.7896] [Medline: 28928109]
47. Snyder M, Dringus L, Maitland Schladen M, Chenail R, Oviawe E. Remote monitoring technologies in dementia care: an interpretative phenomenological analysis of family caregivers' experiences. The Qualitative Report 2020;25. [doi: 10.46743/2160-3715/2020.4127]
48. Christie HL, Bartels SL, Boots LMM, Tange HJ, Verhey FJJ, de Vugt ME. A systematic review on the implementation of eHealth interventions for informal caregivers of people with dementia. Internet Interv 2018 Sep;13:51-59. [doi: 10.1016/j.invent.2018.07.002] [Medline: 30206519]
49. Orlando JF, Beard M, Kumar S. Systematic review of patient and caregivers' satisfaction with telehealth videoconferencing as a mode of service delivery in managing patients' health. PLoS ONE 2019;14(8):e0221848. [doi: 10.1371/journal.pone.0221848] [Medline: 31469865]
50. Dal Bello-Haas VPM, O'Connell ME, Morgan DG, Crossley M. Lessons learned: feasibility and acceptability of a telehealth-delivered exercise intervention for rural-dwelling individuals with dementia and their caregivers. Rural Remote Health 2014;14(3):2715. [doi: 10.22605/RRH2715] [Medline: 25081991]
51. Hautala M, Keränen NS, Leinonen E, Kangas M, Jämsä T. ICT use in family caregiving of elderly and disabled subjects. In: Giokas K, Bokor L, Hopfgartner F, editors. eHealth 360° Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering: Springer; 2017, Vol. 181:42-48. [doi: 10.1007/978-3-319-49655-9_7]
52. Sposaro F, Danielson J, Tyson G. IWander: an android application for dementia patients. Presented at: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology; Aug 31 to Sep 4, 2010; Buenos Aires, Argentina. [doi: 10.1109/IEMBS.2010.5627669]
53. Chi NC, Demiris G. A systematic review of telehealth tools and interventions to support family caregivers. J Telemed Telecare 2015 Jan;21(1):37-44. [doi: 10.1177/1357633X14562734] [Medline: 25475220]
54. Graffigna G, Gheduzzi E, Morelli N, et al. Place4Carers: a multi-method participatory study to co-design, piloting, and transferring a novel psycho-social service for engaging family caregivers in remote rural settings. BMC Health Serv Res 2021 Jun 21;21(1):591. [doi: 10.1186/s12913-021-06563-5] [Medline: 34154592]
55. Morelli N, Barello S, Mayan M, Graffigna G. Supporting family caregiver engagement in the care of old persons living in hard to reach communities: a scoping review. Health Soc Care Community 2019 Nov;27(6):1363-1374. [doi: 10.1111/hsc.12826] [Medline: 31441179]
56. Gaugler JE, Zmora R, Mitchell LL, et al. Six-month effectiveness of remote activity monitoring for persons living with dementia and their family caregivers: an experimental mixed methods study. Gerontologist 2019 Jan 9;59(1):78-89. [doi: 10.1093/geront/gny078] [Medline: 29982413]
57. Lee C. Adoption of smart technology among older adults: challenges and issues. Publ Pol Aging Rep 2014 Mar 6;24(1):14-17. [doi: 10.1093/ppar/prt005]

XSL•FO
RenderX

58. Walker RC, Tong A, Howard K, Darby N, Palmer SC. Patients' and caregivers' expectations and experiences of remote monitoring for peritoneal dialysis: a qualitative interview study. Perit Dial Int 2020 Nov;40(6):540-547. [doi: 10.1177/0896860820927528] [Medline: 33084514]
59. Sixsmith A, Johnson N. A smart sensor to detect the falls of the elderly. IEEE Pervasive Comput 2004;3(2):42-47. [doi: 10.1109/MPRV.2004.1316817] [Medline: 15483305]
60. Visutsak P, Daoudi M. The smart home for the elderly: perceptions, technologies and psychological accessibilities: the requirements analysis for the elderly in thailand. Presented at: 2017 XXVI International Conference on Information, Communication and Automation Technologies (ICAT); Oct 26-28, 2017; Sarajevo, Bosnia and Herzegovina. [doi: 10.1109/ICAT.2017.8171625]
61. Deci EL, Ryan RM. Intrinsic Motivation and Self-Determination in Human Behavior: Springer; 1985. [doi: 10.1007/978-1-4899-2271-7]
62. Tyack A, Mekler ED. Self-determination theory in HCI games research: current uses and open questions. Presented at: 2020 CHI Conference on Human Factors in Computing Systems; Apr 25-30, 2020; Honolulu, HI, USA. [doi: 10.1145/3313831.3376723]
63. Ford M, Wyeth P, Johnson D. Self-determination theory as applied to the design of a software learning system using whole-body controls. Presented at: 24th Australian Computer-Human Interaction Conference; Nov 26-30, 2012; Melbourne, Australia. [doi: 10.1145/2414536.2414562]
64. de Vreede T, Raghavan M, de Vreede GJ. Design foundations for AI assisted decision making: a self determination theory approach. Presented at: 54th Hawaii International Conference on System Sciences; Jan 5-8, 2021; Maui, HI, USA URL: http://hdl.handle.net/10125/70630 [accessed 2025-10-06]
65. Yang X, Aurisicchio M. Designing conversational agents: a self-determination theory approach. Presented at: 2021 CHI Conference on Human Factors in Computing Systems; May 8-13, 2021; Yokohama, Japan. [doi: 10.1145/3411764.3445445]
66. Henkemans OAB, Bierman BPB, Janssen J, et al. Design and evaluation of a personal robot playing a self-management education game with children with diabetes type 1. Int J Hum Comput Stud 2017 Oct;106:63-76. [doi: 10.1016/j.ijhcs.2017.06.001]
67. Huang L. Qualitative analysis of the application of self-determination theory in robotics tournaments. Presented at: Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction; Mar 6-9, 2017; Vienna, Austria. [doi: 10.1145/3029798.3038342]
68. Schwaninger I, Frauenberger C, Fitzpatrick G. Unpacking forms of relatedness around older people and telecare. Presented at: Companion Publication of the 2020 ACM Designing Interactive Systems Conference; Jul 6-10, 2020; Eindhoven, Netherlands. [doi: 10.1145/3393914.3395867]
69. Vansteenkiste M, Ryan RM. On psychological growth and vulnerability: basic psychological need satisfaction and need frustration as a unifying principle. J Psychother Integr 2013;23(3):263-280. [doi: 10.1037/a0032359]
70. Jansen A, Van Mechelen M, Slegers K. Personas and behavioral theories: a case study using self-determination theory to construct overweight personas. Presented at: 2017 CHI Conference on Human Factors in Computing Systems; May 6-11, 2017; Denver, CO, USA. [doi: 10.1145/3025453.3026003]
71. Peters D, Calvo RA, Ryan RM. Designing for motivation, engagement and wellbeing in digital experience. Front Psychol 2018;9:797. [doi: 10.3389/fpsyg.2018.00797] [Medline: 29892246]
72. Blandford A, Furniss D, Makri S. Qualitative HCI Research: Going Behind the Scenes: Springer; 2016. [doi: 10.1007/978-3-031-02217-3]
73. Fan M, Zhao Q, Tibdewal V. Older adults' think-aloud verbalizations and speech features for identifying user experience problems. Presented at: 2021 CHI Conference on Human Factors in Computing Systems; May 8-13, 2021; Yokohama, Japan. [doi: 10.1145/3411764.3445680]
74. Charles Milton M, Madeleine Aldenhoven C, Lee C, D'Ambrosio L, André E, Coughlin J. Mobile interfaces for caregivers and older adults: iterative design of the lifetomorrow ecosystem with aesthetic and functional considerations. Presented at: DRS2024; Jun 23-28, 2024; Boston, MA, USA. [doi: 10.21606/drs.2024.1156]
75. Older adult fire death risk. US Fire Administration. URL: https://www.usfa.fema.gov/statistics/deaths-injuries/older-adults.html [accessed 2025-05-23]
76. Older adult fall prevention. CDC: Centers for Disease Control and Prevention. URL: https://www.cdc.gov/falls/about/index.html [accessed 2023-09-05]
77. Matter: the foundation for connected things. Connectivity Standards Alliance. URL: https://csa-iot.org/all-solutions/matter/ [accessed 2023-09-05]
78. Matter support in iOS. Apple Developer Documentation. URL: https://developer.apple.com/apple-home/matter [accessed 2025-05-23]
79. Age and sex of all people, family members, and unrelated individuals. US Census Bureau. URL: https://www.census.gov/data/tables/time-series/demo/income-poverty/cps-pov/pov-01.html [accessed 2023-09-05]
80. HealthKit: access and share health and ftness data while maintaining the user's privacy and control. Apple Developer Documentation. URL: https://developer.apple.com/documentation/healthkit [accessed 2025-05-23]

XSL•FO
RenderX

81. Zhou C, Dai Y, Huang T, Zhao H, Kaner J. An empirical study on the influence of smart home interface design on the interaction performance of the elderly. Int J Environ Res Public Health 2022 Jul 26;19(15):9105. [doi: 10.3390/ijerph19159105] [Medline: 35897468]
82. Sharma R, Nah FFH, Sharma K, Katta T, Pang N, Yong A. Smart living for elderly: design and human-computer interaction considerations. In: Zhou J, Salvendy G, editors. Human Aspects of IT for the Aged Population Healthy and Active Aging ITAP 2016 Lecture Notes in Computer Science: Springer; 2016:112-122. [doi: 10.1007/978-3-319-39949-2_11]
83. Tsuchiya LD, Braga LF, de Faria Oliveira O, de Bettio RW, Greghi JG, Freire AP. Design and evaluation of a mobile smart home interactive system with elderly users in Brazil. Pers Ubiquit Comput 2021 Apr;25(2):281-295. [doi: 10.1007/s00779-020-01408-0]
84. Greenwood N, Pound C, Brearley S, Smith R. A qualitative study of older informal carers' experiences and perceptions of their caring role. Maturitas 2019 Jun;124:1-7. [doi: 10.1016/j.maturitas.2019.03.006] [Medline: 31097172]
85. Kuckartz U. Qualitative Text Analysis: A Guide to Methods, Practice &Amp; Using Software: SAGE Publications Ltd; 2014. [doi: 10.4135/9781446288719]
86. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
87. Clarke V, Braun V. Successful Qualitative Research: A Practical Guide for Beginners: Sage Publications Ltd; 2013.
88. Barros Pena B, Clarke RE, Holmquist LE, Vines J. Circumspect users: older adults as critical adopters and resistors of technology. Presented at: 2021 CHI Conference on Human Factors in Computing Systems; May 8-13, 2021; Yokohama, Japan. [doi: 10.1145/3411764.3445128]
89. Durick J, Robertson T, Brereton M, Vetere F, Nansen B. Dispelling ageing myths in technology design. Presented at: 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration; Nov 25-29, 2013; Adelaide, Australia. [doi: 10.1145/2541016.2541040]
90. Caldeira C, Nurain N, Connelly K. "I hope i never need one": unpacking stigma in aging in place technology. Presented at: 2022 CHI Conference on Human Factors in Computing Systems; Apr 29 to May 5, 2022; New Orleans, LA, USA. [doi: 10.1145/3491102.3517586]
91. Lifeline medical alert systems. Lifeline. URL: https://www.lifeline.com/medical-alert-systems/ [accessed 2023-09-13]
92. Deci EL, Vansteenkiste M. Self-determination theory and basic need satisfaction: understanding human development in positive psychology. Ricerche di Psicologia 2004;27(1):23-40.
93. Schurgin M, Schlager M, Vardoulakis L, Pina LR, Wilcox L. Isolation in coordination: challenges of caregivers in the USA. Presented at: 2021 CHI Conference on Human Factors in Computing Systems; May 8-13, 2021; Yokohama, Japan. [doi: 10.1145/3411764.3445413]
94. Ghaiumy Anaraky R, Byrne KA, Wisniewski PJ, Page X, Knijnenburg B. To disclose or not to disclose: examining the privacy decision-making processes of older vs. younger adults. Presented at: 2021 CHI Conference on Human Factors in Computing Systems; May 8-13, 2021; Yokohama, Japan. [doi: 10.1145/3411764.3445204]
95. Haugen IK, Englund M, Aliabadi P, et al. Prevalence, incidence and progression of hand osteoarthritis in the general population: the Framingham Osteoarthritis Study. Ann Rheum Dis 2011 Sep;70(9):1581-1586. [doi: 10.1136/ard.2011.150078] [Medline: 21622766]
96. Haase KR, Cosco T, Kervin L, Riadi I, O'Connell ME. Older adults' experiences with using technology for socialization during the COVID-19 pandemic: cross-sectional survey study. JMIR Aging 2021 Apr 23;4(2):e28010. [doi: 10.2196/28010] [Medline: 33739929]
97. Kelley MM, Powell T, Camara D, et al. Mobile health apps, family caregivers, and care planning: scoping review. J Med Internet Res 2024 May 23;26:e46108. [doi: 10.2196/46108] [Medline: 38781588]
98. Lee C, Coughlin JF. PERSPECTIVE: Older adults' adoption of technology: an integrated approach to identifying determinants and barriers. J of Product Innov Manag 2015 Sep;32(5):747-759. [doi: 10.1111/jpim.12176]
99. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
100. Premanandan S, Ahmad A, Cajander Å, Ågerfalk P, van Gemert-Pijnen L. Designing and evaluating IT applications for informal caregivers: scoping review. J Med Internet Res 2024 Oct 23;26:e57393. [doi: 10.2196/57393] [Medline: 39442167]
101. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 2012;36(1):157-178. [doi: 10.2307/41410412]

Abbreviations

HCI: human-computer interaction
MIT: Massachusetts Institute of Technology
UI: user interface

XSL•FO
RenderX

Edited by H Köttl; submitted 20.02.25; peer-reviewed by A Garnett, B Hiebert; revised version received 12.09.25; accepted 22.09.25; published 16.10.25.

Please cite as:
Aldenhoven CM, Milton M, D'Ambrosio L, Lee C, Ashebir S, André E
An Accessible Multifunctional System to Support Safe and Independent Aging in Place: Iterative Development and Qualitative Analysis
JMIR Aging 2025;8:e72579
URL: https://aging.jmir.org/2025/1/e72579
doi:10.2196/72579

© Céline Madeleine Aldenhoven, Matthew Milton, Lisa D'Ambrosio, Chaiwoo Lee, Sophia Ashebir, Elisabeth André. Originally published in JMIR Aging (https://aging.jmir.org), 16.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Technology Supporting Aging-in-Place Using an Equity Lens Through Focus Groups and World Café–Informed Research Agenda: Qualitative Study

Marianne Saragosa[1,2*], PhD; Ibukun-Oluwa Omolade Abejirinde[3,4,5*], MSc, MD, PhD; Evan MacEachern[6], BKin, MScRR; Michelle LA Nelson[2,6*], PhD; Kristina M Kokorelias[7,8,9], PhD; Sidra Bharmal[1*], BSc; Brina Ludwig Prout[10*], BA; Marian Mohamed[1*], BSCN

[1]Science of Care Institute, Sinai Health, 1 Bridgepoint Drive, Toronto, ON, Canada

[10]Hennick Bridgepoint Hospital, Toronto, ON, Canada

[2]Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[3]Institute for Better Health, Trillium Health Centre, Mississauga, ON, Canada

[4]Research and Innovation Institute, Women's College Hospital, Toronto, ON, Canada

[5]Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[6]Lunenfeld-Tanenbaum Research Institute, Sinai Health, Toronto, ON, Canada

[7]Department of Geriatric Medicine, Sinai Health System and University Health Network, Toronto, ON, Canada

[8]Department of Occupational Science and Occupational Therapy, Temerty Faculty of Medicine, Rehabilitation Sciences Institute, University of Toronto, Toronto, ON, Canada

[9]Rehabilitation Sciences Institute, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[*]these authors contributed equally

Corresponding Author:
Marianne Saragosa, PhD
Science of Care Institute, Sinai Health, 1 Bridgepoint Drive, Toronto, ON, Canada

Abstract

Background: Older adults prefer to age in their home or community of choice, which could include naturally occurring retirement communities (NORCs). As a place with a high density of older adults, NORCs could be sites where technology is leveraged to support independence and aging in the right place. However, there is limited research on how technology adoption and use occur in NORCs in ways that support older adults.

Objective: This study aims to cocreate a research agenda on equity-informed technology considerations that help older adults live independently in NORCs.

Methods: This is a 2-phase sequential qualitative descriptive study of 5 community-based focus groups and an in-person World Café event. We use the focus group method to acquire data about older adults' experiences with and perceptions of using technology to support aging-in-place in NORC settings. This data informs the design and facilitation of deliberate dialogues at the World Café event. Three questions helped to guide the small group discussions. The World Café is a creative, collaborative, and conversation-generating method that aims to generate exchanges between people with different views on a particular topic.

Results: In total, 45 NORC residents participated in a focus group about their experience and use of technology. The data revealed 3 central categories that highlight the perception of the use of technology to support the independence of participants in their homes and communities, its challenges, and areas to consider when deploying technology for helping older adults age in place. The subsequent World Café event included 40 participants and a combination of NORC residents, service providers, researchers, technology innovators, and policy makers. Insights drawn from the focus groups and World Café informed a 10-question research agenda about equity-informed technology principles that span accessible support, accessible interfaces, affordable and equitable access, available digital literacy training, accessible data, and accessible partnerships.

Conclusions: Our study explores NORCs as potential environments for offering a transformative opportunity to address equity considerations for technology supporting aging in place. Our findings and research agenda highlight critical areas for consideration, including leveraging partnerships, integrating public and private technology ecosystems, and designing technology with older users that evolves with the population's needs. Notably, by embedding principles of equity, inclusivity, and user-centered design, the collective of developers, researchers, and service providers can ensure that emerging technology serves diverse aging populations equitably and effectively.

XSL•FO
RenderX

(JMIR Aging 2025;8:e71093) doi:10.2196/71093

KEYWORDS

aging in place; community; research agenda; older adult; world cafe; technology; equity

Introduction

Background

As the fastest-growing age demographic in Canada, older adults are expressing a strong desire to age in place [1]. This research defines an "older adult" as a person aged 65 years or older [2]. Aging in place refers to living safely, independently, and comfortably in one's home and community regardless of age, income, or ability level [3]. However, "aging in the right place" has been considered a more appropriate goal, given that place often depends on an older person's personal preference, circumstances, and care needs [4]. Advocates suggest expanding housing options for those unsuitable for long-term care by developing alternatives, such as senior apartments, assisted living residences, and so forth to create safer and more accessible living environments and support the desires of older adults to age in place [4,5]. Such settings could include adaptations made to the residential environments of older people for supporting their independence and improving their quality of life [6]. For example, incorporating digital technologies, such as movement sensors and medication dispensers, is increasingly used to mitigate declining cognitive and physical abilities [7].

One example of an alternative housing model is naturally occurring retirement communities (NORCs). Originally coined in 1986, NORCs are defined as a "housing development that is not planned or designed for older people, but which over time comes to house largely older people" [8]. The NORC Innovation Centre at University Health Network has expanded the conceptualization of NORCs to include that they may consist of communities designed to house many older adults but not purpose-built for this population in the way long-term care homes or retirement homes are [9]. NORCs can provide a model for healthy aging-in-place since many offer supportive service programs based on the needs of residents [8]. While the availability and capacity of government, social, and health service providers and policy makers can create opportunities for supportive programming [10], less is known about how technology is currently being leveraged within NORC environments to support the needs of residents. Technological advancements (eg, remote monitoring systems, wearable health devices, and mobile apps) have shown promise in supporting efforts to age in place [11,12]. For example, countries like the Netherlands and Spain have focused on technology to mitigate the need for additional care and promote independence and autonomy [13]. In particular, low-tech technologies are designed to be as simple as possible and can offer practical, low-cost, and accessible features [14]. Despite considerable public and private interest and investment in aging-in-place technology development and research [15,16], there has been limited adoption of advanced technologies and their impact on people and health care systems [17], including diverse older adults [18,19]. Gaps persist in accessing technology and enabling digital literacy to support aging in the place where one prefers [20].

Later in life, technology acceptance and adoption are often influenced by the digital divide, which is known as the gap between those who do and those who do not have access to or knowledge of how to use advanced forms of technology [21]. While many older adults actively use current technologies (eg, internet, tablets, and smartphones), the digital divide continues to be a barrier for older people with lower access (ie, devices, internet, financial, and digital literacy) [22]. The digital divide is even more pronounced in those who experience marginalization at the intersections of racial and ethnic identity and socioeconomic status [23-25]. Indeed, significant barriers to accessing or engaging with technologies for health-related needs could worsen racial and ethnic health disparities experienced by the aging population [25]. While technology can offer promising solutions to support older adults in Canada and other countries in aging in the right place [19], the digital divide, or the gap between those who do and those who do not have access to new digital and information technology, must be addressed [21].

Objective

With limited research on how technology adoption and use occur in NORCs in ways that support aging in place, this study aimed to cocreate a research agenda on equity-informed technology considerations that help older adults live independently in NORCs.

Methods

Study Design

This 2-phase sequential qualitative descriptive study [26] consisted of community-based in-person focus groups and an intersectoral World Café event. We used the focus group method to gather data about older adults' experiences with and perceptions of using technology to support aging-in-place in NORC settings. The focus group data, in combination with project team discussions, helped develop the overarching questions for the deliberate dialogues at the World Café event [27]. The World Café is a creative, collaborative, and conversation-generating method that aims to generate exchanges between people with different views on a particular topic [28].

Developed in 1995, the World Café is considered a participatory method [29] by shifting to a bottom-up approach by engaging as many different actors as possible in the data collection process [30]. The World Café was the preferred approach in this case, as it has been used extensively to convene community-dwelling older adults and intersectoral participants [31,32], including research prioritization activities [28]. The evidence also supports the World Café as a valuable tool for fostering productive conversations among individuals with lived experience. Finally, with World Cafés, a large and heterogeneous group of people

XSL•FO
RenderX

can be brought together in a systematic and organized manner [30].

We used the following 5 steps of the World Café method to guide our event planning: creating an informal environment, offering a warm welcome and overview from the host, engaging in 3 rounds of conversations among smaller groups of participants on 3 key questions (see), and a final round of sharing the conversation results with the broader group [29]. Typically, participants are encouraged to rotate tables intermittently and meet new people; however, since some participants had mobility differences, we adapted the format such that the "tabletop host" and a notetaker (ie, scribe) rotated tables instead. The following were the World Café questions:

1. What challenges do older adults living in naturally occurring retirement community (NORC) neighborhoods have that might be supported with technology to help them stay in their homes?
2. How might challenges with technology be the same or different for people who are newcomers, are from different cultural backgrounds, education, and income levels?
3. What other things should we think about when developing and using technological (or digital) tools that could help older people stay in their homes?

Sample and Recruitment

Focus Groups

We used purposeful sampling to identify and invite older adults living in urban NORCs in Toronto, Canada, to participate in a focus group. The inclusion criteria were community-dwelling older adults living in a NORC building. We identified NORCs from a publicly available list, including identified apartments, condos, co-operative housing, and social housing buildings within a bounded geographical area [33]. We used personal and professional networks to approach NORC buildings about the study. With agreement and support from the condo board or service providers at each location, we invited residents through word-of-mouth, circulating a recruitment flyer and using gatekeepers (ie, NORC residents or managers who brokered relationships between the researchers and the NORC building).

World Café

We also used purposeful sampling to invite interested focus group participants to the World Café. To ensure adequate representation of NORC residents in numbers, we used snowball sampling, encouraging focus group participants to share information about the World Café event with other residents. Because of the World Café's intersectoral nature, we invited representatives from housing, health and social care, academics and research, and government senior services to attend.

Data Collection

Focus Groups

Focus groups were held on-site in communal activity rooms within 5 different NORC buildings between April 2024 and June 2024. Each focus group involved a facilitator and a cofacilitator who supported the consent process, collected sociodemographic information, took field notes, and provided follow-up questions. The facilitator started each session by welcoming the participants, explaining the discussion's purpose, and inviting them to introduce themselves. With an interview guide, the facilitator guided the discussion to address the following three central questions: (1) tell me about your current care needs; (2) tell me how technology addresses or could address these needs; and (3) tell me about current challenges and benefits with technology. For the focus group participants, we broadly defined "technology" as the various tools, devices, and systems people use daily to facilitate tasks, improve efficiency, and enhance convenience, spanning assistive devices, telemedicine, health monitoring, information and communication technologies, and so forth [34]. The focus groups lasted approximately 45-70 minutes; the discussions were digitally audio-recorded and transcribed verbatim by one of the authors (MM).

World Café

The in-person, full-day World Café event occurred in June 2024 in a large workspace at a tertiary hospital in Toronto, Canada. Upon arrival, participants were asked to complete paperwork (eg, consent forms and demographic survey) and were seated at one of six tables. Each table had a tabletop host and a notetaker. The host and notetaker were assigned one of three questions and rotated at 3 intervals to allow all participants to build on the prior responses to the question and provide additional feedback. Upon completion of the third round, all participants were invited to participate in a group discussion, answering the final question: "What groups need to work together to ensure that new technologies that help older people stay in their homes treat everyone fairly?"

Data Analysis

Focus Groups

We used conventional content analysis to identify cross-cutting categories across the focus groups [35]. Conventional content analysis is most beneficial in the absence of existing research on the phenomenon. Therefore, it uses inductive logic, avoiding applying any predetermined categorical interpretation system [35]. Data analysis started with reading all transcripts, and 2 coders identified salient codes by highlighting exact words from the text to capture thoughts or concepts that appeared to describe an experience or perception of technology within the context of aging in place. Next, the first author created labels that reflected more than one concept, thereby sorting and merging codes into categories. Finally, the emergent categories were used to organize and group all the codes into meaningful clusters. For example, quotes like "[using technology to] connect with people near and far" and "interacted with all the kids and grandkids" were coded as "communication" and categorized under "Uses of technology."

World Café

For the World Café event, we used directed content analysis. This analysis seeks to validate or extend a theory or framework [35]. In this case, we applied the Social Ecological Model (SEM) [36] as an organizing framework to map the factors of technology use among NORC residents, considering the challenges and enablers. The SEM was selected because it

XSL•FO
RenderX

suggests that behavior is integrated into a dynamic network of intrapersonal characteristics, interpersonal processes, and institutional, community, and public policy factors [37]. Using the SEM also allows for identifying specific drivers of technology use or disuse among community members. Handwritten notes were transcribed and consolidated into a Microsoft Word document. In the coding process, we included examples for each of the factors. For instance, older adults' knowledge, attitudes, and behaviors are at the "individual" level of the SEM model. Similar to the focus group analysis, we reviewed the transcribed notes and coded the data using predetermined categories wherever possible. For example, better integrating "sensors in NORC buildings" and coordinating specialized technology training, as a NORC was coded according to the SEM's community level, acknowledges these factors as potential facilitators of aging-in-place. We also coded data that could not be categorized into one of the SEM factors, such as data related to technology devices and their features, design, and interface.

Ethical Considerations

This study was approved by the Mount Sinai Hospital Research Ethics Board (REB reference number 23‐0123-E). Study participants provided written informed consent before engaging in study activities. Three participants in each focus group received a $25 CAD (US $18) gift card as a draw, whereas World Café attendees were offered a $25 CAD (US $18) gift card with the option to decline. Participation was not anonymous due to the focus group and in-person event; however, participation was voluntary. The study findings will be disseminated through presentations at conferences and peer-reviewed publications using deidentified data.

Results

Characteristics of Participants

Focus Groups

In total, 45 people participated in a focus group, with each focus group having between 6 and 25 NORC residents. Of the demographic data we collected, most of the participants identified as female (n=30, 67%), between 70 and 89 years of age (n=33, 74%). The most identified ethnicities were White North American or European (n=29, 66%), Southeast Asian (n=3, 7%), and East Asian (n=3, 7%). Most participants indicated living alone (n=27, 60%), followed by living with a partner (n=14, 31%). The remaining participants lived with others (including family) or preferred to keep their information private. See Table 1 for Focus group participant characteristics.

XSL•FO
RenderX

Table . Focus group participant characteristics.

Characteristic	NORC[a] resident (n=45)
Gender, n (%)	
Women	30 (67)
Men	15 (33)
Age (y), n (%)	
60 - 69	10 (22)
70 - 79	21 (47)
80 - 89	12 (27)
90 - 99	1 (2)
Prefer not to answer	1 (2)
Ethnicity, n (%)	
White Northern American	19 (42)
White European	10 (22)
East Asian	3 (7)
Southeast Asian	3 (7)
Latino or Hispanic	2 (4)
Caribbean	1 (2)
Black African	1 (2)
Middle Eastern	1 (2)
Mixed heritage	1 (2)
Indigenous	1 (2)
Prefer not to answer	3 (7)
Living situation, n (%)	
Live alone	27 (60)
Live with partner	14 (31)
Live with family or friends	2 (4)
Prefer not to answer	2 (4)
Relationship status, n (%)	
Single	22 (49)
In a relationship	13 (29)
Widow	3 (7)
Separated	2 (4)
Prefer not to answer	5 (11)
Employment status, n (%)	
Retired	37 (82)
Unemployed, looking for work	3 (7)
Employed for wages	1 (2)
Prefer not to answer	4 (9)

[a]NORC: naturally occurring retirement community.

World Café

A total of 40 people participated in the World Cafe. Of these, 18 identified themselves as NORC residents; the remaining individuals were system leaders, including service providers, decision-makers, researchers, technology entrepreneurs, innovators, and administrators. Of the NORC residents, most identified as female (n=14, 78%), were between the ages of 70

XSL•FO
RenderX

and 79 (n=9, 50%) years, lived alone (n=9, 50%) or with a partner or family (n=8, 44%), and most identified as White European or North American (n=12, 66%). For the system leaders (N=22), the average years of experience was 11 (SD 4.2) years, ranging from 2 to 17 years. Most participants held a graduate degree (n=15, 68%) and identified as female (n=18, 82%). See Table 2 for World Café participant characteristics.

XSL•FO
RenderX

Table . World Café participant characteristics.

Characteristics	Role	
	NORC[a] resident (n=18)	System leader (n=22)
Gender, n (%)		
Female	14 (78)	18 (82)
Male	4 (22)	4 (18)
Age (y), n (%)		
18 - 29	0 (0)	1 (4)
30 - 39	0 (0)	7 (32)
40 - 49	0 (0)	11 (50)
50 - 59	0 (0)	1 (4)
60 - 69	4 (22)	0 (0)
70 - 79	9 (50)	1 (4)
80 - 89	4 (22)	0 (0)
90 - 99	1 (6)	0 (0)
Prefer not to answer	0 (0)	1 (4)
Ethnicity, n (%)		
White European	8	5 (23)
White North American	4 (22)	5 (23)
East Asian	1 (6)	3 (14)
Southeast Asian	1 (6)	1 (4)
Caribbean	2 (11)	1 (4)
Black African	0 (0)	1 (4)
Middle Eastern	0 (0)	1 (4)
Mixed heritage	0 (0)	2 (9)
Prefer not to answer	2 (11)	3 (14)
Living situation, n (%)		
Live alone	9 (50)	—[b]
Live with partner	8	—
Live with family	1 (6)	—
Relationship status, n (%)		
In a relationship	8	—
Single	6 (33)	—
Widow	1 (6)	—
Prefer not to answer	3 (17)	—
Employment status, n (%)		
Retired	15 (83)	—
Unable to work	1 (6)	—
Self-employed	1 (6)	—
Prefer not to answer	1 (6)	—
Years of experience		
Mean (SD)	—	10.8 (4.2)
Range	—	2 - 17

XSL•FO
RenderX

Characteristics	Role	
	NORC[a] resident (n=18)	System leader (n=22)
Profession, n (%)		
Researcher		11(50)
Service provider	—	4 (18)
Designer/innovator	—	4 (18)
Administrator	—	2 (9)
Policy maker	—	1 (5)
Education level completed, n (%)		
Postsecondary	—	6 (27)
Diploma	—	1 (5)
Graduate degree	—	15 (68)

[a]NORC: naturally occurring retirement community.

[b]Not applicable.

Focus Group Findings

The focus group data revealed 3 central categories that highlight the use of technology to support participants' independence in their homes and communities, its challenges, and areas to consider when deploying technology to help older adults age in place (see Table 3).

Table . Focus group findings.

Category	Open codes	Direct quotes
Using technology	• Life is online • Connecting everyday • Finding information	• "When I leave my appointment everything comes online…medical appointments, email, the whole world." [FG 3] • "If there's something I don't understand how to do, I Google it and it sends me to YouTube where it explains how to do it. But it's so fast-paced that I have to keep rewinding it." [FG 2]
Expressing challenges using technology	• Goofing up the machine (cell phone) • Remembering passwords	• "I do not have a cell phone. I got rid of it a year and a half ago… I can goof the machine up very easy. I always seem to do that with cell phones." [FG 2] • "You have to remember to change your password every six months. We locked out and out because we forgot it." [FG 4]
Identifying technology considerations	• Needing to be taught • Thinking about income	• "My emails right now are coming into trash. And I'm at the point where I won't touch anything." [FG 2] • "We have been waiting for the whole building to be so-called fiber. Well, unfortunately, they have been trying to entice me to join by asking me to spend $90.00 a month. For a retiree, it's not really that simple" [FG 4]

Using Technology

Focus group participants described how they currently use technology to support their independence in their homes and communities. The types of use fell into categories of communication (eg, keeping in contact with family and friends), wellness and activity (eg, tracking upcoming appointments and online banking), information seeking (eg, accessing news, information, and current events), and for personal safety (eg, checking in on others). Some participants recognized the role of existing (eg, printer and video calling) and emerging technology (eg, ChatGPT and wearables) in adding value to their daily activities through access to information and connection to others.

> *I have a new iPhone. And I wanted to use it to make films. So I figured to make a film, you have to have a*

XSL•FO
RenderX

story to start. So, I asked ChatGPT what story I could use. And it gave me a story that was much, much too complicated. So you have a conversation [with ChatGPT] about simplifying it and making it something you can do in two-minute films? [FG 1]

Expressing Challenges Using Technology

Cross-cutting challenges expressed by the participants included the cost associated with technology, such as data or internet plans or fall alert technology; physical changes that make using technology more complex; and concerns related to fraud and data privacy. The latter was expressed in all focus groups, highlighting the vulnerability felt by older adults to financial scams. Another challenge often mentioned was the need for more digital skills or the inability to use technology, such as cell phones, optimally. This challenge was well aligned with the expectation that older adults must become familiar with current technology, as many tasks related to day-to-day societal functioning have become digitalized. This includes accessing lab results or medical records and online banking. Several participants also disclosed individual physical and sensory changes that made navigating technology interfaces more difficult for them:

I think some of us have physical problems. I have a lot of vision problems, so sometimes I can't even do anything at all because I can't see where the arrow is… the other thing is that I'm losing the ability to grasp with my hands, and that is horrible, but it, you know, it makes it difficult [FG 4]

Identifying Technology Considerations

Considering the challenges reported by our sample of NORC residents, many participants often suggested how to enhance their experience with technology. For example, they were intentional about receiving more technology support through education, training, and workshop-style learning. Our sample strongly desired someone to walk them through technology (eg, a digital navigator), especially when devices are seemingly not working and require troubleshooting advice. The desire for education also extended to avoiding cyber scams and using emerging technology, like ChatGPT. However, some participants resisted entirely internet-based systems and wondered about having either nontech or low-tech alternatives or technology that considered the unique needs of older adults, for example, a slower typing speed. Finally, participants also mentioned low-cost technology options aligned with their budgetary realities.

World Café Findings

According to the workshop participants, the SEM framework helped highlight challenges and facilitators of technology at the individual, interprofessional, community, organizational, policy, and technology device levels (see Table 4).

XSL•FO
RenderX

Table . Social Ecological Model framework.

Level	Challenge	Facilitator
Individual (ie, knowledge, attitudes, and behaviors)	• Fatigue with online appointments and expectation of online solutions (ie, "Zoom fatigue") • Hesitancy in asking for help with technology • Need for repeated exploration to acquire and develop digital literacy ("One-and-done doesn't work") • One's formal education does not equate to digital literacy	• Desire to have a "real" interaction • Identified characteristics of those who adapt to technology: healthier, open-minded, able to afford technology-related costs, willingness to learn and acknowledge technology shortcomings
Interpersonal (ie, family, friends, and social networks)	• Customer service supporting access to technology is not always respectful and informative • Technology cannot replace human connection	• Family available who can help facilitate setup; family available to attend virtual meetings • Free of charge or nominal cost-required services equipped to help older people stay connected (eg, Geek Squad and Toronto Public Library)
Community (ie, relationships between organizations)	• Concerns about fraud and scams targeting older adults • Virtual programming may result in challenges (ie, long waitlists, lack of relief to caregivers, or social participation for users) • Ageism toward older people and technology, including them learning how to use technology	• Technology to see what resources or assets are available in the community (eg, grocery delivery and volunteer services) • Power in organizing a "collective idea or request" that supports collective learning and support for integrating technology solutions into living environments • Having people in your building or community who can help and whom you can trust • Building community and technology support ("economies of scale"), which may foster social engagement and participation • Leverage existing, trusted institutions like public libraries
Organizational (ie, organizations and social institutions)	• Customer service or technology support lacks the skills to communicate effectively with older adults (eg, they speak too loudly and give unclear instructions); people would rather live with the problem than deal with technology support	• Technology tutorials or programs (eg, how to sign a document) • Assistive devices in the home (eg, light sensors, passive sensors, and cane sensors) • Existing support system (eg, NORC[a] community coordinator) • Creation of a "digital navigator" to facilitate solution finding and increase user participation
Policy (ie, national, provincial, and local laws)	• Lack of policy requiring technology designed and customized to the needs of older users who consider sensory or cognitive differences • The government seemingly not engaged in supporting use of technology among older adults	• Affordable NORC phone, cell phone, and internet packages • Having technology partners represented at the policy level
Technology device (ie, features, interface, and design)	• Design does not meet the user needs, who may include older adults • Language translation, colors, and fonts need to be unavailable or easily customizable for everyone (eg, bigger font and different colors)	• Use of plain language or avoidance of acronyms • Considering aging-related changes in technology design (eg, mouse speed and color contrast for vision)

[a]NORC: naturally occurring retirement community.

The workshop discussions revealed that technology is not generally tailored to meet the needs of older adults, and the aging-oriented devices that do exist can be improved through more thoughtful and unique designs. According to the workshop participants, co-design is needed, where technology developers work with older adults from the onset of the conceptualization

XSL•FO
RenderX

and design process. Technology should not only be adapted for older adults but also specifically for different demographics and complex needs (eg, hearing impairment). An example of an improvement would be if device applications included multilingual options. Diverse groups of older adults must be consulted in technology development. Therefore, accessibility was frequently discussed, stressing the need for equity in technology design and access. In addition, the lack of affordability of technologies posed a significant barrier to accessibility, significantly disadvantaging lower-income groups. Participants suggested possible improvements, including reduced internet plans or group packages for NORCs.

The widespread adoption of technology to replace traditional services is becoming the default and can complicate many processes for older adults (eg, banking). Instead, there should continue to be a nononline option in addition to technology. The widespread use of technology is often blamed for reducing meaningful human interaction, which is concerning because people naturally seek social connection. Conversely, other participants voiced that information and communication technology allows connecting with others more easily, like family abroad, especially during scenarios of restricted movement, as in the case of the COVID-19 pandemic. Communication technologies (ie, Zoom [Zoom Communications] and Facetime [Apple Inc]) benefit virtual care, where technology can increase access to health care professionals and reduce wait times. Many participants noted they required more support for digital literacy to be comfortable navigating certain technologies, presenting a need for a tutor or someone available to answer technology-related questions, echoing insights from the focus groups.

It was suggested that aging communities, such as NORCs, should partner with programs or community agencies to support older adults in being technologically proficient. Technological literacy is critical in providing NORC residents with user confidence and helps to reduce anxiety around possible scams or security breaches. It was also suggested that instead of being compelled to accept the terms and conditions, users should be able to dictate the information that technology platforms share and with whom. Concurrently, technology could also provide a safety net at home. Many participants already appreciated using technology in remote monitoring and risk identification. Therefore, the ultimate goal to keep in mind when developing technology is how it can provide essential support in maintaining the independence of older adults to age in their preferred environment, which also evolves over an individual's lifetime.

Synthesizing insights drawn from the focus groups and World Café informed a research agenda (Table 5) of questions mapped to equity-informed technology principles.

XSL•FO
RenderX

Table . Research agenda.

Principles	Research question
Accessible support	• How can the NORC[a] environment support access and use of technology among older residents by leveraging existing and new partnerships (ie, City of Toronto, Toronto Public Library, and Connected Canadians)? • How do we leverage public (eg, Remote Care Monitoring) and private technology (eg, Apple Watch) ecosystems to support NORC residents to age in place? • What might be the most impactful ways in which technology could support community building and participation among NORC residents, local service providers, and agencies?
Accessible interfaces	• What are the technology design principles that consider the evolving needs of older adults, and how can these be integrated into current and future applications?
Affordable and equitable access	• How can technology developers be guided to commit to equity and adhere to it during the design and deployment phases of emerging technology? • How can government enact equitable and inclusive policies and funding to support the adoption of senior-friendly technology (ie, AgeTech)? Or policies to support alternatives to technology?
Available digital literacy training	• How do we create community-based co-located senior-friendly technology education and training opportunities to meet the diverse needs of aging adults? • How can education and training offerings be targeted to certain populations, for example, older adults who are new to Canada, those who are multilingual, are underhoused, and persons who have complex care needs?
Accessible data	• How do we measure the impact of technology on NORC residents, and find data-driven solutions?
Accessible partnerships	• How do we create an intersectoral and collaborative network of partners to support NORC residents to access and use technology, and who should be part of the network?

[a]NORC: naturally occurring retirement community.

Discussions

Principal Findings

We conducted a 2-phase sequential qualitative descriptive study of community-based focus groups and a World Café event to develop a research agenda on older adults' informed considerations for technology to support the aging-in-place of persons living in NORCs. We found several cross-cutting categories across study activities. Participants described using technology for various reasons, including connecting with others, wellness and activity, information seeking, and personal safety. However, our sample's technology usage was also impacted by barriers, such as limited access to (or outdated) devices, the cost associated with technology and its use (eg, internet), and a lack of available training or support to help fully use available technology or address low digital literacy within the context of age-related changes (ie, sensory, cognitive, and functional) that could hinder ease of adoption. The final set of 10 potential research questions underscores the importance of creating opportunities for age-friendly technology education, training, and equity in design, including design principles that reflect the evolving needs of older adults. Finally, the questions also emphasize the need to leverage intersectoral partnerships for supporting NORC residents who inhabit a unique living situation that enables the delivery of support within an existing community of persons aging in place. Therefore, accessing and using technology can serve as a mechanism for community building among NORCs and service providers.

Comparison With Prior Work

Our findings on the barriers to using technology among older NORC residents are similar to other research that found challenges across intrinsic (including individual physical and sensory changes) and extrinsic (including inexperience with technology and cost to access) factors [38]. However, some sociodemographic factors may significantly impact specific groups of older adults, like those on lower or fixed incomes, females, ethnic minority groups, those with lower education levels, and those who are older [39,40]. While our study did not examine the intersectional impact of race and sex on access to technology, other research suggests a significant disparity, with non-White older females having the least access to technology [40]. These findings speak to the ongoing digital

XSL•FO
RenderX

divide, referring to the gap between those with access to technology and those without. According to the literature, significant dimensions of the digital divide include digital literacy, affordability (costs), equity-denied group-related content and services (culturally appropriate content, activities, programs, and services), and access (infrastructure) [41].

NORCs, which include buildings or neighborhoods with a high density of older people, are home to highly diverse and socially and medically complex aging demographics [42,43]. Recent studies have found that NORC residents are older, more female, have lower incomes, live with more chronic conditions, and access home and primary care more often than non-NORC residents [43]. Together, these results suggest that there are more significant health needs than those of the general aging population; however, their communal living environment is particularly well-suited for leveraging. In this way, having a large concentration of older adults allows one to address technology needs in the same way health and social programming is delivered in NORCs [44], mainly through partnerships and collective efforts [45]. While we did not collect NORC participants' income levels, medical histories, or educational backgrounds, leveraging the collective power of a NORC environment to arrange, for example, technology training was mentioned during the World Café forum.

Despite the increased interest in equity-informed technology for older adults, limited research exists regarding NORC residents' adoption of technology that supports aging-in-place. Our findings align with previous interventional technology research (eg, sensory-based passive remote monitoring and smart speakers) implemented in NORC-like environments with lower-income and diverse older adults [46,47]. In these examples, access to a "cultural navigator" supported technology adoption through close relationships and information sessions [46]. However, without high-touch support for technology, its use among older residents may be limited to essential functions [47]. Similarly, NORC residents in another study expressed frustration and fear with using technology to access virtual programming during COVID-19 [48].

Strengths and Limitations

Our study uniquely positions NORCs and related collaborative relationships as one approach to consider when addressing the digital divide. Learnings from the NORC context could inform future digital health research, including design and implementation studies focusing more broadly on aging-in-place. The research agenda drawn from this study provides guiding questions that could help identify priority areas when considering equity-informed technology deployment and use in NORCs; they include viewing the opportunity to support aging in place with technology through various lenses. Finally, multiple data sources informed the agenda development, including a rapid review, a focus group, and a World Café. Despite these strengths, this study is not without limitations. For instance, the study data reflect a group of NORC residents and stakeholder participants. Our sample may not be generalized across other NORC settings, such as those in many rural and northern communities and across racial and ethnic minority and low-income older adults. In these settings, limited high-speed broadband internet continues to be problematic [49]. Likewise, since the target population in our study was NORC residents, certain strategies could be used to gather similar information from other population groups. As noted in the literature, partnering with community organizations, relevant clinicians, health service planners, and policy makers can leverage existing networks while establishing new relationships [28]. These networks not only serve to optimize participation among different community members, but they could also support ongoing initiatives catalyzed by the engagement. We also acknowledge that focus groups and World Café participants from NORCs had some degree of technology experience, which may indicate that we did not hear directly from those with minimal technical literacy. Therefore, a more targeted recruitment approach could have identified those significantly impacted by the digital divide. Another consideration is the potential influence that receiving the gift card could have had on participants' decisions regarding participation, for instance, being unduly incentivized to join the World Café. Finally, our final research agenda did not undergo a consensus exercise with our study participants. Future research could involve finalizing the agenda through a more iterative process, asking participants to rank questions on impact and importance.

Future Directions

Future research could tackle one or more of these questions by identifying conventional and nontraditional intersectoral partnerships between, for example, the public library system, academic health institutions, technology start-ups, the public education system, third-sector and charity organizations, and community-based service and housing providers. These new collaborative intersectoral models could help provide more accessible assistance to older adults within a communal environment who seek technology support [50]. Our study has begun to explore NORCs as unique environments that offer a transformative opportunity to address equity considerations in technology supporting aging in place.

Conclusions

Our findings and research agenda highlight critical areas for consideration, including leveraging partnerships, integrating public and private technology ecosystems, and designing technology with older users that evolves with the population's needs. Notably, by embedding principles of equity, inclusivity, and user-centered design, developers, researchers, and service providers can ensure that emerging technology serves diverse aging populations equitably and effectively.

Acknowledgments

The authors thank all the focus group and World Café participants for their invaluable contributions to the study. The Canadian Institutes of Health Research supported this work (award RN503880 - 498841).

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Jakobi J. Aging in Place Reflections from Canadians: Thea Franke Consulting; 2022.
2. Go C. Aging and Chronic Diseases: A Profile of Canadian Seniors: Public Health Agency of Canada Ottawa; 2020.
3. Bosch-Farré C, Malagón-Aguilera MC, Ballester-Ferrando D, et al. Healthy ageing in place: Enablers and barriers from the perspective of the elderly. A qualitative study. Int J Environ Res Public Health 2020 Sep 4;17(18):6451. [doi: 10.3390/ijerph17186451] [Medline: 32899744]
4. Iciaszczyk N, Arulnamby A, Brydges M, Marshall M, Cheng S, Feil C. Ageing in the Right Place: Supporting Older Canadians to Live Where They Want: National Institute on Ageing (NIA); 2022.
5. Golant SM. The quest for residential normalcy by older adults: Relocation but one pathway. J Aging Stud 2011 Aug;25(3):193-205. [doi: 10.1016/j.jaging.2011.03.003]
6. Golant SM. Residential normalcy and the enriched coping repertoires of successfully aging older adults. Gerontologist 2015 Feb;55(1):70-82. [doi: 10.1093/geront/gnu036] [Medline: 24840917]
7. Cheek P, Nikpour L, Nowlin HD. Aging well with smart technology. Nurs Adm Q 2005;29(4):329-338. [doi: 10.1097/00006216-200510000-00007] [Medline: 16260997]
8. E J, Xia B, Susilawati C, Chen Q, Wang X. An overview of Naturally Occurring Retirement Communities (NORCs) for ageing in place. Buildings 2022;12(5):519. [doi: 10.3390/buildings12050519]
9. National Institute on Ageing and NORC Innovation Centre. It's Time to Unleash the Power of Naturally Occurring Retirement Communities in Canada: National Institute on Ageing, Toronto Metropolitan University and NORC Innovation Centre, University Health Network; 2022.
10. DePaul VG, Parniak S, Nguyen P, et al. Identification and engagement of naturally occurring retirement communities to support healthy aging in Canada: a set of methods for replication. BMC Geriatr 2022 Apr 22;22(1):355. [doi: 10.1186/s12877-022-03045-z] [Medline: 35459126]
11. Joe J, Demiris G. Older adults and mobile phones for health: a review. J Biomed Inform 2013 Oct;46(5):947-954. [doi: 10.1016/j.jbi.2013.06.008] [Medline: 23810858]
12. Liu L, Stroulia E, Nikolaidis I, Miguel-Cruz A, Rios Rincon A. Smart homes and home health monitoring technologies for older adults: a systematic review. Int J Med Inform 2016 Jul;91:44-59. [doi: 10.1016/j.ijmedinf.2016.04.007] [Medline: 27185508]
13. Marshall BL, Dalmer NK, Katz S, Loos E, López Gómez D, Peine A. Digitization of aging-in-place: an international comparison of the value-framing of new technologies. Societies (Basel) 2022;12(2):35. [doi: 10.3390/soc12020035]
14. Spacey J. Low tech vs high tech. Simplicable. 2016. URL: https://simplicable.com/new/low-tech-vs-high-tech [accessed 2025-07-07]
15. MacInnes J, Billings J, Dima AL, Farmer C, Nijpels G. Achieving person-centredness through technologies supporting integrated care for older people living at home: an integrative review. J Integr Care 2021 Jul 19;29(3):274-294. [doi: 10.1108/JICA-03-2021-0013]
16. Kim D, Bian H, Chang CK, Dong L, Margrett J. In-home monitoring technology for aging in place: scoping review. Interact J Med Res 2022 Sep 1;11(2):e39005. [doi: 10.2196/39005] [Medline: 36048502]
17. Grigorovich A, Fang ML, Sixsmith J, Kontos P. Defining and evaluating transdisciplinary research: implications for aging and technology. Disabil Rehabil Assist Technol 2019 Aug;14(6):533-542. [doi: 10.1080/17483107.2018.1496361] [Medline: 30318930]
18. Rubeis G, Fang ML, Sixsmith A. Equity in AgeTech for ageing well in technology-driven places: the role of social determinants in designing AI-based assistive technologies. Sci Eng Ethics 2022 Oct 27;28(6):49. [doi: 10.1007/s11948-022-00397-y] [Medline: 36301408]
19. Lee C. Technology and aging: the jigsaw puzzle of design, development and distribution. Nat Aging 2022 Dec;2(12):1077-1079. [doi: 10.1038/s43587-022-00325-6] [Medline: 37118534]
20. Greer V, Canham SL, Hong A, Agutter J, Garcia Zambrana I, Van Natter JM. Aging in place through the COVID-19 pandemic: perspectives from aging service providers. J Appl Gerontol 2023 Jul;42(7):1530-1540. [doi: 10.1177/07334648231159375] [Medline: 36856262]
21. van Dijk J. Digital divide research, achievements and shortcomings. Poetics (Amst) 2006 Aug;34(4-5):221-235. [doi: 10.1016/j.poetic.2006.05.004]
22. Andrey S, Masoodi MJ, Malli N, Dorkenoo S. Mapping Toronto's digital divide 2021. : Ryerson Leadership Lab and Brookfield Institute for Innovation + Entrepreneurship; 2023 Feb 1 URL: https://www.toronto.ca/wp-content/uploads/2021/07/95dc-Mapping-Torontos-Digital-Divide.pdf [accessed 2025-07-24]
23. Fang ML, Siden E, Korol A, Demestihas MA, Sixsmith J, Sixsmith A. A scoping review exploration of the intended and unintended consequences of eHealth on older people: a health equity impact assessment. Hum Technol 2018;14(3):297-323. [doi: 10.17011/ht/urn.201811224835]

XSL•FO
RenderX

24. Kokorelias KM, Nelson ML, Tang T, et al. Inclusion of older adults in digital health technologies to support hospital-to-home transitions: secondary analysis of a rapid review and equity-informed recommendations. JMIR Aging 2022 Apr 27;5(2):e35925. [doi: 10.2196/35925] [Medline: 35475971]
25. Mitchell UA, Chebli PG, Ruggiero L, Muramatsu N. The digital divide in health-related technology use: the significance of race/ethnicity. Gerontologist 2019 Jan 9;59(1):6-14. [doi: 10.1093/geront/gny138] [Medline: 30452660]
26. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [doi: 10.1002/1098-240X(200008)23:4<334::AID-NUR9>3.0.CO;2-G]
27. Boyko JA, Lavis JN, Abelson J, Dobbins M, Carter N. Deliberative dialogues as a mechanism for knowledge translation and exchange in health systems decision-making. Soc Sci Med 2012 Dec;75(11):1938-1945. [doi: 10.1016/j.socscimed.2012.06.016]
28. MacFarlane A, Galvin R, O'Sullivan M, et al. Participatory methods for research prioritization in primary care: an analysis of the World Café approach in Ireland and the USA. Fam Pract 2017 Jun 1;34(3):278-284. [doi: 10.1093/fampra/cmw104] [Medline: 27677298]
29. Monforte J, Netherway J, Smith B. The world café is an unmethod within co-produced research. Qual Res Psychol 2023 Jul 3;20(3):398-419. [doi: 10.1080/14780887.2023.2239728]
30. Löhr K, Weinhardt M, Sieber S. The "World Café" as a participatory method for collecting qualitative data. Int J Qual Methods 2020 Jan 1;19:1609406920916976. [doi: 10.1177/1609406920916976]
31. Khong L, Bulsara C, Hill KD, Hill AM. How older adults would like falls prevention information delivered: fresh insights from a World Café forum. Ageing Soc 2017 Jul;37(6):1179-1196. [doi: 10.1017/S0144686X16000192]
32. Street J, Barrie H, Eliott J, et al. Older adults' perspectives of smart technologies to support aging at home: insights from five World Café forums. Int J Environ Res Public Health 2022 Jun 25;19(13):7817. [doi: 10.3390/ijerph19137817] [Medline: 35805477]
33. Huynh T, Sava N, Hahn-Goldberg S, et al. Mobile on-site COVID-19 vaccination of naturally occurring retirement communities by neighbourhood risk in Toronto. Science Briefs - Ontario COVID-19 Science Advisory Table. 2021. URL: https://covid19-sciencetable.ca/sciencebrief/mobile-on-site-covid-19-vaccination-of-naturally-occurring-retirement-communities-by-neighbourhood-risk-in-toronto/ [accessed 2024-07-24]
34. Ollevier A, Aguiar G, Palomino M, Simpelaere IS. How can technology support ageing in place in healthy older adults? A systematic review. Public Health Rev 2020 Nov 23;41(1):26. [doi: 10.1186/s40985-020-00143-4] [Medline: 33292707]
35. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
36. Stokols D. Establishing and maintaining healthy environments. Toward a social ecology of health promotion. Am Psychol 1992 Jan;47(1):6-22. [doi: 10.1037//0003-066x.47.1.6] [Medline: 1539925]
37. Olaniyan A, Isiguzo C, Hawk M. The Socioecological Model as a framework for exploring factors influencing childhood immunization uptake in Lagos state, Nigeria. BMC Public Health 2021 May 5;21(1):867. [doi: 10.1186/s12889-021-10922-6] [Medline: 33952252]
38. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556. [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
39. Chan DYL, Lee SWH, Teh PL. Factors influencing technology use among low-income older adults: a systematic review. Heliyon 2023 Sep;9(9):e20111. [doi: 10.1016/j.heliyon.2023.e20111] [Medline: 37809586]
40. Suntai Z, Beltran SJ. The intersectional impact of race/ethnicity and sex on access to technology among older adults. Gerontologist 2023 Aug 24;63(7):1162-1171. [doi: 10.1093/geront/gnac178] [Medline: 36477498]
41. Raihan MMH, Subroto S, Chowdhury N, Koch K, Ruttan E, Turin TC. Dimensions and barriers for digital (in)equity and digital divide: a systematic integrative review. Digit Transform Soc 2025 May 6;4(2):111-127. [doi: 10.1108/DTS-04-2024-0054]
42. Sheth MS, Rochon PA, Altaf A, et al. Factors associated with transition to a nursing home in older adults living in naturally occurring retirement communities. J Am Geriatr Soc 2025 Jan;73(1):88-100. [doi: 10.1111/jgs.19038] [Medline: 38888381]
43. Savage RD, Huynh T, Hahn-Goldberg S, et al. A portrait of older adults in naturally occurring retirement communities in Ontario, Canada: a population-based study. J Am Geriatr Soc 2025 Jan;73(1):74-87. [doi: 10.1111/jgs.19278] [Medline: 39618101]
44. Mills CM, Parniak S, Hand C, et al. The impact of a naturally occurring retirement community supportive services program on older adult participants' social networks: semistructured interview study. JMIR Aging 2022 Nov 21;5(4):e37617. [doi: 10.2196/37617] [Medline: 36409533]
45. Parniak S, DePaul VG, Frymire C, DePaul S, Donnelly C. Naturally occurring retirement communities: scoping review. JMIR Aging 2022 Apr 14;5(2):e34577. [doi: 10.2196/34577] [Medline: 35436204]
46. Berridge C, Chan KT, Choi Y. Sensor-based passive remote monitoring and discordant values: qualitative study of the experiences of low-income immigrant elders in the United States. JMIR Mhealth Uhealth 2019 Mar 25;7(3):e11516. [doi: 10.2196/11516] [Medline: 30907741]

XSL•FO
RenderX

47. Chung J, Bleich M, Wheeler DC, et al. Attitudes and perceptions toward voice-operated smart speakers among low-income senior housing residents: comparison of pre- and post-installation surveys. Gerontol Geriatr Med 2021;7:23337214211005869. [doi: 10.1177/23337214211005869] [Medline: 35047655]
48. Fernandes K, Hand C, Laliberte Rudman D, et al. Being and doing together in a naturally occurring retirement community: pandemic experiences of older adults. Can J Aging 2025 Mar;44(1):115-125. [doi: 10.1017/S0714980824000345] [Medline: 39610356]
49. Freeman S, Marston HR, Ross C, et al. Progress towards enhanced access and use of technology during the COVID-19 pandemic: a need to be mindful of the continued digital divide for many rural and northern communities. In: Gates J, editor. Healthc Manage Forum 2022 Sep;35(5):286-290. [doi: 10.1177/08404704221108314] [Medline: 35855623]
50. Golant SM. Commentary: Irrational exuberance for the aging in place of vulnerable low-income older homeowners. J Aging Soc Policy 2008;20(4):379-397. [doi: 10.1080/08959420802131437] [Medline: 19042553]

Abbreviations

NORC: naturally occurring retirement community
SEM: Social Ecological Model

Edited by AF Abdulai; submitted 09.01.25; peer-reviewed by B Davis, J Scott, N Novik; revised version received 23.05.25; accepted 20.06.25; published 30.07.25.

Please cite as:
Saragosa M, Abejirinde IOO, MacEachern E, Nelson MLA, Kokorelias KM, Bharmal S, Ludwig Prout B, Mohamed M
Exploring Technology Supporting Aging-in-Place Using an Equity Lens Through Focus Groups and World Café–Informed Research Agenda: Qualitative Study
JMIR Aging 2025;8:e71093
URL: https://aging.jmir.org/2025/1/e71093
doi:10.2196/71093

© Marianne Saragosa, Ibukun-Oluwa Omolade Abejirinde, Evan MacEachern, Michelle LA Nelson, Kristina M Kokorelias, Sidra Bharmal, Brina Ludwig Prout, Marian Mohamed. Originally published in JMIR Aging (https://aging.jmir.org), 30.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Caregiver Support Model for Informal Caregivers of Frail Older Adults: Randomized Controlled Trial

Cyrus Lap Kwan Leung[1,2*], PhD; Kin-Kit Li[2*], PhD; Dannii Y Yeung[2], PhD; Alice Ming Lin Chong[3,4], PhD; Marcus Yu Lung Chiu[3], PhD; Xin Guan[2,5], PhD; Tit Wing Lo[3], PhD

[1]The Jockey Club School of Public Health and Primary Care, Faculty of Medicine, Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

[2]Department of Social and Behavioural Sciences, College of Liberal Arts and Social Sciences, City University of Hong Kong, Tat Chee Avenue, Kowloon Tong, Hong Kong, China (Hong Kong)

[3]Felizberta Lo Padilla Tong School of Social Sciences, Saint Francis University, Hong Kong, China (Hong Kong)

[4]Caregiving Research and Development Centre, Saint Francis University, Hong Kong, China (Hong Kong)

[5]Department of Social Work and Social Administration, Faculty of Social Sciences, University of Hong Kong, Hong Kong, China (Hong Kong)

[*]these authors contributed equally

Corresponding Author:
Dannii Y Yeung, PhD
Department of Social and Behavioural Sciences, College of Liberal Arts and Social Sciences, City University of Hong Kong, Tat Chee Avenue, Kowloon Tong, Hong Kong, China (Hong Kong)

Abstract

Background: Caregivers of frail older adults face substantial challenges, often managing their own health while providing care. To address these issues, we developed the caregiver support model (CSM), a structured approach that uses systematic assessment, personalized intervention planning, and sustained support to address informal family caregivers' diverse and evolving needs and leverage their resources.

Objective: This study aims to evaluate the effectiveness of CSM.

Methods: A blinded cluster randomized controlled trial was conducted across 8 centers providing services for older adults in Hong Kong. The CSM is a social worker–guided intervention that integrates a structured assessment of caregiver needs and resources, personalized service planning, and ongoing monitoring over 6 months. Meanwhile, the control group continued with their usual procedures without a standardized caregiver assessment. Data were collected at baseline, 3 months, and 6 months.

Results: We recruited 565 informal family caregivers (281/565, 49.7% CSM intervention; 284/565, 50.3% standard care control). Both groups improved over time; compared with the control group, the CSM produced greater reductions in caregiver needs, particularly in role conflict, and greater gains in resources, such as health awareness. Improvements were more pronounced at 6 months compared to 3 months, indicating a lasting effect and consolidation of gains. The intervention was particularly effective for caregivers in other relationships (not spouse or child) and those with higher education than spousal caregivers.

Conclusions: These findings highlight the importance of long-term tailored interventions that adapt to the evolving needs of caregivers through systematic assessment. The CSM offers a promising approach to enhancing the well-being of caregivers and managing the complex demands of caregiving, particularly in an aging population.

Trial Registration: ClinicalTrials.gov NCT04272918; https://www.clinicaltrials.gov/study/NCT04272918

(JMIR Aging 2025;8:e71638) doi:10.2196/71638

KEYWORDS

informal caregiving; caregiver support; randomized controlled trial; caregiving resources; caregiving burden

Introduction

Growing Care Demands and Service Gaps

Hong Kong, like many other high-income international cities, is increasingly grappling with the complex challenges posed by an aging population. The number of adults aged ≥60 years surged from 1.44 million (20.1% of the total population) in 2012 to approximately 2.2 million (29.9%) in 2022, with projections indicating a further rise to 3.21 million (40.1%) by 2050 [1]. This demographic shift has been accompanied by a growing reliance on informal caregivers. In 2021, over three-quarters of older adults needing long-term care (16.6% of those aged ≥60 years) relied on informal caregiving [2]. Given that long-term care needs often stem from age-related frailty, which is associated with increased dependency in mobility, personal care, and supervision, caregivers are frequently tasked with managing complex physical, emotional, and logistical demands,

XSL•FO
RenderX

especially in smaller households where responsibilities often fall on the spouse or adult child [3-5]. These converging trends underscore the pressing need for a well-structured support system that both acknowledges the critical role of informal caregivers and provides resources to enhance their ability to sustain the well-being and dignity of Hong Kong's aging population.

The public health and social welfare systems in Hong Kong currently face significant gaps in providing adequate support for family caregivers, driving many to rely on private care services [6]. While these private services provide essential assistance, they are often expensive for low-income families [4,5]. To bridge this gap, the Social Welfare Department has introduced a range of services to support caregivers [7], including a designated caregiver support hotline, an information portal, and community-based programs for older adults and their caregivers. In addition, local nongovernmental organizations (NGOs) offer various services, such as information dissemination, referrals, skills training, and support groups. Despite these efforts, recent findings highlight persistent structural limitations, such as inadequate workplace flexibility, fragmented service access, and a lack of cross-sector coordination [8,9]. In addition, a critical shortfall remains in the form of unstandardized guidelines for assessing and addressing caregiver needs [10], along with limited training and preparedness among frontline social workers. This structural deficiency hampers the effective management and distribution of resources. To address these challenges, we developed the caregiver support model (CSM), which offers a structured and comprehensive framework for assessment, case management, and training. The CSM is a critical step toward a more integrated approach to supporting caregivers, helping to alleviate the growing strain on informal family caregivers within Hong Kong's health care system. Endorsing an assessment-based, balanced, caregiver-centered, and case-management approach, the CSM has the potential to inform future practices in other countries or cities with similar social services infrastructure.

Dynamic Nature of Caregiving and Theoretical Foundations of the CSM

The CSM is grounded on well-established theoretical frameworks, incorporating key elements from the stress process model [11], the stress and coping model [12,13], and empowerment and strengths-based approaches. The stress process model highlights how primary stressors (eg, functional limitations of the care recipient) and secondary strains (eg, role conflict and psychological demands) shape caregiver burden, while coping strategies and available resources act as mediators. The stress and coping model complements this by emphasizing the process of stress appraisal, coping, and reappraisal over time, illustrating why structured assessment and repeated follow-up are necessary. Together, these models conceptualize caregiving as an evolving process and provide the rationale for interventions that both reduce stressors and strengthen protective resources. Central to the CSM is the empowerment of caregivers, which encourages them to recognize their strengths, advocate for their needs, and adopt sustainable caregiving practices. This is closely aligned with the strengths-based approach [14,15], which focuses on identifying and harnessing caregivers' inherent capacities to overcome challenges, thereby fostering resilience and better adaptation to the caregiving role.

Recognizing caregiving as a dynamic process, the CSM identifies distinct stages of caregiving [16], each presenting unique challenges and needs. Accurately assessing these needs is crucial as they shift alongside the care recipient's illness progression and the evolving role and responsibilities of the caregiver [17,18]. In Hong Kong, several standardized tools are used for caregiver assessment, such as the 12-item Zarit Burden Interview (ZBI-12) [19], the Perceived Stress Scale [20], and the 21-item Depression Anxiety Stress Scale [21]. While these tools offer valuable insights, they often fail to capture the multifaceted nature of caregiving, particularly in areas such as role conflict and the personal needs of caregivers, both of which significantly impact caregiver burden and satisfaction [22-24]. This limitation highlights the need for a comprehensive, standardized tool tailored for caregivers of frail older persons. Such a tool must not only capture caregivers' diverse needs but also evaluate the resources that help them cope with caregiving demands. Needs include physiological, psychological, social, and role-related challenges, while resources encompass personal, relational, and community assets that buffer stress [25,26]. To address this gap, the Caregiver Needs and Resources Assessment (CNRA) [10] was developed as a holistic measure. In the CNRA, needs are assessed through subscales covering physiological strain, psychological distress, social support needs, role conflict, and care recipient needs. Resources are defined to include both personal assets (eg, self-efficacy, spirituality, and health awareness) and external supports (eg, family support, closeness with the care recipient, community resources, and responsibility). This dual focus allows the CNRA to guide interventions that both reduce unmet needs and strengthen supports.

Another key component of the CSM is case management, which is essential for coordinating services and interventions that are tailored to the specific needs and circumstances of caregivers [27,28]. In Hong Kong, however, the lack of standardized processes for assessing caregiver needs often leads to subjective evaluations by social workers or health care professionals. To address this issue, the CSM applies the CNRA. This tool empowers social workers to develop client-centered intervention plans and allocate resources more effectively, ensuring that the support provided is closely aligned with each caregiver's unique situation and needs in performing their caregiving role.

This paper aims to evaluate the effectiveness of the CSM through a clustered randomized controlled trial. The CSM represents a significant innovation in gerontology and social work, offering social workers a scientifically validated, comprehensive, and caregiver-centered framework to assist and empower family caregivers of frail older adults. The model includes a range of tools, such as the CNRA, an online caregiver management system, personalized intervention and monitoring templates, and detailed operational guidelines. By rigorously evaluating the CSM's impact on caregiver support, this study seeks to advance the development of evidence-based strategies in informal care, addressing the complex needs of caregivers of frail older adults in Hong Kong and across the world. We hypothesized that participants in the CSM group would have

XSL•FO
RenderX

more favorable outcomes compared with those in the control group at the end of the intervention trial.

Methods

Study Design

The study used a cluster randomized controlled trial design, with parallel allocation to either intervention or control arms, to evaluate the effectiveness of the CSM. The CSM is a social worker–guided intervention and comprises four key components: (1) the CNRA [10], (2) the formulation of a caregiver intervention plan, (3) service provision and monitoring, and (4) evaluation and termination. Eight centers from the 4 participating NGOs were randomly assigned to either the intervention or control group. Although the service administrators could not be blinded from the intervention, participants of the study were not aware of their assigned conditions. All caregiver participants, regardless of their conditions, completed the CNRA before their first meeting with the social worker. The intervention group followed the CSM protocol, in which the CNRA scores were shared with the social workers, who explained the need and resource scores to the participants and then generated a personalized intervention plan for each of them, followed by recommending appropriate services. The social workers in the control group, without knowledge of the CNRA scores, adhered to their own standard procedures in assessing and differentiating caregivers' needs when they sought assistance. The intervention was delivered between January 1, 2022, and August 31, 2022, with data collection at baseline, 3 months, and 6 months after intervention, concluding by February 28, 2023. The fifth wave of the COVID-19 pandemic led to a temporary pause in recruitment, prompting adaptations in service delivery and subsequent division of the recruitment period into 2 phases (phase 1: January 1, 2022, to April 30, 2022; phase 2: May 1, 2022, to August 31, 2022). The study design and reporting followed the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) guidelines (Checklist 1) to ensure transparency and methodological rigor in the conduct and reporting of this randomized controlled trial.

Participants

Participants in the study were informal family caregivers of frail older adults recruited offline from 8 centers of the 4 participating NGOs in Hong Kong. Eligibility criteria included being an unpaid primary caregiver of a family member aged ≥60 years, being at least the age of 21 years, either newly assuming caregiving responsibilities or facing an urgent need for assistance, scoring ≥13 on the Short-Form ZBI-12 to indicate their needs for receiving service support, as well as with the care recipients scoring ≥3 on at least 1 item of the activities of daily living or instrumental activities of daily living. Exclusion criteria included having received formal psychological intervention within the past 3 months. A total of 580 eligible caregivers were initially recruited, with 565 completing the CNRA at baseline (intervention group: 281/565, 49.7%; control group: 284/565, 50.3%). A total of 66 caregivers did not complete all of the follow-up assessments, resulting in an 11.7% (66/565) dropout rate.

Intervention Protocol

The intervention protocol of the CSM involved a structured process that used the CNRA to assess caregivers' needs and available resources. Once eligible caregivers were identified, a research assistant or a program worker administered the CNRA (T1), and the results were uploaded to an online management system for review by the intervention group's social workers. These social workers then met with the caregivers to discuss their CNRA results and develop intervention plans, using spider web charts to visually illustrate needs and resources, thereby enhancing caregivers' awareness of their own strengths and challenges. To empower caregivers in managing the services they receive, they were encouraged to prioritize their needs based on the discussion surrounding the test scores. The support services received by the caregivers in the intervention group were customized based on their needs and resources, while the participants in the control group received regular support services as usual.

Due to the COVID-19 pandemic, adaptations were made, including extended timelines and the use of mixed methods (face-to-face, video calls, and Zoom [Zoom Communications, Inc] meetings) for the first meeting. Regular follow-ups, either meetings, phone calls, or SMS text messaging, were conducted every month, with each meeting lasting no less than 15 minutes. There was also a midterm evaluation at 3 months (T2) after the baseline assessment, in which the caregivers' needs and resources were assessed again using the CNRA. The intervention plans were then adjusted based on the updated CNRA results. The final CNRA assessment took place 6 months after intake (T3), determining whether the intervention would be continued or terminated. In contrast, social workers in the control group, who were unaware of the CNRA scores, relied on personal judgment and observations to guide their intervention plans. The same CNRA assessment schedule at T2 and T3 was adopted for the control group, even though the scores were not disclosed to the social workers in this group.

The development of the CSM was guided by the stress process model [11] and the stress and coping model [12,13]. With reference to the stress process model, the CNRA was structured to capture primary stressors (eg, care recipient's needs), secondary strains (role conflict, psychological needs, and social support needs), and mediating resources (self-efficacy, family support, spirituality, community resources, and other supportive assets). The stress and coping model informed the intervention process: the baseline CNRA and spider chart feedback facilitated caregivers' appraisal of needs and resources; the individualized service plan supported both problem-focused coping (eg, training and resource referral) and emotion-focused coping (eg, counseling and peer support); and the 3- and 6-month reassessments enabled structured reappraisal and adjustment. Together, these frameworks shaped the CSM as a structured process to assess needs, mobilize resources, and reinforce coping over time. A detailed CSM protocol is provided in Multimedia Appendix 1. This study was registered at ClinicalTrials.gov (identifier: NCT04272918). We adhered to the registered

XSL•FO
RenderX

protocol (version as of February 21, 2020) except for the changes in some of the outcome measures (ZBI-12 and World Health Organization-5 were replaced by CNRA and Positive Aspects of Caregiving Scale [PACS]) after considering the input from participating NGOs. Quality assurance measures included recording and reviewing the first and second meetings with caregivers using a standardized checklist to ensure adherence to the protocol. Adherence rates for both meetings were assessed across centers, with research assistants providing timely feedback to social workers for continuous improvement.

Evaluation and Measurements

The following measurement tools were used in the offline baselines and follow-up assessments for assessing the primary outcomes (that are direct and immediate) and the secondary outcomes (that are more general and distal).

Primary Outcomes

CNRA Tool

The CNRA [10] was used to conduct a comprehensive assessment of the needs and resources of the participating caregivers. A revised version of the original 36-item tool was adopted (Leung et al, unpublished data), incorporating minor wording updates to the closeness with the care recipients subscale. CNRA assesses 5 caregiver needs domains (physiological, role conflict, care recipient's needs, psychological, and social) and 7 resource domains (spirituality, self-efficacy, responsibility, community resources, family support, closeness with the care recipient, and healthy lifestyle). Participants rated each item on a 5-point Likert scale from 1=never to 5=all the time about their caregiving experience in the past month. Higher scores on resource subscales reflect more perceived resources; higher scores on need subscales reflect greater unmet needs. Sample items include "Caring for him/her keeps me from getting enough sleep" (needs) and "Family members help me manage the difficulties of caregiving" (resources). The CNRA demonstrated high internal reliability, with the Cronbach α ranging from 0.74 to 0.97 across different subscales and time points. Two overall need scores are presented: one using the full scale and the other excluding the physiological and psychological needs (designated as "best practice"). This distinction was made because physiological needs are typically addressed through medical interventions, while psychological needs can generally be improved by social workers' active interventions that benefit both intervention and control groups.

PACS Measure

The 11-item PACS was used to measure the positive feelings associated with caregiving, which consisted of the self-affirmation (affirming aspects) and outlook on life (enriching aspects) subscales. It uses a 5-point Likert scale from 1=not at all to 5=all the time, with higher scores indicating greater caregiving satisfaction. Sample items include "Given more meaning to my life" and "Enabled me to learn new skills." The PACS has demonstrated high test-retest reliability and moderate construct validity, with the Cronbach α ranging from 0.82 to 0.89 at 3 time points.

Secondary Outcomes

EQ-5D-5L

The EQ-5D-5L [29] was used to assess the quality of life and well-being across 5 dimensions: mobility, self-care, usual activities, pain or discomfort, and anxiety or depression. It includes a visual analog scale and 5 questions rated on a 5-point Likert scale, with higher scores indicating better health and well-being. The EQ-5D-5L has been validated for use in Hong Kong [30], with Cronbach α ranging from 0.70 to 0.75 in this study.

12-item General Health Questionnaire

The 12-item General Health Questionnaire (GHQ-12) [31] was used to assess the participants' mental health status. Respondents rated 12 statements using a 4-point Likert scale from 0=not at all to 3=all the time, with higher scores reflecting greater psychiatric distress. The GHQ-12 has been validated for the Chinese population and has demonstrated good construct validity and predictive reliability [32], with Cronbach α ranging from 0.86 to 0.88 in this study.

Peace of Mind Scale

The Peace of Mind Scale (PoMS) [33] was used to assess affective well-being. Participants rated 7 items on a 5-point Likert scale from 1=not at all to 5=all the time, with higher scores indicating greater peace of mind. Sample items include "My mind is free and at ease." The PoMS demonstrated good reliability, with Cronbach α ranging from 0.92 to 0.94 in this study.

Statistical Analysis

The demographic characteristics of caregivers and their care recipients in both intervention and control groups were summarized using mean scores and SDs for continuous variables, and percentages for categorical variables. Baseline equivalence between the groups was assessed using 2-tailed *t* tests for continuous variables and Fisher exact test for categorical variables. The 2-tailed *t* tests were also used to compare primary and secondary outcome measures between the intervention and control groups at baseline. The effectiveness of the intervention was evaluated using an intention-to-treat approach, which included all participants according to their original group assignment, with missing data imputed via chained equations [34].

To examine intervention effects, we modeled change scores (T2–T1 and T3–T1) as the primary outcomes. Multilevel regression models were estimated with random intercepts for centers to account for the clustered trial design. Models adjusted for the demographic characteristics of both caregivers and care recipients, including the caregiver's age, sex, education level, long-term illness, marital status, years of caregiving, and caregiving stage, as well as the care recipient's age, sex, and functional status (activities of daily living and instrumental activities of daily living). Robust small-sample corrections (cluster-robust variance estimators, specifically the bias-reduced linearization adjustment) were applied to SEs and *P* values to improve the reliability of inference given the limited number of clusters (in the center level). The potential moderating effects

XSL•FO
RenderX

of caregiver and care recipient characteristics on the primary outcomes were assessed using interaction terms between the intervention condition and each demographic variable (eg, education and caregiver-recipient relationship) in separate multilevel regression models. Each model included the intervention, the moderator, their interaction, and a consistent set of covariates. All models accounted for clustering within centers and were estimated using imputed datasets.

Ethical Considerations

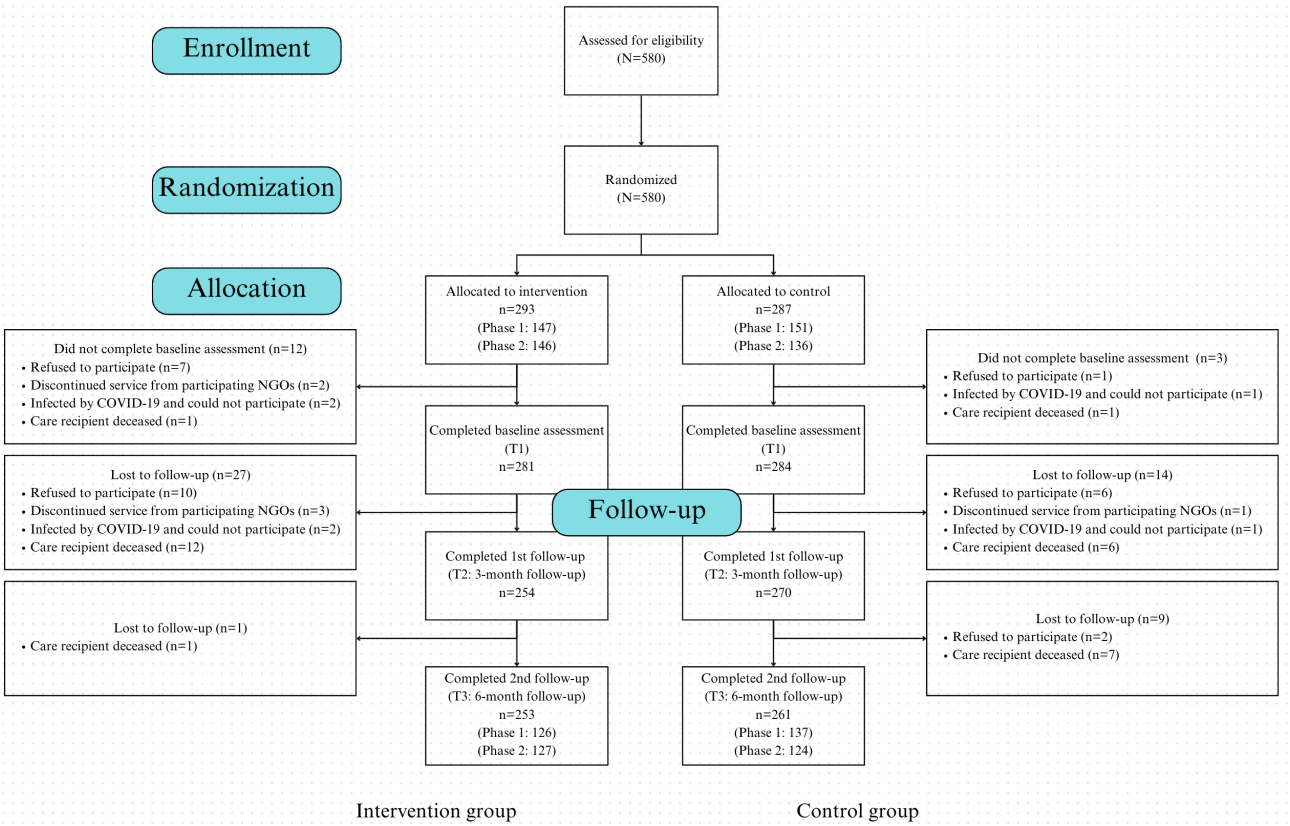
The Human Subject Ethics Subcommittee of the City University of Hong Kong granted ethical approval for the study (reference 3-7-201905-01). Informed consent was obtained from all participants offline, ensuring voluntary participation. Personal identifiers were removed from the dataset, and all data were accessible only to authorized research personnel. Participants were compensated for their time involvement with an HK $150 (US $19.30) supermarket cash voucher upon completing the assessment survey of each round (T1, T2, and T3).

Results

Demographic Characteristics and Baseline Assessment of the Participants

Participant enrollment and allocation are shown in Figure 1. The study sample consisted of caregivers with an average age of 67.0 (SD 11.7) years, the majority of whom were female (478/565, 85%). Educational attainment varied, with the largest group having completed secondary school (234/565, 41%). Most caregivers were married (424/565, 75%), and a significant proportion (439/565, 78%) reported having a chronic disease. The care recipients had an average age of 81.2 (SD 8.9) years, with a majority being male (333/565, 59%). Participant characteristics for the entire sample and by study site information are presented in Table 1 and Table S1 of Multimedia Appendix 1, respectively. No important harms or unintended effects were found in each group.

Figure 1. CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) flow diagram of the 2 study groups. NGO: nongovernmental organization.



XSL•FO
RenderX

Table . Descriptive statistics of the participants (N=565)[a].

	Overall	Intervention (n=281)	Control (n=284)	*P* value
Characteristics of the caregivers				
Age (y), mean (SD; range)	67.0 (11.7; 21 to 96)	66.7 (12.3; 21 to 96)	67.3 (11.1; 26 to 88)	.52
Sex, n (%)				.27
Male	87 (15.4)	48 (17.1)	39 (13.7)	
Female	478 (84.6)	233 (82.9)	245 (86.3)	
Education level, n (%)				.58
No education received	35 (6.2)	19 (6.8)	16 (5.6)	
Private school	20 (43.5)	11 (3.9)	9 (3.2)	
Primary school	174 (30.8)	94 (33.5)	80 (28.2)	
Secondary school	234 (41.4)	113 (40.2)	121 (42.6)	
Certificate or diploma	35 (6.2)	16 (5.7)	19 (6.7)	
Undergraduate or greater	67 (11.9)	28 (10.0)	39 (13.7)	
Marital status, n (%)				.68
Married	424 (75.0)	213 (75.8)	211 (74.3)	
Not married	141 (25.0)	68 (24.2)	73 (25.7)	
Long-term illness, n (%)				.32
With	439 (77.7)	215 (76.5)	224 (78.9)	
Without	126 (22.3)	66 (23.5)	60 (21.1)	
Caregiving stage, n (%)				.29
Stage 1	40 (7.1)	16 (5.7)	24 (8.5)	
Stage 2	66 (11.7)	32 (11.4)	34 (12.0)	
Stage 3	119 (21.1)	59 (21.0)	60 (21.1)	
Stage 4A	126 (22.3)	75 (26.7)	52 (18.3)	
Stage 4B	201 (35.6)	95 (33.8)	106 (37.3)	
Stage 5	12 (2.1)	4 (1.4)	8 (2.8)	
Characteristics of the care recipients				
Age (y), mean (SD; range)	81.2 (8.9; 61 to 106)	80.9 (8.8; 61 to 106)	81.5 (9.0; 61 to 102)	.43
Sex, n (%)				.36
Male	333 (58.9)	171 (60.9)	162 (57)	
Female	232 (41.1)	110 (39.1)	122 (43)	
IADL[b], mean (SD; range)	8.1 (4.4; –3.2 to 18.0)	2.0 (0.6; 0.0 to 18.0)	1.7 (0.5; –3.2 to 18.0)	<.001

XSL•FO
RenderX

	Overall	Intervention (n=281)	Control (n=284)	*P* value
ADL[c], mean (SD; range)	4.1 (4.2; –6.2 to 18.0)	2.2 (0.7; 0.0 to 18.0)	2.0 (0.5; –6.2 to 18.0)	<.001

[a]Imputed data are used.

[b]IADL: instrumental activities of daily living.

[c]ADL: activities of daily living.

At baseline, the intervention group reported a significantly higher overall need score (best practice) compared to the control group (t_{563}=3.20; *P*=.001; mean difference [MD] 0.2, 95% CI 0.1-0.4). In terms of needs subscales, the intervention group scored significantly higher in role conflict (t_{563}=2.96; *P*=.003; MD 0.2, 95% CI 0.1-0.4), and care recipient need than the control group at the baseline (t_{563}=3.20; *P*=.001; MD 0.3, 95% CI 0.1 - 0.4). No significant differences were observed in physiological needs (t_{563}=1.91; *P*=.06; MD 0.2, 95% CI −0.0 to 0.3), psychological needs (t_{563}=1.64; *P*=.10; MD 0.2, 95% CI −0.0 to 0.3), and social support needs (t_{563}=1.73; *P*=.08; MD 0.2, 95% CI −0.0 to 0.4), indicating baseline equivalence between the groups in these domains.

Regarding resources, the intervention group had a lower overall mean score at baseline compared to the control group (t_{563}=−2.49; *P*=.01; MD −0.1, 95% CI −0.2 to −0.0). Subscales such as spirituality (t_{563}=−2.26; *P*=.02; MD −0.2, 95% CI −0.4 to −0.0), family support (t_{563}=−2.53; *P*=.01; MD −0.2, 95% CI −0.4 to −0.0), and closeness with the care recipient (t_{563}=−2.09; *P*=.04; MD −0.2, 95% CI −0.3 to −0.0) also showed significant differences between the 2 groups at baseline.

For secondary outcomes, no significant differences were found in GHQ-12 (t_{563}=1.06; *P*=.29; MD 0.0, 95% CI −0.0 to 0.1) and peace of mind (t_{563}=−1.00; *P*=.32; MD −0.1, 95% CI −0.2 to 0.1), further indicating baseline equivalence. There were no significant differences in the affirming (t_{563}=1.67; *P*=.10; MD 0.1, 95% CI −0.0 to 0.3) and enriching (t_{563}=−1.64; *P*=.10; MD −0.1, 95% CI −0.3 to 0.0) subscales of the PACS between the 2 groups at baseline. However, the intervention group exhibited a significantly lower EQ-5D score compared to the control group (t_{563}=−2.79; *P*=.005; MD 0.0, 95% CI −0.1 to −0.0). Details of baseline equivalence are shown in Table 2.

XSL•FO
RenderX

Table . Baseline comparison on outcome measure between the intervention and control groups (N=565)[a]. The 12-item General Health Questionnaire scores in the range of 0 - 36 can be obtained by multiplying the GHQ-12 mean score by 12.

	Overall, mean (SD; range)	Intervention (n=281), mean (SD)	Control (n=284), mean (SD)	*P* value
Primary outcomes				
Needs (full scale)	2.6 (0.9; 1 to 5)	2.7 (0.9)	2.5 (0.9)	.005[b]
Physiological need	2.4 (1.0; 1 to 5)	2.5 (1.0)	2.3 (1.0)	.06
Role conflict	2.3 (1.0; 1-5)	2.4 (1.0)	2.2 (0.9)	.003[b]
Care recipient need	2.9 (1.0; 1-5)	3.0 (1.0)	2.8 (1.0)	.001[b]
Psychological need	2.7 (1.1; 1 to 5)	2.8 (1.1)	2.7 (1.1)	.10
Social support need	2.8 (1.3; 1 to 5)	2.9 (1.3)	2.7 (1.3)	.08
Needs (best practice)	2.7 (0.9; 1 to 5)	2.8 (0.9)	2.6 (0.9)	.001[b]
Resources	3.3 (0.6; 1 to 5)	3.2 (0.6)	3.4 (0.6)	.01[c]
Spirituality	3.4 (1.1; 1 to 5)	3.3 (1.1)	3.5 (1.0)	.02[c]
Self-efficacy	3.3 (0.8; 1 to 5)	3.3 (0.8)	3.3 (0.8)	.56
Responsibility	4.1 (0.8; 1 to 5)	4.1 (0.7)	4.1 (0.8)	.80
Community resources	2.5 (1.2; 1 to 5)	2.5 (1.2)	2.6 (1.2)	.16
Family support	3.3 (1.0; 1 to 5)	3.2 (1.1)	3.4 (1.0)	.01[c]
Closeness with care recipient	3.4 (0.9; 1 to 5)	3.4 (0.9)	3.5 (0.9)	.04[c]
Health awareness	2.8 (1.1; 1 to 5)	2.8 (1.1)	2.9 (1.0)	.29
PACS[d]	3.2 (0.8; 1 to 5)	3.3 (0.8)	3.2(0.8)	.88
Affirming	3.3 (0.8; 1 to 5)	3.4 (0.8)	3.3 (0.8)	.10
Enriching	3.1 (0.9; 1 to 5)	3.1 (0.9)	3.2 (0.8)	.10
Secondary outcomes				
GHQ-12 (mean score)	1.2 (0.5; 0 to 3)	1.3 (0.5)	1.2 (0.5)	.29
Peace of Mind Scale	3.1 (0.9; 0 to 3)	3.1 (0.9)	3.2 (0.9)	.32
EQ-5D	0.8 (0.2; −0.5 to 1)	0.7 (0.2)	0.8 (0.2)	.005[b]

[a]Imputed data are used.

[b]*P*<.01.

[c]*P*<.05.

[d]PACS: Positive Aspects of Caregiving Scale.

Multilevel Regressions Examining Intervention Effects

Intraclass correlation coefficients (ICC) for the center-level random effects ranged from 0.00 to 0.11 across outcomes (Table 3), indicating that only a small proportion of variance was attributable to clustering. This supports the use of random intercepts for centers to reflect the clustered trial design, while showing that clustering effects were relatively modest.

XSL•FO
RenderX

Table . Multilevel regressions predicting changes in the caregiving needs and resources (N=565)[a].

	T2-T1						T3-T1					
	Control (n=284)		Intervention (n=281)		ICC[b]	*P* value	Control (n=284)		Intervention (n=281)		ICC	*P* value
	Estimated mean (95% CI)	Missing, n	Estimated mean (95% CI)	Missing, n			Estimated mean (95% CI)	Missing, n	Estimated mean (95% CI)	Missing, n		
Primary outcomes												
Needs (full scale)	*−0.26 (−0.39 to −0.13)* [c]	15	*−0.31 (−0.47 to −0.14)*	27	0.01	.60	*−0.21 (−0.42 to −0.00)*	24	*−0.30 (−0.52 to −0.07*	28	0.00	.10
Physiological need	*−0.13 (−0.27 to −0.00)*	15	−0.12 (−0.29 to 0.06)	27	0.00	.75	0.02 (−0.28 to 0.32)	24	0.04 (−0.27 to 0.35)	28	0.004	.84
Role conflict	−0.24 (−0.56 to 0.09)	15	*−0.36 (−0.51 to −0.21)*	27	0.02	.35	−0.20 (−0.41 to 0.01)	24	*−0.42 (−0.65 to −0.19)*	28	0.00	.02[d]
Care recipient need	*−0.18 (−0.34 to −0.03)*	15	*−0.33 (−0.56 to −0.11)*	27	0.005	.21	*−0.26 (−0.43 to −0.09)*	24	*−0.37 (−0.61 to −0.13)*	28	0.00	.15
Psychological need	*−0.41 (−0.66 to −0.16)*	15	−0.34 (−0.68 to 0.00)	27	0.02	.61	*−0.44 (−0.67 to −0.21)*	24	*−0.43 (−0.73 to −0.13)*	28	0.003	.91
Social support need	*−0.35 (−0.59 to −0.11)*	15	*−0.38 (−0.65 to −0.11)*	27	0.00	.72	−0.18 (−0.51 to 0.14)	24	−0.30 (−0.63 to 0.03)	28	0.00	.05
Needs (best practice)	*−0.25 (−0.36 to −0.15)*	15	*−0.36 (−0.50 to −0.22*)	27	0.006	.24	*−0.21 (−0.42 to −0.00)*	24	*−0.36 (−0.57 to −0.15)*	28	0.00	.004[e]
Resources	*0.14 (0.04 to 0.24)*	15	*0.26 (0.12 to 0.41)*	27	0.007	.10	*0.30 (0.16 to 0.44)*	24	*0.45 (0.32 to 0.58)*	28	0.004	.02[d]
Spirituality	0.01 (−0.12 to 0.15)	15	0.21 (−0.05 to 0.53)	27	0.02	.19	*0.33 (0.15 to 0.51)*	24	*0.46 (0.07 to 0.84)*	28	0.02	.42
Self-efficacy	*0.35 (0.16 to 0.54*)	15	*0.42 (0.11 to 0.73)*	27	0.008	.50	*0.42 (0.12 to 0.73)*	24	*0.49 (0.22 to 0.77)*	28	0.001	.40
Responsibility	0.03 (−0.09 to 0.16)	15	0.08 (−0.23 to 0.38)	27	0.02	.69	0.16 (−0.00 to 0.31)	24	*0.24 (0.00 to 0.48)*	28	0.006	.41
Community resources	0.07 (−0.63 to 0.76)	16	0.52 (−0.02 to 1.06)	27	0.06	.17	0.41 (−0.25 to 1.07)	26	*0.88 (0.32 to 1.44)*	28	0.11	.23
Family support	*0.19 (0.05 to 0.33)*	15	*0.22 (0.09 to 0.34)*	27	0.00	.59	*0.25 (0.05 to 0.44*)	24	*0.30 (0.17 to 0.42)*	28	0.00	.47
Closeness with care recipient	*0.18 (0.03 to 0.34)*	15	0.22 (−0.00 to 0.44)	27	0.02	.77	*0.26 (0.03 to 0.48)*	24	*0.46 (0.20 to 0.71)*	28	0.02	.16

XSL•FO
RenderX

	T2-T1						T3-T1					
	Control (n=284)		Intervention (n=281)		ICC[b]	*P* value	Control (n=284)		Intervention (n=281)		ICC	*P* value
	Estimated mean (95% CI)	Missing, n	Estimated mean (95% CI)	Missing, n			Estimated mean (95% CI)	Missing, n	Estimated mean (95% CI)	Missing, n		
Health awareness	*0.21 (0.02 to 0.39)*	15	*0.25 (0.02 to 0.48)*	27	0.00	.47	*0.32 (0.13 to 0.52)*	24	*0.48 (0.29 to 0.68*)	28	0.00	.03[d]
PACS[f]	0.16 (−0.08 to 0.40)	15	0.14 (−0.04 to 0.31)	33	0.01	.77	*0.17 (0.07 to 0.28)*	*24*	*0.26 (0.15 to 0.36)*	34	0.01	.33
Affirming	0.18 (−0.10 to 0.46)	15	0.06 (−0.15 to 0.28)	33	0.02	.34	*0.20 (0.00 to 0.40)*	24	*0.21 (0.06 to 0.36)*	34	0.02	.95
Enriching	0.15 (−0.06 to 0.35)	15	*0.23 (0.05 to 0.42)*	33	0.00	.23	*0.14 (0.03 to 0.25)*	24	*0.32 (0.16 to 0.47)*	34	0.00	.02[d]
Secondary outcomes												
GHQ-12[g]	*−0.20 (−0.38 to −0.02)*	15	*−0.13 (−0.25 to −0.01)*	33	0.02	.35	*−0.20 (−0.31 to −0.09)*	24	*−0.22 (−0.31 to −0.14)*	34	0.00	.43
EQ-5D	−0.03 (−0.10 to 0.03)	15	0.01 (−0.06 to 0.08)	33	0.06	.33	0.01 (−0.07 to 0.09)	24	0.03 (−0.03 to 0.09)	34	0.02	.57
Peace of Mind Scale	0.16 (−0.13 to 0.46)	15	0.05 (−0.11 to 0.21)	33	0.03	.38	*0.31 (0.11 to 0.51)*	24	*0.28 (0.12 to 0.43)*	34	0.00	.64

[a]Imputed data are used. These variables are controlled in the multilevel regression models. Related to caregivers: group allocation, age, sex, education level, married, with full-time job, caregiving stage; related to care recipients: age, sex, activities of daily living, instrumental activities of daily living; relationship between caregivers and recipients using spousal relationship as reference: child and others. Models accounted for clustering at the center level. Intervention effects are reported with bias-reduced linearization (CR2)–adjusted SEs and small-sample corrected degrees of freedom. ICCs quantify the variance attributable to center-level clustering.

[b]ICC: intraclass correlation.

[c]Italicized pairs of estimated means and SEs indicate significant within-group differences between T2 (3-month follow-up) and T1 (baseline assessment) or T3 (6-month follow-up) and T1 at 95% CI.

[d]*P*<.05

[e]*P*<.01

[f]PACS: Positive Aspects of Caregiving Scale.

[g]GHQ-12: 12-item General Health Questionnaire.

T2-T1 Changes

There were significant within-group changes between 3 months of follow-up (T2) and the baseline (T1). Considering changes within the intervention group, there were significant reductions in overall needs (best practice; 95% CI −0.50 to −0.22), role conflict (95% CI −0.51 to −0.21), care recipient need (95% CI −0.56 to −0.11), social support need (95% CI −0.65 to −0.11), as well as GHQ-12 (95% CI −0.25 to −0.01). There were also significant increases in overall resources (95% CI 0.12 - 0.41), specifically self-efficacy (95% CI 0.11 - 0.73), family support (95% CI 0.09 - 0.34), health awareness (95% CI 0.02 - 0.48), and enriching aspect of caregiver (95% CI 0.05 - 0.42). In the control group, there were significant decreases in overall needs (best practice; 95% CI −0.36 to −0.15), physiological need (95% CI −0.27 to −0.00), care recipient need (95% CI −0.34 to −0.03), psychological need (95% CI −0.66 to −0.16), social support need (95% CI −0.59 to −0.11), and GHQ-12 (95% CI −0.38 to −0.02), as well as increases in overall resources (95% CI 0.04 - 0.24), specifically in self-efficacy (95% CI 0.16 - 0.54), family support (95% CI 0.05 - 0.33), closeness with care recipient (95% CI 0.03 - 0.34), and health awareness (95% CI 0.02 - 0.39).

Regarding between-group differences, the multilevel regression revealed no significant intervention effects in T2-T1 changes after adjusting for covariates.

T3-T1 Changes

Considering the within-group changes between the 6-month follow-up (T3) and the baseline, there were significant decreases in overall needs (best practice; 95% CI −0.57 to −0.15) and 3 subdomains, including role conflict (95% CI −0.65 to −0.19), care recipient need (95% CI −0.61 to −0.13), and psychological need (95% CI −0.73 to −0.13). There were also significant increases in overall resources (95% CI 0.32 - 0.58) and all subdomains and positive aspects of caregiving (95% CI 0.15 - 0.36) in the intervention group. We also observed

XSL•FO
RenderX

significant decreases in care recipient (95% CI −0.43 to −0.09) and psychological needs (95% CI −0.67 to −0.21) in the control group, as well as increases in overall resources (95% CI 0.16 - 0.44) and 5 subdomains, including spirituality (95% CI 0.15 - 0.51), self-efficacy (95% CI 0.12 - 0.73), family support (95% CI 0.05 - 0.44), closeness with care recipient (95% CI 0.03 - 0.48), and health awareness (95% CI 0.13 - 0.52). Both groups exhibited increases in peace of mind (intervention: 95% CI 0.12 - 0.43; control: 95% CI 0.11 - 0.51) and a decrease in GHQ-12 (intervention: 95% CI −0.31 to −0.14; control: 95% CI −0.31 to −0.09).

Regarding the between-group differences in the change scores between T3 and T1, the improvements of the intervention group were significantly greater than the control group in overall needs (best practice; b=−0.15; SE=0.03; $t_{5.96}$=−4.54; P=.004; ICC<0.001, intervention: 95% CI −0.57 to −0.15 and control: 95% CI −0.42 to −0.00), role conflict (b=−0.22; SE=0.07; $t_{5.96}$=–3.31; P=.02; ICC<0.001, intervention: 95% CI −0.65 to −0.19 and control: 95% CI −0.41 to 0.01), overall resources (b=0.15; SE=0.05; $t_{5.99}$=2.98; P=.02; ICC=0.004, intervention: 95%CI 0.32 - 0.58 and control: 95% CI 0.16 - 0.44), health awareness (b=0.16; SE=0.06; $t_{5.96}$=2.85; P=.03; ICC<0.001, intervention: 95% CI 0.29 - 0.68 and control: 95% CI 0.13 to 0.52), and enriching aspects of caregiving (b=0.18; SE=0.06; $t_{5.96}$=3.25; P=.02; ICC<0.001, intervention: 95% CI 0.16 - 0.47 and control: 95% CI 0.03 - 0.25). Results of the multilevel regression models are shown in Table 3. The fixed effect full models of significant intervention effects are reported in Table S2 of Multimedia Appendix 1.

Moderating Factors

We tested moderation effects to examine whether the intervention's impact differed across caregiver subgroups (Table 4). There were significant negative interactions between the intervention condition and education level (b=−0.16, SE=0.06, t_{550}=−2.80, P=.005), as well as between the intervention condition and caregiver of other relations (b=−0.45, SE=0.19, t_{550}=−2.40, P=.02) on the T3-T1 needs, indicating a stronger intervention effect in reducing needs among caregivers with higher education levels and in other relationships (nonchild and nonspousal) with the care recipients (as compared to spousal caregivers).

XSL•FO
RenderX

Table . Multilevel regressions examining moderation effect of demographics on T3 (6-month follow-up) to T1 (baseline assessment) changes in primary outcomes (N=565)[a].

	Overall needs (best practice)		Overall resources		Affirming		Enriching	
	b	*P* value	*b*	*P* value	*b*	*P* value	*b*	*P* value
Intervention	−0.15	.03[b]	0.15	.049[b]	0.01	.94	0.18	.009[c]
Carer								
Age (y)	0.00	.89	0.00	.14	0.00	.41	−0.01	.23
Women	0.03	.76	−0.01	.92	0.05	.69	−0.01	.95
Education	−0.01	.87	−0.05	.06	0.01	.83	−0.03	.28
Marital status	−0.13	.19	0.10	.17	0.02	.85	−0.02	.85
Employment	−0.10	.34	0.11	.18	0.17	.12	−0.01	.93
Caregiving stage	0.05	.49	−0.14	.01[b]	−0.10	.17	−0.19	.006[c]
Care recipient								
Age (y)	0.00	.46	−0.01	.15	−0.01	.07	0.00	.48
Women	0.22	.03[b]	−0.05	.54	−0.01	.91	0.11	.27
ADL[d]	0.01	.48	−0.01	.08	0.00	.74	−0.02	.10
IADL[e]	−0.02	.16	0.00	.95	−0.03	.008[c]	−0.02	.03[b]
Child (reference spouse)	−0.15	.27	0.16	.14	0.14	.33	−0.10	.47
Other relations (reference spouse)	−0.21	.10	0.18	.07	−0.04	.75	−0.14	.25
Moderation[f]								
Carer								
Age	0.01	.32	−0.00	.29	−0.01	.33	−0.00	.64
Women	0.08	.66	0.03	.86	0.03	.88	−0.10	.61
Education	−0.16	.005[c]	−0.02	.64	−0.07	.26	−0.08	.17
Marital status	0.09	.57	0.12	.34	0.22	.18	−0.01	.97
Employment	−0.04	.81	0.18	.17	−0.02	.92	−0.10	.58
Caregiving stage	0.10	.46	0.02	.82	−0.09	.53	0.04	.77
Care recipient								
Age	0.00	.93	0.00	.90	−0.00	.69	−0.01	.48
Women	0.02	.89	−0.08	.46	−0.18	.20	0.02	.88
ADL	−0.03	.07	0.01	.68	0.02	.20	−0.02	.18
IADL	−0.01	.61	0.02	.06	0.02	.34	0.01	.47
Child (reference spouse)	−0.05	.74	−0.00	.99	−0.05	.74	0.01	.92
Other relations (reference spouse)	−0.45	.02[b]	−0.18	.20	0.07	.72	−0.23	.21

[a]Imputed data are used. Models accounted for clustering at the center level. Intervention effects are reported with bias-reduced linearization (CR2)–adjusted SEs and small sample–corrected *df*.

[b]*P*<.05.

[c]*P*<.01.

XSL•FO
RenderX

[d]ADL: activities of daily living.

[e]IADL: instrumental activities of daily living.

[f]Only the coefficients relevant to the moderators in the multilevel regressions are shown.

Discussion

Principal Findings

This study demonstrates the effectiveness of the CSM in reducing caregiver burden and enhancing resources among informal caregivers of frail older adults. The predominantly female caregiver population, many managing their own chronic health conditions, underscores the dual challenges these individuals face [35]. Balancing personal health with caregiving responsibilities highlights the necessity for integrated support systems that address both the demands of caregiving and the caregivers' personal well-being.

Intervention Effectiveness

The CSM intervention yielded more pronounced effects at 6-month (T3) follow-ups compared to 3-month (T2) follow-ups, indicating sustained and consolidating benefits over time. Caregivers reported significant improvements in overall needs, suggesting a reduced sense of burden in managing caregiving responsibilities. The reduction in role conflict suggests that caregivers became more adept at balancing caregiving tasks with personal obligations, likely supported by the CSM's structured approach, which offers practical strategies for prioritizing duties and handling competing demands. Personalized intervention plans, guided by the CNRA subscores, enabled targeted support and resource allocation to meet the needs and circumstances of each caregiver, thereby enhancing their ability to address both the practical and emotional needs of frail family members.

The intervention also led to significant increases in overall resources, especially in health awareness. This suggests that the CSM supported caregivers in recognizing and prioritizing their own health needs, which in turn enhances their ability to sustain caregiving over time. By strengthening such personal resources, the intervention contributed to building a more resilient foundation for caregivers' well-being alongside their caregiving responsibilities.

Caregivers also experienced significant enhancements in the enriching aspects of caregiving, reporting greater fulfillment and meaning in their roles. This positive shift in perspective suggests that the intervention helped them recognize the personal growth and skills gained from their caregiving responsibilities. In addition, acknowledgment of their caregiving efforts, both from the intervention team and their wider support network, may also have contributed to these feelings of enrichment.

These findings are consistent with the stress process model [11], which emphasizes the importance of reducing secondary strains such as role conflict, and of strengthening mediating resources. In our results, the intervention reduced role conflict and enhanced caregivers' sense of health awareness, suggesting improved capacity for coping and self-regulation in the caregiving role. The sustained benefits at 6 months are consistent with the stress and coping model [12,13], in which repeated cycles of appraisal (baseline assessment), coping (individualized service planning), and reappraisal (follow-up reviews) enable caregivers to adapt more effectively over time. They also resonate with strengths-based frameworks [14], which emphasize bolstering caregivers' self-efficacy and resourcefulness to manage demands more effectively. By lowering perceived burdens and expanding support systems, the CSM fostered resilience and enhanced caregivers' well-being.

Importantly, the CSM is designed as a personalized intervention in which social workers adjust the content and focus of support according to each caregiver's CNRA scores. This individualized tailoring means that the program emphasizes strengthening caregivers' overall capacity by addressing their most salient and immediate needs, rather than targeting all domains uniformly. Consequently, it is reasonable that the most consistent improvements were observed in overall needs and overall resources, while effects on specific sub-domains varied. This pattern reflects the program's emphasis on adaptive, needs-driven support rather than standardized delivery across all participants.

Moderating Factors and Tailored Interventions

The varying effectiveness of the intervention across different caregiver subgroups highlights the influence of factors such as the caregiver's relationship with the care recipient and education level. Caregivers in other relationships and those with higher education levels experienced more significant benefits from the intervention than the spousal caregivers. Nonspousal caregivers may face fewer emotional complexities, allowing them to engage more readily with practical caregiving strategies provided by the CSM. Conversely, spousal caregivers might experience greater emotional strain when caring for a partner, which can affect their ability to fully use the intervention's resources [36]. Caregivers with higher education levels might find it easier to comprehend and implement the intervention components because of better access to information and learning resources [37]. These findings suggest the need for tailored interventions to support spousal caregivers and those with lower education levels more effectively. Spousal caregivers might benefit from additional emotional support and relational counseling to address the unique challenges they face [38]. Simplified educational materials and training programs could enhance accessibility for caregivers with lower educational backgrounds, ensuring they can fully engage with the intervention. Reflecting ongoing input from frontline practitioners and service users, some refinements have also been made to the wording of CNRA items to improve clarity and accessibility (ie, the 3-item closeness with the care recipient subscale; Leung et al, unpublished data), particularly for caregivers with lower literacy levels.

Practical Implications

The study's findings have significant practical implications for the design and implementation of caregiver support programs. Interventions should be customized to address the specific

XSL•FO
RenderX

challenges faced by diverse caregiver groups, incorporating stress management techniques, health monitoring, and financial planning assistance to alleviate the compounded burdens of caregiving and personal health management [39]. Developing educational resources tailored to varying literacy levels can empower all caregivers with the necessary knowledge and skills, enhancing the overall effectiveness of support programs [40]. The need for sustained improvement among caregivers highlights the importance of long-term, adaptive interventions. Continuous training, regular reassessment of needs, and personalized care plans can help maintain and enhance the benefits of the intervention over time.

Moreover, the findings underscore the importance of policy initiatives and community involvement in supporting caregivers. Policymakers should consider allocating resources to develop and maintain evidence-based caregiver support programs like the CSM, recognizing the critical role caregivers play in the health system. Community engagement is essential for raising awareness about the challenges caregivers face and mobilizing support networks, which can provide additional layers of assistance and relief [41].

Limitations and Future Research

Despite the positive outcomes, the study has several limitations that should be considered. First, the study was conducted during the COVID-19 pandemic, which necessitated adaptations in service delivery, shifting from face-to-face interactions to virtual platforms, which may have influenced the outcomes and limited the generalizability of the findings to nonpandemic contexts. Future research should examine the long-term effectiveness of the CSM in an after-pandemic context to validate these findings. Second, the sample predominantly consisted of female caregivers recruited from 8 centers associated with 4 NGOs in Hong Kong, which may limit the applicability of the findings to male caregivers or those in different regions. Third, the reliance on self-reported measures introduces the possibility of response biases, such as social desirability or recall bias. Fourth, the follow-up period was limited to 6 months, leaving the long-term durability of the intervention's effects uncertain. Future research should extend follow-up periods and incorporate objective measures to validate and expand upon these findings. Fifth, although moderation effects were examined using interaction terms in separate multilevel models, our sample size (N=565) may not have been sufficient to detect small interaction effects, which often require much larger samples (eg, 800 - 1000) to achieve adequate power [42]. Nonsignificant moderation findings should therefore be interpreted with caution. Sixth, while the intervention produced statistically significant improvements, it remains unclear whether these changes meet clinical or practical thresholds. Minimal clinically important differences for most caregiver-reported outcomes, except GHQ-12, have not been established. Referencing the 11/12 threshold used in a local study using the Likert scoring method [43], around 73% (205/281) of caregivers in the intervention group exceeded this cutoff at baseline, indicating high distress levels in the caregiving role, while the percentage fell to 64% (180/281) at T3, suggesting potential benefits of the intervention in alleviating psychological distress, although the difference in the change of GHQ-12 scores between the intervention and control groups is not significant. Future research should define minimal clinically important differences for other key outcomes to improve the interpretation of intervention effects [1]. Finally, future research should also evaluate whether the CSM can reduce frail older adults' health care use, such as hospital admissions or emergency visits, and overall system costs, thereby demonstrating their broader value to health and social care systems.

Conclusions

The CSM offers an effective, structured approach for reducing caregiver burden and enhancing resources among informal caregivers of frail older adults. By offering personalized support and practical tools, the CSM addresses the multifaceted challenges caregivers face. The findings of this study highlight the necessity for long-term adaptable interventions that can evolve with caregivers' changing needs. Implementing such models can significantly improve caregivers' well-being and contribute to a more resilient caregiving system, which is increasingly important in the context of an aging global population.

Acknowledgments

This work was supported by the Simon K.Y. Lee Foundation and the Research Matching Grant Scheme of the University Grants Committee, Hong Kong SAR. Open Access made possible with partial support from the Open Access Publishing Fund of the City University of Hong Kong.

Data Availability

The data supporting this study's findings are available on request from the corresponding authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Detailed intervention protocol of caregiver support model and supplementary results.
[DOCX File, 516 KB - aging_v8i1e71638_app1.docx]

XSL•FO
RenderX

Checklist 1
CONSORT-eHEALTH checklist.
[PDF File, 1131 KB - aging_v8i1e71638_app2.pdf]

References

1. Hong Kong population projections 2020-2069. Census and Statistics Department. 2020. URL: https://www.censtatd.gov.hk/en/data/stat_report/product/B1120015/att/B1120015082020XXXXB0100.pdf [accessed 2025-10-24]
2. Fertility trend in Hong Kong. Census and Statistics Department. 1991. URL: https://www.censtatd.gov.hk/en/data/stat_report/product/FA100090/att/B72302FA2023XXXXB0100.pdf [accessed 2025-10-24]
3. Ringer TJ, Hazzan AA, Kennedy CC, et al. Care recipients' physical frailty is independently associated with subjective burden in informal caregivers in the community setting: a cross-sectional study. BMC Geriatr 2016 Nov 17;16(1):186. [doi: 10.1186/s12877-016-0355-6] [Medline: 27855633]
4. He AJ, Tang VFY. Integration of health services for the elderly in Asia: a scoping review of Hong Kong, Singapore, Malaysia, Indonesia. Health Policy 2021 Mar;125(3):351-362. [doi: 10.1016/j.healthpol.2020.12.020] [Medline: 33422336]
5. Leung VWY, Lam CM, Liang Y. Parents' expectations of familial elder care under the Neoliberal Hong Kong Society. J Fam Issues 2020 Apr;41(4):437-459. [doi: 10.1177/0192513X19879189]
6. Woo J, Mak B, Cheng JOY, Choy E. Identifying service needs from the users and service providers' perspective: a focus group study of Chinese elders, health and social care professionals. J Clin Nurs 2011 Dec;20(23-24):3463-3471. [doi: 10.1111/j.1365-2702.2011.03702.x] [Medline: 22074148]
7. Carer support. Social Welfare Department. 2023. URL: https://www.swd.gov.hk/en/pubsvc/elderly/cat_careersupp/carersuppo/ [accessed 2024-05-23]
8. Tsang WHH, Pang WYC, Lai YYC, Chung CK. Consultancy study on needs and support required of carers of elderly persons and of persons with disabilities in Hong Kong. The Hong Kong Polytechnic University. 2022. URL: https://research.polyu.edu.hk/en/activities/consultancy-study-on-needs-and-support-required-of-carers-of-elde-2/ [accessed 2025-10-24]
9. Lee MMS, Yeoh EK, Wong ELY, et al. Perceptions and experiences of caregiver-employees, employers, and health care professionals with caregiver-friendly workplace policy in Hong Kong: thematic analysis. Interact J Med Res 2025 Feb 10;14(1):e58528. [doi: 10.2196/58528] [Medline: 39928924]
10. Li KK, Leung CLK, Yeung D, et al. Development and validation of the caregiver needs and resources assessment. Front Psychol 2022;14. [doi: 10.3389/fpsyg.2023.1063440]
11. Pearlin LI, Mullan JT, Semple SJ, Skaff MM. Caregiving and the stress process: an overview of concepts and their measures. Gerontologist 1990 Oct;30(5):583-594. [doi: 10.1093/geront/30.5.583] [Medline: 2276631]
12. Haley WE, Levine EG, Brown SL, Bartolucci AA. Stress, appraisal, coping, and social support as predictors of adaptational outcome among dementia caregivers. Psychol Aging 1987;2(4):323-330. [doi: 10.1037/0882-7974.2.4.323]
13. Lazarus RS, Stress FS. Stress, Appraisal, and Coping: Springer Publishing Company; 1984.
14. Berg-Weger M, Rubio DM, Tebb SS. Strengths-based practice with family caregivers of the chronically ill: qualitative insights. Fam Soc 2001 Jun;82(3):263-272. [doi: 10.1606/1044-3894.191]
15. Saleebey D. The strengths perspective in social work practice: extensions and cautions. Soc Work 1996 May;41(3):296-305. [Medline: 8936085]
16. Smith AL. The stages of caregiving. Am J Pastoral Counsel 2001 Aug;4(3):19-40. [doi: 10.1300/J062v04n03_03]
17. Given BA, Given CW, Sherwood PR. Family and caregiver needs over the course of the cancer trajectory. J Support Oncol 2012;10(2):57-64. [doi: 10.1016/j.suponc.2011.10.003] [Medline: 22222251]
18. Lou VW. Holistic carer support: a care management approach. Presented at: 14th HKEC Symposium on Community Engagement: Medical Social Collaboration—We Build We Care; Jun 29, 2019; Hong Kong URL: http://hub.hku.hk/handle/10722/298607 [accessed 2025-10-24]
19. Zarit SH, Reever KE, Bach-Peterson J. Relatives of the impaired elderly: correlates of feelings of burden. Gerontologist 1980 Dec;20(6):649-655. [doi: 10.1093/geront/20.6.649] [Medline: 7203086]
20. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [doi: 10.2307/2136404] [Medline: 6668417]
21. Lovibond PF, Lovibond SH. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behav Res Ther 1995 Mar;33(3):335-343. [doi: 10.1016/0005-7967(94)00075-u] [Medline: 7726811]
22. Guo YL, Liu YJ. Family functioning and depression in primary caregivers of stroke patients in China. Int J Nurs Sci 2015 Jun;2(2):184-189. [doi: 10.1016/j.ijnss.2015.05.002]
23. Iecovich E. Quality of relationships between care recipients and their primary caregivers and its effect on caregivers' burden and satisfaction in Israel. J Gerontol Soc Work 2011;54(6):570-591. [doi: 10.1080/01634372.2011.579691] [Medline: 21780882]
24. Liu HY, Huang LH. The relationship between family functioning and caregiving appraisal of dementia family caregivers: caregiving self-efficacy as a mediator. Aging Ment Health 2018 Apr;22(4):558-567. [doi: 10.1080/13607863.2016.1269148] [Medline: 28001431]

XSL•FO
RenderX

25. Chan CLF, Chui EWT. Association between cultural factors and the caregiving burden for Chinese spousal caregivers of frail elderly in Hong Kong. Aging Ment Health 2011 May 1;15(4):500-509. [doi: 10.1080/13607863.2010.536139]
26. Chow EOW, Ho HCY. The relationship between psychological resources, social resources, and depression: results from older spousal caregivers in Hong Kong. Aging Ment Health 2012 Nov;16(8):1016-1027. [doi: 10.1080/13607863.2012.692769]
27. Dellemain J, Warburton J. Case management in rural Australia: arguments for improved practice understandings. Aust Soc Work 2013 Jun;66(2):297-310. [doi: 10.1080/0312407X.2012.675345]
28. Xun K. Seeing flowers through the mist: how social workers in disability services retrospectively view their educational experiences related to case management in Hong Kong. Chin J Soc Work 2019 Jan 2;12(1):39-55. [doi: 10.1080/17525098.2019.1558741]
29. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736. [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
30. Wong ELY, Ramos-Goñi JM, Cheung AWL, Wong AYK, Rivero-Arias O. Assessing the use of a feedback module to model EQ-5D-5L health states values in Hong Kong. Patient 2018 Apr;11(2):235-247. [doi: 10.1007/s40271-017-0278-0] [Medline: 29019161]
31. Goldberg D, Williams P. A User's Guide to the General Health Questionnaire: NFER-Nelson; 1988. URL: https://cir.nii.ac.jp/crid/1130000795003754752 [accessed 2025-10-24]
32. Chan DW, Chan TSC. Reliability, validity and the structure of the General Health Questionnaire in a Chinese context. Psychol Med 1983 May;13(2):363-371. [doi: 10.1017/s0033291700050996] [Medline: 6878523]
33. Lee YC, Lin YC, Huang CL, Fredrickson BL. The construct and measurement of Peace of Mind. J Happiness Stud 2013 Apr;14(2):571-590. [doi: 10.1007/s10902-012-9343-5]
34. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. Stat Med 2011 Feb 20;30(4):377-399. [doi: 10.1002/sim.4067] [Medline: 21225900]
35. Lam WWY, Nielsen K, Sprigg CA, Kelly CM. The demands and resources of working informal caregivers of older people: a systematic review. Work Stress 2022 Jan 2;36(1):105-127 [FREE Full text] [doi: 10.1080/02678373.2022.2028317]
36. Schulz R, Sherwood PR. Physical and mental health effects of family caregiving. Am J Nurs 2008 Sep;108(9 Suppl):23-27. [doi: 10.1097/01.NAJ.0000336406.45248.4c] [Medline: 18797217]
37. Croog SH, Burleson JA, Sudilovsky A, Baume RM. Spouse caregivers of Alzheimer patients: problem responses to caregiver burden. Aging Ment Health 2006 Mar;10(2):87-100. [doi: 10.1080/13607860500492498] [Medline: 16517483]
38. Griffiths AW, Shoesmith E, Sass C, Nicholson P, Charura D. Relational counselling as a psychosocial intervention for dementia: qualitative evidence from people living with dementia and family members. Dementia (London) 2021 Aug;20(6):2091-2108. [doi: 10.1177/1471301220984912] [Medline: 33382000]
39. Wawrziczny E, Larochette C, Papo D, et al. A customized intervention for dementia caregivers: a quasi-experimental design. J Aging Health 2019 Aug;31(7):1172-1195. [doi: 10.1177/0898264318770056] [Medline: 29665714]
40. Jensen M, Agbata IN, Canavan M, McCarthy G. Effectiveness of educational interventions for informal caregivers of individuals with dementia residing in the community: systematic review and meta-analysis of randomised controlled trials. Int J Geriatr Psychiatry 2015 Feb;30(2):130-143. [doi: 10.1002/gps.4208] [Medline: 25354132]
41. Papastavrou E, Andreou P, Middleton N, Tsangari H, Papacostas S. Dementia caregiver burden association with community participation aspect of social capital. J Adv Nurs 2015 Dec;71(12):2898-2910. [doi: 10.1111/jan.12762] [Medline: 26345604]
42. McClelland GH, Judd CM. Statistical difficulties of detecting interactions and moderator effects. Psychol Bull 1993 Sep;114(2):376-390. [doi: 10.1037/0033-2909.114.2.376] [Medline: 8416037]
43. Cheung LM, Wong WS. The effects of insomnia and internet addiction on depression in Hong Kong Chinese adolescents: an exploratory cross-sectional analysis. J Sleep Res 2011 Jun;20(2):311-317. [doi: 10.1111/j.1365-2869.2010.00883.x] [Medline: 20819144]

Abbreviations

CNRA: Caregiver Needs and Resources Assessment
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
CSM: caregiver support model
GHQ-12: 12-item General Health Questionnaire
ICC: intraclass correlation coefficient
NGO: nongovernmental organization
PACS: Positive Aspects of Caregiving Scale
PoMS: Peace of Mind Scale
ZBI-12: 12-item Zarit Burden Interview

XSL•FO
RenderX

Edited by F Puga; submitted 23.01.25; peer-reviewed by C Maxwell, L Parlett; revised version received 05.10.25; accepted 06.10.25; published 03.11.25.

Please cite as:
Leung CLK, Li KK, Yeung DY, Chong AML, Chiu MYL, Guan X, Lo TW
The Caregiver Support Model for Informal Caregivers of Frail Older Adults: Randomized Controlled Trial
JMIR Aging 2025;8:e71638
URL: https://aging.jmir.org/2025/1/e71638
doi:10.2196/71638

© Cyrus Lap Kwan Leung, Kin-Kit Li, Dannii Y Yeung, Alice Ming Lin Chong, Marcus Yu Lung Chiu, Xin Guan, Tit Wing Lo. Originally published in JMIR Aging (https://aging.jmir.org), 3.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Supporting Dementia Caregiving With a Mobile Care Ecosystem: Development and Mixed Methods Study

Chetna Malhotra[1,2], MD; Yanzhen Yue[3], PhD; Chandrika Ramakrishnan[1], MPH; Shimoni Shah[1], MPH; Wenda Chen[3], PhD; Philip Yap[4], MRCP; Chin Yee Cheong[4], MRCP; Irene Teo[1], PhD; Shiou-Liang Wee[5], PhD; Xiangming Lan[3], MTech; Yi Chen[3], BComp; Chee Seng Chong[3], ME; Xueying Huang[1], BS; Ivy Chua[1], MA

[1]Lien Centre for Palliative Care, Duke NUS Medical School, 8 College Road, Singapore, Singapore

[2]Programme in Health Services Research and Population Health, Duke NUS Medical School, 8 College Road, Singapore, Singapore

[3]Institute of High Performance Computing, Agency for Science, Technology and Research, Singapore, Singapore

[4]Department of Geriatric Medicine, Khoo Teck Puat Hospital, Singapore, Singapore

[5]SR Nathan School of Human Development, Singapore University of Social Sciences, Singapore, Singapore

Corresponding Author:
Chetna Malhotra, MD
Lien Centre for Palliative Care, Duke NUS Medical School, 8 College Road, Singapore, Singapore

Abstract

Background: Dementia presents substantial challenges for informal caregivers. A gap remains in technology-driven personalized support tailored to caregivers' needs.

Objective: This study aimed to develop a theory-driven, multicomponent mobile app specifically designed for caregivers of individuals with dementia and test its usability among end users.

Methods: We developed *CareBuddy*, a mobile care ecosystem based on the stress process model and user-centered design. The app includes personalized assessments and tailored solutions, an artificial intelligence–driven chatbot, GPS-based location monitoring, peer support, a helpline, telemedicine, health care provider integration, and caregiver self-care resources. Development was informed by interviews with caregivers and stakeholders, followed by a 2-phase pilot test involving 18 and 10 participants, respectively, to assess usability and acceptability.

Results: In phase 1, the mean system usability scores increased from 65.4 (SD 11.8) in round 1 to 73.8 (SD 15.9) in round 3, exceeding the benchmark of 68. In phase 2, caregivers rated the app highly, with an overall mean score of 95.4 (SD 8.5) on the Mobile Health App Usability Questionnaire. The domains of ease of use (mean 24.1, SD 2.9), user interface and satisfaction (mean 40.3, SD 3.4), and usefulness (mean 31, SD 3.9) received high Mobile Health App Usability Questionnaire ratings. Participants valued the content focused on dementia management and caregiver well-being. Caregivers appreciated the interactive features: social networking portal, service directory, and conversational large language model. Feedback highlighted areas for improvement, including reducing textual overload and addressing navigational challenges.

Conclusions: *CareBuddy* offers a multifaceted digital solution for dementia caregivers, with high usability and satisfaction. An ongoing trial is evaluating the app's effectiveness in improving caregiver outcomes.

(JMIR Aging 2025;8:e78759) doi:10.2196/78759

KEYWORDS

dementia; caregiving; user-centered design; multicomponent; mobile app; usability; acceptability; artificial intelligence chatbot; AI chatbot

Introduction

Dementia is an increasingly pressing public health issue, driven not only by the rising number of diagnoses in aging populations [1] but also by the substantial burden it places on informal caregivers. Informal caregivers, typically family members or close friends, often dedicate an average of 5 hours per day to caregiving tasks, which range from personal care and managing daily routines to ensuring the safety and well-being of individuals living with dementia [2-4]. Informal caregiving is so integral to dementia care that it accounts for more than half of the total costs associated with dementia care [5], underscoring the critical role these caregivers play within the health care system. However, this immense responsibility comes at a high cost to caregivers, affecting their physical, psychological, and emotional well-being and contributing to loss of labor market productivity [6-12].

XSL•FO
RenderX

To address these challenges, caregivers require multifaceted support, including timely and accurate information, emotional and social support, and access to health care providers [13-15]. Meeting these diverse needs requires comprehensive, multicomponent interventions that can provide not only sustained, holistic care but also immediate assistance when required.

Technology-based platforms have emerged as a promising avenue for supporting dementia caregivers. Systematic reviews of technology-driven interventions have shown that caregivers highly value certain key elements, such as peer support, direct contact with health care providers, and information tailored to their specific needs, rather than generalized content [16-20]. Psychological support is also considered crucial in these interventions [21]. These features are most effective when integrated as multicomponent interventions.

Mobile health apps offer an accessible and convenient platform. However, 2 scoping reviews found that most existing apps, including those commercially available, focused on monitoring, tracking, task management, mental health support, health education, and caregiver communication but fell short of meeting caregivers' needs [22,23]. A systematic review identified only 6 dementia caregiver apps that have undergone research evaluation, most of which provided task-based support [24]. Current apps thus remain limited, often lacking high-quality tailored content, interactivity, personalization, and theoretical grounding in their development [25-27].

A mobile app that is multicomponent, theory driven, and personalized to caregivers' unique needs could provide both sustained support and real-time assistance and serve as a one-stop solution for caregivers, creating an integrated caregiving ecosystem. The effectiveness of such an app would depend on several critical factors, including the app's design, user-friendliness, and the quality of support it offers.

We therefore developed *CareBuddy*, a theory-driven, multicomponent mobile app specifically designed for caregivers of individuals with dementia. The app focuses on enhancing sustained and ongoing support, ensuring continuity of care, and fostering a sustainable caregiving environment. This paper aims to describe the development process of the "*CareBuddy*" app and its multicomponent features and to present the results for its usability and acceptability testing among end users.

Methods

Overview

We used user-centered design principles and iterative processes to develop the *CareBuddy* app [28,29]. We reported the study in accordance with the COREQ (Consolidated Criteria for Reporting of Qualitative Studies) [30] (Checklist 1).

Ethical Considerations

The study was approved by the National University of Singapore Institutional Review Board (NUS-IRB-2022 - 309). All participants signed a written informed consent before participation. We assigned subject codes for participants, and the data were deidentified to ensure anonymity of responses. Participants received S$30 to S$70 (S$1=US $0.77) depending on the phase of the study.

Development Process

Assessment of User Requirements

We conducted qualitative in-depth interviews with family caregivers of people with dementia, policymakers involved in planning dementia programs, clinicians, and community dementia service providers to identify desired features for a mobile app designed for caregivers of individuals with dementia. Using purposive sampling, we recruited family caregivers through referrals from community dementia organizations and an ongoing study, whereas policymakers, clinicians, and service providers were approached through key contacts in relevant organizations. A trained qualitative researcher conducted interviews, which were audio recorded, transcribed, and analyzed on NVivo 11 (Lumivero) through content analysis.

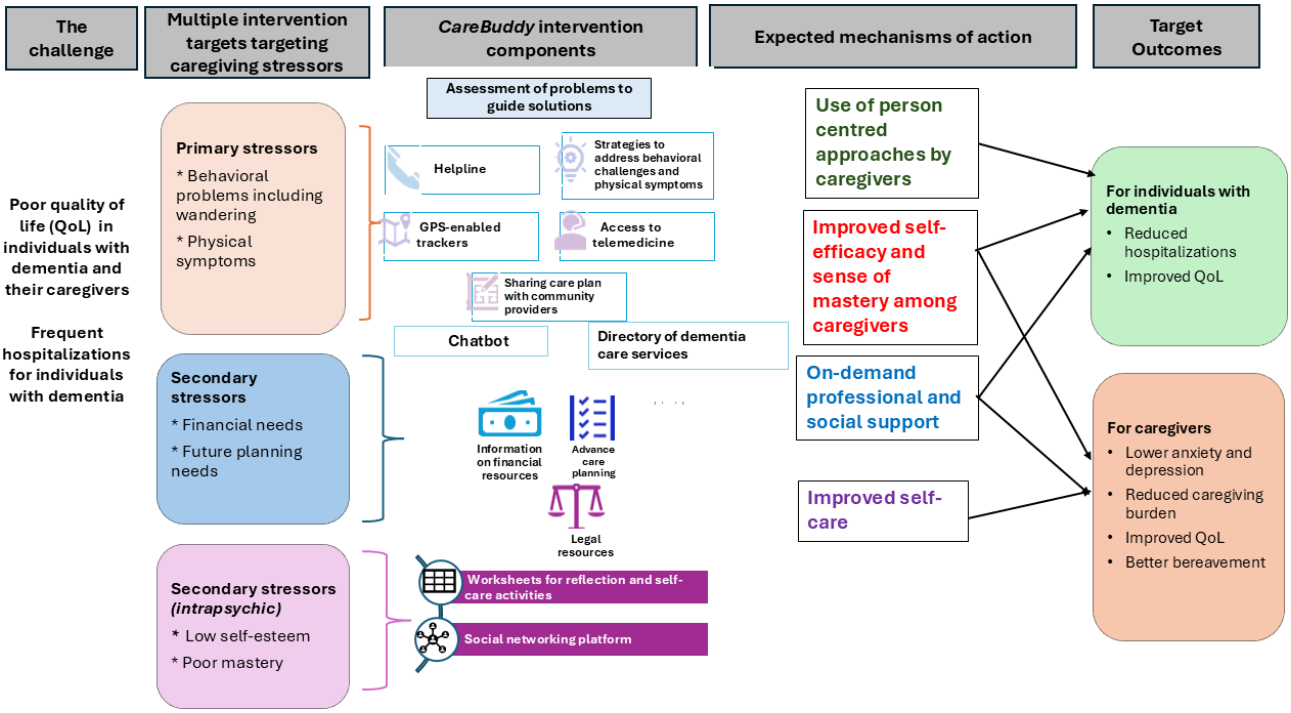
Two informal caregivers of individuals with dementia guided the research team on the design and content of the *CareBuddy* app. Study team members met with them frequently to elicit feedback.

App Features

Guided by the stress process model [31], we developed a logic model to develop the "*CareBuddy*" app (Figure 1). The stress process model, an established framework, describes the relationship between caregiving stressors, psychosocial resources, and caregivers' psychosocial well-being [31,32]. The framework thus provides a foundation to develop an integrated caregiving ecosystem that addresses caregiving stressors and delivers comprehensive support by enhancing caregivers' access to tailored resources, coping strategies, and social support networks.

XSL•FO
RenderX

Figure 1. Logic model for development of the *CareBuddy* mobile care ecosystem.



Pilot Testing

Overview

We conducted pilot testing in 2 phases using a mixed-method design incorporating both qualitative feedback and quantitative assessments for app usability.

Phase 1

In this phase, we aimed to gather feedback on the app's content and functionality. Participants interacted with the prototype version of the app using the interviewer's smartphone. The interviewer then guided the participant through the app's various components using a topic guide to prompt open-ended discussions. This phase involved 3 rounds of in-depth interviews, with the app iteratively updated after each round based on participants' feedback. The interview guide, developed by the study team, explored key aspects of user experience, app functionality, the relevance of the app's solutions for managing behavioral and physical symptoms of individuals with dementia, perceived usefulness, and suggestions for improvement. We also collected demographic data and assessed usability using the 10-item System Usability Scale (SUS) [33].

We purposively sampled family caregivers and health care providers. Eligible caregivers were Singaporean or permanent residents aged ≥21 years, involved in making treatment decisions for a family member with dementia, had been providing care for at least 6 months and were planning to do so for the next 2 months, able to read and write English, and had intact cognition as determined through the Abbreviated Mental Test (only for those aged ≥65 y). Health care providers included those providing health or social care within community organizations or hospitals. The interviews were conducted in English by a trained researcher with no prior relationship with the participants. Sessions were held at a time and place convenient for the participants, lasting 1 to 2 hours.

Phase 2

This phase assessed caregivers' app engagement, usability dimensions including satisfaction and acceptability, and subjective experiences. Caregivers downloaded the app (iOS or Android version) onto their own mobile phones, with a researcher providing an orientation on its features and navigation. If the care recipient was mobile, caregivers additionally received a GPS tracker for tracking their location.

Eligibility criteria for family caregivers remained the same as in phase 1, with the additional requirement of owning an Android or iOS mobile phone. Caregivers of individuals residing in long-term care facilities were excluded.

Following informed consent, we collected demographic data for the caregivers (eg, age, gender, relationship, and education) and the individual with dementia (eg, age and gender) through a survey. At the end of the 1-month trial, app usage was tracked throughout the month, including average engagement time per week, time spent on each component, and frequently used components. We administered the Mobile Health App Usability Questionnaire (MAUQ), English stand-alone patient version, for the intended audience, family caregivers of individuals with dementia [34]. MAUQ, a validated 18-item questionnaire, assesses usability in 3 dimensions (ie, ease of use, interface and satisfaction, and usefulness) on a 7-point Likert scale, ranging from 1 (strongly disagree) to 7 (strongly agree) [34]. Caregivers also participated in open-ended interviews exploring their experiences and subjective feedback on the relevance and perceived usefulness of the app. The open-ended portion was audio recorded and transcribed. Interviews lasted 10 to 25 minutes.

XSL•FO
RenderX

Analysis

Quantitative

We summarized demographics and calculated SUS scores both overall and by interview round to assess the app's usability evolution. For the phase 1 pilot, we compared the SUS scores to the benchmark mean of 68, which indicates good usability [35,36]. In the phase 2 pilot, we calculated total and average scores assessing MAUQ usability dimensions, that is, ease of use, interface and satisfaction, and usefulness, including satisfaction and acceptability, with higher total and average scores indicating greater usability [34]. An average score lower than 4 meant poor usability [37]. We also summarized app usage data in terms of the number of hours used, components used, and time spent on each component in phase 2.

Qualitative

Using NVivo 11 software, 2 team members independently coded transcripts and conducted content analysis [38]. We developed a code book based on the User Version of the Mobile Application Rating Scale [39], categorizing feedback into key domains: engagement, functionality, aesthetics, information quality and quantity, and subjective feedback. We synthesized user feedback within these domains and incorporated suggestions for refining the app.

We triangulated quantitative and qualitative findings on usability and engagement.

Results

Development Process

Assessment of User Requirements

Participant Demographics

We interviewed 22 participants (ie, 11 caregivers, 6 community dementia providers, 3 clinicians, and 2 policymakers), with a mean age of 49 (SD 10.4) years, 73% (n=16) were women, and half (n=11, 50%) were caregivers of individuals with dementia (details in Table 1).

We grouped participants' desired features into three categories: education, communication, and listing/directory.

XSL•FO
RenderX

Table . Participant characteristics.

Item	Assessment of user requirements		Pilot: phase 1		Pilot: phase 2, caregivers (n=10)
	Caregivers (n=11)	Noncaregivers (n=11)	Caregivers (n=11)	Health care providers (n=7)	
Age (y), mean (SD)	53.7 (9.6)	44 (8.9)	58.5 (9.5)	38.9 (11)	58.5 (9.5)
Gender, n (%)					
Male	4 (36.4)	2 (18.2)	4 (36.4)	4 (57.1)	2 (20)
Female	7 (63.6)	9 (81.8)	7 (63.6)	3 (42.9)	8 (80)
Ethnicity, n (%)					
Chinese	9 (81.8)	—[a]	11 (100)	5 (71.4)	9 (90)
Non-Chinese	2 (18.2)	—	0 (0)	2 (28.6)	1 (10)
Primary language, n (%)					
English	—	—	9 (81.8)	7 (100)	—
Chinese	—	—	2 (18.2)	0 (0)	—
Marital status, n (%)					
Married	—	—	10 (90.9)	—	5 (50)
Never married	—	—	1 (9.1)	—	3 (30)
Divorced	—	—	0 (0)	—	2 (20)
Widowed	—	—	0 (0)	—	0 (0)
Education, n (%)					
No formal education	0 (0)	—	0 (0)	—	0 (0)
Secondary	3 (27.3)	—	3 (27.3)	—	1 (10)
Junior college/polytechnic/diploma	3 (27.3)	—	3 (27.3)	—	3 (30)
University and above	5 (45.5)	—	5 (45.5)	—	6 (60)
Relationship with person with dementia, n (%)					
Spouse	1 (9.1)	—	5 (45.5)	—	4 (40)
Children	9 (81.8)	—	5 (45.5)	—	2 (20)
Sibling	0 (0)	—	1 (9.1)	—	1 (10)
Others	1 (9.1)	—	0 (0)	—	3 (30)

[a]Not applicable because data not collected.

Education

Participants desired short videos for caregivers focusing on daily living skills and management of troublesome behaviors among individuals with dementia. They also expressed interest in newsletters, articles, and support group information. Additional requests included information on financial and legal considerations, self-care strategies to address caregiver stress, and self-assessment tools to personalize support. A guide on how to navigate the app was also suggested.

Listing/Directory

Participants wanted directories for various services, including daycare centers, respite care services, dementia-specific community services, and telemedicine providers.

Summary of desired app features with illustrative quotes is detailed in Table S1 in Multimedia Appendix 1.

App Features

We describe the *CareBuddy* app's features under three aspects: app components, user interface design, and system features.

App Components

Overview

On the basis of the user requirements and the stress process framework–based logic model, we developed app components to address both primary and secondary stressors (Figure 1).

Addressing Primary Stressors of Dementia Caregiving

The app targets primary caregiving stressors such as behavioral symptoms (eg, aggression, agitation, mood swings,

XSL•FO
RenderX

hallucinations, disorientation, and refusal of care) and physical challenges (eg, mobility issues, eating difficulties, dry skin, toileting problems, communication difficulties, and sleep disturbances). Upon logging in, caregivers are guided through an assessment using a series of questions to identify current behavioral and physical symptoms, as well as their potential underlying causes. On this input, the app offers tailored, practical management strategies. Specifically, for the care recipient's behaviors, the approach uses the unmet needs model (which views behaviors as expressions of unarticulated needs), the progressively lowered stress threshold model (which emphasizes the reduced capacity of the care recipient to deal with stress or stimuli), and the ABC (antecedent, behavior, and consequences) models. The solutions are further derived using the DICE (describe, investigate, create, evaluate) and P.I.E.C.E.S. (Physical. Intellectual. Emotional. Capabilities. Environment. Social.) frameworks [40,41]. The content was developed in consultation with dementia care providers and from existing resources published globally. All content was reviewed and validated by 2 geriatricians on the study team with expertise in dementia care.

A particularly difficult behavioral symptom, wandering, is addressed through a GPS location monitoring function integrated into the app. This feature allows caregivers to monitor the real-time location of their care recipients using GPS-enabled trackers. Caregivers can view the last detected location of their care recipients directly through Google Maps. Additionally, the app provides step-by-step navigation to the tracked location. The app monitors the GPS tracker's battery level and provides reminders to recharge the battery.

To address primary stressors, *CareBuddy* also facilitates collaborative care management through a care plan function, which supports information sharing between family caregivers and health care professionals. The Care Plan includes a section on the individual's hobbies and interests to promote person-centered care. Caregivers can also use a centralized messaging tool to share updates and discuss care plans with providers. The app also allows caregivers to apply for care center transfers, with approvals managed by health care providers. A secure QR code–based system enables multiple caregivers (eg, family members or friends) to coordinate care efficiently. By streamlining communication and care planning, *CareBuddy* strengthens coordinated, person-centered care.

To ensure timely access to health services, *CareBuddy* integrates telemedicine access via deep links to third-party provider apps. Users can initiate virtual consultations with established general practitioners directly through *CareBuddy*, eliminating the need to search for separate telemedicine apps. This integration enhances continuity within *CareBuddy's* digital ecosystem. Instructional content is also provided to guide caregivers on using these services effectively.

In moments of acute distress, caregivers can also access a crisis support helpline within the app, connecting them directly to trained social workers who provide emotional support, practical guidance, and crisis intervention.

A large language model (LLM)–powered chatbot is integrated into the app to provide personalized, real-time responses to caregivers' queries by offering advice, summarizing key strategies, and encouraging engagement with relevant app components. This chatbot uses GPT-4 as the default language model for generating responses and GPT-3.5 for contextual analysis based on the knowledge database. The system is built on a retrieval-augmented generation architecture, where relevant knowledge items are retrieved from the database and used as context in prompts to the language model. This enables the chatbot to generate responses that are both contextually accurate and conversationally coherent. The chatbot integrates both an internal knowledge repository and external knowledge sources to generate accurate, contextually relevant responses. The internal repository comprises a structured question-and-answer (Q&A) database, curated by the study team to address key caregiving stressors. To enhance response accuracy, we implemented natural language processing to semantically map user queries to the most relevant Q&A entries, ensuring precise information retrieval. A tiered, confidence-based retrieval system further optimizes response generation, with confidence scores derived from cosine similarity between user queries and knowledge-based questions. Queries scoring above 95% are directly matched to Q&A entries with LLMs adding empathetic context, scores between 90% and 95% synthesize information across multiple sources, and scores below 90% engage ChatGPT with domain-specific context for broader, tailored responses. These thresholds were empirically determined through user testing to balance accuracy for specific queries with generalization for broader or ambiguous ones. Testing confirmed that high-confidence matches yield precise answers, whereas lower-confidence matches generate broader, contextually relevant summaries. To ensure safety, the chatbot responds strictly within its knowledge scope at high confidence and, in urgent or sensitive cases, directs users to immediate resources such as helplines rather than attempting diagnostic or therapeutic advice. Personal information is coded, protected, and delinked from individual users, ensuring all chatbot conversations are completely anonymous.

Addressing Secondary Stressors of Dementia Caregiving

To address secondary stressors such as financial burden and future care planning, the app offers curated content on financial resources and subsidies, along with guidance on how to do advance care planning, power of attorney, making a will, and Central Provident Fund nominations. Caregivers also receive information on the anticipatory trajectory of illness for dementia so that they can plan for the evolving needs of their care recipient as the condition progresses.

To reduce internal caregiver stress and improve emotional well-being, *CareBuddy* includes a dedicated self-care module featuring psychologist-developed worksheets. These resources help caregivers reflect on their caregiving identity and journey, recognize signs of burnout, engage in self-care, and build coping strategies for anticipatory grief and emotional resilience.

The app also provides access to the user's phone calendar, which allows caregivers to create personalized event reminders. This function helps users keep track of important dates, appointments, or caregiving tasks and provides timely alerts to ensure they receive the necessary prompts to take action. This feature

XSL•FO
RenderX

responds directly to caregiver feedback and is designed to support daily organization.

Finally, *CareBuddy* facilitates peer support through an in-app online forum, available both on the app and via a web portal. This space allows caregivers to seek help, share experiences, offer support, get connected with other caregivers, and build community. Users can flag inappropriate posts, block others if needed, and benefit from a backend "Censor Curse Words" function that ensures respectful communication. An LLM-based artificial intelligence (AI) bot named "Sparky" is available in the online forum to stimulate user interactions and facilitate engagement with caregivers. The AI bot "Sparky" generates initial replies to caregivers' new posts and provides comments when requested by caregivers.

User Interface Design

Caregiver Interface

The *CareBuddy* app interface is intentionally designed to be user-friendly, with frequently used functions placed in prominent positions and enhanced with high-resolution icons. A text-to-speech feature allows users to listen to information instead of reading, supporting accessibility. Onboarding instructions and reminders are integrated to support user navigation and regular engagement.

Health Care Provider Interface

The *CareBuddy* caregiver app is connected to a dedicated health care provider version that supports them in collaborating with family caregivers. Within the health care provider version interface, users can approve new clients with dementia, update client information, and engage in care planning discussions with caregivers. The health care provider version also enables providers to participate in the online forum, exchanging insights and offering support to caregivers and other stakeholders.

Moderator Interface

The *CareBuddy* caregiver app is also connected to a moderator version. Moderators of the online forum can manage the forum via the *CareBuddy* web portal, with tools to pin or feature posts and share multimedia caregiving resources.

System Features

Overview

From a system development perspective, we implemented a scalable hybrid architecture combining single-service and multiservice systems to efficiently manage traffic capacity and enhance system resilience. Information entered by caregivers is stored securely in the cloud database and synchronized across services to allow continuity across sessions.

System Usage Tracking

CareBuddy includes an embedded system usage log that captures detailed user activity across the app. This feature tracks user actions to better understand caregiver needs and app engagement patterns. In addition, a data-capturing component tracks interactions within the online forum, including posts, threads, and social dynamics among users. These built-in analytics features support future improvements and enable ongoing research into user behavior.

System and Messaging Notifications

CareBuddy integrates a comprehensive notification system using Firebase Cloud Messaging to keep users engaged and informed. The app supports system-scheduled notifications that provide timely nudges for user engagement and updates within the *CareBuddy* environment. In-app notifications are triggered by specific user interactions, such as messaging between family caregivers and health care professionals, ensuring users stay informed of meaningful exchanges and collaborative care efforts. Additionally, push notifications are delivered even when the app is closed, allowing critical information, such as new messages, care plan updates, or forum interactions, to reach users without delay. These notifications are customized based on user activity and context, including upcoming appointments or scheduled reminders.

Chat History Management

This feature was designed to support comprehensive storage, retrieval, and analysis of past user interactions, incorporating user metadata, conversation sessions, and analytical reporting to continuously refine chatbot performance and personalization.

Data Privacy and Security

Data privacy and security measures were implemented for proper handling of sensitive features and data within the app. GPS location data and chatbot records are stored in an anonymized format and not linked to personally identifiable information in the backend database. Thus, data are decoupled from patients' profile data, and identifiable information is separately stored and managed by the research team.

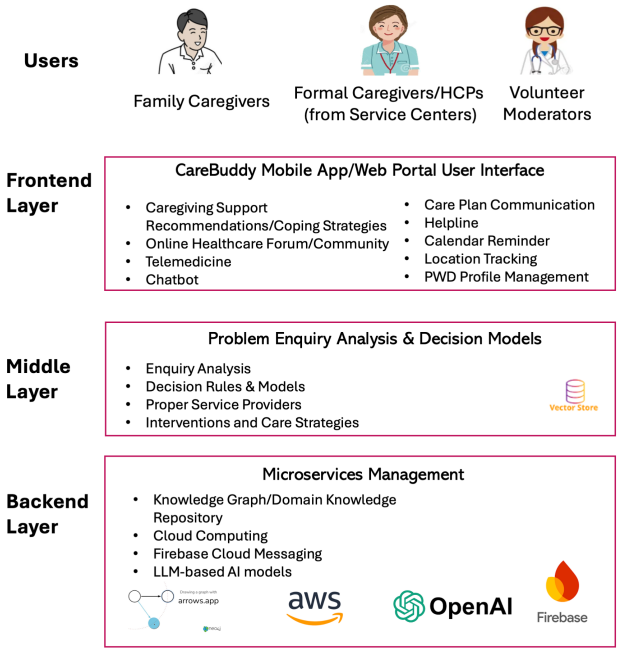
Before the *CareBuddy* mobile app was made available for download (on both the Apple App Store and Google Play Store), the app was reviewed by both platforms. *CareBuddy's* adherence to best practice measures ensured it satisfied both platforms' prevailing compliance and regulatory requirements and ensured privacy and security protection for end users. Furthermore, the privacy policy and End User License Agreement were also made available to end users to elaborate on how content filtering and moderation are implemented to ensure rigor, appropriateness, and security of the app. *CareBuddy* app users also provided their consent and acceptance of the service terms during the account registration process.

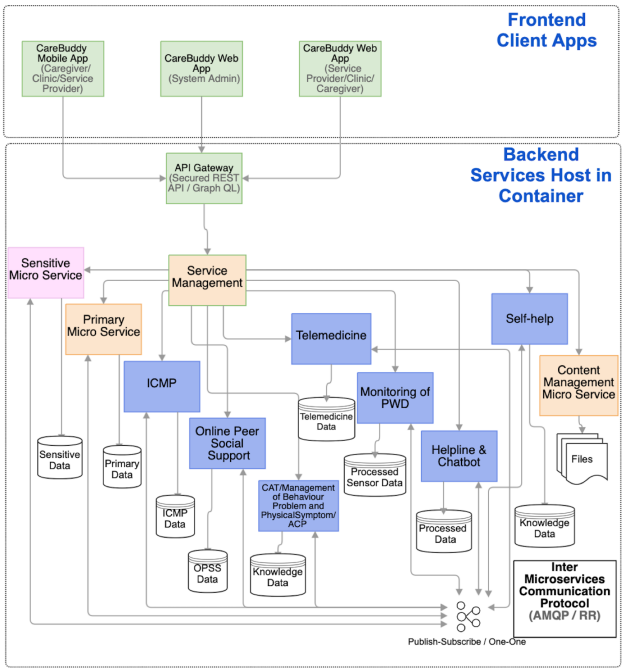
Moreover, all user data and interaction records (including data related to sensitive features) are stored securely in the Amazon Web Services cloud, with its access strictly limited to the designated system administrator from the research team and further protected with multifactor authentication for enhanced security. These measures ensured that user privacy, security, and ethical considerations were upheld throughout the development and pilot testing.

The *CareBuddy* system structure is presented in Figure 2, and sample screenshots of the components, interface, and features are presented in Figure S1 in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 2. *CareBuddy* system structure. ACP: advance care planning; AI: artificial intelligence; AMQP/RR: advanced message queuing protocol/request response; API: application programming interface; CAT: computerized adaptive testing; HCP: health care professional; ICMP: integrated client management portal; LLM: large language model; OPSS: online peer social support; PWD: person with dementia; REST API: representational state transfer application programming interface.





Pilot Testing

A total of 28 participants were involved in the pilot testing, with 18 individuals (11 caregivers and 7 health care providers) in phase 1 and 10 caregivers in phase 2. This sampling aligns with the general rule of thumb, which suggests that 80% of all potential usability problems can be identified by testing with 5 to 10 users [42]. Participant characteristics are described in Table 1.

Quantitative

Usability Scores in Phases 1 and 2

Usability of the app improved in rounds 2 and 3 of phase 1. The overall mean usability scores increased from 65.4 (SD 11.8, range 47.5-80) in round 1 to 77.5 (SD 7.2, range 70-87.5) in round 2 and 73.8 (SD 15.9, range 52.5 - 97.5) in round 3. Notably, rounds 2 and 3 exceeded the benchmark usability score of 68 (Table 2). Overall, 67% (12/18) of participants reported SUS scores above 68, including 73% (8/11) of caregivers and 57% (4/7) of health care providers.

In phase 2, participants rated the app highly, with an overall total mean score of 95.4 (SD 8.5). The ease of use (mean 24.1, SD 2.9), user interface and satisfaction (mean 40.3, SD 3.4), and usefulness (mean 31, SD 3.9) were high (Table 3). Figure 3 shows that the majority of the participants somewhat agreed, agreed, or strongly agreed with all items of MAUQ. All participants either agreed or strongly agreed with "Overall, I am satisfied with this app."

Table . System Usability Scale scores by participant group (phase 1).

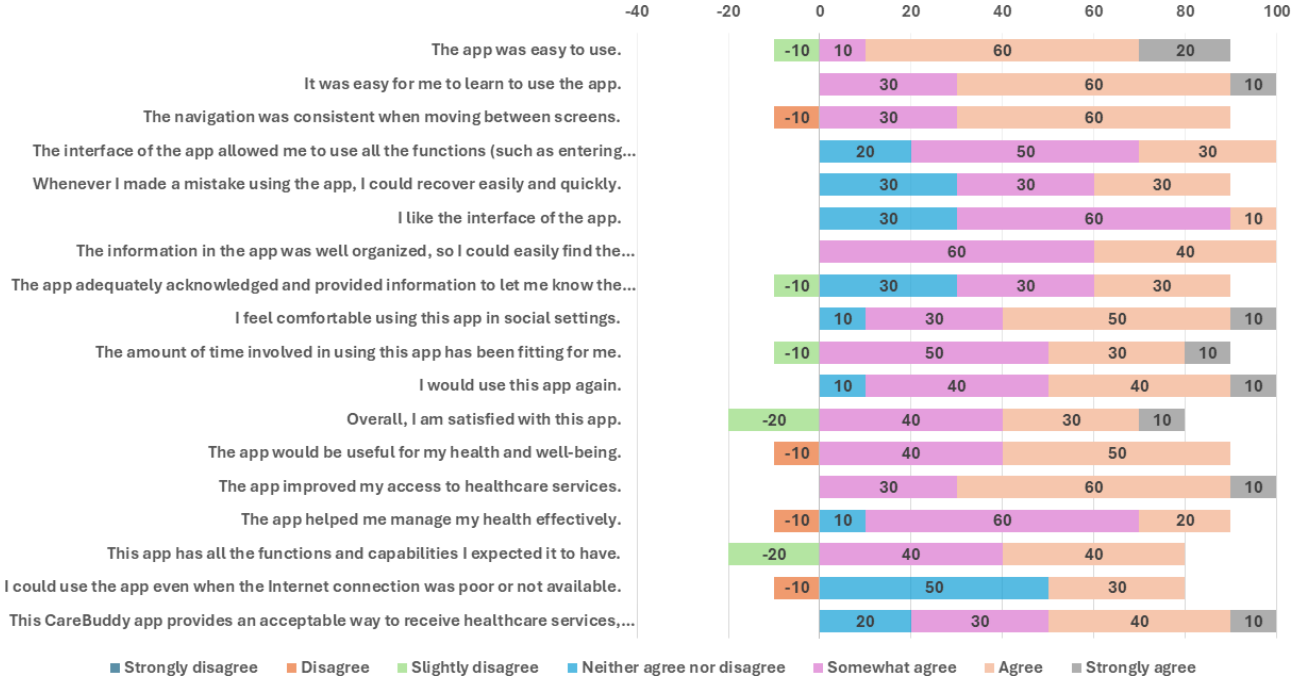
	Caregivers (n=11)	Health care providers (n=7)	Overall (n=18)
Round 1 scores, n	4	2	6
Range	47.5 - 80	57.5 - 67.5	47.5 - 80
Mean (SD)	66.9 (14.3)	62.5 (7.1)	65.4 (11.8)
Round 2 scores, n	4	2	6
Range	70 - 87.5	72.5 - 85	70 - 87.5
Mean (SD)	76.9 (7.7)	78.8 (8.8)	77.5 (7.2)
Round 3 scores, n	3	3	6
Range	52.5 - 97.5	62.5 - 72.5	52.5 - 97.5
Mean (SD)	78.3 (23.2)	69.2 (5.8)	73.8 (15.9)

XSL•FO
RenderX

Table . Usability and satisfaction score among caregivers (n=10; phase 2).

Mobile Health App Usability Questionnaire domains	Sample score range	Score, mean (SD)
Ease of use (range 5-35)	18 - 30	24.1 (2.9)
Interface and satisfaction (range 7-49)	34 - 47	40.3 (3.4)
Usefulness (range 6-42)	23 - 35	31 (3.9)
Overall (range 18-126)	79 - 107	95.4 (8.5)

Figure 3. Usability and satisfaction—Mobile Health App Usability Questionnaire results.



App Usage in Phase 2

The backend system usage tracking data showed that informational resources, which included content on understanding dementia and managing behavioral and physical symptoms, financial assistance, and future planning, and the directory of service providers, were used for a median of 18.8 minutes, with a maximum of 35.5 minutes. Resources focused on caregiver self-care, such as managing stress, caring for themselves, navigating the caregiving journey, and coping with grief, were used for a median of 7.5 minutes with a maximum of 28.5 minutes, and the online social forum for 9.4 minutes with a maximum of 1.7 hours.

Qualitative

Overview

The analysis of substantive interview feedback and suggestions from phases 1 and 2 is summarized under 5 domains of the User Version of the Mobile Application Rating Scale, with representative quotes from participants.

Domain: Information and Content

Participants in phase 1 found the information credible, high quality, and easy to understand, covering relevant caregiving topics. Users in round 1 found the text wordy, suggesting bullet points, visuals, and a ‘view more’ option for details. With progressive iterations and changes made in response to suggestions from earlier rounds, most participants in round 3 were satisfied with content clarity and organization. They perceived the content to be rich, straightforward, containing extensive resources, and providing solutions for managing common behavioral symptoms. One health care provider mentioned:

> *Based on the three different broad categories I have clicked, I think so far, I'm quite impressed. I think that it gives me a quick glimpse of you know why my loved one may exhibit the some behaviour. And then you don't just leave me hanging, you also give me certain solution to work towards.* [Phase 1, participant 16]

Caregivers in phase 2 valued the content on dementia trajectory, management of behavioral and physical symptoms, predeath grief, stress management, and self-care content, finding it relatable. They found the content well organized and easy to use. In comparison to existing local apps, they found *CareBuddy* comprehensive and eliminated the need for search engines such as Google.

Domain: Usability and Functionality

While most participants in phase 1 found the app easy to use, some struggled with navigation and chatbot performance. Participants suggested making the chatbot responses more empathetic. In response, we enhanced navigation by adding

XSL•FO
RenderX

subheaders for each section and incorporating hyperlinks to external resources. We also improved the chatbot's vocabulary, ensuring sensitivity for better responsiveness. A health care provider noted:

> *It is pretty much comprehensive and easy to navigate. These are the common things that most of them [caregivers] will be able to look at.* [Phase 1, participant 09]

In the 1 month of using the app in phase 2, caregivers found the search function in the directory intuitive. They also used the interactive app elements such as a GPS tracker, worksheets, a chatbot, and a directory extensively. Most participants found no issues with navigation; however, some reported missing links to external sites and noted that the tracker function required frequent refreshing to update the care recipient's location.

Domain: Aesthetics and Design

Early feedback from phase 1 cited small fonts, cluttered layout, and dull colors. In response, we made adjustments to text size, formatting, and color contrast for improved readability. We added an audio option for users who preferred listening to the content. Later feedback was very positive. A health care provider reported:

> *I like that you colour coded into the broad category so that, I can instinctively try to find a common theme and then go to it directly. So yeah, I thought that was very nice.* [Phase 1, participant 16, HCP]

Domain: Engagement

Many participants in phase 1 showed interest and eagerness to read the app content and suggested providing empathetic responses to ensure appropriateness. Moreover, providing practical solutions rather than theoretical responses ensured engagement with the app.

In phase 2, at least 40% of caregivers used the app every day. Many participants had positive feedback and felt happy to connect with fellow caregivers on the community portal. Others preferred to just hear others and found information shared by other caregivers to be helpful. They reported that the moderators did a good job of initiating discussions and engaging users. One caregiver reported:

> *In the beginning I was [checking] almost every day. Not only every day like different times of the day. ...I think the moderator did quite good job in like throwing questions like what the activities are you do and then I shared. The last one he asked is what do you do at the social club? It's good to share my knowledge. So, I think that is one of the most useful one (component), the one that I go to a lot.* [Phase 2, participant 01]

Domain: Relevance and Perceived Usefulness

Feedback from participants in early phase 1 was mixed. Some felt the app to be more relevant for early dementia diagnosis, whereas others believed it was suited for later-stage dementia care. In response, we enhanced the content of the app, incorporating more examples of how care can be provided. Caregivers in round 3 perceived the app to be more useful, functioning as a valuable guide and offering real-time support with practical solutions for addressing behavioral and physical symptoms. Some of the participants were keen and excited about the launch of the app due to the lack of a comprehensive app such as *CareBuddy*. A caregiver noted that:

> *I think it's very helpful to care partner like me. When I need help, I know this app is around. When first diagnosed, I was not too sure what to do. Totally lost, feel so scary, sad...upset, worrying, yeah. So, it's like, you know, very helpless, do not know where to go.* [Phase 1, participant 14]

A health care provider also reported that:

> *I am very excited for this app to be launched because right now in the market don't think there is really an app which covers so much information.* [Phase 1, participant 16]

Over a 1-month trial, caregivers found the app solutions helpful in managing behavioral challenges (eg, agitation). The easy access to resources and content enabled a better understanding of dementia progression. Caregivers reported that the curated content saved time, providing quick access to care resources such as nursing homes, respite care, and mental health helplines:

> *I saw the list and I found the ones [service providers] that are nearer to my house. And of course, I'm very happy with the phone numbers. Because very important to have phone numbers I can call immediately.* [Phase 2, participant 06]

Participants appreciated the *CareBuddy* as comprehensive and were keen to recommend the app to others. One of the caregivers reminisced that having the *CareBuddy* app would have been immensely valuable when their loved one had been newly diagnosed with dementia.

The changes made to the app based on participants' suggestions by phases are described in Table S2 in Multimedia Appendix 1, and qualitative analysis with illustrative quotes is detailed in Multimedia Appendix 2.

Discussion

Principal Findings

CareBuddy is a theory-informed mobile intervention purpose-built to support the complex and evolving needs of dementia caregivers. Co-developed by a multidisciplinary team, including health care providers, computer scientists, health services researchers, and caregivers themselves, *CareBuddy* delivers an integrated digital ecosystem that empowers caregivers across the care journey by addressing the practical, emotional, and logistical challenges of caregiving. The app provides continuous, on-demand support through access to educational resources, peer and professional communication channels, and immediate crisis assistance via a built-in helpline and telemedicine access. A mixed methods pilot study demonstrated high usability and satisfaction with user feedback guiding iterative refinements to enhance *CareBuddy*. By offering timely, personalized, and scalable support, *CareBuddy* shows

XSL•FO
RenderX

strong potential to improve caregiver well-being, reduce burden, and promote sustainable, high-quality home care for individuals living with dementia.

In comparison to existing researched apps for caregivers of persons with dementia that have limited features (eg, tracking and caregiver self-care component) [24], *CareBuddy* is grounded in the stress process model and targets caregiving stressors through a conversational AI for caregiver queries, a helpline, telemedicine services, and a moderated social networking portal providing real-time assistance. Additional components on financial and legal information, stress-management exercises, self-care resources, and a directory of dementia care services target secondary stressors supporting long-term caregiver well-being. An interactive care management portal, though not tested in this study, connects caregivers to service providers. Its co-design process, involving caregiver representatives, ensured alignment with real-world needs. Features were informed by interviews with caregivers and health care professionals, ensuring practical utility. Thereafter, an iterative mixed methods design further refined the app based on user input. The app is available on both iOS and Android.

Quantitative usability testing showed strong results. The mean (SUS) scores exceeded the threshold of 68, with 67% of participants scoring above this benchmark. The MAUQ scores indicated high overall satisfaction, with strong ratings for ease of use, user interface, and usefulness (overall mean score 95.4, SD 8.5). Caregivers rated satisfaction statements positively, reinforcing *CareBuddy's* acceptability and usability.

Notably, *CareBuddy* is designed to be a dynamic tool, as new features are continuously being developed and tested, with plans for phased integration into future versions of the app. These include decision aids to help caregivers choose between comfort feeding or tube feeding and enhanced AI-driven support for the online platform.

Future Implications

Importantly, the app is built on an adaptable and modular architecture, allowing for seamless customization to suit different health care systems, cultural contexts, and caregiving environments. Whether implemented in community-based care settings, acute hospitals, or long-term care facilities, the *CareBuddy* platform can be configured to align with local workflows, regulatory requirements, and user preferences. This flexibility ensures that *CareBuddy* is not only scalable but also sustainable across diverse care ecosystems.

Limitations

Despite its promise, this study has limitations. The pilot involved a small sample size and short follow-up duration, consistent with similar studies [43]. While usability and acceptance testing have been essential, further research is needed to assess caregiver engagement and the app's impact on caregiving outcomes. A 2-arm hybrid type 1 randomized controlled trial is planned to evaluate implementation and effectiveness. Additionally, long-term user retention remains a challenge, as prior studies indicate sharp declines in engagement after initial download [44,45]. To mitigate this, we have incorporated a notification system to encourage sustained use and maximize the app's impact. New additional features are being tested and will be added to the app, making the app dynamic. Finally, the app is designed to be adaptable and modular, that is, it can be adapted to different contexts.

Conclusions

CareBuddy, envisioned as a comprehensive digital caregiving hub, holds promise to address the multifaceted challenges of dementia caregiving. Its co-design methodology and iterative feedback loops have ensured that the app is closely aligned with the real-world needs of dementia caregivers, enhancing its usability, accessibility, and relevance. Future research, including a randomized controlled trial, will evaluate the app's long-term effects on caregiver outcomes and key implementation metrics, such as engagement, acceptability, and sustainability. Concurrently, plans are in place to integrate *CareBuddy* into routine care delivery, with adaptations to accommodate different health care systems, languages, and cultural contexts. These efforts aim to extend the app's relevance and scalability across diverse international settings, broadening its reach and impact on dementia caregiving worldwide.

Funding

This work was supported by the National Medical Research Council, Ministry of Health, Singapore (grant MOH-001065; CareEco21-0028). The funders had no role in the study design, data collection, interpretation of the results, or writing of the manuscript.

Data Availability

The data used during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Qualitative themes on desired app features with illustrative quotes, updates made to the CareBuddy app across interview rounds, and user-interface screenshots.
[DOCX File, 519 KB - aging_v8i1e78759_app1.docx]

XSL•FO
RenderX

Multimedia Appendix 2
Qualitative analysis with illustrative quotes.
[DOCX File, 20 KB - aging_v8i1e78759_app2.docx]

Checklist 1
COREQ checklist.
[DOCX File, 23 KB - aging_v8i1e78759_app3.docx]

References

1. Nichols E, Steinmetz JD, Vollset SE. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. Lancet Public Health 2022 Feb;7(2):e105-e125. [doi: 10.1016/S2468-2667(21)00249-8] [Medline: 34998485]
2. Dementia. World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2025-01-16]
3. Ory MG, Hoffman RR III, Yee JL, Tennstedt S, Schulz R. Prevalence and impact of caregiving: a detailed comparison between dementia and nondementia caregivers. Gerontologist 1999 Apr;39(2):177-185. [doi: 10.1093/geront/39.2.177] [Medline: 10224714]
4. Reckrey JM, Bollens-Lund E, Husain M, Ornstein KA, Kelley AS. Family caregiving for those with and without dementia in the last 10 years of life. JAMA Intern Med 2021 Feb 1;181(2):278-279. [doi: 10.1001/jamainternmed.2020.4012] [Medline: 33252607]
5. Lastuka A, Bliss E, Breshock MR, et al. Societal costs of dementia: 204 countries, 2000-2019. J Alzheimers Dis 2024;101(1):277-292. [doi: 10.3233/JAD-240163] [Medline: 39150827]
6. Malhotra C, Chaudhry I, Shah SU, PISCES Study Group. Caregivers' burden and anticipatory grief increases acute health care use in older adults with severe dementia. J Am Med Dir Assoc 2024 Jul;25(7):104981. [doi: 10.1016/j.jamda.2024.03.001] [Medline: 38599241]
7. Malhotra C, Huynh VA, Østbye T, Malhotra R. Caregivers' absenteeism and its association with health shocks and functional impairment among persons with severe dementia. J Aging Soc Policy 2024 Jul 3;36(4):603-620. [doi: 10.1080/08959420.2023.2196232] [Medline: 37026473]
8. Chaudhry I, Poco L, Balasubramanian I, Andres EB, Malhotra C. Caregiving experiences as mediators between caregiving stressors and anticipatory grief in severe dementia: findings from longitudinal path analysis. Am J Geriatr Psychiatry 2025 Feb;33(2):143-152. [doi: 10.1016/j.jagp.2024.07.017] [Medline: 39181837]
9. Balasubramanian I, Chaudhry I, Poco LC, Malhotra C, PISCES Study Group. "I secretly wish..." caregivers' expression of wish for death of persons with severe dementia. Age Ageing 2024 May 1;53(5):afae103. [doi: 10.1093/ageing/afae103] [Medline: 38798114]
10. Balasubramanian I, Poco LC, Andres EB, et al. Caregiver coping mediates the relationship between caregivers' understanding of dementia as terminal and their distress. Alzheimers Dement 2024 Sep;20(9):6606-6614. [doi: 10.1002/alz.14102] [Medline: 38971964]
11. Andres EB, Chaudhry I, Balasubramanian I, et al. Navigating dementia feeding decisions in Asia: a mixed methods study of caregivers. Age Ageing 2024 Oct 1;53(10):afae232. [doi: 10.1093/ageing/afae232] [Medline: 39465492]
12. Zain Z, Mohamad H, Seow D, et al. Caregiver preference for tube feeding in community-dwelling persons with severe dementia. J Am Geriatr Soc 2020 Oct;68(10):2264-2269. [doi: 10.1111/jgs.16641] [Medline: 32559331]
13. Ramakrishnan C, Malhotra C. Challenges to accessing community dementia care services: a qualitative study. BMC Health Serv Res 2025 May 23;25(1):747. [doi: 10.1186/s12913-025-12895-3] [Medline: 40410863]
14. Balasubramanian I, Andres EB, Poco LC, Malhotra C. Prognostic understanding among caregivers of persons with dementia: a scoping review. J Am Geriatr Soc 2025 Mar;73(3):947-955. [doi: 10.1111/jgs.19245] [Medline: 39446015]
15. Poco LC, Andres EB, Balasubramanian I, Malhotra C. Financial difficulties and well-being among caregivers of persons with severe dementia. Aging Ment Health 2025 Jul;29(7):1338-1347. [doi: 10.1080/13607863.2025.2481978] [Medline: 40129248]
16. Ploeg J, Ali MU, Markle-Reid M, et al. Caregiver-focused, web-based interventions: systematic review and meta-analysis (part 2). J Med Internet Res 2018 Oct 26;20(10):e11247. [doi: 10.2196/11247] [Medline: 30368439]
17. Frias CE, Garcia-Pascual M, Montoro M, Ribas N, Risco E, Zabalegui A. Effectiveness of a psychoeducational intervention for caregivers of people with Dementia with regard to burden, anxiety and depression: a systematic review. J Adv Nurs 2020 Mar;76(3):787-802. [doi: 10.1111/jan.14286] [Medline: 31808211]
18. Leng M, Zhao Y, Xiao H, Li C, Wang Z. Internet-based supportive interventions for family caregivers of people with dementia: systematic review and meta-analysis. J Med Internet Res 2020 Sep 9;22(9):e19468. [doi: 10.2196/19468] [Medline: 32902388]

XSL•FO
RenderX

19. McKechnie V, Barker C, Stott J. Effectiveness of computer-mediated interventions for informal carers of people with dementia-a systematic review. Int Psychogeriatr 2014 Oct;26(10):1619-1637. [doi: 10.1017/S1041610214001045] [Medline: 24989249]
20. Parra-Vidales E, Soto-Pérez F, Perea-Bartolomé MV, Franco-Martín MA, Muñoz-Sánchez JL. Online interventions for caregivers of people with dementia: a systematic review. Actas Esp Psiquiatr 2017 May;45(3):116-126. [Medline: 28594057]
21. Hopwood J, Walker N, McDonagh L, et al. Internet-based interventions aimed at supporting family caregivers of people with dementia: systematic review. J Med Internet Res 2018 Jun 12;20(6):e216. [doi: 10.2196/jmir.9548] [Medline: 29895512]
22. Kim E, Baskys A, Law AV, Roosan MR, Li Y, Roosan D. Scoping review: the empowerment of Alzheimer's Disease caregivers with mHealth applications. NPJ Digit Med 2021 Sep 7;4(1):131. [doi: 10.1038/s41746-021-00506-4] [Medline: 34493819]
23. Faieta J, Sheehan J, DiGiovine C. mhealth interventions to improve health and quality of life related outcomes for informal dementia caregivers: a scoping review. Assist Technol 2022 May 4;34(3):362-374. [doi: 10.1080/10400435.2020.1829174] [Medline: 33006526]
24. Ye B, Chu CH, Bayat S, Babineau J, How TV, Mihailidis A. Researched apps used in dementia care for people living with dementia and their informal caregivers: systematic review on app features, security, and usability. J Med Internet Res 2023 Oct 12;25:e46188. [doi: 10.2196/46188] [Medline: 37824187]
25. Wozney L, Freitas de Souza LM, Kervin E, Queluz F, McGrath PJ, Keefe J. Commercially available mobile apps for caregivers of people with Alzheimer disease or other related dementias: systematic search. JMIR Aging 2018 Dec 7;1(2):e12274. [doi: 10.2196/12274] [Medline: 31518255]
26. Werner NE, Brown JC, Loganathar P, Holden RJ. Quality of mobile apps for care partners of people with Alzheimer disease and related dementias: Mobile App Rating Scale evaluation. JMIR mHealth uHealth 2022 Mar 29;10(3):e33863. [doi: 10.2196/33863] [Medline: 35348467]
27. Zou N, Xie B, He D, Hilsabeck R, Aguirre A. mHealth apps for dementia caregivers: systematic examination of mobile apps. JMIR Aging 2024 Nov 20;7:e58517. [doi: 10.2196/58517] [Medline: 39621936]
28. Molina-Recio G, Molina-Luque R, Jiménez-García AM, Ventura-Puertos PE, Hernández-Reyes A, Romero-Saldaña M. Proposal for the user-centered design approach for health apps based on successful experiences: integrative review. JMIR mHealth uHealth 2020 Apr 22;8(4):e14376. [doi: 10.2196/14376] [Medline: 32319965]
29. Norman DA, Draper SW. User Centered System Design: New Perspectives on Human-Computer Interaction: CRC Press; 1986.
30. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
31. Pearlin LI, Mullan JT, Semple SJ, Skaff MM. Caregiving and the stress process: an overview of concepts and their measures. Gerontologist 1990 Oct;30(5):583-594. [doi: 10.1093/geront/30.5.583] [Medline: 2276631]
32. Zhao Y, Hughes M, Teaster PB. Exploring caregiving stressors and informal resources among Alzheimer's caregivers. Aging Health Res 2022 Sep;2(3):100091. [doi: 10.1016/j.ahr.2022.100091]
33. Brooke J. SUS: a quick and dirty usability scale. In: Usability Evaluation in Industry, 1st edition: CRC Press; 1995, Vol. 189.
34. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): development and validation study. JMIR mHealth uHealth 2019 Apr 11;7(4):e11500. [doi: 10.2196/11500] [Medline: 30973342]
35. Hyzy M, Bond R, Mulvenna M, et al. System Usability Scale benchmarking for digital health apps: meta-analysis. JMIR mHealth uHealth 2022 Aug 18;10(8):e37290. [doi: 10.2196/37290] [Medline: 35980732]
36. Lewis JR. The System Usability Scale: past, present, and future. Int J Hum Comput Interact 2018 Jul 3;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
37. Mustafa N, Safii NS, Jaffar A, et al. Malay version of the mHealth App Usability Questionnaire (M-MAUQ): translation, adaptation, and validation study. JMIR mHealth uHealth 2021 Feb 4;9(2):e24457. [doi: 10.2196/24457] [Medline: 33538704]
38. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
39. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the User Version of the Mobile Application Rating Scale (uMARS). JMIR mHealth uHealth 2016 Jun 10;4(2):e72. [doi: 10.2196/mhealth.5849] [Medline: 27287964]
40. Kales HC, Gitlin LN, Stanislawski B, et al. The development of a caregiver-focused, web-based program to assess and manage behavioral and psychological symptoms of dementia. Alzheimer Dis Assoc Disord 2017;31(3):263-270. [doi: 10.1097/WAD.0000000000000177] [Medline: 27849639]
41. Hamilton P, LeClair JK, Collins J, Sturdy-Smith C, O'Connell M. PIECES - guiding collaborative engagement, shared assessment, and supportive care. Pieces Canada. 2020. URL: https://piecescanada.com/pages/pieces-guiding-collaborative-engagement-shared-assessment-and-supportive-care [accessed 2025-12-04]
42. Kushniruk AW, Patel VL. Cognitive and usability engineering methods for the evaluation of clinical information systems. J Biomed Inform 2004 Feb;37(1):56-76. [doi: 10.1016/j.jbi.2004.01.003] [Medline: 15016386]

XSL•FO
RenderX

43. Ali Z, Zhong J, Sadarangani TR. A mixed-methods examination of the acceptability of, CareMOBI, a dementia-focused mHealth app, among primary care providers. Digit Health 2024;10:20552076241287361. [doi: 10.1177/20552076241287361] [Medline: 39502491]
44. Wei Y, Zheng P, Deng H, Wang X, Li X, Fu H. Design features for improving mobile health intervention user engagement: systematic review and thematic analysis. J Med Internet Res 2020 Dec 9;22(12):e21687. [doi: 10.2196/21687] [Medline: 33295292]
45. Vaghefi I, Tulu B. The continued use of mobile health apps: insights from a longitudinal study. JMIR mHealth uHealth 2019 Aug 29;7(8):e12983. [doi: 10.2196/12983] [Medline: 31469081]

Abbreviations

ABC: antecedent, behavior, and consequences
AI: artificial intelligence
COREQ: Consolidated Criteria for Reporting of Qualitative Studies
DICE: describe, investigate, create, evaluate
LLM: large language model
MAUQ: Mobile Health App Usability Questionnaire
P.I.E.C.E.S.: Physical. Intellectual. Emotional. Capabilities. Environment. Social.
Q&A: question and answer
SUS: System Usability Scale

Edited by M O'Connell; submitted 09.06.25; peer-reviewed by B Davis, N Zou; revised version received 18.09.25; accepted 03.11.25; published 30.12.25.

Please cite as:
Malhotra C, Yue Y, Ramakrishnan C, Shah S, Chen W, Yap P, Cheong CY, Teo I, Wee SL, Lan X, Chen Y, Chong CS, Huang X, Chua I
Supporting Dementia Caregiving With a Mobile Care Ecosystem: Development and Mixed Methods Study
JMIR Aging 2025;8:e78759
URL: https://aging.jmir.org/2025/1/e78759
doi:10.2196/78759

© Chetna Malhotra, Yanzhen Yue, Chandrika Ramakrishnan, Shimoni Shah, Wenda Chen, Philip Yap, Chin Yee Cheong, Irene Teo, Shiou-Liang Wee, Xiangming Lan, Yi Chen, Chee Seng Chong, Xueying Huang, Ivy Chua. Originally published in JMIR Aging (https://aging.jmir.org), 30.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessing the Time for Living and Caring (TLC) Study: Mixed-Methods Feasibility Study of a Web-Based Caregiver Intervention to Improve Respite

Amber D Thompson[1], PhD; Alexandra L Terrill[2], PhD; Michael Caserta[3], PhD; Bob Wong[3], PhD; Eli Iacob[3], PhD; Catharine Sparks[2], MSc; Louisa Stark[4], PhD; Rebecca L Utz[1], PhD

[1]College of Social & Behavioral Science, University of Utah, 260 S Central Campus Dr, Salt Lake City, UT, United States

[2]College of Health, University of Utah, Salt Lake City, UT, United States

[3]College of Nursing, University of Utah, Salt Lake City, UT, United States

[4]School of Medicine, University of Utah, Salt Lake City, UT, United States

Corresponding Author:
Amber D Thompson, PhD
College of Social & Behavioral Science, University of Utah, 260 S Central Campus Dr, Salt Lake City, UT, United States

Abstract

Background: Interventions that are self-administered and delivered online are increasingly being seen as a flexible way to support family caregivers. Intervention research should prioritize the measurement of feasibility throughout all of the stages of intervention development and evaluation to provide the essential feedback loop needed for the iterative development and refinement process.

Objective: We describe the methodology and data used to assess the feasibility, usability, and acceptability of the Time for Living and Caring (TLC) intervention, a technology-delivered intervention (app) for dementia caregivers to improve respite time use.

Methods: The feasibility analysis is theoretically guided by a multidimensional definition of feasibility and uses a mixed-methods research design. Stakeholder feedback collected via focus groups during intervention development (n=15), self-reported surveys from participants enrolled in the pilot trial of the intervention (n=163), surveys of a nationwide sample of respite providers (n=57), and end-user statistics, captured passively by Google Analytics from those using the app, were used in the feasibility analysis of the TLC intervention.

Results: The TLC study used an appropriate design and data collection procedures, along with acceptable recruitment capability. Out of 5 intervention features, 4 received favorable ratings (range of 82% - 99%) by intervention participants and respite providers, which, when combined with open-ended recommendations for improvements, indicate a high degree of usability. Acceptability was measured through appraisal of the intervention experience (135/159, 85% positive), potential future use (127/163, 78%), willingness to recommend (148/163, 91%), and perceived benefit (135/163, 83%).

Conclusions: Taken together, the data suggest that the TLC app is a promising intervention that could be implemented as an on-demand resource for respite-using caregivers, irrespective of where they are located or when they choose to access it. Additionally, this paper provides a blueprint for systematically evaluating multiple dimensions of feasibility, using various forms of mixed-methods data collected during intervention development and pilot testing of an intervention, which should help streamline the eventual implementation of effective interventions in real-world settings.

Trial Registration: ClinicalTrials.gov NCT03689179; https://clinicaltrials.gov/study/NCT03689179

(JMIR Aging 2025;8:e71792) doi:10.2196/71792

KEYWORDS

respite care; internet-based intervention; feasibility studies

Introduction

Family Caregivers and Need for Interventions

As the population demand for family caregivers increases [1] and as there is greater public awareness of both the value that caregivers provide [2] and the challenges that they face [3,4], there has been a call for the development and implementation of caregiver support programs [5]. As a result, and not surprisingly, there has been a proliferation of new evidence-based and research-informed caregiver interventions in recent years [6-9].

XSL•FO
RenderX

Respite—defined as a planned break or time away from caregiving [10]—is among the most needed and requested forms of caregiver support [11]; yet, few caregiver interventions address respite use, likely reflecting a lack of understanding of why there are low rates of respite uptake among family caregivers [1] or the mechanisms underlying the negative and mixed results regarding the effectiveness of respite overall [12]. Nevertheless, empirical research showing how caregivers used their respite time and how respite time use affected respite satisfaction and overall well-being [13,14] paved the way for the development of a novel respite-focused intervention called the Time for Living and Caring (TLC). TLC provides caregivers with coaching and educational resources to schedule and plan their respite time use. Based on the well-researched "Selective Optimization with Compensation" model [15], TLC uses weekly goal-setting and goal-review exercises to help caregivers schedule and plan respite and increase self-awareness about their need for respite. It was hypothesized that with better respite time use, caregivers would have better overall well-being and less caregiving-related stress [16].

TLC, like many of the other newer caregiver support interventions [17,18], was developed to be self-administered and delivered via an interactive website or "app"—with the rationale that online delivery is a flexible way to support caregivers who are often time-limited and place-bound, given the demands of providing round-the-clock care to someone in the home [19]. However, many have questioned the appropriateness of using web-based technologies with an older caregiving population (the average age of caregivers in the United States is 49, with 54% 50 y and older) [1]. Older adults are thought to be less interested, able, or willing to use web-based technologies for support and have, on average, less access to the internet and internet-capable devices and lower computer proficiency to access online resources [20], potentially threatening the feasibility of the TLC intervention and reducing its potential to support family caregivers.

Intervention Development and Evaluation Studies

Intervention research starts with an intervention idea, which is often guided by a theoretical model or previous stage-01 empirical research findings. Once developed, interventions should be pilot-tested, followed by more formal testing and evaluation in various controlled and real-world settings. The final stage is implementation into existing clinical or community-based settings (see National Institute of Health stage model) [21,22]. There are a variety of methodological approaches and recommended actions at each stage [23]. For example, current best practices recommend using community-engaged partnerships or stakeholders to help the researchers co-develop the features of the intervention [24], and using randomized controlled trials during the formal evaluation and testing stages. Although there is a clear preference for evidence-informed practices, there exists little consensus on the specific strategies that should be used, the data that should be collected, or the metrics that should be analyzed at each stage of intervention development to maximize the intervention's effectiveness and eventual implementation [25].

Rarely is there a clear delineation of when the development stage ends and when pilot testing and formal evaluation should begin. This may be due, in part, to each stage of intervention research emphasizing the need to formally assess feasibility. *Feasibility* is the broad concept used to describe the possibility, readiness, or likely success of an intervention. Gadke et al [26] provide a comprehensive and multidimensional definition of feasibility, which includes 10 distinct considerations related to (1) design procedures, (2) data collection procedures, (3) recruitment capability, (4) practicality, (5) adaptability, (6) social validity, (7) generalizability, (8) integration into existing systems, (9) implementation, and (10) effectiveness (refer to Table 1 for more information about each dimension). Gadke et al's [26] multi-dimensional definition of feasibility has been used as a framework to guide intervention development and implementation [27-29].

XSL•FO
RenderX

Table . Dimensions of feasibility. Adapted from Gadke et al [26].

Dimension	Conceptual description	Metric or research methodology used to assess
Design procedures	Is the research design appropriate and sensitive to evaluating change?	Randomized controlled trial (RCT), with a modified waitlist control design; See Figure 1. TLC study design
Data collection procedures	Are data collection procedures and outcome measures appropriate and sensitive to change?	Participant self-report surveys (n=6), spaced every 4 weeks over a 20-week time frame
Recruitment capability	Can participants who will benefit from and who will implement the intervention be identified and successfully recruited for study?	See Figure 2 and Table 2
Practicality	Can the intervention be implemented with available resources, time, training, and materials?	See Table 3
Adaptability	Is there sufficient flexibility built into the intervention procedures to accommodate diverse needs?	See Table 3
Social validity	Does the intervention have social significance or relevance? (traditionally conceptualized as acceptability)	See Table 4
Generalizability	Does the intervention generalize to non-research settings, over time, or with diverse samples?	See Table 4
Integration into existing systems	To what extent does the intervention align with the infrastructure of the relevant practice setting or system?	See Table 4 and Table 5
Implementation	Can providers implement the intervention with fidelity?	Online “app” standardizes delivery and fidelity.
Effectiveness	Is there any preliminary evidence of the intervention effect in bringing about positive change?	See Table 6

Table . Practicality and adaptability of TLC intervention (n=163[a]).

Indicator/variable	Mean (SD) or percent	Min	Max
Owned internet capable device	96%	—	—
Had internet access	100%	—	—
Preferred device to access intervention			
Desktop computer	23%		
Laptop computer	36%		
Tablet, such as an iPad	13%		
Smart phone	28%		
Other	<1%		
Computer proficiency (scale range 12 - 60)	56.12 (5.14)	35	60
Number of sessions	20.44 (13.84)	0	67
Average duration per session (in *seconds*)	506 (313)	72	1996
Total Time (in *hours*) per enrollee-recruitment	12.6	-	-
Cost per enrollee-recruitment	$205	-	-

[a]All data come from the TLC sample of dementia caregivers and include Google Analytics, self-reported survey responses by caregivers, and cost analysis.

XSL•FO
RenderX

Table . Feedback about the TLC[a] app from caregivers (n=132) and from providers (n=57).

	Dementia caregivers	Respite providers
Usefulness or importance of specific TLC features[b]		
Introductory video	86%	97%
Calendar	90%	90%
Coaching (goal setting/review)	88%	95%
Dashboard	82%	85%
Resources	67%	98%
Future use and delivery of TLC intervention		
Would recommend app to others	91%	89%
Would use app in future if available	78%	—
App could be used by all caregivers (not just dementia caregivers)	—	75%

[a]TLC: Time for Living and Caring.

[b]Caregivers were asked about the usefulness of each feature; providers were asked about the importance of each intervention feature.

Table . Exemplars of why participants would or would not use a TLC smartphone app.

YES—I would use	NO—I would not use
I rarely use my laptop anymore, just my cell phone. That way, I can get on and do it any time, anywhere.	*I don't like using the small screen for calendaring.*
I ALWAYS have my phone with me.	*I don't like using app's on my phone. I try not to have it glued to me. I like to use different devices and phone is not one of them. I spend the majority of my day at a computer working so picking up a phone after that is not at the top of my list.*
Simpler, easier to incorporate into my existing life. Being a caregiver--for me--means I don't have as much time to spend on the computer.	
It would be much better as an app. Sometimes having to go the extra work of only being able to access the website from my desktop actually added a little more stress.	*My poor old arthritic fingers have to type too much on a tiny iPhone 8+ keyboard already. So much texting. Everyone wants to text! Sometimes a pain. And after many hours, some strain on the eyes. My notebook easier - 13" screen. Bigger keyboard. And I do not want one more app on my phone. But that's just me...*
it would be convenient to access when I'm at the doctor's or am waiting in line. I spend a lot of down time at the doctors.	
Easier for people who don't have or do well on tablets and computers.	*That's actually a "maybe" depending on how well the app works. For me it was useful to sit down at the computer once a week and do the planning.*

Traditionally, a feasibility study is conducted prior to an outcomes-focused pilot study and before any formal full-scale evaluation effort to test an intervention's effectiveness. However, the terms pilot study and feasibility study are often used interchangeably, reinforcing the importance of systematically assessing feasibility during the early stages of intervention development. Pilot studies are used to "try out" an intervention and to develop research procedures that may provide clarity and input on the mechanisms underlying the intervention's effectiveness in general and within specific and unique contexts [30]. This is what Chen [31] calls "viability validity"—evidence that an intervention is practical, affordable, suitable, valuable, and helpful to people using it in the real world. Conducting systematic and rigorous evaluations of an intervention's feasibility can help to avoid wasting time, money, and resources and ultimately lead to an intervention that is feasible, acceptable, or effective for the intended purpose and population [23,25,32]. Conducting a systematic and comprehensive feasibility analysis provides the essential feedback loop needed for the iterative development and refinement process as interventions are developed, tested, and eventually implemented [32,33].

The Present Study

The objective of this paper is to comprehensively assess the feasibility of the TLC respite-focused caregiver intervention (for an Assessing Feasibility of an Online Caregiver Intervention to Improve Respite: The Time for Living and Caring (TLC) Study of TLC intervention outcomes, see Iacob et al [34]). Our analysis is theoretically guided by the multidimensional definition of feasibility offered by Gadke et al [26] and uses multiple sources of data from the National Institute on Aging funded pilot study of the TLC intervention (NIA R01–AG061946). Results provide specific evidence of whether TLC, an online web-based intervention, is a potentially promising new intervention model to support family caregivers. Additionally, a review of our evaluation methods and metrics provides a blueprint for how other researchers can assess feasibility during the early stages of intervention development and pilot testing. Systematic and comprehensive evaluation of feasibility—at all stages of intervention research—is imperative

XSL•FO
RenderX

if the goal is to implement the most effective and potent programs to support and address the needs of the target or intended population [21]—in this case, the 1 in 5 American adults (21.3% or 53 million) who provide an estimated $600 billion worth of unpaid care and support to an adult or child with special needs in our communities [1,2]. Therefore, the aims of this study are twofold: (1) to assess the feasibility of the TLC respite-focused caregiver intervention and (2) to provide an example and process of how to assess feasibility during the early stages of intervention development, refinement, and pilot testing.

Methods

Data

Data come from the TLC research study, a pilot study funded by the National Institute on Aging (R01—AG061946) with specific aims to refine and evaluate the TLC online respite-focused intervention ("app") among a sample of dementia caregivers. In total, 4 data sources are used to conduct our feasibility analysis: (1) stakeholder feedback collected via focus groups conducted during the intervention development stage (n=15 community advisory board members), (2) self-reported surveys from participants enrolled in the pilot trial of the intervention (n=163), (3) end-user statistics captured passively by Google Analytics from those using the app, and (4) surveys of a nationwide sample of respite providers (n=57) (ie, health or social care professionals who are supporting family caregivers with service delivery or advocacy). More information about the TLC study, including results on the efficacy of TLC to improve caregiver outcomes [34] and access to an archive of all de-identified data and study protocols, can be found on the HIVE archive at the University of Utah [35]. A free, publicly available version of the TLC intervention can be found online [36] (see Multimedia Appendix 1for intervention screenshots).

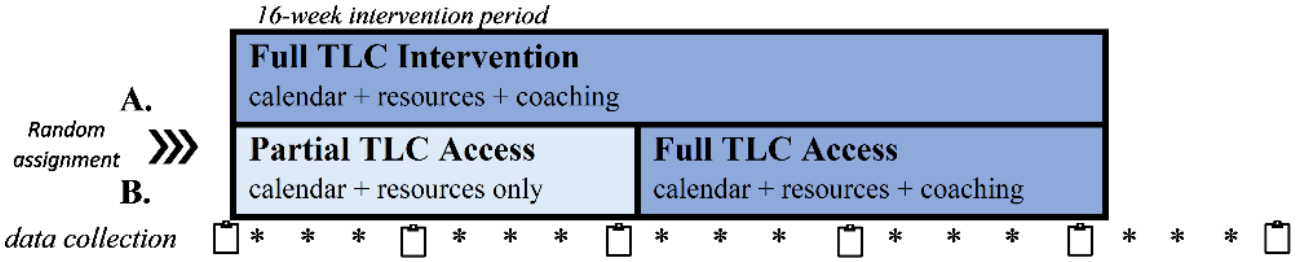
Ethical Considerations

All study procedures were approved by the Institutional Review Board at the University of Utah (# 00120589) and were pre-registered via Clinical Trials (NCT03689179) [37]. Informed consent was obtained from participants after the nature and possible consequences of the studies were explained. To protect participants' privacy and confidentiality, no identifying details have been included.

Measures

The primary outcome for this analysis is feasibility. Using a multidimensional framework for feasibility, Table 1 describes the 10 dimensions of feasibility [26], as well as a description of the specific types of data or metrics we used to assess each dimension: (1) design procedures, (2) data collection procedures, (3) recruitment capability, (4) practicality, (5) adaptability, (6) social validity, (7) generalizability, (8) integration into existing systems, (9) implementation, and (10) effectiveness.

Figure 1. TLC (Time for Living and Caring) study design.



Analytic Strategy

Following recommendations by Aschbrenner et al [38], we used a mixed-methods study design to conduct this feasibility analysis. We first identified which data could be used to assess each of the 10 dimensions of feasibility (eg, rating by study participants of their overall experience with the intervention versus open-ended comments about what they liked most or least). Next, we aligned quantitative and qualitative data sources for the feasibility domain(s) selected (eg, specifying benchmarks and identifying the most relevant participants and sources for the selected feasibility domain(s)). A third strategy was to determine the timing and methodology needed for each assessment (ie, focus groups, study surveys, and provider surveys collected throughout the intervention development and pilot testing). Next, we planned integrative analyses using joint displays to understand feasibility. For example, using a traditional CONSORT diagram, used to describe sample representativeness in a randomized controlled trial study design, along with a descriptive table that compares the study sample to national caregiver population characteristics, allows for a comprehensive and nuanced understanding of the domain of recruitment capability. The last step involved interpreting the quantitative, qualitative, and mixed-methods feasibility results to draw conclusions and meta-inferences about the intervention (eg, the implications and possibilities for intervention refinement that may optimize intervention effectiveness).

Results

Overview

This section provides a description and discussion of the specific metrics, research methodologies, and analyses used to assess the TLC intervention on each of the 10 dimensions of feasibility.

Study Design

The larger TLC study adopted a comprehensive set of mixed-methods research activities and a randomized controlled trial to conduct the formal pilot testing of the TLC intervention (ie, Stage 1A and Stage 1B of the National Institute on Aging stage model for behavioral interventions [21]). This resulted in 4 distinct sources of data that could be used for the feasibility analysis: (1) qualitative feedback from a 15-member community advisory board who worked closely with the research team,

XSL•FO
RenderX

software developers, and creative specialists (writers, media producers, graphic designers) to redevelop and refine the conceptual features of TLC intervention into a self-administered web-based intervention platform (app) that reflected the diverse need and experiences of caregivers, (2) a fully-powered pilot sample of dementia caregivers who had access to the TLC intervention for 16 weeks and who completed self-report questionnaires every week for 20 weeks (n=163), (3) end-user statistics describing engagement with the web-based intervention platform passively captured by Google Analytics for all intervention participants (n=163), and (4) a nationwide sample of respite providers (n=57), who offered an additional layer of quantitative and qualitative data regarding the feasibility and usability of the intervention, with a particular focus on its potential implementation within existing caregiver support services and networks.

As shown in Figure 1, the TLC participant sample used a modified waitlist control design, where participants were assigned to either Group A (the full treatment group, where they received all components of the interventions for the full 16 wk intervention period) or Group B (the partial or delayed access group, where they received the full intervention after an initial 8 wk period during which they only had partial access to the intervention). This randomized controlled trial study design allows each participant to serve as their own control (ie, baseline pre-intervention compared to any other survey), while also preserving the powerful inter-group comparisons that come with randomized controlled designs.

Data Collection

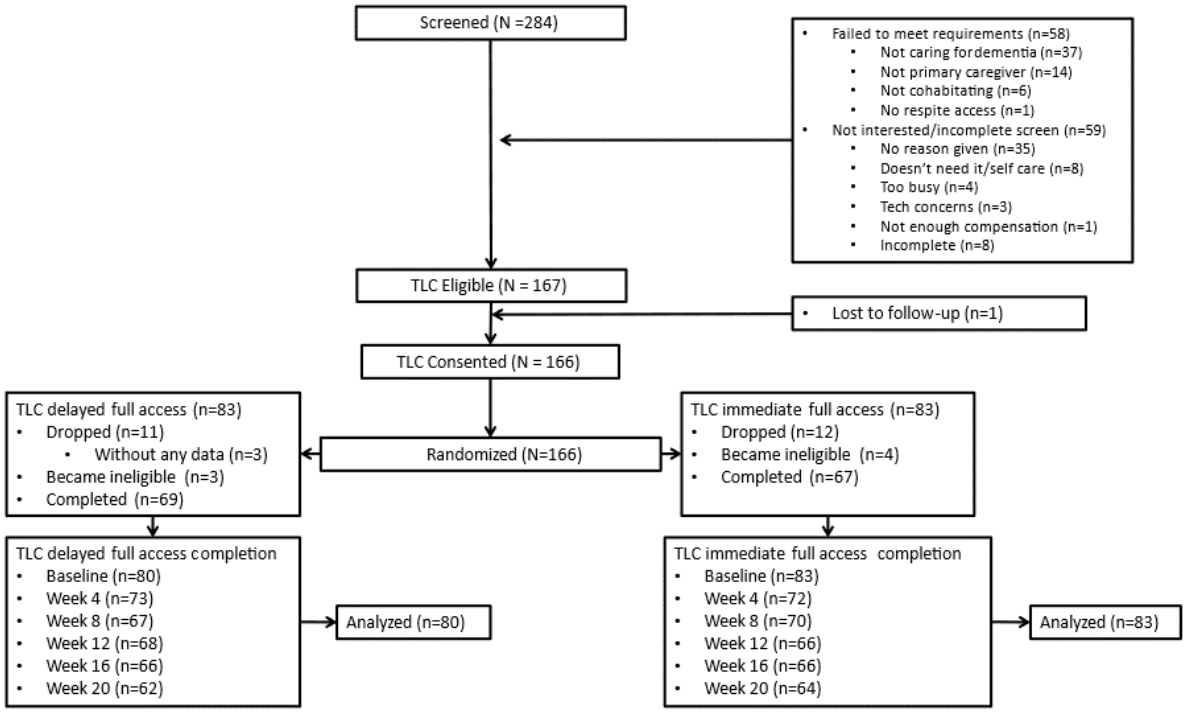
To ensure sensitivity to change, the TLC study used frequent and consistent measurements of outcomes via monthly surveys (baseline, weeks 4, 8, 12, 16, and 20). During enrollment, participants provided information to assess eligibility and provided an email that could be used as a contact throughout the study to send survey links and reminders. Upon enrollment, participants were emailed a baseline survey that included characteristic measures (eg, sociodemographic, caregiving situation, and computer proficiency), pre and post measures (eg, positive affect, desire to institutionalize, and respite characteristics), and repeated measures (eg, depression, anxiety, and burden). We used reliable instruments (eg, the PROMIS Depression short-form questionnaire for adults, an 8-item additive scale that standardizes the distribution of depression-related symptoms on a population distribution) [39], implemented standardized protocols including automated text and targeted email reminders, and collected data across multiple time points. Specifically, participants were emailed a repeated measures survey every 4 weeks after completion of the baseline survey up to week 16. At week 20 (conclusion of the study), participants were emailed a follow-up survey. Additionally, within the app, participants answered weekly questions about respite time amounts and respite satisfaction.

Recruitment Capability

Feedback on the TLC intervention and TLC study protocols was sought from a community advisory board and collected via focus groups [24] before participant recruitment. These relationships provided collaborative opportunities to develop recruitment strategies. Collaborations with community groups and health care professionals, and leveraging online platforms, optimized recruitment efforts of the TLC study [40]. Recruitment of eligible caregivers (n=163) occurred from October 2020 to March 2022 (see Figure 2). The TLC sample included dementia caregivers who were, on average, 61.7 years of age (standard deviation 13.0, min 20, max 92). They were primarily spouses or partners (68.1%) or adult children of care recipients (24.1%). Most were female (78.9%), White (82.5%), non-Hispanic (90.4%), married or living with a partner (83.7%), had a college education (89.7%), and had incomes greater than $50,000 annually (73.6%). The participants recruited for the TLC study were generally comparable to national caregiving populations, with identifiable differences that would be expected based on disease focus and geographic focus (eg, age, ethnicity, caregiver, and recipient age) (see Multimedia Appendix 1).

XSL•FO
RenderX

Figure 2. CONSORT diagram of TLC (Time for Living and Caring) immediate (all feature access) and delayed (initial partial access) randomization arms.



Practicality

Considering "practicality" as a dimension of feasibility helps researchers make informed decisions about whether the study can be realistically executed within the available resources, time, and logistical constraints. The timeline for the intervention pilot was 5 years (2018 - 2023). This included development, recruitment, analysis, and dissemination. Table 2 describes the resources and logistics needed to conduct the TLC pilot study. Per enrollee, it cost US$205 and an average of 12.6 hours of research team time. Access to and willingness to use technology were potential barriers; however, 96% of enrolled participants had an internet-capable device, and 100% had access to the internet. For the pilot study, the intervention was designed for use with desktops, laptops, or tablets, but not for smartphones. Overall, 72% of enrolled participants (117 of 163) preferred the capable devices offered, though the remaining 46 of 163 (28%) preferred smartphone delivery if and when that technology may be available for TLC intervention.

Adaptability

Assessing the sub-domain of "Adaptability" involves empirically observing and documenting the study's capacity to flexibly respond to the diverse needs of participants within the scope of the intervention design. We observed whether participants had different experiences by how familiar they were with technology (see Table 2). The pilot study participants had Computer Proficiency Questionnaire scores [41] that ranged from 35 to 60 (the full range of scores could have been 12 to 60), and an average of 56. Though the sample's average score was high, not all participants were confident using technology and computers. As one participant noted, "It was relatively easy. I'm not very computer savvy, and experienced few problems." Additionally, the intervention included a button where participants could request support or ask questions when they needed technical support; personal support was then provided by email, phone call, or text.

TLC was also adaptable in terms of how and when the participant could engage with it. By creating TLC as an online web-based intervention, participants could access it when it was convenient for them within the weekly schedule. Additionally, participants could log in as many or as few times as desired and for the duration they needed. Over the 16 weeks of active use, the number of participant sessions ranged from 0 to 67, with an average of 20 sessions per participant. The duration of sessions ranged from 1 to 30 minutes, with the average session lasting 8.5 minutes.

Social Validity

Social validity refers to the extent to which the intervention is considered meaningful, relevant, acceptable, and effective by the individuals involved; it concerns the social significance or relevance of intervention goals and is traditionally conceptualized as acceptability [26]. Acceptability is a core concept in digital health in part due to its ability to predict and explain key outcomes of interest [42]. Table 3 describes both participants' and health care providers' feedback on the specific intervention features and potential future uses of the intervention. Participants and providers were generally close in their assessment of how important or useful the features of the intervention were. One exception was the "resources" feature. Overall, 67% (89 of 132 participants who responded to this item) rated the resources feature as useful; however, this includes the 28% (37 of 132) who reported not using this feature. For those participants who used this feature, 94% (89 of 95) rated the resource feature as useful.

XSL•FO
RenderX

Generalizability

Generalizability assesses the extent to which the intervention procedures generalize to non-treatment settings, over time, and with diverse samples. While the TLC study was limited to dementia caregivers, the intervention was developed with an eye toward any caregiver. Respite providers were asked, "Do you think that the TLC app could be used with all types of caregivers?" As shown in Table 3, overwhelmingly, 55 of 57 providers answered that the TLC intervention could, or possibly could, be used for all types of caregivers (75% yes, 21% maybe).

Table 3 shows that 127 of 163 (78%) participants would consider using the TLC website in the future. Reasons for future use varied, and in fact, many participants pointed to the variety of uses they got from the app, as shown by one participant who stated, "I like to have a framework to help me refocus from what my daily routine is. I like the resources section. I feel like it would be nice to go back in and once again review those resources. With time, some types of respite change and new ideas are welcomed." Others mentioned the usefulness of the routine or the "Discipline in keeping track of set goals for caregiver on a daily basis." Another participant noted, "Again, more than the actual schedule (because I tend to use a paper calendar and daily written to-do lists), the reminder each week to take time for personal self-care was really valuable. It kept me on track. I'm pretty ADD and routines are hard, so the accountability of the website kept me constantly reminded and responsible."

Often, for those who did not think they would use TLC in the future, the initial intervention was seen as having accomplished its goals or was beneficial in creating awareness of the need for specific goals. One participant noted, "It has already helped me plan ahead and think of what I will be doing. I don't think I need that assistance anymore but stand to be corrected!" This indicates that future use would be possible if they felt they had stopped planning and taking time for respite. Others may have taken some of the intervention goals, such as being specific, and incorporated them into other systems - "I have an online calendar where everything in my life (and my mom's) is tracked - this would duplicate that effort. BUT what I did like about the site was that I was specific as to what the respite was scheduled for - I don't have that level of detail on my calendar."

Integration

Assessing integration concerns how aligned the intervention is with existing practice and infrastructure. Table 3 shows that 148 of 163 participants and 51 of 57 providers (~90% of each group) would recommend the TLC app to caregivers. However, the method of delivery may be individually and situationally based. In total, 78% (103/132) indicated that they would use TLC optimized for a smartphone. Providers were asked how they preferred to deliver support to caregivers, including as an app or interactive website, workbooks, or worksheets (fillable, printable, pencil and paper), in-person personal coaching, or personal coaching using telephone or video conference. A self-guided app or interactive website and in-person personal coaching tied for first, with 17 of 57 providers (30%) ranking each as first choice. Table 4 provides exemplars of why participants would or would not use the TLC app optimized for a smartphone. While most participants would use TLC as a smartphone app, which would cover those who have it as a preferred method, the ability to use different preferred devices is preferable overall.

Implementation

As noted earlier, providers indicated that they would be willing to recommend the intervention to current and future respite clients, but another critical question is whether the intervention could be implemented by providers in the same manner as the research team. As an online web-based intervention, TLC was implemented with high fidelity (eg, as intended by the developers and researchers; the same messaging and imagery were delivered to participants based on the algorithms underlying the software). A small fraction of participants (<10%) accessed technological support from TLC research staff; this kind of human interaction and support would likely not be available if the app were implemented outside of the research study environment.

Effectiveness

Table 5 describes the metrics we used to identify whether the TLC intervention brought about positive change for the participants. Participants reported an increase in the amount of respite received each week, from an average of 8 hours at the start of the intervention to an average of 13 hours at the end of the intervention. The percentage of participants who felt they were getting enough respite time increased from 10% to 42% (16 participants to 68 participants). Participants were more likely to schedule their respite hours (58% vs 90%) and were happier with what they did during respite (72 or 44% vs 119 or 73%) after the intervention, compared to before. As noted by a participant, "TLC helps me focus on what I want to do with my time. It helps me plan to do things that really nurture me." There were no differences in participants' average depression or burden scores, but there was a slight increase in their perception of positive aspects of caregiving. Additionally, the average scores do not preclude individual benefits that may have been gained by some, as noted by a participant who stated, "Having a goal of respite each week has been helpful and has reduced some of my anxiety."

XSL•FO
RenderX

Table . Measures of intervention effectiveness.

Measure[a]	Pre-intervention	Post-intervention
Caregiver and respite outcomes		
Amount of respite per week (in *hours*), mean (SD)	8 (11.5)	13 (11.7)
I am getting enough respite time, %	10	42
My respite is scheduled in advance, %	58	90
I am happy with what I did during respite, %	44	73
Depression (T-score scale range 37.1 - 81.1), mean (SD) [39]	54.1 (7.4)	53.6 (7.8)
Burden (scale range 0 - 96), mean (SD) [43]	46 (16.6)	45.9 (17)
Positive Aspects of Caring (scale range 9 - 45), mean (SD) [44]	25.5 (9.7)	28.4 (9.4)
Perceived benefits of intervention		
Using the TLC intervention was a positive experience for me, %	—	85
TLC intervention helped me improve my respite time to moderate, %	—	83
I feel happier, healthier, or less stressed after using TLC intervention, %	—	92

[a]All measures were self-reported by caregivers who had used the TLC intervention. The sample used to calculate descriptive statistics for this table was restricted to those who had data at both pre- and post- assessment (depression n=163; burden n=155; positive aspects n=159; perceived benefits n=132).

Overall, of 132 participants who responded to perceived benefits items, 112 (85%) rated their experience using the TLC intervention as positive or very positive. Only 1% (2 of 132) reported a negative experience, while 14% (18 of 132) felt neutral. Participants reported being able to identify times that they may not have considered before. As one participant stated, “I am learning that scheduling respite ahead of time is very helpful. I am actually able to schedule more time than I realized and still give the care needed – we are both happier.” Another participant noted that they were able to focus on using the time available, “Respite becomes more valuable when the time and activities are scheduled. Even if the activity is to disappear and do nothing. Without a respite schedule, the time gets squandered.” The TLC intervention also helped provide validation for some. As one participant stated, “Budgeting time to myself is paramount to being a good caregiver. Seeing the planned times on the calendar gives me value as a person of worth.” Another participant stated, “Honestly, I am taken back about how just calendaring time out has made such a difference – thank you!”

At the end of the 16-week intervention period, a majority of participants (134/159, or 83%) reported that they felt the intervention improved their respite time use, and more than 9 out of 10 participants (92%) reported that using the intervention made them happier, healthier, or less stressed. As one participant stated, “I like that I schedule my respite. I like that the time I do have to myself or the time I scheduled is labeled respite. It makes me appreciate that I really do have carved out time to stop the world and focus on other things I love to do.” The intervention process was noted by another who found, “The process of scheduling time off (and making arrangements to be able to take time off) has been a real boon!”

Discussion

Principal Results

Using a comprehensive mixed-methods approach to feasibility analysis [38], we used pilot data from the TLC study to evaluate a respite-focused caregiver intervention in terms of 10 conceptually distinct measures of feasibility, as outlined by Gadke et al [26] (see Table 1 for definitions of feasibility). Overall, across each domain, we found strong indicators and evidence of feasibility for the TLC intervention. Thus, we conclude that the TLC intervention is feasible. This feedback also provides a glimpse into the potential or promise of the intervention in supporting family caregivers in their use of respite. Providing caregivers with coaching and resources to recognize the value of respite and to maximize the benefit of whatever respite time they have by using goal-setting and goal-review techniques has the potential to provide meaningful support for family caregivers. This respite-focused intervention may be particularly relevant for respite service providers to ensure that their clients are receiving maximum benefit from their respite services (which oftentimes have long waitlists meaning that respite services are fairly limited), as well as for policy makers and caregiver advocates who may need further empirical evidence showing the benefits of respite to advocate for increased funding for respite.

This analysis also provides a case study that can serve as a model for other researchers who may be developing interventions and the study designs that will be used to evaluate the feasibility of the new intervention. Adopting such a comprehensive approach to feasibility assessment allowed us to design a pilot study that included multiple measures of

XSL•FO
RenderX

feasibility, rather than only an isolated set of measures tacked on as an afterthought to a study that focused primarily on the measurement of primary outcomes. For example, we decided to use a fully powered sample for this pilot study because it would allow us to get a better sense of initial efficacy and identification of mechanisms that underlie the intervention's effect. Additionally, by choosing to work with a community advisory board representing the diverse caregiving experiences and multiple cultural perspectives, we were able to make significant refinements during the initial TLC technology build, rather than having to wait until the pilot study was completed to learn about what those refinements should be. Doing a systematic and iterative development process where we co-created the intervention with a community advisory board did cost money and took time, but it provided richer data and a more refined TLC intervention for the pilot study than would have been available without that initial investment and commitment to developing a feasible product from the start.

Pilot studies, although focused primarily on issues of feasibility, can and should be as comprehensive as possible so that we can develop interventions that have the potential to be implemented beyond small-sample pilot studies. By using the Gadke et al [26] 10-dimension perspective, we developed a rich and comprehensive set of data collection instruments and metrics that we could use to assess the complexity of feasibility. For example, in response to the "generalizability" dimension of feasibility, we decided to include a separate data collection effort with a nationwide sample of respite providers. While this additional step added time and effort to the TLC study, it provided us with information and insights about the possible implementation of TLC. Relatedly, intervention research should also engage in cost analyses, assessing whether the benefits of an intervention have measurable cost savings. We assessed how much the intervention cost us to build but have not yet assessed whether using the intervention saved participants or providers money. Future research should include methodology and metrics to assess the cost-effectiveness of interventions [45].

While there are no costs associated with the delivery of the intervention once the web platform is built, there are significant costs associated with the initial build, as well as the refinements that are needed at each stage of evaluation and overall maintenance and general updates and bug fixes. Navigating through the stage-model for intervention development [21]—where each stage may take several years to complete—may be too slow to create an effective intervention, especially for technology-based interventions where the style and functionality change very quickly with the rapid introduction of new apps and features available to the consumer or public. This consideration may help explain the differences for those who reported a neutral experience using TLC, but also reported *any* improvement (from slightly to extremely) in their feeling happier, healthier, or less stressed as a caregiver after using TLC. The style and functionality may have impacted their experience, even though they were able to still see some benefit from the intervention.

Conclusions

Overall, these results suggest that TLC should be further refined and move toward later stages of intervention research, where the most essential features of the intervention can be implemented in less controlled research environments. Based on these results, combined with findings that TLC improves caregiver well-being [34], TLC is likely ready for a large-scale implementation trial "to test real-world transferability and effectiveness" (p. 33) [20]. Moving toward this goal, the key features of the TLC intervention have been made available on a publicly accessible webpage for caregivers to access and use for free [35]. A major goal of this feasibility analysis was to assess whether a fully self-administered, web-based intervention is feasible for family caregivers, who tend to be among older generations that have less computer proficiency and access to internet and internet-capable devices [20]. The TLC pilot study found that there is certainly a subgroup of older caregivers who are responsive to and seeking this kind of support. This helps us shed the outdated tropes that older adults do not want to access support and resources online [46-48].

Acknowledgments

This study was funded by the National Institute on Aging of the National Institutes of Health (R01 AG061946).

Data Availability

De-identified data can be found on the HIVE archive at the University of Utah [35].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Comparison of TLC sample to national profile of family caregivers.
[DOCX File, 17 KB - aging_v8i1e71792_app1.docx]

References

1. National Alliance for Caregiving. Caregiving in the U.S. 2020. AARP 2020. [doi: 10.26419/ppi.00103.001]

XSL•FO
RenderX

2. Reinhard SC, Caldera S, Houser A, Choula R. Valuing the invaluable: 2023 update. : AARP Public Policy Institute; 2023. [doi: 10.26419/ppi.00082.006]
3. Adelman RD, Tmanova LL, Delgado D, Dion S, Lachs MS. Caregiver burden: a clinical review. JAMA 2014 Mar 12;311(10):1052-1060. [doi: 10.1001/jama.2014.304] [Medline: 24618967]
4. Bialon LN, Coke S. A study on caregiver burden: stressors, challenges, and possible solutions. Am J Hosp Palliat Care 2012 May;29(3):210-218. [doi: 10.1177/1049909111416494] [Medline: 21803785]
5. Administration for Community Living. 2022 National strategy to support family caregivers. 2022. URL: http://acl.gov/CaregiverStrategy [accessed 2025-01-09]
6. Schulz R, editor. Handbook on Dementia Caregiving: Evidence-Based Interventions for Family Caregivers: Springer Publishing Company; 2000.
7. Schulz R, Burgio L, Burns R, et al. Resources for Enhancing Alzheimer's Caregiver Health (REACH): overview, site-specific outcomes, and future directions. Gerontologist 2003 Aug;43(4):514-520. [doi: 10.1093/geront/43.4.514] [Medline: 12937330]
8. Cameron JI. Family caregiving research: Reflecting on the past to inform the future. J Spinal Cord Med 2021;44(sup1):S19-S22. [doi: 10.1080/10790268.2021.1970883] [Medline: 34779725]
9. The Benjamin Rose Institute on Aging. Evidence-based and informed programs.: Ben Rose Institute on Aging URL: https://institute.benrose.org/evidence-based-and-informed-programs/ [accessed 2025-01-09]
10. Whitmore KE. The concept of respite care. Nurs Forum 2017 Jul;52(3):180-187. [doi: 10.1111/nuf.12179] [Medline: 27438939]
11. Zarit SH, Bangerter LR, Liu Y, Rovine MJ. Exploring the benefits of respite services to family caregivers: methodological issues and current findings. Aging Ment Health 2017 Mar;21(3):224-231. [doi: 10.1080/13607863.2015.1128881] [Medline: 26729467]
12. Vandepitte S, Van Den Noortgate N, Putman K, Verhaeghe S, Verdonck C, Annemans L. Effectiveness of respite care in supporting informal caregivers of persons with dementia: a systematic review. Int J Geriatr Psychiatry 2016 Dec;31(12):1277-1288. [doi: 10.1002/gps.4504] [Medline: 27245986]
13. Lund DA, Utz R, Caserta MS, Wright SD. Examining what caregivers do during respite time to make respite more effective. J Appl Gerontol 2009 Feb;28(1):109-131. [doi: 10.1177/0733464808323448]
14. Lund DA, Utz RL, Caserta MS, et al. Time for living and caring: an intervention to make respite more effective for caregivers. Int J Aging Hum Dev 2014;79(2):157-178. [doi: 10.2190/AG.79.2.d] [Medline: 25536704]
15. Baltes PB, Baltes MM. Psychological perspectives on successful aging: the model of selective optimization with compensation. In: Baltes PB, Baltes MM, editors. Successful Aging: Perspectives from the Behavioral Sciences: New York, NY: Cambridge University Press; 1990:1-34.
16. Utz RL, Caserta M, Iacob E, et al. Maximizing the benefit of respite for dementia caregivers: a study protocol describing the development & evaluation of the Time for Living & Caring (TLC) intervention. OBM Integr Compliment Med 2023;8(4):040. [doi: 10.21926/obm.icm.2304040] [Medline: 38313766]
17. Lorca-Cabrera J, Grau C, Martí-Arques R, Raigal-Aran L, Falcó-Pegueroles A, Albacar-Riobóo N. Effectiveness of health web-based and mobile app-based interventions designed to improve informal caregiver's well-being and quality of life: a systematic review. Int J Med Inform 2020 Feb;134:104003. [doi: 10.1016/j.ijmedinf.2019.104003] [Medline: 31790857]
18. de-Moraes-Ribeiro FE, Moreno-Cámara S, da-Silva-Domingues H, Palomino-Moral P, Del-Pino-Casado R. Effectiveness of internet-based or mobile app interventions for family caregivers of older adults with dementia: a systematic review. Healthcare (Basel) 2024 Jul 27;12(15):1494. [doi: 10.3390/healthcare12151494] [Medline: 39120195]
19. Hopwood J, Walker N, McDonagh L, et al. Internet-based interventions aimed at supporting family caregivers of people with dementia: systematic review. J Med Internet Res 2018 Jun 12;20(6):e216. [doi: 10.2196/jmir.9548] [Medline: 29895512]
20. Vassilakopoulou P, Hustad E. Bridging digital divides: a literature review and research agenda for information systems research. Inf Syst Front 2023;25(3):955-969. [doi: 10.1007/s10796-020-10096-3] [Medline: 33424421]
21. Onken LS. History and evolution of the NIH stage model. In: Evidence-Based Practice in Action: Bridging Clinical Science and Intervention: New York, NY: Guilford Publications; 2019.
22. Prusaczyk B, Bobitt J. Implementation science: a critical tool for research utilization and policy evaluation. Public Policy Aging Rep 2022 Jan 31;32(1):4-5. [doi: 10.1093/ppar/prab035]
23. O'Cathain A, Croot L, Sworn K, et al. Taxonomy of approaches to developing interventions to improve health: a systematic methods overview. Pilot Feasibility Stud 2019;5(1):41. [doi: 10.1186/s40814-019-0425-6] [Medline: 30923626]
24. Utz RL, Terrill AL, Thompson A. Online interventions to support family caregivers: the value of community-engaged research practices. J Prev Interv Community 2023;51(3):238-253. [doi: 10.1080/10852352.2021.1930817] [Medline: 34053412]
25. Wight D, Wimbush E, Jepson R, Doi L. Six steps in quality intervention development (6SQuID). J Epidemiol Community Health 2016 May;70(5):520-525. [doi: 10.1136/jech-2015-205952] [Medline: 26573236]
26. Gadke DL, Kratochwill TR, Gettinger M. Incorporating feasibility protocols in intervention research. J Sch Psychol 2021 Feb;84:1-18. [doi: 10.1016/j.jsp.2020.11.004] [Medline: 33581765]

XSL•FO
RenderX

27. Xavier A, Vagos P, Palmeira L, et al. The Me and the Us of Emotions: a cluster-randomized controlled trial of the feasibility and efficacy of a compassion-based social-emotional learning program for children. Front Psychol 2023;14:1196457. [doi: 10.3389/fpsyg.2023.1196457] [Medline: 38023010]
28. Parés-Salomón I, Vaqué-Crusellas C, Coffey A, et al. Development of digital strategies for reducing sedentary behavior in a hybrid office environment: modified Delphi study. JMIR Hum Factors 2025 Apr 8;12(1):e59405. [doi: 10.2196/59405] [Medline: 40198908]
29. Harris-Gersten ML, Davagnino JM, Alcorn ER, Hastings SN. Usability and acceptability of social robot pets among community-dwelling veterans living with dementia and their caregivers. Am J Alzheimers Dis Other Demen 2023;38:15333175231200973. [doi: 10.1177/15333175231200973] [Medline: 37688378]
30. Thabane L, Cambon L, Potvin L, et al. Population health intervention research: what is the place for pilot studies? Trials 2019 May 30;20(1):309. [doi: 10.1186/s13063-019-3422-4] [Medline: 31146768]
31. Chen HT. The bottom-up approach to integrative validity: a new perspective for program evaluation. Eval Program Plann 2010 Aug;33(3):205-214. [doi: 10.1016/j.evalprogplan.2009.10.002] [Medline: 19931908]
32. Hoddinott P. A new era for intervention development studies. Pilot Feasibility Stud 2015;1(1):36. [doi: 10.1186/s40814-015-0032-0] [Medline: 27965814]
33. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan 30;17(1):e30. [doi: 10.2196/jmir.4055] [Medline: 25639757]
34. Iacob E, Caserta M, Donaldson G, et al. Evaluating the efficacy of time for living and caring: an online intervention to support dementia caregivers' use of respite. Innov Aging 2024;8(5):igae043. [doi: 10.1093/geroni/igae043] [Medline: 38803611]
35. Utz R. Time for living & caring (TLC) study: virtual coaching to maximize dementia caregivers' respite time-use. . 2024. [doi: 10.7278/S50D-2RGG-4549]
36. Time for Living and Caring-TLC. 2025. URL: www.tlc-respite.com [accessed 2025-01-11]
37. Utz R. Virtual coaching to maximize dementia caregivers respite time-use: a stage 1 pilot test for feasibility and efficacy. clinicaltrials.gov. 2024. URL: https://clinicaltrials.gov/study/NCT03689179 [accessed 2025-01-11]
38. Aschbrenner KA, Kruse G, Gallo JJ, Plano Clark VL. Applying mixed methods to pilot feasibility studies to inform intervention trials. Pilot Feasibility Stud 2022 Sep 26;8(1):217. [doi: 10.1186/s40814-022-01178-x] [Medline: 36163045]
39. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. J Clin Epidemiol 2010 Nov;63(11):1179-1194. [doi: 10.1016/j.jclinepi.2010.04.011] [Medline: 20685078]
40. Sparks C, Hsu A, Neller SA, et al. Comparison of recruitment methodologies for clinical trials: Results from the time for living and caring (TLC) intervention study. Contemp Clin Trials 2024 May;140:107518. [doi: 10.1016/j.cct.2024.107518] [Medline: 38554816]
41. Boot WR, Charness N, Czaja SJ, et al. Computer proficiency questionnaire: assessing low and high computer proficient seniors. Gerontologist 2015 Jun;55(3):404-411. [doi: 10.1093/geront/gnt117] [Medline: 24107443]
42. Perski O, Short CE. Acceptability of digital health interventions: embracing the complexity. Transl Behav Med 2021 Jul 29;11(7):1473-1480. [doi: 10.1093/tbm/ibab048] [Medline: 33963864]
43. Novak M, Guest C. Application of a multidimensional caregiver burden inventory. Gerontologist 1989 Dec;29(6):798-803. [doi: 10.1093/geront/29.6.798] [Medline: 2516000]
44. Cohen CA, Gold DP, Shulman KI, Zucchero CA. Positive aspects in caregiving: an overlooked variable in research. Can J Aging 1994;13(3):378-391. [doi: 10.1017/S071498080000619X]
45. Murray CJ, Evans DB, Acharya A, Baltussen RM. Development of WHO guidelines on generalized cost-effectiveness analysis. Health Econ 2000 Apr;9(3):235-251. [doi: 10.1002/(sici)1099-1050(200004)9:3<235::aid-hec502>3.0.co;2-o] [Medline: 10790702]
46. Coughlin JF. The Longevity Economy: Unlocking the World's Fastest-Growing, Most Misunderstood Market: PublicAffairs; 2017.
47. Anderson M, Perrin A. Tech adoption climbs among older adults pew research center: internet, science & tech. 2017 May 17 URL: https://www.pewresearch.org/internet/2017/05/17/tech-adoption-climbs-among-older-adults/ [accessed 2023-04-26]
48. Berkowsky RW, Sharit J, Czaja SJ. Factors predicting decisions about technology adoption among older adults. Innovation in Aging 2017 Nov 1;1(3):igy002. [doi: 10.1093/geroni/igy002]

Abbreviations

TLC: Time for Living and Caring

XSL•FO
RenderX

Edited by N Schütz; submitted 26.01.25; peer-reviewed by A Takaoka, E Kent; revised version received 20.05.25; accepted 10.06.25; published 18.08.25.

Please cite as:
Thompson AD, Terrill AL, Caserta M, Wong B, Iacob E, Sparks C, Stark L, Utz RL
Assessing the Time for Living and Caring (TLC) Study: Mixed-Methods Feasibility Study of a Web-Based Caregiver Intervention to Improve Respite
JMIR Aging 2025;8:e71792
URL: https://aging.jmir.org/2025/1/e71792
doi:10.2196/71792

© Amber D Thompson, Alexandra L Terrill, Michael Caserta, Bob Wong, Eli Iacob, Catharine Sparks, Louisa Stark, Rebecca L Utz. Originally published in JMIR Aging (https://aging.jmir.org), 18.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Medicaid Education and Eligibility Planning for Caregivers: Website Usability and Validation Study

Marguerite DeLiema[1], PhD; Siyu Gao[1], MSW; Justine Scattarelli[2], MA; Kelly Moeller[3], MPH; Olu Olofinboba[4], MS, MBA

[1]School of Social Work, University of Minnesota, 1404 Gortner Ave., Saint Paul, MN, United States

[2]Department of Educational Psychology, University of Minnesota, Minneapolis, MN, United States

[3]School of Public Health, University of Minnesota, Minneapolis, MN, United States

[4]Moai Technologies (United States), Minneapolis, MN, United States

Corresponding Author:
Marguerite DeLiema, PhD
School of Social Work, University of Minnesota, 1404 Gortner Ave., Saint Paul, MN, United States

Abstract

Background: Most older Americans have not saved enough to cover long-term care costs. Medicaid—a public health care program for low-income individuals—can help Americans with qualifying care needs pay for assistance in a nursing home or for services in the home. Determining financial eligibility for Medicaid is complicated, and the application process is often managed by family caregivers with limited knowledge of Medicaid programs.

Objective: A one-stop digital solution is needed to help family caregivers plan for the cost of long-term care services and learn about getting help paying for services through Medicaid. We aimed to develop a web application that (1) educates informal caregivers about Medicaid programs and eligibility criteria, (2) informs them about the cost of home and institutional care in their local area with and without Medicaid coverage, and (3) uses a custom algorithm to provide personalized financial eligibility information based on the care recipient's income, assets, and monthly spending.

Methods: We first interviewed aging services providers and informal family caregivers, then developed a web application that was refined based on user experience interviews with English- and Spanish-speaking caregivers. In the final validation phase, asynchronous usability sessions were recorded with 109 informal caregivers who completed a series of tasks. Participants viewed and rated animated Medicaid "explainer" videos, input financial information to enable the custom algorithm to determine the care recipient's eligibility for Medicaid, adjusted settings on a care cost calculator to estimate the regional cost of home and institutional care services, and completed a Medicaid knowledge quiz before and after using the website.

Results: After engaging with the website and watching the videos, scores on a Medicaid knowledge quiz increased by 61.2% (2-tailed t_{92}=12.9, P<.001). Participants found it easy to enter the care recipient's financial information to determine Medicaid eligibility (out of 7; mean 5.9, SD 1.3) and perceived the care cost calculator as very helpful (out of 7; mean 6.3, SD 1.2). The website received a very high System Usability Scale rating of 88.3 out of 100 (SD 13.1). Caregivers verbalized wanting more education on complex financial concepts that impact Medicaid eligibility and asset preservation.

Conclusions: A comprehensive Medicaid planning website can significantly improve caregivers' knowledge of Medicaid and provide them with a personalized roadmap for accessing care services. The custom algorithm powering the Medicaid eligibility determination could be further refined to account for state-based exceptions. This application may reduce caregiver burden and help support the long-term care planning process.

(JMIR Aging 2025;8:e77441) doi:10.2196/77441

KEYWORDS

long-term services and supports; long-term care; Medicaid waiver; user-centered design; cost of care; financial education; Medicaid planning; caregiver education; web application

Introduction

Background

Approximately 70% of Americans need long-term services and supports (LTSS) as they age, and 48% use some form of paid care such as skilled nursing, assisted living, home health, and adult day services [1]. The likelihood of needing LTSS and the overall cost are much greater for those living with Alzheimer disease and related dementias (ADRD) [2-4], which affects approximately 1 in 10 adults aged 65 years and older [5]. Based on a 2024 survey of long-term care providers in the United States, the median cost of care in a skilled nursing facility

XSL•FO
RenderX

exceeds US $110,000 per year, whereas assisted living facilities cost approximately US $70,000, and a 44-hour per week home health aide is US $78,000 [6].

Most older adults have not saved enough to cover these high costs, and only 11% have purchased private long-term care insurance [7,8]. In a national survey, Hamel and Montero [7] found that only 23% of adults aged 50-64 years reported that they set aside money to pay for future expenses, and 83% stated that it would be impossible or very difficult to pay for a home health aide or a 1-year stay in a nursing home. As a result, most Americans rely on informal (unpaid) family and friend caregivers until their support needs reach a point when formal LTSS is needed [9].

Medicaid as a Primary Payer of LTSS

Due to the high cost of formal LTSS, many Americans turn to Medicaid, a joint state and federal health insurance program for low-income individuals [10]. Medicaid is the nation's primary payer of LTSS; approximately one-third of total program costs go toward LTSS [11]. In 2020, 62% of nursing home residents were covered by Medicaid [12]. Traditionally, Medicaid reimburses institutional care provided in skilled nursing facilities, but following the 1981 Omnibus Budget Reconciliation Act, Medicaid waivers allowed states to expand coverage to home and community-based services (HCBS), such as home health and personal care aides, case management, caregiver respite, and adult day services. In 2022, 87% of Medicaid LTSS users received HCBS [13].

Unfortunately, many older adults struggle to access Medicaid benefits. For one, Americans hold misconceptions about LTSS insurance coverage; 41% falsely believe that Medicare, the health insurance program for adults aged 65 years and older, covers long-term stays in a nursing home [7]. Older adults may underestimate the cost of residential and home care services or believe they can rely on family members for assistance [14,15]. These assumptions can lead eligible older adults to delay Medicaid enrollment.

Navigating formal LTSS options is a high-stakes, multistep process that often includes touring facilities, interviewing home health aides, calculating costs, and determining whether the care recipient is eligible for public benefits. Because of physical and cognitive impairments, these activities often fall on informal family caregivers who are encountering public benefit programs for the first time. The decision to apply for Medicaid may be triggered by a health crisis, compounding the emotional and cognitive burden of caregiving [15]. When helping the older adult apply for Medicaid, informal caregivers encounter unfamiliar terms like "spousal resource allowance" (the amount of household assets a Medicaid applicant's spouse is allowed to keep) and "asset spend down" (the process of drawing down one's savings and investments to meet eligibility limits).

State-to-state variability in Medicaid eligibility criteria is an additional challenge to understanding who qualifies. For example, in states such as Iowa, Maine, and New Mexico, an individual applying for Medicaid must have less than US $2000 in assets; in New York, the asset limit is US $32,396. California eliminated the asset limit in 2024 [16]. Some states offer exceptions for certain types of assets, and many provide alternative pathways to qualify when assets or income exceed state limits. This lack of standardization in Medicaid policy may cause confusion, especially among informal caregivers who reside in a different state than the care recipient they are helping to enroll.

The demands of the Medicaid application process itself may serve as a significant barrier to enrollment. Moynihan et al [17] argue that "by constructing complex, confusing, and time-consuming application procedures, the state can effectively thwart an individual from accessing benefits, even if eligible by law." Several factors must come together for the application to be processed successfully. First, the care recipient must undergo a functional needs assessment that is scheduled with and administered by the proper local authority. They must submit recent financial and insurance documentation to the state or county. This information can be difficult to obtain by informal caregivers who may lack the legal authority to access accounts. Also, most states have a "look-back" period of 60 months (5 y) that penalizes applicants who gifted or transferred assets to family, friends, or charities before applying, further delaying the eligibility [16].

Researchers have called for digital solutions ("digital nudging") and more automation to reduce information overload and ease the cognitive burden of applying for government programs [18,19]. Digital nudging involves creating user-friendly interface designs that guide behavior in digital choice environments [20]. Effective digital nudges can help caregivers make better choices by simplifying their choice sets; for example, restricting the list of formal LTSS services to those that accept Medicaid clients or only presenting information on spousal resource allowances if the care recipient is married.

Systematic reviews suggest that informal family caregivers are motivated to use technology to help manage caregiving challenges [21-24]. Web and mobile applications have been developed to improve caregiver well-being and resilience [25], motivate physical activity [26], coordinate care [27], and engage with peer support [28], but to date, there is no one-stop solution designed for older adults and their caregivers to get personalized information on offsetting the high cost of institutional LTSS or HCBS by enrolling in Medicaid.

Study Objective

To address this resource gap, the goal of this web application is to (1) educate informal caregivers about Medicaid programs and eligibility criteria, (2) improve caregivers' understanding of the local costs of LTSS with and without Medicaid coverage, and (3) provide personalized eligibility timelines based on the applicant's state of residence, income, assets, and monthly spending. In this study, we present the findings from the asynchronous mixed methods website validation test with 109 current and former older adult caregivers. We assessed their Medicaid program and eligibility knowledge before and after interacting with the website in a pretest-posttest design.

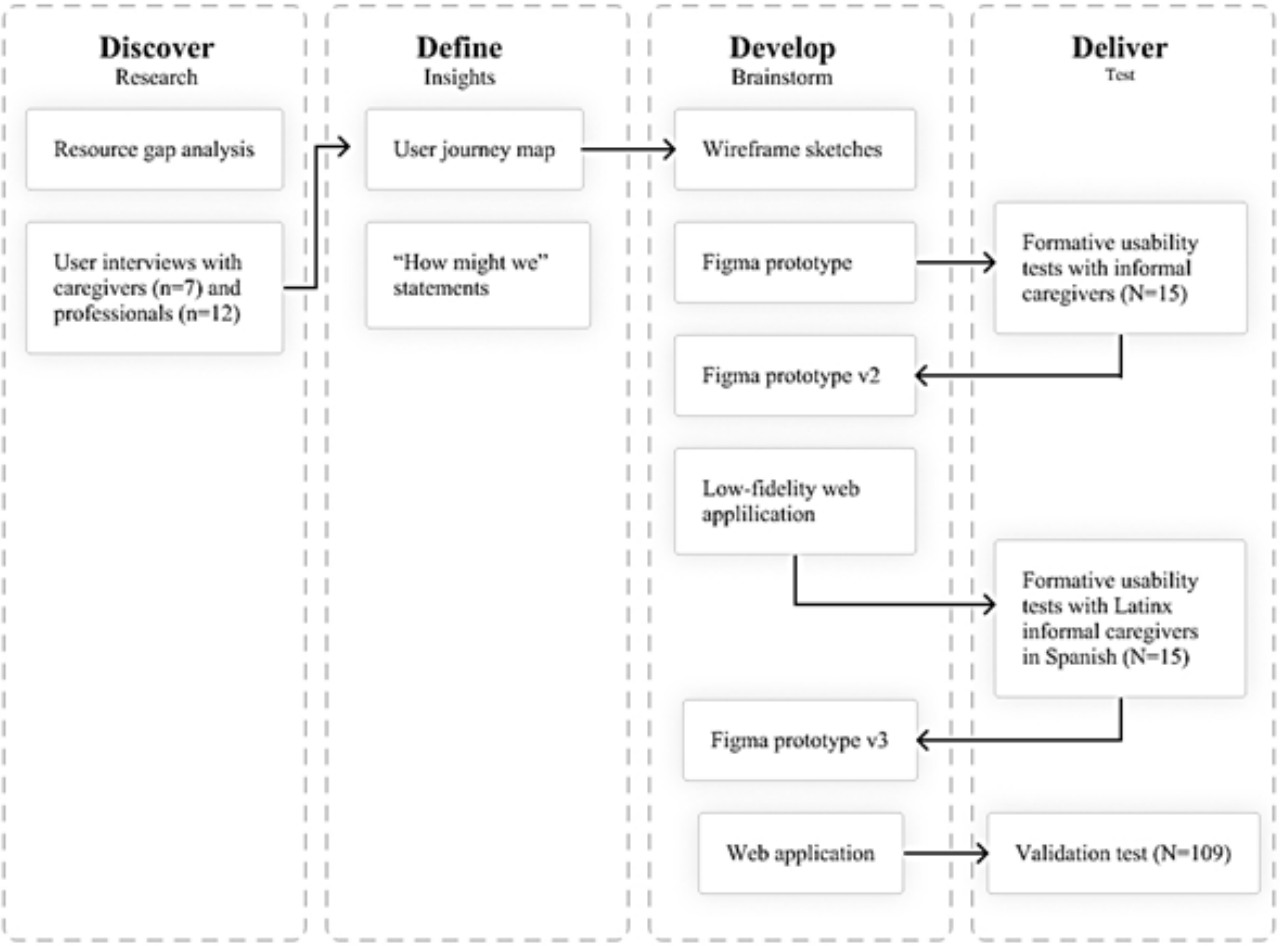
XSL•FO
RenderX

Methods

Development of the Web Application

Website development (Figure 1) followed the Double Diamond design-thinking process, which is a user-centered approach to problem solving comprising 4 iterative phases: discover, define, develop, and deliver [29]. As part of the "discover" stage, the research team interviewed informal caregivers and aging services providers (eg, certified Medicaid planners, county financial workers, geriatric care managers, and social workers) to better understand the challenges of navigating LTSS options, costs, and the process of applying for Medicaid and Medicaid waivers [15]. Involving informal caregivers and direct service providers in the early stages of design is critical to ensuring the end product addresses their stated needs [22].

Figure 1. Research and web application development steps.



In the "define" stage, a larger team of researchers, developers, and designers participated in a synchronous problem definition session and created a user journey map and user personas (Multimedia Appendix 1) to visualize the process of accessing and financing LTSS, as uncovered by our initial research. The team wrote "how might we" statements to further define the problem and begin to conceptualize how a digital tool could support caregivers (eg, "How might we inform caregivers that home care services are an option?").

Alternating "design" and "develop" activities followed. First, team members sketched wireframes of possible solutions and voted on an initial concept. Designers then built a prototype using Figma, a collaborative design software. Researchers tested this first iteration with 15 informal caregivers for older adults with ADRD. The 90-minute formative usability tests were conducted via Zoom (Zoom Communications). Based on these findings, a second iteration of the design was created to include larger information displays, a sidebar menu, and cursor hover capabilities. This version was developed into a dual-language (English and Spanish) web application and was tested by 15 Spanish-speaking informal caregivers. These formative usability tests were moderated by Spanish-speaking interviewers, also using Zoom.

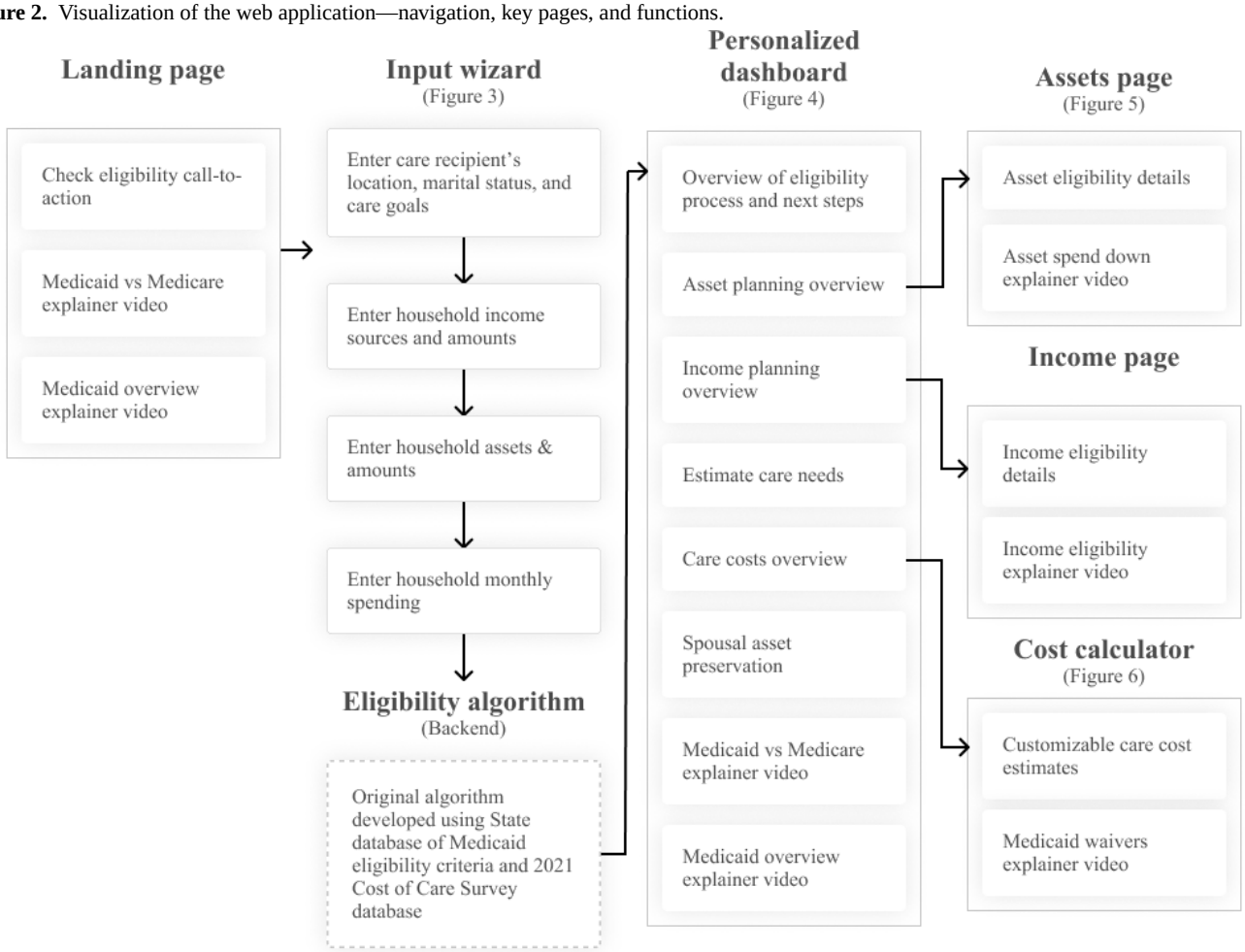
Overview of Web Application

Participant feedback from both user experience test phases was incorporated into a final design. A high-level representation of the final flow of the web application is presented in Figure 2. The landing page welcomes users with a large image and the message: "Get help paying for long-term care. If you don't qualify now, you might in the future. We'll help you get there." From the landing page, users can click the "Check Eligibility" button that takes them through an input wizard where they enter information about the care recipient (ie, applicant): name, location, marital status, and care goals (independent, assisted living, and nursing home). The user is then asked to select the types and enter values (in US $) of the care recipient's income sources (Figure 3), assets, and monthly expenses. The input

XSL•FO
RenderX

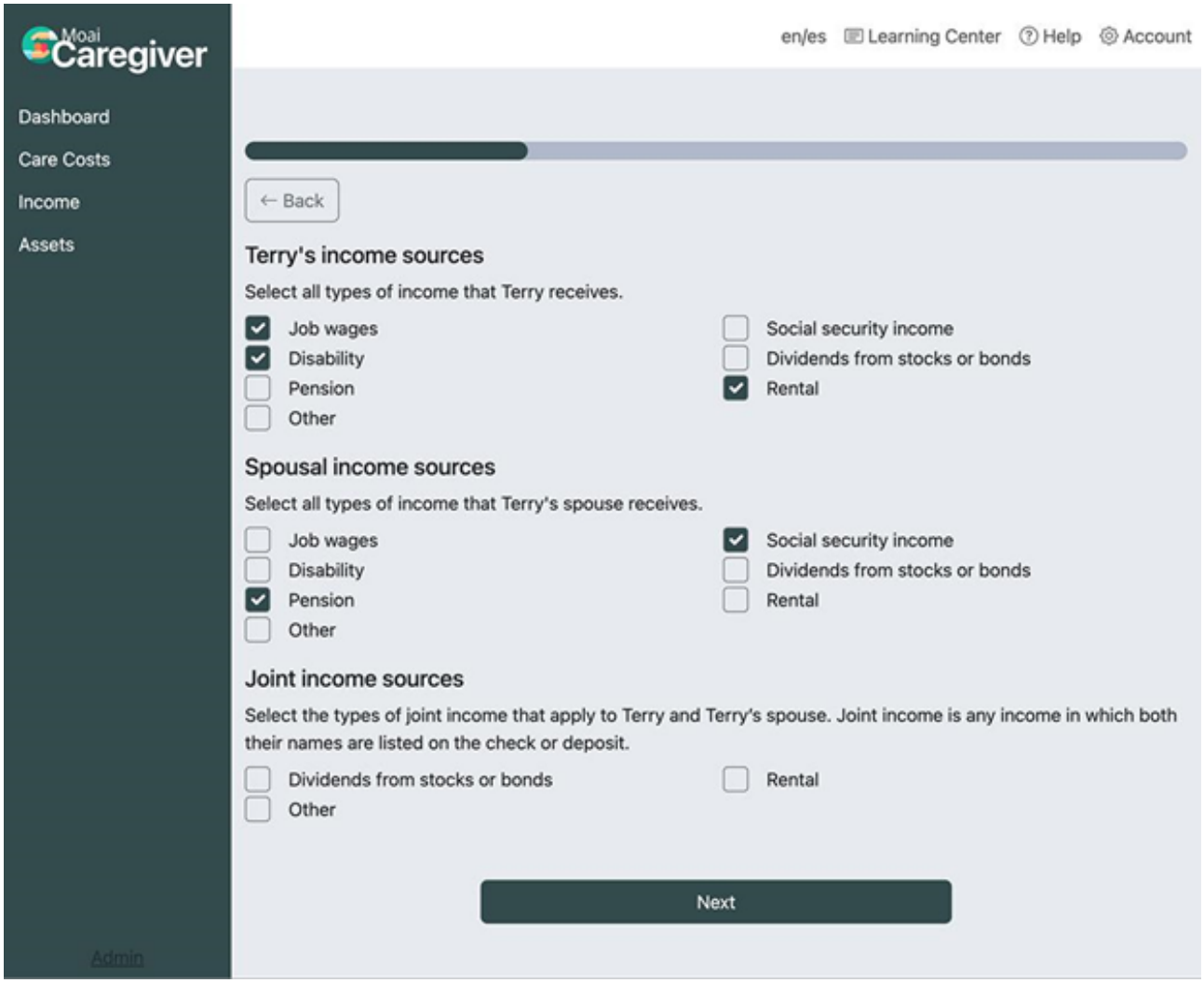
wizard collects information about the applicant's income and assets separately from their spouse's (if applicable) because nearly all states exempt spousal income sources from the applicant's eligibility determination. In some states, such as Delaware and Wisconsin, the spouse's retirement accounts (Individual Retirement Accounts and 401k) are also exempt.

Figure 2. Visualization of the web application—navigation, key pages, and functions.



XSL•FO
RenderX

Figure 3. The “select income sources” step of the input wizard.

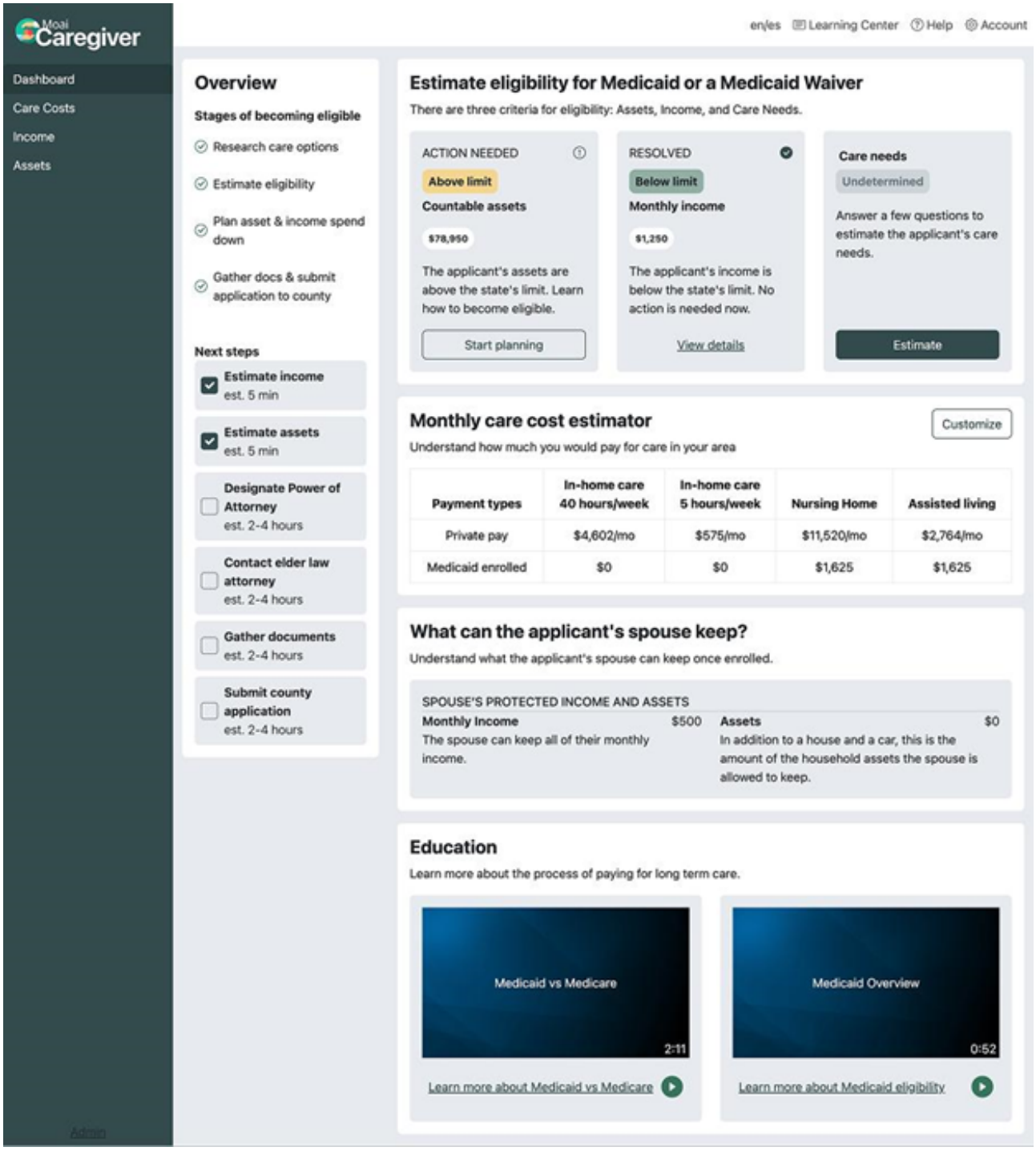


User data collected during the input phase is then passed through an original Medicaid eligibility algorithm that uses a database of state-specific Medicaid eligibility criteria. This database includes information about each state’s income and asset limits, personal needs allowances, home equity amounts, options to qualify when over limits, Medicaid “look-back” periods, and community spouse asset allowances and monthly maintenance needs allowances. The database is also programmed with various state-specific exceptions and thresholds for certain asset types, such as retirement accounts and the equity interest limit of the applicant’s primary home (eg, US $730,000 in most states). The algorithm sums and compares the values entered by the user with the information in the state eligibility database to (1) determine if the applicant is currently eligible for Medicaid because they are below their state’s (a) income limit, and (b) asset limit. If assets are higher than the eligibility limit, the algorithm calculates the amount of the applicant’s asset spend-down obligation (current countable assets minus the state’s asset limit) and the spousal resource allowance, if applicable (either 50% of total household assets or 100% up to the maximum amount allowed by the state, typically US $157,920).

Upon completing the input wizard, users arrive at the Dashboard page that presents a personalized Medicaid eligibility summary (Figure 4). The Dashboard shows whether the care recipient is above or below the state’s income and asset limit. It also provides a summary of average out-of-pocket home and institutional care costs with and without Medicaid enrollment and a checklist of next steps for applying, for example, “contact county Medicaid office” and “gather account statements.” If the care recipient is married, the Dashboard also provides information on the amount of monthly income and assets their community-residing spouse can keep as a resource allowance.

XSL•FO
RenderX

Figure 4. The dashboard of the web application.



From the Dashboard, users can navigate to the “Assets” page (Figure 5), which displays the amount the care recipient must spend down to meet their state’s Medicaid asset limit, excluding exempt assets such as personal belongings and the spousal resource allowance (if applicable). Text with simple graphics informs users about qualified expenses (eg, care services, prepaid burial plots, and home renovations to improve accessibility) and prohibited expenses that violate Medicaid’s “look-back” period for gifts and unauthorized asset transfers.

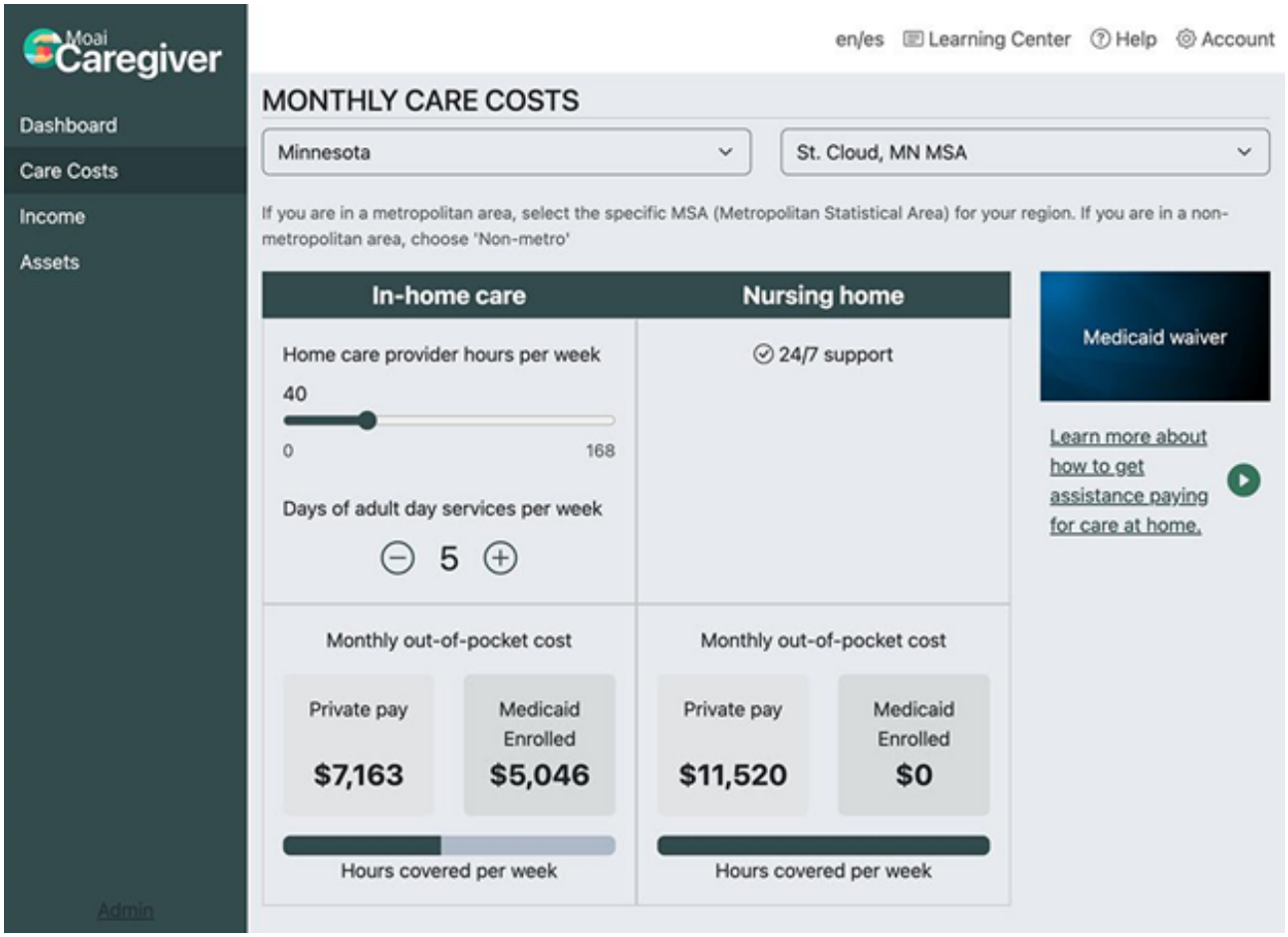
XSL•FO
RenderX

Figure 5. The “Asset Eligibility” page.

XSL•FO
RenderX

Users can also navigate to the "Income" page, which provides personalized information about the state's income eligibility limits, a summary of the care recipient's monthly income sources, and how they can still qualify for Medicaid if their monthly income is above the state's limit. The care cost calculator—accessible from the Dashboard, menu, or the landing page—displays the average monthly cost of different types of LTSS in the applicant's local region (Figure 6). Out-of-pocket costs for private pay are compared to out-of-pocket costs for Medicaid-enrolled individuals. The calculator uses the Federal Long Term Care Insurance Program 2021 Cost of Care Survey database to estimate costs by Metropolitan Statistical Areas (MSAs) for (1) nursing home, private room; (2) home health aide (hourly); and (3) nonmedical adult day care (per day) [30]. For states where adult day care estimates were missing from the Federal Long Term Care Insurance Program database, and for homemaker services, we used an estimate for each state based on the November 2021 Genworth Cost of Care Survey [6]. Monthly out-of-pocket cost estimates are calculated by combining the Medicaid eligibility criteria with data on the cost of care in the region selected by the user. A slider bar allows users to adjust the hours of care provided by a home health aide and add or subtract the number of days of nonmedical adult day care. Average monthly costs of nursing home care (with and without Medicaid coverage) are presented next to the costs of HCBS.

Figure 6. The "care cost calculator" of the web application.



Short, animated videos are contextualized throughout the website to educate users on basic Medicaid facts, including what services may be covered and information about eligibility and spousal resource allowances. Videos were generally less than 1.5 minutes long and covered the following topics: Medicaid versus Medicare, Medicaid eligibility overview, Medicaid waivers, asset spend down, and income eligibility.

After development, the research team led an asynchronous validation study of the final version of the web application with English-speaking informal caregivers of adults aged more than 60 years. This paper presents those results.

Participants

Usability testing was performed using Userfeel (Userfeel Ltd), an online platform where participants record their screens and provide verbal feedback as they navigate websites and complete defined tasks. All Userfeel panel members have screen-capture and audio recording software installed on their computers and are familiar with the testing platform. Sessions were conducted remotely with no live facilitator.

Current and former caregivers of older adults were recruited from Userfeel's participant panel (N=112). Only Userfeel panel members who passed the screening questionnaire were invited to participate. Recruitment criteria included (1) resident of the United States, (2) aged 18 years or older, (3) fluent in English, (4) using a desktop or laptop computer for the session, and (5)

XSL•FO
RenderX

care (or cared) for an older family member or friend aged more than 60 years. Caregiving was defined as "regularly preparing meals, providing transportation, managing household tasks, monitoring medication, helping manage money, and assisting with personal care."

Procedure

Study procedures were reviewed and approved by the University of Minnesota Institutional Review Board. Participants who passed the screening questions and consented to the study were routed to the website. Task instructions and usability questions appeared in a small window in the upper right corner of their computer screens (Multimedia Appendix 2). Participants could minimize the task window but were not able to return to previous tasks or change their answers once they progressed. Participants were instructed to "think aloud" [31] as they navigated the website, providing verbal commentary that was recorded and transcribed by Userfeel. Participants could pause the recording and timer as desired.

Before exploring the landing page, participants answered 11 true or false and multiple-choice questions that assessed their current knowledge of Medicaid LTSS programs and eligibility criteria. These questions appeared in the task window. Participants were told not to refer to the website or to search the internet for answers. After the pretest, participants were presented with a series of tasks which included watching and rating the helpfulness of 4 of the 6 explainer videos, using the intake wizard to enter the care recipient's income, assets, and monthly spending to determine their eligibility for Medicaid, and adjusting the care cost calculator settings to determine average HCBS and institutional care costs if the person was (1) Medicaid-enrolled or (2) a private pay user.

As they followed task instructions and progressed through the website, participants rated their agreement with various statements describing the website's trustworthiness, credibility, attractiveness, the usefulness of specific features, and other characteristics. Responses ranged from 1="strongly disagree" to 5="strongly agree." After answering the user satisfaction questions, participants were presented with a 10-item posttest on Medicaid LTSS programs and eligibility.

The posttest was followed by a brief demographic and caregiver questionnaire that asked participants about their relationship to the care recipient and whether the care recipient was enrolled in Medicaid or a Medicaid waiver program. A final question asked participants to verbalize any final overall impressions of the website. The recording ended after participants closed the final task window or after the timer reached 40 minutes, regardless of whether they finished the tasks.

Participants completed the System Usability Scale (SUS) immediately after the session ended. The SUS is a 10-item scale that asks participants the extent to which they agree or disagree with statements such as, "I thought the system was easy to use" and "I thought there was too much inconsistency in this system." Half of the items are reverse scored, and total scores range from 1 to 100 [32].

Analysis

Researchers reviewed all screen and audio recordings to evaluate session quality, export quotes into a spreadsheet, and document task completion rates (success or fail). Although 112 total sessions were recorded, 3 sessions were dropped from the final sample because they involved the same participant who used a virtual private network (VPN) to enter the study multiple times. Participating in the same usability test more than once is prohibited by Userfeel and the study protocol. Of the remaining 109 participants, 93 people were able to complete all of the Medicaid knowledge pretest and posttest questions, and 89 finished all tasks within the allotted 40 minutes.

Participants' ratings of website features and task completion rates were analyzed descriptively. A paired-samples *t* test was used to determine changes in Medicaid knowledge after engaging with the website and watching the explainer videos. We also analyzed whether the change in knowledge significantly differed between those with and without previous Medicaid experience using an independent samples *t* test. Participants who ran out of time before finishing the Medicaid knowledge posttest did not have their scores included in the analysis. Verbal feedback, including direct quotes, was organized in a spreadsheet according to the website feature.

Ethical Considerations

The protocol was reviewed and approved by the University of Minnesota Institutional Review Board (approval ID STUDY00013179). All participants were presented with the informed consent form during screening. They were required to read and agree to the study procedures before they could begin. No personally identifying information was collected, and all study records were stored securely and accessible only to the research team. All participants were compensated US $30, and the information they entered into the website was deleted from the database after the study period.

Results

Participant Characteristics

Participant sex and age are collected by Userfeel as part of their onboarding process. As illustrated in Table 1, out of 109 participants, 64 (58.7%) identified as female. The average age was 43.6 (SD 12.8) years. Other demographic and caregiving information, which was collected last, is missing from 17 participants (15.6%) who ran out of time before reaching these final questions. Furthermore, 56 participants (51.4%) reported their race as White, 21 (19.3%) as Black or African American, 4 (3.7%) as Asian, 10 (9.2%) as American Indian, multiple races, or "other," and 2 participants (1.8%) preferred not to state. In addition, 12 participants (11%) were Hispanic or Latinx. Of the total, 14 participants (12.8%) reported an annual household income of less than US $25,000, 23 (21.1%) had household incomes between US $25,001 and US $50,000, 35 (32.1%) between US $50,001 and US $100,000, and 20 (18.4%) reported incomes greater than US $100,001.

XSL•FO
RenderX

Table . Participant characteristics (N=109).

Characteristic	Participants
Age (y), mean (SD)	43.6 (12.8)
Sex, n (%)	
Male	44 (40.4)
Female	64 (58.7)
Transgender	1 (0.9)
Race, n (%)	
White or Caucasian	56 (51.4)
Black or African American	21 (19.3)
Asian	4 (3.7)
American Indian or Alaska Native	1 (0.9)
Mixed race or other	9 (8.3)
Prefer not to answer	2 (1.8)
Missing	16 (14.7)
Ethnicity, n (%)	
Hispanic or Latinx	12 (11)
Non-Hispanic or Latinx	81 (74.3)
Missing	16 (14.7)
Income (US $), n (%)	
200,001-300,000	4 (3.7)
150,001-200,000	3 (2.8)
100,001-150,000	13 (11.9)
75,001-100,000	18 (16.5)
50,001-75,000	17 (15.6)
25,001-50,000	23 (21.1)
Under 25,000	14 (12.8)
Missing	17 (15.6)
Care recipient relationship, n (%)	
The care recipient is or was another relative	6 (5.5)
The care recipient is or was my aunt/uncle	3 (2.8)
The care recipient is or was my friend	5 (4.6)
The care recipient is or was my grandparent or grandparent-in-law	30 (27.5)
The care recipient is or was my spouse or partner	1 (0.9)
The care recipient is or was parent or parent-in-law	47 (43.1)
Missing	17 (15.6)
Length of providing care, n (%)	
Greater than 10 years	5 (4.6)
Between 6 and 10 years	10 (9.2)
Between 4 and 6 years	11 (10.1)
Between 2 and 4 years	21 (19.3)
Between 1 and 2 years	22 (20.2)
6 months to 1 year	17 (15.6)
Less than 6 months	6 (5.5)

XSL•FO
RenderX

Characteristic	Participants
Missing	17 (15.6)

In total, 47 participants (43.1%) reported that they care or cared for a parent or parent-in-law, 30 (27.5%) care or cared for a grandparent or grandparent-in-law, 6 (5.5%) care or cared for another relative, and 5 (4.6%) care or cared for a friend. Only one person cares or cared for a spouse. Furthermore, 28 participants responded that the person they care or cared for was enrolled in Medicaid or received a Medicaid waiver for HCBS, 44 (41.3%) were not Medicaid-enrolled, 12 (11%) were unsure, and 9 (14.7%) were missing. Nearly half (n=52) of the participants reported that they assist or assisted the care recipient with money management activities. Userfeel classified 15% (n=17) of participants as "regular users," 77% (n=85) as "power users" (frequent usability testers who are proficient with various online tools), and 7% (n=8) as "web experts" with a deeper understanding of web design.

Medicaid Explainer Videos

Most participants found the animated videos helpful in explaining Medicaid concepts and eligibility rules. The mean helpfulness ratings for the videos ranged from 5.9 to 6.4 (out of 7). Participants commented that the videos provided a useful primer on Medicaid but were confused about some of the details, such as whether a spouse's income is counted toward Medicaid eligibility limits, program differences between states, and rules guiding Medicaid spend down. Others recommended using more images of people:

> *You can probably jazz that up a little bit more with some real people in there…and maybe some clickable buttons where you can find more information 'cuz I still had some questions as I went.* [45-year-old male]

Intake Wizard and Medicaid Eligibility Determination

Table 2 displays mean ratings of various website features. Most found it fairly easy to enter the applicant's financial information requested by the fields (mean 5.9 out of 7, SD 1.34), but based on verbal comments (Multimedia Appendix 3), participants were challenged to enter accurate $ amounts. Some stated that they did not have the care recipient's detailed financial information on hand or did not have access to their accounts at all. However, most strongly agreed that they understood why they were asked to provide this information (mean 4.9 out of 5, SD 0.43). The average length of time it took participants to complete the intake wizard was 6 (SD 3.03) minutes, although many verbalized that they entered approximations or made up the amounts in the interest of privacy.

XSL•FO
RenderX

Table . User ratings on web application features (N=109).

Features	Mean (SD)	Rating scale
Explainer videos		
Medicaid overview video	5.9 (1.08)	1=Extremely unhelpful to 7=Extremely helpful
Medicare versus Medicaid video	6.40 (0.77)	1=Extremely unhelpful to 7=Extremely helpful
Income eligibility video	6.18 (1.24)	1=Extremely unhelpful to 7=Extremely helpful
Asset spend down video	6.09 (1.11)	1=Extremely unhelpful to 7=Extremely helpful
Care cost calculator		
Cost estimator rating	6.32 (1.19)	1=Extremely unhelpful to 7=Extremely helpful
Financial data wizard		
Easy to enter financial information	5.86 (1.34)	1=Extremely difficult to 7=Extremely easy
Average website attribute ratings		
Likely to return in the future	4.32 (1.11)	1=Strongly disagree to 5=Strongly agree
Clean and simple presentation	4.65 (0.59)	1=Strongly disagree to 5=Strongly agree
Attractive	4.04 (0.93)	1=Strongly disagree to 5=Strongly agree
Learned about local care costs	4.43 (0.93)	1=Strongly disagree to 5=Strongly agree
Trust website privacy	4.15 (0.94)	1=Strongly disagree to 5=Strongly agree
Information is trustworthy	4.32 (0.79)	1=Strongly disagree to 5=Strongly agree
Information is credible	4.34 (0.79)	1=Strongly disagree to 5=Strongly agree
Videos helped simplify complex information	4.49 (0.80)	1=Strongly disagree to 5=Strongly agree
Helpful for caregivers	4.85 (0.36)	1=Strongly disagree to 5=Strongly agree
Medicaid eligibility rules are clearer	4.53 (0.63)	1=Strongly disagree to 5=Strongly agree
Did not learn anything useful	1.17 (0.55)	1=Strongly disagree to 5=Strongly agree
Understood why I was asked to enter financial information	4.85 (0.43)	1=Strongly disagree to 5=Strongly agree
Experienced technology issues	1.72 (1.34)	1=Strongly disagree to 5=Strongly agree

On the Dashboard (Figure 4), participants were asked whether the information displayed indicated that the care recipient is eligible for Medicaid based on having income below the state limit (55/109, 51.4% said yes) and assets below the limit (42/109, 39.3% said yes). Several participants commented that they appreciated the list of next steps for LTSS planning and how long these steps would take to complete, such as gathering financial statements for the application. However, many described the Dashboard as being too cluttered and "overwhelming," suggesting a need to simplify the presentation using graphics or reduce the information.

On the Assets page (Figure 5), participants were asked to enter the "spend down" amount displayed or enter US $0 if the care recipient had no spend down obligation because they were currently eligible. Of the total, 89 participants (81.7%) entered the correct number displayed. Of those who failed the task, 2 (1.8% of total) experienced an error message and could not reach the Assets page, 1 (0.9% of total) participant never completed the intake wizard, so the spend down amount could not be calculated, and 11 (10%) could not locate the number displayed.

Care Cost Calculator

Participants completed 2 tasks on the care cost calculator page (Figure 6). First, they were asked to calculate the average monthly cost of 20 hours per week of home health care and 3 days per week of adult day care in Minneapolis-Saint Paul, Minnesota, and then enter the out-of-pocket cost (US $) that was displayed in the window if the care recipient is enrolled in Medicaid. The amount displayed varied across participants because it is based on the care recipient's monthly income. Only 59 participants (54.1%) successfully adjusted the settings and entered the number displayed. Reasons for task failure include not selecting the correct state or MSA (n=36), entering the cost displayed for a nursing home instead of in-home care (n=4), entering the wrong hours or days of care (n=5), or some other reason (n=5).

Using the same settings, participants were asked to enter the average cost displayed for in-home care if the care recipient was NOT enrolled in Medicare (ie, private pay). Successful task completion was similar at 56% (n=61). Errors were largely due to not selecting the correct state or MSA (n=30). In cases where the out-of-pocket care costs for private pay matched the out-of-pocket costs for Medicaid-enrolled, the care recipient's

XSL•FO
RenderX

income exceeded the Medicaid eligibility threshold. Participants verbalized that the slider bar was hard to adjust to an exact number because the selection range was large: 0 to 168 hours. Others commented that they did not understand why the cost of home care was not lower if the care recipient was Medicaid-enrolled (vs private pay).

Despite the challenges participants encountered, most rated the care cost calculator as being very helpful (mean 6.3 out of 7, SD 1.19) and strongly agreed that they learned how much long-term care costs in their area (mean 4.4 out of 5, SD 0.93). Many remarked that they were surprised about the high cost of care.

Improvement in Medicaid Knowledge

Of the total, 93 participants (85%) completed both the Medicaid knowledge pretest and posttest within the 40-minute session. Many participants reasoned through the questions by drawing on their existing knowledge. Those with more experience with Medicaid made comments such as:

> *I think depending on the state, you are allowed to own a home/there is five-year look back [period] for Medicaid.* [60-year-old female]

Using the website and watching 4 of the explainer videos significantly improved participants' knowledge of Medicaid and Medicaid waiver programs and eligibility criteria (t_{92}=12.9, P<.001). Mean performance on the pretest was 42.4% correct (SD 0.19%), and the mean performance on the posttest was 68.4% (SD 0.15), representing a 61.2% average improvement. To assess the effects of previous Medicaid experience on the change in knowledge score, we assessed differences between participants who care or cared for a Medicaid-enrolled person (n=28) versus participants who care or cared for a nonenrolled person or who were unsure if the person was enrolled in Medicaid (n=56). There were no significant differences between these groups (t_{68}=0.05; P=.96), indicating that watching the explainer videos and interacting with the website yielded equal learning gains for those with and without previous experience with Medicaid. However, it is unknown whether providing care to a person enrolled in Medicaid is an adequate proxy for previous knowledge.

User Satisfaction

On a scale of 1 (strongly disagree) to 5 (strongly agree), website attractiveness had a mean rating of 4.0 (SD 0.93). It received a higher average rating for having a "clean and simple presentation" (mean 4.7, SD 0.59). One participant stated:

> *It's a lot of information to process but I think it does a good job of guiding you.* [50-year-old male]

Several participants commented that they wanted to see photographs of diverse older adults to add visual interest to the text-heavy interfaces. The mean rating for trustworthiness of the information was 4.3 (SD 0.79), credibility was 4.3 (SD 0.79), and mean agreement with the statement "I trust the website to keep my responses private" was 4.2 (SD 0.94). In total, 17 participants commented that they were concerned about privacy and were uncertain about how the website intends to use and share the data users enter. A 23-year-old male participant suggested that a privacy disclaimer and a detailed "terms of service" are needed.

Most participants strongly agreed that the website gave them a better understanding of Medicaid eligibility rules (mean 4.5, SD 0.63) and that it would be helpful for caregivers to learn whether a person may be eligible for LTSS benefits through Medicaid (mean 4.9, SD 0.36). Many agreed they would be likely to return sometime in the future (mean 4.3, SD 1.11). Scores were lower among those who shared that the person they cared for is deceased, although some said they would return to explore LTSS options for themselves. Furthermore, 3 participants commented that they wished the website had been available when they were doing care planning for an older relative. Some participants expressed lingering confusion about concepts such as Medicaid spend-down and what assets count toward eligibility limits; however, they recognized that the website covers a challenging topic and providing too many details would be overwhelming. According to one participant,

> *...applying for any kind of government help can be daunting and a little scary, so this was a low-pressure way to determine eligibility.* [66-year-old female]

On average, the total time to complete the user experience testing session was 33 minutes and 43 seconds (SD 5 min 37 s). Participants classified as power users and web experts were more likely to finish all tasks. The average score on the SUS was 88.3 out of 100 (SD 13.1), which is considered very high [33]. Using a slider, participants rated how likely they would be to recommend the website to a friend or colleague. The score was 89.4 out of 100 (SD 18.8), with some stating that they plan to return when they need to apply for Medicaid for themselves or others.

Discussion

Principal Results

Based on a gap analysis and resource needs assessment with ADRD caregivers and aging services providers, we designed a web application that reduces information overload by personalizing the Medicaid eligibility and application planning process to the care recipient's financial circumstances and care preferences. The website received high usability and user satisfaction ratings in the validation and usability test, as well as improved caregivers' knowledge of Medicaid eligibility rules. Participants commented that the intake wizard was simple to use, yet they recognized the importance of entering accurate amounts to get a precise determination of Medicaid eligibility. If the caregiver is not a joint account holder, accessing the care recipient's accounts requires them to be legally appointed as a trustee or attorney-in-fact. Another challenge is that accurate data entry requires a relative understanding of complex financial terms and concepts, such as how the cash value of a life insurance policy is different from the death benefit, and that an annuity can be classified as an income source and an asset. While the current web application provides definitions of selected terms when the user hovers over the text, more education on these concepts is needed. A future design could integrate a conversational artificial intelligence–powered agent

XSL•FO
RenderX

that helps promote financial literacy, but that is trained not to provide legal or financial advice outside the scope of its service.

We found a statistically significant increase in Medicaid knowledge scores from baseline, suggesting that short, animated videos improve understanding of basic Medicaid concepts. Research indicates that short videos are an effective teaching tool and promote retention [34,35]. Participants verbalized that the videos clarified general Medicaid concepts but that some eligibility rules and exceptions needed elaboration, such as what spousal assets and income sources are counted toward a married applicant's eligibility. Future research should assess the effects of other presentation formats on learning outcomes, such as infographics and materials that can be printed, translated into multiple languages, and include hyperlinks to detailed external sources.

Findings from this validation and usability study suggest that caregivers are comfortable entering financial information into the website. This is consistent with research showing that individuals may disclose personal information if the web service offers financial benefits or convenience [36] and also supports the "privacy paradox" whereby individuals' preferences for greater privacy contradict their actual disclosure tendencies [37]. However, some participants expressed concerns about the confidentiality of the information they entered. The development team programmed an account registration feature that was suspended for the validation study to avoid collecting personal identifiers from research participants. Along with providing detailed terms of service, reinstating the login feature is important for securing user data, saving progress, and signaling trustworthiness.

Despite the intake wizard's simple navigation, participants still encountered errors in calculating eligibility and expressed confusion about asset and income limits. They also verbalized confusion about the amount of income the care recipient must contribute to their care costs once enrolled in Medicaid. These findings underscore the complexity of Medicaid program rules and the confusing LTSS landscape. This complexity has led to a robust market for certified Medicaid planners and elder law or estate planning attorneys, but the cost of these services is greater than most middle- and low-income families can afford [16]. Our web application was designed as a solution to this resource gap.

Future Development

While the current website provides proof of concept that a digital platform can help informal caregivers learn about and access Medicaid benefits for the care recipient, future development could further personalize the experience and perhaps facilitate the submission of the Medicaid application to the correct state-level authority. The text could be customized to display the Medicaid program name used in the care recipient's state (eg, "MediCal" for California, "TennCare" for Tennessee) and provide contact information for local care providers that accept Medicaid. For improved accuracy, a more sophisticated eligibility determination algorithm is needed that is automatically updated when state Medicaid policies change and that incorporates additional eligibility exceptions for certain asset types. Many households have complex financial situations, such as assets held in trusts and property located in other states. Determining eligibility for these cases would require the user to enter additional information, adding complexity to the experience. Perhaps a future web application could refer users with complex situations to a local certified Medicaid planner. A future iteration could also integrate a care needs assessment using another intake wizard that collects information on the care recipient's functional abilities, as these also factor into the eligibility determination.

A recent systematic review found that informal caregivers prefer using a smartphone to access digital tools [38]. In the future, this web application can be redesigned as a smartphone application to improve accessibility. Despite the website's limitations, many participants, particularly those aged 50 years and older, expressed that they wished the application had been available when they provided care to an older relative. Others shared that they want to return to the website when they are ready to plan for their own care needs. This feedback suggests that there is a strong interest in personalized digital solutions to simplify Medicaid eligibility determinations and learn about regional LTSS costs.

Limitations

Usability sessions were asynchronous, limiting our ability to ask follow-up questions. Of the total, 2 participants did not complete the intake wizard and encountered error messages that prevented them from completing some of the tasks. This error could be mitigated by limiting navigation options—that is, access to the Dashboard, Assets, and Income pages—before completing the financial intake wizard. Furthermore, 5 participants did not watch all the required videos due to slow loading speeds. This may have affected their satisfaction ratings and scores on the Medicaid knowledge posttest. All multiple-choice and true or false questions were pretested with a separate online panel (n=30), and items were counterbalanced with respect to difficulty and format in the validation study. However, Userfeel does not allow randomization in question and task ordering, which could produce order effects and impact task performance.

Participant confidentiality and an explanation for how data would be used and secured by the researchers were described in the participant consent presented before the usability test. To maintain participant confidentiality, web developers removed the "Create an Account" feature, which would have asked for the participant's name, email, and a password creation. This account creation page described that no data will be shared with third parties. The absence of a secure login page may have negatively impacted participants' perceptions of how their information would be secured and shared.

Many participants expressed that they approximated or "made up" asset, income, and monthly spending amounts, affecting the Medicaid eligibility calculation. While accurate entries were not essential for usability testing, accuracy could be enhanced by requiring users to create a login to save their progress, allowing them to return to the website after gathering the applicant's recent financial statements. The ability to return to the site later without having to start over would also reduce time pressure.

XSL•FO
RenderX

Website validation tests were performed entirely in English by experienced users, limiting the generalizability of the findings. Future validation studies should include inexperienced web users, older caregivers (older than 70 years), and participants with differing levels of experience applying for government programs for themselves or others. A delayed follow-up assessment would also be beneficial to assess knowledge retention of Medicaid concepts.

Although a previous research phase involved usability testing with Spanish-speaking caregivers, further assessment is needed to test the website in other languages. Finally, there were very few spousal caregivers in the sample. Future usability tests should be conducted with spousal caregivers who are typically older than the current participants.

Comparison With Previous Work

Every state's Medicaid office has a website with information on program eligibility and application instructions. The Eldercare Locator, a helpline funded by the Administration on Aging, provides guidance on regional aging services providers and resources, as well as a live chat, but does not offer personalized Medicaid eligibility information [39]. AARP's website offers guidance on getting started with long-term care planning and educates members on Medicare versus Medicaid coverage but does not provide a benefits calculator. The most comprehensive and up-to-date resource on Medicaid LTSS programs was developed by the American Council on Aging [16]. It provides detailed data tables on each state's income and asset eligibility criteria for married and single applicants, information on Medicaid HCBS waivers by state, rules regarding asset spend down, spousal asset protections, Medicaid estate recovery, nursing home costs, and legal strategies to shield assets. The website also offers a simple eligibility test, but it does not provide an eligibility determination. Rather, users are asked to enter their names and emails to be connected with a Medicaid planner who may charge a service fee. We presented this website to interview participants in the initial resource needs and gap analysis. While caregivers found the site incredibly comprehensive, they described being overwhelmed by the details and the text-heavy interface.

To minimize information overload, we designed our website to limit the amount of information on each page and personalized the eligibility information based on the care recipient's state and financial circumstances. The videos used simple animations to depict important concepts that affect planning decisions, like spousal asset preservation and Medicaid spend-down strategies. These videos were intended to make Medicaid more approachable for caregivers experiencing high levels of cognitive load and emotional burden. We find that videos are an effective way to improve Medicaid program knowledge, although some content was omitted or simplified. In the future, caregivers would benefit from a digital tool with customized educational videos that is powered to flag financial circumstances that warrant professional advice from an attorney or certified Medicaid planner.

Conclusions

Medicaid programs provide a critical safety net for millions of older adults and their family caregivers who cannot afford to pay for LTSS out-of-pocket [40]. The application process can be daunting, especially for informal family caregivers with little previous knowledge of Medicaid programs [14,15,19]. Navigating care options and determining how to finance LTSS is very challenging for older adults and their informal caregivers who may be engaging in the process while managing the care recipient's escalating needs. We designed a website that illuminates the out-of-pocket cost of HCBS and institutional LTSS, educates users on Medicaid programs to help offset these costs, and simplifies eligibility determinations. The encouraging results from the website validation study suggest that there is a clear need for a digital solution to support families in the LTSS and Medicaid planning, but that additional features are needed to educate caregivers about state-specific Medicaid rules, program options, and to streamline the application submission process.

Acknowledgments

This project was funded by the National Institute on Aging through a Small Business Innovation Research award 5R44AG074135-03.

Authors' Contributions

Authors MD, JS, and OO designed the web application and conceived of the validation study. OO supervised the development and programming of the web application. SG analyzed the quantitative and qualitative data. MD wrote the first draft of the manuscript and supervised data collection and analysis. JS created the manuscript figures, programmed the validation test in Userfeel, produced the initial wireframes for the web application, and edited the manuscript. KM participated in the qualitative data analysis, created manuscript tables, and assisted with programming the validation test. All authors contributed to and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

The first images display photographs from the researcher and developer synchronous problem definition session. The photographs depict a white board with text and sticky notes denoting user needs and possible solutions. Next is a diagram of the initial user journey map created for the first iteration of the application. The last set of images presents 4 example caregiver user personas.
[DOCX File, 2541 KB - aging_v8i1e77441_app1.docx]

Multimedia Appendix 2
Screenshot from a Userfeel testing session. The task window appears in the upper right corner of the participant's screen. The participant is on the landing page answering a Medicaid knowledge pretest question.
[DOCX File, 382 KB - aging_v8i1e77441_app2.docx]

Multimedia Appendix 3
Verbal feedback and participant quotes.
[DOCX File, 24 KB - aging_v8i1e77441_app3.docx]

References

1. Johnson RW. What is the lifetime risk of needing and receiving long-term services and supports? : Office of the Assistant Sectretary for Planning and Evaluation; 2019 Apr 3 URL: https://aspe.hhs.gov/reports/what-lifetime-risk-needing-receiving-long-term-services-supports-0 [accessed 2025-08-26]
2. Fabius CD, Okoye SM, Mulcahy J, Burgdorf JG, Wolff JL. Associations between use of paid help and care experiences among Medicare-Medicaid enrolled older adults with and without dementia. J Gerontol B Psychol Sci Soc Sci 2022 Dec 29;77(12):e216-e225. [doi: 10.1093/geronb/gbac072] [Medline: 35554530]
3. Kasper JD, Freedman VA, Spillman BC, Wolff JL. The disproportionate impact of dementia on family and unpaid caregiving to older adults. Health Aff (Millwood) 2015 Oct;34(10):1642-1649. [doi: 10.1377/hlthaff.2015.0536] [Medline: 26438739]
4. Deb A, Thornton JD, Sambamoorthi U, Innes K. Direct and indirect cost of managing Alzheimer's disease and related dementias in the United States. Expert Rev Pharmacoecon Outcomes Res 2017 Apr;17(2):189-202. [doi: 10.1080/14737167.2017.1313118] [Medline: 28351177]
5. Manly JJ, Jones RN, Langa KM, et al. Estimating the prevalence of dementia and mild cognitive impairment in the US: the 2016 health and retirement study harmonized cognitive assessment protocol project. JAMA Neurol 2022 Dec 1;79(12):1242-1249. [doi: 10.1001/jamaneurol.2022.3543] [Medline: 36279130]
6. Calculate the cost of long-term care near you. CareScout. URL: https://stg-origin.genworth.com/aging-and-you/finances/cost-of-care.html [accessed 2025-04-28]
7. Hamel L, Montero A. The affordability of long-term care and support services: findings from a KFF survey. Kaiser Family Foundation. 2023 Nov 14. URL: https://www.kff.org/health-costs/poll-finding/the-affordability-of-long-term-care-and-support-services/ [accessed 2025-04-28]
8. Oney M, White L, Coe NB. Out-of-pocket costs attributable to dementia: a longitudinal analysis. J Am Geriatr Soc 2022 May;70(5):1538-1545. [doi: 10.1111/jgs.17746] [Medline: 35278213]
9. Kaye HS, Harrington C, LaPlante MP. Long-term care: who gets it, who provides it, who pays, and how much? Health Aff (Millwood) 2010;29(1):11-21. [doi: 10.1377/hlthaff.2009.0535] [Medline: 20048355]
10. Thach NT, Wiener JM. An overview of long-term services and supports and Medicaid. : Office of the Assistant Secretary for Planning and Evaluation; 2018 URL: https://aspe.hhs.gov/reports/overview-long-term-services-supports-medicaid-final-report-0 [accessed 2025-08-26]
11. Wiener JM, Lepore M, Jones J. What policymakers need to know about long-term services and supports. Public Policy Aging Rep 2018 Apr 17;28(1):29-34. [doi: 10.1093/ppar/pry001]
12. Donohue JM, Cole ES, James CV, Jarlenski M, Michener JD, Roberts ET. The US Medicaid program: coverage, financing, reforms, and implications for health equity. JAMA 2022 Sep 20;328(11):1085-1099. [doi: 10.1001/jama.2022.14791] [Medline: 36125468]
13. Murray C, Stepanczuk C, Carpenter A, Wysocki A. Trends in users and expenditures for home and community-based services as a share of total Medicaid LTSS users and expenditures, 2022. Centers for Medicare and Medicaid Services. 2024 Aug 29. URL: https://www.medicaid.gov/medicaid/long-term-services-supports/downloads/ltss-rebalancing-brief-2022.pdf [accessed 2025-08-26]
14. Long-term care in America: views on who should bear the responsibilities and costs of care. The Long-Term Care Poll. URL: https://www.longtermcarepoll.org/long-term-care-in-america-views-on-who-should-bear-the-responsibilities-and-costs-of-care/ [accessed 2025-04-28]
15. Scattarelli J, Moeller K, Urbanski D, DeLiema M. Perspectives on the challenges of planning for and accessing long-term dementia care services through Medicaid and Medicaid waivers. J Appl Gerontol 2025 Apr 28:7334648251333181. [doi: 10.1177/07334648251333181] [Medline: 40294891]
16. Medicaid eligibility: 2025 income, asset & care requirements for nursing homes & long-term care. American Council on Aging. 2025. URL: https://www.medicaidplanningassistance.org/medicaid-eligibility [accessed 2025-04-28]

XSL•FO
RenderX

17. Moynihan DP, Herd P, Ribgy E. Policymaking by other means: do states use administrative barriers to limit access to Medicaid? Adm Soc 2016;48(4):497-524. [doi: 10.1177/0095399713503540]
18. Fox AM, Stazyk EC, Feng W. Administrative easing: rule reduction and Medicaid enrollment. Public Adm Rev 2020 Jan;80(1):104-117. [doi: 10.1111/puar.13131]
19. Hahn H, Pratt E, Knowles S. Strategies for improving public benefits access and retention. : Urban Institute; 2023 URL: https://digitalgovernmenthub.org/wp-content/uploads/2023/03/Strategies-for-Improving-Public-Benefits-Access-and-Retention.pdf [accessed 2025-08-26]
20. Weinmann M, Schneider C, Brocke JV. Digital nudging. Bus Inf Syst Eng 2016 Dec;58(6):433-436. [doi: 10.1007/s12599-016-0453-1]
21. Ghahramani F, Wang J. Intention to adopt mHealth apps among informal caregivers: cross-sectional study. JMIR Mhealth Uhealth 2021 Mar 17;9(3):e24755. [doi: 10.2196/24755] [Medline: 33729166]
22. Hassan AYI. Challenges and recommendations for the deployment of information and communication technology solutions for informal caregivers: scoping review. JMIR Aging 2020 Jul 29;3(2):e20310. [doi: 10.2196/20310] [Medline: 32723720]
23. Premanandan S, Ahmad A, Cajander Å, Ågerfalk P, van Gemert-Pijnen L. Designing and evaluating IT applications for informal caregivers: scoping review. J Med Internet Res 2024 Oct 23;26:e57393. [doi: 10.2196/57393] [Medline: 39442167]
24. Shen H, Han Y, Shi W, et al. Perspectives and experiences of family caregivers using supportive mobile apps in dementia care: meta-synthesis of qualitative research. JMIR Mhealth Uhealth 2025 Jun 18;13(1):e65983. [doi: 10.2196/65983] [Medline: 40532203]
25. Turk A, Fairclough E, Smith GG, Lond B, Nanton V, Dale J. Exploring the perceived usefulness and ease of use of a personalized web-based resource (care companion) to support informal caring: qualitative descriptive study. JMIR Aging 2019 Aug 20;2(2):e13875. [doi: 10.2196/13875] [Medline: 31518272]
26. Egan KJ, Hodgson W, Dunlop MD, Imperatore G, Kirk A, Maguire R. A novel mobile app (“CareFit”) to support informal caregivers to undertake regular physical activity from home during and beyond COVID-19 restrictions: co-design and prototype development study. JMIR Form Res 2021 Oct 1;5(10):e27358. [doi: 10.2196/27358] [Medline: 34406969]
27. Boutilier JJ, Loganathar P, Linden A, et al. Correction: A web-based platform (CareVirtue) to support caregivers of people living with Alzheimer disease and related dementias: mixed methods feasibility study. JMIR Aging 2022 Sep 14;5(3):e41912. [doi: 10.2196/41912] [Medline: 36103711]
28. Vaughan C, Trail TE, Mahmud A, Dellva S, Tanielian T, Friedman E. Informal caregivers’ experiences and perceptions of a web-based peer support network: mixed-methods study. J Med Internet Res 2018 Aug 28;20(8):e257. [doi: 10.2196/jmir.9895] [Medline: 30154074]
29. Ball J. The Double Diamond: a universally accepted depiction of the design process. Design Council. 2019. URL: https://www.designcouncil.org.uk/our-work/news-opinion/double-diamond-universally-accepted-depiction-design-process/ [accessed 2025-05-04]
30. The Federal Long Term Care Insurance Program 2021 cost of care survey. : The Federal Long Term Care Insurance Program URL: https://cdn.ltcfeds.com/planning-tools/downloads/Cost-of-Care-Survey.pdf [accessed 2025-08-26]
31. Cooke L. Assessing concurrent think-aloud protocol as a usability test method: a technical communication approach. IEEE Trans Profess Commun 2010;53(3):202-215. [doi: 10.1109/TPC.2010.2052859]
32. Lewis JR. The System Usability Scale: past, present, and future. International Journal of Human–Computer Interaction 2018 Jul 3;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
33. Sauro J. Measuring usability with the System Usability Scale (SUS). MeasuringU. URL: https://measuringu.com/sus/ [accessed 2025-05-12]
34. Afify MK. Effect of interactive video length within e-learning environments on cognitive load, cognitive achievement and retention of learning. Turkish Online Journal of Distance Education 2020;21(4):68-89. [doi: 10.17718/tojde.803360]
35. Brame CJ. Effective educational videos: principles and guidelines for maximizing student learning from video content. CBE Life Sci Educ 2016;15(4):es6. [doi: 10.1187/cbe.16-03-0125] [Medline: 27789532]
36. Hann IH, Hui KL, Lee SYT, Png IPL. Overcoming online information privacy concerns: an information-processing theory approach. Journal of Management Information Systems 2007 Oct;24(2):13-42. [doi: 10.2753/MIS0742-1222240202]
37. Kokolakis S. Privacy attitudes and privacy behaviour: a review of current research on the privacy paradox phenomenon. Computers & Security 2017 Jan;64:122-134. [doi: 10.1016/j.cose.2015.07.002]
38. Soares S, Hoffmeister LV, Fernandes MDF, Henriques A, Costa A. The use of digital technologies in the promotion of health literacy and empowerment of informal caregivers: scoping review. JMIR Aging 2024 Apr 29;7:e54913. [doi: 10.2196/54913] [Medline: 38683655]
39. Eldercare locator. Administration for Community Living. URL: https://eldercare.acl.gov/home [accessed 2025-04-28]
40. Thomas KS, Applebaum R. Long-term Services and Supports (LTSS): a growing challenge for an aging America. PPAR 2015;25(2):56-62. [doi: 10.1093/ppar/prv003]

Abbreviations

ADRD: Alzheimer disease and related dementias

XSL•FO
RenderX

HCBS: home and community-based services
LTSS: long-term services and supports
MSA: Metropolitan Statistical Area
SUS: System Usability Scale

Edited by PH Chau; submitted 15.05.25; peer-reviewed by F Yeh, J Oghuan; revised version received 15.07.25; accepted 17.07.25; published 27.08.25.

Please cite as:
DeLiema M, Gao S, Scattarelli J, Moeller K, Olofinboba O
Medicaid Education and Eligibility Planning for Caregivers: Website Usability and Validation Study
JMIR Aging 2025;8:e77441
URL: https://aging.jmir.org/2025/1/e77441
doi:10.2196/77441

© Marguerite DeLiema, Siyu Gao, Justine Scattarelli, Kelly Moeller, Olu Olofinboba. Originally published in JMIR Aging (https://aging.jmir.org), 27.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Culturally Adapted STAR-Caregivers Virtual Training and Follow-Up for Latino Caregivers of People Living With Dementia: Single-Arm Pre-Post Mixed Methods Study

Miguel Angel Mariscal, MPH; Celeste Garcia, MPH; Lily Zavala, BS; Magaly Ramirez, PhD

Department of Health Systems and Population Health, School of Public Health, University of Washington, Hans Rosling Center for Population Health, 3980 15th Avenue NE, Fourth Floor, Box 351621, Seattle, WA, United States

Corresponding Author:
Magaly Ramirez, PhD
Department of Health Systems and Population Health, School of Public Health, University of Washington, Hans Rosling Center for Population Health, 3980 15th Avenue NE, Fourth Floor, Box 351621, Seattle, WA, United States

Abstract

Background: Latino caregivers are at an increased risk of negative health outcomes due to the responsibilities of caring for someone with dementia. Although interventions exist to address caregiver burden, they often do not meet the cultural needs of Latino caregivers.

Objective: This study aimed to pilot test the cultural adaptation of the STAR-Caregivers Virtual Training and Follow-Up (STAR-VTF) intervention. The intervention is an evidence-based training program designed to teach family caregivers strategies to manage behavioral and psychological symptoms of dementia (BPSD). Our research team has conducted past studies to identify and perform culturally relevant adaptations to the training modules of STAR-VTF, and this study aimed to pilot these culturally adapted modules with a sample of Latino caregivers.

Methods: Data on feasibility, usability, and acceptability were collected from a pilot test in which Latino caregivers (n=16) used the training modules of the STAR-VTF intervention over a 7-week period. Participants completed usability surveys following the completion of each module, and acceptability was assessed through semistructured interviews (n=14) postintervention. Preliminary outcome measures were also collected, and a descriptive analysis was conducted. The primary outcomes were the Revised Memory and Behavior Problem Checklist (RMBPC) and the Preparedness for Caregiving Scale.

Results: The pilot study results suggest that it is feasible to deliver the culturally adapted STAR-VTF intervention to Latino caregivers, with 94% (15/16) of participants maintaining enrollment through intervention completion. The intervention's usability was found to be "good" based on an average System Usability Score of 76.7 out of 100 across all training modules. Caregivers were generally satisfied with the training modules. In addition, preliminary outcome results demonstrated a trend of decreased BPSD pre- versus postintervention (RMBPC subscale score: 28.24 to 21.34). Findings also demonstrated decreased caregiver reaction to BPSD pre- versus postintervention (RMBPC subscale score: 40.40 to 37.21) and increased caregiver preparedness based on pre- and postintervention (Preparedness Caregiving Scale score: 1.98 to 2.43).

Conclusions: The pilot study demonstrated that the culturally adapted STAR-VTF intervention is feasible and perceived as easy to use by a small sample of Latino caregivers. We aim to refine the cultural adaptations of the STAR-VTF intervention further based on feedback from study participants. Future studies are necessary to test the efficacy of the intervention and support the broad dissemination of the culturally adapted intervention.

(JMIR Aging 2025;8:e66053) doi:10.2196/66053

KEYWORDS

dementia; gerontology; geriatric; older person; aging; Latina; Latino; caregiver; symptom management; digital technology; digital intervention; digital health application; pilot study

Introduction

In the United States, Latino populations are disproportionately impacted by Alzheimer disease and related dementias (ADRD). Compared to White adults, Latinos are 1.5 times more likely to have ADRD [1]. By 2060, the number of Latinos with ADRD is expected to grow by 832%, which is the steepest increase in ADRD compared to any other racial or ethnic group, and projections show that by 2060, a total of 3.5 million Latinos will be living with dementia [2,3]. Latinos with dementia primarily rely on family members for their care due to the high cost of formal dementia care, legal status, barriers to accessing dementia care services, language preferences, and cultural preferences [4,5].

XSL•FO
RenderX

Latino caregivers comprise 21% of the estimated 40 million family caregivers in the United States [1]. Caregiving for a person living with dementia can be particularly challenging as ADRD can cause changes in personality and behavior, losses in judgment, orientation, and the ability to understand and communicate effectively [6]. Managing dementia symptoms requires intense supervision and physical support, and compared to other populations, Latino families provide more intense caregiving in terms of level of care and hours [1]. Caregiving for a person living with dementia also places Latino caregivers at increased risk of negative health implications, including high levels of stress, depression symptoms, less engagement in physical activity, sleeping problems, and social isolation [4,7]. Although evidence-based caregiver interventions exist to help address the negative impacts of caregiver burden, they often do not meet the unique cultural needs of Latino families [8]. Due to the negative health implications of dementia caregiving, it is crucial to support Latino caregivers through culturally adapted, evidence-based interventions.

STAR Caregivers (STAR-C) is an evidence-based systematic training program designed to teach family caregivers strategies to reduce behavioral distress and manage behavioral and psychological symptoms of dementia [8]. Caregivers are taught effective communication strategies, activator-behavior-consequence (ABC) behavioral problem-solving strategies, and how to identify and use pleasant activities to reduce behavioral disturbances [9]. STAR-C has been demonstrated to significantly reduce caregiver burden and depression and improve the quality of life of persons with ADRD [10]. Since the original randomized controlled trial of STAR-C in 2005 [11], the program has been adopted in various settings to meet caregivers' needs. One of the early adoptions of the STAR-C program in a real-world setting was conducted by the Oregon Department of Human Services-State Unit on Aging and Area Agencies of Aging, which served both rural and urban locations in Oregon [9,12]. In this implementation of STAR-C, the training was provided at home by trained professional health consultants to family caregivers. The program has since been adapted to teach caregivers the STAR-C lessons remotely. Tele-STAR, the telehealth-based adaptation of STAR-C, was created to increase access to anyone with a computer and internet connection and to reach rural populations [13]. Tele-STAR used videoconferencing to connect nurse consultants with family caregivers to guide them through the STAR-C lesson plans. The tele-STAR study found that with the transition from in-person training to remote, the family caregiver burden was reduced, and it retained good program and treatment fidelity to STAR-C [13]. Videoconferencing technology has also been used in the Illinois-based CRIS' Memory Care Program, an adapted STAR-C intervention developed by CRIS Healthy Aging, a nonprofit organization in East Central Illinois that serves older adults primarily in rural communities and of lower income [14]. The latest adaptation of STAR-C in a clinical setting has been STAR-C Virtual Training and Follow-up (STAR-VTF), which was tested at Kaiser Permanente Washington [15]. STAR-VTF is a 6-to-8-week program comprised of six training modules that caregivers complete asynchronously. It also includes 30-minute weekly check-ins with a program coach. Support from program coaches is provided as needed for up to 6 months via secure messaging in the Kaiser Permanente Washington patient portal.

Although STAR-C has been adapted in a variety of ways to expand access to caregivers, it has yet to be tested in large populations of Latino caregivers. Recently, our research team adapted STAR-VTF to improve its cultural relevance and linguistic appropriateness for Latino caregivers [8,15]. We conducted an 8-week pilot study of the culturally adapted training modules from the STAR-VTF intervention. This study reports findings on the feasibility, usability, acceptability, and preliminary outcomes of the pilot study with Latino caregivers. The findings will help improve future iterations of STAR-VTF for Latino caregivers, and more generally, advance our understanding of modifying evidence-based interventions to better serve Latino caregivers of persons with dementia.

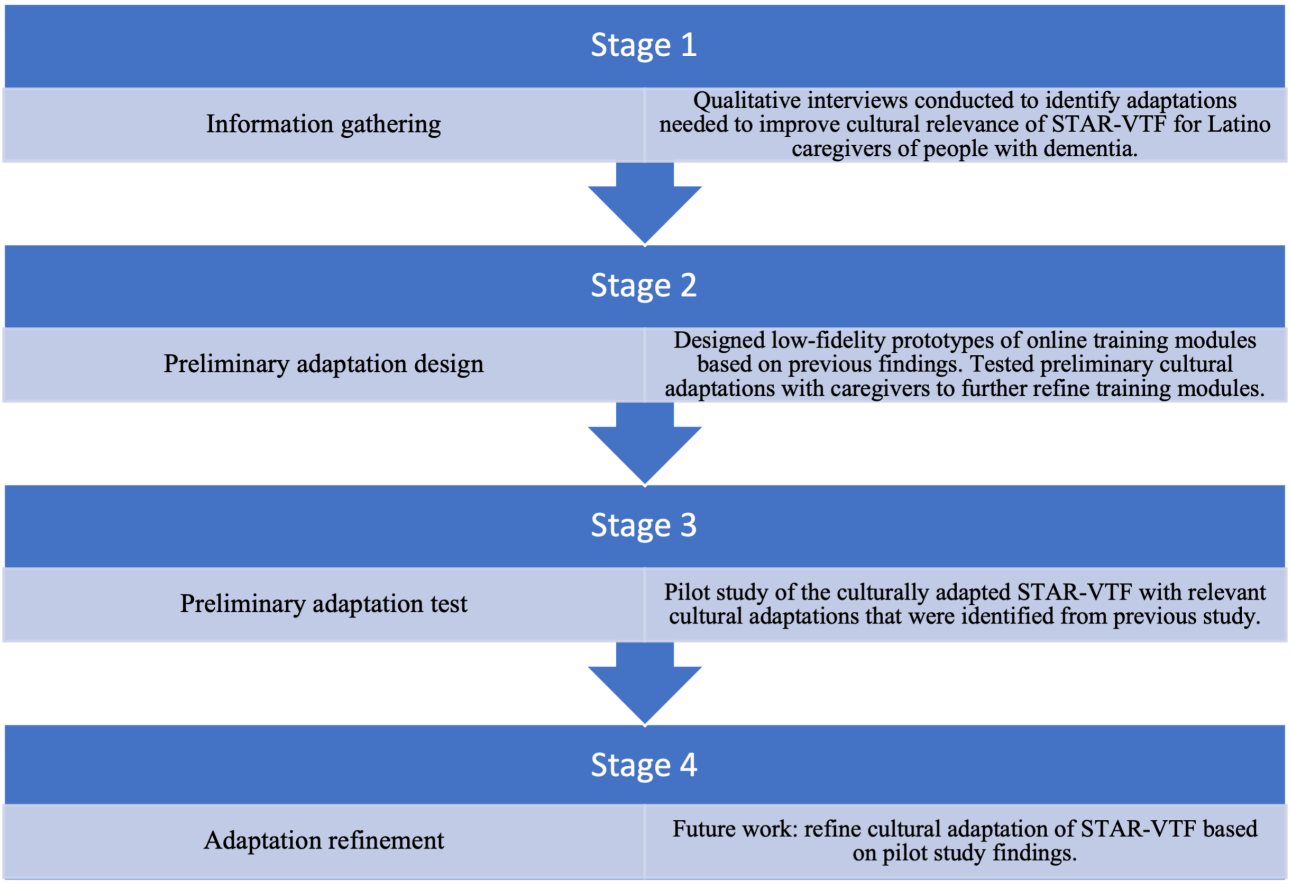
Methods

Design

The framework for the cultural adaptation of evidence-based interventions developed by Barrera and Castro [16] guided the methods of this study. The framework includes the following stages in culturally adapting an evidence-based intervention: (1) gathering information to identify ideas about needed adaptations, (2) conducting preliminary adaptations based on these ideas, (3) conducting pilot studies of the preliminary adaptations, and (4) refining adaptations based on results from pilot studies. Our research team has completed the first two stages [8,15]. This study presents findings in the stage 3 process of pilot-testing cultural adaptations (Figure 1).

XSL•FO
RenderX

Figure 1. STAR-VTF cultural adaptation process. STAR-VTF: STAR-Caregivers Virtual Training and Follow-Up.



This single-arm pilot study tested the feasibility, usability, and acceptability of the culturally adapted STAR-VTF training modules. We also trialed data collection procedures and collected preliminary data on persons living with dementia and caregiver outcomes pre- and postintervention to obtain descriptive estimates. Quantitative and qualitative methods were used in this pre- and posttest study.

Ethical Considerations

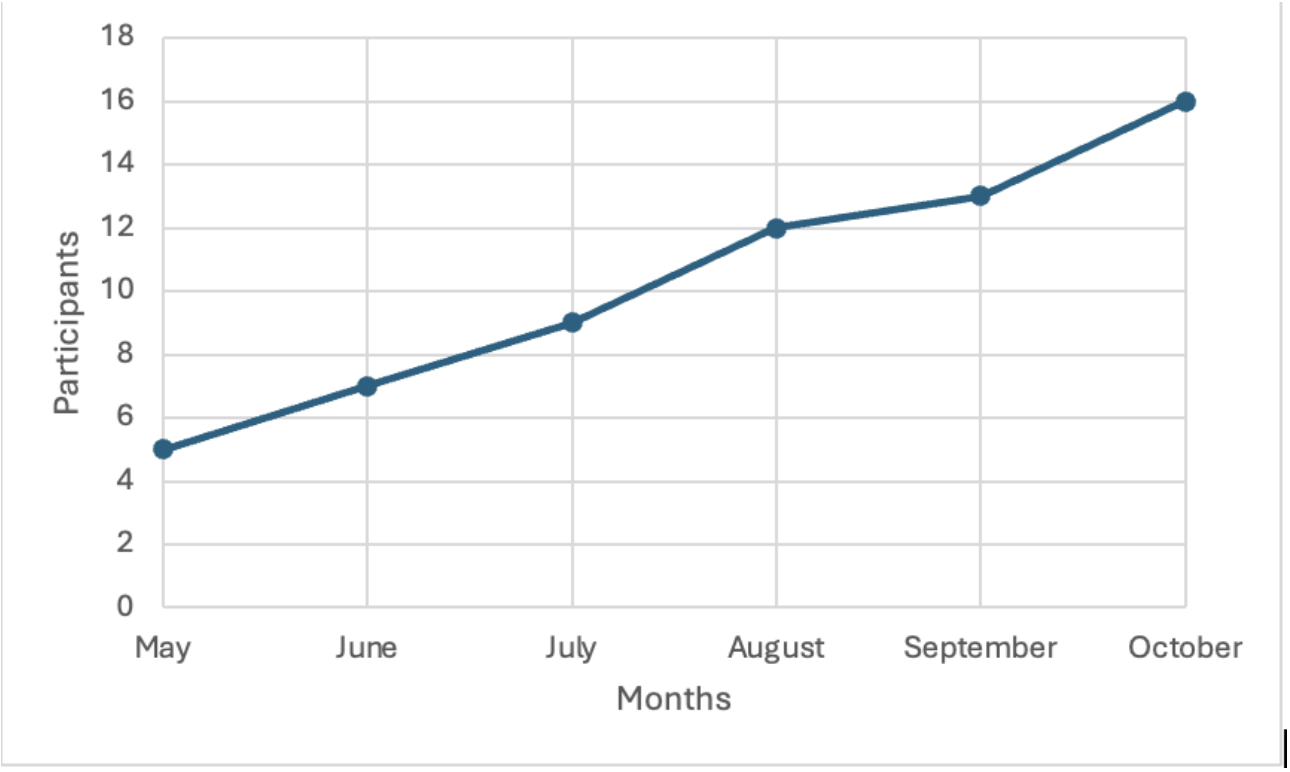
The Advarra Institutional Review Board approved this study (Protocol: National Institute on Aging - FY20_Pilot10_Ramirez). If interested, participants were screened for eligibility, and digital or written informed consent was obtained. Consent was collected from participants via a consent form that was provided in English and Spanish, depending on their preference. Translation of the consent form was done by a certified Spanish translator. Participants were provided a US $40 gift card for their participation. All study data were deidentified and stored on secure, password-protected cloud storage.

Participant Recruitment

Potential study participants (ie, Latino caregivers) were identified using the UW Medicine electronic health record, UW Alzheimer's Disease Research Center Registry, Alzheimer's Prevention Registry maintained by the Banner Alzheimer's Institute, and through the distribution of recruitment flyers through professional networks. Potential study participants were contacted by bilingual staff via phone or email to explain the study and answer their questions about study participation. To be eligible to participate, participants had to identify as Hispanic or Latino, be 18 years of age or older, be taking care of a family member or friend with dementia, live with or within 5 miles of the person living with dementia, spend at least 8 hours a week with them, and indicate that the person they care for is experiencing at least three symptoms related to dementia occurring at least three times in the past week. Figure 2 provides additional information on the number of participants who were enrolled over time for this pilot study.

XSL•FO
RenderX

Figure 2. Enrollment timeline.



Intervention: Culturally Adapted STAR-VTF Training Modules

While the full STAR-VTF intervention consists of weekly training modules and weekly 30-minute phone check-ins with a coach [15], this study only pilot-tested the culturally adapted modules. The adaptations used in this pilot study were based on findings from previous qualitative studies [8,15]. These studies interviewed Latino caregivers of persons with dementia, as well as health care and social service providers of older Latinos, to improve the cultural relevance of STAR-VTF. Based on the findings of prior studies, adaptations included revising language viewed as problematic, expanding content to enhance understanding of dementia, and adding cultural examples that reflect family involvement in caring for relatives with dementia and multigenerational living [8,15]. Additional refinements made to the STAR-VTF training modules, after testing preliminary modules with Latino caregivers, included adding empathetic messaging that highlights the importance of viewing the world from the perspective of the person living with dementia, incorporating additional problem-solving examples to demonstrate diverse challenges, and emphasizing caregiver self-care, to ensure caregiver's mental and emotional health is prioritized [8].

For 7 weeks, caregivers completed training modules asynchronously. They accessed the modules via email or SMS text message. This pilot study was focused on testing the content of the training modules; therefore, caregivers did not receive the coaching component of the STAR-VTF intervention. There were 7 training modules in total. Caregivers were instructed to complete one module per week. The content of the modules is described in Textbox 1.

XSL•FO
RenderX

Textbox 1. Content of the STAR-Caregivers Virtual Training and Follow-Up training modules.

Module 1

- Provides an understanding and overview of dementia.

Module 2

- Introduces caregivers to the behavioral treatment of dementia, realistic expectations, and effective communication.

Module 3

- Covers the antecedents, behaviors, consequences (ABC) approach to problem-solving, including rationale and development of an ABC plan. Hypothetical scenarios of how caregivers successfully used the ABC problem-solving to address behaviors (eg, wandering) are included.

Module 4

- Instructs caregivers to review the ABC plan and revise it as needed.

Module 5

- Covers pleasant events and managing negative thinking.

Module 6

- Instructs caregivers to review the ABC plan, pleasant activities schedule, and to revise as needed.

Module 7

- Covers coping with caregiving and maintaining the use of caregiving strategies.

The length of the training modules ranged between 7 and 16 minutes each. The modules used text, pictures, and illustrations with a voiceover presentation. Caregivers received the training modules in their preferred language (English or Spanish). They also received a printed or electronic workbook to accompany the lessons (the format depended on their preference). The workbook contained additional educational materials, as well as worksheets for caregivers to write down their ABC plans. Intervention materials, including training modules, voiceovers for modules, and workbooks, were translated into Spanish by a certified Spanish translator. Study participants were enrolled in the study for approximately 8 weeks.

Measures

Feasibility

We assessed feasibility based on National Institutes of Health guidance on appropriate feasibility measures for pilot studies [17]. Feasibility measures included details of the success of procedures to recruit eligible participants, retention, and intervention completion.

Usability

To assess usability, we collected data on the System Usability Scale [18] via a REDCap (Research Electronic Data Capture; Vanderbilt University) survey each time a caregiver completed a training module. With the System Usability Scale survey, caregivers rated how easy the training modules were to use. Each item was rated on a 5-point scale ranging from 0=strongly disagree to 5=strongly agree.

Acceptability

Assessment of acceptability was accomplished through semistructured qualitative interviews with participants after they completed 7 weeks of the STAR-VTF training modules. Interview questions explored the participant's experience with the program and solicited feedback on recommendations for further improving the program.

Person Living With Dementia and Caregiver Outcomes

The primary outcomes were the Revised Memory and Behavior Problem Checklist (RMBPC) [19,20] and the Preparedness for Caregiving Scale [21]. The outcomes were collected at baseline and 8 weeks postenrollment via a REDCap survey. For the RMBPC, caregivers rated memory, depression, and disruptive behavior problems in the person living with dementia. The instrument contains 24 items (7 memory-related, 8 depressive, and 9 disruptive) that assess problem behaviors and are rated for frequency of occurrence during the past week and caregiver reaction, which refers to the extent to which caregivers were distressed by the problem behaviors. The Preparedness for Caregiving Scale was used for caregivers to rate how prepared they are for various aspects of caregiving. The instrument contains 8 items that ask caregivers how well prepared they believe they are to provide physical care and emotional support, deal with the stress of caregiving, and set up in-home support services. Each item is rated on a 5-point scale ranging from 0=not at all prepared to 5=very well prepared.

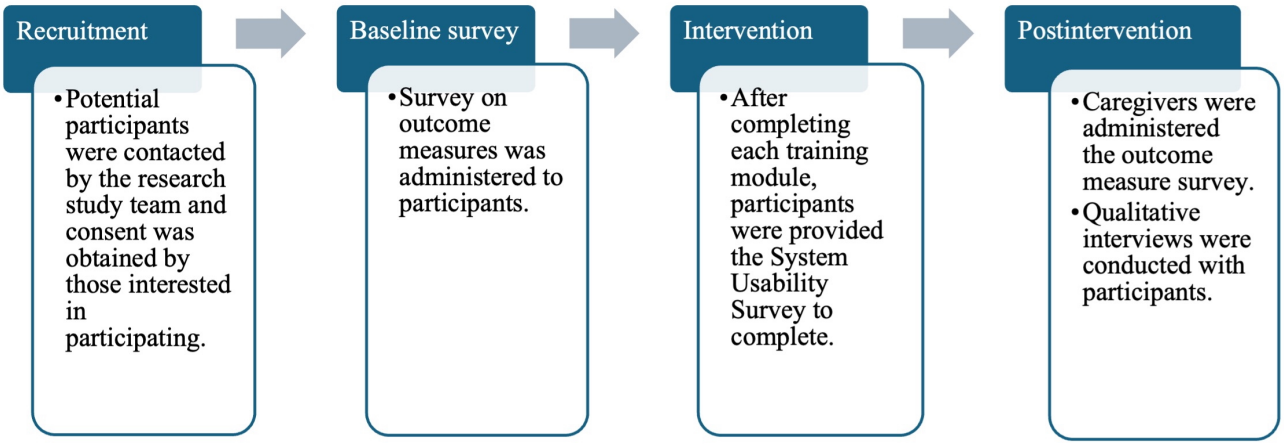
Data Collection

Timeline of Data Collection

The data for this study were collected pre- and postintervention and throughout the intervention period. For reference to the data collection timeline, see Figure 3.

XSL•FO
RenderX

Figure 3. Time of data collection.



Quantitative Data Collection

We collected outcomes data on the RMBPC, Preparedness for Caregiving Scale, and System Usability Scale via REDCap surveys in Spanish and English. The Spanish versions of these scales have been validated and found reliable [20,22,23]. Research assistants on the project delivered a link to the surveys via email or SMS text message to caregivers. These surveys were provided to participants at baseline and postcompletion of the training modules.

Qualitative Data Collection

All participants who reached the end of the study were invited to participate in a semistructured interview conducted in Spanish or English. All interviews with participants were conducted over the phone with bilingual research team members (CG and LZ). The interview guide included questions about their opinions on the program training components, suggestions for improving course content and design, and the quality of the Spanish translation of training modules (Multimedia Appendix 1). The average length of interviews was approximately 30 minutes. Interviews with participants were audio recorded, and the transcription of the interviews was done by a professional transcription service.

Quantitative Data Analysis

We analyzed quantitative data using descriptive statistics. Based on National Institutes of Health guidance, we did not conduct outcome analysis for effect size estimates since an effect size estimated from a pilot study is inherently unstable and would not provide helpful estimation for power calculations [17].

In the descriptive analysis, we only included study participants who completed the outcome assessments at baseline and 8 weeks. If study participants had missing values in their responses (either because they skipped an item or responded with “Don’t Know/Not Applicable”), then we imputed their missing values. The missing data were imputed by taking the highest possible score that participants could obtain if they answered all items and then multiplying that score by the quotient of the participant’s actual score divided by the highest possible score they could have received from the items they answered.

Qualitative Data Analysis

To analyze qualitative data, we used qualitative analysis methods outlined by Miles et al [24]. Interview transcripts were coded and analyzed in their original language (English or Spanish) using Dedoose (version 9.2.007), a cloud app for managing, analyzing, and presenting qualitative and mixed method research data. First, bilingual team members (MAM, CG, and MR) read the transcripts and wrote brief notes of what was in the data. Next, the first (MAM) and second (CG) authors independently coded transcripts using inductive codes. They convened weekly with the senior author (MR) to discuss the application of codes, refine the codebook consisting of inductive codes, and reach a consensus on coding discrepancies. After coding three transcripts, the codebook was comprehensive, and a few additional codes needed to be added. The first author individually coded the remaining transcripts, with guidance from the senior author as needed. Descriptive and in vivo codes were used in the coding process.

The first author identified preliminary themes based on the inductive codes by collating and reviewing excerpts of each code. During this process, relationships among codes and relationships within and among themes were assessed. A collection of relevant excerpts for each theme was pulled from the data while reviewing coded excerpts. Recommended areas of assessment in the third stage of Barrera and Castro’s [16] framework on cultural adaptations of evidence-based interventions, such as satisfaction with intervention elements and suggestions for improvement, were used to help guide theme creation. The preliminary themes were iteratively refined with assistance from the second and third authors (CG and MR) to ensure that subthemes within the themes were relevant to the parent theme and that the final themes were distinguishable.

Results

Participant Characteristics

In total, 16 participants enrolled in the pilot study of the culturally adapted STAR-VTF training modules. Table 1 provides demographic characteristics of the study participants. Demographic characteristics were collected at baseline.

XSL•FO
RenderX

Table . Participant characteristics.

Characteristics	Study participants (n=16)
Age, mean (SD)	51.2 (13.2)
Gender, n (%)	
Man	4 (25)
Woman	12 (75)
Hispanic or Latino origin, n (%)	
Mexican, Mexican American, Chicano or Chicana	10 (62)
Puerto Rican	1 (6)
Other	5 (31)
Race, n (%)	
Black or African American	1 (6)
White	11 (69)
Other	3 (19)
Prefer not to answer	1 (6)
Highest level of education completed, n (%)	
Primary level education	1 (6)
Secondary level education or GED[a] or equivalent	2 (12)
Some college	2 (12)
Associate's degree	2 (12)
Trade or vocational training	1 (6)
Bachelor's degree or higher	8 (50)
Household income before taxes (US $), n (%)	
Less than 15,000	1 (6)
Between 15,000 and 19,999	1 (6)
Between 20,000 and 24,999	1 (6)
Between 25,000 and 34,999	1 (6)
Between 35,000 and 49,999	2 (12)
Between 50,000 and 74,999	2 (12)
100,000 or above	6 (37)
Prefer not to answer	2 (12)
Current occupational status, n (%)	
Employed	9 (56)
Unemployed	3 (19)
Stay-at-home caregiver	2 (12)
Retired	2 (12)
Current marital status, n (%)	
Married	12 (75)
Divorced	1 (6)
Separated	1 (6)
Never married	2 (12)
Primary language spoken at home, n (%)	
English	8 (50)
Spanish	8 (50)

XSL•FO
RenderX

Characteristics	Study participants (n=16)
Received STAR-VTF intervention materials in Spanish, n (%)	7 (44)
Caregiver's health insurance plan, n (%)	
Insurance through employer	8 (50)
Insurance purchased directly	1 (6)
Medicare	3 (19)
Medicaid	1 (6)
No Insurance	3 (19)
Person living with dementia health insurance plan, n (%)	
Insurance purchased directly	1 (6)
Medicare	9 (56)
Medicaid	5 (31)
TRICARE or military health care	1 (6)
Persons living with caregiver, n (%)[b]	
Spouse or partner	12 (75)
Children or grandchildren older than 18 years of age	5 (31)
Children or grandchildren younger than 18 years of age	4 (25)
Relatives (besides spouse or partner and children or grandchildren)	7 (44)
Other	1 (6)
Household size, mean (SD)	3.4 (1.5)
Devices owned, n (%)[b]	
Smartphone	15 (94)
Tablet	11 (69)
Laptop	11 (69)
Computer	7 (44)
Internet access, n (%)	
Cellular data plan	5 (31)
Broadband internet	10 (62)
Satellite internet	1 (6)
Caregiver's relationship to the person living with dementia, n (%)	
Spouse or partner	3 (19)
Adult child	9 (56)
Other close family member	3 (19)
Friend or other nonrelative	1 (6)
Gender of person living with dementia, n (%)	
Man	4 (25)
Woman	12 (75)
Lives with person living with dementia, n (%)	10 (62)
Years in the caregiving role, mean (SD)	4.3 (3.28)
Caregiving hours per week, mean (SD)	60.2 (50.9)

[a]GED: General Educational Development.

[b]Column percent may not add up to 100% because categories are not mutually exclusive.

XSL•FO
RenderX

Feasibility

We achieved satisfactory feasibility based on recruitment, retention, and intervention completion measures. We recruited 16 study participants. The initial recruitment goal was to recruit 20 participants, and we reached 80% of our goal (16/20). Of the participants recruited, 94% (15/16) of participants were retained for all 8 weeks of the pilot study. The proportion of participants who completed all 7 training modules and participated in all study procedures (qualitative exit interviews, pre- and postoutcome assessments, and weekly usability surveys) was 75% (12/16).

Usability

Usability was assessed for each training module. Average scores and SDs for each module were calculated (Table 2). The average System Usability Score across all 7 modules is 76.7 out of 100, which is considered "good" usability, indicating that caregivers perceive the STAR-VTF training modules as easy to use [24,25].

Table . System Usability Scores.

Week	System Usability Score, mean (SD)
1 (n=13)	77.8 (10.8)
2 (n=12)	77.7 (12.9)
3 (n=10)	74.6 (19.2)
4 (n=13)	73.7 (16.3)
5 (n=12)	75.8 (14.3)
6 (n=9)	79.2 (14.5)
7 (n=11)	78.0 (13.7)

Acceptability

The results of the semistructured interviews provided information about the cultural adaptations of the STAR-VTF training modules. The qualitative results also provide a further understanding of what changes are needed to revise the cultural adaptations further.

Theme 1: Latino Caregivers Were Satisfied With the STAR-C Intervention Content and Design Elements of the Training Modules and Accompanying Workbook

Overview

Latino caregivers were generally satisfied with the content covered in the STAR-C intervention, including components on dementia education, ABC problem-solving, and caregiver support strategies. They also expressed satisfaction with the design of the training modules and the accompanying workbook.

Dementia Education Component

Latino caregivers found the dementia education information beneficial in understanding their relatives' dementia condition and how it can progress. There was also a consensus among most caregivers that they liked the empathetic messaging embedded within the dementia education component. Caregivers felt that it was helpful to be reminded that their family member has a condition of the brain and that some of the things their family member says or does are not on purpose, and to be encouraged to have patience and understanding. A caregiver providing care to their father said the following about the messages reminding caregivers to be empathetic:

> *And I felt like it was an important reminder because I guess just like with anything else, even as a parent of young children, I get caught up in the day-to-day and the getting things done, and there's deadlines, and there's work and home life and all sorts of things. And so it is nice to have that reminder that it's not just ... not to take things personally and not to be so quick to react. But just a gentle reminder, there may be another reason why my dad is saying something or not doing something. And yeah, just not to take it personally, so that was a helpful reminder.* [ID 11]

ABC Problem-Solving Component

Many Latino caregivers expressed that the ABC problem-solving strategy of the training was helpful and found the examples used to demonstrate how to implement the strategy to be relatable to their own caregiving experience. One of the training modules demonstrated a hypothetical caregiver using the ABC problem-solving strategy to address a person living with dementia wandering. When asked about their thoughts on this example, one of the caregivers stated:

> *So I thought it was very relatable. And then sort of just kind of showcasing what happened, I think, made it relatable in the sense that you can sort of put yourself in that situation. And then for me, I was like, "Okay, yeah. I could see how this happens, and then what can you do to utilize those ABCs to then basically make it useful for your own personal experience in a similar situation."* [ID 14]

Latino caregivers also shared how the ABC problem-solving strategy was successful in helping them manage their relative's behavioral and psychological symptoms of dementia and provided anecdotes of their experience applying the ABC problem-solving strategy. For example, one of the caregivers described how they used the strategy to address their mother's experience of washing dishes. The caregiver said:

> *So my mom would get really upset and would complain of dizziness around the same time every day. And I thought that was odd. And then when I was doing this and I took the ABC, I realized that the*

XSL•FO
RenderX

reason my mom was getting upset was because she would wash the dishes at the same time every day at, and sometimes I would go by and tell her, "Oh, turn the water higher," or, "Oh, put more soap." And I think me trying to coach her was upsetting her because it made her feel like she didn't know what she was doing. So she would go, and she would say she's dizzy, and she would start to cry. And then because I noticed the timing, what happened before and after, I stopped doing that. I just let her wash the dishes. [ID 27]

Caregiver Support Strategies Component

Latino caregivers overwhelmingly related to the caregiver support strategies, encouraging caregivers to practice self-care. Caregivers expressed the importance of caring for themselves to care for others. They also acknowledged the importance of taking time for themselves. For example, a Spanish-speaking caregiver providing care to his wife explained the reasons why it was important for caregivers to prioritize their own self-care.

It's good that you made us see what we need, also we need to take care of ourselves to be able to provide care. In my case, my wife needed me to be good so I could take care of her well. One needs to sleep well and eat well to give the best help to the sick. (translated quote) [ID 06]

Design of Training Modules and Accompanying Workbook

Latino caregivers were satisfied with the design of the training modules and the accompanying workbook. Some caregivers said they liked the pacing of the training modules and the images used. One caregiver shared the following about the training modules:

I like the pace of it, and I liked the way that it was almost like going through a PowerPoint presentation. I liked that there were slides that kind of bulleted the info that I could either read along to or that it included images that were ... let me see. Images of people of color and not just an elderly White family. [ID 11]

Some of the caregivers described the workbook as a helpful resource to refer to information and strategies that had been covered in the training modules. For example, one of the caregivers planned to continue referencing the workbook after the research study ended. They stated:

So I think for me it's [the workbook] going to be a helpful guide now that I don't have access to the videos to go through it and remind me of things and I can share it with my other family members who are also trying to help with caring with my mom. So I feel like the workbook is kind of the legacy of those classes. [ID 27]

Theme 2: Suggestions for Improving Program Content and Making Training More Engaging

Overview

Although Latino caregivers were generally satisfied with the training modules, some caregivers shared suggestions that they felt would enhance the intervention. The areas for improvement that Latino caregivers suggested are the inclusion of information on additional resources and improving the interactivity of training modules to make videos more engaging.

Including Information on Additional Resources

Latino caregivers expressed wanting additional information on resources outside of the program that they can access to help support people living with dementia. Although some resources were presented in the training modules, caregivers shared they would like more resources on dementia care and community organizations serving people living with dementia and their families. For example, one caregiver shared:

Maybe what might be helpful is if you had ... now this would be just ... I guess you'd have to make it unique for every state, but maybe give them some resources of where they might be able to reach out to. There's different Medicaid programs that might be eligible for. There might be local dementia associations that they could contact, those types of things, right? So resources that people might be able to reach out to, especially ones that offer services in Spanish. [ID 17]

Improve Interactivity to Make Training Modules More Engaging

Some Latino caregivers provided suggestions for improving the interactivity of training modules to make learning program content more engaging. A suggestion to improve the interactivity of program content was to include knowledge checks or short quizzes in which caregivers can test their knowledge of the material covered. As one caregiver shared:

Having something that's a little more interactive where it's not just a video, but almost like a course where you watch a video and then ... having both options like an actual paper workbook or having the option to select multiple choice or just to keep the viewer engaged in the video. So even if it's like, "Here's a summary. Which of these apply?" Select A, B, C, or D, all of the above, or something like that, just so that it really just keeps the viewer engaged. [ID 11]

Latino caregivers also felt there were areas in which the training modules could be improved visually to make them more engaging. A suggestion that was shared by caregivers was to include video reenactments to demonstrate how strategies learned can be used in real-world situations. For example, one caregiver shared they liked the ABC strategy examples in the training module but mentioned that if they were recreated in a video depicting a role-play of the scenario, it would be more engaging. As the caregiver explains:

XSL•FO
RenderX

> *I feel like an actual video example of the dialogue would go a lot further than the pictures with the text-next-to-them pictures. If I remember correctly, that's what they were. They were pictures of the humans, of the woman and the person, her caregiver, and then there was just dialogue text next to them. It made sense. It just seems very dry. It seemed very dry. And I didn't really feel like I was paying attention as much as I should have in those moments.* [ID 24]

Some caregivers also suggested the importance of including testimonials from fellow caregivers of patients with dementia about their experiences with caregiving and how they applied the STAR-VTF training. One caregiver shared how incorporating caregiver testimonials can help demonstrate how to implement strategies covered in the training, such as the importance of caregiver self-care:

> *I think probably maybe interviewing some caregivers and having them share their journey from not knowing what's going on with your mother ... she's crazy ... to understanding what's happening and the realization that it is an illness and examples of how they continued on with taking care of their health. It's very stressful. It's very stressful. I mean, it creates a physical response in your own body. And so I think it's just the affirmation from others that what they've done to manage their own time and their own feelings and examples like that and reassuring us that it's okay. It's okay to leave them and do something fun like go to a movie with your husband instead of feeling guilty about it.* [ID 28]

Theme 3: Caregivers Share Dementia Care Information With Others and Suggest Outreach Strategies

Overview

Latino caregivers shared how they have increased awareness of dementia care strategies by sharing the information they learned in the training modules with relatives. Caregivers generally shared program content by word of mouth or by sharing the link to the training modules that were provided to them. Latino caregivers suggested additional avenues to spread awareness of the STAR-VTF program, such as working with community partners and using social media.

Sharing Information With Relatives

Multiple Latino caregivers shared that they valued the information provided in the training and that they shared this information with family members who assist in the care of the person living with dementia and with family members who are in a similar caregiving role. For example, one caregiver said:

> *So I shared the videos with a group chat for the seven-week period. It was a group chat of my caregivers for my Nana, including myself. So it's going to be my sisters, my two sisters, my two cousins, my father, my uncle, and my aunt. And all of us took part in that group chat, and I shared it with them every time I got a new video because I felt like it was something that we could all benefit from. We all share the same struggle right now. And even though we're not in the same place, as far as mentally, taking care of my grandmother, this is a step to getting us all on the same page. And it seemed like the right thing to do to share the information that I was getting and make sure that my family understood why I was doing what I was doing.* [ID 24]

Community Outreach Strategies

To further improve the STAR-VTF program outreach, some caregivers shared that it is important to partner with trusted Latino community organizations, community health workers, religious centers, and caregiving organizations. As one caregiver explains, to reach Latino communities, it is important to consider partners who have established trust in the community to establish the credibility of the program. The caregiver said:

> *So often Latinos they feel more trusted with people that they know, so family and friends or church or organizations. So if you can tap into those kind of communities to share it, I think, will probably be your best bet in getting the word out. Even my mom, she's very hesitant to go to the doctor. I have to take her because she grew up in a generation where they kind of mistrust doctors, but they trust their sisters or cousins. And if they come at them with information, they're more apt to use it than if it comes from a really formal setting. So to say that word of mouth, that community-centered family groups, community groups, churches, any kind of groups that have a good connection to the Latino community, I think they're more apt to use the information you're providing.* [ID 27]

Some caregivers also suggested advertising the intervention in medical facilities and partnering with health professionals. As one Latino caregiver shared:

> *Working with the community health workers, they can put it out (information on intervention) or maybe have something in the rural area at the doctors. I know they have a lot of doctors. So if it's just one doctor's office, if you have some handout or a little small TV screen that's on loop playing some of that information, it would be good.* [ID 09]

Some Latino caregivers also shared the importance of the STAR-VTF program building a digital presence by posting the training modules on social media, such as Facebook, to reach Latino populations. Caregivers believed delivering the STAR-VTF program through these platforms would increase the reach of the program. For example, one caregiver shared:

> *You can put them (training videos) on Facebook, on social media, and then when you're on social media, you could pop up and ... because you could be on social media on the weekends, and then your study would pop up, or your videos would pop up on the weekends and ... because people are always on social media day, morning, night, weekends, holidays, so if*

XSL•FO
RenderX

you had your content available on social media, then they would be able to view it. [ID 08]

Preliminary Outcomes

The average RMBPC score for a person living with dementia behavior problems changed from 28.24 (SD 18.31) at baseline to 21.34 (SD 16.76) postintervention, indicating a decreased frequency of behavior problems (Table 3). The average RMBPC score for the caregiver's reaction to behavior problems changed from 40.40 (SD 16.68) at baseline to 37.21 (SD 14.60) postintervention, indicating decreased caregiver reaction to behavior problems. Furthermore, the average Preparedness for Caregiving Scale score changed from 1.98 (SD 0.33) at baseline to 2.43 (SD 0.32) postintervention, indicating increased preparedness for caregiving. The findings show a trend toward decreased behavior problems in the person living with dementia, decreased caregiver reaction to behavior problems, and increased caregiver preparedness after versus before the intervention.

Table . Results for RMBPC[a] and Preparedness for Caregiving Scale at weeks 0 and 8.

	Week 0, mean (SD)	Week 8, mean (SD)
RMBPC		
Problem Frequency, overall (n=14)	28.24 (18.31)	21.34 (16.76)
Problem Frequency, memory subscale (n=14)	7.70 (5.35)	4.98 (3.63)
Problem Frequency, disruption subscale (n=14)	10.24 (7.95)	6.39 (6.20)
Problem Frequency, depression subscale (n=14)	10.73 (10.36)	9.23 (10.35)
Caregiver Reaction, overall (n=13)	40.40 (16.68)	37.21 (14.60)
Caregiver Reaction, memory subscale (n=13)	19.21 (5.70)	18.64 (6.52)
Caregiver Reaction, disruption subscale (n=13)	9.35 (6.54)	8.14 (5.27)
Caregiver Reaction, depression subscale (n=13)	11.85 (7.52)	10.43 (5.88)
Preparedness for Caregiving Scale Score (n=14)	1.98 (0.33)	2.43 (0.32)

[a]RMBPC: Revised Memory and Behavior Problem Checklist.

Discussion

Principal Results

This study aimed to pilot-test the culturally adapted STAR-VTF training modules among Latino caregivers of people living with dementia. Delivering the culturally adapted STAR-VTF training modules to Latino caregivers was feasible, with a high proportion of participants completing the study procedures. The training modules were also found to be easy for Latino caregivers to use based on the high average of System Usability Scale scores across modules.

The qualitative analysis of interviews with Latino caregivers who completed the 8-week program revealed their high satisfaction and acceptability with the training modules and components on dementia education, ABC problem solving, and caregiver support strategies. The study also identified areas for improvement in the training modules, such as providing caregivers with information on additional resources and enhancing the engagement of training videos. The study also provided valuable insights into expanding the intervention's reach, including through word of mouth, partnerships with community organizations, and social media. These findings will significantly contribute to the refinement of the culturally adapted STAR-VTF training modules for Latino caregivers. The culturally adapted STAR-VTF intervention was also found to have positive changes in the outcomes of behavioral problems among people living with dementia and caregiver preparedness among the small sample of Latino caregiver participants. Further research is needed to examine the efficacy of the culturally adapted STAR-VTF intervention.

Comparison With Prior Work

This is the first pilot study that tests a cultural adaptation of the STAR-VTF evidence-based intervention for Latino caregivers. Our findings on the feasibility of the culturally adapted STAR-VTF intervention are comparable to other interventions for dementia caregivers delivered digitally, which were also feasible [26,27]. However, unlike these studies, this study consists exclusively of Latino participants, which provides further information on the feasibility of digital interventions for Latino caregivers. Preliminary findings of positive mental health outcomes and improvements in caregiver preparedness are similar to findings from a systematic review of literature on internet-based interventions aimed at supporting family caregivers of people living with dementia [28]. The studies presented in the systematic review of digital dementia caregiving interventions reported reductions in caregiver stress, depression, and strain and improvements in self-efficacy [28]. However, conclusions from these studies and this pilot study should be made with caution due to study design limitations, as these studies consist of quasi-experimental, small randomized controlled trials, pre-post studies, and feasibility studies.

Our findings on acceptability and Latino caregivers' satisfaction with training components in this pilot study were consistent with the identified needs and wants of caregivers who participated in a qualitative study to test low-fidelity versions of the culturally adapted training modules [8]. In the pilot study, Latino caregivers expressed how the module on dementia education was especially beneficial to their understanding of dementia; this finding aligned with caregivers' perspectives of wanting more dementia education in the previous study of

XSL•FO
RenderX

low-fidelity training modules [8]. The desire to learn about ADRD from Latino caregivers of people living with dementia is also consistent with findings from a qualitative study that examined barriers and facilitators to increasing engagement of Hispanics or Latinos in clinical trials on ADRD [29]. In this study, a majority of caregivers expressed an eagerness to receive information about ADRD, including details on the different types of ADRD, specific risk factors for Latino communities, and hereditary characteristics [29].

Another aspect of the training modules that resonated with Latino caregivers, which was highlighted in our previous qualitative study, was the inclusion of reminders for caregivers to be empathetic toward people living with dementia. Our findings on caregiver satisfaction with the caregiver support strategy component are also similar to findings from our previous study in which caregivers highlighted the importance of self-care [8]. The need for information on self-care by Latino caregivers was also highlighted in a qualitative study that explored the needs of Latino caregivers of relatives with early-onset Alzheimer disease [30]. This study found that among their 27 participants, they overwhelmingly expressed the need for additional information concerning self-care, especially mental health care strategies and stress management [28]. Caregivers' high satisfaction with the training components piloted in this study suggests that we effectively integrated feedback from Latino caregivers to meet their needs in the training content.

Regarding our findings on areas of improvement, Latino caregivers' needs for additional dementia care–related resources are consistent with findings from a mixed methods study on the health information–seeking behavior of Latino caregivers of people living with dementia. The study on health-seeking behaviors found that Latino caregivers desire to seek various types of dementia information, including the availability of community resources, health care services, and treatment [31]. Caregivers in this pilot study also suggested improving training videos by making them more engaging with video reenactments of caregiving scenarios, knowledge checks, and caregiver testimonials. This aligns with findings from our prior qualitative study wherein Latino caregivers who viewed the low-fidelity modules expressed the idea of having real people in the videos to depict caregiver scenario examples [8]. Multiple studies have found that using video vignettes in the form of telenovelas and Spanish soap operas is an acceptable and effective method of delivering health information to Latino caregivers [32-34]. Furthermore, a study that evaluated the acceptability of a web-based course for caregivers of older adults found that caregivers liked having postmodule quizzes, and they also had a caregiver suggest the use of caregiver testimonials in course modules [35]. Our findings show that these preferences in web-based training are also warranted among a sample of Latino caregivers.

Limitations

Although preliminary outcomes were positive, this study had several limitations. First, due to the small budget, we were not able to pilot-test the full STAR-VTF intervention, which consists of weekly training modules and weekly 30-minute phone check-ins with a coach. It is during these check-ins that coaches review caregivers' ABC plans. Future pilot studies will include the coaching component, which will also ensure that the ABC plans are being developed appropriately. Second, general limitations of pilot studies apply, such as the inability to determine efficacy and estimate effect sizes for power calculations of a larger-scale study [17]. The target sample size of 20 for this pilot study was determined based on logistical constraints, as it was the maximum feasible within the available budget and study timeline. Third, the pre-poststudy design is subject to limitations due to the lack of a comparison group, which relies on conclusions formed from the temporal relationship of the measurements to the intervention. This type of study does not allow researchers to control for other changes that could have occurred around the same time as the intervention that could have influenced outcome results [36]. A limitation of the qualitative portion of this study is that caregivers had trouble recalling details about earlier training modules. A caregiver shared that since it had been over a month since they viewed the first couple of training modules, it was difficult to remember some of the specific details of the content covered. When conducting interviews with participants, the research assistants (CG and LZ) would sometimes need to remind caregivers about earlier training modules. The time gap between viewing the earlier training modules and being interviewed can result in recall bias, and participants may have omitted important details regarding their experience. Another limitation of this study is that 10 of the 14 interviews were conducted by the same research assistant who recorded the voiceover for the modules. If participants recognized that it was the same person, it may have led to social desirability bias, as participants may have provided answers that would be viewed favorably by the research assistant.

Conclusions

This pilot study found the culturally adapted STAR-VTF intervention to have high levels of feasibility, usability, and acceptability. In general, caregivers found the training modules beneficial and relatable and were highly satisfied with the cultural adaptations to the STAR-VTF intervention. The feedback from the qualitative interviews with Latino caregivers will help inform decisions during the fourth stage (Adaptation refinement) of the Barrera and Castro [16] framework on the cultural adaptation of evidence-based interventions to refine future iterations and meet Latino caregiver needs. Although future research is still necessary to test the efficacy of the culturally adapted STAR-VTF program, this pilot study provided valuable information about the feasibility, usability, and caregiver experiences with the intervention. This study also contributes to the broader knowledge of modifying evidence-based interventions to better serve Latino caregivers of persons with dementia.

XSL•FO
RenderX

Acknowledgments

This work was funded by the National Institute on Aging (NIA) of the National Institutes of Health (NIH) under award U54AG063546, which funds NIA Imbedded Pragmatic Alzheimer's and Alzheimer Disease–Related Dementias Clinical Trials Collaboratory (NIA IMPACT Collaboratory). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Conflicts of Interest

None declared.

Multimedia Appendix 1
English and Spanish interview guides.
[DOCX File, 28 KB - aging_v8i1e66053_app1.docx]

References

1. Martinez IL, Gonzalez EA. Care v. caring: obligation, duty, and love among Latino Alzheimer's family caregivers. J Appl Gerontol 2022 Jul;41(7):1744-1751. [doi: 10.1177/07334648221084998] [Medline: 35484911]
2. Quiroz YT, Solis M, Aranda MP, et al. Addressing the disparities in dementia risk, early detection and care in Latino populations: highlights from the second Latinos & Alzheimer's symposium. Alzheimers Dement 2022 Sep;18(9):1677-1686. [doi: 10.1002/alz.12589] [Medline: 35199931]
3. Vega B, Vega WA, Wu S, Resendez J, Jin H, Aranda M. Latinos and Alzheimer's disease. Innov Aging 2018 Nov 1;2(suppl_1):216-216. [doi: 10.1093/geroni/igy023.794]
4. Balbim GM, Magallanes M, Marques IG, et al. Sources of caregiving burden in middle-aged and older Latino caregivers. J Geriatr Psychiatry Neurol 2020 Jul;33(4):185-194. [doi: 10.1177/0891988719874119] [Medline: 31510848]
5. Martinez IL, Gonzalez EA, Quintero C, Vania MJ. The experience of Alzheimer's disease family caregivers in a Latino community: expectations and incongruences in support services. J Gerontol B Psychol Sci Soc Sci 2022 Jun 1;77(6):1083-1093. [doi: 10.1093/geronb/gbab170] [Medline: 34543386]
6. Alzheimer's Association. 2023 Alzheimer's disease facts and figures. Alzheimers Dement 2023 Apr;19(4):1598-1695. [doi: 10.1002/alz.13016]
7. Luchsinger JA, Tipiani D, Torres-Patiño G, et al. Characteristics and mental health of Hispanic dementia caregivers in New York City. Am J Alzheimers Dis Other Demen 2015 Sep;30(6):584-590. [doi: 10.1177/1533317514568340] [Medline: 25635108]
8. Garcia CN, Duran MC, Ramirez M. Refining cultural adaptations of a behavioral intervention for Latino caregivers of people living with dementia: qualitative interview study in Washington State. JMIR Aging 2024 Jan 11;7(1):e53671. [doi: 10.2196/53671] [Medline: 38206663]
9. McCurry SM, Logsdon RG, Mead J, et al. Adopting evidence-based caregiver training programs in the real world: outcomes and lessons learned from the STAR-C Oregon translation study. J Appl Gerontol 2017 May;36(5):519-536. [doi: 10.1177/0733464815581483] [Medline: 25873454]
10. Lindauer A, Croff R, Mincks K, et al. "It took the stress out of getting help": the STAR-C-telemedicine mixed methods pilot. Care Wkly 2018;2:7-14. [doi: 10.14283/cw.2018.4] [Medline: 30393783]
11. Teri L, McCurry SM, Logsdon R, Gibbons LE. Training community consultants to help family members improve dementia care: a randomized controlled trial. Gerontologist 2005 Dec;45(6):802-811. [doi: 10.1093/geront/45.6.802] [Medline: 16326662]
12. Teri L, McKenzie G, Logsdon RG, et al. Translation of two evidence-based programs for training families to improve care of persons with dementia. Gerontologist 2012 Aug;52(4):452-459. [doi: 10.1093/geront/gnr132] [Medline: 22247431]
13. Lindauer A, McKenzie G, LaFazia D, et al. Using technology to facilitate fidelity assessments: the Tele-STAR caregiver intervention. J Med Internet Res 2019 May 24;21(5):e13599. [doi: 10.2196/13599] [Medline: 31127721]
14. Satz S, White Q, Watkin LG. Utilizing technology and integrated services to support family caregivers in rural communities. Innov Aging 2023 Dec 21;7(Supplement_1):618-618. [doi: 10.1093/geroni/igad104.2015]
15. Ramirez M, Duran MC, Penfold RB, Pabiniak CJ, Hansen KE, Ornelas IJ. STAR-Caregivers Virtual Training and Follow-up: a cultural adaptation for Latino caregivers of people with dementia. Transl Behav Med 2023 Sep 12;13(9):625-634. [doi: 10.1093/tbm/ibad028] [Medline: 37130336]
16. Barrera M, Castro FG. A heuristic framework for the cultural adaptation of interventions. Clin Psychol Sci Pract 2006;13(4):311-316. [doi: 10.1111/j.1468-2850.2006.00043.x]
17. Pilot studies: common uses and misuses. NCCIH. URL: https://www.nccih.nih.gov/grants/pilot-studies-common-uses-and-misuses [accessed 2024-06-05]
18. Brooke J. SUS: a quick and dirty usability scale. In: Usability Evaluation In Industry: CRC Press; 2014:4-7.

XSL•FO
RenderX

19. Teri L, Truax P, Logsdon R, Uomoto J, Zarit S, Vitaliano PP. Assessment of behavioral problems in dementia: the revised memory and behavior problems checklist. Psychol Aging 1992 Dec;7(4):622-631. [doi: 10.1037//0882-7974.7.4.622] [Medline: 1466831]
20. Salvia MG, Dawidowski A, Schapira M, et al. Spanish Revised Memory and Behavior Problems Checklist Scale (SpRMBPC): trans-cultural adaptation and validation of the RMBPC questionnaire. Int Psychogeriatr 2011 Sep;23(7):1160-1166. [doi: 10.1017/S1041610210002073] [Medline: 21294937]
21. Preparedness for Caregiving Scale. HIGN. URL: https://hign.org/consultgeri/try-this-series/preparedness-caregiving-scale [accessed 2024-06-05]
22. Gutierrez-Baena B, Romero-Grimaldi C. Development and psychometric testing of the Spanish version of the Caregiver Preparedness Scale. Nurs Open 2021 May;8(3):1183-1193. [doi: 10.1002/nop2.732] [Medline: 33340272]
23. Sevilla-Gonzalez MDR, Loaeza LM, Lazaro-Carrera LS, et al. Spanish version of the System Usability Scale for the assessment of electronic tools: development and validation. JMIR Hum Factors 2020 Dec 16;7(4):e21161. [doi: 10.2196/21161] [Medline: 33325828]
24. Miles MB, Huberman AM, Saldana J. Qualitative Data Analysis: A Methods Sourcebook: Sage Publishing; 2014.
25. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J User Exper 2009;4(3):114-123 [FREE Full text]
26. Loi SM, Tropea J, Gaffy E, et al. START-online: acceptability and feasibility of an online intervention for carers of people living with dementia. Pilot Feasibility Stud 2022 Feb 16;8(1):41. [doi: 10.1186/s40814-022-00999-0] [Medline: 35172894]
27. Kwok I, Lattie EG, Yang D, et al. Acceptability and feasibility of a socially enhanced, self-guided, positive emotion regulation intervention for caregivers of individuals with dementia: pilot intervention study. JMIR Aging 2023 Sep 6;6(1):e46269. [doi: 10.2196/46269] [Medline: 37672311]
28. Hopwood J, Walker N, McDonagh L, et al. Internet-based interventions aimed at supporting family caregivers of people with dementia: systematic review. J Med Internet Res 2018 Jun 12;20(6):e216. [doi: 10.2196/jmir.9548] [Medline: 29895512]
29. Marquez DX, Perez A, Johnson JK, et al. Increasing engagement of Hispanics/Latinos in clinical trials on Alzheimer's disease and related dementias. Alzheimers Dement (N Y) 2022;8(1):e12331. [doi: 10.1002/trc2.12331] [Medline: 35910673]
30. Withers M, Cortez-Sanchez K, Herrera J, Ringman JM, Segal-Gidan F. "My backpack is so heavy": experiences of Latino caregivers of family with early-onset Alzheimer's. J Am Geriatr Soc 2021 Jun;69(6):1539-1547. [doi: 10.1111/jgs.17091] [Medline: 33713418]
31. Ramirez M, De Anda S, Jin H, Herrera JR, Wu S. Health information-seeking behavior of Latino caregivers of people living with dementia: a mixed-methods study. J Appl Gerontol 2023 Aug;42(8):1738-1748. [doi: 10.1177/07334648231163430] [Medline: 36932723]
32. Cruz-Oliver DM, Malmstrom TK, Fernández N, Parikh M, García J, Sanchez-Reilly S. Education intervention "Caregivers Like Me" for Latino family caregivers improved attitudes toward professional assistance at end-of-life care. Am J Hosp Palliat Care 2016 Jul;33(6):527-536. [doi: 10.1177/1049909115584315] [Medline: 26019262]
33. Gallagher-Thompson D, Tzuang M, Hinton L, et al. Effectiveness of a fotonovela for reducing depression and stress in Latino dementia family caregivers. Alzheimer Dis Assoc Disord 2015;29(2):146-153. [doi: 10.1097/WAD.0000000000000077] [Medline: 25590939]
34. Kajiyama B, Fernandez G, Carter EA, Humber MB, Thompson LW. Helping Hispanic dementia caregivers cope with stress using technology-based resources. Clin Gerontol 2018;41(3):209-216. [doi: 10.1080/07317115.2017.1377797] [Medline: 29236621]
35. Rottenberg S, Williams A. Web-based delivery of the caregiving essentials course for informal caregivers of older adults in Ontario: mixed methods evaluation study. JMIR Aging 2021 Jun 15;4(2):e25671. [doi: 10.2196/25671] [Medline: 34128815]
36. Aggarwal R, Ranganathan P. Study designs: part 4—interventional studies. Perspect Clin Res 2019;10(3):137-139. [doi: 10.4103/picr.PICR_91_19] [Medline: 31404185]

Abbreviations

ABC: activator-behavior-consequence
ADRD: Alzheimer and related dementia
REDCap: Research Electronic Data Capture
RMBPC: Revised Memory Behavior Problem Checklist
STAR-C: STAR Caregivers
STAR-VTF: STAR Virtual Training Follow Up

XSL•FO
RenderX

Edited by A Quialheiro; submitted 04.09.24; peer-reviewed by A Lindauer, K Meyer; revised version received 03.02.25; accepted 21.04.25; published 10.06.25.

Please cite as:
Mariscal MA, Garcia C, Zavala L, Ramirez M
Culturally Adapted STAR-Caregivers Virtual Training and Follow-Up for Latino Caregivers of People Living With Dementia: Single-Arm Pre-Post Mixed Methods Study
JMIR Aging 2025;8:e66053
URL: https://aging.jmir.org/2025/1/e66053
doi:10.2196/66053

© Miguel Angel Mariscal, Celeste Garcia, Lily Zavala, Magaly Ramirez. Originally published in JMIR Aging (https://aging.jmir.org), 10.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Identifying Unmet Needs of Informal Dementia Caregivers in Clinical Practice: User-Centered Development of a Digital Assessment Tool

Olga A Biernetzky[1], BA, MSc, PhD; Jochen René Thyrian[2,3], Prof Dr rer med; Melanie Boekholt[2]; Matthias Berndt[2]; Wolfgang Hoffmann[2,3], MPH, Prof Dr med; Stefan J Teipel[1,4], Prof Dr med; Ingo Kilimann[1,4], Dr med

[1]Deutsches Zentrum fuer Neurodegenerative Erkrankungen, Rostock/Greifswald, Gehlsheimer Strasse 20, Rostock, Germany

[2]Deutsches Zentrum fuer Neurodegenerative Erkrankungen, Rostock/Greifswald, Greifswald, Germany

[3]Institute for Community Medicine, Section Epidemiology and Community Health, University Medicine Greifswald, Greifswald, Germany

[4]Department for Psychosomatic and Psychotherapeutical Medicine, Rostock University Medical Center, Rostock, Germany

Corresponding Author:
Olga A Biernetzky, BA, MSc, PhD
Deutsches Zentrum fuer Neurodegenerative Erkrankungen, Rostock/Greifswald, Gehlsheimer Strasse 20, Rostock, Germany

Abstract

Background: Despite the increasing interventions to support family caregivers of people with dementia, service planning and delivery is still not effective.

Objective: Our study aimed to develop a digitally-supported needs assessment tool for family caregivers of people with dementia that is feasible, time-efficient, understood by users, and can be self-completed in the primary care setting.

Methods: The development of the unmet needs assessment tool was part of a cluster-randomized controlled trial examining the effectiveness of a digitally supported care management programme to reduce unmet needs of family caregivers of people with dementia (GAIN [Gesund Angehörige Pflegen]) and was conducted in 3 phases. Using an iterative participatory approach with informal caregivers, health care professionals including general practitioners, neurologists, psychologists, psychiatrists, nurses, and Alzheimer Society representatives, we developed a digital self-completion unmet needs assessment tool focusing on informal caregivers' biopsychosocial health und quality of life in connection to their caregiver responsibilities. Data were collected through group discussions, written feedback, protocols, think-aloud protocols, and interviews, and analyzed thematically.

Results: Data from 27 caregivers, including caregivers of people with dementia (n=18), health care professionals (n=7), and Alzheimer Society representatives (n=2) were collected. Thematic analysis identified 2 main themes: content of the assessment tool and usability and handling of the digital tablet-based assessment tool. The feedback provided by the stakeholders led to new aspects and changes to make the tool comprehensive, easy to read, and easy to handle. The overall mean completion time was reduced from the initial 37 minutes to 18 minutes, which renders the assessment tool fit to be self-completed in waiting rooms of primary care practices or other settings.

Conclusions: The input of the 3 stakeholder groups has supported the development of the assessment tool ensuring that all aspects considered important were covered and understood and the completion of the assessment procedure was time-efficient and practically feasible. Further validation of the assessment tool will be performed with the data generated as part of the GAIN trial.

Trial Registration: ClinicalTrials.gov NCT04037501; https://clinicaltrials.gov/study/NCT04037501

(JMIR Aging 2025;8:e59942) doi:10.2196/59942

KEYWORDS

unmet needs; assessment development; family caregivers of people with dementia; dementia; need; Alzheimer; self-guided; self-reported; caregiver; informal care; spousal care; interview; qualitative; thematic; usability; mHealth; tablet; self-completed; aging; patient care; health interventions; care giver; digital health; ehealth; digital assessment; memory

Introduction

The World Health Organization (WHO) estimates 55 million people are living with dementia worldwide, with an increase to 78 million by 2030 [1]. In Germany, approximately two-thirds of the 1.8 million people with dementia are cared for at home by an informal caregiver, most often a relative [2]. Numerous studies have shown that caring for a relative with dementia is associated with a multitude of time- and resource-intensive challenges [3-5]. Emotional stress, social isolation, depression,

XSL•FO
RenderX

and financial burden are among the negative consequences that dementia family caregivers have reported [6-10]. To alleviate these consequences, interventions for the reduction of caregivers' burden have been developed [11-14], such as telephone-based as well as multicomponent psychosocial and individualized support programs [11,13]. With the growing consensus that caregiver support is an integral part of the care of patients with dementia, there was also an increase in interventions to support informal caregivers [12]. However, each caregiver's situation is different and requires an individual set of interventions. Individualized interventions are still underused in care practice as they require time and dementia-specific expertise that is often lacking in primary care. A systematic and comprehensive yet individualised unmet needs assessment tool would be useful for health care professionals (HCPs) to better plan services and interventions in clinical settings as it would flag areas where the individual caregiver needs support.

In a mixed methods study, Stirling et al [15] used Bradshaw taxonomy of need [16] to explore the relationship between different types of caregiver service need. Bradshaw and Care [16] defined categories of need as measures of professionally identified carer burden (normative need), service use (expressed need), carer's stated need (felt need) and the comparison of groups using services with groups who do not (comparative need). It is argued that no single measure of the above needs is likely to capture all carer's unmet needs. In the light of person-centred care, the normative need has been increasingly challenged as paternalistic and inappropriate [15,17].

Therefore, a focus on caregivers' stated need (felt need) as well as service use (expressed need) would be an appropriate starting point to consider for an unmet needs assessment tool. An individualized, digitally supported self-assessment instrument can aid in the efficient identification of health and support areas that should be addressed from the perspective of the informal caregivers. This allows collaborative intervention planning and reduces the risk of a paternalistic doctor-patient relationship [18]. If a digital system is to truly ensure collaborative service planning, the digital assessments instrument must be developed with the participation of the users themselves.

The direct input of family caregivers and other stakeholders into the content, structure, and handling of the digital assessment system provides crucial feedback regarding its content validity as well as comprehensiveness and technical usability. Through a greater level of involvement of stakeholders, health care services can be improved and be more applicable in real-world clinical settings and facilitate implementation [19]. A systematic review of instruments used to assess the needs of family dementia caregivers identified 36 instruments that were described in detail or in part in the 70 publications included [20]. The authors reported that only one instrument was partially validated to assess the needs of family dementia caregivers, namely, the Carer's Needs Assessment for Dementia (CNA-D), a semistructured interview not intended for clinical use with a completion time of about an hour. For conceptual clarity, researchers recommended establishing a theoretical model or framework to organize the diverse needs of family dementia caregivers [21]. In addition, researchers argue that a focus on the detection of changes in family caregivers' needs throughout disease progression is important [20,22].

Therefore, the aim of this study was to develop a digital unmet needs assessment tool that can be used in primary care settings. We conducted the study in collaboration with informal and formal caregivers and other stakeholders. Our objectives were to (1) create a user-friendly and family dementia caregiver specific digitally supported assessment tool, (2) develop an assessment process that is understandable and comprehensive yet time-efficient, and (3) create an assessment tool that can be used at multiple locations such as general practitioners' (GPs) offices and memory clinic waiting rooms.

Methods

Overview

The development of the assessment tool was in preparation of the upcoming randomized controlled intervention trial examining the effectiveness of a digitally supported care management intervention for family caregivers (German title: Gesund Angehörige Pflegen [GAIN]). The trial was supported by the Federal Joint Committee (G-BA [Gemeinsamer Bundesausschuss/Innovationsausschuss]). The funding code (FKZ) is 01VSF18030).

Participants

We involved individuals from 3 stakeholder groups, namely, family dementia caregivers, health care professionals (general practitioners, neurologists, psychiatrists, psychologists, and nurses), and Alzheimer Society representatives. We approached HCPs of a medical university and memory clinic staff, as well as Alzheimer Society representatives, and asked for their feedback (phase 1). Family dementia caregivers who were visiting a memory clinic with their relative or friend with dementia were asked whether they would like to provide feedback on the paper-based version of the assessment instrument regarding the content, clarity, and comprehensiveness (phase 2) as well as the usability of the digital version of the assessment tool as a self-completion assessment instrument on a PC tablet (phase 3). No further demographic data were collected.

Procedure

We used an iterative user-centered participatory approach with 3 phases. Previously used questionnaires [20,23,24] assessing unmet needs and care interventions from previous dementia care management studies [4,25-31] in Germany were used as starting point.

Phase 1

In phase 1, the authors developed a list of previously used questionnaires assessing unmet needs in family dementia caregivers as well as a list of available care interventions offered in Germany. The selection of questionnaires was based on a comprehensive review of the literature covering the databases OVID, MEDLINE, and PsycInfo searching for unmet needs, caregiver needs assessments methods and family dementia caregivers. With respect to available care interventions, we used work from previous dementia care management studies

XSL•FO
RenderX

management studies [4,25-31] offering a list of interventions to family dementia caregivers in Germany. We used Germany-based intervention studies as these are health care system specific. Criteria for the inclusion of questionnaires were the content they covered (eg, quality of life, psychosocial factors, and health-related domains), their previous use in research and interventions, and their completion time. We also considered questionnaires based on either their recommendation by the EU Joint Programme in Neurodegenerative Research (JPND) Working Group on Longitudinal Cohorts or common use in larger German trials such as IDemUck [32], DelpHi-MV (Dementia: life- and person-centered help in Mecklenburg-Western Pomerania) [26,28], intersec-CM [33,34], or DemNet-D [35,36]. These questionnaires are validated and allowed comparison with German and international studies. The authors then further discussed these selections in meetings, compared 5 separate versions of a possible unmet needs assessment instrument and checked whether items would cover interventions offered as part of the dementia care management conducted in previous primary care studies [4,26,28,37,38]. With this approach, we wanted to ensure that the operationalization of problem-centered needs were connected to available services within the German health care system and could therefore be addressed in this framework accordingly.

The authors selected different assessment versions aiming to cover all biopsychosocial aspects and compared different scales for caregiver burden. The need items were then compared and matched with the list of informal caregiver interventions from previous studies—these were marked according to a match or no match. This allowed the calculation of percentages of “items covered” from the dementia caregiver intervention list for each version of questionnaires. The five versions of combinations of questionnaires were checked by 3 authors and researchers for plausibility. The results of this item-intervention matching process were used to identify which versions covered most unmet needs with matched possible interventions. These are shown in Table 1.

XSL•FO
RenderX

Table . Proportion of match with family dementia caregiver interventions.

Versions and interventions		Match, %
Version 1		47
	FIMA[a]	
	URN[b]	
	EQ-5D-5L[c]	
	ZBI-7[d]	
	LSNS-6[e]	
Version 2		17
	FIMA	
	URN	
	EQ-5D-5L	
	BIZA-D[f]	
Version 3		96
	CANE[g]	
	EQ-5D-5L	
	ZBI-7	
	LSNS-6	
Version 4		94
	CANE	
	DQoL-OC[h]	
Version 5		39
	HABC[i]	
	DQoL-OC	
	ZBI-7	

[a]FIMA: Questionnaire for the Use of Medical and Non-Medical Services in Old Age [39].

[b]URN: Caregiver unmet resource needs scale [40].

[c]EQ-5D-5L: Health-related quality of life [41].

[d]ZBI-7: Zarit Burden Interview [42,43].

[e]LSNS-6: Lubben Social Network Scale [44,45].

[f]BIZA-D: The Berlin inventory of the burden on relatives - dementia – Module 3,5,and 6 [46].

[g]CANE: Camberwell Assessment of Need for the Elderly [47,48].

[h]DQoL-OC: The Dementia Quality of Life Scale for Older Family Carers [49].

[i]HABC: Healthy Aging Brain Care Monitor—Caregiver Version [50].

Versions 3 and 4 covered most items, whereas versions 1, 2, and 5 covered less than 50% of the dementia-specific items. The two versions covering most unmet needs were then presented and discussed at an advisory board meeting with input from HCPs (eg, neurologists, psychiatrists, psychologists, and nurses) and Alzheimer Society representatives. Meeting minutes were recorded of all points raised and discussed. This input was used to select the final version of the assessment tool.

Phase 2

In phase 2, the chosen assessment version covering most unmet needs was tested in a memory clinic. A tablet computer was deemed most appropriate to serve as digital device due to its size, weight, flexibility, and ease of handling. Based on a previous study on the use of a tablet-based digital expert system [26,28,30], we had taken several decisions on the hardware basis of our system before the participatory part of our study started. The content and usability aspects, however, were developed together with future users. The authors approached family dementia caregivers in a memory clinic to provide feedback and suggestions on a paper-based version of the assessment tool. The changes were primarily focused on the Camberwell Assessment of Need for the Elderly (CANE) questionnaire, as well as the demographic and informal caregiver-specific questions, as these were items that could still be adapted. Family caregivers who were interested in

XSL•FO
RenderX

contributing were given a short version comprising 7 questions of the CANE questionnaire. The researcher noted the caregivers' comments and thoughts in a think-aloud protocol. Together with a researcher, they were asked to read the questions and share their thoughts with respect to each question and their understanding thereof. The focus was on comprehension, words and sentence structure, answer options, descriptions, and additional comments. In the second round, both informal caregivers and health care professionals were approached to provide feedback on the whole printed version of the assessment instrument. Those who agreed to provide feedback were asked to mock-complete the assessment tool and write down their notes and questions. In this round, informal caregivers and HCPs were asked to measure the time needed to complete the assessment procedure. We provided written questions such as "Are the questions clear and easy to understand?"; "Where did you encounter problems when completing the assessment?"; "Did you have issues understanding the questions?"; and "Are there questions or areas that should be in the assessment set but are currently missing?" Based on the feedback received, the first draft of the assessment set was modified resulting in a second draft. When adapting the first draft, our focus was on the assessment procedure being clear and easy to understand for family caregivers in terms of the way questions are phrased. Based on the feedback we received, several parts of the assessment set were adapted and duplicates were removed by the authors resulting in the final draft.

Phase 3

In phase 3, a digitalized tablet-based version of the final version of the assessment tool was tested for its technical usability. In this phase, we asked both informal caregivers and HCPs to test the digital assessment tool with a focus on the handling of the tablet. Data was collected through interviews and think-aloud protocols. Family caregivers in a memory clinic were asked to complete the digital assessment tool in the waiting room and share their thoughts and comments with a study nurse or researcher who protocolled the comments. Similarly, HCPs of a medical university and a memory clinic were asked to complete the digital assessment procedure and write down feedback on the handling and any possible bugs or problems they noticed while testing the digital system.

Analysis

All data collected through written feedback, protocols, think-aloud protocols, and interviews were analyzed thematically following the steps by Braun and Clarke [51]. The six phases involved were (1) familiarizing ourselves with the data, (2) generating initial codes, (3) constructing themes, (4) reviewing themes, (5) defining and naming themes, and (6) producing the report [51].

Ethical Considerations

The study is conducted in accordance with the criteria (valid at present) of the Declaration of Helsinki, the ICH-guidelines for Good Clinical Practice, the Memorandum for Safeguarding Good Scientific Practice (German Research Foundation), and the International Ethical Guidelines for Biomedical Research Involving Human Subjects. Ethical approval has been obtained from the Ethical Committee of the University Medicine Greifswald (BB120/2019) and the Ethical Committee of the University Medicine Rostock (A2020/0013). The trial is registered at ClinicalTrials.gov (NCT04037501).

Results

Results of the development of the assessment tool of the 3 phases are divided in the subsections presented below. The process of phases 2 and 3 was iterative comprising several rounds through which we collected feedback, made changes based on feedback received and then moved to the next round of collecting feedback.

Participants

A total of 27 family caregivers of people with dementia (n=18), HCPs (n=7), and Alzheimer Society representatives (n=2) participated. Among the HCPs were general practitioners, neurologists, psychologists, psychiatrists, and nurses. The mean age of the 27 family caregivers was 57 (SD 13.78; range 35‐88) years and 9 (50%) were women. The HCPs' and representatives' mean age was 41 (SD 8.05; range 28‐50) and 6 out of 9 (67%) were women. As this was a participatory approach, no further demographic data were collected.

Phase 1: Structure and Content of the Assessment Tool

After the input from the advisory board including neurologists, psychiatrists, psychologists, nurses, and Alzheimer Society representatives, the final structure of the assessment instrument was selected. It was the one with the highest percentage of matches between problem-centered needs and the associated interventions (69/72, 96%; version 3; see Table 1).

The unmet needs assessment tool comprised several validated questionnaires with a total of 57 questions examining a person's self-reported demographic information, use of medical and nonmedical services, unmet needs (CANE), health-related quality of life (EQ-5D-5L), caregiver burden (Zarit Burden Interview [ZBI]), and social support (Lubben Social Network Scale [LSNS]). It can be completed using a digital version or a paper-based version. In total, the assessment set covers 51 health- and support-related aspects.

Domains covered by the unmet needs assessment tool are the following: (1) health and care, (2) employment, (3) information and knowledge, (4) emotional support, (5) social support, and (6) caregiver burden.

The structure of the unmet needs assessment tool can be seen in Table 2.

XSL•FO
RenderX

Table . Structure of the unmet needs assessment tool.

Instrument	Domain
Socio-demographic information	—[a]
Caregiver specific information	Individual situation of informal caregiver
CANE[b]	Unmet needs I—Information, physical, mental
EQ-5D-5L	Health-related quality of life
LSNS[c]	Unmet needs II—Social network
ZBI[d]	Unmet needs III—Caregiver burden

[a]Not available.

[b]CANE: Camberwell Assessment of Need for the Elderly.

[c]LSNS: Lubben Social Network Scale.

[d]ZBI: Zarit Burden Interview.

The number of unmet needs addressed the participants' medical needs, home care needs, psychosocial needs, and needs connected to the caregiver role. This needs assessment instrument included selected parts of the CANE [47].

Health-related quality of life was assessed using the EQ-5D-5L [41]. This instrument comprises 5 dimensions: mobility, self-care, usual activities, pain or discomfort, and anxiety or depression. Each dimension has 5 levels varying from no problems to extreme problems. Each level corresponds to a 1-digit number that expresses the level selected for that dimension ranging from 1 to 5 with higher numbers indicating more severe problems. The digits for the 5 dimensions can be converted into an overall index by an not publicly accessible algorithm. The index describes the participant's health-related quality of life. It also contains a vertical visual analogue scale (EQ VAS), with endpoints that are labelled "Best imaginable health state" and "Worst imaginable health state." The VAS reflects the patient's own judgement and can be used as a quantitative measure of health outcome.

Social support was assessed using LSNS-6 [44,45]. This scale is a self-report measure of social engagement including contact to and interaction with family and friends on a 6-item scale. Total scores range from 0 to 30 with an equally weighted sum of the 6 items. The family and friends subscales include questions regarding the number of friends and family one has regular contact with as well as their availability for help and support in private matters. High scores indicate strong social networks.

Informal caregiver burden was assessed using the 7-item version of the Zarit-Burden Interview (ZBI-7). The short version ZBI is a caregiver self-report measure to examine burden, which is associated with functional and behavioral impairments in the social, psychological and physiological context, and home care situation [42,43]. It contains 7 items using a 5-point scale. Response options range from 0 (Never) to 4 (Nearly to Always). Total scores range from 0 indicating no burden to 28 indicating severe burden.

Phases 2 and 3: Feedback Received on the Assessment Tool

Themes

A total of two main themes were derived from the analysis: (1) content of the assessment instrument and (2) digital tablet usability for the assessment tool. Each theme had further subthemes and subtheme categories and were divided in either informal caregivers or HCPs and Alzheimer Society representatives as these groups had different areas they paid special attention to. Almost all subthemes had 1 to 3 subtheme categories, see Table 3.

XSL•FO
RenderX

Table . Themes, subthemes, and subtheme categories.

Theme and subtheme	Subtheme category
1. Content of the assessment procedure	—[a]
Informal caregivers	—
1.1 Comprehension	• 1.1.1 Simplification of questions • 1.1.2 Simplification of answer options • 1.1.3 Separate questions for caregiver health and caregiver care responsibilities in relation to the person with dementia
1.2 Assessment structure	• 1.2.1 No multiple questions in a table • 1.2.2 Questions whether help is needed in a specific domain for clear identification of unmet needs
1.3 Reduction of completion time	• 1.3.1 Fatigue
1.4 Areas important to caregivers	—
Health care professionals	
1.1 Assessment structure	• 1.1.1 Order of instruments • 1.1.2 Focus on caregivers' health and care responsibilities
1.2 Reduction of completion time	—
2. Digital tablet usability of the assessment tool	—
Informal caregivers	
2.1 Simple layout	• 2.1.1 Large font and buttons
2.2 Handling	• 2.2.1 Manual answer selection
Health care professionals	
2.1 Simple layout	• 2.1.1 One question per screen • 2.1.2 Option to go back and forth within the assessment procedure • 2.1.3 Progression bar
2.2 Handling	—

[a]Not applicable.

Content of the Assessment Instrument

Regarding the assessment's content, informal caregivers provided information on 4 subthemes regarding the comprehension, assessment structure, reduction of completion time, and areas important to them.

Comprehension (Informal Caregiver Perspective)

Regarding the clarity of the assessment procedure to them, caregivers commented on the simplification of questions and answer options. Instead of "This is an unmet need" or "This is a need met" (as suggested in the CANE) they preferred to answer "Yes" or "No," for instance. They also commented on being sometimes confused whether a question was referring to themselves or to the person diagnosed with dementia indicating that questions had to be framed specifically.

Assessment Structure

Regarding the assessment's structure, informal caregivers commented that multiple questions in a table format were confusing. In addition, some commented that they would like to self-identify the areas where they needed help rather than having professionals assume that they needed help or support in an area, which was incorporated into the assessment procedure by first asking whether there is a problem in a specific area and secondly asking whether the respondent receives enough help with respect to this problem. A yes or no response to these questions would then indicate the presence or absence of an unmet need.

HCPs and Alzheimer Society representatives considered the order of instruments within the assessment procedure and argued that the instrument covering most unmet needs (CANE) should be moved to the front to prevent that important health care domains received less attention due to fatigue. They also reported that the informal caregivers' own health as well as factors influencing their health due to their caregiver role should be a focus since a caregiver's unmet needs can arise from two sources: either from the personal needs of the family caregiver or from his or her care responsibilities for the people with dementia.

Reduction of Completion Time (Informal Caregiver Perspective)

The majority of the informal caregivers commented on the length of the initial versions of the assessment tool and reported

XSL•FO
RenderX

that they became a bit tired towards the end of the assessment procedure. A caregiver stated: "For the last few questions my focus was a bit lost" (informal caregiver, woman).

Areas Important to Caregivers (Informal Caregiver Perspective)

Caregivers commented on some of the areas that they found particularly important such as psychological support, mobility, social life, and psychoeducation.

A caregiver stated:

> *I would say that there should be a question for psychological support. Self-help groups should be offered locally and it should be asked whether advice or support is needed in this regard. Another thing that I would still find important is to ask how mobile you are and whether you need help in this regard. For example, I don't have a car and have to use public transport to get everywhere. And there are also people who don't even have a driver's license. And if you have to go somewhere and have to pick up things from the pharmacy (eg, incontinence pads), these are sometimes large packages that are also heavy and you have to get them home first. It would be a good idea to ask whether you need help in this regard.* [Informal caregiver, woman]

Another caregiver reported:

> *The social life is very important, my father is hardly involved in it anymore. Because of his swallowing problems, he can no longer even go to a restaurant, which is difficult.* [Informal caregiver, man]

Another cargiver stated:

> *Since we have not yet come into contact with services, I would like a lot of information so that you know what you have to do. And also so that you know why behavior changes, how people then think and what is going on inside you. I'd be very interested in that.* [Informal caregiver, man]

Regarding the content of the assessment instrument, HCPs and Alzheimer Society representatives provided information on 2 subthemes, namely, assessment structure and reduction of completion time.

Digital Tablet Usability of the Assessment Tool

Regarding the assessment's tablet usability, informal caregivers, HCPs, and Alzheimer Society representatives identified 2 subthemes, namely, simple layout and handling with slightly different subtheme categories.

Simple Layout

Informal caregivers commented on the font size, often taking out their reading glasses to be able to read the questions.

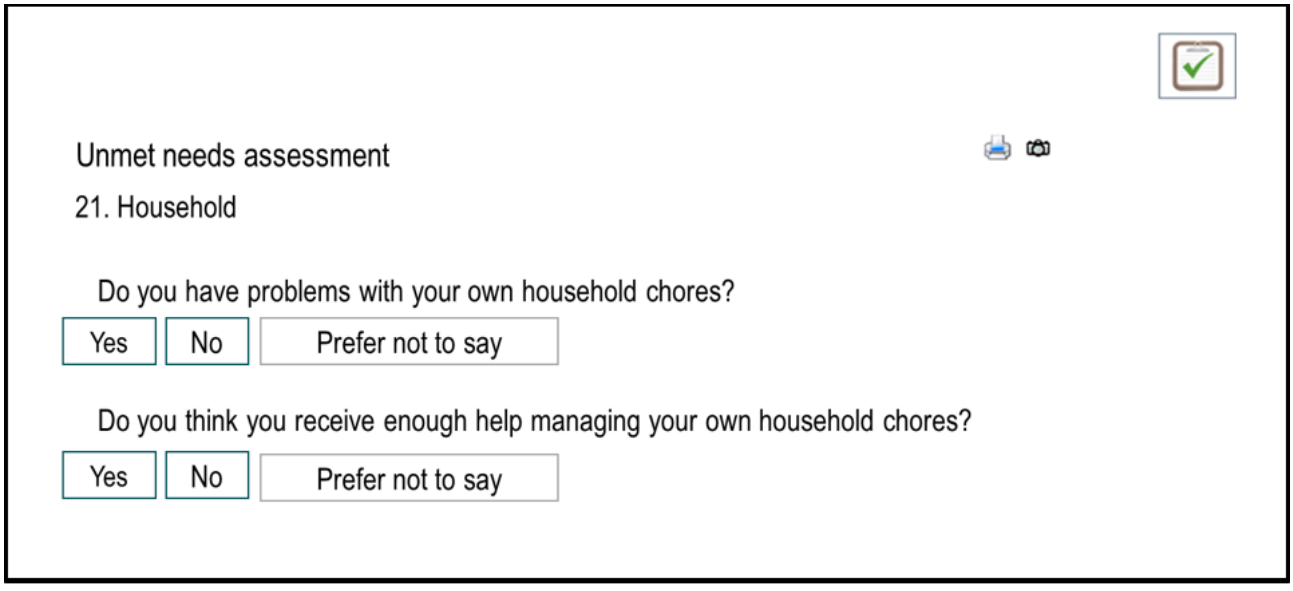
HCPs commented that one question per screen would allow the large font and button size needed. Some clinicians stated that it would be helpful to allow caregivers to go back within the assessment procedure in case they changed their mind regarding a previous answer. Going forward should only be possible if the question on the screen was completed. The majority of HCPs stated that an indication bar within the assessment procedure would be helpful for caregivers to receive information on their progress as well as how much of the assessment procedure is still in front of them.

Handling

The majority of informal caregivers reported that selecting answers by clicking worked well after adjustments were made to the font size and the size of the answer option buttons. Regarding the handling, all HCPs and Alzheimer Society representatives agreed that the selection of answers should be as straight forward as possible allowing caregivers to complete the assessment procedure without technical issues.

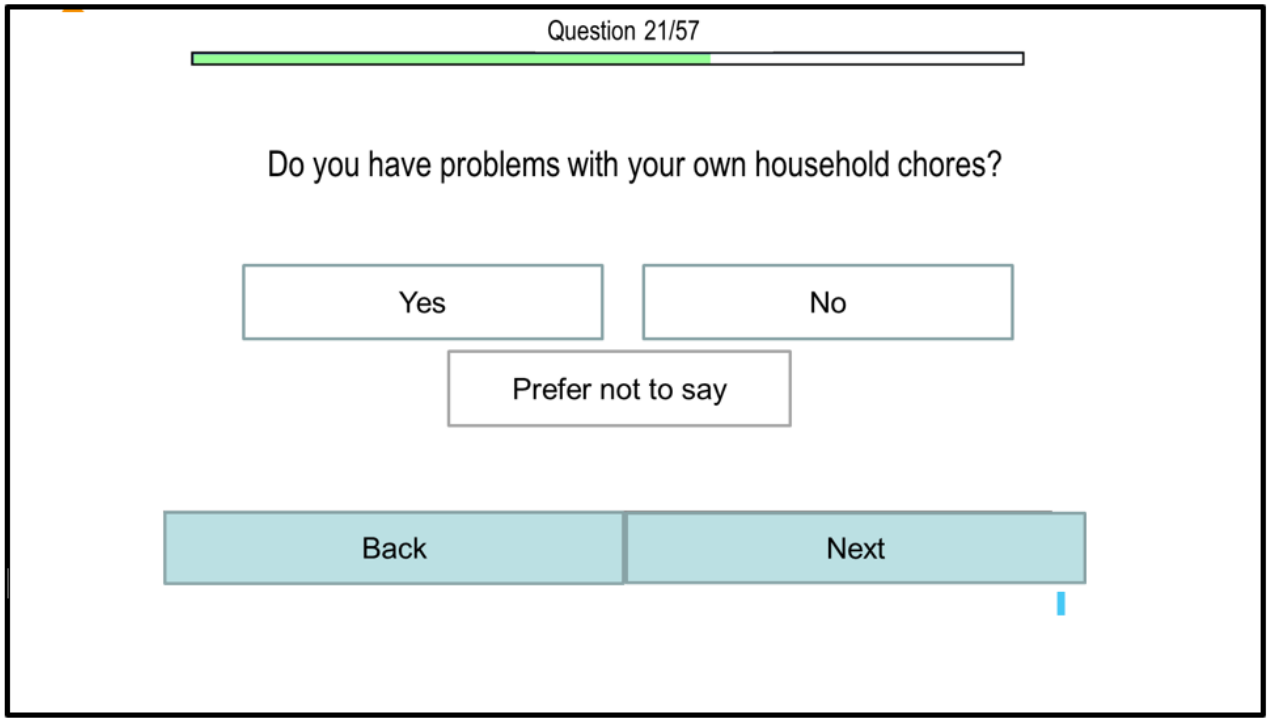
The adaptations made to the digital tablet-based assessment tool are shown in Figures 1 and 2. Figure 1 shows the initial, first version of the tablet-based assessment tool. Figure 2 shows the final tablet-based assessment version after the feedback was incorporated into the development by the IT team.

Figure 1. Outline of the first digital assessment version.



XSL•FO
RenderX

Figure 2. Outline of the final digital assessment version.

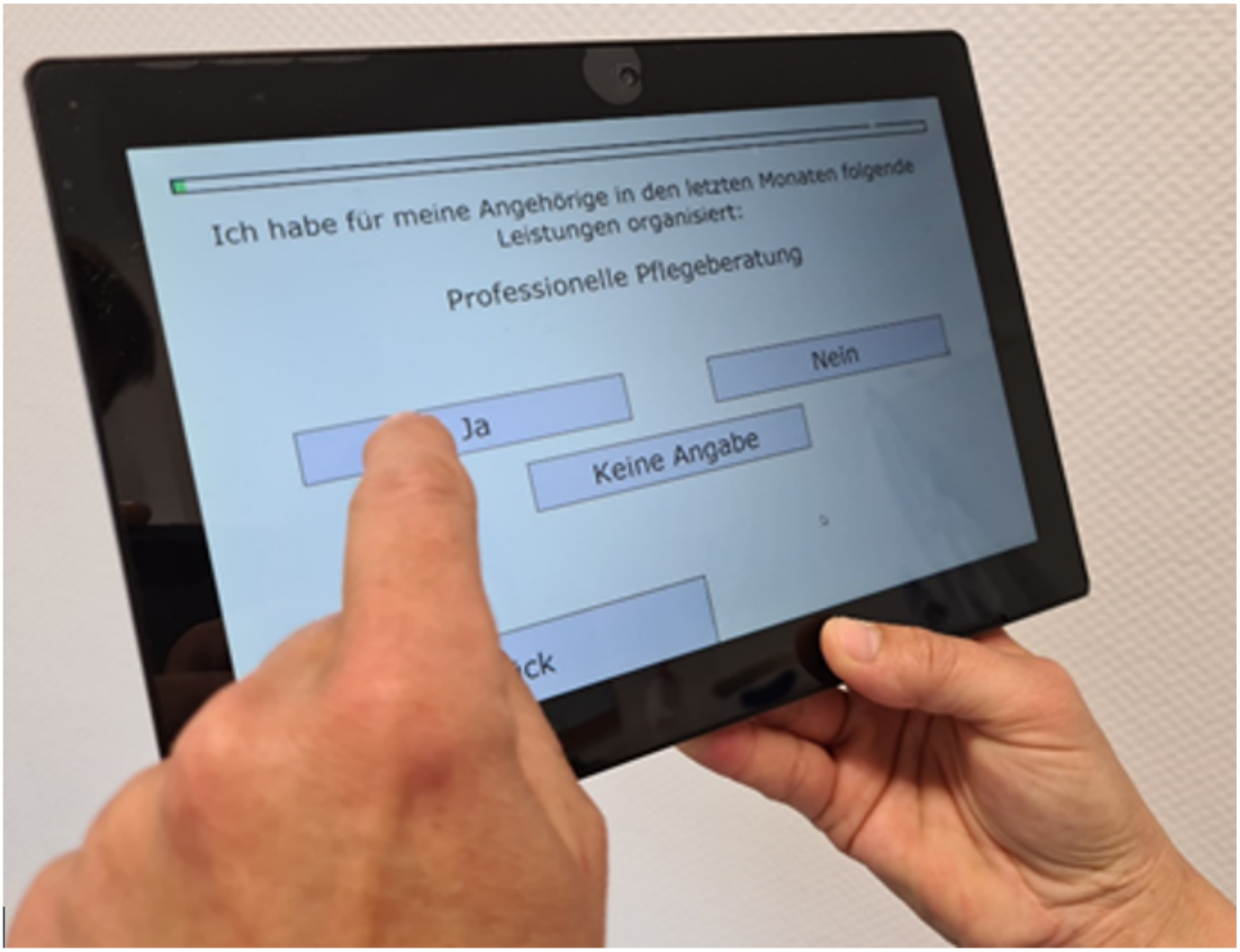


After completion, the outcome of the assessment procedure can be viewed by the study nurse as well as the GP. These professionals can then use the outcome to work together with the informal caregiver toward a reduction of the unmet needs identified. The study nurses provide the informal caregivers with a summary of the consultation where points discussed are listed.

Figure 3 depicts a caregiver filling out the digital version of the assessment tool on a tablet.

XSL•FO
RenderX

Figure 3. An informal caregiver completing the digital tablet-based assessment.



Discussion

Principal Findings

This study describes the phases of an iterative, participatory process we used to develop a digital tablet-based assessment tool designed to comprehensively identify the needs of informal caregivers of people with dementia in clinical practice. Through the collaboration with family dementia caregivers, health care professionals, and Alzheimer Society representatives, we developed a digital unmet needs assessment instrument that consists of a set of standardized instruments but also includes additional questions and adaptations that were incorporated along the iterative development process. The unmet needs assessment tool focuses on needs that can arise from two sources: either from the personal needs of the family caregiver or from the care responsibilities for people with dementia. With the contribution of informal caregivers and other stakeholders, we designed an assessment tool that covers a comprehensive range of needs considered important. At the same time, with the input of the stakeholders, the time to complete the assessment procedure was substantially reduced from the initial 37 minutes to 18 minutes. This indicates that we reached our goal of developing a comprehensive, feasible, and time-efficient assessment device.

With respect to the content of the assessment instrument, we identified important aspects regarding clarity, assessment structure, the reduction of completion time as well as content areas important to informal caregivers. We took the feedback in multiple rounds and incorporated changes along the way. Similar approaches have been conducted successfully by other research teams developing assistive technologies to support self-management of people with dementia [52].

Regarding the digital tablet-based assessment tool, we could incorporate crucial aspects regarding the layout and the handling of the tablet. The difference between our first digital assessment version compared with the final digital assessment version illustrates the importance of involving multiple perspectives, especially of those for whom the assessment instrument or technology is intended. In the literature, there is an increase of studies involving patients and stakeholders in the development or testing of technology [52-54].

The participatory approach allowed us to incorporate aspects into the development of the digital assessment tool that we, as researchers, would have missed otherwise, as can be seen in the first digital tablet-based assessment version in Figure 1 when compared with the final version in Figure 2. The input of informal caregivers, health care professionals, and Alzheimer Society representatives has broadened our perspective and has given us important hints and advice towards aspects that are important for a digital, self-completion unmet needs assessment tool. They identified questions that were unclear, suggested possible answer options, and gave us ideas on how to simplify

XSL•FO
RenderX

the navigation throughout the assessment procedure. Some researchers have used similar procedures for the development of instruments to assess patients' needs [24,47] or to facilitate implementation [19]. A review by Fischer et al [55] reports on the importance of user involvement, particularly the involvement of older adults in technology design. In their review of 40 empirical studies published between 2014 and 2018, these authors examined the consequences of involving older adults in technology design and stated that learning, adjusted design, and improved sense of participation were outcomes frequently stated in studies [55]. Not only did they report that involvement facilitated learning about the needs of older adults but it also led to the generation of new technology ideas such as a companion robot, for instance, which was not in the researchers' mind before involving older adults [55]. Newly gained insights then lead to iterative adjustments of the prototype design [55]. This was also the process that we followed making adjustments to the unmet needs assessment tool. With respect to the quality of technology developed with user involvement, Kopeć and et al [56] noticed that the overall quality of the technology could be improved. They argued that teams who involved users to develop a mobile app were rated higher by an independent jury in a competition than teams who did not. Older adults appreciated the sense of participation and felt that they were being treated as experts on their own lives and equal partners [57]. Researchers reported that involved users enjoyed their experience and some users described their involvement as "happy memories" [58-60].

In contrast to already existing unmet needs assessment instruments, our digital assessment tool identifies a informal caregiver's self-reported unmet needs and ties them directly to interventions most of which are offered within the German health care system. Through the matching process of possible unmet needs with interventions in phase 1 of our assessment development process, we aimed to overcome the pitfalls of identifying problem-centered needs without having support services to meet them at hand. In addition, the digital assessment tool acknowledges the intertwined nature of the dyad, namely the informal caregiver and the person with dementia focusing on not only the health and support needs of the caregiver but also the caregiver's care responsibilities. Although we did not include the CNA-D, we did compare the content and ensured that no important aspect was missed in the assessment procedure.

Limitations

The stakeholders who contributed to the development of the assessment tool were a convenience sample. We have asked informal caregivers, health care professionals and Alzheimer Society representatives who either visited a memory clinic with their significant other, worked at the memory clinic or who were already known to us due to previous collaborations. Furthermore, the sample is limited to people living in a specific region in Northern Germany (Mecklenburg-Western Pomerania). In addition, our level of involvement was limited to stakeholders sharing their feedback on the assessment tool, which is only 1 of the 3 levels of involvement [61]. The tendency to involve stakeholders only in some rather than all levels of involvement was also highlighted by Fischer et al [55] arguing for further research on the actual practices of user involvement. For future projects, a higher level of involvement is sought, especially since current results indicate that an elaborated, theory-based participatory approach including people with dementia, informal caregivers and regional stakeholders might raise the chances of successful implementation in routine care [19].

Conclusion

Our study has shown the feasibility and success of user involvement in the development of an assessment tool for identifying the unmet needs of people with dementia and their informal caregivers. By presenting our approach, we hope to motivate researchers worldwide to expand stakeholder involvement in the early stages of studies. To facilitate this, the questionnaire will be made available to researchers on the homepage of the Network for Translational Dementia Care Research. Further evaluations, such as the acceptance of the tool from the perspective of the study nurses and GPs, are currently being evaluated and will be published separately.

Acknowledgments

The authors would like to acknowledge the input of all stakeholders for their time and perspective. This work was supported by the Federal Joint Committee (G-BA) – Innovation Fund (Gemeinsamer Bundesausschuss/Innovationsausschuss; funding code: 01VSF18030). The ministry had no role in the design of this study and will not have any role during the implementation, analyses, interpretation of the data, or publication of results.

Data Availability

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

Authors' Contributions

Each author has made substantial contributions to this work, has approved the submitted version and has agreed both to be personally accountable for their own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and the resolution documented in the literature.

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Dementia. World Health Organization. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2025-03-25]
2. Blotenberg I, Hoffmann W, Thyrian JR. Dementia in Germany: epidemiology and prevention potential. Dtsch Arztebl Int 2023 Jul 10;120(27-28):470-476. [doi: 10.3238/arztebl.m2023.0100] [Medline: 37226316]
3. Schulz R, Martire LM. Family caregiving of persons with dementia: prevalence, health effects, and support strategies. Am J Geriatr Psychiatry 2004;12(3):240-249. [doi: 10.1097/00019442-200405000-00002] [Medline: 15126224]
4. Zwingmann I, Hoffmann W, Michalowsky B, et al. Supporting family dementia caregivers: testing the efficacy of dementia care management on multifaceted caregivers' burden. Aging Ment Health 2018 Jul;22(7):889-896. [doi: 10.1080/13607863.2017.1399341] [Medline: 29156941]
5. Michalowsky B, Thyrian JR, Eichler T, et al. Economic analysis of formal care, informal care, and productivity losses in primary care patients who screened positive for dementia in Germany. J Alzheimers Dis 2016;50(1):47-59. [doi: 10.3233/JAD-150600] [Medline: 26639964]
6. Murray J, Schneider J, Banerjee S, Mann A. EUROCARE: A cross-national study of co-resident spouse carers for people with Alzheimer's disease: II—A qualitative analysis of the experience of caregiving. Int J Geriat Psychiatry 1999 Aug;14(8):662-667. [doi: 10.1002/(SICI)1099-1166(199908)14:8<662::AID-GPS993>3.0.CO;2-4]
7. Marin D, Amaya K, Casciano R, et al. Impact of rivastigmine on costs and on time spent in caregiving for families of patients with Alzheimer's disease. Int Psychogeriatr 2003 Dec;15(4):385-398. [doi: 10.1017/s1041610203009633] [Medline: 15000418]
8. Baiyewu O, Smith-Gamble V, Akinbiyi A, et al. Behavioral and caregiver reaction of dementia as measured by the neuropsychiatric inventory in Nigerian community residents. Int Psychogeriatr 2003 Dec;15(4):399-409. [doi: 10.1017/s1041610203009645] [Medline: 15000419]
9. Schneider J, Murray J, Banerjee S, Mann A. EUROCARE: a cross-national study of co-resident spouse carers for people with Alzheimer's disease: I--Factors associated with carer burden. Int J Geriatr Psychiatry 1999 Aug;14(8):651-661. [doi: 10.1002/(SICI)1099-1166(199908)14:83.0.CO;2-B] [Medline: 10489656]
10. Brodaty H, Donkin M. Family caregivers of people with dementia. Dialogues Clin Neurosci 2009;11(2):217-228. [doi: 10.31887/DCNS.2009.11.2/hbrodaty] [Medline: 19585957]
11. Berwig M, Heinrich S, Spahlholz J, Hallensleben N, Brähler E, Gertz HJ. Individualized support for informal caregivers of people with dementia – effectiveness of the German adaptation of REACH II. BMC Geriatr 2017 Dec;17(1):286. [doi: 10.1186/s12877-017-0678-y]
12. Williams F, Moghaddam N, Ramsden S, De Boos D. Interventions for reducing levels of burden amongst informal carers of persons with dementia in the community. A systematic review and meta-analysis of randomised controlled trials. Aging Ment Health 2019 Dec 2;23(12):1629-1642. [doi: 10.1080/13607863.2018.1515886]
13. Shata ZN, Amin MR, El-Kady HM, Abu-Nazel MW. Efficacy of A multi-component psychosocial intervention program for caregivers of persons living with neurocognitive disorders, Alexandria, Egypt: A randomized controlled trial. Avicenna J Med 2017;7(2):54-63. [doi: 10.4103/2231-0770.203610] [Medline: 28469987]
14. Lök N, Bademli K. Pilot testing of the "First You Should Get Stronger" program among caregivers of older adults with dementia. Arch Gerontol Geriatr 2017 Jan;68:84-89. [doi: 10.1016/j.archger.2016.09.006]
15. Stirling C, Andrews S, Croft T, Vickers J, Turner P, Robinson A. Measuring dementia carers' unmet need for services - an exploratory mixed method study. BMC Health Serv Res 2010 Dec;10(1):122. [doi: 10.1186/1472-6963-10-122]
16. Bradshaw J, Care M. Taxonomy of social need: problems and progress. In: Medical Care: Essays on Current Research Seventh Series: Nuffield Provincial Hospitals Trust; 1972.
17. Meaney AM, Croke M, Kirby M. Needs assessment in dementia. Int J Geriatr Psychiatry 2005 Apr;20(4):322-329. [doi: 10.1002/gps.1284] [Medline: 15799078]
18. Asadi-Lari M, Tamburini M, Gray D. Patients' needs, satisfaction, and health related quality of life: towards a comprehensive model. Health Qual Life Outcomes 2004 Jun 29;2:32. [doi: 10.1186/1477-7525-2-32] [Medline: 15225377]
19. Seidel K, Quasdorf T, Haberstroh J, Thyrian JR. Adapting a dementia care management intervention for regional implementation: a theory-based participatory barrier analysis. Int J Environ Res Public Health 2022 Apr 30;19(9):5478. [doi: 10.3390/ijerph19095478] [Medline: 35564877]
20. Novais T, Dauphinot V, Krolak-Salmon P, Mouchoux C. How to explore the needs of informal caregivers of individuals with cognitive impairment in Alzheimer's disease or related diseases? A systematic review of quantitative and qualitative studies. BMC Geriatr 2017 Apr 17;17(1):86. [doi: 10.1186/s12877-017-0481-9] [Medline: 28415968]
21. Kipfer S, Pihet S. Reliability, validity and relevance of needs assessment instruments for informal dementia caregivers: a psychometric systematic review. JBI Evid Synth 2020 Apr;18(4):704-742. [doi: 10.11124/JBISRIR-2017-003976] [Medline: 32813339]

XSL•FO
RenderX

22. Bressan V, Visintini C, Palese A. What do family caregivers of people with dementia need? A mixed-method systematic review. Health Soc Care Community 2020 Nov;28(6):1942-1960. [doi: 10.1111/hsc.13048] [Medline: 32542963]
23. Mansfield E, Boyes AW, Bryant J, Sanson-Fisher R. Quantifying the unmet needs of caregivers of people with dementia: a critical review of the quality of measures. Int J Geriatr Psychiatry 2017 Mar;32(3):274-287. [doi: 10.1002/gps.4642] [Medline: 27981694]
24. Wancata J, Krautgartner M, Berner J, et al. The Carers' Needs Assessment for Dementia (CNA-D): development, validity and reliability. Int Psychogeriatr 2005 Sep;17(3):393-406. [doi: 10.1017/s1041610205001699] [Medline: 16252373]
25. Dreier-Wolfgramm A, Michalowsky B, Austrom MG, et al. Dementia care management in primary care: Current collaborative care models and the case for interprofessional education. Z Gerontol Geriatr 2017 May;50(Suppl 2):68-77. [doi: 10.1007/s00391-017-1220-8] [Medline: 28364258]
26. Thyrian JR, Hertel J, Wucherer D, et al. Effectiveness and safety of dementia care management in primary care: a randomized clinical trial. JAMA Psychiatry 2017 Oct 1;74(10):996-1004. [doi: 10.1001/jamapsychiatry.2017.2124] [Medline: 28746708]
27. Zwingmann I, Michalowsky B, Esser A, et al. Identifying unmet needs of family dementia caregivers: results of the baseline assessment of a cluster-randomized controlled intervention trial. J Alzheimers Dis 2019;67(2):527-539. [doi: 10.3233/JAD-180244] [Medline: 30584136]
28. Thyrian JR, Fiß T, Dreier A, et al. Life- and person-centred help in Mecklenburg-Western Pomerania, Germany (DelpHi): study protocol for a randomised controlled trial. Trials 2012 May 10;13(1):56. [doi: 10.1186/1745-6215-13-56] [Medline: 22575023]
29. Dreier A, Thyrian JR, Eichler T, Hoffmann W. Qualifications for nurses for the care of patients with dementia and support to their caregivers: A pilot evaluation of the dementia care management curriculum. Nurse Educ Today 2016 Jan;36:310-317. [doi: 10.1016/j.nedt.2015.07.024] [Medline: 26277428]
30. Thyrian JR, Winter P, Eichler T, et al. Relatives' burden of caring for people screened positive for dementia in primary care : Results of the DelpHi study. Z Gerontol Geriatr 2017 Jan;50(1):4-13. [doi: 10.1007/s00391-016-1119-9] [Medline: 27534949]
31. Eichler T, Thyrian JR, Hertel J, et al. Unmet needs of community-dwelling primary care patients with dementia in Germany: prevalence and correlates. J Alzheimers Dis 2016;51(3):847-855. [doi: 10.3233/JAD-150935] [Medline: 26890767]
32. Köhler L, Meinke-Franze C, Hein J, et al. Does an interdisciplinary network improve dementia care? Results from the IDemUck-study. Curr Alzheimer Res 2014;11(6):538-548. [doi: 10.2174/1567205011666140618100727] [Medline: 24938504]
33. Nikelski A, Keller A, Schumacher-Schönert F, et al. Supporting elderly people with cognitive impairment during and after hospital stays with intersectoral care management: study protocol for a randomized controlled trial. Trials 2019 Aug 30;20(1):543. [doi: 10.1186/s13063-019-3636-5] [Medline: 31470912]
34. Schumacher-Schönert F, Boekholt M, Nikelski A, et al. Versorgungslücken nach dem Krankenhausaufenthalt schließen: Studienergebnisse [intersec-CM] zum Entlass- und Überleitungsmanagement nach § 39 SGB V für Menschen, die im Krankenhaus kognitive demenzielle Beeinträchtigungen zeigen [in German]. Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen 2024 Apr;185:35-44. [doi: 10.1016/j.zefq.2024.01.001]
35. Gräske J, Meyer S, Schmidt A, et al. Regional dementia care networks in Germany--results from the DemNet-D-Study regarding the quality of life of their users. Pflege 2016;29(2):93-101. [doi: 10.1024/1012-5302/a000477] [Medline: 26974281]
36. Laporte Uribe F, Wolf-Ostermann K, Wübbeler M, Holle B. Care arrangements in dementia care networks: findings from the DemNet-D study baseline and 1-year follow-up. J Aging Health 2018 Jul;30(6):882-903. [doi: 10.1177/0898264317696778] [Medline: 28530156]
37. Zülke A, Luck T, Pabst A, et al. AgeWell.de - study protocol of a pragmatic multi-center cluster-randomized controlled prevention trial against cognitive decline in older primary care patients. BMC Geriatr 2019 Aug 1;19(1):203. [doi: 10.1186/s12877-019-1212-1] [Medline: 31370792]
38. Thyrian JR, Eichler T, Pooch A, et al. Systematic, early identification of dementia and dementia care management are highly appreciated by general physicians in primary care - results within a cluster-randomized-controlled trial (DelpHi). J Multidiscip Healthc 2016;9:183-190. [doi: 10.2147/JMDH.S96055] [Medline: 27143912]
39. Seidl H, Bowles D, Bock JO, et al. FIMA – Fragebogen zur Erhebung von Gesundheitsleistungen im Alter: entwicklung und pilotstudie [Article in German]. Gesundheitswesen 2015;77(1):46-52 [FREE Full text] [doi: 10.1055/s-0034-1372618]
40. Higginson IJ, Gao W, Jackson D, Murray J, Harding R. Short-form Zarit Caregiver Burden Interviews were valid in advanced conditions. J Clin Epidemiol 2010 May;63(5):535-542. [doi: 10.1016/j.jclinepi.2009.06.014] [Medline: 19836205]
41. Lubben J, Blozik E, Gillmann G, et al. Performance of an abbreviated version of the Lubben Social Network Scale among three European community-dwelling older adult populations. Gerontologist 2006 Aug;46(4):503-513. [doi: 10.1093/geront/46.4.503]
42. Zank S, Schacke C, Leipold B. Berliner Inventar zur Angehörigenbelastung - Demenz (BIZA-D). Z Klin Psychol Psychother 2006 Oct;35(4):296-305. [doi: 10.1026/1616-3443.35.4.296]
43. Reynolds T, Thornicroft G, Abas M, et al. Camberwell Assessment of Need for the Elderly (CANE). development, validity and reliability. Br J Psychiatry 2000 May;176(5):444-452. [doi: 10.1192/bjp.176.5.444] [Medline: 10912220]

XSL•FO
RenderX

44. Oliveira DC, Vass C, Aubeeluck A. The development and validation of the Dementia Quality of Life Scale for Older Family Carers (DQoL-OC). Aging Ment Health 2018 May;22(5):709-716. [doi: 10.1080/13607863.2017.1293004] [Medline: 28282723]
45. Monahan PO, Alder CA, Khan BA, Stump T, Boustani MA. The Healthy Aging Brain Care (HABC) Monitor: validation of the Patient Self-Report Version of the clinical tool designed to measure and monitor cognitive, functional, and psychological health. Clin Interv Aging 2014;9:2123-2132. [doi: 10.2147/CIA.S64140] [Medline: 25584024]
46. Stein J, Pabst A, Weyerer S, et al. The assessment of met and unmet care needs in the oldest old with and without depression using the Camberwell Assessment of Need for the Elderly (CANE): Results of the AgeMooDe study. J Affect Disord 2016 Mar 15;193:309-317. [doi: 10.1016/j.jad.2015.12.044] [Medline: 26774519]
47. King RB, Hartke RJ, Lee J, Raad J. The stroke caregiver unmet resource needs scale: development and psychometric testing. J Neurosci Nurs 2013 Dec;45(6):320-328. [doi: 10.1097/JNN.0b013e3182a3ce40] [Medline: 24217142]
48. EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [doi: 10.1016/0168-8510(90)90421-9]
49. Gray J, Kim J, Ciesla JR, Yao P. Rasch analysis of the Lubben Social Network Scale-6 (LSNS-6). J Appl Gerontol 2016 May;35(5):508-528. [doi: 10.1177/0733464814560468] [Medline: 25428591]
50. Hébert R, Bravo G, Préville M. Reliability, validity and reference values of the Zarit Burden Interview for assessing informal caregivers of community-dwelling older persons with dementia. Can J Aging 2000;19(4):494-507. [doi: 10.1017/S0714980800012484]
51. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
52. Øksnebjerg L, Woods B, Ruth K, et al. A tablet app supporting self-management for people with dementia: explorative study of adoption and use patterns. JMIR Mhealth Uhealth 2020 Jan 17;8(1):e14694. [doi: 10.2196/14694] [Medline: 31951217]
53. Hettinga M, De Boer J, Goldberg E, Moelaert F. Navigation for people with mild dementia. Stud Health Technol Inform 2009;150:428-432. [doi: 10.3233/978-1-60750-044-5-428] [Medline: 19745347]
54. van de Vijver S, Hummel D, van Dijk AH, et al. Evaluation of a digital self-management platform for patients with chronic illness in primary care: qualitative study of stakeholders' perspectives. JMIR Form Res 2022 Aug 3;6(8):e38424. [doi: 10.2196/38424] [Medline: 35921145]
55. Fischer B, Peine A, Östlund B. The importance of user involvement: a systematic review of involving older users in technology design. Gerontologist 2020 Sep 15;60(7):e513-e523. [doi: 10.1093/geront/gnz163] [Medline: 31773145]
56. Kopeć W, Balcerzak B, Nielek R, Kowalik G, Wierzbicki A, Casati F. Older adults and hackathons: a qualitative study. Empir Software Eng 2018 Aug;23(4):1895-1930. [doi: 10.1007/s10664-017-9565-6]
57. Maaß S, Buchmüller S. The crucial role of cultural probes in participatory design for and with older adults. i-com 2018 Aug 28;17(2):119-135. [doi: 10.1515/icom-2018-0015]
58. Eftring H, Frennert S. Designing a social and assistive robot for seniors. Z Gerontol Geriatr 2016 Jun;49(4):274-281. [doi: 10.1007/s00391-016-1064-7] [Medline: 27220732]
59. Lee HR, Šabanović S, Chang WL, et al. Steps toward participatory design of social robots: mutual learning with older adults with depression. Presented at: Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction; Mar 6-9, 2017; Vienna Austria p. 244-253. [doi: 10.1145/2909824.3020237]
60. Alaoui M, Ting KLH, Lewkowicz M. The urge for empirically-informed design of social-oriented AAL applications – the example of 2 AAL projects. Procedia Comput Sci 2014;27:379-388. [doi: 10.1016/j.procs.2014.02.042]
61. Arnstein SR. A ladder of citizen participation. J Am Inst Plann 1969 Jul;35(4):216-224. [doi: 10.1080/01944366908977225]

Abbreviations

CANE: Camberwell Assessment of Need for the Elderly
CNA-D: Carer's Needs Assessment for Dementia
DelpHi-MV: Dementia: life- and person-centered help in Mecklenburg-Western Pomerania
G-BA: Gemeinsamer Bundesausschuss/Innovationsausschuss
GAIN: Gesund Angehörige Pflegen
GP: general practitioner
HCP: health care professional
JPND: EU Joint Programme in Neurodegenerative Research
LSNS: Lubben Social Network Scale
VAS: visual analog scale
WHO: World Health Organization
ZBI: Zarit Burden Interview

XSL•FO
RenderX

Edited by R Bjarnadottir; submitted 26.04.24; peer-reviewed by C Peterson, M Patrascu; revised version received 02.12.24; accepted 23.12.24; published 07.04.25.

Please cite as:
Biernetzky OA, Thyrian JR, Boekholt M, Berndt M, Hoffmann W, Teipel SJ, Kilimann I
Identifying Unmet Needs of Informal Dementia Caregivers in Clinical Practice: User-Centered Development of a Digital Assessment Tool
JMIR Aging 2025;8:e59942
URL: https://aging.jmir.org/2025/1/e59942
doi:10.2196/59942

© Olga A Klein, Jochen René Thyrian, Melanie Boekholt, Matthias Berndt, Wolfgang Hoffmann, Stefan J Teipel, Ingo Kilimann. Originally published in JMIR Aging (https://aging.jmir.org), 7.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Uncovering Specific Navigation Patterns by Assessing User Engagement of People With Dementia and Family Caregivers With an Advance Care Planning Website: Quantitative Analysis of Web Log Data

Charlèss Dupont[1,2], PhD; Tinne Smets[1,2], PhD; Courtney Potts[3], MSc; Fanny Monnet[1,2], PhD; Lara Pivodic[1,2], PhD; Aline De Vleminck[1,2], PhD; Chantal Van Audenhove[4], PhD; Maurice Mulvenna[5], PhD; Lieve Van den Block[1,2], PhD

[1]End-of-life Care Research Group, Vrije Universiteit Brussel (VUB), Laarbeeklaan 103, Brussels, Belgium

[2]Department of Family Medicine and Chronic Care, Vrije Universiteit Brussel (VUB), Brussels, Belgium

[3]School of Psychology, Ulster University, Coleraine, United Kingdom

[4]LUCAS Center for Care Research and Consultancy, KU Leuven, Leuven, Belgium

[5]School of Computing, Ulster University, Belfast, United Kingdom

Corresponding Author:
Charlèss Dupont, PhD
End-of-life Care Research Group, Vrije Universiteit Brussel (VUB), Laarbeeklaan 103, Brussels, Belgium

Abstract

Background: Web-based tools have gained popularity to inform and empower individuals in advance care planning. We have developed an interactive website tailored to the unique needs of people with dementia and their families to support advance care planning. This website aims to break away from the rigid pathways shown in other tools that support advance care planning, in which advance care planning is shown as a linear process from information to reflection, communication, and documentation.

Objective: This study aimed to assess the website's usage by people with dementia and their family caregivers, identify distinct user engagement patterns, and visualize how users navigated the website.

Methods: We analyzed the website's log data obtained from an 8-week evaluation study of the site. Interactions with the website were collected in log data files and included visited web pages or clicked-on hyperlinks. Distinct user engagement patterns were identified using K-means clustering process mining, a technique that extracts insights from log data to model and visualize workflows, was applied to visualize user pathways through the website.

Results: A total of 52 participants, 21 individuals with dementia and their family caregivers as dyads and 10 family caregivers were included in the study. Throughout the 8-week study, users spent an average of 35.3 (SD 82.9) minutes over 5.5 (SD 3.4) unique days on the website. Family caregivers mostly used the website (alone or with a person with dementia) throughout the 8-week study. Only 3 people with dementia used it on their own. In total, 3 distinct engagement patterns emerged: low, moderate, and high. Low-engagement participants spent less time on the website during the 8 weeks, following a linear path from information to communication to documentation. Moderate- and high-engagement users showed more dynamic patterns, frequently navigating between information pages and communication tools to facilitate exploration of aspects related to advance care planning.

Conclusions: The diverse engagement patterns underscore the need for personalized support in advance care planning and challenge the conventional linear advance care planning representations found in other web-based tools.

(JMIR Aging 2025;8:e60652) doi:10.2196/60652

KEYWORDS

dementia; advance care planning; user engagement; web-based tool; care; website; caregiver; communication; tool; online

Introduction

Advance care planning is a communication process between patients, families, and health care professionals to "define goals and preferences for future care and treatment" [1]. This process holds particular relevance in dementia [2]. The cognitive decline in dementia highlights the need for early initiation of advance care planning, enabling people with dementia to reflect on and express their preferences for future care and treatments [3-5]. While existing definitions of advance care planning often focus on medical care decisions [6], people with dementia and their families have emphasized that it should include exploring what matters now and in the future, including nonmedical aspects of care [7-10]. Furthermore, people with dementia and their

XSL•FO
RenderX

families have expressed a need to discuss future care together [7]. This aligns with the recently introduced public health to advance care planning, emphasizing a shift toward a social focus on "what matters most to people" rather than the current emphasis on end-of-life decision making and underlining the need to support conversations in the family context [11].

To support advance care planning within the family context, interactive web-based tools like websites or apps have been promoted [12]; however, despite their proven benefits in other populations, there is a noticeable absence of tools tailored to the specific needs of people with dementia and their families [13-15]. To address this gap, we developed a website for and with people with dementia and their families [16-19]. This website deviates from the structured linear pathways found in other tools to support advance care planning [15]. Such tools adhere to a stepwise procedure, often commencing with information provision, prompting reflection, followed by communication, and concluding with documentation, commonly in the form of advance directives. For example, in some cases, users must complete or make an explicit effort to skip a step before being able to proceed to the next step, eg, users cannot access certain features such as communication tools or documentation forms without first completing previous tasks, for example, users have to follow a linear pathway when navigating the website [15,20]. In contrast, our website is designed to offer flexibility, allowing users to engage with the content in any way they choose. There are no predefined steps or sequences to follow, and users can access all features, such as communication tools and documentation forms, without needing to complete any previous tasks [16].

Following the development of the website to support advance care planning, we performed an 8-week evaluation study involving people with mild to moderate dementia and their family caregivers [17,18]. The primary focus of this evaluation was to assess the usability, acceptability, feasibility, experiences, and outcomes of using the website [17,18]. Participants in the evaluation study found that the website supports advance care planning [17,18]. After 8 weeks, participants exhibited improved advance care planning knowledge, self-efficacy, and skills [17,18].

Beyond this evaluation, gaining insights into the website's user engagement and usage patterns is crucial to enhance our understanding of how people with dementia and their families use a website to support advance care planning in the family context, how they engage with it and how single users and dyadic users differ (ie, people with dementia together with their family caregiver). This knowledge will inform enhancements to the website's design and functionality by highlighting user needs in digital health interventions, ultimately improving support for families navigating advance care planning. Therefore, this study aims to: explore how people with dementia and their family caregivers used the advance care planning support website during an 8-week evaluation study; explore whether and which user behavior clusters can be typified based on the engagement of people with dementia and their family caregivers with the website; and explore and visualize user pathways of the identified user behavior clusters, that is, how the different user clusters of people with dementia and their family caregivers navigate through the website.

Methods

Study design

This study quantitatively analyzed web log data from a convergent parallel mixed methods evaluation study of a website designed to support people with dementia and their family caregivers in advance care planning. Web log data was collected during an 8-week evaluation study to capture participants' interactions with the website, for example, time spent on the website and what pages were visited. This log data was used to explore the website's usage by people with dementia and their family caregivers, which user behavior clusters can be typified and to visualize user pathways. The protocol, including the objective to explore usage patterns of the website, as well as the main results of the evaluation study, have been published elsewhere [17,18]

The Development of the Advance Care Planning Support Website

The development of the website followed a user-centred, iterative design process, ensuring alignment with the needs of people with dementia and their families. The website was developed to provide information and support for people with dementia and their family caregivers in advance care planning [16]. Furthermore, the website was designed with flexible navigation options, allowing users to explore different sections at their own pace rather than following a linear, step-by-step approach. This was particularly important given the varying levels of readiness among users to engage with sensitive topics such as future care preferences [8,16].

The website includes advance care planning information, information about legal frameworks, communication tips, and documentation sections. Furthermore, it provides accessibility features, such as text-to-speech and text enlargement and 2 interactive communication tools, an "Interactive Card Tool" based on the recently developed paper-based version of the Levenswensen (Life Wishes) cards and a fill-in tool "Thinking Now About Later," that guides users through a reflective process to help users think and talk about and write down their preferences for the present and future (Multimedia Appendices 1 and 2). The results of the evaluation study and experiences with the interactive communication tools, have been published elsewhere [16,18,21].

Participants, Recruitment and Setting

People with mild to moderate dementia, including both early and late onset, along with their family caregivers as dyads or the family caregiver alone, were recruited [17]. To be eligible, participants needed to meet the criteria shown in Textbox 1 [17]
.

XSL•FO
RenderX

Textbox 1. Inclusion criteria for the study.

Both people with dementia and their family caregivers:

- Express interest in and willingness to test the website to support advance care planning.
- Provide consent for study participation.
- Proficient in speaking and understanding Dutch.
- Own a device (laptop, tablet, mobile phone, etc).

People with dementia:

- Diagnosed with young- or late-onset dementia.

Family caregivers:

- Actively involved in the care (physical, emotional, social, etc) of the person with dementia.

For participants recruited as dyads, at least one had to be capable of navigating the website. For instance, the person with dementia and the family caregiver could not both have visual impairments or other disabilities hindering interaction with the advance care planning website. Participants were recruited through dementia care organizations and memory clinic neurologists. Eligible participants with a confirmed dementia diagnosis were identified by health care professionals. People with dementia who were interested in the study contacted the researcher, received study information, and, if still interested after receiving the study information, were offered an appointment to perform the eligibility screening and provide informed consent. The recruitment process is described in more detail in the study protocol that is published elsewhere [17]. The study was conducted in Flanders, the Dutch-speaking region of Belgium.

Data Collection Procedures

Between October 2022 and May 2023, 52 people participated in the evaluation study: 21 dyads, that is, 21 people with dementia and 21 family caregivers, and 10 family caregivers alone. Information on sociodemographic data, including age, gender, computer literacy, type of diagnosis and date, was gathered through a survey assessed at the start of the 8 weeks [17,18]. Participants were granted access to the website after providing informed consent and completing baseline data collection. They were informed that they could use the website freely over an 8-week period, choosing how and when to engage with it according to their preferences [17].

To capture the usage by the participants during the 8-week study, there was a continuous collection of log data. Usage refers to how people with dementia and family caregivers engage with the advance care planning website. The log data captured what pages were visited, time spent on each page, interaction with content or functions, frequency of visits to each page, and the search queries used. Upon accessing the website, participants were prompted to identify themselves through a pop-up question, requiring them to fill in their study ID and specify whether they were a person with dementia, a family caregiver, or engaging together as a dyad. A detailed study protocol of the evaluation study is published elsewhere [8].

Data Analysis

Overview

Log data of the advance care planning website were analyzed using the programming language R (version 4.2.3, R Studio). The log data was saved in 3 files, application and access logs and a file to save the interactions with the interactive communication tools. First, the 3 log datasets were cleaned by eliminating irrelevant data, such as admin and php requests. The application logs were further filtered to include only valid study IDs used by the participants. Study IDs were then added to the access logs based on cross matching with IP addresses. The access and application logs were combined, and additional information from the third dataset regarding the completion of interactive communication tools was incorporated by matching IP addresses to study IDs.

Assessing the Extent to Which the Website Was Used

First, the interactions with the website were summarized by time spent on the website, pages visited, and who visited. This analysis focused on usage patterns, which provided an initial understanding of the variability in engagement.

Identifying Distinct Behavioral Clusters

After, the data underwent K-means clustering. Clustering the data aimed to identify behavioral or interaction patterns that typify user engagement. The 6 features derived from the summary data were used for clustering: total interactions, unique days, duration of use, total clicks on communication, information, and documentation pages. The R package “caret” was used to normalize the data using min-max normalization, converting variables to a range between 0 and 1 [22]. Usage frequency emerged as the key differentiating factor among participants. The “NbClust” package was then used to compare summary statistics and determine the optimal number of clusters, using 30 indices to determine the optimal number of clusters (k) between 2 and 10 based on these metrics [23]. The “NbClust” package indicated that the data could be best classified into 3 clusters, which were named after usage frequency of participants, for example , low, moderate, and high engagement levels. Subsequently, the k-means algorithm was applied to the data for k clusters. Principal component analysis (PCA) was used for data visualization to reduce the multiple features to 2 dimensions. Finally, Kruskal-Wallis rank sum tests

XSL•FO
RenderX

and χ^2 tests were conducted to evaluate the significant differences between features for each cluster, providing further insights into the characteristics of the identified user clusters.

Identification of User Pathways

To visualize how users navigated through the website, we applied process mining techniques using the R package BupaR. This allowed us to map and analyze the various paths users followed across the site's different sections [24]. Log data were filtered to remove redundant information (taken out: change font size, contrast, privacy policy, read speaker, print, and other). First, a process matrix, which is a 2-dimensional matrix showing the flows between activities, was generated to visualize the entire log dataset [25]. After, individual process maps were produced per study ID to show the path taken by participants. Finally, the individual paths were compared with the IDs per cluster to find similarities and differences between the participants' paths in each cluster.

Ethical Considerations

The research protocol was submitted to the Medical Ethics Committee of Brussels university hospital (UZ Brussel) and received ethical approval (BUN 1432022000179). All participants provided informed consent, and the study was conducted in accordance with the relevant ethical guidelines and regulations. To ensure privacy and confidentiality, all participant data were de-identified prior to analysis. Protective measures, such as secure data storage and restricted access to personal information, were implemented to safeguard participant privacy throughout the study.

Results

Overview

In total, 52 participants with 31 study IDs were included (21 dyads and 10 family caregivers), consisting of 21 people with dementia and 31 family caregivers. Ten family caregivers participated alone, either because the person with dementia was unable to provide informed consent (6/10) or chose not to participate (4/10). The average age of people with dementia was 62.8 (SD 10.4) years, with 42.8% (9/21) being female. Among family caregivers, 68% (21/31) were female, with an average age of 62.1 (SD 10.9) years. Family caregivers reported an average computer literacy score of 7.5 (on a scale from 1 to 10, with higher scores indicating greater self-reported computer literacy), while people with dementia had an average score of 4 (Table 1).

Table . Description of the study population.

Variables		Persons with dementia (n=21)	Family caregivers (n=31)
Sex (female), n		9	21
Age (years), mean (SD)		62.1 (10.9)	62.8 (10.4)
Relationship, n			
	Partners[a]	18	25
	Parent (in law)-child	3	6
Profession, n			
	Employed	1	16
	Retired	20[b]	15
Computer literacy[c], mean (SD)		62.1 (10.9)	62.8 (10.4)
Dementia diagnosis[d], n			
	Alzheimer disease	15	20
	Vascular dementia	1	3
	Frontotemporal dementia	3	3
	Lewy body dementia	1	1
	Parkinson dementia	0	1
	I do not know	1	3

[a]Married, living together or in a romantic relationship.

[b]Six of the retired persons with dementia were forced to take early retirement because of their diagnoses.

[c]Self-evaluation on a scale from 1 (no computer skills) to 10 (excellent computer skills).

[d]Dates of diagnoses: January 2013 to December 2022.

Website Usage

Each study ID (n=31) was logged at least once in the log data, indicating at least 1 visit by 1 or both dyad members and by each family caregiver who participated alone. In 10 study IDs, both the family caregiver and dyad used the website. In 15 instances, only the caregiver accessed it and in 3 cases, solely the dyad engaged with the website. In addition, 3 occurrences

XSL•FO
RenderX

involved mixed usage, with indications of the family caregiver, dyad, or person with dementia using the website. The total number of unique interactions by all users (n=31) with the website was 1799 (ie, the total interaction data points), encompassing information searches, clicks, and use of interactive elements. On average, users had 58 (SD 57) interactions over the 8 weeks. The total duration spent on the website during the 8 weeks was 35.3 (SD 82.9) minutes, and, on average, people used the website on 5.5 (SD 3.4) unique days of the 8-week study period.

User interactions revealed that family caregivers had the highest overall number of interactions (757), followed by dyads (235) and people with dementia alone (103). We faced difficulty attributing the information to specific user types for the other 701 interactions documented in the access log file. These unidentified usages per user type arose from difficulties linking interactions with corresponding user types when users left their browsers open for extended periods. The webpage "Advance Care Planning: What Is It?" was visited most (n=304). Followed by the glossary section (209), "Advance Care Planning: Thinking and Talking About Later" (277), "Advance Care Planning: Writing It Down for Later" (259) and the frequently asked questions section (122). The 2 interactive communication tools were used 136 (Thinking Now About Later) and 91 (Levenswensen cards) times.

Identifying Typical User Behaviors by Clustering

To determine the optimal number of clusters, we applied NbClust with 30 indices, revealing that 8 favored 2 clusters and 8 indicated 3 clusters. Following the majority rule, we opted for 3 clusters, each assigned to study IDs reflecting diverse user engagement patterns. Table 2 summarizes each cluster's characteristics. To showcase the most-visited web pages by the clusters, we categorized all pages into 10 categories (Figure 1; Multimedia Appendix 1 visually shows the web pages corresponding to the categories).

Table . Overview of the characteristics of use of the website of each cluster.

Characteristics	Moderate engagement level	Low engagement level	High engagement level	*P* value
Size, study ID/n[a] (%)	15/21 (45)	5/10 (21)	11/16 (34)	—[b]
Total interactions[c], mean (SD)	50 (13)	21 (9)	86 (15)	<0.001[d]
Unique days, mean (SD)	4.4 (1.5)	1.6 (0.5)	8.7 (3.3)	0.004[d]
Range of use (days), mean (SD)[e]	56.7 (19.3)	9.8 (19.2)	65.6 (16.9)	<0.001[d]
Communication pages visited, mean (SD)	17 (7)	5 (3)	28 (10)	<0.001[d]
Documentation pages visited, mean (SD)	7 (4)	6 (4)	18 (4)	<0.001[d]
Information pages visited, mean (SD)	22 (10)	7 (5)	34 (13)	<0.001[d]
Total number of interactions per user type, n				<0.001[g]
Caregiver	420	27	310	
Person with dementia	0	0	103	
Dyad	124	33	88	
Family caregiver and person with dementia[f]	204	54	452	

[a]n is the number of users (N=46 because 5 people with dementia did not use the website or participate anymore).

[b]Not applicable.

[c]Interaction is any movement of the user on the website. This can be clicking on a hyperlink, watching a video, opening a page, or printing the web page.

[d]χ^2 test.

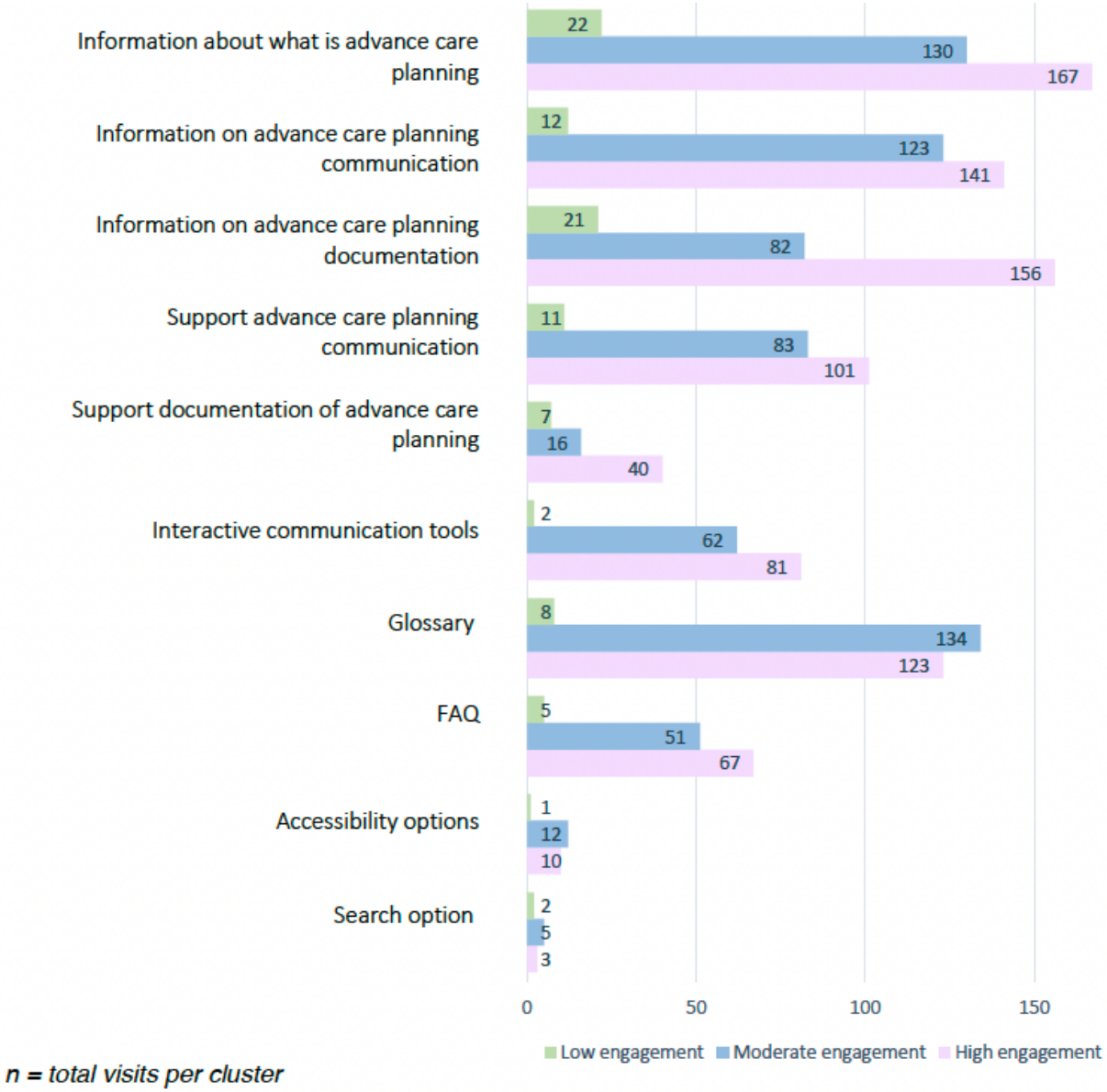
[e]From the first day of use until the last day. For example, cluster moderate engagement started using the website, and, on average, the last usage was 9.8 days later.

[f]When the IP address could not be matched with the date or time of the filled-in study ID.

[g]Kruskal-Wallis rank sum test.

XSL•FO
RenderX

Figure 1. Overview of the most-visited web pages by the 3 clusters. FAQ: frequently asked questions.



Cluster low engagement (5 participants) exhibited the fewest interactions (mean of 21), shorter duration (mean of 9.8 days), and accessed fewer pages. Cluster moderate engagement (15 participants) had a mean of 50 interactions, visited for 4.4 unique days, with a use duration of approximately 56.7 days (Table 2). Cluster high engagement (11 participants) demonstrated the highest engagement, with the highest total interactions (mean of 86), longer duration (mean of 65.6 days), accessing more diverse pages and including the 3 people with dementia who used the website alone. Statistically significant *P* values underscore distinct engagement patterns among the clusters (Table 2).

Participants across all 3 clusters predominantly visited the "what is advance care planning" web page the most. Low-engagement participants mostly focused on the information pages, especially the documentation information and showed limited interest in interactive tools. Moderate engagement participants visited mostly the advance care planning information pages, frequently using the glossary and accessibility features like contrast and text-to-speech, distinguishing their usage from other clusters. In the high-engagement cluster, participants explored information, communication and documentation pages, actively using interactive communication tools.

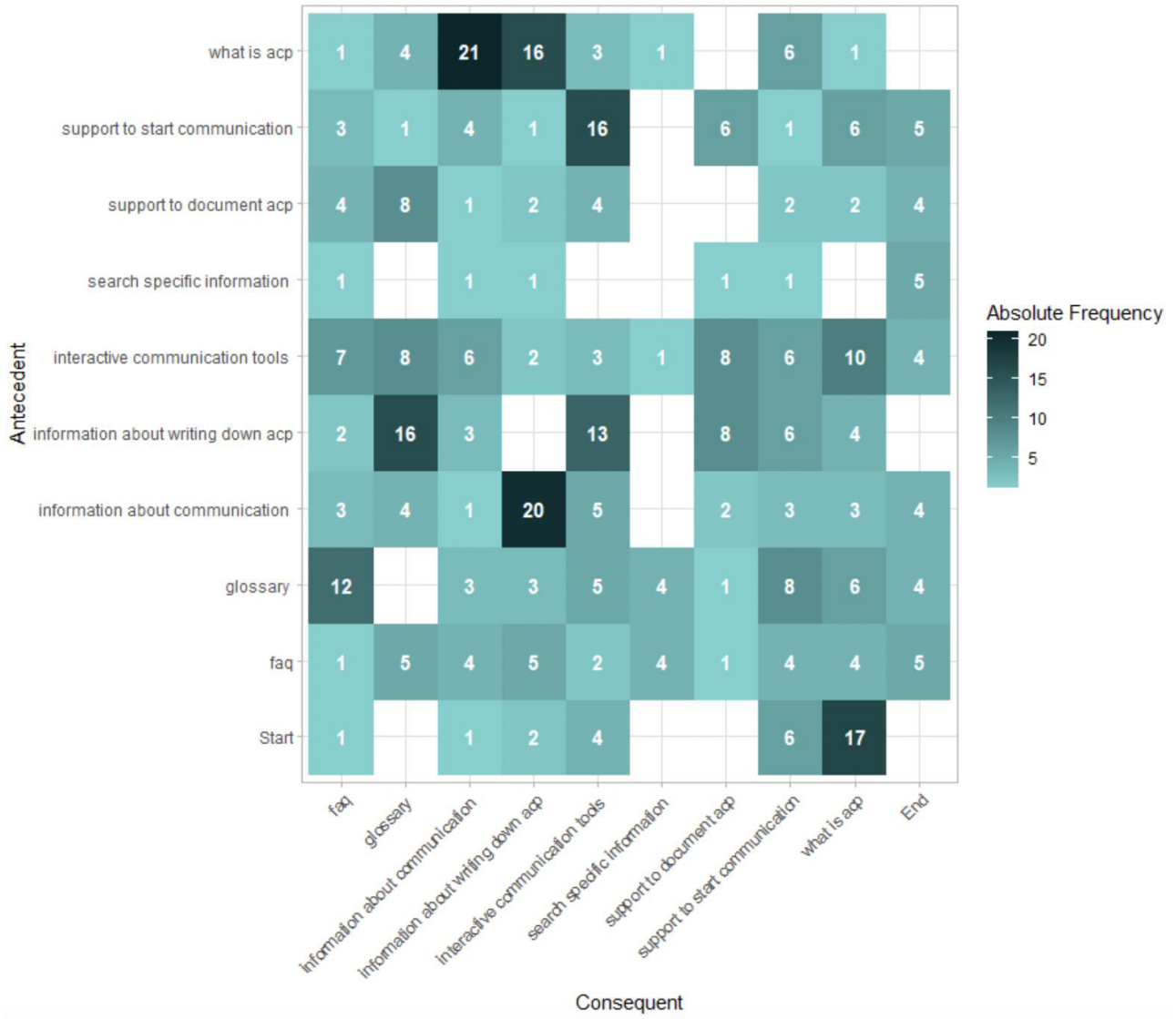
Identification of User Pathways of the Identified User Behavior

To identify the user pathways, we first look at the overall pathways of users using a process matrix (Figure 2). The process matrix is a two-dimensional representation that illustrates the flow between the web pages that users have visited. The matrix is organized with antecedent events followed by the consequent events. The primary pathway, observed in 21 instances, initiates with a visit to the "What is advance care planning" page, followed by navigating to a subsequent page providing "information about communication." Significantly, users frequently started their pathway on informative pages, such as "what is advance care planning," "information on communication," or "Information on documentation." Subsequently, they progress to explore additional information pages. Noteworthy is the observation that users typically visit

XSL•FO
RenderX

pages providing information about communication or documentation before engaging with the interactive communication tools.

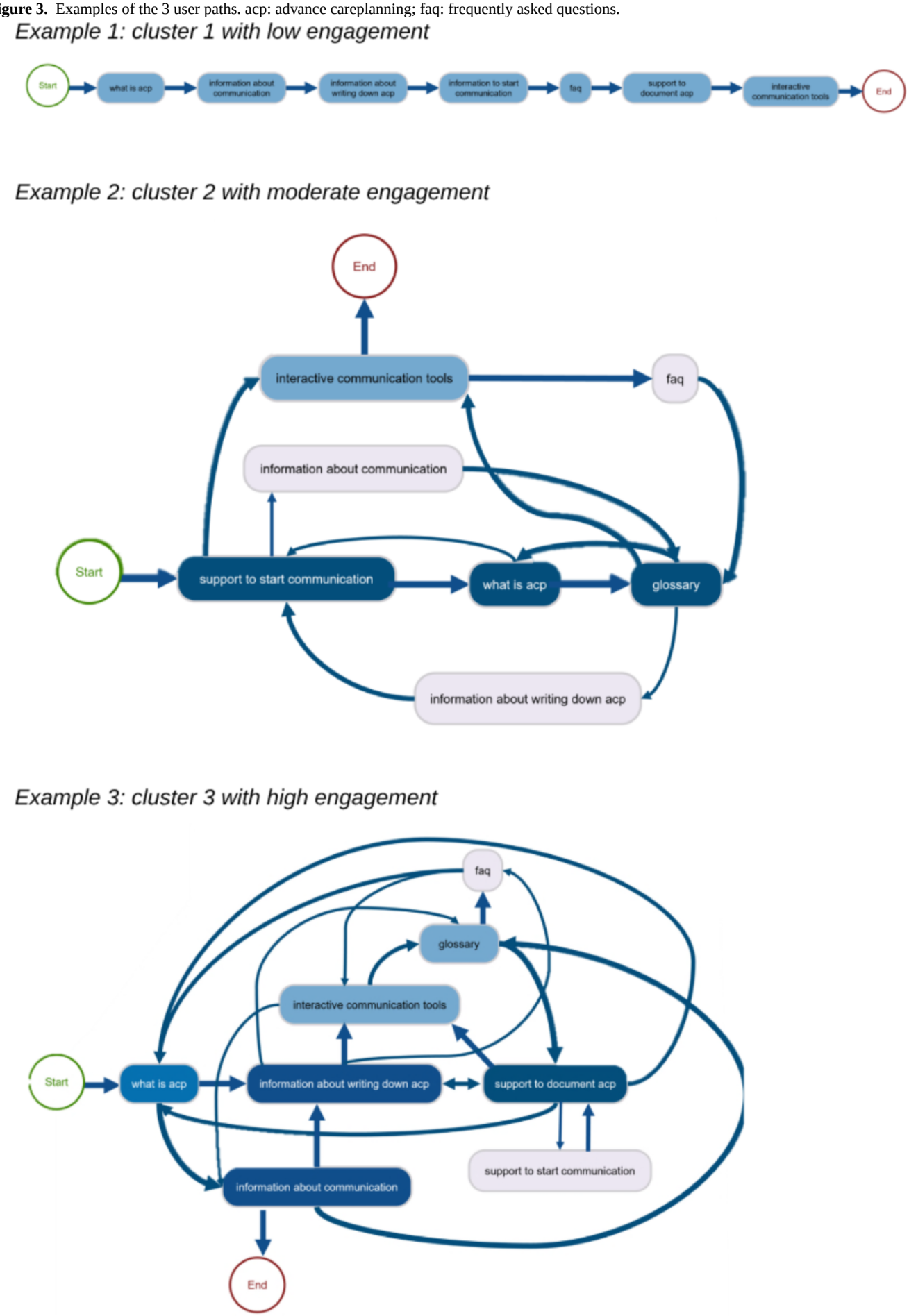
Figure 2. Identification of overall user pathways. acp: advance care planning; FAQ: frequently asked questions.



Then, the user pathways of the identified user behavior clusters were identified (Figure 3). Low-engagement participants (Figure 3, example 1) displayed a linear and direct browsing style, rarely revisiting previous pages during navigation. In contrast, the participants with moderate and high engagement (Figure 3, examples 2 and 3) explored the website by visiting pages sequentially, occasionally revisiting previously viewed pages. High engagement extensively explored various pages, moving between information and guidance pages. Participants with moderate engagement involved frequent transitions between pages but less often than those with high engagement.

XSL•FO
RenderX

Figure 3. Examples of the 3 user paths. acp: advance careplanning; faq: frequently asked questions.



XSL•FO
RenderX

Discussion

Principal Findings

The analyses of the log data of an 8-week evaluation study of a website to support people with dementia and their families in advance care planning showed that family caregivers used the website most often, either alone or with a person with dementia, that is, dyad. Three distinct engagement patterns emerged in this study: low, moderate, and high. Low-engagement users tended to follow a more linear path on the website, while moderate- and high-engagement users displayed a more dynamic engagement, exploring the website in diverse ways.

Flexible user navigation patterns were evident in our study, challenging the conventional linear advance care planning representations found in other web-based tools in which users go through a stepwise process, typically starting with information provision, prompting reflection, moving on to communication, and concluding with documentation in the form of advance directives [15]. This rigid structure, embedded in the design of other ACP websites, may restrict easy access to subsequent features or content until earlier steps are completed. While this structured pathway might work for some users, it assumes that everyone is ready to follow the same route through the ACP process, which may not always be the case [7,8,26]. In our study, low-engagement users tended to follow a more linear pathway. However, those with moderate or high engagement demonstrated more dynamic and varied usage patterns. Some users initiated their engagement by seeking information, while others prioritized documentation or directly accessed communication tools. This flexibility of navigation aligns with feedback from family caregivers during the website's development process, which emphasized the importance of allowing users to navigate at their own pace, rather than following a rigid, linear path [8]. This flexibility allowed for personalized engagement, as different users may be at different stages of readiness to engage with sensitive topics, such as future care planning. Furthermore, this approach aligns with broader technological research [7,10,26], which shows that family caregivers prefer flexible navigation in internet-based tools [27,28]. These findings suggest a need for a more flexible approach, indicating that users should have the freedom to navigate tools that align best with their needs, which is also supported by other research [5,26,29]. This also aligns with Belgian clinical guidelines [30], explicitly mentioning that health care professionals should tailor advance care planning in dementia, including style and content, to the "person's level and rhythm."

Despite participants' expressed interest in advance care planning in the evaluation study, people with dementia rarely engaged with the website on their own. Family caregivers and the person with dementia did engage together, emphasizing the family's importance. This finding is not necessarily surprising as much literature points at the importance of family in a dementia trajectory. A recently published consensus definition on advance care planning in dementia also highlighted family as highly important and specific in this population [6,31]. Regarding the use or uptake of websites among people with dementia, involving family caregivers can play a facilitating role; however, it is essential to acknowledge that not all people with dementia have family caregivers or families directly engaged in their care [32]. Furthermore, overreliance on family caregivers may unintentionally hinder independent usage, potentially undermining autonomy. Therefore, achieving an inclusive environment necessitates balancing involving family caregivers and promoting self-usage.

Strengths and Limitations of the Study

This study has several strengths that contribute to the robustness of our study's findings. Using log data, this study is the first to examine user engagement of people with dementia and their families with a website to support advance care planning. It offers a comprehensive understanding of their specific usage patterns. The elimination of recall bias is another key strength, as log data provided an accurate account of how users engaged with the website. The study also has limitations with regard to the data used. We encountered difficulty identifying the specific type of user for all log data because not all access logs could be matched with the date or time of the filled-in type of user in the application log. In addition, due to a 1-month retention period for log data, a small portion of data (7 days with interactions from 2 users, as indicated in Google Analytics) was lost as it was not downloaded before deletion from the server. Another limitation is the relatively small dataset for our log data analyses. Specifically, the number of participants with dementia who accessed the website independently was small what limits our ability to make comparisons in usage patterns between people with dementia, family caregivers, and dyads. Further research with a larger sample of people with dementia, including both independent and caregiver-supported users, is needed to enhance the robustness of the findings and provide stronger insights into their engagement patterns. Finally, a limitation arises from self-selection bias in user type; for example, users might identify as "family caregivers" while engaging together, and vice versa, introducing variability that could impact the accuracy of findings.

Conclusions

This study offers insights in how people with dementia and their family caregivers use a website designed to support them in advance care planning. The findings show that family caregivers are the website's primary users, often engaging with the website alone or together with the person with dementia. Three distinct user engagement patterns, low, moderate, and high engagement, were identified, with more dynamic navigation observed among high-engagement users, particularly those using interactive communication tools. These findings underscore the need for flexible user pathways in online advance care planning tools, allowing personalized navigation. Future research with a larger and more diverse sample of persons with dementia is necessary to confirm these findings and allow for more detailed comparisons across different user groups.

XSL•FO
RenderX

Acknowledgments

We would like to thank all participants for their time and effort. We thank the neurologists working in memory clinics, the Expertise Centrum Dementia Flanders, Alzheimer Liga Flanders, and Het Ventiel for their efforts in recruiting participants. This study is part of the project "CAPACITY: Flanders Project to Develop Capacity in Palliative Care Across Society," a collaboration between the Vrije Universiteit Brussel, Ghent University, and the Catholic University Leuven, Belgium. This study is supported by a grant from the Research Foundation – Flanders (file number S002219N). The research presented in this paper was carried out as part of the DISTINCT project, a Marie Curie Innovative Training Network (ITN), H2020-MSCA-ITN-2018, under grant agreement number 813196.

Data Availability

The website assessed in this study is available in Dutch [19].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Overview of the website.
[DOCX File, 4672 KB - aging_v8i1e60652_app1.docx]

Multimedia Appendix 2
The interactive communication tools.
[DOCX File, 1664 KB - aging_v8i1e60652_app2.docx]

References

1. Rietjens JAC, Sudore RL, Connolly M, et al. Definition and recommendations for advance care planning: an international consensus supported by the European Association for Palliative Care. Lancet Oncol 2017 Sep;18(9):e543-e551. [doi: 10.1016/S1470-2045(17)30582-X] [Medline: 28884703]
2. Van den Block L. Advancing research on advance care planning in dementia. Palliat Med 2019 Mar;33(3):259-261. [doi: 10.1177/0269216319826411] [Medline: 30758274]
3. Bosisio F, Jox RJ, Jones L, Rubli Truchard E. Planning ahead with dementia: what role can advance care planning play? A review on opportunities and challenges. Swiss Med Wkly 2018 Dec 17;148:w14706. [doi: 10.4414/smw.2018.14706] [Medline: 30594990]
4. van der Steen JT, Radbruch L, Hertogh CMPM, et al. White paper defining optimal palliative care in older people with dementia: a Delphi study and recommendations from the European Association for Palliative Care. Palliat Med 2014 Mar;28(3):197-209. [doi: 10.1177/0269216313493685] [Medline: 23828874]
5. Wendrich-van Dael A, Bunn F, Lynch J, Pivodic L, Van den Block L, Goodman C. Advance care planning for people living with dementia: an umbrella review of effectiveness and experiences. Int J Nurs Stud 2020 Jul;107:103576. [doi: 10.1016/j.ijnurstu.2020.103576] [Medline: 32380259]
6. van der Steen JT, Nakanishi M, Van den Block L, et al. Consensus definition of advance care planning in dementia: a 33-country Delphi study. Alzheimers Dement 2024 Feb;20(2):1309-1320. [doi: 10.1002/alz.13526] [Medline: 37985444]
7. Dickinson C, Bamford C, Exley C, Emmett C, Hughes J, Robinson L. Planning for tomorrow whilst living for today: the views of people with dementia and their families on advance care planning. Int Psychogeriatr 2013 Dec;25(12):2011-2021. [doi: 10.1017/S1041610213001531] [Medline: 24053783]
8. Dupont C, Smets T, Monnet F, et al. Defining the content of a website on advance care planning in dementia: a focus group study with family and health professionals. BMC Med Inform Decis Mak 2023 Nov 8;23(1):254. [doi: 10.1186/s12911-023-02359-1] [Medline: 37940991]
9. Monnet F, Pivodic L, Dupont C, Diaz A, Gove D, Van den Block L. Defining advance care planning from the perspective of people with dementia: focus groups with the European Working Group of People with dementia. Alzheimers Dement 2023 Jun;19(S5). [doi: 10.1002/alz.063913]
10. Van Rickstal R, De Vleminck A, Aldridge MD, et al. Limited engagement in, yet clear preferences for advance care planning in young-onset dementia: an exploratory interview-study with family caregivers. Palliat Med 2019 Oct;33(9):1166-1175. [doi: 10.1177/0269216319864777] [Medline: 31342861]
11. Abel J, Kellehear A, Millington Sanders C, Taubert M, Kingston H. Advance care planning re-imagined: a needed shift for COVID times and beyond. Palliat Care Soc Pract 2020;14:2632352420934491. [doi: 10.1177/2632352420934491] [Medline: 32924012]
12. Green MJ, Levi BH. The era of “e”: the use of new technologies in advance care planning. Nurs Outlook 2012;60(6):376-383. [doi: 10.1016/j.outlook.2012.08.005] [Medline: 23141197]

XSL•FO
RenderX

13. Howard M, Robinson CA, McKenzie M, et al. Effect of an interactive website to engage patients in advance care planning in outpatient settings. Ann Fam Med 2020 Mar;18(2):110-117. [doi: 10.1370/afm.2471] [Medline: 32152014]
14. van der Smissen D, van der Heide A, Sudore RL, Rietjens JAC, Korfage IJ. Measuring advance care planning behavior in Dutch adults: translation, cultural adaptation and validation of the Advance Care Planning Engagement Survey. BMC Med Res Methodol 2021 Sep 25;21(1):194. [doi: 10.1186/s12874-021-01389-5] [Medline: 34563115]
15. Dupont C, Smets T, Monnet F, et al. Publicly available, interactive web-based tools to support advance care planning: systematic review. J Med Internet Res 2022 Apr 20;24(4):e33320. [doi: 10.2196/33320] [Medline: 35442207]
16. Monnet F, Dupont C, Smets T, et al. An advance care planning website for people with dementia and their family caregivers: protocol for a development and usability study. JMIR Preprints. Preprint posted online on Mar 2, 2023. [doi: 10.2196/preprints.46935]
17. Dupont C, Monnet F, Pivodic L, et al. Evaluating an advance care planning website for people with dementia and their caregivers: protocol for a mixed method study. Digit Health 2023;9:20552076231197021. [doi: 10.1177/20552076231197021] [Medline: 37654712]
18. Dupont C, Smets T, Monnet F, et al. A website to support people with dementia and their family caregivers in advance care planning: results of a mixed-method evaluation study. Patient Educ Couns 2024 Oct;127:108357. [doi: 10.1016/j.pec.2024.108357] [Medline: 38943816]
19. Vroegtijdige zorgplanning (VZP) & dementie [Website in Dutch]. URL: https://www.vroegtijdigezorgplanningbijdementie.be/ [accessed 2025-02-05]
20. Fahner JC, Beunders AJM, van der Heide A, et al. Interventions guiding advance care planning conversations: a systematic review. J Am Med Dir Assoc 2019 Mar;20(3):227-248. [doi: 10.1016/j.jamda.2018.09.014] [Medline: 30471947]
21. Dupont C, Smets T, Monnet F, Eneslätt M, Tishelman C, Van den Block L. The cultural adaptation of the go wish card game for use in Flanders, Belgium: a public health tool to identify and discuss end-of-life preferences. BMC Public Health 2022 Nov 17;22(1):2110. [doi: 10.1186/s12889-022-14523-9] [Medline: 36397020]
22. Kuhn M. Building predictive models in R using the caret package. J Stat Softw 2008;28(5). [doi: 10.18637/jss.v028.i05]
23. Charrad M, Ghazzali N, Boiteau V, Niknafs A. NbClust: an R package for determining the relevant number of clusters in a data set. J Stat Soft 2014;61(6). [doi: 10.18637/jss.v061.i06]
24. Janssenswillen G, Depaire B, Swennen M, Jans M, Vanhoof K. bupaR: enabling reproducible business process analysis. Knowl Based Syst 2019 Jan;163:927-930. [doi: 10.1016/j.knosys.2018.10.018]
25. Potts C, Bond RR, Jordan JA, et al. Process mining to discover patterns in patient outcomes in a Psychological Therapies Service. Health Care Manag Sci 2023 Sep;26(3):461-476. [doi: 10.1007/s10729-023-09641-8] [Medline: 37191758]
26. Sellars M, Chung O, Nolte L, et al. Perspectives of people with dementia and carers on advance care planning and end-of-life care: a systematic review and thematic synthesis of qualitative studies. Palliat Med 2019 Mar;33(3):274-290. [doi: 10.1177/0269216318809571] [Medline: 30404576]
27. Núñez-Naveira L, Alonso-Búa B, de Labra C, et al. UnderstAID, an ICT platform to help informal caregivers of people with dementia: a pilot randomized controlled study. Biomed Res Int 2016;2016:5726465. [doi: 10.1155/2016/5726465] [Medline: 28116300]
28. Kajiyama B, Thompson LW, Eto-Iwase T, et al. Exploring the effectiveness of an Internet-based program for reducing caregiver distress using the iCare Stress Management e-Training Program. Aging & Mental Health 2013 Jul;17(5):544-554. [doi: 10.1080/13607863.2013.775641]
29. Ryan T, Amen KM, McKeown J. The advance care planning experiences of people with dementia, family caregivers and professionals: a synthesis of the qualitative literature. Ann Palliat Med 2017 Oct;6(4):380-389. [doi: 10.21037/apm.2017.06.15] [Medline: 28754049]
30. Piers R, Albers G, Gilissen J, et al. Advance care planning in dementia: recommendations for healthcare professionals. BMC Palliat Care 2018 Jun 21;17(1):88. [doi: 10.1186/s12904-018-0332-2] [Medline: 29933758]
31. Bastoni S, Wrede C, da Silva MC, et al. Factors influencing implementation of eHealth technologies to support informal dementia care: umbrella review. JMIR Aging 2021 Oct 8;4(4):e30841. [doi: 10.2196/30841] [Medline: 34623314]
32. Prescop KL, Dodge HH, Morycz RK, Schulz RM, Ganguli M. Elders with dementia living in the community with and without caregivers: an epidemiological study. Int Psychogeriatr 1999 Sep;11(3):235-250. [doi: 10.1017/s1041610299005803] [Medline: 10547124]

Abbreviations

ACP: advance care planning

XSL•FO
RenderX

Edited by AF Abdulai; submitted 18.05.24; peer-reviewed by C Clavien, JD Bacsu; revised version received 17.10.24; accepted 21.11.24; published 11.02.25.

Please cite as:
Dupont C, Smets T, Potts C, Monnet F, Pivodic L, De Vleminck A, Van Audenhove C, Mulvenna M, Van den Block L
Uncovering Specific Navigation Patterns by Assessing User Engagement of People With Dementia and Family Caregivers With an Advance Care Planning Website: Quantitative Analysis of Web Log Data
JMIR Aging 2025;8:e60652
URL: https://aging.jmir.org/2025/1/e60652
doi:10.2196/60652

© Charlèss Dupont, Tinne Smets, Courtney Potts, Fanny Monnet, Lara Pivodic, Aline De Vleminck, Chantal Van Audenhove, Maurice Mulvenna, Lieve Van den Block. Originally published in JMIR Aging (https://aging.jmir.org), 11.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The PDC30 Chatbot—Development of a Psychoeducational Resource on Dementia Caregiving Among Family Caregivers: Mixed Methods Acceptability Study

Sheung-Tak Cheng[1], PhD; Peter H F Ng[2,3], PhD

[1]Department of Health and Physical Education, The Education University of Hong Kong, Tai Po, China (Hong Kong)

[2]Department of Computing, Faculty of Computer and Mathematical Sciences, Hong Kong Polytechnic University, Hung Hom, China (Hong Kong)

[3]Department of Rehabilitation Sciences, Faculty of Health and Social Sciences, Hong Kong Polytechnic University, Hung Hom, China (Hong Kong)

Corresponding Author:
Sheung-Tak Cheng, PhD
Department of Health and Physical Education
The Education University of Hong Kong
10 Lo Ping Road
Tai Po
China (Hong Kong)
Phone: 852 29486563
Email: takcheng@eduhk.hk

Abstract

Background: Providing ongoing support to the increasing number of caregivers as their needs change in the long-term course of dementia is a severe challenge to any health care system. Conversational artificial intelligence (AI) operating 24/7 may help to tackle this problem.

Objective: This study describes the development of a generative AI chatbot—the PDC30 Chatbot—and evaluates its acceptability in a mixed methods study.

Methods: The PDC30 Chatbot was developed using the GPT-4o large language model, with a personality agent to constrain its behavior to provide advice on dementia caregiving based on the Positive Dementia Caregiving in 30 Days Guidebook—a laypeople's resource based on a validated training manual for dementia caregivers. The PDC30 Chatbot's responses to 21 common questions were compared with those of ChatGPT and another chatbot (called Chatbot-B) as standards of reference. Chatbot-B was constructed using PDC30 Chatbot's architecture but replaced the latter's knowledge base with a collection of authoritative sources, including the World Health Organization's iSupport, By Us For Us Guides, and 185 web pages or manuals by Alzheimer's Association, National Institute on Aging, and UK Alzheimer's Society. In the next phase, to assess the acceptability of the PDC30 Chatbot, 21 family caregivers used the PDC30 Chatbot for two weeks and provided ratings and comments on its acceptability.

Results: Among the three chatbots, ChatGPT's responses tended to be repetitive and not specific enough. PDC30 Chatbot and Chatbot-B, by virtue of their design, produced highly context-sensitive advice, with the former performing slightly better when the questions conveyed significant psychological distress on the part of the caregiver. In the acceptability study, caregivers found the PDC30 Chatbot highly user-friendly, and its responses quite helpful and easy to understand. They were rather satisfied with it and would strongly recommend it to other caregivers. During the 2-week trial period, the majority used the chatbot more than once per day. Thematic analysis of their written feedback revealed three major themes: helpfulness, accessibility, and improved attitude toward AI.

Conclusions: The PDC30 Chatbot provides quality responses to caregiver questions, which are well-received by caregivers. Conversational AI is a viable approach to improve the support of caregivers.

(JMIR Aging 2025;8:e63715) doi:10.2196/63715

KEYWORDS

Alzheimer; caregiving; chatbot; conversational artificial intelligence; dementia; digital health; health care technology; psychoeducational; medical innovations; language models; mobile phone

XSL•FO
RenderX

Introduction

Background

How to leverage technological innovation to improve support in the protracted journey of caring for someone with dementia has been identified as one of the top priorities by dementia care researchers, practitioners, and policy makers [1,2]. Dementia family caregiving is widely regarded as a prototype of chronic stress [3,4]. With the increasing loss of abilities in persons with dementia, family members transition from a "companion" to a more heavy-duty caregiver. Initially, family members may want to know more about the condition and what to plan ahead of time. As the relative's condition worsens, tips about handling different behaviors or situations and helping the relative with activities of daily living become more relevant, with end-of-life issues on the horizon further down the road. Thus, caregivers' needs change over time but are often given a time-limited (eg, 8 weekly sessions) standardized training program disregarding their specific concerns [5,6]. While some open-minded caregivers may be curious to find out more about the subject matter, others feel compelled to go through materials irrelevant to their immediate concerns (or things they already know) [7]. Moreover, traditional interventions are packaged to be delivered in regular blocks of time, which may not fit into the busy and often unpredictable schedule of many caregivers [5,6]. In addition, traditional intervention modalities often rely on face-to-face contact, a problem brutally exposed during the COVID-19 pandemic [8,9], but a problem that is generally applicable to those in rural areas.

Although over two-thirds of caregivers in a large UK survey [10] indicated an interest in using technological tools to support their health and well-being, as well as their caregiving role, few such tools are available for dementia family caregivers. A major reason cited by the respondents in support of technological support was "being able to use digital tools quickly to find answers on a regular basis" [10]. Relatedly, this study explores how caregivers react to a chatbot based on generative artificial intelligence (AI) technology, which provides information and advice on demand and can be accessed anytime anywhere.

AI Applications for Dementia Care

AI (especially natural language processing) refers to a collection of algorithms that gets computers or machines to perform autonomous actions that require human intelligence (eg, understanding complex semantics and intent, reasoning, decision-making, and self-correction). In the field of health care, the advent of AI has led to a wide range of applications, such as behavioral monitoring, risk prediction, screening, triage, diagnosis, rehabilitation, health advice chatbot, robotics, and program planning [11-21]. Such applications may be used to assist or replace human activity, or even to improve performance by enhancing efficiency and reducing human error [22-26]. Thus, AI has the potential to fill some service gaps for families with dementia as societies grapple with this public health crisis.

That said, AI is not without drawbacks. For end users, the greatest concerns are the accuracy of the output and the quality of decision-making [27,28]. An AI program is only as good as the algorithm and the data used to train it. For example, a health risk prediction program in the United States classified Black people as having lower risk profiles, compared with White people with comparable health status, resulting in many Black people being denied the proper service. The reason was due to the algorithm using health care cost as the proxy for how sick the person was, without taking into account the fact that Black people have historically been an underserved population [29].

Two reviews provide valuable insights into the development and use of AI chatbots and technology-driven solutions to support dementia care. Hoel et al [30] explored various technology-driven solutions such as social robots and tablet applications that provide activities or interactions to engage persons with dementia, most of whom live in residential care settings. Although the primary targets were persons with dementia, interactions with family caregivers were found to be enhanced through the involvement of the caregiver in the activity (eg, reminiscence therapy). Through more enjoyable interaction, the level of stress felt by family caregivers might be reduced.

The systematic review by Ruggiano et al [9] focused on chatbots, conversational AI agents that use natural language processing and large language models (both machine learning algorithms) to simulate complex human conversation [31,32]. They found 6 commercially available chatbots designed to assist people with dementia and their caregivers. Focusing on the functions and quality of these tools, their review highlighted the potential of chatbots to offer educational content for caregivers (and memory aids for patients) through accessible platforms like mobile apps and voice-activated devices. These chatbots aimed to provide users with timely information and interactive features to engage people with dementia, but a common drawback was the lack of peer-reviewed, evidence-based educational material, which undermines their trustworthiness. As a matter of fact, none of the chatbots were deployed following empirical user evaluation [9]. The information provided was often not rooted in rigorous scientific research, leaving caregivers unsure about the accuracy and reliability of the advice they receive. Based on limited programmed content, these chatbots inevitably constrain interaction with users and provide advice on a narrow range of questions [9].

Another issue is accessibility. Many chatbots require specific phrases or commands to operate effectively, limiting their flexibility and usability for a diverse range of caregivers. This rigidity can be frustrating, especially for those who are not technologically proficient. Additionally, most specialized chatbots provide only standard answers to preset questions and cannot analyze user intent and emotion. They rely heavily on users to provide specific contextual information for continued engagement.

This Study

In view of the limitations of existing chatbots, this study reports on a newly built chatbot, called PDC30 Chatbot, which serves as a care adviser on dementia caregiving using generative AI technology and an evidence-based, comprehensive knowledge base. By simply providing information and advice, without involvement in decision-making, it avoids potential risks of AI applications mentioned above. At the same time, the use of a

XSL•FO
RenderX

validated knowledge base enhances the quality of the answers and trust among users, with its information extracted by generative AI to produce a new and original response to each caregiver enquiry. As such, this new chatbot offers greater flexibility in input handling, allowing caregivers to engage in natural, conversational interactions without the need for rigid commands. Furthermore, through remarkably improved emotional and cognitive support by virtue of generative AI technology, the chatbot is capable of delivering empathetic responses and practical caregiving strategies without requiring constant caregiver input. This makes it a more autonomous and complete tool for addressing the emotional, cognitive, and practical demands of dementia caregiving.

In the following, we describe in detail the construction of the PDC30 Chatbot and its acceptability from a user perspective. A new chatbot would have little value if there is no evident advantage over other resources. Thus, the chatbot's performance in relation to handling common questions concerning dementia and caregiving was examined by comparing its responses to those of two other chatbots including the popular ChatGPT. Upon establishing its performance value, the PDC30 Chatbot was subject to an assessment of acceptability after a 2-week use by dementia family caregivers.

Methods

Chatbot Development

PDC30 Chatbot

We use the latest GPT-4o large language model by OpenAI [33], with the following prompt to define the personality traits of the chatbot (ie, a personality agent in large language models): You are a professional counselor providing advice on caregivers' emotional issues and caregiving challenges. The knowledge base is the Positive Dementia Caregiving in 30 Days Guidebook, an abridged and updated version of the Benefit-Finding Intervention manual (a psychoeducation program with an emphasis on searching for positive meanings in caregiving) which has been found to reduce caregiver burden and depression, with moderate to large effect sizes, up to 12 months [34-39] (the Benefit-Finding Intervention is a workshop-based program with a lengthy instruction manual written for the trainers. The manual was rewritten in simple language and a more concise form for general public consumption and became the knowledge base for this chatbot). After providing an answer, the chatbot is programmed to reference the source of the ideas, so as to encourage users to do more in-depth reading.

In other words, the chatbot would not answer any question unrelated to dementia and caregiving. As a generative AI chatbot, it can formulate answers to a wide range of questions based on these knowledge bases, rather than simply providing preset answers to selected questions. When irrelevant questions are asked, it would say "Sorry, I cannot answer your question." The chatbot, accessible on any device with internet access (including phone, tablet, or computer), was built on botpress.com, an open-source platform for conversational AI solutions. The chatbot serves as one of the components of a self-guided, automated web-based intervention program called Positive Dementia Caregiving in 30 Days (PDC30), which is undergoing evaluation in a global randomized controlled trial (ClinicalTrials.gov identifier NCT06409455). Hence, we call it the PDC30 Chatbot. Currently, this web-based intervention program, the chatbot included, works only in the English language.

ChatGPT and Chatbot-B as Comparators

To evaluate how well the PDC30 Chatbot works, the same questions (see below) were fed into ChatGPT-4o (as a standard of reference) [33] and another self-constructed chatbot (called Chatbot-B for convenience). As ChatGPT is well-known, we focus on introducing Chatbot-B.

Chatbot-B's design is identical to that of the PDC30 Chatbot, but the reference materials are different. The knowledge base for Chatbot-B consists of (1) 154 caregiving-related web pages by the Alzheimer's Association, (2) 17 dementia caregiving-related web pages by the US National Institute on Aging, (3) 14 caregiving topic-based manuals by the Alzheimer's Society, United Kingdom, (4) World Health Organization's iSupport manual [40], and (5) By Us For Us Guides, a collection of 15 Canada-based documents written by people with dementia and their caregivers [41] (if a document was not optimized for automated processing, its content was manually extracted and saved as a plain text file to improve the chatbot's ability to interpret and analyze the information). In other words, Chatbot-B represents a rather comprehensive and authoritative knowledge base from which answers are drawn. In this sense, Chatbot-B is like an assistant surfing the internet for relevant materials on behalf of the user and summarizes the main points for the user in far less time than the user surfing the internet himself or herself.

Together, ChatGPT and Chatbot-B provide strong reference points for assessing the quality of PDC30 responses. ChatGPT is an existing tool that can be used by caregivers in countries where it is available. It is trained on an enormous text retrieved from the internet but is known to have the risk of providing false information when such information exists on the internet. This reliance on internet-based sources makes it vulnerable to inaccuracies, particularly when it draws from unreliable content. Additionally, the chatbot may struggle with context or nuance, offering responses that are overly general or missing critical domain-specific details, which can be problematic in sensitive caregiving scenarios [9]. Chatbot-B, on the other hand, is a new chatbot dedicated to dementia and caregiving topics using authoritative materials. A caveat needs to be mentioned. There are pros and cons of including many reference materials. Up to a certain point, the benefit of including more texts levels off given the redundancy across their content. Processing time may be lengthened as the knowledge base expands in size. Different texts may also offer advice that contradicts each other. Hence, a chatbot based on a single text (ie, the PDC30 Guidebook) may not necessarily fare worse.

Testing Materials and Chatbot Responses

We constructed 21 questions commonly asked by family members by surfing the internet and subjected the three chatbots to the same questions to see how their answers compare. The

XSL•FO
RenderX

responses provided by PDC30 Chatbot and Chatbot-B are reproduced in Multimedia Appendix 1, whereas those by ChatGPT are shown in Multimedia Appendix 2. Note that due to the use of generative AI, the answers provided to the same question will vary slightly from time to time; thus, these are not fixed answers, but rather one sample of a range of possible answers by the respective chatbot. Moreover, it should be mentioned that the boldfaced headings in front of the bullet points were created by the chatbots; they did not exist in the original sources. Interestingly, at one point, the PDC30 Chatbot addressed the first author ("Dear Tak," response to the question "Why is my wife acting like a different person") and that was the only instance any of the chatbots addressed the user by name when answering questions.

For PDC30 Chatbot and Chatbot-B, one can see that despite using proper nouns (eg, mom and husband) and pronouns (eg, her), the chatbots had no problem understanding they were the care recipients. The wording in some questions was intentionally nonspecific (to mimic everyday conversation) and GPT-4o, when constrained to answer questions related to dementia and caregiving, had no difficulty grasping the meaning in the context of dementia caregiving. On the contrary, ChatGPT expectedly needed information on the specific context to construct more relevant responses (for instance, when asked about preparation for the future, it talked about financial investment unless being told that the question concerned a family having a relative newly diagnosed with dementia). Hence questions for ChatGPT were elaborated accordingly. By comparison, questions posed to PDC30 Chatbot and Chatbot-B mimicked natural conversations a lot more.

In terms of answers, occasionally there are bullet points, the relevance of which to the question is not immediately apparent (eg, Chatbot-B's points "utilize online services" and "stay engaged and active" to the question "What preparation should we make for the future?"). This issue appeared to be more common for ChatGPT, including, but not limited to: considering home safety and daily activities for the preparation for future question; encouraging activity involvement for the question on communication skills; discussing financial or legal planning and communication skills for the question on managing personal feelings and stress; asking the caregiver to join support groups, find respite, consider environmental safety, and to take breaks for himself or herself when the question was how to deal with mom's apathy; and getting adult day care, power of attorney, and medical directives when being asked how to allow husband to wander safely. Such answers may well confuse caregivers, limiting the value of ChatGPT as a consultation resource.

Moreover, ChatGPT responses tended to be repetitive, while missing some key advice to caregivers, such as avoiding confrontation and giving due recognition to the disease as the real causal agent for problematic behaviors. On the whole, the responses by PDC30 Chatbot and Chatbot-B were more concise, specific, and to the point. Considering the fact that these chatbots were providing advice on complicated matters, we think their overall performance was quite good—the results support chatbots using generative AI technology as a viable approach to offer advice to caregivers, and the results are better with topic-focused chatbots. For this reason, we focus on the relative performance of PDC30 Chatbot and Chatbot-B below.

The answers provided by these two chatbots were surprisingly similar. In fact, the answers to the question "I am mad with myself; I made so many mistakes" were identical between the two chatbots, though PDC30's answer was probably clearer with more elaboration. Though there is no straightforward way to ascertain the relative quality of the answers, we are of the opinion that compared with PDC30 Chatbot, Chatbot-B's performance was inferior in relation to the questions "I cannot accept the idea that my Mom has dementia and may leave me one day" and "I am so frustrated I can't control my emotions; I would even take out my irritation on her." One difference between the two chatbots does stand out, which is that when it comes to questions related to emotional or mental health issues (technically the above question about not being able to accept mom's condition is also an emotionally laden question), the PDC30 Chatbot provided a greater variety of suggestions. For example, pleasant event scheduling and alternative thinking are important coping strategies but were mentioned by the PDC30 Chatbot only. Other strategies including positive meaning-making and tackling unrealistic assumptions about the care recipient were also emphasized a lot more by the PDC30 Chatbot. These differences are notable, as psychoeducational interventions that include tactics taken from psychotherapy have been found, in a comprehensive meta-analysis of 131 randomized controlled trials, to be much more beneficial for dementia caregivers (especially in terms of relieving depression and promoting self-efficacy and positive gains) than psychoeducation without such tactics [2]. On the whole, the results support the value of the PDC30 Chatbot as a "counselor" for dementia caregivers. Typically, an answer was provided in 20 seconds or less, despite the volume to be scanned, suggesting that a generative AI chatbot is a rather convenient way for caregivers to obtain basic information and advice.

Some brief remarks about Chatbot-B are warranted before moving on to the main acceptability study. In terms of referencing, Alzheimer's Association materials were used in formulating responses to 14 of the 21 questions; Alzheimer's Society and By Us For Us documents were each used for 12 questions; National Institute on Aging web pages, 9 questions; and iSupport, 8 questions. Thus GPT-4o did scan the entire knowledge input before drafting its answers. Note also that Chatbot-B works only if the URLs are up to date. This can make it a less preferable choice as an intervention tool, compared with the PDC Chatbot which has a stable and well-defined knowledge source.

Acceptability Study for PDC30 Chatbot—Design and Procedure

In light of the positive results for the PDC30 Chatbot, we proceeded to the next phase to evaluate the acceptability of the chatbot. A mixed methods study was conducted, in which caregivers were shown the PDC30 Chatbot's responses to the 21 questions above, and asked to use the chatbot (accessible using a hyperlink) for two weeks, at least once a day (as this study was focused on the evaluation of the chatbot, the other components of the PDC30 intervention were not included). At

XSL•FO
RenderX

the end of 2 weeks, participants indicated, on a self-report questionnaire, their frequency of using the chatbot on a scale of 1=almost never, 2=several times a week but less than daily, 3=once a day, and 4=more than once a day. They rated the chatbot using the following questions: (1) the chatbot is easy to use, (2) the answers are easy to understand, (3) the answers are helpful, (4) overall, you are satisfied with the chatbot, and (5) you would recommend the chatbot to other caregivers (all rated from 1=strongly disagree to 5=strongly agree). Participants were also asked to write down any other thoughts they had about the chatbot; this written feedback was subject to thematic analysis [42].

Participants

Participants were recruited through posting notices on campus. The inclusion criteria are (1) aged 18 years or older, (2) self-identification as providing care to a relative with dementia, and (3) self-reported fluency in English. There was no exclusion criterion. All participants were relatives of students and staff who were taking care of a family member with dementia and who provided informed consent to participate. We stopped recruitment after the 21st participant because data saturation had evidently been reached, as the same themes kept repeating. Participants were not financially compensated but were given continuing access to the PDC30 Chatbot after the study.

Ethical Considerations

The study was approved by the Human Research Ethics Committee of The Education University of Hong Kong (reference 2021-2022-0077). Written informed consent was obtained from all participants.

Data Analysis

The two authors independently read the participants' written feedback to first become familiar with the entire set of qualitative data. After highlighting key points and making notes, they generated initial codes for each written feedback. Themes were then extracted by identifying similar codes and patterns using an inductive approach. The two authors compared their work and found that similar themes were identified. The themes were finalized and articulated [42]. As for the Likert-type questions, data were summarized using descriptive statistics.

Results

Overview

Participant characteristics and their ratings of the chatbot can be found in Table 1. Understandably, participants were predominantly child caregivers (10 daughters and 4 sons). Two caregivers were the daughters-in-law of the care recipient. Only four were spouses (one wife and three husbands) and one was a sister.

As can be seen from Table 1, the responses were overwhelmingly positive. A total of 15 out of 21 (71%) participants used the chatbot more than once per day (overall mean 3.62, SD 0.67). Participants thought the chatbot was very easy to use and most gave the highest rating (mean 4.52, SD 0.68). They also thought the advice provided was quite helpful (mean 4.29, SD 0.56) and relatively easy to understand (mean 3.81, SD 0.98). Given such positive experiences, it is no wonder that they felt rather satisfied with the chatbot (mean 4.05, SD 0.59). A total of 16 out of 21 (76%) participants would recommend or strongly recommend it to other caregivers (mean 4.24, SD 0.83). These ratings were more or less substantiated by the participants' written qualitative feedback (Table 2). Thematic analysis of the written comments revealed 3 recurring themes.

XSL•FO
RenderX

Table 1. Acceptability study sample and their ratings of the PDC30 Chatbot.

Caregiver	Age	Sex	Relationship	Q1[a]	Q2[b]	Q3[c]	Q4[d]	Q5[e]	Q6[f,g]
A	40	F	Child	4	5	4	5	4	5
B	58	M	Child	4	4	5	4	4	5
C	51	F	Child	4	5	5	5	5	4
D	36	F	Child	3	5	5	4	4	5
E	73	F	Spouse	2	5	2	3	3	3
F	50	F	Child	4	5	5	4	5	5
G	70	M	Spouse	4	5	3	4	5	4
H	46	M	Child	4	4	4	4	4	5
I	42	F	Child	4	5	4	5	4	5
J	50	F	Child	4	4	3	5	4	5
K	43	F	Child-in-law	4	5	5	4	3	4
L	44	F	Child	3	4	3	4	4	4
M	48	F	Child	2	5	4	5	3	3
N	55	F	Child-in-law	4	5	5	4	4	5
O	50	F	Child	3	4	4	4	5	5
P	72	M	Spouse	4	4	3	4	4	5
Q	71	M	Spouse	4	3	2	4	4	4
R	54	M	Child	4	5	4	5	4	3
S	68	F	Sibling	4	3	3	5	4	3
T	50	M	Child	4	5	4	4	4	3
U	55	F	Child	3	5	3	4	4	4
Mean (SD)	53.62 (11.18)	—[h]	—	3.62 (0.67)	4.52 (0.68)	3.81 (0.98)	4.29 (0.56)	4.05 (0.59)	4.24 (0.83)

[a]Q1: frequency of use.

[b]Q2: whether easy to use.

[c]Q3: answers easy to understand.

[d]Q4: answers helpful.

[e]Q5: satisfaction.

[f]Q6: likelihood to recommend to others.

[g]Except for Q1 which is scored 1-4, all other questions are scored 1-5, with higher scores indicating a better experience.

[h]Not applicable.

XSL•FO
RenderX

Table 2. Participants' written comments (reproduced in full) about the PDC30 Chatbot.

Caregiver	Written comments
A	I love it! I have to work and take care of my mom. My sister do not help. I feel very lonely and not sure what to do at times. Now I can have this chatbot to talk to any time I want - it feels like a friend. And it gives me ideas about how to handle my mom too.
B	Don't feel I should go get help from the centre's social worker again. The chatbot's advise is better! And it works 24 hours a day! I actually feel so stupid now that I never trusted AI. I am really surprised how good it is.
C	The answer refer to a guidebook but we don't have it. That part is a bit confusing. That said, I really like the fact that I can now ask anything any time.
D	Initially I have reservation. I have tried many chatbots (like banks) and they are so annoying because they pretended to help but only gave you standard replies not addressing your query. This one is different. Impressed!
E	I am very busy because my husband fell. So I don't use it much. I appreciate the advise but some suggestions are not easy to follow. But I find it very easy to use. My first thought it is strange talking to a robot. But after using it, I don't feel that anymore. When ok, I will use it more.
F	Can I keep using the chatbot, please? When my nephew mentioned this chatbot to me, I thought it is a joke. Talking to a machine? But I use it every day now. Please let me use it.
G	How come not available earlier? I wish had it 2 years ago when I was most helpless!
H	Did Tom Cruise's boss said drones do not sleep but pilots do? Well, this is a very good drone.
I	This is the first time a chatbot is REALLY answering my questions! There is some wait but waiting time reasonable.
J	Professor, thanks indeed for giving me a chance to use this chatbot. It is very helpful!
K	Most answers are relevant and informative. Quick too. Better than I thought.
L	It say it is not a therapist but it is a therapist. It help me calm down when I was very very frustrated! Thanks. I will try to use it more.
M	A good companion. Answer quality is good. But I use it only when I could not find someone to talk to. It's a backup when no one is available.
N	I can screen cap the answer and send to my husband. Then we discuss how to handle my mother-in-law. I feel more confident because I have both knowledgeable advice and husband support.
O	Frankly, I was very skeptical in the beginning. But I gradually like it more. I have a smart speaker but it does not give good answer to this type of questions.
P	Never used a chatbot. Very interesting. Good information. My wife doesn't know what I do on my phone when I am looking for help. If I call my daughter, she would hear what we said and feel upset.
Q	Some terms are difficult. I need to check dictionary. But very convenient. 24 hour service. My wife has frequent difficult behavior and I like the suggestion to ignore or take a break. Some tips for bathing and serving meal helped.
R	It's a good chatbot. I love it but 2 weeks too short to say how good it really is.
S	Not used to talk to machine. Often I have a problem but forget the chatbot is there. Forget can use my phone. Then afterwards when I see my phone, I remember! Then I ask question and get some helpful answer but sometime too late.
T	I work in IT field and so this is no surprise to me. I am glad something like this is finally available because caregiving is hard! I did not get any nonresponse and response time is very good. The answers are specific with enough elaboration and minimal repetition. These are what determine good user experience.
U	Amazing. It answer all my questions so far. But my father is mild stage. May be my questions not hard enough. Interesting to see if it can help me when things get more tough.

Theme 1: Helpfulness

This is the most dominant theme which was explicitly (caregivers A, B, I, J, L, K, M, N, P, Q, and U) or implicitly (caregivers D, F, G, H, O, R, S, and T) mentioned by 19 (90%) participants. Four (caregivers D, I, O, and T) said the chatbot's answers were appropriate, informative, and specific, unlike other chatbots that provided inadequate, irrelevant, or standardized replies. Caregiver L specifically mentioned that the chatbot helped her to calm down when severely frustrated. Her comments were echoed by caregiver Q who also referred to helpful practical suggestions by the chatbot. Caregiver F, despite being informed at the study's outset, forgot that the chatbot would continue to be available, and asked if she could keep on using it after the testing period, a gesture suggestive of the helpfulness of the chatbot. Another participant (caregiver G), who cared for his wife, lamented the late coming of this resource and imagined what difference it would have made to his situation if it had been available 2 years earlier, again testifying to the perceived helpfulness of the chatbot. By comparing the chatbot to a "very good drone," caregiver H might be saying that the chatbot got the job done (ie, providing good advice), a point verified in a subsequent email conversation with the caregiver. Perhaps a bit exaggerating, caregiver B even complimented the chatbot by saying that its advice was better than the social worker's. One wife caregiver (E), however, thought that some of the suggestions by the chatbot were not easy to follow, which was not entirely surprising as behavioral

XSL•FO
RenderX

and cognitive change takes time and participants did not have the full intervention program to assist them. On the positive side, this comment suggests that the answers were relevant, just not easy to follow through.

Theme 2: Accessibility

In total, 11 (52%) caregivers mentioned, in one way or another, the usability of the chatbot, including being user-friendly (caregiver E), lack of (or low rates of) nonresponse (caregiver T), reasonable waiting time before getting a response (caregivers I, K, and T), and most of all, the chatbot's 24/7 accessibility (caregivers A, B, C, H, and Q). Caregiver H even paraphrased a conversation in the movie "Top Gun: Maverick" to illustrate this point.

Unexpectedly, two caregivers mentioned functional merits unrelated to the chatbot itself. Caregiver N talked about the convenience of saving the chatbot's responses and sharing them with other family members. This action, as alluded to by this caregiver, might encourage more involvement by other family members and support to the main caregiver. In addition, caregiver P made an interesting point about the privacy afforded by using the chatbot on his smartphone. His comment was a reminder of the dilemma faced by many caregivers, especially those in crowded living conditions such as the case in Hong Kong, when trying to seek help when the care recipient is nearby. Conversations, whether over the phone or face-to-face, may be overheard by the care recipient, who, in turn, reacts with more behavior problems. The chatbot operated on the phone offers caregivers private space to get help when necessary.

Theme 3: Attitude Toward AI

Five participants (caregivers B, D, E, F, and O) expressed that they were initially skeptical of AI but the experience changed their perception of it. For example, caregiver B started with a mistrust of AI while caregiver F initially found seeking advice from a machine a ridiculous idea. After trying it, both became excited by the technology and were using the chatbot more than once a day. A related sentiment was expressed by another caregiver, S. She was not used to talking to a machine, and after having access to the chatbot, kept forgetting about it. Yet, the undertone of the feedback was that she found the chatbot responses to be useful and lamented that occasionally she had not taken timely advantage of the resource.

Not everyone was as receptive to the technology. Caregiver M did not reject the chatbot but reserved it for occasions when she could not find people to talk to. Nevertheless, the feedback overall suggested that experience with the chatbot induced a favorable attitude toward it in those who questioned its value to begin with.

Discussion

Principal Findings

This study demonstrates that using a chatbot to help dementia caregivers is a viable approach. Taking advantage of AI development, we were able to build the PDC30 Chatbot that functions as a "counselor" to caregivers (ie, its purpose), by applying a personality agent to constrain the chatbot's behavior and by feeding it with an appropriate knowledge base. The generative AI algorithm summarized the points using "its own" words and organized them using headings to facilitate reading. At first, it might seem that feeding it with more resource materials would help it formulate better responses—for this reason, two other chatbots were included for comparison, namely ChatGPT and Chatbot-B incorporating insider perspectives (those from patients and caregivers) and guidelines by several authoritative agencies. In our testing using 21 common questions by caregivers, there was indeed a tendency for Chatbot-B to generate slightly longer responses than PDC30 Chatbot, though this was not always the case. Nevertheless, length per se does not determine quality, as longer answers may contain irrelevant and repetitive points, as we have seen, especially in the case of ChatGPT. On the whole, we think that PDC30 Chatbot and Chatbot-B were superior to ChatGPT and performed similarly, with PDC30 having a further edge on emotional and mental health issues. In addition, the PDC30 Chatbot was favorably received in an acceptability study of 21 dementia family caregivers after using it more than once a day for two weeks. The chatbot was considered user-friendly, with its responses helpful and easy to understand. Their written feedback about the chatbot centered on three themes: it was helpful and accessible, and it improved their perception of AI as a helping agent. Overall, the experience was rather satisfactory and the participants would strongly recommend it to other caregivers.

Both PDC30 Chatbot and Chatbot-B provide citations of the sources they refer to when formulating responses. A caveat needs to be mentioned. While the citation system may prompt caregivers to do more reading of the relevant materials, it is good only to the extent that the sources can be readily located. As explained in Table 1 footnote, GPT-4o cannot name the author of the web pages cited. If the source is a document, this was handled by the way the document is named, such as "Alzheimer's Society Making Your Home Dementia Friendly"—in other words, the citation is to the name of the file used. If it is just a web page, GPT-4o provides a citation using the first title that appears on the web page; the agency or author producing the web page is omitted—this is primarily due to limitations in GPT-4o's design and functionality. First, GPT-4o cannot read logos or metadata from images on web pages, where authors' names and relevant citation details are often embedded. If the website does not contain alternative text, GPT-4o cannot read the images. Second, while the model processes textual data, it may only effectively distinguish between content and metadata, such as author names or publication details, if explicitly mentioned in the text. The website needs to state the metadata clearly. Additionally, the training data (ie, web page metadata) may not include structured citation formats, making it difficult for the model to recognize and extract authors' names reliably. Furthermore, the design of GPT-4o may include privacy considerations that prevent it from naming individuals unless explicitly included in the training data or provided in a clear, structured format. To address this issue, converting all relevant web pages into text files with clear filenames and structured citation information can help ensure that author names and other citation details are correctly identified and referenced by the model. On the contrary, the PDC30 Chatbot is to be used

XSL•FO
RenderX

together with the Guidebook in actual operation—both are to be hosted on the same intervention website. Thus, the citation problem mentioned above does not apply.

Perhaps a more important consideration in terms of operation is that Chatbot-B requires constant monitoring and updating, should organizations add new web pages, remove certain web pages, or change the URLs of the web pages—there are indeed a very large number of web pages to monitor. Worse, changes to the content of a web page with the same URL will be hard to detect, unless a "last updated date" is provided. By comparison, the PDC30 Chatbot is much easier to maintain, as it is based on only one text.

The PDC30 Chatbot was built with the intention to be included in a randomized controlled trial, and so we subject it to an evaluation of acceptability by caregivers. All but one caregiver provided highly positive feedback about the chatbot; the only exception came from a wife caregiver whose use of the chatbot was disrupted by her husband's fall early in the trial period (had she contacted the research team then, we would have adjusted her trial period so that she would have been able to try the chatbot at a more convenient time. Unfortunately, we were not aware of the issue until she turned in her responses). On the whole, the caregivers found the chatbot easy to use (including the wife caregiver mentioned above) and the answers were informative, relevant to their needs, and easy to understand. Many particularly appreciated the chatbot's helpfulness and 24/7 accessibility, with some further commenting on its lack of nonresponse and reasonable waiting times before responses were generated. As a result, some caregivers acquired a much more favorable attitude toward AI technology because of using the chatbot. The great majority reported high levels of satisfaction with the chatbot and would recommend it to other caregivers. Thus, the PDC30 Chatbot was well-received by caregivers who used it. The chatbot's frequent use was also a testimony to its reception among the caregivers.

Limitations

Despite the encouraging findings, a few limitations need to be mentioned. First, the sample was predominantly adult children. We do not think the paucity of spousal or sibling caregivers was due to the relative lack of digital literacy among older caregivers as there is no skill required to use the chatbot (other than typing questions into the input space). The older, spousal or sibling caregivers who participated generally found the chatbot easy to use. Rather, we think that this reflects cohort differences in the preference for digital material [43]. Furthermore, the composition of the sample is not as one-sided as it seems, as about two-thirds of the caregivers in this community are children [44]. That said, there is no doubt a certain degree of self-selection as not everyone prefers digital resources (we did have two caregivers in the sample who were initially doubtful of the technology). A future study focusing on older caregivers, especially those with little digital literacy, is warranted.

Second, the chatbot was constructed to speak English only, in alignment with the language used in the parent intervention program being evaluated in a global trial. It is probably too harsh to consider this a limitation but in the context of this study, fluent English-speaking persons do not represent the caregiver population in Hong Kong. Due to the time pressure to launch the clinical trial, we also did not recruit caregivers from English-speaking countries. Although the generalizability of the findings may be limited, the PDC30 Chatbot's responses shown in Table 1 suggest that the chatbot should be quite suitable for caregivers in other countries as well.

Conclusions

To the best of our knowledge, this is the first study reporting the development of a psychoeducational chatbot for dementia caregivers and testing its functioning and acceptability among caregivers. The convenience of delivering advice to caregivers by AI-driven chatbots is an approach that needs to be explored further by the field. Services for families with dementia are lacking, especially in resource-poor countries, and even when services are available, many caregivers do not use the services due to lack of time, services not meeting their needs, or simply wanting to locate information themselves [45]. Good chatbots may well fill some of the gaps and get help to these caregivers around the world. There is one more important feature that makes chatbots an advantageous option—only small amounts of data are transmitted each time and hence chatbots are especially suitable for low-income countries and rural areas in general where internet coverage remains an issue.

Acknowledgments

The study was supported by the General Research Fund (18622322) of the Research Grants Council of Hong Kong to the first author. The funder and the sponsor play no roles in conceptual formulation, research design and execution, data acquisition and analysis, and report write-up.

Data Availability

All data generated or analyzed during this study are included in this published article.

Authors' Contributions

STC conceptualized the study and designed the study and the two chatbots. STC and PHFN cosupervised the programmer to set up the chatbots. Both authors contributed to the writing of this manuscript, with STC taking the leadership. Both authors approved the submission and take responsibility for the accuracy and integrity of the work. No professional writer was employed.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Responses provided by PDC30 (Positive Dementia Caregiving in 30 Days) Chatbot and Chatbot-B.
[DOCX File , 55 KB - aging_v8i1e63715_app1.docx]

Multimedia Appendix 2
Responses provided by ChatGPT.
[DOCX File , 57 KB - aging_v8i1e63715_app2.docx]

References

1. Harvath TA, Mongoven JM, Bidwell JT, Cothran FA, Sexson KE, Mason DJ, et al. Research priorities in family caregiving: process and outcomes of a conference on family-centered care across the trajectory of serious illness. Gerontologist 2020;60(Suppl 1):S5-S13 [FREE Full text] [doi: 10.1093/geront/gnz138] [Medline: 32057081]
2. Cheng ST, Li KK, Losada A, Zhang F, Au A, Thompson LW, et al. The effectiveness of nonpharmacological interventions for informal dementia caregivers: an updated systematic review and meta-analysis. Psychol Aging 2020;35(1):55-77. [doi: 10.1037/pag0000401] [Medline: 31985249]
3. Cheng ST. Dementia caregiver burden: a research update and critical analysis. Curr Psychiatry Rep 2017;19(9):64 [FREE Full text] [doi: 10.1007/s11920-017-0818-2] [Medline: 28795386]
4. Vitaliano PP, Young HM, Zhang J. Is caregiving a risk factor for illness? Curr Dir Psychol Sci 2004;13(1):13-16. [doi: 10.1111/j.0963-7214.2004.01301004.x]
5. Cheng ST. Providing Optimal Support to Caregivers. Germany: Springer Science and Business Media LLC; 2021.
6. Cheng ST, Li K, Or PPL, Losada A. Do caregiver interventions improve outcomes in relatives with dementia and mild cognitive impairment? A comprehensive systematic review and meta-analysis. Psychol Aging 2022;37(8):929-953. [doi: 10.1037/pag0000696] [Medline: 35771500]
7. Yu Y, Xiao L, Ullah S, Meyer C, Wang J, Pot AM, et al. The experiences of informal caregivers of people with dementia in web-based psychoeducation programs: systematic review and metasynthesis. JMIR Aging 2023;6:e47152 [FREE Full text] [doi: 10.2196/47152] [Medline: 37247218]
8. Liang J, Aranda MP. The use of telehealth among people living with dementia-caregiver dyads during the COVID-19 pandemic: scoping review. J Med Internet Res 2023;25:e45045 [FREE Full text] [doi: 10.2196/45045] [Medline: 37227755]
9. Ruggiano N, Brown EL, Roberts L, Suarez VF, Luo Y, Hao Z, et al. Chatbots to support people with dementia and their caregivers: systematic review of functions and quality. J Med Internet Res 2021;23(6):e25006 [FREE Full text] [doi: 10.2196/25006] [Medline: 34081019]
10. Egan KJ, Clark P, Deen Z, Dutu CP, Wilson G, McCann L, et al. Understanding current needs and future expectations of informal caregivers for technology to support health and well-being: national survey study. JMIR Aging 2022;5(1):e15413 [FREE Full text] [doi: 10.2196/15413] [Medline: 35084339]
11. Ng PHF, Chen PQ, Sin ZPT, Lai SHS, Cheng ASK. Smart work injury management (SWIM) system: a machine learning approach for the prediction of sick leave and rehabilitation plan. Bioengineering (Basel) 2023;10(2):172 [FREE Full text] [doi: 10.3390/bioengineering10020172] [Medline: 36829666]
12. Cheng ASK, Ng PHF, Sin ZPT, Lai SHS, Law SW. Smart work injury management (SWIM) system: artificial intelligence in work disability management. J Occup Rehabil 2020;30(3):354-361. [doi: 10.1007/s10926-020-09886-y] [Medline: 32236811]
13. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digital Med 2020;3:17 [FREE Full text] [doi: 10.1038/s41746-020-0221-y] [Medline: 32047862]
14. Fernandes M, Vieira SM, Leite F, Palos C, Finkelstein S, Sousa JMC. Clinical decision support systems for triage in the emergency department using intelligent systems: a review. Artif Intell Med 2020;102:101762. [doi: 10.1016/j.artmed.2019.101762] [Medline: 31980099]
15. Bu X, Ng PH, Xu W, Cheng Q, Chen PQ, Cheng AS, et al. The effectiveness of virtual reality-based interventions in rehabilitation management of breast cancer survivors: systematic review and meta-analysis. JMIR Serious Games 2022;10(1):e31395 [FREE Full text] [doi: 10.2196/31395] [Medline: 35225817]
16. Bu X, Ng PHF, Tong Y, Chen PQ, Fan R, Tang Q, et al. A mobile-based virtual reality speech rehabilitation app for patients with aphasia after stroke: development and pilot usability study. JMIR Serious Games 2022;10(2):e30196 [FREE Full text] [doi: 10.2196/30196] [Medline: 35389349]
17. Hird N, Osaki T, Ghosh S, Palaniappan SK, Maeda K. Enabling personalization for digital cognitive stimulation to support communication with people with dementia: pilot intervention study as a prelude to AI development. JMIR Form Res 2024;8:e51732 [FREE Full text] [doi: 10.2196/51732] [Medline: 38227357]

XSL•FO
RenderX

18. Anderson MS, Bankole A, Homdee N, Mitchell BA, Byfield GE, Lach J. Dementia caregiver experiences and recommendations for using the behavioral and environmental sensing and intervention system at home: usability and acceptability study. JMIR Aging 2021;4(4):e30353 [FREE Full text] [doi: 10.2196/30353] [Medline: 34874886]
19. Lee H, Chung MA, Kim H, Nam EW. The effect of cognitive function health care using artificial intelligence robots for older adults: systematic review and meta-analysis. JMIR Aging 2022;5(2):e38896 [FREE Full text] [doi: 10.2196/38896] [Medline: 35672268]
20. Wilmink G, Dupey K, Alkire S, Grote J, Zobel G, Fillit HM, et al. Artificial intelligence-powered digital health platform and wearable devices improve outcomes for older adults in assisted living communities: pilot intervention study. JMIR Aging 2020;3(2):e19554 [FREE Full text] [doi: 10.2196/19554] [Medline: 32723711]
21. Wong KLY, Hung L, Wong J, Park J, Alfares H, Zhao Y, et al. Adoption of artificial intelligence-enabled robots in long-term care homes by health care providers: scoping review. JMIR Aging 2024;7:e55257 [FREE Full text] [doi: 10.2196/55257] [Medline: 39190455]
22. Kilsdonk E, Peute LW, Riezebos RJ, Kremer LC, Jaspers MWM. Uncovering healthcare practitioners' information processing using the think-aloud method: from paper-based guideline to clinical decision support system. Int J Med Inform 2016;86:10-19 [FREE Full text] [doi: 10.1016/j.ijmedinf.2015.11.011] [Medline: 26725690]
23. Dowding D, Mitchell N, Randell R, Foster R, Lattimer V, Thompson C. Nurses' use of computerised clinical decision support systems: a case site analysis. J Clin Nurs 2009;18(8):1159-1167. [doi: 10.1111/j.1365-2702.2008.02607.x] [Medline: 19320785]
24. Ash JS, Sittig DF, Campbell EM, Guappone KP, Dykstra RH. Some unintended consequences of clinical decision support systems. AMIA Annu Symp Proc 2007;2007:26-30 [FREE Full text] [Medline: 18693791]
25. Cook DA, Teixeira MT, Heale BS, Cimino JJ, Fiol GD. Context-sensitive decision support (infobuttons) in electronic health records: a systematic review. J Am Med Inform Assoc 2017;24(2):460-468 [FREE Full text] [doi: 10.1093/jamia/ocw104] [Medline: 27497794]
26. Légat L, Van Laere S, Nyssen M, Steurbaut S, Dupont AG, Cornu P. Clinical decision support systems for drug allergy checking: systematic review. J Med Internet Res 2018;20(9):e258 [FREE Full text] [doi: 10.2196/jmir.8206] [Medline: 30194058]
27. Isbanner S, O'Shaughnessy P, Steel D, Wilcock S, Carter S. The adoption of artificial intelligence in health care and social services in Australia: findings from a methodologically innovative national survey of values and attitudes (the AVA-AI Study). J Med Internet Res 2022;24(8):e37611 [FREE Full text] [doi: 10.2196/37611] [Medline: 35994331]
28. Yeung YYK, Chen PQ, Ng PHF, Cheng ASK. Evaluation of the accuracy of the smart work injury management (SWIM) system to assist case managers in predicting the work disability of injured workers. J Occup Rehabil 2024. [doi: 10.1007/s10926-024-10199-7] [Medline: 38874680]
29. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science 2019;366(6464):447-453. [doi: 10.1126/science.aax2342] [Medline: 31649194]
30. Hoel V, Feunou CM, Wolf-Ostermann K. Technology-driven solutions to prompt conversation, aid communication and support interaction for people with dementia and their caregivers: a systematic literature review. BMC Geriatr 2021 Mar 04;21(1):157. [doi: 10.1186/s12877-021-02105-0] [Medline: 33663401]
31. Rossen B, Lok B. A crowdsourcing method to develop virtual human conversational agents. Int J Hum-Comput Stud 2012;70(4):301-319. [doi: 10.1016/j.ijhcs.2011.11.004]
32. Shan Y, Ji M, Xie W, Qian X, Li R, Zhang X, et al. Language use in conversational agent-based health communication: systematic review. J Med Internet Res 2022;24(7):e37403 [FREE Full text] [doi: 10.2196/37403] [Medline: 35802407]
33. OpenAI. Introducing GPT-4o and more tools to ChatGPT free users. URL: https://openai.com/index/gpt-4o-and-more-tools-to-chatgpt-free/ [accessed 2024-05-20]
34. Cheng ST, Lau RWL, Mak EPM, Ng NSS, Lam LCW. Benefit-finding intervention for Alzheimer caregivers: conceptual framework, implementation issues, and preliminary efficacy. Gerontologist 2014;54(6):1049-1058. [doi: 10.1093/geront/gnu018] [Medline: 24688081]
35. Cheng ST, Lau RW, Mak EP, Ng NS, Lam LC, Fung HH, et al. A benefit-finding intervention for family caregivers of persons with Alzheimer disease: study protocol of a randomized controlled trial. Trials 2012;13(1):98 [FREE Full text] [doi: 10.1186/1745-6215-13-98] [Medline: 22747914]
36. Cheng ST, Fung HH, Chan WC, Lam LCW. Short-term effects of a gain-focused reappraisal intervention for dementia caregivers: a double-blind cluster-randomized controlled trial. Am J Geriatr Psychiatry 2016;24(9):740-750. [doi: 10.1016/j.jagp.2016.04.012] [Medline: 27401052]
37. Cheng ST, Chan WC, Lam LC. Long-term outcomes of the benefit-finding group intervention for Alzheimer family caregivers: a cluster-randomized double-blind controlled trial. Am J Geriatr Psychiatry 2019;27(9):984-994. [doi: 10.1016/j.jagp.2019.03.013] [Medline: 31076215]
38. Cheng ST, Mak EPM, Kwok T, Fung H, Lam LCW. Benefit-finding intervention delivered individually to Alzheimer family caregivers: longer-term outcomes of a randomized double-blind controlled trial. J Gerontol B Psychol Sci Soc Sci 2020;75(9):1884-1893. [doi: 10.1093/geronb/gbz118] [Medline: 31556447]

XSL•FO
RenderX

39. Cheng ST, Mak EPM, Fung HH, Kwok T, Lee DTF, Lam LCW. Benefit-finding and effect on caregiver depression: a double-blind randomized controlled trial. J Consult Clin Psychol 2017;85(5):521-529. [doi: 10.1037/ccp0000176] [Medline: 28287803]
40. Pot AM, Gallagher-Thompson D, Xiao LD, Willemse BM, Rosier I, Mehta KM, et al. iSupport: a WHO global online intervention for informal caregivers of people with dementia. World Psychiatry 2019;18(3):365-366 [FREE Full text] [doi: 10.1002/wps.20684] [Medline: 31496096]
41. Murray Alzheimer ResearchEducation Program. By Us For Us Guides. URL: https://the-ria.ca/resources/by-us-for-us-guides/ [accessed 2024-02-06]
42. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
43. Wu K, Su Y, Chu F, Chen AT, Zaslavsky O. Behavioral change factors and retention in web-based interventions for informal caregivers of people living with dementia: scoping review. J Med Internet Res 2022;24(7):e38595 [FREE Full text] [doi: 10.2196/38595] [Medline: 35797100]
44. Cheng ST, Lam LCW, Kwok T, Ng NSS, Fung AWT. The social networks of Hong Kong Chinese family caregivers of Alzheimer's disease: correlates with positive gains and burden. Gerontologist 2013;53(6):998-1008. [doi: 10.1093/geront/gns195] [Medline: 23371974]
45. Lim J, Goh J, Chionh HL, Yap P. Why do patients and their families not use services for dementia? Perspectives from a developed Asian country. Int Psychogeriatr 2012;24(10):1571-1580. [doi: 10.1017/s1041610212000919]

Abbreviations

AI: artificial intelligence
PDC30: Positive Dementia Caregiving in 30 Days

Edited by Y Jiang; submitted 27.06.24; peer-reviewed by P Ramaswamy, Y Shan; comments to author 22.09.24; revised version received 01.11.24; accepted 12.11.24; published 06.01.25.

Please cite as:
Cheng ST, Ng PHF
The PDC30 Chatbot—Development of a Psychoeducational Resource on Dementia Caregiving Among Family Caregivers: Mixed Methods Acceptability Study
JMIR Aging 2025;8:e63715
URL: https://aging.jmir.org/2025/1/e63715
doi:10.2196/63715
PMID:

©Sheung-Tak Cheng, Peter H F Ng. Originally published in JMIR Aging (https://aging.jmir.org), 06.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Designing a Multimodal and Culturally Relevant Alzheimer Disease and Related Dementia Generative Artificial Intelligence Tool for Black American Informal Caregivers: Cognitive Walk-Through Usability Study

Cristina Bosco[1], MSc; Ege Otenen[1], MSc; John Osorio Torres[1], MSc; Vivian Nguyen[1], BSc; Darshil Chheda[1], MSc; Xinran Peng[1], MSc; Nenette M Jessup[2], MPH; Anna K Himes[2], BS; Bianca Cureton[2], BSPH; Yvonne Lu[2], PhD; Carl V Hill[3], PhD; Hugh C Hendrie[4], MD; Priscilla A Barnes[5], PhD; Patrick C Shih[1], PhD

[1]Luddy School of Informatics, Computing, and Engineering, Indiana University, Bloomington, IN, United States

[2]School of Nursing, Indiana University, Indianapolis, IN, United States

[3]Alzheimer's Association, Chicago, IL, United States

[4]School of Medicine, Indiana University, Indianapolis, IN, United States

[5]School of Public Health, Indiana University, Bloomington, IN, United States

Corresponding Author:
Cristina Bosco, MSc
Luddy School of Informatics, Computing, and Engineering
Indiana University
700 N Woodlawn Ave
Bloomington, IN, 4740
United States
Phone: 1 812 856 5754
Email: cribosco@iu.edu

Abstract

Background: Many members of Black American communities, faced with the high prevalence of Alzheimer disease and related dementias (ADRD) within their demographic, find themselves taking on the role of informal caregivers. Despite being the primary individuals responsible for the care of individuals with ADRD, these caregivers often lack sufficient knowledge about ADRD-related health literacy and feel ill-prepared for their caregiving responsibilities. Generative AI has become a new promising technological innovation in the health care domain, particularly for improving health literacy; however, some generative AI developments might lead to increased bias and potential harm toward Black American communities. Therefore, rigorous development of generative AI tools to support the Black American community is needed.

Objective: The goal of this study is to test Lola, a multimodal mobile app, which, by relying on generative AI, facilitates access to ADRD-related health information by enabling speech and text as inputs and providing auditory, textual, and visual outputs.

Methods: To test our mobile app, we used the cognitive walk-through methodology, and we recruited 15 informal ADRD caregivers who were older than 50 years and part of the Black American community living within the region. We asked them to perform 3 tasks on the mobile app (ie, searching for an article on brain health, searching for local events, and finally, searching for opportunities to participate in scientific research in their area), then we recorded their opinions and impressions. The main aspects to be evaluated were the mobile app's usability, accessibility, cultural relevance, and adoption.

Results: Our findings highlight the users' need for a system that enables interaction with different modalities, the need for a system that can provide personalized and culturally and contextually relevant information, and the role of community and physical spaces in increasing the use of Lola.

Conclusions: Our study shows that, when designing for Black American older adults, a multimodal interaction with the generative AI system can allow individuals to choose their own interaction way and style based upon their interaction preferences and external constraints. This flexibility of interaction modes can guarantee an inclusive and engaging generative AI experience.

(JMIR Aging 2025;8:e60566) doi:10.2196/60566

XSL•FO
RenderX

KEYWORDS

multimodality; artificial intelligence; AI; generative AI; usability; black; African American; cultural; Alzheimer's; dementia; caregivers; mobile app; interaction; cognition; user opinion; geriatrics; smartphone; mHealth; digital health; aging

Introduction

Background

Alzheimer disease and related dementias (ADRD) are brain health conditions impacting cognitive abilities such as memory, language, problem-solving, and executive functions [1,2]. Within the United States, older Black American and Black adults have been identified as being twice as likely as individuals from other ethnic groups to develop ADRD over the course of their lifetime [3-5]. In a study conducted with 59,555 individuals, Black American adults had a 65% greater risk of developing ADRD (with 26.6/1000 Black American adults developing ADRD, compared with rates of 19.6/1000 for Latinos, 19.6/1000 for Pacific Islanders, 19.3/1000 for Whites, and 15.2/1000 for Asian Americans) [6].

However, despite being significantly more likely to develop ADRD, Black Americans exhibit lower levels of ADRD-related health literacy compared with other demographic groups [7]. This disparity impacts informal caregivers who often find themselves unprepared and uninformed about ADRD treatment and diagnosis, as well as about ADRD-related sources and help [8]. Therefore, providing essential medical information to these caregivers is crucial in combating the disease.

Considering the potential of new technologies to aid patients and caregivers with understanding medical language by simplifying content and removing jargon [9], generative artificial intelligence (AI) tools, like ChatGPT developed by OpenAI, could offer innovative opportunities in health care, particularly for improving ADRD-related health literacy. Indeed, recent studies proposed using ChatGPT for reliable and comprehensive health information [10-13].

However, relying solely on generative AI to improve underrepresented communities' health literacy can lead to several issues like providing inaccurate or false health information [14,15]. Additionally, studies have shown that using generative AI exclusively for health diagnosis and treatment can reinforce sexist and racial stereotypes and biases [16-18].

For instance, predictive models trained on historical data that lack representation from diverse demographics can reinforce existing biases in diagnosis and treatment plans [19,20]. The exclusion of certain populations in the studies and design process can lead to misdiagnosis and inappropriate treatment [21-23], exacerbating health inequalities rather than reducing them [20,24,25]. Therefore, it is crucial to ensure greater diversity in model training data and include diverse participants in studies and design processes to improve the generalizability of findings [26] and the fit of the technology use for the target population.

Therefore, given the potential of generative AI to improve health literacy and mitigate health inequalities among underrepresented populations, such as the Black American community, it is imperative to design inclusive ADRD health literacy interventions.

Previous studies, especially in the domain of human-centered design, highlight the importance of integrating cultural insights into health care technologies for this community [27-31]. Hence, it is crucial to tailor generative AI–based health interventions to meet the unique needs of Black American informal caregivers of individuals with ADRD.

Thus, this study identified the needs of Black American caregivers of individuals with ADRD by testing Lola, an ADRD-focused generative AI tool. Lola is a multimodal mobile app that provides access to health information from the Alzheimer's Association online repository [1] via text, chatbot, and voice assistant.

Related Work

Generative AI in Health Care

Generative AI has become a transformative force in health care, offering innovative solutions in diagnostics, personalized medicine, and patient care. Recent studies have highlighted the role of generative models in enhancing diagnostic accuracy by analyzing complex medical data and generating synthetic patient data to train other AI systems. For example, a study by Chen and Esmaeilzadeh [32] demonstrated that generative models like generative adversarial networks could significantly improve the detection and diagnosis of diseases by creating high-quality medical images from existing data. Additionally, variational autoencoders have been used to generate realistic medical imaging data, further aiding disease research and model training. Moreover, the use of large language models (LLMs) such as GPT-4 has shown potential for assisting clinicians by generating comprehensive patient histories and suggesting personalized treatment plans [33].

Moreover, the use of LLMs such as generative pretrained transformer (GPT)-4 has shown potential for assisting clinicians by generating comprehensive patient histories and suggesting personalized treatment plans [33]. Studies have also explored the use of LLMs in generating patient education materials and automating routine documentation tasks [34]. Despite still being in its developmental stages, generative AI holds significant promise for enhancing the accessibility of health care [10-13].

Human-Centered Generative AI

Generative AI models are trained on large data sets that may contain biases, further perpetuating stereotypes and disparities in health care [20,24]. More diversity is needed in both data used to train AI models and participants included in studies to ensure the generalizability of findings [26]. Additionally, addressing ethical and regulatory considerations is crucial when deploying AI in health care.

The integration of human-centered design principles into generative AI applications is crucial to ensure these technologies are accessible, intuitive, and beneficial to end users. Research has emphasized the importance of involving end users in the development process to create AI systems that meet their needs

XSL•FO
RenderX

and preferences. For instance, a study explored the application of ChatGPT for developing digital health interventions, finding that user involvement in the design phase led to higher satisfaction and engagement rates among patients [35]. Specifically, user feedback was gathered through focus groups and iterative testing phases, allowing developers to continuously refine the AI user interface and functionality.

Generative AI and the African American and Black Community

Inequalities have been observed in health care and accessibility to these services for underrepresented communities [36-40]. AI systems that initially aim to reduce the disparities in health care also present challenges for health care accessibility for Black American communities.

Addressing the unique health care challenges faced by Black American communities through generative AI is a growing area of interest. It is essential to ensure that these AI systems do not perpetuate existing biases or disparities. Studies have highlighted the potential of generative AI to provide culturally sensitive health care solutions. For example, a study explored the use of mobile health [41] interventions to mitigate the impact of COVID-19 in Black American communities, emphasizing the need for community-specific content and trusted communication channels to improve health outcomes.

Informal Caregivers of Individuals With ADRD and Their Use of Technology

In 2023, approximately 15 million people served as caregivers for individuals with ADRD [1], with African Americans accounting for one-tenth of these caregivers [1]. This unpaid work amounted to an estimated 16 billion hours [1]. Caregivers often feel unprepared for their roles [42], lacking information on diagnosis, treatment, and available resources [8,43].

Technology has been crucial in supporting these caregivers [44], and technologies have been developed to reduce caregiver stress and anxiety [45] and facilitate access to information for caregivers [46]. However, these technologies targeted different demographics than African American and Black populations. For instance, Hong et al [45] developed and tested a technological intervention aimed at reducing the stress and anxiety related with caregiving by focusing exclusively on Chinese participants. On a similar note, Boutilier et al [46] developed and tested a web app to help informal caregivers document information on their ADRD-diagnosed patients and share important changes with the care team. Boutilier et al [46] tested this system with mainly non-Hispanic White participants, with only 1% of participants being Black Americans.

However, no study has focused on Black American caregivers, which is extremely important and challenging at the same time, considering this population's lower level of technology use and adoption [47]. Additionally, a recent systematic review of technological interventions (mainly mobile apps) [48] found that most of these systems did not present any culturally relevant health information and were characterized by low accessibility and low readability.

Previous studies have shown that online health communities (OHCs) offer emotional and instrumental support to caregivers of individuals with ADRD [49] and chronic diseases [50], providing a platform to share experiences and find support from peers [49,51]. Furthermore, only a few technological support interventions for informal caregivers focus on ADRD-related information. Most aim to enhance caregivers' psychological well-being by providing stress management strategies [52] or developing social games to promote physical activity and mutual support networks [53].

We identified 2 gaps in the previous literature. Although OHCs offer instrumental support by sharing caregiving tips, they often lack localized, actionable ADRD information or personalized support tailored to caregivers' and patients' specific needs. Additionally, there is limited representation of Black American caregivers in research, raising concerns about the ability to understand and address their sociocultural background as well as their unique needs compared with other populations, which is vital for developing health technologies for them [49].

Methods

Recruitment

We conducted recruitment and research sessions in close collaboration with a community engagement team who ensured that participants were all caregivers in some capacity for individuals living with ADRD and fulfilled the inclusion criteria. To meet the inclusion criteria, participants had to be older than 50 years, identify as Black Americans, reside in the target area, and actively provide care for a person with ADRD. Throughout this study, emphasis was placed on collaboration with the Black American community within the region. Consequently, 2 community engagement specialists joined the research team to organize monthly community advisory team board meetings to maintain community involvement and awareness of research insights and findings. Prior to implementing any major design decisions, the advisory team board was consulted to ensure cultural relevance. These measures enabled the research team to make more culturally competent decisions.

Ethical Considerations

Ethical approval was provided by the institutional review board at Indiana University (approval number 12241), All participants received a full explanation of the consent procedures, and any specials needs were accommodated during sessions. Participants were compensated with a US $50 gift card for taking part. Participants' privacy and confidentiality were protected by anonymizing the data, changing the wording of the quotes presented in the Results section, and deleting all the audio recordings after the transcriptions were completed.

Data Collection and Analysis

We performed an adapted cognitive walkthrough with users [54,55]. We conducted 15 sessions in 2 regions in the Midwest region of the United States. Each session was completed in a 1-on-1 meeting that involved 1 researcher and 1 notetaker with the participant. Each session lasted approximately 30 minutes. Although we were aware that longer testing sessions might guarantee the collection of more comprehensive feedback, we

XSL•FO
RenderX

made sure to construct the tasks and testing in a way that allowed the emergence of in-depth insights and findings. Moreover, given the older age of our participants, we did not want to induce cognitive fatigue.

During a session, each participant was introduced to the platform and a set of 3 tasks. Those 3 tasks consisted of searching for an article about brain health, searching for ADRD-related local events, and finally, searching for opportunities to participate in scientific research.

As participants performed the tasks, they were asked to voice their thoughts out loud in real time ("think aloud" method [56]), and after each task was complete, there was a brief semistructured discussion centered around 4 tenets: (1) usability (eg, Did you find anything in this task confusing? In what way?), (2) accessibility (eg, How do you expect Lola to answer you? Did you find the information easy to understand? If not, why?), (3) cultural relevance (eg, Did you find the tone of interactions appropriate? If not, how so? How do you think this platform can be made better for the Black American community?), (4) adoption (eg, What do you think would be a good way of spreading knowledge about this so that people you know can properly adopt it and use it?)

Each session was audio-recorded and transcribed. The generated transcripts were spell-checked. Guided by the principles of usability, accessibility, and relevance, 6 researchers coded the transcripts using the open-source software Taguette [57]. The team took an analytical, iterative approach by discussing the codes and consolidating them into themes [58] through second and third rounds of discussion. The resulting major themes about the needs of the users are described in the Results section.

System Design

Overview

Generative AI is at the forefront of technological advancements in health care applications, specifically in the development of interactive tools such as chatbots and voice assistants. In the context of Lola, an innovative app tailored for Black Americans coping with ADRD, generative AI plays a crucial role. It powers both the chatbot and voice assistant, enabling them to deliver intuitive and empathetic user interactions. These AI-driven tools analyze complex user queries, source relevant information, and articulate responses that are both informative and comforting.

Integration With ChatGPT Technology

The core functionalities of Lola—the chatbot and voice assistant—use the natural language processing prowess of ChatGPT 3.5 Turbo.

Voice Assistant and Chatbot Mechanism

These multimodal features enhance the accessibility of the app, providing users with options for both auditory and textual communication. This is particularly beneficial for users with different levels of technological literacy or physical abilities. The real-time response generation capability of ChatGPT ensures that every interaction is smooth and engaging.

Data Utilization and Information Sourcing

Lola's AI systems are designed to draw information from authoritative sources such as the Alzheimer's Association online repository [1], supplemented by curated recommendations from a team of medical doctors specializing in ADRD. This method ensures that the information provided is not only accurate but also up to date with current medical standards and practices.

Ensuring Clear and Meaningful Communication

The generative process involves a sophisticated workflow in which initial responses, structured using LLM techniques, are refined by ChatGPT for clarity and empathy. This ensures that the communication is not only technically accurate but also tailored to meet the emotional and cognitive needs of the users.

Technical Infrastructure

The comprehensive technical infrastructure of Lola uses Flutter for development of the front end to maintain a consistent and engaging user experience with visually appealing, high-performance interactions across various devices and modalities. The back end is powered by Python's Flask framework, paired with a PostgreSQL database for robust data management and fast data retrieval during the human-app interactions. The deployment of Docker for containerization, along with the use of the research team's university servers, provides a secure and reliable environment for hosting and managing Lola's functionalities. By reducing compatibility issues, this setup ensures consistent system performance and reliability.

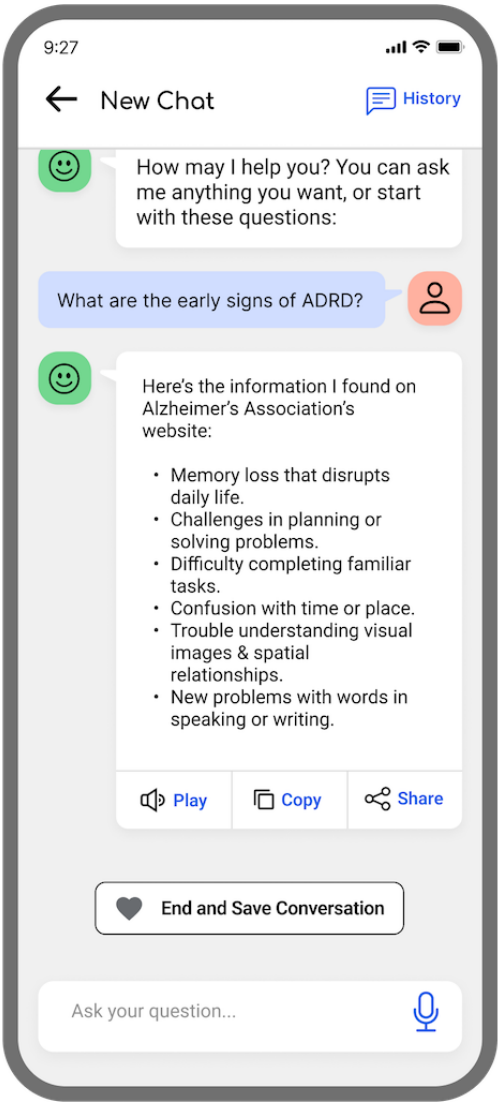
App Features

In response to the challenges faced by older Black Americans regarding the accessibility of credible ADRD information, Lola's interface serves as a platform for accessing ADRD-related articles, local and virtual events, and opportunities to engage in the latest ADRD research. Lola incorporates a chatbot feature represented by a floating green bubble in the bottom right corner, accessible to users throughout their interaction with the app (Figure 1).

Users can interact with the Lola chatbot by clicking the green call-to-action button, which initiates a conversation through text or voice input. Using generative AI, the Lola chatbot provides responses in text format or as voice messages by clicking the "Play" button, facilitating accessibility for users with varying literacy levels or preferences for voice interaction. Users can also copy the response as text or share the response with their contacts. When users choose to end the conversation, the system saves the conversation in the chat history (Figure 2).

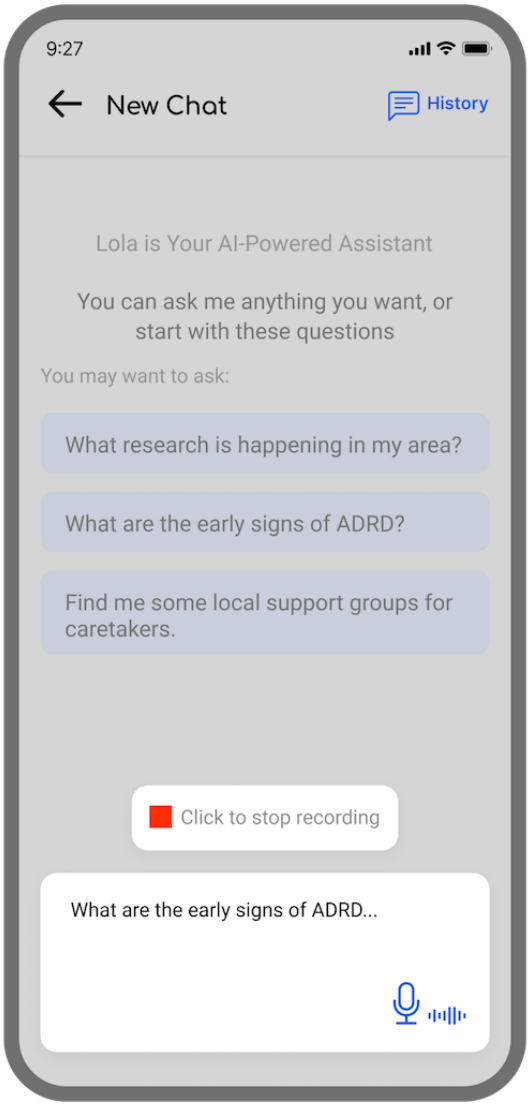
XSL•FO
RenderX

Figure 1. Example of a user having a conversation with the Lola chatbot using the interface designed by the research team.



XSL•FO
RenderX

Figure 2. Example of a user using voice input and the possibility of browsing the history.



Results

Participant Demographic Information

A total of 15 participants were included in this study. All of them were informal caregivers for someone living with ADRD and were Black American. All participants shared that they were born and raised in the United States. Regarding family origins, 5 participants identified having African family origins, 9 were unsure of their family origins, and 1 chose not to disclose. Table 1 summarizes the participants' demographic data. Of the 15 participants, 12 (80%) were 65 years or older. Most participants were female (11/15, 73%), with 2 (2/15, 13%) being male and 2 (2/15, 13%) providing no gender response. Regarding education, 10 participants (10/15, 67%) indicated having at least some college education (partial completion, a bachelor's degree, or an associate degree), while 2 (2/15, 13%) had a graduate degree, 2 (2/15, 13%) indicated a high school education, and 1 (1/15, 7%) did not provide a response. Most participants (11/15, 73%) were retired, only 3 participants (3/15, 20%) worked either part-time or full-time, and 1 (1/15, 7%) did not respond.

XSL•FO
RenderX

Table 1. Participant demographic information.

Participant number	Age (years)	Gender	Education level	Employment	Work sector
1	72	Female	Some college	Retired	Health care
2	74	Male	Some college	Retired	Utilities
3	68	Female	Associate's degree	Retired	NR[a]
4	69	Male	Some college	Retired	Construction
5	54	Female	Associate's degree	Part-time	Health care
6	66	NR	NR	NR	NR
7	69	Female	Graduate degree	Retired	Community advocacy
8	69	Female	Some college	Retired	Health care
9	75	Female	Some college	Retired	Faith-based
10	71	Female	High school	Retired	Customer service
11	69	Female	Associate's degree	Retired	Insurance
12	72	NR	High school	Retired	Health care
13	68	Female	Some college	Retired	Health care
14	63	Female	Graduate degree	Full-time	Employment specialist
15	62	Female	Bachelor's degree	Part-time	Education

[a]NR: not reported.

Overview of Lola

The Lola system includes 3 different models of interaction: (1) accessing information in a traditional touch screen app format, (2) texting with an AI-powered text-based chatbot, and (3) talking to a voice assistant that accompanies the chatbot. Participants experienced all 3 models of interaction during user testing.

Themes

From the analysis of the data collected, 3 main themes emerged: (1) the need for a multimodal AI tool that relies on multiple senses, (2) the need for culturally specific ADRD content, and (3) the importance of physical places and existing social networks to spread Lola among the community. This section will explore these themes in depth and how they relate to the goals of the research.

Theme 1: The Need for a Multimodal AI Tool That Relies on Different Senses

This theme describes participants' need for interacting with the Lola system in different ways aside from the traditional touch screen gesturing predominant in most mobile apps. This multimodal interaction could take several forms, for example, voice interaction that can read content generated as text out loud. This would allow participants with low vision, which is common among older adults, to access the content.

> *Because with the people that I work with on a daily basis, one is not able to read anymore because she's going blind. So, it's not just Alzheimer's dementia that needs to be looked at, at least all the senses need to be looked at because people are going through so many different things at the same time.* [P4]

Consistently, auditory access to information positively influenced participants' interactions with the system. Some participants preferred to engage with the tool using their voice, hence treating the tool as a voice assistant and asking it questions. For the participants, this was a valuable addition to the traditional touch screen interaction. For people not only with low vision but also with lower digital skills, voice interaction resembles a conversational exchange with which they are familiar and comfortable.

> *I think that this was good to be able to speak it. Because, if you're dealing with people who are not computer savvy or internet savvy, being able to speak it, and it automatically pops up or speaks your question and then being able to close that area so it won't keep recording. That was quick and easy. It was simple.* [P9]

Speaking features (as opposed to just reading and typing information) are especially helpful since they would enable people to access their usual methods of dealing with health-related information, such as writing it down on a piece of paper and saving it for later reference while someone is sharing the information. Having an auditory feature to provide health care information in addition to other interaction models can improve the accessibility of relevant information for those with difficulties accessing digital content due to low visual capacity.

> *That would help too if it played. And then they spoke out the information. And there's different ways I can have a pen and a pad. And then while she's speaking, I could write the information down. Because it is printed small. And you have to read all this information. But if she tells me, and the number you can call is one 800, I'd be writing it down.* [P1]

XSL•FO
RenderX

However, solely relying on the sense of hearing could be a limitation for people with auditory impairments, who could get frustrated by the tool.

> *But I couldn't hear it. I was kind of frustrated. So, it kind of forced me to read it in my eyes are strained.* [P14]

One participant highlighted the benefit of a multimodal approach pairing visual and auditory stimuli together, drawing on their experience in church where reading and listening occur simultaneously.

> *For me and people lazy, they don't want to have to go through the reading it would have, because I would like to hear it while I'm reading it. It is better. If it were just speaking it, I probably wouldn't grasp. But it was speaking in I was reading it and I was falling through. It would have been like being in church. And the preacher is preaching a sermon, he tell you what he coming out of and you fall you open your Bible to that. John 316, and you follow him as he's talking.* [P5]

This theme highlights the importance of creating a system with multimodal interaction and pairing auditory and visual inputs in a way that makes the tool more accessible and allows a broader audience to meaningfully engage.

Theme 2: The Need for Context and Culture-Specific ADRD Content

This theme highlighted the need for contextually and culturally specific mobile apps that integrate relevant and accessible ADRD information. When combined with a multimodal AI tool, these tailored apps can significantly enhance participants' access to the information they need, which is not always directly related to health care.

One participant expressed their discontent with the lack of relevant information in the Lola app, stating that they wanted more than just medical information such as actionable lifestyle changes that can improve their health.

> *Definitely more than just medical information. Because I'm not one to push drugs, I don't take none. If a doctor tells me something is going on, I'm asking first and foremost, what can I do? Or how can you help me get to this process? Naturally? Do I need to implement exercise, do I need to change my diet, what vitamins, whatever. But if they talk about medication, it's gonna be a big turn off for me instantly. So that's just me. Some people think a quick fix is the answer. But every Quick Fix has consequences.* [P4]

Additionally, another participant mentioned the need for accessible ADRD mobile apps. Being able to operate and navigate through the app with little or without outside help is ideal.

> *Well, you know I just had therapy at the hospital, okay. And I was so amazed that every time I went to therapy, they gave me a folder with exercises, and one time they gave me an app. A girl was there, and she showed that to me. And I was just shocked. She used the app to show me the exercises, and every time I went back, she added exercises for me to do on the app. So that was amazing to me. So yeah, the apps are important. They really are. And they're to be simple. And they got to be informative. And they got to be to the point where even the agent can navigate through them.* [P10]

Some participants expressed discontent with the use of language in the Lola app, as it was perceived as too technical and not catering to the target population well enough.

> *Now all these big words, that's what we get me. I wouldn't understand them. Okay, so there's too many words, too many big words, I would not understand.* [P15]

Certain participants expressed a desire for the mobile app to cater to their Black American communities by offering culturally specific ADRD information tailored to their needs. Consequently, they sought additional content that would address their unique concerns and circumstances.

> *So I'm thinking on the other side as the creators of it, maybe you don't want to offend other races or whatever. But if Alzheimer's is more prevalent in my community, I feel like we need to have that target. So we will know exactly where to go, filter through some of the other stuff, some of the things but Alzheimer's and dementia are across the board. But maybe some things are just specific toward my race. Sure. And I want to know that, and I want to know that in the most timely fashion possible.* [P9]

Last, some of the participants highlighted the necessity for the mobile app to include community-centric information and information about local resources and organizations for ADRD, which could be tailored to individual users and their community based on their location.

> *I don't travel too far for it. So, I think there needs to be something located in their community.* [P4]
>
> *More websites because a lot of people are interested in a study of Alzheimer's or dementia. And it's national, local. So, I would rather you know, her give me the local if I had to visit it. It's good to know the national, but also there's local organizations as well.* [P1]

This section showed how providing context and culturally relevant ADRD information that is accessible is requested by the Black American community.

Theme 3: Physical Places and Existing Social Networks Are Necessary to Spread Lola Among the Community

This last theme is related to the importance of using physical community-oriented places and pre-existing social networks to spread technological tools among the Black American community. Our participants suggested community-oriented ways to advertise these health-related technologies to be adopted and used by the community. Most of the participants mentioned the importance of doing demonstrations of the Lola app in physical community places such as a town hall.

XSL•FO
RenderX

> *I think if they're going to implement this process, then they should have maybe a town hall meeting to go through the process have a big screen. The people that are interested in finding out more than they'll do it step by step. So I think that will be a demonstration we're having at a community center or, or even in a building like this, and advertise that we have a new application now called Lola that will help.* [P1]

Finally, participants mentioned how, in the advertisement of the Lola app, it is important to leverage pre-existing social networks within the community, such as relying on word of mouth and showing how people within the community are using Lola.

> *More information to let us know, what makes Lola different from the other? That will make me want to go to Lola more. And that could cover word of mouth, that could cover the internet, having more of these to get people involved to know more, maybe part of the marketing or the draw from my community would be Lola specifically, or is more strategic as far as getting information out to Black communities.* [P9]

This section highlighted how physical locations where the community gathers and existing social networks are vital to promote increased use of Lola, as word-of-mouth marketing is effective when performed by trustworthy individuals.

Discussion

Principal Findings

Deployment of generative AI in the health care domain offers the promise of innovation and expands the potential of equitable health technologies. However, erroneous design and development of generative AI could increase bias and potential harm toward historically marginalized and underserved communities. Including underserved populations in the design and development of a technological intervention has been shown to increase the opportunity for representation and inclusivity [27-31]. Following this line of research, our study investigated the experiences of Black American caregivers using a generative AI tool to seek ADRD-related information (such as symptoms, community resources, financial aid).

This is the first study that specifically addresses the needs, demands, and challenges of this specific population from the perspective of caregivers of individuals with ADRD. An in-depth description of the 3 key design implications to create an inclusive and culturally relevant AI-supported platform for health will follow.

Fostering Inclusivity Through Multimodality

We observed how Black American caregivers expressed the need for multimodality of interaction to equitably access what generative AI tools have to offer. Indeed, they expressed the importance of being able to "talk with the tool" and hear back (eg, read content out loud) in the fashion of a conversational voice assistant. Several participants' insights suggested a preference toward multimedia, emphasizing the auditory and visual senses and providing modes (speaking, reading, and listening) to support these interaction styles. These implications may originate from their day-to-day time constraints, requiring short and efficient interactions.

Previous work addressed the efficiency of using multimodal systems that provide health care. Linders et al [59] created a multimodal health assistant to provide health information. Soubutts et al [60] used a multimodal device to support the needs of households for health and care purposes during COVID-19. The need for multimodal interactions for memory aid systems has been also highlighted [61].

Additionally, multimodality can guarantee a larger audience who will interact with the system [62], thereby allowing people with low vision and low hearing to access the system and seek ADRD health-related information. Within the landscape of designing for inclusive and equitable computing, our study took the approach of multimodal interaction. Our findings emphasize the need to tailor technology to accommodate different interaction styles. This can be achieved by designing multimodal technologies that can meet the needs of differently abled users.

Answering the Call for Context and Culture-Specific ADRD Information

Our findings reflected the caregivers' need for contextually and culturally relevant ADRD information that is specific to their communities or personalized based on their needs and preferences. They indicated their need for lifestyle tips and actionable tools for ADRD rather than only medical information in the system. They also did not want to be offered only general ADRD medical information; rather, they wanted ADRD-related information specific to their community, local resources, and geography. Locality and cultural relevance emerged as 2 of the most important factors influencing the perceived usefulness of the information sought. This is in line with previous studies [63,64] showing that providing community-centered information can increase the acceptance and adoption of generative AI tools. Compared with the use of OHC, which provides general emotional and social support, as identified by previous studies [49-51], our study reveals the need for specialized information on health care. Our participants expressed their need for instrumental and informational support specific to their social and cultural context as well as geographic location. This finding complements the benefits of OHCs to meet all the needs of caregivers including instrumental and information support.

Furthermore, our participants wanted to be sure that the information provided was accessible to people with low literacy skills within their own communities. This is in line with the collectivist nature of African American communities observed in a prior study on technology use [30], wherein, for a technological intervention to be accepted, it must be inclusive of all the members of the Black American community.

Keeping Not Only the People But Also the Community in the Loop

Finally, our study offers a novel contribution in generative AI research investigating how generative AI tools can be advertised and spread among the Black American community. To our knowledge, this is the only study that has investigated this aspect of generative AI use and adoption. Indeed, low technology adoption is a striking problem within Black American

XSL•FO
RenderX

communities, especially among people with low economic status, which prevents them from using online tools to access health information and improve their health literacy [65]. However, low adoption does not arise from the lack of digital tools but also from a skepticism toward digital tools [66]. Privacy is also a crucial concern, as Black American populations tend to be more skeptical toward the use of technology based on a fear of privacy intrusions and breaches of privacy [41,63]; thus, generative AI–based technologies for this population should foster a transparent and reliable way of handling user privacy.

Some studies [35,66-68] have investigated how faith-based organizations can be important for health dissemination and assisting Black American adults seeking health information. Our study expands on these previous studies by exploring the strategies suggested by our participants, which could be used to foster generative AI adoption and its use among their communities.

Hence, although most studies have focused on programs designed to disseminate information based upon the creation of seminars and physical events, our study provides evidence of how physical organizations (faith-based organizations, public as well as medical organizations) could be strategic actors in fostering the adoption and use of a generative AI tool for ADRD. Additionally, given that these populations have limited resources and opportunities to interact with AI systems and often have difficulties accessing any technology (eg, software, hardware, internet) due to financial and infrastructural challenges [39], incorporating these physical organizations is significant. Building trust and familiarity by involving faith-based organizations and communities in the dissemination of the tool, as well as in the tool's use, facilitates not only dissemination but also the ability to break the barrier of skepticism toward health technology, by showing the community how it could embrace the tool.

Finally, ethical implications and concerns must be addressed in the discussion of the findings, even though ethical concerns did not explicitly emerge in this study as its focus was on usability, accessibility, cultural relevance, and adoption. However, it is crucial to discuss how topics such as consent, privacy, and bias can impact the use of generative AI for an underrepresented community. As suggested by previous studies, ethical guidelines committed to social good and human well-being should be developed and followed by AI system developers [32]. Consent and data management must be at the core of every generative AI system, allowing users to be in full control of their own data and their interaction with the tool. Furthermore, it is critical that users who interact with or provide data for AI systems fully understand what they are consenting to. This is especially important for populations with lower literacy rates, for whom the clarity and accessibility of consent forms must be carefully considered.

Limitations

This study is subject to several limitations. First, the sample characteristics are not representative of the broader population. The age range of our participants may not adequately capture the unique challenges faced by younger caregivers who juggle multiple responsibilities alongside caregiving. Future studies should aim to recruit a more diverse sample to better understand the challenges different demographics face in accessing health information for caregiving.

Furthermore, our participants demonstrated a high level of education, with the majority (12/15, 80%) having some college education. This could have influenced our findings, potentially skewing the results toward those with higher educational backgrounds. We recommend future research include Black American individuals with lower literacy levels to better understand their specific needs and how they engage with AI-generated health technologies.

Last, the user experience testing sessions for this study were conducted outside of the participants' daily environments, which may not accurately reflect the day-to-day challenges they might encounter when interacting with Lola. This limitation could restrict the scope of identified problems. Future studies should consider conducting tests in settings and scenarios that more closely mimic real-life conditions to gain deeper insights into the tool's practical efficacy and user engagement.

Apart from the opportunities already identified, our study opens the research space on the intersection between generative AI and underrepresented populations living in the United States. One possibility could be investigating the effectiveness of tailoring ADRD content to the community's values; hence, future studies could quantify the effectiveness of such interventions on user engagement as well as on ADRD-related health literacy. Second, future studies could investigate the impact of different interaction modalities on user engagement, thus understanding whether multimodality is preferred over solely having the opportunity of interacting with one modality, and explore the potential benefit of using different modalities over others in groups with various needs.

Conclusions

This study adds to the expanding field of research focused on deploying generative AI to improve health literacy. More specifically, the emphasis of this study was on improving the ADRD-related health literacy of Black American informal caregivers, who often are solely responsible for individuals with ADRD. Our objective in testing our intervention, Lola, was to identify the specific aspects of our design that need to be tailored and modified. This approach aims to develop a tool that can genuinely assist this population, enhance their ADRD-related health literacy, and support them in their caregiving responsibilities.

Our findings showed that ADRD-specific content should be made contextually and culturally relevant to be perceived as useful by informal caregivers, thereby including information on local, geographically dependent resources. This shows how generative AI designed to improve ADRD-related health literacy should be highly personalized and dependent upon the user's social and cultural context. Additionally, concerning the interactions, relying on multimodality can contribute to creating an accessible tool for informal caregivers, by offering them multiple and diverse interaction strategies. Once again, this

XSL•FO
RenderX

finding indicates the need for a generative AI tool that can adapt to the users' physical and cognitive needs and demands.

Finally, through active collaboration and community engagement, this study provides evidence of the importance of including the users and population of interest in the evaluation of a design as a way to guarantee that generative AI technologies will benefit all individuals within society, even underrepresented communities such as Black American populations.

Acknowledgments

Funding was provided by the National Health Institute (NIA 1R24AG071471) and the National Science Foundation (NSF number 2145049).

Authors' Contributions

CB and EO wrote the majority of the manuscript and led the writing process and some of the detailed cognitive walk-through sessions. EO, JOT, VN, DC, and XP wrote and edited portions of the manuscript and led some of the detailed cognitive walk-through sessions. BC, AKH, and NMJ led participant recruitment and community engagement. PAB, YL, CVH, HCH, and PCS initiated and ideated the project.

Conflicts of Interest

None declared.

References

1. Alzheimer's Disease Facts and Figures. Alzheimer's Association. URL: https://www.alz.org/alzheimers-dementia/facts-figures [accessed 2024-12-21]
2. Shojaei F. Uncovering the Impact of Co-Design and Design Thinking in Enhancing the Empowerment of People with Dementia. 2024 Presented at: International Conference on Innovation of Emerging Communication and Information Technology; October 20-23, 2024; Casablanca, Morocco.
3. Barnes LL. Alzheimer disease in African American individuals: increased incidence or not enough data? Nat Rev Neurol 2022 Jan;18(1):56-62 [FREE Full text] [doi: 10.1038/s41582-021-00589-3] [Medline: 34873310]
4. Barnes LL, Bennett DA. Alzheimer's disease in African Americans: risk factors and challenges for the future. Health Aff (Millwood) 2014 Apr;33(4):580-586 [FREE Full text] [doi: 10.1377/hlthaff.2013.1353] [Medline: 24711318]
5. Brewster P, Barnes L, Haan M, Johnson JK, Manly JJ, Nápoles AM, et al. Progress and future challenges in aging and diversity research in the United States. Alzheimers Dement 2019 Jul 18;15(7):995-1003 [FREE Full text] [doi: 10.1016/j.jalz.2018.07.221] [Medline: 30240574]
6. Mayeda ER, Glymour M, Quesenberry CP, Whitmer RA. Inequalities in dementia incidence between six racial and ethnic groups over 14 years. Alzheimers Dement 2016 Mar 10;12(3):216-224 [FREE Full text] [doi: 10.1016/j.jalz.2015.12.007] [Medline: 26874595]
7. Muvuka B, Combs RM, Ayangeakaa SD, Ali NM, Wendel ML, Jackson T. Health literacy in African-American communities: barriers and strategies. Health Lit Res Pract 2020 Jul 16;4(3):e138-e143. [doi: 10.3928/24748307-20200617-01] [Medline: 32674161]
8. Xie B, Champion JD, Kwak J, Fleischmann KR. Mobile health, information preferences, and surrogate decision-making preferences of family caregivers of people with dementia in rural Hispanic communities: cross-sectional questionnaire study. J Med Internet Res 2018 Dec 10;20(12):e11682 [FREE Full text] [doi: 10.2196/11682] [Medline: 30530450]
9. Ayre J, Mac O, McCaffery K, McKay BR, Liu M, Shi Y, et al. New frontiers in health literacy: using ChatGPT to simplify health information for people in the community. J Gen Intern Med 2024 Mar;39(4):573-577 [FREE Full text] [doi: 10.1007/s11606-023-08469-w] [Medline: 37940756]
10. Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, et al. How does ChatGPT perform on the United States Medical Licensing Examination (USMLE)? The implications of large language models for medical education and knowledge assessment. JMIR Med Educ 2023 Feb 08;9:e45312 [FREE Full text] [doi: 10.2196/45312] [Medline: 36753318]
11. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. PLOS Digit Health 2023 Feb;2(2):e0000198 [FREE Full text] [doi: 10.1371/journal.pdig.0000198] [Medline: 36812645]
12. Samaan JS, Yeo YH, Rajeev N, Hawley L, Abel S, Ng WH, et al. Assessing the accuracy of responses by the language model ChatGPT to questions regarding bariatric surgery. OBES SURG 2023 Apr 27;33(6):1790-1796. [doi: 10.1007/s11695-023-06603-5]
13. Walker HL, Ghani S, Kuemmerli C, Nebiker CA, Müller BP, Raptis DA, et al. Reliability of medical information provided by ChatGPT: assessment against clinical guidelines and patient information quality instrument. J Med Internet Res 2023 Jun 30;25:e47479 [FREE Full text] [doi: 10.2196/47479] [Medline: 37389908]

XSL•FO
RenderX

14. Bahrini A, Khamoshifar M, Abbasimehr H, Riggs RJ, Esmaeili M, Majdabadkohne RM. ChatGPT: Applications, Opportunities, and Threats. 2023 Presented at: Systems and Information Engineering Design Symposium (SIEDS); April 27-28, 2023; Charlottesville, VA. [doi: 10.1109/sieds58326.2023.10137850]
15. Biswas SS. Role of Chat GPT in public health. Ann Biomed Eng 2023 May 15;51(5):868-869. [doi: 10.1007/s10439-023-03172-7] [Medline: 36920578]
16. Baldassarre MT, Caivano D, Fernandez Nieto B, Gigante D, Ragone A. The Social Impact of Generative AI: An Analysis on ChatGPT. 2023 Presented at: GoodIT '23: ACM International Conference on Information Technology for Social Good; September 6-8, 2023; Lisbon, Portugal. [doi: 10.1145/3582515.3609555]
17. McDuff D, Song Y, Kapoor A, Ma S. Characterizing bias in classifiers using generative models. 2019 Presented at: NIPS'19: 33rd International Conference on Neural Information Processing Systems; December 8-14, 2019; Vancouver, British Columbia, Canada.
18. Spector-Bagdady K. Generative-AI-generated challenges for health data research. Am J Bioeth 2023 Oct 09;23(10):1-5. [doi: 10.1080/15265161.2023.2252311] [Medline: 37831940]
19. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol 2017 Jun 21;2(4):230-243. [doi: 10.1136/svn-2017-000101]
20. Zhang P, Kamel Boulos MN. Generative AI in medicine and healthcare: promises, opportunities and challenges. Future Internet 2023 Aug 24;15(9):286. [doi: 10.3390/fi15090286]
21. Bautista YJP, Theran C, Aló R, Lima V. Health Disparities Through Generative AI Models: A Comparison Study Using a Domain Specific Large Language Model. In: Lecture Notes in Networks and Systems. Cham, Switzerland: Springer Nature; 2023:220-232.
22. Hoffman KM, Trawalter S, Axt JR, Oliver MN. Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites. Proc Natl Acad Sci U S A 2016 Apr 19;113(16):4296-4301 [FREE Full text] [doi: 10.1073/pnas.1516047113] [Medline: 27044069]
23. Panch T, Mattie H, Atun R. Artificial intelligence and algorithmic bias: implications for health systems. J Glob Health 2019 Dec;9(2):010318. [doi: 10.7189/jogh.09.020318] [Medline: 31788229]
24. Petti U, Nyrup R, Skopek JM, Korhonen A. Ethical considerations in the early detection of Alzheimer's disease using speech and AI. 2023 Presented at: FAccT '23: the 2023 ACM Conference on Fairness, Accountability, and Transparency; June 12-15, 2023; Chicago, IL. [doi: 10.1145/3593013.3594063]
25. Tamkin A, Brundage M, Clark J, Ganguli D. Understanding the Capabilities, Limitations, and Societal Impact of Large Language Models. The Stanford Natural Language Processing Group. URL: https://www-nlp.stanford.edu/pubs/tamkin2021understanding.pdf [accessed 2024-12-21]
26. Mirkin S, Albensi BC. Should artificial intelligence be used in conjunction with neuroimaging in the diagnosis of Alzheimer's disease? Front Aging Neurosci 2023 Apr 18;15:1094233 [FREE Full text] [doi: 10.3389/fnagi.2023.1094233] [Medline: 37187577]
27. Dillahunt TR. Fostering social capital in economically distressed communities. 2014 Presented at: CHI '14: CHI Conference on Human Factors in Computing Systems; April 26-May 1, 2014; Toronto, Ontario, Canada. [doi: 10.1145/2556288.2557123]
28. Dillahunt TR, Kameswaran V, Li L, Rosenblat T. Uncovering the Values and Constraints of Real-time Ridesharing for Low-resource Populations. 2017 Presented at: CHI '17: CHI Conference on Human Factors in Computing Systems; May 6-11, 2017; Denver, CO. [doi: 10.1145/3025453.3025470]
29. Harrington CN, Piper AM. Informing Design through Sociocultural Values: Co-Creation with Low-Income African-American Older Adults. 2018 Presented at: PervasiveHealth '18: 12th EAI International Conference on Pervasive Computing Technologies for Healthcare; May 21-24, 2018; New York, NY. [doi: 10.1145/3240925.3240966]
30. Parker AG, Grinte RE. Collectivistic health promotion tools: Accounting for the relationship between culture, food and nutrition. International Journal of Human-Computer Studies 2014 Feb;72(2):185-206. [doi: 10.1016/j.ijhcs.2013.08.008]
31. Richards OK, Marcu G, Brewer RN. Hugs, Bible Study, and Speakeasies: Designing for Older Adults' Multimodal Connectedness. 2021 Presented at: DIS '21: Designing Interactive Systems Conference 2021; June 28-July 2, 2021; Virtual meeting. [doi: 10.1145/3461778.3462075]
32. Chen Y, Esmaeilzadeh P. Generative AI in medical practice: in-depth exploration of privacy and security challenges. J Med Internet Res 2024 Mar 08;26:e53008 [FREE Full text] [doi: 10.2196/53008] [Medline: 38457208]
33. Mesko B. The ChatGPT (generative artificial intelligence) revolution has made artificial intelligence approachable for medical professionals. J Med Internet Res 2023 Jun 22;25:e48392 [FREE Full text] [doi: 10.2196/48392] [Medline: 37347508]
34. Rodriguez DV, Lawrence K, Gonzalez J, Brandfield-Harvey B, Xu L, Tasneem S, et al. Leveraging generative AI tools to support the development of digital solutions in health care research: case study. JMIR Hum Factors 2024 Mar 06;11:e52885 [FREE Full text] [doi: 10.2196/52885] [Medline: 38446539]
35. Knott CL, Miech EJ, Slade J, Woodard N, Robinson-Shaneman B, Huq M. Evaluation of organizational capacity in the implementation of a church-based cancer education program. Glob Implement Res Appl 2022 Jan 07;2(1):22-33. [doi: 10.1007/s43477-021-00033-0]

XSL•FO
RenderX

36. Etori N, Temesgen E, Gini M. What we know so far: artificial intelligence in African healthcare. arXiv. Preprint posted online on May 10, 2023 .
37. Ganju A, Satyan S, Tanna V, Menezes SR. AI for improving children's health: a community case study. Front Artif Intell 2020;3:544972 [FREE Full text] [doi: 10.3389/frai.2020.544972] [Medline: 33733204]
38. Guo J, Li B. The application of medical artificial intelligence technology in rural areas of developing countries. Health Equity 2018 Aug;2(1):174-181 [FREE Full text] [doi: 10.1089/heq.2018.0037] [Medline: 30283865]
39. Hoodbhoy Z, Hasan B, Siddiqui K. Does artificial intelligence have any role in healthcare in low resource settings? J Med Artif Intell 2019 Jul;2:13-13. [doi: 10.21037/jmai.2019.06.01]
40. Senteio C, Murdock PJ. The Efficacy of Health Information Technology in supporting health equity for Black and Hispanic patients with chronic diseases: systematic review. J Med Internet Res 2022 Apr 04;24(4):e22124 [FREE Full text] [doi: 10.2196/22124] [Medline: 35377331]
41. Harris KM, Mayo Gamble T, Yoo MG, Spell LA, Minor TN, Jones H, et al. Leveraging mHealth to mitigate the impact of COVID-19 in Black American communities: qualitative analysis. JMIR Hum Factors 2023 Dec 22;10:e47294 [FREE Full text] [doi: 10.2196/47294] [Medline: 37874735]
42. Brodaty H, Donkin M. Family caregivers of people with dementia. Dialogues in Clinical Neuroscience 2022 Apr 01;11(2):217-228. [doi: 10.31887/dcns.2009.11.2/hbrodaty]
43. Wozney L, Freitas de Souza LM, Kervin E, Queluz F, McGrath PJ, Keefe J. Commercially available mobile apps for caregivers of people with Alzheimer disease or other related dementias: systematic search. JMIR Aging 2018 Dec 07;1(2):e12274 [FREE Full text] [doi: 10.2196/12274] [Medline: 31518255]
44. Bangerter LR, Griffin J, Harden K, Rutten LJ. Health information-seeking behaviors of family caregivers: analysis of the health information national trends survey. JMIR Aging 2019 Jan 14;2(1):e11237 [FREE Full text] [doi: 10.2196/11237] [Medline: 31518309]
45. Hong Y, Shen K, Han H, Ta Park V, Bagchi P, Lu H, et al. A WeChat-based intervention, Wellness Enhancement for Caregivers (WECARE), for Chinese American dementia caregivers: pilot assessment of feasibility, acceptability, and preliminary efficacy. JMIR Aging 2023 Apr 5;6:e42972. [doi: 10.2196/42972]
46. Boutilier JJ, Loganathar P, Linden A, Scheer E, Noejovich S, Elliott C, et al. A web-based platform (CareVirtue) to support caregivers of people living with Alzheimer disease and related dementias: mixed methods feasibility study. JMIR Aging 2022 Aug 04;5(3):e36975 [FREE Full text] [doi: 10.2196/36975] [Medline: 35925654]
47. Quinn C, Staub S, Barr E, Gruber-Baldini A. Mobile support for older adults and their caregivers: dyad usability study. JMIR Aging 2019 May 23;2(1):e12276. [doi: 10.2196/12276]
48. Guo Y, Yang F, Hu F, Li W, Ruggiano N, Lee HY. Existing mobile phone apps for self-care management of people with Alzheimer disease and related dementias: systematic analysis. JMIR Aging 2020 Jan 24;3(1):e15290 [FREE Full text] [doi: 10.2196/15290] [Medline: 32012045]
49. Johnson J, Arnold V, Piper AM, Hayes GR. "It's a lonely disease": Cultivating Online Spaces for Social Support among People Living with Dementia and Dementia Caregivers. 2022 Presented at: ACM Conference on Human Factors in Computing Systems; April 30-May 5, 2022; New Orleans, LA. [doi: 10.1145/3555133]
50. Smith CE, Levonian Z, Ma H, Giaquinto R, Lein-Mcdonough G, Li Z, et al. "I Cannot Do All of This Alone". ACM Trans. Comput.-Hum. Interact 2020 Aug 17;27(5):1-41. [doi: 10.1145/3402855]
51. Al-Megren S, Majrashi K, Allwihan RM. ”The Dementia Diva Strikes Again!”: A Thematic Analysis of How Informal Carers of Persons with Dementia Use TikTok. 2021 Presented at: MobileHCI '21: 23rd International Conference on Mobile Human-Computer Interaction; September 27 - October 1, 2021; Toulouse, France. [doi: 10.1145/3447527.3474857]
52. Yuan F, Bowland S, Proctor L, Blackburn J, Mukherjee N, Bray R, et al. Robot-assisted Psycho-education to Enhance Alzheimer’s Caregiver Health. 2022 Presented at: IEEE/ACM Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE); November 17-19, 2022; Arlington, VA. [doi: 10.1109/chase52844.2021.00027]
53. Lin XY, Saksono H, Stowell E, Lachman ME, Castaneda-Sceppa C, Parker AG. Go and Grow: an evaluation of a pervasive social exergame for caregivers of loved ones with dementia. Proc ACM Hum Comput Interact 2020 Oct;4(CSCW 2):1 [FREE Full text] [doi: 10.1145/3415222] [Medline: 33150303]
54. Granollers T, Lorés J. Incorporation of users in the Evaluation of Usability by Cognitive Walkthrough. In: Navarro-Prieto R, Vidal J, editors. HCI related papers of Interacción 2004. Dordrecht, The Netherlands: Springer; 2006:243-255.
55. Mahatody T, Sagar M, Kolski C. State of the art on the cognitive walkthrough method, its variants and evolutions. International Journal of Human-Computer Interaction 2010 Jul 30;26(8):741-785. [doi: 10.1080/10447311003781409]
56. Lewis C. Using the "Thinking Aloud" Method in Cognitive Interface Design. Yorktown Heights, NY: IBM Thomas J. Watson Research Division; 1982.
57. Rampin R, Rampin V. Taguette: open-source qualitative data analysis. JOSS 2021 Dec;6(68):3522. [doi: 10.21105/joss.03522]
58. Braun V, Clarke V. What can "thematic analysis" offer health and wellbeing researchers? Int J Qual Stud Health Well-being 2014;9:26152 [FREE Full text] [doi: 10.3402/qhw.v9.26152] [Medline: 25326092]
59. Linders GM, Vaitonytė J, Alimardani M, Mitev KO, Louwerse MM. A realistic, multimodal virtual agent for the healthcare domain. 2022 Presented at: 22nd ACM International Conference on Intelligent Virtual Agents; September 6-9, 2022; Faro, Portugal. [doi: 10.1145/3514197.3551250]

XSL•FO
RenderX

60. Soubutts E, Ayobi A, Eardley R, McNaney R, Cater K, O'Kane AA. Amazon Echo Show as a multimodal human-to-human care support tool within self-isolating older UK households. Proc. ACM Hum.-Comput. Interact 2022 Nov 11;6(CSCW2):1-31. [doi: 10.1145/3555193]
61. Otenen E. Towards Designing for Multimodal Remembering: Findings from an Interview Study. 2024 Presented at: CHI '24: CHI Conference on Human Factors in Computing Systems; May 11-16, 2024; Honolulu, HI URL: https://doi.org/10.1145/3613905.3647965 [doi: 10.1145/3613905.3647965]
62. Kuriakose B, Shrestha R, Sandnes FE. Multimodal navigation systems for users with visual impairments—a review and analysis. MTI 2020 Oct 16;4(4):73. [doi: 10.3390/mti4040073]
63. Hindricks KV, Meyer JJC. Artificial intelligence in health care and medicine: a personalized approach. Acta Scientific Medical Sciences 2019;3(1):71-78. [doi: 10.31080/ASMS.2019.03.0410]
64. Kozyreva A, Lorenz-Spreen P, Hertwig R, Lewandowsky S, Herzog SM. Public attitudes towards algorithmic personalization and use of personal data online: evidence from Germany, Great Britain, and the United States. Humanit Soc Sci Commun 2021 May 14;8(1):1. [doi: 10.1057/s41599-021-00787-w]
65. Seo H, Erba J, Geana M, Lumpkins C. Calling Doctor Google? Technology adoption and health information seeking among low-income African-American older adults. JPIC 2017 Dec 22;1(2):153. [doi: 10.32473/jpic.v1.i2.p153]
66. Santos SLZ, Tagai EK, Scheirer MA, Bowie J, Haider M, Slade J, et al. Adoption, reach, and implementation of a cancer education intervention in African American churches. Implement Sci 2017 Mar 14;12(1):36 [FREE Full text] [doi: 10.1186/s13012-017-0566-z] [Medline: 28292299]
67. Davis CM. Health beliefs and breast cancer screening practices among African American women in California. Int Q Community Health Educ 2020 Jul 13;41(3):259-266. [doi: 10.1177/0272684x20942084]
68. Tagai EK, Scheirer MA, Santos SLZ, Haider M, Bowie J, Slade J, et al. Assessing capacity of faith-based organizations for health promotion activities. Health Promot Pract 2018 Sep 23;19(5):714-723 [FREE Full text] [doi: 10.1177/1524839917737510] [Medline: 29058956]

Abbreviations

ADRD: Alzheimer disease and related dementias
AI: artificial intelligence
GPT: generative pretrained transformer
LLM: large language model
OHC: online health community

Edited by J Wang; submitted 11.06.24; peer-reviewed by M Gasmi , K Bonds Johnson; comments to author 02.10.24; revised version received 15.10.24; accepted 03.12.24; published 08.01.25.

Please cite as:
Bosco C, Otenen E, Osorio Torres J, Nguyen V, Chheda D, Peng X, Jessup NM, Himes AK, Cureton B, Lu Y, Hill CV, Hendrie HC, Barnes PA, Shih PC
Designing a Multimodal and Culturally Relevant Alzheimer Disease and Related Dementia Generative Artificial Intelligence Tool for Black American Informal Caregivers: Cognitive Walk-Through Usability Study
JMIR Aging 2025;8:e60566
URL: https://aging.jmir.org/2025/1/e60566
doi:10.2196/60566
PMID:

©Cristina Bosco, Ege Otenen, John Osorio Torres, Vivian Nguyen, Darshil Chheda, Xinran Peng, Nenette M Jessup, Anna K Himes, Bianca Cureton, Yvonne Lu, Carl V Hill, Hugh C Hendrie, Priscilla A Barnes, Patrick C Shih. Originally published in JMIR Aging (https://aging.jmir.org), 08.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Assessment of Technology Readiness in Norwegian Older Adults With Long-Term Health Conditions Receiving Home Care Services: Cross-Sectional Questionnaire Study

Sverre Bergh[1,2], MD, PhD; Jūratė Šaltytė Benth[1,3,4], PhD; Lisbeth Dyrendal Høgset[1], MHS; Britt Rydjord[5], PhD; Lars Kayser[6], Prof Dr Med

[1]Research Centre for Age-related Functional Decline and Disease, Innlandet Hospital Trust, Ottestad, Norway

[2]Norwegian National Centre for Aging and Health, Vestfold Hospital Trust, Tønsberg, Norway

[3]Institute for Clinical Medicine, University of Oslo, Oslo, Norway

[4]Health Services Research Unit, Akershus University Hospital, Oslo, Norway

[5]Department for Research and Innovation, Innlandet Hospital Trust, Ottestad, Norway

[6]Department of Public Health, University of Copenhagen, Copenhagen, Denmark

Corresponding Author:
Sverre Bergh, MD, PhD
Research Centre for Age-related Functional Decline and Disease
Innlandet Hospital Trust
Pb 68
Ottestad, 2312
Norway
Phone: 47 45679393
Email: sverre.bergh@sykehuset-innlandet.no

Abstract

Background: With the increasing number of older adults globally, there is a constant search for new ways to organize health care services. Digital health services are promising and may reduce workload and at the same time improve patient well-being. A certain level of eHealth literacy is needed to be able to use digital health services. However, knowledge of technology readiness in this target group of older adults is unclear.

Objective: The aim of this study was to understand the technology readiness level of a group of older adults who were provided home care services in order to address the present and future needs of this group in relation to the implementation of digital health care services.

Methods: This quantitative cross-sectional study included 149 older adults from Norway receiving home care services. The participants completed the Readiness and Enablement Index for Health Technology (READHY) instrument, assessments of well-being (World Health Organization-Five Well-Being Index [WHO-5]), and assessments of demographic and clinical variables (sex, age, education, living situation, comorbidity, use of digital devices, and use of IT). Cluster analyses were used to group the users according to their technology readiness.

Results: The mean participant age was 78.6 (SD 8.0) years, and 55.7% (83/149) were women. There was good consistency within the assumed READHY scales (Cronbach α=.61-.91). The participants were grouped into 4 clusters, which differed in terms of READHY scores, demographic variables, and the use of IT in daily life. Participants in cluster 1 (n=40) had the highest scores on the READHY scales, were younger, had a larger proportion of men, had higher education, and had better access to digital devices and IT. Participants in cluster 4 (n=16) scored the lowest on eHealth literacy knowledge. Participants in cluster 1 had relatively high levels of eHealth literacy knowledge and were expected to benefit from digital health services, while participants in cluster 4 had the lowest level of eHealth literacy and would not easily be able to start using digital health services.

Conclusions: The technology readiness level varied in our cohort of Norwegian participants receiving home care. Not all elderly people have the eHealth literacy to fully benefit from digital health services. Participants in cluster 4 (n=16) had the lowest scores in the eHealth Literacy Questionnaire scales in the READHY instrument and should be offered nondigital services or would need extensive management support. The demographic differences between the 4 clusters may inform stakeholders about which older people need the most training and support to take advantage of digital health care services.

XSL•FO
RenderX

(JMIR Aging 2025;8:e62936) doi:10.2196/62936

KEYWORDS

eHealth literacy; digital health services; technology readiness; Readiness and Enablement Index for Health Technology; READHY; chronic conditions

Introduction

In the coming years, the world will have an increasing number of older adults. According to the World Health Organization (WHO), 1 in 6 people in the world will be 60 years or older in 2030 [1]. Today's aging population is more healthy than previous generations, with a longer life expectancy and more years without frailty [2]. Nevertheless, in the future, there will be more older adults with long-term health conditions and frailty [1]. This will further increase the burden on health care services in the coming years. This change in demographics together with a reduced workforce due to a reduction in the birth rate calls for new ways to organize health and care services. There is an urgent need for more efficient ways to provide the needed health care with fewer in-person contacts, the involvement of informal caregivers, an increase in self-management, and a reduction in travel time, particularly in rural areas. Increased digitalization of health care services could be a solution, which could enable and engage more patients, resulting in increased levels of self-efficacy, self-management, and empowerment [3], and thus reduce the burden on the workforce [4]. Moreover, telehealth solutions may increase the number of virtual contacts and provide means for self-monitoring and dialogues with health professionals as an immediate response to the early signs of deterioration in the health condition [5]. The use of telemedicine may be of particular importance in rural areas with scattered settlements, saving travel time and costs for both health personnel and patients. Although evidence for the assumed benefits of the digital transformation of health care services is scarce, particularly among older adults, it will continue in the coming years along with the general digitalization of both private and public services in Norway and many other countries.

In 2021, 96% of Norwegians older than 9 years had access to a smartphone [6]. However, this indicates that 4%, or approximately 200,000 people, do not have a smartphone. A higher number of people probably have an insufficient level of digital health literacy. In a study from 2021, the Norwegian Directorate of Health found that 13% of people aged ≥60 years never used the internet, while 9% used the internet only once a week. Moreover, only 66% of people aged ≥80 years used the internet. Further, 44% of people aged ≥60 years reported no or low experience with digital IDs and 2-factor identification, which are needed to access public web pages [7].

One benefit from the digital health transformation is that citizens and patients may communicate from their homes and receive care and treatment virtually. Services may also be provided by specialists at hospitals in collaboration with registered nurses in municipalities and the person's own general practitioner, offering better coordination. This requires the individuals being cared for to have access to the internet; have devices, such as smartphones, tablets, laptops, and PCs; and have the ability to use these devices. For those not having these resources, it is essential to identify this challenge and plan for the involvement of supporting resources or nondigital services. This calls for new ways to plan for digitally enabled services and usage of digital tools, and to be aware of the specific resources and needs of individuals when offering these.

In this context, awareness of the person's level of digital health literacy [8] and level of technology readiness [9] is important and may be helpful information together with sociodemographic characteristics, including access to and familiarity with digital devices. In particular, older adults have a lower level of digital health literacy, and as lower levels of digital health literacy may be related to economic status, it may be hypothesized that older people living in rural areas with precisely these characteristics may have a low or even insufficient level of digital health literacy. Technology readiness may be a barrier for offering telehealth solutions, which otherwise would be of huge benefit to these people, due to the fact that the remote areas of living may prohibit access to in-person services, particularly outside the summer season. This calls for knowledge about the levels of digital health literacy and technology readiness of different populations.

eHealth literacy, which is currently often used synonymously with digital health literacy, was defined in 2006 by Norman and Skinner as "the ability to seek, find, understand, and appraise health information from electronic sources and apply the knowledge gained to addressing or solving a health problem" [8]. Several factors, such as frailty, low level of digital health literacy, and motoric or cognitive impairment, may be barriers for patients when introducing new technologies [10].

One way to identify barriers is by measuring the level of digital health literacy using a multi-faceted instrument such as the eHealth Literacy Questionnaire (eHLQ) [11]. Another way is to expand on the factors necessary for being able to apply digital health literacy in a context, by using the Readiness and Enablement Index for Health Technology (READHY) [9]. The READHY instrument builds on the eHLQ but includes 4 scales related to self-management from the Health Education Impact Questionnaire (heiQ) [12] and 2 scales related to support from the Health Literacy Questionnaire (HLQ) [13]. The READHY instrument has been proven to inform about variations in needs and skills in relation to technology readiness presented as specific clusters.

The READHY instrument can be used both as a framework for qualitative studies [14] and as an instrument to assess the experience of support, elements of self-management, and levels of digital health literacy. It has been used in a number of contexts, such as cancer survivors [15], type 2 diabetes mellitus (T2DM) patients [16], older hospitalized medical patients, and people living with inflammatory bowel disease [17,18]. The separate scales of the READHY instrument have been translated and validated in the Norwegian language, but the complete

XSL•FO
RenderX

READHY instrument has not yet been published in Norwegian research [19,20].

Despite emerging knowledge within various specific conditions, studies focusing on the technology readiness of older adults are lacking. People receiving home care services may have other needs for digital health services than people with specific diseases, such as cancer and T2DM, and it is important to have an instrument that is broad enough to catch these differences. In our study, the READHY instrument will be tested as an option to elucidate the characteristics of this population of older adults living in a rural area.

Data obtained using the READHY instrument can be clustered to identify segments of users, and this segmentation can be enriched with other characteristics, such as sociodemographics, health conditions, and well-being, forming clusters to help health professionals recognize the need for support or training among the people they serve [21]. This information may help health service organizations to plan for subpopulations based on the identified clusters of the general population and with knowledge about the size of particular groups. It may also help health care professionals to have better conversations when introducing new digital health tools or offering telehealth services to ensure meaning and motivation for older adults [22].

The aim of this study was to explore the technology readiness level in a population of older adults living in a rural area, a group that is hypothesized to be less technology ready. Further, the study attempted to examine their digital resources and behaviors and evaluate associations between these factors and their self-reported health and mental well-being, which may influence their motivation and ability to create meaning in using digital health services and tools.

The results of the study may help service providers to better plan for the specific needs of subgroups and provide health care professionals with insights that may help them to better understand how to introduce new digital health tools and services in this group of older adults living in rural areas. This group is more likely to be frail and at the same time may benefit more from virtual and other contacts that are not in-person. Using clustering analysis, we explored the technology readiness level of older adults receiving home care and their characteristics in terms of resources and digital behaviors (ie, sociodemographic variables and use of digital devices).

Methods

Design

This cross-sectional study included 149 older adults recruited from 7 municipalities in South-Eastern Norway. The data were collected by 1 to 2 health care staff in each of the 7 participating municipalities. The same research nurse trained all data collectors (2-hour training session), and the data collectors were supervised by the research nurse during the study period. The participants completed the questionnaires listed below, either self-reported by themselves or in an interview with the health care staff. The participants were approached in a face-to-face setting where they were informed about the study, signed the consent form, and completed the questionnaires.

Population

Health care staff in the municipal home care services and at the lung department of Innlandet Hospital Trust recruited the participants. Patients admitted to the outpatient clinic at the lung department and people receiving home care services from the municipalities were invited to participate in the study. The inclusion criteria were as follows: (1) living at home, (2) ≥65 years of age, (3) receiving home care services, (4) presence of long-term health conditions, (5) fluency in Norwegian, and (6) written informed consent. The exclusion criteria were as follows: (1) severe dementia (Clinical Dementia Rating scale score of 3), (2) terminal disease with <3 months of expected survival, and (3) severe psychiatric disease, including drug abuse. A total of 211 individuals were eligible, of which 62 were not included as they did not consent to participate. Those not included were on average 83.5 (SD 5.7) years old, and 61.3% were female. The process of participant inclusion was from September 2021 to June 2023.

Included Data and Questionnaires

We collected data on age (years), sex (male/female), living situation (alone, with spouse/partner, or with others), comorbidity, education (compulsory lower secondary school, upper secondary education, master's degree, or more than 4 years university), use of digital devices (laptop, PC, tablet, or smartphone), use of technology to access the general physician and electronic medical record, and digital competence (very poor, poor, average, good, or very good).

The READHY instrument consists of 65 items, with 4 scales from the heiQ [12], 2 scales from the HLQ [13], and all scales from the eHLQ (Table 1) [11]. The heiQ is used to assess patient education and self-management and is validated in Norwegian [12,19]. The HLQ is used to assess health literacy and has been translated to and validated in many languages including Norwegian [13,23]. The eHLQ is used to assess the benefits of telemedicine for individuals, evaluate how telehealth services work together, and examine the telemedicine knowledge of individuals [18]. It has been translated, validated, and psychometrically tested in Norwegian [24]. The full READHY instrument is validated in several languages but not yet in Norwegian. The READHY instrument assesses the person's health technology readiness [9]. Each item in the READHY instrument can be scored from 1 to 4 (strongly disagree, disagree, agree, and strongly agree). The scoring of "strongly agree" in the item heiQ8 *emotional distress* is reversed for the statistical analysis, with the highest score representing the lowest level of distress. The scores of the scales in the READHY instrument are calculated as the mean scores of the included items for each scale. The scores of scales with missing data have been included if the number of missing items is ≤50%.

XSL•FO
RenderX

Table 1. Scales from the Health Education Impact Questionnaire (heiQ), Health Literacy Questionnaire (HLQ), and eHealth Literacy Questionnaire (eHLQ) included in the Readiness and Enablement Index for Health Technology (READHY).

Questionnaire and scale	Description
heiQ[a]	
heiQ3	Self-monitoring and insight
heiQ4	Constructive attitudes and approaches
heiQ5	Skill and technique acquisition
heiQ8	Emotional distress
HLQ[b]	
HLQ1	Feeling understood and supported by health care providers
HLQ4	Social support for health
eHLQ[c]	
eHLQ1	Using technology to process health information
eHLQ2	Understanding of health concepts and language
eHLQ3	Ability to actively engage with digital services
eHLQ4	Feeling safe and in control
eHLQ5	Motivated to engage with digital services
eHLQ6	Access to digital services that work
eHLQ7	Digital services that suit individual needs

[a]heiQ: Health Education Impact Questionnaire.

[b]HLQ: Health Literacy Questionnaire.

[c]eHLQ: eHealth Literacy Questionnaire.

Cutoff values are not meaningful from a psychometric approach, and the values for scoring in the READHY scales have not been calibrated to each other or to outcomes. Despite this, scores below 2.7 on the scales may mirror that the person has a level of "not being sufficient" as not all items in the scales are scored as "strongly agree" or "agree." A score of 2.0 or lower on the scales may be considered problematic for managing digital health services as most items in the scales are scored as "strongly disagree" or "disagree" [25]. The READHY instrument takes 10 to 20 minutes to complete in most populations (including older adults), but it could take up to 30 minutes to complete for individuals with cognitive impairment.

The World Health Organization-Five Well-Being Index (WHO-5) consists of 5 items. Each item is scored from 0 to 5, where 5 indicates *all the time* and 0 indicates *at no time*. The score of each item is multiplied by 4, giving a total score from 0 to 100.

Some additional questions were asked to enrich the data. A question about digital competence was asked: "If someone knowing you well should assess your IT competence, would they say it is very poor/poor/average/good/very good?" A question about the participants' health was taken from the 36-item Short-Form Survey instrument (SF-36): "In general, would you say your health is excellent/very good/good/fair/poor?" Scores range from 1 (excellent) to 5 (poor) [26], and they were treated as a continuous variable.

Statistical Analysis

The participant characteristics have been presented as means and SDs or frequencies and percentages, as appropriate. The READHY scale score was calculated as the mean of the item scores when fewer than half of the items had missing values. This resulted in 2 excluded cases where the scale score could not be calculated. Ideally, one should start with assessing the dimensionality of the READHY instrument by applying a factor analysis. However, due to convergence problems caused by a very small sample size with respect to the number of items, this approach was not possible. We therefore report the Cronbach α for the scales identified by a previous factor analysis of a Danish dataset [9].

To identify groups of participants based on the READHY scales, we first applied a hierarchical cluster analysis with Euclidean squared distance and Ward linkage to identify the possible number of groups [27]. The results of the hierarchical cluster analysis were then used to determine an appropriate number of clusters in the next step, a k-means cluster analysis [28] with Euclidean similarity measure. By examining the dendrogram from the first step, we expected to gain insights into the natural grouping of data, which would inform the subsequent k-means analysis. In this way, we leveraged the strengths of both methods for more effective clustering and based our conclusion on k-means analysis, which is easier to interpret due to each individual being assigned to one cluster. While choosing the final number of clusters, we considered clinical relevance,

XSL•FO
RenderX

cluster size, small within-cluster variability, and between-cluster heterogeneity.

One-way ANOVA was then applied to assess the differences between the identified clusters with respect to the READHY scales. Further, one-way ANOVA, the chi-square test, and the Fisher exact test were applied to compare the profiles of the identified clusters based on the following variables: all scale scores, age, sex, education (≤12 years or >12 years), living situation (alone, with spouse/partner, or with others), comorbidity (number of chronic diagnoses), use of a laptop, PC, or tablet at least once a week, use of a smartphone at least once a week, use of IT in previous work or studies, use of IT to communicate with public services, use of the Norwegian digital identifier to access a webpage, use of IT to communicate with a general practitioner, logging into a national health webpage, use of IT to find health information on the internet or social media, reading medical journals or test results on a national health webpage, IT competence assessed by others (very poor, poor, average, good, or very good), WHO-5 score, and response to the question "In general, would you say your health is excellent/very good/good/fair/poor?" Pairwise comparisons of groups were performed as post hoc analyses. The results of these post hoc analyses were reported as mean differences and SDs for continuous variables, mean differences and SDs in proportions for dichotomous variables, and effect sizes calculated as absolute values of z-statistics from the Mann-Whitney *U* test divided by the square root of the total sample size for ordinal variables, along with *P* values. Only significant pairwise differences were tabulated.

Two linear regression models were estimated to assess the associations of WHO-5 (primary) and IT competence assessed by others (secondary) with prechosen covariates: age, sex, education (≤12 years or >12 years), living situation (alone, with spouse/partner, or with others), comorbidity, use of a laptop, PC, or tablet at least once a week, use of a smartphone at least once a week, use of IT to communicate with a general practitioner, use of IT to find health information on the internet or social media, reading medical journals or test results on a national health webpage, and response to the question "In general, would you say your health is excellent/very good/good/fair/poor?" The regression models were estimated in cases with no missing values for covariates.

Statistical analyses were performed using STATA version 18 (StataCorp). A *P* value of <.05 was considered statistically significant. No adjustment for multiple testing was applied; however, we separated between ANOVA as the main analysis and post hoc analyses considered purely exploratory. Although we have only reported the results of the post hoc analyses with *P*<.05, we emphasize that all pairwise comparisons were performed. Further, the regression analyses have been considered secondary. Moreover, there is clarity regarding the number of tests performed, allowing the reader to consider the reliability of our findings.

Ethical Considerations

Participation in the study was based on informed written consent. The study was considered by The Regional Committee for Medical and Health Research Ethics in Norway as being outside its jurisdiction. The data collection, data storage, data analysis, and publication of the results were approved by the data protection officer at Innlandet Hospital Trust (number: 14832226). Participants could withdraw from the study at any time and request for their data to be deleted.

Results

The descriptive results of the cohort are presented in Table 2. The mean age of the participants was 78.6 (SD 8.0) years, and 55.7% (83/149) were women. Moreover, 19.2% (27/141) had higher education (>12 years) and 69.8% (104/149) were living alone. The mean scores of the READHY scales ranged from 2.0 (SD 0.5) for the eHLQ1 to 3.1 (SD 0.3) for the heiQ3. The heiQ scales assessed patient education and self-management, and the mean scores for the 4 scales ranged from 2.8 (SD 0.6) to 3.1 (SD 0.3), indicating high levels of patient education and self-management. The HLQ scales assessed health literacy, and the mean scores ranged from 2.9 (SD 0.4) to 3.0 (SD 0.5), indicating a high level of health literacy. The eHLQ scales assessed eHealth literacy, and the mean scores ranged from 2.0 (SD 0.5) to 3.0 (SD 0.3), indicating that eHealth literacy was low in our population. Moreover, the scores differed between the scales. Table 3 presents Cronbach α values for the READHY scales. Except for the eHLQ2 scale with a low Cronbach α of .61, the Cronbach α values ranged from .70 to .91 for the READHY scales, showing good consistency within the assumed scales [29].

XSL•FO
RenderX

Table 2. Summarized statistics of the cohort.

Variable	Value (N=149)
Age (years) (n=148), mean (SD)	78.6 (8.0)
Gender (female), n (%)	83 (55.7)
Education (n=141), n (%)	
≤12 years	114 (80.9)
>12 years	27 (19.2)
Living situation, n (%)	
Living alone	104 (69.8)
Living with spouse/partner/others	45 (30.2)
Number of chronic diagnoses (n=144), mean (SD)	2.6 (1.4)
Use of digital devices at least once a week, n (%)	
PC, laptop, or tablet	72 (48.3)
Smartphone	89 (59.7)
IT to communicate with a general practitioner	34 (22.8)
IT to find health information on the internet or social media	17 (11.6)
Reading medical journals or test results on a national health webpage, n (%)	
Never/rarely	116 (78.4)
Sometimes/always	32 (21.6)
In general, would you say your health is excellent/very good/good/fair/poor, mean (SD)	1.2 (0.8)
READHY[a] scales, mean (SD)	
heiQ[b]	
heiQ3	3.1 (0.3)
heiQ4	2.9 (0.4)
heiQ5	2.8 (0.4)
heiQ8	2.8 (0.6)
HLQ[c]	
HLQ1 (n=148)	3.0 (0.5)
HLQ4 (n=148)	2.9 (0.4)
eHLQ[d]	
eHLQ1 (n=147)	2.0 (0.5)
eHLQ2	2.8 (0.4)
eHLQ3 (n=147)	2.2 (0.6)
eHLQ4	3.0 (0.3)
eHLQ5	2.2 (0.5)
eHLQ6	2.4 (0.4)
eHLQ7 (n=148)	2.2 (0.6)

[a]READHY: Readiness and Enablement Index for Health Technology.

[b]heiQ: Health Education Impact Questionnaire.

[c]HLQ: Health Literacy Questionnaire.

[d]eHLQ: eHealth Literacy Questionnaire.

XSL•FO
RenderX

Table 3. Cronbach α values of the Readiness and Enablement Index for Health Technology (READHY) scales.

Scale	Cronbach α
heiQ[a]	
heiQ3: Self-monitoring and insight (6 items)	.69
heiQ4: Constructive attitudes and approaches (5 items)	.77
heiQ5: Skills and technique acquisition (4 items)	.70
heiQ8: Emotional distress (6 items)	.90
HLQ [b]	
HLQ1: Feeling understood and supported by health care providers (4 items)	.83
HLQ4: Social support for health (5 items)	.77
eHLQ[c]	
eHLQ1: Using technology to process health information (5 items)	.86
eHLQ2: Understanding of health concepts and language (5 items)	.61
eHLQ3: Ability to actively engage with digital services (5 items)	.91
eHLQ4: Feeling safe and in control (5 items)	.71
eHLQ5: Motivated to engage with digital services (5 items)	.80
eHLQ6: Access to digital services that work (6 items)	.78
eHLQ7: Digital services that suit individual needs (4 items)	.85

[a]heiQ: Health Education Impact Questionnaire.

[b]HLQ: Health Literacy Questionnaire.

[c]eHLQ: eHealth Literacy Questionnaire.

Based on visually inspecting a dendrogram from the hierarchical cluster analysis (not presented), exploring 3 to 6 cluster solutions by k-means cluster analysis, and considering clinical relevance, within-cluster variability, between-cluster heterogeneity, and reasonable sample size, the 4-cluster model was chosen (Table 4). The details of patient profiles in the 4 clusters are described in Multimedia Appendix 1. The participants in cluster 1 (n=40) had on average the highest scores on the READHY scales (meaning better eHealth literacy), except the HLQ4 scale. The HLQ4 scale assesses the social support for health, where the scores were intermediate. Participants in cluster 1 were younger than the other participants (P=.006), had a larger proportion of men (P<.001), had higher education (P<.001), had better access to smartphones and PCs, laptops, or tablets (P<.001), had used IT more in their previous work (P<.001), and currently used IT to access health information and stay in touch with the health care system (P<.001). The participants in cluster 3 (n=25) had the lowest average scores on the heiQ4, heiQ5, and heiQ8 scales (patient education and self-management) and the 2 HLQ scales (feeling understood and supported by health care providers, and social support for health). The participants in cluster 4 (n=16) had the lowest average scores on the eHLQ scales (meaning low eHealth literacy knowledge), except the eHLQ4 scale (feeling safe and in control), where their scores were comparable to those of the 3 other clusters. Thus, the feelings of being safe and in control were equal in the 4 clusters. Cluster 2 was the largest (n=66) and had intermediate scores on most of the scales. Interestingly, the question “IT competence assessed by others” differentiated between all clusters, except between cluster 2 and cluster 3 (P<.001).

XSL•FO
RenderX

Table 4. Results of k-means clustering with 4 clusters (n=147).

Scale	Cluster 1 (n=40), mean (SD)	Cluster 2 (n=66), mean (SD)	Cluster 3 (n=25), mean (SD)	Cluster 4 (n=16), mean (SD)	*P* value for ANOVA	*P* value for pairwise comparisons, mean difference (SD) and *P* value
heiQ[a]						
heiQ3	3.2 (0.3)	2.9 (0.3)	3.0 (0.3)	3.2 (0.5)	<.001	1 vs 2: 0.3 (0.1), *P*<.001; 1 vs 3: 0.2 (0.1), *P*=.01; 2 vs 4: –0.2 (0.1), *P*=.009
heiQ4	3.1 (0.3)	3.0 (0.3)	2.4 (0.3)	3.0 (0.5)	<.001	1 vs 3: 0.7 (0.1), *P*<.001; 2 vs 3: 0.6 (0.1), *P*<.001; 3 vs 4: –0.6 (0.1), *P*<.001
heiQ5	2.9 (0.4)	2.8 (0.3)	2.6 (0.5)	3.0 (0.5)	.001	1 vs 2: 0.1 (0.1), *P*=.047; 1 vs 3: 0.4 (0.1), *P*=.001; 2 vs 3: 0.3 (0.1), *P*=.004; 3 vs 4: –0.4 (0.2), *P*=.01
heiQ8	2.9 (0.6)	3.0 (0.3)	2.1 (0.5)	2.8 (0.7)	<.001	1 vs 3: 0.8 (0.1), *P*<.001; 2 vs 3: 0.9 (0.1), *P*<.001; 3 vs 4: –0.7 (0.2), *P*=.002
HLQ[b]						
HLQ1	3.2 (0.5)	2.9 (0.4)	2.9 (0.3)	3.0 (0.7)	.003	1 vs 2: 0.3 (0.1), *P*<.001; 1 vs 3: 0.3 (0.1), *P*=.006
HLQ4	3.0 (0.4)	2.8 (0.4)	2.6 (0.4)	3.2 (0.4)	<.001	1 vs 3: 0.3 (0.1), *P*=.004; 2 vs 3: 0.2 (0.1), *P*=.009; 2 vs 4: –0.4 (0.1), *P*=.001; 3 vs 4: –0.6 (0.1), *P*<.001
eHLQ[c]						
eHLQ1	2.5 (0.5)	2.0 (0.2)	2.1 (0.3)	1.1 (0.2)	<.001	1 vs 2: 0.6 (0.1), *P*<.001; 1 vs 3: 0.4 (0.1), *P*=.001; 1 vs 4: 1.4 (0.1), *P*<.001; 2 vs 3: –0.2 (0.1), *P*=.002; 2 vs 4: 0.8 (0.1), *P*<.001; 3 vs 4: 1.0 (0.1), *P*<.001
eHLQ2	3.1 (0.3)	2.6 (0.3)	2.8 (0.2)	2.6 (0.5)	<.001	1 vs 2: 0.4 (0.1), *P*<.001; 1 vs 3: 0.3 (0.1), *P*<.001; 1 vs 4: 0.4 (0.1), *P*<.001; 2 vs 3: –0.2 (0.1), *P*=.02
eHLQ3	2.9 (0.5)	2.1 (0.3)	2.3 (0.4)	1.1 (0.2)	<.001	1 vs 2: 0.8 (0.1), *P*<.001; 1 vs 3: 0.6 (0.1), *P*<.001; 1 vs 4: 1.8 (0.1), *P*<.001; 2 vs 3: –0.2 (0.1), *P*=.004; 2 vs 4: 1.0 (0.1), *P*<.001; 3 vs 4: 1.2 (0.1), *P*<.001
eHLQ4	3.1 (0.4)	3.0 (0.2)	3.0 (0.2)	3.1 (0.4)	.08	1 vs 2: 0.1 (0.1), *P*=.02
eHLQ5	2.7 (0.5)	2.0 (0.2)	2.1 (0.3)	1.5 (0.4)	<.001	1 vs 2: 0.6 (0.1), *P*<.001; 1 vs 3: 0.6 (0.1), *P*<.001; 1 vs 4: 1.2 (0.1), *P*<.001; 2 vs 4: 0.5 (0.1), *P*<.001; 3 vs 4: 0.6 (0.1), *P*<.001
eHLQ6	2.9 (0.3)	2.3 (0.2)	2.3 (0.3)	1.6 (0.3)	<.001	1 vs 2: 0.6 (0.05), *P*<.001; 1 vs 3: 0.5 (0.1), *P*<.001; 1 vs 4: 1.2 (0.1), *P*<.001; 2 vs 4: 0.6 (0.1), *P*<.001; 3 vs 4: 0.7 (0.1), *P*<.001
eHLQ7	2.8 (0.4)	2.0 (0.2)	2.0 (0.3)	1.2 (0.4)	<.001	1 vs 2: 0.8 (0.1), *P*<.001; 1 vs 3: 0.8 (0.1), *P*<.001; 1 vs 4: 1.6 (0.1), *P*<.001; 2 vs 4: 0.8 (0.1), *P*<.001; 3 vs 4: 0.8 (0.1), *P*<.001

[a]heiQ: Health Education Impact Questionnaire.

[b]HLQ: Health Literacy Questionnaire.

[c]eHLQ: eHealth Literacy Questionnaire.

In the bivariate linear regression model, WHO-5 as an outcome (Multimedia Appendix 2) was significantly associated with education and the question “In general, would you say your health is excellent/very good/good/fair/poor?”, while only the latter question remained significant in the multiple model. A 1-point increase in the score of the question “In general, would you say your health is excellent/very good/good/fair/poor?” was associated with a 12.7-point increase in the WHO-5 score.

In the multiple linear regression model with the question “IT competence assessed by others” as an outcome (Multimedia Appendix 3), only “use of a PC, laptop, or tablet at least once in the last week” and “reading medical journals or test results on a national health webpage” were significantly associated with the outcome.

Discussion

Summary of the Main Findings

In this study, we included participants from a rural area in Norway with less education than other parts of Norway, and our recruitment procedure resulted in the inclusion of slightly more women than men. The participants were assessed for technology readiness using the READHY instrument. We found that people receiving home care services in this area had high

XSL•FO
RenderX

levels of patient education and self-management (scores of 2.8-3.1) and high levels of health literacy (scores of 2.9-3.0) but generally low levels of eHealth literacy (scores of 2.0-3.0). The findings were based on the assumption that a score of ≥2.7 (out of a maximum of 4) reflects a sufficient level, as this score reflects overweight of the response "agree" or "strongly agree" to all the items in the scale [25]. The level of eHealth literacy also differed, as scores on the eHLQ scales and between the participants differed, showing less stability than the heiQ and HLQ scales. The eHealth scores were lower in this study than in a previous study by Kayser et al [11], which can be explained by differences in age, education, and study settings. Therefore, in a rural population like the population in our study, the low level of eHealth literacy should be considered when introducing digital health services in the home care setting.

Low eHealth Literacy

The most interesting results for health care services were the characteristics of the participants in cluster 4, with the lowest scores on the eHLQ scales, indicating low eHealth literacy knowledge (Multimedia Appendix 1). These participants were older, were mainly women, and had a lower level of education than participants in cluster 1 (with the highest level of eHealth literacy). Participants in cluster 4 seldom used a PC, laptop, tablet, or smartphone, and they seldom accessed public services generally and health care services specifically compared with participants in cluster 1. The participants in cluster 4 had several diagnoses, although not significantly more than that among participants in cluster 1. We expect their need for care and support from the home care service to be high, and this population group could benefit from digital home monitoring. Nevertheless, the results show that these participants would need more help and facilitation from the health care service than others to take advantage of digital health services. Offering nontechnological services may be a better option. Social resources that could support them in the use of digital health services must be identified if digital health services are offered. If not, they could be discriminated from receiving the same health care services as others.

Screening for eHealth Literacy

We found that the response to the question "If someone knowing you well should assess your IT competence, would they say it is very poor/poor/average/good/very good?" was associated with several variables in the bivariate model, but only the responses to the items "use of PC, laptop, or tablet at least once a week" and "reading medical journals or test results on a national health webpage" remained significant in the multiple model (Multimedia Appendix 3). Surprisingly, no other variables remained significant in the multiple model, but the most likely explanation was the low power in our study. Further, we found a strong association between belonging to 1 of the 4 clusters and the question "If someone knowing you well should assess your IT competence, would they say it is very poor/poor/average/good/very good?" (Multimedia Appendix 1). The answer differentiated between participants in cluster 1 and those in clusters 2, 3, and 4; between participants in cluster 2 and those in cluster 3; and between participants in cluster 3 and those in cluster 4. These associations contribute to the increasing amount of data supporting the relevance and content validity of this single-item scale and indicate that a single question could be used to screen larger populations for the level of digital health skills. Those scoring low on the single question could be further assessed with multidimensional instruments such as the READHY instrument. The READHY instrument consists of 13 scales with a total of 65 items and takes 10 to 20 minutes to complete. It may be exhausting to complete, especially for old frail people with cognitive impairment. Thus, it is preferable if clinicians can scan a population and collect the same information about the technological readiness level using a single question. This single question could potentially help to increase the number of screenings to identify those in need of further investigation. Here, the READHY instrument can be used to identify specific resources and challenges in relation to the use of digitally enabling or assisting services and technologies. However, it is important to emphasize that using 1 screening item will lead to a loss in the understanding of where particular challenges are. Further, we should be aware that it only screens for digital expertise, and self-management and social support are not taken into consideration with this approach.

Well-Being

We analyzed the variables associated with the scores of the WHO-5 (Multimedia Appendix 2). In the bivariate model, there was an association between a higher educational level and higher scores on the WHO-5, which was not significant when adjusting for other variables such as age, sex, living situation, and number of chronic diagnoses. People with a higher level of education had better health, were younger, and had better digital health literacy, which contributed to the loss of the bivariate effect of education on well-being (WHO-5) when adjusting for these variables. Further, we did not find associations of the co-habitant status and number of chronic diagnoses with well-being assessed by the WHO-5. These findings need some consideration, as other European studies have found that living alone is associated with a lower score on the WHO-5 [30-32]. One explanation may be that the context in these studies differs from that in our study (cohort living in a rural area and having long travel distances). The participants in our study were already receiving home care services and thus were in regular contact with a support person and had access to a service they could call, which may be of importance for well-being. The home care service may be equivalent to the effect of living together with others, as people receiving home care services are in regular contact with a support person and have access to a service they can call.

Cluster Profiles

Our analysis resulted in 4 cluster profiles based on the k-means clustering analysis (Table 4). One cluster with high scores in all the READHY scales, 1 with low scores in the eHLQ scales (eHealth literacy), 1 with low scores in the heiQ scales (health education) and the HLQ scales (health literacy), and 1 with intermediate scores. Being male, being younger, having higher education, and having good access to digital devices and knowledge of IT were all associated with a better technology readiness level. This result was comparable to that in a study by Kayser et al [9], which had 1 cluster with the highest scores

XSL•FO
RenderX

in the READHY scales, 1 with low scores in the eHLQ scales, and 2 involving intermediate profiles. In the Danish study by Kayser et al [9], 25% of the participants were in the cluster with the highest READHY scale scores, which is similar to the proportion of 27% in our study. When considering another study [11], comparable proportions of participants were in the group with the lowest eHLQ scale scores (13% in the previous study and 11% in our study). In another Danish study by Thorsen et al [16], it was found that 5 clusters were more meaningful for the population of patients with T2DM. They also described 1 cluster with the highest READHY scale scores and 1 with low scores on the eHLQ scales, with the 3 remaining clusters having intermediate scores [16]. In a study with Danish cancer survivors, individuals in the cluster with the highest READHY scale scores were younger than those in the other clusters [9], as in our study. On the other hand, in the study of Danish patients with T2DM, 1 of the intermediate clusters had the lowest age [16]. A recent Norwegian study recruiting 260 patients from medical and surgical wards found 6 HLQ clusters and 6 eHLQ clusters [33]. They did not collect heiQ data but found that the participants with the lowest HLQ and eHLQ scores had comparable socioeconomic clusters, similar to the participants in cluster 4 in our study [33]. Thus, the 4 profiles based on the k-means clustering analysis of our data are comparable with those in 2 Danish studies and 1 Norwegian study, despite the mean age being higher in our study.

As explained in the statistical analysis subsection in the Methods section, a confirmatory factor analysis to assess the psychometric properties of the READHY instrument was considered. However, it resulted in convergence problems, most likely due to a very small sample size with respect to the number of items. Therefore, we did not succeed in our aim to validate the Norwegian version of the READHY instrument. Instead, we reported the Cronbach α for the scales identified by a previous factor analysis of a Danish dataset [9], which showed good consistency within the assumed scales.

Limitations

The study had several limitations. No formal power calculation was performed prior to the study, and a convenience sample was used. It is suggested that cluster analysis should only be applied when a good group separation is expected, and it is recommended to include a sample size of approximately 20 to 30 individuals per expected group [34]. Based on previous cluster analysis of the READHY instrument [9,16,33], we expected 3 to 5 clusters of individuals in our study. Thus, a sample size of 149 was sufficient according to recommendations. The inclusion of more participants in the study could strengthen our results, could allow us to perform a factor analysis, and might make it possible to report more associations from the regression models (Multimedia Appendix 2 and 3).

The data were collected by 1 to 2 health care staff in each of the 7 participating municipalities. The same research nurse trained all data collectors, and all had experience in assessing and testing the patients. However, we cannot rule out that the use of several data collectors may have caused biases in the data. On the other hand, we used well-established assessment tools that have been found to be reliable and valid. We used a long and extensive case report form (CRF) with several questionnaires. Some of the questionnaires included many questions. The participants may have become tired, especially those with impaired cognitive function, which might have influenced the data quality. The participants were given the opportunity to answer the questionnaires at their own pace and could take breaks if needed. In the end, we found that the data collected were of excellent quality, with limited missing data.

We recruited participants from only 7 municipalities in a rural region in South-Eastern Norway, with a lower education level than that generally in Norway. This fact, in addition to the fact that the participants were recruited from among those receiving home care services, may reduce the generalizability of the study. The READHY instrument should be tested in home care settings in other regions of Norway and in other countries to increase generalizability. The results from future studies and our study could be extrapolated to inform stakeholders in home care services about the need for support when introducing digital health services in their populations. The results may not be generalizable to countries without home care services or with other access to digital devices and IT. On the other hand, the participants in the study represent people who are particularly in need of digital home monitoring because of the travel distances in rural areas, and the results may be important for stakeholders and leaders of health care services in Norway and internationally. Moreover, 62 people did not consent to participate. These people were older and a larger proportion were women when compared with the cohort of participants.

Conclusion

To our knowledge, this is the first study on eHealth literacy in people receiving home care, a population that currently is being offered and in the future will be offered digital technology. Understanding the eHealth literacy knowledge of this population is important to take advantage of digital technology. The technology readiness level varied in the cohort of Norwegian participants receiving home care, and not all people receiving home care services can be expected to benefit from digital health services. Participants in cluster 4 (n=16) had the lowest score in the eHLQ scales in the READHY instrument and should be offered nondigital services or would need extensive management support. The demographic differences between the 4 clusters may inform stakeholders about which older people need the most training and support to be able to take advantage of digital health care services. This study contributes to this important area with information about technology readiness in a population of older adults and with insights into how this is related to their perceived health and digital behaviors. We were not able to report the psychometric properties of the Norwegian version of the READHY instrument, and a larger validation study should be performed.

XSL•FO
RenderX

Acknowledgments

This research was funded by the European Union's Horizon 2020 Research and Innovation Program under grant 101016848 and the Canadian Institutes of Health Research.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Patient profiles within the 4 clusters (N=147).
[DOCX File , 21 KB - aging_v8i1e62936_app1.docx]

Multimedia Appendix 2
Results of linear regression analysis for the World Health Organization-Five Well-Being Index (WHO-5) (N=129).
[DOCX File , 39 KB - aging_v8i1e62936_app2.docx]

Multimedia Appendix 3
Results of linear regression analysis for the question "IT competence assessed by others" (N=131).
[DOCX File , 39 KB - aging_v8i1e62936_app3.docx]

References

1. Decade of Healthy Ageing. World Health Organization. URL: https://apps.who.int/gb/ebwha/pdf_files/EB146/B146_23-en.pdf [accessed 2025-01-20]
2. Chang M, Michelet M, Skirbekk V, Langballe EM, Hopstock LA, Sund ER, et al. Trends in the use of home care services among Norwegians 70+ and projections towards 2050: The HUNT study 1995-2017. Scand J Caring Sci 2023 Sep;37(3):752-765. [doi: 10.1111/scs.13158] [Medline: 36967552]
3. Madanian S, Nakarada-Kordic I, Reay S, Chetty T. Patients' perspectives on digital health tools. PEC Innov 2023 Dec;2:100171 [FREE Full text] [doi: 10.1016/j.pecinn.2023.100171] [Medline: 37384154]
4. Stoumpos AI, Kitsios F, Talias MA. Digital transformation in healthcare: technology acceptance and its applications. Int J Environ Res Public Health 2023 Feb 15;20(4):3407 [FREE Full text] [doi: 10.3390/ijerph20043407] [Medline: 36834105]
5. Hong P, Chen K, Chang Y, Cheng S, Chiang H. Effectiveness of theory-based health information technology interventions on coronary artery disease self-management behavior: a clinical randomized waitlist-controlled trial. J Nurs Scholarsh 2021 Jul;53(4):418-427 [FREE Full text] [doi: 10.1111/jnu.12661] [Medline: 33844425]
6. StatBank. Statistics Norway. URL: https://www.ssb.no/en/ [accessed 2025-01-20]
7. Bjønnes A, Midtbø T, Størset H, Ulven C. Befolkningens digitale kompetanse og deltakelse - Med et ekstra blikk på seniorer og ikke-sysselsatte. Direktoratet for høyere utdanning og kompetanse. 2021. URL: https://hkdir.no/rapporter-undersokelser-og-statistikk/befolkningens-digitale-kompetanse-og-deltakelse-2021 [accessed 2025-01-20]
8. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun 16;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
9. Kayser L, Rossen S, Karnoe A, Elsworth G, Vibe-Petersen J, Christensen JF, et al. Development of the multidimensional readiness and enablement index for health technology (READHY) tool to measure individuals' health technology readiness: initial testing in a cancer rehabilitation setting. J Med Internet Res 2019 Feb 12;21(2):e10377 [FREE Full text] [doi: 10.2196/10377] [Medline: 30747717]
10. Estrela M, Semedo G, Roque F, Ferreira PL, Herdeiro MT. Sociodemographic determinants of digital health literacy: A systematic review and meta-analysis. Int J Med Inform 2023 Sep;177:105124 [FREE Full text] [doi: 10.1016/j.ijmedinf.2023.105124] [Medline: 37329766]
11. Kayser L, Karnoe A, Furstrand D, Batterham R, Christensen KB, Elsworth G, et al. A multidimensional tool based on the eHealth literacy framework: development and initial validity testing of the eHealth Literacy Questionnaire (eHLQ). J Med Internet Res 2018 Feb 12;20(2):e36 [FREE Full text] [doi: 10.2196/jmir.8371] [Medline: 29434011]
12. Osborne RH, Elsworth GR, Whitfield K. The Health Education Impact Questionnaire (heiQ): an outcomes and evaluation measure for patient education and self-management interventions for people with chronic conditions. Patient Educ Couns 2007 May;66(2):192-201. [doi: 10.1016/j.pec.2006.12.002] [Medline: 17320338]
13. Osborne RH, Batterham RW, Elsworth GR, Hawkins M, Buchbinder R. The grounded psychometric development and initial validation of the Health Literacy Questionnaire (HLQ). BMC Public Health 2013 Jul 16;13:658 [FREE Full text] [doi: 10.1186/1471-2458-13-658] [Medline: 23855504]
14. Kikkenborg J, Magelund E, Riise MS, Kayser L, Terp R. Knowledge, skills, and experience with technology in relation to nutritional intake and physical activity among older adults at risk of falls: semistructured interview study. JMIR Hum Factors 2024 May 08;11:e52575 [FREE Full text] [doi: 10.2196/52575] [Medline: 38717810]

XSL•FO
RenderX

15. Rossen S, Kayser L, Vibe-Petersen J, Ried-Larsen M, Christensen JF. Technology in exercise-based cancer rehabilitation: a cross-sectional study of receptiveness and readiness for e-Health utilization in Danish cancer rehabilitation. Acta Oncol 2019 May;58(5):610-618. [doi: 10.1080/0284186X.2018.1562213] [Medline: 30698060]
16. Thorsen IK, Rossen S, Glümer C, Midtgaard J, Ried-Larsen M, Kayser L. Health technology readiness profiles among Danish individuals with type 2 diabetes: cross-sectional study. J Med Internet Res 2020 Sep 15;22(9):e21195 [FREE Full text] [doi: 10.2196/21195] [Medline: 32930669]
17. Nielsen AS, Hanna L, Larsen BF, Appel CW, Osborne RH, Kayser L. Readiness, acceptance and use of digital patient reported outcome in an outpatient clinic. Health Informatics J 2022 Jun 03;28(2):14604582221106000 [FREE Full text] [doi: 10.1177/14604582221106000] [Medline: 35658693]
18. Terp R, Kayser L, Lindhardt T. Older patients' competence, preferences, and attitudes toward digital technology use: explorative study. JMIR Hum Factors 2021 May 14;8(2):e27005 [FREE Full text] [doi: 10.2196/27005] [Medline: 33988512]
19. Wahl AK, Osborne RH, Langeland E, Wentzel-Larsen T, Mengshoel AM, Ribu L, et al. Making robust decisions about the impact of health education programs: Psychometric evaluation of the Health Education Impact Questionnaire (heiQ) in diverse patient groups in Norway. Patient Educ Couns 2016 Oct;99(10):1733-1738. [doi: 10.1016/j.pec.2016.05.001] [Medline: 27211224]
20. Wahl AK, Hermansen A, Osborne RH, Larsen MH. A validation study of the Norwegian version of the Health Literacy Questionnaire: A robust nine-dimension factor model. Scand J Public Health 2021 Jun;49(4):471-478 [FREE Full text] [doi: 10.1177/1403494820926428] [Medline: 32508258]
21. Terp R, Lindhardt T, Kayser L. Theory-driven development of an educative nutritional intervention (ENI) supporting older hospital patients to eat sufficiently, assisted by an eHealth solution: an intervention mapping approach. BMC Health Serv Res 2022 Nov 28;22(1):1435 [FREE Full text] [doi: 10.1186/s12913-022-08679-8] [Medline: 36443804]
22. Krag T, Jørgensen E, Phanareth K, Kayser L. Experiences with in-person and virtual health care services for people with chronic obstructive pulmonary disease: qualitative study. JMIR Rehabil Assist Technol 2023 Aug 14;10:e43237 [FREE Full text] [doi: 10.2196/43237] [Medline: 37578832]
23. Larsen MH, Strumse YS, Andersen MH, Borge CR, Wahl AK. Associations between disease education, self-management support, and health literacy in psoriasis. J Dermatolog Treat 2021 Sep;32(6):603-609 [FREE Full text] [doi: 10.1080/09546634.2019.1688233] [Medline: 31692398]
24. Hermansen A, Andersen MH, Borge CR, Dahl KG, Larsen MH, Lønning K, et al. Preliminary validity testing of the eHealth Literacy Questionnaire (eHLQ): a Confirmatory Factor Analysis (CFA) in Norwegian hospitalized patients. BMC Psychol 2023 Nov 23;11(1):409 [FREE Full text] [doi: 10.1186/s40359-023-01449-z] [Medline: 37996858]
25. Kayser L, Karnoe A, Duminski E, Jakobsen S, Terp R, Dansholm S, et al. Health professionals' eHealth literacy and system experience before and 3 months after the implementation of an electronic health record system: longitudinal study. JMIR Hum Factors 2022 Apr 29;9(2):e29780 [FREE Full text] [doi: 10.2196/29780] [Medline: 35486414]
26. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992 Jun;30(6):473-483. [Medline: 1593914]
27. Ward JH. Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association 1963 Mar;58(301):236-244. [doi: 10.1080/01621459.1963.10500845]
28. Forgy E. Cluster analysis of multivariate data fficiency versus interpretability of classifications. Biometrics 1965;21(3):768-769.
29. Bland JM, Altman DG. Cronbach's alpha. BMJ 1997 Feb 22;314(7080):572-572 [FREE Full text] [doi: 10.1136/bmj.314.7080.572] [Medline: 9055718]
30. De Moortel D, Palència L, Artazcoz L, Borrell C, Vanroelen C. Neo-Marxian social class inequalities in the mental well-being of employed men and women: the role of European welfare regimes. Soc Sci Med 2015 Mar;128:188-200. [doi: 10.1016/j.socscimed.2015.01.027] [Medline: 25621400]
31. Dreger S, Buck C, Bolte G. Material, psychosocial and sociodemographic determinants are associated with positive mental health in Europe: a cross-sectional study. BMJ Open 2014 May 28;4(5):e005095 [FREE Full text] [doi: 10.1136/bmjopen-2014-005095] [Medline: 24871540]
32. Lukaschek K, Vanajan A, Johar H, Weiland N, Ladwig K. "In the mood for ageing": determinants of subjective well-being in older men and women of the population-based KORA-Age study. BMC Geriatr 2017 Jun 16;17(1):126 [FREE Full text] [doi: 10.1186/s12877-017-0513-5] [Medline: 28622764]
33. Andersen MH, Hermansen A, Dahl KG, Lønning K, Meyer KB, Vidnes TK, et al. Profiles of health literacy and digital health literacy in clusters of hospitalised patients: a single-centre, cross-sectional study. BMJ Open 2024 May 20;14(5):e077440 [FREE Full text] [doi: 10.1136/bmjopen-2023-077440] [Medline: 38772592]
34. Dalmaijer ES, Nord CL, Astle DE. Statistical power for cluster analysis. BMC Bioinformatics 2022 May 31;23(1):205 [FREE Full text] [doi: 10.1186/s12859-022-04675-1] [Medline: 35641905]

Abbreviations

eHLQ: eHealth Literacy Questionnaire

XSL•FO
RenderX

heiQ: Health Education Impact Questionnaire
HLQ: Health Literacy Questionnaire
READHY: Readiness and Enablement Index for Health Technology
T2DM: type 2 diabetes mellitus
WHO-5: World Health Organization-Five Well-Being Index

Edited by J Wang; submitted 05.06.24; peer-reviewed by S Kujala, I de Siqueira Silva, L Donelle, X Huang; comments to author 04.07.24; revised version received 26.07.24; accepted 14.01.25; published 07.02.25.

Please cite as:
Bergh S, Benth JŠ, Høgset LD, Rydjord B, Kayser L
Assessment of Technology Readiness in Norwegian Older Adults With Long-Term Health Conditions Receiving Home Care Services: Cross-Sectional Questionnaire Study
JMIR Aging 2025;8:e62936
URL: https://aging.jmir.org/2025/1/e62936
doi:10.2196/62936
PMID:39918862

©Sverre Bergh, Jūratė Šaltytė Benth, Lisbeth Dyrendal Høgset, Britt Rydjord, Lars Kayser. Originally published in JMIR Aging (https://aging.jmir.org), 07.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Experiences of Older Mental Health Patients and Their Care Partners Using a Proxy Account to Access Open Notes: Qualitative Interview Study

Eva Meier-Diedrich[1,2], MSc; Tobias Esch[3], MD; Maria Hägglund[4,5], PhD; Martin Heinze[1,2,6], MD; Stefan Hochwarter[7], PhD; Justin Speck[2], BSc; Marie Wagener[2], MSc; Volker Dahling[1,2,8*], MD; Julian Schwarz[1,2,6*], MD

[1]Department of Psychiatry and Psychotherapy, Center for Mental Health, Immanuel Hospital Rüdersdorf, Brandenburg Medical School Theodor Fontane, Rüdersdorf, Germany

[2]Faculty of Health Sciences Brandenburg, Brandenburg Medical School Theodor Fontane, Neuruppin, Germany

[3]Institute for Integrative Health Care and Health Promotion, Faculty of Health, School of Medicine, Witten/Herdecke University, Witten, Germany

[4]Department of Women's and Children's Health, Uppsala University, Uppsala, Sweden

[5]Uppsala University Hospital, Uppsala, Sweden

[6]Center for Health Service Research Brandenburg, Brandenburg Medical School Theodor Fontane, Rüdersdorf, Germany

[7]Institute for Biomedicine and Health Sciences (HEALTH), Joanneum Research Forschungsgesellschaft m.b.H, Graz, Austria

[8]Research Group Geriatric Psychiatry Research, Brandenburg Medical School Theodor Fontane, Rüdersdorf, Germany

[*]these authors contributed equally

Corresponding Author:
Eva Meier-Diedrich, MSc
Department of Psychiatry and Psychotherapy, Center for Mental Health
Immanuel Hospital Rüdersdorf
Brandenburg Medical School Theodor Fontane
Seebad 82/83
Rüdersdorf, 15562
Germany
Phone: 49 33638 83 501
Email: Eva.Meier-Diedrich@mhb-fontane.de

Abstract

Background: Older patients with serious mental illnesses such as cognitive disorders often rely on family members or spouses (care partners [CPs]) to meet their health care needs. CPs frequently lack essential information to fully understand the patients' illnesses and effectively support their treatment. Open Notes provide patients with digital access to their health care professionals' clinical notes and are associated with many positive outcomes, such as increased adherence and empowerment. However, older patients who use Open Notes may encounter use barriers such as limited digital literacy. Recent developments allow CPs to access Open Notes (proxy access) and receive valuable information, which holds significant potential for improving the care of older patients.

Objective: This study explored the experiences, barriers, and opportunities of older mental health patients and their CPs related to using Open Notes. Furthermore, influencing factors and interdependencies were identified.

Methods: Older patients (n=10) and their CPs (n=10) were provided with web-based proxy access to clinical documentation through a web-based patient portal. In-depth qualitative interviews (N=20) were conducted to explore experiences with this access. Data analysis was conducted in accordance with the constructivist grounded theory approach.

Results: The prerequisites for using Open Notes with proxy access were sufficient digital literacy on the part of the patient or CP, as well as the establishment of a trusting relationship between patients and CPs. Access to Open Notes enabled patients and CPs to gain a deeper understanding of the illness and its treatment while also facilitating enhanced contact with health care professionals. This resulted in greater involvement in the treatment process but may also prompt changes in relationship dynamics—CPs are better equipped to support patients in their health care but may also tend to monitor or control them through Open Notes. As a result, the introduction of Open Notes was accompanied by mixed feelings.

XSL•FO
RenderX

Conclusions: It is of utmost importance to provide older patients with comprehensive access to Open Notes to preserve their health autonomy. However, the involvement of CPs through proxy access is of great value in improving the care of older patients, especially those with cognitive impairments.

(JMIR Aging 2025;8:e66690) doi:10.2196/66690

KEYWORDS

psychiatry; eHealth; mental health; digital literacy; older patients; older adult; care partner; proxy access; open record access; Open Notes; patient portal; artificial intelligence; AI

Introduction

Background

The ongoing digitalization throughout society, coupled with demographic transformations, presents both opportunities and challenges for the health care sector. The global population is living longer, and in particular in industrialized countries, fertility rates are declining, which is significantly altering the age composition of the population [1]. Concurrently, digitalization is gradually permeating all areas of society, including the mental health care sector, thereby providing patients with new opportunities to participate in their treatments [2]. The European Health Data Space proposal follows this development and aims to improve patients' access and control over their personal electronic health data [3]. One such opportunity for patient participation is through online record access (ORA), which enables patients to view their electronic health records (EHRs) via web-based patient portals. When patients are able to also access their health care professionals' (HCPs) notes via ORA, this is referred to as Open Notes [4-6]. In the United States and some Scandinavian countries, Open Notes are already an established practice in health care [3,7]. Patients in the United Kingdom have the option of accessing the clinical notes of their general practitioners [8]. In contrast, in other countries such as Germany, Open Notes have yet to be integrated into the health care system [9,10]. However, current legislation requires all hospitals in Germany to implement patient portals. Therefore, in the near future, it will be technically possible to offer Open Notes [11]. Access to clinical notes and the associated transparency can enhance patient empowerment and engagement, improve medication adherence, and bolster disease management and awareness [12-15]. Studies indicate that vulnerable groups, including older patients, particularly benefit from Open Notes [16]. Furthermore, it can enhance communication between (older) patients and their HCPs, thereby strengthening trust in the treatment process [17,18]. While many older patients express interest in and intent to use digital health technologies, few actually use them [19]. Because older patients are not considered digital natives, they often face several challenges when using digital health technologies, including limited digital health literacy, usability issues, and heightened concerns about data security [19-21]. In addition, older patients are increasingly affected by cognitive deficits (eg, dementia), which further complicates the development of digital health literacy [22]. These barriers lead to a decrease in the use of web-based health services such as ORA with advancing age, especially without (human) guidance [23]. Therefore, it is imperative that digital health literacy is taken into account in the design of patient portals and EHRs to ensure their accessibility and inclusivity for older users [24,25].

In general, a significant proportion of older patients rely on relatives such as family members, partners, or friends (referred to as care partners [CPs]) to meet their health care needs [24]. This is also true for digital health services [26]. Receiving support from CPs has a positive impact on patients' quality of life, quality of care, and health resource use. In addition, patients and CPs would like to have greater involvement in medical care, but this has not been adequately supported by the health care system [24,27]. CPs often lack essential information about the health status and treatment planning of older patients, which can significantly complicate care [24]. With proxy access, Open Notes allow relatives to read medical treatment documentation with the patient's consent [28]. Giving CPs access to Open Notes provides them with important information and facilitates their care. In addition, studies show that the involvement of CPs can also increase engagement in the treatment of (older) patients using ORA [29]. Current research confirms that such access is desired by CPs [22,30] but acceptance is still limited [26]. ORA and Open Notes will become increasingly important in health care, and the associated opportunities should be available to all patients [22,24]. Therefore, it is essential to gain a deeper understanding of the use patterns of older patients, the role of CPs, and the barriers and opportunities associated with them.

Objectives

To date, only a limited number of studies have examined the use of ORA by older patients in the context of proxy access [24,31,32]. Specifically, in the area of mental health, the authors are currently unaware of any studies on this topic [33,34]. In light of the current state of research, the purpose of this study was to explore the experiences, preferences, and needs of older mental health patients and their CPs, as well as the barriers and opportunities related to using Open Notes. In addition, this study aimed to provide recommendations for best practice in this area and sought to identify the factors influencing the impact of CP access to Open Notes.

Methods

Study Design

The *Piloting and evaluation of a participatory patient-accessible electronic health record for geriatric psychiatric patients and their care partners* (PEP.AGE) study is part of the *Piloting and evaluation of a participatory patient-accessible electronic health record in Psychiatry and Somatics* (PEPPPSY) project (2021-2026) [35,36]. In the PEPPPSY project, patients are provided with access to their HCPs' treatment and progress

XSL•FO
RenderX

notes via a dedicated patient portal [37]. Furthermore, the development and implementation of the patient portal are being examined from the dual perspectives of both patients and HCPs. The PEP.AGE study broadens the scope of the PEPPPSY target population by including not only the perspective of older patients but also that of their CPs. Given the exploratory nature of this study, a qualitative design was chosen to ensure a comprehensive and thorough examination of the use of Open Notes by older patients and their CPs.

Ethical Considerations

This study was approved by the ethics committee of Brandenburg Medical School Theodor Fontane (E-01-20210727) and registered with the German Clinical Trials Register (DRKS00030188). Participants were informed of the study content and procedures both verbally and in writing. Informed written consent was then obtained from all participants. All participants had the right to withdraw from the study at any time without any adverse consequences. All data were anonymized. Participants received a compensation of €40 (US $41.39) for their participation.

PEPPPSY App

The patient portal pilot, called PEPPPSY, was initiated as part of a research collaboration between the Norwegian University of Science and Technology and Brandenburg Medical School Theodor Fontane. The portal was developed through an ongoing iterative and participatory process [35,36]. In addition to accessing clinical notes, patients and CPs can respond to HCPs' entries with comments. HCPs are then notified of these comments and can respond to patients (or CPs) within the same thread. In the current second phase of the project, the pilot has been expanded to include access for CPs (proxy access), which will increase the accessibility and utility of the portal for a broader patient population.

Study Setting

This study was conducted in 2 psychiatric outpatient clinics (Rüdersdorf and Strausberg) of the Immanuel Hospital Rüdersdorf in the state of Brandenburg, Germany. Psychiatric outpatient clinics are specialized facilities that provide psychiatric care to patients with severe mental illness. These patients typically require comprehensive and multidisciplinary psychiatric treatment and often lack access to adequate care in other outpatient settings (such as psychiatric or general medical practices) due to the severity or chronicity of their psychiatric conditions.

Recruitment

Eligible participants were enrolled in the PEP.AGE study from June 2023 to January 2024. Participating patients had to be aged ≥60 years, receive treatment at 1 of the 2 designated sites, and be able to provide informed consent. Patients with risk factors such as self-harm or harm to others and severe cognitive impairment were excluded from the study. Participating CPs had to be adults (aged ≥18 years) and able to provide informed consent.

Data Collection

At enrollment, sociodemographic data were collected from both patients and their CPs. Patients and CPs were then introduced to the use of the patient portal by their HCP or a member of the study team. They were provided with a comprehensive, user-friendly manual and the option of one-on-one assistance to set up their accounts and learn how to use Open Notes step by step. Subsequently, the older patients and their CPs (with patient consent) were given access to the patient portal and the HCPs' clinical notes. At the beginning of the intervention phase, participants were randomly contacted to identify and address any barriers to use. At the end of the 3-month intervention phase, semistructured interviews were conducted with patients and CPs using previously developed interview guides to gain deeper insights into their actual experiences with Open Notes (Multimedia Appendix 1). The interviews lasted between 20 and 30 minutes each. Throughout the study (onboarding phase, intervention phase, and interview phase), the study team kept field notes documenting observations and contextual information [38].

Data Analysis

The qualitative interviews were audio recorded, pseudonymized, transcribed, and analyzed by 2 researchers with the computer-assisted analysis software MAXQDA (VERBI GmbH) using the constructivist grounded theory by Charmaz [39]. The selected analytical approach was appropriate to the research subject as this exploratory study aimed to iteratively develop theoretical concepts from the data. In accordance with the approach by Charmaz [39], the data analysis was conducted continuously, commencing with the earliest data gathering (initial interview). The interviews were initially coded line by line to facilitate the conceptualization of ideas and the development of preliminary codes. Subsequently, focused coding was conducted, whereby the most significant and frequent codes were identified, sorted, and synthesized into overarching categories. Following this, relationships between the categories were identified and connected into coherent theoretical concepts (theoretical coding). On the basis of the developed concepts and emerging theory, the research team returned to the field and gathered additional data on specific themes until theoretical saturation was achieved. For quality assurance purposes, the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist was used (Multimedia Appendix 2).

Results

Sociodemographic Data

A total of 10 patients and 10 CPs were interviewed via telephone (by EMD or MW), and their complete sociodemographic data are shown in Table 1. In total, 5 dyads (each consisting of a patient and their respective CP), as well as 5 independent patients and 5 independent CPs, were interviewed. All participants self-identified as White individuals, were born in Germany, and spoke German as their native language. The age of the patients ranged from 62 to 81 years, with a mean of 71.60 (SD 6.43) years. A total of 70% (7/10) of the patients identified as female, and 30% (3/10) identified as male. Most patients

XSL•FO
RenderX

(8/10, 80%) were retired, whereas a minority (2/10, 20%) were still employed. All CPs (10/10, 100%) were family members of the patients (mainly spouses or children). The ages of the CPs were more diverse, ranging from 45 to 81 (mean 61.20, SD 11.02) years. This was due to the participation of spouses (5/10, 50%), children (4/10, 40%), and other family members (1/10, 10%) of the patients as CPs. In total, 40% (4/10) of the CPs were already retired, whereas 60% (6/10) were still employed part time or full time. Before the start of the PEP.AGE study, the vast majority (8/10, 80%) of the older patients had already given their relatives access to their health information. Patients reported medical discussions with their HCPs to their CPs, shared medical correspondence and medication schedules with them, or were accompanied by CPs to medical appointments.

Table 1. Characteristics of the patient–care partner dyads.

Dyad number	Patients			Care partners		
	Age (y)	Sex	Diagnosis (*ICD-10*[a] code)	Age (y)	Relationship to patient	Employment status
1	67	Female	Mild cognitive disorder (F06.7)	62	Spouse	Part time
2	75	Female	Dementia in Alzheimer disease with late onset (F00.1)	51	Child	Full time
3	80	Male	Dementia in Alzheimer disease with late onset (F00.1)	69	Spouse	Retired
4	81	Female	Severe depressive episode without psychotic symptoms (F32.2)	45	Child	Full time
5	67	Female	Social phobias (F40.1)	69	Spouse	Retired
6[b]	69	Female	Recurrent depressive disorder, current episode severe without psychotic symptoms (F33.2)	52	Child	Retired
7[b]	62	Male	Generalized anxiety disorder (F41.1)	53	Child	Full time
8[b]	68	Male	Recurrent depressive disorder, current episode severe without psychotic symptoms (F33.2)	81	Spouse	Retired
9[b]	78	Female	Bipolar affective disorder, current episode mild or moderate depression (F31.3)	69	Spouse	Retired
10[b]	69	Female	Recurrent depressive disorder, current episode severe without psychotic symptoms (F33.2)	61	Other family member	Full time

[a]ICD-10: International Classification of Diseases, 10th Revision.

[b]No dyad; independent care partners and patients.

The patients supported by the independent CPs (CPs 6-10) were aged between 71 and 86 (mean 80.8, SD 6.099) years and retired, and they self-identified as White individuals. Patients were being treated for the following main diagnoses: dementia in Alzheimer disease with late onset (*International Classification of Diseases, 10th Revision* [*ICD-10*], code F00.1); mild cognitive disorder (*ICD-10* code F06.7); recurrent depressive disorder, current episode severe without psychotic symptoms (*ICD-10* code F33.2); and mixed anxiety and depressive disorder (*ICD-10* code F41.2).

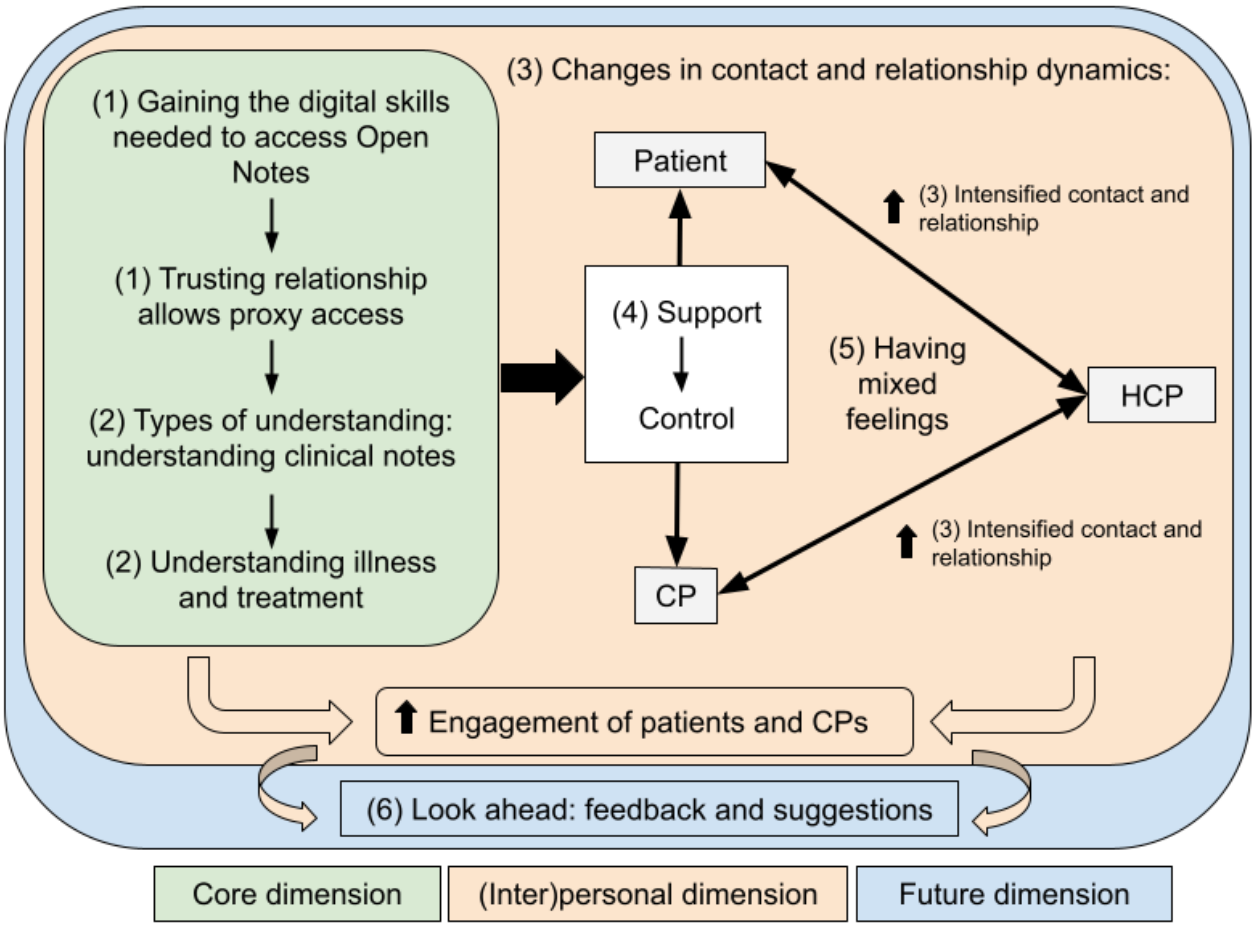
Qualitative Findings

Overview

The results revealed 3 (partially interrelated) dimensions associated with Open Notes with proxy access when used by older patients and their CPs. These dimensions and their interactions are summarized in Figure 1 and described in detail in the following sections. The green core dimension in Figure 1 provides the foundation for the use of Open Notes and proxy access (eg, digital skills and literacy, trust, and understanding of note content). The red (inter)personal dimension encompasses the impact that Open Notes can have on the relationships among patients, CPs, and HCPs. Finally, the blue future dimension offers ideas and recommendations for the further development of Open Notes. Quotations from CPs and patients are identified using the IDs *CP* and *PAT*, respectively.

XSL•FO
RenderX

Figure 1. Interrelations of the qualitative categories and dimensions related to the use of Open Notes with proxy access by older patients and their care partners (CPs). Arrows pointing upward indicate increase or enhancement. HCP: health care professional.



Gaining the Digital Skills Needed to Access Open Notes

The interviews and field observations revealed that the digital health literacy of the older patients and their CPs varied widely—both between dyads (patient-care-partner pairs) and across participant groups. While some dyads required direct assistance from study team members to activate, log into, and use the patient portal, others required no assistance at all. Younger CPs generally found it much easier to navigate the patient portal than older patients. This variability was also evident in the interviews, with approximately half (4/10, 40%) of the patients and most of the CPs (6/10, 60%) reporting that using Open Notes was challenging or even beyond the digital literacy of the patients, as illustrated by the following statements:

> *It [technical difficulties] could very well be because I'm no longer able to do things like that. [I had] Two strokes and then the nerve disease.* [PAT_0420]

> *New technologies are often a challenge for older people. It takes a lot of patience and support from us younger ones, but it's doable if you stick with it and show them they can do it.* [CP_0213]

As the previous quotes illustrate, a significant number of patients (5/10, 50%) relied on the assistance of their CPs to use the patient portal and Open Notes. At the same time, 20% (2/10) of the CPs themselves indicated that using digital health technologies posed a (manageable) challenge to their digital literacy:

> *I had to overcome a few technical hurdles, but with time and some support, I managed to use the application.* [CP_0213]

The 2-factor authentication log-in process proved to be particularly challenging, if not insurmountable, for some patients and CPs. Some participants were able to overcome this barrier on their own or with assistance (from CPs or the study team). However, others became so frustrated that they stopped using the patient portal altogether. CPs reported the following:

> *Well, there's this two-step login process, where you need the SMS PIN, and then—since I mostly used it on mobile—you have to fiddle around a bit to find where the access to the records is, but ultimately it was okay.* [CP_0209]

> *I had a question once because the access didn't work at the beginning, but it was resolved relatively quickly. They sent me a new one, and I was able to use it. Okay, but I don't know exactly if it was my fault or if it was just issues with the program.* [CP_0409]

Although most participants reported using and perceiving benefits from the patient portal, 30% (3/10) of the CPs and 10% (1/10) of the patients indicated that they did not access the patient portal in their daily lives. This was due to a lack of

XSL•FO
RenderX

perceived need to review the information given their regular contact with the HCP and the absence of a crisis situation that would have made accessing the records more relevant:

> *For us, it [Open Notes] doesn't have any everyday use. The idea that there's the possibility to look up and comment on disputed questions is great. But in the six months we've been participating, there hasn't been a situation that required us to intervene or do anything.* [CP_0211]

Most patients (8/10, 80%) agreed that a solid foundation of trust is a prerequisite for disclosing sensitive clinical information to family members. Some CPs (3/10, 30%) expressed a similar view. An open and honest conversation about the advantages and disadvantages regarding the patients' privacy seemed to be particularly important before using Open Notes with a proxy access, as the following quote shows:

> *We had discussed beforehand what this is and what it means. Otherwise, I don't think she [CP] would have agreed to it. When everything is disclosed, you have to be willing to accept that. Some might say, "Oh no, I don't want that, it's too private," depending on who it is and the relationship involved.* [PAT_0419]

Types of Understanding

Overview

The implementation of Open Notes involves several types of understanding. First, it is essential for patients and CPs to comprehend the content of the clinical note, which requires that the notes be written in patient-friendly language. In addition, by reading the clinical notes, older patients and their CPs were able to gain a better understanding of the illness and its treatment. This allowed them to adequately prepare for medical appointments, reducing anxiety and facilitating understanding during treatment sessions.

Understanding Clinical Notes

Most interviewed patients (8/10, 80%) and half (5/10, 50%) of the CPs agreed that the documentation was particularly understandable when it was composed in a manner accessible to patients without a lot of medical or technical jargon:

> *It [Open Notes] was understandable. Without any medical jargon, everything was fine. The way I described it, he [the HCP] wrote it down, more or less in my own words. It was expressed a bit better, but still in normal, understandable terms, I would say.* [PAT_0411]

> *Yes, it was understandable. Of course, there are always medical terms that might be unfamiliar to a layperson at first. But then you remember the conversation and can figure out what it was about and what was meant. So far, I can't say that I didn't understand anything. It was all very understandable.* [CP_0201]

The previous quotes show that, for some users, the information in the Open Notes alone is not sufficient; rather, understanding is built by combining the knowledge gained from the conversation with HCPs during in-person medical appointments and the information provided in the Open Notes.

Some CPs accompanied patients to their medical appointments and were able to recall the content of the conversations. However, for CPs who were unable to attend appointments, the clinical notes seemed to be easy to understand, as the following quote illustrates:

> *Yes, those were his notes. Brief and to the point. Of course, he didn't write long texts, but at least he documented briefly what was discussed, how the medications are, and what the plan is going forward. He wrote it down in a way that was understandable for everyone.* [CP_0416]

Understanding the Illness and Its Treatment

Some patients (3/10, 30%), but especially CPs (6/10, 60%), reported that access to patient-friendly clinical notes provided them with a more complete understanding and awareness of the patient's illness and treatment:

> *I now have a much better overview of the entire treatment process and my mother's current health condition. This makes it easier to make informed decisions and plan the next steps.* [CP_210]

This improved understanding of the illness appeared to serve as a foundation for subsequent developments, including the increased involvement of CPs in the treatment process and the provision of support. Because treatment appointments often leave little time for questions or repetition, patients (5/10, 50%) and CPs (3/10, 30%) found open-ended notes to be a valuable reminder, allowing them to prepare for and follow up on appointments more easily. Participants found it beneficial to have a written record of what was discussed that they could review at their own pace, allowing them to process the information in a way that best met their individual needs:

> *Remembering and understanding important details and conversation points better. This was especially useful for preparing for appointments and following up on recommendations.* [CP_0213]

> *Because for me, it's easier when I see something in writing, read it, and then respond or share my own experiences. This back-and-forth, this exchange with the doctors and staff, it's easier for me in writing than sitting in front of the doctor who might not have much time.* [PAT_0402]

> *If we have an appointment, we review the last one together, summarize the key points, and build the new medical appointment on that.* [CP_0201]

As the preceding quotes illustrate, one patient noted that reviewing the documentation allowed her to better assess and understand her own treatment progress. Another patient reported that he used the clinical notes to confirm that his HCP had understood him correctly during the visit. This gave him the opportunity to address any potential misunderstandings:

> *Exactly, it allows you to see for yourself that everything was conveyed clearly. When you only attend the doctor's visit and then leave, you forget*

XSL•FO
RenderX

> *half of it anyway. It was much better for me to be able to read it again and confirm that I was truly understood.* [PAT_0204]

Changes in Contact and Relationship Dynamics

Nevertheless, many patients (7/10, 70%) and CPs (5/10, 50%) described Open Notes as facilitating communication with HCPs in a variety of ways. Open Notes, with the opportunity to comment on notes and send messages back and forth, made contact with HCPs faster, more direct, more efficient, more accessible, and more frequent. As a result, participants reported an increased sense of involvement in the treatment process and improved collaboration with HCPs:

> *You can communicate quickly. That is important. It's nice, it's accessible.* [CP_0409]

> *Yes, the contact became more frequent, and I felt more involved in the treatment. It was a very positive effect that improved collaboration.* [CP_0210]

In addition, Open Notes allowed participants to contact HCPs outside of office hours, which was particularly convenient for full-time CPs. All participants were aware that this was asynchronous communication and that responses would only be made during the HCPs' working hours. One patient said the following:

> *You can exchange messages, even on the quick...So having a direct line to the doctor [through Open Notes], without having to call during office hours, is relatively quick. You get a prompt response from the doctor to what you write.* [PAT_0204]

In some cases, communication via Open Notes even replaced telephone contact. This was seen as a relief by some participants (1/10, 10% of patients and 4/10, 40% of CPs) as HCPs were often difficult to reach by phone. In this context, one CP highlighted the portal's communication and commenting function as a valuable tool for patients with mental health conditions who may find phone calls challenging:

> *Interviewer: Okay. And compared to calling, did you feel that you could reach your healthcare provider better or faster through the patient portal than by phone or other means?...Patient: Definitely much better, because calling is always tricky. If they're in treatment, you can't reach them. But this way, they responded when they had time, and everything was handled very quickly, so it was totally fine.* [PAT_0204]

> *As I would say, when patients use it themselves, and I am not a patient, but patients generally have underlying issues, often psychological, which make it difficult for them to communicate. Yes, reaching for the phone is challenging, going to the doctor is difficult. But maybe writing is somewhat less personal and might be easier. And it can be done at night or at an inconvenient time without feeling guilty, so I can imagine this is definitely a good option that could continue to be used.* [CP_0409]

Moreover, a slight shift in the relationship dynamics between patients and their CPs was noted with the implementation of Open Notes with proxy access. Now that CPs had access to the clinical notes, some patients (4/10, 40%) felt that their CPs were more understanding of their mental illness. This suggests a developmental process on the part of the CPs initiated by reading the shared notes:

> *Well, my stepson initially had problems because he couldn't imagine it when I said that it's still hard for me to take the bus alone. I get such a racing heart...But as I said, my stepson couldn't really understand it. Maybe you can't fully understand if you are healthy. But he has learned to understand it.* [PAT_0224]

From Support to Control

There was considerable variability in the level of digital health literacy among the older patients (see the *Gaining the Digital Skills Needed to Access Open Notes* section). In particular, patients in the oldest age group and those with cognitive impairments relied heavily on their CPs to help them navigate the digital patient portal. Some patients delegated responsibility for managing their health information in the portal to their CPs alone, whereas others sought to collaborate with their CPs in reading and understanding clinical notes:

> *My mother is now 82, and at that age, she's not likely to engage with apps or registration issues. If anything, I managed it for her or we discussed it.* [CP_0419]

> *I can open this page and then we can read it together. Or I explain to him what my concerns were. But independently, no longer.* [CP_0201]

In contrast, a younger and more digitally literate patient reported that she was able to access her Open Notes independently and only sought assistance from her CP when she encountered issues:

> *Well, I would first read it [Open Notes] on my own because I only have my daughter as a relative. And she has her own problems at the moment, so I don't really need her help unless I have issues. I'll handle it myself first.* [PAT_0402]

The ability to access clinical notes enabled many CPs (6/10, 60%) to gain a deeper understanding of the illness and treatment of the family member with a severe mental health issue. This increased their confidence in providing effective care and managing the illness, allowing them to better support the treatment (such as preparing for medical appointments and adhering to medication plans):

> *It [Open Notes] significantly improved my understanding. I could better follow the treatment processes and medical decisions. This helped me support my mother better and make informed decisions.* [CP_210]

> *It [Open Notes] was sometimes difficult, especially when the reports were not positive. But it helped me be better prepared and respond quickly if something was wrong.* [CP_0213]

XSL•FO
RenderX

As previously indicated (see the *Changes in Contact and Relationship Dynamics* section), this heightened level of involvement and responsibility, in addition to the improved information flow, was partially attributable to more intensive contact with HCPs, as evidenced by the following quote:

> *Yes, the contact became more intense, and I felt more included in the treatment. It was a very positive effect that improved the collaboration.* [CP_0210]

A few CPs (3/10, 30%) reported that Open Notes enabled them to provide more effective support from a greater distance (eg, from another city or country). Furthermore, CPs observed that reading the clinical notes reduced the need to accompany patients to medical appointments, thereby enhancing autonomy for both CPs and patients:

> *This allowed me to monitor their health data and ensure they received the right care even when I couldn't be with her. It gave me a sense of security to always be informed.* [CP_0210]

In addition to increased involvement, responsibility, and support, some CPs (3/10, 30%) also demonstrated a tendency to monitor or control patients through Open Notes. They compared the patient-reported information from medical appointments with the written information to assess the veracity of the reports and identify any potential omissions. This monitoring held the potential for conflict, but in one case, it also led to a more open and honest exchange between a patient and a CP regarding their inner motivations (eg, the withholding of information due to feelings of embarrassment, fear of disempowerment, or memory issues) and, thereby, enhanced mutual understanding:

> *It [Open Notes] allowed me to access all relevant information and better monitor my mother's health. It helped us be better informed and respond more quickly to changes.* [CP_0210]
>
> *There was a moment of surprise when she didn't mention something or had forgotten, but it was actually helpful because it led to a discussion where I could address it. She was honest, and we could discuss things in more detail or I could suggest she pay more attention to certain aspects. So, it wasn't a bad thing; it facilitated further discussion.* [CP_0416]

Having Mixed Feelings

Both CPs (7/10, 70%) and patients (5/10, 50%) provided insights into their emotional perceptions of the Open Notes. Notably, both patients and CPs reported a similar range of emotional experiences, including both positive and negative feelings. Both groups reported feelings of emotional distress associated with reading about deteriorating health or lack of treatment success. One patient even described experiencing persistent worrying thoughts. Some patients (4/10, 40%) also expressed concern that their CPs might experience distress as a result of reading the notes:

> *Well, it's a bit burdensome, I would say, maybe. When something new comes up and then the success doesn't happen.* [PAT_0402]
>
> *It was a mix of relief and concern. Relief because I was informed, and concern when the information wasn't positive. But overall, it helped me to be better prepared.* [CP_0210]
>
> *I am concerned that my relative may be emotionally distressed by reading the entries [Open Notes].* [PAT_0207]

Both CPs and patients noted that Open Notes provided a sense of security regarding the illness and its treatment. This sense of security was derived from 2 sources: first, the ability to access treatment information and, second, the knowledge that this information has been validated by experts:

> *No, for me it's more like the lack of knowledge is stressful. When you have an informed status, you can handle it better.* [CP_0211]
>
> *Yes, actually good, because I know it comes from a competent source and not just from random internet readings where every third person says something different, and so on. So, for me, it's reliable information. Definitely, knowing without having to worry about whether it's true or not or maybe or something, so that's more reassuring for me.* [PAT_0402]

One patient found it motivating and encouraging to read her HCP's notes. In addition, this patient was particularly motivated by proxy access and the fact that her CP also read the notes, which led her to engage in more self-care. This particular finding suggests that access to clinical notes by CPs may also impact treatment outcomes or patient recovery on a personal level beyond the increased involvement of well-informed CPs in the patients' health management:

> *But when you read the family's comments, like, "Hey, you've been letting yourself go lately," or "You seem unmotivated," it motivates you. You realize they are right; there's no reason to just hang around or whatever.* [PAT_0411]

In contrast, for other patients (2/10, 20%) and CPs (2/10, 20%), the clinical notes were less emotionally significant and were perceived more neutrally, as illustrated by the following quote:

> *You perceive it relatively neutrally. You don't get super happy or deeply depressed.* [PAT_0204]

Looking Ahead: Suggestions and Feedback

Several patients and CPs provided feedback on potential modifications to the patient portal that could improve its usability. Typically, suggestions focused on modifying or enhancing existing features within the portal. For instance, 4 participants (n=3, 30% patients and n=1, 10% CPs) expressed a desire for a read receipt feature to confirm when HCPs had received and read their messages. In addition, 20% (2/10) of the patients proposed that they be notified when a response from their HCP had been submitted. This notification feature had already been incorporated into the system and could be enabled by the user, yet these patients were unaware of its availability:

> *Something like that, just like with emails where you can send a confirmation of receipt or read receipt,*

XSL•FO
RenderX

> *so you know it's been received and opened. Sometimes, that's all you need.* [CP_0416]
>
> *And if the HCP has written something, it would be nice to get an email notification so that I know there's a message there. If it's out of the ordinary and you don't check it every day, you might not see it for a few days.* [PAT_0204]

Moreover, one patient expressed a desire for the portal's features to be more appealing and engaging for older patients, with the goal of tailoring the portal's design to the needs of this patient group (eg, by encouraging them to write comments). Both improved guidance and an optimized design were requested by the participants and could facilitate greater accessibility and appeal for the target audience:

> *A little guidance, maybe. Okay, that we start here, with a specific topic being set. I need to know, what should I write?* [CP_0402]
>
> *But the commenting function should be designed in a way that makes you want to use it, that makes you feel like speaking up.* [PAT_0420]

Other participants expressed a preference for integrating additional features into the patient portal beyond simply reading clinical documentation. In total, 10% (1/10) of the patients and 20% (2/10) of the CPs suggested enriching the patient portal with more psychoeducational information about the illness and integrating some type of psychoeducational lexicon or psychiatric frequently asked questions into the patient portal:

> *And I would certainly wish for a way to learn more about the illness, about behavioral strategies, options for the CP, but also for the patient. So you don't have to Google and look for information elsewhere. If you're already in the psychiatric system, maybe you could listen to more. Do you understand? That on this platform, on this level, you could already have specific questions answered.* [CP_0201]

In fact, one CP expanded the original scope of Open Notes by using the commenting feature to document important developments in her mother's health. She used this primarily as a personal reminder while also indirectly facilitating transparency and understanding of progress for HCPs by posting it on the patient portal:

> *At the beginning, I would write down things that I noticed in my daily life with my mom, as a personal reminder. It was helpful to have these notes ready for the next appointment as preparatory points. I definitely find it useful for that.* [CP_0409]

Discussion

Synthesis of the Findings

The results highlight both opportunities and challenges associated with using Open Notes for older patients and their CPs. In addition, the key drivers of proxy access were identified, and their interdependencies were highlighted. Our results show that older patients and their (sometimes older) CPs must first gain (proxy) access to the patient portal to use and benefit from Open Notes. This requires sufficient digital literacy and mutual trust between patients and CPs. In our study, many older patients needed support from their CPs to navigate digital health services. Once access to open records was established, both parties reported feeling more informed about the illness and its treatment and more in touch with HCPs. These 2 factors led to increased health literacy, engagement, and involvement for both patients and CPs. In line with this, our results suggest that access to Open Notes enables CPs to better support patients in their (digital) health management. However, there was also evidence that CPs used Open Notes to control patients, which could lead to conflicts. Finally, recommendations for further developments and feedback emerged.

Our findings show that it is particularly valuable to allow patients, CPs, and HCPs to digitally engage with Open Notes via a comment function, allowing the stakeholders to directly communicate by sending asynchronous messages. This interactivity of the test environment (*PEPPPSY*) in which our study was conducted was frequently used and highly valued by patients and CPs. In particular, it appeared to contribute to the health literacy of patients and CPs by allowing them to ask questions about the content of the notes or the treatment in general. Communication via Open Notes is not a classic feature of Open Notes, nor is it simply a secure messaging function as the digital interaction is not separate from the Open Note itself. This demonstrates that Open Notes serve multiple purposes (such as providing information, facilitating contact, and offering reassurance) depending on the level of interactivity available [40].

Ensuring Accessibility for Older Patients

As evidenced by previous research and observed in our study as well, older patients (and their CPs) predominantly use Open Notes as a memory aid, benefiting from this tool to prepare for or recap medical appointments. This provides both patients and CPs with an increased sense of security in their treatment processes. The positive effects of the implementation of Open Notes with a proxy access shown in this study—such as enhanced patient empowerment and engagement, increased CP involvement, and improved health management—align with those found in previous research [30,41,42]. Nevertheless, for these advantages to be fully realized, the initial challenge must be addressed: ensuring that older patients and their occasionally also older CPs have convenient access to the patient portal. This seems particularly relevant as older adults show interest in using patient portals yet the existence of numerous barriers hinders their ability to do so [19]. While CPs can indeed play a crucial role in compensating for the patients' lack of digital health literacy—as observed in our study—it is equally important to encourage and enable older patients (with sufficient cognitive abilities) to independently access their health information. Older adults are often apprehensive or skeptical about digital health tools, so addressing these concerns is essential [43]. Furthermore, it is important to ensure that older patients are able to comprehend the content of the Open Notes. Consequently, the clinical documentation must be written in patient-friendly language, which is not always the case in clinical practice [6].

XSL•FO
RenderX

On the basis of the study results, it seems important to provide older patients and their CPs with a clear and detailed explanation of the available features (such as the commenting feature and the opt-in notification feature) before they use the patient portal and Open Notes. It is imperative that patient portal interfaces are designed in a manner that is accessible to all age groups and that the technical requirements are kept as user-friendly as possible. For instance, alternative methods of 2-factor authentication should be explored as requiring users to use 2 devices (eg, a phone and a computer) simultaneously can be overwhelming and frustrating and may, ultimately, result in older patients giving up on using the patient portal and Open Notes. At the same time, increased usability must be compatible with high-level data security requirements. Furthermore, the design of patient portal interfaces should adhere to fundamental age-specific design principles, including the use of appropriate fonts, color choices, and audio alternatives and the minimization of text entry requirements [44]. Providing users with the option to select either a *standard* or *older age–accessible* interface design when accessing the patient portal could prove advantageous as it would enable users to customize their experience to align with their specific requirements. Nevertheless, it seems unlikely that merely modifying the interface will be sufficient to significantly increase the adoption of patient portals and Open Notes among older patients, and therefore, a more comprehensive approach is needed.

First, older patients must be made aware of the availability of patient portals (and the possibility of setting up a proxy access) through comprehensive and targeted informational campaigns [45]. Second, older patients need to be encouraged to use patient portals through the aforementioned campaigns and, more importantly, through their general practitioners and other HCPs [46]. As highlighted in the interviews and supported by findings of other studies, human guidance is essential for older patients to use the full range of features available on patient portals [47]. This responsibility should not be borne solely by CPs, particularly given that not all older patients have access to a digitally literate CP [48].

In light of the ongoing digitalization of the health care system, it may be worthwhile to consider the introduction of an institutionalized role dedicated to this task. In the United States, the role of digital navigator is currently being investigated and defined [49]. Digital navigators are HCPs who have undergone specialized training in the area of digital mental health applications. They provide consultative assistance to health care providers and offer continuous guidance to patients in using these applications [50]. To date, this role has been primarily concerned with the use of digital health applications. However, in light of the increasing international adoption and promotion of EHRs and ORA, it may be beneficial to consider expanding the role of digital navigators to encompass these additional tools and consider the integration of artificial intelligence–assisted support—as proposed by Wunderlich et al [48] with the concept of digital case managers. Nevertheless, artificial intelligence solutions should not replace human guidance as one of the primary concerns of older patients is that digital services could potentially supplant personal contact with HCPs [43].

Preserving the Autonomy of Older Patients

In addition to the aforementioned fundamental requirements for adapting patient portals to ensure accessibility for older adults, other aspects must be considered when involving CPs through proxy access. In the course of our study, the themes of trust, support, and control emerged as particularly salient, a finding that is also corroborated by other studies [30]. Despite the fact that CPs are only granted proxy access with the informed consent of the older patients, there is a risk that they may use Open Notes as a tool for control. This issue has significant ethical and practical implications [51]. Patients with dementia are especially reliant on the assistance of CPs in the management of their health [52], which also applies to the use of digital patient portals [53,54]. It could be argued that, particularly in cases in which patients are experiencing significant cognitive decline, such as with dementia, a certain level of control may be necessary and appropriate within the context of their care. Moreover, all participating older patients consented to the involvement of their relatives (and may revoke this consent at any time), thereby indicating their general assent to the sharing of information (and, thus, also to the potential for control). Nevertheless, this controlling behavior represents a substantial limitation of the patients' autonomy and may potentially give rise to conflict in the relationship between patients and CPs. Furthermore, in accordance with the systemic *concept of a good reason*, it can be assumed that all behaviors, including the deliberate withholding of information by patients, are motivated by a good inner reason and represent a more or less constructive coping mechanism in the face of challenges and difficulties (eg, to avoid shame or to maintain independence) [55].

At this point, it is pertinent to re-examine whether these considerations are applicable to patients with age-related cognitive impairments. These reflections could likely be extended indefinitely, leading to a vicious circle. However, it is possible to diverge from this loop and conclude that it is essential to preserve the dignity of older patients (with and without dementia) while using ORA and Open Notes [56]. Therefore, it is crucial to consider how the experience of (controlling and) being controlled can develop into a trusting dependence on the support of CPs [57]. Caine et al [58] and Latulipe et al [30] suggest that patients should be informed precisely about which information CPs can access in the patient portal. It seems particularly important to provide patients with the option of fine-grained access settings, allowing them to decide which information should be shared and which should not [59]. Such fine-grained functionality was available in early versions of the Swedish national patient portal [3,60]. In addition, a *break-glass* access control protocol can be implemented whereby patients can define which information should be released in an emergency (eg, in the event of a significant deterioration in cognitive health or an unexpected hospitalization) [30,61]. Careful attention must be paid to defining the end points meticulously and distinctly (eg, establishing clear criteria for what constitutes significant cognitive decline).

XSL•FO
RenderX

Implications

Older patients can benefit significantly from Open Notes with proxy access in their health care. However, to realize these benefits, older patients (and their CPs) must first be empowered to access the patient portal and understand clinical documentation. This requires adapting the design of patient portals to the needs of the older patient population and supporting the digital literacy of older patients through tailored individual and structural interventions. Enabling patients and CPs to interact with their HCPs through Open Notes seems to be a particularly important new feature. Many older patients rely on the support of CPs to manage their health care, especially when using digital health services. However, to ensure that older patients maintain their autonomy and dignity when using digital health services such as Open Notes, it is crucial to prevent these tools from becoming instruments of control for CPs. Older patients should be able to make granular decisions about what information they want to share with their CPs and what they want to keep private. For emergencies, a *break-glass* access protocol should be established in advance.

Limitations

This study was based on a small number of participants, which limits the generalizability of the results. Furthermore, the group of participants was highly homogeneous with regard to the categories of race and migration status. To obtain generalizable results, larger studies with a more diverse selection of participants are required in the future. Similarly, older patients and their CPs represent a relatively specific participant group, which further limits the generalizability of the results. Younger patients and their CPs (eg, children or adolescents) may have differing user experiences and encounter completely different barriers and opportunities while using Open Notes. It is also necessary to consider the potential influence of social desirability bias. It can be assumed that older patients might have occasionally embellished their statements regarding their own digital literacy and the usability of the patient portal. For instance, a greater number of CPs than patients indicated that the patients experienced difficulties when using the portal. It is difficult to acknowledge one's own shortcomings and limitations. Furthermore, it is particularly challenging to do so in the presence of others, such as CPs and interviewers, as this could lead to embarrassment and perceived loss of status. It should also be noted that the interviews were conducted via telephone. Although the participants were asked to find a quiet and secure place, it cannot be guaranteed that they were undisturbed throughout the interview. However, the telephone interviews allowed the older patients to remain in their homes (ie, they did not have to travel long distances) and have quick access to support from their CPs in case of difficulties (eg, technical problems or comprehension problems due to cognitive deficits). Thus, telephone interviews allowed older patients with physical or cognitive impairments, as well as CPs living in other cities or countries, to take part in the study and reduced barriers to study participation.

Conclusions

Our study suggests that access to Open Notes can facilitate understanding and engagement between patients and their CPs and is associated with improved communication with HCPs. This may influence the dynamics of the triadic relationship among patients, CPs, and HCPs, with potential implications for power dynamics. In summary, no single patient portal can be expected to meet the needs of all patients—one size does not fit all. Individual solutions and adaptations of ORA are clearly needed to ensure acceptance and meaningful use by older patients and their HCPs.

Acknowledgments

This study was funded by the Brandenburg Medical School Theodor Fontane publication fund supported by the Ministry of Science, Research, and Culture of the State of Brandenburg.

Authors' Contributions

All authors contributed to planning, analysis, and the critical interpretation of the findings. JS and EMD contributed to the study design, and JS supervised the study. EMD and MW contributed to data collection. EMD conducted the data analysis. EMD wrote the first draft of the manuscript. EMD, VD, and JS modified successive drafts. JS edited and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guides.
[PDF File (Adobe PDF File), 125 KB - aging_v8i1e66690_app1.pdf]

Multimedia Appendix 2
COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.
[DOCX File , 15 KB - aging_v8i1e66690_app2.docx]

References

XSL•FO
RenderX

1. Bongaarts J. Human population growth and the demographic transition. Philos Trans R Soc Lond B Biol Sci 2009 Oct 27;364(1532):2985-2990 [FREE Full text] [doi: 10.1098/rstb.2009.0137] [Medline: 19770150]
2. Dal Mas F, Massaro M, Rippa P, Secundo G. The challenges of digital transformation in healthcare: an interdisciplinary literature review, framework, and future research agenda. Technovation 2023 May;123:102716. [doi: 10.1016/j.technovation.2023.102716]
3. Hägglund M, Kharko A, Bärkås A, Blease C, Cajander Å, DesRoches C, et al. A Nordic perspective on patient online record access and the European health data space. J Med Internet Res 2024 Jun 27;26:e49084 [FREE Full text] [doi: 10.2196/49084] [Medline: 38935430]
4. Schwarz J, Esch T. [OpenNotes in mental health? Pro]. Nervenarzt 2022 Nov;93(11):1163-1164 [FREE Full text] [doi: 10.1007/s00115-022-01357-4] [Medline: 35913587]
5. Esch T, Mejilla R, Anselmo M, Podtschaske B, Delbanco T, Walker J. Engaging patients through open notes: an evaluation using mixed methods. BMJ Open 2016 Jan 29;6(1):e010034 [FREE Full text] [doi: 10.1136/bmjopen-2015-010034] [Medline: 26826154]
6. Meier-Diedrich E, Lyckblad C, Davidge G, Hägglund M, Kharko A, McMillan B, et al. Impact of patient online record access on documentation: a scoping review. JMIR Preprints Preprint posted online on July 25, 2024 [FREE Full text] [doi: 10.2196/64762]
7. 21st century cures act: the office of the national coordinator: interoperability and information blocking final rule overview for psychiatrists. American Psychiatric Association. URL: https://www.psychiatry.org/psychiatrists/practice/practice-management/health-information-technology/interoperability-and-information-blocking [accessed 2023-10-16]
8. Blease C, Torous J, Dong Z, Davidge G, DesRoches C, Kharko A, et al. Patient online record access in English primary care: qualitative survey study of general practitioners' views. J Med Internet Res 2023 Feb 22;25:e43496 [FREE Full text] [doi: 10.2196/43496] [Medline: 36811939]
9. Schwarz J, Hoetger C, Pluschke LS, Muehlensiepen F, Schneider M, Thoma S, et al. Psychiatrists' perceptions of conditions and consequences associated with the implementation of open notes: qualitative investigation. BMC Psychiatry 2024 Jun 10;24(1):430 [FREE Full text] [doi: 10.1186/s12888-024-05845-6] [Medline: 38858711]
10. Meier-Diedrich E, Neumann K, Heinze M, Schwarz J. [Attitudes and expectations of psychological and medical psychotherapists towards open notes: analysis of qualitative survey responses]. Psychiatr Prax 2024 Oct;51(7):376-384. [doi: 10.1055/a-2320-8929] [Medline: 38810902]
11. Wurster F, Di Gion P, Goldberg N, Hautsch V, Hefter K, Herrmann C, et al. Roger's diffusion of innovations theory and the adoption of a patient portal's digital anamnesis collection tool: study protocol for the MAiBest project. Implement Sci Commun 2024 Jul 15;5(1):74 [FREE Full text] [doi: 10.1186/s43058-024-00614-8] [Medline: 39010236]
12. Hägglund M, McMillan B, Whittaker R, Blease C. Patient empowerment through online access to health records. BMJ 2022 Sep 29;378:e071531 [FREE Full text] [doi: 10.1136/bmj-2022-071531] [Medline: 36175012]
13. Blease CR, Walker J, Torous J, O'Neill S. Corrigendum: sharing clinical notes in psychotherapy: a new tool to strengthen patient autonomy. Front Psychiatry 2020;11:636411 [FREE Full text] [doi: 10.3389/fpsyt.2020.636411] [Medline: 33536954]
14. Bailey SC, Wallia A, Wright S, Wismer GA, Infanzon AC, Curtis LM, et al. Electronic health record-based strategy to promote medication adherence among patients with diabetes: longitudinal observational study. J Med Internet Res 2019 Oct 21;21(10):e13499 [FREE Full text] [doi: 10.2196/13499] [Medline: 31638592]
15. van Kessel R, Ranganathan S, Anderson M, McMillan B, Mossialos E. Exploring potential drivers of patient engagement with their health data through digital platforms: a scoping review. Int J Med Inform 2024 Sep;189:105513 [FREE Full text] [doi: 10.1016/j.ijmedinf.2024.105513] [Medline: 38851132]
16. Blease C, Fernandez L, Bell SK, Delbanco T, DesRoches C. Empowering patients and reducing inequities: is there potential in sharing clinical notes? BMJ Qual Saf 2020 Oct;29(10):1-2. [doi: 10.1136/bmjqs-2019-010490] [Medline: 32188711]
17. Wang X, Zhao YC. Understanding older adults' intention to use patient-accessible electronic health records: based on the affordance lens. Front Public Health 2022;10:1075204 [FREE Full text] [doi: 10.3389/fpubh.2022.1075204] [Medline: 36761325]
18. Meier-Diedrich E, Wördemann J, Speck J, Turvey C, Weibezahl M, Schwarz J. How patients affected by a severe mental health condition and health care professionals communicate via a patient portal: qualitative analysis of secure messages. JMIR Preprints Preprint posted online on June 27, 2024.
19. Nahm ES, Zhu S, Bellantoni M, Keldsen L, Charters K, Russomanno V, et al. Patient portal use among older adults: what is really happening nationwide? J Appl Gerontol 2020 Apr;39(4):442-450 [FREE Full text] [doi: 10.1177/0733464818776125] [Medline: 29779422]
20. Hoogland AI, Mansfield J, Lafranchise EA, Bulls HW, Johnstone PA, Jim HS. eHealth literacy in older adults with cancer. J Geriatr Oncol 2020 Jul;11(6):1020-1022 [FREE Full text] [doi: 10.1016/j.jgo.2019.12.015] [Medline: 31917114]
21. Tinmaz H, Lee YT, Fanea-Ivanovici M, Baber H. A systematic review on digital literacy. Smart Learn Environ 2022 Jun 08;9:e13499. [doi: 10.1186/s40561-022-00204-y]
22. Antonio MG, Petrovskaya O, Lau F. Correction: the state of evidence in patient portals: umbrella review. J Med Internet Res 2021 Aug 16;23(8):e32421 [FREE Full text] [doi: 10.2196/32421] [Medline: 34398799]

XSL•FO
RenderX

23. Crouch E, Gordon NP. Prevalence and factors influencing use of internet and electronic health resources by middle-aged and older adults in a US health plan population: cross-sectional survey study. JMIR Aging 2019 Mar 26;2(1):e11451 [FREE Full text] [doi: 10.2196/11451] [Medline: 31518256]
24. Wolff JL, Dukhanin V, Burgdorf JG, DesRoches CM. Shared access to patient portals for older adults: implications for privacy and digital health equity. JMIR Aging 2022 May 04;5(2):e34628 [FREE Full text] [doi: 10.2196/34628] [Medline: 35507405]
25. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556 [FREE Full text] [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
26. Wolff JL, Berger A, Clarke D, Green JA, Stametz R, Yule C, et al. Patients, care partners, and shared access to the patient portal: online practices at an integrated health system. J Am Med Inform Assoc 2016 Nov;23(6):1150-1158. [doi: 10.1093/jamia/ocw025] [Medline: 27026614]
27. Dahling V, May S, Peter SV, Schwarz J, Ignatyev Y, Scheibner H, et al. [Expectations of patients and relatives for modern hospital care in geriatric psychiatry]. Z Gerontol Geriatr 2022 Oct;55(6):476-481. [doi: 10.1007/s00391-021-01941-4] [Medline: 34251463]
28. Wachenheim D. EHR shared access (Proxy) for family care partners. OpenNotes. URL: https://www.opennotes.org/proxy-access/ [accessed 2024-08-12]
29. Gleason KT, Peereboom D, Wec A, Wolff JL. Patient portals to support care partner engagement in adolescent and adult populations: a scoping review. JAMA Netw Open 2022 Dec 01;5(12):e2248696 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.48696] [Medline: 36576738]
30. Latulipe C, Quandt SA, Melius KA, Bertoni A, Miller DP, Smith D, et al. Insights into older adult patient concerns around the caregiver proxy portal use: qualitative interview study. J Med Internet Res 2018 Nov 02;20(11):e10524 [FREE Full text] [doi: 10.2196/10524] [Medline: 30389654]
31. Wolff JL, Darer JD, Larsen KL. Family caregivers and consumer health information technology. J Gen Intern Med 2016 Jan;31(1):117-121 [FREE Full text] [doi: 10.1007/s11606-015-3494-0] [Medline: 26311198]
32. Jackson SL, Mejilla R, Darer JD, Oster NV, Ralston JD, Leveille SG, et al. Patients who share transparent visit notes with others: characteristics, risks, and benefits. J Med Internet Res 2014 Nov 12;16(11):e247 [FREE Full text] [doi: 10.2196/jmir.3363] [Medline: 25405911]
33. Schwarz J, Meier-Diedrich E, Neumann K, Heinze M, Eisenmann Y, Thoma S. Reasons for acceptance or rejection of online record access among patients affected by a severe mental illness: mixed methods study. JMIR Ment Health 2024 Feb 05;11:e51126 [FREE Full text] [doi: 10.2196/51126] [Medline: 38315523]
34. Schwarz J, Bärkås A, Blease C, Collins L, Hägglund M, Markham S, et al. Sharing clinical notes and electronic health records with people affected by mental health conditions: scoping review. JMIR Ment Health 2021 Dec 14;8(12):e34170 [FREE Full text] [doi: 10.2196/34170] [Medline: 34904956]
35. Hochwarter S, Fehler G, Muente C, Eisenmann Y, Heinze M, Hibsch C, et al. Design of a patient-accessible electronic health record system in mental health. Stud Health Technol Inform 2022 May 25;294:583-584. [doi: 10.3233/SHTI220534] [Medline: 35612155]
36. Hochwarter S, Schwarz J. Facilitating collaboration by opening the common information space in mental health care. In: Proceedings of the 9th International Conference on Infrastructures for Healthcare. 2023 Presented at: InfraHealth 2023; September 11-12, 2023; Siegen, Germany. [doi: 10.1001/jamanetworkopen.2022.48696]
37. Piloting and evaluation of a participatory patient record in psychiatry and somatics (PEPPPSY). PEPPPSY. URL: https://www.pepppsy.de/ [accessed 2025-01-22]
38. Emerson RM, Fretz RI, Shaw LL. Writing Ethnographic Fieldnotes Second Edition. Chicago, IL: The University of Chicago Press; 2011.
39. Charmaz K. Constructing Grounded Theory: A Practical Guide through Qualitative Analysis. Thousand Oaks, CA: SAGE Publications; 2006.
40. Berg M. Practices of reading and writing: the constitutive role of the patient record in medical work. Sociol Health Illn 2008 Jun 28;18(4):499-524 [FREE Full text] [doi: 10.1111/1467-9566.ep10939100]
41. Lam K, Lu AD, Shi Y, Covinsky KE. Assessing telemedicine unreadiness among older adults in the United States during the COVID-19 pandemic. JAMA Intern Med 2020 Oct 01;180(10):1389-1391 [FREE Full text] [doi: 10.1001/jamainternmed.2020.2671] [Medline: 32744593]
42. Reed ME, Huang J, Brand R, Ballard D, Yamin C, Hsu J, et al. Communicating through a patient portal to engage family care partners. JAMA Intern Med 2018 Jan 01;178(1):142-144 [FREE Full text] [doi: 10.1001/jamainternmed.2017.6325] [Medline: 29159402]
43. Landgren S, Cajander Å. Non-use of digital health consultations among Swedish elderly living in the countryside. Front Public Health 2021;9:588583 [FREE Full text] [doi: 10.3389/fpubh.2021.588583] [Medline: 34568247]
44. Liu N, Yin J, Tan SS, Ngiam KY, Teo HH. Mobile health applications for older adults: a systematic review of interface and persuasive feature design. J Am Med Inform Assoc 2021 Oct 12;28(11):2483-2501 [FREE Full text] [doi: 10.1093/jamia/ocab151] [Medline: 34472601]

XSL•FO
RenderX

45. Barysheva GA, Klemasheva EI, Nedospasova OP, Ngoc TT, Thang NC. Involvement of elderly people in the processes of modern digital transformations. Adv Gerontol 2022 Sep 11;12(3):324-330. [doi: 10.1134/S2079057022030043]
46. Latulipe C, Gatto A, Nguyen HT, Miller DP, Quandt SA, Bertoni AG, et al. Design considerations for patient portal adoption by low-income, older adults. Proc SIGCHI Conf Hum Factor Comput Syst 2015 Apr;2015:3859-3868 [FREE Full text] [doi: 10.1145/2702123.2702392] [Medline: 27077140]
47. Grossman LV, Masterson Creber RM, Benda NC, Wright D, Vawdrey DK, Ancker JS. Interventions to increase patient portal use in vulnerable populations: a systematic review. J Am Med Inform Assoc 2019 Aug 01;26(8-9):855-870 [FREE Full text] [doi: 10.1093/jamia/ocz023] [Medline: 30958532]
48. Wunderlich P, Wiegräbe F, Dörksen H. Digital case manager-a data-driven tool to support family caregivers with initial guidance. Int J Environ Res Public Health 2023 Jan 10;20(2):1215 [FREE Full text] [doi: 10.3390/ijerph20021215] [Medline: 36673969]
49. Wisniewski H, Gorrindo T, Rauseo-Ricupero N, Hilty D, Torous J. The role of digital navigators in promoting clinical care and technology integration into practice. Digit Biomark 2020;4(Suppl 1):119-135 [FREE Full text] [doi: 10.1159/000510144] [Medline: 33442585]
50. Perret S, Alon N, Carpenter-Song E, Myrick K, Thompson K, Li S, et al. Standardising the role of a digital navigator in behavioural health: a systematic review. Lancet Digit Health 2023 Dec;5(12):e925-e932 [FREE Full text] [doi: 10.1016/S2589-7500(23)00152-8] [Medline: 38000876]
51. Ganta T, Appel JM, Genes N. Patient portal access for caregivers of adult and geriatric patients: reframing the ethics of digital patient communication. J Med Ethics 2023 Mar;49(3):156-159. [doi: 10.1136/medethics-2021-107759] [Medline: 35437282]
52. Kasper JD, Freedman VA, Spillman BC, Wolff JL. The disproportionate impact of dementia on family and unpaid caregiving to older adults. Health Aff (Millwood) 2015 Oct;34(10):1642-1649 [FREE Full text] [doi: 10.1377/hlthaff.2015.0536] [Medline: 26438739]
53. Gleason KT, Wu MM, Wec A, Powell DS, Zhang T, Gamper MJ, et al. Use of the patient portal among older adults with diagnosed dementia and their care partners. Alzheimers Dement 2023 Dec;19(12):5663-5671 [FREE Full text] [doi: 10.1002/alz.13354] [Medline: 37354066]
54. Riffin C, Cassidy J, Smith JM, Begler E, Peereboom D, Lum HD, et al. Care partner perspectives on the use of a patient portal intervention to promote care partner identification in dementia care. J Appl Gerontol 2025 Jan;44(1):136-146. [doi: 10.1177/07334648241262649] [Medline: 38901834]
55. Kutz A. Systemische Haltung in Beratung und Coaching. Wiesbaden, Germany: Springer; 2020.
56. Tranvåg O, Petersen KA, Nåden D. Dignity-preserving dementia care: a metasynthesis. Nurs Ethics 2013 Dec;20(8):861-880. [doi: 10.1177/0969733013485110] [Medline: 23715554]
57. Wolfe SE, Greenhill B, Butchard S, Day J. The meaning of autonomy when living with dementia: a Q-method investigation. Dementia (London) 2021 Aug;20(6):1875-1890 [FREE Full text] [doi: 10.1177/1471301220973067] [Medline: 33372553]
58. Caine K, Kohn S, Lawrence C, Hanania R, Meslin EM, Tierney WM. Designing a patient-centered user interface for access decisions about EHR data: implications from patient interviews. J Gen Intern Med 2015 Jan;30 Suppl 1(Suppl 1):S7-16 [FREE Full text] [doi: 10.1007/s11606-014-3049-9] [Medline: 25480719]
59. Caine K, Hanania R. Patients want granular privacy control over health information in electronic medical records. J Am Med Inform Assoc 2013 Jan 01;20(1):7-15 [FREE Full text] [doi: 10.1136/amiajnl-2012-001023] [Medline: 23184192]
60. Nurgalieva L, Cajander Å, Moll J, Åhlfeldt RM, Huvila I, Marchese M. 'I do not share it with others. No, it's for me, it's my care': on sharing of patient accessible electronic health records. Health Informatics J 2020 Dec;26(4):2554-2567 [FREE Full text] [doi: 10.1177/1460458220912559] [Medline: 32264734]
61. Marinovic S, Craven R, Ma J, Dulay N. Rumpole: a flexible break-glass access control model. In: Proceedings of the 16th ACM Symposium on Access Control Models and Technologies. 2011 Presented at: SACMAT '11; June 15-17, 2011; Innsbruck, Austria.

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
CP: care partner
EHR: electronic health record
HCP: health care professional
ICD-10: International Classification of Diseases, 10th Revision
ORA: online record access
PEP.AGE: Piloting and evaluation of a participatory patient-accessible electronic health record for geriatric psychiatric patients and their care partners
PEPPPSY: Piloting and evaluation of a participatory patient-accessible electronic health record in Psychiatry and Somatics

XSL•FO
RenderX

Edited by M Gray; submitted 20.09.24; peer-reviewed by F Wurster, EKO Wong; comments to author 15.10.24; revised version received 23.10.24; accepted 30.10.24; published 24.02.25.

Please cite as:
Meier-Diedrich E, Esch T, Hägglund M, Heinze M, Hochwarter S, Speck J, Wagener M, Dahling V, Schwarz J
Experiences of Older Mental Health Patients and Their Care Partners Using a Proxy Account to Access Open Notes: Qualitative Interview Study
JMIR Aging 2025;8:e66690
URL: https://aging.jmir.org/2025/1/e66690
doi:10.2196/66690
PMID:

©Eva Meier-Diedrich, Tobias Esch, Maria Hägglund, Martin Heinze, Stefan Hochwarter, Justin Speck, Marie Wagener, Volker Dahling, Julian Schwarz. Originally published in JMIR Aging (https://aging.jmir.org), 24.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Employers' Perspectives of Caregiver-Friendly Workplace Policies for Caregiver-Employees Caring for Older Adults in Hong Kong: Thematic Analysis

Maggie Man-Sin Lee[1], BBA, MPH, JD, PhD; Eng-Kiong Yeoh[1,2], MBBS; Eliza Lai-Yi Wong[1,2], BSN, RN, MPH, PhD

[1]JC School of Public Health and Primary Care, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

[2]Centre for Health Systems & Policy Research, JC School of Public Health and Primary Care, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

Corresponding Author:
Eliza Lai-Yi Wong, BSN, RN, MPH, PhD
JC School of Public Health and Primary Care
Faculty of Medicine
The Chinese University of Hong Kong
School of Public Health Building, Prince of Wales Hospital
Shatin, New Territories
Hong Kong
China (Hong Kong)
Phone: 852 2252 8772
Fax: 852 2145 7489
Email: lywong@cuhk.edu.hk

Abstract

Background: Caregiver-friendly workplace policies (CFWPs) are rare in Hong Kong. With Hong Kong facing a "silver tsunami" in the near future, it is important to understand the need for such policies and the views of employers for future facilitation.

Objective: This study aimed to identify the support that is currently provided or that could be provided to caregiver-employees (CEs) caring for older adults in Hong Kong and assess the challenge and facilitative support for employers to adopt CFWPs in the specific context of Hong Kong.

Methods: A qualitative research design with semistructured individual in-depth interviews with employers from Hong Kong was adopted for this study. A purposive snowball sampling method was used to recruit participants from the 7 primary industries mentioned in the Hong Kong census and from all 3 employer types (private, public, and nongovernmental organizations), which allowed the inclusion of participants sensitized to the idea and potential of CFWPs. Thematic framework analysis was used to evaluate the data collected during the interviews.

Results: We interviewed 17 employers and managers from 7 major industries in Hong Kong (2.5 to 120,000 employees). There were 4 (24%) male and 13 (76%) female participants, and the participant age ranged from 30 to 50 years. All participants held managerial positions at the time of the interview. Of the 17 participants, 13 were from private companies, 2 were from public institutions, and 2 were from nongovernmental organizations. Four of the companies had a global presence. Four main themes were identified: (1) current support and potential support for CEs (which was limited to discretionary annual leave and unpaid leave when annual leave was exhausted), (2) challenges in adopting CFWPs, (3) facilitating support for adopting CFWPs, and (4) incentives for adopting CFWPs. The participants rated information and resources for CEs (mean 8.56, SD 0.37), bereavement leave (mean 8.47, SD 0.63), flexible working hours (mean 8.32, SD 0.48), and caregiver-inclusive corporate culture (mean 8.32, SD 0.48) as essential CFWPs for CEs in Hong Kong.

Conclusions: While several studies have reported the types of CFWPs and their impacts on CEs, stakeholders' perspectives on CFWPs have been rarely investigated. This study found that although employers consider CFWPs as necessary and see them as a catalyst for a long-term win-win situation, the current support for CEs is discretionary and industry-specific. Government leadership is critical for formulating, piloting, and implementing CFWPs to create a friendly environment that encourages disclosure with trust and respect across industrial sectors in Hong Kong. This study identified the current unmet needs and demands of CEs

XSL•FO
RenderX

from the employer's perspective, the barriers to large-scale adoption of CFWPs, and the path forward to inform further discourse and policy formulation in Hong Kong.

(JMIR Aging 2025;8:e68061) doi:10.2196/68061

KEYWORDS

caregiver; aging; burnout; stress; mental health; employees

Introduction

Filial piety has strong cultural roots in Confucian ideology and is especially prominent in Eastern countries in contrast to children-centered societies, which are more prevalent in Western countries [1-3]. However, there is some evidence of the reduction of the influence of classical Confucian ethos in Asian societies, especially in highly urbanized cities like Beijing [4,5], although the extent of this change is currently unknown. Similarly, the cultural reconfiguration of familial values and expectations has been noted among immigrant Asian families in Western countries [6].

Moreover, emerging research indicates that authoritarian filial piety, which prioritizes social obligations toward parents over one's own needs, varies, but reciprocal filial piety, characterized by genuine gratitude for parents' efforts or sacrifices and voluntary support for them, remains consistent across Eastern and Western cultures [3,7,8]. For instance, a study comparing filial piety among Asian Singaporean and non-Asian Australian adults showed a significantly higher mean score for authoritarian filial piety (mean 21.6, SD 5.9 vs mean 19.4, SD 6.0) on the standardized dual filial piety scale among Singaporean adults, but reciprocal filial piety (mean 34.4, SD 4.8 vs mean 34.2, SD 4.4) was comparable in the 2 populations [8].

Important differences have also been noted in patterns of filial piety within Asian cultures. For example, young adults in Taiwan endorsed strong beliefs in reciprocal and authoritarian filial piety, with both beliefs positively impacting life satisfaction, while young adults in Macau endorsed strong beliefs in reciprocal filial piety and those in Hong Kong endorsed strong beliefs in authoritarian filial piety, with positive impacts on life satisfaction [9]. Interestingly, collectivism and not ethnicity has been identified as a predictor of both forms of filial piety and positive attitudes toward caring for aging parents [7]. Hence, filial piety may be more cross-cultural than previously believed, albeit with regional nuances [7,8]. Indeed, 74% of older adults in the United States National Health and Aging Trends Study (NHATS 2011-2017) reported receiving support from one child caregiver, and 41% of children provided care to aging parents [10].

The rapid increase in the aging population has also increased the caregiving burden of family caregivers, with both being reported across cultures and countries [11,12]. Although several studies have identified caregiver burden as a significant risk factor for burnout and poor mental health outcomes among family caregivers [11,13,14], stronger filial piety may lead to lower caregiver burden [15,16]. A systematic review of 12 studies showed a significant negative correlation (r=–0.23) and association (β=–0.27) between filial piety and caregiver burden among adult children [15]. However, although filial piety is a prominent determinant of caregiver burden, it is not the sole determinant even in Asian societies [17].

While noting mild-to-moderate burden among Asian caregivers of older adults with chronic illnesses, a systematic review observed that specific characteristics of the care recipients (functional dependency, comorbidities, memory, and sleep impairments) and caregivers (advancing age, male gender, spouse as a care recipient, longer care duration, and lack of support/assistance) were associated with higher caregiver burden [16]. However, Asia has a lower old age dependency ratio (population aged ≥65 years to working age population) and health-adjusted dependency ratio (population with the same or higher disease burden to population with a lower disease burden as an average global 65-year-old person) than Europe, North America, or Oceania [18].

East Asia, in particular, has the highest employment-to-population ratio globally (63; for comparison, the corresponding numbers for the Euro area and North America are 58 and 60, respectively) [19], and thus, a substantially large proportion of caregivers in this region likely have a dual role of carer-employee (CE). Studies have shown that those with elderly or child caregiving responsibilities are less favorably viewed by employers in terms of competency (F_2=10.99; P<.001; η^2p=0.092), agency (F_2=15.00; P<.001; η^2p=0.121), commitment (F_2=164.83; P<.001; η^2p=0.602), and availability (F_2=69.01; P<.001; η^2p=0.388) [20]. Hence, CEs were less likely to be hired (F_2=13.65; P<.001; η^2p=0.111) and, if hired, were offered lower salaries (F_2=13.08; P<.001; η^2p=0.107) than nonprimary caregivers [20]. Moreover, the dual role adds extra strain, negatively impacting the mental and physical health of CEs [21,22] and their performance in both roles, which aligns with the role accumulation theory [14,23].

Therefore, the strain of the dual role of CEs will likely offset the low caregiver burden associated with filial piety in Asian populations [8,16,17]. Moreover, the low dependency ratio among Asian populations is likely to be mitigated by the rapid aging of many East Asian populations by 2030, as shown by the China Health and Retirement Longitudinal Study [24]. Thus, several studies have investigated the impact of caregiver-friendly workplace policies (CFWPs) on CEs, which have been associated with the improved overall health of CEs by reducing occupational and overall stress, minimizing work interruptions, and improving performance [22,25-29]. There are also direct economic benefits accruing from adopting CFWPs. For example, educating CEs about their caregiving activities generates a net benefit ranging from US $48,010 to US $675,657 for CEs and employers [30].

XSL•FO
RenderX

CFWPs that help employees manage multiple roles and improve their work-life balance [31] are increasingly being adopted to mitigate some of the caregiving burden of CEs, especially in developed economies [32]. For instance, most employers in the United States and Canada have adopted measures to create a caregiver-friendly workplace environment, such as support services, flexible working hours, financial support, and paid/unpaid leave [33,34]. Similarly, the UK government has set up programs and nongovernmental organizations (NGOs) to help the private sector become more CE friendly [33]. In the specific context of Hong Kong, official caregiver support has not yet been observed, although some private companies provide generic support [35].

Moreover, while several studies have reported the types and impacts of CFWPs [29,33], there is a paucity of studies reporting stakeholders' perspectives on CFWPs. We could identify only 1 study that explored the perspectives of stakeholders (managers working in the Canadian health care sector) on CFWPs [36]. Therefore, this study explored the views and experiences of employers regarding CFWPs for CEs caring for older adults in Hong Kong with the objectives of identifying (1) the support that is currently provided and that could be provided to CEs by employers and (2) the challenge and facilitative support for employers to adopt CFWPs in the specific context of Hong Kong.

Methods

Setting

In Hong Kong, authoritative filial piety emphasizes caring for elderly parents [37]. However, the shift in parental expectations from tangible to emotional support poses challenges for caregiver mental health [38]. Sole children of aging parents face intense burdens, especially in nuclear families, which comprise 64% of households in Hong Kong [39]. One-third of households have at least one elderly member [39], and only 2.3% of the elderly population is entitled to full subsidization [40]. Additionally, given the 56% increase in children with special educational needs in the last half decade [41], support for people in the "sandwich generation" who are caring for both young children and old parents becomes important [28].

The number of CEs, defined as family members (spouse, parents, or siblings) who are employed with monetary rewards and at the same time provide care to another family member [42], in Hong Kong remains undocumented, as general informal caregiver support has only recently started to pick up in Hong Kong. Nevertheless, Hong Kong's highly engaged workforce (50.87% of the overall population) [43] faces additional challenges due to chronic diseases among older adults. Over 65% of the older population in Hong Kong have at least one chronic disease, while one-third have ≥2 chronic diseases [44].

Design

A qualitative research design with semistructured individual in-depth interviews involving employers and company management in Hong Kong was adopted for this study. Independent interviews helped build critical rapport between the researcher and the participants, with minimized privacy issues compared to focus groups, so that respondents would be as open and honest as possible. The COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist (Multimedia Appendix 1) was used as the reporting framework [45].

Participants

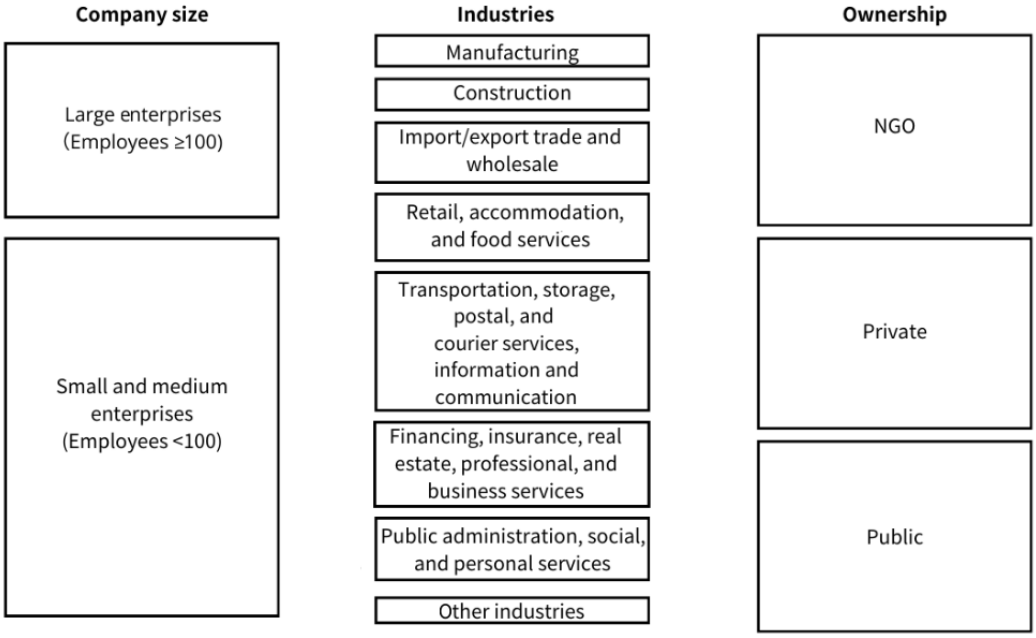
Purposive sampling was used to recruit participants, which allowed the inclusion of participants sensitized to the idea and potential of CFWPs. The list of potential participants or the sampling frame included individuals who had participated in ESG (environmental, social, and governance) policy consultations, and the information was sourced from a local NGO that advocates for working caregivers. Participants were also identified through the snowballing technique via referrals during interviews. Interviews were arranged until data saturation was achieved.

We started recruitment based on the 7 main industries in the Hong Kong census [39] and recruited from all 3 employer types (private, public, and NGO) (Figure 1). We then expanded the list to ensure a good proportion of large enterprises, small- and medium-sized enterprises, NGOs, and public institutions until data saturation was reached. Data saturation was assumed when no new themes emerged during interviews.

Participants were deemed eligible for inclusion if they had served in a managerial role for at least one staff member at a company, had experience working with CEs in a professional or personal context, and were proficient in English or Chinese (Cantonese or Mandarin). Those serving as third-party consultants for a company were excluded.

XSL•FO
RenderX

Figure 1. Map of employers and management recruited in this study. NGO: nongovernmental organization.



Data Collection

The semistructured interviews were conducted using a discussion guide consisting of 17 questions related to four aspects: (1) respondents' personal experiences with caregiver employees, (2) their attitudes and preferences for policies, (3) their perceptions of a caregiver-friendly workplace, and (4) the rating of 10 policies (Multimedia Appendix 2). The discussion guide was developed based on a thorough literature review, the first author's previous Master's thesis [35], and the opinions of 2 expert panels of researchers specializing in the field of caregiving (Professor E Wong, Professor D Dong, and Dr C Chan from the School of Public Health and Primary Care, The Chinese University of Hong Kong).

The first author (MMSL; PhD candidate at the Chinese University of Hong Kong) conducted all interviews in Cantonese face-to-face at the participants' offices or through videoconferencing, with each lasting 60 to 90 minutes. Participants' written consent and demographic data (age, gender, working position, and marital status) were obtained after the interviewer's self-introduction (comprising qualifications, current and past research focus, and experience) and an explanation of the purpose of conducting the research. Conversation starters were used to open up the participants, so they would feel more at ease and provide the most candid answers. Prompts were also used for specific and related questions during the interview to fully explore their experiences and perceptions [46,47].

At the end of the interview, participants were invited to vote for specific CFWPs based on the degree of importance (0 to 10) from the perspective of employers and managers. The list of CFWPs developed from the global literature review included caregiver-inclusive corporate culture; paid caregiver leave; unpaid caregiver leave; bereavement leave; flexible working hours; flexible work locations; switch to part-time work; unpaid leave; medical needs/insurance of employees' parents; and caregiving information, carer skills, and guide to community care resources. No nonparticipants were present during the interviews.

Analysis

Thematic framework analysis in applied policy research was adopted for data analysis. A 5-stage analytical process described by Ritchie and Spencer [48] was used, which comprised (1) familiarization, (2) identifying a thematic framework, (3) indexing, (4) charting, and (5) mapping and interpretation (Figure 2) [48,49].

Audio recordings of all interviews were transcribed and coded using NVivo software (version 12, Lumivero) independently by 2 researchers (MMSL and EW) to ensure all interpretations were agreeable and to enhance the reliability of the study. The transcripts or the findings were not returned to participants for comment or corrections. Any discrepancy in coding was resolved by discussion with supervisors.

The rating of the 10 policies by importance was reported as mean and SD. Inapplicable policies for specific industries were removed during the calculation. For example, there were no flexible work locations for the retail, food, and construction industries. A coding hierarchy chart (Figure 3) was generated to show the visual patterns of the coding.

XSL•FO
RenderX

Figure 2. The 5-stage analytical process used in this study.

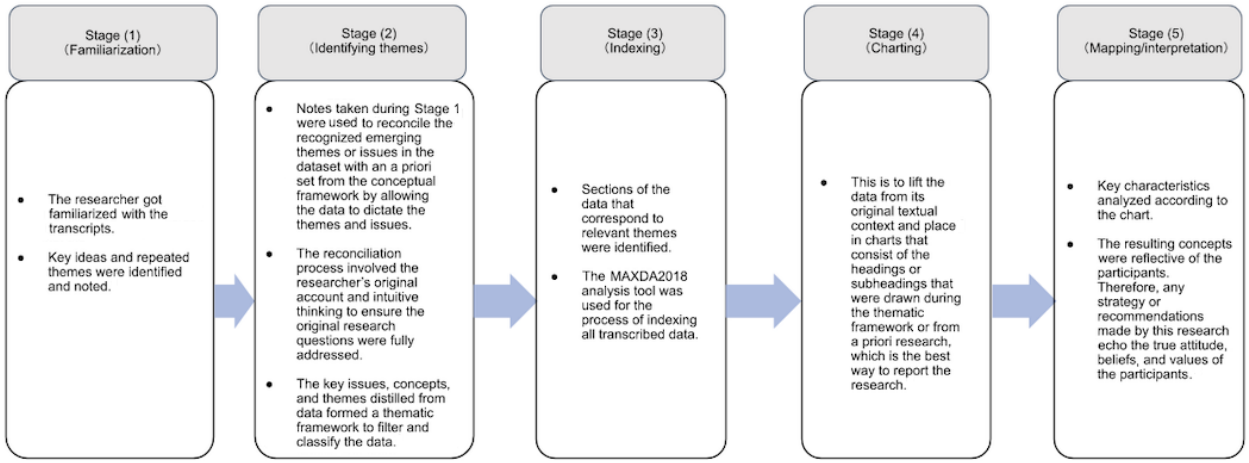
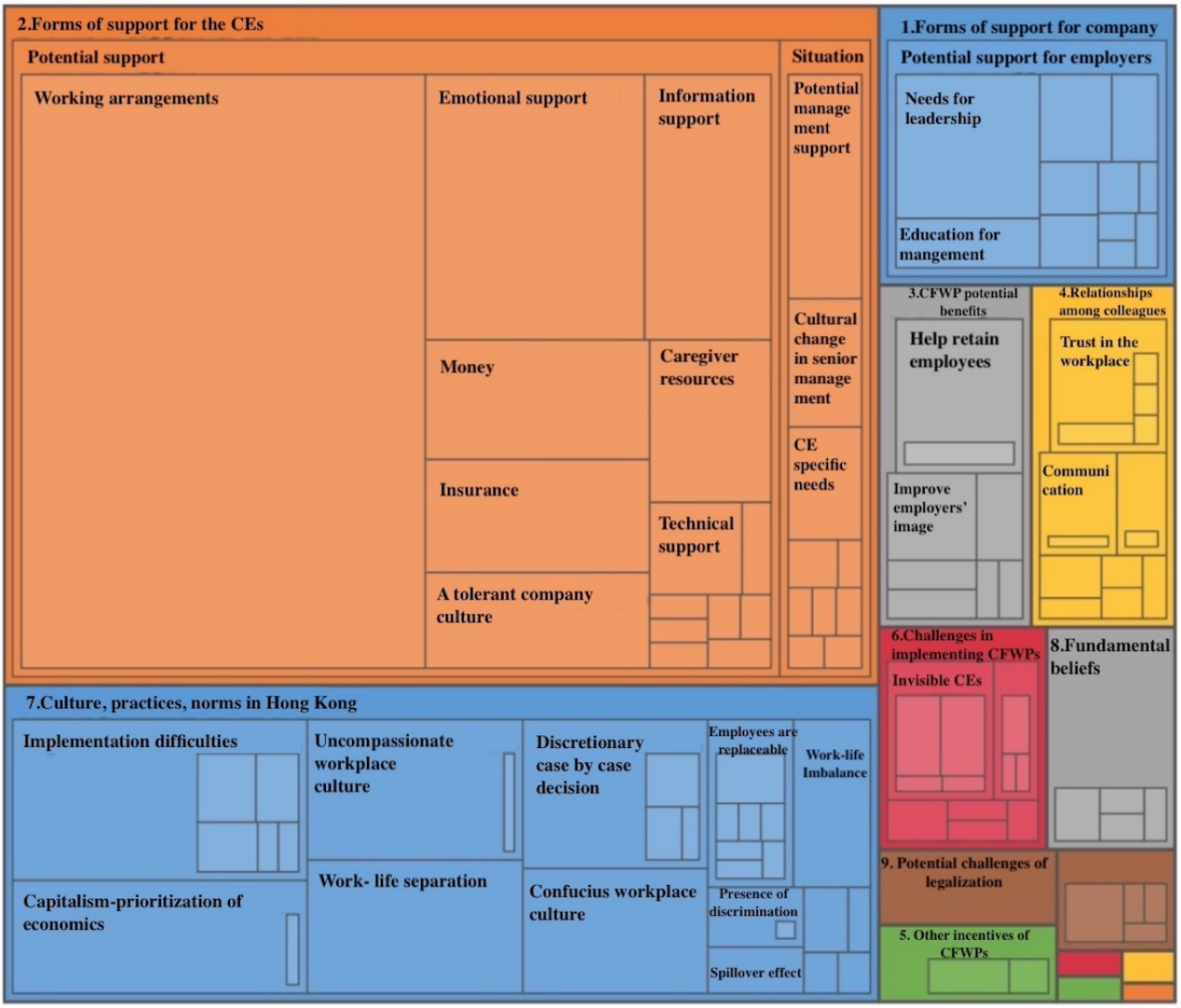


Figure 3. Coding hierarchy chart. CE: caregiver-employee; CFWP: caregiver-friendly workplace policy.



Ethical Considerations

Research ethics approval was granted by The Chinese University of HK Survey and Behavioral Research Ethics Committee (reference number: SBRE(R)-21-011). Written consent was obtained from interview participants, and they were assured of their rights and freedom to withdraw from the study at any time. The information and responses of the participants were treated confidentially. All project data were anonymized and kept in

XSL•FO
RenderX

password-protected folders that were only accessible to the project researchers and supervisors.

Results

Overview

Interviews with 17 participants comprising 4 (24%) male and 13 (76%) female participants were conducted in November 2021, and additional participants were not recruited as data saturation was achieved. Details of the participants are presented in Table 1. The age of the participants ranged from 30 to 50 years, and all held managerial positions at the time of the interview. Of the 17 participants, 13 were from private companies, 2 were from public institutions, and 2 were from NGOs. Given that the proportion of workers employed in public institutions and NGOs is less than 25% of the total workforce in Hong Kong [39], the ratio of 13:2:2 is justified.

The companies and institutions encompassed the 7 major industries in Hong Kong (ranging from 2.5 to 120,000 employees). Four of the companies had a global presence. Data saturation was achieved for NGOs and public institutions with 4 participants, all of whom viewed their industries' organizational policies and culture similarly.

Most participants (16/17) expressed that the concepts of CFWPs had been unheard of before the interviews. However, they all supported the core ideas of CFWPs (enabling the dual roles of CEs as a caregiver and a productive member of the institution) and agreed on the fact that CFWPs were needed, subject to the specific needs of the business, for the betterment of the venture and Hong Kong's future. The participants rated (1) information, caregiving skills, and guide to community care resources for CEs; (2) bereavement leave; (3) flexible working hours; and (4) caregiver-inclusive corporate culture as the most important policies (Table 2).

Despite the unanimity in supporting CFWPs, the operational practicality of CFWPs was of concern to all the participants. Four overarching themes were identified as outcomes of the thematic framework method in the analysis: (1) current support and potential support for CEs, (2) challenges in adopting CFWPs, (3) facilitating support for adopting CFWPs, and (4) incentives for adopting CFWPs (Table 3).

XSL•FO
RenderX

Table 1. Detailed demographics of the participants.

Number	Organization				Position	Age (years)/sex	Marital status[a]
	Industry	Type	Size	Presence			
1	Public administration, social, and personal services	Public	100 (LE[b])	Regional	Chief/director of a primary school	47/male	Married; #2
2	Public administration, social, and personal services	Private	8-10 (SME[c])	Regional	Executive director of a beauty company	60/female	Married; #1
3	Transport, storage, postal, and courier services; information and communication	Private	8-10 (SME)	Regional	Director of an e-platform building firm	40-50/female	Separated; #0
4	Import/export and wholesale	Private	8-10 (SME)	Regional	Director of an import/export company	30-40/female	Married; #1
5	Financing, insurance, real estate, professional, and business services	Private	200-300 (LE)	Regional	Managing director of a financial company	36/female	Single; #0
6	Public administration, social, and personal services	Private	9000 (LE)	Global	Senior marketing director of a music label corporation	48/female	Married; #2
7	Public administration, social, and personal services	Private	2.5 (SME)	Regional	Director of an individual psychiatric clinic	40/male	Married; #1
8	Retail, accommodation, and food services	Private	32 (SME)	Regional	Sales and marketing director	32/female	Single; #0
9	Public administration, social, and personal services	NGO[d]	15 (SME)	Regional	Project manager	35/female	Married; #0
10	Financing, insurance, real estate, professional, and business services	Private	20,000 (LE)	Global	District director	60/male	Married; #2
11	Public administration, social, and personal services	Public	1300 (LE)	Regional	Head of department at a financial regulatory institution	47/female	Married; #2
12	Public administration, social, and personal services	NGO	40 (SME)	Regional	Manager of a district health center	30-40/female	Married; #0
13	Financing, insurance, real estate, professional, and business services	Private	100 (LE)	Regional	Senior marketing manager of a marketing company	32/female	Single; #0
14	Public administration, social, and personal services	Private	120,000 (LE)	Global	Market access manager of a pharmaceutical company	30-40/female	Single; #0
15	Manufacturing	Private	100 (LE)	Global	CEO[e]	30-40/male	Married; #2
16	Construction	Private	10 (SME)	Regional	Owner of a construction company	31/female	Married; #1
17	Retail, accommodation, and food services	Private	11 (SME)	Regional	Owner of a restaurant	31/female	Married; #0

[a]# denotes the number of children.

[b]LE: large enterprise.

[c]SME: small and medium enterprise.

[d]NGO: nongovernmental organization.

[e]CEO: chief executive officer.

XSL•FO
RenderX

Table 2. Participant rating of the importance of the policies.

Rank	Policy	Score, mean (SD)
1	Information/caregiving skills/guide to community care resources	8.56 (0.37)
2	Bereavement leave	8.50 (0.55)
3	Flexible working hours	8.47 (0.63)
4	Caregiver-inclusive corporate culture	8.32 (0.48)
5	Flexible work locations	7.80 (0.61)
6	Paid caregiver leave (especially for those caring for elderly people with chronic diseases, excluding counting maternity leave)	7.72 (0.54)
7	Unpaid caregiver leave	7.60 (0.62)
8	Aiding medical needs/insurance of employees' parents	7.56 (0.75)
9	Switch to part-time work	7.03 (0.44)
10	Unpaid leave	6.78 (0.68)

Table 3. Themes and subthemes.

Main theme	Subthemes
Current support and potential support for CEs[a]	• Leave • Nonfinancial support • Emotional support • Information support • Technical support • Industry-specific support
Challenges in adopting CFWPs[b]	• Invisibility of CEs • Hard to adopt a fair policy for the duration of leave • Concerns about standardization
Facilitating support for adopting CFWPs	• Establishment of a government compensation fund and legal framework for caregiver leave • Government leadership
Incentives for adopting CFWPs	• Build a positive organizational culture • Recruitment of talent and extension of the worldwide network

[a]CE: carer-employee.

[b]CFWP: caregiver-friendly workplace policy.

Theme 1: Current Support and Potential Support for CEs

Six subthemes were identified based on the participants' current and potential support needs.

Leave

All participants had access to discretionary annual leave and unpaid leave when annual leave was exhausted, except participants 16 and 17, who were from the construction and food industries, which grant pay on a daily or hourly basis. All participants stated that they had the flexibility of taking a few hours of leave during working hours at their discretion for an emergency, for instance, a medical emergency related to an older family member at home. However, none of the participants reported specific caregiver policies in their workplaces.

There was bereavement leave in the participants' companies, but as noted by participant 1, such leave "is only for immediate family members." Participant 1 further noted the rigidity of bereavement leave as follows:

> *Like my wife's father, whom she also cared for, was paralyzed for a while and then passed away. I could not take leave for his funeral.*

Nonfinancial Support

Unresolved age-related care challenges created pessimism about the future among 10 participants: rising health care costs and longer life expectancy strain lower- and middle-class families, and single-child families bear immense burdens when caring for elderly people and young people. Without the financial means to hire a helper or buy insurance policies, the quality of life of CEs and their families would drastically drop in the face of immense economic, physical, mental, and emotional strains. Employers recognized the resource-based challenges CEs face, emphasized the need for nonfinancial support, and deemed direct financial support too expensive to be on the agenda for the time being. Participant 4 specifically pointed to the

XSL•FO
RenderX

importance of time as an example of nonfinancial support for working caregivers:

> *I think the most important thing is time. There are different degrees of illness; maybe he (i.e., the care recipient) has a stroke, or he has terminal cancer, a case like mine has only a few months left. And if you have to keep working and cannot spend time with your family. I think that's a great pity.*

Emotional Support

Seven participants pointed to the importance of emotional support for CEs. Participants pointed to the heavy role burdens of CEs, who need to fulfill their roles as parents, team members in the workplace, and caregivers of older adults. For instance, participant 8 witnessed severe psychological distress, possibly an emotional breakdown, during one of the work trips when her colleague was struggling with poor support from her marriage and was struggling in her career while caring for her children and an ailing elderly person. Participant 4 considered that emotional support should apply to all employees (not only CEs):

> *Emotional support should not only apply to CEs. I think the emotions of all employees are important as long as they are in the workforce. Emotional distress is like a cold, very common.*

Participant 12, an occupational therapist specializing in mental health, noted that work stress in Hong Kong is so severe that depression in young workers is often due to work stress. Long working hours and the stigma and discrimination against mental illness would further discourage Hong Kong CEs from taking care of their mental well-being. A sudden imposition of the caregiver role would further challenge the coping skills of the workers themselves and jeopardize their well-being.

Information Support

Four participants pointed out the need for a hotline for CEs to receive information and counseling services. Participant 14, who worked for a multinational pharmaceutical company, reported that they had already set up an international hotline to support employees in potentially all life-related aspects:

> *You can call there to find out about the best schools for children in your neighborhood. Since the hotline is run by a third party, anonymity is maintained to protect privacy, which makes the service even more attractive.*

Technical Support

Four participants also pointed out the problem of time strain (high time demands and little control over timing), which applies to both work and unremitting care work. They therefore pointed to the importance of home-based care services that allow CEs some personal space and time. In this context, participant 6 mentioned that today's open-plan offices prevent CEs from taking personal calls when needed. Therefore, personal space in the workplace, such as a phone box, would support CEs in fulfilling both roles effectively and safely. Participant 6 elaborated as follows:

> *Open-plan offices are very popular today, which I object to because it lacks personal space. A phone box is a personal space where colleagues have to take their own personal calls. For example, if I were to have a conversation with my mother, who is also deaf. I yell, 'No, the diarrhea pill is upstairs.' That can be very embarrassing in an open-plan office. In fact, I think this concept is in line with the idea of work-life separation.*

Industry-Specific Support

Participant 1, one of the directors at a publicly funded primary school, pointed out that staff turnover could severely affect the quality of teaching and that CFWPs were crucial to the school context, given the overly protected privacy in the school's workplace culture. No personal matters were allowed to be revealed that could cause suspicion and doubt among colleagues and between teachers and parents. According to participant 1, in this specific sector, either flexible working hours and locations or a switch to part-time work was possible. Given the bureaucracy in applying for leave and extended working hours, a CFWP culture for school teachers was necessary. Participant 1 commented:

> *Changes should be made at the top policy level of school administration. Because at the moment, even if you are a headmaster, you act according to a set of rules that restrict what you can do and what you can use.*

How wages are paid is also a determinant of whether a policy of flexible working hours is needed. In construction (daily rates), insurance (piece rates), or creative industries (commission rates), the flexibility of working hours is built into the mode of payment. Therefore, there is no substantial need to include it in CFWPs. Flexible work location arrangements are not applicable in construction, retail, food, and personal services where the location of the staff is essential to the purpose of the employment.

Theme 2: Challenges in Adopting CFWPs

Invisibility of CEs

Many challenges of introducing CFWPs are closely linked to the nature of CEs. A key feature of CEs is their invisibility. Unlike maternity leave, which 11 participants said is easier to implement than caregiver leave, no identifiable physical features (such as a swollen and protruding belly) make CEs recognizable. Participant 14, for instance, made the following statement:

> *People are very visual, maternity leave, that the colleague just had a baby. But caregivers, not everyone can see their parents or the needs of their family.*

The invisibility of CEs has enormous implications for the adoption and implementation of CFWPs, as it becomes more difficult to convince everyone of the need to support CEs and make extra efforts to accommodate them, even if maternity, paternity, and caregiving leaves are all based on the same basic need to care for family members. Participant 17 elaborated as follows:

XSL•FO
RenderX

> *[regarding employees requiring maternity leave], we can see the belly sticking out. Also, there is a life in there; you are responsible if something happens to the mother or child. You do not feel so strong if the colleague has a sick parent, for example. If the pregnant colleague feels uncomfortable with her belly and lies down, you have to let her immediately. But if the colleague says that her father is in hospital because of stomach problems, you would not feel so strongly at that moment. Besides, I have experienced that myself—some people might doubt if it is even real.*

According to the participants, CEs disclose their status as caregivers directly to their supervisors when they need to take leave or use other forms of support.

Hard to Adopt a Fair Policy for the Duration of Leave

The ambiguous needs of CEs create many grey areas for institutional arrangements. To be fair to all employees, the rationale for the arrangements should be clear and relevant, but this poses a problem for participants when it is necessary to allocate a certain number of leave days depending on the nature, severity, and stage of medical conditions. On the other hand, participants lacked the medical expertise to assess the CE's situation for a case-by-case scenario, while adopting an all-caregiver-inclusive package for all CEs could be resource-intensive and disruptive to daily operations. Thus, there was no consensus among the participants regarding standardization. Participants 9 and 11, for example, felt that standardization would provide a reference point and guideline for companies, which would help combat discrimination against CEs and give the company an advantage in hiring workers. In contrast, others, such as participant 6, felt the standardization would lack the flexibility to be helpful. This complex and difficult-to-standardize situation of CEs contrasts with maternity and paternity leaves, which are standardized and predictable.

This ambiguous need of CEs further discourages companies from using CFWPs when they entail prohibitive costs in the long run. Participant 15 explained:

> *There is a predictable time frame for maternity and paternity leave... But sickness is a bit longer. For maternity or paternity leave, you would give birth to maybe only two to three children in this lifetime and then stop, so you would only take maternity/paternity leave two to three times; here, there is a fixed number of dates, you can predict it. But if you are taking care of chronically ill or dependent family members, it may take a long time; it may be 5 or 10 years. In between, you may have to take leave every time to accompany them to doctor's appointments, which may happen 10-20 times a year; that possibility actually exists. So, with that uncertainty, it becomes a bit more difficult to implement* [the CFWP policy]

Concerns About Standardization

Standardization also raises concerns that the mechanisms of CFWPs could be abused. Participant 16 pointed to the difficulty of proving caregiver status:

> *Can I just find any aunt or distant relative? That is, the moment I want to take a holiday, the moment I do not want to take this job, I can just find a random aunt. They give me a paper, I give it to the company, then I get the benefits. I feel that these grey areas can be abused very easily... If I introduce care leave as an employer, I may suddenly get them [these medical certificates]. The next day, I wake up and suddenly get a message that I can not go to work [because of a family member's condition]... The wheel would stop turning; we would not know what to do, so I think this thing is more difficult than paternity leave. With paternity leave, you can easily prove that you are the father. Due to the very ambiguous nature of CEs' needs, this can be a major trust issue for management to support CEs.*

Participants commonly expressed that it was difficult to justify the CFWPs giving priority exclusively to workers caring for older individuals, risking favoritism over other needs, such as workers raising their children. The heavy workload of CEs for ailing older individuals would not make caring for needy young individuals less stressful. The cultural preference for newborns in Hong Kong was pointed out by participant 4 as follows:

> *Everyone expects a happy baby, but not everyone would consider the possibility of having a sick parent.*

The comments make the different treatment of CEs with an ailing older relative even more challenging to accept.

Theme 3: Facilitating Support for Adopting CFWPs

Governmental Compensation Fund and Legal Framework for Caregiver Leave

Addressing paid/unpaid leave for caregivers involves considering financial costs, a central issue raised by all participants. Participants proposed government compensation funds for employers to cover the costs of CEs' time off, cashing out the mandatory provident fund (MPF) during a medical crisis, legal frameworks for eligibility, and compulsory employment insurance. Regarding the cash out from the MPF, participant 17 commented:

> *The original intention of the MPF was to maintain a certain quality of life in retirement. But if he (CE) can't get through the hardest time of his life (caring for ailing older family members), it doesn't matter how much money he'll receive in the future.*

Government Leadership

The practical challenges practicing doctors face include a lack of government guidance and reimbursement. Many participants felt that the government must take the initiative to support CEs. Many participants expressed the opinion that "if the government implements it first, we will follow;" however, there is currently no single, consistent CFWP in Hong Kong. Therefore, support for CEs has only been granted on a discretionary and case-by-case basis. If enacted, the policy and legislation supporting CEs would be mandatory for them to follow, and this government initiative would even trigger a shift in

XSL•FO
RenderX

companies and management thinking. Participant 2, for example, said the following:

> *I do not think there was anything like maternity leave before; it's even more similar with paternity, but we all have that now, so I would think, why not? [I am] also very, very supportive* [of CE leave]

Two participants also felt that with the introduction of maternity and paternity leaves, it would be easier to advocate for support for caregivers if the government takes the lead in implementing policies. Participant 2 commented as follows:

> *There is paternity leave now. Before, it was difficult, but now it is possible. If this was applied to CEs, maybe it would be quicker [to get support for them] because people have already taken a step forward.*

This was also affirmed by participants 3 and 8, showing some level of readiness for CFWPs in Hong Kong, but it is crucial for the government to take the lead in getting this process started.

In the context of introducing a legal framework to support the future implementation of CFWPs, participant 11, a department head of a financial regulatory institution, highlighted the earlier success of the pilot project on maternity leave before it was legalized for all companies:

> *The Government could test the leave program for CE and other CFWP in some selected departments to identify the need and the possibilities of implementation. After consultations with various sectors of the economy, I believe it would be readily adopted.*

All participants expressed willingness to introduce some forms of CFWPs if the government could first demonstrate the benefits of the measures and set the framework for practice. Participant 13 felt that a pilot phase within the government is a guarantee for the private sector and is one of the many conditions for local businesses to consider CFWPs.

Four participants were aware of the importance of the government communicating the needs of CFWPs to the business sector, particularly for management understanding, including recognition of the role of CEs, potential caregiver burnout, disease progression, management of older people, and stigma/discrimination against CEs. Participant 6, a director at a multinational company, emphasized that the human resources department could take the lead in advocating for CFWPs in the cultural and operational contexts to ensure sustainability and enforcement. Participants 4 and 6 recommended using case studies by the government or those in business to explore the positive changes associated with CFWPs and motivate other companies to create a caregiver-friendly workplace.

Theme 4: Incentives for Adopting CFWPs

Build a Positive Organizational Culture

The improvement in loyalty and sense of belonging that CFWPs bring in the long term was acknowledged. Participant 10, a district director of a financial institution, noted the following:

> *supporting CEs will result in productivity losses—as an immediate loss to the company, but in the long run, the benefits of improving loyalty and sense of belonging would outweigh the costs overall.*

Participants 2 and 5 noted positive spillover effects, specifically, work engagement improving the caregiving experience. CFWPs aim to facilitate a balance between the dual roles, allowing for healthier management of CEs' circumstances, which is an incentive for adoption. Participant 5 commented:

> *It's better to stay at work than to stop for caregiving because you still get to see what's going on in the world, and your brain keeps moving.*

Round-the-clock caregiving duties could be more stressful than taking on dual roles.

CFWPs could be seen as a confidence-building exercise between the management and employees. Management's effort to cultivate a caregiver-friendly workplace culture is a prerequisite for gaining CEs' trust before disclosing their status and asking for consideration for their individual needs. Participant 6 referred to her experience as follows:

> *When my employee came to my office and asked for a day off to take care of his parents, he must have trusted me that he was asking for help. I am very grateful for that trust.*

To strengthen trust, participant 4 suggested using a medical certificate (similar to a carer's passport scheme in the United Kingdom [50]) confirming the caregiving role of the employee, with constant updates on the care recipient's illness, and indications of possible commitments and the durations of such commitments. Participant 3 was aware of the guilt of CEs about not being able to fulfill the dual roles simultaneously. This sense of guilt could not only hinder productivity but also encourage the suppression of emotions and interfere with communication.

Recruitment of Talent and Extension of the Worldwide Network

Most participants explicitly expressed their belief that CFWPs could give companies an advantage in recruiting talent. Participants considered that CFWPs allow for the retention of highly qualified staff and thus more sustainable productivity, as well as a potential reduction in training and recruitment costs. However, participant 12 stressed that the attractiveness of CFWPs might be limited among younger workers.

Participant 14 emphasized that if nonfinancial reporting is considered, CFWPs could make a company a worthwhile investment for potential investors under ESG standards:

> *Impact investing is now becoming more popular. Many global companies have already included ESG in their reporting; we are already lagging behind a lot, although our company has already set up a hotline for general employee concerns.*

For Hong Kong, there are overall benefits of companies taking up CFWPs, namely keeping workers employable rather than out of the workforce for a significant duration, attracting foreign talent to Hong Kong, and reducing public health costs when CEs perform caregiving duties more efficiently and effectively.

XSL•FO
RenderX

Discussion

Principal Findings

Although support for CEs in Hong Kong is currently discretionary and depends on the context of industries, this study identified strong motivations for Hong Kong employers to adopt CFWPs. While improving mental well-being can lead to a direct improvement in productivity [51-54], the benefits of CFWPs may be more extensive than anticipated. Beyond indirect operational incentives, such as employee loyalty and talent retention, CFWPs could make a company an attractive investment for socially conscious investors under ESG standards [55]. Furthermore, CFWPs can have a positive spillover effect with a potentially significant impact on CEs' mental well-being, especially when caregiver-inclusive corporate culture is more explicit than making decisions on a case-by-case basis.

Although this study identified several gaps and barriers to the large-scale adoption of CFWPs in Hong Kong, there is ample precedence of organizational changes in Hong Kong's public and private sectors. For instance, the 8-step theory of change by Kotter [56], which involves the following steps: (1) create urgency, (2) build a guiding coalition, (3) form a strategic vision, (4) communicate the vision, (5) empower others to act on the vision, (6) generate quick wins, (7) sustain momentum, and (8) institutionalize new approaches, has been shown to be applicable across Hong Kong's monolithic civil service organizations [57], publicly listed corporations [58], and NGOs [59,60], which represent the 3 primary stakeholders for the large-scale adoption of CFWPs in Hong Kong.

In line with the theory by Kotter [56], employers and management participants in this study echoed the prerequisites for the government to take the leadership in raising awareness of the needs of CEs (step 1: create urgency) and advance the discussion on CFWPs among stakeholders (step 2: build a guiding coalition) to develop (step 3: form a strategic vision) and implement (step 4: communicate the vision) goal-oriented strategies. Government leadership is indispensable for small and medium enterprises as, unlike large enterprises, they lack the in-house resources and experiences to formulate relevant policies and account for more than 98% of businesses in Hong Kong [61] (step 5: empower others).

The piloting of CFWPs in the management of public institutions seems to be an appropriate starting point, as this is where the bureaucratic burden is most remarkable (step 6: generate quick wins). CFWPs could be better thought through in this complex environment before being rolled out to the private sector and NGOs. Notably, the government has recently launched a 2-year pilot program to subsidize caregivers from low-income families caring for older individuals with moderate and severe impairments (at least 3 hours of care per day) to the tune of HK $3000 (US $386) per care recipient [62]. However, the number of households benefiting from such high thresholds is questionable, and the policy does not extend to caregivers who continue to work and contribute to the economy.

Furthermore, our participants indicated the need for a general CFWP with some specific policies for certain industries (step 7: sustain momentum). For instance, information, caregiving skills, and guide to community care resources; bereavement leave; and caregiver-inclusive corporate culture can apply to all industries. With Hong Kong well-positioned as a knowledge-based economy [63], these CFWPs may be immediately relevant for human capital and knowledge-intensive industries, such as social and personal services, professional and business services, financing, insurance, and real estate. However, flexible working hours or locations depend on industry-specific payment arrangements and the individual requirements of the working environment. For instance, construction, retail, accommodation, and food service sectors may require special considerations due to their demographics (eg, relatively younger workers) and wage payment mode (eg, piecework in construction and hourly in food services).

According to the rated importance, providing information resources, cultivating a caregiver-inclusive corporate culture, and introducing flexible working hours and locations could be piloted within the government structure to provide feedback for future public implementation (step 8: institutionalize new approaches). In addition, legalizing paid or unpaid caregiver leave is necessary. While the government reimburses companies for maternity and paternity leaves, there are other options to facilitate the future legalization of caregiver leave, such as cashing out from the MPF and introducing a legal framework on the qualifications and durations of leaves for CEs. However, cashing out from the MPF could jeopardize the original intention of securing financial security in retirement. Thus, a greater government-industry engagement focused on CFWPs is essential.

Our study underscores the need for consensus among stakeholders on caregiver eligibility criteria, including the care recipient's age, care forms, employment conditions, relationships, and medical needs. Quantifiable parameters can ensure fairness and prevent policy misuse. For instance, the voluntary leave-sharing program in the United States allows workers to donate their unused annual leaves directly to another worker who is experiencing a personal or family medical emergency and has exhausted available paid leaves [64,65]. However, the program requires a significant income loss (at least 24 hours for full-time employees) due to illness after exhausting paid leaves [64,65]. Such a program with objective and fair parameters may be considered for Hong Kong's different industrial sectors. Furthermore, such programs may be instrumental in breaking the status quo that personal or family matters do not belong in the workplace and may enable generous interworker benefit transfers.

There is some international precedence in identifying the pathway to develop and implement CFWPs in Hong Kong. For instance, the advocacy of CFWPs is typically supported by NGOs working with governments and legislatures to inform and influence policy [66-68]. While such advocacy groups do not exist in Hong Kong yet, they can be a vital conduit for engaging with CEs and negotiating with the government to address their most pressing needs. Similarly, Canada offers caregivers the world's longest unpaid compassionate leave (26 weeks) [69,70]. Caregiver leave to attend to an ailing parent is covered by compulsory employment insurance, which is funded

XSL•FO
RenderX

by worker premiums, without burdening businesses directly [70]. Similarly, the Japanese government has subsidized 93 caregiver leave days for the business sector [71].

Notably, the legalization of caregiver leave in Taiwan and Japan was in response to an alarming number of caregivers committing suicide or murdering elderly members under overwhelming "caring fatigue" [72]. Hong Kong should urgently consider CFWPs before caring for the needs of the aging population takes too heavy a toll on the working population. Although successful precedents in other countries can be a good starting point, it is essential to consider the unique needs and expectations of CEs and their employers in Hong Kong. For instance, while unpaid caregiver leave is legalized for CEs in Taiwan and Japan [68,71], employers in this study emphasized the importance of paid leave over unpaid leave in line with maternity and paternity leaves in Hong Kong.

Limitations

Our study has several limitations that must be noted. First, our participants were predominantly recruited using the chain-referral method, which is known to introduce selection bias [73]. Thus, the findings may have lower generalizability, representativeness, or external validity than research conducted with a random sample [73]. However, our goal was to identify perceived barriers to CFWP implementation and galvanize further research in this domain, which may, in turn, identify additional barriers. Thus, in this type of research, the inherent risk of selection bias due to chain-referral sampling is not expected to skew the findings significantly. Second, some industries may be underrepresented in our study. For instance, while data saturation was achieved with at least two participants for each category by size (ie, small and medium enterprises and large corporations) and constitution (ie, NGOs, public institutions, and private companies), there was at least one participant from each of the main industrial sectors in Hong Kong. Third, there was no "third party" interviewer to reduce the impact of the researchers' biases and beliefs [74] about the phenomenon (ie, CFWPs) on the quality of the data collected. Finally, this research was conducted at the peak of the COVID-19 pandemic, which could have influenced the participants' perspectives and positions on CFWPs [75]. In addition, the pandemic conditions may have led to the unavailability of participants with diverse opinions [75].

Conclusion

Employers and management recognize the importance of CFWPs, but current support for caregivers is discretionary and industry-specific. Government leadership is crucial for the future implementation of CFWPs, and piloting such policies within the government can drive institutional change. Ongoing consultation with the business sector on the specific needs of particular industries is essential before legalizing unpaid or paid caregiver leave in Hong Kong.

Acknowledgments

We are grateful to MOV Data Collection Center Limited for assistance with the recruitment and data collection phase. This study was funded by The Chinese University of Hong Kong and the Carers Hong Kong Company Limited, a licensed charity in Hong Kong.

Conflicts of Interest

None declared.

Multimedia Appendix 1
COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.
[PDF File (Adobe PDF File), 372 KB - aging_v8i1e68061_app1.pdf]

Multimedia Appendix 2
Discussion guide.
[DOCX File , 19 KB - aging_v8i1e68061_app2.docx]

References

1. Ni C. A Comparative Study of Filial Piety and Endowment Patterns Between China and the West. In: Proceedings of the 3rd International Conference on Advances in Management Science and Engineering. 2020 Presented at: 3rd International Conference on Advances in Management Science and Engineering; January 12-13, 2020; Wuhan, China. [doi: 10.2991/aebmr.k.200402.023]
2. Yuan X, Wang Q. A tentative study on differences and integration of Sino-Western filial piety culture. Asian Social Science 2011 Jul 31;7(8):97. [doi: 10.5539/ass.v7n8p97]
3. Gallois C, Giles H, Ota H, Pierson H, Ng SH, Lim TS, et al. Intergenerational Communication across the Pacific Rim: The Impact of Filial Piety. In: Adair JG, Dion KL, Lasry JC, editors. Latest Contributions to Cross-cultural Psychology. England, UK: Routledge; 2020:192-211.
4. Sappor G. A comparative study of filial piety in Beijing and the UK. Journal of Education and Culture Studies 2021 Jan 25;5(1):p22. [doi: 10.22158/jecs.v5n1p22]

XSL•FO
RenderX

5. Lin L, Wang Q. Adolescents' filial piety attitudes in relation to their perceived parenting styles: an urban-rural comparative longitudinal study in China. Front Psychol 2021;12:750751 [FREE Full text] [doi: 10.3389/fpsyg.2021.750751] [Medline: 35140649]
6. Montayre J, Saravanakumar P, Zhao I, Holroyd E, Adams J, Neville S. Holding on and letting go: Views about filial piety among adult children living in New Zealand. J Clin Nurs 2022 Oct;31(19-20):2797-2804. [doi: 10.1111/jocn.16098] [Medline: 34687112]
7. Lim AJ, Lau CYH, Cheng C. Applying the dual filial piety model in the United States: a comparison of filial piety between Asian Americans and Caucasian Americans. Front Psychol 2021 Feb 3;12:786609 [FREE Full text] [doi: 10.3389/fpsyg.2021.786609] [Medline: 35185688]
8. Li WW, Singh S, Keerthigha C. A cross-cultural study of filial piety and palliative care knowledge: moderating effect of culture and universality of filial piety. Front Psychol 2021;12:787724 [FREE Full text] [doi: 10.3389/fpsyg.2021.787724] [Medline: 34925189]
9. Chen WW, Wu CW, Zhou Y, Lo YJ. A within-cultural comparison of filial piety beliefs among college students in Hong Kong, Taiwan and Macau. Families, Relationships and Societies 2020;9(3):361-381. [doi: 10.1332/204674319X15615536876902]
10. Lin I, Wolf DA. Division of parent care among adult children. J Gerontol B Psychol Sci Soc Sci 2020 Nov 13;75(10):2230-2239 [FREE Full text] [doi: 10.1093/geronb/gbz162] [Medline: 31883022]
11. Choi JY, Lee SH, Yu S. Exploring factors influencing caregiver burden: a systematic review of family caregivers of older adults with chronic illness in local communities. Healthcare (Basel) 2024 May 13;12(10):1002 [FREE Full text] [doi: 10.3390/healthcare12101002] [Medline: 38786412]
12. Chang AY, Skirbekk VF, Tyrovolas S, Kassebaum NJ, Dieleman JL. Measuring population ageing: an analysis of the Global Burden of Disease Study 2017. Lancet Public Health 2019 Mar;4(3):e159-e167 [FREE Full text] [doi: 10.1016/S2468-2667(19)30019-2] [Medline: 30851869]
13. Del-Pino-Casado R, Priego-Cubero E, López-Martínez C, Orgeta V. Subjective caregiver burden and anxiety in informal caregivers: A systematic review and meta-analysis. PLoS One 2021;16(3):e0247143 [FREE Full text] [doi: 10.1371/journal.pone.0247143] [Medline: 33647035]
14. Gérain P, Zech E. Do informal caregivers experience more burnout? A meta-analytic study. Psychol Health Med 2021 Feb;26(2):145-161. [doi: 10.1080/13548506.2020.1803372] [Medline: 32816530]
15. Pan Y, Chen R, Yang D. The relationship between filial piety and caregiver burden among adult children: A systematic review and meta-analysis. Geriatr Nurs 2022;43:113-123 [FREE Full text] [doi: 10.1016/j.gerinurse.2021.10.024] [Medline: 34864295]
16. Isac C, Lee P, Arulappan J. Older adults with chronic illness - Caregiver burden in the Asian context: A systematic review. Patient Educ Couns 2021 Dec;104(12):2912-2921. [doi: 10.1016/j.pec.2021.04.021] [Medline: 33958255]
17. Zhang L, Han Y, Ma Y, Xu Z, Fang Y. Eastern perspectives on roles, responsibilities and filial piety: A case study. Nurs Ethics 2021 May;28(3):327-345. [doi: 10.1177/0969733020934143] [Medline: 32666888]
18. Skirbekk V, Dieleman JL, Stonawski M, Fejkiel K, Tyrovolas S, Chang AY. The health-adjusted dependency ratio as a new global measure of the burden of ageing: a population-based study. Lancet Healthy Longev 2022 May;3(5):e332-e338 [FREE Full text] [doi: 10.1016/S2666-7568(22)00075-7] [Medline: 36098308]
19. ILO Modelled Estimates and Projections database (ILOEST). World Bank. URL: https://data.worldbank.org/indicator/SL.EMP.TOTL.SP.ZS [accessed 2025-03-21]
20. Henle CA, Fisher GG, McCarthy J, Prince MA, Mattingly VP, Clancy RL. Eldercare and childcare: how does caregiving responsibility affect job discrimination? J Bus Psychol 2019 Mar 7;35(1):59-83. [doi: 10.1007/s10869-019-09618-x]
21. Gaugler JE, Pestka DL, Davila H, Sales R, Owen G, Baumgartner SA, et al. The complexities of family caregiving at work: a mixed-methods study. Int J Aging Hum Dev 2018 Dec;87(4):347-376 [FREE Full text] [doi: 10.1177/0091415017752936] [Medline: 29345147]
22. Yang J, Wang L, Williams A. The role of carer-friendly workplace policies and social support in relation to the mental health of carer-employees. Health & Social Care in the Community 2023 Mar 13;2023:1-11. [doi: 10.1155/2023/5749542]
23. Kayaalp A, Page KJ, Rospenda KM. Caregiver burden, work-family conflict, family-work conflict, and mental health of caregivers: a mediational longitudinal study. Work Stress 2021 Oct 14;35(3):217-240 [FREE Full text] [doi: 10.1080/02678373.2020.1832609] [Medline: 34483432]
24. Gong J, Wang G, Wang Y, Chen X, Chen Y, Meng Q, et al. Nowcasting and forecasting the care needs of the older population in China: analysis of data from the China Health and Retirement Longitudinal Study (CHARLS). The Lancet Public Health 2022 Dec;7(12):e1005-e1013. [doi: 10.1016/s2468-2667(22)00203-1]
25. Kim J, Ingersoll-Dayton B, Kwak M. Balancing eldercare and employment: the role of work interruptions and supportive employers. J Appl Gerontol 2013 Apr;32(3):347-369. [doi: 10.1177/0733464811423647] [Medline: 25474393]
26. Ding R, Dardas A, Wang L, Williams A. Evaluation of a caregiver-friendly workplace program intervention on the health of full-time caregiver employees: a time series analysis of intervention effects. J Occup Environ Med 2020 Oct;62(10):e548-e558 [FREE Full text] [doi: 10.1097/JOM.0000000000001980] [Medline: 32769780]

XSL•FO
RenderX

27. Jimenez D, Nair S, Fu S, Schulz R, Czaja S. Reducing caregiver burden and stress in the workplace: a pilot intervention for working caregivers. The Gerontologist 2015;55:728-729. [doi: 10.1093/geront/gnv364.05]
28. Turgeman-Lupo K, Toker S, Ben-Avi N, Shenhar-Tsarfaty S. The depressive price of being a sandwich-generation caregiver: can organizations and managers help? European Journal of Work and Organizational Psychology 2020 May 18;29(6):862-879. [doi: 10.1080/1359432x.2020.1762574]
29. Lorenz F, Whittaker L, Tazzeo J, Williams A. Availability of caregiver-friendly workplace policies: an international scoping review follow-up study. IJWHM 2021 Jun 25;14(4):459-476. [doi: 10.1108/ijwhm-10-2019-0136]
30. Mofidi A, Tompa E, Williams A, Yazdani A, Lero D, Mortazavi SB. Impact of a caregiver-friendly workplace policies intervention: a prospective economic evaluation. J Occup Environ Med 2019 Jun;61(6):461-468. [doi: 10.1097/JOM.0000000000001564] [Medline: 31166280]
31. Dich N, Lange T, Head J, Rod NH. Work stress, caregiving, and allostatic load: prospective results from the Whitehall II cohort study. Psychosom Med 2015 Jun;77(5):539-547 [FREE Full text] [doi: 10.1097/PSY.0000000000000191] [Medline: 25984826]
32. Policy support to carers in selected places. Research Office-Legislative Council Secretariat. 2020. URL: https://www.legco.gov.hk/research-publications/english/1920rt07-policy-support-to-carers-in-selected-places-20200309-e.pdf [accessed 2025-03-21]
33. Ireson R, Sethi B, Williams A. Availability of caregiver-friendly workplace policies (CFWPs): an international scoping review. Health Soc Care Community 2018 Jan;26(1):e1-e14 [FREE Full text] [doi: 10.1111/hsc.12347] [Medline: 27071368]
34. Matos K, Galinsky E, Bond JT. 2016 National Study of Employers. SHRM. 2017. URL: https://cdn.sanity.io/files/ow8usu72/production/d73a7246cc3a3fef4ad2ece1e3d5aa4eaec2f263.pdf [accessed 2025-03-21]
35. Lee M, French C, Bai X, Yeung C, Taddese H. A Case Study of Caregiver-Friendly Workplace Policy in Hong Kong. London, UK: Imperial College London; 2018.
36. Sethi B, Williams A, Ireson R. Supporting caregiver employees: managers' perspective in Canada. International Journal of Workplace Health Management 2017 Feb 06;10(1):25-41. [doi: 10.1108/ijwhm-03-2016-0016]
37. Wong O, Chau B. The evolving role of filial piety in eldercare in Hong Kong. Asian J Soc Sci 2006;34(4):600-617. [doi: 10.1163/156853106778917790]
38. Leung VWY, Lam CM, Liang Y. Parents' expectations of familial elder care under the neoliberal Hong Kong society. Journal of Family Issues 2019 Sep 27;41(4):437-459. [doi: 10.1177/0192513x19879189]
39. 2016 Population By Census: Summary Results. Census and Statistics Department-The Government of Hong Kong Special Administrative Region. 2017. URL: https://www.censtatd.gov.hk/en/data/stat_report/product/B1120094/att/B11200942016XXXXB0100.pdf [accessed 2025-03-21]
40. Welfare Services: Statistical Highlights - Residential care services for the elderly. Research Office-Legislative Council Secretariat. 2017. URL: https://www.legco.gov.hk/research-publications/english/1617issh22-residential-care-services-for-the-elderly-20170314-e.pdf [accessed 2025-03-21]
41. Panel on Education Subcommittee on Integrated Education. Legislative Council of the Hong Kong Special Administrative Region. 2014. URL: https://www.legco.gov.hk/yr13-14/english/panels/ed/ed_ie/reports/ed_iecb4-1087-1-e.pdf [accessed 2025-03-21]
42. Brodaty H, Donkin M. Family caregivers of people with dementia. Dialogues in Clinical Neuroscience 2022 Apr 01;11(2):217-228. [doi: 10.31887/dcns.2009.11.2/hbrodaty]
43. Labour Force, Employment and Unemployment. Census and Statistics Department-The Government of Hong Kong Special Administrative Region. 2022. URL: https://www.censtatd.gov.hk/en/scode200.html [accessed 2025-03-21]
44. Wong SY, Zou D, Chung RY, Sit RW, Zhang D, Chan D, et al. Regular source of care for the elderly: a cross-national comparative study of Hong Kong with 11 developed countries. J Am Med Dir Assoc 2017 Sep 01;18(9):807.e1-807.e8. [doi: 10.1016/j.jamda.2017.05.009] [Medline: 28647578]
45. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
46. Corbin J, Strauss A. Basics of qualitative research: Techniques and procedures for developing grounded theory (3rd ed). Thousand Oaks, CA: Sage Publications, Inc; 2008.
47. Creswell JW, Creswell JD. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Thousand Oaks, CA: SAGE Publications, Inc; 2022.
48. Ritchie J, Spencer L. Qualitative Data Analysis for Applied Policy Research. In: Huberman AM, Miles MB, editors. The Qualitative Researcher's Companion. Thousand Oaks, CA: SAGE Publications, Inc; 2002:305-329.
49. Srivastava A, Thomson SB. Framework analysis: a qualitative methodology for applied policy research. JOAAG 2008;4(2):72-79 [FREE Full text]
50. Carer's Passport. Cabinet Office, Government People Group. URL: https://www.gov.uk/government/publications/carers-passport [accessed 2025-03-21]

XSL•FO
RenderX

51. Bondar J, Babich Morrow C, Gueorguieva R, Brown M, Hawrilenko M, Krystal JH, et al. Clinical and financial outcomes associated with a workplace mental health program before and during the COVID-19 pandemic. JAMA Netw Open 2022 Jun 01;5(6):e2216349 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.16349] [Medline: 35679044]
52. Peters P. Case study: awareness raising – Novel Zorg and Wonen, the Netherlands. Eurofound. 2011. URL: https://www.eurofound.europa.eu/en/case-studies-0 [accessed 2025-03-21]
53. Humer Z. Case study: awareness raising – Mercator, Slovenia. Eurofound. 2011. URL: https://www.eurofound.europa.eu/en/case-studies-0 [accessed 2025-03-21]
54. Barroso M. Case study: Work adjustments - Axa, Portugal. Ciência_Iscte. 2011. URL: https://ciencia.iscte-iul.pt/publications/case-study-work-adjustments---axa-portugal/48874 [accessed 2025-03-21]
55. Consultancy Study on Needs and Support Required of Carers of Elderly Persons and of Persons with Disabilities in Hong Kong. The Hong Kong Polytechnic University. 2022. URL: https://research.polyu.edu.hk/en/activities/consultancy-study-on-needs-and-support-required-of-carers-of-elde-2 [accessed 2025-03-21]
56. Kotter JP. Leading Change: Why Transformation Efforts Fail. Harvard Business Review. 1995. URL: https://heeoe.hee.nhs.uk/sites/default/files/leading_change_why_transformation_efforts_fail.pdf [accessed 2025-03-21]
57. Huque AS, Jongruck P. Civil service reforms in Hong Kong and Thailand: similar goals, different paths. Public Administration and Policy 2020 Aug 13;23(2):111-123. [doi: 10.1108/pap-03-2020-0015]
58. Li E. A case study of the critical success factors for organizational change of a public listed corporation. Public Administration and Policy 2018 Nov 07;21(2):152-165. [doi: 10.1108/pap-10-2018-010]
59. Wong M, Yeung D. Permission-based leadership and change management in Hong Kong's nongovernment organisations. Nang Yan Business Journal 2014;3(1):67-91 [FREE Full text] [doi: 10.1515/nybj-2015-0006]
60. Nesbit PL, Lam E. Cultural adaptability and organizational change: a case study of a social service organization in Hong Kong. Contemporary Management Research 2014;10(4):12186. [doi: 10.7903/cmr.12186]
61. Support to Small and Medium Enterprises. Trade and Industry Department-The Government of Hong Kong Special Administrative Region. 2022. URL: https://www.tid.gov.hk/english/smes_industry/smes/smes_content.html [accessed 2025-03-21]
62. Scheme on Living Allowance for Carers of Elderly Persons from Low-income Families. Social Welfare Department-The Government of Hong Kong Special Administrative Region. 2022. URL: https://www.swd.gov.hk/en/pubsvc/elderly/cat_careersupp/allowanceforcarersoftheelderly/ [accessed 2025-03-21]
63. Hong Kong as a Knowledge-based Economy - A Statistical Perspective. Census and Statistics Department-The Government of Hong Kong Special Administrative Region. 2019. URL: https://www.censtatd.gov.hk/en/EIndexbySubject.html?pcode=B1110009&scode=590 [accessed 2025-03-21]
64. Fact Sheet: Voluntary Leave Transfer Program. US Office of Personnel Management. URL: https://www.opm.gov/policy-data-oversight/pay-leave/leave-administration/fact-sheets/voluntary-leave-transfer-program/ [accessed 2025-03-21]
65. Bradbury MD. Voluntary shared leave: an exploratory analysis. Public Personnel Management 2019 Nov 27;49(4):532-546. [doi: 10.1177/0091026019889378]
66. Carers Canada. URL: https://www.carerscanada.ca/ [accessed 2025-03-21]
67. Key facts and figures about caring. Carers UK. 2022. URL: https://www.carersuk.org/news-and-campaigns/press-releases/facts-and-figures [accessed 2025-03-21]
68. About Taiwan Association of Family Caregivers. Taiwan Association of Family Caregivers. URL: https://www.familycare.org.tw/about [accessed 2025-03-21]
69. Compassionate care, family caregiving and bereavement leaves. University of British Columbia. URL: https://hr.ubc.ca/benefits/benefit-plan-details/vacation-leaves/compassionate-care-compassionate-or-bereavement-leave [accessed 2025-03-21]
70. Tucker EM, Vosko LF. Designing Paid and Protected Employment Leaves for Short-Term Sickness and Caregiving. IRPP Insight. 2021. URL: https://irpp.org/research-studies/designing-paid-and-protected-employment-leaves-for-short-term-sickness-and-caregiving/ [accessed 2025-03-21]
71. Act on Childcare Leave, Caregiver Leave, and Other Measures for the Welfare of Workers Caring for Children or Other Family Members. Japanese Law Translation. URL: https://www.japaneselawtranslation.go.jp/en/laws/view/2438/en [accessed 2025-03-21]
72. Yuhara E. Japan's 'Nursing Murders.'. Nippon. 2017. URL: https://www.nippon.com/hk/currents/d00251/?pnum=2# [accessed 2025-03-21]
73. Parker C, Scott S, Geddes A. Snowball Sampling. Thousand Oaks, CA: SAGE Publications Ltd; 2019.
74. Sargeant J. Qualitative research part II: participants, analysis, and quality assurance. J Grad Med Educ 2012 Mar;4(1):1-3 [FREE Full text] [doi: 10.4300/JGME-D-11-00307.1] [Medline: 23451297]
75. Roberts JK, Pavlakis AE, Richards MP. It's more complicated than it seems: virtual qualitative research in the COVID-19 era. International Journal of Qualitative Methods 2021 Mar 25;20:16094069211002959. [doi: 10.1177/16094069211002959]

Abbreviations

CE: carer-employee

XSL•FO
RenderX

CFWP: caregiver-friendly workplace policy
ESG: environmental, social, and governance
MPF: mandatory provident fund
NGO: nongovernmental organization

Edited by P Jansons; submitted 28.10.24; peer-reviewed by E Okoyeocha, Y Jiang; comments to author 15.01.25; revised version received 13.02.25; accepted 11.03.25; published 31.03.25.

Please cite as:
Lee MMS, Yeoh EK, Wong ELY
Employers' Perspectives of Caregiver-Friendly Workplace Policies for Caregiver-Employees Caring for Older Adults in Hong Kong: Thematic Analysis
JMIR Aging 2025;8:e68061
URL: https://aging.jmir.org/2025/1/e68061
doi:10.2196/68061
PMID:

©Maggie Man-Sin Lee, Eng-Kiong Yeoh, Eliza Lai-Yi Wong. Originally published in JMIR Aging (https://aging.jmir.org), 31.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Creating User Personas to Represent the Needs of Dementia Caregivers Who Support Medication Management at Home: Persona Development and Qualitative Study

Anna Jolliff[1], MS, LMHC; Priya Loganathar[2], PhD; Richard J Holden[3], PhD; Anna Linden[2], PhD; Himalaya Patel[3,4], PhD; Jessica R Lee[5], MS; Aaron Ganci[6], MFA, PhD; Noll Campbell[7,8,9], MSc, PharmD; Malaz Boustani[8,9,10], MPH, MD; Nicole E Werner[1], PhD

[1]Center for Research and Innovation in Systems Safety, Department of Anesthesiology, Vanderbilt University Medical Center, Nashville, TN, United States

[2]Department of Industrial & Systems Engineering, University of Wisconsin - Madison, Madison, WI, United States

[3]Department of Health & Wellness Design, School of Public Health, Indiana University Bloomington, Bloomington, IN, United States

[4]Health Systems Research Center for Health Information and Communication, Richard L Roudebush Veterans Affairs Medical Center, United States Department of Veterans Affairs, Indianapolis, IN, United States

[5]Department of Human Systems Engineering, Ira A. Fulton Schools of Engineering, Arizona State University, Mesa, AZ, United States

[6]Herron School of Art + Design, Indiana University Indianapolis, Indiana University Indianapolis, Indianapolis, IN, United States

[7]Department of Pharmacy Practice, Purdue University, West Lafayette, IN, United States

[8]Indiana University Center for Aging Research, Regenstrief Institute, Indianapolis, IN, United States

[9]Center for Health Innovation and Implementation Science, Indiana University, Indianapolis, IN, United States

[10]Department of Medicine, Indiana University, Indiana University School of Medicine, Indianapolis, IN, United States

Corresponding Author:
Anna Jolliff, MS, LMHC
Center for Research and Innovation in Systems Safety
Department of Anesthesiology
Vanderbilt University Medical Center
2525 West End Avenue
Suite 800
Nashville, TN, 37203
United States
Phone: 1 615 936 6598
Email: anna.jolliff@vumc.org

Abstract

Background: Caregiver-assisted medication management plays a critical role in promoting medication adherence and quality of life for people living with Alzheimer disease or related dementias (ADRD). The current landscape of digital and nondigital interventions to support medication management does not meet caregivers' needs, contexts, and levels of technological proficiency. Intervention development can be facilitated using personas or data-driven archetypes that represent end users' traits relevant to solution design.

Objective: This study aims to understand the strategies and unmet needs of ADRD caregivers who manage medications and use this understanding to create personas that can inform customized caregiver interventions.

Methods: Participants were self-identified primary caregivers of people with ADRD living with or near the care recipient. Virtual contextual inquiry was completed in three stages: (1) enrollment interview, (2) virtual observation over a 1-week period, and (3) postobservation interview. Codebook thematic analysis of interview transcripts was used to identify dimensions of caregivers' approaches to medication management. A reflexive, team-based affinity diagramming approach was used to identify attributes within these dimensions and group attributes into personas.

Results: Participants (N=25) were aged 62.32 (SD 11.86) years on average, and 17 (68%) of them were female. Caregivers varied across 6 dimensions relevant to medication management: strategies for medication acquisition, medication storage and organization, medication administration, monitoring the care recipient for symptoms, communication with care network regarding

XSL•FO
RenderX

medication, and acquiring information about medication. Three personas were created to represent the observed strategies, unmet needs, and levels of technology use related to medication management: Checklist Cheryl, in Control; Social Sam, Keeps it Simple; and Responsive Rhonda, Stays Relaxed.

Conclusions: Caregivers in this study demonstrated a range of characteristics and values that informed their approach to medication management. They used a combination of technology-based strategies and strategies situated in their physical environments to manage medications. The personas created can be used to inform interventions, such as digital tools, that address caregivers' unmet needs.

(JMIR Aging 2025;8:e63944) doi:10.2196/63944

KEYWORDS

medication adherence; user-centered design; digital health technology; informal caregivers; dementia; artificial intelligence; AI

Introduction

Overview

Every year, family and friends of people living with Alzheimer disease or related dementias (ADRD) around the globe provide 82 billion hours of care [1]. These family caregivers assist in promoting the health and safety of the persons living with ADRD, including helping them manage complex medication routines [2-5]. A US-based survey of family caregivers found that medication management is a challenging process due to the considerable time involved, the complexity of routines, fear of making a mistake, and care recipients' resistance to taking medication [6]. Caregivers of people living with ADRD may face unique challenges; for example, an Australian study of family caregivers found that people living with ADRD may not recognize their need for assistance with managing medication and may feel embarrassed to accept help [7]. International reviews of the literature further suggest that caregivers of people living with ADRD lack training and resources to support the breadth of activities associated with medication management [5,8]. Thus, ADRD caregivers would benefit from interventions that are tailored to their needs, contexts, and levels of technological proficiency to facilitate safe and successful medication management [9]. Intervention development can be facilitated using personas or data-driven archetypes that represent end users' traits relevant to solution design.

Background

Challenges With Caregiver-Assisted Medication Management

Medication management is a complex, multifaceted process [10]. One aspect of caregiver-assisted medication management (henceforth, medication management) is maintaining a constant supply of medications, which is made more complicated by differences in prescription pack sizes, delays in prescription refill, receiving incomplete prescriptions, and return visits to the pharmacy when concerns or errors are detected [10]. Administering medications is complicated by the intricacies of many medication routines, frequent changes to these routines (including the dose, number, and form of medications), and both caregivers' and care recipients' forgetfulness to take medications [5,10]. Caregivers are also responsible for making judgments, such as when to withhold, increase, decrease, or discontinue medication, which involves monitoring the care recipient's symptoms, medication side effects, and whether medications are having their intended effect [10]. Finally, communication with the care team is a key but ambiguous component of medication management; caregivers may struggle to know when to reach out to prescribers or to feel comfortable reaching out and may receive incomplete information about medications from health care providers (eg, side effects to look out for or changes to a prescription) [10,11]. Thus, medication management can represent significant and taxing work for family caregivers.

Caregiver-Assisted Medication Management and Well-Being

Medication management has implications for the health and well-being of both the care recipient and caregiver. For the care recipient, international research spanning the United States, Europe, Canada, Japan, and Australia has linked suboptimal medication management with poorer health [12], poorer management of comorbid conditions [13], costly hospitalizations [14], and premature institutional placement [15]. With assistance from caregivers, medication adherence rates among those with ADRD can match those of their peers without cognitive impairment [3]. Among caregivers, performing medical tasks has been associated with positive outcomes, such as gaining new skills and feeling closer to and less worried about the care recipient [6]. However, medication management has also been associated with greater caregiver burden and stress in international research [7,16-18]. In the United States, nearly 20% of family caregivers who perform medical tasks for the care recipient describe fear of making a mistake, 32% report continuously monitoring for something to go wrong, and 40% report feeling "down, depressed, or hopeless" as an outcome of performing these tasks [6]. Therefore, it is critical to understand the unmet needs of caregivers managing medications as the first step in maximizing the positive and mitigating the negative outcomes for care recipients and caregivers alike.

Limitations of Existing Medication Management Tools

Many tools already exist to support medication management, from low-tech tools (eg, plastic cups and sandwich bags) to high-tech tools (eg, electronic pillboxes) [9]. Some caregivers' needs are met by their existing approaches, and, for them, no intervention may be needed [9]. However, other caregivers report that current low-tech tools, such as pillboxes and reminder notes, do not meet their needs [19]. Pillboxes may not accommodate more complex medication routines, and reminder notes may be ignored over time [19]. While digital interventions

XSL•FO
RenderX

abound, a recent systematic review of mobile apps available in Google Play or iTunes intended to assist with medication management suggested that these, too, may not meet users' needs [20]. The reviewed apps were largely low quality and missing desirable features, such as refill reminders and information about adherence [20]. In a US-based study of family caregivers, participants perceived existing digital interventions for medication management as too elaborate or overwhelming [9]. Thus, it is evident that current interventions to support medication management do not meet the full range of needs, contexts, and interests in technology of the current population of ADRD caregivers. By examining the strategies currently used by caregivers to manage medication, we can better understand the features of interventions that would support medication management [21].

A Call for Customized Solutions

Each caregiver's medication management needs and contexts are different [22]. Needs may vary based on the nature of the routine [23-25], qualities of the care recipient [6,8,26,27], and caregiver characteristics [28,29]. The routine itself is shaped by the variety of medications prescribed, the number of prescribers, and the use of multiple pharmacies [23-25]. Qualities of the care recipient include the stage and symptoms of dementia [8] and any co-occurring medical conditions [6], both of which may change over time, in turn, prompting changes to their medication routine [26,27]. Caregivers vary in their health literacy, access to information, commitment to self-care, and valuing of the care recipient's independence [28], all of which may affect how they manage medication. Caregivers also differ on the degree to which they receive support from others with caregiving, including medication management [8]. One analysis that identified typologies of caregiver support networks found that male caregivers and low-income caregivers in the United States were more likely to be in low support caregiving networks, that is, networks in which caregivers receive less support for themselves (eg, emotional support and support with housework) or for the care recipient (eg, personal care and transportation) [29]. In summary, the intricacies of medication management intersect with the characteristics of caregivers and care recipients to yield distinct requirements of supportive interventions.

Translating Unique Needs Into Solution Design

There is an outstanding need to understand the variable needs of caregivers managing medication on behalf of people living with ADRD and design solutions to meet those needs. Furthermore, in the design of solutions to address these caregiving challenges, it is critical to involve caregivers themselves [30]. User-centered design (UCD) is an internationally recognized approach for integrating future end users into the design of interventions that aim to meet their unique and complex needs [31]. UCD is defined as "an approach to interactive systems development that aims to make systems usable and useful by focusing on the users, their needs and requirements" [31]. UCD engages representatives of the user population in the design of interventions, resulting in interventions that are usable, useful, relevant, and adaptable to caregivers' needs [21,32]. UCD has proven valuable to the design of solutions in the health care domain across the globe to meet diverse patient population needs [30,33-35].

Within the field of UCD, persona design is a specific method in which real participant data are analyzed and distilled into archetypal end-user representations or personas [36]. Personas can be used to represent the range of key user dimensions regarding a certain process (eg, medication management), including relevant demographic characteristics, strategies they use to do their work, and any unmet needs they encounter [37]. In the context of intervention development, personas have certain advantages over other products of qualitative inquiry, such as theory, themes, or concept maps; they offer concise, data-driven representations of end users, allowing designers to ask and answer the question, *What do our end users need?* [36]. Personas may be particularly helpful for informing digital caregiving designs, including home monitoring systems, automated support with caregiving tasks, social robots that provide comfort to care recipients, and social networks that provide support to caregivers [38]. Persona development is a life cycle; initial personas are somewhat crudely defined by their distinguishing features (conception), enriched with storytelling and stakeholder feedback (gestation), and iteratively refined as they are applied (adulthood) [36]. Previous studies in the United States [39,40], Singapore [41], and Germany [42] have used personas to depict important variations among the target populations of interventions and to infer the necessary intervention features to meet users' unique needs.

Study Purpose

This study aimed to answer the following research questions (RQs):

- RQ1—What are the strategies used by caregivers of people living with ADRD who manage medication?
- RQ2—What are the unmet needs of caregivers of people living with ADRD related to medication management?
- RQ3—How can the answers to RQs 1 and 2 be represented as personas to inform intervention development?

Methods

Ethical Considerations

This study was subject to an expedited ethics review and received approval from the University of Wisconsin–Madison Institutional Review Board (2021-0339). Participants provided verbal consent to participate and were told that they could withdraw consent at any time. All data were deidentified before analysis. Participants received US $25 for completing the enrollment interview and another US $25 for completing the postobservation interview.

Study Design

To identify the approaches and unmet needs related to the daily work of medication management for caregivers of people living with ADRD, we adapted a design research method, contextual inquiry, to a virtual form, which we have termed virtual contextual inquiry (VCI) [43,44]. Traditional contextual inquiry combines principles of ethnography with cognitive task analysis, seeking to understand how tasks are completed in naturalistic

XSL•FO
RenderX

environments, including the implicit and tacit knowledge obtained from repeatedly performing contextualized work [45]. In traditional contextual inquiry, researchers observe people in their natural environments while they perform tasks of interest (eg, medication management) as they normally would. Interview probes are inserted unobtrusively before, during, and after observation. In VCI, participants send the research team multimedia messages (eg, photos and textual descriptions of work) during or soon after completing the work [43]. These multimedia messages allow for nearly real-time observation of tasks and serve as a cognitive aid during the postobservation interview [46,47]. As with in-person contextual inquiry, in VCI, data collection is structured across 3 phases: before, during, and after observation.

In this study, VCI was selected over traditional contextual inquiry for myriad reasons. First, medication management takes place in various physical spaces and may vary depending on the day [10]. In order to gain a more accurate picture of the work, observation must be close and sustained. However, to observe participants in person over multiple days would be costly and time consuming for the researcher as well as inconvenient and even uncomfortable for participants [48]. Virtual observation thus strikes a balance between practicality and comprehensiveness [46,49]. Virtual observation may also improve the accuracy of data; participants who are less conscious of being observed may be less likely to alter their typical behaviors [50,51]. Furthermore, virtual observation inherently creates artifacts that can be repeatedly referenced throughout analysis, meaning researchers are less dependent upon memory or written notes. Finally, VCI allows for widespread recruitment, rather than recruiting solely from areas within geographic proximity to the research institution, and permits scholarship to continue in the face of circumstances that prevent in-person observation (eg, the COVID-19 pandemic) [48].

Conceptual Framework

Our conceptual framework for medication management, which influenced the enrollment and postobservation interview process, was based on the findings of Smith et al [10] related to contexts in which medication-related problems can occur for family caregivers managing medications. We were also guided by previous research on crafting personas for intervention design and thus inquired about participants' unmet needs and technology use [36,52].

Study Sample

This study used 2 different methods to enroll a convenience sample of caregivers of people living with ADRD. In the first method, participants were recruited from the Wisconsin Alzheimer's Disease Research Center's registry of self-identified primary caregivers and in the second method from the research team's ADRD caregiver registry. The eligibility criteria included (1) self-identifying as a primary caregiver for a person living with ADRD; (2) aged ≥18 years; (3) managing medication for a person living with ADRD, as defined by our conceptual framework; (4) currently living with or close to a person living with ADRD; (5) having access to a cellphone during the study period; and (6) speaking English.

Study Procedure

Recruitment

Participants who were recruited from the Wisconsin Alzheimer's Disease Research Center emailed the study team if they were interested in participating. The study information sheet was then shared with the participant, and a phone screening was scheduled. Members of the study team's caregiver registry were first contacted by phone, as they had previously indicated a willingness to be contacted regarding research opportunities. For both recruitment methods, after the screening call was complete and participants provided verbal consent to participate, the enrollment interview was scheduled. Those eligible and willing to participate completed the VCI process, which consisted of three stages: (1) enrollment interview, (2) virtual observation over a 1-week period, and (3) postobservation interview. The interviewer (PL) and participants were unknown to each other before recruitment, and participants knew only the interviewer's professional status as a doctoral student.

Enrollment Interview

The enrollment interview was a 1-hour audio-recorded session that took place over a secure videoconferencing software. During the interview, participants were introduced to the text messaging procedure to be performed during the observation period. Participants were provided with examples of multimedia text messages that would meet study criteria, invited to add the study team's phone number as a contact in their phone, and practiced submitting a message to the study phone. Participants were also provided with an instructional video that detailed how to submit a message. The interviewer (PL) cautioned participants against submitting messages that contained identifying information. Following this, participants indicated 2 times in the morning and evening at which they would like to receive prompts to submit a message from the study team. The interviewer scheduled a final interview that would occur after the observation period and approximately 8 to 10 days from the enrollment interview. Finally, participants completed a demographic survey and were asked to describe their daily, weekly, and monthly medication management activities. On the basis of this information, the interviewer created an initial list of potential persona dimensions (eg, *medication administration*), which were brought to the research team for discussion and refinement.

Observation Period

During the observation period, participants sent multimedia text messages (eg, textual descriptions, still photos, videos, and audio files) that described their medication management to a secure phone dedicated exclusively to study-related communication. Participants were asked to send at least 2 multimedia text messages daily for 7 days, capturing the medication management activities they performed. At the beginning and end of each day, participants received a text message reminding them to send messages. All messages were deidentified to protect the identity of the participants and uploaded to a shared and secure cloud-based drive. After each participant's observation period, the research team met to review their observation data and made notes of any multimedia

XSL•FO
RenderX

messages that needed more clarification in the postobservation interview (eg, if the tasks pictured were not clear or if any essential context was missing). In this meeting, the research team also used the caregivers' messages to summarize their approach to medication management, which would later be presented to the caregiver for feedback in the postobservation interview.

Postobservation Interview

We conducted a 1-hour postobservation interview via videoconferencing software within 8 to 10 days after the enrollment visit. During the interview, an interviewer trained in contextual inquiry reviewed the participant's multimedia messages and probed for detail. The interview followed the standard procedure for contextual inquiry in that questions did not follow a predetermined interview guide but instead were specific to the data supplied by participants [53]. For example, one participant (P15) shared an image of the spreadsheet they used to track medications. The interviewer asked, "Can you talk a little bit about how you started using that [spreadsheet] or how you came up with that?" Another caregiver (P33) shared over text that the care recipient had missed a dose of their medication, which meant that the caregiver had to remind them to take it. The interviewer asked, "How did you know he missed his medication, and then how did you go about reminding him?" Finally, the interviewer presented the participant with their summarized approach to medication management for reflection and refinement. The interview was audio recorded and transcribed for analysis.

Analysis

Overview

There is no single, standardized approach to persona development [39-42,53,54]. Our approach combined positivist and nonpositivist approaches [55]. We first used a structured, positivist approach to identify dimensions of medication management. We used this approach under the assumption that there were a finite number of medication management activities in which caregivers could engage, many of which could have been anticipated before analyses. It was reasonable to assume that 2 coders could identify the same activities, regardless of their perspectives.

With that in mind, we first used Microsoft Excel to conduct a codebook thematic analysis of postobservation interview transcripts to identify dimensions of medication management [55]. Initial dimensions were those generated by the interviewer in the enrollment interview. Four coders (PL, AL, JRL, and Mengwei Tang) coded the first 2 postobservation interview transcripts as a team. During that process, the coders refined the initial codebook, adding or modifying dimensions, definitions, and examples. The initial codebook was then brought to the entire research team and refined through consensus-based discussion. The codebook was then applied separately to the remaining transcripts by the 4 coders, such that each transcript was coded by 2 coders. To promote reliability, coders met weekly to discuss their coding and resolve any discrepancies through consensus-based discussion [56]. The coding team also met regularly with the full research team to discuss coding and resolve any discrepancies [56]. Any emergent codes or proposed changes to the codebook were brought to the entire research team for discussion and approval. The resulting final codebook represented an enumeration of the initial persona dimensions.

We then adopted a nonpositivist approach to move from dimensions to interpretive stories, in this case, toward personas and their attributes. This shift was consistent with a philosophical assumption held by the research team that, while it is possible to identify objective and discrete dimensions of medication management (eg, *medication acquisition*), we were unlikely to identify discrete types of people through our analysis. The enumeration of attributes within these dimensions and the assignment of attributes to personas were assumed to be subjective and interpretive, a product of the minds that performed the work [55].

To move from dimensions to attributes, we identified passages coded to a specific dimension and created a sentence-length paraphrase of each passage. These paraphrases were transferred to virtual sticky notes, color-coded by participant ID, and grouped by perceived similarity in an iterative team-based affinity diagramming process. For example, within the dimension *approach to medication acquisition*, sticky notes perceived as describing a proactive approach to medication acquisition were grouped together to form an attribute. If it was perceived that participants with this proactive approach to medication acquisition tended to use a strategic approach to storing medications, these attributes were spatially positioned near each other on the virtual sticky board. Each participant was assigned a distinct color sticky note. If, for example, three distinct colors appeared in the same regions of the whiteboard, it practically meant that those three participants shared an attribute. When multiple participants seemed to share multiple attributes, it warranted the creation of a persona. A process of crystallization was used to add depth and nuance to the final personas [57]. Transcripts were critically reviewed by a research team member (PL) to look for information that could add detail to the developed personas, ensure all perceived attributes had been represented by a persona, and check that no persona too closely resembled any one participant. The cowriting of this manuscript allowed the multidisciplinary research team to compare understandings of how attributes fit or did not fit within each persona. Member reflections (a Big Q alternative to member checking) were also used to refine personas [58]. The research team shared drafts of personas with 7 ADRD caregivers who were not involved in the study and augmented the personas based on their feedback. This approach was not intended to *verify* personas but rather add further depth and nuance to the research team's interpretations.

Research Team Positioning

Members of the research team held professional positions that inevitably influenced their perceptions of attributes and the creation of personas. Four members of the research team were faculty (RJH and NEW) or doctoral students (PL and AL) in the field of human factors engineering, which meant that attributes likely to shape or pose a barrier to work performance were naturally emphasized, along with attributes with

XSL•FO
RenderX

implications for product design. The lead author of this paper (AJ) was a licensed mental health counselor with training in theories of personality and thus brought assumptions about the likelihood that certain attributes would or would not co-occur. The collaboration of a pharmacist (NC), geriatrician (MB), and a research health scientist specializing in human-computer interaction (HP) added depth of understanding to topics relevant to the RQ.

Results

Overview

We reached out to 40 caregivers, and 25 (62%) of them enrolled in the study.

Caregivers who chose not to enroll largely did not respond to our outreach; of those who did respond, 10% (4/40) were ineligible, 3% (1/40) were unable to find time to schedule the enrollment interview, and 3% (1/40) were not interested in sending multimedia messages.

On average, participants were aged 62.3 (SD 11.9) years; 68% (17/25) of the participants were female, and all (25/25, 100%) identified as White and not of Hispanic or Latino origin. The care recipients, on average, were aged 76.7 (SD 11.1) years, and 12% (3/25) were aged <65 years. Among care recipients, 48% (12/25) were female, 92% (23/25) were White, 4% (1/25) were Black or African American, and 4% (1/25) were of Hispanic or Latino origin. All (25/25, 100%) participants demonstrated multiple medication management activities over the course of the 7-day observation period, and 96% (24/25) of the participants submitted messages on at least 5 of the 7 days. Participant characteristics are summarized in Table 1.

XSL•FO
RenderX

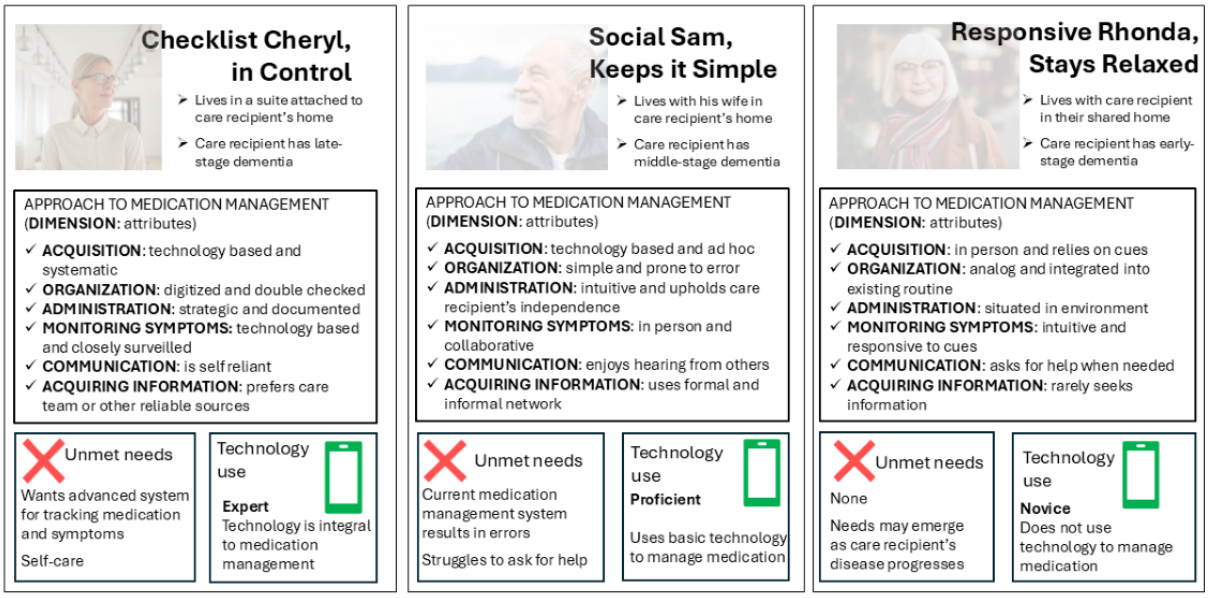
Table 1. Participant demographics (N=25).

Characteristics	Values
Sex, n (%)	
Female	17 (68)
Male	8 (32)
Age (y), mean (SD)	62.3 (11.9)
Race, n (%)	
Black or African American	0 (0)
White	25 (100)
Ethnicity, n (%)	
Hispanic or Latino	0 (0)
Non-Hispanic or Latino	25 (100)
Highest educational attainment, n (%)	
High school diploma or equivalent	1 (4)
Technical school, vocational training, or community college	5 (20)
4-year college	8 (32)
Postcollege education	11 (44)
Employment status, n (%)	
Full time	4 (16)
Part time	3 (12)
Not working	4 (16)
Retired	14 (56)
Annual income (US $), n (%)	
<20,000	1 (4)
20,000-39,999	3 (12)
40,000-$59,999	5 (20)
60,000-79,999	6 (24)
80,000-99,999	5 (20)
>100,000	2 (8)
Undisclosed	3 (12)
Length of time managing medications for care recipient (y), mean (SD)	5.0 (3.1)
State of residence, n (%)	
Wisconsin	16 (64)
California	5 (20)
Other	3 (12)
Undisclosed	1 (4)
Care partner's relationship to the care recipient, n (%)	
Adult child or child-in-law	8 (32)
Spouse or significant other	16 (64)
Other (mother and father)	1 (4)
Care recipient's living arrangement, n (%)	
Lives in the primary caregiver's home	20 (80)
Lives in the care recipient's own home	4 (16)
Other (independent living facility)	1 (4)

XSL•FO
RenderX

We derived 3 personas: Checklist Cheryl, in Control; Social Sam, Keeps it Simple; and Responsive Rhonda, Stays Relaxed. These personas varied across 6 dimensions relevant to medication management: approach to medication acquisition and organization, medication administration, monitoring the care recipient for symptoms, communication with the care network regarding medication, and acquiring information about medication. The presentation of the personas given subsequently is organized by their approach to each dimension of medication management, with illustrative examples and quotations from real study participants. Persona profiles are presented in Figure 1.

Figure 1. Three persona profiles representing caregivers' medication management for patients with Alzheimer disease or related dementias.



Checklist Cheryl, in Control

Overview

Checklist Cheryl, in Control, is characterized by her desire to do things on her own, be in control, and create reliable and precise systems and regimens. Cheryl lives in a suite that is separate from but connected to the home of the care recipient, who has late-stage dementia.

Approach to Acquiring and Organizing Medication

Checklist Cheryl has an organized, technology-based system for acquiring and maintaining a constant supply of medications. She sets up automated refills and mail-based delivery for all medications. She organizes medication in an automated medication dispenser, which she fills every 2 weeks. When filling the dispenser, she verifies the dosage in a notebook that describes each medication as well as a large, printed list of the medications with pictures and a description of each that she keeps on the wall for her reference.

Approach to Administering Medication

Checklist Cheryl continuously improved upon her strategies for medication administration and documents the effectiveness of these strategies in a journal. One participant described iterative strategizing to encourage the care recipient to take her medication:

> *[The care recipient] didn't like [taking medication], and she made bad faces. So, I thought, one day I'll just give her something extra. It started, I mean, we've done like Pringles and different types of chips and crackers and all that kind of stuff.* [P1]

Approach to Monitoring for Symptoms

Checklist Cheryl closely monitors the care recipient for symptoms and writes down her observations in a journal to track changes in behavior. She installs a video surveillance system in the care recipient's home to better monitor any symptoms or changes in the care recipient's behavior:

> *If someone goes in the laundry room, it just brings up my cam and app, and then I can just click on the laundry room, and I can quickly look and see who went in there and what they're doing.* [P1]

Approach to Acquiring Medication Information

Checklist Cheryl seeks formal advice about medications. She contacts the care recipient's health care team with medication questions:

> *The medications he's getting now are from the neurologist. So if I see something is wrong, I call the neurologist for changes or advice.* [P7]

Alternatively, Cheryl may do her own research online. Cheryl does not rely on support groups, friends, or family for medication information.

Approach to Communicating With Care Network

Checklist Cheryl was unlikely to ask her formal care team or network of family and friends for help with medication management:

> *I do all of it. And I order [the medications] myself. Hospice said they would order [the care recipient's medications], but I order [the care recipient's medications] at the same time I order my husband's,*

XSL•FO
RenderX

so I might as well do it all. I'm kind of a control freak. [P40]

Technology Use

Checklist Cheryl is a technology expert. Technology is integral to her medication refill system, medication dispenser, information-seeking approach, and the surveillance system she uses to monitor the care recipient's behavior.

Unmet Needs

Checklist Cheryl wants a better system for tracking the care recipient's changing symptoms and side effects over time. She is also unsure when to communicate symptoms or questions to the formal care team versus relying on the answers she can find through her own research. Finally, she finds that medication management is all consuming and leaves little time for self-care, which often leads to stress.

Social Sam, Keeps It Simple

Overview

Social Sam, Keeps it Simple, is characterized by his attributes of being collaborative, seeking efficiency, and valuing the care recipient's independence. Sam and his wife live in the home of the care recipient, who has middle-stage dementia.

Approach to Acquiring and Organizing Medication

Social Sam has a simple system for monitoring and restocking the medication supply. When he sees a pill bottle running low, he places it near his office computer:

> *[The medication bottles] will be right in front of me on top of the, my desk where my computer is...I'll just get on the website and put in the order right then and there.* [P8]

Sam refills an analog medication organizer on the same evening each week. When possible, Sam incorporates the storage of medication into the care recipient's existing routine. For example, one participant described storing inhalers in the care recipient's backpack:

> *When [the care recipient] used to work, he had that backpack with him 100% of the time, and so that's where the inhalers were. And that, I don't want to change that for him.* [P10]

Approach to Administering Medication

Social Sam's routine for administering medication is so consistent that it has become second nature:

> *It's the first thing that I do when I get up...the first thing I do is grab the pill. I also have medication I'm giving the cats, so I grab that at the same time and put it all out on the coffee table, where I keep it during the day.* [P8]

Because Sam values the care recipient's independence, he is subtle when checking whether medication has been taken:

> *I just pretend I'm doing something else and watch [him take the medication].* [P10]

Approach to Monitoring for Symptoms

Social Sam collaborates with his wife, who is also a member of his care network, to monitor the care recipient for symptoms or changes in behavior:

> *One of the things that's happened is [the care recipient] has gotten GERD, probably from all the medication that she's taking, upsetting her stomach. So she's belching more...And so [my wife] and I talk all the time...There's a constant that we check in, how is she? Is that medication bothering her?* [P25]

To monitor the care recipient's pain symptoms, Sam takes weekly glances at the as-needed pain medication bottle to see if there are fewer pills than there were the previous week.

Approach to Acquiring Medication Information and Communication With the Care Network

Social Sam enjoys receiving information and help from others, both clinicians and nonclinicians. For example, he prefers to pick up medications in person so he can ask the pharmacist questions. Sam uses informal sources of medication information, such as support groups and Facebook:

> *I try to draw on everyone and try to get the best information possible to make the right decision.* [P2]

Technology Use

Social Sam is technology proficient. He uses text messaging to communicate with his wife regarding medication management and uses phone reminders to prompt himself to do medication-related tasks.

Unmet Needs

Social Sam finds his current medication refill reminder system to be too simplistic, not accounting for medications with different refill dates. This insufficiently nuanced system results in Sam occasionally forgetting to refill medications and, consequently, missing doses. Though Sam is naturally collaborative, he finds coordinating medication management activities with other family members and formal care team members to be challenging, as he does not like to bother people.

Responsive Rhonda, Stays Relaxed

Overview

Responsive Rhonda, Stays Relaxed, is characterized by the attributes of being laid-back, nonintrusive, and responding to (rather than preparing for) challenges with medication management as they arise. Rhonda lives with the care recipient, who has early-stage dementia.

Approach to Acquiring Medication

Responsive Rhonda responds to overt reminders and alerts to order medication. For example, she may see the medication supply and notice it is low, or she may look into refilling medications when at the pharmacy for other reasons:

> *Like if I'm going in to pick something up in, at the pharmacy, I will just show them his number, his medical number also and just say, hey, is there anything that's available for pickup?* [P10]

XSL•FO
RenderX

Approach to Organizing and Administering Medication

Responsive Rhonda relies on nontechnological cues to remind herself to organize and administer medications. She stores pills where she regularly performs her nighttime routine:

> *The pill is just part of the setup that I do every night, so it's actually not too hard to remember to put that out.* [P9]

As part of existing daytime routine of Rhonda, she walks past the medication she set out and notices if the care recipient has taken the medications or if they need to be reminded.

Approach to Monitoring for Symptoms

Responsive Rhonda does not have a formal system to monitor the care recipient for symptoms, instead relying on intuition:

> *And if I feel like, this is more intuitive, is if I feel like he's [using Tylenol] much more than normal, then that will trigger me to go check on, you know, to ask, dig in a little more about his pain.* [P1]

Rhonda also relies on the care recipient to report symptoms to her.

Approach to Acquiring Medication Information and Communication With the Care Network

Responsive Rhonda is comfortable relying on others in the care network for help, although at this stage in the care recipient's disease, she rarely needs help. When Rhonda is away, she asks other family members to put the medications out for the care recipient and ensure that they take them. Rhonda seldom sees the need for medication information or advice, but when she does have questions, she is comfortable asking the care recipient's care team.

Technology Use

Responsive Rhonda is a technology novice. Although proficient with mobile phone calls, she does not think of incorporating technology into medication management.

Unmet Needs

Responsive Rhonda does not currently perceive any unmet needs related to medication management.

Discussion

Principal Findings

In this study, we leveraged the UCD method of VCI to understand ADRD caregivers' approaches to and needs for medication management. Caregivers exhibited a range of characteristics and values that informed their approach to medication management.

In addition, caregivers used a combination of technology-based strategies and strategies situated in their physical environments to manage medications. However, these personas must be understood with awareness of the relatively homogenous sample, which may limit generalizability. The average caregiver in this study was a retired or unemployed White woman in her 60s, with college or postcollege education, caring for someone in their 70s who lived in her home in the United States. These personas can be used to inform the UCD of interventions for caregivers with similar attributes by enabling designers to build empathy with future users, remain focused on users' needs, and maintain an empirical foundation for design decisions [36,41,42,54]. Specific implications of designing medication management interventions for the described personas are presented in Table 2.

XSL•FO
RenderX

Table 2. Design implications for Checklist Cheryl, Social Sam, and Responsive Rhonda.

Persona	Unmet needs	Example design implications
Checklist Cheryl, In Control	• No system for tracking changing symptoms and side effects over time • Lacks certainty about when to communicate symptoms and side effects to health care professionals • Medication management is all consuming and leaves little time for self-care	• Adaptive tools to track data (symptoms, side-effects, and strategies) and display trends • AIa-powered database that summarizes academic research • Smartphone apps that combine medication management with self-care for caregivers • Bridge to state and federal caregiver grants
Social Sam, Keeps it Simple	• The current reminder system too simplistic, not accounting for different refill dates • Wants more help and advice from family or friends in his care network	• Text- or email-based reminder systems for reordering different medications • Collaborative platforms for distributing medication management across the care network • Social media platforms, virtual support groups, and online forums for obtaining practical and emotional support from caregiver peers
Responsive Rhonda, Stays Relaxed	• No unmet needs perceived, but may emerge as the care recipient's disease progresses	• Printed educational materials about ADRDb and its progression • Analog tools or infographics that she can situate in her environment • A technology broker who can introduce her to digital medication management supports • Bridge to local aging and disability resource centers

[a]AI: artificial intelligence.

[b]ADRD: Alzheimer disease and related dementias.

The specific attributes of each persona build on the existing literature and extend knowledge in the domain of medication management. In a US-based sample of older adults with chronic heart failure, Holden et al [40] identified the rule-following persona, who has the tendency to respond to uncertainty by seeking reliable rules and clinical expertise. In this study, Checklist Cheryl demonstrates how this tendency extends to medication management processes; Cheryl painstakingly triple checks her medication organization, performs elaborate documentation, and does not trust family or friends to assist in medication management, meaning much of the workload falls on her. Caregivers such as Checklist Cheryl would greatly benefit from state or federal grant programs that pay for respite care or in-home nursing [59]. Such programs require support from the policy makers who set funding priorities and establish budgets. Programs such as these would allow Cheryl to construct a formal support network that she can trust to offset her personal workload.

There are few digital health interventions to support medication management for caregivers, such as Checklist Cheryl, in Control, with some exceptions found internationally [60-62]. As a technology expert, Cheryl may benefit from a multicomponent smartphone app that allows her to, first, view medications and their doses and verify if they have been administered correctly. Second, such an app could enable Checklist Cheryl to document and visualize trends over time, including trends in the care recipient's symptoms, medication side effects, and the effectiveness of strategies [63]. Third, because Cheryl is independent and prefers learning from experts over peers, she may benefit from a feature that uses artificial intelligence to answer medication- and dementia-related questions by synthesizing academic literature [64]. Fourth, Checklist Cheryl needs digital assistance with self-care. A recent international systematic review of mobile apps for family caregivers found only 1 app with a self-care feature [65]. Another systematic review of mobile apps with a caregiver and patient focus found that only 15% of apps included a focus on caregiver self-care [66]. Apps that target self-care needs of other types of busy caregivers (eg, new mothers) could be adapted to fit the self-care needs of caregivers of people living with ADRD [67]. While the typical participant in this study was in their early 60s, younger caregivers are more likely to access the internet and use smartphones and thus may be as interested as Cheryl in digital tools for medication management [68]. Digital tools are also likely to have international applications, as over half of the global population owns a smartphone and about one-third of global smartphone owners have used their phone to access health services [69].

Caregivers such as Social Sam, Keeps it Simple, differ from Checklist Cheryl in that they want simple, rather than complex, medication management systems; deemphasize surveillance in favor of promoting the care recipient's independence; and desire collaboration with other network members as opposed to working alone. While Social Sam's current system is designed for efficiency, it is not sufficiently sophisticated to accommodate polypharmacy, which is increasingly common as the disease progresses [26,70]. Caregivers such as Sam are less technology proficient than those such as Cheryl; he has incorporated the web version of the pharmacy website into his medication acquisition process, but has not used smartphone apps. This may be consistent with other caregivers of Sam's age; research affirms that, though only 61% of those who are aged >65 years own a smartphone, 75% of people in this age range use the internet. Technologically proficient caregivers such as Social

XSL•FO
RenderX

Sam may benefit from websites that support reordering and administering medications, for example, through text- or email-based reminders [71,72].

Social Sam finds his network of family and friends to be helpful, leveraging them to monitor side effects and make clinical judgments. However, similar to other caregivers, those represented by Sam find it challenging to ask others for help [73]. Similar to other male caregivers, in particular, the support network of Sam is small [29]. The current landscape of digital tools for caregivers that support peer interaction is limited [66]. Social Sam would benefit from web-based tools that support distributing caregiving work, and specifically medication management, across a caregiving team [74,75]. Furthermore, Sam would benefit from interventions (eg, virtual support groups and social media platforms) that enable caregivers to connect with peers beyond their geographic area. In the United States, technologies that connect caregivers may have even greater uptake among members of racial and ethnic minority groups, who use social media such as Instagram, Reddit, Snapchat, WhatsApp, and TikTok at higher rates [76]. Reliance on caregiving support from family and friends is also greater among ethnic minority caregivers [77]. More broadly, technologies that connect support networks are needed in low- and middle-income countries, where reliance on unpaid support for dementia caregiving is more common [78].

Caregivers in this study who are represented by Responsive Rhonda, Stays Relaxed, differ from both Checklist Cheryl and Social Sam in that they adopt a fundamentally laid-back management style with little systemization. These caregivers await and respond to medication issues rather than proactively anticipating and addressing them. Similar to the disengaging persona by Holden et al [40], Rhonda does not actively seek out medication information or experiment to improve processes. Responsive Rhonda may represent those caregivers of people in the early stages of dementia, who tend to take fewer medications than those in the later stages [26,27]. As the care recipient's disease progresses, Rhonda would likely benefit from caregiver education and training. Those who are caring for someone in the earlier stages of dementia may be relatively disconnected, with healthcare providers as their sole brokers of support; therefore, health care systems can use personas to create tailored discharge instructions for caregivers such as Responsive Rhonda [79].

While she does not currently report unmet needs related to medication management, as the care recipient's medication regimen becomes more complex, Responsive Rhonda may require a more sophisticated system for medication management. The design of such a system should leverage the tendency of Rhonda to distribute cognition to her physical, rather than digital, environment [80-82]. Rhonda represents caregivers who are novice technology users; therefore, suggestions for Rhonda may also apply to less educated caregivers, who may be less likely and less willing to use digital devices (eg, websites and smartphone apps) to support health, according to a sample of Australian adults [83]. Later in the caregiving journey, caregivers such as Rhonda may benefit from printed materials as well as analog tools (eg, medication organizers) that can be placed as convenient cues throughout the home [80,84]. Rhonda may also benefit from a technology broker who can provide training on using technology to support medication management, for example, by introducing her to specific, informative websites or the mobile app of her preferred pharmacy [85]. This recommendation extends to the 38% of the global population who live in areas in which mobile internet is accessible but who do not currently use it [69].

Limitations and Future Directions

Our study had limitations. While our average participant mirrored the typical US caregiver in many ways [86] and the average care recipient's age matched that of many living with ADRD internationally [87], other sociodemographic characteristics were less representative of caregivers. Our participants had, on average, higher educational attainment, which may result in greater health literacy and enhanced medication management as a result [88]. Furthermore, our sample of caregivers was entirely of White individuals. An analysis of Latino caregivers, for example, would have yielded different personas, as Latino caregivers are more likely to be younger, adult children of care recipients and are less likely to live with the care recipient [86]. One US-based systematic review found that Mexican American, African American, and Korean American caregivers are more committed to keeping care recipients at home rather than long-term care facilities, which may result over time in greater caregiver workload [28]. Moreover, our personas reflect only caregivers who have access to and are willing to participate in research projects, and this group may have unique characteristics (eg, connectedness to academic centers and time to participate in research) [40]. Our personas should not be considered representative of all caregivers managing medication but rather a basis on which future investigations may build [89].

These personas are situated between the conception and gestation stages [36]. That is, while our personas reflect important aspects of our data, have been enriched with storytelling, and deepened through stakeholder involvement, they need further refinement before their full implications can be realized. One limitation of these data is that they reflect only 1 week of caregivers' medication management. A next step in deepening these personas is to collect data about medication management over the course of years, rather than 1 week, to reflect how caregivers' needs change with disease progression. Future work could also incorporate caregivers of people whose cognitive status had been objectively assessed rather than reported by caregivers. While research suggests that family member reports of dementia are often accurate and, when inaccurate, err on the side of underreporting dementia [90,91], it would be important to understand any differences in medication management associated with objectively confirmed dementia diagnoses. Finally, this analysis focused on process- and strategy-related differences in medication management among caregivers. Personas could be further deepened by analyses focused on cognitive, social, or clinical dimensions of medication management [39].

Both our data collection method and the use of personas carry a risk of forgoing the nuance in participant experiences. Our data consist of messages submitted by caregivers and

XSL•FO
RenderX

researchers' follow-up inquiries regarding these messages. When data are collected virtually and participants are asked to provide explanations in retrospect, some degree of detail may be lost (though this is curtailed using messages to aid recall) [92]. Furthermore, in VCI, participants are unlikely to submit messages in inopportune moments, meaning there is a delay not seen in traditional contextual inquiry between performing the task and reporting upon it [46]. While virtual observation may curtail observer effects, it is also possible that participants curate a more flattering self-presentation when communicating digitally versus in person [22]. Last, persona creation inherently emphasizes points of distinction rather than similarities across participants [40]. It is likely that many real caregivers possess qualities of more than one of the personas identified in this study or resemble different personas at different points in their caregiving journey.

Conclusions

This study used a novel methodology, VCI, to understand in rich detail how caregivers manage medications on behalf of those living with ADRD. These data were synthesized into user personas that have direct implications for intervention development. For example, smartphone apps that use artificial intelligence to synthesize information about medication management or meaningfully connect caregivers with others who manage medications could support caregivers who value expertise, such as Checklist Cheryl, or collaboration, such as Social Sam. Before the benefit of medication management personas can be fully realized, these personas must be deepened to reflect nuance and expanded to reflect a more diverse group of caregivers and caregiving experiences. These are vital directions for future research. Once personas have matured in this way, they represent a powerful user-centered strategy to test, evaluate, and disseminate technologies that support medication management among ADRD caregivers.

Acknowledgments

The authors wish to thank Mengwei Tang for her support in formal data analysis. This work was supported by the National Institutes of Health and the National Institute on Aging (grant 1R21AG072418). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

None declared.

References

1. Wimo A, Gauthier S. Global estimates of informal care. Alzheimer's Disease International. URL: https://www.alzint.org/u/global-estimates-of-informal-care.pdf [accessed 2024-05-29]
2. Thorpe JM, Thorpe CT, Kennelty KA, Gellad WF, Schulz R. The impact of family caregivers on potentially inappropriate medication use in noninstitutionalized older adults with dementia. Am J Geriatr Pharmacother 2012 Aug;10(4):230-241 [FREE Full text] [doi: 10.1016/j.amjopharm.2012.05.001] [Medline: 22683399]
3. Cotrell V, Wild K, Bader T. Medication management and adherence among cognitively impaired older adults. J Gerontol Soc Work 2006;47(3-4):31-46. [doi: 10.1300/J083v47n03_03] [Medline: 17062521]
4. Fiss T, Thyrian JR, Fendrich K, van den Berg N, Hoffmann W. Cognitive impairment in primary ambulatory health care: pharmacotherapy and the use of potentially inappropriate medicine. Int J Geriatr Psychiatry 2013 Feb;28(2):173-181 [FREE Full text] [doi: 10.1002/gps.3806] [Medline: 22505357]
5. Gench M, Sawan MJ, Langford A, Gnjidic D. Tools to evaluate medication management for caregivers of people living with dementia: a systematic review. Health Expect 2021 Oct;24(5):1570-1581 [FREE Full text] [doi: 10.1111/hex.13318] [Medline: 34289214]
6. Reinhard SC, Levine C, Samis S. Home alone: family caregivers providing complex chronic care. AARP Public Policy Institute. URL: https://www.aarp.org/content/dam/aarp/research/public_policy_institute/health/home-alone-family-caregivers-providing-complex-chronic-care-rev-AARP-ppi-health.pdf [accessed 2024-10-14]
7. While C, Duane F, Beanland C, Koch S. Medication management: the perspectives of people with dementia and family carers. Dementia (London) 2013 Nov;12(6):734-750. [doi: 10.1177/1471301212444056] [Medline: 24337637]
8. Gillespie R, Mullan J, Harrison L. Managing medications: the role of informal caregivers of older adults and people living with dementia. A review of the literature. J Clin Nurs 2014 Dec;23(23-24):3296-3308 [FREE Full text] [doi: 10.1111/jocn.12519] [Medline: 24354583]
9. LaValley SA, Brady LA, Nguyen M, Jacobs D. Family caregivers of older adults: strategies, tools, and technologies for medication management. J Gerontol Soc Work 2023;66(2):291-302 [FREE Full text] [doi: 10.1080/01634372.2022.2100856] [Medline: 35850634]
10. Smith F, Francis SA, Gray N, Denham M, Graffy J. A multi-centre survey among informal carers who manage medication for older care recipients: problems experienced and development of services. Health Soc Care Community 2003 Mar;11(2):138-145 [FREE Full text] [doi: 10.1046/j.1365-2524.2003.00415.x] [Medline: 14629216]

XSL•FO
RenderX

11. Travis SS, Bernard MA, McAuley WJ, Thornton M, Kole T. Development of the family caregiver medication administration hassles scale. Gerontologist 2003 Jun 01;43(3):360-368. [doi: 10.1093/geront/43.3.360] [Medline: 12810899]
12. Smith D, Lovell J, Weller C, Kennedy B, Winbolt M, Young C, et al. A systematic review of medication non-adherence in persons with dementia or cognitive impairment. PLoS One 2017 Feb 6;12(2):e0170651 [FREE Full text] [doi: 10.1371/journal.pone.0170651] [Medline: 28166234]
13. Hawkins LA, Kilian S, Firek A, Kashner TM, Firek CJ, Silvet H. Cognitive impairment and medication adherence in outpatients with heart failure. Heart Lung 2012;41(6):572-582 [FREE Full text] [doi: 10.1016/j.hrtlng.2012.06.001] [Medline: 22784869]
14. Elliott R, Goeman D, Beanland C, Koch S. Ability of older people with dementia or cognitive impairment to manage medicine regimens: a narrative review. Curr Clin Pharmacol 2015;10(3):213-221 [FREE Full text] [doi: 10.2174/1574884710666150812141525] [Medline: 26265487]
15. Toot S, Swinson T, Devine M, Challis D, Orrell M. Causes of nursing home placement for older people with dementia: a systematic review and meta-analysis. Int Psychogeriatr 2017 Feb;29(2):195-208 [FREE Full text] [doi: 10.1017/S1041610216001654] [Medline: 27806743]
16. Aston L, Hilton A, Moutela T, Shaw R, Maidment I. Exploring the evidence base for how people with dementia and their informal carers manage their medication in the community: a mixed studies review. BMC Geriatr 2017 Oct 18;17(1):242 [FREE Full text] [doi: 10.1186/s12877-017-0638-6] [Medline: 29047339]
17. Maidment ID, Aston L, Moutela T, Fox CG, Hilton A. A qualitative study exploring medication management in people with dementia living in the community and the potential role of the community pharmacist. Health Expect 2017 Oct;20(5):929-942 [FREE Full text] [doi: 10.1111/hex.12534] [Medline: 28105781]
18. Travis SS, McAuley WJ, Dmochowski J, Bernard MA, Kao HF, Greene R. Factors associated with medication hassles experienced by family caregivers of older adults. Patient Educ Couns 2007 Apr;66(1):51-57 [FREE Full text] [doi: 10.1016/j.pec.2006.10.004] [Medline: 17118616]
19. Look KA, Stone JA. Medication management activities performed by informal caregivers of older adults. Res Social Adm Pharm 2018 May;14(5):418-426 [FREE Full text] [doi: 10.1016/j.sapharm.2017.05.005] [Medline: 28528023]
20. Santo K, Richtering SS, Chalmers J, Thiagalingam A, Chow CK, Redfern J. Mobile phone apps to improve medication adherence: a systematic stepwise process to identify high-quality apps. JMIR Mhealth Uhealth 2016 Dec 02;4(4):e132 [FREE Full text] [doi: 10.2196/mhealth.6742] [Medline: 27913373]
21. Evenson S, Muller M, Roth EM. Capturing the context of use to inform system design. J Cogn Eng Decis Mak 2008 Dec 01;2(3):181-203 [FREE Full text] [doi: 10.1518/155534308X377072]
22. Werner NE, Stanislawski B, Marx KA, Watkins DC, Kobayashi M, Kales H, et al. Getting what they need when they need it. Appl Clin Inform 2017 Dec 20;26(01):191-205. [doi: 10.4338/aci-2016-07-ra-0122]
23. Clague F, Mercer SW, McLean G, Reynish E, Guthrie B. Comorbidity and polypharmacy in people with dementia: insights from a large, population-based cross-sectional analysis of primary care data. Age Ageing 2017 Jan 13;46(1):33-39 [FREE Full text] [doi: 10.1093/ageing/afw176] [Medline: 28181629]
24. Mueller C, Molokhia M, Perera G, Veronese N, Stubbs B, Shetty H, et al. Polypharmacy in people with dementia: associations with adverse health outcomes. Exp Gerontol 2018 Jun;106:240-245 [FREE Full text] [doi: 10.1016/j.exger.2018.02.011] [Medline: 29452289]
25. Golchin N, Frank SH, Vince A, Isham L, Meropol SB. Polypharmacy in the elderly. J Res Pharm Pract 2015;4(2):85-88 [FREE Full text] [doi: 10.4103/2279-042X.155755] [Medline: 25984546]
26. Orsel K, Taipale H, Tolppanen AM, Koponen M, Tanskanen A, Tiihonen J, et al. Psychotropic drugs use and psychotropic polypharmacy among persons with Alzheimer's disease. Eur Neuropsychopharmacol 2018 Nov;28(11):1260-1269 [FREE Full text] [doi: 10.1016/j.euroneuro.2018.04.005] [Medline: 30153953]
27. Calvó-Perxas L, Turró-Garriga O, Vilalta-Franch J, Lozano-Gallego M, de Eugenio R, Márquez F, Registry of Dementias of Girona Study Group (ReDeGi Study Group). Trends in the prescription and long-term utilization of antidementia drugs among patients with Alzheimer's disease in Spain: a cohort study using the registry of dementias of Girona. Drugs Aging 2017 Apr;34(4):303-310 [FREE Full text] [doi: 10.1007/s40266-017-0446-x] [Medline: 28258538]
28. Mueller A, Thao L, Condon O, Liebzeit D, Fields B. A Systematic Review of the Needs of Dementia Caregivers Across Care Settings. Home Health Care Manag Pract 2021 Oct 28;34(2):133-147 [FREE Full text] [doi: 10.1177/10848223211056928]
29. Friedman EM, Kennedy DP. Typologies of dementia caregiver support networks: a pilot study. Gerontologist 2021 Nov 15;61(8):1221-1230 [FREE Full text] [doi: 10.1093/geront/gnab013] [Medline: 33585929]
30. Cornet VP, Toscos T, Bolchini D, Rohani Ghahari R, Ahmed R, Daley C, et al. Untold stories in user-centered design of mobile health: practical challenges and strategies learned from the design and evaluation of an app for older adults with heart failure. JMIR Mhealth Uhealth 2020 Jul 21;8(7):e17703 [FREE Full text] [doi: 10.2196/17703] [Medline: 32706745]
31. ISO 9241-210:2019: ergonomics of human-system interaction: part 210: human-centred design for interactive systems. International Organization for Standardization. 2010. URL: https://www.iso.org/standard/77520.html [accessed 2024-05-29]

XSL•FO
RenderX

32. Mackrill J, Marshall P, Payne SR, Dimitrokali E, Cain R. Using a bespoke situated digital kiosk to encourage user participation in healthcare environment design. Appl Ergon 2017 Mar;59(Pt A):342-356 [FREE Full text] [doi: 10.1016/j.apergo.2016.08.005] [Medline: 27890147]
33. Ganeshkumar P, Bhatnagar A, Burka D, Durgad K, Krishna A, Das B, et al. Discovery, development, and deployment of a user-centered point-of-care digital information system to treat and track hypertension and diabetes patients under India Hypertension Control Initiative 2019-2022, India. Digit Health 2024 May 07;10:20552076241250153 [FREE Full text] [doi: 10.1177/20552076241250153] [Medline: 38715975]
34. Young K, Xiong T, Lee R, Banerjee AT, Leslie M, Ko WY, et al. User-centered design and usability of a culturally adapted virtual survivorship care app for Chinese Canadian prostate cancer survivors: qualitative descriptive study. JMIR Hum Factors 2024 Jan 01;11:e49353 [FREE Full text] [doi: 10.2196/49353] [Medline: 38163295]
35. Tandon A, Cobb B, Centra J, Izmailova E, Manyakov N, McClenahan S, et al. A systematic scoping review of studies describing human factors, human-centered design, and usability of sensor-based digital health technologies. medRxiv Preprint posted online February 24, 2024 [FREE Full text] [doi: 10.1101/2024.02.23.24303220]
36. Pruitt J, Adlin T. The Persona Lifecycle: Keeping People in Mind Throughout Product Design. New York, NY: Elsevier; 2010.
37. McCarthy S, O'Raghallaigh P, Woodworth S, Lim YL, Kenny LC, Adam F. An integrated patient journey mapping tool for embedding quality in healthcare service reform. J Decis Syst 2016 Jun 16;25(sup1):354-368 [FREE Full text] [doi: 10.1080/12460125.2016.1187394]
38. Kiselica AM, Hermann GE, Scullin MK, Benge JF. Technology that CARES: enhancing dementia care through everyday technologies. Alzheimers Dement 2024 Dec;20(12):8969-8978 [FREE Full text] [doi: 10.1002/alz.14192] [Medline: 39508340]
39. Holden RJ, Kulanthaivel A, Purkayastha S, Goggins KM, Kripalani S. Know thy eHealth user: development of biopsychosocial personas from a study of older adults with heart failure. Int J Med Inform 2017 Dec;108:158-167 [FREE Full text] [doi: 10.1016/j.ijmedinf.2017.10.006] [Medline: 29132622]
40. Holden RJ, Daley CN, Mickelson RS, Bolchini D, Toscos T, Cornet VP, et al. Patient decision-making personas: an application of a patient-centered cognitive task analysis (P-CTA). Appl Ergon 2020 Sep;87:103107 [FREE Full text] [doi: 10.1016/j.apergo.2020.103107] [Medline: 32310109]
41. Haldane V, Koh JJ, Srivastava A, Teo KW, Tan YG, Cheng RX, et al. User preferences and persona design for an mHealth intervention to support adherence to cardiovascular disease medication in Singapore: a multi-method study. JMIR Mhealth Uhealth 2019 May 28;7(5):e10465 [FREE Full text] [doi: 10.2196/10465] [Medline: 31140445]
42. Engelmann P, Eilerskov N, Thilsing T, Bernardini F, Rasmussen S, Löwe B, et al. Needs of multimorbid heart failure patients and their carers: a qualitative interview study and the creation of personas as a basis for a blended collaborative care intervention. Front Cardiovasc Med 2023;10:1186390 [FREE Full text] [doi: 10.3389/fcvm.2023.1186390] [Medline: 38028443]
43. Werner NE, Campbell NL, Boustani M, Ganci A, Holden RJ. Helping the Helpers - a research protocol for user-centered technology to aid caregiver management of medications for people with Alzheimer's disease and related dementias. Res Social Adm Pharm 2022 Sep;18(9):3680-3686 [FREE Full text] [doi: 10.1016/j.sapharm.2022.03.014] [Medline: 35396151]
44. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340 [FREE Full text] [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g] [Medline: 10940958]
45. Dekker SW, Nyce JM, Hoffman RR. From contextual inquiry to designable futures: what do we need to get there? IEEE Intell Syst 2003 Mar;18(2):74-77 [FREE Full text] [doi: 10.1109/MIS.2003.1193660]
46. Bolger N, Davis A, Rafaeli E. Diary methods: capturing life as it is lived. Annu Rev Psychol 2003 Feb;54(1):579-616. [doi: 10.1146/annurev.psych.54.101601.145030] [Medline: 12499517]
47. Clark RE, Pugh CM, Yates KA, Inaba K, Green DJ, Sullivan ME. The use of cognitive task analysis to improve instructional descriptions of procedures. J Surg Res 2012 Mar;173(1):e37-e42. [doi: 10.1016/j.jss.2011.09.003] [Medline: 22099596]
48. Schlegel EC, Tate JA, Pickler RH, Smith LH. Practical strategies for qualitative inquiry in a virtual world. J Adv Nurs 2021 Oct 04;77(10):4035-4044 [FREE Full text] [doi: 10.1111/jan.15000] [Medline: 34346514]
49. Donnelly T, Rahman-Vyas H, Mandelbaum J, Hatch M, Bhattacharya B. Virtual observation: conducting contextual inquiry remotely. Proc Hum Factors Ergon Soc Annu Meet 2021 Jul 22;10(1):233-237 [FREE Full text] [doi: 10.1177/2327857921101213]
50. Oswald D, Sherratt F, Smith S. Handling the Hawthorne effect: the challenges surrounding a participant observer. Rev Soc Sci 2014;1(1):53-73 [FREE Full text] [doi: 10.21586/ross0000004]
51. Lobo RD, Oliveira MS, Colella JJ, Silva ND, Pastore Junior L, Souza RC. Assessment of the Hawthorne effect during central venous catheter manipulation. Rev Esc Enferm USP 2022;56:e20220125 [FREE Full text] [doi: 10.1590/1980-220X-REEUSP-2022-0125en] [Medline: 36082984]
52. Gaugler JE, Anderson KA, Leach CR, Smith CD, Schmitt FA, Mendiondo M. The emotional ramifications of unmet need in dementia caregiving. Am J Alzheimers Dis Other Demen 2004;19(6):369-380 [FREE Full text] [doi: 10.1177/153331750401900605] [Medline: 15633946]

XSL•FO
RenderX

53. Xhakaj F, Aleven V, McLaren BM. How teachers use data to help students learn: contextual inquiry for the design of a dashboard. In: Proceedings of the 11th European Conference on Technology Enhanced Learning. 2016 Presented at: EC-TEL '16; September 13-16, 2016; Lyon, France p. 340-354 URL: https://link.springer.com/chapter/10.1007/978-3-319-45153-4_26 [doi: 10.1007/978-3-319-45153-4_26]
54. LeRouge C, Ma J, Sneha S, Tolle K. User profiles and personas in the design and development of consumer health technologies. Int J Med Inform 2013 Nov;82(11):e251-e268 [FREE Full text] [doi: 10.1016/j.ijmedinf.2011.03.006] [Medline: 21481635]
55. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a researcher. Int J Transgend Health 2023 Oct 25;24(1):1-6 [FREE Full text] [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
56. O'Connor C, Joffe H. Intercoder reliability in qualitative research: debates and practical guidelines. Int J Qual Methods 2020 Jan 22;19:160940691989922. [doi: 10.1177/1609406919899220]
57. Varpio L, Ajjawi R, Monrouxe LV, O'Brien BC, Rees CE. Shedding the cobra effect: problematising thematic emergence, triangulation, saturation and member checking. Med Educ 2017 Jan 16;51(1):40-50. [doi: 10.1111/medu.13124] [Medline: 27981658]
58. Tracy SJ. Qualitative quality: eight "Big-Tent" criteria for excellent qualitative research. Qual Inq 2010 Oct 01;16(10):837-851. [doi: 10.1177/1077800410383121]
59. How to pay for respite. ARCH National Respite Network and Resource Center. URL: https://archrespite.org/caregiver-resources/how-to-pay-for-respite/ [accessed 2024-10-21]
60. Wagle KC, Skopelja EN, Campbell NL. Caregiver-based interventions to optimize medication safety in vulnerable elderly adults: a systematic evidence-based review. J Am Geriatr Soc 2018 Nov;66(11):2128-2135 [FREE Full text] [doi: 10.1111/jgs.15556] [Medline: 30136714]
61. Kamimura T, Ishiwata R, Inoue T. Medication reminder device for the elderly patients with mild cognitive impairment. Am J Alzheimers Dis Other Demen 2012 Jun;27(4):238-242 [FREE Full text] [doi: 10.1177/1533317512450066] [Medline: 22739031]
62. Shaw CA, Williams KN, Perkhounkova Y, Hein M, Coleman CK. Effects of a video-based intervention on caregiver confidence for managing dementia care challenges: findings from the FamTechCare clinical trial. Clin Gerontol 2020;43(5):508-517 [FREE Full text] [doi: 10.1080/07317115.2020.1729917] [Medline: 32072866]
63. Nie QN, Morrow DG, Azevedo RF, Rogers WA. Developing feedback visualizations to support older adults' medication adherence. Hum Factors Health 2024 Jun;5:100076 [FREE Full text] [doi: 10.1016/j.hfh.2024.100076] [Medline: 40051587]
64. Mostofsky E, Ngo LH, Mittleman MA. The leading AI-powered medical information platform. OpenEvidence. URL: https://openevidence.org [accessed 2024-11-01]
65. Sala-González M, Pérez-Jover V, Guilabert M, Mira JJ. Mobile apps for helping informal caregivers: a systematic review. Int J Environ Res Public Health 2021 Feb 10;18(4):1702 [FREE Full text] [doi: 10.3390/ijerph18041702] [Medline: 33578819]
66. Park JY, Tracy CS, Gray CS. Mobile phone apps for family caregivers: a scoping review and qualitative content analysis. Digit Health 2022 Feb 08;8:20552076221076672 [FREE Full text] [doi: 10.1177/20552076221076672] [Medline: 35154806]
67. Hussain-Shamsy N, Shah A, Vigod SN, Zaheer J, Seto E. Mobile health for perinatal depression and anxiety: scoping review. J Med Internet Res 2020 Apr 13;22(4):e17011 [FREE Full text] [doi: 10.2196/17011] [Medline: 32281939]
68. Faverio M. Share of those 65 and older who are tech users has grown in the past decade. Pew Research Center. URL: https://www.pewresearch.org/short-reads/2022/01/13/share-of-those-65-and-older-who-are-tech-users-has-grown-in-the-past-decade/ [accessed 2024-09-27]
69. Shanahan M, Bahia K. The state of mobile internet connectivity 2024. Global System for Mobile Communications Association. URL: https://www.gsma.com/r/somic/ [accessed 2024-09-27]
70. Chen YH, Lai YC, Wu YC, Sasaki J, Tsai KT, Ho CH. Healthcare utilization in different stages among patients with dementia: a nationwide population-based study. Int J Environ Res Public Health 2021 May 26;18(11):5705 [FREE Full text] [doi: 10.3390/ijerph18115705] [Medline: 34073398]
71. Luong P, Glorioso TJ, Grunwald GK, Peterson P, Allen LA, Khanna A, et al. Text message medication adherence reminders automated and delivered at scale across two institutions: testing the nudge system: pilot study. Circ Cardiovasc Qual Outcomes 2021 May;14(5):e007015 [FREE Full text] [doi: 10.1161/CIRCOUTCOMES.120.007015] [Medline: 33993727]
72. Gautier JF, Boitard C, Michiels Y, Raymond G, Vergez G, Guedon G. Impact of personalized text messages from pharmacists on medication adherence in type 2 diabetes in France: a real-world, randomized, comparative study. Patient Educ Couns 2021 Sep;104(9):2250-2258. [doi: 10.1016/j.pec.2021.02.022] [Medline: 33750593]
73. Dam AE, Boots LM, van Boxtel MP, Verhey FR, de Vugt ME. A mismatch between supply and demand of social support in dementia care: a qualitative study on the perspectives of spousal caregivers and their social network members. Int Psychogeriatr 2018 Jun;30(6):881-892 [FREE Full text] [doi: 10.1017/S1041610217000898] [Medline: 28606195]
74. Linden A, Jolliff A, Gonzalez D, Loganathar P, Elliott C, Zuraw M, et al. "It made me feel like I wasn't alone in the darkness": exploring dementia care network communication and coordination through a digital health platform. J Am Med Inform Assoc 2022 Nov 14;29(12):2003-2013 [FREE Full text] [doi: 10.1093/jamia/ocac172] [Medline: 36166717]

XSL•FO
RenderX

75. Jolliff A, Coller RJ, Kearney H, Warner G, Feinstein JA, Chui MA, et al. An mHealth design to promote medication safety in children with medical complexity. Appl Clin Inform 2024 Jan 17;15(1):45-54 [FREE Full text] [doi: 10.1055/a-2214-8000] [Medline: 37989249]
76. Social media fact sheet. Pew Research Center. URL: https://www.pewresearch.org/internet/fact-sheet/social-media/ [accessed 2024-09-27]
77. Pinquart M, Sörensen S. Ethnic differences in stressors, resources, and psychological outcomes of family caregiving: a meta-analysis. Gerontologist 2005 Feb;45(1):90-106. [doi: 10.1093/geront/45.1.90] [Medline: 15695420]
78. Global action plan on the public health response to dementia 2017 - 2025. World Health Organization. 2017. URL: https://www.who.int/publications/i/item/global-action-plan-on-the-public-health-response-to-dementia-2017---2025 [accessed 2024-04-29]
79. Moore A, Frank C, Chambers LW. Role of the family physician in dementia care. Can Fam Physician 2018 Oct;64(10):717-719 [FREE Full text] [Medline: 30315013]
80. Werner NE, Jolliff AF, Casper G, Martell T, Ponto K. Home is where the head is: a distributed cognition account of personal health information management in the home among those with chronic illness. Ergonomics 2018 Aug;61(8):1065-1078 [FREE Full text] [doi: 10.1080/00140139.2018.1435910] [Medline: 29402181]
81. Mickelson RS, Willis M, Holden RJ. Medication-related cognitive artifacts used by older adults with heart failure. Health Policy Technol 2015 Dec 01;4(4):387-398 [FREE Full text] [doi: 10.1016/j.hlpt.2015.08.009] [Medline: 26855882]
82. Holden RJ, Karanam YL, Cavalcanti LH, Parmar T, Kodthala P, Fowler NR, et al. Health information management practices in informal caregiving: an artifacts analysis and implications for IT design. Int J Med Inform 2018 Dec;120:31-41 [FREE Full text] [doi: 10.1016/j.ijmedinf.2018.09.017] [Medline: 30409344]
83. Lee CM, Thomas E, Norman R, Wells L, Shaw T, Nesbitt J, et al. Educational attainment and willingness to use technology for health and to share health information - the reimagining healthcare survey. Int J Med Inform 2022 Aug;164:104803 [FREE Full text] [doi: 10.1016/j.ijmedinf.2022.104803] [Medline: 35644052]
84. Jolliff AF, Hoonakker P, Ponto K, Tredinnick R, Casper G, Martell T, et al. The desktop, or the top of the desk? The relative usefulness of household features for personal health information management. Appl Ergon 2020 Jan;82:102912 [FREE Full text] [doi: 10.1016/j.apergo.2019.102912] [Medline: 31430599]
85. Hassan AY. Challenges and recommendations for the deployment of information and communication technology solutions for informal caregivers: scoping review. JMIR Aging 2020 Jul 29;3(2):e20310 [FREE Full text] [doi: 10.2196/20310] [Medline: 32723720]
86. Cothran FA, Alto R, Tomaszewski Farias S, Johnson DK, Lara E, Whitmer RA, et al. Caring for the caregivers: an Alzheimer disease research center call to action. Alzheimer Dis Assoc Disord 2022;36(1):92-95 [FREE Full text] [doi: 10.1097/WAD.0000000000000483] [Medline: 34966023]
87. Dementia fact sheet. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2023-09-20]
88. Friis K, Lasgaard M, Rowlands G, Osborne RH, Maindal HT. Health literacy mediates the relationship between educational attainment and health behavior: a Danish population-based study. J Health Commun 2016 Sep 26;21(sup2):54-60. [doi: 10.1080/10810730.2016.1201175] [Medline: 27668691]
89. Young HM, Bell JF, Whitney RL, Ridberg RA, Reed SC, Vitaliano PP. Social determinants of health: underreported heterogeneity in systematic reviews of caregiver interventions. Gerontologist 2020 Feb 14;60(Suppl 1):S14-S28 [FREE Full text] [doi: 10.1093/geront/gnz148] [Medline: 32057083]
90. Cacchione PZ, Powlishta KK, Grant EA, Buckles VD, Morris JC. Accuracy of collateral source reports in very mild to mild dementia of the Alzheimer type. J Am Geriatr Soc 2003 Jun 18;51(6):819-823. [doi: 10.1046/j.1365-2389.2003.51263.x] [Medline: 12757569]
91. Kemp NM, Brodaty H, Pond D, Luscombe G. Diagnosing dementia in primary care: the accuracy of informant reports. Alzheimer Dis Assoc Disord 2002;16(3):171-176. [doi: 10.1097/00002093-200207000-00007] [Medline: 12218648]
92. Kip H, Beerlage-de Jong N, Wentzel J. The contextual inquiry. In: van Gemert-Pijnen L, Kelders SM, Kip H, Sanderman R, editors. eHealth Research, Theory and Development: A Multi-Disciplinary Approach. New York, NY: Routledge; 2018:167-186.

Abbreviations

ADRD: Alzheimer disease or related dementias
RQ: research question
UCD: user-centered design
VCI: virtual contextual inquiry

XSL•FO
RenderX

Edited by F Puga; submitted 03.07.24; peer-reviewed by C Laranjeira, FHC Wong, T Wang; comments to author 02.09.24; revised version received 02.11.24; accepted 07.01.25; published 25.07.25.

Please cite as:
Jolliff A, Loganathar P, Holden RJ, Linden A, Patel H, Lee JR, Ganci A, Campbell N, Boustani M, Werner NE
Creating User Personas to Represent the Needs of Dementia Caregivers Who Support Medication Management at Home: Persona Development and Qualitative Study
JMIR Aging 2025;8:e63944
URL: https://aging.jmir.org/2025/1/e63944
doi:10.2196/63944
PMID:

©Anna Jolliff, Priya Loganathar, Richard J Holden, Anna Linden, Himalaya Patel, Jessica R Lee, Aaron Ganci, Noll Campbell, Malaz Boustani, Nicole E Werner. Originally published in JMIR Aging (https://aging.jmir.org), 25.07.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Stressors Disclosed on Reddit by Caregivers of Older Adults and Social Support Received: Content Analysis

Nova Mengxia Huang[1], MA; Bryan Timothy[2], BA; Shirley S Ho[1], PhD

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

[2]National University of Singapore, Yong Loo Lin School of Medicine, Centre for Behavioural and Implementation Science Interventions, Singapore, Singapore

Corresponding Author:
Shirley S Ho, PhD
Nanyang Technological University
Wee Kim Wee School of Communication and Information
50 Nanyang Ave
Singapore, 639798
Singapore
Phone: 65 6513 8667
Email: tsyho@ntu.edu.sg

Abstract

Background: Informal caregiving of an older adult can be a stressful experience. Anonymous online communities, such as Reddit, provide caregivers with an avenue to disclose their stressors and seek support. However, how caregivers disclose their stressors and the effectiveness of these disclosures in eliciting desired social support remain unclear.

Objective: Guided by the stress process model and optimal matching theory, this study examined stressors disclosed by informal caregivers on Reddit and the types of social support they received in response.

Methods: We conducted a quantitative content analysis of posts and comments in 3 subreddits for informal caregivers of older adults. We identified specific stressors disclosed in the posts, analyzed their co-occurrence, and examined the relationship between these stressors and the presence of different types of social support in corresponding comments.

Results: We collected 737 posts and 3446 comments from 3 subreddits. In the posts, caregivers frequently disclosed care recipients' functional problems (500/737, 67.8%), caregiving relationship strain (279/737, 37.9%), care recipients' emotional problems (195/737, 26.4%), and a scarcity of health and social resources (98/737, 13.3%). Care recipients' functional problems often co-occurred with caregiving relationship strain (χ^2_1=37.2; *P*<.001) and with care recipients' emotional problems (χ^2_1=49.0; *P*<.001). Care recipients' emotional problems also often co-occurred with caregiving relationship strain (χ^2_1=87.2; *P*<.001). Caregiving activity issues (odds ratio [OR] 1.75, 95% CI 1.06-2.89; *P*=.03) and social role conflict (OR 1.64, 95% CI 1.14-2.36; *P*=.008) were positively associated with informational support received in comments, while care recipients' functional problems (OR 1.65, 95% CI 1.32-2.08; *P*<.001) were positively associated with emotional support. Disclosing role overload (OR 1.86, 95% CI 1.03-3.36; *P*=.04) and social role conflict (OR 1.95, 95% CI 1.27-2.77; *P*=.002) increased the chance of receiving esteem support, while disclosing social restriction (OR 2.41, 95% CI 1.03-5.60; *P*=.002) increased the likelihood of receiving network support. In addition, we found that nearly no tangible support (3/3446, 0.09%) was provided in response to self-disclosed stressors.

Conclusions: Caregivers frequently disclosed primary and secondary stressors as well as stressors from the social environment, with various primary stressors usually co-occurring in Reddit posts. These stressors varied in their effectiveness in eliciting different types of social support in comments. Managers and moderators of online communities, such as subreddits, are expected to encourage caregivers' sharing of stressors and promote the exchange of social support with their peers. Health professionals and social workers should develop targeted support interventions to meet the needs of informal caregivers of older adults.

(JMIR Aging 2025;8:e71452) doi:10.2196/71452

KEYWORDS

informal caregivers; Reddit; stressors; social support; optimal matching theory; content analysis

XSL•FO
RenderX

Introduction

Background

Informal caregiving for older adults has become increasingly prevalent due to the growing demand for long-term care in aging societies [1]. Informal caregiving refers to the usually unpaid provision of care to a relative or friend who has a chronic illness, disability, or other long-term health and care needs [2]. Those who assume the role of informal caregivers are most often the person's family members, friends, or neighbors [2,3]. In 2022, a total of 24.1 million American adults served as informal caregivers for family members aged ≥65 years [4]. While informal caregiving is a common experience that may help people gain enrichment and growth in the caregiver role [5,6], individuals face multiple stressors in the process of caregiving. These stressors are aversive life events or conditions that can lead to negative emotions and psychological distress [7]; for instance, informal caregivers of older adults often report financial burdens; work disruptions; and insufficient support from their existing social networks, such as family and friends [8,9]. If not effectively addressed, these stressors can heighten informal caregivers' stress, negatively impacting their well-being [8-10].

The rise of online communities across social media platforms has provided informal caregivers with alternative avenues for coping with stressors [11,12]. Unlike offline support networks, including service providers, families, and friends, online communities offer informal caregivers access to a broader network of individuals with similar experiences, often with greater ease of access [13,14]. Among these, Reddit stands out for its unique features and popularity among informal caregivers. As an assemblage of community forums, Reddit enables users to engage in discussions on a specific topic within peer-to-peer communities (ie, subreddits) [15]. The anonymous nature of Reddit encourages users to discuss sensitive topics freely without fear of judgment, making it easier for caregivers to disclose their challenges [16,17]. Research has shown that Reddit is widely used by informal caregivers to discuss their problems and seek support from peers [18,19]. Therefore, discussions on Reddit provide a novel and valuable source of data for understanding the stressors informal caregivers face and their support needs.

Although a body of empirical studies has investigated informal caregivers' communication on Reddit, such as the topics discussed among informal caregivers for people with Alzheimer disease and related dementias [18,19], research on how informal caregivers of general older adults disclose stressors remains limited. According to the stress process model (SPM), caregivers face distinct stressors, including primary and secondary stressors, across different caregiving contexts [20]. Primary stressors refer to sources of stress directly related to acts of caring for care recipients [20], including objective aspects such as managing care recipients' diseases and symptoms and subjective aspects such as caregivers' feelings of overload or the loss of value [21]. By contrast, secondary stressors are those indirectly related to caregiving tasks but may lead to caregivers' stressful experiences; for example, caregivers often encounter conflict with other family members, financial burdens, and social restrictions due to caregiving [20,22]. Later research also identified additional stressors from the social environment, such as a lack of support and resources from their existing social networks [23]. Given that these various stressors may overwhelm informal caregivers of older adults physically and mentally [20], and there is still limited understanding of specific stressors faced by these informal caregivers, it is critical to examine how this caregiver population discloses these stressors in online communities such as Reddit.

Moreover, the SPM posits that caregiver stressors have a complex and interconnected nature, where one chronic stressor can lead to additional stressors over time as caregivers provide long-term care [20]. Several studies have provided empirical support for this premise; for instance, Koumoutzis et al [24] found that informal caregivers' financial strain was positively associated with their caregiver burden. Similarly, a recent study identified correlations among different challenges disclosed online by informal caregivers of older adults [25]. Despite these findings, further research is needed to deepen our understanding of the relationships among different types of caregiver stressors and to advance the theoretical development of the SPM. In particular, examining how primary, secondary, and social environment stressors interact is essential. A clearer understanding of the co-occurrence and interplay of these stressors can offer a more comprehensive view of the challenges faced by informal caregivers and inform the design of targeted interventions to meet their specific needs.

The SPM also highlights the role of resources, including social (eg, social support) and psychological (eg, self-esteem) resources, in preventing or mitigating the negative impacts of stressors on caregivers' well-being [20]. While past studies have explored the importance of social support in caregiving contexts [26,27], no studies to date have examined how stressors disclosed online by informal caregivers elicit responses from peers. Studies have found that self-disclosed stressors can reveal an individual's difficulties and needs, thus serving as a vehicle for obtaining social support [28]: "the assistance that people provide to others when helping them cope with life changes and situational demands" [29,30]. Cutrona and Russell [31] further developed the optimal matching theory (OMT) to delineate different types of social support, namely action-facilitating support and nurturant support. Action-facilitating support, such as informational (advice and informational input) and tangible (providing goods and services) support, could help remove the stressors people experience, while nurturant support, including emotional (empathy and sympathy), esteem (positive expressions of the self and intrinsic values), and network (enhancing sense of belonging) support, could empower people to cope with the stress experienced [26,27].

The OMT also posits that the acquisition of social support depends on the nature of the stressor disclosed during interpersonal communication. The controllability of disclosed stressful events—the extent to which the stress can be prevented or eliminated—predicts the form of social support provided in intimate relationships [31,32]. Highly controllable stressors are more likely to receive action-facilitating support aimed at eliminating the source of stress. However, recent studies have

XSL•FO
RenderX

challenged the applicability of the OMT in predicting the types of social support exchanged in online contexts; for example, Yip [33] found that online community users often provided immediate support without carefully considering the perceived controllability of personal stressors. As a result, there might be an incongruence between support seekers' needs and the types of support provided by respondents. Given the different stressors faced by informal caregivers and the prevalence of their disclosure of stressors in online communities, there is a need to investigate the relationship between the disclosure of specific stressors and the types of support received in response to infer how well the types of social support provided in comments match stressors disclosed in posts [33].

Objectives

To address these research gaps, this study aims to investigate the interactive dynamics between informal caregivers of older adults on Reddit. This study seeks to examine (1) the different types of stressors disclosed by informal caregivers of older adults in subreddits and their co-occurrence and (2) how the disclosure of various stressors relates to different types of social support received from peers. By conducting a content analysis of caregivers' posts and other caregivers' comments on Reddit, we sought to answer the following research questions (RQs):

- RQ1: What are the caregiver stressors most frequently disclosed by informal caregivers of older adults on Reddit?
- RQ2: What caregiver stressors commonly co-occur in posts by informal caregivers of older adults on Reddit?
- RQ3: How do the caregiver stressors disclosed in posts elicit different types of social support (eg, informational, tangible, emotional, esteem, and network support) in the comments on Reddit?

Methods

Study Overview

To identify the different stressors disclosed by caregivers and the types of social support provided by other caregivers on Reddit, we conducted a quantitative content analysis. Previous research shows that content analysis helps in understanding the narratives shared by people and uncovering underlying meanings. By analyzing the content of narrative posts in online communities, researchers can gain insights into people's experiences and emotions related to specific events [30]. Reddit discussion threads, which included the top-level posts and the comments responding to the top-level posts, were used as the unit of analysis.

Data Collection

We limited our selection to subreddits that discussed the experiences of caring for the general older adult population. While older adults are typically defined as individuals aged ≥60 years, this threshold is not universally fixed in the context of caregiving. Factors such as health status, functional ability, and specific care contexts can influence how "older adult" is defined [34]. Therefore, we did not apply a strict age criterion for older adults in this study.

To identify relevant subreddits, we used the keywords "aging," "older," and "elder" combined with "care" and "caregiver." We then reviewed each subreddit's description to determine whether it served as a forum for informal caregivers of the general older adult population. While we could not verify the exact identities of caregivers or their care recipients from the posts, subreddit descriptions informed our inclusion criteria. Ultimately, we included 3 subreddits: r/AgingParents (approximately 49,000 members), r/eldercare (approximately 9200 members), and r/elderlycare (approximately 2200 members). Among these, r/AgingParents is one of the largest caregiving-related subreddits, ranking in the top 4% of all subreddits by size. It primarily focuses on the experiences of caring for aging parents. The other 2 subreddits—r/eldercare and r/elderlycare—are designed to support informal caregivers of older adults, including those who may not have a child-parent relationship with their care recipients. Due to the anonymous nature of Reddit, we could not confirm personal details such as the nationalities or ages of the posters. However, the frequent use of US-specific medical terms such as "Medicaid" and "Medicare" suggests that a significant portion of users may be based in the United States.

The posts and their corresponding comments from these 3 subreddits were extracted, and we filtered out the posters' responses to other comments as well as comments from other users who responded to comments to the posts from the dataset. In addition, posts that were not written by a caregiver and posts that discussed irrelevant content, such as advertisements or news, were excluded from the final sample. This process resulted in a dataset of 1572 posts with 6984 corresponding comments. Due to the large volume of content, we used simple random sampling—a widely accepted method in human-based content analysis [35,36]—to select a representative subset for in-depth analysis. To achieve this, we used a random number generator in Microsoft Excel to select the posts for analysis. We used a 95% CI and a 3% margin of error to determine the sample size required for the study, which was found to be 636 posts. We oversampled by approximately 20% (n=761) to ensure a sufficient number of posts containing self-disclosures of caregiving experiences. Of these 761 posts, 24 (3.2%) that did not include any self-disclosures of caregiving experiences were excluded from the analysis. The final sample analyzed consisted of 737 posts and 3446 comments.

Coding Scheme

In our content analysis, we examined the stressors from the posts and the types of social support from the comments. The coding scheme for the stressors was based on the SPM and previous works that have examined and identified key stressors that caregivers tend to discuss [20,23]. For the types of social support, we primarily examined the corresponding comments in the dataset. The coding scheme for the different types of social support is based on the social support behavior code in the OMT proposed by Cutrona and Suhr [37], which outlines five different types of social support. Each type was coded as 1 (present) or 0 (absent), as comments and responses can often contain more than one type of social support.

XSL•FO
RenderX

In addition, based on prior studies identifying key factors that influence the receipt of social support [38-40], we also coded for 2 additional variables: support-seeking strategy and the emotional valence of the post. The coding scheme for support-seeking strategy was adapted from Liu et al [39] and included 2 categories: direct requests for support (explicitly asking for help or advice) and indirect requests for support (sharing personal experiences without explicitly asking for help). The emotional valence of the post was categorized into 4 types—positive, negative, neutral, and mixed—based on the sentiments expressed in the posts, as adapted from Li et al [40]. To ensure the fit and suitability of these codes to the dataset, we conducted a preliminary examination with 100 posts and comments. More detailed explanations of the coding scheme, including the definitions and examples of codes, are presented in Tables S1 and S2 in Multimedia Appendix 1.

Coding Procedures

Two coders independently coded a separate set of 150 posts (representing approximately 10% of the total posts and not included in the final analysis) and 684 corresponding comments. Before beginning intercoder reliability coding, the coders met with the authors to discuss the codebook and clarify what each code entailed. After gaining a good understanding of the codes, the coders initiated the reliability coding process. For codes that failed to achieve acceptable intercoder reliability, the coders discussed and came to an agreement on what the code should cover based on its definitions. After discussion and recoding, the coders achieved moderate to high levels of intercoder reliability (κ=0.85, range 0.75-1.00). Intercoder reliability was calculated using the Cohen κ formula in R with the *tidy-comm* package (version 0.4.1), specifically the test_icr function.

Data Analysis

To analyze the frequencies of stressors in posts, we conducted a descriptive analysis, including frequency counts and percentages. To investigate the co-occurrence of the identified stressors, we performed chi-square tests for stressors disclosed in posts. We also calculated the number of posts that contained multiple prominent stressors simultaneously to assess their co-occurrence. In addition, we conducted a series of logistic regression analyses to examine whether the presence of a stressor predicted specific types of social support. Support-seeking strategy and the emotional valence of the post were included as control variables in the logistic regression analysis to account for their potential influence on the receipt of social support.

Ethical Considerations

We obtained approval (IRB-2023-258) from the Nanyang Technological University's integrity review board before data collection. As this study involved a secondary analysis of publicly accessible data from Reddit and did not involve the recruitment of human participants, it was granted exempt status, meaning that informed consent and participant compensation were not required. To ensure confidentiality and anonymity, all identifiable information was removed from the collected data. These procedures adhered strictly to the guidelines of the university's integrity review board to maintain ethical standards in research.

Results

Caregivers' Stressors

RQ1 addressed the types of stressors disclosed by informal caregivers of older adults. The results showed that informal caregivers of older adults disclosed a variety of stressors on Reddit, ranging from primary stressors to secondary stressors, as outlined in the SPM. In addition, users disclosed stressors arising from their social environment. Table 1 elaborates on the stressors identified in the Reddit posts, along with their descriptions.

XSL•FO
RenderX

Table 1. Caregivers' stressors, categories, and descriptions, along with post frequencies and percentages (n=737).

Stressors and categories	Description	Posts, n (%)
Primary stressors		
Care recipients' functional problems	Mention of the care recipient's health problems or diseases, including their cognitive, behavioral, and physical limitations	500 (67.8)
Care recipients' emotional problems	Mention of the care recipient's mental health issues or negative emotions, such as worry, frustration, and anger	195 (26.4)
Caregiving activity issues	Mention of issues related to care recipients' dependence on caregivers regarding caregiving activities, including care recipients' daily needs (activities of daily living) and complex independent living tasks (instrumental activities of daily living)	76 (10.3)
Caregiving relationship strain	Mention of a conflict with the care recipient, such as a medical treatment plan	279 (37.9)
Role overload	Mention of caregivers' perceived hardships that result from excessive demands anchored in caregiving	22 (3)
Secondary stressors		
Social role conflict	Mention of caregiving demands conflicting with job requirements and strains due to other expectations	108 (14.7)
Family role conflict	Mention of conflicts among family members, such as unequal division of older adult care and negative relationships between other family members and care recipients	98 (13.3)
Financial burden	Mention of monetary concerns, including hospital bills, prescription drugs, and physician visit costs	90 (12.2)
Social restriction	Mention of the impact of the disease or caregiving on caregivers' social life, such as missing out on social opportunities with friends, family, and others	48 (6.5)
Stressors from the social environment		
Scarcity of health and social resources	Mention of the unavailability of suitable health and social resources, such as limited day care centers or nursing homes	98 (13.3)

The most frequently disclosed stressor was care recipients' functional problems (500/737, 67.8%), referring to posters' concerns about care recipients' health issues or diseases, including their cognitive, behavioral, and physical problems. The second most frequently disclosed stressor was caregiving relationship strain (279/737, 37.9%), referring to conflict between caregivers and care recipients. Caregivers often mentioned problems in communication or feeling misunderstood by the person they took care of. Care recipients' emotional problems were the third most frequently disclosed stressor (195/737, 26.4%), usually referring to mental health and emotional issues, especially expressions of negative emotions.

The fourth most frequently disclosed stressor was social role conflict (108/737, 14.7%), with caregivers describing how caregiving demands conflicted with their other social roles and led to exhaustion. The scarcity of health and social resources, a stressor from the social environment, was the fifth most prominent stressor disclosed by caregivers, with many posts (98/737, 13.3%) describing the unavailability of suitable health and social resources, such as limited day care centers or nursing homes.

Co-Occurrence of Stressors

RQ2 examined the stressors that often co-occurred in caregivers' posts. Table 2 presents the results of the chi-square analysis. To provide additional context and illustrate the dynamics of highly correlated stressors, representative quotes from relevant Reddit posts are presented in Multimedia Appendix 2.

Care recipients' functional problems often co-occurred with caregiving relationship strain (χ^2_1=37.2; P<.001). Posts that included both care recipients' functional problems and caregiving relationship strain accounted for 30.1% (222/737) of the posts. Care recipients' functional problems often co-occurred with their emotional problems (χ^2_1=49.0; P<.001). Posts that discussed care recipients' functional and emotional problems accounted for 22.9% (169/737) of the posts. Care recipients' emotional problems often co-occurred with caregiving relationship strain (χ^2_1=87.2; P<.001). Posts that discussed care recipients' emotional problems and caregiving relationship strain accounted for 17% (125/737) of the posts. In addition, posts that included care recipients' functional and emotional problems as well as caregiving relationship strain made up 15.3% (113/737) of the posts.

XSL•FO
RenderX

Table 2. Chi-square tests for associations between stressors disclosed in caregivers' posts (n=737).

Variables	Care-recipients' functional problems	Care-recipients' emotional problems	Caregiving activity issues	Role overload	Caregiving relationship strain	Social restriction	Social role conflict	Family role conflict	Financial burden	Scarcity of health and social resources
Care-recipients' functional problems										
Chi-square (*df*)	—[a]	49.0 (1)	0.001 (1)	0.003 (1)	37.2 (1)	5.5 (1)	12.7 (1)	8.2 (1)	0.2 (1)	7.3 (1)
P value	—	<.001	.99	.96	<.001	.02	<.001	.004	.66	.007
Care-recipients' emotional problems										
Chi-square (*df*)	49.0 (1)	—	0.4 (1)	12.2 (1)	87.2 (1)	0.08 (1)	7.3 (1)	1.1 (1)	0.05 (1)	0.5 (1)
P value	<.001	—	.53	<.001	<.001	.78	.007	.30	.83	.47
Caregiving activity issues										
Chi-square (*df*)	0.001 (1)	0.4 (1)	—	0.8 (1)	1.9 (1)	0.02 (1)	0.4 (1)	2.5 (1)	28.4 (1)	26.6 (1)
P value	.99	.53	—	.37	.17	.89	.55	.11	<.001	<.001
Role overload										
Chi-square (*df*)	0.003 (1)	12.2 (1)	0.8 (1)	—	11.5 (1)	5.1 (1)	5.2 (1)	1.8 (1)	2.3 (1)	3.5 (1)
P value	.96	<.001	.37	—	<.001	.02	.02	.19	.13	.06
Caregiving relationship strain										
Chi-square (*df*)	37.2 (1)	87.2 (1)	1.9 (1)	11.5 (1)	—	15.5 (1)	23.7 (1)	3.8 (1)	4.0 (1)	1.7 (1)
P value	<.001	<.001	.17	<.001	—	<.001	<.001	.05	.047	.20
Social restriction										
Chi-square (*df*)	5.5 (1)	0.08 (1)	0.02 (1)	5.1 (1)	15.5 (1)	—	48.3 (1)	19.7 (1)	19.1 (1)	1.6 (1)
P value	.02	.78	.89	.02	<.001	—	<.001	<.001	<.001	.20
Social role conflict										
Chi-square (*df*)	12.7 (1)	7.3 (1)	0.4 (1)	5.2 (1)	23.7 (1)	48.3 (1)	—	16.8 (1)	24.3 (1)	0.1 (1)
P value	<.001	.007	.55	.02	<.001	<.001	—	<.001	<.001	.75
Family role conflict										
Chi-square (*df*)	8.2 (1)	1.1 (1)	2.5 (1)	1.8 (1)	3.8 (1)	19.7 (1)	16.8 (1)	—	6.5 (1)	4.4 (1)
P value	.004	.30	.11	.19	.05	<.001	<.001	—	.01	.04
Financial burden										
Chi-square (*df*)	0.2 (1)	0.05 (1)	28.4 (1)	2.3 (1)	4.0 (1)	19.1 (1)	24.3 (1)	6.5 (1)	—	4.7 (1)
P value	.66	.83	<.001	.13	.047	<.001	<.001	.01	—	.03
Scarcity of health and social resources										

XSL•FO
RenderX

Variables	Care-recipients' functional problems	Care-recipients' emotional problems	Caregiving activity issues	Role overload	Caregiving relationship strain	Social restriction	Social role conflict	Family role conflict	Financial burden	Scarcity of health and social resources
Chi-square (*df*)	7.3 (1)	0.5 (1)	26.6 (1)	3.5 (1)	1.7 (1)	1.6 (1)	0.1 (1)	4.4 (1)	4.7 (1)	—
P value	.007	.47	<.001	.06	.20	.20	.75	.04	.03	—

[a]Not applicable.

The Relationship Between Stressors and Social Support

RQ3 examined the provision of social support in response to various stressors disclosed in caregivers' posts. The results of logistic regression showed whether the presence of specific stressors in posts was significantly associated with the presence of different types of social support in the corresponding comments (Table 3).

Table 3. Logistic regression of types of social support received on disclosed stressors[a].

Independent variables: self-disclosed stressors and categories[b]	Informational support		Emotional support		Esteem support		Network support		Tangible support	
	OR[c] (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Primary stressors										
Care recipients' functional problems	1.01 (0.78-1.31)	.94	1.65 (1.32-2.08)	<.001	1.08 (0.78-1.68)	.49	1.02 (0.48-2.11)	.99	1.00 (0.93-1.08)	.98
Care recipients' emotional problems	0.92 (0.70-1.21)	.54	1.19 (0.97-1.46)	.09	1.15 (0.80-1.66)	.46	0.85 (0.40-1.85)	.69	1.00 (0.92-1.09)	>.99
Caregiving activity issues	1.75 (1.06-2.89)	.03	0.66 (0.46-0.94)	.02	0.28 (0.09-0.70)	.008	1.02 (0.35-3.09)	.94	1.00 (0.90-1.11)	.99
Role overload	0.54 (0.34-0.88)	.01	1.28 (0.86-1.99)	.20	1.86 (1.03-3.36)	.04	1.13 (0.24-5.27)	.88	1.00 (0.83-1.21)	>.99
Caregiving relationship strain	0.88 (0.69-1.13)	.33	1.01 (0.83-1.21)	.99	0.92 (0.61-1.20)	.36	.52 (0.27-1.08)	.08	1.00 (0.93-1.08)	.99
Secondary stressors										
Social restriction	0.94 (0.64-1.38)	.75	1.15 (0.87-1.53)	.34	0.96 (0.57-1.52)	.77	2.41 (1.03-5.60)	.04	1.00 (0.88-1.13)	.99
Social role conflict	1.64 (1.14-2.36)	.008	.88 (0.69-1.12)	.30	1.95 (1.27-2.77)	.002	1.03 (0.45-2.46)	.91	1.00 (0.91-1.10)	.99
Family role conflict	0.79 (0.58-1.10)	.16	1.10 (0.87-1.41)	.42	1.32 (0.87-2.01)	.36	0.41 (0.14-1.21)	.11	1.00 (0.91-1.10)	.99
Financial burden	1.28 (0.86-1.90)	.22	1.28 (0.99-1.67)	.07	0.91 (0.57-1.53)	.76	1.10 (0.46-2.77)	.80	1.00 (0.90-1.11)	.99
Stressors from the social environment										
Scarcity of health and social resources	0.78 (0.55-1.12)	.78	1.24 (0.94-1.63)	.13	1.00 (0.59-1.70)	.99	1.44 (0.59-3.39)	.45	1.00 (0.90-1.11)	.99

[a]Support-seeking strategy and emotional valence of posts were covariates in the models.

[b]The reference category for each stressor is 0 (absent).

[c]OR: odds ratio.

For action-facilitating support, disclosing caregiving activity issues (odds ratio [OR] 1.75, 95% CI 1.06-2.89; *P*=.03) and social role conflict (OR 1.64, 95% CI 1.14-2.36; *P*=.008) significantly positively predicted informational support. However, disclosing role overload (OR 0.54, 95% CI 0.34-0.88; *P*=.01) reduced the likelihood of eliciting informational support.

XSL•FO
RenderX

In addition, we found that nearly no tangible support (3/3446, 0.09%) was provided in response to self-disclosed stressors.

As for nurturant support, disclosing care recipients' functional problems (OR 1.65, 95% CI 1.32-2.08; *P*<.001) significantly positively predicted emotional support. However, disclosing caregiving activity issues (OR 0.66, 95% CI 0.46-0.94; *P*=.02) had a significant negative effect on emotional support provided in user comments.

In addition, disclosing role overload (OR 1.86, 95% CI 1.03-3.36; *P*=.04) and social role conflict (OR 1.95, 95% CI 1.27-2.77; *P*=.002) increased the chance of receiving esteem support. However, disclosing caregiving activity issues (OR 0.28, 95% CI 0.09-0.70; *P*=.008) had a significant negative effect on esteem support provided in user comments, which was similar to their effectiveness in the elicitation of emotional support. In terms of network support, only disclosing social restriction (OR 2.41, 95% CI 1.03-5.60; *P*=.04) increased the likelihood of the receipt of network support in user comments, with other stressors having no significant association.

Discussion

Principal Findings

Using a quantitative content analysis approach, this study examined the stressors disclosed by informal caregivers of older adults on Reddit and the types of social support received in response. Our findings revealed that subreddits were predominantly used by informal caregivers to share primary caregiving stressors, including care recipients' functional problems, caregiving relationship strain, and care recipients' emotional problems. These 3 stressors also often co-occurred in caregivers' posts. Furthermore, the types of stressors disclosed in posts influenced the forms of social support received in the comments. Posts disclosing caregiving activity issues and social role conflicts were positively associated with receiving informational support, whereas posts about care recipients' functional problems were more likely to elicit emotional support. In addition, we identified positive correlations between social role conflict and esteem support as well as between social restriction and network support.

Notably, care recipients' functional problems were mentioned in more than half of the posts (500/737, 67.8%), and their emotional issues were discussed in more than one-fourth of the posts (195/737, 26.4%). This finding aligns with previous studies indicating that caregivers of older adults with Alzheimer disease and related dementias often discussed care recipients' physical and emotional issues on Reddit [18,19]. This implies that care recipients' physical and emotional well-being is a prominent concern among informal caregivers of older adults across different contexts. In addition, the frequent disclosure of caregiving relationship strain also indicated that caregivers struggled with their relationships with care recipients. Aging-related changes in care recipients' behaviors make it harder for caregivers to communicate with them, which impacts the quality of the dyadic relationship [41]. Research has found that these primary stressors can significantly predict caregiver burden and impact the quality of caregiving [42-44]. Therefore, interventions are needed to improve caregivers' health care literacy and communication skills in caregiving.

Social role conflict, a secondary stressor, was also prominent in caregivers' posts. This finding aligns with prior research showing that social role strain is commonly reported among caregivers of older adults both with and without dementia [23,45]. Caregiving demands for older adults consume substantial time and resources, interfering with other daily activities, such as fulfilling job responsibilities and caring for other family members. Interventions aimed at reducing caregiving demands or helping caregivers balance their various social roles are recommended, as unresolved role conflicts can contribute to caregivers' mental health issues [46].

Another key finding is the frequent disclosure of a lack of health and social resources, characterized as a stressor from the social environment. Pearlin et al [20] conceptualized access to caregiving resources as a contextual aspect of the stress process. By contrast, this study identified the prevalence of a scarcity of health and social resources as an additional caregiving stressor, directly contributing to caregiver burden. These findings mirror those of Sun [23] and Giebel et al [47], who found that informal caregivers experienced stress from the social welfare system due to a lack of adequate or suitable supportive services, including long-term care facilities. This finding suggests a need to investigate caregivers' stressors within their social contexts, in addition to primary and secondary stressors.

The co-occurrence of various primary stressors in caregivers' posts, particularly regarding care recipients' functional and emotional problems as well as caregiving relationship strain, supports the premise of the SPM that stressors are often interrelated [20]. Previous empirical studies have demonstrated that caregiver stress is a multifaceted process with multiple contributing factors [48]. In the context of caregiving for the general older adult population, caregivers frequently face interrelated challenges such as managing health-related issues, addressing emotional or behavioral changes, and maintaining interpersonal relationships with care recipients [26]. These challenges can result in significant physical and emotional strain, especially as care recipients' needs become more complex over time [24]. The co-occurrence of multiple stressors underscores the need for further research to explore the specific interactions between caregiver stressors and to develop tailored interventions that address the diverse needs of caregivers for older adults.

When disclosed, caregiver stressors can elicit various forms of social support, such as information, empathy, and encouragement. This aligns with prior research indicating that disclosures of stressful life events on social media platforms are positively associated with receiving social support [49,50]. However, not all stressors showed similar patterns in eliciting social support. According to the OMT, the controllability of disclosed stressful events predicts the type of social support offered [31,32]. Our findings suggest that controllable caregiving-related stressors, such as dealing with logistical issues in caregiving and balancing the caregiver role with other social roles, are more likely to elicit informational support. This might be due to the fact that support providers feel comfortable and knowledgeable enough to provide information based on

XSL•FO
RenderX

their personal experiences. However, other controllable stressors, including financial burden and a scarcity of health and social resources, were not significantly associated with informational support. The reason may stem from Reddit's global community, which limits users' ability to address stressors that require specific local information or familiarity with local social and health services [51]. Similarly, health-related problems, such as care recipients' functional and emotional issues, did not effectively elicit informational support. The nonsignificant relationship between health-related problems and informational support may also be attributed to the low perceived control of the support providers [37]. Many caregivers on Reddit might lack confidence in their knowledge of health issues and therefore refrain from offering information. Instead, these users conveyed emotional support in response to the stressors related to care recipients' health problems.

A notable finding is the imbalance in action-facilitating support provided on Reddit. While informational support was frequently offered in response to caregivers' disclosures, tangible support was almost entirely absent. This highlights a key limitation of online support communities (OSCs): despite their accessibility and ability to connect diverse groups, they struggle to address caregivers' practical needs. This aligns with prior research, which also observed limited tangible support in OSCs [39]. One possible explanation is the weak-tie nature of OSCs, which may limit the trust and rapport required for tangible support exchanges [52]. In addition, the global nature of subreddits makes it difficult for caregivers to exchange tangible support because such interactions often require physical proximity [53]. Given these limitations, offline interactions with traditional support networks, such as health care professionals and friends, remain crucial for meeting caregivers' practical needs.

Investigation into the relationships between caregiver stressors and nurturant support yielded 4 important findings. First, the positive relationship between care recipients' functional problems and emotional support suggests that users tended to express empathy and concern for caregivers struggling with health care issues. One possible explanation is that emotional support is more commonly provided in response to uncontrollable events, such as the progression of an older adult's disease [37]. Second, the negative relationships between caregiving activity issues and emotional and esteem support suggest that the self-disclosure of caregiving activity issues is less likely to elicit emotional and esteem support from peers on subreddits. Taking the positive relationship between caregiving activity issues and informational support into consideration, the results align with the framework of the OMT, which posits that highly controllable stressors are more likely to prompt action-facilitating support rather than emotional support. Caregiving activities, such as daily tasks and logistical issues, may be perceived as routine or inevitable aspects of caregiving rather than emotional challenges requiring empathy or validation. As a result, peers in OSCs may prioritize offering practical advice over emotional reassurance.

Third, caregivers' social role conflicts were more likely to elicit esteem support, which aligns with the needs of support seekers. Previous research has found that conflicts between social roles and the caregiver role are associated with lower self-esteem [54]. The disclosure of social role conflicts in posts may reveal low self-esteem among caregivers, which encourages others to acknowledge their efforts and intrinsic value by providing esteem support. Fourth, caregivers' stressor related to social restrictions elicited network support from others. The disclosure of social restrictions may reveal caregivers' need for social networks, prompting users to provide network support by reminding them of the availability of companionship in subreddits. Thus, online communities such as subreddits have the potential to serve as supplementary social networks for caregivers lacking in-person relationships. However, disclosing other interpersonal relationship issues, including family role conflict, was not effective in eliciting network support. This may be due to the complexity of interpersonal relationships in caregiving, especially those involving family dynamics [55,56]. People might feel uncomfortable providing feedback on such issues due to a lack of background information. These findings highlight the importance of considering the context of support transactions when addressing relational stressors.

Implications

This study has several noteworthy theoretical implications. First, while extant research on informal caregivers' online communication has primarily focused on either their self-disclosure or social support exchange behaviors, we examined how caregiver stressors disclosed online elicited various forms of social support. This approach enriches the research area by delineating the dynamic of communication in anonymous online communities. Second, this study extends the SPM by providing empirical evidence of stressors from the social environment as well as the co-occurrence of various stressors. The results suggest that future research should aim for a comprehensive understanding of caregivers' stress process by considering other stressors from the social environment, such as a scarcity of health and social resources, and by examining the complex dynamics among stressors when investigating caregiver stress.

Third, this study challenges the applicability of the OMT in online contexts by providing insights into the extent to which the types of social support received align with individuals' stressors disclosed on Reddit. The mismatch between disclosed stressors and the social support received, as identified in this research, highlights the importance of recognizing the limitations of online communities in meeting individuals' support needs. The findings also suggest the need to consider potential boundary conditions when examining the applicability of OMT; for example, future research should consider the context of support communication in addressing caregiver stressors, including the characteristics of the stressors and the online communities where supportive communication takes place.

Practically, the findings of this study can inform social workers and public health practitioners about strategies to meet the needs of informal caregivers of older adults. First, it can inform social workers and public health practitioners regarding the development of targeted interventions to address caregivers' various stressors. This can include creating caregiving relationship–focused interventions to support high-quality family caregiving, enhancing health and medical education among

XSL•FO
RenderX

caregivers, and advocating for more appropriate community-based support resources. Second, managers and moderators of OSCs, such as subreddits, can help create spaces and environments that encourage caregivers' sharing of stressors and promote the exchange of social support with their peers. Third and last, they could also introduce professionals into online communities to answer caregivers' questions that peers might not be able to answer, such as those surrounding financial support and health and social resources.

Limitations

Several limitations of this study should be noted. First, the anonymity of Reddit prevents us from determining crucial contextual information, such as caregivers' demographics (eg, age and race), country of origin (eg, within or outside the United States), caregiving situation (eg, sole caregivers or sharing caregiving responsibilities with other family members), and the exact age range of care recipients. Without this information, it is difficult to fully understand the specificity and nuances of the struggles experienced by caregivers across these different contexts. Future research should incorporate more contextual information to allow for a better understanding of these unique struggles, which would aid in tailoring supportive interventions for specific caregiver populations. Second, we acknowledge the ongoing debate regarding the use of search terms such as "elder," which some argue may reflect ageist language [57], although it is commonly used in caregiving contexts. Future research should consider adopting alternative search terms that are inclusive and nonageist. Third, communication patterns on Reddit may differ from those on other platforms, limiting the generalizability of our findings. Future research could compare caregivers' discussions across platforms to provide a more comprehensive understanding of their stressors. Finally, our content analysis provides descriptive insights but does not address caregiver satisfaction with subreddit discussions or their benefits. Future research should understand users' perceptions of subreddits and the outcomes associated with subreddit use through interviews, surveys, or other mixed methods.

Conclusions

This content analysis study found that subreddits for informal caregivers allow users to share their caregiving challenges, including primary and secondary stressors and stressors from the social environment. These disclosures help reveal needs and garner social support from peers. While some controllable caregiving issues effectively elicit informational support, other controllable stressors related to health problems and local policy and service gaps seem to fall short in obtaining the desired informational support. In addition, secondary stressors such as role conflicts and social constraints successfully prompt nurturant support, including esteem and network support, from other users.

Acknowledgments

This research is supported by the Ministry of Education, Singapore, under its Academic Research Fund Tier 1 Grant(RG111/22).

Data Availability

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

Conflicts of Interest

None declared.

Multimedia Appendix 1
Key variables and their descriptions and examples.
[DOCX File , 21 KB - aging_v8i1e71452_app1.docx]

Multimedia Appendix 2
Representative quotes with highly correlated stressors.
[DOCX File , 19 KB - aging_v8i1e71452_app2.docx]

References

1. Ageing and health. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2024-08-30]
2. Revenson TA, Griva K, Luszczynska A, Morrison V, Panagopoulou E, Vilchinsky N, et al. Caregiving in the Illness Context. London, UK: Palgrave Pivot; 2016.
3. Zarzycki M, Morrison V. Getting back or giving back: understanding caregiver motivations and willingness to provide informal care. Health Psychol Behav Med 2021 Jul 13;9(1):636-661 [FREE Full text] [doi: 10.1080/21642850.2021.1951737] [Medline: 34345534]
4. Wolff JL, Cornman JC, Freedman VA. The number of family caregivers helping older US adults increased from 18 million to 24 million, 2011-22. Health Aff (Millwood) 2025 Feb 01;44(2):187-195. [doi: 10.1377/hlthaff.2024.00978] [Medline: 39899774]

XSL•FO
RenderX

5. Mishra N, Datti RS, Tewari A, Sirisety M. Exploring the positive aspects of caregiving among family caregivers of the older adults in India. Front Public Health 2023 May 18;11:1059459 [FREE Full text] [doi: 10.3389/fpubh.2023.1059459] [Medline: 37275495]
6. Pysklywec A, Plante M, Auger C, Mortenson WB, Eales J, Routhier F, et al. The positive effects of caring for family carers of older adults: a scoping review. Int J Care Caring 2020 Aug 01;4(3):349-375. [doi: 10.1332/239788220x15925902138734]
7. Marum G, Clench-Aas J, Nes RB, Raanaas RK. The relationship between negative life events, psychological distress and life satisfaction: a population-based study. Qual Life Res 2014 Mar 13;23(2):601-611. [doi: 10.1007/s11136-013-0512-8] [Medline: 24026629]
8. Friedman EM, Trail TE, Vaughan CA, Tanielian T. Online peer support groups for family caregivers: are they reaching the caregivers with the greatest needs? J Am Med Inform Assoc 2018 Sep 01;25(9):1130-1136 [FREE Full text] [doi: 10.1093/jamia/ocy086] [Medline: 30016449]
9. Committee on Family Caregiving for Older Adults, Board on Health Care Services, Health and Medicine Division, National Academies of Sciences, Engineering, and Medicine. Families Caring for an Aging America. Washington, DC: National Academies Press; 2016.
10. Penning MJ, Wu Z. Caregiver stress and mental health: impact of caregiving relationship and gender. Gerontologist 2016 Dec 17;56(6):1102-1113. [doi: 10.1093/geront/gnv038] [Medline: 26035875]
11. Kelly KJ, Doucet S, Luke A, Azar R, Montelpare W. Exploring the use of a Facebook-based support group for caregivers of children and youth with complex care needs: qualitative descriptive study. JMIR Pediatr Parent 2022 Jun 07;5(2):e33170 [FREE Full text] [doi: 10.2196/33170] [Medline: 35671082]
12. Shoults CC, Rutherford MW, Kemp AS, Addicott MA, Brown A, Greene CJ, et al. Analysis of caregiver burden expressed in social media discussions. Int J Environ Res Public Health 2023 Jan 20;20(3):1933 [FREE Full text] [doi: 10.3390/ijerph20031933] [Medline: 36767297]
13. Tanis M. Health-related on-line forums: what's the big attraction? J Health Commun 2008 Oct 28;13(7):698-714. [doi: 10.1080/10810730802415316] [Medline: 18958781]
14. Tong ST, Heinemann-Lafave D, Jeon J, Kolodziej-Smith R, Warshay N. The use of pro-ana blogs for online social support. Eat Disord 2013 Oct;21(5):408-422. [doi: 10.1080/10640266.2013.827538] [Medline: 24044597]
15. De Choudhury M, De S. Mental health discourse on Reddit: self-disclosure, social support, and anonymity. ICWSM 2014 May 16;8(1):71-80. [doi: 10.1609/icwsm.v8i1.14526]
16. Ammari T, Schoenebeck S, Romero D. Self-declared throwaway accounts on Reddit: how platform affordances and shared norms enable parenting disclosure and support. Proc ACM Hum Comput Interact 2019 Nov 07;3(CSCW):1-30. [doi: 10.1145/3359237]
17. Trepte S, Scharkow M, Dienlin T. The privacy calculus contextualized: the influence of affordances. Comput Hum Behav 2020 Mar;104:106115. [doi: 10.1016/j.chb.2019.08.022]
18. Cheng K, Yin Z. "Please advise": understanding the needs of informal caregivers of people with Alzheimer's disease and related dementia from online communities through a structured topic modeling approach. AMIA Jt Summits Transl Sci Proc 2022;2022:149-158 [FREE Full text] [Medline: 35854737]
19. Pickett AC, Valdez D, Sinclair KL, Kochell WJ, Fowler B, Werner NE. Social media discourse related to caregiving for older adults living with Alzheimer disease and related dementias: computational and qualitative study. JMIR Aging 2024 Jun 19;7:e59294 [FREE Full text] [doi: 10.2196/59294] [Medline: 38896462]
20. Pearlin LI, Mullan JT, Semple SJ, Skaff MM. Caregiving and the stress process: an overview of concepts and their measures. Gerontologist 1990 Oct 01;30(5):583-594. [doi: 10.1093/geront/30.5.583] [Medline: 2276631]
21. Zarit SH, Bottigi K, Gaugler JE. Caregivers, stress and. In: Fink G, editor. Encyclopedia of Stress. Cambridge, UK: Academic Press; 2007.
22. Hilgeman MM, Durkin DW, Sun F, DeCoster J, Allen RS, Gallagher-Thompson D, et al. Testing a theoretical model of the stress process in Alzheimer's caregivers with race as a moderator. Gerontologist 2009 Apr;49(2):248-261 [FREE Full text] [doi: 10.1093/geront/gnp015] [Medline: 19363019]
23. Sun F. Caregiving stress and coping: a thematic analysis of Chinese family caregivers of persons with dementia. Dementia (London) 2014 Nov;13(6):803-818. [doi: 10.1177/1471301213485593] [Medline: 24339084]
24. Koumoutzis A, Cichy KE, Dellmann-Jenkins M, Blankemeyer M. Age differences and similarities in associated stressors and outcomes among young, midlife, and older adult family caregivers. Int J Aging Hum Dev 2021 Jun 14;92(4):431-449. [doi: 10.1177/0091415020905265] [Medline: 32054288]
25. Huang NM, Wong LZ, Ho SS, Timothy B. Understanding challenges and emotions of informal caregivers of general older adults and people with Alzheimer disease and related dementia: comparative study. J Med Internet Res 2025 Feb 28;27:e54847 [FREE Full text] [doi: 10.2196/54847] [Medline: 40053723]
26. Huang NM, Ho SS. Online support community use for informal caregiving: a scoping review across disciplines. Inf Commun Soc 2025 May 22:1-22. [doi: 10.1080/1369118x.2025.2508939]
27. Hall S, Rohatinsky N, Holtslander L, Peacock S. Caregivers to older adults require support: a scoping review of their priorities. Health Soc Care Community 2022 Nov;30(6):e3789-e3809. [doi: 10.1111/hsc.14071] [Medline: 36251430]
28. Derlega VJ, Metts S, Petronio S, Margulis ST. Self-Disclosure. Newbury Park, CA: SAGE Publications; 1993.

XSL•FO
RenderX

29. Xu Y, Burleson BR. Effects of sex, culture, and support type on perceptions of spousal social support: an assessment of the "support gap" hypothesis in early marriage. Human Comm Res 2001 Oct;27(4):535-566. [doi: 10.1111/j.1468-2958.2001.tb00792.x]
30. Zou W, Tang L, Zhou M, Zhang X. Self-disclosure and received social support among women experiencing infertility on Reddit: a natural language processing approach. Comput Hum Behav 2024 May;154:108159. [doi: 10.1016/j.chb.2024.108159]
31. Cutrona CE, Russell DW. Type of social support and specific stress: toward a theory of optimal matching. In: Sarason BR, Sarason IG, Pierce GR, editors. Social Support: An Interactional View. Hoboken, NJ: John Wiley & Sons; 1990.
32. Cutrona CE. Stress and social support—in search of optimal matching. J Soc Clin Psychol 1990 Mar;9(1):3-14. [doi: 10.1521/jscp.1990.9.1.3]
33. Yip JW. Evaluating the communication of online social support: a mixed-methods analysis of structure and content. Health Commun 2020 Sep 03;35(10):1210-1218. [doi: 10.1080/10410236.2019.1623643] [Medline: 31154856]
34. Older persons. UN Refugee Agency. URL: https://emergency.unhcr.org/protection/persons-risk/older-persons [accessed 2025-07-20]
35. Seiter CR, Brophy NS. Social support and aggressive communication on social network sites during the COVID-19 pandemic. Health Commun 2022 Sep 18;37(10):1295-1304. [doi: 10.1080/10410236.2021.1886399] [Medline: 33601993]
36. Whaley B. Research Methods in Health Communication: Principles and Application. Milton Park, UK: Taylor & Francis; 2014.
37. Cutrona CE, Suhr JA. Controllability of stressful events and satisfaction with spouse support behaviors. Commun Res 1992 Apr 01;19(2):154-174. [doi: 10.1177/009365092019002002]
38. Wang YC, Kraut RE, Levine JM. Eliciting and receiving online support: using computer-aided content analysis to examine the dynamics of online social support. J Med Internet Res 2015 Apr 20;17(4):e99 [FREE Full text] [doi: 10.2196/jmir.3558] [Medline: 25896033]
39. Liu Y, Zhu Y, Xia Y. Support-seeking strategies and social support provided in Chinese online health communities related to COVID-19. Front Psychol 2021 Nov 17;12:783135 [FREE Full text] [doi: 10.3389/fpsyg.2021.783135] [Medline: 34867695]
40. Li S, Coduto KD, Song C. Comments vs. one-click reactions: seeking and perceiving social support on social network sites. J Broadcast Electron Media 2020 Dec 09;64(5):777-793. [doi: 10.1080/08838151.2020.1848181]
41. Hailu GN, Abdelkader M, Meles HA, Teklu T. Understanding the support needs and challenges faced by family caregivers in the care of their older adults at home. A qualitative study. Clin Interv Aging 2024 Mar;19:481-490. [doi: 10.2147/CIA.s451833]
42. Huo M, Kim K. Lasting impact of relationships on caregiving difficulties, burden, and rewards. J Soc Pers Relatsh 2022 Aug 07;40(1):288-310. [doi: 10.1177/02654075221118095]
43. Otto AK, Ketcher D, Heyman RE, Vadaparampil ST, Ellington L, Reblin M. Communication between advanced cancer patients and their family caregivers: relationship with caregiver burden and preparedness for caregiving. Health Commun 2021 May 08;36(6):714-721. [doi: 10.1080/10410236.2020.1712039] [Medline: 31910681]
44. Verbakel E, Metzelthin SF, Kempen GI. Caregiving to older adults: determinants of informal caregivers' subjective well-being and formal and informal support as alleviating conditions. J Gerontol B Psychol Sci Soc Sci 2018 Aug 14;73(6):1099-1111. [doi: 10.1093/geronb/gbw047] [Medline: 27130169]
45. Converso D, Sottimano I, Viotti S, Guidetti G. I'll be a caregiver-employee: aging of the workforce and family-to-work conflicts. Front Psychol 2020 Feb 21;11:246 [FREE Full text] [doi: 10.3389/fpsyg.2020.00246] [Medline: 32153465]
46. Kayaalp A, Page KJ, Rospenda KM. Caregiver burden, work-family conflict, family-work conflict, and mental health of caregivers: a mediational longitudinal study. Work Stress 2021 Oct 14;35(3):217-240 [FREE Full text] [doi: 10.1080/02678373.2020.1832609] [Medline: 34483432]
47. Giebel C, Cannon J, Hanna K, Butchard S, Eley R, Gaughan A, et al. Impact of COVID-19 related social support service closures on people with dementia and unpaid carers: a qualitative study. Aging Ment Health 2021 Jul 21;25(7):1281-1288. [doi: 10.1080/13607863.2020.1822292] [Medline: 32954794]
48. Bainbridge D, Krueger P, Lohfeld L, Brazil K. Stress processes in caring for an end-of-life family member: application of a theoretical model. Aging Ment Health 2009 Jul;13(4):537-545. [doi: 10.1080/13607860802607322] [Medline: 19629778]
49. Zhang R. The stress-buffering effect of self-disclosure on Facebook: an examination of stressful life events, social support, and mental health among college students. Comput Hum Behav 2017 Oct;75:527-537. [doi: 10.1016/j.chb.2017.05.043]
50. Lei X, Wu H, Deng Z, Ye Q. Self-disclosure, social support and postpartum depressive mood in online social networks: a social penetration theory perspective. Inf Technol People 2022 Mar 22;36(1):433-453. [doi: 10.1108/itp-12-2020-0825]
51. Duarte F. Reddit user age, gender, and demographics (2025). Exploding Topics. 2025 Jun 23. URL: https://explodingtopics.com/blog/reddit-users [accessed 2025-08-14]
52. Shi J, Chen L, Su Y, Chen M. Offspring caregivers of Chinese women with breast cancer: their social support requests and provision on social media. Telemed J E Health 2019 Aug;25(8):748-755. [doi: 10.1089/tmj.2018.0176] [Medline: 30222524]
53. Walther JB, Boyd S. Attraction to computer-mediated social support. In: Lin CA, Atkin D, editors. Communication Technology and Society: Audience Adoption and Use. Cresskill, NJ: Hampton Press; 2002:153-188.

XSL•FO
RenderX

54. Skaff MM, Pearlin LI. Caregiving: role engulfment and the loss of self. Gerontologist 1992 Oct 01;32(5):656-664. [doi: 10.1093/geront/32.5.656] [Medline: 1427278]
55. Benson JJ, Washington KT, Landon OJ, Chakurian DE, Demiris G, Parker Oliver D. When family life contributes to cancer caregiver burden in palliative care. J Fam Nurs 2023 Aug 16;29(3):275-287. [doi: 10.1177/10748407231167545] [Medline: 37190779]
56. Wittenberg-Lyles E, Washington K, Demiris G, Oliver DP, Shaunfield S. Understanding social support burden among family caregivers. Health Commun 2014;29(9):901-910 [FREE Full text] [doi: 10.1080/10410236.2013.815111] [Medline: 24345081]
57. Morrison S. Don't call me "old": avoiding ageism when writing about aging. National Institute on Aging. 2023 Dec 27. URL: https://www.nia.nih.gov/research/blog/2023/12/dont-call-me-old-avoiding-ageism-when-writing-about-aging [accessed 2025-07-20]

Abbreviations

OMT: optimal matching theory
OR: odds ratio
OSC: online support community
RQ: research question
SPM: stress process model

Edited by G Seçkin; submitted 18.01.25; peer-reviewed by Q Lu, CE Elmore, L Zhang, W Qiaofei; comments to author 22.06.25; revised version received 21.07.25; accepted 04.08.25; published 05.09.25.

Please cite as:
Huang NM, Timothy B, Ho SS
Stressors Disclosed on Reddit by Caregivers of Older Adults and Social Support Received: Content Analysis
JMIR Aging 2025;8:e71452
URL: https://aging.jmir.org/2025/1/e71452
doi:10.2196/71452
PMID:40911350

©Nova Mengxia Huang, Bryan Timothy, Shirley S Ho. Originally published in JMIR Aging (https://aging.jmir.org), 05.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Development and Validation of a Rule-Based Natural Language Processing Algorithm to Identify Falls in Inpatient Records of Older Adults: Retrospective Analysis

Xing Xing Qian[1], PhD; Pui Hing Chau[1], PhD; Daniel Y T Fong[1], PhD; Mandy Ho[1], PhD; Jean Woo[2], MD

[1]School of Nursing, Li Ka Shing Faculty of Medicine, The University of Hong Kong, 5/F, Academic Building, 3 Sassoon Road, Pokfulam, Hong Kong, China (Hong Kong)

[2]Department of Medicine and Therapeutics, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

Corresponding Author:
Pui Hing Chau, PhD
School of Nursing, Li Ka Shing Faculty of Medicine, The University of Hong Kong, 5/F, Academic Building, 3 Sassoon Road, Pokfulam, Hong Kong, China (Hong Kong)

Abstract

Background: In order to address fall underestimation by the *International Classification of Diseases* (*ICD*) in clinical settings, information from clinical notes could be incorporated via natural language processing (NLP) as a possible solution. However, its application to inpatient notes has not been fully investigated.

Objective: This study aims to develop and validate a rule-based NLP algorithm to identify falls based on inpatient admission notes from older patients.

Methods: This retrospective study used 12-year electronic inpatient records of patients aged ≥65 years from public hospitals in Hong Kong. A random sample of 1000 patients was drawn to develop the NLP algorithm. Manual review was the gold standard for assessing the algorithm's performance, with sensitivity, specificity, precision, and F_1-score calculated at the record, episode, and patient levels. In addition, the study compared the number of falls identified by *ICD* codes and clinical notes independently and combined.

Results: Our rule-based NLP algorithm showed excellent performance, with a sensitivity, specificity, precision, and F_1-score of 93.3%, 99.0%, 87.5%, and 0.903 at the record and episode levels, and 92.9%, 98.3%, 89.7%, and 0.912 at the patient level. The combined identification strategy using *ICD* codes and the NLP method provided the most comprehensive capture of fall-related episodes and fallers.

Conclusions: The NLP method proved efficient and accurate in detecting falls from clinical notes in inpatient episodes. For comprehensive capture of fall episodes and fallers, we recommend the combined use of the NLP algorithm and *ICD* codes, which should be applied in future fall epidemiology studies and clinical practice for identifying high-risk groups of fall interventions.

(JMIR Aging 2025;8:e65195) doi:10.2196/65195

KEYWORDS

fall-related admissions; electronic medical records; text mining; case detection; natural language processing

Introduction

Falls are a global public concern and impose considerable health care and economic burdens on the older population [1]. The Decade of Healthy Aging Report from World Health Organization (WHO) highlights fall prevention as an urgent health priority [2]. Fall risk screening is inevitable when identifying vulnerable older individuals for assessment and preventive intervention [3,4]. The history of falls has been widely recognized as a vital component in risk screening [3], given that it doubles the fall risk for older adults [5,6].

In the longitudinal studies on the older population, falls were mainly identified through medical (electronic and nonelectronic) or nonmedical records (eg, retrospective telephone or face-to-face interviews and prospective fall diary or calendar) [6,7]. Compared to nonmedical methods, which are subject to withdrawal or loss of follow-up from participants, medical records are more efficient and cost-saving. Hence, hospitals that provide care for fall-related injuries become the ideal sites to identify and track fallers for interventions. Most studies conducted in clinical settings used the *International Classification of Diseases* (*ICD*) codes to detect falls in the electronic medical records (EMRs) of older adults [6]. Nevertheless, the identification based on the *ICD* codes alone may not completely capture the reported falls. Published studies found that the *International Classification of Diseases, Ninth Revision, Clinical Modification* (*ICD-9-CM*) codes missed 8%

XSL•FO
RenderX

to 20% of falls in patients [8,9]. Involving supplemental records could increase accuracy for fall identification, in which the clinical note review captures the greatest number of events [10]. The traditional manual note review that requires health care professionals as the reviewer is labor-intensive and time-consuming due to the large volume of records. While a relatively expedient and cost-effective approach is searching for the keyword "fall" in the clinical notes, it is important to note that a simple keyword search alone is insufficient to accurately distinguish between fall-related notes and those that do not. This is because notes without actual falls may also contain the term "fall", such as "references to high fall risk," "fall precaution," or "falling asleep," which can result in false positives. To enhance accuracy and overcome this limitation, a novel computer-based approach named natural language processing (NLP) has been applied as an effective solution [11].

NLP algorithms are generally categorized into two approaches: rule based and machine learning (ML) based [11]. While both have their merits, rule-based NLP is selected for this study for several reasons. First, rule-based systems rely on explicitly defined linguistic rules, making them highly interpretable and transparent [12]. This allows developers and users to easily understand and modify the rules, ensuring precise performance. Second, rule-based NLP does not require large annotated datasets, unlike ML-based approaches, which demand extensive training data and computational resource [13]. This makes rule-based systems practical for real-time applications or low-resource environments. Third, rule-based systems efficiently handle edge cases and specific linguistic nuances by directly tailoring rules to unique exceptions—such as distinguishing between actual falls and terms like "falling asleep" in clinical notes—without extensive data or training [12]. While ML-based NLP excels with large-scale datasets and broad contexts [13], it often requires more effort to match the precision and domain-specific adaptability of rule-based systems. In recent years, transformer-based pretrained models have emerged as a powerful ML approach, advancing NLP by capturing long-range dependencies and contextual information in text, even with limited annotated data [14]. Due to their ability to model complex linguistic patterns and contextual relationships, transformer-based models may outperform rule-based systems. Nevertheless, these models still require substantial computational resources, which can be mitigated through cloud-based application programming interfaces (APIs) [15]. In health care systems where patient data confidentiality is strictly restricted, the reliance on internet connectivity to use such models poses a significant challenge. Therefore, for local applications, rule-based approaches remain a pragmatic choice.

However, research on rule-based NLP algorithms for fall identification in clinical notes remains limited. Previous studies have developed rule-based NLP algorithms for inpatient, emergency department (ED), and homecare visit notes [16-18], with the inpatient study specifically focusing on Japanese clinical notes [16]. Although the English-language NLP algorithms for ED and homecare visit notes demonstrated satisfactory performance compared to manual review [17,18], these methods may not be directly transferable to inpatient settings. Hence, more studies are warranted to develop specialized rule-based NLP methods for accurate fall identification in inpatient clinical notes. Furthermore, existing studies primarily compared identification strategies based on single record sources of *ICD* codes or clinical notes [8,9]. There is a lack of studies investigating how the combination of these strategies can enhance fall identification. Therefore, to address these gaps, this study aimed to (1) develop and validate a rule-based NLP algorithm to identify falls based on inpatient admission notes and (2) compare the number of falls identified by independent or combined strategies of *ICD* codes and the NLP method.

Methods

Study Design and Data Source

We performed a retrospective, territory-wide, cohort study using the 12-year (2007-2018) electronic inpatient records of older adults. The dataset used in this study encompasses admissions from all 43 public hospitals of the Hospital Authority (HA), which manage approximately 80% of hospital admissions in Hong Kong and account for nearly 90% of total bed-days [19]. The data were extracted from the Hospital Authority Data Collaboration Lab (HADCL), and the details were described elsewhere [20]. The clinical records were written in English.

Sampling

The sampling frame was a cohort of older adults who (1) resided in Hong Kong, (2) were discharged as an inpatient from public hospitals between January 2007 and December 2017, (3) were aged ≥65 years at admission, (4) had a hospital stay of at least one day for their first inpatient episode, (5) were discharged alive, (6) did not have a hospital discharge from public hospitals 365 days before the first inpatient episode, and (7) were not admitted from the nursing home at their first inpatient episode. More details of the criteria are described elsewhere [20].

To ensure adequate statistical power for developing an NLP algorithm, a sample size of at least 500 is recommended [21]. In this study, a sample of 1000 patients was drawn by random sampling. The electronic clinical notes and diagnosis codes of all inpatient admission episodes from selected patients were extracted for analysis. The dataset contains two types of notes: clinical notes and discharge notes. Only the clinical notes, comprising all textual records associated with linked episodes, were used in this study.

Manual Review

The manual review was considered a gold standard, as consistent with previous studies [16-18]. A total of 2 reviewers independently reviewed each extracted clinical note and determined the corresponding record as being fall related or not, using binary classification. According to the WHO, a fall is defined as an event in which a person unintentionally comes to rest on the ground, floor, or another lower level [1]. An individual who has experienced at least one fall episode is classified as a faller. Reviewers were trained to annotate the clinical notes as fall related if the scenario presented in the notes meet the WHO definition of falls or have specific indicators (eg, "admitted for fall," "presented with fall," "fell onto ground," "slipped and fell," or "fell when..."). Reviewers relied on these

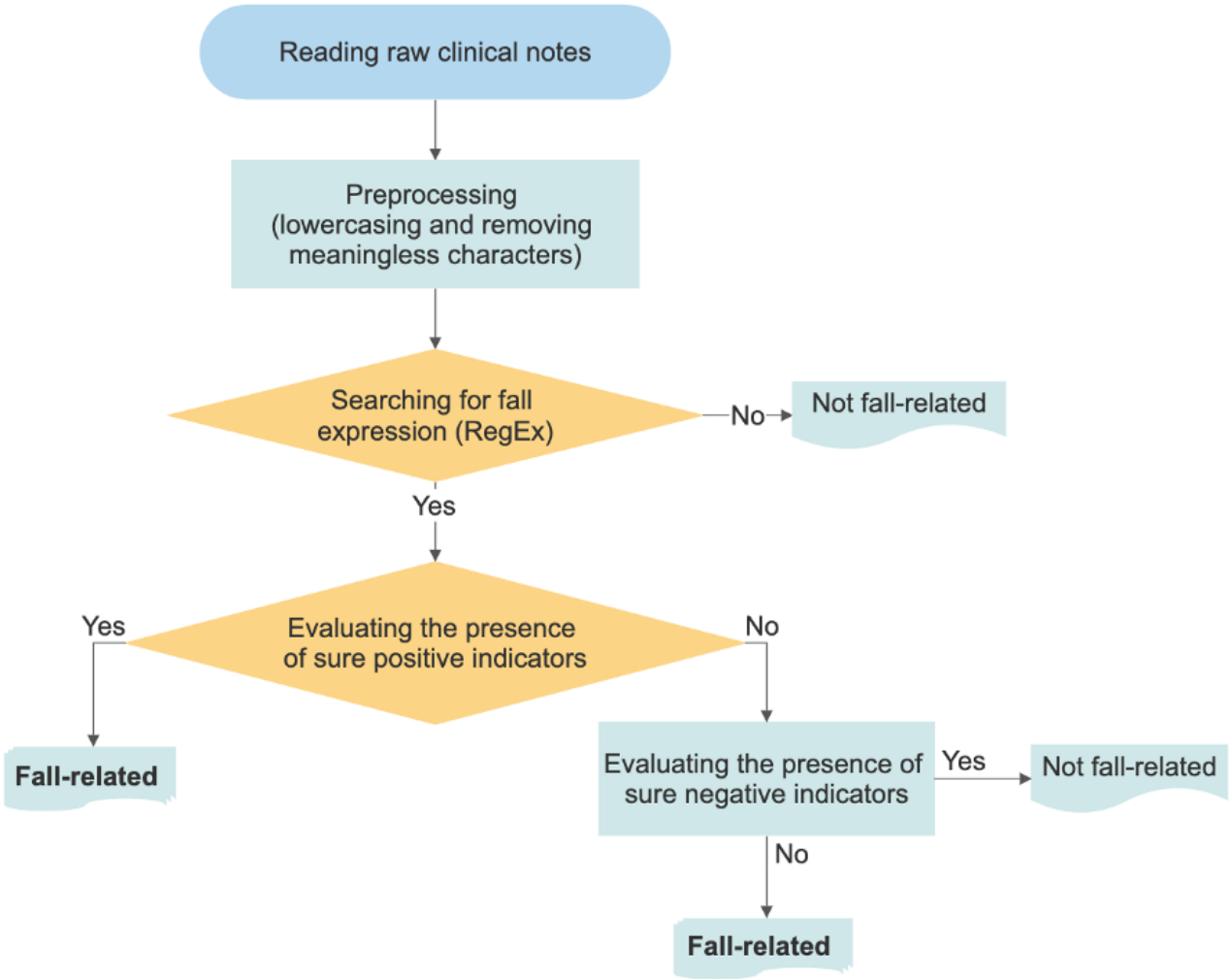
XSL•FO
RenderX

predefined indicators to ensure consistency in their assessments. Any disagreement was resolved by a joint evaluation involving the 2 reviewers and a senior investigator. As each patient could have multiple inpatient admission episodes during the study period and each episode could include multiple records, fall identification was performed at levels of record, episode, and patient. Patients were classified as fallers and episodes were designated as fall related if any associated episodes or records contained falls.

Development and Validation of the Rule-Based NLP Algorithm

Patients were randomly split into the training set for NLP algorithm development and the testing set for validation at a ratio of 80:20 [22]. The development process in the training set included 5 steps (Figure 1). First, clinical notes were preprocessed by lowercasing characters and removing meaningless "Newline Characters" (indicates starting from the next line). Second, the fall expression within the clinical notes was searched by the Regular Expression (RegEx; Section S1 in Multimedia Appendix 1). After filtering out clinical notes with fall expression, the third step evaluated the presence of sure positive indicators (Section S1 in Multimedia Appendix 2) that a fall occurred within the filtered notes. Similarly, the fourth step was evaluating the presence of sure negative indicators (Section S1 in Multimedia Appendix 2) that a fall had not occurred. The sure positive and sure negative indicators were iteratively refined through performance evaluation and rule adjustment. At each cycle, common areas of false positives and negatives were addressed to extend the criteria until the value of sensitivity, specificity, and precision exceeded 90%, and the value of F_1-score exceeded 0.9 at one of the levels for record, episode, or patient [17]. Finally, as fall expression can appear more than once within the same clinical note, the sure positive and negative indicators could coexist in one note. Therefore, clinical notes that (1) had fall expression and (2) had sure positive indicator or did not have sure negative indicator were marked as fall related. Subsequently, the rule-based NLP algorithm was applied to identify falls in clinical notes from the testing set for validation.

Figure 1. Process of the rule-based natural language processing algorithm development. RegEx: Regular Expression.



Comparison of Different Identification Strategies

The number of falls identified by the following methods were compared: (1) manual review of clinical notes, (2) NLP algorithm for clinical notes, (3) *ICD* codes, (4) manual review and *ICD* codes, and (5) NLP algorithm and *ICD* codes. Regarding the strategy of *ICD* codes, a fall was identified by the *ICD-9-CM* E880-E888 for accidental falls or 800 - 909.2,

XSL•FO
RenderX

909.4, 909.9, 910 - 924, 950 - 957, 959 for injury (excluding E000, E800-807, E810-838, E840-858, E860-879, E890-999 for specific external causes of injury) [23,24]. Since not all episodes contain both *ICD* codes and clinical notes—some may lack notes while others may lack *ICD* codes, comprehensive fall identification should consider both data sources. Therefore, the combined strategy, incorporating manual notes review and *ICD* codes, was considered the identification standard to compare the remaining strategies. For combined strategies (4) and (5), the record was identified as fall related when either strategy captured a fall. The comparison was restricted to these 1000 patients, as one of the involved strategies was manual review, which could be time-consuming and labor-intensive if extended to the whole dataset.

Statistical Analysis

The mean and SD values were used to describe the length of clinical notes (number of characters). The falls identified by the developed NLP algorithm were compared with the manual review by confusion matrix to generate standard performance measures of sensitivity, specificity, precision, and F_1-score at the levels of record, episode, and patient in the training and testing sets [25]. Sensitivity refers to the proportion of correctly labeled falls by the NLP algorithm to all falls identified by manual review. Specificity computes the proportion of correctly labeled nonfalls by the NLP algorithm to all nonfalls identified by manual review. Precision reflects the proportion of the correctly labeled falls by the NLP algorithm to all labeled falls by NLP algorithm. F_1-score is the harmonic mean of the precision and sensitivity, with the best value being 1 and the worst being 0. An adequate performance is considered when the value exceeds 90% for sensitivity, specificity, and precision and 0.9 for F_1-score [17]. All data analyses were performed using Python 3.7.6 (CreateSpace).

Ethical Considerations

This study was approved by the Institutional Review Board of the University of Hong Kong and the Hong Kong Hospital Authority West Cluster (approval UW 23 - 046). Informed consent was waived by the Institutional Review Board due to the retrospective, observational nature of the study and the exclusive use of deidentified participant data. Participants were not compensated for their involvement in this study.

Results

In total, 2153 clinical notes (records) of 2095 inpatient admission episodes from 1000 sampled patients were included in the analysis. The mean length of notes was 1072 characters, with an SD of 916. To develop the NLP algorithm, 800 patients (1720 records and 1672 episodes) were randomly assigned to the training set and the remaining 200 patients (433 records and 423 episodes) were assigned to the testing set.

Regarding the training set for development, the manual review identified 140 fall-related records, 135 fall-related episodes, and 127 fallers. In comparison, the NLP algorithm identified 145 fall-related records, 140 fall-related episodes, and 124 fallers. In the testing set for validation, the manual review identified 30 fall-related records, 30 fall-related episodes, and 28 fallers, while the NLP algorithm identified 32 fall-related records, 32 fall-related episodes, and 29 fallers. As shown in Table 1, falls identified by the NLP algorithm were largely concordant with those identified by manual review at record, episode, and patient levels. Performance metrics of the NLP algorithm were satisfactory for both training and testing sets at all 3 levels (Table 2). In the testing set, the NLP algorithm achieved a sensitivity of 93.3%, a specificity of 99.0%, a precision of 87.5%, and an F_1-score of 0.903 at both the record and episode levels. These measures at patient level were 92.9% for sensitivity, 98.3% for specificity, 89.7% for precision, and 0.912 for F_1-score.

Table 3 presents the comparison of different fall identification strategies, with regards to the number of falls identified at the record, episode, and patient levels. Overall, as compared to identification standard of manual notes review combined with *ICD* codes, the NLP algorithm combined with *ICD* codes identified the highest number of falls across all 3 levels (99.4%-104.3%), followed by the NLP algorithm alone (91.1%-93.5%). Specifically, compared to the identification standard, the NLP algorithm combined with *ICD* codes slightly overestimated the numbers of fall-related records and episodes (~4%) yet underestimated the number of fallers (~1%). In contrast, *ICD* codes alone identified the least number of falls (58.7%-60.7%), as 44 inpatient records did not have *ICD-9-CM* codes.

XSL•FO
RenderX

Table . Comparison of fall identification by manual review and the rule-based NLP[a] algorithm.

NLP algorithm and level	Manual review			
	Training set		Testing set	
	Fall, n	Without fall, n	Fall, n	Without fall, n
Record				
Fall	129	16	28	4
Without fall	11	1564	2	399
Episode				
Fall	124	16	28	4
Without fall	11	1521	2	389
Patient				
Fall	116	8	26	3
Without fall	11	665	2	169

[a]NLP: natural language processing.

Table . Performance metrics of the rule-based natural language processing algorithm.

Metric	Training set			Testing set		
	Record level	Episode level	Patient level	Record level	Episode level	Patient level
Sensitivity	92.1%	91.9%	91.3%	93.3%	93.3%	92.9%
Specificity	99%	99%	98.8%	99%	99%	98.3%
Precision	89%	88.6%	93.5%	87.5%	87.5%	89.7%
F_1-score	0.905	0.902	0.924	0.903	0.903	0.912

Table . Number of falls identified by different strategies.

Strategy	Number of falls identified (% of identification standard)		
	Record level	Episode level	Patient level
Manual review and *ICD*[a] codes[b]	193 (100)	184 (100)	168 (100)
NLP[c] algorithm and *ICD* codes	201 (104.1)	192 (104.3)	167 (99.4)
Manual review of clinical notes	170 (88.1)	165 (89.7)	155 (92.3)
NLP algorithm for clinical notes	177 (91.7)	172 (93.5)	153 (91.1)
ICD codes[d]	114 (59.1)	108 (58.7)	102 (60.7)

[a]*ICD*: *International Classification of Diseases*.

[b]Identification standard for comparison.

[c]NLP: natural language processing.

[d]In the whole set, 44 records were without *ICD* codes.

Discussion

Principal Findings

This study developed a rule-based NLP algorithm to identify falls in clinical notes of inpatient admission episodes based on a random, territory-wide sample of older patients, which achieved excellent performance when compared with manual notes review in the testing set. We also compared different fall identification strategies, including the use of *ICD* codes or clinical notes alone or in combination. Our findings provide a feasible and efficient solution for fall screening and shed light on future research and practice related to this field in the clinical setting.

Using NLP algorithms to identify falls can significantly reduce the time and human resources required for the manual notes review. Previous studies have developed NLP methods for fall identification in clinical notes of inpatient wards [16,26,27], and the performance of NLP methods varied. Toyabe [16] derived a rule-based NLP algorithm to detect falls in Japanese progress notes and achieved a sensitivity of 100.0% and a specificity of 98.4%, but with a low precision (6.2%) and an F_1-score of 0.118 at the record level. Shiner et al [26] developed

XSL•FO
RenderX

an ML-based NLP method to capture falls in English progress notes, achieving a relatively low sensitivity of 44.1%, a specificity of 97.0%, a precision of 67.2%, and an F_1-score of 0.533 at the record level. Nakatani et al [27] also developed an ML-based NLP algorithm to differentiate fallers and nonfallers in Japanese nursing notes, with a moderate sensitivity of 76.9% and a specificity of 78.5% at the patient level. In this study's testing set, the sensitivity, specificity, precision, and F_1-score of the rule-based NLP algorithm were 93.3%, 99.0%, 87.5%, and 0.903 at the record level and 92.9%, 98.3%, 89.7%, and 0.912 at the patient level, respectively. The values of these measures were generally higher than previous studies, indicating the promising performance of our rule-based NLP algorithm. Notably, our NLP method was rule based and could be easily applied by researchers or health care professionals with limited programming knowledge to facilitate fall identification in inpatient settings. The rule-based NLP algorithm is also flexible for adjustment based on the different patterns of clinical notes and could be generalized to different systems, such as the ED clinical notes. In addition, future studies may explore the extended application of the NLP method in extracting other fall-related information, such as the circumstance of falls, the mechanism of falls, and resulting injuries, which are important in informing the prognosis and required interventions for patients who have experienced falls [3].

In the epidemiological studies for falls in older adults, the number of fall episodes and fallers are typically the outcomes of interest in estimating the risk (eg, incidence and proportions) and associations. Our findings suggest that the fall identification strategy involving *ICD* codes and clinical notes provides the most comprehensive capture of fall-related episodes and fallers. Identification of falls based on *ICD* codes alone is known to underestimate the number of falls [8,9,28]. Incorporating information from clinical notes is especially beneficial in addressing this issue. Given the limitations of manual review, the NLP algorithm offers an efficient way for applying the combined strategies of clinical notes and *ICD* codes in fall identification on a large scale. Our study found that compared with the combination of manual review and *ICD* codes, the combination of NLP algorithm and *ICD* codes slightly overestimated the number of fall-related episodes (~4%) but not the number of fallers. Therefore, the fall identification strategy combining NLP algorithm and *ICD* codes with feasibility and accuracy has promising application value in risk screening and patient tracking.

The health care workforce shortage is a global problem. Identification of efficient methods to enhance the quality of patient care and safety is a priority agenda in health care field. In terms of implications to practice, there are posthospital health care services or fall prevention programs designed to improve functions and reduce falls for patients recently discharged from hospitals. Fall history is a dominant risk factor for posthospital falls in older patients [6]. Since hospitals are ideal sites for identifying and tracking patients with a history of falls, this combined strategy of NLP method and *ICD* codes is expected to benefit posthospital fall interventions both at baseline and during follow-up. The combined strategy allows for standardized and consistent screening across an extensive EMR dataset, and it can be integrated into existing EMR systems to conduct real-time screening and identification of fall events within minutes. For instance, upon discharge, health care professionals can use it to screen older patients with a history of falls to be recruited in posthospital fall prevention programs. After discharge, patients' fall-related readmission can be tracked and identified timely by this strategy in the EMR system. To establish a more comprehensive surveillance procedure, supplementing nonmedical methods to collect fall incidents that do not present hospitals can be considered. Meanwhile, posthospital health care services, such as the Nurse and Allied Health Clinic (NAHC) for Falls, are designed specifically to prevent falls for patients. Especially for a resource-limited setting, such services can better target patients with a fall history.

The strengths of this study include using 12-year medical records from a territory-wide older population across multiple hospitals, which enhances the reliability and generalizability of the developed NLP algorithm. In addition, we compared the performance of independent and combined fall identification strategies based on *ICD* codes and clinical notes to determine the optimal method in practice. However, several limitations should be acknowledged. First, direct comparison of our NLP algorithm with existing rule-based and ML-based NLP algorithms, including transformer-based pretrained models, was not conducted in this study. Future research should compare the performance of these approaches on the same set of inpatient notes and evaluate their strengths and limitations. Second, a comprehensive analysis of false positives and negatives pattern could not be presented due to data sharing restrictions. Third, due to data availability, this study was limited by the inability to incorporate clinical notes from patients admitted to the ED to derive and compare NLP algorithms for different health care utilizations. Future studies are warranted to compare NLP methods generated from different settings and evaluate their contribution to fall identification. In addition, since the data in this study were drawn from 12-year medical records, it is uncertain whether the writing style of doctors or nurses changed over time. The lack of interrater reliability testing represents another limitation of this study. While the analysis of performance metrics for *ICD* codes would provide a more comprehensive comparison, this analysis was not conducted in this study and remains an important direction for future research. Finally, this study used a cohort developed for another study, which excluded older patients who had hospitalizations within the past year or were admitted from nursing homes, leading to a relatively healthier sample with lower fall risk. Investigations on the performance of our rule-based NLP algorithm in a more general older population are required.

Conclusions

To conclude, this study has demonstrated that the rule-based NLP algorithm could efficiently and accurately detect falls in clinical notes from inpatient admission episodes with excellent performance. The combined identification strategy that incorporates the NLP algorithm and *ICD* codes could be applied to comprehensively capture fall episodes and fallers in future research of fall epidemiology, as well as in clinical practice to facilitate the identification of high-risk groups for fall interventions.

XSL•FO
RenderX

Acknowledgments

The authors would like to thank the Hong Kong Hospital Authority and its Information Technology and Health Informatics Department for providing the data and supporting this study.

Data Availability

The datasets generated during and analyzed during this study are not publicly available as its owned by the Hong Kong Hospital Authority.

Authors' Contributions

XXQ, PHC, DYTF, MH, and JW contributed to the study concept and design; XXQ and PHC contributed to the acquisition, analysis, and interpretation of data; XXQ contributed to the drafting of the manuscript; XXQ, PHC, DYTF, MH, and JW contributed to critical revision of the manuscript for important intellectual content.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Regular expressions used to identify fall events.
[DOCX File, 34 KB - aging_v8i1e65195_app1.docx]

Multimedia Appendix 2
Criteria for sure positive and sure negative fall indicators.
[DOCX File, 37 KB - aging_v8i1e65195_app2.docx]

References

1. Falls. World Health Organization. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/falls [accessed 2024-04-19]
2. Decade of healthy ageing: baseline report. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/9789240017900 [accessed 2024-04-19]
3. Montero-Odasso M, van der Velde N, Martin FC, et al. World guidelines for falls prevention and management for older adults: a global initiative. Age Ageing 2022 Sep 2;51(9):afac205. [doi: 10.1093/ageing/afac205] [Medline: 36178003]
4. Cameron ID, Dyer SM, Panagoda CE, et al. Interventions for preventing falls in older people in care facilities and hospitals. Cochrane Database Syst Rev 2018 Sep 7;9(9):CD005465. [doi: 10.1002/14651858.CD005465.pub4] [Medline: 30191554]
5. Deandrea S, Lucenteforte E, Bravi F, Foschi R, La Vecchia C, Negri E. Risk factors for falls in community-dwelling older people: a systematic review and meta-analysis. Epidemiology 2010 Sep;21(5):658-668. [doi: 10.1097/EDE.0b013e3181e89905] [Medline: 20585256]
6. Qian XX, Chen Z, Fong DYT, Ho M, Chau PH. Post-hospital falls incidence and risk factors among older adults: a systematic review and meta-analysis. Age Ageing 2022 Jan 6;51(1):afab209. [doi: 10.1093/ageing/afab209]
7. Schwenk M, Lauenroth A, Stock C, et al. Definitions and methods of measuring and reporting on injurious falls in randomised controlled fall prevention trials: a systematic review. BMC Med Res Methodol 2012 Apr 17;12:50. [doi: 10.1186/1471-2288-12-50] [Medline: 22510239]
8. Patterson BW, Smith MA, Repplinger MD, et al. Using chief complaint in addition to diagnosis codes to identify falls in the emergency department. J Am Geriatr Soc 2017 Sep;65(9):E135-E140. [doi: 10.1111/jgs.14982] [Medline: 28636072]
9. Patterson BW, Jacobsohn GC, Maru AP, et al. RESEARCHComparing strategies for identifying falls in older adult emergency department visits using EHR data. J Am Geriatr Soc 2020 Dec;68(12):2965-2967. [doi: 10.1111/jgs.16831] [Medline: 32951200]
10. Olsen S, Neale G, Schwab K, et al. Hospital staff should use more than one method to detect adverse events and potential adverse events: incident reporting, pharmacist surveillance and local real-time record review may all have a place. Qual Saf Health Care 2007 Feb;16(1):40-44. [doi: 10.1136/qshc.2005.017616] [Medline: 17301203]
11. Trinh VQN, Zhang S, Kovoor J, et al. The use of natural language processing in detecting and predicting falls within the healthcare setting: a systematic review. Int J Qual Health Care 2023 Oct 17;35(4):mzad077. [doi: 10.1093/intqhc/mzad077] [Medline: 37758209]
12. Liu H, Cocea M. Fuzzy rule based systems for interpretable sentiment analysis. Presented at: 2017 Ninth International Conference on Advanced Computational Intelligence (ICACI); Feb 4-6, 2017; Doha, Qatar p. 129-136. [doi: 10.1109/ICACI.2017.7974497]

XSL•FO
RenderX

13. Spasic I, Nenadic G. Clinical text data in machine learning: systematic review. JMIR Med Inform 2020 Mar 31;8(3):e17984. [doi: 10.2196/17984] [Medline: 32229465]
14. Rahali A, Akhloufi MA. End-to-end transformer-based models in textual-based NLP. AI 2023;4(1):54-110. [doi: 10.3390/ai4010004]
15. Yenduri G, Ramalingam M, Selvi GC, et al. GPT (Generative Pre-Trained Transformer)— a comprehensive review on enabling technologies, potential applications, emerging challenges, and future directions. IEEE Access 2024;12:54608-54649. [doi: 10.1109/ACCESS.2024.3389497]
16. Toyabe SI. Detecting inpatient falls by using natural language processing of electronic medical records. BMC Health Serv Res 2012 Dec 5;12:448. [doi: 10.1186/1472-6963-12-448] [Medline: 23217016]
17. Patterson BW, Jacobsohn GC, Shah MN, et al. Development and validation of a pragmatic natural language processing approach to identifying falls in older adults in the emergency department. BMC Med Inform Decis Mak 2019 Jul 22;19(1):138. [doi: 10.1186/s12911-019-0843-7] [Medline: 31331322]
18. Topaz M, Murga L, Gaddis KM, et al. Mining fall-related information in clinical notes: Comparison of rule-based and novel word embedding-based machine learning approaches. J Biomed Inform 2019 Feb;90:103103. [doi: 10.1016/j.jbi.2019.103103] [Medline: 30639392]
19. Report of the strategic review on healthcare manpower planning and professional development. Health Bureau of the Government of the Hong Kong Special Administrative Region. 2020. URL: https://www.healthbureau.gov.hk/en/press_and_publications/otherinfo/180500_sr/srreport.html [accessed 2024-04-19]
20. Qian XX, Chau PH, Fong DYT, Ho M, Woo J. Post-hospital falls among the older population: the temporal pattern in risk and healthcare burden. J Am Med Dir Assoc 2023 Oct;24(10):1478-1483. [doi: 10.1016/j.jamda.2023.07.014] [Medline: 37591487]
21. Juckett D. A method for determining the number of documents needed for a gold standard corpus. J Biomed Inform 2012 Jun;45(3):460-470. [doi: 10.1016/j.jbi.2011.12.010] [Medline: 22245601]
22. Gholamy A, Kreinovich V, Kosheleva O. Why 70/30 or 80/20 relation between training and testing sets: a pedagogical explanation. Int J Intell Technol Appl Stat 2018;11(2):105-111. [doi: 10.6148/IJITAS.201806_11(2).0003]
23. Lam C, Kang JH, Lin HY, Huang HC, Wu CC, Chen PL. First fall-related injuries requiring hospitalization increase the risk of recurrent injurious falls: a nationwide cohort study in Taiwan. PLoS One 2016;11(2):e0149887. [doi: 10.1371/journal.pone.0149887] [Medline: 26900917]
24. Owens PL, Russo A, Spector W, Mutter R. Emergency department visits for injurious falls among the elderly, 2006. Agency for Healthcare Research and Quality. 2009. URL: https://hcup-us.ahrq.gov/reports/statbriefs/sb80.pdf [accessed 2025-06-06]
25. Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks. Inf Process Manag 2009 Jul;45(4):427-437. [doi: 10.1016/j.ipm.2009.03.002]
26. Shiner B, Neily J, Mills PD, Watts BV. Identification of inpatient falls using automated review of text-based medical records. J Patient Saf 2020 Sep;16(3):e174-e178. [doi: 10.1097/PTS.0000000000000275] [Medline: 27331601]
27. Nakatani H, Nakao M, Uchiyama H, Toyoshiba H, Ochiai C. Predicting inpatient falls using natural language processing of nursing records obtained from Japanese electronic medical records: case-control study. JMIR Med Inform 2020 Apr 22;8(4):e16970. [doi: 10.2196/16970] [Medline: 32319959]
28. Waters TM, Chandler AM, Mion LC, et al. Use of International Classification of Diseases, Ninth Revision, Clinical Modification, codes to identify inpatient fall-related injuries. J Am Geriatr Soc 2013 Dec;61(12):2186-2191. [doi: 10.1111/jgs.12539] [Medline: 24329820]

Abbreviations

API: application programming interface
ED: emergency department
EMR: electronic medical record
HA: Hospital Authority
HADCL: Hospital Authority Data Collaboration Lab
***ICD*:** *International Classification of Diseases*
***ICD-9-CM*:** *International Classification of Diseases, Ninth Revision, Clinical Modification*
ML: machine learning
NAHC: Nurse and Allied Health Clinic
NLP: natural language processing
RegEx: Regular Expression
WHO: World Health Organization

XSL•FO
RenderX

Edited by M O'Connell; submitted 07.08.24; peer-reviewed by A Mitra, H Wang, K Zandbiglari, S Zeng; revised version received 25.03.25; accepted 10.05.25; published 08.07.25.

Please cite as:
Qian XX, Chau PH, Fong DYT, Ho M, Woo J
Development and Validation of a Rule-Based Natural Language Processing Algorithm to Identify Falls in Inpatient Records of Older Adults: Retrospective Analysis
JMIR Aging 2025;8:e65195
URL: https://aging.jmir.org/2025/1/e65195
doi:10.2196/65195

© Xing Xing Qian, Pui Hing Chau, Daniel Y T Fong, Mandy Ho, Jean Woo. Originally published in JMIR Aging (https://aging.jmir.org), 8.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Impact of Vision Impairment on Self-Reported Falls Among Older US Adults: Cross-Sectional and Longitudinal Study

Kasem Seresirikachorn[1,2], MD, PhD; Rachasak Somyanonthanakul[3], PhD; Matthew Johnson[1], BS; Panisa Singhanetr[1,4], MD; Jiraporn Gatedee[5], PhD; David Friedman[1], MD, MPH, PhD; Nazlee Zebardast[1], MSc, MD

[1]Massachusetts Eye and Ear Infirmary, Harvard Medical School, 243 Charles Street, Boston, MA, United States

[2]Department of Ophthalmology, College of Medicine, Rangsit University, Rajavithi Hospital, Bangkok, Thailand

[3]Data Science and Innovation, College of Interdisciplinary Studies, Thammasat University, Pathum Thani, Thailand

[4]Mettapracharak Eye Institute, Mettapracharak (Wat Rai Khing) Hospital, Nakhon Pathom, Thailand

[5]Faculty of Medical Technology, Rangsit University, Pathum Thani, Thailand

Corresponding Author:
Nazlee Zebardast, MSc, MD
Massachusetts Eye and Ear Infirmary, Harvard Medical School, 243 Charles Street, Boston, MA, United States

Abstract

Background: Falls are the leading cause of injury among older adults, with vision impairment recognized as a significant risk factor. However, many existing studies have been limited by small sample sizes, retrospective designs, or insufficient adjustment for confounding factors. To overcome these limitations, we used data from the University of Michigan's Health and Retirement Study (HRS) to analyze the association between self-reported vision and fall risk among older adults in a large, nationally representative sample.

Objective: The objective of this study was to investigate the association between vision impairment and falls and assess whether subjective vision impairment predicts future falls in older adults.

Methods: This cross-sectional and longitudinal analysis used data from the HRS (1996 - 2020) to assess the relationship between self-reported vision, glaucoma history, and falls among US adults aged 65 years and older. HRS uses a biennial, multistage area probability sample survey design, collecting data with community-dwelling individuals followed up every 2 years until death, tracking health, economic, and social outcomes. Multivariate logistic regression was used to analyze associations between self-reported vision and self-reported falls in the past 2 years.

Results: A total of 38,835 respondents contributed 117,834 observations. The weighted proportion of participants reporting falls was 37.9% (95% CI 37.7% - 40.1%). Significant risk factors for falls included overall eyesight impairment (adjusted odds ratio [aOR] 1.36, 95% CI 1.20 - 1.56), distance vision impairment (aOR 1.37, 95% CI 1.32 - 1.42), near vision impairment (aOR 1.33, 95% CI 1.27 - 1.37), and glaucoma (aOR 1.15, 95% CI 1.07 - 1.24). A similar association was observed for serious falls, where overall eyesight impairment (aOR 1.20, 95% CI 1.03 - 1.44), distance vision impairment (aOR 1.14, 95% CI 1.07 - 1.22), near vision impairment (aOR 1.12, 95% CI 1.05 - 1.18), and glaucoma (aOR 1.15, 95% CI 1.05 - 1.26) were significant. In longitudinal analyses, overall vision impairment (aOR 1.23, 95% CI 1.16 - 1.29), distance vision impairment (aOR 1.27, 95% CI 1.20 - 1.38), near vision impairment (aOR 1.23, 95% CI 1.19 - 1.32), and glaucoma (aOR 1.25, 95% CI 1.13 - 1.37) increased the risk of future falls. Reported overall vision was significantly associated with the number of falls in both the same (P<.001) and subsequent (P<.001) survey cycles.

Conclusions: Both distance and near vision impairment, as well as glaucoma, are associated with a higher risk of falls in older adults and present possible areas for intervention and prevention.

(JMIR Aging 2025;8:e68771) doi:10.2196/68771

KEYWORDS

elderly; falls; glaucoma; Health and Retirement Study; vision impairment

Introduction

Falls are the leading cause of injury among adults ages 65 years or older [1]. Nearly 30% of older adults fall each year, and 10% report injury from a fall [2]. In the United States in 2018, there were 36 million reported falls, resulting in 3 million emergency department visits, nearly 1 million hospitalizations, and 32,000 deaths [3]. Approximately 1 in 10 falls in older adults results in serious injury, such as hip fracture or traumatic brain injury, and the annual cost of falls in the United States is more than US $50 billion [4-6]. Globally, falls are the 13th leading cause of death, and the prevalence of falls is increasing as the

XSL•FO
RenderX

population gets older [7]. Fortunately, multifactorial intervention strategies such as medical staff and patient education, providing supplies for hip protection, and sensory optimization by providing hearing aids have all been able to successfully reduce the risk of falling [5,7,8]. Home modifications to improve safety, such as changing floor coverings, removing loose carpets, fitting handrails, maintaining steps and ramps, and improving lighting, can reduce falling hazards in homes.

Falls are significant, life-altering health events with modifiable risk factors such as impaired balance, polypharmacy, and poor vision [9-11]. Impaired vision has been shown to be a risk factor for falls in older adults. In 1998, the Beaver Dam Eye Study showed that impaired visual acuity was associated with an increased risk of falling, and older adults in the worst visual acuity group were significantly more likely to have fallen 2 or more times in the past year [12]. The Blue Mountains Eye Study similarly found a significant relationship between low visual acuity and falls. In addition, they reported that decreased contrast sensitivity and reduced visual field (VF) increased the risk of falls and proposed that the presence of a cataract may explain this association [13]. In the Salisbury Eye Evaluation study, only reduced peripheral VF was associated with the risk of falling, while visual acuity, contrast sensitivity, and stereo acuity were not after adjusting for potential confounders [14]. Other studies support a relationship between impaired vision and falls but differ as to which components of visual function increase fall risk [5,9,15-17].

These studies provided valuable information that broadly supports impaired vision as a modifiable risk factor for falls; however, many of these studies had limitations due to small sample sizes, retrospective analyses, or the inability to adjust for confounders. Even the most extensive studies to date have data spanning only 1 - 2 years and are generally limited to local communities or geographic regions [12-14].

To address these limitations, we used the University of Michigan Health and Retirement Study (HRS) data. The HRS is a diverse, nationally representative longitudinal panel study of US adults ages 50 years and older, followed every 2 years for over 20 years [18]. Data on fall incidence and self-reported vision impairment were collected from over 25,000 respondents over a period of two decades. [7]. Self-reported vision has been established as a practical proxy measure for visual function, demonstrating a strong relationship with various objective visual function assessments [19,20]. This indicates that self-reported vision can serve as a reliable indicator of visual acuity. Furthermore, self-reported vision assessments offer a valuable advantage in under-resourced areas, where accessible and convenient tools for vision evaluation are scarce. By leveraging self-reported vision data, researchers and health care professionals can overcome logistical barriers and better understand vision-related needs in these communities. A better understanding of what aspects of vision impairment increase fall risk will help clarify promising vision interventions to decrease the rate of falls among older adults.

The objective of this study was to investigate the relationship between various aspects of self-reported vision impairment and falls and prospectively assess whether subjective vision impairment is predictive of future falls in a large and diverse longitudinal population-based study of older adults.

Methods

Data Collection

Data were from the HRS database, a longitudinal panel study that surveys a nationally representative sample of Americans aged 50 years and older every two years, with follow-up continuing until death. The HRS uses a biennial, multistage area probability sample survey design, using a combination of in-person and telephone interviews to collect data. The survey focuses on community-dwelling individuals. This longitudinal design enables the tracking of participants' health, economic, and social outcomes over time, providing a comprehensive understanding of aging and retirement in America. The HRS is supported by the National Institute on Aging (NIA U01AG009740) and the Social Security Administration. For this study, we used data from 14 waves of the HRS (1996 - 2020).

The analytic subsample for this study comprises community-dwelling respondents aged 65 years and older from the 1996 - 2020 HRS survey rounds, as the falls questionnaire was only administered to participants at or older than this age threshold, with all included variables having item nonresponse rates below 10%.

Ethical Considerations

Since the HRS database is deidentified and publicly accessible, this study was exempt from the Mass General Brigham institutional review board approval and patient permission was waived. The study adhered to the tenets of the Declaration of Helsinki.

Primary Outcomes and Definitions of Variables

The primary outcome measures included self-reported falls, serious injury falls, and the number of falls. Participants were asked if they had fallen in the past two years. If they answered "yes," they were then asked how many times they had fallen and whether any of the falls were serious. Serious injury falls were defined as those requiring medical treatment.

Ophthalmic parameters were acquired from the HRS survey assessments and included overall eyesight, distance vision, and near vision (see Multimedia Appendix 1), each rated on a scale from 1 to 5, with scores of 4 (fair) and 5 (poor) considered indicative of impaired vision. For overall eyesight, a score of 6 could be assigned for patients with legal blindness. The history of glaucoma was based on whether participants reported ever being treated for glaucoma.

The HRS also collected information on age, gender, race, level of education, marital status, self-reported diagnoses of hypertension, diabetes, stroke, heart conditions (including heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems), arthritis or rheumatism, smoking and alcohol consumption status, psychiatric medication use, exercise activities, and household income.

XSL•FO
RenderX

Study Timepoints of Data Analysis

The data were analyzed cross-sectionally and longitudinally to confirm the correlations between ophthalmic parameters and falls. Each participant observation was analyzed as individual encounters in both analyses. In the cross-sectional analysis, self-reported ophthalmic parameters and falls were ascertained from the same 2-year interview cycle. Each completed survey was counted as 1 entry. In the longitudinal analysis, the self-reported ophthalmic parameters were obtained during the survey and analyzed with the reported falls recorded in the patient's survey 2 years later.

Statistical Analysis

Data are presented as base-year weighted proportions, accounting for the Health and Retirement Study adjustments for attrition and sampling error. Descriptive statistics were used to summarize participant data on falls outcomes, including falls, no falls, serious falls, and nonserious falls. We used *t* tests for continuous variables and χ^2 tests for categorical variables to evaluate differences between outcomes. We used univariable and multivariable logistic regression models to determine the cross-sectional and longitudinal association between ophthalmic parameters and falls, adjusting for age and other potential confounders by calculating the odds ratio (OR) and their corresponding 95% CI. For the cross-sectional model, we assessed the association between vision status and reported falls in the same 2-year survey cycle, while for the longitudinal model, we considered falls reported in the subsequent 2-year survey cycle.

All variables with a *P* value of <.05 in the univariable analyses (including age, sex, Hispanic ethnicity, race, education, marital status, hypertension, diabetes, heart disease, stroke, arthritis, smoking, alcohol status, psychiatric medications, degree of exercise, and household income) were included in multivariable logistic regression models. Separate multivariable regression models were used to determine the relationship between each ophthalmic parameter and falls and serious falls. In addition, the association between the degree of vision impairment and the number of falls was assessed using a 1-way analysis of variance. All models were analyzed using generalized estimating equations for multiple observations from the same participant and multiple imputation techniques to replace the missing data. Data analysis was done with IBM SPSS Statistics 22.0 for Microsoft Windows and *P* values <.05 were considered statistically significant.

Results

Demographics of Study Sample

A total of 38,835 respondents contributed to 117,834 participant-observations (see Tables 1 and 2). The total number of observations in which participants reported a total of 40,477 falls and 13,471 serious falls. The weighted proportion of participants who reported any falls was OR 37.9% (95% CI 37.7% - 40.1%), while the proportion reporting serious falls was OR 30.9% (95% CI 30.7% - 31.1%). Furthermore, the weighted proportion of participants who experienced falls in the future was OR 36.5% (95% CI 35.8% - 37.2%), and OR 32.5% (95% CI 32.2% - 32.9%) for future serious falls. Falls increased with older age, and over half of those aged 90 years and older reported falling. The proportion of falls qualifying as serious also increased with age. White people had the highest fall rate, and a higher proportion of falls and serious falls was observed in females and individuals with lower levels of education. In total, 50% of patients with a history of stroke reported falling, with one-third of the falls being serious falls.

XSL•FO
RenderX

Table . Baseline characteristics of study participants comparing those who reported falls versus no falls (total number of participant-observations).

Characteristic	Participants with any self-reported falls (n=40,477)	Participants with no self-reported falls (n=77,357)	*P* value
Total (%), OR[a] (95% CI)	37.9 (37.7 - 40.1)	62.1 (61.9 - 62.3)	—[b]
Age in years (%), OR (95% CI)			<.001
65 - 69	28.5 (28.2 - 28.7)	71.5 (71.2 - 71.9)	
70 - 74	30.9 (30.5 - 31.4)	69.1 (68.6 - 69.7)	
75 - 79	32.9 (32.3 - 33.7)	67.1 (66.5-67.8)	
80 - 84	37.9 (37.5 - 38.5)	62.1 (61.6-62.9)	
85 - 89	44.8 (44.2 - 45.6)	55.2 (54.8-55.9)	
90+	52.2 (51.6-52.7)	47.8 (46.8-48.4)	
Sex (%), OR (95% CI)			<.001
Male	30.2 (29.8 - 30.8)	69.8 (69.2 - 70.3)	
Female	35.7 (35.3 - 36.1)	64.3 (63.8 - 64.8)	
Ethnicity (%), OR (95% CI)			<.001
Hispanic	33.1 (31.8 - 33.9)	66.9 (65.8 - 67.9)	
Non-Hispanic	34.7 (34.1 - 35.3)	65.3 (65.0 - 65.8)	
Race (%), OR (95% CI)			<.001
White	34.7 (34.1 - 35.3)	65.3 (64.3 - 66.4)	
Black or African American	28.9 (28.0 - 29.9)	71.1 (70.1 - 72.0)	
Others[c]	30.7 (29.9 - 31.8)	69.3 (68.6 - 70.1)	
Education (%), OR (95% CI)			<.001
<High school	35.1 (34.6 - 35.7)	64.9 (63.8 - 65.5)	
High school degree	33.2 (32.8 - 33.5)	66.8 (65.8 - 67.2)	
>High school degree	31.7 (31.1 - 32.3)	68.3 (67.8 - 68.8)	
Marital status (%), OR (95% CI)			<.001
Married	61.0 (60.8-61.2)	39.0 (38.9-39.1)	
Widowed, separated, or divorced	59.3 (59.0-59.8)	40.7 (40.2-41.3)	
Never married	38.7 (38.2-39.2)	61.3 (60.9-61.7)	
Comorbidities (%), OR (95% CI)			
Hypertension	35.7 (35.2-36.1)	64.3 (63.5-64.8)	<.001
Diabetes mellitus	39.7 (39.4-40.2)	60.3 (59.8-60.7)	<.001
Heart	42.0 (41.6-42.5)	58.0 (57.5-58.5)	<.001
Stroke	49.2 (48.3-50.0)	50.8 (50.3-51.5)	<.001
Arthritis (%), OR (95% CI)	36.8 (35.9 - 38.0)	63.2 (61.9 - 64.1)	<.001
Current smoking (%), OR (95% CI)	45.0 (44.3 - 46.0)	56.0 (54.9 - 57.1)	<.001
Active alcohol drinking (%), OR (95% CI)	55.2 (54.8 - 55.7)	44.8 (44.0 - 45.7)	<.001
Psychiatric medication (%), OR (95% CI)	62.1 (61.5 - 62.8)	37.9 (37.1 - 38.6)	<.001
Vigorous exercise (%), OR (95% CI)	70.6 (70.0 - 71.2)	29.4 (28.8 - 29.7)	<.001
Household income (US $), mean (95% CI)	64,319.1 (56,172.2 - 71,816.3)	69,983.7 (45,562.1 - 73,805.4)	<.001

[a]OR: odds ratio.

XSL•FO
RenderX

[b]Not applicable.

[c]Other races: Asian, Latino, and Native American.

XSL•FO
RenderX

Table . Baseline characteristics of study participants comparing those who reported serious falls versus no serious falls (total number of participant-observations).

Characteristics	Participants with any self-reported serious falls (n=13,471)	Participants with no self-reported serious falls (n=27,006)	*P* value
Total (%), OR[a] (95% CI)	30.9 (30.7 - 31.1)	69.1 (68.9 - 69.3)	—[b]
Age in years (%), OR (95% CI)			<.001
65-69	21.4 (21.3 - 21.5)	78.6 (78.3-78.8)	
70-74	31.2 (31.0-31.4)	68.8 (68.4-69.1)	
75-79	40.6 (40.2-40.9)	59.4 (59.0-59.8)	
80-84	42.6 (42.1-43.0)	57.4 (56.9-57.9)	
85-89	43.9 (43.0-44.7)	56.1 (55.7-56.6)	
90+	39.9 (39.3-40.6)	60.1 (59.1-61.1)	
Sex (%), OR (95% CI)			<.001
Male	23.8 (23.7 - 23.9)	76.2 (75.9 - 76.5)	
Female	39.1 (38.9-39.4)	60.9 (60.3-61.5)	
Ethnicity (%), OR (95% CI)			<.001
Hispanic	26.9 (26.3 - 27.5)	73.1 (72.4 - 73.9)	
Non-Hispanic	31.8 (31.2 - 32.4)	68.2 (67.6 - 68.8)	
Race (%), OR (95% CI)			<.001
White	35.8 (35.2 - 36.3)	64.2 (63.7 - 64.8)	
Black or African American	20.0 (19.6 - 20.5)	80.0 (79.7 - 80.4)	
Others[c]	19.9 (19.1 - 20.8)	80.1 (79.6 - 80.9)	
Education (%), OR (95% CI)			<.001
<High school	31.9 (31.3 - 32.7)	68.1 (67.5 - 68.6)	
High school degree	34.1 (33.7 - 34.5)	65.9 (65.5 - 66.6)	
>High school degree	26.9 (26.2 - 27.4)	3.1 (72.7 - 73.6)	
Marital status (%), OR (95% CI)			<.001
Married	31.8 (31.3 - 32.3)	68.2 (67.9 - 68.6)	
Widowed, separated, or divorced	32.8 (32.3 - 33.3)	67.2 (66.8 - 67.9)	
Never married	17.8 (17.3 - 18.2)	82.2 (81.8 - 82.6)	
Comorbidities (%), OR (95% CI)			
Hypertension	32.8 (32.2-33.7)	67.2 (66.9-67.6)	<.001
Diabetes mellitus	32.3 (31.9-33.6)	67.7 (67.0-68.3)	<.001
Heart	35.0 (34.4-35.8)	65.0 (64.6-65.8)	<.001
Stroke	35.6 (34.5-36.3)	64.4 (64.1-64.7)	<.001
Arthritis (%), OR (95% CI)	33.3 (33.3 - 34.3)	66.7 (65.9 - 66.9)	<.001
Current smoking (%), OR (95% CI)	22.2 (21.6 - 22.9)	77.8 (77.0 - 78.7)	<.001
Active alcohol drinking (%), OR (95% CI)	28.6 (28.0 - 29.8)	71.4 (70.8 - 71.9)	<.001
Psychiatric medication (%), OR (95% CI)	35.1 (34.5 - 35.7)	64.9 (64.1 - 65.8)	<.001
Vigorous exercise (%), OR (95% CI)	40.2 (39.4 - 40.9)	59.8 (59.1 - 60.6)	<.001
Household income, mean (95% CI)	62,595.1 (53,582.5 - 70,419.8)	64,289.8 (58,925.3 - 69,643.1)	<.001

[a]OR: odds ratio.

XSL•FO
RenderX

[b]Not applicable.

[c]Other races: Asian, Latino, and Native American.

Ophthalmic Parameters and Falls

Higher rates of any type of falls and serious falls were observed with worse eyesight (see Tables 3 and 4). Approximately half of the participants who reported poor vision or blindness also reported falling, while 42.8% (95% CI 38.3% - 46.7%) of those with blindness had serious injuries associated with the fall. Of the participants with poor distant and near vision, 53.2% (95% CI 52.7% - 54.1%) and 50.2% (95% CI 49.9% - 50.6%) reported falls, respectively, and over a third of the falls were serious. In addition, falls were reported by 39.6% (95% CI 38.8% - 40.5%) of those who stated they had glaucoma.

Table . Ophthalmic parameters of the study sample between patients who reported falls and no falls (total number of participant-observations).

Characteristic	Participants with any self-reported falls (n=40,477), OR[a] (%), (95% CI)	Participants with no self-reported falls (n=77,357), OR (%), (95% CI)	*P* value
Overall	37.9 (37.7 - 40.1)	62.1 (61.9 - 62.3)	—[b]
Overall eyesight			<.001
Excellent	28.7 (28.3-29.7)	71.3 (69.7-71.9)	
Very good	29.1 (28.8-29.5)	70.9 (70.1-71.8)	
Good	32.2 (32.0-32.5)	67.8 (66.9-68.8)	
Fair	38.9 (38.1-40.5)	61.1 (59.7-63.1)	
Poor	51.1 (50.7-51.4)	48.9 (47.6-49.7)	
Blind	48.3 (46.9-49.9)	51.7 (49.1-54.1)	
Distance vision			<.001
Excellent	29.2 (28.1-30.0)	70.8 (69.4-71.9)	
Very good	29.8 (28.8-30.8)	70.2 (69.2-71.2)	
Good	32.8 (31.8-33.8)	67.2 (66.2-67.9)	
Fair	41.8 (41.0-42.9)	58.2 (56.8-59.3)	
Poor	53.2 (52.7-54.1)	46.8 (44.9-47.8)	
Near vision			<.001
Excellent	28.8 (28.3-30.0)	71.2 (69.8-71.9)	
Very good	29.8 (29.1-30.5)	70.2 (69.3-70.9)	
Good	32.2 (32.1-33.0)	67.8 (66.7-68.4)	
Fair	40.2 (39.7-41.0)	59.8 (58.5-60.8)	
Poor	50.2 (49.9-50.6)	49.8 (47.9-51.1)	
Glaucoma			<.001
Yes	39.6 (38.8 - 40.5)	60.4 (58.5 - 62.2)	
No	32.8 (32.0 - 33.8)	67.2 (66.4 - 69.0)	

[a]OR: odds ratio.

[b]Not applicable.

XSL•FO
RenderX

Table . Ophthalmic parameters of the study sample between patients who reported serious falls and no serious falls (total number of participant-observations).

Characteristic	Participants with any self-reported serious falls (n=13,471), OR[a] (%), (95% CI)	Participants with no self-reported serious falls (n=27,006), OR (%), (95% CI)	*P* value
Overall	30.9 (30.7 - 31.1)	69.1 (68.9 - 69.3)	—[b]
Overall eyesight			<.001
Excellent	31.7 (30.1 - 33.7)	68.3 (65.3 - 69.9)	
Very good	30.8 (29.6 - 31.7)	69.2 (67.4 - 70.2)	
Good	31.7 (30.5 - 33.0)	68.3 (67.0 - 69.7)	
Fair	33.3 (32.3 - 34.8)	67.7 (65.1 - 69.1)	
Poor	37.9 (36.8 - 39.6)	62.1 (60.4 - 64.3)	
Blind	42.8 (38.3 - 46.7)	57.2 (51.3 - 63.3)	
Distance vision			<.001
Excellent	30.1 (29.7 - 32.4)	69.9 (67.6 - 71.4)	
Very good	31.7 (30.4 - 33.3)	68.3 (66.7 - 69.5)	
Good	31.8 (29.9 - 32.7)	68.2 (67.3 - 69.2)	
Fair	34.8 (33.5 - 36.2)	65.2 (63.0 - 66.6)	
Poor	39.1 (37.9 - 41.2)	60.9 (58.4 - 62.0)	
Near vision			<.001
Excellent	31.6 (30.5 - 32.6)	68.4 (66.4 - 69.2)	
Very good	31.2 (30.7 - 32.9)	68.8 (67.1 - 69.8)	
Good	32.6 (31.4 - 33.8)	67.4 (67.3 - 68.9)	
Fair	34.2 (33.5 - 35.9)	65.8 (63.1 - 67.6)	
Poor	38.6 (37.6 - 40.6)	61.4 (59.4 - 63.1)	
Glaucoma			<.001
Yes	37.2 (35.4 - 39.5)	62.8 (60.3 - 64.8)	
No	32.3 (32.1 - 33.4)	67.7 (65.7 - 68.9)	

[a]OR: odds ratio.

[b]Not applicable.

Cross-Sectional Analysis

Cross-sectionally, vision impairment was associated with higher odds of falls: impaired overall eyesight (aOR 1.36, 95% CI 1.20 - 1.56; *P*<.001), impaired distance vision (aOR 1.37, 95% CI 1.32 - 1.42; *P*<.001), and impaired near vision (aOR 1.33, 95% CI 1.27 - 1.37; *P*<.001) all conferred significantly increased odds of falls after adjusting for confounding factors (see Table 5 and Figure 1). The odds of serious falls increased in impaired overall eyesight with adjusted ORs of 1.20 (95% CI 1.03 - 1.44; *P*<.001), while having overall eyesight qualifying as legal blindness increased the odds of falls (aOR 1.32, 95% CI 1.26 - 1.37; *P*<.001) and serious falls (aOR 1.16, 95% CI 1.03 - 1.21; *P*<.001) compared to not having impaired vision. Self-reported glaucoma increased the adjusted odds of falls (aOR 1.15, 95% CI 1.07 - 1.24; *P*<.001) and serious falls (aOR 1.15, 95% CI 1.05 - 1.26; *P*<.001).

XSL•FO
RenderX

Table . Cross-sectional model showing the association between falls and ophthalmic parameters.

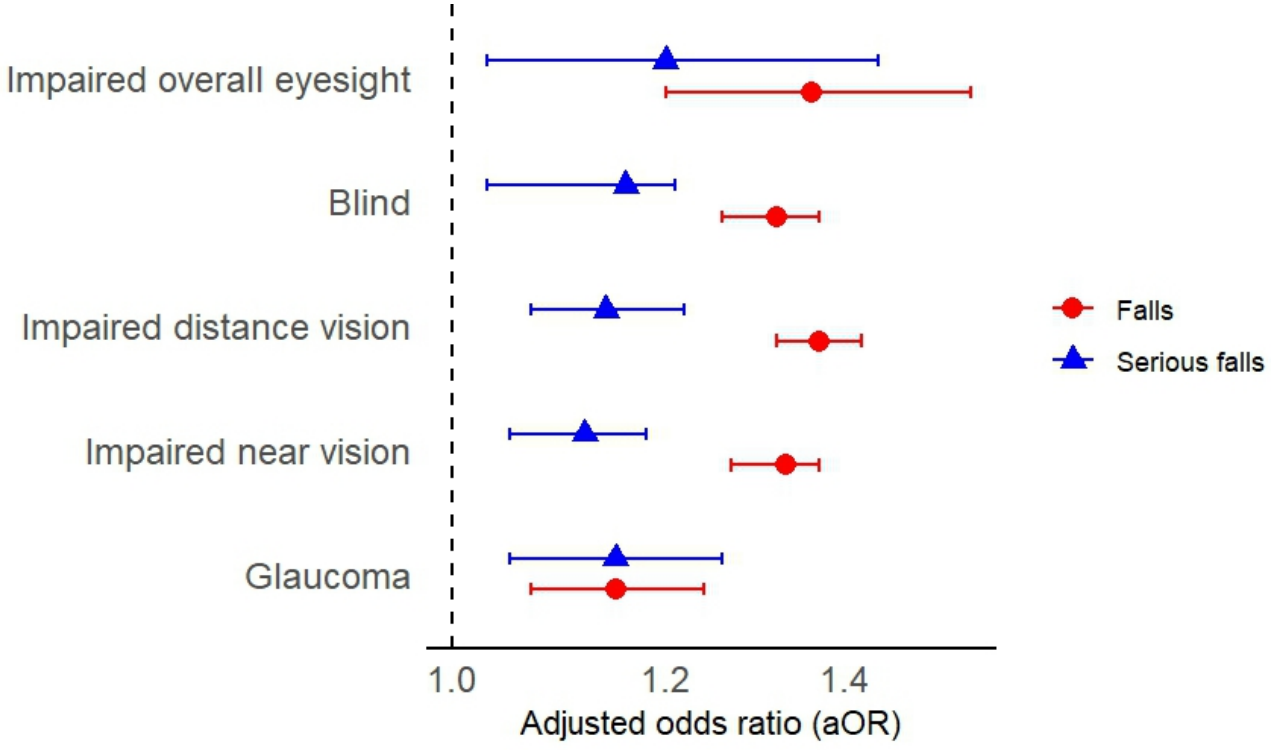
Type of falls	OR[a] (95% CI)	*P* value	Adjusted OR (95% CI)[b]	*P* value
Falls				
Overall eyesight				
Not impaired	Ref[c]	Ref	Ref	Ref
Impaired	1.60 (1.52 - 1.73)	<.001	1.36 (1.20 - 1.56)	<.001
Blind	2.03 (1.83 - 2.44)	<.001	1.32 (1.26 - 1.37)	<.001
Impaired distance vision	1.78 (1.70 - 1.89)	<.001	1.37 (1.32 - 1.42)	<.001
Impaired near vision	1.89 (1.69 - 1.99)	<.001	1.33 (1.27 - 1.37)	<.001
Glaucoma	1.33 (1.27 - 1.41)	<.001	1.15 (1.07 - 1.24)	<.001
Serious falls				
Overall eyesight				
Not impaired	Ref	Ref	Ref	Ref
Impaired	1.23 (1.11 - 1.31)	<.001	1.20 (1.03 - 1.44)	.045
Blind	1.49 (1.24 - 1.89)	<.001	1.16 (1.03 - 1.21)	.005
Impaired distance vision	1.22 (1.19 - 1.27)	<.001	1.14 (1.07 - 1.22)	.007
Impaired near vision	1.21 (1.14 - 1.29)	<.001	1.12 (1.05 - 1.18)	<.001
Glaucoma	1.24 (1.12 - 1.33)	<.001	1.15 (1.05 - 1.26)	.004

[a]OR: odds ratio.

[b]Adjusted for age, sex, Hispanic ethnicity, race, education, marital status, hypertension, diabetes, heart disease, stroke, arthritis, smoking, alcohol, activity, psychiatric medication, vigorous activities, and household income.

[c]Reference.

Figure 1. Forest plot illustrating the association between falls and ophthalmic parameters, based on a cross-sectional multivariable model using generalized estimating equations.



XSL•FO
RenderX

Longitudinal Analysis

In the longitudinal analysis of the association between ophthalmic parameters with falls and serious falls, impaired overall eyesight, impaired distance vision, impaired near vision, and self-reported glaucoma all increased the odds of future falls with an OR of 1.23 (95% CI 1.16 - 1.29; *P*<.001), 1.27 (95% CI 1.20 - 1.38; *P*<.001), 1.23 (95% CI 1.19 - 1.32; *P*<.001), and 1.25 (95% CI 1.13 - 1.37; *P*<.001), respectively (see Table 6 and Figure 2). Impaired overall eyesight, distance vision, and near vision also increased the risk of serious falls in the subsequent 2 years with an OR of 1.18 (95% CI 1.05 - 1.23; *P*<.001), 1.13 (95% CI 1.04 - 1.21; *P*<.001), and 1.15 (95% CI 1.07 - 1.22; *P*<.001), respectively. There was no association between self-reported glaucoma and the odds of future serious falls (*P*=.32).

Table . Longitudinal model showing the association between falls and ophthalmic parameters.

Type of falls	OR[a] (95% CI)	*P* value	Adjusted OR (95% CI)[b]	*P* value
Falls				
Overall eyesight				
Not impaired	Ref[c]	Ref	Ref	Ref
Impaired	1.59 (1.53 - 1.67)	<.001	1.23 (1.16 - 1.29)	<.001
Blind	2.03 (1.65 - 2.68)	<.001	1.53 (1.14 - 2.07)	.005
Impaired distance vision	1.73 (1.63 - 1.88)	<.001	1.27 (1.20 - 1.38)	.001
Impaired near vision	1.69 (1.63 - 1.75)	<.001	1.23 (1.19 - 1.32)	<.001
Glaucoma	1.40 (1.23 - 1.52)	<.001	1.25 (1.13 - 1.37)	<.001
Serious falls				
Overall eyesight				
Not impaired	Ref	Ref	Ref	Ref
Impaired	1.17 (1.10 - 1.23)	<.001	1.18 (1.05 - 1.23)	<.001
Blind	1.45 (1.17 - 1.93)	<.001	1.31 (1.04 - 1.66)	.046
Impaired distance vision	1.20 (1.13 - 1.26)	<.001	1.13 (1.04 - 1.21)	.004
Impaired near vision	1.19 (1.13 - 1.26)	<.001	1.15 (1.07 - 1.22)	<.001
Glaucoma	1.15 (1.13 - 1.17)	.039	1.03 (0.89 - 1.15)	.32

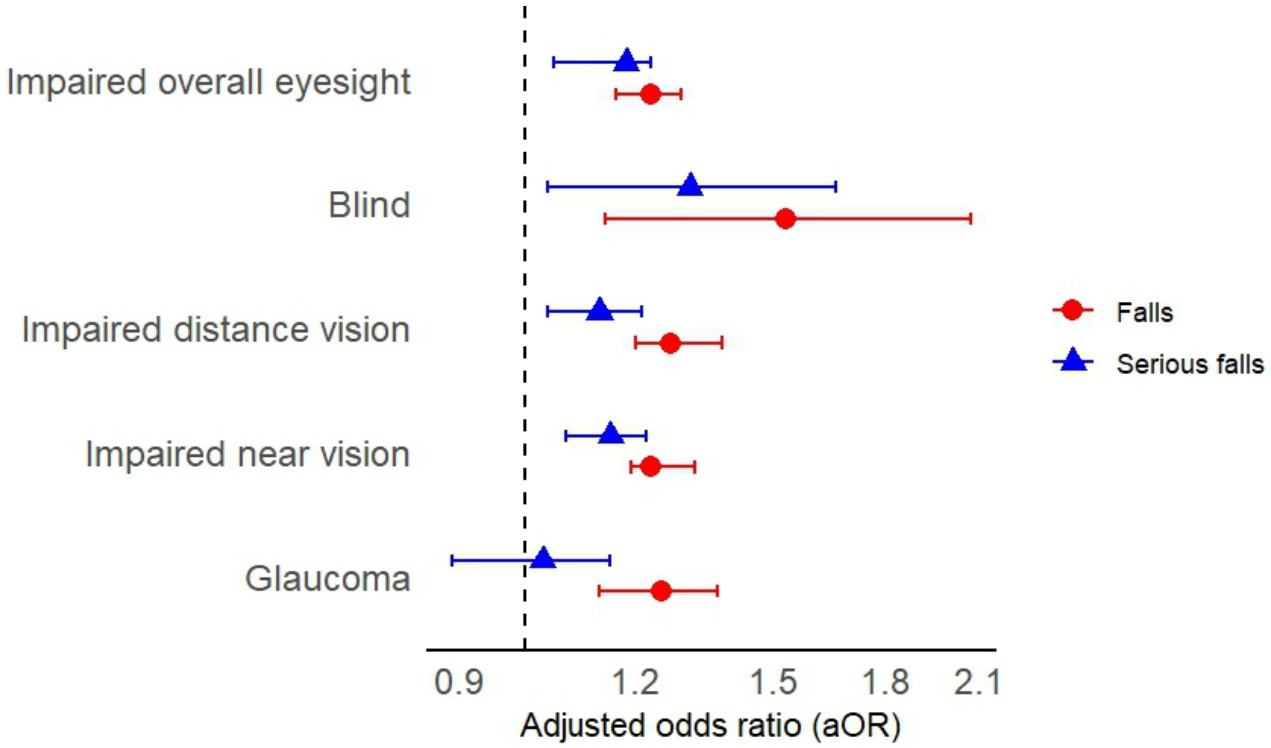
[a]OR: odds ratio.

[b]Adjusted for age, sex, Hispanic ethnicity, race, education, marital status, hypertension, diabetes, heart disease, stroke, arthritis, smoking, alcohol, activity, psychiatric medication, vigorous activities, and household income.

[c]Reference.

XSL•FO
RenderX

Figure 2. Forest plot illustrating the association between falls and ophthalmic parameters, based on a longitudinal multivariable model using Generalized Estimating Equation.



Analysis Between Mean Number of Falls and Eyesight

The mean number of falls was associated with the degree of self-reported vision impairment (see Table 7). The level of eyesight reported in the questionnaire was associated with the reported number of falls in the same survey cycle as well as the subsequent cycle (P<.001 for both). Approximately 2.57 (SD 4.58) falls per person who fell were observed in the group with excellent current eyesight compared to 3.70 (SD 5.33) for those with poor reported vision. Similarly, self-reported excellent overall eyesight had an average of 2.78 (SD 4.17) falls per person while poor overall eyesight had approximately 3.65 (SD 5.40) falls per person in the subsequent cycle.

Table . The association of the degree of visual impairment and mean number of falls.

Degree of impairment	Excellent	Very good	Good	Fair	Poor	*P* value
Cross-sectional, mean (SD)						
Overall eyesight	2.57 (4.58)	2.58 (3.84)	2.74 (4.05)	3.15 (4.59)	3.70 (5.33)	<.001
Distal eyesight	2.66 (4.16)	2.70 (4.04)	2.79 (4.05)	3.34 (4.94)	3.73 (5.26)	<.001
Near eyesight	2.59 (3.86)	2.65 (4.23)	2.77 (4.04)	3.2 (4.52)	3.88 (5.79)	<.001
Longitudinal, mean (SD)						
Overall eyesight	2.78 (4.17)	2.71 (4.10)	2.84 (4.18)	3.19 (4.72)	3.65 (5.40)	<.001
Distal eyesight	2.74 (4.13)	2.75 (3.95)	2.93 (4.40)	3.27 (4.89)	3.59 (5.08)	<.001
Near eyesight	2.66 (3.92)	2.72 (4.01)	2.91 (4.37)	3.28 (4.92)	3.62 (5.10)	<.001

Discussion

Principal Findings

In this large and diverse longitudinal population-based study of older adults, poor vision, including both distance and near vision, increased the odds of falling, and more severe vision loss was related to the number of falls. We additionally found that the presence of self-reported glaucoma was associated with falls.

Comparison With Previous Studies

Many population-based studies have found associations between self-reported vision impairment and falls. The Survey of Health, Ageing, and Retirement in Europe (SHARE) study found that self-reported distance and near vision impairment were cross-sectionally and longitudinally associated with increased risk of falls [21]. The Beaver Dam Eye Study reported that people with worse binocular acuity had higher odds of 2 or more falls [12], while the EPIC-Norfolk eye study found significant associations between poor self-reported distance vision and falls

XSL•FO
RenderX

independent of visual acuity [20]. Similarly, we demonstrate here that the presence of visual impairment, including overall eyesight, near vision, and distance vision, is all associated with significantly increased odds for falls and severe falls in both our cross-sectional and longitudinal analyses. Our results suggest that self-reported visual status is predictive of falls and serious falls.

Importantly, the degree of self-reported visual impairment was associated with the number of falls even after adjusting for potential confounders. When looking at the mean number of falls that occur per person for each level of reported vision, we found that though people with excellent vision had falls, the number increased significantly as vision worsened, with those with blindness having the highest risk of falls. This is true for both falls that occurred 2 years before and those that occurred in the following 2 years of the reported vision and does not take into account the fact that blindness likely reduces the number of steps taken.

This study also subclassified falls as serious based on whether a fall had any associated injuries requiring medical treatment. The incidence of serious falls increased with age and deterioration of eyesight, and we found the risk factors for falls and serious falls to be similar. Other studies, including a Finnish population-based study, have examined risk factors of major injurious falls and similarly found poor distant visual acuity to be a risk factor for major falls among disabled elder adults [22].

Visual input plays a critical role in coordinating and planning movement, as well as maintaining balance [23]. Individuals with poor vision are more likely to lose stability, alter their gait to avoid obstacles, and ultimately increase their risk of slips, trips, and falls [24]. Our study differs from previous research in this area, which has largely relied on data from previous falls and visual function assessments. By using self-reported data, our research offers a fresh perspective and highlights the potential of self-reported visual function data as a predictive tool for falls risk.

We also found that glaucoma is a risk factor for serious falls resulting in injury despite the fact that our definition of glaucoma was imperfect (based on self-report only). Any nondifferential misclassification in diagnosing glaucoma would have driven our associations toward the null, and this implies that the associations found were real and may have been underestimated. Previous studies have also reported glaucoma as an independent risk factor for falling [13]. Glaucoma severity has been reported as a risk factor for falls, with advanced glaucoma patients having the highest risk of falls compared to controls with no glaucoma and mild or moderate glaucoma [25]. While central vision and, therefore, visual acuity are often preserved in glaucoma until severe stages of disease, glaucoma patients experience changes in multiple other aspects of vision, including VF loss, decrease in contrast sensitivity, and loss of depth perception. All these visual measures have been shown to independently increase the risk of falling [13,14,26]. Indeed, the Salisbury Eye Evaluation found that loss in peripheral VF is a more important risk factor for falls than loss of central VF [14]. Patients with glaucoma may experience one or all visual deficits simultaneously, further increasing the risk of falling.

Studies have suggested various interventions to reduce the risk of falls, including mobility training, vision interventions to improve vision (eye examinations, cataract surgery, and prescription glasses) [27], risk factor assessment, and environmental modifications to make living environment less hazardous [28]. Visual impairment leads to physical impairment of mobility and balance control, increasing the risk of falls [29]. Glaucoma initially affects peripheral vision through retinal nerve fiber layer loss, which is associated with decreased postural stability; therefore, glaucoma patients may benefit from mobility training to reduce falling risks [30,31].

Clinical Implications

Our results provide evidence to support the development of multifaceted fall prevention strategies. Specifically, our findings suggest that self-reported vision screening could be a valuable tool in identifying individuals at risk of falls. This screening could be easily integrated into telemedicine platforms or incorporated into standard clinical guidelines. Furthermore, we propose the development of a simple, score-based questionnaire that can be used to predict falls risk in older adults. Those with higher scores could be targeted for earlier intervention, thereby reducing their risk of falls.

Limitations

There were several limitations to our study. All data used in this study were from self-reported questionnaires administered every 2 years and are susceptible to recall bias and recall error. Specifically, participants may have difficulty recalling their visual function or falls history, which could lead to inaccurate reporting. In addition, while self-reported visual function has been shown to correlate closely with objective visual function, it is also understood as a multidimensional measure that reflects some, but not all components of vision [19]. Since no eye examinations were done and objective visual function was not measured, we cannot determine which visual components contributed to participants' self-reported visual function. This study is also limited by the specific questions collected as part of the HRS study. No questions were asked about the other leading causes of blindness and visual impairment in this older adult population, such as age-related macular degeneration and other neurological conditions that could affect falls, such as Parkinson disease, dementia, or epilepsy. Finally, we could not explore other known vision-related risk factors for falls such as contrast sensitivity, VF deterioration, depth perception, and stereoacuity [26]. Nonvision factors such as cognitive ability, balance, motor control, and degree of mobility, which are known to be associated with falls, were not evaluated since they were not included in the questionnaire. The omission of these factors may limit the generalizability of our findings and highlights the need for future studies to incorporate a more comprehensive set of measures.

Conclusions

Our analysis of data from a large, longitudinal national survey of older adults demonstrated that vision-related risk factors, particularly the presence and degree of visual impairment and glaucoma, increase the rate of falls and serious falls. We additionally show that self-reported vision status correlates with

XSL•FO
RenderX

the number of fall events both at the time of and occurring after the survey date. Our results support screening for the presence of visual impairment and signs of glaucoma in older adults to identify individuals at the highest risk of falling who might most benefit from fall prevention interventions.

Acknowledgments

The Health and Retirement Study is supported by the National Institute on Aging, supplemented by the Social Security Agency, and operated from the Institute for Social Research at the University of Michigan.

Data Availability

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

Conflicts of Interest

None declared.

Multimedia Appendix 1
Self-reported visual function and falls questions in the Health and Retirement Study (HRS).
[DOCX File, 15 KB - aging_v8i1e68771_app1.docx]

References

1. Step safely: strategies for preventing and managing falls across the life-course. : World Health Organization; 2021 URL: https://www.who.int/publications/i/item/978924002191-4 [accessed 2025-07-02]
2. Bergen G, Stevens MR, Burns ER. Falls and fall injuries among adults aged ≥65 years - United States, 2014. MMWR Morb Mortal Wkly Rep 2016 Sep 23;65(37):993-998. [doi: 10.15585/mmwr.mm6537a2] [Medline: 27656914]
3. Moreland B, Kakara R, Henry A. Trends in nonfatal falls and fall-related injuries among adults aged ≥65 years - United States, 2012-2018. MMWR Morb Mortal Wkly Rep 2020 Jul 10;69(27):875-881. [doi: 10.15585/mmwr.mm6927a5] [Medline: 32644982]
4. Rubenstein LZ, Josephson KR. The epidemiology of falls and syncope. Clin Geriatr Med 2002 May;18(2):141-158. [doi: 10.1016/s0749-0690(02)00002-2] [Medline: 12180240]
5. Tinetti ME. Clinical practice. Preventing falls in elderly persons. N Engl J Med 2003 Jan 2;348(1):42-49. [doi: 10.1056/NEJMcp020719] [Medline: 12510042]
6. Facts about falls. : CDC; 2023 URL: https://www.cdc.gov/falls/data-research/facts-stats/index.html [accessed 2025-07-02]
7. Wang H, Naghavi M, Allen C, et al. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. The Lancet 2016 Oct;388(10053):1459-1544. [doi: 10.1016/S0140-6736(16)31012-1]
8. Close J, Ellis M, Hooper R, Glucksman E, Jackson S, Swift C. Prevention of falls in the elderly trial (PROFET): a randomised controlled trial. Lancet 1999 Jan 9;353(9147):93-97. [doi: 10.1016/S0140-6736(98)06119-4] [Medline: 10023893]
9. Ambrose AF, Paul G, Hausdorff JM. Risk factors for falls among older adults: a review of the literature. Maturitas 2013 May;75(1):51-61. [doi: 10.1016/j.maturitas.2013.02.009] [Medline: 23523272]
10. Dhargave P, Sendhilkumar R. Prevalence of risk factors for falls among elderly people living in long-term care homes. Journal of Clinical Gerontology and Geriatrics 2016 Sep;7(3):99-103. [doi: 10.1016/j.jcgg.2016.03.004]
11. Patil S, Suryanarayana S, Shivraj N, Murthy N, Dinesh R. Risk factors for falls among elderly: a community-based study. Int J Health Allied Sci 2015;4(3):135. [doi: 10.4103/2278-344X.160867]
12. Klein BEK, Moss SE, Klein R, Lee KE, Cruickshanks KJ. Associations of visual function with physical outcomes and limitations 5 years later in an older population: the Beaver Dam eye study. Ophthalmology 2003 Apr;110(4):644-650. [doi: 10.1016/S0161-6420(02)01935-8] [Medline: 12689880]
13. Ivers RQ, Cumming RG, Mitchell P, Attebo K. Visual impairment and falls in older adults: the Blue Mountains Eye Study. J Am Geriatr Soc 1998 Jan;46(1):58-64. [doi: 10.1111/j.1532-5415.1998.tb01014.x] [Medline: 9434666]
14. Freeman EE, Muñoz B, Rubin G, West SK. Visual field loss increases the risk of falls in older adults: the Salisbury eye evaluation. Invest Ophthalmol Vis Sci 2007 Oct;48(10):4445-4450. [doi: 10.1167/iovs.07-0326] [Medline: 17898264]
15. Nevitt MC, Cummings SR, Kidd S, Black D. Risk factors for recurrent nonsyncopal falls. A prospective study. JAMA 1989 May 12;261(18):2663-2668. [Medline: 2709546]
16. Campbell AJ, Borrie MJ, Spears GF. Risk factors for falls in a community-based prospective study of people 70 years and older. J Gerontol 1989 Jul;44(4):M112-M117. [doi: 10.1093/geronj/44.4.m112] [Medline: 2738307]
17. Lord SR, Ward JA, Williams P, Anstey KJ. Physiological factors associated with falls in older community‐dwelling women. J American Geriatrics Society 1994 Oct;42(10):1110-1117. [doi: 10.1111/j.1532-5415.1994.tb06218.x]

XSL•FO
RenderX

18. Fisher GG, Ryan LH. Overview of the health and retirement study and introduction to the special issue. Work Aging Retire 2018 Jan;4(1):1-9. [doi: 10.1093/workar/wax032] [Medline: 29423243]
19. El-Gasim M, Munoz B, West SK, Scott AW. Associations between self-rated vision score, vision tests, and self-reported visual function in the Salisbury Eye Evaluation Study. Invest Ophthalmol Vis Sci 2013 Sep 27;54(9):6439-6445. [doi: 10.1167/iovs.12-11461] [Medline: 23812494]
20. Yip JLY, Khawaja AP, Broadway D, et al. Visual acuity, self-reported vision and falls in the EPIC-Norfolk Eye study. Br J Ophthalmol 2014 Mar;98(3):377-382. [doi: 10.1136/bjophthalmol-2013-304179] [Medline: 24338086]
21. Ogliari G, Ryg J, Qureshi N, Andersen-Ranberg K, Scheel-Hincke LL, Masud T. Subjective vision and hearing impairment and falls among community-dwelling adults: a prospective study in the Survey of Health, Ageing and Retirement in Europe (SHARE). Eur Geriatr Med 2021 Oct;12(5):1031-1043. [doi: 10.1007/s41999-021-00505-4] [Medline: 34003480]
22. Koski K, Luukinen H, Laippala P, Kivelä SL. Risk factors for major injurious falls among the home-dwelling elderly by functional abilities. A prospective population-based study. Gerontology 1998;44(4):232-238. [doi: 10.1159/000022017] [Medline: 9657085]
23. Manchester D, Woollacott M, Zederbauer-Hylton N, Marin O. Visual, vestibular and somatosensory contributions to balance control in the older adult. J Gerontol 1989 Jul;44(4):M118-M127. [doi: 10.1093/geronj/44.4.m118] [Medline: 2786896]
24. Buckley JG, Panesar GK, MacLellan MJ, Pacey IE, Barrett BT. Changes to control of adaptive gait in individuals with long-standing reduced stereoacuity. Invest Ophthalmol Vis Sci 2010 May;51(5):2487-2495. [doi: 10.1167/iovs.09-3858] [Medline: 20335609]
25. Bhorade AM, Perlmutter MS, Sabapathypillai SL, Goel M, Wilson B, Gordon MO. Rate of falls, fear of falling, and avoidance of activities at-risk for falls in older adults with glaucoma. Am J Ophthalmol 2021 Jul;227:275-283. [doi: 10.1016/j.ajo.2021.02.017] [Medline: 33626364]
26. Mehta J, Czanner G, Harding S, Newsham D, Robinson J. Visual risk factors for falls in older adults: a case-control study. BMC Geriatr 2022 Feb 17;22(1):134. [doi: 10.1186/s12877-022-02784-3] [Medline: 35177024]
27. Zhang XY, Shuai J, Li LP. Vision and relevant risk factor interventions for preventing falls among older people: a network meta-analysis. Sci Rep 2015 May 28;5(1):10559. [doi: 10.1038/srep10559] [Medline: 26020415]
28. Sotimehin AE, Yonge AV, Mihailovic A, et al. Locations, circumstances, and outcomes of falls in patients with glaucoma. Am J Ophthalmol 2018 Aug;192:131-141. [doi: 10.1016/j.ajo.2018.04.024] [Medline: 29750950]
29. Zwierko T, Jedziniak W, Florkiewicz B, et al. The consequences of glaucoma on mobility and balance control in the older adults: a cross-sectional study. J Aging Phys Act 2021 Jun 1;29(3):372-381. [doi: 10.1123/japa.2020-0079] [Medline: 32994380]
30. O'Connell C, Mahboobin A, Drexler S, et al. Effects of acute peripheral/central visual field loss on standing balance. Exp Brain Res 2017 Nov;235(11):3261-3270. [doi: 10.1007/s00221-017-5045-x] [Medline: 28765993]
31. Black AA, Wood JM, Lovie-Kitchin JE, Newman BM. Visual impairment and postural sway among older adults with glaucoma. Optom Vis Sci 2008 Jun;85(6):489-497. [doi: 10.1097/OPX.0b013e31817882db] [Medline: 18521027]

Abbreviations

aOR: adjusted odds ratio
HRS: Health and Retirement Study
OR: odds ratio
VF: visual field

Edited by R Bjarnadottir; submitted 24.11.24; peer-reviewed by P Roberts, Q Mao, Y Gu; revised version received 11.03.25; accepted 21.04.25; published 31.07.25.

Please cite as:
Seresirikachorn K, Somyanonthanakul R, Johnson M, Singhanetr P, Gatedee J, Friedman D, Zebardast N
The Impact of Vision Impairment on Self-Reported Falls Among Older US Adults: Cross-Sectional and Longitudinal Study
JMIR Aging 2025;8:e68771
URL: https://aging.jmir.org/2025/1/e68771
doi:10.2196/68771

© Kasem Seresirikachorn, Rachasak Somyanonthanakul, Matthew Johnson, Panisa Singhanetr, Jiraporn Gatedee, David Friedman, Nazlee Zebardast. Originally published in JMIR Aging (https://aging.jmir.org), 31.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is

XSL•FO
RenderX

properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

AI-Assisted Dynamic Postural Control Screening to Improve Functional Mobility in Older Adult Populations: Quasi-Experimental Study

Kai-Chih Lin[1*], MSc; Rong-Jong Wai[1*], PhD; Hung-Yu Chang Chien[2*], MSHP

[1]Department of Electronic and Computer Engineering, National Taiwan University of Science and Technology, No. 43, Sec. 4, Keelung Rd., Da'an Dist., Taipei, Taiwan

[2]Department of Health Promotion and Health Education, National Taiwan Normal University, Taipei, Taiwan

[*]all authors contributed equally

Corresponding Author:
Rong-Jong Wai, PhD
Department of Electronic and Computer Engineering, National Taiwan University of Science and Technology, No. 43, Sec. 4, Keelung Rd., Da'an Dist., Taipei, Taiwan

Abstract

Background: Falls are a major cause of disability among older adults, and early identification of functional decline is essential for prevention. Artificial intelligence (AI) systems may enhance mobility screening by providing objective, real-time feedback.

Objective: This study aimed to evaluate whether AI-assisted dynamic postural control screening combined with adaptive training improves functional mobility outcomes in older adult populations.

Methods: A quasi-experimental study was conducted with 2005 older adults recruited from community centers and health care institutions in Keelung, Taiwan. Participants were assigned to either an experimental group (n=1451), which underwent AI-assisted screening with adaptive exercise prescriptions, or a control group (n=554), which completed follow-ups through regular physical assessments with standard care without AI-tailored training. The AI system integrated skeletal tracking with the Short Physical Performance Battery to assess balance, gait speed (4-m walk), and sit-to-stand performance. Independent-samples 2-tailed *t* tests and repeated-measures ANOVA were applied, and effect sizes (Cohen *d* and η^2) with 95% CIs were reported.

Results: The experimental group demonstrated significantly greater improvements compared with the control group in Short Physical Performance Battery scores (Δ=0.8 vs 0.3; t_{2003}=3.41; *P*=.001; Cohen *d*=0.45, 95% CI 0.18 - 0.72), gait speed (Δ=15 cm/s vs 5 cm/s; t_{2003}=4.85; *P*<.001; Cohen *d*=0.62, 95% CI 0.35 - 0.88), and sit-to-stand time (Δ=–1.4 s vs –0.6 s; t_{2003}=3.12; *P*=.002; Cohen *d*=0.39, 95% CI 0.12 - 0.65). Here "Δ" refers to the change score, calculated as post-intervention minus baseline (ie, the amount of improvement during the study period). Participation rate was strongly associated with outcomes, with 1-way ANOVA showing significant group differences ($F_{2,1448}$=8.74 - 12.21; *P*<.001; η^2=0.07 - 0.10).

Conclusions: AI-assisted dynamic postural control screening combined with adaptive training substantially improved functional performance in mobility, balance, and gait among older adults. While fall incidence was not directly measured, these functional gains may have implications for fall risk reduction. Future longitudinal studies with extended follow-up (12 - 24 mo) and prospective fall incidence tracking across diverse populations are required to validate whether these improvements translate into actual reductions in fall risk.

(JMIR Aging 2025;8:e73290) doi:10.2196/73290

KEYWORDS

artificial Intelligence; fall risk assessment; dynamic postural control; older adult mobility; older adult fall prevention

Introduction

Background on Fall Risks in the Older Adult Population

Falls are one of the leading causes of disability, hospitalization, and mortality among older adults [1,2]. Early detection of functional decline is critical for fall prevention; however, conventional assessments such as the Short Physical Performance Battery (SPPB) rely heavily on manual observation, which is time-consuming and subject to interrater variability [3,4]. With the rapid advancement of digital health technologies, artificial intelligence (AI) has emerged as a promising tool for objective, real-time mobility screening [5-7].

Recent studies have demonstrated that AI-based motion analysis can improve the accuracy and scalability of functional assessments in older adult populations [3,6,8-10]. Unlike static

XSL•FO
RenderX

paper-based tests, AI can dynamically track skeletal movement, quantify gait and balance parameters, and provide continuous feedback. These innovations enable early detection of mobility decline and offer opportunities for personalized interventions [5,7,9-11]. Despite these advances, limited large-scale evidence exists regarding whether AI-assisted screening, when combined with adaptive training, leads to measurable improvements in functional mobility outcomes among older adults.

Recent *JMIR* studies have also highlighted the feasibility of integrating SPPB with broader fall risk assessment tools [12], the application of explainable AI for fall prediction [13], and the potential of digital and telemedicine-based prevention programs [14,15]. Moreover, wearable technology is increasingly considered a viable option for health care monitoring [16].

Research Objectives and Contributions

This study investigates the effectiveness of AI-assisted dynamic postural control screening combined with adaptive training in improving mobility outcomes among older adults. The research makes three key contributions:

1. Demonstrating the feasibility of integrating AI skeletal tracking with the standardized SPPB to objectively assess gait speed, balance, and sit-to-stand performance.
2. Evaluating the short-term functional improvements attributable to AI-generated personalized training prescriptions.
3. Providing large-scale evidence on the relationship between participation rate and functional gains in an older adult population.

Research Question

This study addressed the following research question: can AI-assisted dynamic postural control screening combined with adaptive training significantly improve functional mobility outcomes (SPPB score, gait speed, and sit-to-stand performance) compared with standard care in older adult populations?

Hypotheses

This study had the following research hypotheses:

1. Participants in the experimental group will demonstrate significantly greater improvements in SPPB scores compared with the control group.
2. Participants in the experimental group will show significantly faster improvements in gait speed and sit-to-stand time compared with the control group.

While prior studies have demonstrated that the SPPB is a reliable predictor of functional decline [12,13], its traditional administration is limited by subjectivity and reliance on manual timing or scoring [14]. AI-assisted motion analysis offers a scalable solution by automating the assessment process, increasing measurement precision, and enabling real-time feedback [15,16].

This study builds upon existing evidence by integrating AI-based skeletal tracking with SPPB assessments and extending the application to a large older adult cohort. The primary aim is not only to demonstrate feasibility but also to evaluate whether AI-assisted screening, when combined with adaptive training prescriptions, leads to measurable improvements in mobility outcomes. Imran et al [10] and Santos et al [11] specifically support the dual rationale of integrating AI into both functional assessment and individualized prescription, underscoring the novelty of this approach.

By conducting a quasi-experimental study with more than 2000 participants, this research provides large-scale evidence on the effectiveness of AI-assisted interventions. Unlike prior work that focused mainly on prediction or assessment, our approach tested whether real-time AI-driven prescriptions translated into tangible improvements in balance, gait, and sit-to-stand performance.

Methods

Study Design and Participants

This quasi-experimental study was conducted in Keelung, Taiwan, over a 6-month period. A total of 2005 older adults (≥65 years) were recruited through stratified sampling from 12 community centers and 9 health care institutions.

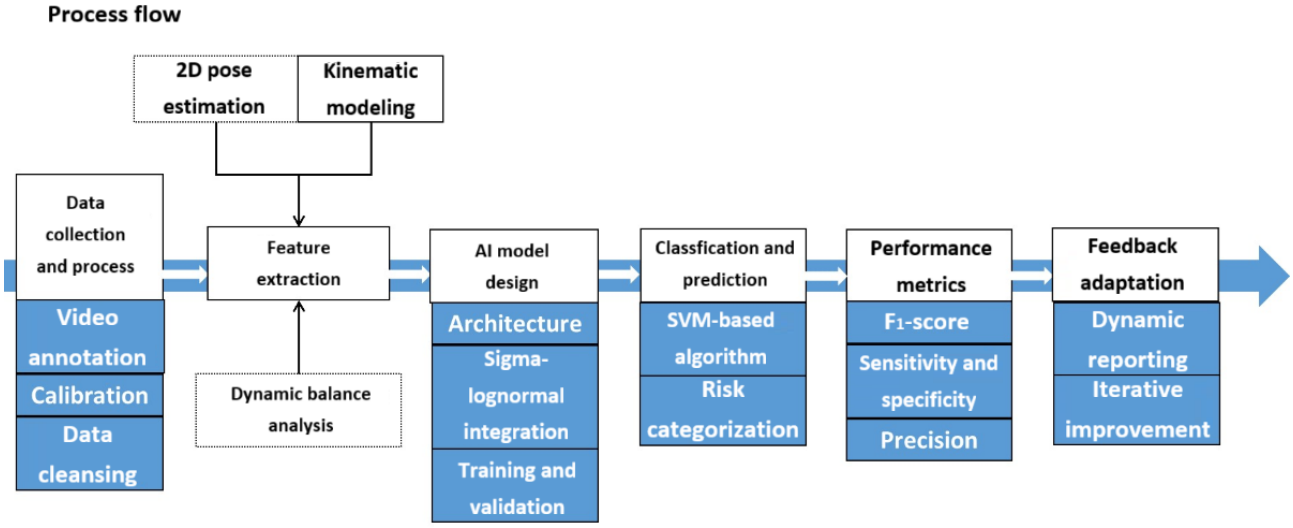
The experimental group (n=1451) was recruited primarily from community centers. Participants underwent AI-assisted SPPB assessments and received individualized exercise prescriptions dynamically adapted by the AI system. Prescriptions were adjusted at subsequent assessments according to participants' performance.

The control group (n=554) was recruited from health care institutions. Participants underwent AI-assisted assessments and continued with standard care (ie, routine health advice and monitoring by their institutions) but did not receive AI-tailored prescriptions. This ensured ethical parity while isolating the effect of AI-enabled personalization.

Importantly, baseline differences reflected recruitment sources: the experimental group had a mean SPPB score of 11.2 (SD 1.8), whereas the control group had a mean score of 10.4 (SD 1.8). These differences were expected, given the healthier status of community-dwelling participants compared with clinical recruits. Independent-sample tests confirmed that baseline differences did not bias the interpretation of intervention outcomes. A summary of participant recruitment, group allocation, and follow-up assessments is presented in Figure 1.

XSL•FO
RenderX

Figure 1. Flow diagram of participant recruitment, group allocation, intervention procedures, and follow-up assessments. AI: artificial intelligence; SVM: support vector machine.



Sample Size and Power Analysis

Although not prespecified, a post hoc power analysis using G*Power (Heinrich Heine University Düsseldorf) indicated that a minimum of 788 participants per group would be required to detect a medium effect size (Cohen *d*=0.4) with 90% power at α=.05. Our actual sample sizes (1451 and 554) exceeded this requirement, ensuring sufficient statistical power and confirming that the large cohort was methodologically and ethically justified.

Assessment Tools and AI Integration

Functional mobility was assessed using the SPPB, which is widely validated as a predictor of functional decline and fall risk [4,11-13]. Wearable and sensor-based technologies further demonstrate feasibility for continuous monitoring and objective evaluation [16]. The SPPB included (1) balance tests, (2) 4-meter gait speed, and (3) the 5-times sit-to-stand test. Assessments were conducted with AI-assisted skeletal tracking to improve accuracy and objectivity.

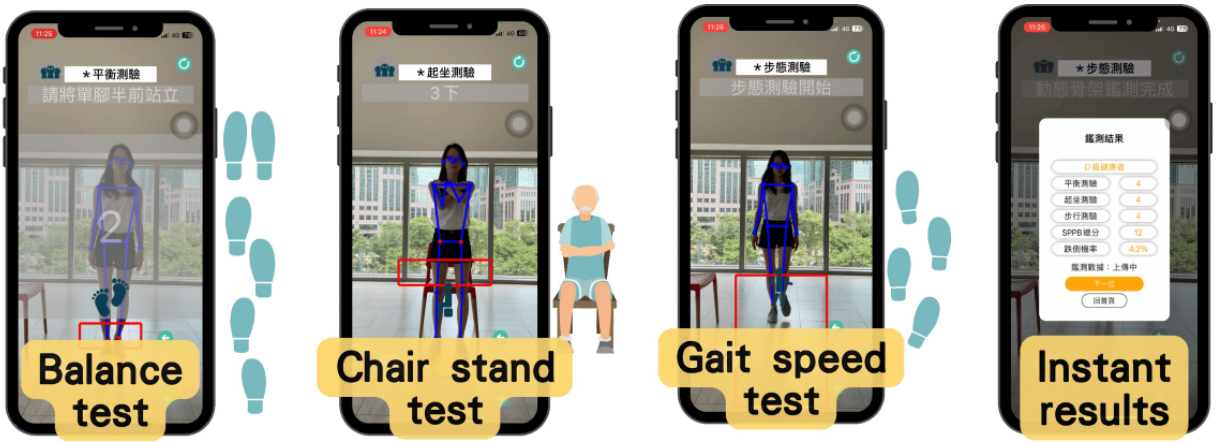
The AI system replaced traditional manual scoring by extracting quantitative motion features in real time. This consolidated statement replaces prior repeated descriptions, improving readability while preserving scientific accuracy. Key features included the following:

- Real-time monitoring: continuous capture of postural stability, gait patterns, and transitional movements.
- Automated data processing: machine learning algorithms compared participant movements against validated mobility benchmarks.
- Immediate feedback: results were processed instantly, producing individualized reports for trainers and health care staff.

System Validation

In a prior pilot study of 320 older adults, the AI algorithm achieved greater than 90% accuracy in classifying functional mobility categories compared with physiotherapist ratings. Test-retest reliability was high (intraclass correlation coefficient >0.88) for gait speed and sit-to-stand measures, supporting the robustness of the system. The architecture and workflow of the AI system are summarized in Figure 2, showing the process from skeletal tracking and feature extraction to automatic scoring and feedback generation.

Figure 2. Architecture and workflow of the artificial intelligence–assisted skeletal tracking system, including video capture, landmark extraction, feature quantification, and automatic scoring with feedback.



XSL•FO
RenderX

Ethical Considerations

This study was approved by the Institutional Review Board of En Chu Kong Hospital (ECKIRB1071204). All participants were informed about the study objectives and procedures and provided written informed consent before enrollment. No financial incentives or compensation were provided for participation. All data were anonymized, and confidentiality and data security measures were strictly implemented to protect participant privacy in accordance with institutional and national ethical guidelines.

AI-Enabled Training Adaptation

In the experimental group, the AI system generated individualized exercise prescriptions that were adapted at subsequent assessments. Algorithms adjusted the type, intensity, and progression of exercises based on gait speed thresholds, sit-to-stand performance, and balance stability metrics.

This ensured tailored interventions beyond generic training exposure. The control group used the AI system solely for assessment, without adaptive prescriptions. The adaptive prescription framework is presented in Figure 3, which illustrates how SPPB-derived performance metrics (ie, balance, gait, and sit-to-stand performance) informed AI-driven adjustments in exercise type, intensity, and progression.

Figure 3. Framework of artificial intelligence–enabled adaptive training prescriptions, showing how balance, gait speed, and sit-to-stand performance metrics informed real-time adjustments to exercise type, intensity, and progression.



Data Collection Procedures

Participants in both groups were assessed at baseline and 6-month follow-up. Monthly interim assessments were conducted to monitor progress. All assessments were video recorded through the AI skeletal tracking system and processed automatically.

Data collected included balance scores from SPPB balance tasks; gait speed (cm/s) from the 4-meter walk; sit-to-stand time (s) from the 5-repetition test; and demographics such as age, sex, and recruitment source.

AI-generated feedback reports were shared with trainers to guide intervention adjustments in the experimental group.

Statistical Analysis

All statistical analyses were performed using SPSS (version 27.0; IBM Corp). Descriptive statistics (means and SDs) were calculated for all measures. Between-group comparisons were conducted using independent-samples *t* tests to compare improvements in SPPB, gait speed, and sit-to-stand performance between the experimental and control groups. Within-group changes were examined using repeated-measures ANOVA to examine changes from baseline to follow-up within each group. Participation-level analysis was performed using one-way ANOVA to compare outcomes among subgroups with high (>75%), moderate (50% - 75%), and low (<50%) participation rates, to explore the relationship between adherence and functional outcomes. Effect sizes were calculated using Cohen *d* for pairwise comparisons, and η^2 was reported for ANOVA models; 95% CIs were reported for all effect size estimates. Statistical significance was set at a 2-tailed α=.05.

Results

Baseline Characteristics

A total of 2005 participants were analyzed, including 1451 in the experimental group and 554 in the control group. The experimental group had a mean age of 73.4 (SD 6.1) years, with 841 (57.9%) female participants. The control group had a mean age of 74.1 (SD 6.4) years, with 338 (61.0%) female participants.

Baseline differences in walking speed were noted. As shown in Table 1, the experimental group demonstrated a mean baseline gait speed of 117 cm/s, compared with 105 cm/s in the control group. These values align with cut-off thresholds for functional decline in mobility. Baseline functional measures showed expected differences due to recruitment sources (community vs health care institutions):

- Experimental group: SPPB=11.2, gait speed=117 cm/s, and sit-to-stand=10.5 seconds.
- Control group: SPPB=10.4, gait speed=105 cm/s, and sit-to-stand=11.2 seconds.

An SPPB score of ≤9 was considered indicative of high fall risk, while scores ≥10 represented normal to moderate function, consistent with prior validation studies. Independent-samples tests confirmed that although baseline functional status differed, these differences did not bias the interpretation of intervention outcomes.

XSL•FO
RenderX

Table . Comparison of artificial intelligence–assisted and Standard Short Physical Performance Battery (SPPB) Results.

Metric	Experimental group (baseline), mean (SD 1.8)	Experimental group (end), mean (SD 1.6)	Control group (baseline), mean (SD 1.8)	Control group (end), mean (SD 1.7)	Improvement difference	*P* value	Cohen *d*
SPPB score	11.2	12.0	10.4	10.7	0.5	.008	0.45
Gait speed (cm/s)	117	132	105	110	10	<.001	0.62
Sit-to-stand (s)	10.5	9.1	11.2	10.6	–0.8	.006	0.39

Intervention Outcomes

Overview

The improvements in SPPB, gait speed, and sit-to-stand performance observed in this study are consistent with findings from prior comparative validation studies using SPPB and Grupo de Desenvolvimento Latino-Americano para a Maturidade batteries [6,17]. Our AI-assisted skeletal tracking showed high accuracy and reliability, similar to other automated approaches such as camera-based SPPB evaluation [18] and sensor insole–based gait analysis [19].

SPPB Scores

The experimental group's mean SPPB score increased from 11.2 to 12.0 (Δ=0.8), while the control group's mean SPPB score increased from 10.4 to 10.7 (Δ=0.3). The improvement difference was 0.5 points (t_{2003}=3.41; *P*=.001; Cohen *d*=0.45, 95% CI 0.18-0.72).

Gait Speed

The experimental group's gait speed improved from 117 cm/s to 132 cm/s (Δ=15 cm/s), compared with the control group's improvement from 105 cm/s to 110 cm/s (Δ=5 cm/s). The improvement difference was 10 cm/s (t_{2003}=4.85; *P*<.001; Cohen *d*=0.62, 95% CI 0.35-0.88).

Sit-to-Stand Performance

The experimental group reduced their sit-to-stand time from 10.5 s to 9.1 s (Δ=–1.4 s), compared with the control group's reduction from 11.2 s to 10.6 s (Δ=–0.6 s). The improvement difference was –0.8 s (t_{2003}=3.12; *P*=.002; Cohen *d*=0.39, 95% CI 0.12-0.65).

Participation-Level Analysis

Participation rates strongly influenced outcomes [17]. The differences across participation levels are summarized in Table 2. The experimental group was divided into high (>75%), moderate (50% - 75%), and low (<50%) attendance subgroups.

One-way ANOVA confirmed significant differences across participation levels for all three functional outcomes. As shown in Table 2, higher participation was associated with greater improvements in SPPB scores, gait speed, and sit-to-stand performance.

Table . Summary of participation impact with statistical tests.

Metric	High participation (>75%; n=482), n (%)	Moderate participation (50% - 75%; n=515), n (%)	Low participation (<50%; n=454), n (%)	*F* test (*df*)	*P* value	η^2
SPPB[a] score (Δ)	1.2 (10.0)	0.8 (7.0)	0.3 (2.0)	8.74 (2, 1448)	<.001	0.07
Gait speed (Δ, cm/s)	20 (17.0)	12 (10.0)	5 (4.0)	12.21 (2, 1448)	<.001	0.10
Sit-to-stand (Δ, s)	–1.8 (15.0)	–1.0 (9.0)	–0.4 (4.0)	9.63 (2, 1448)	<.001	0.08

[a]SPPB: Short Physical Performance Battery.

[b]Values in parentheses indicate percentage change relative to each subgroup's baseline mean

Discussion

Summary of Key Findings

This study demonstrated that AI-assisted dynamic postural control screening combined with adaptive training significantly improved functional performance among older adults over a 6-month period. Improvements were observed in overall SPPB scores, gait speed, and sit-to-stand performance, with effect sizes ranging from moderate to large (Cohen *d*=0.39 - 0.62). Furthermore, the participation rate was positively associated with functional gains, indicating a dose-response relationship.

Comparison With Prior Work

Our results align with previous AI- and sensor-based fall risk assessment studies [3,5,9,13,18,19]. Importantly, our findings expand on recent *JMIR* research exploring digital health interventions for fall prevention [14,15]. Classic reviews and epidemiological studies have long established that falls among older adults are preventable through multifactorial interventions [20]. In addition, the screening component itself warrants emphasis. Early and objective detection of deficits in mobility, balance, and gait is a prerequisite for timely intervention, particularly as conventional screening tools such as the SPPB are subject to variability and may not scale efficiently in community or clinical settings [4,12]. By embedding AI-based

XSL•FO
RenderX

screening in our design, we demonstrated not only intervention efficacy but also the feasibility of scalable, objective, and standardized risk detection, which justifies the inclusion of the screening part of the study. This represents a shift from using AI solely for risk identification toward using it as a tool for delivering individualized, dynamic interventions. While conventional screening tools such as the SPPB are validated, they are limited by interrater variability and scalability issues. Few existing AI interventions have incorporated both real-time skeletal tracking and adaptive exercise prescriptions, making our approach distinct from most prior AI studies. In contrast, our study demonstrates that AI-assisted screening combined with adaptive prescriptions can not only identify deficits but also drive functional improvements in a large older adult cohort. A structured comparison of prior AI applications and this study is summarized in Table 3, highlighting how our work advances beyond assessment toward intervention.

XSL•FO
RenderX

Table . Comparison of studies using artificial intelligence (AI) and Short Physical Performance Battery (SPPB) to assess physical function in older adults.

Study	Title	Research focus	Methods used	Key findings	Contribution to the field
Friedrich et al [3]	A Deep Learning Approach for TUG[a] and SPPB Score Prediction of (Pre-) Frail Older Adults on Real-Life IMU Data	Predicting TUG and SPPB scores in prefrail and frail older adults using real-life IMU[b] data	• IMU sensor on the right hip • SPPB and TUG tests • LSTM[c]+CNN[d] models • Train-validation-test split	• TUG prediction accuracy: 95.9% • SPPB prediction accuracy: 94.3%	• Demonstrates feasibility of predicting physical performance in older adults using real-life IMU data • Supports early detection and intervention for physical decline in the older adults
González-Castro et al [5]	The Applications of Artificial Intelligence for Assessing Fall Risk: Systematic Review	Analyzing AI applications in assessing postural control and fall risk	• Literature review of 22 studies • Accelerometers, video, EHR[e], and big data • Various ML[f] algorithms (SVM[g], RF[h], LSTM, etc)	• AI model accuracy: >70% • Gait speed and posture are important predictors • Accelerometers are the most common data source	• Confirms AI as a useful tool for fall risk screening • Future focus should be on integrating and validating AI models in clinical practice
Kouno et al [7]	Introduction of AI Technology for Objective Physical Function Assessment	Reviewing current research on AI-based objective physical function assessment	• Literature review of 26 studies • Video, sensors, EHR, and questionnaires • ML regression or classification (SVM, RF, LSTM, etc)	• Sensors are the main input for gait speed prediction • Video is increasingly used for TUG and gait analysis	• Validates AI for predicting SPPB, TUG, gait speed, grip strength, and other physical function indices • Recommends incorporating AI-assisted physical function assessment into future clinical routines
Kraus et al [19]	Prediction of Physical Frailty in Orthogeriatric Patients Using Sensor Insole–Based Gait Analysis and Machine Learning Algorithms: Cross-sectional Study	Comparing sensor insoles, questionnaires, and conventional assessments in predicting physical frailty in older patients	• 57 older adult outpatients • Sensor insoles • SARC-F[i] questionnaire • SPPB and TUG tests • Recursive feature elimination • KNN[j] and RF	• Sensor insoles+ML models outperform SARC-F and TUG alone • Gait parameters are the most predictive features	• Supports integrating sensors and ML into fall risk screening • Recommends large-scale validation of this approach
Welch et al [4]	The Short Physical Performance Battery (SPPB): A Quick and Useful Tool for Fall Risk Stratification Among Older Primary Care Patients	Evaluating SPPB as a fall risk stratification tool in older adults and exploring its integration with the STEADI[k] tool	• 417 community-dwelling older adults • SPPB • Fall history • Balance confidence • Multivariable negative binomial regression models	• SPPB total and component scores predict 1-year fall risk • In the STEADI-positive group, low SPPB scores are associated with higher fall rates	• Supports SPPB as a fall risk stratification tool for older adults • SPPB shows promise in complementing STEADI screening but requires large-scale study validation

XSL•FO
RenderX

Study	Title	Research focus	Methods used	Key findings	Contribution to the field
Marcos-Pardo et al [6]	GDLAM and SPPB batteries for screening sarcopenia in community-dwelling Spanish older adults: Healthy-age network study	Comparing the diagnostic ability of GDLAM[l] and SPPB batteries to classify sarcopenia	• 382 community-dwelling older adults • GDLAM battery • SPPB • ROC[m] curve analysis	• GDLAM has moderate to high sensitivity (60.1% - 72.2%) and moderate to very high specificity (57.6% - 90.7%) • SPPB lacks ability to classify sarcopenia	• GDLAM battery can screen for sarcopenia in community-dwelling older adults • GDLAM is more reliable and valid than SPPB for sarcopenia assessment
Duncan et al [18]	Automated Camera-Based Assessment of Short Physical Performance Battery (SPPB) for Older Adults with Cancer	Introducing a camera system for automated SPPB evaluation in older patients with cancer	• Motorized camera system • SPPB tests • Template matching, Haar cascades, and CSRT[n] tracking • In vivo testing with 8 volunteers	• Gait speed accuracy: >95% • Balance and 5TSS[o] time accuracy: >97%	• Automates SPPB evaluation for older patients with cancer • Reliable, efficient, and overcomes limitations of manual assessment
Santos et al [11]	Short Physical Performance Battery (SPPB) score as a discriminator of dynapenic abdominal obesity among community-dwelling older adults	Examining SPPB score as a discriminator of dynapenic abdominal obesity in older adults	• 382 community-dwelling older adults • SPPB • Hand grip strength • Waist circumference • ROC curve analysis	• SPPB scores inversely associated with DAO[p] • SPPB cut-off of ≤9 discriminates DAO	• SPPB can screen for DAO in older adults • SPPB is a useful tool for early DAO detection and management

[a]TUG: Timed Up and Go.

[b]IMU: inertial measurement unit.

[c]LSTM: long short-term memory.

[d]CNN: convolutional neural network.

[e]EHR: electronic health record.

[f]ML: machine learning.

[g]SVM: support vector machine.

[h]RF: random forest.

[i]SARC-F: Simple Questionnaire to Rapidly Diagnose Sarcopenia.

[j]KNN: k-nearest neighbor.

[k]STEADI: Stopping Elderly Accidents, Deaths, and Injuries.

[l]GDLAM: Grupo de Desenvolvimento Latino-Americano para a Maturidade.

[m]ROC: receiver operating characteristic.

[n]CSRT: channel and spatial reliability tracking.

[o]5TSS: 5 times sit-to-stand.

[p]DAO: dynapenic abdominal obesity.

Practical Implications

The findings suggest that AI-assisted systems have potential for broad application in community and clinical settings. By automating data capture and analysis, the AI platform provides objective, real-time, and scalable assessments that reduce reliance on subjective manual scoring. The adaptive prescription framework further allows for individualized exercise programs that can be implemented in day care centers, senior fitness programs, and long-term care facilities. Such systems may enhance efficiency for health care staff while promoting functional independence among older adults.

Limitations

Several limitations should be noted. First, fall incidence was not directly measured; therefore, conclusions should be interpreted as functional performance improvements rather than confirmed reductions in fall events [21]. Second, the recruitment of the experimental group from community centers and the control group from health care institutions introduced baseline

XSL•FO
RenderX

differences in functional status. Although statistical analyses confirmed that these differences did not bias the outcomes, they may limit comparability. Third, the study was conducted in Keelung, Taiwan, which may limit generalizability to populations with different cultural, socioeconomic, or rural backgrounds. Fourth, the follow-up period of 6 months reflects short-term to midterm outcomes; longer-term effects remain unknown.

Future Research

Future studies should extend follow-up to 12 to 24 months to examine the sustainability of functional gains and to prospectively track fall incidence [22]. Multisite validation across diverse populations, including rural communities, lower-income groups, and culturally varied settings, is required to establish generalizability. Further work should also investigate adherence, engagement, and the accessibility of AI technologies, particularly among populations considered vulnerable who may face barriers to digital health adoption.

Conclusions

This study demonstrated that AI-assisted dynamic postural control screening, when combined with adaptive training, significantly improved functional mobility among older adults. Improvements were evident in SPPB scores, gait speed, and sit-to-stand performance, with effect sizes in the moderate-to-large range. These results, based on a large sample of more than 2000 participants, provide strong support for the feasibility of implementing AI-assisted programs at scale. Importantly, participation levels showed a dose-response relationship with outcomes, highlighting the role of consistent engagement in maximizing functional benefits.

While the findings are encouraging, several considerations remain. Fall incidence was not directly measured; therefore, conclusions should be interpreted as functional improvements rather than confirmed reductions in fall risk. The 6-month study duration limits insights into long-term sustainability, and the recruitment of participants primarily from community centers and health care institutions in Keelung, Taiwan, may constrain generalizability. Furthermore, the accessibility and usability of AI-based systems for older adults in rural or low-resource settings require further examination.

In summary, this study contributes large-scale evidence that AI-assisted dynamic postural control screening with adaptive training can enhance functional performance in older adults. Future research should extend longitudinal follow-up to 12 to 24 months, prospectively measure fall events, and conduct multisite trials across diverse cultural and socioeconomic contexts. Addressing adherence, technology accessibility, and digital literacy will be essential to ensure equitable adoption and long-term impact of AI-assisted older adult care solutions.

Data Availability

The datasets generated and analyzed during this study are not publicly available due to ethical and institutional restrictions imposed by the Institutional Review Board of En Chu Kong Hospital. No data can be shared outside the approved research protocol.

Conflicts of Interest

None declared.

References

1. Step safely: strategies for preventing and managing falls across the life-course. : World Health Organization; 2021 URL: https://www.who.int/publications/i/item/978924002191-4 [accessed 2025-11-19]
2. Falls. World Health Organization. 2021 Apr 26. URL: https://www.who.int/news-room/fact-sheets/detail/falls [accessed 2025-08-26]
3. Friedrich B, Lau S, Elgert L, Bauer JM, Hein A. A deep learning approach for TUG and SPPB score prediction of (pre-) frail older adults on real-life IMU data. Healthcare (Basel) 2021 Feb 2;9(2):149. [doi: 10.3390/healthcare9020149] [Medline: 33540555]
4. Welch SA, Ward RE, Beauchamp MK, Leveille SG, Travison T, Bean JF. The Short Physical Performance Battery (SPPB): a quick and useful tool for fall risk stratification among older primary care patients. J Am Med Dir Assoc 2021 Aug;22(8):1646-1651. [doi: 10.1016/j.jamda.2020.09.038] [Medline: 33191134]
5. González-Castro A, Leirós-Rodríguez R, Prada-García C, Benítez-Andrades JA. The applications of artificial intelligence for assessing fall risk: systematic review. J Med Internet Res 2024 Apr 29;26:e54934. [doi: 10.2196/54934] [Medline: 38684088]
6. Marcos-Pardo PJ, González-Gálvez N, Carbonell-Baeza A, Jiménez-Pavón D, Vaquero-Cristóbal R. GDLAM and SPPB batteries for screening sarcopenia in community-dwelling Spanish older adults: healthy-age network study. Exp Gerontol 2023 Feb;172:112044. [doi: 10.1016/j.exger.2022.112044] [Medline: 36509296]
7. Kouno N, Takahashi S, Komatsu M, et al. Introduction of AI technology for objective physical function assessment. Bioengineering (Basel) 2024 Nov 16;11(11):1154. [doi: 10.3390/bioengineering11111154] [Medline: 39593814]
8. Lin KC, Wai RJ. A feasible fall evaluation system via artificial intelligence gesture detection of gait and balance for sub-healthy community- dwelling older adults in Taiwan. IEEE Access 2021;9:146404-146413. [doi: 10.1109/ACCESS.2021.3123297]

XSL•FO
RenderX

9. Yang H, Chang J, He W, Wee CF, Yit JST, Feng M. Frailty modeling using machine learning methodologies: a systematic review with discussions on outstanding questions. IEEE J Biomed Health Inform 2025 Jan;29(1):631-642. [doi: 10.1109/JBHI.2024.3430226] [Medline: 39024091]
10. Imran R, Khan SS. A systematic review on the efficacy of artificial intelligence in geriatric healthcare: a critical analysis of current literature. BMC Geriatr 2025 Apr 11;25(1):248. [doi: 10.1186/s12877-025-05878-w] [Medline: 40217136]
11. Santos EP, Silva CF, Ohara DG, Matos AP, Pinto AC, Pegorari MS. Short Physical Performance Battery (SPPB) score as a discriminator of dynapenic abdominal obesity among community-dwelling older adults. Geriatr Nurs 2021;42(2):467-472. [doi: 10.1016/j.gerinurse.2021.02.016] [Medline: 33714904]
12. Tanaka S, Imaizumi T, Morohashi A, et al. In-hospital fall risk prediction by objective measurement of lower extremity function in a high-risk population. J Am Med Dir Assoc 2023 Dec;24(12):1861-1867. [doi: 10.1016/j.jamda.2023.07.020] [Medline: 37633314]
13. Liang HW, Ameri R, Band S, et al. Fall risk classification with posturographic parameters in community-dwelling older adults: a machine learning and explainable artificial intelligence approach. J NeuroEngineering Rehabil 2024;21(1):21. [doi: 10.1186/s12984-024-01310-3]
14. Rein DB, Hackney ME, Haddad YK, et al. Telemedicine-based risk program to prevent falls among older adults: protocol for a randomized quality improvement trial. JMIR Res Protoc 2024 Mar 26;13:e54395. [doi: 10.2196/54395] [Medline: 38346180]
15. Jacobson CL, Foster LC, Arul H, Rees A, Stafford RS. A digital health fall prevention program for older adults: feasibility study. JMIR Form Res 2021 Dec 23;5(12):e30558. [doi: 10.2196/30558] [Medline: 34837492]
16. Laurens MC, Pieterse ME, Brusse-Keizer M, et al. Alcohol avoidance training as a mobile app for problem drinkers: longitudinal feasibility study. JMIR Mhealth Uhealth 2020 Apr 14;8(4):e16217. [doi: 10.2196/16217] [Medline: 32286235]
17. Muhammad T, Saravanakumar P, Sharma A, Srivastava S, Irshad CV. Association of food insecurity with physical frailty among older adults: study based on LASI, 2017-18. Arch Gerontol Geriatr 2022;103:104762. [doi: 10.1016/j.archger.2022.104762] [Medline: 35841798]
18. Duncan L, Zhu S, Pergolotti M, et al. Automated camera-based assessment of Short Physical Performance Battery (SPPB) for older adults with cancer. In: Domínguez-Morales M, Luna-Perejón F, editors. Human Gait - Recent Findings and Research: IntechOpen; 2024. [doi: 10.5772/intechopen.112899]
19. Kraus M, Saller MM, Baumbach SF, et al. Prediction of physical frailty in orthogeriatric patients using sensor insole-based gait analysis and machine learning algorithms: cross-sectional study. JMIR Med Inform 2022 Jan 5;10(1):e32724. [doi: 10.2196/32724] [Medline: 34989684]
20. Rubenstein LZ, Josephson KR. Falls and their prevention in elderly people: what does the evidence show? Med Clin North Am 2006 Sep;90(5):807-824. [doi: 10.1016/j.mcna.2006.05.013] [Medline: 16962843]
21. Bischoff-Ferrari HA, Dawson-Hughes B, Willett WC, et al. Effect of Vitamin D on falls: a meta-analysis. JAMA 2004 Apr 28;291(16):1999-2006. [doi: 10.1001/jama.291.16.1999] [Medline: 15113819]
22. Verghese J, Lipton RB, Hall CB, Kuslansky G, Katz MJ, Buschke H. Abnormality of gait as a predictor of non-Alzheimer's dementia. N Engl J Med 2002 Nov 28;347(22):1761-1768. [doi: 10.1056/NEJMoa020441] [Medline: 12456852]

Abbreviations

AI: artificial intelligence
SPPB: Short Physical Performance Battery

Edited by M Mardini; submitted 01.03.25; peer-reviewed by J Zhang, R Boumans, SJ Diwan, WK Lim; revised version received 07.09.25; accepted 10.11.25; published 28.11.25.

Please cite as:
Lin KC, Wai RJ, Chang Chien HY
AI-Assisted Dynamic Postural Control Screening to Improve Functional Mobility in Older Adult Populations: Quasi-Experimental Study
JMIR Aging 2025;8:e73290
URL: https://aging.jmir.org/2025/1/e73290
doi:10.2196/73290

© Kai-Chih Lin, Rong-Jong Wai, Hung-Yu Chang Chien. Originally published in JMIR Aging (https://aging.jmir.org), 28.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium,

XSL•FO
RenderX

provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Building Strong Foundations: Nonrandomized Interventional Study of a Novel, Digitally Delivered Fall Prevention Program for Older Adults

David Wing[1,2], MS; Jeanne F Nichols[1,2], PhD; Hava Shoshana Barkai[1,2], MS; Olivia Culbert[1,2], MS; Daniel Moreno[1,2], MS; Michael Higgins[1,2], MS; Anna O'Brien[1,2], BS; Mariana Perez[2], BS; Hope Davey[2], BS; Ryan Moran[2,3], MPH, MD

[1]Herbert Wertheim School of Public Health and Human Longevity, University of California, San Diego, San Diego, CA, United States

[2]Exercise and Physical Activity Resource Center, University of California, San Diego, San Diego, CA, United States

[3]School of Medicine, University of California, San Diego, San Diego, CA, United States

Corresponding Author:
David Wing, MS
Exercise and Physical Activity Resource Center
University of California, San Diego
9500 Gilman Drive #0811
CALIT2/Atkinson Hall Room 3504
San Diego, CA, 92093
United States
Phone: 1 858 534 9315
Fax: 1 858 534 9404
Email: dwing@eng.ucsd.edu

Abstract

Background: Injuries from falls are a major concern among older adults. Targeted exercise has been shown to improve fall risk, and recommendations for identifying and referring older adults for exercise-based interventions exist. However, even when very inexpensive or free, many do not use available fall prevention programs, citing barriers related to convenience and safety. These issues are even greater among older adults residing in rural areas where facilities are less abundant. These realities highlight the need for different approaches to reducing falls in novel ways that increase reach and are safe and effective. Web-based delivery of exercise interventions offers some exciting and enticing prospects.

Objective: Our objective was to assess the efficacy of the Strong Foundations exercise program to change markers of physical function, posture, balance, strength, and fall risk.

Methods: Strong Foundations is a once weekly (60 minutes), 12-week iterative program with 3 core components: postural alignment and control, balance and mobility, and muscular strength and power. We used a quasi-experimental design to determine changes in physical function specific to balance, postural control, and muscular strength among older adults at low or moderate risk of falling.

Results: A total of 55 low-risk and 37 moderate-risk participants were recruited. Participants significantly improved on the 30-second Chair Stand (mean change of 1, SD 3.3 repetitions; P=.006) and Timed Up and Go (mean change of 0.2, SD 0.7 seconds; P=.004), with the moderate-risk group generally improving to a greater degree than the low-risk group. Additionally, Short Physical Performance Battery performance improved significantly in the moderate-risk category (P=.02). The majority of postural measures showed statistically significant improvement for both groups (P<.05). Measures of “relaxed” posture showed improvements between 6% and 27%. When an “as tall as possible” posture was adopted, improvements were ~36%.

Conclusions: In this 12-week, iterative, web-based program, we found older adults experienced improvement not only in measures used in clinical contexts, such as the 30-second Chair Stand and Timed Up and Go, but also contextualized gains by providing deeper phenotypical measurement related to posture, strength, and balance. Further, many of the physical improvements were attenuated by baseline fall risk level, with those with the highest level of risk having the greater gains, and, thus, the most benefit from such interventions.

(JMIR Aging 2025;8:e68957) doi:10.2196/68957

XSL•FO
RenderX

KEYWORDS

exercise; older adults; digital intervention; Zoom; balance; posture; strength; fall prevention

Introduction

The absolute number of adults older than 65 years of age has been steadily growing for decades and now outnumbers the number of children younger than 18 years of age [1]. In addition, estimates indicate that older adults will represent greater than 23% of the population in the United States by 2035 [1]. Injuries resulting from falls are the largest cause of accidental death and mobility-related disability among older adults. This is likely because nearly one-quarter of community-residing adults older than 65 years of age fall annually, and that number rises to almost half of those older than 80 years of age [2]. Particularly concerning is the fact that once an individual starts to fall, there are often additional falls. Indeed, evidence from numerous sources suggests that between 10% and 44% of older adult patients who have fallen will sustain additional falls in the following year [3-7].

Targeted strength and balance exercises have consistently been shown to improve fall risk, and accordingly, recommendations for identifying and referring older adults for exercise-based interventions have been developed. Public health and clinically oriented authorities, including the Centers for Disease Control and Prevention (CDC) and the US Preventive Services Task Force, have acknowledged the importance of risk assessment and physical activity (PA) in improving fall risk for older adults. The CDC's Stopping Elderly Accidents and Deaths Initiative (STEADI) has also established an evidenced-based questionnaire that assesses known risk factors to help clinicians stratify individuals by risk category to identify those who should be best served by fall prevention programs [8-10].

Many exercise programs for older adults exist. While some health insurance may include access to programs such as Silver Sneakers and EnhanceFitness PA at no additional charge [11], many older adults may have to pay substantial fees to participate. Further, even when very inexpensive or free, many eligible older adults do not use these programs, citing several barriers [11,12]. These include personal preferences (dislike of gymnasiums and lack of knowledge regarding appropriate activity), environmental factors (difficulty in reaching gymnasiums or classes, concerns about bad weather and driving, concerns about being "old" in a gymnasium environment), and structural factors (limited number of facilities and concerns about instructor expertise with older adults) [12]. These issues are even greater among older adults residing in rural areas. Rural residents have fewer facilities for guided PA, more limited transportation infrastructure [13], and less exercise expertise than their urban counterparts [14]. Confounding these difficulties, ongoing concerns exist surrounding the COVID-19 transmissibility alongside the availability, and desirability, of in-person programs. Since the height of the pandemic, there has been a growth in available options for in-person exercise programs. However, while vaccines and exposure have decreased the likelihood of serious infection, there remains some risk associated with (large) in-person gatherings, particularly among the more vulnerable older adult population.

These realities highlight the need for different approaches in addressing the public health challenge of reducing the likelihood of falls in older adults in novel ways that increase reach and are safe and effective. Digital delivery of exercise interventions offers some exciting and enticing prospects, as not only can this platform eliminate barriers related to difficulty with transportation and concerns about the gymnasium environment, but it can also bring highly qualified experts to individuals who are the most in need. These include those experiencing specific conditions that require special instruction and those living in rural and geographically remote areas. Given that older adults are increasingly confident and competent using the internet and videoconferencing tools [15,16], developing interventions that leverage this technology to improve access to fall-risk reduction exercise programs, including balance, posture, and strength training, is both worthwhile and feasible.

Despite these possibilities, there remains limited research on digital group–based exercise programs, particularly those focused on fall prevention. Further, the available evidence is generally drawn from relatively small samples [17-19], or the studies have been conducted in populations with specific medical conditions or comorbidities such as diabetes [20,21] or mild cognitive impairment [22]. To the authors' knowledge, there are only 2 randomized studies with a sample size greater than 20 that have focused on healthy older adults. Although the exercise modalities differed, following the intervention, participants did have generally positive findings regarding improvements in strength and balance [23,24]. However, neither intervention included exercises specific to postural control, a factor strongly hypothesized to influence fall risk [25,26]. Moreover, instruction during exercise sessions in nearly all studies we reviewed, regardless of the health status of the participants, was generalized to the entire class due to the difficulty of the instructor speaking to any 1 person to correct technique without interrupting the flow of the class.

Therefore, we developed and digitally deployed an evidence-based exercise program, Strong Foundations (SF), with instruction provided in real time, to enhance physical function and ultimately prevent falls among older adults. The aim of this study was to evaluate the SF program's effectiveness in improving the posture, balance, and strength of older adults using valid, reliable, and widely used measures of physical function.

Methods

Ethical Considerations

This study was approved by the institutional review board (#806696) at the University of California, San Diego, and informed consent was gathered from all participants. They were further informed that they did not have to complete any measurements and were free to withdraw from the exercise sessions at any time. Participants were additionally provided with US $20 for each laboratory visit to cover travel costs and were given resistance bands to be used during the exercise

XSL•FO
RenderX

intervention (value ~US $10). All data were de-identified prior to analyses with a code to participant numbers and identities held only by the PI and study coordinator and maintained in protected file servers.

Study Design and Recruitment

We used a quasi-experimental design to determine the efficacy of digitally delivered exercise for improving physical function specific to balance, postural control, and muscular strength among older adults at low or moderate risk of falling. Several strategies were used to recruit participants: (1) posters displayed at older adult living facilities and community centers; (2) emails sent to existing lists of older adults who previously expressed interest in research studies in general, and in fracture prevention, specifically; and (3) announcements made following educational presentations on fall and fracture prevention delivered digitally or in person to various groups of older adults in the community. Those who expressed interest were contacted by study personnel and screened by phone for eligibility. Inclusion criteria included being 60 years and older, having an internet-enabled device with a screen of 7+ inches (tablet, laptop, or similar), and scoring less than 8 on the CDC's STEADI questionnaire. Participants were further stratified based on their answers to the STEADI questionnaire into low- and moderate-risk groups with a maximal score of 3 and 7, respectively [27].

Intervention Design

Designed by 2 exercise physiologists working in consultation with both a medical doctor and a doctor of physical therapy, the SF program is a once weekly (60 minutes), 12-week iterative program with 3 core components: postural alignment and control, balance and mobility, and muscular strength and power. All instructors leading the training had, at minimum, a bachelor's degree in an associated field (kinesiology, exercise physiology, etc) and licensing as a personal trainer from an accredited national institution (American College of Sports Medicine, American Council on Exercise, and National Academy of Sports Medicine). In addition to the weekly group class, participants received both printed and recorded instruction regarding how to safely complete the exercises without supervision and were "assigned" homework materials and encouraged to practice 2 or more additional times per week preferably with a family member or friend present for safety.

All of the exercises introduced throughout the course were designed to be appropriate for an older adult population and were standardized so that participants received the same basic instruction, but the level of difficulty was scaled (and coached) based on individual capability, experience, and musculoskeletal limitations. Each week, new foundational exercises identified as being important to maintaining strength, postural control, balance, and mobility, and also relevant to specific daily activities (eg, picking up an object from the floor to place it on a countertop) were added to the class. In addition, over time, most exercises initially taught as isolated movements were expanded to more complicated, multijoint and multiplanar compound movements, once again with the intent to mimic daily movement.

Home Setup and Safety Concerns

One to two weeks before the first exercise class, participants were emailed instructions for setting up their exercise space at their residence. This included having, at minimum, a 6×6 ft uncluttered exercise space on either a nonslippery floor or an area with wall-to-wall carpet (no loose rugs) and adequate open space surrounding the designated exercise space. In addition, if possible, they were encouraged to have a small amount of wall space immediately behind their exercise area to allow for intermittent wall exercises. Participants were also provided with a set of resistance bands with the level of difficulty determined based on their self-described strength and were instructed to ensure that they had a chair without wheels (and preferably without arms) available for chair-based exercises and to assist with balance during standing exercise.

Intervention Delivery

One of the novel features of this program is the delivery of semi-individualized instruction in real time within a small group setting. To this end, classes were designed to have 10-15 participants. A single lead instructor provided verbal instruction while demonstrating each exercise. Simultaneously, at least 1, and sometimes 2, additional instructors provided individualized instruction to participants who were performing an exercise incorrectly or those who were ready to move on to a more difficult progression, with a particular emphasis on exercise form and safety. This allowed all participants to develop competency with key exercises while allowing those with experience and confidence to progress to more advanced movements in a timely manner.

In-person classes were offered once per week and lasted slightly longer than 1 hour. Classes began with approximately 5 minutes of aerobic warm-up during, which participants were encouraged to move somewhat vigorously in order to raise their heart rate and elicit blood flow to the working muscles. Following the warm-up, each class had different foci as shown in Table 1. Regardless of the focus of the week, 5-10 minutes were spent teaching the week's foundational exercises with particular emphasis on explaining proper form. In the early weeks of the intervention, participants were instructed to complete the exercises slowly, paying particular attention to ensure that they were feeling the target muscles activate or engage in the appropriate anatomical area. Individualized feedback from an assistant instructor correcting errors in technique was emphasized during this time, making sure that each participant received feedback from an instructor at least once during the class. As the weeks progressed, the lead instructor cued participants to increase the resistance or increase the speed, or increase both the speed and resistance concurrently during strength training exercises, as tolerated, and also increase balance challenges if they felt safe doing so.

XSL•FO
RenderX

Table 1. Weekly foci and foundational exercises.

Week	Foundational exercise 1	Foundational exercise 2	Foundational exercise 3	Foundational exercise 4	Focus of the week
1	Neutral spine or marionette pose	Pelvic tucks or tilts	Hip hikes	Hip hinge	Posture
2	Quick feet drills	Review hip hinge	"Traditional abs"	Self-test: static balance	Balance
3	Front and lateral arm raises (with band)	Hip hinge (review) with squat or chair sit or stand	Reverse lunge	Heel toe raises	Strength
4	Neutral spine and pelvic tuck or tilt—standing	Static balance	Reverse lunge with progression	Heel toe raises with tennis ball	Combine 3 pillars
5	Sit and stand transitions	Around the clock steps	Hip hinge with choice of W/T/Y/I arm extensions	Modified or regular jumping jacks	Floor to stand transition—floor exercises
6	Hip hinge (review) with squat or chair sit or stand	Lunge with chair progression	Single leg heel raises	Hing hinge+picking up objects	Compound strength movements
7	Drinking bird	Squat to lateral leg raise	Single leg heel raises	Toe raise walk around chair	Increase balance challenge
8	Warrior 2 with chair sit	Wall sit with wood chop	Introduce pivot	Quick feet multidirectional	Odd impact or multidirectional movements
9	Weight transfer with split stance	Squat into a high knee and arm reach	Pivot	Head rotations in a tandem stance	Combine balance with compound movements
10	Single leg stance with band pull downs and serial 7's	Lunge with chair support and head rotation	Standing superman with cognitive challenge	Goal post with arm slides	Introduce cognitive challenge
11	Reaching squats	Lateral jacks	Vertical push-ups	Knee drivers	Speed or power movements
12	N/A[a]	N/A	N/A	N/A	Collaborative workout

[a]N/A: not applicable.

For the remainder of the class, a variety of exercises focused on strength, balance, and posture (alone and in combination) in line with the focus of the week were completed in sets of 3 to 5 with pauses between sets to cue the next set of exercises. During the final 5 minutes, participants were guided through a brief cool down that incorporated stretching of the muscles used (most) during the exercise session.

Participants were also provided a secure link to a public-facing web-based video library that hosted a recording of the exercise session and a written description of the foundational exercises emphasized each week. Participants were further encouraged to complete 2 additional exercise sessions weekly. They also were provided a "mid-week challenge" 3 days after the weekly class session. These challenges included video presentations on the importance of posture in preventing falls, selected yoga-based movements for balance, and other short (5-10 minutes) presentations designed to keep participants engaged with the larger fall prevention focus of SF. Finally, participants were reminded of the importance of cardiovascular exercise and encouraged to walk or do other aerobic activities regularly.

Laboratory Measures

Overview

All measures were completed at baseline, prior to the first exercise class. Follow-up visits were completed anytime from the day after the 11th class through 2 weeks following the completion of the last (12th) exercise session. Participants completed their physical measures in the following order:

Balance and Physical Function

In-laboratory testing of balance and physical function included the Short Physical Performance Battery (SPPB) [28] combined with computerized dynamic posturography to quantify postural sway and the Timed Up and Go (TUG) [29].

Participants completed the standard SPPB, which includes (1) a standing balance challenge, in which the participant is asked to stand as still as possible with an increasingly narrow base of support; (2) regular walking speed at a "regular" pace on a 4-m course; and (3) leg strength gathered from the time taken to complete 5 chair stands. Continuous scores were normalized with a maximal score of 12 (4 points per measurement category) [28]. In an effort to address possible ceiling effects associated with the balance component of the SPPB, the BTrackS Balance Plate (version 7.5; Balance Tracking Systems) was used to capture overall postural sway, measured in centimeters.

Participants completed the TUG at both a normal walking speed and as fast as possible. Participants began in a seated position and were instructed to rise without using their arms to push off, walk 3 m, negotiate around an obstacle (rubber cone), return to their chair, and sit back down. Participants were allowed 1 practice trial and then completed each test twice with the best time (ie, lowest) score used for statistical analysis.

XSL•FO
RenderX

Postural Assessments

Posture was measured using a variety of methods designed to identify areas of postural deficiency. Specifically, for general posture, standing height was measured in two ways using a stadiometer (SECA 213) while participants were instructed to (1) "stand as you normally do" and (2) "stand as tall as possible." Both measurements were taken at the top of the breath cycle. To identify cervical and upper thoracic postural decline, the occiput to wall distance (OWD) [30] was measured in the same 2 manners (ie, regular and tall stance). More specifically, OWD was measured by having participants stand against a wall with their heels and buttocks touching the wall or as close as they could get comfortably. They were then instructed to look straight ahead (neutral head position) and hold still while the distance from the occiput to the wall was measured in centimeters. Finally, the Debrunner kyphometer and flexiruler were used to measure thoracic and lumbar curvature using previously published methodology [31,32]. Briefly, the Debrunner kyphometer provided an angle of curvature of the thoracic spine when placed in the joint space between T2 and T3 on one end and T12 and L1 on the other while participants stood in a normal or relaxed manner. The flexiruler is moldable and is placed with one end at the base of C7 and the other in the joint space between L5 and S1. The molded ruler is then traced onto a piece of paper, from which the kyphotic index was calculated by 2 independent raters as a function of the thoracic length and thoracic width using standardized procedures previously described (width/length×100) [31]. The average of the 2 raters was used for analyses.

Functional Muscular Strength

All participants completed a 30-second Chair Stand (30CS) and grip strength test. A subset (initial 3 cohorts of recruited participants) also completed isometric strength testing of hamstrings or quadriceps and trunk or lumbar muscle groups (procedure described in Isometric Muscular Strength section).

The 30CS used a chair with a seat height of 17 inches. Participants were instructed to keep their feet flat on the floor with their arms folded across their chest and touching their chest throughout the test. To be counted, the participant had to rise to a fully upright position and then return to a seated position while maintaining a fairly vertical body position (avoiding excessive forward lean). If participants did not maintain proper form, they were coached, but time was not stopped.

Hand Grip Strength

Hand grip strength was measured using an adjustable grip strength dynamometer (Jamar Plus Digital Hand Dynamometer). During the baseline visit, the grip or handlebar of the dynamometer was adjusted so the second joint of the fingers fits around the handle with handle size (1-5) recorded and used during subsequent visits. Participants were familiarized with the measurement by performing one submaximal effort on each hand. After becoming comfortable with the procedure, participants were instructed to hold the dynamometer with their arms at their side. They were then coached to take a deep breath in and squeeze as hard as possible as they exhaled. Measurement staff provided encouragement throughout each attempt. The measurement was repeated twice on each hand, alternating between the dominant and nondominant hand, with the highest score for each hand recorded in kilograms.

Isometric Muscular Strength

Isometric strength of the hamstrings, quadriceps, and spinal extensor muscles was measured using the Biodex System 4 PRO dynamometer (version 4.60; Biodex Medical Systems). Specifically, isometric strength of the hamstring and quadriceps muscle groups was measured with the leg held in at 45, 75, and 90 degrees of knee flexion. Participants performed 3 repetitions of 5-second maximal isometric contraction, extension followed by flexion, with 5 seconds of rest between repetitions at each angle. A 2-minute rest period was given between sets or leg positions. To decrease participant burden, only 1 leg was tested with participants indicating if either of their legs or knees had any previous injuries, or any current pain, after which the "healthiest" underwent testing. The average maximal contraction at each angle was recorded, and scores were combined into a flexion and extension composite score.

The Biodex Dual Position Back Ex/Flex Attachment, seated-compressed variation (isolated lumbar position), was used to measure the isometric strength of the trunk at 0 degrees of spinal flexion-extension (seated up-right). One set of 3 repetitions of 5-second maximal isometric contractions for both trunk flexion and extension was completed by participants. A 10-second rest was given between repetitions. To minimize the risk of injury, participants were instructed to slowly and gradually increase their muscle engagement to a volitional maximum and not attempt any sudden or explosive movement against the resistance.

Aerobic Fitness

Participants were asked to walk continuously for 2.5 minutes at their normal walking speed on a 50-m corridor [33]. During the walk, participants wore a chest strap–based heart rate monitor (Polar H10). Total distance, average heart rate, and peak heart rate were recorded.

Statistical Analyses

All statistical analyses were conducted using SPSS (version 27; IBM Corp). As this was largely designed as a feasibility study to determine the acceptability of this novel intervention delivery method and the degree to which the intervention could improve metrics of strength, balance, and posture, power calculations were not conducted. Instead, the sample size was primarily determined by safety concerns involving class size, which we settled on 10-15 per class after conducting 3 brief, 4-week pilot sessions beginning with ~6 participants, increasing to 8-12 to keep class size manageable. The sample size was also restricted by fiscal constraints.

Descriptive statistics (percentages, means, and SDs) were used to characterize demographic variables and identify potential outliers. Change scores were derived by subtracting baseline values from follow-up values on an individual level.

Independent 2-tailed *t* tests were conducted to assess differences across groups at baseline. Paired sample *t* tests were conducted to evaluate the effects of the intervention on the measures of

XSL•FO
RenderX

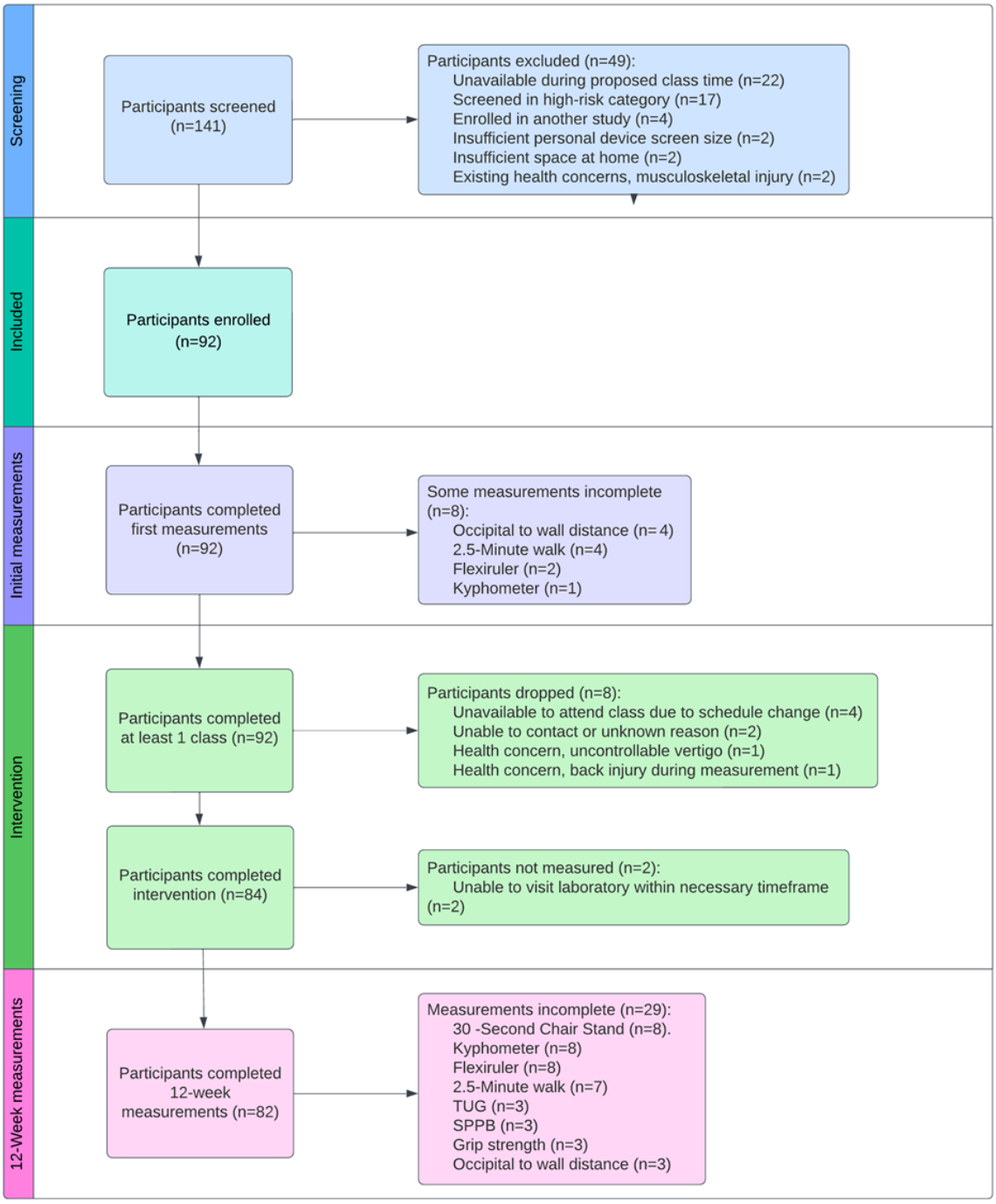
interest, with analyses conducted both for the total population and also with participants divided into fall-risk groups, which were assessed independently of each other. Associations between changes in measures of strength, balance, and posture and number of classes attended were examined using Pearson correlation without controlling for any covariates. When associations were observed, univariate linear modeling was conducted, with sex, chronological age, and baseline score included as covariates. These were included based on the known systematic declines in performance in older individuals, the possibility of systematic differences between male and female individuals, and the differences in potential change associated with baseline performance.

Results

Overview

The number of participants screened, enrolled, and followed through the final measurement is shown in Figure 1.

Figure 1. Participant recruitment flow. The sum of the breakdown of excluded participants at the initial and 12-week measurements exceeds the total number of excluded participants, as some individuals missed more than one assessment. SPPB: Short Physical Performance Battery; TUG: Timed Up and Go.



XSL•FO
RenderX

Population Descriptive Data

A total of 92 participants were enrolled across 8 cohorts with 55 low-risk and 37 moderate-risk individuals assigned to their own risk-group classes (n=4 classes per risk group). Key demographics and population descriptive data are shown in Table 2. The majority of the population was female (n=77, 84%) and White (n=74, 80%). Participants were largely well educated, with all participants having some college degree, and 91% (n=84) having a bachelor's degree or beyond. Additionally, although very few were currently employed (n=15, 16%), the majority (n=55, 60%) of participants were at a high socioeconomic status (>US $100,000 income), with 19 (21%) individuals earning more than US $200,000 per year.

As expected, based upon the group stratification, there was a difference in STEADI fall risk score between groups. Additionally, those at moderate risk (STEADI >3 but <8) were slightly heavier (low risk: mean 64.2, SD 11.3 kg and medium risk: mean 71.6, SD 17.7 kg; *P*=.008) and with a higher BMI (low risk: mean 23.8, SD 3.9 kg/m^2 and medium risk: mean 26.5, SD 5.2 kg/m^2; *P*=.003).

Performance metrics at baseline and following the 12-week intervention are shown in Table 3. There was a difference at baseline between the low- and moderate-risk groups on the 30CS (*P*<.001), TUG (*P*=.002), walking distance at normal speed over 2.5 minutes (*P*<.001), SPPB (*P*<.001), and the OWD measures of posture (normal stance: *P*=.007 and tall stance: *P*=.005). In all of these cases, the moderate-risk group had worse performance compared to the low-risk group. However, strength and balance and some measures of posture (flexiruler and Debrunner kyphometer) were not different by risk category.

Table 2. Demographic data of participants by fall-risk group.

	Total (N=92)	Low risk (n=55)	Moderate risk (n=37)	*P* value
Female, n (%)	77 (84)	47 (85)	30 (81)	.58
Race, n (%)				N/A[a]
Asian	8 (9)	5 (9)	3 (8)	
Black	2 (2)	1 (2)	1 (3)	
Hispanic	3 (3)	1 (2)	2 (5)	
White	74 (80)	45 (82)	29 (78)	
Other	5 (5)	3 (5)	2 (5)	
Age (years), mean (SD)	72.0 (6.6)	71.3 (6.9)	73.0 (6.2)	.25
Height (cm), mean (SD)	163.9 (7.7)	164.0 (6.0)	163.7 (9.9)	.43
Weight (kg), mean (SD)	67.1 (14.6)	64.2 (11.3)	71.6 (17.7)	*.008* [b]
BMI (kg/cm^2), mean (SD)	24.9 (4.6)	23.8 (3.9)	26.5 (5.1)	*.003*
Income (US $ per year), n (%)				
<50,000	12 (13)	4 (7)	8 (22)	.11
>50,000 but <100,000	24 (26)	15 (27)	9 (24)	N/A
>100,000	55 (60)	35 (64)	20 (54)	N/A
STEADI[c] risk score (0-12 range), mean (SD)	3.3 (2.0)	2.0 (1.4)	5.2 (1.3)	*<.001*
Classes attended, mean (SD)	9.7 (3.1)	9.4 (3.5)	10.3 (2.3)	.08

[a]N/A: not applicable.

[b]Values in italics format indicate statistical significance at or below *P*<.05.

[c]STEADI: Stopping Elderly Accidents and Deaths Initiative.

XSL•FO
RenderX

Table 3. Intervention effects on performance-based measures by fall-risk category.

	Total (n=82 except where noted)			Low (n=48 except where noted)			Moderate (n=34 except where noted)		
	Preintervention, mean (SD)	Postintervention, mean (SD)	*P* value	Preintervention, mean (SD)	Postintervention, mean (SD)	*P* value	Preintervention, mean (SD)	Postintervention, mean (SD)	*P* value
Kyphotic index (%)	11.4 (3.4)	10.7 (3.5)	*.007* [a]	11.4 (3.5)	10.8 (3.3)	.20	11.4 (3.3)	10.5 (3.8)	*.008*
Kyphotic angle (degrees)	40.3 (11.9)	37.7 (13.7)	*.005*	40.3 (10.9)	38 (11.2)	*.02*	40.2 (13.6)	37.3 (17.3)	*.008*
OWD[b]—normal stance (cm)	4.8 (3.8)	3.5 (3.8)	*<.001*	3.9 (3.4)	2.7 (3.2)	*.004*	6.1 (4)	4.7 (4.3)	*<.001*
OWD—tall stance (cm)	2.5 (3.0)	1.6 (2.7)	*<.001*	1.8 (2.7)	0.8 (2.1)	*.002*	3.6 (3.3)	2.8 (2.9)	*.01*
30-Second Chair Stand (repetitions)	12.9 (4.9)	14.3 (5.3)	*.006*	14.7 (4.4)	15.6 (5.1)	.44	10.1 (4.3)	12.3 (5.0)	*.002*
Grip combined (kg)	47.8 (13.8)	47.6 (14.4)	.54	47.9 (14.1)	47.1 (14.1)	.18	47.8 (13.7)	48.1 (15.0)	.60
Isometric leg strength extension (kg) (n=26 with/19 low and 7 moderate)	199 (80.6)	206 (79.1)	*.02*	196.1 (84.6)	193.7 (65.7)	.07	213.4 (62.1)	253.3 (112.6)	.11
Isometric leg strength flexion (kg) (n=26 with/19 low and 7 moderate)	93.9 (42.9)	98.7 (41.1)	.13	94.2 (45)	97.0 (37.7)	.29	92.3 (34.5)	105.2 (56.0)	.20
Isometric trunk strength extension (kg) (n=26 with/19 low and 7 moderate)	70.3 (28.1)	83.0 (40.0)	*.03*	72.5 (28.5)	90.9 (38.0)	*.01*	64.6 (28.1)	61.8 (40.6)	.75
Isometric trunk strength flexion (kg) (n=26 with/19 low and 7 moderate)	43.9 (21.1)	51.8 (20.4)	*<.001*	45.5 (19.8)	54.1 (17.9)	*.002*	39.7 (25.4)	45.9 (26.4)	.16
SPPB[c] (scored from 0 to 12)	11 (1.4)	11.2 (1.2)	.13	11.5 (0.9)	11.5 (1.0)	.40	10.2 (1.6)	10.8 (1.4)	*.02*
Timed Up and Go (seconds)	6.9 (1.7)	6.7 (1.5)	*.004*	6.3 (1.2)	6.1 (1.2)	*.02*	7.9 (1.9)	7.5 (1.6)	.09
Postural sway (cm)	91.4 (31.1)	92.6 (35)	.43	88.4 (27)	86.8 (25.7)	.98	96.7 (37.2)	101.5 (44.8)	.16
Distance during 2.5-minute walk (m)	196.3 (30.7)	193.1 (31.6)	.18	209.2 (29.2)	202.5 (29.9)	.05	179.9 (24.5)	181.0 (33.4)	.75

[a]Values in italics format indicate statistically significant values with a *P* value of less than .05.

[b]OWD: occiput to wall distance.

[c]SPPB: Short Physical Performance Battery.

Balance and Physical Function

On average, at baseline, participants were able to complete slightly more than the minimum number of chair stands (n<12) to be considered free from disability [34]. However, 38 (41%) individuals were below that level. In contrast, on average, participants were well above the minimum threshold to identify disability for the TUG and SPPB (>13.5 seconds and <9, combined score respectively) [35,36], with 0 for TUG and 14 (16%) for SPPB below the threshold. In terms of the effects of the intervention, there were no significant changes in postural sway during static balance stances associated with the SPPB.

XSL•FO
RenderX

However, overall participants significantly improved on the 30CS and TUG (mean change of 1.4 repetitions: *P*=.006 and 0.2 seconds: *P*=.004, respectively). While not always statistically significant, the moderate-risk group improved to a greater degree on both of these assessments than the low-risk group (although both groups improved). Additionally, SPPB performance improved significantly in the moderate-risk category (*P*=.02).

Posture

Although there were no inclusion or exclusion criteria related to posture, there were a meaningful number that displayed suboptimal posture. Indeed, depending upon the measurement tool and associated categorical cut-points used, between 7% and 26% of the total sample were hyperkyphotic (flexiruler: n=7, 8%; kyphometer: n=12, 13%; OWD [normal stance]: n=24, 26%). The majority of postural measures showed statistically significant improvement following the intervention. Indeed, measures of "relaxed" posture showed improvements between 6% (n=5; Debrunner kyphometer) and 27% (n=22; OWD). When an "as tall as possible" posture was adopted for the OWD, the improvements were greater at ~36% change from baseline. These changes are even more impressive when individuals who had no opportunity for improvement are removed (ie, individuals who scored 0 on OWD at baseline). Specifically, when the 18 individuals standing normally and 47 as tall as possible who scored 0 are removed, overall change scores with normal stance were 1.7 cm, and with tall stance were 2.2 cm.

Muscular Strength

We did not observe significant differences in grip strength (*P*=.54) However, as mentioned earlier, 30CS is a marker both of functional movement but also muscular strength (and endurance) and had significant changes across the population (*P*=.006) Finally, in the subgroup that completed isometric testing, there was a significant improvement in knee extension (*P*=.02) and both trunk extension and flexion (*P*=.03 and <.001, respectively).

Aerobic Fitness

There were no significant changes in measures of aerobic fitness measured by the 2.5-minute walk at normal speed.

Associations Between Class Attendance and Functional Measures

When exploring associations between change scores and rates of attendance, the TUG (*r*=–0.255; *P*=.03) and 30CS (*r*=0.27; *P*=.02) were significantly correlated; however, there were no significant associations between class attendance and changes in physical function on the majority of other variables (*P*>.05). Regression analysis controlling for sex, age, baseline measurement, and number of classes attended indicated that attendance was not a significant predictor of change in TUG, although including it did slightly improve the model's predictive value (from *R*=0.45 to 0.49). However, attendance was a significant predictor of improvement in the 30CS (unstandardized β=.502; *P*=.02). This indicates that for each class attended, there was a 0.5 increase in the change score in terms of number of repetitions. Model summary and individual contributions of variables and covariates are shown in Tables 4 and 5.

Table 4. Linear regression analysis of functional fitness and attendance change in the Timed Up and Go[a].

Predictors	Unstandardized β (SE)	Standardized β coefficient	*t* test (*df*=72)	*P* value
Constant	–.382 (0.981)	N/A[b]	–0.389	.70
Sex[c]	–.066 (0.19)	–0.036	–0.346	.73
Chronological age (years)	.031 (0.012)	0.297	2.657	*.01*[d]
Baseline TUG[e] (seconds)	–.174 (0.047)	–0.417	–3.72	*<.001*
Attendance	–.073 (0.04)	–0.191	–1.81	.07

[a]Model summary: *R*=0.485; R^2=0.235; adjusted R^2=0.193; standard error of the estimate=0.6118; *P*=.07.

[b]N/A: not applicable.

[c]Male: n=1 and female: n=2.

[d]Values in italics format indicate significance at, or below, the level of .05.

[e]TUG: Timed Up and Go.

XSL•FO
RenderX

Table 5. Linear regression analysis of functional fitness and attendance change in the 30-second Chair Stand[a].

Predictors	Unstandardized β (SE)	Standardized β coefficient	*t* test (*df*=72)	*P* value
Constant	6.172 (5.689)	N/A[b]	1.085	.28
Sex[c]	–.785 (0.991)	–0.089	–0.792	.43
Chronological age (years)	–.096 (0.06)	–0.185	–1.6	.11
Baseline 30-second Chair Stand	–.156 (0.079)	–0.227	–1.978	.05
Attendance	.502 (0.21)	0.266	2.393	*.02*[d]

[a]Model summary: R=0.374; R^2=0.14; adjusted R^2=0.092; standard error of the estimate=3.173; P=.02.

[b]N/A: not applicable.

[c]Male: n=1 and female: n=2.

[d]Values in italics format indicate significance at, or below, the level of .05.

Adverse Events

Based upon weekly tracking, there were several minor adverse events throughout the intervention period. However, none were related to the exercise intervention, and only 1 minor event was possibly related to the measurements. Additionally, there were no falls associated with the intervention, either while supervised during class or performed on nonclass days.

Discussion

Principal Findings

In this 12-week, iterative, web-based program, we found that older adults experienced an improvement in several measures widely used to evaluate fall risk and mobility-related disability [7,28,34]. While the magnitude of most of the changes was small, they were primarily consistent in the expected direction, which is not surprising because the program was designed to be largely instructional rather than a progressive training program. Instructors emphasized proper form and execution while maintaining good posture. Increases in the intensity of resistance exercises and challenges to balance and mobility exercises were only introduced once instructors observed consistent execution of movements. Importantly, no injuries were sustained during the remote delivery of our program or during the recommended home exercise practice.

Most muscular strength and physical function measures showed favorable changes in both the low- and moderate-risk groups, although not all results reached statistical significance. Nevertheless, the observed changes have clinical relevance. For instance, the intervention induced meaningful improvements in measures commonly used in clinical settings, such as the 30CS and TUG tests. Notably, many improvements varied by risk level, with greater gains seen in those at higher fall risk, who arguably had more to gain from the intervention. Specifically, participants in the moderate-risk group completed an additional 2 chair rises (a 21% improvement) by the end of the 12 weeks. Given that both leg strength and muscular endurance are essential for scoring well on the 30CS, the observed improvement in chair stands may be attributed to the focused practice of squat and related exercises in class as well as encouragement for similar home practice. While upper body or arm strength is important to overall good health and function, they do not substantially contribute to fall prevention and were therefore not heavily emphasized in our home practice recommendations. Similarly, the lack of significant changes in aerobic capacity was expected based on the program's goals and the unique challenges of aerobic activities conducted via a Zoom format (Zoom Video Communications).

Additionally, these findings also contextualize participant gains by providing deeper phenotypical measurements related to posture, more sophisticated measures of strength, and balance. For example, there were positive changes recorded in measures of thoracic kyphosis and OWD across both risk groups. This is important because posture is often overlooked in many group exercise programs for both younger and older adults. In this study population, posture measured as the kyphotic angle using a Debrunner kyphometer decreased 2-3 degrees (P=.02 and .07 for low- and moderate-risk groups, respectively), while the OWD decreased between 1 and 1.4 cm (all P<.05) among participants in both groups over the once-weekly, 12-week exercise program. Considering that thoracic kyphosis typically worsens steadily over time beginning as early as age 40 years [37-39], observing a reversal of this condition within just 12 weeks of targeted exercise is highly encouraging. This is particularly meaningful, given the observed links between hyperkyphosis and both fall risk [26,40] and physical function [41,42].

Although there were no significant differences observed in grip strength, it is likely that the prescribed exercises contributed to an enhancement of the participant's back strength [43] and chest flexibility, which helped alleviate the tightness in chronically shortened pectoral muscles [26]. These improvements not only reflected positively in participants' attempts at maintaining "good" posture (as measured by OWD while standing tall) but also in their "normal" posture as assessed by various tools. These changes in both "good" and "normal" posture likely stemmed from the early introduction of postural control as a foundational element of the program as well as ongoing reminders throughout the course for participants to practice maintaining good posture during daily activities. It is worth noting that postural control is seldom included in group exercise training programs, making this approach particularly novel and engaging, thereby providing participants with more opportunities to learn and improve compared to more traditional exercise modalities such as aerobic activities or strength training.

XSL•FO
RenderX

While there have been some similar findings regarding the possibility of providing digital exercise instruction to older adults [17-19], this is, to our knowledge, the first study to robustly measure a large population of older adults across multiple metrics of physical health and function. Our findings clearly demonstrate the potential to deliver an effective fall prevention program through a technological interface. The program was well-received, as indicated by an attendance rate exceeding 80% (mean attendance 9.7 of 12), and no adverse events were reported in association with the exercise regimen. Additionally, the observed improvements in leg strength, postural control, and overall mobility, both independently and synergistically, show promise for reducing falls among older adults, including those at substantial risk.

Limitations

Despite several strengths of this study, certain limitations must be acknowledged. Chief among them is the absence of a control group, which limits our ability to ascertain whether observed changes were influenced by external factors. A control group would also help clarify the significance of nonsignificant changes observed in some metrics, particularly in light of the expected performance declines over time. Additionally, although there were improvements in most metrics of strength and posture, balance as measured in this context was not improved. This may be a function of the method of measuring static balance (ie, postural sway during the SPPB). However, it may be that the program's focus on balance was not sufficient to induce changes over 12 weeks. Furthermore, unlike many exercise interventions for adults of all ages, we did not emphasize cardiovascular health other than recommending daily general PA. While recognizing the importance of aerobic activity for overall health and longevity, instructional focus on this area was minimal; encouragement for at-home exercise emphasized practicing postural, strength, and balance-based exercises. Finally, the relatively affluent nature of the population limits the generalizability of both the results and the ability to disseminate this program to a wider population.

With the success of this program and the limitations noted earlier in mind, future research should explore the possibility of deploying this intervention in populations that are (1) of a lower overall socioeconomic status and (2) more remote from the location of intervention deployment. In addition, developing these materials into other languages in a culturally appropriate manner offers substantial potential to expand the reach to other populations who would benefit from the opportunity to receive at-home fall-risk training.

Conclusions

The 12-week SF program improved physical function and posture in ways consistent with reduced fall risk. The real-time instruction also helps ensure program safety and adherence. As such, this program shows promise in digitally delivering needed public health interventions targeting fall risk among older adults.

Acknowledgments

Study data were collected and managed using REDCap (Research Electronic Data Capture) electronic data capture tools hosted at the University of California, San Diego (UCSD). The REDCap software system provided by the UCSD Clinical and Translational Research Center is supported by (award UL1TR001442) the National Center for Research Resources. The authors would like to thank Ms Ashley Gauthie and Ms Sharmaine Olivarez for their work in managing the technical elements of Zoom-based intervention deployment. The authors would also like to thank their funder, the UCSD Clinical and Translation Research Center.

Data Availability

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

Conflicts of Interest

None declared.

References

1. The graying of America: more older adults than kids by 2035. United States Census Bureau. 2018. URL: https://www.census.gov/library/stories/2018/03/graying-america.html [accessed 2024-10-03]
2. Stevens J. A CDC compendium of effective fall interventions: what works for community-dwelling older adults. In: Injury Prevention, 2nd Edition. Scotts Valley, CA: CreateSpace Independent Publishing; 2010.
3. Lundin-Olsson L, Nyberg L, Gustafson Y. Attention, frailty, and falls: the effect of a manual task on basic mobility. J Am Geriatr Soc 1998;46(6):758-761. [doi: 10.1111/j.1532-5415.1998.tb03813.x] [Medline: 9625194]
4. Tinetti ME, Speechley M, Ginter SF. Risk factors for falls among elderly persons living in the community. N Engl J Med 1988;319(26):1701-1707. [doi: 10.1056/NEJM198812293192604] [Medline: 3205267]
5. Baloh RW, Corona S, Jacobson KM, Enrietto JA, Bell T. A prospective study of posturography in normal older people. J Am Geriatr Soc 1998;46(4):438-443. [doi: 10.1111/j.1532-5415.1998.tb02463.x] [Medline: 9560065]
6. Espinosa de Los Monteros K, Gallo LC, Elder JP, Talavera GA. Individual and area-based indicators of acculturation and the metabolic syndrome among low-income Mexican American women living in a border region. Am J Public Health 2008;98(11):1979-1986 [FREE Full text] [doi: 10.2105/AJPH.2008.141903] [Medline: 18799765]

XSL•FO
RenderX

7. Gill TM. Assessment of function and disability in longitudinal studies. J Am Geriatr Soc 2010;58(Suppl 2):S308-S312 [FREE Full text] [doi: 10.1111/j.1532-5415.2010.02914.x] [Medline: 21029059]
8. Stevens JA, Phelan EA. Development of STEADI: a fall prevention resource for health care providers. Health Promot Pract 2013;14(5):706-714 [FREE Full text] [doi: 10.1177/1524839912463576] [Medline: 23159993]
9. Sarmiento K, Lee R. STEADI: CDC's approach to make older adult fall prevention part of every primary care practice. J Safety Res 2017;63:105-109 [FREE Full text] [doi: 10.1016/j.jsr.2017.08.003] [Medline: 29203005]
10. US Preventive Services Task Force, Grossman DC, Curry SJ, Owens DK, Barry MJ, Caughey AB, et al. Interventions to prevent falls in community-dwelling older adults: US Preventive Services Task Force recommendation statement. JAMA 2018;319(16):1696-1704 [FREE Full text] [doi: 10.1001/jama.2018.3097] [Medline: 29710141]
11. Belza B, Shumway-Cook A, Phelan EA, Williams B, Snyder SJ, LoGerfo JP. The effects of a community-based exercise program on function and health in older adults: the EnhanceFitness program. J Appl Gerontol 2006 Aug 01;25(4):291-306. [doi: 10.1177/0733464806290934]
12. Bethancourt HJ, Rosenberg DE, Beatty T, Arterburn DE. Barriers to and facilitators of physical activity program use among older adults. Clin Med Res 2014;12(1-2):10-20 [FREE Full text] [doi: 10.3121/cmr.2013.1171] [Medline: 24415748]
13. Pelletier CA, White N, Duchesne A, Sluggett L. Barriers to physical activity for adults in rural and urban Canada: a cross-sectional comparison. SSM Popul Health 2021;16:100964 [FREE Full text] [doi: 10.1016/j.ssmph.2021.100964] [Medline: 34841038]
14. Kakara R, Bergen G, Burns E, Stevens M. Nonfatal and fatal falls among adults aged ≥65 years—United States, 2020–2021. MMWR Morb Mortal Wkly Rep 2023;72(35):938-943 [FREE Full text] [doi: 10.15585/mmwr.mm7235a1] [Medline: 37651272]
15. Mace RA, Mattos MK, Vranceanu A. Older adults can use technology: why healthcare professionals must overcome ageism in digital health. Transl Behav Med 2022;12(12):1102-1105 [FREE Full text] [doi: 10.1093/tbm/ibac070] [Medline: 36073770]
16. Jacobson CL, Foster LC, Arul H, Rees A, Stafford RS. A digital health fall prevention program for older adults: feasibility study. JMIR Form Res 2021;5(12):e30558 [FREE Full text] [doi: 10.2196/30558] [Medline: 34837492]
17. Yerlikaya T, Öniz A, Özgüren M. The effect of an interactive tele rehabilitation program on balance in older individuals. Neurol Sci Neurophys 2021;38(3):180-186 [FREE Full text] [doi: 10.4103/nsn.nsn_91_21]
18. Ing JBM, Tan MP, Whitney J, Tiong IK, Singh DKA. Acceptability, feasibility, and effectiveness of WE-SURF™: a virtual supervised group-based fall prevention exercise program among older adults. Aging Clin Exp Res 2024;36(1):125 [FREE Full text] [doi: 10.1007/s40520-024-02759-x] [Medline: 38836944]
19. Langeard A, Bigot L, Maffiuletti NA, Moussay S, Sesboüé B, Quarck G, et al. Non-inferiority of a home-based videoconference physical training program in comparison with the same program administered face-to-face in healthy older adults: the MOTION randomised controlled trial. Age Ageing 2022;51(3):afac059. [doi: 10.1093/ageing/afac059] [Medline: 35290431]
20. Blioumpa C, Karanasiou E, Antoniou V, Batalik L, Kalatzis K, Lanaras L, et al. Efficacy of supervised home-based, real time, videoconferencing telerehabilitation in patients with type 2 diabetes: a single-blind randomized controlled trial. Eur J Phys Rehabil Med 2023;59(5):628-639 [FREE Full text] [doi: 10.23736/S1973-9087.23.07855-3] [Medline: 37350165]
21. Terkes N, Aksu NT, Yamac SU. The effect of an online-supervised exercise program in older people with diabetes on fasting blood sugar, psychological resilience and quality of life: a double blind randomised controlled trial. Int J Older People Nurs 2023;18(5):e12564. [doi: 10.1111/opn.12564] [Medline: 37559428]
22. Li F, Harmer P, Fitzgerald K, Winters-Stone K. A cognitively enhanced online Tai Ji Quan training intervention for community-dwelling older adults with mild cognitive impairment: a feasibility trial. BMC Geriatr 2022;22(1):76 [FREE Full text] [doi: 10.1186/s12877-021-02747-0] [Medline: 35078407]
23. Pitre V, Sénéchal M, Bouchard DR. Comparing online and in-person delivery of a fall prevention exercise program for older adults. J Aging Phys Act 2024;32(1):55-61. [doi: 10.1123/japa.2023-0070] [Medline: 37741634]
24. Granet J, Peyrusqué E, Ruiz F, Buckinx F, Abdelkader LB, Dang-Vu TT, et al. Web-based physical activity interventions are feasible and beneficial solutions to prevent physical and mental health declines in community-dwelling older adults during isolation periods. J Gerontol A Biol Sci Med Sci 2023;78(3):535-544 [FREE Full text] [doi: 10.1093/gerona/glac127] [Medline: 35675174]
25. Kado DM, Huang M, Nguyen CB, Barrett-Connor E, Greendale GA. Hyperkyphotic posture and risk of injurious falls in older persons: the Rancho Bernardo Study. J Gerontol A Biol Sci Med Sci 2007;62(6):652-657. [doi: 10.1093/gerona/62.6.652] [Medline: 17595423]
26. van der Jagt-Willems HC, de Groot MH, van Campen JPCM, Lamoth CJC, Lems WF. Associations between vertebral fractures, increased thoracic kyphosis, a flexed posture and falls in older adults: a prospective cohort study. BMC Geriatr 2015;15:34 [FREE Full text] [doi: 10.1186/s12877-015-0018-z] [Medline: 25888399]
27. Rubenstein LZ, Vivrette R, Harker JO, Stevens JA, Kramer BJ. Validating an evidence-based, self-rated Fall Risk Questionnaire (FRQ) for older adults. J Safety Res 2011;42(6):493-499. [doi: 10.1016/j.jsr.2011.08.006] [Medline: 22152267]

XSL•FO
RenderX

28. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994;49(2):M85-M94. [doi: 10.1093/geronj/49.2.m85] [Medline: 8126356]
29. Podsiadlo D, Richardson S. The timed "Up and Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991;39(2):142-148. [doi: 10.1111/j.1532-5415.1991.tb01616.x] [Medline: 1991946]
30. Wiyanad A, Chokphukiao P, Suwannarat P, Thaweewannakij T, Wattanapan P, Gaogasigam C, et al. Is the occiput-wall distance valid and reliable to determine the presence of thoracic hyperkyphosis? Musculoskelet Sci Pract 2018;38:63-68. [doi: 10.1016/j.msksp.2018.09.010] [Medline: 30278368]
31. Milne JS, Lauder IJ. The relationship of kyphosis to the shape of vertebral bodies. Ann Hum Biol 1976;3(2):173-179 [FREE Full text] [doi: 10.1080/03014467600001281] [Medline: 1275439]
32. Ohlén G, Spangfort E, Tingvall C. Measurement of spinal sagittal configuration and mobility with Debrunner's kyphometer. Spine (Phila Pa 1976) 1989;14(6):580-583. [doi: 10.1097/00007632-198906000-00006] [Medline: 2749372]
33. Rossman I. Normal human aging: the Baltimore Longitudinal Study of Aging. JAMA 1986;255(7):960. [doi: 10.1001/jama.1986.03370070114046]
34. Rikli R, Jones C. Functional fitness normative scores for community-residing older adults, ages 60-94. J Aging Phys Act 1999;7(2):129-161 [FREE Full text] [doi: 10.1123/japa.7.2.129]
35. Rose D, Jones C, Lucchese N. Predicting the probability of falls in community-residing older adults using the 8-foot up-and-go: a new measure of functional mobility. J Aging Phys Act 2002;10(4):466-475 [FREE Full text] [doi: 10.1123/japa.10.4.466]
36. de Fátima Ribeiro Silva C, Ohara DG, Matos AP, Pinto ACPN, Pegorari MS. Short Physical Performance Battery as a measure of physical performance and mortality predictor in older adults: a comprehensive literature review. Int J Environ Res Public Health 2021;18(20):10612 [FREE Full text] [doi: 10.3390/ijerph182010612] [Medline: 34682359]
37. Fon GT, Pitt MJ, Thies AC. Thoracic kyphosis: range in normal subjects. Am J Roentgenol 1980;134(5):979-983. [doi: 10.2214/ajr.134.5.979] [Medline: 6768276]
38. Katzman WB, Wanek L, Shepherd JA, Sellmeyer DE. Age-related hyperkyphosis: its causes, consequences, and management. J Orthop Sports Phys Ther 2010;40(6):352-360 [FREE Full text] [doi: 10.2519/jospt.2010.3099] [Medline: 20511692]
39. Ensrud KE, Black DM, Harris F, Ettinger B, Cummings SR. Correlates of kyphosis in older women. The Fracture Intervention Trial Research Group. J Am Geriatr Soc 1997;45(6):682-687. [doi: 10.1111/j.1532-5415.1997.tb01470.x] [Medline: 9180660]
40. McDaniels-Davidson C, Davis A, Wing D, Macera C, Lindsay SP, Schousboe JT, et al. Kyphosis and incident falls among community-dwelling older adults. Osteoporos Int 2018;29(1):163-169 [FREE Full text] [doi: 10.1007/s00198-017-4253-3] [Medline: 29018904]
41. Katzman WB, Vittinghoff E, Ensrud K, Black DM, Kado DM. Increasing kyphosis predicts worsening mobility in older community-dwelling women: a prospective cohort study. J Am Geriatr Soc 2011;59(1):96-100 [FREE Full text] [doi: 10.1111/j.1532-5415.2010.03214.x] [Medline: 21198460]
42. Katzman WB, Huang M, Lane NE, Ensrud KE, Kado DM. Kyphosis and decline in physical function over 15 years in older community-dwelling women: the study of osteoporotic fractures. J Gerontol A Biol Sci Med Sci 2013;68(8):976-983 [FREE Full text] [doi: 10.1093/gerona/glt009] [Medline: 23633167]
43. Katzman WB, Vittinghoff E, Lin F, Schafer A, Long RK, Wong S, et al. Targeted spine strengthening exercise and posture training program to reduce hyperkyphosis in older adults: results from the study of hyperkyphosis, exercise, and function (SHEAF) randomized controlled trial. Osteoporos Int 2017;28(10):2831-2841 [FREE Full text] [doi: 10.1007/s00198-017-4109-x] [Medline: 28689306]

Abbreviations

30CS: 30-second Chair Stand
CDC: Centers for Disease Control and Prevention
OWD: occiput to wall distance
PA: physical activity
SF: Strong Foundations
SPPB: Short Physical Performance Battery
STEADI: Stopping Elderly Accidents and Deaths Initiative
TUG: Timed Up and Go

XSL•FO
RenderX

Edited by M Gray; submitted 21.11.24; peer-reviewed by D Kunkel, E Okoyeocha; accepted 02.01.25; published 26.02.25.

Please cite as:
Wing D, Nichols JF, Barkai HS, Culbert O, Moreno D, Higgins M, O'Brien A, Perez M, Davey H, Moran R
Building Strong Foundations: Nonrandomized Interventional Study of a Novel, Digitally Delivered Fall Prevention Program for Older Adults
JMIR Aging 2025;8:e68957
URL: https://aging.jmir.org/2025/1/e68957
doi:10.2196/68957
PMID:

©David Wing, Jeanne F Nichols, Hava Shoshana Barkai, Olivia Culbert, Daniel Moreno, Michael Higgins, Anna O'Brien, Mariana Perez, Hope Davey, Ryan Moran. Originally published in JMIR Aging (https://aging.jmir.org), 26.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Development of a Predictive Dashboard With Prescriptive Decision Support for Falls Prevention in Residential Aged Care: User-Centered Design Approach

S Sandun Malpriya Silva[1], BSc, PhD; Nasir Wabe[1], BPharm, MScClinPharm, MScEpid, PhD; Amy D Nguyen[1], BMedSc, PhD; Karla Seaman[1], BPharm, MClinEpid, PhD; Guogui Huang[1], BA, MA, PhD; Laura Dodds[2], BA, MPH; Isabelle Meulenbroeks[1], BPhysio, MPH, PhD; Crisostomo Ibarra Mercado[1], BSN, MPH; Johanna I Westbrook[1], BAppSc, GradDipAppEpid, MHA, PhD

[1]Centre for Health Systems and Safety Research, Australian Institute of Health Innovation, Macquarie University, North Ryde, NSW, Australia

[2]The MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Sydney, New South Wales, Australia

Corresponding Author:
S Sandun Malpriya Silva, BSc, PhD
Centre for Health Systems and Safety Research
Australian Institute of Health Innovation
Macquarie University
75 Talavera Road
North Ryde, NSW, 2113
Australia
Phone: 61 29850 4045
Email: sandunmalpriya.silva@mq.edu.au

Abstract

Background: Falls are a prevalent and serious health condition among older people in residential aged care facilities, causing significant health and economic burdens. However, the likelihood of future falls can be predicted, and thus, falls can be prevented if appropriate prevention programs are implemented. Current fall prevention programs in residential aged care facilities rely on risk screening tools with suboptimal predictive performance, leading to significant concerns regarding resident safety.

Objective: This study aimed to develop a predictive, dynamic dashboard to identify residents at risk of falls with associated decision support. This paper provides an overview of the technical process, including the challenges faced and the strategies used to overcome them during the development of the dashboard.

Methods: A predictive dashboard was co-designed with a major residential aged care partner in New South Wales, Australia. Data from resident profiles, daily medications, fall incidents, and fall risk assessments were used. A dynamic fall risk prediction model and personalized rule-based fall prevention recommendations were embedded in the dashboard. The data ingestion process into the dashboard was designed to mitigate the impact of underlying data system changes. This approach aims to ensure resilience against alterations in the data systems.

Results: The dashboard was developed using Microsoft Power BI and advanced R programming by linking data silos. It includes dashboard views for those managing facilities and for those caring for residents. Data drill-through functionality was used to navigate through different dashboard views. Resident-level change in daily risk of falling and risk factors and timely evidence-based recommendations were output to prevent falls and enhance prescriptive decision support.

Conclusions: This study emphasizes the significance of a sustainable dashboard architecture and how to overcome the challenges faced when developing a dashboard amid underlying data system changes. The development process used an iterative dashboard co-design process, ensuring the successful implementation of knowledge into practice. Future research will focus on the implementation and evaluation of the dashboard's impact on health processes and economic outcomes.

International Registered Report Identifier (IRRID): RR2-https://doi.org/10.1136/bmjopen-2021-048657

(JMIR Aging 2025;8:e63609) doi:10.2196/63609

XSL•FO
RenderX

KEYWORDS

falls prevention; dashboard architecture; predictive; sustainability; challenges; decision support; falls; aged care; geriatric; older adults; economic burden; prevention; electronic health record; EHR; intervention; decision-making; patient safety; risks; older people; monitoring

Introduction

Falls are the second leading cause of unintentional injury deaths worldwide, with those older than 60 years having the highest risk of death or serious injury [1]. People living in residential aged care facilities (RACFs; ie, nursing homes and care homes) are three times more likely to fall than their community-dwelling peers and are ten times more likely to sustain fall-related injuries [2]. Studies have found that falls are often predictable and preventable and that the health and economic burden of falls could be minimized if effective prevention strategies are implemented [3-5].

Falls prevention is an issue of major importance in RACFs. Various fall risk assessment tools (FRATs) are used by RACF staff internally to identify residents at greatest risk of falls [6]. However, when assessed, the predictive performance of these tools has generally shown to be poor [6,7]. Moreover, FRATs are most often completed when residents are admitted to a facility or after a fall incident, and thus, only represent data collected at a single point in time, resulting in a static prediction that does not reflect changing falls risk factors (eg, if a resident's medication use changes) [6].

As a growing number of RACFs implement electronic health record (EHR) systems, new opportunities have emerged to develop a personalized, dynamic approach to predicting residents' fall risk by taking advantage of multiple potential contributory factors [8]. Some studies have integrated routinely collected EHR data including vital signs into the development of fall prediction tools through the application of machine learning models [9]. Despite these tools exhibiting superior performance in contrast to conventional FRATs, the risk predictions generated are often difficult to interpret and apply in practice. Thus, relying solely on these models will not enhance decision-making to provide a personalized care strategy for fall prevention and management. Health dashboards that can integrate these predictive models in combination with visualization features may be a more effective intervention for supporting staff in RACFs.

A dashboard can be defined as an at-a-glance real-time or near real-time processing interface, showing a graphical presentation of the current or recent status, along with historical trends or organizational key performance indicators to support decision-making [10]. There have been some attempts to develop a dashboard for fall prevention purposes globally. The falls dashboard, developed by the Network of Patient Safety Databases and managed by the US Department of Health and Human Services, collects nonidentifiable patient-related fall incident data from health care service providers. These data are then aggregated to a national level, analyzed, and visualized using the Network of Patient Safety Database dashboards and chartbooks to identify and track patient safety risks nationally [11]. Similarly, the New Zealand Health Quality and Safety Commission has established a falls and fracture outcomes dashboard to help the health sector evaluate the benefits of the services provided to older people [12]. Such aggregated dashboards are helpful for monitoring, although they offer limited assistance in predicting and preventing falls in routine care settings. To the best of our knowledge, no previous study has reported the development of a dashboard designed to predict fall risks for individual residents and aid decision-making in preventing and managing falls. As part of a comprehensive research program aimed at enhancing fall prevention and the well-being of older adults in RACFs [8], this paper details the design and development process of a predictive analytics dashboard with prescriptive decision support for fall risk prediction, prevention, and management. We discuss the challenges encountered and the strategies used to overcome them during its development.

Methods

Participants

This project was conducted in collaboration with a large, aged care provider that operates 24 RACFs in Sydney, Australia. The development process of the dashboard spanned from January 1, 2022, to December 31, 2023. For the dashboard views and study sample, this study used data from both respite and permanent residents who were present in a RACF between January 1, 2021, and December 31, 2021. The dashboard was designed using participant data that was extracted onto Excel sheets from the provider. After the development, the dashboard was connected to the provider's live data sources through a data lake.

Ethical Considerations

All methods and analyses in this study were conducted in accordance with the principles of the Declaration of Helsinki. The study received ethics approval from the Macquarie University Human Research Ethics Committee (52019614412614). As a retrospective cohort study using deidentified, routinely collected aged care data, a waiver of informed consent was granted by the committee, in line with the Australian National Statement on Ethical Conduct in Human Research (Sect. 3.3.14(a)), which allows for the use of previously collected deidentified data.

Dashboard Development Process

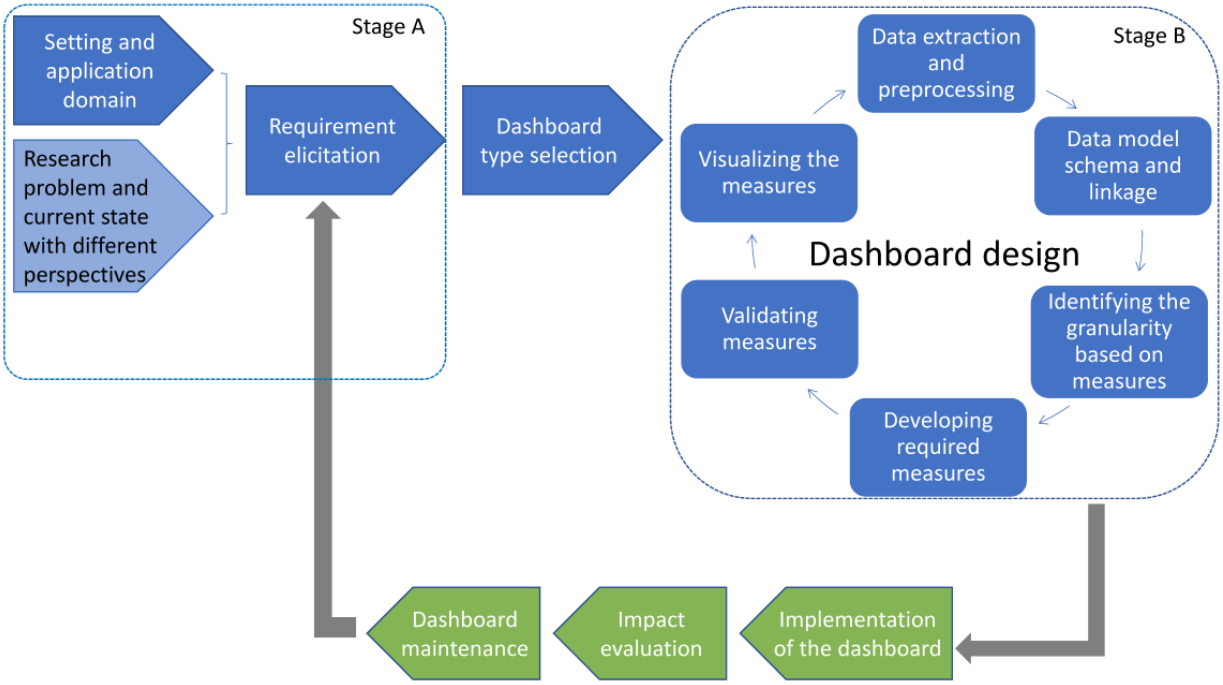
Overview

We followed an iterative process to co-design the dashboard, with intended dashboard end users (ie, RACF staff, general practitioners, and consumers at RACFs), along with domain experts in the field of falls prevention and management. Figure 1 highlights the dashboard development process followed during the study. The stakeholders were involved continuously in Stage A and Stage B through regular discussions, interviews, and

XSL•FO
RenderX

workshops, which will be described in detail elsewhere. The steps colored in blue are described in this paper and the final steps relating to implementation and evaluation colored in green are in progress and will be reported in subsequent papers.

Figure 1. The dashboard development process.



Setting and Application Domain

Domain knowledge can be defined as the knowledge about the environment in which the proposed system will operate [13]. Domain knowledge was obtained from the aged care provider through discussions and analysis of data related to falls incidents and the Peninsula Health Falls Risk Assessment Tool (PH-FRAT) risk assessments (ie, to determine the current falls rate across facilities, the characteristics of falls incidents and contributing factors listed in falls incidents). PH-FRAT was developed in 1999 by Peninsula Health in Victoria, Australia, and is a validated and user-friendly tool for screening and assessing fall risk, as well as managing strategies to reduce this risk [14]. Information about predictors of falls and strategies used to prevent falls in the RACFs was obtained by systematically reviewing published literature on predictive models for fall prevention, and the effectiveness of fall prevention interventions [15,16]. The identified predictors of falls in older people have been incorporated into the development of the embedded predictive model, which is reported elsewhere [17].

Research Problem and Current State With Different Perspectives

Considering the suboptimal predictive performance of current tools [6], there is a need for a dashboard embedded with a dynamic predictive model and decision support (research problem). Moreover, falls prevention and management requires action at different levels of an organization's hierarchy from management to clinical staff. Therefore, it was important to gather different perspectives. For this purpose, a series of discussions were held between researchers and RACF stakeholders across a 2-stage process (Figure 1). In Stage A discussions, different perspectives on falls prevention and management in RACFs were gathered while identifying the dashboard stakeholders and their information requirements. In Stage B, dashboard stakeholders engaged in rounds of iterative feedback workshops to support the dashboard designing process. A qualitative analysis of these interviews will be reported elsewhere.

Requirement Elicitation

The requirement elicitation stage was based on three fundamental tasks: (1) identifying the individual stakeholders; (2) eliciting stakeholder requirements depending on the role of the stakeholder; and (3) integrating, refining, and organizing the requirements identified [18].

In Stage A, baseline interviews were conducted to understand the current utilization of health care data and information by potential dashboard end users. The aim was to identify the information requirements essential for enhancing decision-making in a clinical dashboard. In this process, the individual stakeholders and information providers of the dashboard were identified along with their high-level user expectations [10]. This also facilitated a precise understanding of the individuals with different perspectives of falls prevention and management, who would be accessing each dashboard view. The information collected from these interviews was used to define measures that should be generated in the dashboard backend.

From these identified measures, the designers then evaluated the variability of information and organized the information requirements based on the level of reporting in the dashboard

XSL•FO
RenderX

(ie, resident level, facility level, or organizational level). This further supported the dashboard design stage in arranging visual plots and tables within views.

Dashboard Type Selection

Power BI (PBI; Microsoft Corp) was selected on the grounds of digital data visualization, easy implementation at the user end, and stable R programming integration for advanced statistical analysis and plots. In this step, the initial prototype of the dashboard was created and shared with our partner organization and refined through collaborative efforts.

Dashboard Design

Overview

The aim of this step was to meet the user requirements for preventing and managing falls by effectively presenting resident data and fall risk information to support informed decision-making. As in Figure 1, this step included an iterative process that included data extraction, preprocessing, and linkage along with identification of the granularity of data, development, validation, and visualization of the measures on the front end.

Data Extraction and Preprocessing

Access to four datasets was obtained: the resident profile (includes residents' demographics and admission information); all medications administered to residents; the organization's PH-FRATs (this aged care providers used the PH-FRAT) to obtain information related to falls risk assessments; and incident dataset (includes information related to all fall incidents and pressure injuries; Table S1 in Multimedia Appendix 1).

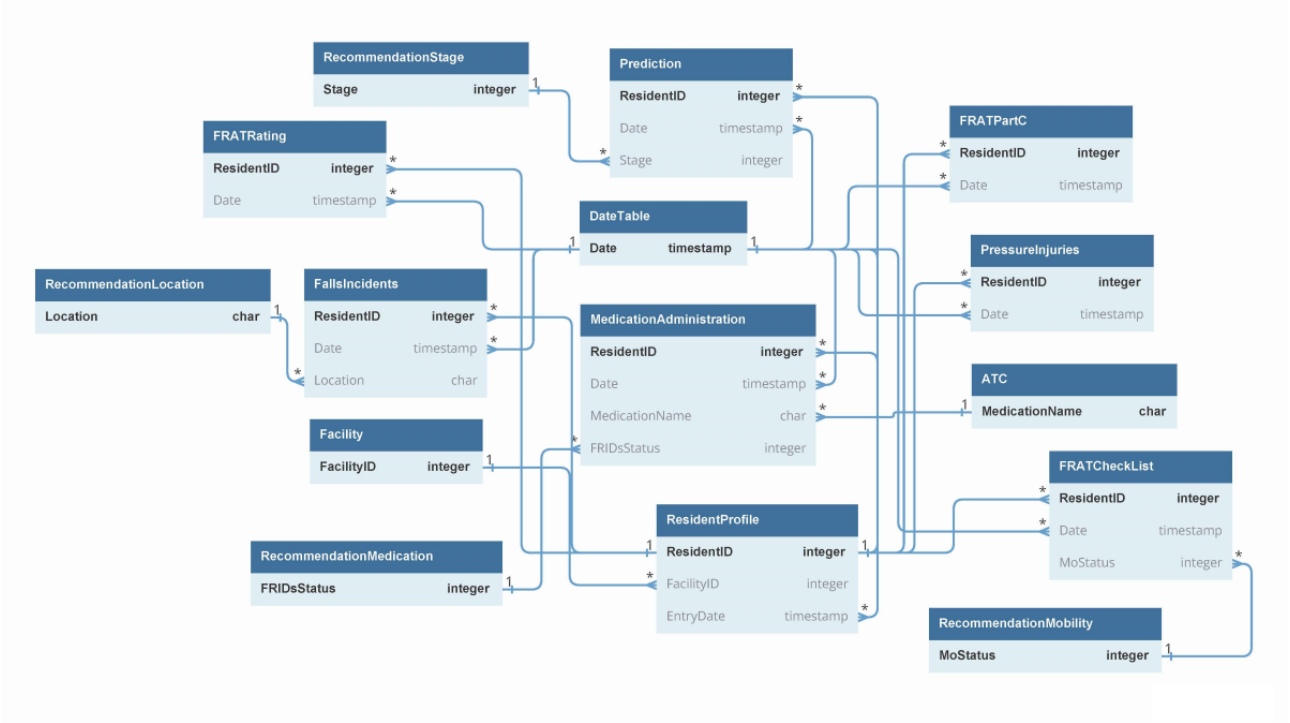
The profile dataset included a free text field that reported the comorbidities along with other special needs of the patient at admission (ie, health status). From this field, comorbidities present at admission were identified using the R-programmed version of the "aged care health status algorithm" [19] within PBI. The algorithm identifies the health conditions using free text fields from EHRs [20]. All medications in the dataset were coded using the Anatomical Therapeutic Chemical codes [21]. These datasets were then linked in the dashboard backend.

The extracted datasets for the study period underwent an external analysis for this study using R programming language (version 4.3.3; R Core Team) and were also subsequently used for its intended purpose within PBI. A descriptive analysis of the admission-related information from the resident profile dataset is reported appropriately (Table S2 in Multimedia Appendix 1). Resident characteristics from the resident profile dataset, FRAT dataset, and daily medication administration data are recorded in Table S3 in Multimedia Appendix 1.

Data Model Schema and Linkage

A star schema was adopted as much as possible where the fact table is centered and surrounded by the multiple dimension tables. The resident profile dataset was used as the main fact table that contains quantitative information (facts) about the residents [22]. Other datasets were used as dimension tables, which describe the events that residents experience during their stay [22]. The data tables were then linked as presented in the simplified form of the database schema as shown in Figure 2. The real schema on the backend of the dashboard is an extension of this and is more advanced.

Figure 2. Simplified view of the data model schema. ATC: Anatomical Therapeutic Chemical code; FRAT: falls risk assessment tool.



Identifying Granularity Based on Measures

To determine the granularity of the data tables, we considered the dashboard end users, information requirements (measures), and level of reporting that were preidentified. This step was done in parallel or a back-and-forth manner with the data model schema (ie, preceding step) and development of required measures (ie, proceeding step). A higher granularity was

XSL•FO
RenderX

preferred for most of the datasets to maintain sustainable analytical power to answer new user requirements within the designing stage [23]. Subsequently, the data tables were transformed to simply calculate the measures, or measure calculations were adjusted in a way to meet the form and granularity in the extracted tables. In certain instances, highly granular data tables were summarized to a lower-level granularity to support the prediction model and avoid intricate calculations during measure development.

Developing the Required Measures

In this step, measures were calculated on the transformed tables using Data Analysis Expression (DAX). Measures related to some figures such as funnel plots were created using embedded R programming in the PBI backend.

Validating the Measures

To ensure the accuracy and reliability of the information presented, the calculated measures were validated with provider-extracted datasets. The user input filters on each view were also tested during the validation process to identify whether there were any errors. If errors were identified in the data, remedial action was taken after a thorough examination of the underlying causes, in data extraction, data linkage, data preprocessing, or DAX formula.

Visualizing the Measures

The calculated measures were appropriately visualized using statistical plots, tables, and responsive figures on the dashboard, which comprises several dashboard views intended for various users at different levels in the hierarchy. The above-mentioned steps within the dashboard design stage were enclosed in a co-design process as explained below.

Dashboard Co-Design and Work Process Features (Stage B)

The core constructs of the integrated Promoting Action on Research Implementation in Health Services [24] framework were considered to facilitate innovation in the development of the dashboard: (1) innovation refers to the alignment of new insights including the predictive risk of falls captured by the dashboard with the current priorities and practices for falls prevention at RACFs, (2) recipients refer to the RACFs staff and other external stakeholders who will be affected by and influence the implementation of the dashboard, (3) context refers to the RACFs including the organization (inner layer) and wider health system (outer layer) around falls prevention and management, and (4) facilitation refers to the process that activates the implementation through assessing the characteristics of the three previously mentioned constructs.

The core constructs were integrated into the dashboard development process and the design was meticulously refined through a series of workshops (ie, facilitation) involving diverse teams of experts from various layers. These teams included general practitioners, geriatricians, allied health staff from RACFs (including nurses and clinical care managers), and human-computer interaction experts (ie, recipients). The collaborative workshops aimed to identify and implement enhancements to the dashboard, ensuring that it caters to the needs and preferences of all relevant stakeholders (ie, innovation). The design features (eg, color, graph types, positioning, filters, slices, and data views) and content for decision-making (ie, information included or missing on the dashboard and predictive model output) were considered. In these workshops, the development stage prototypes were presented to gather information on decision-making on client care and to identify the dashboard views that required improvements in managing falls. Current priorities and practices involved in managing falls at RACFs (ie, context) were also identified.

Descriptive and Diagnostic Analytics: Measures, Indicators, and Statistical Tests

All falls, injurious falls, and falls requiring hospitalization were identified as the main indicators in the dashboard. The total number of falls, the crude incident rate (ie, falls rate per 1000 resident days), the percentage of injurious falls, and the percentage of hospitalization were calculated for each outcome measure to enhance the decision support capability at an organizational level.

Funnel plots were used to visualize crude incident rates with 95% and 99% control limits, relative to the number of admissions. A time series plot was used to visualize the percentage of injurious falls and hospitalization due to falls.

Predictive Analytics With Resident Tailored Recommendations: Dynamic Falls Predictive and Monitoring Model

A dynamic predictive tool was developed to classify each resident's risk of falls using a technique called stratified landmarking and reported in detail elsewhere [17]. Initially, 116 variables were screened, and the most relevant variables on the fall outcome were selected using Collett's [25] variable selection approach. Two separate models based on dementia status were developed [17]. The final two models were based on patient demographics, dementia status, fall risk–increasing comorbidities such as cerebrovascular accident, visual impairment, osteoporosis or fracture, medication use, falls history, other risk factors such as psychological status, mobility status, behavior status, and activities of daily living [17].

The predictive model used in this study was developed and internally validated using retrospective data from the same provider [17]. The scores based on model coefficients and baseline hazards were then embedded into the dashboard backend and calculated the resident's probability of falling in a near real-time manner. The predictions generated by the model are updated every 24 hours given the modifiable risk factors such as medication use, previous fall incidents, and comorbidities.

As mentioned in the introduction, relying solely on a predictive model will not improve decision-making for developing a personalized care strategy for fall prevention and management. Additionally, the dashboard that visualizes risk factors at both the resident and facility levels should be used to enhance the interpretability of these predictions and other associated risk factors of falls. This dashboard visualizes the risk factors for falls, along with changes in risk and risk stages derived from

XSL•FO
RenderX

the predictive model. Furthermore, to offer a more prescriptive approach based on these predictions, evidence-based, resident-tailored recommendations have been developed as follows.

Resident-tailored recommendations are generated using a rule-based system in the dashboard's backend. The system identifies residents' mobility issues from PH-FRAT data and detects the use of fall risk–increasing drugs (FRIDs) from daily medication records. Fall incident data are used to determine the most frequent fall locations, while a predictive model assesses the current fall risk. Based on these factors, rule-based recommendations [26-28] are created and displayed in the resident fall view (Table S4 in Multimedia Appendix 1). Additionally, strategies identified through clinical judgment in the PH-FRATs are also displayed to enhance the personalization.

Results

Descriptive Profile of the Aged Care Provider, Facilities, and Residents

The aged care provider managed 24 RACFs. During the period of dashboard development, these RACFs housed 3686 resident admissions from 3057 unique residents. The median facility size was 124 (IQR 86-160) residents. Of the resident admissions, 71% (n=2622) were permanent. The median length of stay was 373 (IQR 41-1266) days (Tables S1 and S2 in Multimedia Appendix 1).

Dashboard Evolution Within the Design Stage

The Stage A interviews and Stage B workshops identified themes and findings related to content, design, functionality, and decision support. This led the dashboard to evolve through a series of iterations (Multimedia Appendix 2). Two dashboard views namely, organizational falls and facility falls were designed for organizational managers, while another two views were designed for care staff and focused on residents, resident falls, and resident-level medications in detail. The intervention recommendations gathered on fall prevention and management from published randomized controlled trials, government reports, systematic reviews, and meta-analyses [26-28] are included in the recommendation panel for education purposes.

Measures Developed in Dashboard Views

Textbox 1 presents the measures developed for each dashboard view. Given the interrelatedness of these views, certain measures used at the organizational level (such as total falls and falls rate [per 1000 resident days]) were also included at the facility level, while certain measures used at the resident level (ie, FRIDs use in last three days and polypharmacy indicator) were included in the medications in detail view.

XSL•FO
RenderX

Textbox 1. Measures available in the dashboard.

Organizational falls

- Total falls
- Total falls requiring hospitalization
- Total injurious falls
- Falls rate (per 1000 resident days)
- Falls requiring hospitalization rate (per 1000 resident days)
- Injurious falls rate (per 1000 resident days)
- Number of admissions per selected period
- Resident days per selected period

Resident falls

- Current falls risk
- Current risk category
- Daily percentage change in risk
- Falls in last 6 months
- Fall risk–increasing drugs (FRIDs) use in the last three days
- Change in FRIDs use in last three days
- Polypharmacy indicator
- Antipsychotic indicator
- Recent falls risk assessment tool (FRAT) assessment date
- Recent cognitive status identified by Peninsula Health-FRAT
- Recent psychological status

Facility falls

- Total number of residents
- Residents having stage 1 pressure injury
- Residents having stage 2 pressure injury
- Residents having stage 3 pressure injury
- Residents having stage 4 pressure injury
- Residents having unstageable pressure injuries
- Residents having deep tissue pressure injuries
- Residents with a pressure injury

Medications in detail

- Polypharmacy indicator – without pro re nata and short course
- Daily FRIDs indicator
- Daily change in FRIDs
- Daily change in opioid use
- Daily change in anxiolytics
- Daily change in hypnotic and sedatives
- Daily change in beta blocking agents
- Daily change in vasodilators
- Daily change in antidepressants

XSL•FO
RenderX

Dashboard Architecture

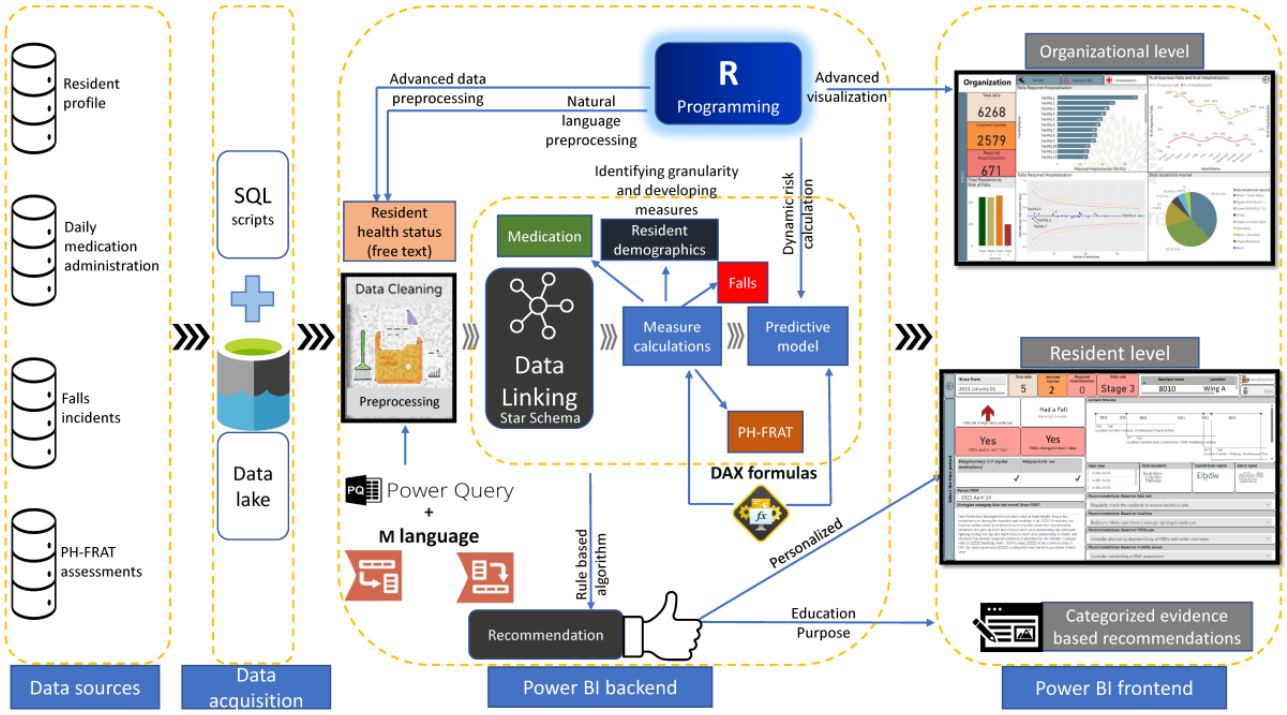
Data flows through 4 layers, to formulate the necessary information (Figure 3).

The first layer includes the data sources where the aged care provider collects and stores data. The next layer is the data acquisition layer. It includes data extraction scripts such as SQL scripts, which collect data from the previous layer. As the dashboard was developed without a direct connection to the above data silos (sources), the data extractions at the beginning were done using SQL scripts for each data silo, including the original variable names and formats. This layer is also helpful in maintaining the reliability of the dashboard when underlying IT systems change by storing the original form of data that is compatible with the dashboard backend, which is the next layer. At implementation, this layer will be replaced with a data lake that feeds provider-level data to the dashboard backend.

Then the data extracted from the data lake are preprocessed and linked at the backend of the dashboard using the data model schema discussed (Figure 2). Previously identified measures and indicators are developed in this layer, to support the visualization of data at the front end of the dashboard. For this purpose, we used power query M language, DAX, and R programming. Risk calculations of the prediction model were carried out in this stage using R programming.

In the fourth layer, the information requirement of stakeholders is satisfied by visualization and reporting. The front end of the dashboard includes five views namely organizational falls, facility falls, resident falls, medications in detail, and recommendation panel. The filters are applied to various dashboard views depending on user needs as identified during the requirement elicitation and dashboard design stages. Moreover, statistical plots are used to enhance valid comparisons of indicators between facilities and to support decision-making.

Figure 3. The dashboard architecture with four layers. DAX: Data Analysis Expression; PH-FRAT: Peninsula Health falls risk assessment tool.



Dashboard View: Organizational Falls and Facility Falls

The organizational falls and facility falls are accessible by organizational and facility managers. Please note that some information about these views has been removed or modified to deidentify the residents and facility details.

The organizational falls view (Figure 4) includes organizational and facility-level information relevant to fall incidence. The time period and facility names are included as the filters. The information on this view can be updated with the values selected for these filters. The total number of falls along with injuries, and falls that required hospitalization are reported at the organizational level. The funnel plot in the view visualizes the crude fall rate. Users can visualize this plot for the three outcome measures namely all falls, injurious falls, and falls-related hospitalization. Funnel plots with the crude incident rates for each outcome measure enable valid comparisons between the facilities, as it can identify the outlying facilities with higher fall rates related to the number of admissions during the study period. The time series plot with a percentage of injurious falls and a percentage of hospitalization enables the user to investigate overall patterns of falls occurring in selected facilities or at an organizational level. This feature is helpful in identifying the longitudinal trend of injurious falls and hospitalization due to falls in facilities. The time series plot with the most frequent time of falls enables the user to investigate the time pattern of fall incidents. The most frequent locations of falls are linked with the time of falls by a Sankey plot. These plots are dynamic and responsive to the facility selected from the filters. Furthermore, the table with fall risk stages indicates the number of residents within each risk classification predicted by the

XSL•FO
RenderX

embedded predictive model. The most frequent injured body locations are also reported in a table.

The facility falls view (Figure 5) provides more details of the facility-specific falls along with facility characteristics. This view can also be accessed through the organizational level view using the "drill through" functionality in PBI. This view highlights the facility-specific fall rate per 1000 resident days for a specified time. To enhance the interpretability of the predictive model, resident-specific risk factors extracted from the model are reported in this view. Importantly, residents classified under each risk classification along with changes in fall risk compared to the previous day can be accessed in this view for effective fall prevention and management. The view also identifies comorbidities that increase fall risk (eg, dementia) and modifiable risk factors (eg, recent falls within the last six months, frequent locations and times of falls, mobility issues, psychological status, and the use of antipsychotic, opioid, and analgesic medications), all color-coded based on the level of risk identified by the predictive model. Additionally, facility-specific monthly indicators including four national mandatory quality indicators namely the percentage of care recipients: who experienced one or more falls; who were prescribed nine or more medications; who received antipsychotic medications; and with pressure injuries are also reported in this view. This view will also be included in a daily email notification to users once the dashboard is implemented.

Figure 4. Organizational falls view.

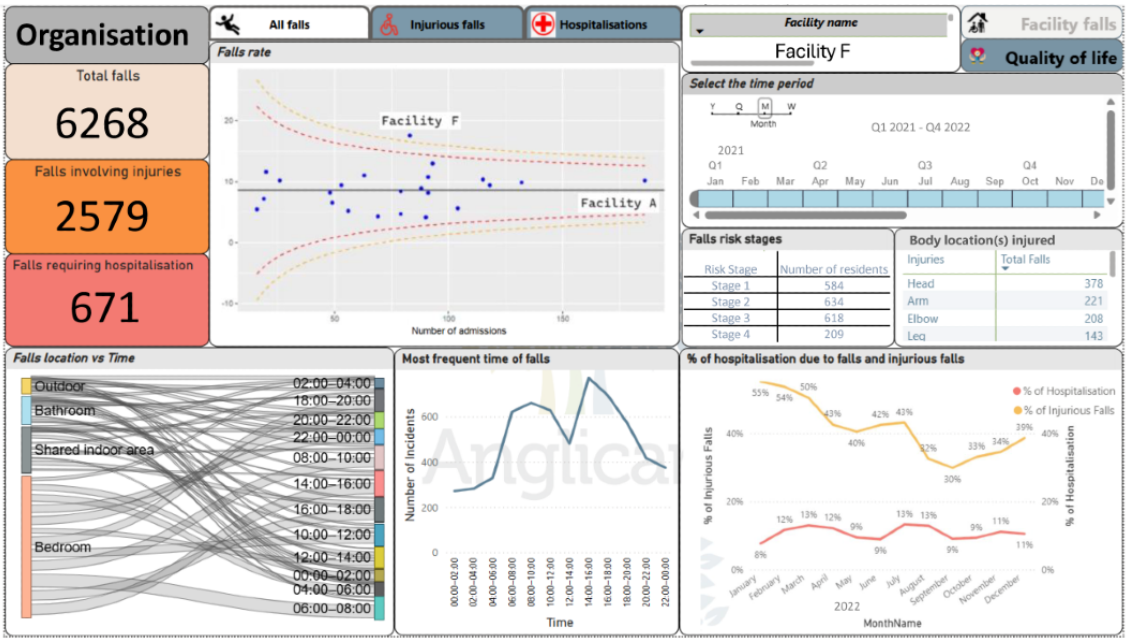
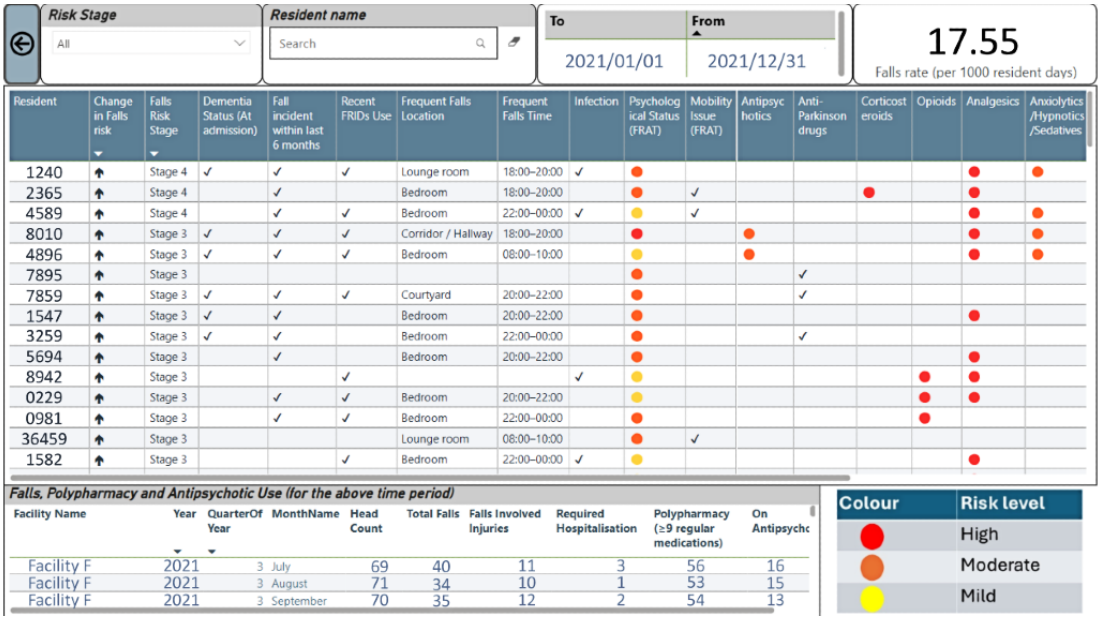


Figure 5. Facility falls view. FRAT: falls risk assessment tool.



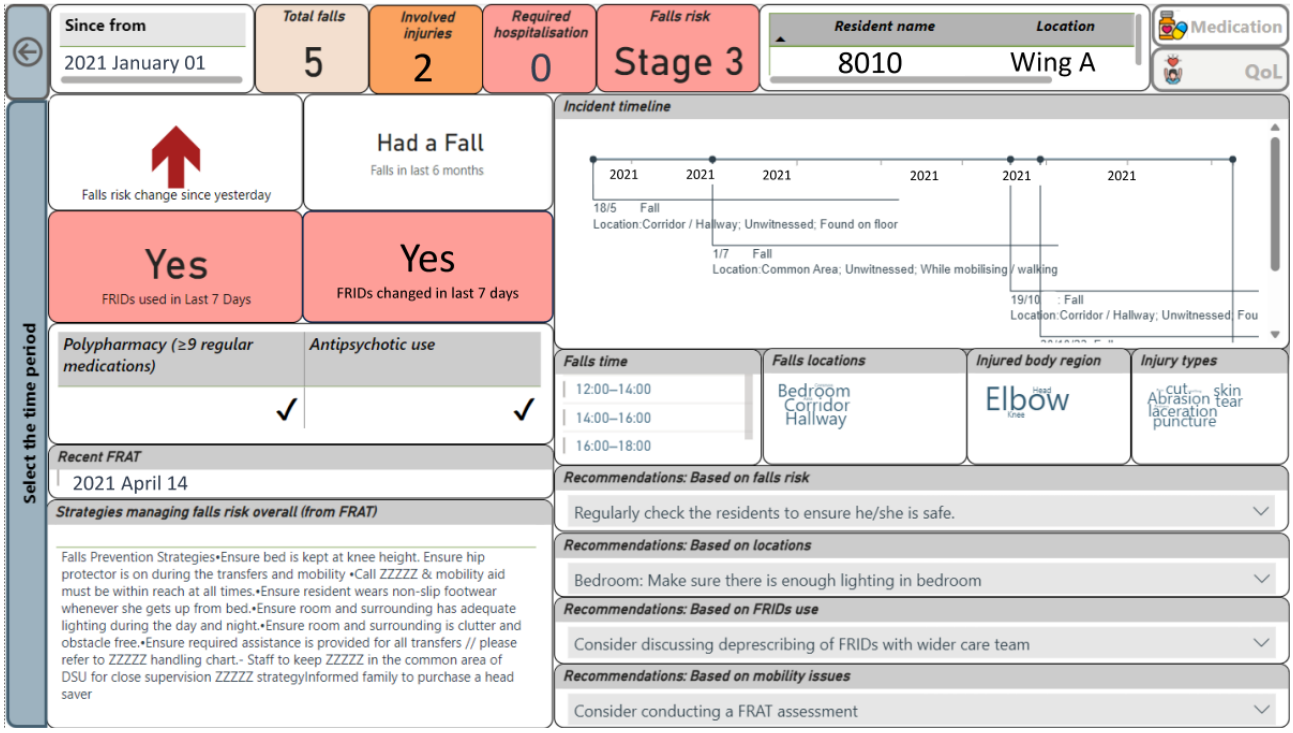
XSL•FO
RenderX

Dashboard View: Resident Falls and the Detailed Medications

Care staff will have access to the resident falls view, which includes the time period, facility name, falls risk classification (ie, derived from the predictive model), and resident identifier as filters.

The resident falls view focuses on the resident-specific falls information along with recommendations to prevent falls (Figure 6). This can also be accessed using the facility falls view using drill-through functionality in PBI. The polypharmacy, FRIDs, and antipsychotic medication use are identified using the resident's daily medication administration data. This view presents the current risk category along with the total number of falls faced by the resident during their stay. Risk classification derived from the predictive model is shown using color codes (ie, green to red) and text (ie, stages 1 to 4). Most importantly, the number of falls faced within the last 6 months is shown in this view. Patient-specific fall history is also visualized using a timeline along with the details of the events (ie, event type, injury status, and witnessed or unwitnessed). The sequence and the gaps between the fall incidents are highlighted in this plot. Furthermore, the word clouds visualize the injured body regions, injury types, locations, and time of the falls. At the bottom of the view, personalized recommendations are provided to the resident, which aligns with falls risk categorization, falls history, medication use, and mobility issues.

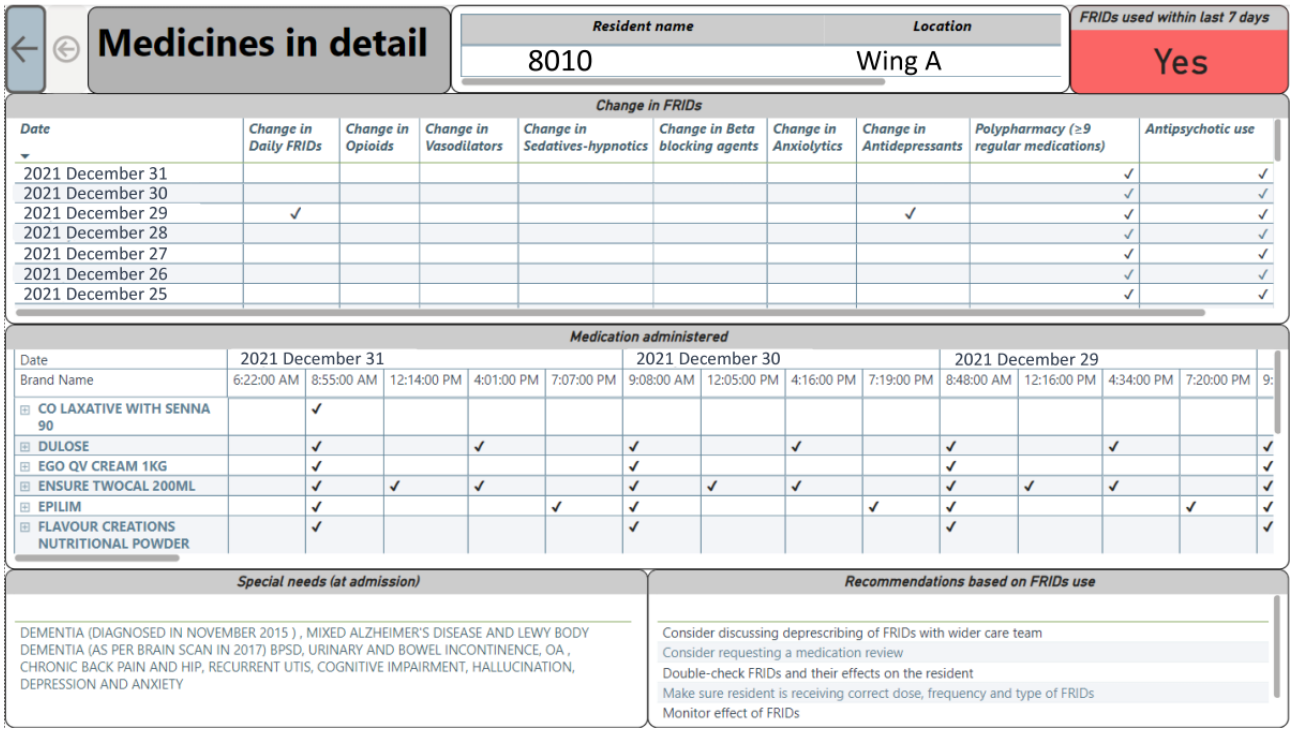
Figure 6. Resident falls view. FRAT: falls risk assessment tool; FRID: fall risk–increasing drug.



The medications in detail view (Figure 7) focuses on the resident-specific daily medication use. This view can be assessed using the "drill through" functionality enabled in the resident falls view. Here, changes in FRIDs (ie, whether the resident has previously used the drug, whether a medication was omitted, or a new drug added to the medication profile) within the last three days are identified and displayed, along with polypharmacy and antipsychotic medication indicators. The resident's use of FRIDs is highlighted at the top. Daily medication administration records can also be accessed using this view. The special needs (ie, comorbidities and allergies) identified at the admission are included in the view to highlight the comorbidities, special needs, or risk factors of the resident.

XSL•FO
RenderX

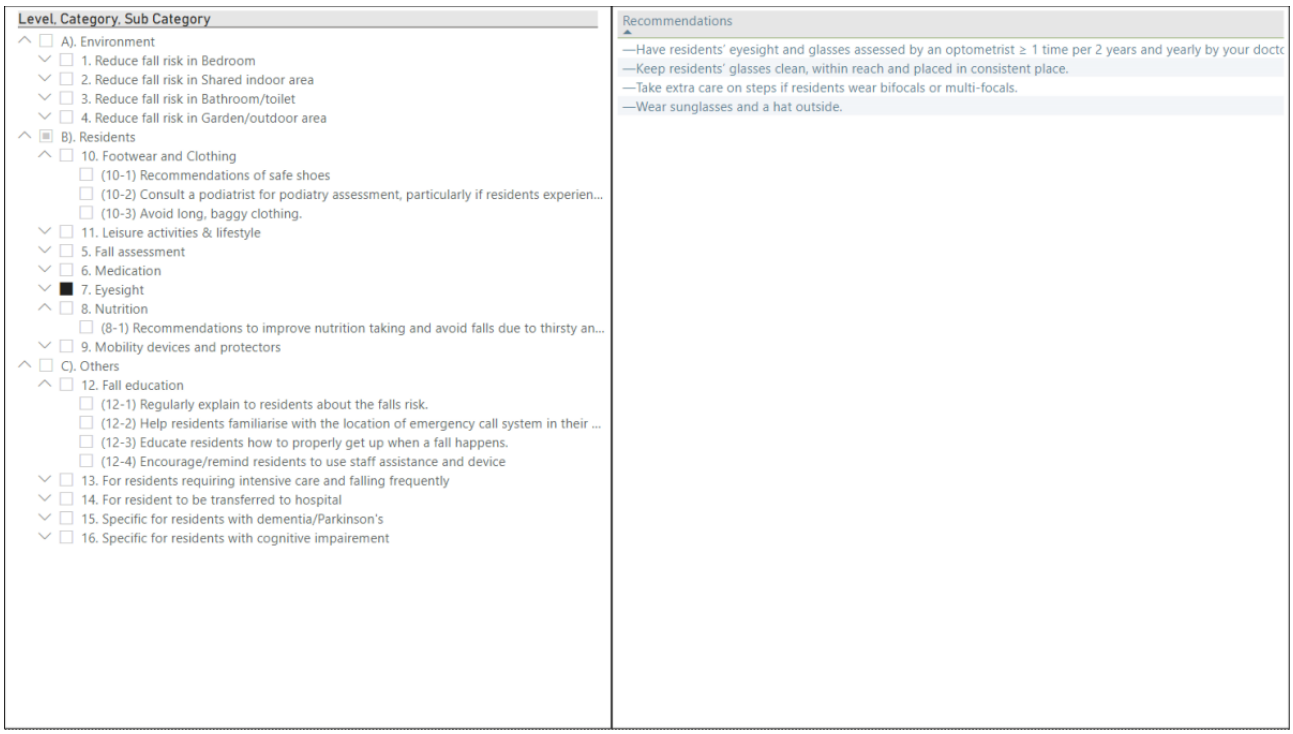
Figure 7. Medicines in detail. FRID: fall risk–increasing drug.



Dashboard View: Recommendation Panel

As falls are affected by intrinsic and extrinsic factors [29], recommendations are grouped and reported in the recommendations panel (Figure 8) in three groups namely: resident (intrinsic), environment (extrinsic), and others followed by subcategorizations.

Figure 8. Recommendation panel.



Illustrative Use Scenario

Organizational Managers

A total of 6268 fall incidents were reported from the 3057 unique residents (729,410 resident days; Table S3 in Multimedia Appendix 1) between January 1, 2021, and December 31, 2021 (Figure 4). Of these, 2579 (41%) falls were injurious and 671 (11%) falls required hospitalization. The average fall rate was 8.36 per 1000 resident days for the study period (funnel plot). The percentage of injurious falls shows an overall gradual decline from 56% to 38%, whereas the percentage of falls requiring hospitalization is stable between 7% and 13% between

XSL•FO
RenderX

January 1, 2021, and December 31, 2021. The majority of falls occurred in residents' rooms. Injurious falls most often involved injuries to the head and face followed by upper limbs. When considering all falls, Facility A reported the highest number of falls. Facility F had the sixth-highest number of falls. However, the rate was above the 99% upper control limit (funnel plot), indicating a higher fall rate (per 1000 resident days) compared to other facilities with a similar number of admissions. Therefore, organizational managers can prioritize Facility F in intervening. They also can select Facility F and drill through to the facility falls view for a more detailed investigation.

Facility Managers

A total of 83 admissions were reported for facility F during the study period, with a fall rate of 17.55 per 1000 resident days (Figure 5). As of December 31, 2021, 3 (4%) residents were categorized with a fall risk of stage 4, while 13 (18%) residents were categorized with a risk of stage 3. Most of these residents had experienced a fall within the last six months and were on FRIDs. The resident bedroom was identified as the most frequent location for falls among the majority of residents. The majority of residents in this facility had cognitive impairment and dementia. It was clear that 54 (77%) and 13 (18%) residents were on polypharmacy and antipsychotics, respectively, for September 2021. Organizational managers and facility managers can use the dashboard information to identify facility-level risk factors and assist them with mandatory quality indicator reporting.

Nurse or Personal Carer

On the facility falls view, a stage 3 resident (ie, resident: 8010) was selected and drilled down to the resident falls view to analyze the resident-level risk factors, fall incident characteristics, and personalized recommendations to prevent and manage future fall risk (Figures 5 and 6). Risk factors: resident 8010 was in stage 3 (as of December 31, 2021) with delirium, dementia, and stroke (Figure 7). The resident was on polypharmacy and antipsychotic medications as of December 30, 2021. There was an increase in fall risk compared to December 30, 2021, which was generated by the embedded predictive model. The resident had five falls during the study period, of which, two were injurious. When considering the resident's incidents timeline, four fall incidents had occurred in the last six months. The resident was on polypharmacy along with FRIDs (as of December 31, 2021) where he had a change in FRIDs in the last three days (ie, December 29-31, 2021). Characteristics of fall incidents: most of these incidents occurred during the afternoon in the resident's bedroom including the hallway and caused injuries to the elbow, head, and knee. The most recent FRAT assessment was conducted in April 2021. Recommendations to prevent future falls: according to the recent FRAT, the identified strategies in managing fall risk were displayed along with prescriptive recommendations such as "regular monitoring," "assessing environmental risk in toilet or bathroom and bedroom," and "assessing FRIDs use."

For a further detailed view of medication use, the resident's medications in detail view (Figure 7) can be assessed. Moreover, carers can use the information provided in the recommendation panel when assessing the fall risk for the resident, based on the resident-level risk factors (eg, dementia), environmental risk factors (ie, bedroom), etc.

Discussion

Principal Findings

Our dashboard adapted and refined the dashboard development process outlined by Staron [10], with an emphasis on understanding the specific domain of residential aged care and their current work practices around fall prevention and management. This was achieved through continuous stakeholder discussions (ie, Stages A and B), quantitative data analysis on the extracted datasets, and reviewing the published literature. This work produces important new resources for the current state of falls prevention in RACFs along with the extent of evidence-based prevention strategies. The dashboard design stage was an iterative process that required close collaboration with the users and its stakeholders (ie, Stage B). This was the most important stage of the dashboard development as it identified the data presentation for dashboard views along with the quantitative measures to be included for fall prevention. To enhance the technology implementation process within the context of fall prevention and management, the four constructs of the integrated Promoting Action on Research Implementation in Health Services framework were followed throughout the dashboard development process.

The core innovation of this study comes from several novel aspects. The Royal Commission into Aged Care Quality and Safety has noted that the residential aged care sector has lagged behind other health care sectors in adopting and implementing technology [30]. It is also found that information systems and processes in these settings are both underdeveloped and inadequately integrated [31]. Furthermore, successfully connecting multiple disparate data sources in aged care is uncommon. This may be due to various factors, such as data being collected by different systems and third-party software vendors being inflexible in incorporating unique identifiers (ie, resident identifiers) across systems. Therefore, this study emphasizes the value of linking existing datasets and visualizing the linked information to support falls prevention. The resident falls view in the dashboard presents the linked data from four data sources and supports to identify resident-level risk factors, fall incident characteristics, and personalized recommendations for fall prevention. This allows users to access crucial information at a glance, rather than navigating across different systems, thus demonstrating significant added value in preventing falls.

The dashboard provides staff with easy access to synthesized, evidence-based recommendations for effective fall prevention strategies. In the dashboard, external evidence on fall prevention and management (eg, published prevention recommendations) has been aligned with current priorities and practices at RACFs (such as PH-FRATs). Current practices, such as the fall prevention strategies identified through PH-FRATs, were embedded, along with rule-based recommendations. Additionally, a separate recommendation panel was established to support decision-making around fall risk assessments.

XSL•FO
RenderX

Another novel aspect of the dashboard is the embedded predictive model. Predicting fall risk is challenging because of the numerous contributing factors [32,33], including medical conditions, medications, functional status, behavior, physiological aspects, and environmental conditions. Many of these factors, such as medication use, can vary over time [34]. The dynamic predictive model embedded in the dashboard is therefore an important part of the dashboard. It uses near real-time data and helps to classify the resident risk of falls on a daily basis. The incorporation of the predictive algorithm in the dashboard goes beyond what has been possible with existing fall risk assessments as it dynamically predicts the risk and links it to residents' fall risk factors and personalized strategies. This provides timely and meaningful risk prediction along with evidence-based recommendations based on resident profile, to the caring staff to facilitate person-centered fall prevention and improve outcomes.

Finally, the dashboard is innovative because it adopts a holistic approach to managing falls in aged care residents by integrating multiple data sources and making this information easily accessible to different stakeholders including managers and carers. The granular resident-level data are aggregated to the facility and organizational levels, allowing users to navigate seamlessly between these levels with just a few clicks. This structure enables users to drill down from the organizational level to individual residents or vice versa, providing a comprehensive and detailed view of fall risk and prevention strategies.

Challenges Faced and Strategies to Overcome

There were many challenges faced when developing this dashboard. Initially, the dashboard was developed using datasets that were extracted from the provider rather than directly connecting to the database. This leads to the integration process becoming more challenging unless the variables are in standard format as in the database. Moreover, our partner has changed its incident reporting system and daily medication management system within the development process. More robust processes were required to be implemented on the backend of the dashboard to overcome these challenges. The data acquisition layer discussed in dashboard architecture was important in this aspect. This layer acts as a data reservoir. Here, the initial data are extracted from original sources, stored in the original format, and then transformed into the data tables that the dashboard requires. This layer enabled the dashboard to be sustainable in the long run amid the underlying system changes.

The data model in the backend of the dashboard also plays a critical role in facilitating the above process. This is crucial as the measures were developed using the linkages between the tables. Adapting a star schema as much as possible within the dashboard backend has promoted several advantages within the study. Simplicity and ease of use of the data model during various stages of the development process can be noted. Visualization and embedding of analytical measures within the dashboard were simplified without complex data integration. This has streamlined the dashboard development during the co-design stage while accelerating the time of incorporating various views on fall prevention and management. Flexibility and scalability can be noted as another advantage where the schema was helpful in adapting the dashboard design to changing RACF requirements in managing falls while incorporating additional data fields and sources with minimum modifications to the existing structure. Therefore, overall, adapting the star schema as much as possible within the dashboard development process was an excellent choice due to its ability to incorporate changing user requirements in the context of fall management at RACFs.

Furthermore, for the daily calculation of changes in fall risk, this study uses R programming. This decision stems from PBI's inability to retain and save daily risk calculations. To address this limitation, the designers have integrated R programming code into the Power Query Editor, enabling the storage of these risk calculations on a shared drive. The code automatically saves these calculations on a daily basis during the data refresh cycle, occurring every 24 hours. Subsequently, on a daily basis, these stored files (ie, previous-day risk calculations) are retrieved into the PBI backend to calculate the difference between current and previous risks at the resident level.

The integration of R programming into the dashboard backend has introduced a technical skills gap between the partner organization and the dashboard designers. Consequently, workshops were organized with the partner organization's IT teams to provide training on R programming, bridging the gap in statistical analysis using R programming. Additionally, line-by-line code interpretation was incorporated into the embedded codes to facilitate a seamless transition during dashboard implementation.

Implications for Policy, Practice, and Future Research

This study highlights the potential of using a fall prevention and management dashboard to enhance resident outcomes in aged care through the effective use of information systems. Health dashboards facilitate decision-making and are required to be used in conjunction with clinical judgment, staff education, training, and individualized tailored responses. Future policy and practice must ensure the interoperability of information systems in aged care, involve key stakeholders in development and implementation, and allocate appropriate funding for these processes. Future research should also focus on expanding the fall prediction dashboard and incorporating other care needs in residential aged-care facilities, such as hospitalization and wound management.

Conclusions

In this study, a comprehensive process for developing a fall prevention and management dashboard with descriptive, predictive, and prescriptive analytics for RACFs was thoroughly discussed along with a sustainable architecture. The key focus was to provide a holistic approach to managing falls by linking existing data sources at the provider level. This highly applied and translational research demonstrated how EHRs can be used along with technology to deliver evidence-based information to aged care allied health workers and managers, to drive better aged care quality and safety.

XSL•FO
RenderX

Acknowledgments

We thank our partners and collaborators including Anglicare, Northern Sydney Local Health District, Sydney North Primary Health Network, the Deeble Institute for Health Policy Research, and the Australian Aged Care Quality and Safety Commission.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to the privacy of the individuals who participated in the study but are available from the corresponding author upon reasonable request.

Authors' Contributions

The dashboard was technically designed by SSMS. The co-design process of the dashboard involved SSMS, NW, ADN, KS, LD, IM, CIM, and GH. The manuscript was drafted by SSMS and revised by NW, ADN, KS, GH, and IM. Both the manuscript draft and system were reviewed by JIW.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables.
[DOCX File , 28 KB - aging_v8i1e63609_app1.docx]

Multimedia Appendix 2
Dashboard evolution within the design stage.
[PNG File , 534 KB - aging_v8i1e63609_app2.png]

References

1. Falls. World Health Organization. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/falls [accessed 2022-11-25]
2. Falls and fractures: effective interventions in health and social care. Department of Health. 2009. URL: https://laterlifetraining.co.uk/wp-content/uploads/2011/12/FF_Effective-Interventions-in-health-and-social-care.pdf [accessed 2024-05-05]
3. Ward JA, Harden M, Gibson RE, Byles JE. A cluster randomised controlled trial to prevent injury due to falls in a residential aged care population. Med J Aust 2010;192(6):319-322. [doi: 10.5694/j.1326-5377.2010.tb03530.x] [Medline: 20230348]
4. Waldron N, Hill AM, Barker A. Falls prevention in older adults—assessment and management. Aust Fam Physician 2012 Dec;41(12):930-935. [doi: 10.3316/informit.998232410954897] [Medline: 23210114]
5. Cameron ID, Dyer SM, Panagoda CE, Murray GR, Hill KD, Cumming RG, et al. Interventions for preventing falls in older people in care facilities and hospitals. Cochrane Database Syst Rev 2018;9(9):CD005465 [FREE Full text] [doi: 10.1002/14651858.CD005465.pub4] [Medline: 30191554]
6. Wabe N, Siette J, Seaman KL, Nguyen AD, Raban MZ, Close JCT, et al. The use and predictive performance of the Peninsula Health Falls Risk Assessment Tool (PH-FRAT) in 25 residential aged care facilities: a retrospective cohort study using routinely collected data. BMC Geriatr 2022;22(1):271 [FREE Full text] [doi: 10.1186/s12877-022-02973-0] [Medline: 35365078]
7. Park SH. Tools for assessing fall risk in the elderly: a systematic review and meta-analysis. Aging Clin Exp Res 2018;30(1):1-16. [doi: 10.1007/s40520-017-0749-0] [Medline: 28374345]
8. Ludlow K, Westbrook J, Jorgensen M, Lind KE, Baysari MT, Gray LC, et al. Co-designing a dashboard of predictive analytics and decision support to drive care quality and client outcomes in aged care: a mixed-method study protocol. BMJ Open 2021 Aug 25;11(8):e048657 [FREE Full text] [doi: 10.1136/bmjopen-2021-048657] [Medline: 34433599]
9. Thapa R, Garikipati A, Shokouhi S, Hurtado M, Barnes G, Hoffman J, et al. Predicting falls in long-term care facilities: machine learning study. JMIR Aging 2022 Apr 01;5(2):e35373 [FREE Full text] [doi: 10.2196/35373] [Medline: 35363146]
10. Staron M. Dashboard development guide—how to build sustainable and useful dashboards to support software development and maintenance. 2015. URL: https://core.ac.uk/download/pdf/43560952.pdf [accessed 2025-03-08]
11. What is the network of patient safety databases? Agency for Healthcare Research and Quality. 2023. URL: https://www.ahrq.gov/npsd/what-is-npsd/index.html [accessed 2025-03-08]
12. Live Stronger. 2023. URL: https://www.livestronger.org.nz/assets/Uploads/falls-and-fractures-indicator-methodology.pdf [accessed 2025-03-08]
13. Hjørland B, Albrechtsen H. Toward a new horizon in information science: domain-analysis. J Am Soc Inf Sci 1995;46(6):400-425. [doi: 10.1002/(sici)1097-4571(199507)46:6<400::aid-asi2>3.0.co;2-y]

XSL•FO
RenderX

14. Falls risk assessment tool (FRAT) instructions. Department of Health & Human Services. 2016. URL: https://training.aacs.com.au/wp-content/uploads/2016/08/Falls-Risk-Assessment-Tool-FRAT.pdf [accessed 2024-05-01]
15. Meulenbroeks I, Mercado C, Gates P, Nguyen A, Seaman K, Wabe N, et al. Effectiveness of fall prevention interventions in residential aged care and community settings: an umbrella review. BMC Geriatr 2024;24(1):75 [FREE Full text] [doi: 10.1186/s12877-023-04624-4] [Medline: 38243175]
16. Seaman K, Ludlow K, Wabe N, Dodds L, Siette J, Nguyen A, et al. The use of predictive fall models for older adults receiving aged care, using routinely collected electronic health record data: a systematic review. BMC Geriatr 2022;22(1):210 [FREE Full text] [doi: 10.1186/s12877-022-02901-2] [Medline: 35291948]
17. Wabe N, Meulenbroeks I, Huang G, Silva S, Gray L, Close J, et al. Development and internal validation of a dynamic fall risk prediction and monitoring tool in aged care using routinely collected electronic health data: a landmarking approach. J Am Med Inform Assoc 2024;31(5):1113-1125 [FREE Full text] [doi: 10.1093/jamia/ocae058] [Medline: 38531675]
18. Pacheco C, Garcia I, Calvo-Manzano JA, Reyes M. Measuring and improving software requirements elicitation in a small-sized software organization: a lightweight implementation of ISO/IEC/IEEE 15939:2017—systems and software engineering—measurement process. Requir Eng 2022;28(2):257-281. [doi: 10.1007/s00766-022-00394-4]
19. Lind KE, Raban MZ, Brett L, Jorgensen ML, Georgiou A, Westbrook JI. Measuring the prevalence of 60 health conditions in older Australians in residential aged care with electronic health records: a retrospective dynamic cohort study. Popul Health Metr 2020;18(1):25 [FREE Full text] [doi: 10.1186/s12963-020-00234-z] [Medline: 33032628]
20. Aged care health status algorithms. Macquarie University, Australia. 2020. URL: https://www.mq.edu.au/research/research-centres-groups-and-facilities/healthy-people/centres/australian-institute-of-health-innovation/Research-Streams/the-aged-care-evaluation-and-research-acer-team/aged-care-health-status-algorithm [accessed 2023-12-01]
21. Anatomical therapeutic chemical (ATC) classification. World Health Organization. URL: https://www.who.int/tools/atc-ddd-toolkit/atc-classification [accessed 2022-01-05]
22. Kimball R, Ross M. The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling. Indianapolis, IN: Wiley; 2013.
23. Deckler G, Powell B, Gordon L. Mastering Microsoft Power BI: Expert Techniques to Create Interactive Insights for Effective Data Analytics and Business Intelligence. Birmingham, England: Packt; 2022.
24. Harvey G, Kitson A. PARIHS revisited: from heuristic to integrated framework for the successful implementation of knowledge into practice. Implement Sci 2016 Mar 10;11:33 [FREE Full text] [doi: 10.1186/s13012-016-0398-2] [Medline: 27013464]
25. Collett D. Modelling Survival Data in Medical Research. Boca Raton, USA: CRC Press LLC; 2003.
26. Vieira ER, Palmer RC, Chaves PHM. Prevention of falls in older people living in the community. BMJ 2016;353:i1419. [doi: 10.1136/bmj.i1419] [Medline: 27125497]
27. Montero-Odasso M, van der Velde N, Martin FC, Petrovic M, Tan MP, Ryg J, et al. World guidelines for falls prevention and management for older adults: a global initiative. Age Ageing 2022;51(9):afac205 [FREE Full text] [doi: 10.1093/ageing/afac205] [Medline: 36178003]
28. Don't fall for it, falls can be prevented!: a guide to preventing falls for older people. Australian Government. 2011. URL: https://www.health.gov.au/sites/default/files/documents/2021/04/don-t-fall-for-it-falls-can-be-prevented.pdf [accessed 2025-03-08]
29. Statewide Action Plan: Falls Prevention in Older People 2002-2006. Brisbane: Queensland Health; 2002.
30. Final report: care, dignity and respect. Royal Commission into Aged Care Quality and Safet. 2021. URL: https://www.royalcommission.gov.au/aged-care/final-report [accessed 2025-03-08]
31. Davis J, Morgans A, Burgess S. Information management in the Australian aged care setting. Health Inf Manage J 2017;46(1):3-14. [doi: 10.1177/1833358316639434] [Medline: 27105477]
32. Graafmans WC, Ooms ME, Hofstee HM, Bezemer PD, Bouter LM, Lips P. Falls in the elderly: a prospective study of risk factors and risk profiles. Am J Epidemiol 1996;143(11):1129-1136. [doi: 10.1093/oxfordjournals.aje.a008690] [Medline: 8633602]
33. Klenk J, Becker C, Palumbo P, Schwickert L, Rapp K, Helbostad JL, et al. Conceptualizing a dynamic fall risk model including intrinsic risks and exposures. J Am Med Dir Assoc 2017;18(11):921-927. [doi: 10.1016/j.jamda.2017.08.001] [Medline: 28916290]
34. Rajagopalan R, Litvan I, Jung T. Fall prediction and prevention systems: recent trends, challenges, and future research directions. Sensors (Basel) 2017 Nov 01;17(11):2509 [FREE Full text] [doi: 10.3390/s17112509] [Medline: 29104256]

Abbreviations

DAX: Data Analysis Expression
EHR: electronic health record
FRAT: falls risk assessment tool
FRID: fall risk–increasing drug
PBI: Power BI

XSL•FO
RenderX

PH-FRAT: Peninsula Health Falls Risk Assessment Tool
RACF: residential aged care facility

Edited by Y Jiang; submitted 26.06.24; peer-reviewed by J Bhowmik, J Walsh; comments to author 20.09.24; revised version received 10.10.24; accepted 28.02.25; published 07.04.25.

Please cite as:
Silva SSM, Wabe N, Nguyen AD, Seaman K, Huang G, Dodds L, Meulenbroeks I, Mercado CI, Westbrook JI
Development of a Predictive Dashboard With Prescriptive Decision Support for Falls Prevention in Residential Aged Care: User-Centered Design Approach
JMIR Aging 2025;8:e63609
URL: https://aging.jmir.org/2025/1/e63609
doi:10.2196/63609
PMID:

©S Sandun Malpriya Silva, Nasir Wabe, Amy D Nguyen, Karla Seaman, Guogui Huang, Laura Dodds, Isabelle Meulenbroeks, Crisostomo Ibarra Mercado, Johanna I Westbrook. Originally published in JMIR Aging (https://aging.jmir.org), 07.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Evaluating the User Experience and Usability of Game-Based Cognitive Assessments for Older People: Systematic Review

Rhys Mantell[1,2], BEc, MPH; Ye In Jane Hwang[1,2], PhD; Matthew Dark[1], BSc, MPH; Kylie Radford[2,3,4], PhD; Michael Kasumovic[5,6], PhD; Lauren Monds[7], PhD; Peter W Schofield[8,9], MSc, MD; Tony Butler[1], PhD; Adrienne Withall[2,4], PhD

[1]School of Population Health, Faculty of Medicine and Health, University of New South Wales (UNSW), Sydney, Australia

[2]Ageing Futures Institute, University of New South Wales, Sydney, Australia

[3]Neuroscience Research Australia (NeuRA), Sydney, Australia

[4]School of Psychology, Faculty of Science, University of New South Wales, Sydney, Australia

[5]School of Biological, Earth & Environmental Sciences, Faculty of Science, University of New South Wales, Sydney, Australia

[6]Arludo Holdings Pty Ltd, Sydney, Australia

[7]Specialty of Addiction Medicine, Central Clinical School, The University of Sydney, Sydney, Australia

[8]Neuropsychiatry Service, Hunter New England Health, Newcastle, Australia

[9]School of Medicine & Public Health, University of Newcastle, Newcastle, Australia

Corresponding Author:
Rhys Mantell, BEc, MPH
School of Population Health
Faculty of Medicine and Health
University of New South Wales (UNSW)
Level 2, Samuels Building, Gate 11, Botany Street
Kensington
Sydney, 2052
Australia
Phone: 61 9385 3136
Email: r.mantell@unsw.edu.au

Abstract

Background: Game-based cognitive assessments (GBCAs) have the potential to transform the field of cognitive testing by enabling more effective screening of age-related cognitive decline. However, we lack a strong understanding of the usability and overall user experience of these games. This is a risk because the primary target users for GBCAs, older people, are seldom involved in game design research and development.

Objective: This study aims to address this gap by investigating the usability, acceptability, and enjoyability of GBCAs for older people.

Methods: This study followed established practices for undertaking evidence-based systematic reviews.

Results: The initial database search returned 15,232 records. After a thorough screening process, 8 studies remained for extraction and analysis. A synthesis of the included papers identified 2 overlapping yet distinct areas of focus: system usability and subjective user experience. Usability scores were mostly positive across the studies included. However, in several of the game studies, older adults and those with cognitive impairment tended to find GBCAs less usable. This trend was observed even when the games were explicitly designed for these populations, and the tasks were simplistic and representative of basic daily activities. In our second focus area, user experience, we identified the importance of perceived challenge in mediating gameplay experience across groups. That is, generating the appropriate level of difficulty for each user is important for positive user experiences, specifically enjoyment.

Conclusions: On the basis of these findings, we identified key learnings for researchers interested in designing and developing GBCAs. These include (1) recognizing that validity is essential but not sufficient on its own; (2) clearly defining the intended user; (3) designing games that align with the unique preferences and needs of older people; and (4), whenever possible, providing each user with their optimal level of challenge.

Trial Registration: PROSPERO CRD42023433298; https://www.crd.york.ac.uk/PROSPERO/view/CRD42023433298

XSL•FO
RenderX

(JMIR Aging 2025;8:e65252) doi:10.2196/65252

KEYWORDS

serious games; gamification; cognitive assessment; usability; user experience; aging

Introduction

Background

In recent years, there has been growing enthusiasm about the use of serious game–based technology. The primary objective of these games is to increase user engagement and motivation, thereby offering people a more pleasant environment to complete "serious" objectives that go beyond entertainment [1]. One area identified as being particularly well suited to the application of serious game technology is the field of cognitive assessment [2].

Serious games for cognitive assessment have typically been developed in 1 of 2 ways [3]. The first is "gamification," the process of building game-like features onto a cognitive task. This includes adding points, time limits, or appealing graphics to the foundational design of a traditional task. The second is to create or use an "actual" game, where the entire environment has been purposefully designed to be a game. This approach, more commonly associated with the term *serious game*, offers more creative development options, player choice, in-game competition, overarching narratives, and integrated or blended design elements. This approach is commonly based on theories of cognition or reimagining commercial games rather than reproducing a traditional cognitive task [3]. However, there remains significant confusion and inconsistency across the literature regarding what constitutes a serious game versus a gamified task. Therefore, unless otherwise noted, this paper refers to both approaches collectively as game-based cognitive assessments (GBCAs).

Traditional Cognitive Screening and the Case for Change

Traditional cognitive assessments are used by health professionals to evaluate a person's cognitive, or "thinking," functions to identify cognitive impairments or age-related declines, such as Alzheimer's disease (AD) [4,5]. Most clinicians initially use short cognitive screening tests (eg, 10-15 min) when assessing these functions. Although there are clear and well-evidenced benefits of using traditional cognitive assessment and screening tools (eg, the Montreal Cognitive Assessment has been shown to have high sensitivity [6]), there are also several issues that reduce their overall effectiveness. One concern often cited by users is that traditional screening tests can be boring and repetitive, thereby reducing user motivation and engagement during assessment. This is problematic because data obtained from individuals who are not motivated to perform a task to the best of their abilities may not be representative of their optimal performance, and this can cause misleading performance data (eg, false positives) [1]. Brief assessment tools also lack ecologic validity, meaning tasks may not translate to real-life circumstances [1,3,7]. In addition, assessments can sometimes create feelings of shame and distress [8]. Consequently, some people report significant test anxiety as well as self-stigma related to low literacy or education levels when completing cognitive tests [9,10]. Another limitation is the concerning association between cognitive performance and education level found with many screening tools [5]. Finally, some widely used neuropsychological tests are not culturally appropriate [11].

An initial attempt to address some of the concerns with traditional screening methods included digital software suites, such as CogTest and the Cambridge Neuropsychological Test Automated Battery [12]. Although these tools enable more consistent assessment and data capture, they are essentially digitized versions of traditional cognitive tests and, thus, are not gamified. Tests of this nature have encountered validation issues [12] as well as similar problems to traditional testing, such as a lack of user motivation when performing somewhat uninteresting tasks on a computer [2].

GBCAs Offer Exciting Alternatives, but They Must Be Valid, Acceptable, and Motivating

Games have recently been seen as a viable alternative to address many of these issues and improve both testing engagement and validity [1,2,12]. The benefits of using GBCAs include increased participant motivation and engagement relative to traditional tasks [2,3]. GBCAs can also offer shorter, more cost-effective, and scalable assessments [13,14]. Unlike traditional "pen-and-paper" assessments, games do not need to be initiated by a care provider and can be designed for self-administration or delivery by nonclinicians [1,14]. In addition, they may support more ecologically valid assessment through realistic context and gameplay, thereby engaging cognitive processes in ways that closely mirror real-life situations [15,16]. GBCAs can also enable consistent tracking of an individual's cognitive performance data to support the detection of subtle changes or variations in cognitive processes over time [14]. The richness of this type of data can provide invaluable population-level insight into subtle aggregate patterns, changes, or important impacts (eg, effects of exercise or smoking) on cognitive health across groups. This is an especially attractive prospect, given the increasing sophistication of machine learning models.

However, to justify their use, GBCAs must function as effective and reliable cognitive assessment tools (ie, psychometrically validated) while also being usable, acceptable, and motivating. Much of the current literature on GBCAs assumes that the simple process of gamification improves user engagement and creates a more satisfying and less tedious experience than traditional pen-and-paper or digitized neurological tests. Although many games have been found to achieve this [1-3], this is not always the case. For instance, Birk et al [17] found that adding game elements to standard cognitive tests (eg, go/no go; n-back tasks) did not improve user engagement when compared with the original task. Gamification actually reduced test performance for the go/no go task. Tong et al [18] reported difficulties in testing a carnival whack-a-mole game with older

XSL•FO
RenderX

users, also based on the go/no go paradigm, because of ergonomic issues with mobile and touchscreen devices (eg, the difficulty holding a touch device). It is also important to recognize that there may be restrictive settings or social contexts where GBCAs may be impractical, and traditional approaches may be more appropriate.

Overall, when games are not appropriately designed for their target users or contexts, they risk being neither acceptable nor motivating. Similarly, games might be skillfully designed and psychometrically validated but not adequately evaluated from users' perspectives. A systematic review of game-based interventions for neuropsychological assessment, training, and rehabilitation found that user experience evaluation was performed in <25% of interventions, with usability and enjoyability assessed in only 13% and 18%, respectively [19]. For a technology specifically aimed at increasing user engagement and enjoyment, this lack of user design input and evaluation appears to be a barrier to achieving the technology's intended outcomes.

GBCAs for Older Users

The group of interest for this study is older people (aged ≥50 years) who have or may be at risk of cognitive decline and, consequently, neurocognitive disorders (eg, dementia, such as AD). Recent research on aging has suggested that groups considered marginalized (eg, people experiencing homelessness or incarceration) are at risk of age-related conditions, such as frailty and cognitive impairment [20,21], once they reach the age of 50. Thus, to ensure adequate inclusion of people considered marginalized who are at risk of cognitive decline, we define older people as those aged ≥50 years.

Concerns about user engagement are exacerbated for game assessments aimed at detecting cognitive decline such as dementia because the target group for such tests is typically older people. This represents a challenge because games are rarely designed for or tested with older people. By contrast, there is promising—but preliminary—research to suggest that people who either have or are at risk of cognitive decline can play, attend to, and enjoy digital games, even when they do not have previous experience with touchscreen devices [22]. However, their ability to successfully complete gamified tasks appears to depend on the specific application and mechanics used within a game [22]. More broadly, the game preferences of older people seem to be different from those of younger cohorts, with studies reporting the preferences of older users gravitating toward intellectually stimulating games (ie, puzzle, educational, and strategy games) [23] that enable them to compete for high scores, require only a single player, and emphasize intellectual challenge over quick reflexes [22]. According to Blocker et al [24], games with violent content, fantasy characteristics, or interactive web-based components are not preferred by older users. A qualitative study investigating the cognitive game preferences of older people in prison reported similar findings [25], with participants suggesting a need to avoid "childish" game design, which some found condescending.

The Useful Concepts of Usability and User Experience

Across the design literature, usability and user experience are cited as major determinants of the successful adoption of any information system. Therefore, both are useful concepts when considering the acceptability, motivation, and engagement generated by any GBCA. Usability is the capacity of an object (ie, a digital game-based application) to functionally serve its intended purpose (ie, valid cognitive assessment) through qualities ranging from technical efficiency to conformity and configuration to ease of use [26]. According to the International Organization for Standardization (ISO: 9241-210:2019) [27], this refers to "the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use." According to Hassenzahl and Tractinsky [28], usable systems or products enable users to achieve pragmatic objectives successfully and efficiently without obvious barriers.

On the other hand, user experience refers to the feelings a person has while interacting with a product under particular conditions [26]. This means that it goes beyond meeting the instrumental (pragmatic) need for usefulness and efficiency that good usability provides. User experience includes the need to achieve hedonistic goals [28,29], which involve "users' emotions, beliefs, preferences, perceptions, comfort, behaviors, and accomplishments that occur before, during and after use" [27]. Although there is a clear overlap, the distinction between the concepts is that good usability makes it easy for someone to use something, whereas good user experience makes using that thing feel enjoyable or satisfying.

In the context of gamified technology, usability is focused on technical features such as a seamless interface, ergonomics, clear instructions, minimal in-game bugs or errors, and other features that facilitate use, whereas user experience is concerned with impacting a person's subjective experience about the contents of a game, that is, a compelling narrative, an exciting challenge, motivation toward certain objectives, and so on [26]. There are more detailed and stratified definitions of usability and user experience across the user design literature (eg, refer to the 10 Usability Heuristics for User Interface Design by Nielsen [30]). However, consistent with our simplified definition of GBCAs, we intentionally use the basic demarcations of usability and user experience defined earlier to enable a consumable preliminary synthesis and analysis of this still-developing area of inquiry (ie, user evaluations of GBCAs for older people).

Arguably, the critical success factor for any GBCA is a coupling of user evaluation and psychometric validity. This is particularly true for older people at risk of, or presenting with, cognitive decline who are likely to have different preferences and experiences from other cohorts. However, much of the current literature is skewed toward investigating psychometric validity before or without evaluating usability and user experience. For instance, although several reviews in recent years have investigated GBCAs [2,3,31-34] and some have reported summary results on user evaluation methods (ie, how many studies conducted user evaluation) [19,35], there remains a limited understanding of the importance of user evaluation

XSL•FO
RenderX

findings. This appears to be a risk, as even the most clinically useful game in the world is unlikely to succeed at scale if it is neither functional nor engaging.

The Aims of This Study

This systematic review attempts to address some of these concerns by investigating the usability, acceptability, and enjoyability of GBCAs that have undergone psychometric validation. The intention is to better understand the nexus of clinical usefulness, usability, and user experience and, in turn, report preliminary lessons learned from user evaluations of validated GBCAs.

Methods

This study (protocol registered on PROSPERO [CRD42023433298]) followed established practices for undertaking evidence-based systematic reviews, including the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist (Multimedia Appendix 1) and the Scottish Intercollegiate Guidelines Network methodology checklist for study quality appraisal (Multimedia Appendix 2 [16,36-43]).

Data Collection, Extraction, and Quality Assessment

Independent and systematic database searches of PsycINFO, Embase, IEEE, and MEDLINE were conducted by the first author (RM) based on the predefined search terms summarized in Table 1. Broad search parameters were set, given the disparity of terminology used across the GBCA research. Relevant studies were then imported into Covidence, ready for initial screening. Title screening was manually completed by RM, where obvious exclusions and duplicates were removed. Duplicates were removed automatically via Covidence's duplicate removal function. This was followed by independent and systematic screening of relevant abstracts by 2 reviewers (RM and MD), who applied predefined inclusion and exclusion criteria to assess all abstracts.

Table 1. Database search terms. Note: the truncation symbol * (asterisk) searches for multiple variants of a word (eg, singular, plural, different conjugations, etc) all at once (eg, cogn* includes searches for cognition, cognitive, cognitively, cognizant, etc). MCI: mild cognitive impairment; VR: virtual reality.

Topic area	Search terms
Serious games	game* OR gami* OR virtual reality OR VR
	AND
Cognitive	cogn* OR MCI OR dementia OR neurocog* OR neuropsych*
	AND
Assessment	assess* OR test* OR measur* OR collect* OR detect*

Textbox 1 presents the inclusion and exclusion criteria. Important definitions and justifications relating to the inclusion and exclusion criteria are discussed more thoroughly in the Discussion of important terms subsection. When disagreements arose at the abstract screening level, the second author (YIJH) facilitated a resolution through a 3-way vote. Full-text screening was then conducted by RM and MD. Conflicts at this stage were again handled through 3-way voting consensus involving YIJH, RM, and MD. At the full-text screening stage, there were disagreements with 12 (15%) of 83 papers, which necessitated additional consensus voting. Finally, a thorough data extraction process was conducted by RM, including a risk of bias assessment. We used the Scottish Intercollegiate Guidelines Network "Methodology Checklist 4: Case Control Studies" [36] guidance to assess the risk of bias for each included study. A detailed risk of bias assessment can be found in Multimedia Appendix 2. This final extraction process was reviewed by all authors to ensure consensus and quality control.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- User evaluated
- People aged ≥50 years
- Psychometrically validated
- Assessment or screening
- Cognitive functioning
- Primary research
- 2015 onward
- Digital serious games and gamification of task

Exclusion criteria

- Did not include target population or outcome
- Intervention with high immersion (via head-mounted display, etc)
- Training or rehabilitation focused without assessment
- No psychometric validation
- No user evaluation
- No data collection performed (ie, game performance data not collected)
- Inappropriate study design (reviews, conference abstracts, etc)
- Focus on social cognition only

Discussion of Important Terms

This section summarizes and justifies several key concepts relevant to our inclusion criteria in Textbox 1, as well as our overall search strategy. It is important to emphasize that (1) our search strategy was intentionally broad to minimize the risk of missing any relevant GBCA studies and (2) our inclusion criteria were intentionally narrow to address our research aim of investigating the usability, acceptability, and enjoyability of GBCAs that have undergone psychometric validation.

User-Evaluated Study

The key criterion in this review was that a GBCA must have been user evaluated through primary research. This required the explicit inclusion of methods to measure usability, user experience, user motivation, user preference, enjoyment, or acceptability of the game.

GBCA Overview

Our definition of GBCA mirrors the broad definition included in the Introduction section. That is, both simple task gamification and more “from scratch” serious games were included. In addition, our focus was technology-based GBCAs; thus, we only included digital GBCAs (ie, games that could be played on computers, tablets, phones, etc).

Target Group

The target group for this review was older people (aged ≥50 years) who had, or were at risk of, cognitive decline and, consequently, neurocognitive disorders (eg, dementia, such as AD). Although there are numerous GBCAs that have been designed and tested to detect cognitive impairment associated with other disorders across the life course (eg, attention-deficit/hyperactivity disorder, depression, traumatic brain injury, drug- and alcohol-related cognitive deficits, etc), this paper is interested specifically in the assessment of cognitive decline in older people. This demarcation enables a manageable and specific grouping of findings.

Validated Study

To be included in this review, a GBCA must have been psychometrically validated with a sample of older users (aged ≥50 years). This validation could have occurred in the same study as the user evaluation or in a previous validation of the same game. In other words, evidence of validation was a prerequisite for inclusion in this review, but validation of the game was not necessarily required within the actual studies included in this review. According to previous definitions [3], validation included comparisons against a traditional cognitive screening tool (eg, Montreal Cognitive Assessment), a comprehensive neuropsychological battery, or a previous clinical diagnosis of cognitive impairment.

Cognitive Domains Included

A broad definition of cognitive functioning was applied for this study. Cognition refers to the brain’s ability to perceive, assimilate, organize, store, and manipulate information [44]. As such, cognitive functioning is an umbrella term that encompasses a range of integrated skills and domains that allow us to process and respond to information within our environment [44]. These domains include (but are not limited to) basic functioning (eg, processing speed and attention) and memory (eg, episodic memory, semantic memory, and prospective memory), visuospatial functions (eg, color perception and mental

XSL•FO
RenderX

rotation), and executive functions (eg, planning, and decision-making). Relevant GBCAs testing any (or all) of these domains were included in this review, with the notable exception of social cognition. The clinical assessment of social cognition and decline in older people is not as well established as these other cognitive domains and was therefore intentionally excluded. However, investigating social cognitive domains should be seen as a critical consideration for future research.

Low-Immersion Interventions

This review made a distinction between low- and high-immersion games when deciding if a study should be included. Games that were deemed highly immersive from a technological perspective, such as those requiring headset technology and immersive 3D environments, were excluded. This review focused on GBCAs in a digital format that enabled accessibility through touch or mouse click capability. The main reason for this distinction was to ensure uniformity in findings. Although highly immersive games have significant potential benefits with respect to ecologic validity and integrated assessment of multiple cognitive functions [45], they have distinctive technology requirements when compared to computer or tablet-based game systems. In addition, factors such as motion sickness [46], which can affect enjoyability, are specific to virtual 3D headset environments and would have added unwanted complexity to this synthesis. Importantly, due to the disparity in the definition of serious games, gamification, and virtual reality across the literature, some interventions have been included in this review even when the authors refer to their interventions as *3D* or *virtual reality* game applications. However, these interventions aligned with our definition of a tablet-based and low-immersion GBCA and were thus included.

2015 Onward

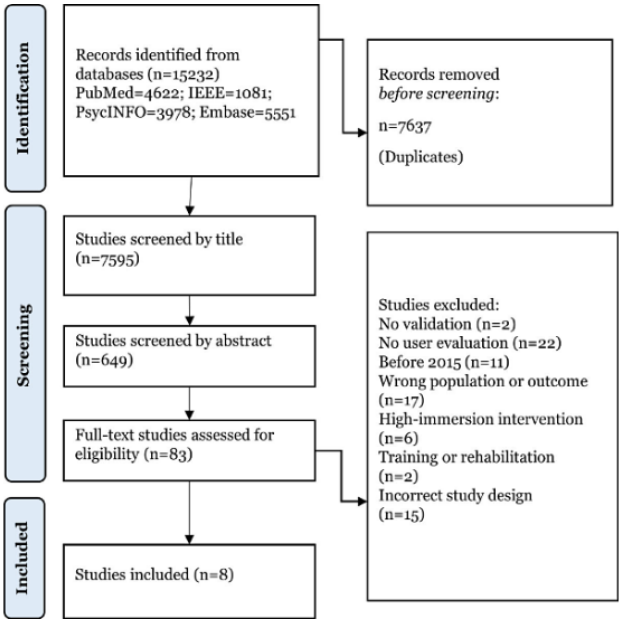
There is limited literature before 2015 that investigated games that were both psychometrically validated and user evaluated. There are 2 notable reviews that looked at validity before 2015 [2,34], but these reviews concluded that little work had been done in user evaluation of GBCAs before this time. In addition, a preliminary scan of literature before 2015 by the first author confirmed a lack of usability data from psychometrically validated games. Given our very broad search strategy, the significant growth of literature on GBCAs in the last 5 to 7 years, and a lack of user evaluation data before this time, a cut-off of 2015 for selection appeared justified.

Results

Overview

The initial database search on December 9, 2022, identified 15,232 records. There were 7595 articles left after removing duplicates. A total of 6946 articles were removed during the title screening for obvious exclusions, leaving 649 articles for abstract screening. Full texts were retrieved for 83 articles, of which 8 studies remained for extraction and analysis. More detail is provided in the PRISMA flow diagram in Figure 1, including reasons for exclusion at the full-text phase. The key study characteristics for all included papers are presented in Table 2. A summary of key psychometric validation results can be found in Multimedia Appendix 3 [15,47,48]. All studies included psychometric validity testing as part of their analysis except the Virtual Supermarket Test (VST) [37], where validity testing had been completed previously [15,47,48]. All games were designed and developed by the researchers involved in the included papers, with the exception of the Smart Aging Serious Game (SASG) [38], which was developed by Zucchella et al [49]. None of the included papers involved testing commercially available games; however, 2 of the games [39,40] were experimentally controlled versions of existing games. In addition, 4 of the 8 game studies were noted by the authors as being “virtual reality” interventions, which attempted to replicate real-life activities to assess cognition. However, all these games had low immersion and adhered to our definition of a digital GBCA (refer to the Discussion of important terms subsection for further justification).

Figure 1. PRISMA flow diagram.



XSL•FO
RenderX

Table 2. Summary of key study characteristics.

Studies	Game	Sample, n	Age, mean	Female (%)	Study aim
Cabinio et al [38], 2020	SASG[a]	Total: 139; controls: 107; MCI[b]: 32	Controls: 76.47; MCI: 76.75	Controls: 49.5; MCI: 46.8	Tested the usability of SASG in a cohort of participants with MCI and determined the validity of the SASG in discriminating between a preclinical population with MCI and healthy controls.
Chesham et al [40], 2019	SMT[c]	Total: 52; young (18-35 y): 28; older (65-85 y): 13; oldest (≥85 y): 11	Young: 21.68; older: 70.54; oldest: 89.27	Young: 71.4; older: 53.9; oldest: 81.8	Examined the initial validity and usability of an experimentally controlled version of a popular TMM3 puzzle video game—the SMT.
Manera et al [41], 2015	Kitchen and Cooking	Total: 21; MCI: 9; AD[d]: 12	MCI: 75.8; AD: 80.3	MCI: 77.8; AD: 86.7	Conducted a feasibility study of patients with MCI and AD and related disorders using the game "Kitchen and Cooking."
Nef et al [39], 2020	Numberlink puzzle task	Total: 55; young (18-31 y): 18; older (64-79 y): 14; oldest (86-98 y): 14; PD[e]: 4; HD[f]: 5	Young:21.83; older: 71.4; oldest: 89.4; PD: 67.5; HD: 49.4	Young: 66.7; older: 57.1; oldest: 85.7; PD: 50; HD: 50	Evaluated the feasibility and preliminary validity of a maze-like NL puzzle video game as a tool to assess cognitive and motor differences in adults and patients with neurodegenerative disorders.
Valladares-Rodriguez et al [16], 2017	Episodix	Total: 16; controls: 8; MCI: 3; AD: 5	Controls: 68.3; MCI: 75.8; AD: 75	Total: 75	Investigated if a video game could be designed and developed to assess episodic memory and predict early cognitive impairments in an ecologic setting with low immersion.
Vallejo et al [42], 2017	Virtual games	Total: 38; controls: 20; AD: 18	Controls: 74.6; AD: 77.8	Controls: 40; AD: 50	Evaluated the usability and the screening potential of several low-immersion virtual game tasks for patients with AD.
Wang et al [43], 2022	GBCA[g]	Total: 124; controls: 57; NCDs[h]: 67	Controls: 72.31; NCDs: 74.78	Controls: 69; NCDs: 63	Developed a game-based tool and evaluated its validity as an early screening for patients with cognitive impairment.
Zygouris et al [37], 2022	VST[i]	Total: 57; SCD[j]: 24; MCI: 33	SCD: 65.58; MCI: 68.39	SCD: 79; MCI:70	Assessed the VST's usability in a sample of older adults with MCI and SCD. The paper analyzed how usability is affected by age, education, diagnosis, in-game performance, and familiarity with touch devices.

[a]SASG: Smart Aging Serious Game.

[b]MCI: mild cognitive impairment.

[c]SMT: search and match task.

[d]AD: Alzheimer's disease.

[e]PD: Parkinson disease.

[f]HD: Huntington disease.

[g]GBCA: game-based cognitive assessment.

[h]NCD: neurocognitive disorder.

[i]VST: Virtual Supermarket Test.

[j]SCD: subjective cog decline.

User Evaluation Findings

Overview

Preliminary analysis of the included papers identified 2 distinct areas of focus: usability and user experience. As such, our proceeding synthesis is demarcated based on usability and user experience (applying the simple definitions of these constructs provided in the Introduction section). Reporting findings from these 2 overlapping yet distinctive areas provides a holistic perspective of user evaluation by incorporating both system-level usability results as well as more individualized user experiences of serious gameplay. A detailed overview of relevant study findings and key statistical results is presented in Multimedia Appendix 4 [50-56].

GBCA Usability

A total of 5 studies [37-40,42] included system-level usability results. All these studies involved participant self-report measures but varied in the level of detail participants were asked to report. Overall, the average usability, from a systems perspective, of the games presented ranged from good to excellent. Familiarity with computers or touch screens did not appear to have a significant effect on system usability for computers or touchscreen devices, which is a positive finding. However, there were distinct differences in system usability across age groups and cognitive functioning levels in some

XSL•FO
RenderX

studies. An interesting relationship was also found between usability and game performance in the VST study [37]. These are discussed further in the subsequent sections.

In the SASG study [38], participants were simply asked to respond to some questions about their familiarity with computers and touch screens. Results showed that familiarity with computer systems did not influence the SASG game performance score. Although this is not in itself an assessment of usability, it revealed an important usability finding: a lack of digital experience did not impact overall cognitive performance scores in a sample of older game users with low digital literacy and experience. A potential reason for this result may be the cautious and simplistic design of the SASG interface, which supported usability for the target cohort. The game was built specifically for older and nonexpert users, and according to the authors, it did not require skilled digital abilities. Using a low-immersion, first-person perspective, the game required users to complete gamified versions of everyday tasks, such as watering flowers with the radio on and dialing a phone number. Furthermore, the authors suggested a key factor in the usability of their serious game intervention was the decision to use a touch screen instead of a mouse.

The remaining 4 studies used the system usability scale (SUS) to measure and analyze usability [57]. The SUS is a brief questionnaire (10 questions) using a 5-point Likert scale with a possible aggregate score ranging from 0 to 100. The SUS includes questions such as “I thought the system was easy to use” and “I needed to learn a lot of things before I could get going with this system.” According to previous studies [58,59], the average SUS score (50th percentile) is 68, a good score is between 71 and 84, and an excellent usability score is ≥85.

In the VST study [37], the average SUS score was 83.11 (SD14.6). No statistically significant differences were found between the participant groups, people with subjective cognitive impairment and mild cognitive impairment (MCI), regarding SUS scores. However, there was a significant correlation (r=−0.496; P<.001) found between the SUS score and the average time needed to complete the VST test trials, independent of the participant group. This suggested a relationship between game completion and usability: the lower the participants rated the game’s usability, the longer they took to complete it. Although this was an interesting finding, the direction of causality is unknown. As the authors noted [37], “It is unclear whether participants who took longer to complete the VST test trials provided a lower usability score as they were frustrated by the long time it took them to complete the exercise or if there are underlying usability issues that affected their performance thus resulting in a longer time to complete the VST test trials.” It is also worth noting that in the VST study, participants were asked to complete the VST exercise on their own. The examiner prepared the tablet and launched the VST game but did not assist the participants. In fact, researchers intentionally provided little support to participants once gameplay started. The reason for this was to assess the usability of the system in an environment where a user may need to self-administer the game without support.

In the Numberlink (NL) puzzle task [39], the SUS average was 93.38 (SD 5.72) for young adults, 93.33 (SD 6.25) for older adults, 83.39 (SD 11.79) for the oldest adults, 81.88 (SD 9.66) for people with PD, and 83.12 (SD 10.68) for those with HD. These scores reflected a statistically significant difference in usability ratings between groups (P=.02). There were similar findings in the search and match task (SMT) study, albeit with lower reported usability overall [40]. The average SUS score was 88.67 (SD 7.28) for younger adults, 79.09 (SD 15.50) for older adults, and 68.25 (SD 18.78) for the oldest adults. This indicated a significant effect of age group (P=.01), with usability ratings decreasing with increasing participant age. Finally, the virtual game study [42] reported a mean SUS score of 83.5 (SD 11.16) for the control group and 83.75 (SD 9.82) for people with AD. As such, similar to the VST study and contrary to the NL puzzle task and SMT studies, there was no significant group difference found in usability. There was also no significant difference in SUS scores found between participants with and without touch device familiarity.

GBCA User Experience

Five studies [16,39-41,43] examined user experiences of gameplay, specifically focusing on participant engagement with game elements, perceived motivation, and enjoyment. Overall, all the games presented were well received by users regarding motivation, enjoyability, and satisfaction. However, there were some important differences across the study findings. For instance, some studies identified a relationship between being cognitively healthy and improved user experience [40,43], whereas an opposite relationship was found in another study [41]. Furthermore, 2 studies identified the role of a health professional as essential to the success of their respective game interventions [16,41], which contrasted with the intent in studies to test implementing games in a self-administering environment [37,41].

In the Episodix study [16], a user experience questionnaire based on game usefulness and user motivation was completed by participants. The questionnaire was completed twice, once before and once after gameplay. Before playing the game, most participants reported low interest, motivation, and perceived usefulness of serious games; however, their experiences playing Episodix changed these perceptions. For instance, the average participant’s motivation to play video games increased by 40% after gameplay. Participants also indicated that the Episodix game seemed more engaging than the California Verbal Learning Test (CVLT) pen-and-paper test, which the game was based on. One potential reason for the game seeming more engaging than the CVLT may have been the comprehensive user-centered design process used in developing the game. A combination of validity testing, expert input, and user focus groups underpinned the development process. Through the design process, participants from both focus groups and the pilot experiment indicated a preference for a touch interface rather than a traditional keyboard and mouse setup. This finding was similar to that of the SASG game study. Technical developments were made to facilitate these preferences. The Episodix study [16] also reported important qualitative findings from user focus groups. The game was perceived as very useful by users, but they preferred it to be administered by a health

XSL•FO
RenderX

professional rather than self-administered, regardless of whether they were confident interacting with the game on their own. Participants agreed that if both instruments (the CVLT and the game) were demonstrated to perform the same cognitive assessment, they would prefer to play the game. However, participants preferred the digital intervention to be referred to as a test because they perceived the term *video game* to have negative and nonserious connotations.

Both the NL puzzle task and SMT studies [39,40] assessed user experience through the Perception of Game Training Questionnaire [50]. Participants rated the extent to which they found playing the games enjoyable, challenging, frustrating, and motivating. Both games were developed based on experimentally controlled versions of commercially available puzzle games. In addition, both games offered different difficulty levels to create an appropriate level of challenge for a broad range of users. In the NL puzzle task study, ratings of enjoyment (P=.43), challenge (P=.07), frustration (P=.06), and motivation (P=.67) for the game did not differ significantly between groups. The users generally rated the game experience positively in both the healthy aging group and the group with cognitive impairment. However, there was a slight, albeit statistically nonsignificant, increase in rating the game session as challenging and frustrating among older participants and/or those with impairment, particularly participants with PD and HD. According to the authors, further adapting the difficulty levels to a person's performance might partially reverse this trend. In the SMT study, there were significant differences between the 3 age groups regarding ratings of challenge (P<.001). Overall, young adults perceived the SMT as significantly less challenging than older adults (P=.02). There were also significant age group differences in average difficulty ratings for both the short version played by the young, older, and oldest adults (P<.001) and the long version (P<.001) played by the young and older adults. Despite this, there were no significant group differences regarding enjoyment, frustration, and motivation while playing the SMT task. One possible reason why challenge and difficulty levels varied across groups while enjoyment and motivation levels remained constant is that the SMT provided a total of 71 possible difficulty levels. This means that the game could be responsive to a user's gameplay and adjust the level of perceived challenge accordingly.

In the GBCA study [43], the user experience questionnaire included questions on game ease-of-use (not technically user experience), whether users felt the game stories were interesting and familiar, users' reactions to the interface design (typesetting, color, instruction, etc), and questions aimed at comparing the GBCA intervention with other cognitive assessment tools. The responses of the control group suggested that they were very satisfied with the game. By contrast, the average responses of participants with neurocognitive disorders were significantly lower than those of controls. Interestingly, lower scores given by the participants on the user experience questionnaire corresponded to a higher clinical dementia rating, older age, and a lower educational level. Higher levels of cognitive impairment appeared to adversely affect the participants' ability to use the tablet to complete the GBCA intervention, as indicated by the poor ratings given to questions regarding the design of the GBCA intervention. This finding raises questions about the functional usability of the game as well as the more subjective user experience. Furthermore, older age had a negative influence on the participants' satisfaction regarding the pictures, text, and story used in the game. This was somewhat surprising, given that the user interface of the GBCA intervention was intentionally designed to be acceptable for older users and included large pictures and buttons as well as spoken instructions so users could operate the app easily. The game could also be played by tapping buttons, drawing by dragging one's fingers over the touch screen, and speaking into a microphone (using low-cost speech recognition technology).

The Kitchen and Cooking study [41] also captured user experiences. The participants completed a questionnaire about the game experience at 2 points during the study. Specifically, satisfaction, interest, motivation (intrinsic and extrinsic), emotional experience of gameplay, and fatigue levels were assessed. The overall results of the self-report questionnaire showed that the participants were highly satisfied, interested, and motivated by the game experience. Intrinsic motivation was significantly higher than extrinsic motivation. Average user experience scores did not change between the 2 sessions, thus confirming the overall positive evaluation of the game after repeated gameplay. However, it is worth noting that scores did not trend upward after the second session regarding motivation, as observed in the Episodix study [16]. Of particular interest in this study is that participants with AD reported being significantly more satisfied with the game than participants with MCI (P=.04). This contrasts with the findings from the NL puzzle task, SMT, and GBCA studies, where the inverse trend appeared. Finally, the authors suggested that a critical factor in the success of the intervention was the presence of a clinician. The game was designed as a tool to help patients train outside clinical consultations. However, the periodic supervision of the clinician helped to explain the functioning of the game to the patients and their families, keep track of the evolution of the performance, adapt the intervention to the patients' changing needs, and maintain user motivation.

Discussion

Overview

The aim of this review was to investigate studies that evaluated the experiences of older people using GBCAs. Although the combined findings are from only 8 included papers, they establish promising preliminary user evaluation results for the use of GBCAs, albeit with some interesting implications for future design and development practices. Our review provides a much-needed synthesis of user evaluation findings from GBCAs, which have undergone psychometric validity testing. As such, we offer new insights regarding the relationship between the validity, usability, and user experience of GBCAs. This is particularly important with respect to the experiences of our target study cohort, older people (aged ≥50 years), whose input is greatly needed yet seldom included in GBCA design and development.

The first area of focus in this review was system usability. Usability scores were mostly positive across the studies

XSL•FO
RenderX

included. However, in some of the games presented, there were trends toward lower usability scores for participants with cognitive impairment. This was observed even when games were explicitly designed for older people and/or those with impairment, and tasks were simplistic and representative of basic daily activities, such as in the VST. In the VST, a game that has been validated in multiple contexts [15,47,48], no significant differences were found between the groups with subjective cognitive impairment and MCI in usability scores; however, participants who took longer to complete the test trials had lower self-reported usability. Usability challenges were further identified in the NL puzzle task and the SMT. Although both games had good usability scores on average, they both identified significant differences in usability ratings not only across cognitive health but also age range. On average, in both games, the older or more impaired a person was, the less usable they found the respective game system.

These collective findings on usability have some interesting implications. Traditionally, it has been assumed that system usability can be increased by improving the technical interactions a person has with a system, the intuitiveness of that system, and by reducing any bugs or errors that generate needless complexity within that system. For example, Joddrell and Astell [60] significantly improved the usability of solitaire for people with dementia (mean age 84.17, SD 8.35; range 66-102) by adding accessibility features such as additional card control options, clearer gameplay layout, and audio-visual feedback cues. However, this review highlights that while these technical and functional system improvements may be important, they are not necessarily sufficient to produce usable interventions for this target group of users. Individual variations in age, cognitive function, and gameplay performance may all impact functional usability, even when the system has been technically adapted to be suitable for older users and/or those with impairment.

The second focus area, regarding subjective user experience, identified the importance of perceived challenge in mediating gameplay experience across groups. In the GBCA study, the healthy control group was more satisfied with their experience of using the GBCA intervention than the group of participants with cognitive impairment, whereas the opposite trend appeared in the Kitchen and Cooking game. In the Kitchen and Cooking game, a simplistic interface based on a basic cooking task, game satisfaction was higher among the participants with AD than among participants with MCI. In the SMT task, there was a statistically significant difference in how challenging older and younger people found the game. However, levels of enjoyment and motivation remained the same. These contrasting results may be related to perceived gameplay challenges. For instance, the NL puzzle task, the SMT, and the GBCA were, on average, more complex than the Kitchen and Cooking game, with the NL puzzle task and SMT offering different levels of complexity for different users.

According to the broader literature on game design preferences, people want to be challenged to a level that will satisfy their need for competence [25]. Applying self-determination theory, it has been shown that a player will experience low competence if the challenges are too great (eg, game controls are overly complex or enemies are too numerous), and this will have a negative impact on a person's ability to enjoy the game [61]. If a game is too difficult, it can become frustrating and discouraging, reducing a user's motivation, engagement, and enjoyment. The optimal challenge (ie, the sweet spot) for any player is mediated by that individual's threshold for difficulty [62,63]. This may explain why basic gamified versions of simple tasks, such as in the Kitchen and Cooking game, were less motivating or enjoyable for younger participants and those without cognitive impairment. It may also explain why more "from scratch" serious games that offered more difficulty levels, such as the NL puzzle task and SMT, had constantly high enjoyment and motivation scores across the younger, older, and oldest groups of users.

Another noteworthy finding in this review relates to the role of a health professional during game-based assessment. Two studies identified the role of a health professional as essential to the success of their respective game interventions. In the Episodix game, although users demonstrated improved appreciation for the value of GBCAs after gameplay, this increased appreciation appeared to stem from the value of the game as a health tool. The presence of a clinician seemed to underpin this, as the game was perceived as very useful by users, but they preferred it to be administered by a health professional rather than self-administered. Likewise, in the Kitchen and Cooking study, the authors suggested a critical success factor for the intervention was the presence of a clinician. These findings raise questions about self-administration, which is seen as one of the most exciting (and valuable) applications of serious game technology [15]. Self-administration of GBCAs has been recently identified as a way to overcome some of the issues with formal assessment [15], such as significant test anxiety as well as self-stigma regarding low literacy or education levels when performing cognitive tests [9,10]. Self-administration may also empower people to take control of their own health earlier and more proactively than standard approaches. However, the findings of this review suggest that this autonomy and control may come with some unintended consequences, such as a lack of confidence in undertaking assessments without support, as well as a lower perceived value or legitimacy of the assessment. This is important to recognize as a potential downside for self-administrated GBCAs.

The Objectives of GBCAs and Subsequent Tensions With Usability and User Experience

One approach often taken when designing and developing cognitive games for older people and/or individuals with cognitive impairment is to make the game easy. This appears to have been the design strategy in games such as SASG, VST, and Kitchen and Cooking, where the games were based on simple daily living tasks, and the design and development of the games were tailored to people with cognitive decline. These games were quite successful at engaging older people with significant cognitive impairment. However, given that the objective of a GBCA is to assess maximal cognitive capacity, adapting the floor and ceiling to make games easier may not always be desirable. This is because the reason for adding complexity to a GBCA, besides increasing user engagement, is to improve the sensitivity of a game to detect cognitive

XSL•FO
RenderX

functioning more accurately. In other words, a game that is too simple or too domain specific may not be relevant for healthier older adults or those with milder or more complex forms of cognitive impairment, even if it is more appropriate for people with more significant cognitive decline. For instance, the basic requirement of shopping for groceries in the VST may be suitable for people with moderate neurocognitive disorders; however, it may not detect subtle impairments or small cognitive changes over time in healthier participants. More simplistic games are also unlikely to be very enjoyable or challenging (ie, satisfying a need for competence) for many people without neurocognitive disorders (as was observed in some of the studies included). This reduces their viability as a scalable screening tool. By adding more challenge and complexity, it becomes feasible to engage a broader user group and assess multiple cognitive domains. This, of course, comes at the risk of marginalizing the most vulnerable and arguably important cohort: those with significant impairments. It is also important to note that people with more significant neurocognitive disorders are unlikely to be the intended target of most GBCAs, or even traditional cognitive screening. This is a tension that has been seldom discussed or investigated in the growing body of literature on GBCAs.

Overall, the goals of usability and user experience—foundational to human-centered design approaches—and the intention of GBCAs and cognitive testing, in general, reveal a potentially problematic misalignment that warrants further discussion. For instance, according to ISO, the basic requirement of usability is that a product can be used (by target users) to achieve specific goals with effectiveness, efficiency (ie, in minimum time and with minimum errors), and satisfaction in a specified context [27]. However, is it reasonable, desirable, or even possible to meet this standard when developing a GBCA? If the primary goal of a cognitive test is the detection of cognitive deficits, the implication is that people with cognitive impairment will likely (1) fail to achieve certain tasks, (2) take longer to complete objectives, and (3) make more errors than cognitively healthy people.

In fact, reliable detection of these "failures" is essential to ensure a valid and clinically meaningful test of cognitive performance. Thus, intentionally producing user failure states is a central design tenet when developing GBCAs, and this, as we have identified through this review, may cause reductions in usability and subjective user experiences.

However, while this may risk departing somewhat from the principles of user-centered design, it is important to recognize that the "serious" objective of a GBCA is not necessarily to maximize unconstrained usability and user experience; rather, it is to produce a cognitive assessment that is more motivating and engaging than a traditional task. As traditional tasks are sometimes reported to be tedious, boring, anxiety provoking, and stigmatizing, the objective of a GBCA is to improve the subjective user experience relative to traditional approaches while maintaining functional usability. This implies that a GBCA may not necessarily need to be maximally fun but simply "fun enough" while being motivating, engaging, and technically usable. Furthermore, what "usable" means in this context is potentially hard to reconcile with the ISO aim of effectiveness, efficiency, and satisfaction. For instance, if an intervention is too easy to use, can it adequately produce the failure states that are the primary objective of a GBCA?

These open questions become even more interesting when we consider that the intended application of GBCAs, at least in the near future, is likely to be a relatively short (eg, 10-20 min) user experience (ie, replacing a cognitive screen), which is only "played" by users sporadically. The implication is that traditional user-centered game design that focuses on intrinsically motivating features, such as narrative and social connectedness, may not be feasible or desirable. This, in many ways, is misaligned with the overarching goals of human-centered design, yet it is arguably more consistent with the preferences of older people taking a cognitive test. In fact, this was what we reported in a recent qualitative study investigating the GBCA preferences of older people in prison [25]. We found that GBCAs with numerous immersive game features risked being perceived as too childish for the serious context of a cognitive test. This was also likely the case in the Episodix study sample, where participants preferred the digital intervention to be referred to as a test because they perceived the term game to be nonserious [16]. However, we argue that these tensions do not reduce the value of incorporating usability and user experience principles into the development of GBCAs; rather, they simply reframe what the feasible intention of a GBCA ought to be in the context of user experience and usability. Perhaps "good enough" usability and user experience are good enough?

Our conclusion is that there is an inherent tension between usability, user experience, and GBCAs that (1) in the meantime should be carefully considered when designing and developing GBCAs, especially for older people; and (2) requires additional research to further explore and potentially resolve. We offer some key lessons in the Implications for Practice subsection. Regarding future research, we suggest an urgent need for further mixed methods research that incorporates the qualitative preferences of older people using GBCAs with rigorous validity testing. In addition, there is a need to explore what "good enough" usability and user experience may mean in this context and how these concepts can be tested and incorporated into game design, development, and improvement.

Limitations

This review has some limitations. Many of the participant samples were small, and the studies involved slightly different target user groups. Most studies focused on user evaluation as a secondary outcome, with psychometric validity testing being the primary aim in most papers. All the games included in this review were distinct from one another regarding design and development processes and, as a result, the end products. Combined with the disparities in sampled participants, this heterogeneity makes it inappropriate to make specific recommendations regarding which precise techniques work for whom in what contexts. This is further complicated by the fact that most of the games did not seem to be publicly available, creating uncertainty about how well developed and sophisticated the games actually are in practice. This lack of access to the games, beyond what was described and visualized in the primary research papers, also inhibited a more comprehensive review

XSL•FO
RenderX

of specific design choices based on a more technical analysis of mechanics and processes. However, despite these limitations, we were still able to interpret the findings in the context of the broader literature on serious games and provide general lessons learned for designing and developing GBCAs.

In addition, there were also some limitations related to the methods applied for user evaluation. Most user evaluation data were collected via self-report and are thus exposed to the well-established biases associated with this collection method. Furthermore, although the SUS is well validated, it is short and generic. Thus, it does not present a comprehensive assessment of game usability, and its validity for people with neurocognitive disorders has not been established. Other reported usability and user experience findings came from disparate tools, some of which were customized from other scales or created based on theory. This created additional barriers when attempting to synthesize the findings quantitatively and prevented a meta-analysis. Some of these limitations were evident in the quality appraisal, in which a case-control checklist was used to assess the risk of bias. This basic checklist was deemed appropriate because most of the study papers included games that were still in the pilot stage of development and, thus, not ready for more sophisticated diagnostic appraisal. Regardless, it is hoped that a more sophisticated and suitable appraisal tool can be used in future research to adequately assess the quality of papers investigating the psychometric validity and user experiences of GBCAs.

Implications for Practice

Overview

This systematic review aimed to synthesize the user evaluation data from psychometrically validated GBCAs played by older people. As part of this process, we have identified key lessons for developing potentially transformative assessments using gamified technology. These preliminary key lessons have been developed based on our systematic review of the included studies, our own primary research experiences (eg, Mantell et al [25]), and a broader synthesis of relevant theory and literature (eg, self-determination theory). Given that this is still a relatively new area of scientific inquiry, we hope that these lessons can serve as guidance for clinicians, researchers, and game developers who find themselves navigating the complexity of creating GBCAs that are both clinically effective and offer "good enough" user experiences and usability. More broadly, we hope this review can support and inform more evidence-based and rigorous design, development, and evaluation of GBCAs for older users in the future.

Validity Is Very Necessary but Not Sufficient

A game may function as an assessment tool in theory, but if it is not usable or engaging, it is unlikely to work in practice. Input from both users and game designers, at the outset and throughout the design process, is the best way to ensure a game is suitable. This is usually best done alongside an expert reference group working to ensure the game has psychometric validity [13]. The more the target users, developers, and cognitive experts collaborate, the more likely the game will succeed. Furthermore, researchers should always plan to evaluate game usability, as well as user experiences and preferences, to formally investigate the success of their game (and identify any scope for improvement).

Clarify the Intended User

There is a complex relationship between cognitive decline, usability, and user experience as people age. Many older people with cognitive impairment will respond better to simple gamification of realistic tasks. More complex serious games risk becoming less appealing as people become older or more impaired, and this will impact usability and user experience. However, the more rudimentary the game-based task, the less likely it is to engage younger adults or those with better cognitive function. To proactively address this tension, developers should begin with a critical question: who is this intervention for, and what exactly is the game intended to assess? For instance, "good enough" usability and user experience are probably good enough for a short and sporadic cognitive screener targeting older adults at risk of impairment. However, this may not be the case for a more expansive game trying to also engage younger cohorts and track more subtle cognitive changes over time.

Create Games That Align With the Unique Preferences and Needs of Older People

Older people appear to prefer games that avoid childlike features, are intellectually focused (eg, puzzles), and/or are based on the replication of real-life tasks or challenges. Avoiding fantasy, violent content, and multiplayer interaction is recommended. Large digital touchscreen devices, such as tablets, also seem to be more usable and intuitive for older people than smaller devices, mouses, or virtual reality [16,38,64]. However, tablets can also create ergonomic difficulties for older people who may be frail or have specific neurological conditions (eg, difficulty holding the device steadily, dyspraxia, or visuospatial impairment), and these factors need to be accounted for [18].

Providing Each User With Their Optimal Challenge Is Key

Tailoring the level of difficulty (known as difficulty balancing) to an individual's capacity should be attempted wherever possible in the design of GBCAs. It has been shown that this personalization does not need to be highly sophisticated to be useful in increasing user experience [37,39]. Introducing various levels or randomly generating different design elements or challenges per round can also make it possible to detect more subtle impairments and reduce learning effects [37]. In addition, it can also alleviate some of the tension between providing games that are suitable for older people or those with more impairment and maintaining sensitivity for more subtle cognitive decline, as more levels mean a lower floor and higher ceiling.

XSL•FO
RenderX

Acknowledgments

This work was funded through an Australian National Health and Medical Research Council Ideas Grant (grant GA143456; ASCAPE: Audio App-delivered Screening for Cognition and Age-related Health in Older Prisoners).

Conflicts of Interest

MK is the chief executive officer of Arludo.

Multimedia Appendix 1
PRISMA checklist.
[DOCX File , 22 KB - aging_v8i1e65252_app1.docx]

Multimedia Appendix 2
Risk of bias.
[DOCX File , 24 KB - aging_v8i1e65252_app2.docx]

Multimedia Appendix 3
Validity summary.
[DOCX File , 26 KB - aging_v8i1e65252_app3.docx]

Multimedia Appendix 4
User evaluation summary.
[DOCX File , 25 KB - aging_v8i1e65252_app4.docx]

References

1. Friehs MA, Dechant M, Vedress S, Frings C, Mandryk RL. Effective gamification of the stop-signal task: two controlled laboratory experiments. JMIR Serious Games 2020 Sep 08;8(3):e17810 [FREE Full text] [doi: 10.2196/17810] [Medline: 32897233]
2. Lumsden J, Edwards EA, Lawrence NS, Coyle D, Munafò MR. Gamification of cognitive assessment and cognitive training: a systematic review of applications and efficacy. JMIR Serious Games 2016 Jul 15;4(2):e11 [FREE Full text] [doi: 10.2196/games.5888] [Medline: 27421244]
3. Wiley K, Robinson R, Mandryk RL. The making and evaluation of digital games used for the assessment of attention: systematic review. JMIR Serious Games 2021 Aug 09;9(3):e26449 [FREE Full text] [doi: 10.2196/26449] [Medline: 34383674]
4. About dementia. Dementia-Australia. URL: https://www.dementia.org.au/information/about-dementia [accessed 2025-05-29]
5. Manning V. Managing cognitive impairment in AOD treatment guidelines. Turning Point. URL: https://www.turningpoint.org.au/treatment/clinicians/Managing-Cognitive-Impairment-in-AOD-Treatment-Guidelines [accessed 2025-05-29]
6. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
7. Pedersen MK, Díaz CM, Wang QJ, Alba-Marrugo MA, Amidi A, Basaiawmoit RV, et al. Measuring cognitive abilities in the wild: validating a population-scale game-based cognitive assessment. Cogn Sci 2023 Jun;47(6):e13308. [doi: 10.1111/cogs.13308] [Medline: 37354036]
8. National Collaborating Centre for Mental Health. Dementia: A NICE-SCIE Guideline on Supporting People With Dementia and Their Carers in Health and Social Care. London, UK: British Psychological Society; 2007.
9. Dorenkamp MA, Vik P. Assessment-related anxiety among older adults: development of a measure. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2021 Sep 27;28(5):781-795. [doi: 10.1080/13825585.2020.1826398] [Medline: 32985347]
10. Dorenkamp MA, Vik PW. Neuropsychological assessment anxiety: a systematic review. Pract Innov 2018 Sep;3(3):192-211. [doi: 10.1037/pri0000073]
11. Byrd DA, Rivera Mindt MM, Clark US, Clarke Y, Thames AD, Gammada EZ, et al. Creating an antiracist psychology by addressing professional complicity in psychological assessment. Psychol Assess 2021 Mar;33(3):279-285. [doi: 10.1037/pas0000993] [Medline: 33779204]
12. Tong T, Chignell M, Tierney MC, Lee J. A serious game for clinical assessment of cognitive status: validation study. JMIR Serious Games 2016 May 27;4(1):e7 [FREE Full text] [doi: 10.2196/games.5006] [Medline: 27234145]
13. Khaleghi A, Aghaei Z, Mahdavi MA. A gamification framework for cognitive assessment and cognitive training: qualitative study. JMIR Serious Games 2021 May 18;9(2):e21900 [FREE Full text] [doi: 10.2196/21900] [Medline: 33819164]

XSL•FO
RenderX

14. Leduc-McNiven K, White B, Zheng H, D McLeod R, R Friesen M. Serious games to assess mild cognitive impairment: 'the game is the assessment'. Res Rev Insights 2018;2(1):52. [doi: 10.15761/rri.1000128]
15. Zygouris S, Iliadou P, Lazarou E, Giakoumis D, Votis K, Alexiadis A, et al. Detection of mild cognitive impairment in an at-risk group of older adults: can a novel self-administered serious game-based screening test improve diagnostic accuracy? J Alzheimers Dis 2020;78(1):405-412 [FREE Full text] [doi: 10.3233/JAD-200880] [Medline: 32986676]
16. Valladares-Rodriguez S, Perez-Rodriguez R, Facal D, Fernandez-Iglesias MJ, Anido-Rifon L, Mouriño-Garcia M. Design process and preliminary psychometric study of a video game to detect cognitive impairment in senior adults. PeerJ 2017;5:e3508 [FREE Full text] [doi: 10.7717/peerj.3508] [Medline: 28674661]
17. Birk MV, Mandryk RL, Bowey J, Buttlar B. The effects of adding premise and backstory to psychological tasks. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 2024 Presented at: CHI '15; April 18-23, 2015; Seoul, Republic of Korea p. 2015 URL: http://gamification-research.org/wp-content/uploads/2014/11/GAMICHI15_birk_et_al.pdf
18. Tong T, Chignell M, Lam P, Tierney MC, Lee J. Designing serious games for cognitive assessment of the elderly. Proc Int Sympt Hum Factors Ergon Health Care 2014 Jul 22;3(1):28-35. [doi: 10.1177/2327857914031004]
19. Ferreira-Brito F, Fialho M, Virgolino A, Neves I, Miranda AC, Sousa-Santos N, et al. Game-based interventions for neuropsychological assessment, training and rehabilitation: which game-elements to use? A systematic review. J Biomed Inform 2019 Oct;98:103287 [FREE Full text] [doi: 10.1016/j.jbi.2019.103287] [Medline: 31518700]
20. Mantell R, Hwang YI, Radford K, Perkovic S, Cullen P, Withall A. Accelerated aging in people experiencing homelessness: a rapid review of frailty prevalence and determinants. Front Public Health 2023 Mar 16;11:1086215 [FREE Full text] [doi: 10.3389/fpubh.2023.1086215] [Medline: 37006541]
21. Withall A, Mantell R, Hwang YI. Background paper: issues facing older people who are leaving prison. Australian Association of Gerontology. 2022. URL: https://aag.asn.au/libraryviewer?ResourceID=74 [accessed 2025-05-29]
22. Astell AJ, Joddrell P, Groenewoud H, de Lange J, Goumans M, Cordia A, et al. Does familiarity affect the enjoyment of touchscreen games for people with dementia? Int J Med Inform 2016 Jul;91:e1-e8. [doi: 10.1016/j.ijmedinf.2016.02.001] [Medline: 26897552]
23. Chesham A, Wyss P, Müri RM, Mosimann UP, Nef T. What older people like to play: genre preferences and acceptance of casual games. JMIR Serious Games 2017 Apr 18;5(2):e8 [FREE Full text] [doi: 10.2196/games.7025] [Medline: 28420601]
24. Blocker KA, Wright TJ, Boot WR. Gaming preferences of aging generations. Gerontechnology 2014;12(3):174-184 [FREE Full text] [doi: 10.4017/gt.2014.12.3.008.00] [Medline: 29033699]
25. Mantell R, Withall A, Radford K, Kasumovic M, Monds L, Hwang YI. Design preferences for a serious game-based cognitive assessment of older adults in prison: thematic analysis. JMIR Serious Games 2023 Apr 17;11:e45467 [FREE Full text] [doi: 10.2196/45467] [Medline: 37067850]
26. de Paula DF, Menezes BH, Araújo CC. Building a quality mobile application: a user-centered study focusing on design thinking, user experience and usability. In: Proceedings of the 3rd International Conference on Design, User Experience, and Usability: User Experience Design for Diverse Interaction Platforms and Environments. 2014 Presented at: DUXU '14; June 22-27, 2014; Crete, Greece p. 313-322 URL: https://link.springer.com/chapter/10.1007/978-3-319-07626-3_29 [doi: 10.1007/978-3-319-07626-3_29]
27. ISO 9241-210:2019: ergonomics of human-system interaction: part 210: human-centred design for interactive systems. International Organization for Standardization. 2024. URL: https://www.iso.org/standard/77520.html [accessed 2025-05-29]
28. Hassenzahl M, Tractinsky N. User experience - a research agenda. Behav Inf Technol 2006 Mar;25(2):91-97. [doi: 10.1080/01449290500330331]
29. Gaver B, Martin H. Alternatives: exploring information appliances through conceptual design proposals. In: Proceedings of the 2000 SIGCHI conference on Human Factors in Computing Systems. 2000 Presented at: CHI '00; April 1-6, 2000; The Hague, The Netherlands p. 209-216 URL: https://dl.acm.org/doi/10.1145/332040.332433 [doi: 10.1145/332040.332433]
30. Nielsen J. Enhancing the explanatory power of usability heuristics. In: Proceedings of the 1994 ACM Conference on Human Factors in Computer Systems. 1994 Presented at: CHI '94; April 24-28, 1994; Boston, MA p. 152-158 URL: https://static.aminer.org/pdf/PDF/000/089/679/enhancing_the_explanatory_power_of_usability_heuristics.pdf
31. Kletzel SL, Sood P, Negm A, Heyn PC, Krishnan S, Machtinger J, et al. Effectiveness of brain gaming in older adults with cognitive impairments: a systematic review and meta-analysis. J Am Med Dir Assoc 2021 Nov;22(11):2281-8.e5 [FREE Full text] [doi: 10.1016/j.jamda.2021.05.022] [Medline: 34146517]
32. Ning H, Li R, Ye X, Zhang Y, Liu L. A review on serious games for dementia care in ageing societies. IEEE J Transl Eng Health Med 2020;8:1-11. [doi: 10.1109/jtehm.2020.2998055]
33. Sabermahani F, Almasi-Dooghaee M, Sheikhtaheri A. Alzheimer's disease and mild cognitive impairment serious games: a systematic analysis in smartphone application markets. Stud Health Technol Inform 2022 May 16;293:179-186. [doi: 10.3233/SHTI220366] [Medline: 35592979]
34. Valladares-Rodríguez S, Pérez-Rodríguez R, Anido-Rifón L, Fernández-Iglesias M. Trends on the application of serious games to neuropsychological evaluation: a scoping review. J Biomed Inform 2016 Dec;64:296-319 [FREE Full text] [doi: 10.1016/j.jbi.2016.10.019] [Medline: 27815228]

XSL•FO
RenderX

35. Mezrar S, Bendella F. A systematic review of serious games relating to cognitive impairment and dementia. J Digit Inf Manag 2022;20:1. [doi: 10.6025/jdim/2022/20/1/01-09]
36. Sleith C. Methodology checklist 4: case control studies healthcare improvement Scotland. Scottish Intercollegiate Guidelines Network. 2012. URL: https://www.sign.ac.uk/using-our-guidelines/methodology/checklists/ [accessed 2025-05-29]
37. Zygouris S, Segkouli S, Triantafyllidis A, Giakoumis D, Tsolaki M, Votis K, et al. Usability of the virtual supermarket test for older adults with and without cognitive impairment. J Alzheimers Dis Rep 2022 May 02;6(1):229-234. [doi: 10.3233/adr-210064]
38. Cabinio M, Rossetto F, Isernia S, Saibene FL, Di Cesare M, Borgnis F, et al. The use of a virtual reality platform for the assessment of the memory decline and the hippocampal neural injury in subjects with mild cognitive impairment: the validity of smart aging serious game (SASG). J Clin Med 2020 May 06;9(5):1355 [FREE Full text] [doi: 10.3390/jcm9051355] [Medline: 32384591]
39. Nef T, Chesham A, Schütz N, Botros AA, Vanbellingen T, Burgunder J, et al. Development and evaluation of maze-like puzzle games to assess cognitive and motor function in aging and neurodegenerative diseases. Front Aging Neurosci 2020 Apr 21;12:87 [FREE Full text] [doi: 10.3389/fnagi.2020.00087] [Medline: 32372942]
40. Chesham A, Gerber SM, Schütz N, Saner H, Gutbrod K, Müri RM, et al. Search and match task: development of a Taskified match-3 puzzle game to assess and practice visual search. JMIR Serious Games 2019 May 09;7(2):e13620 [FREE Full text] [doi: 10.2196/13620] [Medline: 31094325]
41. Manera V, Petit P, Derreumaux A, Orvieto I, Romagnoli M, Lyttle G, et al. 'Kitchen and cooking,' a serious game for mild cognitive impairment and Alzheimer's disease: a pilot study. Front Aging Neurosci 2015 Mar 17;7:24 [FREE Full text] [doi: 10.3389/fnagi.2015.00024] [Medline: 25852542]
42. Vallejo V, Wyss P, Rampa L, Mitache AV, Müri RM, Mosimann UP, et al. Evaluation of a novel serious game based assessment tool for patients with Alzheimer's disease. PLoS One 2017 May 4;12(5):e0175999 [FREE Full text] [doi: 10.1371/journal.pone.0175999] [Medline: 28472049]
43. Wang CS, Wu J, Hsu W, Chien P, Chen P, Huang Y, et al. Using self-administered game-based cognitive assessment to screen for degenerative dementia: a pilot study. J Alzheimers Dis 2022;86(2):877-890. [doi: 10.3233/JAD-215142] [Medline: 35147533]
44. Manning V, Piercy H, Garfield JB, Clark SG, Andrabi MN, Lubman DI. A personalized approach bias modification smartphone app ("SWiPE") to reduce alcohol use: open-label feasibility, acceptability, and preliminary effectiveness study. JMIR Mhealth Uhealth 2021 Dec 10;9(12):e31353 [FREE Full text] [doi: 10.2196/31353] [Medline: 34890355]
45. McWilliams EC, Barbey FM, Dyer JF, Islam MN, McGuinness B, Murphy B, et al. Feasibility of repeated assessment of cognitive function in older adults using a wireless, mobile, Dry-EEG headset and tablet-based games. Front Psychiatry 2021 Jun 25;12:574482 [FREE Full text] [doi: 10.3389/fpsyt.2021.574482] [Medline: 34276428]
46. Palmisano S, Allison RS, Kim J. Cybersickness in head-mounted displays is caused by differences in the user's virtual and physical head pose. Front Virtual Real 2020 Nov 12;1:66. [doi: 10.3389/frvir.2020.587698]
47. Paliokas I, Kalamaras E, Votis K, Doumpoulakis S, Lakka E, Kotsani M, et al. Using a virtual reality serious game to assess the performance of older adults with frailty. Adv Exp Med Biol 2020;1196:127-139. [doi: 10.1007/978-3-030-32637-1_13] [Medline: 32468314]
48. Zygouris S, Giakoumis D, Votis K, Doumpoulakis S, Ntovas K, Segkouli S, et al. Can a virtual reality cognitive training application fulfill a dual role? Using the virtual supermarket cognitive training application as a screening tool for mild cognitive impairment. J Alzheimers Dis 2015;44(4):1333-1347. [doi: 10.3233/JAD-141260] [Medline: 25428251]
49. Zucchella C, Sinforiani E, Tassorelli C, Cavallini E, Tost-Pardell D, Grau S, et al. Serious games for screening pre-dementia conditions: from virtuality to reality? A pilot project. Funct Neurol 2014;29(3):153-158 [FREE Full text] [Medline: 25473734]
50. Boot WR, Champion M, Blakely DP, Wright T, Souders DJ, Charness N. Video games as a means to reduce age-related cognitive decline: attitudes, compliance, and effectiveness. Front Psychol 2013;4:31 [FREE Full text] [doi: 10.3389/fpsyg.2013.00031] [Medline: 23378841]
51. Lee Y, Kozar KA, Larsen KR. The technology acceptance model: past, present, and future. Commun Assoc Inf Syst 2003;12:50. [doi: 10.17705/1cais.01250]
52. Rivas Costa C, Fernández Iglesias MJ, Anido Rifón LE, Gómez Carballa M, Valladares Rodríguez S. The acceptability of TV-based game platforms as an instrument to support the cognitive evaluation of senior adults at home. PeerJ 2017;5:e2845 [FREE Full text] [doi: 10.7717/peerj.2845] [Medline: 28070464]
53. Gourlan M, Sarrazin P, Trouilloud D. Motivational interviewing as a way to promote physical activity in obese adolescents: a randomised-controlled trial using self-determination theory as an explanatory framework. Psychol Health 2013 Nov;28(11):1265-1286 [FREE Full text] [doi: 10.1080/08870446.2013.800518] [Medline: 23756082]
54. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 1988 Jun;54(6):1063-1070. [doi: 10.1037//0022-3514.54.6.1063] [Medline: 3397865]
55. Gledhill JA, Rodary C, Mahé C, Laizet C. [French validation of the revised Piper Fatigue Scale]. Rech Soins Infirm 2002 Mar(68):50-65. [Medline: 12001628]

XSL•FO
RenderX

56. Bottiroli S, Cavallini E. Can computer familiarity regulate the benefits of computer-based memory training in normal aging? A study with an Italian sample of older adults. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2009 Jul 29;16(4):401-418. [doi: 10.1080/13825580802691763] [Medline: 19253069]
57. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry. Boca Raton, FL: CRC Press; 1996:189-194.
58. Brooke J. SUS: a retrospective. J Usability Stud 2013;8:29-40 [FREE Full text]
59. Bangor AP, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4:114-123 [FREE Full text]
60. Joddrell P, Astell AJ. Implementing accessibility settings in touchscreen apps for people living with dementia. Gerontology 2019 Apr 24;65(5):560-570. [doi: 10.1159/000498885] [Medline: 31018201]
61. Przybylski AK, Rigby CS, Ryan RM. A motivational model of video game engagement. Rev Gen Psychol 2010 Jun 01;14(2):154-166. [doi: 10.1037/a0019440]
62. Uysal A, Yildirim IG. Self-determination theory in digital games. In: Bostan B, editor. Gamer Psychology and Behavior. Cham, Switzerland: Springer; 2016:123-135.
63. Vaz de Carvalho C. Dynamic serious games balancing. In: Proceedings of the 5th International Conference Serious Games, Interaction, and Simulation. 2016 Presented at: SGAMES '15; September 16-18, 2015; Novedrate, Italy p. 33 URL: https://eudl.eu/doi/10.1007/978-3-319-29060-7_5 [doi: 10.1007/978-3-319-29060-7_5]
64. Plechatá A, Sahula V, Fayette D, Fajnerová I. Age-related differences with immersive and non-immersive virtual reality in memory assessment. Front Psychol 2019 Jun 11;10:1330 [FREE Full text] [doi: 10.3389/fpsyg.2019.01330] [Medline: 31244729]

Abbreviations

AD: Alzheimer's disease
CVLT: California Verbal Learning Test
GBCA: game-based cognitive assessment
ISO: International Organization for Standardization
MCI: mild cognitive impairment
NL: Numberlink
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SASG: Smart Aging Serious Game
SMT: search and match task
SUS: system usability scale
VST: Virtual Supermarket Test

Edited by R Yang; submitted 09.08.24; peer-reviewed by M Friehs, N Vita-Barrull, A Katz; comments to author 02.01.25; revised version received 11.02.25; accepted 29.03.25; published 11.06.25.

Please cite as:
Mantell R, Hwang YIJ, Dark M, Radford K, Kasumovic M, Monds L, Schofield PW, Butler T, Withall A
Evaluating the User Experience and Usability of Game-Based Cognitive Assessments for Older People: Systematic Review
JMIR Aging 2025;8:e65252
URL: https://aging.jmir.org/2025/1/e65252
doi:10.2196/65252
PMID:

©Rhys Mantell, Ye In Jane Hwang, Matthew Dark, Kylie Radford, Michael Kasumovic, Lauren Monds, Peter W Schofield, Tony Butler, Adrienne Withall. Originally published in JMIR Aging (https://aging.jmir.org), 11.06.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Dynamic Interactions Among Sleep Duration, Cognitive Function, and Depressive Symptoms in Middle-Aged and Older Chinese Adults: Temporal Network Analysis From CHARLS

Furong Chen[1*], PhD; Jiaying Li[2*], PhD; Junchen Guo[3], PhD; Ying Xiong[4], MD; Zengjie Ye[1], Prof Dr

[1]School of Nursing, Guangzhou Medical University, No.1 Xinzao Road, Panyu District, Guangzhou, Guangdong Province, China

[2]The Nethersole School of Nursing, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong SAR, China

[3]School of Nursing, University of Wollongong, Wollongong, Australia

[4]School of Nursing, Guangzhou University of Chinese Medicine, Guangzhou, Guangdong Province, China

[*]these authors contributed equally

Corresponding Author:
Zengjie Ye, Prof Dr
School of Nursing, Guangzhou Medical University, No.1 Xinzao Road, Panyu District, Guangzhou, Guangdong Province, China

Abstract

Background: While bidirectional associations among sleep duration, cognitive function, and depression are established, the symptom-level temporal interactions among these factors in China's aging population, which is experiencing unprecedented growth, remain poorly characterized.

Objective: We aim to use a novel temporal network analysis to clarify these dynamics and guide targeted interventions, with a focus on sex-specific dynamic pathways.

Methods: We conducted a longitudinal temporal network analysis on 3136 Chinese adults aged ≥45 years from the China Health and Retirement Longitudinal Study (CHARLS) across 5 waves (2011, 2013, 2015, 2018, and 2020). A graphical vector autoregressive (GVAR) model delineated the interdependencies among sleep duration, cognitive performance (assessed via the Mini-Mental State Examination [MMSE]), and depressive symptoms (evaluated with the 10-item Center for Epidemiologic Studies Depression Scale [CESD-10]). We also examined sex-specific differences in network structures.

Results: The symptom "bothered" was found to predict all other CESD-10 symptoms. There were significant predictive links between sleep and the CESD-10 node (ie, bothered, drained, and depressed), along with sleep and the MMSE functions (ie, numerical ability). Furthermore, sleep duration served as a bridge between depression symptoms and cognitive functions. There were significant differences in longitudinal network structure between sexes. Sex-specific analyses revealed distinct network patterns. Among female participants, the "bothered" node significantly predicted several outcomes over time. In contrast, the temporal network for male participants was sparser, with the "stuck" node in the depression domain being predominantly influenced by other nodes.

Conclusions: Our study revealed that emotional distress, especially the "bothered" symptom, plays a central role in depressive symptoms and cognitive decline. The bridging effect of short sleep duration underscores the potential of interventions targeting both sleep and emotional distress for alleviating depressive symptoms and delaying cognitive deterioration in older adults.

(JMIR Aging 2025;8:e76210) doi:10.2196/76210

KEYWORDS

CHARLS; sleep; cognition; depression; temporal network analysis; middle-aged and older people; China Health and Retirement Longitudinal Study

Introduction

The population is aging at an unprecedented rate in China, driven by socioeconomic advancements and improved health care [1]. This demographic shift poses critical challenges to sustaining the cognitive well-being and mental health of older adults [2]. Studies have reported that the prevalence of depressive symptoms in the middle-aged and older population in China ranges from 24.1% to 31.8% [3,4], and the prevalence of cognitive impairment is 26.7% [5]. This population also experiences more frequent and severe sleep issues [6]. For the middle-aged and older population, these 3 elements could significantly affect their physical and psychological status and quality of life. Therefore, clarification of the relationship among the 3 could facilitate the development of targeted interventions and improvement of health care policies.

XSL•FO
RenderX

Impaired cognition and sleep disturbances are prevalent in aging populations, and accumulating evidence suggests that sleep deficits may mediate up to 15% of cognitive decline [7-9]. In addition, depression is a common comorbidity in older adults, which further exacerbates cognitive risks [10]. The dual impact of sleep deprivation and mood dysregulation may compromise hippocampal-dependent cognition, while sleep fragmentation can further amplify the hyperactivity of the limbic system characteristic of depression [11-13]. Despite these advances, several limitations persist in existing studies. The relationship between depression and cognitive impairment is complex and bidirectional, yet its predictive nature remains unclear. Moreover, most studies have relied on cross-sectional designs, which do not capture the temporal directionality of associations among sleep, cognition, and depression. While studies by Zhou et al [11] and Sun et al [14] have investigated bidirectional cognition-depression links and sleep-depression interactions in Chinese cohorts, the potential simultaneous interactions and between-subject effects among these 3 variables remain insufficiently explored. Furthermore, although previous research has investigated sex differences in these relationships [15-18], further investigation into sex-based differences in middle-aged and older populations using longitudinal predictive frameworks is needed to address gaps in current research.

Traditional analytical approaches, while establishing bidirectional associations among sleep, cognition, and depression, often fail to capture symptom-level dynamics. Temporal network analysis, a novel approach in psychosomatic research, offers a dynamic framework to model bidirectional and triadic relationships over time [19]. Unlike traditional cross-lagged panel models, temporal network analysis distinguishes between within-person (lagged) and between-person (contemporaneous) effects, offering greater mechanistic clarity than conventional cross-sectional or longitudinal designs [20]. In this study, temporal network analysis enables us to examine how sleep, cognitive function, and depressive symptoms interact and predict each other over time, while also disentangling within-person and between-person effects.

This study aims to explore the dynamic connections between sleep duration, cognitive function, and depressive symptoms in middle-aged and older Chinese adults using large-scale longitudinal data and temporal network analysis. We hypothesized that (1) emotional distress is associated with increased depression symptoms, (2) sleep duration acts as a link between depression and cognitive decline, and (3) there are significant sex-specific differences in those temporal networks.

Methods

Study Design and Data Source

This temporal network analysis used longitudinal data from the China Health and Retirement Longitudinal Study (CHARLS), a nationally representative survey of Chinese adults aged 45 years and older. Funded by Peking University, CHARLS uses a stratified, probability-proportional-to-size sampling method to ensure comprehensive geographic coverage. The survey spanned 28 provinces, including autonomous regions and municipalities directly governed by the central government, covering 150 counties and 450 communities.

This analysis used data from 5 waves collected in 2011, 2013, 2015, 2018, and 2020, initially recruiting 17,705 participants in 2011. Participants were excluded if they had (1) missing baseline data on cognition or depression, (2) missing demographic variables, (3) age <45 years at baseline, and (4) missing follow-up data on cognition and depression.

Measurement

Demographic Variables

The participants provided information regarding their age, sex, education level, retirement status, and marital status.

Depression Assessment

Depressive symptoms were measured using the 10-item Center for Epidemiologic Studies Depression Scale (CESD-10) [21]. This instrument was well validated, demonstrating excellent interrater reliability and specificity in Chinese middle-aged and older Chinese adults [22,23]. The scale comprised 10 items, with the final score calculated as the sum of individual item scores, ranging from 0 to 30. A total score of 10 or greater was used to indicate the presence of depressive symptoms. Measurements were conducted by trained CHARLS interviewers during face-to-face home visits as part of the core questionnaire.

Cognition Assessment

Cognitive function was evaluated using the Mini-Mental State Examination (MMSE) [24]. This assessment focused on 2 primary domains: episodic memory and mental status. Episodic memory was assessed using a 10-word recall test. Participants were read a random word list and asked to immediately recall as many words as possible (immediate recall score: 0 - 10). After a 10-minute interval, delayed recall of the same list was tested (delayed recall score: 0 - 10). The episodic memory score (range 0 - 10) was calculated as the average of these 2 trials. Mental intactness was evaluated using adapted components from the Telephone Interview for Cognitive Status (TICS). This included serial-7 subtractions from 100 (maximum 5 trials), temporal orientation (date, day, and season), and intersecting pentagon copying. Responses were summed into a composite score ranging 0 - 11. The global cognitive score (range 0 - 21) represented the sum of episodic memory and mental intactness scores. All cognitive tests were administered directly by trained CHARLS interviewers during in-home visits following standardized protocols.

Sleep Duration

Nighttime sleep duration was assessed by asking respondents, “On average, how many hours per night did you actually sleep in the past month?” Responses were recorded verbatim by interviewers during face-to-face home visits.

Statistical Analyses

Overall, the analytical approach comprised several key steps designed to rigorously model and interpret the longitudinal dynamics among sleep duration, cognitive function, and depressive symptoms. We first preprocessed the data to ensure normality and stationarity, then selected appropriate nodes to

XSL•FO
RenderX

represent distinct constructs, and finally used a graphical vector autoregressive (GVAR) model for network estimation. All analyses were conducted using R (version 4.4.1; R Core Team).

Data Preparation

Before performing the network analyses, we undertook a series of data preprocessing steps to ensure that the assumptions underlying dynamic network models were satisfied [25,26]. First, we confirmed that each participant provided enough observations for the longitudinal analysis. Next, to verify that our data were suitable for parametric analysis, we applied the Shapiro-Wilk normality test. This test indicated that the distribution of our key variables did not deviate substantially from normality; given the large sample size, this justified the use of methods that assume normally distributed data. Finally, we evaluated the stationarity of the time series data using the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test [26]. Stationarity was essential for dynamic models because it ensures that the statistical properties of the variables (such as mean and variance) remain constant over time. In our analysis, the KPSS test results were all nonsignificant (False), confirming that the data met the stationarity assumption required for reliable longitudinal analysis.

Node Selection via Redundancy Checking

First, to avoid redundancy, we used the 5 subscales of the MMSE as separate nodes in the network analysis, rather than a single item, since the entries under these dimensions share similar characteristics [27]. Second, each depression symptom and the sleep duration score were treated as individual nodes to provide a comprehensive representation of these factors. In total, 16 nodes (10 depression symptoms, 5 MMSE subscales, and 1 sleep scale) were included in the network based on the available data. Node redundancy was assessed at each time point using the goldbricker function [27], with no suggested reductions. The node abbreviations, description, and items for CESD-10, MMSE, and Sleep are provided in Table S1 in Multimedia Appendix 1.

Network Estimation and Visualization

The GVAR model was constructed using the psychonetrics package to explore the longitudinal relationships between the study variables [28].

The panel GVAR captures temporal dependencies as partial correlations between the deviations from the individual mean of a variable at one time point and the deviations at the following time point, while adjusting for all variables at the preceding time point. This creates a matrix of regression coefficients that can be used to plot a directed lag-1 network (ie, the temporal or longitudinal network), which represents the generalized temporal within-subject effects between variables [28]. The panel GVAR also estimates 2 additional networks through 2 Gaussian graphical models (GGMs)—a contemporaneous network that interprets the relationships among variables within a person at a specific time point, adjusting for temporal effects, and a between-person network that captures associations between the stationary means across all participants.

The GVAR model was estimated using full information maximum likelihood (FIML) estimation. We searched for the optimal model (ie, pruned model) by minimizing the Bayesian information criterion and threshold at α=.05. To assess model fit, we used the confirmatory fit index (CFI) and the Tucker-Lewis index (TLI) and the root mean squared error of approximation (RMSEA), with values >0.85 (CFI and TLI) and <0.05 (RMSEA) indicating good fit [29].

The psychonetrics package uses the qgraph package [30] to create graphical network models displaying the estimated coefficients. For both contemporaneous and between-subject networks, the conservative “AND-rule” was used to retain significant edges, meaning that an edge was included only if both regressions it relied on were significant. Node centrality is interpreted through visual inspection of the network structures [31].

Ethical Considerations

This study used publicly available data from the CHARLS survey, without any new sample recruitment or additional data collection. In that survey, informed consent was obtained from all participants. During the face-to-face household interviews, trained interviewers thoroughly explained the study objectives, interview content, potential risks and benefits, data confidentiality protocols, and the voluntary nature of participation (including the right to refuse or withdraw at any stage). Participants provided written informed consent. For participants who were illiterate, the consent form was read aloud in the presence of a witness, and participants provided a fingerprint or signature as confirmation. The CHARLS study underwent a thorough review and received approval from the Ethics Review Board of Peking University (IRB00001052-11015), ensuring adherence to the highest ethical standards.

The analysis in this study used this de-identified, publicly available data that had already received institutional ethics approval for the original data collection. Our analysis falls within the scope of research purposes for which the data were originally collected and approved. In addition, co-author Junchen Guo obtained formal authorization to access and use the CHARLS dataset through the official CHARLS website. All data usage complies with the CHARLS dataset’s terms of use and citation requirements established by the National School of Development, Peking University.

Results

Demographic Characteristics of Participants

After applying the inclusion criteria, the final analytic sample comprised 3136 individuals (refer to Figure 1 for the screening flowchart). Among the 3136 eligible individuals, ages ranged from 45 to 81 years, with a mean age of 55.73 (SD 7.47) years. In addition, 56.8% (1782/3136) of the participants were male. Detailed demographics and characteristics are presented in Table 1.

XSL•FO
RenderX

Figure 1. Study flowchart. The figure shows 2 different types of missing data: Within-wave missing data: for example, in Wave 3, 1628 participants had missing cognition or depression data within that wave. Between-wave attrition: 2147 participants were lost between Wave 3 and Wave 4. The mathematics is correct: Wave 3 (5799) - attrition to Wave 4 (2147) = Wave 4 participants (3652). Similarly, Wave 4 (3652) - attrition to Wave 5 (516) = Wave 5 participants (3136). The numbers we referenced (1628 for Wave 3) represent missing data within that specific wave, not the attrition between waves.

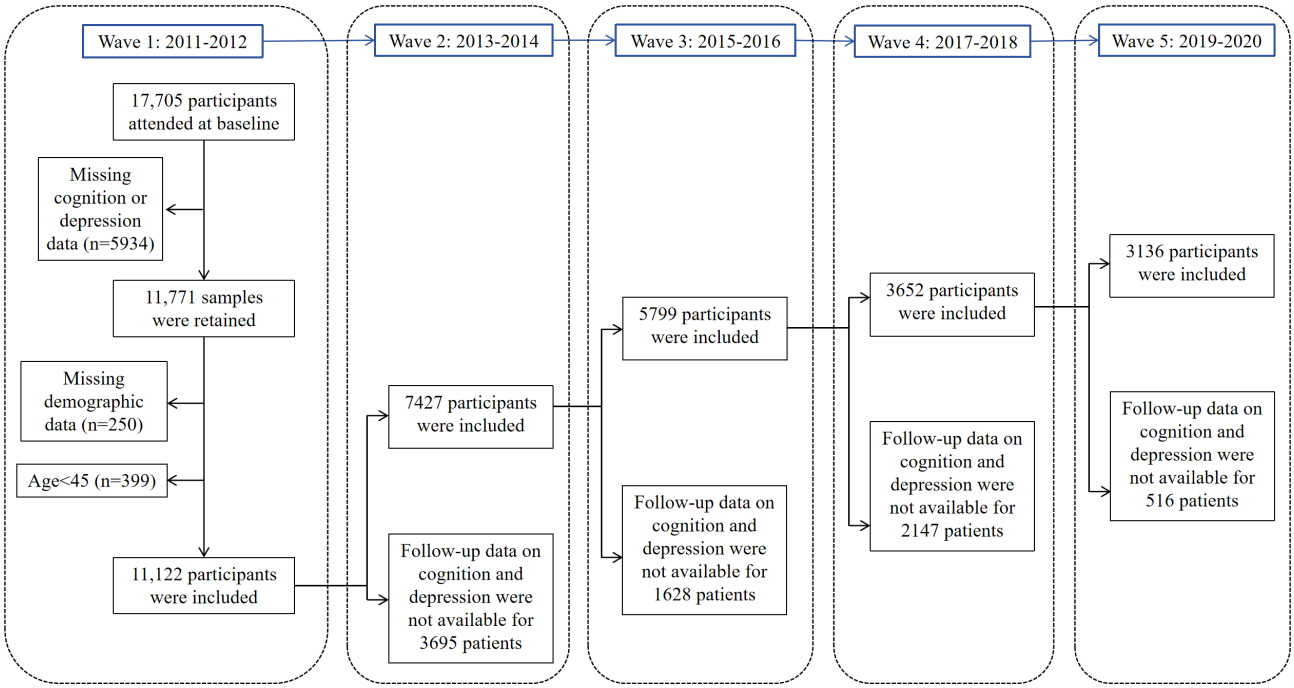


Table . Demographics and characteristics of the sample (N=3136).

Characteristics	Values
Age (years), mean (SD)	55.73 (7.47)
Sex, n (%)	
Male	1782 (56.8)
Female	1354 (43.2)
Marital status, n (%)	
Married	2963 (94.5)
Unmarried	12 (0.4)
Divorced	22 (0.7)
Widowed	139 (4.4)
Education, n (%)	
Primary school and below	1363 (43.5)
Junior high school	1098 (35)
High school or Middle professional school	598 (19.1)
University and above	77 (2.4)
Retirement status	
Yes	838 (26.7)
No	2298 (73.3)

Model Fitting Results

We selected the step-up model as our final framework, with detailed fit indices presented in Table 2. The model demonstrated acceptable fit to the data (χ^2_{3135}=9373.40; *P*<.001; RMSEA=0.025; 95% CI 0.025-0.026; CFI=0.89; TLI=0.89). Complete parameter matrices for the estimated temporal, contemporaneous, and between-subject networks are provided in Tables S2-S4 in Multimedia Appendix 1, and Figures S1-S3 in Multimedia Appendices 2-4 illustrate the node centrality of these networks.

XSL•FO
RenderX

Table . Model fit indices across model types.

Model type	χ^2 (*df*)	RMSEA[a] (95% CI)	CFI[b]	TLI[c]	AIC[d]	BIC[e]
Baseline	7712.06 (2776)	0.024 (0.023 - 0.024)	0.89	0.87	656,172.84	659,464.43
Pruned (α=.05)	11,638.07 (3171)	0.029 (0.029 - 0.030)	0.81	0.80	659,308.86	660,210.41
Stepup	9373.40 (3135)	0.025 (0.025 - 0.026)	0.89	0.89	657,116.18	658,235.56

[a]RMSEA: root mean squared error of approximation.

[b]CFI: confirmatory fit index.

[c]TLI: Tucker-Lewis index.

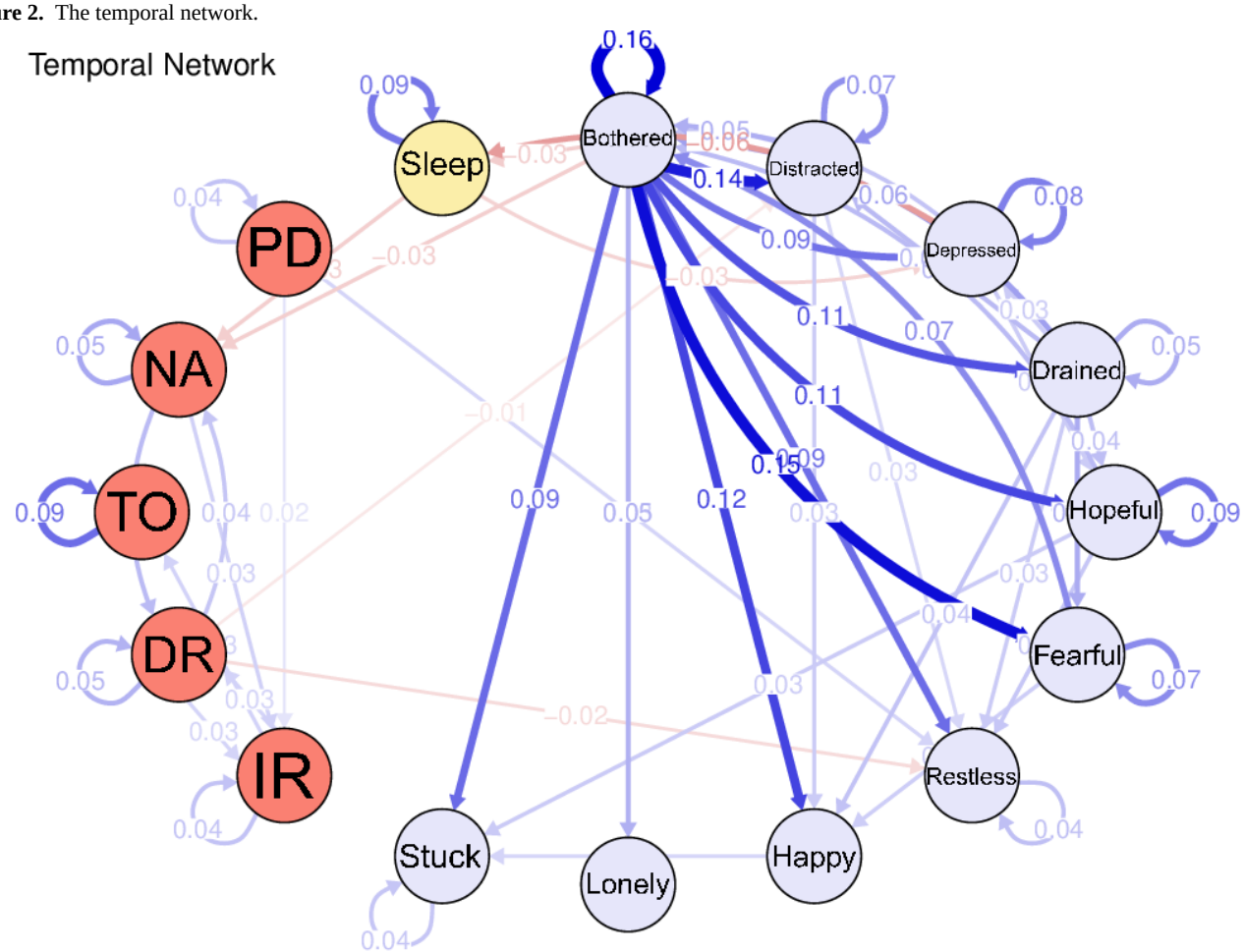
[d]AIC: Akaike information criterion.

[e]BIC: Bayesian information criterion.

Temporal Network

Figure 2 illustrates the temporal network, highlighting the within-subject temporal predictions between nodes. Bothered ("I was bothered by things that usually don't bother me.") was the core node, predicting all other nodes in the CESD-10 scale positively and significantly. Moreover, bothered had a negative and significant effect on "NA" (numerical ability, β=−.030) in the MMSE scale. Higher Bothered (β=−.031) and Depressed ("I felt depressed.", β=−.062) predicted shorter sleep duration. In addition, longer sleep duration predicted lower levels of NA (β=−.031).

Figure 2. The temporal network.



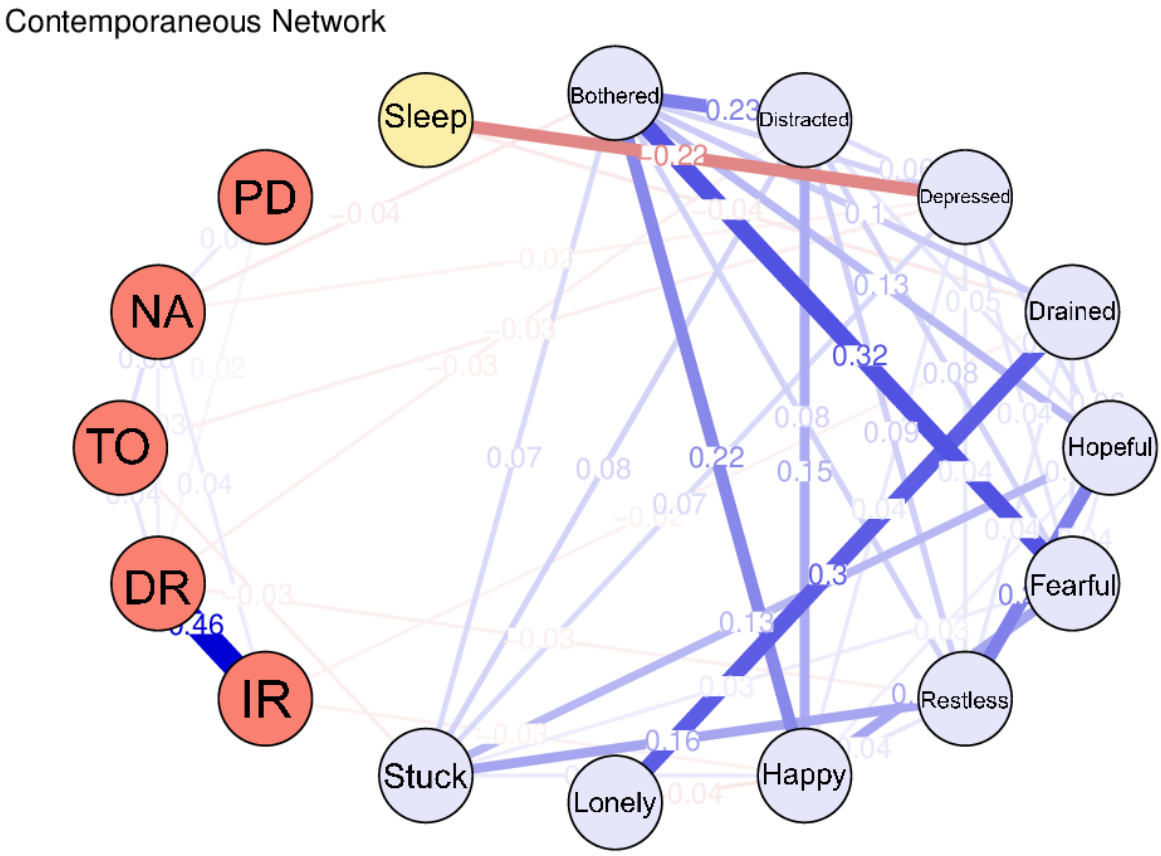
Contemporaneous Network

The contemporaneous network (Figure 3; Table S3 in Multimedia Appendix 1) reflects cross-sectional associations at individual time points after adjusting for temporal effects. Most connections occurred within the same measurement domain; however, 2 notable cross-domain links were observed. Specifically, the "depressed" node in the depression symptom

XSL•FO
RenderX

group was negatively correlated with the time orientation node (r=−0.029), the NA node in cognition (r=−0.023), and the sleep duration node (r=−0.221). In addition, drained ("I felt that everything I did was an effort.") in CESD-10 correlated negatively with immediate recall (r=−0.023) in MMSE and sleep duration (r=−0.038).

Figure 3. The contemporaneous network.

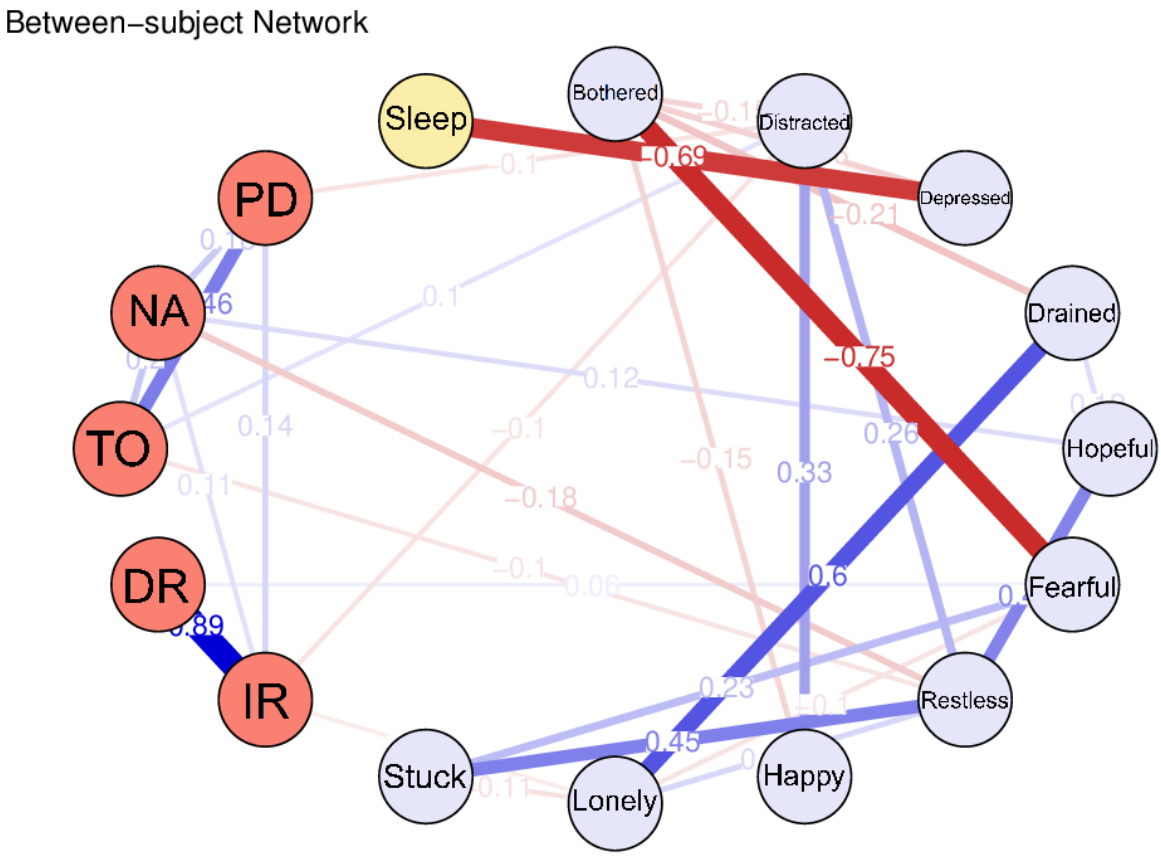


Between-Subject Network

In the between-subject network (Figure 4), the "bothered" node exhibited the most extensive interscale associations. For example, this node was negatively correlated with the "fearful" (r=−0.749) and "drained" (r=−0.208) nodes within the depression symptoms. In addition, the cognition "NA" (numerical ability) node was positively associated with the depression "hopeful" node (r=0.121) and negatively associated with the "restless" node (r=−0.179). Notably, the "sleep duration" node maintained a negative association with the depression "depressed" node (r=−0.692).

XSL•FO
RenderX

Figure 4. The between-subject network.



Sex Differences

Sex-specific analyses revealed distinct network patterns. Among female participants (Multimedia Appendix 5), the "bothered" node significantly predicted several outcomes over time, including the "fearful" node (β=.149) in the depression domain. In contrast, the temporal network for male participants (Multimedia Appendix 6) was sparser, with the "stuck" node in the depression domain being predominantly influenced by other nodes. The contemporaneous networks (Multimedia Appendices 7 and 8) demonstrated similar cross-sectional linkages across sexes. Interestingly, the between-person network for males (Multimedia Appendix 9 for male participants; Multimedia Appendix 10 for female participants) exhibited atypical associations, whereby higher "bothered" scores were linked with greater levels of "happy" (r=0.374) and "hopeful" (r=0.270). Details are provided in Tables S5-S10 in Multimedia Appendix 1.

Discussion

Principal Findings

To the best of our knowledge, this is the first predictive longitudinal study of sleep, cognition, and depression in a large Chinese middle-aged and older population.

In our temporal network analysis, the node "bothered" emerged as central, significantly predicting all other depressive symptoms and showing a negative association with numerical ability—a key cognitive measure. This finding reinforces the central role of emotional distress in depression [32]. Although the corresponding β value is low, this reflects the nature of symptoms as collections of moment-by-moment experiences, with these micro-level processes forming the foundation of psychopathology. While the effect of a single instance may be limited, the cumulative reduction in "bothered" moments, whether through mindfulness or environmental adjustments, may lead to clinically meaningful improvements in negative affect.

Moreover, both "bothered" and "depressed" were negatively associated with shorter sleep duration, aligning with previous findings [33]. Intriguingly, longer sleep duration predicted lower numerical ability, perhaps because middle-aged and older individuals, despite extended sleep periods, often experience frequent nocturnal awakenings that compromise sleep quality [34]. This impaired sleep quality may contribute to accelerated atrophy in the frontal, temporal, and parietal regions—areas critical for memory and numerical processing [35,36], ultimately leading to cognitive decline [37].

In addition, these results suggest that sleep may serve as a crucial intermediary between depression and cognitive function. Depression can adversely affect cognition directly or indirectly through elevated harmful cytokine levels [38,39] and reduced hippocampal volume [40]. Previous studies have highlighted that the association between sleep duration and cognitive function is substantially influenced by the severity of depressive symptoms [12,13,41]. Given that depressive symptoms are

XSL•FO
RenderX

closely associated with sleep disturbances [42] and that inadequate sleep may lower the secretion of brain-derived neurotrophic factor, thereby impairing cognitive ability [43]. Similar findings have been observed in US adolescent populations, showing that functional connectivity between the cortex and basal ganglia mediates the effects of insufficient sleep on depression and thought problems, while structural properties of the anterior temporal lobe mediate its effects on crystallized intelligence [44]. Therefore, improving sleep quality might not only alleviate emotional distress and depressive symptoms but also play an important role in delaying cognitive decline.

The contemporaneous network analysis revealed that depressive symptoms not only predict other symptoms over time but also exert an immediate influence on cognitive function and sleep quality. These findings reinforce the notion that depression is closely linked with declines in memory, attention, and overall cognitive performance [45]. A cohort study of older adults in the United Kingdom similarly demonstrated that reduced sleep was associated with depression and that baseline depression predicted subsequent sleep reduction [46]. Moreover, depression appears to diminish an individual's adaptability to external changes, thereby exacerbating sleep disturbances [42]. In the between-subject network, the "bothered" node emerged as a critical hub connecting symptom dissemination across individuals, suggesting that social factors may influence the spread of emotional distress among middle-aged and older adults [47].

In the temporal network specific to females, the "bothered" node was the most central emotion, significantly impacting other depressive symptoms. This observation aligns with evidence that women tend to express emotions more readily and may be more vulnerable to negative affective states [48]. These patterns are further amplified by Chinese cultural norms of emotional labor, where women bear primary responsibility for familial emotional regulation, intensifying affective sensitivity through sustained practice. Conversely, in men, depressive symptoms appeared to manifest more independently, potentially reflecting a tendency toward emotional suppression or less overt emotional expression [49]. This dissociation may stem from Confucian masculinity norms that stigmatize emotional disclosure while sanctioning somatic codification of distress [50]. The difference can be explained by brain structure differences [51]. Variations in hippocampal volume and corpus callosum structure have been reported between sexes [52]. These differences might affect cognitive processing and interhemispheric communication, further modulating how emotional information is integrated and expressed [53].

In women, enhanced reactivity within the limbic system, particularly in the amygdala, and stronger connectivity between the amygdala and prefrontal cortex have been associated with heightened sensitivity to emotional stimuli and increased vulnerability to mood disturbances [54]. By contrast, men often exhibit differences in the structure and function of the prefrontal cortex and anterior cingulate cortex, which may confer a greater capacity for emotional regulation and influence the more discrete manifestation of depressive symptoms [55]. Notably, the between-subject network for males revealed instances where stress or emotional distress was associated with positive emotional responses. This phenomenon may be attributable to sex-specific emotion regulation strategies, such as self-encouragement or attentional shifting [56], highlighting the potential benefit of incorporating positive emotion regulation techniques in interventions targeted at Chinese men experiencing emotional distress.

The temporal dynamics observed in this network analysis carry significant implications for public health strategies targeting the Chinese aging population. Given the centrality of sleep duration in driving depressive symptoms and cognitive decline, community health centers could prioritize scalable sleep screening protocols, such as incorporating the Pittsburgh Sleep Quality Index (PSQI) into annual geriatric check-ups to identify high-risk individuals before downstream mental health deterioration occurs. Building on the bidirectional sleep-depression pathway, we further advocate for training primary care physicians in brief behavioral interventions for insomnia, leveraging the Chinese tiered health care system to deliver stepped-care models where digital platforms (eg, WeChat-based sleep trackers) provide first-line support, with psychiatric referrals reserved for complex cases. Crucially, the mediating role of cognitive function suggests that integrating cognitive preservation components into existing sleep programs, particularly for those with mild cognitive impairment, may amplify intervention benefits.

This study has several strengths. First, we examined both intraindividual and interindividual longitudinal networks for sleep duration, cognitive function, and depression symptoms. This approach addresses a growing need for research at the symptom level, representing a shift from traditional cross-sectional studies to analyses of complex, dynamic systems within individuals and across subpopulations. Second, we extended previous research to examine sex-specific network patterns in a Chinese context. Third, our innovative integration of sleep duration, cognition, and depression into a unified network analysis has enriched our understanding of the dynamic relationships among these factors, which is critical for mental well - being and cognitive health in the aging population.

Limitations

However, our study has limitations that offer directions for future research. First, the current data do not capture immediate variable relationships; future studies might use ecological momentary assessment to more accurately capture dynamic changes in sleep, cognition, and depression within natural environments [57]. Second, another limitation is the use of a single-item sleep duration measure, which does not capture sleep architecture or the presence of sleep disorders. Future network studies should prioritize multimodal sleep assessments (eg, actigraphy combined with clinical screening tools) to enhance measurement accuracy. Third, the sources of sex-specific network heterogeneity remain unclear. It is uncertain whether these differences stem solely from inherent group characteristics or from a combination of group differences and other psychosocial factors. Prospective experimental designs over extended periods are needed to clarify the causality of these sex differences. In addition, statistical comparisons of

XSL•FO
RenderX

sex-stratified networks were not performed due to the absence of validated methods for testing longitudinal network invariance [58]. Finally, replicating and extending this study in a Western context would help determine whether the interaction patterns among sleep, cognition, and depression observed in Chinese middle-aged and older adults are consistent across cultural settings.

Conclusions

This study pioneers the investigation of predictive relationships among sleep duration, cognitive function, and depression in a large cohort of middle-aged and older Chinese individuals. Our findings indicate that emotional distress—most notably the "bothered" symptom—plays a central role in predicting depressive symptoms and impairing cognitive function, while poor sleep duration serves as a critical link between depression and cognitive decline. These results emphasize the importance of addressing emotional distress and improving sleep duration as potential strategies to mitigate depressive symptoms and slow cognitive decline in older adults.

Acknowledgments

The authors thank all collaborators for their support in this study.

Data Availability

Data are deposited in publicly available repositories (further details and data are available at [59]).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Node abbreviation index, complete network analysis results, with sex-specific results for female and male groups.
[DOCX File, 85 KB - aging_v8i1e76210_app1.docx]

Multimedia Appendix 2
Expected Influence of network nodes among temporal network.
[PNG File, 21 KB - aging_v8i1e76210_app2.png]

Multimedia Appendix 3
Centrality of network nodes among contemporaneous network, including Expected Influence (A), Closeness (B), Strength (C), and Betweenness (D).
[PNG File, 45 KB - aging_v8i1e76210_app3.png]

Multimedia Appendix 4
Centrality of network nodes among between-subject network, including Expected Influence (A), Closeness (B), Strength (C), and Betweenness (D).
[PNG File, 49 KB - aging_v8i1e76210_app4.png]

Multimedia Appendix 5
Temporal network models for female participants.
[PDF File, 276 KB - aging_v8i1e76210_app5.pdf]

Multimedia Appendix 6
Temporal network models for male participants.
[PDF File, 280 KB - aging_v8i1e76210_app6.pdf]

Multimedia Appendix 7
Contemporaneous network models for male participants.
[PDF File, 269 KB - aging_v8i1e76210_app7.pdf]

Multimedia Appendix 8
Contemporaneous network models for female participants.
[PDF File, 259 KB - aging_v8i1e76210_app8.pdf]

XSL•FO
RenderX

Multimedia Appendix 9
Between-subject network models for male participants.
[PDF File, 271 KB - aging_v8i1e76210_app9.pdf]

Multimedia Appendix 10
Between-subject network models for female participants.
[PDF File, 277 KB - aging_v8i1e76210_app10.pdf]

References

1. Luo Y, Su B, Zheng X. Trends and challenges for population and health during population aging - China, 2015-2050. China CDC Wkly 2021 Jul 9;3(28):593-598. [doi: 10.46234/ccdcw2021.158] [Medline: 34594944]
2. Reynolds CF 3rd, Jeste DV, Sachdev PS, Blazer DG. Mental health care for older adults: recent advances and new directions in clinical practice and research. World Psychiatry 2022 Oct;21(3):336-363. [doi: 10.1002/wps.20996] [Medline: 36073714]
3. Wu Y, Lei P, Ye R, Sunil TS, Zhou H. Prevalence and risk factors of depression in middle-aged and older adults in urban and rural areas in China: a cross-sectional study. Lancet 2019 Oct;394:S53. [doi: 10.1016/S0140-6736(19)32389-X]
4. Fan X, Guo X, Ren Z, et al. The prevalence of depressive symptoms and associated factors in middle-aged and elderly Chinese people. J Affect Disord 2021 Oct 1;293:222-228. [doi: 10.1016/j.jad.2021.06.044] [Medline: 34217959]
5. Zhang NJ, Qian ZD, Zeng YB, Gu JN, Jin Y, Li W. Incidence and risk factors associated with progression to mild cognitive impairment among middle aged and older adults. Eur Rev Med Pharmacol Sci 2022 Dec;26(23):8852-8859. [doi: 10.26355/eurrev_202212_30557] [Medline: 36524504]
6. Wu Y, Chen Z, Cheng Z, et al. Effects of chronic diseases on health related quality of life is mediated by sleep difficulty in middle aged and older adults. Sci Rep 2025 Jan 23;15(1):2987. [doi: 10.1038/s41598-025-86420-1] [Medline: 39849013]
7. Bubu OM, Brannick M, Mortimer J, et al. Sleep, cognitive impairment, and alzheimer's disease: a systematic review and meta-analysis. Sleep 2017 Jan 1;40(1). [doi: 10.1093/sleep/zsw032] [Medline: 28364458]
8. Benito-León J, Louis ED, Bermejo-Pareja F. Cognitive decline in short and long sleepers: a prospective population-based study (NEDICES). J Psychiatr Res 2013 Dec;47(12):1998-2003. [doi: 10.1016/j.jpsychires.2013.09.007] [Medline: 24094933]
9. Lowe CJ, Safati A, Hall PA. The neurocognitive consequences of sleep restriction: a meta-analytic review. Neurosci Biobehav Rev 2017 Sep;80:586-604. [doi: 10.1016/j.neubiorev.2017.07.010] [Medline: 28757454]
10. Butters MA, Young JB, Lopez O, et al. Pathways linking late-life depression to persistent cognitive impairment and dementia. Dialogues Clin Neurosci 2008;10(3):345-357. [doi: 10.31887/DCNS.2008.10.3/mabutters] [Medline: 18979948]
11. Zhou L, Ma X, Wang W. Relationship between cognitive performance and depressive symptoms in Chinese older adults: the China Health and Retirement Longitudinal Study (CHARLS). J Affect Disord 2021 Feb 15;281:454-458. [doi: 10.1016/j.jad.2020.12.059] [Medline: 33360747]
12. Guan Q, Hu X, Ma N, et al. Sleep quality, depression, and cognitive function in non-demented older adults. J Alzheimers Dis 2020;76(4):1637-1650. [doi: 10.3233/JAD-190990] [Medline: 32675408]
13. Riemann D, Krone LB, Wulff K, Nissen C. Sleep, insomnia, and depression. Neuropsychopharmacology 2020 Jan;45(1):74-89. [doi: 10.1038/s41386-019-0411-y] [Medline: 31071719]
14. Sun Y, Shi L, Bao Y, Sun Y, Shi J, Lu L. The bidirectional relationship between sleep duration and depression in community-dwelling middle-aged and elderly individuals: evidence from a longitudinal study. Sleep Med 2018 Dec;52:221-229. [doi: 10.1016/j.sleep.2018.03.011] [Medline: 29861378]
15. Huang D, Liu Z, Ma S, Liu M, Liu C, Liu A. The relationship between night sleep duration, sleep quality and depressive symptoms in Chinese elderly couples. Geriatr Nurs 2024;59:623-629. [doi: 10.1016/j.gerinurse.2024.07.024] [Medline: 39182443]
16. Hajali V, Andersen ML, Negah SS, Sheibani V. Sex differences in sleep and sleep loss-induced cognitive deficits: the influence of gonadal hormones. Horm Behav 2019 Feb;108:50-61. [doi: 10.1016/j.yhbeh.2018.12.013] [Medline: 30597139]
17. Zhou L, Kong J, Li X, Ren Q. Sex differences in the effects of sleep disorders on cognitive dysfunction. Neurosci Biobehav Rev 2023 Mar;146:105067. [doi: 10.1016/j.neubiorev.2023.105067] [Medline: 36716906]
18. White KH, Rumble ME, Benca RM. Sex differences in the relationship between depressive symptoms and actigraphic assessments of sleep and rest-activity rhythms in a population-based sample. Psychosom Med 2017 May;79(4):479-484. [doi: 10.1097/PSY.0000000000000434] [Medline: 27922568]
19. Jordan DG, Winer ES, Salem T. The current status of temporal network analysis for clinical science: considerations as the paradigm shifts? J Clin Psychol 2020 Sep;76(9):1591-1612. [doi: 10.1002/jclp.22957] [Medline: 32386334]
20. Silva VF, Silva ME, Ribeiro P, Silva F. Time series analysis via network science: concepts and algorithms. WIREs Data Min & Knowl 2021 May;11(3). [doi: 10.1002/widm.1404]
21. Andresen EM, Malmgren JA, Carter WB, Patrick DL. Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). Am J Prev Med 1994;10(2):77-84. [Medline: 8037935]

XSL•FO
RenderX

22. Chen H, Mui AC. Factorial validity of the center for epidemiologic studies depression scale short form in older population in China. Int Psychogeriatr 2014 Jan;26(1):49-57. [doi: 10.1017/S1041610213001701] [Medline: 24125553]
23. Sakal C, Li J, Xiang YT, Li X. Development and validation of the Chinese Geriatric Depression Risk calculator (CGD-risk): a screening tool to identify elderly Chinese with depression. J Affect Disord 2022 Dec 15;319:428-436. [doi: 10.1016/j.jad.2022.09.034] [Medline: 36184985]
24. Chan RCK, Hoosain R, Lee TMC. Reliability and validity of the cantonese version of the test of everyday attention among normal Hong Kong Chinese: a preliminary report. Clin Rehabil 2002 Dec;16(8):900-909. [doi: 10.1191/0269215502cr574oa] [Medline: 12501953]
25. Epskamp S, Waldorp LJ, Mõttus R, Borsboom D. The Gaussian graphical model in cross-sectional and time-series data. Multivariate Behav Res 2018;53(4):453-480. [doi: 10.1080/00273171.2018.1454823] [Medline: 29658809]
26. Borsboom D, Deserno MK, Rhemtulla M, et al. Network analysis of multivariate data in psychological science. Nat Rev Methods Primers 2021;1(1):58. [doi: 10.1038/s43586-021-00055-w]
27. Fried EI, Cramer AOJ. Moving forward: challenges and directions for psychopathological network theory and methodology. Perspect Psychol Sci 2017 Nov;12(6):999-1020. [doi: 10.1177/1745691617705892] [Medline: 28873325]
28. Epskamp S. Psychometric network models from time-series and panel data. Psychometrika 2020 Mar;85(1):206-231. [doi: 10.1007/s11336-020-09697-3] [Medline: 32162233]
29. Marsh HW, Hau KT, Wen Z. In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. Struct Equ Modeling 2004 Jul;11(3):320-341. [doi: 10.1207/s15328007sem1103_2]
30. Epskamp S, Costantini G, Cramer AOJ, Waldorp LJ. Qgraph: graph plotting methods, psychometric data visualization and graphical model estimation. 2016 URL: https://cran.r-project.org/web/packages/qgraph/qgraph.pdf [accessed 2025-08-11]
31. Bringmann LF, Pe ML, Vissers N, et al. Assessing temporal emotion dynamics using networks. Assessment 2016 Aug;23(4):425-435. [doi: 10.1177/1073191116645909] [Medline: 27141038]
32. Geraghty AWA, Santer M, Beavis C, et al. "I mean what is depression?" A qualitative exploration of UK general practitioners' perceptions of distinctions between emotional distress and depressive disorder. BMJ Open 2019 Dec 15;9(12):e032644. [doi: 10.1136/bmjopen-2019-032644] [Medline: 31843841]
33. Berger AT, Wahlstrom KL, Widome R. Relationships between sleep duration and adolescent depression: a conceptual replication. Sleep Health 2019 Apr;5(2):175-179. [doi: 10.1016/j.sleh.2018.12.003] [Medline: 30928118]
34. Gulia KK, Kumar VM. Sleep disorders in the elderly: a growing challenge. Psychogeriatrics 2018 May;18(3):155-165. [doi: 10.1111/psyg.12319] [Medline: 29878472]
35. Fjell AM, Sørensen Ø, Amlien IK, et al. Self-reported sleep relates to hippocampal atrophy across the adult lifespan: results from the lifebrain consortium. Sleep 2020 May 12;43(5):zsz280. [doi: 10.1093/sleep/zsz280] [Medline: 31738420]
36. Xanidis N, Brignell CM. The association between the use of social network sites, sleep quality and cognitive function during the day. Comput Human Behav 2016 Feb;55:121-126. [doi: 10.1016/j.chb.2015.09.004]
37. Spira AP, Gonzalez CE, Venkatraman VK, et al. Sleep duration and subsequent cortical thinning in cognitively normal older adults. Sleep 2016 May 1;39(5):1121-1128. [doi: 10.5665/sleep.5768] [Medline: 26951390]
38. Majd M, Saunders EFH, Engeland CG. Inflammation and the dimensions of depression: a review. Front Neuroendocrinol 2020 Jan;56:100800. [doi: 10.1016/j.yfrne.2019.100800] [Medline: 31654681]
39. Price RB, Duman R. Neuroplasticity in cognitive and psychological mechanisms of depression: an integrative model. Mol Psychiatry 2020 Mar;25(3):530-543. [doi: 10.1038/s41380-019-0615-x] [Medline: 31801966]
40. Han S, Fang K, Zheng R, et al. Gray matter atrophy is constrained by normal structural brain network architecture in depression. Psychol Med 2024 May;54(7):1318-1328. [doi: 10.1017/S0033291723003161] [Medline: 37947212]
41. Zeki Al Hazzouri A, Caunca MR, Nobrega JC, et al. Greater depressive symptoms, cognition, and markers of brain aging: Northern Manhattan study. Neurology (ECronicon) 2018 Jun 5;90(23):e2077-e2085. [doi: 10.1212/WNL.0000000000005639] [Medline: 29743209]
42. Barnett KJ, Cooper NJ. The effects of a poor night sleep on mood, cognitive, autonomic and electrophysiological measures. J Integr Neurosci 2008 Sep;7(3):405-420. [doi: 10.1142/s0219635208001903] [Medline: 18988299]
43. Alitalo O, González-Hernández G, Rosenholm M, et al. Linking hypothermia and altered metabolism with TrkB activation. ACS Chem Neurosci 2023 Sep 6;14(17):3212-3225. [doi: 10.1021/acschemneuro.3c00350] [Medline: 37551888]
44. Yang FN, Xie W, Wang Z. Effects of sleep duration on neurocognitive development in early adolescents in the USA: a propensity score matched, longitudinal, observational study. Lancet Child Adolesc Health 2022 Oct;6(10):705-712. [doi: 10.1016/S2352-4642(22)00188-2] [Medline: 35914537]
45. Sutter C, Zöllig J, Allemand M, Martin M. Sleep quality and cognitive function in healthy old age: the moderating role of subclinical depression. Neuropsychology 2012 Nov;26(6):768-775. [doi: 10.1037/a0030033] [Medline: 22984799]
46. Didikoglu A, Guler ES, Turk HK, et al. Depression in older adults and its associations with sleep and synaptic density. J Affect Disord 2024 Dec 1;366:379-385. [doi: 10.1016/j.jad.2024.08.186] [Medline: 39216641]
47. Choi HS, Lee JE. Factors affecting depression in middle-aged and elderly men living alone: a cross-sectional path analysis model. Am J Mens Health 2022;16(1):15579883221078134. [doi: 10.1177/15579883221078134] [Medline: 35184578]

XSL•FO
RenderX

48. Simpson PA, Stroh LK. Gender differences: emotional expression and feelings of personal inauthenticity. J Appl Psychol 2004 Aug;89(4):715-721. [doi: 10.1037/0021-9010.89.4.715] [Medline: 15327356]
49. Rogier G, Garofalo C, Velotti P. Is emotional suppression always bad? A matter of flexibility and gender differences. Curr Psychol 2019 Apr;38(2):411-420. [doi: 10.1007/s12144-017-9623-7]
50. Shi P, Yang A, Zhao Q, Chen Z, Ren X, Dai Q. A hypothesis of gender differences in self-reporting symptom of depression: implications to solve under-diagnosis and under-treatment of depression in males. Front Psychiatry 2021;12:589687. [doi: 10.3389/fpsyt.2021.589687] [Medline: 34759845]
51. Lindquist KA, Jackson JC, Leshin J, Satpute AB, Gendron M. The cultural evolution of emotion. Nat Rev Psychol 2022;1(11):669-681. [doi: 10.1038/s44159-022-00105-4]
52. Mohammadi S, Seyedmirzaei H, Salehi MA, et al. Brain-based sex differences in depression: a systematic review of neuroimaging studies. Brain Imaging Behav 2023 Oct;17(5):541-569. [doi: 10.1007/s11682-023-00772-8] [Medline: 37058182]
53. Wang Y, Chen Z, Ma G, et al. A frontal transcallosal inhibition loop mediates interhemispheric balance in visuospatial processing. Nat Commun 2023 Aug 25;14(1):5213. [doi: 10.1038/s41467-023-40985-5] [Medline: 37626171]
54. Sun S, Yu H, Yu R, Wang S. Functional connectivity between the amygdala and prefrontal cortex underlies processing of emotion ambiguity. Transl Psychiatry 2023 Oct 28;13(1):334. [doi: 10.1038/s41398-023-02625-w] [Medline: 37898626]
55. Levy R. The prefrontal cortex: from monkey to man. Brain (Bacau) 2024 Mar 1;147(3):794-815. [doi: 10.1093/brain/awad389] [Medline: 37972282]
56. Kaur A, Kailash SZ, Sureshkumar KCS, Rumaisa N. Gender differences in emotional regulation capacity among the general population. Int Arch Integr Med 2022;9(1):22-28 [FREE Full text]
57. Stone AA, Schneider S, Smyth JM. Evaluation of pressing issues in ecological momentary assessment. Annu Rev Clin Psychol 2023 May 9;19:107-131. [doi: 10.1146/annurev-clinpsy-080921-083128] [Medline: 36475718]
58. van Borkulo CD, van Bork R, Boschloo L, et al. Comparing network structures on three aspects: a permutation test. Psychol Methods 2023 Dec;28(6):1273-1285. [doi: 10.1037/met0000476] [Medline: 35404628]
59. CHARLS. China Health and Retirement Longitudinal Study [Website in Chinese]. URL: https://charls.pku.edu.cn [accessed 2025-08-11]

Abbreviations

CESD-10: 10-item Center for Epidemiologic Studies Depression Scale
CFI: confirmatory fit index
CHARLS: China Health and Retirement Longitudinal Study
FIML: full information maximum likelihood
GGMs: Gaussian graphical models
GVAR: graphical vector autoregressive
KPSS: Kwiatkowski-Phillips-Schmidt-Shin
MMSE: Mini-Mental State Examination
PSQI: Pittsburgh Sleep Quality Index
RMSEA: root mean squared error of approximation
TICS: Telephone Interview for Cognitive Status
TLI: Tucker-Lewis index

Edited by H LaMonica; submitted 18.04.25; peer-reviewed by C Sakal, L Yang, V Rashedi; revised version received 15.06.25; accepted 30.06.25; published 16.09.25.

Please cite as:
Chen F, Li J, Guo J, Xiong Y, Ye Z
Dynamic Interactions Among Sleep Duration, Cognitive Function, and Depressive Symptoms in Middle-Aged and Older Chinese Adults: Temporal Network Analysis From CHARLS
JMIR Aging 2025;8:e76210
URL: https://aging.jmir.org/2025/1/e76210
doi:10.2196/76210

©Furong Chen, Jiaying Li, Junchen Guo, Ying Xiong, Zengjie Ye. Originally published in JMIR Aging (https://aging.jmir.org), 16.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium,

XSL•FO
RenderX

provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Impact of Social Stress and Healthy Lifestyle on the Mortality of Chinese Older Adults: Prospective Cohort Study

Jin Yang[1*], MD, PhD; Jilong Huang[1*], MPH; Qingmei Huang[1], MD, PhD; Jian Gao[1], MD, PhD; Dan Liu[1], MD, PhD; Zhihao Li[1], MD, PhD; Yuebin Lv[2], MD, PhD; Xiaoming Shi[2], MD, PhD; Chen Mao[1,3], MD, PhD

[1]Department of Epidemiology, School of Public Health, Southern Medical University, No. 1023-1063, Shatai South Road, Baiyun District, Guangzhou, Guangdong, China

[2]China CDC Key Laboratory of Environment and Population Health, National Institute of Environmental Health,, Chinese Center for Disease Control and Prevention, Beijing, China

[3]National Institute of Health Data Science of China, Southern Medical University, Guangzhou, Guangdong, China

[*]these authors contributed equally

Corresponding Author:
Chen Mao, MD, PhD
Department of Epidemiology, School of Public Health, Southern Medical University, No. 1023-1063, Shatai South Road, Baiyun District, Guangzhou, Guangdong, China

Abstract

Background: With social progress, social stress (SS) has become a key factor affecting health. Unhealthy lifestyles may exacerbate these effects. However, the relationship between SS, lifestyle, and older adults' mortality rate still needs to be studied.

Objective: This study aimed to explore the relationship between SS and all-cause mortality in Chinese older adults, as well as the influence of healthy lifestyle factors.

Methods: Three groups of SS were defined through latent class analysis: low, medium, and high. We created a healthy lifestyle index based on smoking, alcohol consumption, physical activity, and diet. Multivariable Cox proportional hazards models, interaction analyses, and mediation analyses were conducted.

Results: The Chinese Longitudinal Healthy Longevity Survey (CLHLS) datasets included participants from 806 cities and counties across 23 provinces in China from 1998 to 2018. In this study, participants were recruited from 4 waves of the CLHLS (2005, 2008, 2011, and 2014). Finally, 19,236 participants were included in this study, of which 6891 (35.8%) had low SS, 11,662 (60.6%) had medium SS, and 683 (3.6%) had high SS. In the fully adjusted model, the hazard ratio (HR) for medium SS was 1.16 (95% CI 1.11 - 1.20; *P*<.001), and for high SS, it was 1.28 (95% CI 1.18 - 1.40; *P*<.001) compared to the low SS group. For individuals aged ≥80 years, the medium SS group had a 28% (HR 1.28, 95% CI 1.22 - 1.34; *P*<.001) increased mortality risk, and the high SS group had a 38% (HR 1.38, 95%CI 1.26 - 1.52; *P*<.001) increased risk compared to the low SS group. Approximately 7% of the association between SS and mortality was mediated through the healthy lifestyle. Under different SS, the lower the healthy lifestyle score, the higher the risk of mortality.

Conclusions: SS was an independent predictor of all-cause mortality in Chinese older adults. The healthy lifestyle mediated this effect to some extent. Unhealthy lifestyle behaviors were associated with a higher risk of mortality at all SS levels.

(JMIR Aging 2025;8:e75942) doi:10.2196/75942

KEYWORDS

social stress; healthy lifestyle; mortality; older adults; cohort study

Introduction

As social progress advances, increasing attention has been given to the impact of chronic social stress (SS) on human health. SS is a complex and multifaceted construct, typically encompassing 3 key areas, such as interpersonal relationships, economic conditions, and the living environment. Interpersonal stress arises from demands and expectations within family, friendships, and work relationships [1]. Economic stress is primarily linked to concerns over income or job stability [2] while environmental stress involves factors, such as housing stress and access to social services [3]. All of these factors can contribute to an individual's overall level of SS. Clinical and biological studies have demonstrated that long-term exposure to stress can disrupt the normal functioning of biological systems, leading to negative health outcomes, such as cardiovascular disease [4], gastrointestinal disorders [5], and immune dysfunction [6]. In addition, unhealthy lifestyle behaviors, such as poor diet, physical inactivity, and smoking, are well-documented risk

XSL•FO
RenderX

factors for a range of chronic diseases and premature mortality [7], and may further compound the adverse effects of SS.

China is undergoing a significant demographic shift, with a rapidly aging population. As such, the health and well-being of older adults have become key areas of focus. While stress affects individuals at all stages of life, its impact on older adults warrants particular attention. Studies have shown that stress may be linked to poorer health in older adults [8-10]. However, more research is needed to understand how SS affects the health of older adults and the impact of lifestyle.

This study aimed to investigate the impact of different levels of SS on mortality among Chinese adults older than 65 years, using data from the Chinese Longitudinal Healthy Longevity Survey (CLHLS). In addition, the study investigated the possible impact of a healthy lifestyle on the relationship between SS and all-cause mortality.

Methods

Study Population

The data for this study were drawn from the Chinese Longitudinal Healthy Longevity Survey (CLHLS), a prospective cohort study described in detail in previous publications [11,12]. In brief, the CLHLS was based on a dynamic, multistage stratified sampling design, which includes participants from 806 cities and counties across 23 provinces in China from 1998 to 2018. The survey covers 85% of the Chinese population, with a focus on individuals aged 65 years and older [13]. Data collected include information on demographic characteristics, socioeconomic status, lifestyle factors, and health-related variables. For this analysis, 27,585 participants were recruited from 4 waves of the CLHLS (2005, 2008, 2011, and 2014). During the follow-up surveys, the survival status of the participants was ascertained as surviving, dead, or lost to follow-up. Participants were excluded if they met any of the following criteria: (1) lost to follow-up or dead before (n=5285), (2) age younger than 65 years at baseline (n=201), and (3) missing baseline information (n=2863). After applying these exclusions, the final sample consisted of 19,236 participants.

Assessment of SS

Existing research primarily examines the impact of SS on health from 3 aspects, namely, interpersonal relationships, economic conditions, and housing conditions [14-16]. Accordingly, our study assessed SS among the older adult population in the CLHLS based on these 3 dimensions. Based on the investigation of relationships with marriage, the health of children, and social support, these questions further transformed into 3 questions, such as “Do you live with your spouse?,” “Are your children still alive?,” and “Do you have someone with whom you can talk, confide in, or seek help when needed?” Economic status was measured by the question, “Is all of the financial support sufficient to pay for daily expenses?” Housing stress was evaluated using the question, “How many people are living with you?” Given the living arrangements typical for older adults in China, those living alone or with more than 5 people were categorized as experiencing housing stress (Table S1 in Multimedia Appendix 1).

Assessment of Healthy Lifestyle Score

A healthy lifestyle was defined by 4 factors, such as smoking, alcohol consumption, physical activity, and diet. Participants who had never smoked or had quit smoking for at least 30 years were assigned 1 point [17]. Men who consume less than 62 g of alcohol per day or women who consume less than 41 g per day were assigned 1 point [18]. Physical activities included regular exercise (both aerobic and anaerobic), housework, tasks, outdoor activities, gardening, pet care, reading, playing cards or mahjong, watching television or listening to the radio, and attending social events. The frequency of “almost every day” was scored as 2, “occasionally” was scored as 1, and “rarely or never” was scored as 0. A total physical activity score was calculated by summing the points across these categories and then standardized to a range of 0-1. A score below 0.6 was assigned 0 points while a score of 0.6 or above was assigned 1 point (Table S2 in Multimedia Appendix 1). The dietary score was determined using a standardized food frequency questionnaire that includes fresh vegetables, fresh fruits, legumes, meat, eggs, fish and seafood, salty vegetables, tea, and garlic. Similar to physical activity, the score for each food group was based on frequency of consumption, and the total dietary score is calculated accordingly (Table S3 in Multimedia Appendix 1). The final lifestyle score ranged from 0 to 4, with higher scores indicating a healthier lifestyle [19].

Outcome

Death was confirmed either by official death certificates or through information provided by the participant’s next of kin, local physician, or resident physician. The follow-up period was defined as the time from study enrollment to the date of death or the last follow-up, whichever came first. Loss to follow-up was defined as the inability to contact the participant after at least 3 documented attempts.

Covariates

Several variables were considered as potential confounding factors. The first set included demographic variables, namely, age (y), sex (male or female), residence (urban or rural), education, and occupation prior to retirement. Considering that the participants included in this study were 65 years old and above, 0 years of education was defined as low, less than 5 years of education was defined as medium, and 5 years of education or more was defined as high. Occupations before retirement were divided into farming, employed, and unemployed. The second set included the history of chronic diseases, such as hypertension (yes or no), diabetes (yes or no), heart disease (yes or no), stroke and cerebrovascular disease (yes or no), and cancer (yes or no). The history of chronic diseases was measured with the following question: “Have you been diagnosed by a doctor with the conditions listed below?”

Statistical Analysis Methods of the Results

Continuous variables were summarized as means and SDs while categorical variables were presented as counts (N) and percentages (%). Differences between groups were assessed using analysis of variance for continuous variables and the chi-square test for categorical variables. Cox proportional hazard regression models were used to estimate the hazard ratios (HRs)

XSL•FO
RenderX

and 95% CI for mortality associated with SS and healthy lifestyle scores. The proportional hazards assumption was tested using Schoenfeld residuals. Model 1 was unadjusted; Model 2 adjusted for age and sex; Model 3 further included education level, occupation, and residence; and Model 4 additionally adjusted for healthy lifestyle factors. Mediation analysis was used to further explore the role of healthy lifestyle in the relationship between SS and all-cause mortality. In addition, the product term of SS and healthy lifestyle score was added to the model to examine their multiplicative interaction. The relative excess risk due to interaction and its corresponding 95% CI were used to assess the interaction on the additive scale. This was calculated using the coefficients and SEs of the product term, SS, and healthy lifestyle score, along with the covariance matrix.

Latent class analysis (LCA) was used to classify participants into distinct SS groups. Convergence was defined as a maximum absolute deviation of 0.000001 between parameter estimates across successive iterations. Iterations terminated when this deviation threshold was reached. Model selection was guided by criteria, including Akaike Information Criterion, Bayesian Information Criterion, likelihood ratio statistic (G^2), and entropy value. Entropy, ranging from 0 to 1, assesses the accuracy of classification in LCA, with higher values indicating better model classification accuracy. In this study, after convergence failed for the 6-class model, 5 latent class groups were selected for comparison, with optimal model performance observed at 3 latent classes based on the highest entropy values (Tables S4-S6 in Multimedia Appendix 1).

Furthermore, when we analyze the impact of a healthy lifestyle on mortality under different SS conditions, the small number of healthy lifestyle groups in the high SS group may affect the results. Therefore, we combined the medium SS group and the high SS group into the medium-high SS group for analysis. We conducted a subgroup analysis stratified by age, sex, education, residence, and occupation to present the differences. We performed several sensitivity analyses to test the robustness of our results: (1) Excluded deaths that occurred within the first 2 years of follow-up. (2) Excluded the mediating effect of each lifestyle factor on the relationship between SS and all-cause mortality. (3) The results were evaluated again by weighting the lifestyle scores. (4) Due to the low proportion of participants with lifestyle scores of 2 and 3, they were combined into 1 group for analysis.

All analyses were based on R version 4.3.2 (R Foundation for Statistical Computing). All statistical tests were 2-sided, with $P<.05$ considered to be indicative of statistical significance.

Ethical Considerations

This study was approved by the Ethics Committee of Peking University (IRB00001052 −13074), and informed consent was obtained from all participants. All data were anonymized and deidentified to protect participant privacy. Participants provided written consent for the publication of their anonymized data and quotes in this study.

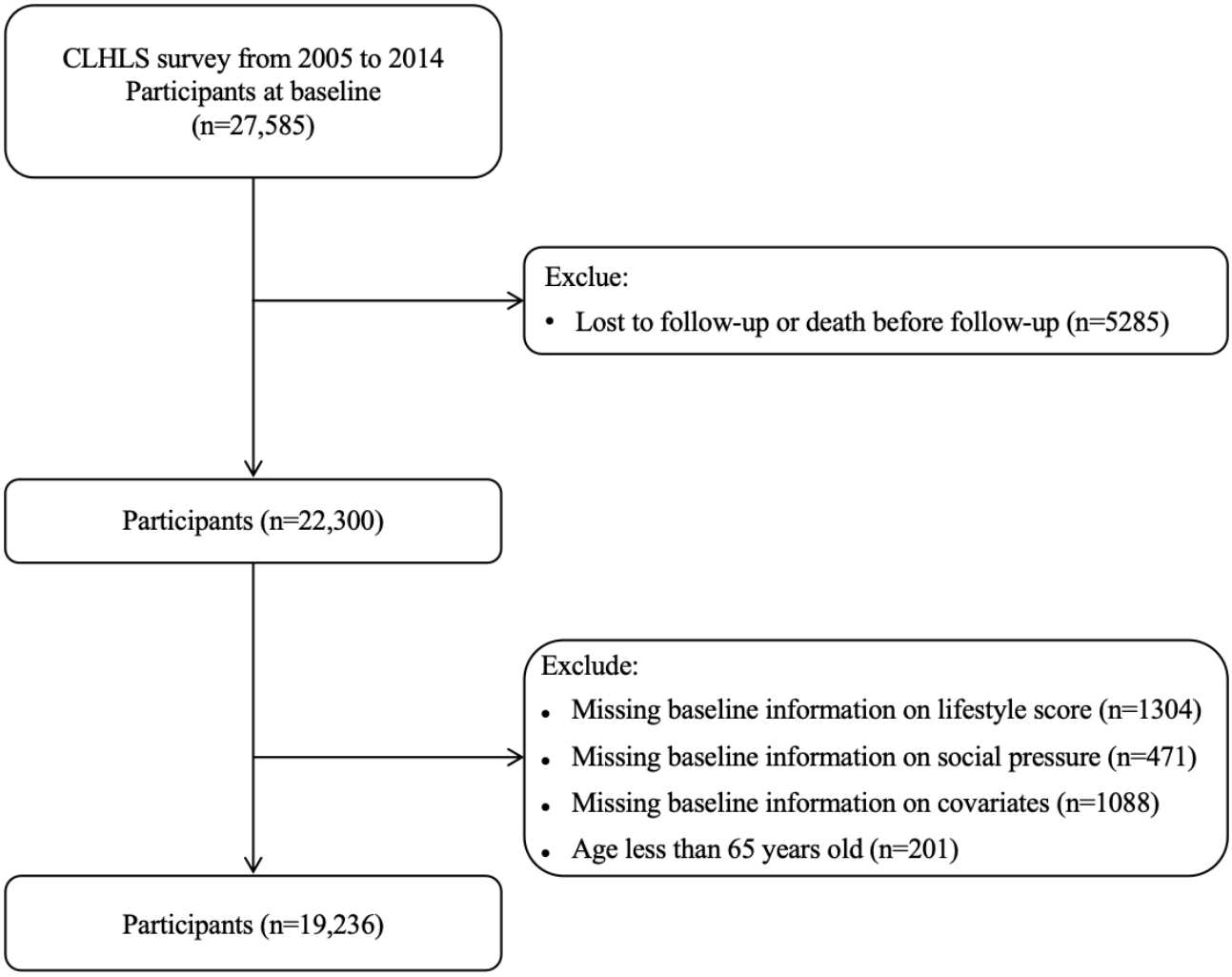
Results

Baseline Characteristics

A total of 19,236 Chinese adults aged 65 years and older were included in this prospective cohort study (Figure 1). Among them, 6891 (35.8%) people had low SS, 11,662 (60.6%) people had medium SS, and 683 (3.6%) people had high SS. The overall mean age of the participants was 88 (SD 12) years. Overall, females, rural residents, medium education, and agriculture make up a larger proportion of the total population. The high SS group exhibited a higher proportion of individuals experiencing several adverse conditions, including the loss of children, lack of social support, living without a spouse, insufficient financial resources to cover daily expenses, and housing stress. Overall, 6827 (35.5%) participants had a healthy lifestyle score of 0, 7017 (36.5%) participants scored 1, 2633 (13.7%) participants scored 2, and 2759 (14.3%) participants scored ≥3. Chronic disease conditions and other details can be seen in Table 1.

XSL•FO
RenderX

Figure 1. Flowchart of participant selection. CLHLS: Chinese Longitudinal Healthy Longevity Survey.



XSL•FO
RenderX

Table . Baseline characteristics of the study participants.

Variables	Overall (N=19,236)	Social stress		
		Low (n=6891)	Medium (n=11,662)	High (n=683)
Age (year), mean (SD)	88 (12)	81 (11)	91 (10)	91 (10)
Sex (female), n (%)	11195 (58.2)	2792 (40.5)	7964 (68.3)	439 (64.3)
Residence (urban), n (%)	2338 (12.2)	908 (13.2)	1368 (11.7)	62 (9.1)
Education[a], n (%)				
High	3436 (17.9)	1950 (28.3)	1429 (12.3)	57 (8.3)
Medium	12487 (64.9)	3432 (49.8)	8513 (73.0)	542 (79.4)
Low	3313 (17.2)	1509 (21.9)	1720 (14.7)	84 (12.3)
Occupation, n (%)				
Agriculture	13042 (67.8)	4382 (63.6)	8140 (69.8)	520 (76.1)
Unemployment	2560 (13.3)	671 (9.7)	1785 (15.3)	104 (15.2)
Employment	3634 (18.9)	1838 (26.7)	1737 (14.9)	59 (8.6)
Are your children alive? (Yes), n (%)	18519 (96.3)	6810 (98.8)	11232 (96.3)	477 (69.8)
Do you have someone with whom you can talk, confide in, or seek help when needed? (Yes), n (%)	17676 (91.9)	6750 (98.0)	10778 (92.4)	148 (21.7)
Do you live with your spouse? (Yes), n (%)	5637 (29.3)	5616 (81.5)	0 (0.0)	19 (3.1)
Is all of the financial support sufficient to pay for daily expenses? (Yes), n (%)	14914 (77.5)	4381 (63.6)	10532 (90.3)	1 (0.1)
Are you living alone or with ≥5 people? (No), n (%)	11141 (57.9)	5545 (80.5)	5333 (45.7)	263 (38.5)
Healthy lifestyle score[b], n (%)				
≥3	2759 (14.3)	1190 (17.3)	1540 (13.2)	29 (4.2)
2	2633 (13.7)	1040 (15.1)	1460 (12.5)	133 (19.5)
1	7017 (36.5)	2234 (32.4)	4409 (37.8)	374 (54.8)
0	6827 (35.5)	2427 (35.2)	4253 (36.5)	147 (21.5)
Hypertension (yes), n (%)	4981 (25.9)	2032 (29.5)	2829 (24.3)	120 (17.6)
Diabetes (yes), n (%)	1964 (10.2)	824 (12.0)	1107 (9.5)	33 (4.8)
Heart disease (yes), n (%)	2935 (15.3)	1214 (17.6)	1657 (14.2)	64 (9.4)
Stroke and cerebrovascular disease (yes), n (%)	2158 (11.2)	898 (13.0)	1206 (10.3)	54 (7.9)
Cancer (yes), n (%)	1042 (5.4)	404 (5.9)	618 (5.3)	20 (2.9)
Mortality, n (%)	13999 (72.8)	3986 (57.8)	9425 (80.8)	588 (86.1)

[a]Education: low (0 y), medium (<5 y), high (≥5 y).

[b]Healthy lifestyle was defined by 4 factors: smoking, alcohol consumption, physical activity, and diet. Healthy lifestyle score ranges from 0 to 4, with higher scores indicating a healthier lifestyle.

Association of SS and All-Cause Mortality

In the unadjusted analysis (Model 1), compared with the low SS group, the results showed the HRs of the medium and high SS groups were 2.05 (95% CI 1.98 - 2.13) and 2.47 (95% CI 2.26 - 2.69). After adjusting for age and sex (Model 2), the HRs of the medium and high SS groups were 1.15 (95% CI 1.11 - 1.20) and 1.37 (95% CI 1.26 - 1.50). In Model 3, which further adjusted for local education, occupation, and healthy lifestyle score based on Model 2, the HRs of the medium and high SS groups were 1.15 (95% CI 1.11 - 1.20) and 1.28 (95% CI 1.17 - 1.39). In the fully adjusted Model 4, the mortality risk for medium SS was 1.16 (95% CI 1.11 - 1.20), and for high

XSL•FO
RenderX

SS, it was 1.28 (95% CI 1.18 - 1.40) compared to the low SS groups (Table 2).

Table . Association of social stress and all-cause mortality.

Social stress	Model 1[a] HR[b] (95% CI)	Model 2[c] HR (95% CI)	Model 3[d] HR (95% CI)	Model 4[e] HR (95% CI)	*P* value
Low[f]	1 (Reference)	1 (Reference)	1 (Reference)	1 (Reference)	—[g]
Medium	2.05 (1.98-2.13)	1.15 (1.11-1.20)	1.15 (1.11-1.20)	1.16 (1.11-1.20)	<.001
High	2.47 (2.26-2.69)	1.37 (1.26-1.50)	1.28 (1.17-1.39)	1.28 (1.18-1.40)	<.001

[a]Model 1: unadjusted.

[b]Hazard ratio

[c]Model 2: adjusted for age and sex.

[d]Model 3: adjusted for residence, education, occupation, and healthy lifestyle score based on Model 2.

[e]Model 4: adjusted for hypertension, diabetes, heart disease, cancer, stroke, and cerebrovascular disease based on Model 3.

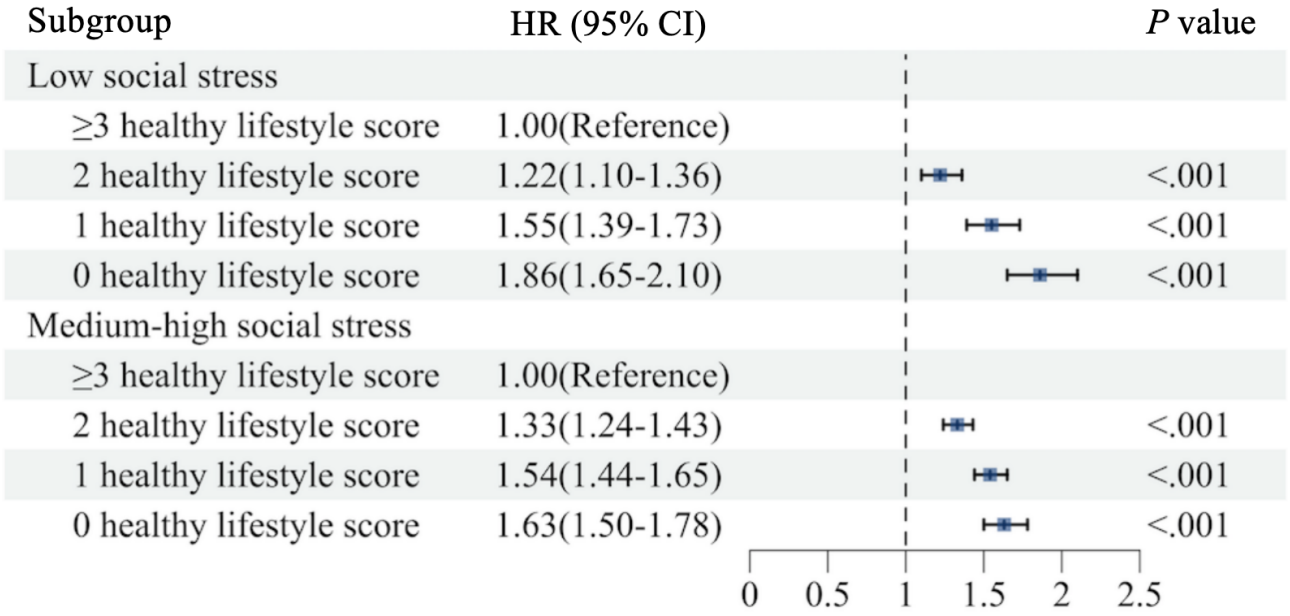
[f]The group is set as a reference group.

[g]Not applicable.

Impact of a Healthy Lifestyle on All-Cause Mortality Across Different SS Groups

The lower the healthy lifestyle score, the higher the risk of death compared with a healthy lifestyle score ≥3 (Table S7 in Multimedia Appendix 1). For the low SS group, individuals with a higher healthy lifestyle score had a lower mortality risk. Specifically, compared to those with a healthy lifestyle score of 3 or more, individuals with a score of 2 had a 22% increased risk of mortality (HR 1.22, 95% CI: 1.10 - 1.36), those with the score of 1 had a 55% increased risk (HR 1.55, 95% CI 1.39 - 1.73), and those with the score of 0 had an 86% increased risk (HR 1.86, 95% CI 1.65 - 2.10). Further combining the medium and high SS groups, the results found that individuals with a score of 2 had a 33% higher mortality risk (HR 1.33, 95% CI 1.24 - 1.43) compared to those with a score of 3 or more. The score of 1 was associated with a 54% higher risk (HR 1.54, 95% CI 1.44 - 1.65), and the score of 0 was associated with a 63% higher risk (HR 1.63, 95% CI 1.50 - 1.78) (Figure 2).

Figure 2. Impact of a healthy lifestyle on all-cause mortality across different social stress groups. All covariates were adjusted. HR: hazard ratio.



Mediation and Interaction Analysis of a Healthy Lifestyle on Associations of SS With All-Cause Mortality

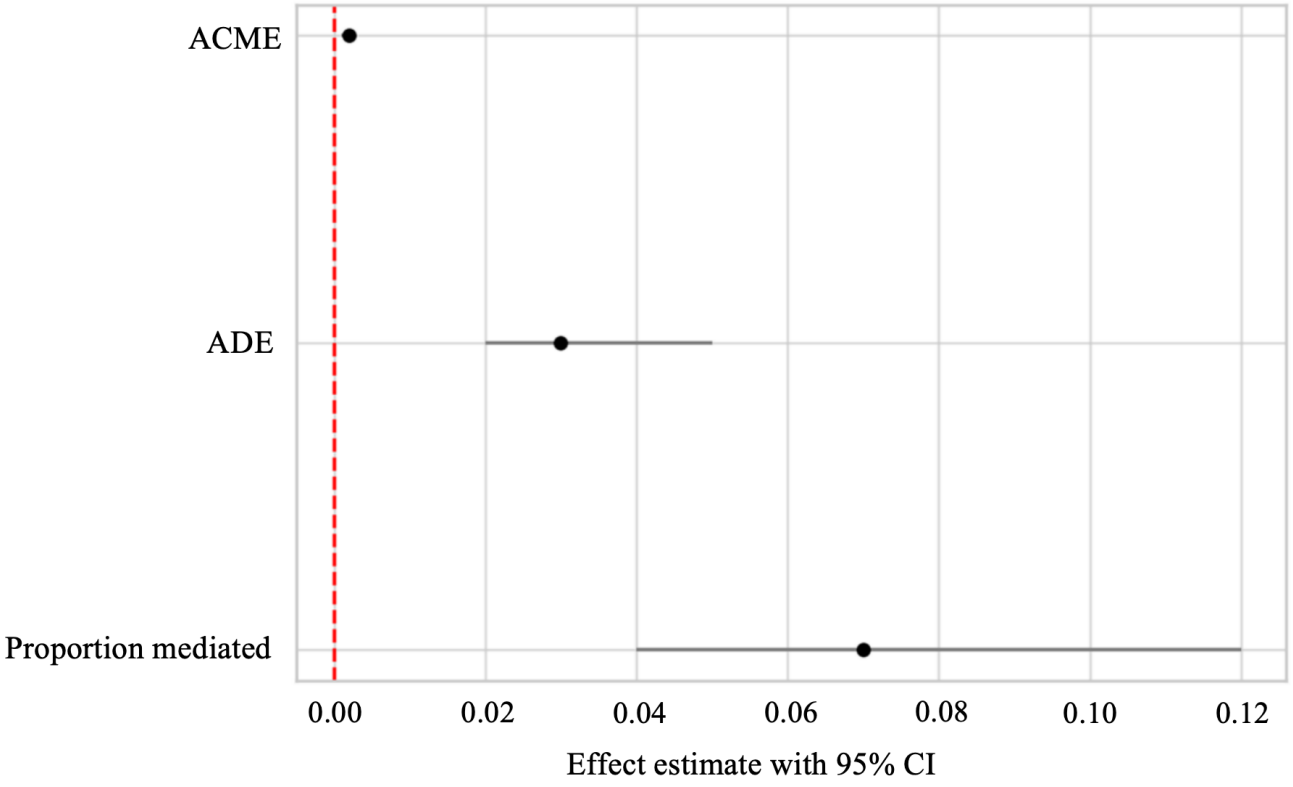
The proportion of the total effect of SS on all-cause mortality that is mediated through a healthy lifestyle was 0.07 (95% CI 0.04-0.12), with a *P* value <.001. This suggests that approximately 7% of the association between SS and mortality is explained by the mediation effect of a healthy lifestyle. The average direct effect of SS on all-cause mortality, independent of the mediation through a healthy lifestyle, was 0.03 (95% CI 0.02-0.05), with a *P* value <.001. The average causal mediation effect was 0.002 (95% CI 0.001-0.002), with a *P* value <.001. These findings demonstrate that a healthy lifestyle partially mediates the relationship between SS and all-cause mortality (Figure 3).

After multiplicative interaction, the interaction term exponentiated coefficient (exp [coef]) is 0.941, and the *P* value

XSL•FO
RenderX

was less than .001, indicating that there was a significant multiplicative interaction between social pressure and a healthy lifestyle. The additive interaction results showed that the relative excess risk due to interaction was measured to be 0.127 (95% CI 0.013 - 0.240), indicating the existence of a positive additive interaction. The attributable proportion due to interaction was found to be 0.071 (95% CI 0.017 - 0.124), indicating that 7.1% of the combined mortality risk caused by social pressure and unhealthy lifestyle can be attributed to the synergistic effect of both factors. The Synergy Index was calculated to be 1.189 (95% CI 1.060 - 1.334), indicating that when SS coexists with unhealthy lifestyles, the risk of death is 1.189 times higher than the sum of their individual effects.

Figure 3. The mediating effect of a healthy lifestyle in the relationship between social stress and all-cause mortality. ACME: average causal mediation effects; ADE: average direct effects.



Stratified and Sensitivity Analyses

We analyzed the effects of SS and a healthy lifestyle on the oldest- and nonoldest-old. The results showed that there was a significant association between SS and mortality risk in both the ≥80 years and <80 years age groups (Table S8 in Multimedia Appendix 1). In the older adults and nonolder adults groups, regardless of stress level, the mortality risk (HR) increased significantly as the healthy lifestyle score decreased (from ≥3 points to 0 points) (Figure S1 and Table S9 in Multimedia Appendix 1). In the older adults and nonolder adults groups, healthy lifestyle partially mediated the relationship between SS and mortality risk (Table S10 in Multimedia Appendix 1). The associations of SS, healthy lifestyle, and all-cause mortality were generally similar across sex (Table S11 in Multimedia Appendix 1), residence (Table S12 in Multimedia Appendix 1), occupation (Table S13 in Multimedia Appendix 1), and education (Table S14 in Multimedia Appendix 1). The associations of SS, healthy lifestyle, and all-cause mortality did not change appreciably when we excluded participants who died within 1 or 2 years of follow-up (Table S15 in Multimedia Appendix 1). The mediating effects of different components of a healthy lifestyle (smoking, alcohol consumption, physical activity, and diet scores) on the relationship between SS and mortality were further analyzed (Table S16 in Multimedia Appendix 1). Association of SS with all-cause mortality after adjustment for weighted lifestyle score and other covariates (Table S17 in Multimedia Appendix 1). Since the population with lifestyle score=2 and lifestyle score=3 each accounted for 1/3 of the other population (lifestyle score=0 or 1), we combined the population with score 2 and score 3 for sensitivity analysis, and the results remained robust (Table S18 in Multimedia Appendix 1).

Discussion

Principal Findings

In this large prospective cohort study, we investigated the impact of SS and a healthy lifestyle on mortality among older adults in China. Our findings showed that both medium and high levels of SS were associated with increased mortality risks, with HRs of 1.16 and 1.28, respectively, compared to the low SS group after full adjustment for covariates. For individuals aged <80 years and ≥80 years, the HRs of the high SS group were 2.09 and 1.38. SS has a greater impact on the mortality of older adults younger than 80 years. These results underscore the significant impact of SS on health outcomes in older adults. Approximately 7% of the association between SS and mortality was mediated through the healthy lifestyle. Under different SS, the lower the healthy lifestyle score, the higher the risk of mortality. These

XSL•FO
RenderX

results provide valuable insights into the complex interplay between SS and health behaviors in shaping mortality risks among older adults in China.

Chronic SS has been shown to induce a range of physiological and behavioral changes that contribute to the development or progression of various health conditions [20]. Previous studies have demonstrated that prolonged exposure to stress elevates plasma cortisol levels, which may lead to memory impairment [21]. Mild stress could help improve cognitive function, but if the intensity of stress exceeded a predetermined threshold (which is different for each person), it could lead to cognitive impairment [22]. Stress could affect the function of the immune system by regulating the processes of the central nervous system and the neuroendocrine system [23]. In addition, stress could affect the occurrence of cardiovascular disease, gastrointestinal disease, endocrine disease, and other diseases [24-26]. A study from China explored the relationship between SS and health inequality and confirmed that SS has a greater impact on people with low economic status [14]. However, most existing studies focus on the impact of SS on psychological or psychiatric diseases. In addition, there is currently a lack of research on the impact of SS on the survival of Chinese older adults and the role of lifestyle in this association. With the aging trend in China, it is necessary to explore the relationship between SS and the health of older adults [27,28]. In our study, we found that compared with low SS, moderate and high SS increased the risk of death in Chinese people aged 65 years and more than 65 years to varying degrees, and lifestyle played a partial mediating role. This study provides causal support for the relationship between SS and the health of older adults through prospective analysis.

Although stress and its effects occur at all stages of an individual's life, with the rapid development of China's aging population, the social pressure faced by older adults has increased, mainly including pressure from interpersonal relationships, the economy, and the environment [29,30]. These components are interrelated and should be examined collectively in relation to health outcomes in older adults. Our study divided these 3 variables into 3 levels, such as low, medium, and high, based on LCA analysis. The results showed that among adults aged 65 years and older, compared with low SS, medium SS increased the risk of death by 16%, and high SS increased the risk of death by 28%. Approximately 7% of the association between SS and mortality is explained by the mediation effect of a healthy lifestyle. This indicates that SS directly influences mortality risk among older adults, and although a healthy lifestyle partially mediates this relationship, the effect size is small. We speculated that there may be several reasons. First, for older adults, the impact of SS may have already had a cumulative effect on their physical health earlier in life, and adopting a healthy lifestyle may not be sufficient to fully offset this long-term negative impact. Second, with age, declining physical functioning, accumulated chronic diseases, and a weakened immune system may reduce older adults' ability to benefit from a healthy lifestyle. Third, social pressure is a complex stressor that may affect the health of older adults' pathways, such as mental well-being, immune function, quality of life, and other mechanisms. Older adults may face higher levels of loneliness and reduced social support, and a healthy lifestyle may not sufficiently mitigate the physiological and psychological effects of these pressures.

Furthermore, participants aged ≥80 years were analyzed separately from younger older adults. We found that SS had a greater impact on mortality risk among older adults aged less than 80 years. Compared with the low SS group, mortality risk increased by 35% and 109% in the medium and high SS groups, respectively. Among adults aged 80 years and older than 80 years, the risk of death increased by 28% and 38% in the medium and high SS groups, respectively, compared to the low SS group. This suggests that SS may pose a greater threat to the survival of younger-old adults. Several factors may help explain this phenomenon. First, younger-old adults may be in a critical stage of lifestyle change, such as facing adaptation after retirement, changes in social roles, etc. Stress may accelerate the occurrence or aggravation of cardiovascular diseases, diabetes, and other chronic diseases, potentially leading to premature death [31]. Second, younger-older adults may not fully cope with SS, resulting in a surge in emotions, such as anxiety and depression [32]. Third, younger-older adults may experience greater economic pressure, including ongoing family expenses and housing challenges [33]. Fourth, in response to a rapidly changing social environment, younger-old adults may adopt unhealthy coping strategies, such as smoking or alcohol consumption, which can exacerbate health risks [34]. Therefore, our findings suggest that age-specific interventions may be necessary to reduce the impact of SS on the health and survival of older adults. This study investigated the interaction between SS and healthy lifestyle on all-cause mortality, using both multiplicative and additive interaction models. The findings revealed a statistically significant interaction effect, suggesting that the relationship between SS and mortality is modified by a healthy lifestyle. The additive interaction highlights the importance of addressing both SS and unhealthy lifestyles simultaneously. Interventions targeting stress reduction alongside lifestyle promotion could yield greater mortality benefits than addressing either factor in isolation.

Studies have shown that SS affects health through a variety of biological pathways. For example, long-term SS may increase cortisol levels in the body, which may lead to decreased immune function and enhanced inflammatory response, thereby increasing the risk of illness or death [35]. Our study suggests that reducing SS among older adults may help improve their quality of life. Moreover, our study found that social pressure had a greater impact on mortality risk among males, as well as among older adults who lived in urban areas, had never been employed, and had received more than 5 years of education. These findings offer valuable insights for future research on SS among older adults.

Our study has several major strengths, including a large sample size and a prospective cohort study. To our knowledge, no previous study has provided a detailed analysis of SS and mortality among the older adult population in China. Nevertheless, our study also has some limitations. First, data related to SS and data related to healthy lifestyles were based on self-reports, which, despite strict collection criteria, are still prone to measurement errors. Second, cause-of-death data were

XSL•FO
RenderX

not comprehensively recorded, limiting our ability to assess the impact of SS on cause-specific mortality among older adults. Third, residual confounding could not be fully eliminated due to limited covariates, such as the absence of genetic data and information on other health conditions. Fourth, this study only focuses on the social pressure composed of interpersonal relationships, economic status, and living environment and may not have considered other related factors. Fifth, although our lifestyle score included key behavioral factors, such as smoking, diet, physical activity, and alcohol consumption, other potentially important components, such as sleep duration and quality, perceived stress, and sedentary behavior, were not included due to data limitations. Sixth, we did not consider the potential changes in lifestyle factors during this period, which may affect the observed results. Therefore, more relevant studies are needed to provide evidence for the present findings.

Conclusions

The results showed that SS was an important risk factor for all-cause mortality in the Chinese older adults. This association remained strong even after adjusting for various sociodemographic and health-related factors, highlighting the importance of SS as a determinant of mortality risk in the Chinese older adults. A healthy lifestyle only partially mediated the relationship between SS and survival among older adults. However, an unhealthy lifestyle was associated with an increased risk of mortality, regardless of SS level.

Acknowledgments

The authors express their gratitude to the participants and staff involved in data collection and management in the Chinese Longitudinal Healthy Longevity Study. This study was supported by the National Natural Science Foundation of China (82425052) and the Postdoctoral Fellowship Program of CPSF (GZC20231052).

Data Availability

Data are available from the Chinese Longitudinal Healthy Longevity Study upon request [36]. Deidentified datasets will be provided following approval, with no restrictions on noncommercial use.

Authors' Contributions

XS, YL, and CM collected the data. JY and JH analyzed and interpreted the data. JY and JH wrote the manuscript. QH, JG, DL, and ZL checked the data and results. CM and XS were co-senior authors for this article.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional data supporting the primary analysis.
[DOCX File, 5723 KB - aging_v8i1e75942_app1.docx]

References

1. Avison WR, Turner RJ. Stressful life events and depressive symptoms: disaggregating the effects of acute stressors and chronic strains. J Health Soc Behav 1988 Sep;29(3):253-264. [Medline: 3241066]
2. Farnworth M, Leiber MJ. Strain theory revisited: economic goals, educational means, and delinquency. Am Sociol Rev 1989 Apr;54(2):263. [doi: 10.2307/2095794]
3. La Gory M, Ritchey FJ, Mullis J. Depression among the homeless. J Health Soc Behav 1990 Mar;31(1):87-102. [Medline: 2313079]
4. Franklin BA, Rusia A, Haskin-Popp C, Tawney A. Chronic stress, exercise and cardiovascular disease: placing the benefits and risks of physical activity into perspectiv. Int J Environ Res Public Health 2021 Sep 21;18(18):9922. [doi: 10.3390/ijerph18189922] [Medline: 34574843]
5. Mangiaterra S, Marker L, Cerquetella M, et al. Chronic stress-related gastroenteric pathology in cheetah: relation between intrinsic and extrinsic factors. Biology (Basel) 2022 Apr 15;11(4):606. [doi: 10.3390/biology11040606] [Medline: 35453805]
6. Fali T, Vallet H, Sauce D. Impact of stress on aged immune system compartments: overview from fundamental to clinical data. Exp Gerontol 2018 May;105:19-26. [doi: 10.1016/j.exger.2018.02.007] [Medline: 29427753]
7. Li Y, Schoufour J, Wang DD, et al. Healthy lifestyle and life expectancy free of cancer, cardiovascular disease, and type 2 diabetes: prospective cohort study. BMJ 2020 Jan 8;368:l6669. [doi: 10.1136/bmj.l6669] [Medline: 31915124]
8. Tehee M, Buchwald D, Booth-LaForce C, Omidpanah A, Manson SM, Goins RT. Traumatic stress, social support, and health among older American Indians: the native elder care study. J Gerontol B Psychol Sci Soc Sci 2019 Jun 14;74(5):908-917. [doi: 10.1093/geronb/gbx163] [Medline: 29304244]
9. Chokkanathan S. Stressors social support and elder mistreatment. Aging Ment Health 2017 Feb;21(2):125-132. [doi: 10.1080/13607863.2015.1081151] [Medline: 26370609]

XSL•FO
RenderX

10. de Frias CM, Whyne E. Stress on health-related quality of life in older adults: the protective nature of mindfulness. Aging Ment Health 2015;19(3):201-206. [doi: 10.1080/13607863.2014.924090] [Medline: 24940847]
11. Wang J, Chen C, Zhou J, et al. Healthy lifestyle in late-life, longevity genes, and life expectancy among older adults: a 20-year, population-based, prospective cohort study. Lancet Healthy Longev 2023 Oct;4(10):e535-e543. [doi: 10.1016/S2666-7568(23)00140-X] [Medline: 37804845]
12. Zeng Y, Feng Q, Hesketh T, Christensen K, Vaupel JW. Survival, disabilities in activities of daily living, and physical and cognitive functioning among the oldest-old in China: a cohort study. Lancet 2017 Apr;389(10079):1619-1629. [doi: 10.1016/S0140-6736(17)30548-2]
13. Chinese Longitudinal Healthy Longevity and Happy Family Study (CLHLS-HF). Peking University Open Research Data. URL: https://opendata.pku.edu.cn/dataverse/CHADS [accessed 2024-11-13]
14. Zhang Y. Social stress and health inequality: an empirical study based on CLHLS [Article in Chinese]. South China Population 2023;38(1):1-9. [doi: 10.3969/j.issn.1004-1613.2023.01.001]
15. Birmingham WC, Holt-Lunstad J. Social aggravation: understanding the complex role of social relationships on stress and health-relevant physiology. Int J Psychophysiol 2018 Sep;131:13-23. [doi: 10.1016/j.ijpsycho.2018.03.023] [Medline: 29626504]
16. Engeln M, Fox ME, Lobo MK. Housing conditions during self-administration determine motivation for cocaine in mice following chronic social defeat stress. Psychopharmacology (Berl) 2021 Jan;238(1):41-54. [doi: 10.1007/s00213-020-05657-y] [Medline: 32914243]
17. Tobacco. World Health Organization. URL: https://www.who.int/health-topics/tobacco#tab=tab_1 [accessed 2024-11-14]
18. International guide for monitoring alcohol consumption and related harm. World Health Organization. URL: https://iris.who.int/handle/10665/66529 [accessed 2024-11-14]
19. Zhang YB, Chen C, Pan XF, et al. Associations of healthy lifestyle and socioeconomic status with mortality and incident cardiovascular disease: two prospective cohort studies. BMJ 2021 Apr 14;373:n604. [doi: 10.1136/bmj.n604] [Medline: 33853828]
20. Yaribeygi H, Panahi Y, Sahraei H, Johnston TP, Sahebkar A. The impact of stress on body function: a review. EXCLI J 2017;16:1057-1072. [doi: 10.17179/excli2017-480] [Medline: 28900385]
21. Lupien SJ, McEwen BS, Gunnar MR, Heim C. Effects of stress throughout the lifespan on the brain, behaviour and cognition. Nat Rev Neurosci 2009 Jun;10(6):434-445. [doi: 10.1038/nrn2639] [Medline: 19401723]
22. Sandi C. Stress and cognition. Wiley Interdiscip Rev Cogn Sci 2013 May;4(3):245-261. [doi: 10.1002/wcs.1222] [Medline: 26304203]
23. Segerstrom SC, Miller GE. Psychological stress and the human immune system: a meta-analytic study of 30 years of inquiry. Psychol Bull 2004 Jul;130(4):601-630. [doi: 10.1037/0033-2909.130.4.601] [Medline: 15250815]
24. Wu ZS. Epidemiological studies on the relationship between psychosocial factors and cardiovascular disease. Zhongguo Yi Xue Ke Xue Yuan Xue Bao 2001 Feb;23(1):73-77. [Medline: 12905825]
25. Konturek PC, Brzozowski T, Konturek SJ. Stress and the gut: pathophysiology, clinical consequences, diagnostic approach and treatment options. J Physiol Pharmacol 2011 Dec;62(6):591-599. [Medline: 22314561]
26. Charmandari E, Tsigos C, Chrousos G. Endocrinology of the stress response. Annu Rev Physiol 2005;67:259-284. [doi: 10.1146/annurev.physiol.67.040403.120816] [Medline: 15709959]
27. Jin S, Li C, Cao X, Chen C, Ye Z, Liu Z. Association of lifestyle with mortality and the mediating role of aging among older adults in China. Arch Gerontol Geriatr 2022 Jan;98:104559. [doi: 10.1016/j.archger.2021.104559]
28. Chen Y, Sun B, Wang C, Zhang K, Lu W, Huang G. Associations of types of grains and lifestyle with all-cause mortality among Chinese adults aged 65 years or older: a prospective cohort study. J Transl Med 2023;21(1):88. [doi: 10.1186/s12967-023-03927-9]
29. Sharifi S, Babaei Khorzoughi K, Khaledi-Paveh B, Rahmati M. Association of intergenerational relationship and supports with cognitive performance in older adults: a systematic review. Geriatr Nurs (Lond) 2023 Jul;52:146-151. [doi: 10.1016/j.gerinurse.2023.05.014]
30. Fangfang H, Xiao H, Shuai Z, et al. Living environment, built environment and cognitive function among older Chinese adults: results from a cross-sectional study. J Prev Alzheimers Dis 2022 Jan;9(1):126-135. [doi: 10.14283/jpad.2021.59] [Medline: 35098983]
31. Schultz WM, Kelli HM, Lisko JC. Socioeconomic status and cardiovascular outcomes: challenges and interventions. Circulation 2018 May 15;137(20):2166-2178. [doi: 10.1161/CIRCULATIONAHA.117.029652] [Medline: 29760227]
32. Thoits PA. Stress and health: major findings and policy implications. J Health Soc Behav 2010;51 Suppl:S41-S53. [doi: 10.1177/0022146510383499] [Medline: 20943582]
33. More than 1 in 10 people in Japan are aged 80 or over. Here's how its ageing population is reshaping the country. World Economic Forum. URL: https://www.weforum.org/stories/2023/09/elderly-oldest-population-world-japan/ [accessed 2024-12-18]
34. Villiers-Tuthill A, Copley A, McGee H, Morgan K. The relationship of tobacco and alcohol use with ageing self-perceptions in older people in Ireland. BMC Public Health 2016 Jul 22;16:627. [doi: 10.1186/s12889-016-3158-y] [Medline: 27448397]

XSL•FO
RenderX

35. McEwen BS. Protective and damaging effects of stress mediators: central role of the brain. Dialogues Clin Neurosci 2006 Dec 31;8(4):367-381. [doi: 10.31887/DCNS.2006.8.4/bmcewen]
36. The Chinese Longitudinal Healthy Longevity Survey (CLHLS)-Longitudinal Data (1998-2018). Peking University Open Research Data. URL: http://opendata.pku.edu.cn/dataset.xhtml?persistentId=doi:10.18170/DVN/WBO7LK [accessed 2025-06-05]

Abbreviations

CLHLS: Chinese Longitudinal Healthy Longevity Survey
HR: hazard ratio
LCA: latent class analysis
SS: social stress

Edited by J Job; submitted 13.04.25; peer-reviewed by H Duan, Y Wei; revised version received 11.06.25; accepted 13.06.25; published 12.08.25.

Please cite as:
Yang J, Huang J, Huang Q, Gao J, Liu D, Li Z, Lv Y, Shi X, Mao C
The Impact of Social Stress and Healthy Lifestyle on the Mortality of Chinese Older Adults: Prospective Cohort Study
JMIR Aging 2025;8:e75942
URL: https://aging.jmir.org/2025/1/e75942
doi:10.2196/75942

© Jin Yang, Jilong Huang, Qingmei Huang, Jian Gao, Dan Liu, Zhihao Li, Yuebin Lv, Xiaoming Shi, Chen Mao. Originally published in JMIR Aging (https://aging.jmir.org), 12.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Edentulousness and the Likelihood of Becoming a Centenarian: Longitudinal Observational Study

Xindi Wei[1*], MD; Longfei Zhuang[2*], PhD; Yuan Li[1], PhD; Junyu Shi[1], PhD; Yijie Yang[1], MD; Hongchang Lai[1], PhD; Beilei Liu[1], PhD

[1]Shanghai Ninth People's Hospital, School of Medicine, Shanghai Jiao Tong University, 639 Zhizaoju Road, Shanghai, China

[2]Huishang Dental Clinic, Shanghai, China

[*]these authors contributed equally

Corresponding Author:
Beilei Liu, PhD
Shanghai Ninth People's Hospital, School of Medicine, Shanghai Jiao Tong University, 639 Zhizaoju Road, Shanghai, China

Abstract

Background: In recent decades, the global life expectancy has risen notably to approximately 73.5 years worldwide, coinciding with a rapid growth in the older adult population, which presents a significant public health challenge in promoting healthy aging and longevity.

Objective: This study aimed to prospectively investigate the link between edentulousness and the likelihood of reaching centenarian status among individuals aged 80 years and older.

Methods: Data from the Chinese Longitudinal Healthy Longevity Survey were analyzed. Logistic regression models were used to assess the relationship between edentulousness and the likelihood of becoming a centenarian. Demographic characteristics, lifestyle habits, and disease histories were adjusted as confounding factors. Several sensitivity analyses, including propensity score matching and 2-year lag analyses, were conducted to further assess the association between edentulousness and the likelihood of becoming a centenarian. The correlation between the number of natural teeth as a continuous variable and the likelihood of becoming a centenarian was evaluated as well.

Results: The study included 4239 participants aged 80-100 years. After adjusting for all covariates, the likelihood for becoming a centenarian increased in the nonedentulous group compared to the edentulous group (odds ratio [OR] 1.384, 95% CI 1.093 - 1.751). The relationship persisted after propensity score matching analysis (OR 1.272, 95% CI 1.037 - 1.561). The association remained statistically significant after excluding participants with a follow-up duration of less than 2 years (OR 1.522, 95% CI 1.083 - 2.140; *P*=.02). Furthermore, a significant positive association between the number of natural teeth and the likelihood of becoming a centenarian was found after adjusting for all covariates (OR 1.022, 95% CI 1.002 - 1.042; *P*=.03), which aligned with the main results of the study.

Conclusions: The findings revealed that the presence of natural teeth was linked to an increased probability of becoming a centenarian, underscoring the importance of maintaining oral health even in advanced age.

(JMIR Aging 2025;8:e68444) doi:10.2196/68444

KEYWORDS

public health; edentulous; oral-systemic disease; epidemiology; cohort studies

Introduction

Over the past few decades, the global life expectancy at birth has significantly risen to approximately 73.5 years worldwide and 77.6 years in mainland China [1,2]. Additionally, the proportion of individuals aged 65 years and older is rapidly expanding, constituting 13.5% of the population in China [3]. This increase in life expectancy has coincided with a rapid expansion of the aging population, presenting a substantial public health challenge in fostering healthy aging and longevity. However, focusing solely on traditional determinants of longevity is insufficient, underscoring the importance of identifying other modifiable risk factors.

Oral diseases are widespread globally, affecting over 3.5 billion individuals [4,5], yet their influence on longevity is often overlooked. In addition to sociodemographic, genetic, and lifestyle factors [6], tooth loss—the ultimate outcome of oral diseases and a surrogate for overall health status—is essential for healthy aging as well. The retention of natural teeth over a lifetime is a fundamental aspect of overall well-being [7]. Tooth loss can be a potential risk factor for longevity that is independently related to the onset of disability and mortality in old age [7-9]. For instance, the 6-year mortality rate of

XSL•FO
RenderX

individuals with edentulousness (the lack of teeth) who do not use dentures was significantly higher than that of the individuals with ≥20 teeth [10]. Another study based on Baltimore Longitudinal Study of Aging indicated that being edentulous or having fewer than 20 teeth was independently associated with the mortality of older adults [11]. Centenarians comprise those who successfully age and have good resilience, and the number of centenarians is increasing worldwide [12-14]. Centenarians and their offspring demonstrated better oral health, suggesting the potential relationship between natural teeth retention and longevity [8]. Moreover, possessing ≤20 natural teeth was an independent risk factor for frailty among centenarians [3]. However, there is no cohort study that focuses on the impact of natural teeth retention on achieving centenarian status; thus, leveraging aging-focused, big data analytics to understand the intricate relationship among the older adult population becomes imperative.

The objective of this study was to prospectively investigate the association between edentulousness and the likelihood of becoming a centenarian in individuals aged 80 years and older, using data from the Chinese Longitudinal Healthy Longevity Survey (CLHLS)—an aging-focused, nationally representative cohort of the older Chinese population. These findings have the potential to inform interventions aimed at maximizing life expectancy.

Methods

Study Design and Population

The CLHLS is a comprehensive nationwide study that uses random sampling techniques to select participants from half of the counties and cities in 22 out of the 31 provinces across mainland China, covering approximately 85% of the total population [15]. The CLHLS collected data from 8 waves of surveys carried out in 1998, 2000, 2002, 2005, 2008, 2011, 2014, and 2018. Each survey round involved follow-ups with existing participants and the recruitment of new participants.

This cohort study used the baseline data from 1998, and the mortality follow-up data were from 1998, 2000, 2002, 2005, 2008, 2011, 2014, and 2018. As long as the individuals were aged 80 years or older in 1998, they could potentially live to 100 years in 2018. Thus, 6675 participants (aged ≥80 years and <100 years) who had the potential to age to 100 years or older were first included in this study. The participants who were lost to follow-up or had incomplete information on confounders were then excluded, and 4239 participants were included in the final analyses.

This study adheres to the STROBE (Standards for Reporting of Observational Studies in Epidemiology) guidelines [16].

Exposure Variable

The number of natural teeth of the participants was collected in the survey. Participants were divided into two categories based on the existence of natural teeth. The edentulous group was defined as participants with the complete loss of all dentition, while the nonedentulous group was defined as participants with at least 1 tooth.

Outcome Variable

Interviews were conducted with a close family member of the participants who had passed away between the previous wave's interview and the subsequent survey. The year of death and the age at death were collected. Individuals reaching the age of 100 years were defined as centenarians.

Covariates

Age in 1998, sex (male or female), ethnicity (Han Chinese or other), residence (urban or rural), marital status (currently married and living with a spouse, widowed, separated, divorced, or never married), exercise habits (yes or no), smoking status (nonsmoker, former smoker, or current smoker), alcohol use (nondrinker, former drinker, or current drinker), denture use (yes or no), diabetes (yes or no), hypertension (yes or no), stroke or cardiovascular disease (CVD; yes or no), and cancer (yes or no) were selected as covariates [6]. Denture use included the use of complete dentures and removable partial dentures [17]. All data were obtained by a face-to-face interview via a questionnaire by well-trained interviewers from the local Centers for Disease Control and Prevention.

Statistical Analysis

In the descriptive statistics, median and IQR were used for continuous variables, and frequency distributions were used for categorical variables. Continuous variables were compared by ANOVA tests for variables meeting the assumptions of normal distribution and homogeneity of variance, and by Kruskal-Wallis H test for those not meeting these assumptions. Categorical variables were compared using χ^2 tests. The univariate and multivariate logistic regression models were adopted to evaluate the odd ratios (ORs) and 95% CIs pertaining to the association between edentulousness and the likelihood of becoming a centenarian. To control for confounders, three logistic regression models were constructed to eliminate the influence of covariates. No variables were adjusted in model 1. Model 2 adjusted for the age in 1998, sex, ethnicity, marital status, smoking status, alcohol use, and exercise habits. Model 3 further incorporated adjustments for denture use, diabetes, hypertension, stroke or CVD, and cancer. Several sensitivity analyses were performed. First, a 1:1 propensity score matching (PSM) analysis was conducted to balance the differences between the edentulous and nonedentulous groups using the *MatchIt* R package [18], which adjusted for the age in 1998, sex, ethnicity, marital status, exercise habits, smoking status, alcohol use, denture use, diabetes, hypertension, stroke or CVD, and cancer. The data after PSM were then analyzed using logistic regression to confirm the association between edentulousness and the likelihood of becoming a centenarian, with only the covariates that were still significant after PSM being adjusted. Second, a 2-year lag analysis was conducted by excluding individuals whose follow-ups were less than 2 years. Third, the association between the number of natural teeth as a continuous variable and the likelihood of becoming a centenarian was explored by logistic regression models.

A 2-sided *P* value <.05 was considered statistically significant in all analyses. All statistical analyses were performed by R (version 4.3.1; R Foundation for Statistical Computing).

XSL•FO
RenderX

Ethical Considerations

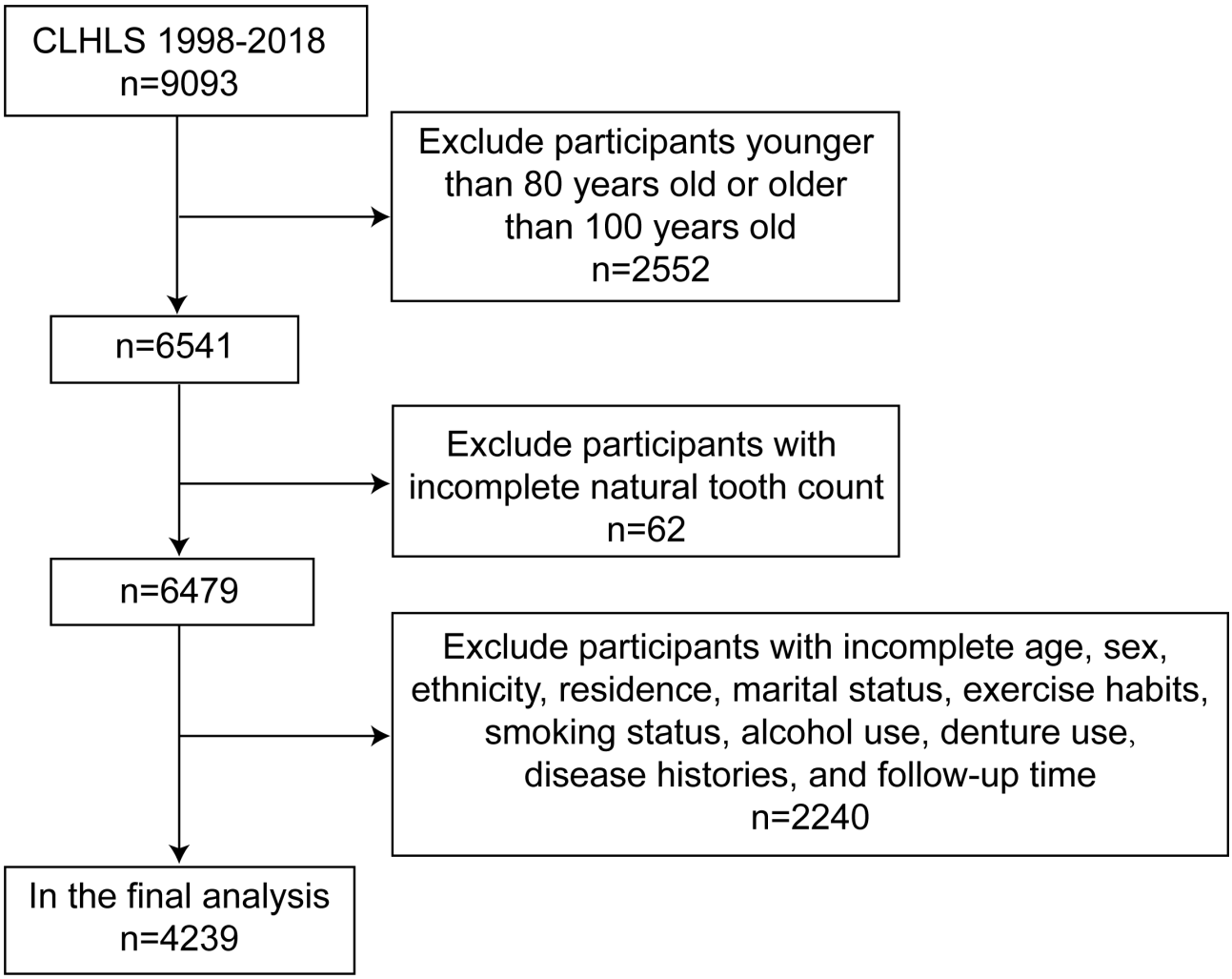
Ethical approval was granted by the Biomedical Ethics Committee of Peking University (IRB00001052-13,074), and all participants or their proxy respondents provided informed consent without receiving financial compensation. The data have been anonymized to protect the privacy of participants.

Results

Baseline Characteristics of the Study Sample

As depicted in Figure 1, this study included 4239 participants from the CLHLS 1998‐2018. Table 1 demonstrated the baseline characteristics of these participants. Significant differences were detected among the participants in various factors, including age (P<.001), sex (P<.001), ethnicity (P=.03), denture use (P<.001), smoking status (P<.001), alcohol use (P<.001), and hypertension (P=.02). There were no significant differences between the two groups in terms of residence (P=.09), exercise habits (P=.87), diabetes (P>.99), stroke or CVD (P=.17), and cancer (P=.47). Overall, 607 (14.3%) of the 4239 individuals became centenarians during follow-up, and 264 (43.5%) of these 607 centenarians were edentulous.

Figure 1. Flowchart of the process for the selection of eligible participants. CLHLS: Chinese Longitudinal Healthy Longevity Survey.



XSL•FO
RenderX

Table . Characteristics of the participants (n=4239).

Characteristics	Noncentenarians (n=3632)	Centenarians (n=607)	Statistic (*df*)	*P* value
Age (years), median (IQR)	88 (84-93)	98 (96-99)	1056.900 (1)	<.001[a]
Sex, n (%)			40.334 (1)	<.001[b]
Male	1795 (49.4)	215 (35.4)		
Female	1837 (50.6)	392 (64.6)		
Ethnicity, n (%)			4.596 (1)	.03[b]
Han Chinese	3336 (91.9)	541 (89.1)		
Other	296 (8.1)	66 (10.9)		
Residence, n (%)			2.796 (1)	.09[b]
Urban	540 (14.9)	74 (12.2)		
Rural	3092 (85.1)	533 (87.8)		
Marital status, n (%)			40.658 (4)	<.001[b]
Currently married and living with a spouse	759 (20.9)	62 (10.2)		
Separated	51 (1.4)	8 (1.3)		
Divorced	22 (0.6)	3 (0.5)		
Widowed	2751 (75.7)	529 (87.1)		
Never married	49 (1.3)	5 (0.8)		
Exercise habits, n (%)			0.027 (1)	.87[b]
Yes	990 (27.3)	168 (27.7)		
No	2642 (72.7)	439 (72.3)		
Smoking status, n (%)			32.136 (2)	<.001[b]
Current smoker	799 (22)	78 (12.9)		
Former smoker	600 (16.5)	89 (14.7)		
Nonsmoker	2233 (61.5)	440 (72.5)		
Alcohol use, n (%)			21.517 (2)	<.001[b]
Current drinker	943 (26)	120 (19.8)		
Former drinker	441 (12.1)	52 (8.6)		
Nondrinker	2248 (61.9)	435 (71.7)		
Self-reported diabetes, n (%)			<0.001 (1)	>.99[b]
Yes	23 (0.6)	4 (0.7)		
No	3609 (99.4)	603 (99.3)		
Self-reported hypertension, n (%)			5.042 (1)	.02[b]
Yes	502 (13.8)	63 (10.4)		
No	3130 (86.2)	544 (89.6)		
Self-reported stroke or CVD[c], n (%)			1.849 (1)	.17[b]
Yes	133 (3.7)	15 (2.5)		
No	3499 (96.3)	592 (97.5)		
Self-reported cancer, n (%)			0.528 (1)	.47[b]
Yes	17 (0.5)	1 (0.2)		
No	3615 (99.5)	606 (99.8)		

XSL•FO
RenderX

Characteristics	Noncentenarians (n=3632)	Centenarians (n=607)	Statistic (*df*)	*P* value
Edentulousness, n (%)			23.847 (1)	<.001[b]
Yes	1206 (33.2)	264 (43.5)		
No	2426 (66.8)	343 (56.5)		
Denture use, n (%)			15.562 (1)	<.001[b]
Yes	784 (21.6)	88 (14.5)		
No	2848 (78.4)	519 (85.5)		

[a]Kruskal-Wallis H test.

[b]Chi-square test.

[c]CVD: cardiovascular disease.

Correlation Between Edentulousness and the Likelihood of Becoming a Centenarian

A significant association between edentulousness and the likelihood of becoming a centenarian was found among the participants. After adjusting for all covariates in model 3, the likelihood for becoming a centenarian increased in the nonedentulous group compared to the edentulous group (odds ratio [OR] 1.384, 95% CI 1.093 - 1.751; *P*=.007; Table 2).

Table . Association between edentulousness and the likelihood of becoming a centenarian.

Edentulousness	Participants (N=4239), n (%)	Model 1[a]		Model 2[b]		Model 3[c]	
		OR[d] (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Yes	1470 (34.7)	Ref[e]		Ref		Ref	
No	2769 (65.3)	0.646 (0.542-0.769)	<.001	1.311 (1.042-1.648)	.02	1.384 (1.093-1.751)	.007

[a]Model 1: unadjusted.

[b]Model 2: model 1 plus additional adjustment for the age in 1998, sex, ethnicity, marital status, smoking status, alcohol use, and exercise habits.

[c]Model 3: model 2 plus additional adjustment for denture use and disease histories (including diabetes, hypertension, stroke or cardiovascular disease, and cancer).

[d]OR: odds ratio.

[e]Ref: reference.

Sensitivity Analyses

The robustness of the observed association was confirmed in the sensitivity analyses. PSM analysis was used to further evaluate the association between edentulousness and the likelihood of becoming a centenarian. Participants were divided into two groups for PSM analysis—the edentulous group and the nonedentulous group—and nearly all covariates were not significantly different after PSM (Multimedia Appendix 1). A significant increase in the likelihood of becoming a centenarian in the nonedentulous group was observed compared to the reference group (OR 1.272, 95% CI 1.037 - 1.561; *P*=.02; Table 3). Even after excluding participants who were followed up for less than 2 years, the association remained significant (OR 1.522, 95% CI 1.083 - 2.140; *P*=.02; Table 4). Regarding the association between the number of natural teeth as a continuous variable and the likelihood of becoming a centenarian, a notable positive association was found after adjusting for all covariates (OR 1.022, 95% CI 1.002 - 1.042; *P*=.03; Multimedia Appendix 2), which was comparable with the main results of this study.

Table . Association between edentulousness and the likelihood of becoming a centenarian after propensity score matching.

Edentulousness	Participants (n=2560), n (%)	OR[a] (95% CI)	*P* value
Yes	1280 (50)	Ref[b]	
No	1280 (50)	1.272 (1.037-1.561)	.02

[a]OR: odds ratio.

[b]Ref: reference.

XSL•FO
RenderX

Table . Association between edentulousness and the likelihood of becoming a centenarian in a 2-year lag analysis.

Edentulousness	Participants (n=2310), n (%)	Model 1[a]		Model 2[b]		Model 3[c]	
		OR[d] (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Yes	410 (17.7)	Ref[e]		Ref		Ref	
No	1900 (82.3)	0.712 (0.572-0.885)	.002	1.414 (1.017-1.966)	.04	1.522 (1.083-2.140)	.02

[a]Model 1: unadjusted.

[b]Model 2: model 1 plus additional adjustment for the age in 1998, sex, ethnicity, marital status, smoking status, alcohol use, and exercise habits.

[c]Model 3: model 2 plus additional adjustment for denture use and disease histories (including diabetes, hypertension, stroke or cardiovascular disease, and cancer).

[d]OR: odds ratio.

[e]Ref: reference.

Discussion

Principal Findings

In this large-scale cohort study, the association between edentulousness and the likelihood of becoming a centenarian among populations aged 80 years or older was comprehensively analyzed. The likelihood of becoming a centenarian increased for the nonedentulous group compared to the edentulous group. Importantly, the correlation persisted after PSM analysis and was confirmed by the 2-year lag analysis. A significant positive association was found between the number of natural teeth and the likelihood of becoming a centenarian, aligning with the primary findings of this investigation.

The outcomes of this study generally support previous research, suggesting that retaining natural teeth is linked to reduced mortality rates among individuals aged 80 years or older. A previous study indicated that a majority of centenarians expressed contentment with their oral health [19]. A cross-sectional study using data from the New England Centenarian Study unveiled a lower prevalence of tooth loss among centenarians compared to their peers at the ages of 65-74 years, hinting at tooth loss as a potential indicator of decreased longevity. Notably, both centenarians and their descendants exhibited superior oral health compared to the control group [8]. Another cross-sectional study that included 1034 centenarians from the CLHLS dataset indicated that having ≤20 natural teeth was an independent risk factor for frailty among centenarians [3], underscoring the significance of dental health for this age group. A population-based survey from Finland disclosed that even a small number of missing teeth could signify a heightened risk of overall mortality, emphasizing the link between tooth loss and mortality [20]. These findings align with the results of this study, and this research further enhances existing knowledge by showcasing that possessing natural teeth correlates with an increased likelihood of reaching centenarian status through extensive follow-ups and several sensitivity analyses. The robustness of this study was reaffirmed by a series of sensitivity analyses, as the association persisted after PSM analysis and was confirmed by the 2-year lag analysis. Upon examining the relationship between the number of natural teeth and the probability of attaining centenarian status, a significant positive correlation emerged. This association remained robust even after adjusting for all covariates, aligning with the principal findings of this study. These results underscore the importance of managing oral diseases, suggesting that older adults with natural teeth have a greater chance of becoming centenarians.

In this study, the OR changed from 0.646 to 1.384 from model 1 to model 3 for the correlation between edentulousness and the likelihood of becoming a centenarian, suggesting the existence of potential confounding factors. By adjusting the variables that were included in the model, age in 1998 was found to be the variable causing the OR to invert. Given that the outcome of this study was centenarian status, the age of the participants at enrollment could profoundly influence the outcome. Older individuals at enrollment were more likely to reach the age of 100 years. Therefore, to further mitigate this influence, a 2-year lag analysis was conducted in the *Sensitivity Analyses* section, excluding older individuals whose follow-ups were less than 2 years.

The precise mechanism underlying the higher likelihood of reaching centenarian status among individuals with natural teeth remains incompletely understood. Previous studies indicated an association between the number of teeth and masticatory function [21], and reduced chewing ability was related to premature death [22]. Individuals with edentulousness may have an unbalanced food selection, consuming inadequate amounts of fruits and vegetables, and their nutritional status may dispose these individuals to more chronic diseases [7,23], as the dietary intake pattern influences the microbial compositions and systemic inflammation [24].

This study demonstrates notable strengths by using longitudinal data from a sizable, nationally representative cohort of older Chinese individuals, enabling a prospective assessment of the link between edentulousness and achieving centenarian status in older adults. Furthermore, the study used various models that adjusted for multiple variables and conducted several sensitivity analyses to enhance the reliability of the results. Significantly, this research unveils, for the first time, the influence of edentulousness on the likelihood of becoming a centenarian. Nevertheless, this study is subject to various limitations. First, most lifestyle behaviors were self-reported, introducing potential measurement errors. Second, medical conditions were also self-reported, despite detailed explanations provided by trained interviewers during data collection. This method likely led to

XSL•FO
RenderX

an underestimation of disease prevalence, possible misclassification, and residual confounding due to unmeasured medical conditions [6]. Third, despite adjusting for numerous confounding factors, the association could still be influenced by unmeasured or residual confounders, such as nutrition status that may affect both dental conditions and lifespan. Fourth, there may be reverse causality in this study. Individuals in better health status are more likely to retain their teeth, potentially biasing the interpretation of the findings. Finally, the survivorship bias may exist, as the participants included were aged ≥80 years at baseline, which may represent a population in relatively better health and may not be fully representative of all older adults.

Conclusions

In this cohort study involving individuals aged 80 years or older in China, the presence of natural teeth was linked to an increased probability of reaching the age of 100 years, emphasizing the significance of preserving oral health even in advanced age. Implementing targeted intervention strategies for oral health to enhance overall well-being could potentially contribute to longevity. Further prospective research and basic research experiments are essential to validate these findings and illuminate the underlying mechanisms.

Acknowledgments

This study was supported by the Shuguang Program of Shanghai Education Development Foundation and Shanghai Municipal Education Commission (22SG14), the Shanghai Rising Star Project (22QA1405200), the Biomaterials Joint Research Project (2022LHB03), the Biobank Project of the Ninth People's Hospital–affiliated Shanghai Jiao Tong University School of Medicine (YBKA202207), and the National Natural Science Foundation of China (82201094).

Data Availability

The Chinese Longitudinal Healthy Longevity Survey data of this study are available on the web [25].

Authors' Contributions

HL and BL are co-corresponding authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Characteristics of the participants after propensity score matching (n=2560).
[DOCX File, 24 KB - aging_v8i1e68444_app1.docx]

Multimedia Appendix 2
Association between the number of natural teeth and the likelihood of becoming a centenarian.
[DOCX File, 17 KB - aging_v8i1e68444_app2.docx]

References

1. Yang J, Bai R, Wang L, Feng X, Liu Y, Zhou M. A projection of life expectancy based on the Global Burden of Disease Study 2019 - China, 1990-2030. China CDC Wkly 2023 Jan 13;5(2):40-44. [doi: 10.46234/ccdcw2023.008] [Medline: 36776687]
2. GBD 2019 Demographics Collaborators. Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950-2019: a comprehensive demographic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1160-1203. [doi: 10.1016/S0140-6736(20)30977-6] [Medline: 33069325]
3. Zhang J, Xu L. Frailty and associated factors among Chinese centenarians. J Nutr Health Aging 2022;26(8):806-813. [doi: 10.1007/s12603-022-1831-7] [Medline: 35934826]
4. Botelho J, Mascarenhas P, Viana J, et al. An umbrella review of the evidence linking oral health and systemic noncommunicable diseases. Nat Commun 2022 Dec 9;13(1):7614. [doi: 10.1038/s41467-022-35337-8] [Medline: 36494387]
5. Peres MA, Macpherson LMD, Weyant RJ, et al. Oral diseases: a global public health challenge. Lancet 2019 Jul 20;394(10194):249-260. [doi: 10.1016/S0140-6736(19)31146-8] [Medline: 31327369]
6. Li Y, Wang K, Jigeer G, et al. Healthy lifestyle and the likelihood of becoming a centenarian. JAMA Netw Open 2024 Jun 3;7(6):e2417931. [doi: 10.1001/jamanetworkopen.2024.17931] [Medline: 38900423]
7. Friedman PK, Lamster IB. Tooth loss as a predictor of shortened longevity: exploring the hypothesis. Periodontol 2000 2016 Oct;72(1):142-152. [doi: 10.1111/prd.12128] [Medline: 27501497]

XSL•FO
RenderX

8. Kaufman LB, Setiono TK, Doros G, et al. An oral health study of centenarians and children of centenarians. J Am Geriatr Soc 2014 Jun;62(6):1168-1173. [doi: 10.1111/jgs.12842] [Medline: 24889721]
9. Holm-Pedersen P, Schultz-Larsen K, Christiansen N, Avlund K. Tooth loss and subsequent disability and mortality in old age. J Am Geriatr Soc 2008 Mar;56(3):429-435. [doi: 10.1111/j.1532-5415.2007.01602.x] [Medline: 18194226]
10. Shimazaki Y, Soh I, Saito T, et al. Influence of dentition status on physical disability, mental impairment, and mortality in institutionalized elderly people. J Dent Res 2001 Jan;80(1):340-345. [doi: 10.1177/00220345010800010801] [Medline: 11269726]
11. Padilha DMP, Hilgert JB, Hugo FN, Bós AJG, Ferrucci L. Number of teeth and mortality risk in the Baltimore Longitudinal Study of Aging. J Gerontol A Biol Sci Med Sci 2008 Jul;63(7):739-744. [doi: 10.1093/gerona/63.7.739] [Medline: 18693229]
12. Borras C, Ingles M, Mas-Bargues C, et al. Centenarians: an excellent example of resilience for successful ageing. Mech Ageing Dev 2020 Mar;186:111199. [doi: 10.1016/j.mad.2019.111199] [Medline: 31899226]
13. Zeng Y, Shen K. Resilience significantly contributes to exceptional longevity. Curr Gerontol Geriatr Res 2010;2010:525693. [doi: 10.1155/2010/525693] [Medline: 21197075]
14. GBD 2017 Mortality Collaborators. Global, regional, and national age-sex-specific mortality and life expectancy, 1950-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1684-1735. [doi: 10.1016/S0140-6736(18)31891-9] [Medline: 30496102]
15. Zeng Y, Feng Q, Hesketh T, Christensen K, Vaupel JW. Survival, disabilities in activities of daily living, and physical and cognitive functioning among the oldest-old in China: a cohort study. Lancet 2017 Apr 22;389(10079):1619-1629. [doi: 10.1016/S0140-6736(17)30548-2] [Medline: 28285816]
16. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007 Oct 20;370(9596):1453-1457. [doi: 10.1016/S0140-6736(07)61602-X]
17. Corrado OJ. Dentures. BMJ 1990 Dec 1;301(6763):1265-1268. [doi: 10.1136/bmj.301.6763.1265] [Medline: 2271830]
18. Ho D, Imai K, King G, Stuart EA. MatchIt: nonparametric preprocessing for parametric causal inference. J Stat Softw 2011 Jun 14;42(8):1-28. [doi: 10.18637/jss.v042.i08]
19. Rupel K, Tettamanti M, Vella F, et al. What do we learn from the clinical and biological evaluation of the oral cavity in centenarians? Maturitas 2021 Mar;145:31-37. [doi: 10.1016/j.maturitas.2020.12.005] [Medline: 33541560]
20. Liljestrand JM, Havulinna AS, Paju S, Männistö S, Salomaa V, Pussinen PJ. Missing teeth predict incident cardiovascular events, diabetes, and death. J Dent Res 2015 Aug;94(8):1055-1062. [doi: 10.1177/0022034515586352] [Medline: 25991651]
21. Ueno M, Yanagisawa T, Shinada K, Ohara S, Kawaguchi Y. Masticatory ability and functional tooth units in Japanese adults. J Oral Rehabil 2008 May;35(5):337-344. [doi: 10.1111/j.1365-2842.2008.01847.x] [Medline: 18405269]
22. Wei X, Zhang X, Chen R, et al. Diet-related diseases mediate the effect of masticatory function on premature death in older adults. J Periodontal Res 2024 Aug 13. [doi: 10.1111/jre.13335] [Medline: 39136162]
23. Smit MJ, Nijholt W, Bakker MH, Visser A. The predictive value of masticatory function for adverse health outcomes in older adults: a systematic review. J Nutr Health Aging 2024 May;28(5):100210. [doi: 10.1016/j.jnha.2024.100210] [Medline: 38489994]
24. Agurs-Collins T, Alvidrez J, ElShourbagy Ferreira S, et al. Perspective: nutrition health disparities framework: a model to advance health equity. Adv Nutr 2024 Apr;15(4):100194. [doi: 10.1016/j.advnut.2024.100194] [Medline: 38616067]
25. Peking University Open Research Data Platform. URL: https://opendata.pku.edu.cn/ [accessed 2025-03-14]

Abbreviations

CLHLS: Chinese Longitudinal Healthy Longevity Survey
CVD: cardiovascular disease
OR: odd ratio
PSM: propensity score matching
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by F Puga; submitted 06.11.24; peer-reviewed by C Liu, P Yin; revised version received 18.01.25; accepted 05.03.25; published 21.03.25.

Please cite as:
Wei X, Zhuang L, Li Y, Shi J, Yang Y, Lai H, Liu B
Edentulousness and the Likelihood of Becoming a Centenarian: Longitudinal Observational Study
JMIR Aging 2025;8:e68444
URL: https://aging.jmir.org/2025/1/e68444
doi:10.2196/68444

XSL•FO
RenderX

© Xindi Wei, Longfei Zhuang, Yuan Li, Junyu Shi, Yijie Yang, Hongchang Lai, Beilei Liu. Originally published in JMIR Aging (https://aging.jmir.org), 21.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Association Between Age and Severity at Disability Onset and All-Cause Mortality: Longitudinal Observational Study From the Health and Retirement Study

Anying Bai[1*], PhD; Cuie Liu[2*], MD; Yu Jiang[1,3], PhD; Weihao Xu[4], MD; Jian Cao[2], MD

[1]School of Population Medicine and Public Health, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[2]The 4th Healthcare Department of the Second Medical Center, Chinese PLA General Hospital, National Clinical Research Center for Geriatric Diseases, No. 28 Fuxing Road, Haidian District, Beijing, China

[3]School of Health Policy and Management, Chinese Academy of Medical Science and Peking Union Medical College, Beijing, China

[4]Department of Geriatrics, Guangdong Provincial Geriatrics Institute, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Southern Medical University, Guangzhou, China

[*]these authors contributed equally

Corresponding Author:
Jian Cao, MD
The 4th Healthcare Department of the Second Medical Center, Chinese PLA General Hospital, National Clinical Research Center for Geriatric Diseases, No. 28 Fuxing Road, Haidian District, Beijing, China

Abstract

Background: Disability is a global public health challenge, with its prevalence increasing, particularly among older adults, and it exerts a profound impact on both health outcomes and mortality rates.

Objective: This study investigates the associations between age at disability onset, severity at disability onset, and all-cause mortality in community-dwelling adults.

Methods: We analyzed data from waves 10 to 16 (2010 - 2023) of the Health and Retirement Study, a nationally representative longitudinal survey of US adults aged ≥51 years. Participants without disabilities in activities of daily living (ADLs) or instrumental activities of daily living (IADLs) from the Health and Retirement Study were followed biennially until December 31, 2023. During the follow-up period, 4500 participants developed ADL disability and 4260 developed IADL disability. For each case participant, a control participant matched for age (+1 to −1 y) and sex was randomly selected. Multivariable Cox proportional hazards models were used to assess hazard ratios (HRs) for all-cause mortality among participants with new-onset disabilities, stratified by age groups and severity at disability onset.

Results: Over a median follow-up duration of 8.58 years, 1709 (37.98%) deaths occurred in the ADL group and 1832 (43%) deaths occurred in the IADL group. Individuals who developed ADL disability before the age of 55 years exhibited the highest all-cause mortality risk compared to matched controls (HR 3.12, 95% CI 1.85 - 5.26), which further increased with severe disability (HR 4.07, 95% CI 2.03 - 8.19). The mortality risk was inversely associated with age at onset. A parallel trend was identified in the IADL cohort. Notably, men demonstrated a significantly elevated mortality risk compared to women, emphasizing the need for gender-specific interventions.

Conclusions: Early and severe disability onset significantly increases mortality risk, with men experiencing a disproportionately higher risk. Preventive strategies aimed at addressing early-onset and severe disability, with consideration of gender differences, are essential for improving long-term outcomes in affected populations.

(JMIR Aging 2025;8:e73254) doi:10.2196/73254

KEYWORDS

disability onset; mortality; older population; severity; matched control

Introduction

Background

The World Health Organization (WHO) estimates that approximately 1.3 billion individuals globally experience some form of disability [1]. Between 2016 and 2021, the prevalence of disability among noninstitutionalized adults in the United States increased from 24% to over 25% [2], driven by an aging population and enhanced disclosure of disability status following advances in societal acceptance and legal protections [3]. Findings from the Cognitive Function and Aging Study spanning from 1991 to 2011 indicated an increase in both total life expectancy and disability-free life expectancy; however, the

XSL•FO
RenderX

proportion of life spent without disability has decreased [4]. Population aging in high-income countries is projected to persist, but evidence regarding trends in morbidity and disability prevalence remains inconclusive, with studies in the United States and Europe reporting increases, decreases, or stability in disability prevalence over time [5-8]. These disparities may be attributed to variations in the types of impairments and the prevalence of risk factors, including advanced age, unsafe living conditions, exposure to violence, malnutrition, poverty, and unhealthy behaviors [9,10]. Previous research has consistently reported health disadvantages among individuals who are disabled relative to their counterparts who are not disabled [11-13], as well as an elevated risk of mortality among those with disabilities [14,15].

Existing literature indicates that disability prevalence is significantly higher in older populations and often considered inevitable at the end of life [16,17]. Individuals who are disabled are known to have shorter lifespans compared to the general population [18,19]. Nevertheless, most studies rely on cross-sectional assessments of disability, lacking information on subsequent mortality risk. Longitudinal investigations are limited and tend to focus on incident disability rather than the longitudinal progression of disability after onset, often due to constraints in statistical power [20]. Both age at disability onset and the duration of disability can influence health outcomes [21]. While extensive research has been conducted on disability onset, the mortality risk associated with the duration of disability remains underexplored, and current findings regarding the mortality implications of disability onset age are inconsistent. On the one hand, emerging evidence suggests that early-onset disability is associated with a significantly higher risk of all-cause mortality compared to disability that develops later in life [22]. Individuals who become disabled at younger ages often experience prolonged exposure to functional limitations, comorbidities, and social disadvantages, which can accelerate physiological decline and increase the risk of death [17,23-25]. On the other hand, some studies have proposed that early-onset disability may allow more time for psychological adaptation and the development of coping strategies, which could potentially improve self-perceived health outcomes [26-28]. Despite these contrasting perspectives, most existing research estimates health outcomes based primarily on prevalent disability age, without fully accounting for the duration or severity of disability at onset, which may play a crucial role in shaping long-term survival.

Objective

The objective of this study is to assess the all-cause mortality risk among individuals with new-onset activities of daily living (ADLs) or instrumental activities of daily living (IADLs) disabilities, compared to a control population, stratified by age at onset and disability severity in a representative national cohort.

Methods

Study Participants

This study used data from waves 10 to 16 (2010 - 2023) of the Health and Retirement Study (HRS), a comprehensive longitudinal survey that examines the aging experiences of Americans aged ≥51 years. The HRS uses a multistage probability sampling design to yield a nationally representative sample of US adults in this age group [29]. Attrition rates varied from 10% to 19% across waves (Table S1 in Multimedia Appendix 1). The study collected self-reported data on a range of demographic, health, and functional indicators, including chronic health conditions, daily activities, disability status, and other determinants of health, at baseline and biennially thereafter. In 2006, the HRS expanded its methodology to include enhanced in-person interviews, which incorporated physical performance assessments, biomarker collections, and a psychosocial leave-behind questionnaire. A random half of the participants underwent these enhanced assessments in 2006, while the remaining participants were assessed in 2008, with this process continuing in subsequent waves. Detailed descriptions of the HRS's recruitment methods and design have been published elsewhere [29].

For the ADL disability analysis, we included 28,062 participants from waves 10 to 16. We excluded 10,198 individuals for the following reasons: 2253 (22.09%) who did not attend the examination and already had a disability, 2709 (26.56%) with missing ADL information, and 5236 (51.34%) who were diagnosed with ADL disability at the time of interview. This left 17,864 eligible participants, among whom 5076 (28.41%) developed new-onset ADL disability during a median follow-up of 3.46 years (interquartile range [IQR], 0.54 - 4.88 years). Participants free of ADL disability at baseline were 1:1 matched with those who developed new-onset ADL disability, based on age (±1 y) and sex. After further excluding 1152 participants with missing data on key covariates, the final analytic sample for ADL disability consisted of 9000 individuals (4500 matched case-control pairs).

For the IADL disability analysis, 21,381 participants remained after excluding 6681 individuals for similar reasons (nonattendance with disability, missing IADL information, or IADL disability at interview). Among these 21,381 participants, 4759 (22.26%) developed new-onset IADL disability over a median follow-up of 3.42 years (IQR, 0.54–4.88 years). Following the same age- and sex-matching procedure and excluding 998 participants with missing covariate data, the final IADL analysis included 8520 participants (4260 matched case-control pairs). To minimize bias due to unequal follow-up time or delayed outcome ascertainment, we implemented a matched cohort design. For each case identified as developing disability in a specific year, a control participant of similar age (+1 or –1 y) and sex, who was confirmed to be free of disability in the same wave, was selected. Follow-up for both the case and matched control commenced in that same year of disability onset for the case. For example, if a male participant aged 51 years developed ADL disability in 2010, a control was selected from male participants aged between 49 and 51 years who were disability-free at that time, and both were followed starting in 2010.

Assessment of New-Onset Disability

In this study, disability was operationally defined as the participant's inability to independently perform essential basic

XSL•FO
RenderX

ADLs and IADLs. ADLs were measured using the Katz Index, which evaluates fundamental physical tasks such as dressing, bathing, feeding, transferring from bed to chair, toileting, and maintaining continence [30]. IADLs were assessed based on the Lawton scale, encompassing more complex tasks necessary for independent living, including preparing hot meals, medication management, financial management, grocery shopping, and telephone use [31]. Since wave 10 of the HRS, standardized questionnaires have systematically collected disability data by asking participants whether they experience difficulty performing these activities due to physical, mental, emotional, or memory problems, excluding any difficulties expected to last less than 3 months. Responses were coded dichotomously as *yes* (1) or *no* (0). Composite scores for ADLs and IADLs were calculated by summing individual item responses, resulting in a range from 0 to 5, where higher scores indicated greater difficulty severity. For both ADLs and IADLs, a score of ≥2 indicated severe difficulty, a score of 0 to 2 denoted mild difficulty, and any score >0 signified the presence of challenges in performing ADLs or IADLs. A score of 0 represented no reported difficulties.

Outcomes

Mortality outcomes, including the year and month of death, were ascertained through exit interviews or via core interviews with the spouse or partner of the deceased participant.

Data Collection

Sociodemographic characteristics, lifestyle factors (smoking status and physical activity), and medical history (including stroke, heart disease, and depressive symptoms) were collected via face-to-face interviews using a standardized survey protocol. Data collection occurred biennially, as previously described in the study by Fisher and Ryan [32]. Sociodemographic variables encompassed age (in years), sex (male or female), educational attainment (categorized as primary school or below, high school or equivalent, and college or above), and marital status (married vs unmarried). Current smoking status was determined using the question, “Do you currently smoke cigarettes?” Physical activity levels were classified into tertiles based on the International Physical Activity Questionnaire [33]. Depressive symptoms were assessed using the 8-item Center for Epidemiologic Studies Depression Scale, with the presence of 4 or more symptoms indicative of depression [34]. Heart disease and stroke diagnoses were confirmed by a physician during in-person evaluations using a structured questionnaire administered by the study personnel.

Statistical Analysis

Study participants, comprising case participants with newly developed disabilities and their matched controls, were stratified into 4 onset age groups: <55 years, 55 - 65 years, 65 - 75 years, and ≥75 years. Continuous variables were presented as mean (SD), while categorical variables were expressed as percentages. Baseline characteristics between the case and control groups were compared using the Student *t* test (2-tailed) for continuous variables and the chi-square test for categorical variables. A trend analysis for individual factors across disability onset age groups was conducted by assigning the median value of each age group as a continuous variable within a separate model. General linear models were used for continuous variables, and logistic regression models were applied to categorical variables. The *P* values for trends were derived using the Wald test.

To examine the effect of disability onset on all-cause mortality across age groups, Cox proportional hazards regression analysis was used. The observation period spanned from the index date—the date of incident disability for the case (and the same time point for the matched control)—to either the date of death or the end of the follow-up period. Adjusted hazard ratios (AHRs) with 95% CIs for all-cause mortality were computed for participants with new-onset disabilities relative to their matched controls within each age group. All models were adjusted for sociodemographic variables, lifestyle factors, and relevant medical histories.

All survival analyses accounted for the complex survey design of the HRS, including sample weights, stratification, and clustering. To ensure nationally representative estimates and accurate variance estimation, wave-specific respondent-level weights were applied based on participants’ timing of entry into the study. Analyses were adjusted for the survey’s multistage sampling structure using appropriate design variables, and models were fitted using survey-weighted Cox proportional hazards regression with Taylor series linearization to estimate robust SEs.

Sensitivity analyses were conducted to validate the robustness of the findings. First, we excluded outcome events occurring within the initial 1 or 2 years of follow-up to mitigate potential reverse causation. Second, participants diagnosed with cancer by a physician or receiving active treatment for cancer were excluded to control for the confounding effects of severe cancer on mortality rather than disability. Third, we excluded participants with less than 2 years of follow-up to address potential biases introduced by short observation periods and ensure sufficient time for outcome ascertainment. Fourth, we performed E-value analyses to evaluate the potential impact of unmeasured confounding. Finally, we included participants with missing covariate data by implementing a 2-level multiple imputation by chained equations method under the assumption of missing at random. Statistical significance was defined as a *P* value of <.05. All statistical analyses were performed using Stata (version 17; StatCorp).

Ethical Considerations

The HRS was approved by the institutional review boards of the University of Michigan and the National Institute on Aging. All participants provided written informed consent before data collection. The study procedures were conducted in accordance with the ethical standards of the responsible committees on human experimentation at both institutions, as well as with the principles of the Declaration of Helsinki. Participant privacy and confidentiality were strictly maintained throughout the study.

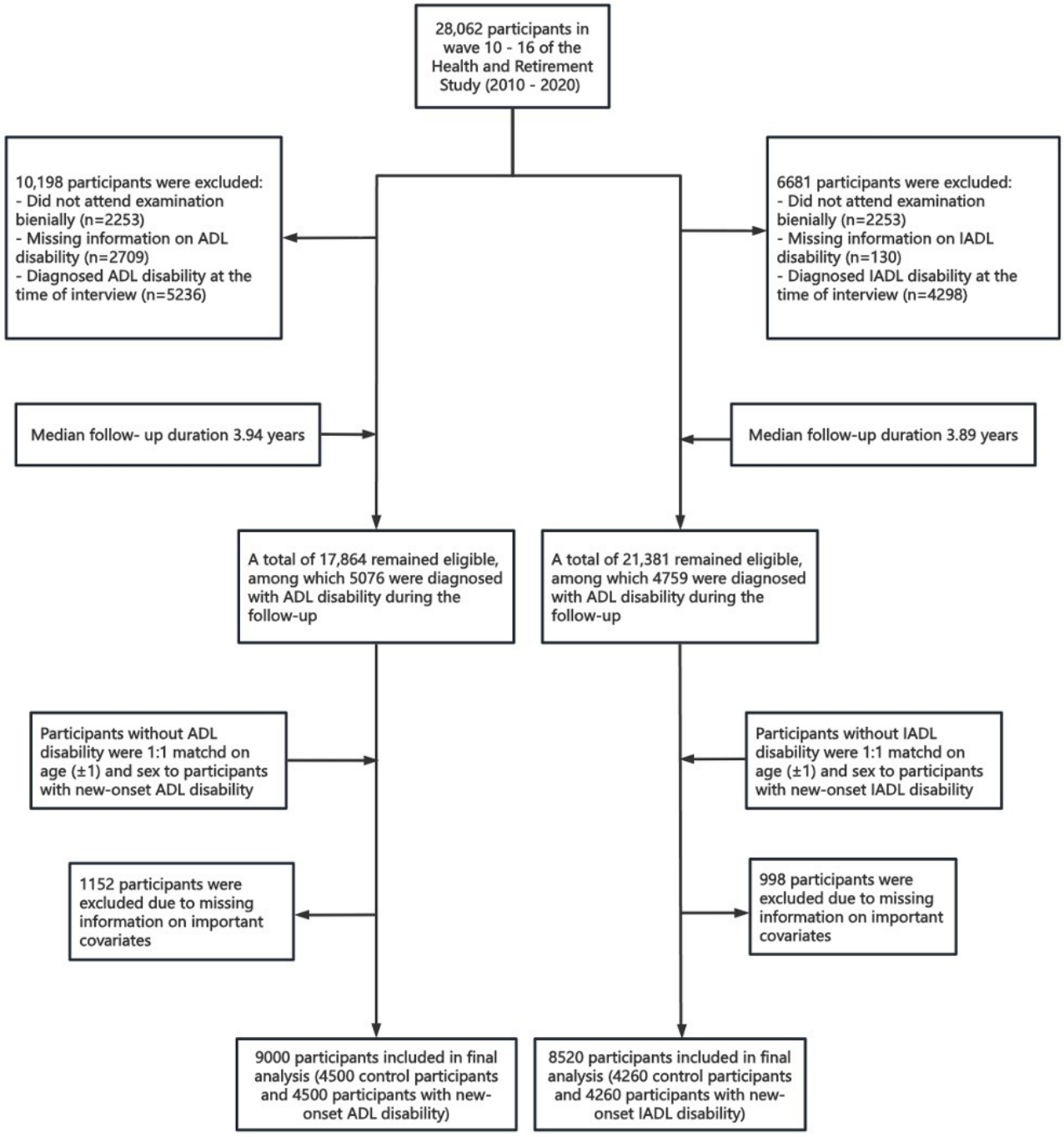
XSL•FO
RenderX

Results

Baseline Characteristics

Detailed participant screening and enrollment procedures are presented in Figure 1. Baseline characteristics of the study population are summarized in Tables 1 and 2. The ADL analysis included a total of 9000 participants, among whom 4500 (50%) developed ADL disability during the follow-up period. Similarly, the IADL analysis comprised 8520 participants, with 4260 (51.64%) experiencing IADL disability during follow-up. The mean age at onset of ADL and IADL disability was 71.08 (SD 11.68) years and 71.74 (SD 12.09) years, respectively. Compared to control participants, individuals with new-onset ADL or IADL disability had lower levels of educational attainment, lower rates of marriage, a higher prevalence of current smoking, and a greater tendency to engage in light physical activity. In addition, a higher prevalence of cardiovascular disease (including heart disease and stroke) and depressive symptoms was observed in participants with new-onset disabilities (Table 1). Among those with younger-onset ADL or IADL disability, the likelihood of being a current smoker, experiencing depressive symptoms, and engaging in vigorous physical activity was higher relative to those with an older onset (Table 1).

Figure 1. Study flowchart of participant selection. ADL: activity of daily living; IADL: instrumental activities of daily living.



XSL•FO
RenderX

Table . Baseline characteristics of participants included in the analysis of new-onset activity of daily living (ADL) disability.

	New-onset ADL disability	Control participants	*P* value	ADL disability onset age (y)				*P* value for trend
				<55	55 - 65	65 - 75	≥75	
Participants, n	4500	4500	—[a]	313	1239	967	1982	—
Age (y), mean (SD)	71.08 (11.68)	71.44 (11.36)	.51	50.74 (4.25)	59.55 (2.77)	69.51 (2.92)	82.26 (4.85)	<.001
Sex (male), n (%)	1788 (39.72)	1788 (39.74)	.99	100 (31.95)	505 (40.76)	388 (40.12)	795 (40.11)	.11
Educational background, n (%)			<.001					.14
Illiterate	1145 (25.44)	694 (15.43)		89 (28.43)	289 (23.33)	256 (26.47)	511 (25.78)	
Primary school or above	2538 (56.39)	2601 (57.81)		179 (57.19)	748 (60.37)	529 (54.71)	1082 (54.59)	
Secondary school or above	818 (18.17)	1204 (26.76)		45 (14.38)	202 (16.3)	182 (18.82)	389 (19.63)	
Married, n (%)	2093 (46.5)	2672 (59.39)	<.001	177 (56.55)	557 (44.96)	520 (53.77)	839 (42.33)	<.001
Current smoker, n (%)	691 (15.35)	447 (9.94)	<.001	112 (35.78)	323 (26.07)	148 (15.31)	108 (5.45)	<.001
Physical exercise, n (%)			<.001					<.001
Light	2178 (48.39)	1058 (23.52)		110 (35.14)	469 (37.85)	424 (43.85)	1175 (59.28)	
Moderate	1336 (29.68)	1510 (33.56)		97 (30.99)	437 (35.27)	309 (31.95)	493 (24.87)	
Vigorous	987 (21.93)	1931 (42.92)		106 (33.87)	333 (26.88)	234 (24.2)	314 (15.84)	
History of heart disease, n (%)	1539 (34.19)	1151(25.58)	<.001	64 (20.45)	270 (21.79)	339 (35.06)	866 (43.69)	<.001
History of stroke, n (%)	645 (14.33)	276 (6.13)	<.001	27 (8.63)	120 (9.69)	150 (15.51)	348 (17.56)	<.001
Depressive symptoms, n (%)	657 (14.6)	237 (5.27)	<.001	87 (27.8)	303 (24.46)	128 (13.24)	139 (7.01)	<.001

[a]Not available.

XSL•FO
RenderX

Table . Baseline characteristics of participants included in analysis of new-onset instrumental activity of daily living (IADL) disability.

	New-onset IADL disability	Control participants	*P* value	IADL disability onset age (y)				*P* value for trends
				<55	55-65	65-75	≥75	
Participants, n	4260	4260	—[a]	876	962	1036	1377	—
Age (y), mean (SD)	71.74 (12.09)	72.23 (11.75)	.03	54.92 (4.21)	64.10 (2.88)	75.03 (2.82)	85.30 (3.94)	<.001
Sex (male), n (%)	1724 (40.56)	1712 (40.1)	.67	330 (37.67)	398 (41.37)	427 (41.22)	569 (41.32)	.13
Educational background, n (%)			001					.06
Illiterate	1165 (27.41)	630 (14.76)		227 (25.91)	282 (29.31)	296 (28.57)	360 (26.14)	
Primary school or above	2377 (55.92)	2393 (56.06)		512 (58.45)	533 (55.41)	590 (56.95)	742 (53.89)	
Secondary school or above	709 (16.68)	1246 (29.29)		137 (15.64)	147 (15.28)	150 (14.48)	275 (19.97)	
Married, n (%)	1918 (45.12)	2502 (58.61)	<.001	396 (45.21)	459 (47.71)	538 (51.93)	525 (38.13)	<.001
Current smoker, n (%)	670 (15.76)	371 (8.69)	<.001	292 (33.33)	204 (21.21)	124 (11.97)	50 (3.63)	<.001
Physical exercise, n (%)			<.001					<.001
Light	2185 (51.4)	1066 (24.97)		332 (37.90)	448 (46.57)	556 (53.67)	849 (61.66)	
Moderate	1264 (29.73)	1322 (30.97)		293 (33.45)	315 (32.74)	289 (27.9)	367 (26.65)	
Vigorous	802 (18.87)	1881 (44.06)		251 (28.65)	199 (20.69)	191 (18.44)	161 (11.69)	
History of heart disease, n (%)	1523 (35.83)	1110 (26)	<.001	159 (18.15)	315 (32.74)	435 (41.99)	614 (44.59)	<.001
History of stroke, n (%)	668 (15.71)	273 (6.39)	<.001	78 (8.90)	130 (13.51)	200 (19.31)	260 (18.88)	<.001
Depressive symptoms, n (%)	651 (15.31)	193 (4.52)	<.001	259 (29.57)	186 (19.33)	114 (11)	92 (6.68)	<.001

[a]Not available.

Associations Between Disability Onset Age and the Risk of Mortality

Moreover, during the follow-up period, 25.79% (1161/4500) of participants with new-onset ADL disability and 28.35% (1208/4260) of those with new-onset IADL disability died, compared with 18.2% (819/4500) and 20.58% (877/4260) of their corresponding controls. After adjusting for key sociodemographic variables (age, sex, educational level, and marital status), lifestyle factors (smoking status and physical activity), and medical history (stroke, heart disease, and depressive symptoms), individuals with new-onset ADL or IADL disability were found to have a significantly increased risk of all-cause mortality compared to their matched controls across all age groups. Notably, participants with IADL disability onset between the ages of 55 and 65 years demonstrated the highest relative risk of all-cause mortality, with an AHR of 4.12 (95% CI 2.99 - 5.68). In addition, this elevated risk was more pronounced among male participants (AHR 4.52, 95% CI 2.82 - 7.17) compared to female participants (AHR 3.61, 95% CI 2.33 - 5.60). The incremental risks gradually decreased with each 10-year increase of onset age. Similarly, participants with ADL disability onset between ages 55 and 65 years exhibited the highest relative risks, with an AHR of 5.38 (95% CI 3.37 - 8.58) for severe IADL disability onset between the ages of 55 and 65 years. The relative risk of mortality also progressively declined with increasing age at disability onset (Figure 2), and male participants generally demonstrated higher risks than female participants across most age groups (Figure 3). We also graphically illustrated the incidence rates of all-cause mortality in different age groups in Figure S1 in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 2. Adjusted hazard ratios (HRs) with 95% CIs for all-cause mortality in patients with new-onset activity of daily living (ADL) disability compared to control participants, stratified by age group and disability severity at onset. All models were adjusted for sociodemographic characteristics (age, sex, educational attainment, and marital status); lifestyle factors (smoking status and physical activity); and medical history (including stroke, heart disease, and depressive symptoms). IADL: instrumental activities of daily living.

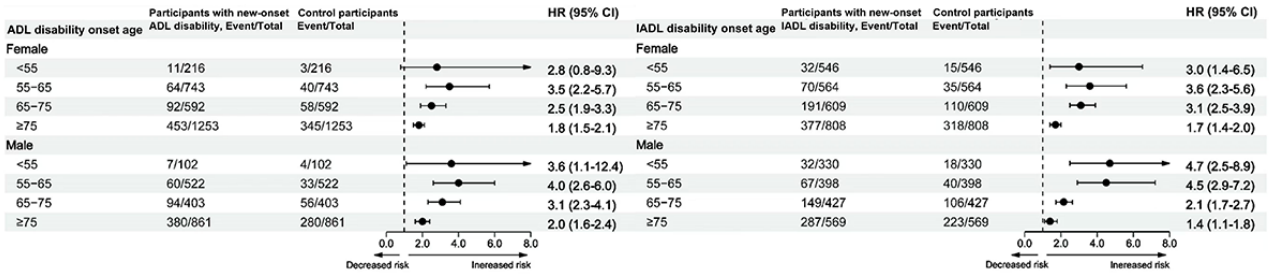
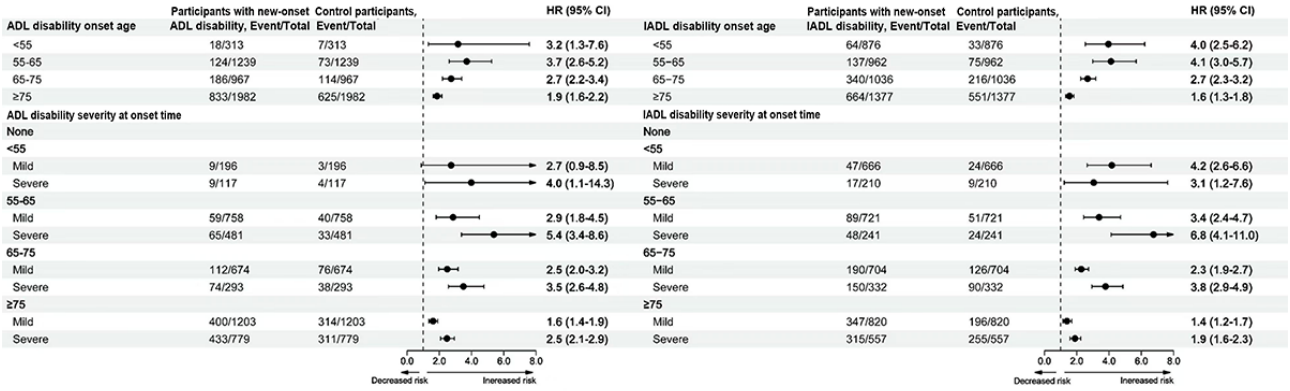


Figure 3. Adjusted hazard ratios (HRs) with 95% CIs for all-cause mortality in patients with new-onset activity of daily living (ADL) disability compared to controls, stratified by age group and gender. All models were adjusted for sociodemographic characteristics (age, sex, educational attainment, and marital status); lifestyle factors (smoking status and physical activity); and medical history (including stroke, heart disease, and depressive symptoms). IADL: instrumental activities of daily living.



Associations Between Disability Onset Severity and Risk of Mortality

Compared to participants with new-onset mild disability, those with new-onset severe ADL or IADL disability consistently exhibited higher relative risks across all age groups (Figure 2) and in the overall study population (Table S2 in Multimedia Appendix 1). For each onset age group, the AHRs for severe disability onset were consistently higher than for mild disability onset. However, the AHR for severe IADL disability onset below the age of 55 years was 3.05 (95% CI 1.22 - 7.62), which was lower than that for mild IADL disability onset (AHR 4.17, 95% CI 2.64 - 6.6). The incidence rates of all-cause mortality across various age groups, stratified by the severity of disability onset, are illustrated in Figure S2 in Multimedia Appendix 1.

Sensitivity Analyses

Sensitivity analyses demonstrated that the association between new-onset disability and the risk of all-cause mortality remained robust across all age groups. This consistency persisted even after excluding participants who experienced outcome events within the first or second year of follow-up (Tables S3-S5 in Multimedia Appendix 1). Furthermore, the results remained unchanged upon excluding individuals with a history of cancer or those currently receiving treatment (Tables S3-S5 in Multimedia Appendix 1).

Discussion

Principal Findings

In this study, we observed that female participants diagnosed with ADL disability before the age of 55 years exhibited the highest relative risks for all-cause mortality when compared to age- and sex-matched controls. Participants with severe ADL or IADL disability at the time of diagnosis demonstrated greater mortality risks across all age groups compared to those with mild disability. Notably, the relative risk of mortality declined as the age of disability onset increased.

Previous literature has consistently identified disability as an independent predictor of mortality, even after adjusting for confounding factors such as cardiovascular disease [35], depressive symptoms [36], physical activity [37], and socioeconomic status [38]. The association between mortality and disability appears to be particularly pronounced for individuals with severe disabilities [39]. Infurna and Wiest [40] reported that the severity of disability plays a critical role in determining life satisfaction after onset, potentially due to the maintenance of an active lifestyle (eg, engagement in hobbies) or increased dependency on others, leading to a loss of independence. However, these studies primarily focused on participants with prevalent disabilities and did not initiate follow-up at the time of disability diagnosis. Although the precise mechanisms underlying the elevated mortality risk

XSL•FO
RenderX

among individuals with disabilities remain unclear, our findings are biologically plausible. The relationship between disability and increased mortality risk in older adults is multifactorial, with age-related frailty being a key factor influencing both physical disability and mortality [41-43]. Decreases in muscle strength [42], chronic inflammation [44], neurological disorders [45], and chronic conditions [46] are interrelated factors that contribute to disability and concurrently accelerate mortality risk [42]. Multimorbidity is also closely associated with increased risks of both disability and mortality in older populations [47]. In addition, physical frailty and loss of independence often complicate the management of chronic conditions, leading to worse health outcomes. [48]

Our findings underscore the importance of early preventive interventions for individuals who experience disability onset before the age of 65 years, particularly those with severe functional impairments. Although disability onset has been extensively examined, research on how age at onset and severity influence mortality risk remains limited. Similar to our results, a population-based cohort study with over 10 years of follow-up [49] found that younger individuals with ADL impairments experience a significant reduction in survival. Only 2 previous studies have explicitly considered age as a timescale when assessing mortality risk [39,50]. Lamarca et al [50] demonstrated that earlier onset of ADL disability was associated with a higher mortality risk compared to older adults, with women less likely to recover from disability compared to men. Majer et al [39] similarly found that the mortality risk associated with disability decreased with increasing age, suggesting that younger individuals are more adversely affected by disability in terms of mortality risk. Nonetheless, some studies have suggested that younger onset of disability has a less detrimental effect on mortality in older populations, a finding that contrasts with our results [40,51,52]. These discrepancies may be attributable to differences in age stratification, disability classification [53,54], and the duration of postdiagnosis care [55] across studies.

The early identification of disability is of critical importance, as it has been consistently linked to an elevated risk of mortality. Individuals who experience disability onset at younger ages, such as those aged <55 years or aged between 55 and 65 years, demonstrate a heightened risk of mortality compared to those with onset between 65 and 75 years or those aged >75 years. This disparity in risk can be attributed to multiple factors. First, younger individuals with disabilities typically endure prolonged durations of illness and its associated complications, contributing to increased mortality rates [56]. In addition, the earlier onset of disability may signal the presence of more aggressive underlying pathological processes, thereby influencing adverse outcomes. Furthermore, disability at a younger age may significantly affect quality of life and psychosocial well-being, potentially resulting in detrimental health behaviors or insufficient social support, both of which compound health risks. Our findings also indicate that among individuals of the same gender, those aged <65 years with IADL disability exhibit a higher mortality risk than those with ADL impairment. In contrast, in populations aged ≥65 years, the pattern reverses: individuals with ADL disability face a greater mortality risk than those with IADL limitations. This crossover effect may reflect the different implications of functional limitations at various life stages. In middle-aged adults, IADL disability could be an early marker of accelerated biological aging or cognitive decline, often preceding ADL limitations and signaling broader systemic deterioration [57,58]. These individuals may still appear physically capable but experience subtle impairments in executive functioning or complex task management, which have been linked to excess mortality [59]. However, in older adults, ADL impairment reflects more severe physical frailty and dependency, often closely associated with terminal decline and higher short-term mortality risk. [60,61]

Our study further elucidates gender-specific differences in the association between disability onset and all-cause mortality. Specifically, we observed stronger mortality risks among men than women, corroborating existing literature [39,62-64] and aligning with the established gender paradox in morbidity and mortality [65,66]. Women tend to experience more frequent and prolonged periods of disability, which may contribute to their lower recovery and mortality rates relative to men [67,68]. Conversely, men are more susceptible to fatal conditions such as cardiovascular disease and stroke, which elevates their mortality risk [69]. This observed gender disparity in disability and mortality may not be completely explained by gender-specific health conditions, body composition, or the operational definitions of disability [70]. Therefore, additional investigation into alternative explanatory pathways is warranted.

This study offers several notable strengths. It is the first large-scale prospective study to examine the impact of disability onset age on all-cause mortality, with participants tracked from the time of diagnosis. Moreover, the inclusion of age- and sex-matched controls without disability allows for a comprehensive consideration of potential confounders. We also accounted for the severity of disability at onset, distinguishing between mild and severe cases, thereby capturing the influence of both duration and severity on mortality outcomes. Furthermore, repeated clinical examinations provided access to a wide range of covariates, enabling the minimization of confounding bias.

Limitations

Despite the strengths, several limitations must be acknowledged. First, the onset age of disability was approximated based on the date of initial investigation minus the participant's birthdate, which may differ slightly from the precise diagnosis age. Nevertheless, given that participants were interviewed biennially, such discrepancies are unlikely to substantially impact the findings. Second, although adjustments were made for a range of factors, residual confounding cannot be entirely ruled out, especially from unmeasured factors such as nutritional status or exposure to violence. However, our E-value analysis provides quantitative reassurance regarding the validity of our findings (Table S6 in Multimedia Appendix 1). Third, disability status and severity were based on self-report, which may be subject to recall or interpretation biases. Notably, the HRS explicitly instructs respondents to exclude any difficulties expected to last fewer than 3 months when reporting activity limitations. This reduces the risk of misclassifying temporary impairments, such as those due to acute illness or injury, as

XSL•FO
RenderX

new-onset disability. Moreover, while self-reported difficulty does not necessarily imply complete dependence, it has been shown to be a meaningful indicator of early functional decline in aging research. Fourth, because disability is a dynamic process with potential for recovery, our approach—which identifies incident disability at the first report of difficulty—may include cases that later resolve. However, requiring confirmation across multiple waves would risk underestimating meaningful early disability and may introduce survivor bias by excluding individuals who recover or die before a second measurement. Moreover, our use of a standardized and widely accepted definition of ADL or IADL disability, along with multiple sensitivity analyses, supports the robustness of our findings. Finally, we acknowledge the potential for survivor bias, as individuals who develop disability at older ages must have survived without severe illness to that point. However, this bias is mitigated by our use of age-at-onset stratification and incident (rather than prevalent) disability, which allows for fairer comparisons within age groups and reduces the overrepresentation of unusually resilient individuals.

Conclusions

This study demonstrates that the relative risks of all-cause mortality vary significantly depending on the age at which disability is diagnosed, with a notably higher risk observed among individuals diagnosed at a younger age. The identification and precise quantification of the elevated mortality risk associated with early-onset disability provide critical insights for risk stratification at earlier stages of the condition. Such stratification may inform the development of targeted interventions, including intensive therapeutic approaches and tailored lifestyle modifications, aimed at mitigating the risk of complication-related morbidity and mortality in this growing population of individuals with early-onset disability.

Acknowledgments

This study was supported by the Natural Science Foundation of Guangdong Province of China (2023A1515011259).

Data Availability

The datasets analyzed during this study are available from the Health and Retirement Study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Detailed incidence rates, attrition rates, main and sensitivity analyses, and robustness checks of the associations between disability onset, severity, and all-cause mortality among older adults in the Health and Retirement Study.
[DOCX File, 598 KB - aging_v8i1e73254_app1.docx]

References

1. Global report on health equity for persons with disabilities. World Health Organization. 2022 Dec 2. URL: https://www.who.int/publications/i/item/9789240063600 [accessed 2025-09-23]
2. Okoro CA, Hollis ND, Cyrus AC, Griffin-Blake S. Prevalence of disabilities and health care access by disability status and type among adults - United States, 2016. MMWR Morb Mortal Wkly Rep 2018 Aug 17;67(32):882-887. [doi: 10.15585/mmwr.mm6732a3] [Medline: 30114005]
3. Varadaraj V, Deal JA, Campanile J, Reed NS, Swenor BK. National prevalence of disability and disability types among adults in the US, 2019. JAMA Netw Open 2021 Oct 1;4(10):e2130358. [doi: 10.1001/jamanetworkopen.2021.30358] [Medline: 34673966]
4. Jagger C, Matthews FE, Wohland P, et al. A comparison of health expectancies over two decades in England: results of the Cognitive Function and Ageing Study I and II. Lancet 2016 Feb 20;387(10020):779-786. [doi: 10.1016/S0140-6736(15)00947-2] [Medline: 26680218]
5. Moe JO, Hagen TP. Trends and variation in mild disability and functional limitations among older adults in Norway, 1986-2008. Eur J Ageing 2011 Mar;8(1):49-61. [doi: 10.1007/s10433-011-0179-3] [Medline: 21475398]
6. Freedman VA, Spillman BC, Andreski PM, et al. Trends in late-life activity limitations in the United States: an update from five national surveys. Demography 2013 Apr;50(2):661-671. [doi: 10.1007/s13524-012-0167-z] [Medline: 23104207]
7. Lin SF, Beck AN, Finch BK, Hummer RA, Masters RK. Trends in US older adult disability: exploring age, period, and cohort effects. Am J Public Health 2012 Nov;102(11):2157-2163. [doi: 10.2105/AJPH.2011.300602] [Medline: 22994192]
8. van Gool CH, Picavet HS, Deeg DJ, et al. Trends in activity limitations: the Dutch older population between 1990 and 2007. Int J Epidemiol 2011 Aug;40(4):1056-1067. [doi: 10.1093/ije/dyr009] [Medline: 21324941]
9. Banks LM, Kuper H, Polack S. Poverty and disability in low- and middle-income countries: a systematic review. PLoS ONE 2017 Dec 21;12(12):e0189996. [doi: 10.1371/journal.pone.0189996] [Medline: 29267388]
10. Marmot M. Social determinants of health inequalities. Lancet 2005 Mar;365(9464):1099-1104. [doi: 10.1016/S0140-6736(05)71146-6] [Medline: 15781105]

XSL•FO
RenderX

11. Bahk J, Kang HY, Khang YH. The life expectancy gap between registered disabled and non-disabled people in Korea from 2004 to 2017. Int J Environ Res Public Health 2019 Jul 20;16(14):2593. [doi: 10.3390/ijerph16142593] [Medline: 31330839]
12. Dammeyer J, Chapman M. A national survey on violence and discrimination among people with disabilities. BMC Public Health 2018 Mar 15;18(1):355. [doi: 10.1186/s12889-018-5277-0] [Medline: 29544470]
13. Glover G, Williams R, Heslop P, Oyinlola J, Grey J. Mortality in people with intellectual disabilities in England. J Intellect Disabil Res 2017 Jan;61(1):62-74. [doi: 10.1111/jir.12314] [Medline: 27489075]
14. Plana-Ripoll O, Pedersen CB, Agerbo E, et al. A comprehensive analysis of mortality-related health metrics associated with mental disorders: a nationwide, register-based cohort study. Lancet 2019 Nov 16;394(10211):1827-1835. [doi: 10.1016/S0140-6736(19)32316-5] [Medline: 31668728]
15. Warren MD, Knight R. Mortality in relation to the functional capacities of people with disabilities living at home. J Epidemiol Community Health 1982 Sep;36(3):220-223. [doi: 10.1136/jech.36.3.220] [Medline: 6216299]
16. Smith AK, Walter LC, Miao Y, Boscardin WJ, Covinsky KE. Disability during the last two years of life. JAMA Intern Med 2013 Sep 9;173(16):1506-1513. [doi: 10.1001/jamainternmed.2013.8738] [Medline: 23836120]
17. Gill TM, Gahbauer EA, Han L, Allore HG. Trajectories of disability in the last year of life. N Engl J Med 2010 Apr 1;362(13):1173-1180. [doi: 10.1056/NEJMoa0909087] [Medline: 20357280]
18. Iburg KM, Charalampous P, Allebeck P, et al. Burden of disease among older adults in Europe-trends in mortality and disability, 1990-2019. Eur J Public Health 2023 Feb 3;33(1):121-126. [doi: 10.1093/eurpub/ckac160] [Medline: 36421036]
19. Kuper H, Rotenberg S, Azizatunnisa' L, Banks LM, Smythe T. The association between disability and mortality: a mixed-methods study. Lancet Public Health 2024 May;9(5):e306-e315. [doi: 10.1016/S2468-2667(24)00054-9] [Medline: 38702095]
20. Nusselder WJ, Peeters A. Successful aging: measuring the years lived with functional loss. J Epidemiol Community Health 2006 May;60(5):448-455. [doi: 10.1136/jech.2005.041558] [Medline: 16614337]
21. Boström K, Ahlström G. Quality of life in patients with muscular dystrophy and their next of kin. Int J Rehabil Res 2005 Jun;28(2):103-109. [doi: 10.1097/00004356-200506000-00002] [Medline: 15900179]
22. Reynolds SL, Silverstein M. Observing the onset of disability in older adults. Soc Sci Med 2003 Nov;57(10):1875-1889. [doi: 10.1016/s0277-9536(03)00053-4] [Medline: 14499512]
23. Stolz E, Gill TM, Mayerl H, Rásky É, Freidl W. Trajectories of late-life disability vary by the condition leading to death. J Gerontol A Biol Sci Med Sci 2021 Jun 14;76(7):1260-1264. [doi: 10.1093/gerona/glaa234] [Medline: 32939547]
24. Seeman TE, Charpentier PA, Berkman LF, et al. Predicting changes in physical performance in a high-functioning elderly cohort: MacArthur studies of successful aging. J Gerontol 1994 May;49(3):M97-M108. [doi: 10.1093/geronj/49.3.m97] [Medline: 8169338]
25. Freedman VA, Martin LG, Schoeni RF, Cornman JC. Declines in late-life disability: the role of early- and mid-life factors. Soc Sci Med 2008 Apr;66(7):1588-1602. [doi: 10.1016/j.socscimed.2007.11.037] [Medline: 18222580]
26. Bishop M. Quality of life and psychosocial adaptation to chronic illness and disability: preliminary analysis of a conceptual and theoretical synthesis. Rehabil Couns Bull 2005;48(4):219-231. [doi: 10.1177/00343552050480040301]
27. Livneh H. Psychosocial adaptation to chronic illness and disability: a conceptual framework. Rehabil Couns Bull 2001 Apr;44(3):151-160. [doi: 10.1177/003435520104400305]
28. Finch J, Robinson M. Aging and late-onset disability: addressing workplace accommodation. (Aging and late-onset disability). J Rehabil 2003;69(2):38-43 [FREE Full text]
29. Heeringa SG, Connor JH. Technical description of the health and retirement survey sample design. University of Michigan. 1995 May. URL: https://hrsonline.isr.umich.edu/sitedocs/userg/HRSSAMP.pdf [accessed 2025-09-24]
30. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged. The index of ADL: a standardized measure of biological and psychosocial function. JAMA 1963 Sep 21;185:914-919. [doi: 10.1001/jama.1963.03060120024016] [Medline: 14044222]
31. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179-186. [doi: 10.1093/geront/9.3_Part_1.179] [Medline: 5349366]
32. Fisher GG, Ryan LH. Overview of the health and retirement study and introduction to the special issue. Work Aging Retire 2018 Jan;4(1):1-9. [doi: 10.1093/workar/wax032] [Medline: 29423243]
33. Hagströmer M, Oja P, Sjöström M. The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. Public Health Nutr 2006 Sep;9(6):755-762. [doi: 10.1079/phn2005898] [Medline: 16925881]
34. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977;1(3):385-401. [doi: 10.1177/014662167700100306]
35. Kattainen A, Reunanen A, Koskinen S, Martelin T, Knekt P, Aromaa A. Disability predicted mortality in men but not women with coronary heart disease. J Clin Epidemiol 2004 May;57(5):513-521. [doi: 10.1016/j.jclinepi.2003.10.004] [Medline: 15196622]
36. van den Brink CL, Tijhuis M, van den Bos GA, Giampaoli S, Nissinen A, Kromhout D. The contribution of self-rated health and depressive symptoms to disability severity as a predictor of 10-year mortality in European elderly men. Am J Public Health 2005 Nov;95(11):2029-2034. [doi: 10.2105/AJPH.2004.050914] [Medline: 16195527]

XSL•FO
RenderX

37. Tager IB, Haight TJ, Hollenberg M, Satariano WA. Physical functioning and mortality in older women: an assessment of energy costs and level of difficulty. J Clin Epidemiol 2003 Aug;56(8):807-813. [doi: 10.1016/s0895-4356(03)00149-5] [Medline: 12954475]
38. Fried LP, Kronmal RA, Newman AB, et al. Risk factors for 5-year mortality in older adults: the Cardiovascular Health Study. JAMA 1998 Feb 25;279(8):585-592. [doi: 10.1001/jama.279.8.585] [Medline: 9486752]
39. Majer IM, Nusselder WJ, Mackenbach JP, Klijs B, van Baal PH. Mortality risk associated with disability: a population-based record linkage study. Am J Public Health 2011 Dec;101(12):e9-e15. [doi: 10.2105/AJPH.2011.300361] [Medline: 22021307]
40. Infurna FJ, Wiest M. The effect of disability onset across the adult life span. J Gerontol B Psychol Sci Soc Sci 2018 Jun 14;73(5):755-766. [doi: 10.1093/geronb/gbw055] [Medline: 27166123]
41. Wu LW, Chen WL, Peng TC, et al. All-cause mortality risk in elderly individuals with disabilities: a retrospective observational study. BMJ Open 2016 Sep 13;6(9):e011164. [doi: 10.1136/bmjopen-2016-011164] [Medline: 27625055]
42. Lavie CJ, Ozemek C, Carbone S, Katzmarzyk PT, Blair SN. Sedentary behavior, exercise, and cardiovascular health. Circ Res 2019 Mar;124(5):799-815. [doi: 10.1161/CIRCRESAHA.118.312669] [Medline: 30817262]
43. Afilalo J, Alexander KP, Mack MJ, et al. Frailty assessment in the cardiovascular care of older adults. J Am Coll Cardiol 2014 Mar 4;63(8):747-762. [doi: 10.1016/j.jacc.2013.09.070] [Medline: 24291279]
44. Friedman E, Shorey C. Inflammation in multimorbidity and disability: an integrative review. Health Psychol 2019 Sep;38(9):791-801. [doi: 10.1037/hea0000749] [Medline: 31436464]
45. Mat Rosly M, Mat Rosly H, Davis Oam GM, Husain R, Hasnan N. Exergaming for individuals with neurological disability: a systematic review. Disabil Rehabil 2017 Apr;39(8):727-735. [doi: 10.3109/09638288.2016.1161086] [Medline: 27108475]
46. GBD 2017 Risk Factor Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1923-1994. [doi: 10.1016/S0140-6736(18)32225-6] [Medline: 30496105]
47. Dugravot A, Fayosse A, Dumurgier J, et al. Social inequalities in multimorbidity, frailty, disability, and transitions to mortality: a 24-year follow-up of the Whitehall II cohort study. Lancet Public Health 2020 Jan;5(1):e42-e50. [doi: 10.1016/S2468-2667(19)30226-9] [Medline: 31837974]
48. Schoeller SD, Lima DK, Martins MM, et al. Protocol for a scoping review on nursing care and the autonomy of disabled persons. BMJ Open 2018 Oct 8;8(10):e022106. [doi: 10.1136/bmjopen-2018-022106] [Medline: 30297347]
49. Delva F, Touraine C, Joly P, et al. ADL disability and death in dementia in a French population‐based cohort: new insights with an illness‐death model. Alzheimers Dementia 2016 Aug;12(8):909-916. [doi: 10.1016/j.jalz.2016.03.007] [Medline: 27103260]
50. Lamarca R, Ferrer M, Andersen PK, Liestol K, Keiding N, Alonso J. A changing relationship between disability and survival in the elderly population: differences by age. J Clin Epidemiol 2003 Dec;56(12):1192-1201. [doi: 10.1016/s0895-4356(03)00201-4] [Medline: 14680670]
51. Jamoom EW, Horner-Johnson W, Suzuki R, Andresen EM, Campbell VA, RRTC Expert Panel on Health Status Measurement. Age at disability onset and self-reported health status. BMC Public Health 2008 Jan 9;8:10. [doi: 10.1186/1471-2458-8-10] [Medline: 18184437]
52. Fraire N, Infurna FJ. The effect of disability on subjective well-being across adulthood: the moderating role of age. Innov Aging 2017 Jun 30;1(suppl_1):1144. [doi: 10.1093/geroni/igx004.4179]
53. Picavet HS, van den Bos GA. Comparing survey data on functional disability: the impact of some methodological differences. J Epidemiol Community Health 1996 Feb;50(1):86-93. [doi: 10.1136/jech.50.1.86] [Medline: 8762361]
54. Klijs B, Mackenbach JP, Kunst AE. Disability occurrence and proximity to death. Disabil Rehabil 2010;32(21):1733-1741. [doi: 10.3109/09638281003746049] [Medline: 20373858]
55. Weaver F, Stearns SC, Norton EC, Spector W. Proximity to death and participation in the long-term care market. Health Econ 2009 Aug;18(8):867-883. [doi: 10.1002/hec.1409] [Medline: 18770873]
56. Smythe T, Kuper H. The association between disability and all-cause mortality in low-income and middle-income countries: a systematic review and meta-analysis. Lancet Glob Health 2024 May;12(5):e756-e770. [doi: 10.1016/S2214-109X(24)00042-1] [Medline: 38614629]
57. Stuck AE, Walthert JM, Nikolaus T, Büla CJ, Hohmann C, Beck JC. Risk factors for functional status decline in community-living elderly people: a systematic literature review. Soc Sci Med 1999 Feb;48(4):445-469. [doi: 10.1016/s0277-9536(98)00370-0] [Medline: 10075171]
58. Njegovan V, Hing MM, Mitchell SL, Molnar FJ. The hierarchy of functional loss associated with cognitive decline in older persons. J Gerontol A Biol Sci Med Sci 2001 Oct;56(10):M638-M643. [doi: 10.1093/gerona/56.10.m638] [Medline: 11584037]
59. Johnson JK, Lui LY, Yaffe K. Executive function, more than global cognition, predicts functional decline and mortality in elderly women. J Gerontol A Biol Sci Med Sci 2007 Oct;62(10):1134-1141. [doi: 10.1093/gerona/62.10.1134] [Medline: 17921427]
60. Verbrugge LM, Jette AM. The disablement process. Soc Sci Med 1994 Jan;38(1):1-14. [doi: 10.1016/0277-9536(94)90294-1] [Medline: 8146699]

XSL•FO
RenderX

61. Yang Y, Du Z, Liu Y, Lao J, Sun X, Tang F. Disability and the risk of subsequent mortality in elderly: a 12-year longitudinal population-based study. BMC Geriatr 2021 Nov 23;21(1):662. [doi: 10.1186/s12877-021-02611-1] [Medline: 34814844]
62. Doblhammer G, Hoffmann R. Gender differences in trajectories of health limitations and subsequent mortality. A study based on the German Socioeconomic Panel 1995-2001 with a mortality follow-up 2002-2005. J Gerontol B Psychol Sci Soc Sci 2010 Jul;65(4):482-491. [doi: 10.1093/geronb/gbp051] [Medline: 19564209]
63. Drumond Andrade FC, Guevara PE, Lebrão ML, de Oliveira Duarte YA, Santos JL. Gender differences in life expectancy and disability-free life expectancy among older adults in São Paulo, Brazil. Womens Health Issues 2011;21(1):64-70. [doi: 10.1016/j.whi.2010.08.007] [Medline: 21185991]
64. Tiainen K, Luukkaala T, Hervonen A, Jylhä M. Predictors of mortality in men and women aged 90 and older: a nine-year follow-up study in the Vitality 90+ study. Age Ageing 2013 Jul;42(4):468-475. [doi: 10.1093/ageing/aft030] [Medline: 23486170]
65. Nathanson CA. Illness and the feminine role: a theoretical review. Soc Sci Med (1967) 1975 Feb;9(2):57-62. [doi: 10.1016/0037-7856(75)90094-3] [Medline: 1093257]
66. Verbrugge LM. Sex differentials in health. Public Health Rep 1982;97(5):417-437. [Medline: 6750677]
67. Dunlop DD, Hughes SL, Manheim LM. Disability in activities of daily living: patterns of change and a hierarchy of disability. Am J Public Health 1997 Mar;87(3):378-383. [doi: 10.2105/ajph.87.3.378] [Medline: 9096537]
68. Hardy SE, Allore HG, Guo Z, Gill TM. Explaining the effect of gender on functional transitions in older persons. Gerontology 2008;54(2):79-86. [doi: 10.1159/000115004] [Medline: 18230952]
69. Antonogeorgos G, Panagiotakos DB, Pitsavos C, et al. Understanding the role of depression and anxiety on cardiovascular disease risk, using structural equation modeling; the mediating effect of the Mediterranean diet and physical activity: the ATTICA study. Ann Epidemiol 2012 Sep;22(9):630-637. [doi: 10.1016/j.annepidem.2012.06.103] [Medline: 22831995]
70. Pongiglione B, De Stavola BL, Kuper H, Ploubidis GB. Disability and all-cause mortality in the older population: evidence from the English Longitudinal Study of Ageing. Eur J Epidemiol 2016 Aug;31(8):735-746. [doi: 10.1007/s10654-016-0160-8] [Medline: 27177908]

Abbreviations

ADL: activity of daily living
AHR: adjusted hazard ratio
HR: hazard ratio
HRS: Health and Retirement Study
IADL: instrumental activity of daily living
WHO: World Health Organization

Edited by Y Du; submitted 28.02.25; peer-reviewed by KT Pritchard, S Fernandes; revised version received 20.07.25; accepted 13.09.25; published 03.10.25.

Please cite as:
Bai A, Liu C, Jiang Y, Xu W, Cao J
Association Between Age and Severity at Disability Onset and All-Cause Mortality: Longitudinal Observational Study From the Health and Retirement Study
JMIR Aging 2025;8:e73254
URL: https://aging.jmir.org/2025/1/e73254
doi:10.2196/73254

© Anying Bai, Cuie Liu, Yu Jiang, Weihao Xu, Jian Cao. Originally published in JMIR Aging (https://aging.jmir.org), 3.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Artificial Intelligence-Driven Biological Age Prediction Model Using Comprehensive Health Checkup Data: Development and Validation Study

Chang-Uk Jeong[1*], BS; Jacob S Leiby[1*], PhD; Dokyoon Kim[1,2], PhD; Eun Kyung Choe[3,4], MD, PhD

[1]Department of Biostatistics, Epidemiology & Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

[2]Institute for Biomedical Informatics, University of Pennsylvania, Philadelphia, PA, USA

[3]Department of Surgery, Seoul National University Hospital Healthcare System Gangnam Center, 38-40FL Gangnam Finance Center(prior Star Tower) 152, Teheran-ro, Gangnam-gu, Seoul, Republic of Korea

[4]Department of Surgery, College of Medicine, Seoul National University, Seoul, South Korea

[*]these authors contributed equally

Corresponding Author:
Eun Kyung Choe, MD, PhD
Department of Surgery, Seoul National University Hospital Healthcare System Gangnam Center, 38-40FL Gangnam Finance Center(prior Star Tower) 152, Teheran-ro, Gangnam-gu, Seoul, Republic of Korea

Abstract

Background: The global increase in life expectancy has not shown a similar rise in healthy life expectancy. Accurate assessment of biological aging is crucial for mitigating diseases and socioeconomic burdens associated with aging. Current biological age prediction models are limited by their reliance on conventional statistical methods and constrained clinical information.

Objective: This study aimed to develop and validate an aging clock model using artificial intelligence, based on comprehensive health check-up data, to predict biological age and assess its clinical relevance.

Methods: We used data from Koreans who underwent health checkups at the Seoul National University Hospital Gangnam Center as well as from the Korean Genome and Epidemiology Study. Our model incorporated 27 clinical factors and employed machine learning algorithms, including linear regression, least absolute shrinkage and selection operator, ridge regression, elastic net, random forest, support vector machine, gradient boosting, and K-nearest neighbors. Model performance was evaluated using adjusted R^2 and the mean squared error (MSE) values. Shapley Additive exPlanation (SHAP) analysis was conducted to interpret the model's predictions.

Results: The Gradient Boosting model achieved the best performance with a mean (SE) MSE of 4.219 (0.14) and a mean (SE) R^2 of 0.967 (0.001). SHAP analysis identified significant predictors of biological age, including kidney function markers, gender, glycated hemoglobin level, liver function markers, and anthropometric measurements. After adjusting for the chronological age, the predicted biological age showed strong associations with multiple clinical factors, such as metabolic status, body compositions, fatty liver, smoking status, and pulmonary function.

Conclusions: Our aging clock model demonstrates a high predictive accuracy and clinical relevance, offering a valuable tool for personalized health monitoring and intervention. The model's applicability in routine health checkups could enhance health management and promote regular health evaluations.

(JMIR Aging 2025;8:e64473) doi:10.2196/64473

KEYWORDS

biological age; aging clock; mortality; artificial intelligence; machine learning; record; history; health checkup; clinical relevance; gerontology; geriatric; older; elderly; aging; prediction; predictive; life expectancy; AI

Introduction

Over the past several decades, global life expectancy has increased remarkably, rising from 66.8 years in 2000 to 73.4 years in 2019, according to the World Health Organization. However, healthy life expectancy has not kept pace, increasing only from 58.3 years to 63.7 years during the same period [1]. This demographic shift toward an aging population has led to increased health care dependency and associated social costs. The medical industry related to aging and the social costs thereof are continuously increasing [2]. Accurately assessing biological aging is a critical first step in mitigating age-related diseases and their socioeconomic impact.

XSL•FO
RenderX

Biological age refers to an estimation of an individual's physiological and functional status, reflecting the cumulative effects of genetic, environmental, and lifestyle factors on the aging process [3]. An individual with a biological age younger than their chronological age may have a lower risk of developing age-related diseases, while an older biological age could indicate a higher vulnerability to such conditions. This highlights the clinical significance of accurately estimating the biological age for personalized health interventions and monitoring. While numerous studies have explored the human lifespan [4,5], their evaluation is challenging due to the required long-term observations and limited clinical applicability.

Applying findings from academic studies to clinical practice remains challenging. Current biological age prediction models, primarily based on conventional statistical methods such as multivariate regression analysis, rely on limited clinical data, restricting their predictive power and insights into the aging process [5-8]. Recent advances have led to models using omics data [9], including DNA methylation [10], transcriptome [11], metabolome [12], and telomere data [9]. However, these models face implementation challenges in clinical settings due to their complexity and the difficulty in measuring omics markers. The model requires multiple molecular modalities and functional data to exhibit a superior performance [9]. In addition, it is challenging to quantify the dynamic effects of environmental, lifestyle, behavioral, and interventional factors on biological age.

In Asian countries, the health checkup industry has been growing substantially [13-15], with individuals regularly monitoring their health status. However, these checkups typically only indicate normal or abnormal conditions for individual tests, lacking comprehensive health status indicators. Providing biological age predictions through an aging clock could serve as a valuable tool for the health screening of patients, offering a comprehensive health status measure and encouraging regular checkup participation.

This study investigated the clinical relevance of artificial intelligence (AI)-predicted biological age in the Korean population using comprehensive health checkup data, examining its relationship with various clinical characteristics.

Methods

Participants and Datasets

The study investigated the healthy population participating in comprehensive health checkups at the Seoul National University Hospital Gangnam center, from 2003 to 2016. The initial baseline data were used, and the participants included a total of 81,211 Koreans, who comprised the Health and Prevention Enhancement (H-PEACE) cohort. The details of the H-PEACE cohort have been described previously [16]. To summarize, each participant completed a questionnaire on their past medical history and underwent anthropometric measurements and laboratory tests after at least 10 hours of fasting on the same day. We also used the data from the Korean Genome and Epidemiology Study (KoGES) from the Korean Center for Disease Control and Prevention as the replication set. From the KoGES data, we used the health examination cohort (KoGES HEXA data), which included past medical history, anthropometric measurement, and laboratory data [17]. Briefly, the KoGES HEXA cohort is a national health examinee registry, consisting of 173,357 urban Korean adults who underwent health checkup programs. We used clinical factors overlapping between factors from the H-PEACE cohort data and the KoGES HEXA data. The exclusion criteria to define the healthy adult super-control cohort were as follows: (1) participants diagnosed with diabetes, hypertension, or dyslipidemia; (2) participants drinking alcohol more than 14 g/week; (3) current or previous smokers; (4) those aged less than 30 years; and (5) those having a history of malignant disease. To check the clinical implication of the predicted biological age, we performed multiclinical feature association study in the gene-environmental interaction and phenotype (GENIE) study [16], which consisted of 123 clinical factors and gene datasets. We used the 116 clinical factors, excluding the 27 factors used to predict biological age, to determine the multiple associations with predicted biological age.

Development of a Biological Age Prediction Model

This study aimed to develop a widely applicable biological age model using only basic health screening parameters. The clinical features used to construct the biological age model were based on routine clinical measurements, including standard demographic features, blood test results, and anthropometric measurements, rather than the findings from expensive specialized examinations. The selection of clinical factors was constrained by the requirement that they be present in both the training dataset (H-PEACE cohort) and the replication dataset (KoGES HEXA cohort). Through this feature selection process, we used 27 clinical factors as inputs for predicting the biological age, which are gender, anthropometric measurements (height, weight, BMI, and waist circumference), metabolic status (levels of fasting glucose, glycated hemoglobin [HbA1c], uric acid, total cholesterol, triglyceride, high-density lipoprotein cholesterol, and low-density lipoprotein cholesterol), liver functions (albumin, total bilirubin, alkaline phosphatase, aspartate aminotransferase, alanine aminotransferase, and gamma-glutamyl transferase), complete blood cell counts (white blood cell count, red blood cell count, hemoglobin, hematocrit, and platelet count), calcium and renal function (levels of blood urea nitrogen and creatinine, and the glomerular filtration rate calculated using the chronic kidney disease epidemiology collaboration (CKD-EPI) equation [18]).

The true labels were defined as the chronological age of the super-control population, based on the assumption that chronological age aligns with biological age in physiologically standard individuals. This cohort was carefully selected to exclude individuals with pathological conditions, such as metabolic diseases or malignant diseases, as well as those exposed to environmental factors known to influence biological age, including smoking and alcohol consumption [19].

The baseline data of the H-PEACE super-control cohort was split into 80% as the training set and 20% as the testing set with stratification based on both age and sex. Specifically, age stratification was conducted by categorizing participants into

XSL•FO
RenderX

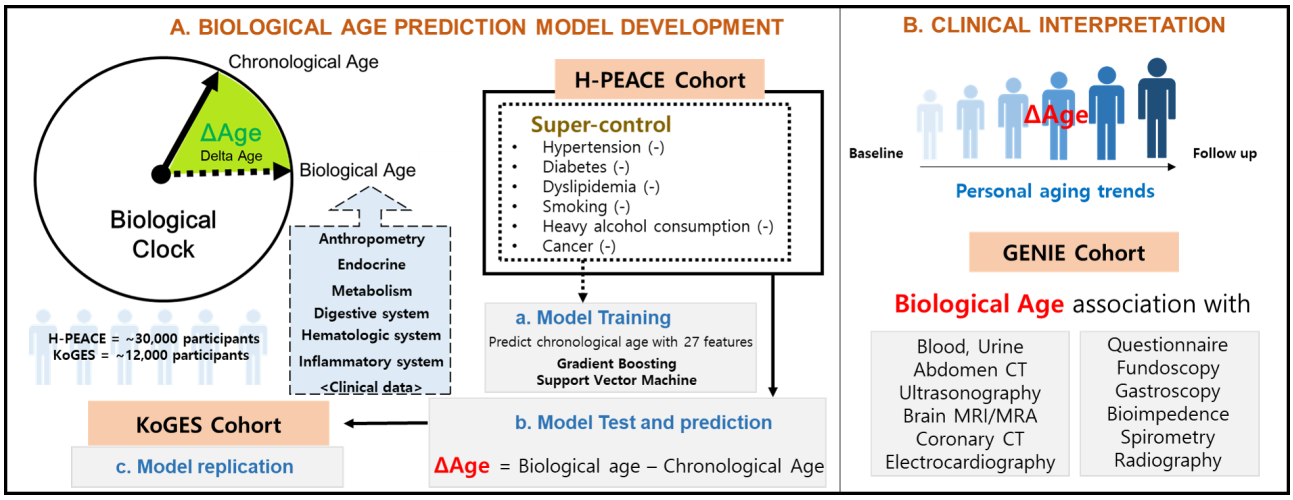
decade-based intervals, which were 20 - 29, 30 - 39, 40 - 49, 50 - 59, 60 - 69, 70 - 79, and more than 80 years. Five-fold cross-validation was performed to find the best hyperparameter for each model. We employed various machine learning models such as linear regression, least absolute shrinkage and selection operator (LASSO) regression, ridge regression, elastic net, random forest, support vector machine (SVM), gradient boosting, and K-nearest neighbors. These models were chosen for their ability to handle diverse relationships in data, including linear, nonlinear, and complex interactions. Hyperparameter optimization for each model was conducted using a grid search. To evaluate the performances of the developed models to predict biological age, we used adjusted R^2 and mean squared error (MSE) values. The evaluation results reported in this study are from 10 iterative experiments. A replication study was conducted in the KoGES HEXA dataset with the same experimental setting as the H-PEACE dataset. We interpreted the biological age prediction results using Shapley Additive exPlanation (SHAP) [20]. SHAP analysis was performed to elucidate the roles and impacts of different biological markers in predicting biological age.

Investigation of the Clinical Relevance of the Predicted Biological Age

We conducted linear regression analysis for multiple clinical factors using the GENIE study dataset with the predicted biological age, adjusting for chronological age. All reported *P* values were corrected for multiple tests using the Bonferroni correction.

The overview of the study is shown in Figure 1.

Figure 1. Overview of the study. H-PEACE: Health and Prevention Enhancement, KoGES: Korean Genome and Epidemiology Study, GENIE: gene-environmental interaction and phenotype, CT: computed tomography, MRI: magnetic resonance imaging, MRA: magnetic resonance angiography.



Ethical Considerations

The Institutional Review Board of the Seoul National University Hospital approved the study protocol and waived the need for informed consent (IRB number H-2005-223-1129). The study was performed in accordance with the Declaration of Helsinki. All patient data were deidentified before analysis, and strict confidentiality measures were implemented throughout the data collection, storage, and analysis processes. Access to the data was restricted to authorized research personnel only. The data management procedures complied with relevant data protection regulations to safeguard participants' privacy. No personally identifiable information was included in the final dataset or results.

Statistical and Computational Analyses

All analyses and calculations were performed using Python version 3.11.11 (Python Software Foundation). Multiple evaluation indices, including adjusted R^2 and MSE, β, and *P* value, were used to comprehensively evaluate the performances of the models and the significance of the associations. The statistical significance was based on a two-tailed *P* value of <.05.

Results

General Characteristics of the Participants

After applying the exclusion criteria to define the super-control cohort and removing the missing variables, we developed a model using a dataset collected from 28,417 individuals who underwent comprehensive health checkups at the Seoul National University Hospital Gangnam Center (ie, the H-PEACE cohort). The enrollment process is shown in Figure S1 in Multimedia Appendix 1. The mean (SD) participant age was 44.22 (11.26) years for 6467 men and 21,950 women. The baseline characteristics according to the aging of the study participants are shown in Table S1 in Multimedia Appendix 2. There were 1005 participants who were more than 65 years old and 27,412 who were 65 years old or less. Figure S2 in Multimedia Appendix 1 shows the chronological age distribution for the respective genders. Figure S3 in Multimedia Appendix 1 shows the distribution of gender in the training and test sets as well as the chronological age group distribution in the training and test sets.

Development and Performance of the Biological Age Prediction Model

The models were trained using 27 clinical variables on the full training dataset (80% of the super-control cohort), then predicted

XSL•FO
RenderX

on the test set (20%). For generalizability, 5-fold cross-validation with 10 iterative experiments were performed. The model was replicated using the KoGES HEXA data of 11,968 super-controls. Among 8 machine learning algorithms, the model showing the best performance was gradient boosting, for which the mean (SE) MSE value was 4.219 (0.140) and the mean (SE) R^2 value was 0.967 (0.001). The hyperparameters for the gradient boosting model were α=0.9, complexity parameter α=0, learning rate=0.1, maximum depth=5, and number of trees=500.

The second-best performing algorithm was the SVM model, with a mean (SE) MSE of 8.244 (0.210) and mean (SE) R^2 value of 0.935 (0.002). The performances of the 8 machine learning models in the test set are shown in Table 1.

Table . Comparison of the performances of 8 machine learning models to predict biological age in the test set (N=5684).

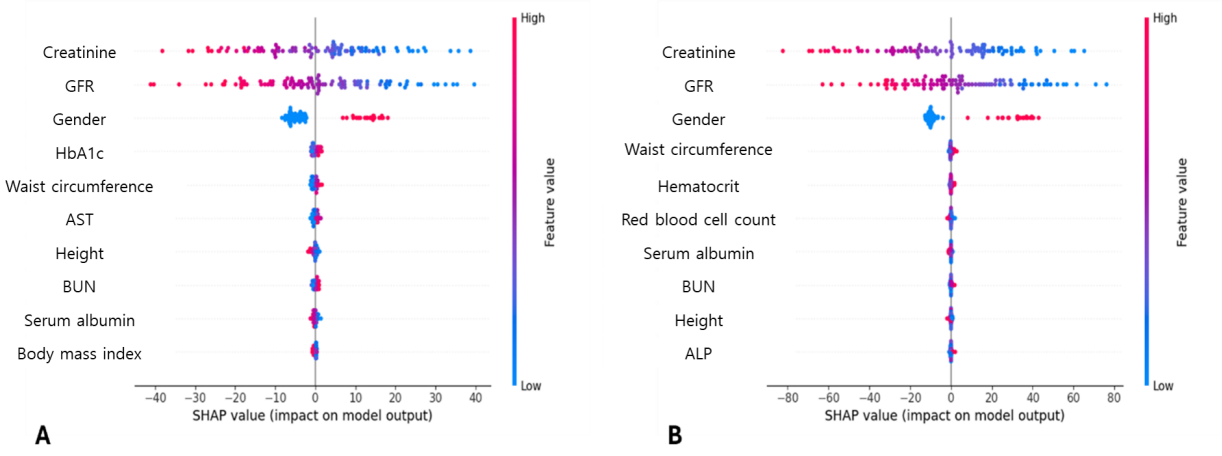
Model	Mean squared error, mean (standard error)	R^2, mean (standard error)
K nearest neighbor	63.829 (0.176)	0.497 (0.002)
Elastic net	50.518 (2.251)	0.602 (0.018)
Linear regression	50.314 (5.165)	0.603 (0.040)
LASSO[a]	50.271 (3.477)	0.604 (0.027)
Ridge	50.235 (5.047)	0.604 (0.039)
Random forest	20.941 (1.020)	0.835 (0.008)
Support vector machine	8.244 (0.210)	0.935 (0.002)
Gradient boosting	4.219 (0.140)	0.967 (0.001)

[a]LASSO: Least Absolute Shrinkage and Selection Operator.

The SHAP values from the gradient boosting and SVM models were analyzed to interpret their predictions of biological age. The SHAP values offer feature importance, providing the interpretation of the model's decision-making by quantifying the contribution of each feature to the model's output [20]. The corresponding visualizations are presented in Figure 2.

In the predictions of biological age generated by the gradient boosting model, the markers of kidney function, gender, HbA1c level, liver function, and anthropometric measurements were highlighted as significant predictors. In the SVM model, the SHAP summary plot revealed that kidney function markers, gender, liver function markers, red blood cell indices, and anthropometric measurements were the most influential predictors of biological age. These findings underscored the multifaceted nature of aging and highlighted the importance of maintaining optimal kidney function, metabolic status, and body composition in mitigating biological aging.

Figure 2. Visualization of the feature importance with Shapley Additive exPlanation (SHAP) values. SHAP summary plots for the model from the gradient boosting model (A) and the support vector machine (SVM) model (B) are visualized. Features with broader spreads and higher SHAP values have a more significant impact to predict biological age and values with the color gradient indicate whether higher or lower feature values are associated with increased biological age predictions. GFR: glomerular filtration rate, HbA1c: glycated hemoglobin, AST: aspartate aminotransferase, BUN: blood urea nitrogen, ALP: alkaline phosphatase.



Subgroup Analysis by Gender

The performances of the gradient boosting and SVM models were evaluated in the respective genders, male versus female participants. Applying the gradient boosting model to male participants, the mean (SE) MSE value was 5.258 (0.490). The SHAP value was significantly influenced by renal function, metabolic status, red blood cell indices, and anthropometric measurements. In female participants, the mean (SE) MSE value was 2.743 (0.099) and the SHAP values with the highest impact were similar to those for the male participants. The performances in the SVM and SHAP plots for male and female participants

XSL•FO
RenderX

are shown in Table 2 and Figure S4 in Multimedia Appendix 1.

Table . Comparison of the performances of the test set and replication set in all ages and in age above 65 years.

Dataset	Subgroup	Gradient boosting		Support Vector Machine	
		All ages	Age above 65 years	All ages	Age above 65 years
		MSE[a]	MSE[a]	MSE[a]	MSE[a]
Test set	Male and Female	4.219 (0.140)	18.942 (1.250)	8.244 (0.210)	18.077 (1.094)
Replication set	Male and Female	2.406 (0.083)	2.190 (0.085)	7.838 (0.065)	3.116 (0.148)
Test set	Male	5.258 (0.490)	20.357 (2.493)	15.476 (0.652)	20.678 (1.504)
Replication set	Male	2.032 (0.110)	2.096 (0.113)	6.337 (0.053)	3.143 (0.205)
Test set	Female	2.743 (0.099)	17.153 (0.819)	7.955 (0.170)	18.928 (1.698)
Replication set	Female	16.050 (0.467)	3.746 (0.366)	37.301 (0.777)	3.832 (0.510)

[a]Mean squared error (MSE) values are shown as mean (standard error).

Replication of the Developed Biological Age Prediction Model

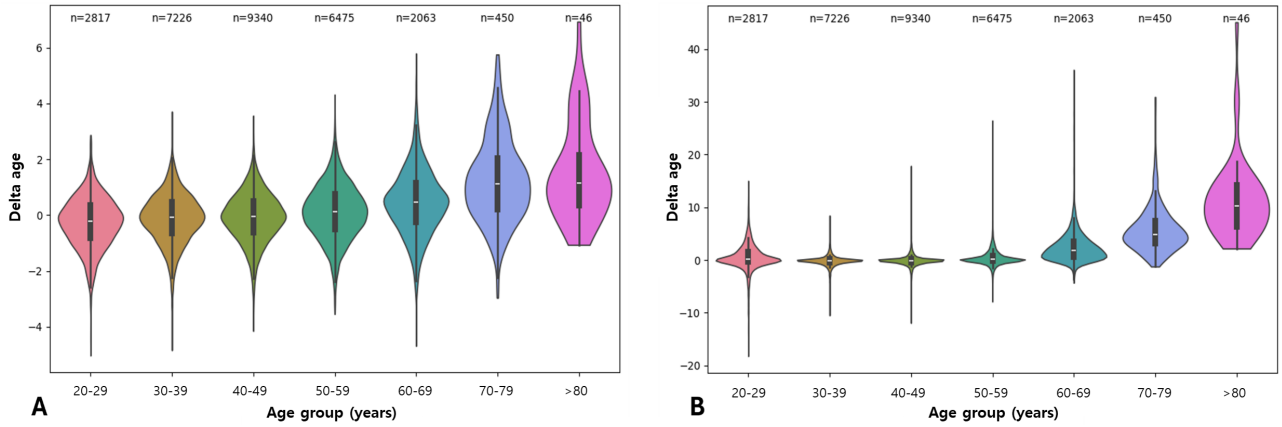
We replicated the model with the gradient boosting and SVM models using the KoGES HEXA data for participants of all ages and those under 65 years of age. The prediction model was better replicated in male participants and those aged above 65 years. The replicated results are shown in Table 2.

Delta Age Evaluation Using Baseline and Follow-Up Data

The model predicted the biological age in the test set (20% of the total dataset; 5684 individuals). The delta age, which is the difference between the biological age and the predicted biological age, was calculated for all individuals in the test dataset.

Figure 3 shows the delta age distribution across the different age groups, and Figure S5 in Multimedia Appendix 1 shows the delta age distribution across different age groups in each gender.

Figure 3. The delta age distribution across different age groups. The distribution of the delta age across different age groups in the test set at baseline (total n=5684) is shown for the gradient boosting (A) and support vector machine (B) models.



With the increase in the age groups, there was an observable pattern in the variability of delta age. In the younger age group (the 1st to the 3rd groups), there was increased variability, suggesting diverse aging processes influenced by multiple factors such as genetics, lifestyle, and health conditions. Older groups (the 7th and 8th groups) showed less variability and more positive values of delta age, possibly reflecting a selection of healthier individuals who have managed to reach older ages. The results were consistent in each gender group as well.

Among 5684 baseline participants, there were 2022 participants who had follow-up data, with follow-up duration ranging from 1 to 10 years (with 0 meaning the same age at two consecutive visits).

Figure 4 shows the projectiles of the changes in delta age during follow-up across different age groups in the test set (n=2022) for the gradient boosting and SVM models. While the younger age groups exhibited stable biological aging trajectories, the middle-aged and older groups showed increased variability and accelerated aging over the follow-up period. The observed trends were consistent across both the gradient boosting and SVM models, providing robust evidence for the described aging patterns.

XSL•FO
RenderX

Figure 4. The delta age trajectory during follow-up across different age groups. The projectiles of the changes in delta age during follow-up across different age groups in the test set (n=2022) are shown for the gradient boosting model (A) and SVM model.

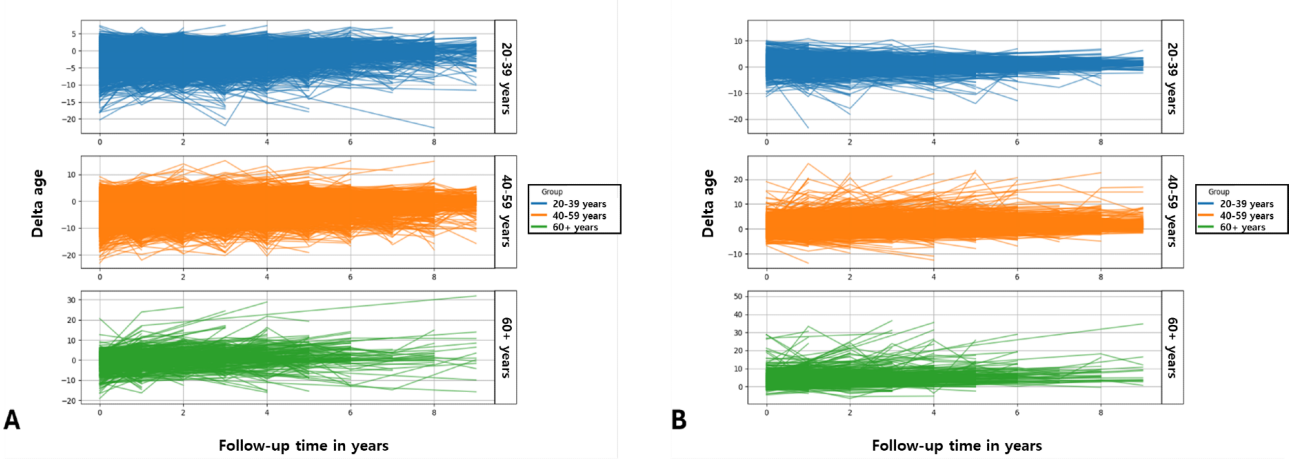


Figure S6 in Multimedia Appendix 1 shows the trajectory of the changes in delta age during follow-up across different gender groups in the test set at baseline (n= 5684) for the gradient boosting model and SVM model.

Clinical Interpretation Using a Multiple Clinical Factor Association Study

To explore the clinical relevance of the predicted biological age, we performed a phenome-wide association analysis between 116 phenotypes corroborated by comprehensive health checkups from the GENIE study dataset and the output of the predicted biological age model (data for both genders are shown in Table S2 in Multimedia Appendix 2, data for the male gender are shown in Table S3 in Multimedia Appendix 2, and data for the female gender are shown in Table S4 in Multimedia Appendix 2).

We found that the predicted biological age was significantly associated with 50 clinical factors after Bonferroni correction (Tables S2-S4 in Multimedia Appendix 2). Notably, the most significant associations were observed with metabolic status, body composition, fatty liver, smoking status, and pulmonary function, after adjusting for the chronological age. Table 3 shows the top 10 significant results in both genders, in male participants, and in female participants. The correlations among the variables are shown in Figure S7 in Multimedia Appendix 1. These findings suggest that the predicted biological age, developed from multiple clinical factors from comprehensive health checkups, may also possess predictive capabilities for aging in various organs.

XSL•FO
RenderX

Table . Top 10 significant clinical factors associated with the predicted biological age.

Clinical factors	Study modality	N	β	*P* value
In both genders				
Diabetes diagnosis	Questionnaire	10,351	−.002	1.75×10^{-270}
Skeletal muscle mass	Bioelectrical impedance	10,231	−.053	7.87×10^{-256}
Fatty liver	Ultrasonography	10,287	−.006	5.20×10^{-195}
Diabetes medication	Questionnaire	10,351	−.001	4.38×10^{-178}
Metabolic syndrome	Questionnaire	10,351	−.003	1.35×10^{-174}
Smoking status	Questionnaire	8995	−.004	8.43×10^{-170}
Visceral fat area	Abdominal computed tomography	6183	−46.952	2.86×10^{-153}
Forced vital capacity (liters)	Spirometry	10,138	−.005	6.32×10^{-141}
Forced expiratory volume in 1 s (liters)	Spirometry	10,138	−.004	5.71×10^{-135}
Chloride level	Blood	9964	.011	4.18×10^{-77}
In male participants				
Diabetes diagnosis	Questionnaire	4292	−.001	1.91×10^{-56}
Fatty liver	Ultrasonography	4270	−.004	2.74×10^{-46}
Metabolic syndrome	Questionnaire	4292	−.002	1.72×10^{-41}
Diabetes medication	Questionnaire	4292	−.001	9.63×10^{-36}
Mean corpuscular hemoglobin	Blood	4264	.013	3.61×10^{-35}
Mean corpuscular volume	Blood	4264	.026	1.54×10^{-27}
Visceral fat area	Abdominal computed tomography	2157	−26.868	3.61×10^{-24}
Total fat area	Abdominal computed tomography	2157	−48.823	1.23×10^{-14}
Mean corpuscular hemoglobin concentration	Blood	4264	.005	5.10×10^{-14}
Heart rate	Electrocardiography	3297	−.040	1.02×10^{-12}
In female participants				
Diabetes diagnosis	Questionnaire	6059	−.002	4.13×10^{-146}
Diabetes medication	Questionnaire	6059	−.002	2.49×10^{-98}
Metabolic syndrome	Questionnaire	6059	−.002	1.93×10^{-38}
Fatty liver	Ultrasonography	6017	−.004	5.03×10^{-31}
Potassium level	Blood	5839	.001	4.02×10^{-24}
Sodium level	Blood	5839	.007	1.59×10^{-23}
Chloride level	Blood	5839	.007	3.07×10^{-15}
Forced vital capacity (liters)	Spirometry	5956	.002	3.97×10^{-15}
Forced vital capacity percent	Spirometry	5956	.032	4.71×10^{-14}
Visceral fat area	Abdominal computed tomography	4026	−17.835	9.80×10^{-14}

XSL•FO
RenderX

Discussion

This study underscores the clinical relevance of biological age as predicted by AI using comprehensive health checkup data. Our findings elucidate the clinical relevance of biological age—as assessed through machine learning models—which exhibits strong associations with multiple clinical factors, thereby providing valuable insights into the aging process and its implications on health. The result demonstrated that the best-performing model, gradient boosting, achieved a mean (SE) MSE value of 4.219 (0.140) and mean (SE) R^2 value of 0.967 (0.001), indicating acceptable predictive accuracy. Additionally, the SHAP analysis highlighted markers such as kidney function, gender, HbA1c level, liver function, and anthropometric measurements as significant predictors of biological age. These findings suggest that biological age is a multifaceted construct influenced by various physiological factors and can serve as a robust indicator of the overall health status. The identification of these markers is significant because it supports the importance of maintaining optimal kidney function, metabolic health, and body composition in mitigating biological aging. This comprehensive understanding can lead to more targeted and effective interventions aimed at improving overall health and longevity.

Our study findings align with existing research findings that emphasize the utility of biological age as a comprehensive health indicator [21-23]. Previous studies have also identified the importance of factors such as kidney function [24], metabolic health [25,26], and inflammatory markers [27,28] in the aging process. However, our use of machine learning models such as the gradient boosting model and the SVM model provides a more nuanced understanding of how these factors interact to influence biological age, setting our research apart from traditional statistical methods. For instance, while existing literature has demonstrated the importance of individual biomarkers, our approach integrates multiple clinical variables to provide a holistic prediction model.

The use of multiple machine learning algorithms allowed for a comprehensive evaluation of their predictive capabilities. The gradient boosting model outperformed other algorithms, with performance R^2 of 0.967 and MSE of 4.219, due to its ability to handle nonlinear relationships and feature interactions, which are crucial in modeling complex biological systems [29]. Unlike linear models such as the ridge or LASSO regression, the gradient boosting model iteratively optimizes residual errors, capturing intricate dependencies between features. Its robustness to outliers and noise further enhances performance, making it particularly suited for real-world health checkups [30,31]. The generalizability of gradient boosting, demonstrated in both the training and replication datasets, underscores its potential for clinical application. Future studies should explore hybrid approaches to integrate the strengths of the gradient boosting model with more efficient models for broader use.

In the SHAP analysis, kidney function emerged as a significant predictor of biological age, consistent with its established role as an indicator of systemic health [32]. The decrease in kidney function is closely associated with aging-related changes, such as a reduced glomerular filtration rate and increased risk of chronic kidney disease [33]. These conditions often signify cumulative damage from metabolic stressors, hypertension, and other age-related factors [34]. The inclusion of kidney function markers, such as creatinine and the estimated glomerular filtration rate, in our model highlights their critical contribution to capturing the physiological aging process. Maintaining optimal kidney function could therefore serve as a target for mitigating biological aging and reducing the burden of associated comorbidities.

The body composition, including markers such as waist circumference and BMI, was another key contributor identified on SHAP analysis. This finding underscores the impact of adiposity and muscle mass on aging trajectories. Increased visceral fat and reduced skeletal muscle mass are hallmark features of sarcopenic obesity, a condition linked to metabolic dysfunction and accelerated biological aging [35,36]. Such changes exacerbate systemic inflammation and insulin resistance, further compounding aging-related risks [37,38]. The identification of body composition as a significant predictor emphasizes the importance of lifestyle interventions, such as exercise and dietary modifications, to preserve muscle mass and manage body fat. By integrating these markers into the prediction model, our study highlights their multifaceted roles in biological aging. Kidney function and body composition serve not only as indicators of systemic health but also as modifiable factors, offering potential avenues for personalized health interventions aimed at slowing biological aging and promoting longevity. This enhanced understanding enriches the clinical utility of our model and provides actionable insights for practitioners.

Visualization of the delta age distribution across different age groups shows that there is increased variability in the younger age group and less variability and a more positive value of delta age in the older group. This pattern underscores the importance of middle age as a critical period for implementing health interventions to manage biological aging effectively, and the need for focused care strategies to manage aging-related health issues in older adults.

In the multiple clinical factor association study for predicted biological age, we found that the predicted biological age was significantly associated with 50 clinical factors even after the adjustment of chronological age. The most significant associations were observed with the metabolic status, body composition, fatty liver, smoking status, and pulmonary function. These associations highlight the clinical relevance of biological age as a comprehensive marker of an individual's health status and suggest potential pathophysiological mechanisms accompanying the aging process. Reduced pulmonary function is known to be an indicator of systemic aging and has been linked to increased mortality and morbidity in older adults [39,40]. This suggests that maintaining optimal lung function could be crucial in mitigating the effects of biological aging and improving overall health outcomes.

The association between smoking status and predicted biological age is well-documented in various studies. For example, research using AI to analyze blood data from 149,000 adults revealed

XSL•FO
RenderX

that smokers exhibit a faster rate of biological aging compared to nonsmokers [41]. Remarkably, female smokers were predicted to have a biological age twice that of their chronological age [42]. These findings emphasize the detrimental impact of smoking on biological age. Additionally, exposure to tobacco during early life has been linked to accelerated biological aging in adulthood [43]. The underlying mechanisms include inflammation [44,45], epigenetic changes [46], and chromosomal damage [47]. These results underline the potential for biological recovery after cessation, aligning with prior research and emphasizing both the dangers of smoking and the benefits of quitting.

Body composition changes, including increased body fat, decreased muscle mass, and higher BMI, were also significantly associated with biological age. These findings align with previous studies that have shown that sarcopenia (loss of muscle mass) and obesity are critical factors in the aging process [48]. Excess body fat, particularly visceral fat, is associated with metabolic syndrome, cardiovascular diseases, and insulin resistance, all of which contribute to accelerated biological aging [49,50]. Conversely, maintaining muscle mass is essential for physical function and metabolic health, highlighting the importance of exercise and nutrition in managing the biological age. These findings also provide insights into the pathophysiological aspects of aging, suggesting that targeted interventions in these areas could potentially mitigate the adverse effects of aging and improve health outcomes.

The results have significant implications for health care and aging research. By providing a comprehensive measure of biological age, health care providers can better assess an individual's overall health status and identify potential risks for age-related diseases. This approach could also facilitate more personalized health care strategies, improving patient outcomes and quality of life. Furthermore, the findings contribute to the broader understanding of the aging process, offering valuable insights for researchers aiming to develop innovative strategies to overcome age-related diseases and enhance healthy aging.

Despite the robust findings, there are limitations to consider. The study population was predominantly Korean, which may limit the generalizability of the results to other ethnic groups. Additionally, the reliance on clinical data from health checkups may not capture all factors influencing biological age, such as genetic predispositions or epigenetic clocks and environmental influences. Future studies should aim to include more diverse populations and incorporate additional variables to enhance the comprehensiveness and applicability of the models. Furthermore, longitudinal studies are needed to validate the long-term predictive capabilities of the models and their effectiveness in different health care settings.

In conclusion, our study demonstrates the clinical relevance of biological age as predicted by machine learning models. The findings provide valuable insights into the aging process and highlight the potential of biological age as a comprehensive health indicator. Future research should focus on refining these models and exploring their applicability in diverse populations to further enhance their utility in clinical practice. The integration of advanced machine learning techniques with comprehensive health data holds great promise for advancing our understanding of aging and improving health care outcomes.

Acknowledgments

This study received funding from the Seoul National University Hospital Research Fund (grant 04-2020-2100).

Data Availability

Data are not publicly available due to restrictions (institutional policy to protect the privacy of research participants), but are available from the corresponding author on reasonable request.

Authors' Contributions

EKC and DK equally contributed as corresponding authors, and DK can be reached at dokyoon.kim@pennmedicine.upenn.edu for this purpose.
Data curation: EKC
Conceptualization and methodology: EKC and DK
Writing, analysis and interpretation of data, and software: EKC, CJ, and JSL
Funding acquisition: EKC
Review and editing and visualization: EKC, DK, CJ, and JSL
Data acquisition: EKC
All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary figures.
[PDF File, 1208 KB - aging_v8i1e64473_app1.pdf]

XSL•FO
RenderX

Multimedia Appendix 2
Supplementary tables.
[XLSX File, 36 KB - aging_v8i1e64473_app2.xlsx]

References

1. World Health Organization. GHE: Life expectancy and healthy life expectancy. URL: https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghe-life-expectancy-and-healthy-life-expectancy [accessed 2024-06-20]
2. Chen J, Zhao M, Zhou R, Ou W, Yao P. How heavy is the medical expense burden among the older adults and what are the contributing factors? A literature review and problem-based analysis. Front Public Health 2023;11:1165381. [doi: 10.3389/fpubh.2023.1165381] [Medline: 37397714]
3. Wang Q, Wang Z, Mizuguchi K, Takao T. Biological age prediction using a novel DNN model based on steroid metabolic pathways. Bioinformatics. . [doi: 10.1101/2024.11.07.622443]
4. Sun Q, Yu D, Fan J, et al. Healthy lifestyle and life expectancy at age 30 years in the Chinese population: an observational study. Lancet Public Health 2022 Dec;7(12):e994-e1004. [doi: 10.1016/S2468-2667(22)00110-4] [Medline: 35926549]
5. Wang J, Chen C, Zhou J, et al. Healthy lifestyle in late-life, longevity genes, and life expectancy among older adults: a 20-year, population-based, prospective cohort study. Lancet Healthy Longev 2023 Oct;4(10):e535-e543. [doi: 10.1016/S2666-7568(23)00140-X] [Medline: 37804845]
6. Voitenko VP, Tokar AV. The assessment of biological age and sex differences of human aging. Exp Aging Res 1983;9(4):239-244. [doi: 10.1080/03610738308258458] [Medline: 6667707]
7. Takeda H, Inada H, Inoue M, Yoshikawa H, Abe H. Evaluation of biological age and physical age by multiple regression analysis. Med Inform (Lond) 1982;7(3):221-227. [doi: 10.3109/14639238209010720] [Medline: 7162237]
8. Krøll J, Saxtrup O. On the use of regression analysis for the estimation of human biological age. Biogerontology 2000;1(4):363-368. [doi: 10.1023/a:1026594602252] [Medline: 11708217]
9. Rutledge J, Oh H, Wyss-Coray T. Measuring biological age using omics data. Nat Rev Genet 2022 Dec;23(12):715-727. [doi: 10.1038/s41576-022-00511-7] [Medline: 35715611]
10. Bell CG, Lowe R, Adams PD, et al. DNA methylation aging clocks: challenges and recommendations. Genome Biol 2019 Nov 25;20(1):249. [doi: 10.1186/s13059-019-1824-y] [Medline: 31767039]
11. Mao S, Su J, Wang L, Bo X, Li C, Chen H. A transcriptome-based single-cell biological age model and resource for tissue-specific aging measures. Genome Res 2023 Aug;33(8):1381-1394. [doi: 10.1101/gr.277491.122] [Medline: 37524436]
12. Lassen JK, Wang T, Nielsen KL, Hasselstrøm JB, Johannsen M, Villesen P. Large-scale metabolomics: predicting biological age using 10,133 routine untargeted LC-MS measurements. Aging Cell 2023 May;22(5):e13813. [doi: 10.1111/acel.13813] [Medline: 36935524]
13. Lu J. Ningen Dock: Japan's unique comprehensive health checkup system for early detection of disease. Glob Health Med 2022 Feb 28;4(1):9-13. [doi: 10.35772/ghm.2021.01109] [Medline: 35291200]
14. Gong W, Cheng KK. Challenges in screening and general health checks in China. Lancet Public Health 2022 Dec;7(12):e989-e990. [doi: 10.1016/S2468-2667(22)00207-9] [Medline: 36462520]
15. Kang HT. Current status of the national health screening programs in South Korea. Korean J Fam Med 2022 May;43(3):168-173. [doi: 10.4082/kjfm.22.0052] [Medline: 35610963]
16. Lee C, Choe EK, Choi JM, et al. Health and Prevention Enhancement (H-PEACE): a retrospective, population-based cohort study conducted at the Seoul National University Hospital Gangnam Center, Korea. BMJ Open 2018 Apr 19;8(4):e019327. [doi: 10.1136/bmjopen-2017-019327] [Medline: 29674364]
17. Kim Y, Han BG, GESg K. Cohort profile: the Korean Genome and Epidemiology Study (KoGES) consortium. Int J Epidemiol 2017 Apr 1;46(2):1350. [doi: 10.1093/ije/dyv316] [Medline: 27085081]
18. Fu EL, Levey AS, Coresh J, et al. Accuracy of GFR estimating equations based on creatinine, cystatin C or both in routine care. Nephrol Dial Transplant 2024 Mar 27;39(4):694-706. [doi: 10.1093/ndt/gfad219] [Medline: 37813817]
19. Husted KLS, Brink-Kjær A, Fogelstrøm M, et al. A model for estimating biological age from physiological biomarkers of healthy aging: cross-sectional study. JMIR Aging 2022 May 10;5(2):e35696. [doi: 10.2196/35696] [Medline: 35536617]
20. Wang K, Tian J, Zheng C, et al. Interpretable prediction of 3-year all-cause mortality in patients with heart failure caused by coronary heart disease based on machine learning and SHAP. Comput Biol Med 2021 Oct;137:104813. [doi: 10.1016/j.compbiomed.2021.104813] [Medline: 34481185]
21. Kang YG, Suh E, Lee JW, Kim DW, Cho KH, Bae CY. Biological age as a health index for mortality and major age-related disease incidence in Koreans: National Health Insurance Service - Health screening 11-year follow-up study. Clin Interv Aging 2018;13:429-436. [doi: 10.2147/CIA.S157014] [Medline: 29593385]
22. Jia Q, Chen C, Xu A, et al. A biological age model based on physical examination data to predict mortality in a Chinese population. iScience 2024 Mar 15;27(3):108891. [doi: 10.1016/j.isci.2024.108891] [Medline: 38384842]
23. Bafei SEC, Shen C. Biomarkers selection and mathematical modeling in biological age estimation. NPJ Aging 2023 Jul 1;9(1):13. [doi: 10.1038/s41514-023-00110-8] [Medline: 37393295]

XSL•FO
RenderX

24. Peng S, Xu R, Wei K, Liu N, Lv Y, Lin Y. Association between kidney function and biological age: a China health and retirement longitudinal study. Front Public Health 2023;11:1259074. [doi: 10.3389/fpubh.2023.1259074] [Medline: 38164447]
25. Bae CY, Piao M, Kim M, et al. Biological age and lifestyle in the diagnosis of metabolic syndrome: the NHIS health screening data, 2014-2015. Sci Rep 2021 Jan 11;11(1):444. [doi: 10.1038/s41598-020-79256-4] [Medline: 33431923]
26. Devrajani T, Abid S, Shaikh H, et al. Relationship between aging and control of metabolic syndrome with telomere shortening: a cross-sectional study. Sci Rep 2023 Oct 19;13(1):17878. [doi: 10.1038/s41598-023-44715-1] [Medline: 37857729]
27. Kozlov M. "Inflammation clock" can reveal body's biological age. Nature New Biol 2021 Jul 13. [doi: 10.1038/d41586-021-01915-x] [Medline: 34257449]
28. Li X, Li C, Zhang W, Wang Y, Qian P, Huang H. Inflammation and aging: signaling pathways and intervention therapies. Signal Transduct Target Ther 2023 Jun 8;8(1):239. [doi: 10.1038/s41392-023-01502-8] [Medline: 37291105]
29. Alghushairy O, Ali F, Alghamdi W, Khalid M, Alsini R, Asiry O. Machine learning-based model for accurate identification of druggable proteins using light extreme gradient boosting. J Biomol Struct Dyn 2024;42(22):12330-12341. [doi: 10.1080/07391102.2023.2269280] [Medline: 37850427]
30. Yarkın Yıldız A, Kalayci A. Gradient boosting decision trees on medical diagnosis over tabular data. arXiv. Preprint posted online on Oct 11, 2024. [doi: 10.48550/arXiv.2410.03705]
31. Singh R, Kumar A, Singh P, Sharma M. Improving cancer prediction in recommender systems through gradient boosting. Presented at: 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT); Jun 24-26, 2024; Kamand, India p. 1-6. [doi: 10.1109/ICCCNT61001.2024.10724973]
32. Xu L, Yu C, Chen A, Li C, Mao Y. Longitudinal analysis of renal function changes in elderly populations: health status evaluation and risk factor assessment. Clin Interv Aging 2024;19:1217-1224. [doi: 10.2147/CIA.S450388] [Medline: 38974512]
33. Muglia L, Di Dio M, Filicetti E, et al. Biomarkers of chronic kidney disease in older individuals: navigating complexity in diagnosis. Front Med (Lausanne) 2024;11:1397160. [doi: 10.3389/fmed.2024.1397160] [Medline: 39055699]
34. Fossel M. Aging: How Aging Works, How We Reverse Aging, and Prospects for Curing Aging Diseases, 1st edition: Academic Press; 2024:1-17. [doi: 10.1016/B978-0-443-15500-0.00001-3]
35. Fermín-Martínez CA, Márquez-Salinas A, Guerra EC, et al. AnthropoAge, a novel approach to integrate body composition into the estimation of biological age. Aging Cell 2023 Jan;22(1):e13756. [doi: 10.1111/acel.13756] [Medline: 36547004]
36. Wei S, Nguyen TT, Zhang Y, Ryu D, Gariani K. Sarcopenic obesity: epidemiology, pathophysiology, cardiovascular disease, mortality, and management. Front Endocrinol (Lausanne) 2023;14:1185221. [doi: 10.3389/fendo.2023.1185221] [Medline: 37455897]
37. Aging Biomarker Consortium, Bao H, Cao J, et al. Biomarkers of aging. Sci China Life Sci 2023 May;66(5):893-1066. [doi: 10.1007/s11427-023-2305-0] [Medline: 37076725]
38. Briand M, Raffin J, Gonzalez-Bautista E, et al. Body composition and aging: cross-sectional results from the INSPIRE study in people 20 to 93 years old. Geroscience 2025 Feb;47(1):863-875. [doi: 10.1007/s11357-024-01245-6] [Medline: 39028455]
39. Sharma G, Goodwin J. Effect of aging on respiratory system physiology and immunology. Clin Interv Aging 2006;1(3):253-260. [doi: 10.2147/ciia.2006.1.3.253] [Medline: 18046878]
40. Thomas ET, Guppy M, Straus SE, Bell KJL, Glasziou P. Rate of normal lung function decline in ageing adults: a systematic review of prospective cohort studies. BMJ Open 2019 Jun 27;9(6):e028150. [doi: 10.1136/bmjopen-2018-028150] [Medline: 31248928]
41. Mamoshina P, Kochetov K, Cortese F, et al. Blood biochemistry analysis to detect smoking status and quantify accelerated aging in smokers. Sci Rep 2019 Jan 15;9(1):142. [doi: 10.1038/s41598-018-35704-w] [Medline: 30644411]
42. Skjodt N, Mamoshina P, Kochetov K, et al. Smoking causes early biological aging: a deep neural network analysis of common blood test results. 2018 Sep 15 Presented at: ERS International Congress 2018 abstracts; Sep 15, 2018. [doi: 10.1183/13993003.congress-2018.OA3809]
43. Cui F, Tang L, Li D, et al. Early-life exposure to tobacco, genetic susceptibility, and accelerated biological aging in adulthood. Sci Adv 2024 May 3;10(18):eadl3747. [doi: 10.1126/sciadv.adl3747] [Medline: 38701212]
44. Babizhayev MA, Yegorov YE. Smoking and health: association between telomere length and factors impacting on human disease, quality of life and life span in a large population-based cohort under the effect of smoking duration. Fundam Clin Pharmacol 2011 Aug;25(4):425-442. [doi: 10.1111/j.1472-8206.2010.00866.x] [Medline: 20698892]
45. Huang J, Okuka M, McLean M, Keefe DL, Liu L. Effects of cigarette smoke on fertilization and embryo development in vivo. Fertil Steril 2009 Oct;92(4):1456-1465. [doi: 10.1016/j.fertnstert.2008.07.1781] [Medline: 19019360]
46. Klopack ET, Carroll JE, Cole SW, Seeman TE, Crimmins EM. Lifetime exposure to smoking, epigenetic aging, and morbidity and mortality in older adults. Clin Epigenetics 2022 May 28;14(1):72. [doi: 10.1186/s13148-022-01286-8] [Medline: 35643537]
47. Mitiku T, Abebe B, Birhane N, Getie B. DNA Damage, Transposable Element Expression and Their Associated Factors in Aging: Preprints; 2024. [doi: 10.20944/preprints202410.1563.v1]

XSL•FO
RenderX

48. Cruz-Jentoft AJ, Bahat G, Bauer J, et al. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing 2019 Jan 1;48(1):16-31. [doi: 10.1093/ageing/afy169] [Medline: 30312372]
49. Bergman RN, Kim SP, Hsu IR, et al. Abdominal obesity: role in the pathophysiology of metabolic disease and cardiovascular risk. Am J Med 2007 Feb;120(2 Suppl 1):S3-S8. [doi: 10.1016/j.amjmed.2006.11.012] [Medline: 17296343]
50. Shah RV, Murthy VL, Abbasi SA, et al. Visceral adiposity and the risk of metabolic syndrome across body mass index: the MESA Study. JACC Cardiovasc Imaging 2014 Dec;7(12):1221-1235. [doi: 10.1016/j.jcmg.2014.07.017] [Medline: 25440591]

Abbreviations

AI: artificial intelligence
CKD-EPI: chronic kidney disease epidemiology collaboration
GENIE: gene-environmental interaction and phenotype
H-PEACE: Health and Prevention Enhancement
HbA1c: glycated hemoglobin
HEXA: health examination
KoGES: Korean Genome and Epidemiology Study
LASSO: least absolute shrinkage and selection operator
MSE: mean squared error
SHAP: Shapley Additive exPlanation
SVM: support vector machine

Edited by A Quialheiro; submitted 18.07.24; peer-reviewed by M Gasmi, M Malhotra, RS Gomaa Mahmoud, Y Cheng, Y Pan; revised version received 29.01.25; accepted 10.03.25; published 11.04.25.

Please cite as:
Jeong CU, Leiby JS, Kim D, Choe EK
Artificial Intelligence-Driven Biological Age Prediction Model Using Comprehensive Health Checkup Data: Development and Validation Study
JMIR Aging 2025;8:e64473
URL: https://aging.jmir.org/2025/1/e64473
doi:10.2196/64473

© Chang-Uk Jeong, Jacob S Leiby, Dokyoon Kim, Eun Kyung Choe. Originally published in JMIR Aging (https://aging.jmir.org), 11.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Peer Volunteers' Journeys Through Training and Engagement in Older Adult Communities: Descriptive Qualitative Study

Shefaly Shorey[1], PhD; Jeremy Soon Leong Seow[1]; Rathi Mahendran[2]; Geok Hua Wee[3]; Lee Gan Goh[4]; Ee Heok Kua[2]

[1]Alice Lee Centre for Nursing Studies, Yong Loo Lin School of Medicine, National University of Singapore, 10 Medical Dr, Singapore, Singapore

[2]Department of Psychological Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

[3]Principle Mindful Awareness Trainer (MAP), Age-Well Everyday Programme (AWE), Singapore, Singapore

[4]Division of Family Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

Corresponding Author:
Shefaly Shorey, PhD
Alice Lee Centre for Nursing Studies, Yong Loo Lin School of Medicine, National University of Singapore, 10 Medical Dr, Singapore, Singapore

Abstract

Background: The rising prevalence of mental health conditions such as depression and anxiety among the aging population underscores the need for accessible and effective psychosocial support, particularly for community-dwelling older adults who face barriers like social stigma and limited mental health literacy. Peer volunteers have emerged as a promising resource to support these individuals; yet, they often lack the requisite training for effective intervention.

Objective: This study aims to explore the experiences of peer volunteers who participated in a Psychological First Aid training program.

Methods: Using a descriptive qualitative research design, semistructured interviews were conducted with 13 older adults between September and October 2024, and data were thematically analyzed.

Results: Three themes were identified: (1) dimensions of volunteerism from motivations to resistance, (2) empowerment through collaborative learning, and (3) recommendations for designing inclusive, holistic training programs.

Conclusions: The findings of this study showed positive outcomes such as personal growth and strengthened social connections among participants. However, enhancements in teaching methods, logistical arrangements, and session regularity are recommended to optimize the Psychological First Aid program. These insights can guide the development of more robust training models to support both peer volunteers and the older adult communities they serve.

(JMIR Aging 2025;8:e71810) doi:10.2196/71810

KEYWORDS

psychological first aid; PFA training; peer volunteers; mental health; psychosocial support; older adults

Introduction

The global population of older adults has grown significantly in recent years. In Singapore, the proportion of citizens aged 65 years and older increased to 19.1% in 2023 and is projected to reach 24.1% by 2030 [1]. Globally, approximately 14% of older adults live with mental health conditions, with depression and anxiety being the most common [2]. This trend is mirrored in Singapore, where issues such as depression, dementia, and subsyndromal depression are increasingly prevalent among older adults [3]. Social isolation, multimorbidity, and low socioeconomic status have been identified as key contributors to the development of mental health conditions in older adults. These factors are associated with severe consequences, including increased suicide risk, diminished quality of life, and heightened susceptibility to cognitive disorders such as dementia [4-7]. Unfortunately, mental health conditions in this population often remain undiagnosed or underreported due to stigma and limited mental health literacy, leading to delays in treatment and intervention [8,9].

Volunteerism has been recognized as an effective approach to improving the mental well-being of older adults. Programs such as AmeriCorps Seniors and Age UK internationally, and SG Cares and Lions Befrienders locally, provide platforms for older adults to engage socially and meaningfully, helping to reduce social isolation and promote active aging [10-15]. Despite these benefits, studies conducted in Portugal and the United Kingdom reveal that volunteers often face challenges due to insufficient training to support individuals with mental health conditions [16,17]. Structured training programs have been increasingly recognized for their ability to prepare peer volunteers to provide effective psychosocial support. Peer volunteers' training has been widely used in various settings to help individuals cope with potentially traumatic events, reduce distress, and foster a

XSL•FO
RenderX

sense of safety while supporting others [18,19]. Such programs have also proven beneficial not only in humanitarian aid situations but also in enhancing volunteers' ability to offer timely, empathetic, and informal support [20,21]. Formal training equips volunteers with the knowledge and skills necessary to provide effective support to individuals in distress while also addressing the volunteers' own needs [22].

However, most studies examining peer volunteer training have been conducted in Western contexts, highlighting a critical gap in understanding its application and effectiveness in Asian settings [23-25]. Given that Asian cultures are deeply rooted in collectivism and family-centered caregiving, there is strong evidence that Western models shaped by individualism and structured volunteer systems may not fully translate to these contexts [26]. Considering the growing prevalence of mental health conditions among Singapore's older adults and the potential benefits of formal training programs, there is an urgent need to implement and evaluate such initiatives locally.

This study aims to explore the experiences of peer volunteers who participated in a Psychological First Aid (PFA) training program designed to support community-dwelling older adults in Singapore. Specifically, it addresses the research question: "What are the experiences of peer volunteers who completed the PFA programme?" The findings will inform the refinement of culturally relevant and effective training strategies to enhance psychosocial support for both volunteers and the older adult population in Singapore.

Methods

Study Design

A descriptive qualitative research design was used to explore the experiences of peer volunteers who participated in a PFA program. The use of a qualitative design is prudent as it allows participants to share "thick" or "rich" descriptions, providing deeper insights into subjective topics [27]. The Consolidated Criteria for Reporting Qualitative Research (COREQ) checklist guided the reporting of this qualitative study [28] (Table S1 in Multimedia Appendix 1).

Sampling Strategy and Eligibility

Convenience sampling was used to recruit eligible participants who met the following criteria: (1) older adults aged 50 to 95 years, (2) participated in the PFA program, and (3) spoke and read English.

PFA Program

The PFA program was developed based on foundational literature [29-31] and insights from a multidisciplinary research team. This team comprised psychiatrists, psychologists, nurses, and peer volunteers with extensive experience supporting community-dwelling older adults. Their prior primary and secondary research studies formed the basis of the PFA intervention design.

The program consisted of 4 half-day workshops, each lasting 4 hours, and conducted face-to-face over consecutive weekends. The workshops were held in a conducive community space and addressed a range of topics, including common mental health disorders in Singapore, psychosocial needs of older adults and peer volunteers, skills in empathetic listening, low-intensity psychological support, mindfulness training, identifying individuals at risk, and strategies for facilitating connections to professional care.

Workshops were delivered by a diverse team of trainers, including psychiatrists, nurses, experienced peer volunteers, and mindfulness instructors. Peer volunteers involved in earlier training sessions also contributed as facilitators, adding a practical perspective to the training content. The primary objective of the PFA program was to enable early detection of mental health issues among community-dwelling older adults. By equipping volunteers with these skills, the program also aimed to enhance the resilience and coping strategies of both the older adults and the volunteers themselves.

Ethical Considerations

The study received ethical approval from the institutional review board of the National University of Singapore (NUS-IRB-2024 - 506), which serves as the ethics board for the participating institution. Prior to recruitment, participants were provided with detailed information sheets outlining the study's objectives and their roles. Written informed consent was obtained, and voluntary participation was emphasized. To ensure anonymity and confidentiality, each participant's data were coded using unique serial numbers ranging from 1 to 13.

Data Collection

Data collection for the study was conducted between September and October 2024. All 20 peer volunteers who participated in the first batch of the PFA program were invited to join the study. Invitations were extended by a study administrator independent of the research team. Volunteers who expressed interest were introduced to the research team. A male research assistant (independent of the PFA program and participants) explained the study details to the interested participants, provided the participant information sheet, and obtained written informed consent. Subsequently, one-on-one interviews were arranged. Participants could choose between web-based or in-person interviews based on their preference. Before the interview, sociodemographic data such as age, gender, citizenship, ethnicity, religion, marital status, number of children, education level, employment status, and volunteering duration were collected.

A male research assistant (JS) with a BSc (Hons) in Nursing and no dependent relationship with the participants, developed the semistructured interview guide guided by literature review [32,33] and conducted the interviews to ensure consistency. Additionally, the same research assistant recorded field notes during each session. The guide included open-ended questions focusing on participants' experiences with the PFA program and their volunteerism journey (eg, "Would you like to share any personal experiences or stories that inspired you to join this PFA programme?" and "How was your experience with the PFA programme?"). All interviews were conducted via Zoom (Zoom Video Communications), and consent was obtained from participants for audio and video recording. Transcripts were transcribed verbatim from the recordings. To ensure rigor, the

XSL•FO
RenderX

research assistant (JS) received qualitative research training under the guidance of a PhD-trained researcher (SS) with extensive qualitative experience. A pilot interview was conducted with one peer volunteer to evaluate the interview guide's flow and relevance, aligning it with the research objectives; no modifications were needed, and the data from this pilot interview was included in the analysis as it contributed valuable insights to the study. A total of 13 interviews were conducted with no attrition, with durations ranging from 36 to 101 minutes (mean of 59.5 min). Data saturation was reached after the 11th interview, as no new concepts emerged [34]. Two additional interviews were conducted to confirm saturation. The interview guide is provided in Table S2 in Multimedia Appendix 1.

Data Analysis

The data were analyzed using Braun and Clarke's [35] 6-step framework for inductive thematic analysis. The process began with the verbatim transcripts being imported into Microsoft Word for organization and management. Two researchers (JS and SS) independently familiarized themselves with the data by reading and rereading the transcripts to gain a thorough understanding. Initial codes were generated by identifying meaningful units of information and relevant concepts within the data. These codes were systematically applied across all transcripts. Through an iterative process, the researchers grouped the codes into broader categories to develop initial themes and subthemes. These themes reflected the key patterns and narratives emerging from the data (Table S3 in Multimedia Appendix 1). The coding and thematic development were continuously refined through team discussions to ensure accuracy and depth. Any discrepancies in interpretation were resolved through consensus, ensuring a comprehensive and cohesive representation of the findings.

Rigor

To ensure the rigor of this study's findings, credibility, transferability, dependability, and confirmability were prioritized, as outlined by Cypress [36]. Credibility was maintained through prolonged engagement with participants during data collection and investigator triangulation during data analysis. This allowed for multiple perspectives to be considered and verified against the data. Transferability was established by providing thick, rich descriptions of participants' experiences with the PFA program, supported by direct quotes. This approach enables readers to evaluate the applicability of the findings to other contexts. Dependability was reinforced by creating an audit trail, which included comprehensive documentation of interview transcripts, coding, and the analytical process. This ensures that the study can be independently reviewed and verified. Confirmability was achieved using a reflexive journal throughout the research process. Reflexivity, as highlighted by Teh and Lek [37], is a critical aspect of qualitative research that minimizes the influence of unconscious researcher biases during data interpretation. By maintaining this journal, the researchers ensured their reflections and decisions were transparent, bolstering the trustworthiness of the study [38].

Results

Overview

A total of 13 community-dwelling older adults participated in this study, and their sociodemographic characteristics are presented in Table 1. The mean age of the participants was 61.2 (SD 6.9) years. Most participants were female (9/13, 69%) and married (9/13, 69%), with all of them were Singaporeans (13/13, 100%). The group was predominantly composed of Malay (6/13, 46%) and Chinese (5/13, 38%) ethnic backgrounds. The participants' employment status varied, with being retired (4/13, 31%), employed full-time (4/13, 31%), employed part-time (2/13, 15%), unemployed (2/13, 15%), and self-employed (1/13, 8%). The group had a balanced distribution of volunteerism experience, with 54% (7/13) having up to 1 year of experience, and 46% (6/13) having more than 1 year of experience.

Three themes were identified from the data analysis, highlighting older adults' experiences of the PFA program and the value it added to their volunteerism journey: (1) dimensions of volunteerism from motivations to hurdles, (2) empowerment through collaborative learning, and (3) recommendations for designing inclusive, holistic training programs. These themes were further supported by 9 subthemes (Figure 1).

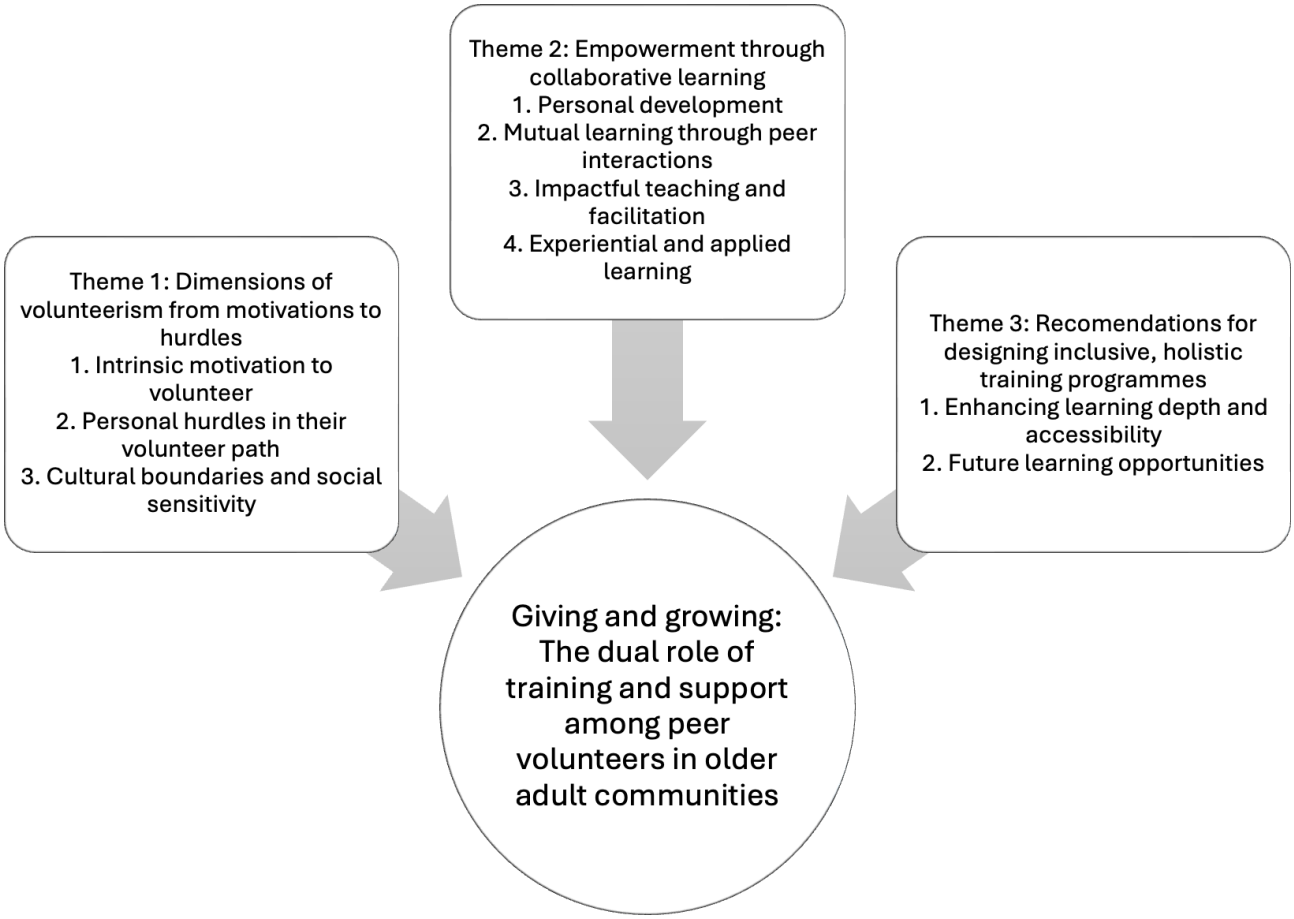
XSL•FO
RenderX

Table . Participants' sociodemographic characteristics (n=13).

Sociodemographic variables	Values
Age (years), mean (SD)	61.2 (6.9)
Sex, n (%)	
Female	9 (69)
Male	4 (31)
Nationality, n (%)	
Singaporean	13 (100)
Ethnicity, n (%)	
Chinese	5 (38)
Malay	6 (46)
Indian	1 (8)
Javanese	1 (8)
Marital status, n (%)	
Married	9 (69)
Single	3 (23)
Divorced	1 (8)
Length of marriage (years), n (%)	
0 - 10	4 (31)
11 - 20	1 (8)
21 - 30	2 (15)
31 - 40	3 (23)
>41	3 (23)
Highest education level, n (%)	
Secondary school	6 (46)
Polytechnic	2 (15)
University	5 (38)
Employment status, n (%)	
Employed full-time	4 (31)
Employed part-time	2 (15)
Self-employed	1 (8)
Unemployed	2 (15)
Retired	4 (31)
Volunteered as a peer volunteer (years), n (%)	
Up to 1	7 (54)
More than 1	6 (46)

XSL•FO
RenderX

Figure 1. Older adults' experience on the Psychological First Aid program and personal volunteering experience.



Theme 1: Dimensions of Volunteerism From Motivations to Hurdles

Overview

This theme highlighted older adults' experiences from motivations to hurdles, which shaped both their initial decision to volunteer and their continued commitment. These were categorized into 3 subthemes: intrinsic motivation to volunteer, personal hurdles in their volunteer path, and cultural boundaries and social sensitivity.

Intrinsic Motivation to Volunteer

Most older adults shared their personal motivations for volunteerism. As the majority of older adults were either employed part-time or retirees, they had free time available. Hence, they expressed a strong desire to actively engage with the community to enhance their personal well-being by giving back to society and fostering friendships. Many older adults also shared the perceived benefits of volunteerism, from keeping their minds active and serving as a protection to prevent the onset of dementia.

> *I volunteer as It's to help myself to prevent having dementia because I'm learning something new...* [K3, Female peer volunteer]

> *I volunteer as It's to help myself to prevent having dementia because I'm learning something new...* [K3, Female peer volunteer]

Some older adults attributed their volunteerism motivations to the cultural values instilled by their families. They were raised by their parents to be benevolent and gracious, which fostered a strong desire to give back to society. As a result, the concept of volunteerism became deeply embedded in their upbringing.

> *I see my father as a figure who do all this volunteer work in his life.… so it's already in my blood* [K5, Female peer volunteer]

Older adults attributed their decision to join the PFA program to the influence of their peers, highlighting the positive impact of fostering a sense of purpose and building meaningful friendships. For instance, one participant, who became a single mother during her teenage years, recounted the emotional struggles she faced, such as limited support from her family and peers, which negatively affected her emotional health. Having developed emotional resilience through overcoming these challenges, she felt inspired to volunteer and provide emotional support to others navigating similar difficulties.

> *To help people because I had gone through a period of time [her teenage years] and from that time, it makes me want to touch more and especially on single moms* [K13, Female peer volunteer]

Older adults shared that establishing meaningful engagement with their community residents was a significant challenge. Many encountered resistance when attempting to approach and offer support. Observing the reluctance of fellow older adults to seek help underscored for them the importance of acquiring

XSL•FO
RenderX

proper training. They recognized that such training was crucial to effectively support and connect with others in their communities.

> *I think it's the resistance of the people like when we try to talk or whatever, there is certain resistance… that led me to learn more from this programme* [K11, Female peer volunteer]

Personal Hurdles in Their Volunteer Path

Older adults highlighted the struggle of physical limitations, such as the lack of strength, despite being motivated to attend training courses to support others. This is an important consideration, as most community residents are older adults who may need assistance moving from one place to another. One participant shared that, due to her age and physical build, she is not as strong as she was in her youth and therefore needs to be more cautious when interacting with the older community residents or attending any courses.

> *I don't have much strength, things are very physical. I would not be able to help because … we may need to provide them physical support, like to support them so that they don't fall down...maybe emotional support would do better [for me]…* [K3, Female peer volunteer]

In addition, one participant shared that he was diagnosed with aphasia. This gravely impacted his communication abilities with both his peers and community residents. As a result, he struggled to express his needs efficiently.

> *Sometimes my train of thoughts go out... sometimes when people talk to me, I can't interpret properly… that sometime dampen me from opening up…* [K10, Male peer volunteer]

Cultural Boundaries and Social Sensitivity

Many older adults identified cultural perceptions as a significant barrier in their volunteerism experiences. They faced cultural taboos, such as Muslims' discomfort with certain terms used around older adults, including "forgetful" and concerns about interacting too much with the opposite gender. As a result, the volunteers, especially those of different ethnicities, had to be mindful of their behavior when approaching the community-dwelling older adults. Additionally, some volunteers highlighted that Muslim residents in the community often exhibit caution towards "outsiders" and show a preference for staying within their own social and ethnic circles, which limits their willingness to engage with others. These cultural dynamics present challenges in fostering broader social connections and offering support.

> *For the Malay elderly how we want to get them, how are we going to approach them to ask them that, we can really give them help. That's our challenge now.* [K13, Female peer volunteer]

Gender dynamics were frequently discussed by older adults, who reported that supporting members of the opposite gender was often challenging and required careful consideration, particularly regarding body language and posture. Many volunteers found it easier to assist individuals of the same gender, such as a female peer volunteer supporting a female older adult, as they could more easily interpret nonverbal cues like body language and gestures. One female peer volunteer shared her discomfort when a male community resident exhibited inappropriate behavior, such as touching without consent. These gender-related challenges highlight the complexities involved in volunteer interactions and the need for sensitive handling of such situations in community support roles.

> *Not easy to understand a male, what he's thinking. A female, when you look into her eyes, you talk to her, you can understand more… not easy with males….* [K1, Female peer volunteer]

Many felt the need for a chaperone, especially when interacting with the opposite gender, as a safety measure and to prevent unnecessary conflict. In one instance, one female peer volunteer experienced negative experiences from a wife due to her close rapport with the community resident.

> *There is a gender difference [in our local setting] that we need to be more mindful. For female, we can touch the female, can consult but for male, we need to be more careful, more mindful on our gesture.* [K11, Female peer volunteer]

Theme 2: Empowerment Through Collaborative Learning

Overview

This theme highlighted elements that contributed to the empowerment of older adults through collaborative learning during the PFA program. These were supported by 4 subthemes as follows: personal development, mutual learning through peer interactions, impactful teaching and facilitation, and experiential and applied learning.

Personal Development

Older adults highlighted the benefits of participating in the PFA program. Many expressed how the knowledge acquired through the PFA program helped them in real-life situations, enabling them to better showcase their skills as peer volunteers. Topics such as grief and mindfulness were essential in their role as peer volunteers. This, in turn, fostered a sense of empowerment and personal growth among the older adults.

> *It Was Nice, It Was Very Knowledgeable. I Must Say We Learn a Lot* [K1, Female Peer Volunteer]

Additionally, the older adults reported maintaining a positive attitude toward learning. They expressed that learning is a continuous journey, and the knowledge they gained expanded their viewpoints.

> *The… workshop where it really broadened our mind and then we learn more about things [eg how to support others better].* [K2, Male peer volunteer]

Mutual Learning Through Peer Interactions

Older adults emphasized that the PFA program highlighted the importance of maintaining open and effective communication among peers. Through interactions during class and breaks,

XSL•FO
RenderX

they had the opportunity to share personal experiences, which many participants found invaluable. This peer exchange allowed them to learn from each other's experiences and mentally prepare for similar real-life situations in the future. The older adults also expressed that the learning process was reciprocal, with everyone eager to contribute and learn from one another.

> *You get different ideas from them [peers attending PFA programme]. You learn more things from them, they learn from you.* [K1, Male peer volunteer]

Additionally, some older adults shared that the class, which included individuals from various ethnic backgrounds such as Chinese, Malay, and Indian, provided a valuable opportunity to engage with and learn about different cultural practices. This allowed them to conceptualize and compare the differences or similarities between the cultures.

> *The other group are Malays so then cultural wise, it's quite different. I think to understand each other's cultural difference… that was helpful* [K3, Female peer volunteer]

Impactful Teaching and Facilitation

Dedicated facilitators consistently worked to build rapport with the older adults, helping to ease initial discomfort and encouraging open communication. Older adults highlighted that the facilitators were highly knowledgeable in their respective fields. They were committed to sharing their knowledge on various topics such as handling difficult situations and mindfulness, which made the older adults feel comfortable asking questions whenever they had any doubts. With that, this fostered communication and the older adults were able to maximize the knowledge gained throughout this PFA program.

> *Think the way they [facilitators] present and they deliver the topic well, yeah and they are very knowledgeable on the topic… you feel motivated.* [K11, Female peer volunteer]

Additionally, many older adults highlighted that they were inspired by the facilitators' presentation style. The facilitators' consistent engagement, well-paced sessions, body language, tone, and delivery made the learning process more accessible and understandable for older learners. The older adults also valued the facilitators' sharing of personal experiences, as it allowed them to learn from real-life situations.

> *I learn from it and [blinded] because she's a mindfulness teacher and I'm learning to teach mindfulness. So of course I also want to learn all the right things from her.* [K3, Female peer volunteer]

Experiential and Applied Learning

Many older adults favored experiential learning over traditional lecture-style delivery. During the PFA program, they were able to actively participate in hands-on activities such as role-plays and this helped them to better visualize and retain the concepts taught.

> *The role play, it showed you like reality what could happen or what will happen or what can happen. That was really like an experience, very good because learning from slides and all is quite normal.* [K1, Female peer volunteer]

Additionally, participants reported that the role-playing exercises played a crucial role in boosting their confidence in handling real-life situations. By simulating real-world scenarios, they were able to "experience" these situations in a controlled environment. This allowed them to observe nonverbal cues and body language, which heightened their awareness of their own nonverbal communication when offering support. The interactive nature of the role-playing activities also facilitated communication among the older adults, helping to break the ice and foster connections.

> *When you have a role play, you really understand it more. You sometimes put yourself into that shoes and see how it is like.* [K1, Female peer volunteer]

Theme 3: Recommendations for Designing Inclusive, Holistic Training Programs

Overview

This theme highlighted recommendations to enhance the PFA program by addressing the varied needs, backgrounds, abilities, and learning styles of the participants. This theme is supported by 2 subthemes as follows: depth and quality of learning, and future learning opportunities.

Depth and Quality of Learning

A few older adults expressed that the content covered was too brief. They desired more in-depth knowledge to be taught, enabling them to gain additional insights that could be applied throughout their volunteerism journey. However, they noted that the brief content might be a result of the short session duration. Therefore, they expressed a preference for longer sessions.

> *I think it can go a bit deeper. It's just that the depth can be, in my personal opinion, if time allowed, the depth can be more* [K9, Female peer volunteer]

Furthermore, the older adults highlighted that the program venue was inconvenient. Some had to get up as early as 7 in the morning to arrive on time for the session. Despite the travel challenges, they expressed a desire for longer sessions to facilitate deeper learning.

> *I find it's really short because we [travel] from the east [and] we spend only two hours… the timing is not enough to share with us.* [K04, Female Peer Volunteer]

Many older adults envisage more hands-on practice during the PFA program. Though older adults caught up during breaks and role-plays, having a specific time for small-group discussions regularly was highlighted by some older adults.

> *I was thinking maybe it's good to have maybe small group discussion during each session.* [K11, Female peer volunteer]

A few older adults suggested incorporating various learning tools, such as case studies and better audio-visual materials for learning. Those older adults with less volunteer experience felt

XSL•FO
RenderX

the need for more case studies to be discussed so that they are more prepared to handle the situations in real life.

> *The other techniques like case study so examples of how we can approach the people… I never counsel anyone so I wouldn't know, what [is] the real situation outside... I am bit worried to handle if someone suddenly cry…* [K3, Female peer volunteer]

In addition, older adults highlighted several considerations that influenced their learning experience when engaging with older learners. They encountered issues with audio and visual clarity, such as some participants finding the projected fonts too small, requiring them to squint to read. Furthermore, the facilitators' lack of use of a microphone led to poor hearing, making it difficult for a few participants to fully comprehend the content taught.

> *As we age, our eyesight is not so good, we have to strain our eyes. I had to strain my eyes to read [the slides].* [K1, Female peer volunteer]

> *If you're seated behind far away, you can't hear much. I'm trying to figure out what he's saying. So if you have a mic, then at least, you know, we can hear, everybody can hear.* [K4, Female peer volunteer]

Future Learning Opportunities

Older adults highlighted the potential for future learning opportunities. Many expressed a desire for regular "top-up" sessions, either annually or semiannually, to revise the knowledge. Furthermore, they showed interest in gaining new knowledge through these regular sessions. A few older adults stressed the importance of regular sessions so that more peer volunteers from their respective organizations could participate and benefit from the PFA program.

> *It will be good if there's every six months or a year, we come back do skill upgrading.* [K9, Male peer volunteer]

> *I think we need to have more of this courses, for more people to be trained because now Singapore is aging.* [K7, Male peer volunteer]

In addition, they valued the regular sessions as an opportunity to reconnect with their peers. They were eager to meet and share their recent volunteerism experiences, fostering mutual learning through casual conversations.

> *A small group, then we will have regular meetups for sharing and to also impart certain message.* [K11, Female peer volunteer]

Discussion

Principal Findings

This study explored the experiences of peer volunteers in the PFA program in Singapore, providing insights into their motivations, challenges, and the overall impact of their participation. All interviewed peer volunteers participated in all 4 sessions of the program. The sample was primarily composed of Malay participants (n=6), followed by Chinese participants (n=5), with one participant from each of the Indian and Javanese communities. While Singapore's population is predominantly Chinese, Malay, and Indian, future studies could aim for a more balanced racial distribution to better reflect broader societal trends [39].

The older adults in this study cited various reasons for their continued involvement in volunteerism, including cultural values instilled by their families, peer influence, and the promotion of healthy aging. Through volunteerism, the older adults in this study were also able to actively engage with the community, fostering new friendships and a sense of purpose. This aligns with existing literature, with studies conducted in Japan, Australia, and the United Kingdom suggesting that the motivations for volunteerism in older adulthood were driven by community engagement, building new friendships, positive influences on personal well-being, and the desire to foster a sense of purpose [40-42]. Interestingly, altruistic attitudes, such as cultural values of graciousness instilled in them, were a key motivation for the older adults in our study. This contrasts with existing literature, where volunteers report motivations driven by egoistic reasons such as the development of a new skill and personal satisfaction [43,44]. This difference may be attributed to the strong Asian values that emphasize family cohesion and collective well-being [45]. As such, future studies can compare and explore volunteer motivations across different cultural contexts.

While the older adults in this study demonstrated a strong commitment to volunteerism, they also faced personal barriers, including difficulties in engaging with the community and physical limitations such as ongoing medical conditions or frailty. Chronic illnesses or frailty restrict their ability to participate in physically demanding tasks, limiting their ability to contribute as fully as they may have wished. This is consistent with studies conducted on older adults in Spain and Norway, which found that maintaining good health was a key factor enabling volunteerism as it was highly prioritized by older adults [46,47]. These challenges highlight the need for support systems that prioritize older adults' health to ensure their ongoing participation in volunteer activities, ultimately improving both their physical well-being and their ability to contribute meaningfully to the community.

Cultural and social sensitivity emerged as another key theme in this study. Older adults encountered cultural taboos, particularly within the Muslim community, where there was a preference for interactions within one's own ethnic or social circle. This limited broader community engagement, highlighting the importance of cultural sensitivity in volunteer programs. Additionally, gender dynamics were cited as a challenge, with female participants feeling more comfortable providing support to those of the same gender. These findings are consistent with research indicating that cultural and gender factors can influence interpersonal dynamics in volunteer settings [48]. This may be attributed to Singapore's diverse ethnic composition, where cultural norms are critical. Furthermore, in the Asian context, where gender dynamics are often less openly discussed, it is possible that the older adults in this study feel more at ease interacting with individuals of the same gender [49]. This suggests the importance of addressing these sensitivities in designing inclusive volunteer

XSL•FO
RenderX

programs that accommodate diverse cultural and gender preferences. Future research could explore how these factors shape volunteer experiences and inform the development of more culturally inclusive training programs.

Personal development emerged as another significant theme, with participants noting the value of peer interactions, role-plays, and experiential learning during the PFA program. These activities not only reinforced the content but also helped participants apply the concepts to real-world situations. This contrasts with a study of clinicians in the United States, where participants expressed a preference against interactive training methods [50]. The difference may be attributed to the older adult population's preference for active, hands-on learning approaches, which has been noted in other studies [51]. Facilitators also played a crucial role, providing reassurance and fostering a supportive learning environment. This finding is consistent with research indicating that facilitators' knowledge and support are integral to positive learning experiences [52,53]. Participants in this study also expressed a desire for more knowledge and confidence to handle real-life situations, aligning with findings from Japan and Haiti, where peer volunteer training boosted confidence in applying training strategies safely [54,55]. This highlights the importance of designing volunteer training programs that incorporate both theoretical knowledge and practical, experiential learning.

Designing inclusive training programs involves a holistic approach that considers accessibility, logistical support, practical learning experiences, and efficient course delivery, all of which contribute to enriched learning experiences and future improvement opportunities. Our study highlighted that the participants in our study desired greater comprehensive coverage of the content, with additional hands-on practice and small group discussions to enhance the training experience and add greater value to their volunteerism journey. Additionally, they expressed interest in incorporating other educational tools, such as case studies, to better assimilate real-world scenarios and mentally prepare themselves. This is consistent with studies conducted in Gaza, Korea, and Singapore, which emphasized that PFA trainers desire more in-depth training and new delivery approaches to psychologically equip themselves, enabling them to better support the recipients [25,52,56]. Additionally, older adults in our study emphasized the importance of regular sessions to stay informed about current information about PFA. Moreover, our study found that holding regular sessions would facilitate enrollment for prospective participants, providing them with comprehensive knowledge. These findings align with studies conducted in the United States and the United Kingdom, where participants envisage refresher sessions for PFA to further equip themselves with the latest knowledge [32,50]. Furthermore, several participants in our study, drawing on their prior volunteerism experience, expressed a desire for advanced volunteerism courses offering enhanced learning opportunities. This aligns with findings from studies conducted in the United States and Singapore, which highlight that individuals with volunteerism experience gain greater benefits from comprehensive training programs compared with basic ones [50,52]. Thus, a holistic approach to program design, which considers accessibility, practical learning opportunities, and continuous support, can enhance the overall success and impact of volunteer programs, and stakeholders can consider these recommendations in providing better support to the peer volunteers and, in turn, to the community-dwelling older adults.

Limitations

This study has a few limitations. First, this study was conducted on English-speaking older adults and may not be transferable to non–English-speaking older adults in Singapore. Furthermore, Indians were underrepresented in this study. Hence, future studies can consider exploring the purposive sampling technique to study participants who are non–English-speaking and from diverse ethnic backgrounds to get a holistic understanding of the phenomenon of interest. Second, interviews were conducted with a convenience sample of older adults who volunteered to participate and hence could have harbored particularly positive views from those participants who were motivated to share about the PFA program. Third, a potential limitation is that participants were not explicitly asked to provide feedback on the interview guide during the pilot interviews, specifically on aspects such as question clarity, sequencing, and overall structure. Such feedback could have offered valuable insights for refining the guide and enhancing its suitability. This should be considered in future research.

Implications for Future Research and Practice

Future research should explore the influence of cultural norms and gender dynamics in volunteer programs, particularly in diverse and multiethnic contexts like Singapore. Studies can investigate how cultural and gender considerations affect volunteer engagement and explore strategies for overcoming these barriers. Future studies should focus on the role of cultural sensitivity training and its impact on improving interactions among volunteers and community members. The physical limitations of older adults, including chronic illnesses or frailty, were identified as significant barriers to volunteerism. Future research should delve deeper into how health status influences volunteer participation and explore ways to accommodate volunteers with varying levels of physical ability. Moreover, longitudinal studies could assess the long-term health benefits of volunteerism and how these programs might mitigate the physical limitations of older adults. Stakeholders working with older adults, such as community partners and health care providers, could benefit from considering the cultural and age-related concerns highlighted in this study. Understanding the unique cultural values, gender dynamics, and health challenges faced by older adults is crucial for fostering more effective engagement and providing tailored support. This study highlighted the importance of hands-on learning, role-playing, and peer interactions in enhancing volunteers' understanding and confidence. Future research should focus on evaluating the effectiveness of different interactive learning methods across various populations. Additionally, exploring the impact of peer-led sessions versus professional-led training could help refine approaches to fostering confidence and competency in volunteers. Research into the most effective formats for delivering such programs—whether through in-person, via web, or hybrid methods—could contribute valuable knowledge to improve volunteer retention and engagement. This study

XSL•FO
RenderX

emphasizes that individuals with prior volunteerism experience benefit from a more comprehensive training curriculum. Hence, future research can explore the potential of offering both basic and advanced PFA programs for learners at different proficiency levels to enhance individual learning outcomes. While this study focused on short-term engagement with the PFA program, future research should explore the long-term effects of participation on older adults' well-being, volunteer retention, and community impact.

Conclusions

This study explored the experiences of older adult volunteers who participated in a PFA program in Singapore. Findings highlighted the complex interplay of cultural, social, and personal factors that influence the experiences of older adult volunteers. The findings suggest that training programs could be designed with cultural sensitivity, provide experiential learning opportunities, and support participants' ongoing personal development to maximize their engagement and impact. Future research should further explore these findings to develop more inclusive, effective, and sustainable volunteer programs that empower older adults to contribute meaningfully to their communities.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional material.
[DOCX File, 61 KB - aging_v8i1e71810_app1.docx]

References

1. Population in brief. National Population and Talent Division. URL: https://www.population.gov.sg/media-centre/articles/population-in-brief-2023-key-trends/ [accessed 2024-11-22]
2. Mental health of older adults. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/mental-health-of-older-adults [accessed 2024-12-02]
3. Institute of Mental Health. Well-being of the Singapore Elderly (WiSE) 2021 Study. URL: https://corp.nhg.com.sg/Media%20Releases/[illegible].pdf [accessed 2025-03-19]
4. Desai R, Tsipa A, Fearn C, et al. Suicide and dementia: a systematic review and meta-analysis of prevalence and risk factors. Ageing Res Rev 2024 Sep;100:102445. [doi: 10.1016/j.arr.2024.102445] [Medline: 39127443]
5. Low LL, Kwan YH, Ko MSM, et al. Epidemiologic characteristics of multimorbidity and sociodemographic factors associated with multimorbidity in a rapidly aging Asian country. JAMA Netw Open 2019 Nov 1;2(11):e1915245. [doi: 10.1001/jamanetworkopen.2019.15245] [Medline: 31722030]
6. Oh DJ, Han JW, Bae JB, et al. Chronic subsyndromal depression and risk of dementia in older adults. Aust N Z J Psychiatry 2021 Aug;55(8):809-816. [doi: 10.1177/0004867420972763] [Medline: 33198490]
7. Teo RH, Cheng WH, Cheng LJ, Lau Y, Lau ST. Global prevalence of social isolation among community-dwelling older adults: a systematic review and meta-analysis. Arch Gerontol Geriatr 2023 Apr;107:104904. [doi: 10.1016/j.archger.2022.104904] [Medline: 36563614]
8. Lutz J, Van Orden KA. Sadness and worry in older adults: differentiating psychiatric illness from normative distress. Med Clin North Am 2020 Sep;104(5):843-854. [doi: 10.1016/j.mcna.2020.05.001] [Medline: 32773049]
9. National population health survey. Ministry of Health. URL: https://www.moh.gov.sg/others/resources-and-statistics/nphs-2023 [accessed 2024-12-02]
10. Volunteer with older people. Age UK. URL: https://www.ageuk.org.uk/get-involved/volunteer/ [accessed 2024-12-11]
11. Bringing out the best of America. AmeriCorps. URL: https://americorps.gov/about/our-impact [accessed 2024-12-11]
12. Kim ES, Whillans AV, Lee MT, Chen Y, VanderWeele TJ. Volunteering and subsequent health and well-being in older adults: an outcome-wide longitudinal approach. Am J Prev Med 2020 Aug;59(2):176-186. [doi: 10.1016/j.amepre.2020.03.004] [Medline: 32536452]
13. Volunteer opportunities. Lions Befrienders Service Association. URL: https://www.lionsbefrienders.org.sg/volunteer-opportunities/ [accessed 2024-12-11]
14. Morrow-Howell N, Hinterlong J, Rozario PA, Tang F. Effects of volunteering on the well-being of older adults. J Gerontol 2003 May;58(3):S137-S145. [doi: 10.1093/geronb/58.3.S137]
15. Singapore cares - building a caring and inclusive society. SG Cares. URL: https://www.sg/sgcares [accessed 2024-12-11]
16. Pinto da Costa M, Aviron P. Comparing the views of mental health professionals and volunteers on volunteering in mental health care in the UK. Discov Health Systems 2023 Dec 5;2(1):46. [doi: 10.1007/s44250-023-00056-4]

XSL•FO
RenderX

17. Pinto da Costa M, Oliveira J. Views on volunteering in mental health: a focus group study with mental health professionals and volunteers in Portugal. Discov Ment Health 2023 Jun 22;3(1):12. [doi: 10.1007/s44192-023-00038-1] [Medline: 37861821]
18. Brymer M, Jacobs A, Layne C, et al. Psychological First Aid: Field Operations Guide, 2nd edition: National Child Traumatic Stress Network and National Center for PTSD; 2006.
19. Vernberg EM, Steinberg AM, Jacobs AK, et al. Innovations in disaster mental health: psychological first aid. Prof Psychol: Res Pract 2008;39(4):381-388. [doi: 10.1037/a0012663]
20. Kelly EL, Fenwick KM, Brekke JS, Novaco RW. Sources of social support after patient assault as related to staff well-being. J Interpers Violence 2021 Jan;36(1-2):NP1003-NP1028. [doi: 10.1177/0886260517738779] [Medline: 29294965]
21. Pratt SD, Jachna BR. Care of the clinician after an adverse event. Int J Obstet Anesth 2015 Feb;24(1):54-63. [doi: 10.1016/j.ijoa.2014.10.001] [Medline: 25499810]
22. Wang L, Norman I, Xiao T, Li Y, Leamy M. Psychological first aid training: a scoping review of its application, outcomes and implementation. Int J Environ Res Public Health 2021 Apr 26;18(9):4594. [doi: 10.3390/ijerph18094594] [Medline: 33926108]
23. Bakes-Denman L, Mansfield Y, Meehan T. Supporting mental health staff following exposure to occupational violence - staff perceptions of “peer” support. Int J Ment Health Nurs 2021 Feb;30(1):158-166. [doi: 10.1111/inm.12767] [Medline: 33241642]
24. de Freitas Girardi J, Miconi D, Lyke C, Rousseau C. Creative expression workshops as Psychological First Aid (PFA) for asylum-seeking children: an exploratory study in temporary shelters in Montreal. Clin Child Psychol Psychiatry 2020 Apr;25(2):483-493. [doi: 10.1177/1359104519891760] [Medline: 31814432]
25. Schafer A, Snider L, Sammour R. A reflective learning report about the implementation and impacts of Psychological First Aid (PFA) in Gaza. Disaster Health 2016;3(1):1-10. [doi: 10.1080/21665044.2015.1110292] [Medline: 28229011]
26. Hartanto A, Sim L, Lee D, Majeed NM, Yong JC. Cultural contexts differentially shape parents' loneliness and wellbeing during the empty nest period. Commun Psychol 2024 Nov 3;2(1):105. [doi: 10.1038/s44271-024-00156-8] [Medline: 39489810]
27. Doyle L, McCabe C, Keogh B, Brady A, McCann M. An overview of the qualitative descriptive design within nursing research. J Res Nurs 2020 Aug;25(5):443-455. [doi: 10.1177/1744987119880234] [Medline: 34394658]
28. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
29. Launch of Psychological First Aid (PFA) training for SGsecure responders. Ministry of Home Affairs. URL: https://www.mha.gov.sg/mediaroom/speeches/launch-of-psychological-first-aid-training-for-sgsecure-responders-closing-remarks-by-ms-sun-xueling-minister-of-state-ministry-of-home-affairs-ministry-of-social-and-family-development/ [accessed 2024-11-25]
30. Ng TKS, Feng L, Fam J, et al. Mindfulness Awareness Practice (MAP) to prevent dementia in older adults with mild cognitive impairment: protocol of a randomized controlled trial and implementation outcomes. Int J Environ Res Public Health 2021 Sep 28;18(19):10205. [doi: 10.3390/ijerph181910205] [Medline: 34639513]
31. Ruzek JI, Brymer MJ, Jacobs AK, Layne CM, Vernberg EM, Watson PJ. Psychological first aid. J Ment Health Couns 2007 Jan 1;29(1):17-49. [doi: 10.17744/mehc.29.1.5racqxjueafabgwp]
32. Cook A, Keyte R, Sprawson I, Matharu A, Mantzios M. Mental health first aid experiences: a qualitative investigation into the emotional impact of mental health first aid responsibilities and the significance of self-compassion. SN Soc Sci 2024 Nov 9;4(9):171. [doi: 10.1007/s43545-024-00962-1]
33. Horn R, O'May F, Esliker R, et al. The myth of the 1-day training: the effectiveness of psychosocial support capacity-building during the Ebola outbreak in West Africa. Glob Ment Health (Camb) 2019;6:e5. [doi: 10.1017/gmh.2019.2] [Medline: 31143466]
34. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907. [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
35. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
36. Cypress BS. Rigor or reliability and validity in qualitative research: perspectives, strategies, reconceptualization, and recommendations. Dimens Crit Care Nurs 2017;36(4):253-263. [doi: 10.1097/DCC.0000000000000253] [Medline: 28570380]
37. Teh YY, Lek E. Culture and reflexivity: systemic journeys with a British Chinese family. J Fam Ther 2018 Nov;40(4):520-536. [doi: 10.1111/1467-6427.12205]
38. Buetow S. Apophenia, unconscious bias and reflexivity in nursing qualitative research. Int J Nurs Stud 2019 Jan;89:8-13. [doi: 10.1016/j.ijnurstu.2018.09.013] [Medline: 30316055]
39. Rocha ZL, Yeoh BSA. Measuring race, mixed race, and multiracialism in singapore. In: Rocha ZL, Aspinall PJ, editors. The Palgrave International Handbook of Mixed Racial and Ethnic Classification: Springer International Publishing; 2020:629-647.

XSL•FO
RenderX

40. Hotta K, Ishimaru M. The meaning in group volunteering and its influence on the well-being of community-dwelling older volunteers in Japan: a qualitative study. Arch Gerontol Geriatr Plus 2024 Dec;1(4):100086. [doi: 10.1016/j.aggp.2024.100086]
41. Same A, McBride H, Liddelow C, Mullan B, Harris C. Motivations for volunteering time with older adults: a qualitative study. PLOS ONE 2020;15(5):e0232718. [doi: 10.1371/journal.pone.0232718] [Medline: 32365081]
42. Withall J, Thompson JL, Fox KR, et al. Participant and public involvement in refining a peer-volunteering active aging intervention: project ACE (active, connected, engaged). Gerontologist 2018 Mar 19;58(2):362-375. [doi: 10.1093/geront/gnw148] [Medline: 27927733]
43. Barlow J, Hainsworth J. Volunteerism among older people with arthritis. Ageing Soc 2001 Mar;21(2):203-217. [doi: 10.1017/S0144686X01008145]
44. Okun MA, Michel J. Sense of community and being a volunteer among the young-old. J Appl Gerontol 2006 Apr;25(2):173-188. [doi: 10.1177/0733464806286710]
45. Welzel C. The Asian values thesis revisited: evidence from the world values surveys. Jpn J Political Sci 2011 Apr;12(1):1-31. [doi: 10.1017/S1468109910000277]
46. Hansen T, Slagsvold B. An “Army of Volunteers”? engagement, motivation, and barriers to volunteering among the baby boomers. J Gerontol Soc Work 2020;63(4):335-353. [doi: 10.1080/01634372.2020.1758269] [Medline: 32364438]
47. Serrat-Graboleda E, González-Carrasco M, Casas Aznar F, Malo Cerrato S, Cámara Liebana D, Roqueta-Vall-Llosera M. Factors favoring and hindering volunteering by older adults and their relationship with subjective well-being: a mixed-method approach. Int J Environ Res Public Health 2021 Jun 22;18(13):6704. [doi: 10.3390/ijerph18136704] [Medline: 34206359]
48. Best DL, Puzio AR. Gender and culture. In: Matsumoto D, Hwang HC, editors. The Handbook of Culture and Psychology: Oxford Academic; 2019:235-291. [doi: 10.1093/oso/9780190679743.003.0009]
49. Seow D. The role of gender, patriarchy and culture in the asian female travel experience. : Bournemouth University; 2020.
50. Hambrick EP, Rubens SL, Vernberg EM, Jacobs AK, Kanine RM. Towards successful dissemination of psychological first aid: a study of provider training preferences. J Behav Health Serv Res 2014 Apr;41(2):203-215. [doi: 10.1007/s11414-013-9362-y]
51. Ahmad NA, Abd Rauf MF, Mohd Zaid NN, Zainal A, Tengku Shahdan TS, Abdul Razak FH. Effectiveness of instructional strategies designed for older adults in learning digital technologies: a systematic literature review. SN Comput Sci 2022;3(2):130. [doi: 10.1007/s42979-022-01016-0] [Medline: 35039803]
52. Shorey S, Kua EH, Tam W, et al. “Where-There-Is-No-Psychiatrist Integrated Personal Therapy” among community-dwelling older adults: a randomized pilot study. Int J Environ Res Public Health 2021 Sep 9;18(18):9514. [doi: 10.3390/ijerph18189514] [Medline: 34574438]
53. Terry J. Experiences of instructors delivering the mental health first aid training programme: a descriptive qualitative study. J Psychiatr Ment Health Nurs 2010 Sep;17(7):594-602. [doi: 10.1111/j.1365-2850.2010.01557.x] [Medline: 20712682]
54. Schafer A, Snider L, van Ommeren M. Psychological first aid pilot: Haiti emergency response. Intervention (Amstelveen) 2010 Nov;8(3):245-254. [doi: 10.1097/WTF.0b013e32834134cb]
55. Semlitz L, Ogiwara K, Weissbecker I, et al. Psychological first aid training after Japan’s triple disaster: changes in perceived self competency. Int J Emerg Ment Health 2013;15(3):181-196. [Medline: 24558747]
56. Choi YJ. Psychological first-aid experiences of disaster health care workers: a qualitative analysis. Disaster Med Public Health Prep 2020 Aug;14(4):433-436. [doi: 10.1017/dmp.2019.87] [Medline: 31587684]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
PFA: Psychological First Aid

Edited by Y Jiang; submitted 27.01.25; peer-reviewed by K Machin, L Perrier; revised version received 02.04.25; accepted 07.04.25; published 13.06.25.

Please cite as:
Shorey S, Seow JSL, Mahendran R, Wee GH, Goh LG, Kua EH
Peer Volunteers’ Journeys Through Training and Engagement in Older Adult Communities: Descriptive Qualitative Study
JMIR Aging 2025;8:e71810
URL: https://aging.jmir.org/2025/1/e71810
doi:10.2196/71810

© Shefaly Shorey, Jeremy Soon Leong Seow, Rathi Mahendran, Geok Hua Wee, Lee Gan Goh, Ee Heok Kua. Originally published in JMIR Aging (https://aging.jmir.org), 13.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and

XSL•FO
RenderX

reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Knowledge, Attitudes, and Behaviors Related to Dementia Prevention and Caregiving Among Korean Americans (the KIMCHI Project): Pre- and Posttest Evaluation Study

Van Ta Park[1], MPH, PhD; Bora Nam[1], MSN, PhD; Daren Huang[1], MPH; Stacy W Yun[2], PhD; Nicole Phan[1], BS; Eun Jeong Lee[3,4], PhD; Hye-Won Shin[5,6], PhD

[1]Department of Community Health Systems, School of Nursing, University of California, San Francisco, San Francisco, CA, United States

[2]Psychology Department, University of Colorado Colorado Springs, Colorado Springs, CO, United States

[3]Asian American Resource and Information Network, Inc, Wood Ridge, NJ, United States

[4]Department of Human Services, NYC College of Technology, City University of New York, New York, NY, United States

[5]Somang Society, Cypress, CA, United States

[6]University of California, Irvine Institute for Memory Impairments and Neurological Disorders, Irvine, CA, United States

Corresponding Author:
Van Ta Park, MPH, PhD
Department of Community Health Systems
School of Nursing
University of California, San Francisco
490 Illinois St
San Francisco, CA, 94158
United States
Phone: 1 415 514 3318
Email: Van.Park@ucsf.edu

Abstract

Background: Many Korean American older adults have limited English proficiency, have low socioeconomic status, and are immigrants. The availability and accessibility of linguistic and culturally appropriate dementia-related health care and caregiving resources for this population are limited. This is concerning given that Korean American older adults are a rapidly growing population, and Alzheimer disease and related dementias (ADRD) represent a significant public health issue.

Objective: We aimed to assess changes in pre- and posttest knowledge, attitudes, and behaviors regarding ADRD prevention and caregiving among participants in the Koreans Invested in Making Caregivers' Health Important (KIMCHI) before and after their participation in KIMCHI workshop presentations.

Methods: A community engagement dissemination project, KIMCHI, was developed by academic and community partners to culturally and linguistically tailor selected evidence-based research from the Patient-Centered Outcomes Research Institute for Korean American older adults, caregivers, and other stakeholders. Dissemination activities were conducted in-person (as workshops) and on digital platforms. Through partnerships with 1 academic institution and 2 community organizations that serve Korean Americans, 211 participants participated in the KIMCHI in-person workshop presentations, and 134 participants participated asynchronously online (via fact sheets on the project website, YouTube videos, and other social media, such as Facebook and X [formerly known as Twitter]). Pre- and postparticipation tests on knowledge, attitudes, and behaviors for ADRD and caregiving were conducted with workshop participants. We administered satisfaction surveys to all participants (workshop and online) and conducted two-tailed paired-sample 2-tailed *t* tests to assess mean changes in the pre- and postparticipation tests.

Results: Among the workshop participants (N=211), most (114/204, 55.9%) were older adults (mean age 69, SD 12.1 y; range 24-90 y), female (n=148, 70.1%), and foreign-born (n=203, 96.2%). Many reported having limited English proficiency (167/211, 79.1%). Significant changes were observed in posttest ADRD knowledge (mean 11.51, SD 2.64), attitudes (mean 7.13, SD 2.83), and behaviors (mean 8.88, SD 2.72) compared to pretest scores (knowledge: mean 10.34, SD 2.67; t_{210}=1.17; P<.001; attitudes: mean 6.33, SD 2.44; t_{210}=0.8; P<.001; and behaviors: mean 8.11, SD 2.68; t_{210}=0.76; P<.001). Workshop participants reported high satisfaction (196/209, 93.8%) with KIMCHI, with the workshop presentations being perceived as culturally relevant and applicable (196/209, 93.8%). Most workshop participants expressed interest in learning more about ADRD-related health topics

XSL•FO
RenderX

(186/209, 89%). Similarly, the online participants (N=134) expressed high satisfaction (115/134, 85.8%) and agreed that the topics and content were culturally relevant and applicable (116/133, 87.2%) and learned new information (110/134, 82.1%).

Conclusions: The findings indicate that KIMCHI may have a positive impact on improving ADRD knowledge, attitudes, and behaviors among Korean Americans. Academic-community collaborations should continue to culturally tailor the programs and studies to help ensure greater representation of Korean Americans in research and community engagement projects.

(JMIR Aging 2025;8:e72147) doi:10.2196/72147

KEYWORDS

Alzheimer disease and related dementias; Korean Americans; knowledge; attitudes; behaviors; caregiving; community-based dissemination

Introduction

Background

According to the US Census Bureau and Pew Research Center data, Korean Americans, who encompass 8% of all Asian American subgroups, have rapidly expanded to approximately 2 million in the United States as of 2023 and are the fifth largest Asian origin in the United States [1,2]. Following the Immigration and Nationality Act of 1965, which removed national immigration quotas, the Korean immigrant population steadily rose, contributing to the development of Korean communities across the United States [3].

Despite their growing presence, Korean Americans face a range of challenges on dementia-related knowledge gaps. The 2015 Asian American Quality of Life survey, which included 471 Korean Americans, found that lack of literacy led to concerns, worries, and unpreparedness if one has Alzheimer disease (AD) or is a caregiver for individuals with AD [4]. In addition, researchers who conducted a cross-sectional survey of AD knowledge among Korean Americans (N=268) in the Greater Washington metropolitan area found the main topic participants reported to be the least knowledgeable about was caregiving [5]. In a scoping review that examined the caregiving experiences of Korean American caregivers of persons living with dementia [6], the authors described how caregivers experience challenges in finding culturally relevant AD and related dementias (ADRD) resources, which limits their ability to learn new information about dementia caregiving.

Dementia remains a major health concern among older Korean American adults and their families. Studies indicate that Korean American caregivers experience higher levels of caregiving distress and fewer coping mechanisms compared to the general population, partly due to a common belief that dementia is a natural part of aging, which discourages seeking support [7,8]. Notably, Korean American caregivers have difficulty in confronting their family members about memory loss due to stigma associated with memory loss and the belief that it is better to keep the issue in the family [7]. Koreans living in the United States tend to engage in family caregiving rather than relying on formal support services [9]. In addition, only 7% of older Korean Americans with probable dementia seek medical advice [10]. While informal social support helps reduce caregiving stress, these findings underscore the need for culturally appropriate dementia education and support within this community [11]. While caregiver stress is a concern, this project does not directly measure or address it. However, to bridge the dementia-related knowledge gaps, we launched the Koreans Invested in Making Caregivers' Health Important (KIMCHI) project, a national bilingual project designed to address the misconceptions about ADRD and empower Korean Americans with information on healthy cognitive aging and caregiving resources. To the best of our knowledge, there is 1 previous educational study [12] of Korean Americans with dementia and their caregivers; however, the study did not implement an immediate pre- and posttest approach nor include a digital platform for engagement. By incorporating core Korean cultural values that emphasize family-based support, KIMCHI aims to provide tools for informed health care decision-making while respecting these cultural traditions. We hypothesized that participation in the KIMCHI project would lead to improved ADRD-related knowledge, attitudes, and behaviors (KAB) among Korean American participants, and an increased willingness to learn more about ADRD and be involved in future projects supported by the Patient-Centered Outcomes Research Institute (PCORI).

Objectives

KIMCHI was supported by a PCORI community dissemination award to share recent patient-centered research findings with the Korean American community (eg, caregivers, older adults, and general public) in culturally and linguistically appropriate ways, as well as to enhance participants' KAB regarding ADRD, healthy brain aging, and caregiving before and after attending the KIMCHI workshop presentations.

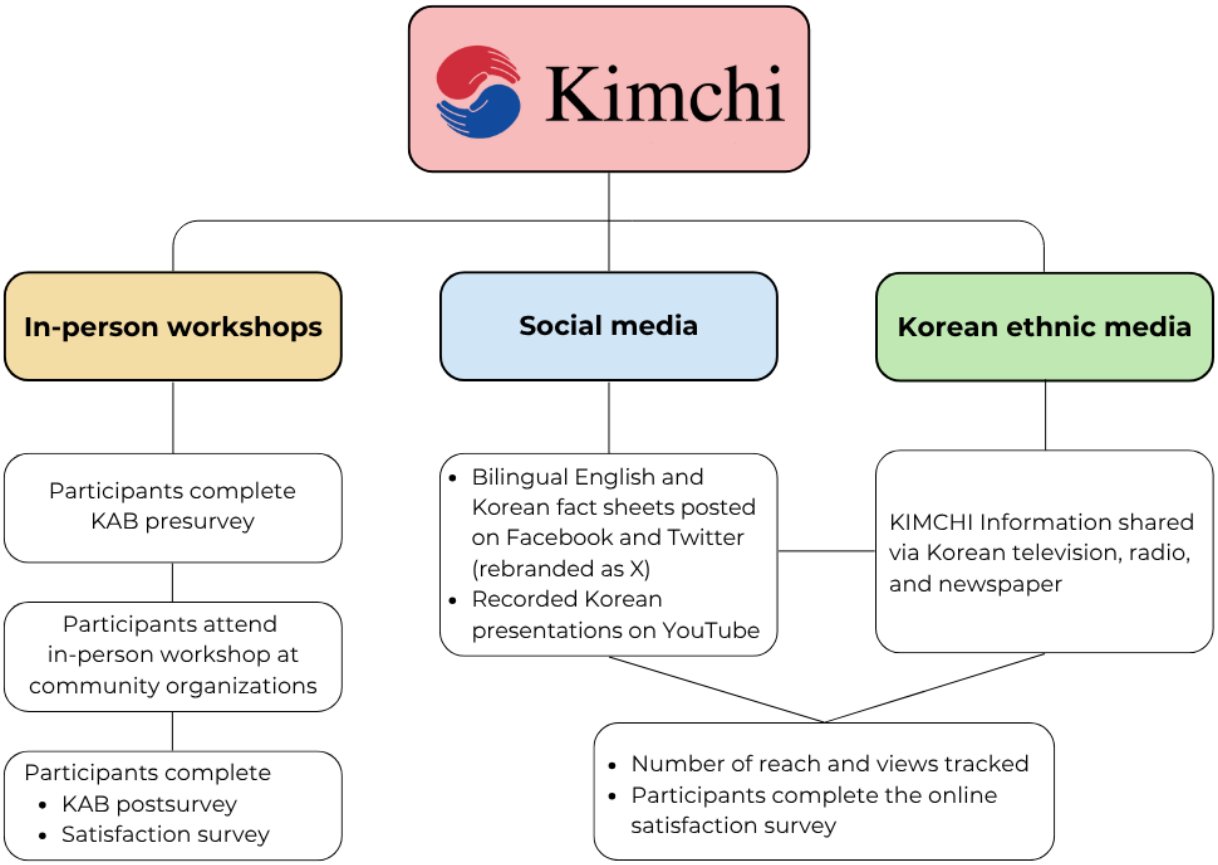
Methods

Design and Methods

The KIMCHI project culturally and linguistically tailored PCORI evidence findings and then disseminated this to Korean American stakeholders, such as patients with dementia, family members, caregivers, health care providers, and community-based organizations. Our dissemination strategy involved 3 main approaches (Figure 1; Textbox 1).

XSL•FO
RenderX

Figure 1. Flowchart of the 3 main dissemination strategies of the Koreans Invested in Making Caregivers' Health Important (KIMCHI) project. KAB: knowledge, attitudes, and behaviors.



Textbox 1. Dissemination strategies.

In-person workshops

- We presented the content via in-person workshops that were conducted in Korean. We administered pretest, posttest, and satisfaction surveys to the participants.

Social media

- We posted the bilingual English and Korean fact sheets on Facebook and Twitter (rebranded X), as well as the recorded Korean presentations on YouTube. We examined engagement via the number of views on YouTube, Facebook, and Twitter (rebranded as X).

Korean ethnic media

- We shared information about the Koreans Invested in Making Caregivers' Health Important (KIMCHI) project (eg, what it is, website, bilingual fact sheets, as appropriate) via major Korean media, including television, radio, and newspapers. The ethnic media provided us with the estimated reach for each type of medium. For both social and ethnic media, we encouraged the viewers and listeners to complete the satisfaction survey after reviewing the bilingual fact sheet.

Participants and Recruitment

Overview

Participants were eligible to participate in KIMCHI activities if they identified as Korean American, were aged ≥18 years, and could read or speak Korean or English. Although ADRD predominately affects older individuals and caregivers, an adult of any age may be a caregiver to older adults (eg, grandparent), particularly among Asian American (including Korean American) families who live in multigenerational households (24% vs 13% White individuals in 2021) [13] and who are family-centric. For in-person project activities, participants were either asked to confirm their eligibility by phone or email before the scheduled activities. Recruitment efforts primarily leveraged community-based organizations, such as the Somang Society and the Asian American Resource and Information Network (AARIN), that serve Korean Americans in the United States. In addition to these organizations, recruitment strategies included flyers; Korean ethnic media outlets (major Korean

XSL•FO
RenderX

television channels, Seoul Broadcasting System-USA, The Korean Channel news, and Radio Korea); social media; and word-of-mouth referrals. KIMCHI was also shared with the Korean participants in the Collaborative Approach for Asian Americans, Native Hawaiians, and Pacific Islanders (AANHPI) Research and Education registry, a registry aimed toward improving the representation of AANHPI populations in aging, ADRD, and caregiving research [14,15].

Initially designed to include 4 web-based presentations, the project expanded based on community feedback and technological barriers faced by older adults. The community's main counterargument to our initial web-based presentation was that potential Korean American participants would have trouble joining Zoom on their phone or computer, and participants stated that "they do not want to have to bother their children or grandchildren to assist them due to their lack of technological knowledge," which will deter them from enrolling into KIMCHI. Therefore, the KIMCHI team mitigated these barriers and hosted in-person workshops to accommodate the Korean American community.

Selected PCORI Evidence

The KIMCHI project was primarily supported by a PCORI community engagement dissemination award, which supports projects that help organizations and communities share pertinent PCORI-funded research findings to patients, families, caregivers, and the broader health care community in ways that will command their attention and interest and encourage use of this information in their health care decisions [16].

Findings from 2 PCORI-funded studies [17,18] informed our community education efforts to reduce stigma and facilitate discussions on ADRD caregiving. The rationale for selecting these studies was that they were the most relevant educational material on ADRD caregiving for Korean Americans among the list of potential PCORI-funded studies to disseminate. In the studies by Smith [17] and Jennings et al [18], the authors recommended that patients with dementia identify key personal goals (physical functioning, meaningful activities, and not being a burden to family) and prioritize these goals to improve their well-being. To do that, there has to be community education on dementia that informs caregivers how to set goals, including (1) the safety of patients with dementia, (2) goals with health care providers, and (3) managing caregivers' stress [18]. The patients and caregivers need to readdress the goals of both sides if the patient with dementia progresses, articulating the need for caregivers to adapt and change their goals to accommodate the patient [18]. The KIMCHI project helps to provide caregivers with the necessary knowledge and skills to better care for their loved ones. In addition, both studies emphasized the need to provide community members with knowledge on healthy cognitive aging and what can be done to prevent dementia by offering examples of physical exercise, brain fitness, and social engagements [17,18]. Thus, as an example, the KIMCHI project informs the participants of (1) the knowledge component, which is the positive benefits of physical exercise, brain fitness, and social engagement for ADRD prevention; (2) the attitudes component, where examples of physical exercise, brain fitness, and social engagement activities will be discussed, which will affect their attitude or intention to practice these activities; and (3) behaviors component, where participants will indicate a plan or intention to change their behavior.

Additional studies also emphasized the benefits of behavioral interventions for caregivers and patients with dementia. Research on cognitive aging, dementia prevention, and goal setting for patients and caregivers highlights evidence-based strategies, including physical exercise and cognitive training for social and emotional quality of life, potentially reducing the dementia risk [17,18]. In addition, findings on dementia prevention interventions, such as wellness activities and yoga, improve the mood and lower the burden of caregivers and their partners with amnestic mild cognitive impairment [19]. On the basis of these findings, our dissemination content emphasizes brain health and dementia prevention strategies, outlining risk factors and protective factors while reinforcing a holistic approach to cognitive aging, conveying the message that what is good for the heart is good for the brain. In addition, it highlights the need for community support and culturally tailored education for Korean Americans. We also focus on 5 key domains in dementia caregiving—medical care, physical and emotional well-being, access to services, and caregiver support—stressing the importance of safety, stress management, and meaningful activities.

Cultural and Linguistic Tailoring

A culturally sensitive dissemination strategy was used, including bilingual materials and messaging tailored to the Korean American community. Initially, the KIMCHI team drafted dissemination materials, including presentation slides and fact sheets, in English. The materials were reviewed by the KIMCHI-wide team, including scientific advisers, who provided feedback on the accuracy, clarity, and interpretation of the PCORI research findings; their input informed several key improvements, such as replacing technical terms such as "psychotherapy" with more accessible language such as "engaging in counseling," tailoring content for the intended audience, incorporating best practices to support comprehension, and enhancing clarity through simplified phrasing and the addition of concrete examples. Once approved, the materials underwent a structured translation process, guided by the World Health Organization's (WHO's) translation and adaptation guidelines [20]. These guidelines were selected because they are specifically designed for translating and adapting public health materials across diverse linguistic and cultural contexts. The WHO approach emphasizes conceptual equivalence rather than direct linguistic translation, ensuring that key health messages remain accurate, culturally appropriate, and accessible to target populations [20]. The translation process followed a 3-step approach, informed by key principles of the WHO translation and adaptation guidelines, with the goal of ensuring both cultural relevance and linguistic accuracy. First, bilingual members of our research team conducted the initial translation, prioritizing conceptual clarity over direct linguistic equivalence. Second, the translated materials underwent review by community-based organization advisers to evaluate cultural appropriateness and clarity. Third, the content was verified through an iterative process involving bilingual staff and

XSL•FO
RenderX

outreach team members, who ensured alignment with the lived experiences and language norms of the target communities. To ensure linguistic and cultural appropriateness, we also cross-checked key terminology and phrasing with native Koreans living in South Korea, which helped confirm both accuracy and resonance with Korean cultural norms. This iterative process ensured that all dissemination materials were both accessible and contextually meaningful to our participants.

We anticipated that content that was both culturally respectful and carefully targeted would foster acceptance and engagement with new knowledge, motivating behavioral change. In alignment with cognitive behavioral theory [21], the project facilitated discussions around cultural perspectives and traditional practices related to dementia and caregiving. The KAB model is based on the cognitive behavioral theory, and the educational program helps participants reflect on their values, habits, needs, and wants and empowers them to explore and adopt desirable attitudes and behaviors about health care practices to improve their health. Following the dissemination activity, we aimed to motivate and empower participants to apply their newfound knowledge to their lives and caregiving after attending the training and educational workshops with experts and trusted leaders, or from reading the how-to-materials fact sheets.

Dissemination Activities

KIMCHI conducted in-person workshop presentations and online dissemination activities.

Workshop Presentations

Key findings from the studies by Smith [17] and Jennings et al [18] were synthesized and culturally adapted to resonate with the specific needs, values, and concerns of the Korean American community. Specifically, we incorporated Korean-language metaphors and idiomatic expressions to enhance relevance and comprehension and reflected culturally specific caregiving practices commonly observed in Korean American families, such as the value of filial piety and multigenerational caregiving structures. We also integrated religious and spiritual elements, including the influence of Christianity and ancestral rites, which shape caregiving beliefs and behaviors.

Moreover, insights from Somang Society's dementia caregiver support groups informed our cultural adaptation process. Through their work with Korean American caregivers, they identified strategies used by older adults to conceal early signs of dementia due to stigma or fear of judgment. For example, individuals experiencing unusual forgetfulness may avoid admitting memory lapses by using general honorifics such as "Elder [Name]" instead of recalling specific names of fellow church members. This behavior reflects a broader cultural tendency to hide symptoms and reinforces the need for sensitive, community-informed approaches in addressing early detection and support.

All cultural adaptations were developed through iterative consultations with bilingual community partners and cultural experts to ensure the materials were accessible and meaningfully aligned with the lived experiences of Korean American participants. The workshop was facilitated by bilingual community partners from AARIN and the Somang Society, each holding a PhD—one in biomedical and biochemical engineering, the other in social work. Both facilitators possess extensive expertise in dementia and caregiving research and have demonstrated longstanding leadership in serving older Korean American communities. They were involved in the translation process to ensure that the content delivered was appropriate and understandable for an audience composed primarily of older Korean Americans, reviewed the key PCORI findings, collaborated with other partners in the network, and created culturally sensitive presentation slides to present to the participants.

To evaluate changes in KAB related to ADRD and cognitive aging, participants were asked to complete a pretest questionnaire before and after the workshop. Participant satisfaction was also evaluated through a brief survey. From March 2023 to May 2023, we conducted 11 workshops across 5 community sites: Korean Community Services of Metropolitan New York, Korean Community Services New Jersey, Chicago Hanul Family Alliance, Asian Women's Christian Association New Jersey, and La Palma Church in California. Each workshop lasted approximately 1 hour. Participants spent 10-15 minutes completing pre- and postparticipation tests and an additional 5 minutes on a satisfaction survey.

Online Dissemination

Fact Sheets

Two KIMCHI fact sheets, titled "Healthy Cognitive Aging" and "Caregiving for Loved Ones with Dementia," were created in both English [22] and Korean [23] (Multimedia Appendices 1 [17,18] and 2 [17,18], respectively). These fact sheets aimed to empower the Korean American community with practical tips and comprehensive information on enhancing cognitive health and improving caregiving practices. The fact sheets were disseminated through various channels, including Somang Society's printed newsletters, emails to participants in the Collaborative Approach for AANHPI Research and Education registry [14,15], and in-person at the Irvine Korean Cultural Festival. The 2-page flyer included a QR code and a link to the survey.

Social Media Outreach

KIMCHI also established and used social media channels, including Facebook [24], X (formerly known as Twitter) [25], and YouTube [26], for dissemination. Messages were crafted to engage the Korean American community effectively, using culturally relevant content in both English and Korean to promote ongoing involvement with the KIMCHI project. This broad range of engagement styles aimed to increase the reach toward Korean Americans nationwide.

Ethical Considerations

We would like to clarify that the University of California, San Francisco institutional review board determined that this project is not a formal research study, but rather a program evaluation; thus, this project does not require institutional review board approval or informed consent. Nevertheless, data were securely collected and stored in University of California, San Francisco's

XSL•FO
RenderX

REDCap (Research Electronic Data Capture; Vanderbilt University), which is compliant with the Health Insurance Portability and Accountability Act. All 211 attendees of the KIMCHI workshop completed the pre- and posttest surveys and were awarded a gift card as appreciation for their participation and time. Online participants were also compensated with a gift card.

Measures

Sociodemographic Information

Participants' self-reported demographic information was also collected, including age, sex, birthplace (including the year of arrival to the United States, if applicable), education, and household income. In addition, participants were asked to rate their proficiency in speaking, reading, and writing English on a 5-point Likert scale from "very well" to "not at all." Participants who reported "some," "a little bit," or "not at all" were categorized as having limited English proficiency (LEP). Participants were also asked about their experience with individuals affected by ADRD, history of receiving education or training in dementia care, and interest in further learning about dementia.

KAB Survey

Pre- and posttest surveys on KAB regarding cognitive aging and caregiving were conducted to assess changes in participants' KAB regarding cognitive aging and caregiving. The surveys consisted of 21 items based on validated questionnaires [27,28] as well as additional culturally appropriate items that were developed for the KIMCHI project (Textbox 2 [27,28]).

Textbox 2. Details of the Knowledge, Attitudes, and Behaviors pre- and postparticipation survey.

Knowledge

- On the basis of the Alzheimer Disease Knowledge Scale from Carpenter et al [27], 9 knowledge component items were designed as true or false questions, and the instrument can be found in Multimedia Appendix 3 [27,28]. Knowledge item 1 was developed by the Koreans Invested in Making Caregivers' Health Important (KIMCHI) team. The total sum score from the scale was added, where the higher score demonstrated greater knowledge related to Alzheimer disease and related dementias (ADRD). The standardized Cronbach α for the knowledge items in the pre- and postparticipation tests was 0.55 and 0.70, respectively.

Attitude

- Referencing the Dementia Attitudes scale-6 from Clark et al [28], 6 attitude items were rated on a 3-point Likert scale from "agree to neutral to disagree," to assess participants' feelings about caregiving for individuals with ADRD (Multimedia Appendix 3). The total sum score from the scale was computed, where the higher score yielded more positive attitudes regarding ADRD. The standardized Cronbach α for the attitude items in the pre- and postparticipation tests was 0.63 and 0.77, respectively.

Behavior

- Six behavior component items, rated on a 3-point Likert scale from "agree to neutral to disagree," were used to assess participants' engagement in activities related to cognitive aging and ADRD prevention (Multimedia Appendix 3). The behavior items were developed by the KIMCHI team and supported by the Patient-Centered Outcomes Research Institute advisers and the KIMCHI experts as appropriate to use. The total sum score from the scale was calculated, where the higher score showed greater positive engagement related to ADRD prevention and cognitive aging. The standardized Cronbach α for the behavior items in the pre- and postparticipation tests was 0.77 and 0.82, respectively.

After the workshop presentation, participants completed the same 21 items on KAB and a satisfaction survey (refer to the subsequent section). The standardized Cronbach α for the overall KAB items in the pre- and postparticipation tests was 0.70 and 0.79, respectively. Both surveys were not anonymized to connect participants' pre- and posttest responses, and they were available in English and Korean, as well as electronic and paper formats. However, all participants opted to complete the survey in Korean and in paper format. Volunteers and staff members were available to assist participants with questions during the survey completion process.

Online Dissemination Satisfaction Survey

To evaluate participant experiences with the KIMCHI project's educational materials, an online satisfaction survey, available in Korean and English, was administered and opened to anyone. We did not have a variable that collected participants' language preference. This survey was designed to gather feedback on participants' satisfaction with the dissemination activities, the relevance of the content, and their learning outcomes. The survey included questions assessing overall satisfaction, knowledge gained, and interest in future educational opportunities. Respondents rated their satisfaction with the workshop on a 5-point Likert scale from "very satisfied" to "very unsatisfied." Participants were also asked if they learned something new from the program and if they were interested in participating in future community educational events related to cognitive aging, dementia caregiving, and advanced health care directives. For specific outreach channels where pretests were not feasible (eg, radio or social media), tailored questions were included. For Korean radio listeners, participants were asked about their understanding of ADRD prevention and care before and after listening to the program. Social media users were similarly inquired about their perceptions before and after engaging with content on platforms such as Facebook. Observations on what worked well and areas for improvement were discussed during the KIMCHI meetings that convened biweekly. The survey tool was administered to participants who engaged with the program through various platforms, including YouTube recordings, fact sheets, and social media posts. To check for the genuineness of entries, participants had to provide their email addresses and zip codes, and the KIMCHI team evaluated each survey to verify and ensure there were no duplicate entries from the same person. Each participant (via

XSL•FO
RenderX

email) was provided with only one survey participation incentive. All online survey responses were collected, pooled, and analyzed to measure the effectiveness of the educational materials and content and identify areas for improvement.

Statistical Analysis

Data from the pretest and posttest questionnaires were analyzed to assess the effectiveness and significance of the KIMCHI educational community workshop presentations in improving participants' KAB regarding ADRD prevention and dementia care. Because KAB has 3 possible categorical outcomes of either "true, false, or do not know" or "agree, neutral, or disagree," we used Bhapkar test [29], which can be thought of as a generalization of McNemar test (which is only appropriate for dichotomous outcomes), to test whether there is marginal homogeneity between one or more of the row marginal proportions of the pretest and the column proportions of the posttest. Missing knowledge responses were included as "don't know" in the analysis, and missing attitude and behavior responses were included as "neutral." The pre- and posttest cores for knowledge were calculated as 2 points if answered correctly, 1 for do not know, and 0 for incorrect responses. Knowledge items 1, 3, 5, 6, and 9 were reverse-coded. In contrast, for attitude and behavior scores, 2 points were assigned for agree, 1 for neutral, and 0 for disagree responses. Total scores were combined in their perspective component to calculate the mean pre- and posttest scores. Overall weighted scores consider that the knowledge component has 3 more items than the attitudes and behaviors component, whereas all 3 subscales were equal in overall unweighted scores. Weighting was applied to ensure the total weighted score was contributed proportionally from each KAB item and prevent the knowledge items from being overpowered because it has 3 more items than the attitude and behavior items. Using paired samples *t* tests, we examined the overall mean changes in participants' KAB and 95% CI regarding ADRD pre- and posttest scores. We used a Bonferroni corrected *P* value threshold to correct for multiple comparisons. As suggested by Cohen [30], we interpreted the effect size between the 2 sample means as small (d=0.2), medium (d=0.5), or large (d=0.8).

Overall, descriptive statistics summarized participants' demographic characteristics. In contrast, inferential statistics, such as paired *t* tests, were used to compare pretest and posttest scores and assess the significance of any observed changes. To test for differences in the 5 pre-post score test components while controlling for additional factors, as a sensitivity analysis, we fit a linear mixed model for each of the test components with a random effect for individual and fixed effects for age, sex, LEP, education, and income. Subgroups with the highest number were the reference group. Statistical analyses were performed using SAS (version 9.4; SAS Institute Inc) software.

Results

In-Person Workshop Participant Characteristics

Table 1 presents the demographic characteristics of the KIMCHI participants. The average age of participants was 69 (SD 12.1) years, with 70.1% (148/211) of the participants identifying as female. The vast majority (203/211, 96.2%) were born outside of the United States and had LEP (167/211, 79.1%). More than half of the participants (117/211, 55.4%) had earned a bachelor degree or higher and had an annual household income of US $75,000 or less in the last year (104/211, 49.3%). In addition, more than a third of the participants (78/211, 37%) knew or worked with someone who has ADRD, and among them, 61% (47/77) had informal or family caregiving experience, and 39% (29/74) had professional experience. Notably, 83.4% (176/211) of the participants reported they had never received education or training in dementia care before.

XSL•FO
RenderX

Table 1. Overview of the sample characteristics among Koreans Invested in Making Caregivers' Health Important (KIMCHI) workshop participants (N=211).

Characteristics	Values
Age (y), mean (SD; range)	69 (12.1; 24-90)
Age range (y), n (%)	
<50	12 (5.7)
50-59	30 (14.2)
60-69	52 (24.6)
70-79	70 (33.2)
≥80	40 (19)
Did not report	7 (3.3)
Sex, n (%)	
Female	148 (70.1)
Male	59 (28)
Did not report	4 (1.9)
Born in the United States, n (%)	
Yes	2 (0.9)
No	203 (96.2)
Did not report	6 (2.8)
If not born in the United States, year arrived in the United States (n=203), n (%)	
1925-1974	21 (10.3)
1975-1999	127 (62.6)
2000-2023	43 (21.2)
Did not report	12 (5.9)
Limited English proficiency[a], n (%)	
Yes	167 (79.1)
No	35 (16.6)
Did not report	9 (4.3)
Education level, n (%)	
High school or less	44 (20.9)
Some college or technical school	40 (19)
Bachelor degree	80 (37.9)
Master degree or higher	37 (17.5)
Prefer not to answer	10 (4.7)
Annual household income (US $), n (%)	
≤25,000	41 (19.4)
25,001 to 75,000	63 (29.9)
75,001 to 150,000	40 (19)
≥150,001	21 (10)
Prefer not to answer	46 (21.8)
“Have you ever known or worked with someone who has Alzheimer disease or dementia,” n (%)	
Yes	78 (37)
No	130 (61.6)
Did not report	3 (1.4)

XSL•FO
RenderX

Characteristics	Values
"I have informal or family caregiving experience[b]" (n=77), n (%)	
Yes	47 (61)
No	30 (39)
"I have professional caregiving experience[b]" (n=74), n (%)	
Yes	29 (39.2)
No	45 (60.8)
"Have you ever received education or training in dementia care," n (%)	
Yes	33 (15.6)
No	176 (83.4)
Did not report	2 (0.9)

[a]"No" refers to those who can speak, read, or write in English very well or well. "Yes" refers to those who can speak, read, or write some, a little bit, or not at all in English.

[b]Supplemental questions on informal or family caregiving and professional experiences were asked if participants answered "yes" to "Have you ever known or worked with someone who has Alzheimer's disease or dementia?"

Testing Marginal Homogeneity

Regarding the marginal homogeneity between the pre- and posttest responses, 48% (10/21) of the KAB items indicated significant differences (P<.002; Bonferroni correction for 21 tests; Table 2).

In the knowledge component, the posttest responses of 44% (4/9) of the survey items differed significantly from the pretest responses of the same participants: (1) "Having high blood pressure may increase a person's risk of developing Alzheimer's disease and related dementias" (P<.001), (2) "Once people have Alzheimer's disease and related dementias, they are no longer capable of making informed decision about their own care" (P<.001), (3) "It has been scientifically proven that mental exercise can potentially help alleviate symptoms of Alzheimer's disease and related dementias" (P<.001), and (4) "People whose Alzheimer's disease and related dementias are not yet severe can benefit from psychotherapy for depression and anxiety" (P=.001).

In the attitudes component, the distribution of posttest responses in 50% (3/6) of the survey items differed significantly from the pretest responses of the same participants: (1) "People with Alzheimer's disease and related dementias can be creative" (P<.001), (2) "It is possible to enjoy interacting with people with Alzheimer's disease and related dementias" (P<.001), and (3) "People with Alzheimer's disease and related dementias can enjoy life" (P<.001).

In the behaviors component, the distribution of posttest responses in 50% (3/6) of the survey items differed significantly from the pretest responses of the same participants: (1) "I engage in physical exercise to promote new brain cell growth" (P<.001), (2) "I engage in healthy diet to lower Alzheimer's disease and related dementias risks" (P=.001), and (3) "I participate in research or other activities (dementia caregiving support group and/or classes) to learn more about Alzheimer's disease and related dementias prevention and care" (P<.001).

The complete distribution of all 21 items with pre- and posttest KAB responses is displayed in Table 2.

XSL•FO
RenderX

Table 2. Distribution of Knowledge, Attitudes, and Behavior (KAB) survey responses, stratified by pre- and postparticipation tests, among workshop participants (N=211).

KAB questions	Pretest, n (%)	Posttest, n (%)	*P* value[a]
Knowledge			
1. "Everyone develops dementia when he or she becomes old."			.50
True	35 (16.6)	36 (17.1)	
False	135 (64)	140 (66.4)	
Do not know	41 (19.4)	35 (16.6)	
2. "Poor nutrition can make the symptoms of Alzheimer's disease or dementia worse."			.003
True	141 (66.8)	163 (77.3)	
False	23 (10.9)	21 (10)	
Do not know	47 (22.3)	27 (12.8)	
3. "When people with Alzheimer's disease or dementia begin to have difficulty taking care of themselves, caregiver should take over right away."			.65
True	171 (81)	170 (80.6)	
False	17 (8.1)	21 (10)	
Do not know	23 (10.9)	20 (9.5)	
4. "Having high blood pressure may increase a person's risk of developing Alzheimer's disease and related dementias."			*<.001* [b]
True	117 (55.5)	186 (88.2)	
False	25 (11.9)	10 (4.7)	
Do not know	69 (32.7)	15 (7.1)	
5. "Most people with Alzheimer's disease and related dementias live in nursing homes."			.27
True	69 (32.7)	60 (28.4)	
False	99 (46.9)	107 (50.7)	
Do not know	43 (20.4)	44 (20.9)	
6. "Once people have Alzheimer's disease and related dementias, they are no longer capable of making informed decision about their own care."			*<.001*
True	146 (69.2)	117 (55.5)	
False	36 (17.1)	63 (29.9)	
Do not know	29 (13.7)	31 (14.7)	
7. "It has been scientifically proven that mental exercise can potentially help alleviate symptoms of Alzheimer's disease and related dementias."			*<.001*
True	159 (75.4)	184 (87.2)	
False	5 (2.4)	9 (4.3)	
Do not know	47 (22.3)	18 (8.5)	
8. "People whose Alzheimer's disease and related dementias are not yet severe can benefit from psychotherapy for depression and anxiety."			*.001*
True	178 (84.4)	197 (93.4)	
False	2 (0.9)	3 (1.4)	
Do not know	31 (14.7)	11 (5.2)	
9. "Eventually, a person with Alzheimer's disease and related dementias will need 24-hours supervision."			.04
True	148 (70.1)	143 (67.8)	
False	25 (11.9)	38 (18)	
Do not know	38 (18)	30 (14.2)	
Attitudes			

XSL•FO
RenderX

KAB questions	Pretest, n (%)	Posttest, n (%)	*P* value[a]
10. "It is rewarding to care for people who have Alzheimer's disease and related dementias."			.90
Agree	163 (77.3)	166 (78.7)	
Neutral	37 (17.5)	35 (16.6)	
Disagree	11 (5.2)	10 (4.7)	
11. "I am comfortable touching people with Alzheimer's disease and related dementias."			.30
Agree	92 (43.6)	81 (38.4)	
Neutral	91 (43.1)	100 (47.4)	
Disagree	28 (13.3)	30 (14.2)	
12. "I feel relaxed around people with Alzheimer's disease and related dementias."			.009
Agree	15 (7.1)	24 (11.4)	
Neutral	97 (46)	106 (50.2)	
Disagree	99 (46.9)	81 (38.4)	
13. "People with Alzheimer's disease and related dementias can be creative."			*<.001*
Agree	33 (15.6)	55 (26.1)	
Neutral	44 (20.9)	53 (25.1)	
Disagree	134 (63.5)	103 (48.8)	
14. "It is possible to enjoy interacting with people with Alzheimer's disease and related dementias."			*<.001*
Agree	31 (14.7)	64 (30.3)	
Neutral	90 (42.7)	83 (39.3)	
Disagree	90 (42.7)	64 (30.3)	
15. "People with Alzheimer's disease and related dementias can enjoy life."			*<.001*
Agree	131 (62.1)	152 (72)	
Neutral	46 (21.8)	43 (20.4)	
Disagree	34 (16.1)	16 (7.6)	
Behaviors			
16. "I engage in cognitively stimulating everyday activities like reading, doing puzzles, learning a new skill or hobby."			.004
Agree	119 (56.4)	139 (65.9)	
Neutral	72 (34.1)	60 (28.4)	
Disagree	20 (9.5)	12 (5.7)	
17. "I engage in physical exercise to promote new brain cell growth."			*<.001*
Agree	109 (51.7)	137 (64.9)	
Neutral	85 (40.3)	63 (29.9)	
Disagree	17 (8.1)	11 (5.2)	
18. "I engage in healthy diet to lower Alzheimer's disease and related dementias risks."			*.001*
Agree	90 (42.7)	112 (53.1)	
Neutral	97 (46)	82 (38.9)	
Disagree	24 (11.4)	17 (8.1)	
19. "I engage in social activities to lower Alzheimer's disease and related dementias risks."			.003
Agree	85 (40.3)	105 (49.8)	
Neutral	100 (47.4)	87 (41.2)	
Disagree	26 (12.3)	19 (9)	
20. "I engage in positive thinking to maintain good mental health."			.95

XSL•FO
RenderX

KAB questions	Pretest, n (%)	Posttest, n (%)	*P* value[a]
Agree	163 (77.3)	164 (77.7)	
Neutral	45 (21.3)	43 (20.4)	
Disagree	3 (1.4)	4 (1.9)	
21. "I participate in research or other activities (dementia caregiving support group and/or classes) to learn more about Alzheimer's disease and related dementias prevention and care."			*<.001*
Agree	51 (24.2)	72 (34.1)	
Neutral	79 (37.4)	80 (37.9)	
Disagree	81 (38.4)	59 (28)	

[a]*P* values were calculated with the Bhapkar chi-square test.

[b]Italicized *P* values indicate statistical significance (*P*<.002, a Bonferroni correction for 21 tests).

Paired-Sample t Tests

As shown in Table 3, we observed a positive increase between the post- and prescores among the overall, weighted (t_{210}=0.95, 95% CI 0.77-1.13) and unweighted (t_{210}=0.91, 95% CI 0.74-1.09) KAB scores. Weighted and unweighted analyses led to the same conclusions. Specifically, in the knowledge summary score, the mean score was 1.17 (SD 2.29; 95% CI 0.86-1.48) points higher at the posttest than at the pretest. Participants collectively showed increased changes in their attitude summary score (t_{210}=0.80, SD 2.22; 95% CI 0.50-1.10) at the posttest than at the pretest. Similarly, for the behavior summary score, the mean score was 0.76 (SD 2.03; 95% CI 0.49-1.04) points higher at the posttest than at the pretest. The knowledge component exhibited the highest average score in both the pre- and postparticipation tests, attributable to the inclusion of 9 knowledge items compared to 6 items each in the attitudes and behaviors components. There was a moderate effect size as the overall weighted and unweighted knowledge component between the pre- and posttest scores differed by 0.71 and 0.51 SDs, respectively. We detected small effect sizes in the components of attitudes (*d*=0.36) and behaviors (*d*=0.38). Through the paired *t* test analysis, all the mean changes and 95% CIs yielded a significance level of *P*<.001.

Table 3. Unweighted and weighted pre- and postsummary scores on overall knowledge, attitudes, and behaviors for workshop participants (N=211).

Survey	Items, n	Prescore, mean (SD; range)	Postscore, mean (SD; range)	*t* test (*df;* SD)	95% CI	Cohen *d*	*P* value[a]
Overall, weighted[b]	21	8.56 (1.75; 5-14)	9.51 (1.73; 6-14)	0.95, (210; 1.33)	(0.77-1.13)	0.71	*<.001*
Overall, unweighted[c]	21	8.26 (9.51; 4-13)	9.17 (1.73; 5-13)	0.91, (210; 1.29)	(0.74-1.09)	0.71	*<.001*
Knowledge	9	10.34 (2.67; 3-18)	11.51 (2.64; 6-18)	1.17, (210; 2.29)	(0.86-1.48)	0.51	*<.001*
Attitudes	6	6.33 (2.44; 0-12)	7.13 (2.83; 0-12)	0.80, (210; 2.22)	(0.50-1.10)	0.36	*<.001*
Behaviors	6	8.11 (2.68; 0-12)	8.88 (2.72; 1-12)	0.76, (210; 2.03)	(0.49-1.04)	0.38	*<.001*

[a]Italicized *P* values indicate statistical significance (*P*<.01; Bonferroni correction for 5 tests).

[b]Overall weighted scores consider that the knowledge component has 3 more items than attitudes and behaviors.

[c]Overall unweighted scores consider that all 3 subscales are equal.

In the linear mixed model results, the difference between the post- and pretest scores for the 5 test components remained unchanged, compared to the paired *t* test (Table 3), after adjusting for age, sex, LEP, education, and income (Multimedia Appendix 4).

In-Person Workshop Satisfaction Survey

Table 4 presents the findings from the in-person participant workshop satisfaction survey. A notable 93.8% (196/209) of participants expressed satisfaction or high satisfaction with the workshop presentation, affirming that the topics, content, and presentation style were culturally relevant and applicable. Furthermore, 89.4% (185/207) of respondents indicated that they learned something new from this presentation and expressed interest in gaining further education on ADRD-related topics.

XSL•FO
RenderX

Table 4. In-person participant satisfaction survey following the dissemination activity (N=211).

Posttest survey questions	Participants, n (%)
"Are you satisfied with the presenter's presentation/workshop today" (n=209)[a]	
Very satisfied	98 (46.9)
Satisfied	98 (46.9)
Neutral	13 (6.2)
Unsatisfied	0 (0)
Very unsatisfied	0 (0)
"The topics, content, and presentation style are culturally relevant and applicable" (n= 209)[a]	
Strongly agree	100 (47.8)
Agree	96 (45.9)
Neutral	12 (5.7)
Disagree	1 (0.5)
Strongly disagree	0 (0)
"Did you learn something new from this presentation" (n=207)[a]	
Yes	185 (89.4)
No	22 (10.6)
"Are you interested in learning more about dementia" (n=209)[a]	
Yes	186 (89)
No	23 (11)

[a]Totals may not add up to 211 due to missing values.

Online Participant Characteristics and Satisfaction Survey

Table 5 displays the findings of the total nonpresentation or online evaluation survey (N=134) from the KIMCHI fact sheets (n=66, 49.3%), YouTube videos (n=63, 47%), and Facebook social media posts (n=5, 3.7%). The average age of the participants was 53 (SD 16.4) years; 64.7% (86/133) of the participants identified as female, while 35.3% (47/133) identified as male; and 79.9% (107/134) of the participants were born outside of the United States. Slightly more than half (75/134, 56%) of the participants demonstrated LEP in reading, writing, and speaking. Participants were also highly educated, with 82.7% (110/133) having a bachelor degree or higher and 46.3% (62/134) with a yearly household income of US $75,000 or higher. Regarding caregiving experiences, 46.6% (62/133) of the participants had known or cared for someone who had ADRD, and among this subgroup, 60% (37/62) had formal or family caregiving experiences, and 23% (14/62) had professional experiences. Although only 19.5% (26/133) of the participants had received education or training in dementia care, 79.1% (106/134) showed interest in learning more about dementia.

The overall satisfaction and content of the evaluation survey were encouraging. Of the 134 participants, 115 (85.8%) expressed that they were satisfied or very satisfied with the KIMCHI fact sheets, YouTube videos, and social media content on Facebook, and 87.2% (116/133) agreed or strongly agreed that the topics, content, and presentation style were culturally relevant and applicable. Finally, 82.1% (110/134) of the participants learned something new from the nonpresentation contents, and 77.6% (104/134) would be interested in learning more about future advanced directives on health, cognitive aging, and dementia caregiving.

Our social media outreach efforts capture significant reach, attention, and participation of our target audience. As of April 15, 2024, our Facebook page has garnered 426 viewers, while our Twitter (later rebranded as X) account has reached an audience of 2453 viewers. Our YouTube channel has attracted 1511 viewers and garnered at least 4390 total views. In addition, we have reached 100,000 listeners from the ethnic media, with these numbers representing total views. Finally, we have distributed 12,500 copies of fact sheets, exceeding our initial goal of 3000 copies, through community partner sites, newsletters, and in-person Korean cultural events.

XSL•FO
RenderX

Table 5. Sociodemographic characteristics of online participants (N=134).

Characteristics	Values
Age (y), mean (SD; range)	53 (16.4; 22-87)
Age range (y), n (%)	
<40	33 (24.6)
40-49	28 (20.9)
50-59	22 (16.4)
60-69	22 (16.4)
≥70	29 (21.6)
Sex (n=133) [a] , n (%)	
Male	47 (35.3)
Female	86 (64.7)
Born in the United States, n (%)	
Yes	27 (20.1)
No	107 (79.9)
Year arrived in the United States (n=107), n (%)	
1919-1999	48 (44.9)
2000-2023	38 (35.5)
Did not report	21 (19.6)
Limited English proficiency, n (%)	
Yes	75 (56)
No	59 (44)
Education level (n=133)[a], n (%)	
High school or less	12 (9)
Some college or technical school	11 (8.3)
Bachelor degree	70 (52.6)
Master degree or higher	40 (30.1)
Annual household income in the last year (US $), n (%)	
≤25,000	13 (9.7)
25,001 to 75,000	39 (29.1)
75,001 to 150,000	34 (25.4)
150,001 to 200,000	11 (8.2)
≥200,001	17 (12.7)
Prefer not to answer	20 (14.9)
“Have you ever known or cared for someone who has Alzheimer’s disease or dementia[a]” (n=133), n (%)	
Yes	62 (46.6)
No	71 (53.4)
“I have informal/family caregiving experience[b]” (n=62), n (%)	
Yes	37 (59.7)
No	25 (40.3)
“I have professional caregiving experience[b]” (n=62), n (%)	
Yes	14 (22.6)
No	48 (77.4)

XSL•FO
RenderX

Characteristics	Values
"Have you ever received education or training in dementia care[a]" (n=133), n (%)	
Yes	26 (19.5)
No	107 (80.5)
"Do you have an interest in learning more about dementia," n (%)	
Yes	106 (79.1)
No	28 (20.9)
"Are you satisfied with the KIMCHI[c]fact sheets, YouTube video you watched or social media content you read," n (%)	
Very satisfied	66 (49.3)
Satisfied	49 (36.6)
Neutral	16 (11.9)
Unsatisfied	3 (2.2)
Very unsatisfied	0 (0)
"The topics, content, and presentation style are culturally relevant and applicable[a]" (n=133), n (%)	
Strongly agree	52 (39.1)
Agree	64 (48.1)
Neutral	15 (11.3)
Disagree	2 (1.5)
Strongly disagree	0 (0)
"Did you learn something new from our fact sheets, YouTube video, or social media content," n (%)	
Yes	110 (82.1)
No	24 (17.9)
"Do you want to learn more about future community educational or learning opportunities related to healthy cognitive aging, dementia caregiving, and advance healthcare directive, such as this research," n (%)	
Yes	104 (77.6)
No	30 (22.4)

[a]Totals do not add up to 134 due to missing values.

[b]Supplemental questions on informal or family caregiving and professional experiences were asked if participants answered "yes" to "Have you ever known or worked with someone who has Alzheimer's disease or dementia?"

[c]KIMCHI: Koreans Invested in Making Caregivers' Health Important.

Discussion

Principal Findings

The KIMCHI project disseminated the PCORI evidence on healthy cognitive aging and dementia care within the Korean American community through a variety of platforms, reaching a wide audience and achieving high participant engagement. The final dissemination effort included 11 in-person workshops, and attrition was a nonfactor as 211 participants who initially enrolled completed pre- and postsurveys. The results revealed significant improvements and changes in participants' KAB (P<.001) regarding cognitive aging and caregiving. In addition, 134 online satisfaction surveys indicated exceptionally high satisfaction, with most respondents reporting newly acquired knowledge and a strong interest in future educational opportunities on cognitive aging and dementia caregiving.

The project's multimodal dissemination approach—combining in-person presentations, ethnic media, and social media—effectively broadened its reach. Each method contributed to engagement as shown by comparative participation rates across platforms. The high satisfaction ratings, with 93.8% (196/209) and 85.8% (115/134) of respondents scoring 3 or higher on a 5-point Likert scale for in-person workshop and online dissemination, respectively, along with statistically significant changes in KAB scores (P<.001), underscore the project's impact. Although we did not survey online participants with KAB items, social media and online resources provided valuable avenues for individuals who could not attend in-person presentations, ensuring the dissemination of critical information on cognitive health and caregiving practices.

In this project, cultural sensitivity refers to the deliberate effort to align content with the linguistic, social, and belief systems commonly observed among Korean American communities,

XSL•FO
RenderX

particularly first-generation immigrants. To make ADRD content culturally appealing and meaningful, several adaptations were implemented. These included using Korean-language idioms and culturally familiar caregiving scenarios, integrating values, such as filial piety and respect for elders, and addressing stigma-related barriers, such as the perception of dementia as a natural part of aging or a source of family shame. We also incorporated religious and spiritual elements, such as Christian-based perspectives and ancestral rites, which are prevalent within segments of the Korean American population. Insights from bilingual community partners and organizations, such as the Somang Society and the AARIN, were instrumental in shaping the content. Their direct experiences with Korean American caregivers highlighted subtle yet impactful behaviors—such as avoiding the use of specific names to hide memory loss—that informed how we addressed stigma and early detection in a culturally responsive way.

The KIMCHI project highlights the impact of culturally adapted, community-based educational programs in addressing knowledge gaps and shifting attitudes and behaviors related to dementia caregiving among Korean Americans. The project's findings resonate with previous research on health education in ethnic minority populations, suggesting that culturally and linguistically tailored programs are instrumental in reducing stigma and enhancing engagement in dementia care and prevention.

Comparison With Previous Work

A limited number of previous studies have examined dementia-related KAB among Korean Americans. The KIMCHI project builds upon this foundation using a structured pre- and posttest KAB model to quantitatively assess changes within this population. KIMCHI differs from previous projects due to its larger sample size (or greater reach), bilingual resources, broader dissemination strategies, and same-day implementation of the pre- and postparticipation tests. This approach is consistent with research on culturally tailored programs, which have shown promise in enhancing dementia literacy among ethnic minority groups. For example, a pilot study [12] tested a dementia literacy intervention for Korean Americans with dementia and their caregivers, in which they assessed long-term knowledge retention, whereas KIMCHI measured same-day posttest changes, reporting a modest increase in dementia literacy scores by an average of 1 point (SD 2.4) at a 12-week follow-up. Similarly, the KIMCHI project observed a significant mean knowledge component change of 1.17 (SD 2.29; $P<.001$) in the pre- and postparticipation tests. These findings collectively underscore the importance of culturally responsive programs, such as KIMCHI, in enhancing dementia-related awareness and caregiving tools among Korean Americans.

Another relevant study [4] explored concerns and planning behaviors related to AD among Asian American subgroups and identified higher levels of concern and planning intent among Korean Americans compared to Chinese Americans. The study found Korean Americans had higher odds of worrying about developing AD (OR 2.65, 95% CI 1.86-3.79), having the possibility of caregiving for someone with AD (OR 1.60, 95% CI 1.16-2.23), and planning for future AD needs (OR 1.56, 95% CI 1.01-2.42) [4]. While previous research [4] found higher levels of AD concern, they did not assess actual engagement in their educational programs compared to the KIMCHI project, which used evaluation surveys after the workshop and digital platform engagement.

In addition, studies involving broader AANHPI communities have reported similar challenges and educational needs. One study [31] implemented a KAB focus group about healthy aging, ADRD, and dementia among older Chinese Americans (N=18) in New York City but did not conduct the pre- and postparticipation tests. The qualitative results disclosed their cultural beliefs about ADRD and dementia, how aging can lead to cognitive decline, and challenges to incorporating daily healthy diets, which were topics discussed in the KIMCHI presentations [31]. Another study [32] also focused on the knowledge, attitudes, and perceptions of ADRD and healthy aging among AANHPI communities without pre- and postparticipation tests. In their interview series with AANHPI community stakeholders (N=11), structural barriers to health care, cultural norms, and the lack of social support for AANHPI communities were identified, with a determination to provide culturally focused programs to combat these challenges, early detection, and ADRD prevention [32].

Digital Technology Engagement

The KIMCHI project used digital platforms—social media, the project website, and ethnic media (radio and television)—to reach individuals who could not attend in-person workshops. The rationale for this broad range of engagement styles was to provide stakeholders with flexible access to ADRD resources; we accumulated over 4390 total social media views and 100,000 ethnic media viewers. We were unable to track the total number of fact sheet downloads on the participant's end. This approach was particularly effective for individuals seeking to engage with materials at their convenience, including those who may have otherwise missed in-person presentations. To ensure continued accessibility, we plan to promote our KIMCHI resources, such as fact sheets [22,23] and presentation recordings, through our social media channels, including Facebook [24] and X (formerly known as Twitter) [25] pages, as well as YouTube videos [26]. Up-to-date content, event announcements, and future engagement opportunities will remain accessible on the KIMCHI website [33], ensuring a sustained digital presence (Table 5) and ongoing community engagement.

As a potential next step to explore in future work, we may consider designing a culturally and linguistically appropriate ADRD generative artificial intelligence tool for Korean American family caregivers. This potential next step can be guided by research with other heterogeneous older adult populations. For example, in a recent study [34], researchers tested a culturally relevant artificial intelligence mobile app for Black American informal caregivers and reported some positive findings, including an engaging and inclusive experience.

Limitations

There are several limitations that should be considered when interpreting the findings. One limitation was the timing of the data collection, with pre- and postparticipation tests administered

XSL•FO
RenderX

immediately before and after each workshop. This short interval may have limited participants' ability to fully absorb and reflect on the information presented, potentially leading to an underestimation of immediate long-term understanding or behavior change. In addition, the timing of some workshops, particularly those scheduled after Sunday worship services, may have affected participation and attention levels, especially among older adults who may have been fatigued from an extended period of engagement. Future studies could adjust the timing of workshops to avoid such issues.

The KIMCHI project used a pre- and posttest survey approach to assess short-term changes in KAB among participants following dissemination activities. While this approach offered valuable quantitative insights, we acknowledge the limitations of not including qualitative methods (eg, focus groups or interviews) or a long-term follow-up component. However, these elements were not incorporated due to the defined scope and objectives of the project, which focused on outreach and education rather than formal research. As such, the design prioritized feasible, scalable data collection methods aligned with the project's goals of raising awareness and gathering immediate community feedback.

We recognize the limitations of treating Korean Americans as a monolithic group. The Korean American community is diverse in terms of generation, immigration history, acculturation level, language proficiency, education, and religious beliefs. While our materials were primarily developed for older, first-generation Korean immigrants, we acknowledge that this approach may not fully capture the needs or perspectives of younger generations or more acculturated individuals. Future efforts should strive for a more segmented approach that reflects this intragroup diversity. Overall, while our culturally tailored approach was grounded in community input and relevant cultural frameworks, we recognize the importance of continued refinement and differentiation to reflect the full spectrum of experiences within the Korean American community.

Another limitation was the potential ambiguity of some KAB survey items and response bias. For example, knowledge item 3, "When people with Alzheimer's disease or dementia begin to have difficulty taking care of themselves, the caregiver should take over right away," may not have been clear in specifying whether it referred to professional support or family member caregivers. Because the survey was conducted in person with volunteers and staff members present, this may have led to social desirability bias in responses. The KIMCHI team reassured the participants that their responses would be kept confidential and protected. Future studies should refine survey items to ensure clarity and avoid misinterpretation and should not require participants to provide their names.

There are sociodemographic factors, such as marital status or health care access, that influence ADRD burden and the development of dementia. In the Health and Retirement Study [35], divorced (OR=2.05; P<.001), widowed (OR=1.52; P<.001), cohabiting (OR=1.55; P<.05), and never married (OR=1.6; P<.005) respondents had significantly higher odds of developing dementia than their married counterparts, after controlling for age, sex, race, education, and proxy report. Moreover, having access to health insurance could decrease the burden of ADRD and reduce mortality [36]. Limited studies have explored the link between marital status and health care access and ADRD knowledge, and future studies should investigate the associations.

We acknowledge that there may be a potential for ceiling effects, as some participants answered the knowledge component questions correctly. Among the 211 participants, 25 (11.8%) answered at least 78% (7/9) of the knowledge questions correctly in the pretest, while 3 (1.4%) participants had perfect scores. Although 2 of the 3 participants received perfect scores, they stated "Yes" to the "Did you learn something new from this presentation?" question in the posttest survey. Overall, most of the participants in the workshop did not have high baseline KAB scores, which did not limit the ability to detect improvements in knowledge or changes in attitudes and behaviors in the posttest.

In addition, inconsistencies in measurement scales may have influenced response patterns. While knowledge items were listed in the "true-false-do not know" format, attitudes and behaviors items were measured using a 3-point Likert scale: "agree-neutral-disagree." This discrepancy may have affected comparability and response consistency across domains. Standardizing all KAB items to a uniform scale would enhance internal consistency and improve comparability in future studies.

The KIMCHI project could not provide comparative data on knowledge retention between in-person and online participants. This was one of the challenges in digital outreach as we did not implement the same pre- and posttest program for online participants to examine their KAB components after reading the fact sheets, watching our YouTube videos, or viewing our Facebook posts. The online participants did not have the opportunity to engage with our health care professionals and experts and ask any follow-up questions. The online participants were younger (mean age 53, SD 16.4 y) than the in-person workshop participants (mean age 69 years, SD 12.1), had less LEP (56% online vs 79% in-person), and had obtained a bachelor degree or higher (83% online vs 56% in-person). Studies that intend to conduct online surveys might foresee younger participants rather than older adults, particularly those with higher English proficiency, higher education (at least a bachelor's degree), and more advanced technology skills.

Although fact sheets and social media generated an extensive reach in our online dissemination, our response rate regarding survey completion was low. With potential nonresponse bias, online participants were unable or unwilling to participate because they came across the fact sheets or social media posts randomly on their page and were unaware of what KIMCHI is about. In addition, web page surveys could lead to spam risks, as we received surveys (n=23) that were not genuine due to duplicates of the same responses in multiple surveys or contact information of non-Korean names. Therefore, we could not filter out the total view count to those that were genuine, which could be one reason the total views were higher than the response rate. Another reason for the low response rate is that online respondents need basic computer skills or an internet connection to complete the survey. Nonetheless, future studies

XSL•FO
RenderX

should ensure they accurately differentiate between genuine and spam survey respondents when analyzing results. In this project, we presented the descriptive statistics of the online dissemination data as a comparison group to our workshop participants instead of making inferences. Finally, community-based and convenience sampling limits the generalizability of our findings and may not be representative of broader populations with different demographic or socioeconomic characteristics. Future studies should consider strategies to enhance the representativeness of the sample.

Despite these limitations, there are several notable strengths. One of our project's strengths is to evaluate trends and patterns in KAB in the tests before and after participation in the workshops. In addition, the culturally tailored dissemination activities facilitated social engagement and fostered greater community participation, making the program more accessible and impactful. Importantly, the evaluation findings will contribute to refining future KIMCHI dissemination strategies, improving implementation processes, and enhancing the effectiveness of community educational dissemination activities on cognitive aging and dementia caregiving.

Conclusions

The KIMCHI project's community-based, culturally adapted dissemination model demonstrated effectiveness in changing KAB regarding healthy cognitive aging and caregiving among Korean Americans, in addition to increasing the willingness to learn more about ADRD and being involved in PCORI type of projects. By leveraging PCORI-funded research through a culturally tailored approach, caregivers and stakeholders were empowered to adopt new caregiving strategies that align with traditional Korean values of family and community support. Health care systems are encouraged to consider adopting similar culturally sensitive strategies to promote health equity and access to dementia-related resources for underserved populations. Future studies should continue to explore the efficacy of tailored programs, incorporating real-life case studies and other culturally relevant elements to enhance outcomes and strengthen the representation of Korean Americans in health research.

Acknowledgments

The authors would like to thank all the KIMCHI participants for their participation and the staff involved in the coordination of the workshop presentations. The authors would also like to thank Dr Thomas J Hoffmann for his consultation on the statistical analysis and Bobin Kim for her assistance with Figure 1 of this project. The Koreans Invested in Making Caregivers' Health Important (KIMCHI) project was funded by a Patient-Centered Outcomes Research Institute (PCORI) Eugene Washington PCORI Engagement Award (EACB 24814) and the National Institute on Aging of the National Institutes of Health (awards R24AG063718 and R01AG083926).

Authors' Contributions

VTP conceptualized the project. VTP and BN supervised the project. EJL and HWS led the community educational presentations and outreach efforts, ensuring that the presentation materials and publications were culturally competent. DH analyzed the data. VTP, BN, and DH contributed to the manuscript's writing and revision. All the authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Online fact sheet for English readers.
[PDF File (Adobe PDF File), 784 KB - aging_v8i1e72147_app1.pdf]

Multimedia Appendix 2
Online fact sheet for Korean readers.
[PDF File (Adobe PDF File), 962 KB - aging_v8i1e72147_app2.pdf]

Multimedia Appendix 3
Survey items on knowledge, attitudes, and behaviors.
[DOCX File , 26 KB - aging_v8i1e72147_app3.docx]

Multimedia Appendix 4
Results of the linear mixed effects analysis for the weighted and unweighted pre- and posttest summary scores on knowledge, attitudes, and behaviors (N=211).
[DOCX File , 32 KB - aging_v8i1e72147_app4.docx]

XSL•FO
RenderX

References

1. Krogstad JM, Im C. Key facts about Asians in the U.S. Pew Research Center. 2025 May 1. URL: https://www.pewresearch.org/short-reads/2025/05/01/key-facts-about-asians-in-the-us/ [accessed 2025-05-01]
2. Korean Americans: a survey data snapshot. Pew Research Center. URL: https://www.pewresearch.org/2024/08/06/korean-americans-a-survey-data-snapshot/ [accessed 2024-12-17]
3. Shin KR, Shin C. The lived experience of Korean immigrant women acculturating into the United States. Health Care Women Int 1999 Nov;20(6):603-617. [doi: 10.1080/073993399245494] [Medline: 10889638]
4. Jang Y, Yoon H, Park NS, Rhee MK, Chiriboga DA. Asian Americans' concerns and plans about Alzheimer's disease: the role of exposure, literacy and cultural beliefs. Health Soc Care Community 2018 Mar 04;26(2):199-206 [FREE Full text] [doi: 10.1111/hsc.12509] [Medline: 28980423]
5. Lee SE, Hong M, Casado BL. Alzheimer's disease (AD) knowledge in Korean Americans: identifying knowledge gaps and misconceptions and examining predictors of AD knowledge. Ethn Health 2023 Apr 01;28(3):431-445. [doi: 10.1080/13557858.2022.2045907] [Medline: 35229696]
6. Ta Park VM, Ly Q, von Oppenfeld J, Lee Y, Joo Y, Shin HW, et al. A scoping review of dementia caregiving for Korean Americans and recommendations for future research. Clin Gerontol 2023 Oct 21;46(2):223-239 [FREE Full text] [doi: 10.1080/07317115.2022.2133907] [Medline: 36268979]
7. Watari KF, Gatz M. Pathways to care for Alzheimer's disease among Korean Americans. Cultur Divers Ethnic Minor Psychol 2004 Feb;10(1):23-38. [doi: 10.1037/1099-9809.10.1.23] [Medline: 14992628]
8. Chun M, Knight BG, Youn G. Differences in stress and coping models of emotional distress among Korean, Korean-American and White-American caregivers. Aging Ment Health 2007 Jan;11(1):20-29. [doi: 10.1080/13607860600736232] [Medline: 17164154]
9. Kim HJ, Kehoe P, Gibbs LM, Lee JA. Caregiving experience of dementia among Korean American family caregivers. Issues Ment Health Nurs 2019 Feb;40(2):158-165 [FREE Full text] [doi: 10.1080/01612840.2018.1534909] [Medline: 30620625]
10. Lee HB, Han HR, Huh BY, Kim KB, Kim MT. Mental health service utilization among Korean elders in Korean churches: preliminary findings from the Memory and Aging Study of Koreans in Maryland (MASK-MD). Aging Ment Health 2014 Jul 26;18(1):102-109 [FREE Full text] [doi: 10.1080/13607863.2013.814099] [Medline: 23889338]
11. Lee M, Yoon E, Kropf NP. Factors affecting burden of South Koreans providing care to disabled older family members. Int J Aging Hum Dev 2007 Apr 01;64(3):245-262. [doi: 10.2190/c4u5-078n-r83l-p1mn]
12. Han HR, Choi S, Wang J, Lee HB. Pilot testing of a dementia literacy intervention for Korean American elders with dementia and their caregivers. J Clin Transl Res 2021;7(6):1. [doi: 10.18053/jctres.07.202106.001]
13. Cohn D, Horowitz JM, Minkin R, Fry R, Hurst K. The demographics of multigenerational households. Pew Research Center. 2022 Mar 24. URL: https://www.pewresearch.org/social-trends/2022/03/24/the-demographics-of-multigenerational-households/ [accessed 2025-05-15]
14. Ta Park VM, Meyer OL, Tsoh JY, Kanaya AM, Tzuang M, Nam B, et al. The Collaborative Approach for Asian Americans and Pacific Islanders Research and Education (CARE): a recruitment registry for Alzheimer's disease and related dementias, aging, and caregiver-related research. Alzheimers Dement 2023 Feb;19(2):433-443 [FREE Full text] [doi: 10.1002/alz.12667] [Medline: 35420258]
15. Grill JD, Tsoh JY, Nam B, Tzuang M, Huang D, Fara-On G, et al. Increasing representation of Asian American, Native Hawaiian, and Pacific Islander communities in aging, dementia, and caregiving research: an update from the CARE registry. Alzheimers Dement 2025 Apr 10;21(4):e70144 [FREE Full text] [doi: 10.1002/alz.70144] [Medline: 40207420]
16. Eugene Washington PCORI Engagement Award Program. Patient-Centered Outcomes Research Institute. URL: https://www.pcori.org/engagement-research/eugene-washington-pcori-engagement-award-program [accessed 2025-05-28]
17. Smith GE. Healthy cognitive aging and dementia prevention. Am Psychol 2016;71(4):268-275. [doi: 10.1037/a0040250] [Medline: 27159433]
18. Jennings LA, Palimaru A, Corona MG, Cagigas XE, Ramirez KD, Zhao T, et al. Patient and caregiver goals for dementia care. Qual Life Res 2017 Mar;26(3):685-693 [FREE Full text] [doi: 10.1007/s11136-016-1471-7] [Medline: 28000094]
19. Amofa PA, Locke DE, Chandler M, Crook JE, Ball CT, Phatak V, et al. Comparative effectiveness of behavioral interventions to prevent or delay dementia: one-year partner outcomes. J Prev Alzheimers Dis 2021 Jan;8(1):33-40. [doi: 10.14283/jpad.2020.59] [Medline: 33336222]
20. WHO guidelines on translation and adaptation of instruments: process of translation and adaptation of instrument. World Health Organization. URL: https://www.scribd.com/document/533869240/WHO-Guidelines-on-Translation-and-Adaptation-of-Instruments [accessed 2025-01-10]
21. Gonyea JG, López LM, Velásquez EH. The effectiveness of a culturally sensitive cognitive behavioral group intervention for Latino Alzheimer's caregivers. Gerontologist 2016 Apr 22;56(2):292-302 [FREE Full text] [doi: 10.1093/geront/gnu045] [Medline: 24855313]
22. EN KIMCHI fact sheets. KIMCHI. URL: https://kimchi.ucsf.edu/sites/g/files/tkssra8136/f/EN%20KIMCHI%20Fact%20Sheets.pdf [accessed 2024-12-01]

XSL•FO
RenderX

23. KO KIMCHI fact sheets. KIMCHI. URL: https://kimchi.ucsf.edu/sites/g/files/tkssra8136/f/KO%20KIMCHI%20Fact%20Sheets.pdf [accessed 2024-12-01]
24. Kimchi: Koreans invested in making caregivers health important Facebook page. Facebook. URL: https://m.facebook.com/UCSF.KIMCHI/ [accessed 2024-12-01]
25. UCSF KIMCHI. Twitter. URL: https://twitter.com/ucsf_kimchi [accessed 2024-12-01]
26. Somang society. YouTube. URL: https://www.youtube.com/@somangsociety9750 [accessed 2024-12-01]
27. Carpenter BD, Balsis S, Otilingam PG, Hanson PK, Gatz M. The Alzheimer's Disease Knowledge Scale: development and psychometric properties. Gerontologist 2009 Apr;49(2):236-247 [FREE Full text] [doi: 10.1093/geront/gnp023] [Medline: 19363018]
28. Clark MS, Ebert AR, Hicks Patrick J. The DAS-6: a short form of the Dementia Attitudes Scale. Int J Aging Hum Dev 2023 Jun 03;96(4):488-500. [doi: 10.1177/00914150221106094] [Medline: 35656738]
29. Sun X, Yang Z. Generalized McNemar’s test for homogeneity of the marginal distributions. SAS Global Forum. 2008. URL: https://support.sas.com/resources/papers/proceedings/pdfs/sgf2008/382-2008.pdf [accessed 2024-12-04]
30. Cohen J. Statistical Power Analysis for the Behavioral Sciences. New York, NY: Routledge; 1988.
31. Robbins R, Chong S, Liang A, Chanko N, Trinh-Shevrin C, Kwon S. Knowledge, attitudes, and cultural beliefs about healthy aging and Alzheimer's disease among older Chinese Americans in New York City. J Cross Cult Gerontol 2022 Jun 17;37(2):161-180 [FREE Full text] [doi: 10.1007/s10823-022-09450-6] [Medline: 35579786]
32. Wong JA, Min DK, Kranick J, Ushasri H, Trinh-Shevrin C, Kwon SC. Exploring community knowledge, attitudes and perceptions of Alzheimer's Disease/Alzheimer's disease-related dementias and healthy ageing in Asian American, Native Hawaiian and Pacific Islanders. Health Soc Care Community 2022 Nov 14;30(6):e5946-e5958 [FREE Full text] [doi: 10.1111/hsc.14025] [Medline: 36102595]
33. KIMCHI homepage. KIMCHI. URL: https://kimchi.ucsf.edu/ [accessed 2024-12-01]
34. Bosco C, Otenen E, Osorio Torres J, Nguyen V, Chheda D, Peng X, et al. Designing a multimodal and culturally relevant Alzheimer disease and related dementia generative artificial intelligence tool for Black American informal caregivers: cognitive walk-through usability study. JMIR Aging 2025 Jan 08;8:e60566 [FREE Full text] [doi: 10.2196/60566] [Medline: 39778201]
35. Liu H, Zhang Z, Choi SW, Langa KM. Marital status and dementia: evidence from the health and retirement study. J Gerontol B Psychol Sci Soc Sci 2020 Sep 14;75(8):1783-1795 [FREE Full text] [doi: 10.1093/geronb/gbz087] [Medline: 31251349]
36. Wang JT, Xu G, Ren RJ, Wang Y, Tang R, Huang Q, et al. The impacts of health insurance and resource on the burden of Alzheimer's disease and related dementias in the world population. Alzheimers Dement 2023 Mar 12;19(3):967-979. [doi: 10.1002/alz.12730] [Medline: 35820032]

Abbreviations

AANHPI: Asian Americans, Native Hawaiians, and Pacific Islanders
AARIN: Asian American Resource and Information Network
AD: Alzheimer disease
ADRD: Alzheimer disease and related dementias
KAB: knowledge, attitudes, and behaviors
KIMCHI: Koreans Invested in Making Caregivers’ Health Important
LEP: limited English proficiency
PCORI: Patient-Centered Outcomes Research Institute
REDCap: Research Electronic Data Capture
WHO: World Health Organization

Edited by A Garcia; submitted 07.02.25; peer-reviewed by R Venkatesh, C Bosco; comments to author 24.03.25; revised version received 19.05.25; accepted 03.06.25; published 15.08.25.

Please cite as:
Ta Park V, Nam B, Huang D, Yun SW, Phan N, Lee EJ, Shin HW
Knowledge, Attitudes, and Behaviors Related to Dementia Prevention and Caregiving Among Korean Americans (the KIMCHI Project): Pre- and Posttest Evaluation Study
JMIR Aging 2025;8:e72147
URL: https://aging.jmir.org/2025/1/e72147
doi:10.2196/72147
PMID:

XSL•FO
RenderX

©Van Ta Park, Bora Nam, Daren Huang, Stacy W Yun, Nicole Phan, Eun Jeong Lee, Hye-Won Shin. Originally published in JMIR Aging (https://aging.jmir.org), 15.08.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Associations Between Dietary Habits and Accelerated Aging and the Establishment of an Accelerated Aging Interpretable Risk Prediction Model via Shapley Additive Explanations: Cross-Sectional Study From Two Representative Populations

Zhengyang Wu[1,2,3,4*], MBBS; Ning Zhang[1,2,3*], MBBS; Haiwei Wang[4*], MBBS; Yang Yang[1,2,3], MBBS; Houhao Shen[4], MBBS; Qidi Zhang[4], MBBS; Yunxia Cao[1,2,3], MD; Yinan Du[4], PhD; Dongmei Ji[1,2,3], MD

[1]Department of Obstetrics and Gynecology, NHC Key Laboratory of Study on Abnormal Gametes and Reproductive Tract, First Affliated Hospital of Anhui Medical University, Hefei, China

[2]Engineering Research Center of Biopreservation and Artificial Organs, Ministry of Education, Hefei, China

[3]Key Laboratory of Population Health Across Life Cycle (Anhui Medical University), Ministry of Education of the People's Republic of China, Hefei, China

[4]School of Basic Medical Sciences, Anhui Medical University, Hefei, China

[*]these authors contributed equally

Corresponding Author:
Dongmei Ji, MD
Department of Obstetrics and Gynecology
NHC Key Laboratory of Study on Abnormal Gametes and Reproductive Tract
First Affliated Hospital of Anhui Medical University
No.218 Jixi Road
Hefei, 230022
China
Phone: 86 551 6590 8445
Email: jidongmei@ahmu.edu.cn

Abstract

Background: Research has revealed potential links between specific dietary habits and accelerated aging. However, most studies focus only on singular diets or lack ethnic diversity.

Objective: This study aimed to investigate the associations between 5 dietary indices and the risk of accelerated aging and develop an interpretable machine learning (ML) model for accelerated aging prediction.

Methods: We explored associations between dietary indices and the risk of accelerated aging using data from the US National Health and Nutrition Examination Survey (NHANES) and the UK Biobank. A weighted linear regression analysis was used to determine whether accelerated aging was linked to dietary habits, and the covariates were gradually adjusted to ensure that the association was stable. Nonlinear correlations were explored using restricted cubic spline curves. In addition, multiple ML algorithms were used to build predictive models of accelerated aging risk.

Results: Except for the Dietary Inflammation Index (β=0.35, 95% CI 0.23-0.74), the other 4 dietary indices (Alternative Healthy Eating Index, Alternative Mediterranean Diet, Healthy Eating Index-2020, and Dietary Approaches to Stop Hypertension) were negatively associated with the risk of accelerated aging in NHANES participants. Similar results were observed in UK Biobank participants. Nine ML algorithms were used to develop risk prediction models, among which the gradient boosting decision tree model showed the best overall performance. A web-based prediction platform was developed and made publicly available.

Conclusions: Significant associations between accelerated aging and dietary indices were observed. High compliance with the Dietary Inflammation Index had a promoting effect on accelerated aging, while high compliance with the Alternative Healthy Eating Index, Alternative Mediterranean Diet, Healthy Eating Index-2020, and Dietary Approaches to Stop Hypertension showed varying degrees of protection against accelerated aging.

(JMIR Aging 2025;8:e72020) doi:10.2196/72020

XSL•FO
RenderX

KEYWORDS

dietary indices; accelerated aging; machine learning; National Health and Nutrition Examination Survey; NHANES; UK Biobank

Introduction

Biological Aging and Its Measurement

Aging is a highly complex biological process, typically attributed to molecular alterations and the accumulation of detrimental biomarkers that compromise tissue and organ function and regeneration, ultimately resulting in disease and mortality [1-3]. However, chronological age does not always accurately reflect biological aging [4]. Accelerated aging occurs when biological age exceeds chronological age. Ideally, the measurement of biological aging should reflect aging status across multiple levels of biological systems. Current studies have proposed several methods for measuring biological aging, ranging from individual biomarkers (such as telomere length) to integrated algorithms from epigenetic, proteomic, metabolomic, and other omics data [5,6]. However, these methods are prohibitively expensive for practical use, and their accuracy remains contentious. In this context, Klemera and Doubal [7] and Levine et al [8] proposed new biological age algorithms, the Klemera-Doubal method (KDM) and the PhenoAge algorithm, respectively. These algorithms have since been widely used for predicting biological age and are considered among the most accurate measures currently available [9,10].

Dietary Patterns, Inflammation, and Accelerated Aging

Dietary habits directly influence the proportion of nutrients absorbed by the human body, thereby impacting endocrine processes and subsequently affecting normal metabolism. Several studies have shown that high-quality diets are associated with reduced risks of mortality, cancer, and cardiovascular disease [11-13]. Adherence to the typical Mediterranean diet was found to effectively reduce oxidative stress levels [12]. In addition, dietary indices such as the Dietary Approaches to Stop Hypertension (DASH), the Alternative Healthy Eating Index (AHEI), the Healthy Eating Index-2020 (HEI-2020), and the Dietary Inflammatory Index (DII) have been extensively studied and proven to effectively evaluate the healthiness of individuals' dietary habits [14]. Recent research has also explored the relationship between dietary indices and accelerated aging. One study showed that inflammatory diets could significantly accelerate the aging process [15]. In addition, a large cohort study suggested that healthy diet habits might mitigate the adverse effects of pollutants on aging [16,17]. However, previous studies on the relationship between dietary indices and accelerated aging have had limitations, such as a focus on specific dietary indices or a lack of diversity in cohort demographics. These constraints highlight the need for further research using comprehensive datasets and diverse dietary indices to provide more robust and generalizable insights.

In this study, we simultaneously analyzed data from 2 representative populations in Europe and the United States to assess the relationship between dietary indices and accelerated aging. Weighted multivariable-adjusted linear regression was used to explore associations between 5 dietary indices and accelerated aging. Subset analyses were conducted to ensure the robustness of the results. In addition, we developed a machine learning (ML)–based web-based prediction tool for estimating accelerated aging risk based on individual dietary habits.

Methods

Study Population

The National Health and Nutrition Examination Survey (NHANES) is a nationally representative, population-based survey designed to assess the health and nutritional status of adults and children in the United States. Data from 7 consecutive survey cycles from 2005 to 2018 were included. We excluded participants younger than 20 years, as recommended by the NHANES for adult analyses. The UK Biobank is a prospective cohort study that enrolled individuals across the United Kingdom between 2006 and 2010 [18]. At baseline, participants completed questionnaires covering socioeconomic status, lifestyle, and health history, followed by a nurse-led interview to verify health information and perform standardized physical measurements. We analyzed cross-sectional data from the baseline assessment of the UK Biobank. Individuals were excluded if they lacked essential dietary intake data, blood biochemical measurements, or key covariates. We further excluded individuals with extreme energy intake (male individuals: <600 or >4800 kcal/d; female individuals: <500 or >3600 kcal/d) to minimize potential bias from aberrant energy consumption. A total of 10,940 eligible participants from the NHANES and 8594 participants from the UK Biobank were included in the study.

Dietary Assessment and Calculation of Dietary Indices

We obtained detailed dietary intake information from NHANES participants to estimate energy, nutrient, and other food component consumption during the 24 hours preceding the interview. All participants were eligible for two 24-hour dietary recall interviews. The first interview was conducted in person at the NHANES mobile examination center, and the second was administered by telephone 3 to 10 days later. Only participants who completed both recalls were included in the analysis. Within the UK Biobank, participants recruited between 2006 and 2010 completed the Oxford WebQ dietary questionnaire during their baseline assessment visit. This questionnaire retrospectively assessed the consumption of up to 206 food items and 32 beverage types over the preceding 24 hours. The total intake of each food or beverage was calculated by multiplying the standard portion size by the reported frequency of consumption [19]. Only participants who completed at least 2 Oxford WebQ assessments were included in the analysis.

Dietary indices were calculated using the R package [20] *dietaryindex* [21]. This package can calculate various dietary indices of high interest in epidemiology, public health, nutrition, and clinical research, including the AHEI, DASH, HEI-2020, DII, and Alternative Mediterranean Diet (aMED). A detailed

XSL•FO
RenderX

description of these dietary indices is provided in Multimedia Appendix 1 [22-24].

Calculation of Accelerated Aging

KDM and PhenoAge were calculated using the R package *BioAge*. PhenoAge was calculated based on chronological age and 13 biomarkers. These biomarkers include albumin, alkaline phosphatase, blood urea nitrogen, creatinine, C-reactive protein, glycated hemoglobin, total cholesterol, uric acid, white blood cell count, lymphocyte percentage, mean cell volume, and red cell distribution width. PhenoAge represents an individual's expected age, consistent with the estimated mortality risk based on their biochemical biomarker profile. KDM models biological age as the average biological state associated with a specific chronological age in the reference population, assuming that biological age increases linearly over time [25]. The biochemical indicators required by the KDM and PhenoAge algorithms are more easily obtainable than those required by traditional DNA methylation–based aging algorithms [9,26,27].

Covariates

The survey data included demographic information, health status, and lifestyle factors, including age, sex (male or female), race (White, Black, or other), educational level, marital status, deprivation index, family size, BMI, history of hypertension, physical activity (PA), smoking history, alcohol consumption, and daily energy intake [28]. For NHANES participants, the deprivation index was categorized based on the poverty income ratio (PIR) as higher (most deprived; PIR <1.5), middle (3.5>PIR≥1.5), and lower (least deprived; PIR ≥3.5) [29]. PA was measured on the basis of the Global Physical Activity Questionnaire, which includes questions related to work-related, leisure-time, and sedentary activities. In the UK Biobank, participants' deprivation index was calculated based on the Townsend Deprivation Index, which was constructed from 4 key variables: unemployment, household overcrowding, lack of car ownership, and lack of home ownership. A higher index indicated a higher level of deprivation. PA was estimated as metabolic equivalent task (MET) minutes using items derived from the International Physical Activity Questionnaire [30]. For participants from both the NHANES and the UK Biobank, PA levels were classified according to MET minutes per week into 3 categories: inactive (≤499 MET-min/wk), moderately active (500-1554 MET-min/wk), and regular exercisers (≥1555 MET-min/wk) [25]. Observations with missing data for any covariate were excluded from the analysis.

Development of ML Models

Nine ML algorithms were used for estimating accelerated aging risk: gradient boosting decision tree (GBDT), random forest, decision tree, tabular prior-data fitted networks, logistic regression, light gradient boosting machine, categorical boosting, deep learning, and extreme gradient boosting based on the area under the receiver operating characteristic curve (AUROC) and the area under the precision-recall curve (AUPRC). We used the NHANES dataset as the training set for the model. Multivariate factors such as demographics, lifestyle, disease history, and dietary indices were used to predict accelerated aging risk. To validate the predictive effect, we used 70% of the data for model training and 30% for testing. To assess model stability and prediction accuracy, we conducted 5-fold cross-validation on the training set. Owing to imbalanced data distribution, we used synthetic minority oversampling technique for balancing. In addition to internal cross-validation, we included baseline participants from the UK Biobank dataset as external validation data. To apply the established early aging risk prediction model, a web-based application was developed using the Python *Streamlit* library.

Statistical Analysis

Group comparisons between participant characteristics and accelerated aging values were conducted using 2-tailed *t* tests for continuous variables and chi-square or Fisher exact tests for categorical variables, as appropriate. Participants' characteristics were presented as means and SDs for continuous variables and as percentages for categorical variables. Subsequently, the relationships between the 5 dietary indices (HEI-2020, aMED, DII, DASH, and AHEI) and accelerated aging (calculated by the PhenoAge and KDM algorithms) were assessed using linear regression, with adjustments for covariates. BMI, history of hypertension, educational level, and smoking history from the covariates were considered stratification variables to stratify participants. Further regression analysis was conducted to determine the impact of diet on accelerated aging across different strata. All analyses involving NHANES data in this study used survey weights for weighted analysis, balancing the unequal selection probabilities arising from the NHANES sampling survey design, to obtain nationally representative estimates. The study followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) reporting guideline. All data analyses and modeling processes were completed using Stata (version 17.0; StataCorp LLC), Python (version 3.10; Python Software Foundation), or R (version 4.3.0; R Foundation for Statistical Computing). Statistical significance was set at 2-tailed $P<.05$.

Ethical Considerations

NHANES investigators obtained appropriate ethical approvals for the publicly available data. The study was approved by the National Center for Health Statistics Ethics Review Board (Protocols: #2005-06, Continuation of Protocol #2005-06, #2011-17, Continuation of Protocol #2011-17, and #2018-01). The research was conducted in accordance with the Declaration of Helsinki. Informed consent was obtained by NHANES investigators from all participants involved in the research. Furthermore, the UK Biobank received ethics approval from the North West Multicenter Research Ethics Committee. Our study used UK Biobank resources (application 211452).

Results

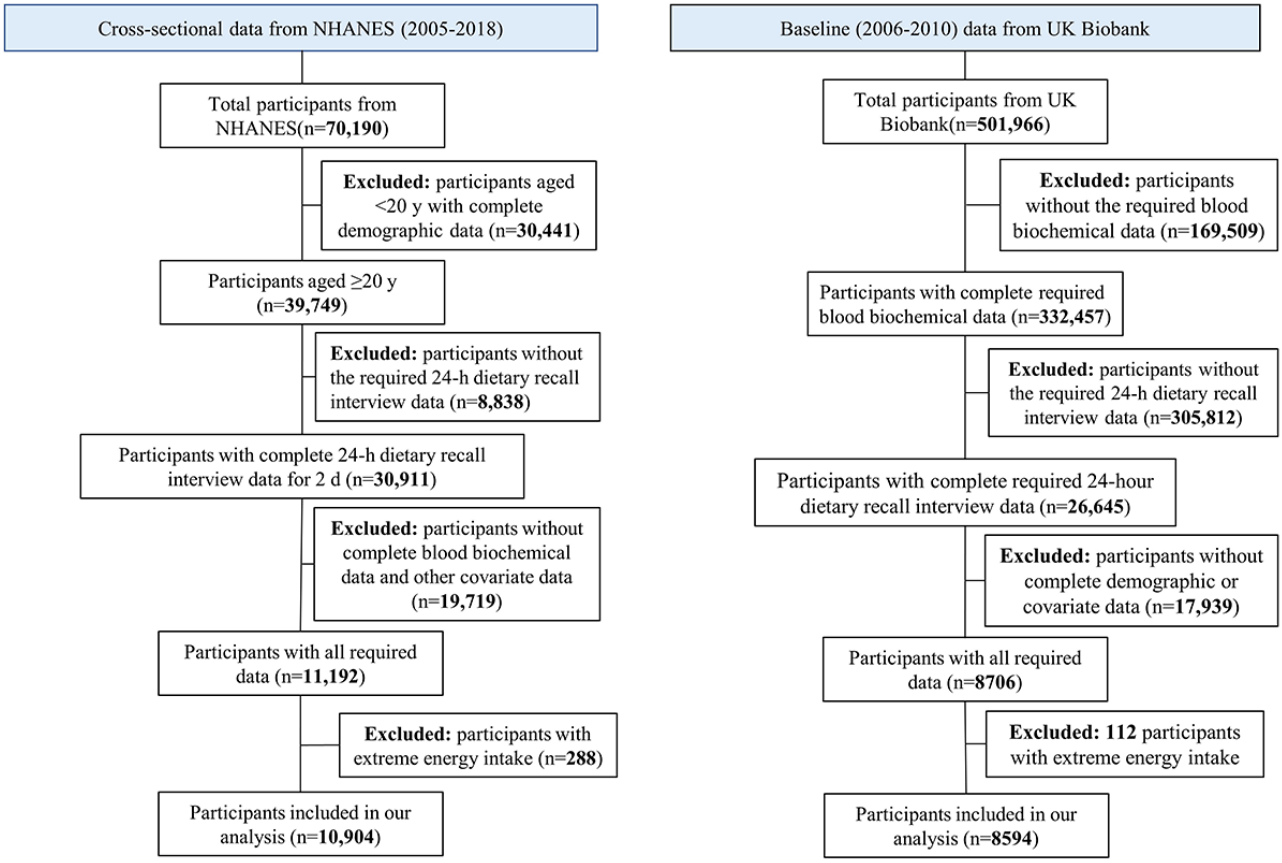
Participant Characteristics

From the NHANES population, 10,904 participants, representing an estimated 476,381,485 American adults, were included (Figure 1). The overall characteristics of the participants are summarized in Table 1. Their average age was 48.61 (SD 17.40) years; 52.49% (5724/10,904) were female, 72.48% (7903/10,904) were White, and 79.9% (8712/10,904) had an

XSL•FO
RenderX

educational level higher than high school. The mean age acceleration was –5.23 (SD 14.56) years, indicating that participants' KDM age was on average 5.23 years lower than their chronological age. Moreover, the mean senescence time for accelerated aging calculated with PhenoAge was −0.85 (SD 6.45) years. The scores for the 5 dietary indices approximately followed normal distributions (Figure S1 in Multimedia Appendix 1).

Figure 1. Participant selection flow chart.



XSL•FO
RenderX

Table 1. Demographic characteristics of US National Health and Nutrition Examination Survey participants.

Characteristics	Overall (n=10,904)	KDM[a] age		*P* value[b]	PhenoAge		*P* value[b]
		Accelerated aging (n=7258)	Nonaccelerated aging (n=3646)		Accelerated aging (n=6847)	Nonaccelerated aging (n=4057)	
Age (years), mean (SD)	48.61 (17.40)	51.09 (16.92)	43.68 (17.29)	*<.001* [c]	47.58 (17.11)	50.35 (17.74)	*<.001*
Sex, n (%)				.84			*<.001*
Female	5724 (52.49)	3805 (52.42)	1919 (52.63)		3821 (55.81)	1903 (46.91)	
Male	5180 (47.51)	3453 (47.58)	1727 (47.37)		3026 (44.19)	2154 (53.09)	
Race, n (%)				*<.001*			*<.001*
Black	2029 (18.61)	1150 (15.84)	879 (24.11)		1031 (15.06)	998 (24.6)	
White	7903 (72.48)	5458 (75.2)	2445 (67.06)		5203 (75.99)	2700 (66.55)	
Other	972 (8.91)	650 (8.96)	322 (8.83)		613 (8.95)	359 (8.85)	
Education, n (%)				*<.001*			*.002*
High school graduate or GED[d] or equivalent	1318 (12.09)	822 (11.33)	496 (13.6)		769 (11.23)	549 (13.53)	
Less than high school	874 (8.02)	633 (8.72)	241 (6.61)		558 (8.15)	316 (7.79)	
More than high school	8712 (79.9)	5803 (79.95)	2909 (79.79)		5520 (80.62)	3192 (78.68)	
Marital status, n (%)				*<.001*			*<.001*
Divorced or living alone	4194 (38.46)	2594 (35.74)	1600 (43.88)		2470 (36.07)	1724 (42.49)	
Other	6710 (61.54)	4664 (64.26)	2046 (56.12)		4377 (63.93)	2333 (57.51)	
Deprivation Index, n (%)				*<.001*			*<.001*
Higher (most deprived)	3686 (33.8)	2657 (36.61)	1029 (28.22)		2555 (37.32)	1131 (27.88)	
Middle	3689 (33.83)	2423 (33.38)	1266 (34.72)		2272 (33.18)	1417 (34.93)	
Lower (least deprived)	3529 (32.36)	2178 (30.01)	1351 (37.05)		2020 (29.5)	1509 (37.19)	
Family size, n (%)				*<.001*			.06
1-3	6914 (63.41)	4728 (65.14)	2186 (59.96)		4287 (62.61)	2627 (64.75)	
4-5	2923 (26.81)	1862 (25.65)	1061 (29.1)		1887 (27.56)	1036 (25.54)	
≥6	1067 (9.79)	668 (9.2)	399 (10.94)		673 (9.83)	394 (9.71)	
BMI (kg/m^2, n (%)				*<.001*			*<.001*
<20	447 (4.1)	355 (4.89)	92 (2.52)		334 (4.88)	113 (2.79)	
20-25	2550 (23.39)	1922 (26.48)	628 (17.22)		1896 (27.69)	654 (16.12)	
>25	7907 (72.51)	4981 (68.63)	2926 (80.25)		4617 (67.43)	3290 (81.09)	
Hypertension, n (%)				*<.001*			*<.001*
No	7221 (66.22)	5014 (69.08)	2207 (60.53)		4939 (72.13)	2282 (56.25)	
Yes	3683 (33.78)	2244 (30.92)	1439 (39.47)		1908 (27.87)	1775 (43.75)	
Physical activity, n (%)				*<.001*			*<.001*
Exercise regularly	5099 (46.76)	3242 (44.67)	1857 (50.93)		3267 (47.71)	1832 (45.16)	
Moderately active	1930 (17.7)	1365 (18.81)	565 (15.5)		1289 (18.83)	641 (15.8)	
Inactive	3875 (35.54)	2651 (36.53)	1224 (33.57)		2291 (33.46)	1584 (39.04)	
Smoking status, n (%)				*<.001*			*<.001*
Current smoker	2142 (19.64)	1297 (17.87)	845 (23.18)		1097 (16.02)	1045 (25.76)	
Ever smoker	2598 (23.83)	1831 (25.23)	767 (21.04)		1595 (23.29)	1003 (24.72)	
Never smoker	6164 (56.53)	4130 (56.9)	2034 (55.79)		4155 (60.68)	2009 (49.52)	

XSL•FO
RenderX

Characteristics	Overall (n=10,904)	KDM[a] age		*P* value[b]	PhenoAge		*P* value[b]
		Accelerated aging (n=7258)	Nonaccelerated aging (n=3646)		Accelerated aging (n=6847)	Nonaccelerated aging (n=4057)	
Alcohol consumption, n (%)				*<.001*			*.049*
No	1481 (13.58)	1044 (14.38)	437 (11.99)		964 (14.08)	517 (12.74)	
Yes	9423 (86.42)	6214 (85.62)	3209 (88.01)		5883 (85.92)	3540 (87.26)	
Daily energy intake (kcal), mean (SD)	2010.68 (708.38)	2003.76 (701.24)	2024.44 (722.28)	.17	2011.27 (699.13)	2009.67 (723.81)	.80
PhenoAge (years), mean (SD)	−0.85 (6.45)	−2.33 (5.63)	2.09 (6.96)	*<.001*	−4.66 (3.00)	5.59 (5.53)	*<.001*
KDM age (years), mean (SD)	−5.23 (14.56)	−13.05 (9.22)	10.36 (10.01)	*<.001*	−8.90 (12.35)	0.97 (15.85)	*<.001*
AHEI[e] scores, mean (SD)	38.82 (11.63)	40.00 (11.75)	36.46 (11.02)	*<.001*	40.06 (11.69)	36.72 (11.22)	*<.001*
aMED[f] scores, mean (SD)	3.51 (1.38)	3.65 (1.39)	3.23 (1.33)	*<.001*	3.67 (1.38)	3.23 (1.34)	*<.001*
DASH[g] scores, mean (SD)	22.49 (5.06)	23.08 (5.07)	21.31 (4.84)	*<.001*	23.17 (5.03)	21.34 (4.90)	*<.001*
DII[h] scores, mean (SD)	1.15 (1.68)	1.04 (1.71)	1.37 (1.60)	*<.001*	1.01 (1.71)	1.38 (1.62)	*<.001*
HEI-2020[i] scores, mean (SD)	51.58 (11.97)	52.77 (11.97)	49.20 (11.61)	*<.001*	52.83 (11.92)	49.46 (11.75)	*<.001*

[a]KDM: Klemera-Doubal method.

[b]Wilcoxon rank sum test for continuous variables; Pearson chi-square test for categorical variables.

[c]Italicization indicates values that met the threshold for significance (*P*<.05).

[d]GED: General Educational Development Test.

[e]AHEI: Alternative Healthy Eating Index.

[f]aMED: Alternative Mediterranean Diet.

[g]DASH: Dietary Approaches to Stop Hypertension.

[h]DII: Dietary Inflammatory Index.

[i]HEI-2020: Healthy Eating Index-2020.

From the UK Biobank cohort, 8594 participants were included, of whom 54.56% (n=4689) were female, 97.68% (n=8395) had White European ancestry, and 64.96% (n=8594) had attained higher education (Table S1 in Multimedia Appendix 1). The participants' PhenoAge scores were on average 5.47 years lower than their chronological age, whereas the average age acceleration derived from KDM age was 17.16 years. Compared with the NHANES participants, those in the UK Biobank were older, had higher educational attainment, and exhibited lower BMI levels.

In addition, we noticed an overall positive correlation among the DASH, AHEI, HEI-2020, and aMED scores, except for the DII scores, which exhibited negative correlations with other indices, in both the NHANES and the UK Biobank. Among the NHANES participants, the strongest association was between the DASH and AHEI scores (weighted Pearson *r*=0.85), whereas among the UK Biobank participants, the strongest association was between the DII and AHEI scores (Pearson *r*=−0.78; Figure S2 in Multimedia Appendix 1).

Correlations Between Different Dietary Indices and Accelerated Aging

For the analysis of NHANES participants, we constructed 3 weighted linear regression models and adjusted the confounders to find the correlations between different dietary indices and accelerated aging. Model 1 was only adjusted for age, sex, and race; model 2 was adjusted for model 1's covariates, family size, educational level, deprivation index, and marital status; and model 3 was adjusted for model 2's covariates, BMI, smoking status, alcohol consumption, PA, daily energy intake, and history of hypertension.

In model 1, the AHEI, DASH, aMED, and HEI-2020 scores were negatively correlated with accelerated aging calculated by both PhenoAge and KDM, whereas the DII scores were positively correlated with them (DII score for PhenoAge: β=0.35, 95% CI 0.23 to 0.46; DII score for KDM: β=0.55, 95% CI 0.36 to 0.74; Table 2). In models 2 and 3, the associations between the dietary indices and accelerated aging calculated by both PhenoAge and KDM remained significant (*P*<.05; Table 2). These results revealed correlations between the different dietary indices and accelerated aging.

XSL•FO
RenderX

Table 2. Weighted linear regression analysis of dietary indices and accelerated aging among US National Health and Nutrition Examination Survey participants.

Dietary indices	Model 1[a]				Model 2[b]				Model 3[c]			
	PhenoAge		KDM[d] age		PhenoAge		KDM age		PhenoAge		KDM age	
	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value
AHEI[e]	–0.09 (–0.11 to –0.07)	<.001	–0.16 (–0.19 to –0.13)	<.001	–0.08 (–0.09 to –0.06)	<.001	–0.15 (–0.18 to –0.12)	<.001	–0.06 (–0.08 to –0.04)	<.001	–0.11 (–0.13 to –0.08)	<.001
aMED[c]	–0.83 (–0.96 to –0.69)	<.001	–1.51 (–1.71 to –1.31)	<.001	–0.74 (–0.87 to –0.61)	<.001	–1.43 (–1.65 to –1.22)	<.001	–0.57 (–0.70 to –0.44)	<.001	–1.07 (–1.28 to –0.86)	<.001
DASH[d]	–0.24 (–0.27 to –0.20)	<.001	–0.44 (–0.50 to –0.38)	<.001	–0.21 (–0.25 to –0.18)	<.001	–0.42 (–0.49 to –0.36)	<.001	–0.16 (–0.20 to –0.13)	<.001	–0.32 (–0.38 to –0.26)	<.001
DII[e]	.53 (0.41 to 0.64)	<.001	.84 (0.65 to 1.03)	<.001	.45 (0.34 to 0.57)	<.001	.76 (0.56 to 0.96)	<.001	.35 (0.23 to 0.46)	<.001	.55 (0.36 to 0.74)	<.001
HEI-2020[f]	–0.09 (–0.10 to –0.07)	<.001	–0.15 (–0.17 to –0.13)	<.001	–0.08 (–0.10 to –0.06)	<.001	–0.14 (–0.17 to –0.12)	<.001	–0.06 (–0.08 to –0.05)	<.001	–0.10 (–0.13 to –0.08)	<.001

[a]Adjusted for age, sex, and race.

[b]Model 1+adjusted for family size, educational level, deprivation index, and marital status.

[c]Model 2+adjusted for BMI, smoking status, alcohol consumption, physical activity, daily energy intake, and history of hypertension.

[d]KDM: Klemera-Doubal method.

[e]AHEI: Alternative Healthy Eating Index.

[f]aMED: Alternative Mediterranean Diet.

[g]DASH: Dietary Approaches to Stop Hypertension.

[h]DII: Dietary Inflammatory Index.

[i]HEI-2020: Healthy Eating Index-2020.

In the analysis of UK Biobank participants too, we conducted linear regression analysis using 3 models, with covariate adjustments identical to those applied in the NHANES analyses. The results showed that the associations of AHEI, DASH, aMED, and DII scores with age acceleration were consistent with those observed in the NHANES analyses, whereas HEI-2020 scores showed no significant association with age acceleration in any of the 3 models (DII score for PhenoAge: *P*=.85; DII score for KDM: *P*=.83; Table S2 in Multimedia Appendix 1). In the fully adjusted model 3, higher aMED scores showed the strongest protective association with age acceleration, regardless of whether the KDM or PhenoAge algorithm was used (aMED score for PhenoAge: β=−0.08, 95% CI −0.15 to −0.01; aMED score for KDM: β=−0.39, 95% CI −0.55 to −0.23; Table S2 in Multimedia Appendix 1).

Correlations Between Quartiles of 5 Dietary Indices and Accelerated Aging

In NHANES participants, linear regression analysis across quartiles of each of the 5 dietary indices showed significant correlations with accelerated aging, as calculated by both PhenoAge and KDM, with DII scores being positively correlated and the scores of the other 4 dietary indices being negatively correlated (Table 3). All correlations exhibited strong linear trends (*P* for trend<.001). We observed similar results in the analysis of UK Biobank participants, but the association between HEI-2020 scores and age acceleration was not significant (Table S3 in Multimedia Appendix 1).

XSL•FO
RenderX

Table 3. Quartile regression and linear trend analysis among US National Health and Nutrition Examination Survey participants.

Dietary indices and variables	PhenoAge		KDM[a] age	
Characteristic	β (95% CI)	*P* value	β (95% CI)	*P* value
AHEI[b]				
Q1	Reference	—[c]	Reference	—
Q2	–0.74 (–1.10 to –0.37)	<.001[d]	–2.24 (–2.94 to –1.55)	<.001
Q3	–1.04 (–1.42 to –0.67)	<.001	–3.62 (–4.64 to –2.60)	<.001
Q4	–1.52 (–2.04 to –1.00)	<.001	–5.42 (–6.37 to –4.47)	<.001
P for linear trend	—	<.001	—	<.001
aMED[e]				
Q1	Reference	—	Reference	—
Q2	–0.22 (–0.61 to 0.16)	.232	–1.27 (–2.33 to –0.20)	<.001
Q3	–0.67 (–1.09 to –0.24)	.002	–3.79 (–4.88 to –2.70)	<.001
Q4	–0.99 (–1.52 to –0.46)	<.001	–5.04 (–6.03 to –4.05)	<.001
P for linear trend	—	<.001	—	<.001
DASH[d]				
Q1	Reference	—	Reference	—
Q2	–0.99 (–1.43 to –0.55)	<.001	–2.80 (–3.88 to –1.73)	<.001
Q3	–1.23 (–1.63 to –0.82)	<.001	–4.09 (–4.86 to –3.31)	<.001
Q4	–1.73 (–2.24 to –1.23)	<.001	–6.76 (–7.72 to –5.80)	<.001
P for linear trend	—	<.001	—	<.001
HEI-2020[e]				
Q1	Reference	—	Reference	—
Q2	–0.75 (–1.08 to –0.42)	<.001	–2.08 (–2.98 to –1.17)	<.001
Q3	–1.14 (–1.59 to –0.69)	<.001	–3.60 (–4.52 to –2.67)	<.001
Q4	–1.5 (–2.00 to –0.99)	<.001	–5.13 (–6.11 to –4.14)	<.001
P for linear trend	—	<.001	—	<.001
DII[f]				
Q1	Reference	—	Reference	—
Q2	0.44 (0.07 to 0.80)	.012	1.57 (0.69 to 2.45)	.003
Q3	0.83 (0.33 to 1.32)	<.001	3.09 (2.11 to 4.06)	<.001
Q4	1.11 (0.55 to 1.67)	<.001	2.26 (1.38 to 3.15)	<.001
P for linear trend	—	<.001	—	<.001

[a]KDM: Klemera-Doubal method.

[b]AHEI: Alternative Healthy Eating Index.

[c]Not applicable.

[d]Italicization indicates values that met the threshold for significance (*P*<.05).

[e]aMED: Alternative Mediterranean Diet.

[f]DASH: Dietary Approaches to Stop Hypertension.

[g]HEI-2020: Healthy Eating Index-2020.

[h]DII: Dietary Inflammatory Index.

XSL•FO
RenderX

Subgroup Analysis

To ascertain whether the association between accelerated aging and dietary indices was consistent across the general population, we conducted analyses stratified by BMI, history of hypertension, educational level, and smoking history. First, we performed the subgroup analyses in people with different BMI levels, examining the relationships between the dietary indices and accelerated aging calculated by both PhenoAge and KDM in NHANES participants. In the analysis of KDM age acceleration, higher aMED and DASH scores exhibited stronger protective effects in individuals who were obese (*P* for interaction aMED=.02; *P* for interaction DASH=.01), a finding that is consistent with the results from the UK Biobank analysis (interaction aMED: *P*=.002; interaction DASH: *P*=.03; Figure S3 in Multimedia Appendix 1). Subgroup analyses across other stratifications revealed no significant between-group heterogeneity (Figures S4-S10 in Multimedia Appendix 1).

Nonlinear Relationship and Inflection Point in the Effects of Dietary Indices on Accelerated Aging

We used restricted cubic spline curves to explore the relationships between accelerated aging and the dietary indices. Our results remained unchanged after both primary and additional adjustments for potential confounders; therefore, only the most adjusted model (model 3) is reported. In the analysis of NHANES participants, we observed highly significant linear trends in the associations between the dietary indices and accelerated aging, quantified using both PhenoAge and KDM (overall association: *P*<.001; nonlinearity: *P*>.05). We observed similar results among participants in the UK Biobank, except that HEI-2020 scores showed no significant linear association with the 2 accelerated aging metrics in the UK Biobank analysis (Figures S11-S12 in Multimedia Appendix 1).

Testing the Performance of ML Models in Identifying Accelerated Aging

From the previous analysis, we determined that there was a significant correlation between dietary indices and accelerated aging. We constructed an ML model and used Shapley additive explanations (SHAP) values to elucidate the relative importance of key predictive features. The NHANES data were split into training and testing sets in a 0.7 ratio, while the UK Biobank data were used as an external validation cohort. Model performance was quantified using the area under the receiver operating characteristic curve, while the area under the precision-recall curve characterized the precision-recall trade-off across classification thresholds. We evaluated 9 ML algorithms, including GBDT; categorical boosting achieved the highest AUROC (0.729) in the training set, whereas tabular prior-data fitted networks yielded the best performance (AUROC=0.757) in the test set (Figure S13 in Multimedia Appendix 1; Table S4 in Multimedia Appendix 1). However, to prioritize the stability, generalization, and interpretability of the model, we chose GBDT as the final modeling algorithm. This algorithm performed stably and well in both the training and test sets. The precision-recall curve (AUPRC=0.976) for the test set indicated very good prediction performance on imbalanced samples (Figure S14 in Multimedia Appendix 1). Specifically, the learning rate of this model was set to 0.1 to balance convergence and generalization. The maximum number of iterations of the learner was 58, and the maximum tree depth was 4 to constrain model complexity. The minimum segmentation requirement for each internal node was 50 samples to reduce overfitting. All experiments were conducted with fixed random seeds to ensure reproducibility. Furthermore, we evaluated the inclusion of different dietary indices to assess their influence on model performance. The hyperparameter information of other models is presented in Table S5 in Multimedia Appendix 1. The results demonstrated that the inclusion of the AHEI enhanced model performance in both the training (AUROC: 0.711 vs 0.682) and test (AUROC: 0.706 vs 0.680) sets. Incorporating aMED, DII, DASH, or HEI-2020 individually improved training set performance but compromised the model's generalizability (Figure S15 in Multimedia Appendix 1). A web-based platform was developed and can be accessed [31].

Interpretation of Personalized Predictions

The SHAP plot reveals the effect of each variable in the model on identifying accelerated aging cases within the test dataset. The honeycomb diagram (Figure S16A in Multimedia Appendix 1) arranges data points horizontally to visualize feature value distributions and their corresponding contributions to model predictions. Points positioned to the right reflect increased positive contributions to the prediction, whereas those to the left indicate stronger negative influences. In the feature importance bar chart (Figure S16B in Multimedia Appendix 1), the bar length is proportional to the relative importance of each feature. The analysis identified age, marital status, and AHEI score as the 3 most influential features in the model. In addition, force plots were generated separately for participants with accelerated aging and those without (Figures S16C and S16D in Multimedia Appendix 1). The bold value denotes the model's predicted log-odds, f(x), for an individual, whereas the base value represents the expected prediction in the absence of any feature inputs; f(x) corresponds to the log-odds of accelerated aging for each observation. Features shown in red increase the predicted risk of accelerated aging, whereas those shown in blue decrease it. Arrow length visually encodes the magnitude of each feature's contribution: longer arrows indicate stronger effects on the prediction.

Discussion

Previous research has explored the associations between various dietary patterns and diseases such as cancer, hypertension, and Alzheimer disease [27,32,33]. This study found that 5 dietary indices had differing impacts on the acceleration of aging in the NHANES cohorts. DII scores were positively correlated with accelerated aging, while the AHEI, aMED, HEI-2020 and DASH scores were negatively correlated. Furthermore, we rigorously evaluated the final model's performance using training, internal validation, and external validation cohorts comprising diverse populations from the United States and Europe. Integrated analyses of ROC and precision-recall curves unequivocally demonstrated the model's superior predictive accuracy.

Diet is a crucial and easily modifiable risk factor for various chronic diseases [34]. Increasing evidence suggests that different

XSL•FO
RenderX

dietary indices impact aging to varying degrees. The aMED and the DASH diet have been shown to have anti-inflammatory and cardioprotective effects, which may contribute to preventing dementia associated with neuroinflammation [35,36]. Our cross-sectional findings further suggest that a healthier dietary pattern could significantly reduce the risk of accelerated aging. Diet can regulate the immune system to some extent. Evidence suggests that the intake of certain nutrients and bioactive components can affect the process of neuroinflammation in animals, thereby affecting aging [37]. In addition, long-term consumption of a proinflammatory diet may reduce gut microbiota diversity, particularly decreasing the abundance of butyrate-producing *Firmicutes* [37,38]. Furthermore, interactions between the gut microbiota and the host immune system may result in prolonged low-grade inflammatory exposure in the host, thereby promoting the development of accelerated aging [39]. This finding is further corroborated by the positive association between the DII scores and an increased risk of accelerated aging observed in our study.

An important part of healthy aging involves screening for nutritional intake as well as developing and improving lifestyle intervention measures to support health behavior changes focused on aging in different populations [40]. In addition to dietary intake, BMI is considered an important indicator of health in older adults, with obesity being a major risk factor for unhealthy aging [41]. In subgroup analyses, we also observed that maintaining a good diet was more protective against the risk of accelerated aging in individuals with higher levels of obesity.

Subgroup analyses revealed that individuals with higher degrees of obesity exhibited increased susceptibility to both the age-accelerating effects of high DII eating patterns and the age-mitigating impacts of elevated DASH, aMED, HEI-2020, and AHEI scores. We hypothesize that this heightened sensitivity may stem from the generally less healthful dietary habits associated with obesity [42]. Interestingly, similar trends (*P* values for all 5 dietary indices for both PhenoAge and KDM age were <.05) were observed among nonsmokers and participants with lower educational attainment.

Currently, no widely recognized predictive model exists for accelerated aging. Previous studies have often focused on the predictive power of accelerated aging for common chronic diseases [43-45]. Among existing models, the one developed by Liu et al [46], based on free tri-iodothyronine, performs best, but it requires the collection of free tri-iodothyronine data from participants, which increases costs and hinders widespread adoption of the model. Several other predictive models have limited sample sizes, lack external validation, and exhibit AUROC values lower than those of our model [47,48]. The predictive model developed in this study integrates dietary indices and common demographic data; it achieves high prediction accuracy and enables data collection via questionnaires, demonstrating strong clinical feasibility. From a clinical application perspective, this web-based tool can be widely adopted in community primary care settings, allowing health care providers to rapidly perform initial screening for accelerated aging risk among residents or individuals undergoing routine health examinations. Especially during routine community health check-ups, medical staff can conveniently and quickly identify individual risk factors based on SHAP values provided by the web-based tool and offer corresponding recommendations.

The first strength of our study is that we included multiple biological age and dietary indices for comparison and incorporated sociological factors to stratify the population for further analysis among different groups, which verified the stability of the research results [15]. This study helps to further establish the connection between dietary health and accelerated aging, linking an essential factor of daily life, diet, with aging, and provides effective dietary adjustment strategies for aging prevention. In addition, many covariates related to lifestyle habits or personal health status were considered in our study, such as BMI, smoking history, history of hypertension, and PA.

Our study also has some limitations. First, the early aging risk prediction model developed in this study exhibited relatively poor predictive efficiency for positive results in the confusion matrix analysis due to a certain degree of bias in the training samples. Second, this is a cross-sectional study and could demonstrate an association between dietary habits and accelerated aging, but it could not provide evidence of causation. Further research is needed to elucidate the association and underlying mechanisms between eating habits and accelerated aging. Moreover, the cross-sectional nature of this study inherently limits the model's predictive capacity and causal inference. We will conduct further cohort studies and use follow-up data to further test and optimize the performance of our model.

In summary, the findings from this cohort study of American and European adults suggest that healthy dietary habits may be associated with a lower risk of accelerated aging. However, further studies on the effects of dietary habits on accelerated aging are needed to determine the long-term effects of dietary habits.

Acknowledgments

This work was funded by the Science Fund for Distinguished Young Scholars of Anhui province (grant 2022AH020072), the Center for Big Data and Population Health of IHM (Institute of Health and Medicine; grant JKS2023015), the National College Students' Innovation and Entrepreneurship Training Program (grant 202410366002), Key Research Program of Anhui Science and Technology Innovation Platform (202305a12020016), University Natural Foundation of Anhui Educational Committee (2022AH010072) and the Early Contact with Research Training Program of Anhui Medical University (grant 2023-ZQKY-075). The authors thank the Undergraduate Interdisciplinary Medical Research Association of Anhui Medical University for providing valuable guidance, support, research opportunities, and resources.

XSL•FO
RenderX

Data Availability

The data used in this study are available from the National Center for Health Statistics of the Centers for Disease Control and Prevention [49].

Authors' Contributions

ZW, NZ, and HW were responsible for data curation, software, formal analysis, investigation, validation, visualization, and writing the original draft. YY, HS, and QZ were responsible for data curation, formal analysis, validation, software, and visualization. DJ, YD, and YC were responsible for conceptualization, funding acquisition, investigation, methodology, project administration, resources, supervision, validation, and reviewing and editing the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Subgroup analyses and details on model development for this study.
[DOCX File , 6772 KB - aging_v8i1e72020_app1.docx]

References

1. Kennedy BK, Berger SL, Brunet A, Campisi J, Cuervo AM, Epel ES, et al. Geroscience: linking aging to chronic disease. Cell 2014 Nov 06;159(4):709-713 [FREE Full text] [doi: 10.1016/j.cell.2014.10.039] [Medline: 25417146]
2. Kirkwood TB. Understanding the odd science of aging. Cell 2005 Feb 25;120(4):437-447 [FREE Full text] [doi: 10.1016/j.cell.2005.01.027] [Medline: 15734677]
3. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. Cell 2013 Jun 06;153(6):1194-1217 [FREE Full text] [doi: 10.1016/j.cell.2013.05.039] [Medline: 23746838]
4. Sebastiani P, Thyagarajan B, Sun F, Schupf N, Newman AB, Montano M, et al. Biomarker signatures of aging. Aging Cell 2017 Apr;16(2):329-338 [FREE Full text] [doi: 10.1111/acel.12557] [Medline: 28058805]
5. Ferrucci L, Gonzalez-Freire M, Fabbri E, Simonsick E, Tanaka T, Moore Z, et al. Measuring biological aging in humans: a quest. Aging Cell 2020 Feb;19(2):e13080 [FREE Full text] [doi: 10.1111/acel.13080] [Medline: 31833194]
6. Kothari M, Belsky DW. Unite to predict. Elife 2021 Feb 12;10:e66223 [FREE Full text] [doi: 10.7554/eLife.66223] [Medline: 33576740]
7. Klemera P, Doubal S. A new approach to the concept and computation of biological age. Mech Ageing Dev 2006 Mar;127(3):240-248. [doi: 10.1016/j.mad.2005.10.004] [Medline: 16318865]
8. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, et al. An epigenetic biomarker of aging for lifespan and healthspan. Aging (Albany NY) 2018 Apr 18;10(4):573-591 [FREE Full text] [doi: 10.18632/aging.101414] [Medline: 29676998]
9. Belsky DW, Moffitt TE, Cohen AA, Corcoran DL, Levine ME, Prinz JA, et al. Eleven telomere, epigenetic clock, and biomarker-composite quantifications of biological aging: do they measure the same thing? Am J Epidemiol 2018 Jun 01;187(6):1220-1230 [FREE Full text] [doi: 10.1093/aje/kwx346] [Medline: 29149257]
10. Li X, Ploner A, Wang Y, Magnusson PK, Reynolds C, Finkel D, et al. Longitudinal trajectories, correlations and mortality associations of nine biological ages across 20-years follow-up. Elife 2020 Feb 11;9:e51507 [FREE Full text] [doi: 10.7554/eLife.51507] [Medline: 32041686]
11. Jun S, Cowan AE, Dodd KW, Tooze JA, Gahche JJ, Eicher-Miller HA, et al. Association of food insecurity with dietary intakes and nutritional biomarkers among US children, National Health and Nutrition Examination Survey (NHANES) 2011-2016. Am J Clin Nutr 2021 Sep 01;114(3):1059-1069 [FREE Full text] [doi: 10.1093/ajcn/nqab113] [Medline: 33964856]
12. Nani A, Murtaza B, Sayed Khan A, Khan NA, Hichami A. Antioxidant and anti-inflammatory potential of polyphenols contained in Mediterranean diet in obesity: molecular mechanisms. Molecules 2021 Feb 12;26(4):985 [FREE Full text] [doi: 10.3390/molecules26040985] [Medline: 33673390]
13. Reedy J, Krebs-Smith SM, Miller PE, Liese AD, Kahle LL, Park Y, et al. Higher diet quality is associated with decreased risk of all-cause, cardiovascular disease, and cancer mortality among older adults. J Nutr 2014 Jun;144(6):881-889 [FREE Full text] [doi: 10.3945/jn.113.189407] [Medline: 24572039]
14. Varraso R, Chiuve SE, Fung TT, Barr RG, Hu FB, Willett WC, et al. Alternate Healthy Eating Index 2010 and risk of chronic obstructive pulmonary disease among US women and men: prospective study. BMJ 2015 Feb 03;350:h286 [FREE Full text] [doi: 10.1136/bmj.h286] [Medline: 25649042]
15. Wang X, Sarker SK, Cheng L, Dang K, Hu J, Pan S, et al. Association of dietary inflammatory potential, dietary oxidative balance score and biological aging. Clin Nutr 2024 Jan;43(1):1-10. [doi: 10.1016/j.clnu.2023.11.007] [Medline: 37992632]

XSL•FO
RenderX

16. Xu X, Zheng J, Li J, Shen Y, Zhu L, Jin Y, et al. Phthalate exposure and markers of biological aging: the mediating role of inflammation and moderating role of dietary nutrient intake. Ecotoxicol Environ Saf 2024 Aug;281:116649 [FREE Full text] [doi: 10.1016/j.ecoenv.2024.116649] [Medline: 38954910]
17. Yang Z, Pu F, Cao X, Li X, Sun S, Zhang J, et al. Does healthy lifestyle attenuate the detrimental effects of urinary polycyclic aromatic hydrocarbons on phenotypic aging? An analysis from NHANES 2001-2010. Ecotoxicol Environ Saf 2022 Jun 01;237:113542 [FREE Full text] [doi: 10.1016/j.ecoenv.2022.113542] [Medline: 35468442]
18. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015 Mar 31;12(3):e1001779 [FREE Full text] [doi: 10.1371/journal.pmed.1001779] [Medline: 25826379]
19. Perez-Cornago A, Pollard Z, Young H, van Uden M, Andrews C, Piernas C, et al. Description of the updated nutrition calculation of the Oxford WebQ questionnaire and comparison with the previous version among 207,144 participants in UK Biobank. Eur J Nutr 2021 Oct;60(7):4019-4030 [FREE Full text] [doi: 10.1007/s00394-021-02558-4] [Medline: 33956230]
20. Kwon D, Belsky DW. A toolkit for quantification of biological age from blood chemistry and organ function test data: BioAge. Geroscience 2021 Dec;43(6):2795-2808 [FREE Full text] [doi: 10.1007/s11357-021-00480-5] [Medline: 34725754]
21. Zhan JJ, Hodge RA, Dunlop AL, Lee MM, Bui L, Liang D, et al. Dietaryindex: a user-friendly and versatile R package for standardizing dietary pattern analysis in epidemiological and clinical studies. bioRxiv. Preprint posted online on August 09, 2023 2023:-. [doi: 10.1016/j.ajcnut.2024.08.021]
22. Kresovich JK, Park YM, Keller JA, Sandler DP, Taylor JA. Healthy eating patterns and epigenetic measures of biological age. Am J Clin Nutr 2022 Jan 11;115(1):171-179 [FREE Full text] [doi: 10.1093/ajcn/nqab307] [Medline: 34637497]
23. Chiuve SE, Fung TT, Rimm EB, Hu FB, McCullough ML, Wang M, et al. Alternative dietary indices both strongly predict risk of chronic disease. J Nutr 2012 Jun;142(6):1009-1018 [FREE Full text] [doi: 10.3945/jn.111.157222] [Medline: 22513989]
24. Zhan JJ, Hodge RA, Dunlop AL, Lee MM, Bui L, Liang D, et al. Dietaryindex: a user-friendly and versatile R package for standardizing dietary pattern analysis in epidemiological and clinical studies. bioRxiv. Preprint posted online on August 9, 2023 2023 Aug 09 [FREE Full text] [doi: 10.1101/2023.08.07.548466] [Medline: 37609152]
25. Gao X, Geng T, Jiang M, Huang N, Zheng Y, Belsky DW, et al. Accelerated biological aging and risk of depression and anxiety: evidence from 424,299 UK Biobank participants. Nat Commun 2023 Apr 20;14(1):2277 [FREE Full text] [doi: 10.1038/s41467-023-38013-7] [Medline: 37080981]
26. Graf GH, Crowe CL, Kothari M, Kwon D, Manly JJ, Turney IC, et al. Testing Black-White disparities in biological aging among older adults in the United States: analysis of DNA-methylation and blood-chemistry methods. Am J Epidemiol 2022 Mar 24;191(4):613-625 [FREE Full text] [doi: 10.1093/aje/kwab281] [Medline: 34850809]
27. Liu Z, Kuo PL, Horvath S, Crimmins E, Ferrucci L, Levine M. A new aging measure captures morbidity and mortality risk across diverse subpopulations from NHANES IV: a cohort study. PLoS Med 2018 Dec 31;15(12):e1002718 [FREE Full text] [doi: 10.1371/journal.pmed.1002718] [Medline: 30596641]
28. Cao C, Friedenreich CM, Yang L. Association of daily sitting time and leisure-time physical activity with survival among US cancer survivors. JAMA Oncol 2022 Mar 01;8(3):395-403 [FREE Full text] [doi: 10.1001/jamaoncol.2021.6590] [Medline: 34989765]
29. Wang D, Morton JI, Salim A, Shaw JE, Magliano DJ. Association between diabetes and anemia: evidence from NHANES and the UK Biobank. Diabetes Care 2025 May 01;48(5):816-826. [doi: 10.2337/dc24-2535] [Medline: 40116818]
30. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
31. Machine learning-based premature aging prediction platform. Streamlit. URL: http://mitusml.com:9538/ [accessed 2025-05-29]
32. Banach M, Mazidi M, Mikhailidis DP, Toth PP, Jozwiak J, Rysz J, et al. Association between phenotypic familial hypercholesterolaemia and telomere length in US adults: results from a multi-ethnic survey. Eur Heart J 2018 Oct 21;39(40):3635-3640 [FREE Full text] [doi: 10.1093/eurheartj/ehy527] [Medline: 30165413]
33. Ge Y. Vascular contributions to healthy aging and dementia. Aging Dis 2024 Aug 01;15(4):1432-1437. [doi: 10.14336/AD.2023.1719] [Medline: 39059424]
34. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2224-2260 [FREE Full text] [doi: 10.1016/S0140-6736(12)61766-8] [Medline: 23245609]
35. Appel LJ, Moore TJ, Obarzanek E, Vollmer WM, Svetkey LP, Sacks FM, et al. A clinical trial of the effects of dietary patterns on blood pressure. DASH Collaborative Research Group. N Engl J Med 1997 Apr 17;336(16):1117-1124. [doi: 10.1056/NEJM199704173361601] [Medline: 9099655]
36. Casas R, Sacanella E, Urpí-Sardà M, Corella D, Castañer O, Lamuela-Raventos RM, et al. Long-term immunomodulatory effects of a Mediterranean diet in adults at high risk of cardiovascular disease in the PREvención con DIeta MEDiterránea

XSL•FO
RenderX

(PREDIMED) randomized controlled trial. J Nutr 2016 Sep;146(9):1684-1693 [FREE Full text] [doi: 10.3945/jn.115.229476] [Medline: 27440261]
37. Calder PC, Bosco N, Bourdet-Sicard R, Capuron L, Delzenne N, Doré J, et al. Health relevance of the modification of low grade inflammation in ageing (inflammageing) and the role of nutrition. Ageing Res Rev 2017 Nov;40:95-119 [FREE Full text] [doi: 10.1016/j.arr.2017.09.001] [Medline: 28899766]
38. Costa LM, Mendes MM, Oliveira AC, Magalhães KG, Shivappa N, Hebert JR, et al. Dietary inflammatory index and its relationship with gut microbiota in individuals with intestinal constipation: a cross-sectional study. Eur J Nutr 2022 Feb;61(1):341-355. [doi: 10.1007/s00394-021-02649-2] [Medline: 34351455]
39. Barroso-Batista J, Demengeot J, Gordo I. Adaptive immunity increases the pace and predictability of evolutionary change in commensal gut bacteria. Nat Commun 2015 Nov 30;6:8945 [FREE Full text] [doi: 10.1038/ncomms9945] [Medline: 26615893]
40. Roberts SB, Silver RE, Das SK, Fielding RA, Gilhooly CH, Jacques PF, et al. Healthy aging-nutrition matters: start early and screen often. Adv Nutr 2021 Jul 30;12(4):1438-1448 [FREE Full text] [doi: 10.1093/advances/nmab032] [Medline: 33838032]
41. Jensen MD, Ryan DH, Apovian CM, Ard JD, Comuzzie AG, Donato KA, American College of Cardiology/American Heart Association Task Force on Practice Guidelines, Obesity Society. 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and The Obesity Society. Circulation 2014 Jun 24;129(25 Suppl 2):S102-S138 [FREE Full text] [doi: 10.1161/01.cir.0000437739.71477.ee] [Medline: 24222017]
42. Stanek A, Brożyna-Tkaczyk K, Zolghadri S, Cholewka A, Myśliński W. The role of intermittent energy restriction diet on metabolic profile and weight loss among obese adults. Nutrients 2022 Apr 05;14(7):1509 [FREE Full text] [doi: 10.3390/nu14071509] [Medline: 35406122]
43. Gao Y, Gao K, Shi R, Huang X, Dang P, Liu H, et al. Association between phenotypic age and in-hospital outcomes in patients with acute myocardial infarction: a retrospective observational study. Int J Cardiol Heart Vasc 2025 Mar 29;58:101670 [FREE Full text] [doi: 10.1016/j.ijcha.2025.101670] [Medline: 40235940]
44. Li S, Zhou J, Zhang D, Du Q. Association of dietary inflammatory index and vigorous physical activity on phenotypic age acceleration: a cross-sectional study with machine learning. Front Nutr 2025 Jul 28;12:1602821 [FREE Full text] [doi: 10.3389/fnut.2025.1602821] [Medline: 40791234]
45. Zhu X, Wang X, Tian X, Kong Y. Association between biological aging and cardiovascular health: combined evidence based on cross-sectional and prospective study. Arch Gerontol Geriatr 2025 May;132:105785. [doi: 10.1016/j.archger.2025.105785] [Medline: 39983447]
46. Liu C, Han XC, Hua L, Xin Z. Age-specific relationships between free triiodothyronine and biological aging. Hormones (Athens) (Forthcoming) 2025 May 07. [doi: 10.1007/s42000-025-00670-2] [Medline: 40332749]
47. Chen X, Chen X, Shi F, Yu W, Gao C, Gou S, et al. Associations between anthropometric indices and biological age acceleration in American adults: insights from NHANES 2009-2018 data. Sci Rep 2025 Jul 02;15(1):22691 [FREE Full text] [doi: 10.1038/s41598-025-08559-1] [Medline: 40596327]
48. Xia T, Xing G, Shan G, Xing S. Vitamin E intake mediates the associations of triglyceride-glucose index and its related parameters with phenotypic age acceleration in middle-aged and elderly population. Sci Rep 2025 Jul 05;15(1):24011 [FREE Full text] [doi: 10.1038/s41598-025-10436-w] [Medline: 40617862]
49. NHANES questionnaires, datasets, and related documentation. Center for Disease Control and Prevention (CDC). URL: https://wwwn.cdc.gov/nchs/nhanes/ [accessed 2025-05-29]

Abbreviations

AHEI: Alternative Healthy Eating Index
aMED: Alternative Mediterranean Diet
AUPRC: area under the precision-recall curve
AUROC: area under the receiver operating characteristic curve
DASH: Dietary Approaches to Stop Hypertension
DII: Dietary Inflammatory Index
GBTD: gradient boosting decision tree
HEI-2020: Healthy Eating Index-2020
IHM: Institute of Health and Medicine
KDM: Klemera-Doubal method
MET: metabolic equivalent task
ML: machine learning
NHANES: National Health and Nutrition Examination Survey
PA: physical activity
PIR: poverty income ratio

XSL•FO
RenderX

SHAP: Shapley additive explanations
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by J Job; submitted 01.02.25; peer-reviewed by Y Hu, L Song; comments to author 30.04.25; revised version received 03.10.25; accepted 18.10.25; published 01.12.25.

Please cite as:
Wu Z, Zhang N, Wang H, Yang Y, Shen H, Zhang Q, Cao Y, Du Y, Ji D
Associations Between Dietary Habits and Accelerated Aging and the Establishment of an Accelerated Aging Interpretable Risk Prediction Model via Shapley Additive Explanations: Cross-Sectional Study From Two Representative Populations
JMIR Aging 2025;8:e72020
URL: https://aging.jmir.org/2025/1/e72020
doi:10.2196/72020
PMID:

©Zhengyang Wu, Ning Zhang, Haiwei Wang, Yang Yang, Houhao Shen, Qidi Zhang, Yunxia Cao, Yinan Du, Dongmei Ji. Originally published in JMIR Aging (https://aging.jmir.org), 01.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital Isolation and Depression Risk in Older Adults Using the National Health and Aging Trends Study Database: 8-Year Longitudinal Study

Zhili He[1], PhD; Shijun Yang[2], PhD; Wei Tong[3], PhD; Wei Zhang[1], MD

[1]Department of Neurology, The Central Hospital of Wuhan, Tongji Medical College, Huazhong University of Science and Technology, 26 Shengli Street, Jiang'an District, Wuhan, China

[2]Department of Cardiology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[3]Department of Orthopaedics, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

Corresponding Author:
Wei Zhang, MD
Department of Neurology, The Central Hospital of Wuhan, Tongji Medical College, Huazhong University of Science and Technology, 26 Shengli Street, Jiang'an District, Wuhan, China

Abstract

Background: The rapid advancement of digital technologies has profoundly transformed communication practices. However, this technological revolution has also led to "digital isolation," a form of social disconnection caused by limited or absent engagement with digital communication tools, including smartphones, computers, email, and the internet. This issue is particularly concerning for older adults, as it may increase their likelihood of developing mental health disorders, with depression being a primary concern. Although digital isolation has been studied less frequently than traditional social isolation, it may be a significant contributor to both the initiation and progression of depression in this population.

Objective: This investigation seeks to assess longitudinal relationships between multidimensional digital disengagement (encompassing 4 dimensions: mobile device use, computer interaction, electronic correspondence, and web-based engagement) and incident depression among older adults, using longitudinal data from the nationally representative National Health and Aging Trends Study (NHATS).

Methods: The analysis was conducted based on the NHATS dataset, a nationally representative longitudinal survey using multistage sampling to represent community-dwelling Medicare beneficiaries aged 65 years and older in the United States. We analyzed data from 2011 (Round 1) to 2018 (Round 8), including 8199 participants in the discovery and validation cohorts. Digital isolation was measured using a 4-item index based on self-reported nonuse of mobile phones, computers, email, and the internet. Participants were categorized into high (aggregate score ≥3) or low (aggregate score ≤2) digital isolation groups. Weighted Cox regression models with proportional hazards assumptions were used to quantify longitudinal associations between the digital isolation index (and its individual components) and incident depression, incorporating multivariable adjustment for sociodemographic characteristics (age, sex, and race or ethnicity), socioeconomic indicators (education level, family income, and marital status), and clinical profiles (tobacco use history and multimorbidity burden). Time-to-event analyses were visualized through Kaplan-Meier estimators, complemented by prespecified subgroup analyses evaluating effect modification patterns through interaction term testing.

Results: A high level of digital isolation, as measured by the composite index, was associated with a significantly greater risk of incident depression (fully adjusted model: hazard ratio 1.35, 95% CI 1.18 - 1.55; P<.001). Furthermore, analysis of the individual components showed that nonuse of computers, email, and the internet was each significantly associated with a higher depression risk, whereas mobile phone isolation had a weaker, nonsignificant association.

Conclusions: The study revealed a robust association between increased digital isolation and a higher likelihood of depression in the older population. These results underscore the importance of implementing tailored public health strategies to address digital isolation, especially for older adults. To minimize its detrimental effects on mental health, policymakers should encourage digital literacy programs and strengthen mental health services.

(JMIR Aging 2025;8:e75174) doi:10.2196/75174

KEYWORDS

digital isolation; depression; older adults; longitudinal cohort study; National Health and Aging Trends Study; NHATS; mental health; social isolation; technology use; risk factors; public health interventions

XSL•FO
RenderX

Introduction

In today's digital era, the rapid advancement of information technology has profoundly reshaped human lifestyles. The increasing reliance on digital technologies for communication and social connection has led to the emergence of a phenomenon known as "digital isolation" [1]. It is important to distinguish the concept of "digital isolation" from the more widely recognized "digital divide" [2]. The digital divide typically refers to disparities in access to information and communication technologies and the necessary digital skills across populations [3]. In this study, we use the term "digital isolation" to describe a related but conceptually distinct psychosocial phenomenon: a state of social disconnection resulting from limited or absent engagement with essential digital communication tools, regardless of access [4]. This term refers to social disconnection that occurs when individuals either lack access to or refrain from using modern communication tools, such as smartphones, computers, email, and the internet. Although designed to foster human interaction, these technologies can paradoxically exacerbate loneliness and social isolation when misused or unavailable [4,5]. This paradox is of particular concern due to its potential negative impact on mental health, especially depression.

Depression, a common global mental health disorder, substantially impairs quality of life and social functioning. According to the World Health Organization, depression affects approximately 280 million individuals, or 3.8% of the global population, making it a significant public health concern [6]. Women are disproportionately affected compared to men, and prevalence increases significantly among older adults, where it is a leading cause of disability and suicide [7,8]. Beyond its psychological dimensions, depression is closely linked to physical health conditions such as cardiovascular disease and diabetes, contributing to increased mortality [9,10]. Therefore, identifying risk factors for depression and developing effective prevention and treatment strategies are crucial for public health.

The urgency of this issue is further underscored by the broader public health crisis of social isolation among older adults. A landmark 2020 report by the National Academies of Sciences, Engineering, and Medicine identified social isolation and loneliness as serious yet underrecognized public health risks, noting that approximately one-quarter of community-dwelling Americans aged 65 years and older are socially isolated [11]. The report emphasizes that social isolation poses a major risk for premature mortality, comparable to well-known risk factors such as smoking or obesity [12].

In the era of rapid digitalization, the long-standing problem of social isolation is acquiring a new and critical dimension. As social networks, health care resources, and daily services increasingly move online, limited digital engagement may become a key driver of social exclusion. Consequently, digital isolation—a form of social disconnection arising from nonuse of digital tools—may serve as a crucial pathway through which modern societal structures adversely affect mental health in vulnerable older adults [13,14]. While overuse of digital media poses risks for some, older adults are more often affected by digital underuse or exclusion, which may compound traditional social isolation and increase the risk of depression [15]. Research has begun to explore how digital isolation contributes to depressive symptoms among older adults, but evidence remains limited, particularly regarding the combined effects of multiple forms of digital isolation and their cumulative impact on mental health [16,17].

We conceptualize digital isolation as a latent psychosocial state; to make it observable, we use a 4-domain nonuse composite (mobile phone, computer, email, and internet) as an observable proxy, distinct from the digital divide and digital literacy. Within this framework, this study aims to assess the longitudinal association between multidimensional digital isolation—defined by the use or nonuse of mobile phones, computers, email, and the internet—and the risk of incident depression in a large, nationally representative cohort of older adults. Using data from a large-scale retrospective cohort, we aim to examine both the individual and combined effects of these types of digital disengagement on depression risk. The findings are expected to yield new insights into the mental health consequences of digital disconnection and support the development of targeted public health interventions.

Methods

Study Population

The National Health and Aging Trends Study (NHATS) formed our analytical sample, a longitudinal survey that uses a multistage probability sampling technique to mirror the US population of community-living Medicare beneficiaries aged 65 years and older [18].

We analyzed data collected from NHATS starting in 2011 (Round 1). For the discovery cohort, we used data from Round 1 (2011) to Round 8 (2018), encompassing 7 waves of follow-up. The initial wave enrolled 8245 participants. To ensure analytical precision, we excluded 2230 participants lacking crucial information on key variables such as digital isolation and depressive symptoms, which were central to our research question. Additionally, we excluded 846 participants with a baseline diagnosis of depression to focus on the de novo development of depression. This resulted in a discovery cohort of 5169 participants.

For external validation, we established an independent verification cohort comprising 4182 newly enrolled participants during Wave 5 (2015) of NHATS data collection, ensuring temporal and spatial independence from the discovery sample. For these participants, we used data from Round 5 to Round 12 (2015 - 2022), also spanning 7 waves of follow-up. The validation cohort was formed by excluding 723 individuals who had missing data on crucial variables and an additional 429 participants who were diagnosed with depression at the baseline assessment. Consequently, a total of 3030 participants constituted the final validation cohort.

In the final analysis, we combined the discovery and validation cohorts, yielding a total sample of 8199 participants. Merging these cohorts enhanced the robustness and generalizability of our findings.

XSL•FO
RenderX

Digital Isolation

Digital isolation was defined as limited engagement with modern communication technologies. Based on NHATS data, we assessed 4 domains of digital engagement: mobile phone ownership, computer use, email or text messaging, and internet activity. We consider multidomain nonuse as an observable proxy for the latent state of digital isolation, mirroring the common practice in social isolation research where concrete behaviors are used to represent abstract psychosocial states. This multidomain operationalization aligns with evidence framing older adults' digital engagement as a continuum and supports using composite, multidomain indicators to capture disengagement [19]. Each domain was coded as 1 (nonuse or infrequent use) or 0 (regular use), yielding a cumulative score ranging from 0 to 4. The specific NHATS questionnaire items, variable names, and precise coding logic used to construct each of these 4 components are detailed in Table S4 (Multimedia Appendix 1). Higher scores reflected greater levels of digital isolation. Participants with scores ≥3 were classified as highly digitally isolated, in line with prior NHATS-based gerontechnology research that stratified older adults by multidomain technology nonuse [5]. This threshold approximately corresponded to the top quartile of the score distribution and indicated substantial disengagement from the digital environment.

Depression

The NHATS used the Patient Health Questionnaire-2 (PHQ-2) to evaluate depressive symptoms. This concise screening instrument consists of 2 questions that assess the frequency of such symptoms experienced over the previous month: (1) a lack of interest or enjoyment in activities and (2) persistent feelings of sadness, hopelessness, or depression. Respondents can choose from a scale of "not at all" (assigned 1 point) to "nearly every day" (4 points), resulting in a total PHQ-2 score between 2 and 8. Higher scores on this scale indicate more severe depressive symptoms [20,21]. Notably, NHATS applies a 1-month reference period for the PHQ-2, differing from the standard 2-week timeframe, to align with the reference periods used for other functional measures, such as self-care and mobility assessments [22]. Based on the foundational work of Kroenke et al [20], we used a threshold of 3 points or more to define clinically significant depressive symptoms.

Covariates

Our analytical approach comprehensively adjusted for potential confounders spanning sociodemographic determinants, socioeconomic gradients, behavioral risk profiles, and clinical comorbidities. The analysis focused on several variables known from previous research to be linked with depression and mental health. These variables comprised age, categorized as ≤75 years and >75 years; gender, distinguished as male and female; race or ethnicity, comparing non-Hispanic Whites with other groups; marital status, with comparisons made between married individuals and others; educational attainment, contrasted as less than high school versus high school or above; income level, separated into <50,000; smoking status, classified as nonsmokers and smokers; presence or absence of sleep disorders; dementia status; presence or absence of anxiety; and chronic disease burden, defined as having <3 or ≥3 conditions.

Age, gender, and race or ethnicity may affect individuals' use and adaptation to digital technologies, influencing their social interaction patterns and mental well-being. Marital status, education level, and income level were included as proxies for socioeconomic status, a well-documented determinant of mental health outcomes [23,24]. Additionally, health behaviors and medical history, such as smoking status, sleep disorders, dementia, and anxiety, have shown strong associations with the risk of depression in the existing literature [25,26].

Statistical Analysis

We used survey-weighted analytical methods to address the intricate design of the NHATS dataset. These approaches incorporated adjustments for varying selection probabilities, strategies to minimize nonresponse bias, and targeted oversampling of the specific population segments. Specifically, all analyses incorporated the cross-sectional analytic weights provided in the NHATS public use files to generate nationally representative estimates [18]. Weighted means with SE were used to describe continuous measures, while categorical variables were depicted as adjusted proportions accompanied by 95% CIs. Intergroup comparisons used the chi-square test. Survival analyses used inverse probability-weighted Cox proportional hazards models, with hierarchical adjustment for: (1) sociodemographic factors (age, sex, race or ethnicity, and education), (2) health behaviors (tobacco exposure and sleep dysfunction), and (3) clinical profiles (multimorbidity burden, dementia status, and anxiety comorbidity). To clarify the independent effects of digital isolation components (eg, phone, computer, email, and online isolation) on depression risk, additional analyses estimated hazard ratios (HRs) for each component.

Kaplan-Meier (KM) survival and stratified analyses were conducted after combining the discovery and validation cohorts. KM curves visualized depression incidence differences between low- and high-isolation groups, with log-rank tests evaluating their statistical significance. To examine the appropriateness of combining discovery and validation cohorts, we constructed a dummy variable to indicate cohort membership and tested its interaction with digital isolation in Cox regression models. Nonsignificant interaction terms suggested that the cohort did not modify the effect, thus supporting the pooled analysis. In stratified analyses, the low-isolation group served as the reference, and associations between digital isolation and depression incidence were examined across subgroups based on age (<75 vs ≥75 years), gender, race or ethnicity, smoking status, sleep disorders, marital status, and chronic disease burden (<3 vs ≥3). To examine potential nonlinear associations between digital isolation scores and depression risk, we used restricted cubic splines (RCS) in our analysis. Cox regression models with 3 knots were fitted to assess nonlinear trends in digital isolation scores (0 - 4). Risk curves visualized the progressive impact of higher digital isolation scores on depression risk.

To evaluate the robustness of the observed associations, 2 sets of sensitivity analyses were performed. First, the digital isolation score was recategorized into 3 levels—no isolation (0 points),

XSL•FO
RenderX

moderate isolation (1‐2 points), and high isolation (3‐4 points)—and Cox proportional hazards models were re-estimated in the discovery, validation, and pooled cohorts. Second, to address potential confounding by offline social disengagement, we constructed an In-person Social Isolation Score (range: 0‐4) based on 4 dimensions: absence of a confidant, lack of participation in social or religious activities, and living alone. This variable was included in fully adjusted models to assess the independent effect of digital isolation. All analyses were performed using R software (version 4.4.0; R Core Team), with 2-sided *P* values <.05 considered statistically significant.

Ethical Considerations

This study used deidentified, publicly available data from the NHATS. The original NHATS study protocol was approved by the Johns Hopkins University Institutional Review Board, and all participants provided written informed consent at enrollment. For the present secondary data analysis, the research protocol was additionally reviewed and approved by the Institutional Review Board of The Central Hospital of Wuhan, which is affiliated with Tongji Medical College and Huazhong University of Science and Technology. Because the NHATS public-use dataset contains no identifiable private information, no additional consent was required for this analysis. All procedures adhered to institutional, federal, and journal guidelines for research involving human participants.

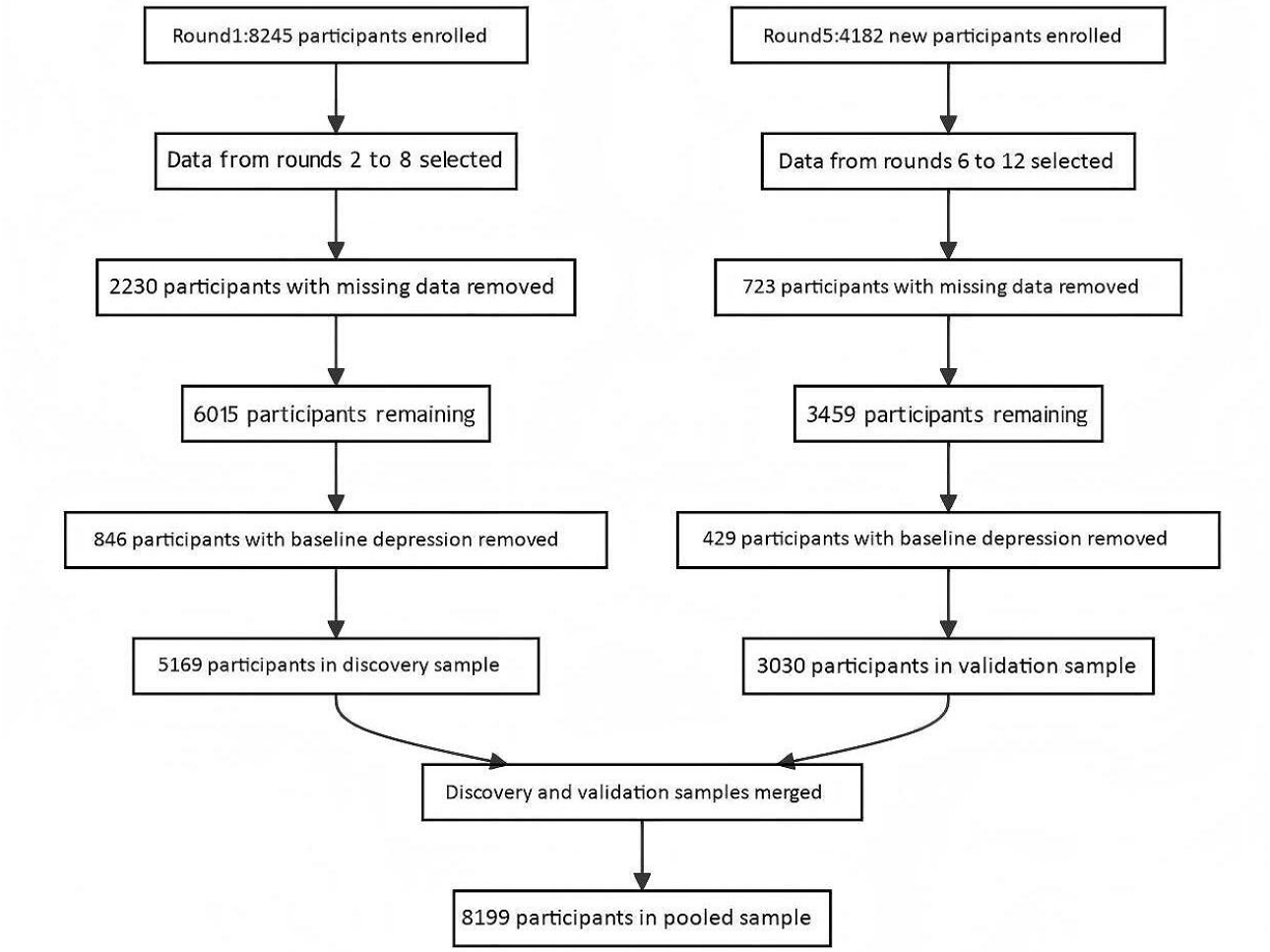
Results

Baseline Characteristics of Participants

As shown in Figure 1, a total of 8199 participants were included after applying the predefined inclusion and exclusion criteria. Table 1 presents the baseline sociodemographic and clinical profiles stratified by analytical cohort (discovery: n=5169, validation: n=3030, pooled: n=8199). The participant pool comprised 44.72% males and 55.28% females. In terms of age, the majority, 68.39%, were 75 years or younger, while the remaining 31.61% were older than 75 years. The majority (70.39%) identified as non-Hispanic White. Educational attainment indicated that 80.12% of participants had completed at least high school, while 17.89% had not. Income data revealed that 35.63% of participants had an annual income below US $50,000, whereas 21.20% reported an income above this threshold. Lifestyle analysis showed that 48.31% were nonsmokers, and 51.59% were smokers. Health assessments found that 80.89% of participants did not report sleep disorders, 97.46% were dementia-free, and 93.19% showed no symptoms of anxiety. Regarding chronic disease burden, 78.86% of participants had fewer than 3 chronic conditions, while 21.14% had 3 or more. For digital isolation, 69.01% of participants were categorized as low isolation and 30.99% as high isolation. Among specific components, 8.59% experienced phone isolation, 30.00% computer isolation, 52.14% email isolation, and 51.09% online isolation. Chi-square tests identified significant group differences in categorical variables, including gender, age group, education, income, smoking status, sleep disorders, dementia, anxiety symptoms, chronic disease burden, and components of digital isolation.

XSL•FO
RenderX

Figure 1. Flowchart of participant selection and data integration process for discovery, validation, and pooled samples.



XSL•FO
RenderX

Table . Baseline characteristics of participants.[a]

Variables[b]	Discovery group (n=5169)	Validation group (n=3030)	Pooled group (n=8199)
Sex, n (%)			
Male	2303 (44.55)	1364 (45.02)	3667 (44.72)
Female	2866 (55.45)	1666 (54.98)	4532 (55.28)
Age group, n (%)[c]			
≤75 years	3440 (66.55)	2167 (71.52)	5607 (68.39)
>75 years	1729 (33.45)	863 (28.48)	2592 (31.61)
Race or ethnicity, n (%)[c]			
White, non-Hispanic	3699 (71.56)	2072 (68.38)	5771 (70.39)
Other[d]	1413 (27.34)	841 (27.76)	2254 (27.49)
Marital status, n (%)[c]			
Married	2836 (54.87)	1609 (53.1)	4445 (54.21)
Other	2327 (45.02)	1417 (46.77)	3744 (45.66)
Education, n (%)[c]			
Low (<high school)	997 (19.29)	470 (15.51)	1467 (17.89)
High (≥high school)	4118 (79.67)	2451 (80.89)	6569 (80.12)
Income, n (%)[c]			
Low (<50,000)	1919 (37.13)	1002 (33.07)	2921 (35.63)
High (≥50,000)	1019 (19.71)	719 (23.73)	1738 (21.2)
Smoking status, n (%)[c]			
Nonsmoker	2453 (47.46)	1508 (49.77)	3961 (48.31)
Smoker	2713 (52.49)	1517 (50.07)	4230 (51.59)
Sleep disorder, n (%)[c]			
No	4184 (80.94)	2448 (80.79)	6632 (80.89)
Yes	975 (18.86)	579 (19.11)	1554 (18.95)
Dementia, n (%)[e]			
No	5033 (97.37)	2958 (97.62)	7991 (97.46)
Yes	134 (2.59)	70 (2.31)	204 (2.49)
Anxiety, n (%)[c]			
No	4808 (93.02)	2833 (93.5)	7641 (93.19)
Yes	361 (6.98)	197 (6.5)	558 (6.81)
Chronic disease, n (%)[c,f]			
Less than 3 diseases	4060 (78.55)	2406 (79.41)	6466 (78.86)
3 or more diseases	1109 (21.45)	624 (20.59)	1733 (21.14)
Digital isolation group, n (%)[c]			
Low isolation	3408 (65.93)	2250 (74.26)	5658 (69.01)
High isolation	1761 (34.07)	780 (25.74)	2541 (30.99)
Items of digital isolation, n (%)			
Phone isolation[e]	533 (10.31)	171 (5.64)	704 (8.59)
Computer isolation[c]	1579 (30.55)	881 (29.08)	2460 (30)

XSL•FO
RenderX

Variables[b]	Discovery group (n=5169)	Validation group (n=3030)	Pooled group (n=8199)
Email isolation[c]	2973 (57.52)	1302 (42.97)	4275 (52.14)
Online isolation[c]	2856 (55.25)	1333 (43.99)	4189 (51.09)

[a]Percentages for some variables, such as Income and Marital Status, may not sum to 100% due to missing data. Participants with missing data on any key variable were excluded from the final analysis as described in the Methods.

[b]We conducted chi-square tests for all categorical variables to assess the association between each variable and the occurrence of depression.

[c]Variables with *P* values <.05 in chi-square tests are marked with superscript "c."

[d]The "Other" category in the race or ethnicity classification includes Black or African American, American Indian, Alaska Native, Asian, Native Hawaiian, Pacific Islander, and other unspecified races or ethnicities. The "Other" category for Marital Status includes widowed, divorced, separated, and never married.

[e]Variables with *P* values <.001 in chi-square tests are marked with superscript "e."

[f]The assessment of chronic disease burden is based on the cumulative number of the following conditions: heart disease, hypertension, arthritis, osteoporosis, diabetes, lung disease, stroke, dementia or Alzheimer disease, cancer, and others.

Association Between Digital Isolation and Depression

Table 2 demonstrates a significant correlation between heightened digital isolation and a greater likelihood of depression. The Cox proportional hazards analysis indicated that participants in the high-isolation category had a notably higher risk of depression compared to those in the low-isolation category. In the discovery cohort, the unadjusted HR for the high-isolation group was 1.73 (95% CI 1.54 - 1.94; *P* <.001), which reduced to an adjusted HR of 1.33 (95% CI 1.12 - 1.57; *P*=.001) upon controlling for potential confounders. The validation cohort displayed comparable results, with an unadjusted HR of 1.81 (95% CI 1.55 - 2.12; *P*<.001) and an adjusted HR of 1.37 (95% CI 1.07 - 1.75; *P*=.012). When analyzing the combined sample, the unadjusted HR was 1.75 (95% CI 1.60 - 1.92; *P*<.001), and the adjusted HR was 1.35 (95% CI 1.18 - 1.55; *P*<.001). Figure 2 visually supports these findings, depicting a significantly higher cumulative incidence of depression in the high-isolation group throughout the follow-up period. Figure 3 presents the stratified analyses examining the association between digital isolation and depression across various subgroups. Formal tests for interaction revealed that the effect of digital isolation on depression risk was significantly modified by age, education level, and chronic disease burden. Specifically, the association was stronger in individuals aged 75 years or younger compared to those older than 75 years (*P* for interaction <.001), in those with higher education (*P* for interaction=.003), and in those with fewer than 3 chronic diseases (*P* for interaction=.04). No significant interaction effects were observed for other subgroups, including gender, race or ethnicity, and marital status, suggesting the association was broadly consistent across these groups. Collectively, the data from Table 2 and Figures 2 and 3 strongly suggest that increased digital isolation is significantly linked to a higher risk of depression, a relationship that persists even after accounting for multiple potential confounding factors. This underscores the necessity of incorporating digital isolation reduction strategies into public health initiatives to mitigate the burden of depression.

Table . The observed link between digital isolation and the likelihood of depression.

Samples	Event and total number (%)	Model 1 HR (95% CI)[a,b]	*P* value	Model 2 HR (95% CI)[b,c]	*P* value
Discovery					
Low isolation (reference)	3408/5169 (65.93)	1 (reference)		1 (reference)	
High isolation	1761/5169 (34.07)	1.73 (1.54-1.94)	<.001	1.33 (1.12-1.57)	.001
Validation					
Low isolation (reference)	2250/3030 (74.26)	1 (reference)		1 (reference)	
High isolation	780/3030 (25.74)	1.81 (1.55-2.12)	<.001	1.37 (1.07-1.75)	.01
Pooled					
Low isolation (reference)	5658/8199 (69.01)	1 (reference)		1 (reference)	
High isolation	2541/8199 (30.99)	1.75 (1.60-1.92)	<.001	1.35 (1.18-1.55)	<.001

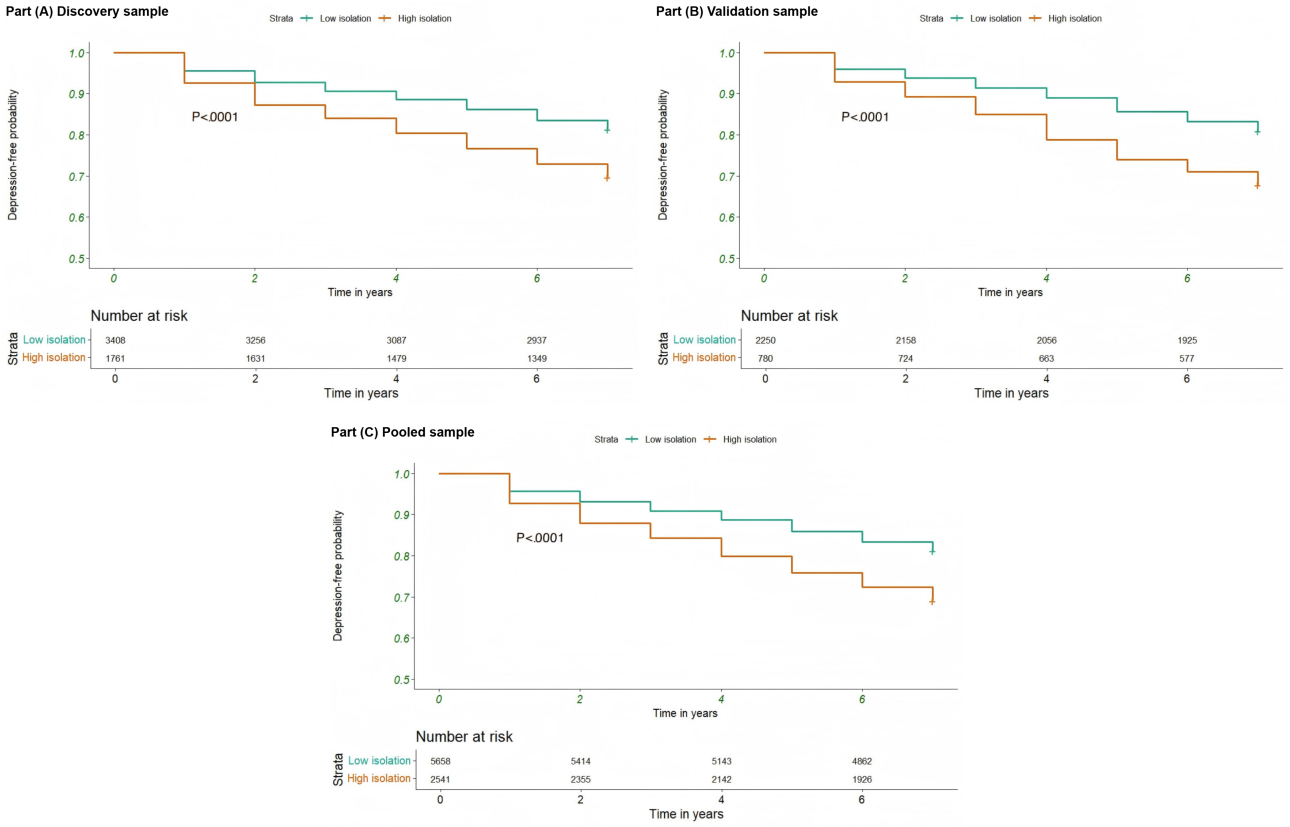
[a]Model 1: unadjusted Cox proportional hazards model.

[b]HR: hazard ratio.

[c]Model 2: Cox proportional hazards model adjusted for age, gender, race, education, income, marital status, sleep disorder, smoking status, dementia, anxiety, and chronic disease.

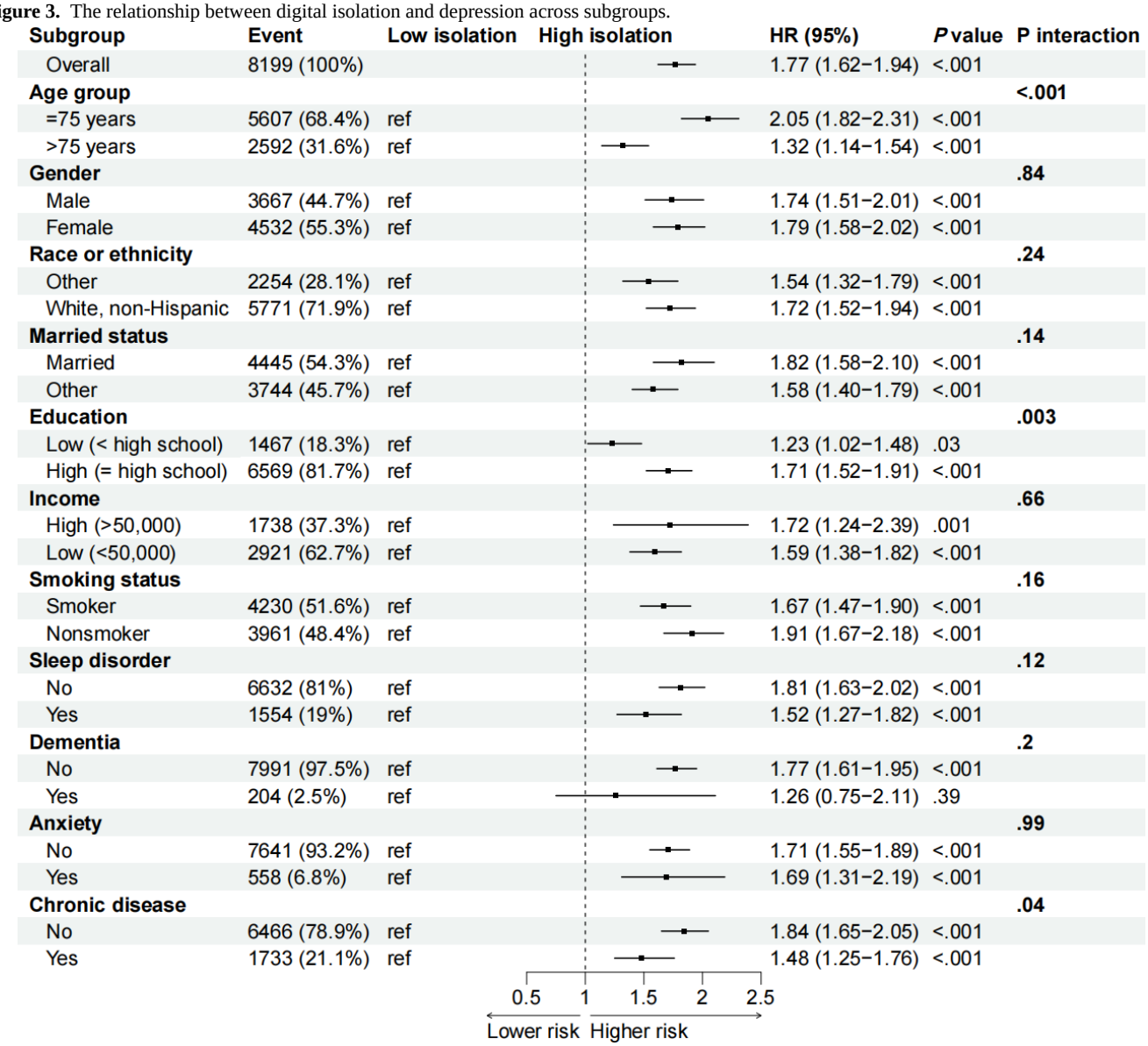
XSL•FO
RenderX

Figure 2. Kaplan-Meier curves depicting the identified association between digital isolation and the risk of depression.



XSL•FO
RenderX

Figure 3. The relationship between digital isolation and depression across subgroups.



Association Between Different Items of Digital Isolation and Depression

An overview of the correlation between the various components of digital isolation and the likelihood of depression is provided in Table 3. Adjusted Cox models revealed significant associations for computer isolation, email isolation, and online isolation, but not for phone isolation. In the pooled sample, adjusted HRs were 1.23 (95% CI 1.08 - 1.41; *P*=.003) for computer isolation, 1.52 (95% CI 1.32 - 1.75; *P*<.001) for email isolation, and 1.32 (95% CI 1.15 - 1.52; *P*<.001) for online isolation. In contrast, the adjusted HR for phone isolation was 1.19 (95% CI 0.98 - 1.44; *P*=.08), which was not statistically significant. These findings suggest that isolation related to more complex digital activities, such as computer, email, and internet use, has a stronger and more significant association with depression risk in comparison to the weaker and nonsignificant association observed for mobile phone isolation.

XSL•FO
RenderX

Table . Association between different items of digital isolation and depression.

Sample or variables	Model 1 HR[a] (95% CI)	*P* value	Model 2 HR (95% CI)	*P* value
Discovery sample				
Phone isolation	1.29 (1.09-1.54)	.004	1.18 (0.94-1.47)	.15
Computer isolation	1.55 (1.38-1.75)	<.001	1.19 (1.01-1.40)	.04
Email isolation	1.81 (1.60-2.05)	<.001	1.48 (1.24-1.77)	<.001
Online isolation	1.76 (1.56-1.99)	<.001	1.35 (1.13-1.61)	.001
Validation sample				
Phone isolation	1.34 (1.00-1.79)	.05	1.24 (0.85-1.83)	.27
Computer isolation	1.74 (1.49-2.02)	<.001	1.28 (1.02-1.61)	.03
Email isolation	1.99 (1.71-2.31)	<.001	1.59 (1.25-2.02)	<.001
Online isolation	1.86 (1.60-2.17)	<.001	1.26 (0.99-1.60)	.06
Pooled sample				
Phone isolation	1.30 (1.12-1.51)	<.001	1.19 (0.98-1.44)	.08
Computer isolation	1.62 (1.47-1.78)	<.001	1.23 (1.08-1.41)	.003
Email isolation	1.86 (1.69-2.04)	<.001	1.52 (1.32-1.75)	<.001
Online isolation	1.79 (1.63-1.97)	<.001	1.32 (1.15-1.52)	<.001

[a]HR: hazard ratio.

Nonlinear Relationship Between Digital Isolation and Depression Risk

Multimedia Appendix 2 presents results from RCS analyses examining nonlinear relationships between digital isolation scores (0‐4) and depression risk. Across all cohorts—discovery, validation, and pooled—a consistent nonlinear trend was observed, with depression risk increasing progressively with higher digital isolation scores. This trend reinforces the role of digital isolation as a significant contributor to depression risk.

Sensitivity Analysis

Sensitivity analyses confirmed the robustness of the main findings. Using a multilevel categorization of digital isolation (no isolation, moderate isolation, and high isolation; Table S1 in Multimedia Appendices 1 and 3), the results were consistent across discovery, validation, and pooled samples. To address potential concerns regarding the combination of 2 cohorts, interaction analyses between digital isolation and cohort membership were conducted. As shown in Table S2 (Multimedia Appendix 1), no significant interaction was found in either unadjusted or adjusted models (all $P>.05$), confirming the consistency of digital isolation effects across both cohorts. The adjusted HRs in the combined analysis revealed a significant association between isolation and adverse outcomes. Specifically, moderate isolation was associated with an HR of 1.63 (95% CI 1.44‐1.84; $P<.001$), whereas high isolation carried an even higher HR of 2.23 (95% CI 1.99‐2.51; $P<.001$), both in comparison to the absence of isolation. KM survival curves further supported these findings, showing significantly higher depression incidence over time in the high isolation group. The sensitivity analyses emphasized the strong and consistent link between digital isolation and depression risk, thereby underscoring the reliability of the findings across different analytical methods and stratifications. After further adjustment for in-person social isolation, the association between digital isolation and depression remained robust across the discovery, validation, and pooled cohorts (Table S3 in Multimedia Appendix 1). Although the HRs were slightly attenuated in Model 3, the associations remained statistically significant, supporting the independent contribution of digital isolation to depression risk.

Discussion

Principal Results

The study demonstrates a strong association between digital isolation and the increased likelihood of developing depression. Using Cox proportional hazards models, researchers observed that individuals in the high-isolation category exhibited a significantly elevated risk of depression compared to those in the low-isolation group, specifically in terms of computer isolation, email isolation, and online isolation. Additionally, the analysis using KM survival curves confirmed these findings, revealing a notably higher cumulative incidence of depression over time within the high-isolation group. These outcomes emphasize digital isolation as a crucial risk factor for depression.

This study's results support existing research that demonstrates a connection between digital isolation and negative mental health consequences. Notably, prior investigations have consistently found that social isolation and restricted access to digital resources are strongly linked to increased rates of depression [14,27,28]. Our study builds on these findings, highlighting the pronounced effects of computer, email, and online isolation. Notably, in contrast to some earlier research, we observed that the impact of mobile phone isolation was

XSL•FO
RenderX

relatively limited. This may be due to the high prevalence of mobile phone adoption among older adults. As shown in Table 1, only 8.6% of the pooled sample experienced phone isolation, compared to 30.0% for computer isolation and more than 50% for email and online isolation. Because the vast majority of older adults in this cohort use mobile phones, the small group that does not may be too heterogeneous to show a consistent effect. The relationship between digital isolation and depression may operate through several psychosocial mechanisms. First, digital isolation can foster social isolation by reducing interpersonal interactions and social support, thereby increasing vulnerability to depression. Second, restricted access to online information may limit awareness and use of mental health resources, exacerbating depressive symptoms. In particular, computer and online isolation may more severely restrict access to information and social networks, potentially explaining their stronger associations with depression observed in this study. Notably, our subgroup analyses revealed a more complex relationship than anticipated. The detrimental effect of digital isolation was paradoxically stronger in individuals who were younger (≤75 years), had higher education, and were healthier (fewer than 3 chronic diseases). These unexpected findings challenge a simplistic view of vulnerability, suggesting that the psychological reliance on digital connectivity—and the subsequent impact of being disconnected—may differ substantially across demographic groups. The mechanisms driving these specific interactions represent a critical area for future investigation.

Strengths

One of the key strengths of our research lies in the adoption of a longitudinal cohort design. This approach permitted us to monitor participants' progress over an extended period, thereby enhancing our ability to establish a clearer causal link between digital isolation and depression. Additionally, we used a large sample size and performed dual analysis using discovery and validation cohorts to ensure the robustness and reproducibility of our results. This design enhances the external validity of the study, augmenting the generalizability of the findings. We also adjusted for a variety of potential confounding factors, such as demographic characteristics, health status, and lifestyle factors, further enhancing the accuracy of the analysis. Further analysis revealed a nonlinear association between digital isolation scores and depression risk. The RCS analysis (Multimedia Appendix 2) showed a nonlinear upward trend in depression risk as digital isolation scores increased, underscoring the complexity of how digital isolation affects mental health. These results support the conclusion that higher digital isolation levels substantially elevate depression risk. Sensitivity analyses (Table S1 in Multimedia Appendices 1 and 3) confirmed the robustness of these findings, demonstrating consistent results across different grouping strategies. These additional analyses further strengthen the link between digital isolation and depression risk, thereby enhancing the credibility and reliability of the study's conclusions.

Limitations

Despite its strengths, including a large, longitudinal, and nationally representative sample, this study has several limitations. First, our measures have inherent constraints. The digital isolation index, while based on prior research, has not been formally validated, and its binary (use vs. nonuse) operationalization does not capture the nuances of engagement frequency or quality. Furthermore, both digital use and depressive symptoms were self-reported, which may be subject to recall bias. Besides, the use of a 1-month recall period for the PHQ-2, a deviation from the standard 2-week timeframe for which the instrument was validated, may affect the comparability of our findings and could potentially influence symptom reporting. Second, challenges to causal inference remain. Despite adjusting for in-person social isolation, the potential for unmeasured confounding from factors such as mobility limitations, cognitive decline, or health care accessibility persists. Additionally, while our longitudinal design mitigates the risk, we cannot fully rule out reverse causality, where subclinical depressive symptoms might precede digital withdrawal. Future research using advanced methods such as cross-lagged panel models is needed to further disentangle this relationship. Finally, the generalizability of our findings may be limited. The results are based on community-dwelling older adults in the United States and may not apply to institutionalized individuals, younger populations, or those in different cultural and technological contexts. This nonuse-based proxy primarily captures the quantity and frequency of participation, but it does not directly assess the quality of interactions or subjective experiences of digital loneliness. Additionally, some instances of nonuse may be voluntary.

Future Directions and Public Health Relevance

Our findings have direct public health implications, indicating that promoting digital literacy among older adults is essential for safeguarding their mental health through targeted policies and interventions. To strengthen this observational evidence, future studies should prioritize establishing causality through experimental designs, such as randomized controlled trials of digital literacy interventions. Methodologically, advanced techniques such as cross-lagged panel models are needed to further address the possibility of reverse causality. Finally, the research scope should be expanded to explore the effects of digital isolation on additional outcomes, including anxiety and cognitive decline, as well as to examine the complex role of emerging digital platforms.

Conclusions

The study reveals a strong connection between digital isolation and the likelihood of depression, with specific emphasis on computer isolation, email detachment, and internet seclusion. Though mobile phone isolation seems to have a lesser effect, the overall outcomes emphasize the vital need to combat digital isolation in order to foster improved mental health. These findings hold substantial relevance for the development of public health policies and interventions, particularly targeting older individuals and other vulnerable groups, to alleviate digital isolation and its associated risks.

XSL•FO
RenderX

Acknowledgments

We extend our heartfelt gratitude to the participants and staff of the National Health and Aging Trends Study (NHATS) for their indispensable role in data collection and management. The NHATS team is commendable for preserving this vital resource, which is accessible to the public via their website. We acknowledge their efforts in making these data available for research purposes.

We acknowledge the DeepSeek R1 large language model for its assistance in language translation and manuscript polishing. The authors carefully reviewed and edited the content as necessary and take full responsibility for the final publication.

Funding

This study was supported by the Wuhan Science and Technology Bureau Knowledge Innovation Special Dawn Plan Fund (grant number: 2022020801020541). The funding agency provided financial support for data analysis and manuscript preparation but had no role in the study design, data collection, or the decision to submit the manuscript for publication.

Authors' Contributions

Conceptualization: ZH, SY, WZ
Methodology: ZH, SY
Formal analysis: ZH, SY
Investigation: ZH
Data curation: ZH, SY
Writing – original draft: ZH
Writing – review & editing: SY, WT, WZ
Supervision: WZ
Project administration: WZ

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional analyses, including sensitivity analyses, interaction effects, item-level associations of digital isolation, and detailed construction of the digital isolation index used in the study.
[DOCX File, 20 KB - aging_v8i1e75174_app1.docx]

Multimedia Appendix 2
Nonlinear association between digital isolation score and depression risk in discovery, validation, and pooled samples (restricted cubic splines analysis).
[DOCX File, 142 KB - aging_v8i1e75174_app2.docx]

Multimedia Appendix 3
Kaplan-Meier survival curves showing depression risk by digital isolation levels (sensitivity analysis).
[DOCX File, 171 KB - aging_v8i1e75174_app3.docx]

References

1. Digital isolation: the paradox of disconnecting in a connected world. Self Drive Psychology. 2025. URL: https://selfdrivepsychology.com/digital-isolation [accessed 2025-07-31]
2. Li H, Glecia A. Impact of social isolation and digital divide on mental health and wellbeing in patients with mental health disorders during COVID-19: a multiple case study. Issues Ment Health Nurs 2023 Apr;44(4):313-320. [doi: 10.1080/01612840.2023.2189957] [Medline: 36989507]
3. The Digital Divide: what it is, and what's being done to close it. Investopedia. 2025. URL: https://www.investopedia.com/the-digital-divide-5116352 [accessed 2025-07-31]
4. Fleischmann K. The Online Pandemic in Design Courses: Design Higher Education in Digital Isolation: Proud Pen; 2020. [doi: 10.51432/978-1-8381524-0-6_1]
5. Deng C, Shen N, Li G, Zhang K, Yang S. Digital isolation and dementia risk in older adults: longitudinal cohort study. J Med Internet Res 2025 Feb 19;27:e65379. [doi: 10.2196/65379] [Medline: 39969956]
6. Depressive disorder (depression). World Health Organization. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/depression [accessed 2024-08-20]

XSL•FO
RenderX

7. Zenebe Y, Akele B, W/Selassie M, Necho M. Prevalence and determinants of depression among old age: a systematic review and meta-analysis. Ann Gen Psychiatry 2021 Dec 18;20(1):55. [doi: 10.1186/s12991-021-00375-x] [Medline: 34922595]
8. Sadek J, Diaz-Piedra B, Saleh L, MacDonald L. A narrative review: suicide and suicidal behaviour in older adults. Front Psychiatry 2024;15:1395462. [doi: 10.3389/fpsyt.2024.1395462] [Medline: 38800059]
9. Rajan S, McKee M, Rangarajan S, et al. Association of symptoms of depression with cardiovascular disease and mortality in low-, middle-, and high-income countries. JAMA Psychiatry 2020 Oct 1;77(10):1052-1063. [doi: 10.1001/jamapsychiatry.2020.1351] [Medline: 32520341]
10. Inoue K, Beekley J, Goto A, Jeon CY, Ritz BR. Depression and cardiovascular disease events among patients with type 2 diabetes: a systematic review and meta-analysis with bias analysis. J Diabetes Complications 2020 Dec;34(12):107710. [doi: 10.1016/j.jdiacomp.2020.107710] [Medline: 32921574]
11. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division; Board on Behavioral, Cognitive, and Sensory Sciences, Division of Behavioral and Social Sciences and Education. Social Isolation and Loneliness in Older Adults: Opportunities for the Health Care System: National Academies Press (US); 2020. [doi: 10.17226/25663] [Medline: 32510896]
12. Donovan NJ, Blazer D. Social isolation and loneliness in older adults: review and commentary of a National Academies report. Am J Geriatr Psychiatry 2020 Dec;28(12):1233-1244. [doi: 10.1016/j.jagp.2020.08.005] [Medline: 32919873]
13. Wang J, Lu X, Ngai SBC, et al. Digital exclusion and depressive symptoms among older people: findings from five aging cohort studies across 24 countries. Health Data Sci 2025;5:0218. [doi: 10.34133/hds.0218] [Medline: 39790099]
14. Yan R, Liu X, Xue R, et al. Association between internet exclusion and depressive symptoms among older adults: panel data analysis of five longitudinal cohort studies. EClinicalMedicine 2024 Sep;75:102767. [doi: 10.1016/j.eclinm.2024.102767] [Medline: 39188711]
15. Li Y, Bai X, Chen H. Social isolation, cognitive function, and depression among Chinese older adults: examining internet use as a predictor and a moderator. Front Public Health 2022;10(809713):809713. [doi: 10.3389/fpubh.2022.809713] [Medline: 35359786]
16. Welch V, Ghogomu ET, Barbeau VI, et al. Digital interventions to reduce social isolation and loneliness in older adults: an evidence and gap map. Campbell Syst Rev 2023 Dec;19(4):e1369. [doi: 10.1002/cl2.1369] [Medline: 38024780]
17. Saha I, Sundström C, Kandasamy A, et al. Digital interventions for common mental health problems among older adults in low- and middle-income countries: a scoping review. BMJ Glob Health 2025 Jun 24;10(6):e017836. [doi: 10.1136/bmjgh-2024-017836] [Medline: 40555507]
18. Freedman VA, Kasper JD. Cohort profile: the National Health and Aging Trends Study (NHATS). Int J Epidemiol 2019 Aug 1;48(4):1044-1045g. [doi: 10.1093/ije/dyz109] [Medline: 31237935]
19. Kebede AS, Ozolins LL, Holst H, Galvin K. Digital engagement of older adults: scoping review. J Med Internet Res 2022 Dec 7;24(12):e40192. [doi: 10.2196/40192] [Medline: 36477006]
20. Kroenke K, Spitzer RL, Williams JBW. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care 2003 Nov;41(11):1284-1292. [doi: 10.1097/01.MLR.0000093487.78664.3C] [Medline: 14583691]
21. Li C, Friedman B, Conwell Y, Fiscella K. Validity of the Patient Health Questionnaire 2 (PHQ-2) in identifying major depression in older people. J Am Geriatr Soc 2007 Apr;55(4):596-602. [doi: 10.1111/j.1532-5415.2007.01103.x] [Medline: 17397440]
22. Freedman VA, Schrack JA, Skehan ME. National Health and Aging Trends Study User Guide: Rounds 1-14 Beta Release. 2025 Sep URL: https://www.nhats.org/sites/default/files/2025-09/NHATSR14BetaUserGuide_09242025.pdf [accessed 2025-11-29]
23. Ghorbani S, Ghavidel F, Abdollahi S, et al. Socioeconomic inequality in mental health disorders: a cross-sectional study from the Tehran University of Medical Sciences employees' cohort study. Sci Rep 2025 May 22;15(1):17796. [doi: 10.1038/s41598-025-02192-8] [Medline: 40404803]
24. Kirkbride JB, Anglin DM, Colman I, et al. The social determinants of mental health and disorder: evidence, prevention and recommendations. World Psychiatry 2024 Feb;23(1):58-90. [doi: 10.1002/wps.21160] [Medline: 38214615]
25. Wallensten J, Ljunggren G, Nager A, et al. Stress, depression, and risk of dementia - a cohort study in the total population between 18 and 65 years old in Region Stockholm. Alzheimers Res Ther 2023 Oct 2;15(1):161. [doi: 10.1186/s13195-023-01308-4] [Medline: 37779209]
26. Hahad O, Beutel M, Gilan DA, et al. The association of smoking and smoking cessation with prevalent and incident symptoms of depression, anxiety, and sleep disturbance in the general population. J Affect Disord 2022 Sep 15;313:100-109. [doi: 10.1016/j.jad.2022.06.083] [Medline: 35777492]
27. Zhao X, Han X, Lou VWQ, Zhang Z. Influence of new internet usage on depressive symptoms among older adults: does the effect vary in people with different economic status? J Gerontol Soc Work 2025 Jan;68(1):37-52. [doi: 10.1080/01634372.2024.2428254] [Medline: 39560404]
28. Guo L, Li Y, Cheng K, Zhao Y, Yin W, Liu Y. Impact of internet usage on depression among older adults: comprehensive study. J Med Internet Res 2025 Jan 31;27:e65399. [doi: 10.2196/65399] [Medline: 39890098]

XSL•FO
RenderX

Abbreviations

HR: hazard ratio
KM: Kaplan-Meier
NHATS: National Health and Aging Trends Study
PHQ-2: Patient Health Questionnaire-2
RCS: restricted cubic splines

Edited by R Yang; submitted 29.03.25; peer-reviewed by C Garcia-Pena, K Zhang, Y Matsuoka; revised version received 06.09.25; accepted 29.09.25; published 12.12.25.

Please cite as:
He Z, Yang S, Tong W, Zhang W
Digital Isolation and Depression Risk in Older Adults Using the National Health and Aging Trends Study Database: 8-Year Longitudinal Study
JMIR Aging 2025;8:e75174
URL: https://aging.jmir.org/2025/1/e75174
doi:10.2196/75174

© Zhili He, Shijun Yang, Wei Tong, Wei Zhang. Originally published in JMIR Aging (https://aging.jmir.org), 12.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Technology Activities and Cognitive Trajectories Among Community-Dwelling Older Adults: National Health and Aging Trends Study

Erh-Chi Hsu[1*], RN, MPH; Erin M Spaulding[1,2,3*], RN, PhD; Eric Jutkowitz[4*], PhD

[1]School of Nursing, Johns Hopkins University, 525 N Wolfe St, Baltimore, MD, United States

[2]Digital Health Innovation Laboratory, Ciccarone Center for the Prevention of Cardiovascular Disease, Division of Cardiology, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, United States

[3]Welch Center for Prevention, Epidemiology, and Clinical Research, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

[4]School of Public Health, University of Minnesota, Minneapolis, MN, United States

[*]all authors contributed equally

Corresponding Author:
Erh-Chi Hsu, RN, MPH
School of Nursing, Johns Hopkins University, 525 N Wolfe St, Baltimore, MD, United States

Abstract

Background: While the positive effects of digital technology on cognitive function are established, the specific impacts of different types of technology activities on distinct cognitive domains remain underexplored.

Objective: This study aimed to examine the associations between transitions into and out of various technology activities and trajectories of cognitive domains among community-dwelling older adults without dementia.

Method: Data were drawn from 5566 community-dwelling older adults without dementia who participated in the National Health and Aging Trends Study from 2015 to 2022. Technology activities assessed included online shopping, banking, medication refills, social media use, and checking health conditions online. The cognitive domains measured were episodic memory, executive function, and orientation. Asymmetric effects models were used to analyze the associations between technology activity transitions and cognitive outcomes, adjusting for demographic, socioeconomic, and health-related covariates. Lagged models were applied for sensitivity analysis.

Results: In the asymmetric effects models, the onset of online shopping (β=.046, *P*=.02), medication refills (β=.073, *P*<.001), and social media use (β=.065, *P*=.01) was associated with improved episodic memory. The cessation of online shopping was associated with faster episodic memory decline (β=−.023, *P*=.047). In contrast, the cessation of online banking (β=−.078, *P*=.01) and social media use (β=−.066, *P*=.003) was associated with decreased episodic memory. The initiation of instrumental, social, and health-related technology activities was associated with slower cognitive decline in orientation. The lagged models further emphasized the effects of stopping online banking and starting online medication refills in relation to episodic memory, as well as the positive associations between online shopping and social media use and orientation. All significant effects were of small magnitude.

Conclusions: Combining findings from the main and sensitivity analyses, results suggest that interventions designed to support episodic memory in older adults should emphasize promoting the use of online medication refill services and sustaining engagement with online banking, particularly among those who have already established these habits. To support orientation, strategies should focus on facilitating adoption of online shopping and social media use, helping older adults become comfortable navigating these platforms. Future trials are needed to assess the clinical relevance of targeted interventions for specific cognitive domains, to promote the initiation and maintenance of digital activities to help mitigate domain-specific cognitive decline in aging populations.

(JMIR Aging 2025;8:e77227) doi:10.2196/77227

KEYWORDS

aging; cognitive function; technology use; digital activities; NHATS; National Health and Aging Trends Study

Introduction

As technology use becomes increasingly prevalent among older adults, its potential impact on cognitive health warrants attention. Between 2018 and 2021, smartphone ownership among individuals aged 65 years and older in the United States increased from 46% to 61%, outpacing growth in younger age groups [1]. Similarly, tablet ownership increased from 6% to 44%, and social media use surged from 16% to 45% over the

XSL•FO
RenderX

past decade [1]. This rise in digital engagement holds the potential to influence various aspects of older adults' lives, including their cognitive function. Research indicates that digital technology use is associated with higher social well-being and lower depressive symptoms in older adults [2,3]. In addition, technology can aid in managing instrumental activities of daily living (IADLs), which are crucial for maintaining independence (eg, grocery shopping online) [4]. While cognitive health is crucial for preserving autonomy and quality of life in aging populations, and technology appears to support these goals, the specific relationship between technology use and cognitive function remains understudied, particularly among older adults without dementia.

Cognitive health, encompassing domains such as episodic memory, executive function, and orientation, is a crucial predictor of health outcomes, quality of life, and cognitive trajectories in older adults [5-10]. Declines in cognitive function are associated with diminished quality of life [11] and greater financial strain on individuals and their families [12]. In addition, cognitive health is closely linked to the ability to perform daily activities, including those involving digital technologies. From a neurobiological standpoint, cognitive functions are supported by intricate brain networks. The hippocampus is essential for episodic memory, allowing flexible retrieval of information and aiding social interactions by integrating memory with context [13,14]. The prefrontal cortex manages goal-directed actions through top-down control, coordinating activity across different brain regions, which underlies executive function [15]. Moreover, spatial and temporal orientation involve the temporoparietal junction, a region crucial for navigating through time and space [16].

Many digital tasks, such as managing medications online, navigating shopping sites, or using social media, require a combination of these cognitive skills. Performing these activities involves remembering steps, following instructions, organizing information, engaging with digital interfaces, and maintaining awareness of dates and times. Behaviorally, cognitive decline is strongly associated with a reduced ability to perform basic and IADLs [17]. Recent research using network analysis has demonstrated strong associations between nondigital IADLs (eg, banking and shopping) and specific cognitive domains such as memory and orientation [18]. These findings underscore the importance of studying digital equivalents of such tasks.

Most studies have focused on general internet use and global cognitive functioning, without examining the nuanced impact of transitions into and out of specific technology activities on different cognitive domains. For example, a recent study using the National Health and Aging Trends Study (NHATS) indicated that cessation of internet use predicted declines in episodic memory, executive function, and orientation [19]. However, that study treated internet use as a single broad category, without differentiating between specific online activities. Similarly, while Li et al [18] found connections between instrumental activities and specific cognitive domains, it remains unclear whether these relationships translate to their digital counterparts. Other research using the Health and Retirement Study has examined different types of internet use—such as social, informational, and instrumental purposes—and found associations with better global cognitive scores [20]. Together, these studies highlight a key gap: although digital engagement appears to influence cognitive health, it is not yet clear which types of technology activities are related to specific cognitive domains. Addressing this gap is crucial for informing resource allocation and intervention design, ensuring that limited resources are used effectively and that interventions target the cognitive domains most likely to benefit from technology use. This study aimed to address the following 2 research questions:

1. What are the associations between technology activities (ie, online shopping, banking, medication refills, social media use, and checking health information) and cognitive domains (ie, episodic memory, executive function, and orientation) among community-dwelling older adults without dementia?
2. Does the start and stop of using specific technology activities moderate the rate of cognitive decline over time in community-dwelling older adults without dementia?

Methods

Data Source and Study Population

This project involved a secondary quantitative data analysis using data from the NHATS collected between 2015 and 2022 (Rounds 5-12). NHATS, supported through a cooperative agreement with the National Institute on Aging (U01AG032947), is a nationally representative longitudinal study of Medicare beneficiaries aged 65 years and older [21]. Annual interviews have been conducted since 2011 (Round 1), with sample replenishment in 2011, 2015, and 2022 (Rounds 1, 5, and 12 [21]). A detailed cohort profile is available in Freedman and Kasper [21], published by the NHATS study team.

The sample consisted of community-dwelling individuals aged 65 years or older without a diagnosis of dementia or mild cognitive impairment at baseline [5]. The original analytic sample included 5596 community-dwelling individuals aged 65 years and older without dementia. We identified 30 out of 5596 participants who were classified as having no dementia but lacked cognitive test scores. After consulting with the NHATS team, they indicated that these participants, or their proxy respondents, refused to complete the cognition test. In such cases, NHATS relied solely on self-reports of dementia instead of the tests. Since cognition scores were the main outcome measure, the level of missingness was minimal (0.5%), and these cases were excluded from the analysis. The final sample size in 2015 (Round 5) included 5566 older adults, with 31,038 observations from 2015 to 2022 (Rounds 5-12). The median number of follow-up observations per round was 3631.

Measures

Independent Variables

Five technology activities were the primary variables of interest in this study: (1) online shopping, (2) online banking, (3) online medication refills, (4) visiting social networking sites, and (5) checking health information online. These questions were asked across all 8 rounds of data collection. Each activity was

XSL•FO
RenderX

measured using dichotomous (yes or no) questions, as phrased in Table 1.

Instead of using raw binary variables for technology activities, we created new variables to capture cumulative transitions into and out of each activity over time. Specifically, a Cumulative Transition In variable was incremented by one each time an individual's response changed from "no=0" to "yes=1" for a given technology activity. Similarly, a Cumulative Transition Out variable was incremented by one each time a response shifted from "yes=1" to "no=0." This approach allowed us to focus on the effects of transitions (onset or cessation of use) rather than the mere status of technology use. This within-person asymmetric effect method has been used in previous studies, such as Kim and Han and Ghaiumy Anaraky et al [19,22]. Multimedia Appendix 1 shows how cumulative transitions were calculated for online shopping as an example.

Table . Exposure measurements.

Topics	Questions
Online shopping	In the last month, have you gone on the internet or online to shop for groceries or personal items?
	If needed: by personal items, we mean things such as toiletries that you can buy at the grocery or drug store.
Online banking	In the last month, have you gone on the internet or online to pay bills or do banking?
Online medication refills	In the last month, have you gone on the internet or online to order or refill prescriptions?
Visit social network sites	In the last month, have you gone on the internet or online to visit social network sites?
	If needed, these include sites where you can keep in touch with friends, such as Facebook or LinkedIn.
Check health information online	In the last year, have you gone on the internet or online to get information about your health conditions?

Dependent Variables

The primary outcome, cognitive function, was assessed across 3 domains. Episodic memory, the ability to recall personal experiences that are tied to specific times and locations, was scored from 0 to 20 based on immediate and delayed recall of 10 words. Executive function, which involves decision-making and problem-solving, was scored from 0 to 5 through the clock-drawing test. Orientation, the ability to recognize one's identity, spatial context, and time, was scored from 0 to 8 based on knowledge of the current location, time, date, and the names of the president or vice president. To enable comparisons across these differently scaled domains, each score was standardized by converting it to a *z* score (original score minus mean, divided by SD).

Covariates

Covariates included survey year (2015 - 2022), sex (female or male), age at baseline (continuous), race (non-Hispanic White, non-Hispanic Black, non-Hispanic Asian American and Pacific Islander, or Hispanic), education (no college degree, college degree, or beyond), living arrangement (alone or with others), number of devices owned (cell phone, computer, or tablet; ranging 0 - 3), rurality (metropolitan or nonmetropolitan), number of activities of daily living difficulties (discrete), number of IADL difficulties (discrete), self-rated health (poor, fair, good, very good, and excellent). Covariates were selected based on established associations with cognitive function in older adults. Demographic factors (sex, age, race, education, and living arrangement) were included as supported by Ghaiumy Anaraky et al [19]. Device ownership reflects digital access and has been linked to slower cognitive decline [23]. Activities of daily living and IADLs were included due to their strong connections to cognitive health [18,24]. Rurality and self-rated health are known predictors of cognitive outcomes [25,26].

Weighting

In our study, we applied analytic weights using the weighting variables from the sample person file [27]. Since the 2015 cohort was chosen as the foundation of our target population, we tackled the problem of multiple (repeated) observations through a multilevel modeling strategy that incorporated the NHATS sample design. The weights for each round were used to account for differential probabilities of selection and nonresponse, as well as for clustering and stratification variables. Adhering to the guidelines and code detailed in NHATS Technical Paper 23 [28], we used round-specific analytic weights and developed a new weight variable to represent individual-level weights within the multilevel models [28].

Statistical Analysis

We conducted generalized linear mixed-effects models separately for the 3 cognitive domains, both unadjusted and adjusted for covariates. The predictors were various types of technology activities, and the outcomes were the different domains of cognition, including episodic memory, executive function, and orientation.

General linear mixed-effects models are particularly suited for analyzing longitudinal or hierarchical data, where repeated measurements are nested within individuals. Given the nature of cognitive function may change over time due to aging and may also be influenced by exposure to different types of

XSL•FO
RenderX

technology, the use of general linear mixed-effects models provided an appropriate statistical framework for this analysis.

Two models were specified: the main effect model (Model A), which assessed the association between transitions in technology use and cognitive function, and the moderation effect model (Model B), which indicated the effects of technology use on the rate of change in cognitive functioning. For example, the main effect model (Model A) for episodic memory and online shopping was specified as follows:



The moderation effects model (Model B) added interaction terms for transitions and time, as follows:



The index *i* represents the individual sample ID, and *j* denotes the time point (ie, *j*=2015, 2016, ..., 2022).

We first estimated whether the start (transition in) and stop (transition out) of technology activities were associated with cognitive functioning at a given time (Model A). We then added an interaction term between transitions and time to explore whether starting or stopping technology activities could change the rate of cognitive decline over time (Model B).

Since we evaluated multiple outcomes, we applied the Benjamini-Hochberg false discovery rate correction to reduce the risk of type I errors [29]. To minimize the possibility of a reverse causal relationship between digital activities and cognitive domains, we conducted sensitivity analyses using 1-year lagged data for both Model A and Model B to ensure the associations hold over time.

Ethical Considerations

This study was exempted by the Johns Hopkins Medicine Institutional Review Board (IRB00460714), where informed consent is not required. The NHATS public use data is de-identified.

Results

Characteristics of the Study Sample at Baseline

The study sample characteristics at baseline are provided in Table 2. At baseline, the study sample (N=5566; representing an estimated 32,929,570 older adults living in communities in the United States) had an average age of 76.48 (SD 7.04) years and no dementia. The sample was predominantly female (3190, 57.31%), with well-distributed racial and ethnic groups. Most participants had no college degree (3738, 67.16%), lived with someone else (3803, 68.33%), and resided in metropolitan areas (4476, 80.42%). Furthermore, most participants owned one or more digital devices (eg, cell phone, computer, or tablet), and around a quarter of them had performed some technology activities in the past month.

XSL•FO
RenderX

Table . Study sample characteristics at baseline (N=5566).

Characteristic	Value
Age (years), mean (SD)	76.48 (7.04)
Episodic memory (0 - 20), mean (SD)	9.13 (2.93)
Executive function (0 - 5), mean (SD)	3.91 (0.93)
Orientation (0 - 8), mean (SD)	6.88 (1.24)
Sex, n (%)	
Male	2376 (42.69)
Female	3190 (57.31)
Race and ethnicity, n (%)	
Non-Hispanic White	4005 (71.95)
Non-Hispanic Black	1021 (18.34)
Non-Hispanic AAPI[a]	145 (2.66)
Hispanic	274 (4.92)
Missing	121 (2.17)
Number of device ownership (cell phone, computer, and tablet), n (%)	
0	376 (6.76)
1	1421 (25.53)
2	1949 (35.02)
3	1820 (32.70)
Education, n (%)	
No college degree	3738 (67.16)
College degree or beyond	1828 (32.84)
Living arrangement, n (%)	
Alone	1763 (31.67)
Living with someone	3803 (68.33)
Rurality, n (%)	
Metropolitan	4476 (80.42)
Non-metropolitan	1090 (19.58)
Number of difficulties in ADL[b], mean (SD)	0.63 (1.27)
Number of difficulties in IADL[c], mean (SD)	0.42 (0.86)
Self-rated health, n (%)	
Excellent	724 (13.01)
Very good	1717 (30.86)
Good	1917 (34.45)
Fair	975 (17.52)
Poor	231 (4.15)
Technology activities, n (%)	
Online shopping	1129 (20.3)
Online banking	1553 (27.9)
Online medication refills	665 (11.9)
Visit social network sites	1414 (25.4)
Checking health information online	1198 (21.5)

XSL•FO
RenderX

[a]AAPI: Asian American and Pacific Islander.

[b]ADL: activities of daily living.

[c]IADL: instrumental activities of daily living.

The Associations Between Technology Activities and Cognitive Domains

Table 3 shows the beta coefficients for the initiation and cessation of each technology activity across various cognitive domains. These findings stem from Model A and address research question 1—the associations between technology activities (online shopping, banking, medication refills, social media use, and checking health information) and cognitive domains (episodic memory, executive function, and orientation). *P* values are marked after the false discovery rate correction using the Benjamini-Hochberg method. The *P* values before and after adjustment are provided in Multimedia Appendix 2.

Transitions into online shopping were associated with an increase in the score of episodic memory at a given wave (=.046, *P*=.02; 95% CI 0.013-0.080; Model A), as well as an increase in the score of executive function (=.041, *P*=.043; 95% CI 0.007-0.075; Model A) and orientation at a given wave (=.091, *P*<.001; 95% CI 0.055-0.126; Model A).

Starting online banking was associated with an increase in executive function at a given wave (=.061, *P*=.043; 95% CI 0.010-0.112; Model A) and orientation (=.077, *P*=.004; 95% CI 0.032-0.123; Model A). Stopping online banking was associated with decreasing scores of episodic memory (=−.078, *P*=.01; 95% CI −0.129 to −0.026; Model A).

Transitions into online medication refills were associated with increasing episodic memory (=.073, *P*<.001; 95% CI 0.038-0.109; Model A) and increasing orientation at a given wave (=.097, *P*<.001; 95% CI 0.054-0.140; Model A).

Starting to visit social networking sites online was associated with increasing episodic memory (=.065, *P*=.01; 95% CI 0.022-0.109; Model A) and orientation at a given wave (=.080, *P*<.001; 95% CI 0.041-0.119; Model A). Stopping visits to social network sites online was associated with decreasing episodic memory (=−.066, *P*=.004; 95% CI −0.105 to −0.028; Model A). Transitions into checking health information online were associated with increasing orientation (=.098, *P*<.001; 95% CI 0.068-0.127; Model A) at a given wave.

Among all the technology activities examined, online shopping and online banking showed the widest impact across various areas of cognitive function. Combining our main findings with results from lagged models in sensitivity analyses (Multimedia Appendix 3), older adults who started using online shopping and social media were related to better orientation; those who began managing prescriptions digitally showed cognitive benefits in episodic memory; and those who stopped using online banking were linked to a decline in episodic memory. Notably, the effects were statistically significant but small in magnitude.

Table . Effects of starting and stopping technology activities on cognitive domains (Model A).

Technology activity	Episodic memory *z* score, β (95% CI)		Executive function *z* score, β (95% CI)		Orientation *z* score, β (95% CI)	
	Start	Stop	Start	Stop	Start	Stop
Online shopping	0.046[a](0.013 to 0.080)	0.019 (−0.020 to 0.058)	0.041[b] (0.007 to 0.075)	0.004 (−0.062 to 0.069)	0.091[c] (0.055 to 0.126)	0.010 (−0.026 to 0.046)
Online banking	0.049 (−0.003 to 0.100)	−0.078[a] (−0.129 to 0.026)	0.061[b] (0.010 to 0.112)	−0.059[b] (−0.117 to 0.000)	0.077[a] (0.032 to 0.123)	−0.062 (−0.126 to 0.003)
Medication refills	0.073[c] (0.038 to 0.109)	−0.014 (−0.068 to 0.040)	0.026 (−0.014 to 0.067)	0.011 (−0.038 to 0.059)	0.097[c] (0.054 to 0.140)	−0.007 (−0.055 to 0.040)
Social media	0.065[a] (0.022 to 0.109)	−0.066[a] (−0.105 to 0.028)	0.050 (−0.002 to 0.101)	−0.045 (−0.094 to 0.004)	0.080[c] (0.041 to 0.119)	−0.024 (−0.068 to 0.021)
Check health information	0.047 (0.005 to 0.089)	0.027 (−0.014 to 0.067)	−0.011 (−0.054 to 0.031)	0.060[c] (0.032 to 0.088)	0.098[c] (0.068 to 0.127)	0.048[b] (0.010 to 0.085)

[a]*P*<.01.

[b]*P*<.05.

[c]*P*<.001.

Effects of Technology Activities on the Rate of Cognitive Decline

In Model B, we added an interaction term between transitions of technology activities and time to assess their effect on the rate of cognitive decline. Table 4 shows the beta coefficients for the interaction terms between shifts in technology activities and time, reflecting the rate of cognitive decline. A positive value indicates a reduction in cognitive decline, while a negative score suggests an increase in cognitive decline.

Stopping online shopping was associated with a faster rate of cognitive decline in episodic memory (=−.023, *P*=.047; 95% CI −0.041 to−0.005; Model B). The transition into online shopping was associated with a slower cognitive decline in the orientation domain (=.049, *P*<.001; 95% CI 0.035-0.063; Model B). The onset of online banking was associated with a

XSL•FO
RenderX

mitigated rate of orientation decline (=.049, *P*<.001; 95% CI 0.035-0.063; Model B). The onset of online medication refill was associated with a slower orientation decline (=.052, *P*<.001; 95% CI 0.036-0.068; Model B). The onset of social media use was associated with a slower decline in the orientation domain (=.039, *P*<.001; 95% CI 0.025-0.053; Model B). The transition into checking health information online was associated with a mitigated rate of orientation decline (=.036, *P*<.001; 95% CI 0.022-0.051; Model B). When incorporating findings from lagged models in sensitivity analyses (Multimedia Appendix 3), regarding the interaction with time, beginning to use any of the studied digital technologies was linked to significantly better but small preservation of orientation.

The summary of all models, stratified by technology activities, is provided in Multimedia Appendix 4. Several noteworthy confounders emerged in the analysis. A higher number of device owners was consistently associated with better cognitive functioning across all technology activities. Similarly, individuals with higher educational attainment (college degree or beyond) demonstrated superior performance across all cognitive domains. Females exhibited stronger effects specifically in episodic memory, regardless of the technology activity.

Table . The interaction coefficients of starting and stopping technology activities on cognitive domains (Model B).

Technology activity	Episodic memory *z* score, β (95% CI)		Executive function *z* score, β (95% CI)		Orientation *z* score, β (95% CI)	
	Start	Stop	Start	Stop	Start	Stop
Online shopping	0.014 (0.000 to 0.027)	−0.023 [a](−0.041 to 0.005)	0.004 (−0.011 to 0.018)	−0.000 (−0.016 to 0.016)	0.049[b] (0.035 to 0.063)	−0.011 (−0.026 to 0.004)
Online banking	0.016 (−0.001 to 0.032)	−0.010 (−0.030 to 0.011)	0.020 (−0.003 to 0.043)	−0.017 (−0.045 to 0.010)	0.044[b] (0.021 to 0.067)	0.005 (−0.016 to 0.026)
Medication refills	−0.001 (−0.017 to 0.015)	−0.012 (−0.037 to 0.012)	0.015 (−0.006 to 0.037)	−0.013 (−0.042 to 0.016)	0.052[b] (0.036 to 0.068)	0.005 (−0.012 to 0.021)
Social media	0.010 (−0.005 to 0.026)	−0.022 (−0.040 to 0.004)	−0.008 (−0.026 to 0.009)	0.008 (−0.014 to 0.030)	0.039[b] (0.025 to 0.053)	−0.002 (−0.018 to 0.015)
Check health information	0.001 (−0.014 to 0.015)	−0.006 (−0.021 to 0.009)	0.004 (−0.014 to 0.022)	0.002 (−0.018 to 0.022)	0.036[b] (0.022 to 0.051)	0.020[c] (0.005 to 0.034)

[a]*P*<.05.

[b]*P*<.001.

[c]*P*<.01.

Discussion

Principal Findings

This study highlights the distinct effects of technology use on specific cognitive domains among older adults. Compared with other digital activities, the use of online shopping and online banking services showed the most comprehensive influence on the 3 cognitive domains. Similar to findings from a previous study, online shopping and banking involve a series of actions such as memorizing accounts and passwords, thinking about recipes and corresponding products to purchase, or financial plans to arrange, and remaining oriented to dates, time, and their own names to process transactions [18,30].

Older adults who began using online shopping and social media showed a small improvement in orientation. These activities may enhance spatial and temporal awareness by requiring users to follow delivery schedules, navigate websites, and engage with time-stamped social content. Starting to manage prescriptions online was linked to better episodic memory. This task likely activates memory systems through routine recall of medications and schedules, reinforcing memory through repeated use and the integration of new and existing information. In contrast, stopping online banking was associated with a decline in episodic memory. Online banking involves complex memory tasks such as recalling passwords and managing finances. Disengaging from such cognitively demanding activities may reduce mental stimulation, contributing to memory decline. However, all observed effects were small in magnitude, and their clinical significance remains uncertain. Future trials are needed to evaluate the clinical relevance of targeted digital engagement interventions on specific cognitive domains.

Limitations

This study used observational data, which introduces the possibility of reverse causation. It is plausible that older adults with lower cognitive performance were less likely to adopt or continue using technology. However, findings from Choi et al [31], using a cross-lagged panel analysis of NHATS data, suggest that technology use predicts cognitive improvements over time, particularly in episodic memory and executive function. While episodic memory also predicted subsequent technology use, executive function did not, and orientation was not assessed in their study [31]. Similarly, Hartanto et al [32], using data from the Midlife in the United States study, found that computer use significantly predicted positive changes in executive function. Their analysis showed consistent associations between computer use and improvements across all executive function tasks, though no significant effect was observed for episodic memory. Collectively, these findings support a stronger directional relationship from technology use

XSL•FO
RenderX

to cognitive function, although a bidirectional relationship cannot be ruled out.

Some limitations stemmed from the instrument used by NHATS. For example, the question about social media use simply asked whether the participant had gone online to visit social network sites (eg, Facebook and LinkedIn) in the last month. The specific platform, content on the sites, and the level of engagement were not asked. We recognize the heterogeneity of social media and the potential varied associations with cognitive function. In addition, digital activity questions were designed to help participants recall their technology use behaviors over the past month or year, depending on the activity. This study only assessed changes in technology use and cognition over 8 years, with the smallest time unit being a year. Thus, one instance of use within a given year was coded as "yes," regardless of frequency. We acknowledge that this annual resolution limits our ability to capture short-term fluctuations in digital engagement or the precise timing of technology use relative to cognitive assessments.

Attrition in this study occurred through 3 primary mechanisms: death, nonresponse, and transitions from community to institutional settings. While the NHATS cohort maintained high response rates (>90%) with relatively low attrition per round [21], we used weighted analyses to account for potential attrition bias. Analysis of NHATS data from 2011 to 2018 revealed that 86.2% of older adults remained in community settings, 9% transitioned to residential care, and 4.9% to nursing homes [33]. Although only a small proportion of older adults changed settings, these transitions predominantly occurred among older adults with greater functional and cognitive limitations, potentially limiting our ability to capture technology access and use patterns among this more vulnerable subgroup.

Another limitation concerns the last few years of the study period, which coincided with the COVID-19 pandemic. Social isolation during this time was particularly acute for older adults [34]. Evidence from NHATS data shows that older adults who reduced their use of emails, texts, and phone calls during the pandemic, especially those with lower socioeconomic status, showed a limited increase in video calls to stay connected with family and friends [35]. The pandemic's potential impact on technology use trends and its subsequent influence on cognitive function remains unclear, introducing uncertainty to the study's results.

In considering generalizability, it is important to note that wired high-speed internet access remains more prevalent in urban than rural areas—available in 76% of noninstitutional households in urban settings compared with 62% in rural areas as of 2023 in the United States [36]. While our findings highlight the potential cognitive benefits of certain digital activities, real-world implementation may be limited by disparities in local infrastructure and household financial capacity. Without reliable internet access or the resources to obtain and maintain digital devices, many older adults, particularly in underserved areas, may be unable to engage in these beneficial activities.

Comparison With Previous Work

The observed association between social media use and cognitive function aligns with previous research. For example, Byrne and Ghaiumy Anaraky [37] found that greater use of social technology was linked to enhanced cognitive performance among socially isolated older adults, though their outcome measure was a global cognitive score rather than disaggregated domains. Anaraky et al [19] reported that stopping the use of computers, tablets, the internet, texts, and emails predicted declines in episodic memory, executive function, and orientation. Similarly, Kim and Han [22] reported that initiating internet use was associated with improved cognitive functioning and a slower rate of cognitive decline over time, while ceasing internet use was linked to worse cognitive outcomes and accelerated decline. These findings align with this study's results, which show that both initiating and maintaining technology use are crucial for cognitive health. Our study provides more granular insights, demonstrating the unique effects of various technology activities on episodic memory, executive function, and orientation.

Conclusions

Our findings suggest that interventions aiming to maintain episodic memory in older adults should focus on encouraging the adoption of online medication refill services and preventing disengagement from online banking, particularly among those already accustomed to these activities. To preserve orientation, efforts should prioritize supporting older adults in initiating and maintaining online shopping and social media use. More broadly, all types of digital engagement examined in this study—including instrumental, social, and informational activities—were associated with a slower decline in orientation over time. Teaching older adults new technology skills and providing resources to prevent disengagement from technology activities may help preserve cognitive health. Furthermore, when designing technology-based interventions, researchers can use these findings to target specific cognitive domains with appropriate technology activities. By tailoring interventions to match the cognitive benefits of different activities, resources can be allocated more effectively to maximize their impact on cognitive well-being in older adults.

Acknowledgments

EJ reported grants from RGF Environmental Group outside the submitted work; is a cofounder and member of the board of directors of Plans4Care Inc, a digital health company that provides personalized, on-demand dementia care; serves as a consultant to the Lewin Group under their Centers for Medicare & Medicaid Services (CMS) GUIDE contract; and is a member of the board of directors for PACE-RI. EMS reports personal consulting fees from Corrie Health. Outside the submitted work, EMS reports support from the American Heart Association (grant nos 20SFRN35380046 and 878924), the Patient - Centered Outcomes Research Institute (award no IHS - 2021C3 - 24 147), and the National Institutes of Health (grant no U01HL096812).

XSL•FO
RenderX

Funding

This work was supported by the 2024 Pilot Grant from the Center for Equity in Aging at the Johns Hopkins University School of Nursing.

Authors' Contributions

ECH contributed to conceptualization, data curation, formal analysis, methodology, validation, visualization, and writing of the original draft. EJ contributed to methodology, supervision, and writing, including review and editing. ES contributed to supervision and writing, including review and editing.

Conflicts of Interest

EMS reports receiving personal consulting fees from Corrie Health. EJ is a co-founder and member of the board of directors of Plans4Care Inc., a digital health company that provides personalized dementia care on demand. He also serves on the finance committee for PACE-RI.

Multimedia Appendix 1
An example of data structure.
[DOCX File, 22 KB - aging_v8i1e77227_app1.docx]

Multimedia Appendix 2
Raw and false discovery rate (FDR)-adjusted *P* values for Models A and B.
[DOCX File, 29 KB - aging_v8i1e77227_app2.docx]

Multimedia Appendix 3
Sensitivity analyses using lagged models.
[DOCX File, 23 KB - aging_v8i1e77227_app3.docx]

Multimedia Appendix 4
Asymmetric random-effects models of technology activities and cognitive domains (*z* score).
[DOCX File, 76 KB - aging_v8i1e77227_app4.docx]

References

1. Faverio M. Share of tech users among Americans 65 and older grew in past decade. Pew Research Center. 2022 Jan 13. URL: https://www.pewresearch.org/short-reads/2022/01/13/share-of-those-65-and-older-who-are-tech-users-has-grown-in-the-past-decade/ [accessed 2024-04-10]
2. Lee MA, Ferraro KF, Kim G. Digital technology use and depressive symptoms among older adults in Korea: beneficial for those who have fewer social interactions? Aging Ment Health 2021 Oct 3;25(10):1839-1847. [doi: 10.1080/13607863.2020.1839863]
3. Sen K, Prybutok G, Prybutok V. The use of digital technology for social wellbeing reduces social isolation in older adults: a systematic review. SSM Popul Health 2022 Mar;17:101020. [doi: 10.1016/j.ssmph.2021.101020] [Medline: 35024424]
4. Schepens Niemiec SL, Lee E, Saunders R, Wagas R, Wu S. Technology for activity participation in older people with mild cognitive impairment or dementia: expert perspectives and a scoping review. Disabil Rehabil Assist Technol 2023 Nov;18(8):1555-1576. [doi: 10.1080/17483107.2022.2116114] [Medline: 36067094]
5. Kasper JD, Freedman VA, Spillman B. Classification of persons by dementia status in the national health and aging trends study. Technical Paper #5.: Johns Hopkins University School of Public Health; 2013 URL: https://www.nhats.org/sites/default/files/inline-files/DementiaTechnicalPaperJuly_2_4_2013_10_23_15.pdf
6. Cipolli GC, de Assumpção D, Borim FSA, et al. Cognitive impairment predicts sarcopenia 9 years later among older adults. J Am Med Dir Assoc 2023 Aug;24(8):1207-1212. [doi: 10.1016/j.jamda.2023.05.008] [Medline: 37311558]
7. Ho HT, Lin SI, Guo NW, Yang YC, Lin MH, Wang CS. Executive function predict the quality of life and negative emotion in older adults with diabetes: a longitudinal study. Prim Care Diabetes 2022 Aug;16(4):537-542. [doi: 10.1016/j.pcd.2022.05.002] [Medline: 35659729]
8. Martin P, Poon LW, Lee G, et al. A 13-year time-lagged description of general cognitive and functional abilities in older men: a cross-lagged panel model. J Aging Health 2023 Jun;35(5-6):335-344. [doi: 10.1177/08982643221130381] [Medline: 36194185]
9. Numbers K, Crawford JD, Kochan NA, Draper B, Sachdev PS, Brodaty H. Participant and informant memory-specific cognitive complaints predict future decline and incident dementia: findings from the Sydney Memory and Ageing Study. PLoS ONE 2020;15(5):e0232961. [doi: 10.1371/journal.pone.0232961] [Medline: 32396544]

XSL•FO
RenderX

10. Schaeverbeke JM, Gabel S, Meersmans K, et al. Baseline cognition is the best predictor of 4-year cognitive change in cognitively intact older adults. Alzheimers Res Ther 2021 Apr 7;13(1):75. [doi: 10.1186/s13195-021-00798-4] [Medline: 33827690]
11. Stites SD, Harkins K, Rubright JD, Karlawish J. Relationships between cognitive complaints and quality of life in older adults with mild cognitive impairment, mild Alzheimer disease dementia, and normal cognition. Alzheimer Dis Assoc Disord 2018;32(4):276-283. [doi: 10.1097/WAD.0000000000000262] [Medline: 29944474]
12. Zhu CW, Sano M, Ferris SH, Whitehouse PJ, Patterson MB, Aisen PS. Health-related resource use and costs in elderly adults with and without mild cognitive impairment. J Am Geriatr Soc 2013 Mar;61(3):396-402. [doi: 10.1111/jgs.12132] [Medline: 23414481]
13. Rubin RD, Watson PD, Duff MC, Cohen NJ. The role of the hippocampus in flexible cognition and social behavior. Front Hum Neurosci 2014;8(SEP):742. [doi: 10.3389/fnhum.2014.00742] [Medline: 25324753]
14. Eichenbaum H. Prefrontal-hippocampal interactions in episodic memory. Nat Rev Neurosci 2017 Sep;18(9):547-558. [doi: 10.1038/nrn.2017.74] [Medline: 28655882]
15. Funahashi S, Andreau JM. Prefrontal cortex and neural mechanisms of executive function. J Physiol Paris 2013 Dec;107(6):471-482. [doi: 10.1016/j.jphysparis.2013.05.001] [Medline: 23684970]
16. Kaski D, Quadir S, Nigmatullina Y, Malhotra PA, Bronstein AM, Seemungal BM. Temporoparietal encoding of space and time during vestibular-guided orientation. Brain (Bacau) 2016 Feb;139(Pt 2):392-403. [doi: 10.1093/brain/awv370] [Medline: 26719385]
17. Lee MT, Jang Y, Chang WY. How do impairments in cognitive functions affect activities of daily living functions in older adults? PLoS ONE 2019;14(6):e0218112. [doi: 10.1371/journal.pone.0218112] [Medline: 31173607]
18. Li J, He R, Hsu EC, Li J. Network analysis of key instrumental activities of daily living and cognitive domains for targeted intervention in US older adults without dementia: cross-sectional study. JMIR Aging 2025 Mar 19;8:e67632. [doi: 10.2196/67632] [Medline: 40106829]
19. Ghaiumy Anaraky R, Schuster AM, Cotten SR. Can changes in older adults' technology use patterns be used to detect cognitive decline? Gerontologist 2024 Jul 1;64(7):gnad158. [doi: 10.1093/geront/gnad158] [Medline: 37974489]
20. Wang K, Chen XS, Kang SY, Smith BD, Gu D. Older adults' online activities and cognition: investigating the psychological mechanisms and age and gender differences. Soc Sci Med 2024 Jul;352:116988. [doi: 10.1016/j.socscimed.2024.116988] [Medline: 38820692]
21. Freedman VA, Kasper JD. Cohort profile: the National Health and Aging Trends Study (NHATS). Int J Epidemiol 2019 Aug 1;48(4):1044-1045g. [doi: 10.1093/ije/dyz109] [Medline: 31237935]
22. Kim YK, Han SH. Internet use and cognitive functioning in later life: focus on asymmetric effects and contextual factors. Gerontologist 2022 Mar 28;62(3):425-435. [doi: 10.1093/geront/gnab149] [Medline: 34614179]
23. Jin Y, Jing M, Ma X. Effects of digital device ownership on cognitive decline in a middle-aged and elderly population: longitudinal observational study. J Med Internet Res 2019 Jul 29;21(7):e14210. [doi: 10.2196/14210] [Medline: 31359864]
24. Fauth EB, Schwartz S, Tschanz JT, Østbye T, Corcoran C, Norton MC. Baseline disability in activities of daily living predicts dementia risk even after controlling for baseline global cognitive ability and depressive symptoms. Int J Geriatr Psychiatry 2013 Jun;28(6):597-606. [doi: 10.1002/gps.3865] [Medline: 22968965]
25. Glauber R. Rural depopulation and the rural-urban gap in cognitive functioning among older adults. J Rural Health 2022 Sep;38(4):696-704. [doi: 10.1111/jrh.12650] [Medline: 35257439]
26. Bond J, Dickinson HO, Matthews F, Jagger C, Brayne C, MRC CFAS. Self-rated health status as a predictor of death, functional and cognitive impairment: a longitudinal cohort study. Eur J Ageing 2006 Nov 17;3(4):193-206. [doi: 10.1007/s10433-006-0039-8]
27. Freedman VA, Kasper JD. National health and aging trends study development of round 5 survey weights. NHATS Technical Paper #14. 2016. URL: https://www.nhats.org/ [accessed 2024-07-10]
28. Freedman VA, Hu M, Dematteis J, Kasper JD. Accounting for sample design in NHATS and NSOC analyses: frequently asked questions. URL: https://www.nhats.org/ [accessed 2024-07-10]
29. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B 1995 Jan 1;57(1):289-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
30. Okahashi S, Seki K, Nagano A, Luo Z, Kojima M, Futaki T. A virtual shopping test for realistic assessment of cognitive function. J NeuroEngineering Rehabil 2013;10(1):59. [doi: 10.1186/1743-0003-10-59]
31. Choi EY, Wisniewski KM, Zelinski EM. Information and communication technology use in older adults: a unidirectional or bi-directional association with cognitive function? Comput Human Behav 2021 Aug;121:106813. [doi: 10.1016/j.chb.2021.106813] [Medline: 33986562]
32. Hartanto A, Yong JC, Toh WX, Lee STH, Tng GYQ, Tov W. Cognitive, social, emotional, and subjective health benefits of computer use in adults: a 9-year longitudinal study from the Midlife in the United States (MIDUS). Comput Human Behav 2020 Mar;104:106179. [doi: 10.1016/j.chb.2019.106179]
33. Chyr LC, Drabo EF, Fabius CD. Patterns and predictors of transitions across residential care settings and nursing homes among community-dwelling older adults in the United States. Gerontologist 2020 Nov 23;60(8):1495-1503. [doi: 10.1093/geront/gnaa070] [Medline: 32598465]

XSL•FO
RenderX

34. Lazzari C, Rabottini M. COVID-19, loneliness, social isolation and risk of dementia in older people: a systematic review and meta-analysis of the relevant literature. Int J Psychiatry Clin Pract 2022 Jun;26(2):196-207. [doi: 10.1080/13651501.2021.1959616] [Medline: 34369248]
35. Zhang Y, Luo A, Hou L, et al. Crisis response during the COVID-19 pandemic: changes in social contact and social participation of older Americans. PLoS ONE 2023;18(7):e0284935. [doi: 10.1371/journal.pone.0284935]
36. NTIA data explorer. National Telecommunications and Information Administration. URL: https://www.ntia.gov/data/explorer#sel=pcOrTabletUser&demo=metro&pc=isPersonP&disp=chart [accessed 2025-09-11]
37. Byrne KA, Ghaiumy Anaraky R. Identifying racial and rural disparities of cognitive functioning among older adults: the role of social isolation and social technology use. J Gerontol B Psychol Sci Soc Sci 2022 Oct 6;77(10):1779-1790. [doi: 10.1093/geronb/gbac055] [Medline: 35312775]

Abbreviations

IADL: instrumental activities of daily living
NHATS: National Health and Aging Trends Study

Edited by H Köttl; submitted 09.05.25; peer-reviewed by A Hartanto, H Sasai, IC Lai; revised version received 26.09.25; accepted 16.10.25; published 25.11.25.

Please cite as:
Hsu EC, Spaulding EM, Jutkowitz E
Technology Activities and Cognitive Trajectories Among Community-Dwelling Older Adults: National Health and Aging Trends Study
JMIR Aging 2025;8:e77227
URL: https://aging.jmir.org/2025/1/e77227
doi:10.2196/77227

© Erh-Chi Hsu, Erin M Spaulding, Eric Jutkowitz. Originally published in JMIR Aging (https://aging.jmir.org), 25.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using Virtual Reality to Assess Spatial Navigation Ability in Individuals With Mild Cognitive Impairment and Older Adults: Cross-Sectional Study

Ming-Chyi Pai[1,2,3], MD, PhD; Yuh-Ting Lin[3], MS; Chung-Yao Hsiao[1], MD; Chia-Hung Lai[4]; Chih-Jen Chen[4]; Chaur-Jong Hu[5,6], MD; Cheng-Yu Chen[7,8,9,10,11], MD

[1]Division of Behavioral Neurology, Department of Neurology, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, 138 Sheng-Li Road, North District, Tainan, Taiwan

[10]Department of Radiology, National Defense Medical Center, Taipei, Taiwan

[11]Neuroscience Research Center, Taipei Medical University, Taipei, Taiwan

[2]Alzheimer's Disease Research Center, National Cheng Kung University Hospital, Tainan, Taiwan

[3]Institute of Gerontology, College of Medicine, National Cheng Kung University, Tainan, Taiwan

[4]Department of Medicine, National Cheng Kung University Hospital, Tainan, Taiwan

[5]Department of Neurology, Medical College, Taipei Medical University, Taipei, Taiwan

[6]Department of Neurology, Shuang-Ho Hospital, Taipei Medical University, New Taipei, Taiwan

[7]Department of Medical Imaging, Taipei Medical University Hospital, Taipei, Taiwan

[8]Department of Radiology, School of Medicine, Taipei Medical University, Taipei, Taiwan

[9]Translational Imaging Research Center, Taipei Medical University Hospital, Taipei, Taiwan

Corresponding Author:
Ming-Chyi Pai, MD, PhD
Division of Behavioral Neurology, Department of Neurology, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, 138 Sheng-Li Road, North District, Tainan, Taiwan

Abstract

Background: Spatial navigation impairment is prevalent in people with Alzheimer disease (AD) and may appear in its initial clinical stage. Detecting this deficit in people at risk may not only help prevent them from getting lost or going missing but also provide a useful clinical aid to accurate diagnosis. Traditional assessments for spatial navigation impairment include questionnaires, paper-and-pencil and maze tests, or video games. While a real-world setting is more valid, direct, and accurate, it is plagued by unpredictable conditions such as weather, obstacles, or accidents. Owing to modern technology, virtual reality (VR) offers a new way to test spatial navigation impairment.

Objective: The aims of this study were to test the feasibility of a VR setting to assess sense of location in people with mild cognitive impairment (MCI) and the power of VR to discriminate among groups with different clinical conditions.

Methods: We used the Pai-Jan virtual reality (PJVR) device to test spatial navigation ability in those who were cognitively unimpaired (CU) and those who experienced subjective cognitive decline (SCD) and MCI. The PJVR device is the VR version (VIVE Pro Eye head-mounted display) of the Pai-Jan device, which has demonstrated its power to discriminate among CU, AD MCI, and mild AD dementia. With a map provided and using joysticks or handles, participants were asked to reach 5 points on a 660-m path. Linear deviation (LD; in meters) from each target point and vector deviation (in degrees) from the direction to the start point at each location were treated as the variables for comparison.

Results: A total of 113 participants provided informed consent to initiate the study. Of these 113 participants, 93 (82.3%) completed the trials, including 22 (24%) who were CU, 39 (42%) with SCD, and 32 (34%) with MCI. In total, 17.7% (20/113) failed the trials due to cybersickness. The mean LD of the CU, SCD, and MCI groups was 38.2 (SD 39.5), 50.4 (SD 40.7), and 100.4 (SD 46.2) meters, respectively (P<.001). The MCI group showed greater vector deviation (mean 63.2, SD 42.4 degrees) than either the SCD (mean 39.4, SD 33.0 degrees) or CU (mean 38.6, SD 37.4 degrees; P=.02) group. The LD of the PJVR device was correlated with the total scores on the caregiver version of the Questionnaire on Everyday Navigational Ability (P<.001), indicating good ecological validity.

Conclusions: The PJVR device is feasible for older adults and participants with MCI. It can detect spatial navigation deficits related to AD pathology, and the results show a close correlation with real-world navigation ability.

(JMIR Aging 2025;8:e75952) doi:10.2196/75952

XSL•FO
RenderX

KEYWORDS

virtual reality; mild cognitive impairment; spatial navigation; dementia; Alzheimer disease

Introduction

Spatial navigation impairment (SNI) is frequently encountered in people with Alzheimer disease (AD) and may appear in the early clinical stage of AD [1]. These patients usually compensate for their difficulties by using geographic or literal cues, taking fixed routes, or reducing their range of ambulation. As the disease advances, their ability to construct allocentric spatial representations and the translation between spatial reference frames becomes progressively impaired [2]. At this stage, they are no longer able to adopt alternatives when they are disoriented and may get lost [3,4]. Getting lost may result in going missing when efficient problem-solving strategies are absent or when timely help from pedestrians is unavailable. Although getting lost or going missing usually occurs in people with AD, it may also occur in those with neurodegenerative disorders due to non-AD pathology (MC Pai, unpublished data, April 2024). Under such conditions, the manifestations may be different from those of AD. People with Lewy body diseases, for example, may lose their way due to impaired egocentric spatial memory, executive dysfunction, or attentional deficit [5].

Impaired episodic memory is arguably a powerful indicator of hippocampal dysfunction and, hence, a valuable clinical feature for a diagnosis of AD. Notably, SNI or even getting lost may appear before episodic memory impairment in some patients with AD as one of the incipient symptoms [1]. In focal brain region models, the medial temporal lobe, especially the hippocampus, reflected the activity of allocentric navigation and episodic memory [6]. As the brain regions damaged by initial AD pathology overlap with the brain areas essential for human spatial navigation, SNI has been proposed as a marker for the diagnosis of AD [7-10]. Hence, the detection of SNI can not only prevent instances of getting lost or going missing but also identify early symptoms of clinical AD.

Currently, many tests have been developed to assess spatial navigation abilities. A real-world and large-scale setting is definitely more valid, direct, and accurate. However, unpredictable or unfavorable conditions may arise during testing, such as bad weather, physical discomfort, obstacles, or even accidents [11]. Nowadays, the various forms of virtual reality (VR) are mature and ready to assess spatial cognition in humans [12,13] (Table 1). For example, Tu et al [14] designed an indoor-environment VR test that was used to test spatial navigation abilities and episodic memory to distinguish AD from frontotemporal dementia, whereas Rekers and Finke [15] provided a novel design that reduced the confounding factors of fine motor skills to test the two abilities. However, many patients with AD lost their way in the outdoor environment, causing great life-threatening issues and social burden. Therefore, we hoped to design a large-scale virtual world to test SNI based on our previous model conducted in the real world.

XSL•FO
RenderX

Table . Examples of tests to assess spatial navigation ability.

Category	Examples
Function assessments	• SBSOD[a] [16] • QuENA[b] [17] • WQ[c] [18]
Pencil-and-paper or maze test	• FMT[d] [19]
Desktop VR[e] test	• CG[f] Arena [20,21] • SHQ[g] [22] • vYSA[h] [23] • The Navigation Test Suite [24,25]
Real-world settings	• PJ[i] test [26] • DNT[j] [27]
Immersive VR	• VST[k] [14] • Triangle completion test [12] • VIENNA[l] [15] • PJVR[m] [13] • VR Morris Water Maze [28] • VENLab[n] [29] • Loop closure test [30] • Y-maze task [31]
Others	• TCTs[o] [32] • HGT[p] [33]

[a]SBSOD: Santa Barbara Sense of Direction Scale.

[b]QuENA: Questionnaire on Everyday Navigational Ability.

[c]WQ: Wayfinding Questionnaire.

[d]FMT: floor maze test.

[e]VR: virtual reality.

[f]CG: computer generated.

[g]SHQ: Sea Hero Quest.

[h]vYSA: virtual Y-maze strategy assessment.

[i]PJ: Pai-Jan.

[j]DNT: detour navigation test.

[k]VST: virtual supermarket task.

[l]VIENNA: virtual environments navigation assessment.

[m]PJVR: virtual reality version of the Pai-Jan device.

[n]VENLab: Virtual Environment Navigation Lab.

[o]TCT: triangle completion task.

[p]HGT: hidden goal task.

As a result, the aims of this study were to test the feasibility of a VR setting to assess sense of location in people with mild cognitive impairment (MCI) and the power of VR to discriminate among groups with different clinical conditions. We hypothesized that the MCI group would have the poorest navigational performance compared to the cognitively unimpaired (CU) and subjective cognitive decline (SCD) groups, whereas the CU group would perform better than the SCD group.

Methods

Participants

We conducted a prospective study on individuals with cognitive impairment or concerns who regularly visited a special clinic in a university hospital. Participants included individuals with MCI and SCD and individuals who were CU. A diagnosis of MCI was made by a senior behavioral neurologist, MCP, based on patients' clinical manifestations, the results of neuropsychological examinations, brain magnetic resonance imaging, and single-photon emission computed tomography

XSL•FO
RenderX

[34]. Some participants received tests for AD biomarkers (Table 2). All participants with MCI met the Petersen criteria for MCI [35]: (1) memory concerns, preferably corroborated by an informant; (2) impaired memory function for their age and educational level; (3) preserved general cognitive function; (4) intact activities of daily living; and (5) absence of dementia. Similarly, a diagnosis of SCD was based on the chief concern of the patients and the scores on the Subjective Cognitive Decline Questionnaire (SCD-Q) [36].

Table . The results of plasma pTau217 levels and amyloid positron emission tomography (PET).

	CU[a] (n=22), n (%)	SCD[b] (n=39), n (%)	MCI[c] (n=32), n (%)	Total (N=93), n (%)
Plasma pTau217				
Low level	10 (45)	12 (31)	7 (22)	29 (31)
Medium level	1 (5)	3 (8)	3 (9)	7 (8)
High level	3 (14)	4 (10)	15 (47)	22 (24)
Amyloid PET				
Negative	1 (5)	17 (44)	2 (6)	20 (22)
Positive	0 (0)	1 (3)	2 (6)	3 (3)
Unknown	7 (32)	2 (5)	3 (9)	12 (13)
ADNC[d]	3 (14)	5 (13)	17 (53)	25 (27)

[a]CU: cognitively unimpaired.

[b]SCD: subjective cognitive decline.

[c]MCI: mild cognitive impairment.

[d]ADNC: Alzheimer disease neuropathologic changes, including high plasma pTau217 level or positive amyloid PET.

Because of the novelty of using this device to quantify spatial disorientation and the absence of established normative benchmarks for key outcome measures such as linear deviations (LDs) and vector deviations (VDs), a priori sample size calculation was based on effect sizes reported in related literature [13]. Assuming a large effect size (Cohen *f*=0.40), a significance level of .05, and a statistical power of 80%, we estimated that a total sample size of 84 participants would be sufficient to detect statistically significant group differences using one-way ANOVA. To account for potential exclusions due to VR-related adverse effects or incomplete data, we planned for an initial recruitment target of 90 to 100 participants, anticipating an attrition rate of up to 20%.

Basic Measurements

Basic measurements included the Cognitive Abilities Screening Instrument [37], Clinical Dementia Rating Scale [38], Geriatric Depression Scale [39], SCD-Q [36], and Questionnaire on Everyday Navigational Ability (QuENA) [17]. The QuENA consists of patient (QuENA-P) and caregiver (QuENA-CG) versions. The definition of caregivers in this study was the spouses, children, or key care partners of the participants. We used perceived ease of use [40], the Computer Anxiety Scale [41], and perceived enjoyment [42,43] to assess domains from the technology acceptance model (TAM). Moreover, we inquired and recorded any symptoms of cybersickness, including nausea, oculomotor discomfort, and disorientation. Most of the questionnaires were administered before the VR test, whereas several questions regarding cybersickness and technological acceptance were administered after the test to find out the feasibility and practicality of this device.

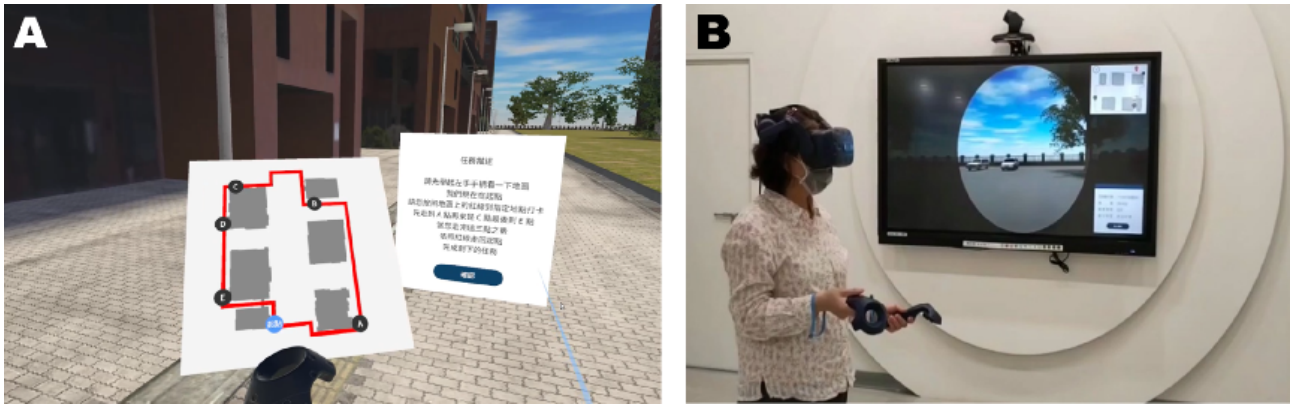
Assessment of Spatial Navigation Abilities

As indicated in our previous research [13], we remodeled the Tzu-Chiang campus of National Cheng Kung University to adapt it to our VR devices due to its ecological validity and for further comparisons. The Pai-Jan virtual reality (PJVR) device set included a VIVE Pro Eye head-mounted display, joysticks, and sensors, which were contract-manufactured by HTC International Corporation, whereas the software was from SteamVR (Valve) and the Unity game engine (Unity Technologies). In the pretest phase, the participants were instructed to approach the designated punch-in points A, C, and E in the first round and D and B in the second round by manipulating joysticks, which allowed them to move forward and backward, punch in, and select the answers to the questions. The selection of punch-in points was based on visual blockage from the buildings in this model (Figure 1). The other scenario details, such as the exterior of the buildings, the texture of the asphalt road, and the addition of trees, cars, and other materials, strived to be the same as those of the real world. Two parameters, LD and VD, were tested in this study. LD was defined as a relative value (in meters) based on the coordinates established in the system determining the degree of difference between the location where the participant punched in and the correct punching point. VD was defined as the angular difference (in degrees) between 2 vectors originating from the starting point: one connecting the starting point to the target checkpoint (the correct location) and the other representing the participant's pointing direction toward the presumed start point. LD was tested using the computer system, which calculated the relative distance between the standing point of the participants in the VR world and the correct punch-in point. We also asked the participants to point at the starting point (Chi-Mei building, the building with white exterior), and the computer determined

XSL•FO
RenderX

the angular differences through the in-built system. The VR tests concluded after 2 rounds, and the remaining questionnaires mentioned previously were then administered. All the examinations were conducted in a classroom with great care for security and comfort to reduce other factors that may affect the results of the tests.

Figure 1. The Pai-Jan virtual reality test. The figure shows (A) a map in the virtual space and the blue light beam for functional manipulation and (B) the testing environment from the examiner's views. The map clearly demonstrated 5 spots, A, C, E, D, and B, and the light blue spot represented the starting point. Participants could read written instructions on a billboard, and at the same time, the recording repeated the words through earphones on the head-mounted display.



Data Analysis and Statistics

All data were properly stored in an ethical manner. Continuous variables are presented as means and SDs, whereas categorical variables are summarized as counts and percentages. To compare demographic characteristics across the 3 groups, we first assessed the normality of the continuous variables. On the basis of the results, one-way ANOVA was applied to variables with a normal distribution, whereas the Kruskal-Wallis test was used for those that did not meet the normality assumption. For categorical variables, the chi-square test was used to examine group differences when expected cell counts were adequate. When the overall group comparison was statistically significant, post hoc pairwise comparisons were conducted using least significant difference to identify specific between-group differences. A correlation was analyzed between the total QuENA score and LD and VD. The main purpose of this study was not to concentrate on over- or underestimation judgments of distance or vector, so the VDs in the statistical analysis were processed in absolute values.

All statistical analyses were conducted using SPSS (version 17.0; IBM Corp), with a 2-sided significance level of α=.05.

Ethical Considerations

This study was approved by the ethical review board of National Cheng Kung University Hospital located in Tainan, Taiwan (B-BR-110-046). All participants gave their written informed consent to take part. The data were deidentified during analytic processes and were stored in an ethical and secure manner to prevent data leakage.

Results

A total of 128 participants provided informed consent. However, of these 128 participants, 15 (11.7%) were not tested using the PJVR device because 10 (67%) changed their minds and 5 (33%) had medical issues. Of the 113 participants who initiated the trials, 20 (17.7%) developed cybersickness and were unable to complete the PJVR test, including 4 (20%) who were CU, 7 (35%) with SCD, and 9 (45%) with MCI. Eventually, 93 participants finished the procedure, including 22 (24%) who were CU, 39 (42%) with SCD, and 32 (34%) with MCI. The demographic data are shown in Table 3.

XSL•FO
RenderX

Table . Demographic data and basic assessments. Least significant difference tests were conducted as the post hoc test.

	CU[a] (n=22)	SCD[b] (n=39)	MCI[c] (n=32)	*F* test (*df*)	*P* value	Post hoc test	95% CI of the mean
Sex (male), n (%)	12 (55)	22 (56)	15 (47)	—[d]	.71	ns[e]	—
Age (y), mean (SD)	67.7 (5.7)	68.5 (4.2)	70.0 (3.7)	1.818 (2, 90)	.17	ns	67.9-69.8
Education (y), mean (SD)	13.3 (1.7)	13.1 (2.5)	12.0 (3.1)	2.355 (2, 90)	.10	ns	12.2-13.3
CDR[f] (n=15/15/32; range 0-3), mean (SD)	0.1 (0.2)	0.3 (0.2)	0.5 (0.0)	23.579 (2, 90)	<.001	MCI>SCD>CU	0.3-0.4
GDS[g] (range 0-15), mean (SD)	0.5 (0.7)	0.5 (1.0)	0.7 (1.6)	0.204 (2, 90)	.82	ns	0.3-0.8
SCD-Q-P[h] (range 0-24), mean (SD)	5.2 (3.2)	9.0 (5.4)	8.2 (5.0)	4.622 (2, 90)	.01	MCI=SCD>CU	6.8-8.9
SCD-Q-CG[i] (range 0-24), mean (SD)	4.8 (4.8)	9.7 (5.0)	14.3 (2.6)	26.742 (2, 90)	<.001	MCI>SCD>CU	9.2-11.9
MMSE[j,k] (range 0-30), mean (SD)	27.5 (1.9)	28.1 (1.6)	22.7 (2.9)	58.953 (2, 90)	<.001	CU=SCD>MCI	25.4-26.8
CASI[l] total score (range 0-100), mean (SD)	91.8 (4.3)	92.5 (3.8)	77.7 (9.5)	53.769 (2, 90)	<.001	CU=SCD>MCI	85.3-89.2
Remote memory	10.0 (0.0)	10.0 (0.3)	9.6 (0.9)	3.584 (2, 90)	.03	CU=SCD>MCI	9.7-9.9
Recent memory	10.4 (1.3)	11.2 (1.0)	6.3 (3.0)	59.903 (2, 90)	<.001	CU=SCD>MCI	8.7-10.0
Attention	7.7 (0.6)	7.7 (0.6)	7.2 (0.9)	4.999 (2, 90)	.009	CU=SCD>MCI	7.4-7.7
Mental manipulation	9.2 (1.3)	9.2 (1.2)	8.4 (2.2)	2.447 (2, 90)	.09	SCD>MCI	8.6-9.3
Orientation	17.8 (0.5)	17.9 (0.3)	13.8 (3.6)	38.021 (2, 90)	<.001	CU=SCD>MCI	15.9-17.1
Abstract thinking	8.6 (1.5)	8.6 (1.5)	7.7 (1.3)	4.911 (2, 90)	.009	CU=SCD>MCI	8.0-8.6
Language	9.7 (0.5)	9.9 (0.4)	9.6 (0.8)	2.675 (2, 90)	.07	SCD>MCI	9.6-9.8
Drawing	9.9 (0.4)	9.6 (0.8)	8.8 (1.5)	7.374 (2, 90)	.001	CU=SCD>MCI	9.2-9.6
Verbal fluency	8.5 (2.0)	8.5 (1.7)	6.3 (2.3)	12.073 (2, 90)	<.001	CU=SCD>MCI	7.3-8.2
QuENA[m] (range 0-30), mean (SD)							
QuENA-P[n] total	1.6 (1.6)	3.6 (3.1)	4.0 (3.4)	4.728 (2, 90)	.01	MCI=SCD>CU	2.7-3.9
LSA[o]	0.1 (0.4)	1.1 (1.5)	1.1 (1.4)	4.432 (2, 90)	.02	MCI=SCD>CU	0.6-1.1
ED[p]	0.2 (0.4)	0.5 (0.9)	0.9 (1.1)	3.54 (2, 90)	.03	MCI>CU	0.4-0.8
INA[q]	1.0 (1.0)	1.6 (1.1)	1.2 (1.3)	1.974 (2, 90)	.15	ns	1.1-1.6
HD[r]	0.2 (0.4)	0.4 (0.8)	0.9 (1.3)	3.668 (2, 90)	.03	MCI>CU=SCD	0.3-0.8
QuENA-CG[s] total	2.3 (2.1)	3.2 (2.8)	5.7 (4.1)	6.492 (2, 90)	.003	MCI>CU=SCD	3.3-5.0
LSA	0.1 (0.4)	0.7 (1.0)	1.4 (1.6)	6.271 (2, 90)	.003	MCI>CU=SCD	0.6-1.2

XSL•FO
RenderX

	CU[a] (n=22)	SCD[b] (n=39)	MCI[c] (n=32)	*F* test (*df*)	*P* value	Post hoc test	95% CI of the mean
ED	0.5 (0.9)	0.7 (1.0)	1.4 (1.3)	3.944 (2, 90)	.02	MCI>CU=SCD	0.7-1.2
INA	1.5 (1.0)	1.2 (1.0)	1.5 (1.6)	0.354 (2, 90)	.70	ns	1.1-1.7
HD	0.3 (0.6)	0.7 (1.2)	1.5 (1.5)	5.247 (2, 90)	.008	MCI>CU=SCD	0.6-1.3

[a]CU: cognitively unimpaired.

[b]SCD: subjective cognitive decline.

[c]MCI: mild cognitive impairment.

[d]Not applicable.

[e]ns: not significant.

[f]CDR: Clinical Dementia Rating Scale.

[g]GDS: Geriatric Depression Scale.

[h]SCD-Q-P: patient version of the Subjective Cognitive Decline Questionnaire.

[i]SCD-Q-CG: caregiver version of the Subjective Cognitive Decline Questionnaire.

[j]MMSE: Mini-Mental State Examination.

[k]Extracted and modified from the Cognitive Abilities Screening Instrument.

[l]CASI: Cognitive Abilities Screening Instrument.

[m]QuENA: Questionnaire on Everyday Navigational Ability.

[n]QuENA-P: patient version of the QuENA.

[o]LSA: landmark and scene agnosia.

[p]ED: egocentric agnosia.

[q]INA: inattention.

[r]HD: heading disorientation.

[s]QuENA-CG: caregiver version of the QuENA.

Table 4 shows the results for the TAM domains and cybersickness. Regarding the TAM, the CU group expressed more enjoyment than the SCD group, whereas the MCI group manifested less perceived ease of use than the other 2 groups, which warrants further investigation. Regarding cybersickness, no difference was detected in any symptoms among the 3 groups. This is an important finding for future application.

XSL•FO
RenderX

Table . Technology acceptance and cybersickness results. Least significant difference tests were conducted for the post hoc test.

Item	CU[a] (n=22), mean (SD)	SCD[b] (n=39), mean (SD)	MCI[c] (n=32), mean (SD)	*F* test (*df*)	*P* value	Post hoc test	95% CI of the mean
PEOU[d] 1	6.0 (1.1)	6.0 (1.1)	5.0 (1.9)	5.723 (2, 90)	.005	CU=SCD>MCI	5.4-6.0
PEOU 2	5.1 (2.0)	4.9 (2.3)	5.0 (1.7)	0.117 (2, 90)	.89	ns[e]	4.6-5.4
PEOU 3	6.0 (0.9)	5.3 (1.3)	4.9 (1.7)	4.468 (2, 90)	.01	CU>MCI	5.0-5.6
PEOU 4	5.9 (1.0)	5.4 (1.5)	5.0 (1.6)	2.228 (2, 90)	.11	ns	5.1-5.7
CANX[f] 1	7.0 (0.2)	6.3 (1.8)	6.1 (2.1)	1.916 (2, 90)	.15	ns	6.0-6.7
CANX 2	6.1 (1.8)	5.2 (2.2)	5.2 (2.2)	1.712 (2, 90)	.19	ns	5.0-5.8
CANX 3	5.5 (2.1)	5.2 (2.2)	5.8 (1.9)	0.891 (2, 90)	.41	ns	5.1-5.9
CANX 4	6.7 (1.3)	6.1 (1.8)	6.1 (1.8)	0.907 (2, 90)	.41	ns	5.9-6.6
ENJ[g] 1	5.3 (1.5)	4.4 (1.7)	4.8 (1.6)	2.164 (2, 90)	.12	ns	4.4-5.1
ENJ 2	5.5 (1.1)	4.6 (1.5)	5.1 (1.4)	3.553 (2, 90)	.03	CU>SCD	4.7-5.3
ENJ 3	5.6 (1.2)	4.7 (1.4)	5.0 (1.6)	3.132 (2, 90)	.048	CU>SCD	4.7-5.3
Nausea	17.8 (26.5)	21.3 (32.2)	7.8 (12.5)	2.540 (2, 90)	.08	ns	10.4-21.2
Oculomotor discomfort	8.3 (11.7)	18.1 (27.0)	9.7 (20.3)	1.914 (2, 90)	.15	ns	8.3-17.4
Disorientation	18.3 (29.3)	33.6 (59.2)	16.5 (30.1)	1.518 (2, 90)	.23	ns	14.9-33.3
Total score	16.0 (23.3)	26.3 (40.7)	12.3 (22.0)	1.868 (2, 90)	.16	ns	12.5-25.6

[a]CU: cognitively unimpaired.

[b]SCD: subjective cognitive decline.

[c]MCI: mild cognitive impairment.

[d]PEOU: perceived ease of use.

[e]ns: not significant.

[f]CANX: Computer Anxiety Scale.

[g]ENJ: perceived enjoyment.

Findings of the Core Test

The MCI group took more time to complete the test than the SCD or CU groups in the counterclockwise A-C-E sequence, but no difference was detected in the clockwise B-D sequence.

The mean LD of the CU, SCD, and MCI groups was 38.2 (SD 39.5), 50.4 (SD 40.7), and 100.4 (SD 46.2) meters, respectively (P<.001; Table 5). LD showed a large group effect ($F_{2, 90}$=17.87; P<.001; ω^2=0.27; partial η^2=0.28; f=0.63). In the same way, the MCI group yielded a greater VD than either the SCD or CU group (P=.02). VD showed a small to medium group effect ($F_{2, 90}$=4.33; P=.02; partial η^2=0.088; ω^2=0.067; Cohen f=0.31).

XSL•FO
RenderX

Table . Time to complete the test, linear deviation, and vector deviation. A, B, C, D, and E indicate the 5 points in the setting. Least significant difference tests were conducted for the post hoc test.

	CU[a] (n=22), mean (SD)	SCD[b] (n=39), mean (SD)	MCI[c] (n=32), mean (SD)	*F* test (*df*)	*P* value	Post hoc test	95% CI of the mean
Time to complete (min)							
A-C-E	23.4 (8.8)	25.5 (7.4)	32.7 (11.7)	7.855 (2, 90)	.001	MCI>CU=SCD	25.4-29.5
B-D	19.3 (4.6)	19.8 (6.1)	22.5 (10.2)	1.524 (2, 90)	.22	ns[d]	19.1-22.2
Linear deviation (m)							
A	31.9 (61.2)	25.4 (45.7)	92.3 (80.1)	11.088 (2, 90)	<.001	MCI>CU=SCD	35.6-64.3
B	29.5 (45.4)	54.7 (66.9)	116.4 (64.8)	14.920 (2, 90)	<.001	MCI>CU=SCD	55.5-84.8
C	66.5 (93.1)	90.2 (94.7)	131.7 (101.6)	3.229 (2, 90)	.04	MCI>CU	78.5-119.3
D	22.2 (24.3)	30.4 (35.5)	63.5 (50.9)	9.050 (2, 90)	<.001	MCI>CU=SCD	31.0-48.7
E	41.3 (55.5)	49.3 (54.4)	98.3 (75.3)	7.314 (2, 90)	.001	MCI>CU=SCD	50.5-78.0
Mean	38.2 (39.5)	50.4 (40.7)	100.4 (46.2)	17.866 (2, 90)	<.001	MCI>CU=SCD	54.5-74.9
Vector deviation (degrees)							
A	26.3 (58.0)	20.1 (42.3)	67.7 (93.8)	4.799 (2, 90)	.01	MCI>CU=SCD	23.5-52.4
B	42.4 (56.5)	42.2 (53.4)	66.9 (83.6)	1.447 (2, 90)	.24	ns	37.1-64.6
C	46.8 (92.1)	37.7 (51.4)	47.8 (50.7)	0.271 (2, 90)	.76	ns	30.4-56.2
D	40.3 (66.9)	34.8 (64.0)	49.9 (54.7)	0.533 (2, 90)	.59	ns	28.6-53.9
E	37.0 (48.7)	61.1 (66.3)	83.4 (81.6)	3.015 (2, 90)	.05	ns	48.6-77.5
Mean	38.6 (37.4)	39.4 (33.0)	63.2 (42.4)	4.333 (2, 90)	.02	MCI>CU=SCD	39.4-55.3

[a]CU: cognitively unimpaired.

[b]SCD: subjective cognitive decline.

[c]MCI: mild cognitive impairment.

[d]ns: not significant.

The LD of the PJVR device was correlated with the total scores on the QuENA-CG (P<.001), indicating good ecological validity. The correlation with the scores on the QuENA-P was not significant (P=.07). As expected, the MCI group yielded a larger LD and VD than the other groups in judging their current location and path integration, the functions of which are highly related to the entorhinal cortex [44].

Discussion

Feasibility

The PJVR device is theory-driven and carefully designed to detect deficits in spatial cognition in people with cerebral AD pathology [13,26]. The fact that 82.3% (93/113) of the participants completed the procedures with satisfactory technology acceptance supports its feasibility for use in older adults and individuals with MCI. Our recent research showed that people at the stage of all-cause MCI were at risk of getting lost or going missing. Hence, a test to detect SNI in this group is important. This study is part of the limited research that has used an immersive VR setting to assess spatial cognition in MCI in Asian countries, and the results may encourage more researchers to develop similar devices in this field.

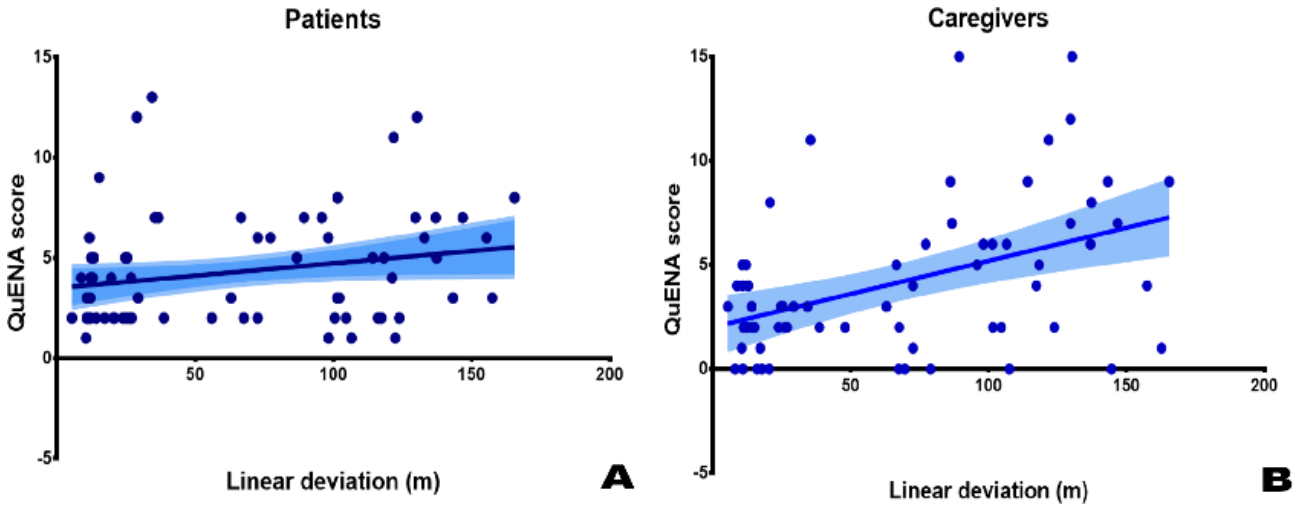
Principal Findings

The CU and SCD groups shared similar features, mostly in demographics, cognitive assessment, and functions of sense of location and path integration. The SCD group was similar to the MCI group in SCD-Q and QuENA-P scores as these 2 questionnaires were self-reported by patients and the SCD group may have downplayed their spatial ability and exaggerated their cognitive impairment out of anxiety.

As expected, the MCI group spent more time completing the test as their spatial navigation ability was worse than that of the other 2 groups (Table 5). The correlation was high between the LD and the QuENA-CG score but not between the LD and the QuENA-P score (Figure 2). The QuENA-CG assesses explicit navigation behavior observed by others and shows good predictive accuracy for risk of getting lost [17,45]. Hence, it reflects the SNI of these patients and is useful for identifying people at risk of getting lost or going missing. However, the MCI group may have overestimated their navigational ability, and the greater the overestimation, the greater the likelihood of getting lost [17,45]. This may explain the poor correlation between the LD and the QuENA-P score. This was also true for the correlation between VD and the total scores on the QuENA-P and QuENA-CG.

XSL•FO
RenderX

Figure 2. The correlation between the Questionnaire on Everyday Navigational Ability (QuENA) total score and linear deviation: (A) total score on the patient version of the QuENA (0.01232 × linear deviation [meters] + 3.493; R^2=0.0483; *P*=.07) and (B) total score on the caregiver version of the QuENA (0.03180 × linear deviation [meters] + 2.020; R^2=0.1968; *P*<.001).



In terms of the comparison between the CU and SCD groups, both showed similar navigational abilities and mental functions, which were supported by the LD and VD results and the Cognitive Abilities Screening Instrument scores in our study. However, in the QuENA-P, the SCD group showed similar patterns of getting lost to those of patients with MCI in several aspects.

The absence of a difference in LD or VD between the CU and SCD groups did not support our hypothesis. One of the reasons is that SNI does not appear before the clinical stage of MCI. However, this was a cross-sectional study rather than a longitudinal one, so it was not possible to recognize the gradual deterioration of navigational abilities in patients with MCI who had been through the CU and SCD periods. Another possibility is that there was a difference in SNI between these 2 groups, but the current assessment was not sensitive enough to detect it. On the other hand, the SCD group showed mismatched results in comparison with patients with MCI in subjective getting lost experience and the standardized PJVR testing environment, which may be due to distinct severity of spatial disorientation.

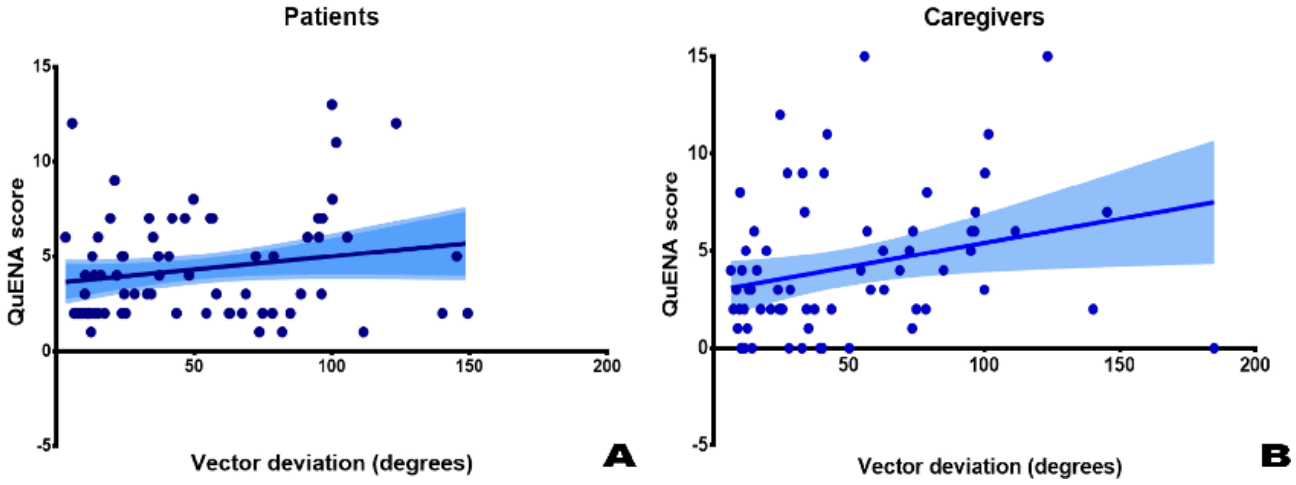
Clinical Implications

Daily navigation in a familiar environment usually takes only a few minutes to half an hour. When traveling from one place to another, events that occur along the way may distract the navigator. Keeping the destination or sequential stops along the journey in mind is a challenge to those with poor spatial orientation and episodic memory. However, traditional neuropsychological assessments are not good at predicting the risk of getting lost or going missing for individuals at the stage of MCI or very mild AD dementia [3].

The QuENA examines navigation deficits in daily life explicitly and is superior to many similar questionnaires [17,46,47]. In this study, the correlations between the QuENA-CG total score and both LD and VD on the PJVR device were significant (Figures 2 and 3; for LD, *P*<.001; for VD, *P*=.03). On the basis of this, for those with poor performance on the PJVR device, the risk of getting lost or going missing is high when timely help from others is unavailable and their problem-solving ability is also impaired. The definition of poor performance on the PJVR device using LD or VD cutoff values needs further study to verify its external validity.

XSL•FO
RenderX

Figure 3. The correlation between the Questionnaire on Everyday Navigational Ability (QuENA) total score and vector deviation: (A) total score on the patient version of the QuENA (0.01391 × vector deviation [degrees] + 3.607; R^2=0.03648; *P*=.11) and (B) total score on the caregiver version of the QuENA (0.02469 × vector deviation [degrees] + 2.943; R^2=0.07384; *P*=.03).



The QuENA-CG assesses the risk of getting lost or going missing [45], and the PJVR device simulates real-world navigation scenarios. Thus, a poor performance on the latter indicates a higher risk of getting lost or going missing. However, different people may adopt different methods to overcome their disorientation and prevent themselves from going missing. Most older adults, for example, use an egocentric instead of an allocentric strategy to compensate for the impaired hippocampal function [47]. In this regard, the QuENA-CG score shows the risk of getting lost or going missing, and the PJVR device reveals evidence of SNI.

From a clinical viewpoint, patients are often aware of their deficits in spatial navigation but are unwilling to report the problem to others because of pride. Another possibility is that they cannot retain these disorientation or getting lost experiences in memory long enough to inform others. Gradually, difficulties in wayfinding increase until getting lost or going missing eventually occurs. Studies have revealed that baseline spatial navigation ability can predict clinical progression in patients in their midlife or with cognitive impairment [20,48-50]. Thus, we recommend a test or a questionnaire, such as the PJVR device or the QuENA, to assess spatial navigation ability in people with cognitive impairment, especially for those with suspected cerebral amyloid pathology. Moreover, many investigators have recommended SNI as a marker for diagnosing early AD [7-10,26]. Modification of the PJVR device to a simpler version is needed because the era of unsupervised digital cognitive assessment is imminent.

Comparison With Other Similar VR Settings

In comparison with other immersive VR settings used to evaluate spatial navigation ability, the PJVR device has several advantages. First, it was remodeled from the real-world Pai-Jan device, and head-to-head comparison is possible. During the PJVR test, participants wear VR headsets and orient themselves by turning their heads, and navigation is typically performed using a monitor and joystick. However, due to space limitations, participants used handheld controllers to move rather than walking physically. As a result, proprioceptive input such as joint position and actual horizontal displacement was not engaged, making the PJVR device relatively less immersive in this respect compared to the original real-world Pai-Jan device or other tasks conducted in smaller physical spaces where participants can walk, such as the triangle completion test [12] or Morris Water Maze [28] or setups such as the Virtual Environment Navigation Lab [29]. Second, as a VR spatial navigation test, the PJVR system features a testing environment modeled after a real university campus, enhancing its ecological validity compared to other VR-based navigation tasks such as the triangle completion test [12], Morris Water Maze [28], loop closure test [30], or Y-maze task [31]. These settings are computer-generated or indoors or involve nature or wild scenarios rather than human environments. As mentioned previously, an outdoor scenario is more akin to real-world environments where people with cognitive impairment would get lost or go missing.

Third, the PJVR device is a theory-driven design based on neuroscience evidence and safer than real-world outdoor tasks. The PJVR device assesses not only distance errors during navigation but also sense of direction (eg, the ability to return to the starting point) and navigational efficiency (measured via completion time). As participants must pass through multiple checkpoints in sequence, navigation errors can accumulate progressively. In other words, the PJVR device captures sequential errors that occur during long-route spatial navigation, which is an aspect that is rarely examined in other VR-based navigation tasks. Given the spatial complexity of the PJVR environment, executive functions may also be required for task completion. To reduce the cognitive demands related to direction selection and route planning, participants follow a predefined path. Notably, the design ensures that checkpoints are obstructed by buildings, preventing direct visual access. This reduces reliance on landmark or beacon cues and visuospatial memory, thereby emphasizing the use of allocentric spatial navigation strategies. Finally, similar to other VR settings, the PJVR device is a standardized, well-designed program, and the researchers could record parameters of interest precisely for analysis. More detailed comparisons can be found in a book to be published [51].

XSL•FO
RenderX

In summary, the PJVR device primarily evaluates allocentric spatial navigation ability, achieves a notable degree of ecological validity, and measures multiple dimensions of spatial navigation performance. Its unique design allows for the detection of more detailed navigational behaviors. With appropriate application, the PJVR device holds potential for advancing new discoveries or the development of novel hypotheses in spatial cognition and dementia research.

Limitations

In this study, participants with MCI were carefully selected based on clinical manifestation, neuropsychological tests, and neuroimages. Regarding the concept of early and late MCI, participants with late MCI might show worse performance in sense of location as assessed using the PJVR device. However, the number of participants in this study was insufficient to divide them into early and late MCI groups. Future studies with more participants and using amyloid positron emission tomography scans or fluid biomarkers to verify cerebral amyloid pathology are needed to answer this question. In addition, due to the innovation of our study, only a few studies could serve as the foundation, which might lead to insufficient explanatory power of the results. Nevertheless, this research provides a protocol for future studies to refer to, and the results presented in this paper could be verified by other researchers.

In conclusion, the PJVR device is feasible for assessing spatial navigation ability in people with MCI and can distinguish them from those who are CU or with SCD. Poor performers in the PJVR test have a higher risk of getting lost or going missing, and their care partners must take action to prevent unfavorable events.

Funding

This research was funded by the National Science and Technology Council of Taiwan (grants NSTC 113-2321-B-006-017, NSTC 112-2321-B-006-015, NSTC 113-2634-F-038-001, NCST 112-2634-F-038-001, NSTC 111-2634-F-038-001, MOST 110-2634-F-038-006, MOST 109-2514-S-006-005, MOST 108-2514-S-006-002, and MOST 107-2514-S-006-005).

Conflicts of Interest

None declared.

References

1. Pai MC, Hsiao S. Incipient symptoms of Alzheimer's disease and effect of education on the onset age: a study of 155 Taiwanese patients. Acta Neurol Taiwan 2002;11(2):66-69 [FREE Full text]
2. Pai MC, Yang YC. Impaired translation of spatial representation in young onset Alzheimer's disease patients. Curr Alzheimer Res 2013 Jan;10(1):95-103. [doi: 10.2174/1567205011310010013] [Medline: 23036016]
3. Pai MC, Jacobs WJ. Topographical disorientation in community-residing patients with Alzheimer's disease. Int J Geriatr Psychiatry 2004 Mar;19(3):250-255. [doi: 10.1002/gps.1081] [Medline: 15027040]
4. Tu MC, Pai MC. Getting lost for the first time in patients with Alzheimer's disease. Int Psychogeriatr 2006 Sep;18(3):567-570. [doi: 10.1017/S1041610206224025] [Medline: 16867206]
5. Tuena C, Riva G, Murru I, et al. Contribution of cognitive and bodily navigation cues to egocentric and allocentric spatial memory in hallucinations due to Parkinson's disease: a case report. Front Behav Neurosci 2022 Oct 13;16:992498. [doi: 10.3389/fnbeh.2022.992498] [Medline: 36311858]
6. Ekstrom AD, Hill PF. Spatial navigation and memory: a review of the similarities and differences relevant to brain models and age. Neuron 2023 Apr 5;111(7):1037-1049. [doi: 10.1016/j.neuron.2023.03.001] [Medline: 37023709]
7. Lithfous S, Dufour A, Després O. Spatial navigation in normal aging and the prodromal stage of Alzheimer's disease: insights from imaging and behavioral studies. Ageing Res Rev 2013 Jan;12(1):201-213. [doi: 10.1016/j.arr.2012.04.007] [Medline: 22771718]
8. Serino S, Riva G. Getting lost in Alzheimer's disease: a break in the mental frame syncing. Med Hypotheses 2013 Apr;80(4):416-421. [doi: 10.1016/j.mehy.2012.12.031] [Medline: 23374425]
9. Hort J, Laczó J, Vyhnálek M, Bojar M, Bures J, Vlcek K. Spatial navigation deficit in amnestic mild cognitive impairment. Proc Natl Acad Sci U S A 2007 Mar 6;104(10):4042-4047. [doi: 10.1073/pnas.0611314104] [Medline: 17360474]
10. Schöberl F, Pradhan C, Irving S, et al. Real-space navigation testing differentiates between amyloid-positive and -negative aMCI. Neurology (ECronicon) 2020 Feb 25;94(8):e861-e873. [doi: 10.1212/WNL.0000000000008758] [Medline: 31896617]
11. Schöberl F, Zwergal A, Brandt T. Testing navigation in real space: contributions to understanding the physiology and pathology of human navigation control. Front Neural Circuits 2020 Mar 6;14:6. [doi: 10.3389/fncir.2020.00006] [Medline: 32210769]
12. Howett D, Castegnaro A, Krzywicka K, et al. Differentiation of mild cognitive impairment using an entorhinal cortex-based test of virtual reality navigation. Brain 2019 Jun 1;142(6):1751-1766. [doi: 10.1093/brain/awz116] [Medline: 31121601]
13. Lai CH, Pai MC. The feasibility and practicality of auxiliary detection of spatial navigation impairment in patients with mild cognitive impairment due to Alzheimer's disease by using virtual reality. Heliyon 2024 Jan 20;10(3):e24748. [doi: 10.1016/j.heliyon.2024.e24748] [Medline: 38317980]

XSL•FO
RenderX

14. Tu S, Wong S, Hodges JR, Irish M, Piguet O, Hornberger M. Lost in spatial translation - a novel tool to objectively assess spatial disorientation in Alzheimer's disease and frontotemporal dementia. Cortex 2015 Jun;67:83-94. [doi: 10.1016/j.cortex.2015.03.016] [Medline: 25913063]
15. Rekers S, Finke C. Translating spatial navigation evaluation from experimental to clinical settings: the virtual environments navigation assessment (VIENNA). Behav Res Methods 2024 Mar;56(3):2033-2048. [doi: 10.3758/s13428-023-02134-0] [Medline: 37166580]
16. Hegarty M, Richardson AE, Montello DR, Lovelace K, Subbiah I. Development of a self-report measure of environmental spatial ability. Intelligence 2002;30(5):425-447. [doi: 10.1016/S0160-2896(02)00116-2]
17. Pai MC, Lee CC, Yang YC, et al. Development of a questionnaire on everyday navigational ability to assess topographical disorientation in Alzheimer's disease. Am J Alzheimers Dis Other Demen 2012 Feb;27(1):65-72. [doi: 10.1177/1533317512436805] [Medline: 22467415]
18. van der Ham IJ, Kant N, Postma A, Visser-Meily JMA. Is navigation ability a problem in mild stroke patients? Insights from self-reported navigation measures. J Rehabil Med 2013 May;45(5):429-433. [doi: 10.2340/16501977-1139] [Medline: 23615778]
19. Tangen GG, Engedal K, Bergland A, Moger TA, Hansson O, Mengshoel AM. Spatial navigation measured by the Floor Maze Test in patients with subjective cognitive impairment, mild cognitive impairment, and mild Alzheimer's disease. Int Psychogeriatr 2015 Aug;27(8):1401-1409. [doi: 10.1017/S1041610215000022] [Medline: 25644091]
20. Jheng SS, Pai MC. Cognitive map in patients with mild Alzheimer's disease: a computer-generated arena study. Behav Brain Res 2009 Jun 8;200(1):42-47. [doi: 10.1016/j.bbr.2008.12.029] [Medline: 19162077]
21. Thomas KG, Hsu M, Laurance HE, Nadel L, Jacobs WJ. Place learning in virtual space. III: investigation of spatial navigation training procedures and their application to fMRI and clinical neuropsychology. Behav Res Methods Instrum Comput 2001 Feb;33(1):21-37. [doi: 10.3758/bf03195344] [Medline: 11296717]
22. Coughlan G, Coutrot A, Khondoker M, Minihane AM, Spiers H, Hornberger M. Toward personalized cognitive diagnostics of at-genetic-risk Alzheimer's disease. Proc Natl Acad Sci U S A 2019 May 7;116(19):9285-9292. [doi: 10.1073/pnas.1901600116] [Medline: 31015296]
23. Rodgers MK, Sindone 3rd JA, Moffat SD. Effects of age on navigation strategy. Neurobiol Aging 2012 Jan;33(1):202.e15-22. [doi: 10.1016/j.neurobiolaging.2010.07.021] [Medline: 20832911]
24. Wiener JM, Carroll D, Moeller S, et al. A novel virtual-reality-based route-learning test suite: assessing the effects of cognitive aging on navigation. Behav Res Methods 2020 Apr;52(2):630-640. [doi: 10.3758/s13428-019-01264-8] [Medline: 31236900]
25. Laczó M, Wiener JM, Kalinova J, et al. Spatial navigation and visuospatial strategies in typical and atypical aging. Brain Sci 2021 Oct 27;11(11):1421. [doi: 10.3390/brainsci11111421] [Medline: 34827423]
26. Pai MC, Jan SS. Have I been here? Sense of location in people with Alzheimer's disease. Front Aging Neurosci 2020 Dec 9;12:582525. [doi: 10.3389/fnagi.2020.582525] [Medline: 33362529]
27. Puthusseryppady V, Morrissey S, Spiers H, Patel M, Hornberger M. Predicting real world spatial disorientation in Alzheimer's disease patients using virtual reality navigation tests. Sci Rep 2022 Aug 4;12(1):13397. [doi: 10.1038/s41598-022-17634-w] [Medline: 35927285]
28. Iggena D, Jeung S, Maier PM, Ploner CJ, Gramann K, Finke C. Multisensory input modulates memory-guided spatial navigation in humans. Commun Biol 2023 Nov 14;6(1):1167. [doi: 10.1038/s42003-023-05522-6] [Medline: 37963986]
29. Kearns MJ, Warren WH, Duchon AP, Tarr MJ. Path integration from optic flow and body senses in a homing task. Perception 2002;31(3):349-374. [doi: 10.1068/p3311] [Medline: 11954696]
30. Chrastil ER, Nicora GL, Huang A. Vision and proprioception make equal contributions to path integration in a novel homing task. Cognition 2019 Nov;192:103998. [doi: 10.1016/j.cognition.2019.06.010] [Medline: 31228680]
31. Bécu M, Sheynikhovich D, Ramanoël S, et al. Landmark-based spatial navigation across the human lifespan. Elife 2023 Mar 13;12:e81318. [doi: 10.7554/eLife.81318] [Medline: 36912888]
32. Adamo DE, Briceño EM, Sindone JA, Alexander NB, Moffat SD. Age differences in virtual environment and real world path integration. Front Aging Neurosci 2012;4:26. [doi: 10.3389/fnagi.2012.00026] [Medline: 23055969]
33. Bažadona D, Fabek I, Babić Leko M, et al. A non-invasive hidden-goal test for spatial orientation deficit detection in subjects with suspected mild cognitive impairment. J Neurosci Methods 2020 Feb 15;332:108547. [doi: 10.1016/j.jneumeth.2019.108547] [Medline: 31830545]
34. Waragai M, Yamada T, Matsuda H. Evaluation of brain perfusion SPECT using an easy Z-score imaging system (eZIS) as an adjunct to early-diagnosis of neurodegenerative diseases. J Neurol Sci 2007 Sep 15;260(1-2):57-64. [doi: 10.1016/j.jns.2007.03.027] [Medline: 17493637]
35. Petersen RC, Caracciolo B, Brayne C, Gauthier S, Jelic V, Fratiglioni L. Mild cognitive impairment: a concept in evolution. J Intern Med 2014 Mar;275(3):214-228. [doi: 10.1111/joim.12190] [Medline: 24605806]
36. Rami L, Mollica MA, García-Sanchez C, et al. The Subjective Cognitive Decline Questionnaire (SCD-Q): a validation study. J Alzheimers Dis 2014;41(2):453-466. [doi: 10.3233/JAD-132027] [Medline: 24625794]

XSL•FO
RenderX

37. Teng EL, Hasegawa K, Homma A, et al. The Cognitive Abilities Screening Instrument (CASI): a practical test for cross-cultural epidemiological studies of dementia. Int Psychogeriatr 1994;6(1):45-58. [doi: 10.1017/s1041610294001602] [Medline: 8054493]
38. Morris JC. Clinical dementia rating: a reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. Int Psychogeriatr 1997;9 Suppl 1:173-176. [doi: 10.1017/s1041610297004870] [Medline: 9447441]
39. Hoyl MT, Alessi CA, Harker JO, et al. Development and testing of a five-item version of the Geriatric Depression Scale. J Am Geriatr Soc 1999 Jul;47(7):873-878. [doi: 10.1111/j.1532-5415.1999.tb03848.x] [Medline: 10404935]
40. Adams DA, Nelson RR, Todd PA. Perceived usefulness, ease of use, and usage of information technology: a replication. MIS Q 1992 Jun;16(2):227-247. [doi: 10.2307/249577]
41. Lester D, Yang B, James S. A short computer anxiety scale. Percept Mot Skills 2005 Jun;100(3 Pt 2):964-968. [doi: 10.2466/pms.100.3c.964-968] [Medline: 16158683]
42. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manage Sci 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
43. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
44. Hafting T, Fyhn M, Molden S, Moser MB, Moser EI. Microstructure of a spatial map in the entorhinal cortex. Nature New Biol 2005 Aug 11;436(7052):801-806. [doi: 10.1038/nature03721] [Medline: 15965463]
45. Pai MC, Lee CC. The incidence and recurrence of getting lost in community-dwelling people with Alzheimer's disease: a two and a half-year follow-up. PLoS ONE 2016 May 16;11(5):e0155480. [doi: 10.1371/journal.pone.0155480] [Medline: 27183297]
46. Laczó M, Svatkova R, Lerch O, et al. Spatial navigation questionnaires as a supportive diagnostic tool in early Alzheimer's disease. iScience 2024 Apr 26;27(6):109832. [doi: 10.1016/j.isci.2024.109832] [Medline: 38779476]
47. Tangen GG, Engedal K, Persson K, et al. Validation of Spatial Orientation Screening questionnaire for use in memory clinic patients. J Alzheimers Dis 2025 Feb;103(4):1013-1022. [doi: 10.1177/13872877241308877] [Medline: 39814544]
48. Levine TF, Allison SL, Stojanovic M, Fagan AM, Morris JC, Head D. Spatial navigation ability predicts progression of dementia symptomatology. Alzheimers Dement 2020 Mar;16(3):491-500. [doi: 10.1002/alz.12031] [Medline: 32043719]
49. Čepukaitytė G, Newton C, Chan D. Early detection of diseases causing dementia using digital navigation and gait measures: a systematic review of evidence. Alzheimers Dement 2024 Apr;20(4):3054-3073. [doi: 10.1002/alz.13716] [Medline: 38425234]
50. Newton C, Pope M, Rua C, et al. Entorhinal-based path integration selectively predicts midlife risk of Alzheimer's disease. Alzheimers Dement 2024 Apr;20(4):2779-2793. [doi: 10.1002/alz.13733] [Medline: 38421123]
51. Pai MC, Yang SH. Using digital tools to detect spatial navigation impairment in early Alzheimer's disease (chapter 10). In: Nguyen TA, editor. Digital Convergence to Provide Better Solutions for People with Dementia: CRC Press; Forthcoming 2025.

Abbreviations

AD: Alzheimer disease
CU: cognitively unimpaired
LD: linear deviation
MCI: mild cognitive impairment
PJVR: Pai-Jan virtual reality
QuENA: Questionnaire on Everyday Navigational Ability
QuENA-CG: caregiver version of the Questionnaire on Everyday Navigational Ability
QuENA-P: patient version of the Questionnaire on Everyday Navigational Ability
SCD: subjective cognitive decline
SCD-Q: Subjective Cognitive Decline Questionnaire
SNI: spatial navigation impairment
TAM: technology acceptance model
VD: vector deviation
VR: virtual reality

XSL•FO
RenderX

Edited by Q Fan; submitted 21.04.25; peer-reviewed by JH Park, S Serino; revised version received 20.10.25; accepted 23.10.25; published 04.12.25.

Please cite as:
Pai MC, Lin YT, Hsiao CY, Lai CH, Chen CJ, Hu CJ, Chen CY
Using Virtual Reality to Assess Spatial Navigation Ability in Individuals With Mild Cognitive Impairment and Older Adults: Cross-Sectional Study
JMIR Aging 2025;8:e75952
URL: https://aging.jmir.org/2025/1/e75952
doi:10.2196/75952

© Ming-Chyi Pai, Yuh-Ting Lin, Chung-Yao Hsiao, Chia-Hung Lai, Chih-Jen Chen, Chaur-Jong Hu, Cheng-Yu Chen. Originally published in JMIR Aging (https://aging.jmir.org), 4.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Critical Moderating Role of Cognitive Function in Digital Inclusion: Data Analysis Study on Depression Risk Among Older Adults

Gang Xiao[1], PhD; Tingting Nie[2], MD

[1]School of Mathematics and Statistics, Hanshan Normal University, Chaozhou, Guangdong province, China

[2]Department of Radiology, Hubei Cancer Hospital, Tongji Medical College, Huazhong University of Science and Technology, No 116 Zhuodaoquan South Load, Hongshan District, Wuhan, Hubei, China

Corresponding Author:
Tingting Nie, MD
Department of Radiology, Hubei Cancer Hospital, Tongji Medical College, Huazhong University of Science and Technology, No 116 Zhuodaoquan South Load, Hongshan District, Wuhan, Hubei, China

Abstract

Background: Digital inclusion has become increasingly important in promoting healthy aging, yet its association with mental health among older adults appears complex and heterogeneous. The role of cognitive function as a moderator and the underlying mechanisms remain understudied.

Objective: This study aims to examine cognitive function's moderating role in the relationship between digital inclusion and depression risk among older adults, and to investigate multiple pathways of association.

Methods: Using data from the 2020 wave of the China Health and Retirement Longitudinal Study, we analyzed 18,673 adults aged 60 years and above (mean age 68.4 y, SD 6.5; 50.8% male participants [n=9486], 49.2% female participants [n=9187]). We constructed interaction effect models to test the moderation hypothesis and employed path analysis with bootstrapped 95% confidence intervals (2000 iterations) to investigate multiple pathways through which digital inclusion is associated with depression.

Results: Cognitive function significantly moderated the digital inclusion-depression relationship (β=−.002, *P*=.03). The association was not statistically significant at low cognitive function (β=−.137, *P*=.33) but strongly protective at high cognitive function (β=−.517, *P*<.001), revealing a "cognitive threshold effect." Path analysis identified 3 significant pathways: direct effects (66.7% of total effect), cognitive enhancement (8.3%), and social participation (8%). Importantly, higher digital inclusion was not found to be associated with increased depression risk at any cognitive function level.

Conclusions: Our findings suggest that older adults require adequate cognitive resources to derive mental health benefits from digital participation, though no harmful effects were observed at lower cognitive levels. This asymmetric pattern has important implications for designing cognitive-informed digital inclusion programs that integrate digital skills training with cognitive enhancement strategies for promoting mental health in aging populations.

(JMIR Aging 2025;8:e80700) doi:10.2196/80700

KEYWORDS

cognitive aging; cognitive function; depression; digital inclusion; moderation effect; multiple pathways; older adults

Introduction

Background

The rapid expansion of digital technologies in everyday life, coupled with global population aging, has increased focus on how digital participation is associated with older adults' well-being. Digital inclusion—conceptualized as access to, engagement with, and effective utilization of digital technologies [1]—has emerged as a critical dimension of social inclusion for older adults. In China, home to the world's largest aging population and a rapidly evolving digital landscape, the intersection of digital technology and aging presents unique research opportunities. According to recent surveys, only 43.2% of Chinese adults aged 60 years and above use the internet, significantly below the national average of 74.4% [2]. Within this population exists considerable heterogeneity in digital participation, with distinct patterns influenced by factors including age, education, urbanicity, and, notably, cognitive capabilities.

Research on digital inclusion's mental health associations has yielded inconsistent findings. Some studies report positive associations between digital technology use and psychological well-being among older adults, suggesting reduced depression risk through increased social connectedness, access to health

XSL•FO
RenderX

information, and enhanced autonomy [3,4]. However, other research identifies potential negative effects, including technology-related stress, information overload, and displacement of face-to-face interactions [5,6]. These contradictory findings suggest complex relationships likely influenced by moderating factors and operating through multiple pathways.

Cognitive function represents a particularly important yet understudied moderating factor in this relationship. Cognitive aging, characterized by normative age-related changes in cognitive abilities, profoundly affects how older adults interact with technology. Processing speed declines, working memory capacity diminishes, and selective attention narrows—all critical resources for technology interaction [7]. Existing research on technology adoption among older adults identifies cognitive abilities as significant predictors of digital technology uptake [8], but few studies have explicitly examined cognitive function's moderating role in the relationship between digital inclusion and depression risk.

Digital Inclusion and Mental Health Among Older Adults

Digital inclusion has been conceptualized in multiple ways across the literature, evolving from narrow definitions focused on physical access to digital technologies toward more comprehensive frameworks encompassing skills, usage patterns, and perceived benefits [9,10]. Current conceptualizations recognize digital inclusion's multidimensionality, encompassing access, skills, motivation, and effective usage in daily life [1]. Among older adults, digital inclusion presents unique challenges and opportunities, with substantial heterogeneity in adoption patterns and usage behaviors [11,12].

Research examining digital inclusion's mental health associations among older adults has yielded mixed findings. Several studies report positive mental health associations, with Chen and Schulz's [4] systematic review identifying social connectedness, informational support, and entertainment as key mechanisms through which information and communication technology interventions reduce social isolation. Cotten et al's [3] longitudinal analysis found that internet use was associated with reduced depression risk among retired older adults, with effects partially mediated by reduced loneliness and increased social connection. Similarly, Chopik [13] reported that social technology use was associated with better self-rated health, fewer chronic conditions, and reduced depression, with benefits mediated through decreased loneliness.

However, other research highlights potential negative effects. Nimrod [5] introduced the concept of "technostress" among older adults, demonstrating that technological complexity, rapid change, and information overload created anxiety and frustration for many seniors. Reinecke et al [6] found that communication load and internet multitasking were associated with increased stress and reduced psychological well-being across the lifespan, with some evidence suggesting heightened vulnerability among older adults. These contradictory findings suggest a more complex relationship than initially conceptualized, likely influenced by individual differences, social context, and specific usage patterns.

Cognitive Aging, Cognitive Reserve, and Technology Engagement

Cognitive aging theory provides important insights into how age-related cognitive changes might influence technology interaction patterns and outcomes. Salthouse's [7] processing-speed theory highlights how slowed cognitive processing in aging is associated with complex task performance, potentially increasing cognitive demands of technology use. Similarly, Craik and Byrd's [14] resource reduction hypothesis suggests that reduced attentional and working memory resources with age might be related to the ability to navigate complex digital environments. These cognitive aging processes create potential barriers to effective technology use among older adults, as documented in Czaja et al's [8] Center for Research and Education on Aging and Technology Enhancement (CREATE) study, which identified cognitive abilities—particularly processing speed and working memory—as significant predictors of technology adoption and successful use.

Complementing cognitive aging perspectives, the cognitive reserve hypothesis provides a theoretical framework for understanding potential bidirectional relationships between cognitive function and digital engagement. Stern [15] proposed that engaging in cognitively stimulating activities builds neural resources that protect against age-related cognitive decline and pathology. Several studies suggest that digital activities may contribute to cognitive reserve, with Tun and Lachman [16] finding that computer use was associated with better episodic memory and executive function among older adults, particularly for those with lower education levels. Similarly, Small et al [17] documented altered neural activation patterns following internet search training among older adults, suggesting that digital engagement might directly stimulate neural plasticity.

From a cognitive load perspective, technology interaction imposes varying cognitive demands depending on design features, task complexity, and user experience [18]. Charness and Boot [19] highlighted how age-related changes in fluid intelligence, working memory, and attention regulation create challenges for older adults navigating digital interfaces, potentially increasing cognitive load during technology interaction. This suggests that older adults with varying cognitive capacities might experience digital engagement differently—as manageable cognitive challenge for those with strong cognitive resources, but as overwhelming cognitive demand for those with limited resources. However, few studies have explicitly examined cognitive function as a moderator of digital technology's psychological effects, creating a significant gap in understanding differential associations across cognitive capacity levels.

Multiple Pathways Linking Digital Inclusion to Mental Health

Theoretical frameworks and empirical evidence suggest multiple potential pathways through which digital inclusion might be associated with mental health outcomes among older adults. Francis et al [20] proposed a multidimensional framework emphasizing how digital technologies reshape autonomy, competence, and connectedness in later life, highlighting both

XSL•FO
RenderX

direct psychological effects and mediated pathways through cognitive and social domains. This framework aligns with emerging research suggesting at least 3 distinct but potentially interconnected pathways.

First, direct psychological pathways involve immediate associations of digital engagement on psychological well-being, independent of cognitive or social mediators. These include enhanced sense of mastery and environmental control [21], reduced perceived vulnerability through access to health information [22], and enjoyable leisure experiences [23]. Experimental studies demonstrate that successful technology interaction generates positive affect and enhances perceived self-efficacy among older adults, potentially counteracting depression-linked feelings of helplessness and loss of control [24].

Second, cognitive pathways involve digital engagement's associations with cognitive function, which in turn is related to depression risk. Longitudinal research documents bidirectional relationships between cognitive function and depression in later life, with cognitive decline predicting increased depression risk and depressive symptoms accelerating cognitive deterioration [25]. If digital engagement is associated with enhanced cognitive function through stimulation and learning, as suggested by cognitive reserve research [16], this might create indirect protective effects against depression. Slegers et al [26] found that computer training and usage were associated with improved aspects of cognitive function among older adults, particularly executive control and memory, domains also implicated in depression pathophysiology.

Third, social pathways involve digital technologies' facilitation of social connection and participation, which protect against depression through enhanced social integration. Santini et al's [27] systematic review documented strong protective effects of social relationships against depression, with both structural (network size, frequency of contact) and functional (perceived support, relationship quality) aspects contributing to protection. Digital communication tools may help maintain existing relationships despite mobility limitations or geographical separation, while online communities might provide new social participation opportunities. Several studies document how social media and communication technologies reduce loneliness and enhance social connections among older adults, potentially mediating technology's mental health benefits [13].

While these pathways have theoretical and empirical support, few studies have simultaneously tested multiple mechanisms or examined their relative contributions to overall effects. Additionally, potential interactions between pathways—such as cognitive function enabling effective social digital use—remain largely unexplored. Comprehensive examination of multiple pathways would advance understanding of digital inclusion's complex associations with mental health in later life, potentially informing more targeted interventions addressing specific mechanisms.

Theoretical Framework and Hypotheses

Integrating insights from cognitive aging theory, cognitive reserve hypothesis, and social integration theory, we propose a comprehensive theoretical framework examining cognitive function's moderating role in digital inclusion's associations with depression risk, and multiple pathways of relationships. This framework acknowledges bidirectional relationships between digital engagement and cognitive function while examining how cognitive capacities moderate digital inclusion's psychological associations through multiple potential mechanisms.

Cognitive aging theory [7,14] suggests that age-related changes in processing speed, working memory, and attentional control are associated with technology interaction, potentially increasing cognitive demands of digital participation. The cognitive reserve hypothesis [15,28] proposes that engaging in cognitively stimulating activities builds neural resources that protect against pathological processes, suggesting that digital engagement may both require and enhance cognitive resources. Social integration theory [29] emphasizes how social relationships are associated with health through multiple pathways, including social support, access to resources, and social influence, with digital technologies potentially reshaping social connection patterns in later life.

Drawing on these theoretical perspectives and previous research, we propose the following hypotheses:

H1: Cognitive function significantly moderates the relationship between digital inclusion and depression risk, with stronger protective associations among older adults with higher cognitive function.

H2: Digital inclusion is associated with depression risk through multiple significant pathways:

H2a: A direct pathway linking digital inclusion directly to depression risk;

H2b: A cognitive enhancement pathway, where digital inclusion is positively associated with cognitive function, which in turn is related to reduced depression risk;

H2c: A social integration pathway, where digital inclusion is associated with enhanced social participation, which in turn is related to reduced depression risk.

H3: Cognitive function and social participation operate sequentially, with digital inclusion related to cognitive function, which is associated with social participation, which in turn is related to depression risk.

We test these hypotheses using data from the 2020 wave of the China Health and Retirement Longitudinal Study (CHARLS), a nationally representative survey of Chinese adults aged 45 years and above, focusing on respondents aged 60 years and older (N=18,673). We employ interaction effect models to examine the moderation hypothesis and path analysis to investigate multiple pathways, providing a comprehensive examination of how digital inclusion is associated with depression risk among older adults with varying cognitive capabilities.

This study makes several important contributions to the literature. First, it systematically examines cognitive function's moderating role in the digital inclusion-depression relationship,

XSL•FO
RenderX

addressing a significant gap in understanding differential associations across cognitive capacity levels. Second, it employs path analysis to simultaneously test multiple mechanisms, providing a more nuanced understanding of how digital participation is associated with mental health outcomes. Third, it develops and tests an integrated theoretical framework combining cognitive aging, cognitive reserve, and social integration perspectives. Finally, it offers practical implications for designing digital inclusion programs and mental health interventions that account for cognitive heterogeneity among older adults. These contributions advance the understanding of digital technology's role in promoting healthy aging in increasingly digitalized societies.

Methods

Data Source and Sample

This study utilized publicly available data from the 2020 wave of the CHARLS, which does not involve direct interaction with participants and ensures anonymity. The CHARLS is a nationally representative survey of adults aged 45 years and above in China, employing a multistage stratified probability-proportional-to-size sampling method, covering 28 provinces and 150 counties or districts. The survey includes comprehensive modules on demographic characteristics, socioeconomic status, health conditions, cognitive assessments, and digital device usage, providing an ideal data foundation for this research.

We applied the following sample selection criteria: (1) adults aged 60 years and above, (2) exclusion of proxy-answered questionnaires to ensure the validity of cognitive assessments and self-reported data, and (3) elimination of cases with missing data on core variables (digital inclusion, depression scores, cognitive function, social participation). After screening, the final effective sample consisted of 18,673 participants, with male participants accounting for 50.8% (n=9481) and female participants accounting for 49.2% (n=9192) . The average age was 68.4 years (SD 6.5 y), ranging from 60 to 108 years.

Ethical Considerations

As the data were anonymized and publicly available, no informed consent or ethics committee approval was required for this study. The authors declare that all procedures followed were in accordance with relevant ethical guidelines, ensuring participants' privacy and confidentiality. The study was exempted from requiring ethics approval by the affiliated institution's ethical review board. No compensation was provided to the participants, as the data were part of a publicly available survey.

Measures

Digital Inclusion

Based on multiple items regarding digital technology use in the CHARLS survey, we constructed a multidimensional digital inclusion index. The index included the following four dimensions:

1. Digital device access (0 - 3 points): Measuring basic digital technology access, including internet use (da040) and device usage (mobile phones or computers, da041)
2. Breadth of digital functionality use (0 - 3 points): Measuring diversity of digital technology features used, including mobile payment, social media use, information gathering, and so forth
3. Depth of digital use (0 - 3 points): Measuring complexity of digital technology usage, including frequency and proficiency of advanced functions in social, financial, and entertainment domains
4. Subjective perception dimension (0 - 1 point, normalized): Based on digital confidence ratings, standardized to 0 - 1 point

A weighted combination method was used to construct a comprehensive digital inclusion index, represented as:

$$DIi=0.3\times Accessi+0.3\times Breathi+0.2\times Depthi+0.2\times perceptioni$$

where DI_i represents the comprehensive digital inclusion index for individual i, $Access_i$ represents the device access score, $Breadth_i$ represents the functionality use breadth score, $Depth_i$ represents the use depth score, and $Perception_i$ represents the subjective perception score. Weight allocation was derived from both theoretical frameworks and empirical validation:

1. Theoretical basis: van Dijk's [9] digital divide framework identifies access as the foundational layer (weight=0.3); Kim and Kim's [30] research on older adults shows functional breadth has stronger associations with well-being than depth (breadth=0.3, depth=0.2).
2. Empirical validation: We conducted principal component analysis (PCA) to examine the variance explained by each dimension. The eigenvalue contributions were as follows: Access=35%, Breadth=28%, Depth=20%, Perception=17%, which informed our final weight allocation (0.3, 0.3, 0.2, 0.2).
3. Sensitivity analysis: Alternative weighting schemes (equal weights, PCA-derived weights) yielded substantively similar results (correlation with final index r>0.95), confirming robustness.

The final result was a 0 - 10 point comprehensive digital inclusion index, with higher scores indicating higher levels of digital inclusion.

To verify the construct validity of the measurement, we conducted PCA, which showed that all 4 dimensions loaded onto a single principal component, explaining 68.7% of the total variance, with a Kaiser-Meyer-Olkin value of 0.78, indicating that the data were suitable for factor analysis. The internal consistency of the comprehensive index was good (Cronbach α=0.81).

Cognitive Function

Based on the cognitive test questions in the CHARLS, we constructed a comprehensive cognitive function index. The CHARLS cognitive assessment includes the following aspects:

1. Immediate memory ability: Assessed through word recall tests, including immediate word recall (participants recall

XSL•FO
RenderX

10 words immediately after hearing them) and delayed word recall (recalling again several minutes later), with 1 point for each correct recall
2. Temporal or spatial orientation: Assessed through recognition of current year, month, date, day of the week, and location, with 1 point for each correct answer
3. Executive function: Assessed through mathematical calculations (eg, serial subtraction of 7 from 100) and time management tasks, with 1 point for correct completion
4. Visuospatial ability: Assessed through drawing tasks, with 1 point for correct completion

We first standardized each cognitive test item, then used PCA to extract common factors and construct a comprehensive cognitive function index:

$$CF_i = w_1 Z_{\text{Immediate Recall}} + w_2 Z_{\text{Delayed Recall}} + w_3 Z_{\text{Orientation}} + w_4 Z_{\text{Executive}} + w_5 Z_{\text{Visuospatial}}$$

where CF_i represents the comprehensive cognitive function score for individual i, Z represents standardized scores, and w represents the weights of each item on the first principal component. The scores were then linearly transformed to a 0‐1000 point scale, with higher scores indicating better cognitive function. All cognitive test items showed good internal consistency (Cronbach α=0.79), and PCA analysis showed that the first principal component explained 57.2% of the total variance, with a Kaiser-Meyer-Olkin value of 0.82, confirming the rationality of using a single comprehensive index.

Social Participation

Social participation was constructed based on respondents' frequency of participation in various social activities, including (1) mutual aid groups or community activities; (2) leisure and cultural entertainment activities; (3) sports and exercise; (4) club activities; (5) educational or training courses; (6) volunteer service activities; and (7) caring for relatives, friends, or neighbors. Each activity participation frequency was rated (0=never, 1=not often, 2=often) and calculated as follows:

$$SP_i = \sum_{j=i}^{7} \text{Activity}_{ij} / (7 \times 2)$$

where SP_i represents the social participation score for individual i, $Activity_{ij}$ represents the participation frequency score of individual i in activity j, and the denominator standardizes the score to the 0‐1 interval. Higher scores indicate higher levels of social participation. The measurement showed acceptable internal consistency (Cronbach α=0.76).

Depression Risk

Depression risk was measured using the 10-item depression symptom scale (CES-D-10 simplified version) in the CHARLS. This scale assesses the frequency of depressive symptoms over the past week, including feeling depressed, finding things effortful, sleeping poorly, and 7 other items. Each item is scored based on frequency (0=rarely or never, 1=sometimes, 2=often, 3=most or all of the time), with a total score range of 0‐30 points. Higher scores indicate greater depression risk. The scale showed good internal consistency (Cronbach α=0.83).

Due to the severely right-skewed distribution of depression scores, we applied Winsorized processing to the original depression scores (truncated at the 1% and 99% percentiles) to balance data structure preservation and outlier impact control. As a sensitivity analysis, we also used alternative methods such as logarithmic transformation [log(depression score +1)], Box-Cox transformation (λ=−0.12), and quantile transformation to normal distribution.

Control Variables

To reduce the influence of confounding factors, we included the following control variables:

1. Demographic variables: Gender (1=male participants, 2=female participants), age (continuous variable), marital status (1=married or cohabiting, 2=separated or divorced, 3=widowed, 4=never married), education level (1=no education, 2=primary school, 3=middle school, 4=high school or secondary school, 5=college and above), and household registration type (1=agricultural, 2=non-agricultural)
2. Socioeconomic factors: Logarithm of household income (calculated based on annual total income), retirement status (1=retired, 0=not retired), and subjective economic status (1=very poor, 5=very good)
3. Health status: Self-rated health status (1=very poor, 5=very good), number of chronic diseases (calculated based on self-reported status of 14 common chronic diseases), activities of daily living, and instrumental activities of daily living limitation scores

Analytical Strategy

Analysis of Cognitive Function Moderation Effect

To test the moderating effect of cognitive function on the relationship between digital inclusion and depression risk, we constructed interaction effect models. The basic model is as follows:

$$\text{Depression}_i = \beta_0 + \beta_1 \text{DigitalInclusion}_i + \beta_2 \text{Cognition}_i + \beta_3 (\text{DigitalInclusion}_i \times \text{Cognition}_i) + \beta_4 \text{Controls}_i + \varepsilon_i$$

where Depression_i represents the depression risk score for individual i, $\text{DigitalInclusion}_i$ represents the digital inclusion score, Cognition_i represents the cognitive function score, $(\text{DigitalInclusion}_i \times \text{Cognition}_i)$ is the interaction term between digital inclusion and cognitive function, Controls_i includes control variables, and ε_i is the random error term.

We adopted a stepwise model building strategy: Model 1 included only main effects (digital inclusion and cognitive function), Model 2 added the interaction term, and Model 3 further controlled for demographic and socioeconomic variables. The significance and direction of the interaction term coefficient β_3 reflect the moderating effect of cognitive function.

To intuitively interpret the interaction effect, we calculated simple slopes of digital inclusion on depression at different levels of cognitive function (mean ±1 standard deviation):

$$\text{SimpleSlope}_{\text{Cognition}_k} = \beta_1 + \beta_3 \times k$$

where k represents a specific cognitive function level. Additionally, we plotted 3D interaction graphs to visually demonstrate the complex relationship between digital inclusion, cognitive function, and depression risk.

XSL•FO
RenderX

Multiple Pathway Analysis

To explore the multiple pathways through which digital inclusion is associated with depression risk, we employed structural equation modeling for path analysis. The basic path model included the following system of equations:

$$Cognition_i = \alpha_0 + \alpha_1 DigitalInclusion_i + \alpha_2 Controls_i + \varepsilon_{1i}$$

$$SocialParticipation_i = \gamma_0 + \gamma_1 DigitalInclusion_i + \gamma_2 Cognition_i + \gamma_3 Controls_i + \varepsilon_{2i}$$

$$Depression_i = \delta_0 + \delta_1 DigitalInclusion_i + \delta_2 Cognition_i + \delta_3 SocialParticipation_i + \delta_4 Controls_i + \varepsilon_{3i}$$

This system of equations allowed us to estimate: direct effects (δ_1), indirect effects through cognitive function ($\alpha_1 \times \delta_2$), indirect effects through social participation ($\gamma_1 \times \delta_3$), and sequential indirect effects through cognitive function to social participation ($\alpha_1 \times \gamma_2 \times \delta_3$).

We used the Bootstrap method (2000 repetitions) to construct 95% confidence intervals to assess the statistical significance of indirect effects and calculated the proportion of each pathway in the total effect:

$$Proportion_{path_k} = \frac{Effect_{path_k}}{Effect_{total}} \times 100\%$$

where $Effect_{path_k}$ represents the effect size of pathway k and $Effect_{total}$ represents the total effect.

Sensitivity Analysis

To assess the robustness of results, we conducted multiple aspects of sensitivity analysis: (1) separate analysis of the 4 constituent dimensions of digital inclusion to test moderation effects of different dimensions; (2) stratified analysis by gender, age group, and education level to explore potential influences of demographic characteristics; and (3) use of alternative statistical models, including tobit regression (handling censoring of depression scores), quantile regression (examining effects at different depression levels), and robust regression (controlling potential outlier influences).

Additionally, we checked for multicollinearity (calculating variance inflation factors, all variables variance inflation factor<2.5), residual heteroscedasticity (through Breusch-Pagan test), and model goodness of fit (χ^2 test, comparative fit index [CFI], Tucker-Lewis index [TLI], root mean square error of approximation [RMSEA], and standardized root mean square residual [SRMR]). For path analysis models, we adopted common fit criteria: CFI and TLI>0.95, RMSEA<0.06, SRMR<0.08 indicating good fit [31].

Data cleaning and analysis were conducted using R (version 4.1.0; developed by Yves Rosseel), primarily relying on the lavaan package for path analysis and structural equation modeling. All continuous variables were centered before constructing interaction terms to reduce multicollinearity issues. All statistical tests were 2-tailed, with significance level set at .05. The study was approved by the ethics committee of the affiliated institution, and all analyses used anonymized data.

Results

Descriptive Statistics and Correlation Analysis

Table 1 presents descriptive statistics and correlation coefficients for key variables in the study. The mean comprehensive digital inclusion index score was 2.35 (SD 2.54), reflecting the overall low level of digital participation among Chinese older adults. The raw cognitive function score averaged 18.72 (SD 7.83, range 3 - 31 on a 0 - 31 scale), which after PCA transformation yielded a mean of 562.35 (SD 235.78, range 85 - 980 on a 0 - 1000 scale). The mean social participation score was 0.35 (SD 0.24), and the Winsorized depression score was 7.46 (SD 5.94, range 0 - 27), indicating mild depressive symptomatology on average. Notably, the raw depression score before Winsorization was 7.82 (SD 6.15), with extreme values truncated at the first and 99th percentiles for all subsequent analyses.

XSL•FO
RenderX

Table . Descriptive statistics and correlation matrix (n=18,673)[a].

Variable	Mean (SD)	Range	1	2	3	4
1. Digital inclusion (0 - 10)	2.35 (2.54)	0 - 10	1			
2. Cognitive function						
2a. Raw score (0 - 31)	18.72 (7.83)	3 - 31	0.42[b]	1		
2b. Transformed score (0 - 1000)[c]	562.35 (235.78)	85 - 980	—[d]	—		
3. Social participation (0 - 1)	0.35 (0.24)	0 - 1	0.29[b]	0.31[b]	1	
4. Depression (CES-D-10)						
4a. Raw score (0 - 30)[e]	7.82 (6.15)	0 - 30	–0.26[b]	–0.34[b]	-0.21[b]	1
4b. Winsorized score (0 - 30)[f]	7.46 (5.94)	0 - 27	–0.27[b]	–0.35[b]	-0.22[b]	—

[a]Values in parentheses in variable names indicate theoretical score ranges. Raw cognitive function score represents the sum of all cognitive test items: immediate recall (0-10)+delayed recall (0-10)+orientation (0-5)+executive function (0-5)+visuospatial ability (0-1).

[b]*P*<.001.

[c]Transformed cognitive function score was derived through principal component analysis and linearly transformed to a 0-1000 scale. This transformed score was used in all subsequent analyses.

[d]—: missing or unavailable data.

[e]Raw CES-D-10 score represents the unprocessed sum of 10 depression items before any transformation.

[f]Winsorized CES-D-10 score was truncated at the first and 99th percentiles (values of 0 and 27, respectively) to reduce the influence of extreme values. This winsorized score was used as the dependent variable in all regression and path analyses.

Correlation analysis showed significant associations between all core variables. Digital inclusion was moderately positively correlated with cognitive function (r=0.42, P<.001), low to moderately positively correlated with social participation (r=0.29, P<.001), and low to moderately negatively correlated with depression (r=−0.27, P<.001). The negative correlation between cognitive function and depression (r=−0.35, P<.001) was stronger than the negative correlation between social participation and depression (r=−0.22, P<.001). These preliminary association patterns support our hypotheses about potential relationships between variables but require more complex multivariate analysis to control for potential confounding factors and test interaction effects.

Moderating Role of Cognitive Function

To test the moderating role of cognitive function in the relationship between digital inclusion and depression risk, we constructed 3 nested regression models, with results shown in Table 2. Model 1 included only the main effects of digital inclusion and cognitive function, showing that both were significantly negatively associated with depression risk ($\beta_{\text{digital inclusion}}$=−0.352, P=.008; $\beta_{\text{cognitive function}}$=−0.005, P<.001). Model 2 added the interaction term between the two, with results showing that the interaction term coefficient was significantly negative (β=−.002, P=.03), indicating that cognitive function significantly moderates the relationship between digital inclusion and depression. Model 3 further controlled for demographic and socioeconomic factors, with the interaction effect remaining significant (β=−.002, P=.03), confirming the robustness of the moderation effect.

XSL•FO
RenderX

Table . Regression results of cognitive function moderation effect[a].

Variable	Model 1 (main effects)	Model 2 (with interaction)	Model 3 (full model)
Digital inclusion	−0.352[b] (0.121)	−0.352[b] (0.129)	−0.327[b] (0.124)
Cognitive function	−0.005[c] (0.001)	−0.005[c] (0.001)	−0.004[c] (0.001)
Digital inclusion×cognitive function	—[d]	−0.002[e] (0.001)	−0.002[e] (0.001)
Control variables	No	No	Yes
R^2	0.067	0.071	0.085
F[f] test	127.45[c]	92.18[c]	37.56[c]
ΔR^2	—	0.004[e]	0.014[c]

[a]Dependent variable is Winsorized depression score; robust standard errors in parentheses. Control variables include age, gender, marital status, education level, residence, log income, health status, number of chronic diseases, and ADL/IADL scores.

[b]P<.01.

[c]P<.001.

[d]—: missing or unavailable data

[e]P<.05.

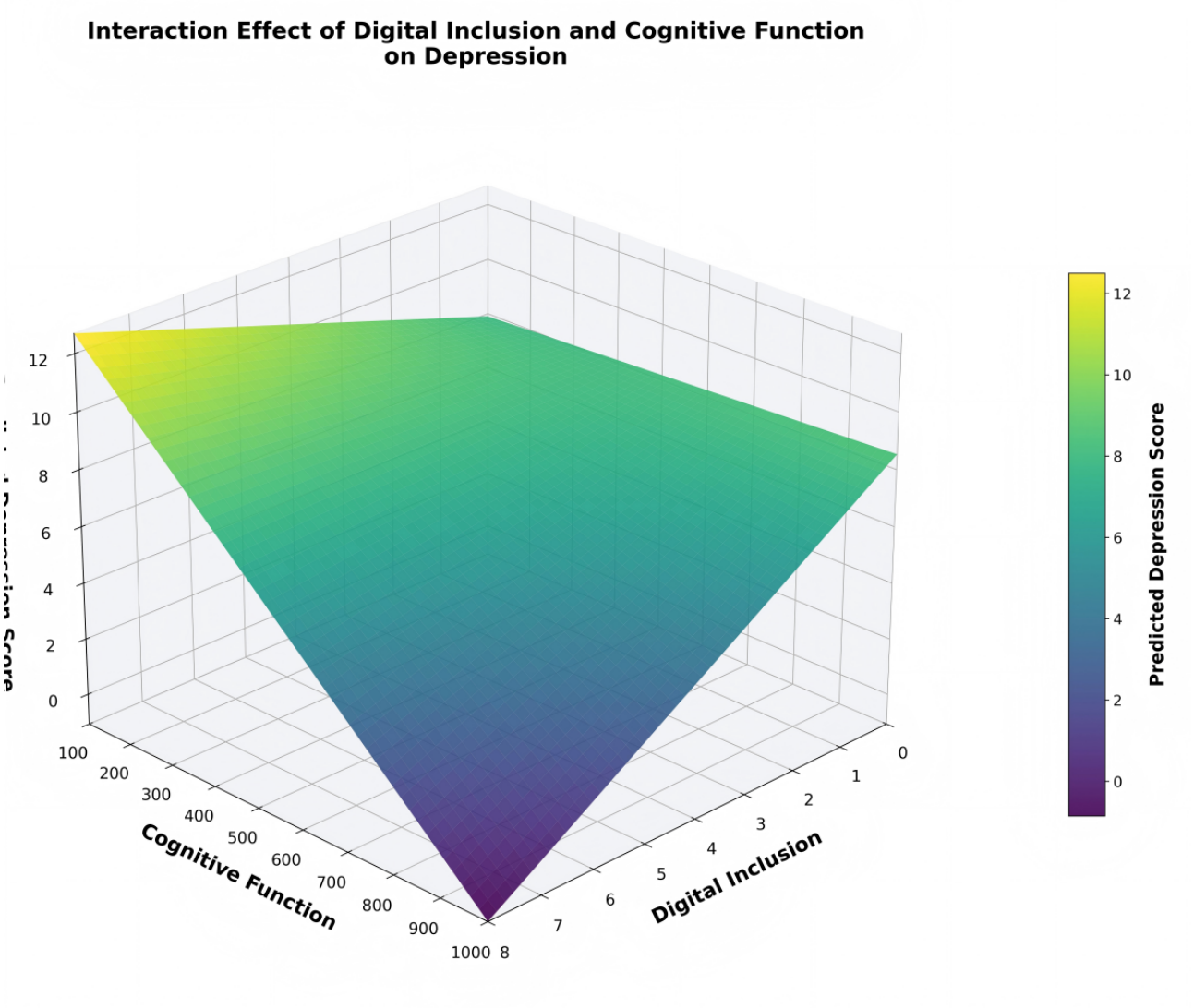
[f]Overall model significance: Model 1, $F(2, 18670)=127.45$, P<.001; Model 2, $F(3, 18669)=92.18$, P<.001; Model 3, $F(18, 18654)=37.56$, P<.001.

To more intuitively understand the moderation effect, we calculated simple slopes at different cognitive function levels. At low cognitive function level (mean –1 standard deviation, ie, 326.57 points), the association between digital inclusion and depression was not significant (β=−.137, P=.33); at average cognitive function level (562.35 points), digital inclusion was significantly negatively associated with depression (β=−.327, P=.008); at high cognitive function level (mean +1 standard deviation, ie, 798.13 points), the association was stronger (β=−.517, P<.001). This pattern indicates that the protective association of digital inclusion against depression strengthens as cognitive function improves.

Figure 1's 3D interaction graph visually demonstrates this moderation effect. In the figure, colors from dark to light represent depression scores from low to high, showing that when cognitive function is higher (right side of the figure), the relationship between increased digital inclusion and reduced depression risk is more significant. This pattern demonstrates a "cognitive threshold effect," suggesting that older adults need to reach a certain cognitive level to derive notable psychological benefits from digital participation.

XSL•FO
RenderX

Figure 1. Moderating effect of cognitive function on the relationship between digital inclusion and depression. The 3D surface shows predicted depression risk (CES-D-10 scores, 0 - 30 scale) as a function of digital inclusion (0 - 10 scale) and cognitive function (0 - 1000 scale). Colors from dark purple (lower depression) to yellow (higher depression) represent the gradient of predicted scores. The surface demonstrates a "cognitive threshold effect," characterized by a relatively flat relationship between digital inclusion and depression at lower cognitive function levels (left side), transitioning to a steeper protective association at higher cognitive function levels (right side). This pattern indicates that older adults require adequate cognitive resources to derive mental health benefits from digital participation. Predictions are based on the interaction model presented in Table 2 (Model 3), with all covariates held at sample means.

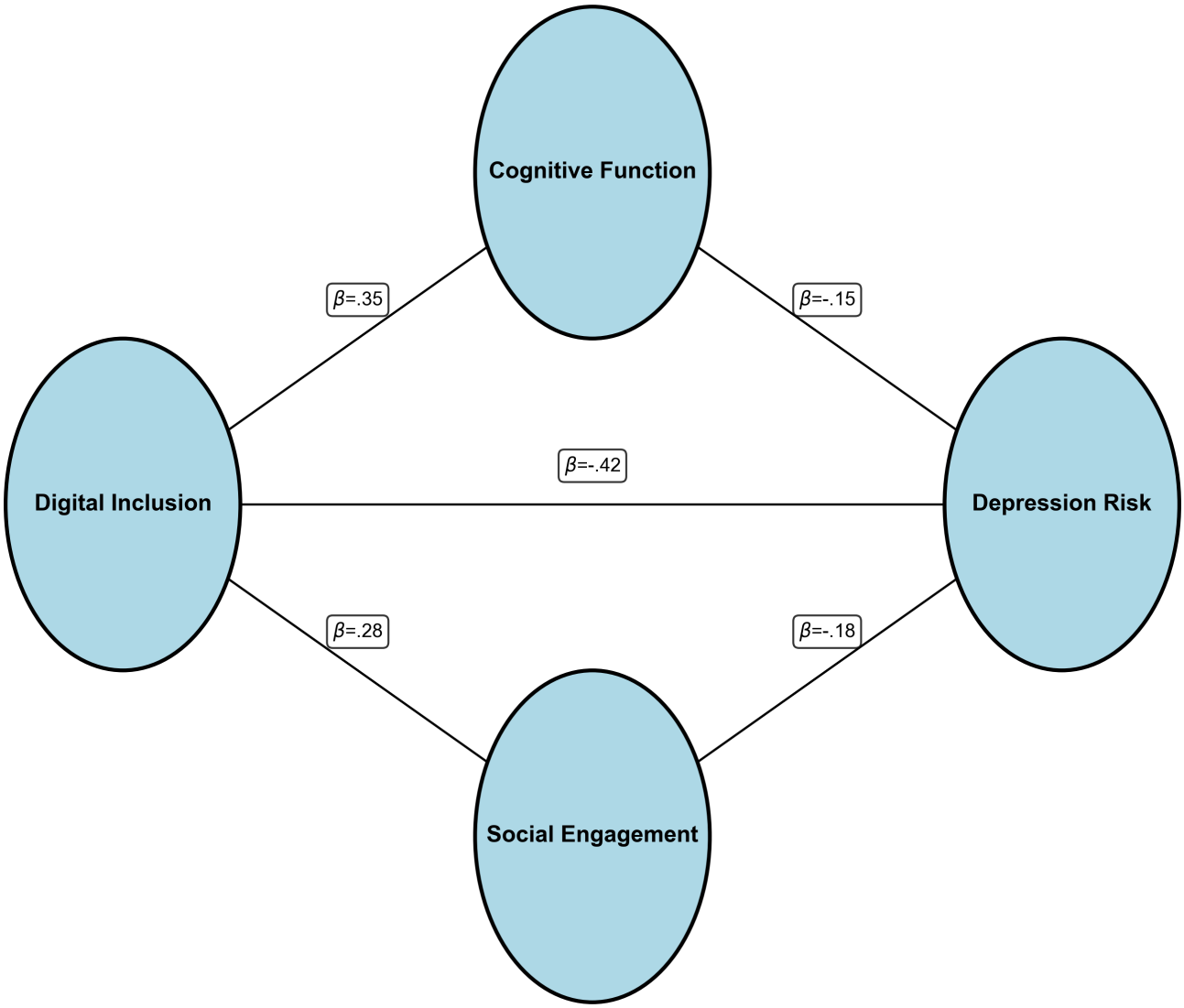


Multiple Pathway Analysis

To explore the mechanisms through which digital inclusion is associated with depression risk, we conducted path analysis to test multiple mediation effects. The model fit was good (χ^2_3=14.27, *P*=.009; CFI=0.996; TLI=0.989; RMSEA=0.032; SRMR=0.011). Table 3 and Figure 2 present the path analysis results.

XSL•FO
RenderX

Figure 2. Path analysis model of digital inclusion's associations with depression risk. The figure shows digital inclusion being associated with depression risk through direct effects and through 2 mediation pathways: cognitive function and social participation. Path coefficients represent standardized effects (SE in parentheses). Line thickness is proportional to effect size. Control variables (age, gender, education, marital status, health status) are included but not shown. *P<.05, **P<.01, ***P<.001. Model fit: χ^2_3=14.27, P=.009; CFI=0.996; TLI=0.989; RMSEA=0.032; SRMR=0.011.



XSL•FO
RenderX

Table . Path analysis: Multiple pathways of digital inclusion's associations with depression[a].

Pathway	Path coefficient	Standard error	95% CI	*P* value	Proportion of effect (%)
Direct effect (Digital inclusion → Depression)	−0.42[b]	0.06	[−0.53,−0.31]	<.001	66.7 [60% - 73%]
Indirect effect 1 (via Cognitive function)	−0.05[c]	0.02	[−0.09,−0.01]	.03	8.3 [2% - 15%]
Indirect effect 2 (via Social participation)	−0.05[d]	0.02	[−0.08,−0.02]	.003	8.0 [3% - 14%]
Indirect effect 3 (Cognitive function → Social participation)	−0.02[c]	0.01	[−0.04,−0.01]	.03	2.8 [1% - 6%]
Unexplained effect	-0.09	—[e]	—	—	14.2
Total effect	-0.63[b]	0.05	[−0.72,−0.54]	<.001	100

[a]95% CI derived from 2000 bootstrap iterations; all pathways control for age, gender, education level, marital status, and health status.

[b]*P*<.001.

[c]*P*<.05.

[d]*P*<.01.

[e]—: missing or unavailable data.

Digital inclusion had a significant direct association with depression risk (β=−.42, *P*<.001), while also exerting associations through multiple indirect pathways: (1) digital inclusion was positively associated with cognitive function (β=.35, *P*<.001), which in turn was negatively related to depression (β=−.15, *P*=.03), forming a cognitive function mediation pathway; (2) digital inclusion promoted social participation (β=.28, *P*=.008), which was associated with reduced depression risk (β=−.18, *P*=.009), forming a social participation mediation pathway; and (3) digital inclusion also exerted associations through a sequential mediation pathway by first being related to cognitive function, then being associated with social participation, and finally relating to depression.

Decomposition of the total effect showed that the direct effect accounted for 66.7% of the total effect, cognitive function mediation effect for 8.3%, social participation mediation effect for 8.0%, sequential mediation effect for 2.8%, and the remaining 14.2% may involve unmeasured pathways. This indicates that digital inclusion primarily is associated with depression through direct mechanisms, but cognitive enhancement and social integration are also important pathways of association.

Cognitive Domain Specificity Analysis

To test whether specific cognitive domains were more critical for moderating the relationship between digital inclusion and depression, we utilized the itemized measurements from CHARLS cognitive tests to decompose cognitive function into 4 domains: immediate memory, delayed memory, orientation, and executive function. Table 4 presents the stratified moderation effects of different cognitive domains.

Table . Moderation effect analysis of different cognitive domains[a].

Cognitive domain	Interaction effect coefficient	Standard error	*P* value
Immediate memory	−0.003[b]	0.001	.02
Delayed memory	−0.002	0.001	.08
Orientation	−0.001	0.001	.32
Executive function	−0.004[c]	0.001	.002

[a]All models control for the same covariates.

[b]**P*<.05.

[c]***P*<.01.

The results showed that executive function exhibited the strongest moderation effect (β=−.004, *P*=.007), followed by immediate memory (β=−.003, *P*=.03), while the moderation effects of delayed memory (β=−.002, *P*=.08) and orientation (β=−.001, *P*=.32) were not significant. This suggests that executive function (including planning, working memory, and task-switching abilities) and immediate memory are particularly important for the protective association of digital technology use, providing more detailed evidence for the heterogeneous moderation effect of cognitive function.

XSL•FO
RenderX

Sensitivity Analysis and Methodological Robustness

To verify the robustness of the results, we conducted various sensitivity analyses. First, we examined the moderation effects of different dimensions of digital inclusion. The digital use depth dimension showed the strongest cognitive moderation (β=−.003, *P*=.008), followed by functional use breadth (β=−.002, *P*=.04), while the moderation of the device access dimension was not significant (β=−.001, *P*=.18). This suggests that cognitive function may play a more critical moderating role in more complex digital activities.

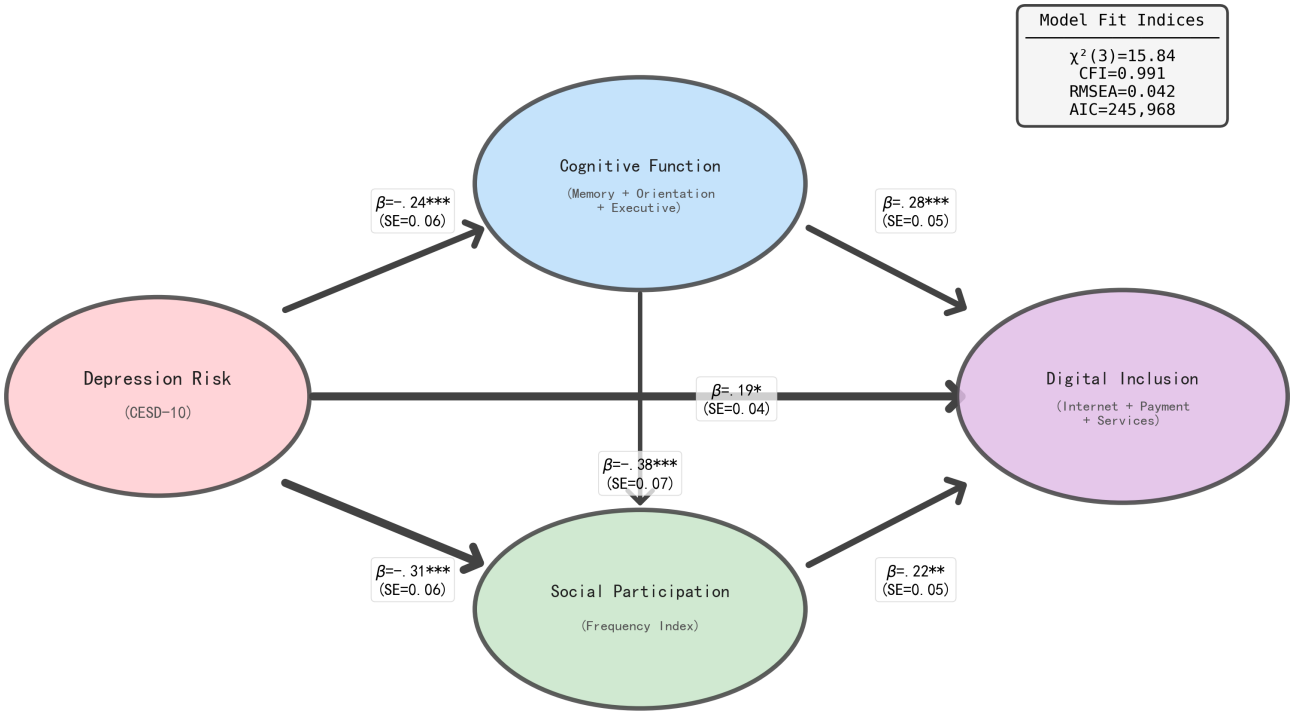
Second, we conducted demographic characteristic stratified analysis. The cognitive moderation effect was slightly stronger in female participants (β=−.003, *P*=.007) than in male participants (β=−.002, *P*=.04), stronger in younger elderly (60 - 70 y) (β=−.003, *P*=.009) than in older elderly (>70 y) (β=−.001, *P*=.09), and stronger in educated older adults (β=−.003, *P*<.01) than in uneducated ones (β=−.001, *P*=.17). This indicates that the cognitive moderation effect itself may be influenced by demographic characteristics.

Finally, we employed alternative model specifications—including tobit regression to handle the censoring of depression scores, quantile regression to analyze effects at different depression levels, and robust regression to handle potential outliers. All alternative specifications supported the significant moderating role of cognitive function, confirming the robustness of the research findings.

Alternative Model Specification

To examine the directionality of relationships, we tested an alternative model where depression risk was specified as the predictor of digital inclusion, cognitive function, and social participation. This reverse-causation model was compared to our hypothesized model using standard fit indices. The path structure and coefficients for this alternative model (Model 2: Depression → Digital Inclusion) are presented in Figure 3.

Figure 3. Alternative model specification: Depression risk as a predictor of digital inclusion, cognitive function, and social participation. *P<.05; **P<.01; ***P<.001. AIC: Akaike information criterion; CFI: comparative fit index; RMSE: root mean square error.



As shown in Figure 3, the alternative model demonstrated acceptable fit indices (χ^2_3=15.84, *P*=.007; CFI=0.991; RMSEA=0.042), with all 6 path coefficients reaching statistical significance. In this reverse-causation model, depression was negatively associated with cognitive function (β=−.24, *P*<.001) and social participation (β=−.31, *P*<.001), while showing a positive association with digital inclusion (β=.19, *P*<.05). However, systematic comparison of model fit revealed that the hypothesized model provided significantly better fit to the data.

Path coefficients represent standardized effects (SE in parentheses). Line thickness is proportional to effect size. This reverse-causation model (Model 2) examines whether depression predicts subsequent digital inclusion and related variables, providing a test of alternative causal directions to the hypothesized model (Figure 2). In this model, depression risk (measured by CES-D-10 scores) is positioned as the independent variable, with arrows indicating hypothesized directional relationships to cognitive function (memory, orientation, and executive function), social participation (frequency index of community activities), and digital inclusion (internet and payment services usage). The model fit values were as follows: χ^2_3=15.84, *P*=.007; CFI =0.991; RMSEA=0.042; AIC=245,968. Control variables (age, gender, education, marital status, household registration, income, retirement status, health status, number of chronic diseases, and activities of daily living or instrumental activities of daily living scores) are included in the model but not shown in the figure for clarity. The hypothesized model (Model 1, Figure 2) demonstrated superior fit across all indices (AIC=245,832, CFI =0.996, RMSEA =0.032), with a significant χ^2 difference ($\Delta\chi^2$=136, *df*=0,

XSL•FO
RenderX

P<.001). The lower AIC value for Model 1 indicates better balance between model fit and parsimony. While both models show acceptable fit (CFI>0.99, RMSEA<0.05), the consistently better performance of Model 1 across all fit criteria (lower AIC and RMSEA, higher CFI) provides empirical support for the theorized direction from digital inclusion to depression rather than reverse causation. However, given the cross-sectional nature of the data, bidirectional relationships cannot be definitively ruled out, and longitudinal research is needed to establish temporal precedence and causal directionality.

The results showed that Model 1 (our hypothesized model) provided significantly better fit to the data (ΔAIC=136; $\Delta\chi^2$=136, *P*<.001), suggesting that while both directions are theoretically plausible with cross-sectional data, the pattern of associations is more consistent with digital inclusion as a predictor of depression rather than vice versa. The hypothesized model demonstrated superior CFI (0.996 vs 0.991) and lower RMSEA (0.032 vs 0.042), supporting our theoretical framework. Although the alternative model also showed acceptable fit indices (CFI>0.99, RMSEA<0.05), the consistent superiority of the hypothesized model across all fit criteria (lower AIC/BIC, higher CFI, lower RMSEA) provides empirical support for our theoretical direction. However, bidirectional relationships cannot be ruled out given the cross-sectional nature of the data, and longitudinal research is needed to establish temporal precedence.

Discussion

Principal Findings

This study systematically examined the moderating role of cognitive function in the relationship between digital inclusion and depression risk among older adults and explored the multiple pathways through which digital inclusion is associated with mental health outcomes. Our cross-sectional findings reveal a significant cognitive moderation effect, demonstrating that the protective associations of digital inclusion are contingent on cognitive capacities. Additionally, we identified 3 significant pathways—direct effects, cognitive enhancement, and social integration—through which digital inclusion is associated with depression risk. These findings have important theoretical and practical implications for understanding and promoting digital inclusion among older adults.

The Cognitive Threshold Effect: Theoretical Implications

The significant moderating role of cognitive function in the digital inclusion-depression relationship constitutes our most novel and theoretically important finding. In this cross-sectional study, we observed that the association between digital inclusion and depression was not statistically significant below certain cognitive function levels, but increasingly strong protective associations above this threshold—a pattern we term the "cognitive threshold effect." As demonstrated by the simple slope analysis (β=−.137, *P*=.33 at 326.57 points vs β=−.517, *P*<.001 at 798.13 points), the association between digital inclusion and depression varied substantially across cognitive function levels. This finding extends current theoretical understanding in several important ways.

First, this cognitive threshold effect advances cognitive aging theory by demonstrating how age-related cognitive changes are associated with technology's psychological impacts. Traditional cognitive aging frameworks focus primarily on how cognitive changes are related to technology adoption and usage patterns [8,19], with limited attention to how these same cognitive changes might moderate technology's psychological outcomes. Our findings suggest that the processing resource limitations central to cognitive aging theory [14] not only create barriers to technology use but also fundamentally alter the psychological consequences of that use. Our results indicate that for older adults with limited cognitive resources, digital engagement may be associated with significant attentional demands; conversely, those with stronger cognitive resources may navigate digital environments more efficiently, experiencing technology as empowering rather than burdensome.

Second, our findings provide nuanced support for the cognitive reserve hypothesis in the digital context. While previous research has documented associations between digital activities and cognitive function [16,32], our study reveals important bidirectionality in this relationship. Cognitive function not only is associated with benefits from digital engagement but also moderates digital participation's mental health associations, suggesting a complex, mutually reinforcing relationship. The cognitive reserve hypothesis typically emphasizes how cognitive stimulation builds neural resources that protect against pathology; our findings suggest that these same neural resources may determine whether digital engagement functions as beneficial cognitive stimulation or overwhelming cognitive demand.

Third, the cognitive threshold effect suggests important theoretical refinements to existing digital inclusion frameworks. Current conceptualizations of digital inclusion often implicitly assume uniform effects across user populations or simple linear relationships between digital participation and outcomes [9,10]. Our findings call for more nuanced theoretical models that account for individual difference factors—particularly cognitive capacities—that moderate digital participation's effects. This perspective aligns with emerging person-environment fit approaches in gerontechnology, which emphasize the match between individual capabilities and technological demands [18].

Multiple Pathways: Understanding Mechanisms of Association

Our path analysis revealed 3 significant pathways through which digital inclusion is associated with depression risk: direct effects (66.7% of total effect), cognitive enhancement (8.3%), and social integration (8.0%), with an additional sequential pathway through cognitive function to social participation (2.8%). This multiple pathway structure advances theoretical understanding of the mechanisms linking digital participation to psychological outcomes.

The direct pathway's predominance suggests that immediate psychological associations of digital engagement—including sense of mastery, environmental control, access to information resources, and enjoyable leisure experiences—constitute the primary mechanism through which digital inclusion is associated with depression risk. This finding aligns with previous research

XSL•FO
RenderX

highlighting technology's role in enhancing perceived autonomy and environmental control among older adults [20,21]. Digital engagement may function as an important source of self-efficacy experiences for older adults, generating feelings of competence and mastery that directly counteract depression-linked feelings of helplessness and loss of control. Additionally, access to health information, entertainment resources, and practical tools may directly address specific depression risk factors common in late life, including health anxiety, boredom, and practical life difficulties.

The cognitive enhancement pathway's significance (8.3% of total effect) provides important empirical support for the cognitive reserve hypothesis in the digital context. This pathway suggests that digital engagement is positively associated with cognitive function, which in turn is related to reduced depression. This mechanism aligns with emerging research on the cognitive associations of digital activities [16,17] and established links between cognitive function and depression risk [25]. Importantly, the strength of this pathway likely depends on the specific types of digital activities undertaken. Cognitively challenging digital pursuits—such as information searching, online learning, and strategic communication—may provide greater cognitive stimulation than passive consumption activities, potentially generating stronger cognitive enhancement effects.

The social integration pathway's substantial contribution (8% of total effect) confirms the importance of social mechanisms in digital inclusion's psychological effects. This pathway suggests that digital technologies facilitate social connections and participation, enhancing social resources that are associated with protection against depression. This mechanism aligns with established social integration theory [29] and emerging research on digital social connection among older adults [13,23]. Digital communication tools may be particularly valuable for older adults facing mobility limitations, geographical separation from family, or social role losses, helping maintain existing relationships and develop new connections. Additionally, online communities organized around shared interests may provide valuable alternative social participation opportunities for older adults with limited local options.

No Evidence of Harmful Effects

An important finding worthy of emphasis is that regardless of cognitive function level, higher digital inclusion was not found to be associated with higher depression risk. In the low cognitive function group, the association was not statistically significant (β=−.137, P=.33) rather than positive. This contrasts with literature documenting technostress and information overload as potential negative consequences of digital engagement [5,6].

Our findings suggest that among older adults who choose to engage digitally, the potential risks may be minimal, while the benefits are substantial for those with adequate cognitive resources. This asymmetric pattern—where benefits depend on cognitive function, but no harmful effects were detected—has important implications for policy. It suggests that promoting digital inclusion is unlikely to harm older adults with lower cognitive function, though they may not derive the same mental health benefits as their higher functioning peers.

Practical Implications

The cognitive threshold effect has profound implications for digital inclusion policy and practice. Most significantly, it suggests that cognitive assessment should be integrated into digital inclusion program design. Current digital inclusion initiatives typically adopt standardized approaches without systematically assessing or accommodating cognitive differences [33]. Our findings suggest that cognitive screening before digital training could identify individuals requiring additional cognitive support and inform appropriate instructional approaches. Simple cognitive assessments measuring processing speed, working memory, and executive function could help tailor digital training to match cognitive capacities, ensuring that participants engage with appropriately challenging but manageable digital activities.

These recommendations align with recent intervention research. For example, Dodge et al [34] conducted a randomized controlled trial among socially isolated older adults, demonstrating that video chat-based conversational interactions improved both cognitive and emotional outcomes. Their findings provide experimental evidence that digital interventions can simultaneously target multiple pathways—social integration and cognitive enhancement—supporting our theoretical framework of interconnected mechanisms. Such multicomponent digital interventions may be particularly effective for older adults at risk of social isolation and cognitive decline.

Our finding that cognitive function both moderates and is associated with digital inclusion's effects suggests promising opportunities for integrated digital skills and cognitive training interventions. Rather than viewing digital skills training and cognitive training as separate domains, practitioners could develop synergistic interventions that simultaneously enhance digital skills and cognitive abilities, creating virtuous cycles of improvement. Such integrated interventions might incorporate cognitive training principles into digital skills instruction. For example, digital training could systematically progress from simpler to more complex tasks, gradually increasing cognitive demands while building technological skills.

Our identification of multiple significant pathways suggests that comprehensive interventions addressing multiple mechanisms simultaneously may be most effective. Rather than focusing exclusively on digital skills acquisition, interventions should explicitly address cognitive enhancement, social connection, and direct psychological benefits. For the direct pathway, interventions should emphasize experiences that enhance sense of mastery, control, and self-efficacy. For the cognitive enhancement pathway, interventions should prioritize digital activities with greater cognitive stimulation potential. For the social integration pathway, interventions should explicitly address both technical skills and social strategies for digital communication.

Limitations and Future Directions

Most critically, the cross-sectional nature of our data precludes causal inference. While we use terms such as "moderation" and "mediation" to describe statistical relationships, these should be interpreted as associations rather than established causal pathways. The observed patterns are consistent with our

XSL•FO
RenderX

theoretical framework, but alternative causal directions cannot be ruled out. For example, depression risk may influence both cognitive function and digital inclusion rather than vice versa. While our alternative model specification (see the *Alternative Model Specification* section) showed that the hypothesized model fit the data better than the reverse-causation model, this provides only suggestive evidence given the cross-sectional design. Longitudinal research examining temporal sequences and potentially bidirectional relationships is essential to establish true causal effects and understand the dynamic interplay between digital inclusion, cognitive function, and mental health over time.

Second, our measurement of digital inclusion, while multidimensional, did not fully capture qualitative aspects of digital engagement. The specific content, purposes, and contexts of digital activities likely are associated with both cognitive demands and psychological outcomes. Third, while our path analysis identified significant mediating pathways, alternative model specifications might yield different conclusions about pathway structure and relative importance. Fourth, our sample, while large and nationally representative, derived from a single country with unique cultural, technological, and social characteristics. Finally, our focus on depression as the primary outcome measure may not capture the full range of psychological effects associated with digital inclusion.

Building on our findings, particularly promising directions for future research include (1) investigating potential critical periods when cognitive function most strongly moderates digital inclusion's associations, particularly during early adoption phases versus established use; (2) exploring whether cognitive training interventions can raise individuals above the cognitive threshold, enabling greater psychological benefits from digital participation; (3) examining whether certain digital design features might reduce cognitive demands, potentially lowering the cognitive threshold for benefit; and (4) developing and testing integrated cognitive-digital interventions based on our theoretical framework. Such research would advance both theoretical understanding and practical applications of our findings.

Conclusions

This study provides significant new insights into the complex relationship between digital inclusion and depression risk among older adults, revealing cognitive function's critical moderating role and identifying multiple pathways of associations. The cognitive threshold effect demonstrates that digital inclusion's psychological benefits are not universally accessible but contingent on cognitive capacities. This finding challenges simplistic approaches to digital inclusion and calls for more nuanced, cognitively informed interventions that match digital activities to individual cognitive resources.

The multiple pathways we identified—direct effects, cognitive enhancement, and social integration—provide a comprehensive framework for understanding digital inclusion's psychological associations. This framework suggests that comprehensive interventions addressing multiple mechanisms simultaneously may be most effective, particularly approaches integrating digital skills development with cognitive enhancement and social connection strategies.

As digital technologies become increasingly central to daily functioning, ensuring that older adults can participate meaningfully in digital society becomes an important public health concern. Our findings suggest that simply promoting technology adoption without considering cognitive factors may be insufficient and potentially counterproductive for some older adults. Instead, cognitively informed digital inclusion strategies—tailoring approaches to individual cognitive capacities, integrating cognitive and digital skills training, and addressing multiple pathways simultaneously—offer more promising approaches for leveraging digital technologies to promote mental health in an aging society.

These insights have particular relevance in post-pandemic contexts, where digital participation has become increasingly essential for accessing health care, maintaining social connections, and meeting daily needs. By illuminating the critical role of cognitive function and multiple mechanisms of association, this research provides a stronger scientific foundation for digital inclusion policies and programs that effectively promote mental health among increasingly diverse older populations in our rapidly digitalizing world.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability

The data that support the findings of this study are available from the China Health and Retirement Longitudinal Study, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of the China Health and Retirement Longitudinal Study.

Conflicts of Interest

None declared.

References

XSL•FO
RenderX

1. Reisdorf BC, Rhinesmith C. Digital inclusion as a core component of social inclusion. Soc Incl 2020;8(2):132-137. [doi: 10.17645/si.v8i2.3184]
2. The 49th statistical report on China's internet development. : China Internet Network Information Center; 2022 URL: https://www1.cnnic.cn/ [accessed 2025-11-08]
3. Cotten SR, Ford G, Ford S, Hale TM. Internet use and depression among retired older adults in the United States: a longitudinal analysis. J Gerontol B Psychol Sci Soc Sci 2014 Sep;69(5):763-771. [doi: 10.1093/geronb/gbu018] [Medline: 24671896]
4. Chen YRR, Schulz PJ. The effect of information communication technology interventions on reducing social isolation in the elderly: a systematic review. J Med Internet Res 2016 Jan 28;18(1):e18. [doi: 10.2196/jmir.4596] [Medline: 26822073]
5. Nimrod G. Technostress: measuring a new threat to well-being in later life. Aging Ment Health 2018 Aug;22(8):1080-1087. [doi: 10.1080/13607863.2017.1334037] [Medline: 28562064]
6. Reinecke L, Aufenanger S, Beutel ME, et al. Digital stress over the life span: the effects of communication load and internet multitasking on perceived stress and psychological health impairments in a German probability sample. Media Psychol 2017 Jan 2;20(1):90-115. [doi: 10.1080/15213269.2015.1121832]
7. Salthouse TA. Consequences of age-related cognitive declines. Annu Rev Psychol 2012;63:201-226. [doi: 10.1146/annurev-psych-120710-100328] [Medline: 21740223]
8. Czaja SJ, Charness N, Fisk AD, et al. Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). Psychol Aging 2006 Jun;21(2):333-352. [doi: 10.1037/0882-7974.21.2.333] [Medline: 16768579]
9. van Dijk JAGM. The Deepening Divide: Inequality in the Information Society: Sage Publications; 2005. [doi: 10.4135/9781452229812]
10. Helsper EJ. A corresponding fields model for the links between social and digital exclusion. Commun Theor 2012 Nov;22(4):403-426. [doi: 10.1111/j.1468-2885.2012.01416.x]
11. Quan-Haase A, Williams C, Kicevski M, Elueze I, Wellman B. Dividing the grey divide: deconstructing myths about older adults' online activities, skills, and attitudes. American Behavioral Scientist 2018 Aug;62(9):1207-1228. [doi: 10.1177/0002764218777572]
12. Vroman KG, Arthanat S, Lysack C. "Who over 65 is online?" Older adults' dispositions toward information communication technology. Comput Human Behav 2015 Feb;43:156-166. [doi: 10.1016/j.chb.2014.10.018]
13. Chopik WJ. The benefits of social technology use among older adults are mediated by reduced loneliness. Cyberpsychol Behav Soc Netw 2016 Sep;19(9):551-556. [doi: 10.1089/cyber.2016.0151] [Medline: 27541746]
14. Craik FIM, Byrd M. Aging and cognitive deficits: the role of attentional resources. In: Craik FIM, Trehub S, editors. Aging and Cognitive Processes: Plenum; 1982:191-211. [doi: 10.1007/978-1-4684-4178-9_11]
15. Stern Y. Cognitive reserve in ageing and Alzheimer's disease. Lancet Neurol 2012 Nov;11(11):1006-1012. [doi: 10.1016/S1474-4422(12)70191-6] [Medline: 23079557]
16. Tun PA, Lachman ME. The association between computer use and cognition across adulthood: use it so you won't lose it? Psychol Aging 2010 Sep;25(3):560-568. [doi: 10.1037/a0019543] [Medline: 20677884]
17. Small GW, Moody TD, Siddarth P, Bookheimer SY. Your brain on Google: patterns of cerebral activation during internet searching. Am J Geriatr Psychiatry 2009 Feb;17(2):116-126. [doi: 10.1097/JGP.0b013e3181953a02] [Medline: 19155745]
18. Rogers WA, Fisk AD. Toward a psychological science of advanced technology design for older adults. J Gerontol B Psychol Sci Soc Sci 2010 Nov;65(6):645-653. [doi: 10.1093/geronb/gbq065] [Medline: 20833690]
19. Charness N, Boot WR. Aging and information technology use: potential and barriers. Curr Dir Psychol Sci 2009;18(5):253-258. [doi: 10.1111/j.1467-8721.2009.01647.x]
20. Francis J, Ball C, Kadylak T, Cotten SR. Aging in the digital age: conceptualizing technology adoption and digital inequalities. In: Ageing and Digital Technology: Springer; 2019:35-49. [doi: 10.1007/978-981-13-3693-5_3]
21. Czaja SJ, Boot WR, Charness N, Rogers WA, Sharit J. Improving social support for older adults through technology: findings from the PRISM randomized controlled trial. Gerontologist 2018 May 8;58(3):467-477. [doi: 10.1093/geront/gnw249] [Medline: 28201730]
22. Chou KL, Cronin A, Chi I. Concerns about health and independence: a predictive problem for depression among older Chinese. J Gerontol Soc Work 2012;55(6):540-555. [doi: 10.1080/01634372.2012.675836]
23. Szabo A, Allen J, Stephens C, Alpass F. Longitudinal analysis of the relationship between purposes of internet use and well-being among older adults. Gerontologist 2019 Jan 9;59(1):58-68. [doi: 10.1093/geront/gny036] [Medline: 29688332]
24. Myhre JW, Mehl MR, Glisky EL. Cognitive benefits of online social networking for healthy older adults. J Gerontol B Psychol Sci Soc Sci 2017 Sep 1;72(5):752-760. [doi: 10.1093/geronb/gbw025] [Medline: 26984523]
25. Köhler S, van Boxtel MPJ, van Os J, et al. Depressive symptoms and cognitive decline in community-dwelling older adults. J Am Geriatr Soc 2010 May;58(5):873-879. [doi: 10.1111/j.1532-5415.2010.02807.x] [Medline: 20374396]
26. Slegers K, van Boxtel M, Jolles J. Effects of computer training and internet usage on cognitive abilities in older adults: a randomized controlled study. Aging Clin Exp Res 2009 Feb;21(1):43-54. [doi: 10.1007/BF03324898] [Medline: 19225269]
27. Santini ZI, Koyanagi A, Tyrovolas S, Mason C, Haro JM. The association between social relationships and depression: a systematic review. J Affect Disord 2015 Apr 1;175:53-65. [doi: 10.1016/j.jad.2014.12.049] [Medline: 25594512]

XSL•FO
RenderX

28. Lövdén M, Bäckman L, Lindenberger U, Schaefer S, Schmiedek F. A theoretical framework for the study of adult cognitive plasticity. Psychol Bull 2010 Jul;136(4):659-676. [doi: 10.1037/a0020080] [Medline: 20565172]
29. Berkman LF, Glass T, Brissette I, Seeman TE. From social integration to health: Durkheim in the new millennium. Soc Sci Med 2000 Sep;51(6):843-857. [doi: 10.1016/s0277-9536(00)00065-4] [Medline: 10972429]
30. Quan-Haase A, Martin K, Schreurs K. Interviews with digital seniors: ICT use in the context of everyday life. Inf Commun Soc 2016 May 3;19(5):691-707. [doi: 10.1080/1369118X.2016.1140217]
31. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
32. Xavier AJ, d'Orsi E, de Oliveira CM, et al. English longitudinal study of aging: can Internet/E-mail use reduce cognitive decline? J Gerontol A Biol Sci Med Sci 2014 Sep;69(9):1117-1121. [doi: 10.1093/gerona/glu105] [Medline: 25116923]
33. Seifert A, Cotten SR, Xie B. A double burden of exclusion? Digital and social exclusion of older adults in times of COVID-19. J Gerontol B Psychol Sci Soc Sci 2021 Feb 17;76(3):e99-e103. [doi: 10.1093/geronb/gbaa098] [Medline: 32672332]
34. Dodge HH, Yu K, Wu CY, et al. Internet-based conversational engagement randomized controlled clinical trial (I-CONECT) among socially isolated adults 75+ years old with normal cognition or mild cognitive impairment: topline results. Gerontologist 2024 Apr 1;64(4):gnad147. [doi: 10.1093/geront/gnad147] [Medline: 37935416]

Abbreviations

CFI: comparative fit index
CHARLS: China Health and Retirement Longitudinal Study
PCA: principal component analysis
RMSEA: root mean square error of approximation
SRMR: standardized root mean square residual
TLI: Tucker-Lewis index

Edited by PH Chau; submitted 16.07.25; peer-reviewed by H Dodge, T Sun; revised version received 18.10.25; accepted 19.10.25; published 25.11.25.

Please cite as:
Xiao G, Nie T
The Critical Moderating Role of Cognitive Function in Digital Inclusion: Data Analysis Study on Depression Risk Among Older Adults
JMIR Aging 2025;8:e80700
URL: https://aging.jmir.org/2025/1/e80700
doi:10.2196/80700

© Gang Xiao, Tingting Nie. Originally published in JMIR Aging (https://aging.jmir.org), 25.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Research Status and Trends in Virtual Reality Technology for Older Adults: Bibliometric and Visual Analysis

Jing Xu[1], MS; Wenjin Zhang[2], MS; WenLi Liu[3], MS

[1]School of Nursing, Shandong Xiandai University, Jinan, China

[2]Shanxi Bethune Hospital, Shanxi Academy of Medical Sciences, Third Hospital of Shanxi Medical University, Tongji Shanxi Hospital, No. 99 Longcheng Street, Xiaodian District, Taiyuan, Shanxi, China

[3]Linfen Central Hospital, Linfen, China

Corresponding Author:
Wenjin Zhang, MS
Shanxi Bethune Hospital, Shanxi Academy of Medical Sciences, Third Hospital of Shanxi Medical University, Tongji Shanxi Hospital, No. 99 Longcheng Street, Xiaodian District, Taiyuan, Shanxi, China

Abstract

Background: Virtual reality (VR) technology is increasingly applied in aging-related research. Although existing bibliometric studies have focused on specific applications, such as older adults' acceptance of VR and its use in cognitive rehabilitation, no comprehensive mapping of the global research landscape on VR for older populations has been conducted. This study fills this gap by providing a holistic bibliometric and thematic analysis of VR applications in older adults, mapping research trends, intellectual structures, and emerging frontiers.

Objective: This study aims to explore the current applications, potential benefits, and future directions of virtual reality technology for older adults, based on literature published between January 1, 2015, and April 30, 2025.

Methods: This study used bibliometric methods to systematically examine the current status and developmental trends in VR research for older adults. We searched the Web of Science Core Collection for research articles and reviews published in English. A total of 1609 publications were included in the final analysis. Using CiteSpace and VOSviewer, we conducted coauthorship network analysis, keyword clustering, and burst detection to map research hot spots, academic collaboration patterns, and emerging trends in the field.

Results: Our analysis of 1609 publications revealed a steady growth in the application of VR technology for older adults. The predominant research areas included meta-analysis, rehabilitation, dementia, and gait. The United States and China were the two most productive countries, with Tel Aviv University emerging as the leading institution. *Frontiers in Aging Neuroscience* and *Applied Sciences Basel* were the most prolific journals, each publishing 40 papers. The most cited article evaluated the effects of VR-based physical and cognitive training on executive function and dual-task gait performance in older adults with mild cognitive impairment. Emerging research themes include artificial intelligence, association, and depression.

Conclusions: VR research for older adults is rapidly expanding and globally collaborative. Although applications span multiple geriatric domains, future efforts should prioritize mental health, disease integration, and artificial intelligence–enhanced VR technologies.

(JMIR Aging 2025;8:e76609) doi:10.2196/76609

KEYWORDS

virtual reality; older adults; bibliometric analysis; CiteSpace; visual analysis

Introduction

Background

According to the World Health Organization, the global population aged 60 years and older reached 1.1 billion in 2023 and is projected to double to 2.1 billion by 2050 [1,2]. Declines in physical and mental health, loss of functional capacity, and social isolation hinder active aging among older adults. These factors not only threaten their quality of life but also pose significant challenges to global public health [3]. In this context, virtual reality (VR) has emerged as a promising interactive technology, demonstrating broad application potential in the field of older adults [4-8].

VR systems are defined as highly interactive 3D digital media platforms [9], with applications across diverse environments, including health care innovation [10]. They enhance user-environment interactions, offer real-time performance-based feedback, and improve accessibility and cost-efficiency. Empirical studies demonstrate their potential to prevent social isolation, encourage health and well-being,

XSL•FO
RenderX

and enhance quality of life [11,12]. VR-related research is rapidly expanding, with advancements in VR-based rehabilitation and cognitive intervention yielding substantial clinical evidence [7,13-16].

However, systematic analyses of knowledge structures, trends, and evolutionary patterns in geriatric applications are still scarce. A critical gap exists in leveraging bibliometric and visualization techniques to map disciplinary collaboration networks and developmental trajectories within this domain.

In recent years, bibliometrics has become a key method for assessing academic literature using mathematical and statistical techniques [17,18], and it is recognized as a robust method for unveiling disciplinary development trajectories [19]. Such analyses provide data-driven insights that help researchers identify emerging research frontiers, trace disciplinary evolution, and refine strategic directions [20]. As the most widely used scientific citation database, Web of Science Core Collection (WoSCC) serves as a crucial resource for bibliometric data analysis.

This study used CiteSpace and VOSviewer for bibliometric and visual analysis. The analysis aimed to identify emerging trends and interdisciplinary connections, as well as to examine the impact of VR as a multifunctional intervention tool on research priorities and clinical translation. Specifically, we identified key research hot spots, influential authors, leading institutions, and evolving trends in this field.

Objectives

This study aims to provide a more comprehensive understanding of VR applications in aging research by integrating multiple analytical dimensions, extending beyond the scope of prior bibliometric reviews. The findings are intended to guide researchers and practitioners in the development and implementation of VR technologies for older adults.

Methods

Data Sources and Search Strategy

We searched the WoSCC on May 7, 2025, and completed the search in one day to avoid bias from database updates. The search formula was “TS=(“virtual reality” OR “virtual medicine” OR “augmented reality” OR “mixed reality” OR “virtual simulation”) AND TS=(“older” OR “older adults” OR “elderly” OR “elder person” OR “older people” OR “oldest old” OR “elderly people” OR “elderly patient”). The search was restricted to articles and review articles published in English from January 1, 2015, to April 30, 2025.

We focused our literature screening on the application of VR technology to older adults, reviewing both abstracts and full-text articles. Duplicate and irrelevant topics were systematically removed from the dataset. The exclusion criteria also specified source types, including conference papers, reviews, data papers, edited materials, and retractions. After the initial search, all literature was screened and assessed separately by 3 researchers (JX, WZ, and WL) to ensure the relevance of the selected papers to the research topic. Any disagreements during the analysis process were resolved through internal discussions within the research team and consultations with experts to reach a consensus.

Analysis Tools

In this study, Microsoft Excel (2013; Microsoft Corp) was used for organizing data and performing statistical analyses on the retrieved bibliometric data, including the generation of charts that depict the quantity and growth trends of publications. For the visualization and bibliometric analysis, CiteSpace (version 6.4.R2) and VOSviewer (version 1.6.20) were used to examine various aspects: author productivity and collaboration patterns, journal publication volumes, national and institutional cooperation networks, highly cited references, burst keywords, and keyword co-occurrence patterns. This comprehensive approach has facilitated an in-depth overview of the research progress and developmental trajectories concerning the application of VR technology among the older adult population.

Ethical Considerations

The data were retrieved from WoSCC. This study involved no direct patient or public participation, as it was based solely on published literature. This study was conducted and reported in accordance with the BIBLIO (Guideline for Reporting Bibliometric Reviews of the Biomedical Literature) guidelines (Checklist 1). The recommendations of BIBLIO were followed to ensure full compliance with the applicable research design and reporting framework.

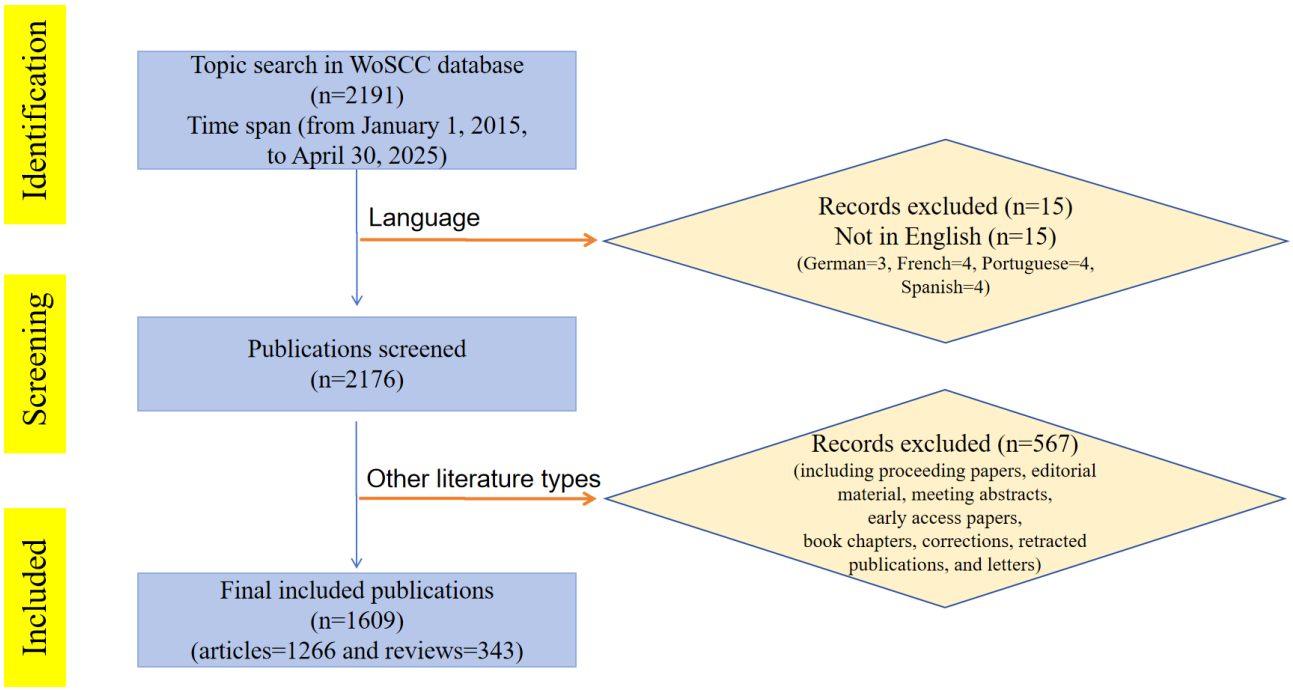
Results

Literature Screening Workflow

Ultimately, 1609 articles met the inclusion criteria for this study. These articles were exported in their entirety, including full records and cited references, and saved as plain-text files in .txt format. The process of searching and filtering is illustrated in Figure 1.

XSL•FO
RenderX

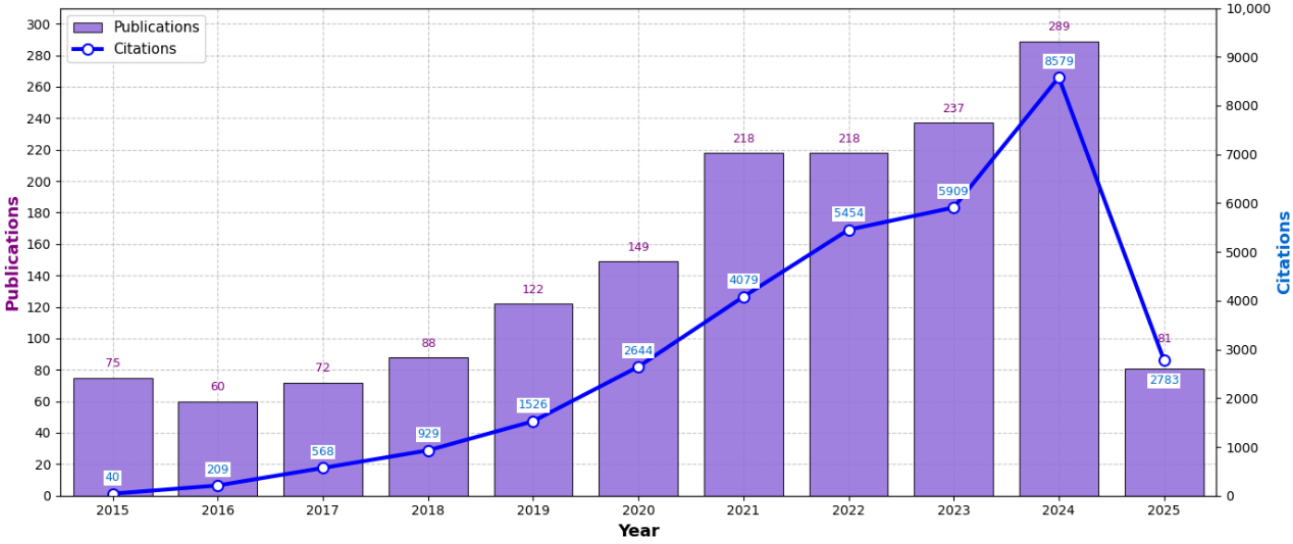
Figure 1. Literature screening workflow for the bibliometric analysis of virtual reality applications in older adults. WoSCC: Web of Science Core Collection.



Number of Publications and Growth Trend

The number of publications on VR applications for older adults has demonstrated consistent growth since 2022, reaching its peak at 289 articles in 2024 (Figure 2). This upward trajectory reflects a growing scholarly interest in VR technologies for aging populations worldwide.

Figure 2. Trends in annual publications and citations on virtual reality applications for older adults. Global bibliometric analysis showed a peak of 289 publications in 2024; note that the 2025 data includes up to April 30.

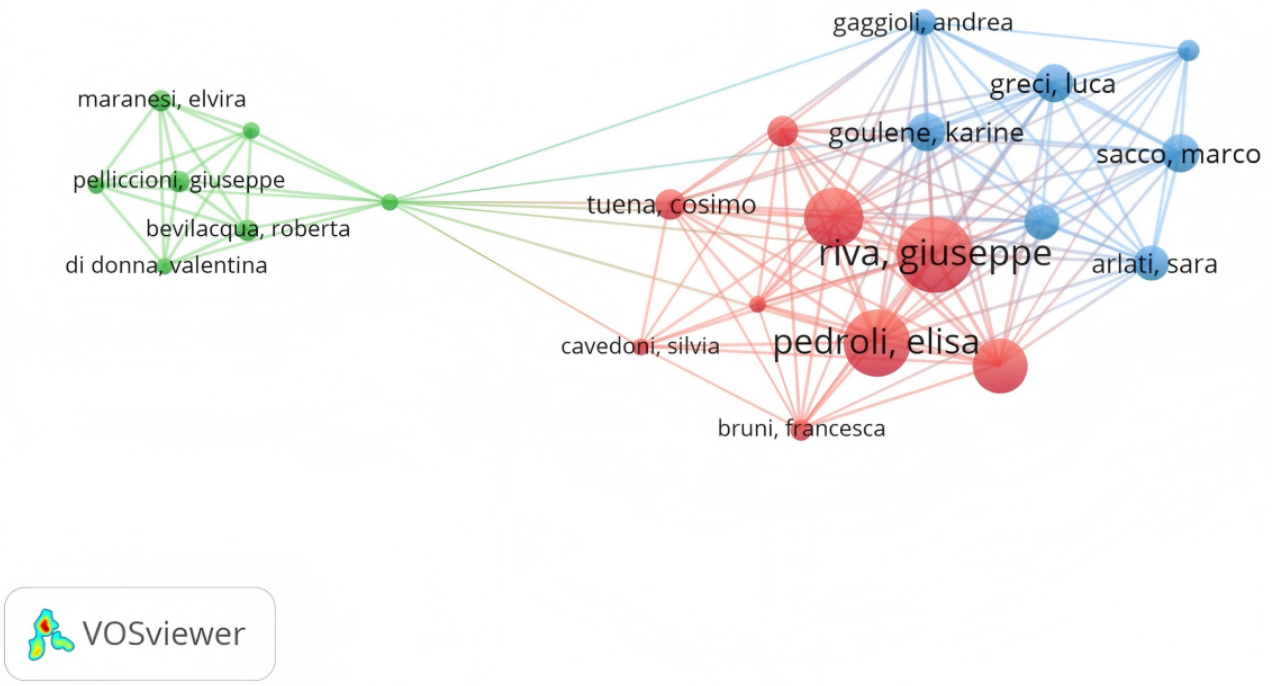


Principal Authors and Coauthorship Trends

Our research identifies that a total of 7291 authors contributed to the study of VR for older adults from January 1, 2015, to April 30, 2025. Furthermore, our observations suggest that there are particularly prolific authors who have established a collaborative network (Figure 3), indicating the development of several high-output research teams in this area.

XSL•FO
RenderX

Figure 3. Global author collaboration network in virtual reality research for older adults.



Publication Journal of VR Research for Older Adults

Our bibliometric analysis revealed that VR-related gerontological research is widely distributed across multiple academic journals. As shown in Table 1, the top 5 most productive journals collectively published 192 articles, accounting for 11.93% of the total publications in this field. *Frontiers in Aging Neuroscience* and *Applied Sciences Basel* led the group, with 40 publications each on VR applications for aging populations, followed by the *International Journal of Environmental Research and Public Health*, which had 39 articles.

Table . Top 5 journals ranked by publication count (as of April 30, 2025).

Number	Journal name	Country	Number of papers	Number of citations	Average citations per publication
1	*Frontiers in Aging Neuroscience*	Switzerland	40	1144	28.6
2	*Applied Sciences Basel*	Switzerland	40	332	8.3
3	*International Journal of Environmental Research and Public Health*	Switzerland	39	768	19.69
4	*Scientific Reports*	England	37	589	15.93
5	*JMIR Serious Games*	Canada	36	693	19.25

Analysis of Countries and Institutes

Our bibliometric analysis reveals contributions from 79 countries/regions and 2452 institutions, collectively producing 1609 publications. As illustrated in Figure 4A, the collaborative network highlights the United States as the leading contributor (n=338), followed by China (n=289) and Italy (n=134). The visualization reveals strong research collaborations between the United States and other leading nations, including Australia, Switzerland, and England.

XSL•FO
RenderX

Figure 4. National and institutional contributions. (A) Distribution of publications from different countries. (B) Global institutional collaboration network in virtual reality research for older adults.



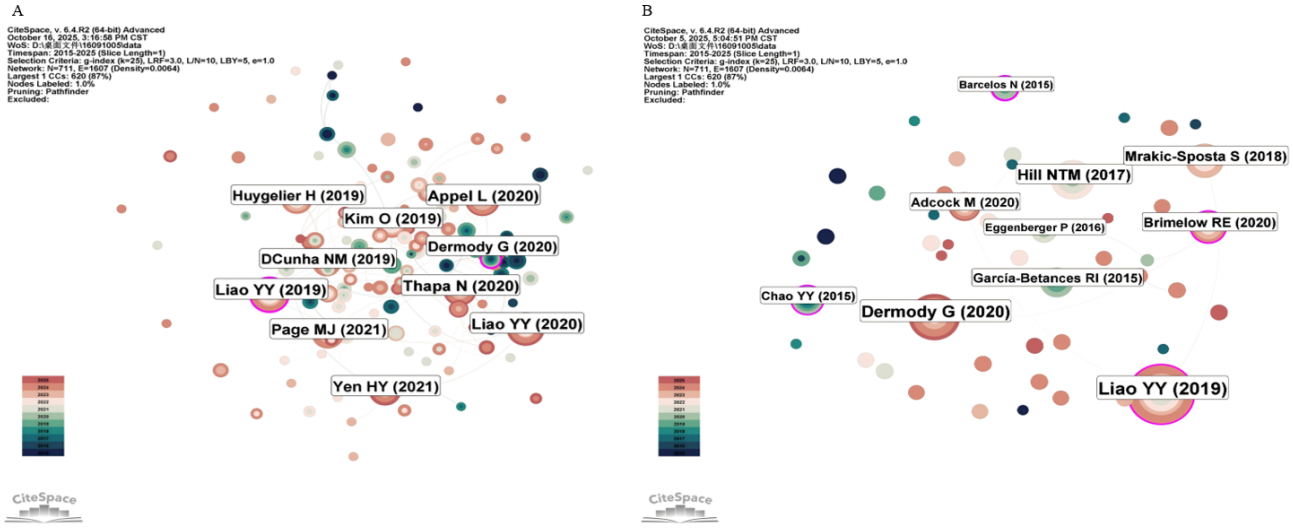
Institutional analysis showed that the top 10 institutions account for 185 publications (11.5% of the total), with Tel Aviv University being the most productive (n=27), followed closely by Hong Kong Polytechnic University (n=25). Figure 4B demonstrates increasingly strong interinstitutional connections.

Literature Cocitation and Cluster Analysis

A citation is a vital bibliometric indicator, with frequently cited studies greatly influencing their research areas [14]. Citation analysis serves as a critical bibliometric tool for assessing scholarly impact, where highly cited publications effectively identify research priorities within a field [21]. Table 2 lists the top 10 most influential articles ranked by total citation count and provides a visual analysis of their citation frequencies, as illustrated in Figure 5A.

XSL•FO
RenderX

Figure 5. An analysis of the literature's impact on the field. (A) Cocitation network visualization based on citation frequency. (B) Centrality map identifying the most influential papers. (C) Top 10 references with the strongest citation bursts.



C

Top 10 References	Year	Strength	Begin	End	2015-2025
1. Rendon AA, 2012, AGE AGEING, V41, P549, DOI 10.1093/ageing/afs053	2012	**11.64**	2015	2017	
2. Jorgensen MG, 2013, J GERONTOL A-BIOL, V68, P845, DOI 10.1093/gerona/gls222	2013	**10.53**	2015	2017	
3. Miller KJ, 2014, AGE AGEING, V43, P188, DOI 10.1093/ageing/aft194	2014	**9.38**	2015	2019	
4. Bieryla KA, 2013, CLIN INTERV OLDER ADULTS, V8, P775, DOI 10.2147/CIA.S46164	2013	**8.62**	2015	2018	
5. Cohen J, 2013, STATISTICAL POWER ANALYSIS FOR THE BEHAVIORAL SCIENCES, VO, PO	2013	**8.62**	2015	2018	
6. Molina KI, 2014, J NEUROENG REHABIL, VI1, P0, DOI 10.1186/1743-0003-11-156	2014	**9.57**	2016	2019	
7. Allain P, 2014, J INT NEUROPSYCH SOC, V20, P468, DOI 10.1017/S1355617714000344	2014	**8.78**	2017	2019	
8. Skjeret N,2016, INT J MED INFORM, V85, P1, DOI 10.1016/j.ijmedinf.2015.10.008	2016	**10.36**	2018	2021	
9. Manera V, 2016, PLOS ONE, V11, P0, DOI 10.1371/journal.pone.0151487	2016	**9.99**	2019	2021	
10. Page MJ, 2021.BMJ-BRIT MED J, V372, P0, DOI 10.1136/bmj.n160	2021	**11.33**	2022	2025	

XSL•FO
RenderX

Table . The top 10 most-cited articles on virtual reality–based interventions in older adults, ranked by citation frequency.

TC[a]	Title	Authors	Year	Journal	DOI
63	Effects of virtual reality-based physical and cognitive training on executive function and dual-task gait performance in older adults with mild cognitive impairment: a randomized control trial [22]	Liao et al	2019	*Frontiers in Aging Neuroscience*	10.3389/fnagi.2019.00162
63	Using virtual reality-based training to improve cognitive function, instrumental activities of daily living and neural efficiency in older adults with mild cognitive impairment [23]	Liao et al	2020	*European Journal of Physical and Rehabilitation Medicine*	10.23736/S1973-9087.19.05899-4
60	Older adults with cognitive and/or physical impairments can benefit from immersive virtual reality experiences: a feasibility study [24]	Appel et al	2020	*Frontiers in Medicine*	10.3389/fmed.2019.00329
57	The effect of a virtual reality-based intervention program on cognition in older adults with mild cognitive impairment: a randomized control trial [25]	Thapa et al	2020	*Journal of Clinical Medicine*	10.3390/jcm9051283
51	The effectiveness of virtual reality for people with mild cognitive impairment or dementia: a meta-analysis [26]	Kim et al	2019	*BMC Psychiatry*	10.1186/s12888-019-2180-x
47	Virtual reality exergames for improving older adults' cognition and depression: a systematic review and meta-analysis of randomized control trials [27]	Yen and Chiu	2021	*Journal of the American Medical Directors Association*	10.1016/j.jamda.2021.03.009
47	PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews [28]	Page et al	2021	*BMJ* (*British Medical Journal*)	10.1136/bmj.n160
46	Acceptance of immersive head-mounted virtual reality in older adults [29]	Huygelier et al	2019	*Scientific Reports*	10.1038/s41598-019-41200-6

XSL•FO
RenderX

TC[a]	Title	Authors	Year	Journal	DOI
45	A mini-review of virtual reality-based interventions to promote well-being for people living with dementia and mild cognitive impairment [30]	D'Cunha et al	2019	*Gerontology*	10.1159/000500040
43	The role of virtual reality in improving health outcomes for community-dwelling older adults: systematic review [31]	Dermody et al	2020	*Journal of Medical Internet Research*	10.2196/17331

[a]TC: total number of citations.

In network analysis, betweenness centrality reflects the importance of nodes within a network; higher values indicate greater influence [32]. Figure 5B presents the top 10 publications ranked by betweenness centrality. A visual analysis, with pink nodes highlighting those with high betweenness centrality, reveals key structural influencers in the citation network. The study by Barcelos et al [33], published in the *Journal of the International Neuropsychological Society*, exhibited the highest betweenness centrality (0.25), followed by Brimelow et al in *Cyberpsychology, Behavior, and Social Networking* [34] (0.12) and Liao et al in *Frontiers in Aging Neuroscience* [22] (0.11). These publications serve as central hubs for knowledge dissemination in VR research for older adults, helping to identify and anticipate key developments and emerging trends in the field.

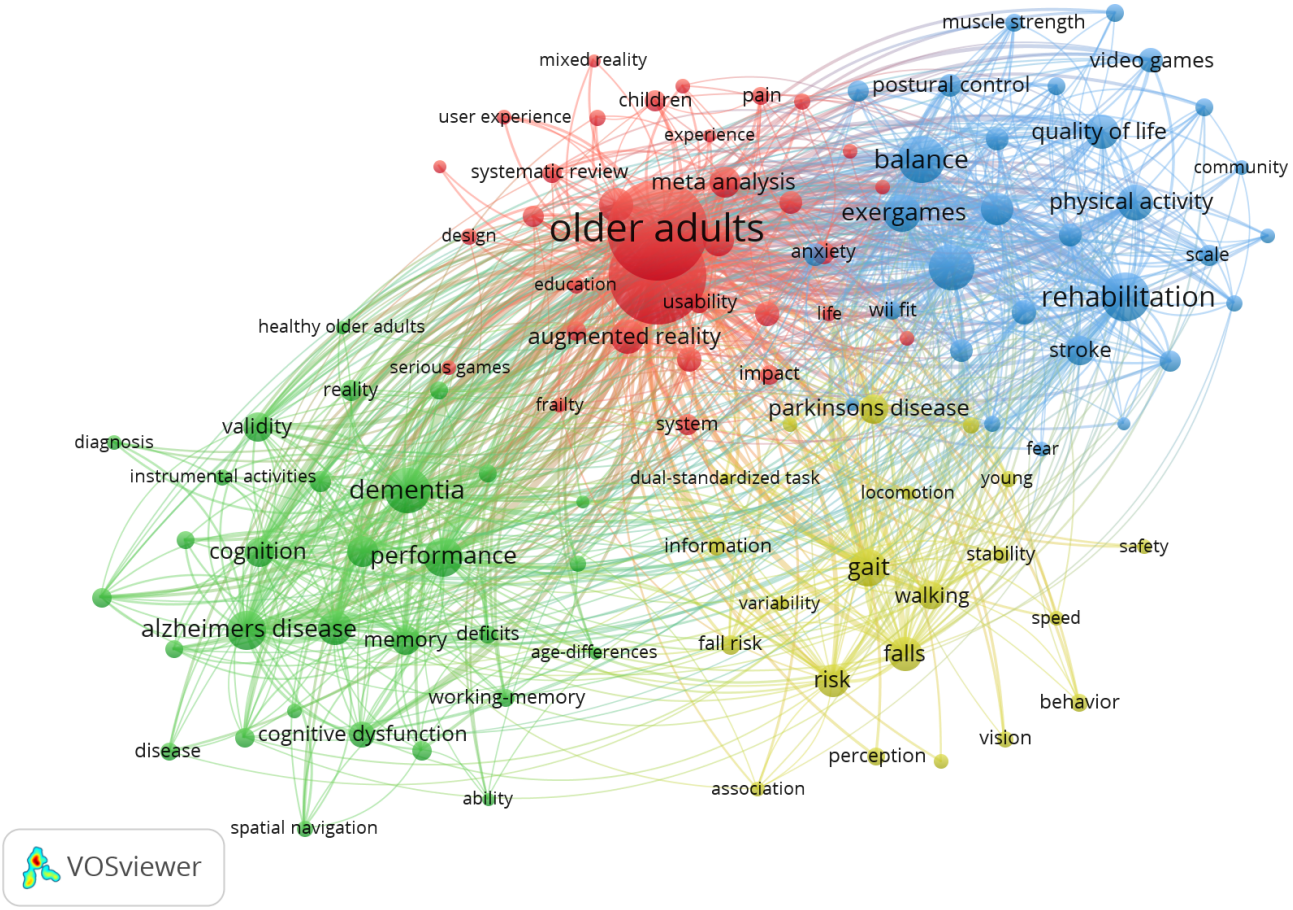
Additionally, we analyzed citation burst strength and identified the top 10 articles with the strongest bursts (Figure 5C). The study by Rendon et al [35], published in *Age and Aging*, exhibited the highest burst strength (11.64). The article by Page et al [28], published in *BMJ*, also showed a strong citation burst (11.33), reflecting rapid academic attention.

Keyword Visualization and Cluster Analysis

Through a systematic keyword analysis of 1609 publications using VOSviewer and CiteSpace, 4 key research clusters were identified in the application of VR for older adults (Figure 6). Cluster 1 focuses on meta-analysis, emphasizing methodological synthesis and evidence-based approaches. Cluster 2 is characterized by terms such as rehabilitation, balance, and exergames, highlighting the role of VR in improving physical function and mobility. Cluster 3 includes keywords like dementia, Alzheimer disease, and cognition, indicating the potential of VR in the diagnosis and management of neurodegenerative conditions. Cluster 4, marked by terms such as gait, falls, and stability, addresses the utility of VR in fall prevention and gait rehabilitation. Together, these clusters outline the current research landscape of VR applications for older adults.

XSL•FO
RenderX

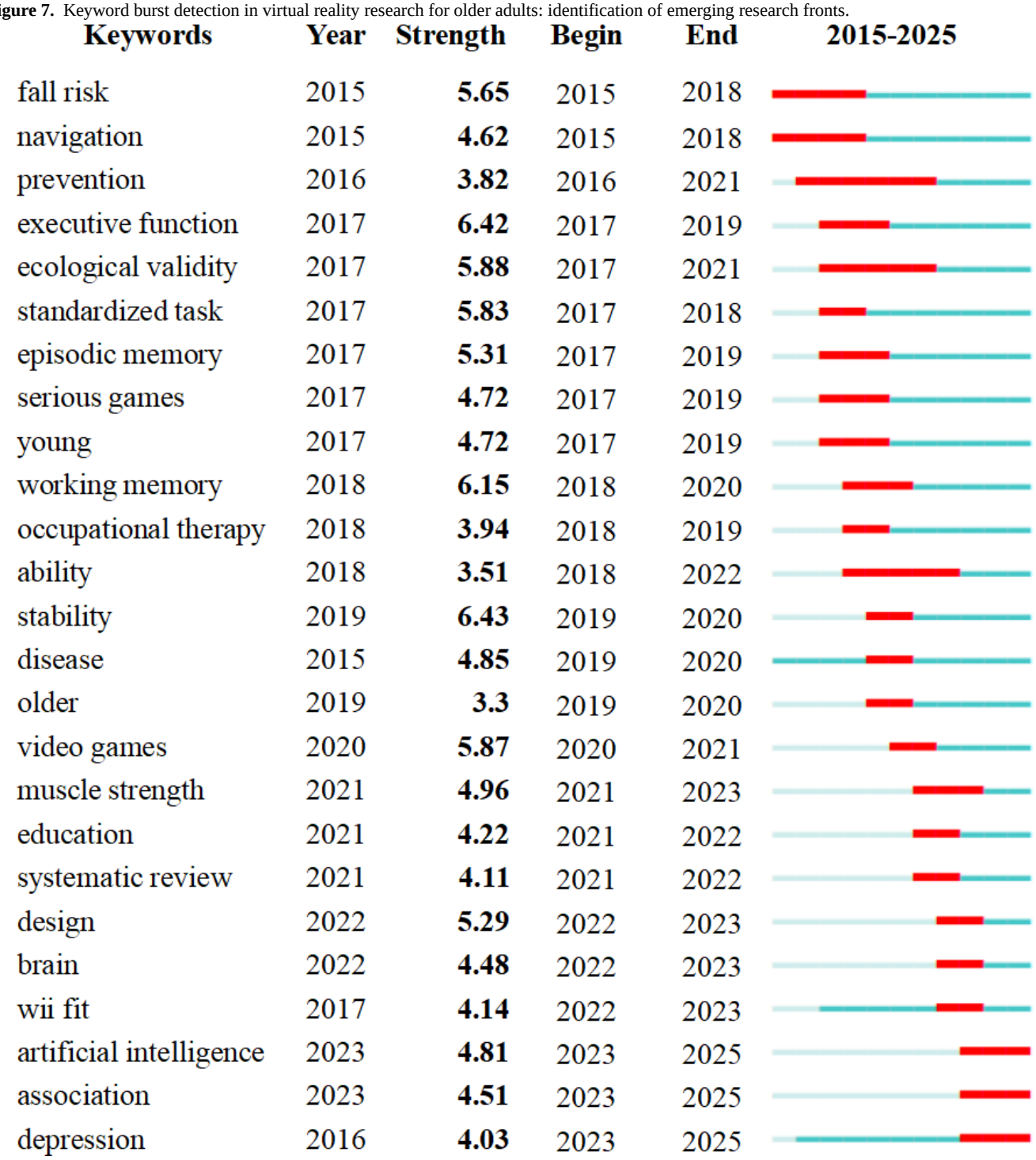
Figure 6. Keyword co-occurrence network analysis (2015 - 2025) of 1609 articles on virtual reality for older adults.



The identification of keyword bursts, which refer to sudden increases in citation frequency over short periods, serves as a critical indicator of emerging research frontiers [36]. Figure 7 presents the 25 most prominent keyword bursts from 2015 to 2025, with thick red bars indicating periods of heightened activity. The term “stability” exhibited the strongest burst intensity (6.43), followed closely by “executive function” (6.42). Notably, “prevention” showed the longest sustained burst duration (6 y). Recently emerged keywords include “artificial intelligence,” “association,” and “depression,” collectively reflecting current research trajectories that may shape future investigative priorities in this field.

XSL•FO
RenderX

Figure 7. Keyword burst detection in virtual reality research for older adults: identification of emerging research fronts.



Discussion

Principal Findings

The accelerating global demographic shift toward population aging has elevated geriatric health to a pivotal concern at the intersection of public health and technological innovation. Our bibliometric analysis of 1609 scholarly publications revealed that research on VR applications for older adults has evolved from isolated intervention trials into a well-defined, globally collaborative, and rapidly expanding field characterized by sustained growth and increasing interdisciplinary integration. Persistent geographical disparities in research advancement call for transnational alliances and structured knowledge exchange, strengthened collaboration between institutions in developing and advanced countries, and equitable co-development of VR technologies for aging.

We conducted a cocitation analysis to map the knowledge structure and thematic evolution of the field. The results revealed a dual emphasis on cognitive and emotional health in VR applications for older adults. Liao et al [22] demonstrated in a randomized controlled trial that VR improves executive function and dual-task gait in individuals with mild cognitive impairment, highlighting its potential for cognitive rehabilitation. Brimelow et al [34] used a mixed methods design in nursing homes and

XSL•FO
RenderX

found that immersive VR significantly reduces apathy, underscoring its value as a psychosocial intervention.

The keyword co-occurrence network revealed 4 distinct clusters, reflecting the evolving research landscape of VR applications for older adults. Notably, the red cluster captured a pivotal shift in the field—from initial exploration of clinical feasibility to a growing emphasis on methodological rigor and evidence synthesis.

This transition was further corroborated by citation burst analysis of references, which showed that early research was dominated by studies on VR's efficacy in balance improvement and fall prevention [35], reflecting a formative phase focused on proof-of-concept and immediate outcomes. In contrast, recent citation surges highlighted methodological frameworks and reporting standards—particularly PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [28]—indicating a maturation toward transparent, reproducible, and evidence-informed research practices.

Systematic reviews provide comprehensive and transparent evaluations of existing literature, supporting clinical decision-making and guiding future research directions. For instance, a recent systematic review of commercially available VR applications found that VR interventions hold potential for improving health outcomes in older adults [31]. However, the studies included in such reviews generally did not reach high certainty or quality according to the GRADE (Grading of Recommendations Assessment, Development and Evaluation) framework, and authors emphasized the need for more rigorous scientific methodologies to robustly evaluate and validate VR technologies.

The blue cluster centered on rehabilitation, balance, and exergames, reflecting VR's role in addressing physical and cognitive decline in older adults. Exergames—VR systems combining exercise with gamified feedback—have been integrated into cognitive training with promising engagement [37,38]. Yet, a systematic review found limited high-quality evidence of sustained benefits in long-term care settings [39], underscoring the need for larger, rigorous trials that consider cost-effectiveness and user-centered design.

The green cluster focused on cognitive decline and dementia, highlighting an emerging trend in the application of VR for neurocognitive rehabilitation. VR technology offers immersive, simulated environments that can replicate real-life scenarios, providing patients with safe, controlled, and repeatable opportunities for functional training. A recent meta-analysis further confirmed the dual benefits of VR: not only does it enhance cognitive processing and sensorimotor coordination, but it also positively promotes mental health by increasing user engagement and autonomy [40].

The yellow cluster underscored VR's role in enhancing mobility and preventing falls in older adults. VR-based interventions simulate dynamic environments, enabling safe and controlled training of adaptive motor responses. Evidence shows these immersive experiences improve gait and dynamic balance in older adults with Parkinson disease through engagement in rich, complex, and personalized virtual scenarios [41].

Research Hot Spots and Frontiers

Burst keyword analysis in this study revealed that the first prominent burst was associated with AI, which demonstrated a strong growth trend from 2023 to 2025. This indicates that artificial intelligence (AI) is emerging as an increasingly central driver in aging-related research. AI encompasses core technologies such as machine learning, natural language processing, and computer vision, all of which show significant potential in addressing unmet needs among older adult populations.

For instance, innovative systems integrating AI and VR have been developed to create immersive social virtual environments that simulate interpersonal interaction scenarios, aiming to improve emotional well-being and cognitive function in older adults [42]. Additionally, AI-powered transformations of hospital waiting rooms—enabled by the integration of Internet of Things and AI technologies—are being explored to enhance health care efficiency, optimize patient experience, and support early disease detection [43]. Recently, a study introduced an adaptive AI-driven content generation system based on user feedback, capable of dynamically adjusting the difficulty and thematic content of cognitive training tasks. This system delivers personalized and engaging experiences for older users, effectively slowing cognitive decline and significantly improving quality of life, emotional stability, and subjective well-being [44].

In summary, future research should further advance the embedded application of AI within VR environments, with particular emphasis on two critical directions [3]: (1) ensuring safety and privacy protection mechanisms for AI systems deployed in home settings and (2) bridging the "digital divide" by addressing older adults' technological barriers and trust concerns in using AI-driven tools, thereby promoting equitable access to digital health innovations.

The keyword "association" showed a significant increase during 2023-2025, reflecting growing scholarly attention to the relationships between VR interventions and health outcomes in older adults. This trend indicates a shift in research focus—from primarily validating the efficacy of interventions toward exploring underlying influencing factors and mechanistic pathways. This transition is empirically supported by Oliveira et al [45], who developed the ECO-VR multitask assessment system. Their findings revealed that task performance was significantly associated with age, education level, Mini-Mental State Examination scores, and verbal fluency, with age emerging as the strongest predictor. This underscores the critical role of demographic variables in shaping VR-based outcomes. The study further confirms that VR can serve as an ecologically valid tool for assessing multidimensional cognitive associations in aging populations.

This growing emphasis on association aligns with the broader trend in digital health toward personalization and precision medicine. As AI and machine learning are increasingly integrated into VR systems, it becomes feasible to capture user behavioral data in real time and identify dynamic associations with cognitive decline or fall risk. Predictive models derived from large-scale data analysis may uncover causal pathways

XSL•FO
RenderX

between clinical indicators, enabling early risk detection and timely intervention.

The third research hot spot is depression, characterized by persistent sadness and loss of interest in daily activities, necessitating early detection for effective treatment and intervention [46]. It significantly impairs quality of life, physical health, and cognitive function, while elevating risks of social isolation, suicidal ideation, and health care utilization [47]. This burden is particularly pronounced among long-term care residents, where environmental constraints, loneliness, and limited social engagement act as key contributing factors [48]. In this context, VR technology has emerged as a promising nonpharmacological intervention [27,49,50].

Challenges to Clinical Translation

Despite growing academic interest in VR for older adults, significant translational barriers persist in implementing VR within community and home-based care settings. These barriers operate across 3 interconnected domains. First, the absence of a standardized scientific framework for developing and evaluating VR-based interventions impedes evidence-based implementation. Second, high equipment costs and low digital literacy limit VR adoption among older adults [51], while implementation costs often exceed marginal health benefits [52]. Third, commercial VR platforms often feature complex interfaces and immersive environments, which may pose potential safety risks.

First, establishing a standardized scientific framework for VR-based therapy development and evaluation [53] is essential. Concurrently, deploying voice-guided VR interfaces on low-cost or refurbished devices could mitigate financial and usability barriers. Technological innovation must be integrated with real-world economic constraints, ensuring designs align with older adults' needs and preferences—empowering them to independently engage with technology [54]. Alternatively, family-mediated device guidance may enhance user experience and reduce adoption barriers.

Limitations and Future Directions

Our study has several limitations that should be acknowledged. First, literature retrieval based on topic fields relies on partial keyword matching, which may retrieve records with weak relevance to the research topic, thereby introducing noise and compromising retrieval accuracy. Second, the restriction to English-language publications and specific document types may have introduced selection bias and limited the representativeness of the findings. Third, the reliance on a single database and a fixed time window constrains the comprehensiveness and generalizability of the results. Finally, automatic term extraction in CiteSpace and VOSviewer may have resulted in synonym redundancy or conceptual overlap, potentially affecting the accuracy of topic clustering. Future studies should refine search terms, use multidatabase sourcing, and validate keywords to enhance reliability.

Conclusions

Addressing health challenges associated with population aging has become a critical component of global health strategies. Although VR technologies demonstrate considerable promise for geriatric applications, rigorous empirical investigations remain essential to establish the efficacy and feasibility of various intervention approaches. Future studies should prioritize the development of more targeted VR applications while systematically evaluating their adaptability across diverse older adult populations. Furthermore, enhanced international collaboration among scholars and institutions is warranted to accelerate the development and implementation of VR-AI integrated technologies, collectively addressing global challenges posed by population aging.

Acknowledgments

After the paper was written, ChatGPT was used to correct and polish the grammar of the manuscript. Following the use of this tool, the authors reviewed and edited the content as necessary, and the authors take full responsibility for the content of the original manuscript.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: JX and WZ. Data curation: JX, WZ, and WL. Data analysis: JX and WL. Funding acquisition: WZ Visualization: JX, WZ, and WL. Writing—original draft: JX. Writing—review and editing: JX, WZ, and WL.

Conflicts of Interest

None declared.

Checklist 1

XSL•FO
RenderX

The BIBLIO checklist for reporting bibliometric reviews of the biomedical literature.
[PDF File, 116 KB - aging_v8i1e76609_app1.pdf]

References

1. Mental health of older adults. World Health Organization. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/mental-health-of-older-adults [accessed 2025-12-10]
2. World population prospects 2024. United Nations Department of Economic and Social Affairs, Population Division. URL: https://population.un.org/wpp/downloads [accessed 2025-10-12]
3. Wang J, Liang Y, Cao S, Cai P, Fan Y. Application of artificial intelligence in geriatric care: bibliometric analysis. J Med Internet Res 2023 Jun 23;25:e46014. [doi: 10.2196/46014] [Medline: 37351923]
4. Li YJ, Wilke C, Shiyanov I, Muschalla B. Impact of virtual reality-based group activities on activity level and well-being among older adults in nursing homes: longitudinal exploratory study. JMIR Serious Games 2024 Mar 29;12:e50796. [doi: 10.2196/50796] [Medline: 38551635]
5. Lu Z, Wang W, Yan W, Kew CL, Seo JH, Ory M. The application of fully immersive virtual reality on reminiscence interventions for older adults: scoping review. JMIR Serious Games 2023 Oct 6;11:e45539. [doi: 10.2196/45539] [Medline: 37801360]
6. Everard G, Declerck L, Lejeune T, et al. A self-adaptive serious game to improve motor learning among older adults in immersive virtual reality: short-term longitudinal pre-post study on retention and transfer. JMIR Aging 2025 Mar 3;8:e64004. [doi: 10.2196/64004] [Medline: 40053708]
7. Mao Q, Zhao Z, Yu LS, Zhao Y, Wang HL. The effects of virtual reality-based reminiscence therapies for older adults with cognitive impairment: systematic review. J Med Internet Res 2024 Nov 12;26:e53348. [doi: 10.2196/53348] [Medline: 39531267]
8. Raffegeau TE, Young WR, Fino PC, Williams AM. A perspective on using virtual reality to incorporate the affective context of everyday falls into fall prevention. JMIR Aging 2023 Jan 11;6:e36325. [doi: 10.2196/36325] [Medline: 36630173]
9. Jin S, Tan Z, Liu T, et al. Preference of virtual reality games in psychological pressure and depression treatment: discrete choice experiment. JMIR Serious Games 2023 Jan 16;11:e34586. [doi: 10.2196/34586] [Medline: 36645698]
10. Shu S, Woo BKP. Pioneering the Metaverse: the role of the metaverse in an aging population. JMIR Aging 2023 Jan 20;6:e40582. [doi: 10.2196/40582] [Medline: 36662547]
11. Skurla MD, Rahman AT, Salcone S, et al. Virtual reality and mental health in older adults: a systematic review. Int Psychogeriatr 2022 Feb;34(2):143-155. [doi: 10.1017/S104161022100017X] [Medline: 33757619]
12. Yu D, Li X, Lai FHY. The effect of virtual reality on executive function in older adults with mild cognitive impairment: a systematic review and meta-analysis. Aging Ment Health 2023 Apr;27(4):663-673. [doi: 10.1080/13607863.2022.2076202] [Medline: 35635486]
13. He D, Cao S, Le Y, Wang M, Chen Y, Qian B. Virtual reality technology in cognitive rehabilitation application: bibliometric analysis. JMIR Serious Games 2022 Oct 19;10(4):e38315. [doi: 10.2196/38315] [Medline: 36260388]
14. Fan T, Wang X, Song X, Zhao G, Zhang Z. Research status and emerging trends in virtual reality rehabilitation: bibliometric and knowledge graph study. JMIR Serious Games 2023 Mar 6;11:e41091. [doi: 10.2196/41091] [Medline: 36877556]
15. Ho KY, Cheung PM, Cheng TW, Suen WY, Ho HY, Cheung DSK. Virtual reality intervention for managing apathy in people with cognitive impairment: systematic review. JMIR Aging 2022 May 11;5(2):e35224. [doi: 10.2196/35224] [Medline: 35544317]
16. Kwan RYC, Liu J, Sin OSK, et al. Effects of virtual reality motor-cognitive training for older people with cognitive frailty: multicentered randomized controlled trial. J Med Internet Res 2024 Sep 11;26:e57809. [doi: 10.2196/57809] [Medline: 39259959]
17. Han Q, Shi J, Liu J, et al. Decoding the research landscape of drug resistance and therapeutic approaches in head and neck cancer: a bibliometric analysis from 2000 to 2023. Front Pharmacol 2024;15:1375110. [doi: 10.3389/fphar.2024.1375110] [Medline: 38645557]
18. Fu Y, Han Q, Wang F, Dong XY. Bibliometric analysis of youth myocardial infarction research (1980-2023). Front Cardiovasc Med 2024;11:1478158. [doi: 10.3389/fcvm.2024.1478158] [Medline: 39660115]
19. Liao H, Tang M, Luo L, Li C, Chiclana F, Zeng XJ. A bibliometric analysis and visualization of medical big data research. Sustainability 2018;10(1):166. [doi: 10.3390/su10010166]
20. Guler AT, Waaijer CJF, Palmblad M. Scientific workflows for bibliometrics. Scientometrics 2016;107(2):385-398. [doi: 10.1007/s11192-016-1885-6] [Medline: 27122644]
21. Donath L, Rössler R, Faude O. Effects of virtual reality training (exergaming) compared to alternative exercise training and passive control on standing balance and functional mobility in healthy community-dwelling seniors: a meta-analytical review. Sports Med 2016 Sep;46(9):1293-1309. [doi: 10.1007/s40279-016-0485-1] [Medline: 26886474]
22. Liao YY, Chen IH, Lin YJ, Chen Y, Hsu WC. Effects of virtual reality-based physical and cognitive training on executive function and dual-task gait performance in older adults with mild cognitive impairment: a randomized control trial. Front Aging Neurosci 2019;11:162. [doi: 10.3389/fnagi.2019.00162] [Medline: 31379553]

XSL•FO
RenderX

23. Liao YY, Tseng HY, Lin YJ, Wang CJ, Hsu WC. Using virtual reality-based training to improve cognitive function, instrumental activities of daily living and neural efficiency in older adults with mild cognitive impairment. Eur J Phys Rehabil Med 2020 Feb;56(1):47-57. [doi: 10.23736/S1973-9087.19.05899-4] [Medline: 31615196]
24. Appel L, Appel E, Bogler O, et al. Older adults with cognitive and/or physical impairments can benefit from immersive virtual reality experiences: a feasibility study. Front Med (Lausanne) 2019;6:329. [doi: 10.3389/fmed.2019.00329] [Medline: 32010701]
25. Thapa N, Park HJ, Yang JG, et al. The effect of a virtual reality-based intervention program on cognition in older adults with mild cognitive impairment: a randomized control trial. J Clin Med 2020 Apr 29;9(5):1283. [doi: 10.3390/jcm9051283] [Medline: 32365533]
26. Kim O, Pang Y, Kim JH. The effectiveness of virtual reality for people with mild cognitive impairment or dementia: a meta-analysis. BMC Psychiatry 2019 Jul 12;19(1):219. [doi: 10.1186/s12888-019-2180-x] [Medline: 31299921]
27. Yen HY, Chiu HL. Virtual reality exergames for improving older adults' cognition and depression: a systematic review and meta-analysis of randomized control trials. J Am Med Dir Assoc 2021 May;22(5):995-1002. [doi: 10.1016/j.jamda.2021.03.009] [Medline: 33812843]
28. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ 2021 Mar 29;372:n160. [doi: 10.1136/bmj.n160] [Medline: 33781993]
29. Huygelier H, Schraepen B, van Ee R, Vanden Abeele V, Gillebert CR. Acceptance of immersive head-mounted virtual reality in older adults. Sci Rep 2019 Mar 14;9(1):4519. [doi: 10.1038/s41598-019-41200-6] [Medline: 30872760]
30. D'Cunha NM, Nguyen D, Naumovski N, et al. A mini-review of virtual reality-based interventions to promote well-being for people living with dementia and mild cognitive impairment. Gerontology 2019;65(4):430-440. [doi: 10.1159/000500040] [Medline: 31108489]
31. Dermody G, Whitehead L, Wilson G, Glass C. The role of virtual reality in improving health outcomes for community-dwelling older adults: systematic review. J Med Internet Res 2020 Jun 1;22(6):e17331. [doi: 10.2196/17331] [Medline: 32478662]
32. Synnestvedt MB, Chen C, Holmes JH. CiteSpace II: visualization and knowledge discovery in bibliographic databases. AMIA Annu Symp Proc 2005;2005:724-728. [Medline: 16779135]
33. Barcelos N, Shah N, Cohen K, et al. Aerobic and Cognitive Exercise (ACE) pilot study for older adults: executive function improves with cognitive challenge while exergaming. J Int Neuropsychol Soc 2015 Nov;21(10):768-779. [doi: 10.1017/S1355617715001083] [Medline: 26581789]
34. Brimelow RE, Dawe B, Dissanayaka N. Preliminary research: virtual reality in residential aged care to reduce apathy and improve mood. Cyberpsychol Behav Soc Netw 2020 Mar;23(3):165-170. [doi: 10.1089/cyber.2019.0286] [Medline: 31829729]
35. Rendon AA, Lohman EB, Thorpe D, Johnson EG, Medina E, Bradley B. The effect of virtual reality gaming on dynamic balance in older adults. Age Ageing 2012 Jul;41(4):549-552. [doi: 10.1093/ageing/afs053] [Medline: 22672915]
36. Chen C. CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature. J Am Soc Inf Sci 2006 Feb;57(3):359-377. [doi: 10.1002/asi.20317]
37. Anderson-Hanley C, Arciero PJ, Brickman AM, et al. Exergaming and older adult cognition: a cluster randomized clinical trial. Am J Prev Med 2012 Feb;42(2):109-119. [doi: 10.1016/j.amepre.2011.10.016] [Medline: 22261206]
38. Huang KT. Exergaming executive functions: an immersive virtual reality-based cognitive training for adults aged 50 and older. Cyberpsychol Behav Soc Netw 2020 Mar;23(3):143-149. [doi: 10.1089/cyber.2019.0269] [Medline: 31794673]
39. Chu CH, Quan AML, Souter A, Krisnagopal A, Biss RK. Effects of exergaming on physical and cognitive outcomes of older adults living in long-term care homes: a systematic review. Gerontology 2022;68(9):1044-1060. [doi: 10.1159/000521832] [Medline: 35290983]
40. Stavropoulou I, Sakellari E, Barbouni A, Notara V. Community-based virtual reality interventions in older adults with dementia and/or cognitive impairment: a systematic review. Exp Aging Res 2025;51(2):162-189. [doi: 10.1080/0361073X.2024.2377438] [Medline: 38972047]
41. Canning CG, Allen NE, Nackaerts E, Paul SS, Nieuwboer A, Gilat M. Virtual reality in research and rehabilitation of gait and balance in Parkinson disease. Nat Rev Neurol 2020 Aug;16(8):409-425. [doi: 10.1038/s41582-020-0370-2] [Medline: 32591756]
42. Kosti MV, Benayoun M, Georgakopoulou N, et al. Connecting the elderly using VR: a novel art-driven methodology. Appl Sci (Basel) ;14(5):2217. [doi: 10.3390/app14052217]
43. Spoladore D, Mondellini M, Mahroo A, et al. Smart waiting room: a systematic literature review and a proposal. Electronics (Basel) 2024;13(2):388. [doi: 10.3390/electronics13020388]
44. Li HH, Liao YH. Application and effectiveness of adaptive AI in elderly healthcare. Psychogeriatrics 2025 Jan;25(1):e13214. [doi: 10.1111/psyg.13214] [Medline: 39537559]
45. Oliveira CR, Lopes Filho BJP, Esteves CS, et al. Neuropsychological assessment of older adults with virtual reality: association of age, schooling, and general cognitive status. Front Psychol 2018;9:1085. [doi: 10.3389/fpsyg.2018.01085] [Medline: 30008689]

XSL•FO
RenderX

46. Nejadshamsi S, Karami V, Ghourchian N, et al. Development and feasibility study of HOPE model for prediction of depression among older adults using Wi-Fi-based motion sensor data: machine learning study. JMIR Aging 2025 Mar 3;8:e67715. [doi: 10.2196/67715] [Medline: 40053734]
47. Qiu T, Zhang G, Zhou F, Jiang H. Application of virtual reality to enhance therapeutic Tai Chi for depression in elderly people. Acta Psychol (Amst) 2024 Aug;248:104316. [doi: 10.1016/j.actpsy.2024.104316] [Medline: 39025032]
48. Tan JDL, Maneze D, Montayre J, Ramjan LM, Wang D, Salamonson Y. Family visits and depression among residential aged care residents: an integrative review. Int J Nurs Stud 2023 Oct;146:104568. [doi: 10.1016/j.ijnurstu.2023.104568] [Medline: 37597458]
49. Buele J, Avilés-Castillo F, Del-Valle-Soto C, Varela-Aldás J, Palacios-Navarro G. Correction: effects of a dual intervention (motor and virtual reality-based cognitive) on cognition in patients with mild cognitive impairment: a single-blind, randomized controlled trial. J NeuroEngineering Rehabil 2024;21(1):149. [doi: 10.1186/s12984-024-01443-5]
50. Zhu KY, Zhang QY, He BW, Huang MZ, Lin R, Li H. Immersive virtual reality-based cognitive intervention for the improvement of cognitive function, depression, and perceived stress in older adults with mild cognitive impairment and mild dementia: pilot pre-post study. JMIR Serious Games 2022 Feb 21;10(1):e32117. [doi: 10.2196/32117] [Medline: 35188466]
51. Iqbal AI, Aamir A, Hammad A, et al. Immersive technologies in healthcare: an in-depth exploration of virtual reality and augmented reality in enhancing patient care, medical education, and training paradigms. J Prim Care Community Health 2024;15:21501319241293311. [doi: 10.1177/21501319241293311] [Medline: 39439304]
52. Wilding R, Barbosa Neves B, Waycott J, et al. Introducing virtual reality to older adults: a qualitative analysis of a co-design innovation with care staff. Arch Gerontol Geriatr 2024 Oct;125:105505. [doi: 10.1016/j.archger.2024.105505] [Medline: 38851090]
53. Birckhead B, Khalil C, Liu X, et al. Recommendations for methodology of virtual reality clinical trials in health care by an international working group: iterative study. JMIR Ment Health 2019 Jan 31;6(1):e11973. [doi: 10.2196/11973] [Medline: 30702436]
54. Koh WQ, Ludlow K, Liddle J, Pachana NA. Selection, optimization, and compensation strategies used by older adults to live well with technology: qualitative study. JMIR Aging 2025 Sep 19;8:e75019. [doi: 10.2196/75019] [Medline: 40971580]

Abbreviations

AI: artificial intelligence
BIBLIO: Guideline for Reporting Bibliometric Reviews of the Biomedical Literature
GRADE: Grading of Recommendations Assessment, Development and Evaluation
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
VR: virtual reality
WoSCC: Web of Science Core Collection

Edited by A Quialheiro; submitted 27.04.25; peer-reviewed by A Abdullah, J Xun, M Tomita, T Tai, X Luo, Y Fu, Z Zhang; accepted 09.11.25; published 02.12.25.

Please cite as:
Xu J, Zhang W, Liu W
Research Status and Trends in Virtual Reality Technology for Older Adults: Bibliometric and Visual Analysis
JMIR Aging 2025;8:e76609
URL: https://aging.jmir.org/2025/1/e76609
doi:10.2196/76609

© Jing Xu, Wenjin Zhang, WenLi Liu. Originally published in JMIR Aging (https://aging.jmir.org), 2.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Double-Edged Sword of Digital Engagement—How Digital Access and Internet Use Reshape Sleep Schedules and Underlying Mechanisms in Older Adults: Longitudinal Observational Study

Chi Zhang[1], DM; Longxuan Lin[1], MA; Li Wang[2], MSW; Han Hu[3], DM; Heyang Zhang[1], BA

[1]Lanzhou University, No. 222, Tianshui South Road, Chengguan District, Lanzhou, Gansu Province, China

[2]School of Philosophy and Sociology, Northwest Normal University, Lanzhou, China

[3]School of Public Policy and Administration, Xi'an Jiaotong University, Xi'an, China

Corresponding Author:
Chi Zhang, DM
Lanzhou University, No. 222, Tianshui South Road, Chengguan District, Lanzhou, Gansu Province, China

Abstract

Background: Given the rapid development of the digital economy and the sustained proliferation of the internet, digital engagement in older adults has garnered mounting attention from the academic community. However, research has yet to systematically examine the impact of digital engagement on sleep in this demographic.

Objective: This study aims to examine the association of digital engagement—operationalized as digital access and internet use duration—with the sleep schedules (nocturnal sleep duration, afternoon nap duration, and sleep onset time) of older adults in China, using longitudinal data and robust statistical modeling to explore longitudinal associations and potential mechanisms.

Methods: Data were derived from 4 waves (2014, 2016, 2018, and 2020) of the China Family Panel Studies, involving 16,784 older adults (≥60 y). We used panel fixed effects models and a random-effects ordered logit model to analyze the effects on continuous outcomes (nocturnal and nap sleep duration), controlling for time-invariant individual characteristics. As sleep onset time is an ordinal variable, a random-effects ordered logit model was used for this outcome. Moderation analyses were conducted by introducing interaction terms (digital engagement×sex and digital engagement×residence) into the models to examine heterogeneity across subgroups (urban or rural, men or women). Mediation analyses were performed using the Sobel test with year-fixed effects and the nonparametric bootstrap method (1000 resamples) to assess the significance of indirect effects via mechanistic pathways (nonfarm employment, protein intake, memory, depressive mood, and instrumental activities of daily living).

Results: The study included a total of 16,784 older adults, with an average age of 69 (SE 6.946) years, including 9100 (54.22%) women and 7684 (45.78%) men. The results showed that both digital access (β=−.15, 95% CI −.25 to −.06; *P*=.002) and internet use time (β=−.07, 95% CI −.13 to −.01; *P*=.027) were significantly associated with significantly shorter sleep duration of older adults. Digital access was significantly associated with a significant reduction in the length of afternoon naps among older adults, while internet use did not have this effect; both digital access and internet use were significantly associated with a significant delay in older adults' sleep onset time. Digital access was associated with older adults' sleep schedules through its correlations with nonfarm employment, protein intake, memory, depressive mood, and instrumental activities of daily living. Digital access had a greater and more significant impact on men and urban older adults, while internet use had a greater and more significant impact on women and urban older adults.

Conclusions: The study indicates that digital engagement, such as the use of electronic devices, is associated with a reduction in both daily and nap sleep duration, as well as a delay in sleep onset, among older adults.

(JMIR Aging 2025;8:e79731) doi:10.2196/79731

KEYWORDS

digital health; digital engagement; sleep duration; afternoon nap; sleep onset time; mediation effect; cohort study

Introduction

The digital engagement of older adults has become a subject of increasing interest and concern within the academic community. This phenomenon is precipitated by the rapid development of the digital economy and the continuous popularization of the internet [1-3]. In Europe and the United States, the proportion of individuals aged ≥65 years who use the internet has increased

XSL•FO
RenderX

from less than 30% (the United States in 2004 and Europe in 2010) [4,5] in the early years to approximately 60% - 67% in recent years (the United States in 2016 and Europe in 2019) [5,6]. In a similar vein, forecasts suggest that the internet penetration rate among individuals aged ≥60 years in China will reach 51.07% in 2023 [7,8]. Despite the numerous benefits of internet use for older adults [9-11], concerns have been raised regarding its impact on sleep. Specifically, internet use has been shown to potentially disrupt the duration and quality of sleep by occupying sleep time [12], interfering with physiological rhythms [13], increasing brain excitation [14], and interfering with the sleep environment, among other factors. The impact of digital engagement on the sleep of older adults warrants further in-depth study.

Previous studies have shown that sleep is influenced by multiple factors including physiology [15,16], psychology [17,18], environment [19-21], and behavior [22-24], but this study focuses on the digital engagement factors among them. The digital engagement is significantly associated with 4 primary factors that influence sleep in older individuals: the use of LED-backlit screens on electronic devices has been demonstrated to emit high-intensity, short-wavelength blue light, which has been shown to inhibit melatonin secretion [25,26]. A growing body of research has indicated that online social media use can reduce feelings of loneliness and depression in older adults [27-29]. In addition, digital engagement has been found to disrupt sleep environments, such as the intrusion of electronic devices into bedrooms [30]. A sedentary lifestyle and lack of exercise due to internet access have also been shown to be detrimental to healthy sleep. The paucity of physical activity and sedentary lifestyle that is concomitant with Internet access has also been demonstrated to be deleterious to the quality of sleep in older adults [31]. However, there is a paucity of research on the impact of digital engagement on the sleep of older people. Digital delivery of cognitive behavioral therapies has been shown to have a positive effect on insomnia symptoms in older adults, with studies indicating a reduction in sleep effort [32,33]. In addition, community-based digital mental health monitoring platforms have the potential to provide real-time tracking of sleep fragmentation, efficiency, and other indicators, thereby assisting older adults in optimizing their sleep through a feedback mechanism [34]. However, it is important to note that nighttime use of digital screens by older adults may potentially exacerbate sleep disorders due to the suppression of melatonin by blue light [35]. A paucity of studies exists on the effects of digital engagement on older adults' sleep and the mechanisms of its influence. Furthermore, these studies are unsystematic. However, given the increasing prevalence of digital engagement among older adults, it is imperative to investigate its association with sleep quality.

In summary, given the limitations of extant studies, this study endeavors to use the 4-period tracking data from the China Family Tracking Survey from 2014 to 2020. The study uses panel fixed effects (FE) models and a random-effects ordered logit model to investigate the impact of digital engagement (digital access and internet duration) on the sleep schedules of Chinese older adults aged ≥60 years. The dependent variables (sleep schedules) include sleep duration, afternoon nap duration, and nocturnal sleep onset time. The study also explores the underlying mechanisms by which these effects occur.

Methods

Design and Study Population

This study is based on data from the China Family Panel Studies (CFPS), a nationally representative, longitudinal survey launched in 2010 by the Institute of Social Science Survey at Peking University. The CFPS adopts a multistage probability proportional to size sampling strategy, covering 25 provinces, municipalities, and autonomous regions in China, which represent about 95% of the national population. The survey collects extensive information at the individual, household, and community levels, including socioeconomic status, health, education, family structure, and digital engagement. It is conducted biennially, thereby enabling longitudinal analysis of social, economic, and health transitions among Chinese residents. The CFPS is widely regarded as one of the most authoritative and comprehensive datasets for social science and public health research in China, and it has been extensively used in previous studies to examine issues related to aging, health, and digital behavior. The database under consideration includes respondents' digital engagement and sleep schedules, which meet the needs of this study. The data used in this study to examine the impact of digital engagement on sleep schedules of older adults are drawn from 4 periods of CFPS in 2014, 2016, 2018, and 2020, comprising a total of 142,544 samples. This cohort study was conducted in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guidelines (Checklist 1) [36].

Ethical Considerations

The protocols followed in the CFPS were aligned with the principles of the Declaration of Helsinki [37]. This research was reviewed and approved by the Beijing University Biomedical Ethics Committee (IRB00001052-14010). All participants in the CFPS provided written informed consent. The participants received a physical examination report for their participation. The data were deidentified and treated for confidentiality by anonymization and coding.

Sleep Schedule

The explained variable of this study is the sleep schedule of older adults, which is systematically operationalized as sleep duration, afternoon nap duration, and sleep onset time. The prevalence of regular napping among older adults has been well-documented [38-40]. However, the incorporation of this behavior into the broader context of sleep scheduling has been a notable oversight in previous research. The fixed-distance variables in hours include sleep duration and nap duration. Sleep onset time is an ordinal variable, combining the data at midnight to midnight the following day, and assigning the data from 1 AM to 4 AM to 5 AM to 8 AM for the sake of analyzing convenience. Concurrently, the atypical samples characterized by nocturnal sleep duration between 5 AM and 6 PM were eliminated from the analysis (out of 16,784 samples, 358 samples were thus excluded). With regard to the treatment of

XSL•FO
RenderX

extant studies, sleep schedules were classified into 3 categories during the descriptive statistics stage [41].

Digital Engagement

The core explanatory variables of this study are the digital engagement of older adults. Informed by prior research and the data structure of the CFPS, we operationalized digital engagement through two distinct but complementary measures that allowed us to distinguish between the binary state of being connected and the intensity of use among connected individuals:

1. Digital access: this is a binary variable indicating whether an older adult has fundamental access to the digital world. It was constructed based on responses to the CFPS questionnaire regarding the use of digital devices to access the internet. Respondents who reported using either a mobile phone or a computer to go online were assigned a value of 1. Those who used neither of these devices for internet access were assigned a value of 0.
2. Online time: this continuous variable measures the intensity of digital engagement among those with access. It is based on the self-reported average number of hours per week spent on the internet. To address the substantial right-skewness in its distribution (where most users report low-to-moderate hours, and a few report very high hours), we applied a natural logarithmic transformation to this variable. This transformation improves the linearity of its relationship with the sleep outcome variables and mitigates the influence of extreme values, thereby yielding more robust and interpretable regression coefficients.

It is important to note the analytical implication of this 2-pronged approach: the variable online time was only analyzed for the subsample of older adults who had digital access (ie, where digital access=1). For respondents without digital access, data on online time were logically defined as missing and excluded from analyses involving this specific variable.

Covariates

The following covariates were used in this study: sex (assigned a value of 0 for females and 1 for males); marital status (assigned a value of 1 for those in marriage and 0 for those not in marriage); years of education; domicile (assigned a value of 0 for those in rural areas and 1 for those in urban areas); and age. The participants were asked to self-assess their health using a scale from 1 to 5, with 1=“very good health” and 5=“very unhealthy.” The higher the value assigned, the more unhealthy the participant’s health. The participants were also asked to indicate whether they had a religion (with the option of assigning a value of 1 for those who had a religion and 0 for those who did not) and chronic disease (presence of a chronic disease is assigned a value of 1, otherwise 0).

The selection of covariates was guided by their established associations with sleep patterns in the existing literature, aiming to minimize potential confounding bias. Sex, marital status, education, domicile (urban or rural), age, self-rated health, and chronic disease status are well-documented demographic and socioeconomic determinants of sleep. Although religious belief was reported by a small minority (542/16784, 3.23%) of our sample, it was retained as a covariate due to its theoretical relevance as a sociodemographic marker that can influence lifestyle, social support, and psychological well-being, all of which are linked to sleep quality and duration. We also considered including working status (eg, retired); however, this variable in the CFPS is primarily defined around formal urban retirement systems. Its inclusion would have drastically reduced the analyzable sample size and shifted the analysis to a nonrepresentative subpopulation, as the vast majority of respondents, particularly those in rural areas or with informal employment backgrounds, were not covered by this definition. Therefore, to maintain sample representativeness and statistical power, we opted not to include this variable.

Statistical Analysis

Descriptive statistics were presented as means (with SDs) for continuous variables and counts (with percentages) for categorical variables. To estimate the longitudinal association of digital engagement with sleep schedules, we used panel regression models to control for unobserved time-invariant individual heterogeneity. For the continuous dependent variables (nocturnal sleep duration and afternoon nap duration), we estimated panel FE models. The FE model effectively controls for all time-invariant confounders by leveraging within-individual variation over time. This process inherently accounts for within-individual correlation by effectively comparing each person to themselves across different survey waves. Year-FE were included in all models to account for common temporal shocks affecting all respondents in a given survey year. For the ordinal dependent variable (sleep onset time), a random-effects ordered logit model was used. This model was chosen because the conditional FE estimator for ordered outcomes is subject to the incidental parameters problem in short panels, which leads to biased coefficient estimates [42]. The random-effects model accounts for within-individual correlation across survey waves by including an individual-specific random intercept, which follows a logistic distribution.

We conducted mediation analyses to test the hypothesized mechanisms through which digital access affects sleep schedules. We employed separate mediation models for each mediator to isolate its unique indirect effect. The significance of the mediation (indirect) effect was tested using two methods: (1) the Sobel test and (2) the nonparametric bootstrap method with 1000 resamples to generate bias-corrected 95% CIs. Mediation was considered statistically significant if the bootstrap CI did not include 0. All mediation models controlled for year-FE and the full set of covariates listed in the Covariates section.

To examine whether the effects of digital engagement differed across subpopulations, we performed heterogeneity analyses by introducing interaction terms between the digital engagement variables (digital access and internet use) and 2 key moderators: sex and urban-rural residence. Specifically, we added the interaction terms “digital engagement×sex and digital engagement×household” to our baseline models. Stratified models were then used to estimate and report the subgroup-specific coefficients.

XSL•FO
RenderX

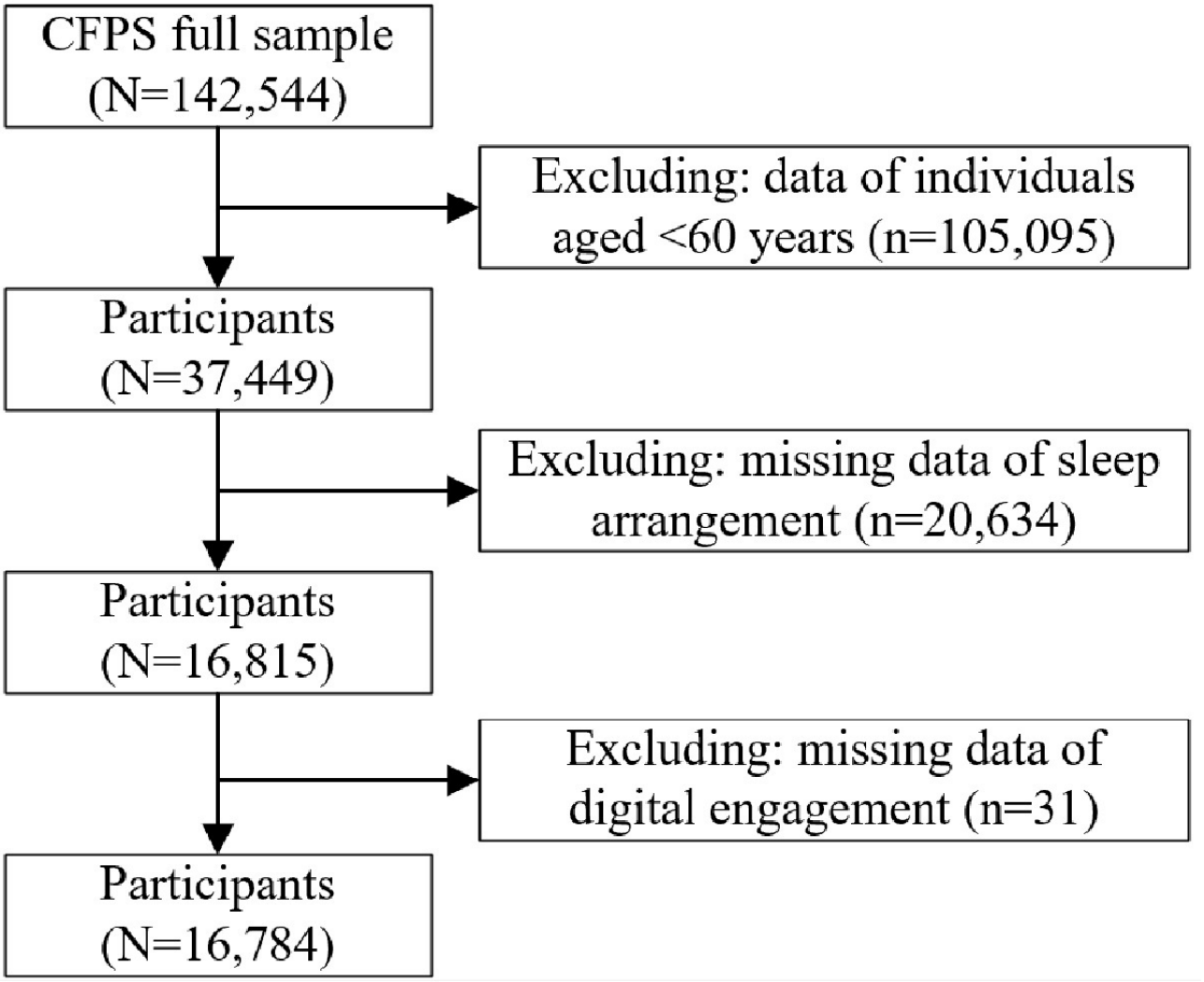
Of note, the variable internet use duration was only applicable to and analyzed for the subsample of older adults who had digital access (ie, digital access=1). For those without digital access, data on internet use duration were treated as missing and excluded from the respective analyses.

Results

Sample Characteristics

A total of 16,784 older adult samples were included in the study, including 9100 (54.22%) females and 7684 (45.78%) males with an average age of 69.726 (SD 6.95) years, after gradually excluding samples aged <60 years and samples missing digital engagement or sleep schedule (Figure 1).

Figure 1. Flowchart of the study sample inclusion process. CFPS: China Family Panel Studies.



Of the 2266 digital accesses and 15,418 digital divides, it is evident that there is considerable potential for further digital penetration among the older adult population in China. The digital access samples spent an average of 8.56 hours per week on the internet, suggesting a relatively balanced use pattern. Table 1 presents the fundamental characteristics of the samples according to their sleep schedules. The findings indicate that, in contrast to older adults experiencing digital exclusion, those who attained digital access demonstrated a reduced prevalence of extended or diminished sleep duration. Notably, the proportion of individuals experiencing protracted sleep duration was considerably lower in the study population compared to older adults within the digital divide. With regard to the duration of the afternoon nap, older adults with digital access predominantly demonstrate a tendency toward significantly abbreviated breaks, exhibiting a frequency approximately 13% higher than older adults encountering barriers to digital engagement. With regard to nocturnal somnolence, older adults with digital access exhibited a more balanced nocturnal sleep pattern, characterized by a reduction in very early sleep onset times and an increase in very late sleep onset times. In general, older adults who were male, had a high level of education, were married, had a rural household registration, and had no religion exhibited more balanced sleep duration. Similarly, older adults who were male, were married, had an urban household registration, and had no religion demonstrated more balanced afternoon nap duration. Finally, older adults who were female, had a high level of education, had an urban household registration, and had a religious affiliation exhibited healthier nocturnal sleep onset time.

XSL•FO
RenderX

Table . Baseline characteristics of participants according to sleep arrangement.

Characteristics	Total sample	Sleep per day (h)			Afternoon nap per day[a] (h)			Sleep onset time[b]		
		≤5 (n=2283)	Approximately 5-9 (n=11,039)	≥9 (n=3462)	≤0.5 (n=2379)	Approximately 0.5-1.5 (n=5166)	≥1.5 (n=3258)	≤8 PM (n=3007)	Approximately 8 PM to 12 AM (n=12,783)	≥12 AM (n=623)
Digital access, n (%)										
Yes	2266 (13.50)	244 (10.77)	1843 (81.33)	179 (7.90)	510 (33.03)	802 (51.94)	232 (15.03)	123 (5.49)	1916 (85.54)	201 (8.97)
No	14,518 (86.50)	2039 (14.04)	9196 (63.34)	3283 (22.61)	1869 (20.19)	4364 (47.13)	3026 (32.68)	2884 (20.35)	10,867 (76.67)	422 (2.98)
Online time[c], mean (SD)	2.147 (1.255)	2.165 (1.180)	2.174 (1.226)	1.805 (1.619)	2.083 (1.237)	2.152 (1.202)	2.159 (1.326)	2.012 (1.492)	2.082 (1.251)	2.711 (0.932)
Sex, n (%)										
Female	9100 (54.22)	1518 (16.68)	5679 (62.41)	1903 (20.91)	1433 (25.55)	2586 (46.11)	1589 (28.33)	1635 (18.39)	6970 (78.39)	287 (3.23)
Male	7684 (45.78)	765 (9.96)	5360 (69.76)	1559 (20.29)	946 (18.21)	2580 (49.66)	1669 (32.13)	1372 (18.24)	5813 (77.29)	336 (4.47)
At marriage[d], n (%)										
Yes	12,608 (75.12)	1565 (12.41)	8673 (68.79)	2370 (18.80)	1804 (21.98)	4028 (49.09)	2374 (28.93)	2051 (16.62)	9801 (79.42)	488 (3.95)
No	4176 (24.88)	718 (17.19)	2366 (56.66)	1092 (26.15)	575 (22.14)	1138 (43.82)	884 (34.04)	956 (23.47)	2982 (73.21)	135 (3.31)
Education, mean (SD)	5.002 (4.829)	4.467 (4.656)	5.614 (4.868)	3.370 (4.375)	5.667 (5.044)	5.501 (4.951)	4.152 (4.521)	3.102 (4.162)	5.358 (4.866)	6.995 (4.653)
Household, n (%)										
Rural	9321 (55.63)	1268 (13.60)	5577 (59.83)	2476 (26.56)	1107 (18.69)	2692 (45.45)	2124 (35.86)	2089 (22.90)	6831 (74.89)	201 (2.20)
Urban	7433 (44.37)	1008 (13.56)	5444 (73.24)	981 (13.20)	1266 (26.02)	2467 (50.71)	1132 (23.27)	911 (12.54)	5930 (81.65)	422 (5.81)
Age (y), mean (SD)	69.726 (6.946)	70.655 (7.147)	69.060 (6.649)	71.234 (7.417)	69.350 (6.647)	69.745 (6.868)	70.922 (7.190)	72.437 (7.393)	69.182 (6.695)	67.066 (5.644)
Self-rated health[e], mean (SD)	3.622 (1.184)	3.959 (1.108)	3.501 (1.159)	3.785 (1.247)	3.612 (1.138)	3.575 (1.169)	3.716 (1.229)	3.779 (1.229)	3.584 (1.169)	3.596 (1.207)
Religious belief, n (%)										
Yes	542 (3.23)	88 (16.24)	324 (59.78)	130 (23.99)	98 (26.42)	158 (42.59)	115 (31.00)	93 (17.42)	430 (80.52)	11 (2.06)
No	16,228 (96.77)	2193 (13.51)	10,709 (65.99)	3326 (20.50)	2278 (21.85)	5008 (48.03)	3140 (30.12)	2912 (18.35)	12,343 (77.79)	612 (3.86)
Chronic diseases[f], mean (SD)	0.339 (0.473)	0.432 (0.495)	0.316 (0.465)	0.350 (0.477)	0.337 (0.473)	0.355 (0.478)	0.367 (0.482)	0.349 (0.477)	0.336 (0.472)	0.340 (0.474)

[a]Missing data for 5981 (35.64%) participants.

[b]Missing data for 371 (2.21%) participants. Time points were recorded as before 8 PM (≤8 PM), 8 PM to 12 AM, and after 12 AM (≥12).

[c]Online time was measured as the natural logarithm of self-reported weekly hours online. This variable was only applicable to and calculated for the subsample with digital access (n=2266).

[d]Marital status: “married” includes those currently in marriage; “not married” includes those separated, divorced, widowed, or never married.

[e]Self-rated health was assessed on a 5-point scale from 1=very good to 5=very poor. A higher score indicates poorer self-perceived health.

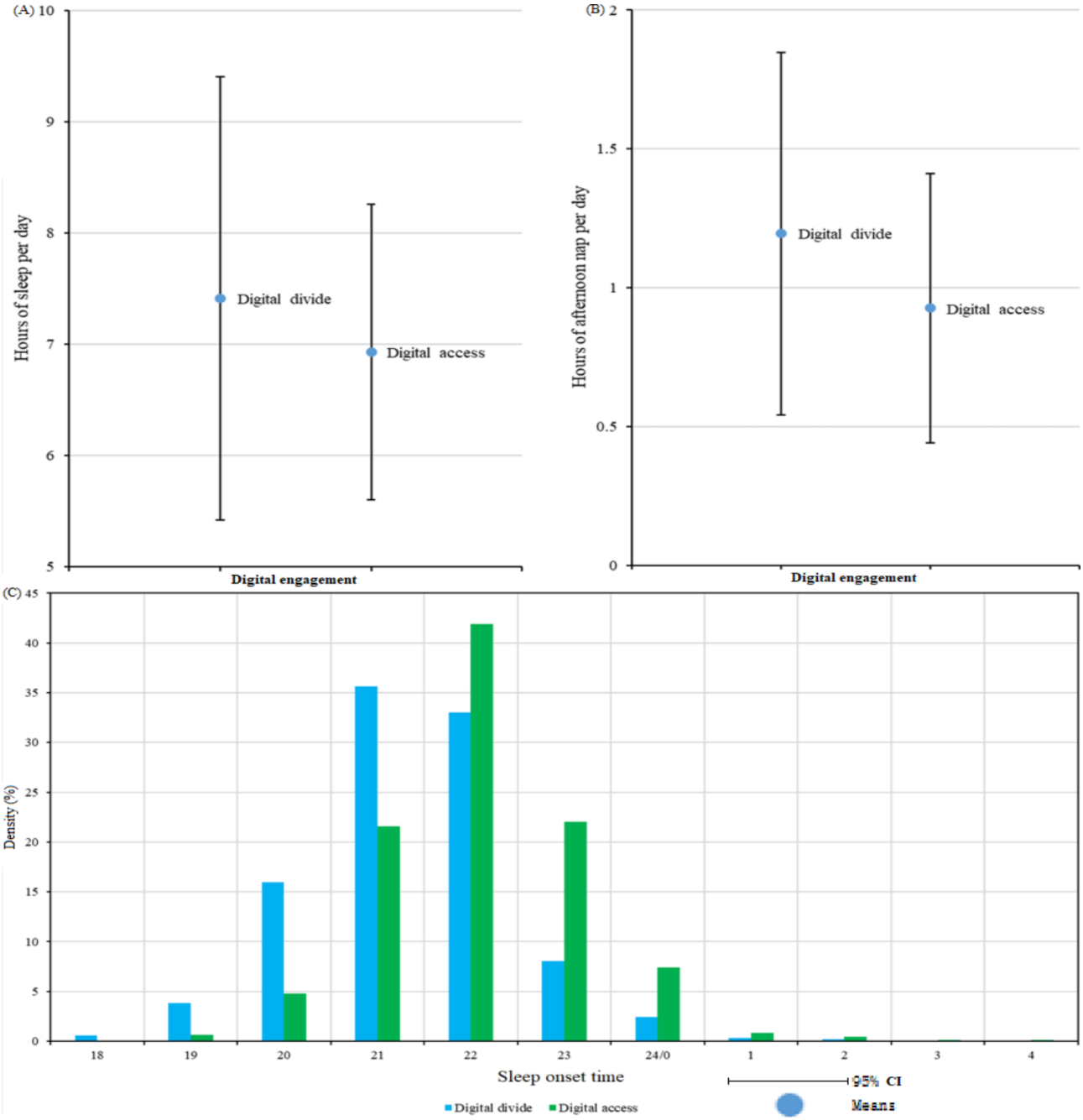
[f]Chronic disease is presented as the count and its corresponding percentage of the column total for “yes.”

XSL•FO
RenderX

Condition for Sleep Schedules According to Digital Engagement

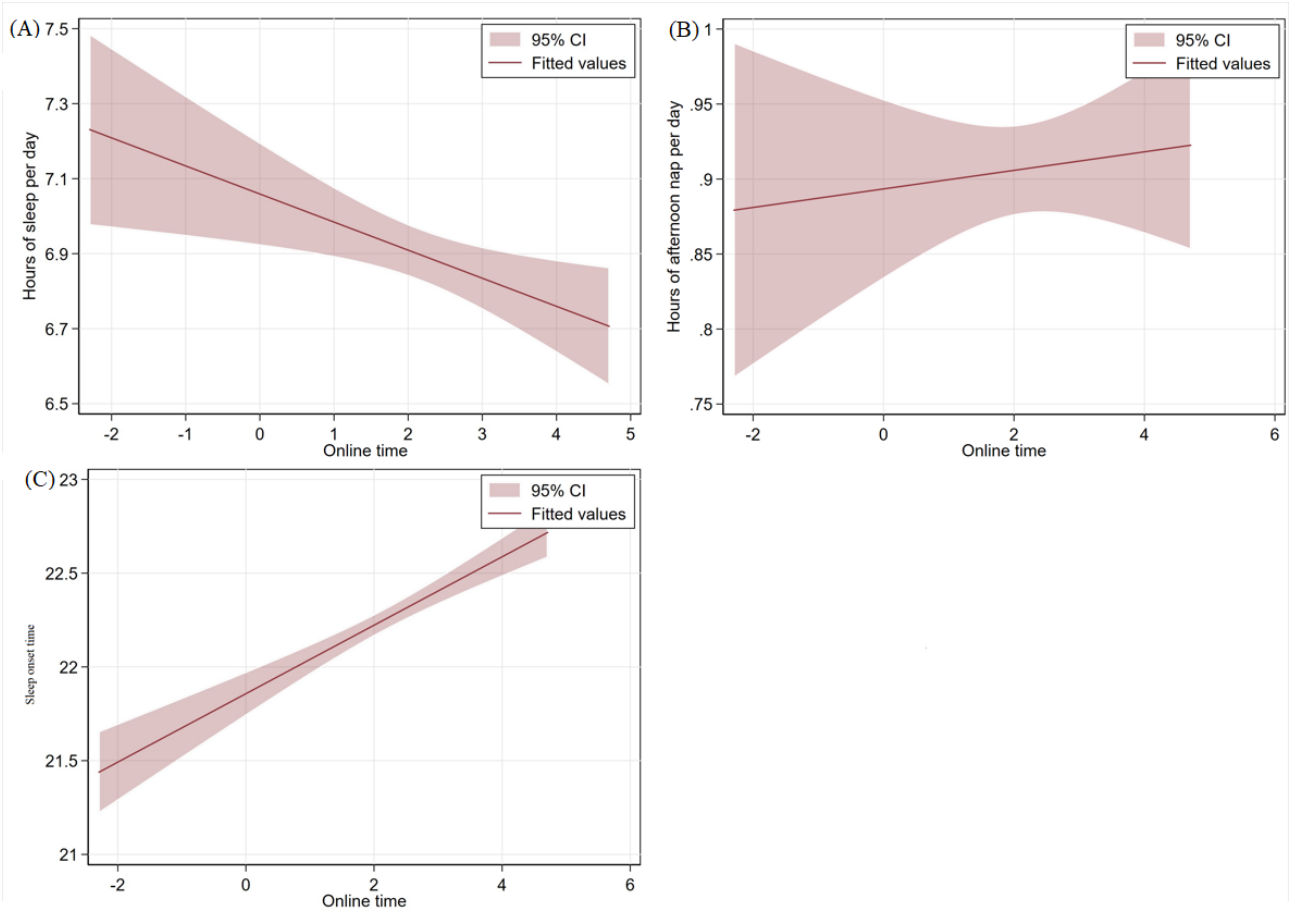
Figure 2 presents the sleep duration, afternoon nap duration, and nocturnal sleep onset time for older adults with varying digital access conditions. Figure 3 presents the correlation between online time and sleep schedules of older adults (nocturnal sleep onset time is an ordinal variable, but it can be approximated as a fixed-distance variable due to the large number of categories), with the sleep duration, afternoon nap duration, and nocturnal sleep onset time. The findings indicated that older adults experiencing the digital divide exhibited a greater number of daily sleep hours (mean 7.413, SD 1.990 h vs mean 6.929, SD 1.328 h) and more afternoon nap hours (mean 1.196, SD 0.653 h vs mean 0.926, SD 0.485 h) compared to older adults with digital access. With regard to the phenomenon of falling asleep, the overall tendency among older adults with limited digital access was to retire for the night at an earlier hour than their counterparts with digital access. The duration of time spent online exhibited a comparable tendency, with total sleep duration demonstrating an inverse relationship with the amount of time spent online. For older adults, the time of nocturnal sleep onset exhibited a positive correlation with the duration of time spent online. However, the length of the afternoon nap was proportional to the duration of internet access.

Figure 2. Condition for sleep schedules according to digital access. The density represents the probability distribution of each variable, allowing for visual comparison of the shape and central tendency of sleep schedules between older adults with (green) and without (blue) digital access. A higher density at a particular value indicates a greater proportion of the sample around that value. (A) Hours of sleep per day. (B) Hours of afternoon nap per day. (C) Sleep onset time.



XSL•FO
RenderX

Figure 3. The correlation between online time and sleep schedules.



Benchmark Regression

Table 2 reports the results of the baseline regressions (due to space limitations; Table S1 in the Multimedia Appendix 1), where model 1 is the effect of digital access on the length of daily sleep of older adults and model 2 examines the relationship between internet access duration and sleep duration among older adults. Model 3 explores the impact of digital access on the duration of the afternoon nap for this demographic. Model 4 analyzes the effect of internet access duration on the duration of afternoon nap for older adults. Model 5 investigates the impact of digital access on the timing of sleep onset for older adults during nighttime hours. Model 6 investigates the impact of internet access duration on the timing of sleep onset for older adults. The findings indicate that older adults with digital access were associated with a mean sleep duration of 0.154 hours less per day compared to those who may be experiencing a digital divide. In addition, a significant inverse correlation is observed between the duration spent online and the duration of sleep per day, suggesting that increased digital engagement may be associated with reduced sleep duration. A study of older adults revealed that those with digital access were associated with an average of 0.156 fewer afternoon nap hours than their counterparts in the digital divide. The association of internet access with this phenomenon was found to be nonsignificant. It has been demonstrated that older adults who have access to the internet are associated with a later sleep onset time than their counterparts who do not have access to the digital divide. Furthermore, the duration of time spent online has been shown to have a significant association with the timing of nighttime sleepiness among older adults.

XSL•FO
RenderX

Table . Associations of digital engagement with sleep schedules among older adults: results from panel regression models (brief).

Variables	Total sleep duration (h)		Afternoon nap duration (h)		Nocturnal sleep onset time[a]	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Digital access, β (SE)	−0.154[b] (0.049)	—[c]	−0.156[b] (0.020)	—	0.664[b] (0.076)	—
Online time[d], β (SE)	—	−0.068[e] (0.031)	—	0.009 (0.014)	—	0.372[b] (0.080)
Sex, β (SE)	0.355[b] (0.032)	0.317[b] (0.075)	0.122[b] (0.013)	0.035 (0.033)	−0.133[e] (0.067)	0.159 (0.257)
At marriage, β (SE)	−0.029 (0.038)	0.245[e] (0.111)	−0.009 (0.016)	0.102[e] (0.051)	−0.098 (0.073)	0.021 (0.355)
Education, β (SE)	−0.022[b] (0.004)	−0.003 (0.011)	−0.010[b] (0.002)	−0.002 (0.005)	0.093[b] (0.008)	0.020 (0.035)
Household, β (SE)	−0.423[b] (0.035)	−0.318[e] (0.131)	−0.138[b] (0.014)	−0.087 (0.058)	0.938[b] (0.067)	0.511 (0.396)
Age (y), β (SE)	0.005[e] (0.002)	0.002 (0.007)	0.005[b] (0.001)	0.003 (0.003)	−0.090[b] (0.005)	−0.129[b] (0.023)
Self-rated health, β (SE)	−0.008 (0.013)	−0.087[e] (0.038)	0.015[b] (0.006)	−0.002 (0.016)	−0.038[f] (0.021)	0.004 (0.101)
Religious belief, β (SE)	−0.015 (0.085)	0.067 (0.217)	−0.017 (0.034)	−0.033 (0.086)	0.128 (0.132)	−0.423 (0.553)
Chronic diseases, β (SE)	−0.078[e] (0.033)	−0.037 (0.079)	0.039[b] (0.014)	0.089[b] (0.034)	0.026 (0.048)	−0.063 (0.189)
Constant term, β (SE)	7.203[b] (0.183)	7.131[b] (0.534)	0.848[b] (0.076)	0.675[b] (0.229)	—	—
Year-fixed effect	Yes	Yes	Yes	Yes	—	—
Year random effect	—	—	—	—	Yes	Yes
Sample size	15,747	1306	10,183	853	15,402	1281
R^2	0.0351	0.0394	0.0531	0.0225	—	—
Wald chi-square test (*df*)	—	—	—	—	1226.42 (9)	55.17 (9)

[a]Nocturnal sleep onset time is an ordinal variable. Models 5 and 6 report coefficients from a random-effects ordered logit model.

[b]$P<.01$.

[c]Not applicable.

[d]Online time was measured as the natural logarithm of weekly hours and was only analyzed for the subsample with digital access. This explains the smaller sample sizes in models 2, 4, and 6.

[e]$P<.05$.

[f]$P<.10$.

Analysis of the Intermediary Mechanism

To systematically examine the mechanisms through which digital access influences sleep schedules, we tested a conceptual framework that organizes five mediators into three distinct yet potentially interconnected pathways: (1) the time displacement pathway, (2) the psychosocial resource pathway, and (3) the physiological-cognitive pathway. This framework posits that digital engagement reshapes sleep by altering individuals' daily routines, mental well-being, and physical health.

Specifically, within the time displacement pathway, we posit that digital access facilitates nonfarm payroll employment [43,44]. The time demands of such employment are hypothesized to directly compete with and reduce the time available for both nocturnal sleep and afternoon naps. Within the psychosocial resource pathway, informed by theories of aging and resource conservation, we propose that digital access enhances psychosocial well-being by reducing feelings of depression through improved social connectivity [27,28] and by maintaining functional autonomy, as measured by instrumental activities of daily living (IADLs). We hypothesize that these improvements in mental health and functional capacity may, paradoxically, lead to a reduced inclination for prolonged rest, thereby shortening sleep duration and nap time. Finally, the physiological-cognitive pathway captures the potential benefits of digital engagement on core biological and mental functions. We test whether digital access supports memory through cognitive stimulation [45,46] and improves nutrition by increasing protein intake via the convenience of e-commerce. These enhancements, while beneficial for overall health, may exert complex, and sometimes opposing, influences on sleep regulation, potentially explaining the masking effects observed in our analysis. As a prerequisite for mediation, Table S2 in Multimedia Appendix 1 confirms that digital access is

XSL•FO
RenderX

significantly associated with all 5 mediators after adjusting for covariates. The following results delineate the specific indirect effects operating through these hypothesized pathways.

As shown in Table 3, the results of the mediating mechanism tests are reported herein. The results of the analysis indicated that none of the mediating effect tests for the duration of internet access on sleep schedules were significant. Therefore, the data in the table are indicative of the mediating mechanisms for the effect of digital access on the sleep schedules of older adults. The findings indicate that digital access has a significant association with the sleep patterns of older adults, evidenced by a reduction in sleep duration, afternoon nap duration, and delayed nocturnal sleep onset time. The findings suggest that the masking effect of older adults' dietary protein intake on the attenuating effect of digital access on older adults' sleep duration implies that digital access promotes older adults' protein food intake, which in turn promotes sleep duration. The findings revealed a masking effect of memory on the association between digital access and sleep duration, as well as a delaying effect on sleep onset. This suggests that digital access may have a positive impact on memory in older adults. In addition, the study found a strong correlation between sleep quality and memory in older adults, supporting the hypothesis that sleep quality is a crucial factor in memory function [47,48]. Levels of depression in older adults had a masking effect on the attenuating effects of digital access on sleep duration and afternoon nap duration.

Table . Mediation analysis of the effect of digital access on sleep schedules via hypothesized mechanisms.

Sleep schedule and mediating variables	Effect	SE[a]	Percentage of effect (%)	95% CI	*P* value
Sleep duration per day					
Nonfarm payroll	−0.033	0.010	32.84	−0.052 to −0.013	<.001
Protein intake	0.009	0.004	−4.90	0.001 to 0.016	0.03
Memory	0.009	0.004	−5.76	0.001 to 0.017	0.02
Depression	0.014	0.003	−9.26	0.007 to 0.021	<.001
IADLs[b]	−0.014	0.004	9.49	−0.021 to −0.007	<.001
Afternoon nap					
Nonfarm payrolls	−0.015	0.005	8.50	−0.023 to −0.006	0.003
Depression	-0.002	0.001	1.25	−0.004 to 0.000	0.02
IADLs	−0.004	0.001	2.94	−0.007 to −0.002	<.001
Sleep onset time					
Nonfarm payrolls	0.017	0.006	6.05	0.006 to 0.029	0.02
Memory	−0.010	0.002	−2.69	−0.015 to −0.005	<.001
IADLs	0.007	0.002	1.89	0.004 to 0.010	<.001

[a]The significance of the indirect effects was tested using the nonparametric bootstrap method with 500 resamples. An effect is considered statistically significant if the 95% bias-corrected confidence interval (95% bootstrapped CI) does not include 0.

[b]IADL: instrumental activity of daily living. A higher score indicates greater difficulty, hence the negative association with digital access in the prerequisite model (Table S2 in Multimedia Appendix 1).

Heterogeneity Analysis

Given the substantial heterogeneity in sleep behavior across sexes [49,50], and the significant heterogeneity in sleep behavior and digital access among urban and rural older adults [51-54], this study examined the associations of digital engagement with sleep schedules among older adults, with a focus on sex and urban-rural heterogeneity. As shown in Table 4, the results of the heterogeneity test were reported. It should be noted that control variables were incorporated into the model. The findings indicate that, with respect to sleep duration, digital access exerted a negligible association with the sleep patterns of female older adults. However, among male older adults, digital access demonstrated a substantial association with sleep duration. Conversely, internet access duration exhibited a significant association with female older adults, while having no discernible associations with male older adults. The associations of digital access with sleep duration were examined for 2 distinct groups of older adults: those residing in urban areas and those residing in rural areas. The findings revealed a significant association of digital access with sleep duration for the urban older adult population. However, this association was not observed among the rural older adult population. In addition, the study investigated the association between internet access duration and sleep duration. The results indicated that for rural older adults, an increase in internet access duration was associated with a decrease in sleep duration. Conversely, this association was not observed among urban older adults. In terms of sleep onset time, digital access has a more significant effect on delaying sleep onset time for older men and older adults in rural areas, while online time has a more significant effect on delaying sleep onset time for older women and older adults in urban areas, but has no significant effect on older adults in rural areas.

XSL•FO
RenderX

Table . Heterogeneity test.

Variable	Female	Male	Rural	Urban
Heterogeneity test: effects on sleep duration				
Digital access				
Coefficient	–0.067	–0.218	–0.099	–0.259
SE	0.076	0.062	0.099	0.052
P value	.38	<.001	.32	<.001
Sample size, n (%)	8492 (53.93)	7255 (46.07)	8679 (55.12)	7068 (44.88)
R^2	0.0339	0.0344	0.0195	0.0267
Online time				
Coefficient	–0.099	–0.028	–0.175	–0.044
SE	0.045	0.043	0.088	0.033
P value	.03	.51	.051	.19
Sample size, n (%)	545 (41.73)	761 (58.27)	122 (9.34)	1184 (90.66)
R^2	0.0648	0.0272	0.1746	0.0306
Year-fixed effect	Yes	Yes	Yes	Yes
Control variable	Yes	Yes	Yes	Yes
Heterogeneity test: effects on afternoon nap				
Digital access				
Coefficient	–0.123	–0.177	–0.148	–0.158
SE	0.031	0.025	0.039	0.022
P value	<.001	<.001	<.001	<.001
Sample size, n (%)	5257 (51.63)	4926 (48.37)	5555 (54.55)	4628 (45.45)
R^2	0.0486	0.0545	0.0210	0.0505
Online time				
Coefficient	0.024	−0.002	0.026	0.006
SE	0.021	0.018	0.045	0.015
P value	.25	.93	.56	.69
Sample size, n (%)	323 (37.87)	530 (62.13)	75 (8.79)	778 (91.21)
R^2	0.0726	0.0248	0.1888	0.0193
Year-fixed effect	Yes	Yes	Yes	Yes
Control variable	Yes	Yes	Yes	Yes
Heterogeneity test: effects on sleep onset time				
Digital access				
Coefficient	0.649	0.678	0.664	0.582
SE	0.110	0.107	0.128	0.101
P value	<.001	<.001	<.001	<.001
Sample size, n (%)	8298 (53.88)	7104 (46.12)	8498 (55.17)	6904 (44.83)
Wald chi-square test (*df*)	669.81 (9)	558.26 (9)	374.40 (9)	450.92 (9)
Online time				
Coefficient	0.519	0.258	0.258	0.386
SE	0.119	0.110	0.157	0.089

XSL•FO
RenderX

Variable	Female	Male	Rural	Urban
P value	<.001	<.001	.10	<.001
Sample size, n (%)	534 (41.69)	747 (58.31)	119 (9.29)	1162 (90.71)
Wald chi-square test (*df*)	28.93 (9)	32.58 (9)	6.17 (9)	51.40 (9)
Year-fixed effect	Yes	Yes	Yes	Yes
Control variable	Yes	Yes	Yes	Yes

Discussion

Principal Results

This study examined the association of digital engagement with sleep schedules among older adults, drawing upon 4 periods of CFPS data from 2014 to 2020. The findings indicated that digital engagement was associated with a reduction in sleep duration and an advancement in nocturnal sleep onset time for older adults. In addition, digital access was associated with a decrease in the length of afternoon nap duration for this demographic. The results of the heterogeneity analysis suggest the presence of internal heterogeneity in the effects of digital engagement on older adults' sleep schedules. That is to say, the effects of digital access are not equivalent to those of hours of internet use. The association of digital access with sleep schedules among older adults was found to be more pronounced among older male adults compared to their female counterparts residing in urban areas. Similarly, the impact of internet use hours on sleep schedules was found to be more significant for older female adults compared to their male counterparts residing in urban areas. These findings suggest that digital engagement may act as a "double-edged sword" for older adults' sleep: while digital access is associated with more balanced sleep schedules, excessive internet use may undermine sleep duration and delay sleep onset.

Limitations

Despite the fact that the tracking data facilitated causal inference, this study was not without its limitations. First, and most importantly, the observational nature of our data prevents definitive causal conclusions. Although we used panel FE models and a random-effects ordered logit model to control for time-invariant confounders, reverse causality remains a plausible alternative explanation. It is possible that older adults who experience insomnia or have poorer sleep quality may turn to digital devices during wakeful periods, which could partly explain the observed associations between digital engagement and delayed or shorter sleep.

Second, due to the nature of the CFPS as a comprehensive household survey, rather than a survey specifically targeting older adults, the data are not well-targeted, and the sample size for the digitally engaged subgroup is relatively small. Third, for the same reasons as the previous point, the clinical significance of this study is relatively poor; for example, the study's limitations include its failure to consider the physiological mechanisms through which digital engagement affects older adults' sleep schedules. Finally, the lack of a new database from the CFPS for years beyond 2020, despite the study's completion, has hindered the identification of potentially evolving effects of digital engagement. Despite having completed the survey, the timeliness of this study is less than ideal, as evidenced by the year in which it was conducted: 2020.

The generalizability of our findings may be influenced by the relatively low rate of digital access (13.5%) in our analytical sample, which is an average across the survey waves from 2014 to 2020. This rate is substantially lower than the figures reported in contemporaneous Western populations [4,5]. It is, however, representative of the digital inclusion landscape for older adults in China during this specific period, reflecting a historically lower but rapidly growing penetration rate. Consequently, our findings are most directly generalizable to contexts with similar levels of digital adoption among older adults. As global digital engagement continues to rise, the associations observed here may evolve.

Comparison With Prior Work

A body of research has examined the impact of digital engagement on sleep duration in older adults [55]. However, the existing literature is not sufficiently comprehensive to determine the influence of digital engagement on the overall sleep behavior of this demographic. A number of studies have examined the efficacy of automated, web-based insomnia intervention programs in promoting sleep in older adults [56]. However, these studies have not directly addressed the impact of digital engagement on sleep quality in this demographic. A paucity of research has been conducted on the direct impact of digital engagement on the sleep patterns of older adults.

The mediation analysis unveils the complex mechanisms underpinning this "double-edged sword" effect, revealing competing pathways through which digital access influences sleep. On one hand, digital engagement appears to erode sleep time through time-displacement and activity-enabling pathways. Specifically, it facilitates nonfarm employment and maintains functional capacity (as measured by IADLs), both of which compete with time available for rest, leading to shorter nocturnal sleep, reduced napping, and later bedtimes. On the other hand, digital access also seems to promote psychosocial and physiological resources that are generally beneficial for health, yet these exert countervailing (masking) effects on sleep. For instance, it is associated with improved memory, higher protein intake, and reduced depressive symptoms—factors that are themselves positively correlated with better sleep quality or duration. The net negative association observed in our main results suggests that the sleep-disrupting effects of time displacement and heightened activity may outweigh the

XSL•FO
RenderX

protective effects conferred by these enhanced resources in the current cohort of older adults.

In light of the mounting popularity of digital devices and the expanding scope of digital engagement among older adults, it is imperative to examine the impact of digital engagement on sleep quality in this demographic. This research not only addresses a significant practical concern but also contributes to the theoretical underpinnings of related research domains. Furthermore, our mediation analysis provides empirical evidence for the potential mechanisms underpinning this "double-edged sword" effect, which can be interpreted through several theoretical lenses. The finding that digital access reduces sleep duration through increased engagement in nonfarm employment aligns with theories of time displacement, where digital connectivity fosters economic opportunities that directly compete with time traditionally allocated for rest [57]. Conversely, the observed masking effects of improved protein intake, memory, and reduced depressive mood suggest a more complex relationship. This can be framed within a social-ecological model: digital access appears to enhance psychosocial and physiological resources at the individual level [58], which are generally protective for health, yet simultaneously introduces competing time demands and cognitive stimulation at the behavioral level that disrupt sleep patterns. The net negative effect on sleep duration suggests that the time-displacing and cognitively arousing effects of digital engagement may outweigh its potential benefits for sleep in this population. Finally, the mechanism via IADL improvement—whereby better functional capacity leads to reduced sleep—resonates with the activity theory of aging, proposing that greater engagement in life activities can reduce the time and need for sleep [59]. These competing mechanisms highlight the nuanced role of digital technology in older adults' lives, acting simultaneously as a tool for resource enhancement and a source of behavioral disruption.

In this study, we operationalized digital engagement as the length of digital access and internet use, and we defined older adults' sleep as the sleep schedule, that is, the length of daily sleep, the length of afternoon nap duration, and the time of falling asleep at night. This study systematically and comprehensively focuses on the impact of digital engagement on the sleep schedule of older adults, which meets the needs of the real world and fills in the shortcomings of related research.

Implications and Contribution

The findings of the mediating mechanism test suggest that digital access has a significant impact on the total sleep duration of older adults. It does so by promoting nonfarm employment and reducing their IADLs. However, this effect is obscured by factors such as depression levels, protein intake, and memory levels. Furthermore, digital access has been observed to curtail the naptime of older adults by fostering nonfarm employment, mitigating depression, and reducing IADLs. In addition, it has been demonstrated that digital access delays nocturnal sleep onset time in older adults by promoting nonfarm employment and reducing IADLs. This effect is, however, masked by depression levels in older adults.

The findings of the mediating effect unveil two issues that have not been the focal point of prior research. First, nonfarm employment resulting from digital access, while facilitating older adults' access to economic resources, is detrimental to their overall achievement of healthier sleep schedules. Second, digital access promotes older adults' physiological well-being, whereas superior physiological well-being permits older adults to engage in more activities that may instead result in shorter sleep duration and falling asleep too late.

Subgroup characterization of sample characteristics suggests that digital access may contribute to a more balanced sleep duration, nap duration, and healthier sleep onset time for older adults overall, but also contribute to later sleep onset time and less sleep duration than older adults across the digital divide. Studies have shown that appropriate afternoon nap duration length is beneficial to cognitive health in older adults [60,61], and appropriate daily sleep length is beneficial to physical and mental health in older adults [62-64], while going to bed too late increases health risks in older adults [65], so it is possible to identify heterogeneity in the impact of digital access and internet use on older adults' sleep schedules, and the different health implications: digital access has been shown to promote healthier sleep schedules among older adults. However, excessive internet use can have adverse health consequences. While digital access contributes to healthier sleep schedules, excessive internet use can hinder the achievement of a healthy sleep schedule among older adults, leading to digital addiction.

The results of the heterogeneity analysis demonstrate the impact of sex and urban-rural inequalities in digital access. Disadvantaged groups, in their efforts to access digital resources, require greater social resources. This, in turn, could lead to these groups being more likely to be overly digitally engaged, which, in turn, leads to less healthy overall sleep schedules.

Conclusions

In this study, digital engagement was associated with shorter sleep duration and afternoon nap duration, as well as later sleep onset time in older adults. However, the associations between digital engagement and older adults' sleep schedules exhibit internal heterogeneity, with digital access favoring healthier sleep schedules and excessive hours of internet use potentially resulting in less favorable outcomes.

In summary, it is imperative for policymakers to acknowledge the dual impact of digital engagement on the sleep patterns of older adults. On the one hand, promoting digital penetration, which is associated with more balanced sleep schedules in some contexts, remains important. On the other hand, there is a need to popularize sleep hygiene and health habits among older adults who have already achieved digital penetration through digital media.

XSL•FO
RenderX

Acknowledgments

The authors thank the Aging Research Base of Lanzhou University for their help in this study. The authors also thank the Gansu Province Philosophy and Social Science Planning Project (2024QN005); National Social Science Foundation of China (22CRK014); Soft Science Special Project of the Third Batch of Provincial Science and Technology Plans in 2023 (23JRZA369); and Lanzhou University “Artificial Intelligence +" Philosophy and Social Sciences Research (LZUAIYJYB09).

Data Availability

All the data used in this study are from the data published by the China Family Tracking Survey Project Team on their official website. The datasets generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: CZ, HH
Data curation: CZ, LL
Formal analysis: CZ, LL
Funding acquisition: CZ
Methodology: CZ
Project administration: CZ
Supervision: HH
Visualization: CZ
Writing – original draft: LL, WL, HZ
Writing – review & editing: LL, WL, HH, HZ

Conflicts of Interest

None declared.

Multimedia Appendix 1
Baseline characteristics, regression models, mediation, and heterogeneity analyses examining the associations between digital engagement and sleep patterns in older adults.
[DOCX File, 21 KB - aging_v8i1e79731_app1.docx]

Checklist 1
STROBE checklist.
[DOC File, 84 KB - aging_v8i1e79731_app2.doc]

References

1. Kim J, Gray JA, Ciesla JR, Yao P. The impact of an internet use promotion programme on communication, internet use, and the extent of social networks among low-income older adults. Ageing Int 2022;47(2):348-371. [doi: 10.1007/s12126-021-09422-0] [Medline: 34305212]
2. Shapira N, Barak A, Gal I. Promoting older adults’ well-being through Internet training and use. Aging & Mental Health 2007 Sep;11(5):477-484. [doi: 10.1080/13607860601086546]
3. Zhou Y, Bai Y, Wang J. The impact of internet use on health among older adults in China: a nationally representative study. BMC Public Health 2024 ;24(1):1065. [doi: 10.1186/s12889-024-18269-4]
4. Malta S. Love, Sex and Intimacy in New Late-Life Romantic Relationships [Thesis]. : Swinburne University of Technology; 2013 URL: https://www.semanticscholar.org/paper/Love%2C-sex-and-intimacy-in-new-late-life-romantic-Malta/60f654d3ba7b971fb4807601e24954aa790b36f7 [accessed 2025-10-29]
5. Kemp S, Erades Pérez N. Consumer fraud against older adults in digital society: examining victimization and its impact. Int J Environ Res Public Health 2023 Apr 5;20(7):5404. [doi: 10.3390/ijerph20075404] [Medline: 37048017]
6. Rochat L, Wilkosc-Debczynska M, Zajac-Lamparska L, et al. Internet use and problematic use in seniors: a comparative study in Switzerland and Poland. Front Psychiatry 2021;12:609190. [doi: 10.3389/fpsyt.2021.609190] [Medline: 34177634]
7. Hargittai E, Piper AM, Morris MR. From internet access to internet skills: digital inequality among older adults. Univ Access Inf Soc 2019 Nov;18(4):881-890. [doi: 10.1007/s10209-018-0617-5]
8. Choi NG, Dinitto DM. The digital divide among low-income homebound older adults: Internet use patterns, eHealth literacy, and attitudes toward computer/internet use. J Med Internet Res 2013 May 2;15(5):e93. [doi: 10.2196/jmir.2645] [Medline: 23639979]
9. Wang J, Gao L, Wang G, Hu B. The impact of internet use on old-age support patterns of middle-aged and older adults. Front Public Health 2023 Jan 13;10. [doi: 10.3389/fpubh.2022.1059346]

XSL•FO
RenderX

10. Guo E, Li J, Luo L, Gao Y, Wang Z. eThe effect and mechanism of Internet use on the physical health of the older people—mpirical analysis based on CFPS. Front Public Health 2022;10:952858. [doi: 10.3389/fpubh.2022.952858] [Medline: 36339140]
11. Zhang Q, Li Z. The impact of internet use on the social networks of the elderly in China—the mediating effect of social participation. Int J Environ Res Public Health 2022 Jan;19(15):9576. [doi: 10.3390/ijerph19159576]
12. Kim SY, Kim MS, Park B, Kim JH, Choi HG. Lack of sleep is associated with internet use for leisure. PLoS One 2018;13(1):e0191713. [doi: 10.1371/journal.pone.0191713] [Medline: 29360882]
13. Asarnow LD, Gasperetti CE, Gumport NB, Harvey AG. Internet use and its impact on internalizing disorder symptoms and sleep in adolescents with an evening circadian preference. J Clin Sleep Med 2021 Oct 1;17(10):2019-2027. [doi: 10.5664/jcsm.9380] [Medline: 34606439]
14. Qiu WF, Ma JP, Xie ZY, Xie XT, Wang CX, Ye YD. Online risky behavior and sleep quality among Chinese college students: the chain mediating role of rumination and anxiety. Curr Psychol 2023 Jun;42(16):13658-13668. [doi: 10.1007/s12144-021-02641-3]
15. Vaz Fragoso CA, Gill TM. Sleep complaints in community-living older persons: a multifactorial geriatric syndrome. J Am Geriatr Soc 2007 Nov;55(11):1853-1866. [doi: 10.1111/j.1532-5415.2007.01399.x] [Medline: 17916123]
16. Kim DE, Yoon JY. Factors that influence sleep among residents in long-term care facilities. Int J Environ Res Public Health 2020 Mar 14;17(6):1889. [doi: 10.3390/ijerph17061889] [Medline: 32183274]
17. Thichumpa W, Howteerakul N, Suwannapong N, Tantrakul V. Sleep quality and associated factors among the elderly living in rural Chiang Rai, northern Thailand. Epidemiol Health 2018 May 14;40:e2018018. [doi: 10.4178/epih.e2018018]
18. De Jesús R, Fishbein W. Risk and resiliency factors associated with poor sleep quality in elderly populations. MOJ Gerontology & Geriatrics 2021 May 13;6(3):64-67. [doi: 10.15406/mojgg.2021.06.00270]
19. Kinnunen KM, Rapaport P, Webster L, et al. A manual-based intervention for carers of people with dementia and sleep disturbances: an acceptability and feasibility RCT. Health Technol Assess 2018;22(71):1-408. [doi: 10.3310/hta22710]
20. Okamoto-Mizuno K, Tsuzuki K, Mizuno K. Effects of mild heat exposure on sleep stages and body temperature in older men. Int J Biometeorol 2004 Sep;49(1):32-36. [doi: 10.1007/s00484-004-0209-3] [Medline: 15173935]
21. Tsuzuki K, Sakoi T, Sakata Y. Effect of seasonal ambient temperature on sleep and thermal comfort in older people living in public elderly facilities. Buildings 2021 Dec;11(12):574. [doi: 10.3390/buildings11120574]
22. Solikhah FK, Nursalam N, Ulfiana E. The effect of sleep hygiene on the sleep quality in elderly. Presented at: Proceedings of the 8th International Nursing Conference on Education, Practice and Research Development in Nursing (INC 2017); Apr 8-9, 2017; East Java, Indonesia p. 101-103. [doi: 10.2991/inc-17.2017.26]
23. Melanie F, Noviani W. Application of sleep hygene and elderly gymnastics against sleep pattern disorders in the elderly. KESANS International Journal of Health and Science 2023 Jul 20;2(10):845-854. [doi: 10.54543/kesans.v2i10.206]
24. de Carvalho KM, do Livramento Fortes Figueiredo M, Neto NMG, de Moura Sá GG, Silva C, Mendes PN. Comparison of the effectiveness of two educational interventions on sleep quality in older adults: a randomized clinical trial. Rev Esc Enferm USP 2023 Jan 16;56:e20220326. [doi: 10.1590/1980-220x-reeusp-2022-0326pt]
25. Cajochen C, Frey S, Anders D, et al. Evening exposure to a light-emitting diodes (LED)-backlit computer screen affects circadian physiology and cognitive performance. J Appl Physiol (1985) 2011 May;110(5):1432-1438. [doi: 10.1152/japplphysiol.00165.2011] [Medline: 21415172]
26. Höhn C, Hahn MA, Gruber G, Pletzer B, Cajochen C, Hoedlmoser K. Effects of evening smartphone use on sleep and declarative memory consolidation in male adolescents and young adults. Brain Commun 2024;6(3):fcae173. [doi: 10.1093/braincomms/fcae173] [Medline: 38846535]
27. Chen E, Wood D, Ysseldyk R. Online social networking and mental health among older adults: a scoping review. Can J Aging 2022 Mar;41(1):26-39. [doi: 10.1017/S0714980821000040] [Medline: 33736725]
28. Chen W, Ding M, Wang X. The contribution of the internet to promoting mental health for older adults: cross-sectional survey in China. J Med Internet Res 2023 Dec 19;25(1):e40172. [doi: 10.2196/40172] [Medline: 38113096]
29. Wei N, Sun D, Huang W. Effects of WeChat use on the subjective health of older adults. Front Psychol 2022;13. [doi: 10.3389/fpsyg.2022.919889]
30. Gao B, Han X, Ye F, Li Y. Internet usage and household electricity consumption. Front Energy Res 2023 Nov 29;11. [doi: 10.3389/fenrg.2023.1323077]
31. Hanifah L, Nasrulloh N, Sufyan DL. Sedentary behavior and lack of physical activity among children in Indonesia. Children (Basel) 2023 Jul 26;10(8):1283. [doi: 10.3390/children10081283] [Medline: 37628282]
32. Richardson MX, Ehn M, Stridsberg SL, Redekop K, Wamala-Andersson S. Nocturnal digital surveillance in aged populations and its effects on health, welfare and social care provision: a systematic review. BMC Health Serv Res 2021 Jun 30;21(1):622. [doi: 10.1186/s12913-021-06624-9] [Medline: 34187472]
33. Kutzer Y, Whitehead L, Quigley E, Stanley M. Changes in sleep effort mediate insomnia severity in older adults following online cognitive behavioural therapy. Psychogeriatrics 2024 Mar;24(2):303-311. [doi: 10.1111/psyg.13074] [Medline: 38168883]

XSL•FO
RenderX

34. Song S, Seo Y, Hwang S, Kim HY, Kim J. Digital phenotyping of geriatric depression using a community-based digital mental health monitoring platform for socially vulnerable older adults and their community caregivers: 6-week living lab single-arm pilot study. JMIR Mhealth Uhealth 2024 Jun 17;12(1):e55842. [doi: 10.2196/55842] [Medline: 38885033]
35. Matos LC, Rebellato C, Lopes CBB, Nascimento JS. Impacts of sleep changes on occupational performance: perceptions of older people. Cad Bras Ter Ocup 2024 Sep 27;32:e3728. [doi: 10.1590/2526-8910.ctoao286837281]
36. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008 Apr;61(4):344-349. [doi: 10.1016/j.jclinepi.2007.11.008]
37. Goodyear MDE, Krleza-Jeric K, Lemmens T. The Declaration of Helsinki. BMJ 2007 Sep 29;335(7621):624-625. [doi: 10.1136/bmj.39339.610000.BE]
38. Yang Y, Liu W, Ji X, et al. Extended afternoon naps are associated with hypertension in women but not in men. Heart Lung 2020 Jan;49(1):2-9. [doi: 10.1016/j.hrtlng.2019.09.002]
39. Xin C, Zhang B, Fang S, Zhou J. Daytime napping and successful aging among older adults in China: a cross-sectional study. BMC Geriatr 2020 Jan 2;20(1):2. [doi: 10.1186/s12877-019-1408-4] [Medline: 31898552]
40. Jing R, Xu T, Rong H, Lai X, Fang H. Longitudinal association between sleep duration and depressive symptoms in Chinese elderly. Nat Sci Sleep 2020 Oct 13;12:737-747. [doi: 10.2147/NSS.S269992] [Medline: 33117009]
41. Luo M, Dong Y, Fan B, et al. Sleep duration and functional disability among Chinese older adults: cross-sectional study. JMIR Aging 2024 Jun 10;7:e53548. [doi: 10.2196/53548] [Medline: 38771907]
42. Greene W. Convenient estimators for the panel probit model: further results. Empir Econ 2004 Jan 1;29(1):21-47. [doi: 10.1007/s00181-003-0187-z]
43. Wang Y, Qi Y, Li Y. How does digital inclusive finance influence non-agricultural employment among the rural labor force?—Evidence from micro-data in China. Heliyon 2024 Jul;10(13):e33717. [doi: 10.1016/j.heliyon.2024.e33717]
44. Biddle JE, Hamermesh DS. Sleep and the allocation of time. J Polit Econ 1990 Oct;98(5, Part 1):922-943. [doi: 10.1086/261713]
45. Zhang S, Wang S. Digital treatment: base on the mobile interface for memory improvement of elderly. Presented at: 2023 4th International Symposium on Artificial Intelligence for Medicine Science; Oct 20-22, 2023; Chengdu, China p. 574-577. [doi: 10.1145/3644116.3644211]
46. Shelton BE, Uz C. Immersive technology and the elderly: a mini-review. Gerontology 2015;61(2):175-185. [doi: 10.1159/000365754] [Medline: 25502624]
47. Sonni A, Spencer RMC. Sleep protects memories from interference in older adults. Neurobiol Aging 2015 Jul;36(7):2272-2281. [doi: 10.1016/j.neurobiolaging.2015.03.010] [Medline: 25890819]
48. Wilson JK, Baran B, Pace-Schott EF, Ivry RB, Spencer RMC. Sleep modulates word-pair learning but not motor sequence learning in healthy older adults. Neurobiol Aging 2012 May;33(5):991-1000. [doi: 10.1016/j.neurobiolaging.2011.06.029] [Medline: 22244090]
49. van den Berg JF, Miedema HME, Tulen JHM, Hofman A, Neven AK, Tiemeier H. Sex differences in subjective and actigraphic sleep measures: a population-based study of elderly persons. Sleep 2009 Oct;32(10):1367-1375. [doi: 10.1093/sleep/32.10.1367] [Medline: 19848365]
50. Kováčová K, Stebelová K. Sleep characteristics according to gender and age measured by wrist actigraphy. Int J Environ Res Public Health 2021 Dec 15;18(24):13213. [doi: 10.3390/ijerph182413213] [Medline: 34948821]
51. Wu CY, Miller L, Wall R, Beattie Z, Kaye J, Silbert L. Physical inactivity in older adult couples: a dyadic analysis using continuous in-home monitoring. Innov Aging 2020 Dec 16;4(Supplement_1):412-412. [doi: 10.1093/geroni/igaa057.1327]
52. Hu H, Xu W. Socioeconomic differences in digital inequality among Chinese older adults: results from a nationally representative sample. PLoS One 2024;19(4):e0300433. [doi: 10.1371/journal.pone.0300433] [Medline: 38564613]
53. Zhang J, Li D, Gao J. Health disparities between the rural and urban elderly in China: a cross-sectional study. Int J Environ Res Public Health 2021 Jan;18(15):8056. [doi: 10.3390/ijerph18158056]
54. Delbari A, Ahmadi F, Zar A, Zandvakili A, Sadeghipour HR, Sims J. Living in urban or rural environments affect the sleep quality of the elderly in Bushehr (Southern Iran): emphasizing the active and inactive of the elderly. BMC Public Health 2024 May 18;24(1):1340. [doi: 10.1186/s12889-024-18747-9] [Medline: 38760757]
55. Fan ZY, Yin RY, Tang L, Zhang CH, Zhang F. Relationships between internet use and sleep duration in Chinese adults: a cross-sectional study. Int J Gen Med 2021 Aug 20;14:4677-4685. [doi: 10.2147/IJGM.S317658] [Medline: 34447262]
56. Shaffer KM, Smith JG, Glazer JV, et al. Effects of an internet-delivered insomnia intervention for older adults: a secondary analysis on symptoms of depression and anxiety. J Behav Med 2022 Oct;45(5):728-738. [doi: 10.1007/s10865-022-00353-y] [Medline: 35932397]
57. Vanden Abeele MMP, Vandebosch H, Koster EHW, et al. Why, how, when, and for whom does digital disconnection work? A process-based framework of digital disconnection. Commun Theory 2024 Feb 23;34(1):3-17. [doi: 10.1093/ct/qtad016]
58. Buckingham S, Tu G, Elliott L, et al. Digital competence and psychological wellbeing in a social housing community: a repeated survey study. BMC Public Health 2023 Oct 13;23(1):2002. [doi: 10.1186/s12889-023-16875-2]
59. Monjan AA. Perspective on sleep and aging. Front Neur 2010 Sep 28;1. [doi: 10.3389/fneur.2010.00124]

XSL•FO
RenderX

60. Li J, Cacchione PZ, Hodgson N, et al. Afternoon napping and cognition in Chinese older adults: findings from the China Health and Retirement Longitudinal Study baseline assessment. J Am Geriatr Soc 2017 Feb;65(2):373-380. [doi: 10.1111/jgs.14368] [Medline: 27995615]
61. Li J, Chang YP, Riegel B, et al. Afternoon naps may preserve cognition in Chinese older adults. J Gerontol A Biol Sci Med Sci 2018 Mar 2;73(3):360-366. [doi: 10.1093/gerona/glx069]
62. Ding L, Zhang L, Cui Y, et al. The association of sleep duration and quality with depressive symptoms in older Chinese women. PLoS One 2022;17(3):e0262331. [doi: 10.1371/journal.pone.0262331] [Medline: 35290372]
63. He M, Deng X, Zhu Y, Huan L, Niu W. The relationship between sleep duration and all-cause mortality in the older people: an updated and dose-response meta-analysis. BMC Public Health 2020 Jul 28;20(1):1179. [doi: 10.1186/s12889-020-09275-3] [Medline: 32723316]
64. da Silva AA, de Mello RGB, Schaan CW, Fuchs FD, Redline S, Fuchs SC. Sleep duration and mortality in the elderly: a systematic review with meta-analysis. BMJ Open 2016 Feb 17;6(2):e008119. [doi: 10.1136/bmjopen-2015-008119] [Medline: 26888725]
65. Fan Y, Wu Y, Peng Y, et al. Sleeping late increases the risk of myocardial infarction in the middle-aged and older populations. Front Cardiovasc Med 2021;8:709468. [doi: 10.3389/fcvm.2021.709468] [Medline: 34631815]

Abbreviations

CFPS: China Family Panel Studies
FE: fixed effects
IADL: instrumental activity of daily living

Edited by PH Chau; submitted 26.06.25; peer-reviewed by L Ji, M Mattos; accepted 15.10.25; published 05.11.25.

Please cite as:
Zhang C, Lin L, Wang L, Hu H, Zhang H
The Double-Edged Sword of Digital Engagement—How Digital Access and Internet Use Reshape Sleep Schedules and Underlying Mechanisms in Older Adults: Longitudinal Observational Study
JMIR Aging 2025;8:e79731
URL: https://aging.jmir.org/2025/1/e79731
doi:10.2196/79731

© Chi Zhang, Longxuan Lin, Li Wang, Han Hu, Heyang Zhang. Originally published in JMIR Aging (https://aging.jmir.org), 5.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Aesthetic Experience in the Acceptance of Wearable Technology for People With Dementia: Critical Interpretive Synthesis

Yixuan Wei[1], MA; John Ratcliffe[1,2], PhD; Dag Aarsland[3], PhD; Wei Liu[1], PhD

[1]Department of Engineering, King's College London, Strand, London, United Kingdom

[2]School of Healthcare, University of Leeds, Leeds, United Kingdom

[3]Institute of Psychiatry, Psychology and Neuroscience, King's College London, United Kingdom

Corresponding Author:
Wei Liu, PhD
Department of Engineering, King's College London, Strand, London, United Kingdom

Abstract

Background: The prevalence of dementia has led to a growing interest in wearable technologies to assist dementia care. Despite their potential, these technologies face low adoption rates, often attributed to poor aesthetic design and insufficient consideration of user experience.

Objective: This study aims to (1) explore how the aesthetic design of wearable devices relates to their adoption and user experience in dementia care and (2) critically examine the ways in which aesthetic elements shape people with dementia's perceptions of acceptability and inform future design considerations.

Methods: A critical interpretive synthesis with a systematic search was conducted across 2 databases, namely Web of Science and Scopus on August 22, 2024. Studies were included if they reported on the current use of wearable technologies in dementia care or provided value in qualitative studies addressing attitudes from people with dementia and their caregivers toward the wearable product. Two authors independently screened the abstracts and full texts to extract data, and additional studies were included from web searches, owing to their conceptual contributions to offering insights into the emergence of wearable technology, including the factors driving its commercial value and appraisal.

Results: A total of 63 studies were included in this study. Findings suggest that aesthetically considered designs are preferred by users when concerning their acceptance toward wearable devices, particularly when devices symbolize empowerment and support personal engagement. The objects that evoke comfort, emotional connection, and personal meaning are more likely to be accepted by people with dementia. Improved aesthetics may also support caregivers through more consistent and effective data collection.

Conclusions: This study uncovers a significant gap in the aesthetic design of wearable technologies for dementia care, limiting user acceptance and emotional engagement. By synthesizing key themes focusing on the interaction between user and product, this review proposes a conceptual framework for dementia care, emphasizing the importance of aesthetics in enabling more meaningful, inclusive, and human-centered design.

(JMIR Aging 2025;8:e72082) doi:10.2196/72082

KEYWORDS

dementia care; wearable technology; aesthetics; health care design; assistive technology

Introduction

Background

Dementia is an incurable disease that causes brain damage from various conditions, contributing to disability and severe dependency, particularly in older adults [1]. In 2023, over 37 million people were diagnosed with dementia, with a projection for the number to rise to 10 million annually, reaching over 300 million by 2050 [1]. As dementia diagnoses increase, greater strain is placed on caregivers and health care systems, affecting the mental and physical well-being of those involved [2]. To date, smart wearables equipped with sensors to monitor daily health data or provide support are considered promising tools to ease these burdens [3,4]. Although a growing number of studies have shown an interest in testing the feasibility of wearable technologies in dementia care, there remains a low adoption rate among patients in accepting wearable devices [5]. While aesthetics is a critical factor that often directly influences the adoption of devices [6], its role in wearable technology design, particularly in dementia care, remains underexplored. This study addresses this gap by investigating how aesthetic factors contribute to user engagement and experience in an exploratory context in dementia care.

XSL•FO
RenderX

Dementia

Dementia encompasses a range of progressive symptoms characterized by the decline of cognitive and functional abilities, often associated with neuronal loss in the hippocampus and cerebral cortex [7]. Common forms include Alzheimer disease and vascular dementia, both of which significantly impact daily functioning [8]. As the condition progresses, individuals often experience memory loss, impaired communication, and behavioral changes, which are frequently accompanied by behavioral and psychological symptoms, such as the emergence of agitation, aggression, and anxiety [9]. While research on wearable technologies' application in dementia care has shown significant potential in early-stage disease detection [10,11], GPS location tracking [12], and daily activity monitoring [13-15], most have prioritized feasibility over design considerations. As a result, issues of stigma directed toward people with dementia have persisted, potentially influencing their willingness and ability to use these products [16,17]. Aesthetics play a crucial factor in users' acceptance and adoption of these devices, encouraging individuals to try and adopt devices intended for continuous health monitoring and long-term well-being [14].

Aesthetic Experience in Dementia Care

As emerging technologies, such as virtual reality and wearable devices, become increasingly integrated into health care and daily life, there is growing recognition that their success depends not only on functional performance but also on the quality of user experience [18,19]. Aesthetic experience, thus, can be addressed to help understand how individuals perceive, engage with, and emotionally respond to a product [20].

Aesthetics contributes to a sense of delight and appreciation when engaging with a product, often through sensations of beauty, amusement, and pleasure derived from sensory experiences, such as touch, sight, or sound [20]. The impressions left by aesthetics are deeply tied to how individuals physically interact with objects in their environment. Beyond the level of appeal from the appearance, design often carries symbolic meaning, where objects also evoke the development of cultural and personal associations. This symbolic layer of aesthetics reflects how users interpret and internalize a product's presence, which can influence not only their comfort and engagement but also their sense of identity and connection. Although dementia progressively impairs memory and the processing of selfhood, it is argued that the selfhood and one's aesthetic awareness can persist despite cognitive decline, such as how one dresses or behaves in a social context [21]. While stigma has been identified as a significant concern for people with dementia, who may face discrimination and ageism, the preservation of self-image, thus, plays a crucial role in shaping how individuals perceive their appearance and social standing [22]. To this extent, understanding both the physical engagement and symbolic meaning in aesthetics from the interaction between people with dementia and their surroundings offers valuable insights for designing and conceptualizing technologies that are not only usable but also emotionally and socially resonant.

Research Aim

This research aims to critically investigate (1) how wearable technologies are currently applied in dementia care and (2) how these technologies are experienced in terms of aesthetics and long-term engagement from the perspectives of people with dementia. In response to the limited theoretical development in this area, this review adopts a critical interpretive synthesis (CIS) approach to engage with the existing literature. This method enables a critical exploration of whether and how aesthetic values are considered by people with dementia through their interactions with the devices. By emphasizing the role of aesthetics in user acceptance, this study seeks to inform the design of wearable technologies that prioritize comfort, beauty, and personal identity, thereby supporting sustained use, enhancing quality of life, and increasing the impact of assistive technologies in dementia care.

Methods

Design

This review used CIS to critically examine existing literature on the application of wearable technology in dementia care, focusing on aesthetic design. CIS is instrumental in generating innovative theories by synthesizing evidence for deeper insights into the topic. It merges systematic review methods with qualitative analysis to examine both empirical and nonempirical studies [23]. A flow diagram is generated to assess literature eligibility and illustrate the clarity of the data selection process, with 2 reviewers involved to ensure diverse perspectives. This approach enabled the identification of key insights and future research directions, particularly on inclusive wearable design and its impact on patient adoption and data collection in dementia care.

Search Strategy

This review collected studies from 2 well-grounded databases, namely Scopus and Web of Science. The search was conducted on August 22, 2024. An example of the search term is (jewel* OR wear* OR accessory* OR fashion* OR craft*) AND dementia* (TOPIC) on Web of Science, and the language is limited to English (LANGUAGE). The year limit and journal types have not been set in the literature search to allow flexibility to offer a broad view of theories and practices of wearable devices and their application in dementia. Altogether, 3 additional studies [24-26] were added to the inclusion list of the reviewed studies from a web search, owing to their conceptual contributions to offering insights into the emergence of wearable technology, including the factors driving its commercial value and appraisal [24-26].

Study Selection

The inclusion criteria of studies are considered relevant based on the following: (1) studies relevant to technology that physically touches and interacts with people with dementia, (2) a feasibility study on applied wearable technology, and (3) interviews or discussions toward the existing or preferred wearables from the perspectives of people with dementia. Studies in different disciplines and perspectives (eg, discussion of aesthetics for assistive devices, the commercial market of

XSL•FO
RenderX

wearable technology, and product design framework) are considered helpful and offer insights into data analysis. The exclusion of studies involving irrelevant topics or unclear presentations of the wearable prototype.

Quality Appraisal

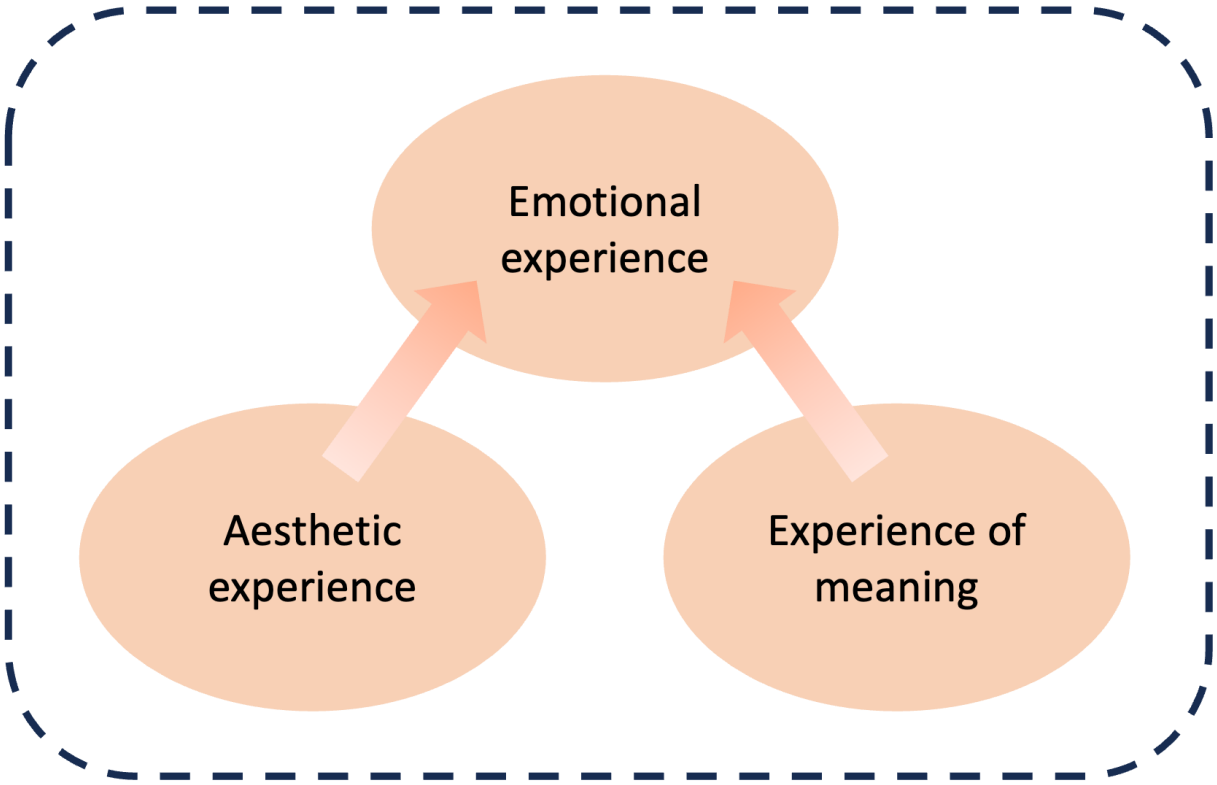
This study adopted a combination of quality criteria to guide the literature review process. The initial selection of literature from 2 databases followed a structured and transparent filtration procedure, documented through a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)-style flowchart. Following the principles of CIS, additional conceptually significant sources were purposefully sampled beyond the initial search, including gray literature, such as a website of a product and materials identified via Google Scholar. These were included for their value in illuminating the symbolic positioning and public perception of wearable technologies in contemporary culture [23,27]. Two reviewers have been assigned to conduct a full-text review independently to determine the final inclusion of studies in the review.

Data Extraction

Data extraction was conducted by the lead author using Microsoft Excel, designed to systematically capture information relevant to the research aims. Extracted details included study authors, domain, design, country, body location of wearable application, types or brands of wearable technologies, and whether the study addressed practical feasibility, aesthetic value, or both.

The inclusion of aesthetic value was guided by Desmet and Hekkert's [20] conceptual framework, selected for its relevance in examining how aesthetics shapes user-product interaction (Figure 1). The term product was interpreted broadly to include artifacts with tangible features (eg, shape, texture, and color) designed for user engagement.

Figure 1. Product experience framework (adapted from Desmet and Hekkert [20], which was originally based on the earlier work of Hekkert [28]). Published under Creative Commons Attribution 4.0 License [CC BY 4.0] [29].



Data Synthesis

Further coding of studies identified as engaging with the aesthetic value of the product was carried out by 2 reviewers using the 3 themes. The framework served as a sensitizing tool, providing a conceptual lens while allowing interpretive flexibility in line with CIS. For instance, discussions of device appearance (eg, preferred color and weight) informed the aesthetic experience dimension, while references to personal significance or identity informed the experience of meaning dimension. Given the interrelated nature of the themes, studies were allowed to span multiple categories where appropriate. Each reviewer conducted independent coding, followed by a collaborative discussion to resolve discrepancies and ensure coherence in categorizing the themes.

Findings were presented through a combination of tables, visuals, and narrative synthesis. One table summarized all included studies, while a second focused specifically on the distribution of aesthetic themes. Thematic analysis, which was initially guided by the referenced framework, was then explored under the dementia care context. Insights gained from this process informed the development of a new conceptual framework to better capture the role of aesthetic values in the

XSL•FO
RenderX

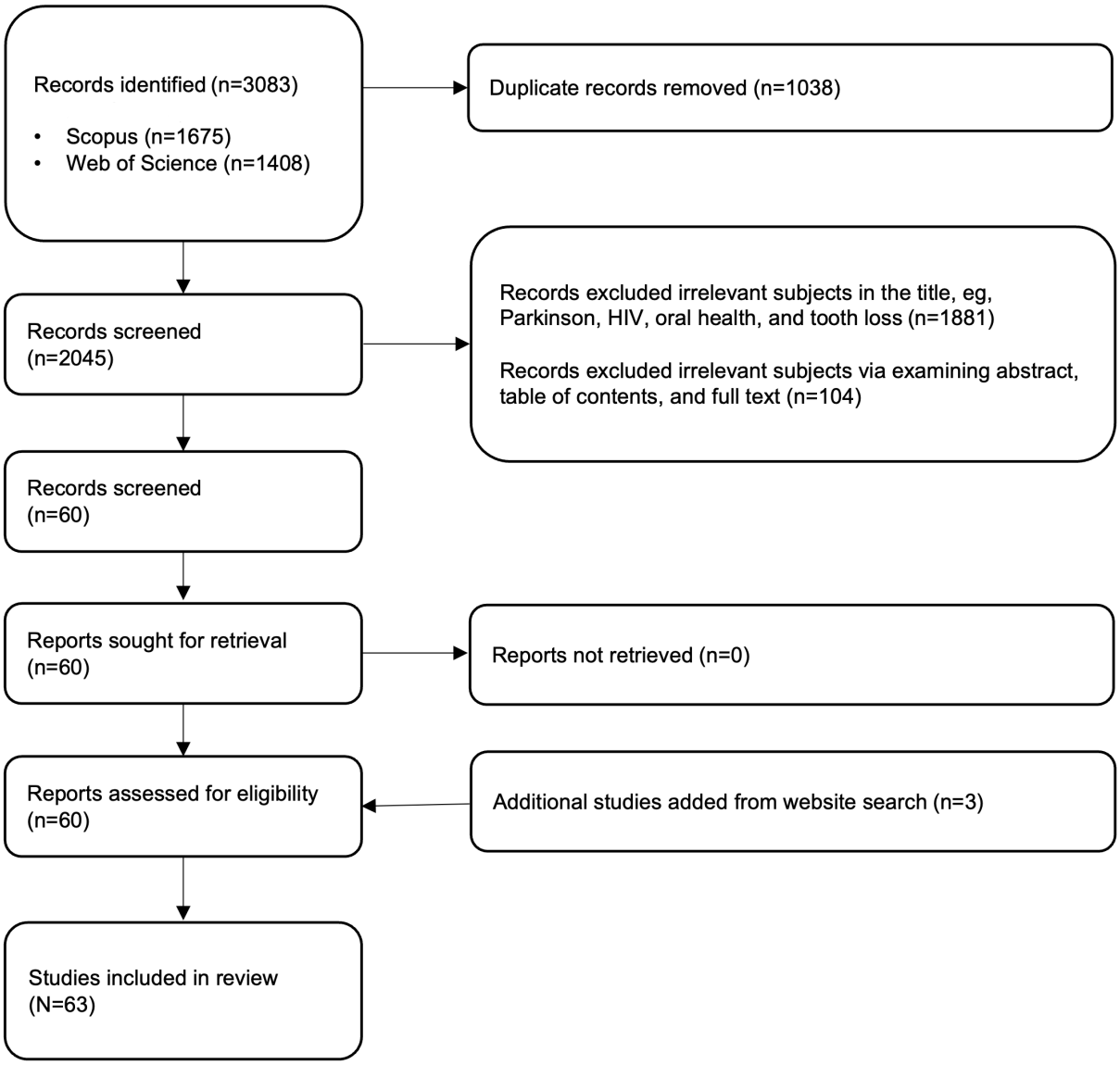
design and acceptance of wearable technologies for people with dementia.

Results

Search Results and Analysis of the Wearables

The initial search returned 3083 results (1675 from Scopus and 1408 from Web of Science). Following the application of exclusion criteria, 60 studies [4,5,11-15,18,30-81] were included. An additional 3 studies [24-26] were purposefully added for their conceptual relevance to the aesthetic and commercial dimensions of smart wearable devices, bringing the total to 63 studies [4,5,11-15,18,24-26,30-81] (Figure 2). The complete list is presented in Table 1.

Figure 2. Flow diagram of reviewed work.



XSL•FO
RenderX

Table . Reviewed studies distribution chart.

Literature reference	Study domain	Study design	Country	Body location	Wearable technology applied	Practical application	Aesthetics value
Apple [26]	N/A[a]	Website or Gray literature	United States	Wrist	Apple Watch Series 6		✓
Alam et al [14]	Dementia, agitation	Iterative study	United States	Wrist	The Pebble Watch	✓	
Alam et al [15]	Dementia	Transdisciplinary study	United States	Wrist	The Pebble Watch	✓	
Amato et al [30]	Alzheimer	Mixed methods	Italy	Wrist	Empatica E4, the Eclipse	✓	✓
Anwar et al [11]	Vascular dementia	Simulation and experimental design	United Kingdom	Eye	smart Radio Frequency (RF) glasses (proposed)	✓	
Appel et al [31]	Dementia, BPSD[b]	Pilot Study	Canada	Head	Samsung Gear VR[c] HMD[d] with Sennheiser HD 221 headphones	✓	✓
Mc Ardle et al [32]	Alzheimer, mild cognitive impairment	Feasibility study	United Kingdom	Wrist	AX3 Axivity	✓	
Mc Ardle et al [33]	Alzheimer	Pilot study	United Kingdom	Waist	AX3 Axivity	✓	✓
Badawi et al [34]	Dementia, BPSD	Mixed methods	Canada	Wrist	Empatica E4	✓	
Boletsis et al [35]	Dementia	Qualitative feasibility study	Norway	Wrist	Basis B1	✓	
Chan et al [36]	Dementia	Feasibility and validity	Australia	Wrist	AX3 Axivity	✓	
Cheung et al [37]	Dementia	Qualitative study	United Kingdom	Head	Meta Quest 2	✓	✓
Cheung et al [38]	Dementia, agitation	Scoping review	China	Wrist	N/A	✓	✓
Choi and Kim [24]	N/A	N/A	South Korea	Wrist	N/A		✓
Chung et al [39]	Dementia	Dyadic case study	United States	Wrist	Garmin Vivoactive HR[e]	✓	✓
Cullen et al [4]	Dementia	Systematic review	United Kingdom	Wrist	N/A	✓	
Davidoff et al [40]	Dementia, agitation	Multimodal system design	Belgium	Wrist	Chill+	✓	
Davis and Sikorskii [41]	Alzheimer	Feasibility Study	United States	Head	Eye tracking glasses with joystick	✓	✓
Farina et al [42]	Dementia	Feasibility study	United Kingdom	Wrist	GENEactive Original	✓	✓
Favela et al [43]	BPSD	Mixed methods	Mexico	Wrist	Fitbit Charge 2 HR	✓	✓
Felber et al [44]	Dementia	Qualitative study	Switzerland	N/A	N/A		✓
Guu et al [45]	Dementia, Alzheimer	Review study	United Kingdom	Wrist	N/A	✓	✓
Guu et al [46]	Dementia	Feasibility study	United Kingdom	Wrist	GENEactive	✓	✓

XSL•FO
RenderX

Literature reference	Study domain	Study design	Country	Body location	Wearable technology applied	Practical application	Aesthetics value
Harper and Ghali [47]	Agitation and aggression, BPSD	Systematic review	United Kingdom	Wrist	Empatica E4, The Philips Discrete Tension Indicator 2 (DTI-2), The Pebble watch, The Eclipse	✓	
Hassan et al [48]	Dementia	Qualitative Study	United Kingdom	Wrist	Fitbit Charge HR		✓
Imtiaz et al [49]	Dementia, BPSD	Design experiment	United States	Wrist	Empatica E4	✓	
Iwata et al [50]	Dementia	Control study	Japan	Wrist	Omron wristband	✓	
Kamil et al [51]	Alzheimer	Feasibility study	United States	Waist	Opal	✓	
Kikhia et al [52]	Dementia, BPSD	Feasibility study	Sweden	Wrist	The Philips DTI-2	✓	
Kikhia et al [53]	Dementia, BPSD	Feasibility study	Sweden	Wrist	The Philips DTI-2	✓	
Klimek et al [54]	Dementia, BPSD	Scoping review	The Netherlands	N/A	N/A	✓	✓
Kuzmik et al [55]	Dementia	Feasibility study	United States	Wrist	MotionWatch 8	✓	
Larnyo et al [56]	Dementia	Cross-sectional survey	China	N/A	N/A		✓
Lastovicka-Medin [57]	Dementia	Research design	Montenegro	N/A	e-Textile (self-developed)	✓	✓
Lee et al [58]	Dementia	Qualitative Study	South Korea	Head	VR Device with EEG[f] Headset	✓	
Li et al [59]	Alzheimer	Research design	China	N/A	e-Textile (self-developed)	✓	✓
Mahoney and Mahoney [60]	Alzheimer	Interviews	United States	N/A	N/A		✓
Marcén et al [61]	Dementia, agitation	Feasibility study	Spain	Wrist	GENEactive Original	✓	
Megges et al [18]	Dementia	Comparative study	Germany	Wrist	HIMATIC GPS Uhr Alpha; Re-SOS-2—Die Notfalluhr;	✓	✓
Mohamedali and Matoorian [62]	Dementia	Research study	United Kingdom	N/A	N/A		✓
Mohammed and Mohammed [63]	Dementia	Design research	United States	N/A	N/A	✓	✓
Murphy et al [64]	Dementia	Cross-sectional study	United Kingdom	Arm	Sensewear Armband	✓	
Musaeus et al [65]	Alzheimer	Feasibility study	Denmark	Ear	Ear EEG (self-developed)	✓	✓
Nieroda et al [25]	N/A	Qualitative study	United Kingdom	Wrist	N/A		✓
Oh and Gross [66]	Dementia	Design research	United States	Feet	Awareable steps (self-developed)	✓	✓
O'Sullivan et al [67]	Dementia	Mixed methods	The Netherlands	Wrist	Fitbit Charge 3	✓	✓
Peeters et al [68]	Dementia, BPSD	Qualitative study	The Netherlands	Wrist	Empatica E4	✓	✓

XSL•FO
RenderX

Literature reference	Study domain	Study design	Country	Body location	Wearable technology applied	Practical application	Aesthetics value
Rings et al [69]	Dementia	Qualitative Study	Germany	Head	VR Device with crafted Camera	✓	✓
Schneider and Henneberger [12]	Dementia	N/A	Austria	Wrist	Laipac S911	✓	✓
Silva et al [70]	Alzheimer	Experimental study	Portugal	Chest	SenseCam	✓	
Spasojevic et al [13]	Agitation and aggression, BPSD	Pilot study	Canada	Wrist	Empatica E4	✓	
Staab et al [71]	Dementia	Experimental study	Germany	Wrist	Apple Watch Series 6	✓	
Stavropoulos et al [72]	Alzheimer	Public involvement activities	Greece	Wrist	N/A	✓	✓
Stavropoulos et al [73]	Dementia	Design study	Greece	Wrist	The Philips DTI-2, UP24	✓	
Tervonen et al [74]	Dementia	Design study	Finland	Neck	Self-developed sensor	✓	
Thorpe et al [75]	Dementia	Feasibility study	Denmark	Wrist	Sony SmartWatch 3	✓	
Thorpe et al [76]	Dementia	User-centered approach	Denmark	Wrist	Sony SmartWatch 3	✓	
Thorpe et al [77]	Dementia	Mobility measurement module	Denmark	Wrist	Sony SmartWatch 3	✓	
Treadaway and Kenning [78]	Dementia	Design research	United Kingdom	N/A	e-Textile (self-developed)	✓	✓
Wallace et al [79]	Dementia	Cocreative, design-led	United Kingdom	N/A	Self-developed jewelry		✓
Woodberry et al [5]	Alzheimer	Participatory study	United Kingdom	Neck	SenseCam	✓	
Yokoi et al [80]	Dementia, BPSD	Exploratory study	Japan	Finger	Ring		✓
Zeng et al [81]	Dementia	End-to-end deep learning method	Japan	Head	EEG cap (self-developed)	✓	✓

[a]N/A: not applicable.

[b]BPSD: behavioral and psychological symptoms of dementia.

[c]VR: virtual reality.

[d]HMD: head-mounted display.

[e]HR: heart rate.

[f]EEG: electroencephalogram.

A total of 63 studies [4,5,11-15,18,24-26,30-81] met the eligibility criteria, showcasing a wide variety of wearable devices used on different parts of the body from head to toe. The wrist was the most frequently targeted area, with 37 studies [4,12-15,18,24-26,30,32,34-36,38-40,42,43,45-50,52,53,55,61,67,68,71-73,75-77] focusing on wrist-worn devices that often feature display screens and 3-axis accelerometers, mainly used to record motion data [47]. Among these products, the Empatica E4 was the most tested device for monitoring physiological signals like heart rate variability and electrodermal activity [47], especially for everyday activity tracking and agitation detection [13,30,34,47,49,68]. The Philips Discrete Tension Indicator has also been appraised in dementia care to monitor physical indicators and stress levels [47,52,53,73]. Likewise, the AX3 Axivity device has been studied for its feasibility in dementia care, especially for gait analysis [32,33,36]. Beyond wrist devices, an armband with a triaccelerometer has been used to monitor daily nutrition intake in people with dementia [64]. Waist-worn sensors have been tested for detecting wandering behaviors [51]. Body-worn SenseCams, worn on the neck or chest, show promise for memory stimulation in patients with Alzheimer disease [5,71]. In addition, a neck-worn laser device

XSL•FO
RenderX

aids navigation [74] while footwear sensors have been explored for alerting and monitoring purposes [66]. For headwear, virtual reality (VR) devices are seen as merging innovative technologies that often offer rich sensory engagement, promoting active physical interaction for people with dementia, such as reaching out with hands or feet and rotating the body while exploring an immersive 3D environment [31,37,41,58,69].

Around 35 studies [12,18,24-26,30-32,37-39,41-46,48,54,56,57,59,60,62,63,65-69,72,78-81] were assessed for alignment with 3 core themes of the framework, namely aesthetic experience, experience of meaning, and emotional experience [20]. These themes were developed to examine how wearables engage users beyond functionality, encompassing aesthetic involvement, self-image, and emotional connection. They were informed by theoretical foundations in product experience and patterns identified in the extracted data.

A detailed breakdown of the studies that addressed esthetic considerations is presented in Table 2, with a clear distribution presented under three referenced thematic categories.

XSL•FO
RenderX

Table . Distribution of studies related to aesthetic value.

Literature reference	Applied product	Aesthetics experience	Experience of meaning	Emotional experience
Schneider and Henneberger [12]	• Laipac S911BL • Xexun TK202 • Xexun TK203 • Xexun TK102-2 • CRT 19N CY 2130 • Teletonika GH 3000	• Recommended size: 50×50×20 mm, ≤100 g. • Panic button essential. • Bracelet or watch design preferred	—[a]	—
Megges et al [18]	• HIMATIC GPS Uhr Alpha, • ReSOS-2	• Preferred: button-based, minimal design.	• Design perceived as stigmatizing.	—
Choi and Kim [24]	• Smartwatch	—	• Smartwatches are seen as innovative, luxury fashion items.	—
Nieroda et al [25]	• Wearable devices	—	• Wearables are viewed as hybrid products: smart and innovative.	—
Apple [26]	• Apple Watch Series 6	—	• Apple Watch linked to a high-tech, healthy lifestyle. • Fashion ties reinforce the stylish image.	—
Amato et al [30]	• Empatica E4	• Tactile features aid engagement. • E4 is seen as thick, tight, and uncomfortable.	—	—
Appel et al [31]	• Samsung Gear VR[b]	• Device linked to heaviness, nausea, and discomfort • Rich sensory engagement with movement aided by sound cues.	—	• Reported relaxation, happiness, stress relief, and benefits linked to brief wear rather than extended use.
Mc Ardle et al [32]	• Axivity AX3	• The device is comfortable and easy to wear.	—	—
Cheung et al [37]	• Meta Quest 2	• Immersive art or crafting experience with colorful tools.	—	• Enabled expression and engagement with personalized, meaningful aspects.
Cheung et al [38]	• Wrist-worn accelerometer devices	• The issue of comfort has been addressed.	—	—
Chung et al [39]	• Garmin Vivoactive HR[c]	• Preferred watch-like monitor with on-screen display.	—	—
Davis and Sikorskii [41]	• VR Glasses or Goggles • Applied Science Laboratory	• Clear glasses with joysticks supported active body engagement.	—	—
Farina et al [42]	• GeneActiv	• Comfort, size, weight, usability, safety, security, and durability valued.	• Watch-like design seen as nondistinctive, causing issues for existing watch wearers.	—

XSL•FO
RenderX

Literature reference	Applied product	Aesthetics experience	Experience of meaning	Emotional experience
Favela et al [43]	• Fitbit Charge 2 • Fitbit Alta	—	• Preferred discreet design. • Lacked self-image value.	—
Felber et al [44]	• Wearable devices	• Existing wearables seen as unappealing and clunky.	—	• Wearables caused feelings of incapacity. • Complex functions discouraged users.
Guu et al [45]	• Wearable devices	—	—	• Older adults underrepresented in wearables. • Ageism needs addressing.
Guu et al [46]	• GeneActiv	• Suggested: a more feminine device with a time display.	• Watch-like design confused participants already wearing a watch.	—
Hassan et al [48]	• Axivity AX3 • Fitbit Charge HR; MOTO 360 • Garmin Vivogit 2 • Misfit Speedo Shine • Withings Activite Pop • Withings Pulse Ox	• Water resistance and ease of use were identified as critical factors.	• The device may draw unwanted attention.	• The device may cause the wearer to feel incapable.
Klimek et al [54]	• Empatica E4	• Watch-like design recommended.	—	—
Larnyo et al [56]	• Wearable devices	• Discomfort was reported as an issue by wearers.	—	• Reported issues: frustration, confusion, embarrassment, and anxiety.
Lastovicka-Medin [57]	• Self-developed sensored textile	• e-Textiles responded to touch, fostering engagement and connecting past and future knowledge.	—	—
Li et al [59]	• Self-developed sensored textile	• Soft, colorful textiles promoted communication and tactile engagement.	—	—
Mahoney and Mahoney [60]	• Wearable devices	• Recommended: waterproof, jewelry-like, easy to clean, and hypoallergenic.	• Products were perceived as institutional and intrusive.	• The products were perceived as diminishing wearer dignity.
Mohamedali and Matoorian [62]	• Wearable devices	• Preferred: comfort and varied straps or mounts (colors, sizes, and materials).	• Preferred: less obtrusive, discreet design.	• Wearable presentation undermines users' sense of capability.
Mohammed and Mohammed [63]	• Wearable devices	• A "watch-like" design is preferred.	—	—
Musaeus et al [65]	• Self-proposed ear EEG[d]	• Recommended: • Smaller amplifier, improved textures. • Mild ear discomfort reported.	• Wearing the device may draw unwanted attention.	• Wearables may reveal vulnerability, discouraging use.

XSL•FO
RenderX

Literature reference	Applied product	Aesthetics experience	Experience of meaning	Emotional experience
Oh and Gross [66]	• Self-developed shoes	• Soft leathers were incorporated for comfort and visual appeal.	—	—
O'Sullivan et al [67]	• Fitbit Charge 3	—	• Participants may forget or avoid the watch-like device. • Adds little user value.	—
Peeters et al [68]	• Empatica E4	• Preferred: water-resistant, easy-to-use, comfortable, personalized (colors or jewelry). • E4 caused tightness and discomfort.	—	• Device's appearance raised fear of unwanted public attention.
Rings et al [69]	• VR Device with Crafted Camera	• Wearers performed active movements, including body exercises.	—	• Game-based designs offered immersive experiences, supporting personal stories and memories.
Stavropoulos et al [72]	• Wearable devices	• Preferred: water-resistant, soft materials, and varied colors.	—	—
Treadaway and Kenning [78]	• Self-developed textile	• Rich-textured, soft design evoked attachment and affection.	—	—
Wallace et al [79]	• Self-developed jewelry	• Self-made jewelry with personalized forms and varied materials.	—	• Personalized jewelry fostered emotional attachment.
Yokoi et al [80]	• Jewelry	• Aesthetically designed stone rings.	—	• Stone rings enhanced confidence, beauty, and dignity.
Zeng et al [81]	• Self-developed head EEG	• Textile EEG cap designed for stable skin contact and comfort.	—	—

[a]Not applicable.

[b]VR: virtual reality.

[c]HR: heart rate.

[d]EEG: electroencephalogram.

Aesthetic Experience

Aesthetic experiences engage the senses and shape user interactions with the product [20]. On this level, the aesthetic connections start by giving pleasure to the wearer's sensual feelings, whether it is an attractive appearance or an amusing sound. In the reviewed studies, wrist-worn wearables, mostly designed as smartwatches with screens and strap bands, such as Basis B1, Sony SmartWatch 3, Apple Watch Series 6, and Pebble Watch, are the most commonly used devices in dementia care [35,40,47,50,71,75-77], although some have watch-like forms but without screens [55,61]. The aesthetics of having the wrist-worn devices with a "watch-like" design are also recommended by several studies [12,39,54,63]. In the considerations of tactile senses, materials with waterproof quality have been discussed as necessary to cope with different environments [48,60,72]. Comfort is another critical factor that greatly influences users' opinions toward assistive technology while the product is intimately engaging with the wearer's body [38,56]. It is also suggested that materials should be hypoallergenic and washable to ensure safety and comfort [60,72]. However, overly secure fastenings, while beneficial for data consistency, can increase anxiety or agitation, emphasizing the need for user-friendly designs [68]. Regarding the wearable's weight and size, it is mentioned that the artifact should weigh under 100 g and not exceed dimensions of 50x50x20 mm [12], with miniaturized designs preferred for discretion [18,60,65]. Tight-fitting devices, such as the Empatica E4, have been criticized for causing discomfort and skin marks [68]. Concerning the personalization of wearable design

XSL•FO
RenderX

aesthetics, feminine appearance [46], different colors and various material options are preferred [12,44,62,68,72].

In the development of VR experiences, comfort has also been a concern, with reports of simulation sickness and heaviness, including mild nausea associated with wearing headsets [31]. Despite these issues, to further enhance sensory immersion, studies have combined VR headsets with joysticks, tangible objects, and earphones, allowing users to experience not only visual pleasures but also tactile and auditory stimuli, fostering a more embodied and present experience [31,41,69].

Soft materials and objects with textured surfaces are highly valued in the responses of individuals with dementia [57,59,66,78,81]. While footwear naturally becomes part of people's daily outdoor routines, devices like smart shoes made from soft brown leather demonstrate how fashionable designs can seamlessly blend into everyday life while meeting functional needs [66]. For textiles, a study has begun designing a hat using soft materials to cover EEG sensors for emotional detection [81]. A co-design study's examination revealed that interacting with tactile objects enhances emotional well-being and social connections for those affected [78]. Personal reminiscence can also be triggered through the crafting process with textiles [57]. It is also found that a handkerchief with varied colors, designed to assist pain communication for individuals with Alzheimer disease, has been decorated and accepted [59].

Experience of Meaning

A meaningful experience involves a user's interaction with a product that engages deeper cognitive processes, such as interpretation, memory recall, and the formation of personal associations. Symbolism in design influences how users envision their lives with an object, extending its impact beyond simple sensory pleasure. For example, luxury homeware may symbolize comfort and a high standard of living [20]. Similarly, wrist-worn devices, especially those styled like watches, often represent innovation and intelligence by combining functions, such as time display with health tracking [24,25]. Unlike traditional devices, smartwatches like the Apple Watch are more closely integrated with the body. The design, featuring customizable bands and collaborations with brands like Hermès, has strengthened their dual role as both technical tools and luxury fashion accessories [26]. In dementia care, however, this symbolic link to health and lifestyle may pose new challenges.

In the context of dementia, individuals often prefer discreet designs with smaller displays, mainly to avoid unwanted attention and public anxiety [43]. Instead of conveying wellness or a high quality of life as intended by watch-like wearable designs, such devices can cause embarrassment and discomfort due to their visibility in public settings [60,62]. This reveals a key conflict in the symbolic meaning of smart technologies. While they are often linked to innovation and well-being, for those already dealing with health issues, these devices can express fear and a wish to hide. A visible device may be seen by others as a sign of incapacity, making wearers feel vulnerable or stigmatized [65]. The perception that these devices give an "institutional look" to people with dementia has reinforced negative symbolism and may reduce the user's sense of dignity and autonomy [60]. Moreover, for many who already own watches with personal significance, the addition of a second watch focused solely on health monitoring can cause confusion and resistance, especially when users are unsure why they are wearing 2 watches [42,46,67].

Emotional Experience

Emotional experience arises from an individual's interpretation of a product rather than from the product itself [20]. In dementia care, preferences for discreet presentation reflect concerns that negative emotional responses to smart wearables are less about the devices themselves and more about what they represent, including monitoring, dependency, and declining health. Wearable technologies have thus been associated with symbols of vulnerability and constant surveillance [65]. As quality of life in later years is closely linked to autonomy and social connectedness, the current design of wearable devices can act as a social barrier by unintentionally signaling a need for assistance and thereby diminishing confidence and willingness to engage socially [31,44,45,48,56,60,62,65]. This underscores the importance of exploring how potential emotionally threatening associations could be improved or transformed through more positive human-product interactions by examining cases where people with dementia experience pleasant emotional responses to wearable objects.

In studies reporting positive emotional responses among people with dementia, both aesthetic experience and personal meaning appear to play crucial roles in shaping the likelihood of wearable adoption. From an aesthetic perspective, recent evidence highlights a preference for wearables designed as fashion accessories, particularly jewelry-like forms, which are perceived as more acceptable than watch-like designs [60,68]. While emotions play a key role in forming attachment to products, studies have shown that women with dementia displayed increased grace and self-awareness when wearing gemstone rings provided during the study [80]. Although based on a small sample, participants expressed positive feelings by using appreciative words, such as "fancy" and "beautiful," highlighting the empowering potential of aesthetics through jewelry.

In relation to the experience of meaning, products that incorporate personalized stories or narratives have been shown to elicit positive emotional responses. For example, a co-design project developed personalized jewelry pieces for a woman with dementia, who described them as emotionally meaningful and deeply personal. This case highlights the importance of aesthetic and symbolic connections in the design of wearable objects [79]. As a head-mounted technology, VR has shown promise in providing immersive experiences for people with dementia, supporting momentary engagement and self-expression through interaction with virtual reality environments [37,69]. Studies have enabled users to personalize scenes, such as cocreating meaningful spaces from physical to virtual reality presentation [37] or having the flexibility to take photos that one likes under a built environment [69].

While centered on fostering positive experiences for people with dementia through the aesthetic dimension, these cases highlight the importance of wearable technologies that are both visually appealing and capable of personalization. Such design approaches may encourage positive emotional connections and

XSL•FO
RenderX

facilitate the development of long-term attachment to the product within dementia care.

Discussion

Principal Findings

This study underscores the crucial role of aesthetics in influencing how people with dementia adopt wearable devices. Key factors include the sensory experience of the product, symbolic associations that shape perceptions of empowerment versus stigma, and the significance of personal narratives in fostering emotional connections with the device. All of these should be carefully considered in the design for dementia and inclusive care.

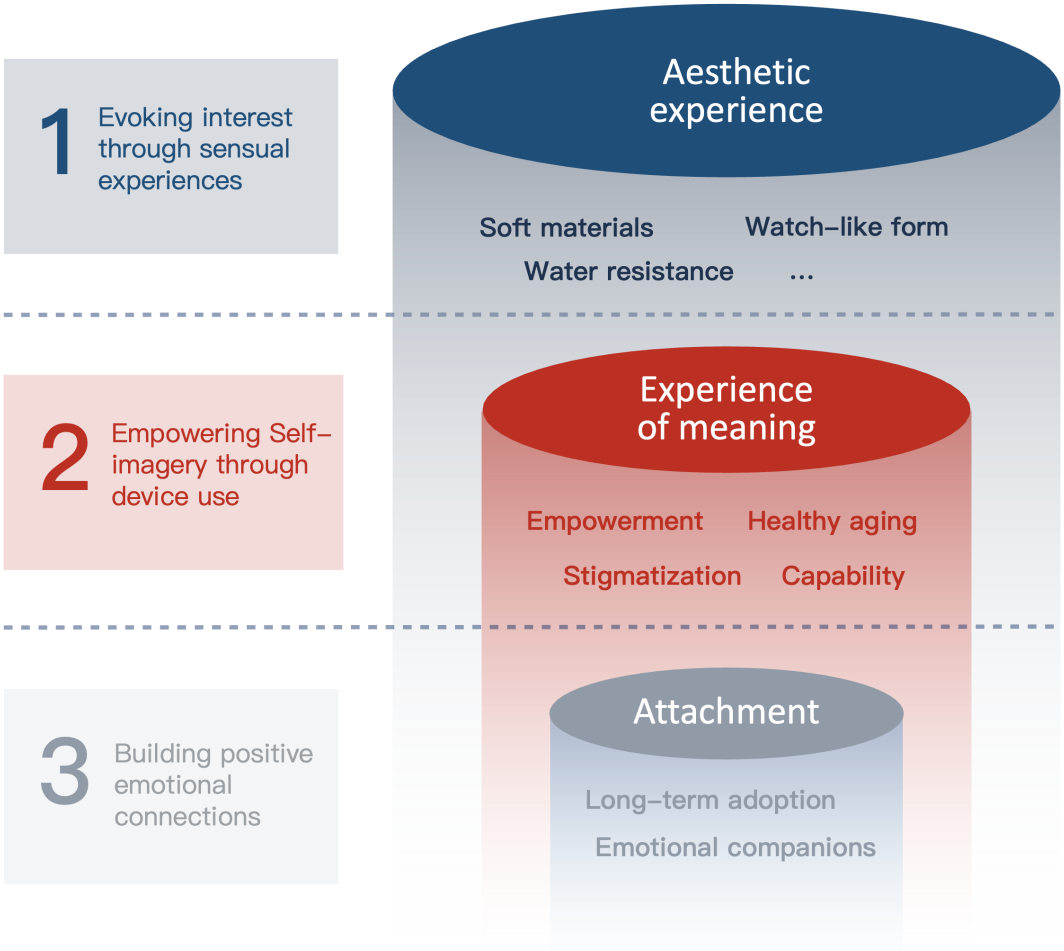
New Conceptual Framework for Dementia

Building on an existing framework that links aesthetic, emotional, and meaningful experiences in human-product interaction [20], this review introduces a new conceptual model (Figure 3) based on the perspectives of people with dementia toward current wearable technologies and studies exploring the aesthetics of wearable objects in dementia care. While the original framework highlights the potential interlinking of the 3 elements, our findings indicate that they need to be directly interconnected in studying and designing for a better experience between people with dementia and wearable technologies. This begins with generating interest through aesthetic appeal and addressing concerns about uncomfortable or unappealing designs. It then moves to shaping a positive self-image that fosters empowerment rather than incapacity. Finally, it culminates in building lasting emotional connections with the device, encouraging continued use and supporting both physical and mental well-being. To encourage acceptance, long-term use, and emotional engagement, future design of wearable devices should consider both aesthetic appeal and the personal significance they may hold for individuals with dementia.

As illustrated in Figure 3, the interaction between people with dementia and wearable devices is conceptualized in 3 key stages. The first stage involves an initial attraction driven by the product's sensory appeal, where the aesthetic experience is primarily based on sensual appreciation. This early interest encourages deeper engagement in the second stage, where individuals begin to form a personal connection with the device. At this point, the product should empower users by supporting health, offering comfort, and providing clear benefits without reinforcing stigma or perceptions of incapability. Together, these first 2 stages create space for personal interpretation and self-expression, allowing individuals to build meaningful narratives of their own with the device and experience emotional resonance. Ideally, this process culminates in a third stage of attachment, in which the individual expresses a sense of companionship with the product.

XSL•FO
RenderX

Figure 3. User-wearable product framework for dementia care.



Implications for Future Wearable Technologies

While this study has identified key preferences, such as the size and weight of wrist-worn devices, from the perspectives of people with dementia, the wider aesthetic impacts of wearable technology remain underexplored. This includes how design choices across different body locations and device types (eg, VR-based technologies or crafted objects) may support therapeutic experiences, such as through artistic engagement [37] and memory recollection [82]. While some females with dementia express concerns about the lack of feminine appeal in current wearable designs, studies examining aesthetic rings and customized jewelry may reveal potential. This furthermore highlights the need for culturally appropriate design, as not all will identify with “feminine” aesthetics for inclusive design. In addition, soft, richly textured materials have shown promise in offering comfort and sensory reassurance, especially as memory loss advances and tactile needs increase [83]. Integrating these materials into long-term wearable designs may lessen the sense of detachment or discomfort often linked to medicalized or institutional-looking devices, instead fostering familiarity and emotional connection.

Together, these findings highlight the need for interdisciplinary collaboration between artists, designers, and technologists to develop devices that are not only functionally effective but also personally meaningful and aesthetically engaging, fostering long-term attachment and acceptance among users with dementia.

The Need to Build a Positive Image of the Self

As shown in Figure 3, although current wearable technologies have increasingly focused on aesthetic enhancement and incorporated various design elements, such as the use of rich materials and multisensory features explored through innovative research on soft textiles and immersive VR tools, there remains limited understanding of how purely sensory experiences contribute to sustained emotional attachment or long-term adoption among individuals with dementia. A key next step in advancing these technologies is to deepen the understanding of how individuals interpret and form personal associations with such devices.

From a contrasting perspective, it is important to consider how people with dementia may perceive the idealized imagery of smart technologies and healthy aging narratives as alienating or even stigmatizing [43]. This highlights the need to evaluate

XSL•FO
RenderX

the value and benefits of wearable technologies from the perspective of the wearers. To this end, co-design and participatory workshops may offer valuable opportunities to explore how people with dementia articulate their needs for monitoring sensors and technologies with varied purposes, their preferences regarding the form and feel of devices, and how these technologies can better support their lived experiences [33,84,85]. Rather than focusing solely on improving the feasibility of the products, developers and companies must consider where on the body these devices are preferred, how they should be integrated, and how much control individuals with dementia wish to retain, particularly concerning assistive functions, such as location tracking and close body monitoring.

The Continuance of Self and Personalization

The disruption of selfhood is a longstanding concern in dementia care, with memory loss often seen as the key marker of identity decline. However, recent perspectives highlight that selfhood can persist through embodied actions, including everyday gestures, routines, and social behaviors, that continue even as memory fades [21]. These expressions, along with interactions with personal objects, reflect a form of identity beyond words.

In this context, the standardized, clinical aesthetics of wrist-worn wearables reveal a gap in supporting personal identity and emotional resonance. While crafted artifacts, such as bespoke jewelry, can foster attachment and convey empowering symbolism, most data-driven wearables prioritize function over individual expression [79]. This raises the need to reimagine wearables as more than passive sensing tools. Inspired by how VR enables users to construct personalized narratives and environments, wearable technologies could similarly support self-expression and identity continuity. Rather than static objects, they could become symbolic carriers of personal meaning, affirming agency and dignity through participatory, aesthetic engagement shaped from the moment of wearing.

Strengths and Limitations

This review adopted a multidatabase search to capture a broad range of stakeholder perspectives on wearable technologies in dementia care. Two reviewers were involved to support a more flexible interpretation of the literature. However, as a critical interpretive synthesis, early-stage conceptual studies may have been excluded, potentially limiting insights into future design directions.

Conclusion

This study conducted a critical interpretive synthesis, revealing a critical gap in the aesthetic design of wearable technologies for dementia care, which affects user acceptance and emotional engagement. Despite functional advances, adoption remains limited due to a lack of engagement in considering building personal meaning and sensory appeal. By analyzing 3 key themes influencing user-product interaction, this study proposes a new conceptual framework for understanding the role of aesthetics in dementia-focused wearable design. The findings point to the potential of aesthetic and inclusive design in shaping more engaging, acceptable, and human-centered wearable technologies for dementia care.

Acknowledgments

The authors used AI-based language models to assist with grammar and language refinement of the final manuscript. All content and interpretations are the authors' own.

Data Availability

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

Authors' Contributions

YXW led the conceptualization, data curation, formal analysis, investigation, methodology development, resource management, software use, validation, visualization, and drafting of the original manuscript. JR contributed to the methodology design and was responsible for reviewing and editing the manuscript. DA supported supervision of the project. WL contributed to support data analysis and validation, provided overall supervision, review, and editing of the manuscript.

Conflicts of Interest

None declared.

References

1. Dementia. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2025-01-07]
2. David MCB, Kolanko M, Del Giovane M, et al. Remote monitoring of physiology in people living with dementia: an observational cohort study. JMIR Aging 2023 Mar 9;6:e43777. [doi: 10.2196/43777] [Medline: 36892931]
3. Desai C, Dove E, Nanthakumar J, et al. Assistive technology to support dementia management: Protocol for a scoping review of reviews. JMIR Res Protoc 2024 Nov 11;13:e57036. [doi: 10.2196/57036] [Medline: 39527790]

XSL•FO
RenderX

4. Cullen A, Mazhar MKA, Smith MD, Lithander FE, Ó Breasail M, Henderson EJ. Wearable and portable GPS solutions for monitoring mobility in dementia: a systematic review. Sensors (Basel) 2022 Apr 27;22(9):3336. [doi: 10.3390/s22093336] [Medline: 35591026]
5. Woodberry E, Browne G, Hodges S, Watson P, Kapur N, Woodberry K. The use of a wearable camera improves autobiographical memory in patients with Alzheimer's disease. Memory 2015 Apr 3;23(3):340-349. [doi: 10.1080/09658211.2014.886703]
6. Andrade-Moraes CH, Oliveira-Pinto AV, Castro-Fonseca E, et al. Cell number changes in Alzheimer's disease relate to dementia, not to plaques and tangles. Brain (Bacau) 2013 Dec;136(Pt 12):3738-3752. [doi: 10.1093/brain/awt273] [Medline: 24136825]
7. Shu S, Woo BKP. Use of technology and social media in dementia care: current and future directions. World J Psychiatry 2021 Apr 19;11(4):109-123. [doi: 10.5498/wjp.v11.i4.109] [Medline: 33889536]
8. Machado SD, Tavares JDR, Barbosa JLV. Technologies for monitoring patients with Alzheimer's disease: a systematic mapping study and taxonomy. J Ambient Intell Smart Environ 2024;16(1):3-22. [doi: 10.3233/AIS-220407]
9. Nesbitt C, Gupta A, Jain S, Maly K, Okhravi HR. Reliability of wearable sensors to detect agitation in patients with dementia. 2018 May 16 Presented at: ICBBT '18; May 16-18, 2018; Amsterdam Netherlands p. 73-77 URL: https://dl.acm.org/doi/proceedings/10.1145/3232059 [accessed 2025-10-17] [doi: 10.1145/3232059.3232067]
10. Perez-Valero E, Morillas C, Lopez-Gordo MA, Minguillon J. Supporting the detection of early Alzheimer's disease with a four-channel EEG analysis. Int J Neur Syst 2023 Apr;33(4):2350020. [doi: 10.1142/S0129065723500211]
11. Anwar U, Arslan T, Hussain A, Russ TC, Lomax P. Design and evaluation of a wearable multimodal RF sensing system for vascular dementia detection. IEEE Trans Biomed Circuits Syst 2023 Oct;17(5):928-940. [doi: 10.1109/TBCAS.2023.3282350] [Medline: 37267143]
12. Schneider C, Henneberger S. Electronic spatial assistance for people with dementia: choosing the right device. Technologies (Basel) 2014;2(2):96-114. [doi: 10.3390/technologies2020096]
13. Spasojevic S, Nogas J, Iaboni A, et al. A pilot study to detect agitation in people living with dementia using multi-modal sensors. J Healthc Inform Res 2021 Sep;5(3):342-358. [doi: 10.1007/s41666-021-00095-7] [Medline: 35419507]
14. Alam R, Gong J, Hanson M, et al. Motion biomarkers for early detection of dementia-related agitation. 2017 Jun 23 Presented at: DigitalBiomarkers '17: Proceedings of the 1st Workshop on Digital Biomarkers; Niagara Falls, NY p. 15-20 URL: https://dl.acm.org/doi/proceedings/10.1145/3089341 [accessed 2025-10-17] [doi: 10.1145/3089341.3089344]
15. Alam R, Anderson M, Bankole A, Lach J. Inferring physical agitation in dementia using smartwatch and sequential behavior models. Presented at: 2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI); Mar 4-7, 2018; Las Vegas, NV p. 170-173. [doi: 10.1109/BHI.2018.8333396]
16. Bastoni S, Wrede C, da Silva MC, et al. Factors influencing implementation of eHealth technologies to support informal dementia care: umbrella review. JMIR Aging 2021 Oct 8;4(4):e30841. [doi: 10.2196/30841] [Medline: 34623314]
17. Chien SY, Zaslavsky O, Berridge C. Technology usability for people living with dementia: concept analysis. JMIR Aging 2024 Jul 3;7:e51987. [doi: 10.2196/51987] [Medline: 38959053]
18. Megges H, Freiesleben SD, Rösch C, Knoll N, Wessel L, Peters O. User experience and clinical effectiveness with two wearable global positioning system devices in home dementia care. A&D Transl Res & Clin Interv 2018 Jan;4(1):636-644. [doi: 10.1016/j.trci.2018.10.002]
19. Peterson CM, St Louis RM, Flannagan C. Enhancing enrollment and adherence in long-term wearable research on dementia: qualitative systematic review and meta-synthesis. JMIR Aging 2025 Jul 31;8(1):e63768. [doi: 10.2196/63768] [Medline: 40743521]
20. Desmet PMA, Hekkert P. Framework of product experience. Int J Des 2007;1(1):57-66 [FREE Full text]
21. Kontos PC. Embodied selfhood in Alzheimer's disease: rethinking person-centred care. Dementia (London) 2005 Nov;4(4):553-570. [doi: 10.1177/1471301205058311]
22. Bacsu JD, Fraser S, Chasteen AL, et al. Using Twitter to examine stigma against people with dementia during COVID‐19: infodemiology study. JMIR Aging 2022 Mar 31;5(1):e35677. [doi: 10.2196/35677] [Medline: 35290197]
23. Dixon-Woods M, Cavers D, Agarwal S, et al. Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. BMC Med Res Methodol 2006 Jul 26;6(1):35. [doi: 10.1186/1471-2288-6-35] [Medline: 16872487]
24. Choi J, Kim S. Is the smartwatch an IT product or a fashion product? A study on factors affecting the intention to use smartwatches. Comput Human Behav 2016 Oct;63:777-786. [doi: 10.1016/j.chb.2016.06.007]
25. Nieroda ME, Mrad M, Solomon MR. How do consumers think about hybrid products? Computer wearables have an identity problem. J Bus Res 2018 Aug;89:159-170. [doi: 10.1016/j.jbusres.2018.04.024]
26. Apple watch series 6 delivers breakthrough wellness and fitness capabilities. Apple. 2020 Sep 15. URL: https://www.apple.com/uk/newsroom/2020/09/apple-watch-series-6-delivers-breakthrough-wellness-and-fitness-capabilities/ [accessed 2025-01-07]
27. Depraetere J, Vandeviver C, Keygnaert I, Beken TV. The critical interpretive synthesis: an assessment of reporting practices. Int J Soc Res Methodol 2021 Nov 2;24(6):669-689. [doi: 10.1080/13645579.2020.1799637]
28. Hekkert P. Design aesthetics: Principles of pleasure in design. Psychology Science 2006;48(2):157-172 [FREE Full text]

XSL•FO
RenderX

29. Attribution 4.0 international (CC BY 4.0). Creative Commons. URL: https://creativecommons.org/licenses/by/4.0/ [accessed 2025-10-06]
30. Amato F, Bianchi S, Comai S. CLONE: a promising system for the remote monitoring of alzheimer's patients: an experimentation with a wearable device in a village for alzheimer's care. Presented at: Proceedings of the 4th EAI International Conference on Smart Objects and Technologies for Social Good; Nov 28-30, 2018; Bologna, Italy p. 255-260 URL: https://dl.acm.org/doi/10.1145/3284869.3284906 [accessed 2025-10-17] [doi: 10.1145/3284869.3284906]
31. Appel L, Kisonas E, Appel E, et al. Administering virtual reality therapy to manage behavioral and psychological symptoms in patients with dementia admitted to an acute care hospital: results of a pilot study. JMIR Form Res 2021 Feb 3;5(2):e22406. [doi: 10.2196/22406] [Medline: 33533720]
32. Mc Ardle R, Del Din S, Galna B, Thomas A, Rochester L. Differentiating dementia disease subtypes with gait analysis: feasibility of wearable sensors? Gait & Posture 2020 Feb;76:372-376. [doi: 10.1016/j.gaitpost.2019.12.028]
33. Mc Ardle R, Morris R, Hickey A, et al. Gait in mild Alzheimer's disease: feasibility of multi-center measurement in the clinic and home with body-worn sensors: a pilot study. J Alzheimers Dis 2018;63(1):331-341. [doi: 10.3233/JAD-171116] [Medline: 29614664]
34. Badawi A, Elgazzar K, Ye B, et al. Investigating multimodal sensor features importance to detect agitation in people with dementia. Presented at: 2023 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE); Sep 24-27, 2023; Regina, Canada p. 77-82. [doi: 10.1109/CCECE58730.2023.10288764]
35. Boletsis C, McCallum S, Landmark BF. The use of smartwatches for health monitoring in home-based dementia care. In: Zhou J, Salvendy G, editors. Human Aspects of IT for the Aged Population Design for Everyday Life Lecture Notes in Computer Science: Springer; 2015, Vol. 9194:15-26. [doi: 10.1007/978-3-319-20913-5_2]
36. Chan LLY, Lord SR, Brodie MA. Is maximal or usual walking speed from large scale wrist sensor data better at predicting dementia, depression and death?*. Presented at: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society; Jul 24-27, 2023; Sydney, Australia p. 1-4. [doi: 10.1109/EMBC40787.2023.10340255]
37. Cheung E, Ppali S, Otkhmezuri B, Thomas CL, Covaci A. Co-creating meaningful spaces: stepping into virtual worlds crafted by people living with dementia. Presented at: DIS '23: Designing Interactive Systems Conference; Jul 10-14, 2023; Pittsburgh, PA p. 81-85. [doi: 10.1145/3563703.3596645]
38. Cheung JCW, So BPH, Ho KHM, Wong DWC, Lam AHF, Cheung DSK. Wrist accelerometry for monitoring dementia agitation behaviour in clinical settings: a scoping review. Front Psychiatry 2022;13:913213. [doi: 10.3389/fpsyt.2022.913213] [Medline: 36186887]
39. Chung J, Boyle J, Pretzer-Aboff I, Knoefel J, Young HM, Wheeler DC. Using a GPS watch to characterize life-space mobility in dementia: a dyadic case study. J Gerontol Nurs 2021 Oct;47(10):15-22. [doi: 10.3928/00989134-20210908-03] [Medline: 34590977]
40. Davidoff H, Van den Bulcke L, Vandenbulcke M, et al. Toward quantification of agitation in people with dementia using multimodal sensing. Innov Aging 2022;6(7):igac064. [doi: 10.1093/geroni/igac064] [Medline: 36600807]
41. Davis R, Sikorskii A. The feasibility of using virtual reality and eye tracking in research with older adults with and without Alzheimer's disease. Front Aging Neurosci 2021;13:607219. [doi: 10.3389/fnagi.2021.607219] [Medline: 34248596]
42. Farina N, Sherlock G, Thomas S, Lowry RG, Banerjee S. Acceptability and feasibility of wearing activity monitors in community-dwelling older adults with dementia. Int J Geriatr Psychiatry 2019 Apr;34(4):617-624. [doi: 10.1002/gps.5064] [Medline: 30701592]
43. Favela J, Cruz-Sandoval D, Morales-Tellez A, Lopez-Nava IH. Monitoring behavioral symptoms of dementia using activity trackers. J Biomed Inform 2020 Sep;109:103520. [doi: 10.1016/j.jbi.2020.103520] [Medline: 32783922]
44. Felber NA, Lipworth W, Tian YJA, Roulet Schwab D, Wangmo T. Informing existing technology acceptance models: a qualitative study with older persons and caregivers. Eur J Ageing 2024 Mar 29;21(1):12. [doi: 10.1007/s10433-024-00801-5] [Medline: 38551677]
45. Guu TW, Muurling M, Khan Z, et al. Wearable devices: underrepresentation in the ageing society. Lancet Digit Health 2023 Jun;5(6):e336-e337. [doi: 10.1016/S2589-7500(23)00069-9] [Medline: 37236695]
46. Guu TW, Brem AK, Albertyn CP, Kandangwa P, Aarsland D, Ffytche D. Wrist-worn actigraphy in agitated late-stage dementia patients: a feasibility study on digital inclusion. Alzheimers Dement 2024 May;20(5):3211-3218. [doi: 10.1002/alz.13772] [Medline: 38497216]
47. Harper M, Ghali F. A systematic review of wearable devices for tracking physiological indicators of dementia related difficulties. 2020 Presented at: 2020 13th International Conference on Developments in eSystems Engineering (DeSE); Dec 14-17, 2020; Liverpool, United Kingdom p. 406-411. [doi: 10.1109/DeSE51703.2020.9450793]
48. Hassan L, Swarbrick C, Sanders C, et al. Tea, talk and technology: patient and public involvement to improve connected health 'wearables' research in dementia. Res Involv Engagem 2017;3:12. [doi: 10.1186/s40900-017-0063-1] [Medline: 29062537]
49. Imtiaz D, Anwar Y, Khan A. Wearable sensors and a multisensory music and reminiscence therapies application: to help reduce behavioral and psychological symptoms in person with dementia. Smart Health (2014) 2020 Nov;18:100140. [doi: 10.1016/j.smhl.2020.100140]

XSL•FO
RenderX

50. Iwata A, Kowa H, Tsuji S. Monitoring daily life activity shows less activity among female dementia patients. Neurol Clin Neurosci 2013 May;1(3):91-95. [doi: 10.1111/j.2049-4173.2013.00029.x]
51. Kamil RJ, Bakar D, Ehrenburg M, et al. Detection of wandering behaviors using a body‐worn inertial sensor in patients with cognitive impairment: a feasibility study. Front Neurol 2021;12:529661. [doi: 10.3389/fneur.2021.529661] [Medline: 33776875]
52. Kikhia B, Stavropoulos TG, Andreadis S, et al. Utilizing a wristband sensor to measure stress level for people with dementia. Sensors (Basel) 2016 Nov 24;16(12):1989. [doi: 10.3390/s16121989] [Medline: 27886155]
53. Kikhia B, Stavropoulos TG, Meditskos G, et al. Utilizing ambient and wearable sensors to monitor sleep and stress for people with BPSD in nursing homes. J Ambient Intell Human Comput 2018 Apr;9(2):261-273. [doi: 10.1007/s12652-015-0331-6]
54. Klimek A, Mannheim I, Schouten G, Wouters EJM, Peeters MWH. Wearables measuring electrodermal activity to assess perceived stress in care: a scoping review. Acta Neuropsychiatr 2025;37:1-11. [doi: 10.1017/neu.2023.19]
55. Kuzmik A, Resnick B, Cacchione P, Arendacs R, Boltz M. Physical activity in hospitalized persons with dementia: feasibility and validity of the MotionWatch 8. J Aging Phys Act 2021 Oct 1;29(5):852-857. [doi: 10.1123/japa.2020-0275] [Medline: 33863850]
56. Larnyo E, Dai B, Larnyo A, et al. Impact of actual use behavior of healthcare wearable devices on quality of life: a cross-sectional survey of people with dementia and their caregivers in Ghana. Healthcare (Basel) 2022 Jan 30;10(2):275. [doi: 10.3390/healthcare10020275] [Medline: 35206890]
57. Lastovicka-Medin G. Journey through dementia with e-textile as medium for assisting memory recall: tactile dialogues. In: Jozwiak L, Stojanovic R, editors. Presented at: 2019 8th Mediterranean Conference on Embedded Computing (MECO); Jun 10-14, 2019; Budva, Montenegro p. 624-627. [doi: 10.1109/MECO.2019.8760169]
58. Lee B, Lee T, Jeon H, et al. Synergy through integration of wearable EEG and virtual reality for mild cognitive impairment and mild dementia screening. IEEE J Biomed Health Inform 2022 Jul;26(7):2909-2919. [doi: 10.1109/JBHI.2022.3147847] [Medline: 35104235]
59. Li Y, Bai L, Mao Y, et al. Rethinking pain communication of patients with Alzheimer's disease through E-textile interaction design. Front Physiol 2023;14:1248893. [doi: 10.3389/fphys.2023.1248893]
60. Mahoney EL, Mahoney DF. Acceptance of wearable technology by people with Alzheimer's disease: issues and accommodations. Am J Alzheimers Dis Other Demen 2010 Sep;25(6):527-531. [doi: 10.1177/1533317510376944] [Medline: 20702501]
61. Marcén AC, Carro J, Monasterio V. Wearable monitoring for the detection of nocturnal agitation in dementia. 2016 Presented at: International Conference on Pervasive and Embedded Computing; Jul 25-27, 2016; Lisbon, Portugal p. 63-69. [doi: 10.5220/0005938500630069]
62. Mohamedali F, Matoorian N. Support dementia: using wearable assistive technology and analysing real-time data. Presented at: 2016 International Conference on Interactive Technologies and Games (iTAG); Oct 26-27, 2016; Notthingham, United Kingdom p. 50-54. [doi: 10.1109/iTAG.2016.15]
63. Mohammed A, Mohammed GA. An artificially intelligent wearable device for dementia patients. Presented at: 2019 International Conference on Machine Learning and Data Engineering (iCMLDE); Dec 2-4, 2019; Taipei, Taiwan p. 25-29. [doi: 10.1109/iCMLDE49015.2019.00016]
64. Murphy J, Holmes J, Brooks C. Measurements of daily energy intake and total energy expenditure in people with dementia in care homes: the use of wearable technology. J Nutr Health Aging 2017;21(8):927-932. [doi: 10.1007/s12603-017-0870-y] [Medline: 28972246]
65. Musaeus CS, Waldemar G, Andersen BB, et al. Long-term EEG monitoring in patients with Alzheimer's disease using Ear-EEG: a feasibility study. J Alzheimers Dis 2022;90(4):1713-1723. [doi: 10.3233/JAD-220491] [Medline: 36336927]
66. Oh H, Gross MD. Awareable steps: functional and fashionable shoes for patients with dementia. Presented at: UbiComp '15: The 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing; Sep 7-11, 2015; Osaka, Japan p. 579-584 URL: https://dl.acm.org/doi/10.1145/2800835.2801666 [accessed 2025-10-17] [doi: 10.1145/2800835.2801666]
67. O'Sullivan G, Whelan B, Gallagher N, et al. Challenges of using a Fitbit smart wearable among people with dementia. Int J Geriatr Psychiatry 2023 Mar;38(3):e5898. [doi: 10.1002/gps.5898] [Medline: 36814072]
68. Peeters MWH, Schouten G, Wouters EJ. Wearables for residents of nursing homes with dementia and challenging behaviour: values, attitudes, and needs. Gerontechnology 2021 Jan 1;20(2):1-13 [FREE Full text] [doi: 10.4017/gt.2021.20.2.7.06]
69. Rings S, Steinicke F, Picker T, Prasuhn C. Memory journalist: creating virtual reality exergames for the treatment of older adults with dementia. Presented at: 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW); Mar 22-26, 2020; Atlanta, GA p. 687-688 URL: https://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=9086528 [accessed 2025-10-17] [doi: 10.1109/VRW50115.2020.00194]
70. Silva AR, Pinho MS, Macedo L, Moulin CJA. The cognitive effects of wearable cameras in mild Alzheimer disease: an experimental study. Curr Alzheimer Res 2017;14(12):1270-1282. [doi: 10.2174/1567205014666170531083015] [Medline: 28558637]

XSL•FO
RenderX

71. Staab S, Luderschmidt J, Martin L. Recognition of usual similar activities of dementia patients via smartwatches using supervised learning. Presented at: 2021 IEEE International Conference on Progress in Informatics and Computing (PIC); Dec 17-19, 2021; Shanghai, China p. 397-402. [doi: 10.1109/PIC53636.2021.9687025]
72. Stavropoulos TG, Lazarou I, Diaz A, et al. Wearable devices for assessing function in Alzheimer's disease: a european public involvement activity about the features and preferences of patients and caregivers. Front Aging Neurosci 2021;13:643135. [doi: 10.3389/fnagi.2021.643135] [Medline: 33912025]
73. Stavropoulos TG, Meditskos G, Kompatsiaris I. DemaWare2: integrating sensors, multimedia and semantic analysis for the ambient care of dementia. Pervasive Mob Comput 2017 Jan;34:126-145. [doi: 10.1016/j.pmcj.2016.06.006]
74. Tervonen J, Asghar MZ, Yamamoto G, Pulli P. A navigation aid for people suffering from dementia using a body worm laser device. 2013 Presented at: 2013 7th International Symposium on Medical Information and Communication Technology (ISMICT 2013); Mar 6-8, 2013; Tokyo, Japan p. 178-182. [doi: 10.1109/ISMICT.2013.6521725]
75. Thorpe J, Forchhammer BH, Maier AM. Adapting mobile and wearable technology to provide support and monitoring in rehabilitation for dementia: feasibility case series. JMIR Form Res 2019 Oct 17;3(4):e12346. [doi: 10.2196/12346] [Medline: 31625951]
76. Thorpe JR, R nn ‐ Andersen KVH, Bień P, Özkil AG, Forchhammer BH, Maier AM. Pervasive assistive technology for people with dementia: a UCD case. Healthcare Tech Letters 2016 Dec;3(4):297-302. [doi: 10.1049/htl.2016.0057]
77. Thorpe JR, Forchhammer BH, Maier AM. Development of a sensor-based behavioral monitoring solution to support dementia care. JMIR Mhealth Uhealth 2019 May 30;7(6):e12013. [doi: 10.2196/12013] [Medline: 31199304]
78. Treadaway C, Kenning G. Sensor e-textiles: person centered co-design for people with late stage dementia. Working with Older People 2016 Jun 13;20(2):76-85. [doi: 10.1108/WWOP-09-2015-0022]
79. Wallace J, Wright PC, McCarthy J, Green DP, Thomas J, Olivier P. A design-led inquiry into personhood in dementia. 2013 Apr 27 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Apr 27 to May 2, 2013; Paris, France p. 2617-2626 URL: https://dl.acm.org/doi/proceedings/10.1145/2470654 [accessed 2025-10-17] [doi: 10.1145/2470654.2481363]
80. Yokoi T, Okamura H, Yamamoto T, et al. Effect of wearing fingers rings on the behavioral and psychological symptoms of dementia: an exploratory study. SAGE Open Med 2017;5:2050312117726196. [doi: 10.1177/2050312117726196] [Medline: 28856006]
81. Zeng F, Lin Y, Siriaraya P, Choi D, Kuwahara N. Emotion detection using EEG and ECG signals from wearable textile devices for elderly people. J Text Eng 2020;66(6):109-117. [doi: 10.4188/jte.66.109]
82. Ampofo P, Vitali A, Morganti F, Rizzi C, Regazzoni D. A Preliminary Evaluation of a VR Application for Associative Memory and Cognitive Rehabilitation. Stud Health Technol Inform 2025 Apr 24;324:1-6. [doi: 10.3233/SHTI250148] [Medline: 40270372]
83. Fleetwood-Smith R, Tischler V, Robson D. Aesthetics and dementia: exploring the role of everyday aesthetics in dementia care settings. Des Health 2022 Jan 2;6(1):91-113. [doi: 10.1080/24735132.2022.2074207]
84. Zhu D, Al Mahmud A, Liu W. Digital storytelling intervention for enhancing the social participation of people with mild cognitive impairment: co ‐ design and usability study. JMIR Aging 2024 Jan 17;7:e54138. [doi: 10.2196/54138] [Medline: 38231541]
85. Goerss D, Köhler S, Rong E, et al. Smartwatch-based interventions for people with dementia: user-centered design approach. JMIR Aging 2024 Jun 7;7:e50107. [doi: 10.2196/50107] [Medline: 38848116]

Abbreviations

CIS: critical interpretive synthesis
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
VR: virtual reality

Edited by H Köttl; submitted 03.02.25; peer-reviewed by A Shnall, A Vitali; revised version received 12.09.25; accepted 22.09.25; published 06.11.25.

Please cite as:
Wei Y, Ratcliffe J, Aarsland D, Liu W
Aesthetic Experience in the Acceptance of Wearable Technology for People With Dementia: Critical Interpretive Synthesis
JMIR Aging 2025;8:e72082
URL: https://aging.jmir.org/2025/1/e72082
doi:10.2196/72082

XSL•FO
RenderX

©Yixuan Wei, John Ratcliffe, Dag Aarsland, Wei Liu. Originally published in JMIR Aging (https://aging.jmir.org), 6.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Factors Associated With the Ability To Keep Up With Technology Developments: Findings From a National Multigenerational Cross-Sectional Survey in Sweden

Jens Offerman[1], PhD; Sofi Fristedt[1,2], PhD; Steven M Schmidt[1], PhD; Susanne Iwarsson[1], PhD

[1]Department of Health Sciences, Faculty of Medicine, Lund University, Solvegatan 19, Lund, Sweden

[2]School of Health and Welfare, Jonkoping University, Jonkoping, Sweden

Corresponding Author:
Jens Offerman, PhD
Department of Health Sciences, Faculty of Medicine, Lund University, Solvegatan 19, Lund, Sweden

Abstract

Background: Digital technologies are increasingly central to supporting autonomy, health, and social participation in later life. However, disparities persist in the ability to keep up with technological developments, affecting individuals' opportunities to benefit from digital health and social innovations.

Objective: This study aimed to investigate factors associated with individuals' self-reported ability to keep up with technological developments, focusing on generational differences, attitudes toward digital tools, and sociodemographic characteristics.

Methods: We conducted a national cross-sectional online survey in Sweden with 2121 respondents aged 30 to 39 years, 50 to 59 years, and 70 to 79 years. Logistic regression analyses were used to identify associations between self-reported ability to keep up with technology and independent variables, including attitudes toward information and communication technology, gender, education, self-rated economic situation, and general health.

Results: Most respondents reported being able to keep up with technological developments. Compared to the oldest generation (70-79 years), participants aged 30 to 39 years had 188% higher odds (odds ratio [OR] 2.88, 95% CI 1.84-4.53) of reporting they kept up with technology developments, and women had lower odds than men (OR 0.52, 95% CI 0.39-0.70). Positive attitudes toward information and communication technology being user-friendly (OR 1.81, 95% CI 1.21-2.73), timesaving (OR 2.03, 95% CI 1.44-2.87), and increasing independence (OR 1.99, 95% CI 1.33-2.96) were also significantly associated with keeping up.

Conclusions: These findings suggest that digital inclusion in aging societies is shaped by complex and intersecting factors that go beyond age. Promoting equitable digital engagement requires addressing attitudinal, economic, and gender-related barriers and fostering inclusive technology design and support systems for both current and future generations of older adults.

(JMIR Aging 2025;8:e77930) doi:10.2196/77930

KEYWORDS

digital inclusion; aging; technology adoption; digital literacy; generational differences; health equity; ICT attitudes; information and communication technology

Introduction

Over the last decade, research on aging and technology has expanded extensively, and potential benefits of technology to support active and healthy aging have been identified [1-3]. Current and future generations of older adults are all immersed in the process of adopting, choosing, and using various technologies available today. The role of technology in promoting active and healthy aging is becoming increasingly relevant for people of all ages [4,5]. However, research on aging and technology has typically targeted one generation at a time, neglecting to consider the similarities and differences among and across generations. For instance, findings from a meta-analysis suggest that chronological age in itself only predicts lower technology acceptance in cases where perceived usefulness is unclear, such as with social media [6]. Instead, older adults—such as other age groups—are more likely to adopt technologies when they perceive them as valuable and relevant to their lives [7]. To facilitate technology adoption among forthcoming generations of older adults, more knowledge is imperative to nurture the development of effective services and interventions.

The decision to adopt technology is influenced by multiple factors, including perceived value, ease of use, affordability, confidence, and individual interests [7-9]. The Unified Theory of Acceptance and Use of Technology (UTAUT) provides a framework for understanding these factors, highlighting key determinants such as performance expectancy, effort expectancy, social influence, and facilitating conditions [10]. According to

XSL•FO
RenderX

UTAUT, people are more likely to adopt technology when they perceive it as beneficial, easy to use, socially supported, and have access to the necessary resources and skills. This model can be particularly insightful in understanding how different generations approach technology adoption, as it considers both intrinsic factors (eg, personal motivation) and extrinsic factors (eg, social influences and accessibility) [10].

Interest in technology is not necessarily age-specific but rather depends on individual previous experiences and interests [11]. Evidence indicates that both younger and older adults report difficulties with information and communication technologies (ICTs) necessary for everyday activities [12,13]. However, older adults often report lower confidence and greater concerns about complexity, cost, and privacy, which can slow adoption [14,15]. These mixed findings point to the importance of going beyond chronological age as a factor determining technology adoption.

Still, keeping up with technology can become a challenge as we age [1]. Research has shown that people who take an early interest in technology are more likely to continue using and keeping up with technology later in life [16]. While younger generations are more inclined to be early adopters, it is important to recognize that older generations are increasingly using technology [17]. According to Lee et al [18], age disparities likely diminish over time, implying that attitudes toward and use of technology will become more similar across generations. Cross-national studies further highlight how adoption patterns vary by context. In countries with strong digital infrastructures, older adults report higher levels of technology use, whereas in other settings, barriers linked to affordability and access remain more pronounced [19,20]. Even in high-income countries such as Sweden, where overall internet use is high, 1 in 5 adults report need of assistance with digital tasks [21]. Noteworthy, this need is not limited to older adults; younger age groups are increasingly reporting difficulties with everyday digital tasks [21].

In a previous study, Offerman et al [22] identified similarities and differences between generations in their attitudes toward technology and the types of technology they wanted for active and healthy aging. In another study, Offerman et al [23] found similarities regarding what types of ICT products were used more often during the early phase of the COVID-19 pandemic. Across 3 age groups, most respondents in both studies described having the necessary technological knowledge to meet their everyday needs and to keep up with technology developments. Such findings suggest that the digital divide is not necessarily between younger and older generations, but rather between groups of users with different characteristics and their interest and ability to keep up with technology. Understanding the underlying factors associated with the ability to keep up with technology can help refine technological design. This, in turn, enables the development of products and services that better suit the diverse needs of people of all ages and ensure that everyone can keep up with technological advancements.

Accordingly, there is a need to study factors that influence the adoption and usage of digital technology, such as people's self-perceptions of keeping up with technological developments, and to identify how these perceptions are associated with attitudes toward technology and sociodemographic characteristics. Taken together, this suggests that technology adoption is shaped not only by age but also by confidence, motivation, and contextual circumstances. However, most prior studies have examined one generation at a time, providing limited insight into similarities and differences across cohorts. Existing evidence is contradictory because some studies predict that generational gaps in technology adoption will diminish over time [18], whereas others suggest that persistent disparities in resources, confidence, and motivation will sustain inequality [19,24]. It remains underexplored how individuals across different generations perceive their ability to keep up with technological developments and which sociodemographic and attitudinal factors are associated with these perceptions.

In this study, *keeping up with technology developments* refers to individuals' self-reported ability to follow and adapt to ongoing technological changes in everyday life. The phenomenon at target captures aspects of perceived competence and confidence, as well as willingness to engage with new technologies, rather than objectively measured skills or behaviors. This construct is conceptually linked to digital competence and perceived digital inclusion as described in previous research [19]. Addressing this phenomenon allows us to better understand multigenerational patterns of technology adoption and inform inclusive strategies for active and healthy aging.

In this study, the term digital technology refers to ICT (eg, computers, smartphones, internet services, and applications) and other digital tools that support health and everyday activities. This study builds on an earlier study from the same survey [22], which examined generational attitudes toward technology in a broad sense. While that work mapped overall attitudes, less is known about the specific factors associated with people's self-reported ability to keep up with ongoing technological developments. Accordingly, the purpose of this study was defined based on the findings from the previous study from the GenerationTech project. Focusing on sociodemographic characteristics and attitudinal factors, this study builds on the UTAUT framework [10] and prior research showing that these variables are central for technology adoption. The purpose was to investigate factors associated with the self-reported ability to keep up with technology developments among people from 3 age groups representing different generations, taking attitudes toward technology and sociodemographics into consideration. The analyses were guided by the following research question: Which sociodemographic (age, gender, education, health, and economic resources) and attitudinal factors (perceived usefulness, ease of use, time-saving potential, and independence support of ICT and digital household devices) are associated with the self-reported ability to keep up with technological developments?

Methods

Respondents, Sampling, and Recruitment

Operationalizing generational categories as 3 distinct age groups to avoid overlap, a random sample was drawn from the Swedish

XSL•FO
RenderX

State Personal Address Register, representing men and women stratified in 3 age cohorts (30 - 39, 50 - 59, and 70 - 79 years).

Kantar Sifo (KS; a company with documented experience from large-scale data collection) acquired 10,000 addresses from Swedish State Personal Address Register in August 2019. Different numbers of addresses were included for each age cohort to compensate for the fact that younger people tend to have lower response rates based on KS's data collection experience. The survey was conducted between August and October 2019.

KS used a comprehensive strategy to achieve the highest possible response rates. Postal letters containing informed consent information, survey web links, and unique login details to an online questionnaire were sent to potential respondents. Nonresponders were followed up with a postal reminder after 1 week, and if they still did not respond, they were contacted up to 8 times by trained staff via telephone. During these calls, respondents were reminded about the survey and given the alternatives of completing it via a telephone interview or receiving a postal version.

These efforts resulted in a final sample of 2121 respondents (n=1081 men, 51%; and n=1040 women, 49%). The response rate was 22% for men and 21% for women across 3 age groups (for details, see a study by Offerman et al [22]). Most of the 2121 respondents were born in Sweden and had at least completed compulsory school education. The youngest generation had a higher education level than the older generations and reported better general health, whereas the oldest generation rated their general health and economy lower than the other generations. Table 1 provides detailed sample characteristics.

Table . Characteristics of survey sample (n=2121).

Characteristics	Age 30 - 39 y (n=639), n (%)	Age 50 - 59 y (n=703), n (%)	Age 70 - 79 y (n=779), n (%)
Sex			
Men	316 (49)	345 (49)	420 (54)
Women	323 (51)	358 (51)	359 (46)
Education			
Compulsory school	14 (2)	30 (4)	217 (28)
High school	153 (24)	244 (35)	111 (14)
Polytechnic	85 (14)	83 (12)	130 (17)
University	383 (60)	341 (49)	314 (41)
Self-rated economy for technology needs			
Good	333 (52)	382 (55)	305 (40)
Fairly good	228 (36)	239 (34)	338 (44)
Fairly bad	55 (9)	51 (7)	76 (10)
Bad	21 (3)	25 (4)	52 (6)
Self-rated general health[a]			
Excellent	130 (21)	120 (17)	64 (8)
Very good	264 (41)	275 (39)	242 (32)
Good	182 (29)	213 (31)	302 (39)
Fair	52 (8)	69 (10)	143 (19)
Poor	9 (1)	21 (3)	17 (2)
Self-rated life satisfaction[a]			
Excellent	96 (15)	115 (17)	120 (16)
Very good	290 (46)	307 (44)	301 (39)
Good	183 (29)	193 (28)	258 (34)
Fair	53 (8)	66 (9)	80 (10)
Poor	12 (2)	13 (2)	6 (1)

[a]Assessed using a self-reported measure based on the SF-36 Health Survey [25].

In this study, the term "generation" was used pragmatically to refer to 3 age groups: 30 to 39 years, 50 to 59 years, and 70 to 79 years. These groups were chosen to reflect meaningful differences in historical exposure to digital technologies and societal transitions while avoiding overlapping of adjacent cohorts. That is, the age groups represent distinct life-course

XSL•FO
RenderX

positions rather than fixed generational identities, capturing variation in digital familiarity, health needs, and societal expectations regarding aging. Generational categories, defined by birth years, assume shared experiences that shape attitudes and behaviors. These categories highlight broad patterns, such as how historical and cultural contexts influence technology adoption. For example, generational cohorts may share exposure to specific technologies or societal shifts that have shaped their comfort and confidence with technologies.

Data Collection

Data were collected through a structured questionnaire developed for the GenerationTech project [22], based on qualitative findings [11] involving the same age groups as this study and relevant scientific literature, focusing on technology use, attitudes, and adoption. The survey was constructed to align with the research objectives and included questions grounded in theories of technology adoption and digital inclusion [10,26]. The survey included 24 questions about attitudes toward and adoption of technology, for example, "I have no problem keeping up with technology developments." Questions addressed experiences and attitudes toward ICT and household devices, for instance, if the respondent perceives these save time, are useful, support independence, and are user-friendly. The survey also included 7 questions about respondent characteristics, such as education, occupation, housing, civil state, and country of birth, as well as self-rated general health, life satisfaction, and self-rated economy to cover technology needs (Multimedia Appendix 1). To ensure data quality and usability, a pilot study, followed by minor revisions to the questions, was conducted before full-scale data collection. It took approximately 10 to 15 minutes to complete the survey.

KS performed regular quality control of data during data collection, focusing on correct, complete, and logical recording in the database. Researchers monitored the process to identify potential errors, and KS communicated with the research team when needed. For more details, see the Method section of the study by Offerman et al [22].

Data Analyses

Descriptive statistics were used for sample demographics. Binary logistic regression was implemented to analyze the factors associated with the respondents' reported ability to keep up with technological developments. The dependent variable was whether respondents reported keeping up with technology developments. This question had 4 response alternatives, ranging from strongly disagree to strongly agree. To facilitate the data analysis because of the nonnormal distribution and skewness of the data, the response alternatives were merged and dichotomized into agreeing or not agreeing (yes or no). In addition, this approach was chosen to enhance the interpretability for a broad audience, facilitate consistency across related studies, and strengthen the comparability of outcomes. Independent variables included gender, belonging to a specific generation (ie, age group), attitudes toward ICT and digital household devices (eg, perceieved usefulness, user-friendly, timesaving, and a means to increase independence), and sociodemographic characteristics, such as education, self-rated health, self-rated economy for technology, and life satisfaction. To identify potential multicollinearity, a correlation matrix was established. None of the independent variables met the criterion for exclusion (cutoff value >0.7). Logistic regressions were first computed univariably, followed by a multivariable model, including all independent variables. Cases with missing values on the dependent variable were excluded. For independent variables, respondents with incomplete data were excluded from the regression models where it was deemed relevant. As the proportion of missing data was low, no imputation was performed. Data were not normally distributed but met the assumptions of the statistical tests used. SPSS Statistics 29 (IBM) was used for the data analyses. The alpha level was set to $P<.05$.

Ethical Considerations

This study was based on a quantitative, cross-sectional survey conducted as part of the GenerationTech project. The Swedish Ethical Review Authority (No: 2019 - 02072) approved the study. KS performed the sampling, recruitment, and data collection on behalf of and in collaboration with the research team (for details, see a study by Offerman et al [22]). All respondents gave their informed consent.

Results

A majority of the respondents reported having the ability to keep up with technology developments. In the univariable models, compared to the oldest generation, the odds of reporting having no problems keeping up with technology developments were 308% higher among the youngest generation and 72% higher in the middle-aged generation. Attitudes toward ICT all showed a significant positive association with the dependent variable. Having a university education compared to compulsory school-level education was associated with 187% higher odds of reporting having no problems keeping up with technology developments (Table 2).

XSL•FO
RenderX

Table . Univariable and multivariable logistic regression models of factors associated with the reported ability to keep up with technology developments (n=2121). Note: The dependent variable is reported ability to keep up with technology developments. Independent variables are generation, gender, self-rated life satisfaction, self-rated health, self-rated economy to cover technology needs, attitudes toward ICT[a], and household devices. Model fit (Nagelkerke R^2)=0.218. Numbers are rounded to nearest integer.

Factor	Univariable model		Multivariable model	
	OR[b] (95% CI)	*P* value	OR (95% CI)	*P* value
Generation				
30 - 39 y	4.08 (2.71-6.12)	<.001	2.88 (1.84-4.53)	<.001
50 - 59 y	1.72 (1.27-2.33)	<.001	1.36 (0.95-1.93)	.09
70 - 79 y	Ref[c]	Ref	Ref	Ref
Gender				
Male	Ref	Ref	Ref	Ref
Female	0.55 (0.42-0.73)	<.001	0.52 (0.39-0.70)	<.001
Attitudes toward ICT				
Perceieved usefulness (ref no)	1.79 (1.36-2.35)	<.001	1.33 (0.96-1.84)	.08
User friendly (ref no)	2.96 (2.09-4.19)	<.001	1.81 (1.21-2.73)	.004
Timesaving (ref no)	2.72 (2.05-3.61)	<.001	2.03 (1.44-2.87)	<.001
Make me independent (ref no)	2.37 (1.69-3.32)	<.001	1.99 (1.33-2.96)	<.001
Attitudes toward household devices				
Perceieved usefulness (ref no)	1.15 (0.87-1.51)	.32	0.75 (0.53-1.05)	.09
User friendly (ref no)	1.31 (0.99-1.73)	.06	0.84 (0.59-1.18)	.32
Timesaving (ref no)	1.36 (1.03-1.80)	.03	0.89 (0.63-1.27)	.51
Make me independent (ref no)	0.94 (0.69-1.28)	.69	0.68 (0.46-0.99)	.04
Education level				
Compulsory school	Ref	Ref	Ref	Ref
High school	2.34 (1.56-3.52)	<.001	1.32 (0.82-2.12)	.25
Polytechnic	2.19 (1.38-3.49)	<.001	1.55 (0.94-2.58)	.09
University	2.87 (1.99-4.13)	<.001	1.45 (0.95-2.21)	.09
General health				
Excellent	1.95 (1.21-3.15)	.006	1.31 (0.74-2.30)	.35
Very good	1.83 (1.30-2.57)	<.001	1.47 (0.99-2.17)	.05
Good	Ref	Ref	Ref	Ref
Bad	0.63 (0.44-0.90)	.01	0.85 (0.56-1.29)	.44
Life satisfaction				
Excellent	1.73 (1.09-2.74)	.02	1.19 (0.69-2.04)	.52
Very good	1.47 (1.07-2.03)	.02	1.06 (0.73-1.53)	.77
Good	Ref	Ref	Ref	Ref
Bad	0.73 (0.49-1.11)	.14	0.89 (0.55-1.44)	.65
My economy covers my technology need				
Good	1.08 (0.79-1.47)	.62	0.71 (0.50-0.99)	.04
Fairly good	Ref	Ref	Ref	Ref
Bad	0.47 (0.33-0.68)	<.001	0.55 (0.36-0.83)	.004

XSL•FO
RenderX

[a]ICT: information and communication technology.

[b]OR: odds ratio.

[c]Ref: reference.

In the multivariable model, compared to the oldest generation, the odds in the youngest generation were 188% higher of reporting having no problems keeping up with technology developments. Women were less likely (odds ratio [OR] 0.52, 95% CI 0.39-0.70) to report themselves as keeping up with technology developments compared to men. In the multivariable model, attitudes such as user-friendly (OR 1.81, 95% CI 1.21-2.73), time-saving (OR 2.03, 95% CI 1.44-2.87), and technology as a means to increase independence (OR 1.99, 95% CI 1.33-2.96) all showed significant associations with the dependent variable, reporting having no problems to keep up with technology developments. Education showed no significant associations in the multivariable model. Respondents who rated their economy as good (OR 0.71, 95% CI 0.50-0.99) or bad (OR 0.55, 95% CI 0.36, 0.83) compared to those who rated their economy fairly good were less likely to report keeping up with technology developments (Table 2). The final multivariable model showed a Nagelkerke R^2 of 0.218.

Discussion

Principal Findings

Our investigation of factors associated with the reported ability to keep up with technology among current and future generations of older adults reveals essential insights into the dynamics of technology adoption and its perceived benefits. Both the youngest and the middle-aged generations were more likely than the oldest generation to report having the ability to keep up with technology developments, with a majority across all generations indicating the same. This study highlights key factors associated with whether individuals report keeping up with technology, for instance, attitude toward technology, self-rated economy for technology needs, age, and gender.

Generational differences in technology adoption deserve attention. Younger individuals, particularly those aged 30 to 39 years, are more likely to perceive themselves as technologically up to date compared to older adults, such as those aged 70 to 79 years. UTAUT is a useful framework to explore how generational differences in constructs, such as social influence, effort expectancy, and facilitating conditions, may explain these disparities in technology adoption [26]. In addition, these differences could stem from varying levels of exposure to technology, particularly through education and professional experiences and environments, which allow younger individuals to integrate technology more seamlessly into their lives [27]. In contrast, older generations often require more time and support to adapt to new technologies, facing unique challenges due to limited prior exposure [15,16,26]. However, Lee et al [18] suggested that these age-related disparities in technology adoption may diminish over time, as attitudes and usage patterns across generations gradually converge. In line with this, in a recent study [23], we found that respondents across 3 generations reported having the necessary technological knowledge to meet their everyday needs, although this competence was more pronounced in the youngest generation. Investigating this further, in this study, most respondents reported keeping up with technological developments. Taken together, these findings suggest that age-related disparities in technology adoption may be influenced not only by exposure but also by individual perceptions of the benefits and usability of technology. According to UTAUT, people are more likely to adopt technology when they perceive it as beneficial to their daily lives [26]. This perceived benefit, related to performance expectancy, likely contributes to the reported ability to keep up with technological developments. Thus, the digital divide may not solely be between younger and older adults, but between user groups with varying levels of perceived usefulness and expected performance outcomes.

The results show that people with a positive attitude toward ICT are more likely to report keeping up with technology developments. Previous research has shown that older adults are more likely to adopt new technology when they perceive it as easy to use and beneficial for their independence and daily life [16,28]. Similarly, our study shows that individuals who view technology as user-friendly, time-saving, and a means to increase independence are more inclined to report having the ability to keep up with technology developments. These attitudes are crucial motivators for learning and engaging with technology [1,29] and critical determinants of technology adoption [10]. UTAUT helps explain this by suggesting that as facilitating conditions (eg, access to technology and support) improve, older adults may experience fewer barriers to adopting and using technology [26].

The proactive aging research approach, introduced by Iwarsson et al [30], aligns with these findings by recognizing the dynamic interactions between individuals and their environments throughout the aging process. This approach considers aspects throughout the life course, addressing both current and future generations of older adults, from earlier stages of aging into later life. It focuses on the potential and challenges associated with aging, including cognitive, psychological, physical, and social factors, all of which can influence technology adoption among older adults. Promoting positive attitudes toward technology through campaigns and programs that highlight its benefits can support goals toward proactive aging [30] and encourage broader acceptance and engagement with technology among older adults [31,32].

Surprisingly and somewhat difficult to interpret, both those who responded that their economy covers their technology needs as good or bad are less likely to report keeping up with technology compared to those who report their economy as fair. This could indicate that individuals at economic extremes may either feel less motivation to engage with technology due to financial security and make other priorities in life, or they lack the resources and motivation due to financial constraints. However, previous literature indicates that individuals with higher socioeconomic status tend to have better access to technology [33]. Our finding adds nuance, showing that access might not be enough to elicit feelings of being able to keep up with the

XSL•FO
RenderX

rapid technological developments. Our findings suggest that individuals with financial security may be slower to adopt new technologies, whereas those with limited means often struggle to access them. This gap can significantly deepen existing disparities, as financial stability can postpone technology adoption, thereby widening the technological divide and further marginalizing those who are less fortunate in an increasingly digital landscape. It is also possible that participants with financial security adopt alternative solutions or invest their resources in other valued areas (eg, travel, sports, or social activities), making digital engagement less of a priority. In this sense, digital exclusion may stem not only from economic constraints but also from economic choice, reinforcing that motivation, interest, and perceived usefulness are central to understanding technology adoption across different groups. However, we would like to emphasize that our interpretation of this unexpected finding remains cautious and should be further explored in future research.

Gender differences play a significant role in technology engagement, with women less likely than men to report keeping up with technological developments. Similarly, a cross-national meta-analysis found that men consistently rate their technology abilities higher than women, although this difference is narrowing over time [34]. This disparity may stem from differences in access to technology, confidence levels, and societal norms that influence technology use and career choices [35]. In fact, men tend to overestimate their technological abilities, whereas women underestimate theirs [36]. Research suggests this is partly due to exclusive design processes and inadequate testing involving women [37]. Addressing these disparities is crucial to reduce the digital divide between genders. Enhancing women's confidence in technology requires more inclusive design processes and equitable testing across genders. Ensuring that digital technologies are designed inclusively and tested equitably can improve accessibility, usability, and overall engagement, ultimately reducing gender disparities in technology adoption [37].

Moreover, while higher education was influential in univariate analyses, more than doubling the odds of reporting keeping up with technology developments, it did not remain significant in the multivariable model. Individuals with higher education levels are more likely to report keeping up with technology compared to those with only compulsory education. This suggests that education equips individuals with critical thinking skills and technical knowledge necessary for engaging with new technologies [27,38]. Higher education fosters critical thinking abilities, which are crucial for understanding and adapting to new technologies. This educational advantage translates into greater ease in adopting and using new technologies, reinforcing the perception of keeping up with technological developments [38]. Again, there are notable differences among subgroups of the aging population deserving explicit attention regarding the prerequisites for technology adoption.

Implications

Addressing generational diversity and socioeconomic disparities in technology adoption should be a priority in digitalization policies. Our findings highlight the importance of fostering positive attitudes toward technology, ensuring equitable access, and addressing disparities in perceived usefulness and competence. Targeted strategies are essential to support people in keeping up with technology along the life course and as they age, particularly those from older generations [31].

Support systems tailored to meet individual needs within different generations, due to socioeconomic backgrounds, are essential to prevent further digital exclusion [8]. Examples include digital literacy programs embedded in community services, intergenerational mentoring initiatives, and locally available "digital helpdesks" where individuals can receive hands-on support. Policies should be informed by individuals' actual needs, recognizing that both financial constraints and financial security can influence technology engagement. Providing economic support and incentives for technology purchases and training can help individuals with lower economic means access and adopt new technologies [8]. Such support can ensure that financial barriers do not hinder technological engagement, fostering a more inclusive digital society [32]. Furthermore, inclusive design and testing processes must be implemented to reduce gender disparities and enhance usability for all [37]. This means involving users from diverse age groups and socioeconomic backgrounds in co-design processes and ensuring that interfaces are intuitive, adaptable, and responsive to different skill levels. This, in turn, could help reduce the intimidation factor associated with new technology [39]. Policies and interventions should go beyond access to promote confidence, motivation, and meaningful digital participation, with sustained investment in digital skills development across the life course [19]. Without such targeted strategies and interventions, there is a risk that digital transformation will reinforce existing inequalities, leaving certain groups behind as they struggle to keep up with technological developments.

Strengths and Limitations

It should be kept in mind that the cross-sectional design restricts the ability to disentangle age, cohort, and period effects. Generational categories were used to describe different sociotechnical contexts, and the observed differences are best understood as cross-sectional associations rather than causal generational effects. Longitudinal research is required to clarify these dynamics.

The strength of this study lies in the nationwide sample, which was randomly selected from the Swedish population. However, despite extensive efforts to use a broad sampling frame and issue multiple reminders, the response rate was low, particularly among younger generations. Nevertheless, the sample reflects the demographic composition of the Swedish population. For instance, 78% of respondents had at least a high school education, compared to 85% in the general population. A substantial majority of respondents (90%) were born in Sweden, and approximately 75% resided in larger cities. The diversity in socioeconomic backgrounds and nationalities among respondents contributed to a heterogeneous survey sample. Still, it is also possible that individuals who are less digitally engaged or less confident in technology use were underrepresented, which may have led to an overestimation of the perceived ability to keep up with technology and an underestimation of the

XSL•FO
RenderX

challenges faced by the least digitally included groups. Therefore, while the findings remain generally relevant, they should be interpreted with consideration of these factors and potential bias.

Another inherent limitation is that participation in a survey on digital technology may have favored individuals with greater interest or competence in digital technology, potentially leading to overestimation of engagement. Nonresponse bias may also have occurred, with less confident users underrepresented. As with all self-reported data, recall and social desirability bias are possible. Although survey weights were applied to enhance representativeness, some subgroups may still be over- or under-represented.

It would have been beneficial to include questions about the respondents' actual use and previous experience with technology. However, designing the survey required balancing comprehensiveness with an acceptable respondent burden. In retrospect, adding these questions would have made it easier to draw conclusions about the respondents' level of interest, knowledge, and competence. It would also have made it possible to study whether there is an association between using and keeping up with technology, which is an interesting question for future research.

The dichotomization of the outcome variable is another limitation deserving comment because it entailed a loss of nuance and may have obscured subtle variations in perceptions. Despite these trade-offs, we ultimately deemed this the most appropriate analytical strategy. In addition to enhancing communicative clarity, dichotomization supported consistency across studies and facilitated comparability of outcomes, thereby strengthening the coherence of the overall analysis.

Conclusions

This study underscores the multifaceted factors associated with the ability to keep up with technology among current and future generations of older adults. Applying a proactive aging approach to research on aging and technology highlights the value of adopting a generational perspective to understand similarities and distinct challenges faced by different age groups. Across generations, perceived ease of use and benefits such as enhanced independence and time-saving capabilities are essential motivators for technology adoption. Addressing educational, economic, gender, and attitudinal barriers, fostering positive attitudes toward technology, promoting user-friendly designs, and encouraging inclusivity across demographic divides are critical steps toward bridging the digital divide and fostering a more inclusive environment for technological engagement.

The results underscore the importance of encouraging individuals to prepare for technological and environmental changes to enhance their quality of life as they age. By implementing targeted strategies that consider the differences between generations, society can ensure that all individuals have the opportunity to keep up with technology and benefit from its potential to promote active and healthy aging for current as well as future generations of older adults. This includes policies that support lifelong digital skill development, interventions such as intergenerational mentoring and community-based digital support, and design practices that prioritize accessibility and co-creation with diverse users. Taken together, these efforts can help translate research insights into concrete actions that promote equitable digital inclusion across the life course.

Acknowledgments

This work was supported by the Swedish Research Council for Health, Working Life and Welfare (FORTE; 2017 - 01614 to SI), and the Ribbingska Foundation in Lund, Sweden.

Data Availability

The data used in this study contain sensitive information about the study participants who did not provide consent for public data sharing. The current approval by the Swedish Ethical Review Authority (ref. no. 2019 - 02072) does not include data sharing. The manuscript has been read and approved for submission by all the named authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire used for the survey study.
[DOCX File, 87 KB - aging_v8i1e77930_app1.docx]

References

1. Czaja SJ, Charness N, Rogers WA, Sharit J, Moxley JH, Boot WR. The benefits of technology for engaging aging adults: findings from the PRISM 2.0 trial. Innov Aging 2024;8(6):igae042. [doi: 10.1093/geroni/igae042] [Medline: 38854853]
2. Nimrod G. Aging well in the digital age: technology in processes of selective optimization with compensation. J Gerontol B Psychol Sci Soc Sci 2020 Oct 16;75(9):2008-2017. [doi: 10.1093/geronb/gbz111] [Medline: 31504873]

XSL•FO
RenderX

3. Siegel C, Dorner TE. Information technologies for active and assisted living-Influences to the quality of life of an ageing society. Int J Med Inform 2017 Apr;100:32-45. [doi: 10.1016/j.ijmedinf.2017.01.012] [Medline: 28241936]
4. Chen C, Ding S, Wang J. Digital health for aging populations. Nat Med 2023 Jul;29(7):1623-1630. [doi: 10.1038/s41591-023-02391-8] [Medline: 37464029]
5. De' R, Pandey N, Pal A. Impact of digital surge during Covid-19 pandemic: a viewpoint on research and practice. Int J Inf Manage 2020 Dec;55:102171. [doi: 10.1016/j.ijinfomgt.2020.102171] [Medline: 32836633]
6. Hauk N, Hüffmeier J, Krumm S. Ready to be a silver surfer? A meta-analysis on the relationship between chronological age and technology acceptance. Comput Human Behav 2018 Jul;84:304-319. [doi: 10.1016/j.chb.2018.01.020]
7. Berkowsky RW, Sharit J, Czaja SJ. Factors predicting decisions about technology adoption among older adults. Innov Aging 2018 Jan;2(1):igy002. [doi: 10.1093/geroni/igy002] [Medline: 30480129]
8. Chan DYL, Lee SWH, Teh PL. Factors influencing technology use among low-income older adults: a systematic review. Heliyon 2023 Sep;9(9):e20111. [doi: 10.1016/j.heliyon.2023.e20111] [Medline: 37809586]
9. Zhou J, Zhang B, Tan R, Tseng ML, Zhang Y. Exploring the systematic attributes influencing gerontechnology adoption for elderly users using a meta-analysis. Sustainability 2020;12(7):2864. [doi: 10.3390/su12072864]
10. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
11. Fristedt S, Svärdh S, Löfqvist C, Schmidt SM, Iwarsson S. "Am I representative (of my age)? No, I'm not"-Attitudes to technologies and technology development differ but unite individuals across rather than within generations. PLoS ONE 2021;16(4):e0250425. [doi: 10.1371/journal.pone.0250425] [Medline: 33886658]
12. Zingmark M, Nordeström F, Iwarsson S. Challenges related to self-assessment of active ageing during the Covid-19 pandemic in Sweden. BMC Res Notes 2022 May 13;15(1):171. [doi: 10.1186/s13104-022-06059-3] [Medline: 35562784]
13. Olsson T, Samuelsson U, Viscovi D. At risk of exclusion? Degrees of ICT access and literacy among senior citizens. Inf Commun Soc 2019 Jan 2;22(1):55-72. [doi: 10.1080/1369118X.2017.1355007]
14. Boot WR, Moxley JH, Roque NA, et al. Exploring older adults' video game use in the PRISM computer system. Innov Aging 2018 Jan;2(1):igy009. [doi: 10.1093/geroni/igy009] [Medline: 30480133]
15. Vaportzis E, Clausen MG, Gow AJ. Older adults perceptions of technology and barriers to interacting with tablet computers: a focus group study. Front Psychol 2017 Oct 4;8:1687. [doi: 10.3389/fpsyg.2017.01687] [Medline: 29071004]
16. Mitzner TL, Savla J, Boot WR, et al. Technology adoption by older adults: findings from the PRISM trial. Gerontologist 2019 Jan 9;59(1):34-44. [doi: 10.1093/geront/gny113] [Medline: 30265294]
17. Internet, Broadband Fact Sheet. Pew Research Center. URL: https://www.pewresearch.org/internet/fact-sheet/internet-broadband/ [accessed 2022-03-16]
18. Lee CC, Czaja SJ, Moxley JH, et al. Attitudes toward computers across adulthood from 1994 to 2013. Gerontologist 2019 Jan 9;59(1):22-33. [doi: 10.1093/geront/gny081] [Medline: 29982458]
19. Scheerder A, van Deursen A, van Dijk J. Determinants of internet skills, uses and outcomes. A systematic review of the second- and third-level digital divide. Telemat Informat 2017 Dec;34(8):1607-1624. [doi: 10.1016/j.tele.2017.07.007]
20. Yusif S, Soar J, Hafeez-Baig A. Older people, assistive technologies, and the barriers to adoption: a systematic review. Int J Med Inform 2016 Oct;94:112-116. [doi: 10.1016/j.ijmedinf.2016.07.004] [Medline: 27573318]
21. The Swedes and the internet 2024. : Swedish Internet Foundation; 2024 URL: https://svenskarnaochinternet.se/rapporter/svenskarna-och-internet-2024/ [accessed 2025-03-03]
22. Offerman J, Fristedt S, Schmidt SM, Lofqvist C, Iwarsson S. Attitudes related to technology for active and healthy aging in a national multigenerational survey. Nat Aging 2023 May;3(5):617-625. [doi: 10.1038/s43587-023-00392-3] [Medline: 37118552]
23. Offerman J, Schmidt SM, Iwarsson S, Fristedt S. Self-reported changes in use of and attitudes toward ICT in three generations in Sweden during the early phase of the COVID-19 pandemic. Gerontol Geriatr Med 2024;10:23337214241228109. [doi: 10.1177/23337214241228109] [Medline: 38283763]
24. Friemel TN. The digital divide has grown old: determinants of a digital divide among seniors. New Media & Society 2016 Feb;18(2):313-331. [doi: 10.1177/1461444814538648]
25. Ware JE, Sherbourne CD. The MOS 36-ltem Short-Form Health Survey (SF-36). Med Care 1992 Jun;30(6):473-483. [doi: 10.1097/00005650-199206000-00002]
26. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 2012;36(1):157. [doi: 10.2307/41410412]
27. Rogers EM. Diffusion of Innovations: Simon and Schuster; 2010. URL: https://www.simonandschuster.com/books/Diffusion-of-Innovations-4th-Edition/Everett-M-Rogers/9781451602470 [accessed 2025-11-07]
28. Harris MT, Blocker KA, Rogers WA. Older adults and smart technology: facilitators and barriers to use. Front Comput Sci 2022;4. [doi: 10.3389/fcomp.2022.835927]
29. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
30. Iwarsson S, Jönson H, Deierborg T, et al. "Proactive aging" is a new research approach for a new era. Nat Aging 2023 Jul;3(7):755-756. [doi: 10.1038/s43587-023-00438-6] [Medline: 37291219]

XSL•FO
RenderX

31. LoBuono DL, Leedahl SN, Maiocco E. Teaching technology to older adults: modalities used by student mentors and reasons for continued program participation. J Gerontol Nurs 2020 Jan 1;46(1):14-20. [doi: 10.3928/00989134-20191118-02] [Medline: 31895957]
32. Miller LMS, Callegari RA, Abah T, Fann H. Digital literacy training for low-income older adults through undergraduate community-engaged learning: single-group pretest-posttest study. JMIR Aging 2024 May 14;7:e51675. [doi: 10.2196/51675] [Medline: 38599620]
33. Pirhonen J, Lolich L, Tuominen K, Jolanki O, Timonen V. "These devices have not been made for older people's needs" – Older adults' perceptions of digital technologies in Finland and Ireland. Technol Soc 2020 Aug;62:101287. [doi: 10.1016/j.techsoc.2020.101287]
34. Christensen MA. Tracing the gender confidence gap in computing: a cross-national meta-analysis of gender differences in self-assessed technological ability. Soc Sci Res 2023 Mar;111:102853. [doi: 10.1016/j.ssresearch.2023.102853] [Medline: 36898792]
35. Shaouf A, Altaqqi O, Al-Zintan University. The impact of gender differences on adoption of information technology and related responses: a review. IJMAR 2018;5(1):22-41 [FREE Full text] [doi: 10.18646/2056.51.18-003]
36. Huffman AH, Whetten J, Huffman WH. Using technology in higher education: the influence of gender roles on technology self-efficacy. Comput Human Behav 2013 Jul;29(4):1779-1786. [doi: 10.1016/j.chb.2013.02.012]
37. Stanney K, Fidopiastis C, Foster L. Virtual reality is sexist: but it does not have to be. Front Robot AI 2020;7:4. [doi: 10.3389/frobt.2020.00004] [Medline: 33501173]
38. Schenck A. Examining relationships between technology and critical thinking: a study of South Korean EFL learners. Educ Sci 2024;14(6):652. [doi: 10.3390/educsci14060652]
39. Fleury S, Chaniaud N. Multi-user centered design: acceptance, user experience, user research and user testing. Theor Issues Ergon Sci 2024 Mar 3;25(2):209-224. [doi: 10.1080/1463922X.2023.2166623]

Abbreviations

ICT: information and communication technology
KS: Kantar Sifo
OR: odds ratio
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by A Garcia; submitted 22.05.25; peer-reviewed by D Verran, Q Li, T Vagg, T Abah; revised version received 16.10.25; accepted 25.10.25; published 14.11.25.

Please cite as:
Offerman J, Fristedt S, Schmidt SM, Iwarsson S
Factors Associated With the Ability To Keep Up With Technology Developments: Findings From a National Multigenerational Cross-Sectional Survey in Sweden
JMIR Aging 2025;8:e77930
URL: https://aging.jmir.org/2025/1/e77930
doi:10.2196/77930

© Jens Offerman, Sofi Fristedt, Steven M Schmidt, Susanne Iwarsson. Originally published in JMIR Aging (https://aging.jmir.org), 14.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Patient-Related Barriers to Digital Technology Adoption in Alzheimer Disease: Systematic Review

Andrea Panzavolta[1], MS; Andrea Arighi[2], MD; Emanuele Guido[1], MD; Luigi Lavorgna[3], MD, PhD; Francesco Di Lorenzo[4], MD, PhD; Alessandra Dodich[5], PhD; Chiara Cerami[1,6], MD

[1]IUSS Cognitive Neuroscience ICoN Center, Scuola Universitaria Superiore IUSS, Piazza della Vittoria, 15, Pavia, Italy

[2]Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy

[3]Division of Neurology, Department of Advanced Medical and Surgical Sciences, AOU-University of Campania, Naples, Italy

[4]Experimental Neuropsychophisiology Unit, Fondazione Santa Lucia IRCCS, Rome, Italy

[5]Center for Mind/Brain Sciences-CIMeC, University of Trento, Rovereto, Italy

[6]Brain e-Health Aging (BeA) Laboratory, Department of Neurorehabilitation, Istituti Clinici Scientifici Maugeri IRCCS, Milan, Italy

Corresponding Author:
Andrea Panzavolta, MS
IUSS Cognitive Neuroscience ICoN Center, Scuola Universitaria Superiore IUSS, Piazza della Vittoria, 15, Pavia, Italy

Abstract

Background: Digital technology in dementia is an area of great development with varying experiences across countries. However, novel digital solutions often lack a patient-oriented perspective, and several relevant barriers prevent their use in clinics.

Objective: In this study, we reviewed the existing literature on knowledge, familiarity, and competence in using digital technology and on attitude and experiences with digital tools in Alzheimer disease. The main research question is whether digital competence and attitudes of patients and caregivers may affect the adoption of digital technology.

Methods: Following the PRISMA guidelines, a literature search was conducted by two researchers in the group. Inter-rater reliability was calculated with Cohen κ statistics. The risk of bias assessment was also recorded.

Results: Of 597 initial records, only 18 papers were considered eligible. Analyses of inter-rater reliability showed good agreement levels. Significant heterogeneity in study design, sample features, and measurement tools emerged across studies. Quality assessment showed a middle-high overall quality of evidence. The main factors affecting the adoption of digital technology in patients and caregivers are severity of cognitive deficits, timing of adoption, and the availability of training and support. Additional factors are age, type of digital device, and ease of use of the digital solution.

Conclusions: Adoption of digital technology in dementia is hampered by many patient-related barriers. Improving digital competence in patient-caregiver dyads and implementing systematic, patient-oriented strategies for the development and use of digital tools are needed for a successful incorporation of digital technology in memory clinics.

(JMIR Aging 2025;8:e64324) doi:10.2196/64324

KEYWORDS

digital technology; digital e-health; accessibility; user-friendliness; neurocognitive disorders; Alzheimer disease; dementia

Introduction

The use of digital technology for the prevention, diagnosis, management, and assistance of patients with cognitive disorders and their caregivers has witnessed a dramatic increase in recent years. Telemedicine and eHealth technology services have proven to be valuable instruments for remote support and care in neurocognitive disorders, as consistently shown during the COVID-19 pandemic (see [1,2] for examples). The severe constraints on health care resources and the reduced time available to chronic patients during the pandemic have demonstrated the urgent need for the reorganization of memory clinic services to ensure adequate diagnosis and care [3]. Accordingly, many initiatives have tested digital technology solutions and telemedicine services in dementia settings [4,5], particularly in Alzheimer disease (AD), which represents one of the highest-risk and fastest-growing burdens on the health care system [6,7]. Literature evidence showed good levels of feasibility and effectiveness, particularly in individuals living in remote or underserved geographical areas or where in-person access to care facilities is limited [8]. Overall, patients with cognitive disorders and caregivers reported good satisfaction rates for digital technology, suggesting a general propensity toward adopting it routinely [8-10].

The literature highlights various solutions, including digital diagnostic instruments, tools for active and passive monitoring, and digital technologies supporting cognitive and motor rehabilitation. These tools have shown promising results not

XSL•FO
RenderX

only in optimizing diagnostic pathways, disease management, and treatments, in avoiding dysfunctional and harmful disease trajectories, but also in helping the empowerment of patients and caregivers through active engagement in their care pathways [11,12]. However, despite the growing evidence suggesting the potential benefit of digital technology tools in cognitive patients, their integration into the daily routines of memory clinics is still limited [13].

Preliminary evidence suggested the presence of enabling factors such as familiarity, acceptability, and a positive attitude toward digital technology that could facilitate its adoption [9,13,14]. Conversely, inadequate experience with technology, poor digital literacy, and low education levels, as well as insufficient accommodation for motor and sensory impairments, may cause reduced acceptability and engagement in older adults [9,15]. However, a clear understanding of factors promoting or hampering the use of digital tools in memory clinics is still lacking.

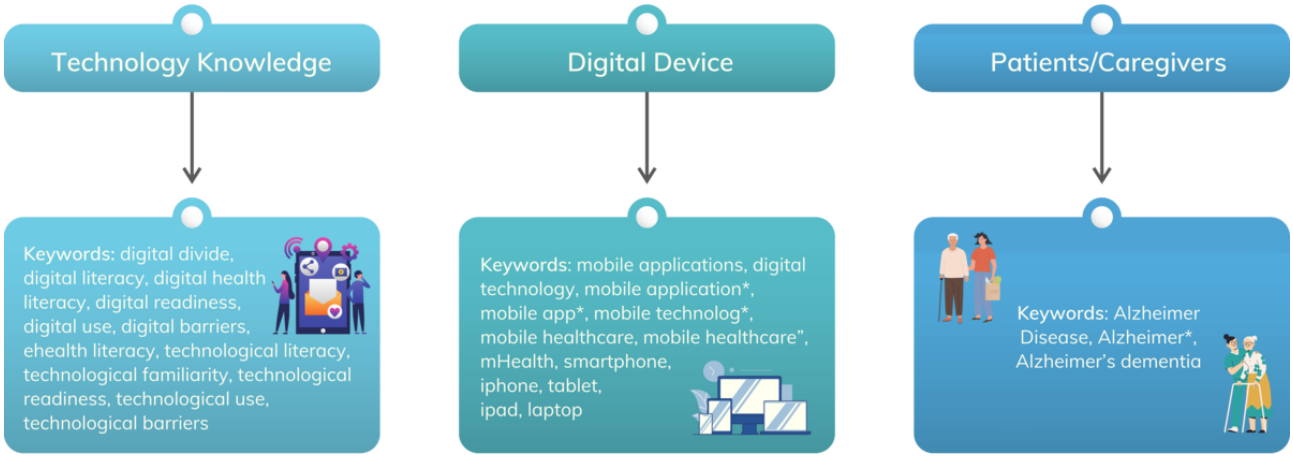
Given these considerations, we reviewed existing literature with a particular interest in studies reporting features facilitating the use of digital technology in users (eg, competence, attitudes) or evaluating the performance of digital solutions (eg, usability). The main research question is whether digital competence and the attitudes of patients and caregivers may affect the adoption of digital technology. The final goal is to explore real-world needs, facilitators, and barriers in the field and to provide patient-oriented guidance for using and implementing novel digital technology tools in memory clinics.

Methods

Search Strategy and Selection Criteria

A systematic search was conducted by two researchers of the group using PubMed, Scopus, Embase, and CINAHL (Cumulated Index in Nursing and Allied Health Literature) databases, following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [16]. According to our research hypothesis, the search strategy included a combination of MeSH (Medical Subject Headings) and relevant text terms focused on assessing familiarity and competence with digital technology and attitudes and experience regarding the use of digital solutions. See Figure 1 and Multimedia Appendix 1 for the full list of search terms. Given our theoretical framework and the growing body of literature on digital technologies in dementia, the search was focused on AD studies, aiming to improve consistency among findings.

Figure 1. Keywords for the literature search.



Inclusion criteria for paper selection were as follows: (1) paper: full original research articles (excluding conference abstracts, case reports, reviews, and book chapters); (2) population of interest: individuals with AD (ranging from mild cognitive impairment to moderate-severe dementia stages) and their caregivers; (3) measurement: papers reporting quantitative and qualitative measures (ie, questionnaires, scales, or interviews) for assessing knowledge or familiarity with digital technology and attitude or experience in using digital solutions; (4) language and time span: papers published in English from 2010 up to October 15, 2023. Articles published before 2010 were excluded due to their reliance on outdated technologies, such as personal digital assistants, which may be less relevant for a contemporary audience.

Study Selection and Data Extraction

The final set of records was uploaded onto Rayyan, a free web and mobile app designed to facilitate the initial screening of abstracts and titles through semiautomation, ensuring a high level of usability [17]. Data extraction was performed encompassing study identifiers (eg, authors, year of publication), sample features (eg, participants, diagnosis, group sizes), description of study methods (eg, type of digital solution and measurement), and findings. To ensure consistency in study selection and data extraction, the Cohen κ statistic was used to assess inter-rater reliability during the screening and eligibility phases [18]. Discrepancies were resolved through discussion, with input from an external senior expert when necessary.

Risk of Bias Assessment

A risk of bias assessment was conducted using the Appraisal tool for Cross-Sectional Studies questionnaire [19] and the National Institutes of Health (NIH) Study Quality Assessment (SQA) Tool [20] for observational cohort and cross-sectional studies. These questionnaires evaluate the quality of human research studies, providing a global score as the rate of positive answers out of either 20 or 14 questions, and indicate overall quality and bias for each study. Considering the percentage of

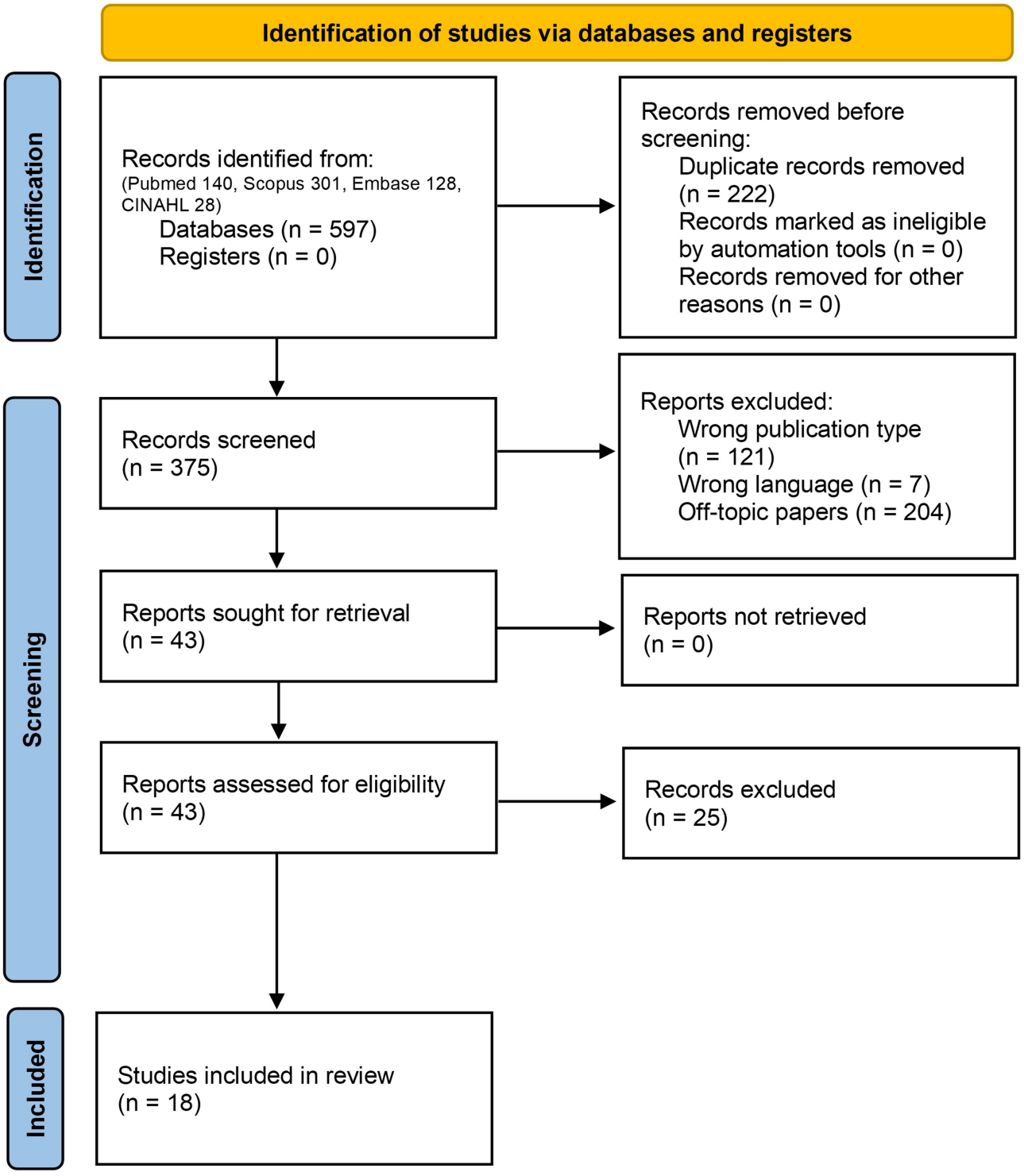
XSL•FO
RenderX

positive answers, categorical ratings were devised as follows: high quality (>70%), middle (70% - 50%), and low (≤49%) quality.

Results

Out of the 597 articles initially retrieved, 554 papers were excluded during the screening phase (ie, 222 duplicates, 121 nonoriginal papers, 204 off-topic articles, and 7 articles not in English). A detailed review of the full texts led to the exclusion of 25 articles, with 18 papers meeting the eligibility criteria for this study. See Figure 2 for details of the PRISMA flow diagram. Analyses of inter-rater reliability showed good agreement levels for both the screening (κ=0.76; 94.46% agreement) and the eligibility (κ=0.75; 87.76% agreement) phases. Quality assessment of the included studies showed a middle-high quality of evidence.

Figure 2. PRISMA flowchart reporting literature search strategy. CINAHL: Cumulated Index in Nursing and Allied Health Literature.



Retained papers showed high heterogeneity in terms of study design (ie, qualitative vs quantitative), sample features (eg, patients and caregivers with various disease severity), and data type (eg, unstructured vs structured data). Given the high

XSL•FO
RenderX

methodological and population-based heterogeneity across studies, results are presented as a narrative synthesis, focusing on shared themes and distinct sources of variability rather than on statistical aggregation.

See Tables 1 and 2 for details of included papers and the quality assessment.

XSL•FO
RenderX

Table . Summary of papers exploring knowledge, familiarity, and competence with digital technology in patients with Alzheimer disease and caregivers.

Authors and country	Year	Sample	Measurements (qualitative/quantitative)	Study findings	Quality assessment
Guzman-Parra et al [21]; Spain, Sweden	2020	1086 MCI[a] and dementia patient-caregiver dyads	Ad hoc-developed questionnaire to assess use and familiarity with touchscreen mHealth[b] technology (TechPH questionnaire [22])	Low technophilia in patients, young, male, highly educated, living with children, having better mental health	High quality (NIH[c] Study QA[d] Tool Score=75%)
Jacobs et al [23]; US	2021	68 MCI patients 121 ADD[e] patients	Ad hoc-developed web survey	Higher willingness to remote cognitive testing in MCI; significantly less access to video chat–capable technology in ADD; old age and low educational level influence access to technology and willingness to participate in remote cognitive assessment	High quality (NIH Study QA Tool Score=70%)
Christiansen et al [24]; Belgium, Spain, Sweden	2021	1082 MCI patients	Ad hoc-developed questionnaire to assess user experience with mHealth technology	Good to excellent QoL[f] in participants with moderately or high technical skills; variation in technical skills and internet use was a relevant obstacle	High quality (NIH Study QA Tool Score=70%)
Wójcik et al [25]; Poland	2021	102 ADD caregivers	Ad hoc-developed questionnaires to assess acceptance of mHealth technology and laptop/PC use	Age, gender, and education level impact on technology acceptance; old age impact on computer but not on smartphone use; digital technology was perceived as useful for daily caring (locomotion, toileting, and meals)	High quality (NIH Study QA Tool Score=80%)
Albers et al [26]; US	2022	20 people with mild-to-moderate memory complaints patient-caregiver dyads	Semistructured interviews on digital technology use pre- and post-COVID-19 pandemic, adoption of digital technology during the pandemic, facilitators, and barriers to digital technology adoption	Technology use may benefit on maintaining social connections, alleviating boredom, and fostering caregiver relief; patient dependence and low technological literacy prevented the use of digital technology	High quality (NIH Study QA Tool Score=70%)
Talbot and Briggs [27]; UK	2022	19 ADD, mixed and related dementia patients	Semistructured interviews on digital technology use and experience during the COVID-19 pandemic	Low usability and cognitive fatigue in the use of digital technologies; need for training on the use and for a better engagement	Middle quality (NIH Study QA Tool Score=60%)
Nguyen et al [28]; US	2022	124 ADD and related dementia caregivers	Ad hoc-developed web survey to assess experience with iN2L Samsung Galaxy Tablet	Benefit on caregiver wellbeing, alleviating stress, increasing satisfaction, and improving access to supportive programs	Middle quality (NIH Study QA Tool Score=64%)

XSL•FO
RenderX

Authors and country	Year	Sample	Measurements (qualitative/quantitative)	Study findings	Quality assessment
Wilson et al [29]; UK	2023	29 MCI, ADD and related dementia patient-caregiver dyads	Videoconferencing or phone interviews of mhealth technology use	Smart devices are valuable, versatile tools for essential and meaningful activities, and necessary devices to participate in modern life; need for training on the use of smart devices	Middle quality (NIH Study QA Tool Score=50%)

[a]MCI: Mild cognitive impairment.

[b]mHealth: mobile health.

[c]NIH: National Institutes of Health.

[d]QA: Quality assessment.

[e]ADD: Alzheimer disease dementia.

[f]QoL: Quality of Life.

XSL•FO
RenderX

Table . Summary of papers exploring use, acceptability, and usability of digital solutions in AD patients and caregivers.

Authors and country	Year	Sample	Digital solution	Measurements (qualitative/quantitative)	Study findings	Quality assessment
Lim et al [30]; Australia	2013	21 unspecified dementia patient-caregiver dyads	iPad and 11 iPad apps	Ad hoc-developed questionnaires to assess experience, ability to use, engagement and utility	Half of dementia patients were able to engage with and use the iPad independently	Middle quality (NIH[a] Study QA[b] Tool Score=60%)
Zmily et al [31]; Jordan	2014	10 early ADD[c] patients	Samsung Galaxy Tablet and Android app (ADcope)	Ad hoc-developed questionnaires, NASA-TLX[d] index workload assessment [32]	Low workload scores; good post-task satisfaction; successful use even without any prior experience	Middle quality (NIH Study QA Tool Score=50%)
Brown et al [33]; US	2016	11 ADD caregivers	Android app (CareHeroes App)	Ad hoc-developed web-based survey	Perceived easiness to perform tasks despite medium-low proficiency with technology	High quality (NIH Study QA Tool Score=80%)
Killin et al [34]; UK	2018	10 ADD, vascular or mixed dementia patient-caregiver dyads	Digital Support Platform (DSP)	Semistructured interviews	Caregiver use was better than that of patients; High interest in learning to use technology more effectively and enjoyed having their own tablet devices; Need of training in the use of new technology	Middle quality (NIH Study QA Tool Score=50%)
Ruggiano et al [35]; US	2019	36 ADD and related dementia caregivers	Android app (Care IT)	Interviews and focus group	eHealth and individual technologies may not fully meet the needs of caregivers; need for more effective, easy-to-use, and more widely disseminated – especially for caregivers from a disadvantaged background	High quality (NIH Study QA Tool Score=70%)
Øksnebjerg et al [36]; Denmark	2020	112 ADD, vascular dementia patients 98 caregivers	ReACT (Rehabilitation in Alzheimer disease using Cognitive support Technology) app	Ad hoc-developed web-based survey (including USEdem questionnaire)	Need for timely introduction of digital technology; need of caregiver support for the adoption of digital solution	High quality (NIH Study QA Tool Score=85%)
Evans et al [37]; UK	2021	26 ADD and related dementia patient-caregiver dyads	App-based prompter for a touchscreen Tablet	Semistructured interviews on experience and usefulness	Attitudes to technology, perceived utility, and emotional impact of needing help impact the acceptance	High quality (NIH Study QA Tool Score=80%)

XSL•FO
RenderX

Authors and country	Year	Sample	Digital solution	Measurements (qualitative/quantitative)	Study findings	Quality assessment
Berge et al [38]; Norway	2022	24 ADD patient-caregiver dyads	iOS app (Alight)	SUS[e] scale [39] modified; ad hoc-developed questionnaire to explore adoption, user-friendliness, usefulness, and impact	High adoption and feasibility; Need of timely introduction; 50% of the accepting dyads had difficulties independently managing the digital solution	Middle quality (NIH Study QA Tool Score=65%)
Skirrow et al [40]; UK, US	2022	73 MCI/mild[f] ADD patients 78 healthy control subjects	Story recall task app	Ad hoc-developed questionnaires to assess usability and task engagement	Technical problems, easy use of the app, and a broad interest in the tasks; Modest improvement of recall	High quality (NIH Study QA Tool Score=79%)
Rossetto et al [41]; Italy	2023	11 MCI, 19 mild AD patient-caregiver dyads	Telerehabilitation app and web-based software (ABILITY)	SUS scale [39], percentage of sessions attended within 6 weeks (adherence)	High adherence rate, usability, and treatment efficacy on global cognitive level in the digital-treated group	High quality (NIH Study QA Tool Score=93%)

[a]NIH: National Institutes of Health.

[b]QA: Quality assessment.

[c]ADD: Alzheimer disease dementia.

[d]NASA TLX: NASA Task load index.

[e]SUS: System usability scale.

[f]MCI: Mild cognitive impairment.

Evidence From Papers Reporting Knowledge, Familiarity, and Competence With Digital Technology

Among studies investigating factors affecting knowledge, familiarity, and competence in digital mobile health (mHealth) technology use [21,23-29], the degree of cognitive impairment was the first factor significantly affecting digital competence in the adoption of technology solutions. Individuals with milder cognitive deficits showed higher competence in the use of digital devices and the internet and consequently, greater willingness levels to participate in studies compared to those with more advanced cognitive decline [23,24]. Overall, users exhibited greater openness to digital technology when they had prior experience and familiarity and good digital literacy [21,23,24]. Significant challenges were reported in studies related to limited technical skills of both patients and caregivers [26-28]. Nonetheless, while overall caregivers showed greater competence and enthusiasm for digital technology than patients, other factors, as male sex, higher education, and younger age, impact at the individual level on familiarity and willingness to participate to studies [21,25].

Whereas smartphone acceptance is consistent across age groups, computer use declined with age and in patients versus caregivers, indicating an age-related barrier to some technologies [27-29]. For example, the use of tablets represented a significant barrier in the VITAL at Home project [28], in which caregivers required substantial guidance to effectively use the software implemented in tablets.

Evidence From Papers Reporting Acceptability and Usability of Digital Solutions

Studies examining the use of digital solutions among AD patients and caregivers [30,31,33-38,40,41] consistently found that IT system accessibility and ease of use of the proposed solution are crucial factors influencing adoption rates. Ease of use of the digital health solutions and attitude to technology represent key factors enhancing acceptance of digital tools in daily routines [30,31,33,35,37,40]. The likelihood of sustained use is related to the timing of their adoption within the disease course, with earlier implementation correlating with better long-term engagement for both patients and caregivers [34,36]. The provision of informal support and training is another enabling factor ensuring long-term and better use [31,36,41]. As the disease progresses, the severity of cognitive decline and the perceived cognitive workload pose relevant challenges for constant engagement and use [31,37,38].

Discussion

Digital technology in dementia is an area of rapid development marked by diverse approaches across countries. A huge step forward has been made during the COVID-19 pandemic, acting as a catalyst for extensive research and development of eHealth

XSL•FO
RenderX

and mHealth solutions, favoring the testing of digital solutions in memory clinics and promoting their implementation in practice [2,4,42]. In this study, we reviewed the existing literature on digital competence, attitude to technology, and use of digital tools in AD, with the final aim to underline main factors affecting the adoption of digital solutions and provide useful information for their future development and use in memory clinics.

Literature evidence revealed several barriers in adopting digital technology in dementia. Among the main factors, the severity of cognitive deficits is the first key element. As adherence to digital technology interventions decreases with the severity of cognitive decline [23,24,31,37,38], clinicians and researchers should embrace technology-based solutions at an earlier stage of the care pathway with the aim to improve use and effectiveness. With the progression of the disease, poor verbal and motor initiative may affect patient engagement with digital technology, reducing interest, initiative, and participation [27]. In such cases, behavioral or sensory-based interventions, such as sensory stimulation delivered through digital devices, represent an appropriate choice. Therefore, an in-depth assessment of the behavioral and cognitive profile should always guide the specialist to choose the right digital tool to use and address patients accordingly.

Age, digital literacy, and special needs (eg, visual or auditory impairments) are additional variables that limit the access or use of digital technology both for patients and caregivers, without affecting the will to participate in telecare services [21,25,43]. These aspects are of particular relevance in memory clinics, where final users are often patients and caregivers in old adulthood [44]. The type of digital device may also impact the feasibility and real-life implementation of digital protocols, with smartphones being more accessible compared to computers or other devices [27,29,30]. Consequently, smartphones show higher acceptability scores in patients [45,46], although some challenges in the use of touchscreen and software updates are reported, particularly among those with poor digital literacy and visual impairments [24].

Another crucial barrier to the stable use of eHealth and mHealth solutions is the need for training and support [27,29,30,34,36]. The availability of formal support in the technology setup or training has been shown to positively impact willingness and acceptability [30,34,41]. As individuals may be unlikely to embrace burdensome technologies [26], preliminary training at health care facilities before the in-home adoption should be recommended to overcome potential technophobia in patients and caregivers [13,21]. The evaluation of specific digital needs in each single patient-caregiver dyad may guarantee a more inclusive and practical approach.

Therefore, the lack of active involvement of final users in the design, development, and testing of novel digital solutions crucially contributes to the poor applicability of these instruments in real-life scenarios [47].

In addition, developing technologies without user burden is pivotal for their success. Minimizing the burden on caregivers is of paramount importance, as they are often responsible for managing digital tools (eg, calendars for medication or televisits). The mitigation of stress and anxiety due to the use of technologies is crucial for caregivers [26,35,37,38], particularly for those individuals with disadvantageous socioeconomical backgrounds, who face additional challenges in adapting to the tools and using them effectively [35]. Co-design with end-users and adaptation of the digital tool to real-life needs are therefore critical [47].

Several limitations characterize the body of research analyzed in this review. Many studies involved small sample sizes that were not fully representative of the AD population, often excluding individuals who were unable or unwilling to participate for various reasons. Availability of high-speed internet connection [23,24], recruitment or screening modality (eg, through email, social media or videoconference) [26,27], and socioeconomic status of participants [23,37] are additional variables affecting results, while study design (eg, lack of control group or randomized approach [28]) and measurement issues (eg, lack of pre-post measures [36]) are crucial aspects that hamper the reliability of data findings.

In conclusion, digital tools have the potential to significantly influence the quality of dementia services acting on different dimensions [13]. They can increase effectiveness by enabling faster access to specialist care, better diagnoses and treatments, and preventing avoidable hospitalizations. They improve timeliness, reducing waiting lists and unnecessary travel. They ensure patient-centeredness and safe care, providing care tailored to individual needs and values and treating patients more appropriately. They provide integrated, more efficient, and equitable care by taking an interdisciplinary approach involving multiple specialists, improving cost-effectiveness, and overcoming geographic barriers addressing cultural diversity. However, the adoption of digital technology is limited by many patient-related barriers. Improving digital competence in the patient-caregiver dyads and implementing a systematic patient-oriented strategy for the development and use of digital tools (eg, by promoting participated design, early timing of solution adoption and availability of training and technical support) remain critical factors to consider for the successful incorporation of digital eHealth and mHealth solutions and services into future memory clinics.

Data Availability

Data are supplied in supporting files available for download along with the published manuscript.

Authors' Contributions

Conceptualization: CC, AP

XSL•FO
RenderX

Methodology, CC, LL, AP
Software: AP, EG
Supervision: CC
Writing—original draft preparation, CC, EG, AP
Writing—review and editing: CC, AA, LL, FDL, AD, AP

Conflicts of Interest

None declared.

Multimedia Appendix 1
Combination of Medial Subject Headings (MeSH) and text words used for literature search in different databases.
[DOCX File, 22 KB - aging_v8i1e64324_app1.docx]

Checklist 1
PRISMA 2020 checklist for reporting systematic reviews and meta-analyses.
[DOCX File, 24 KB - aging_v8i1e64324_app2.docx]

References

1. Cuffaro L, Di Lorenzo F, Bonavita S, Tedeschi G, Leocani L, Lavorgna L. Dementia care and COVID-19 pandemic: a necessary digital revolution. Neurol Sci 2020 Aug;41(8):1977-1979. [doi: 10.1007/s10072-020-04512-4] [Medline: 32556746]
2. Lukas H, Xu C, Yu Y, Gao W. Emerging telemedicine tools for remote COVID-19 diagnosis, monitoring, and management. ACS Nano 2020 Dec 22;14(12):16180-16193. [doi: 10.1021/acsnano.0c08494] [Medline: 33314910]
3. Cagnin A, Di Lorenzo R, Marra C, et al. Behavioral and psychological effects of coronavirus disease-19 quarantine in patients with dementia. Front Psychiatry 2020;11(578015):578015. [doi: 10.3389/fpsyt.2020.578015] [Medline: 33033486]
4. Owens AP, Ballard C, Beigi M, et al. Implementing remote memory clinics to enhance clinical care during and after COVID-19. Front Psychiatry 2020;11:579934. [doi: 10.3389/fpsyt.2020.579934] [Medline: 33061927]
5. Tu CC, Weng SY, Hsieh NC, Cheng WC, Alizargar J, Chang KS. Increasing use of telemedicine for neurological disorders during the COVID-19 pandemic: a mini-review. J Multidiscip Healthc 2023;16:411-418. [doi: 10.2147/JMDH.S390013] [Medline: 36820220]
6. Kourtis LC, Regele OB, Wright JM, Jones GB. Digital biomarkers for Alzheimer's disease: the mobile/ wearable devices opportunity. NPJ Digit Med 2019;2(1):9. [doi: 10.1038/s41746-019-0084-2] [Medline: 31119198]
7. Staffaroni AM, Tsoy E, Taylor J, Boxer AL, Possin KL. Digital cognitive assessments for dementia. Pract Neurol (Fort Wash Pa) 2020(24) [FREE Full text] [Medline: 33927583]
8. Sekhon H, Sekhon K, Launay C, et al. Telemedicine and the rural dementia population: a systematic review. Maturitas 2021 Jan;143:105-114. [doi: 10.1016/j.maturitas.2020.09.001] [Medline: 33308615]
9. Wilson SA, Byrne P, Rodgers SE, Maden M. A systematic review of smartphone and tablet use by older adults with and without cognitive impairment. Innov Aging 2022;6(2):igac002. [doi: 10.1093/geroni/igac002] [Medline: 35243008]
10. Mendez KJW, Labrique AB, Budhathoki C, et al. Use of mobile applications and health technologies among dementia caregivers with chronic conditions: a cross-sectional study. Digit Health 2023;9(20552076231181213):20552076231181213. [doi: 10.1177/20552076231181213] [Medline: 37361436]
11. Yousaf K, Mehmood Z, Awan IA, et al. A comprehensive study of mobile-health based assistive technology for the healthcare of dementia and Alzheimer's disease (AD). Health Care Manag Sci 2020 Jun;23(2):287-309. [doi: 10.1007/s10729-019-09486-0] [Medline: 31218511]
12. Bradley L, Shanker S, Murphy J, Fenge LA, Heward M. Effectiveness of digital technologies to engage and support the wellbeing of people with dementia and family carers at home and in care homes: a scoping review. Dementia (London) 2023 Aug;22(6):1292-1313. [doi: 10.1177/14713012231178445] [Medline: 37235791]
13. Angelopoulou E, Papachristou N, Bougea A, et al. How telemedicine can improve the quality of care for patients with Alzheimer's disease and related dementias? A narrative review. Medicina (Kaunas) 2022 Nov 22;58(12):1705. [doi: 10.3390/medicina58121705] [Medline: 36556907]
14. Stronge AJ, Rogers WA, Fisk AD. Human factors considerations in implementing telemedicine systems to accommodate older adults. J Telemed Telecare 2007;13(1):1-3. [doi: 10.1258/135763307779701158] [Medline: 17288650]
15. Bossen AL, Kim H, Williams KN, Steinhoff AE, Strieker M. Emerging roles for telemedicine and smart technologies in dementia care. Smart Homecare Technol Telehealth 2015;3(49–57):49-57. [doi: 10.2147/SHTT.S59500] [Medline: 26636049]
16. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097. [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]

XSL•FO
RenderX

17. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
18. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med 2012;22(3):276-282. [doi: 10.11613/BM.2012.031] [Medline: 23092060]
19. Downes MJ, Brennan ML, Williams HC, Dean RS. Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). BMJ Open 2016 Dec 8;6(12):e011458. [doi: 10.1136/bmjopen-2016-011458] [Medline: 27932337]
20. Study quality assessment tools. NHLBI, NIH. 2021. URL: https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools [accessed 2023-04-12]
21. Guzman-Parra J, Barnestein-Fonseca P, Guerrero-Pertiñez G, et al. Attitudes and use of information and communication technologies in older adults with mild cognitive impairment or early stages of dementia and their caregivers: cross-sectional study. J Med Internet Res 2020 Jun 1;22(6):e17253. [doi: 10.2196/17253] [Medline: 32442136]
22. Anderberg P, Eivazzadeh S, Berglund JS. A novel instrument for measuring older people's attitudes toward technology (TechPH): development and validation. J Med Internet Res 2019 May 23;21(5):e13951. [doi: 10.2196/13951] [Medline: 31124467]
23. Jacobs DM, Peavy GM, Banks SJ, Gigliotti C, Little EA, Salmon DP. A survey of smartphone and interactive video technology use by participants in Alzheimer's disease research: implications for remote cognitive assessment. Alzheimers Dement (Amst) 2021;13(1):e12188. [doi: 10.1002/dad2.12188] [Medline: 34027018]
24. Christiansen L, Sanmartin Berglund J, Anderberg P, et al. Associations between mobile health technology use and self-rated quality of life: a cross-sectional study on older adults with cognitive impairment. Gerontol Geriatr Med 2021;7:23337214211018924. [doi: 10.1177/23337214211018924] [Medline: 34104685]
25. Wójcik D, Szczechowiak K, Konopka P, et al. Informal dementia caregivers: current technology use and acceptance of technology in care. Int J Environ Res Public Health 2021 Mar 19;18(6):3167. [doi: 10.3390/ijerph18063167] [Medline: 33808644]
26. Albers EA, Mikal J, Millenbah A, et al. The use of technology among persons with memory concerns and their caregivers in the United States during the COVID-19 pandemic: qualitative study. JMIR Aging 2022 Mar 17;5(1):e31552. [doi: 10.2196/31552] [Medline: 35134748]
27. Talbot CV, Briggs P. The use of digital technologies by people with mild-to-moderate dementia during the COVID-19 pandemic: a positive technology perspective. Dementia (London) 2022 May;21(4):1363-1380. [doi: 10.1177/14713012221079477] [Medline: 35333111]
28. Nguyen LT, Prophater LE, Fazio S, et al. Project VITAL at home: impact of technology on loneliness and well-being of family caregivers of people with dementia. Clin Gerontol 2024;47(1):149-160. [doi: 10.1080/07317115.2022.2123726] [Medline: 36111824]
29. Wilson SA, Byrne P, Rodgers SE. "I'd be lost without my smartphone": a qualitative analysis of the use of smartphones and tablets by people living with dementia, mild cognitive impairment, and their caregivers. Aging Ment Health 2024 Apr;28(4):595-603. [doi: 10.1080/13607863.2023.2205585] [Medline: 37139944]
30. Lim FS, Wallace T, Luszcz MA, Reynolds KJ. Usability of tablet computers by people with early-stage dementia. Gerontology 2013;59(2):174-182. [doi: 10.1159/000343986] [Medline: 23257664]
31. Zmily A, Mowafi Y, Mashal E. Study of the usability of spaced retrieval exercise using mobile devices for Alzheimer's disease rehabilitation. JMIR Mhealth Uhealth 2014 Aug 14;2(3):e31. [doi: 10.2196/mhealth.3136] [Medline: 25124077]
32. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. Human mental workload 1988:139-183. [doi: 10.1016/S0166-4115(08)62386-9]
33. Brown EL, Ruggiano N, Page TF. CareHeroes web and android. Res Gerontol Nurs 2016;9(4):193-203. [doi: 10.3928/19404921-20160229-02] [Medline: 29977440]
34. Killin LOJ, Russ TC, Surdhar SK, et al. Digital Support Platform: a qualitative research study investigating the feasibility of an internet-based, postdiagnostic support platform for families living with dementia. BMJ Open 2018 Apr 12;8(4):e020281. [doi: 10.1136/bmjopen-2017-020281] [Medline: 29654028]
35. Ruggiano N, Brown EL, Shaw S, et al. The potential of information technology to navigate caregiving systems: perspectives from dementia caregivers. J Gerontol Soc Work 2019;62(4):432-450. [doi: 10.1080/01634372.2018.1546786] [Medline: 30422754]
36. Øksnebjerg L, Woods B, Ruth K, et al. A tablet app supporting self-management for people with dementia: explorative study of adoption and use patterns. JMIR Mhealth Uhealth 2020 Jan 17;8(1):e14694. [doi: 10.2196/14694] [Medline: 31951217]
37. Evans N, Boyd H, Harris N, et al. The experience of using prompting technology from the perspective of people with Dementia and their primary carers. Aging Ment Health 2021 Aug;25(8):1433-1441. [doi: 10.1080/13607863.2020.1745145] [Medline: 32223428]
38. Berge LI, Gedde MH, Torrado Vidal JC, et al. The acceptability, adoption, and feasibility of a music application developed using participatory design for home-dwelling persons with dementia and their caregivers. The "Alight" app in the LIVE@Home.Path trial. Front Psychiatry 2022;13:949393. [doi: 10.3389/fpsyt.2022.949393] [Medline: 36061298]

XSL•FO
RenderX

39. Brooke J. SUS: A quick and dirty usability scale. Usability Eval Ind 1995 Nov 30 [FREE Full text]
40. Skirrow C, Meszaros M, Meepegama U, et al. Validation of a remote and fully automated story recall task to assess for early cognitive impairment in older adults: longitudinal case-control observational study. JMIR Aging 2022 Sep 30;5(3):e37090. [doi: 10.2196/37090] [Medline: 36178715]
41. Rossetto F, Isernia S, Realdon O, et al. A digital health home intervention for people within the Alzheimer's disease continuum: results from the Ability-TelerehABILITation pilot randomized controlled trial. Ann Med 2023 Dec;55(1):1080-1091. [doi: 10.1080/07853890.2023.2185672] [Medline: 36929703]
42. León-Salas B, González-Hernández Y, Infante-Ventura D, et al. Telemedicine for neurological diseases: a systematic review and meta-analysis. Eur J Neurol 2023 Jan;30(1):241-254. [doi: 10.1111/ene.15599] [Medline: 36256522]
43. Conway A, Ryan A, Harkin D, Mc Cauley C, Goode D. A review of the factors influencing adoption of digital health applications for people living with dementia. Digit Health 2023;9:20552076231162985. [doi: 10.1177/20552076231162985] [Medline: 36937696]
44. Canevelli M, Di Pucchio A, Marzolini F, et al. A national survey of centers for cognitive disorders and dementias in Italy. J Alzheimers Dis 2021;83(4):1849-1857. [doi: 10.3233/JAD-210634] [Medline: 34459403]
45. Joe J, Demiris G. Older adults and mobile phones for health: a review. J Biomed Inform 2013 Oct;46(5):947-954. [doi: 10.1016/j.jbi.2013.06.008] [Medline: 23810858]
46. Koo BM, Vizer LM. Mobile technology for cognitive assessment of older adults: a scoping review. Innov Aging 2019 Jan;3(1):igy038. [doi: 10.1093/geroni/igy038] [Medline: 30619948]
47. Bitkina O, Kim HK, Park J. Usability and user experience of medical devices: an overview of the current state, analysis methodologies, and future challenges. Int J Ind Ergon 2020 Mar;76:102932. [doi: 10.1016/j.ergon.2020.102932]

Abbreviations

AD: Alzheimer disease
CINAHL: Cumulated Index in Nursing and Allied Health Literature
NIH: National Institutes of Health
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SQA: Study Quality Assessment

Edited by AF Abdulai; submitted 15.07.24; peer-reviewed by F Yu, L Cuffaro, RS Gomaa Mahmoud; revised version received 09.12.24; accepted 01.01.25; published 10.04.25.

Please cite as:
Panzavolta A, Arighi A, Guido E, Lavorgna L, Di Lorenzo F, Dodich A, Cerami C
Patient-Related Barriers to Digital Technology Adoption in Alzheimer Disease: Systematic Review
JMIR Aging 2025;8:e64324
URL: https://aging.jmir.org/2025/1/e64324
doi:10.2196/64324

© Andrea Panzavolta, Andrea Arighi, Emanuele Guido, Luigi Lavorgna, Francesco Di Lorenzo, Alessandra Dodich, Chiara Cerami. Originally published in JMIR Aging (https://aging.jmir.org), 10.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Frequency of Electronic Personal Health Record Use in US Older Adults: Cross-Sectional Study of a National Survey

Lavlin Agrawal[1*], PhD; Richelle Oakley DaSouza[1*], PhD; Pavankumar Mulgund[2*], PhD; Pankaj Chaudhary[1*], PhD

[1]Department of Business Information Systems and Analytics, Willie A. Deese College of Business and Economics, North Carolina Agricultural and Technical State University, 1601 E. Market Street, Greensboro, NC, United States

[2]Department of Management Information Systems, Fogelman College of Business & Economics, University of Memphis, Memphis, United States

[*]all authors contributed equally

Corresponding Author:
Lavlin Agrawal, PhD
Department of Business Information Systems and Analytics, Willie A. Deese College of Business and Economics, North Carolina Agricultural and Technical State University, 1601 E. Market Street, Greensboro, NC, United States

Abstract

Background: Electronic personal health records (ePHRs) hold significant potential to improve health management for older adults by enhancing access to medical information and facilitating communication with health care providers. However, usage remains low among individuals aged 65 and older. While existing research has identified barriers such as low self-efficacy, limited digital literacy, and usability challenges, the specific factors influencing the use of ePHRs among older adults are not yet fully understood.

Objective: This study integrates the Aging and Technology framework with the Patient Technology Acceptance Model to examine key predictors of ePHR use among older adults, including age, education, issue involvement, performance expectancy, effort expectancy, and self-efficacy, while controlling for demographic factors such as gender, race, and income.

Methods: This study utilizes data from the Health Information National Trends Survey (HINTS 5 cycle 3), which includes 532 respondents representing 13,136,180 US adults aged 65 years and older, after applying survey weights. Structural equation modeling was used to analyze the factors influencing the frequency of ePHR access over the past 12 months.

Results: The findings indicate that older adults with higher self-efficacy used ePHRs more frequently. Additionally, issue involvement, performance expectancy, and effort expectancy were positively associated with ePHR use. Notably, self-efficacy partially mediated the relationship between age and the frequency of ePHR use.

Conclusions: These findings suggest that enhancing self-efficacy, improving usability, and increasing the perceived benefits of ePHRs are critical for boosting usage among older adults. The study underscores the need for targeted interventions to support older users, simplify digital interfaces, and provide accessible educational resources, ultimately contributing to better health outcomes and an improved quality of life for older adults.

(JMIR Aging 2025;8:e71460) doi:10.2196/71460

KEYWORDS

electronic personal health record; aging population; self-efficacy; technology adoption; health IT; patient engagement

Introduction

Background

Health information technologies (HITs), such as electronic health records (EHRs) and electronic medical records (EMRs), are pivotal tools in modernizing health care by enhancing efficiency and improving the quality of care [1-3]. Beyond these general applications, HIT has the potential to address the unique medical needs of the older population. Extensive research demonstrates that HIT use can significantly improve health outcomes for older adults [4], particularly through tools such as EMRs and electronic personal health records (ePHRs). For example, EMRs in hospital settings help health care professionals identify at-risk patients, such as those prone to falls, by analyzing factors such as age and medication use [5]. Similarly, ePHRs empower older adults by providing easier access to health information, fostering improved communication with health care providers, and encouraging active engagement in personal health care management [6].

Luo et al [7] highlighted that, among older adults, the use of ePHRs was positively associated with the ease of understanding health information. ePHRs enable patients to track their health history to better understand the progress or deterioration of their health conditions. Patients with access to their health information experience increased medical data transparency, which reduces medical errors, improves trust in care providers, and enhances

XSL•FO
RenderX

patient satisfaction [8]. As such, in the United States, the adoption of ePHRs has steadily increased. This growth has been supported by policies such as the Health Information Technology for Economic and Clinical Health Act of 2009 [9], which incentivized health care providers to adopt EHR systems through financial rewards, indirectly creating an infrastructure supportive of ePHRs. The subsequent Meaningful Use Program set specific objectives for health care providers to demonstrate effective use of EHRs, including patient-access capabilities. Further support came from the 21st Century Cures Act of 2016 [10], which enhanced patient access to electronic health information and prohibited information blocking, thereby facilitating smoother integration and utilization of ePHRs.

While the benefits of HIT for older persons are well-documented, the adoption and sustained use of these technologies face notable barriers, such as lower levels of technology usage among older adults [11]. Recent data indicate that approximately 4 in 10 individuals aged 65 years and older utilize ePHRs, a rate significantly lower than that of younger populations [12]. Furthermore, a previous study reported that approximately 54% of US adults have been offered access to their online medical records, and among those offered, 57% accessed them at least once, translating to about 31% of the adult population accessing their online medical records [13]. However, specific statistics for individuals aged 65 years and older are not provided in this source. Given that older adults are generally less likely to use the internet, it is reasonable to infer that their rates of ePHR usage may be lower than those of younger populations [12].

Further, researchers have found that older age is negatively associated with HIT [14]. Reduced cognitive and motor abilities, common among older adults, often complicate interactions with modern digital applications [15,16]. van der Vaart et al [8] discussed how older age is associated with the nonuse of web-based portals. Czaja and Lee [17] argued that although older adults are willing to use technology, many report usability problems with existing systems, which may, in part, be due to the cognitive and perceptual demands placed on the user. Their findings show that self-efficacy is an important predictor of general use of technology and that people with lower self-efficacy are less likely to use technology. However, individual-level barriers such as confidence and usability are only part of the picture. A growing body of research suggests that technologies are often not designed with the specific needs of older adults in mind [18,19]. Studies have also shown that older patients may not be offered digital health tools for various reasons, including—but not limited to—age-related stereotypes held by health care professionals or concerns about overburdening older patients [20,21]. Together, these individual and environmental barriers create a complex landscape that can inhibit older adults' use of health technologies.

Despite advancements in technology, significant gaps remain in understanding the factors that influence the regular use of ePHRs among older adults. Existing research has focused on general barriers and facilitators, but limited knowledge exists about the interplay between these factors and how they shape the frequency of ePHR usage. This knowledge gap calls for a deeper examination of the factors affecting ePHR use in older populations. Additionally, limited electronic health literacy remains a critical challenge that hinders ePHR use in older populations [22]. However, the potential benefits cannot be overlooked. ePHRs offer older adults opportunities to coordinate their health care by sharing critical health information with providers and other stakeholders, an approach particularly crucial for managing chronic conditions that require multifaceted treatment strategies [7].

There are still unanswered questions about how older adults use ePHRs, particularly regarding the specific factors that facilitate or hinder ePHR use and the extent to which these factors influence it. To address this gap, this study explores the following research questions:

- What are the key factors influencing the use of ePHRs among older adults?
- How do age and education mediate the impact of self-efficacy on the frequency of ePHR use in older populations?
- What roles do facilitators such as issue involvement, performance expectancy, and effort expectancy of ePHRs play in encouraging regular use?

To answer these questions, we build on established theoretical frameworks such as the Patient Technology Acceptance Model (PTAM) [23], which provides insights into technology adoption and use processes. We extend the PTAM by incorporating factors specifically relevant to older persons, such as usability challenges, limited education, and issue involvement, to develop a comprehensive understanding of the dynamics at play. This study aims to contribute to the literature by offering empirical evidence and actionable insights into how ePHRs can be effectively designed and implemented to support the unique health care needs of aging populations.

Hypotheses and Proposed Model

Overview

This study aims to deepen the understanding of the frequency of PHR use among the older population by addressing critical individual factors that influence usage. A nuanced understanding of these factors is essential for improving ePHR-related outcomes, such as patient empowerment, health care accessibility, and overall health management. Specifically, this research investigates the determinants influencing the frequency of ePHR use among individuals aged 65 years and older.

Based on the existing literature, we contend that theoretical frameworks grounded in the Technology Acceptance Model (TAM) [24], including the PTAM [23], are particularly suitable for investigating frequency of use, aligning well with our study's specific outcomes. Previous research applying the TAM has effectively examined ongoing usage rather than merely initial adoption. For instance, McCloskey [25] utilized the TAM to evaluate sustained electronic commerce (e-commerce) behaviors, specifically the frequency of online purchasing, and found that perceived usefulness directly influenced the frequency of technology use. Similarly, Lederer et al [26] applied the TAM to explore frequent usage of the World Wide Web, explicitly addressing ongoing use rather than initial acceptance. Moreover, Martín-García et al [27] extended the TAM framework to

XSL•FO
RenderX

examine older adults' frequency of using technological devices, demonstrating the applicability of TAM constructs beyond initial intentions to sustained technological engagement.

Building upon this foundation, we integrate Czaja et al's [28] conceptual model of aging and technology with the PTAM [23]. Our selection of the PTAM specifically addresses our research context—older patients managing multiple chronic conditions at home using digital health tools. Unlike broader acceptance models such as the TAM [24], the Unified Theory of Acceptance and Use of Technology (UTAUT) [29], and the Health Belief Model [30], PTAM explicitly considers patient-centric variables critical for older populations managing complex health conditions.

For example, while the TAM broadly emphasizes general technological acceptance through perceived ease of use and perceived usefulness, it lacks explicit patient-oriented or health outcome considerations [31]. Similarly, the UTAUT primarily addresses technology acceptance within general workplace contexts rather than the nuanced health care needs of older patients with chronic conditions [32]. The Health Belief Model likewise falls short by inadequately addressing the interactive dynamics unique to older patient populations [33].

By contrast, PTAM's incorporation of patient-specific constructs—such as perceived health improvements, patient-provider relationships, and health care contextual factors—offers a more precise and comprehensive understanding of older patients' sustained engagement with digital health interventions. This tailored approach provides richer theoretical insights and clearer practical implications, significantly enhancing explanatory power and applicability compared with general acceptance models.

Age

The utilization of ePHRs and patient portals exhibits notable age-related disparities. Despite being frequent consumers of health care services, older adults are significantly less likely to use these digital tools compared with younger individuals [34]. This gap in use can be attributed to age-related barriers, such as limited digital literacy and difficulties in navigating technological interfaces, which become increasingly pronounced with age [34,35].

A systematic review of patient acceptance of HIT underscored this trend, finding a negative correlation between age and HIT acceptance [14]. Similarly, Heart and Kalderon [36] provided compelling evidence that older adults demonstrate lower use of health-related information and communication technologies. Additionally, van der Vaart et al [8] highlighted that older individuals are generally less likely to use patient web portals. Based on these findings, we propose the following hypothesis:

- *H1a: Age is negatively associated with the frequency of ePHR use.*

Research has consistently shown that as individuals age, their electronic self-efficacy—the belief in their ability to effectively use technology—tends to decline. This decrease in self-efficacy is often attributed to a combination of factors, including reduced exposure to new technologies, cognitive decline, and negative stereotypes associated with older adults and technology use. For example, studies suggest that older adults are more likely to perceive technology as complex and overwhelming, which can lower their confidence in using digital tools such as EMRs or patient portals [28]. With age, physiological capacity and the ability to respond to environmental stresses decline, further contributing to the reduction in electronic self-efficacy [37]. Hence, we hypothesize:

- *H1b: Age is negatively associated with self-efficacy.*

Education

Smith et al [38] argued that individuals with higher levels of education are more likely to register for and utilize patient portals. Empirical evidence supports this claim, indicating that higher educational attainment is positively correlated with increased patient portal usage [39]. Educated individuals are more likely to engage with digital platforms for managing their health information, as they typically possess greater health literacy and familiarity with technology. Furthermore, a recent literature review highlights that lower educational levels act as a barrier to ePHR use [40]. Similarly, individuals with higher education are more likely to perceive the benefits of ePHRs, leading to higher usage rates [41]. Based on these findings, we propose the following hypothesis:

- *H2a: Education is positively associated with the frequency of ePHR use*.

Research consistently demonstrates that education and training play a pivotal role in enhancing individuals' electronic self-efficacy, particularly in the context of using digital tools such as patient portals and EMRs. Czaja et al [28] highlighted a positive association between education and self-efficacy, emphasizing the importance of foundational skills acquired through educational experiences. These skills, including problem-solving, critical thinking, and basic computer literacy, are critical for successfully navigating digital platforms [42]. Furthermore, education has been shown to empower individuals across all age groups to overcome barriers to technology use. Based on this, we propose the following hypothesis:

- *H2b: Education is positively associated with self-efficacy.*

Issue Involvement

Issue involvement in the health care domain refers to "how relevant a specific health issue is to a patient" [43,44]. A more involved patient typically has a severe health condition and frequently visits health care providers [45]. Ross et al [46] argued that issue involvement has a significant positive impact on the use of patient-accessible medical records. This finding is further confirmed by Angst and Agarwal [45] and Abdelhamid et al [43]. A more involved patient is more likely to use ePHRs, as they help patients better prepare for upcoming visits with physicians by enhancing their knowledge of their medical condition, increasing their sense of control, and allowing them to seek clarification about treatment. Hence, we hypothesize:

- *H3: Issue involvement is positively associated with the frequency of ePHR use.*

XSL•FO
RenderX

Performance Expectancy

Performance expectancy is defined as the degree to which a person feels that using a system will help them perform a job more efficiently [29]. In keeping with this understanding, we refer to performance expectancy as the degree to which the patient believes that using ePHRs helps them monitor their health. Venkatesh et al [29] theorized that performance expectancy drives the intention to use information systems. Several researchers have also identified performance expectancy as one of the critical predictors of eHealth acceptance and use [40,47-49]. Patient portals help improve patient engagement and empower individuals to access their health information anytime and anywhere [50]. Thus, we hypothesize:

- *H4: Performance expectancy is positively associated with the frequency of ePHR use.*

Effort Expectancy

Venkatesh et al [29] defined effort expectancy as the degree of comfort associated with system use. Consistent with Venkatesh et al [29], we define effort expectancy as the degree of ease associated with understanding the health information in the online medical record. Effort expectancy is another key variable that drives use intentions [46]. Many researchers have also identified effort expectancy as one of the critical predictors of health technology use [2,23,37,48]. Therefore, we also propose the following hypothesis:

- *H5: Effort expectancy is positively associated with the frequency of ePHR use.*

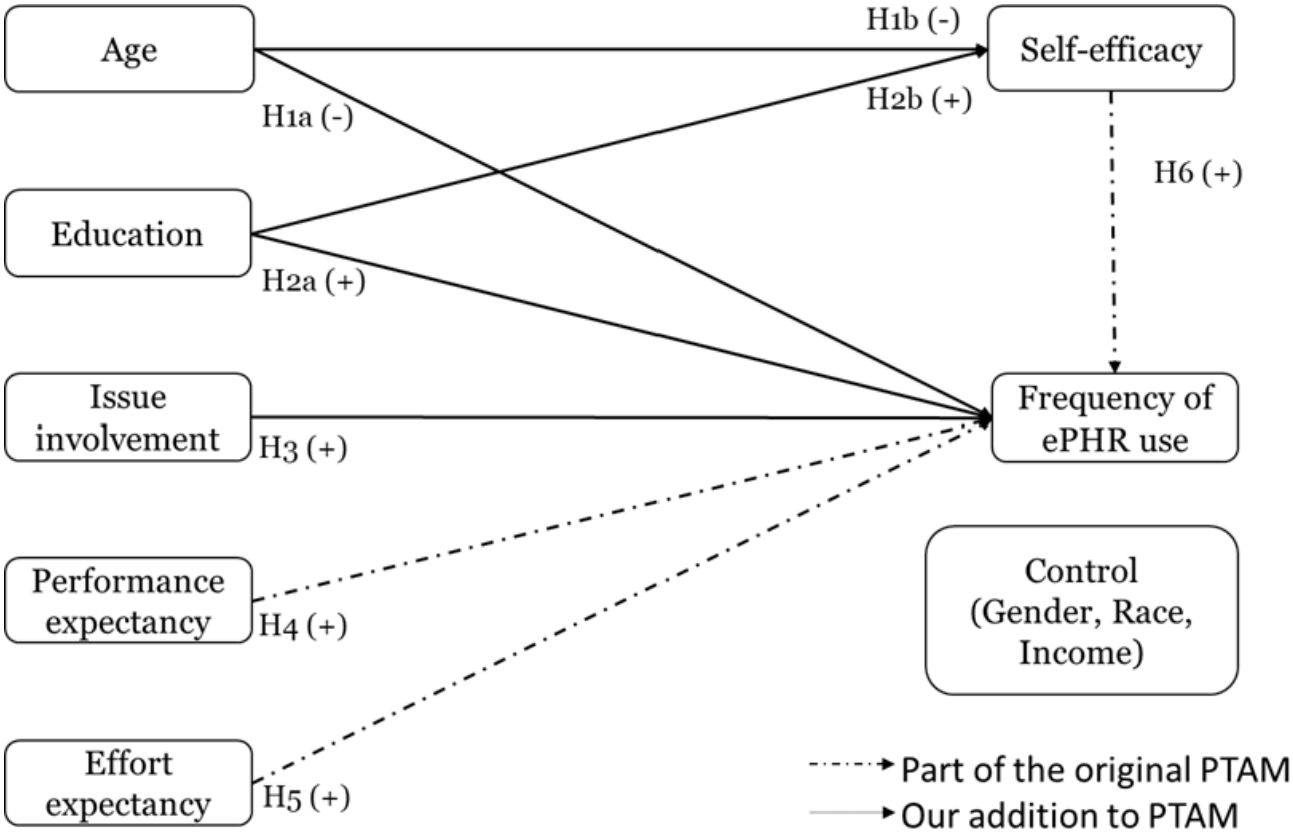
Self-Efficacy

Research has consistently demonstrated that electronic self-efficacy—the belief in one's capability to effectively use technology—declines with age due to factors such as reduced exposure to new technologies, cognitive decline, diminished physiological capacity, and negative stereotypes about older adults' technology use [28,37]. Consequently, older adults often perceive technology, including ePHRs, as complex and overwhelming, undermining their confidence in engaging with such digital tools. Therefore, we selected self-efficacy as a mediator to explicitly capture the psychological mechanism linking age with ePHR use. Rather than age directly limiting ePHR engagement, we posit that reduced electronic self-efficacy partially explains why older adults use ePHRs less frequently. This approach emphasizes the central role of self-efficacy in technology use among aging populations [36,51]. Hence, we hypothesize:

- *H6: Self-efficacy is positively associated with the frequency of ePHR use.*

Our proposed model is depicted in Figure 1. Additionally, we control for demographic factors such as gender, race, and income to account for their potential influence on the frequency of ePHR use.

Figure 1. The research model. ePHR: electronic personal health record; PTAM: Patient Technology Acceptance Model.



XSL•FO
RenderX

Methods

Data Source

This study used publicly available cross-sectional data collected in 2019 by the National Cancer Institute (Health Information National Trends Survey [HINTS] 5 cycle 3). The cross-sectional analysis for this study was limited to a subset of the original dataset and included participants aged 65 years or older. The total number of individuals aged 65 years or older was 1961. Of these, 1234 had not used an ePHR at least once in the past 12 months and were therefore excluded from the analysis. Further, 195 responses were removed due to missing values for key variables. The final sample consisted of 532 respondents, representing 13,136,180 US adults aged 65 and older when survey weights were applied.

Ethical Considerations

The HINTS 5 survey, conducted with the general population, underwent expedited review and received approval from the Westat Institutional Review Board on March 28, 2016 (project number 6048.14). This analysis used deidentified, publicly available data from HINTS, which did not constitute human research as defined by 45 CFR 46.102 and, therefore, did not require Institutional Review Board review.

Measurements

In this study, the primary dependent variable, “frequency of ePHR use,” is operationalized using a single-item measure assessing how frequently patients accessed their online medical records within the past 12 months. The response options are categorized as follows: 1-2 times, 3-5 times, 6-9 times, and 10 or more times. Individuals who reported not accessing their online medical records during the past year were excluded from the analysis, as critical variables pertinent to these respondents were not captured.

Performance expectancy, representing the perceived usefulness of online medical records for health monitoring, is measured using a 5-point Likert scale, ranging from “5=very useful” to “1=don’t use.” Effort expectancy, reflecting the perceived effort necessary to understand health information in online medical records, is assessed using a 4-point Likert scale ranging from “4=very easy” to “1=very difficult.” Issue involvement captures an individual’s engagement with personal health management and is operationalized as the number of interactions a respondent has had with health care providers over the past year. Self-efficacy in accessing electronic health records is assessed through a binary-choice question indicating whether respondents have utilized any electronic method to access their medical records during the past 12 months.

Single-item measures are used for several constructs in this study. Such measures are considered acceptable when questions are clear, unambiguous, and not prone to multiple interpretations [52]. Additionally, single-item measures are prevalent and widely validated within information systems research, particularly in studies applying structural equation modeling (SEM) in health care contexts [43,44].

Gender, race, and income are incorporated as control variables, consistent with prior studies [43,44]. For comprehensive details on questionnaire items, scales, and specific variable operationalizations, please refer to Multimedia Appendix 1.

Statistical Analysis

This study uses SEM to conduct a path analysis. Although SEM is predominantly used to model latent variables, it is also commonly applied to conduct a path analysis in mediation models. We examine 2 mediating relationships: first, self-efficacy mediates the relationship between age and the frequency of ePHR use; second, self-efficacy also mediates the relationship between education and the frequency of ePHR use. Accordingly, we use SEM to test the model, consistent with prior research [53,54]. We used SEM with weighted least squares mean and variance adjusted estimation to test the hypotheses. Weighted least squares mean and variance adjusted is well-suited for models with ordinal outcome variables [55,56]. The analysis was conducted in R (R Foundation) using the “lavaan.survey” package. We incorporated HINTS-supplied survey weights and used jackknife variance estimation techniques to account for the complex HINTS sampling design and to calculate nationally representative estimates [57].

Results

Descriptive Statistics

Table 1 shows the descriptive statistics of the survey respondents. The survey included questions about the frequency of participants’ ePHR use, as well as questions related to the model variables, including performance expectancy, effort expectancy, issue involvement, self-efficacy, age, education, gender, race, and income.

XSL•FO
RenderX

Table . Descriptive statistics.

Demographic characteristics	Sample size (N=532)[a]	With survey weights (N=13,136,180)[b]
Gender		
Male	249 (46.8)	5,729,371
Female	283 (53.2)	7,406,809
Race		
White	458 (86.1)	11,961,412
Non-White	74 (13.9)	1,174,768
Education		
Less than high school	7 (1.3)	278,445
High school or higher	525 (98.7)	12,857,735
Income (US $)		
Less than 20,000	57 (10.7)	1,191,896
20,000 to <35,000	68 (12.8)	2,009,787
35,000 to <50,000	86 (16.2)	2,100,148
50,000 to <75,000	122 (22.9)	3,110,545
75,000 or higher	199 (37.4)	4,723,804
Age (years)		
Range	65-97	65-97
Mean (SD)	71.68 (5.66)	71.90 (6.09)

[a]The sample size for gender, race, education, and income is presented as n (%).

[b]Data are presented as n for gender, race, education, and income.

Reliability and Validity

Table 2 presents the correlations between all variables of interest. Correlation coefficients are important because high correlations among independent variables may indicate multicollinearity, which can introduce bias into the model results. Multicollinearity is not a concern in this analysis, as all correlations fall within the acceptable threshold of 0.6 [58], with the highest correlation being 0.37 between performance expectancy and effort expectancy. Table 2 also provides the means and SDs for the principal variables.

Table . Correlation matrix.

Correlation	Mean (SD)	Performance expectancy	Effort expectancy	Issue involvement	Self-efficacy
Performance expectancy	4.15 (1.03867)	1.00	—[a]	—	—
Effort expectancy	3.30 (0.67763)	0.37	1.00	—	—
Issue involvement	3.80 (1.49733)	−0.01	−0.14	1.00	—
Self-efficacy	0.82 (0.38126)	0.17	0.06	0.09	1.00

[a]Not applicable.

Common Method Variance

As data are self-reported and collected through a single survey, they may suffer from common method variance (CMV), which hampers the relationship between the variables [59]. We used the marker variable technique [60] to check if the data are suffering from CMV. Marker variable is a variable that is theoretically unrelated to 1 or more of the principal variables measured in the study and typically has a low correlation with the central variables.

The theoretically unrelated construct “morning person or night person” was used as a marker variable. The correlations between the marker variable morning person or night person and other principal variables were below the threshold of 0.1 (performance expectancy: −0.08; effort expectancy: −0.08; self-efficacy: 0.03) [60], except for issue involvement (0.17). The low correlation of the marker variable with the principal variables in the model indicates that CMV is not a problem.

XSL•FO
RenderX

Data Analysis

As the National Cancer Institute administered both a paper-based questionnaire and an online questionnaire to survey participants, we first assessed whether the mode of survey introduced any biases. We regressed the dependent variable, "frequency of ePHR use," on the mode of survey and found that the relationship between the 2 was nonsignificant (*P*=.48). Thus, the mode of survey administration did not introduce any bias in the main outcome variable, that is, "frequency of ePHR use."

The SEM results are presented in Table 3. The overall fit statistics (χ^2_{17}=13.230, *P*=.004; Comparative Fit Index=0.950, Tucker-Lewis Index=0.715, root-mean-square error of approximation=0.080, standardized root-mean-square residual=0.022) indicate a good model fit [61]. Table 4 presents the results of the mediation analysis for self-efficacy with age and education.

The results show that age is negatively associated with self-efficacy (β=−.9890, *P*=.03), suggesting that as age increases, self-efficacy decreases, supporting H1b. However, no significant (*P*=.85) relationship was found between self-efficacy and education (H2b) in our research context.

Further, the results show that age is positively associated with the frequency of ePHR use (β=1.4950, *P*=.03), suggesting that as age increases, ePHR use also increases. This is the opposite of what we hypothesized for H1a. This counterintuitive trend may reflect increased health care needs among older adults, leading to greater engagement with ePHRs for managing chronic conditions, medication schedules, and communication with physicians. No significant (*P*=.88) relationship was found between education and the frequency of ePHR use (H2a) in our research context. H3 proposed a positive relationship between issue involvement and the frequency of ePHR use. The path coefficient was positive and statistically significant (β=.2560, *P*<.001), suggesting that higher issue involvement leads to higher ePHR use.

Our analysis also revealed a significant positive relationship between performance expectancy and the frequency of ePHR use (β=.2470, *P*<.001), as well as between effort expectancy and the frequency of ePHR use (β=.1850, *P*=.04). These findings suggest that higher performance expectancy and higher effort expectancy both lead to increased ePHR use, supporting H4 and H5. We also found that self-efficacy is positively associated with the frequency of ePHR use (β=.4990, *P*<.001), supporting H6.

Further analysis confirms that self-efficacy partially mediates the relationship between age and the frequency of ePHR use among older adults, as evidenced by the statistical significance of both indirect (*P*=.03) and direct effects (*P*=.03). Specifically, self-efficacy accounts for 49.2% (−0.493/1.002) of the total effect, underscoring its substantial role as a mediator in this relationship. However, we did not find any statistically significant mediating effect of self-efficacy on the relationship between education and the frequency of ePHR use.

Table . Results of structural equation modeling.

Dependent variable and hypothesis	Variables	Estimate	95% CI	*P* value	Significant
Self-efficacy					
H1b	Log (age)	−0.9890	−1.8870 to −0.0900	.03	Yes
H2b	High school or more	−0.0360	−0.4070 to 0.3350	.85	No
Control	Female	−0.0190	−0.1300 to 0.0910	.73	No
Control	White	−0.0040	−0.1460 to 0.1370	.95	No
Control	Income	0.0120	−0.0200 to 0.0430	.47	No
Dependent variable: frequency of electronic patient health record use					
H1a	Log (age)	1.4950	0.1460 to 2.8440	.03	Yes
H2a	High school or more	0.0320	−0.3770 to 0.4410	.88	No
H3	Issue involvement	0.2560	0.1810 to 0.3300	<.001	Yes
H4	Performance expectancy	0.2470	0.1340 to 0.3600	<.001	Yes
H5	Effort expectancy	0.1850	0.0070 to 0.3640	.04	Yes
H6	Self-efficacy	0.4990	0.3260 to 0.6730	<.001	Yes
Control	Female	0.1970	−0.0410 to 0.4340	.10	No
Control	White	0.0590	−0.3050 to 0.4220	.75	No
Control	Income	−0.0300	−0.1290 to 0.0680	.54	No

XSL•FO
RenderX

Table . Mediation results of structural equation modeling.

Variables	Estimate	95% CI	*P* value	Significant
Log (age)				
Indirect through self-efficacy	−0.4930	−0.9410 to −0.0460	.03	Yes
Direct	1.4950	0.1460 to 2.8440	.03	Yes
Total	1.0020	−0.3170 to 2.3200	.14	No
High school or more				
Indirect through self-efficacy	−0.0180	−0.2040 to 0.1680	.85	No
Direct	0.0320	−0.3770 to 0.4410	.88	No
Total	0.0140	−0.4860 to 0.5140	.96	No

Discussion

Determinants of ePHR Use

Our study aimed to update the PTAM by examining the impact of performance expectancy, effort expectancy, and self-efficacy on the frequency of ePHR Use. By using an integrated framework, this research provides actionable insights into the factors driving ePHR usage among older adults, incorporating variables that have not been widely explored in prior studies. This study makes both a significant theoretical and practical contribution.

Theoretical Implications

This study contributes to the literature on ePHR use by building on and extending existing theoretical frameworks, including the Aging and Technology framework [28] and the PTAM [23]. Prior research has predominantly highlighted barriers to ePHR use among older adults, such as reduced self-efficacy, usability challenges, and low digital literacy [15,17]. While these studies provide an important foundation, our findings challenge and refine these perspectives by demonstrating a positive association between age and ePHR use, with self-efficacy serving as a partial mediator in this relationship.

Contrary to earlier work suggesting older age is a barrier to HIT use [8,14], this study reveals that age positively influences ePHR use when health needs increase, as chronic illnesses and frequent health care interactions necessitate greater reliance on digital tools. This finding aligns with studies emphasizing the contextual nature of technology use, where health-related motivations can offset age-related challenges [7]. By demonstrating that older adults use ePHRs more frequently despite lower self-efficacy, this study extends theoretical models by incorporating health-related drivers, such as issue involvement, into the broader narrative of technology adoption.

Self-efficacy, identified as a strong predictor and mediator, validates and expands prior work in the field. Studies have consistently highlighted self-efficacy as a key determinant of technology use [62]. Our findings reinforce this and reveal that self-efficacy mediates the impact of both age and education on ePHR use. This underscores the critical role of psychological factors, particularly in populations facing cognitive and physical challenges, and aligns with theories from the PTAM [23] and the UTAUT [29].

The study also affirms the relevance of performance expectancy and effort expectancy as critical drivers of ePHR usage, consistent with the PTAM and UTAUT frameworks [47,48]. However, by identifying issue involvement as a significant predictor, this study contributes a novel dimension to the literature. This finding supports prior arguments that perceived relevance and personal health involvement significantly enhance engagement with health technologies [43,45].

Overall, this research advances theoretical understanding by integrating psychological, motivational, and contextual factors into established frameworks for HIT use. It bridges gaps in the existing literature by demonstrating that age, when considered alongside mediating factors such as self-efficacy and health involvement, can positively influence ePHR use. These insights offer a nuanced understanding of older adults' interactions with health technologies, offering a robust foundation for future research and practical interventions.

Practical Implications

Overview

The findings of this study underscore several actionable strategies for health care providers, policy makers, and technology developers to enhance ePHR usage among older adults. By addressing barriers and leveraging the facilitators identified, targeted interventions can empower older adults to better manage their health through ePHRs, contributing to improved health outcomes and broader system-wide benefits, as outlined below.

Building Self-Efficacy Through Tailored Training Programs

To enhance older adults' self-efficacy in using ePHRs, health care organizations can establish targeted training and education programs aligned with the recommendations of the Office of the National Coordinator for Health Information Technology [63]. These programs could include workshops in collaboration with senior community centers and local health care providers, focusing on essential tasks such as medication tracking, viewing test results, and scheduling appointments [64]. Additionally, peer-led mentoring initiatives, consistent with community-based

XSL•FO
RenderX

health literacy efforts promoted by Healthy People 2030, can encourage skill sharing among older adults [65].

Supporting Continuous Improvement and User-Centered Innovation

Continuous evaluation through user feedback mechanisms should be institutionalized to guide the ongoing improvement of ePHR technologies [66], particularly for older adults. Actively encouraging the participation of older adults in innovation processes can help health care organizations advance the user-centered innovation strategies promoted by federal health technology initiatives [63]. Collaboration among academia, health care providers, and technology developers can further drive innovation tailored to the evolving needs of older adults.

Promoting Health Involvement Through Guided Support

Health care providers can encourage older adults' active participation in managing their health through targeted education campaigns that highlight the practical benefits of ePHRs. Public health strategies that emphasize patient-provider communication, administrative ease, and improved health care outcomes can be strengthened by incorporating user testimonials and success stories [67]. Additionally, guided tutorials embedded within ePHRs can further support patient engagement, aligning with strategies recommended by the Agency for Healthcare Research and Quality for patient empowerment [68].

Implementing these recommendations, grounded in existing US health care policies and strategies, can significantly enhance ePHR adoption among older adults, help reduce health disparities, and promote a more efficient and patient-centered health care system.

Limitations

This study has several limitations that should be considered. First, the use of secondary data from the HINTS limited the analysis to variables available in the dataset. Important factors, such as detailed measures of digital literacy or prior technology experience, could not be examined. Further, this study included only individuals who reported using ePHRs at least once in the past 12 months, as key variables of interest were not captured for those who had not used ePHRs. Additionally, some variables in this study were measured using single-item scales, which may lack the robustness of multi-item measures. However, single-item measures are considered acceptable when the constructs are straightforward and unambiguous [52]. We acknowledge that single-item measures inherently do not allow for traditional assessments of internal consistency reliability, which presents a limitation in this study. This limitation necessitates cautious interpretation of the findings. Future research should aim to incorporate multi-item scales where feasible and conduct reliability analyses to ensure the stability and consistency of these measures.

Second, the reliance on self-reported data introduces the possibility of CMV, which could inflate relationships between variables [59]. To address this concern, the study applied the marker variable technique [53], which confirmed that CMV was not a significant issue. Third, the cross-sectional design of the study limits the ability to establish causal relationships among the variables. Longitudinal research is needed to explore how factors such as self-efficacy and performance expectancy influence the use of ePHRs over time. Fourth, although the use of survey weights enhances the representativeness of the sample for the United States, the relatively small sample size remains a limitation that future research with a more targeted approach could address. Additionally, we did not control for factors such as income, digital literacy, and prior technology experience, as these variables were not accessible in the dataset. The absence of these controls may limit the completeness of the study. Future research should incorporate these variables to provide a more comprehensive understanding of the frequency of HIT use among older adults. Finally, the findings are specific to the US context and may not fully generalize to other populations or health care systems. Despite these limitations, the study offers valuable insights and lays important groundwork for future research.

Conclusions

This study emphasizes the important role of ePHRs in empowering older adults to manage their health and maintain independence. By utilizing survey weights, the findings can be generalized to the broader US population, making them particularly relevant as the aging demographic continues to grow. Unlike previous research, this study reveals a positive relationship between age and the level of ePHR use, challenging the common assumption that older adults are reluctant to adopt health technologies.

Key factors driving ePHR usage are performance expectancy, effort expectancy, self-efficacy, and issue involvement, all of which offer actionable pathways for increasing the frequency of use among older adults. Raising awareness of the practical benefits and ease of use of ePHRs can encourage more frequent engagement. Additionally, addressing usability concerns and emphasizing the relevance of ePHRs to individual health needs can further promote their adoption. Frequent use of ePHRs empowers older adults to manage chronic conditions, access vital health information, and make informed decisions, ultimately enhancing their quality of life.

As health care systems strive to meet the challenges posed by an aging population, integrating ePHRs into routine care and ensuring equitable access can provide both social and economic benefits. This study lays the groundwork for targeted interventions aimed at bridging the digital divide and fostering a more inclusive, health-empowered society for older adults.

Acknowledgments

The authors thank the reviewers and editor for constructive feedback.

XSL•FO
RenderX

Disclaimer

Generative artificial intelligence tools such as Grammarly and ChatGPT were used to assist with language editing and clarity in certain sections of the manuscript. The article is an original work by the authors, who take full responsibility for its content.

Data Availability

The dataset used in this study is publicly available from the National Cancer Institute's Health Information National Trends Survey (HINTS) portal [69].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Operationalization of constructs.
[DOCX File, 22 KB - aging_v8i1e71460_app1.docx]

References

1. What are the advantages of electronic health records? Office of the National Coordinator for Health Information Technology. 2024. URL: https://www.healthit.gov/faq/what-are-advantages-electronic-health-records [accessed 2025-07-13]
2. Dutta B. What motivates people to adopt personal health record? Understanding the effect of innovation and context-specific threat on personal health record adoption intention. International Journal of Human–Computer Interaction 2024 Sep;40(17):4831-4844. [doi: 10.1080/10447318.2023.2223821]
3. Chaudhry B, Wang J, Wu S, et al. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. Ann Intern Med 2006 May 16;144(10):742-752. [doi: 10.7326/0003-4819-144-10-200605160-00125] [Medline: 16702590]
4. Davaridolatabadi N, Sadoughi F, Ahmadi M. Health information system to improve elderly health: study of health information system in us to setup standard for developing countries. Medical Gazette/Medicinski Glasnik 2013;18:50. [doi: 10.5937/medgla1350018D]
5. Weber V, White A, McIlvried R. An electronic medical record (EMR)-based intervention to reduce polypharmacy and falls in an ambulatory rural elderly population. J Gen Intern Med 2008 Apr;23(4):399-404. [doi: 10.1007/s11606-007-0482-z] [Medline: 18373136]
6. Perotti L, Stamm O, Dietrich M, Buchem I, Müller-Werdan U. The usability and user experience of an interactive e-learning platform to empower older adults when using electronic personal health records: an online intervention study. Universal Access in the Information Society 2025;24:1193-1208.
7. Luo RZ, Lu Q, Sun Y, Sun H, Liu YH, Zhao Y. Investigation on risk factors of chronic diseases among community residents: a study based on health management systems supported by mobile phones. Nurs Open 2023 Aug;10(8):5117-5128. [doi: 10.1002/nop2.1747] [Medline: 37071702]
8. van der Vaart R, Drossaert CHC, Taal E, Drossaers-Bakker KW, Vonkeman HE, van de Laar M. Impact of patient-accessible electronic medical records in rheumatology: use, satisfaction and effects on empowerment among patients. BMC Musculoskelet Disord 2014 Mar 26;15(1):102. [doi: 10.1186/1471-2474-15-102] [Medline: 24673997]
9. HITECH Act enforcement interim final rule. US Department of Health and Human Services. 2009. URL: https://www.hhs.gov/hipaa/for-professionals/special-topics/hitech-act-enforcement-interim-final-rule/index.html [accessed 2025-07-13]
10. 21st Century Cures Act. US Food and Drug Administration. 2016. URL: https://www.fda.gov/regulatory-information/selected-amendments-fdc-act/21st-century-cures-act [accessed 2025-07-13]
11. An S, Cheung CF, Lo YT. Improving older adults' technology adoption on mobile map: a gamification approach. Int J Hum Comput Interact 2024:1-19. [doi: 10.1080/10447318.2024.2426914]
12. Share of those 65 and older who are tech users has grown in the past decade. Pew Research Center. 2022 Jan 13. URL: https://www.pewresearch.org/short-reads/2022/01/13/share-of-those-65-and-older-who-are-tech-users-has-grown-in-the-past-decade/ [accessed 2025-07-13]
13. Trivedi N, Patel V, Johnson C, Chou WYS. Barriers to accessing online medical records in the United States. Am J Manag Care 2021 Jan;27(1):33-40. [doi: 10.37765/ajmc.2021.88575] [Medline: 33471460]
14. Or CKL, Karsh BT. A systematic review of patient acceptance of consumer health information technology. J Am Med Inform Assoc 2009;16(4):550-560. [doi: 10.1197/jamia.M2888] [Medline: 19390112]
15. Taha J, Czaja SJ, Sharit J, Morrow DG. Factors affecting usage of a personal health record (PHR) to manage health. Psychol Aging 2013 Dec;28(4):1124-1139. [doi: 10.1037/a0033911] [Medline: 24364414]
16. Hourcade JP, Chrischilles EA, Gryzlak BM, et al. Design lessons for older adult personal health records software from older adults. Presented at: Universal Access in Human-Computer Interaction Users Diversity: 6th International Conference, UAHCI 2011, Held as Part of HCI International 2011; Jul 9-14, 2011; Orlando, FL p. 176-185. [doi: 10.1007/978-3-642-21663-3_19]

XSL•FO
RenderX

17. Czaja SJ, Lee CC. The impact of aging on access to technology. Univ Access Inf Soc 2007 Mar 19;5(4):341-349. [doi: 10.1007/s10209-006-0060-x]
18. Peine A, Neven L. From intervention to co-constitution: new directions in theorizing about aging and technology. Gerontologist 2019 Jan 9;59(1):15-21. [doi: 10.1093/geront/gny050] [Medline: 29850812]
19. Gallistl V, Rohner R, Seifert A, Wanka A. Configuring the older non-user: between research, policy and practice of digital exclusion. SI 2020;8(2):233-243. [doi: 10.17645/si.v8i2.2607]
20. Mannheim I, Wouters EJM, van Boekel LC, van Zaalen Y. Attitudes of health care professionals toward older adults' abilities to use digital technology: questionnaire study. J Med Internet Res 2021 Apr 21;23(4):e26232. [doi: 10.2196/26232] [Medline: 33881408]
21. Spann A, Stewart E. Barriers and facilitators of older people's mHealth usage: a qualitative review of older people's views. Human Technology 2018;14(3):264-296. [doi: 10.17011/ht/urn.201811224834]
22. Noblin AM, Wan TTH, Fottler M. The impact of health literacy on a patient's decision to adopt a personal health record. Perspect Health Inf Manag 2012;9(Fall):1-13. [Medline: 23209454]
23. Or CKL, Karsh BT, Severtson DJ, Burke LJ, Brown RL, Brennan PF. Factors affecting home care patients' acceptance of a web-based interactive self-management technology. J Am Med Inform Assoc 2011;18(1):51-59. [doi: 10.1136/jamia.2010.007336] [Medline: 21131605]
24. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
25. McCloskey D. Evaluating electronic commerce acceptance with the Technology Acceptance Model. JCIS 2004;44(2):49-57. [doi: 10.1080/08874417.2004.11647566]
26. Lederer AL, Maupin DJ, Sena MP, Zhuang Y. The technology acceptance model and the World Wide Web. Decis Support Syst 2000 Oct;29(3):269-282. [doi: 10.1016/S0167-9236(00)00076-2]
27. Martín-García AV, Redolat R, Pinazo-Hernandis S. Factors influencing intention to technological use in older adults. The TAM model aplication. Res Aging 2022;44(7-8):573-588. [doi: 10.1177/01640275211063797] [Medline: 34962846]
28. Czaja SJ, Charness N, Fisk AD, et al. Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). Psychol Aging 2006 Jun;21(2):333-352. [doi: 10.1037/0882-7974.21.2.333] [Medline: 16768579]
29. Venkatesh V, Morris M, Davis G, Davis F. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
30. Rosenstock IM. The Health Belief Model and preventive health behavior. Health Educ Monogr 1974 Dec;2(4):354-386. [doi: 10.1177/109019817400200405]
31. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172. [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
32. Venkatesh V, Thong JY, Xu X. Consumer acceptance and use of information technology: extending the Unified Theory of Acceptance and Use of Technology. MIS Q 2012;36(1):157. [doi: 10.2307/41410412]
33. Champion VL, Skinner CS. The health belief model. In: Health Behavior and Health Education: Theory, Research, and Practice: Jossey-Bass; 2008, Vol. 4:45-65.
34. Khadjesari Z, Houghton J, Brown TJ, Jopling H, Stevenson F, Lynch J. Contextual factors that impact the implementation of patient portals with a focus on older people in acute care hospitals: scoping review. JMIR Aging 2023;6:e31812. [doi: 10.2196/31812] [Medline: 36735321]
35. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: an integrative systematic review. BMC Public Health 2024;24. [doi: 10.1186/s12889-024-18036-5]
36. Heart T, Kalderon E. Older adults: are they ready to adopt health-related ICT? Int J Med Inform 2013 Nov;82(11):e209-e231. [doi: 10.1016/j.ijmedinf.2011.03.002] [Medline: 21481631]
37. Troen BR. The biology of aging. Mt Sinai J Med 2003 Jan;70(1):3-22. [Medline: 12516005]
38. Smith SG, O'Conor R, Aitken W, Curtis LM, Wolf MS, Goel MS. Disparities in registration and use of an online patient portal among older adults: findings from the LitCog cohort. J Am Med Inform Assoc 2015 Jul;22(4):888-895. [doi: 10.1093/jamia/ocv025] [Medline: 25914099]
39. Dendere R, Slade C, Burton-Jones A, Sullivan C, Staib A, Janda M. Patient portals facilitating engagement with inpatient electronic medical records: a systematic review. J Med Internet Res 2019 Apr 11;21(4):e12779. [doi: 10.2196/12779] [Medline: 30973347]
40. Niazkhani Z, Toni E, Cheshmekaboodi M, Georgiou A, Pirnejad H. Barriers to patient, provider, and caregiver adoption and use of electronic personal health records in chronic care: a systematic review. BMC Med Inform Decis Mak 2020 Jul 8;20(1):153. [doi: 10.1186/s12911-020-01159-1] [Medline: 32641128]
41. Scott Kruse C, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: a systematic review. J Telemed Telecare 2018 Jan;24(1):4-12. [doi: 10.1177/1357633X16674087] [Medline: 29320966]
42. van Deursen AJ, van Dijk JA. The digital divide shifts to differences in usage. New Media & Society 2014 May;16(3):507-526. [doi: 10.1177/1461444813487959]

XSL•FO
RenderX

43. Abdelhamid M, Gaia J, Sanders GL. Putting the focus back on the patient: how privacy concerns affect personal health information sharing intentions. J Med Internet Res 2017 Sep 13;19(9):e169. [doi: 10.2196/jmir.6877] [Medline: 28903895]
44. Agrawal L, Ndabu T, Mulgund P, Sharman R. Factors affecting the extent of patients' electronic medical record use: an empirical study focusing on system and patient characteristics. J Med Internet Res 2021 Oct 28;23(10):e30637. [doi: 10.2196/30637] [Medline: 34709181]
45. Angst C, Agarwal R. Adoption of electronic health records in the presence of privacy concerns: the elaboration likelihood model and individual persuasion. MIS Q 2009;33(2):339. [doi: 10.2307/20650295]
46. Ross SE, Todd J, Moore LA, Beaty BL, Wittevrongel L, Lin CT. Expectations of patients and physicians regarding patient-accessible medical records. J Med Internet Res 2005 May 24;7(2):e13. [doi: 10.2196/jmir.7.2.e13] [Medline: 15914460]
47. de Veer AJE, Peeters JM, Brabers AEM, Schellevis FG, Rademakers J, Francke AL. Determinants of the intention to use e-Health by community dwelling older people. BMC Health Serv Res 2015 Mar 15;15:103. [doi: 10.1186/s12913-015-0765-8] [Medline: 25889884]
48. Tavares J, Oliveira T. Electronic health record portal adoption: a cross country analysis. BMC Med Inform Decis Mak 2017 Jul 5;17(1):97. [doi: 10.1186/s12911-017-0482-9] [Medline: 28679423]
49. Mukhopadhyay S, Basak R, Carpenter D, Reithel B. Patient use of online medical records: an application of technology acceptance framework. Information & Computer Security 2019;28(1):97-115 [FREE Full text] [doi: 10.1108/ICS-07-2019-0076]
50. Goel MS, Brown TL, Williams A, Cooper AJ, Hasnain-Wynia R, Baker DW. Patient reported barriers to enrolling in a patient portal. J Am Med Inform Assoc 2011 Dec;18 Suppl 1(Suppl 1):i8-12. [doi: 10.1136/amiajnl-2011-000473] [Medline: 22071530]
51. Chen K, Chan AHS. A review of technology acceptance by older adults. Gerontechnology 2011;10(1). [doi: 10.4017/gt.2011.10.01.006.00]
52. Wanous JP, Reichers AE, Hudy MJ. Overall job satisfaction: how good are single-item measures? J Appl Psychol 1997 Apr;82(2):247-252. [doi: 10.1037/0021-9010.82.2.247] [Medline: 9109282]
53. Santibáñez-Andrade G, Castillo-Argüero S, Vega-Peña EV, Lindig-Cisneros R, Zavala-Hurtado JA. Structural equation modeling as a tool to develop conservation strategies using environmental indicators: the case of the forests of the Magdalena river basin in Mexico City. Ecol Indic 2015 Jul;54:124-136. [doi: 10.1016/j.ecolind.2015.02.022]
54. MacKinnon D, Cheong J, Pirlott A. Statistical Mediation Analysis: American Psychological Association; 2012.
55. DiStefano C, Morgan GB. A comparison of diagonal weighted least squares robust estimation techniques for ordinal data. Struct Equ Modeling 2014 Jul 3;21(3):425-438. [doi: 10.1080/10705511.2014.915373]
56. Nguyen TQ, Webb-Vargas Y, Koning IM, Stuart EA. Causal mediation analysis with a binary outcome and multiple continuous or ordinal mediators: simulations and application to an alcohol intervention. Struct Equ Modeling 2016;23(3):368-383. [doi: 10.1080/10705511.2015.1062730] [Medline: 27158217]
57. Moser RP, Naveed S, Cantor D, et al. Integrative analytic methods using population-level cross-sectional data. National Cancer Institute. 2013. URL: https://hints.cancer.gov/docs/hints_ida_report.pdf [accessed 2025-07-13]
58. Hair J, Anderson R, Tatham R, Black W. Multivariate Data with Readings: Prentice Hall; 1995.
59. Podsakoff PM, MacKenzie SB, Lee JY, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J Appl Psychol 2003 Oct;88(5):879-903. [doi: 10.1037/0021-9010.88.5.879] [Medline: 14516251]
60. Lindell MK, Whitney DJ. Accounting for common method variance in cross-sectional research designs. J Appl Psychol 2001 Feb;86(1):114-121. [doi: 10.1037/0021-9010.86.1.114] [Medline: 11302223]
61. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
62. Sarkar U, Karter AJ, Liu JY, et al. The literacy divide: health literacy and the use of an internet-based patient portal in an integrated health system-results from the diabetes study of northern California (DISTANCE). J Health Commun 2010;15 Suppl 2(Suppl 2):183-196. [doi: 10.1080/10810730.2010.499988] [Medline: 20845203]
63. Federal health IT strategic plan. Office of the National Coordinator for Health Information Technology (ONC). 2021. URL: https://www.healthit.gov [accessed 2025-07-13]
64. Charness N, Boot WR. A grand challenge for psychology: reducing the age-related digital divide. Curr Dir Psychol Sci 2022;31(2):187-193. [doi: 10.1177/09637214211068144] [Medline: 35754678]
65. Healthy people 2030. US Department of Health and Human Services. 2020. URL: https://health.gov/healthypeople [accessed 2025-07-13]
66. Czaja SJ, Boot WR, Charness N, Rogers WA. Designing for Older Adults: Principles and Creative Human Factors Approaches, 3rd edition: CRC Press; 2019. [doi: 10.1201/b22189]
67. Krist AH, Tong ST, Aycock RA, Longo DR. Engaging patients in decision-making and behavior change to promote prevention. Stud Health Technol Inform 2017;240:284-302. [doi: 10.3233/ISU-170826] [Medline: 28972524]
68. Patient empowerment strategies. Agency for Healthcare Research and Quality (AHRQ). 2019. URL: https://digital.ahrq.gov/program-overview/research-stories/engaging-empowering-patients [accessed 2025-07-13]

XSL•FO
RenderX

69. How to use HINTS data. Health Information National Trends Survey (HINTS). URL: https://hints.cancer.gov/data/download-data.aspx#H5C3 [accessed 2025-07-13]

Abbreviations

CMV: common method variance
e-commerce: electronic commerce
EMR: electronic medical record
ePHR: electronic personal health record
HINTS: Health Information National Trends Survey
HIT: health information technology
PTAM: Patient Technology Acceptance Model
SEM: structural equation model
TAM: Technology Acceptance Model
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by H Köttl; submitted 18.01.25; peer-reviewed by G Mccormack, P Setthasuravich, S Choemprayong; revised version received 23.05.25; accepted 26.05.25; published 28.07.25.

Please cite as:
Agrawal L, DaSouza RO, Mulgund P, Chaudhary P
Frequency of Electronic Personal Health Record Use in US Older Adults: Cross-Sectional Study of a National Survey
JMIR Aging 2025;8:e71460
URL: https://aging.jmir.org/2025/1/e71460
doi:10.2196/71460

© Lavlin Agrawal, Richelle Oakley DaSouza, Pavankumar Mulgund, Pankaj Chaudhary. Originally published in JMIR Aging (https://aging.jmir.org), 28.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Ethics and Equity Challenges in Telerehabilitation for Older Adults: Rapid Review

Mirella Veras[1,2,3], PT, PhD; Louis-Pierre Auger[2,4], PhD; Jennifer Sigouin[2], PhD; Nahid Gheidari[5], PhD; Michelle LA Nelson[6,7], PhD; William C Miller[8,9], PhD; Anne Hudon[10], PhD; Dahlia Kairy[2,10], PhD

[1]Department of Physical Therapy, University of Manitoba, R115-771 McDermot Avenue, Winnipeg, MB, Canada
[10]Physiotherapy Program, School of Rehabilitation, Faculty of Medicine, University of Montreal, Montreal, QC, Canada
[2]The Centre for Interdisciplinary Research in Rehabilitation of Greater Montreal (CRIR), Montreal, QC, Canada
[3]Centre on Aging, University of Manitoba, Winnipeg, MB, Canada
[4]School of Physical and Occupational Therapy, Faculty of Medicine and Health Sciences, McGill University, Montreal, QC, Canada
[5]Clinical Research Unit, Montreal Neurological Institute/Hospital, Montreal, QC, Canada
[6]Lunenfeld-Tanenbaum Research Institute, Sinai Health, Toronto, ON, Canada
[7]Dalla Lana School of Public Health, University of Toronto, Toronto, Canada
[8]Occupational Science & Occupational Therapy, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada
[9]Centre for Aging SMART (Solutions for Mobility, Activity, Rehabilitation and Technology), Rehabilitation Research Program, Vancouver, BC, Canada

Corresponding Author:
Mirella Veras, PT, PhD
Department of Physical Therapy, University of Manitoba, R115-771 McDermot Avenue, Winnipeg, MB, Canada

Abstract

Background: The integration of technology in rehabilitation is transforming health care delivery for older adults, especially through telerehabilitation, which addresses barriers to in-person care.

Objective: This rapid review explores the ethical and equity concerns associated with telerehabilitation for older adults, focusing on challenges such as internet access, technology adoption, and digital literacy.

Methods: Conducted according to Cochrane Rapid Review guidelines, this review used the Metaverse Equitable Rehabilitation Therapy framework, focusing on equity and ethics. Studies included telerehabilitation services for adults aged 55 years and older, published between 2010 and 2023. Screening was conducted independently by 2 researchers using Rayyan (Qatar Computing Research Institute, Hamad Bin Khalifa University), with full-text review by additional team members. Searches were performed in Medline and CINAHL

Results: From 323 papers retrieved, 49 studies met the inclusion criteria. The included studies were published between 2013 and 2023. Disparities in socioeconomic status, geographic location, and racial and ethnic backgrounds were found to impact telerehabilitation use. Additionally, ethical concerns around privacy, security, and autonomy were often inadequately addressed.

Conclusions: This review emphasizes the need for culturally appropriate, accessible, and inclusive telerehabilitation services that integrate ethical and equity considerations into their design and delivery.

(JMIR Aging 2025;8:e69660) doi:10.2196/69660

KEYWORDS

aging; older adults; equity; ethics; telerehabilitation; telehealth; review; rehabilitation; digital health; geriatric; virtual care

Introduction

The use of technology in rehabilitation is increasing worldwide, including its application in delivering services to the older adults population [1]. This is transforming how health care services are delivered to the aging population [2]. Telerehabilitation, which involves the evaluation and treatment of patients through technology, has emerged as an attractive option for older adults who may have multiple comorbidities [1,3]. Before the COVID-19 pandemic, numerous adults faced challenges accessing rehabilitation treatments due to diverse clinical limitations and geographic circumstances. Telerehabilitation has assumed greater importance in addressing the needs of older adults, as the COVID-19 pandemic exacerbated difficulties by further restricting face-to-face interactions with implementation of societal measures aimed at lowering disease prevalence [4].

When older adults experience complex medical needs, whether due to multiple chronic conditions or acute illnesses, it can lead to increased dependence on others and diminish the quality of life for both the individuals and their caregivers [5]. Health care

XSL•FO
RenderX

technologies can help overcome these challenges and enhance the well-being of the aging population by improving access to rehabilitation [5]. In this context, telerehabilitation provides an alternative platform for delivering health education and care to clients in clinical, community, or home care settings [6]. However, challenges in delivering telerehabilitation to older adults have been reported in the literature and include issues with internet access, adoption of technology, self-efficacy, experience with technology, frequency of usage, and reliance on guidance and digital literacy [7,8]. Other barriers reported in a systematic review about the use of telehealth use among older adults were trust of internet information, reading small fonts, information overload, clicking small icons, interacting with scroll bars, lack of understanding of app capabilities, fear of others overhearing, security, need for support to use, cost, privacy concerns, among others [8].

Despite the numerous benefits of telerehabilitation for older adults, this is recognized as an extremely heterogeneous population; therefore, there is a pressing need to investigate and understand the ethical and equity concerns associated with these services. No reviews to date have thoroughly explored these aspects, which is critical given the reported challenges. Knowing about these problems is important to help address the effectiveness and accessibility of telerehabilitation. Ethical concerns may arise regarding the confidentiality of patient information, while equity issues could stem from disparities in access to technology and the internet among different socioeconomic groups, for example. Addressing these concerns is essential to ensure that telerehabilitation services are delivered fairly and effectively, promoting better health outcomes and quality of life for all older adults. This review aims to examine these concerns and explore telerehabilitation for older adults over the past 10 years.

Methods

Study Design

Rapid reviews are a form of synthesis that accelerates the traditional systematic review process, facilitating the dissemination of literature in a resource-efficient manner, particularly relevant in the context of rapidly advancing technology research [9]. This rapid review aimed to explore and summarize the current literature on the equity and ethical concerns associated with delivering telerehabilitation services, following the Cochrane Rapid Review guidelines [9]. To guide the data extraction and analysis of this review, the Metaverse Equitable Rehabilitation Therapy (MERTH) framework [10] was applied, adapted to telerehabilitation, with a focus on equity (accessibility, inclusivity, diversity, fairness, and cultural relevance), health services integration (responsiveness, continuity of care, and autonomy to participate in health-related decisions), interoperability, and humanization (communication, person-centered care, and empowerment) [10]. For this article, we are focusing on equity (accessibility, inclusivity, diversity, fairness, cultural relevance) and empowerment.

We included studies that focused on telerehabilitation services for both older adults and middle-aged individuals aged 55 years and above. The inclusion criteria focused on reviews (systematic reviews, scoping reviews, meta-analyses, and narrative reviews with descriptions of included studies) involving older adults aged 55 years and older, which includes the older working-age population (55 to 64 y) [11]. This age range was selected to include individuals who may benefit from telerehabilitation services, considering their potential health care needs and technological engagement during this phase of life. These studies were selected if published between 2010 and 2023. Editorials, conference abstracts, and papers without full text available were excluded. Two researchers (MV and NG) independently completed the title and abstract screening using Rayyan [12], an online app that accelerates initial screening of abstracts and titles for managing reviews. One reviewer (MV) performed the full-text screening, and any doubts regarding inclusion were discussed with MV, DK, and JS. DK, JS, NG, and MV developed a data extraction tool and refined it throughout the extraction process. For added rigor, a sample of the data, such as 10% (13/125) of the reviews at the abstract phase, was independently reviewed by another team member (DK or JS).

To conduct the search, we consulted a specialized health sciences library. We tailored a search strategy for 2 databases: Medline and CINAHL.

Detailed search strategies can be found in Multimedia Appendix 1.

Ethical Considerations

This study is a review of existing literature and does not involve any primary data collection or direct interaction with human participants. Therefore, ethical approval and consent for publication are not required. All sources cited have been appropriately referenced per academic standards.

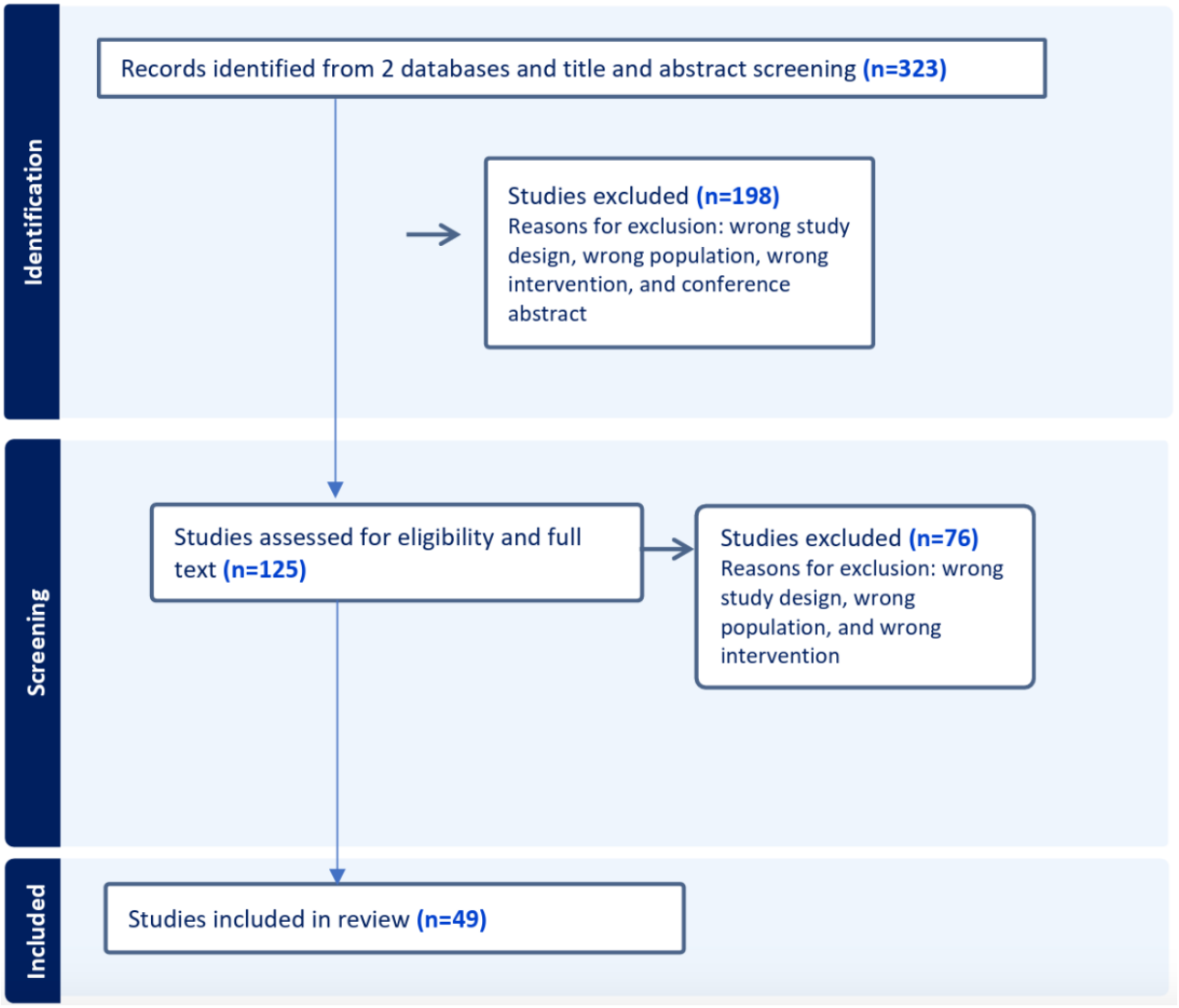
Results

Overview

Figure 1 provides an overview of the review process. Initially, 323 papers were retrieved through the search. After title and abstract screening, followed by full-text review, 49 studies met the criteria for extraction and final evaluation.

XSL•FO
RenderX

Figure 1. Flow diagram of study selection.



Publications Characteristics

The included review studies were published between 2013 and 2023. Approximately 39% (n=19) of the studies were published in 2021. The included studies predominantly consisted of systematic reviews (22 studies), systematic reviews and meta-analyses (11 studies) followed by scoping reviews (8 studies; Table S1 in Multimedia Appendix 2 [6,13-60]).

Morbidities

Cardiorespiratory, neurological, and musculoskeletal conditions are the primary morbidities present in the studies identified in this review. Other key categories include cognitive impairment and mental health, cancer, and diabetes. Less frequently studied morbidities include vision and hearing problems, balance disorders, risk for falls, and pelvic floor disorders. The distribution of included studies (n=49) across morbidity categories revealed that the most commonly addressed condition was cardiorespiratory, featured in 33/49 (67%) of the studies. Neurological and musculoskeletal conditions followed this, each reported in 17/49 (35%) of the studies. Studies categorized under "other" comprised 15/49 (31%). Cognitive impairment and mental health, and cancer were each represented in 9/49 (19%) of the studies. Diabetes was addressed in 6/49 (13%), while vision and hearing problems, balance disorders and risk for falls, and pelvic floor disorders were the least represented, each occurring in 2/49 (5%) of the included studies. These findings indicate a predominant research focus on cardiorespiratory and neuromusculoskeletal morbidities.

It is important to note that individual morbidities can be counted in more than 1 study within these reviews and that each study could include participants with several morbidities. This repetition highlights the prevalence and importance of specific categories of health issues within the reviewed literature.

Type of Technologies Used in Telerehabilitation for Older Adults

Within the included reviews, the categorization of communication technologies in telerehabilitation included several key areas, each highlighted by different examples and applications. Communication technologies included eHealth platforms such as telehealth, telemedicine, and telerehabilitation, which used tools such as videoconferencing (Skype [Microsoft Corp] and Zoom [Zoom Communications, Inc]), telephones, text messaging, emails, and web-based platforms. Telemonitoring and remote monitoring were used through health

XSL•FO
RenderX

monitoring systems, wearable sensors, and mobile health apps for tracking health metrics. Virtual and augmented reality were applied in digital rehabilitation systems and exergames such as Wii Fit (Nintendo Entertainment Analysis & Development) and Kinect (Microsoft). Wearable technology included sensors, smartwatches, and fitness trackers. Education and support tools provided digital health education via e-learning apps and digital pain coaches. Digital and assistive technologies enhanced patient engagement and supported independent living through interactive internet platforms and emergency assistance systems. Finally, home-based rehabilitation and exercise facilitated remote rehabilitation exercises with personalized programs and activity trackers (Table 1)

Based on the studies included in this rapid review, the most frequent categories of technologies used are communication technologies, including telemonitoring and remote monitoring, which appear in 73% of the included studies. In contrast, the least used categories are virtual and augmented reality and interactive and assistive technologies, each appearing in 12% of the studies. Specifically, virtual and augmented reality were found in 6 studies, which corresponds to approximately 12% of the total studies. The total percentage exceeds 100% due to the overlap of studies across multiple categories (Table 1).

XSL•FO
RenderX

Table . Study characteristics: technologies used and examples.

Categorization	Examples	References
Communication technologies	• eHealth platforms: telehealth interventions, telemedicine, and telerehabilitation platforms • Telehealth devices: remote supervision tools, home-health monitoring systems, and teleconsultation platforms Videoconferencing: Skype (Microsoft Corp), Zoom (Zoom Communications, Inc), FaceTime (Apple Inc), Adobe Connect, and Microsoft Teams • Telephone and mobile phones: standard calls, smartphones, and telephone support • Text messaging and emails: SMS and email • Web-based platforms: online multimedia, discussion boards, blogs, and web-based health management portals	[6,13-42,60]
Telemonitoring and remote monitoring	• Health monitoring systems: heart health monitoring system, wearable sensors, pulse oximeters, and blood pressure monitors • Telemonitoring devices: ECG[a], heart monitors, Bluetooth-powered sensors, and wireless transmitters • Mobile health apps: apps for tracking health metrics and remote wound care applications	[6,13,18,21-23,26,28,30,35,37,39,43-50]
Virtual and augmented reality	• Virtual rehabilitation systems: virtual reality–based rehabilitation and interactive virtual reality exercises • Exergames: Wii Fit (Nintendo Entertainment Analysis & Development), Kinect (Microsoft), VRc[b] video game dancing, and interactive gaming systems	[6,19,23,51,52,61]
Wearable technology	• Sensors and monitors: inertial sensors, accelerometers, gyroscopes, and pedometers • Wearable devices: smartwatches, fitness trackers, and wearable pressure	[6,19,21,41,43-46,53,59]
Education and support tools	• Provides health education and support through digital tools. Examples: e-learning apps, digital pain coaches.	[6,14,15,17,21,23,26,30,36,38,45,46,50,54-56]
Interactive and assistive technologies	• Enhances patient engagement and supports independent living. Examples: interactive internet platforms, emergency assistance	[17,22,39,50,53]
Home-based rehabilitation and exercise	• Facilitates rehabilitation exercises at home with remote guidance. Examples: personalized exercise programs, activity trackers.	[6,15,21,22,24,35,43-45,47,48,50,52,53,55]

[a]ECG: electrocardiogram.

[b]VR: virtual reality.

Equity

Sex or Gender

Twenty-nine reviews included at least 1 study that reported findings on gender (Table S1 in Multimedia Appendix 2). The reviews that reported sex or gender in their results did not provide a full analysis based on these variables. However, 2 studies included some comments related to their gender or sex results.

One review reported that the findings regarding gender and the use of eHealth are mixed. “Out of 22 articles, only five suggest that gender plays a role in influencing eHealth usage, while eight studies found no connection between gender and eHealth use.” The studies did not define gender, underscoring the importance of clearly specifying sex, gender, and sexual orientation within the included population. Additionally, 3 studies highlight that women tend to be more engaged with and satisfied by eHealth apps, using them more frequently than men

XSL•FO
RenderX

[13]. In 1 review, the authors mentioned a gender bias, with more male participants than female participants [6]. In another study, the authors noted that most of the studies included more women than men. Four studies included only women. One study explained this decision as "to avoid the influence of gender differences on the risk of falling." Two studies excluded the few male participants, and another study did not provide a reason for the exclusive participation of women. Additionally, 1 study did not describe the age range or gender distribution of the participants [43].

Ethnicity

Racial differences in telehealth care use were observed, with cultural minority groups. For instance, 1 study found no significant effect of telehealth interventions on older Hispanic patients with chronic heart failure [14]. Overall, most studies that reported ethnicity included predominantly White participants. The term "Caucasian" is used as a more formal and historically established term to describe people of European descent. However, it can be viewed as outdated or less inclusive compared to simply using "White" or more specific ethnic identifiers (Table S1 in Multimedia Appendix 2).

Geographic Location

Older adults with chronic diseases residing in rural areas are less likely to use eHealth services compared to their urban counterparts, primarily due to limited access and lower socioeconomic status. Rural inhabitants face greater barriers to the effective implementation of eHealth, often related to insufficient infrastructure and resources. Contrarily, some regions, such as certain cities in South Korea with limited medical facilities, have shown a higher adoption of eHealth compared to more urbanized areas, highlighting the potential benefits of these technologies in underserved locations [13] (Table S1 in Multimedia Appendix 2).

Socioeconomic Status

One study highlighted that despite lower-income groups often facing barriers in accessing eHealth technologies due to economic constraints, their satisfaction levels with using these technologies were found to be comparable to those of higher-income groups once they had access (Reiners et al, 2019) [13]. This suggests that while income initially affects access to eHealth services, it may not impact the perceived benefits and satisfaction derived from using them once accessibility barriers are overcome.

Barriers

Digital barriers: the following summaries highlight the main challenges and associated studies for each category concerning the adoption of telerehabilitation among older adults:

Economic barriers: affordability of technology is an obstacle for older adults and those with lower educational backgrounds (Bertolazzi et al, 2024 [57] and Bhattarai and Phillips, 2017 [15]).

Usability challenges: difficulty in using mHealth technologies hinders adoption, exacerbated by lack of exposure (Johnson et al, 2021) [44].

Digital literacy disparities: lower digital literacy among racial or ethnic minorities, low-income groups, and people with sensory and manual dexterity deficits impact technology acceptance (Yi et al, 2021 [58]; Johnson et al, 2021 [44]; Jonker et al, 2020 [45]; Bertolazzi et al, 2024 [57]; Bostrom et al, 2020 [46]; Beckie, 2019 [16]; Gaspar and Lapão, 2021 [43]);

Support from family members: family encouragement aids acceptance and use of eHealth technologies (Jonker et al, 2020) [45].

Technology design and training: user-friendly designs and proper training enhance technology adoption (Johnson et al, 2021 [44] and Bhattarai and Phillips, 2017 [15]).

Inclusion and accessibility: addressing barriers such as income and education levels is crucial for equitable access (Bertolazzi et al, 2024) [57].

Health care provider involvement: clinician support and effective information flow facilitate technology adoption (Bhattarai and Phillips, 2017) [15].

Costs

The results related to digital health technology costs are structured into several key categories: costs, benefits, equity considerations, and policy implications. First, the section on cost details the financial implications of adopting digital health technologies, highlighting initial investments, ongoing maintenance costs, and cost analyses. Second, the benefits category outlines the positive impacts of these technologies on telerehabilitation, such as improved patient outcomes, enhanced efficiency, and reduced health care disparities. Third, equity considerations address how digital health technologies can potentially bridge gaps in health care access and outcomes among the older adult population. Lastly, the section on policy implications describes regulatory frameworks, privacy concerns, and recommendations for effective implementation strategies (Table S1 in Multimedia Appendix 2).

Overview of Costs in Digital Health Technology

Background of Costs in Digital Health Technology

Digital health technologies include a range of costs associated with implementation and operation. The following provides a summary of findings concerning cost considerations, supplemented with insights from the studies reviewed. For detailed examples and other comments, examples can be found in Table S1 in Multimedia Appendix 2.

General Cost Implications

Telehealth technologies are often highlighted as cost-effective strategies, particularly in reducing hospital and nursing home expenses.

Cost-Effective Strategies

For instance, telehealth has been proposed as cost-effective for delivering health education and promoting self-monitoring behaviors among older adults with chronic conditions.

XSL•FO
RenderX

Cost Types

Costs in digital health include implementation activities and operational expenses related to eHealth applications.

Specific Cost and Benefit Types

Patient Outcomes

Studies predominantly capture patient outcomes through quality-of-life considerations and physical health status indicators (summary of results of the included studies).

Resource Use

Emphasis is placed on reducing home care visits and hospital usage through remote monitoring, aimed at preventing unnecessary hospital admissions (results of included studies).

Equity and Access Considerations

Gender and Socioeconomic Variables

These variables influence the usability and adoption of digital health technologies, highlighting disparities in access and usage (comment on discussion of included studies).

Racial and Ethnic Differences

Studies note disparities in telehealth use among different racial and ethnic groups, suggesting cultural acceptance and access barriers.

Geographic and Socioeconomic Factors

Access to and effectiveness of telerehabilitation are influenced by factors such as geographic location and socioeconomic status, with rural and lower-income populations facing greater challenges.

Recommendations and Future Directions Related to Costs

Addressing Barriers

Recommendations include designing culture-friendly and accessible telehealth services to enhance engagement among diverse populations.

Policy Implications

Policy makers are urged to minimize cost barriers, provide subsidies for technology access, and promote equitable health technology adoption.

Ethics

Overview

Twenty-five studies reported some ethical aspects of the included studies. The following ethical aspects of the included studies can be categorized into safety, adverse events, privacy, empowerment, respect and knowledge of cultural diversity, and inclusion of ethics in research design.

Safety and Adverse Events

Safety concerns focus on ensuring that interventions do not cause harm to participants. Studies highlight that most interventions were safe, with no substantial adverse events reported (Huang et al, 2020 [17]; Malaguti et al, 2021 [18]; Johnson et al, 2021 [44]; Kraaijkamp et al, 2021 [19]; Bostrom et al, 2020 [46]; Gaspar and Lapão, 2021 [43]; Reeder et al, 2016 [20]; Solis-Navarro et al, 2022 [31]). The adverse events category pertains to the reporting and management of negative outcomes associated with interventions. Some studies mentioned the importance of feasibility testing to identify potential adverse events (Dennette et al, 2021 [21]; Kraaijkamp et al, 2021 [19]; Devi et al, 2015 [22]; Dequanter et al, 2021 [23]; Ambrens et al, 2022 [24]; Del Pino et al, 2022 [47]; and Ding et al, 2023 [25]).

Privacy

Concerns about maintaining the confidentiality and privacy of participant data were noted, emphasizing the need for secure data handling and adherence to privacy regulations (Yi et al, 2021 [58] and Jonker et al, 2020 [45]).

Respect and Knowledge of Cultural Diversity

This category highlights the importance of culturally appropriate interventions, especially for Indigenous older adults, and includes the use of culturally relevant content and rewards (Choukou et al, 2021 [26]).

Empowerment

In 1 study, the authors highlighted the empowerment divide, noting that despite having access to the necessary hardware and skills to use eHealth apps, some individuals did not use these opportunities because they did not feel personally empowered or believe they would benefit. While the reviewed papers primarily addressed economic and usability divides, the authors emphasized that future research should specifically focus on overcoming the empowerment divide (Reiners et al, 2019 [13]).

Inclusion of Ethics in the Design of the Interventions

Several studies highlighted in their discussion the importance of incorporating ethical considerations in the design and implementation of interventions. This includes respecting client preferences, obtaining informed consent, and addressing ethical and privacy concerns (Saito and Izawa, 2021 [27]; Sulz et al, 2021 [28], Su et al, 2020 [29]; Theodoros et al, 2019 [30]; and Tao et al, 2018 [42]).

Our review showed that 26 studies did not address the ethical aspects of their interventions. However, among those that did, several key ethical considerations emerged. Safety was a primary focus, with studies generally reporting no adverse events associated with the interventions (Huang et al, 2020 [17]; Malaguti et al, 2021 [18]; Johnson et al, 2021 [44]; Kraaijkamp et al, 2021 [19]; Bostrom et al, 2020 [46]; Gaspar and Lapão, 2021 [43]; Reeder et al, 2016 [20]; Solis-Navarro et al, 2022 [31]). Adverse event management was emphasized through feasibility testing (Dennette et al, 2021 [21]; Kraaijkamp et al, 2021 [19]; Devi et al, 2015 [22]; Dequanter et al, 2021 [23]; Ambrens et al, 2022 [24]; Del Pino et al, 2022 [47]; Ding et al, 2023 [25]).

Privacy concerns were frequently mentioned, highlighting the need for secure data handling and compliance with privacy regulations (Yi et al, 2021 [58]; Jonker et al, 2020 [45]). Empowerment was also critical, with studies stressing the importance of making participants feel capable and confident in using eHealth technologies (Reiners et al, 2019 [13]). Respect

XSL•FO
RenderX

and knowledge of cultural diversity were crucial, particularly for indigenous older adults, calling for culturally appropriate content and rewards (Choukou et al, 2021) [26]. Finally, the inclusion of ethics in research design was recommended, advocating for respect for client preferences, informed consent, and addressing ethical and privacy issues throughout the intervention process (Saito and Izawa, 2021 [27]; Sulz et al, 2021 [28]; Su Jing, 2020 [29]; Theodoros et al, 2019 [30]; and Tao et al, 2018 [42]). These findings emphasize the necessity of integrating comprehensive ethical considerations into digital health interventions to ensure safety, privacy, and respect for all participants.

Discussion

Principal Findings

The distribution of morbidities highlights areas where research is concentrated and indicates potential gaps in the literature that warrant further investigation, such as vision and hearing problems, balance disorders, risk of falls, and pelvic floor disorders. This distribution might also reflect conditions that health professionals are more comfortable managing. Therefore, there is a need to develop strategies to offer remote rehabilitation services for these conditions as well.

Older adults have a higher prevalence of chronic obstructive pulmonary disease, a trend that is expected to rise substantially in the coming decades due to aging populations and prolonged exposure to risk factors [62]. It has been estimated that between 55% and 98% of older adults aged 60 years and older have at least 2 chronic diseases, known as multimorbidity. Among these, cardiovascular diseases are the most common [63]. This aligns with our findings that cardiorespiratory issues were the most frequently studied morbidity type.

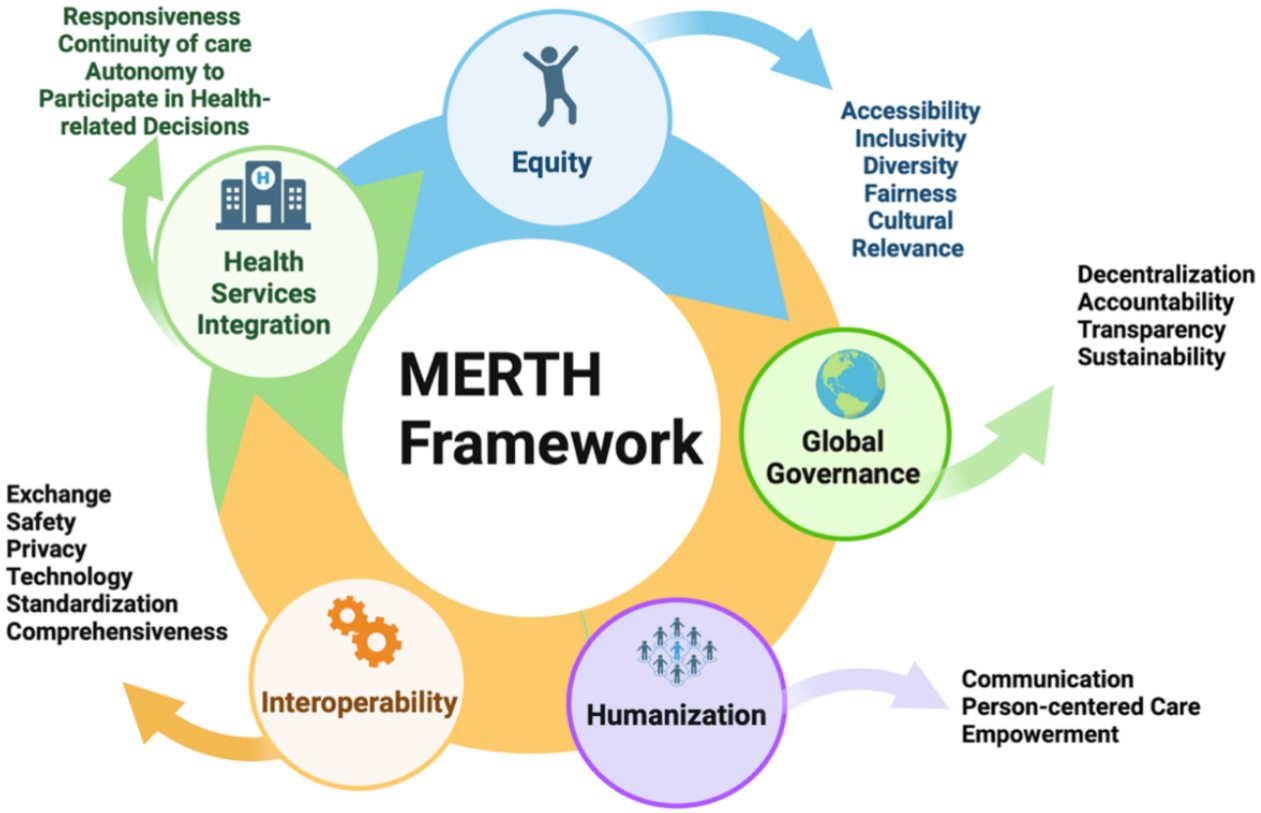
When considering the Health Services Integration of the MERTH framework (Figure 2), the need for continuity of care and responsiveness becomes even more evident [10]. To effectively address the complex health needs of older adults, particularly those with multiple chronic conditions, health systems must be able to provide integrated and coordinated care across various settings and providers. This not only involves traditional clinical care but also includes social support systems and digital resources, which are increasingly important for managing aging-related health issues [10].

Incorporating telerehabilitation could be a key part of achieving this integration. These digital platforms can help bridge gaps in health care access and ensure that patients receive consistent care, no matter their location or mobility [10]. They also offer an innovative way to enhance autonomy in decision-making, allowing patients to engage in their rehabilitation process more actively and with greater flexibility. This is especially important as older adults may have difficulties attending in-person sessions due to transportation issues or other physical limitations.

In the rapid review, communication technologies, telemonitoring, and remote monitoring were the most frequently included categories, totaling 38 reviews. In contrast, virtual and augmented reality, as well as interactive and assistive technologies, were less frequently cited, with only 6 reviews each. This distribution indicates a stronger focus on technologies that enable direct communication and continuous health monitoring, while immersive and interactive technologies, though valuable, are less commonly studied.

XSL•FO
RenderX

Figure 2. MERTH framework [10]. MERTH: Metaverse Equitable Rehabilitation Therapy.



This review highlights several key factors influencing the use and effectiveness of telerehabilitation services across different demographics. Racial disparities are evident, with cultural minority groups often showing lower acceptance of telehealth. For instance, 1 study found no significant impact of telehealth interventions on older Hispanic patients with chronic heart failure, indicating that cultural differences may influence outcomes. There is a substantial research gap in addressing racial differences in telerehabilitation, and the current literature primarily identifies barriers without proposing solutions to promote racial equity in telerehabilitation use. To ensure successful interventions, health care providers must design and deliver culturally friendly telerehabilitation services tailored to their target patients. Future research should focus on exploring these racial differences further and developing strategies to enhance the acceptance and effectiveness of telerehabilitation among diverse racial and ethnic groups.

Geographic location also plays a crucial role, with older adults in rural areas less likely to use telerehabilitation services due to limited access and lower socioeconomic status. These rural barriers often stem from insufficient infrastructure and resources, although some regions, such as certain South Korean cities with limited medical facilities, have shown higher eHealth adoption, demonstrating the potential of these technologies in underserved areas [13]. The overall trend emphasizes a persistent urban-rural divide in telerehabilitation usage, necessitating targeted efforts to improve access and infrastructure in rural areas to ensure equitable health care delivery.

Socioeconomic status further influences telerehabilitation use; lower-income groups face initial access barriers but report satisfaction levels comparable to higher-income groups once they gain access, indicating that economic constraints primarily affect initial accessibility rather than ongoing satisfaction and perceived benefits. The study on gender disparities in telehealth use among older adults in the United States during the COVID-19 pandemic provides relevant results for comparison with our review. It showed a significant increase in telehealth use among females during the pandemic, influenced by factors such as multimorbidity, tablet ownership, and learning new technologies. Males with similar factors also had higher odds of telehealth use [64]. Although some reviews in our study mentioned gender differences, they did not provide much detailed information. These findings emphasize the necessity for targeted interventions to enhance older adults' access to telerehabilitation services and mitigate digital disparities, aligning with our review's emphasis on the need for equitable access across diverse populations.

The reviews included in this analysis did not provide a detailed discussion on gender, sexual orientation, or the distinction between sex and gender. While the term "gender" may have been used in some instances, it was not consistently defined or explored in the context of the studies reviewed. In fact, many of the reviews primarily focused on sex-based differences, and in some cases, the term "gender" may have been used interchangeably with sex without clarifying the distinction. Additionally, there was no substantial examination of sexual orientation, and these factors were not adequately addressed or analyzed across the included studies. To provide a more comprehensive understanding of how these variables influence telerehabilitation outcomes, future research should ensure a clear and explicit discussion of gender, sex, and sexual

XSL•FO
RenderX

orientation, and their potential impact on the accessibility and effectiveness of telerehabilitation services.

Digital literacy plays a significant role in the adoption of telerehabilitation among older adults, with several key challenges highlighted in this review. Economic barriers, such as the affordability of technology, pose significant obstacles, particularly for older adults and those with lower educational backgrounds. Usability challenges, often exacerbated by a lack of exposure to health technologies, further hinder adoption. Studies show that enhancing internet use among older adults can significantly improve their access to health information and their ability to manage their health [65,66].

Disparities in digital literacy are particularly pronounced among racial or ethnic minorities, low-income groups, and individuals with sensory and manual dexterity deficits, affecting technology acceptance and use. Family support has been shown to aid in the acceptance and use of health technologies, while user-friendly design and proper training can significantly enhance adoption rates. A randomized controlled trial on technophilia (enthusiasm toward new technologies) among people with dementia or mild cognitive impairment and their caregivers demonstrated technophilia was lower among persons with disabilities influenced by factors such as younger age, male gender, higher education, better health, and depression. For caregivers, lower burden and better quality of life also contributed [67]. The sensory and manual dexterity deficits can significantly impact how individuals accept and use technology. These challenges highlight the importance of designing inclusive technology solutions that accommodate a diverse range of physical abilities, ensuring equitable access and usability for all users. Addressing all these barriers is crucial for ensuring equitable access to these technologies.

In terms of accountability, as per the MERTH framework, telerehabilitation services must maintain high standards of care despite these barriers. One area where accountability becomes particularly important is in ensuring that health care providers are properly trained and equipped to offer culturally competent and accessible digital care [10]. As there is a disparity in the acceptance and effectiveness of telerehabilitation across different demographics—such as rural versus urban populations, and racial or ethnic minorities—health care providers must be held accountable for creating inclusive services. This includes adapting technologies to be more accessible, as well as implementing feedback mechanisms to assess the effectiveness of interventions.

Cost considerations include initial investments and ongoing maintenance, but digital health technologies are generally seen as cost-effective, particularly in reducing hospital and nursing home expenses [14,68,69]. To maximize benefits and ensure equity, it is important to design culturally friendly and accessible telerehabilitation services, minimize cost barriers, and promote equitable technology adoption across diverse populations [14]. These findings support our review results on the need for targeted interventions to improve access to telerehabilitation services and reduce digital disparities among older adults.

Regarding ethical aspects, 26 studies did not address the ethical aspects of their interventions. This gap in addressing ethical considerations highlights a substantial shortcoming in the research, underscoring the need for more rigorous ethical evaluation in future studies. However, for the studies that did address ethics, several key considerations came to light. Safety was reported as a comment or result with no significant adverse events associated with the interventions. Adverse event management was emphasized through feasibility testing. A recent scoping review of risks, adverse effects, and mitigation strategies in delivering mental health services via telehealth, emphasizes the growing recognition of the importance of adverse event reporting in telehealth. However, it also highlights the lack of standardized protocols for this purpose [61]. Challenges persist in the telerehabilitation regarding the systematic collection and evaluation of adverse events. The study urges improved reporting of near-misses and actual incidents to enhance risk management. It also advocates for comprehensive clinical training to prevent adverse events and proposes the establishment of robust reporting mechanisms to promote ongoing enhancements in telehealth practices [61]. There is a clear and pressing need to establish standardized protocols for reporting adverse events in telerehabilitation, as highlighted by the findings. This will enable better risk management and improve the overall quality and safety of telerehabilitation services.

The review authors emphasized the importance of addressing privacy concerns and noted a lack of attention to these issues in the included studies. This highlights the need for secure data handling and adherence to privacy regulations. The studies reviewed here briefly addressed privacy and security concerns but did not thoroughly investigate these aspects. Privacy risks in telerehabilitation practice include environmental factors (such as inadequate private spaces for sensitive consultations), technology issues (including data security and limited internet access), and operational challenges [70]. The privacy concerns raised in the review align closely with elements of the MERTH framework: interoperability, particularly regarding exchange, safety, privacy, technology standardization, and comprehensiveness [10]. The challenges highlighted, such as inadequate private spaces, data security issues, and limited internet access are foundational to the safe and effective exchange of information in telerehabilitation. These issues underscore the importance of prioritizing robust data handling and adherence to privacy regulations to safeguard sensitive information of older adults [10].

Empowerment was also critical, with studies stressing the importance of making participants feel capable and confident in using health technologies. There is a noticeable gap between the digital technologies developed for older adults and what they actually need. Aging is often viewed negatively, with older adults stereotypically seen as frail and incompetent [20]. Consequently, many technologies designed for this demographic focus primarily on care rather than empowerment. These negative stereotypes commonly result in excluding older adults from the research and design processes of digital technology. Mannheim et al [71] argue that including older adults in these processes is crucial for creating technologies that genuinely enhance their well-being. Age limits are frequently justified by ageist assumptions that question the competence and reliability

XSL•FO
RenderX

of older adults. Social exclusion, driven by these stereotypes, is evident in various areas, including access to services, social relations, and research participation. Arbitrary upper age limits, which more than half of the relevant studies impose, often result in the unjustified exclusion of older adults from clinical research and randomized controlled trials. Although the use of these exclusion criteria may be declining, selective exclusion remains common [71]. These practices highlight ethical challenges in recognizing older adults' ability to participate meaningfully in research and design, underscoring the impact of stereotypes on their involvement.

This issue aligns closely with the MERTH framework's element humanization, which emphasizes the importance of person-centered care, communication, and empowerment in health care [10]. The humanization of health care, especially in the context of telerehabilitation, means creating environments where patients are treated as whole individuals, not just as a collection of symptoms or conditions. This approach fosters respect, dignity, and empowerment, encouraging patients to engage actively in their own care and decision-making. In the case of telerehabilitation and digital health technologies, this translates to providing digital spaces that not only cater to physical rehabilitation but also consider emotional, social, and psychological needs [10].

Respect and knowledge of cultural diversity were crucial, particularly for Indigenous older adults, calling for culturally appropriate content and rewards [26]. There is a lack of studies that include Indigenous older adult populations. A study conducted with Indigenous older adults in Saskatchewan showed that many express a desire to learn more about technology and recognize its value in supporting healthy aging [26]. The study highlighted that some researchers emphasized the importance of Indigenous language apps and the revitalization of cultural teachings, seeking technology's assistance in facilitating these cultural learning processes. It emphasizes the importance of considering the role of families in engaging older adults with technology, given the crucial cultural role families play within Indigenous communities [26].

The findings of this review underline the necessity of integrating comprehensive ethical considerations into telerehabilitation interventions to ensure safety, privacy, and respect for all populations.

Study Strengths and Limitations

The strengths of this rapid review include its comprehensive synthesis of a wide range of literature within a condensed timeframe, providing timely insights into telerehabilitation for older adults. This efficiency allows for a quick overview of key findings and trends, aiding health care practitioners and policy makers in making informed decisions promptly. Additionally, the review's focus on ethics and equity in telerehabilitation ensures that critical considerations around patient safety, privacy, and inclusivity are highlighted, promoting ethical standards in telerehabilitation delivery. Finally, by emphasizing the integration of diverse perspectives and the identification of gaps in current research, this rapid review contributes to advancing knowledge and guiding future investigations aimed at enhancing telerehabilitation effectiveness and accessibility for older adults.

Study limitations include language constraints, as searches were limited to English and French publications. Future research should explore strategies to access studies in all languages, particularly in equity contexts, to ensure broader inclusivity. Additionally, the review did not ascertain the specific rehabilitation professionals involved in the studies, as individual study details were not accessed, and this information was, for the most part, not presented in the reviews. Furthermore, data extraction was conducted by a single reviewer, which may have introduced potential biases. To address this, the single reviewer discussed the results with the team and engaged in iterative discussions to mitigate any potential biases. Additionally, to further ensure accuracy, 10% of the reviews, including abstracts, were independently cross-checked by another team member. These limitations highlight the need for comprehensive approaches in future reviews to enhance inclusivity and rigor in synthesizing evidence across diverse contexts and disciplines.

Conclusions

In conclusion, this review has shed light on the distribution of studies across various health conditions within telerehabilitation contexts, emphasizing a predominant focus on cardiorespiratory, neurological, and musculoskeletal conditions. These findings highlight both the areas of concentrated research and gaps, particularly in less-studied morbidities such as vision and hearing problems, balance disorders, and pelvic floor disorders.

The review also emphasizes disparities in technology adoption and research participation among different demographic groups among older and middle-aged adults, particularly highlighting racial disparities and socioeconomic influences on telerehabilitation use. Challenges persist in addressing these disparities, with rural areas facing barriers due to limited infrastructure and resources. Moreover, the review has identified ethical considerations as paramount, emphasizing the need for standardized protocols in adverse event reporting, improved privacy, and inclusive research practices that recognize and respect cultural diversity, including Indigenous older adults. Moving forward, addressing these gaps and challenges is essential to advancing equitable access to effective telerehabilitation services for older adults across diverse populations.

Acknowledgments

We thank the Canadian Institutes for Health Research, Association Médicale Canadienne, and Réseau Québécois de Recherche en Adaptation-Réadaptation (REPAR).

XSL•FO
RenderX

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during this study.

Authors' Contributions

MV wrote the first and final version of this paper. MV, JS, NG, AH, and DK were involved in the methodology, data screening, extraction, and analysis of the results. All authors (MV, JS, NG, LPA, MLAN, WCM, AH, and DK) contributed critically to all phases of this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy for this paper.
[DOCX File, 130 KB - aging_v8i1e69660_app1.docx]

Multimedia Appendix 2
Study characteristics of included studies.
[DOCX File, 129 KB - aging_v8i1e69660_app2.docx]

References

1. Marzaleh MA, Peyravi M, Azhdari N, Bahaadinbeigy K, Sarpourian F. Application of telerehabilitation for older adults during the COVID-19 pandemic: a systematic review. Disaster Med Public Health Prep 2023;17:e402. [doi: 10.1017/dmp.2022.219] [Medline: 36004522]
2. Roots RK, Brown H, Bainbridge L, Li LC. Rural rehabilitation practice: perspectives of occupational therapists and physical therapists in British Columbia, Canada. Rural Remote Health 2014;14:2506. [Medline: 24528153]
3. Pol M, Qadeer A, van Hartingsveldt M, Choukou MA. Perspectives of rehabilitation professionals on implementing a validated home telerehabilitation intervention for older adults in geriatric rehabilitation: multisite focus group study. JMIR Rehabil Assist Technol 2023 Jul 18;10:e44498. [doi: 10.2196/44498] [Medline: 37463040]
4. Omboni S, McManus RJ, Bosworth HB, et al. Evidence and recommendations on the use of telemedicine for the management of arterial hypertension: an international expert position paper. Hypertension 2020 Nov;76(5):1368-1383. [doi: 10.1161/HYPERTENSIONAHA.120.15873] [Medline: 32921195]
5. Gokalp H, de Folter J, Verma V, Fursse J, Jones R, Clarke M. Integrated telehealth and telecare for monitoring frail elderly with chronic disease. Telemed J E Health 2018 Dec;24(12):940-957. [doi: 10.1089/tmj.2017.0322] [Medline: 30129884]
6. Appleby E, Gill ST, Hayes LK, Walker TL, Walsh M, Kumar S. Effectiveness of telerehabilitation in the management of adults with stroke: a systematic review. PLoS One 2019;14(11):e0225150. [doi: 10.1371/journal.pone.0225150] [Medline: 31714924]
7. Md Fadzil NH, Shahar S, Rajikan R, et al. A scoping review for usage of telerehabilitation among older adults with mild cognitive impairment or cognitive frailty. Int J Environ Res Public Health 2022 Mar 28;19(7):4000. [doi: 10.3390/ijerph19074000] [Medline: 35409683]
8. Kruse C, Fohn J, Wilson N, Nunez Patlan E, Zipp S, Mileski M. Utilization barriers and medical outcomes commensurate with the use of telehealth among older adults: systematic review. JMIR Med Inform 2020 Aug 12;8(8):e20359. [doi: 10.2196/20359] [Medline: 32784177]
9. Garritty C, Gartlehner G, Nussbaumer-Streit B, et al. Cochrane Rapid Reviews Methods Group offers evidence-informed guidance to conduct rapid reviews. J Clin Epidemiol 2021 Feb;130:13-22. [doi: 10.1016/j.jclinepi.2020.10.007] [Medline: 33068715]
10. Veras M, Labbé DR, Furlano J, et al. A framework for equitable virtual rehabilitation in the metaverse era: challenges and opportunities. Front Rehabil Sci 2023;4:1241020. [doi: 10.3389/fresc.2023.1241020] [Medline: 37691912]
11. Kim YM, Jang SN, Cho SI. Working hours, social engagement, and depressive symptoms: an extended work-life balance for older adults. BMC Public Health 2024;23(1):23. [doi: 10.1186/s12889-023-17072-x]
12. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
13. Reiners F, Sturm J, Bouw LJW, Wouters EJM. Sociodemographic factors influencing the use of ehealth in people with chronic diseases. Int J Environ Res Public Health 2019 Feb 21;16(4):645. [doi: 10.3390/ijerph16040645] [Medline: 30795623]
14. Guo Y, Albright D. The effectiveness of telehealth on self-management for older adults with a chronic condition: a comprehensive narrative review of the literature. J Telemed Telecare 2018 Jul;24(6):392-403. [doi: 10.1177/1357633X17706285] [Medline: 28449619]

XSL•FO
RenderX

15. Bhattarai P, Phillips JL. The role of digital health technologies in management of pain in older people: an integrative review. Arch Gerontol Geriatr 2017;68:14-24. [doi: 10.1016/j.archger.2016.08.008] [Medline: 27584871]
16. Beckie TM. Utility of home-based cardiac rehabilitation for older adults. Clin Geriatr Med 2019 Nov;35(4):499-516. [doi: 10.1016/j.cger.2019.07.003] [Medline: 31543181]
17. Huang J, Han Y, Wei J, et al. The effectiveness of the Internet-based self-management program for cancer-related fatigue patients: a systematic review and meta-analysis. Clin Rehabil 2020 Mar;34(3):287-298. [doi: 10.1177/0269215519889394] [Medline: 31793340]
18. Malaguti C, Dal Corso S, Janjua S, Holland AE. Supervised maintenance programmes following pulmonary rehabilitation compared to usual care for chronic obstructive pulmonary disease. Cochrane Database Syst Rev 2021 Aug 17;8(8):CD013569. [doi: 10.1002/14651858.CD013569.pub2] [Medline: 34404111]
19. Kraaijkamp JJM, van Dam van Isselt EF, Persoon A, Versluis A, Chavannes NH, Achterberg WP. eHealth in geriatric rehabilitation: systematic review of effectiveness, feasibility, and usability. J Med Internet Res 2021 Aug 19;23(8):e24015. [doi: 10.2196/24015] [Medline: 34420918]
20. Reeder B, Chung J, Stevens-Lapsley J. Current telerehabilitation research with older adults at home: an integrative review. J Gerontol Nurs 2016 Oct 1;42(10):15-20. [doi: 10.3928/00989134-20160201-02] [Medline: 26870988]
21. Dennett EJ, Janjua S, Stovold E, Harrison SL, McDonnell MJ, Holland AE. Tailored or adapted interventions for adults with chronic obstructive pulmonary disease and at least one other long-term condition: a mixed methods review. Cochrane Database Syst Rev 2021 Jul 26;7(7):CD013384. [doi: 10.1002/14651858.CD013384.pub2] [Medline: 34309831]
22. Devi R, Singh SJ, Powell J, Fulton EA, Igbinedion E, Rees K. Internet-based interventions for the secondary prevention of coronary heart disease. Cochrane Database Syst Rev 2015 Dec 22;2015(12):CD009386. [doi: 10.1002/14651858.CD009386.pub2] [Medline: 26691216]
23. Dequanter S, Gagnon MP, Ndiaye MA, et al. The effectiveness of e-Health solutions for aging with cognitive impairment: a systematic review. Gerontologist 2021 Sep 13;61(7):e373-e394. [doi: 10.1093/geront/gnaa065] [Medline: 32525977]
24. Ambrens M, Alley S, Oliveira JS, et al. Effect of eHealth-delivered exercise programmes on balance in people aged 65 years and over living in the community: a systematic review and meta-analysis of randomised controlled trials. BMJ Open 2022 Jun 10;12(6):e051377. [doi: 10.1136/bmjopen-2021-051377] [Medline: 35688586]
25. Ding J, Yang Y, Wu X, Xiao B, Ma L, Xu Y. The telehealth program of occupational therapy among older people: an up-to-date scoping review. Aging Clin Exp Res 2023 Jan;35(1):23-40. [doi: 10.1007/s40520-022-02291-w] [Medline: 36344805]
26. Choukou MA, Maddahi A, Polyvyana A, Monnin C. Digital health technology for Indigenous older adults: a scoping review. Int J Med Inform 2021 Apr;148:104408. [doi: 10.1016/j.ijmedinf.2021.104408] [Medline: 33609927]
27. Saito T, Izawa KP. Effectiveness and feasibility of home-based telerehabilitation for community-dwelling elderly people in Southeast Asian countries and regions: a systematic review. Aging Clin Exp Res 2021 Oct;33(10):2657-2669. [doi: 10.1007/s40520-021-01820-3] [Medline: 33765258]
28. Sülz S, van Elten HJ, Askari M, Weggelaar-Jansen AM, Huijsman R. eHealth applications to support independent living of older persons: scoping review of costs and benefits identified in economic evaluations. J Med Internet Res 2021 Mar 9;23(3):e24363. [doi: 10.2196/24363] [Medline: 33687335]
29. Su JJ, Yu DSF, Paguio JT. Effect of eHealth cardiac rehabilitation on health outcomes of coronary heart disease patients: a systematic review and meta-analysis. J Adv Nurs 2020 Mar;76(3):754-772. [doi: 10.1111/jan.14272] [Medline: 31769527]
30. Theodoros D, Aldridge D, Hill AJ, Russell T. Technology-enabled management of communication and swallowing disorders in Parkinson's disease: a systematic scoping review. Int J Lang Commun Disord 2019 Mar;54(2):170-188. [doi: 10.1111/1460-6984.12400] [Medline: 29923267]
31. Solis-Navarro L, Gismero A, Fernández-Jané C, et al. Effectiveness of home-based exercise delivered by digital health in older adults: a systematic review and meta-analysis. Age Ageing 2022 Nov 2;51(11):11. [doi: 10.1093/ageing/afac243] [Medline: 36346736]
32. Tschoepe R, Benfield A, Posey R, Mercer V. A systematic review of the effects of community transition programs on quality of life and hospital readmissions for adults with traumatic spinal cord injury. Arch Phys Med Rehabil 2022 May;103(5):1013-1022. [doi: 10.1016/j.apmr.2021.08.002] [Medline: 34464608]
33. van Egmond MA, van der Schaaf M, Vredeveld T, et al. Effectiveness of physiotherapy with telerehabilitation in surgical patients: a systematic review and meta-analysis. Physiotherapy 2018 Sep;104(3):277-298. [doi: 10.1016/j.physio.2018.04.004] [Medline: 30030037]
34. Whitfield T, Barnhofer T, Acabchuk R, et al. The effect of mindfulness-based programs on cognitive function in adults: a systematic review and meta-analysis. Neuropsychol Rev 2022 Sep;32(3):677-702. [doi: 10.1007/s11065-021-09519-y] [Medline: 34350544]
35. Geraedts H, Zijlstra A, Bulstra SK, Stevens M, Zijlstra W. Effects of remote feedback in home-based physical activity interventions for older adults: a systematic review. Patient Educ Couns 2013 Apr;91(1):14-24. [doi: 10.1016/j.pec.2012.10.018] [Medline: 23194823]

XSL•FO
RenderX

36. Curry J, Patterson M, Greenley S, Pearson M, Forbes CC. Feasibility, acceptability, and efficacy of online supportive care for individuals living with and beyond lung cancer: a systematic review. Support Care Cancer 2021 Nov;29(11):6995-7011. [doi: 10.1007/s00520-021-06274-x] [Medline: 34008080]
37. Brick R, Turner R, Bender C, et al. Impact of non-pharmacological interventions on activity limitations and participation restrictions in older breast cancer survivors: a scoping review. J Geriatr Oncol 2022 Mar;13(2):132-142. [doi: 10.1016/j.jgo.2021.09.010] [Medline: 34598902]
38. Guan T, Qan'ir Y, Song L. Systematic review of illness uncertainty management interventions for cancer patients and their family caregivers. Support Care Cancer 2021 Aug;29(8):4623-4640. [doi: 10.1007/s00520-020-05931-x] [Medline: 33495851]
39. Chan JKY, Klainin-Yobas P, Chi Y, Gan JKE, Chow G, Wu XV. The effectiveness of e-interventions on fall, neuromuscular functions and quality of life in community-dwelling older adults: a systematic review and meta-analysis. Int J Nurs Stud 2021 Jan;113:103784. [doi: 10.1016/j.ijnurstu.2020.103784] [Medline: 33120138]
40. Bacanoiu MV, Danoiu M. New strategies to improve the quality of life for normal aging versus pathological aging. J Clin Med 2022 Jul 20;11(14):14. [doi: 10.3390/jcm11144207] [Medline: 35887969]
41. Zhang J, Yang M, Ge Y, Ivers R, Webster R, Tian M. The role of digital health for post-surgery care of older patients with hip fracture: a scoping review. Int J Med Inform 2022 Apr;160:104709. [doi: 10.1016/j.ijmedinf.2022.104709] [Medline: 35144101]
42. Tao KFM, Brennan-Jones CG, Capobianco-Fava DM, et al. Teleaudiology services for rehabilitation with hearing aids in adults: a systematic review. J Speech Lang Hear Res 2018 Jul 13;61(7):1831-1849. [doi: 10.1044/2018_JSLHR-H-16-0397] [Medline: 29946688]
43. Gaspar AGM, Lapão LV. eHealth for addressing balance disorders in the elderly: systematic review. J Med Internet Res 2021;23(4):e22215. [doi: 10.2196/22215]
44. Johnson AJ, Palit S, Terry EL, et al. Managing osteoarthritis pain with smart technology: a narrative review. Rheumatol Adv Pract 2021;5(1):rkab021. [doi: 10.1093/rap/rkab021] [Medline: 33928214]
45. Jonker LT, Haveman ME, de Bock GH, van Leeuwen BL, Lahr MMH. Feasibility of perioperative eHealth interventions for older surgical patients: a systematic review. J Am Med Dir Assoc 2020 Dec;21(12):1844-1851. [doi: 10.1016/j.jamda.2020.05.035] [Medline: 32694000]
46. Bostrom J, Sweeney G, Whiteson J, Dodson JA. Mobile health and cardiac rehabilitation in older adults. Clin Cardiol 2020 Feb;43(2):118-126. [doi: 10.1002/clc.23306] [Medline: 31825132]
47. Del Pino R, Díez-Cirarda M, Ustarroz-Aguirre I, et al. Costs and effects of telerehabilitation in neurological and cardiological diseases: a systematic review. Front Med (Lausanne) 2022;9:832229. [doi: 10.3389/fmed.2022.832229] [Medline: 36523783]
48. Linn N, Goetzinger C, Regnaux JP, et al. Digital health interventions among people living with frailty: a scoping review. J Am Med Dir Assoc 2021 Sep;22(9):1802-1812. [doi: 10.1016/j.jamda.2021.04.012] [Medline: 34000266]
49. Marzano G, Lubkina V. Usability in social telerehabilitation systems for elderly users. Public Health 2017 Mar;144:1-3. [doi: 10.1016/j.puhe.2016.11.011] [Medline: 28274368]
50. Amirova A, Fteropoulli T, Williams P, Haddad M. Efficacy of interventions to increase physical activity for people with heart failure: a meta-analysis. Open Heart 2021 Jun;8(1):e001687. [doi: 10.1136/openhrt-2021-001687] [Medline: 34108272]
51. Velayati F, Ayatollahi H, Hemmat M. A systematic review of the effectiveness of telerehabilitation interventions for therapeutic purposes in the elderly. Methods Inf Med 2020 May;59(2-03):104-109. [doi: 10.1055/s-0040-1713398] [Medline: 32629502]
52. Emmanouilidis S, Hackney ME, Slade SC, Heng H, Jazayeri D, Morris ME. Dance is an accessible physical activity for people with Parkinson's disease. Parkinsons Dis 2021;2021:7516504. [doi: 10.1155/2021/7516504] [Medline: 34721836]
53. Ollevier A, Aguiar G, Palomino M, Simpelaere IS. How can technology support ageing in place in healthy older adults? A systematic review. Public Health Rev 2020 Nov 23;41(1):26. [doi: 10.1186/s40985-020-00143-4] [Medline: 33292707]
54. Allida S, Du H, Xu X, et al. mHealth education interventions in heart failure. Cochrane Database Syst Rev 2020 Jul 2;7(7):CD011845. [doi: 10.1002/14651858.CD011845.pub2] [Medline: 32613635]
55. Xie SH, Wang Q, Wang LQ, Wang L, Song KP, He CQ. Effect of internet-based rehabilitation programs on improvement of pain and physical function in patients with knee osteoarthritis: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2021 Jan 5;23(1):e21542. [doi: 10.2196/21542] [Medline: 33399542]
56. Alaimo C, Campana E, Stoppelli MR, et al. Cognitive tele-enhancement in healthy older adults and subjects with subjective memory complaints: a review. Front Neurol 2021;12:650553. [doi: 10.3389/fneur.2021.650553] [Medline: 34290660]
57. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: an integrative systematic review. BMC Public Health 2024 Feb 16;24(1):506. [doi: 10.1186/s12889-024-18036-5] [Medline: 38365698]
58. Yi JS, Pittman CA, Price CL, Nieman CL, Oh ES. Telemedicine and dementia care: a systematic review of barriers and facilitators. J Am Med Dir Assoc 2021 Jul;22(7):1396-1402. [doi: 10.1016/j.jamda.2021.03.015] [Medline: 33887231]
59. Johnsen TL, Johansen T, Momsen AMH, et al. eHealth interventions to facilitate work participation: a scoping review. JBI Evid Synth 2021 Oct;19(10):2739-2759. [doi: 10.11124/JBISRIR-D-19-00433] [Medline: 34224522]

XSL•FO
RenderX

60. Antunes TPC, de Oliveira ASB, Hudec R, et al. Assistive technology for communication of older adults: a systematic review. Aging Ment Health 2019 Apr;23(4):417-427. [doi: 10.1080/13607863.2018.1426718] [Medline: 29451399]
61. Martiniuk A, Toepfer A, Lane-Brown A. A review of risks, adverse effects and mitigation strategies when delivering mental health services using telehealth. J Ment Health 2024 Jun;33(3):415-438. [doi: 10.1080/09638237.2023.2182422] [Medline: 36866784]
62. Matera MG, Hanania NA, Maniscalco M, Cazzola M. Pharmacotherapies in Older Adults with COPD: Challenges and Opportunities. Drugs Aging 2023 Jul;40(7):605-619. [doi: 10.1007/s40266-023-01038-0] [Medline: 37316689]
63. Marengoni A, Angleman S, Meinow B, et al. Coexisting chronic conditions in the older population: variation by health indicators. Eur J Intern Med 2016 Jun;31:29-34. [doi: 10.1016/j.ejim.2016.02.014] [Medline: 26944564]
64. Chen XS, Wang K. Gender disparities in telehealth use among older adults in the United States during the COVID-19 pandemic. IJPS 2024;10(4):114-124. [doi: 10.36922/ijps.1817]
65. Arcury TA, Sandberg JC, Melius KP, et al. Older adult internet use and eHealth literacy. J Appl Gerontol 2020 Feb;39(2):141-150. [doi: 10.1177/0733464818807468] [Medline: 30353776]
66. Delello JA, McWhorter RR. Reducing the digital divide: connecting older adults to iPad technology. J Appl Gerontol 2017 Jan;36(1):3-28. [doi: 10.1177/0733464815589985] [Medline: 26084479]
67. Guzman-Parra J, Barnestein-Fonseca P, Guerrero-Pertiñez G, et al. Attitudes and use of information and communication technologies in older adults with mild cognitive impairment or early stages of dementia and their caregivers: cross-sectional study. J Med Internet Res 2020 Jun 1;22(6):e17253. [doi: 10.2196/17253] [Medline: 32442136]
68. Veras M, Stewart J, Deonandan R, et al. Cost analysis of a home-based virtual reality rehabilitation to improve upper limb function in stroke survivors. GJHS 2020;12(2):98. [doi: 10.5539/gjhs.v12n2p98]
69. Kairy D, Lehoux P, Vincent C, Visintin M. A systematic review of clinical outcomes, clinical process, healthcare utilization and costs associated with telerehabilitation. Disabil Rehabil 2009;31(6):427-447. [doi: 10.1080/09638280802062553] [Medline: 18720118]
70. Houser SH, Flite CA, Foster SL. Privacy and security risk factors related to telehealth services - a systematic review. Perspect Health Inf Manag 2023;20(1):1f. [Medline: 37215337]
71. Mannheim I, Schwartz E, Xi W, et al. Inclusion of older adults in the research and design of digital technology. Int J Environ Res Public Health 2019 Oct 2;16(19):19. [doi: 10.3390/ijerph16193718] [Medline: 31581632]

Abbreviations

MERTH: Metaverse Equitable Rehabilitation Therapy

Edited by C Lee; submitted 05.12.24; peer-reviewed by LK Chua, T Adeniji; revised version received 15.05.25; accepted 03.07.25; published 13.08.25.

Please cite as:
Veras M, Auger LP, Sigouin J, Gheidari N, Nelson MLA, Miller WC, Hudon A, Kairy D
Ethics and Equity Challenges in Telerehabilitation for Older Adults: Rapid Review
JMIR Aging 2025;8:e69660
URL: https://aging.jmir.org/2025/1/e69660
doi:10.2196/69660

© Mirella Veras, Louis-Pierre Auger, Jennifer Sigouin, Nahid Gheidari, Michelle LA Nelson, William C Miller, Anne Hudon, Dahlia Kairy. Originally published in JMIR Aging (https://aging.jmir.org), 13.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Best Practices for Implementing Electronic Care Records in Adult Social Care: Rapid Scoping Review

Martha Snow[1], MA; Wagner Silva-Ribeiro[1], PhD; Mary Baginsky[2], PhD; Sonya Di Giorgio[3], MA; Nicola Farrelly[4], PhD; Cath Larkins[4], PhD; Karen Poole[3], MSc; Nicole Steils[2], PhD; Joanne Westwood[4], PhD; Juliette Malley[1], PhD

[1]Care Policy and Evaluation Centre, London School of Economics and Political Science (LSE), London, United Kingdom

[2]NIHR Health & Social Care Workforce Research Unit, King's College London, London, United Kingdom

[3]Libraries & Collections, King's College London, London, United Kingdom

[4]School of Health, Social Work and Sport, University of Central Lancashire, Preston, Lancashire, United Kingdom

Corresponding Author:
Martha Snow, MA
Care Policy and Evaluation Centre
London School of Economics and Political Science (LSE)
PAN 8.01J
Houghton Street
London, WC2A 2AE
United Kingdom
Phone: 44 (0)20 7405 7686
Email: m.snow@lse.ac.uk

Abstract

Background: In the past decade, the use of digital or electronic records in social care has risen worldwide, capturing key information for service delivery. The COVID-19 pandemic accelerated digitization in health and social care. For example, the UK government created a fund specifically for adult social care provider organizations to adopt digital social care records. These developments offer valuable learning opportunities for implementing digital care records in adult social care settings.

Objective: This rapid scoping review aimed to understand what is known about the implementation of digital care records in adult social care and how implementation varies across use cases, settings, and broader contexts.

Methods: A scoping review methodology was used, with amendments made to enable a rapid review. Comprehensive searches based on the concepts of digital care records, social care, and interoperability were conducted across the MEDLINE, EmCare, Web of Science Core Collection, HMIC Health Management Information Consortium, Social Policy and Practice, and Social Services Abstracts databases. Studies published between 2018 and 2023 in English were included. One reviewer screened titles and abstracts, while 2 reviewers extracted data. Thematic analysis mapped findings against the nonadoption, abandonment, scale-up, spread, and sustainability (NASSS) framework.

Results: Our search identified 2499 references. After screening titles and abstracts, 71 records were selected for full-text review, resulting in 31 references from 29 studies. Studies originated from 11 countries, including 1 multicountry study, with the United Kingdom being the most represented (10/29, 34%). Studies were most often conducted in nursing homes or facilities (7/29, 24%) with older people as the target population (6/29, 21%). Health records were the most investigated record type (12/29, 41%). We identified 45 facilitators and 102 barriers to digital care record implementation across 28 studies, spanning 6 of the 7 NASSS framework domains and aligning with 5 overarching themes that require greater active management regarding implementation. Intended or actual implementation outcomes were reported in 17 (59%) of the 29 studies.

Conclusions: The findings suggest that implementation is complex due to a lack of consensus on what digital care records and expected outcomes and impacts should look like. The literature often lacks clear definitions and robust study designs. To be successful, implementation should consider complexity, while studies should use robust frameworks and mixed methods or quantitative designs where appropriate. Future research should define the target population, gather data on carer or service user experiences, and focus on digital care records specifically used in social care.

(JMIR Aging 2025;8:e60107) doi:10.2196/60107

XSL•FO
RenderX

KEYWORDS

digital care records; adult social care; digitization; domiciliary care; care homes; electronic care records; PRISMA

Introduction

Background

The demand for adult social care is vast. Global demographic changes throughout the 20th century have led to substantial population aging, decreased mortality and communicable diseases, and increased chronic noncommunicable diseases. Consequently, more adults and older people have long-term care needs, particularly in high-income countries where the epidemiological transition began earlier [1]. Technology has been proposed to help manage this increasing demand in health and social care by improving efficiency, care quality, and effectiveness [2-4]. Digital care records are one such innovation in adult social care.

In this paper, adult social care refers to long-term, aged, or disability care, including care homes; support in the home; domiciliary care (eg, personal care, practical tasks, and crisis support); community-based support such as inclusive arts programs; and social relationships that aim to keep people independent, active, and living well.

The use of digital care records has increased across various adult social care settings and countries since 2012 [5]. These records capture key information for service delivery, including individuals' characteristics, the care they receive, and how they respond to it. They monitor service users and track service delivery, supporting care planning, medication, and assessments [6-10]. In addition, they serve administrative purposes [8,11,12], support compliance with data documentation regulations [13], and inform care delivery decisions [14,15]. Different terms are used to describe digital care records in social care. In the United Kingdom, the term digital social care records (DSCRs) is common. In North America, parts of Europe, and Australia, terms such as electronic health records [16-20], electronic patient records [6], or electronic medical records [9,19] are often used. Digital care records can be part of health information exchange initiatives, which facilitate data sharing across health and social care to improve care continuity and efficiency [16-20].

Despite the increasing use of digital care records, much of the existing literature focuses on their implementation in nursing homes or approaches the topic from a social work perspective, failing to capture the full scope of adult social care. One systematic review of electronic health records identified that health information exchange is facilitated by workflow integration and flexible organizational culture and impeded by incomplete data, inefficiency, and unfavorable market conditions [21]. Another review found that electronic health records support health outcomes, clinical documentation management, and decision-making [2]. To the best of our knowledge, the only previous review of DSCRs that looked at the benefits of implementation was a review by Greenstock [22]. This literature review highlighted improved documentation and health outcomes as well as increased collaboration and communication, efficiency, quality of care, client or family involvement, and risk management [22]. It identified less evidence regarding financial benefits and increased workforce satisfaction [22]. It is unclear how many benefits were realized versus anticipated [22]. A scoping review of electronic information systems in social care found that they can negatively affect social workers' priorities and do not meet sector needs [23].

These reviews predate the COVID-19 pandemic, which accelerated digital system development in health and social care [24]. For example, the UK government injected funds during the COVID-19 pandemic to drive digitization and has since continued these efforts. A specific fund for adult social care provider organizations supports DSCR adoption, with the most recent government target of 80% adoption across adult social care provider organizations in England by March 2025.

Objectives

The intensity of the activity discussed above presented an opportunity to learn about DSCR implementation and impact through evaluation. Considering recent developments, this rapid scoping review sought to assess what is known about DSCR implementation in adult social care settings and identify evidence gaps to inform a rapid evaluation of DSCR implementation. While this purpose has influenced decisions around the methods, such as a rapid approach and more intensive searching for UK literature, the review considers the international literature on DSCR implementation and draws out implications for an international audience.

We mapped our findings against the nonadoption, abandonment, scale-up, spread, and sustainability (NASSS) framework by Greenhalgh et al [25]. Designed in 2017 as an evidence-based, theory-informed, and pragmatic tool, it helps predict and evaluate the success of technology-supported health or social care programs. As it focuses on adoption, nonadoption, and abandonment of technologies as well as the challenges associated with the scale-up, spread, and maintenance of digital systems, it was deemed appropriate for capturing the field's complexity. The framework was particularly useful during data analysis. Most of the literature retrieved identified large numbers of facilitators of and barriers to DSCR adoption. The NASSS framework helped to position these within an interrelated system and organize them in a way that could provide guidance in areas requiring active management of complexity. As the NASSS framework has been applied more often to health care settings, this review was also an opportunity to explore its value for technology adoption in social care.

Methods

Overview

The rapid scoping review followed the 6-stage framework outlined by Arksey and O'Malley [26], which was later refined by Levac et al [27] and the Joanna Briggs Institute [28]. Following the study by Tricco et al [29], we made some amendments to enable a rapid review. The review is reported in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for

XSL•FO
RenderX

Scoping Reviews) guidelines (Multimedia Appendix 1) [30]. The search strategy is reported in accordance with the PRISMA-S (Preferred Reporting Items for Systematic reviews and Meta-Analyses literature search extension) checklist [31]. A protocol for this review was developed using the PRISMA-ScR and registered prospectively with the Open Science Framework on August 9, 2023 [32].

Identifying the Research Question

We used the Joanna Briggs Institute's population, concept, and context framework [28] to formulate the following scoping review questions: (1) What is known about the implementation of DSCRs in social care settings? and (2) How does implementation vary across use cases, social care settings, and the broader context? The subquestions were as follows:

- What DSCR is being used?
- What situation or setting is the DSCR being used in, and which actors are involved?
- What is the broader context within which DSCRs are being implemented or used?
- What is the use case for the DSCR, and what are the intended outcomes and benefits?
- How has the implementation of DSCRs been evaluated or researched, and what theoretical framings have been used?
- What are the intended or actual outcomes and benefits of DSCR implementation?
- What helps or gets in the way of the implementation of DSCRs?

Identifying Relevant Studies

A librarian with experience in undertaking reviews (KP) designed the search in consultation with the research team. The search was undertaken between August 2, 2023, and August 11, 2023, by 2 librarians (KP and SDG) across MEDLINE (through Ovid; KP), EmCare (through Ovid; SDG), Web of Science Core Collection (Clarivate; KP), HMIC Health Management Information Consortium (through Ovid; KP), Social Policy and Practice (through Ovid; KP), and Social Services Abstracts (through ProQuest; SDG) databases.

The search strategy used 3 concepts: digital care records, social care, and interoperability. These concepts were combined in the search string as (Digital Care Records AND Social Care) OR (Social Care AND Interoperability). The interoperability concept was included, as it is central to policy narratives surrounding the implementation of DSCRs in England, with expectations that DSCRs will facilitate data sharing with general practitioners and hospitals. The initial search strategy was developed on MEDLINE (Ovid) by one of the librarians (KP) and run in each database by KP and SDG. Publications were limited to those published in or after 2018 until 2023. The results were limited to the English language. The databases were searched using keywords and controlled vocabulary (eg, Medical Subject Headings or Emtree) where appropriate and adapted according to the requirements of each database. The full search strategy for each database can be found in Multimedia Appendix 2.

There were 3466 results in total. The results were exported to EndNote (Clarivate), and 993 duplicates were removed following a structured process [33], leaving 2473 unique results. These were exported as a research information systems file to Covidence (Veritas Health Innovation Ltd) software [34] for title and abstract screening as well as for full-text review.

In addition, given the intention of informing an evaluation in the context of the English language, we searched key English websites to capture gray literature not identified through the databases. The chosen websites were the Local Government Association [35]; King's Fund [36]; Social Care Institute of Excellence [37]; Centre for Care [38]; Digital Care Hub, formerly Digital Social Care [39]; and TEC Service Association [40]. Searches were also performed on Google, and we contacted experts identified through the review. From these searches and reference checking, 27 references were identified. Of these 27 references, 1 (4%) duplicate was removed, and 1 (4%) reference that reported results from a study already included was merged with the main reference. One reference recommended by an expert was also included. This resulted in 26 references retrieved through our gray literature search.

Study Selection

We included studies that (1) took place within adult social care settings; (2) involved the implementation of a DSCR, which may be referred to by other labels, such as electronic care records and electronic information systems; (3) were carried out using any study design (eg, experimental, quasi-experimental, and observational, including quantitative and qualitative studies); and (4) were published from 2018 onward. This decision was made on the basis that existing reviews have captured the literature on DSCRs up until the end of 2017.

Following rapid review methodology guidance [29,41], all references retrieved from our search were screened by a reviewer with expertise in systematic reviews (WSR). Initial screening was based on titles and abstracts. References were selected for full-text review if they met our inclusion criteria or if it was unclear that they did. The same reviewer (WSR) performed the full-text review. A second reviewer, who is an expert in adult social care research (JM), cross-checked references that were excluded in this phase. Disagreement was discussed until a consensus was reached.

Charting the Data

A data extraction template was developed by the team using Excel (Microsoft Corp). The form included key characteristics of included studies, such as the population, concept, context, study design, and methods, and key findings that were relevant to the review questions. In total, 2 reviewers (MS and WSR) performed the data extraction. Due to the heterogeneity of studies and following best practice, the extraction form was piloted and iteratively adapted through discussions between the 2 reviewers and a third reviewer (JM), who oversaw the extraction process.

During the data extraction, we discovered that 2 publications [42,43] reported results from the same study. Another publication [44] was a preprint version of 1 peer-reviewed article [45], which was also included in the review. All publications were included to ensure we used the information available, but

XSL•FO
RenderX

to avoid duplication of information, we extracted information at the study level rather than the publication level.

Collating, Reporting, and Summarizing the Results

There were several steps involved in collating and reporting the results. We first created a summary of the included studies, categorizing the papers according to relevant study characteristics, such as study design, population, context, methods of data collection and analysis, and theoretical perspectives. We then worked inductively to identify intended or actualized benefits and outcomes and barriers to and facilitators of implementation raised in the papers. Using thematic analysis, we compiled a descriptive overview of the unique barriers and facilitators identified in the papers, including frequency distributions.

We then used the NASSS framework as a sensitizing framework and worked deductively to ensure we had not missed anything of relevance to the NASSS domains. In this process, further barriers and facilitators were identified, and these were mapped alongside those identified from the inductive process to the NASSS domains and subdomains. Where a category was associated with >1 NASSS domain, it was mapped against the domain perceived as most affected.

To synthesize our findings, we then grouped the barriers and facilitators into themes capturing complex aspects of the adoption process. Complexity was determined using the NASSS framework, which defines implementation as simple (ie, few components and predictable), complicated (ie, many components but still largely predictable), or complex (ie, many components interacting in a dynamic and unpredictable way) [46]. The more complexity there is in the system, the less likely the technology is to achieve sustained adoption across the system, and the more likely it is to be abandoned [46]. The themes draw attention to areas that require greater active management with respect to implementation [25].

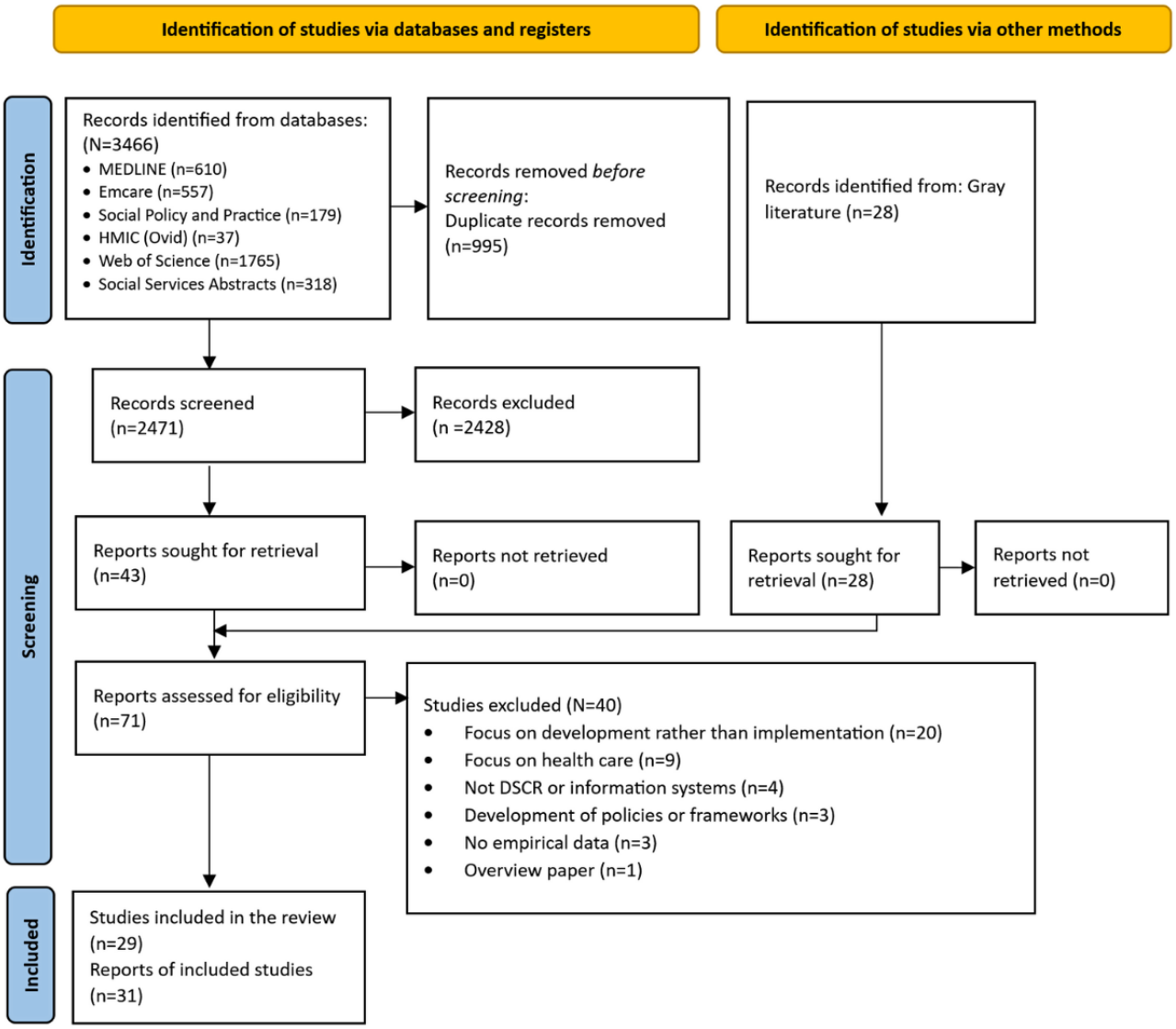
Results

Overview

Our search resulted in 2471 references after duplicates were removed. An additional 28 references were identified through the gray literature search, resulting in 2499 references. After screening references based on titles and abstracts, 71 records were selected for full-text review, of which 31 references were included from 29 different studies (ie, 2 pairs of papers reported on the same studies). The article selection process and reasons for exclusion are presented in Figure 1 [47].

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flowchart. DSCR: digital social care record.



Study Characteristics

As Table 1 presents, of the 29 studies, 10 (34%) were conducted in the United Kingdom—7 (24%) in England [3,7,10,20,48-50], 2 (7%) in Scotland [43,45], and 1 (3%) [51] in multiple UK countries. Of the 29 studies, 5 (17%) were conducted in the United States [9,14-16,18], 3 (10%) in Finland [52-54], 2 (7%) in Australia [11,13], 2 (7%) in Canada [19,55], and 2 (7%) in Sweden [12,56]. The remaining studies were conducted in Switzerland [8], Japan [57], Austria [17], and Italy [58] or involved multiple countries [6].

Studies in the United Kingdom were conducted in care homes (3/29, 10%) [43,45,48], health and social care provider organizations (4/29, 14%) [3,7,20,49], multiple social care settings (1/29, 3%) [51], councils (1/29, 3%) [50], and a continuing health care team (1/29, 3%) [10]. More than one-third of studies from other countries were conducted in nursing homes and facilities (7/29, 24%) [6,8,9,14,15,17,18]. Other settings included home care (5/29, 17%) [12,55-58], care homes (1/29, 3%) [13], long-term care facilities (1/29, 3%) [11], an acute care hospital and its neighboring long-term care home (1/29, 3%) [19], assisted living communities (1/29, 3%) [16], health centers in Finland (2/29, 7%) [52,53], and social care services (1/29, 3%) [54].

XSL•FO
RenderX

Table 1. Key characteristics of the included studies (N=29).

Characteristic	Studies, n (%)
Design	
Cohort	1 (3)
Cross-sectional	6 (21)
Mixed methods	8 (28)
Qualitative	14 (48)
Country	
Australia	2 (7)
Austria	1 (3)
Canada	2 (7)
Finland	3 (10)
Italy	1 (3)
Japan	1 (3)
Sweden	2 (7)
Switzerland	1 (3)
United Kingdom	10 (35)
United States	5 (17)
Belgium, Czech Republic, and Spain	1 (3)
Aim	
Barriers and facilitators	9 (31)
Prevalence of use of DSCRs[a]	7 (24)
Professionals' perceptions about DSCRs	4 (14)
Impact of DSCRs on professionals' work	4 (14)
Services' readiness to implement DSCRs	3 (10)
Strategies to improve DSCRs	1 (3)
Impact of DSCRs on health outcomes	1 (3)
Setting	
Assisted living community	1 (3)
Care homes	4 (14)
Continuing or community health care	3 (10)
Councils	1 (3)
Home care	5 (17)
Multisector (ie, health and social care)	6 (21)
Municipal social services	1 (3)
Nursing homes or facilities	7 (24)
Social care provider organizations	1 (3)
Population	
Older people	6 (21)
Adults with care needs	2 (7)
People with dementia	1 (3)
Not specified	20 (69)
Respondents[b]	

XSL•FO
RenderX

Characteristic	Studies, n (%)
Social care staff	10 (35)
Health care staff	8 (28)
Social care managers	8 (28)
Health care managers	3 (10)
Users or carers	2 (7)
Regional stakeholders	2 (7)
National stakeholders	1 (3)
Technology providers or vendors	1 (3)
Not specified	6 (21)
Types of technology	
Health records	12 (41)
Health and social care records	8 (28)
Social care records	4 (14)
Interoperability	4 (14)
Not specified	1 (3)
Theoretical framework used[b]	
The DeLone and McLean model of information systems success	1 (3)
The Wang and Strong quality framework	1 (3)
Activity theory	1 (3)
Sociotechnical systems theory	1 (3)
Computer-supported cooperative work	1 (3)
Design thinking	1 (3)
Nolan stage model	1 (3)
Normalization process theory	2 (7)
Implementation process framework	1 (3)
Unified theory of acceptance and use of technology	1 (3)
Not specified or applicable	20 (69)

[a]DSCR: digital social care record.

[b]Total is >100% because some studies collected information with different types of informants.

Studies investigated several different types of digital care records—from health information technology in general (4/29, 14%) [18,48,55,57] to electronic medical records or electronic health records specifically (12/29, 41%) [6,8,9,11,12,14-18,53,56]. Some were specific systems commissioned by or developed for care providers, local authorities, or regions, such as the aged care ecosystem [13], the Edotto regional information system [58], CareFirst [10], the PASSsystem [51], or CareCentric [3], among others. Of the 29 studies, 4 (14%) focused on interoperability [10,19,20,57]. Studies can be grouped into those which aimed to (1) identify barriers or facilitators to the implementation of DSCRs and information exchange systems (9/29, 31%) [3,10,13,19,20,48-50,53], (2) assess the proportion and prevalence of services using DSCRs and information exchange systems and how these are being used (7/29, 24%) [14-18,45,56], (3) investigate how digital systems affect the work of care professionals or care providers (4/29, 14%) [7,11,51,54], (4) assess care professionals' perceptions about the use of digital systems (4/29, 14%) [6,8,52,55], (5) map services' readiness or maturity and care professionals' capability to adopt DSCRs and information exchange systems (3/29, 10%) [9,12,42], (6) assess potential strategies to improving existing DSCRs or information exchange systems (1/29, 3%) [58], and (7) assess the impact of use of information communication technology (ICT) on health outcomes (1/29, 3%) [57].

To achieve these aims, most studies used a qualitative design (14/29, 48%) [3,6,7,11-13,17,48-50,52,53,56,58]. Of the 29 studies, 8 (28%) used mixed methods [9,10,19,20,45,51,54,55] and 7 (24%) used a quantitative design [8,14-16,18,43,57]. None of the studies that aimed to identify barriers or facilitators, to investigate how digital systems affect work routines, or to assess potential strategies to improve digital systems used quantitative methods. Among the studies that aimed to assess professionals' perceptions of digital systems, only 1 was quantitative [8]. Of

XSL•FO
RenderX

the studies that aimed to assess the proportion and prevalence of digital systems, most (4/29, 14%) used quantitative methods [14-16,18]. The study assessing the impact of ICT on health outcomes was also quantitative [57]. The studies that aimed to map services' readiness and maturity varied between qualitative [12], quantitative [42], and mixed methods approaches [9].

Most studies (20/29, 69%) did not use a theoretical framework to interpret their results [3,8-10,14-17,19,42,45,48-52,54-57]. Of those that did, theories included normalization process theory (2/29, 7%) [20,53], sociotechnical systems theory (1/29, 3%) [6], the Nolan stage model (1/29, 3%) [18], an implementation process framework (1/29, 3%) [13], the DeLone and McLean model of information systems success and the Wang and Strong quality framework (1/29, 3%) [58], activity theory (1/29, 3%) [11], design thinking (1/29, 3%) [12], computer-supported cooperative work (1/29, 3%) [7], and the unified theory of acceptance and use of technology (1/29, 3%) [20].

A detailed list of the characteristics of all studies included in this review is provided in Multimedia Appendix 3 [3,6-20,42-45,48-58].

Summary of Facilitators of and Barriers to the Implementation of DSCRs

Overview

Of the 29 studies, 28 (97%) identified 45 facilitators of and 102 barriers to digital implementation. These were then coded into 32 categories that aligned with the NASSS framework domains: 18 contained facilitators and 24 contained barriers (the total is >32 because some categories contained both facilitators and barriers). The most frequent barriers were related to the digital system lacking interoperability, which was found in 10 of the 29 studies (34%). They also related to insufficient funding or financial incentives and high costs of implementation (9/29, 31%), and technology not matching the context of use (9/29, 31%). Most facilitators were associated with building interorganizational trust and collaborative relationships (5/29, 17%); adequate training (5/29, 17%); anticipating, frontloading, and resourcing the work required to clarify information governance (4/29, 14%); skillful leadership enhancing an organization's digital readiness and capacity for change (4/29, 14%); and high usability of the digital system (4/29, 14%).

Regarding the NASSS framework domains, most facilitators were related to the organization (24/45, 53%). This was followed by the adopter system (8/45, 18%), the technology (6/45, 13%), the value proposition (4/45, 9%), the interaction between domains and adaptation over time (2/45, 4%), and the wider context (1/45, 2%). Most barriers were also related to the organization (52/102, 51%). This was followed by the technology (25/102, 24.5%), the wider context (14/102, 13.7%), the value proposition (6/102, 5.9%), the adopter system (4/102, 3.9%), and the interaction between domains and adaptation over time (1/102, 1%). No barriers or facilitators were related to the condition domain.

The categories containing facilitators and barriers were then organized into five broad themes: (1) the legal and institutional context for holding and sharing data and its effect on the ability and willingness to share data, (2) digital readiness and organizational capacity for change, (3) using and sharing recorded information within technical constraints, (4) alignment between care practices and digital recording practices, and (5) differences between what is expected and what is achievable with digital systems.

A summary of how barriers and facilitators identified in each study were mapped to categories, themes, and the NASSS framework domains and subdomains is provided in Multimedia Appendix 4 [3,6-20,42-45,48-56,58]. The 5 themes are summarized in greater detail in the subsequent sections.

Legal and Institutional Context for Holding and Sharing Data and Its Effect on Ability and Willingness to Share Data

A key challenge to DSCR implementation involved information governance concerns about holding and sharing data. These issues arose from vague legislation, market competition, conflicting priorities, poor internal and external coordination, and low cross-organizational trust. Building trust and adequately resourcing digital change facilitated implementation.

Commercial and Regulatory Context in Which Care Providers Operate

A total of 3 studies conducted in the United Kingdom [3,42,50] identified barriers related to market competition among social care provider organizations and digital suppliers and a lack of national regulation and standards. Private sector care providers were concerned that the commercial sensitivity of data could compromise their competitive advantage [42]. Vendor lock-in also occurred, as technology suppliers hesitated to share data with other suppliers [3].

The governance and ethics framework for social care data in the United Kingdom is less developed compared to that for National Health Service (NHS) data. There is no established system for the governance of care home data, which are held by private companies, care regulators, and health and social care provider organizations [42]. This context made data sharing challenging [42]. Despite regulatory progress, councils found new national frameworks inadequate on data and interoperability standards [50].

Interorganizational Trust and Relationships

In 5 studies [3,9,42,48,49], 4 of which were UK based, a lack of trust between providers and other organizations hampered information governance and data sharing. Clinical and health care partners were particularly reluctant to share data with social care [3,9,50] due to misunderstandings about their role and concerns about sharing information with staff who were not registered social workers [50]. Ownership of a large volume of patient data and responsibility for confidentiality also fostered a risk-averse attitude among general practitioners [3].

Four studies conducted in the United Kingdom [3,42,49,50] found that building trust and collaboration between organizations facilitated implementation. Care homes were more willing to share information when they had well-established relationships with local authorities [42]. Scotland's regional "data safe havens," led by trusted partners such as the NHS, academic

XSL•FO
RenderX

institutions, and government agencies, represented a centralized approach to managing, storing, and handling access requests to health care data that encouraged relationships between health and social care provider organizations [42]. They were an example of data being handled respectfully, professionally, and securely [42]. Ambiguous governance frameworks necessitated clarifying information governance requirements and building mutual trust in systems. In the United Kingdom, local authority and provider staff needed to dedicate significant resources upfront to ensure safe data handling processes [3]. Setting up information sharing agreements that specified data flows between organizations could be intensive, involving unexpected time and effort that was often related to building relationships and engaging numerous actors with data sharing plans [3]. Undertaking this work early on in projects facilitated implementation, locating expertise and capacity, and building trust across organizations [3]. Leaders who fostered positive working relationships between decision makers facilitated shared priority setting [49], helping them to circumvent barriers stemming from organizational fragmentation [49].

Organizational Coordination to Clarify Information Governance

In 4 studies [3,9,42,49], a general lack of coordination hindered the clarification of information governance processes needed to implement digital systems. Divisions between and within organizations created siloed data systems, resulting in residents' records being stored in different systems across multiple services [49]. Poor coordination was linked to information governance professionals, who managed personal data for single organizations, lacking the capacity to handle additional responsibilities for cross-organization information governance and data sharing [3]. This issue was compounded by provider leaders' lack of understanding of information governance [3].

Lack of a shared, standardized understanding of information governance and data ownership across organizations also created confusion among staff. Nursing home leaders in the United States [9], for example, raised concerns about transparency and maintaining control of residents' health data that were viewed as belonging to the patients, leading to fear of lawsuits regarding data sharing [9]. In the United Kingdom, there was a lack of shared understanding with confusion about consent, which was related to social care and local government starting from a different position to NHS partners when it came to information sharing [50].

While a lack of organizational coordination was a barrier to implementation, 4 studies [3,10,49,58] identified that prioritizing and adequately resourcing the work required to define information governance was a facilitator to data sharing and in turn improved service quality. In the implementation of information systems across home care services in Italy [58], agreements could be reached on hardware and software once information governance had been properly defined. The synergies resulting from integrating information systems from different organizations then positively affected service quality. In another study, health and social care managers also acknowledged that undertaking considerable work together to agree on what could be shared helped to implement a shared electronic record between nursing and adult social care practitioners [10]. While fostering cross-organizational relationships was important, substantial resources were required to develop and sustain these relationships [49].

Digital Readiness and Organizational Capacity for Change

The importance of investing in the necessary groundwork and anticipating the work involved in digital implementation is linked to an organization's digital readiness and capacity for new technology more generally. Facilitators and barriers within this theme were related to hardware and internet connectivity issues, funding issues in the sector, organizational infrastructure, and resourcing the work required for digital change, including leadership and training.

Hardware and Internet Connectivity Issues

Hardware issues hindered implementation in 4 studies [6,8,11,55] and negatively impacted care quality in 2 studies [6,11]. Problems included a lack of computers and handheld devices for timely patient data documentation in nursing homes [6,8], ergonomic challenges in home care [55], and poor battery lives on portable devices in home care [11,55]. In Australia, residential aged care nurses and care workers relied on memory when portable devices ran out of battery during medication rounds, reducing patient safety [11]. Sharing limited devices in nursing homes also delayed access to updated care plans in a cross-country nursing home study [6]. Hardware issues implied a failure to commit the upfront investment needed to install the hardware required to successfully implement digital systems, reflecting a lack of organizational capacity and readiness [11,55]. Internet connectivity issues were identified as barriers in 5 studies [13,42,48,55,56]. Reliable internet was often deemed essential for digital implementation, and poor connectivity indicated insufficient organizational resources. This was problematic in home care, where mobile internet access was inconsistent [55,56], and in care homes with poor Wi-Fi in old buildings [13,48]. For instance, 18% of care homes in a southeast Scotland project experienced regular internet interruption, and 27% of care homes had limited internet access [42].

Funding Issues in the Sector

In total, 9 studies [3,9,14,16,18,48-50,56] identified insufficient funding or financial incentives and high costs as barriers. Four studies were conducted in the United States [9,14,16,18], 4 studies were conducted in England [3,48-50], and 1 study was conducted in Sweden [56]. In England, short-term funding pushed organizations toward unambitious digital solutions [3]. The financial pressures often forced providers to adopt a short-term view on the finances needed to implement and sustain digital records, constraining the scale of change and preventing it from being embedded [3]. Where funding was available, finding, requesting, and receiving it was not always straightforward [48]. Small care homes in England faced issues such as poor communication from funders, complicated application procedures, and delays in receiving funds [48].

XSL•FO
RenderX

Organizational Infrastructure and Resourcing the Work Required for Digital Change

Barriers related to organizational infrastructure were noted in 5 studies [19,20,48,49,56]. Issues included insufficient ICT and human resources staff [48,56], high senior staff turnover [49], poor internal communication that left staff unaware of implementation [19,20], and inadequate leadership [20]. One English study [3] highlighted that successful implementation required clear planning and resource allocation; for example, phased deployment of resources demonstrated providers' competence in managing digital change, making it easier for them to secure further funding [3].

Four studies [3,10,50,53], 3 of which were conducted in England, noted the importance of skillful leadership in enhancing digital readiness. Identifying leaders with the right skills was crucial for managing large-scale digital projects [3]. The type of leadership required depended on context, with some providers preferring leaders who could balance risk and reward in deploying resources, while others sought leaders who were respected by their peers to help foster engagement among staff [3]. Senior staff functioning as "change agents" also motivated practitioners to review their practices [10]. Successful councils had strong leadership support for digital initiatives [50]. In England, councils successful in implementing data standards and interoperability had strong leadership support [50], with directors of social care, chief information officers, and elected members all prioritizing digital working and integrated care [50].

Adequate Training

Absent or inadequate training was a barrier in 4 studies [6,19,20,48]. Issues included a lack of tailored training [20] and inappropriate content [19]. Conversely, 5 studies identified high-quality training as a facilitator [3,6,11,49,53]. One multicountry study identified both facilitators and barriers across the different contexts [6]. Effective training was tailored to practitioners' skills and tasks [6,53] and included on-the-job and context-specific training [6], ongoing sessions [55], follow-up visits [53], and continued onsite support from suppliers [6,11]. High-quality training that was tailored, targeted, and practical aligned care practices with the new practices required by digital systems.

Using and Sharing Recorded Information Within Technical Constraints

This theme included issues with technical interoperability of digital systems, their level of usability and user-friendliness, and the extent to which they had been appropriately adapted for social care from other settings, which were often acute or primary care.

Interoperability

Interoperability is understood as a technology's capacity to electronically share patient information between different systems and to use the information that has been shared [59]. Lack of interoperability was identified in 10 studies as a barrier to sharing recorded information [3,6,9,15,17,42,49,50,52,54], being reported by 57% of 491 respondents in the US [15] nursing facilities with electronic health records. Care professionals and managers in Finland [52] and senior health and care leaders in England [49] also criticized information systems for not always "communicating" with each other. While providers were adopting digital solutions, these were not necessarily increasing interoperability and risked creating new data silos [15].

In some studies, interoperability barriers were attributed to the multitude of systems used by different organizations. Across 9 nursing homes in Austria [17], managers exchanged information with at least 18 other organizations, most of which were not part of the same electronic health record system. In the United States, while 95.1% (775/815) of nursing homes had electronic medical records, only 45.8% (373/815) had some capability for information exchange with other organizations. The variety and sheer number of systems used by different providers was a concern for 8 (67%) out of 12 staff members in subsequent interviews [9].

In England [50], interoperability issues presented as systems being unable to store identification data such as the NHS number. However, local authorities were often unaware, at the procurement stage, of which digital options could store such information. There was also confusion among councils and suppliers about the possibilities and limitations of NHS number tracing. This was linked to low organizational readiness and capacity, with providers not knowing which technological features they needed when choosing a system [50]. It also related to the downstream value suppliers promised providers in terms of being transparent about what their products could offer [50].

Staff in all 3 nursing homes in a multicountry study also complained that the electronic patient records lacked interoperability and options to adjust features to meet specific needs. This implied a contradiction between customizability and interoperability, with customizable systems more likely to meet care provider needs but less likely to be compatible with other systems than off-the-shelf technology [6].

Usability and User-Friendliness

A total of 8 studies [6,11,12,51,52,54-56] reported barriers related to this theme, 3 of which [52,54,56] were based in Nordic countries. These barriers were more closely associated with using, rather than sharing, recorded information within technical constraints.

A total of 4 studies [51,52,54,55] found problems with the system being slow, crashing, and having unscheduled downtime. Others pointed to features that made staff work routines more inefficient, such as the example from US home care nurses needing to click 22 times to get into each individual's medical record, a cumbersome process that had to start again when they moved on to the next patient [12]. In a multicountry study, care home staff disliked being forced to enter narrative text into the electronic patient record and preferred drop-down menus [6]. An inefficient information retrieval process within an Australian electronic Health record system meant that staff in long-term care facilities had to perform lengthy manual searches to identify wound charts, with the system also failing to alert them if they were duplicating charts that already existed [11].

XSL•FO
RenderX

In 5 studies [8,11,13,20,55], the high usability and usefulness of digital systems facilitated implementation. In 3 cases [13,20,55], systems offered easy access to information, improving the immediacy of care provision and documentation. In some instances, they enhanced the accuracy of care documentation through better information visibility [13,55] or by automating tasks that were previously manual and prone to human error [11]. Digital systems with flagging features also supported resource prioritization and management decisions [13]. These facilitators aided implementation by increasing task efficiency and supporting the knowledge generated or made visible by the technology, thereby improving data accuracy and decision-making.

Adapting Technology From Other Settings

Barriers in 5 studies [6,20,45,54,56] were related to digital systems that had been maladapted from other settings and were consequently deemed inappropriate for social care. In England, social care workers were less likely to perceive health information exchange systems as useful compared to health care workers and experienced issues with the user interface [20]. Staff noted that the system looked unfamiliar compared to other systems they used, as the health information exchange was primarily designed for acute and primary care settings, with little consideration given to social and community services [20].

An Australian study [13] reported successful adaptation of a digital system originally designed for an acute hospital setting to a care home involving staff at all levels, residents, and their relatives that helped to make the product appropriate for the care home setting [13]. This co-designed process facilitated implementation and increased the likelihood of success.

Alignment Between Care Practices and Digital Recording Practices

Overview

Barriers related to digital systems not matching the context of use were identified in 9 studies [3,6,7,12,45,50,54-56]. These barriers referred to misalignments between care practices within the social care sector and recording practices demanded by new digital systems. They included reduced interactions between clients and practitioners, conflicts with preferred data input methods, and exacerbation of existing organizational issues. Staff perceptions of improved care quality increased the likelihood of accepting the technology.

Care Quality and the Relational Nature of Social Care

A total of 5 studies [6,7,9,49,52] highlighted barriers where digital systems decreased the relational nature of social care. Problems arose when care staff experienced disruption to their relational work and viewed the technology as depersonalizing care. In Finland, new information systems increased technical tasks at the expense of relational tasks performed physically close to clients [52]. In England, digital records influenced the nature of the clinical encounter for occupational therapists. By focusing on data collection and adherence to standard procedures, they reduced opportunities for building rapport with clients [7]. Concerns also existed that technology use close to clients was intrusive and reduced care quality [6,55]. Defining the problem as a preference for “high touch” over “high tech,” a US study found that 5 out of 12 nursing home leaders feared that technology might detract from the personal experience they aimed to provide [9].

Technological features, such as prescriptive data fields, also imposed work routines that prioritized clinical data and processes. In a Scottish study, data systems in care homes promoted a task-oriented culture over resident-focused care [45]. Prescripted data fields limited the recording of social and emotional activities and care provision, leading to an overly clinical focus in the data [45].

Only 1 study found that a digital system aligned well with the relational nature of social work, facilitating implementation [13]. In Australia, an aged care ecosystem that was co-designed with staff and residents allowed care workers to multitask and spend more time with residents. This saved time for staff and improved care quality, encouraging acceptance of the system [13]. Managers noted that the technology provided prompts for tasks such as repositioning residents, better aligning care with resident needs [13]. In England, 2 studies found that perceived care quality improvements increased staff acceptance of digital systems [3,20]. Demonstrating the technology’s value to different professionals helped staff “buy into” digital change [3]. Administrative staff adopted technology for time-saving benefits, while practitioners focused on its impact on care [3]. Perceived improvements to patient safety also increased the likelihood of adopting digital systems [20].

Preexisting Organizational Problems

A total of 2 studies [12,19] identified barriers where digital systems exacerbated preexisting organizational problems, such as the numerous communication channels in home care organizations [12]. The lack of standardization required nurses to adapt to various communication methods, for example, contacting physicians through primary care nurses or by fax [12]. They often only discovered that their request had reached doctors through changes made to patients’ medicines [12]. Rather than standardizing processes, the new digital system added more communication channels. While this issue presented as inappropriate technology, it was rooted in inefficient work routines that predated the technology’s introduction.

Conflicts Between Data Recording Practices and Digital Systems

A total of 4 UK-based studies [3,7,42,50] identified barriers due to conflicts between data recording practices preferred by care providers and those permitted by digital systems. The lack of systematic data collection in care homes made it difficult to capture the complexity of care for individuals with multiple conditions and high support needs [42]. Frontline practitioners preferred narrative text input, while digital systems often emphasized coded data entry [3,50]. In 1 study, social workers entering free-text information sometimes included data about third parties without consent [3]. Such issues were linked to a lack of understanding about data quality in social care [50], requiring retraining on the importance of proper data collection and recording practices [3,50]. An English study found that conflicts between recording preferences and the recording permitted by digital systems were due to a mismatch between digital care records and occupational therapy concerns [7].

XSL•FO
RenderX

Therapists had to recode their interventions to fit the system's structure, suggesting that the technology did not align with sector needs, rather than indicating poor recording practices.

Differences Between What Is Expected and What Is Achievable With Digital Systems

Overview

The final theme related to the gap between organizational expectations and realistic achievements with digital systems. Guidance on available technology was often inadequate, and care providers lacked internal consensus about the technology's capabilities and what they wanted to gain from implementation. Creating a shared digital vision and adopting digital systems as part of wider cultural changes facilitated implementation.

Guidance on the Technology Available

Insufficient guidance on available technology was a barrier in 1 study in English care homes [48]. The overwhelming number of suppliers created an "unregulated tech product maze," making it difficult to choose the best option and avoid paying for unsuitable technology [48]. Care homes criticized NHS England's "assured suppliers list" of DSCR suppliers, which was introduced to aid decision-making [48]. Although suppliers on the list met a set of standards, some care homes complained that suppliers did not meet their needs and requirements, while others reported poor experiences with suppliers on the list and found themselves locked into contracts despite consistent software malfunctions [48].

A Shared Digital Vision

Creating a shared vision for collectively understanding the technology involved building organizational consensus on its potential while remaining realistic about its limitations. A total of 4 studies [3,17,20,52] found that care provider staff disagreed about the purpose of digital systems, and awareness of potential benefits for care delivery was low. There were tensions between 2 distinct staff groups with different expectations [3]. One group represented a technical and managerial culture that often initiated digital change projects and was primarily interested in the information captured by digital systems. The other was a clinical culture that was concerned with how technology could help deliver care and was more skeptical of changes to practice that lacked certain types of evidence [3]. Managers were generally more positive about implementation but lacked awareness of some of its negative effects on employees' work [52]. Staff anticipated unrealistic benefits and were often unaware of the technology's value [17,20].

A total of 2 studies found facilitators to creating a shared vision [13,53]. They highlighted the importance of co-design and inclusive implementation by gathering suggestions from staff, residents, and their relatives [13] or by conducting monitoring based on staff's feedback to system developers [53]. Involving different groups as partners in the process helped envision a digital system that benefited everyone [13]. Professionals praised comprehensive and continuous communication that helped them make sense of a new service, with information delivered through multiple channels to reach as many employees as possible, including shift workers [53].

Implementing Digital Change as a Cultural Change

Framing digital implementation as a cultural change program facilitated success in 3 studies [3,10,13]. In an Australian care home, co-designing the system, establishing a shared vision across the workforce, and providing training and feedback loops instigated a culture change that improved service delivery and problem-solving [13]. In England, barriers to scaling digital changes in health and social care were mitigated by treating them as part of a wider technology-supported clinical transformation program, rather than an ICT project [3], or as part of a larger cultural change program to improve administrative efficiency [10].

Summary of the Intended and Actual Outcomes

Outcomes of digital implementation, either intended or actual, were identified in 17 studies [7-11,13,17,19,20,48,50-52,54,55,57,58], although they were the focus of only 1 study [57]. The full details of the benefits and outcomes are provided in Table 2.

A total of 3 studies [9,17,48] identified the outcomes that participants hoped to achieve through adopting digital systems. Improved information accessibility, information sharing, and quality of care records were identified in 2 studies [17,48], making them the most frequent intended outcomes. Examples of the improved quality of records included more complete and readily available patient-related information and less documents being lost during patient transitions between different institutions [17]. Improved efficiency [17] and time savings [48] were identified as intended outcomes in 1 study.

Three studies [9,48,50], 2 of which were based in England [48,50], cited poor awareness about the benefits of digital systems for social care or concerns that they would not benefit the sector. In England, information sharing initiatives were often focused on health care and hospitals, with less attention paid to the potential benefits for councils or social care [50]. This made it difficult for social care staff and care home residents to see the benefits that digital systems could bring [48].

A total of 13 studies identified positive outcomes realized through digital record implementation [7,8,10,11,13,19,20,50-52,55,57,58]. Improved efficiency was the most frequent actual outcome (8/13, 62%) [10,11,13,19,20,50-52], achieved through the automation of previously manual processes [11], reduced duplication of procedures [20], and the increased availability [19] and immediacy [13] of information improving decision-making and care planning. These outcomes were associated with increased staff capacity [50] and productivity [52]. Impacts on efficiency were not always clear. In 1 study [11], while automatic data entry in patient records was beneficial, the system did not completely align with work processes, and staff needed to record some data twice.

XSL•FO
RenderX

Table 2. The intended and actual outcomes of digital social care record implementation (N=17).

Theme	Studies, n (%)
Intended outcomes	
Improved quality of data records	2 (12)
Improved information sharing	2 (12)
Improved information accessibility	2 (12)
Improved efficiency	1 (6)
Time savings	1 (6)
Improved care quality or planning	1 (6)
Improved communication or collaboration	1 (6)
Improved information accuracy	1 (6)
Space savings (less paper)	1 (6)
Actual outcomes	
Improved efficiency	8 (47)
Perceived time savings	7 (41)
Improved information accessibility	5 (29)
Workarounds (viewed negatively)	4 (24)
Improved communication or collaboration	3 (18)
Improved information security and risk management	3 (18)
Additional time burdens	3 (18)
Improved care quality or planning	2 (12)
Increased face-to-face work with patients	2 (12)
Improved information sharing	2 (12)
Improved information accuracy	2 (12)
Improved transparency and accountability	2 (12)
Increased staff or patient satisfaction	1 (6)
Workarounds (viewed positively)	1 (6)
Decreased communication or collaboration	1 (6)
Decreased efficiency	1 (6)
Decreased care quality	1 (6)
Decreased face-to-face work with patients	1 (6)
Lack of financial benefits	1 (6)
Rationing care documentation	1 (6)

Perceived time savings were reported in 7 studies, although the findings were not conclusive [10,13,19,20,50,52,58]. Some studies reported staff spending less time retrieving and documenting information for decision-making [13,58] and chasing other organizations for patients' whereabouts [50]. One study found time savings of up to 45 minutes for long-term care staff when completing medication reconciliation [19]. However, 2 studies found time savings in some areas and additional time burdens in others [10,52]. In 1 case, disagreements between managers and their staff arose regarding whether the digital system created time savings [52]. Managers and employees agreed that moving from phone calls to digital messaging had freed up staff time for other tasks [52]. However, employees felt that the new tasks, such as responding to clients through messages, required extra time. This additional time was not always recognized by management, nor were additional resources provided [52].

A total of 3 studies [8,13,19,20,52] found that digital systems made information more accessible. In one case, this enabled person-centered care, with easily accessible information on individual backgrounds helping staff to "see the person first and the diagnosis second" [13]. In another case, improved visibility of information facilitated medication tracking and therefore supported patient safety [19]. A total of 3 studies also highlighted improved communication and collaboration [10,52,55] and improved information security and risk management [51,52,58]. Electronic information sharing

XSL•FO
RenderX

improved partnership working, enhancing collaboration and increasing the timeliness, efficiency, and quality of care [10,52,55]. Improved information security and risk management were linked to secure information transfer and storage [51,52,58], better client monitoring [52], and increased data accuracy [51,58].

Workarounds, identified in 5 studies [7,9,11,54,55], were the most common negative outcome. Workarounds involve the implementation, by end users, of temporary practices or behaviors to overcome the limitations of a technological system [60]. Staff developed workarounds for various reasons. These included circumventing the system to share health data with residents [9] and accessing case-based information [54]. While workarounds could be beneficial [7] and support task completion [54], they also threatened data security [54].

Discussion

Principal Findings

This study investigated what is known about the implementation of digital records in adult social care settings. The literature was diverse in terms of the type of digital system, setting, and use case studied. Most of the studies used a qualitative design (14/29, 48%), particularly those looking at facilitators and barriers, how digital systems affect work routines, and potential strategies to improve digital systems. Studies were most frequently based within the United Kingdom (10/29, 34%).

Most studies focused on facilitators of and barriers to digital implementation. Many facilitators and barriers were interlinked and associated with multiple NASSS framework domains, which compounded the complexity of implementing digital systems. The 5 themes we identified using the NASSS framework are particularly complex areas that require more active management and consideration when implementing DSCRs in social care contexts.

While our findings suggest that implementing digital systems is an inherently complex process, this review did identify some strategies to manage complexity, which could constitute "good practice." In terms of digital readiness and organizational capacity for change, high-quality training was found to increase implementation success. Where training was tailored, practical, and ongoing, it helped align care practices with new practices required by the technology, thereby increasing employees' ability and willingness to adopt and continue to use the system. Although high-quality training depended on care provider leaders anticipating the financial resources needed, it seems a worthwhile investment for successful digital implementation. This finding echoes the results from a previous scoping review, which highlighted training as a key factor influencing the use of electronic information systems [23].

Implementing digital systems as part of wider cultural change projects also addressed multifaceted complexity. An example of this was the project in which implementation was co-designed with staff [13]. This approach enabled a shared vision of the technology to be created across the care home among residents and staff at different levels. The sense of ownership this instilled addressed complexity in the adopter system domain, with all users more likely to support the technology and view it as "business as usual." Co-design also addressed complexity in the technology domain, with the digital system more likely to align with the needs and practices of its user group. While incorporating digital implementation as part of broader transformation required significant resources, where there was sufficient organizational readiness and capacity for comprehensive rollout, implementation seemed to have greater potential for sustainability, scaling, and spread.

Complexity related to data sharing and information governance seemed to be more difficult to address. Trust and relationship building across organizations could help establish data sharing agreements at a localized level and therefore address complexity within the organization domain. However, fundamental barriers were associated with complexity around regulations and standards in the wider context domain, over which care providers had no direct control. Until there is primary or secondary legislative change, the governance and regulatory context will continue to impede cross-organizational data sharing efforts.

Although 17 studies identified intended or actual outcomes, they more often focused on identifying facilitators of and barriers to implementation. Improved efficiency, accuracy, and time savings were the most common positive outcomes realized through digital adoption, while workarounds and additional time burdens were the most frequently cited negative outcomes. Some of the positive outcomes reflect the results presented in the review by Greenstock [22], which also found efficiency and productivity to be a benefit of DSCRs. However, the limited detail in outcome reporting and variations in the extent to which different benefits are observed suggest that this topic would benefit from future research. Specifically, there seems to be a need for studies that quantify outcomes and pay greater attention to the necessary conditions for positive benefits to be realized.

Limitations of Studies

Most studies (20/29, 69%) lacked a clear theoretical or methodological framework. This meant it was often unclear which type of digital system or record was being implemented as well as the context, setting, and use case. While studies mentioned >100 facilitators and barriers to implementation, they did not provide any objective parameters or measures to assess how they impact implementation or social care practices. This hinders a more comprehensive comparison between the barriers and facilitators.

Some digital systems were simply described as ICT, electronic digital systems or health information technology [18,55,57], or digitalization or digital change generally [3,48,52], without definitions of these terms. Some studies appeared to use the same vocabulary to describe different systems. However, this was difficult to determine as most studies (20/29, 69%) did not specify their target population clearly. Many studies also lacked detail regarding care settings and other relevant information, which limited the possibility of performing more comprehensive comparative analysis. Future studies should pay greater attention to how they report which digital systems were implemented, the target population for the system, the setting, and the roles of the professionals involved to facilitate comparisons between

XSL•FO
RenderX

studies. Standardized reporting guidelines, such as the template for intervention description and replication checklist and guide [61], may facilitate describing digital projects or systems.

Of the 29 studies, only 2 (7%) included carers or service users as respondents, while most studies included staff (n=18, 62%) or managers (n=11, 38%). Future research may therefore benefit from incorporating the perspectives of people drawing on care to cover this gap in the literature.

Although studies included in our review mention the potential impacts of DSCRs, none provide quantifiable parameters to estimate such impacts, such as potential time savings or cost-effectiveness metrics. New studies that are appropriately designed to measure such outcomes are needed to fill this important knowledge gap in the literature on DSCR implementation.

Methodological Limitations

Due to the prevalence of qualitative designs and a lack of clear theoretical or methodological frameworks among the studies reviewed, we used the NASSS framework as a structured approach to categorizing and interpreting heterogeneous data. As this was a rapid review, the framework served as a tool to guide our data interpretation and triangulation, especially given the large number of barriers and facilitators and the varied ways these issues were described across different studies. For example, it directed our analysis of hardware and internet connectivity issues. While the studies reviewed often attributed these to technological problems, the framework enabled us to trace the associated complexity back to the organization domain, with care providers lacking the awareness, readiness, and capacity to prepare for digital implementation and adopt appropriate systems.

However, the NASSS framework carried some limitations for our analysis. No facilitators or barriers were associated with the condition domain. While the framework was developed for both health and social care, the focus of this domain on comorbidities and clinical aspects of a patient's condition may be more appropriate for health care technologies. For social care technologies, it may be more useful to approach the condition domain in terms of whether digital systems are appropriate for particular groups of clients, such as older people or people with learning disabilities, rather than specific illnesses. Alternatively, the lack of relevance of the condition domain may reflect limited attention to diversity and inclusion considerations within the studies reviewed. Another limitation of the NASSS framework was related to the final domain (ie, interaction between domains and adaptation over time). As most of the complexity we identified was multifaceted, we found it more useful to iteratively analyze the interactions between domains instead of restricting them to 1 domain. Rather than viewing complexity as belonging to separate domains, we suggest using this final domain to provide an overarching perspective of how complexity constantly intersects and interacts across domains at every stage of digital implementation.

Considering the rapid nature of this scoping review, we simplified some review procedures, such as screening and full-text assessment, which always carry the risk of missing relevant studies. To minimize such risks, all review procedures were undertaken by researchers who are experts in systematic review methods and social care research. As with every review, the choice of databases is also a limitation, as relevant studies may have been uniquely indexed in databases that were not included. However, our research was able to identify all relevant studies that were suggested by experts in the field. Moreover, we performed a comprehensive gray literature search to reduce the likelihood of missing key studies.

Despite the limitations, we believe that our review provides a comprehensive picture of the state of the literature on DSCRs. It builds on 4 previous reviews, which, when taken together, captured the literature about digital records until the end of 2017 [2,21-23]. Our review has updated and added to these findings, covering both academic and gray literature up until 2023 and using a robust theoretical framework to draw out complexity in terms of sustainability, scaling, spread, nonadoption, and abandonment of digital care records.

Conclusions

Our findings suggest that the implementation of digital care records is particularly complex due to the lack of a common language and consensus about what DSCRs should look like as well as expected outcomes and impacts. This is reflected in the scientific literature, which often lacks operationalization of key constructs and robust study designs. To be successful, implementation should consider complexity, while studies should use a robust theoretical framework and use mixed methods or quantitative designs where appropriate. We also suggest that future studies define the target population, consider gathering data on the experiences of carers and service users, and focus on digital care records specifically being used in social care, such as DSCRs.

Acknowledgments

The authors would like to acknowledge the support of the Social Care Rapid Evaluation Team coordinator Jo Ward, as well as Andra Fry, the London School of Economics and Political Science library information specialist. This study is funded by the National Institute for Health and Care Research (NIHR) Health and Social Care Delivery Research Programme (NIHR163639) as part of the Social Care Rapid Evaluation Team (NIHR153673), and open access fees were covered by the grant. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

XSL•FO
RenderX

Data Availability

No new data were generated for this scoping review. All data were obtained from publicly available published sources and are included in this published article and its supplementary information files. We have included the search strategy in Multimedia Appendix 2, summaries of the extracted articles in Multimedia Appendix 3, and the thematic analysis process in Multimedia Appendix 4 to provide transparency on extractions and analysis.

Authors' Contributions

MS contributed to screening and data collection and was responsible for data analysis and preparation of the first draft of the manuscript. WS-R contributed to study conception and design, was responsible for screening and data collection, and contributed to data analysis and preparation of some sections of the manuscript. JM conceived and designed the research; oversaw the screening, data collection, and data analysis; and contributed to drafting of the manuscript. KP and SDG were responsible for developing and running the search strategy and drafting the paragraphs on the search methods. MB and CL contributed to the study conception and design. WS-R, JM, MB, JW, CL, NF, and NS critically reviewed the manuscript, adding important intellectual content. All authors read and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist.
[PDF File (Adobe PDF File), 162 KB - aging_v8i1e60107_app1.pdf]

Multimedia Appendix 2
Search Strategy, August 14, 2023.
[DOCX File , 32 KB - aging_v8i1e60107_app2.docx]

Multimedia Appendix 3
Detailed characteristics of the included studies (N=29).
[DOCX File , 43 KB - aging_v8i1e60107_app3.docx]

Multimedia Appendix 4
Thematic analysis of the facilitators and barriers using the nonadoption, abandonment, scale-up, spread, and sustainability framework.
[DOC File , 64 KB - aging_v8i1e60107_app4.doc]

References

1. Bongaarts J. Human population growth and the demographic transition. Philos Trans R Soc Lond B Biol Sci 2009 Oct 27;364(1532):2985-2990 [FREE Full text] [doi: 10.1098/rstb.2009.0137] [Medline: 19770150]
2. Kruse CS, Mileski M, Vijaykumar AG, Viswanathan SV, Suskandla U, Chidambaram Y. Impact of electronic health records on long-term care facilities: systematic review. JMIR Med Inform 2017 Sep 29;5(3):e35 [FREE Full text] [doi: 10.2196/medinform.7958] [Medline: 28963091]
3. Maguire D, Evans H, Honeyman M, Omojomolo D. Digital change in health and social care. The King's Fund. 2018 Jun 20. URL: https://www.kingsfund.org.uk/insight-and-analysis/reports/digital-change-health-social-care [accessed 2025-02-05]
4. Health at a glance 2023. Organisation for Economic Co-operation and Development. 2023 Nov 7. URL: https://www.oecd.org/en/publications/2023/11/health-at-a-glance-2023_e04f8239.html [accessed 2025-02-05]
5. Slawomirski L, Lindner L, de Bienassis K, Haywood P, Hashiguchi TC, Steentjes M, et al. Progress on implementing and using electronic health record systems: developments in OECD countries as of 2021. Organisation for Economic Co-operation and Development. 2023 Sep 13. URL: https://tinyurl.com/4fspm6ry [accessed 2025-02-05]
6. Shiells K, Diaz Baquero AA, Štěpánková O, Holmerová I. Staff perspectives on the usability of electronic patient records for planning and delivering dementia care in nursing homes: a multiple case study. BMC Med Inform Decis Mak 2020 Jul 13;20(1):159 [FREE Full text] [doi: 10.1186/s12911-020-01160-8] [Medline: 32660474]
7. Sugarhood P, Rouncefield M. Occupational therapists and digital care records: exploring health and social care integration through ethnography - summary of key findings. UK Occupational Researcher Foundation. 2018. URL: https://www.rcot.co.uk/sites/default/files/UKOTRF%20Key%20findings_Dr%20Paul%20Sugarhood.pdf [accessed 2025-02-05]
8. Ausserhofer D, Favez L, Simon M, Zúñiga F. Electronic health record use in Swiss nursing homes and its association with implicit rationing of nursing care documentation: multicenter cross-sectional survey study. JMIR Med Inform 2021 Mar 02;9(3):e22974 [FREE Full text] [doi: 10.2196/22974] [Medline: 33650983]

XSL•FO
RenderX

9. Powell KR, Deroche CB, Alexander GL. Health data sharing in US nursing homes: a mixed methods study. J Am Med Dir Assoc 2021 May;22(5):1052-1059 [FREE Full text] [doi: 10.1016/j.jamda.2020.02.009] [Medline: 32224261]
10. Chester H, Hughes J, Bowns I, Abendstern M, Davies S, Challis D. Electronic information sharing between nursing and adult social care practitioners in separate locations: a mixed-methods case study. J Long Term Care 2021 Jan 12:1-11 [FREE Full text] [doi: 10.31389/jltc.16]
11. Qian S, Yu P, Bhattacherjee A. Contradictions in information technology mediated work in long-term care: an activity theoretic ethnographic study. Int J Nurs Stud 2019 Oct;98:9-18. [doi: 10.1016/j.ijnurstu.2019.05.017] [Medline: 31238234]
12. Persson J, Larsson R, Erlingsdottir G, Rydenfält C. How digital systems are used in Swedish home care nursing practice: a qualitative interview study to identify challenges and opportunities. Comput Inform Nurs 2023 Aug 01;41(8):586-594 [FREE Full text] [doi: 10.1097/CIN.0000000000001006] [Medline: 36649078]
13. Bail K, Gibson D, Hind A, Strickland K, Paterson C, Merrick E, et al. 'It enables the carers to see the person first': qualitative evaluation of point-of-care digital management system in residential aged care. J Clin Nurs 2023 Jan;32(1-2):174-190 [FREE Full text] [doi: 10.1111/jocn.16285] [Medline: 35285557]
14. Felix H, Dayama N, Morris ME, Pradhan R, Bradway C. Organizational characteristics and the adoption of electronic health records among nursing homes in one southern state. J Appl Gerontol 2021 May 21;40(5):481-488. [doi: 10.1177/0733464820906685] [Medline: 32081058]
15. Vest JR, Jung HY, Wiley KJ, Kooreman H, Pettit L, Unruh MA. Adoption of health information technology among US nursing facilities. J Am Med Dir Assoc 2019 Aug;20(8):995-1000.e4 [FREE Full text] [doi: 10.1016/j.jamda.2018.11.002] [Medline: 30579920]
16. Lin SC, Tunalilar O. Rapid adoption of electronic health record and health information exchange among assisted living communities, 2010-2018. J Am Med Inform Assoc 2022 Apr 13;29(5):953-957 [FREE Full text] [doi: 10.1093/jamia/ocac021] [Medline: 35187569]
17. Schaller M, Dornauer V, Hackl WO, Lechleitner G, Überegger M, Ammenwerth E. Implementing national electronic health records in nursing homes in Tyrol: a nursing management perspective. Stud Health Technol Inform 2020 Jun 23;271:240-247. [doi: 10.3233/SHTI200102] [Medline: 32578569]
18. Alexander GL, Liu J, Powell KR, Stone PW. Examining structural disparities in US nursing homes: national survey of health information technology maturity. JMIR Aging 2022 Aug 23;5(3):e37482 [FREE Full text] [doi: 10.2196/37482] [Medline: 35998030]
19. Aird T, Holditch C, Culgin S, Vanderheyden M, Rutledge G, Encinareal C, et al. An analysis of a novel Canadian pilot health information exchange to improve transitions between hospital and long-term care/skilled nursing facility. J Integr Care 2022 Sep 19;30(4):399-412. [doi: 10.1108/jica-03-2022-0022]
20. Watkinson F, Dharmayat KI, Mastellos N. A mixed-method service evaluation of health information exchange in England: technology acceptance and barriers and facilitators to adoption. BMC Health Serv Res 2021 Jul 25;21(1):737 [FREE Full text] [doi: 10.1186/s12913-021-06771-z] [Medline: 34303379]
21. Kruse CS, Marquez G, Nelson D, Palomares O. The use of health information exchange to augment patient handoff in long-term care: a systematic review. Appl Clin Inform 2018 Oct;9(4):752-771 [FREE Full text] [doi: 10.1055/s-0038-1670651] [Medline: 30282094]
22. Greenstock J. Benefits of digital social care records. The Strategy Unit. 2021 Aug. URL: https://tinyurl.com/3p87s34y [accessed 2025-02-05]
23. Ylönen K. The use of Electronic Information Systems in social work. A scoping review of the empirical articles published between 2000 and 2019. Eur J Soc Work 2022 Apr 18;26(3):575-588. [doi: 10.1080/13691457.2022.2064433]
24. Peek N, Sujan M, Scott P. Digital health and care: emerging from pandemic times. BMJ Health Care Inform 2023 Oct;30(1):e100861 [FREE Full text] [doi: 10.1136/bmjhci-2023-100861] [Medline: 37832967]
25. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
26. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
27. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69 [FREE Full text] [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
28. Peters MD, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
29. Tricco AC, Langlois EV, Straus SE. Rapid reviews to strengthen health policy and systems: a practical guide. World Health Organization. 2017. URL: https://iris.who.int/bitstream/handle/10665/258698/9789241512763-eng.pdf [accessed 2025-02-05]
30. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]

XSL•FO
RenderX

31. Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ, et al. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. Syst Rev 2021 Jan 26;10(1):39 [FREE Full text] [doi: 10.1186/s13643-020-01542-z] [Medline: 33499930]
32. Ribeiro WS, Malley J, Baginsky M, Larkins C. Good practice in the implementation of digital social care records (DSCRs): a rapid scoping review. Open Science Framework Home. 2023 Aug 9. URL: https://osf.io/d5pcj/ [accessed 2024-04-25]
33. Falconer J. Removing duplicates from an EndNote library. Library, Archives & Open Research Services at the London School of Hygiene & Tropical Medicine. 2018 Dec 7. URL: https://blogs.lshtm.ac.uk/library/2018/12/07/removing-duplicates-from-an-endnote-library/ [accessed 2025-02-05]
34. Covidence systematic review software. Veritas Health Innovation. URL: https://www.covidence.org [accessed 2025-01-22]
35. Local Government Association homepage. Local Government Association. URL: https://www.local.gov.uk [accessed 2023-08-11]
36. The King's Fund homepage. The King's Fund. URL: https://www.kingsfund.org.uk [accessed 2023-08-11]
37. The Social Care Institute for Excellence homepage. The Social Care Institute for Excellence. URL: https://www.scie.org.uk [accessed 2023-08-11]
38. The Centre for Care homepage. The Centre for Care. URL: https://centreforcare.ac.uk [accessed 2023-08-11]
39. Digital Care Hub homepage. Digital Care Hub. URL: https://www.digitalcarehub.co.uk [accessed 2023-08-11]
40. TEC Services Association homepage. TEC Services Association. URL: https://www.tsa-voice.org.uk [accessed 2023-08-11]
41. Tricco AC, Antony J, Zarin W, Strifler L, Ghassemi M, Ivory J, et al. A scoping review of rapid review methods. BMC Med 2015 Sep 16;13:224 [FREE Full text] [doi: 10.1186/s12916-015-0465-6] [Medline: 26377409]
42. Johnston L, Koikkalainen H, Anderson L, Lapok P, Lawson A, Shenkin SD. Foundation level barriers to the widespread adoption of digital solutions by care homes: insights from three Scottish studies. Int J Environ Res Public Health 2022 Jun 16;19(12):7407 [FREE Full text] [doi: 10.3390/ijerph19127407] [Medline: 35742667]
43. Johnston L, Kokkalainen H, Anderson L, Lapok P, Lowson A, Shenkin S. Digital and data readiness of care homes for older people in South East Scotland. Edinburghdinburgh Napier University. 2022 Apr. URL: https://edwebcontent.ed.ac.uk/sites/default/files/atoms/files/ladder_study_2022.pdf [accessed 2025-02-05]
44. Johnston L, Hockley J, Henderson DA, Shenkin SD. The development of a care home data platform in Scotland: insights from the care home innovation partnership, Lothian. medRxiv Preprint posted online on August 20, 2020 [FREE Full text] [doi: 10.1101/2020.08.17.20176503]
45. Shenkin SD, Johnston L, Hockley J, Henderson DA. Developing a care home data platform in Scotland: a mixed methods study of data routinely collected in care homes. Age Ageing 2022 Dec 05;51(12):afac265 [FREE Full text] [doi: 10.1093/ageing/afac265] [Medline: 36580390]
46. Greenhalgh T, Abimbola S. The NASSS framework - a synthesis of multiple theories of technology implementation. Stud Health Technol Inform 2019 Jul 30;263:193-204. [doi: 10.3233/SHTI190123] [Medline: 31411163]
47. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
48. Supporting small care home providers and their managers on their journey towards digitalisation. National Institute for Health and Care Research. 2022 Nov 30. URL: https://tinyurl.com/bddvrtmf [accessed 2025-02-05]
49. Ingram E, Cooper S, Beardon S, Körner K, McDonald HI, Hogarth S, et al. Barriers and facilitators of use of analytics for strategic health and care decision-making: a qualitative study of senior health and care leaders' perspectives. BMJ Open 2022 Feb 17;12(2):e055504 [FREE Full text] [doi: 10.1136/bmjopen-2021-055504] [Medline: 35177457]
50. Local government social care data standards and interoperability. Local Government Association. 2019 Sep 20. URL: https://www.local.gov.uk/publications/local-government-social-care-data-standards-and-interoperability [accessed 2025-02-05]
51. Gatawa T, Swift N, Gibson S, Lindsay-Walters F. Improving social care through digital care planning: an evaluation of the PASSsystem. Social Care Institute for Excellence and everyLIFE Technologies. 2019 Mar. URL: https://everylifetechnologies.com/wp-content/uploads/2020/11/everyLIFE-SCIE-full-report-compressed.pdf [accessed 2025-02-05]
52. Kaihlanen AM, Laukka E, Nadav J, Närvänen J, Saukkonen P, Koivisto J, et al. The effects of digitalisation on health and social care work: a qualitative descriptive study of the perceptions of professionals and managers. BMC Health Serv Res 2023 Jun 30;23(1):714 [FREE Full text] [doi: 10.1186/s12913-023-09730-y] [Medline: 37386423]
53. Nadav J, Kaihlanen AM, Kujala S, Laukka E, Hilama P, Koivisto J, et al. How to implement digital services in a way that they integrate into routine work: qualitative interview study among health and social care professionals. J Med Internet Res 2021 Dec 01;23(12):e31668 [FREE Full text] [doi: 10.2196/31668] [Medline: 34855610]
54. Salovaara S, Ylönen K. Client information systems' support for case-based social work: experiences of Finnish social workers. Nordic Soc Work Res 2021 Nov 10;12(3):364-378. [doi: 10.1080/2156857x.2021.1999847]
55. Ibrahim S, Donelle L, Regan S, Sidani S. A qualitative content analysis of nurses' comfort and employment of workarounds with electronic documentation systems in home care practice. Can J Nurs Res 2020 Mar 14;52(1):31-44. [doi: 10.1177/0844562119855509] [Medline: 31200603]
56. Rydenfält C, Persson J, Erlingsdottir G, Johansson G. eHealth services in the near and distant future in Swedish home care nursing. Comput Inform Nurs 2019 Jul;37(7):366-372. [doi: 10.1097/CIN.0000000000000536] [Medline: 31135467]

XSL•FO
RenderX

57. Yoshimoto T, Nawa N, Uemura M, Sakano T, Fujiwara T. The impact of interprofessional communication through ICT on health outcomes of older adults receiving home care in Japan - a retrospective cohort study. J Gen Fam Med 2022 Jul 20;23(4):233-240 [FREE Full text] [doi: 10.1002/jgf2.534] [Medline: 35800645]
58. Bianchi P, Trimigno M. How does information system success come about in inter-organizational networks of public services? Public Money Manag 2019 Sep 25;41(3):236-245. [doi: 10.1080/09540962.2019.1665361]
59. Bates DW, Samal L. Interoperability: what is it, how can we make it work for clinicians, and how should we measure it in the future? Health Serv Res 2018 Oct;53(5):3270-3277 [FREE Full text] [doi: 10.1111/1475-6773.12852] [Medline: 29527678]
60. Cresswell KM, Worth A, Sheikh A. Integration of a nationally procured electronic health record system into user work practices. BMC Med Inform Decis Mak 2012 Mar 08;12:15 [FREE Full text] [doi: 10.1186/1472-6947-12-15] [Medline: 22400978]
61. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014 Mar 07;348:g1687 [FREE Full text] [doi: 10.1136/bmj.g1687] [Medline: 24609605]

Abbreviations

DSCR: digital social care record
ICT: information communication technology
NASSS: nonadoption, abandonment, scale-up, spread, and sustainability
NHS: National Health Service
PRISMA-S: Preferred Reporting Items for Systematic reviews and Meta-Analyses literature search extension
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by H LaMonica; submitted 03.05.24; peer-reviewed by Y Chu, MP Gagnon; comments to author 07.07.24; revised version received 08.08.24; accepted 13.11.24; published 14.02.25.

Please cite as:
Snow M, Silva-Ribeiro W, Baginsky M, Di Giorgio S, Farrelly N, Larkins C, Poole K, Steils N, Westwood J, Malley J
Best Practices for Implementing Electronic Care Records in Adult Social Care: Rapid Scoping Review
JMIR Aging 2025;8:e60107
URL: https://aging.jmir.org/2025/1/e60107
doi:10.2196/60107
PMID:

©Martha Snow, Wagner Silva-Ribeiro, Mary Baginsky, Sonya Di Giorgio, Nicola Farrelly, Cath Larkins, Karen Poole, Nicole Steils, Joanne Westwood, Juliette Malley. Originally published in JMIR Aging (https://aging.jmir.org), 14.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Implementation of a Web-Based Program for Advance Care Planning and Evaluation of its Complexity With the Nonadoption, Abandonment, Scale-Up, Spread, And Sustainability (NASSS) Framework: Qualitative Evaluation Study

Doris van der Smissen[1*], PhD; Maud A Schreijer[1*], MSc; Lisette J E W C van Gemert-Pijnen[2], PhD; Rudolf M Verdaasdonk[3], PhD; Agnes van der Heide[1], MD, PhD; Ida J Korfage[1*], PhD; Judith A C Rietjens[1,4*], PhD

[1]Department of Public Health, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

[2]Psychology, Health & Technology, Faculty of Behavioural, Management and Social Sciences, University of Twente, Enschede, The Netherlands

[3]Health Technology Implementation, Faculty of Science and Technology, University of Twente, Enschede, The Netherlands

[4]Department of Design, Organisation and Strategy, Faculty of Industrial Design Engineering, Delft University of Technology, Delft, The Netherlands

[*]these authors contributed equally

Corresponding Author:
Ida J Korfage, PhD
Department of Public Health
Erasmus MC, University Medical Center Rotterdam
Dr. Molewaterplein 40
Rotterdam, 3015 GD
The Netherlands
Phone: 31 107038460
Email: i.korfage@erasmusmc.nl

Abstract

Background: The implementation of eHealth applications often fails. The NASSS (nonadoption, abandonment, scale-up, spread, and sustainability) framework aims to identify complexities in eHealth applications; the more complex, the more risk of implementation failure.

Objective: This study aimed to analyze the implementation of the web-based advance care planning (ACP) program "Explore Your Preferences for Treatment and Care" using the NASSS framework.

Methods: The NASSS framework enables a systematic approach to improve the implementation of eHealth tools. It is aimed at generating a rich and situated analysis of complexities in multiple domains, based on thematic analysis of existing and newly collected data. It also aims at supporting individuals and organizations to handle these complexities. We used 6 of 7 domains of the NASSS framework (ie, condition, technology, value proposition, adopters, external context, and embedding and adaptation over time) leaving out "organization," and analyzed the multimodal dataset of a web-based ACP program, its development and evaluation, including peer-reviewed publications, notes of stakeholder group meetings, and interviews with stakeholders.

Results: This study showed that the web-based ACP program uses straightforward technology, is embedded in a well-established web-based health platform, and in general appears to generate a positive value for stakeholders. A complexity is the rather broad target population of the program. A potential complexity considers the limited insight into the extent to which health care professionals adopt the program. Awareness of the relevance of the web-based ACP program may still be improved among target populations of ACP and among health care professionals. Furthermore, the program may especially appeal to those who value individual autonomy, self-management, and an explicit and direct communicative approach.

Conclusions: Relatively few complexities were identified considering the implementation of the web-based ACP program "Explore Your Preferences for Treatment and Care." The program is evidence-based, freestanding, and well-maintained, with straightforward, well-understood technology. The program is expected to generate a positive value for different stakeholders. Complexities include the broad target population of the program and sociocultural factors. People with limited digital literacy may need support to use the program. Its uptake might be improved by increasing awareness of ACP and the program among a wider population of potential users and among health care professionals. Addressing these issues may guide future use and sustainability of the program.

XSL•FO
RenderX

(JMIR Aging 2025;8:e49507) doi:10.2196/49507

KEYWORDS

eHealth; web-based intervention; implementation; sustainability; advance care planning; NASSS framework; nonadoption, abandonment, scale-up, spread, and sustainability framework; health communication; patient education; patient-centered care

Introduction

eHealth tools can be useful for health promotion, enhancing self-management skills for people with long-term conditions, and patient-physician communication [1]. The web-based program "Explore Your Preferences for Treatment and Care" [2] is aimed at supporting users in engaging in advance care planning (ACP). ACP is a communication process that enables persons to think about their goals and preferences for future treatment and care, to discuss these with their relatives or health care professionals, and to record these if appropriate [3]. The web-based ACP program guides users through 3 steps of ACP—exploration, discussion, and recording of preferences for future treatment and care (Textbox 1). A before and after study showed that the web-based ACP program increased ACP engagement among persons with a chronic disease and was perceived as usable, attractive, and comprehensible [4]. The program was embedded as a decision aid in the Dutch platform "Thuisarts" [5] in April 2020 (English version "GPinfo website" is in development). This platform is owned and hosted by the Dutch College of General Practitioners and provides evidence-based health-related information to the public.

Textbox 1. Main characteristics of the web-based advance care planning program "Explore Your Preferences for Treatment and Care" [2,4].

Content

- Information about advance care planning (ACP).
- Thinking about values and quality of life.
- Communication about preferences with relatives and health care professionals.
- Appointing a health care representative.
- Recording of preferences in an advance directive.
- Reviewing the advance directive.
- References to information about specific diseases, patient organizations, and peer support opportunities.
- The content was based on a scoping review that explored the content, feasibility, and effectiveness of web-based ACP programs [6]; an interview study that identified information needs for web-based ACP of patients with chronic diseases and their relatives [7]; and meetings with a stakeholder group containing relevant stakeholders for the web-based ACP program including patients, relatives, and patient organizations [4].

Structure

- Interactive program; users can watch videos and click on additional information, users are asked questions regarding ACP and can save a document containing their responses.
- Stepwise approach to guide users through the ACP process.
- Embedment in information platform (Thuisarts website [English version: GPinfo website]) which is hosted and owned by the Dutch College of General Practitioners, and provides evidence-based health information.
- Accessible for free.
- Hyperlinks to external websites.
- Text-to-speech option.

Development

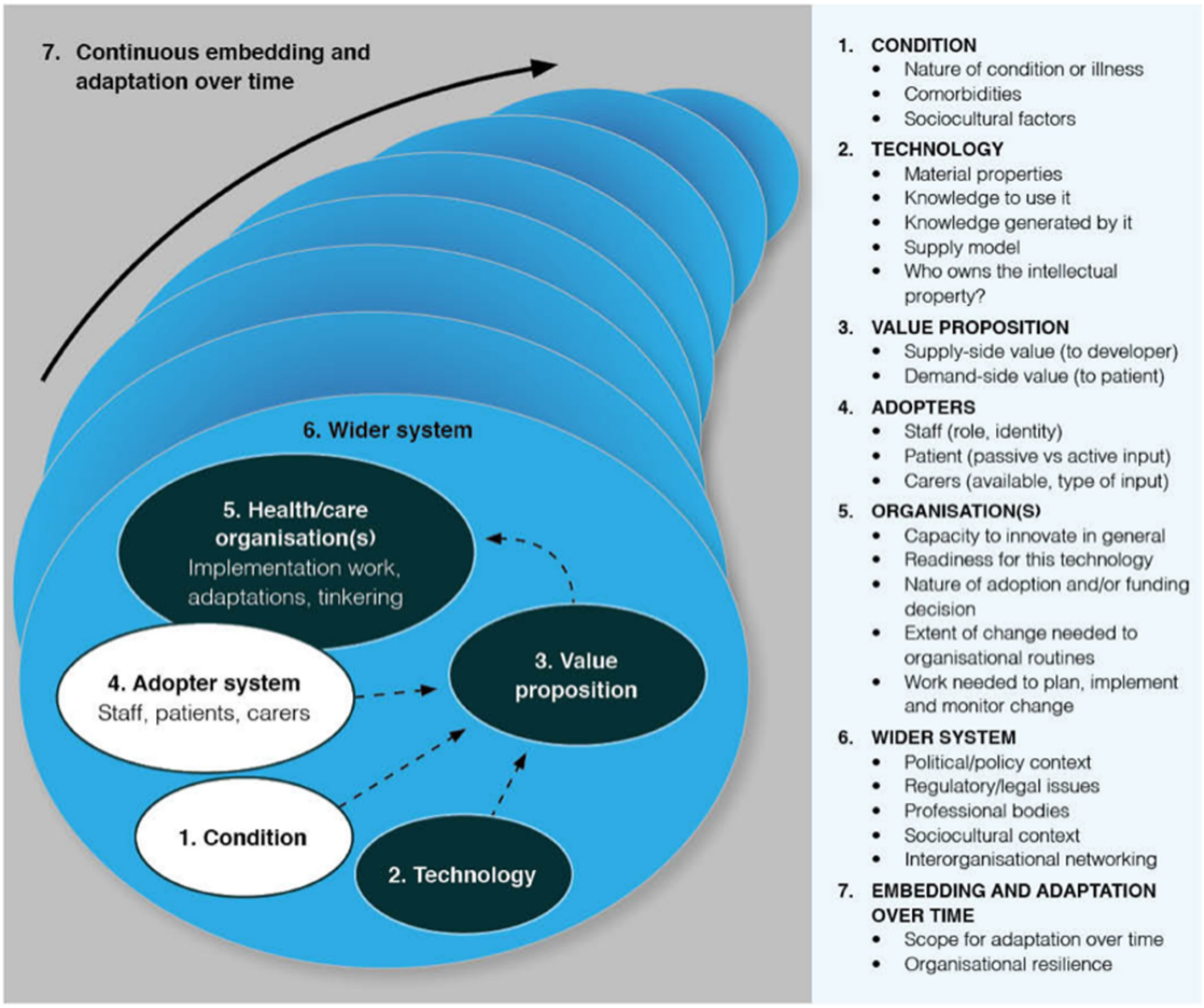
- Developed by researchers (DvdS, IJK, JACR, and AvdH) with expertise in shared decision-making, care at the end of life, and eHealth, in collaboration with stakeholder group including 1 patient, 2 relatives, representatives of the Dutch College of General Practitioners (Nederlands Huisartsen Genootschap), the Dutch Association for Kidney Patients (Nierpatiënten Vereniging Nederland), the Dutch Patient Association (Nederlandse Patiëntenvereniging), Agora (organization to promote the palliative approach), 1 expert in health communication of the Nivel (Netherlands institute for health services research), 1 expert in eHealth of the University of Twente (LvG-P), and 1 representative of Vital Innovators, an organization that conducts Social Return of Investment analyses.
- The development was funded by the Netherlands Health Organisation for Health Research and Development.

Although many eHealth tools are developed, when they are implemented (ie, made available to end users), their sustainable implementation often fails [8]. To identify issues that may hamper the implementation of eHealth tools early, the NASSS (nonadoption, abandonment, scale-up, spread, and sustainability) framework has been developed by Prof T Greenhalgh and her team at Oxford University. The framework is aimed to encourage timely reflecting eHealth applications in health care

XSL•FO
RenderX

and to systematically explore the chances of successful implementation of these eHealth applications by identifying complexities and reducing them [8]. NASSS stands for "nonadoption, abandonment, scale-up, spread, and sustainability over time," which addresses the 5 possible reasons for eHealth implementation to fail [8]. The NASSS framework enables a systematic approach to explore complexities of eHealth tools regarding their implementation; the more complex, the more risk of uptake failure [8]. NASSS focuses at 7 key domains; condition, technology, value proposition, adopters, organization, external context, and embedding and adaptation over time (Figure 1) [8]. A domain can be classified as "simple" when it is straightforward, predictable, and only contains a few components (as in making a sandwich), "complicated" when it contains multiple interacting components (as in building a rocket), and "complex" when it is dynamic, unpredictable, and not easily disaggregated into constituent components (as in raising a child) [8,9]. The NASSS framework postulates that technologies with most domains classified as "simple" are more likely to be successfully implemented than technologies where most domains are classified as "complex" or "complicated." Complexities in different domains can be interdependent, for example, when organizations have difficulty to adopt an eHealth technology, this may complicate adoption of the eHealth tool by staff as well [10,11]. We used the NASSS framework to identify complexities that may hamper sustainable implementation of the web-based ACP program.

Figure 1. The nonadoption, abandonment, scale-up, spread, and sustainability (NASSS) framework [8] published under Creative Commons Attribution 4.0 International License [12].



Methods

Data Analysis

We evaluated the complexity of the web-based ACP program "Explore Your Preferences for Treatment and Care" [2] using the NASSS framework, following the approach of Abimbola et al [11] who also conducted an ex post (retrospective) thematic analysis.

First, we developed a multimodal dataset consisting of different existing, relevant data sources (Table 1). The dataset includes peer-reviewed and published studies describing the development and evaluation of the program [4,6,7], notes of stakeholder group meetings, and a social return on investment analysis in which the social costs (input) and benefits for stakeholders (outcomes) of the program were mapped.

XSL•FO
RenderX

Next, 2 researchers (DvdS an MAS) developed a data extraction form based on the nonadoption, abandonment, scale-up, spread, and sustainability complexity assessment tool (NASSS-CAT; IRIHS group); the NASSS CAT-LONG interview guide [13]. They used this form to extract relevant pieces of text from the data sources per NASSS domain, that is, (1) condition, (2) technology, (3) value proposition, (4) adopters, (5; previously 6) external context, (6; previously 7) embedding and adaptation over time. We considered the original domain 5, "organization," not applicable because the web-based ACP program is not implemented in a specific health care organization and the hosting organization, the Dutch College of General Practitioners, did not have to change their routines to embed the program. Next, DvdS and MAS together with the wider research group conducted an initial thematic analysis of the extracted information to identify whether sufficient information per domain was available. They conducted 3 additional semistructured interviews to complete the dataset (1) with a representative of the Dutch college of general practitioners (the program host), (2) a general practitioner, and (3) a representative of the organization that funded the development of the web-based ACP program. The interview questions were based on the NASSS CAT-LONG interview guide [13]. Interviewees provided written informed consent for participating in an online interview in which questions were asked about attitudes toward the web-based ACP program related to the domains of the NASSS framework. Interviews were audio-recorded and transcribed verbatim.

Subsequently, DvdS and MAS together with the wider research group conducted a thematic analysis of the final dataset, using the NASSS framework as a lens, leading to a thematic description of complexities per domain. We used relevant ACP literature to further underpin the findings. The conclusion whether a domain is simple, complicated, or complex was based on discussion and eventually consensus among the members of the research group, where we weighed the potential impact and relevance of the different complexities.

Complexities in each of the 6 domains were classified as "simple," "complicated," or "complex." Per domain, all relevant aspects were taken into consideration. An overall classification as "simple" does not require a total absence of room for improvement for a domain; however, most of its elements should be considered as simple. Based on the multimodal dataset and the classification in Multimedia Appendix 1 [8,9], DvdS and MAS prepared a classification per domain, which was discussed during meetings with all researchers (MAS, DvdS, JACR, IJK, and AvdH), until consensus was reached.

XSL•FO
RenderX

Table 1. Overview of the multimodal dataset used to analyze complexities of the web-based advance care planning program "Explore Your Preferences for Treatment and Care."

Sources of information	NASSS[a] framework domains					
	Condition	Technology	Value proposition	Adopters	External context	Embedding and adaptation over time
Results and data from studies:						
Scoping review [6]: • Aim: to determine the feasibility and effectiveness of 11 web-based ACP[b] programs in 27 studies.			✓[c]		✓	
Interview study [7]: • Aim: identifying information needs for web-based ACP. • Participants: 9 patients with a chronic disease and 7 relatives.			✓	✓		
Pilot study [4]: • Aim: assessing the usability and feasibility of the web-based ACP program. • Participants: 3 health care professionals (2 general practitioners and 1 vascular surgeon) and 6 patients with a chronic disease.		✓	✓	✓		
Evaluation study [4]: • Aim: evaluation of ACP engagement (before using the web-based program vs 2 months after completion), usability, and users' satisfaction (including comprehensibility). • Participants: 147 patients with a chronic disease.		✓	✓	✓		
Other available resources:						
Notes of stakeholder group meetings: • Including researchers, patient organizations, patients, and relatives. • Decisions were made concerning the program (eg, definition of target group and content of the program).	✓			✓	✓	✓
SROI[d] analysis: • The social costs and benefits for stakeholders, and mapped investments (input) and revenues (outcomes) of the program. • The SROI methodology followed 9 steps that led to the complete cost-benefit model and the final calculated SROI ratio. • Interviews were conducted with health care organizations, stakeholder group, patients and relatives, and a health insurance company.			✓	✓		✓
Content of the web-based program: • Content, structure, layout, or assembly. • Page with additional information (information about target group and purpose of ACP).	✓	✓				
Platform of the web-based program ("Thuisarts"): • Requirements of the program, consultations with the Thuisarts website. • Guidelines of the Thuisarts website (eg, not saving personal data and maintenance web-based program).	✓	✓		✓	✓	✓

XSL•FO
RenderX

Sources of information	NASSS[a] framework domains					
	Condition	Technology	Value proposition	Adopters	External context	Embedding and adaptation over time
Publicity web-based program: • Interviews with researchers and involved patients and relatives in magazines and the news. • News articles.				✓	✓	✓
• Number of visits to web-based ACP program				✓	✓	
Financial information: • Funder of the development of the web-based ACP program. • Thuisarts website.			✓			
Relevant literature	✓	✓	✓	✓	✓	✓
Newly collected data						
Interviews with 3 additional key stakeholders	✓	✓	✓	✓	✓	✓

[a]NASSS: nonadoption, abandonment, scale-up, spread, and sustainability.

[b]ACP: advance care planning.

[c]✓: included in the dataset.

[d]SROI: social return on investment.

Ethical Considerations

This study was approved by the Medical Research Ethics Committee of the Erasmus MC, University Medical Center Rotterdam on October 21, 2019 (MEC-2019-0590), confirming that the rules laid down in the Medical Research Involving Human Subjects Act do not apply to the interviews as conducted in this study. The study conforms with the International Committee of Medical Journal Editors' recommendations for the conduct, reporting, editing, and publication and for the protection of research participants. The interview participants were provided with information about the aim and content of the interview. They provided written informed consent. Interview participants did not receive compensation. The authors confirm that all patient or personal identifiers have been removed or disguised so the persons described are not identifiable and cannot be identified through the details of the story.

Results

The Condition or Illness

The domain "condition or illness" considers to what extent the condition or illness is well-characterized, well-understood, and predictable and to what extent the condition or illness is influenced by comorbidities or sociocultural factors.

Target Group

The target group of the web-based ACP program is described as follows: "people can engage in ACP when they are healthy, but also when they become older or (chronically) ill, after having an accident or when nearing the end of life" [2]. The target group is hence not strictly demarcated, which is in line with recent recommendations [3,14,15]. During the development of the program, demarcating the target population was a recurring issue with some stakeholders promoting a focus on people with advanced illness and others promoting a broader scope. The target group ambiguity is also visible at the Thuisarts website, where the program is embedded: internal links to the program are included in pages about topics such as end-of-life care and dementia.

Sociocultural Factors Affecting Engagement in ACP

The web-based ACP program is aimed at individuals who are interested in ACP, regardless of their religion or culture. The program contains images and videos of both men (n=5) and women (n=3), persons with various ethnic backgrounds, and a wide range of ages (middle age to older adults) [2]. One video in the program includes a reference to a religious belief.

Nevertheless, the program is currently only available in the Dutch language and was only evaluated among Dutch-speaking persons. Multilingual support is not offered, which limits the accessibility of the program to non–Dutch-speaking individuals. The program fosters autonomy and self-management, and uses an explicit and direct communicative style [2]. This approach may be less appealing to people who consider the role of the family to be central in medical decision-making, and to those who prefer a less direct communication style [16]. This may limit the appeal of the program for those who consider the role of the family to be central in medical decision-making, and to those who prefer a less direct communication style [16].

Classification and Conclusion

The "condition or illness" domain is classified as complex given the rather undemarcated target population of the web-based ACP program and because the program may not address the ACP preferences of all potentially eligible persons.

XSL•FO
RenderX

The Technology

The domain "the technology" encompasses a description of the web-based ACP program and the Thuisarts website, as well as user experiences with the technology.

The Web-Based ACP Program Embedded in Thuisarts Website

The web-based ACP program guides users through the process of ACP in three steps, that are (1) thinking about preferences for future treatment and care, (2) discussing these preferences with relatives and health care professionals, and (3) recording these preferences. Users are asked questions, and the answers to these can be saved in a document and printed. Users of the program are recommended to use this document during a conversation with a relative or a consultation with a health care professional. Users do not need to complete the program at once; they can access it at any preferred moment. The program is freely accessible without the need to register and data of users are not saved. Content maintenance of the program occurs through critical review of the program every 3 years by a stakeholder group of patients, relatives, health care professionals, and patient organizations.

The web-based ACP program is embedded in the Thuisarts website, a web-based platform containing information about health and disease, based on evidence-based guidelines (English version "GPinfo website" is in development). The platform is owned and maintained by the Dutch College of General Practitioners [5]. Approximately two-thirds of the Dutch population are familiar with the platform [17], and it has 6.6 million monthly visitors [18]. The content is written on the B1 level (easily readable and concise), and is offered in standardized formats [19], reading text out loud is enabled, and information is summarized in videos with subtitles. Every year, experts test whether the platform has the recommended level for web-based information (level AA) [20] and identified problems are solved [21].

User Experiences With the Technology

A qualitative pilot study showed that the web-based ACP program was acceptable and feasible for 6 interviewed patients with a chronic disease [4]. In an evaluation study among 147 members of an online research portal (who are expected to have at least some digital skills), including people with levels of low, middle, and high health literacy, the program was perceived as user-friendly (mean score of 70, SD 13, scale 0-100), attractive (mean 3.8, SD 0.7, scale 1-5), and comprehensible (mean 4.2, SD 0.6, scale 1-5) [4]. To support people with limited health- or digital literacy in using the program, people can watch videos and use a text-to-speech option to read the text aloud [4]. Furthermore, a clear and simple structure is used and the number of topics covered is not too large [4]. The texts are written on the B1 level (easily readable and concise) [19].

Classification and Conclusion

The "technology" domain is classified as simple because the web-based ACP program is freestanding and well-maintained, with straightforward, well-understood technology. An evaluation study showed that the program is perceived as user-friendly and comprehensible. However, people with limited health literacy or digital literacy might require help to use the program.

The Value Proposition

The "value proposition" domain concerns whether a new technology is worth developing and for whom it may generate value. It includes demand-side value (value for the users) as well as supply-side value (value for the developers).

Value for Stakeholders

In total, 3 interviewed health care professionals indicated they considered the program as valuable to patients [4] and an evaluation study showed that after use of the web-based ACP program, people contemplated about ACP more often and indicated to feel more ready for ACP [4]. They considered the program to be user-friendly and were satisfied with the program [4]. This is in line with findings from a scoping review, which concluded that participants of web-based ACP programs generally consider these programs as easy to use, not burdensome, and feasible [6]. This review also indicated that web-based ACP programs are effective to improve ACP knowledge, communication about ACP, and documentation of ACP among program users [6].

Regarding the value for health care professionals, the funder of the web-based ACP program project and a general practitioner indicated in the additional interviews to consider the program as helpful to start ACP for health care professionals:

> *Some basic things are in there [in the program] which could be discussed: (…) like what is important to someone? That can help as a guidance in your conversation.* [general practitioner]

Costs

The web-based ACP program was developed in the context of a publicly funded research program and has no commercial purpose. The Dutch College of General Practitioners invests in long-term maintenance of the Thuisarts website and they will maintain the web-based ACP program technically as well.

Cost and Benefit Analysis

The web-based ACP program is freely accessible. The Thuisarts website is not aimed at making a profit. It is financed by the Dutch College of General Practitioners and receives no money from companies seeking financial gain, such as pharmaceutical companies. To better understand the relationship between the total investment (costs) and the expected social effects (benefits) of the web-based ACP program, a so-called social return on investment analysis was conducted [22]. This analysis showed that each investment in terms of money, time, and effort, is expected to generate a social return of approximately 1.7. This means that the program is expected to have a positive value for patients, relatives, health care professionals, and the health care system in terms of money or societal value, such as increased quality of life.

Classification and Conclusion

The "value proposition" domain is classified as simple because the technology is expected to generate a positive value for stakeholders (patients, relatives, health care professionals, and

XSL•FO
RenderX

the health care system) while the benefits of the program are expected to exceed the costs. The Dutch College of General Practitioners invests in long-term maintenance of the Thuisarts website and they will maintain the web-based ACP program technically as well.

The Adopters

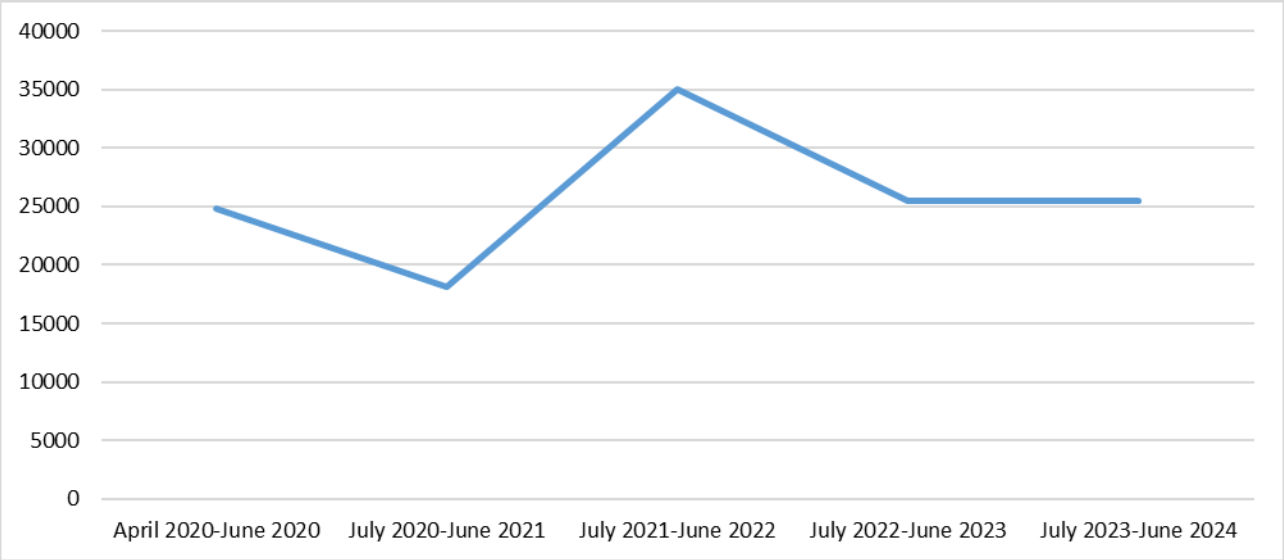
The domain "adopters" focuses on the adoption and continued use of the technology by the groups of people who are intended to actually use the program or to refer potential users to the program, and what they need to use the program or to refer to it. The potential adopters of the web-based ACP program include patients, relatives, health care professionals, and the public in general.

Patients and the Public

Patients with chronic disease and their relatives were asked about their information needs for ACP in a qualitative interview study [7]. They indicated the need for guidance in ACP, information about their disease and care, and information on how to communicate their preferences to relatives and health care professionals [7]. These needs were taken into account in the development of the web-based ACP program. In an evaluation study, users of the web-based ACP program indicated that the program helped them to consider ACP [4]. Users were generally satisfied with the amount of information in the program [4]. A few users mentioned that it was confronting to complete the program because of the difficulty of the topic and a large amount of information. In the first 3 months after its launch (April 1-June 30, 2020) the program was visited 24,849 times. From July 1, 2020, to June 30, 2021, it was visited 18,160 times; from July 1, 2021, to June 30, 2022, it was visited 35,026 times; from July 1, 2022, to June 30, 2023, it was visited 25,423 times; and from July 1, 2023, to June 30, 2024, it was visited 25,510 times. In total, it was visited 128,986 times (Figure 2).

More than 40 medical and patient organizations refer to the web-based ACP program on their websites. However, as described in the domain "condition or illness," not everyone may be interested in engaging in ACP because of a preference to focus on their health and life in the present [23] or valuation of other principles than those central to the web-based ACP program, for example, valuing a central role of the family in medical decision making [16]. While the web-based ACP program is considered useful, it may require substantial commitment and effort of the user in terms of reflecting on their values and preferences for treatment and care.

Figure 2. Visits to the web-based advance care planning program "Explore Your Preferences for Treatment and Care" from its launch (April 1, 2020) to June 30, 2024.



Health Care Professionals

The web-based ACP program recommends participants to involve their health care professionals in the ACP process, for instance, to explain the medical situation, discuss patients' preferences, or file an advance directive. It is known that most people access ACP information through health care professionals [24] and expect that the professional initiates ACP [25]. The Thuisarts website is a frequently used platform among health care professionals; 90% (1751 of 1946) of general practitioners are familiar with it and inform patients about it, and 73% (1421 of 1946) use it for information provision during their consultations [26]. Health care professionals were included in the stakeholder group and provided input during the development of the web-based ACP program. However, opinions about the web-based ACP program of health care professionals have not been extensively evaluated [4]. In total, 3 interviewed health care professionals considered the program to be of value [4]. However, several other studies have described that health care professionals may experience barriers in the initiation of ACP. ACP may be time-consuming, health care professionals report to lack the skills in initiating sensitive ACP conversations [25,27], and they may fear taking away patients' hope [25,27]. Whether the web-based ACP program may support health care professionals to overcome these barriers remains to be studied.

Classification and Conclusion

The "adopters" domain is classified as simple. However, we identified a few complexities. While the web-based ACP

XSL•FO
RenderX

program is potentially useful, it may require substantial commitment and effort of users. Furthermore, it remains to be studied whether the program can support health care professionals to overcome well-known barriers to ACP such as lack of time and fear of taking away patients' hope.

The External Context

The "external context" domain encompasses (1) the policy and political climate, (2) professional organizations and patient organizations, (3) the regulatory context, and (4) the economic context.

Policy and Political Climate

ACP is generally encouraged by organizations of health care professionals in the Netherlands, such as general practitioners, home care organizations, and hospitals. Many initiatives are taken to implement ACP in health care practice such as the development of ACP tools and manuals, websites for the public as well as health care professionals such as Palliaweb website, and ACP modules for electronic patient files [3,28-30]. In addition, the Dutch Ministry of Health, Welfare and Sport has spent €51 million (US $52.6 million) in the period 2014-2020 on research on palliative care including ACP [30] and initiated another research program on ACP in 2022 [31]. Furthermore, the Ministry has organized a public information campaign in 2022 to raise awareness for the importance of timely talking about death and dying [32,33].

Professional Organizations and Patient Organizations

Several patient organizations were involved in the development, evaluation, and implementation of the web-based ACP program. More than 40 organizations, including The Netherlands Patient Federation and the Dutch Association for Kidney Patients, as well as general practitioners and care organizations, referred to the web-based ACP program on their websites and in newsletters.

Regulatory Context

The regulatory context of ACP in general in the Netherlands is supportive regarding ACP. Several professional bodies have developed and provided guidelines for ACP [34], toolkits [35], and manuals for doctors [29,36]. Some of these specifically refer to the web-based ACP program [29,36]. Some materials are specifically developed for patients [37]. However, these ACP materials have different target groups, for instance, health care providers [29,34,36], patients in the palliative phase or at the end of life [3,14,15], or broader populations [3], and therefore it may not always be clear to whom they apply. Regarding the regulatory context of the web-based ACP program, users' privacy is protected in compliance with data security and privacy requirements, by not requiring any log-in procedure and not recording any details of users.

Economic Context

The reimbursement of ACP for health care professionals is not clearly regulated.

Classification and Conclusion

The "external context" domain is classified as simple. Professional bodies are supportive toward the web-based ACP program and ACP in general. We found some complexities, such as different available manuals and tools for ACP containing different recommendations, and the unavailability of a treatment code for health care professionals. However, these external conditions are not likely to complicate the adoption of the web-based ACP program to a large extent.

Embedding and Adaptation Over Time

Overview

The scope of the Thuisarts website is expanding. Information provision was focused on primary care in the past, but currently also includes information on secondary care. The Thuisarts website, including the web-based ACP program, is likely to be maintained in the future, with both technical and content updates. Content maintenance of the program occurs through a critical review of the program every 3 years by a stakeholder group of patients, relatives, health care professionals, and patient organizations. They will review compliance with (evolving) data security regulations, privacy requirements, and policies. If necessary, the program will be adapted. At this time, no changes in data protection regulations or policies are expected in the near future that would affect the program.

Our evaluation study showed that in general, participants had a positive user experience and were satisfied with the web-based ACP program [4]. To further improve this, we are systematically collecting informal feedback on the program and visitor numbers. These will be discussed during each 3-yearly review of the program by the stakeholder group. If necessary, they will make recommendations to adapt the program. To do so, the web-based ACP program will be adapted using an iterative design strategy of ongoing collecting informal feedback and the editorial team of the Thuisarts website adapting the program accordingly. Based on the feedback, user experience and satisfaction are aimed to be enhanced by improving the usability and feasibility of the web-based program.

Classification and Conclusion

We classify the "embedding and adaptation over time" domain as simple, since it is unlikely that the web-based ACP program as embedded in the Thuisarts website or its value are significantly going to change in the next years.

Discussion

Principal Results

To analyze the implementation of the web-based ACP program "Explore Your Preferences for Treatment and Care," the NASSS framework was used to identify complexities within 6 domains. The analysis revealed no or little complexity in the domains "technology," "value proposition," "external context," and "embedding and adaptation over time." The program is evidence-based, freestanding, and well-maintained, with straightforward, well-understood technology. The program is expected to generate a positive value for different stakeholders as well as for the Thuisarts website. The analyses revealed some complexities in the "condition" domain, including the broad and rather undemarcated target population of ACP and the web-based ACP program and sociocultural factors that may

XSL•FO
RenderX

limit users' engagement. In the "adopters" domain, complexity entails that the program requires substantial commitment and effort by users with respect to reflecting on their preferences for future treatment and care. Some people with limited digital literacy may need support to use the program. Furthermore, it is yet unstudied to what extent health care professionals adopt the web-based ACP program. The program is embedded in the general practitioners' platform "Thuisarts," which is frequently used by general practitioners and the general public. The fact that the web-based ACP program is not embedded in a specific health care organization simplified its implementation, since no alignment regarding content or layout was required. On the other hand, embedding the program in the routine practice of a health care organization, could have boosted its use. Finally, we found some complexity in the "external context" domain. While the political and policy climate, regulatory context, professional organizations, and patient organizations are generally supportive regarding ACP, a national guideline for ACP and a treatment code to reimburse ACP could further increase willingness and ability of health care providers to engage in ACP. Overall, the analysis showed that the program has good potential for sustainable implementation, as it is expected to be continuously used in practice [8,38,39], will be updated if necessary, and will be maintained long-term.

Comparison With Previous Work

The results of this study showed that the web-based ACP program "Explore Your Preferences for Treatment and Care" has relatively few complexities regarding the implementation. This may have several reasons. First, the user-centered design of the program included an extensive preparatory phase, including a scoping review to learn from existing web-based ACP programs [6], identification of the needs of patients with chronic disease and their relatives for web-based ACP in an interview study [7], and close collaboration with stakeholders including patients and relatives. Second, during the development of the program, we paid due attention to sustainable implementation in an existing, well-used platform—"Thuisarts." The team of the Thuisarts website has expertise in developing information for the public with different levels of health literacy. It is remarkable that other web-based ACP programs often seem to have not paid due attention to implementation. Three recent reviews about web-based ACP programs [6,40,41] showed that studies have mainly focused on the development of web-based ACP programs and evaluation of their feasibility, usability, acceptability, and effectiveness, but most did not address their implementation. This is in line with reviews about implementing eHealth in general showing that research has mainly focused on the content and evaluation of eHealth tools, but less so took into account the external context and sustainable adoption in health care systems [8,10].

The results of the current study also indicated that health care professionals are an important group of adopters for the web-based ACP program. The web-based ACP program, following existing guidelines, recommends an active role of health care professionals in the ACP process of patients, for instance, to discuss diagnosis, prognosis, and relevant treatment and care options with the patient, and to document the ACP process in the medical file [3,4]. In the literature, several barriers are described that health care professionals may experience in the initiation of ACP, such as lack of time, lack of skills in conducting ACP conversations, and fear of taking away patients' hope [25,27]. In the development of the web-based ACP program, the role of health care professionals was to a limited extent taken into account. First, the program is embedded in the general practitioners' platform "Thuisarts," which is frequently used by general practitioners [26] and the general public [17,18]. Furthermore, health care professionals were included in the stakeholder group and provided input during the development of the web-based ACP program. In the program, the user is encouraged to discuss and share their preferences with their health care professionals at several points [2]. However, the actual use of the web-based ACP program among health care professionals and their views and experiences regarding the program, are unstudied. Given the barriers that health care professionals may experience in the initiation of ACP [25,27], a patient-centered and community approach of ACP is important, in which patients can also take the initiative to engage in ACP themselves. This is central to the program, as the program supports the initiation of ACP, self-management of users, and encourages conversation with health care professionals about values, goals, and preferences.

The results indicated complexities considering the broad target population of ACP in the web-based ACP program. This complexity is a reflection of the evolving concept of ACP in the literature. Initially, ACP was conceptualized to be used for patients in the palliative phase and at the end of life [3,14,15]. More recently, an international consensus study on ACP recommended that individuals can engage in ACP in any stage of life, but that ACP can be more targeted when individuals' health condition worsens or when their age increases [3]. This evolving concept of ACP aligns with the evolving concept of palliative care, whose definition has widened from a sole focus on the end of life to also including chronic illnesses [42] and a shift from in-hospital to community-based care [43]. Reaching the community with high-quality information and guidance about ACP through the Thuisarts website may therefore fit these developments.

Strengths and Limitations

A strength of this study is our systematic approach of evaluating the web-based ACP program using the NASSS framework. It provided insight in its complexities that might hamper its implementation. Furthermore, we had a large dataset including publications about the development and evaluation of the web-based ACP program, notes of stakeholder meetings, and interviews with the program host, a general practitioner, and a funding agency representative. Insight into the complexities may be used to improve future sustainable implementation of the program, and may be used to improve other web-based programs as well.

A limitation is the start of the NASSS evaluation after the tool was developed and embedded in the Thuisarts website, which has limited the opportunity to adapt the program during its development. Some of the authors (DvdS, JACR, AvdH, and IJK) were the developers of the web-based ACP program, and therefore they could provide in-depth insight in the

XSL•FO
RenderX

developmental process. However, although the authors aimed to analyze the implementation of the web-based ACP program objectively, there can be a bias in the evaluation.

Recommendations to Improve Complexities of the Web-Based Advance Care Planning Program

To enhance sustainable implementation, we provide several recommendations to improve the identified complexities of the web-based ACP program.

First, to make the program available in other languages than Dutch. This may further enhance the accessibility of the program to people with different languages. For instance, an English version of the program could be placed on the GPinfo website, which is the English version of the Thuisarts website that is still in development. The web-based ACP program may be translated to other languages as well, to enhance its accessibility to people who are not proficient in Dutch. To improve cultural inclusiveness, future research should explore the needs of people from various cultural backgrounds with regard to ACP and if desired, the program could be culturally adapted or different versions of the program could be developed.

Second, to enhance accessibility of the web-based ACP program to persons with very limited health or digital literacy. The program was developed to assist individuals with limited health or digital literacy, by offering readable texts, text-to-speech options, and videos. However, individuals with very limited health or digital literacy may still experience barriers to use the program. To enhance accessibility of the program for these users, recommendations to complete the program together with a relative could be added to the program. Furthermore, a PDF version of the web-based ACP program could be developed, that could be printed by a relative or health care professional, to make the program also available "offline."

Third, to inform health care professionals about the variance in readiness for ACP among patients. Not everyone may want to engage in ACP [44], for instance, due to different cultural backgrounds and different needs to consider treatment and care preferences. In addition, we recommend to inform and support health care professionals in how to conduct ACP with their patients. The web-based ACP program may help to overcome lack of time, since health care professionals can refer to the program, patients can prepare themselves for an ACP discussion on beforehand, and then the discussion with the health care professional can be scheduled.

Fourth, the concept of ACP as an ongoing process of considering preferences (also including healthy individuals) [3], could be embedded in national guidelines, and be routinely integrated in health care practice to reach a broader target group for ACP. National guidelines and availability of a treatment code may lower barriers in ACP for health care professionals, and may consequently lead to increased engagement with ACP for patients. Websites aimed at the general public could refer to the web-based ACP program, for instance, pages about shared decision-making and pages within the Thuisarts website about chronic diseases. Furthermore, we recommend to conduct public awareness campaigns about ACP (including web-based ACP programs), as these have been found to enhance awareness of ACP [45]. These may also help to enhance accessibility of ACP for people beyond the medical health care setting.

Finally, to ensure adaptation of the program to evolving data protection requirements. During critical review of the program, that will take place every 3 years by a stakeholder group of patients, relatives, health care professionals, and patient organizations, compliance with (evolving) data security regulations, privacy requirements, and policies should be carefully reviewed. If necessary, the program should be adapted. In addition, the strategy of the Thuisarts website, the host of the web-based ACP program, is aimed at adapting the information on the Thuisarts website to the evidence-based guidelines for general practitioners as developed by the Dutch College of General Practitioners. When guidelines change, the information at the Thuisarts website is adapted accordingly by their editorial team. Furthermore, the strategy of the Thuisarts website focuses on collaborating with health care professionals and policy makers, and on including patients' perceptions in the information, to ensure the content is adapted to significant changes in health care landscapes or policies.

Since the domains of the NASSS framework may be interdependent, reducing complexity in a particular domain may consequently also reduce the complexity in other domains. For instance, in case of national guidelines on ACP, health care professionals may be more willing to initiate ACP with their patients and patients may be more ready to initiate ACP.

Recommendations for Future Research

The following recommendations, based on the results and conclusions of the current study, can aid future implementation of comparable web-based ACP programs.

In order to improve the sustainable implementation of eHealth tools, in particular web-based ACP programs, first, we recommend to systematically and throughout the process (from idea to implementation) assess their complexities, with careful attention for all potential adopters, their external context, and their embedment and adaptation over time. The NASSS framework is suitable to conduct such evaluation, as well as other models that can be used for implementation or sustainability evaluation, such as the Normalization Process Theory (a framework to understand how interventions become embedded or "normalized" in health care settings) [46] and the Centre for eHealth Research Roadmap (CeHRes) [47,48]. These frameworks consider similar domains as the NASSS framework, sometimes ordered slightly differently, and can provide guidance for the sustainable development and implementation of eHealth technologies [46-48]. We expect that others who aim to implement similar web-based ACP programs and who apply comparable implementation frameworks, will benefit from the findings of this study and the recommendations. Yet, these should be contextualized for each application (eg, aligned with local policy frameworks and targeted to the target groups central in that program).

Second, we recommend identifying all relevant target groups of the eHealth tools, and explore their needs and preferences to ensure the tool meets their needs. For the web-based ACP program in particular, we recommend exploring health care

XSL•FO
RenderX

professionals' perspectives as well as their actual use and awareness of the program.

Third, we recommend assessing as well as improving the attractiveness of eHealth tools to potential users, including a culturally diverse population and to groups with low digital literacy.

Fourth, we recommend that research includes a long-term perspective regarding the implementation of eHealth tools. This may include evaluating the long-term effects as well as the identification of new scientific evidence that necessitates an update of the content of eHealth tools.

Conclusions

Relatively few complexities were identified considering the implementation of the web-based ACP program "Explore Your Preferences for Treatment and Care." The program is evidence-based, freestanding, and well-maintained, with straightforward, well-understood technology. The program is expected to generate a positive value for different stakeholders. Complexities include the broad target population of the program and sociocultural factors. People with limited digital literacy may need support to use the program. Its uptake might be improved by increasing awareness of ACP and the program among a wider population of potential users and among health care professionals. Addressing these issues may guide future use and sustainability of the program.

Acknowledgments

The authors would like to thank the interviewees for participating in this study. The authors would also like to thank the developers of the NASSS framework Prof Dr Trish Greenhalgh and Dr Chrysanthi Papoutsi for their valuable insights regarding the NASSS framework and their permission to reproduce the original NASSS Figure 1 and Multimedia Appendix 1.

This work was supported by the Netherlands Organisation for Health Research and Development (grant 555003026). The funding source had no involvement in the study design, collection, analysis, and interpretation of data, in the writing of the report, and in the decision to submit the paper for publication. The authors declare that no generative artificial intelligence was used in any portion of the manuscript writing.

Data Availability

Research data will be available for verification for at least 15 years. All data will be retained and stored at the central network server of the Erasmus MC, University Medical Center Rotterdam, and can be accessed by those specifically granted access.

Authors' Contributions

All authors contributed to the concept or design of the work or acquisition, analysis, or interpretation of data; drafted the paper or revised it critically for important intellectual content; approved the version to be published; and have participated sufficiently in the work to take public responsibility for appropriate portions of the content.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Domains and questions in the Non-adoption, abandonment, scale-up, spread, and sustainability (NASSS) framework (reproduced from Greenhalgh T et al [8], published under Creative Commons Attribution 4.0 International License [12].
[PDF File (Adobe PDF File), 165 KB - aging_v8i1e49507_app1.pdf]

References

1. Murray E. eHealth: where next? Br J Gen Pract 2014;64(624):325-326 [FREE Full text] [doi: 10.3399/bjgp14X680365] [Medline: 24982470]
2. Thuisarts.nl. Verken uw wensen voor zorg en behandeling. The Dutch College of General Practitioners. 2020 Apr 01. URL: https://www.thuisarts.nl/keuzehulp/verken-uw-wensen-voor-zorg-en-behandeling [accessed 2022-10-11]
3. Rietjens JAC, Sudore RL, Connolly M, van Delden JJ, Drickamer MA, Droger M, European Association for Palliative Care. Definition and recommendations for advance care planning: an international consensus supported by the European association for palliative care. Lancet Oncol 2017;18(9):e543-e551 [FREE Full text] [doi: 10.1016/S1470-2045(17)30582-X] [Medline: 28884703]
4. van der Smissen D, Rietjens JAC, van Dulmen S, Drenthen T, Vrijaldenhoven-Haitsma FRMD, Wulp M, et al. The web-based advance care planning program "Explore Your Preferences for Treatment and Care": development, pilot study, and before-and-after evaluation. J Med Internet Res 2022;24(12):e38561 [FREE Full text] [doi: 10.2196/38561] [Medline: 36459410]

XSL•FO
RenderX

5. Thuisarts.nl. The Dutch College of General Practitioners. Over Thuisarts.nl. 2022. URL: https://www.thuisarts.nl/over-thuisartsnl [accessed 2022-10-11]
6. van der Smissen D, Overbeek A, van Dulmen S, van Gemert-Pijnen L, van der Heide A, Rietjens JA, et al. The feasibility and effectiveness of web-based advance care planning programs: scoping review. J Med Internet Res 2020;22(3):e15578 [FREE Full text] [doi: 10.2196/15578] [Medline: 32181750]
7. van der Smissen D, Rietjens JAC, van Gemert-Pijnen L, van Dulmen S, van der Heide A, Korfage IJ. Information needs of patients with chronic diseases and their relatives for web-based advance care planning: a qualitative interview study. BMC Palliat Care 2021;20(1):77 [FREE Full text] [doi: 10.1186/s12904-021-00770-x] [Medline: 34053435]
8. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
9. Plsek PE, Greenhalgh T. Complexity science: the challenge of complexity in health care. BMJ 2001;323(7313):625-628 [FREE Full text] [doi: 10.1136/bmj.323.7313.625] [Medline: 11557716]
10. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. Milbank Q 2004;82(4):581-629 [FREE Full text] [doi: 10.1111/j.0887-378X.2004.00325.x] [Medline: 15595944]
11. Abimbola S, Patel B, Peiris D, Patel A, Harris M, Usherwood T, et al. The NASSS framework for ex post theorisation of technology-supported change in healthcare: worked example of the TORPEDO programme. BMC Med 2019;17(1):233 [FREE Full text] [doi: 10.1186/s12916-019-1463-x] [Medline: 31888718]
12. Attribution 4.0 International (CC BY 4.0). Creative Commons. URL: https://creativecommons.org/licenses/by/4.0/ [accessed 2025-01-30]
13. Greenhalgh T, Maylor H, Shaw S, Wherton J, Papoutsi C, Betton V, et al. The NASSS-CAT tools for understanding, guiding, monitoring, and researching technology implementation projects in health and social care: protocol for an evaluation study in real-world settings. JMIR Res Protoc 2020;9(5):e16861 [FREE Full text] [doi: 10.2196/16861] [Medline: 32401224]
14. Ott B, van Thiel GJ, de Ruiter CM, van Delden HJ. [Timing of advance care planning in frail elderly patients: when to start?] Kwetsbare ouderen en advance care planning: Wanneer beginnen? Ned Tijdschr Geneeskd 2015;159:A8295. [Medline: 25650032]
15. Billings JA, Bernacki R. Strategic targeting of advance care planning interventions: the Goldilocks phenomenon. JAMA Intern Med 2014;174(4):620-624. [doi: 10.1001/jamainternmed.2013.14384] [Medline: 24493203]
16. Cheng SY, Lin CP, Chan HYL, Martina D, Mori M, Kim SH, et al. Advance care planning in Asian culture. Jpn J Clin Oncol 2020;50(9):976-989. [doi: 10.1093/jjco/hyaa131] [Medline: 32761078]
17. Bijna tweederde Nederlanders bekend met Thuisarts.nl. ICTHealth.nl. 2021. URL: https://www.icthealth.nl/nieuws/bijna-twee-derde-nederlanders-bekend-met-thuisarts-nl/ [accessed 2022-07-21]
18. Forse stijging aantal bezoeken Thuisarts.nl. Federatie Medisch Specialisten. 2022. URL: https://demedischspecialist.nl/nieuwsoverzicht/nieuws/forse-stijging-aantal-bezoeken-thuisartsnl [accessed 2023-02-15]
19. Springvloet L, Bos N, Jong J, Friele R, Boer D. De Transparantiemonitor 2018. Thuisarts.nl: Hoe draagt Thuisarts.nl bij aan transparantie en de best passende zorg?. 2019. URL: https://www.nivel.nl/nl/publicatie/de-transparantiemonitor-2018-thuisartsnl-hoe-draagt-thuisartsnl-bij-aan-transparantie-en [accessed 2025-01-03]
20. DigiToegankelijk. Logius, ministerie van binnenlandse zaken en koninkrijkrelaties. WCAG: Richtlijnen voor webcontent. 2022. URL: https://www.digitoegankelijk.nl/uitleg-van-eisen/wat-wcag [accessed 2022-12-02]
21. Thuisarts. Thuisarts.nl ook bruikbaar voor mensen met beperking. The Dutch College of General Practitioners. URL: https://www.thuisarts.nl/thuisartsnl-ook-bruikbaar-voor-mensen-met-beperking [accessed 2022-10-11]
22. Nicholls J, Lawlor E, Neitzert E, Goodspeed T. A Guide to Social Return on Investment. London: The Cabinet Office; 2012.
23. McIlfatrick S, Slater P, Bamidele O, Muldrew D, Beck E, Hasson F. 'It's almost superstition: If I don't think about it, it won't happen'. Public knowledge and attitudes towards advance care planning: A sequential mixed methods study. Palliat Med 2021;35(7):1356-1365 [FREE Full text] [doi: 10.1177/02692163211015838] [Medline: 34000901]
24. Rodi H, Detering K, Sellars M, Macleod A, Todd J, Fullerton S, et al. Exploring advance care planning awareness, experiences, and preferences of people with cancer and support people: an Australian online cross-sectional study. Support Care Cancer 2021;29(7):3677-3688. [doi: 10.1007/s00520-020-05878-z] [Medline: 33184712]
25. Jabbarian LJ, Zwakman M, van der Heide A, Kars MC, Janssen DJA, van Delden JJ, et al. Advance care planning for patients with chronic respiratory diseases: a systematic review of preferences and practices. Thorax 2018;73(3):222-230. [doi: 10.1136/thoraxjnl-2016-209806] [Medline: 29109233]
26. Drenthen T, Beijaert RPH, Jansen PWM, Korevaar JC, Smeele IJM. [What do you think of Thuisarts.nl? Experiences after 3 years of www.Thuisarts.nl] Thuisarts.nl, hoe bevalt dat? Ervaringen na 3 jaar Thuisarts.nl. Ned Tijdschr Geneeskd 2014;158:A8282. [Medline: 25515391]
27. Jimenez G, Tan WS, Virk AK, Low CK, Car J, Ho AHY. Overview of systematic reviews of advance care planning: Summary of evidence and global lessons. J Pain Symptom Manage 2018;56(3):436-459.e25 [FREE Full text] [doi: 10.1016/j.jpainsymman.2018.05.016] [Medline: 29807158]

XSL•FO
RenderX

28. Palliaweb. Advance care planning. Palliaweb. 2022. URL: https://palliaweb.nl/zoeken?searchText=advance+care+planning [accessed 2022-12-12]
29. Boddaert M, Caffarel J, Dermois M, Kodde A, Reyners AKL. Leidraad voor het proces en uniform vastleggen van proactieve zorgplanning (advance care planning, acp) naar aanleiding van de covid-19-pandemie. 2020. URL: https://palliaweb.nl/getmedia/1e1ce663-6f7d-46ac-a47a-c8cb589597b7/Leidraad-proactieve-zorgplanning-(ACP)_22102020_new.pdf [accessed 2020-10-22]
30. Programma palliantie. Meer dan zorg. ZonMw. 2022. URL: https://www.zonmw.nl/nl/over-zonmw/onderwijs/programmas/programma-detail/palliantie-meer-dan-zorg/ [accessed 2022-12-12]
31. ZonMw. Inzet van proactieve zorgplanning in de praktijk (uitgewerkte subsidieaanvragen). ZonMw. 2022. URL: https://www.zonmw.nl/nl/subsidies/openstaande-subsidieoproepen/detail/item/inzet-van-proactieve-zorgplanning-in-de-praktijk-uitgewerkte-subsidieaanvragen/ [accessed 2023-03-03]
32. SIRE. Campagne 'de dood, praat erover' campaign 'death, talk about it'. URL: https://sire.nl/campagnes/de-dood-praat-erover/ [accessed 2022-10-27]
33. Rijksoverheid. Rijksoverheid [Dutch Government]. Campagnetoolkits palliatieve zorg. 2019. URL: https://www.campagnetoolkits.nl/palliatieve-zorg [accessed 2022-10-27]
34. Rietjens J, Sudore R, Deliens L, Ott B, Connolly M, van Delden J, et al. Internationale definitie van advance care planning. Huisarts Wet 2018;61(8):22-29 [FREE Full text] [doi: 10.1007/s12445-018-0224-y]
35. Toolkit advance care planning mbt het levenseinde. Frank Guldemond, Brenda Ott, Annet Wind, kaderhuisartsen ouderengeneeskunde. 2017. URL: https://www.laego.nl/wp-content/uploads/2018/04/Toolkit-Advance-Care-Planning-2017.pdf [accessed 2022-12-12]
36. KNMG. Patiëntenfederatie Nederland. Handreiking voor artsen om met patiënten in gesprek te gaan tijdig praten over het levenseinde. KNMG and Patiëntenfederatie Nederland. URL: https://data.maglr.com/3230/issues/30306/393465/downloads/handreiking-tijdig-praten-over-het-levenseinde.pdf [accessed 2022-12-12]
37. KNMG. Patiëntenfederatie Nederland. Praat op tijd over je levenseinde over passende zorg in je laatste levensfase KNMG and Patiëntenfederatie Nederland. URL: https://www.knmg.nl/advies-richtlijnen/dossiers/praten-over-het-levenseinde-1.htm [accessed 2022-12-12]
38. Chambers DA, Glasgow RE, Stange KC. The dynamic sustainability framework: addressing the paradox of sustainment amid ongoing change. Implement Sci 2013;8:117 [FREE Full text] [doi: 10.1186/1748-5908-8-117] [Medline: 24088228]
39. Aarons GA, Hurlburt M, Horwitz SM. Advancing a conceptual model of evidence-based practice implementation in public service sectors. Adm Policy Ment Health 2011;38(1):4-23 [FREE Full text] [doi: 10.1007/s10488-010-0327-7] [Medline: 21197565]
40. Dupont C, Smets T, Monnet F, Pivodic L, De Vleminck A, Van Audenhove C, et al. Publicly available, interactive web-based tools to support advance care planning: Systematic review. J Med Internet Res 2022;24(4):e33320 [FREE Full text] [doi: 10.2196/33320] [Medline: 35442207]
41. Butler M, Ratner E, McCreedy E, Shippee N, Kane RL. Decision aids for advance care planning: an overview of the state of the science. Ann Intern Med 2014;161(6):408-418 [FREE Full text] [doi: 10.7326/M14-0644] [Medline: 25069709]
42. Radbruch L, De Lima L, Knaul F, Wenk R, Ali Z, Bhatnaghar S, et al. Redefining palliative care-A new consensus-based definition. J Pain Symptom Manage 2020;60(4):754-764 [FREE Full text] [doi: 10.1016/j.jpainsymman.2020.04.027] [Medline: 32387576]
43. Sibbald B, McDonald R, Roland M. Shifting care from hospitals to the community: A review of the evidence on quality and efficiency. J Health Serv Res Policy 2007;12(2):110-117. [doi: 10.1258/135581907780279611] [Medline: 17407662]
44. Zwakman M, Jabbarian LJ, van Delden J, van der Heide A, Korfage IJ, Pollock K, et al. Advance care planning: a systematic review about experiences of patients with a life-threatening or life-limiting illness. Palliat Med 2018;32(8):1305-1321 [FREE Full text] [doi: 10.1177/0269216318784474] [Medline: 29956558]
45. Seymour J. The impact of public health awareness campaigns on the awareness and quality of palliative care. J Palliat Med 2018;21(S1):S30-S36 [FREE Full text] [doi: 10.1089/jpm.2017.0391] [Medline: 29283867]
46. May CR, Mair F, Finch T, MacFarlane A, Dowrick C, Treweek S, et al. Development of a theory of implementation and integration: Normalization process theory. Implement Sci 2009;4:29 [FREE Full text] [doi: 10.1186/1748-5908-4-29] [Medline: 19460163]
47. van Gemert-Pijnen L, Kelders SM, Kip H, Sanderman R. Ehealth research, theory and development: A multi-disciplinary approach. Oxfordshire: Routledge; 1999.
48. CEHRES Roadmap Toolkit. Centre for eHealth & Wellbeing Research, University of Twente. 2018. URL: https://www.utwente.nl/en/bms/ehealth/cehres-roadmap-toolkit/ [accessed 2023-05-18]

Abbreviations

ACP: advance care planning
CeHRes: Centre for eHealth Research Roadmap
NASSS: nonadoption, abandonment, scale-up, spread, and sustainability

XSL•FO
RenderX

NASSS-CAT: nonadoption, abandonment, scale-up, spread, and sustainability complexity assessment tool

Edited by Y Jiang; submitted 01.06.23; peer-reviewed by F Monnet, P Delir Haghighi, Y Chu; comments to author 25.10.23; revised version received 16.02.24; accepted 06.11.24; published 04.03.25.

Please cite as:
van der Smissen D, Schreijer MA, van Gemert-Pijnen LJEWC, Verdaasdonk RM, van der Heide A, Korfage IJ, Rietjens JAC Implementation of a Web-Based Program for Advance Care Planning and Evaluation of its Complexity With the Nonadoption, Abandonment, Scale-Up, Spread, And Sustainability (NASSS) Framework: Qualitative Evaluation Study
JMIR Aging 2025;8:e49507
URL: https://aging.jmir.org/2025/1/e49507
doi:10.2196/49507
PMID:40053753

©Doris van der Smissen, Maud A Schreijer, Lisette J E W C van Gemert-Pijnen, Rudolf M Verdaasdonk, Agnes van der Heide, Ida J Korfage, Judith A C Rietjens. Originally published in JMIR Aging (https://aging.jmir.org), 04.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Determinants of Telehealth Adoption Among Older Adults: Cross-Sectional Survey Study

Siow-Hooi Tan[1], PhD; Yee-Yann Yap[1], MPhil; Siow-Kian Tan[2], PhD; Chee-Kuan Wong[3], MD, MRCPE

[1]Faculty of Management, Multimedia University, Cyberjaya, Malaysia

[2]School of Economics and Management, Xiamen University Malaysia, Sepang, Malaysia

[3]Department of Medicine, Faculty of Medicine, Universiti Malaya, Kuala Lumpur, Malaysia

Corresponding Author:
Siow-Hooi Tan, PhD
Faculty of Management, Multimedia University
Persiaran Multimedia
Cyberjaya, 63100
Malaysia
Phone: 60 38312 ext 5658
Email: shtan@mmu.edu.my

Abstract

Background: The aging population and the accompanying rise in chronic diseases have intensified the need to study the adoption of telehealth services. However, the success of telehealth services depends not only on their ease and usefulness but also on addressing broader concerns. Despite being a substantial user group in traditional health services, older adults may encounter barriers to adopting telehealth services. Increasing the adoption of telehealth among the older adult population is crucial for enhancing their access to care and managing the challenges of aging effectively.

Objective: We aimed to explore factors influencing the adoption of telehealth services among older adults in Malaysia, going beyond the conventional framework by incorporating transition cost and subjective well-being as additional constructs.

Methods: A cross-sectional survey was conducted among 119 adults aged ≥60 years in Malaysia, using 39 survey items adapted from existing studies. Data analysis was performed using partial least squares structural equation modeling, with both the measurement model and structural model being evaluated. To determine the predictive relevance of the model, PLSpredict was applied. In addition, importance-performance map analysis was conducted to further expand on the structural model results by assessing the performance of each variable.

Results: Of the 119 participants, 52 (43.7%) were women and 67 (56.3%) were men. The study found that subjective well-being (β=0.448; *P*<.001) was the most significant factor, followed by attitude (β=0.242; *P*<.001), transition cost (β=−0.163; *P*<.001), and perceived usefulness (β=0.100, *P*=.02) in influencing telehealth service intention. Furthermore, perceived ease of use (β=0.271; *P*<.001), availability (β=0.323; *P*<.001), subjective well-being (β=0.261; *P*<.001), and trust (β=0.156, *P*=.004) positively influenced perceived usefulness, while inertia (β=0.024, *P*=.22) did not. In addition, availability (β=0.420; *P*<.001) and subjective well-being (β=0.260; *P*<.001) were positively related to perceived ease of use, with inertia (β=−0.246; *P*<.001) having a negative impact. The importance-performance map analysis results showed that subjective well-being (importance=0.532) was the most crucial factor for older adult users, while availability (importance=70.735) had the highest performance in telehealth services.

Conclusions: This research underscores the importance of catering to the subjective well-being of older adults and optimizing the availability of telehealth services to encourage adoption, ultimately advancing health care accessibility and quality for this vulnerable demographic.

(JMIR Aging 2025;8:e60936) doi:10.2196/60936

KEYWORDS

telehealth services adoption; survey; questionnaire; telehealth; older adult population; subjective well-being; transition cost; technology acceptance model; importance-performance map analysis; IPMA

XSL•FO
RenderX

Introduction

Background

Telehealth refers to the delivery of long-distance clinical health care services by health care professionals using electronic information and telecommunications technologies. Due to the growth of the internet and communication infrastructure, telehealth has gradually developed into a practical and secure way for patients to obtain reliable information and medical consultation [1]. Using telehealth has several advantages, such as eliminating the need for direct patient–health care provider interaction during regular treatment. Telehealth can also provide remote care, which can reduce the need for medical center resources and increase the accessibility of care.

Previous telehealth studies have developed various concepts to address how telehealth could fulfill the needs of older adults, such as in the context of chronic disease management, enhancing independent living and improving their overall well-being [2]. Telehealth can be useful for older adults with chronic diseases to monitor their conditions at home. For instance, with the use of telehealth, older adult patients can prevent unnecessary hospitalization and still ensure they receive emergency treatment in a cost-efficient manner [3,4]. In addition, Chou et al [5] found that telehealth can improve the well-being of older adults by enhancing their quality of life. Furthermore, telehealth can promote independent living at home among older adults [6].

Older adults could benefit from telehealth as they are the fragile groups who may need this service sooner or later. Nevertheless, they are also the group most concerned about technology. The literature provides evidence that older adults can receive several advantages with the use of telehealth services, such as health monitoring and care, disease prevention, improved quality of life, and independent living. Despite all the advantages, the adoption of telehealth technology may be challenging for older adults because they are slower and more resistant to adopting new technology as they tend to be more traditional, cautious, risk-averse, and suspicious toward innovations [7]. There is a lack of knowledge about what factors individuals will consider when accepting telehealth [8]. The market response indicates that acceptance of technology by older adults is a complex problem that is impacted by a variety of factors rather than just the technology's performance or cost [9].

In the context of Malaysia, telehealth is becoming increasingly important in the health care system, especially with the older adult population expected to exceed 15% by 2030 [10]. This demographic shift will place added pressure on the health care system, particularly due to a rise in chronic diseases [11]. Malaysia's health care system offers accessible services but faces a workforce shortage to meet growing demands. With 2.4 physicians per 1000 people—fewer than those in Singapore, Japan, and Australia—Malaysia faces an aging population and staffing shortages leading to overcrowded public hospitals and strained health care capacity [11]. This has drastically burdened the health care system in Malaysia.

In response, telehealth offers a vital solution to address the growing imbalance between health care supply and demand as health care needs continue to rise [12]. The Malaysian government has been actively exploring technological solutions and launching telehealth initiatives to address the rising health care needs of its aging population. For instance, Malaysia's Ministry of Health initiated a teleconsultation at public hospitals to improve health care access and reduce congestion. Despite these efforts and the recent surge in telehealth-related studies, there remains a scarcity of research to investigate telehealth adoption in emerging economies, especially from the older adults' perspective [11,13]. Investigating telehealth adoption by older adults across different countries is essential, as varying levels of technological development and cultural contexts substantially influence their attitudes and behaviors [13,14].

Previous studies on technology adoption among older adults span various cultural contexts, revealing distinct factors influencing their behavior. In Canada, a study by Ahmed et al [15] found that >half of older adults adopted new technology for online social interactions. Despite having the knowledge to stay connected, they faced challenges like limited access and motivation. In a cross-cultural survey by Elimelech et al [16], older adults in Israel, France, and Spain exhibited different perceptions of technology use, emphasizing the need for culturally tailored adaptations. In Australia, Catapan et al [17] reported that patients had a high level of confidence and trust in the use of telehealth. Meanwhile, in China, Lin et al [18] showed that telehealth's ease of use and usefulness played a substantial role in affecting its adoption. In a developed country such as Singapore, Zhang et al [19] found that telehealth effectively supports the health-seeking behavior of older adults, challenging the belief that they resist technology and lack proficiency. Similarly, Haimi and Sergienko [20] found that telehealth uses among older adults in Israel remained elevated after the COVID-19 pandemic, indicating their ability to effectively learn and use digital health services.

However, in Malaysia, older adults may have different perceptions of telehealth. According to Ting et al [21], many older adults still prefer face-to-face interactions due to a cultural preference for personal consultations. Reservations about the impersonal nature of telehealth remain a substantial obstacle. However, the behavior of older adults toward telehealth in emerging economies, such as Malaysia, remains underexplored. There is a need for research focused on telehealth adoption within the Malaysian context to gain better grasp of this context.

Theory

Researchers have introduced various theoretical models to explain consumer behavior in the context of technology adoption. Well-established models for predicting technology acceptance among consumers include the theory of planned behavior by Ajzen [22], the technology acceptance model (TAM) by Davis et al [23], TAM2 by Venkatesh and Davis [24], the unified theory of acceptance and use of technology by Venkatesh et al [25], and TAM3 by Venkatesh and Bala [26], among others.

A growing body of research highlights the reliability and effectiveness of the TAM in explaining technology adoption. TAM, introduced by Davis et al [23], uses the concepts of perceived usefulness and perceived ease of use to elucidate how

XSL•FO
RenderX

technological factors influence a consumer's intent to adopt a specific technology. For the adoption of health-related technologies, the TAM and the unified theory of acceptance and use of technology have emerged as 2 prominent models, as demonstrated in studies by Harst et al [27], Heinsch et al [28], Rouidi et al [29], and Lin et al [18]. In the context of the aging population, a recent systematic literature review by Yap et al [30] confirmed that the TAM is the most widely used theory for explaining technology adoption among older adults.

Despite the widespread use of the TAM in the literature, previous studies argued that solely using the fundamental TAM to identify the consumer's technology adoption is insufficient [31,32]. Similarly, Attié and Meyer-Waarden [33] criticized that functional and utilitarian benefits, such as perceived ease of use, are insufficient to explain technology acceptance. Therefore, it can be found that previous studies have frequently extended the original TAM with additional variables or other theories to better reflect the technology's acceptance. For instance, Zhou et al [34] extended the original TAM by incorporating perceptions of medical service quality and information quality into the model for predicting telehealth acceptance among older adults. Telehealth acceptance, in turn, is influenced by older adults' perceptions of telehealth and their current behavioral intentions toward telehealth services. In addition, Rho et al [8] extended the TAM with perceived incentives, self-efficacy, and accessibility of patients' medical records, while Klingberg et al [35] included image, self-efficacy, voluntariness, compatibility, and anxiety in the TAM fundamental framework to enhance the explanatory power of the TAM. Moreover, integration of theories and perspectives has been performed as well, such as in the study by Tsai et al [36] that integrated the TAM with the status quo bias and technology anxiety concept to explain the telehealth intention. Therefore, it is suggested that there is a need for research to expand the TAM by including additional variables to provide a more comprehensive explanation and understanding regarding the telehealth intention of older adults.

Objectives

The objectives of this study are as follows:

- To investigate the coexistence and possible effects of TAM constructs, transition cost, and subjective well-being on telehealth service adoption among older adults
- To examine how inertia, availability, subjective well-being, and trust relate to the TAM's key antecedents

Methods

Research Model and Hypothesis Development

Overview

Telehealth is an effective and advanced alternative method for delivering health care services. In the context of telehealth adoption, the fundamental TAM might be insufficient to explain the telehealth adoption among older adults. In particular, the older adult population might often hold negative opinions about technology's inaccuracies that would influence their intention to use it [37]. Considering the evolving landscape of telehealth adoption among older adults in Malaysia and the limited existing literature, this study contributes to the expansion of the TAM. It does so by introducing and exploring a set of key factors aimed at comprehending the older adult population's willingness to embrace telehealth. In addition to attitude and perceived usefulness, this study proposes 2 additional constructs: transition cost and subjective well-being, to evaluate older adults' intention to use telehealth.

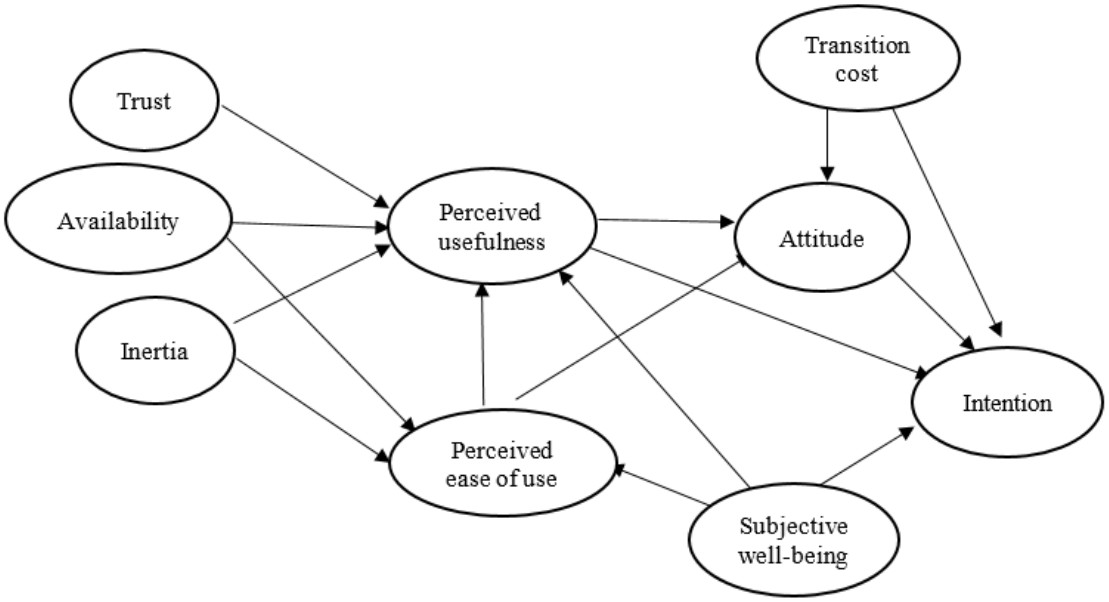
First, subjective well-being is included to address critiques of the traditional TAM. Recent studies [11,32,33] argue that the TAM's focus solely on utilitarian benefits is insufficient to fully explain consumer technology adoption. On the basis of transformative consumer studies and the uses and gratification theory [38], it is evident that beyond utilitarian benefits like usefulness and ease of use, consumers also seek affective elements, such as well-being when adopting a new technology [33]. The literature suggests enhancing the TAM by incorporating affective elements like subjective well-being to more accurately predict technology adoption [11]. Nevertheless, subjective well-being has received minimal attention in telehealth adoption studies, particularly from the perspective of the older adult population. Therefore, research proposes that subjective well-being be considered as one of the predictors of older adults' intention to use telehealth. If older adults perceive that telehealth can enhance their well-being, they are more likely to adopt it, as people naturally seek experiences that improve their overall quality of life [39,40].

The inclusion of transition cost as an extension to the TAM is grounded in the status quo bias theory, which posits that individuals tend to prefer maintaining their current routines over adopting change [41]. This is especially relevant for older adults, who often find transitions, such as shifting to telehealth, burdensome due to the perceived effort, time, and disruption of familiar health care practices like face-to-face consultations. Many older adults grew up in a time when technological innovations were not widespread, leading them to develop long-standing routines that provide comfort and predictability [42]. As a result, this study proposes that transition costs are particularly impactful for this group, as they are more likely to resist switching to telehealth in favor of maintaining their familiar health care practices.

Furthermore, the research incorporates availability, trust, inertia, and subjective well-being as factors influencing the TAM constructs. Subsequent sections of this study will provide a detailed analysis of the significance of these constructs within the unique context of this research. Figure 1 shows the research model of this study. The proposed hypotheses based on the developed research model are discussed in subsequent sections.

XSL•FO
RenderX

Figure 1. Research model.



Attitude

Attitude is an evaluation of effect, which refers to a person's positive or negative feelings regarding performing the respective behavior [43]. The impact of attitude on behavioral intention is a significant relationship in the theory of reasoned action, the theory of planned behavior, and the TAM. The connection between one's attitude and their intention signifies that individuals are more inclined to embrace technology if they hold a favorable perception of it, as noted by Davis et al [23]. The concept of attitude has played an important role in numerous studies seeking to gain deeper insights into consumers' willingness to adopt health care technologies. Previous research has consistently confirmed that a positive attitude toward health care–related technologies among consumers has a substantial impact on their intent to use such technologies, as evidenced by studies conducted by Park et al [44], Papa et al [45], Rajak and Shaw [46], and Ahn and Park [47].

Furthermore, Tsai et al [36] unveiled that the older adults' attitudes exert a positive influence on their intention to adopt telehealth. In addition, prior research has shown that attitude can serve as a mediator between beliefs and behavioral intent, as demonstrated by the work of Yang and Yoo [48]. Consequently, this study proposed the following hypothesis:

- Hypothesis 1: there is a positive association between attitude and the intention to adopt telehealth.

Perceived Ease of Use

Perceived ease of use is defined as the degree to which an older adult perceives that using telehealth technology would be free of effort. Perceived ease of use is one of the key constructs in the TAM. According to the TAM, perceived ease of use is related to attitudes and perceived usefulness of new technologies, which also influences intention [24]. Previous research has found that perceived ease of use influences consumers' attitudes and perceived usefulness toward new technologies [31,36,46,47]. For instance, Lazaro et al [49] confirmed that perceived ease of use positively affects the older adults' perceived usefulness and their attitude toward wearable health care technology. Hence, we proposed the following hypothesis:

- Hypothesis 2a: there is a positive association between perceived ease of use and perceived usefulness.
- Hypothesis 2b: there is a positive association between perceived ease of use and attitude.

Perceived Usefulness

Perceived usefulness is defined as the degree to which an individual feels that using a certain technology will improve their job performance [23]. Researchers have supported perceived usefulness as a crucial factor that predicts various types of technology adoption among older adults, such as the internet [50], social networking sites [51], health monitoring wearable technologies [52], telehealth [53], and automation technology [54]. Besides, previous studies revealed that perceived usefulness was highly relevant in predicting telehealth acceptance among the older adult population [34,55]. In addition, perceived usefulness influences consumers' attitudes toward new technologies [36,46,47]. Therefore, the following hypothesis is proposed:

- Hypothesis 3a: there is a positive association between perceived usefulness and attitude.
- Hypothesis 3b: there is a positive association between perceived usefulness and intention to adopt telehealth.

Transition Cost

According to Kim and Kankanhalli [56], transition costs refer to the user's perceived disutility that they incur when switching from the status quo to a new technology. When older adults consider using new technology, such as telehealth, the transition costs involved are essential. Transition costs that are uncertain might become a barrier and negatively influence a person's attitude [57]. In addition, Hsieh [58] revealed that the transition cost that might be incurred when using health-related technology

XSL•FO
RenderX

is one of the key concerns for technology adoption. An individual is likely to continue and remain with an existing system if the transition costs involved, such as effort and time, to learn to use a new technology, are deemed to be high. In the context of telehealth adoption, Tsai et al [36] claimed that transition costs negatively affect older adults' attitudes and hence affect their telehealth adoption. Therefore, if the transition cost to use a new technology is high, older adults will have a negative attitude toward it and not intend to use it. On the basis of the discussions, the following hypothesis is proposed:

- Hypothesis 4a: there is a negative association between transition cost and attitude.
- Hypothesis 4b: there is a negative association between transition cost and intention to adopt telehealth.

Subjective Well-Being

Subjective well-being refers to an individual's perception of an experience positively by using affective reactions and cognitive judgment instead of objective facts [59]. When adopting new technology, people always seek pleasurable experiences that enhance their well-being [39,40]. Previous studies have revealed that well-being influences consumer technology adoption [60]. Al-Jabri and Sohail [61] revealed that technology's characteristics contribute to a person's well-being. Similarly, well-being could act as a determinant that influences technology use [40]. Wu and Lu [62] have highlighted that positive emotions become a motivator for technology adoption when a person uses technology. In addition, findings show that well-being toward technology positively influences an individual's perceived ease of use, usefulness, and intention across all consumer adoption phrases [33,63,64]. Consumers will develop positive emotions toward technology when they perceive that using it will enhance their well-being. Therefore, positive emotions will positively influence the perceived benefits, such as perceived usefulness and ease of using the technology [33].

Therefore, this study proposes that older adults perceive that telehealth will enhance their well-being, which consequently will influence their perceived ease of use, usefulness, and intention to use telehealth. Therefore, we hypothesize the following:

- Hypothesis 5a: there is a positive association between well-being and perceived ease of use.
- Hypothesis 5b: there is a positive association between well-being and perceived usefulness.
- Hypothesis 5c: there is a positive association between well-being and intention to adopt telehealth.

Inertia

Inertia refers to the degree of a person's willingness to continue using traditional physical products despite knowing that better options are available [41]. Even when better alternatives or switching incentives are available, consumers remain attached to and steadfast in their use of existing technologies [65]. Hence, the greater an individual's attachment toward a thing that he or she is familiar with, the less inclined they are to explore new experiences. In addition, Bem [66] and Petty and Cacioppo [67] claimed that people usually depend on their prior behavior and hence fail to recognize a new technology's advantages. In line with this assumption, older adults always seek to maintain their existing internal and external structures when making adaptive decisions, as they would prefer to continue to engage in similar activities or behaviors as they did throughout their previous experiences [68]. The older adult population grew up in an era when technological innovation was not commonly used. As a result of their early experiences, they might have long been accustomed to receiving health care services physically at hospitals or clinics. Older adults might perceive telehealth as not useful and not easy to operate as they prefer to continue to engage in the behavior they are more familiar with to minimize feelings of anxiety. It is hypothesized that inertia will negatively influence the older adults' behavioral perceptions of a new technology and hence create lower inclinations to use new technologies. Individuals who have high inertia tend to reduce the variety of technologies that are available to them and rely on prior behavior to influence their perceptions and intentions [36]. Therefore, the following hypothesis is proposed:

- Hypothesis 6a: there is a negative association between inertia and perceived ease of use.
- Hypothesis 6b: there is a negative association between inertia and perceived usefulness.

Availability

On the basis of the study by Venkatesh [69], availability refers to the extent to which consumers perceive that they can obtain a technological service or product without barriers, along with the presence of organizational support to help them overcome any challenges in using the technology. A person's control belief about the availability of organizational resources and support structures to enable technology use is related to facilitating conditions [26]. Many previous studies have proved the positive impact of facilitating conditions on technology adoption [70,71]. In the telehealth adoption context, the determinant associated with external control might involve the availability of manufacturer's assistance, where the firms provide consumers assistance to overcome the difficulties of using a new technology. Telehealth technologies can make medical resources available to health care professionals, caregivers, and older adults at any time and from any location, allowing a considerable improvement in patient health care. Wu et al [72] revealed that availability positively influences the perceived usefulness of the telehealth care technology. In addition, previous studies also provided support on the impact of availability in the context of telehealth adoption [73,74]. Similarly, a recent study by Tsai et al [36] revealed that availability is an important predictor in determining perceived ease of use and usefulness of telehealth among the older adult population. According to the existing evidence, this study hypothesizes that availability would increase the older adults' perceived ease of use and usefulness of telehealth.

- Hypothesis 7a: there is a positive association between availability and perceived ease of use.
- Hypothesis 7b: there is a positive association between availability and perceived usefulness.

XSL•FO
RenderX

Trust

Trust is evidently a crucial determinant in health care adoption. On the basis of the study by Gefen et al [75], trust refers to a sense of confidence in the trustworthiness and integrity of the other party. When it comes to technological adoption, trust is essential, especially when the technology is relatively new and might involve risks and uncertainties for the older adult population. Despite trust not being included in the original TAM, it has been incorporated into several of the study contexts. Many existing studies demonstrate that trust has a strong positive influence on technology adoption [45,76]. Previous studies also revealed that trust plays an important role in predicting adoption of health-related technologies (eg, telehealth) [76,77].

In the context of this study, telehealth can be a difficult and complex task, especially for older adults, as its use requires a good understanding of devices to communicate effectively with the health care providers [78,79]. Hence, people might lack trust in using telehealth due to the risks incurred, including unclear regulatory authority in place to deal with issues like confidentiality, misconduct, and liability in telehealth. Older adults, who used to obtain services physically, might have trust issues with telehealth's ability to replace in-person consultations and physical health assessments. Particularly in the IT setting, Li et al [80] revealed that trust is crucial, as people must overcome the perceived risk before technology adoption. Previous research has established the importance of trust in determining health-related technology adoption [46]. For instance, Catapan et al [17], Chew et al [81], and Orrange et al [82] showed that trust substantially affects telehealth adoption. Hence, this study hypothesizes that trust would increase older adults' perceived usefulness of telehealth.

- Hypothesis 8: there is a positive association between trust and perceived usefulness.

Research Instrument Development

The measurement scales were adjusted in accordance with existing literature and tailored to suit our research context. The questionnaire was divided into 2 sections. The first part consisted of the 11 constructs used in this study: attitude, availability, transition cost, perceived ease of use, perceived usefulness, inertia, trust, subjective well-being, and intention to adopt. The second part comprised a survey focusing on demographic characteristics, such as gender and ethnicity. An overview of the instrument is provided in Multimedia Appendix 1 [32,36,83,84].

The scales for attitude, availability, transition cost, perceived ease of use, perceived usefulness, and inertia were taken from the study of Tsai et al [36] and Zhang and Zaman [83]. The measures for subjective well-being were based on the work of Yap et al [32]. Trust and intention were measured according to the work of Wu et al [84]. In total, 39 items were assessed using a 5-point Likert scale, which ranged from *completely disagree* (score=1) to *completely agree* (score=5).

To ensure the questionnaire's reliability, a pilot survey was carried out involving 10 older adults. This pretest, which involved contacting 10 older adults before conducting the web-based survey, was conducted to validate the instrument. On the basis of the feedback received, minor adjustments were made to the questionnaire to improve its effectiveness.

Research Sample and Data Collection Procedure

This study focused on the factors influencing the intention to adopt telehealth among adults aged ≥60 years in Malaysia. This demographic was chosen as research participants due to their increased susceptibility to chronic diseases, as outlined by Tsai et al [36]. Chronic diseases can affect individuals of all ages, but the risk escalates as people advance in age, justifying the selection of this population for our research.

In our data collection process, we used a survey approach to investigate the intent of older adults to adopt telehealth services. Recognizing the challenges in directly reaching this demographic, we engaged students as intermediaries to connect with their family members who were older. This method combines elements of convenience and snowball sampling, as students were encouraged to distribute surveys within their social networks, primarily targeting their older family members. We encouraged student intermediaries to recruit participants from diverse geographic regions and socioeconomic backgrounds to enhance sample diversity where possible.

The rationale behind using student intermediaries was 2-fold: first, it provided a practical and effective means of reaching older adult participants, a group that may not be as digitally connected or comfortable with technology. The student intermediaries, who were trained to administer the survey, acted as trusted conduits, facilitating communication and engagement with older adult participants in a familiar, trustworthy, and less intimidating environment. This helped to overcome potential barriers related to accessibility and comprehension of the survey, ensuring that the older adult participants felt comfortable and supported throughout the process. Hence, it was essential for the student intermediaries to have a thorough understanding of both the study's purpose and the questionnaire to effectively guide their older family members. To ensure this, students underwent a comprehensive briefing before the survey, equipping them with a solid grasp of the questionnaire and its objectives.

Data Analysis

This study conducted data analysis using partial least squares structural equation modeling (PLS-SEM). We used SmartPLS 3.2.8 [85] for PLS-SEM, as it is well-suited for analyzing measurement and structural models without the need for normality assumptions. This is particularly useful because survey research data are often not normally distributed [86]. Furthermore, PLS-SEM offers a higher explanatory power compared to covariance-based structural equation modeling.

Ethical Considerations

This study received formal approval from the Research Ethics Committee of Multimedia University (EA2882021). Informed written consent was obtained from all survey respondents before participation. Respondents were required to read the ethical statement at the top of the survey form and proceed only if they agreed to participate. All collected data are treated with the utmost confidentiality, ensuring anonymity and used solely for

XSL•FO
RenderX

research purposes. Additionally, consent to publish was obtained from all participants.

Results

Respondent Characteristics

The survey was conducted from March 1, 2022, to August 31, 2022, resulting in 125 received samples, of which 119 were considered valid. While this approach offers advantages for reaching a challenging-to-access population, researchers should remain vigilant about potential biases arising from the familial and social connections of the student intermediaries. In the sample, 52 (43.7%) of the 119 respondents were women, and 67 (56.3%) were men. Additional demographic details are provided in Multimedia Appendix 2.

Reliability and Validity Tests

The evaluation of the measurement model involved assessing reliability and validity. In Table 1, all variables exhibited factor loadings >0.7 [87], and Cronbach α exceeded 0.7, signifying a high level of reliability. Both the composite reliability and average variance extracted surpassed 0.7 and 0.5, respectively, indicating strong convergent validity [88]. In the second step of the analysis, we evaluated discriminant validity using the heterotrait-monotrait (HTMT) ratio criterion, as suggested by Henseler et al [89] and Franke and Sarstedt [90]. As indicated in Table 2, the results of the HTMT criterion demonstrate that all HTMT values fall below the threshold of 0.85 for the more stringent criterion. This indicated that the respondents recognized the distinctiveness of all the constructs. The combination of these 2 validity tests affirms that the measurement items exhibited both validity and reliability.

XSL•FO
RenderX

Table 1. Measurement model and cross-validated redundancy.

Constructs and items	Loadings	Composite reliability	Average variance extracted
Attitude		0.896	0.743
ATT[a]1	0.833		
ATT2	0.850		
ATT3	0.900		
Perceived ease of use		0.963	0.897
EOU[b]1	0.938		
EOU2	0.954		
EOU3	0.948		
Perceived usefulness		0.928	0.762
USE[c]1	0.841		
USE2	0.857		
USE3	0.902		
USE4	0.890		
Transition cost		0.953	0.872
COST[d]1	0.920		
COST2	0.939		
COST3	0.942		
Inertia		0.889	0.728
INE[e]1	0.889		
INE2	0.811		
INE3	0.858		
Availability		0.886	0.723
AVAI[f]1	0.862		
AVAI2	0.888		
AVAI3	0.798		
Trust		0.900	0.751
TRUST[g]1	0.886		
TRUST2	0.808		
TRSUT3	0.903		
Subjective well-being		0.959	0.885
SWB[h]1	0.934		
SWB2	0.948		
SWB3	0.941		
Intention		0.932	0.821
INT[i]1	0.911		
INT2	0.885		
INT3	0.923		

[a]ATT: attitude.

[b]EOU: perceived ease of use.

[c]USE: perceived usefulness.

[d]COST: transition cost.

XSL•FO
RenderX

[e]INE: inertia.
[f]AVA: availability.
[g]TRUST: trust.
[h]SWB: subjective well-being.
[i]INT: intention.

Table 2. Discriminant validity (heterotrait-monotrait 0.85 criterion).

	ATT[a]	EOU[b]	USE[c]	COST[d]	INE[e]	AVA[f]	TRUST[g]	SWB[h]	INT[i]
ATT	—[j]	—	—	—	—	—	—	—	—
EOU	0.677	—	—	—	—	—	—	—	—
USE	0.769	0.719	—	—	—	—	—	—	—
COST	0.549	0.623	0.498	—	—	—	—	—	—
INE	0.390	0.427	0.261	0.523	—	—	—	—	—
AVA	0.665	0.654	0.781	0.453	0.140	—	—	—	—
TRUST	0.688	0.530	0.683	0.389	0.210	0.527	—	—	—
SWB	0.755	0.577	0.715	0.170	0.374	0.545	0.753	—	—
INT	0.753	0.674	0.701	0.570	0.432	0.579	0.677	0.818	—

[a]ATT: attitude.
[b]EOU: perceived ease of use.
[c]USE: perceived usefulness.
[d]COST: transition cost.
[e]INE: inertia.
[f]AVA: availability.
[g]TRUST: trust.
[h]SWB: subjective well-being.
[i]INT: intention.
[j]Not applicable.

Results of the Structural Model

Following the guidance of Hair et al [91], we reported the path coefficients, SEs, 1-tailed *t* test values, and *P* values using a bootstrapping procedure with 5000 resamples [92]. This large resample size ensures result stability, according to Hair et al [91] and Hair and Alamer [93]. In addition, Hahn and Ang [94] emphasized that *P* values alone were insufficient for hypothesis significance and recommended combining *P* values, effect sizes, and bias-corrected interval for a more comprehensive evaluation. Table 3 presents the evaluation of the hypotheses. Moreover, the results of bias-corrected interval, effect sizes, and variance inflation factor are presented in Multimedia Appendix 3.

XSL•FO
RenderX

Table 3. Hypothesis testing direct effects[a].

Hypothesis	Relationship	Standard β	SD	*t* test value	*P* value
Hypothesis 1	Attitude → intention	0.242	0.058	4.168	<.001
Hypothesis 2a	Perceived ease of use → perceived usefulness	0.271	0.048	5.693	<.001
Hypothesis 2b	Perceived ease of use → attitude	0.209	0.056	3.734	<.001
Hypothesis 3a	Perceived usefulness → attitude	0.456	0.054	8.450	<.001
Hypothesis 3b	Perceived usefulness → intention	0.100	0.050	1.997	.02
Hypothesis 4a	Transition cost → attitude	−0.153	0.055	2.779	.003
Hypothesis 4b	Transition cost → intention	−0.163	0.037	4.357	<.001
Hypothesis 5a	Subjective well-being → perceived ease of use	0.260	0.053	4.927	<.001
Hypothesis 5b	Subjective well-being → perceived usefulness	0.261	0.057	4.531	<.001
Hypothesis 5c	Subjective well-being → intention	0.448	0.047	9.582	<.001
Hypothesis 6a	Inertia → perceived ease of use	−0.246	0.041	5.972	<.001
Hypothesis 6b	Inertia → perceived usefulness	0.024	0.034	0.778	.22
Hypothesis 7a	Availability → ease of use	0.420	0.048	8.725	<.001
Hypothesis 7b	Availability → perceived usefulness	0.323	0.047	6.818	<.001
Hypothesis 8	Trust → perceived usefulness	0.156	0.060	2.620	.004

[a]We used 95% CI with a bootstrapping of 5000.

First, we examined the influence of the 4 predictors on intention. Attitude (β=0.242; *P*<.001), perceived usefulness (β=0.100, *P*=.02), transition cost (β=−0.163; *P*<.001), and subjective well-being (β=0.448; *P*<.001) were all associated with intention. This means that hypotheses 1, 3b, 4b, and 5c were supported.

Second, the impact of the 3 predictors on attitude was investigated. Perceived ease of use (β=0.209; *P*<.001), perceived usefulness (β=0.456; *P*<.001), and transition cost (β=−0.153; *P*<.001) were linked to attitude, affirming support for hypotheses 2b, 3a, and 4a.

Third, perceived usefulness was assessed in relation to 5 predictors. Perceived ease of use (β=0.271; *P*<.001), availability (β=0.323; *P*<.001), subjective well-being (β=0.261; *P*<.001), and trust (β=0.156; *P*<.001) exhibited positive associations, while inertia (β=.024, *P*=.22) showed no significant relationship with perceived usefulness. Consequently, hypotheses 2a, 5b, 7b, and 8 were supported, while hypothesis 6b was not.

Finally, we examined the impact of availability (β=.420; *P*<.001) and subjective well-being (β=0.260; *P*<.001) on perceived ease of use, finding that both were positively related, while inertia (β=−0.246; *P*<.001) was negatively related. This supported hypotheses 5a, 6a, and 7a.

Consistent with the results of previous hypothesis testing, the bias-corrected 95% CIs for all hypotheses (except 6a) confirmed that they did not encompass 0, indicating support for these hypotheses (Multimedia Appendix 3). Furthermore, the variances for intention to adopt (R^2=0.650), attitude (R^2=0.502), perceived usefulness (R^2=.654), and perceived ease of use (R^2=0.475) were generally above 33%. This suggests the model possesses a moderate predictive capacity.

To assess the predictive relevance of the model, Shmueli et al [95] suggested the use of PLSpredict, a holdout sample-based method that produces case-level predictions at either the item or construct level. We used PLSpredict with a 10-fold procedure to assess predictive capability. As demonstrated in Table 4, all the errors associated with the PLS model were lower compared to the linear model, indicating the robust predictive power of our model. Therefore, we can confidently assert that our model exhibits strong predictive capacity.

Table 4. PLSpredict results.

Items for the construct intention	PLS[a]	LM[b]	PLS-LM
INT1[c]	0.734	0.735	−0.001
INT2	0.856	0.866	−0.010
INT3	0.767	0.769	−0.002

[a]PLS: partial least squares.

[b]LM: linear model.

[c]INT: intention.

XSL•FO
RenderX

Results of the Importance-Performance Map Analysis

The importance-performance map analysis (IPMA) of older adults' intention to adopt telehealth services aimed to expand upon the results of the structural model by evaluating the performance of each variable. As mentioned by Hair et al [91], areas requiring management attention are those with high importance but poor performance on a specific endogenous latent variable. In our research, we assessed the impact of latent exogenous factors on the endogenous variable (ie, intention to adopt) in terms of their significance and performance. The results of this analysis are presented in Table 5.

Table 5. Importance-performance map analysis.

Constructs	Performance	Importance
Attitude	63.245	0.241
Perceived ease of use	60.779	0.108
Perceived usefulness	64.078	0.210
Transition cost	52.758	0.200
Inertia	69.450	0.021
Availability	70.735	0.113
Trust	64.529	0.033
Subjective well-being	53.330	0.532

The IPMA results showed that the most important factor was subjective well-being (0.532), followed by attitude (0.241), perceived usefulness (0.210), transition cost (0.200), availability (0.113), perceived ease of use (0.108), trust (0.033), and inertia (0.021).

On the basis of performance, availability (70.735) was the highest, followed by inertia (69.45), trust (64.53), perceived useful (64.078), attitude (63.22), ease of use (60.779), subjective well-being (53.33), and transition cost (52.758). Therefore, it is evident that subjective well-being, attitude, transition cost, and perceived usefulness play crucial roles in influencing the intention of older adults to adopt telehealth, as these constructs exhibit relatively higher total effects (importance) compared to other factors in the model. However, the performance of well-being, attitude, transition cost, and usefulness were relatively lower compared to other factors like availability, inertia, and trust.

In summary, to enhance the intention to adopt telehealth among older adults, managerial efforts should be primarily focused on addressing and emphasizing subjective well-being, while continuing to support and maintain positive attitudes and perceived usefulness. On the other hand, perceived ease of use should be given lower priority.

Discussion

Principal Findings

This study aims to enhance the existing TAM for the adoption of telehealth services among older adults in the Malaysian context. Apart from attitude and perceived usefulness, this study proposed 2 additional constructs, namely subjective well-being and transition cost, for assessing the intention of older adults to use telehealth. Furthermore, this study introduced availability, trust, inertia, and subjective well-being as antecedents of TAM constructs. The study found that subjective well-being is the most important factor in telehealth adoption, followed by attitude, transition cost, and perceived usefulness. Perceived ease of use, perceived usefulness, and transition cost substantially affected attitude. Perceived ease of use, availability, subjective well-being, and trust positively influenced perceived usefulness, while inertia did not. In addition, availability and subjective well-being were positively related to perceived ease of use, with inertia having a negative impact. IPMA results showed that subjective well-being was the most crucial factor for older adult users, while availability had the highest performance in telehealth services.

As expected, this study found that attitude positively affects intention, which is in line with the existing health care–related studies, such as those by Park et al [44], Papa et al [45], Rajak and Shaw [46], and Ahn and Park [47]. As per the results, the key TAM construct, perceived ease of use, positively influences older adults' perceived usefulness and attitude toward the telehealth system. These outcomes corroborate the findings of Rajak and Shaw [46] and Ahn and Park [47], which indicate that older adults will have a positive attitude toward telehealth and perceive it to be useful to them if telehealth is easy to use. On the other hand, perceived usefulness also positively influences older adults' attitude and intention toward telehealth. Specifically, perceived usefulness has a significantly larger influence on attitudes regarding adopting telehealth than perceived ease of use (0.456 vs 0.209). The rationale behind these findings may be due to the relevance of using telehealth for managing older adults' health. Hence, the usefulness of telehealth is the priority as compared to its ease of use. In other words, older adults are more likely to use telehealth services if they can provide useful features to them, such as improving their quality of life and offering better health care services.

Consistent with previous findings [36,57,58], these findings show that transition costs are driving forces that have a negative impact on the older adults' attitude and intention to use telehealth. According to the findings, older adults do not intend to use telehealth and will continue using the traditional method of obtaining health care services if they believe the time and effort required to learn telehealth is too high.

XSL•FO
RenderX

Besides, inertia was found to negatively influence the perceived ease of use but had no impact on the perceived usefulness. The significance of inertia's influence on perceived ease of use stems from the fact that ease of use is directly linked to how simple or user-friendly telehealth appears. When older adults experience inertia, they tend to resist adopting telehealth because of the additional effort required to learn and adapt to new technologies. This resistance (inertia) makes new systems seem more complex, thereby reducing the perceived ease of use of telehealth among older adults. However, the significance of inertia on perceived ease of use contradicts the findings of Tsai et al [36], where inertia had no effect on ease of use. This discrepancy may stem from the fact that the majority of respondents in the study by Tsai et al [36] were younger (>40 years) with higher digital literacy, while our study focuses on those aged ≥60 years. In contrast, the insignificant impact of inertia on perceived usefulness in this study is consistent with the findings of Tsai et al [36]. This is because perceived usefulness is more about the functional benefits the user gains from the technology. Hence, even if older adults perceive the system to be hard to use due to inertia, they might still acknowledge that the system provides value or could be beneficial in terms of performance and efficiency. This can explain why inertia does not affect perceived usefulness to the same extent.

Availability showed a positive influence on ease of use and perceived usefulness, which is consistent with existing studies, such as those by Chang et al [74] and Tsai et al [36]. However, these findings partially conflict with the findings of Wu et al [72], which revealed that availability only influences the perceived usefulness and has no impact on the perceived ease of use. This revealed different perspectives between patients and hospital professionals toward the importance of availability in determining telehealth adoption. From the older adult users' perspective, the study's findings revealed that the availability of assistance, such as providing responses to assist them in overcoming obstacles when using telehealth, is very important, which would affect their perceived ease of use and usefulness of telehealth.

This study revealed that well-being positively influences perceived ease of use, perceived usefulness, and intention toward telehealth. These outcomes are expected as people always seek pleasurable experiences that can enhance their well-being when adopting new technology [39,40]. These findings are also consistent with previous results, which showed that an individual's perceived well-being toward technology will positively influence their perceived ease of use, usefulness, and intention [32,33,63,64].

Furthermore, this study also revealed that older adults' trust positively influences the perceived usefulness and attitude of telehealth. This finding is predictable because most older adults prefer to receive services in-person and may have trust issues with telehealth's ability to replace in-person consultations and physical health assessments, so they may perceive high risks and lower trust toward telehealth. The study validates the findings of Rajak and Shaw [46], who revealed the crucial role of trust in the health care technology adoption among older adults.

Conclusions and Implications

Given the rapid growth of the older adult health care industry, improving health care–related services is a promising opportunity. Hence, this study was conducted to investigate the factors influencing the intention of older adults to use telehealth services by extending the TAM model. The model has been stretched to incorporate factors like availability, transition cost, trust, inertia, and well-being. From the literature, it was observed that a considerable number of studies have been conducted on telehealth. However, there is no evidence in the literature of incorporating these factors and further assessing the model in the Malaysian context. In addition, although several papers have adopted the TAM to investigate telehealth adoption, there are limited studies investigating the formation of the TAM key constructs (perceived ease of use and perceived usefulness) in depth from the older adults' perspective. With the introduction of these constructs into the TAM model, this study can contribute to the literature by providing a better understanding of factors affecting older adults' intention to use telehealth in the Malaysian context.

This study has several practical implications. First, the positive impact of attitude toward older adults' intention on telehealth use provides valuable implications to the health care centers and managers. They should create an environment that can ensure older adults have a positive attitude toward telehealth. On the basis of the findings, positive attitude toward telehealth can be enhanced in several ways, for example, when the older adults perceive it to be useful and easy to use. Hence, the telehealth developer and health care personnel should carefully assess telehealth's usefulness in assisting older adults' health care before it is introduced to them.

In addition, telehealth developers should also design an older adult–friendly interface that is easier for them to navigate and understand, as the older adults are not as digitally literate as the younger generation. For example, a case study by Pires et al [96] emphasizes the importance of simplicity and ease of use in successfully implementing the VITASENIOR-MT telehealth system for older adult users. By using television as the primary interface, the telehealth system became more accessible and easier by integrating familiar technology into the homes of older adults. Health professionals provided valuable feedback on design and usability, enhancing the system's effectiveness as they remotely monitored patients. This case exemplifies the best practices in telehealth design, highlighting the significance of user-centered development that prioritizes ease of use for older adults. In addition, health professionals are encouraged to actively engage in the development process to further improve the system's effectiveness. A clear step-by-step video tutorial, easy interactive dialogue, and straightforward click-through procedures for making health care appointments can be introduced. Moreover, the government or relevant authorities can provide older adults with training on how to use telehealth services to improve their skills and confidence in accessing online services.

The results from this study revealed that if the transition cost required when using telehealth is perceived as high, it will cause people to have a negative attitude and intention toward

XSL•FO
RenderX

telehealth. Furthermore, results also indicate that older adults tend to keep using their habitual ways to obtain health care services, which is known as inertia in this study. This inertia negatively affects their perceived ease of use and usefulness of telehealth. The rationale behind these results might be related to the fact that most commonly available smartphone devices, such as tablet computers and smartphones, as well as mobile internet subscription plans, are generally sold at a high price. Therefore, older adult mobile users only seek to access the internet through free Wi-Fi networks due to the high cost of using mobile devices. Given these considerations, it is recommended that telecommunications companies collaborate with government agencies to launch programs that offer special low prices, rebates, or government-funded subsidies for older adults when they purchase a mobile device and sign up for mobile internet packages. For instance, telecommunication companies in Taiwan have effectively held a promotional campaign for older adults to access the internet for free for a certain period [97]. Through a short, free trial, older adults' increasing internet accessibility is an effective way to encourage them to learn and become familiar with mobile apps and to reduce their inertia by providing a greater understanding of the mobile services' value and usefulness, particularly the telehealth services.

In addition to utilitarian-oriented benefits like usefulness and ease of use, which have been frequently discussed in earlier literature, policy makers and managers must also highlight the affective components like well-being [33]. Surprisingly, the result revealed that well-being is the most important factor that affects older adults' intention for telehealth use. Hence, practitioners should pay more attention to developing a sense of well-being enhancement if older adults use telehealth services. For instance, relevant authorities can create an online community of innovative users where the older adults can share their pleasant, positive, and healthy experiences to highlight the well-being provided by telehealth, and hence indirectly generate positive word of mouth that motivates the potential users to use telehealth. On the other hand, findings also revealed that availability positively influences older adults' perceived ease of use and usefulness of telehealth. Therefore, technology developers, manufacturers, or organizations should always be prepared to respond and assist the older adult in overcoming challenges and obstacles when using telehealth services, particularly during the early stages of adoption.

Furthermore, the study's findings also revealed that older adults' trust positively affects the perceived usefulness. Incorporating trust as a key determinant of perceived usefulness is particularly important for older adults adopting telehealth. The literature suggests that trust in telehealth can be fostered through mechanisms, such as confidence in the competence of health care providers [17,81,82] and the reliability of the information provided [98]. In practice, trust can be strengthened by implementing ongoing professional development and telehealth-specific training for health care providers. This can ensure they remain updated on best practices and patient care, thus boosting patient confidence in their competence. In addition, regularly updating health information and cross-referencing it with evidence-based guidelines reinforces the reliability of the information provided on telehealth platforms.

Moreover, government authorities and telehealth platform developers play a critical role in improving trust in using telehealth services. As a result, it is suggested that government agencies implement clear and detailed support policies for telehealth services. For instance, the government should strengthen the laws to ensure that the platform developers ensure the confidentiality of patient data and data security while health care institutions and staff are qualified. The introduction of national policies shows the government's support for telehealth services, which could increase the public's trust toward them [99]. In addition, managers should assist in the development of telehealth platforms by implementing and promoting features and guidelines on patient data confidentiality and privacy protection to establish customer trust. Furthermore, the use of secondary data and security policies should be communicated and advertised by companies to increase trust. In addition, the government should also provide sufficient budget allocation and training for health care professionals to use telehealth services so they can provide better telehealth services for patients and, in turn, increase their trust in telehealth services.

Limitations and Suggestions for Future Research

Like all research endeavors, this study is not immune to limitations. The first constraint becomes evident through the limited number of determinants examined concerning the intention to use telehealth services. We recognize that numerous potential factors remain unexplored, particularly within the realm of customization.

Second, our research adopts a cross-sectional survey design approach. Given the rapid evolution of technology and the increasing familiarity of consumers with telehealth, this approach may impact their evolving perceptions of telehealth. In addition, consumers' lifestyles are dynamic, undergoing continuous changes across different life stages. Consequently, a longitudinal study would yield valuable insights into the sustained use of telehealth services. In addition, future research could explore how different types of telehealth services (eg, video consultations vs remote monitoring) specifically impact adoption rates among older adults.

This study used convenient and snowball sampling to reach a hard-to-access population. Although the data represent various ethnicities and regions across Malaysia, the uneven distribution of responses may still introduce a certain level of bias. To enhance representativeness, we recommend using random or stratified sampling methods in future research.

Another limitation is the potential for self-report biases, as participants may have provided socially desirable responses rather than their true thoughts or behaviors. While we mitigated this by ensuring survey anonymity, self-report biases may still persist and should be considered when interpreting the findings. Future research should incorporate objective measures, such as tracking actual telehealth use data, to complement self-reported data and provide a more accurate assessment.

Although this study meets the sample size requirements based on power analysis using G-Power, it may not fully capture the

XSL•FO
RenderX

diversity of the older adult population, potentially limiting the generalizability of the findings. Caution is advised when applying these results to the broader older adult population. Future studies should include larger, more diverse samples across different socioeconomic backgrounds, geographic locations, and health conditions to strengthen the robustness and applicability of the findings.

Acknowledgments

The authors acknowledge financial support from the Ministry of Higher Education, Malaysia (FRGS/1/2019/SS06/MMU/02/3), and Multimedia University (MMUI/220112).

Authors' Contributions

SHT and YYY: conceptualization, methodology, writing, analysis and interpretation of data, and software. SHT: funding acquisition. SKT: review and editing and visualization. CKW: acquisition of data. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Overview of the instrument.
[DOCX File , 21 KB - aging_v8i1e60936_app1.docx]

Multimedia Appendix 2
Respondents' information (N=119).
[DOCX File , 15 KB - aging_v8i1e60936_app2.docx]

Multimedia Appendix 3
Bias-corrected interval, f2, and variance inflation factor.
[DOCX File , 17 KB - aging_v8i1e60936_app3.docx]

References

1. Kruse CS, Krowski N, Rodriguez B, Tran L, Vela J, Brooks M. Telehealth and patient satisfaction: a systematic review and narrative analysis. BMJ Open 2017 Aug 03;7(8):e016242 [FREE Full text] [doi: 10.1136/bmjopen-2017-016242] [Medline: 28775188]
2. Khosravi P, Ghapanchi AH. Investigating the effectiveness of technologies applied to assist seniors: a systematic literature review. Int J Med Inform 2016 Jan;85(1):17-26. [doi: 10.1016/j.ijmedinf.2015.05.014] [Medline: 26216463]
3. Pinto A, Almeida JP, Pinto S, Pereira J, Oliveira AG, de Carvalho M. Home telemonitoring of non-invasive ventilation decreases healthcare utilisation in a prospective controlled trial of patients with amyotrophic lateral sclerosis. J Neurol Neurosurg Psychiatry 2010 Nov;81(11):1238-1242. [doi: 10.1136/jnnp.2010.206680] [Medline: 20826878]
4. Gellis ZD, Kenaley B, McGinty J, Bardelli E, Davitt J, Ten Have T. Outcomes of a telehealth intervention for homebound older adults with heart or chronic respiratory failure: a randomized controlled trial. Gerontologist 2012 Aug;52(4):541-552. [doi: 10.1093/geront/gnr134] [Medline: 22241810]
5. Chou CC, Chang CP, Lee TT, Chou HF, Mills ME. Technology acceptance and quality of life of the elderly in a telecare program. Comput Inform Nurs 2013 Jul;31(7):335-342. [doi: 10.1097/NXN.0b013e318295e5ce] [Medline: 23728446]
6. Nikou S, Agahari W, Keijzer-Broers W, de Reuver M. Digital healthcare technology adoption by elderly people: a capability approach model. Telemat Inform 2020 Oct;53:101315. [doi: 10.1016/j.tele.2019.101315]
7. Hanson VL. Influencing technology adoption by older adults. Interact Comput 2010 Nov;22(6):502-509. [doi: 10.1016/j.intcom.2010.09.001]
8. Rho MJ, Choi IY, Lee J. Predictive factors of telemedicine service acceptance and behavioral intention of physicians. Int J Med Inform 2014 Aug;83(8):559-571. [doi: 10.1016/j.ijmedinf.2014.05.005] [Medline: 24961820]
9. Lee C, Coughlin JF. PERSPECTIVE: older adults' adoption of technology: an integrated approach to identifying determinants and barriers. J Product Innov Manag 2014 Jun 03;32(5):747-759. [doi: 10.1111/jpim.12176]
10. Current population estimates, Malaysia, 2023. Ministry of Economy, Department of Statistics Malaysia. 2023. URL: https://www.dosm.gov.my/portal-main/release-content/current-population-estimates-malaysia----2023 [accessed 2024-04-12]
11. Tan SH, Wong CK, Yap YY, Tan SK. Factors influencing telemedicine adoption among physicians in the Malaysian healthcare system: a revisit. Digit Health 2024 Jun 06;10:20552076241257050 [FREE Full text] [doi: 10.1177/20552076241257050] [Medline: 38854922]

XSL•FO
RenderX

12. Lawrence K, Nov O, Mann D, Mandal S, Iturrate E, Wiesenfeld B. The impact of telemedicine on physicians' after-hours electronic health record "work outside work" during the COVID-19 pandemic: retrospective cohort study. JMIR Med Inform 2022 Jul 28;10(7):e34826 [FREE Full text] [doi: 10.2196/34826] [Medline: 35749661]
13. Teh PL, Kwok AO, Cheong WL, Lee S. Insights into the use of a digital healthy aging coach (AGATHA) for older adults from Malaysia: app engagement, usability, and impact study. JMIR Form Res 2024 May 21;8:e54101 [FREE Full text] [doi: 10.2196/54101] [Medline: 38772022]
14. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: an integrative systematic review. BMC Public Health 2024 Feb 16;24(1):506 [FREE Full text] [doi: 10.1186/s12889-024-18036-5] [Medline: 38365698]
15. Ahmed MH, Awol SM, Kanfe SG, Hailegebreal S, Debele GR, Dube GN, et al. Willingness to use telemedicine during COVID-19 among health professionals in a low income country. Inform Med Unlocked 2021;27:100783 [FREE Full text] [doi: 10.1016/j.imu.2021.100783] [Medline: 34778509]
16. Elimelech OC, Ferrante S, Josman N, Meyer S, Lunardini F, Gómez-Raja J, et al. Technology use characteristics among older adults during the COVID-19 pandemic: a cross-cultural survey. Technol Soc 2022 Nov;71:102080 [FREE Full text] [doi: 10.1016/j.techsoc.2022.102080] [Medline: 35991080]
17. Catapan SD, Haydon HM, Hickman IJ, Webb L, Isbel N, Johnson D, et al. Trust and confidence in using telehealth in people with chronic kidney disease: a cross-sectional study. J Telemed Telecare 2023 Nov 26;29(10_suppl):16S-23S. [doi: 10.1177/1357633x231202275]
18. Lin Y, Xu X, Liu Y, Alias H, Hu Z, Wong LP. Perception and acceptance of telemedicine use in health care among the general public in China: web-based cross-sectional survey. J Med Internet Res 2024 Jul 16;26:e53497 [FREE Full text] [doi: 10.2196/53497] [Medline: 39012687]
19. Zhang Y, Lee EW, Teo WP. Health-seeking behavior and its associated technology use: interview study among community-dwelling older adults. JMIR Aging 2023 May 04;6:e43709 [FREE Full text] [doi: 10.2196/43709] [Medline: 36996003]
20. Haimi M, Sergienko R. Adoption and use of telemedicine and digital health services among older adults in light of the COVID-19 pandemic: repeated cross-sectional analysis. JMIR Aging 2024 Apr 24;7:e52317 [FREE Full text] [doi: 10.2196/52317] [Medline: 38656768]
21. Ting CY, Abdul Halim NH, Ling JN, Tiong IK, Ahmad Shauki NI, Lee YF, et al. The use of a multi-disciplinary geriatric telemedicine service (TELEG) and its acceptance at a tertiary care centre in Malaysia. BMC Geriatr 2024 Feb 05;24(1):133 [FREE Full text] [doi: 10.1186/s12877-024-04676-0] [Medline: 38317117]
22. Ajzen I. From intentions to actions: a theory of planned behavior. In: Kuhl J, Beckmann J, editors. Action Control. Berlin, Heidelberg: Springer; 1985.
23. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manag Sci 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
24. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manag Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
25. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478. [doi: 10.2307/30036540]
26. Venkatesh V, Bala H. Technology Acceptance Model 3 and a research agenda on interventions. Decis Sci 2008 May 09;39(2):273-315. [doi: 10.1111/j.1540-5915.2008.00192.x]
27. Harst L, Lantzsch H, Scheibe M. Theories predicting end-user acceptance of telemedicine use: systematic review. J Med Internet Res 2019 May 21;21(5):e13117 [FREE Full text] [doi: 10.2196/13117] [Medline: 31115340]
28. Heinsch M, Wyllie J, Carlson J, Wells H, Tickner C, Kay-Lambkin F. Theories informing eHealth implementation: systematic review and typology classification. J Med Internet Res 2021 May 31;23(5):e18500 [FREE Full text] [doi: 10.2196/18500] [Medline: 34057427]
29. Rouidi M, Elouadi AE, Hamdoune A, Choujtani K, Chati A. TAM-UTAUT and the acceptance of remote healthcare technologies by healthcare professionals: a systematic review. Inform Med Unlock 2022;32:101008. [doi: 10.1016/j.imu.2022.101008]
30. Yap YY, Tan SH, Choon SW. Elderly's intention to use technologies: a systematic literature review. Heliyon 2022 Jan;8(1):e08765 [FREE Full text] [doi: 10.1016/j.heliyon.2022.e08765] [Medline: 35128090]
31. Kamal SA, Shafiq M, Kakria P. Investigating acceptance of telemedicine services through an extended technology acceptance model (TAM). Technol Soc 2020 Feb;60:101212. [doi: 10.1016/j.techsoc.2019.101212]
32. Yap YY, Tan SH, Tan SK, Choon SW. Integrating the capability approach and technology acceptance model to explain the elderly's use intention of online grocery shopping. Telematics Inform 2022 Aug;72:101842. [doi: 10.1016/j.tele.2022.101842]
33. Attié E, Meyer-Waarden L. The acceptance and usage of smart connected objects according to adoption stages: an enhanced technology acceptance model integrating the diffusion of innovation, uses and gratification and privacy calculus theories. Technol Forecast Soc Change 2022 Mar;176:121485. [doi: 10.1016/j.techfore.2022.121485]

XSL•FO
RenderX

34. Zhou M, Zhao L, Kong N, Campy KS, Qu S, Wang S. Factors influencing behavior intentions to telehealth by Chinese elderly: an extended TAM model. Int J Med Inform 2019 Jun;126:118-127. [doi: 10.1016/j.ijmedinf.2019.04.001] [Medline: 31029253]
35. Klingberg A, Sawe HR, Hammar U, Wallis LA, Hasselberg M. m-Health for burn injury consultations in a low-resource setting: an acceptability study among health care providers. Telemed J E Health 2020 Apr 01;26(4):395-405 [FREE Full text] [doi: 10.1089/tmj.2019.0048] [Medline: 31161967]
36. Tsai JM, Cheng MJ, Tsai HH, Hung SW, Chen YL. Acceptance and resistance of telehealth: the perspective of dual-factor concepts in technology adoption. Int J Inf Manag 2019 Dec;49:34-44. [doi: 10.1016/j.ijinfomgt.2019.03.003]
37. Steele R, Lo A, Secombe C, Wong YK. Elderly persons' perception and acceptance of using wireless sensor networks to assist healthcare. Int J Med Inform 2009 Dec;78(12):788-801. [doi: 10.1016/j.ijmedinf.2009.08.001] [Medline: 19717335]
38. Katz E, Blumler JG, Gurevitch M. Utilization of mass communication by the individual. In: Blumler JG, Katz E, editors. The Uses of Mass Communications: Current Perspectives on Gratifications Research. Thousand Oaks, CA: Sage Publications; 1974:19-31.
39. Li C, Shi X, Dang J. Online communication and subjective well-being in Chinese college students: the mediating role of shyness and social self-efficacy. Comput Hum Behav 2014 May;34:89-95. [doi: 10.1016/j.chb.2014.01.032]
40. Yoon S. Does social capital affect SNS usage? A look at the roles of subjective well-being and social identity. Comput Hum Behav 2014 Dec;41:295-303. [doi: 10.1016/j.chb.2014.09.043]
41. Samuelson W, Zeckhauser R. Status quo bias in decision making. J Risk Uncertainty 1988 Mar;1(1):7-59. [doi: 10.1007/bf00055564]
42. Yap YY, Tan SH, Tan SK, Choon SW. Online grocery shopping intention: elderly's perspective in Malaysia. Heliyon 2023 Oct;9(10):e20827 [FREE Full text] [doi: 10.1016/j.heliyon.2023.e20827] [Medline: 37916123]
43. Ajzen I, Fishbein M. A Bayesian analysis of attribution processes. Psychol Bull 1975 Mar;82(2):261-277. [doi: 10.1037/h0076477]
44. Park E, Kim KJ, Kwon SJ. Understanding the emergence of wearable devices as next-generation tools for health communication. Inf Technol People 2016 Nov 07;29(4):717-732. [doi: 10.1108/itp-04-2015-0096]
45. Papa A, Mital M, Pisano P, Del Giudice M. E-health and wellbeing monitoring using smart healthcare devices: an empirical investigation. Technol Forecast Soc Change 2020 Apr;153:119226 [FREE Full text] [doi: 10.1016/j.techfore.2018.02.018]
46. Rajak M, Shaw K. An extension of technology acceptance model for mHealth user adoption. Technol Soc 2021 Nov;67:101800. [doi: 10.1016/j.techsoc.2021.101800]
47. Ahn H, Park E. Determinants of consumer acceptance of mobile healthcare devices: An application of the concepts of technology acceptance and coolness. Telematics Inform 2022 May;70:101810. [doi: 10.1016/j.tele.2022.101810]
48. Yang HD, Yoo Y. It's all about attitude: revisiting the technology acceptance model. Decis Support Syst 2004 Oct;38(1):19-31. [doi: 10.1016/S0167-9236(03)00062-9]
49. Lazaro MJ, Lim J, Kim SH, Yun MH. Wearable technologies: acceptance model for smartwatch adoption among older adults. In: Proceedings of the 6th International Conference on Human Aspects of IT for the Aged Population. Technologies, Design and User Experience. 2020 Presented at: ITAP 2020; July 19-24, 2020; Copenhagen, Denmark. [doi: 10.1007/978-3-030-50252-2_23]
50. Pan S, Jordan-Marsh M. Internet use intention and adoption among Chinese older adults: from the expanded technology acceptance model perspective. Comput Hum Behav 2010 Sep;26(5):1111-1119. [doi: 10.1016/j.chb.2010.03.015]
51. Braun MT. Obstacles to social networking website use among older adults. Comput Hum Behav 2013 May;29(3):673-680. [doi: 10.1016/j.chb.2012.12.004]
52. Li J, Ma Q, Chan AH, Man SS. Health monitoring through wearable technologies for older adults: smart wearables acceptance model. Appl Ergon 2019 Feb;75:162-169. [doi: 10.1016/j.apergo.2018.10.006] [Medline: 30509522]
53. Hoque R, Sorwar G. Understanding factors influencing the adoption of mHealth by the elderly: an extension of the UTAUT model. Int J Med Inform 2017 May;101:75-84. [doi: 10.1016/j.ijmedinf.2017.02.002] [Medline: 28347450]
54. Motamedi S, Masrahi A, Bopp T, Wang J. Different level automation technology acceptance: older adult driver opinion. Transp Res Part F 2021 Jul;80:1-13. [doi: 10.1016/j.trf.2021.03.010]
55. Cimperman M, Makovec Brenčič M, Trkman P. Analyzing older users' home telehealth services acceptance behavior-applying an extended UTAUT model. Int J Med Inform 2016 Jun;90:22-31. [doi: 10.1016/j.ijmedinf.2016.03.002] [Medline: 27103194]
56. Kim HW, Kankanhalli A. Investigating user resistance to information systems implementation: a status quo bias perspective. MIS Q 2009;33(3):567-582. [doi: 10.2307/20650309]
57. Kim H. The effects of switching costs on user resistance to enterprise systems implementation. IEEE Trans Eng Manag 2011 Aug;58(3):471-482. [doi: 10.1109/tem.2010.2089630]
58. Hsieh PJ. An empirical investigation of patients' acceptance and resistance toward the health cloud: the dual factor perspective. Comput Hum Behav 2016 Oct;63:959-969. [doi: 10.1016/j.chb.2016.06.029]
59. Diener E. Subjective well-being. Psychol Bull 1984;95(3):542-575. [doi: 10.1037/0033-2909.95.3.542]
60. Diener E, Chan MY. Happy people live longer: subjective well-being contributes to health and longevity. Appl Psychol Health Well Being 2011;3(1):1-43. [doi: 10.1111/j.1758-0854.2010.01045.x]

XSL•FO
RenderX

61. Al-Jabri I, Sohail MS. Mobile banking adoption: application of diffusion of innovation theory. J Electron Commer Res 2012;13(4):379-391.
62. Wu J, Lu X. Effects of extrinsic and intrinsic motivators on using utilitarian, hedonic, and dual-purposed information systems: a meta-analysis. J Assoc Inf Syst 2013 Mar;14(3):153-191. [doi: 10.17705/1jais.00325]
63. Andreasen AR, Goldberg ME, Sirgy MJ. Foundational research on consumer welfare. In: Mick DG, Pettigrew S, Pechmann C, Ozanne JL, editors. Transformative Consumer Research for Personal and Collective Well-Being. New York, NY: Routledge; 2011.
64. Davis B, Pechmann C. Introduction to the special issue on transformative consumer research: developing theory to mobilize efforts that improve consumer and societal well-being. J Bus Res 2013 Aug;66(8):1168-1170. [doi: 10.1016/j.jbusres.2012.08.008]
65. Polites GL, Karahanna E. Shackled to the status quo: the inhibiting effects of incumbent system habit, switching costs, and inertia on new system acceptance. MIS Q 2012;36(1):21-42. [doi: 10.2307/41410404]
66. Bem DJ. Self-perception theory. In: Gawronski B, editor. Advances in Experimental Social Psychology. Mahwah, NJ: Lawrence Erlbaum; 1972:1-62.
67. Petty RE, Cacioppo JT. Attitudes And Persuasion: Classic and Contemporary Approaches. Boulder, CO: Westview Press Inc; 1996.
68. Atchley RC. A continuity theory of normal aging. Gerontologist 1989 Apr 01;29(2):183-190. [doi: 10.1093/geront/29.2.183] [Medline: 2519525]
69. Venkatesh V. Determinants of perceived ease of use: integrating control, intrinsic motivation, and emotion into the technology acceptance model. Inf Syst Res 2000 Dec;11(4):342-365. [doi: 10.1287/isre.11.4.342.11872]
70. Zhou T. Examining location-based services usage from the perspectives of unified theory of acceptance and use of technology and privacy risk. J Electron Commer Res 2012;13(2):135-144.
71. Alaiad A, Zhou L. The determinants of home healthcare robots adoption: an empirical investigation. Int J Med Inform 2014 Nov;83(11):825-840. [doi: 10.1016/j.ijmedinf.2014.07.003] [Medline: 25132284]
72. Wu IL, Li JY, Fu CY. The adoption of mobile healthcare by hospital's professionals: an integrative perspective. Decis Support Syst 2011 Jun;51(3):587-596. [doi: 10.1016/j.dss.2011.03.003]
73. Hong S, Tam K. Understanding the adoption of multipurpose information appliances: the case of mobile data services. Inf Syst Res 2006 Jun;17(2):162-179 [FREE Full text] [doi: 10.1287/isre.1060.0088]
74. Chang YZ, Ko CY, Hsiao CJ, Chen RJ, Yu CW, Cheng YW, et al. Understanding the determinants of implementing telehealth systems: a combined model of the theory of planned behavior and the technology acceptance model. J Appl Sci 2015 Jan 15;15(2):277-282. [doi: 10.3923/jas.2015.277.282]
75. Gefen D, Karahanna E, Straub DW. Trust and TAM in online shopping: an integrated model. MIS Q 2003;27(1):51-90. [doi: 10.2307/30036519]
76. Alam MZ, Hu W, Kaium MA, Hoque MR, Alam MM. Understanding the determinants of mHealth apps adoption in Bangladesh: a SEM-neural network approach. Technol Soc 2020 May;61:101255. [doi: 10.1016/j.techsoc.2020.101255]
77. Naik N, Ibrahim S, Sircar S, Patil V, Hameed BM, Rai BP, et al. Attitudes and perceptions of outpatients towards adoption of telemedicine in healthcare during COVID-19 pandemic. Ir J Med Sci 2022 Aug 16;191(4):1505-1512 [FREE Full text] [doi: 10.1007/s11845-021-02729-6] [Medline: 34402031]
78. Ateriya N, Saraf A, Meshram VP, Setia P. Telemedicine and virtual consultation: the Indian perspective. Natl Med J India 2018;31(4):215-218. [doi: 10.4103/0970-258x.258220]
79. Scott Kruse C, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: a systematic review. J Telemed Telecare 2018 Jan;24(1):4-12 [FREE Full text] [doi: 10.1177/1357633X16674087] [Medline: 29320966]
80. Li X, Hess TJ, Valacich JS. Why do we trust new technology? A study of initial trust formation with organizational information systems. J Strateg Inf Syst 2008 Mar;17(1):39-71. [doi: 10.1016/j.jsis.2008.01.001]
81. Chew E, Teo SH, Tang WE, Ng DW, Koh GC, Teo VH. Trust and uncertainty in the implementation of a pilot remote blood pressure monitoring program in primary care: qualitative study of patient and health care professional views. JMIR Hum Factors 2023 Jan 05;10:e36072 [FREE Full text] [doi: 10.2196/36072] [Medline: 36602847]
82. Orrange S, Patel A, Mack WJ, Cassetta J. Patient satisfaction and trust in telemedicine during the COVID-19 pandemic: retrospective observational study. JMIR Hum Factors 2021 Apr 22;8(2):e28589 [FREE Full text] [doi: 10.2196/28589] [Medline: 33822736]
83. Zhang X, Zaman BU. Adoption mechanism of telemedicine in underdeveloped country. Health Informatics J 2020 Jun 30;26(2):1088-1103 [FREE Full text] [doi: 10.1177/1460458219868353] [Medline: 31566450]
84. Wu D, Gu H, Gu S, You H. Individual motivation and social influence: a study of telemedicine adoption in China based on social cognitive theory. Health Policy Technol 2021 Sep;10(3):100525. [doi: 10.1016/j.hlpt.2021.100525]
85. Ringle CM, Wende S, Becker JM. SmartPLS 3. SmartPLS GmbH. 2015. URL: https://www.smartpls.com [accessed 2023-10-12]

XSL•FO
RenderX

86. Chin WW, Marcolin BL, Newsted PR. A partial least squares latent variable modeling approach for measuring interaction effects: results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. Inf Syst Res 2003 Jun;14(2):189-217. [doi: 10.1287/isre.14.2.189.16018]
87. Wu ML. Structural Equation Model: Operation and Application of AMOS. Chongqing, China: Chongqing University Press; 2009.
88. Hair J, Hult GT, Ringle CM, Sarstedt M. A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). Thousand Oaks, CA: SAGE Publications; 2016.
89. Henseler J, Ringle CM, Sarstedt M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. J Acad Mark Sci 2014 Aug 22;43(1):115-135. [doi: 10.1007/s11747-014-0403-8]
90. Franke G, Sarstedt M. Heuristics versus statistics in discriminant validity testing: a comparison of four procedures. Internet Res 2019 Jun 03;29(3):430-447. [doi: 10.1108/intr-12-2017-0515]
91. Hair JF, Risher JJ, Sarstedt M, Ringle CM. When to use and how to report the results of PLS-SEM. Eur Bus Rev 2019 Jan 14;31(1):2-24. [doi: 10.1108/ebr-11-2018-0203]
92. Ramayah T, Cheah JH, Chuah F, Ting H, Memon MA. Partial Least Squares Structural Equation Modeling (PLS-SEM) Using SmartPLS 3.0: An Updated and Practical Guide to Statistical Analysis. Kuala Lumpur, Malaysia: Pearson; 2016.
93. Hair J, Alamer A. Partial Least Squares Structural Equation Modeling (PLS-SEM) in second language and education research: guidelines using an applied example. Res Methods Appl Linguist 2022 Dec;1(3):100027. [doi: 10.1016/j.rmal.2022.100027]
94. Hahn ED, Ang SH. From the editors: new directions in the reporting of statistical results in the Journal of World Business. J World Bus 2017 Feb;52(2):125-126. [doi: 10.1016/j.jwb.2016.12.003]
95. Shmueli G, Sarstedt M, Hair JF, Cheah JH, Ting H, Vaithilingam S, et al. Predictive model assessment in PLS-SEM: guidelines for using PLSpredict. Eur J Mark 2019 Nov 11;53(11):2322-2347. [doi: 10.1108/ejm-02-2019-0189]
96. Pires G, Lopes A, Correia P, Almeida L, Oliveira L, Panda R, et al. Usability of a telehealth solution based on TV interaction for the elderly: the VITASENIOR-MT case study. Univers Access Inf Soc 2023 Jan 17;22(2):525-536 [FREE Full text] [doi: 10.1007/s10209-021-00859-3] [Medline: 35069063]
97. Yang HL, Lin SL. The reasons why elderly mobile users adopt ubiquitous mobile social service. Comput Hum Behav 2019 Apr;93:62-75. [doi: 10.1016/j.chb.2018.12.005]
98. Ghani M, Adler C, Yeung H. Patient factors associated with interest in teledermatology: cross-sectional survey. JMIR Dermatol 2021 May 10;4(1):e21555 [FREE Full text] [doi: 10.2196/21555] [Medline: 37625162]
99. Goins RT, Kategile U, Dudley KC. Telemedicine, rural elderly, and policy issues. J Aging Soc Policy 2001 Oct 02;13(4):53-71. [doi: 10.1300/j031v13n04_04] [Medline: 12418382]

Abbreviations

HTMT: heterotrait-monotrait
IPMA: importance-performance map analysis
PLS-SEM: partial least squares structural equation modeling
TAM: technology acceptance model

Edited by Y Luo; submitted 26.05.24; peer-reviewed by F Abdelfattah, S Mohezar; comments to author 16.07.24; revised version received 24.10.24; accepted 29.11.24; published 24.03.25.

Please cite as:
Tan SH, Yap YY, Tan SK, Wong CK
Determinants of Telehealth Adoption Among Older Adults: Cross-Sectional Survey Study
JMIR Aging 2025;8:e60936
URL: https://aging.jmir.org/2025/1/e60936
doi:10.2196/60936
PMID:40126531

©Siow-Hooi Tan, Yee-Yann Yap, Siow-Kian Tan, Chee-Kuan Wong. Originally published in JMIR Aging (https://aging.jmir.org), 24.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Gender Differences in the Digital Divide, Digital Back-Feeding, and Health-Related Quality of Life Among Rural Older Adults: Cross-Sectional Study

Xin Che[1,2,3,4], PhD; Shujun Chai[5], PhD; Dan Zhao[1,2,3,4], PhD; Shirong Chen[1,2,3,4], MSc; Chengchao Zhou[1,2,3,4,6], PhD

[1]Department of Social Medicine and Health Management, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan, China
[2]NHC Key Lab of Health Economics and Policy Research, Shandong University, Jinan, China
[3]Center for Health Management and Policy Research, Shandong University, Jinan, China
[4]Institute of Health and Elderly Care, Shandong University, Jinan, China
[5]School of Population and Health, Renmin University of China, Beijing, China
[6]Advanced Medical Research Institute, Shandong University, Jinan, China

Corresponding Author:
Chengchao Zhou, PhD
Department of Social Medicine and Health Management
School of Public Health, Cheeloo College of Medicine
Shandong University
44 Wen-hua-xi Road
Jinan, 250012
China
Phone: 86 531 8838 1567
Fax: 86 531 8838 2553
Email: zhouchengchao@sdu.edu.cn

Abstract

Background: The digital divide has loomed as a global public issue in recent years. However, evidence is limited regarding whether the digital divide is associated with health-related quality of life (HRQOL) and whether digital back-feeding would buffer this association.

Objective: This study aims to explore the role of digital back-feeding in the relationship between the digital divide and HRQOL among older men and women living in rural China.

Methods: We used data from wave 3 of the Shandong Rural Elderly Health Cohort, conducted in 2022. A total of 3242 (n=1946, 60.02% women) rural older adults were included in the analysis. Moderating effect analysis was performed using Tobit regression models and margins plots.

Results: A total of 71.01% (2302/3242) of the participants reported experiencing digital divide. Participants experiencing digital divide were significantly associated with lower HRQOL as measured by EQ-5D-5L scores (β=–0.020; *P*<.001). We found that digital back-feeding buffered the relationship between digital divide and HRQOL (β=0.024; *P*=.02). Furthermore, gender-stratified analyses revealed divergent moderation patterns; a significant buffering role was observed in women (β=0.031; *P*=.02), whereas no substantially significant moderating role emerged in men.

Conclusions: Our study established a significant inverse association between the digital divide and HRQOL among rural adults. Digital back-feeding emerged as a measurable protective buffer mitigating this adverse relationship. Furthermore, this buffering effect was only observed among older women. Policy implications underscore the necessity of gender-tailored digital inclusion strategies, particularly advocating for technology-proficient adult offsprings to prioritize digital engagement with their mothers in digitally marginalized rural communities.

(JMIR Aging 2025;8:e75925) doi:10.2196/75925

KEYWORDS

digital divide; digital back-feeding; health-related quality of life; HRQOL; gender difference

XSL•FO
RenderX

Introduction

Background

Population aging has become a global issue of wide concern. As one of the countries in the world experiencing the fastest population aging, China had approximately 280 million people aged more than 60 years in 2022, accounting for 19.8% of its total population [1]. This number is expected to rise to 509 million, accounting for 38.81% of its total population by 2050 [2]. As the population ages and life expectancy increases, improving the health-related quality of life (HRQOL) of older adults has attracted attention from academic and social fields. HRQOL is a multidimensional concept that includes physical functioning, mental health, and socially related roles as perceived by an individual over time [3]. It has been increasingly used as a comprehensive health indicator in health surveys, as poor HRQOL is associated with various adverse outcomes, including loss of functional independence, mental health deterioration, and mortality [4,5]. To promote HRQOL among older people, it is important to identify factors associated with this key health-related outcome.

Digital access is emerging as a social determinant of health. Research has demonstrated that digital technologies directly impact specific aspects of HRQOL in older adults through multiple pathways. For instance, telehealth access has reduced health care barriers through remote monitoring and internet-based consultations [6]. Digital platforms have combated social isolation through communication tools [7]. Furthermore, online health information seeking has empowered informed health care decision-making [8]. People with better access to digital technologies tend to have improved mental health, physical health, and medical decision-making skills [9-12]. Existing literature has predominantly examined quality of life (QOL) in older populations [13-15], and these findings hold critical implications for HRQOL. However, the disparity in different groups' access to and use of information and communications technologies (ICTs) and the internet often leads to inequalities known as the *digital divide* [16,17]. People affected by the digital divide may miss out on many of the described benefits, causing inequities and disparities in health [18,19], thus reducing QOL. Previous studies have shown that individuals aged more than 60 years account for 41.9% of the total number of nonnetizens [20], and the rate of digital divide could reach up to 90.4% [21] in China. Notably, due to the unequal economic development and resource allocation, the digital divide seems to pose an even greater challenge for older individuals residing in rural areas [18,21,22]. However, the association between the digital divide and HRQOL among rural older adults remains unclear. Understanding the digital challenges faced by rural older adults is crucial for developing strategies to bridge the digital divide and improve their health outcomes.

Digital back-feeding, a new form of postfigurative culture, refers to the teaching behaviors of the younger generation toward the older generation on digital access, use, and literacy [23-25]. Unlike other forms of intergenerational support, such as emotional encouragement or financial aid, digital back-feeding is distinct in its targeted focus on transferring digital skills and knowledge, enabling older adults to navigate technology independently [22,26]. Family members, especially the adult offsprings, are usually considered warm experts who participate in the daily lives of older adults and assist them in solving problems encountered while using technology [27]. While digital support could theoretically come from various family members, including grandchildren or other relatives, adult offsprings typically emerge as the primary source of such assistance in rural areas [28]. This pattern reflects cultural norms, such as filial piety, closer geographic proximity [29], and a more favorable age gap that facilitates effective intergenerational communication and learning [30].

Digital back-feeding is crucial in addressing the digital divide and its impact on HRQOL, as it may bridge generational gaps in digital access and empower older adults to mitigate the negative effects of digital exclusion. Previous studies showed that intergenerational interaction and support can promote internet use and digital literacy among older adults [31,32]. Receiving digital back-feeding could facilitate the integration of older adults into the digital era, foster harmonious parent-child relationships, bolster self-efficacy, and improve subjective well-being [33-37]. A previous study demonstrated that contact with family plays a significant role in the relationship between IT use and psychological well-being [38]. Khalaila and Vitman-Schorr [14] found that social capital, such as being accompanied by family, moderates the positive association between internet use and QOL. With regard to our hypotheses, digital back-feeding is posited as a key moderator that strengthens the relationship between digital literacy and improved HRQOL by facilitating informed health decisions and reducing social isolation. However, the existing studies on digital back-feeding have mainly focused on qualitative research, and the role of digital back-feeding as a moderator between the digital divide and HRQOL has not been convincingly demonstrated.

In theory, the buffering effect model of social support posits that social support enables individuals to alleviate the detrimental effects of stressful events on their physical and mental well-being, thereby improving their overall health [39]. Specifically, while social support does not modify the stressful event itself, it can adjust individuals' perceptions of the stressors and their own circumstances, thus providing a protective effect. Simultaneously, being profoundly influenced by traditional Chinese culture, support from the younger generation is perceived as a filial duty. Various forms of intergenerational assistance from juniors are seen as forms of repayment to older adults, thereby enhancing the subjective well-being of older adults [40]. Digital back-feeding, which was considered the most crucial means to deal with the digital divide [41], may mitigate the effects of the digital divide on HRQOL.

Another issue deserving more attention is whether the moderating mechanism of the digital divide and HRQOL would differ by gender. Compelling evidence from previous research shows that there is a gender disparity in both access to and use of ICTs, with women consistently lagging behind men [42-44]. Nevertheless, women are commonly perceived as more relationship-oriented; therefore, they may derive greater benefits

XSL•FO
RenderX

from social participation compared to men [45,46]. The digital back-feeding within familial contexts also reveals a discernible gender disparity [47]. Specifically, older women demonstrate a more pronounced breadth and depth of digital back-feeding in comparison to men. These disparities may stem from and lead to differential effects of digital back-feeding on HRQOL through differences in social roles, support-seeking behaviors, and health management patterns. For instance, existing literature indicated that women often show greater embeddedness in family networks [48], engage more actively in help-seeking within families [49], and participate more proactively in health information seeking [50,51]. Given this, we speculate that the moderating effect of digital back-feeding between the digital divide and HRQOL among rural older adults might differ by gender.

Objectives

The aims of this study were as follows: (1) to determine the prevalence of digital divide among older adults in rural China, (2) to explore the association between digital divide and HRQOL, and (3) to examine the moderating role of digital back-feeding in the digital divide–HRQOL relationship and gender difference in this moderating mechanism. This study provides a new perspective to help rural older adults bridge the digital divide.

Methods

Data Collection

The data used in this study were obtained from wave 3 of the Shandong Rural Elderly Health Cohort, which was designed to investigate the health status of the older population aged 60 years or more in rural areas of Shandong province. A multistage stratified random sampling method was used to select participants. Details of the study methodology are described in detail elsewhere [52,53]. The baseline survey was completed in June 2019, and follow-up surveys were conducted in 2020 and 2022. In the third wave, 3468 questionnaires were distributed, and 3242 valid questionnaires with complete information were included in this analysis.

Measures

HRQOL Measure

The HRQOL in this study was measured by the EQ-5D-5L questionnaire, which consists of 5 dimensions, including mobility, self-care, usual activities, pain or discomfort, and anxiety or depression, each of which has 5 levels of response (1=no problems, 2=slight problems, 3=moderate problems, 4=severe problems, and 5=extreme problems). The EQ-5D-5L utility values are generated by weighting each dimension of HRQOL using the time trade-off model set for the Chinese general population [54], ranging from –0.391 to 1.000. Higher scores indicate better health.

Digital Divide

The participants were asked two questions [55,56]: (1) how frequently they use the internet (including smartphones and computers), and the responses ranged from 1(“never”) to 5 (“almost every day”) and (2) whether they have a smartphone with internet access. If the answer to the former question was “never” or “less than 1 month” and the answer to the latter question was “no,” the rural older adults were classified as “1” denoting digital divide; otherwise, they were classified as “0” denoting no digital divide. This measurement approach captured the “first-level digital divide” by focusing on basic access to digital technologies, which has been widely used in digital divide research among the older adult population [18,57]. Given that basic internet access remains a primary barrier for rural older adults in China [21], this binary classification was selected for its simplicity and practical applicability. Moreover, digital literacy–based measurements tend to underperform in this population due to their limited baseline digital skills and experience [58].

Digital Back-Feeding

Whether the older adults received digital back-feeding was a dichotomous variable represented by the question “Whether your child helps you when you encounter difficulties in using digital devices?” If the answer was yes, the participants were classified as “1,” denoting they received digital back-feeding.” If the answer was no, they were classified as “0,”denoting they did not receive digital back-feeding.”

Covariate Variable

Covariate variables included sociodemographic characteristics, life behaviors, and health status. Sociodemographic characteristics were measured by gender, age, education, marital status, and household income. Educational attainment was classified into 3 categories: illiterate or elementary school, junior high school, and high school or above. Marital status was divided into 2 categories: single and married, of which the single category included those who were unmarried, divorced, or widowed. Household income was classified into 4 categories: quartile 1, quartile 2, quartile 3, and quartile 4 according to the quartile methods, and the higher quartile represented a higher family economic status. Life behaviors included current smoking status, current drinking status, and performing physical exercise. Health status was measured using the number of chronic diseases experienced. Chronic disease status was divided into 3 categories: 0, 1, and 2 or more.

Statistical Analysis

We used Stata/MP (version 17.0; Stata Corp) to examine the data and conducted descriptive statistics and correlation analysis. Descriptive statistics were provided to analyze older adults’ characteristics. Continuous variables were summarized using the means and SDs. Categorical variables were reported as frequencies and percentages. Independent samples 2-tailed *t* tests and 1-way ANOVA were used to test for differences in HRQOL between groups. Tobit regression models were used to examine the association between the digital divide and HRQOL because the distribution of the EQ-5D utility index was censored at 1. Compared to ordinary least squares regression, tobit models provided more accurate parameter estimates when dealing with bounded outcomes [59,60]. We evaluated several progressive models. Model 1 of the regression analysis included digital divide and digital back-feeding as predictors of HRQOL, and model 2 was examined to explore

XSL•FO
RenderX

whether digital back-feeding is a potential moderator by adding the interaction term (digital divide×digital back-feeding). The interaction term tested whether the slope of the association between digital divide and HRQOL differed across levels of digital back-feeding. A significant interaction suggested that the strength or direction of the relationship between digital divide and HRQOL varied depending on participants' level of digital back-feeding. All confounders were included to adjust the model. Regarding gender differences in the moderating mechanism, we performed multiple group analyses by fitting the models separately by gender (models 3-6). In addition, the margins plot was used to illustrate the prediction of HRQOL by digital divide and digital back-feeding. The reported CIs were calculated at the 95% level, and *P* values less than.05 were considered statistically significant.

Ethical Considerations

This study was approved by the ethics committee of Shandong University (20181228). All participants signed informed consent before the survey. The manuscript contains no identifiable features of research participants. All participants received modest household commodities (eg, laundry detergent and towels) valued at approximately $3 USD per session as compensation for their time and participation.

Results

Descriptive Statistics

This study included 3242 older people. Among them, 2302 (71.01%) rural older adults experienced digital divide, and the prevalence was higher among older women (1407/1946, 72.3%). Rural older adults who were younger, men, married, had higher educational levels, and did not have chronic diseases reported higher EQ-5D scores. More information on the participants' characteristics for the whole sample and gender subsamples is presented in Table 1.

XSL•FO
RenderX

Table 1. Description and univariate analysis of EQ-5D for all participants, men, and women (N=3242).

Variables	All		Men (n=1296)		Women (n=1946)	
	Participants, n (%)	EQ-5D, mean (SD)	Participants, n (%)	EQ-5D, mean (SD)	Participants, n (%)	EQ-5D, mean (SD)
Total (N=3242)	3242 (100)	0.896 (0.165)	1296 (39.98)	0.905 (0.167)	1946 (60.02)	0.891 (0.163)
Digital divide						
No	940 (28.99)	0.928 (0.112)[a]	401 (30.94)	0.942 (0.111)[a]	539 (27.7)	0.918 (0.112)[a]
Yes	2302 (71.01)	0.883 (0.180)	895 (69.06)	0.888 (0.185)	1407 (72.3)	0.881 (0.177)
Digital back-feeding						
No	1132 (34.92)	0.894 (0.156)	474 (36.57)	0.905 (0.157)	658 (33.81)	0.886 (0.155)
Yes	2110 (65.08)	0.898 (0.169)	822 (63.43)	0.904 (0.173)	1288 (66.19)	0.893 (0.167)
Age (y)						
60-69	1098 (33.87)	0.923 (0.130)[a]	414 (31.94)	0.933 (0.134)[a]	684 (35.15）	0.917 (0.127)[a]
70-79	1762 (54.35)	0.894 (0.165)	720 (55.56)	0.904 (0.166)	1042 (53.55)	0.886 (0.164)
≥80	382 (11.78)	0.833 (0.222)	162 (12.5)	0.834 (0.219)	220 (11.31)	0.831 (0.225)
Marital status						
Single[b]	744 (22.95)	0.880 (0.186)[c]	181 (13.97)	0.883 (0.193)	563 (28.93)	0.879 (0.184)[d]
Married	2498 (77.05)	0.901 (0.157)	1115 (86.03)	0.908 (0.162)	1383 (71.07)	0.896 (0.153)
Education						
Illiterate or elementary school	2542 (78.41)	0.888 (0.172)[a]	845 (65.2)	0.891 (0.181)[a]	1697 (87.2)	0.887 (0.167)[a]
Junior high school	517 (15.95)	0.927 (0.124)	330 (25.46)	0.931 (0.133)	187 (9.61)	0.920 (0.106)
High school or above	183 (5.64)	0.922 (0.144)	121 (9.34)	0.930 (0.134)	62 (3.19)	0.906 (0.161)
Chronic diseases, n						
0	489 (15.08)	0.961 (0.083)[a]	254 (19.6)	0.964 (0.088)[a]	235 (12.08)	0.959 (0.077)[a]
1	858 (26.47)	0.919 (0.138)	369 (28.47)	0.935 (0.118)	489 (25.13)	0.907 (0.150)
≥2	1895 (58.45)	0.869 (0.184)	673 (51.93)	0.866 (0.200)	1222 (62.8)	0.871 (0.175)
Household income [e]						
Quartile 1	820 (25.29)	0.880 (0.179)[a]	287 (22.15)	0.874 (0.192)[a]	533 (27.39)	0.884 (0.171)[a]
Quartile 2	801 (24.71)	0.890 (0.169)	351 (27.08)	0.898 (0.167)	450 (23.12)	0.884 (0.171)
Quartile 3	811 (25.02)	0.902 (0.168)	331 (25.54)	0.918 (0.167)	480 (24.67)	0.891 (0.168)
Quartile 4	810 (24.98)	0.913 (0.138)	327 (25.23)	0.926 (0.138)	483 (24.82)	0.905 (0.137)
Current smoking status						
No	2562 (79.03)	0.892 (0.168)[c]	776 (59.88)	0.901 (0.177)	1786 (91.78)	0.888 (0.165)[d]
Yes	680 (20.97)	0.912 (0.148)	520 (40.12)	0.910 (0.151)	160 (8.22)	0.918 (0.137)
Current drinking status						
No	2414 (74.46)	0.887 (0.175)[a]	642 (49.54)	0.883 (0.200)[a]	1772 (91.06)	0.889 (0.164)[d]
Yes	828 (25.54)	0.924 (0.127)	654 (50.46)	0.926 (0.123)	174 (8.94)	0.914 (0.142)
Physical exercise						
No	1366 (42.13)	0.860 (0.199)[a]	569 (43.9)	0.871 (0.203)[a]	797 (40.96)	0.852 (0.197)[a]
Yes	1876 (57.87)	0.923 (0.127)	727 (56.1)	0.931 (0.127)	1149 (59.04)	0.918 (0.127)

[a]*P*<.001.

[b]Singles included those who were unmarried (n=28, 0.86%), divorced (n=3, 0.09%), and widowed (n=713, 21.99%) among the whole sample.

XSL•FO
RenderX

[c]*P*<.01.

[d]*P*<.05.

[e]Quartile 1 refers to the lowest income, and quartile 4 refers to the highest income.

Association Between the Digital Divide and HRQOL

As shown in Table 2, older adults experiencing digital divide were associated with poorer EQ-5D scores compared to those not experiencing digital divide (β=–0.020; *P*<.001), when control variables were included among the whole sample in model 1. Similar associations were observed in the grouped regression (models 3 and 5).

XSL•FO
RenderX

Table 2. Standardized coefficients of explanatory variables on health-related quality of life among all participants, men, and women (N=3242).

Exploratory variables	All		Men (n=1296)		Women (n=1946)	
	Model 1[a], β (SE)	Model 2[b], β (SE)	Model 3[c], β (SE)	Model 4[d], β (SE)	Model 5[c], β (SE)	Model 6[d], β (SE)
Digital divide (reference: no)						
Yes	–0.02 (0.005)[e]	–0.036 (0.008)[e]	–0.025 (0.009)[f]	–0.032 (0.012)[f]	–0.017 (0.007)[g]	–0.038 (0.01)[e]
Digital back-feeding (reference: no)						
Yes	0.001 (0.006)	–0.016 (0.007)[g]	–0.008 (0.009)	–0.016 (0.011)	0.007 (0.007)	–0.016 (0.009)
Digital divide × digital back-feeding	—[h]	0.024 (0.01)[g]	—	0.011 (0.016)	—	0.031 (0.013)[g]
Gender (reference: men)						
Women	0.008 (0.008)	0.008 (0.008)	—	—	—	—
Age (y; reference: 60-70 y)						
70-80	–0.011 (0.006)	–0.011 (0.006)	–0.002 (0.009)	–0.002 (0.009)	–0.018 (0.007)[g]	–0.018 (0.007)[g]
80	–0.073 (0.012)[e]	–0.073 (0.012)[e]	–0.065 (0.018)[e]	–0.065 (0.018)[e]	–0.077 (0.016)[e]	–0.077 (0.016)[e]
Marital status (reference: single[i])						
Married	0.006 (0.007)	0.007 (0.007)	0.014 (0.014)	0.014 (0.014)	0.005 (0.009)	0.005 (0.009)
Education (reference: illiterate or elementary school)						
Junior high school	0.023 (0.007)[e]	0.023 (0.007)[e]	0.022 (0.01)[g]	0.022 (0.01)[g]	0.025 (0.009)[f]	0.025 (0.009)[f]
High school or above	0.006 (0.011)	0.007 (0.011)	0.014 (0.013)	0.014 (0.013)	–0.012 (0.018)	–0.011 (0.018)
Chronic diseases (n; reference: 0)						
1	–0.04 (0.006)[e]	–0.04 (0.006)[e]	–0.031 (0.008)[e]	–0.03 (0.008)[e]	–0.047 (0.008)[e]	–0.047 (0.008)[e]
≥2	–0.085 (0.006)[e]	–0.085 (0.006)[e]	–0.094 (0.009)[e]	–0.094 (0.009)[e]	–0.081 (0.007)[e]	–0.08 (0.007)[e]
Household income[j] (reference: quartile 1)						
Quartile 2	–0.002 (0.009)	–0.002 (0.009)	0.017 (0.014)	0.018 (0.014)	–0.014 (0.011)	–0.014 (0.011)
Quartile 3	0.005 (0.009)	0.005 (0.009)	0.027 (0.014)	0.027 (0.014)	–0.009 (0.012)	–0.01 (0.011)
Quartile 4	0.01 (0.008)	0.01 (0.008)	0.029 (0.013)[g]	0.028 (0.013)[g]	–0.001 (0.01)	–0.002 (0.01)
Smoke (reference: no)						
Yes	0.002 (0.007)	0.002 (0.007)	–0.004 (0.009)	–0.004 (0.009)	0.015 (0.011)	0.015 (0.012)
Drinking (reference: no)						
Yes	0.025 (0.007)[e]	0.025 (0.007)[e]	0.031 (0.008)[e]	0.031 (0.008)[e]	0.011 (0.011)	0.01 (0.011)
Physical exercise (reference: no)						
Yes	0.062 (0.006)[e]	0.062 (0.006)[e]	0.061 (0.009)[e]	0.061 (0.009)[e]	0.064 (0.008)[e]	0.063 (0.008)[e]
Constant	0.925 (0.015)[e]	0.936 (0.015)[e]	0.908 (0.024)[e]	0.913 (0.023)[e]	0.94 (0.015)[e]	0.956 (0.015)[e]

[a]Model 1 adjusted all cofounders for the whole sample.

[b]Model 2 adjusted all cofounders and the interaction term of digital divide and digital back-feeding for the whole sample.

[c]Models 3 and 5 adjusted all cofounders for the gender subsamples.

[d]Models 4 and 6 adjusted all cofounders and the interaction term of digital divide and digital back-feeding for the gender subsamples.

[e]*P*<.001.

[f]*P*<.01.

[g]*P*<.05.

[h]Data not available.

[i]Singles included those who were unmarried (n=28, 0.86%), divorced (n=3, 0.09%), and widowed (n=713, 21.99%) among the whole sample.

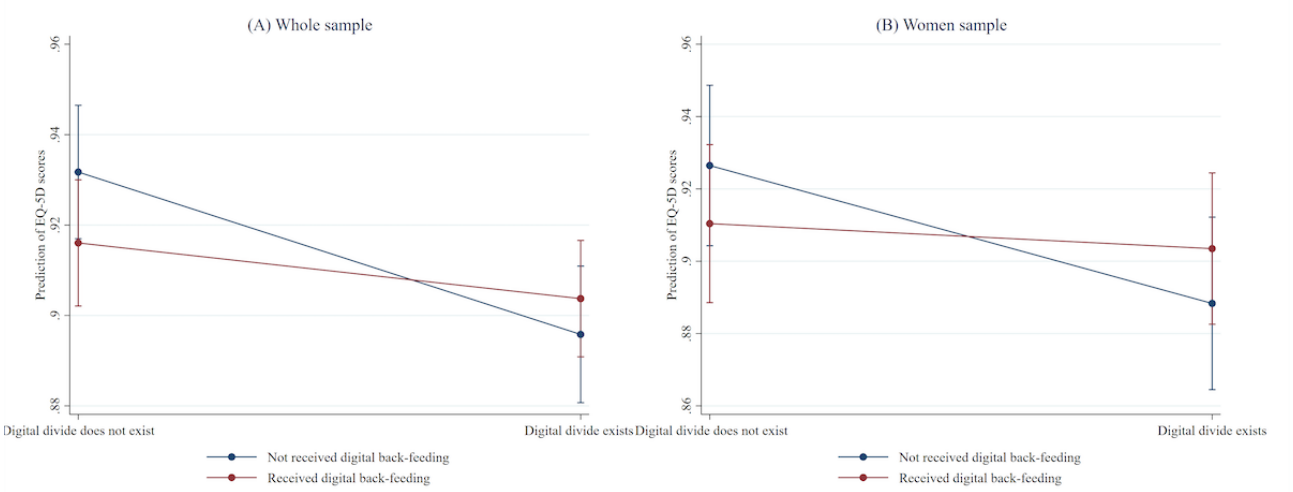
[j]Quartile 1 refers to the lowest income, and quartile 4 refers to the highest income.

XSL•FO
RenderX

Moderating Role of Digital Back-Feeding Between the Digital Divide and HRQOL

Model 2 in Table 2 included the interaction term (digital divide × digital back-feeding) to explore whether intergenerational digital support played a moderating role in the relationship between digital divide and HRQOL. The significant interaction term (β=0.024; *P*=.02) suggested that digital back-feeding was associated with a weaker negative relationship between the digital divide and HRQOL. In the men subsample, no significant moderating effect was observed in the relationship between the digital divide and HRQOL. However, within the women subsample, a significant moderating effect was observed (β=0.031; *P*=.02). Figure 1 illustrates the moderating pattern of digital back-feeding on the association between the digital divide and HRQOL among the whole sample and women subsample.

Figure 1. Interaction between digital divide and digital back-feeding in the prediction of EQ-5D scores among the whole and women samples.



Discussion

Principal Findings

This study first explored the association between the digital divide and HRQOL as well as the role of digital back-feeding as a moderator in this relationship among rural older adults. Our results revealed that the digital divide was associated with poorer HRQOL, and digital back-feeding alleviated this negative association. Moreover, the buffer effect of digital back-feeding between the digital divide and HRQOL was observed in older women participants but not in older men participants. The findings were beneficial to facilitate the development of targeted interventions for individuals to improve HRQOL in rural older adults.

In this study, 71.01% (2302/3242) of the participants experienced digital divide. Although the internet gradually penetrates the middle and older age groups over time [61,62], it remains lower in China and some other low- and middle-income countries [56]. The possible reason was that the advantages of the internet were constrained due to the concurrent occurrence of economic growth and population aging. Older adults face numerous physical or cognitive challenges when learning and using the internet [63]. Furthermore, decreased familiarity, the lack of digital literacy, issues of trust, and concerns about privacy restrict the use of the internet among older adults [64]. Compared with urban areas, rural areas lag in terms of internet infrastructure, education levels, and digital literacy [65]. Thus, as the primary population experiencing "information poverty," rural older adults face barriers to enjoying the dividends brought by the development of digital technology and require more assistance.

Results from our study indicated that the digital divide was negatively associated with HRQOL among Chinese rural older adults, which is consistent with previous studies [13,14]. The possible explanation is that the internet enables older adults to access more health-related information and medical services [66,67]. Rural older adults with poor internet accessibility are blocked from accessing health knowledge, which affects their health-related behaviors affected, deteriorates their physical and mental health, and ultimately reduces HRQOL. Moreover, the internet provides older adults with additional options for social participation and access to information related to their personal interests. Rural older people who encounter the digital divide may feel socially isolated and have reduced social interaction, resulting in feelings of loneliness, unhappiness, and boredom [68]. These findings imply that promoting internet use to bridge the digital divide could be beneficial for older people's HRQOL, leading to better health outcomes.

Our study also revealed that digital back-feeding moderated the association between digital divide and HRQOL. Specifically, digital back-feeding buffers the negative impact of digital divide on HRQOL among rural older adults. There are several possible explanations for this finding. First, support with digital technology from their offspring can help older adults better adapt to digital life and build better connections with society [69]. It can strengthen their sense of social belonging, enhance their self-worth, enrich their spiritual life, and be beneficial for their mental health. Second, digital back-feeding can improve health outcomes by increasing digital health literacy among older adults. Previous studies found that better digital health literacy is associated with positive health behaviors and better HRQOL [11,70], while digital back-feeding can improve parents' digital literacy [71]. Third, family constitutes the pivotal

XSL•FO
RenderX

domain for digital feedback [72]. For rural older adults, households function as the fundamental units for health promotion and significantly contribute to their overall well-being. Family support and intergenerational interactions help mitigate feelings of loneliness and reduce the risk of psychological diseases by fostering stronger bonds across different age groups [37,38,73]. These supportive relationships ultimately enhance the HRQOL of older adults. Thus, digital back-feeding could mitigate the possibility of rural older adults experiencing the digital divide having worse HRQOL.

This study examined gender differences in the moderation mechanism. Digital back-feeding buffered the negative correlation between the digital divide and HRQOL for older women after adjusting for other confounding factors but not among men. Consistent with previous studies, women are less likely to access and use ICTs than men [42,43,74]. This disparity can be somewhat elucidated through a life course perspective [75]. Men and women have encountered distinct opportunities over the years, especially in the realms of educational attainment, paid employment, social concepts, and caregiving responsibilities, which have subsequently influenced their exposure to ICTs. Therefore, digital back-feeding activities aimed at women are more likely to show positive impacts. Compared to men, women's social worlds were more confined to their locality, with the domestic and emotional labor of "making" a home being central to their identity [76]. This distribution of duties tends to increase their reliance on family members for support, particularly when confronted with new technologies. Conversely, men may participate less in these intrafamilial support activities and thus benefit less from digital support from their offsprings. Another explanation was that older women are more likely to articulate and share their emotions and demonstrate superior effectiveness in establishing informal networks [77]. These gender-specific findings have important implications for targeted interventions to bridge the digital divide among rural older adults. Given that women showed greater responsiveness to digital back-feeding, family-integrated digital literacy programs should be prioritized. Adult offsprings can serve as digital mentors in these programs, providing collaborative support that matches women's relational orientations. For men, digital back-feeding demonstrated limited benefits due to socialized independence; therefore, interventions should emphasize autonomous learning approaches. Peer-led training or individual tutorials focusing on practical applications may be more acceptable while building self-efficacy.

The findings of this study provide new inspiration for rural older adults to enhance HRQOL, especially highlighting the need for gender-tailored interventions. It is crucial to mobilize social support from both formal and informal organizations. First, policymakers should safeguard digital rights for older adults by strengthening the retrofitting of digital platforms for aging and enhancing digital health literacy. For example, local governments can establish digital literacy training in community centers, offering family-oriented sessions where adult offsprings guide older women in using smartphones for health management and communication. Second, rural communities should promote digital development by improving internet infrastructure and collaborating with telecommunication providers to offer affordable smartphones and internet plans, ensuring that all older adults can easily access online services. Third, adult offsprings should be encouraged to spend more time and patience assisting older adults, especially older women, to use digital technologies for communication or entertainment. Last but not least, social groups and organizations should actively promote digital inclusion by providing volunteer support and technical assistance, while internet service providers should develop age-friendly interfaces to reduce access barriers. By implementing these targeted and practical measures, rural areas can more effectively bridge the digital divide and improve the health and QOL of all older adults, with particular attention to the unique needs of women.

Limitations

Despite the contributions made to the existing literature, there are several limitations in this study. First, this study used cross-sectional data, which may introduce recall bias. In the future, longitudinal designs will be used to verify this relationship. Second, our measurement approaches have inherent constraints. The digital divide was assessed primarily through basic access metrics rather than comprehensive digital literacy or quality of internet use, which may have overlooked nuanced variations in digital engagement affecting HRQOL. This limitation would be remedied in the future study. Third, the findings' generalizability should be interpreted cautiously, as data were self-reported and drawn exclusively from Shandong Province, China. Broader, multiregional studies with objective measures are recommended to validate and extend these results.

Conclusions

In conclusion, 71.01% (2302/3242) of the participants reported having experienced the digital divide. Our findings demonstrated that the digital divide was negatively associated with HRQOL, and digital back-feeding moderated this relationship among rural older adults. Moreover, the buffering effect was observed in older women but not in older men. Recognizing the relationship between the digital divide and HRQOL provides insights for bridging the digital divide and enhancing HRQOL among older adults. This study suggests that broad measures should be implemented, and efforts should be taken to help the rural older population bridge the digital divide and integrate into the digital society.

Acknowledgments

This work was supported by the National Science Foundation of China (grants 72274109, 71774104, and 71974117). The study sponsor has no role in study design, data analysis and interpretation, writing the manuscript, or the decision to submit the paper for publication.

XSL•FO
RenderX

Authors' Contributions

CZ was involved in conceptualizing and designing the study. Analysis and interpretation of data were performed by XC, S Chai, DZ, and S Chen. The manuscript was drafted by XC. Critical revision of the manuscript for important intellectual content was performed by CZ, DZ, and S Chai. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. 2022 national aging development bulletin. National Development and Reform Commission. URL: https://www.gov.cn/lianbo/bumen/202312/content_6920261.htm [accessed 2025-05-29]
2. World population prospects 2022: summary of results. United Nations. 2022 Aug. URL: https://www.un.org/development/desa/pd/content/World-Population-Prospects-2022 [accessed 2025-09-17]
3. Atlanta G. Measuring healthy days: population assessment of health-related quality of life. Centers for Disease Control and Prevention. URL: https://stacks.cdc.gov/view/cdc/6406 [accessed 2025-05-29]
4. Thompson WW, Zack MM, Krahn GL, Andresen EM, Barile JP. Health-related quality of life among older adults with and without functional limitations. Am J Public Health 2012 Mar;102(3):496-502. [doi: 10.2105/AJPH.2011.300500] [Medline: 22390514]
5. Masel MC, Ostir GV, Ottenbacher KJ. Frailty, mortality, and health-related quality of life in older Mexican Americans. J Am Geriatr Soc 2010 Nov;58(11):2149-2153 [FREE Full text] [doi: 10.1111/j.1532-5415.2010.03146.x] [Medline: 21054296]
6. Pinero de Plaza MA, Gulyani A, Bulto LN, Allande-Cussó, R, Pearson V, Lange, B, et al. Telehealth as a Care Solution for Homebound People: Systematic Review and Meta-Analysis of Healthcare Utilization, Quality of Life, and Well-Being Outcomes. Health & Social Care in the Community 2014 Jul 25;2025(1):-. [doi: 10.1155/hsc/7224151]
7. Grigorovich A, Marcotte AA, Colobong R, Szabo M, MacNeill C, Blais D, et al. Using voice-activated technologies to enhance well-being of older adults in long-term care homes. Innov Aging 2024 Nov 04;8(12):igae102 [FREE Full text] [doi: 10.1093/geroni/igae102] [Medline: 39664605]
8. Tennant B, Stellefson M, Dodd V, Chaney B, Chaney D, Paige S, et al. eHealth literacy and Web 2.0 health information seeking behaviors among baby boomers and older adults. J Med Internet Res 2015 Mar 17;17(3):e70 [FREE Full text] [doi: 10.2196/jmir.3992] [Medline: 25783036]
9. Wallinheimo AS, Evans SL. More frequent internet use during the COVID-19 pandemic associates with enhanced quality of life and lower depression scores in middle-aged and older adults. Healthcare (Basel) 2021 Apr 01;9(4):393 [FREE Full text] [doi: 10.3390/healthcare9040393] [Medline: 33916054]
10. Withers EM. The digital divide and health: examining digital access as a social determinant of health. Portland State University. URL: https://www.detroitseniorsolution.org/app/uploads/2022/02/Withers_The_Digital_Divide-as-a-SDOH.pdf [accessed 2025-09-17]
11. Choukou MA, Sanchez-Ramirez DC, Pol M, Uddin M, Monnin C, Syed-Abdul S. COVID-19 infodemic and digital health literacy in vulnerable populations: a scoping review. Digit Health 2022 Feb 10;8:20552076221076927 [FREE Full text] [doi: 10.1177/20552076221076927] [Medline: 35223076]
12. Mitchell UA, Chebli PG, Ruggiero L, Muramatsu N. The digital divide in health-related technology use: the significance of race/ethnicity. Gerontologist 2019 Jan 09;59(1):6-14. [doi: 10.1093/geront/gny138] [Medline: 30452660]
13. Aggarwal B, Xiong Q, Schroeder-Butterfill E. Impact of the use of the internet on quality of life in older adults: review of literature. Prim Health Care Res Dev 2020 Dec 02;21:e55 [FREE Full text] [doi: 10.1017/S1463423620000584] [Medline: 33263273]
14. Khalaila R, Vitman-Schorr A. Internet use, social networks, loneliness, and quality of life among adults aged 50 and older: mediating and moderating effects. Qual Life Res 2018 Feb;27(2):479-489. [doi: 10.1007/s11136-017-1749-4] [Medline: 29210015]
15. Yuan H. Internet use and mental health problems among older people in Shanghai, China: the moderating roles of chronic diseases and household income. Aging Ment Health 2021 Apr;25(4):657-663. [doi: 10.1080/13607863.2020.1711858] [Medline: 31928208]
16. World telecommunication development conference (WTDC-02): final report. International Telecommunication Union (ITU). URL: https://www.itu.int/pub/D-TDC-WTDC-2002 [accessed 2025-05-29]
17. National Telecommunications and Information Administration. Falling through the net: defining the digital divide. In: Compaine BM, editor. The Digital Divide: Facing a Crisis or Creating a Myth?. Cambridge, MA: The MIT Press; 2001:15.
18. Hong YA, Zhou Z, Fang Y, Shi L. The digital divide and health disparities in China: evidence from a national survey and policy implications. J Med Internet Res 2017 Sep 11;19(9):e317 [FREE Full text] [doi: 10.2196/jmir.7786] [Medline: 28893724]

XSL•FO
RenderX

19. Eruchalu CN, Pichardo MS, Bharadwaj M, Rodriguez CB, Rodriguez JA, Bergmark RW, et al. The expanding digital divide: digital health access inequities during the COVID-19 pandemic in New York City. J Urban Health 2021 Apr;98(2):183-186 [FREE Full text] [doi: 10.1007/s11524-020-00508-9] [Medline: 33471281]
20. The 52nd statistical report on China's internet development. China Internet Network Information Center. 2022. URL: https://www.cnnic.com.cn/IDR/ReportDownloads/202311/P020231121355042476714.pdf [accessed 2025-09-17]
21. Li L, Jin G, Guo Y, Zhang Y, Jing R. Internet access, support, usage divides, and depressive symptoms among older adults in China: a nationally representative cross-sectional study. J Affect Disord 2023 Feb 15;323:514-523. [doi: 10.1016/j.jad.2022.12.001] [Medline: 36496102]
22. Hong J, Li X. The “cultural feedback” of we chat in rural families: a case study of Chen Qu Village. Chin J Journal Commun 2019;41(10):50-74. [doi: 10.13495/j.cnki.cjjc.2019.10.003]
23. Margaret M. Culture and Commitment: A Study of the Generation Gap. New York, NY: Doubleday; 1970.
24. Lu D, Lin H. An empirical study on digital feedback behavior of young people in COVID-19 with health belief model and UTAUT model. Mob Inf Syst 2022 Sep 21;2022:1-13. [doi: 10.1155/2022/9095003]
25. Ma J, Fang G, Guo K. Bridging the digital divide: the influence of digital feedback on the digital capabilities of the rural elderly. Inf Dev 2023 Dec 28;41(6):E41-E50. [doi: 10.1177/02666669231222208] [Medline: 22861175]
26. Taipale S. Intergenerational Connections in Digital Families. Cham, Switzerland: Springer; 2019.
27. Olsson T, Viscovi D. Warm experts for elderly users: who are they and what do they do? Hum Technol 2018 Nov 30;14(3):324-342. [doi: 10.17011/ht/urn.201811224836]
28. Zhang H, He R. A study on digital inclusion of Chinese rural older adults from a life course perspective. Front Public Health 2022;10:974998 [FREE Full text] [doi: 10.3389/fpubh.2022.974998] [Medline: 36187630]
29. Zhao D, Li J, Fu P, Hao W, Yuan Y, Yu C, et al. Cognitive frailty and suicidal ideation among Chinese rural empty-nest older adults: Parent-child geographic proximity as a possible moderator? J Affect Disord 2021 Mar 01;282:348-353. [doi: 10.1016/j.jad.2020.12.111] [Medline: 33421862]
30. Friemel TN. The digital divide has grown old: determinants of a digital divide among seniors. New Media Soc 2014 Jun 12;18(2):313-331. [doi: 10.1177/1461444814538648]
31. Tsai HS, Shillair R, Cotten SR. Social support and "playing around": an examination of how older adults acquire digital literacy with tablet computers. J Appl Gerontol 2017 Jan;36(1):29-55 [FREE Full text] [doi: 10.1177/0733464815609440] [Medline: 26491029]
32. Wang S, Dong H. Research on the impact of Internet use on the health of the elderly: based on the perspective of gender differences and intergenerational support. J Popul Sci 2024;46(02):77-92 [FREE Full text] [doi: 10.16405/j.cnki.1004-129X.2024.02.005]
33. Sheng M, Wang J, Zhu X, Zhang P. An exploratory study of digital reverse mentoring based on grounded theory from the youth's perspective. Doc Inf Knowl 2023;40(04):52-61. [doi: 10.13366/J.DIK.2023.04.052]
34. Wu L, Zhu B. Intergenerational cooperation and behavior choice of rural digital feedback. Acad J Zhongzhou 2022(5):57-65 [FREE Full text]
35. Jiang Y. Media literacy factors in family socialization perspective: a survey of the media literacy of middle school students and their parents in NanNing. Communication University of China [dissertation] 2023:1-167 [FREE Full text]
36. Zhou X. The mobile phone's technical feedback, parent-child communication and parenting style: integrating theories of technology acceptance and use. Journal Res 2018(4):108-119 [FREE Full text]
37. Li S, Lei S, She C, Wei R. Friend support matters in bridging digital divide: the impact of social support on the subjective well-being in older adults. Chin J Commun 2023;45(11):81-104. [doi: 10.13495/j.cnki.cjjc.2023.11.003]
38. Fang Y, Chau AK, Wong A, Fung HH, Woo J. Information and communicative technology use enhances psychological well-being of older adults: the roles of age, social connectedness, and frailty status. Aging Ment Health 2018 Nov;22(11):1516-1524. [doi: 10.1080/13607863.2017.1358354] [Medline: 28777010]
39. Cohen S, Wills TA. Stress, social support, and the buffering hypothesis. Psychol Bull 1985 Sep;98(2):310-357. [doi: 10.1037/0033-2909.98.2.310]
40. Jang J. Inter-generational support and the subjective well-being of the elderly: an analysis based on the age difference. South Chin Popul 2016;31(02):60-70.
41. Xin L, Fang Z. The identity and construction of the elderly in digital back-feeding. In: Proceedings of the 2021 International Conference on Culture-oriented Science & Technology. 2021 Presented at: ICCST '21; November 18-21, 2021; Beijing, China p. 623-627 URL: https://ieeexplore.ieee.org/document/9637564 [doi: 10.1109/iccst53801.2021.00135]
42. Hilbert M. Digital gender divide or technologically empowered women in developing countries? A typical case of lies, damned lies, and statistics. Womens Stud Int Forum 2012;34(6):479-489. [doi: 10.1111/j.1552-6909.2012.01399.x]
43. Antonio A, Tuffley D. The gender digital divide in developing countries. Future Internet 2014 Oct 31;6(4):673-687. [doi: 10.3390/fi6040673]
44. Mumporeze N, Prieler M. Gender digital divide in Rwanda: a qualitative analysis of socioeconomic factors. Telematic Inform 2012;34(7):1285-1293. [doi: 10.1111/j.1552-6909.2012.01399.x] [Medline: 22861175]
45. Wang K, Chen XS, Kang SY, Smith BD, Gu D. Older adults' online activities and cognition: investigating the psychological mechanisms and age and gender differences. Soc Sci Med 2024;352:116988. [doi: 10.1016/j.socscimed.2024.116988]

XSL•FO
RenderX

46. Choi E, Han KM, Chang J, Lee YJ, Choi KW, Han C, et al. Social participation and depressive symptoms in community-dwelling older adults: emotional social support as a mediator. J Psychiatr Res 2021 May;137:589-596. [doi: 10.1016/j.jpsychires.2020.10.043] [Medline: 33168196]
47. Wang Q. Digital divide and digital feedback: the status and impact of digital media use among the older adults. Res Radio Telev 2024:49-54 [FREE Full text]
48. Ebimgbo SO, Chukwu NE, Okoye UO. Gender differences in family support to older adults and implications for social work in south-east Nigeria. J Aging Stud 2021 Dec;59:100979. [doi: 10.1016/j.jaging.2021.100979] [Medline: 34794723]
49. Meagley K, Schriver B, Geary RS, Fielding-Miller R, Stein AD, Dunkle KL, et al. The gender dimensions of social networks and help-seeking behaviors of young adults in Soweto, South Africa. Glob Health Action 2016;9:31138 [FREE Full text] [doi: 10.3402/gha.v9.31138] [Medline: 27265147]
50. Bidmon S, Terlutter R. Gender differences in searching for health information on the internet and the virtual patient-physician relationship in Germany: exploratory results on how men and women differ and why. J Med Internet Res 2015 Jun 22;17(6):e156 [FREE Full text] [doi: 10.2196/jmir.4127] [Medline: 26099325]
51. Mackenzie CS, Gekoski WL, Knox VJ. Age, gender, and the underutilization of mental health services: the influence of help-seeking attitudes. Aging Ment Health 2006 Nov;10(6):574-582. [doi: 10.1080/13607860600641200] [Medline: 17050086]
52. Yuan Y, Li J, Fu P, Jing Z, Wang Y, Zhou C. Association between physical activity and falls among older adults in rural China: are there gender and age related differences? BMC Public Health 2022 Feb 19;22(1):356 [FREE Full text] [doi: 10.1186/s12889-022-12773-1] [Medline: 35183149]
53. Wang Y, Li J, Fu P, Jing Z, Zhou C. Social engagement and sleep quality among older adults in rural China: is there a sex difference? J Am Med Dir Assoc 2021 Oct;22(10):2140-2146. [doi: 10.1016/j.jamda.2020.11.014] [Medline: 33309649]
54. Luo N, Liu G, Li M, Guan H, Jin X, Rand-Hendriksen K. Estimating an EQ-5D-5L value set for China. Value Health 2017 Apr;20(4):662-669 [FREE Full text] [doi: 10.1016/j.jval.2016.11.016] [Medline: 28408009]
55. Li Y, Li S. Offspring's education, feedback paths and the digital access divide of the elderly. Doc Inf Knowl 2023;40(04):62-71. [doi: 10.13366/j.dik.2023.04.062]
56. Lu X, Yao Y, Jin Y. Digital exclusion and functional dependence in older people: Findings from five longitudinal cohort studies. eClinicalMedicine 2022 Dec;54(6):101708-101750. [doi: 10.1016/j.eclinm.2022.101708] [Medline: 22861175]
57. Cui Y, He Y, Xu X, Zhou L, Nutakor JA, Zhao L. Cultural capital, the digital divide, and the health of older adults: a moderated mediation effect test. BMC Public Health 2024 Jan 25;24(1):302 [FREE Full text] [doi: 10.1186/s12889-024-17831-4] [Medline: 38273305]
58. Shao Y, Yang X, Chen Q, Guo H, Duan X, Xu X, et al. Determinants of digital health literacy among older adult patients with chronic diseases: a qualitative study. Front Public Health 2025 Mar 25;13:1568043. [doi: 10.3389/fpubh.2025.1568043] [Medline: 40231178]
59. Austin PC. Bayesian extensions of the Tobit model for analyzing measures of health status. Med Decis Making 2002;22(2):152-162. [doi: 10.1177/0272989X0202200212] [Medline: 11958497]
60. Jing Z, Li J, Wang Y, Yuan Y, Zhao D, Hao W, et al. Association of smoking status and health-related quality of life: difference among young, middle-aged, and older adults in Shandong, China. Qual Life Res 2021 Feb 28;30(2):521-530. [doi: 10.1007/s11136-020-02645-9] [Medline: 32989682]
61. The 51st statistical report on China's internet development. China Internet Network Information Center. 2023. URL: https://www.cnnic.net.cn/n4/2023/0303/c88-10757.html [accessed 2025-09-17]
62. Anderson M, Perrin A. Tech adoption climbs among older adults. Pew Research Center. 2017 May 17. URL: https://www.pewresearch.org/internet/2017/05/17/tech-adoption-climbs-among-older-adults/ [accessed 2025-09-17]
63. Prensky M. Digital natives, digital immigrants part 1. OTH 2001 Sep;9(5):1-6. [doi: 10.1108/10748120110424816]
64. Fischer SH, David D, Crotty BH, Dierks M, Safran C. Acceptance and use of health information technology by community-dwelling elders. Int J Med Inform 2014 Sep;83(9):624-635 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.06.005] [Medline: 24996581]
65. Zhuqing X. Digital divided or digital provided? The effective supply of information and the farm-gate price: an empirical study from micro-level. China Economic Quarterly 2013:-. [doi: 10.13821/j.cnki.ceq.2013.04.015]
66. Zajac IT, Flight IH, Wilson C, Turnbull D, Cole S, Young G. Internet usage and openness to internet-delivered health information among Australian adults aged over 50 years. Australas Med J 2012;5(5):262-267 [FREE Full text] [doi: 10.4066/AMJ.2012.1065] [Medline: 22848321]
67. Gatto SL, Tak SH. Computer, internet, and e-mail use among older adults: benefits and barriers. Educ Gerontol 2008 Aug 22;34(9):800-811. [doi: 10.1080/03601270802243697]
68. Mellor D, Firth L, Moore K. Can the internet improve the well-being of the elderly? Ageing Int 2008 Feb 9;32(1):25-42. [doi: 10.1007/s12126-008-9006-3]
69. Jin Y, Hu W, Feng Y. Internet use and the life of older adults aged 50 and above in digital era: findings from a national survey. Population Research 2024;48(01):40-55.

XSL•FO
RenderX

70. Liu S, Lu Y, Wang D, He X, Ren W, Kong D, et al. Impact of digital health literacy on health-related quality of life in Chinese community-dwelling older adults: the mediating effect of health-promoting lifestyle. Front Public Health 2023 Jun 21;11:1200722 [FREE Full text] [doi: 10.3389/fpubh.2023.1200722] [Medline: 37415711]
71. Wen Z, Mo M, Xu J. The internal mechanism and micro-evidence of digital feedback driving the participation of rural elderly in smart home care. Mathematics 2025;13(11):105-116. [doi: 10.16582/j.cnki.dzzw.2024.03.009]
72. Zhou Y, Ding H. The overview and influential factors of digital backfeeding among the three generations of Chinese families. Chinese Journal of Journalism & Communication 2020;42(03):6-31. [doi: 10.13495/j.cnki.cjjc.2020.03.001]
73. Phang JK, Kwan YH, Yoon S, Goh H, Yee WQ, Tan CS, et al. Digital intergenerational program to reduce loneliness and social isolation among older adults: realist review. JMIR Aging 2023 Jan 04;6:e39848 [FREE Full text] [doi: 10.2196/39848] [Medline: 36598801]
74. Bhandari A. Gender inequality in mobile technology access: the role of economic and social development. Information, Communication & Society 2019 Jan 10;22(5):678-694. [doi: 10.1080/1369118X.2018.1563206]
75. Kim J, Lee HY, Christensen MC, Merighi JR. Technology access and use, and their associations with social engagement among older adults: do women and men differ? J Gerontol B Psychol Sci Soc Sci 2017 Sep 01;72(5):836-845. [doi: 10.1093/geronb/gbw123] [Medline: 28073816]
76. Russell C. What do older women and men want? Gender differences in the 'lived experience' of ageing. Current Sociology 2012;55(2):173-192. [doi: 10.1177/0011392107073300]
77. Bai X, Guo Y, Fu YY. Self-image and intergenerational relationships as correlates of life satisfaction in Chinese older adults: will gender make a difference? Ageing Soc 2017 Feb 27;38(7):1502-1519. [doi: 10.1017/s0144686x17000113]

Abbreviations

HRQOL: health-related quality of life
ICT: information and communication technology
QOL: quality of life

Edited by PH Chau; submitted 13.04.25; peer-reviewed by J Chen, C Liu; comments to author 07.07.25; revised version received 19.08.25; accepted 11.09.25; published 02.10.25.

Please cite as:
Che X, Chai S, Zhao D, Chen S, Zhou C
Gender Differences in the Digital Divide, Digital Back-Feeding, and Health-Related Quality of Life Among Rural Older Adults: Cross-Sectional Study
JMIR Aging 2025;8:e75925
URL: https://aging.jmir.org/2025/1/e75925
doi:10.2196/75925
PMID:

©Xin Che, Shujun Chai, Dan Zhao, Shirong Chen, Chengchao Zhou. Originally published in JMIR Aging (https://aging.jmir.org), 02.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Digital Pathways to Reducing Depression Among Aging Populations Through the “Broadband China” Pilot Program: Quasi-Natural Experiment

Zhiying Li[1], PhD

School of Government, University of Chinese Academy of Social Sciences, Beijing, China

Corresponding Author:
Zhiying Li, PhD
School of Government
University of Chinese Academy of Social Sciences
11 Changyu Street
Fangshan District
Beijing, 102488
China
Phone: 86 15653730005
Email: 124671336@qq.com

Abstract

Background: With the rise of digital technology, infrastructure development has become vital for social welfare and public health. However, evidence on its effects on depressive symptoms among middle-aged and older adults remains limited.

Objective: This study evaluates the impact of digital infrastructure development on depressive symptoms among middle-aged and older adults, focusing on underlying mechanisms, heterogeneous effects, and health inequalities.

Methods: We use longitudinal data from the China Health and Retirement Longitudinal Study (CHARLS), 2011-2020 (N=56,211). Exploiting the quasi-natural experiment of the “Broadband China” pilot policy, we apply a difference-in-differences approach to estimate the effect on depressive symptoms. Mediation analysis follows the Baron-Kenny 3-step procedure, with bootstrap tests (95% CI) for robustness, and causal interpretation relies on standard assumptions for observational data. Subgroup analyses explore heterogeneity across age, education, and sex groups.

Results: Our findings indicate that the “Broadband China” pilot significantly reduces depressive symptoms among middle-aged and older adults (−0.33, P<.01). The positive effect is primarily mediated through strengthened social networks, including increased family connection, close social interactions, and greater social participation. Heterogeneity analysis shows that the benefits for depression reduction are more pronounced among women (−0.38, P<.01), middle-aged adults (−0.41, P<.01), and those with lower levels of education (−0.33, P<.01). Moreover, the results suggest that digital infrastructure plays a compensatory role in mitigating health disparities, thereby reducing inequalities in depression outcomes (−0.01, P<.01).

Conclusions: Digital infrastructure reduces depressive symptoms among aging populations mainly by strengthening social networks. Embedding infrastructure into long-term strategies, enhancing digital literacy, and integrating digital health services are key to promoting healthy aging and reducing inequalities.

(JMIR Aging 2025;8:e79260) doi:10.2196/79260

KEYWORDS

digital infrastructure; depression; social networks; health inequality; middle-aged and older adults; difference-in-differences

Introduction

Background

Population aging has become a significant challenge to global socioeconomic development [1]. According to the United Nations’ 2024 World Population Prospects, by the end of the 21st century, the global population aged 65 years and older is expected to reach 2.2 billion, surpassing the population of children younger than 18 years, and this proportion will continue to rise. China’s aging process is particularly rapid, with the population aged 65 years and older accounting for 15.6% of the total in 2024, and this share is expected to accelerate in the coming decades. In this context, the health—particularly the

XSL•FO
RenderX

mental health dimension, such as depression—of middle-aged and older adults is not only crucial for individual well-being but also profoundly impacts health care resource allocation, the stability of social security systems, and sustainable economic development.

In recent years, improvements in social security systems and advancements in medical technologies have somewhat enhanced the physical health of middle-aged and older adults, but depression has still not received adequate attention [2]. Depression is one of the most common psychological disorders among older adults. Studies show that the global prevalence of major depression is as high as 9.7% [3], while the depression rate for middle-aged and older adults in China is 24.1% [4], with severe depression accounting for 2.3% [5]. Depressive symptoms not only increase individual medical expenditures and social health care burdens [6] but may also impair the social functioning of middle-aged and older adults, reduce their quality of life, and further affect social welfare systems and macroeconomic development [7]. However, current health evaluation systems often underestimate the long-term socioeconomic impacts of depression, leading to its neglect in public policy and resource allocation [8]. In this context, how to effectively reduce depressive symptoms among middle-aged and older adults has become a key issue for advancing healthy aging.

Social networks are considered one of the most important factors affecting the depression status of middle-aged and older adults. As retirement, children's migration for work, and the increased mobility of social networks take place, interactions with the outside world may decrease, leading to an increased sense of social isolation, which exacerbates depressive symptoms [9]. Social networks not only include direct face-to-face interactions but also extend to remote social connections facilitated by digital technology [10]. Existing studies show that middle-aged and older adults with lower levels of social networks are more prone to anxiety and depression, while actively maintaining social relationships helps reduce depressive symptoms [11]. The promotion of digital infrastructure development and the widespread availability of the internet have provided middle-aged and older adults with new social channels, alleviating social isolation to some extent. The application of digital technology can help reduce depression by enhancing social networks and increasing social support [12]. On the one hand, the internet allows middle-aged and older adults to maintain closer contact with their children and friends, reducing social alienation and thus lowering feelings of loneliness and risk of depression [13]. On the other hand, the use of social media and web-based interaction platforms expands their social networks, enhancing their sense of social participation and helping to alleviate depressive symptoms [14].

However, there is still controversy regarding the impact of digital technology on depression among middle-aged and older adults. One view holds that digital technology can compensate for the social interaction and participation deficits of middle-aged and older adults, thus reducing depressive symptoms [15]. Another view, based on the "substitution effect of presence," argues that excessive reliance on digital technology may reduce real-world social interactions, weakening the emotional support provided by face-to-face communication, and thus negatively affecting depression outcomes [16]. Moreover, excessive use of social media may also lead to issues such as information overload and internet fraud, increasing anxiety and psychological stress [17]. Although these studies provide valuable insights into the relationship between digital technology and depression, there is still no consensus in the academic community regarding how digital infrastructure impacts the depressive symptoms of middle-aged and older adults. Current research mainly focuses on internet use at the individual level, with less attention paid to the systemic impact of digital infrastructure development as an exogenous variable. Additionally, the role of social networks in this process remains underexplored. Therefore, examining how digital infrastructure development influences the depressive symptoms of middle-aged and older adults through the enhancement of social networks is not only of significant academic value but also has practical implications for policy-making on healthy aging.

Based on this, this paper uses the quasi-natural experiment of the "Broadband China" pilot policy to systematically examine the effect of digital infrastructure development on depressive symptoms among middle-aged and older adults and explore the mechanisms through social networks. Compared with existing studies, the main contributions of this paper are as follows: First, rigorous identification leveraging policy shocks. Existing research has focused on the correlation between individual internet use and depression, making it difficult to overcome endogeneity issues. This paper uses the "Broadband China" pilot, applying the difference-in-differences (DID) method to estimate the effect of digital infrastructure development on depressive symptoms, providing more externally valid empirical evidence for policy optimization. Second, it reveals the role of social networks. Although social networks are considered an important factor influencing depression, few studies have focused on how digital infrastructure reshapes the social patterns of middle-aged and older adults and their implications for depressive symptoms. This paper focuses on how digital infrastructure enhances family connection, close social interactions, and social participation, improving social interaction levels, thereby alleviating loneliness and reducing depressive symptoms, filling a gap in the literature. Third, expanding the perspective of health inequality. This paper further examines the heterogeneous effects of digital infrastructure development across different sexes, ages, and socioeconomic groups and, by combining the Kakwani relative deprivation index, reveals how the digital divide affects inequality in depression outcomes among middle-aged and older adults. This analysis not only deepens the understanding of the social effects of digital infrastructure development but also provides empirical support for the government to develop targeted healthy aging policies.

Policy Background and Theoretical Framework

Policy Background

As an essential component of digital infrastructure, broadband networks are not only the cornerstone of information technology development but also have a profound impact on residents' psychological well-being and risk of depression. To accelerate

XSL•FO
RenderX

the spread of broadband networks, the State Council released the "Broadband China Strategy and Implementation Plan" in August 2013, which emphasized promoting the nationwide construction of high-speed, secure, and green digital infrastructure through regional pilot projects. From 2014 to 2016, pilot cities were selected in 3 batches, gradually expanding the scope of broadband access services [18]. The core objective of this strategy was to improve network coverage and transmission speed to meet the demands of socioeconomic entities for high-quality internet services, while also promoting the integration of broadband networks with other infrastructure and public services through increased fiscal investment, government guidance, interdepartmental collaboration, and technological innovation [19].

In the field of depression and psychological well-being, the popularization of broadband networks has provided essential support for the social connections and emotional health of middle-aged and older adults. High-quality internet access not only enables them to maintain closer contact with family and friends, reducing feelings of social isolation, but also lays the foundation for the widespread use of web-based social platforms, community interaction, and remote psychological services. Through the application of digital technologies, older adults can more easily participate in community activities, social organizations, and volunteer programs, enhancing their sense of social participation and belonging, which in turn helps reduce depressive symptoms [20]. Therefore, the "Broadband China" strategy is not only an important initiative in advancing the nation's informatization construction but also provides a natural quasi-experiment for researching how digital infrastructure influences depression among middle-aged and older adults.

Theoretical Framework

The theory of resocialization suggests that individuals must continually adapt to new social environments, roles, and norms throughout their lives to maintain physical and psychological health, particularly in preventing depression, and realize self-worth. However, traditional social structure changes, adjustments in intergenerational relationships, and role transitions after retirement present challenges for middle-aged and older adults, such as shrinking social networks and reduced social participation, which may exacerbate feelings of loneliness and psychological imbalance [21]. In this context, digital infrastructure construction, as an essential support for social modernization, provides middle-aged and older adults with new pathways for socialization [22], enabling them to access information more conveniently and expand their social interactions, thus reducing depression risk [23]. Existing studies have shown that internet use can effectively alleviate social isolation among middle-aged and older adults and improve their social adaptability and health [24]. By optimizing information accessibility and the convenience of social interactions, digital infrastructure offers middle-aged and older adults broader social channels, enhancing their adaptability to social changes. This not only reduces loneliness but also increases a sense of belonging and self-efficacy, thereby lowering depressive symptoms [25].

The theory of social support further elucidates the mechanism through which digital infrastructure construction affects depression outcomes of middle-aged and older adults. Social support includes emotional support, belonging, being valued, practical help, as well as information and guidance, all of which are vital external resources affecting mental health [26]. The improvement of digital infrastructure has facilitated efficient interactions among family members, allowing middle-aged and older adults to communicate more smoothly with children and friends, thus alleviating emotional loneliness due to generational gaps or physical separation [27]. The development of digital services, such as smart payment, web-based government services, and telemedicine, has reduced the uncertainty in daily life, enabling middle-aged and older adults to gain greater autonomy and security in social activities, health care, and financial transactions, helping individuals better understand and adjust their psychological states and enhancing their ability to cope with psychological stress. Based on this, the following hypothesis is proposed:

- **H1:** Digital infrastructure development reduces depressive symptoms among middle-aged and older adults.

Social networks refer to stable relational systems formed during individuals' social interactions [28], typically consisting of both strong and weak ties. Strong ties are primarily found among close family members and intimate friends, characterized by strong emotional connections and stability [29], while weak ties include neighbors, community members, or interest groups with lower interaction frequencies. Although emotional bonds are weaker, weak ties offer greater information heterogeneity and facilitate resource acquisition and social integration [30]. An individual's social network is not only a platform for exchanging information and resources but also a significant source of emotional support and social recognition. However, as people age, the social networks of middle-aged and older adults gradually shrink, leading to fewer social opportunities and increased loneliness, which negatively affects their depression outcomes [31]. Research indicates that emotional support has a significant impact on depressive symptoms among older adults [32], and the widespread adoption of digital infrastructure provides middle-aged and older adults with low-cost and convenient social channels, allowing them to reconstruct social connections through digital platforms such as social media and web-based communities, expanding their social circles and maintaining interactions within their existing social networks. On the one hand, this strengthens the strong-tie network between middle-aged and older adults and their families and friends, enabling them to maintain more frequent emotional connections despite geographical and temporal limitations [13]. On the other hand, it expands weak-tie networks, allowing them to connect with like-minded peers through web-based communities and interest forums, enhancing social interaction and social support [33]. This dual optimization mechanism not only improves emotional support but also increases social participation, effectively alleviating loneliness and psychological stress, thereby reducing depressive symptoms. Based on this, the following hypotheses are proposed:

XSL•FO
RenderX

- **H2:** Digital infrastructure development improves the depression outcomes of middle-aged and older adults by enhancing their social networks.
- **H2a:** Digital infrastructure development alleviates depressive symptoms among middle-aged and older adults by strengthening their family connections.
- **H2b:** Digital infrastructure development alleviates depressive symptoms among middle-aged and older adults by enhancing their close social interactions.
- **H2c:** Digital infrastructure development alleviates depressive symptoms among middle-aged and older adults by promoting their social participation.

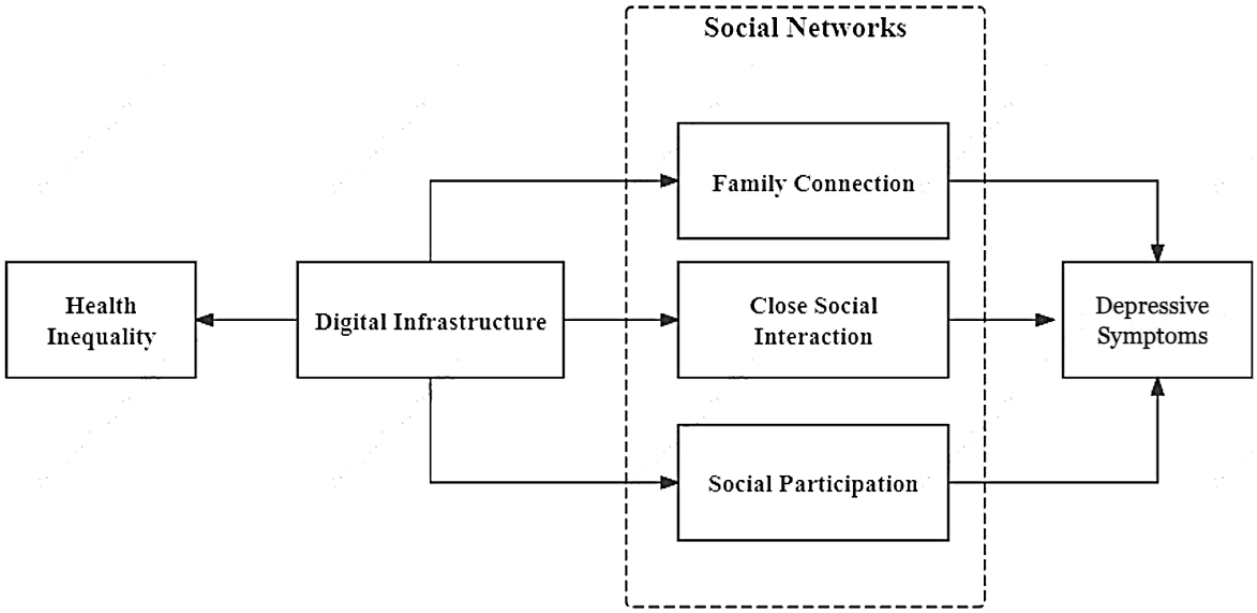
As digital technologies continue to develop, the digital divide has become an important manifestation of social inequality, exacerbating economic inequality [34], knowledge inequality [35], and educational inequality [36], as well as having a profound impact on depression among different groups. Particularly in rural areas with underdeveloped digital infrastructure and among middle-aged and older adults with low incomes and educational levels, inadequate access to information often prevents them from obtaining necessary resources for depression prevention and intervention in a timely manner. Numerous studies have shown that information asymmetry is a key factor in the deterioration of depressive symptoms [37], especially regarding knowledge about depression, intervention measures, and access to medical resources [38]. The construction of digital infrastructure, by providing more convenient internet access and information channels, can help these disadvantaged groups break through information barriers and gain timely knowledge and web-based support related to depression, thus effectively narrowing the inequalities in depression outcomes caused by information asymmetry and regional differences.

The construction of digital infrastructure is not only a tool for the flow of information but also provides new avenues for the accumulation of social capital. Social capital theory posits that an individual's depression status is closely related to the breadth and depth of their social support network [39]. For middle-aged and older adults, a lack of social support and feelings of isolation are major factors contributing to depressive symptoms, especially for economically disadvantaged groups. Traditional social support networks, constrained by geographical, cultural, and generational differences, often fail to meet the needs of these groups. The widespread use of digital technologies, particularly the internet and social media, offers middle-aged and older adults more social platforms and interaction channels, enabling them to break the limitations of traditional social networks and establish digital connections with other groups, gaining more emotional support and psychological guidance [40]. Through these digital platforms, older adults can form new social support networks, significantly reducing the inequalities in depression caused by differences in social capital. Based on this, the following hypothesis is proposed:

- **H3:** Digital infrastructure development helps reduce inequalities in depressive symptoms among middle-aged and older adults.

In summary, the research framework for this study is illustrated in Figure 1.

Figure 1. Research framework. This figure presents the conceptual framework of the study. All variables are measured in each wave of the panel data (2011, 2013, 2015, 2018, and 2020).



Methods

Data Source

The data for this study are drawn from the China Health and Retirement Longitudinal Study (CHARLS), a nationally representative survey conducted by the National School of Development at Peking University. Since 2011, CHARLS has carried out follow-up surveys every 2 to 3 years, covering 150 districts and counties nationwide through a multi-stage, probability-proportional-to-size sampling design [41,42]. The survey repeatedly collects extensive information on

XSL•FO
RenderX

respondents—including demographic characteristics, socioeconomic conditions, health status, and family structure—ensuring comparability across waves.

This study uses 5 survey waves (2011, 2013, 2015, 2018, and 2020). The dependent variable, depressive symptoms (measured by the Center for Epidemiologic Studies Depression Scale [CES-D] 10 scale), is consistently available across all waves. Time-varying individual-level covariates—such as age, sex, marital status, education, family size, pension status, insurance coverage, and chronic disease status—are also updated in each wave. The key explanatory variable, digital infrastructure development, is constructed by linking city-level "Broadband China" pilot policy data to respondents based on their place of residence in each survey year.

The original pooled dataset contained more than 60,000 person-wave observations. To handle missing data, we applied listwise deletion, excluding respondents with incomplete information on the dependent variable or any covariates. The final analytic sample comprises 56,211 person-wave observations of adults aged 45 years and older. Because the proportion of missing cases was modest and not systematically patterned, the resulting sample remains broadly representative, and the risk of bias due to missing data is minimal.

Ethical Considerations

The CHARLS project was reviewed and approved by the Institutional Review Board of Peking University (IRB00001052-11015). All participants, including middle-aged and older adults, provided written informed consent prior to participation. This study used deidentified, publicly available data, and all procedures complied with the ethical standards of the Declaration of Helsinki. Participants' privacy and confidentiality were maintained throughout the study. No further ethics approval was required for this secondary data analysis.

Variable Definitions

The core explanatory variable of this study is digital infrastructure development. According to the "Broadband China" Strategy and Implementation Plan issued by the State Council in 2013, the core goals of the strategy include enhancing broadband network coverage, optimizing network service quality, and promoting the development of digital applications to improve regional digital infrastructure. The implementation of the "Broadband China" strategy not only represents the government's systematic investment in digital infrastructure but also has created significant regional infrastructure disparities, providing a suitable policy experiment environment for this study. Therefore, following the approaches of Peng et al [43], He et al [44], and Jia and Li [45], this study uses the "Broadband China" pilot policy as a proxy for digital infrastructure development. If middle-aged and older adult individual i in city c was included in the "Broadband China" pilot program in year t, the binary variable did is assigned a value of 1; otherwise, it is assigned a value of 0. For individuals who changed residence during the observation period, the assignment of did was based on the policy status of their current city in each survey wave, thereby dynamically capturing the actual local policy environment experienced by each respondent. Since the proportion of cross-city movers in the sample was very small, the potential influence of such cases on the estimation results is negligible.

The dependent variable in this study is depressive symptoms among middle-aged and older adults. Based on the work of Ma et al [46], depressive symptoms were measured using the CES-D. This scale includes 10 questions related to core symptoms and common manifestations of depression, such as: "I get upset over little things," "I have trouble concentrating," "I feel downhearted," "I feel that things are an effort," "I am hopeful about the future," "I feel frightened," "I have trouble sleeping," "I feel happy," "I feel lonely," and "I feel like I can't continue my life." CHARLS asks respondents about their feelings and behaviors during the previous week, with response options ranging from "Rarely or none of the time (<1 day)" to "Most of the time (5-7 days)." Each item is scored from 0 to 3, with 2 items, "I am hopeful about the future" and "I feel happy," being reverse-coded to align with the other items. The total score for these 10 items is the CES-D score, which serves as the depression outcome variable in this study, with a range from 0 to 30. A higher score indicates worse depressive symptoms. The CES-D has been widely applied in studies of middle-aged and older adults, and previous research has demonstrated its good reliability and validity [47,48].

The mechanism variable in this study is the social network of middle-aged and older adults, which includes family connections, close social interactions, and social participation. According to social network theory, the network is not only characterized by the breadth of relationships but also by their intimacy [30]. Therefore, the 3 dimensions selected in this study reflect varying degrees of intimacy in social relationships. Family connection, as the most intimate dimension of social contact, is measured by the frequency of contact with children. This frequency is based on the response to the question "How often do you see your children when you don't live with them?" The response scale covers multiple time frames, ranging from "almost every day," "two to three times per week," "once a week," "once every two weeks," "once a month," "once every three months," "once every six months," and "once a year," to "almost never." Thus, the measure captures the regularity of parent-child contact across daily, weekly, monthly, and annual intervals. Values range from 1 to 9, where a higher score indicates less frequent contact and reflects weaker family connections. For middle-aged and older adults with multiple children, the mean frequency across all children is used. Close social interactions represent looser social relationships and are measured by the response to "Have you visited friends or socialized in the past month?" A response of "Yes" is assigned a value of 1, and "No" is assigned a value of 0. Close social interactions reflect the frequency of interaction with nonfamily members, which typically offers emotional support and psychological relief, while indicating interaction with a more distant social circle. Social participation reflects broader social connections and indicates the extent of involvement in public and external social arenas. Social participation is measured by responses to whether, in the past month, the individual participated in activities such as "playing mahjong, chess, cards, or attending community events," "joining social or volunteer

XSL•FO
RenderX

activities," or "attending educational or training courses." A response of "Yes" is scored as 1, otherwise 0. The social participation score is the sum of these 4 activities, ranging from 0 to 4, indicating the individual's involvement in social interaction and public affairs.

This study also controls for other factors that may affect the depressive symptoms of middle-aged and older adults, including individual characteristics (such as residence, sex, age, education level, ethnicity, pension insurance, health insurance, and chronic diseases) and family characteristics (such as marital status and household size). Furthermore, the study controls for time fixed effects and city fixed effects. All data management and statistical analyses were performed using Stata (version 18.0; Stata Corp). Descriptive statistics for the relevant variables are shown in Table 1, with additional descriptive statistics by survey wave reported in Multimedia Appendix 1.

Table 1. Variable definitions and descriptive statistics.

Variable type and variable name	Variable definition and assignment	Mean (SD)	Min	Max
Dependent variable				
Depressive symptoms	A continuous variable is constructed based on the CES-D[a] (higher score indicates more severe depressive symptoms)	8.223 (6.221)	0	30
Core explanatory variable				
Did	"Broadband China" pilot policy × year	0.208 (0.406)	0	1
Control variables				
Residence	Respondent's residence (Urban = 0, Rural = 1)	0.597 (0.491)	0	1
Sex	Respondent's sex (Female = 0, Male = 1)	0.483 (0.5)	0	1
Age (years)	Respondent's actual age at the time of the survey	61.41 (10.008)	46	120
Education	Respondent's education (Below Primary = 1, Primary = 2, Secondary = 3, High School and Above = 4)	2.018 (1.062)	1	4
Ethnicity	Ethnicity of Respondents (Non-Han = 0, Han = 1)	0.924 (0.266)	0	1
Pension	Respondent's pension insurance status (No pension insurance = 0, Any pension insurance = 1)	0.59 (0.492)	0	1
Health insurance	Respondent's health insurance status (No health insurance = 0, Any health insurance = 1)	0.948 (0.222)	0	1
Chronic disease	Whether the respondent has chronic diseases (No = 0, Yes = 1)	0.764 (0.425)	0	1
Marital status	Respondent's marital status (Not married = 0, Married = 1)	0.856 (0.351)	0	1
Household size	Number of family members of the respondent	3.176 (1.639)	1	16

[a]CES-D: Center for Epidemiologic Studies Depression Scale.

Model Specification

In economic and social research, the DID method is commonly used to assess policy effects. The fundamental approach involves treating the policy pilot as an exogenous "quasi-natural experiment." The implementation of the Broadband China pilot policy may generate pre- and postimplementation differences within the pilot regions, as well as differences between pilot and nonpilot regions at the same point in time. By performing regression analysis on these 2 types of differences, the net effect of the policy on the depressive symptoms of middle-aged and older adults can be effectively identified. Therefore, this study treats the Broadband China policy as a quasi-natural experiment and uses the DID method to examine its impact. The baseline regression model is specified as follows:

$$Depression_{ict}=\alpha_0+\beta_1 DID_{ct}+\beta_2 Controls_{ict}+\gamma_c+\delta_t+\varepsilon_{ict} \quad (1)$$

Where the subscripts *i*, *c*, and *t* represent individual, city, and time, respectively. $Depression_{ict}$ is the dependent variable, indicating the level of depressive symptoms of middle-aged and older adult individual i in city c at year t. DID_{ct} is the core explanatory variable, representing whether city c implemented the "Broadband China" pilot in year t. It is the interaction term between the city-level treatment group indicator and the pre and postpilot implementation dummy variables for the "Broadband China" initiative. $Controls_{ict}$ represents the individual and household-level control variables. γ_c and δ_t are the city fixed effects and time fixed effects, respectively. ε_{ict} is the random error term.

The validity of the estimation in equation 1 relies on the parallel trend assumption, which asserts that, in the absence of policy intervention, the dependent variable trends of the treatment and control groups should exhibit similar patterns. To test this assumption and explore the dynamic effects of the policy, the study adopts an event study approach, following the methodology of Beck et al [49]. The model for the dynamic effect analysis is as follows:

XSL•FO
RenderX



Where β_t represents the estimated values corresponding to the years 2011-2020. The definitions of other variables are consistent with those in equation 1.

The Kakwani relative deprivation index [50] effectively quantifies disparities in depressive symptoms. According to the theory of relative deprivation, individuals with poorer health experience greater health deprivation and higher levels of health inequality. This study follows the methodologies of Li et al [41] and Turguttopbaş [51] and uses the Kakwani index to measure inequality in depressive symptoms. Specifically, since our data are panel in nature, the index was computed on a wave-by-wave basis (2011, 2013, 2015, 2018, and 2020). In each wave, the depressive symptom levels of middle-aged and older adults were ranked, and the relative deprivation value of each individual was calculated within that year's reference group. As a result, every respondent has a time-varying relative deprivation value, reflecting their relative health position in each survey year. Formally, let Y represent the reference group in year t, with a sample size of n, and depressive symptom distribution vector $y=(y_1, y_2, y_3, y_{n-1}, y_n)$, where $y_1 \le y_2 \le y_3 \le y_{n-1} \le y_n$. Therefore, the relative deprivation index RD (y_j, y_i) of middle-aged and older adult individual i compared with individual j is defined as:



Based on equation 4, the average relative deprivation index of depressive symptoms for individual i in year t is calculated as follows:



Where μ_y is the average depression score of all samples in the reference group Y, n_{yi}^{+} is the number of middle-aged and older adults in the reference group Y whose health exceeds the health level of individual i (denoted as y_i), and μ_{yi}^{+} is the average health score of middle-aged and older adults in the reference group Y whose health exceeds y_i.

To address potential multicollinearity, we calculated variance inflation factors (VIF) for all explanatory variables. All VIF values were well below the conventional threshold of 10 (mean VIF=1.17), suggesting that multicollinearity is not a concern. Model fit is evaluated using adjusted R^2 (within), reported in each regression table.

Results

Baseline Regression

Table 2 presents the baseline regression results on the impact of digital infrastructure development on the depressive symptoms of middle-aged and older adults. In Model 1, which includes only the core explanatory variables, the coefficient of the interaction term is negative and statistically significant. Models 2 and 3 include individual and household characteristics, respectively, and the coefficient of the interaction term remains significant, indicating the robustness of the results. Model 4 further incorporates time and regional fixed effects. The coefficient of the interaction term is –0.3349 and is statistically significant at the 1% significance level. These results suggest that digital infrastructure development is associated with a significant reduction in depressive symptoms among middle-aged and older adults, thus supporting Hypothesis 1.

XSL•FO
RenderX

Table 2. Baseline regression results.

Variable	Model 1	Model 2	Model 3	Model 4
Did				
β (95% CI)	–0.6073 (–0.7288 to –0.4858)	–0.4647 (–0.5824 to –0.3469)	–0.4799 (–0.5976 to –0.3622)	–0.3349 (–0.5055 to –0.1644)
P value	<.001	<.001	<.001	<.001
Sex				
β (95% CI)	__[a]	–1.7454 (–1.8462 to –1.6445)	–1.6422 (–1.7435 to –1.5410)	–1.7217 (–1.8661 to –1.5772)
P value	—	<.001	<.001	<.001
Residence				
β (95% CI)	—	1.2259 (1.1219 to 1.3299)	1.2385 (1.1347 to 1.3422)	0.8625 (0.6738 to 1.0513)
P value	—	<.001	<.001	<.001
Age (years)				
β (95% CI)	—	0.0219 (0.0154 to 0.0283)	0.0068 (0.0002 to 0.0135)	0.0130 (0.0045 to 0.0215)
P value	—	<.001	.05	.003
Education				
β (95% CI)	—	–0.7438 (–0.7942 to –0.6933)	–0.7327 (–0.7830 to –0.6823)	–0.6442 (–0.7184 to –0.5701)
P value	—	<.001	<.001	<.001
Ethnicity				
β (95% CI)	—	–0.4684 (–0.6942 to –0.2426)	–0.4512 (–0.6764 to –0.2261)	–0.2496 (–0.7019 to 0.2026)
P value	—	<.001	<.001	.28
Pension				
β (95% CI)	—	–0.2344 (–0.3517 to –0.1172)	–0.2155 (–0.3326 to –0.0984)	–0.3700 (–0.4959 to –0.2441)
P value	—	<.001	<.001	<.001
Health insurance				
β (95% CI)	—	–0.7111 (–0.9560 to –0.4661)	–0.6147 (–0.8592 to –0.3702)	–0.5423 (–0.8237 to –0.2609)
P value	—	<.001	<.001	<.001
Chronic disease				
β (95% CI)	—	2.2477 (2.1329 to 2.3624)	2.2626 (2.1481 to 2.3771)	2.0488 (1.9152 to 2.1824)
P value	—	<.001	<.001	<.001
Marital status				
β (95% CI)	—	—	–1.3494 (–1.5068 to –1.1921)	–1.2600 (–1.4946 to –1.0255)
P value	—	—	<.001	<.001
Household size				
β (95% CI)	—	—	–0.0339 (–0.0645 to –0.0032)	–0.0669 (–0.1041 to –0.0296)
P value	—	—	.03	<.001
_cons				
β (95% CI)	8.1897 (8.1325 to 8.2470)	7.9973 (7.4786 to 8.5160)	10.0024 (9.4298 to 10.5750)	9.7019 (8.8907 to 10.5132)
P value	<.001	<.001	<.001	<.001
Time-fixed effects	No	No	No	Yes

XSL•FO
RenderX

Variable	Model 1	Model 2	Model 3	Model 4
Regional-fixed effects	No	No	No	Yes
N	56,211	56,211	56,211	56,211
Adjusted R^2	0.0017	0.0952	0.1001	0.1421

[a]Not available.

Robustness Check

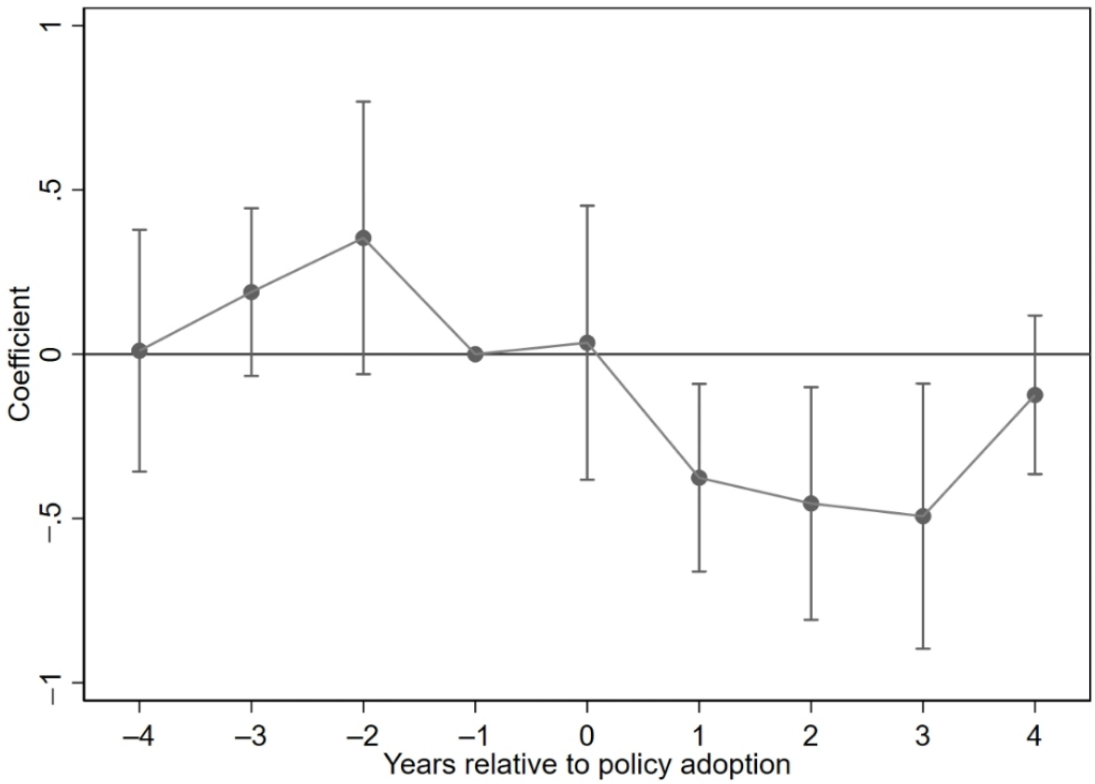
Overview

To ensure the robustness of the research findings, multiple methods were used for robustness checks, including the parallel trend test, placebo test, propensity score matching difference-in-differences (PSM-DID), substitution of the dependent variable, and sample period adjustment.

Parallel Trend Test

The DID method relies on the parallel trend assumption, which requires that the depressive symptom trends of the treatment and control groups are similar prior to policy implementation. To examine this, we use an event study approach using the prepolicy period immediately before implementation (event=−1) as the reference. As a result, no coefficient or CI is reported for this baseline period. Figure 2 plots the estimated coefficients and their 95% CIs, showing that the prepolicy coefficients fluctuate around zero and are not statistically significant, suggesting no systematic differences in depressive symptom trajectories between the 2 groups.

Figure 2. Parallel trend test.



The corresponding regression results are reported in Table 3. Furthermore, a joint F test of the prepolicy coefficients (event –4, event –3, and event –2) yields $F_{3,\ 16,981}$ of 1.81 with a P value of .143, indicating that they are not jointly significant. Taken together, these results support the validity of the parallel trend assumption.

XSL•FO
RenderX

Table 3. Event study estimates for the parallel trends test.

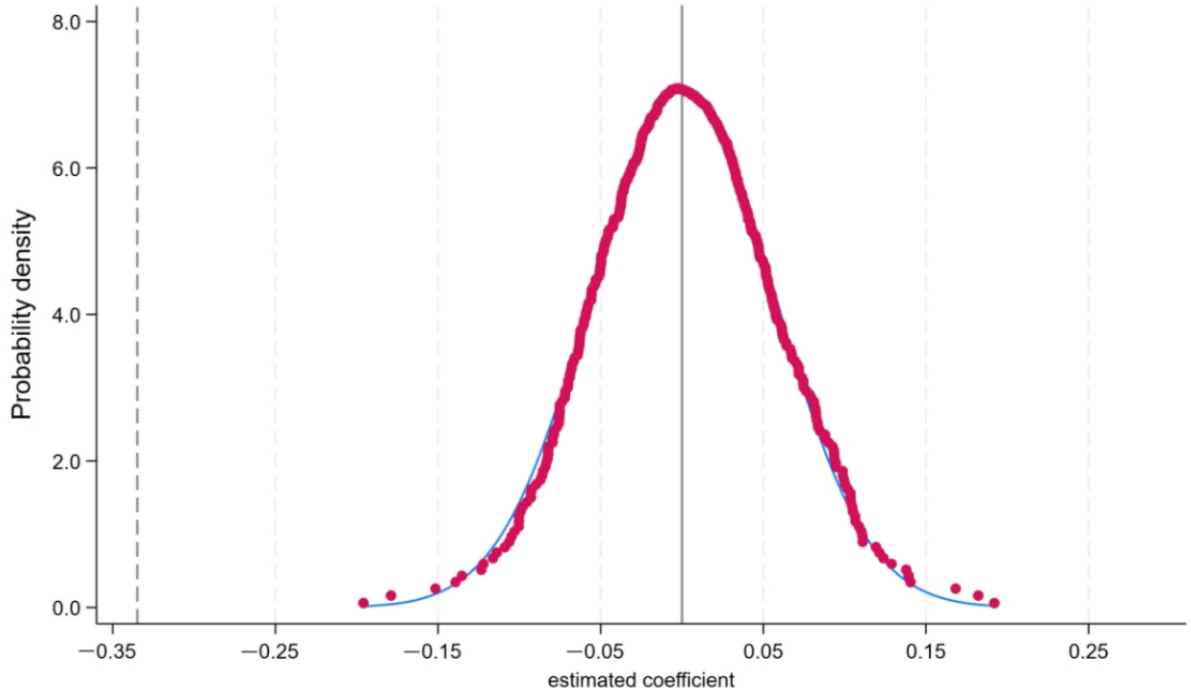
Variables	Coefficient	SE	*T*	*P* value	95% CI
Event –4	0.010	0.188	0.06	.96	–0.358 to 0.379
Event –3	0.189	0.130	1.45	.15	–0.066 to 0.444
Event –2	0.354	0.212	1.67	.09	–0.061 to 0.769
Event –1	__[a]	—	—	—	—
Event 0	0.035	0.213	0.16	.87	–0.382 to 0.452
Event 1	–0.376	0.146	–2.58	.01	–0.662 to –0.090
Event 2	–0.455	0.181	–2.51	.01	–0.809 to –0.100
Event 3	–0.493	0.206	–2.40	.02	–0.897 to –0.090
Event 4	–0.124	0.123	–1.01	.31	–0.366 to 0.117
Control variables	Yes	Yes	Yes	Yes	Yes
_cons	9.744	0.428	22.77	<.001	8.905 to 10.583

[a]Omitted.

Placebo Test

To further verify the robustness of the policy effect and eliminate the interference of unobservable factors, this study randomly selected the experimental group and the policy pilot period to generate a simulated interaction term. A placebo test was conducted by randomly sampling the “Broadband China” policy interaction term 500 times. Figure 3 presents the kernel density distribution of the 500 random regression coefficients and the scatter plot of the *P* values. The results show that the mean of the simulated interaction term coefficients is close to zero, with a distribution that approximates normality. Most *P* values are greater than 0.1 and do not reach statistical significance. This indicates that the effect of the “Broadband China” policy is not influenced by sample selection bias, supporting the robustness of the baseline regression results.

Figure 3. Placebo test.



PSM-DID

To control for selection bias between pilot and nonpilot regions, this study uses the PSM-DID method for robustness checks. Using nearest neighbor matching (1:4) and local linear regression matching, samples from nonpilot regions with similar socioeconomic characteristics are selected as the control group. DID regression is then implemented on the matched samples. The results from Models 1-2 in Table 4 show that the DID estimate coefficient for the Broadband China pilot policy remains significantly negative, further confirming the beneficial effect of digital infrastructure development on depressive symptoms among middle-aged and older adults.

XSL•FO
RenderX

Table 4. Robustness checks: propensity score matching difference-in-differences (PSM-DID), substitution of the dependent variable, and sample period adjustment.

Variable	Model 1 (PSM-DID)	Model 2 (PSM-DID)	Model 3 (substitution of the dependent variable)	Model 4 (sample period adjustment)
Did				
β (95% CI)	–0.3713 (–0.5480 to –0.194)	–0.4219 (–0.6010 to –0.2429)	0.0266 (0.0028 to 0.0504)	–0.4312 (–0.6136 to –0.2488)
P value	<.001	<.001	.03	<.001
Control variables	Yes	Yes	Yes	Yes
_cons				
β (95% CI)	9.6373 (8.6974 to 10.5770)	9.7868 (8.6852 to 10.8885)	3.3162 (3.2140 to 3.4184)	9.9433 (9.0575 to 10.8290)
P value	<.001	<.001	<.001	<.001
Time-fixed effects	Yes	Yes	Yes	Yes
Regional-fixed effects	Yes	Yes	Yes	Yes
N	50,852	48,048	55,405	42,933
Adjusted R^2	0.1425	0.1428	0.0548	0.1358

Substitution of the Dependent Variable

Although life satisfaction and depression symptoms are conceptually distinct, prior studies have shown that life satisfaction is strongly correlated with mental health outcomes and is often treated as an important dimension of psychological well-being [52,53]. To verify the robustness of the results, this study substitutes life satisfaction indicators for the CES-D scale to measure the mental health status of middle-aged and older adults. The range of this indicator is from 1 to 5, with higher values indicating higher life satisfaction. The results from Model 3 in Table 4 show that after replacing the CES-D with life satisfaction, digital infrastructure development remains significantly associated with better psychological outcomes among middle-aged and older adults. While this substitution should be interpreted with caution, the consistency of results across different indicators enhances confidence in the robustness of our findings.

Sample Period Adjustment

The COVID-19 pandemic has had a profound impact on global public health systems and individual health conditions, potentially interfering with the identification of policy effects. To exclude this potential impact, this study excludes data from 2020 and reassesses the effect of digital infrastructure development on the depressive symptoms of middle-aged and older adults. The regression results from Model 4 in Table 4 show that after excluding the 2020 sample, the estimated coefficient of the "Broadband China" pilot policy is –0.4312, and it remains significant at the 1% level, further confirming the positive association between digital infrastructure development and reduced depressive symptoms among middle-aged and older adults.

Mechanism Testing

To explore the mechanisms through which digital infrastructure development affects the depressive symptoms of middle-aged and older adults, this study introduces social network as a mediator, dividing it into 3 dimensions: family connection, close social interactions, and social participation. The regression results in Table 5 show that digital infrastructure development significantly promotes these 3 dimensions of the social network, thereby further reducing depressive symptoms among middle-aged and older adults, confirming the mediating effect.

To test the significance of the mediation effect, this study uses the Bootstrap method. If the CI of the indirect effect does not contain zero, it indicates that the effect is significant. Table 6 shows that the indirect effects of family connection (95% CI –0.0256 to –0.0038), close social interactions (95% CI –0.0162 to –0.0030), and social participation (95% CI –0.0290 to –0.0029) are all significant. These results are consistent with the previous regression analysis, indicating that digital infrastructure development effectively improves psychological well-being by mitigating depressive symptoms through strengthened social networks in various dimensions.

XSL•FO
RenderX

Table 5. Mechanism test.

Variable	Model 1 (family connection)	Model 2 (mental health)	Model 3 (close social interaction)	Model 4 (mental health)	Model 5 (social participation)	Model 6 (mental health)
Did						
β (95% CI)	–0.1004 (–0.1718 to –0.0290)	–0.3644 (–0.5765 to –0.1523)	0.0235 (0.0078 to 0.0392)	–0.3259 (–0.4963 to –0.1554)	0.0209 (0.0052 to 0.0366)	–0.4158 (–0.5981 to –0.2335)
P value	.006	.001	.003	<.001	.009	<.001
Family connection						
β (95% CI)	__[a]	0.1466 (0.1120 to 0.1811)	—	—	—	—
P value	—	<.001	—	—	—	—
Close social interactions						
β (95% CI)	—	—	—	–0.4090 (–0.5179 to –0.3000)	—	—
P value	—	—	—	<.001	—	—
Social participation						
β (95% CI)	—	—	—	—	—	–0.7632 (–0.8843 to –0.6422)
P value	—	—	—	—	—	<.001
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
_cons						
β (95% CI)	7.1214 (6.7989 to 7.4439)	8.6123 (7.6721 to 9.5525)	0.5679 (0.5102 to 0.6255)	9.9346 (9.1209 to 10.7484)	0.2741 (0.2023 to 0.3458)	10.1532 (9.2701 to 11.0363)
P value	<.001	<.001	<.001	<.001	<.001	<.001
Time-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Regional-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	42,364	42,364	56,209	56,209	42,931	42,931
Adjusted R^2	0.1960	0.1390	0.0393	0.1431	0.0779	0.1392

[a]Not available.

Table 6. Mediation effect test based on the bootstrap method.

Mediator variable and effect type	Coefficient	SE	95% CI
Family connections			
Indirect effect	–0.0147	0.0056	–0.0256 to –0.0038
Direct effect	–0.3644	0.1237	–0.6069 to –0.1219
Intimate social interaction			
Indirect effect	–0.0096	0.0034	–0.0162 to –0.0030
Direct effect	–0.3259	0.0880	–0.4984 to –0.1534
Social participation			
Indirect effect	–0.0159	0.0067	–0.0290 to –0.0029
Direct effect	–0.4158	0.1048	–0.6212 to –0.2104

XSL•FO
RenderX

Heterogeneity Analysis

Given the considerable differences in socioeconomic characteristics, health conditions, and levels of digital technology acceptance among middle-aged and older adults, the impact of digital infrastructure development on depressive symptoms is likely to vary across subgroups. Table 7 reports the heterogeneous effects. The results indicate that the policy significantly reduces depressive symptoms among the middle-aged (Model 1) and individuals with lower levels of education (Model 3), with coefficients of –0.4084 (P<.01) and –0.3300 (P<.01), respectively. In contrast, the effects for older adults (Model 2) and those with higher education (Model 4) are weaker and statistically insignificant, with coefficients of –0.2599 (P>.1) and –0.3769 (P>.05), suggesting that the policy's benefits for depressive symptom reduction are not robust in these groups.

Table 7. Heterogeneity analysis.

Variable	Model 1 (age≤65 years)	Model 2 (age>65 years)	Model 3 (low education)	Model 4 (high education)	Model 5 (female)	Model 6 (male)
Did						
β (95% CI)	–0.4084 (–0.6175 to –0.1993)	–0.2599 (–0.6059 to 0.0860)	–0.3300 (–0.5156 to –0.1443)	–0.3769 (–0.7894 to 0.0355)	–0.3794 (–0.6341 to –0.1247)	–0.2931 (–0.5189 to –0.0673)
P value	<.001	.14	<.001	.07	.004	.01
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
_cons						
β (95% CI)	8.6117 (7.5430 to 9.6805)	11.5659 (9.7401 to 13.3917)	9.8586 (8.9596 to 10.7576)	6.6861 (4.8701 to 8.5021)	10.4365 (9.2335 to 11.6395)	7.4745 (6.3568 to 8.5922)
P value	<.001	<.001	<.001	<.001	<.001	<.001
Time-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Regional-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	38,331	17,880	49,216	6995	28,436	27,775
Adjusted R^2	0.1501	0.1355	0.1334	0.0937	0.1304	0.1028

Turning to sex, digital infrastructure development significantly reduces depressive symptoms among both women and men, though the effect is stronger for women (–0.3794, P<.01) than for men (–0.2931, P<.05). Taken together, these findings highlight the differentiated impacts of the "Broadband China" policy, showing that middle-aged adults, less-educated individuals, and women derive relatively greater reductions in depressive symptoms, thereby underscoring the policy's potential to mitigate health inequalities.

Further Analysis

Table 8 presents the impact of digital infrastructure development on health inequality among older adults. Based on the Kakwani relative deprivation index, this study first reversely encodes the depressive symptoms variable, where higher scores indicate better health. Regression analysis is then conducted. Model 1 does not include control variables, while Models 2 and 3 introduce individual and family characteristics, respectively. Model 4 further considers time and regional fixed effects. In Model 4, the coefficient for the core explanatory variable did is –0.0089 (P<.01), indicating that digital infrastructure development significantly reduces inequalities in depressive symptoms among middle-aged and older adults. This finding confirms the role of digital infrastructure development in alleviating health inequality among middle-aged and older adults.

XSL•FO
RenderX

Table 8. The impact of digital infrastructure development on health inequality.

Variable	Model 1	Model 2	Model 3	Model 4
Did				
β (95% CI)	–0.0119 (–0.0154 to –0.0083)	–0.0095 (–0.0130 to –0.0061)	–0.0103 (–0.0138 to –0.0068)	–0.0089 (–0.0140 to –0.0039)
P value	<.001	<.001	<.001	.001
Sex				
β (95% CI)	__[a]	–0.0513 (–0.0543 to –0.0483)	–0.0482 (–0.0512 to –0.0452)	–0.0488 (–0.0531 to –0.0446)
P value	—	<.001	<.001	<.001
Residence				
β (95% CI)	—	0.0240 (0.0209 to 0.0270)	0.0243 (0.0212 to 0.0273)	0.0242 (0.0187 to 0.0297)
P value	—	<.001	<.001	<.001
Age (years)				
β (95% CI)	—	0.0006 (0.0004 to 0.0008)	0.0001 (–0.0001 to 0.0003)	0.0003 (0.0000 to 0.0005)
P value	—	<.001	.15	.03
Education				
β (95% CI)	—	–0.0181 (–0.0196 to –0.0166)	–0.0179 (–0.0194 to –0.0164)	–0.0177 (–0.0198 to –0.0156)
P value	—	<.001	<.001	<.001
Ethnicity				
β (95% CI)	—	–0.0126 (–0.0193 to –0.0059)	–0.0120 (–0.0187 to –0.0054)	–0.0097 (–0.0235 to 0.0041)
P value	—	<.001	<.001	.17
Pension				
β (95% CI)	—	–0.0054 (–0.0088 to –0.0019)	–0.0050 (–0.0085 to –0.0016)	–0.0088 (–0.0125 to –0.0050)
P value	—	.002	<.001	<.001
Health insurance				
β (95% CI)	—	–0.0220 (–0.0292 to –0.0147)	–0.0190 (–0.0262 to –0.0118)	–0.0170 (–0.0258 to –0.0082)
P value	—	<.001	<.001	<.001
Chronic disease				
β (95% CI)	—	0.0590 (0.0557 to 0.0624)	0.0594 (0.0560 to 0.0628)	0.0557 (0.0519 to 0.0594)
P value	—	<.001	<.001	<.001
Marital status				
β (95% CI)	—	—	–0.0396 (–0.0443 to –0.0350)	–0.0397 (–0.0471 to –0.0323)
P value	—	—	<.001	<.001
Household size				
β (95% CI)	—	—	–0.0021 (–0.0030 to –0.0012)	–0.0020 (–0.0032 to –0.0009)
P value	—	—	<.001	<.001
_cons				
β (95% CI)	0.1429 (0.1413 to 0.1446)	0.1432 (0.1279 to 0.1585)	0.2080 (0.1911 to 0.2249)	0.2003 (0.1757 to 0.2248)
P value	<.001	<.001	<.001	<.001
Time-fixed effects	No	No	No	Yes

XSL•FO
RenderX

Variable	Model 1	Model 2	Model 3	Model 4
Regional-fixed effects	No	No	No	Yes
N	56,211	56,211	56,211	56,211
Adjusted R^2	0.0008	0.0750	0.0804	0.0946

[a]Not available.

Discussion

Principal Findings

This study uses panel data from CHARLS 2011-2020 and adopts the DID method, based on the "Broadband China" pilot policy, to assess the impact of digital infrastructure development on the depressive symptoms of middle-aged and older adults, analyze its mechanisms, and explore its effects on health inequality in this group. The findings are as follows:

First, digital infrastructure development has been shown to reduce depressive symptoms of middle-aged and older adults. The findings of this study remain robust even after controlling for selection bias and unobserved factors, providing further evidence of the association effect of digital infrastructure development on the depressive symptoms of older adults. This result is consistent with existing literature, which highlights that digital technologies can effectively alleviate depression [54], a conclusion that applies to various populations, such as children [55], adolescents [56], and older adults [24]. However, the innovation of this study lies in its departure from the traditional focus on individual or household-level perspectives, instead examining the impact of digital infrastructure development on depressive symptoms from a macro policy perspective. This approach reveals the positive role of government-driven policy interventions aimed at expanding digital technology access in enhancing the well-being of the older adult population, further emphasizing the critical role of policy orientation in improving health outcomes.

Second, digital infrastructure development helps to alleviate depressive symptoms among middle-aged and older adults by promoting the improvement of their social networks. This finding is consistent with social network theory, which highlights the critical role of social connections in individual psychological health [57]. Existing research has confirmed that social networks provide emotional support, reduce feelings of loneliness, and enhance social identity [58]. However, while many studies focus on the impact of social networks on psychological health, few explore how digital infrastructure development can promote the formation and enhancement of social networks at a macro level. This study fills this gap by revealing the role of digital infrastructure development in strengthening the social networks of middle-aged and older adults, thereby mitigating depressive symptoms. It further underscores the systematic effect of policy interventions on promoting psychological well-being.

Further analysis shows that improvements in depressive symptoms among middle-aged and older adults are primarily achieved through 3 social network pathways: enhancing family connection, close social interactions, and social participation. The strengthening of these pathways effectively reinforces the social support system, ultimately leading to reduced depressive symptoms. Strengthening family connections is likely due to the reduction in communication costs provided by digital technologies, enabling more frequent and intimate interactions between middle-aged and older adults and their children or relatives [59]. Regarding close social interactions, the proliferation of digital platforms, especially social media, provides a more convenient channel for older adults to stay in touch with friends and expand their social circles, thereby alleviating feelings of loneliness and enhancing social support [60]. Digital technologies have also significantly promoted the social participation of middle-aged and older adults, particularly in community activities, web-based learning, and public affairs, which not only increases their sense of social role but also strengthens their sense of belonging [33]. These findings indicate that digital infrastructure development not only represents an investment in material resources but also has a profound influence on social relationships, highlighting the critical role of digital policies in reducing depressive symptoms among middle-aged and older adults.

Third, the effects of digital infrastructure development on depressive symptoms vary significantly across different subgroups. Specifically, the positive association between digital infrastructure development and improved psychological well-being is more pronounced among women, middle-aged adults, and those with lower education levels. Consistent with existing research, the health improvements in these groups may stem from the promotion of social welfare and individual lifestyle changes facilitated by digital infrastructure development [61,62]. However, this study further reveals the underlying reasons why these groups benefit more.

The significant benefits observed in middle-aged adults suggest that the impact of digital infrastructure development on this group is closely related to the awakening of their health awareness and changes in behavior. Middle-aged adults are typically at a critical transitional stage in life, facing pressures related to health management and the dual burdens of work and family. Digital infrastructure can provide this group with more convenient access to health information and services, which helps mitigate depressive symptoms. Therefore, health interventions targeted at middle-aged adults should focus more on the application of digital tools to enhance health awareness and promote behavior change.

The significant benefits observed among individuals with lower education levels reveal the key role of education in digital technology acceptance and use. Lower-educated groups often face higher health risks, and the simple and user-friendly health tools and educational content provided by digital technologies can effectively reduce cognitive and technological barriers [63].

XSL•FO
RenderX

These groups' ability to access health information and manage their health is enhanced through digital infrastructure. Therefore, digital health interventions targeted at lower-educated groups should focus more on designing health tools that align with their cognitive abilities and technical skills to ensure effective participation and benefit.

At the same time, this study also identifies significant sex differences. Digital infrastructure development has a positive impact on reducing depressive symptoms for both men and women, but the effect is more pronounced among women. This finding may be linked to the multiple responsibilities and pressures that women often bear in both family and social roles [64]. The improvement of digital infrastructure not only provides more opportunities for social interaction but also facilitates access to health resources and psychological support, thereby helping to alleviate women's emotional distress and mental burden to some extent. Therefore, in advancing digital infrastructure development, greater attention should be directed toward addressing the specific needs of women, making use of digital tools to enhance their health and psychological support, and ultimately promoting sex equity in health outcomes.

In contrast, the effects on older adults and those with higher education levels are relatively weaker, which may be closely related to their existing health resources and methods of accessing information. Older adults often face multiple challenges with their health, but they may have higher barriers to adapting to new technologies [65]. Therefore, although digital infrastructure can provide convenience for older adults, issues such as technological adaptability and the learning curve may affect their health benefits. Additionally, urban residents and those with higher education levels are generally more easily able to access medical services and health information, meaning the marginal benefits of digital infrastructure development are lower in these groups.

Fourth, digital infrastructure development has alleviated health inequality in depressive symptoms among middle-aged and older adults to some extent. Previous studies have shown that health inequality is influenced by various factors, including education level [66] and geographic location [67]. However, most existing literature focuses on how traditional socioeconomic factors exacerbate health inequality, with less emphasis on how digital infrastructure can bridge gaps among different social groups by optimizing the allocation of public resources and improving the accessibility of services related to depression. This study finds that digital infrastructure development has had a significant positive effect on disadvantaged groups, demonstrating a certain "compensatory effect." A possible explanation is that traditional health services are often limited by spatial, temporal, and financial constraints, whereas digital infrastructure can provide remote medical care, health education, and information access, significantly reducing the geographical and economic barriers faced by disadvantaged groups and improving their access to health services. Moreover, the ease of use of digital technologies and the convenience of information delivery help reduce the cognitive burden faced by groups with lower health literacy during health decision-making, thereby enhancing their ability to manage depressive symptoms independently. This "compensatory effect" suggests that under the promotion of digital infrastructure development, traditionally disadvantaged groups can better access health services and improve their health outcomes. This not only helps reduce health disparities among groups but also demonstrates the enormous potential of digital technology in promoting health equity.

This study has several strengths. First, it innovatively identifies the associations between digital technology and depressive symptoms. Based on the natural experiment of the "Broadband China" pilot policy, the study uses the DID method to address endogeneity issues and provides external validity evidence of the effects of digital infrastructure development on the depressive symptoms of middle-aged and older adults, filling the gap in existing research regarding the impact of macro digital policy interventions on depression. Second, it explores the social network mechanism in depth. The study reveals how digital infrastructure promotes family connections, close social interactions, and social participation to mitigate depressive symptoms among middle-aged and older adults, enriching the theoretical framework of the interaction between digital infrastructure and social networks. It provides empirical evidence for policymakers to improve the well-being of middle-aged and older adults while promoting social network development. Third, it expands the perspective of health inequality. By measuring disparities in depressive symptoms and conducting heterogeneity analysis from multiple perspectives, the study explores the "compensatory effect" of digital infrastructure development on low-resource groups, providing strong empirical support for healthy aging policies and emphasizing the tremendous potential of digital technology in reducing health disparities.

However, the study has some limitations. First, the data used in this study were primarily collected and assessed through self-reports or proxy reports, which may introduce recall bias and subjectivity, especially concerning depressive symptoms and social network-related issues. Nevertheless, the CHARLS survey follows standardized protocols, ensuring the reliability and comparability of cross-wave data, thus effectively reducing potential biases. Second, due to limitations in data updates, this study was unable to obtain the CHARLS database after 2020. Therefore, future research could extend the time frame and further explore the long-term effects of digital infrastructure on depressive symptoms among middle-aged and older adults. Third, this study did not analyze other potential factors that may affect depression, such as the expansion of long-term care pilot projects and the improvement of health literacy. These factors may influence depressive symptoms to some extent. Future research should incorporate more diversified data sources to further assess the associations between these factors and depression among middle-aged and older adults.

Conclusions

This study, based on panel data from CHARLS 2011-2020, uses a quasi-natural experiment based on the "Broadband China" pilot policy to evaluate the impact of digital infrastructure development on the depressive symptoms of middle-aged and older adults. It also explores the role of social networks and subgroup heterogeneity. The key conclusions are as follows:

XSL•FO
RenderX

First, digital infrastructure development significantly improves the depressive symptoms of middle-aged and older adults, confirming the effectiveness of digital technology in reducing depression. The study analyzes the positive role of government-driven digital infrastructure development from a macro policy perspective, emphasizing the critical role of policy guidance in promoting reductions in depressive symptoms and improving psychological well-being. Second, digital infrastructure development alleviates depressive symptoms by strengthening the social networks of middle-aged and older adults. The study finds that digital infrastructure effectively promotes family connection, close social interactions, and social participation, thereby enhancing social support systems. This indicates that digital infrastructure not only impacts material aspects but also has profound effects at the social relationship level. Third, the impact of digital infrastructure development on depressive symptoms across different subgroups exhibits significant heterogeneity, particularly for women, middle-aged adults, and low-education groups. This suggests that digital infrastructure development helps these groups overcome barriers to health information access and health management, thereby reducing depression levels. Fourth, digital infrastructure development plays an important role in alleviating health inequality. By increasing the accessibility of health services, it significantly reduces depressive symptoms among low-resource groups, thereby improving their psychological well-being, demonstrating the "compensatory effect" of digital infrastructure.

Based on these conclusions, the study proposes the following policy recommendations. First, digital infrastructure should be embedded in long-term national strategies, with coordinated governance and financial investment to reduce regional disparities. Such efforts resonate with the World Health Organization's Global Strategy on Digital Health and highlight the importance of aligning national initiatives with global frameworks. Second, improving digital literacy across the life course is equally crucial. Lifelong, tailored training—especially for women, middle-aged adults, and low-education groups—should combine web-based platforms with community-based programs, drawing on international experiences such as the EU's Active and Healthy Ageing initiatives. Third, strengthening social networks through digital platforms can further alleviate depressive symptoms. Public-private partnerships that encourage family interaction, community participation, and intergenerational support have proven effective in other contexts, such as Japan and Europe, and can be adapted locally. Finally, digital health services should be fully integrated into primary care systems. Expanding reimbursement for telemedicine, embedding digital depression-related services into community health centers, and prioritizing underserved regions, as demonstrated in Australia's telehealth reforms, can promote inclusivity and reduce health inequalities.

Acknowledgments

Language polishing of this manuscript was assisted by ChatGPT. All research-related tasks, including study design, data collection, analysis, and interpretation, were conducted exclusively by the authors. No funding was received for this study.

Data Availability

The data used in this study are available from the China Health and Retirement Longitudinal Study (CHARLS) repository [68]. Researchers may request access following the procedures described on the official website.

Authors' Contributions

ZL designed the study, conducted the analyses, and wrote the manuscript. The author has read and approved the final manuscript and is accountable for all aspects of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Descriptive statistics by survey wave.
[DOCX File , 29 KB - aging_v8i1e79260_app1.docx]

References

1. Partridge L, Deelen J, Slagboom PE. Facing up to the global challenges of ageing. Nature 2018;561(7721):45-56. [doi: 10.1038/s41586-018-0457-8] [Medline: 30185958]
2. Webb LM, Chen CY. The COVID-19 pandemic's impact on older adults' mental health: contributing factors, coping strategies, and opportunities for improvement. Int J Geriatr Psychiatry 2022;37(1) [FREE Full text] [doi: 10.1002/gps.5647] [Medline: 34729802]

XSL•FO
RenderX

3. Abdoli N, Salari N, Darvishi N, Jafarpour S, Solaymani M, Mohammadi M, et al. The global prevalence of major depressive disorder (MDD) among the elderly: a systematic review and meta-analysis. Neurosci Biobehav Rev 2022;132:1067-1073. [doi: 10.1016/j.neubiorev.2021.10.041] [Medline: 34742925]
4. Fan X, Guo X, Ren Z, Li X, He M, Shi H, et al. The prevalence of depressive symptoms and associated factors in middle-aged and elderly Chinese people. J Affect Disord 2021;293:222-228. [doi: 10.1016/j.jad.2021.06.044] [Medline: 34217959]
5. Wang F, Zhang Q, Zhang L, Ng CH, Ungvari GS, Yuan Z, et al. Prevalence of major depressive disorder in older adults in China: a systematic review and meta-analysis. J Affect Disord 2018;241:297-304. [doi: 10.1016/j.jad.2018.07.061] [Medline: 30142588]
6. Proudman D, Greenberg P, Nellesen D. The growing burden of major depressive disorders (MDD): implications for researchers and policy makers. Pharmacoeconomics 2021;39(6):619-625 [FREE Full text] [doi: 10.1007/s40273-021-01040-7] [Medline: 34013439]
7. Arias D, Saxena S, Verguet S. Quantifying the global burden of mental disorders and their economic value. EClinicalMedicine 2022;54:101675 [FREE Full text] [doi: 10.1016/j.eclinm.2022.101675] [Medline: 36193171]
8. Vigo D, Jones L, Atun R, Thornicroft G. The true global disease burden of mental illness: still elusive. Lancet Psychiatry 2022;9(2):98-100. [doi: 10.1016/S2215-0366(22)00002-5] [Medline: 35026138]
9. Banerjee A, Duflo E, Grela E, McKelway M, Schilbach F, Sharma G, et al. Depression and loneliness among the elderly in low- and middle-income countries. J Econ Perspect 2023;37(2):179-202. [doi: 10.1257/jep.37.2.179]
10. Marinucci M, Pancani L, Aureli N, Riva P. Online social connections as surrogates of face-to-face interactions: a longitudinal study under Covid-19 isolation. Comput Hum Behav 2022;128:107102. [doi: 10.1016/j.chb.2021.107102]
11. Li Z. Analysis of the health effects of multiple social networks on the older adult: the substitution role of labor participation. Front Public Health 2024;12:1501597 [FREE Full text] [doi: 10.3389/fpubh.2024.1501597] [Medline: 39687726]
12. Boardman J, Killaspy H, Mezey G. Social Inclusion and Mental Health. Cambridge: Cambridge Univ Press; 2022.
13. Li C, Ning G, Xia Y, Guo K, Liu Q. Does the internet bring people closer together or further apart? the impact of internet usage on interpersonal communications. Behav Sci (Basel) 2022;12(11):425 [FREE Full text] [doi: 10.3390/bs12110425] [Medline: 36354402]
14. Zhang K, Burr J, Mutchler J, Lu J. Internet use and loneliness among urban and non-urban Chinese older adults: the roles of family support, friend support, and social participation. J Gerontol B Psychol Sci Soc Sci 2024;79(7):gbae081. [doi: 10.1093/geronb/gbae081] [Medline: 38742464]
15. Chai Y, Xian G, Wang M, Guo L, Luo S. Aging wisely: the impact of internet use on older adults' mental health. J Affect Disord 2024;364:139-145. [doi: 10.1016/j.jad.2024.08.076] [Medline: 39147146]
16. Litwin H, Levinsky M. Social networks and mental health change in older adults after the Covid-19 outbreak. Aging Ment Health 2022;26(5):925-931. [doi: 10.1080/13607863.2021.1902468] [Medline: 33749448]
17. Hardy BW, Castonguay J. The moderating role of age in the relationship between social media use and mental well-being: an analysis of the 2016 general social survey. Comput Hum Behav 2018;85:282-290. [doi: 10.1016/j.chb.2018.04.005]
18. Li X, He P, Liao H, Liu J, Chen L. Does network infrastructure construction reduce urban–rural income inequality? based on the “Broadband China” policy. Technol Forecast Soc Change 2024;205:123486. [doi: 10.1016/j.techfore.2024.123486]
19. Xu Q, Li X, Dong Y, Guo F. How digital infrastructure development affects residents' health: a quasi-natural experiment based on the “Broadband China” strategy. Cities 2025;157:105611. [doi: 10.1016/j.cities.2024.105611]
20. Zhou Y, Bai Y, Wang J. The impact of internet use on health among older adults in China: a nationally representative study. BMC Public Health 2024;24(1):1065 [FREE Full text] [doi: 10.1186/s12889-024-18269-4] [Medline: 38632588]
21. Ouden WV, van Boekel L, Janssen M, Leenders R, Luijkx K. The impact of social network change and health decline: a qualitative study on experiences of older adults who are ageing in place. BMC Geriatr 2021;21(1):480 [FREE Full text] [doi: 10.1186/s12877-021-02385-6] [Medline: 34481476]
22. Lapshin I. Digital resocialization of elderly people. In: Proceedings of the International Conference on Contemporary Education, Social Sciences and Ecological Studies (CESSES 2018). Hong Kong: Atlantis Press; 2018:990-992.
23. Tsang APL, Lee CK, Chan SCY. A cross-lagged panel analysis of social participation in the relationship between functional limitations and cognitive functioning: evidence from CHARLS. J Appl Gerontol 2025;44(4):551-560. [doi: 10.1177/07334648241282756] [Medline: 39291869]
24. Anderer S. Internet use may boost mental health benefits in older adults. JAMA 2025;333(3):195. [doi: 10.1001/jama.2024.25506] [Medline: 39705027]
25. Yang H, Zhang S, Cheng S, Li Z, Wu Y, Zhang S, et al. A study on the impact of internet use on depression among Chinese older people under the perspective of social participation. BMC Geriatr 2022;22(1):701 [FREE Full text] [doi: 10.1186/s12877-022-03359-y] [Medline: 35999498]
26. Drageset J. Social support. In: Health Promotion in Health Care – Vital Theories and Research. Cham: Springer; 2021:137-144.
27. Hajek A, König HH. Frequency of contact with friends and relatives via internet and psychosocial factors in middle-aged and older adults during the COVID-19 pandemic. findings from the German ageing survey. Int J Geriatr Psychiatry 2022;37(1) [FREE Full text] [doi: 10.1002/gps.5623] [Medline: 34505322]

XSL•FO
RenderX

28. Wrzus C, Hänel M, Wagner J, Neyer FJ. Social network changes and life events across the life span: a meta-analysis. Psychol Bull 2013;139(1):53-80. [doi: 10.1037/a0028601] [Medline: 22642230]
29. Kennedy Terry KM. At the intersection of SLA and sociolinguistics: the predictive power of social networks during study abroad. Mod Lang J 2022;106(1):245-266. [doi: 10.1111/modl.12763]
30. Granovetter MS. The strength of weak ties. Am J Sociol 1973;78(6):1360-1380. [doi: 10.1086/225469]
31. Jaspal R, Breakwell GM. Socio-economic inequalities in social network, loneliness and mental health during the COVID-19 pandemic. Int J Soc Psychiatry 2022;68(1):155-165 [FREE Full text] [doi: 10.1177/0020764020976694] [Medline: 33287610]
32. Wu Z, Penning MJ. Children and the mental health of older adults in China: what matters? Popul Res Policy Rev 2018;38(1):27-52. [doi: 10.1007/s11113-018-9495-z]
33. Sun K, Zhou J. Understanding the impacts of internet use on senior citizens' social participation in China: evidence from longitudinal panel data. Telemat Inform 2021;59:101566. [doi: 10.1016/j.tele.2021.101566]
34. Peng Z, Dan T. Digital dividend or digital divide? digital economy and urban-rural income inequality in China. Telecommun Policy 2023;47(9):102616. [doi: 10.1016/j.telpol.2023.102616]
35. Francis DV, Weller CE. Economic inequality, the digital divide, and remote learning during COVID-19. Rev Black Polit Econ 2022;49(1):41-60 [FREE Full text] [doi: 10.1177/00346446211017797] [Medline: 35291320]
36. van de Werfhorst HG, Kessenich E, Geven S. The digital divide in online education: inequality in digital readiness of students and schools. Comput Educ Open 2022;3:100100 [FREE Full text] [doi: 10.1016/j.caeo.2022.100100] [Medline: 40478103]
37. Cheshmehzangi A, Zou T, Su Z. The digital divide impacts on mental health during the COVID-19 pandemic. Brain Behav Immun 2022;101:211-213 [FREE Full text] [doi: 10.1016/j.bbi.2022.01.009] [Medline: 35041937]
38. Li H. Asymmetric information in the field of healthcare. Adv Econ Manag Political Sci 2024;100(1):28-34. [doi: 10.54254/2754-1169/100/20241089]
39. Sundvall M, Titelman D, DeMarinis V, Borisova L, Çetrez Ö. Safe but isolated - an interview study with Iraqi refugees in Sweden about social networks, social support, and mental health. Int J Soc Psychiatry 2021;67(4):351-359. [doi: 10.1177/0020764020954257] [Medline: 32907462]
40. Chen E, Wood D, Ysseldyk R. Online social networking and mental health among older adults: a scoping review. Can J Aging 2022;41(1):26-39. [doi: 10.1017/S0714980821000040] [Medline: 33736725]
41. Li Z, Zheng L, Xu H. From caregiving burden to income transformation: intergenerational employment effects of long-term care insurance. Front Public Health 2025;13:1606448 [FREE Full text] [doi: 10.3389/fpubh.2025.1606448] [Medline: 40556916]
42. Chu J, Li Y, Wang X, Xu Q, Xu Z. Development of a longitudinal model for disability prediction in older adults in China: analysis of CHARLS data (2015-2020). JMIR Aging 2025;8:e66723 [FREE Full text] [doi: 10.2196/66723] [Medline: 40247464]
43. Peng H, Ling K, Zhang Y. The carbon emission reduction effect of digital infrastructure development: evidence from the broadband China policy. J Clean Prod 2024;434:140060. [doi: 10.1016/j.jclepro.2023.140060]
44. He J, Mu Y, Wang C, Mao Y. Impact of digital infrastructure construction on financial development: evidence from the "broadband China" strategy. Heliyon 2024;10(15):e35262 [FREE Full text] [doi: 10.1016/j.heliyon.2024.e35262] [Medline: 39161820]
45. Jia X, Li X. Has digital infrastructure accelerated enterprises' specialization? evidence from China. Bus Process Manag J 2024;31(4):1560-1578. [doi: 10.1108/bpmj-06-2024-0448]
46. Ma N, Ji X, Shi Y, Wang Q, Wu J, Cui X, et al. Adverse childhood experiences and mental health disorder in China: a nationwide study from CHARLS. J Affect Disord 2024;355:22-30 [FREE Full text] [doi: 10.1016/j.jad.2024.03.110] [Medline: 38548193]
47. Zhang X, Xue M, Zhang Z, Gao Z, Li C, Wu J, et al. Impact of social, familial and personal factors on depressive symptoms in middle-aged and older adults from the national CHARLS cohort. BMC Public Health 2024;24(1):2669 [FREE Full text] [doi: 10.1186/s12889-024-20159-8] [Medline: 39350109]
48. Chen H, Mui AC. Factorial validity of the center for epidemiologic studies depression scale short form in older population in China. Int Psychogeriatr 2014;26(1):49-57 [FREE Full text] [doi: 10.1017/S1041610213001701] [Medline: 24125553]
49. Beck T, Levine R, Levkov A. Big bad banks? the winners and losers from bank deregulation in the United States. J Finance 2010;65(5):1637-1667. [doi: 10.1111/j.1540-6261.2010.01589.x]
50. Kakwani N, Wagstaff A, van Doorslaer E. Socioeconomic inequalities in health: measurement, computation, and statistical inference. J Econom 1997;77(1):87-103. [doi: 10.1016/s0304-4076(96)01807-6]
51. Turguttopbaş N. Analysis of the equity of health care financing system in Turkey. Turk Stud 2020;15(3):1639-1653. [doi: 10.47644/TurkishStudies.42036]
52. Jeong HJ, Kim S, Lee J. Mental health, life satisfaction, supportive parent communication, and help-seeking sources in the wake of COVID-19: first-generation college students vs. non-first-generation college students. J Coll Stud Psychother 2021;37(2):71-86. [doi: 10.1080/87568225.2021.1906189]
53. Headey B, Kelley J, Wearing A. Dimensions of mental health: life satisfaction, positive affect, anxiety and depression. Soc Indic Res 1993;29(1):63-82. [doi: 10.1007/bf01136197]

XSL•FO
RenderX

54. Li J. Digital technologies for mental health improvements in the COVID-19 pandemic: a scoping review. BMC Public Health 2023;23(1):413 [FREE Full text] [doi: 10.1186/s12889-023-15302-w] [Medline: 36859184]
55. Cardoso-Leite P, Buchard A, Tissieres I, Mussack D, Bavelier D. Media use, attention, mental health and academic performance among 8 to 12 year old children. PLoS One 2021;16(11):e0259163 [FREE Full text] [doi: 10.1371/journal.pone.0259163] [Medline: 34788306]
56. Metherell TE, Ghai S, McCormick EM, Ford TJ, Orben A. Digital access constraints predict worse mental health among adolescents during COVID-19. Sci Rep 2022;12(1):19088 [FREE Full text] [doi: 10.1038/s41598-022-23899-y] [Medline: 36352002]
57. Wellman B. The community question: the intimate networks of east yorkers. Am J Sociol 1979;84(5):1201-1231. [doi: 10.1086/226906]
58. Kim H, Kwak S, Youm Y, Chey J. Social network characteristics predict loneliness in older adults. Gerontology 2022;68(3):309-320 [FREE Full text] [doi: 10.1159/000516226] [Medline: 34182553]
59. Lu H, Kandilov IT. Does mobile internet use affect the subjective well-being of older chinese adults? an instrumental variable quantile analysis. J Happiness Stud 2021;22(7):3137-3156. [doi: 10.1007/s10902-021-00365-6]
60. Yu B, Leung YW. Establishing ties or strengthening friendships? students' use of online social networks in intercultural friendship development. Inf Technol People 2023;37(2):815-834. [doi: 10.1108/itp-01-2022-0035]
61. Peláez AL, Suh SM, Zelenev S. Digital Transformation and Social Well-Being: Promoting an Inclusive Society. London: Taylor & Francis; 2022.
62. Hatuka T, Zur H, Mendoza JA. The urban digital lifestyle: an analytical framework for placing digital practices in a spatial context and for developing applicable policy. Cities 2021;111:102978. [doi: 10.1016/j.cities.2020.102978]
63. Weßner A, Mohajerzad H, Fliegener L, Bernhard-Skala C, Rohs M. Consensus in uncertainty. a group Delphi study on the impact of digitalisation on the continuing education of low-qualified adults in Germany. Int J Lifelong Educ 2024;43(4):355-370. [doi: 10.1080/02601370.2024.2356812]
64. Mussida C, Patimo R. Women's family care responsibilities, employment and health: a tale of two countries. J Fam Econ Issues 2021;42(3):489-507 [FREE Full text] [doi: 10.1007/s10834-020-09742-4] [Medline: 33324030]
65. Tomczyk L, Mascia ML, Gierszewski D, Walker C. Barriers to digital inclusion among older people: a intergenerational reflection on the need to develop digital competences for the group with the highest level of digital exclusion. Int J Technol Educ Innov 2023;9(1):5-26. [doi: 10.24310/innoeduca.2023.v9i1.16433]
66. Popham F, Iannelli C. Does comprehensive education reduce health inequalities? SSM Popul Health 2021;15:100834 [FREE Full text] [doi: 10.1016/j.ssmph.2021.100834] [Medline: 34189241]
67. Akter M, Kabir H. Health inequalities in rural and urban Bangladesh: the implications of digital health. Mayo Clin Proc Digit Health 2023;1(2):201-202 [FREE Full text] [doi: 10.1016/j.mcpdig.2023.04.003] [Medline: 40206725]
68. China Health and Retirement Longitudinal Study. URL: https://charls.pku.edu.cn/ [accessed 2025-10-15]

Abbreviations

CES-D: Center for Epidemiologic Studies Depression Scale
CHARLS: China Health and Retirement Longitudinal Study
DID: difference-in-differences
PSM-DID: propensity score matching difference-in-differences
VIF: variance inflation factors

Edited by M Antoniou; submitted 17.06.25; peer-reviewed by CY Chen, TY Liu; comments to author 24.08.25; revised version received 04.09.25; accepted 06.10.25; published 27.10.25.

Please cite as:
Li Z
Digital Pathways to Reducing Depression Among Aging Populations Through the "Broadband China" Pilot Program: Quasi-Natural Experiment
JMIR Aging 2025;8:e79260
URL: https://aging.jmir.org/2025/1/e79260
doi:10.2196/79260
PMID:

©Zhiying Li. Originally published in JMIR Aging (https://aging.jmir.org), 27.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is

XSL•FO
RenderX

properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

The Impact of Digital Technology on the Physical Health of Older Workers: Scoping Review

Jeroen JA Spijker[1,2], PhD; Hande Barlın[3], PhD; Melina Dritsaki[4], PhD; Yang Gu[5], MSc; Aija Klavina[6,7,8], PhD; Nilufer Korkmaz Yaylagul[9], PhD; Gunilla Kulla[10], DrPH; Murat Anil Mercan[11], PhD; Eda Orhun[12], PhD; Anna Sevcikova[13], PhD; Brigid Unim[14], PhD; Gunay Yildizer[15], PhD; Cristina Maria Tofan[16,17], PhD

[1]Institute for Advanced Family Studies, Universitat Internacional de Catalunya, Barcelona, Spain

[2]Centre for Demographic Studies (CED-CERCA), Bellaterra, Spain

[3]Department of Economics, Gebze Technical University, Gebze, Turkey

[4]Department of Economics, University of Western Macedonia, Kastoria, Greece

[5]Department of Work, Employment, Management & Organization, University of Leicester, Leicester, United Kingdom

[6]Laboratory of Sports and Nutrition Research, Riga Stradiņš University, Riga, Latvia

[7]Department of Health Promotion and Rehabilitation, Lithuanian Sports University, Kaunas, Lithuania

[8]Klaipėdos Valstybinė Kolegija / Higher Education Institution, Klaipeda, Lithuania

[9]Department of Gerontology, Faculty of Health Sciences, Akdeniz University, Antalya, Turkey

[10]Department of Health and Caring Sciences, Western Norway University of Applied Sciences, Forde, Norway

[11]Department of Economics, Yıldız Technical University, Istanbul, Turkey

[12]School of Business, Al Akhawayn University, Ifrane, Morocco

[13]Faculty of Social Studies, Masaryk University, Brno, Czech Republic

[14]Department of Cardiovascular, Endocrine-Metabolic-Diseases and Ageing, Italian National Institute of Health, Rome, Italy

[15]Faculty of Sport Sciences, Eskisehir Technical University, Eskişehir, Turkey

[16]Psychology and Educational Sciences Department “Gheorghe Zane”, Institute for Economic and Social Research, Romanian Academy, Iasi, Romania

[17]Department of Sociology, Social Work and Human Resources, Faculty of Philosophy and Social-Political Sciences, Alexandru Ioan Cuza University, Iasi, Romania

Corresponding Author:
Jeroen JA Spijker, PhD
Institute for Advanced Family Studies
Universitat Internacional de Catalunya
Terré 11-19
Barcelona, 08017
Spain
Phone: 34 932541800
Email: jjaspijker@uic.es

Abstract

Background: Digital technologies are increasingly present in workplaces; however, their impact on the physical health of older workers remains unclear.

Objective: This scoping review aims to examine and summarize the scientific evidence on how digital technology affects the physical health of older workers.

Methods: Following PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analysis extension for Scoping Reviews) guidelines, we conducted a scoping review of English-language peer-reviewed studies extracted from MEDLINE, Cochrane, ProQuest, Web of Science, Scopus, APA PsycInfo, and ERIH PLUS. The review followed the population, concept, and context (PCC) framework, including studies on workers aged 50 years or older, any form of digital technology (eg, teleworking and the use of digital tools at work), and its impact on physical health (eg, vision loss and musculoskeletal disorders). Studies that focused only on mental health were excluded. A 13-member research team screened studies in 3 stages, namely title and abstract screening, full-text review, and data extraction. Each study was independently reviewed by at least 2 researchers, and disagreements were resolved through discussion. Data extraction and synthesis were conducted using the web-based systematic review platform Covidence (Veritas Health Innovation Ltd).

XSL•FO
RenderX

Results: In total, 18 studies were selected, published between 2012 and 2024, with most conducted in Europe (n=8) and Asia (n=6), followed by North America (n=2), Oceania (n=1), and Africa (n=1). We identified 6 key physical health areas impacted by digital technology in older workers, including eye health, musculoskeletal health, metabolic and cardiovascular health, workplace sound levels, and user experiences of new technologies. Findings showed mixed effects, with notable negative impacts on eye strain, musculoskeletal disorders, and hearing health issues, but positive effects on weight management, cardiovascular health, physical activity, and perceived physical well-being.

Conclusions: Digital technology presents both risks and benefits for the physical health of older workers. While prolonged screen use and digital work environments contribute to eye strain, musculoskeletal issues, and hearing concerns, other technologies support better weight management, cardiovascular health, and increased physical activity. These findings also underscore the need for workplace intervention to reduce health risks.

International Registered Report Identifier (IRRID): RR2-10.2196/59900

(JMIR Aging 2025;8:e78406) doi:10.2196/78406

KEYWORDS

digital tools; digital technology; physical health; mobility; vision loss; musculoskeletal disorders; older workers; older population; aging; scoping review

Introduction

Physical health can be defined as a "dynamic state, the process of preserving and developing its biological, physiological, and mental functions, optimal work capacity, and social activity with the maximum life expectancy" [1]. During the last decade, research literature presents growing concern not only in the context of individual aging but also in workplace aging, as working life and retirement age continue to extend [2]. For older workers, maintaining physical health involves not only performing job-related tasks effectively but also preventing and managing age-related conditions to support overall well-being. Significant physiological changes associated with aging affect various capabilities, including sensory function, muscle function, cardiovascular and respiratory function, neurological function, and immune response [3]. However, despite the inevitable age-related physiological decline, a significant proportion of older workers demonstrate adaptation abilities that enable them to maintain work performance [4].

From 2004 to 2019, the share of workers aged 55 years and older in the European Union (EU) workforce increased from 12% to 20% [5]. However, 21.6% of these employees reported more than 2 work-related physical health problems [6]. As the global workforce ages and digital technologies increasingly penetrate the occupational landscape, new job roles and work conditions have emerged. The labor market and traditional work processes have undergone significant transformations through practices, such as remote work, blended work, teleworking, and the use of digital applications, including mobile apps. These changes have introduced new physical and psychological demands [7], bringing both advantages and disadvantages. For instance, digital tools offer greater flexibility, enabling employees to work from home and reduce reliance on physically demanding tasks, thereby lowering the risk of musculoskeletal injuries. However, these benefits can only be realized if workstations are ergonomically designed and regular breaks are taken [8].

Physical health has always been a concern for older adults, but with extended working lives and delayed retirement, it is increasingly critical for older workers. A large study examining work-related health risks among older and younger workers found that perceptions of risk differ by age group, often deviating from actual exposure (eg, injury) [6]. Additionally, workplace digitalization presents new physical health challenges for older workers, such as prolonged screen exposure causing eye strain and headaches, extended periods of sitting without physical activity increasing the risk of cardiovascular diseases, and the stress of constant connectivity contributing to spinal and postural issues [8-10].

Moreover, difficulty adapting to technology also affects older workers' health and productivity. Therefore, continuous training, regular biometric and physical health monitoring [11,12], maintaining fitness, managing chronic diseases, and workplace adaptations are key strategies to enhance the safety, health, and productivity of older workers. Likewise, digital health training programs can assist older employees in maintaining health through retirement transition, positively influencing physical well-being [13].

Despite an aging workforce and the digital transformation of work environments, comprehensive research on how digitalization affects the physical health of older workers remains limited. Older workers are often assumed to be more affected than their younger peers [14]. While some recent studies have explored the psychological consequences of workplace digitalization—technostress [15,16], burnout, and psychological strain [11,17,18]—as well as the role of digital technologies in health and disease management [6,12,13,19-24], the physical health consequences for older employees remain largely unexplored. This scoping review aims to bridge that gap by analyzing existing research on this crucial health issue in older workers.

Methods

Guidance Frameworks

To guarantee a systematic and coherent process, this scoping review followed the recommendations of Levac et al [25], which were built on the Arksey and O'Malley [26] framework. The process included the following stages: (1) identifying the

XSL•FO
RenderX

research question, (2) identifying relevant studies, (3) selecting studies, (4) charting and collating the data, and (5) summarizing and reporting the results. The findings are reported according to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analysis extension for Scoping Reviews) guidelines [27,28].

Protocol and Registration

The protocol was published [27] and registered under the International Registered Report Identifier (IRRID): PRR1–10.2196/59900 [28] before conducting the full electronic literature search. A search for existing scoping reviews found no studies on the impact of digitalization on physical health, though several focused on its effects on mental health [27]. No changes were made between the published protocol and this scoping review.

Stage 1: Identifying the Research Question

We posed the main research question, "How do digital technologies in the workplace affect the physical health of older workers?" Given the wide range of ways in which digital technology can affect an older worker's physical health (from vision loss to musculoskeletal disorders) and the diversity of workplace digitalization (from working online from home to the use of robotics in a car assembly line), we formulated the following subquestions:

1. What are the most common physical health issues that older workers face as a result of workplace digitalization?

2. Which industries most strongly affect older workers' physical health as a result of workplace digitalization?

To refine our focus, we applied the population, concepts, and context (PCC) framework [25] (Multimedia Appendix 1 [29,30]). In summary, our review examines studies on older workers aged 50 years and older (Population). Digital technologies (Concept) refer to electronic devices or tools used for data manipulation, storage, and transmission, specifically in work-related activities. The Context focuses on how these technologies impact physical health, excluding nonwork or nondigital applications.

Stage 2: Identifying Relevant Studies

Following the Joanna Briggs Institute Manual for Evidence Synthesis [31] search strategy, we searched for peer-reviewed studies written in English in the following databases: MEDLINE, Cochrane, ProQuest, Web of Science, Scopus, APA PsycInfo, and ERIH PLUS. We used the web-based systematic review platform Covidence (Veritas Health Innovation Ltd) to import the references and remove duplicates. The search terms used are listed in Textbox 1.

Textbox 1. Search terms.

(physical adj (health OR condition* OR issue* OR impairment* OR fitness OR wellbeing OR (well adj being) OR integrit* OR state* OR stress) OR disease* OR vision OR mobility OR obes* OR overweight OR "Body Mass Index") AND (digital OR app* OR web OR internet OR tech* OR (social adj media) OR chat OR online* OR cyber OR virtual OR computerized OR computerised OR electronic OR ICT) AND ((old* or elder* or ageing or ageing or senior*) adj1 (work* OR employee* OR profession* OR labor OR labour OR colleague* OR staff* OR cowork* OR personnel)).*

Stage 3: Study Selection and Eligibility Criteria

A total of 7290 references were imported into Covidence for screening. After removing 461 duplicates, 6638 studies proceeded to title and abstract screening, followed by a full-text review. Through this process, 17 relevant studies addressing the main and secondary research questions were selected [32-48]. Additionally, one more study was added post hoc, as Covidence did not identify it due to the use of "middle-aged" instead of (synonyms for) "older workers" in the title and abstract, despite the sample age range, including individuals aged 35-65 years [9]. Given the lack of a standard definition of "older workers," we included studies with participants aged 50 years and older, even if younger workers were also part of the sample and analyzed (eg, age group >45 years). Similarly, we included studies across diverse occupational settings, ranging from health care professionals and office workers to farmers and self-employed individuals, as well as studies examining a wide variety of digital technologies. For the purpose of this review, we considered both studies that explicitly examined digital tools (eg, personal computers, smartphones, and specialized apps) and those where technology use was implicit in the work arrangement (eg, telework). This broad approach ensured that potentially relevant evidence was not excluded, though it introduced heterogeneity in populations, contexts, and technologies. A detailed overview of the study characteristics, including design, sample, occupational setting, and type of digital technology examined, is provided in Multimedia Appendix 2 for ease of reference.

All coauthors participated in the selection process, with each study reviewed by at least 2 researchers. Conflicts identified by Covidence regarding study inclusion or exclusion were resolved through discussion or by consulting a third reviewer.

Eligibility was determined based on the PCC framework and additional criteria shown in Textbox 2, while the selection process is synthesized in a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) flowchart. Notably, Covidence allows only one exclusion reason per record during the full-text screening phase. However, many excluded studies met multiple exclusion criteria (eg, lacking both relevant physical health outcomes and age-specific data), which is why only the categories of exclusion—not the number of studies per reason—are reported in the flowchart.

No restrictions were placed on publication year or study location, as the absence of a prior scoping review on this specific topic in a peer-reviewed journal justified including all years, while excluding countries would have required further justification.

XSL•FO
RenderX

Textbox 2. Eligibility criteria for the scoping review.

Inclusion criteria

- Population: older workers (participants more than 50 years of age were included in the study)
- Concept: digital technologies or tools related to work
- Context: physical health outcomes
- Setting: nonclinical and in the work sphere
- Study type: original studies with any design or data type (quantitative, qualitative, and mixed)
- Publication status: published in a peer-reviewed journal
- Publication language: English
- Full text available

Exclusion criteria

- Population: younger workers (participants more than 50 years of age were not included in the study)
- Concept: digital technologies or tools not related to work (eg, for health management)
- Context: non-physical health outcomes (eg, mental health)
- Setting: clinical and not in the work sphere
- Study type: other study types (eg, protocols, narrative reviews, or systematic reviews)
- Publication status: published without peer review, dissertations, books, conference papers, letters, and editorials.
- Publication language: written in a language other than English
- Full text not available

Stages 4 and 5: Charting the Data and Collating, Summarizing, and Reporting Results

From each selected publication in the final Covidence phase, the main outcomes of interest were extracted, including bibliographic information, study designs and aims, population characteristics, the type of digital technology or tools analyzed (concept), the type of physical health addressed (context), and the main findings (Multimedia Appendix 2). These findings were analyzed, evaluated, and discussed following the PRISMA-ScR [49] guidelines.

We assessed study quality using the Mixed Methods Appraisal Tool (MMAT) [50], which evaluates five study design categories: (1) qualitative studies, (2) quantitative randomized controlled trials, (3) quantitative nonrandomized studies, (4) quantitative descriptive studies, and (5) mixed methods studies. The MMAT includes a 2-part checklist: an initial screening with 2 questions to confirm the study's empirical nature (clarity of research questions and feasibility of addressing them), followed by 5 design-specific criteria guiding the assessment process.

Three reviewers independently appraised each study to minimize bias and ensure reliability. Discrepancies were resolved through discussion. Each study was assessed using 2 initial screening questions ("Yes"=1, "No"=0) and 5 design-specific criteria ("Yes"=1, "No"=0, "Don't know"=0), evaluating methodology, data collection, relevance of measurements, and coherence between data sources and analysis methods. We subsequently calculated the percentage of "Yes" responses for each study to assess methodological quality, and no studies were excluded based on this assessment. The quality score was calculated by dividing the number of "Yes" responses by the total applicable criteria (2 for screening, 5 for study design) and multiplying by 100. These scores (0%-100%) represent the proportion of MMAT criteria met. We then averaged quality scores across the screening items and each study design category to provide an overview. This quantitative measure contextualizes the reliability and validity of the findings within our review.

Results

Qualitative Assessment

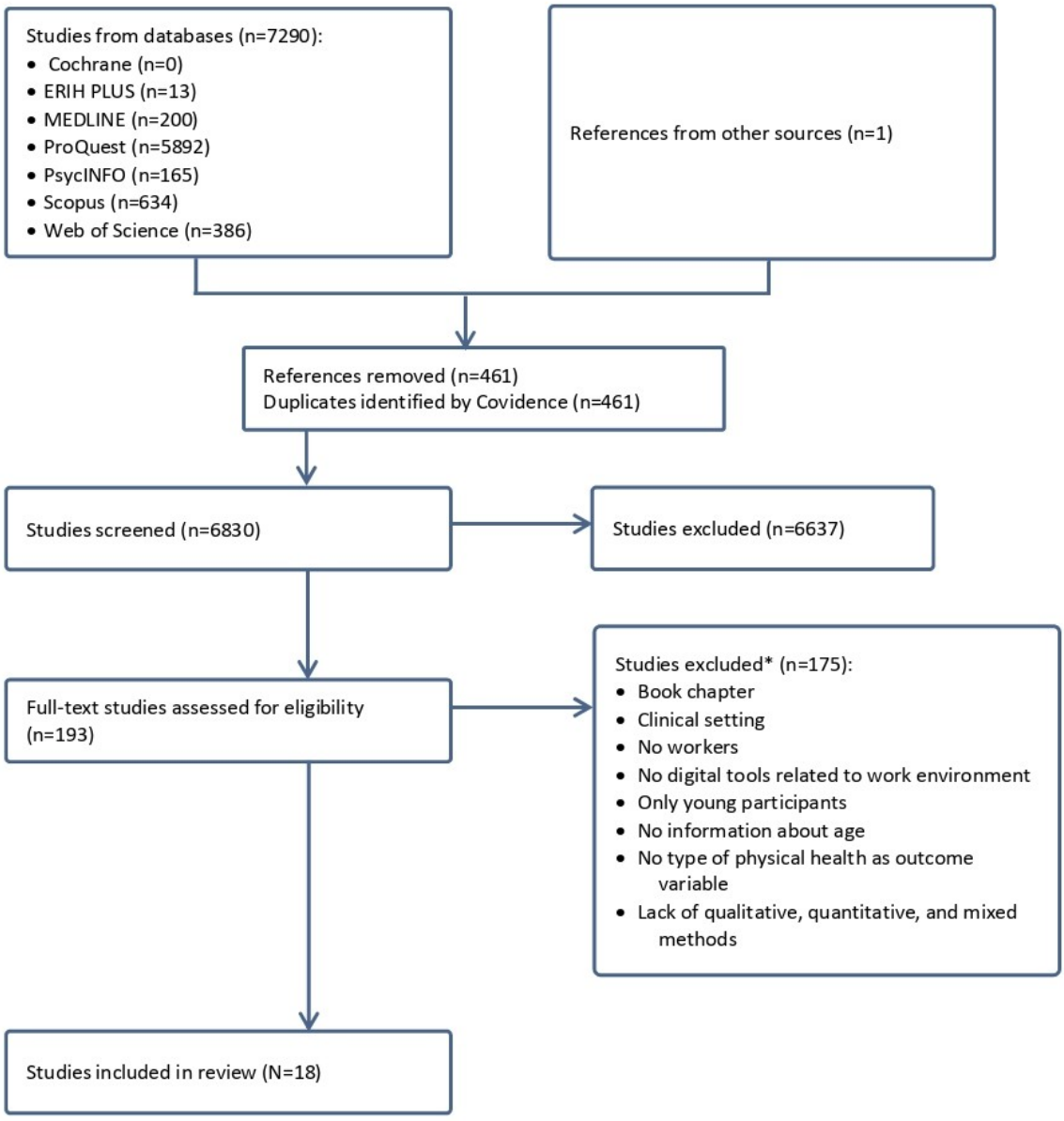
All 18 selected studies (Figure 1) stated clear research questions and collected appropriate data to address them (Multimedia Appendix 3). Although not all studies explicitly described their study design, we identified 9 quantitative nonrandomized studies based on the MMAT user guide [50], with an average quality score of 76% (SD 19%) on the MMAT criteria. These included 1 prospective study [32], 3 cohort studies [33-35], 3 cross-sectional studies [9,36,37], 1 quasi-experimental study [38], and 1 retrospective observational study [39]. Furthermore, we identified 4 quantitative descriptive studies (mean score 65%, SD 30%), including 2 prevalence studies [40,41] and 2 survey-based studies [42,43], 2 quantitative randomized controlled trials (mean score 70%, SD 14%) [44,45], 1 qualitative study (mean score 100%) [46], and 2 mixed method studies (mean score 60%, SD 0%) [47,48].

The lower quality of the quantitative descriptive studies was partly due to the lack of reporting on nonresponses in all studies. Similarly, the lower quality of the mixed methods studies stemmed from a lack of integration between qualitative and quantitative findings. For the 2 quantitative randomized

XSL•FO
RenderX

controlled trials, the outcome assessors were not blinded, or the blinding process was not clearly reported.

Figure 1. PRISMA flowchart of studies through the scoping review process. *No values provided, as Covidence allows only one reason per entry while there were often several reasons to exclude studies.



Characteristics of Included Studies

Overall, 4 [9,38,45,46] of the 18 selected studies were published in 2010, and the other 14 studies were published between 2021 and 2024. Around 12 studies used data collected between 2011 and 2023, including 4 studies [34,35,43,47] that explicitly addressed the effect of the COVID-19 pandemic. A total of 6 studies [37,39,41,44-46] did not state the year of data collection.

Regarding the geographic location, 2 studies [42,48] were conducted in Germany, 2 [9,35] in Japan, and 2 [40,47] in Italy, with the latter also analyzing a sample from the Netherlands. The remaining 12 studies were performed in Australia, Canada, China, Denmark, Israel, Latvia, Lebanon, Libya, Norway, Saudi Arabia, the United Kingdom, and the United States. Europe was the most represented region, with 8 studies, followed by 6 in Asia, 2 in North America, 1 in Oceania, and 1 in Africa.

Sample sizes ranged from 91 to 33,087, although 6 studies [36,39,40,43,45,46] did not report participant numbers. Research covered various economic sectors. Around 4 studies [32,36,41,46] focused on health sector–related occupations, including health care workers, home care service workers, radiation technologists, and radiation therapists. Most other studies examined occupations in sectors outside health care—particularly public institutions—or unspecified fields. These included upper-level nonmanual workers (eg, managers and professionals) [35], office, administrative, white collar, and public sector workers [37-40,45], production line employees [38], and computer users [44]. A Lebanese study examined self-employed individuals, freelancers, and temporary workers

XSL•FO
RenderX

in any sector [43]. Overall, 1 [9] of 2 studies conducted in Japan focused on employees in electronics-related companies performing generally representative (not industry-specific) jobs, including administration, design, research and development, sales, and manufacturing. In total, 1 study [48] focused on farmers, gardeners, and foresters, while another included a mixed sample of health care and non–health care workers, students, and unemployed individuals [36]. Altogether, 4 studies [33,34,42,47] did not explicitly state respondents' occupation or economic sector.

Regarding socioeconomic characteristics, age was analyzed differently across studies. A total of 6 studies [33,34,36,37,40,44] presented results for different age categories, while 2 studies [44,45] treated age as a continuous variable, 1 did both [35], and others studied 2 birth cohorts [42]. The remaining 7 studies [32,38,39,43,46-48] did not assess age differences, although they included employed participants aged 50 years and older. Only 3 studies [42,46,47] exclusively analyzed older workers, while all the others included both younger and older workers. Two studies [32,39] did not report gender, and 1 study [9] included both men and women but only analyzed men due to a low female sample size. The rest incorporated both sexes in their analyses. Regarding job position, only 5 studies [35,41,43,46,48] collected such data.

Types of Digital Technologies Assessed

Work-Related Hardware and Communication Tools

Six studies explicitly investigated common workplace digital technologies. Three studies [9,36,37] studied the intensity of using computers, laptops, tablets, smartphones, mobile phones, emails, and websites. One [9] of those studies examined, besides occupational use, private use also (eg, gaming, email, and web browsing). Another study focused on video display units—covering mobile phones, laptops, computers, tablets, and electronic readers—in relation to computer vision syndrome (CVS) [40]. The other 2 studies [33,42] explored the impact of daily internet use on the physical health of older adults and older workers.

Teleworking

In 2 studies [35,43], digital technologies were not directly specified but were inferred, linking outcomes to remote or hybrid working arrangements where the role of digital technologies was implicit rather than explicitly described. One study [34] combined both perspectives by examining working from home alongside technical organizational support as a mediating factor

Workplace-Specific Technical Tools

Two studies examined technologies unique to specialized occupational settings: the presence and use of personal computers and picture archiving and communication systems among radiation technologists [41] and technical changes and adaptations in work processes, as identified through recommendations by mature radiation therapists in face-to-face interviews [46].

Digital Health and Support Interventions

The remaining 7 studies evaluated digital support programs and apps, namely (1) the Roczen Program, a medically-led, technology-enabled weight management initiative [39]; (2) digital support systems for managing neck pain [32]; (3) guided internet-based interventions, email communication, and electronic couch support (via telephone and internal messaging) for pain management [48]; (4) the AgeWell Digital Coach, a smartphone app with an avatar-based interface, motivational messaging, and activity tracking, designed to promote physical activity, mental well-being, and social engagement during the transition to retirement [47]; (5) an ergonomic learning program delivered digitally to promote healthier workplace practices [38]; (6) a novel "EYE ROLL" device connected to office computers to support vision relaxation [44]; and (7) a self-modeling webcam photos to reduce musculoskeletal risk among office workers using computers [45].

Physical Health Assessment Instruments

Overview

To examine the physical health effects of digital technology use among employed (older) adults, various instruments were used across the selected studies, ranging from self-reported measures to clinical assessments.

Self-Reported Symptoms and Pain

Eight studies used self-reported questionnaires to assess physical symptoms associated with digital technology use. These included gathered self-reported ratings on headaches, back pain, neck pain, carpal tunnel syndrome, and CVS (ocular symptoms) [36] and chronic pain [48]. Neck pain intensity was assessed by asking participants whether they had been "bothered" by pain based on a previously published symptom checklist [32], and musculoskeletal pain or discomfort for 5 body regions was evaluated using response options ranging from never (0) to almost always (4) for frequency and mild (1) to severe (3) for severity [34]. Other studies used different scales, for example, a 0-10 point scale for musculoskeletal pain [38], an ergonomic survey [41], and the Rapid Upper Limb Assessment (RULA) to evaluate an innovative self-modeling photo training method for reducing musculoskeletal risk among office workers using computers [45]. Finally, 1 study [43] used a web-based cross-sectional self-report questionnaire with yes or no responses conducted via Google Forms on teleworking-related health problems among teleworkers to capture physical health outcomes (neck, shoulder, wrist, hand, and back pain), alongside chronic disease indicators (hypertension and diabetes).

Standardized Physical Health Surveys

Two studies [42,47] used the internationally validated SF-12 **(**12-Item Short-Form Health Survey) questionnaire to assess self-rated health status based on the 2 subscales, physical and mental health.

Computer Vision and Ocular Health

A qualitative study using an interpretive phenomenological approach conducted face-to-face interviews with radiation therapists aged more than 50 years to explore the ocular health effects of workplace technology [46]. In contrast, 4 quantitative

XSL•FO
RenderX

studies assessed visual function and symptomatology using a range of clinical and self-reported measures. These include the application of the Computer Vision Syndrome-Q IT scale to capture the frequency and intensity of 16 visual symptoms, as well as tear stability (break-up time), corneal staining presence, and tear quantity using the Schirmer II test [40]. Other studies (1) collected self-reported symptom ratings, such as eye burning, itching, and eye pain, to calculate severity scores [36]; (2) conducted comprehensive vision examinations, including saccadic eye movement recordings and objective accommodation measurements [44]; and (3) measured ocular axial length (AL) of both eyes using an optical AL meter [9].

Clinical Biomarkers and Physiological Measures

One study evaluated the impact of a digitally enabled time-restricted eating intervention—The Roczen Program—by incorporating objective physiological measurements, including body weight (kg), BMI (kg/m²), waist circumference (cm), hemoglobin A_{1c} (HbA_{1c}, mmol/mol), and blood pressure (BP); systolic and diastolic, mm Hg) [39]. Another study used objective and self-reported instruments to assess the impact of the indoor environment on individual well-being and physiological outcomes, including heart rate variability and BP. Physical activity levels were assessed by the ECGMove3's triaxial accelerometer sensor [37].

Chronic Conditions and Cardiovascular Health

Two studies investigated the relationship between work and chronic health outcomes using self-reported data. One study [33] collected self-reported diagnoses of hypertension, arthritis, and asthma confirmed by medical professionals. Another study examined the association between cardiovascular risks and remote work, categorized into different levels of medical treatment and disease progression [35].

These diverse instruments highlight the multifaceted nature of physical health assessments in relation to digital technology use, encompassing subjective experiences, standardized scales, and clinical health indicators.

Physical Health Domains

Overview

The studies on physical health outcomes included in the scoping review can be broadly grouped into 5 physical health domains: eye health, musculoskeletal health, metabolic and cardiovascular health, workplace sound levels, and user experiences of new digital technologies. Several studies span multiple categories. While some focus on the effects of work-related digital technology use on physical health, others evaluate interventions.

Eye Health

Eye health studies reported on visual strain, CVS, and eye discomfort. High rates of CVS symptoms, such as blurred vision, headaches, and burning sensations, were observed, particularly among workers using screens for more than 6 hours per day. Interventions like guided vision relaxation exercises showed promise in reducing symptoms. More specifically, 6 studies contained analysis regarding eye health–related issues. One study [36] reported a 38.6% prevalence of CVS among participants in Libya, identifying higher risk among individuals aged 45 years or older, students, and those using computers for more than 6 hours per day. Besides burning eyes, itching eyes, and increased sensitivity to light, other frequently reported physical health symptoms include headaches, back pain, neck pain, and shoulder pain. A study conducted among Italian office workers reported a higher prevalence of CVS (67.2%), with symptoms, including blurred vision (63.5%), perceived visual deterioration (61.8%), headaches (56.3%), and burning sensations in the eyes (54.2%). Although no significant age-related differences in CVS prevalence were found, women, individuals with intensive digital device use at work (>6 h/d) and those with optical correction were more affected [40]. In a study of male employees in Japanese electronics-related companies, use of work-related digital technology was associated with ocular AL elongation, especially among older workers [9]. Longitudinal data from China similarly suggested that regular internet use was associated with a lower risk of vision impairment [33]. A qualitative study involving medical radiation technologists aged 50 years and older revealed that many experienced deteriorating eyesight that negatively affected their work, especially when using progressive lenses, along with hearing loss [46]. Another study focused on remedies for eye problems connected with computer use by analyzing the effectiveness of an eye roll device designed for guided vision relaxation exercises in open workspaces. The intervention was found to alleviate eye-related complaints, such as eye fatigue and vision discomfort [44]. Although the quality score of included studies was low (mean score 63%, SD 29%), their findings consistently indicate that the duration and intensity of digital device use often have a greater impact on visual health than chronological age itself.

Musculoskeletal Health

Seven studies addressed musculoskeletal disorders, often associated with prolonged sitting or suboptimal workstation ergonomics. Neck, back, and shoulder pain were common, especially among teleworkers and those in health care roles. Ergonomic training programs, photo-based posture correction, and participatory interventions demonstrated efficacy in reducing musculoskeletal symptoms. One study [34] found that older workers who teleworked more than their preferred number of days had a higher likelihood of experiencing musculoskeletal pain. Another study identified back pain as a common drawback of telework among employees in Lebanon [43]. A study in Saudi Arabia focused on radiographic technologists and reported that 27.3% of participants experienced musculoskeletal complaints, with a higher prevalence among female and older workers. Contributing factors included the use of imaging equipment, particularly tasks involving patient transfer, wearing lead aprons, and moving heavy equipment [41]. A large-scale longitudinal study from China found that daily internet usage decreased the risk of hypertension, chronic lung disease, stroke, digestive disease, memory-related disorders, arthritis or rheumatism, asthma, depression, and vision impairment among both working and nonworking adults older than 45 years [33]. Three other studies [32,38,45] analyzed the effects of employers' training on musculoskeletal disorders or symptom risks. According to a Norwegian study using data collected from home care service

XSL•FO
RenderX

workers, the introduction of new technologies in the work environment, accompanied by user training, resulted in lower levels of neck pain reported by participants [32]. In a randomized controlled trial that compared 2 interventions: one based on conventional ergonomic training and workstation adjustment, and the other based on self-modeling photo training accompanied by conventional ergonomic training [45]. While both interventions had beneficial effects in improving work posture and reducing the risk of musculoskeletal disorders—particularly among older workers—only the intervention involving photo training has lasting effects, with stronger outcomes observed among women. In another study that evaluated the effects of a learning program focused on musculoskeletal health, specifically addressing low-strain working techniques, physical training, accompanied by participatory ergonomics, led to increased work productivity among participants [38]. Overall, the findings related to musculoskeletal health are derived from studies with an average quality score of 74% (SD 10%).

Metabolic and Cardiovascular Health

A total of 3 studies [33,35,39] dealt with chronic health conditions. A UK study assessed the impact of a health-related intervention implementing time - restricted eating through the Roczen Program [39]. The authors found that after 1 year, study participants—most of whom were public sector workers—experienced weight loss and reductions in BMI and waist circumference. However, no statistically significant effects were observed for systolic and diastolic BP. A nationwide cohort study in Japan found no significantly increased risks of coronary heart disease among managers and professionals at older ages [35]. The authors concluded that while remote work (teleworking) elevates the risk for younger managers and professionals, it did not increase the risk for older counterparts. Finally, a Chinese study that analyzed 4 waves of the China Health and Retirement Longitudinal Study (CHARLS)—which included both working and nonworking participants aged 45 and older—found that daily internet use was associated with a lower incidence of several chronic conditions, including hypertension, arthritis, asthma, stroke, and vision impairment. However, the authors noted that these outcomes could be influenced by a sample selection effect, as healthier or more advantaged individuals are more likely to use the internet [33]. Overall positive outcomes of the aforementioned studies included reduced BMI and waist circumference through technology-enabled programs. Daily internet use was associated with decreased incidence of several chronic conditions, though some studies found no significant effects on BP or coronary heart disease risk among older professionals. Despite the few identified studies on metabolic and cardiovascular health, their average quality score was relatively high (mean score 87%, SD 12%).

Workplace Environment and Sound

Three studies examined workplace sound levels. One study [37] analyzed the effects of sound levels on public institution workers in the United States, finding that sound levels at or above 50 decibels (A-weighted) increased variability in anxiety levels. Contributing factors to interpersonal differences in the sound–well-being association included age, BMI, high BP, and intensive computer use. Supporting this finding, a Canadian study offered a qualitative perspective, in which a radiation technologist expressed frustration over constant workplace noise and its negative impact on the work environment [46]. In another study addressing workplace health risks, digital technology use did not negatively impact workability or physical health across different occupational sectors in Germany, provided that the use was not perceived as digital work intensification [42]. Overall, elevated noise levels were linked to increased physiological stress, particularly in office settings with dense computer use. Participant narratives also revealed frustration related to environmental noise. The mean quality score of the 3 studies was 93% (SD 12%).

User Experience and Engagement

Finally, 2 studies focused on users' perspectives, barriers, and facilitators rather than direct health outcomes of digital technology use among older workers. One specifically on the design and user experience of a virtual coaching app to support healthy aging among older workers approaching retirement. Findings suggested the need to enhance physical activity and self-efficacy, and increase emotional and informational support during the transition to retirement [47]. The other study highlighted engagement challenges and user perceptions of internet-based interventions (IBIs) designed to prevent depression and manage pain-related disability in green professions, such as farming, gardening, and forestry. Key barriers identified included mental health stigma (especially in rural settings), low prioritization of mental well-being compared to physical health, financial stress, and the physically demanding nature of agricultural work. Only 22.2% of participants completed at least 80% of intervention modules, with farmers facing greater challenges due to the work-related pain and financial strain. However, the autonomy and confidentiality of IBIs appeal to rural users, and peer or family support enhances participation. Thus, IBIs hold potential, addressing engagement barriers crucial for effectiveness [48]. Nevertheless, it should be noted that the results from these 2 intervention studies should be cautiously approached due to their moderate quality score of 60% (SD 0%).

Discussion

Addressing a Research Gap: Physical Health in the Digital Workplace

The increasing integration of digital technologies into daily life and work environments has raised the need for researchers to examine their impact on the physical health of older workers. While previous scoping and systematic reviews have predominantly focused on psychosocial outcomes, such as technostress and burnout [27], evidence on physical health consequences has remained fragmented. This scoping review addresses that gap by mapping the physical health impacts of digital technologies among older workers across diverse occupational sectors. Our synthesis shows that digitalization has multifaceted effects on eye health, musculoskeletal health, metabolic and cardiovascular health, and workability, with risks and benefits often varying by age, context, and work design.

XSL•FO
RenderX

The integration of digital technologies into the workplace also presents both opportunities and risks. On the one hand, digital tools can automate physically demanding tasks, optimizing work processes and thereby reducing physical strain for older workers [51,52]. However, despite the frequent use of technologies like smartphones, laptops, and tablets, comprehensive studies directly linking digital work environments to physical health outcomes among older workers remain scarce. This highlights critical gaps in the evidence base—including the need for longitudinal designs, sector-specific analyses, and user-centered interventions—that should guide future research. In this way, the review contributes to an expanding body of evidence aiming to ensure that the digital transformation of work supports rather than undermines healthy aging.

Thematic Insights Into Physical Health Domains

This review highlights several key domains in which digital technologies affect physical health outcomes:

1. Eye health: Multiple studies reported a heightened prevalence of CVS in older workers exposed to prolonged screen use. Risk factors included longer exposure duration, older age, female gender, and pre-existing optical correction needs. Importantly, interventions, such as guided vision relaxation exercises, demonstrated efficacy, indicating that workplace-based preventive strategies can mitigate visual strain.
2. Musculoskeletal health: Increased technology use linked to remote and hybrid work arrangements was consistently associated with musculoskeletal complaints (neck, back, and shoulder pain). Evidence supports ergonomic interventions, such as participatory training and workstation modification, as effective countermeasures. These findings confirm earlier work on sedentary behavior but extend it by situating musculoskeletal risks specifically within digitalized work environments.
3. Metabolic and cardiovascular health: Evidence in this domain was more mixed. While teleworking was associated with elevated coronary heart disease risk among younger professionals, no such association was observed in older workers. Conversely, daily internet use appeared protective, reducing the likelihood of chronic disease among adults aged 45 years and older. These findings suggest that digital engagement may have differential effects across age groups, pointing to the need for more nuanced, longitudinal research.
4. Workability and career transitions: digital interventions, such as virtual coaching apps and internet-based support systems, demonstrated promise in enhancing physical activity, self-efficacy, and emotional well-being during key career transitions, including preparation for retirement. When tailored to the needs of older workers, these tools were generally well received. However, barriers to engagement, such as low digital literacy, rural isolation, and physical fatigue, particularly in physically demanding sectors, were frequently noted. These challenges highlight the importance of cultural tailoring, digital literacy support, and user-centered design in the development of effective workplace digital interventions.

Together, these thematic insights suggest that digitalization in later working life is neither uniformly beneficial nor uniformly harmful but rather context-dependent, shaped by technology type, work design, and worker characteristics.

Designing Digital Interventions for Healthy Aging at Work

Based on our findings, we emphasize the critical role of employers in mitigating the physical health risks associated with digital work environments. Organizations should proactively implement measures to safeguard the physical well-being of older employees in digitalized work settings. Potential strategies include the following:

1. Adaptive ergonomic workstations with adjustable desks, chairs, and screen positions to meet the health needs of aging workers, specifically related to musculoskeletal health and visual functions.
2. Wearable health devices that monitor posture, track movement, and screen time to provide real-time feedback and encourage regular breaks and stretching exercises.
3. Hiring of occupational health specialists to assess individual health risks and customize intervention strategies for older workers.
4. Employers to support physical activity either by introducing active-break routines at work or by arranging external events like weekend activities and retreats, which focus on physical exercises to benefit musculoskeletal health.
5. Workplace wellness programs, such as (virtual) physiotherapy, guided stretching sessions, and eye health workshops, can help to prevent musculoskeletal and digital eye strain.

Employers who integrate technology with human-centered design principles can create digital workspaces that promote both the health and inclusion of older employees, fostering a sustainable and age-inclusive workforce.

Limitations

A potential limitation of this scoping review is that, despite an extensive and systematic search across 7 major academic databases, some relevant studies may have been missed. Non-English language publications were excluded due to resource constraints, which may have introduced language bias. Given the rapid pace of digital innovation, newer studies may also have been published after our search was completed. Our analysis reflects only the studies retrieved through our search strategy, with most conducted in Europe and Asia. Therefore, the findings are more applicable to these regions, and future studies from underrepresented parts of the world are needed for a more global understanding. Moreover, cultural and regional workplace differences may influence the applicability of findings. Occupational health frameworks, ergonomic standards, digital literacy, and cultural attitudes toward aging and technology vary across regions, meaning that experiences of digital work in, for example, Norway or Japan may not be directly comparable to contexts with fewer resources or less employer support. These contextual differences highlight the need for future studies to examine workplace cultures and policy environments when assessing the transferability of interventions.

XSL•FO
RenderX

Another limitation concerns the varying definitions of digital technology. Some studies examined explicit tools (eg, personal computers, mobile devices, and specialized apps), while others focused on implicit contexts (eg, telework) where technology use was assumed but not specified. By including both types, the breadth of digitalization in the workplace was captured, but this variability may limit the comparability of findings across studies.

Regarding study design, most intervention-based studies (including 8 of the 9 quantitative nonrandomized, both quantitative randomized controlled trials, and both mixed method studies) were limited in exposure duration, ranging from just 6 weeks to 2.5 years, while the remaining studies were primarily cross-sectional or descriptive.

Our review also highlights broader methodological limitations in the current literature. Many studies rely on cross-sectional designs, restricting the ability to establish causal relationships between technology use and physical health outcomes. A few exceptions point to promising directions; for instance, 1 study [37] used repeated measures and found that both low and high sound levels negatively affected physiological well-being. Such findings suggest that workplace designs addressing sensory environments—for example, providing quiet workspaces for older employees—may represent an important avenue for improving health.

Future studies should therefore emphasize longer-term, longitudinal, and workplace-based intervention designs to better capture the sustained impact of digital technology use on the physical health of older workers.

While our review identifies several ergonomic and digital health programs that may benefit older workers' physical health [eg, 37,43,44]—demonstrating both the feasibility and potential benefits of workplace-based strategies—the strength of evidence is limited by small samples, heterogeneous designs, and short-term effects. As with all scoping reviews, our goal was to map the existing literature rather than formally evaluate intervention effectiveness. To strengthen our approach, we followed the PRISMA-ScR checklist (Multimedia Appendix 4) and critically appraised all included studies using the MMAT, which allowed us to assess the methodological quality of both quantitative and qualitative evidence and provide a more informed synthesis (Multimedia Appendix 3). Looking ahead, systematic reviews or meta-analyses should apply formal evidence grading frameworks to more rigorously assess the robustness of specific interventions and inform workplace policy.

Conclusions

To the best of our knowledge, this scoping review is the first to provide a comprehensive synthesis of the physical health effects associated with digital technologies in the workplace, offering valuable insights into how digitalization influences the well-being of older workers. The findings highlight a diverse range of physical health implications, spanning eye health, musculoskeletal conditions, metabolic and cardiovascular health, and workplace ergonomics. These results underscore the pressing need for targeted workplace interventions and employer-driven initiatives to ensure a healthier work environment for aging employees.

Although intervention studies remain scarce, existing examples—such as ergonomic learning programs, vision relaxation exercises, and posture improvement interventions—suggest promising avenues. Future research should build on these findings by developing and evaluating evidence-based interventional programs tailored to older workers and diverse workplace contexts, while also conducting longitudinal studies to examine causal relationships between digital technology use and physical health outcomes in greater detail.

Additionally, the design and implementation of digital technologies should adopt an inclusive and adaptive approach, acknowledging age-related changes in sensory, cognitive, and physical functions. Participatory design models—in which older workers actively contribute to the development of digital workplace solutions—will be critical for ensuring usability, acceptance, and sustained engagement.

In conclusion, this review reinforces the necessity of adopting a holistic and integrated approach to safeguarding the physical health of older workers in an increasingly digitalized labor market. By addressing the physical, ergonomic, and psychosocial challenges of modern work environments and leveraging adaptive technologies, it is possible to support healthy aging at work, extend career longevity, and maintain the well-being and productivity of older employees in a sustainable manner.

Acknowledgments

The research was partially financed through the COST Action CA21107, "Work inequalities in later life redefined by digitalization (DIGI-net)," supported by the European Cooperation in Science and Technology (COST), and the project "Research of Excellence on Digital Technologies and Wellbeing" (CZ.02.01.01/00/22_008/0004583"), co-financed by the European Union under the Czech Excellent research grant.

Authors' Contributions

JJAS contributed to conceptualization. JJAS and CMT devloped the methodology. JJAS, HB, MD, YG, AK, NKY, GK, MAM, EO, AS, BU, GY, and CMT conducted the investigation and validation. JJAS performed visualization and writing–original draft preparation. JJAS, HB, MD, YG, AK, NKY, GK, MAM, EO, AS, BU, GY, and CMT handled writing–review and editing.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Definitions of the inclusion criteria regarding population, concept, and context in the scoping review.
[PDF File (Adobe PDF File), 153 KB - aging_v8i1e78406_app1.pdf]

Multimedia Appendix 2
Description of studies selected (N = 18).
[PDF File (Adobe PDF File), 185 KB - aging_v8i1e78406_app2.pdf]

Multimedia Appendix 3
Qualitative assessment (Mixed Methods Appraisal Tool) of selected studies (N = 18).
[PDF File (Adobe PDF File), 223 KB - aging_v8i1e78406_app3.pdf]

Multimedia Appendix 4
PRISMA-ScR checklist.
[PDF File (Adobe PDF File), 236 KB - aging_v8i1e78406_app4.pdf]

References

1. Koipysheva E, Lebedinsky V, Koipysheva M. Physical health (definition, semantic content, study prospects). In: Ardashkin IB, Iosifovich BV, Martyushev NV, editors. Research Paradigms Transformation in Social Sciences, vol 50. European Proceedings of Social and Behavioural Sciences. Irkutsk, Russia: Future Academy; 2018:601-605.
2. Poscia A, Moscato U, La Milia DI, Milovanovic S, Stojanovic J, Borghini A, et al. Workplace health promotion for older workers: a systematic literature review. BMC Health Serv Res 2016;16(Suppl 5):329 [FREE Full text] [doi: 10.1186/s12913-016-1518-z] [Medline: 27609070]
3. Maertens JA, Putter SE, Chen PY, Diehl M, Huang YH. Physical capabilities and occupational health of older workers. In: The Oxford handbook of work and aging. New York, NY: Oxford University Press; 2012:215-235.
4. Phillip JM, Aifuwa I, Walston J, Wirtz D. The mechanobiology of aging. Annu Rev Biomed Eng 2015;17:113-141 [FREE Full text] [doi: 10.1146/annurev-bioeng-071114-040829] [Medline: 26643020]
5. Eurostat. Ageing Europe: Looking at the lives of older people in the EU. Luxembourg: Publications Office of the European Union; 2020.
6. Jones MK, Latreille PL, Sloane PJ, Staneva AV. Work-related health risks in Europe: are older workers more vulnerable? Soc Sci Med 2013;88:18-29. [doi: 10.1016/j.socscimed.2013.03.027] [Medline: 23702206]
7. Cijan A, Jenič L, Lamovšek A, Stemberger J. How digitalization changes the workplace. Dyn Relatsh Manag J 2019;8(1):3-21.
8. Milaković M, Koren H, Bradvica-Kelava K, Bubaš M, Nakić J, Jeličić P, et al. Telework-related risk factors for musculoskeletal disorders. Front Public Health 2023;11:1155745 [FREE Full text] [doi: 10.3389/fpubh.2023.1155745] [Medline: 37465175]
9. Honda T, Nakagawa T, Watanabe Y, Hayashi T, Nakano T, Horie S, et al. Association between information and communication technology use and ocular axial length elongation among middle-aged male workers. Sci Rep 2019;9(1):17489 [FREE Full text] [doi: 10.1038/s41598-019-53423-8] [Medline: 31767931]
10. Dunstan DW, Howard B, Healy GN, Owen N. Too much sitting--a health hazard. Diabetes Res Clin Pract 2012;97(3):368-376. [doi: 10.1016/j.diabres.2012.05.020] [Medline: 22682948]
11. Arvola R, Tint P, Kristjuhan, Siirak V. Impact of telework on the perceived work environment of older workers. Sci Ann Econ Bus 2017;64(2):199-214 [FREE Full text] [doi: 10.1515/saeb-2017-0013]
12. Ranasinghe U, Tang LM, Harris C, Li W, Montayre J, de Almeida Neto A, et al. A systematic review on workplace health and safety of ageing construction workers. Saf Sci 2023;167:106276. [doi: 10.1016/j.ssci.2023.106276]
13. Stara V, Santini S, Kropf J, D'Amen B. Digital health coaching programs among older employees in transition to retirement: systematic literature review. J Med Internet Res 2020;22(9):e17809 [FREE Full text] [doi: 10.2196/17809] [Medline: 32969827]
14. Komp-Leukkunen K, Poli A, Hellevik T, Herlofson K, Heuer A, Norum R, et al. Older workers in digitalizing workplaces: a systematic literature review. J Aging Soc Change 2022;12(2):37-59.
15. Marchiori DM, Mainardes EW, Rodrigues RG. Do individual characteristics influence the types of technostress reported by workers? Int J Hum-Comput Interact 2018;35(3):218-230. [doi: 10.1080/10447318.2018.1449713]
16. Tams S, Grover V, Thatcher J, Ahuja M. Grappling with modern technology: interruptions mediated by mobile devices impact older workers disproportionately. Inf Syst E-Bus Manage 2021;20(4):635-655. [doi: 10.1007/s10257-021-00526-3]

XSL•FO
RenderX

17. Venz L, Wöhrmann AM. Always on Call: is there an age advantage in dealing with availability and response expectations? Work Aging Retire 2023;9(4):342-357 [FREE Full text] [doi: 10.1093/workar/waac034]
18. Mauno S, Minkkinen J, Tsupari H, Huhtala M, Feldt T. Do older employees suffer more from work intensification and other intensified job demands? Evidence from upper white-collar workers. Scand J Work Organ Psychol 2019;4(1):3.
19. Matthew-Maich N, Harris L, Ploeg J, Markle-Reid M, Valaitis R, Ibrahim S, et al. Designing, implementing, and evaluating mobile health technologies for managing chronic conditions in older adults: a scoping review. JMIR Mhealth Uhealth 2016;4(2):e29 [FREE Full text] [doi: 10.2196/mhealth.5127] [Medline: 27282195]
20. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021;21(1):1556 [FREE Full text] [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
21. Kim BY, Lee J. Smart devices for older adults managing chronic disease: a scoping review. JMIR Mhealth Uhealth 2017;5(5):e69 [FREE Full text] [doi: 10.2196/mhealth.7141] [Medline: 28536089]
22. Chiu C, Hu J, Lo Y, Chang E. Health promotion and disease prevention interventions for the elderly: a scoping review from 2015-2019. Int J Environ Res Public Health 2020;17(15):5335 [FREE Full text] [doi: 10.3390/ijerph17155335] [Medline: 32722162]
23. Zaman SB, Khan RK, Evans RG, Thrift AG, Maddison R, Islam SMS. Exploring barriers to and enablers of the adoption of information and communication technology for the care of older adults with chronic diseases: scoping review. JMIR Aging 2022;5(1):e25251 [FREE Full text] [doi: 10.2196/25251] [Medline: 34994695]
24. De Santis KK, Mergenthal L, Christianson L, Zeeb H. Digital technologies for health promotion and disease prevention in older people: protocol for a scoping review. JMIR Res Protoc 2022;11(7):e37729 [FREE Full text] [doi: 10.2196/37729] [Medline: 35862187]
25. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010;5:69 [FREE Full text] [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
26. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005;8(1):19-32 [FREE Full text] [doi: 10.1080/1364557032000119616]
27. Spijker JJA, Barlın H, Grad DA, Gu Y, Klavina A, Korkmaz Yaylagul N, et al. The impact of digital technology on the physical health of older workers: protocol for a scoping review. JMIR Res Protoc 2024;13:e59900 [FREE Full text] [doi: 10.2196/59900] [Medline: 39325529]
28. Spijker JJA, Barlın H, Grad DA, Gu Y, Klavina A, Korkmaz Yaylagul N, et al. The impact of digitalization on the physical health of older workers: scoping review protocol summary. OSF Registries. 2024 Sep 19. URL: https://doi.org/10.17605/OSF.IO/DJ34A
29. McCarthy J, Heraty N, Cross C, Cleveland J. Who is considered an 'older worker'? Extending our conceptualisation of 'older' from an organisational decision maker perspective. Human Res Mgmt Journal 2014 May 27;24(4):374-393. [doi: 10.1111/1748-8583.12041]
30. LaMeres BJ. Introduction to Logic Circuits & Logic Design with VHDL. 2nd edition. Cham: Springer; 2019.
31. Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z. JBI Manual for Evidence Synthesis. JBI. URL: https://doi.org/10.46658/JBIMES-24-01 [accessed 2025-10-23]
32. Christensen J, Johannessen H. Is new tech a pain in the neck? The impact of introducing new technologies in home-care on neck pain: a prospective study. BMC Public Health 2024;24(1):734 [FREE Full text] [doi: 10.1186/s12889-024-18252-z] [Medline: 38454440]
33. Li P, Zhang C, Gao S, Zhang Y, Liang X, Wang C, et al. Association between daily internet use and incidence of chronic diseases among older adults: prospective cohort study. J Med Internet Res 2023;25:e46298 [FREE Full text] [doi: 10.2196/46298] [Medline: 37459155]
34. Oakman J, Lambert KA, Weale VP, Stuckey R, Graham M. The effect of preference and actual days spent working from home on stress and musculoskeletal pain in older workers. Int Arch Occup Environ Health 2023;96(8):1113-1121 [FREE Full text] [doi: 10.1007/s00420-023-01992-7] [Medline: 37464225]
35. Zaitsu M, Ishimaru T, Tsushima S, Muramatsu K, Ando H, Nagata T, et al. Incidence of coronary heart disease among remote workers: a nationwide web-based cohort study. Sci Rep 2024;14(1):8415 [FREE Full text] [doi: 10.1038/s41598-024-59000-y] [Medline: 38600223]
36. Alturaiki H, Alnajjar J, Alibrahim I, Almuhaysin F, El Gaddafi MW, Almarzoq M, et al. Computer vision syndrome among the general population in the eastern region of libya: prevalence and risk factors. Cureus 2023;15(11):e48777 [FREE Full text] [doi: 10.7759/cureus.48777] [Medline: 38098910]
37. Srinivasan K, Currim F, Lindberg CM, Razjouyan J, Gilligan B, Lee H, et al. Discovery of associative patterns between workplace sound level and physiological wellbeing using wearable devices and empirical Bayes modeling. NPJ Digit Med 2023;6(1):5 [FREE Full text] [doi: 10.1038/s41746-022-00727-1] [Medline: 36639725]
38. Sell L, Lund H, Holtermann A, Søgaard K. The effect on work ability of a tailored ergonomic learning program. Work 2015;53(2):357-366. [doi: 10.3233/WOR-152228] [Medline: 26835856]
39. Brown A, Dobbie LJ, Falvey L, Patel DC, Kwan JTC, Steinmo S, et al. Real-world data of a digitally enabled, time-restricted eating weight management program in public sector workers living with overweight and obesity in the United Kingdom:

XSL•FO
RenderX

a service evaluation of the Roczen program. Obes Sci Pract 2024;10(1):e730 [FREE Full text] [doi: 10.1002/osp4.730] [Medline: 38344678]
40. Cantó-Sancho N, Porru S, Casati S, Ronda E, Seguí-Crespo M, Carta A. Prevalence and risk factors of computer vision syndrome-assessed in office workers by a validated questionnaire. PeerJ 2023;11:e14937 [FREE Full text] [doi: 10.7717/peerj.14937] [Medline: 36890870]
41. Shubayr N, Alashban Y. Musculoskeletal symptoms among radiation technologists in Saudi Arabia: prevalence and causative factors. Acta Radiol 2022;63(4):497-503. [doi: 10.1177/02841851211004428] [Medline: 33779321]
42. Borle P, Boerner-Zobel F, Voelter-Mahlknecht S, Hasselhorn HM, Ebener M. The social and health implications of digital work intensification. Associations between exposure to information and communication technologies, health and work ability in different socio-economic strata. Int Arch Occup Environ Health 2021;94(3):377-390 [FREE Full text] [doi: 10.1007/s00420-020-01588-5] [Medline: 33084928]
43. Haddad C, Zakhour M, Sacre H, Karam J, El Khatib S, Akel M, et al. Perceived advantages and drawbacks of teleworking during the economic crisis and COVID-19: a cross-sectional study among university degree holders. Arch Environ Occup Health 2024;79(2):91-105. [doi: 10.1080/19338244.2024.2366815] [Medline: 38912891]
44. Svede A, Semjonova S, Ganebnaya A, Puhova L, Baig KF, Kucika A, et al. Application of a new device for vision relaxation in computer users. Vision (Basel) 2024;8(3):40 [FREE Full text] [doi: 10.3390/vision8030040] [Medline: 39051226]
45. Taieb-Maimon M, Cwikel J, Shapira B, Orenstein I. The effectiveness of a training method using self-modeling webcam photos for reducing musculoskeletal risk among office workers using computers. Appl Ergon 2012;43(2):376-385. [doi: 10.1016/j.apergo.2011.05.015] [Medline: 21745654]
46. Wilson J. Old hands, new terrain: examining the concerns and challenges of the mature radiation therapist in British Columbia. J Med Imaging Radiat Sci 2014;45(3):283-290. [doi: 10.1016/j.jmir.2014.05.002] [Medline: 31051980]
47. Santini S, Fabbietti P, Galassi F, Merizzi A, Kropf J, Hungerländer N, et al. The impact of digital coaching intervention for improving healthy ageing dimensions among older adults during their transition from work to retirement. Int J Environ Res Public Health 2023;20(5):4034 [FREE Full text] [doi: 10.3390/ijerph20054034] [Medline: 36901045]
48. Braun L, Freund J, Thielecke J, Baumeister H, Ebert DD, Titzler I. Barriers to and facilitators of engaging with and adhering to guided internet-based interventions for depression prevention and reduction of pain-related disability in green professions: mixed methods study. JMIR Ment Health 2022;9(11):e39122 [FREE Full text] [doi: 10.2196/39122] [Medline: 36350684]
49. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
50. Hong QN, Fàbregues S, Bartlett G, Boardman F, Cargo M, Dagenais P, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. Educ Inf 2018;34(4):285-291. [doi: 10.3233/EFI-180221]
51. Hoonakker P. Information and communication technology and quality of working life: backgrounds, facts, and figures. In: Korunka C, Hoonakker P, editors. The impact of ICT on quality of working life. Dordrecht, Netherlands: Springer; 2014:9-23.
52. Hudomiet P, Willis RJ. Computerization, obsolescence and the length of working life. Labour Econ 2022;77:102005 [FREE Full text] [doi: 10.1016/j.labeco.2021.102005] [Medline: 37035254]

Abbreviations

AL: axial length
BP: blood pressure
CHARLS: China Health and Retirement Longitudinal Study
EU: European Union
HbA1c: hemoglobin A1c
IBIs: internet-based interventions
CVS: computer vision syndrome
IRRID: International Registered Report Identifier
MMAT: Mixed Method Appraisal Tool
PCC: population, concepts, and context
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
RULA: Rapid Upper Limb Assessment
SF-12: 12-Item Short-Form Health Survey

XSL•FO
RenderX

Edited by H Köttl; submitted 03.06.25; peer-reviewed by Y Zhang, M Beltrão, M Okonkwo, MS Mekbib, J Oghuan; comments to author 07.07.25; revised version received 23.09.25; accepted 06.10.25; published 18.11.25.

Please cite as:
Spijker JJA, Barlın H, Dritsaki M, Gu Y, Klavina A, Korkmaz Yaylagul N, Kulla G, Mercan MA, Orhun E, Sevcikova A, Unim B, Yildizer G, Tofan CM
The Impact of Digital Technology on the Physical Health of Older Workers: Scoping Review
JMIR Aging 2025;8:e78406
URL: https://aging.jmir.org/2025/1/e78406
doi:10.2196/78406
PMID:

©Jeroen JA Spijker, Hande Barlın, Melina Dritsaki, Yang Gu, Aija Klavina, Nilufer Korkmaz Yaylagul, Gunilla Kulla, Murat Anil Mercan, Eda Orhun, Anna Sevcikova, Brigid Unim, Gunay Yildizer, Cristina Maria Tofan. Originally published in JMIR Aging (https://aging.jmir.org), 18.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Association of Insulin Resistance, Sarcopenia, and Risk of Cardiovascular Disease: Findings From the China Health and Retirement Longitudinal Study

Huilin Liu[1], MS; Sijing Li[1], MD; Ximin Zhang[1], BS; Wenjie Long[1], MD; Huili Liao[1], MD; Lu Lu[2], MD; Shihao Ni[2], MD; Zhongqi Yang[1], MD

[1]Department of Geriatrics, First Affiliated Hospital of Guangzhou University of Chinese Medicine, No. 12 Airport Road, Baiyun District, Guangzhou, Guangdong Province, China

[2]Lingnan Medical Research Center, Guangzhou University of Chinese Medicine, Guangzhou, Guangdong Province, China

Corresponding Author:
Zhongqi Yang, MD
Department of Geriatrics, First Affiliated Hospital of Guangzhou University of Chinese Medicine, No. 12 Airport Road, Baiyun District, Guangzhou, Guangdong Province, China

Abstract

Background: Cardiovascular disease (CVD) is the main cause of death in middle-aged and older people in China. The interplay between sarcopenia and insulin resistance (IR) in driving CVD risk has not been fully understood, particularly regarding sarcopenia severity and IR heterogeneity.

Objective: This study aimed to investigate the relationship between IR and sarcopenia and the risk of new-onset CVD.

Methods: Using data from the China Health and Retirement Longitudinal Study (CHARLS). Cox proportional hazards models were used to assess associations of sarcopenia status (nonsarcopenia, possible sarcopenia, sarcopenia, and severe sarcopenia) and 6 IR indices (triglyceride-glucose, TyG; TyG-BMI; TyG-waist circumference; TyG-waist-to-height ratio; triglyceride/high-density lipoprotein cholesterol, TG/HDL-C; and metabolic score for insulin resistance, METS-IR) with incident CVD. Additive and multiplicative interaction analyses and subgroup analyses by age and sex were performed. Receiver operating characteristic analysis was used to determine clinically relevant cutoffs.

Results: In this study, during a median 9-year follow-up, we included 5514 middle- and older-aged (≥45 y) residents, of whom 550 presented with CVD incidence. Participants with possible sarcopenia and high IR exhibited 1.24 - 1.85-fold higher CVD risk versus nonsarcopenia and low-IR counterparts (P<.05) after adjustment for potential confounders. While TyG-BMI and TyG-waist circumference were the strongest independent predictors, formal interaction analysis revealed that the TG/HDL-C ratio and METS-IR demonstrated the most consistent synergistic effects with possible sarcopenia (relative excess risk due to interaction=0.139 and 0.074, respectively). In subgroups of different ages and sexes, the combination of IR and sarcopenia is associated with the highest risk of CVD. Receiver operating characteristic analysis provided clinically applicable cutoffs for these indices, including TG/HDL-C ≥2.09 and METS-IR ≥34.26.

Conclusions: We found that IR and sarcopenia, especially early-stage sarcopenia, synergistically increase the incidence of CVD in older adults. These findings advocate for dual-targeted CVD interventions (muscle preservation and IR mitigation) in aging societies, particularly during the transitional phase of possible sarcopenia.

(JMIR Aging 2025;8:e80115) doi:10.2196/80115

KEYWORDS

sarcopenia; insulin resistance; cardiovascular disease; China health and retirement longitudinal study; aging

Introduction

Aging populations represent a significant public health challenge globally, especially in Asian countries transitioning toward higher proportions of older individuals [1]. Previous research has indicated that age is a primary risk factor for cardiovascular disease (CVD), a leading cause of mortality in older adults. The global burden of CVD is substantial, with deaths rising from 13.1 million in 1990 to 19.2 million in 2023, of which ischemic heart disease and stroke constituted 81.8% [2]. This trend is markedly evident in China, where CVD-related deaths surged by 89.1%, from 2.4 million to 4.6 million, between 1990 and 2019 [3]. Critically, this burden is projected to rise continuously with the rapid growth of the aging population [4]. The ADDIN EN.CITE [5-7] susceptibility to CVD among older adults is further exacerbated by biological factors associated with aging, including hormonal changes, immunosenescence, impaired

XSL•FO
RenderX

autophagy, oxidative stress, and mitochondrial dysfunction [8]. Despite addressing well-established CVD risk factors, such as hypertension, diabetes, smoking, and obesity, CVD remains a main cause of mortality and disability among middle-aged and old-aged populations in China, indicating the presence of additional, less recognized risk factors. This study focuses on 2 such potential contributors: insulin resistance (IR) and sarcopenia.

IR is a critical pathological process associated with type 2 diabetes (T2D) and is linked to hypertension and coronary heart disease (CHD), contributing to CVD risk by increasing vascular stiffness [9-11]. However, standard methods for diagnosing IR are complex and time-consuming [12], prompting interest in alternative indices derived from biochemical markers and anthropometric factors, such as triglyceride-glucose (TyG) [13] and TyG-BMI [14]. While studies have established a connection between TyG levels and CVD risk [15,16], further inquiry into other IR indicators is warranted.

Sarcopenia, marked by an age-related decline in muscle mass and strength, threatens the quality of life for older individuals by increasing the risk of falls, disability, and mortality. Recent studies highlight the link between sarcopenia and CVD, with conditions such as CHD [17,18], atrial fibrillation, heart failure (HF) [19,20], and stroke [21] being associated. Chin et al [22] identified a higher CVD risk in older adults with sarcopenia (odds ratio [OR] 1.768, 95% CI 1.075 - 2.909) among 1578 participants aged 65 and older using the Korean population data. A meta-analysis of 38 studies found that sarcopenia prevalence in patients with CVD is 35%, double that of the general population's 13%, with the highest risks observed in acute decompensated HF (61%) and chronic HF (32%) [8]. Fukuda et al [23] also found a significant relationship between sarcopenic obesity and CVD events, including unstable angina and stroke, in individuals with T2D. There is evidence that sarcopenia is linked to IR, with obesity, inflammation, and hormonal changes possibly explaining their interconnection. IR may disrupt muscle protein balance, leading to decreased muscle mass and functionality, while sarcopenia can impair glucose metabolism [24]. The coexistence of IR and sarcopenia may worsen physical decline through interconnected pathways, increasing CVD risk, particularly in middle-aged and older adults facing additional risk factors. Understanding the combined effects of IR and sarcopenia on CVD risk in older populations is critical for developing targeted prevention and intervention strategies, which could lead to improved health outcomes tailored for this demographic.

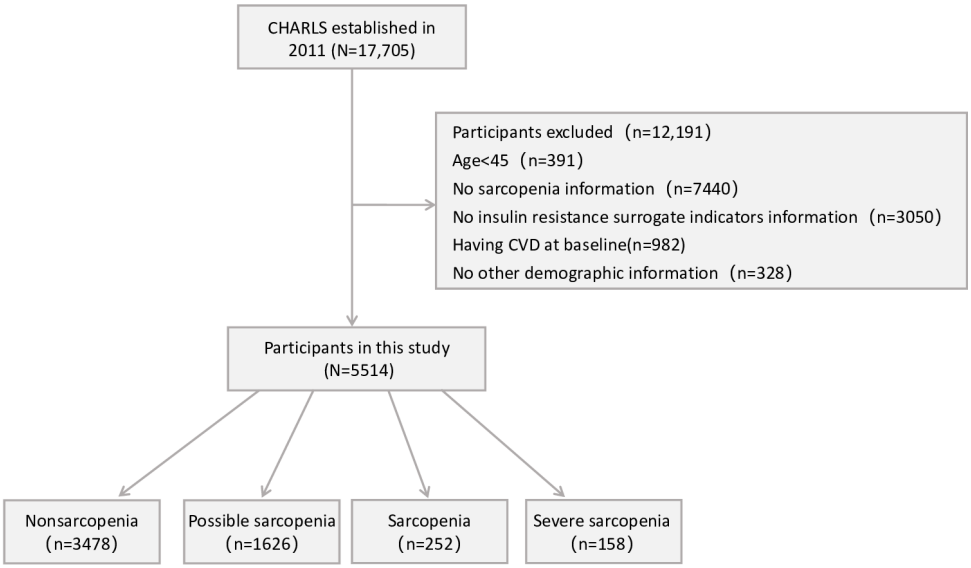
Methods

Study Design and Participants

This study used a longitudinal cohort design to examine the combined effects of IR and sarcopenia on the incidence of CVD in a middle- and older-aged population. Participants were selected from the China Health and Retirement Longitudinal Study (CHARLS), which spanned 150 counties and 450 communities across 28 provinces between 2011 and 2012. CHARLS has conducted 5 rounds of data collection, including the initial survey and 4 annual follow-ups until 2020 [25].

For this study, participants aged over 45 without prior CVD at baseline were included, drawn from the CHARLS data collected between 2011 and 2020. Exclusions were made for individuals with incomplete follow-up data, missing baseline sarcopenia status, absent blood samples, or insufficient demographic information. Ultimately, 5514 participants without self-reported CVD at baseline were included, with detailed selection criteria shown in Figure 1.

Figure 1. Flowchart for participants included in the study. CHARLS: China Health and Retirement Longitudinal Study; CVD: cardiovascular disease.



XSL•FO
RenderX

Assessment of Sarcopenia Status

Sarcopenia was defined according to the 2019 Asian Working Group for Sarcopenia criteria [26]. Handgrip strength was measured with the Yuejian WL-1000 dynamometer (Nantong Yuejian Physical Measurement Instrument Co. Ltd), with 2 measurements taken for both hands and the maximum value recorded. Diagnostic thresholds were set at <28 kg for men and <18 kg for women. Muscle mass was estimated using the appendicular skeletal muscle mass formula, which considers weight, height, sex, and age, showing concordance with dual-energy X-ray absorptiometry [27]. Participants were classified as having low muscle mass if height-adjusted appendicular skeletal muscle mass fell below 20% of the population, with criteria of <4.90 kg/m² for women and <6.79 kg/m² for men. Low physical performance was defined as a walking speed <1 m/s or a completion time of the 5-time chair stand test exceeding 12 seconds. Sarcopenia was diagnosed when low muscle mass coexisted with either low physical performance or low muscle strength, while those with only one of these indicators were classified as having possible sarcopenia [28-30]. Participants were categorized into 4 groups: nonsarcopenia (n=3478), possible sarcopenia (n=1626), sarcopenia (n=252), and severe sarcopenia (n=158).

Assessment of CVD Events

The study's primary outcome was the incidence of CVD, encompassing heart disease and stroke. The time to event was calculated from the baseline assessment to the first recorded CVD event, censoring, or study end, whichever came first. CVD events were identified through self-reported physician diagnoses using specific questions from the CHARLS questionnaire regarding heart disease and stroke diagnoses, such as "Have you been told by a doctor that you have been diagnosed with a heart attack, angina, CHD, HF, stroke, or other heart problems?" This method aligns with previous research [31,32], ensuring consistency and accuracy in CVD identification.

Calculation of IR Surrogate Indicators

Alongside CVD assessment, surrogate indicators of IR were evaluated. Fasting plasma glucose (FPG), triglycerides, and high-density lipoprotein cholesterol (HDL-C) levels were measured through venous blood samples taken after an overnight fast. Participants' height, weight, and waist circumference (WC) were measured using standard equipment. In total, 6 surrogate indicators of IR were computed based on validated equations from prior literature [33-36]:

1. TyG=Ln [(TG × FPG) / 2]
2. TyG-BMI=TyG × BMI (BMI=weight/ height², (kg/m²))
3. TyG-WC=TyG × WC
4. TyG-WHtR (waist-to-height ratio)=TyG × WHtR (WHtR=WC (cm) / height (cm))
5. TG/HDL-C=TG/HDL-C
6. METS-IR (metabolic score for insulin resistance)=Ln [(2× FPG)+ TG] × BMI/Ln (HDL-C)

Assessment of Covariates

Baseline data on participants were collected by trained professionals, covering demographic information, health behaviors, chronic disease history, medication use, and blood test results. Demographics included age, sex, and education level. Health characteristics, such as height, weight, BMI, WC, and WHtR, were assessed, along with participants' smoking and drinking habits. Chronic diseases and medication use were based on self-reported physician diagnoses, consistent with methodologies for evaluating CVD events in the CHARLS database. No biochemical indicators were used to define these conditions. Data collection adhered to rigorous scientific and ethical standards to ensure study integrity and participant privacy.

Statistical Analysis

In this study, we first assessed the distributional characteristics of continuous variables. For skewed continuous variables, data were reported as median (IQR), while normally distributed variables were represented as mean (SD). Categorical variables were expressed as counts and percentages. Participants were classified into four groups based on sarcopenia status and IR surrogate index levels at baseline: (1) nonsarcopenia with low-IR levels, (2) high-IR levels only, (3) possible sarcopenia, sarcopenia, and severe sarcopenia only, and (4) coexistence of sarcopenia and high-IR levels. Follow-up time was calculated from the date of the baseline interview (2011 - 2012) to the date of the first CVD event, the date of the last follow-up interview (up to 2020), or the date of censoring, whichever occurred first. The median follow-up time for the cohort was 9 years.

The Chi-square test was used for categorical variables, while the Mann-Whitney *U* test was applied to compare continuous variables across groups. We used the Cox proportional hazards model to assess the association between baseline sarcopenic status, IR levels, and incident CVD, calculating hazard ratios (HRs) with 95% CI values. Follow-up time was defined as the interval from the last interview to either the date of CVD diagnosis or the last interview date (March 2019).

Overall, 3 stepwise-adjusted Cox regression models were constructed. Model 1 was unadjusted; Model 2 was adjusted for age and sex; and Model 3 included additional adjustments for smoking status (current and former), alcohol consumption, obesity, hypertension, diabetes, dyslipidemia, and medications for hypertension, diabetes, and dyslipidemia. Variables significant in univariate analysis (P<.05) and clinically relevant were included in the multivariate analysis.

Additionally, we performed comprehensive interaction analyses between IR and sarcopenia. Primary analyses focused on possible sarcopenia based on its clinical relevance as an early intervention window and adequate sample size for robust estimation. Secondary exploratory analyses examined interactions with other sarcopenia severity levels (sarcopenia and severe sarcopenia). For multiplicative interaction assessment, product terms (eg, high-IR status × possible sarcopenia status) were included in Cox proportional hazards models. Additive interaction was quantified using the relative excess risk due to interaction (RERI), attributable proportion (AP), and synergy index (SI), which were calculated from HRs derived from Cox models that included a 4-category exposure variable (nonsarcopenia and low-IR, high-IR only, possible sarcopenia only, and both conditions).

XSL•FO
RenderX

To establish clinically applicable cutoff values for identifying high-risk individuals, we performed receiver operating characteristic (ROC) curve analysis. The optimal cutoff points for each IR index and sarcopenia-related measure were determined by maximizing the Youden index (J=sensitivity + specificity – 1), which provides the optimal balance between sensitivity and specificity for predicting incident CVD. This analysis was conducted using the pROC package in R, and the resulting cutoff values are presented in Table S6 in Multimedia Appendix 1. All statistical analyses were conducted using R statistical software (version 4.3.2; R Core Team), with a *P* value of less than .05 considered statistically significant.

Ethical Considerations

The CHARLS study protocol was approved by the Biomedical Ethics Committee of Peking University (IRB 00001052 - 11015). All participants provided written informed consent at the time of the original data collection. This study involved a secondary analysis of publicly available, deidentified CHARLS data. No direct contact with participants occurred, and no individual-level identifiable information was accessed. Therefore, additional ethics approval and informed consent were not required, as the original consent and institutional review board approval covered secondary use of the dataset. All data used in this study were fully anonymized to ensure participant privacy and confidentiality. Participants in the original CHARLS survey received compensation according to the study's standard protocol.

Results

Overview

The baseline characteristics of the participants stratified by the presence of CVD are presented in Table 1. A total of 5514 participants were included in the analysis, with 4529 classified as non-CVD and 985 as CVD cases. The median age of participants was 62 years, and 2157 (39%) were female. Among those with CVD, 550 (56%) were female. During a median follow-up of 9 years, 985 new cases of CVD were recorded, constituting 17.86% of the total. All 6 IR indicators (TyG, TyG-BMI, TyG-WC, TyG-WHtR, TG/HDL-C, and METS-IR) exhibited significantly elevated levels in patients with new-onset CVD compared to the control group without CVD ($P<.001$). The prevalence rates of sarcopenia by group were 63% for nonsarcopenia, 29% for possible sarcopenia, 4.6% for sarcopenia, and 2.9% for severe sarcopenia.

XSL•FO
RenderX

Table . Baseline patient characteristics grouped by outcomes.

Variables	All (N=5514)	Non-CVD[a] (n=4529)	CVD (n=985)	*P* value
Age (years)	62 (59-68)	62 (59-68)	62 (59-67)	.40
Sex, n (%)				<.001
Female	2707 (49)	2157 (48)	550 (56)	
Male	2807 (51)	2372 (52)	435 (44)	
Education, n (%)				.60
College or university	62 (1.1)	48 (1.1)	14 (1.4)	
Illiterate	1880 (34)	1537 (34)	343 (35)	
Primary school	2530 (46)	2078 (46)	452 (46)	
Secondary or high school	1032 (19)	858 (19)	174 (18)	
Systolic blood pressure (mm Hg), median (IQR)	130 (117-144)	129 (116-143)	134 (121-149)	<.001
Diastolic blood pressure (mm Hg), median (IQR)	74 (67-81)	74 (67-81)	76 (68-84)	<.001
HbA_{1c}[b], median (IQR)	5.20 (4.90-5.40)	5.20 (4.90-5.40)	5.20 (4.90-5.50)	.11
TG[c] (mg/dL), median (IQR)	103 (74-148)	100 (73-146)	115 (81-160)	<.001
TC[d] (mg/dL), median (IQR)	192 (168-217)	191 (167-216)	196 (173-222)	<.001
HDL-C[e] (mg/dL), median (IQR)	50 (41-61)	51 (41-61)	48 (40-60)	.002
LDL-C[f] (mg/dL), median (IQR)	116 (95-139)	115 (94-138)	119 (99-143)	<.001
WC[g] (cm), median (IQR)	85 (78-92)	85 (78-92)	87 (79-95)	<.001
Hypertension, n (%)	2395 (44)	1845 (41)	550 (56)	<.001
Diabetes, n (%)	415 (9.8)	305 (8.9)	110 (14)	<.001
Dyslipidemia, n (%)	1009 (25)	789 (24)	220 (29)	.002
Smoking, n (%)				<.001
Ex-smoker	466 (8.6)	378 (8.5)	88 (9.1)	
Nonsmoker	3197 (59)	2577 (58)	620 (64)	
Smoker	1735 (32)	1481 (33)	254 (26)	
Drinking, n (%)	1895 (34)	1584 (35)	311 (32)	.046
TyG[h], median (IQR)	8.59 (8.23-9.02)	8.56 (8.21-9)	8.71 (8.32-9.16)	<.001
Obesity, n (%)	2113 (38)	1660 (37)	453 (46)	<.001
WHtR[i], median (IQR)	0.54 (0.49-0.59)	0.54 (0.49-0.59)	0.55 (0.50-0.61)	<.001
TyG-BMI , median (IQR)	198 (173-226)	196 (172-223)	207 (182-236)	<.001
TyG-WC, median (IQR)	731 (650-815)	725 (646-806)	754 (673-851)	<.001
TyG-WHtR , median (IQR)	4.66 (4.14-5.25)	4.62 (4.10-5.20)	4.85 (4.33-5.45)	<.001
TG/HDLC, median (IQR)	2.05 (1.28-3.40)	1.97 (1.26-3.36)	2.31 (1.46-3.70)	<.001
METS-IR[j], median (IQR)	34 (29-40)	34 (29-39)	36 (31-42)	<.001
Sarcopenia state, n (%)				<.001
Nonsarcopenia	3478 (63)	2912 (64)	566 (57)	
Possible sarcopenia	1626 (29)	1273 (28)	353 (36)	
Sarcopenia	252 (4.6)	211 (4.7)	41 (4.2)	

XSL•FO
RenderX

Variables	All (N=5514)	Non-CVD[a] (n=4529)	CVD (n=985)	*P* value
Severe sarcopenia	158 (2.9)	133 (2.9)	25 (2.5)	

[a]CVD: cardiovascular disease.

[b]HbA_{1c}: glycosylated hemoglobin.

[c]TG: triglyceride.

[d]TC: total cholesterol.

[e]HDL-C: high-density lipoprotein cholesterol.

[f]LDL-C: low-density lipoprotein cholesterol.

[g]WC: waist circumference.

[h]TyG: triglyceride glucose.

[i]WHtR: waist-to-height ratio.

[j]METS-IR: metabolic score for insulin resistance.

Compared to participants without a history of CVD, those diagnosed with CVD had higher prevalence rates of obesity, increased WC, and higher levels of triglyceride, total cholesterol, and low-density lipoprotein cholesterol. Additionally, the prevalence of hypertension, diabetes, and dyslipidemia was significantly greater in the CVD group (*P*<.001). A detailed account of these demographic and clinical characteristics is illustrated in Table 1. Due to the results of the normality test indicating non-normal distribution, continuous variables are expressed as median (IQR) in Table 1, providing a more accurate representation of central tendency and variability.

Cox Regression Analysis of Baseline IR and Sarcopenia on Follow-Up CVD

Table 2 and Table S1 in Multimedia Appendix 1 detailed the association between the 6 IR surrogate indicators (TyG, TyG-BMI, TyG-WC, TyG-WHtR, TG/HDL-C, and METS-IR) and the risk of CVD, alongside the status of sarcopenia. In multivariate analysis, a significant association was found between TyG (per SD increment) and the risk of incident CVD (HR 1.176, 95% CI 1.030-1.343). Other IR surrogates, such as TyG-BMI (HR 1.004, 95% CI 1.001-1.007) and TyG-WC (HR 1.001, 95% CI 1.000-1.002), also demonstrated statistically significant relationships with increased CVD risk. Conversely, TG/HDL-C showed an attenuated result, with no significant relationship (HR 1.06, 95% CI 0.99-1.14). In mutual adjustment models, TyG, TyG-BMI, and TyG-WC remained statistically significant.

When sarcopenia status was categorized into 4 types, possible sarcopenia and severe sarcopenia were associated with an increased risk of CVD compared to the nonsarcopenia population, after adjusting for other confounders (HR 1.320, 95% CI 1.125-1.549) for possible sarcopenia and 1.709 (1.075-1.763) for severe sarcopenia. However, the risk of CVD associated with the sarcopenia group did not reach statistical significance (HR 1.181, 95% CI 0.791-1.763). Furthermore, significant associations were observed between possible sarcopenia and severe sarcopenia with the risk of incident CVD in multivariable analyses, even after adjustments for IR surrogate indicators. These findings indicate that both IR and sarcopenia play critical roles in predicting CVD risk.

XSL•FO
RenderX

Table . Relationship of insulin resistance surrogate indicators or sarcopenia status with cardiovascular disease risk.

Characteristics	Crude model, HR[a] (95% CI)	*P* value	Adjusted model[b], HR (95% CI)	*P* value	*P* for interaction[c]
IR[d] surrogate[e] indicators					
TyG[f] (per SD)	1.312 (1.186 - 1.451)	<.001	1.176 (1.030 - 1.343)	.02	.003
TyG-BMI (per SD)	1.006 (1.004 - 1.008)	<.001	1.004 (1.001 - 1.007)	.02	.01
TyG-WC[g] (per SD)	1.002 (1.001 - 1.002)	<.001	1.001 (1 - 1.002)	.02	.06
TyG-WHtR[h] (per SD)	1.275 (1.174 - 1.385)	<.001	1.096 (0.985 - 1.219)	.09	.04
TG[i]/HDL-C[j] (per SD)	1.035 (1.013 - 1.057)	<.001	1.007 (0.943 - 1.082)	.26	<.001
METS-IR[k] (per SD)	1.024 (1.015 - 1.033)	<.001	1.011 (0.980 - 1.035)	.12	.001
Sarcopenia status					
Nonsarcopenia	1 (Reference)	—[l]	1 (Reference)	—	—
Possible sarcopenia	1.462 (1.280 - 1.670)	<.001	1.320 (1.125 - 1.549)	.001	—
Sarcopenia	1.171 (0.853 - 1.608)	.33	1.181 (0.791 - 1.763)	.42	—
Severe sarcopenia	1.311 (0.878 - 1.957)	.19	1.709 (1.075 - 2.715)	.02	—

[a]HR: hazard ratio.

[b]Adjusted model: adjusted for age, sex, smoking, drinking, hypertension, diabetes, dyslipidemia, hypertension medications, diabetes medications, and dyslipidemia medications.

[c]An interaction between IR surrogate indicators and sarcopenia status.

[d]IR: insulin resistance.

[e]IR surrogate indicators are analyzed as a continuous variable in the table.

[f]TyG: triglyceride glucose.

[g]WC: waist circumference.

[h]WHtR: waist-to-height ratio.

[i]TG: triglyceride.

[j]HDL-C: high-density lipoprotein cholesterol.

[k]METS-IR: metabolic score for insulin resistance.

[l]Not available.

The Combination of IR and Sarcopenia Increased the Risk

Table 3 illustrates the crude and multivariable adjusted associations between the primary exposures and CVD risk. To illustrate, consider the combination of TyG and sarcopenia status, the incident CVD rates showed the following HRs in the unadjusted model: 1.44 (1.19-1.75) for nonsarcopenia and high TyG, 1.43 (1.12-1.82) for possible sarcopenia and low TyG, 2.01 (1.63-2.48) for possible sarcopenia and high TyG, and 1.86 (1.08-3.20) for severe sarcopenia and low TyG groups, relative to nonsarcopenia and low TyG. These associations remained statistically significant in Model 2, adjusted for age and sex, except for the severe sarcopenia/low TyG group.

After adjusting for potential confounders, including smoking and drinking habits, and chronic disease history (Model 3), we found that possible sarcopenia and high TyG is associated with a 1.53-fold higher hazard of experiencing CVD (HR 1.53, 95% CI 1.20-1.96), while sarcopenia and low TyG was associated with a 1.37-fold higher hazard (HR 1.37, 95% CI 1.07-1.76). The combination of other IR surrogates with sarcopenia also demonstrated similar patterns, with participants experiencing heightened CVD risk across various combinations of IR indicators and sarcopenia statuses (eg, HR 1.32, 95% CI 1.02 - 1.71) for possible sarcopenia with low TyG, 1.79 (1.30 - 2.46) for possible sarcopenia with high METS-IR). These findings emphasize the significance of considering both IR and sarcopenia in assessing CVD risk.

XSL•FO
RenderX

Table . Association of sarcopenia status and insulin resistance surrogate indicators with new onset of cardiovascular disease.

Groups[a]	Model 1[b], HR[c] (95% CI)	Model 2[d], HR (95% CI)	Model 3[e], HR (95% CI)
Nonsarcopenia and low TyG[f]	1 (Reference)	1 (Reference)	1 (Reference)
Nonsarcopenia and high TyG	1.44 (1.19 - 1.75)[g]	1.41 (1.17 - 1.71)[g]	1.18 (0.96 - 1.46)
Possible sarcopenia and low TyG	1.43 (1.12 - 1.82)[h]	1.37 (1.07 - 1.76)[i]	1.32 (1.02 - 1.71)[i]
Possible sarcopenia and high TyG	2.01 (1.63 - 2.48)[g]	1.83 (1.47 - 2.28)[g]	1.53 (1.20 - 1.96)[g]
Severe sarcopenia and low TyG	1.86 (1.08 - 3.20)[i]	1.35 (0.74 - 2.48)	1.68 (0.89 - 3.19)
Nonsarcopenia and low TyG-BMI	1 (Reference)	1 (Reference)	1 (Reference)
Nonsarcopenia and high TyG-BMI	1.66 (1.37 - 2.02)[g]	1.65 (1.35 - 2.01)[g]	1.25 (0.94 - 1.67)
Possible sarcopenia and low TyG-BMI	1.58 (1.22 - 2.03)[g]	1.50 (1.15 - 1.95)[h]	1.47 (1.12 - 1.92)[h]
Possible sarcopenia and high TyG-BMI	2.18 (1.75 - 2.70)[g]	2.02 (1.61 - 2.52)[g]	1.69 (1.20 - 2.39)[h]
Sarcopenia and low TyG-BMI	1.50 (1.03 - 2.20)[i]	1.21 (0.80 - 1.83)	1.13 (0.72 - 1.76)
Severe sarcopenia and low TyG-BMI	2.12 (1.35 - 3.31)[g]	1.52 (0.90 - 2.56)	1.50 (0.89 - 2.54)
Nonsarcopenia and low TyG-WC[j]	1 (Reference)	1 (Reference)	1 (Reference)
Nonsarcopenia and high TyG-WC	1.64 (1.35 - 1.99)[g]	1.62 (1.33 - 1.96)[g]	1.26 (0.98 - 1.61)
Possible sarcopenia and low TyG-WC	1.78 (1.40 - 2.28)[g]	1.71 (1.33 - 2.20)[g]	1.67 (1.29 - 2.17)[g]
Possible sarcopenia and high TyG-WC	1.95 (1.57 - 2.41)[g]	1.80 (1.44 - 2.25)[g]	1.30 (0.97 - 1.74)
Severe sarcopenia and low TyG-WC	2.02 (1.25 - 3.24)[h]	1.38 (0.79 - 2.39)	1.52 (0.86 - 2.68)
Nonsarcopenia and low TyG-WHtR[k]	1 (Reference)	1 (Reference)	1 (Reference)
Nonsarcopenia and high TyG-WHtR	1.64 (1.35 - 1.99)[g]	1.59 (1.31 - 1.94)[g]	1.15 (0.89 - 1.48)
Possible sarcopenia and low TyG-WHtR	1.80 (1.40 - 2.30)[g]	1.74 (1.36 - 2.24)[g]	1.71 (1.32 - 2.21)[g]
Possible sarcopenia and high TyG-WHtR	1.92 (1.56 - 2.37)[g]	1.90 (1.51 - 2.39)[g]	1.29 (0.97 - 1.73)
Sarcopenia and low TyG-WHtR	1.56 (1.06 - 2.30)[i]	1.29 (0.84 - 1.97)	1.27 (0.80 - 2.01)
Severe sarcopenia and low TyG-WHtR	1.88 (1.11 - 3.19)[i]	1.40 (0.77 - 2.56)	1.62 (0.88 - 2.96)
Severe sarcopenia and high TyG-WHtR	2.64 (1.24 - 5.62)[i]	2.25 (1.00 - 5.07)	1.64 (0.67 - 4.04)
Nonsarcopenia and low TG[l]/HDL-C[m]	1 (Reference)	1 (Reference)	1 (Reference)
Nonsarcopenia and high TG/HDL-C	1.46 (1.21 - 1.76)[g]	1.45 (1.20 - 1.75)[g]	1.27 (1.03 - 1.56)[i]
Possible sarcopenia and low TG/HDL-C	1.36 (1.06 - 1.75)[i]	1.31 (1.01 - 1.69)[g]	1.28 (0.99 - 1.67)
Possible sarcopenia and high TG/HDL-C	2.09 (1.70 - 2.57)[g]	1.92 (1.55 - 2.38)[g]	1.67 (1.32 - 2.13)[g]
Sarcopenia and high TG/HDL-C	1.92 (1.04 - 3.53)[i]	1.47 (0.76 - 2.82)	1.73 (0.86 - 3.48)

XSL•FO
RenderX

Groups[a]	Model 1[b], HR[c] (95% CI)	Model 2[d], HR (95% CI)	Model 3[e], HR (95% CI)
Severe sarcopenia and low TG/HDL-C	1.76 (1.02 - 3.03)[i]	1.27 (0.69 - 2.34)	1.66 (0.89 - 3.08)
Severe sarcopenia and high TG/HDL-C	2.30 (1.13 - 4.66)[i]	1.70 (0.79 - 3.63)	1.68 (0.73 - 3.85)
Nonsarcopenia and low METS-IR[n]	1 (Reference)	1 (Reference)	1 (Reference)
Nonsarcopenia and high METS-IR	1.61 (1.33 - 1.96)[g]	1.59 (1.31 - 1.94)[g]	1.23 (0.94 - 1.61)
Possible sarcopenia and low METS-IR	1.53 (1.19 - 1.97)[h]	1.43 (1.10 - 1.86)[h]	1.38 (1.05 - 1.81)[i]
Possible sarcopenia and high METS-IR	2.15 (1.74 - 2.66)[g]	2.00 (1.60 - 2.49)[g]	1.79 (1.30 - 2.46)[g]
Sarcopenia and high METS-IR	2.06 (1.12 - 3.79)[i]	1.46 (0.76 - 2.82)	1.52 (0.76 - 3.05)
Severe sarcopenia and low METS-IR	2.10 (1.34 - 3.27)[h]	1.43 (0.85 - 2.41)	1.49 (0.88 - 2.52)

[a]IR surrogate category is divided into high-value and low-value groups based on the median.

[b]Hazard ratios for combinations of sarcopenia and IR that were statistically significant (P<.05) in the unadjusted Model 1 combinations not significant in Model 1 are not shown. Model 1: unadjusted.

[c]HR: hazard ratio.

[d]Model 2: adjusted for age and sex.

[e]Model 3: adjusted as model 2 with further adjustment for smoking, drinking, hypertension, diabetes, dyslipidemia, hypertension medications, diabetes medications, and dyslipidemia medications.

[f]TyG: triglyceride glucose.

[g]P<.001.

[h]P<.01.

[i]P<.05.

[j]WC: waist circumference.

[k]WHtR: waist-to-height ratio.

[l]TG: triglyceride.

[m]HDL-C: high-density lipoprotein cholesterol.

[n]METS-IR: metabolic score for insulin resistance.

Interaction Analyses Between IR and Sarcopenia

To quantitatively characterize the joint effects of IR and sarcopenia on CVD risk, we conducted both multiplicative and additive interaction analyses using median-based cutoffs for IR indices.

Multiplicative interaction analysis revealed significant interactions for 2 indices: TyG-WC (HR 0.728, 95% CI 0.533 - 0.995; P=.046) and TyG-WHtR (HR 0.715, 95% CI 0.523 - 0.979; P=.04). For other IR indices, multiplicative interaction terms were not statistically significant (all P>.05). Additive interaction analysis, assessed using RERI, AP, and S, demonstrated distinct patterns: TG/HDL-C ratio showed clear synergistic effects (RERI=0.139, AP=0.083, SI=1.263); METS-IR demonstrated synergistic effects (RERI=0.039, AP=0.025, SI=1.075); TyG exhibited mild synergistic effects (RERI=0.045, AP=0.029, SI=1.089); Other indices (TyG-BMI, TyG-WC, and TyG-WHtR) showed antagonistic effects (negative RERI values). Complete results of both multiplicative and additive interaction analyses are presented in Table S2-5, in Multimedia Appendix 1.

Subgroup Analysis

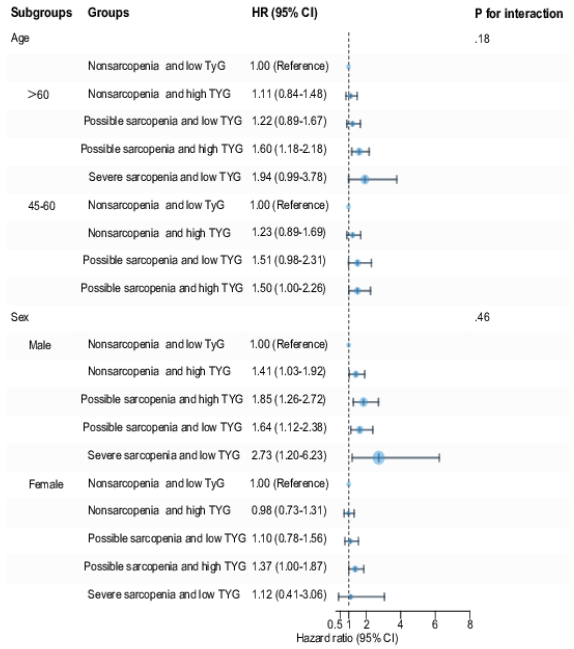
Figures 2-4 display the significant interplay between IR levels and sarcopenia status on CVD risk across different age and sex subgroups. Stratified Cox regression analyses, referenced against the nonsarcopenia with low IR (non_sp.IR_L) group, revealed distinct age- and sex-specific associations between sarcopenia severity and IR profiles. Notably, possible sarcopenia combined with high IR consistently represented the highest risk profile, particularly in younger adults (45 - 60 y). Even in low-IR states, those with possible sarcopenia and low TyG-WHtR independently increased their risk by 2.5-fold (HR 2.51, 95% CI 1.69 - 3.73), indicating potential sarcopenia-driven pathways beyond mere metabolic dysregulation. In male participants, possible sarcopenia combined with high IR, particularly lipid-centric markers such as elevated TG/HDL-C, conferred the greatest risk. In female participants, combinations of possible sarcopenia with high-IR indices linked to central adiposity (eg, TyG-WC) significantly elevated risk, with even low-IR subgroups exhibiting independent risks driven by sarcopenia.

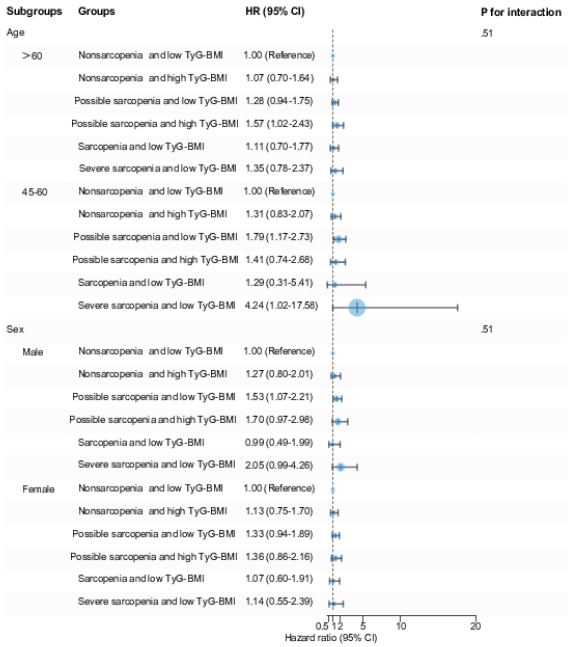
Due to small subgroup sizes for sarcopenia (n=252) and severe sarcopenia (n=158), analyses were limited in precision for extreme risk estimates. Models adjusted for age, sex, smoking

XSL•FO
RenderX

status, drinking, obesity, hypertension, diabetes, dyslipidemia, and medications confirmed the robustness of findings, underscoring the dual contribution of sarcopenia and IR to CVD pathogenesis. These results advocate for integrated clinical strategies targeting muscle health and metabolic dysregulation, particularly in high-risk subgroups characterized by possible sarcopenia and elevated IR.

Figure 2. Subgroup analysis of insulin resistance surrogate indicators and sarcopenia status with cardiovascular disease risk. Hazard ratios were adjusted for age, sex, smoking, drinking, hypertension, diabetes, dyslipidemia, hypertension medications, diabetes medications, and dyslipidemia medications. HR: hazard ratio; TyG: triglyceride glucose.





XSL•FO
RenderX

Figure 3. Subgroup analysis of insulin resistance surrogate indicators and sarcopenia status with cardiovascular disease risk. Hazard ratios were adjusted for age, sex, smoking, drinking, hypertension, diabetes, dyslipidemia, hypertension medications, diabetes medications, and dyslipidemia medications. HR: hazard ratio; TyG: triglyceride glucose; WC: waist circumference; WHtR: waist-to-height ratio.

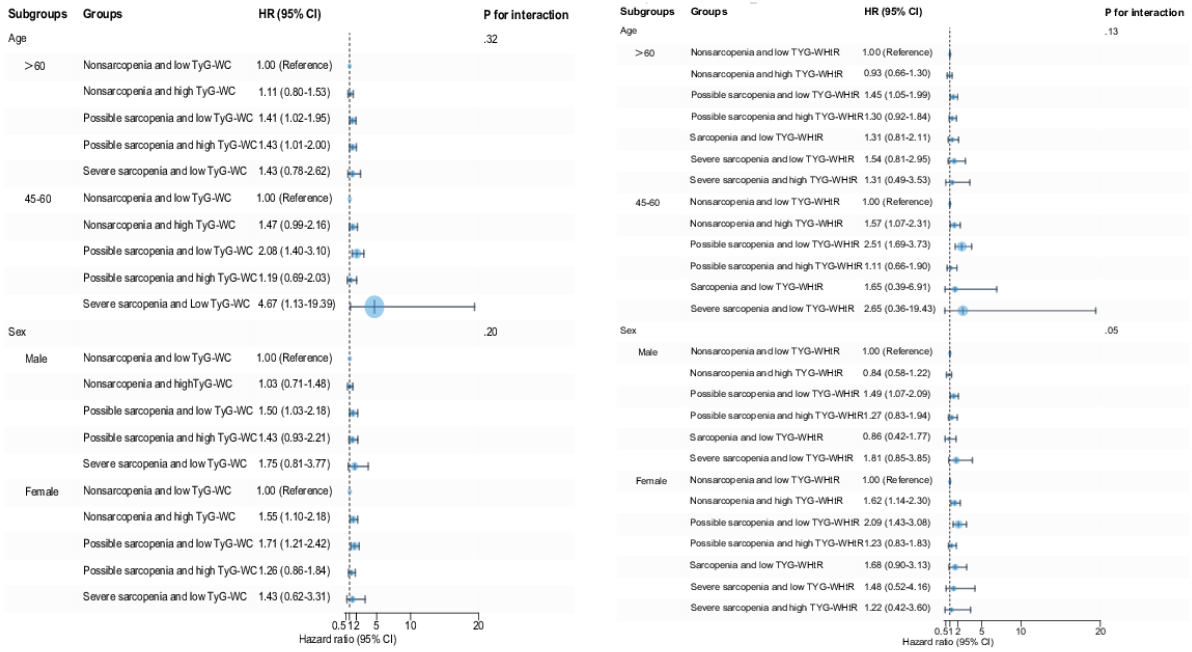
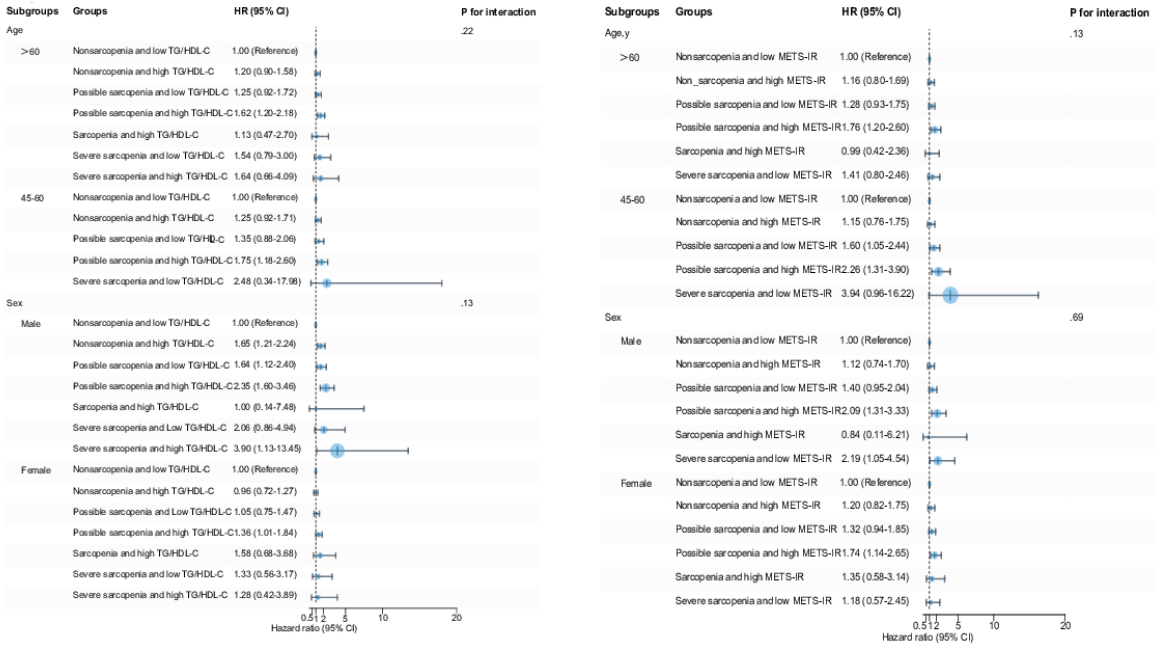


Figure 4. Subgroup analysis of insulin resistance surrogate indicators and sarcopenia status with cardiovascular disease risk. Hazard ratios were adjusted for age, sex, smoking, drinking, hypertension, diabetes, dyslipidemia, hypertension medications, diabetes medications, and dyslipidemia medications. HR: hazard ratio; METS-IR: metabolic score for insulin resistance; TG/HDL-C: triglyceride/high-density lipoprotein cholesterol; TyG: triglyceride glucose.



XSL•FO
RenderX

Clinically Applicable Cutoff Values

To facilitate the identification of high-risk individuals in clinical practice, we determined optimal cutoff values for all IR indices and sarcopenia measures using ROC analysis. Key cutoff values include: TG/HDL-C ratio 2.09 (sensitivity 58.8% and specificity 53.4%), METS-IR 34.26 (sensitivity 58.2% and specificity 54%), TyG-BMI 192.47 (sensitivity 65.9% and specificity 46.4%), handgrip strength 26.85 kg, and chair stand test time 9.41 seconds. The complete results, including cutoff values, sensitivity, specificity, and area under the curve, are presented in Table S6 in Multimedia Appendix 1.

Discussion

Principal Findings

In this longitudinal study, we observed a synergistic interplay between IR and sarcopenia regarding the risk of CVD among middle-aged and older individuals. Notably, the combination of high IR and possible sarcopenia exhibited the highest cardiovascular risk, particularly in those with elevated METS-IR levels, when compared to individuals in the nonsarcopenia and low METS-IR group. The prospective design ensures that assessments of sarcopenia and IR at baseline preceded CVD onset during follow-up, strengthening the inference of a causal pathway and reducing concerns of reverse causality. This research systematically correlates multidimensional IR indices with sarcopenia staging within a large national cohort, offering new insights into the confluence of IR and sarcopenia's impact on CVD risk.

Previous studies have demonstrated the significant positive association between the METS-IR index and CVD risk. Duan et al [37] found that METS-IR exhibited a stronger association with cardiovascular mortality compared to the TyG index, TG/HDL-C, and Homeostatic Model Assessment for Insulin Resistance, particularly in individuals under 65 years old. Another prospective cohort study comprising 47,270 participants from the Kailuan Study also indicated that elevated METS-IR levels increased CVD risk. The TyG index and TG/HDL-C ratio are increasingly recognized as surrogate markers of IR, confirming their relevance in CVD risk assessment. Previous research supports this association; for example, Salazar et al [38] conducted a 9-year prospective study involving 796 participants, finding a higher TG/HDL-C ratio as a predictor of incident CVD events. Similarly, Che et al [39] presented evidence of a significant correlation between elevated TyG and TG/HDL-C indices and increased CVD risks using the UK Biobank population. Further supporting these findings, a cross-sectional study involving 13,706 participants indicated that the TyG index independently predicts 10-year CVD risk, outperforming traditional IR measures such as Homeostatic Model Assessment for Insulin Resistance [14]. Additionally, Hong et al [40] also reported a robust link between elevated TyG indices and atherosclerotic CVD risk in older adults, highlighting the necessity of identifying optimal TyG levels for early intervention. In a retrospective cohort analysis involving 14,653 adults from the NHANES database (2001 - 2018), Li et al [41] conducted a longitudinal study based on the CHARLS database (2011 - 2015) and established a significant positive association between cumulative average TyG-BMI and incident CVD. Moreover, another longitudinal study from the CHARLS database involving 7115 participants demonstrated that TyG, TyG-BMI, TyG-WC, and TyG-WHtR were all significantly associated with new-onset CVDs. Notably, the TyG index outperformed the modified TyG indices in identifying CVD risk over a 7-year follow-up period [42]. In this study, we found that TyG, TyG-BMI, and TyG-WC functioned as robust predictors of CVD risk, independent of traditional confounders. Conversely, the associations of TyG-WHtR, TG/HDL-C, and METS-IR attenuated to nonsignificance after adjusting for factors such as age, sex, smoking, drinking habits, hypertension, diabetes, dyslipidemia, and obesity. These results highlight the critical role of specific IR markers in predicting CVD risk and support the relevance of the TyG index and its derivatives in clinical assessments and interventions aimed at improving cardiovascular health.

Our findings revealed that sarcopenia, alongside IR, serves as a significant predictor of CVD risk. We demonstrated that individuals with possible sarcopenia were at a higher risk for CVD compared to those without sarcopenia. In line with this, Gao et al [29] noted the association between possible sarcopenia and increased CVD risk, while Chai et al [43] large Taiwanese cohort study confirmed that sarcopenia correlated with a higher prevalence of stroke in patients with T2D. Xia et al [18] further showed that sarcopenia independently increased the risks of myocardial infarction and atrial fibrillation in middle-aged and older adults. However, our study presented sarcopenia status in 4 categories, revealing no statistically significant difference concerning CVD risk among sarcopenic groups, possibly attributable to limited sample sizes. Thus, more prospective multicenter studies are warranted for further validation. Given these findings, targeted interventions for patients diagnosed with sarcopenia, especially those with possible sarcopenia, are crucial. These interventions may include physical therapy to enhance muscle strength and function, along with medications to improve muscle health, thereby potentially reducing the incidence of CVD.

The mechanisms underlying this synergistic effect remain unclear; however, previous studies provide valuable insights. Skeletal muscle is pivotal for insulin-mediated glucose disposal postmeal [44,45]. With IR, glucose metabolism in muscle is disrupted, leading to decreased protein synthesis and increased breakdown, contributing to sarcopenia development [45]. Previous studies have found that IR can impair PI3K-Akt signaling, resulting in reduced glucose uptake, suppression of mTOR-mediated protein synthesis, and activation of the ubiquitin-proteasome system responsible for protein degradation [46]. Aging exacerbates these issues, as adipose tissue inflammation can lead to fat accumulation in the intra-abdominal region (visceral fat). Fatty infiltration in skeletal muscle leads to diminished strength and functionality [47]. Increased production of adipokines due to visceral fat promotes muscle loss through pro-inflammatory cytokines, which counteract anabolic signals from myokines, fostering a chronic inflammatory state that aggravates both IR and muscle degeneration [48,49]. Together, these processes contribute to CVD's onset and progression through mechanisms involving

XSL•FO
RenderX

metabolic disturbances, inflammation, oxidative stress, and vascular impairment [50-52].

In this study, we specifically demonstrated that individuals with both possible sarcopenia and high IR had a substantially enhanced CVD risk from 1.24-1.85 fold compared to the independent effects of these conditions. While prior studies typically approached sarcopenia dichotomously, our stratification by severity revealed a nonlinear relationship: early-stage sarcopenia may represent a critical window for metabolic intervention, while advanced muscle loss risks appear to follow a different risk trajectory. This pattern was reinforced in our subgroup analyses, which showed that the combination of possible sarcopenia and high IR consistently conferred robust and statistically significant risk across demographic strata. Interestingly, we observed numerically elevated, though statistically imprecise, HRs for severe sarcopenia even in subgroups with low-IR levels (Figure 2). This suggests that advanced muscle loss may itself constitute an independent risk factor, potentially operating through pathways distinct from metabolic dysregulation. However, the instability of these estimates, attributable to limited sample sizes in severe sarcopenia subgroups, necessitates cautious interpretation and underscores the need for future studies with larger cohorts to definitively characterize these relationships.

Formal interaction analyses further provided quantitative evidence supporting a synergistic relationship between IR and sarcopenia. Although interaction patterns varied across IR indices, the TG/HDL-C ratio and METS-IR consistently demonstrated synergistic effects on both additive and multiplicative scales. These composite measures, which integrate lipid metabolism, glucose homeostasis, and adiposity, may be particularly sensitive to the pathophysiological interplay between metabolic health and muscle quality. Their consistent performance highlights the biological complexity of IR and its connection to sarcopenia, while reinforcing our primary finding that the combination of possible sarcopenia and high IR confers the greatest CVD risk.

The ROC-derived cutoff values established in this study bridge the gap between epidemiological evidence and clinical practice by providing tangible thresholds for risk identification. Although single indices had modest discrimination (area under the curve 0.50‐0.58), the ROC-derived cutoffs nevertheless provide practical screening thresholds to identify individuals who warrant combined metabolic and muscle health assessment. Particularly noteworthy is the alignment between the cutoff values for TG/HDL-C ratio (2.09) and METS-IR (34.26) and our interaction analyses, which identified these indices as exhibiting the most consistent synergistic effects with possible sarcopenia. These findings support a practical 2-step approach to CVD risk stratification in middle-aged and older adults: (1) initial screening using the established cutoff values to identify individuals with high IR, impaired muscle health, or both; (2) focused intervention for those with both conditions, especially the combination of high TG/HDL-C (≥2.09) or high METS-IR (≥34.26) with possible sarcopenia, as this profile confers the highest CVD risk. The accessibility and calculability of these indices enhance their use in primary care and community settings where complex diagnostic procedures may be unavailable. Furthermore, the early identification of possible sarcopenia, a potentially reversible stage, combined with metabolic assessment, creates a critical window for implementing dual-targeted interventions aimed at preserving muscle mass while improving metabolic health.

Collectively, these findings emphasize the necessity of tailored management strategies addressing both sarcopenia and IR, particularly during the early stages of muscle decline. Proactive intervention during this critical window may improve clinical outcomes and attenuate CVD progression.

While this study elucidates the synergistic effect of sarcopenia staging and IR on CVD risk within a nationally representative older adults cohort, it is subject to certain limitations. First, the reliance on self-reported CVD diagnoses may introduce misclassification bias. We were unable to validate incident CVD outcomes against medical records, administrative claims, or mortality registries. Consequently, outcome misclassification may have occurred. If such misclassification is nondifferential with respect to the exposures studied, it would likely bias HR estimates toward the null, implying that the true associations might be stronger than those observed. However, differential misclassification (eg, systematic underreporting among older, less educated, or rural participants) could bias estimates in either direction, and we were not able to quantify the magnitude of these potential biases in the current dataset. Future studies that link CHARLS participants to hospital records, insurance claims, or death registries, or that apply adjudicated outcome ascertainment, would substantially strengthen the validity of outcome assessment. Second, the wide CI values in severe sarcopenia subgroups due to small sample sizes necessitate cautious interpretation, limiting the generalizability of findings to a Chinese-specific context. Moreover, we were unable to monitor changes in IR levels and sarcopenia status over time, restricting our understanding of their dynamic roles in CVD development. Future studies should incorporate additional potential confounders and explore the long-term effects of changes in IR and sarcopenia status on CVD risk, guiding precision interventions.

Conclusions

In conclusion, this study has evidenced that IR and sarcopenia, particularly in their early possible stage, interact synergistically to elevate CVD risk in middle-aged and older adults. The findings underscore the urgent need for early identification and intervention strategies, especially for individuals experiencing possible sarcopenia alongside high-IR levels. Such proactive approaches could significantly influence patient outcomes, reducing cardiovascular event incidence and enhancing prognosis and quality of life for middle-aged and older adults. However, further prospective studies are essential to confirm these findings and clarify the mechanisms behind the interaction between IR and sarcopenia. Addressing these gaps will be critical in developing effective clinical strategies to manage these conditions and improve health outcomes in this vulnerable population.

XSL•FO
RenderX

Acknowledgments

The authors thank all participants and staff of the China Health and Retirement Longitudinal Study.

Funding

This work was supported by the National Natural Science Foundation of China (grant number 82374406), Science and Technology Projects in Guangzhou (grant number 2024B03J1298).

Data Availability

The datasets analyzed during this study are available on the China Health and Retirement Longitudinal Study website repository [53].

Authors' Contributions

HL designed the research, completed statistical analysis, and drafted the article. SL carried out data collection and provided statistical support. XZ assisted in statistical analysis and assisted in writing articles. WL and HL assisted in writing the article. LL and SN reviewed the analysis process and checked the results. ZY revised the article and supervised the project.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Analyses of the associations and interactions between insulin resistance surrogate indicators, sarcopenia, and cardiovascular disease risk.
[DOCX File, 35 KB - aging_v8i1e80115_app1.docx]

References

1. The Lancet Public Health. Ageing: a 21st century public health challenge? Lancet Public Health 2017 Jul;2(7):e297. [doi: 10.1016/S2468-2667(17)30125-1] [Medline: 29253452]
2. Global Burden of Cardiovascular Diseases and Risks 2023 Collaborators. Global, regional, and national burden of cardiovascular diseases and risk factors in 204 countries and territories, 1990-2023. J Am Coll Cardiol 2025 Dec 2;86(22):2167-2243. [doi: 10.1016/j.jacc.2025.08.015] [Medline: 40990886]
3. National Center For Cardiovascular Diseases The Writing Committee Of The Report On Cardiovascular Health And Diseases In China 1. Report on cardiovascular health and diseases in China 2023: an updated summary. Biomed Environ Sci 2024 Sep 20;37(9):949-992. [doi: 10.3967/bes2024.162] [Medline: 39401992]
4. Zhou M, Zhao G, Zeng Y, Zhu J, Cheng F, Liang W. Aging and cardiovascular disease: current status and challenges. Rev Cardiovasc Med 2022 Apr;23(4):135. [doi: 10.31083/j.rcm2304135] [Medline: 39076212]
5. GBD 2016 Causes of Death Collaborators. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017 Sep 16;390(10100):1151-1210. [doi: 10.1016/S0140-6736(17)32152-9] [Medline: 28919116]
6. Dunbar SB, Khavjou OA, Bakas T, et al. Projected costs of informal caregiving for cardiovascular disease: 2015 to 2035: a policy statement from the American Heart Association. Circulation 2018 May 8;137(19):e558-e577. [doi: 10.1161/CIR.0000000000000570] [Medline: 29632217]
7. Corbett S, Courtiol A, Lummaa V, Moorad J, Stearns S. The transition to modernity and chronic disease: mismatch and natural selection. Nat Rev Genet 2018 Jul;19(7):419-430. [doi: 10.1038/s41576-018-0012-3] [Medline: 29743650]
8. Zuo X, Li X, Tang K, et al. Sarcopenia and cardiovascular diseases: a systematic review and meta-analysis. J Cachexia Sarcopenia Muscle 2023 Jun;14(3):1183-1198. [doi: 10.1002/jcsm.13221] [Medline: 37002802]
9. Castro L, Brant L, Diniz MDF, et al. Association of hypertension and insulin resistance in individuals free of diabetes in the ELSA-Brasil cohort. Sci Rep 2023 Jun 10;13(1):9456. [doi: 10.1038/s41598-023-35298-y] [Medline: 37301876]
10. Pyörälä M, Miettinen H, Halonen P, Laakso M, Pyörälä K. Insulin resistance syndrome predicts the risk of coronary heart disease and stroke in healthy middle-aged men: the 22-year follow-up results of the Helsinki Policemen Study. Arterioscler Thromb Vasc Biol 2000 Feb;20(2):538-544. [doi: 10.1161/01.atv.20.2.538] [Medline: 10669654]
11. Hill MA, Yang Y, Zhang L, et al. Insulin resistance, cardiovascular stiffening and cardiovascular disease. Metab Clin Exp 2021 Jun;119:154766. [doi: 10.1016/j.metabol.2021.154766] [Medline: 33766485]
12. Shulman GI, Rothman DL, Jue T, Stein P, DeFronzo RA, Shulman RG. Quantitation of muscle glycogen synthesis in normal subjects and subjects with non-insulin-dependent diabetes by 13C nuclear magnetic resonance spectroscopy. N Engl J Med 1990 Jan 25;322(4):223-228. [doi: 10.1056/NEJM199001253220403] [Medline: 2403659]

XSL•FO
RenderX

13. Guerrero-Romero F, Simental-Mendía LE, González-Ortiz M, et al. The product of triglycerides and glucose, a simple measure of insulin sensitivity. Comparison with the euglycemic-hyperinsulinemic clamp. J Clin Endocrinol Metab 2010 Jul;95(7):3347-3351. [doi: 10.1210/jc.2010-0288] [Medline: 20484475]
14. Guo W, Zhu W, Wu J, et al. Triglyceride glucose index is associated with arterial stiffness and 10-year cardiovascular disease risk in a Chinese population. Front Cardiovasc Med 2021;8:585776. [doi: 10.3389/fcvm.2021.585776] [Medline: 33816569]
15. Lopez-Jaramillo P, Gomez-Arbelaez D, Martinez-Bello D, et al. Association of the triglyceride glucose index as a measure of insulin resistance with mortality and cardiovascular disease in populations from five continents (PURE study): a prospective cohort study. Lancet Healthy Longev 2023 Jan;4(1):e23-e33. [doi: 10.1016/S2666-7568(22)00247-1] [Medline: 36521498]
16. Zhang N, Chi X, Zhou Z, et al. Triglyceride-glucose index is associated with a higher risk of stroke in a hypertensive population. Cardiovasc Diabetol 2023 Dec 13;22(1):346. [doi: 10.1186/s12933-023-02082-1] [Medline: 38093283]
17. Santana NDM, Mendes RML, Silva ND, Pinho CPS. Sarcopenia and sarcopenic obesity as prognostic predictors in hospitalized elderly patients with acute myocardial infarction [Article in English, Portuguese]. Einstein (Sao Paulo) 2019 Aug 19;17(4):eAO4632. [doi: 10.31744/einstein_journal/2019AO4632] [Medline: 31433007]
18. Xia MF, Chen LY, Wu L, et al. Sarcopenia, sarcopenic overweight/obesity and risk of cardiovascular disease and cardiac arrhythmia: a cross-sectional study. Clin Nutr 2021 Feb;40(2):571-580. [doi: 10.1016/j.clnu.2020.06.003] [Medline: 32593523]
19. Curcio F, Testa G, Liguori I, et al. Sarcopenia and heart failure. Nutrients 2020 Jan 14;12(1):211. [doi: 10.3390/nu12010211] [Medline: 31947528]
20. Suzuki T, Palus S, Springer J. Skeletal muscle wasting in chronic heart failure. ESC Heart Fail 2018 Dec;5(6):1099-1107. [doi: 10.1002/ehf2.12387] [Medline: 30548178]
21. Li W, Yue T, Liu Y. New understanding of the pathogenesis and treatment of stroke-related sarcopenia. Biomed Pharmacother 2020 Nov;131:110721. [doi: 10.1016/j.biopha.2020.110721] [Medline: 32920517]
22. Chin SO, Rhee SY, Chon S, et al. Sarcopenia is independently associated with cardiovascular disease in older Korean adults: the Korea National Health and Nutrition Examination Survey (KNHANES) from 2009. PLoS ONE 2013;8(3):e60119. [doi: 10.1371/journal.pone.0060119] [Medline: 23533671]
23. Fukuda T, Bouchi R, Takeuchi T, et al. Sarcopenic obesity assessed using dual energy X-ray absorptiometry (DXA) can predict cardiovascular disease in patients with type 2 diabetes: a retrospective observational study. Cardiovasc Diabetol 2018 Apr 10;17(1):55. [doi: 10.1186/s12933-018-0700-5] [Medline: 29636045]
24. Hong SH, Choi KM. Sarcopenic obesity, insulin resistance, and their implications in cardiovascular and metabolic consequences. Int J Mol Sci 2020 Jan 13;21(2):494. [doi: 10.3390/ijms21020494] [Medline: 31941015]
25. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). Int J Epidemiol 2014 Feb;43(1):61-68. [doi: 10.1093/ije/dys203] [Medline: 23243115]
26. Chen LK, Woo J, Assantachai P, et al. Asian working group for sarcopenia: 2019 consensus update on sarcopenia diagnosis and treatment. J Am Med Dir Assoc 2020 Mar;21(3):300-307. [doi: 10.1016/j.jamda.2019.12.012] [Medline: 32033882]
27. Delmonico MJ, Harris TB, Lee JS, et al. Alternative definitions of sarcopenia, lower extremity performance, and functional impairment with aging in older men and women. J Am Geriatr Soc 2007 May;55(5):769-774. [doi: 10.1111/j.1532-5415.2007.01140.x] [Medline: 17493199]
28. Alexandre TDS, Duarte YDO, Santos JLF, Wong R, Lebrão ML. Sarcopenia according to the European working group on sarcopenia in older people (EWGSOP) versus dynapenia as a risk factor for mortality in the elderly. J Nutr Health Aging 2014;18(8):751-756. [doi: 10.1007/s12603-014-0540-2] [Medline: 25286455]
29. Gao K, Cao LF, Ma WZ, et al. Association between sarcopenia and cardiovascular disease among middle-aged and older adults: findings from the China health and retirement longitudinal study. EClinicalMedicine 2022 Feb;44:101264. [doi: 10.1016/j.eclinm.2021.101264] [Medline: 35059617]
30. Wu X, Li X, Xu M, Zhang Z, He L, Li Y. Sarcopenia prevalence and associated factors among older Chinese population: findings from the China Health and Retirement Longitudinal Study. PLoS ONE 2021;16(3):e0247617. [doi: 10.1371/journal.pone.0247617] [Medline: 33661964]
31. Li H, Zheng D, Li Z, et al. Association of depressive symptoms with incident cardiovascular diseases in middle-aged and older Chinese adults. JAMA Netw Open 2019 Dec 2;2(12):e1916591. [doi: 10.1001/jamanetworkopen.2019.16591] [Medline: 31800066]
32. Xie W, Zheng F, Yan L, Zhong B. Cognitive decline before and after incident coronary events. J Am Coll Cardiol 2019 Jun 25;73(24):3041-3050. [doi: 10.1016/j.jacc.2019.04.019] [Medline: 31221251]
33. Bello-Chavolla OY, Almeda-Valdes P, Gomez-Velasco D, et al. METS-IR, a novel score to evaluate insulin sensitivity, is predictive of visceral adiposity and incident type 2 diabetes. Eur J Endocrinol 2018 May;178(5):533-544. [doi: 10.1530/EJE-17-0883] [Medline: 29535168]
34. Er LK, Wu S, Chou HH, et al. Triglyceride glucose-body mass index is a simple and clinically useful surrogate marker for insulin resistance in nondiabetic individuals. PLoS ONE 2016;11(3):e0149731. [doi: 10.1371/journal.pone.0149731] [Medline: 26930652]

XSL•FO
RenderX

35. Lim J, Kim J, Koo SH, Kwon GC. Comparison of triglyceride glucose index, and related parameters to predict insulin resistance in Korean adults: an analysis of the 2007-2010 Korean National Health and Nutrition Examination Survey. PLoS ONE 2019;14(3):e0212963. [doi: 10.1371/journal.pone.0212963] [Medline: 30845237]
36. Liu H, Liu J, Liu J, Xin S, Lyu Z, Fu X. Triglyceride to high-density lipoprotein cholesterol (TG/HDL-C) ratio, a simple but effective indicator in predicting type 2 diabetes mellitus in older adults. Front Endocrinol (Lausanne) 2022;13:828581. [doi: 10.3389/fendo.2022.828581] [Medline: 35282431]
37. Duan M, Zhao X, Li S, et al. Metabolic score for insulin resistance (METS-IR) predicts all-cause and cardiovascular mortality in the general population: evidence from NHANES 2001-2018. Cardiovasc Diabetol 2024 Jul 10;23(1):243. [doi: 10.1186/s12933-024-02334-8] [Medline: 38987779]
38. Salazar MR, Carbajal HA, Espeche WG, et al. Identifying cardiovascular disease risk and outcome: use of the plasma triglyceride/high-density lipoprotein cholesterol concentration ratio versus metabolic syndrome criteria. J Intern Med 2013 Jun;273(6):595-601. [doi: 10.1111/joim.12036] [Medline: 23331522]
39. Che B, Zhong C, Zhang R, et al. Triglyceride-glucose index and triglyceride to high-density lipoprotein cholesterol ratio as potential cardiovascular disease risk factors: an analysis of UK biobank data. Cardiovasc Diabetol 2023 Feb 16;22(1):34. [doi: 10.1186/s12933-023-01762-2] [Medline: 36797706]
40. Hong S, Han K, Park CY. The triglyceride glucose index is a simple and low-cost marker associated with atherosclerotic cardiovascular disease: a population-based study. BMC Med 2020 Nov 25;18(1):361. [doi: 10.1186/s12916-020-01824-2] [Medline: 33234146]
41. Li F, Wang Y, Shi B, et al. Association between the cumulative average triglyceride glucose-body mass index and cardiovascular disease incidence among the middle-aged and older population: a prospective nationwide cohort study in China. Cardiovasc Diabetol 2024 Jan 6;23(1):16. [doi: 10.1186/s12933-023-02114-w] [Medline: 38184577]
42. Cui C, Qi Y, Song J, et al. Comparison of triglyceride glucose index and modified triglyceride glucose indices in prediction of cardiovascular diseases in middle aged and older Chinese adults. Cardiovasc Diabetol 2024 May 29;23(1):185. [doi: 10.1186/s12933-024-02278-z] [Medline: 38812015]
43. Chai KC, Chen WM, Chen M, Shia BC, Wu SY. Association between preexisting sarcopenia and stroke in patients with type 2 diabetes mellitus. J Nutr Health Aging 2022;26(10):936-944. [doi: 10.1007/s12603-022-1846-0] [Medline: 36259582]
44. DeFronzo RA, Tripathy D. Skeletal muscle insulin resistance is the primary defect in type 2 diabetes. Diabetes Care 2009 Nov;32 Suppl 2(Suppl 2):S157-S163. [doi: 10.2337/dc09-S302] [Medline: 19875544]
45. Kalyani RR, Corriere M, Ferrucci L. Age-related and disease-related muscle loss: the effect of diabetes, obesity, and other diseases. Lancet Diabetes Endocrinol 2014 Oct;2(10):819-829. [doi: 10.1016/S2213-8587(14)70034-8] [Medline: 24731660]
46. Sartori R, Romanello V, Sandri M. Mechanisms of muscle atrophy and hypertrophy: implications in health and disease. Nat Commun 2021 Jan 12;12(1):330. [doi: 10.1038/s41467-020-20123-1] [Medline: 33436614]
47. Li CW, Yu K, Shyh-Chang N, et al. Pathogenesis of sarcopenia and the relationship with fat mass: descriptive review. J Cachexia Sarcopenia Muscle 2022 Apr;13(2):781-794. [doi: 10.1002/jcsm.12901] [Medline: 35106971]
48. Abete I, Konieczna J, Zulet MA, et al. Association of lifestyle factors and inflammation with sarcopenic obesity: data from the PREDIMED-Plus trial. J Cachexia Sarcopenia Muscle 2019 Oct;10(5):974-984. [doi: 10.1002/jcsm.12442] [Medline: 31144432]
49. Bilski J, Pierzchalski P, Szczepanik M, Bonior J, Zoladz JA. Multifactorial mechanism of sarcopenia and sarcopenic obesity. role of physical exercise, microbiota and myokines. Cells 2022 Jan 4;11(1):160. [doi: 10.3390/cells11010160] [Medline: 35011721]
50. Kernan WN, Inzucchi SE, Viscoli CM, Brass LM, Bravata DM, Horwitz RI. Insulin resistance and risk for stroke. Neurology (ECronicon) 2002 Sep 24;59(6):809-815. [doi: 10.1212/wnl.59.6.809] [Medline: 12349850]
51. Kim TN, Choi KM. The implications of sarcopenia and sarcopenic obesity on cardiometabolic disease. J Cell Biochem 2015 Jul;116(7):1171-1178. [doi: 10.1002/jcb.25077] [Medline: 25545054]
52. Korytowski W, Wawak K, Pabisz P, et al. Impairment of macrophage cholesterol efflux by cholesterol hydroperoxide trafficking: implications for atherogenesis under oxidative stress. Arterioscler Thromb Vasc Biol 2015 Oct;35(10):2104-2113. [doi: 10.1161/ATVBAHA.115.306210] [Medline: 26315403]
53. China Health and Retirement Longitudinal Study [Web page in Chinese]. CHARLS. URL: https://charls.pku.edu.cn/en/ [accessed 2025-10-10]

Abbreviations

AP: attributable proportion
CHARLS: China Health and Retirement Longitudinal Study
CHD: coronary heart disease
CVD: cardiovascular disease
FPG: fasting plasma glucose
HDL-C: high-density lipoprotein cholesterol
HF: heart failure

XSL•FO
RenderX

HR: hazard ratio
IR: insulin resistance
METS-IR: metabolic score for insulin resistance
RERI: relative excess risk due to interaction
ROC: receiver operating characteristic
SI: synergy index
T2D: type 2 diabetes
TyG: triglyceride glucose
WC: waist circumference
WHtR: waist-to-height ratio

Edited by Y Du; submitted 06.07.25; peer-reviewed by B Liu, Y Zhao; revised version received 12.10.25; accepted 07.11.25; published 31.12.25.

Please cite as:
Liu H, Li S, Zhang X, Long W, Liao H, Lu L, Ni S, Yang Z
Association of Insulin Resistance, Sarcopenia, and Risk of Cardiovascular Disease: Findings From the China Health and Retirement Longitudinal Study
JMIR Aging 2025;8:e80115
URL: https://aging.jmir.org/2025/1/e80115
doi:10.2196/80115

© Huilin Liu, Sijing Li, Ximin Zhang, Wenjie Long, Huili Liao, Lu Lu, Shihao Ni, Zhongqi Yang. Originally published in JMIR Aging (https://aging.jmir.org), 31.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Risk Factors Related to Low Calf Circumference in Older Adults With Multimorbidity: Cross-Sectional Latent Class Analysis

Xilin Peng[1,2*], MD; Xudong Chen[1,2*], MD; Ruihao Zhou[1,2], MD; Fanfan Shi[3], MD; Tao Zhu[1,2], MD; Guo Chen[1,2], MD

[1]Department of Anesthesiology, National Clinical Research Center for Geriatrics, West China Hospital, Sichuan University, No. 37 Guoxue Lane, Wuhou District, Chengdu, Sichuan, China

[2]The Research Units of West China (2018RU012)–Chinese Academy of Medical Sciences, West China Hospital, Sichuan University, Chengdu, Sichuan, China

[3]Department of Clinical Research and Management, Center of Biostatistics, Design, Measurement and Evaluation (CBDME), West China Hospital, Sichuan University, Chengdu, Sichuan, China

[*]these authors contributed equally

Corresponding Author:
Guo Chen, MD
Department of Anesthesiology, National Clinical Research Center for Geriatrics, West China Hospital, Sichuan University, No. 37 Guoxue Lane, Wuhou District, Chengdu, Sichuan, China

Abstract

Background: As the global population continues to age, the prevalence of sarcopenia is gradually increasing, and the loss of skeletal muscle mass is one of the manifestations of sarcopenia. Low calf circumference (CC) is often used as a predictor of poor skeletal muscle mass or sarcopenia. Older adults usually have a combination of multiple chronic diseases. There is a lack of evidence to explore the risk factors for low CC with multimorbidity in Chinese, community-dwelling, older adults.

Objective: This study aimed to explore the risk factors and potential categories in older adult patients with low CC and multimorbidity from an individual-centered perspective.

Methods: We selected 15,874 participants from the Chinese Longitudinal Healthy Longevity Survey in 2018 and screened for low CC in older adult patients. The individual-centered latent class analysis was used to classify potential multimorbidity groups. Multiple logistic regression was used to explore the risk factors associated with low CC and multimorbidity by applying the elastic net to screen for reliable risk variables.

Results: A total of 7956 older individuals were eligible for the study, of whom 3960 (49.8%) were aged >90 years and 2166 (27.2%) had multimorbidity with low CC. The prevalence of multimorbidity increases between the ages of 65 and 89 years. However, the majority of older adults remain in reasonably good health beyond the age of 90 years. Five multimorbidity groups were identified by latent class analysis: multisystem morbidity diseases (78/2166, 3.6%), arthritis-rheumatism or rheumatoid diseases (400/2166, 18.47%), diabetes-hypertension diseases (330/2166, 15.23%), respiratory-heart diseases (347/2166, 16.02%), and cardiovascular diseases (1011/2166, 46.68%). Through 12 variables screened by the elastic net, multiple logistic regression showed different impacts on multimorbidity groups, including demographic background, behavioral characteristics, and physical and mental health factors. In particular, older patients who self-report poor health and live in urban areas need more attention.

Conclusions: The results revealed that low CC is a common phenomenon among community-dwelling older adults, and a substantial proportion also present with multimorbidity. In the older adult population with low CC, the proportion of multimorbidity does not simply increase with age. Multimorbidity in low CC has been identified in 5 potential groups. Different groups have distinctive risk factors. Public health authorities should pay attention to low CC in older adult patients with multimorbidity and carry out targeted interventions, thereby enhancing health outcomes.

(JMIR Aging 2025;8:e68760) doi:10.2196/68760

KEYWORDS

low calf circumference; multimorbidity; sarcopenia; older adults; China

XSL•FO
RenderX

Introduction

Background

It is projected that approximately 200 million individuals worldwide will develop sarcopenia by 2050 [1]. Muscle fiber degeneration typically begins around the age of 50 years, with up to 50% of fibers diminished by the age of 80 years. Notably, this decline is also observed in physically active individuals, including athletes [2]. Calf circumference (CC) is a commonly used measurement to assess lower limb anthropometrics; it outperforms BMI or waist circumference and exhibits moderate to high sensitivity and specificity in predicting sarcopenia or poor skeletal muscle mass [3,4]. According to the Asian Working Group for Sarcopenia 2019 guidelines, low CC can be used to identify patients with sarcopenia in advance among older adult individuals in China. Increasing evidence suggests that low CC is capable of detecting patients with sarcopenia across various physical conditions [5,6]. In addition, the research found that low CC was associated with disability in people aged >60 - 70 years [7] and could even predict the mortality risk among older adults in both the United States and Taiwan [8,9].

The prevalence of type 2 diabetes, cerebrovascular diseases, cancer, and neurodegenerative disorders significantly increases with age [10]. Older adults often have more than one chronic disease. Multisystem disorders can shorten life expectancy and exacerbate already existing imbalances in the distribution of medical resources [11]. Approximately 42.4% of people in China aged >50 years live with multiple diseases [12]. Nearly 70% of all fatalities in China are attributed to the nation's rapidly aging population and the high prevalence of chronic illnesses [12]. Studies have confirmed that multimorbidity is not a simple random combination but has some underlying physiological and pathological connections [13]. Identifying the combination of different categories in multimorbidity will be beneficial in achieving an individual-centered medical care pattern and improving the clinical outcomes of patients with multimorbidity.

The evidence from UK Biobank participants suggests a link between probable sarcopenia and long-term conditions as well as multimorbidity [14]. A study identified 5 distinct multimorbidity patterns associated with lower limb muscle loss in China [15]. However, there is a lack of data to explore the relationship between older adults with low CC and multimorbidity, which is essential for the management of healthy aging.

Objectives

Our study aimed to investigate the epidemiology of older adults with low CC and multimorbidity in the Chinese community, including potential multimorbidity categories, and to explore related risk factors, thereby establishing the theoretical foundation for implementing prevention and management measures to achieve healthy aging in the older adult population with low CC and multimorbidity.

Methods

Participants

The Chinese Longitudinal Healthy Longevity Survey (CLHLS) is a comprehensive and representative longitudinal study on the health and lifespan of older people in China. Following a baseline survey conducted in 1998, the CLHLS conducted 7 waves of surveys every 2 to 3 years from 2000 to 2018, covering 22 of China's 31 provinces and approximately 85% of the population. The study used the CLHLS 2018 data, which included a total of 15,874 older adult participants. This is in accordance with the design of our study, includes an effective sample size, and reflects the most recent health behaviors of older adults in China.

Ethical Considerations

The Peking University Research Ethics Committee authorized the CLHLS study (IRB00001052-13074), and all participants or proxy respondents gave written informed consent. Each respondent who agreed to participate in the survey signed a consent form confirming that they were aware of the objective of the research. All processes were executed following the relevant rules and standards. The CLHLS data used in our study were completely anonymized and deidentified. All personal identifiers were removed from the dataset and replaced with unique identification codes that cannot be traced back to individuals. The dataset contains no personally identifiable information, such as names, detailed addresses, or ID numbers. According to Article 32 of the Ethical Review Measures for Life Science and Medical Research Involving Human Beings issued by China in 2023, scientific research based on legally obtained anonymized public data, which does not involve biological samples and does not violate the commercial interests of others, can be exempted from ethical review. As the CLHLS followed strict protocols to protect participant confidentiality and privacy, no second ethical review was conducted for our study, and no additional informed consent was obtained due to the anonymity of the participants. Further, based on available public documentation about the CLHLS, there is no explicit information regarding specific compensation provided to participants. The survey primarily focused on collecting data about older adult health and living conditions to support academic research and policy development.

Calf Circumference

CC (measured in centimeters) was rounded up to an integer in the CLHLS 2018 wave. Trained investigators measured CC with participants seated, with their hips and knees bent at 90° and their feet flat on the floor. A nonelastic tape was placed around the calf without compressing the tissue and moved along its length to identify the maximum circumference. Each leg was measured twice, and the average was recorded. The final value was the mean of both legs [16]. According to the content of the 2018 questionnaire, CC was rounded to the whole number and entered into the dataset. All investigators were strictly trained and followed standard operating procedures. Additional quality control details are available on the CLHLS website. The diagnosis of our study adopted the cutoff of low CC advised by Asian Working Group for Sarcopenia 2019: CC<34 cm for men

XSL•FO
RenderX

and CC<33 cm for women [17]. These cutoffs were derived from multicenter cohorts of community-dwelling older adults in Asia to support standardized sarcopenia screening in primary care and public health.

Multimorbidity

Every individual was asked if they had experienced any of the following 17 chronic diseases: hypertension, heart disease, bronchitis, emphysema, pneumonia or asthma, tuberculosis, gastric or duodenal ulcer, dyslipidemia, cholecystitis or cholelith disease, chronic nephritis, hepatitis, stroke or cardiovascular disease (CVD), dementia, epilepsy, Parkinson disease, arthritis, rheumatism or rheumatoid disease, cancer, and diabetes. The 17 chronic diseases were selected from the CLHLS questionnaire by excluding sex-specific conditions and retaining those most representative of chronic disease patterns among older Chinese adults, based on previous literature [18,19]. The co-occurrence of at least 2 chronic diseases in 1 person is referred to as multimorbidity [20].

Multidimensional Risk Factors

On the basis of certifications from previous studies [21,22], we identified relevant risk variables that represent the characteristics of China's older adult population. Demographic data included age, sex, place of residence, marital status, and educational level. Behavioral characteristics consisted of current smoking status, current alcohol consumption, physical activity, and sleep quality. No physical activity implies that the participants do not engage in hobbies such as reading, watching television, listening to the radio, or making friends, among other things. Physical and psychological health characteristics included scores from the Mini-Mental State Examination (0 - 30 points) for the cognitive function test, Center for Epidemiological Studies-Depression Scale 10-item (CES-D-10; 0 - 30 points), and Generalized Anxiety Disorder 7-item (GAD-7; 0 - 21 points) in the past 2 weeks; additional measures included fall experience in the past year, BMI, self-reported health, and difficulty with activities of daily living (ADLs). ADLs included 6 items: bathing, dressing, simple indoor activities (getting in and out of bed or sitting up), eating, using the toilet, and controlling urination and defecation. Difficulty with at least 1 of the abovementioned items was classified as a *yes*. All variables were collected by trained interviewers through standardized questionnaires as part of the CLHLS survey, ensuring data quality and consistency.

Statistical Analysis

Descriptive Statistics

The normality of continuous variables was checked using the Kolmogorov-Smirnov test. Continuous variables with normal distributions were represented by means and SDs, while those with abnormal distributions were represented by medians and IQRs. Categorical variables and frequencies (%) were used. Two-tailed independent sample *t* tests, Mann–Whitney U tests, and chi-square tests were used, as appropriate, to compare the baseline characteristics between participants with multimorbidity and those who were relatively healthy. Moreover, we used the MissForest algorithm for missing value imputation. MissForest effectively imputes missing values in multidimensional and mixed continuous and categorical data. This method has lower imputation errors and no parameter adjustment or data distribution assumptions [23].

Latent Class Analysis

Latent class analysis (LCA) has been widely used in health behavior research to categorize latent features based on apparent behaviors and identify the proportion of each class to execute intervention approaches for each subclass. The 17 types of multimorbidity were answered with a yes or no response (1 or 0). The Akaike information criterion (AIC), Bayesian information criterion (BIC), adjusted BIC (aBIC), Lo-Mendell-Rubin, bootstrapped likelihood ratio test, and entropy model fit indices were used to find the best LCA model. The AIC imposes a lighter penalty on model complexity, tending to favor more complex models. In contrast, the BIC uses a stricter penalty and is generally regarded as more robust for determining the number of classes. The aBIC strikes a balance between AIC and BIC, providing an intermediate penalty that is particularly suitable for balanced model selection with moderate sample sizes. The Lo-Mendell-Rubin and bootstrapped likelihood ratio test statistically evaluate whether the current model demonstrates significant improvement over a model with one fewer class. A statistically significant *P* value from these tests indicates that the current model exhibits a substantially better fit compared to the simpler alternative model. An entropy value >0.8 indicates that 90% of individuals are correctly identified, whereas a value of 0.6 suggests that >20% are incorrectly classified. The best-fitting model is chosen based on a thorough evaluation of the aforementioned indicators and clinical interpretability.

Elastic Net

The elastic net (EN), proposed by Zou and Hastie [24], can eliminate characteristics unrelated to the dependent variable, such as least absolute shrinkage and selection operator regression, and choose a critical feature, such as ridge regression. By adjusting the parameters to balance the strength of the L1 and L2 regularization, an ideal model can be obtained. In this study, we used the EN (*glmnet* [25] and *ggplot2* [26] R packages) for feature selection and optimized the key parameters. A 10-fold cross-validation was used to prevent overfitting. The risk factor coefficients that were not closely related to multimorbidity were excluded by compressing them to 0, and the closely related possible risk factors that were not 0 were retained. The best and strict risk variables generally choose the lambda.1se model.

Multiple Logistic Regression

On the basis of the results of LCA and EN, multiple logistic regression (MLR) was used to identify the risk factors associated with different multimorbidity groups in the population with low CC. The statistical analysis was conducted using R (version 4.1.1; R Foundation for Statistical Computing) and Mplus (version 7.4; Muthén & Muthén). The regression study yielded data in the form of odds ratios accompanied by 95% CIs. The statistical significance criterion was determined using 2-sided *t* tests and a *P* value of .05.

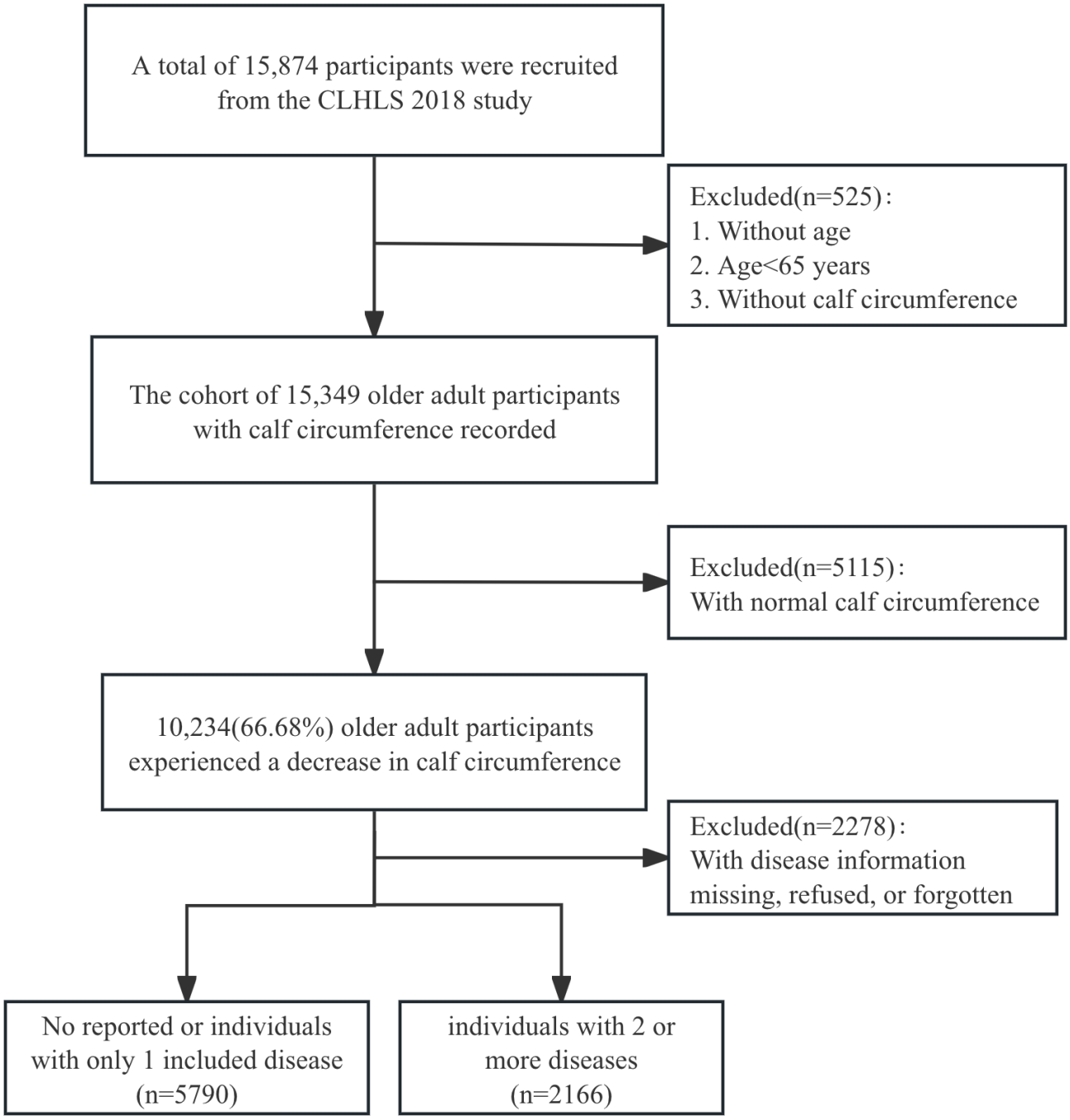
XSL•FO
RenderX

Results

Descriptive Characteristics

In this study, we included 7956 older adults with low CC (Multimedia Appendix 1; the detailed screening flow chart is shown in Figure 1). Their average age was 89 (range 79-99) years, with 49.77% (3960/7956) aged >90 years, and 39.19% (2269/5790) were men. Compared to the relatively healthy group, the multimorbidity group was younger, more likely to live in urban areas, more often partnered, and better educated. The multimorbidity group was likely to have a history of smoking and drinking habits (little or always). Over half of the people with multimorbidity dislike engaging in physical activities and sleep fairly or badly. Scores on the Mini-Mental State Examination, CES-D-10, GAD-7, and BMI were notably higher in the relatively healthy group. They may have greater difficulty with ADLs and overall health conditions. The proportions of distinct prevalent chronic diseases are shown in Table 1. The prevalence of multimorbidity increased between the ages of 65 and 89 years. However, after the age of 90 years, the proportion of multimorbidity decreased, and the majority of older adults were relatively healthy (Figure 2).

Figure 1. The flowchart of study design. CLHLS: Chinese Longitudinal Healthy Longevity Survey.

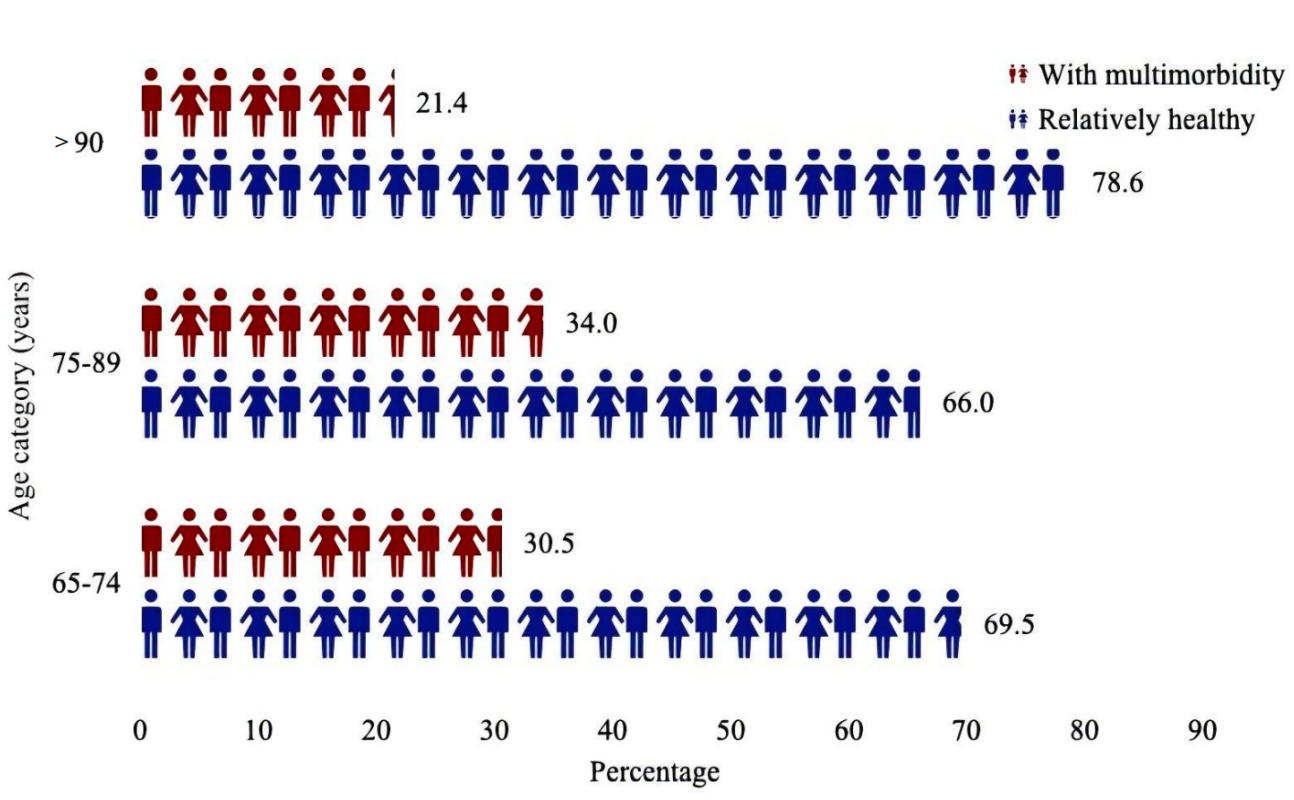


XSL•FO
RenderX

Table . The proportion of multimorbidity among older adult patients with low calf circumference (CC; N=7956).

Chronic diseases	Low CC, n (%)	Low CC and multimorbidity (n=2166), n (%)	Low CC and relatively healthy (n=5790), n (%)	*P* value
Hypertension	2787 (35.0)	1556 (71.8)	1231 (21.3)	<.001
Heart disease	1109 (13.9)	883 (40.8)	226 (3.9)	<.001
Bronchitis, emphysema, pneumonia, or asthma	819 (10.3)	506 (23.4)	313 (5.4)	<.001
Tuberculosis	61 (0.8)	49 (2.3)	12 (0.2)	<.001
Gastric or duodenal ulcer	384 (4.8)	292 (13.5)	92 (1.6)	<.001
Dyslipidemia	251 (3.2)	234 (10.8)	17 (0.3)	<.001
Cholecystitis or cholelith disease	282 (3.5)	224 (10.3)	58 (1.0)	<.001
Chronic nephritis	79 (1.0)	79 (3.7)	0 (0.0)	<.001
Hepatitis	27 (0.3)	21 (1.0)	6 (0.1)	<.001
Stroke or cardiovascular disease	735 (9.2)	576 (26.6)	159 (2.8)	<.001
Dementia	210 (2.6)	128 (5.9)	82 (1.4)	<.001
Epilepsy	22 (0.3)	20 (0.9)	2 (0.0)	<.001
Parkinson disease	58 (0.7)	45 (2.1)	13 (0.2)	<.001
Arthritis	735 (9.2)	590 (27.2)	145 (2.5)	<.001
Rheumatism or rheumatoid disease	360 (4.5)	294 (13.6)	66 (1.1)	<.001
Cancer	100 (1.3)	81 (3.7)	19 (0.3)	<.001
Diabetes	532 (6.7)	449 (20.7)	83 (1.4)	<.001

Figure 2. The distribution of multimorbidity in different age groups.

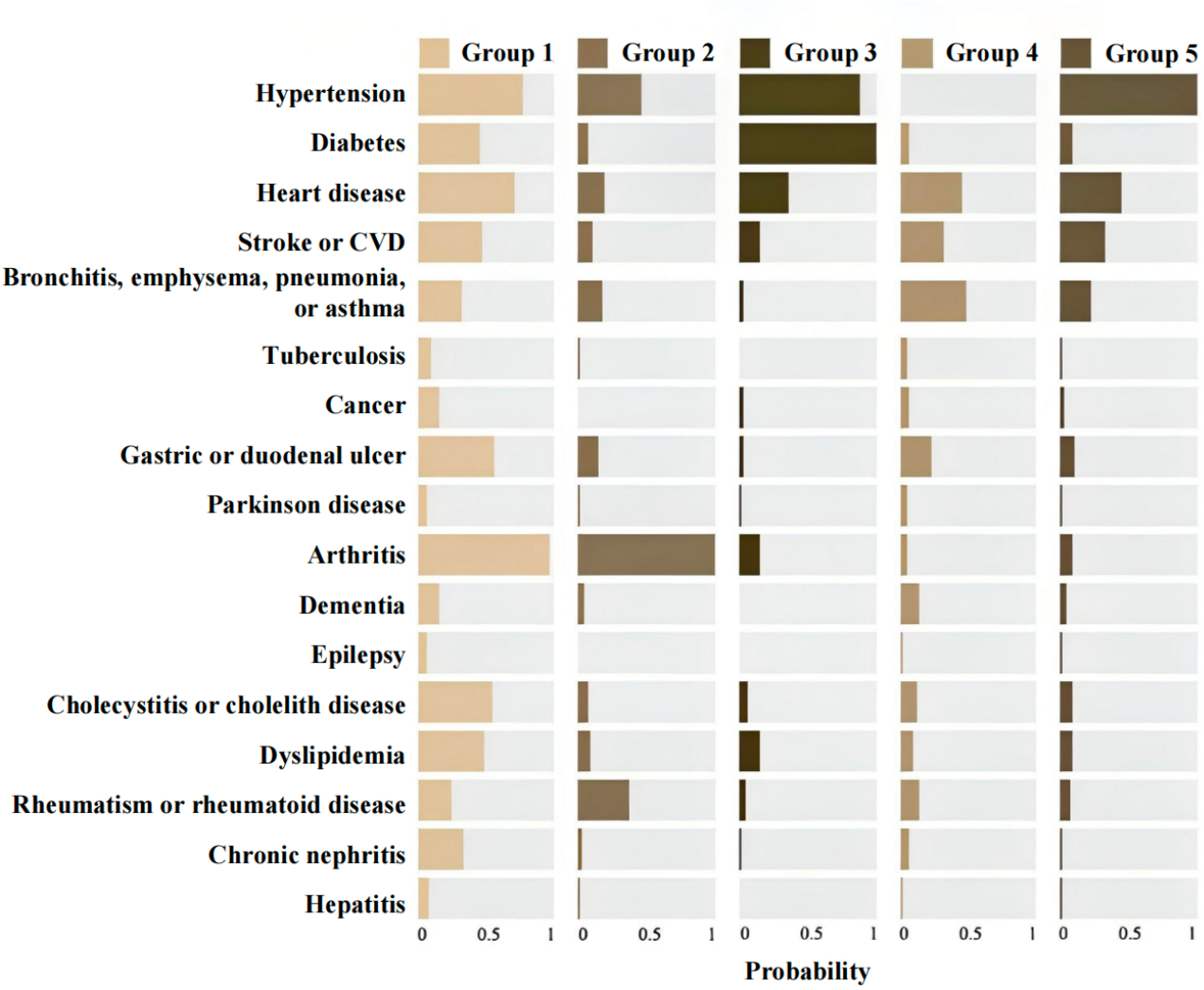


XSL•FO
RenderX

Latent Class Analysis

Multimedia Appendix 2 compares fit statistics for models with 1 to 8 classes. AIC, BIC, and aBIC were low for class 5. Entropy (class 5=0.9) >0.8 suggests 90% correct categorization. In Figure 3, the membership probability is the "conditional prevalence" for each maintained class. We chose class 5 and named groups based on a combination of diseases and clinical evidence: multisystem morbidity diseases (group 1: 78/2166, 3.6%), arthritis-rheumatism or rheumatoid diseases (group 2: 400/2166, 18.47%), diabetes-hypertension diseases (group 3: 330/2166, 15.24%), respiratory-heart diseases (group 4: 347/2166, 16.02%), and CVDs (group 5: 1011/2166, 46.67%).

Figure 3. Cluster-specific conditional probabilities for the specific subgroups of older adults with low calf circumference and multimorbidity across 5 latent class models. CVD: cardiovascular disease.

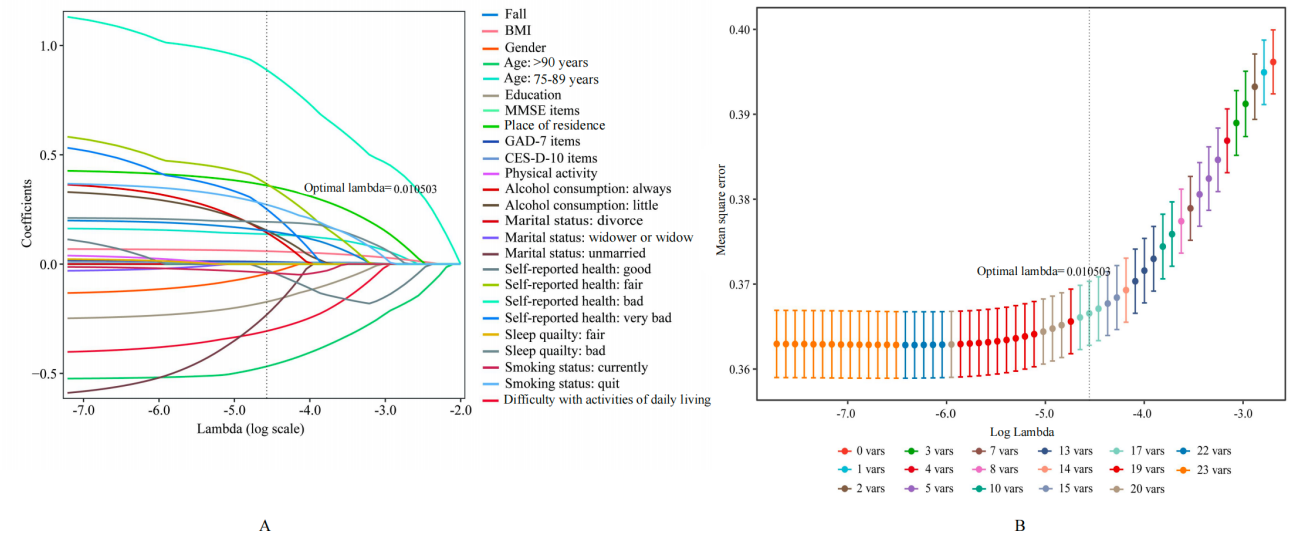


Elastic Net

The EN analysis included 25 variables, and the multicategorical variables were transformed using one-hot encoding (Figure 4A shows the process of variable selection as the penalty term increased). The best model was constructed using the 10-fold cross-validation approach, with α=.5 and λ with an SD of 0.010503 as the ideal parameters (Figure 4B shows the cross-validation plot for the penalty term). Finally, 12 strong risk variables and their association coefficients (Table 2) were obtained, including age, place of residence, education, difficulty with ADLs, smoking, current alcohol consumption, self-reported health, sleep quality, fall, BMI, CES-D-10, and GAD-7.

XSL•FO
RenderX

Figure 4. Plots for elastic net coefficients over different values of the penalty parameter. CES-D-10: Center for Epidemiological Studies-Depression Scale 10-item; GAD-7: Generalized Anxiety Disorder 7-item; MMSE: Mini-Mental State Examination.



A B

Table . The coefficient in the filter variables by the elastic net.

Variables	Coefficient
Age: >90 years	–0.43
Difficulty with activities of daily living	–0.25
Education	–0.10
Self-reported health: good	–0.09
Smoking status: currently	–0.03
CES-D-10[a] items	0.00
GAD-7[b] items	0.01
Current alcohol consumption: always	0.02
BMI	0.05
Self-reported health: very bad	0.07
Age: 75-89 years	0.10
Fall	0.11
Sleep quality: bad	0.18
Smoking status: quit	0.20
Self-reported health: fair	0.25
Place of residence	0.31
Self-reported health: bad	0.77

[a]CES-D-10: Center for Epidemiological Studies-Depression Scale 10-item.

[b]GAD-7: Generalized Anxiety Disorder 7-item.

Multidimensional Risk Factors for Differences

The MLR identified risk variables for 5 latent categories, using the relatively healthy group as the reference class (Multimedia Appendix 3).

Demographic Data

The older adults aged 75 to 89 years were more likely to have CVDs than the relatively healthy group, whereas those aged >90 years were healthier. Urbanites had higher rates of multimorbidity, notably in the multisystem morbidity, diabetes-hypertension, CVDs, and respiratory-heart disease groups. Multisystem diseases, diabetes-hypertension, and CVDs are more common among educated older adults.

Behavioral Characteristics

Smoking history may enhance the chance of an older adult being diagnosed with respiratory-heart disease and CVDs. Current smoking does not appear to be a multimorbidity risk factor. Little and often alcohol use may be detrimental for

XSL•FO
RenderX

diabetes-hypertension, CVDs, and respiratory-heart disease. Arthritis-rheumatism and CVDs were more common in older adults with poor sleep.

Physical and Psychological Health Characteristics

According to the MLR, multisystem morbidity increased the risk of depression. Assessment using the GAD-7, which evaluates anxiety symptoms over the preceding 2 weeks, indicated that the CVDs group may have had higher anxiety levels than the healthy group in the past 2 weeks. A higher BMI is a potential risk factor for falling across all multimorbidity groups, excluding respiratory-heart disease. Older adults with low CC who fell in the last year are more than twice as likely to develop arthritis-rheumatism or rheumatoid illness. The diabetes-hypertension, CVDs, and respiratory-heart disease groups were more likely to include older adults with ADL difficulties. The multimorbidity groups self-reported their health as fair, bad, or very bad.

Discussion

Principal Findings

In China, we found that low CC is a common phenomenon in community-dwelling older adults. LCA categorized these older patients into 5 groups: multisystem morbidity diseases, arthritis-rheumatism or rheumatoid diseases, diabetes-hypertension diseases, respiratory-heart diseases, and CVDs. MLR used 12 reliable risk variables established by EN and revealed that older adults with low CC and multimorbidity exhibited both comparable and dissimilar demographic, behavioral, and physical and psychological health characteristics.

Sarcopenia is a common geriatric syndrome characterized by the progressive loss of skeletal muscle mass and strength. It can be classified into primary sarcopenia, which is age related, and secondary sarcopenia, which is caused by factors such as malnutrition, physical inactivity, and chronic inflammation. The pathophysiology of sarcopenia involves disruptions in protein metabolism, neuromuscular degeneration, and abnormal fat infiltration, all of which are closely linked to the aging process [27,28]. Previous studies have demonstrated that CC, as a simple and feasible proxy for muscle mass, shows a declining trend with advancing age, with a notable inflection point occurring around 65 years. This pattern aligns with the natural progression of sarcopenia. Moreover, research suggests that calf muscle circumference, which was calculated by subtracting the calf subcutaneous fat thickness from the total CC, exhibits a stronger correlation with skeletal muscle mass, indicating its potential value as a more precise marker for early detection of sarcopenia. Given its ease of measurement in community and primary care settings, CC deserves further consideration as a practical screening tool. In addition, muscle atrophy associated with sarcopenia often coexists with bone mineral density loss and changes in lower limb morphology, which may indirectly influence CC measurements. Therefore, future studies should further investigate the multidimensional associations between CC, muscle mass, bone health, and functional status in older adults. In addition, there are chronic diseases such as arthritis, diabetes, varicose veins, neuromuscular diseases, and other diseases that can affect the CC. Prior research has established a correlation between low CC and chronic conditions such as hypertension and diabetes, as well as overall mortality [29,30]. Emerging evidence suggests that mitochondrial dysfunction in T cells may be a central mechanistic driver of multimorbidity. T cells exhibiting mitochondrial homeostatic imbalance synergistically exacerbate multiorgan pathophysiological cascades through chronic inflammatory activation, dysregulated metabolic interplay, and propagation of senescence-associated secretory phenotype signaling [31].

Our study identified CVDs as the most prevalent comorbidity pattern, providing a theoretical foundation for optimizing preventive care and therapeutic strategies. The marked disparity in CVD-related mortality between rural and urban areas in China (46.7% vs 44.3%) underscores the necessity for targeted implementation of mobile health programs in rural regions with limited health care resources [32]. The findings revealed that CC <33 cm is a risk factor for stroke and predisposes women to a higher risk of carotid plaque development [33]. After adjusting for potential confounders, a significant association between low CC and hypertension was identified. The underlying etiology of low CC may involve reduced capillary density and elevated resting heart rate [34]. Elevated heart rate is a consequential result of reduced stroke volume and reduced ability to tolerate exercise, which heightens the likelihood of cardiovascular incidents and perhaps leads to chronic hypertension. Integrating routine CC measurements into primary care screenings, coupled with patient education to enhance self-assessment awareness of modifiable risks such as resting heart rate and exercise tolerance, could mitigate the progression of CVDs in both populations.

Arthritis-rheumatism or rheumatoid disease is second only to CVDs as a comorbid pattern, which is consistent with the results of previous research [35]. Muscle atrophy is a common problem in joint diseases, such as osteoarthritis, and can also lead to articular muscle inhibition. Abnormal joint afferent nerves trigger the release of neurotransmitters at the spinal cord level, inhibiting the activity of α-motor neurons and resulting in decreased muscle activity, which leads to secondary muscle atrophy. A decrease in muscle measurement indicators, such as low CC, reflects reduced muscle cross-sectional area and increased fat infiltration, indicating muscle atrophy [36].

In this study, the combination of respiratory-heart disease was identified as the third most common comorbidity. Respiratory conditions such as chronic obstructive pulmonary disease or asthma can cause chronic hypoxemia, leading to pulmonary hypertension, increased right ventricular load, and eventually right heart dysfunction. Hypoxia also activates the sympathetic nervous system, elevating heart rate and blood pressure, which may further strain the cardiovascular system—potentially explaining this comorbidity pattern. Hypertension and diabetes commonly co-occur, largely due to overlapping pathophysiological mechanisms. Insulin resistance activates the sympathetic nervous system and the renin-angiotensin-aldosterone system, contributing to elevated blood pressure. Hyperglycemia accelerates endothelial dysfunction and atherosclerosis. Chronic inflammation,

XSL•FO
RenderX

oxidative stress, and metabolic disturbances also play key roles in the shared development of both conditions. These comorbidity patterns have been established in separate investigations [37,38]. Our study sample consisted of older adults with low CC, which was a major source of discrepancy in comorbid patterns compared to other studies.

Our results confirmed the findings of prior studies, demonstrating that the prevalence of comorbidity was higher among younger older individuals than among those who lived longer [12]. One important source of potential bias is selection bias, particularly survivorship bias, which refers to the fact that individuals who reach exceptional longevity (eg, age ≥90 y) are more likely to possess inherent physiological resilience or to have consistently engaged in health-promoting behaviors, such as lifelong physical activity and balanced nutrition. Consequently, this population often appears as a cohort with low prevalence of comorbidities in observational studies. However, this phenomenon underscores a fundamental methodological limitation in aging research—namely, that the survival process inherently filters out individuals with greater vulnerability to chronic diseases, thereby potentially underestimating the true burden of morbidity in the general older adult population. Another factor that might explain this phenomenon is remembering bias: they are more prone to having both cognitive impairment and chronic diseases and may not accurately record their chronic conditions. There is a notable disparity in the prevalence of chronic diseases among older adults living in urban and rural areas in China, largely reflecting the unequal distribution of health care resources [39]. These differences may also arise from variations in lifestyle, health care access, and health awareness. Urban residents generally have better access to medical services but are more likely to engage in sedentary behavior and consume high-fat diets, increasing their risk of chronic conditions. In contrast, rural residents tend to engage in more physical activity but often face limited access to timely diagnosis and treatment. Socioeconomic and educational inequalities further contribute to these urban-rural disparities [12,40].

Smoking is widely recognized as an independent behavioral risk factor for respiratory diseases and CVDs [41]. In our cross-sectional analysis, we observed that personal smoking and drinking history were associated with increased likelihood of coexisting respiratory-heart diseases, CVDs, and diabetes-hypertension patterns in older adults. While these associations cannot establish causal relationships, they highlight potential modifiable lifestyle factors that warrant attention in older adults.

Rheumatoid arthritis, often accompanied by joint pain, swelling, and stiffness, may contribute to sleep disturbances, which are also commonly reported in older adults with multimorbidity. Moreover, both the disease itself and the side effects of medications such as nonsteroidal anti-inflammatory drugs and immunosuppressants may impair sleep quality [42]. Similarly, in individuals with CVDs (eg, coronary heart disease), nocturnal chest discomfort and medication use (eg, diuretics and beta-blockers) may further disrupt sleep. Importantly, joint dysfunction may lead to decreased physical activity and subsequent muscle disuse, potentially contributing to low CC. Conversely, muscle weakness related to sarcopenia may exacerbate joint instability and functional limitations [43]. Given the potential bidirectional relationship between them, clinicians should pay close attention to musculoskeletal function in older adults with joint diseases. Early identification of decreased lower limb muscle mass in patients with arthritis may help prevent further functional decline.

Regarding physical and psychological health characteristics, one study confirmed that people with multiple illnesses were 2 to 3 times more prone to developing depression than those without multimorbidity. Our findings indicate that older adults with multimorbidity are more likely to report psychological distress, including depressive and anxiety symptoms, consistent with prior cross-sectional and longitudinal studies [44]. A survey reported that the severity of anxiety symptoms is linked to a higher risk of cardiovascular problems [45]. Although the directionality of these associations remains uncertain, the co-occurrence of mental and physical health conditions underscores the need for integrated care. Obesity is a common risk factor for several diseases, and an observational study found that a higher BMI is one of the risk factors for complex and frequent morbidity [46]. Arthritis-rheumatoid joint diseases may cause calf muscle atrophy and lower limb muscle weakness, increasing the likelihood of falls [47]. Consistent with the existing evidence, loss of ADLs is a risk factor for multiple morbidities in older adults [48]. Higher body weight, joint disorders, and reduced physical function suggest that clinical efforts in older adults should focus on weight management, appropriate rehabilitation or strength training to maintain musculoskeletal health, and regular functional assessments. These measures may help identify high-risk individuals early and facilitate timely interventions to improve overall health outcomes.

Self-rated health, as a simple and cost-effective measure, was associated with multimorbidity in our study. Older adults with negative perceptions of their health may be more likely to present with multiple chronic conditions. Studies conducted in several countries have similarly supported this finding, showing that patients with 3 or 4 disorders are significantly more likely to report poor health [49]. This highlights the potential value of incorporating self-rated health into routine assessments to help identify individuals at higher health risk.

Strengths and Limitations

Our study has several strengths. First, our results are more applicable because we used a sizable cohort of long-lived Chinese citizens, which is nationally representative. Second, the EN develops a variable model specifically suited to this population, identifying significant risk factors. Third, due to the potential existence of mixed factors in the study population, we did not simply use the disease classification method of combining 2, 3, or more diseases; instead, we selected LCA for analysis to obtain different representative comorbidity groups. More significantly, this is the first study to look at risk factors for multimorbidity classifications in older adults with low CC, which has significant ramifications for clinical practice and health policy.

XSL•FO
RenderX

Our study also has some limitations. Although it was not feasible to draw firm conclusions about causality from the cross-sectional study, EN allowed us to develop a trustworthy risk model. According to studies, CC measurements in individuals with lower limb edema yielded erroneous findings, and some patients with heart failure were excluded from the study due to their swollen calves. Noncommunicable illnesses are currently China's top cause of death for older adults. Some infectious illnesses have high fatality rates, which might be the reason we did not find any evidence of their prevalence.

Conclusions

This study found that individuals with low CC and multimorbidity constituted a large group that cannot be ignored in the older adult population of the Chinese community, and potential risk factors were revealed in different combinations of multimorbidity. These findings underscore the imperative to operationalize precision public health frameworks—integrating geriatric biomarkers such as low CC into multidisciplinary care models—to mitigate adverse outcomes in high-risk older adults. Public health efforts should prioritize community-based interventions such as chronic disease self-management, coordinated primary care, and preventive lifestyle support. Future research is needed to evaluate the effectiveness and scalability of such approaches to reduce the burden of multimorbidity and promote healthy aging.

Acknowledgments

We appreciate all Chinese Health Longevity Longitudinal Survey (CLHLS) survey participants, recorders, and the CLHLS team for supplying the data. The authors of this manuscript certify that they comply with the ethical guidelines for authorship and publishing in *JMIR Aging*. This study was supported by the Noncommunicable Chronic Diseases-National Science and Technology Major Project (2023ZD0501801 to RZ and GC), the Sichuan Province Health Research Project (ZH2025-103 to TZ), Key Project of Educational Teaching Reform Research in Sichuan Province (YJGXM24-B029, GSSCU2024027 to GC), and the CAMS Innovation Fund for Medical Sciences (2023-I2M-C&T-B-122 to TZ, 2023-I2M-C&T-B-121 to GC).

Authors' Contributions

XP and XC conceived the project and planned the investigation. XP got data access. XP, FS, and XC conducted statistical analyses and result interpretations. XP drafted the first version of the manuscript. XC, RZ, GC, and TZ revised it for significant intellectual substance. All authors reviewed the final draft and gave their approval. TZ served as the co-corresponding author for this study and contributed equally to the supervision of research. TZ can be reached via email at 739501155@qq.com.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Characteristics of older adults with low calf circumference and multimorbidity.
[DOCX File, 23 KB - aging_v8i1e68760_app1.docx]

Multimedia Appendix 2
Comparison of model fit statistics for latent class analysis across 1 to 8 classes.
[DOCX File, 14 KB - aging_v8i1e68760_app2.docx]

Multimedia Appendix 3
Multidimensional risk factors associated with different groups.
[DOCX File, 25 KB - aging_v8i1e68760_app3.docx]

References

1. Cruz-Jentoft AJ, Landi F, Topinková E, Michel JP. Understanding sarcopenia as a geriatric syndrome. Curr Opin Clin Nutr Metab Care 2010 Jan;13(1):1-7. [doi: 10.1097/MCO.0b013e328333c1c1] [Medline: 19915458]
2. Papadopoulou SK. Sarcopenia: a contemporary health problem among older adult populations. Nutrients 2020 May 1;12(5):1293. [doi: 10.3390/nu12051293] [Medline: 32370051]
3. Hwang AC, Liu LK, Lee WJ, Peng LN, Chen LK. Calf circumference as a screening instrument for appendicular muscle mass measurement. J Am Med Dir Assoc 2018 Feb;19(2):182-184. [doi: 10.1016/j.jamda.2017.11.016] [Medline: 29306606]
4. Wang X, Yang F, Zhu M, et al. Development and assessment of assisted diagnosis models using machine learning for identifying elderly patients with malnutrition: cohort study. J Med Internet Res 2023 Mar 14;25:e42435. [doi: 10.2196/42435] [Medline: 36917167]

XSL•FO
RenderX

5. Borges K, Artacho R, Jodar-Graus R, Molina-Montes E, Ruiz-López MD. Calf circumference, a valuable tool to predict sarcopenia in older people hospitalized with hip fracture. Nutrients 2022 Oct 12;14(20):4255. [doi: 10.3390/nu14204255] [Medline: 36296937]
6. Wu SE, Chen WL. Calf circumference refines sarcopenia in correlating with mortality risk. Age Ageing 2022 Feb 2;51(2):afab239. [doi: 10.1093/ageing/afab239] [Medline: 35134846]
7. Pérez-Zepeda MU, Gutiérrez-Robledo LM. Calf circumference predicts mobility disability: a secondary analysis of the Mexican health and ageing study. Eur Geriatr Med 2016 Jun;7(3):262-266. [doi: 10.1016/j.eurger.2016.01.004] [Medline: 27656259]
8. Tsai AC, Chang TL. The effectiveness of BMI, calf circumference and mid-arm circumference in predicting subsequent mortality risk in elderly Taiwanese. Br J Nutr 2011 Jan;105(2):275-281. [doi: 10.1017/S0007114510003429] [Medline: 21129232]
9. Wu CJ, Kao TW, Lin YY, et al. Examining the association between anthropometric parameters and telomere length and mortality risk. Oncotarget 2017 May 23;8(21):34057-34069. [doi: 10.18632/oncotarget.15976] [Medline: 28423661]
10. Bellantuono I. Find drugs that delay many diseases of old age. Nature New Biol 2018 Feb;554(7692):293-295. [doi: 10.1038/d41586-018-01668-0] [Medline: 32094723]
11. Sum G, Hone T, Atun R, et al. Multimorbidity and out-of-pocket expenditure on medicines: a systematic review. BMJ Glob Health 2018;3(1):e000505. [doi: 10.1136/bmjgh-2017-000505] [Medline: 29564155]
12. Yao SS, Cao GY, Han L, et al. Prevalence and patterns of multimorbidity in a nationally representative sample of older Chinese: results from the China health and retirement longitudinal study. J Gerontol A Biol Sci Med Sci 2020 Sep 25;75(10):1974-1980. [doi: 10.1093/gerona/glz185] [Medline: 31406983]
13. Olson JE, Takahashi PY, St Sauver JM. Understanding the patterns of multimorbidity. Mayo Clin Proc 2018 Jul;93(7):824-825. [doi: 10.1016/j.mayocp.2018.05.016] [Medline: 29976373]
14. Dodds RM, Granic A, Robinson SM, Sayer AA. Sarcopenia, long-term conditions, and multimorbidity: findings from UK Biobank participants. J Cachexia Sarcopenia Muscle 2020 Feb;11(1):62-68. [doi: 10.1002/jcsm.12503] [Medline: 31886632]
15. Yang K, Yang S, Chen Y, et al. Multimorbidity patterns and associations with gait, balance and lower extremity muscle function in the elderly: a cross-sectional study in Northwest China. Int J Gen Med 2023;16:3179-3192. [doi: 10.2147/IJGM.S418015] [Medline: 37533839]
16. Kawakami R, Murakami H, Sanada K, et al. Calf circumference as a surrogate marker of muscle mass for diagnosing sarcopenia in Japanese men and women. Geriatr Gerontol Int 2015 Aug;15(8):969-976. [doi: 10.1111/ggi.12377] [Medline: 25243821]
17. Chen LK, Woo J, Assantachai P, et al. Asian Working Group for Sarcopenia: 2019 consensus update on sarcopenia diagnosis and treatment. J Am Med Dir Assoc 2020 Mar;21(3):300-307. [doi: 10.1016/j.jamda.2019.12.012] [Medline: 32033882]
18. Chen ZT, Wang XM, Zhong YS, Zhong WF, Song WQ, Wu XB. Association of changes in waist circumference, waist-to-height ratio and weight-adjusted-waist index with multimorbidity among older Chinese adults: results from the Chinese longitudinal healthy longevity survey (CLHLS). BMC Public Health 2024 Jan 29;24(1):318. [doi: 10.1186/s12889-024-17846-x] [Medline: 38287292]
19. Sun X, Liu X, Wang X, Pang C, Yin Z, Zang S. Association between residential proximity to major roadways and chronic multimorbidity among Chinese older adults: a nationwide cross-sectional study. BMC Geriatr 2024 Jan 29;24(1):111. [doi: 10.1186/s12877-024-04712-z] [Medline: 38287240]
20. Skou ST, Mair FS, Fortin M, et al. Multimorbidity. Nat Rev Dis Primers 2022 Jul 14;8(1):48. [doi: 10.1038/s41572-022-00376-4] [Medline: 35835758]
21. Yuan S, Larsson SC. Epidemiology of sarcopenia: prevalence, risk factors, and consequences. Metab Clin Exp 2023 Jul;144:155533. [doi: 10.1016/j.metabol.2023.155533] [Medline: 36907247]
22. Cheng L, Sit JWH, Chan HYL, et al. Sarcopenia risk and associated factors among Chinese community-dwelling older adults living alone. Sci Rep 2021 Nov 15;11(1):22219. [doi: 10.1038/s41598-021-01614-7] [Medline: 34782685]
23. Stekhoven DJ, Bühlmann P. MissForest--non-parametric missing value imputation for mixed-type data. Bioinformatics 2012 Jan 1;28(1):112-118. [doi: 10.1093/bioinformatics/btr597] [Medline: 22039212]
24. Zou H, Hastie T. Regularization and variable selection via the elastic net. J R Stat Soc Series B Stat Methodol 2005 Apr 1;67(2):301-320. [doi: 10.1111/j.1467-9868.2005.00503.x]
25. Engebretsen S, Bohlin J. Statistical predictions with glmnet. Clin Epigenetics 2019 Aug 23;11(1):123. [doi: 10.1186/s13148-019-0730-1] [Medline: 31443682]
26. Ginestet C. ggplot2: elegant graphics for data analysis. J R Stat Soc Ser A Stat Soc 2011 Jan 1;174(1):245-246. [doi: 10.1111/j.1467-985X.2010.00676_9.x]
27. Pedersen BK, Febbraio MA. Muscles, exercise and obesity: skeletal muscle as a secretory organ. Nat Rev Endocrinol 2012 Apr 3;8(8):457-465. [doi: 10.1038/nrendo.2012.49] [Medline: 22473333]
28. Nelke C, Dziewas R, Minnerup J, Meuth SG, Ruck T. Skeletal muscle as potential central link between sarcopenia and immune senescence. EBioMedicine 2019 Nov;49:381-388. [doi: 10.1016/j.ebiom.2019.10.034] [Medline: 31662290]

XSL•FO
RenderX

29. Patiño-Villada FA, Estrada-Restrepo A, Aristizábal J. Handgrip strength in older adults from Antioquia-Colombia and comparison of cutoff points for dynapenia. Sci Rep 2023 Jan 31;13(1):1718. [doi: 10.1038/s41598-023-28898-1] [Medline: 36720934]
30. Wang X, Ying Y, Pei M, et al. Calf circumference change and all-cause mortality among community-dwelling Chinese older people. Clin Nutr 2023 Mar;42(3):277-281. [doi: 10.1016/j.clnu.2023.01.002] [Medline: 36724725]
31. Desdín-Micó G, Soto-Heredero G, Aranda JF, et al. T cells with dysfunctional mitochondria induce multimorbidity and premature senescence. Science 2020 Jun 19;368(6497):1371-1376. [doi: 10.1126/science.aax0860] [Medline: 32439659]
32. The Writing Committee Of The Report On Cardiovascular Health And Diseases In China, Hu SS. Report on cardiovascular health and diseases in China 2021: an updated summary. J Geriatr Cardiol 2023 Jun 28;20(6):399-430. [doi: 10.26599/1671-5411.2023.06.001] [Medline: 37416519]
33. Debette S, Leone N, Courbon D, et al. Calf circumference is inversely associated with carotid plaques. Stroke 2008 Nov;39(11):2958-2965. [doi: 10.1161/STROKEAHA.108.520106] [Medline: 18703804]
34. Hedman A, Andersson PE, Reneland R, Lithell HO. Insulin-mediated changes in leg blood flow are coupled to capillary density in skeletal muscle in healthy 70-year-old men. Metab Clin Exp 2001 Sep;50(9):1078-1082. [doi: 10.1053/meta.2001.25604] [Medline: 11555842]
35. Zhang R, Lu Y, Shi L, Zhang S, Chang F. Prevalence and patterns of multimorbidity among the elderly in China: a cross-sectional study using national survey data. BMJ Open 2019 Aug 18;9(8):e024268. [doi: 10.1136/bmjopen-2018-024268] [Medline: 31427309]
36. Wiewiorski M, Dopke K, Steiger C, Valderrabano V. Muscular atrophy of the lower leg in unilateral post traumatic osteoarthritis of the ankle joint. Int Orthop 2012 Oct;36(10):2079-2085. [doi: 10.1007/s00264-012-1594-6] [Medline: 22722542]
37. Zhao X, Zhang Q, Ma C, Liu H, Chen Y. Association between multimorbidity patterns and healthcare costs among middle-aged and older adults in China. Arch Gerontol Geriatr 2023 Jun;109:104959. [doi: 10.1016/j.archger.2023.104959] [Medline: 36804649]
38. Zhang Q, Han X, Zhao X, Wang Y. Multimorbidity patterns and associated factors in older Chinese: results from the China health and retirement longitudinal study. BMC Geriatr 2022 Jun 1;22(1):470. [doi: 10.1186/s12877-022-03154-9] [Medline: 35641904]
39. Zhang X, Dupre ME, Qiu L, Zhou W, Zhao Y, Gu D. Urban-rural differences in the association between access to healthcare and health outcomes among older adults in China. BMC Geriatr 2017 Jul 19;17(1):151. [doi: 10.1186/s12877-017-0538-9] [Medline: 28724355]
40. Bonnell LN, Clifton J, Rose GL, Waddell EN, Littenberg B. Urban-rural differences in mental and physical health among primary care patients with multiple chronic conditions: a secondary analysis from a randomized clinical trial. Int J Environ Res Public Health 2022 Nov 24;19(23):15580. [doi: 10.3390/ijerph192315580] [Medline: 36497657]
41. Freund O, Shetrit A, Bar-Shai A, et al. Smoking and respiratory diseases in patients with coronary microvascular dysfunction. Am J Med 2024 Jun;137(6):538-544. [doi: 10.1016/j.amjmed.2024.02.034] [Medline: 38485108]
42. Fitzcharles MA, Lussier D, Shir Y. Management of chronic arthritis pain in the elderly. Drugs Aging 2010 Jun 1;27(6):471-490. [doi: 10.2165/11536530-000000000-00000] [Medline: 20524707]
43. Prado CM, Landi F, Chew STH, et al. Advances in muscle health and nutrition: a toolkit for healthcare professionals. Clin Nutr 2022 Oct;41(10):2244-2263. [doi: 10.1016/j.clnu.2022.07.041] [Medline: 36081299]
44. Read JR, Sharpe L, Modini M, Dear BF. Multimorbidity and depression: a systematic review and meta-analysis. J Affect Disord 2017 Oct 15;221:36-46. [doi: 10.1016/j.jad.2017.06.009] [Medline: 28628766]
45. van Zutphen EM, Kok AAL, Muller M, et al. Cardiovascular risk indicators among depressed persons: a special case? J Affect Disord 2023 May 15;329:335-342. [doi: 10.1016/j.jad.2023.02.092] [Medline: 36842656]
46. Kivimäki M, Strandberg T, Pentti J, et al. Body-mass index and risk of obesity-related complex multimorbidity: an observational multicohort study. Lancet Diabetes Endocrinol 2022 Apr;10(4):253-263. [doi: 10.1016/S2213-8587(22)00033-X] [Medline: 35248171]
47. Bennett JL, Pratt AG, Dodds R, Sayer AA, Isaacs JD. Rheumatoid sarcopenia: loss of skeletal muscle strength and mass in rheumatoid arthritis. Nat Rev Rheumatol 2023 Apr;19(4):239-251. [doi: 10.1038/s41584-023-00921-9] [Medline: 36801919]
48. Sieber S, Roquet A, Lampraki C, Jopp DS. Multimorbidity and quality of life: the mediating role of ADL, IADL, loneliness, and depressive symptoms. Innov Aging 2023;7(4):igad047. [doi: 10.1093/geroni/igad047] [Medline: 37435089]
49. Ishizaki T, Kobayashi E, Fukaya T, Takahashi Y, Shinkai S, Liang J. Association of physical performance and self-rated health with multimorbidity among older adults: results from a nationwide survey in Japan. Arch Gerontol Geriatr 2019;84:103904. [doi: 10.1016/j.archger.2019.103904] [Medline: 31302504]

Abbreviations

aBIC: adjusted Bayesian information criterion
ADL: activity of daily living

XSL•FO
RenderX

AIC: Akaike information criterion
BIC: Bayesian information criterion
CC: calf circumference
CES-D-10: Center for Epidemiological Studies-Depression Scale 10-item
CLHLS: Chinese Longitudinal Healthy Longevity Survey
CVD: cardiovascular disease
EN: elastic net
GAD-7: Generalized Anxiety Disorder 7-item
LCA: latent class analysis
MLR: multiple logistic regression

Edited by G Seçkin; submitted 13.11.24; peer-reviewed by J Wang, JV Neto, O Afolayan, W Yang; revised version received 18.06.25; accepted 04.09.25; published 02.10.25.

Please cite as:
Peng X, Chen X, Zhou R, Shi F, Zhu T, Chen G
Exploring Risk Factors Related to Low Calf Circumference in Older Adults With Multimorbidity: Cross-Sectional Latent Class Analysis
JMIR Aging 2025;8:e68760
URL: https://aging.jmir.org/2025/1/e68760
doi:10.2196/68760

© Xilin Peng, Xudong Chen, Ruihao Zhou, Fanfan Shi, Tao Zhu, Guo Chen. Originally published in JMIR Aging (https://aging.jmir.org), 2.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using Machine Learning to Predict Cognitive Decline in Older Adults From the Chinese Longitudinal Healthy Longevity Survey: Model Development and Validation Study

Hao Ren[1,2*], PhD; Yiying Zheng[3*], MSc; Changjin Li[2*], MSc; Fengshi Jing[2], PhD; Qiting Wang[4], MSc; Zeyu Luo[2], BSc; Dongxiao Li[5], BSc; Deyi Liang[6], MSc; Weiming Tang[1,7,8], PhD; Li Liu[4], PhD; Weibin Cheng[1,2,9], PhD

[1]Institute for Healthcare Artificial Intelligence Application, The Affiliated Guangdong Second Provincial General Hospital of Jinan University, No. 466 Xingangzhong Road, Haizhu District, Guangzhou, China

[2]Faculty of Data Science, City University of Macau, Macao SAR, China

[3]The Affiliated Traditional Chinese Medicine Hospital, Guangzhou Medical University, Guangzhou, China

[4]School of Public Health, Guangdong Pharmaceutical University, Guangzhou, China

[5]Hainan International College, Minzu University of China, Beijing, China

[6]Guangdong Women and Children Hospital, Guangzhou, China

[7]Institute for Global Health and Infectious Disease, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[8]University of North Carolina at Chapel Hill Project-China, Guangzhou, China

[9]College of Computing, City University of Hong Kong, Hong Kong SAR, China

[*]these authors contributed equally

Corresponding Author:
Weibin Cheng, PhD
Institute for Healthcare Artificial Intelligence Application, The Affiliated Guangdong Second Provincial General Hospital of Jinan University, No. 466 Xingangzhong Road, Haizhu District, Guangzhou, China

Abstract

Background: Cognitive impairment, indicative of Alzheimer disease and other forms of dementia, significantly deteriorates the quality of life of older adult populations and imposes considerable burdens on families and health care systems worldwide. The early identification of individuals at risk for cognitive impairment through a convenient and rapid method is crucial for the timely implementation of interventions.

Objective: The objective of this study was to explore the application of machine learning (ML) to integrate blood biomarkers, life behaviors, and disease history to predict the decline in cognitive function.

Methods: This approach uses data from the Chinese Longitudinal Healthy Longevity Survey. A total of 2688 participants aged 65 years or older from the 2008‐2009, 2011‐2012, and 2014 Chinese Longitudinal Healthy Longevity Survey waves were included, with cognitive impairment defined as a Mini-Mental State Examination (MMSE) score below 18. The dataset was divided into a training set (n=1331), an internal test set (n=333), and a prospective validation set (n=1024). Participants with a baseline MMSE score of less than 18 were excluded from the cohort to ensure a more accurate assessment of cognitive function. We developed ML models that integrate demographic information, health behaviors, disease history, and blood biomarkers to predict cognitive function at the 3-year follow-up point, specifically identifying individuals who are at risk of experiencing significant declines in cognitive function by that time. Specifically, the models aimed to identify individuals who would experience a significant decline in their MMSE scores (less than 18) by the end of the follow-up period. The performance of these models was evaluated using metrics including accuracy, sensitivity, and the area under the receiver operating characteristic curve.

Results: All ML models outperformed the MMSE alone. The balanced random forest achieved the highest accuracy (88.5% in the internal test set and 88.7% in the prospective validation set), albeit with a lower sensitivity, while logistic regression recorded the highest sensitivity. SHAP (Shapley Additive Explanations) analysis identified instrumental activities of daily living, age, and baseline MMSE scores as the most influential predictors for cognitive impairment.

Conclusions: The incorporation of blood biomarkers, along with demographic, life behavior, and disease history into ML models offers a convenient, rapid, and accurate approach for the early identification of older adult individuals at risk of cognitive impairment. This method presents a valuable tool for health care professionals to facilitate timely interventions and underscores the importance of integrating diverse data types in predictive health models.

(JMIR Aging 2025;8:e67437) doi:10.2196/67437

XSL•FO
RenderX

KEYWORDS

older adults; cognitive decline; Alzheimer disease; machine learning; blood biomarkers; disease history; Mini-Mental State Examination; MMSE; Chinese Longitudinal Healthy Longevity Survey; CLHLS

Introduction

Alzheimer disease (AD), the most prevalent form of dementia, is a progressive condition primarily characterized by memory loss [1]. It is estimated that approximately 50 million people worldwide are currently living with AD [2]. This condition not only deteriorates the quality of life for older adult individuals but also imposes significant burdens on families and health care systems, especially as the global population continues to age [3-6]. Mild cognitive impairment (MCI) is recognized as an intermediary stage between normal aging and the more severe cognitive decline observed in dementia. Early detection of MCI through diagnostic tools, such as magnetic resonance imaging (MRI), can facilitate timely interventions aimed at reducing the risk of progression to AD [7]. However, cognitive impairment, which encompasses a broader spectrum of decline, including MCI, is a critical concern that requires early identification. Predicting cognitive impairment, including MCI and other forms of decline, is essential for implementing preventive strategies and improving long-term health outcomes. Therefore, identifying individuals at potential risk for cognitive impairment is crucial.

While MRI and cerebrospinal fluid biomarkers, such as amyloid β, are significant indicators of AD [8], we aim to develop an algorithm that allows for the precise identification of individuals at risk using a more convenient and rapid approach, without relying on complex analyses like MRI or cerebrospinal fluid biomarkers. Recent studies have established a correlation between blood biomarkers and various factors, including diseases such as hypertension, diabetes, heart disease, and cerebrovascular disease, and lifestyle behaviors such as smoking, physical activity, and living conditions, with cognitive impairment [9,10]. High-dimensional data analysis has proven effective in capturing features that are crucial for identifying health issues [11,12]. Hence, the incorporation of these relevant factors into a sophisticated model allows for a practical application both in clinical settings and at home, enabling doctors or families to use existing data to train and monitor patients or older adult individuals, thereby identifying high-risk groups for further treatment and intervention.

In recent years, machine learning (ML), a branch of artificial intelligence, has been increasingly used for the prediction of disease outcome [13,14]. Unlike traditional methods that rely heavily on statistical significance, ML leverages algorithms to process existing factors and develop optimized models [15]. While there has been significant research into understanding the pathogenesis and influencing factors of MCI through ML, most of these studies have primarily focused on imaging techniques such as MRI [16]. Although MRI is a powerful diagnostic tool, its high cost and the inconvenience it poses limit its practicality for widespread use. In contrast, cognitive function assessments, such as the Mini-Mental State Examination (MMSE), offer a more accessible and feasible option for large-scale population screenings [17,18].

In this study, we used follow-up data from the Chinese Longitudinal Healthy Longevity Survey (CLHLS) collected during the years 2008‐2009, 2011‐2012 and 2014, encompassing a total of 2688 participants. Cognitive impairment was classified based on the MMSE scores, with a cutoff point set at 18, to determine cognitive status after 3 years. We included routine blood indices, lifestyle behaviors, and disease history from the baseline data in the ML model for training, aiming to predict the occurrence of cognitive impairment.

Methods

Study Participants

The cohort for this study is selected from CLHLS, which is a comprehensive longitudinal survey co-orchestrated by Peking University and the China Aging Science Research Center [19,20]. The survey targets Chinese seniors aged 65 years or older and includes detailed information about their living conditions, socioeconomic status, and health profiles [21]. The CLHLS initiates its baseline survey in 1998 and subsequently conducts follow-up surveys at regular intervals, with the relevant cohorts for this study being those from 2008‐2009, 2011‐2012, and 2014. The CLHLS study maintains ethical standards with approval from the Research Ethics Committee of Peking University (IRB00001052-13074), and all participants or their legal proxies provide written informed consent.

The initial samples from these specified years consist of 8418, 6066, and 3441 participants, respectively. Baseline subjects were screened by MMSE, and cognitive impairment was defined as an MMSE score <18 points [22]. The participants without missing data on MMSE scores at baseline and follow-up while with biomarker data were included in this study. This results in a narrowed-down research sample, with 602 participants from the 2008‐2009 cohort, 1263 from 2011‐2012, and 1116 from 2014, leaving a final sample size of 2688 subjects. Among the remaining participants, those from the 2008‐2009 and 2011‐2012 waves (n=1664) were further divided into a training set (n=1331) and an internal test set (n=333). The 1024 participants from the 2014 wave were used as a prospective validation set.

Predictors

Demographic predictors include age, gender, and BMI, calculated as weight in kilograms divided by height in meters squared (kg/m^2). Data on life behaviors and disease history are collected from the CLHLS questionnaire. Life behaviors account for living status (living alone or not), current smoking and drinking habits, exercise practices, marital status, and overall activity ability and mental health. Activity ability is assessed through activities of daily living (ADL) and instrumental activities of daily living (IADL). ADL was assessed by 6 indicators, including bathing, dressing, toileting, indoor transfer, continence, and eating. If all 6 items can be completely self-care, it means that daily life activities can take care of themselves. If

XSL•FO
RenderX

one or more items cannot be completely self-care, it means the daily life activities cannot take care of themselves completely (0=normal, 1=disability) [23,24]. IADL was assessed by 8 indicators, including visiting neighbors, going shopping, cooking a meal, washing clothing, walking continuously for 1 km at a time, lifting a weight of 5 kg, continuously crouching and standing up 3 times, and taking public transportation. If 8 items can be completed independently, such as the evaluation method of ADL, it means the instrumental daily life activities can be completed by themselves (0=normal or 1=disability) [23,25,26].

The development of the model in this study encompasses 3 categories of predictors derived from the baseline survey data: biomarkers, life behaviors, and disease history. The set of biomarkers comprises both routine blood examination indices and plasma biochemical examination indices. Routine blood indices include white blood cell count, red blood cell count, hemoglobin, erythrocyte hematocrit, erythrocyte mean corpuscular volume, erythrocyte mean corpuscular hemoglobin, erythrocyte mean corpuscular hemoglobin concentration, platelet count, plateletcrit, mean platelet volume, lymphocyte count, percentage of lymphocytes, and platelet distribution width. The plasma biochemical indices include high-density lipoprotein cholesterol (HDL), uric acid, plasma creatinine, glucose, triglyceride, total cholesterol (CHO), high-sensitivity c-reactive protein, malondialdehyde (MDA), and superoxide dismutase activity.

Mental health evaluation incorporates 7 questions, with 4 positively framed inquiries (regarding optimism, neatness, decision-making, and happiness relative to youth) and 3 negatively framed ones (concerning fear, loneliness, and feelings of decreased self-worth with age). Responses are scored on a scale, with higher scores correlating with poorer mental health. Disease history captures the presence or absence of hypertension, diabetes, heart disease, stroke, cancer, and arthritis. These multifaceted predictors collectively contribute to the ML model, providing a comprehensive profile for the assessment of cognitive impairment risk.

Data Preprocessing and Model Configuration

The data preprocessing phase involved addressing missing values and mitigating class imbalance to ensure a robust foundation for model training. Missing values in the dataset were imputed using the mean of the respective columns, which ensured completeness and preserved the statistical properties of the data [27]. In order to address the issue of class imbalance, the SMOTE (Synthetic Minority Over-Sampling Technique) was applied to the training set. SMOTE effectively generated synthetic samples for the minority class, enhancing the model's ability to learn from the imbalanced data distribution [28].

The selection of ML models was driven by a desire to compare different algorithms' performance and assess their robustness in dealing with imbalanced datasets. We chose 5 widely used algorithms: random forest (RF) [29], Extreme Gradient Boosting (XGBoost) [30], logistic regression [31], support vector machines (SVM) [32], and balanced random forest (BRF) [33]. These models were selected for their varied approaches to classification and their effectiveness in handling different types of data. RF and XGBoost are ensemble models that excel in handling high-dimensional data and capturing nonlinear relationships. Logistic regression and SVM are classic algorithms for binary classification, with SVM known for its ability to handle high-dimensional spaces effectively. The addition of BRF was made specifically to address the class imbalance issue in the RF model. The main difference between RF and BRF is how they handle class imbalance.

Each model was configured as follows:

- XGBoost (XGBClassifier): a learning rate of 0.03 and a maximum depth of 4 were chosen for the XGBoost model, with class weights incorporated to address class imbalance.
- Logistic regression: configured with a maximum of 1000 iterations and class weights set to "balanced" to account for class imbalance.
- SVM: probability estimation was enabled and class weights were balanced to ensure fairness across experiments.
- BRF: class weights were set to "balanced" to address the class imbalance within the dataset.

Outcomes

The assessment of cognitive function among participants was conducted using the MMSE, administered at both baseline and during follow-up sessions. The MMSE encompasses evaluations across various cognitive domains including orientation, registration, attention and calculation, recall, and language abilities, with a maximum achievable score of 30 points. In this study, the primary outcome for the ML model is the determination of cognitive impairment, defined as an MMSE score of less than 18 points at follow-up.

Model Construction, Evaluation, and Interpretation

The construction of each model followed a systematic approach to ensure rigorous evaluation and validation. Initially, the CLHLS 2008 - 2009 and CLHLS 2011 - 2012 datasets were merged. From this combined dataset, 1331/1664 (80%) of the data was randomly allocated as the training set, while the remaining 333/1664 (20%) was reserved as an internal test set for model evaluation. And then, the CLHLS 2014 dataset was used as a prospective validation set, with the model training based exclusively on the merged data from CLHLS 2008 - 2009 and CLHLS 2011 - 2012. Standard performance metrics—including accuracy, sensitivity, specificity, and the area under the receiver operating characteristic curve—were calculated to assess model effectiveness. All statistical analyses are conducted using Python (version 3.8; Python Software Foundation)

SHAP (Shapley Additive Explanations) [34] is a method for interpreting the output of ML models by assigning a contribution value to each feature, allowing us to understand the impact of individual predictors on a model's decision. In our study, SHAP was used to explain the predictive model for cognitive impairment in older adult individuals. SHAP values decompose the model's prediction into individual contributions from each feature, making it possible to attribute the output to the various risk factors in a transparent and interpretable manner.

XSL•FO
RenderX

Ethical Considerations

The CLHLS was approved by the Duke University Institutional Review Board (Pro00062871) and the Peking University Biomedical Ethics Committee (IRB00001052–13074). All participants provided written informed consent. The data used in this study were deidentified to protect participant privacy and confidentiality. No compensation wasprovided to participants.

Results

Participant Characteristics

Table 1 summarizes the baseline characteristics of the 2688 participants included in this study, stratified into 3 subsets: the training set (n=1331), the internal test set (n=333), and the prospective validation set (n=1024). Overall, the mean age was 79.73 (SD 10.97) years and slightly over half of the participants were male (53.24%). The mean BMI was 21.81 (SD 3.8 kg/m^2). Most participants lived with others (79.49%), did not smoke (77.79%), or drink (80.96%) at present, and had normal ADL (95.02%). Additionally, more than half (55.28%) showed abnormal mental health status, while 51.45% were married and living with a spouse. Several biochemical indicators were measured. The mean white blood cell count was 5.85 (SD 1.74×10^9)/L, red blood cell count was 4.64 (SD 1.68×10^{12})/L, and hemoglobin was 131.56 (SD 30.20) g/L. Participants had a mean HDL level of 1.31 (SD 0.38) mmol/L, and other biomarkers (eg, uric acid, creatinine, glucose, triglycerides, CHO, high-sensitivity c-reactive protein, MDA, and superoxide dismutase) were also assessed. Regarding disease history, hypertension was most prevalent (27.18%), followed by arthritis (11.31%), heart disease (8.01%), stroke (6.29%), and diabetes (2.75%). Cancer was reported by 0.42% of participants. Detailed distributions of these characteristics across the training, internal test, and prospective validation sets are presented in Table 1. A detailed flow diagram of participant inclusion and exclusion is provided in Figure 1.

XSL•FO
RenderX

Table . Characteristics of study subjects at baseline.

Predictors		Overall (n=2688)	Train set (n=1331)	Internal test set (n=333)	Prospective validation set (n=1024)
Demographics and life behaviors					
Age, year, mean (SD)		79.73 (10.97)	80.24 (11.43)	79.69 (11.6)	79.73 (10.07)
Gender, n (%)					
	Male	1431 (53.24)	685 (51.47)	173 (51.95)	573 (55.96)
	Female	1257 (46.76)	646 (48.53)	160 (48.05)	451 (44.04)
BMI, kg/m^2, mean (SD)		21.81 (3.8)	21.51 (3.86)	21.32 (3.67)	22.31 (3.64)
Living alone, n (%)					
	Yes	543 (20.51)	267 (20.37)	62 (18.9)	214 (21.21)
	No	2105 (79.49)	1044 (79.63)	266 (81.1)	795 (78.79)
Smoke at present, n (%)					
	Yes	593 (22.21)	317 (23.94)	72 (21.82)	204 (20.08)
	No	2077 (77.79)	1007 (76.06)	258 (78.18)	812 (79.92)
Drink at present, n (%)					
	Yes	509 (19.04)	265 (19.98)	57 (17.17)	187 (18.42)
	No	2164 (80.96)	1061 (80.02)	275 (82.83)	828 (81.58)
Exercise at present, n (%)					
	Yes	2134 (81.08)	1056 (81.23)	267 (81.65)	811 (80.7)
	No	498 (18.92)	244 (18.77)	60 (18.35)	194 (19.3)
ADL[a], n (%)					
	Normal	2479 (95.02)	1232 (93.97)	313 (95.72)	934 (96.19)
	Disability	130 (4.98)	79 (6.03)	14 (4.28)	37 (3.81)
IADL[b], n (%)					
	Normal	1593 (59.53)	742 (55.87)	198 (59.64)	653 (64.27)
	Disability	1083 (40.47)	586 (44.13)	134 (40.36)	363 (35.73)
Marital status, n (%)					
	Married and living with spouse	1370 (51.45)	654 (49.32)	165 (49.55)	551 (54.88)
	Others	1293 (48.55)	672 (50.68)	168 (20.45)	453 (45.12)
Mental health, n (%)					
	Normal	1122 (44.72)	542 (43.78)	139 (44.84)	441 (45.89)
	Abnormal	1387 (55.28)	696 (56.22)	171 (55.16)	520 (54.11)
Biomarkers, mean (SD)					
WBC[c], 10^9/L		5.85 (1.74)	5.63 (1.76)	5.64 (1.89)	6.19 (1.61)
RBC[d], 10^{12}/L		4.64 (1.68)	4.82 (2.11)	4.98 (1.9)	4.3 (0.56)
HGB[e], g/L		131.56 (30.2)	129.98 (26.56)	132.23 (22.69)	133.36 (36)
HCT[f], %		36.31 (15.05)	33.77 (18.01)	34.25 (18.1)	40.07 (7.23)
MCV[g], fL		93.95 (10.24)	93.51 (12.24)	92.94 (9.76)	94.81 (7.03)
MCH[h], pg		29.46 (7.24)	28.38 (5.25)	28.16 (5.75)	31.24 (9.22)
MCHC[i], g/L		312.18 (42.25)	304.34 (48.89)	301.38 (54.41)	325.58 (18.74)
PLT[j], 10^9/L		207.32 (99.24)	216.44 (109.67)	213.75 (111.53)	193.65 (77.22)

XSL•FO
RenderX

Predictors			Overall (n=2688)	Train set (n=1331)	Internal test set (n=333)	Prospective validation set (n=1024)
	PCT[k], %		0.27 (3.29)	0.33 (4.42)	0.2 (0.1)	0.18 (0.06)
	MPV[l], fL		9.54 (5.14)	9.43 (2.87)	9.97 (6.53)	9.55 (6.65)
	LYMPH[m], 10^9/L		14.56 (15.81)	23.13 (15.58)	21.28 (15.48)	2.03 (1.25)
	LYM%[n], %		19.55 (16.81)	10.49 (14.48)	11.59 (15.01)	33.06 (9.32)
	PDW[o], fL		15.47 (4.37)	14.86 (2.38)	14.87 (1.99)	16.44 (6.27)
	HDL[p], mmol/L		1.31 (0.38)	1.26 (0.37)	1.24 (0.38)	1.41 (0.37)
	UA[q], umol/L		291 (85.39)	285.55 (85.47)	288.51 (87.31)	298.83 (84.13)
	CRE[r], mmol/L		81.36 (24.19)	82.49 (25.97)	81.68 (22.2)	79.81 (22.28)
	GLU[s], mmol/L		5.06 (2.05)	4.83 (2.29)	4.85 (1.87)	5.42 (1.71)
	TG[u], mmol/L		1.29 (0.95)	1.22 (0.95)	1.44 (1.24)	1.34 (0.82)
	CHO[u], mmol/L		4.34 (1.17)	4.06 (1.18)	3.94 (1.22)	4.84 (0.95)
	CRPHS[v], mg/L		4.17 (18.46)	5.46 (23.21)	6.15 (29.35)	2.49 (5.29)
	MDA[w], nmol/ml		5.64 (3.4)	5.43 (2.76)	5.2 (2.63)	5.93 (4.04)
	SOD[x], IU/mL		55.53 (10.33)	55.53 (11.9)	53.34 (11.61)	56.04 (8.13)
Disease history, n (%)						
	Hypertension					
		Yes	716 (27.18)	316 (24.09)	88 (26.91)	312 (31.36)
		No	1918 (72.82)	996 (75.91)	239 (73.09)	683 (68.64)
	Diabetes					
		Yes	73 (2.75)	31 (2.34)	11 (3.33)	31 (3.11)
		No	2577 (97.25)	1292 (97.66)	319 (96.67)	966 (96.89)
	Heart disease					
		Yes	212 (8.01)	91 (6.9)	28 (8.54)	93 (9.29)
		No	2436 (91.99)	1228 (93.1)	300 (91.46)	908 (90.71)
	Stroke					
		Yes	167 (6.29)	77 (5.82)	20 (6.08)	70 (6.97)
		No	2490 (93.71)	1247 (94.18)	309 (93.92)	934 (93.03)
	Cancer					
		Yes	11 (0.42)	3 (0.23)	1 (0.3)	7 (0.73)
		No	2588 (99.58)	1301 (99.77)	329 (99.7)	958 (99.27)
	Arthritis					
		Yes	300 (11.31)	183 (13.86)	55 (16.72)	62 (6.18)

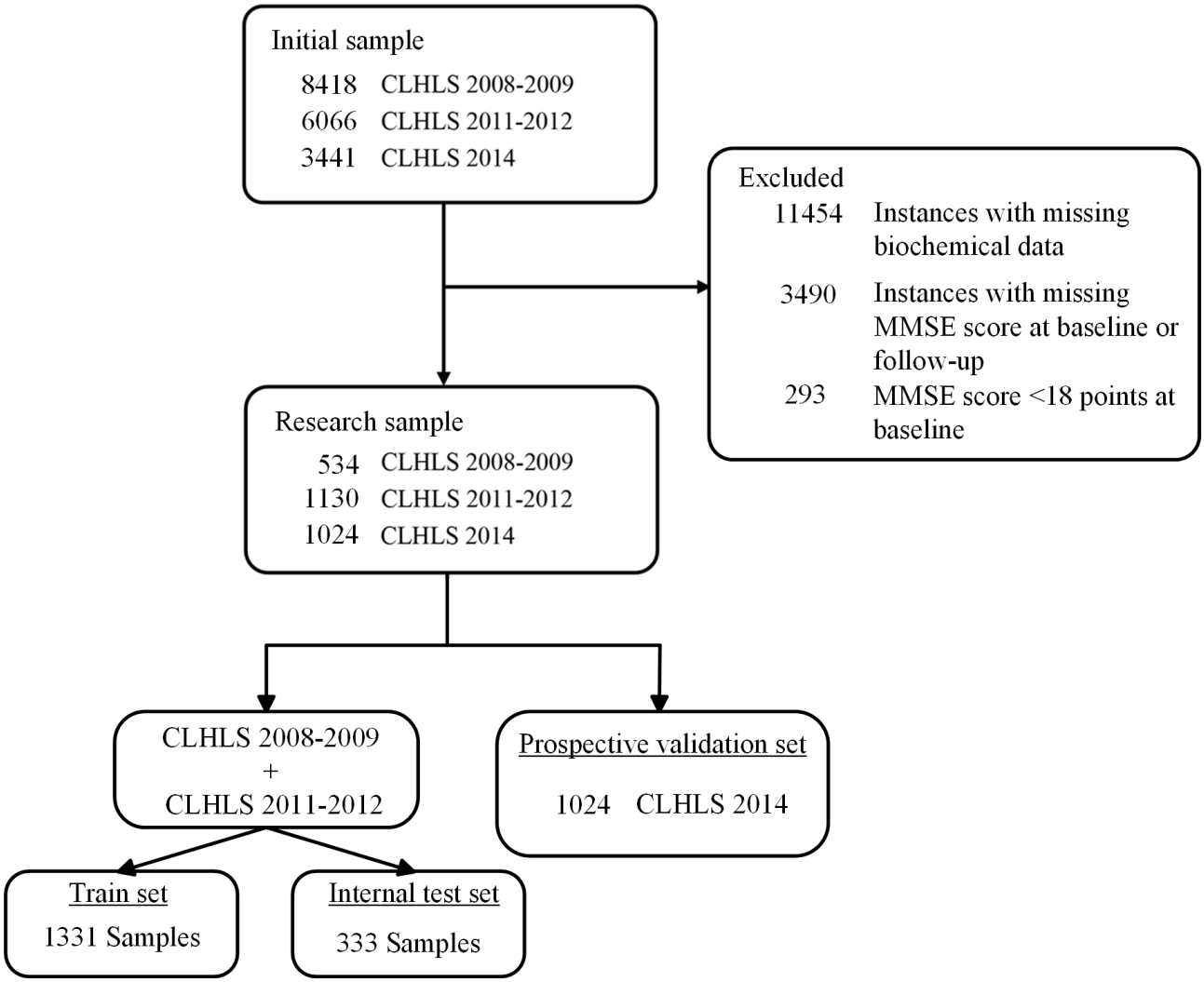
XSL•FO
RenderX

Predictors		Overall (n=2688)	Train set (n=1331)	Internal test set (n=333)	Prospective validation set (n=1024)
	No	2352 (88.69)	1137 (86.14)	274 (83.28)	941 (93.82)

[a]ADL: activities of daily living.
[b]IADL: instrumental activities of daily living.
[c]WBC: white blood cell count.
[d]RBC: red blood cell count.
[e]HGB: hemoglobin.
[f]HCT: erythrocyte hematocrit.
[g]MCV: erythrocyte mean corpuscular volume.
[h]MCH: erythrocyte mean corpuscular hemoglobin.
[i]MCHC: erythrocyte mean corpuscular hemoglobin concentration.
[j]PLT: platelet count.
[k]PCT: plateletocrit.
[l]MPV: mean platelet volume.
[m]LYMPH: lymphocyte count.
[n]LYM%: percentage of lymphocytes.
[o]PDW: platelet distribution width.
[p]HDL: high-density lipoprotein cholesterol.
[q]UA: urea acid.
[r]CRE: plasma creatine.
[s]GLU: glucose.
[t]TG: triglyceride.
[u]CHO: total cholesterol.
[v]CRPHS: high-sensitivity c-reactive protein.
[w]MDA: malondialdehyde.
[x]SOD: superoxide dismutase activity.

XSL•FO
RenderX

Figure 1. Flow diagram of the subject selection. CLHLS: Chinese Longitudinal Healthy Longevity Survey; MMSE: Mini-Mental State Examination.



Model Performance

The performance of the ML models in predicting follow-up cognitive impairment was evaluated using 5 algorithms—RF, XGBoost, logistic regression, SVM, and BRF—on both the internal test set and prospective validation set. Each model's accuracy, sensitivity, and area under the receiver operating characteristic curve were assessed. The detailed results are presented in Table 2 and visually depicted in Figure 2. Additionally, we used the MMSE as an input for ML prediction. After testing several ML models, the overall performance remained suboptimal. For illustration, we selected 1 representative result, as shown in Table 2.

XSL•FO
RenderX

Table . The prediction results for ML[a] models.

	Internal test			Prospective validation		
	Accuracy (95% CI)	Sensitivity (95% CI)	AUC[b] (95% CI)	Accuracy (95% CI)	Sensitivity (95% CI)	AUC (95% CI)
RF[c]	0.828 (0.806-0.854)	0.688 (0.633-0.742)	0.81 (0.769-0.852)	0.813 (0.79-0.84)	0.684 (0.63-0.739)	0.806 (0.767-0.848)
XGBoost[d]	0.849 (0.825-0.87)	0.674 (0.621-0.728)	0.811 (0.769-0.851)	0.836 (0.812-0.857)	0.682 (0.629-0.733)	0.808 (0.768-0.848)
Logistic regression	0.778 (0.751-0.803)	0.715 (0.66-0.761)	0.803 (0.753-0.847)	0.771 (0.744-0.795)	0.745 (0.696-0.789)	0.804 (0.753-0.847)
SVM[e]	0.69 (0.661-0.719)	0.681 (0.629-0.731)	0.777 (0.725-0.827)	0.672 (0.643-0.701)	0.676 (0.623-0.725)	0.777 (0.726-0.826)
Balanced RF classifier	0.885 (0.866-0.903)	0.58 (0.54-0.626)	0.809 (0.765-0.849)	0.887 (0.867-0.905)	0.616 (0.572-0.664)	0.811 (0.767-0.849)
MMSE[f]	N/A[g]	N/A	0.571 (0.485-0.653)	N/A	N/A	0.558 (0.494-0.62)

[a]ML: machine learning.

[b]AUC: area under the curve.

[c]RF: random forest.

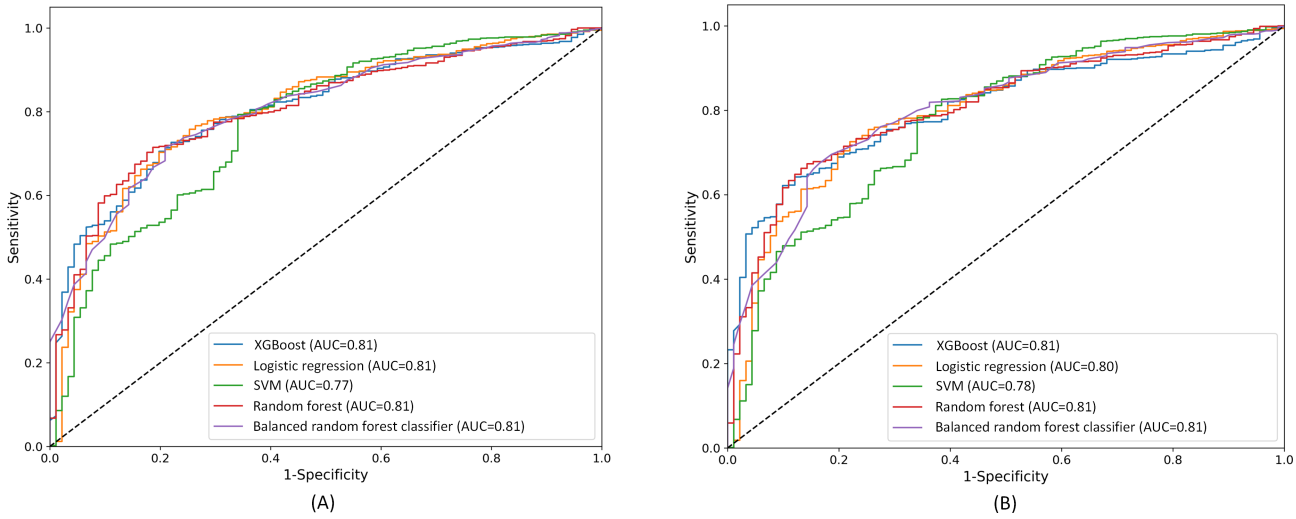
[d]XGBoost: Extreme Gradient Boosting.

[e]SVM: support vector machines.

[f]MMSE: Mini-Mental State Examination.

[g]N/A: not applicable.

Figure 2. ROC curves with AUC values for machine learning models: (A) internal test set and (B) prospective validation set. AUC: area under the curve; ROC: receiver operating characteristic; SVM: support vector machines; XGBoost: Extreme Gradient Boosting.



In the internal test set, the BRF achieved the highest accuracy at 88.5% but had lower sensitivity at 58%. RF and XGBoost provided balanced results with accuracies of 82.8% and 84.9% and sensitivities of 68.8% and 67.4%, respectively. Logistic regression yielded a moderate accuracy of 77.8% but the highest sensitivity at 71.5%, while SVM had the lowest performance with 69% accuracy and 68.1% sensitivity. Similar patterns were observed in the prospective validation set, with BRF at 88.7% accuracy, logistic regression reaching 74.5% sensitivity, and SVM again showing the lowest overall performance. Additionally, the performance of all ML models was superior to that of the MMSE.

Figure 3 offers a detailed exposition of the XGBoost for cognitive impairment in older adult individuals, featuring the 20 most influential risk predictors. The model prioritizes IADL, age, and baseline MMSE scores as the top determinants, with marital status, living alone, and hypertension also providing significant predictive value. Other important factors include biological markers such as MDA and HDL, alongside lifestyle factors such as current exercise, smoking, and drinking habits.

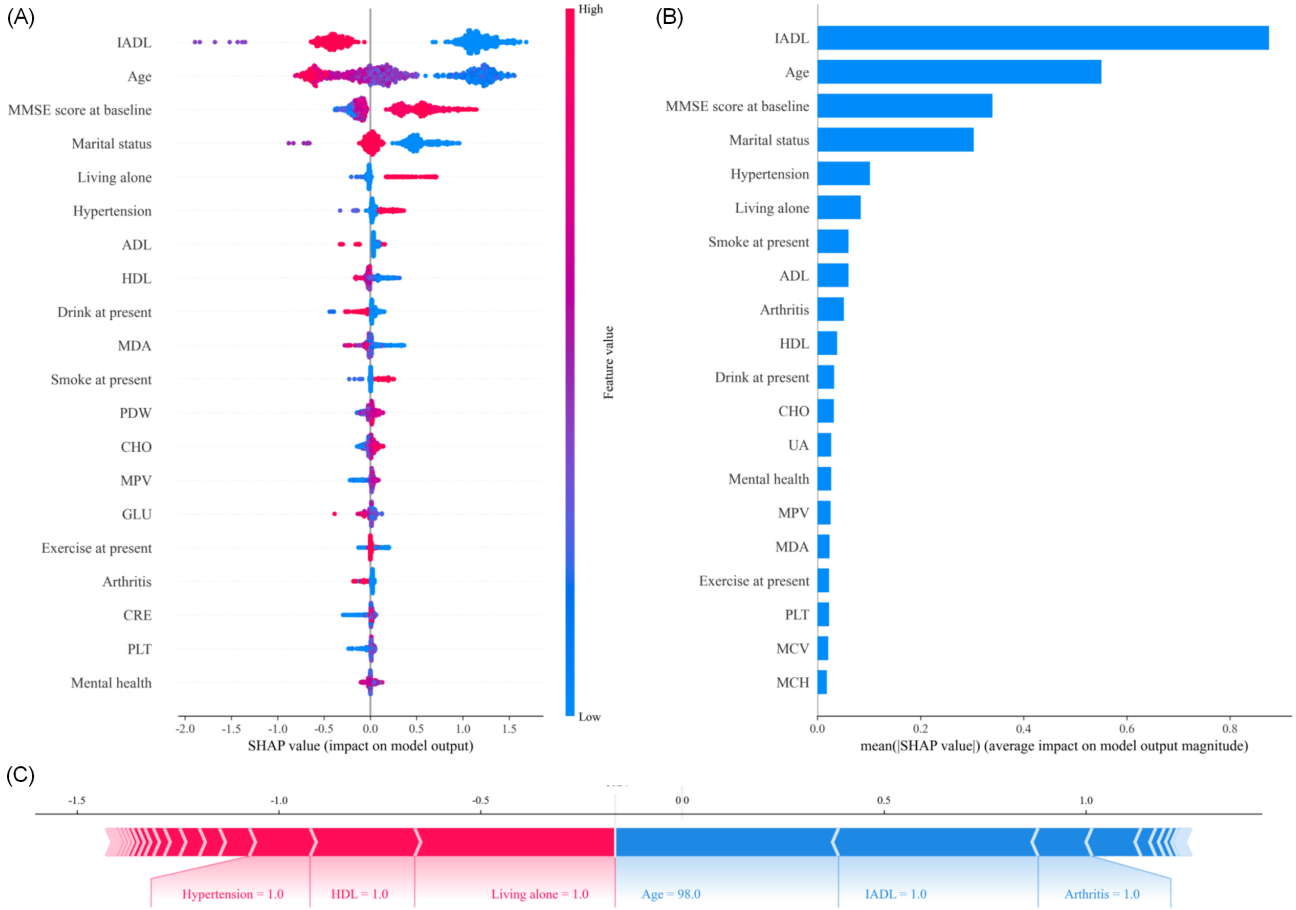
Figure 3A illustrates the distribution of SHAP values for these predictors, indicating their impact on the XGBoost's output. In this plot, each dot represents a single instance of a feature in the dataset, and the horizontal axis shows the SHAP value of that feature. The SHAP value indicates the contribution of each feature to the model's prediction, with positive values suggesting that the feature increases the likelihood of cognitive impairment, and negative values indicating a decrease. The color of the dots, ranging from blue to red, represents the feature's value, with blue corresponding to lower values and red indicating higher values of the predictor. This color scheme helps highlight how different values of each predictor influence the model's outcome.

XSL•FO
RenderX

A high SHAP value for a given feature corresponds to a high level of importance in the predictive model. Figure 3B features a bar plot that quantifies the average impact of each predictor, measured by the mean absolute SHAP value. The length of each bar represents the average contribution of a feature to the model's output across all data points. Here, the most important features in the model—such as IADL, age, and MMSE score at baseline—are easily identified, as they have the longest bars, indicating that they have the highest average impact on the model's predictions.

Figure 3C illustrates a specific case study, showing how the SHAP values for a particular individual (in this case, an adult aged 98 years) contribute to the XGBoost's prediction of cognitive impairment. Each feature is shown with its value (eg, hypertension=1), and the arrows indicate how these values shift the model's prediction. Features such as hypertension and living alone appear to have a red color, indicating they push the prediction toward a higher risk of cognitive impairment. Similarly, age (with a value of 98) and IADL further emphasize the risk in this individual's profile. The interaction of these predictors is visualized through their SHAP values, which collectively guide the prediction model's decision, offering an individualized risk profile for this person.

Figure 3. Explanation of the interpretability of the XGBoost (Extreme Gradient Boosting) model (the best-performing model) for predicting older adult mortality. (A) and (B) show the top 20 risk predictors for prediction of cognitive impairment subjects, and (C) shows the SHAP plots of a subject. ADL: activities of daily living; CHO: total cholesterol; CRE: plasma creatine; GLU: glucose; HDL: high-density lipoprotein cholesterol; IADL: instrumental activities of daily living; MCH: erythrocyte mean corpuscular hemoglobin; MCV: erythrocyte mean corpuscular volume; MDA: malondialdehyde; MMSE: Mini-Mental State Examination; MPV: mean platelet volume; PDW: platelet distribution width; PLT: platelet count; SHAP: Shapley Additive Explanations; UA: uric acid.



Discussion

Principal Findings

This study uses the capabilities of ML to integrate a diverse set of 39 predictors for forecasting the decline in cognitive function over a 3-year period, yielding significant findings that resonate with the existing body of literature. Our ML model underscores the importance of IADL, age, baseline MMSE scores, marital status, living arrangements, hypertension, arthritis, and general lifestyle habits as pivotal factors influencing cognitive function. These determinants are consistent with findings from prior research, thereby affirming the reliability and relevance of our analytical approach [12,22].

Consistent with prior studies, advanced age has been identified as a significant risk factor for cognitive impairment [35,36]. The risk of MCI in the older adults aged 65 years or older is as high as 10% - 20% [37]. The limitation of older adult individuals' ability to perform activities, as assessed by ADL and IADL, restricts their range of activities, diminishes social interactions, and consequently reduces the cerebral stimulation necessary for maintaining cognitive functions [38,39]. Furthermore, the impact of living conditions on cognitive health

XSL•FO
RenderX

is evidenced by the predictive value of marital status and living alone [22,40-42]. Older adult individuals residing with a spouse typically exhibit healthier brain functions due to increased communication and maintenance of a normal life.

Additionally, this study underscores the detrimental effects of unhealthy lifestyle behaviors such as smoking and drinking on cognitive function [43-45], as well as the risk of cognitive impairment caused by pre-existing health conditions [46-49]. For instance, an Indian study highlighted that older smokers were 24% more likely to experience cognitive impairment compared to nonsmokers [50]. Similarly, Sabia et al [45] reported that abstaining from alcohol or consuming more than 14 units per week in middle age escalates the risk of AD. Moreover, recent research suggests that up to 3% of dementia cases could be averted by enhancing physical activity levels [51,52]. Concerning health conditions, a community-based cohort study illustrated that hypertension is associated with an increased risk of both all-cause MCI and nonamnestic MCI, with hazard ratios of 1.4 and 1.7, respectively, after age and sex adjustments [49]. Additionally, Appenzeller et al [53] found that patients with rheumatoid arthritis exhibited a significantly higher incidence of cognitive impairment compared to healthy controls.

In this study, in addition to age, lifestyle behaviors, and disease history—which can be evaluated through questionnaires or scales—biomarkers were specifically included in the ML model to enhance the predictive accuracy of clinical risk assessments for cognitive decline. Prior research has indicated that count elevated levels of specific biomarkers, including MDA, HDL, platelets, mean platelet volume, platelet distribution width, mean corpuscular hemoglobin, CHO, lymphocyte percentage, and plasma creatinine, are associated with an increased risk of cognitive decline [54-65]. For clinicians, integrating a patient's lifestyle behaviors with blood biochemical markers can aid in the assessment of cognitive function. For communities, these indicators can help identify residents who may be at high risk. For family members, this model enables the evaluation of older adult relatives who may be reluctant to acknowledge their cognitive decline, thereby facilitating timely medical intervention. The model developed in this study is versatile and offers valuable insights for the identification of cognitive impairment across various settings.

The ML model developed in this study has the potential to significantly improve clinical practice and primary care by providing a rapid, efficient, and accessible tool for identifying individuals at risk of cognitive decline. Traditional methods for cognitive function assessment, such as imaging techniques such as MRI, can be time-consuming and resource-intensive, especially in resource-limited settings. By leveraging routine data such as blood biomarkers, demographic information, and lifestyle factors, this model offers a cost-effective approach to identify individuals who may require further clinical evaluation or early intervention. In primary care settings, where health care professionals often manage large volumes of patients, the model can serve as a valuable screening tool to detect early cognitive decline and facilitate referrals for specialized care. Furthermore, by integrating this model into electronic health records, health care providers can make timely and informed decisions, improving patient outcomes through proactive management. In essence, the model has the potential to transform early detection and intervention strategies, shifting the focus toward preventative care and better allocation of health care resources.

Compared to previous studies using ML for cognitive impairment prediction, our study offers several distinct contributions. For example, while studies such as those by Hu et al [22] and Gao et al [66] have successfully developed ML models to predict cognitive impairment among Chinese community-dwelling older adult individuals, they often focused on a more limited set of predictors—typically emphasizing demographic factors and neuropsychological assessments. In contrast, our study integrates a comprehensive set of 39 predictors, including both routine blood biomarkers (eg, MDA and HDL) and detailed lifestyle and disease history data. This broader approach not only enhances predictive accuracy but also provides a rapid, cost-effective tool that can be easily applied in community and clinical settings. Moreover, while some previous work [67,68] has primarily relied on imaging data or traditional statistical methods, our use of advanced ensemble ML techniques (such as BRF and XGBoost) combined with SHAP-based interpretability offers a clearer understanding of individual risk factors. This interpretability is crucial for clinicians to tailor early intervention strategies. In summary, our study advances the field by delivering a more inclusive and interpretable model that effectively tracks cognitive decline over 3 years, thereby addressing gaps in existing research and offering tangible benefits for early detection and intervention.

Nevertheless, this study presents certain limitations that warrant consideration. Primarily, the dependence on MMSE scores as the sole measure of cognitive impairment may not fully represent the broad spectrum of cognitive health, as the MMSE mainly evaluates specific cognitive domains and does not address emotional or psychological aspects. Furthermore, there is potential bias due to the generally younger age and better overall function of this study's participants compared to those who were lost to follow-up, which could skew the results. However, the representation of the dataset at a national level does provide a measure of balance, helping to partially offset these biases.

Conclusions

In conclusion, this study validates the efficacy of a ML model integrating demographic data, lifestyle factors, and biomarkers to predict cognitive impairment in older adults. It underscores the significance of traditional risk factors such as age and daily functional abilities while highlighting the role of solitary living conditions and unhealthy habits in cognitive decline. By including a broad spectrum of biomarkers, the model enriches the predictive framework, offering clinicians, communities, and families a valuable tool for early identification and intervention in cognitive impairment, which could have far-reaching implications for public health and the well-being of the aging population.

Limitations

While the CLHLS is a large-scale longitudinal study that primarily focuses on individuals aged 65 years and older, to assess the health status and longevity of the older adult

XSL•FO
RenderX

population in China, it is important to acknowledge the limitations of the dataset concerning its representativeness. The CLHLS sample is designed to represent the health conditions of the older adult population in China, but it may not fully capture the global demographics of AD or other forms of dementia. Specifically, older adult populations in other countries or regions may differ by genetic background, lifestyle factors, and health risks, which could influence the development and progression of cognitive impairment. Future studies incorporating diverse, multinational cohorts would be beneficial in enhancing the generalizability and robustness of cognitive decline prediction models.

Acknowledgments

We thank CLHLS for providing data, and all participants for supporting this study. This study was supported by Guangzhou Science and Technology Plan Project (2024A03J1074 and 2024A03J1067) and General Project of Guangdong Provincial Natural Science Foundation (2025A1515012013).

Data Availability

The data used in this study are stored at Peking University and available upon request.

Authors' Contributions

YZ and CL contributed equally to the work. YZ, CL, HR, QW, ZL, D Li, D Liang, WT, and LL performed the data collection and analysis. YZ, CL, HR, and QW contributed to visualization. YZ and HR were responsible for data validation. YZ, CL and HR wrote this paper, and review editing was performed by HR and WC. All the authors read and approved the final paper.

Conflicts of Interest

None declared.

References

1. Hane FT, Robinson M, Lee BY, Bai O, Leonenko Z, Albert MS. Recent progress in Alzheimer's disease research, part 3: diagnosis and treatment. J Alzheimers Dis 2017;57(3):645-665. [doi: 10.3233/JAD-160907] [Medline: 28269772]
2. Scheltens P, De Strooper B, Kivipelto M, et al. Alzheimer's disease. Lancet 2021 Apr 24;397(10284):1577-1590. [doi: 10.1016/S0140-6736(20)32205-4] [Medline: 33667416]
3. 2023 Alzheimer's disease facts and figures. Alzheimers Dement 2023 Apr;19(4):1598-1695. [doi: 10.1002/alz.13016]
4. Prince M, Wimo A, Ali GC, Wu YT, Prina M. World Alzheimer report 2015: the global impact of dementia: an analysis of prevalence, incidence, cost and trends. : Alzheimer's Disease International; 2015.
5. Takizawa C, Thompson PL, van Walsem A, Faure C, Maier WC. Epidemiological and economic burden of Alzheimer's disease: a systematic literature review of data across Europe and the United States of America. JAD 2014;43(4):1271-1284. [doi: 10.3233/JAD-141134]
6. Jia J, Wei C, Chen S, et al. The cost of Alzheimer's disease in China and re-estimation of costs worldwide. Alzheimers Dement 2018 Apr;14(4):483-491. [doi: 10.1016/j.jalz.2017.12.006] [Medline: 29433981]
7. Anderson ND. State of the science on mild cognitive impairment (MCI). CNS Spectr 2019 Feb;24(1):78-87. [doi: 10.1017/S1092852918001347] [Medline: 30651152]
8. Tan CC, Yu JT, Tan L. Biomarkers for preclinical Alzheimer's disease. J Alzheimers Dis 2014;42(4):1051-1069. [doi: 10.3233/JAD-140843] [Medline: 25024325]
9. Dong X, Nao J, Shi J, Zheng D. Predictive value of routine peripheral blood biomarkers in Alzheimer's disease. Front Aging Neurosci 2019;11:332. [doi: 10.3389/fnagi.2019.00332] [Medline: 31866854]
10. Jia L, Du Y, Chu L, et al. Prevalence, risk factors, and management of dementia and mild cognitive impairment in adults aged 60 years or older in China: a cross-sectional study. Lancet Public Health 2020 Dec;5(12):e661-e671. [doi: 10.1016/S2468-2667(20)30185-7] [Medline: 33271079]
11. Jing F, Cheng M, Li J, et al. Social, lifestyle, and health status characteristics as a proxy for occupational burnout identification: a network approach analysis. Front Psychiatry 2023;14:1119421. [doi: 10.3389/fpsyt.2023.1119421] [Medline: 37124263]
12. Tan WY, Hargreaves C, Chen C, Hilal S. A machine learning approach for early diagnosis of cognitive impairment using population-based data. J Alzheimers Dis 2023;91(1):449-461. [doi: 10.3233/JAD-220776] [Medline: 36442196]
13. Aguayo GA, Zhang L, Vaillant M, et al. Machine learning for predicting neurodegenerative diseases in the general older population: a cohort study. BMC Med Res Methodol 2023 Jan 11;23(1):8. [doi: 10.1186/s12874-023-01837-4] [Medline: 36631766]
14. Tan M, Xiao Y, Jing F, et al. Evaluating machine learning-enabled and multimodal data-driven exercise prescriptions for mental health: a randomized controlled trial protocol. Front Psychiatry 2024;15:1352420. [doi: 10.3389/fpsyt.2024.1352420] [Medline: 38287940]

XSL•FO
RenderX

15. Navarro CLA, Damen JAA, Takada T, et al. Completeness of reporting of clinical prediction models developed using supervised machine learning: a systematic review. BMC Med Res Methodol 2022 Jan 13;22(1):12. [doi: 10.1186/s12874-021-01469-6] [Medline: 35026997]
16. Frizzell TO, Glashutter M, Liu CC, et al. Artificial intelligence in brain MRI analysis of Alzheimer's disease over the past 12 years: a systematic review. Ageing Res Rev 2022 May;77:101614. [doi: 10.1016/j.arr.2022.101614] [Medline: 35358720]
17. Tombaugh TN, McIntyre NJ. The Mini-Mental State Examination: a comprehensive review. J Am Geriatr Soc 1992 Sep;40(9):922-935. [doi: 10.1111/j.1532-5415.1992.tb01992.x] [Medline: 1512391]
18. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6]
19. The Chinese Longitudinal Healthy Longevity Survey (CLHLS)—Longitudinal Data (1998–2018). Peking University Open Research Data Platform. 1998. URL: https://opendata.pku.edu.cn/dataset.xhtml?persistentId=doi:10.18170/DVN/WBO7LK [accessed 2025-04-19]
20. Zeng Y, Vaupel JW. Chinese Longitudinal Healthy Longevity Survey (CLHLS), biomarkers datasets, 2009, 2012, 2014 (ICPSR 37226). National Archive of Computerized Data on Aging. 2009. URL: https://www.icpsr.umich.edu/web/NACDA/studies/37226/versions/V1 [accessed 2025-04-19]
21. Wu J, Li J. The impact of social participation on older people's death risk: an analysis from CLHLS. China Popul Dev Stud 2018 Jun;2(2):173-185. [doi: 10.1007/s42379-018-0008-x]
22. Hu M, Shu X, Yu G, Wu X, Välimäki M, Feng H. A risk prediction model based on machine learning for cognitive impairment among chinese community-dwelling elderly people with normal cognition: development and validation study. J Med Internet Res 2021 Feb 24;23(2):e20298. [doi: 10.2196/20298] [Medline: 33625369]
23. Mlinac ME, Feng MC. Assessment of activities of daily living, self-care, and independence. Arch Clin Neuropsychol 2016 Sep;31(6):506-516. [doi: 10.1093/arclin/acw049] [Medline: 27475282]
24. Farias ST, Mungas D, Reed BR, Harvey D, Cahn-Weiner D, Decarli C. MCI is associated with deficits in everyday functioning. Alzheimer Dis Assoc Disord 2006;20(4):217-223. [doi: 10.1097/01.wad.0000213849.51495.d9] [Medline: 17132965]
25. Kiosses DN, Alexopoulos GS. IADL functions, cognitive deficits, and severity of depression: a preliminary study. Am J Geriatr Psychiatry 2005 Mar;13(3):244-249. [doi: 10.1176/appi.ajgp.13.3.244] [Medline: 15728756]
26. Graf C. The Lawton instrumental activities of daily living scale. AJN 2008;108(4):52-62. [doi: 10.1097/01.NAJ.0000314810.46029.74]
27. Little RJA, Rubin D. Statistical Analysis with Missing Data: John Wiley & Sons; 2019.
28. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling Technique. JAIR 2002;16:321-357. [doi: 10.1613/jair.953]
29. Breiman L. Random forests. Mach Learn 2001;45(1):5-32. [doi: 10.1023/A:1010933404324]
30. Chen T, Guestrin C. XGBoost. Presented at: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; Aug 13-17, 2016; San Francisco, CA, United States.
31. Sturdivant RX, Hosmer DW, Lemeshow S. Applied Logistic Regression: John Wiley & Sons; 2013. [doi: 10.1002/9781118548387]
32. Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995 Sep;20(3):273-297. [doi: 10.1007/BF00994018]
33. Khalilia M, Chakraborty S, Popescu M. Predicting disease risks from highly imbalanced data using random forest. BMC Med Inform Decis Mak 2011 Jul 29;11:1-13. [doi: 10.1186/1472-6947-11-51] [Medline: 21801360]
34. Mangalathu S, Hwang SH, Jeon JS. Failure mode and effects analysis of RC members based on machine-learning-based Shapley Additive Explanations (SHAP) approach. Eng Struct 2020 Sep;219:110927. [doi: 10.1016/j.engstruct.2020.110927]
35. Cong L, Ren Y, Wang Y, et al. Mild cognitive impairment among rural-dwelling older adults in China: a community-based study. Alzheimers Dement 2023 Jan;19(1):56-66. [doi: 10.1002/alz.12629] [Medline: 35262288]
36. Grueso S, Viejo-Sobera R. Machine learning methods for predicting progression from mild cognitive impairment to Alzheimer's disease dementia: a systematic review. Alzheimers Res Ther 2021 Sep 28;13(1):162. [doi: 10.1186/s13195-021-00900-w] [Medline: 34583745]
37. Langa KM, Levine DA. The diagnosis and management of mild cognitive impairment: a clinical review. JAMA 2014 Dec 17;312(23):2551-2561. [doi: 10.1001/jama.2014.13806] [Medline: 25514304]
38. Njegovan V, Hing MM, Mitchell SL, Molnar FJ. The hierarchy of functional loss associated with cognitive decline in older persons. J Gerontol A Biol Sci Med Sci 2001 Oct;56(10):M638-M643. [doi: 10.1093/gerona/56.10.m638] [Medline: 11584037]
39. Jekel K, Damian M, Wattmo C, et al. Mild cognitive impairment and deficits in instrumental activities of daily living: a systematic review. Alzheimers Res Ther 2015;7(1):17. [doi: 10.1186/s13195-015-0099-0] [Medline: 25815063]
40. Sundström A, Westerlund O, Kotyrlo E. Marital status and risk of dementia: a nationwide population-based prospective study from Sweden. BMJ Open 2016 Jan 4;6(1):e008565. [doi: 10.1136/bmjopen-2015-008565] [Medline: 26729377]
41. Sommerlad A, Ruegger J, Singh-Manoux A, Lewis G, Livingston G. Marriage and risk of dementia: systematic review and meta-analysis of observational studies. J Neurol Neurosurg Psychiatry 2018 Mar;89(3):231-238. [doi: 10.1136/jnnp-2017-316274] [Medline: 29183957]

XSL•FO
RenderX

42. Ahn IS, Kim JH, Kim S, et al. Impairment of instrumental activities of daily living in patients with mild cognitive impairment. Psychiatry Investig 2009 Sep;6(3):180-184. [doi: 10.4306/pi.2009.6.3.180] [Medline: 20046393]
43. Tyas SL, White LR, Petrovitch H, et al. Mid-life smoking and late-life dementia: the Honolulu-Asia aging study. Neurobiol Aging 2003;24(4):589-596. [doi: 10.1016/s0197-4580(02)00156-2] [Medline: 12714116]
44. Xu W, Wang H, Wan Y, et al. Alcohol consumption and dementia risk: a dose-response meta-analysis of prospective studies. Eur J Epidemiol 2017 Jan;32(1):31-42. [doi: 10.1007/s10654-017-0225-3] [Medline: 28097521]
45. Sabia S, Fayosse A, Dumurgier J, et al. Alcohol consumption and risk of dementia: 23 year follow-up of Whitehall II cohort study. BMJ 2018 Aug 1;362:k2927. [doi: 10.1136/bmj.k2927] [Medline: 30068508]
46. Vitturi BK, Nascimento BAC, Alves BR, de Campos FSC, Torigoe DY. Cognitive impairment in patients with rheumatoid arthritis. J Clin Neurosci 2019 Nov;69:81-87. [doi: 10.1016/j.jocn.2019.08.027] [Medline: 31447371]
47. Meade T, Manolios N, Cumming SR, Conaghan PG, Katz P. Cognitive impairment in rheumatoid arthritis: a systematic review. Arthritis Care Res (Hoboken) 2018 Jan;70(1):39-52. [doi: 10.1002/acr.23243] [Medline: 28371512]
48. Aronow WS. Hypertension and cognitive impairment. Ann Transl Med 2017 Jun;5(12):259. [doi: 10.21037/atm.2017.03.99] [Medline: 28706927]
49. Reitz C, Tang MX, Manly J, Mayeux R, Luchsinger JA. Hypertension and the risk of mild cognitive impairment. Arch Neurol 2007 Dec;64(12):1734-1740. [doi: 10.1001/archneur.64.12.1734] [Medline: 18071036]
50. Muhammad T, Govindu M, Srivastava S. Relationship between chewing tobacco, smoking, consuming alcohol and cognitive impairment among older adults in India: a cross-sectional study. BMC Geriatr 2021 Jan 29;21(1):85. [doi: 10.1186/s12877-021-02027-x] [Medline: 33514331]
51. Liang JH, Lu L, Li JY, et al. Contributions of modifiable risk factors to dementia incidence: a Bayesian network analysis. J Am Med Dir Assoc 2020 Nov;21(11):1592-1599. [doi: 10.1016/j.jamda.2020.04.006] [Medline: 32563753]
52. Dominguez LJ, Veronese N, Vernuccio L, et al. Nutrition, physical activity, and other lifestyle factors in the prevention of cognitive decline and dementia. Nutrients 2021 Nov 15;13(11):4080. [doi: 10.3390/nu13114080] [Medline: 34836334]
53. Appenzeller S, Bertolo MB, Costallat LTL. Cognitive impairment in rheumatoid arthritis. Methods Find Exp Clin Pharmacol 2004 Jun;26(5):339-343. [doi: 10.1358/mf.2004.26.5.831324] [Medline: 15319812]
54. Keller JN, Schmitt FA, Scheff SW, et al. Evidence of increased oxidative damage in subjects with mild cognitive impairment. Neurology (ECronicon) 2005 Apr 12;64(7):1152-1156. [doi: 10.1212/01.WNL.0000156156.13641.BA] [Medline: 15824339]
55. Liu Z, Liu Y, Tu X, et al. High serum levels of malondialdehyde and 8-OHdG are both associated with early cognitive impairment in patients with acute ischaemic stroke. Sci Rep 2017 Aug 25;7(1):9493. [doi: 10.1038/s41598-017-09988-3] [Medline: 28842715]
56. van Exel E, de Craen AJM, Gussekloo J, et al. Association between high-density lipoprotein and cognitive impairment in the oldest old. Ann Neurol 2002 Jun;51(6):716-721. [doi: 10.1002/ana.10220] [Medline: 12112077]
57. Atzmon G, Gabriely I, Greiner W, Davidson D, Schechter C, Barzilai N. Plasma HDL levels highly correlate with cognitive function in exceptional longevity. J Gerontol A Biol Sci Med Sci 2002 Nov;57(11):M712-M715. [doi: 10.1093/gerona/57.11.m712] [Medline: 12403798]
58. Crichton GE, Elias MF, Davey A, Sullivan KJ, Robbins MA. Higher HDL cholesterol is associated with better cognitive function: the Maine-Syracuse study. J Int Neuropsychol Soc 2014 Nov;20(10):961-970. [doi: 10.1017/S1355617714000885] [Medline: 25382185]
59. Zhang Y, Liu J, Wei Z, et al. Elevated serum platelet count inhibits the effects of brain functional changes on cognitive function in patients with mild cognitive impairment: A resting-state functional magnetic resonance imaging study. Front Aging Neurosci 2023;15:1088095. [doi: 10.3389/fnagi.2023.1088095] [Medline: 37051376]
60. Sun D, Wang Q, Kang J, et al. Correlation between serum platelet count and cognitive function in patients with atrial fibrillation: a cross-sectional study. Cardiol Res Pract 2021;2021:9039610. [doi: 10.1155/2021/9039610] [Medline: 33824766]
61. Wang RT, Jin D, Li Y, Liang QC. Decreased mean platelet volume and platelet distribution width are associated with mild cognitive impairment and Alzheimer's disease. J Psychiatr Res 2013 May;47(5):644-649. [doi: 10.1016/j.jpsychires.2013.01.014] [Medline: 23395109]
62. Winchester LM, Powell J, Lovestone S, Nevado-Holgado AJ. Red blood cell indices and anaemia as causative factors for cognitive function deficits and for Alzheimer's disease. Genome Med 2018 Jun 28;10(1):51. [doi: 10.1186/s13073-018-0556-z] [Medline: 29954452]
63. Guo Y, Li P, Ma X, et al. Association of circulating cholesterol level with cognitive function and mild cognitive impairment in the elderly: a community-based population study. Curr Alzheimer Res 2020;17(6):556-565. [doi: 10.2174/1567205017666200810165758] [Medline: 32781960]
64. Magaki S, Yellon SM, Mueller C, Kirsch WM. Immunophenotypes in the circulation of patients with mild cognitive impairment. J Psychiatr Res 2008 Feb;42(3):240-246. [doi: 10.1016/j.jpsychires.2007.01.004] [Medline: 17320906]
65. Yoshida M, Higashi K, Kuni K, et al. Distinguishing mild cognitive impairment from Alzheimer's disease with acrolein metabolites and creatinine in urine. Clinica Chimica Acta 2015 Feb;441:115-121. [doi: 10.1016/j.cca.2014.12.023]
66. Gao H, Schneider S, Hernandez R, et al. Early identification of cognitive impairment in community environments through modeling subtle inconsistencies in questionnaire responses: machine learning model development and validation. JMIR Form Res 2024 Nov 13;8:e54335. [doi: 10.2196/54335] [Medline: 39536306]

XSL•FO
RenderX

67. Graham SA, Lee EE, Jeste DV, et al. Artificial intelligence approaches to predicting and detecting cognitive decline in older adults: a conceptual review. Psychiatry Res 2020 Feb;284:112732. [doi: 10.1016/j.psychres.2019.112732] [Medline: 31978628]
68. Zhou Y, Song Z, Han X, Li H, Tang X. Prediction of Alzheimer's disease progression based on magnetic resonance imaging. ACS Chem Neurosci 2021 Nov 17;12(22):4209-4223. [doi: 10.1021/acschemneuro.1c00472] [Medline: 34723463]

Abbreviations

AD: Alzheimer disease
ADL: activities of daily living
BRF: balanced random forest
CHO: total cholesterol
CLHLS: Chinese Longitudinal Healthy Longevity Survey
HDL: high-density lipoprotein cholesterol
IADL: instrumental activities of daily living
MCI: mild cognitive impairment
MDA: malondialdehyde
ML: machine learning
MMSE: Mini-Mental State Examination
MRI: magnetic resonance imaging
RF: random forest
SHAP: Shapley Additive Explanations
SMOTE: Synthetic Minority Over-Sampling Technique
SVM: support vector machines
XGBoost: Extreme Gradient Boosting

Edited by H Köttl; submitted 19.10.24; peer-reviewed by A Garikipati, Y Wu; revised version received 05.03.25; accepted 01.04.25; published 30.04.25.

Please cite as:
Ren H, Zheng Y, Li C, Jing F, Wang Q, Luo Z, Li D, Liang D, Tang W, Liu L, Cheng W
Using Machine Learning to Predict Cognitive Decline in Older Adults From the Chinese Longitudinal Healthy Longevity Survey: Model Development and Validation Study
JMIR Aging 2025;8:e67437
URL: https://aging.jmir.org/2025/1/e67437
doi:10.2196/67437

© Hao Ren, Yiying Zheng, Changjin Li, Fengshi Jing, Qiting Wang, Zeyu Luo, Dongxiao Li, Deyi Liang, Weiming Tang, Li Liu, Weibin Cheng. Originally published in JMIR Aging (https://aging.jmir.org), 30.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Relationship Between Sleep Disorders and Combination of Diabetes and Sarcopenia in Adults Aged 45 Years or Older: 10-Year Nationwide Prospective Cohort Study

Shugang Li[1], PhD; Yimi Wang[2], MBBS; Linzhi Li[2], MBBS; Hao Wu[1], MSci

[1]School of General Practice and Continuing Education, Capital Medical University, No.10 Xitoutiao, Youanmenwai, Fengtai District, Beijing, China
[2]School of Medicine, Shihezi University, Shihezi, China

Corresponding Author:
Hao Wu, MSci
School of General Practice and Continuing Education, Capital Medical University, No.10 Xitoutiao, Youanmenwai, Fengtai District, Beijing, China

Abstract

Background: With changes in lifestyle, the issue of sleep disorders is becoming increasingly common. Diabetes and sarcopenia have been found to be independently associated with sleep disorders. However, fewer studies have explored the interaction between the combination of diabetes and sarcopenia at different stages and sleep disorders.

Objective: This study aimed to explore the relationship between the combination of diabetes and sarcopenia and the incidence of sleep disorders in adults aged 45 years and older.

Methods: Based on data from the CHARLS (China Health and Retirement Longitudinal Study), we selected participants with comprehensive diagnostic information on diabetes and sarcopenia from 2011 who had normal sleep duration at baseline and checked their follow-up information of sleep duration from 2013, 2015, 2018, and 2020. Diabetes was classified into diabetes (D), prediabetes (PD), and nondiabetes (ND), and sarcopenia was divided into sarcopenia (S), possible sarcopenia (PS), and nonsarcopenia (NS). The participants were divided into DS, DPS, DNS, PDS, PDPS, PDNS, NDS, NDPS, and NDNS groups. Kaplan-Meier survival curves, the log-rank test, Cox proportional hazards regression, and restricted cubic spline models were used for statistical analysis.

Results: A total of 4936 participants were included in this study. The DS group had the highest incidence of sleep disorders: 49.32%, 28.57%, 36.36%, and 80.00% in 2013, 2015, 2018, and 2020 respectively. In the crude model, compared with the NDNS group, the risk of sleep disorders was increased in the DS group (hazard ratio [HR] 1.707, 95% CI 1.196 - 2.437), PDS (HR 1.599, 95% CI 1.235 - 2.071), NDS (HR 1.465, 95% CI 1.282 - 1.674), and DPS group (HR 1.318, 95% CI 1.097 - 1.583). The risk was increased but not statistically significant in the PDPS group (HR 1.160, 95% CI 0.987 - 1.365). After adjusting for covariates, the risk of sleep disorders remained statistically significant in the DS group (HR 1.515, 95% CI 1.059 - 2.167) and was significantly higher in the PDS (HR 1.423, 95% CI 1.096 - 1.847) and NDS (HR 1.279, 95% CI 1.113 - 1.468) groups than that in the NDNS group. The nonlinear associations between appendicular skeletal muscle mass, grip strength, 5-time chair test, fasting plasma glucose, and sleep disorders were observed and described.

Conclusions: The combination of diabetes and sarcopenia significantly increases the risk of sleep disorders in adults aged 45 years and older. and the implementation of progression control of both diabetes and sarcopenia may be helpful to prevent sleep disorders in this population.

(JMIR Aging 2025;8:e66372) doi:10.2196/66372

KEYWORDS

diabetes; sarcopenia; combined effect; sleep disorders; elderly; cohort study

Introduction

With changes in lifestyle, the issue of sleep disorders is becoming increasingly common. Studies report the worldwide prevalence of obstructive sleep apnea, poor sleep quality, other sleep problems, insomnia, and excessive daytime sleepiness to be 46%, 10%, 37%, 29%, and 19% respectively [1]. Even among adolescents, the prevalence of short and disturbed sleep is as high as 34.1% [2]. The impact of sleep disorders on the health of the older adults is more obvious, which can lead to cardiovascular disease [3], dementia [4], depression, and many other health problems. With the acceleration of aging, sleep disorders in older adults has become a global public health problem. A large number of studies have focused on the factors influencing sleep disorders. In addition to psychological [5], environmental [6], lifestyle [7] factors, chronic diseases, such

XSL•FO
RenderX

as diabetes [8] and sarcopenia [9] are found to be closely related to sleep disorders. According to the International Diabetes Federation report, a total of 537 million adults worldwide had diabetes in 2021, which is expected to increase by 46% to 783 million by 2045 [10]. The prevalence of diabetes in people aged ≥45 years in China is 10.6% [11], and there are significant differences between men and women, and individuals in urban and rural areas. With the increase in age, human function significantly declines, and coupled with the impact of metabolic diseases, sarcopenia is becoming more common in older adults. With a reported prevalence of 27% in individuals over 60 years of age [12], sarcopenia has become one of the common comorbidities in older adults.

Diabetes is an important factor affecting sleep disorders [13]. Studies have found a significant association between obstructive sleep apnea risk and type 2 diabetes (odds ratio [OR] 2.44, 95% CI 1.78-3.35; *P*<.001) [14]. Compared with patients without diabetes, both patients with type 1 (hazard ratio [HR] 1.41, 95% CI 1.37-1.46,) and type 2 diabetes (HR 1.40, 95% CI 1.37 - 1.44) have a significantly increased risk of sleep disorders [15]. In addition, studies have shown that sarcopenia is a chronic disease that induces sleep disorders [16], and people with both sarcopenia and diabetes have a significantly increased risk of sleep abnormalities [17]. Even self-reported sleep duration may provide the possibility for early prediction of sarcopenia. These findings suggest a complicated relationship between diabetes, sarcopenia, and sleep disorders.

However, these existing studies have not elucidated the risk of sleep disorders in populations with comorbid diabetes and sarcopenia; most studies have analyzed the relationships between diabetes and sleep disorders, as well as the impact of sarcopenia on sleep disorders. Another important issue is that both diabetes guidelines and the expert consensus on sarcopenia emphasize prediabetes and possible sarcopenia, which are similar to large icebergs beneath the surface—and their contribution to sleep disorders also warrants attention. With aging, the prevalence of diabetes and sarcopenia gradually increases in the older population, which prompts us to consider whether an interaction between diabetes and sarcopenia at different stages significantly increases the risk of sleep disorders. Previous studies have mostly explored the relationship between sarcopenia, diabetes, and sleep disorders using a cross-sectional design. On the one hand, there is a lack of evidence from longitudinal cohort designs; on the other hand, the interaction of the risk between diabetes and sarcopenia at different stages and sleep disorders has not been elucidated. Based on this scientific question, the present study used the China Health and Retirement Longitudinal Study (CHARLS) to analyze the impact of the interaction of diabetes and sarcopenia on the occurrence of sleep disorders over four longitudinal follow-up visits (ie, 2013, 2015, 2018, 2020) from the 2011 baseline population, to provide a reference for the prevention and treatment of sleep disorders in adults aged ≥45 years.

Methods

Participants and Study Design

The participants for this study were derived from the CHARLS, a nationally representative longitudinal survey initiated in 2011. It covered 150 county-level units, 450 village-level units, and 10,257 households across 28 provinces. These participants were followed up every 2 to 3 years. CHARLS covers basic personal information, household structure, health status, physical measurements, health service utilization, health insurance, and basic community characteristics. Four waves of follow-up data collection have been completed so far, and detailed information on CHARLS has been published [18].

Data were collected from CHARLS during 2011 - 2020. All information related to diabetes, sarcopenia, and sleep disorders among participants aged ≥45 years could be downloaded from 2011, 2013, 2015, 2018, and 2020.

We selected participants according to the following steps and excluded based on the following criteria: (1) lack of age information or age<45 years in 2011; 2) missing height and weight measurements at baseline; (3) height and weight data beyond mean (3 SD); (4) missing gender information; (5) missing sarcopenia-related information; (6) missing FPG (fasting plasma glucose) information; (7) FPG <3.9 mmol/L, (8) no sleep status at baseline, (9) nightly sleep duration <6 or >8 hours; (10) questionnaire lacked diabetes self-report information, (11) normal FPG but choosing a self-reported diagnosis of “diabetes or high blood sugar” for the question “Have you been diagnosed with by a doctor?”.

Subsequently, all the participants selected at baseline were matched by ID to the database of 2013, 2015, 2018, and 2020, respectively; those without necessary information about sleep duration were excluded. Only the first record of occurrence of sleep disorder was retained, if one ID was found repeatedly across years.

Measures

The measurements of exposure variables (ie, diabetes, sarcopenia), outcome variables (ie, sleep disorder, change of sleep duration or status), and covariates (ie, age, gender, education level, habitation, current marital status, current smoking, and drinking status) were respectively assessed across the cohort at each survey wave.

Assessment of Diabetes

All participants were classified as having diabetes, prediabetes, and nondiabetes according to the Guidelines for the Prevention and Treatment of type 2 Diabetes Mellitus in China (2020 edition) [19]. Since only FPG in venous blood was measured in the CHARLS, the diagnosis in this study was based only on FPG values. A venous FPG value ≥7.0 mmol/L was classified as diabetes, FPG ≥6.1 mmol/L and <7.0 mmol/L as prediabetes, and FPG <6.0 mmol/L as nondiabetes. Although the baseline survey included the question “Have you been diagnosed with by a doctor?”, the participants answering “diabetes or high blood sugar, (including impaired glucose tolerance and elevated fasting

XSL•FO
RenderX

blood sugar)," were excluded, as it was not possible to distinguish between diabetes and prediabetes.

Assessment of Sarcopenia

The Asian Working Group for Sarcopenia: 2019 Consensus Update on Sarcopenia Diagnosis and Treatment (AWGS 2019) was used to assess the status of the sarcopenia and classify participants into sarcopenia, possible sarcopenia, and nonsarcopenia groups. The average of maximum values of handgrip strength was measured using either or both hands to evaluate the skeletal muscle strength. According to AWGS 2019, the threshold for low handgrip strength is <18 kg for women and <28 kg for men. The 5-time chair stand test was used to evaluate physical performance of the participants; a time ≥12 s of the 5-time chair stand test was defined as low physical performance. Individuals who could not complete the test were excluded. The muscle mass of the participants was evaluated using a previously validated formula [20]:

ASM=0.193×body weight (kg)+0.107×height (cm)−4.157×sex (1=male, 2=female)−0.037×age (years)−2.631

Following the estimation of appendicular skeletal muscle mass index (ASM) using the above formula, the appendicular skeletal muscle mass index (ASMI) was calculated using the ASM divided by the square of height in meters (ASMI = ASM/height^2 m^2). Therefore, the cut-off values for ASMI used in this study were based on the percentile value of the lowest 20% of the study population; low muscle mass was defined as ASMI <7.05 kg/m^2 for men and <5.36 kg/m^2 for women. Sarcopenia was diagnosed if muscle mass, muscle strength, or physical performance were low, and possible sarcopenia was diagnosed if muscle strength or physical performance was low.

Covariates

Based on the literature review, some common factors correlated with sleep disorders, including demographic factors (eg, age and gender), area of residence (ie, city or rural), current marital status (ie, married or unmarried), and education level with a cut-off of junior high school graduation were selected as covariates to divide the participants into two groups. Health–related factors included current smoking and alcohol consumption status.

Statistical Analysis

The categorical variables such as gender, education level, and habitation were expressed as percentages. The χ^2 test was used for comparison between groups and Fisher exact test was used for analyses when the χ^2 test was not appropriate. For quantitative indicators such as age, FPG, and muscle mass index, the mean (SD) was used for normally distributed variables. When the variables were not normally distributed, median with interquartile range was used calculated. The 2-tailed *t* test, ANOVA, or Kruskal-Wallis tests were used for comparison among groups, as appropriate. Kaplan-Meier survival analysis and log-rank test were used to compare the incidence of sleep disorders. Based on the different adjustment covariates, three Cox proportional hazards regression models were used to analyze the combined effect of diabetes and sarcopenia on sleep disorders. The restricted cubic spline (RCS) models were used to analyze the exposure-response and nonlinear relationships between FPG, muscle mass index, grip strength, 5-time chair stand test, and sleep disorders. Microsoft Excel (2021), SPSS software (version 27.0; IBM Corp) and R software (version 4.3.2; R Foundation for Statistical Computing) were used to conduct data collation, statistical testing, and RCS modelling respectively. A $P<.05$ was considered statistically significant.

Ethical Considerations

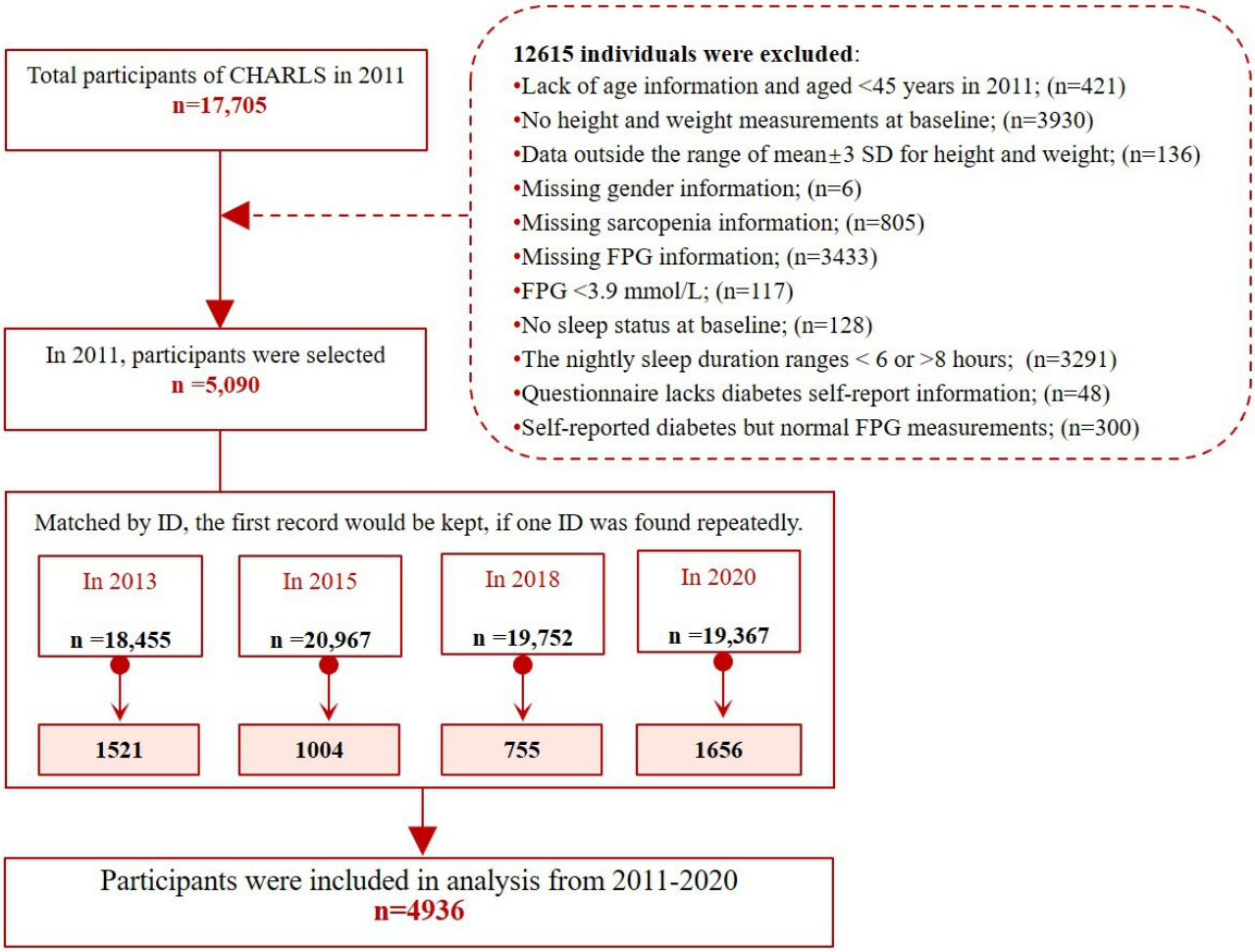
In this study, all data were extracted from CHARLS, which received approval from Peking University Biomedical Ethics Committee (IRB00001052-11015) prior to conducting the original field investigation and all the participants had signed the informed consent form before enrollment. The secondary analysis was approved and anonymized data was used for analysis.

Results

A total of 17,705 participants were included at baseline. Of these, 12,615 were gradually excluded based on the exclusion criteria, and matched with follow-up data from 2013, 2015, 2018, and 2020 according to the participant ID. The first occurrence of sleep disorders was retained in the database. Finally, 4936 participants were included in the analysis (Figure 1).

XSL•FO
RenderX

Figure 1. The flow chart of screening for enrolled individuals. CHARLS: China Health and Retirement Longitudinal Study; FPG: fasting plasma glucose.



A total of 4936 participants were selected and followed up on their sleep during 4 follow-up visits from 2011 to 2020, including 2427 male (49.2%) and 2509 female (50.8%) participants. Further comparison of the combination of diabetes and sarcopenia at different stages—NDNS, NDPS, NDS, PDNS, PDPS, PDS, DNS, DPS, and DS were statistically significant for age, sex, marital status, place of residence, education level, smoking, alcohol consumption, review status, weight, grip strength, skeletal muscle mass, 5-time chair stand test, and blood glucose levels *(*all *Ps*<.05; Table 1).

XSL•FO
RenderX

Table . Sociodemographic characteristics and critical variables of participants.

Variables	NDNS (Nondiabetes and Nonsarcopenia), n=2288	NDPS (Nondiabetes and Possible sarcopenia), n=841	NDS (Nondiabetes and Sarcopenia), n=324	PDNS (Prediabetes and Nonsarcopenia), n=599	PDPS (Prediabetes and Possible sarcopenia), n=230	PDS (Prediabetes and Sarcopenia), n=75	DNS (Diabetes and Nonsarcopenia), n=380	DPS (Diabetes and Possible sarcopenia), n=162	DS (Diabetes and Sarcopenia), n=37	H (*df*)/χ2 (*df*)	*P* value
Age (years), median (IQR)	55 (49-61)	58 (52-64)	68 (60-74)	57 (50-63)	60 (55-65)	68 (60-75)	57 (51-63)	59 (55-65)	70 (62-75)	528.833[a] (8)	<.001
Gender, n (%)										37.412[b] (8)	<.001
Male	1168 (51.0)	349 (41.5)	146 (45.1)	323 (53.9)	111 (48.3)	39 (52.0)	203 (53.4)	68 (42.0)	20 (54.1)		
Female	1120 (49.0)	492 (58.5)	178 (54.9)	276 (46.1)	119 (51.7)	36 (48.0)	177 (46.6)	94 (58.0)	17 (45.9)		
Marriage, n (%)											<.001[c]
Yes	2120 (92.7)	772 (91.8)	261 (80.6)	556 (92.8)	198 (86.1)	58 (77.3)	355 (93.4)	143 (88.3)	28 (75.7)		
No	168 (7.3)	69 (8.2)	63 (19.4)	43 (7.2)	32 (13.9)	17 (22.7)	25 (6.6)	19 (11.7)	9 (24.3)		
Habitation, n (%)											<.001[c]
City	2098 (91.7)	796 (94.6)	314 (96.9)	543 (90.7)	205 (89.1)	73 (97.3)	341 (90.0)	149 (92.0)	35 (94.6)		
Rural	190 (8.3)	45 (5.4)	10 (3.1)	56 (9.3)	25 (10.9)	2 (2.7)	38 (10.0)	13 (8.0)	2 (5.4)		
Education, n (%)										169.139[b] (8)	<.001
Above of junior high school graduation	946 (41.3)	225 (26.8)	45 (13.9)	235 (39.2)	68 (29.6)	12 (16.0)	159 (42.0)	40 (24.7)	6 (16.2)		
Others	1342 (58.7)	616 (73.2)	278 (86.1)	364 (60.8)	162 (70.4)	63 (84.0)	220 (58.0)	122 (75.3)	31 (83.8)		
Smoking	935 (40.9)	298 (35.5)	117 (36.1)	277 (46.2)	90 (39.1)	38 (50.7)	160 (42.1)	62 (38.3)	17 (45.9)	24.306[b] (8)	.002
Drinking	864 (37.8)	234 (27.8)	91 (28.1)	248 (41.4)	72 (31.3)	24 (32.0)	165 (43.4)	41 (25.3)	9 (24.3)		<.001
Height(cm), median (IQR)	159.5(153.4-165.5)	157.8(152.4-163.9)	154.9 (146.8-161.0)	159.5 (153.3-165.8)	158.5 (153.5-164.8)	154.0 (146.0-154.0)	159.5 (152.9-165.3)	156.0(151.7-162.3)	157.0(151.0-163.2)	138.643[a] (8)	<.001
Weight (kg), median (IQR)	58.7 (52.1-66.2)	60.2 (55.1-67.0)	45.6 (42.2-49.3)	61.0 (54.6-69.4)	63.3 (57.3-69.2)	46.6 (42.3-50.3)	61.6 (53.3-69.9)	61.4 (55.8-68.7)	46.4 (42.5-51.2)	878.573[a] (8)	<.001
Handgrip strength (kg), median (IQR)	34.0 (28.0-41.0)	26.5 (20.0-34.4)	22.5 (17.0-27.7)	34.0 (28.5-41.6)	27.5 (19.3-35.2)	22.6(17.5-29.5)	34.0 (28.1-41.3)	25.0 (18.5-32.5)	21.6 (17.0-29.3)	818.091[a] (8)	<.001
ASM[d] (kg/m^2), median (IQR)	17.8 (14.3-20.7)	16.7 (14.4-20.3)	12.9 (10.6-17.1)	18.1 (14.9-21.5)	18.2 (14.9-21.1)	13.9 (10.6-17.7)	18.3 (14.6-21.3)	16.7 (14.4-20.4)	14.7 (11.6-17.3)	343.836[a] (8)	<.001

XSL•FO
RenderX

Variables	NDNS (Nondiabetes and Nonsarcopenia), n=2288	NDPS (Nondiabetes and Possible sarcopenia), n=841	NDS (Nondiabetes and Sarcopenia), n=324	PDNS (Prediabetes and Nonsarcopenia), n=599	PDPS (Prediabetes and Possible sarcopenia), n=230	PDS (Prediabetes and Sarcopenia), n=75	DNS (Diabetes and Nonsarcopenia), n=380	DPS (Diabetes and Possible sarcopenia), n=162	DS (Diabetes and Sarcopenia), n=37	H (*df*)/χ2 (*df*)	*P* value
5-time chair stand test (s), median (IQR)	8.6 (7.2-10.2)	13.5 (12.3-15.4)	13.3 (11.2-15.5)	8.9 (7.4-10.3)	13.5 (12.2-16.3)	13.3 (11.3-16.3)	9.0 (7.5-10.3)	13.0 (12.3-15.4)	13.6 (11.2-17.3)	2336.735[a] (8)	<.001
FPG[e] (mmol/L), median (IQR)	5.5 (5.2-5.8)	5.5 (5.1-5.8)	5.4 (5.1-5.7)	6.4 (6.3-6.6)	6.4 (6.2-6.6)	6.4 (6.2-6.6)	8.0 (7.3-9.5)	8.0 (7.4-9.4)	7.8 (7.3-8.9)	3208.235[a] (8)	<.001

[a] Kruskal-Wallis H test.

[b]χ^2: chi-square statistic.

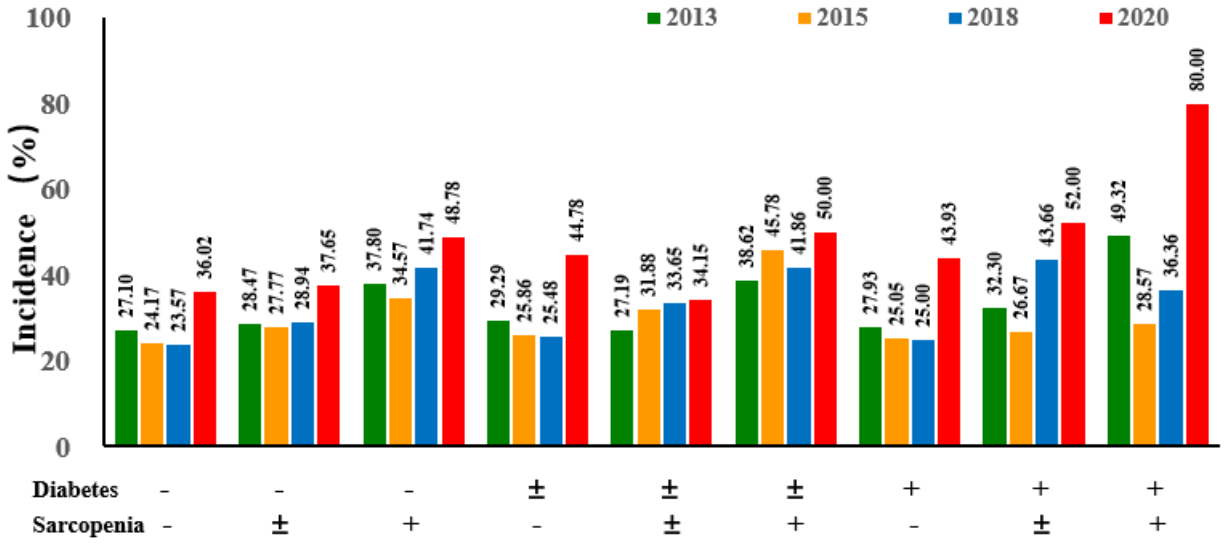
[c]Fisher's exact test.

[d]ASM: appendicular skeletal muscle mass.

[e]FPG: fasting plasma glucose.

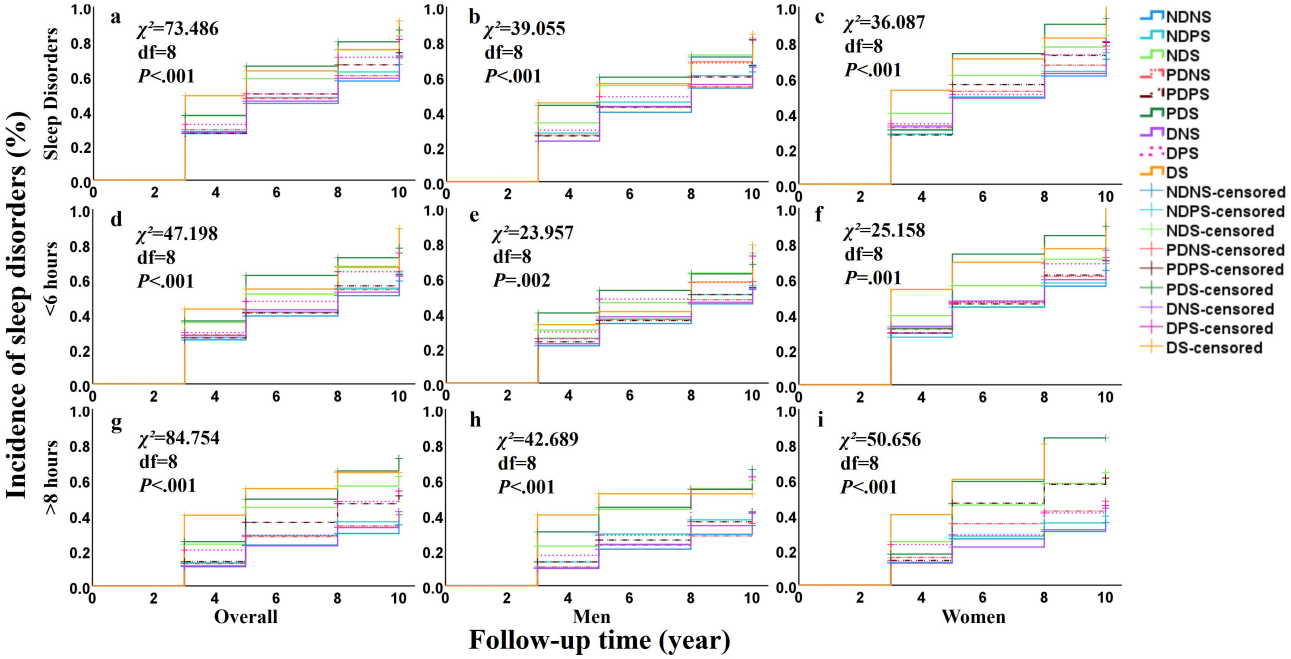
Based on survival analysis, it was observed that with the change in follow-up time, there was a significant difference in the incidence of sleep disorders among all groups (χ^{28}=73.486, *P*<.001; Figure 2). In 2020, the DS group had the highest incidence of sleep disorders (80. and the incidence in the DPS, PDS and NDS groups reached 52.00%, 50.00%, and 48.78%, respectively; whereas, the incidence of sleep disorders was the lowest in the NDNS group (36.02%). Furthermore, the incidence of sleep disorders in the DS group was 49.32%, 28.57%, 36.36%, and 80.00% in 2013, 2015, 2018, and 2020, respectively, showing a significantly increasing trend. Significant differences were also observed between men and women, and the incidence of short sleep duration (<6 hours) and long sleep duration (>8 hours), as shown in Figure 3.

Figure 2. The incidence of sleep disorders of the co-exposure of diabetes and sarcopenia at different stages. In the diabetes group, (+) represents diabetes, (±) represents prediabetes, and (-) depicts nondiabetes. In the sarcopenia group, (+) represents sarcopenia, (±) indicates possible sarcopenia, and (-) depicts nonsarcopenia.



XSL•FO
RenderX

Figure 3. Incidence of sleep disorders related to co-exposure of diabetes and sarcopenia among groups. Follow-up time depicts the number of years of follow-up in 2013, 2015, 2018, 2020, respectively. χ^2 indicates the statistic from the log-rank test. NDNS: Nondiabetes and Nonsarcopenia, NDPS: Nondiabetes and Possible Sarcopenia, NDS: Nondiabetes and Sarcopenia, PDNS: Prediabetes and Nonsarcopenia, PDPS: Prediabetes and Possible Sarcopenia, PDS: Prediabetes and Sarcopenia, DNS: Diabetes and Nonsarcopenia, DPS: Diabetes and Possible Sarcopenia, DS: Diabetes and Sarcopenia; SD: sleep disorders (including sleep duration <6h and >8h); <6 hours: sleep duration at the final follow-up was <6 hours; >8 hours: sleep duration at the final follow-up was >8 hours.



According to Cox regression analysis, there was a significant association between the combined status of diabetes mellitus, sarcopenia, and sleep disorders. In Model 1, no covariates were adjusted, compared with the NDNS group, the risk of sleep disorders was increased in the DS group (HR 1.707, 95% CI 1.196 - 2.437, *P*=.003), PDS (HR 1.599, 95% CI 1.235 - 2.071, *P*<.001), NDS (HR 1.465, 95% CI 1.282 - 1.674, *P*<.001), and DPS (HR 1.318, 95% CI 1.097 - 1.583, *P*=.003). The risk was increased but not statistically significant in the PDPS group. After adjusting for the variables (Model 3), the risk of sleep disorders in the DS group (HR 1.515, 95% CI 1.059 - 2.167; *P*=.02) was still statistically significant, and the risk in the PDS (HR 1.423, 95% CI 1.096 - 1.847; *P*=.008) and NDS (HR 1.279,95% CI 1.113 - 1.468; *P*<.001) groups were significantly higher than that in the NDNS group (Table 2).

XSL•FO
RenderX

Table . Association between sleep disorders and co-exposure to diabetes and sarcopenia at different status. Symbols indicate existing condition: diabetes (+), nondiabetes (–), prediabetes (±) in the diabetes group, and sarcopenia (+), possible sarcopenia (–), and no sarcopenia (±) in the sarcopenia group.

Group		Model 1[a], HR[b] (95% CI)	Model 2[c], HR (95% CI)	Model 3[d], HR (95% CI)
Diabetes	Sarcopenia			
Sleep disorders[e]				
-	-	Reference	Reference	Reference
-	±	1.105 (1.003-1.218)	1.058 (0.959-1.167)	1.033 (0.936-1.139)
-	+	1.465 (1.282-1.674)	1.344 (1.171-1.541)	1.279 (1.113-1.468)
±	-	1.107 (0.992-1.235)	1.100 (0.986-1.277)	1.098 (0.984-1.225)
±	±	1.160 (0.987-1.365)	1.114 (0.946-1.311)	1.110 (0.943-1.307)
±	+	1.599 (1.235-2.071)	1.501 (1.157-1.946)	1.423 (1.096-1.847)
+	-	1.067 (0.935-1.219)	1.061 (0.929-1.212)	1.070 (0.936-1.223)
+	±	1.318 (1.097-1.583)	1.251 (1.040-1.504)	1.228 (1.021-1.477)
+	+	1.707 (1.196-2.437)	1.594 (1.115-2.279)	1.515 (1.059-2.167)
Sleep duration <6 hours				
-	-	Reference	Reference	Reference
-	±	1.100 (0.981-1.234)	1.046 (0.932-1.174)	1.023 (0.911-1.149)
-	+	1.455 (1.233-1.716)	1.343 (1.134-1.591)	1.286 (1.084-1.525)
±	-	1.120 (0.985-1.273)	1.120 (0.985-1.273)	1.118 (0.983-1.271)
±	±	1.086 (0.884-1.334)	1.059 (0.861-1.302)	1.061 (0.863-1.305)
±	+	1.596 (1.139-2.235)	1.503 (1.071-2.110)	1.417 (1.008-1.992)
+	-	1.073 (0.915-1.258)	1.070 (0.913-1.255)	1.075 (0.916-1.261)
+	±	1.366 (1.099-1.697)	1.287 (1.035-1.601)	1.260 (1.012-1.568)
+	+	1.713 (1.123-2.612)	1.623 (1.062-2.481)	1.546 (1.010-2.365)
Sleep duration >8 hours				
-	-	Reference	Reference	Reference
-	±	1.208 (1.007-1.450)	1.137 (0.946-1.367)	1.073 (0.892-1.290)
-	+	2.132 (1.696-2.681)	1.791 (1.415-2.268)	1.601 (1.262-2.030)
±	-	1.197 (0.972-1.474)	1.173 (0.952-1.445)	1.160 (0.942-1.429)
±	±	1.594 (1.222-2.080)	1.465 (1.121-1.914)	1.429 (1.093-1.869)
±	+	2.545 (1.699-3.814)	2.256 (1.502-3.389)	2.033 (1.351-3.058)
+	-	1.198 (0.941-1.525)	1.162 (0.913-1.480)	1.195 (0.937-1.524)
+	±	1.665 (1.184-2.343)	1.516 (1.076-2.136)	1.502 (1.065-2.118)
+	+	2.739 (1.414-5.306)	2.340 (1.205-4.543)	1.970 (1.011-3.840)
Decreased sleep duration[f]				
-	-	Reference	Reference	Reference
-	±	1.122 (1.014-1.242)	1.073 (0.969-1.189)	1.052 (0.949-1.166)
-	+	1.368 (1.175-1.593)	1.274 (1.090-1.490)	1.224 (1.046-1.433)
±	-	1.076 (0.958-1.208)	1.069 (0.952-1.201)	1.067 (0.950-1.199)
±	±	1.121 (0.933-1.349)	1.091 (0.907-1.313)	1.101 (0.914-1.325)
±	+	1.735 (1.261-2.386)	1.617 (1.174-2.229)	1.525 (1.105-2.104)
+	-	1.063 (0.923-1.224)	1.058 (0.918-1.219)	1.069 (0.928-1.232)
+	±	1.251 (1.018-1.537)	1.188 (0.966-1.461)	1.167 (0.948-1.435)
+	+	1.546 (1.015-2.357)	1.452 (0.951-2.216)	1.391 (0.910-2.124)

XSL•FO
RenderX

Group		Model 1[a], HR[b] (95% CI)	Model 2[c], HR (95% CI)	Model 3[d], HR (95% CI)
Diabetes	Sarcopenia			
Increased sleep duration[g]				
-	-	Reference	Reference	Reference
-	±	1.253 (1.078-1.457)	1.212 (1.042-1.409)	1.167 (1.002-1.358)
-	+	1.674 (1.363-2.056)	1.468 (1.188-1.814)	1.385 (1.119-1.713)
±	-	1.115 (0.937-1.325)	1.099 (0.924-1.308)	1.081 (0.908-1.286)
±	±	1.377 (1.094-1.733)	1.295 (1.028-1.631)	1.259 (0.998-1.588)
±	+	1.952 (1.367-2.787)	1.815 (1.270-2.595)	1.697 (1.185-2.429)
+	-	1.104 (0.901-1.353)	1.078 (0.879-1.322)	1.088 (0.886-1.335)
+	±	1.270 (0.941-1.714)	1.196 (0.885-1.616)	1.201 (0.888-1.624)
+	+	2.067 (1.167-3.662)	1.852 (1.043-3.288)	1.641 (0.920-2.924)

[a]Model 1 represents no covariates were adjusted.

[b]HR: hazard ratio.

[c]Model 2 represents adjusting for gender and age.

[d]Model 3 represents adjustment for gender, age, education, current marital status and habitation, smoking, and drinking.

[e]Sleep disorders including sleep duration <6 hours and >8 hours.

[f]Decreased indicates difference of sleep duration between baseline and follow-up <0.

[g]Increased indicates the difference of sleep duration between baseline and follow-up >0.

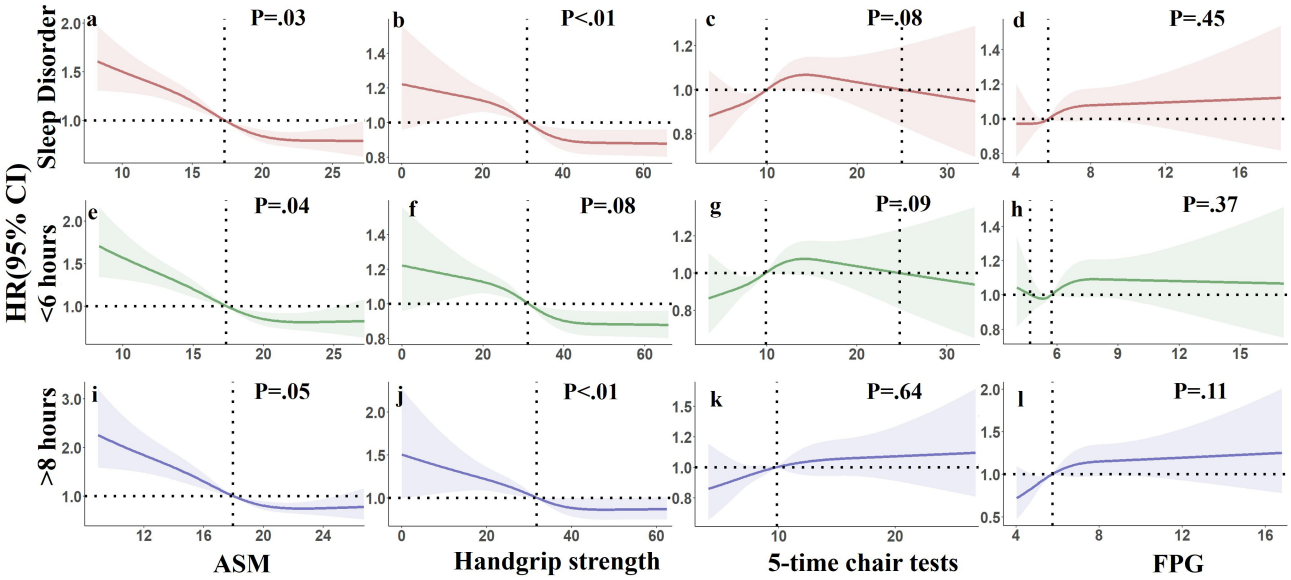
After a stratified analysis of sleep duration (ie, <6 hours and >8 hours), the risk of short sleep duration (< 6 hours) was significantly increased in the DS group without adjusting for variables (Model 1: HR 1.713, 95% CI 1.123 - 2.612) and remained statistically significant after adjusting for variables (Model 3: HR 1.546, 95% CI 1.010 - 2.365). The risk of long sleep duration (> 8 hours) was significantly increased in the DS group without adjusting for variables (Model 1: HR 2.739, 95% CI 1.414 - 5.306), which remained statistically significant after adjusting for variables (Model 3: HR 1.970, 95% CI 1.011 - 3.840; Table 2).

To analyze the changes in sleep duration, we stratified participants based on the difference in sleep duration between the final follow-up and baseline. If the difference was <0, it was defined as “Decrease,” while a difference >0, was defined as “Increase in sleep duration.” Cox regression analysis revealed that the risk of “Decrease” in sleep duration in the NDS group without adjusting for the variable (Model 1) is significantly higher (HR 1.735, 95% CI 1.261 - 2.386), and this remained statistically significant after adjusting for covariates (Model 3: (HR 1.525, 95% CI 1.105 - 2.104). After adjusting for covariates, the risk of “Increase” in the PDS (Model 3: HR 1.697, 95% CI 1.185 - 2.429) and NDS (Model 3: HR 1.385, 95% CI 1.119 - 1.713) groups also increased significantly (Table 2).

To further assess the relationship between diabetes, sarcopenia, and sleep disorders, the RCS model was used to describe important factors related to sarcopenia—ASM, handgrip strength, 5-time chair stand test performance—and the exposure response relationship between FPG levels, and sleep disorders. Figure 4 illustrates these association of sleep disorders with ASM (χ^2=45.04, P<.001, P for nonlinearity=.0284), handgrip strength (χ^2=16.62, P<.008, P for nonlinearity<.0017), and 5-time chair stand tests (χ^2=8.78, P=.03, P for nonlinearity=.0806). A nonlinear correlation was observed between sleep disorders and both ASM and handgrip strength. The relationship between FPG and long sleep duration (>8 hours) was also significant (χ^2=9.44, P=.02, P for nonlinearity=.10731).

XSL•FO
RenderX

Figure 4. Exposure-response relationship among diabetes, sarcopenia, and sleep disorders.. The restricted cubic spline (RCS) was used to model expose-response curves after adjusting for factors including gender, age, education, current marital status, and habitation, smoking, and drinking. ASM: appendicular skeletal muscle mass; FPG: fasting plasma glucose; HR: hazard ratio; SD: sleep disorders; <6 h: sleep duration was <6 hours; >8 h: sleep duration was >8 hours.



Discussion

With increasing age, the issue of sleep disorders in older adults is becoming more severe [21]. Common chronic conditions including diabetes mellitus and sarcopenia which are highly prevalent could affect sleep quality. In particular, prediabetes and possible sarcopenia are now being diagnosed at an early age. If this status is not addressed during reversible stages, the risk of sleep disorders in adults aged ≥45 years is even more severe. Although studies have reported a relationship between diabetes [22] or sarcopenia [23] and sleep disorders, there is a lack of studies on the relationship between co-exposure at varying stages of diabetes and sarcopenia. This study found that participants with diabetes and sarcopenia had the highest risk of sleep disorders; groups with prediabetes and sarcopenia, nondiabetes with sarcopenia, and diabetes with possible sarcopenia had a significantly higher risk of sleep disorders. Additionally, there was an exposure-response relationship between ASM, handgrip strength, 5-time chair test, FPG levels, and sleep disorders. These findings can provide a reference for the prevention of sleep disorders in adults aged ≥45 years and the development of early screening and personalized intervention.

Diabetes, as a chronic disease with high incidence in older adults, is an important factor contributing to sarcopenia [24] and sleep disorders [25]. Our findings showed that participant groups with both diabetes and prediabetes have an increased risk of sleep disorders. This may be explained by the following mechanisms. On one hand, diabetes can affect the function of the hypothalamic-pituitary-adrenal axis, resulting in abnormal secretion of hormones such as cortisol [26]. The rhythm of these hormones can affect the sleep-wake cycle, leading to the sleep disorders [27]. On the other hand, hyperglycemia may lead to the secretion of melatonin, which plays a key role in regulating sleep rhythms [28]. Its altered secretion can lead to difficulty falling asleep and poor sleep quality [29]. In addition, patients with sarcopenia are prone to muscle fatigue and damage due to the loss of mass and strength of several muscles [30]. During sleep, changes in body posture may strain already fragile muscles, causing pain and discomfort, resulting in sleep disruption. Muscle strength in the chest and abdomen decreases in patients with sarcopenia, resulting in decreased respiratory function [31]. During sleep, weakness in respiratory muscles may cause poor breathing, snoring and even apnea, thus interfering with sleep [32]. It is also important to note that diabetes can cause sarcopenia. Long-term hyperglycemia may lead to peripheral neuropathy [33], causing a decrease in muscle blood supply, affecting the nutrition and innervation of nerves to muscles [34], which in turn leads to muscle waste atrophy. Patients with diabetes and prediabetes have widespread insulin resistance, which blocks insulin signaling pathways such as PI3K/Ak [35] and Ras-MAPK [36], thus cannot effectively promote the uptake of amino acids by muscle cells, thereby inhibiting the synthesis of muscle proteins; this results in a decrease in muscle mass. These mechanisms reveal the relationship between FPG, ASM, handgrip strength, 5-time chair stand test, and sleep disorders, as identified by the RCS model.

This study found that participants with comorbid diabetes and sarcopenia had the highest risk of sleep disorders, and the groups with prediabetes and sarcopenia, nondiabetes and sarcopenia, and diabetes with possible sarcopenia also had a significantly higher risk of sleep disorders. The potential mechanisms of sleep disorders caused by the comorbidity of diabetes and sarcopenia may be related to the following: (1) patients with diabetes already have abnormal blood sugar regulation; sarcopenia will further affect the uptake and utilization of glucose by muscles, resulting in more frequent and difficult-to-control blood sugar fluctuations [37]. This may result in the superposition of metabolic disorders, triggering energy metabolism imbalance [38], making the body unable to obtain enough energy at night to maintain normal physiological

XSL•FO
RenderX

functions and sleep status. (2) Neuropathy caused by diabetes, including damage to peripheral and autonomic nerves can affect the sensory and motor functions of muscles [39]. Sarcopenia aggravates muscle weakness and atrophy, further impairing the function of the neuromuscular junction [40]. This neuromuscular dysfunction can lead to nocturnal muscle twitches, spasms, or pain that interfere with sleep. Autonomic dysfunction is common in both diabetes and sarcopenia; it affects the regulation of the cardiovascular and respiratory systems, resulting in abnormal fluctuations in blood pressure and heart rate at night, poor breathing, and increases the risk of sleep disruption [41]. Both diabetes and sarcopenia are associated with abnormal hormone secretion [42], which may lead to reduced secretion of testosterone, estrogen, and insulin-like growth factor 1 (IGF-1) among others. Significant gender differences were observed in our study, which may further support this phenomenon. Furthermore, diabetes and sarcopenia are both accompanied by chronic inflammation and increased levels of oxidative stress [43]. Tumor necrosis factor α (TNF-α), interleukin 6 (IL-6) promote muscle catabolism and may affect the central nervous system function, resulting in disorders in sleep regulation centers. Reactive oxygen species produced during oxidative stress can damage both neural and muscular systems, aggravate physical discomfort, and then affect sleep [44].

Our findings have several implications. First, special attention should be paid to adults aged ≥45 years with sleep disorders. To improve sleep duration and quality, it is necessary to effectively manage diabetes, perform physical exercises to improve muscle mass and function, and continuously monitor sleep quality. Second, for people with prediabetes or possible sarcopenia, early intervention should be carried out to prevent the occurrence of diabetes or sarcopenia, which can effectively reduce the risk of sleep disorders. Third, the community should build appropriate sports facilities and provide an environment that promotes physical exercise to improve the health of residents. In addition, regular screening for diabetes and sarcopenia should be implemented to detect patients with diabetes and sarcopenia at an early stage, and comprehensive interventions should be actively used to prevent sleep disorders.

The advantages of this study include the following aspects. It is a large-scale, nationally representative cohort study of community residents, using survival analysis and Cox regression models to explore the co-exposure of diabetes and sarcopenia—as well as the impact of prediabetes with latent risk and possible sarcopenia—on sleep duration and sleep status. However, there are also a few limitations within our study: First, due to limitations in the CHARLS data, we could not fully explore the relationship between diabetes status, progression trends in sarcopenia from 2011 to 2020, and specific types of sleep disorders. Our analysis was limited to examining the impact of diabetes and sarcopenia status on sleep disorders in 2011; the relationship between the factors affecting progression and sleep disorders is complex. Concurrently, due to limited data, we included a smaller number of participants aged ≥75 years in our study, and the differences in population distribution across age groups may limit the extrapolation of our findings. Second, prediabetes is an intermediate stage that can revert toward normoglycemia and diabetes. Therefore, our study only analyzed the effect of prediabetes exposure at baseline on sleep disorders. Since the CHARLS did not provide data on FPG levels for the four follow-up waves , it was impossible to explore the role of changes in prediabetes status on sleep disorders. However, our findings still suggest that patients with prediabetes were at an increased risk of sleep disorders. Therefore, family-based and both individual health education [45], and community-level behavioral interventions play an important role in the prevention of sleep disorders [46]. Fourth, there is a mutual relationship among diabetes, sarcopenia, and sleep disorders. Although both diabetes and sarcopenia can lead to sleep disorders, the interaction between diabetes and sarcopenia on sleep disorders could not be analyzed due to lack of sufficient data. Fifth, this study found a potential relationship between the co-exposure to various diabetes and sarcopenia statuses, and sleep disorders. Diabetes and sarcopenia may have a synergistic effect on the occurrence of sleep disorders; however, causality cannot be based solely on our findings. Future research should focus on the progression and changes of diabetes, sarcopenia, and sleep disorders, including sleep efficiency [47] to explore the relationship among them.

In conclusion, this study found that diabetes combined with sarcopenia in adults aged 45 years and older may increase the risk of sleep disorders. To reduce this risk, it is essential to implement effective disease management for both diabetes and sarcopenia, and to further evaluate joint intervention measures for individuals with prediabetes or possible sarcopenia.

Acknowledgments

This study was funded by the Beijing High level Public Health Technical Talents Training Plan (2022-1-005, Key DisciplineMember-02- ‐ 44). We are grateful to the CHARLS (China Health and Retirement Longitudinal Study), which provided data for this research. We also thank all the participants, staff members, volunteers devoted to the CHARLS project.

Authors' Contributions

Conceptualization: HW, SL
Data curation: SL, YW
Formal analysis: SL, YW
Funding acquisition: HW, SL
Investigation: SL, YW
Methodology: SL, YW

XSL•FO
RenderX

Supervision: HW, SL
Validation: LL
Writing – original draft: SL, YW
Writing – review & editing: HW, LL, SL, YW

Conflicts of Interest

All authors declare no conflicts of interest.

References

1. Canever JB, Zurman G, Vogel F, et al. Worldwide prevalence of sleep problems in community-dwelling older adults: a systematic review and meta-analysis. Sleep Med 2024 Jul;119:118-134. [doi: 10.1016/j.sleep.2024.03.040] [Medline: 38669835]
2. Yue L, Cui N, Liu Z, Jia C, Liu X. Patterns of sleep problems and internalizing and externalizing problems among Chinese adolescents: a latent class analysis. Sleep Med 2022 Jul;95:47-54. [doi: 10.1016/j.sleep.2022.04.008] [Medline: 35561474]
3. Liu SH, Lin FJ, Kao YH, et al. Chronic partial sleep deprivation increased the incidence of atrial fibrillation by promoting pulmonary vein and atrial arrhythmogenesis in a rodent model. Int J Mol Sci 2024 Jul 11;25(14):7619. [doi: 10.3390/ijms25147619] [Medline: 39062858]
4. Robbins R, Quan SF, Weaver MD, Bormes G, Barger LK, Czeisler CA. Examining sleep deficiency and disturbance and their risk for incident dementia and all-cause mortality in older adults across 5 years in the United States. Aging (Albany NY) 2021 Feb 11;13(3):3254-3268. [doi: 10.18632/aging.202591] [Medline: 33570509]
5. Haddadi A, Matinnia N, Yazdi-Ravandi S. The relationship between corona disease anxiety and sleep disturbances and suicidal ideation in medical staff: The mediating role of resiliency and cognitive flexibility: a cross-sectional study. Health Sci Rep 2024 Aug;7(8):e2282. [doi: 10.1002/hsr2.2282] [Medline: 39144407]
6. Wang J, Gueye-Ndiaye S, Castro-Diehl C, et al. Associations between indoor fine particulate matter (PM2.5) and sleep-disordered breathing in an urban sample of school-aged children. Sleep Health 2024 Aug 1:S2352-7218(24)00133-5. [doi: 10.1016/j.sleh.2024.06.004] [Medline: 39095254]
7. Zheng YB, Huang YT, Gong YM, et al. Association of lifestyle with sleep health in general population in China: a cross-sectional study. Transl Psychiatry 2024 Aug 4;14(1):320. [doi: 10.1038/s41398-024-03002-x] [Medline: 39098892]
8. Algethami A, Alfahmi FK, Alhusayni MA, et al. Evaluation of sleep quality among people living with type 2 diabetes mellitus in Taif, Saudi Arabia: a cross-sectional study. Cureus 2024 Jul;16(7):e64124. [doi: 10.7759/cureus.64124] [Medline: 39119415]
9. Liu K, Luo J, Chen Y, et al. Association between sarcopenia and sleep disorders: a cross-sectional population based study. Front Nutr 2024;11:1415743. [doi: 10.3389/fnut.2024.1415743]
10. Diabetes around the world in 2021. International Diabetes Federation. URL: https://diabetesatlas.org/ [accessed 2025-08-05]
11. Lingling G, Lin Y, Jiao Y. Prevalence and influencing factors of diabetes for the 45 years and older population in China. Modern Preventive Medicine 2019;46:2890-2894 [FREE Full text]
12. Petermann-Rocha F, Balntzi V, Gray SR, et al. Global prevalence of sarcopenia and severe sarcopenia: a systematic review and meta-analysis. J Cachexia Sarcopenia Muscle 2022 Feb;13(1):86-99. [doi: 10.1002/jcsm.12783] [Medline: 34816624]
13. Woo J, Lehrer HM, Tabibi D, Cebulske L, Tanaka H, Steinhardt M. The association of multidimensional sleep health with hba1c and depressive symptoms in African American Adults with type 2 diabetes. Psychosom Med 2024 May 1;86(4):307-314. [doi: 10.1097/PSY.0000000000001298] [Medline: 38724038]
14. Taimah M, Ahmad A, Al-Houqani M, et al. Association between obstructive sleep apnea risk and type 2 diabetes among Emirati adults: results from the UAE healthy future study. Front Endocrinol (Lausanne) 2024;15:1395886. [doi: 10.3389/fendo.2024.1395886] [Medline: 39081790]
15. Ni MH, Yang YS, Huang JY, Lo SC, Huang CN, Kornelius E. The association of depression and sleep disorders in patients with type 1 diabetes in Taiwan. Medicine (Baltimore) 2024 Jul 19;103(29):e38969. [doi: 10.1097/MD.0000000000038969] [Medline: 39029051]
16. Lv X, Peng W, Jia B, Lin P, Yang Z. Longitudinal association of sleep duration with possible sarcopenia: evidence from CHARLS. BMJ Open 2024 Mar 23;14(3):e079237. [doi: 10.1136/bmjopen-2023-079237] [Medline: 38521528]
17. Ida S, Kaneko R, Nagata H, et al. Association between sarcopenia and sleep disorder in older patients with diabetes. Geriatr Gerontol Int 2019 May;19(5):399-403. [doi: 10.1111/ggi.13627] [Medline: 30773802]
18. Zhao Y, Strauss J, Yang G, Giles J. 2011-2012 national baseline user's guide [chinese]. China Health and Retirement Longitudinal Study. 2013. URL: https://charls.charlsdata.com/Public/ashelf/public/uploads/document/2011-charls-wave1/application/Chinese_users__guide_20130407_.pdf [accessed 2024-07-20]
19. Chinese Diabetes Society, Zhu D. Guideline for the prevention and treatment of type 2 diabetes mellitus in China (2020 edition) [Chinese]. Chin J Diabetes Mellitus 2021;13(4):315-409 [FREE Full text] [doi: 10.3969/j.issn.1006-6187.2022.01.002]
20. Wen X, Wang M, Jiang CM, Zhang YM. Anthropometric equation for estimation of appendicular skeletal muscle mass in Chinese adults. Asia Pac J Clin Nutr 2011;20(4):551-556. [Medline: 22094840]

XSL•FO
RenderX

21. Delbari A, Ahmadi F, Zar A, Zandvakili A, Sadeghipour HR, Sims J. Living in urban or rural environments affect the sleep quality of the elderly in Bushehr (Southern Iran): emphasizing the active and inactive of the elderly. BMC Public Health 2024 May 18;24(1):1340. [doi: 10.1186/s12889-024-18747-9] [Medline: 38760757]
22. Tenda ED, Henrina J, Cha JH, et al. Obstructive sleep apnea: overlooked comorbidity in patients with diabetes. World J Diabetes 2024 Jul 15;15(7):1448-1460. [doi: 10.4239/wjd.v15.i7.1448] [Medline: 39099813]
23. Smith L, Shin JI, Veronese N, et al. Sleep duration and sarcopenia in adults aged ≥ 65 years from low and middle-income countries. Aging Clin Exp Res 2022 Jul;34(7):1573-1581. [doi: 10.1007/s40520-022-02074-3] [Medline: 35103953]
24. Chowdhury SR, Chakrabarti A, Datta PK. Sarcopenia in patients with diabetes mellitus, an overlooked perioperative condition. Br J Anaesth 2023 Sep;131(3):e65-e67. [doi: 10.1016/j.bja.2023.06.032]
25. Kuroda H, Yeung SLA, Fujii R, Iwagami M, Goto A. Investigating the non-linear association between sleep duration and type 2 diabetes: conventional and Mendelian randomization analyses from the UK Biobank. J Diabetes Investig 2025 Apr 3:40181521. [doi: 10.1111/jdi.70039] [Medline: 40181521]
26. Marissal-Arvy N, Moisan MP. Diabetes and associated cognitive disorders: role of the hypothalamic-pituitary adrenal axis. Metabol Open 2022 Sep;15:100202. [doi: 10.1016/j.metop.2022.100202] [Medline: 35958117]
27. Ba-Ali S, Brøndsted AE, Andersen HU, Sander B, Jennum PJ, Lund-Andersen H. Assessment of diurnal melatonin, cortisol, activity, and sleep-wake cycle in patients with and without diabetic retinopathy. Sleep Med 2019 Feb;54:35-42. [doi: 10.1016/j.sleep.2018.10.018] [Medline: 30529775]
28. Onaolapo AY, Adebisi EO, Adeleye AE, Olofinnade AT, Onaolapo OJ. Dietary melatonin protects against behavioural, metabolic, oxidative, and organ morphological changes in mice that are fed high-fat, high-sugar diet. Endocr Metab Immune Disord Drug Targets 2020;20(4):570-583. [doi: 10.2174/1871530319666191009161228] [Medline: 32138638]
29. van Maanen A, Meijer AM, Smits MG, van der Heijden KB, Oort FJ. Effects of melatonin and bright light treatment in childhood chronic sleep onset insomnia with late melatonin onset: a randomized controlled study. Sleep 2017 Feb 1;40(2):28364493. [doi: 10.1093/sleep/zsw038] [Medline: 28364493]
30. Zhang Y, Chen X, Hou L, et al. Prevalence and risk factors governing the loss of muscle function in elderly sarcopenia patients: a longitudinal study in China with 4 years of follow-up. J Nutr Health Aging 2020;24(5):518-524. [doi: 10.1007/s12603-020-1361-0] [Medline: 32346691]
31. Nagano A, Wakabayashi H, Maeda K, et al. Respiratory sarcopenia and sarcopenic respiratory disability: concepts, diagnosis, and treatment. J Nutr Health Aging 2021;25(4):507-515. [doi: 10.1007/s12603-021-1587-5] [Medline: 33786569]
32. Peng X, Zhou R, Liu C, Chen X, Zhu T, Chen G. Abnormal sleep duration is associated with sarcopenia in older Chinese people: A large retrospective cross-sectional study. Open Med (Wars) 2024;19(1):20240938. [doi: 10.1515/med-2024-0938] [Medline: 38584821]
33. Albariqi MM, Versteeg S, Brakkee EM, et al. Human IAPP is a contributor to painful diabetic peripheral neuropathy. J Clin Invest 2023 Apr 17;133(8):e156993. [doi: 10.1172/JCI156993] [Medline: 36917177]
34. van Sloten TT, Savelberg HHCM, Duimel-Peeters IGP, et al. Peripheral neuropathy, decreased muscle strength and obesity are strongly associated with walking in persons with type 2 diabetes without manifest mobility limitations. Diabetes Res Clin Pract 2011 Jan;91(1):32-39. [doi: 10.1016/j.diabres.2010.09.030] [Medline: 20965601]
35. Feng Y, Ren Y, Zhang X, et al. Metabolites of traditional Chinese medicine targeting PI3K/AKT signaling pathway for hypoglycemic effect in type 2 diabetes. Front Pharmacol 2024;15:1373711. [doi: 10.3389/fphar.2024.1373711]
36. Kasowanjete P, Abrahamse H, Houreld NN. Photobiomodulation at 660 nm stimulates in vitro diabetic wound healing via the Ras/MAPK Pathway. Cells 2023 Apr 4;12(7):1080. [doi: 10.3390/cells12071080] [Medline: 37048153]
37. Norlin S, Axelsson J, Ericsson M, Edlund H. O304 ameliorates hyperglycemia in mice by dually promoting muscle glucose effectiveness and preserving β-cell function. Commun Biol 2023 Aug 25;6(1):877. [doi: 10.1038/s42003-023-05255-6] [Medline: 37626210]
38. Al-Sayyar A, Hammad MM, Williams MR, Al-Onaizi M, Abubaker J, Alzaid F. Neurotransmitters in type 2 diabetes and the control of systemic and central energy balance. Metabolites 2023 Mar 4;13(3):384. [doi: 10.3390/metabo13030384] [Medline: 36984824]
39. Zilliox LA. Diabetes and peripheral nerve disease. Clin Geriatr Med 2021 May;37(2):253-267. [doi: 10.1016/j.cger.2020.12.001] [Medline: 33858608]
40. Miao Y, Xie L, Song J, et al. Unraveling the causes of sarcopenia: roles of neuromuscular junction impairment and mitochondrial dysfunction. Physiol Rep 2024 Jan;12(1):e15917. [doi: 10.14814/phy2.15917] [Medline: 38225199]
41. Mizukami K. Autonomic dysfunction in dementia with Lewy bodies: focusing on cardiovascular and respiratory dysfunction. PCN Rep 2023 Sep;2(3):e129. [doi: 10.1002/pcn5.129] [Medline: 38867816]
42. He Y, Fu Q, Sun M, et al. Phosphoproteome reveals molecular mechanisms of aberrant rhythm in neurotransmitter-mediated islet hormone secretion in diabetic mice. Clin Transl Med 2022 Jun;12(6):e890. [doi: 10.1002/ctm2.890] [Medline: 35758323]
43. Muvhulawa N, Mazibuko-Mbeje SE, Ndwandwe D, et al. Sarcopenia in a type 2 diabetic state: reviewing literature on the pathological consequences of oxidative stress and inflammation beyond the neutralizing effect of intracellular antioxidants. Life Sci 2023 Nov 1;332:122125. [doi: 10.1016/j.lfs.2023.122125] [Medline: 37769808]
44. Singh G, Singh K, Sinha RA, Singh A, Kumar A. Japanese encephalitis virus infection causes reactive oxygen species - mediated skeletal muscle damage. Eur J of Neuroscience 2024 Sep;60(5):4843-4860. [doi: 10.1111/ejn.16469]

XSL•FO
RenderX

45. Amerzadeh M, Shafiei Kisomi Z, Senmar M, Khatooni M, Hosseinkhani Z, Bahrami M. Self-care behaviors, medication adherence status, and associated factors among elderly individuals with type 2 diabetes. Sci Rep 2024 Aug 18;14(1):19118. [doi: 10.1038/s41598-024-70000-w] [Medline: 39155329]
46. Van Ancum JM, Meskers CGM, Reijnierse EM, et al. Lack of knowledge contrasts the willingness to counteract sarcopenia among community-dwelling adults. J Aging Health 2020;32(7-8):787-794. [doi: 10.1177/0898264319852840] [Medline: 31156027]
47. Sakal C, Li T, Li J, Yang C, Li X. Association between sleep efficiency variability and cognition among older adults: cross-sectional accelerometer study. JMIR Aging 2024 Apr 4;7:e54353. [doi: 10.2196/54353] [Medline: 38596863]

Abbreviations

ASM: appendicular skeletal muscle mass
ASMI: appendicular skeletal muscle mass index
AWGS: Asian Working Group for Sarcopenia: 2019 Consensus Update on Sarcopenia Diagnosis and Treatment
CHARLS: China Health and Retirement Longitudinal Study
FPG: fasting plasma glucose
HR: hazard ratio
RCS: restricted cubic spline

Edited by H Köttl; submitted 11.09.24; peer-reviewed by WQ Chen, Z Yang; revised version received 13.04.25; accepted 17.04.25; published 02.06.25.

Please cite as:
Li S, Wang Y, Li L, Wu H
The Relationship Between Sleep Disorders and Combination of Diabetes and Sarcopenia in Adults Aged 45 Years or Older: 10-Year Nationwide Prospective Cohort Study
JMIR Aging 2025;8:e66372
URL: https://aging.jmir.org/2025/1/e66372
doi:10.2196/66372

© Shugang Li, Yimi Wang, Linzhi Li, Hao Wu. Originally published in JMIR Aging (https://aging.jmir.org), 2.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Sarcopenia and Risk of Cognitive Impairment: Cohort Study and Mendelian Randomization Analysis

Tingting Sha[1,2,3], PhD; Yuqing Zhang[4,5], DSc; Jie Wei[1,2,3,6,7], PhD; Changjun Li[3,6,8], MD, PhD; Chao Zeng[1,2,3,6,7], MD, PhD; Guanghua Lei[1,2,3,6], MD, PhD; Yilun Wang[1,2,3], MD, PhD

[1]Department of Orthopaedics, Xiangya Hospital, Central South University, 87 Xiangya Road, Changsha, China

[2]Hunan Key Laboratory of Joint Degeneration and Injury, Changsha, China

[3]Key Laboratory of Aging-related Bone and Joint Diseases Prevention and Treatment of Ministry of Education, Changsha, China

[4]Division of Rheumatology, Allergy, and Immunology, Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, United States

[5]The Mongan Institute, Massachusetts General Hospital, Harvard Medical School, Boston, MA, United States

[6]National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, China

[7]Department of Epidemiology and Health Statistics, Xiangya School of Public Health, Central South University, Changsha, China

[8]Department of Endocrinology, Endocrinology Research Center, Xiangya Hospital of Central South University, Changsha, China

Corresponding Author:
Yilun Wang, MD, PhD
Department of Orthopaedics, Xiangya Hospital, Central South University, 87 Xiangya Road, Changsha, China

Abstract

Background: Over half the people over 60 years of age experience cognitive impairment, with limited treatment options, making it crucial to identify risk factors. Studies have examined the association between sarcopenia and cognitive impairment; however, the evidence is inconclusive and cannot be used to make causal inferences.

Objective: This study aims to appraise the causal association of sarcopenia with cognitive impairment by triangulating the data from a cohort study and Mendelian randomization (MR) analysis.

Methods: Using UK Biobank data, we first examined the associations of sarcopenia and its indices (appendicular lean mass [ALM], handgrip strength, and gait speed) with cognitive function (fluid intelligence and prospective memory) by using mixed-effects regression models. Then, we explored the causal associations of genetically predicted sarcopenic indices with cognitive function through a 2-sample MR, and examined potential mediation by omega-3 fatty acids, vitamin D levels, physical inactivity, falls, frailty, sleep disorders, anxiety, depression, stroke, metabolic syndrome, and type 2 diabetes.

Results: A total of 34,457 participants, with a mean age of 56.4 (SD 7.6) years, 51.1% (n=17,620) of which were female, completed baseline cognitive tests between 2006 and 2010 and attended at least 1 follow-up visit in 2012, 2014, or 2019, and were included in the observational analysis. The cohort study revealed that sarcopenia was significantly associated with cognitive impairment, which was evidenced by reduced fluid intelligence scores (β=−0.91, 95% CI −1.68 to −0.15; *P*=.02). Each of the sarcopenic indices also exhibited significant associations with either fluid intelligence or prospective memory (all *P*<.05). MR analyses yielded compelling evidence of positive associations between the genetically predicted increases in ALM (β=0.09, 95% CI 0.07-0.12; *P*<.001), handgrip strength (β=0.18, 95% CI 0.08-0.29; *P*<.001) and gait speed (β=0.78, 95% CI 0.53-0.29; *P*<.001) and improved cognitive function. The effects of ALM and handgrip strength on cognitive function were partially mediated by genetically predicted physical activity, with indirect effects of 0.01 (95% CI 0.00-0.02) for ALM and 0.02 (95% CI 0.00-0.05) for handgrip strength.

Conclusions: Our study suggests that sarcopenia is a potential causal risk factor for cognitive impairment, with physical activity acting as a modifiable mediator in this relationship.

(JMIR Aging 2025;8:e66031) doi:10.2196/66031

KEYWORDS

sarcopenia; musculoskeletal disease; cognitive impairment; cognition; mediation; Mendelian randomization; genetic variation; cohort study

XSL•FO
RenderX

Introduction

An impairment in cognitive function is a neurodegenerative process that affects 10% - 20% of people aged 65 years and older [1], and that can lead to adverse health consequences, including diminished quality of life [2] and increased risk of hospitalization [3], as well as mortality [4]. However, by the time of diagnosis, the pathological changes related to cognitive impairment have often become irreversible. The available treatments are limited and target symptoms, with very low efficacy [5]. Thus, identifying the potential factors for predicting the risk of subsequent cognitive impairment has become a public health priority [6,7]. Such measurable indices can help recognize high-risk populations to explore potential disease-modifying therapies.

Sarcopenia, characterized by accelerated loss of muscle mass and deterioration in function, is a prevalent skeletal muscle disorder in the elderly [8]. As predicted by statistics, about 2 billion individuals globally will be affected by sarcopenia by 2050 [9]. Through the progression of physical inactivity, patients with sarcopenia tend to experience an increased risk of disability, which has been associated with cognitive impairment due to its impact on neurogenesis and cerebral blood vessel formation [10,11]. Previous cohort studies have reported a possible relationship between sarcopenia and cognitive impairment [12-17]. However, this relationship is still not fully understood and may even be controversial [18]. Some studies have even rejected any significant association between sarcopenia itself [19], or its related indices [20,21] and cognitive impairment. Discrepancies in definitions of sarcopenia and sample sizes among studies may underlie these conflicting outcomes. Hence, a more intricate exploration is warranted to establish the precise link between sarcopenia and cognitive impairment.

Furthermore, current available human evidence on such association is mostly based on observational research, and therefore cannot be relied on to derive causal inferences due to inherent limitations (eg, reverse causality and unmeasured confounders) [22]. Examining the causal association between sarcopenia and cognitive impairment and identifying potentially modifiable intervention targets along the causal pathway is of important public health significance for preventing cognitive impairment [18]. Mendelian randomization (MR) is a method that involves using genetic variants as proxies for the targeted exposure. Since, genetic variants are conditional on parental genotypes and randomly allocated at conception, the results from MR studies are more resistant to reverse causality and confounding than those derived from conventional observational studies [22].

Using a large population-based cohort, this study was initiated by evaluating the observational data that links the consensus definition of sarcopenia, which was established by the European Working Group on Sarcopenia in Older People in 2019 (EWGSOP2), as well as its 3 defining indices (ie, appendicular lean mass [ALM], handgrip strength, and gait speed), to the risk of cognitive impairment. To best assess causality, we performed multiple MR analyses to explore the potential causal association between sarcopenia and cognitive function. Subsequently, we further used MR mediation analysis to examine the degree to which 11 putative mediators (ie, omega-3 fatty acids, vitamin D levels, physical inactivity, falls, frailty, sleep disorders, anxiety, depression, stroke, metabolic syndrome, and type 2 diabetes) may impact the effects of sarcopenia, to identify potential intervention targets in the causal pathway. This study aims to elucidate the causal relationship between sarcopenia and cognitive impairment while identifying actionable intervention targets within the causal pathway through the evaluation of these mediators.

Methods

Study Design

We examined the associations of sarcopenia and its indices with cognitive function in a cohort of approximately 35,000 participants from the United Kingdom Biobank. We then conducted multiple MR analyses to explore the potential causal relations between sarcopenic indices (ie, ALM, handgrip strength, and gait speed) and cognitive function. Finally, we used mediation MR analysis to quantify to what extent physical inactivity, depression, anxiety, falls, frailty, and vitamin D use mediated the effects of sarcopenic indices on cognitive function. This research has been conducted using the UK Biobank Resource under Application 77,646.

Cohort Study

Study Participants

The UK Biobank is a comprehensive prospective cohort that enrolled ≥500,000 participants aged 40 - 69 from 22 centers over the United Kingdom between 2006 and 2010. Neuropsychological tests were performed to assess cognitive function assessment in a subset of participants at baseline as well as 3 subsequent visits via a touch screen. Approximately 170,000 individuals participated in the baseline visit (2006 - 2010), and ~20,000, ~60,000, and ~8000 individuals participated in the first (2012 - 2013), second (2014+), and third (2019+) follow-up visits, respectively. For each follow-up visit dataset, consistent diagnostic criteria were applied for both sarcopenia and cognitive function. In our analysis, we included 35,000 participants who underwent cognitive function tests at baseline and had at least one subsequent follow-up visit (see Figure S1 in Multimedia Appendix 1).

Exposure and Outcome Measurements

Sarcopenia was defined according to the EWGSOP2, which uses the detection of low grip strength and low muscle mass to confirm sarcopenia, with poor physical performance indicating severe sarcopenia [23]. ALM was evaluated by bioelectrical impedance analysis with a Tanita BC418MA body composition analyzer and expressed in kilograms. We used the cut points as recommended in the EWGSOP2 definition of <7 kg/m^2 in men and 5.5 kg/m^2 in women. Grip strength was measured with a hydraulic handheld dynamometer (Jamar J00105). Specifically, 3 measurements were recorded for the maximum strength of both hands, and the highest values were used for analysis, expressed in kilograms. The cut-points for low grip strength

XSL•FO
RenderX

recommended by EWGSOP2 were <27 kg in men and <16 kg in women. Gait speed was acquired directly from the participants through the question "How would you characterize your usual walking speed?" (options: "slow" [<3 miles per hour], "steady/average" [3 - 4 miles per hour], "fast" [>4 miles per hour], or "prefer not to answer" [regarded as missing data]). Participants who reported that they were unable to walk or walked at a slow pace had low physical performance.

Cognitive function was assessed through 2 neuropsychological tests, namely fluid intelligence and prospective memory, by using a touch screen during visits to the UK Biobank assessment center. The fluid intelligence test was designed as a 13-item problem-solving task aiming at logic and reasoning abilities. Prospective memory, which is a form of episodic memory, was used as an indicator to measure the participants' capacity to remember future tasks. The participants were instructed to reproduce a figure on a touch screen following a single instruction to be recalled later in the session. The responses were judged as either "correct on the first attempt" or otherwise, indicating a lapse in prospective memory (eg, "instruction not remembered, either skipped or incorrect" or "correctly recalled on the second attempt").

Covariate Measurements

The selection of potential confounders was based on previous literature [12,13], which was assessed through a touch-screen questionnaire at baseline. These variables included: age (continuous), sex (male or female), education level (college or university; A levels (Advanced Level), AS level (Advanced Subsidiary Level) or equivalent; O level, GCSEs (General Certificate of Secondary Education) or equivalent; CSEs or equivalent; NVQ (National Vocational Qualification), HND (Higher National Diploma), HNC (Higher National Certificate) or equivalent; other professional qualifications), self-reported race (White, Asian, Black, Mixed or other), assessment center, Townsend Deprivation Index at recruitment (continuous), BMI (continuous), presence of long-standing illness (no or yes), overall health rating (excellent, good, fair or poor), smoking status (never, former, or current), frequency of alcohol intake (daily or almost daily; 3 or 4 times a week; once or twice a week; 1 to 3 times a month or special occasions only), sleep duration (<7 h/day, 7 - 9 h/day, or >9 h/day), and TV viewing duration (continuous). Specifically, the presence of long-standing illness was assessed using the following question, "Do you have any long-standing illness, disability, or infirmity". Overall health rating was collected by asking "In general how would you rate your overall health?" All categorical variables strictly adhered to UK Biobank classification standards, with no additional category merging or processing.

Statistical Analyses

To address the repeated measurements of the 2 cognitive tests, we used mixed-effects linear and logistic regression models to evaluate the associations of sarcopenia and its indices with fluid intelligence scores and with the risk of prospective memory impairment, respectively. All the mixed-effects models featured a random individual-specific intercept with a fixed slope. In regression models, we adjusted for sociodemographic, socioeconomic factors, and the number of cognitive assessments, and then further added health-related factors, including BMI, long-standing illness, and health rating. Finally, lifestyle factors, including smoking, alcohol intake frequency, sleep duration, and duration of TV viewing were added. We used a missing indicator approach in the primary analysis and conducted a sensitivity analysis by excluding missing data to evaluate its impact.

MR Analyses

Data Source of Sarcopenic Indices and General Cognitive Function

Given the lack of genome-wide association studies (GWASs) specifically focused on sarcopenia, we used the publicly available summary data on sarcopenic indices and cognitive function from the UK Biobank [24,25] and the COGENT Consortium [26]. The GWAS summary data for ALM (n=450,243), handgrip strength (n=461,089), and gait speed (n=461,089) were derived from the largest public GWASs of individuals with European ancestry in the UK Biobank [24,25]. The summary genetic statistics for the associations of general cognitive function were extracted from a comprehensive GWAS conducted through the UK Biobank and COGENT Consortium [26]. The summary phenotype source and outcome descriptions are provided in Table S1 in Multimedia Appendix 1.

Genetic Instruments for Sarcopenic Indices

The single nucleotide polymorphism (SNPs) meeting the following criteria were selected as instrumental variables: (1) significantly associated with ALM, handgrip strength, or usual gait speed ($P<5\times10^{-8}$), (2) independent (linkage disequilibrium $r^2<.001$ within 10,000 kb), (3) with appropriate effect allele frequencies (≥1%), and (4) not palindromic (adenine/thymine or cytosine/guanine). To correct for multiple comparisons (3 exposures), the Bonferroni method was used. Associations with $P<.016$ (0.05/3) were deemed statistically significant.

Data Analyses

The 2-Sample MR Analyses

The associations between genetically predicted sarcopenic indices and general cognitive function were examined by multiplicative random-effects inverse variance weighted (IVW) analysis, which can provide the most accurate and unbiased estimates [27]. Furthermore, we performed MR-Egger method, MR-PRESSO method, and RadialMR method to pinpoint potential violations of MR assumptions and assess the robustness of primary results [27,28]. The RadialMR method identifies outliers influencing MR analysis, and the results are reanalyzed after their removal [29]. A consistent estimate across multiple sensitivity analyses indicates strengthened causal evidence. We also assessed the bias and type 1 error rate for sample overlap using an internet-based calculator [30]. In addition, we reanalyzed the data using the MRlap method, which is robust to biases caused by sample overlap, winner's curse, and weak instruments [31].

Mediation MR Analyses

We performed univariable MR to estimate the effects of sarcopenic indices on 11 genetically predicted putative

XSL•FO
RenderX

mediators, ie, omega-3 fatty acids, vitamin D levels, moderate-to-vigorous intensity physical activity (MVPA) during leisure time, falls, frailty, sleep disorders, anxiety, depression, stroke, metabolic syndrome, and type 2 diabetes. Mediators showing causal evidence were selected for multivariable Mendelian randomization (MVMR) analysis to estimate the indirect effect of sarcopenic indices on cognitive function mediated by each. The mediation proportion was calculated as the indirect effect divided by the total effect on cognitive function, with standard errors estimated by the delta method [32]. If an inconsistent mediation was observed, where the direct effect opposes the indirect effect, no mediation proportion would be estimated [33].

Complementary Analyses

We performed bidirectional MR analysis to partially explore the potential reverse causality. The SNPs from GWAS that were significantly associated with general cognitive function ($P<5\times10^{-8}$) were selected as instrumental variables (selection criteria remained consistent with the ones mentioned earlier). In addition, we used the MR-Steiger method to examine the directionality of the relationship [34].

All analyses were conducted using the R (R Foundation for Statistical Computing) packages *TwoSampleMR* (version 0.5.6), *MVMR* (version 0.3), *MRPRESSO* (version 1.0), and *MRlap* (version 0.0.3.0) in R (version 4.3.0). A *P* value <.05 was considered significant. IVW estimates were deemed causal if consistent with at least one sensitivity analysis and showed no pleiotropy (Egger intercept *P*>.05). Results were reported as odds ratios (ORs), β coefficients, or proportions with 95% confidence intervals (CIs).

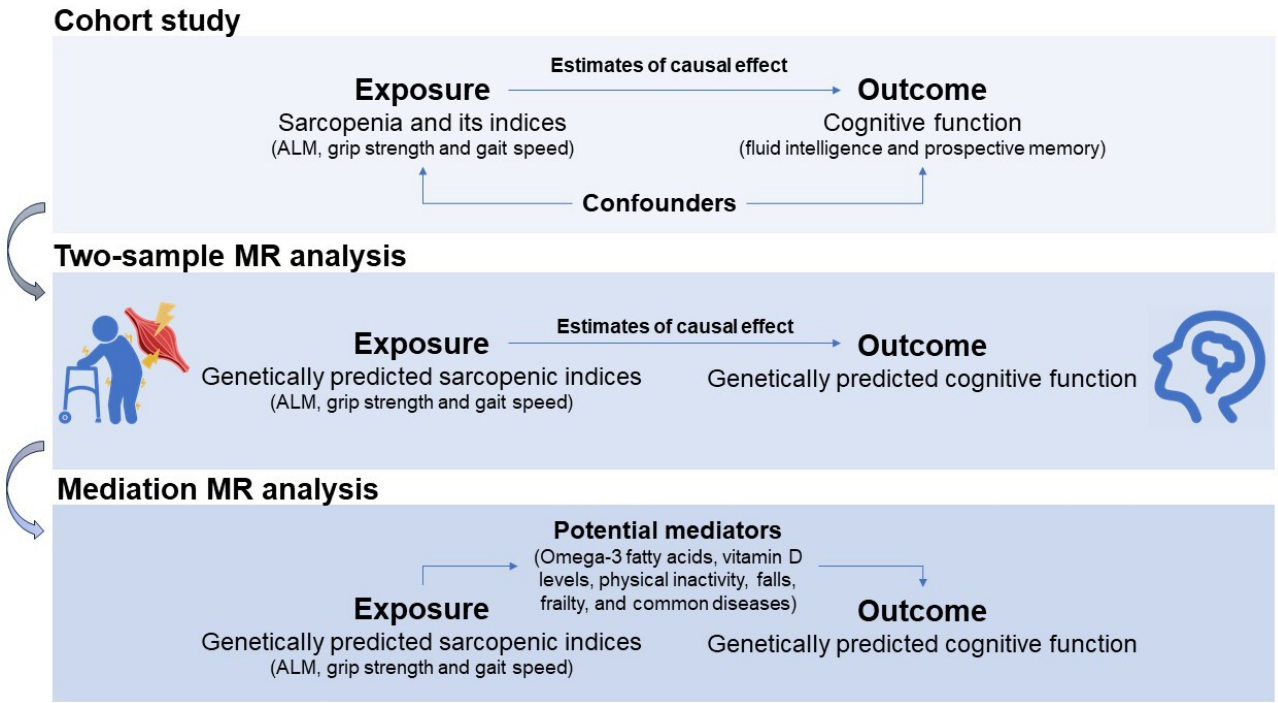
Ethical Considerations

This study had been granted with the UK Biobank research approval by the North West Centre for Research Ethics Committee (11/NW/0382) and informed consent was obtained from all participants. For GWAS datasets, ethical review and approval can be accessed in the original studies. The data used were anonymized to ensure privacy and confidentiality. No compensation was provided to participants.

Results

Figure 1 provides an overview of the study design. The baseline characteristics of the cohort study are summarized in Table 1. Among 34,457 participants, 17,620 (51.1%) were women, with a mean age of 56.4 (SD 7.6) years.

Figure 1. Study design flowchart. ALM: appendicular lean mass; MR: Mendelian randomization.



XSL•FO
RenderX

Table . Baseline characteristics of the cohort study.

Characteristics	Results
Sociodemographics	
Total sample, n	34,457
Age (years), mean (SD)	56.4 (7.6)
Gender (female), n (%)	17,620 (51.1)
Deprivation index, mean (SD)	−1.92 (2.7)
Race, n (%)	
White	33,286 (96.6)
Asian	466 (1.4)
Black	240 (0.7)
Mixed or other	162 (0.5)
Unknown	303 (0.9)
Education level, n (%)	
College or university degree	15,190 (44.1)
A[a] levels, AS[b] levels, or equivalent	4355 (12.6)
O levels, GCSEs,[c] or equivalent	6712 (19.5)
CSEs[d] or equivalent	1345 (3.9)
NVQ[e], HND[f], HNC,[g] or equivalent	1985 (5.8)
Other professional qualifications	1747 (5.0)
Unknown	3123 (9.1)
Health-related factors	
BMI, n (%)	
Underweight	158 (0.5)
Normal weight	12,792 (37.1)
Overweight	14,684 (46.2)
Obesity	6750 (19.6)
Unknown	73 (0.2)
Long-standing illness, n (%)	
No	24,258 (70.4)
Yes	9571 (27.8)
Unknown	628 (1.8)
Overall health rating, n (%)	
Excellent	7082 (20.6)
Good	20,969 (60.9)
Fair	5541 (16.1)
Poor	796 (2.3)
Unknown	69 (0.2)
Lifestyle behaviors	
Smoking status, n (%)	
Never	20,501 (59.5)
Former	11718 (34.1)
Current	2163 (6.3)

XSL•FO
RenderX

Characteristics	Results
Unknown	75 (0.2)
Alcohol intake frequency, n (%)	
Daily or almost daily	7768 (22.5)
Three or four times a week	9408 (27.3)
Once or twice a week	8694 (25.2)
One to three times a month	3755 (10.9)
Special occasions only	3069 (8.9)
Never	1749 (5.1)
Unknown	14 (0.0)
Sleep duration, n (%)	
Short sleep (<7 h/day)	7504 (21.8)
Normal (7 - 9 h/day)	26,465 (76.8)
Long sleep (>9 h/day)	386 (1.1)
Unknown	102 (0.3)
TV viewing (h/days) mean (SD)	1.8 (3.0)

[a]A: Advanced.
[b]AS: Advanced Subsidiary.
[c]GCSE: General Certificate of Secondary Education
[d]CSE: Certificate of Secondary Education.
[e]NVQ: National Vocational Qualification.
[f]HND: Higher National Diploma.
[g]HNC: Higher National Certificate.

Cohort Study

As shown in Table 2, participants with sarcopenia had lower fluid intelligence scores than those without sarcopenia, and the multivariable-adjusted difference in the mean level of fluid intelligence scores was 0.91 (95% CI −1.68 to −0.15; *P*=.02). However, no statistically significant association with prospective memory loss was detected (OR 1.36, 95% CI 0.31-5.92; *P*=.68). Each 5-kg increase in ALM was found to be associated with an increased fluid intelligence score (β=0.27, 95% CI 0.21-0.32; *P*<.001) and a decreased risk of prospective memory loss (OR 0.87, 95% CI 0.78-0.97; *P*<.001). Likewise, each 5-kg increase in handgrip strength was positively associated with fluid intelligence (β=0.02, 95% CI 0.01-0.04; *P*<.001) and negatively associated with the prospective memory loss (OR 0.89, 95% CI 0.86-0.93; *P*<.001). Furthermore, slow gait speed was associated with a lower fluid intelligence score (β=−0.10, 95% CI −0.23 to 0.03; *P*=.15) and an increased risk of prospective memory loss (OR 1.65, 95% CI 1.23-2.19; *P*<.001). The results of sensitivity analyses, after excluding cases with missing data, were consistent with the overall analyses (see Table S2 in Multimedia Appendix 1).

XSL•FO
RenderX

Table . The associations of baseline sarcopenia and its indices with follow-up cognitive function.

Sarcopenia indices	Cognitive function	
	Fluid intelligence, β (95% CI)	Prospective memory loss, OR[a] (95% CI)
Sarcopenia		
Model 1[b]	−0.87 (−1.62 to −0.12)[c]	1.67 (0.47-1.78)
Model 2[d]	−0.86 (−1.61 to −0.11)[c]	1.59 (0.39-6.50)
Model 3[e]	−0.91 (−1.68 to −0.15)[c]	1.36 (0.31-5.92)
Appendicular lean mass		
Model 1[b]	0.17 (0.13-0.21)[c]	0.87 (0.80-0.94)[c]
Model 2[d]	0.27 (0.22-0.32)[c]	0.82 (0.74-0.91)[c]
Model 3[e]	0.27 (0.21-0.32)[c]	0.87 (0.78-0.97)[c]
Handgrip strength		
Model 1[b]	0.03 (0.02-0.05)[c]	0.93 (0.89-0.96)[c]
Model 2[d]	0.03 (0.01-0.04)[c]	0.91 (0.88-0.94)[c]
Model 3[e]	0.02 (0.01-0.04)[c]	0.89 (0.86-0.93)[c]
Slow gait speed		
Model 1[b]	−0.20 (−0.32 to −0.08)[c]	1.64 (1.27-2.12)[c]
Model 2[d]	−0.14 (−0.27 to −0.01)[c]	1.71 (1.30-2.25)[c]
Model 3[e]	−0.10 (−0.23 to 0.03)	1.65 (1.23-2.19)[c]

[a]OR: odds ratio.

[b]Model 1 was adjusted for age, sex, race, aged race, education, deprivation index, and the number of cognitive assessments.

[c]*P*<.05.

[d]Model 2 was additionally adjusted for health-related factors including BMI and long-standing illness.

[e]Model 3 was additionally adjusted for smoking, alcohol intake, sleep duration, and TV viewing.

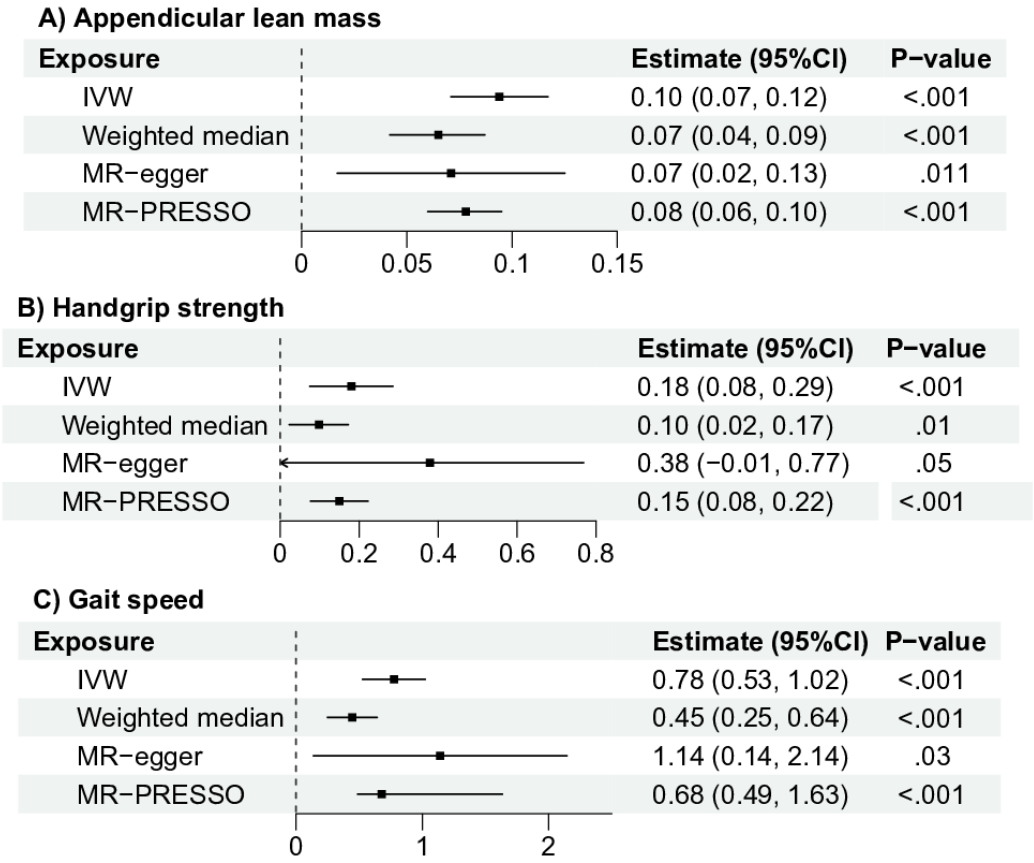
MR Analyses

The 2-Sample MR Analyses

As shown in Figure 2, for an increase of each one unit in genetically predicted ALM, handgrip strength, and gait speed, the general cognitive function score was increased by 0.10 (95% CI 0.07-0.12; *P*<.001), 0.18 (95% CI 0.08-0.29; *P*<.001), and 0.78 (95% CI 0.53-1.02; *P*<.001), respectively. To assess the consistency of directional causation, the effect estimates of 3 different sensitivity methods, that is, weighted median, MR-Egger, and MR-PRESSO were examined and plotted, confirming the directional causation between sarcopenic indices to general cognitive function (see Figure 2 and Figure S2 in Multimedia Appendix 1). There was some evidence of heterogeneity in the SNP effects although the MR-Egger intercepts indicated limited evidence of directional pleiotropy (see Table S3 in Multimedia Appendix 1). In radialMR analyses, outliers were detected (see Figure S3 in Multimedia Appendix 1). MR results remained consistent, showing slightly smaller effects after removing outliers (see Table S4 in Multimedia Appendix 1). For all sarcopenic indices and cognitive function phenotypes, the type 1 error rate was robustly controlled below 0.05 and bias estimates were confined to a narrow range of −0.01 to 0.01 (see Table S5 in Multimedia Appendix 1). Validation through the MRlap method further confirmed that sample overlap did not substantially influence the causal inferences (see Table S6 in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 2. Mendelian randomization results for the relationship of sarcopenic indices with cognitive function. IVW: inverse variance weighted analysis; MR: Mendelian randomization.



Mediation MR Analyses

Figure 3 illustrates the effects of sarcopenic indices on 11 potential mediators. Univariable IVW MR analysis revealed that genetically predicted ALM exhibited protective effects on MVPA (β=0.07, 95% CI 0.05-0.09; *P*<.001) and negative associations with omega-3 fatty acids (β=−0.08, 95% CI −0.12 to −0.05; *P*<.001), falls (β=−0.03, 95% CI −0.05 to -0.003; *P*=.03), and frailty (β=−0.05, 95% CI −0.07 to −0.03; *P*<.001). Genetically predicted handgrip strength showed a positive association with MVPA (β=0.10, 95% CI 0.02-0.19; *P*=.02) and negative associations with falls (β=−0.15, 95% CI −0.24 to −0.07; *P*<.001), and frailty (β=−0.22, 95% CI −0.30 to −0.14; *P*<.001), depression (β=−0.09, 95% CI −0.16 to −0.01; *P*=.02), and stroke (β=−0.41, 95% CI −0.69 to −0.14; *P*=.003). Genetically predicted gait speed was positively associated with MVPA (β=0.83, 95% CI 0.65-1.01; *P*<.001). It also demonstrated significant protective effects against sleep disorders, falls, frailty, anxiety, depression, metabolic syndrome, and type 2 diabetes (with βs ranging from −0.01 to −2.32, all *P* values <.05). Further mediation analysis revealed that the total effect of genetically predicted ALM on general cognitive function decreased from 0.10 (95% CI 0.07-0.12) to 0.09 (95% CI 0.06-0.11) after adjusting for MVPA in MVMR analysis (Table 3). Similarly, the total effect of genetically predicted handgrip strength on general cognitive function attenuated from 0.18 (95% CI 0.08-0.29) to 0.16 (95% CI 0.05-0.27) with adjustment in MVMR analysis. MVPA mediated 8.2% of the total direct effect of ALM on cognitive function, and 10.6% of the total direct effect of handgrip strength on cognitive function. No apparent mediation effect was observed through omega-3 fatty acids, vitamin D level, falls, frailty, sleep disorders, anxiety, depression, stroke, metabolic syndrome, or type 2 diabetes. The type 1 error rate due to sample overlap between sarcopenic indices and the mediators remained below 0.05 (with bias estimates under 0.01) for all phenotypes (see Table S5 in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 3. Effects of genetically predicted sarcopenic indices on potential mediators. MVPA: moderate-to-vigorous intensity physical activity.

XSL•FO
RenderX

Table . The mediation effect of sarcopenic indices on cognitive function via potential mediator in Mendelian randomization (MR) analyses.

Exposures	Mediator	Total effects, β (95% CI)	Direct effects, β (95% CI)	Indirect effects, β (95% CI)	Mediated proportion, % (95% CI)
Appendicular lean mass	MVPA[a]	0.10 (0.07-0.12)[b]	0.09 (0.06-0.11)[b]	0.01 (0.00-0.02)[b]	8.2 (0-16.7)
Handgrip strength	MVPA[a]	0.18 (0.08-0.29)[b]	0.16 (0.05-0.27)[b]	0.02 (0.00-0.05)[b]	10.6 (0-29.6)

[a]MVPA: moderate-to-vigorous intensity physical activity during leisure time.

[b]$P<.05$.

Complementary Analyses

We conducted bidirectional MR analysis to investigate the potential reverse causality between sarcopenia and cognitive function. Table S7 in Multimedia Appendix 1 shows evidence supporting a causal effect of genetically predicted cognitive function on all sarcopenic indices. The effect estimates demonstrated a general consistency across different sensitivity methods (see Figure S2 in Multimedia Appendix 1). Some heterogeneity in the SNP effects was observed, although the MR-Egger intercepts suggested no evidence of directional pleiotropy (see Table S7 in Multimedia Appendix 1). In radialMR analyses, outliers were detected (see Figure S4 in Multimedia Appendix 1). The MR results remained consistent both before and after outlier correction and are presented in Tables S7 and S8 in Multimedia Appendix 1.

Discussion

Principal Findings

This study revealed that sarcopenia and its defining indices (appendicular lean mass, handgrip strength, and gait speed) are associated with cognitive function based on observational data. MR analyses further established a causal relationship between higher levels of sarcopenic indices and better general cognitive function. In addition, physical activity was identified as a significant mediator in the causal pathway linking sarcopenic indices to cognitive function. Our findings suggest sarcopenia as a risk factor and potential biomarker for cognitive impairment, with physical activity offering a therapeutic approach to delay or prevent cognitive decline.

Comparison With Previous Work

The association between sarcopenia and the risk of cognitive impairment has been investigated by several longitudinal studies, but the results were conflicting. Some authors indicated that individuals with sarcopenia [12-17], reduced muscle strength [13,17,35], decreased muscle mass [35], or compromised physical performance [35,36] were associated with an elevated risk of cognitive impairment. However, some others reported no significant association between sarcopenia or its indicators and the risk of cognitive impairment [19-21]. The conflicting results could stem from differences in how sarcopenia is defined and the variations in sample sizes. Our observational analyses addressed these discrepancies by focusing on the more recent EWGSOP2 definition of sarcopenia and its 3 defining indices within a substantial sample size (>200,000 participants). As a result, we found that sarcopenia, as a construct, and its 3 distinct indices were closely correlated with cognitive impairment. Furthermore, the discrepancies observed in previous observational studies might also be attributed to factors such as residual confounding and measurement errors. To mitigate these limitations in our study, we used MR, a genetic epidemiological technique using genetic variants as proxies for the exposure [37]. This approach is less vulnerable to the aforementioned limitations since the genetic variants are accurately measured and documented and are randomly allocated during gamete formation and conception. Therefore, this method can minimize the potential for measurement errors and reduce the likelihood of being influenced by confounding variables [37].

Possible Explanations

Our study demonstrates that physical inactivity is a potential mediating factor in the causal pathway between sarcopenia and cognitive impairment. On one hand, individuals with sarcopenia might have a low level of physical activity [38], which could be attributed to the fact that weakened muscles might hinder the ability to exercise regularly. On the other hand, physical activity positively impacts brain health [39]. First, it can stimulate the formation of new neurons, enhance neuronal survival [10], increase resistance to brain injuries, and facilitate synaptic development and plasticity [40]. Second, physical activity promotes better blood vessel formation in the brain, which is associated with increased learning capabilities [10,11]. Third, it also activates specific gene expression profiles benefiting brain plasticity and cognitive function [41]. Fourth, engaging in physical activity has been associated with reduced amyloid deposition in the brains of cognitively normal elderly adults [42]. Furthermore, physical activity can reduce systemic inflammatory markers [43] and boost the production of neuroprotective proteins like brain-derived neurotrophic factor (BDNF), supporting the growth and survival of neurons [44]. In addition, it positively influences energy balance and glucose metabolism by modulating AMP kinase and insulin signaling, potentially aiding Aβ clearance [45]. These multiple benefits of physical activity contribute to safeguarding cognitive function and underscore the importance of maintaining an active lifestyle in individuals with sarcopenia to support brain health.

Limitations

The triangulation of findings through complementary cohort and MR approaches significantly bolstered the confidence in our drawn inferences. However, several limitations warrant consideration. First, despite using multiple MR methods to resist pleiotropy-related confounding, we could not eliminate residual confounding, which is a known limitation of the MR approach. Second, due to a mere 5% response rate and healthy volunteer bias in the UK Biobank, whether our findings can be generalized

XSL•FO
RenderX

to the broader UK population remains uncertain, despite the large sample size. Third, although we have accounted for common modifiable lifestyle factors and preventable diseases to inform public health policies, this study does not encompass all potential mediation pathways. Fourth, only self-reported gait speed was collected, lacking objective measurements for more accurate estimates. Fifth, the inability to perform stratified analyses by age and gender due to data constraints limits insights into these factors' roles in the sarcopenia-cognition relationship. Sixth, the overlap between GWAS datasets in the MR analysis could bias results and inflate Type 1 error rates. To address this, we assessed the error rate and applied the MRlap method, confirming the robustness of our findings and minimizing the influence of sample overlap on causal associations. Finally, to maintain homogeneity, we focused on individuals of European ancestry in our sample selection, which might influence the generalization of the results, although the sample size was substantial.

Clinical and Research Implications

Our findings provide a promising approach to alleviate the burden of cognitive impairment by identifying and intervening in sarcopenia. Specifically, we have recognized sarcopenia as a risk factor for future cognitive impairment, making it a potential clinical biomarker to screen adults at risk of late-life cognitive impairment. Although no pharmaceutical treatment has been specifically approved for sarcopenia, implementing non-pharmacological interventions, such as physical activity, can serve as a therapeutic approach to proactively delay or prevent the onset of cognitive impairment in affected individuals. Our research indicates that physical activity can mediate the effect of sarcopenia on cognitive function, offering valuable insights that complement the prevailing emphasis on intellectual pursuits as the primary means of exercising the brain [46]. Promoting physical activity may yield a dual positive impact, addressing both sarcopenia and cognitive impairment simultaneously. This approach can potentially enhance the overall health and well-being of those affected by sarcopenia.

Conclusions

Our study suggests that sarcopenia is a causal risk factor for cognitive impairment. Physical activity, a modifiable factor, is capable of measuring the effect of sarcopenia on cognitive function.

Acknowledgments

This work was supported by the National Key Research and Development Program (2022YFC3601900 and 2022YFC2505500), the National Natural Science Foundation of China (81930071, 81902265, 82072502, 82302771, and 82402882), the National Natural Science Foundation Regional Innovation and Development Joint Fund (U21A20352), the Natural Science Foundation of Hunan Province (2022JJ40821, 2024JJ6667), the Project Program of National Clinical Research Center for Geriatric Disorders (Xiangya Hospital; 2020LNJJ03 and 2021LNJJ06), the Scientific Research Program of the Education Department of Hunan Province (296840), and the Natural Science Foundation of Changsha (kq2403017). The funding source had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and the decision to submit the manuscript for publication.

Data Availability

The primary data from the UK Biobank resource are accessible upon application from the UK Biobank repository. The genome-wide association study (GWAS) summary statistics for sarcopenic indices, cognitive function, and potential mediators are publicly available and can be directly accessed through the original publications cited in the reference list [24,26,47-59].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary methods and results for the study.
[DOCX File, 1104 KB - aging_v8i1e66031_app1.docx]

References

1. Langa KM, Levine DA. The diagnosis and management of mild cognitive impairment: a clinical review. JAMA 2014 Dec 17;312(23):2551-2561. [doi: 10.1001/jama.2014.13806] [Medline: 25514304]
2. Pan CW, Wang X, Ma Q, Sun HP, Xu Y, Wang P. Cognitive dysfunction and health-related quality of life among older Chinese. Sci Rep 2015 Nov 25;5:17301. [doi: 10.1038/srep17301] [Medline: 26601612]
3. Fogg C, Griffiths P, Meredith P, Bridges J. Hospital outcomes of older people with cognitive impairment: an integrative review. Int J Geriatr Psychiatry 2018 Jun 26;33(9):1177-1197. [doi: 10.1002/gps.4919] [Medline: 29947150]
4. Rait G, Walters K, Bottomley C, Petersen I, Iliffe S, Nazareth I. Survival of people with clinical diagnosis of dementia in primary care: cohort study. BMJ 2010 Aug 5;341:c3584. [doi: 10.1136/bmj.c3584] [Medline: 20688840]
5. Cooper C, Li R, Lyketsos C, Livingston G. Treatment for mild cognitive impairment: systematic review. Br J Psychiatry 2013 Sep;203(3):255-264. [doi: 10.1192/bjp.bp.113.127811] [Medline: 24085737]

XSL•FO
RenderX

6. Sakal C, Li T, Li J, Li X. Identifying predictive risk factors for future cognitive impairment among Chinese older adults: longitudinal prediction study. JMIR Aging 2024 Mar 22;7:e53240. [doi: 10.2196/53240] [Medline: 38534042]
7. Scarmeas N, Anastasiou CA, Yannakoulia M. Nutrition and prevention of cognitive impairment. Lancet Neurol 2018 Nov;17(11):1006-1015. [doi: 10.1016/S1474-4422(18)30338-7] [Medline: 30244829]
8. Petermann-Rocha F, Balntzi V, Gray SR, et al. Global prevalence of sarcopenia and severe sarcopenia: a systematic review and meta‐analysis. J Cachexia Sarcopenia Muscle 2022 Feb;13(1):86-99. [doi: 10.1002/jcsm.12783] [Medline: 34816624]
9. Cruz-Jentoft AJ, Baeyens JP, Bauer JM, et al. Sarcopenia: European consensus on definition and diagnosis: report of the European Working Group on sarcopenia in older people. Age Ageing 2010 Jul;39(4):412-423. [doi: 10.1093/ageing/afq034] [Medline: 20392703]
10. van Praag H, Christie BR, Sejnowski TJ, Gage FH. Running enhances neurogenesis, learning, and long-term potentiation in mice. Proc Natl Acad Sci U S A 1999 Nov 9;96(23):13427-13431. [doi: 10.1073/pnas.96.23.13427] [Medline: 10557337]
11. Black JE, Isaacs KR, Anderson BJ, Alcantara AA, Greenough WT. Learning causes synaptogenesis, whereas motor activity causes angiogenesis, in cerebellar cortex of adult rats. Proc Natl Acad Sci U S A 1990 Jul;87(14):5568-5572. [doi: 10.1073/pnas.87.14.5568] [Medline: 1695380]
12. Lin A, Wang T, Li C, et al. Association of sarcopenia with cognitive function and dementia risk score: a national prospective cohort study. Metabolites 2023 Feb 8;13(2):245. [doi: 10.3390/metabo13020245] [Medline: 36837864]
13. Peng TC, Chiou JM, Chen TF, Chen YC, Chen JH. Grip strength and sarcopenia predict 2-year cognitive impairment in community-dwelling older adults. J Am Med Dir Assoc 2023 Mar;24(3):292-298. [doi: 10.1016/j.jamda.2022.10.015] [Medline: 36435272]
14. Salinas-Rodríguez A, Palazuelos-González R, Rivera-Almaraz A, Manrique-Espinoza B. Longitudinal association of sarcopenia and mild cognitive impairment among older Mexican adults. J Cachexia Sarcopenia Muscle 2021 Dec;12(6):1848-1859. [doi: 10.1002/jcsm.12787] [Medline: 34535964]
15. Hu Y, Peng W, Ren R, Wang Y, Wang G. Sarcopenia and mild cognitive impairment among elderly adults: the first longitudinal evidence from CHARLS. J Cachexia Sarcopenia Muscle 2022 Dec;13(6):2944-2952. [doi: 10.1002/jcsm.13081] [Medline: 36058563]
16. Nishiguchi S, Yamada M, Shirooka H, et al. Sarcopenia as a risk factor for cognitive deterioration in community-dwelling older adults: a 1-year prospective study. J Am Med Dir Assoc 2016 Apr 1;17(4):372. [doi: 10.1016/j.jamda.2015.12.096] [Medline: 26897591]
17. Beeri MS, Leugrans SE, Delbono O, Bennett DA, Buchman AS. Sarcopenia is associated with incident Alzheimer's dementia, mild cognitive impairment, and cognitive decline. J Am Geriatr Soc 2021 Jul;69(7):1826-1835. [doi: 10.1111/jgs.17206] [Medline: 33954985]
18. Amini N, Ibn Hach M, Lapauw L, et al. Meta-analysis on the interrelationship between sarcopenia and mild cognitive impairment, Alzheimer's disease and other forms of dementia. J Cachexia Sarcopenia Muscle 2024 Aug;15(4):1240-1253. [doi: 10.1002/jcsm.13485] [Medline: 38715252]
19. Abellan van Kan G, Cesari M, Gillette-Guyonnet S, et al. Sarcopenia and cognitive impairment in elderly women: results from the EPIDOS cohort. Age Ageing 2013 Mar;42(2):196-202. [doi: 10.1093/ageing/afs173] [Medline: 23221099]
20. van Kan GA, Cesari M, Gillette-Guyonnet S, Dupuy C, Vellas B, Rolland Y. Association of a 7-year percent change in fat mass and muscle mass with subsequent cognitive dysfunction: the EPIDOS-Toulouse cohort. J Cachexia Sarcopenia Muscle 2013 Sep;4(3):225-229. [doi: 10.1007/s13539-013-0112-z] [Medline: 23888380]
21. Taekema DG, Ling CHY, Kurrle SE, et al. Temporal relationship between handgrip strength and cognitive performance in oldest old people. Age Ageing 2012 Jul;41(4):506-512. [doi: 10.1093/ageing/afs013] [Medline: 22374646]
22. Skrivankova VW, Richmond RC, Woolf BAR, et al. Strengthening the reporting of observational studies in epidemiology using mendelian randomisation (STROBE-MR): explanation and elaboration. BMJ 2021 Oct 26;375:n2233. [doi: 10.1136/bmj.n2233] [Medline: 34702754]
23. Cruz-Jentoft AJ, Bahat G, Bauer J, et al. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing 2019 Jul 1;48(4):601-601. [doi: 10.1093/ageing/afz046]
24. Pei YF, Liu YZ, Yang XL, et al. The genetic architecture of appendicular lean mass characterized by association analysis in the UK Biobank study. Commun Biol 2020 Oct 23;3(1):608. [doi: 10.1038/s42003-020-01334-0] [Medline: 33097823]
25. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. Elife 2018 May 30;7:e34408. [doi: 10.7554/eLife.34408] [Medline: 29846171]
26. Lee JJ, Wedow R, Okbay A, et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nat Genet 2018 Jul 23;50(8):1112-1121. [doi: 10.1038/s41588-018-0147-3] [Medline: 30038396]
27. Burgess S, Bowden J, Fall T, Ingelsson E, Thompson SG. Sensitivity analyses for robust causal inference from mendelian randomization analyses with multiple genetic variants. Epidemiology (Sunnyvale) 2017 Jan;28(1):30-42. [doi: 10.1097/EDE.0000000000000559] [Medline: 27749700]
28. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet 2018 May;50(5):693-698. [doi: 10.1038/s41588-018-0099-7] [Medline: 29686387]

XSL•FO
RenderX

29. Bowden J, Spiller W, Del Greco M F, et al. Improving the visualization, interpretation and analysis of two-sample summary data mendelian randomization via the radial plot and radial regression. Int J Epidemiol 2018 Aug 1;47(4):1264-1278. [doi: 10.1093/ije/dyy101] [Medline: 29961852]
30. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample mendelian randomization. Genet Epidemiol 2016 Nov;40(7):597-608. [doi: 10.1002/gepi.21998] [Medline: 27625185]
31. Mounier N, Kutalik Z. Bias correction for inverse variance weighting mendelian randomization. Genet Epidemiol 2023 Jun;47(4):314-331. [doi: 10.1002/gepi.22522] [Medline: 37036286]
32. MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V. A comparison of methods to test mediation and other intervening variable effects. Psychol Methods 2002 Mar;7(1):83-104. [doi: 10.1037/1082-989x.7.1.83] [Medline: 11928892]
33. Valeri L, Vanderweele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. Psychol Methods 2013 Jun;18(2):137-150. [doi: 10.1037/a0031034] [Medline: 23379553]
34. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. PLoS Genet 2017 Nov;13(11):e1007081. [doi: 10.1371/journal.pgen.1007081] [Medline: 29149188]
35. Amini N, Dupont J, Lapauw L, et al. Sarcopenia-defining parameters, but not sarcopenia, are associated with cognitive domains in middle-aged and older European men. J Cachexia Sarcopenia Muscle 2023 Jun;14(3):1520-1532. [doi: 10.1002/jcsm.13229] [Medline: 37021434]
36. Moon JH, Moon JH, Kim KM, et al. Sarcopenia as a predictor of future cognitive impairment in older adults. J Nutr Health Aging 2016;20(5):496-502. [doi: 10.1007/s12603-015-0613-x] [Medline: 27102786]
37. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med 2008 Apr 15;27(8):1133-1163. [doi: 10.1002/sim.3034]
38. Chiba I, Lee S, Bae S, et al. Difference in sarcopenia characteristics associated with physical activity and disability incidences in older adults. J Cachexia Sarcopenia Muscle 2021 Dec;12(6):1983-1994. [doi: 10.1002/jcsm.12801] [Medline: 34612020]
39. Mc Ardle R, Taylor L, Cavadino A, Rochester L, Del Din S, Kerse N. Characterizing walking behaviors in aged residential care using accelerometry, with comparison across care levels, cognitive status, and physical function: cross-sectional study. JMIR Aging 2024 Jun 4;7:e53020. [doi: 10.2196/53020] [Medline: 38842168]
40. Lu B, Chow A. Neurotrophins and hippocampal synaptic transmission and plasticity. J Neurosci Res 1999 Oct 1;58(1):76-87. [Medline: 10491573]
41. Escorihuela RM, Tobeña A, Fernández-Teruel A. Environmental enrichment and postnatal handling prevent spatial learning deficits in aged hypoemotional (Roman high-avoidance) and hyperemotional (Roman low-avoidance) rats. Learn Mem 1995;2(1):40-48. [doi: 10.1101/lm.2.1.40] [Medline: 10467565]
42. Brown BM, Peiffer JJ, Martins RN. Multiple effects of physical activity on molecular and cognitive signs of brain aging: can exercise slow neurodegeneration and delay Alzheimer's disease? Mol Psychiatry 2013 Aug;18(8):864-874. [doi: 10.1038/mp.2012.162] [Medline: 23164816]
43. Ford ES. Does exercise reduce inflammation? Physical activity and C-reactive protein among U.S. adults. Epidemiology 2002 Sep;13(5):561-568. [doi: 10.1097/00001648-200209000-00012] [Medline: 12192226]
44. Nichol K, Deeny SP, Seif J, Camaclang K, Cotman CW. Exercise improves cognition and hippocampal plasticity in APOE epsilon4 mice. Alzheimers Dement 2009 Jul;5(4):287-294. [doi: 10.1016/j.jalz.2009.02.006] [Medline: 19560099]
45. De Felice FG, Vieira MNN, Bomfim TR, et al. Protection of synapses against Alzheimer's-linked toxins: insulin signaling prevents the pathogenic binding of abeta oligomers. Proc Natl Acad Sci U S A 2009 Feb 10;106(6):1971-1976. [doi: 10.1073/pnas.0809158106] [Medline: 19188609]
46. Hillman CH, Erickson KI, Kramer AF. Be smart, exercise your heart: exercise effects on brain and cognition. Nat Rev Neurosci 2008 Jan;9(1):58-65. [doi: 10.1038/nrn2298] [Medline: 18094706]
47. Hand grip strength (right). IEU open GWAS project. URL: https://gwas.mrcieu.ac.uk/datasets/ukb-b-10215/ [accessed 2025-05-21]
48. Usual walking pace. IEU open GWAS project. URL: https://gwas.mrcieu.ac.uk/datasets/ukb-b-4711/ [accessed 2025-05-21]
49. UK Biobank. URL: https://www.ukbiobank.ac.uk/ [accessed 2025-05-21]
50. Wang Z, Emmerich A, Pillon NJ, et al. Genome-wide association analyses of physical activity and sedentary behavior provide insights into underlying mechanisms and roles in disease prevention. Nat Genet 2022 Sep;54(9):1332-1344. [doi: 10.1038/s41588-022-01165-1] [Medline: 36071172]
51. Non-cancer illness code, self-reported: anxiety/panic attacks. IEU open GWAS project. URL: https://gwas.mrcieu.ac.uk/datasets/ukb-b-17243/ [accessed 2015-05-21]
52. Trajanoska K, Seppala LJ, Medina-Gomez C, et al. Genetic basis of falling risk susceptibility in the UK Biobank Study. Commun Biol 2020 Sep 30;3(1):543. [doi: 10.1038/s42003-020-01256-x] [Medline: 32999390]
53. Atkins JL, Jylhävä J, Pedersen NL, et al. A genome-wide association study of the frailty index highlights brain pathways in ageing. Aging Cell 2021 Sep;20(9):e13459. [doi: 10.1111/acel.13459] [Medline: 34431594]
54. Jiang X, O'Reilly PF, Aschard H, et al. Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. Nat Commun 2018 Jan 17;9(1):260. [doi: 10.1038/s41467-017-02662-2] [Medline: 29343764]

XSL•FO
RenderX

55. Study: GCST009602. GWAS catalog. URL: https://www.ebi.ac.uk/gwas/studies/GCST009602 [accessed 2025-05-21]
56. Omega-3 fatty acids. IEU open GWAS project. URL: https://gwas.mrcieu.ac.uk/datasets/met-d-Omega_3/ [accessed 2025-05-21]
57. Sleep disorders (combined). IEU open GWAS project. URL: https://gwas.mrcieu.ac.uk/datasets/finn-b-SLEEP/ [accessed 2025-05-21]
58. Malik R, Traylor M, Pulit SL, et al. Low-frequency and common genetic variation in ischemic stroke: the METASTROKE collaboration. Neurology (ECronicon) 2016 Mar 29;86(13):1217-1226. [doi: 10.1212/WNL.0000000000002528] [Medline: 26935894]
59. Mahajan A, Wessel J, Willems SM, et al. Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes. Nat Genet 2018 Apr;50(4):559-571. [doi: 10.1038/s41588-018-0084-1] [Medline: 29632382]

Abbreviations

A level: Advanced Level
ALM: appendicular lean mass
AS level: Advanced Subsidiary Level
EWGSOP2: European Working Group on Sarcopenia in Older People in 2019
GCSE: General Certificate of Secondary Education
GWAS: genome-wide association study
HNC: Higher National Certificate
HND: Higher National Diploma
IVW: inverse variance weighted analysis
MR: Mendelian randomization
MVMR: multivariable Mendelian randomization
MVPA: moderate-to-vigorous intensity physical activity
NVQ: National Vocational Qualification
OR: odds ratio
SNP: single nucleotide polymorphism

Edited by R Bjarnadottir; submitted 02.09.24; peer-reviewed by G Feng, Z Wang; revised version received 12.03.25; accepted 21.04.25; published 11.06.25.

Please cite as:
Sha T, Zhang Y, Wei J, Li C, Zeng C, Lei G, Wang Y
Sarcopenia and Risk of Cognitive Impairment: Cohort Study and Mendelian Randomization Analysis
JMIR Aging 2025;8:e66031
URL: https://aging.jmir.org/2025/1/e66031
doi:10.2196/66031

© Tingting Sha, Yuqing Zhang, Jie Wei, Changjun Li, Chao Zeng, Guanghua Lei, Yilun Wang. Originally published in JMIR Aging (https://aging.jmir.org), 11.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Feasibility of a Cinematic–Virtual Reality Program Educating Health Professional Students About the Complexity of Geriatric Care: Pilot Pre-Post Study

Elizabeth A Beverly[1,2*], PhD; Samuel Miller[3*], BA; Matthew Love[2*], MFA; Carrie Love[2], MFA

[1]Department of Primary Care, Ohio University Heritage College of Osteopathic Medicine, 1 Ohio University, Athens, OH, United States

[2]The Diabetes Institute, Ohio University, Athens, OH, United States

[3]Department of Medicine, Ohio University Heritage College of Osteopathic Medicine, Athens, OH, United States

[*]these authors contributed equally

Corresponding Author:
Elizabeth A Beverly, PhD
Department of Primary Care, Ohio University Heritage College of Osteopathic Medicine, 1 Ohio University, Athens, OH, United States

Abstract

Background: The US population is aging. With this demographic shift, more older adults will be living with chronic conditions and geriatric syndromes. To prepare the next generation of health care professionals for this aging population, we need to provide training that captures the complexity of geriatric care.

Objective: This pilot study aimed to assess the feasibility of the cinematic–virtual reality (cine-VR) training in the complexity of geriatric care. We measured changes in attitudes to disability, self-efficacy to identify and manage elder abuse and neglect, and empathy before and after participating in the training program.

Methods: We conducted a single-arm, pretest-posttest pilot study to assess the feasibility of a cine-VR training and measure changes in attitudes to disability, self-efficacy to identify and manage elder abuse and neglect, and empathy. Health professional students from a large university in the Midwest were invited to participate in 1 of 4 cine-VR trainings. Participants completed 3 surveys before and after the cine-VR training. We performed paired *t* tests to examine changes in these constructs before and after the training.

Results: A total of 65 health professional students participated in and completed the full cine-VR training for 100% retention. Participants did not report any technological difficulties or adverse effects from wearing the head-mounted displays or viewing the 360-degree video. Out of the 65 participants, 48 completed the pre- and postassessments. We observed an increase in awareness of discrimination towards people with disability (t_{47}=−3.97; P<.001). In addition, we observed significant improvements in self-efficacy to identify and manage elder abuse and neglect (t_{47}=−3.36; P=.002). Finally, we observed an increase in participants' empathy (t_{47}=−2.33; P=.02).

Conclusions: We demonstrated that our cine-VR training program was feasible and acceptable to health professional students at our Midwestern university. Findings suggest that the cine-VR training increased awareness of discrimination towards people with disabilities, improved self-efficacy to identify and manage elder abuse and neglect, and increased empathy. Future research using a randomized controlled trial design with a larger, more diverse sample and a proper control condition is needed to confirm the effectiveness of our cine-VR training.

(JMIR Aging 2025;8:e64633) doi:10.2196/64633

KEYWORDS

virtual reality; VR; aging; geriatric syndromes; diabetes; elder abuse and neglect; gerontology; geriatrics; older; elderly; education; student; cinematic; video; head mounted; feasibility; experience; attitude; opinion; perception; elder abuse; chronic conditions; older adult care; health intervention; randomized controlled trial

Introduction

The US population is aging. Over the next 3 decades, the number of Americans aged 65 and older is projected to increase from 58 million in 2022 to 82 million in 2050, representing a 47% increase [1]. With this demographic shift, more older adults will be living with chronic conditions. According to the National Council on Aging, 94.9% of adults aged 60 years and older have at least one chronic condition and 78.7% have two or more [2]. Furthermore, 42% of adults aged 60 years and older have

XSL•FO
RenderX

obesity, which increases the risk for cardiovascular disease, type 2 diabetes, and certain cancers [3]. This translates to older Americans requiring more health care and access to health care professionals. Thus, training the next generation of health care professionals in the complexity of care for older adults is an urgent need.

Despite this need, the education system lacks adequate training in geriatric care [4]. Health professional training should address age-related physiological, physical, cognitive, and psychosocial changes that may negatively impact quality of life, disability, morbidity, and mortality [5,6]. Furthermore, this training should emphasize the early detection of and treatment for geriatric syndromes (eg, polypharmacy, falls, frailty, and incontinence) given their high prevalence and frequent under-recognition in older adults [7]. Finally, health professional training should draw attention to implicit and explicit biases in aging as well as ethical issues related to elder abuse and neglect and end-of-life care [8].

Clinical simulation may be an innovative approach to teaching health professional students about the complexity of care in older adults. Previous research demonstrates that clinical simulation is an effective methodology for health professional students [9]. The value of clinical simulation is that it replaces or enhances real experiences with guided experiences that replicate the real world in a fully interactive manner [10]. Furthermore, clinical simulation is repeatable, which allows for the repetition of acquired skills with appropriate feedback [11]. Clinical simulation can be delivered in different modalities, including role play, the use of standardized patients, computer screen-based simulation, and the use of electronic patient records [11]. Electronic patients are most commonly mannequin-based or replicas of clinical sites; however, new advancements in technology have led to more sophisticated and realistic virtual reality (VR) based simulations [12].

Cinematic-VR (cine-VR) is one type of VR-based simulation. Specifically, cine-VR leverages 360-degree video and spatialized audio through film. It differs from traditional VR, which uses computer-generated characters and worlds. In cine-VR, filmmakers apply the techniques of cinema (ie, narrative storytelling, script writing, actors, lighting, and postproduction) to create an immersive, visually compelling environment [13,14]. We created a cine-VR training program to depict the complexity of geriatric care. Our clinical simulation depicted an older adult with multimorbidity, disability, and several geriatric syndromes. In addition, we wove in themes of intersectionality with regards to ageism, ableism, and racism. We conducted a pilot study to assess the feasibility and acceptability of the cine-VR training program with health professional students. We also measured changes in health professional students' attitudes toward disability, self-efficacy to identify and manage elder abuse and neglect, and empathy. The pilot study was designed to assess the feasibility of our data collection methods, willingness of participants to complete the cine-VR training program, sample size estimation, and refinement of measurements. We will use the findings to inform the development of a future randomized controlled trial to compare the cine-VR training program to traditional instruction.

Methods

Research Design

We conducted a single-arm, pretest-posttest study to assess changes in attitudes to disability, self-efficacy to identify and manage elder abuse and neglect, and empathy.

Ethical Considerations

We obtained ethics approval from the Ohio University Office of Research Compliance Institutional Review Board (approval number: #22-X-153). We ensured our pilot study met the requirements set forth in the regulations on public welfare in Part 46 of Title 45 of the Code of Federal Regulations (45 CFR 46) by complying with federal, state, and local laws and regulations for human subjects; the principles set forth in "The Belmont Report," and the Helsinki Declaration of 1975. All participants provided electronic informed consent. Informed consent was completed before participation in the study.

Cine-VR Episodes

The cine-VR episodes were designed to educate health professional students about complex health conditions, social drivers of health, and implicit bias. We conveyed these objectives through our patient, John Chen, an 80-year-old Chinese American man with a 14-year history of type 2 diabetes and comorbid hypertension. Mr. Chen is a person who is hard of hearing with a mobility disability. His most recent HBA_{1c} level was 9.7%. His medications include metformin and glyburide for his diabetes and lisinopril for his blood pressure. He also has osteoarthritis, decreased renal function (eGFR=58 mL/min/1.73 m^2), elevated triglycerides due to his glucose levels (299 mg/dL), and a body mass index of 23.1 kg/m^2. He has not received diabetes self-management education and support, despite being enrolled in both Medicare and Medicaid. Recently, Mr Chen moved in with his son's family, but he is having a difficult time adjusting to his new home. Although he is able to move around on his own, his ambulation is unsteady, and he falls often. He also experiences frequent urinary incontinence, which frustrates his son and daughter-in-law because they have to clean him. Furthermore, Mr Chen does not drive, shop, cook, or bathe on his own. Therefore, he is reliant on his family to meet most of his daily needs. This leaves Mr Chen feeling isolated, dependent, and ashamed of his physical limitations.

In the cine-VR training program, participants watched 6 episodes, ranging in length from 2 minutes to 5 minutes. The episodes captured 3 separate patient-provider interactions, including visits with a primary care physician, an urgent care physician, and an emergency department physician. Participants were encouraged to identify issues with his medical care as well as concerns about the caregiving he received at home. Specifically, participants were challenged to figure out why Mr Chen was experiencing so many falls, injuries, and health emergencies.

The fifth and sixth episodes of the cine-VR training program were "guided simulations," or prerecorded cine-VR face-to-face conversations with the emergency department physician and Mr Chen. The guided simulations were designed to simulate a

XSL•FO
RenderX

high-stakes conversation, to give participants an opportunity to practice difficult conversations without the pressure of causing harm. Participants were instructed to speak predetermined dialogue that appeared at the bottom of the 360-degree video visible within the headset. The first "guided simulation" was a conversation between the participant and the emergency department physician. The participant assumed the role of a clinical colleague and encouraged the physician to inquire further about the bruises observed on John's body and to call social work for suspicion of elder abuse and neglect. The second "guided simulation" was a conversation between the emergency department physician, Mr Chen, and a social worker. In this "guided simulation," the participant took on the role of the emergency department physician and asked Mr Chen difficult questions about his injuries, his family, and life at home.

Cine-VR Training Curricular Content

We developed curricular content to be taught in tandem with the cine-VR episodes. The curriculum included 6 debriefs or reflections that reinforced key takeaways from each cine-VR episode. The key takeaways from each episode focused on the following content: (1) Diabetes, disability, and the aging population, (2) bias toward disability, (3) association between disability and elder abuse and neglect, (4) recognizing elder abuse and neglect, (5) identifying risk factors for elder abuse and neglect, and (6) reporting elder abuse and neglect. An experienced behavioral diabetes researcher (EAB) delivered both trainings. The integrity of the curricular content was ensured by written materials, a peer-review process of all materials, delivery by one trained behavioral diabetes expert, and team member observation of cine-VR training.

Cine-VR Technology

Participants viewed the cine-VR episodes with Pico G2 4K head-mounted displays. These head-mounted displays allowed participants to move their head and body in any direction to choose what aspects they paid attention to during each cine-VR episode. We also synchronized all cine-VR episodes from a central computer using VR Sync software so all participants viewed the same content in the head-mounted display at the same time.

Power Analysis

This was a pilot study so we did not conduct an a priori power analysis. Using the guidance of Lancaster et al [15], we recruited a minimum sample size of 30 participants.

Recruitment

Health professional students were recruited from a large university in the Midwest. Eligibility criteria for participating in the pilot study included English-speaking and reading adults aged 18 years and older who were enrolled during the 2022 - 2023 academic year and majoring in one of the following programs: premedicine program, speech-language pathology, audiology, nutrition and dietetics, physical therapy, nursing (any level), social work (any level), athletic training, exercise physiology, pre-physician assistant program, psychology, public health administration, or public health. There were no other exclusion criteria. Specifically, we hosted 4 in-person cine-VR training sessions on campus in classrooms and invited students from those health professional programs to participate. Sessions were hosted on October 17, 2022, October 24, 2022, April 5, 2023, and June 6, 2023.

Measures

To assess the feasibility of the cine-VR training program, we measured recruitment, retention, length of time required to recruit, rate of completion of the cine-VR training, and feasibility of the data collection measures. In addition, participants provided demographic factors (age, gender, race, ethnicity, year in program, and program) and completed the following measures:

Attitudes to Disability Scale

The Attitudes to Disability Scale [16] is a 16-item validated scale that assesses attitudes toward disability and attitudes toward people with a disability. The 16 items were answered on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The scale included 4 domains or subscales with good internal consistency: Inclusion—focuses on attitudes toward the inclusion and exclusion for people with disabilities as well as the burden on families and society (Cronbach α=0.76), Discrimination—focuses on awareness of unfair and prejudicial treatment to people with disabilities (Cronbach α=0.74), Gains—focuses on positives of having a disability (Cronbach α=0.75), and Prospects—focuses on future hopes and whether disability impacts those hopes (Cronbach α=0.72) [16].

Responding to Elder Abuse in GERiAtric Care—Provider Questionnaire

The Responding to Elder Abuse in GERiAtric Care—Provider Questionnaire [17] is a validated 25-item scale that assesses health care provider preparedness to identify and manage elder abuse and neglect. For the purposes of this study, we used 8 of the 25 items. The items we used addressed self-efficacy to identify and manage elder abuse and neglect on a 0 to 10 scale. The 17 excluded items were written for practicing health care professionals. In our sample, the 8 self-efficacy questions demonstrated excellent internal consistency for identifying and managing elder abuse and neglect (Cronbach α=0.95) [17].

Jefferson Scale of Empathy Health Professionals Students Version

The Jefferson Scale of Empathy Health Professionals Students Version [18] is a 20-item validated scale that assesses empathy among health professional students. The 20 items were answered on a 7-point Likert scale, ranging from 1 (strongly disagree) to 7 (strongly agree). Scores are summed for a composite score, ranging from 20 to 140. The scale demonstrated good internal consistency (Cronbach α=0.85) [18].

Data Collection

Participants completed the assessments through Qualtrics, an electronic questionnaire service. To access the preassessment, participants scanned a QR code from a Microsoft PowerPoint slide that directed them to the Qualtrics website. The Qualtrics website included a brief description of the study, the informed consent form, demographic questions, and the 3 measures. All participants provided informed consent without a signature in Qualtrics before completing the measures. After the cine-VR

XSL•FO
RenderX

training, participants scanned a different QR code that directed them to the postassessment measures.

We collected data anonymously. To link participants' pre- and postassessment responses, we included 3 questions at the beginning of the pre- and postassessments, which served as a unique identifier (ie, model of first car, high school mascot, and the day of the month on which they were born [eg, 31]). Total time to complete the pre- and postmeasures took approximately 15 - 20 minutes. Participants received a US $25 gift card for human subject compensation. To maintain anonymity, participants clicked on a new Qualtrics link that was not connected to their pre- or postmeasures to receive a gift card; it was possible that participants received compensation without completing all pre- and postmeasures.

Statistical Analysis

First, we assessed participants' demographic factors using descriptive statistics and presented them as means and SDs or sample size and percentages. Next, we examined the distribution of the data to ensure they met statistical test assumptions. To assess changes pre- and post–cine-VR training, we conducted paired *t* tests; all test statistics are presented with α values and degrees of freedom. We also calculated effect sizes using Cohen *d*, with a small effect=0.2, medium effect=0.5, and a large effect=0.8. We defined statistical significance as a *P* value less than .05 and conducted analyses in IBM SPSS statistical software (version 29.0).

Results

Feasibility

A total of 65 health professional students participated in and completed the full cine-VR training for 100% retention. Participants did not report any technological difficulties or adverse effects from wearing the head-mounted displays or viewing the 360-degree video. We scheduled 4 cine-VR trainings in a 6-month period to recruit the 65 participants, and all 4 trainings lasted approximately 60 minutes. For data collection, 65 participants completed the preassessment measures; however, only 48 completed the postassessment measures, resulting in a 74% completion rate. The 74% rate of completion suggests that we may want to consider adaptations or refinements to our selected measures. Overall, the ease of recruitment, length of time required to recruit, retention in the cine-VR trainings, and feasibility of data collection methods suggest that the cine-VR training was feasible.

Demographics

A total of 48 health professional students consented to participate in the pilot study and completed all pre- and postassessments. The mean age of participants was 22.5 (SD 3) (see Table 1). A total of 41 participants (85.4%) self-identified as women, 6 (12.5%) self-identified as men, and 1 (2.1%) self-identified as nonbinary. The participants self-identified their race as follows: 6.3% (3/48) Asian or Pacific Islander, 18.8% (9/48) Black or African American, 4.2% (2/48) Mixed Race, and 70.8% (34/48) White. For ethnicity, 2 participants (4.2%) self-identified as Hispanic, Latino, or Spanish origin and 2 (4.2%) self-identified as Mexican, Mexican American, or Chicano. The majority of participants were in the final year of their program (37.5%, 18/48), with premedicine as the most reported major (42.6%, 20/48).

XSL•FO
RenderX

Table . Participant demographic characteristics in the cinematic-virtual reality (cine-VR) study (N=48).

Variable	Results
Age (years), mean (SD)	22.5 (3)
Gender, n (%)	
Woman	41 (85.4)
Man	6 (12.5)
Nonbinary	1 (2.1)
Transgender	0 (0)
Genderqueer	0 (0)
An identity not listed	0 (0)
Ethnicity, n (%)	
Hispanic, Latino, or Spanish origin	2 (4.2)
Mexican, Mexican American, or Chicano	2 (4.2)
Non-Hispanic, Non-Latino, or Non-Spanish	44 (91.6)
Race, n (%)	
American Indian or Pacific Islander	0 (0)
Asian or Pacific Islander	3 (6.3)
Black or African American	9 (18.8)
Middle Eastern	0 (0)
Mixed Race	2 (4.2)
White	34 (70.8)
Program [a], n (%)	
Child Life Specialist	18 (38.3)
Exercise Physiology	1 (2.1)
Nutrition	4 (8.5)
Pre-Medicine	20 (42.6)
Pre-Physician Assistant	2 (4.3)
Psychology	2 (4.3)
Year in Program [b], n (%)	
Year 1	9 (18.8)
Year 2	10 (20.8)
Year 3	9 (18.8)
Year 4	18 (37.5)

[a]Missing values “Program” (n=1).

[b]"Year in Program” (n=2).

Attitudes Toward Disability Findings

Pre- and postdomain scores for the Attitudes to Disability Scale can be found in Table 2. Post–cine-VR training, we observed significant changes in 1 of the 4 Attitudes to Disability Subscales: Discrimination Domain (t_{47}=−3.97, P<.001). This change had a Cohen d=0.55, indicating a medium effect. This finding suggests that the cine-VR training may have increased participants’ awareness of discrimination towards people with a disability. In the Discrimination Domain, 3 of the 4 items showed a significant change with more participants “agreeing” and “strongly agreeing” with the following statements posttraining: “People often make fun of disabilities” (t_{47}=−3.92; P<.001); “People tend to become impatient with those with a disability” (t_{47}=−3.19; P=.003); and “People tend to treat those with a disability as if they have no feelings” (t_{47}=−2.69; P=.01). Post-cine-VR training, we did not observe significant changes in the Inclusion Domain (t_{47}=−.16; P=.88), Gains Domain (t_{47}=−1.76; P=.08), or Prospects Domain (t_{47}=−.30; P=.77).

XSL•FO
RenderX

Table . Participants' attitudes to disability scale mean scores before and after the cinematic-virtual reality (cine-VR) training program (N=48).

Questions	Pre-VR[a], mean (SD)	Post-VR, mean (SD)	*t* test (*df*)	*P* value
Inclusion domain				
People with a disability find it harder than others to make new friends.	3.9 (0.6)	3.8 (0.8)	.948 (47)	.35
People with a disability have problems getting involved in society.	3.5 (1.2)	3.8 (0.8)	−2.047 (47)	.046
People with a disability are a burden on society.	1.2 (0.4)	1.3 (0.6)	−1.000 (47)	.32
People with a disability are a burden on their family.	1.9 (0.9)	1.7 (1.0)	1.944 (47)	.06
Total	2.6 (0.6)	2.6 (0.5)	−.156 (47)	.88
Discrimination domain				
People often make fun of disabilities.	3.9 (0.6)	4.2 (0.6)	−3.923 (47)	<.001
People tend to become impatient with those with a disability.	3.9 (0.8)	4.3 (0.7)	−3.186 (47)	.003
People tend to treat those with a disability as if they have no feelings.	3.5 (1.1)	3.9 (0.8)	−2.687 (47)	.01
People with a disability are easier to take advantage of (exploit or treat badly) compared with other people.	3.7 (0.9)	3.9 (1.0)	−1.095 (47)	.28
Total	3.8 (0.6)	4.1 (0.6)	−3.967 (47)	<.001
Gains domain				
Having a disability can make someone a stronger person.	3.8 (0.7)	3.9 (0.8)	−1.219 (47)	.23
Having a disability can make someone a wiser person.	3.7 (0.8)	3.7 (0.8)	.227 (47)	.82
Some people achieve more because of their disability (eg, they are more successful).	3.0 (0.8)	3.2 (0.8)	−1.741 (47)	.09
People with a disability are more determined than others to reach their goals.	3.0 (0.7)	3.1 (0.6)	−1.401 (47)	.17
Total	3.3 (0.5)	3.5 (0.6)	−1.764 (47)	.08
Prospects domain				
Sex should not be discussed with people with disabilities.	1.6 (0.9)	1.4 (0.7)	1.752 (47)	.09
People should not expect too much from those with a disability.	1.5 (0.5)	1.6 (0.8)	−1.182 (47)	.24
People with a disability should not be optimistic (hopeful) about their future	1.3 (0.7)	1.5 (0.9)	−1.543 (47)	.13
People with a disability have less to look forward to than others.	1.6 (0.7)	1.6 (0.7)	.423 (47)	.67

XSL•FO
RenderX

Questions	Pre-VR[a], mean (SD)	Post-VR, mean (SD)	*t* test (*df*)	*P* value
Total	1.5 (0.5)	1.5 (0.6)	−.300 (47)	.77

[a]VR: virtual reality.

Self-Efficacy to Identify and Manage Elder Abuse and Neglect Findings

Post—cine-VR training, we observed significant changes in participants' self-efficacy to identify and manage elder abuse and neglect (t_{47}=−3.36, *P*=.002, Table 3). This change had a Cohen *d*=0.49, approaching a medium effect.

XSL•FO
RenderX

Table . Participants' perceived self-efficacy for identifying and managing elder abuse and neglect before and after the cinematic-virtual reality (cine-VR) training program (N=48).

Questions	Pre-VR,[a] mean (SD)	Post-VR, mean (SD)	*t* test (*df*)	*P* value	Cohen *d*
Asking questions about abuse to an older patient who has clear indications of now being, or having previously been, subjected to abuse.	6.2 (2.4)	7.2 (2.4)	−3.354 (47)	.002	.48
Asking questions about abuse to an older patient who has no clear indications of now being or having previously been, subjected to abuse.	5.7 (2.6)	5.9 (2.4)	−.585 (47)	.56	.08
Ensuring you are able to ask questions about abuse in private to an older patient who has a relative who insists on being present during all contact.	6.1 (2.7)	7.0 (2.4)	−2.911 (47)	.005	.42
In conversation, providing support to an older patient who tells about abuse.	7.4 (2.6)	8.0 (2.1)	−2.438 (47)	.02	.35
Helping an older patient subjected to abuse on to the right body in health care, or to the right support function in society.	6.6 (2.9)	7.7 (2.1)	−3.174 (47)	.003	.46
Helping an older patient subjected to abuse to make a report to the police or social services.	6.9 (2.8)	7.6 (2.2)	−1.984 (47)	.05	.29
Helping and supporting an older patient subjected to abuse, who does not currently want to change his or her situation.	5.5 (2.9)	6.3 (2.6)	−2.433 (47)	.02	.35
Handling the meeting with an older patient who says no to questions about abuse, but where you still have strong suspicions that the patient is subjected to abuse.	5.7 (2.8)	6.6 (2.2)	−2.601 (47)	.01	.38
Total	6.3 (2.4)	7.0 (1.8)	−3.364 (47)	.002	.49

[a]VR: virtual reality.

In total, 6 of the 8 items demonstrated significant improvements in perceived self-efficacy. Specifically, participants reported increases in their self-efficacy to ask questions about abuse to an older individual who has clear indications of being abused (t_{47}=−3.354; *P*=.002; Cohen *d*=0.48) and to ensure that they ask questions about abuse in private (t_{47}=−2.911; *P*=.005; Cohen *d*=0.42). We also observed increases in self-efficacy for providing support to an older individual who discloses abuse (t_{47}=−2.438; *P*=.02; Cohen *d*=0.35) and directing an older adult experiencing abuse to the right support person or service

XSL•FO
RenderX

(t_{47}=−3.174; P=.003; Cohen d=0.46). Finally, participants showed improvements in their self-efficacy for helping older adults experiencing abuse who do not want to change their situation (t_{47}=−2.433; P=.02; Cohen d=0.35) and handling an older person who says no to questions about abuse despite strong suspicions of abuse (t_{47}=−2.601; P=.01; Cohen d=0.38).

One item showed a trend in improving self-efficacy to help an older adult make a report to the police or social services (t_{47}=−1.984; P=.05; Cohen d=0.29). Participants did not show an improvement in their self-efficacy to ask questions about abuse to an older adult who has no clear indications of presently or previously experiencing abuse (t_{47}=−.585; P=.56; Cohen d=0.08). This finding highlights an area to be addressed in future research.

Empathy Findings

Pre–cine-VR training empathy scores ranged from 47 to 128, with a mean of 105.7 (SD 16.6). Post–cine-VR training empathy scores ranged from 80 to 131, with a mean of 110.4 (SD 12.1). We observed a significant increase in participants' empathy scores post–cine-VR training (mean change=4.8; t_{47}=−2.329; P=.02), with a Cohen d=0.34 indicating a small effect. This finding indicates a noticeable difference in empathy before and after participating in the cine-VR training. While the effect size is small, an increase in empathy may have practical significance in training health professional students about the complexity of diabetes management. More research is needed to confirm this finding.

Discussion

Principal Findings

In this pilot study, we assessed the feasibility of implementing a cine-VR training program with health professional students. Our findings showed that we were able to recruit, implement, and retain participants for the full cine-VR training. Furthermore, participants reported no adverse effects or issues with the cine-VR technology. In total, 3-quarters of the participants completed the pre- and postmeasures for the pilot study. Overall, we observed significant improvements in participants' awareness of discrimination toward people with disability, self-efficacy to identify and manage elder abuse and neglect, and empathy. These improvements suggest that the cine-VR training program may be an effective teaching modality to educate health professional students about the complexity of care in older adults. More research is needed with a larger, more diverse sample and a proper attention control condition to confirm its effectiveness.

An objective of the cine-VR training program was to raise awareness about discrimination towards aging and disabilities. The cine-VR training portrayed multiple interactions among Mr Chen, his family, and health care providers to raise awareness about discrimination. In one simulation, for instance, Mr Chen's grandson is depicted yelling "Good morning!" and laughing, a reaction to his grandfather's hearing loss. Later in the same simulation, his daughter-in-law appears visibly frustrated following an incident of urinary incontinence. These interactions were included to show that sometimes people, including family members, make fun of disabilities or become impatient with people who have a disability. Similarly, the cine-VR training included episodes involving the primary care physician and urgent care physician, directing their questions about Mr Chen's medical history to his son and daughter-in-law, rather than speaking to Mr Chen directly. This was done intentionally to suggest that the providers did not think Mr Chen was capable of discussing his own health. In addition, we included an insensitive comment from the urgent care physician directed at Mr Chen poking fun of his situation. These scenarios were designed to illustrate various types of discrimination and elicit emotional reactions from the participants. Learning that involves emotions increases memory retention through emotional engagement, enhanced cognitive processing, and memory consolidation [19,20]. The brain tends to prioritize emotionally significant information for long-term memory over neutral information, which may explain the increased awareness of discrimination observed in the findings [21].

The cine-VR training program also emphasized recognizing the signs of elder abuse and neglect, how to report abuse, and the subsequent actions taken after a report is made. Each episode uncovered new signs of elder abuse and neglect involving Mr Chen (eg, withholding fluids throughout the day to prevent urinary incontinence, not refilling Mr Chen's medications, and making Mr Chen live in the garage while the house is being renovated), alongside behaviors commonly exhibited by abusers. The follow-up debriefs and reflections concentrated on the responsibilities of mandated reporters, the reporting process, and the steps that follow a report. This content likely contributed to the participants' increased self-efficacy in asking older adults about abuse, ensuring questions asked about abuse are asked in private, providing emotional support to an older adult who tells you about abuse, helping an older adult subjected to abuse to the right authorities, helping an older adult subjected to abuse who does not want to change circumstances, and handling an older adult who denies abuse but you have strong suspicions that they are subjected to abuse. Additional research is needed to confirm if these increases in self-efficacy are sustained over time.

Comparison With Previous Work

Medical training uses a variety of simulation types to prepare health professional students for clinical practice. The most common types of simulation include manikins and advanced patient simulators, standardized patients, and task trainers [22]. The first manikin-based simulation can be traced back to the 1960s with the development of Resusci Anne for CPR training [23]. Since then, health professional students have used manikins and advanced patient simulators to replicate complex medical scenarios, including cardiac arrest, childbirth delivery, and trauma resuscitation [23]. The advantages of using manikins and advanced patient simulators include a risk-free setting for repeated practice of clinical skills and standardized training experiences [24-26]. Drawbacks of using manikins and advanced patient simulators include the high purchase price, ongoing maintenance costs, and the technological limitations that prevent them from fully replicating the complexity and unpredictability of human behavior [27,28]. Conversely, standardized patients

XSL•FO
RenderX

provide a more realistic interaction for students. Standardized patients have been shown to improve communication skills, critical thinking, self-efficacy, and clinical examination skills [29]. Despite training, standardized patients may show individual differences that can lead to variations in role portrayal, impacting the consistency of the simulated learning experience [30]. Unlike manikins, advanced patient simulators, and standardized patients, a task trainer is a specific device designed to help train health professional students on how to perform specific procedures, such as inserting an intravenous line or performing a lumbar puncture. Task trainers allow for focused learning, skill mastery, and standardization of training in a safe learning environment equipped to provide feedback and assessment; however, they are limited in scope, lack human interaction, and require regular maintenance and upkeep to remain functional [22].

Cine-VR builds on these traditional simulation methods by integrating narrative storytelling and interactive visual experiences [13,14,31,32]. Unlike manikins, advanced patient simulators, and task trainers, cine-VR offers an immersive environment that shows real-life scenarios, situational complexity, and human emotions [33]. Furthermore, cine-VR can provide both physical and decision-based feedback through interactive episodes, whereas task trainers only provide physical feedback for procedural skills. Furthermore, standardized patients offer realistic encounters with patients; however, actors can vary in their portrayal of patients, which affects the consistency of the training. Cine-VR can standardize patient interactions while also allowing for variations in patient interactions to balance consistency with the flexibility to adapt to human unpredictability. Altogether, cine-VR incorporates elements from manikins, advanced patient simulators, standardized patients, and task trainers into a single comprehensive training tool [33].

Our cine-VR training program has several advantages compared with other simulation approaches. First, our cine-VR was standardized, safe, and reusable. If needed, participants can repeat the training as often as desired to learn or reinforce knowledge and clinical skills. In addition, participants could practice difficult, high-stakes conversations through our "guided simulations," without any risk or harm to the patient, Mr Chen. Next, the multisensory experience of cine-VR engages emotional learning. Research on emotional learning shows that it increases long-term memory as well as increased acquisition of skills [34-36]. Finally, cine-VR training offers participants a glimpse into the lives of their patients, thereby increasing their awareness of another person's feelings and experiences. This form of perspective-taking increases empathy similar to the findings in our pilot study [32,37,38]. A 2017 study by Schutte and Silinovic [39] compared the impact of VR and traditional didactic learning on empathy among undergraduate students. They found participants in the VR group reported greater increases in empathy compared to the didactic presentation group. Thus, cine-VR has the potential to increase empathy in health professional students; more research is needed to assess empathy and other prosocial behaviors (ie, helping, sharing, and comforting) in future research. Considering patients value empathy from their health care professionals, often as much or even more than clinical expertise, simulation training that enhances empathy is crucial [40,41].

Limitations

Limitations of this study included the small sample, data collected from one site, heterogeneity of health care specialties, selection bias, subject bias, and lack of a control group. We recruited a sample of 65 participants for the training, and of those, 48 completed the pre- and postassessment. While the sample of 48 participants was small, it was sufficient to pilot the measures for our pilot study. Next, we collected data from one university in the Midwest, which limits the generalizability of our findings to all health professional students. Furthermore, participants represented 6 different health professional fields. Geriatric care is inherently interprofessional and typically requires collaboration from multiple health care professionals. However, the small sample size across 6 health professional programs limits the ability to draw conclusions about the cine-VR training program in these fields. In addition, students who volunteered to participate may have been more willing or motivated to participate in this cine-VR training on care for older adults. Thus, the findings may be susceptible to subject bias and social desirability bias. Finally, we did not include a control condition as a comparison group. Future research with a large sample size must include a control condition to determine the effectiveness of the cine-VR training. In future work, we plan to conduct a randomized controlled trial comparing the effectiveness of the cine-VR training to simulated patients. Simulated patients are used extensively in medical training to enhance the educational experience of health professional students; however, simulated patients can be costly and time-consuming. If we demonstrate that the cine-VR training is more effective or equally effective as simulated patients, cine-VR training may be a more sustainable, cost-effective approach to refine health professional students' clinical and communication skills without the risk of harming real patients.

Conclusions

We demonstrated that our cine-VR training program was feasible and acceptable to health professional students at our university in the Midwestern United States. Our findings suggest that the training increased awareness of discrimination towards people with disability, improved self-efficacy to identify and manage elder abuse and neglect, and increased empathy. More research is needed to confirm the effectiveness of this cine-VR training program on the complexity of care for older adults. Future research should use a randomized controlled trial to compare the cine-VR to an attention control condition, such as simulated patients. If confirmed, the cine-VR training program may be a new, effective approach to learning about multimorbidity, geriatric syndromes, and biases related to aging.

XSL•FO
RenderX

Acknowledgments

The research was funded by the Osteopathic Heritage Foundation Ralph S Licklider, DO Endowed Professor awarded to EAB. The creation of the virtual reality simulations was part of the Medicaid Care Experience Simulation funded by the Ohio Department of Medicaid and administered by the Ohio Colleges of Medicine Government Resource Center. The views expressed in the virtual reality simulations of this publication are solely those of the creators and do not represent the views of the state of Ohio or federal Medicaid programs.

Authors' Contributions

EAB, SM, ML, and CL provided substantial contributions to conception and design, acquisition of data, and analysis and interpretation of data. EAB, SM, ML, and CL drafted the article or revised it critically for important intellectual content. EAB, SM, ML, and CL gave final approval of the version of the article to be published. EAB, SM, ML, and CL agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Conflicts of Interest

None declared.

References

1. Mather M, Scommegna P. Fact sheet: aging in the United States. : Population Reference Bureau; 2024 URL: https://www.prb.org/resources/fact-sheet-aging-in-the-united-states/
2. Tavares JL, Cohen MA, Silberman S, Popham L. Chronic inequities: measuring disease burden among older adults in the US: a health and retirement study analysis.: National Council on Aging; 2022.
3. Stierman B, Afful J, Carroll MD, et al. National Health and Nutrition Examination Survey 2017-March 2020 Prepandemic Data Files-Development of Files and Prevalence Estimates for Selected Health Outcomes. Natl Health Stat Report 2021 Jun 14(158). [doi: 10.15620/cdc:106273] [Medline: 39380201]
4. Siew AL, Wong JW, Chan EY. Effectiveness of simulated patients in geriatric education: A scoping review. Nurse Educ Today 2021 May;100:104856. [doi: 10.1016/j.nedt.2021.104856] [Medline: 33740706]
5. Urtamo A, Jyväkorpi SK, Strandberg TE. Definitions of successful ageing: a brief review of a multidimensional concept. Acta Biomed 2019 May 23;90(2):359-363. [doi: 10.23750/abm.v90i2.8376] [Medline: 31125022]
6. Noto S. Perspectives on aging and quality of life. Healthcare (Basel) 2023 Jul 26;11(15):2131. [doi: 10.3390/healthcare11152131] [Medline: 37570372]
7. Sanford AM, Morley JE, Berg-Weger M, et al. High prevalence of geriatric syndromes in older adults. PLoS One 2020;15(6):e0233857. [doi: 10.1371/journal.pone.0233857] [Medline: 32502177]
8. Arai H, Chen LK. Aging populations and perspectives of geriatric medicine in Japan. Glob Health Med 2024 Feb 29;6(1):1-5. [doi: 10.35772/ghm.2024.01001] [Medline: 38450117]
9. O'Rourke LA, Morrison M, Grimsley A, Cotter VT. High-fidelity simulation and nurse clinical competence-an integrative review. J Clin Nurs 2023 May;32(9-10):1549-1555. [doi: 10.1111/jocn.16028] [Medline: 34453385]
10. Gaba DM. The future vision of simulation in health care. Qual Saf Health Care 2004 Oct;13 Suppl 1(Suppl 1):i2-10. [doi: 10.1136/qhc.13.suppl_1.i2] [Medline: 15465951]
11. Herrera-Aliaga E, Estrada LD. Trends and innovations of simulation for twenty first century medical education. Front Public Health 2022;10:619769. [doi: 10.3389/fpubh.2022.619769] [Medline: 35309206]
12. Sapkaroski D, Mundy M, Dimmock MR. Virtual reality versus conventional clinical role-play for radiographic positioning training: A students' perception study. Radiography (Lond) 2020 Feb;26(1):57-62. [doi: 10.1016/j.radi.2019.08.001] [Medline: 31902456]
13. Beverly E, Rigot B, Love C, Love M. Perspectives of 360-degree cinematic virtual reality: interview study among health care professionals. JMIR Med Educ 2022 Apr 29;8(2):e32657. [doi: 10.2196/32657] [Medline: 35486427]
14. Beverly EA, Love C, Love M, Williams E, Bowditch J. Using virtual reality to improve health care providers' cultural self-efficacy and diabetes attitudes: pilot questionnaire study. JMIR Diabetes 2021 Jan 27;6(1):e23708. [doi: 10.2196/23708] [Medline: 33502335]
15. Lancaster GA, Dodd S, Williamson PR. Design and analysis of pilot studies: recommendations for good practice. J Eval Clin Pract 2004 May;10(2):307-312. [doi: 10.1111/j..2002.384.doc.x] [Medline: 15189396]
16. Power MJ, Green AM, WHOQOL-DIS Group. The Attitudes to Disability Scale (ADS): development and psychometric properties. J Intellect Disabil Res 2010 Sep;54(9):860-874. [doi: 10.1111/j.1365-2788.2010.01317.x] [Medline: 20712699]
17. Simmons J, Wenemark M, Ludvigsson M. Development and validation of REAGERA-P, a new questionnaire to evaluate health care provider preparedness to identify and manage elder abuse. BMC Health Serv Res 2021 May 19;21(1):473. [doi: 10.1186/s12913-021-06469-2] [Medline: 34006271]
18. Hojat M, Gonnella J, Maxwell K. Jefferson Scales of Empathy (JSE) Professional Manual & User 's Guide: Jefferson Medical College; 2009.

XSL•FO
RenderX

19. Grassini S, Laumann K, Rasmussen Skogstad M. The use of virtual reality alone does not promote training performance (but sense of presence does). Front Psychol 2020;11:1743. [doi: 10.3389/fpsyg.2020.01743] [Medline: 32765384]
20. Slater M, Wilbur S. A Framework for Immersive Virtual Environments (FIVE): speculations on the role of presence in virtual environments. Presence: Teleoperators & Virtual Environments 1997 Dec;6(6):603-616. [doi: 10.1162/pres.1997.6.6.603]
21. LeDoux JE. Emotion circuits in the brain. Annu Rev Neurosci 2000;23:155-184. [doi: 10.1146/annurev.neuro.23.1.155] [Medline: 10845062]
22. Levine AI, Demaria S, Schwartz AD, Sim AJ. The Comprehensive Textbook of Healthcare Simulation: Springer; 2013.
23. Elendu C, Amaechi DC, Okatta AU, et al. The impact of simulation-based training in medical education: A review. Medicine (Baltimore) 2024 Jul 5;103(27):e38813. [doi: 10.1097/MD.0000000000038813] [Medline: 38968472]
24. Okuda Y, Bryson EO, DeMaria S Jr, et al. The utility of simulation in medical education: what is the evidence? Mt Sinai J Med 2009 Aug;76(4):330-343. [doi: 10.1002/msj.20127] [Medline: 19642147]
25. Issenberg SB, McGaghie WC, Petrusa ER, Lee Gordon D, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. Med Teach 2005 Jan;27(1):10-28. [doi: 10.1080/01421590500046924] [Medline: 16147767]
26. Ziv A, Wolpe PR, Small SD, Glick S. Simulation-based medical education: an ethical imperative. Acad Med 2003 Aug;78(8):783-788. [doi: 10.1097/00001888-200308000-00006] [Medline: 12915366]
27. Cooper JB, Taqueti VR. A brief history of the development of mannequin simulators for clinical education and training. Postgrad Med J 2008 Nov;84(997):563-570. [doi: 10.1136/qshc.2004.009886] [Medline: 19103813]
28. Gordon JA, Wilkerson WM, Shaffer DW, Armstrong EG. "Practicing" medicine without risk: students' and educators' responses to high-fidelity patient simulation. Acad Med 2001 May;76(5):469-472. [doi: 10.1097/00001888-200105000-00019] [Medline: 11346525]
29. Hillier M, Williams TL, Chidume T. Standardization of Standardized Patient Training in Medical Simulation: StatPearls Treasure Island (FL); 2025.
30. May W, Park JH, Lee JP. A ten-year review of the literature on the use of standardized patients in teaching and learning: 1996-2005. Med Teach 2009 Jun;31(6):487-492. [doi: 10.1080/01421590802530898] [Medline: 19811163]
31. Beverly EA, Love C, Love M, Lammert L, Bowditch J. Cinematic virtual reality for educating health care providers about type 2 diabetes, disability, and elder abuse and neglect: a pilot study. J Diabetes Sci Technol 2023 Sep;17(5):1160-1171. [doi: 10.1177/19322968231171586] [Medline: 37114917]
32. McCalla MM, Jones D, Grice R, et al. Feasibility of a cinematic-virtual reality training program for health professional students: a single-arm pre-post study. J Diabetes Sci Technol 2023 Sep;17(5):1181-1189. [doi: 10.1177/19322968231171136] [Medline: 37138539]
33. The Power of Virtual Reality Cinema for Healthcare Training: A Collaborative Guide for Medical Experts and Media Professionals: Routledge; 2022.
34. Tyng CM, Amin HU, Saad MNM, Malik AS. The influences of emotion on learning and memory. Front Psychol 2017;8:1454. [doi: 10.3389/fpsyg.2017.01454] [Medline: 28883804]
35. Steadman RH, Coates WC, Huang YM, et al. Simulation-based training is superior to problem-based learning for the acquisition of critical assessment and management skills. Crit Care Med 2006 Jan;34(1):151-157. [doi: 10.1097/01.ccm.0000190619.42013.94] [Medline: 16374169]
36. Kyaw BM, Saxena N, Posadzki P, et al. Virtual reality for health professions education: systematic review and meta-analysis by the Digital Health Education Collaboration. J Med Internet Res 2019 Jan 22;21(1):e12959. [doi: 10.2196/12959] [Medline: 30668519]
37. Tong X, Gromala D, Kiaei Ziabari SP, Shaw CD. Designing a virtual reality game for promoting empathy toward patients with chronic pain: feasibility and usability study. JMIR Serious Games 2020 Aug 7;8(3):e17354. [doi: 10.2196/17354] [Medline: 32763883]
38. Herrera F, Bailenson J, Weisz E, Ogle E, Zaki J. Building long-term empathy: A large-scale comparison of traditional and virtual reality perspective-taking. PLoS One 2018;13(10):e0204494. [doi: 10.1371/journal.pone.0204494] [Medline: 30332407]
39. Schutte NS, Stilinović EJ. Facilitating empathy through virtual reality. Motiv Emot 2017 Dec;41(6):708-712. [doi: 10.1007/s11031-017-9641-7]
40. Hirsch EM. The role of empathy in medicine: a medical student's perspective. Virtual Mentor 2007 Jun 1;9(6):423-427. [doi: 10.1001/virtualmentor.2007.9.6.medu1-0706] [Medline: 23218048]
41. Halpern J. What is clinical empathy? J Gen Intern Med 2003 Aug;18(8):670-674. [doi: 10.1046/j.1525-1497.2003.21017.x] [Medline: 12911651]

Abbreviations

cine-VR: cinematic–virtual reality
VR: virtual reality

XSL•FO
RenderX

Edited by H LaMonica; submitted 22.07.24; peer-reviewed by J Wardian, X Jing, Y Weng; revised version received 21.01.25; accepted 21.01.25; published 12.02.25.

Please cite as:
Beverly EA, Miller S, Love M, Love C
Feasibility of a Cinematic–Virtual Reality Program Educating Health Professional Students About the Complexity of Geriatric Care: Pilot Pre-Post Study
JMIR Aging 2025;8:e64633
URL: https://aging.jmir.org/2025/1/e64633
doi:10.2196/64633

© Elizabeth A Beverly, Samuel Miller, Matthew Love, Carrie Love. Originally published in JMIR Aging (https://aging.jmir.org), 12.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Dance as a Therapeutic Approach for Parkinson Disease Through the Social Robotics for Active and Healthy Ageing (SI-Robotics): Results From a Technical Feasibility Study

Roberta Bevilacqua[1], MScPH; Elvira Maranesi[1], PhD; Marco Benadduci[1], BPT; Gabriella Cortellessa[2], PhD; Alessandro Umbrico[2], PhD; Francesca Fracasso[2], PhD; Giovanni Melone[3], MSc; Arianna Margaritini[1], MScPH; Angela La Forgia[3], MSIS; Pierpaolo Di Bitonto[4], BCS; Ada Potenza[5], MSc; Laura Fiorini[6,7], PhD; Carlo La Viola[6], PhD; Filippo Cavallo[6,7], PhD; Alessandro Leone[8], BSc; Andrea Caroppo[8], MSCE; Gabriele Rescio[8], PhD; Mauro Marzorati[9], PhD; Amedeo Cesta[2], PhD; Giuseppe Pelliccioni[10], MD; Giovanni Renato Riccardi[11], MD; Lorena Rossi[1], MEng

[1]Scientific Direction, IRCCS INRCA, Via Santa Margherita 5, Ancona, Italy

[10]Clinical Unit of Neurology, Istituto di Ricovero e Cura a Carattere Scientifico Istituto Nazionale Ricovero e Cura per Anziani (IRCCS INRCA), Ancona, Italy

[11]Clinical Unit of Physical Rehabilitation, Istituto di Ricovero e Cura a Carattere Scientifico Istituto Nazionale Ricovero e Cura per Anziani (IRCCS INRCA), Ancona, Italy

[2]National Research Council–Institute of Sciences and Technologies for Cognition, Rome, Italy

[3]Exprivia S.p.a., Innovation Lab, Innovation, Marketing & Technology, Molfetta, Italy

[4]Department of Soil, Plant and Food Science, University of Bari Aldo Moro, Bari, Italy

[5]Grifo multimedia Srl, Bari, Italy

[6]Department of Industrial Engineering, University of Florence, Florence, Italy

[7]The BioRobotics Institute, Scuola Superiore Sant'Anna, Pontedera, Italy

[8]National Research Council of Italy – Institute for Microelectronics and Microsystems, Lecce, Italy

[9]Institute of Biomedical Technologies, National Research Council, Segrate, Italy

Corresponding Author:
Elvira Maranesi, PhD
Scientific Direction, IRCCS INRCA, Via Santa Margherita 5, Ancona, Italy

Abstract

Background: Parkinson disease (PD) is a progressive neurodegenerative disorder characterized by motor symptoms. Recently, dance has started to be considered an effective intervention for people with PD. Several findings in the literature emphasize the necessity for deeper exploration into the synergistic impacts of dance therapy and exergaming for PD management. Moreover, socially engaging robotic platforms equipped with advanced interaction and perception features offer potential for monitoring patients' posture and enhancing workout routines with tailored cues.

Objective: This paper presents the results of the Social Robotics for Active and Healthy Ageing (SI-Robotics) project, aimed at designing an innovative rehabilitation program targeted at seniors affected by (early-stage) PD. This study therefore aims to assess the usefulness of a dance-based rehabilitation program enriched by artificial intelligence–based exergames and contextual robotic assistance in improving motor function, balance, gait, and quality of life in patients with PD. The acceptability of the system is also investigated.

Methods: The study is designed as a technical feasibility pilot to test the SI-Robotics system. For this study, 20 patients with PD were recruited. A total of 16 Irish dance–based rehabilitation sessions of 50 minutes were conducted (2 sessions per week, for 8 wks), involving 2 patients at a time. The designed rehabilitation session involves three main actors: (1) a therapist, (2) a patient, and (3) a socially interacting robot. To stimulate engagement, sessions were organized in the shape of exergames where an avatar shows patients the movements they should perform to correctly carry out a dance-based rehabilitation exercise.

Results: Statistical analysis reveals a significant difference on the Performance-Oriented Mobility Assessment scale, both on balance and gait aspects, together with improvements in Short Physical Performance Battery, Unified Parkinson Disease Rating Scale–III, and Timed Up and Go test, underlying the usefulness of the rehabilitation intervention on the motor symptoms of PD. The analysis of the Unified Theory of Acceptance and Use of Technology subscales provided valuable insights into users' perceptions and interactions with the system.

Conclusions: This research underscores the promise of merging dance therapy with interactive exergaming on a robotic platform as an innovative strategy to enhance motor function, balance, gait, and overall quality of life for patients grappling with PD.

XSL•FO
RenderX

Trial Registration: ClinicalTrials.gov NCT05005208; https://clinicaltrials.gov/study/NCT05005208

(JMIR Aging 2025;8:e62930) doi:10.2196/62930

KEYWORDS

Parkinson disease; rehabilitation; Irish dancing; balance; gait; socially interacting robot

Introduction

Parkinson disease (PD) is a progressive neurodegenerative disorder characterized by motor symptoms, such as tremors, bradykinesia, rigidity, and postural instability, as well as nonmotor symptoms, including cognitive impairment, mood disorders, and autonomic dysfunction [1]. These symptoms can significantly impact patients' quality of life and independence as the disease progresses. While pharmacological interventions, such as dopamine replacement therapy, remain the mainstay of PD management, their efficacy may diminish over time and they often fail to address the nonmotor symptoms adequately [2]. Physical therapy represents the key adjuvant treatment for PD, offering potential benefits in improving motor function, balance, gait, and overall mobility [3]. Among various forms of physical therapy, dance-based interventions have gained attention due to their multifaceted approach targeting motor, cognitive, and psychosocial aspects of the disease [4]. Dance therapy engages patients in rhythmic movements, music, and social interaction, promoting coordination, flexibility, and emotional well-being [5].

Exergaming, which involves interactive video games requiring physical movement, has emerged as a novel approach to delivering dance-based therapy in both home-based and clinical settings [6]. Exergames offer the advantages of personalized, adaptable, and enjoyable exercise programs, potentially enhancing patient motivation and adherence [7]. Recent research has highlighted the potential of dance therapy and exergaming to address the complex needs of PD patients. A randomized controlled trial by Duncan and Earheart [4] demonstrated significant improvements in motor function and balance among patients with PD participating in community-based dancing compared with a control group. Furthermore, a study by Barry et al [7] found that exergaming interventions were effective in improving mobility and reducing fall risk in patients with PD.

These findings underscore the importance of further investigating the combined effects of dance therapy and exergaming in PD management. Furthermore, the reliability and performance of current sensing technologies provide physiological data that are useful to monitor the state and the quality of the exercises objectively performed by patients. For example, commercial wearable devices could be easily integrated during rehabilitation sessions to gather information about the heart or breath rate of involved patients and objectively evaluate the metabolic effort. Socially interactive robotic platforms with their interaction and perception capabilities (eg, onboard cameras) are well suited to monitor the posture of patients and support the execution of the exercises through tailored stimuli. A more comprehensive approach may integrate advanced technologies, such as sensor technology for real-time monitoring, artificial intelligence (AI) techniques for personalized difficulty adjustments, and robotic platforms to deliver personalized feedback and enhance the rehabilitation experience.

In this context, the Social Robotics for Active and Healthy Ageing (SI-Robotics) project designed an innovative rehabilitation program targeted at seniors affected by (early stage) PD [8]. The approach relies on the integration of several AI technologies ranging from knowledge representation and reasoning, for user modeling and personalization to machine learning and automated planning (AP) for continuous proactive and adaptive assistance [9]. Leveraging the mentioned interventions based on different types of dance [10,11], the idea is to develop a technology-enhanced, dance-based rehabilitation program where a socially interacting robot and several sensing devices (wearable sensors and 3D cameras mainly) support therapists and patients during the execution of the exercises. This study therefore aims to assess the usefulness of a dance-based physical therapy program enriched by AI-based exergames and contextual robotic assistance in improving motor function, balance, gait, and quality of life in patients with PD. We hypothesized that the proposed technology-enhanced dance therapy would lead to greater improvements in PD symptoms.

Methods

The study is designed as a technical feasibility pilot to test the SI-Robotics system. The entire protocol, including the description of scales, the platform functioning, the training program, and procedures has been previously described in detail [8].

Subjects

A total of 20 patients with PD were selected by the outpatient department at the Clinical Unit of Physical Rehabilitation, Istituto di Ricovero e Cura a Carattere Scientifico Istituto Nazionale Ricovero e Cura per Anziani (IRCCS INRCA). The patients were recruited if they were over 65 years old; able to provide informed consent; had a stage of Hoen and Yahr scale between 1 and 2 [12]; had a Functional Ambulation Category score ≥2 [13]; had a Rankin Scale score ≤3 [14]; had stability of drug treatment for at least 1 month; had a Geriatric Depression Scale 4-item score ≤1 [15]; had a Mini-Mental State Examination ≥24 [16]; and could maintain an upright posture≥30 seconds, evaluated by a trained physiotherapist during the recruitment. The evaluation of compliance with the inclusion and exclusion criteria was performed during the recruitment session. Once we completed this phase, informed consent was obtained. The patient's clinical assessment was performed at the start and the end of the treatment. In particular, the evaluation consisted of the administration of the following scale: measurement of functional state with the Barthel Index [17]; physical performance with Tinetti's Performance-Oriented

XSL•FO
RenderX

Mobility Assessment (POMA) [18], the Short Physical Performance Battery (SPPB) [19], the 6-Minute Walking Test [20], and the Timed Up and Go test (TUG) [21]; evaluation of the quality of life with the 12-Item Short-Form Health Survey (SF-12) [22]; fear of falling with Falls Efficacy Scale-International (FES-I) [23]; and the assessment of the prognosis of PD with the Unified Parkinson Disease Rating Scale – III (UPDRS-III) [24].

Intervention

Figure 1 shows the experimental setting in a protected environment (gymnasium) and the positioning of the technologies, the patient, and the physiotherapist, during the dance-based rehabilitation sessions.

Figure 1. Technological components and actors of Social Robotics for Active and Healthy Ageing (SI-Robotics) rehabilitation sessions.



A total of 16 therapy sessions of 50 minutes were conducted (2 sessions per wk, for 8 wks), involving 2 patients at a time. Cardiac and respiratory activity was monitored during dancing treatments to detect heart rate and breathing frequency. The participants were required to complete at least 80% of the sessions. Each session involved the following activities: (1) breathing, relaxation, and postural harmonization exercises (5 min); (2) active mobility and stretching exercises (5 min); (3) the SI-Robotics intervention (35 min) consisting of the AI-enhanced “Let’s Dance” game, and a socially interacting robot monitoring the execution of the physical exercises of patients; and (4) relaxation exercises (5 min).

After selecting a difficulty level, players are presented as dancers on the screen, along with footprints that suggest the movement to be done. Each task can vary from simple aerobic exercises (side steps, arm raising, hand clapping, etc) to choreographies. At the end of the game session, a score is presented to the patient. This would make a patient aware of the quality of the performance. This score can also be used by the therapist to assess the patient’s performance and the possible increase in the difficulty level of the game.

The next sections describe the technological components of the session in more detail.

The AI-Based Rehabilitation Session

The designed rehabilitation session involves three main actors: (1) a therapist, (2) a patient, and (3) a socially interacting robot. To stimulate engagement, sessions are organized in the shape of exergames where an avatar shows patients the movements they should perform to correctly carry out a rehabilitation exercise (eg, dancing steps). Given a set of data about the needs of a group of patients and about stimulation capabilities of known exercise, a game engine integrates an AI planner [25] and synthesizes exercises by selecting a (sub)set of movements (ie, dancing steps) that best fit the clinical objectives of the session (ie, of the current patient). The exergame is enriched with AP to synthesize suitable physical exercises, contextualized to the clinical objectives of the rehabilitation sessions [25]. Many works in the literature investigate the use of AI in health care and PD [11,26] AI is used for example to predict the wearing-off of symptoms [27], support decisions [28], or early diagnosis of PD [29]. The majority of these works adopt AI solutions based on deep or machine learning and focus on the diagnosis of the disease. Few works investigate the use of AP to support therapists in the synthesis of rehabilitation programs. The study by Gonzalez et al [30], for example, integrates AP into a control architecture to allow a social robot to physically show motions to a patient during physical rehabilitation. Baschieri et al [31] uses AP to synthesize simulation scenarios within a serious game for cognitive rehabilitation. This work pursues an objective similar to ours but in a different clinical scenario.

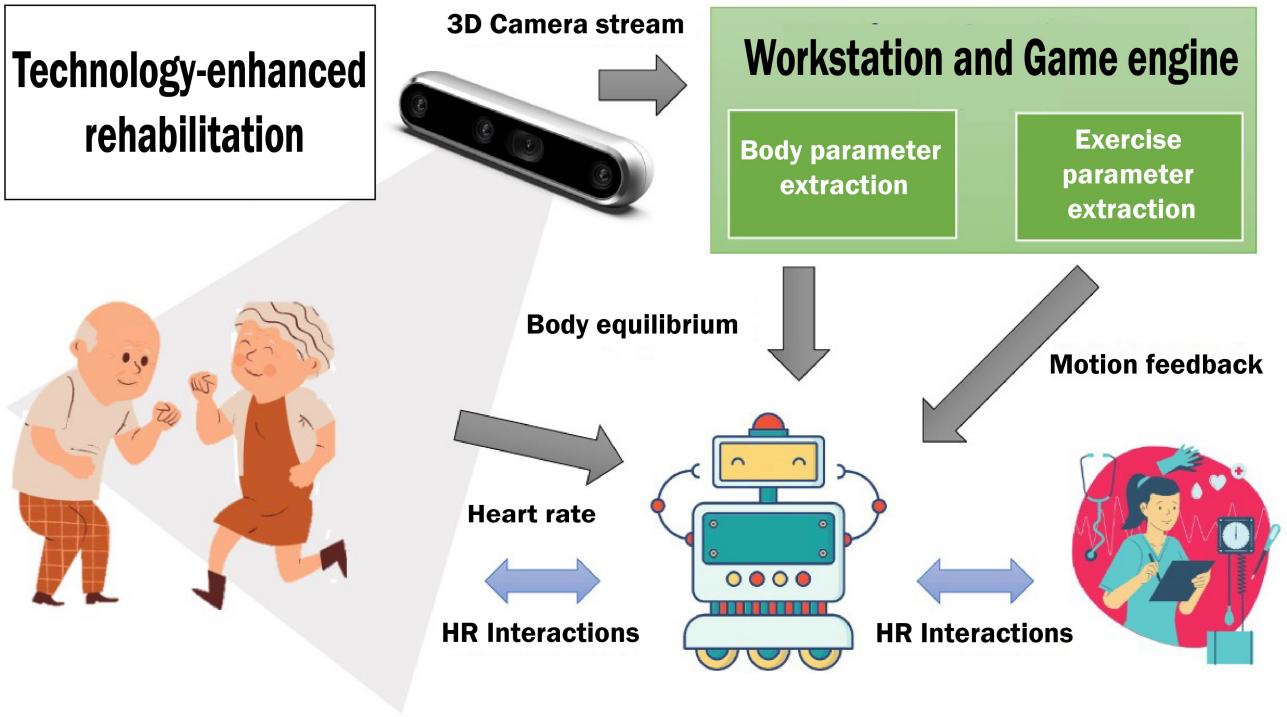
As shown in Umbrico et al [25] SI-Robotics proposes AP to support the synthesis of physical rehabilitation programs for patients with PD . A key aspect of the developed planning framework is the combined reasoning about spatial and clinical effects of stimuli (ie, motions or dancing steps) needed to synthesize plans that are technically valid and effective from a clinical point of view. Compared with a manual definition of the rehabilitation session, the use of a planner aims to improve the quality, accuracy, and engagement of the resulting rehabilitation programs. A therapist provides a planner with data about the rehabilitation session (song time, song rhythm, and difficulty level) and the clinical objectives. Given this input, the planner decides a sequence of steps, optimized according to an objective function encoding the specified clinical objective. Planned steps are thus chosen according to the rehabilitation needs of the participating patient (personalization).

In addition, several devices enrich the session to extract useful data about the health conditions and the performance of patients.

XSL•FO
RenderX

The generated data streams are gathered by the robot that embodies the AI-based reasoning modules. Overall, three types of perception components were considered: (1) wearable sensing devices (eg, sensorized shirts) that constantly produce data streams about physiological parameters of patients (eg, heart rate) [32]; (2) a video-processing component elaborating 3D camera data to extract kinematics about motions and produce data about body posture and body equilibrium of patients and; (3) a video-processing component elaborating 3D camera data to analyze patients' behaviors and produce data about the correctness of performed motions (ie, feedback). Figure 2 below shows a conceptual overview of the designed session pointing out a subset of the main hardware and software components and the data streams generated during the execution of a rehabilitation session.

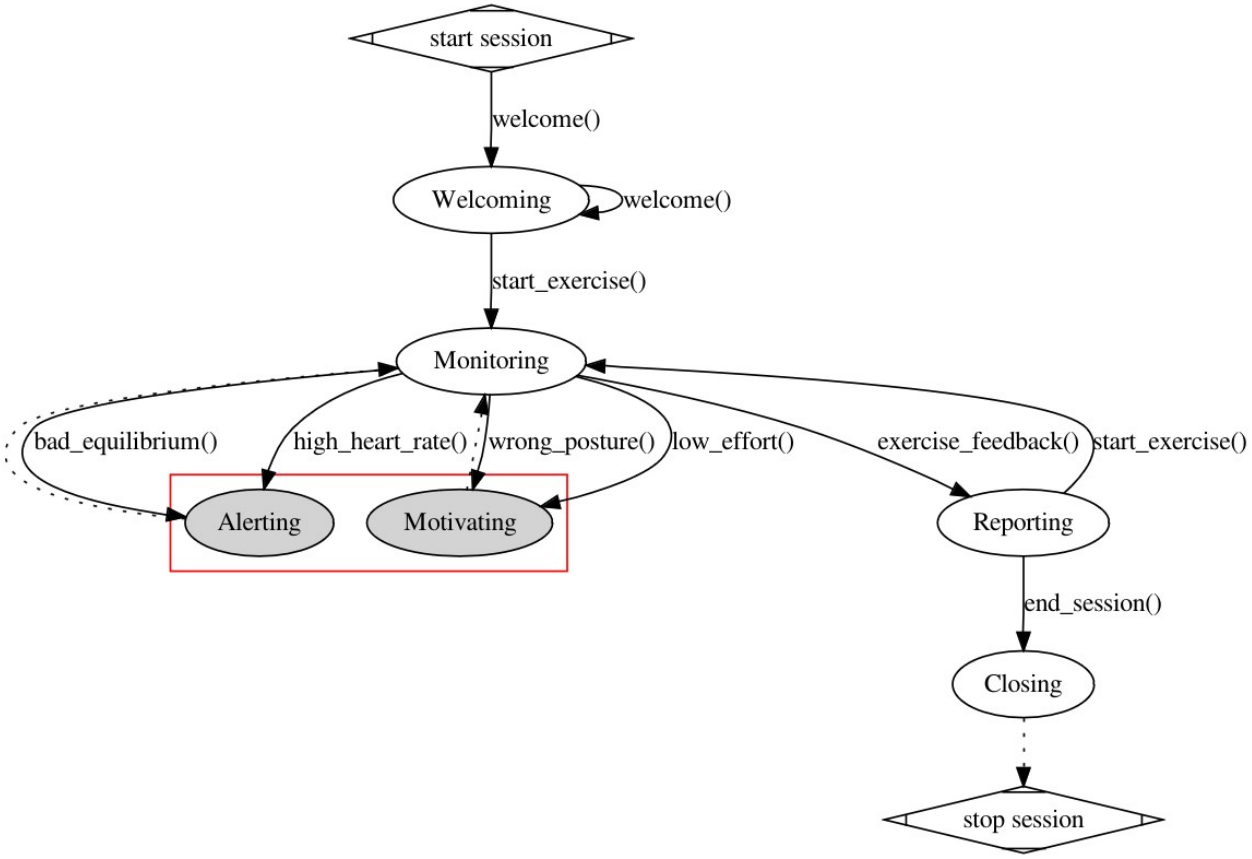
Figure 2. Automata describing the assistive behavior of the robot during a rehabilitation session. HR: Human-Robot



In addition, a socially interacting robot should support the execution of a session by recognizing the current state of a session according to data gathered from the environment, autonomously detecting events that would trigger different phases, and setting the behavioral goals needed to support a session. The robotic platform consists of a novel social robotic platform designed within SI-Robotics and developed by Co-Robotics [33]. Figure 3 shows the lifecycle determining the assistive behavior of the robot during each rehabilitation session.

XSL•FO
RenderX

Figure 3. Architectural structure of the artificial intelligence–based cognitive control of the robot.



A rehabilitation session starts with a welcoming state where the robot introduces patients to the session. Patients do not know the structure of the rehabilitation and are not familiar with the involved technology. The objective of this state is therefore to prepare patients for the therapy by explaining the structure of the session and the technology used. When all patients are ready to start a session, the robot enters into a "monitoring" state gathering data from the environment to monitor the performance and health state of patients. When a rehabilitation exercise ends the robot enters into the "reporting" state to interact with the therapist to comment on the exercise and the observed performance. This state is especially useful to enrich the subjective experience of the therapist by enhancing the awareness of the therapist with objective feedback on the quality of the exercises. It is indeed meant as a support for the decisions on how to proceed with the session (eg, change or repeat the exercises, difficulty of the next exercise, etc). If the therapist decides to start a new exercise, then the robot enters again into the monitoring state. Otherwise, the robot enters into the "closing" state to finalize the session. In this latter case, the robot interacts with patients to ask their personal feelings about the session and to show a qualitative assessment of their session.

Different types of interaction are foreseen during an exercise, depending on the type of situation recognized by the robot. Two types of interactions have been considered characterized by the motivating and the alerting states in the automata. The motivating phase allows the robot to play the role of a coach during the session stimulating patients to perform better or correcting their behavior. The robot is supposed to enter the motivating phase when it detects a heart rate below some minimum "training threshold" or when it detects a wrong posture of the body, for example, bad alignment or asymmetrical motion of lower or upper limbs.

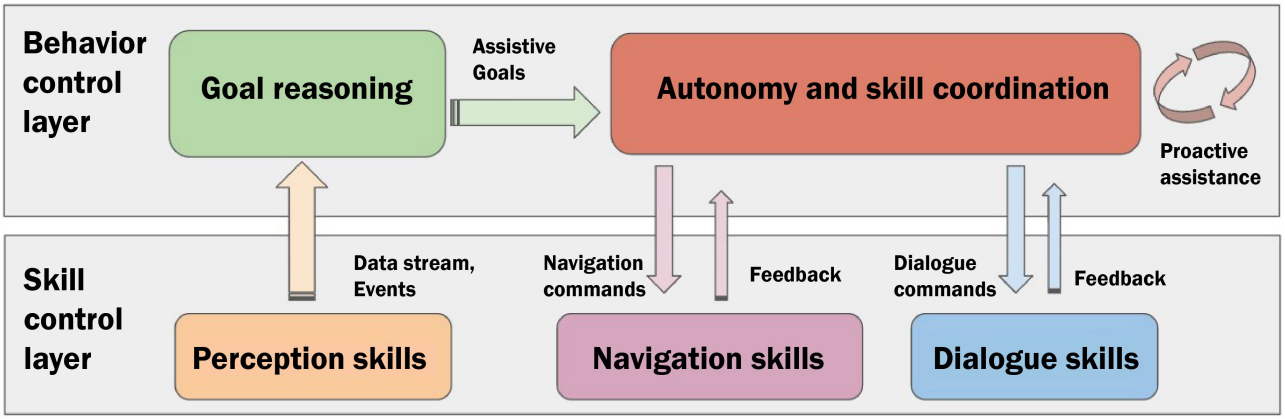
The alerting phase allows the therapist (and the robot) to intervene when some critical situation concerning the health state of a patient is detected during the session. In such case, the robot is supposed to warn the therapist about the danger and simultaneously interact with the involved patient, even interrupting the exercise if necessary. Specifically, the robot enters the alerting phase when it detects the risk of a fall or an anomalous heart rate for a patient. Both these situations are critical for the considered target of patients and need a prompt intervention.

AI-Based Control of Robotic Skills

A key aspect (and still an open challenge) in the design of autonomous robotic agents is the integration of different AI technologies and Robotics [34,35]. SI-Robotics proposes an AI-based control architecture with twofold objectives: (1) to support the abstraction and reasoning capabilities needed to recognize health-related situations and assistive objectives and (2) to coordinate robotic skills to "act" in the environment and (autonomously) achieve contextualized assistive objectives. Figure 4 shows an overview of the designed architecture with the main functional components involved in the control process. Two main architectural layers are considered and are organized according to a cognitive architecture inspired by the Dual Process Psychological Theory [36].

XSL•FO
RenderX

Figure 4. Designed architecture.



A skill control layer encapsulates the control components responsible for providing the robot with the "low-level" capabilities needed to perceive the environment and interact with the environment. The Navigation Skills component encapsulates functionalities that allow a robot to robustly navigate within the environment and autonomously reach known locations. The Dialog Skills encapsulate natural language processing functionalities that allow a robot to talk with patients and therapists. This component in particular enables an affective interaction [37] by adapting the way the robot talks with a patient to the detected emotional state and known personality traits.

While some level of perception is needed for both navigation and dialogue skills, the Perception Skills component emphasizes the additional capability needed to process data from deployed environmental devices. It encapsulates low-level data processing to provide the higher decisional level with refined information. A behavior control layer supports AI-based control features to contextualize observations, trigger assistive goals, and coordinate underlying skills to implement the desired assistance. The Autonomy and Skill Coordination component encapsulates an AI plan–based controller that supports behavior continuity coordinating robotic skills. The Goal Reasoning component instead is in charge of contextualizing observations to automatically set assistive goals. The combination of goal reasoning and a plan-based controller is the key architectural feature enabling proactive and autonomous goal-oriented behaviors of the robot [38].

The AI-based architectural elements composing the coordination layer of Figure 4 have been embedded into the robotic platform and integrated through the Robot Operating System (ROS) using the ROS Melodic distribution [39]. The coordination component responsible for the synthesis of the assistive behavior has been implemented in the shape of an AI plan–based controller relying on AP. The component uses the open-source framework ROXANNE [40], which implements goal-oriented acting capabilities into an ROS [41]. ROXANNE allows the robot to receive high-level assistive goals and (autonomously) synthesize and execute navigation and dialogue commands or tasks to realize a desired assistive behavior. The goal reasoning component correlates data received from the perception component with the autonomous behavior of the robot. It endows the robot with the cognitive capabilities needed to understand the evolution of a rehabilitation session autonomously select suitable planning goals and consequently decide the assistance needed in a particular context.

Embedded AP models the behavioral constraints needed to correctly support a rehabilitation session. Behavioral constraints are specified in the shape of temporal constraints and require interacting tasks performed through available skills. In this regard, Table 1 shows the modeled assistive goals and the parameters defined to contextualize interactions. Furthermore, the table briefly describes the behavioral constraints specified in the model to correctly implement the desired assistance and support the associated goal. It is worth noting that AI planning facilitates robot programming since it supports an "easy specification" of desired behavioral dynamics without hard-coding the implemented skills.

XSL•FO
RenderX

Table . Acting goals supported by the robot-embedded controller to synthesize assistive behaviors during a rehabilitation session.

Goal	Interaction parameters	Interaction parameters
Welcoming	• Therapist: string • Patient: string • Gender: string • Novel: boolean	The robot moves from its current location to the welcoming area to call the specified patient and take patient to the dressing area to wear the sensorized shirt. Then it guides the patient to the rehabilitation area to start the session. If the patient is new to the session (flag novel set to true) the robot explains the organization of the rehabilitation and the functioning of the devices. Otherwise, the robot shows data about the last session to motivate the patient for the current one.
Exercise start	—[a]	When an exercise begins the robot starts moving around the rehabilitation area and monitoring patients' state by collecting physiological, body posture, and performance data.
Exercise report	• Patient: string • Gender: string • Score: double • Past_score: double • Global_score: double	When an exercise ends the robot moves close to a patient to show a brief report about the patient performance. The robot shows the achieved score and (if available) compares it to the average score of the patient in the past and the average score of other patients. This would make a patient aware of the quality of the performance.
Low heart rate and high heart rate	• Therapist: string • Patient: string • Gender: string • Heart_rate: double	The robot periodically analyzes trends of gathered data about the heart rate of a patient. Thresholds are parametric and computed according to the age of the patient as follows: HR_max=80% (220 – age) HR_min=40% h_max The objective of the session is to keep the heart rate of patients within HR_max and HR_min to stimulate a proper metabolic effort. In the case of low heart rate (i.e., value below HR_min) the robot implements a motivational behavior asking a patient to improve the performance. In case of high heart rate (i.e., value above HR_max), the root implements an assistive behavior alerting the therapist and asking the patient to reduce the performance
Bad posture	• Patient: string • Gender: string • Tilt: integer in [0, 1] • Arms: integer in [–1, 1] • Legs: integer in [–1, 1]	The robot constantly monitors the motions and the inclination of patients. If anomalous kinematics parameters are detected, for example, wrong inclinations of the body or wrong alignment of arms or legs, the robot implements a motivational behavior to correct the body posture of a patient and thus stimulate correct motions.
Bad equilibrium	• Therapist: string • Patient: string • Gender: string • Lateral: integer in [–1,1] • Frontal: integer in [–1,1]	In addition to the body posture, the robot constantly monitors the equilibrium of patients to prevent dangerous falls during physical therapy. If dangerous equilibrium conditions are detected from camera data (eg, the center of mass is outside the patient's base) the robot implements an assistive behavior to promptly notify the therapist and warn the patient about the danger or falling.

[a]Not applicable.

Outcomes

All assessment procedures adhere to a standardized protocol. Specifically, the primary focus of the study revolves around enhancing balance, and gait, and alleviating the fear of falling among elderly patients with PD. This is gauged through the utilization of the 3 POMA scales (POMA balance, POMA Gait, and POMA Total) after the 10 treatment sessions, as a result of the Irish dance intervention. In addition, secondary measures include evaluating the gait speed of elderly patients with PD, their fear of falling (assessed via FES-I), their physical

XSL•FO
RenderX

performance (SPPB), their autonomy in daily living activities (UPDRS-III), and their overall physical and psychological well-being (SF-12) together with the evaluation of the Unified Theory of Acceptance and Use of Technology (UTAUT) that is composed by 10 subscales (anxiety; attitude; facilitating conditions; intention to use; perceived adaptability; perceived enjoyment; perceived ease of use; perceived usefulness; social influence; and trust).

Statistical Analysis

Continuous variables were presented as either mean and SD or median and IQR, depending on their distribution, which was determined using the Kolmogorov-Smirnov test. Categorical variables were expressed as absolute numbers and percentages. To test statistically significant differences ($P<.050$) between pre- and postconditions, Wilcoxon signed-rank test (for non-normal distribution) or paired test (for normal distribution) were used, alongside simple descriptive statistics such as means, medians, and SDs as appropriate. The statistical analysis was performed using SPSS software.

Ethical Considerations

The study was approved by the Ethics Committee of the Istituto Nazionale Ricovero e Cura per Anziani, (IRCCS INRCA) on June 17, 2021 (CE INRCA 21004). The protocol is registered on ClinicalTrials.gov with trial registration number NCT05005208 (October 23 , 2023). All participants signed the informed consent and data processing consent. The data are anonymised so that the identity of the subject cannot be traced. No compensation is provided for participation in the study.

Results

Demographic and clinical data of the patients are reported in Table 2. Two participants dropped out because they did not complete the rehabilitation program.

The CONSORT (Consolidated Standards of Reporting Trials) flowchart is shown in Figure 5.

Gender differences were not statistically significantly different in all scales used to select the sample.

Table . Baseline demographic and clinical profile.

		Total (n=18)	Male (n=10)	Female (n=8)	*P* value
Age (years), mean (SD)		75.3 (5.5)	73.7 (5.4)	77.2 (5.2)	.18
Marital status, n (%)					.48
	Married	15 (83.3)	9 (90)	6 (75)	
	Single	1 (5.5)	0 (0)	1 (12.5)	
	Widowed	2 (11.2)	1 (10)	1 (12.5)	
Educational level, n (%)					.17
	Primary education	4 (22.2)	1 (10)	3 (37.5)	
	Secondary education	10 (55.6)	6 (60)	4 (50)	
	University or more	4 (22.2)	3 (30)	1 (12.5)	
Hoehn and Yahr score, mean (SD)		1.8 (0.3)	1.9 (0.3)	1.7 (0.4)	.42
Rankin Scale score, mean (SD)		1.6 (0.9)	1.4 (0.8)	1.8 (1.1)	.33
GDS[a], mean (SD)		3.0 (0.7)	2.9 (0.7)	3.2 (0.7)	.33
FAC[b], mean (SD)		4.6 (0.6)	4.7 (0.4)	4.5 (0.74)	.50
MMSE[c], mean (SD)		29.4 (0.9)	29.7 (0.6)	29.1 (0.7)	.22
Barthel Index score, mean (SD)		95.5 (5.6)	94.5 (6.8)	96.8 (3.7)	.39

[a]GDS: Geriatric Depression Scale.

[b]FAC: Functional Ambulation Category.

[c]MMSE: Mini-Mental State Examination.

XSL•FO
RenderX

Figure 5. The CONSORT (Consolidated Standards of Reporting Trials) flowchart.

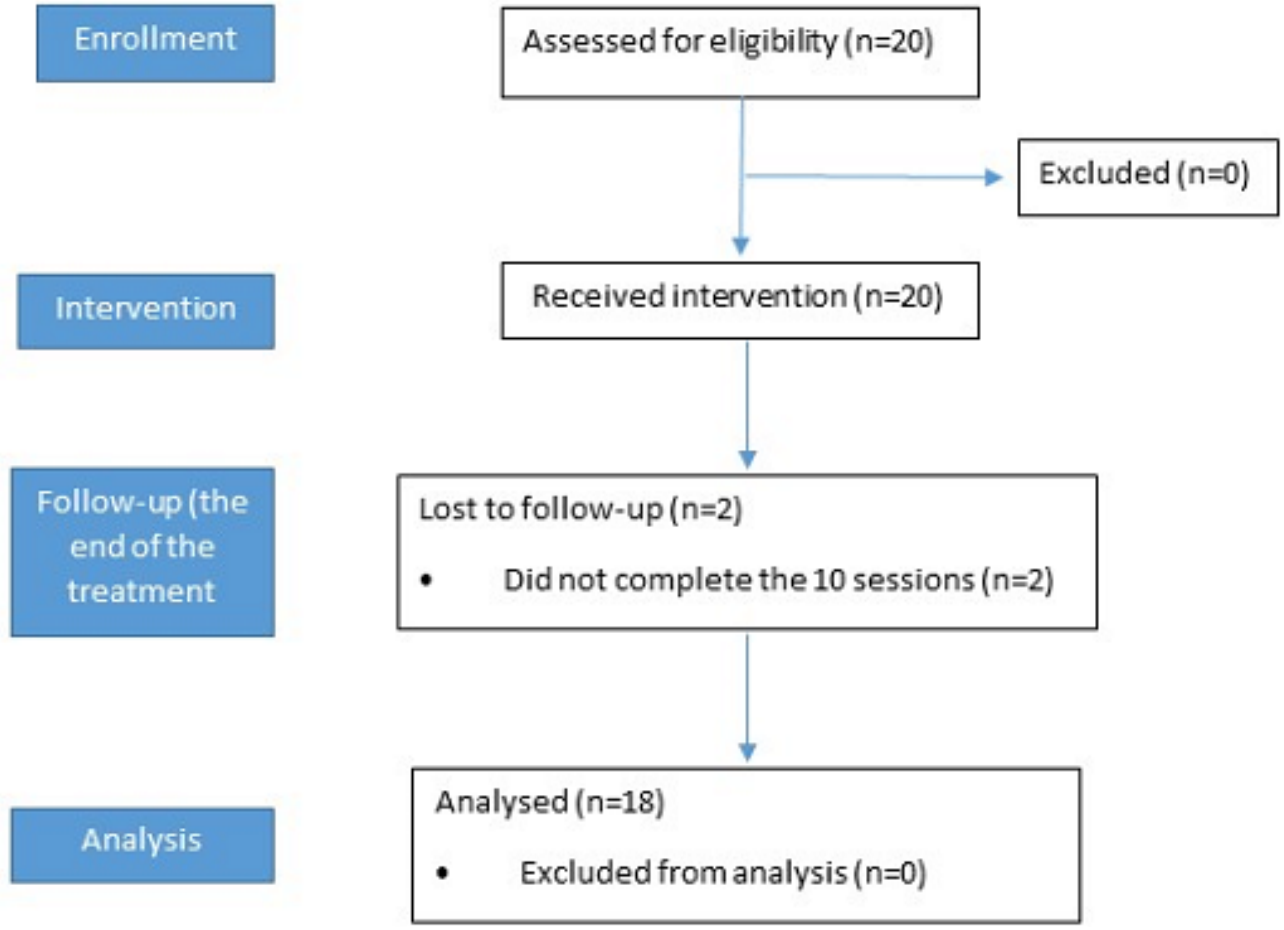


Table 3 shows differences between pre- and postintervention on the functional state scales with the UPDRS-III and the SPPB, gait and balance performance on Tinetti POMA Gait and POMA Balance, evaluation of quality of life with SF-12 and its subscores (physical component score and mental component score), fear of falling (FES-I) together with the TUG execution time, the meters covered during the 6-Minute Walking Test and the gait speed.

Statistical analysis revealed a significant effect on UPDRS-III and TUG together with improvements in the POMA scale, both on balance and gait aspects, and in SPPB. However, for the POMA and SPPB scales, there was a low Cohen *d* coefficient, underlining a probable effect of the low sample size on these results. In contrast, the effect size for UPDRS-III and TUG was high (>0.70), highlighting the usefulness of the rehabilitation intervention on the motor symptoms of PD also in our small sample.

XSL•FO
RenderX

Table . Mean (SD) values of the mean of pre- and postintervention scores on the UPDRS-III[a], SPPB[b], POMA[c] (total, gait and balance), SF-12[d] (total, physical, and mental component score), FES-I[e], TUG[f] execution time, 6MWT[g], and gait speed. Pre-post and between-group comparisons are reported for each score (P<.05). The effect size (Cohen d) is also reported.

		Preintervention scores, mean (SD)	Postintervention scores, mean (SD)	P value	Cohen d
UPDRS-III		13.56 (1.85)	13.83 (1.85)	.01[h]	−0.73
SPPB		8.83 (2.61)	10.39 (2.45)	.02[h]	0.10
POMA					
	POMA Total	24.11 (3.92)	26.28 (2.74)	.001[h]	−0.07
	POMA Gait	9.83 (2.30)	10.94 (1.55)	.01[h]	−0.02
	POMA Balance	14.28 (2.16)	15.33 (1.37)	.01[h]	−0.10
SF-12					
	SF-12-Tot[i]	31.28 (1.99)	32.33 (2.42)	.10	−0.20
	PCS-12[j]	13.56 (1.84)	13.83 (1.85)	.49	−0.23
	MCS-12[k]	17.72 (1.22)	18.50 (1.91)	.12	−0.16
FES-I		10.78 (1.92)	10.67 (2.22)	.83	−0.29
TUG (s)		11.0 (3.91)	9.19 (3.14)	.002[h]	−0.82
6MWT (m)		361.50 (99.33)	373.72 (98.67)	.40	−0.19

[a]UPDRS-III: Unified Parkinson Disease Rating Scale – III.

[b]SPPB: Short Physical Performance Battery.

[c]POMA: Performance Oriented Mobility Assessment.

[d]SF-12: 12-Item Short-Form Health Survey.

[e]FES-I: Falls Efficacy Scale–International.

[f]TUG: Timed Up and Go test.

[g]6MWT: 6-Minute Walking test.

[h]P values from matched-pairs Student t test.

[i]SF-12-Tot: 12-Item Short-Form Health Survey total score.

[j]PCS-12: 12-Item Short-Form Health Survey physical component score

[k]MCS-12: 12-Item Short-Form Health Survey mental component score.

Table 4 reports the scores of the subscales of UTAUT.

The analysis of the UTAUT subscales provided valuable insights into users' perceptions and interactions with the system. In particular, attitude, perceived adaptability, enjoyment, ease of use, usefulness, and social influence were generally positive.

XSL•FO
RenderX

Table . Mean (SD) of the mean of the Unified Theory of Acceptance and Use of Technology (UTAUT) subscales scores.

UTAUT subscales		Score, mean (SD)	Range
Anxiety (ANX)	Evoking anxious or emotional reactions when using the system.	7.4 (4.2)	0 - 20
Attitude (ATT)	Positive or negative feelings about the appliance of the technology	11.2 (2.5)	0 - 15
Facilitating conditions (FC)	Objective factors in the environment that facilitate using the system	2.6 (1.7)	0 - 15
Intention to use (ITU)	The outspoken intention to use the system over a longer period in time	3 (0)	0 - 10
Perceived adaptability (PAD)	The perceived ability of the system to be adaptive to the changing needs of the user.	10.5 (3.0)	0 - 15
Perceived enjoyment (PENJ)	Feelings of joy or pleasure associated by the user with the use of the system.	18.2 (2.2)	0 - 25
Perceived ease of use (PEOU)	The degree to which the user believes that using the system would be free of effort	18.1 (3.8)	0 - 25
Perceived usefulness (PU)	The perceived ability of the system to perform sociable behavior.	11.4 (2.9)	0 - 15
Social influence (SI)	The user's perception of how people who are important to him think about him using the system	8 (2.4)	0 - 10
Trust	The belief that the system performs with personal integrity and reliability.	7.5 (1.8)	0 - 10

Discussion

Principal Findings

This study aims to investigate the usefulness of the SI-Robotics intervention based on AI robotic assistance and exergame stimulation, in the context of PD. This study aims to assess the ability of a dance-based physical therapy program enriched by AI-based exergames and contextual robotic assistance in improving motor function, balance, gait, and quality of life in patients with PD. The findings of this study demonstrate the potential usefulness of a dance-based physical therapy program using a robotic platform to deliver exergames in improving motor function, balance, gait, and quality of life in patients with PD. Results suggest that integrating dance therapy with engaging technology may lead to greater improvements in PD symptoms compared with conventional physical therapy alone. Although the statistical analysis revealed significant improvements in motor function, balance, and gait (eg, UPDRS-III, POMA, and TUG scores), the effect sizes for several outcome measures were relatively small (Cohen d<0.2 for POMA and SPPB scales). This suggests that, while the observed changes were statistically significant, the practical impact of the intervention on certain motor and functional outcomes may be limited. These small effect sizes could be partly attributed to the small sample size, which may have reduced the power to detect larger effects.

This study adds to the growing body of research supporting the use of dance therapy and exergaming as adjunctive treatments for PD. Previous studies have shown that both dance therapy and exergaming can independently improve motor function, balance, and mobility in patients with PD [4,7]. By combining these approaches, our intervention aimed to target the multifaceted nature of PD symptoms, including motor, cognitive, and psychosocial aspects.

The results are consistent with previous findings demonstrating the benefits of dance therapy and exergaming in PD management. For example, Duncan and Earheart [4] reported significant improvements in motor function and balance among patients with PD participating in community-based dancing, while Barry et al [7] found that exergaming interventions were effective in improving mobility and reducing fall risk in patients with PD. These studies, along with ours, highlight the potential of integrating dance therapy with exergaming to address the complex needs of patients with PD.

One novel aspect of our intervention is the use of the SI-Robotics system, a robotics-based platform designed to engage patients with PD in Irish set dancing. This innovative approach combines the physical and cognitive benefits of dance therapy with the interactive and adaptable nature of exergaming. By incorporating personalized avatars and choreography, the "Let's Dance" game provided a stimulating and enjoyable exercise experience for participants, potentially enhancing motivation and adherence to the intervention.

Regarding the analysis of technology acceptance among the population studied, the results indicate a generally positive perception and interaction with technology. In particular, users displayed a positive attitude towards the technology, indicating

XSL•FO
RenderX

favorable feelings toward it. They also found the system adaptable, enjoyable, easy to use and useful showing that users experienced pleasure while using the system. Finally, trust in the system was moderately high, indicating that users believed it to be reliable and trustworthy. These highlight how the system successfully meets user expectations by offering an engaging, straightforward, and beneficial experience.

Our study focused on elderly patients with PD with mild to moderate disease severity, as reflected by their scores on the Hoen and Yahr scale, Functional Ambulation Category, Rankin Scale, Geriatric Depression Scale, and Mini-Mental State Examination. These inclusion criteria aimed to ensure that participants were physically and cognitively capable of engaging in the intervention safely and effectively. Future research could explore the applicability of our intervention to a broader range of patients with PD , including those with more advanced disease stages.

Limitations of our study include the small sample size and lack of a control group for comparison. While our results provide preliminary evidence of the usefulness of the intervention, larger-scale randomized controlled trials are needed to confirm these findings and establish the optimal dosage and timing of dance-based physical therapy with exergames for PD management. In addition, longer-term follow-up assessments could help evaluate the sustainability of the intervention effects over time.

Conclusions

In conclusion, this study highlights the potential of integrating dance therapy with exergaming integrated with an interactive robotic platform as a novel approach to improving motor function, balance, gait, and quality of life in patients with PD. The SI-Robotics intervention offers a promising avenue for future research and clinical practice in the management of PD symptoms. To strengthen the validity of the current findings and ensure their broader applicability, future research will focus on replicating the study with a larger sample. This will help to verify the generalizability of the results providing a more comprehensive understanding of the observed effects. Moreover, further studies are warranted to explore the long-term effects and feasibility of implementing this intervention in diverse clinical settings.

Acknowledgments

The study has received funding under the project Social Robotics for Active and Healthy Ageing (SI-Robotics (ARS01_01120), funded by the Italian Ministry of Education, Universities and Research, under the National Operational Program Area, Technologies for Living Environments.

Conflicts of Interest

None declared.

References

1. Postuma RB, Berg D, Stern M, et al. MDS clinical diagnostic criteria for Parkinson's disease. Mov Disord 2015 Oct;30(12):1591-1601. [doi: 10.1002/mds.26424] [Medline: 26474316]
2. Olanow CW, Stern MB, Sethi K. The scientific and clinical basis for the treatment of Parkinson disease (2009). Neurology (ECronicon) 2009 May 26;72(21 Suppl 4):S1-136. [doi: 10.1212/WNL.0b013e3181a1d44c] [Medline: 19470958]
3. Tomlinson CL, Patel S, Meek C, et al. Physiotherapy versus placebo or no intervention in Parkinson's disease. Cochrane Database Syst Rev 2014;2014(6):CD002815. [doi: 10.1002/14651858.CD002815]
4. Duncan RP, Earhart GM. Randomized controlled trial of community-based dancing to modify disease progression in Parkinson disease. Neurorehabil Neural Repair 2012 Feb;26(2):132-143. [doi: 10.1177/1545968311421614] [Medline: 21959675]
5. Hackney ME, Earhart GM. Effects of dance on movement control in Parkinson's disease: a comparison of Argentine tango and American ballroom. J Rehabil Med 2010;42(7):662-669. [doi: 10.2340/16501977-0362]
6. Mirelman A, Rochester L, Reelick M, et al. V-TIME: a treadmill training program augmented by virtual reality to decrease fall risk in older adults: study design of a randomized controlled trial. BMC Neurol 2011;11(1):115. [doi: 10.1186/1471-2377-13-15]
7. Barry G, Galna B, Rochester L, Godfrey A. Delivery of community-based exercise programs for people with Parkinson's disease: a mixed-methods systematic review. J Neurol 2014;261(7):1321-1334.
8. Bevilacqua R, Benadduci M, Bonfigli AR, et al. Dancing with Parkinson's disease: the SI-ROBOTICS study protocol. Front Public Health 2021;9:780098. [doi: 10.3389/fpubh.2021.780098] [Medline: 34993171]
9. Cortellessa G, De Benedictis R, Fracasso F, Orlandini A, Umbrico A, Cesta A. AI and robotics to help older adults: revisiting projects in search of lessons learned. Paladyn J Behav Robot 2021 Feb 10;12(1):356-378. [doi: 10.1515/pjbr-2021-0025]
10. Carapellotti AM, Stevenson R, Doumas M. The efficacy of dance for improving motor impairments, non-motor symptoms, and quality of life in Parkinson's disease: a systematic review and meta-analysis. PLoS ONE 2020;15(8):e0236820. [doi: 10.1371/journal.pone.0236820] [Medline: 32756578]

XSL•FO
RenderX

11. Volpe D, Signorini M, Marchetto A, Lynch T, Morris ME. A comparison of Irish set dancing and exercises for people with Parkinson's disease: a phase II feasibility study. BMC Geriatr 2013 Jun 4;13:54. [doi: 10.1186/1471-2318-13-54] [Medline: 23731986]
12. Holden MK, Gill KM, Magliozzi MR. Gait assessment for neurologically impaired patients. Standards for outcome assessment. Phys Ther 1986 Oct;66(10):1530-1539. [doi: 10.1093/ptj/66.10.1530] [Medline: 3763704]
13. Holden MK, Gill KM, Magliozzi MR. Gait assessment for neurologically impaired patients. Phys Ther 1986 Oct 1;66(10):1530-1539. [doi: 10.1093/ptj/66.10.1530]
14. van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. Stroke 1988 May;19(5):604-607. [doi: 10.1161/01.str.19.5.604] [Medline: 3363593]
15. Rinaldi P, Mecocci P, Benedetti C, et al. Validation of the five-item geriatric depression scale in elderly subjects in three different settings. J Am Geriatr Soc 2003 May;51(5):694-698. [doi: 10.1034/j.1600-0579.2003.00216.x] [Medline: 12752847]
16. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
17. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. Md State Med J 1965 Feb;14:61-65. [Medline: 14258950]
18. Tinetti ME. Performance-oriented assessment of mobility problems in elderly patients. J Am Geriatr Soc 1986 Feb;34(2):119-126. [doi: 10.1111/j.1532-5415.1986.tb05480.x] [Medline: 3944402]
19. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994 Mar;49(2):M85-M94. [doi: 10.1093/geronj/49.2.m85] [Medline: 8126356]
20. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991 Feb;39(2):142-148. [doi: 10.1111/j.1532-5415.1991.tb01616.x] [Medline: 1991946]
21. Balke B. A simple field test for the assessment of physical fitness. Rep 63-6. Rep Civ Aeromed Res Inst. 1963. URL: https://www.faa.gov/data_research/research/med_humanfacs/oamtechreports/1960s/media/am63-06.pdf
22. Ware J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996 Mar;34(3):220-233. [doi: 10.1097/00005650-199603000-00003] [Medline: 8628042]
23. Ruggiero C, Mariani T, Gugliotta R, et al. Validation of the Italian version of the falls efficacy scale international (FES-I) and the SHORT FES-I in community-dwelling older persons. Arch Gerontol Geriatr 2009;49 Suppl 1:211-219. [doi: 10.1016/j.archger.2009.09.031] [Medline: 19836635]
24. Goetz CG, Fahn S, Martinez-Martin P, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): process, format, and clinimetric testing plan. Mov Disord 2007 Jan;22(1):41-47. [doi: 10.1002/mds.21198] [Medline: 17115387]
25. Umbrico A, Benadduci M, Bevilacqua R, et al. Combining clinical and spatial constraints into temporal planning to personalize physical rehabilitation. ICAPS 2023;33(1):532-540. [doi: 10.1609/icaps.v33i1.27234]
26. Belić M, Bobić V, Badža M, Šolaja N, Đurić-Jovičić M, Kostić VS. Artificial intelligence for assisting diagnostics and assessment of Parkinson's disease—a review. Clin Neurol Neurosurg 2019 Sep;184:105442. [doi: 10.1016/j.clineuro.2019.105442] [Medline: 31351213]
27. Victorino JN, Shibata Y, Inoue S, Shibata T. Predicting wearing-off of Parkinson's disease patients using a wrist-worn fitness tracker and a smartphone: a case study. Appl Sci (Basel) 2021;11(16):7354. [doi: 10.3390/app11167354]
28. Devarajan JP, Sreedharan VR, Narayanamurthy G. Decision making in health care diagnosis: evidence from Parkinson's Disease via hybrid machine learning. IEEE Trans Eng Manage 2023;70(8):2719-2731. [doi: 10.1109/TEM.2021.3096862]
29. Parisi L, RaviChandran N, Manaog ML. Feature-driven machine learning to improve early diagnosis of Parkinson's disease. Expert Syst Appl 2018 Nov;110:182-190. [doi: 10.1016/j.eswa.2018.06.003]
30. González JC, Pulido JC, Fernández F. A three-layer planning architecture for the autonomous control of rehabilitation therapies based on social robots. Cogn Syst Res 2017 Jun;43:232-249. [doi: 10.1016/j.cogsys.2016.09.003]
31. Baschieri D, Gaspari M, Zini F. A planning-based serious game for cognitive rehabilitation in multiple sclerosis. Presented at: Proceedings of the 4th EAI International Conference on Smart Objects and Technologies for Social Good; Nov 28, 2018; Bologna Italy. [doi: 10.1145/3284869.3284916]
32. Leone A, Rescio G, Diraco G, Manni A, Siciliano P, Caroppo A. Ambient and wearable sensor technologies for energy expenditure quantification of ageing adults. Sensors (Basel) 2022 Jun 29;22(13):4893. [doi: 10.3390/s22134893] [Medline: 35808387]
33. Co-Robotics. URL: http://www.corobotics.eu [accessed 2024-12-23]
34. Rajan K, Saffiotti A. Towards a science of integrated AI and robotics. Artif Intell 2017 Jun;247:1-9. [doi: 10.1016/j.artint.2017.03.003]
35. Ingrand F, Ghallab M. Deliberation for autonomous robots: a survey. Artif Intell 2017 Jun;247:10-44. [doi: 10.1016/j.artint.2014.11.003]
36. Umbrico A, De Benedictis R, Fracasso F, Cesta A, Orlandini A, Cortellessa G. A mind-inspired architecture for adaptive HRI. Int J Soc Robot 2023;15(3):371-391. [doi: 10.1007/s12369-022-00897-8] [Medline: 35910297]
37. Beale R, Creed C. Affective interaction: how emotional agents affect users. Int J Hum Comput Stud 2009 Sep;67(9):755-776. [doi: 10.1016/j.ijhcs.2009.05.001]

XSL•FO
RenderX

38. Umbrico A, Cesta A, Cortellessa G, Orlandini A. A holistic approach to behavior adaptation for socially assistive robots. Int J Soc Robotics 2020 Jul;12(3):617-637. [doi: 10.1007/s12369-019-00617-9]
39. ROS melodic morenia. ROS.org. URL: http://wiki.ros.org/melodic [accessed 2024-12-23]
40. Pstlab/roxanne_rosjava. GitHub. URL: https://github.com/pstlab/roxanne_rosjava.git [accessed 2024-12-02]
41. Umbrico A, Cesta A, Orlandini A. Human-aware goal-oriented autonomy through ROS-integrated timeline-based planning and execution. Presented at: The 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN); Aug 28-31, 2023; Busan, Republic of Korea. [doi: 10.1109/RO-MAN57019.2023.10309516]

Abbreviations

AI: artificial intelligence
BI: Barthel Index
FES-I: Falls Efficacy Scale-International
PD: Parkinson disease
POMA: Tinetti's Performance-Oriented Mobility Assessment
ROS: Robot Operating System
SF-12: 12-Item Short-Form Health Survey
SI-Robotics: Social Robotics for Active and Healthy Ageing
SPPB: Short Physical Performance Battery
TUG: Timed Up and Go
UPDRS-III : Unified Parkinson Disease Rating Scale – III

Edited by PH Chau; submitted 05.06.24; peer-reviewed by D Marks, F Noveletto, G Barresi; revised version received 05.09.24; accepted 09.09.24; published 14.01.25.

Please cite as:
Bevilacqua R, Maranesi E, Benadduci M, Cortellessa G, Umbrico A, Fracasso F, Melone G, Margaritini A, La Forgia A, Di Bitonto P, Potenza A, Fiorini L, La Viola C, Cavallo F, Leone A, Caroppo A, Rescio G, Marzorati M, Cesta A, Pelliccioni G, Riccardi GR, Rossi L
Exploring Dance as a Therapeutic Approach for Parkinson Disease Through the Social Robotics for Active and Healthy Ageing (SI-Robotics): Results From a Technical Feasibility Study
JMIR Aging 2025;8:e62930
URL: https://aging.jmir.org/2025/1/e62930
doi:10.2196/62930

© Roberta Bevilacqua, Elvira Maranesi, Marco Benadduci, Gabriella Cortellessa, Alessandro Umbrico, Francesca Fracasso, Giovanni Melone, Arianna Margaritini, Angela La Forgia, Pierpaolo Di Bitonto, Ada Potenza, Laura Fiorini, Carlo La Viola, Filippo Cavallo, Alessandro Leone, Andrea Caroppo, Gabriele Rescio, Mauro Marzorati, Amedeo Cesta, Giuseppe Pelliccioni, Giovanni Renato Riccardi, Lorena Rossi. Originally published in JMIR Aging (https://aging.jmir.org), 14.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Development and Validation of a Predictive Model Based on Serum Silent Information Regulator 6 Levels in Chinese Older Adult Patients: Cross-Sectional Descriptive Study

Yuzi You[1], MD; Wei Liang[2*], PhD; Yajie Zhao[2*], PhD

[1]Department of General Practice, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[2]Department of Geriatrics, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, 197 Ruijin Er Road, Shanghai, China

[*]these authors contributed equally

Corresponding Author:
Wei Liang, PhD
Department of Geriatrics, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, 197 Ruijin Er Road, Shanghai, China

Abstract

Background: Serum levels of silent information regulator 6 (SIRT6), a key biomarker of aging, were identified as a predictor of coronary artery disease (CAD), but whether SIRT6 can distinguish severity of coronary artery lesions in older adult patients is unknown.

Objectives: This study developed a nomogram to demonstrate the functionality of SIRT6 in assessing severity of coronary artery atherosclerosis.

Methods: Patients aged 60 years and older with angina pectoris were screened for this single-center clinical study between October 1, 2022, and March 31, 2023. Serum specimens of eligible patients were collected for SIRT6 detection by enzyme-linked immunosorbent assay. Clinical data and putative predictors, including 29 physiological characteristics, biochemical parameters, carotid artery ultrasonographic results, and complete coronary angiography findings, were evaluated, with CAD diagnosis as the primary outcome. The nomogram was derived from the Extreme Gradient Boosting (XGBoost) model, with logistic regression for variable selection. Model performance was assessed by examining discrimination, calibration, and clinical use separately. A 10-fold cross-validation technique was used to compare all models. The models' performance was further evaluated on the internal validation set to ensure that the obtained results were not due to overoptimization.

Results: Eligible patients (n=222) were divided into 2 cohorts: the development cohort (n=178) and the validation cohort (n=44). Serum SIRT6 levels were identified as both an independent risk factor and a predictor for CAD in older adults. The area under the receiver operating characteristic curve (AUROC) was 0.725 (95% CI 0.653 - 0.797). The optimal cutoff value of SIRT6 for predicting CAD was 546.384 pg/mL. Predictors included in this nomogram were serum SIRT6 levels, triglyceride glucose (TyG) index, and apolipoprotein B. The model achieved an AUROC of 0.956 (95% CI 0.928 - 0.983) in the development cohort. Similarly, in the internal validation cohort, the AUROC was 0.913 (95% CI 0.828 - 0.999). All models demonstrated satisfactory calibration, with predicted outcomes closely aligning with actual results.

Conclusions: SIRT6 shows promise in predicting CAD, with enhanced predictive abilities when combined with the TyG index. In clinical settings, monitoring fluctuations in SIRT6 and TyG may offer valuable insights for early CAD detection. The nomogram for CAD outcome prediction in older adult patients with angina pectoris may aid in clinical trial design and personalized clinical decision-making, particularly in institutions where SIRT6 is being explored as a biomarker for aging or cardiovascular health.

(JMIR Aging 2025;8:e64374) doi:10.2196/64374

KEYWORDS

aging; coronary artery disease; nomogram; SIRT6; TyG index; silent information regulator 6; triglyceride glucose index

Introduction

The global burden of ischemic heart disease from 1990 to 2019 reached staggering figures, with an estimated 9.14 million deaths attributed to coronary artery disease (CAD) in 2019 alone, affecting approximately 197 million individuals worldwide [1]. China witnessed a notable surge in CAD-associated mortality during this period, accounting for 38.2% of the global rise. By 2017, CAD-related deaths in China had skyrocketed by 1.12 million, representing an astonishing 184.1% increase compared with that of 1990 [2,3]. As of 2019, China was the global leader in CAD-related deaths, with 1.87 million reported cases [4]. By 2029, a 64% increase in CAD cases is expected in China

XSL•FO
RenderX

compared with 2020. As a dynamically evolving condition, CAD significantly contributes to the development of major adverse cardiovascular events, such as myocardial infarction, stroke, and cardiovascular mortality [5].

Aging, an independent risk factor for CAD, is characterized by a progressive decline in coronary artery and microvascular function, resulting in altered myocardial perfusion and increased myocardial injury [6]. CAD predominantly affects middle-aged and older adult populations, particularly people aged 60 years and older, with a prevalence exceeding 70%. Notably, the older adult population in China, aged 65 years and older, reached 220 million in 2023, constituting 15.4% of the nation's population and approximately 26.8% of the global older adult population [7]. The aging demographic in China is expanding rapidly, with nearly 40 million individuals aged 80 years and older in 2022, accounting for approximately 2.7% of the total population of China. Hence, China is predicted to enter a super-aged society by 2030, with the older adult population comprising more than 20% of the total population [8]. By 2084, it is estimated that half of China's population will be older adults. A higher mortality rate has been observed among Chinese older adults with CAD, with the majority of deaths occurring in people aged 75 years and older [9]. Despite these trends, research on CAD has often overlooked the specific needs of the older adult population. In China, approximately 300 million people suffer from chronic diseases, with half of them being aged 65 years or older. The older adult population often presents with atypical clinical symptoms, multiple comorbidities, and prolonged use of medications, all of which amplify the risk of cardiovascular disease [10,11]. Previous studies have indicated that therapeutic interventions in the early stages of CAD can reduce its incidence and improve prognosis [12,13], underscoring the necessity for noninvasive methods for identifying predictive factors for CAD.

Silent information regulator 6 (SIRT6), a member of the nicotinamide adenine dinucleotide–dependent histone deacetylase family, plays a crucial role in aging and aging-related diseases by maintaining telomerase stability and metabolic homeostasis, as well as regulating oxidative stress [14]. Recent evidence has highlighted the protective properties of SIRT6 against CAD in preserving endothelial function, inhibiting inflammatory responses, and regulating glucose and lipid metabolism. Hence, a decline in serum SIRT6 levels has emerged as an independent risk factor for CAD [15], warranting further investigation of its diagnostic and predictive value in older adult patients. Existing risk assessment models, such as the Pooled Cohort Equations model and the Systematic Coronary Risk Estimation model, were primarily developed based on Caucasian and African American populations and lack specific biomarkers and demonstrate poor calibration when applied to the Chinese population [16]. Given these ethnic differences, there is an urgent need for tailored risk assessment models in cardiovascular disease prevention, particularly in diverse populations such as China.

In this study, we aimed to develop a nomogram for predicting outcomes in older adults presenting with clinical symptoms suggestive of CAD. We hypothesized that a combination of baseline SIRT6 levels with clinical parameters could improve the evidence-based selection of candidates for this marker and facilitate clinical decision-making, resulting in its potential implementation in clinical trials.

Methods

Study Design and Participants

In this single-center study, a nomogram was developed for predicting the outcomes of potential cases with CAD and was validated using data from Ruijin Hospital, Shanghai Jiao Tong University School Of Medicine. Patients aged 60 years or older who were diagnosed with angina pectoris were screened between October 1, 2022, and March 31, 2023. The Judkins method [17] was used for performing coronary angiography (CAG) via the radial or femoral artery. The angiographic results underwent joint assessment by 3 experienced cardiovascular specialists (WFS , ZBZ, and JWN). The severity of stenosis in the major coronary arteries, including the left anterior descending branch (LAD), left circumflex artery, right coronary artery, and left main coronary artery, was evaluated, with CAD defined as ≥50% stenosis in any 1 vessel and coronary atherosclerosis (CAS) as <50% stenosis in all vessels.

Eligible participants were individuals aged 60 years or older who presented with chest pain and had been evaluated by a specialist, resulting in a preliminary diagnosis of CAD with an indication for CAG. Participants must have no prior history of CAG, percutaneous coronary intervention, or coronary artery bypass grafting. Furthermore, informed consent must be obtained for the collection of biological samples, with blood samples obtained prior to CAG. Participants must also have coronary artery stenosis as confirmed by the CAG results. Exclusion criteria included recent acute infections, gastrointestinal bleeding, surgical procedures, or trauma within the past 6 months. Individuals were also excluded if they had positive viral markers, including hepatitis B surface antigen, hepatitis B core antibody with hepatitis B virus–DNA of the detection threshold or greater, positive hepatitis C virus antibody with hepatitis C virus–RNA of the upper limit of normal or greater, or positive HIV antibody. Severe cardiac conditions also warranted exclusion, including decompensated heart failure, significant valvular heart disease within the past 6 months, notable electrocardiographic abnormalities (eg, any degree of atrial fibrillation, second-degree type II or third-degree atrioventricular block, or corrected QT interval exceeding 470 milliseconds in females or 450 milliseconds in males), uncontrolled symptomatic arrhythmias, cerebrovascular events or transient ischemic attacks within the past 6 months, a history of malignancy or autoimmune diseases, or severe liver or renal disorders unrelated to the study condition.

Upon admission, patients were assessed for CAD severity using the Gensini score, an angiographic tool for grading coronary artery lesions [18]. Serum specimens were collected from eligible patients for the detection of SIRT6 levels by enzyme-linked immunosorbent assay. Clinical data and the complete CAG inspection results of these participants were also collected.

XSL•FO
RenderX

Ethical Considerations

The study protocol was approved by the human ethics committee of Ruijin Hospital (KY2021-108 Ruijin Hospital). It adhered to strict data confidentiality measures in compliance with the Helsinki Declaration and the institutional guidelines and reporting studies conducted using routinely collected observational data. Participants provided written informed consent at the time of data collection.

Selection Bias

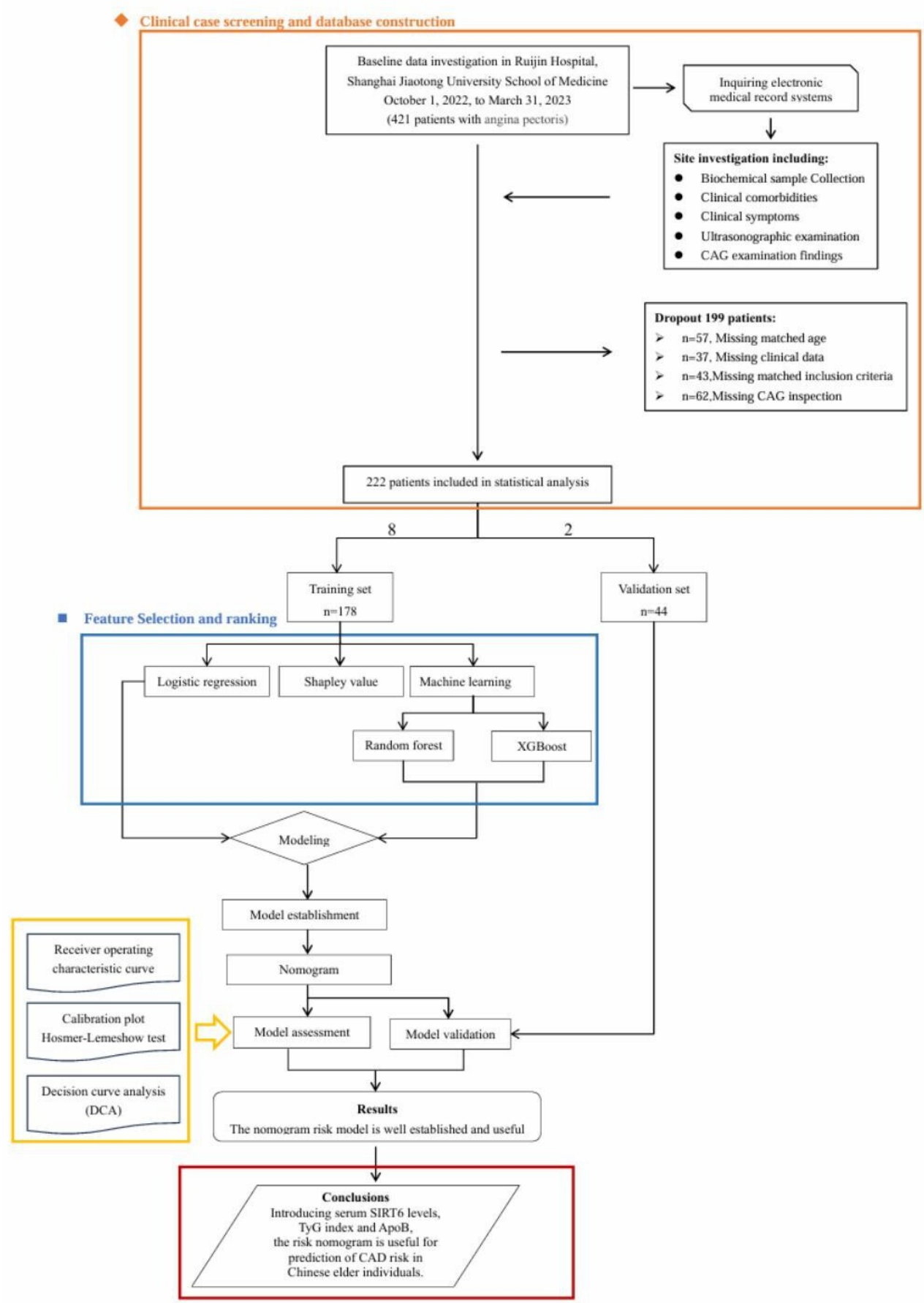
Several rigorous measures were implemented to address potential selection bias in this study. Stringent inclusion and exclusion criteria were carefully defined and applied to ensure a well-characterized and homogeneous study population. Furthermore, random sampling techniques were used where appropriate, and blinding in outcome assessment was ensured to minimize bias in both participant selection and data interpretation. These strategies were designed to enhance the internal validity and robustness of our study results.

Data Collection

All recruited patients were divided into 2 cohorts, the development cohort and the validation cohort, in an approximate ratio of 8:2. Then, predefined criteria were applied to ensure cohort comparability. A total of 29 pretherapeutic parameters were collected, including demographic characteristics, initial symptoms, history of hypertension and diabetes mellitus, baseline clinical status, and baseline laboratory test results. Finally, 3 out of 26 collected parameters were tested as putative predictors for outcomes in the models. Three putative predictors allowed for 9-10 events per predictor for the primary outcome in the training cohort, meeting the recommended minimal number of events per predictor. These putative predictors were selected based on previous research demonstrating their potential prognostic value in CAD and the investigators' (WFS, ZYC, and JWN) clinical experience with CAD. The study flowchart is shown in Figure 1.

XSL•FO
RenderX

Figure 1. Study flowchart. ApoB: apolipoprotein B; CAD: coronary artery disease; CAG:coronary angiography; SIRT6: silent information regulator 6; TyG: triglyceride glucose; XGBoost: Extreme Gradient Boosting.



XSL•FO
RenderX

Statistical Analysis

Variable Selection

In this study, IBM SPSS Statistics (version 26.0; IBM Corp) software was used for baseline description and logistic regression (LR) analysis. Continuous variables were presented as mean (SD), while categorical variables were expressed as percentages (n [%]). Normality was assessed using the Shapiro-Wilk test, with $P>.05$ indicating adherence to normal distribution. For normally distributed variables, an intergroup analysis was conducted using the *t* test, while nonnormally distributed data were compared using the Wilcoxon rank sum test. Features with >30% missing values were removed. The remaining 29 features were collected for further processing. For features with missing values of <10%, median or mean imputation was used. Those with 10% - 20% missing values were imputed by the "MissForest" package in R (R Foundation for Statistical Computing). Random forest and Extreme Gradient Boosting (XGBoost) algorithms were implemented via R statistical software (version 4.3.2; R Foundation for Statistical Computing). Receiver operating characteristic (ROC) curves were plotted using the "pROC" package, while nomograms were constructed using the "rms" and "regplot" packages. The "val.prob" function was used for refining calibration curve plots and Hosmer-Lemeshow tests, while the "dcurves" function facilitated decision curve analysis (DCA). All statistical tests were 2-tailed, with $P<.05$ considered statistically significant. Feature ranking was obtained using Shapley Additive Explanation, Gini, and Gain values, respectively. A 10-fold cross-validation strategy was applied to develop the data set into training, validation, and test sets.

Model Building and Visualization:

Three methodologies, including LR, random forest, and XGBoost, were used for model building. All analyses were conducted using the statistical software package R (version 4.3.2). The data from the Ruijin Hospital of Shanghai Jiao Tong Medical University were used as the training set and internal validation set for model development and verification. Binary LR, random forest, and XGBoost models were constructed using the training data to classify patients into those with CAS and those with CAD. The XGBoost model exhibited the highest comprehensive discriminant ability [19] and therefore was selected for further analysis. The final model was visualized as a column line chart to address issues of poor machine learning interpretability and consequent low clinical use.

Model Comparison

Accuracy, precision, recall, and F_1-score were used to evaluate the performance of the multiclassification model. The model error was further analyzed using a confusion matrix. Discrimination, calibration, and clinical use between the nomogram and the variables incorporated into the nomogram were assessed in both the training and validation sets, respectively. ROC curves, calibration curves, and DCA plots were plotted, and the areas under the ROC curves (AUROCs) were compared using the Delong test. Calibration was evaluated using the Hosmer-Lemeshow test.

Results

Characteristics of the Training and Validation Sets

The final study cohort comprised 222 patients, with 178 allocated to the training set and 44 to the validation set. The demographic characteristics and clinical results of patients in both sets are summarized in Table 1. There were 23.87% (53/222) of patients in the CAS group and 76.13% (169/222) of patients in the CAD group. There were significant differences in the low-density lipoprotein cholesterol (LDL-C), apolipoprotein B (ApoB), apolipoprotein E (ApoE), fasting blood glucose (FBG), hemoglobin A_{1c} (HbA_{1c}), Gensini score, TyG index, and atherogenic index of plasma (AIP) between the CAS and CAD groups ($P<.05$). Specifically, the CAD group exhibited a higher TyG index (12.4 [SD 1.38]) and AIP (0.10 [SD 0.31]) than the CAS group (10.9 [SD 0.54], 0.00 [SD 0.27]), as shown in Table 1. Significant differences were also observed in categorical variables, such as sex, history of diabetes, clinical symptoms, segmental vascular lesions (ie, LAD, left circumflex artery, left main coronary artery, and right coronary artery), and 10-year cardiovascular risk between the CAS and CAD groups ($P<.05$).

XSL•FO
RenderX

Table . Baseline characteristics of participants[a].

Variables	CAS[b] (n=53)	CAD[c] (n=169)	*P* value
Age (years), mean (SD)	69.3 (5.33)	68.9 (5.12)	.56
Sex, n (%)			.03*
Female	28 (52.8)	59 (34.9)	
Male	25 (47.2)	110 (65.1)	
BMI, mean (SD)	25.1 (4.46)	24.6 (2.64)	.50
Drinking, n (%)			1.00
No	47 (88.7)	151 (89.3)	
Yes	6 (11.3)	18 (10.7)	
Smoking, n (%)			.68
No	47 (88.7)	144 (85.2)	
Yes	6 (11.3)	25 (14.8)	
SBP[d], mean (SD)	145 (18.3)	146 (17.8)	.82
DBP[e], mean (SD)	77.6 (10.1)	78.5 (10.8)	.60
Diabetes mellitus, n (%)			.006**
No	43 (81.1)	100 (59.2)	
Yes	10 (18.9)	69 (40.8)	
Hypertension, n (%)			.65
No	20 (37.7)	56 (33.1)	
Yes	33 (62.3)	113 (66.9)	
hs-CRP[f], mean (SD)	2.19 (3.37)	2.15 (3.32)	.94
CysC[g], mean (SD)	1.00 (0.16)	1.59 (7.08)	.28
eGFR[h], mean (SD)	84.1 (9.63)	82.9 (10.6)	.47
mALB[i], mean (SD)	1.40 (1.78)	2.03 (5.04)	.17
cTnI[j], mean (SD)	7.60 (16.1)	25.3 (114)	.05
proBNP[k], mean (SD)	170 (338)	305 (1764)	.35
LDL-C[l], mean (SD)	2.36 (0.89)	1.95 (0.69)	.003**
Lp(a)[m], mean (SD)	0.26 (0.32)	0.31 (0.35)	.35
sLDL[n], mean (SD)	0.76 (0.35)	0.66 (0.31)	.08
ApoA1[o], mean (SD)	1.35 (0.25)	1.28 (0.24)	.10
ApoB[p], mean (SD)	0.76 (0.23)	0.68 (0.19)	.02*
ApoE[q], mean (SD)	3.77 (0.52)	3.60 (0.62)	.05
FBG[r], mean (SD)	5.67 (1.00)	6.10 (1.72)	.03*
HbA_{1c}[s], mean (SD)	6.00 (0.66)	6.46 (1.28)	.001**
FINS[t], mean (SD)	9.87 (5.37)	11.7 (11.3)	.11
CP[u], mean (SD)	2.61 (1.16)	2.72 (1.23)	.56
Carotid plaque, n (%)			.049*
<50%	37 (69.8)	138 (81.7)	
>50%	1 (1.89)	9 (5.33)	

XSL•FO
RenderX

Variables	CAS[b] (n=53)	CAD[c] (n=169)	*P* value
No	15 (28.3)	22 (13)	
Ten-year cardiovascular risk, n (%)			.009**
High	21 (39.6)	107 (63.3)	
Low	7 (13.2)	15 (8.88)	
Middle	25 (47.2)	47 (27.8)	
LAD [v] , n (%)			.01*
No	11 (20.8)	12 (7.10)	
Yes	42 (79.2)	157 (92.9)	
RCA [w] , n (%)			<.001***
No	24 (45.3)	27 (16)	
Yes	29 (54.7)	142 (84)	
LCX [x] , n (%)			<.001***
No	38 (71.7)	57 (33.7)	
Yes	15 (28.3)	112 (66.3)	
LM [y] , n (%)			.01*
No	53 (100)	147 (87)	
Yes	N/A[z]	22 (13)	
SIRT6[aa], mean (SD)	866 (510)	495 (443)	<.001***
TyG[ab], mean (SD)	10.9 (0.54)	12.4 (1.38)	<.001***

XSL•FO
RenderX

Variables	CAS[b] (n=53)	CAD[c] (n=169)	*P* value
AIP[ac], mean (SD)	0.00 (0.27)	0.10 (0.31)	.02*

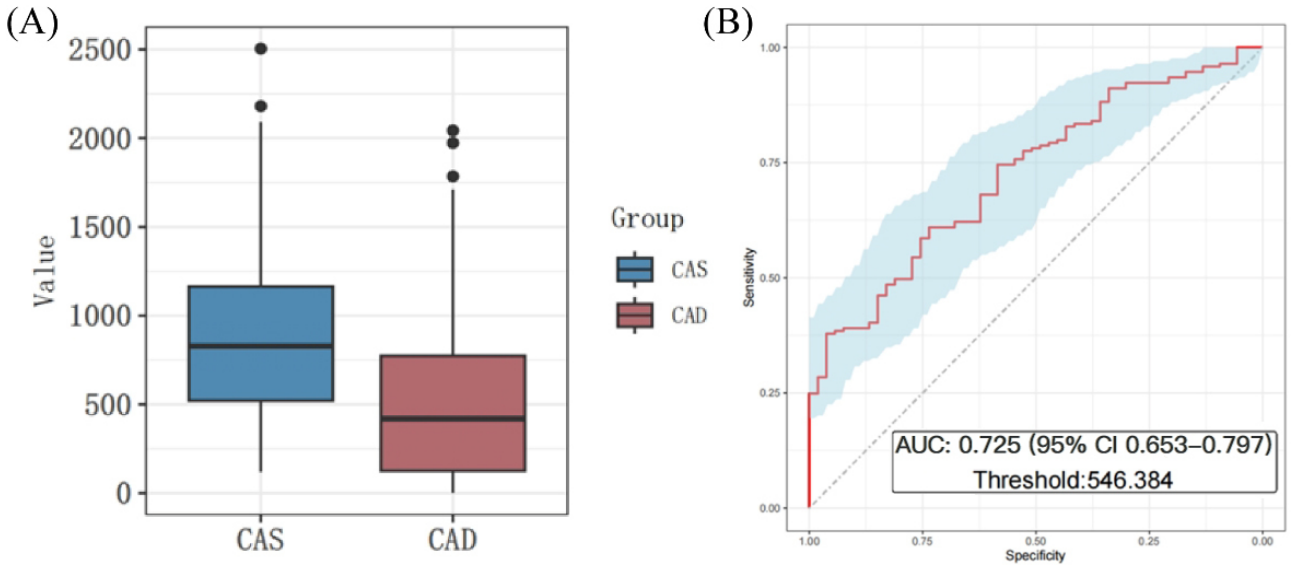
[a]Variables of significance (*$P \leq .05$, ** $P \leq .01$, and *** $P \leq .001$).
[b]CAS, coronary atherosclerosis.
[c]CAD: coronary artery disease.
[d]SBP: systolic pressure.
[e]DBP: diastolic blood pressure.
[f]hs-CRP: hypersensitive C-reactive protein.
[g]CysC: cystatin c.
[h]eGFR: estimated glomerular filtration rate.
[i]mALB: microalbumin.
[j]cTnI: troponin I.
[k]proBNP: pro-B-type natriuretic peptide.
[l]LDL-C: low-density lipoprotein cholesterol.
[m]Lp(a): lipoprotein(a).
[n]sLDL: small dense low-density lipoprotein.
[o]ApoA1: apolipoprotein A1.
[p]ApoB: apolipoprotein B.
[q]ApoE: apolipoprotein E.
[r]FBG: fasting blood glucose.
[s]HbA_{1c}: hemoglobin A_{1c}.
[t]FINS: fasting insulin.
[u]CP: C-peptide.
[v]LAD: left anterior descending branch.
[w]RCA: right coronary artery.
[x]LCX: left circumflex artery.
[y]LM: left main coronary artery.
[z]N/A: not applicable.
[aa]SIRT6: silent information regulator 6.
[ab]TyG: triglyceride glucose.
[ac]AIP: atherogenic index of plasma.

Circulating SIRT6 Levels in Patients

To investigate the potential difference in serum SIRT6 levels between the CAD and CAS groups, the enzyme-linked immunosorbent assay was used to determine the SIRT6 levels in serum samples. A significant reduction in serum SIRT6 levels was observed in the CAD group (495 [SD 443] pg/mL) compared with the CAS group (866 [SD 510] pg/mL) (Figure 2A). To further assess the diagnostic potential of serum SIRT6 as a biomarker for distinguishing between CAS and CAD, ROC analysis was performed. A serum SIRT6 level of 546.384 pg/mL was identified as the optimal cutoff value for discriminating between CAS and CAD (AUROC 0.725, 95% CI 0.653 - 0.797) (Figure 2B).

XSL•FO
RenderX

Figure 2. Serum levels of silent information regulator 6 (SIRT6) and receiver operating characteristic (ROC) curve. (**A**) Serum levels of SIRT6 were lower in patients with CAD and patients with CAS. (**B**) ROC curve analysis of SIRT6 in predicting the diagnosis of CAD (AUROC 0.725, 95% CI 0.653 - 0.797). AUROC: area under the receiver operating characteristic curve; CAD: coronary artery disease; CAS: coronary atherosclerosis.

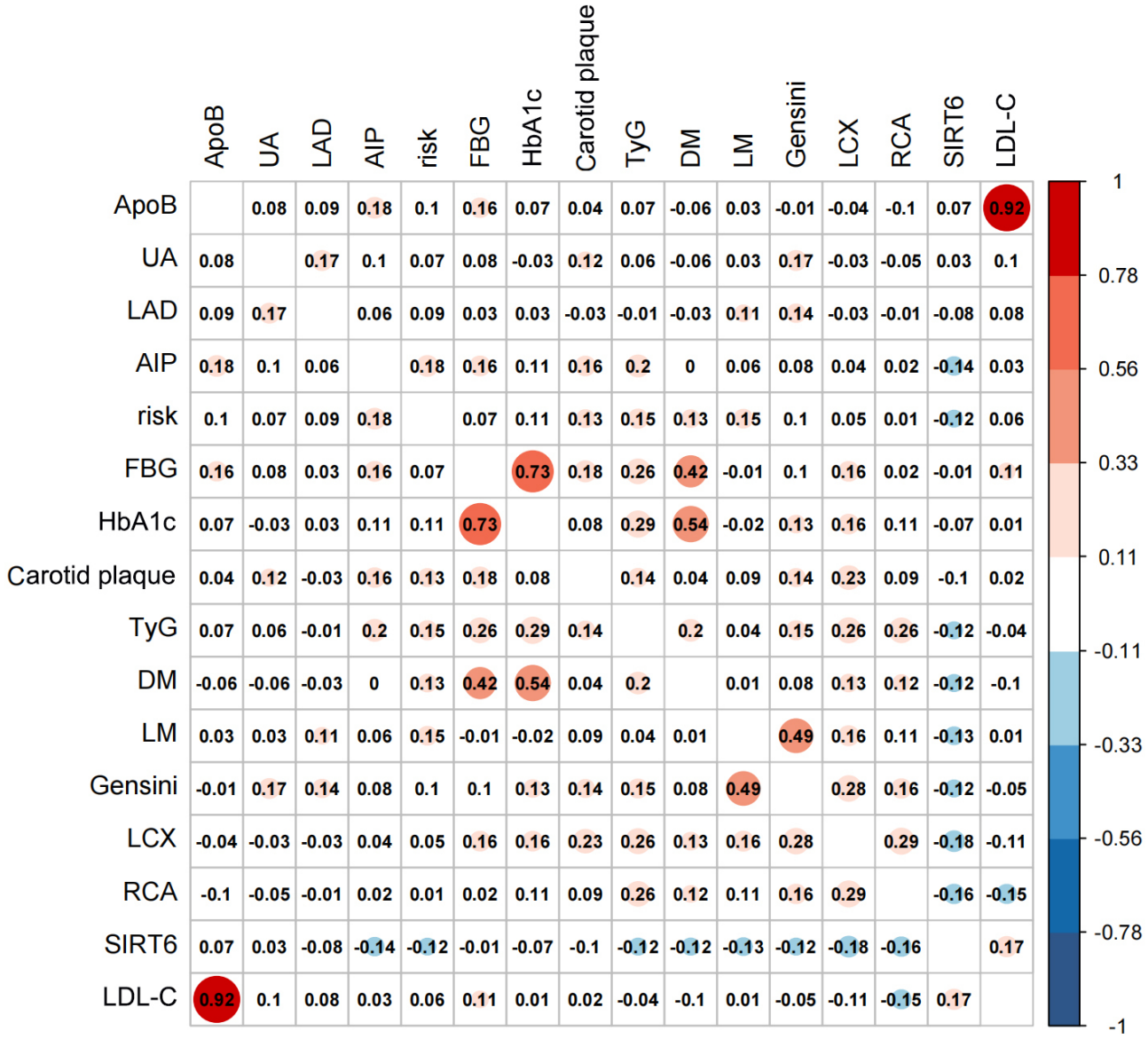


Exploration of Variable Correlations

Factors with significant differences in intergroup analysis, including ApoB, unstable angina pectoris, LDL-C, AIP, FBG, HBA_{1c}, TyG index, history of diabetes mellitus, SIRT6, LAD, AIP, 10-year cardiovascular risk, carotid plaque burden, segmental vascular lesions, and Gensini score, were selected for further interfactor correlation analysis. To illustrate the correlations among these predictive indicators, a correlation coefficient matrix graph was constructed. While serum SIRT6 did not exhibit positive correlations with other indicators, aside from LDL-C, it demonstrated negative correlations with the Gensini score, 10-year cardiovascular risk, and carotid plaque burden (Figure 3).

XSL•FO
RenderX

Figure 3. Correlation heatmap analysis. AIP: atherogenic index of plasma; ApoB: apolipoprotein B; DM: diabetic mellites; FBG: fasting blood glucose; HbA_{1c}: hemoglobin A_{1c}; LAD: left anterior descending branch; LCX: left circumflex artery; LDL-C: low-density lipoprotein cholesterol; LM: left main coronary artery; RCA: right coronary artery; Risk: 10-year cardiovascular risk; SIRT6: silent information regulator 6; TyG: triglyceride glucose; UA: unstable angina pectoris.



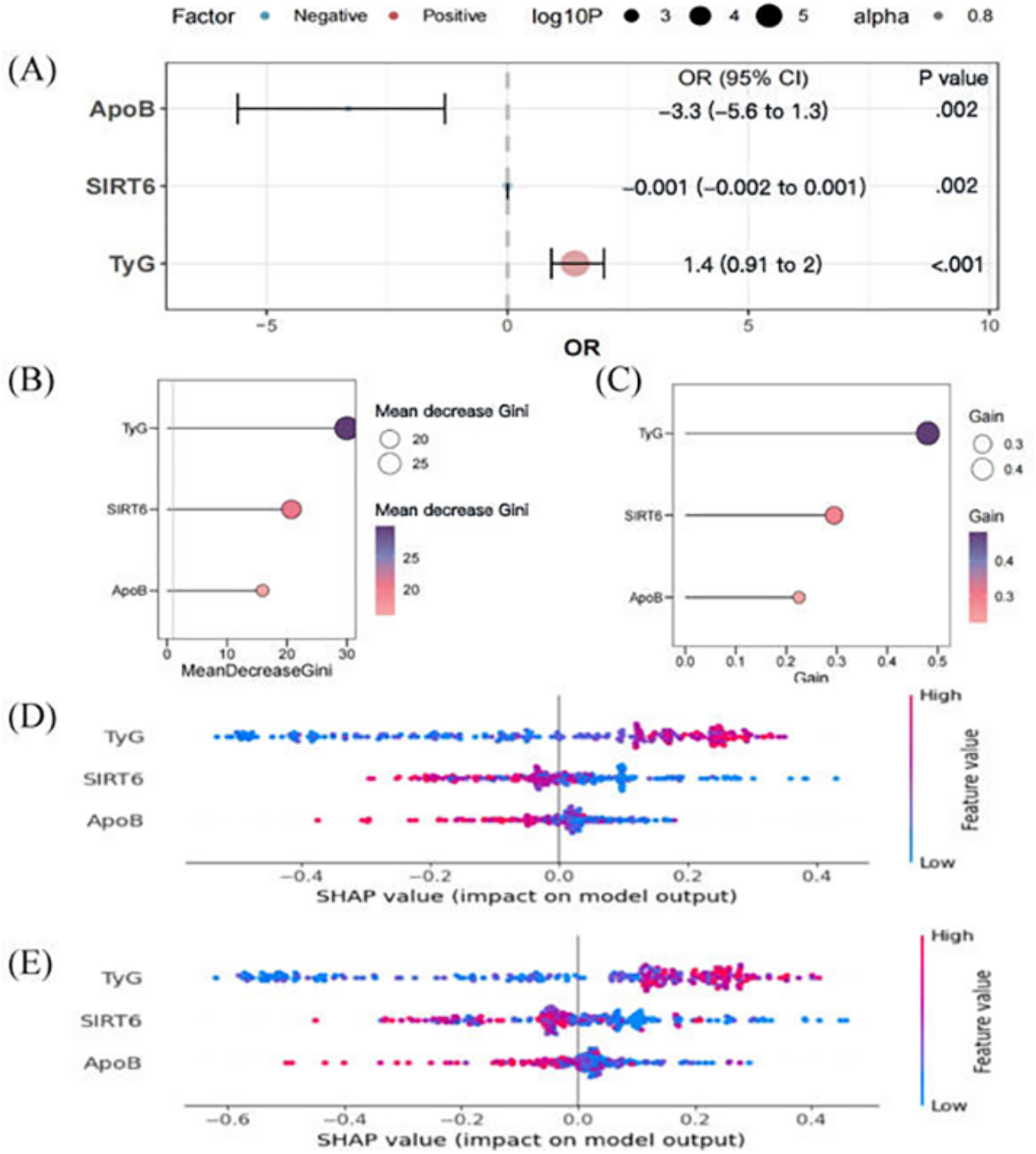
A correlation coefficient matrix graph was constructed. Darker shades represent stronger correlations, with positive correlations indicated by positive values and negative correlations by negative values. The predictive indicators included in this analysis are primarily independent predictive factors.

Feature Selection and Ranking

The significance of variables was preliminarily evaluated by LR analysis. Ultimately, serum SIRT6 levels, TyG index, and ApoB were incorporated into the model construction (Figure 4A). Feature importance ranking was determined using Gini (Figure 4B), Gain (Figure 4C), and Shapley Additive Explanation values (Figure 4D and E), which yielded similar results (Figure 4).

XSL•FO
RenderX

Figure 4. Feature importance. (**A**) Logistic regression results; (**B**) random forest results; (**C**) XGBoost results; (**D**) random forest SHAP; and (**E**) XGBoost SHAP. ApoB: apolipoprotein B; OR: odds ratio; SHAP: Shapley Additive Explanation; SIRT6: silent information regulator 6; TyG: triglyceride glucose; Extreme Gradient Boosting XGBoost.

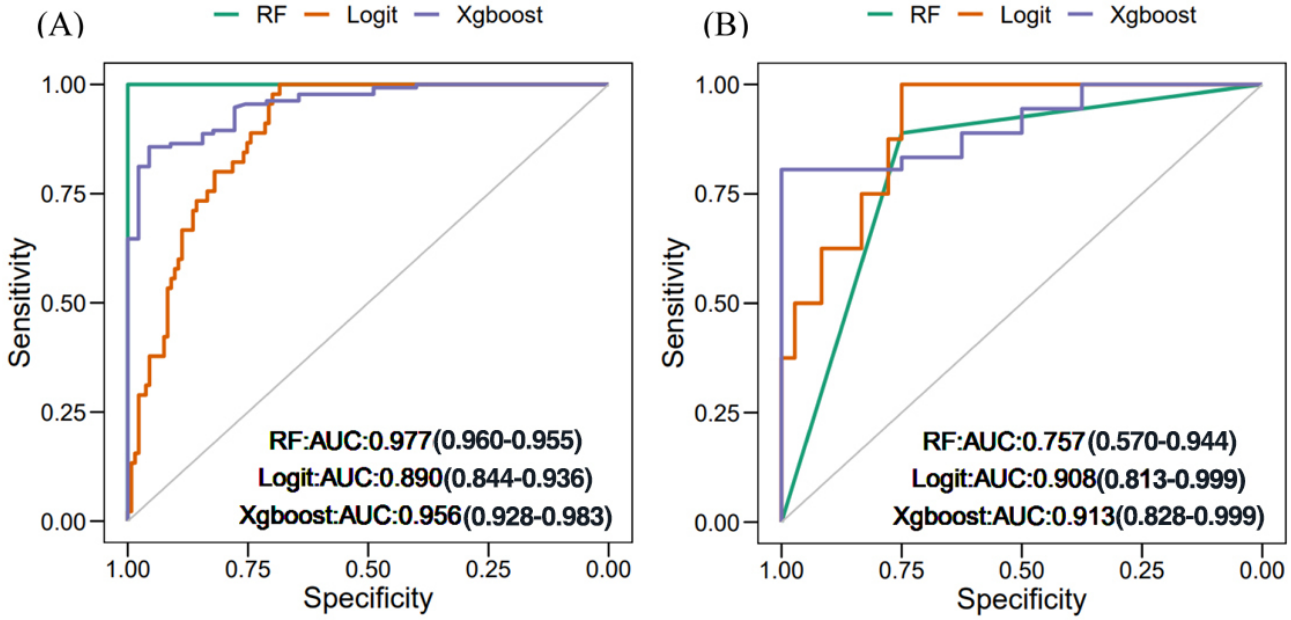


Model Performance Comparisons

An LR model and 2 machine learning models were constructed to predict the development of CAD in older adult patients. The discriminative performance of the 3 models is shown via ROC curves in Figure 5.

XSL•FO
RenderX

Figure 5. Receiver operating characteristic curves for predicting different classes using various models. (**A**) AUROC of the training set and (**B**) AUROC of the validation set. AUROC: area under the receiver operating characteristic curve; RF: random forest; Xgboost: Extreme Gradient Boosting.

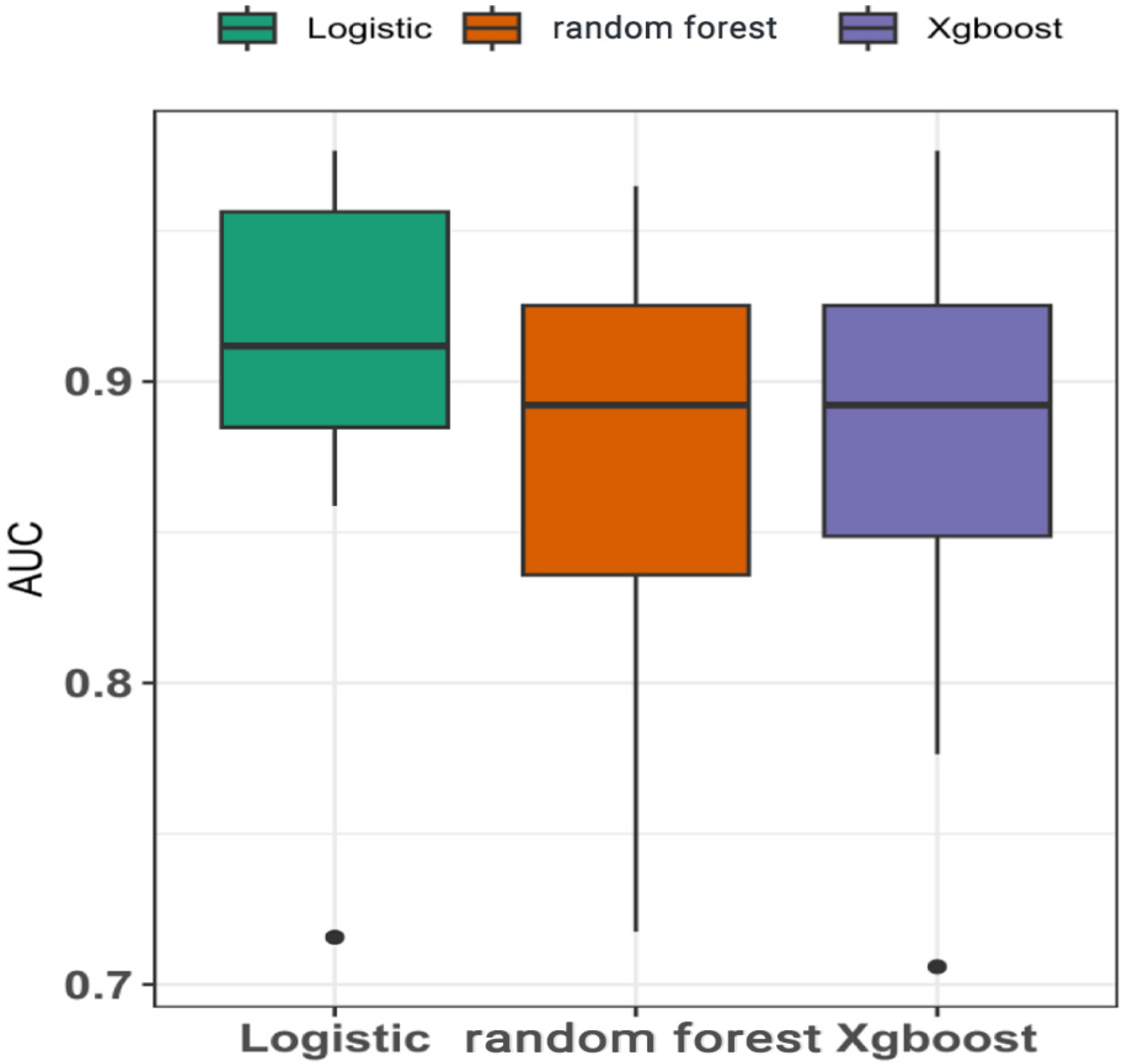


In the training set, the random forest model demonstrated the best predictive ability for CAD in older adult patients (AUROC 0.977, 95% CI 0.960 - 0.995), followed by the XGBoost (AUROC 0.956) and LR (AUROC 0.890) models (Figure 5A). However, in the validation set, the classification performance of the random forest model was noticeably inferior to that of the XGBoost model. The AUROC values of XGBoost, LR, and random forest models were 0.913, 0.908, and 0.757, respectively (Figure 5B).

In cross-validation, the LR model demonstrated the best predictive ability for CAD in older adult patients (AUROC 0.912), followed by the XGBoost (AUROC 0.892) and random forest (AUROC 0.892) models (Figure 6).

XSL•FO
RenderX

Figure 6. Cross-validation AUROC of various models. AUROC: area under the receiver operating characteristic curve; Xgboost: Extreme Gradient Boosting.



Detailed performance metrics of the 3 models are shown in Table 2. The XGBoost model exhibited the best discrimination, with the highest recall (1) and accuracy (0.863), the second-highest precision (0.833), and the second-highest F_1-score (0.909). Although the XGBoost model may not perform as well as the other 2 models in certain aspects, its overall stability and balanced performance on both the training and the validation sets make it a more reliable choice. As a results, we developed a CAD prediction model for the older adults using the XGBoost algorithm and visualized the results through a nomogram. The nomogram based on the serum SIRT6-level model is shown in Figure 7.

Table . Model performance metrics.

Model	Accuracy	Precision	Recall	F_1-score
LR[a]	0.795	0.75	1	0.857
RF[b]	0.863	0.941	0.889	0.914
XGBoost[c]	0.863	0.833	1	0.909

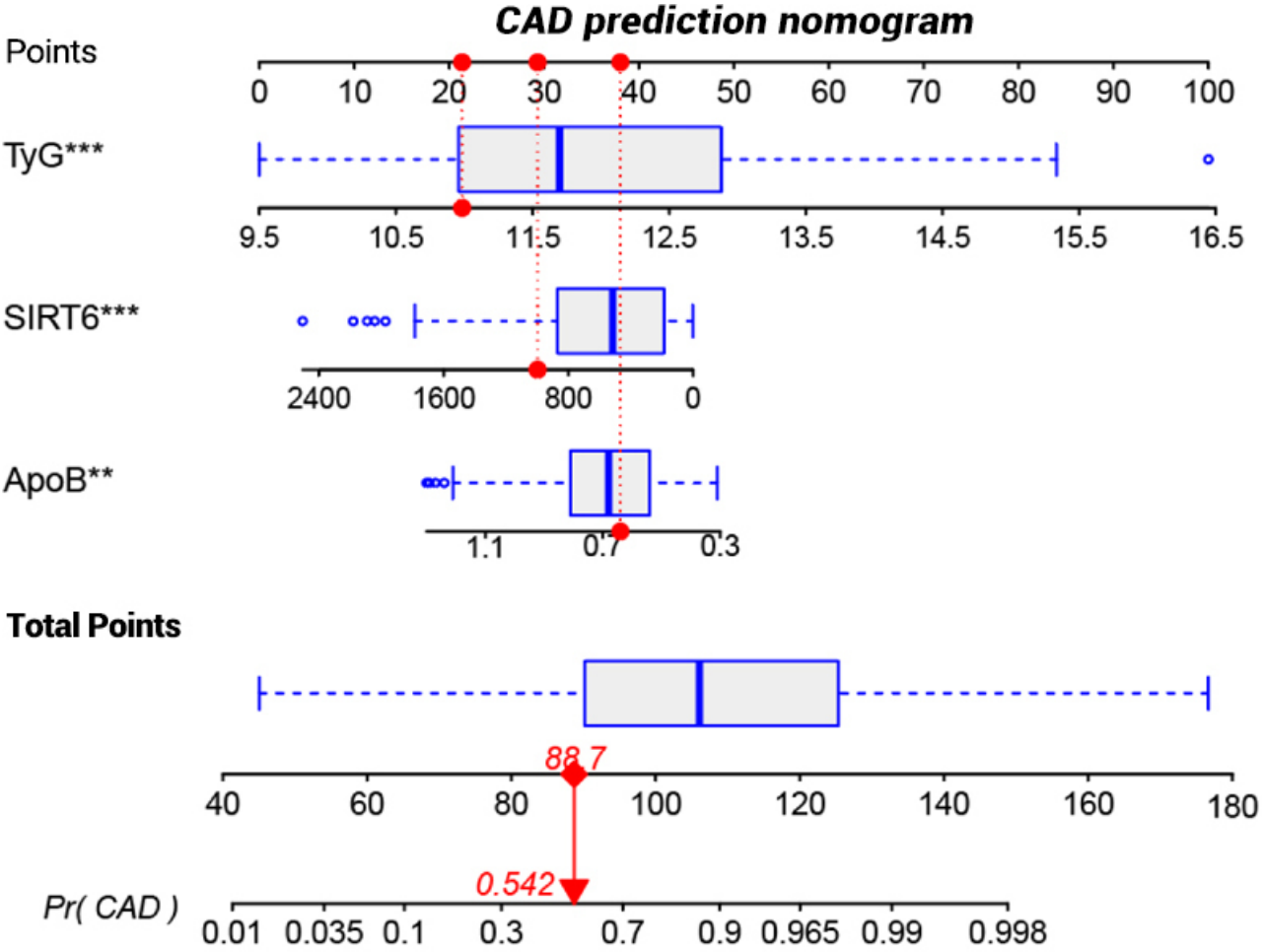
[a]LR: logistic regression.

[b]RF: random forest.

[c]XGBoost: Extreme Gradient Boosting.

XSL•FO
RenderX

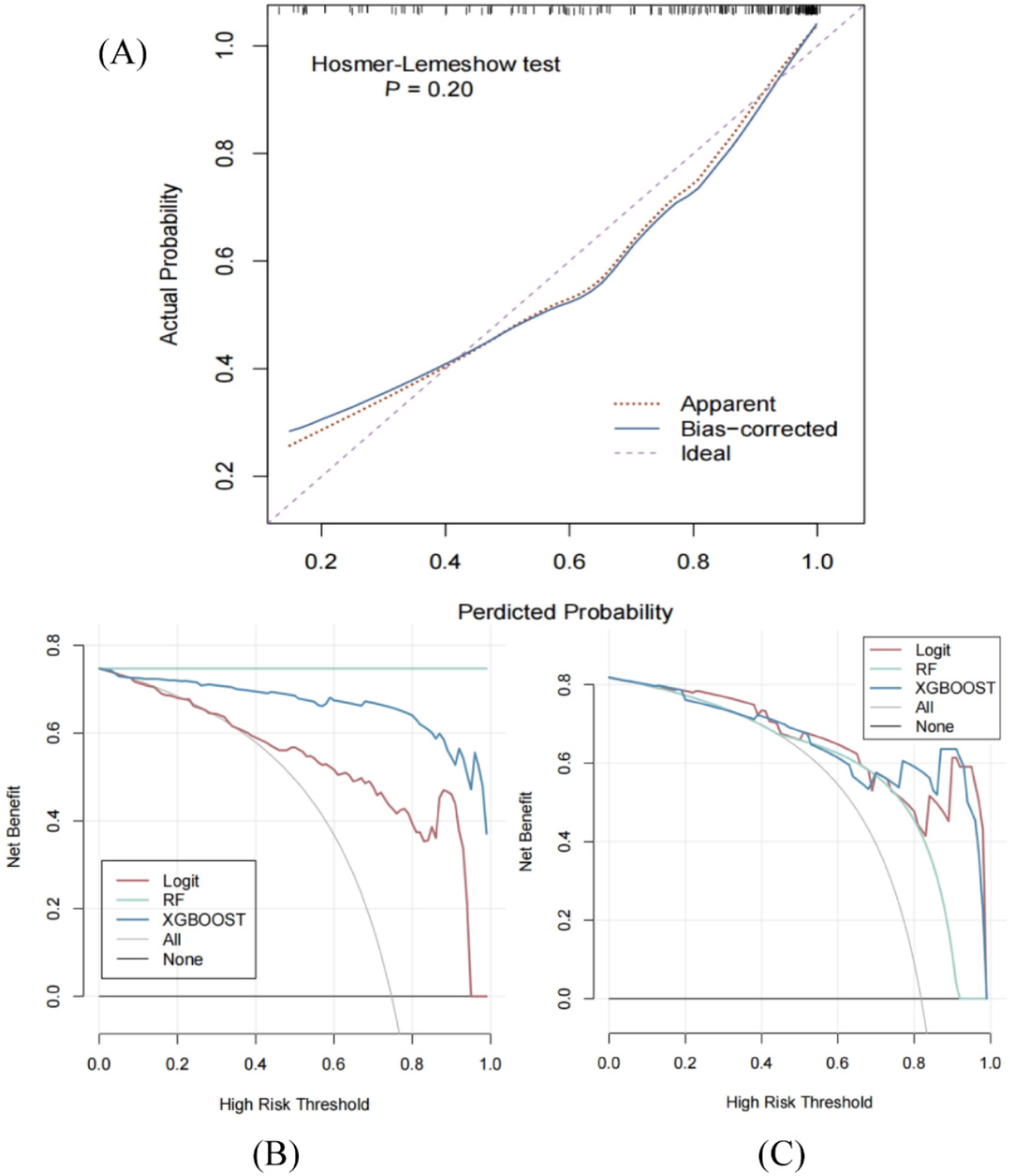
Figure 7. Nomogram for predicting the probability of CAD. Variables of significance (**P≤.01 and ***P≤.001). ApoB: apolipoprotein B; CAD: coronary artery disease; SIRT6: silent information regulator 6; TyG: triglyceride glucose.



The calibration curves (Figure 8A) showed a similar trend among the 3 models, and the Hosmer-Lemeshow test results (χ^{28}=11.001; P=.20) indicated no significant difference between the predicted and observed values. These data suggest that the XGBoost model has good calibration ability.

XSL•FO
RenderX

Figure 8. Calibration curve and decision curve analyses for predicting coronary artery disease. (**A**) Calibration curves of the training cohort; (**B**) decision curve analysis (DCA) of the training cohort; and (**C**) DCA of the validation cohort. RF: random forest; XGBoost: Extreme Gradient Boosting.



The DCA revealed that the XGBoost model exhibited greater net benefit along with the threshold probability, indicating its superior clinical use compared with other models (Figure 8B and C).

Discussion

Principal Findings

An accurate risk assessment is critical for patient-centered clinical decision-making. The nomogram presented in this study may serve as an important clinical decision aid tool, assisting in determining whether invasive CAG should be performed, or informing patients as a basis for joint decision-making. In this study, traditional LR and 2 common machine learning

XSL•FO
RenderX

algorithms were used to establish a prediction and evaluation model of coronary artery stenosis in older adult patients with suspected CAD. Through a comprehensive analysis of models, the XGBoost algorithm–based model with excellent prediction performance and good internal verification ability was selected. In addition, we developed a nomogram for CAD prediction by combining traditional clinical variables (eg, ApoB) with 2 novel variables relevant to the older adult population, serum SIRT6 levels and the TyG index, which allowed for assessing the risk of coronary artery stenosis more conveniently and interactively (Figure 6). This, to some extent, helps address the issue of poor interpretability, which often hampers the implementation of machine learning models.

Given its significant role in genomic stability, DNA repair, and telomere function, SIRT6 has been considered to have therapeutic potential in aging and aging-related diseases. Moreover, SIRT6 plays a crucial role in the development and progression of CAD due to its involvement in oxidative stress, inflammation, and energy metabolism [20]. Previous studies have highlighted the importance of SIRT6 in protecting blood vessels and the heart from endothelial dysfunction, atherosclerosis, myocardial fibrosis, and ischemia or reperfusion injury [21,22].

In this study, we observed lower serum SIRT6 levels in older adult patients with CAD than in those with CAS. Correlation analysis revealed a negative association between serum SIRT6 levels and the Gensini score, suggesting that reduced serum SIRT6 levels may exacerbate the severity of coronary artery lesions. Furthermore, serum SIRT6 levels were negatively correlated with segmental vascular lesions and carotid plaque burden, indicating that serum SIRT6 may reflect the overall burden of atherosclerotic plaques in blood vessels. Dyslipidemia is an independent risk factor for the occurrence and development of CAD, while diabetes mellitus doubles the risk of cardiovascular disease and is a leading cause of mortality in patients with dyslipidemia [23,24].

LDL-C levels are considered a determinant of the absolute risk of major cardiovascular events [25]. Previous evidence has revealed that liver-specific knockout of SIRT6 significantly increases the proprotein convertase subtilisin/kexin type 9 (*PCSK9*) gene expression and plasma LDL-C levels, while SIRT6 overexpression improves lipid metabolism and reduces atherosclerosis risk by lowering plasma LDL-C and *PCSK9* levels [26]. Although a previous clinical study in a Chinese population of all ages found no correlation between serum SIRT6 and LDL-C levels [15], in this study, we found a positive correlation between serum SIRT6 and LDL-C levels. Future investigations are needed to explore the mechanisms by which SIRT6 affects cardiovascular function through lipid metabolism in the older adult population.

Disruption of glucose homeostasis is another recognized risk factor for CAD. Animal models have shown that SIRT6 plays an important role in glucose production and uptake, insulin signaling, and metabolism [27]. Our data revealed that circulating SIRT6 levels were negatively correlated with FBG and the comorbidity of diabetes, which is consistent with previous findings [28]. Whether glucose metabolism disruption mediates the relationship between SIRT6 and CAD warrants further investigation. In addition, our study identified serum SIRT6 as an independent biomarker for assessing the severity of coronary artery lesions in older adults, with the optimal cutoff value for classifying atherosclerosis and CAD being 546.384 pg/mL.

Abnormal blood lipids are established independent risk factors for CAD. Unlike previous studies that commonly included LDL-C, HDL-C, and other conventional measures, our study incorporated ApoB as one of the predictive factors. ApoB is a crucial component of LDL-C, and its deposition within the arterial wall is a fundamental step driving the progression of atherosclerosis, from initial lipid deposition to the development of acute complex events, such as plaque rupture [29]. Compared with other individual lipid markers, ApoB serves as a more accurate biomarker for cardiovascular risk assessment [30,31]. Elevated ApoB levels, representing LDL levels, are associated with the occurrence of CAD, and ApoB levels are positively correlated with the Gensini score, indicating the degree of atherosclerosis and arterial narrowing [32]. In addition, circulating ApoB levels are considered more predictive than plasma cholesterol levels for early-onset CAD risk [33]. Correlation analysis indicated a significant positive correlation between ApoB and LDL-C levels. Moreover, in all 3 statistical methods used for feature selection, ApoB consistently emerged as significant, which is in line with previous findings [31]. However, in an intergroup analysis, we observed higher levels of ApoB in the CAS group than in the CAD group, which is contrary to previous results [33].

Insulin resistance further disturbs glucose and lipid metabolism in patients with diabetes, promoting chronic inflammation, disrupting normal endothelial function, and accelerating the development of complications, such as atherosclerosis-related cardiovascular diseases. The persistent prevalence of insulin resistance is considered a major contributor to the high mortality rate in atherosclerosis-related cardiovascular diseases worldwide [34]. The TyG index is a quantitative measure based on blood glucose and triglycerides that assesses insulin resistance. A clinical study has shown an association between the TyG index and coronary artery calcification [35]. However, in cohort studies, the TyG index was found to be correlated with carotid atherosclerosis [36] but unrelated to the incidence of CAD [37]. In our study, intergroup analysis revealed that the TyG index was significantly higher in the CAD group than in the CAS group. Correlation analysis indicated positive associations between the TyG index and the Gensini score, carotid plaque burden, and segmental vascular lesions. Furthermore, in multifactorial regression analysis and machine learning feature selection, the TyG index consistently demonstrated a strong correlation with CAD, suggesting that it may be an independent risk factor for CAD. Furthermore, we observed a negative correlation between serum SIRT6 levels and the TyG index. Future research is needed to explore the impact of SIRT6 on the TyG index and to elucidate the relationship between these factors in terms of glucose and lipid metabolism disturbance and the development of CAD.

XSL•FO
RenderX

Strengths and Limitations

Our prediction model addresses the challenge of poor interpretability associated with machine learning models. By incorporating readily obtainable clinical and physiological indicators, this nomogram not only provides prediction results and probabilities directly but also facilitates personalized intervention through probability curves. Tailored patient treatment may reduce medical costs and unnecessary invasive tests. Particularly in resource-constrained medical settings, our model may assist in disease screening and alleviate the medical burden.

Although the overall performance of the model was good, the sample distribution in our study was not balanced, and the model was not validated externally with independent data sets. Therefore, the generalizability and extrapolation to the overall population cannot be currently estimated. Second, the model was more inclined to classify patients without the disease as having the disease, while the proportion of patients with the disease misclassified as being disease-free was very low. The rate of missed diagnosis was also low. Third, the limited small sample size and single-center, cross-sectional design constrained the generalizability of the results. Furthermore, this study specifically focused on the older adult population; however, there is currently no consensus on age-related changes in circulating ApoB levels. Furthermore, some studies have suggested that the ApoB–apolipoprotein A1 (ApoA1) ratio has a higher predictive value for atherosclerosis and intima-media thickness than individual lipid markers. Therefore, in future research, expanding the sample size and using the ApoB-ApoA1 ratio as a composite indicator are needed to further evaluate the correlation between ApoB and the severity of coronary artery lesions.

Conclusions

Based on the LR, random forest, and XGBoost algorithms, we developed a predictive model for CAD occurrence in the older adult population, incorporating SIRT6, the TyG index, and ApoB. The model was comprehensively evaluated for discrimination, calibration, clinical applicability, and internal validation. The XGBoost predictive model exhibited favorable predictive performance and clinical use, which may facilitate early CAD screening and diagnosis in older adult patients, particularly for identifying high-risk individuals. Furthermore, the model may reduce unnecessary invasive examinations in negative patients and minimize missed diagnoses in positive patients. The personalized probability curve generated by the model may offer targeted intervention guidance. Our findings also underscore the importance of considering risk factors such as SIRT6, ApoB, and TyG levels in this population.

Acknowledgments

This research is supported by a grant from the National Natural Science Foundation of China (grant 82103079) and the Shanghai Municipal Health Commission Clinical Research in Health Industry Project (grant 202140052).

Authors' Contributions

YY conceptualized and designed the study, collected and analyzed data, drafted the manuscript, and implemented the study. WL and YZ provided the study materials instrumentation and other analysis tools and checked literature. All authors had full access to all the data in the study and accepted responsibility to submit for publication. All authors revised the manuscript and approved the final manuscript as submitted.

Conflicts of Interest

None declared.

References

1. Mensah GA, Fuster V, Murray CJL, Roth GA, Global Burden of Cardiovascular Diseases and Risks Collaborators. Global burden of cardiovascular diseases and risks, 1990-2022. J Am Coll Cardiol 2023 Dec 19;82(25):2350-2473. [doi: 10.1016/j.jacc.2023.11.007] [Medline: 38092509]
2. Crea F. The burden of cardiovascular risk factors: a global perspective. Eur Heart J 2022 Aug 7;43(30):2817-2820. [doi: 10.1093/eurheartj/ehac430]
3. Liu S, Li Y, Zeng X, et al. Burden of cardiovascular diseases in China, 1990-2016: findings from the 2016 Global Burden of Disease Study. JAMA Cardiol 2019 Apr 1;4(4):342-352. [doi: 10.1001/jamacardio.2019.0295] [Medline: 30865215]
4. Roth GA, Mensah GA, Fuster V. The global burden of cardiovascular diseases and risks: a compass for global action. J Am Coll Cardiol 2020 Dec 22;76(25):2980-2981. [doi: 10.1016/j.jacc.2020.11.021] [Medline: 33309174]
5. Goyal P, Kwak MJ, Al Malouf C, et al. Geriatric cardiology: coming of age. JACC Adv 2022 Aug;1(3):100070. [doi: 10.1016/j.jacadv.2022.100070] [Medline: 37705890]
6. Wan X, Ren H, Ma E, Yang G. Mortality trends for ischemic heart disease in China: an analysis of 102 continuous disease surveillance points from 1991 to 2009. BMC Public Health 2017 Jul 25;18(1):52. [doi: 10.1186/s12889-017-4558-3] [Medline: 28743245]
7. The Lancet. Population ageing in China: crisis or opportunity? Lancet 2022 Nov;400(10366):1821. [doi: 10.1016/S0140-6736(22)02410-2]

XSL•FO
RenderX

8. He X, Song M, Qu J, et al. Basic and translational aging research in China: present and future. Protein Cell 2019 Jul;10(7):476-484. [doi: 10.1007/s13238-019-0617-0]
9. Bjarnason-Wehrens B. Elderly patients with ischaemic heart disease need our attention!. Eur J Prev Cardiol 2018 Feb;25(3):276-277. [doi: 10.1177/2047487317750497] [Medline: 29313365]
10. Nanna MG, Chen ST, Nelson AJ, Navar AM, Peterson ED. Representation of older adults in cardiovascular disease trials since the inclusion across the lifespan policy. JAMA Intern Med 2020 Nov 1;180(11):1531-1533. [doi: 10.1001/jamainternmed.2020.2750] [Medline: 32897289]
11. Forman DE, Maurer MS, Boyd C, et al. Multimorbidity in older adults with cardiovascular disease. J Am Coll Cardiol 2018 May 15;71(19):2149-2161. [doi: 10.1016/j.jacc.2018.03.022] [Medline: 29747836]
12. Mendieta G, Pocock S, Mass V, et al. Determinants of progression and regression of subclinical atherosclerosis over 6 years. J Am Coll Cardiol 2023 Nov 28;82(22):2069-2083. [doi: 10.1016/j.jacc.2023.09.814] [Medline: 37993199]
13. Cho SMJ, Koyama S, Honigberg MC, et al. Genetic, sociodemographic, lifestyle, and clinical risk factors of recurrent coronary artery disease events: a population-based cohort study. Eur Heart J 2023 Sep 21;44(36):3456-3465. [doi: 10.1093/eurheartj/ehad380] [Medline: 37350734]
14. Guo Z, Li P, Ge J, Li H. SIRT6 in aging, metabolism, inflammation and cardiovascular diseases. Aging Dis 2022 Dec 1;13(6):1787-1822. [doi: 10.14336/AD.2022.0413] [Medline: 36465178]
15. Yan Z, Wang X, Liu YS, Xing XW, Zhang XG, Lu QH. Decreased serum SIRT6 as a novel predictor of coronary artery disease. Eur Rev Med Pharmacol Sci 2021 Nov;25(21):6660-6669. [doi: 10.26355/eurrev_202111_27110] [Medline: 34787870]
16. Piccininni M, Rohmann JL, Huscher D, et al. Performance of risk prediction scores for cardiovascular mortality in older persons: external validation of the SCORE OP and appraisal. PLoS One 2020;15(4):e0231097. [doi: 10.1371/journal.pone.0231097] [Medline: 32271825]
17. Rao SV, Stone GW. Arterial access and arteriotomy site closure devices. Nat Rev Cardiol 2016 Nov;13(11):641-650. [doi: 10.1038/nrcardio.2016.133] [Medline: 27558003]
18. Rampidis GP, Benetos G, Benz DC, Giannopoulos AA, Buechel RR. A guide for Gensini Score calculation. Atherosclerosis 2019 Aug;287:181-183. [doi: 10.1016/j.atherosclerosis.2019.05.012] [Medline: 31104809]
19. Han JW, Lee SK, Kwon JH, et al. A machine learning algorithm facilitates prognosis prediction and treatment selection for Barcelona clinic liver cancer stage C hepatocellular carcinoma. Clin Cancer Res 2024 Jul 1;30(13):2812-2821. [doi: 10.1158/1078-0432.CCR-23-3978] [Medline: 38639918]
20. Guo J, Wang Z, Wu J, et al. Endothelial SIRT6 is vital to prevent hypertension and associated cardiorenal injury through targeting Nkx3.2-GATA5 signaling. Circ Res 2019 May 10;124(10):1448-1461. [doi: 10.1161/CIRCRESAHA.118.314032] [Medline: 30894089]
21. Zhu Y, Hu S, Pan X, et al. Hepatocyte Sirtuin 6 protects against atherosclerosis and steatohepatitis by regulating lipid homeostasis. Cells 2023 Aug 5;12(15):2009. [doi: 10.3390/cells12152009] [Medline: 37566087]
22. Zhang Q, Tu W, Tian K, et al. Sirtuin 6 inhibits myofibroblast differentiation via inactivating transforming growth factor - β1/Smad2 and nuclear factor - κB signaling pathways in human fetal lung fibroblasts. J of Cell Biochem 2019 Jan;120(1):93-104. [doi: 10.1002/jcb.27128]
23. Pastromas S, Terzi AB, Tousoulis D, Koulouris S. Postprandial lipemia: an under-recognized atherogenic factor in patients with diabetes mellitus. Int J Cardiol 2008 May 7;126(1):3-12. [doi: 10.1016/j.ijcard.2007.04.172] [Medline: 17689745]
24. McAuley PA, Artero EG, Sui X, Lavie CJ, Almeida MJ, Blair SN. Fitness, fatness, and survival in adults with prediabetes. Diabetes Care 2014 Feb;37(2):529-536. [doi: 10.2337/dc13-1347] [Medline: 24062333]
25. Ference BA, Graham I, Tokgozoglu L, Catapano AL. Impact of lipids on cardiovascular health: JACC health promotion series. J Am Coll Cardiol 2018 Sep 4;72(10):1141-1156. [doi: 10.1016/j.jacc.2018.06.046] [Medline: 30165986]
26. Tao R, Xiong X, DePinho RA, Deng CX, Dong XC. FoxO3 transcription factor and Sirt6 deacetylase regulate low density lipoprotein (LDL)-cholesterol homeostasis via control of the proprotein convertase subtilisin/kexin type 9 (Pcsk9) gene expression. J Biol Chem 2013 Oct 11;288(41):29252-29259. [doi: 10.1074/jbc.M113.481473] [Medline: 23974119]
27. Mostoslavsky R, Chua KF, Lombard DB, et al. Genomic instability and aging-like phenotype in the absence of mammalian SIRT6. Cell 2006 Jan 27;124(2):315-329. [doi: 10.1016/j.cell.2005.11.044] [Medline: 16439206]
28. Wu X, Liu H, Brooks A, et al. SIRT6 mitigates heart failure with preserved ejection fraction in diabetes. Circ Res 2022 Nov 11;131(11):926-943. [doi: 10.1161/CIRCRESAHA.121.318988] [Medline: 36278398]
29. Borén J, Williams KJ. The central role of arterial retention of cholesterol-rich apolipoprotein-B-containing lipoproteins in the pathogenesis of atherosclerosis: a triumph of simplicity. Curr Opin Lipidol 2016 Oct;27(5):473-483. [doi: 10.1097/MOL.0000000000000330] [Medline: 27472409]
30. Ference BA, Kastelein JJP, Ray KK, et al. Association of triglyceride-lowering LPL variants and LDL-C-lowering LDLR variants with risk of coronary heart disease. JAMA 2019 Jan 29;321(4):364-373. [doi: 10.1001/jama.2018.20045] [Medline: 30694319]
31. Welsh C, Celis-Morales CA, Brown R, et al. Comparison of conventional lipoprotein tests and apolipoproteins in the prediction of cardiovascular disease. Circulation 2019 Aug 13;140(7):542-552. [doi: 10.1161/CIRCULATIONAHA.119.041149]

XSL•FO
RenderX

32. Campeau L, Enjalbert M, Lespérance J, et al. The relation of risk factors to the development of atherosclerosis in saphenous-vein bypass grafts and the progression of disease in the native circulation. A study 10 years after aortocoronary bypass surgery. N Engl J Med 1984 Nov 22;311(21):1329-1332. [doi: 10.1056/NEJM198411223112101] [Medline: 6333635]
33. Quispe R, Martin SS, Michos ED, et al. Remnant cholesterol predicts cardiovascular disease beyond LDL and ApoB: a primary prevention study. Eur Heart J 2021 Nov 7;42(42):4324-4332. [doi: 10.1093/eurheartj/ehab432]
34. Bornfeldt KE, Tabas I. Insulin resistance, hyperglycemia, and atherosclerosis. Cell Metab 2011 Nov 2;14(5):575-585. [doi: 10.1016/j.cmet.2011.07.015] [Medline: 22055501]
35. Lee SB, Ahn CW, Lee BK, et al. Association between triglyceride glucose index and arterial stiffness in Korean adults. Cardiovasc Diabetol 2018 Mar 21;17(1):41. [doi: 10.1186/s12933-018-0692-1] [Medline: 29562908]
36. Irace C, Carallo C, Scavelli FB, et al. Markers of insulin resistance and carotid atherosclerosis. A comparison of the homeostasis model assessment and triglyceride glucose index. Int J Clin Pract 2013 Jul;67(7):665-672. [doi: 10.1111/ijcp.12124] [Medline: 23758445]
37. Vega GL, Barlow CE, Grundy SM, Leonard D, DeFina LF. Triglyceride-to-high-density-lipoprotein-cholesterol ratio is an index of heart disease mortality and of incidence of type 2 diabetes mellitus in men. J Investig Med 2014 Feb;62(2):345-349. [doi: 10.2310/JIM.0000000000000044] [Medline: 24402298]

Abbreviations

AIP: atherogenic index of plasma
ApoA1: apolipoprotein A1
ApoB: apolipoprotein B
ApoE: apolipoprotein E
AUROC: area under the receiver operating characteristic curve
CAD: coronary artery disease
CAG: coronary angiography
CAS: coronary atherosclerosis
DCA: decision curve analysis
FBG: fasting blood glucose
HbA_{1c}: hemoglobin A_{1c}
LAD: left anterior descending branch
LDL-C: low-density lipoprotein cholesterol
LR: logistic regression
***PCSK9*:** proprotein convertase subtilisin/kexin type 9
ROC: receiver operating characteristic
SIRT6: silent information regulator 6
TyG: triglyceride glucose
XGBoost: Extreme Gradient Boosting

Edited by Y Jiang; submitted 17.07.24; peer-reviewed by D Hu, XM Tian; revised version received 27.09.24; accepted 04.10.24; published 15.01.25.

Please cite as:
You Y, Liang W, Zhao Y
Development and Validation of a Predictive Model Based on Serum Silent Information Regulator 6 Levels in Chinese Older Adult Patients: Cross-Sectional Descriptive Study
JMIR Aging 2025;8:e64374
URL: https://aging.jmir.org/2025/1/e64374
doi:10.2196/64374

© Yuzi You, Wei Liang, Yajie Zhao. Originally published in JMIR Aging (https://aging.jmir.org), 15.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Wearable Smartphone-Based Multisensory Feedback System for Torso Posture Correction: Iterative Design and Within-Subjects Study

Amanda Polin Pereira[1*], BSc, MSc; Olibario Jose Machado Neto[2*], BSc, MSc, PhD; Valeria Meirelles Carril Elui[1*], BSc, MSc, DSC, HDip; Maria da Graca Campos Pimentel[3*], BSc, MSc, DPhil, HDip, Prof Dr

[1]Faculty of Medicine of Ribeirão Preto, University of São Paulo, Ribeirão Preto SP, Brazil

[2]State University of Mato Grosso do Sul, Nova Andradina MS, Brazil

[3]Institute of Mathematics and Computer Sciences, University of São Paulo, São Carlos SP, Brazil

[*]all authors contributed equally

Corresponding Author:
Maria da Graca Campos Pimentel, BSc, MSc, DPhil, HDip, Prof Dr
Institute of Mathematics and Computer Sciences
University of São Paulo
Avenida Trabalhador São-carlense, 400
1203
São Carlos SP, 13566-590
Brazil
Phone: 55 1633739671
Email: mgp@icmc.usp.br

Abstract

Background: The prevalence of stroke is high in both males and females, and it rises with age. Stroke often leads to sensor and motor issues, such as hemiparesis affecting one side of the body. Poststroke patients require torso stabilization exercises, but maintaining proper posture can be challenging due to their condition.

Objective: Our goal was to develop the Postural SmartVest, an affordable wearable technology that leverages a smartphone's built-in accelerometer to monitor sagittal and frontal plane changes while providing visual, tactile, and auditory feedback to guide patients in achieving their best-at-the-time posture during rehabilitation.

Methods: To design the Postural SmartVest, we conducted brainstorming sessions, therapist interviews, gathered requirements, and developed the first prototype. We used this initial prototype in a feasibility study with individuals without hemiparesis (n=40, average age 28.4). They used the prototype during 1-hour seated sessions. Their feedback led to a second prototype, which we used in a pilot study with a poststroke patient. After adjustments and a kinematic assessment using the Vicon Gait Plug-in system, the third version became the Postural SmartVest. We assessed the Postural SmartVest in a within-subject experiment with poststroke patients (n=40, average age 57.1) and therapists (n=20, average age 31.3) during rehabilitation sessions. Participants engaged in daily activities, including walking and upper limb exercises, without and with app feedback.

Results: The Postural SmartVest comprises a modified off-the-shelf athletic lightweight compression tank top with a transparent pocket designed to hold a smartphone running a customizable Android app securely. This app continuously monitors sagittal and frontal plane changes using the built-in accelerometer sensor, providing multisensory feedback through audio, vibration, and color changes. Patients reported high ratings for weight, comfort, dimensions, effectiveness, ease of use, stability, durability, and ease of adjustment. Therapists noted a positive impact on rehabilitation sessions and expressed their willingness to recommend it. A 2-tailed t-test showed a significant difference (P<.001) between the number of the best-at-the-time posture positions patients could maintain in 2 stages, without feedback (mean 13.1, SD 7.12) and with feedback (mean 4.2, SD 3.97), demonstrating the effectiveness of the solution in improving posture awareness.

Conclusions: The Postural SmartVest aids therapists during poststroke rehabilitation sessions and assists patients in improving their posture during these sessions.

(JMIR Aging 2025;8:e55455) doi:10.2196/55455

XSL•FO
RenderX

KEYWORDS

stroke rehabilitation; posture; postural balance; wearable technology; multisensory feedback; smartphone; stroke; mHealth; mobile health; digital health; digital technology; digital intervention; wearable technology; gerontology

Introduction

The 2019 Global Burden of Disease report indicates that stroke is a major global health issue, with around 12.2 million new cases and 6.6 million deaths worldwide that year. This makes stroke the second leading cause of death and the third leading cause of disability [1], with its occurrence increasing with age and affecting both males and females [2]. In Brazil, stroke is the leading cause of death, and mortality from stroke has increased in recent years. Registering 103,000 deaths in 2019, the number rose to over 112,000 by 2023 [3].

One of the most common and challenging poststroke conditions is hemiparesis, which affects a substantial number of stroke survivors globally [4,5]. For these patients, controlling trunk movement becomes a fundamental motor skill essential for performing various functional tasks [6,7]. For instance, research shows that maintaining proper posture correlates with walking ability in patients undergoing acute stroke rehabilitation [8].

In individuals without hemiparesis, posture typically involves symmetric and balanced alignment of all body parts, including the head, shoulders, spine, hips, and limbs, following the body's natural curves. In contrast, a person with hemiparesis, characterized by muscle weakness on one side of the body due to brain injuries or stroke, may exhibit an asymmetric posture, resulting in an imbalanced or tilted position. We refer to the term "best-at-the-time posture" to describe the optimal posture achievable by a hemiparetic patient, considering their motor limitations (we use the term "correct posture" interchangeably). The goal is to achieve the highest level of alignment and balance possible.

The literature registers many efforts to support stroke survivors with respect to improving trunk stability [9], trunk compensation [10], motor control [11-13], and accessing mobility [14]. Furthermore, a systematic review has confirmed the significantly positive impact of trunk training on various aspects of trunk control, sitting and standing balance, and mobility [15].

In health care, literature acknowledges the promise of smartphones [14,16] and wearables [17-20] in empowering individuals, aiding diagnosis, promoting behavior change, and enabling self-monitoring. Wearable and rehabilitation devices for different body parts, such as head, limbs, and torso, enhance training outcomes through valuable feedback [21-24]. Moreover, numerous studies have affirmed the feasibility of inertial sensors for balance and gait assessment [25-28], as corroborated by comprehensive literature reviews [29,30], even for individuals with chronic stroke [31].

In the context of wearable technology for poststroke patient support, researchers have delved into various aspects. For instance, studies demonstrated the feasibility of using step activity monitors for patients with recent stroke [32,33]. One study involved the use of intelligent insoles for analyzing the gait of patients with hemiparesis [34]. Another result is a home-based rehabilitation system using wearable sensors in smartwatches, allowing therapists to monitor patients remotely [35]. In another study, the authors evaluated the impact of haptic nudging delivered via a wrist-worn wearable device on upper limb movement during inpatient stroke rehabilitation [36]. Additionally, a recent study reported a test for assessing kinematic parameters in chronic stroke survivors. This test used a standardized mobility assessment with a simple smartphone attached to the lumbar spine using an elastic band to measure participants' kinematics [14].

Existing literature explores a range of wearable and mobile device-based solutions for postural monitoring. In the context of hemiparesis, researchers have investigated the effects of rhythmic haptic cueing on spatial and temporal gait characteristics using haptic feedback [37]. In a broader context, Smart Pose leverages a smartphone's camera, accelerometer, and magnetometer to detect poor neck posture during smartphone use, providing feedback through vibrations, text messages, and alarms [38]. Additionally, a 3-axis accelerometer biofeedback system corrected neck posture during prolonged computer use and effectively reduced inappropriate neck angles [28]. Other innovations include elastic t-shirts with embedded sensors for posture feedback [39], wearable devices for spine posture monitoring [40], and posture differentiation systems [41].

However, it is essential to recognize that effective functional rehabilitation requires both body awareness and torso control to ensure upper limb functionality [4,7,9,42]. Additionally, concerns regarding the cost and availability of treatment [43-46] highlight the need to explore alternatives that use ubiquitous, low-cost smartphones. Despite the growing number of mHealth mobile apps for patients with stroke [47], few address trunk control as a primary focus [27,48,49]. Furthermore, while both intrinsic and extrinsic feedback are crucial for motor learning after stroke [50-52], and therapists spend a significant amount of time providing this feedback [53,54], few efforts used multisensory feedback from smartphones for upper body rehabilitation [24]. Moreover, while smartphone-based multisensory feedback has proven effective for postural monitoring in healthy individuals [55], similar solutions for patients with stroke are still lacking. This highlights the need for research into new wearable technologies to enhance rehabilitation support and improve posture and trunk awareness in patients with stroke.

To address this challenge, we aimed to develop Postural SmartVest. This affordable wearable technology takes advantage of low-cost smartphone resources by (1) leveraging the built-in accelerometer sensor to monitor sagittal and frontal plane changes continuously and (2) exploiting the device's visual, tactile, and auditory feedback to guide patients in performing the movements required to return to their best posture at the time.

XSL•FO
RenderX

Methods

Overview

The research group comprises 2 occupational therapists (authors APP and VMCE) and 2 computer scientists (authors OJMN and MDGCP). To design the initial software architecture, the group received input from other health and computing professionals [56].

The study was conducted in a specialized rehabilitation center that requires both ethical committee approval and verification of ethical and safety aspects for studies conducted on its premises. Therapists participating had at least 1 year of experience in stroke rehabilitation and received training on device use before working with patients. The smartphones used in the feasibility study were either from the research team, having been tested and shown to function properly, or were participants' own devices. To minimize malfunction risks, no smartphone was allowed to be charged during use by participants. The Informed Consent Form informed participants about potential discomfort due to heat and the vest and assured them that, in the unlikely event of a smartphone malfunction causing an explosion, immediate assistance and emergency services would be provided. Data acquisition and storage were handled securely, with all data kept anonymized and stored in accounts requiring login via a secure network. Photographs were authorized with face identification removed. Participants consented to the use of data for the study and were informed of their right to withdraw at any time without coercion or obligation. Participation was voluntary and did not include any financial benefit, in compliance with Brazilian regulations.

The study comprised design and development iterations involving brainstorming meetings, interviews with therapists and patients, prototype development, feasibility studies, and experimental sessions (see workflow in Textbox 1).

Textbox 1. Research workflow.

Literature review

Iteration 1

- Brainstorm meetings
- Interview with therapists
- Prototype #1
- Feasibility study

Iteration 2

- Prototype #2
- Pilot study

Iteration 3

- Prototype #3
- Kinematic assessment
- Postural SmartVest

Within-subjects study

- Presession interviews with patients
- Experimental session
- Postsession interviews with patients
- Postsession interviews with therapists

Throughout the study, our literature review was ongoing and dynamically aligned with emerging themes. Initially, it concentrated on the specific needs of patients with stroke [4] and the effectiveness of trunk training [15]. Next, we reviewed work on novel wearable [21] and affordable smartphone-based solutions [31], including reports in literature reviews [29,57], and noticed the involvement of healthy adults in feasibility studies [26,58], as well as control participants [27]. These works are representative of the broader ongoing review process, which continued to cover these and additional relevant areas throughout the study.

We used the Google Sheets software for statistical analysis relative to the within-subject study. We used thematic analysis [59] for categorizing interview responses. For the kinematic assessment, we compared angular momentum from values calculated by Vicon and the smartphone’s sensors.

Participants

Our study involved 3 groups: poststroke participants, therapists, and healthy participants. We used interviews and questionnaires to collect requirements, demographic data, and their impressions of Postural SmartVest.

XSL•FO
RenderX

For the interviews with therapists, the inclusion criterion was having at least 1 year of experience in stroke rehabilitation. We invited professionals from a rehabilitation center using printed flyers and email invitations. In the feasibility study, we recruited participants via university email: the inclusion criteria were being 18 years or older and owning a smartphone. The exclusion criterion was self-reported physical or cognitive impairments. For the pilot and within-subjects studies, the inclusion criteria were having chronic hemiparesis, being 18 years or older, walking independently, having cognitive ability for communication, and attending rehabilitation sessions at least twice a week for a minimum of 1 month.

Iteration 1: Brainstorming Meetings, Interviews With Therapists, and Feasibility Study

For the brainstorming meetings, the research team held three 2-hour sessions to gather functional and nonfunctional requirements for the solution to develop the initial prototype comprising the vest and the app, Prototype #1 (Figure 1).

For the interviews with therapists, one team member (APP) conducted semi-structured interviews with therapists. The interview adhered to a structured protocol comprising 8 questions categorized into 4 categories: resources for posture improvement, general technology acceptance, specific technology requirements, and smartphone use (Multimedia Appendix 1: App A).

Figure 1. Prototype #1 in use in the deployment phase of the feasibility study.



For the feasibility study, we used the first version of the app. We created 8 vests and conducted sessions with 8 participants at a time. We provided smartphones to participants who could not use their own. Each session lasted 2 hours and followed a 3-phase protocol: preparation, deployment, and follow-up. In the preparation phase, therapists adjusted the vests and calibrated the seating posture for each participant.

In the deployment phase, participants set their devices to airplane mode and wore headphones to ensure privacy for their feedback. They then engaged in smartphone-free activities like reading or using their laptops (Figure 1). The deployment phase consisted of two 30-minute segments: feedback from the app was turned off during the first part and turned on during the second. The app continuously recorded the participants' postural data throughout both segments in this phase.

In the follow-up phase, participants responded to 2 questionnaires, one with scaled questions and the other with open-ended inquiries. These questionnaires aimed to assess usability and satisfaction. One questionnaire evaluated Postural SmartVest as an assistive technology for posture correction (Multimedia Appendix 1: App B), while the second assessed Postural SmartVest as an assistive technology in general (Multimedia Appendix 1: App C). For the latter questionnaire, we adapted the Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST; version 2.0), a 12-item assessment designed to measure user satisfaction with devices and services [60]. We used a translated and validated version [61] and focused on the device-related section of the questionnaire. This section included 8 scale questions that assessed size, weight, adjustability, safety, durability, ease of use, comfort, and overall effectiveness of the technology.

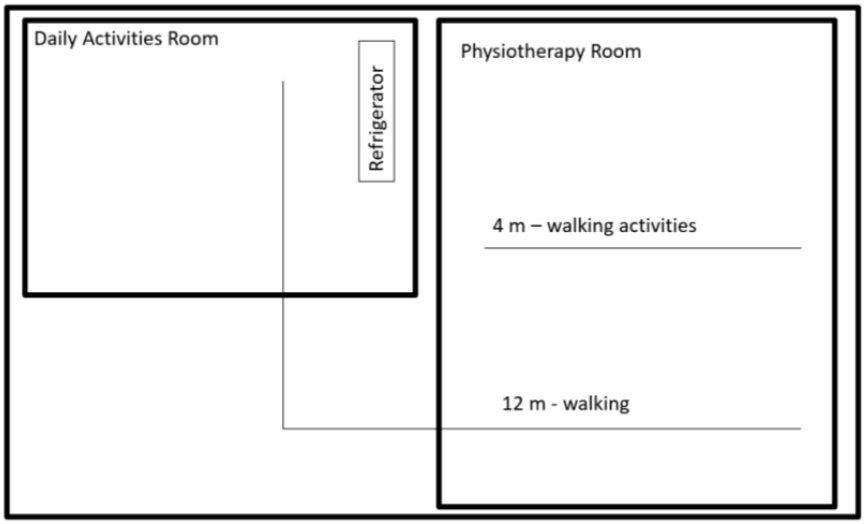
Iteration 2: Pilot Study

The feedback from the previous iteration guided the design of Prototype #2. We then used this prototype in a pilot study involving 1 patient with hemiparesis to observe the solution and detail the within-subjects study design. In the pilot study and the following studies, the smartphone used was a Motorola G with Android (version 5.1).

XSL•FO
RenderX

During the pilot study, the poststroke patient followed a protocol that included 5 ambulant-based daily activities conducted in the therapy room (Figure 2). The protocol consisted of 4 distinct phases. In the first phase, the participant walked 8 m, incorporating forward, sideways, and backward walking. In the second phase, they proceeded with a 12-m free walk to a refrigerator. In the third phase, the participant engaged in an upper limb activity that entailed opening the refrigerator, taking a glass, and returning it to the fridge. Finally, in the fourth phase, they completed a 12-m free walk back to the starting point.

Figure 2. Layout of the therapy room.



Iteration 3: Kinematic Assessment

The feedback from the pilot study informed the development of a new prototype (Prototype #3). At this stage, we checked the accelerometer readings obtained by the app running on Motorola G with Android (version 5.1).

On the one hand, sensor type and placement, activity characteristics, and population-specific conditions can influence the accuracy of sensor outcome estimates, as reported in a study that used research-grade wearable sensors on the upper arms, waist, and ankles and included poststroke patients [62]. On the other hand, the literature reports positive results regarding using state-of-the-art smartphone accelerometer technology. For instance, researchers investigated the comparative performance of 3 different commercially available smartphone accelerometers among themselves and to a gold-standard Vicon MX motion capture system, indicating that the devices are valid and reliable measuring instruments for estimating linear accelerations [63].

We assessed the angle calculations of the third prototype by comparing them with the Vicon Gait Plug-in system, a gold-standard in full-body kinematic and kinetic modeling. The Vicon system includes software and fixed cameras that capture signals from a moving target, providing visual feedback through animated vectors and plans, and it also calculates quantitative data, including distances and angles between selected planes.

Under the guidance of one occupational therapist (APP), author OJMN used the prototype to perform 6 movements: trunk flexion, trunk extension, left and right-side bending, and left and right trunk rotation (Figure 3).

The therapist calibrated the subject's optimal posture position, and for all movements, the subject began in an upright position, executed the movement, and returned to the initial position. The Vicon system and the app prototype collected data simultaneously, following an initial manual synchronization process, with the Vicon recording frames at 250 frames per second and the app prototype recording angular moments at each second along with coordinate values. To reconstruct the trunk segment, we placed markers on the clavicle, sternum, cervical vertebra C7, thoracic vertebra T10, and a point located on the medial border of the right scapula. We conducted 6 sessions using 39 markers on a full-body model (Figure 4).

Following the kinematic assessment, Prototype #3 was referred to as SmartVest and was used for the remainder of the study (Figure 5).

XSL•FO
RenderX

Figure 3. Movements used for simultaneous data collection by Vicon and the app.

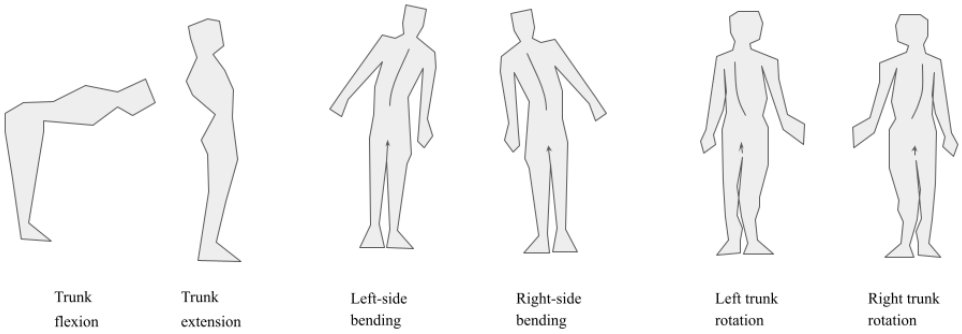


Figure 4. Simultaneous data collection by Vicon and the app.



Figure 5. The Postural SmartVest solution comprises a customizable smartphone app placed in a transparent pocket on an athletic, lightweight compression tank top.



Within-Subjects Study

The within-subjects study included individual sessions with poststroke patients, each accompanied by their therapists. Before that, therapists used the Postural SmartVest for about 30 minutes while the researchers provided guidelines to ensure the professionals understood how the solution worked. Only afterward did the therapists apply Postural SmartVest to their patients during their rehabilitation sessions. All sessions used the same smartphone (Motorola G with Android; version 5.1).

XSL•FO
RenderX

In the presession interview with each patient, we gathered information about their posture control during activities, awareness of posture issues, and any potential biases toward one side of the body. We also inquired about their strategies to correct posture problems when they noticed them (Multimedia Appendix 1: App D).

At the beginning of each session, the therapist adjusted the vest on the patient, calibrated the app to recognize the patient's best-at-the-time standing posture, and initiated the monitoring process (Textbox 2). Subsequently, the patient performed the activities defined in our pilot study twice, in consecutive sessions, each with an estimated duration of 10 minutes (Figures 6 and 7). The app did not provide feedback in the first session but registered all the movements. In the subsequent session, the app provided feedback through screen color changes (green and red), vibrations, and audio guidance.

Afterward, we asked the poststroke patients to fill out 2 questionnaires, one focusing on Postural SmartVest as an assistive technology for posture correction (Multimedia Appendix 1: App B) and the other assessing Postural SmartVest as a general assistive technology (Multimedia Appendix 1: App C).

Additionally, the therapists shared their feedback after working with the patients using a tailored version of the Postural SmartVest QUEST (version 2.0) questionnaire (Multimedia Appendix 1: App E). This study helped us understand the effectiveness of Postural SmartVest during stroke rehabilitation, both from the patient's and therapist's perspectives.

Textbox 2. Calibration workflow.

Adjust the vest on patient

Start the app

Access settings

- Enter the tolerance threshold angle for frontal movement (in degrees)
- Enter the tolerance threshold angle for lateral movement (in degrees)
- Enter the allowable duration for temporary deviations from the calibrated posture (in seconds)
- Ensure the option for vibration feedback is selected
- Ensure the option for audio feedback is selected
- Guide the patient to assume his best posture at that moment
- Press the "Calibrate" button to configure the current position as the best posture for monitoring
- Return to main screen

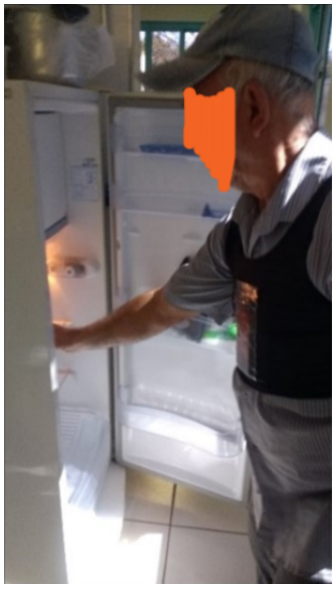
Press play or pause to initiate monitoring

Figure 6. Walking activity.



XSL•FO
RenderX

Figure 7. Upper limb activity.





Ethics Approval

The study received prior approval from the Hospital of the Clinics of the Ribeirão Preto School of Medicine at the University of São Paulo, Brazil, identifier 57234816.3.0000.5440. As detailed in the text approved by the ethics committee, the study adhered to ethical principles by ensuring beneficence through rigorous monitoring and support for participants, justice by offering equitable access and fair treatment, and respect by obtaining informed consent and addressing potential discomfort and safety concerns proactively.

Results

Participants

For the interviews, we enlisted 20 therapists: 10 occupational therapists and 10 physiotherapists. All participants were female, aged 26-40, with varying years since graduation (Table 1).

For the feasibility study, we recruited 28 healthy participants, aged 20-62, with an average age of 30.9. The group included 17 males and 11 females (Table 1). Of the participants, 25 used their own smartphones, while 3 used devices provided by us. The study covered 22 different smartphone models and 7 versions of the Android operating system.

For the pilot and within-subjects studies, we recruited poststroke participants from patients attending rehabilitation sessions at least twice a week for a month at the exact center. Inclusion criteria included chronic hemiparesis from stroke, aged 18 years or older, independent walking ability, cognitive ability for communication, and participation in rehabilitation sessions for at least a month.

We enrolled 1 patient for the pilot study and 40 patients for the within-subjects study (27 males, 67.5%), with a mean age of 57.1 (SD 11.42) years, and functional independence categorized as independent (n=39) of moderate assistance (n=1) on the functional independence measure [64,65]. Among them, 24 had right hemispheric lesions, 16 had left hemispheric lesions, and 38 had ischemic strokes. The duration of their condition ranged from 6 months to 19 years, with a majority between 1 and 4 years (Table 1).

XSL•FO
RenderX

Table 1. Participants' demographics.

Characteristics	Values
Therapists (n=20; all female), n	
Age range (years)	
26-30	8
31-40	12
Years since graduation[a]	
1-2	2
3-5	6
6-9	9
10+	3
Healthy participants (n=28; 17 male)	
Age (years)	
Range	20-62
Mean	30.9
Smartphone used, n	
Participant’s own	25
Ours	3
Poststroke patients in the within-subjects study (n=40; 27 male)	
Age range (years)	
20-50	8
51-60	16
61-70	11
71-80	5
Time since stroke (months; mean 43.9 months), n	
0-12	16
13-24	5
25-36	4
37-48	5
48+	10
Affected hemisphere	
Right	24
Left	16
Type of stroke	
Ischemic	38
Hemorrhagic	2
FIM[b]	
Moderate assistance	1
Moderate or complete independence	39

[a]Having at least 1 year of experience in stroke rehabilitation.

[b]Functional independence measure.

XSL•FO
RenderX

Iterative Design

The Postural SmartVest solution comprises a customizable smartphone app for Android devices securely placed in a transparent pocket on an athletic, lightweight compression tank top (Figure 5).

Vest Component Evolution

The development of the vest component went through several iterations (Figure 8), with 6 prototypes created based on therapist's input (see responses in Multimedia Appendix 1: App F-G) and feedback gathered during the feasibility study (see responses in the Multimedia Appendix 1: App H). The final version of the garment used in the Postural SmartVest uses an off-the-shelf lightweight compression tank top made from polyamide with elastane. We added a perforated plastic pocket to the top front. The pocket serves 3 primary purposes: securely holding the smartphone, allowing for easy headset connection, and facilitating heat dissipation.

Figure 8. Vest prototypes developed during the iterative design.



App Component Evolution

The smartphone app also underwent significant changes during the iterative design process. Initially conceptualized following brainstorming sessions that defined both functional and nonfunctional requirements (Textbox 3 and Multimedia Appendix 1: App F), the app's preliminary software architecture was informed by input from health and computing professionals [56].

Overall, the app was guided by simplicity, feasibility, and adaptability principles, in line with cost-effective assistive technology solutions from our previous research [66]. Refinements and improvements were made based on therapist interviews and feedback obtained during the feasibility study

XSL•FO
RenderX

(Textbox 4). The final version has a main screen and a sliding panel for configuration options (Figure 9).

The app records every activity along with its corresponding timestamp. This record encompasses posture changes, feedback delivery, and adjustments to application settings. During the design, we evolved the app to use schemes suggested by the OpenMHealth organization [67] when saving log data.

Textbox 3. App requirements summary.

Functional requirements

- Therapist can calibrate best-at-the-time posture with 1 touch.
- Therapist defines threshold values for small movements.
- Therapist defines threshold values to indicate how long the patient can be off best-at-the-time posture before the app starts guiding the patient.
- The app guides the patient back to their calibrated best-at-the-time posture using screen color, vibration, and audio messages.
- The app must record and store all interactions for later export and analysis.

Nonfunctional requirements

- The app should work on a low-end smartphone.
- The smartphone should be attached to a vest within a transparent pocket.
- The app should be compatible with Android, the most common operating system in Brazil.
- The app should work with the screen off to conserve battery if needed.

Textbox 4. App development iterations

Prototype #1:

- Implements all functional and nonfunctional requirements
- Features a single screen displaying x, y, and z coordinates, directions, settings, and a calibration button
- Calibration settings are not stored for future use
- Provides 5 guiding messages in Brazilian Portuguese (1 for each direction and 1 for achieving the best-at-the-time posture)

Prototype #2:

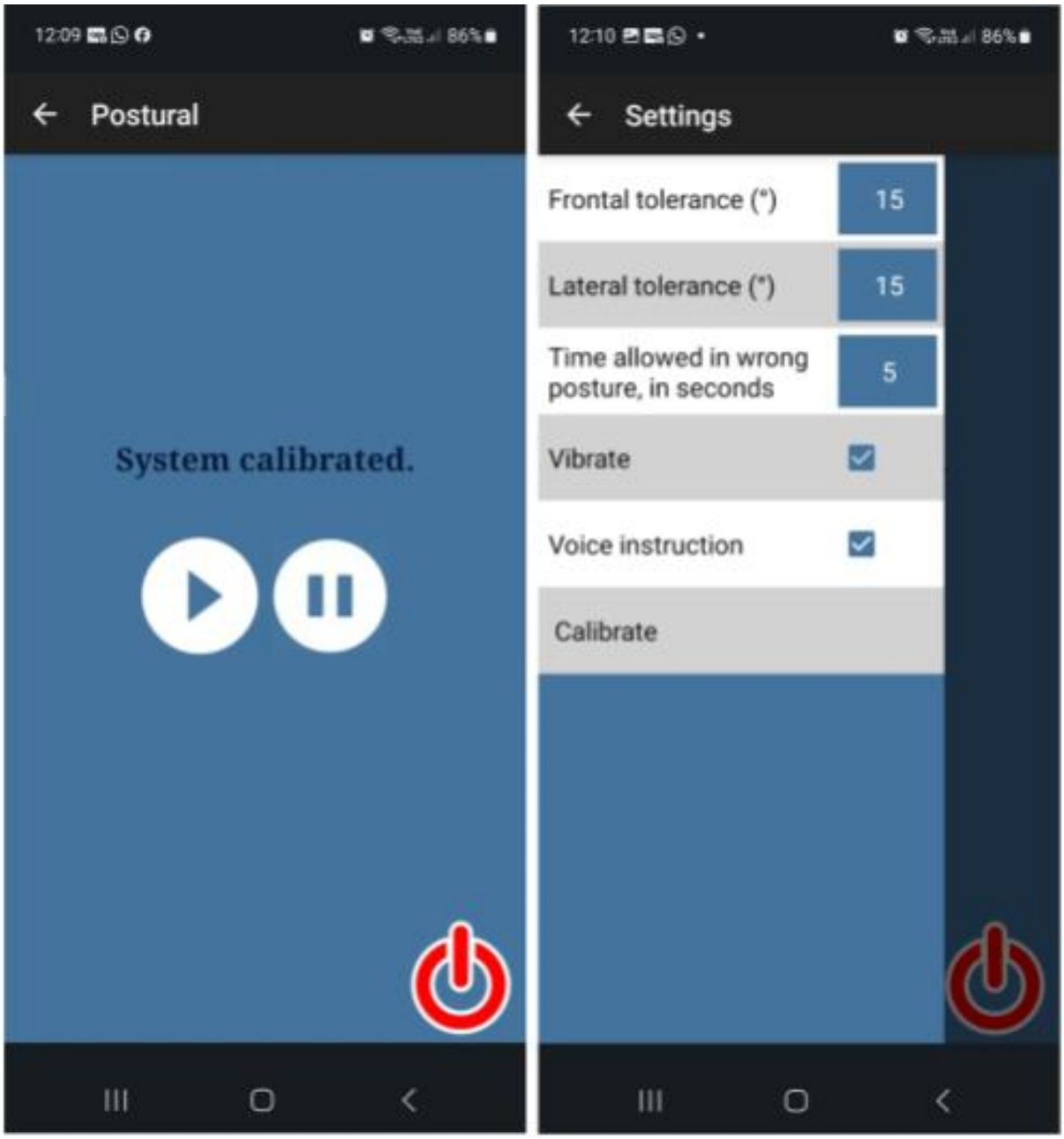
- Replaces coordinates and directions with information on sagittal and frontal planes, using values in degrees
- Introduces a new configuration screen for settings and calibration
- Stores and uses calibration settings until a new calibration is performed
- Stores all coordinate readings and feedback provided (version for kinematic assessment)
- Includes play and pause buttons for monitoring, along with an exit button on the home screen
- Expands the number of audio instructions, available in both English and Spanish
- Offers audio feedback every 5 minutes of maintaining the best-at-the-time posture

Prototype #3:

- Information on planes renamed to frontal and lateral planes
- Includes a help option with instructions on how to use the app

XSL•FO
RenderX

Figure 9. App main screen (left) and configuration panel (right).



App Component Interface

The app's main screen provides large buttons for easy access to pausing, resuming, and closing the application (Figure 9, left side). By clicking on "Postural" in the top bar on the left side of Figure 9, therapists can access 6 customization settings, as shown on the figure's right side. From top to bottom in the figure, the first 5 customization parameters allow therapists to configure thresholds for frontal and lateral movements (in degrees), to define the allowable duration for temporary deviations from the calibrated posture (in seconds), and to personalize feedback to the patient, including options for vibration and voice instructions. The therapist uses the last option, Calibrate, to set the best-at-the-time posture that the patient with hemiparesis can achieve at that moment. When calibration is activated, the app registers the current coordinates as the target posture for the patient to achieve and provides audio feedback.

After calibrating, the therapist can return to the main screen using the arrow in the top bar. From that point onward, the therapist can use the play and pause buttons to activate or deactivate posture monitoring. Additionally, we have included a large "off" button to allow the therapist to exit the application easily.

Monitoring starts when the therapist hits the play button, and the whole screen changes to green. If a patient's posture deviation exceeded the predefined thresholds for lateral and frontal angles relative to the calibrated posture, the app responded immediately with visual feedback, transitioning the screen color from green to red, and tactile feedback through vibrations. Furthermore, the app introduced audio feedback if a patient deviated from the best-at-the-time posture beyond the

XSL•FO
RenderX

allowed time for temporary deviations from the calibrated posture. Additionally, the app gives positive feedback every 5 minutes if the patient is maintaining the best-at-the-time posture. This comprehensive feedback system provided patients with real-time guidance tailored to their unique needs and specific postural challenges.

The app provides audio instructions: “please lean forward,” “please lean backward,” “please lean to the right,” and “please lean to the left.” We implemented enhancements throughout the iterations, such as adding positive feedback messages to encourage users to maintain the best-at-the-time posture and random variations in the messages that congratulate users for returning to their best-at-the-time posture.

To meet a requirement identified during the feasibility study, where participants used devices set to various languages, we improved the app to support users in English, Spanish, and Portuguese. It dynamically adapts its interface and voice instructions based on the language settings of the user’s device.

Kinematic Assessment

The kinematic assessment results, which involved comparing angular momentum values calculated by the Vicon system with those obtained from the smartphone's sensors, are presented graphically in the Multimedia Appendix 1: App I. The results consistently demonstrated that smartphone accelerometers offer reliable data readings, and our mobile application accurately interprets these values. These outcomes align with existing literature [24,28,38,40,58], reinforcing the utility of our approach within the context of our application. Moreover, this positive evaluation of the system's trustworthiness is in line with the feedback received from therapists regarding its performance and reliability, as detailed in the section *Postsession Interviews With Therapists*.

Pilot Study

The pilot study confirmed the proposed activity protocol for the within-subjects study, with participants safely completing the walking circuit within the estimated time frame.

Within-Subjects Study

Presession Interviews With Patients

Findings are summarized in Table 2. Out of the 40 participants with hemiparesis, 34 (85%) acknowledged experiencing challenges in balancing one side of their body. These participants reported various triggers for posture correction, including pain, the sensation of weight, and fatigue. Specifically, 6 participants stated that they only adjusted their posture when they felt pain, 3 did so in response to a sense of heaviness on the side affected by hemiparesis, and 2 did it to alleviate fatigue. Participants mentioned that they recognized the need for correction only when viewing photos (2 participants), standing in front of a mirror (2 participants), or following verbal advice from a family member (6 participants). In total, 4 participants admitted that they often neglected to correct their posture due to forgetfulness, with one of them explaining, “It is hard to remember, and correcting myself all day is tiring.”

When providing opinions on assistive technologies, participants considered the technologies “viable for use” and noted that they provide “progress to believe in improvement.” They also mentioned that these technologies “make rehabilitation more precise” and “help more than just me and the therapist.” Additionally, it was observed that they “stimulate people's lives even with minimal gain.” The repetition of these responses highlights the consistent perception of the technologies' value and potential impact.

Regarding the necessary requirements for the technology, 11 participants agreed that it could assist in the perception and control of the trunk. They envisioned the technology as a “chair that helps maintain correct posture” or “something that allows for posture visualization.” Some participants suggested a device that “corrects the arm and leg to prevent them from falling” and another that “reminds me to correct my posture.” They also proposed features such as “notifying to stay straight” or having the ability to “pull to the side.” Other suggestions included a “vest to correct posture” or a device that “attaches to the trunk and holds it in place.” Additional requirements included providing a “signal on the body,” “holding the shoulders, reminding to keep the back straight, and being fixed on the bra,” and being “close to the body,” “comfortable, practical, and unobtrusive,” and “adaptable to the body.” Participants emphasized that the technology should be “integrated with therapy,” “usable over time,” and “inexpensive.”

XSL•FO
RenderX

Table 2. Patients' report on posture, assistive technology, and smartphone usage.

Category	Patients, n
Maintenance of correct posture	
Unable to maintain posture	22
Able to maintain posture	17
Able to maintain posture sitting, but not standing	1
Perception of correct posture	
Perceive correct posture	36
Do not perceive correct posture	4
Difficulties and postural correction	
Remembering to correct and trying to adjust	8
Unable to correct posture	6
Need to leave sitting posture to correct	1
Triggers for posture correction	
After perceiving pain	6
After request from family	6
After perceiving weight on the side of hemiparesis	2
After feeling tired	2
After observing photos	2
After seeing oneself in the mirror	2
Opinion on assistive technology	
Believe it can help but consider it expensive	13
Knowledge from TV but no contact	2
Should be offered by the public health system	1
Would be good if recommended by professionals	1
Smartphone usage	
Participants with smartphones	26
Made accessibility adjustments	6
Experience difficulties	9
Use daily	19
Use sporadically	2
Use only for calls	4
Use only for WhatsApp, music, and camera	1
Other uses (WhatsApp, Facebook, YouTube, and internet banking)	12

Session Data

The average duration of the stages with and without feedback was 9.14 (SD 5.36) minutes and 8.52 (SD 3.21) minutes, respectively. A 2-tailed *t*-test indicated that these differences were not statistically significant.

We compared the number of the best-at-the-time posture positions patients could maintain in the 2 stages: without feedback and with feedback. The 2-tailed *t*-test results indicated a statistically significant difference between these 2 conditions. Specifically, in the stage with feedback, patients exhibited a higher ability to maintain correct positions (mean 13.1, SD 7.12) compared to the stage without feedback (mean 4.2, SD 3.97). This significant difference ($P<.001$) underscores that patients could achieve and sustain their best-at-the-time postures more frequently when Postural SmartVest provided feedback.

Additionally, we examined the average number of movements performed by patients to attain their best-at-the-time posture in both the stages without feedback and with feedback. In the stage without feedback, patients averaged 8.09 (SD 6.88) movements. In contrast, in the stage with feedback, they averaged 4.49 (SD 1.49) movements. The 2-tailed *t*-test results indicated no statistically significant difference between the values obtained in these 2 stages, with a *P* value of .11.

XSL•FO
RenderX

Postsession Interviews With Poststroke Patients

In the evaluation of satisfaction with Postural SmartVest, 37 (92.5%) patients reported being satisfied or very satisfied with the support provided by Postural SmartVest to help them maintain proper posture. In total, 2 respondents had neutral feelings, and 1 expressed dissatisfaction with the solution. In total, 39 (97.5%) patients said they would recommend the solution to others with hemiparesis due to stroke, and one indicated that they might recommend it.

When asked if they were able to maintain their correct posture, 17 patients answered “yes,” 2 responded with “sometimes,” and 21 answered “no.” When asked if they noticed that their body often tilted to a particular side, 18 reported that their body tilted to the right side, 16 indicated that their body tilted to the left, 3 stated that their body tilted backward, and 3 could not respond.

Responses to the questionnaire evaluating Postural SmartVest as an assistive technology using a 1-5 scale (Multimedia Appendix 1: App E) are summarized in Table 3. Regarding open-ended questions directed to the patients, we highlight in Table 4 some critical opinions and the number of users who expressed them.

Poststroke patients provided rich feedback on Postural SmartVest’s usability and effectiveness (Table 4). They suggested design improvements, such as incorporating openings with zippers and using breathable materials for cell phone accommodation. Some recommended extended usage periods, while others emphasized early implementation for gradual posture improvement. Software suggestions included customizable message frequency and varied vibration patterns. Participants highlighted the positive impact of audio feedback, likening it to therapist interactions during sessions. These insights, shared by a significant portion of the 40 patients, showcase the multifaceted utility of Postural SmartVest and its potential to address various poststroke rehabilitation needs. The integration of smartphones into thoracic clothing proved safe during our study, with no observed risks during continuous 1-hour use.

Table 3. Patients’ evaluation of postural SmartVest as an assistive technology.

Aspect	Favorable responses, %
Weight	100
Comfort	100
Dimension	97.5
Effectiveness	97.5
Ease of Use	87.5
Stability	82.5
Durability	72.5
Ease of Adjustment	67.5

Table 4. Patients’ responses to open-ended questions. The last column indicates the number of participants reporting.

	Responses, n
“I would suggest improvements to the garment, such as an opening Velcro or zipper and the use of breathable material to accommodate the cell phone.”	10
“The vest should be more discreet so that it could be worn more imperceptibly under the shirt.”	10
“The garment should be worn for a longer period, over one hour.”	4
“The solution should be introduced to me as early as the first rehabilitation session, immediately after a stroke, so I could get used to perceiving and correcting my posture earlier.”	4
“I would like to adjust the audio volume of the messages through the app.”	3
“The frequency of the messages should be customizable.”	3
“I would like different vibration patterns for different body movements.”	3
“Listening to the app's directions reminds me of you (the therapist) talking to me during the rehabilitation sessions.”	1
“I would use the solution to prevent falls.”	1
“The solution helps me walk.”	1
“The solution helps me improve my motor perception.”	1
“I believe that the vest alone can play an important role in postural perception.”	1
“I think that using the solution after some rehabilitation sessions is better because right after the stroke, I could not minimally perceive or correct my posture.”	1

XSL•FO
RenderX

Postsession Interviews With Therapists

Responses to the adapted QUEST (version 2.0) questionnaire evaluating the solution as an assistive technology (Multimedia Appendix 1: App E) are summarized in Table 5. In total, 2 therapists provided a neutral rating for "Support to Posture" and "Audio guidance accuracy." They explained that patients with spatial orientation difficulties still required assistance from therapists despite using the solution. In total, 5 therapists provided neutral responses regarding "Solution trustworthiness." In total, 2 therapists recommended that the app recognize other body parts and suggested improvements in the instructions, noting that the app reinforced the user's reference sides.

We summarize the responses to evaluating satisfaction with Postural SmartVest on a 1-10 scale (Multimedia Appendix 1: App E) in Table 6. The responses indicated a highly positive impact on patients' posture control, with 95% rating it as "A lot." Additionally, 90% of the therapists believed that the app significantly contributed to achieving therapy goals. Moreover, 60% reported conducting fewer patient posture-related interventions after using the app. In terms of ease of use, 70% of the therapists found it "Very easy" for patients to follow the app's guidance.

Table 5. Therapists' responses to the adapted QUEST (version 2.0) questionnaire (1-5 scale).

Response	Rating 4 or higher, %	Indication
Support to Posture	90	Very satisfied
Audio guidance accuracy	90	Very satisfied
Device comfort	95	Very comfortable
Solution trustworthiness	83	Very reliable
Safety during use	100	Very safe
Likelihood to recommend	100	Very likely

Table 6. Therapists' satisfaction with SmartVest (1-10 scale).

Question	Percentage: answer	Indicating
Did you notice a positive impact on patients' posture control?	95%: 8 or higher	A lot
Do you think the app contributed to the therapy goals?	90%: 8 or higher	A lot
Did you notice if you conducted more or less posture-related interventions for patients?	60%: 3 or less	Fewer
How difficult was it for the patients to follow the app guidance?	70%: 3 or less	Very easy

Discussion

Principal Results

We designed Postural SmartVest, a solution consisting of a chest garment holding a smartphone running an Android application. When the user wears the Postural SmartVest in a rehabilitation session, the therapist first calibrates the app to the best posture the patient can achieve at the time. The app uses the smartphone accelerometer to identify variations in the user's body movement in the sagittal and frontal planes relative to the calibrated position. Based on these variations, the app provides multisensorial feedback that guides the patient to perform the movements required to return to the calibrated position.

The iterative development of Postural SmartVest, informed by therapist interviews and feasibility study feedback, tailored the final product to user needs. The app reliably identifies posture changes and provides corresponding multisensorial feedback, as substantiated by kinematic assessment.

Postural SmartVest significantly improved patients' ability to maintain their best-at-the-time postures, with patients showing a higher capacity for correct positions when using feedback. Although the analysis found no significant difference in the number of movements between stages, the increased movement during the feedback stage suggests an active engagement with the system for postural adjustments. This engagement aligns with rehabilitation goals, emphasizing motor skill enhancement and proprioception. These findings highlight Postural SmartVest's potential as an assistive technology for poststroke patients, emphasizing the importance of feedback in rehabilitation tools.

Postural SmartVest's flexibility meets each patient's unique needs and progress, with a highly customizable app enabling therapists to personalize feedback and adjust thresholds. This multisensory approach enhances engagement and may result in more effective posture correction. The app offers a comprehensive feedback system, including visual cues, tactile vibrations, and audio instructions.

The technology received positive feedback regarding its design and usability, with patients and therapists giving favorable ratings for various aspects of Postural SmartVest, including weight, comfort, dimensions, and effectiveness. Patients believed in its utility for walking assistance and motor perception improvement, and their suggestions aligned with their approval. Many patients were willing to recommend Postural SmartVest to other patients with poststroke hemiparesis, highlighting its perceived value. Healthy users also expressed satisfaction with Postural SmartVest's posture support.

XSL•FO
RenderX

Limitations

Our study used a specific sample of poststroke patients, which may not fully represent the diverse stroke survivor population. This highlights the need for including a broader and more diverse sample in future research. Although the findings indicate high satisfaction among both patients and therapists, individual preferences and responses varied. Patient feedback identified several limitations that suggest potential improvements to the app's personalization features, such as the ability to adjust the audio volume while using the vest and the inclusion of customizable vibration patterns and prompt frequencies. Additionally, enabling therapists to record guidance prompts in their own voices could enhance the patient experience, as one patient noted: “Listening to the app's directions reminds me of you (the therapist) talking to me during the rehabilitation sessions.”

The study’s duration might have limited the assessment of long-term effects, as improvements in posture maintenance could potentially change over time. A longer follow-up period could provide insights into whether the benefits observed are sustained in the long term. Additionally, the self-controlled approach used in the study—where participants served as their own controls by using the app with feedback turned off—may introduce bias related to user expectations and behavior. Future research should consider alternative control group designs to minimize potential biases and enhance the validity of the results. Another limitation of this study is the relatively unstructured approach to coding and scoring responses from interviews and feedback sessions. The multidimensional nature of the questions, ranging from levels of agreement to broader attitudes and preferences, presented challenges in achieving uniform analysis. To address this, we plan to use more structured, software-assisted procedures in future research to standardize data collection and analysis [68,69], thereby improving the consistency and depth of understanding gained from participant feedback.

Future research should also explore categorizing participants based on injury time, dominance, or body laterality to better understand the varying impacts of the Postural SmartVest. Including studies that involve patients using the device at home could provide insights into its real-world effectiveness and long-term impact in poststroke rehabilitation outside of clinical settings. Understanding how patients integrate the Postural SmartVest into their daily routines and the effects of extended, unsupervised use could reveal valuable information about its benefits in home-based rehabilitation. Furthermore, while Postural SmartVest records a range of activities, including posture changes and feedback delivery, it currently lacks features for therapists to access and use this data effectively. Future research should focus on developing these functionalities to enhance patient monitoring and rehabilitation planning. Additionally, addressing the requirements identified during initial therapist interviews, such as including a photo and video gallery and a dashboard for progress monitoring, could improve Postural SmartVest’s functionality and user satisfaction.

Comparison With Prior Work

Postural SmartVest is a modified athletic compression tank top with a transparent pocket for a smartphone running a customizable Android app. This app offers tactile vibrations, dynamic visual feedback, and clear audio instructions for posture monitoring and guidance. It is designed to be minimally intrusive, replicable, and useful for both therapists and poststroke patients, using readily available Android smartphones and avoiding the need for expensive specialized equipment.

The development of Postural SmartVest reflects the evolution of cognitive rehabilitation from classical methods to personalized systems with kinematic analysis and user-centered design [70]. It incorporates feedback from patients with stroke and therapists, aligning with recommendations for wearable technology design [71]. Unlike wearable rehabilitation systems that prioritize sensor accuracy over comfort [24], Postural SmartVest is distinguished by its comfort and effective smartphone placement, addressing both functional and user experience aspects.

Research on user perspectives for wearable systems in upper extremity stroke rehabilitation highlights the importance of monitoring both upper extremity and trunk movements and evaluating their quality and quantity [39]. In our experimental sessions, we included activities such as position maintenance, lateral trunk movements, trunk rotation, forward-backward trunk motions, and object retrieval, aligning with tasks assessed by various trunk ability evaluation tools [72]. While some studies have investigated smartphone-based training to improve balance and trunk performance among seated patients with stroke—using modified balance boards with visual and audio feedback in an external monitor [48,49] or smartphones attached to harnesses for screen feedback [25]—Postural SmartVest supports dynamic ambulatory movements and provides visual feedback directly on the smartphone screen, using a single color (red or green) to indicate posture status. This approach is particularly effective in therapy rooms with mirror walls and accommodates individuals with visual deficiencies, which are common after a stroke [73].

Postural SmartVest uses the smartphone’s accelerometer on the patient’s trunk to assist with trunk posture, while many studies use sensors on different body parts for various purposes [74-77]. While postural SmartVest uses smartphone-generated sensor data to identify trunk postures and facilitate posture correction, existing literature often combines data from multiple sensors placed on the user’s body for posture detection [78].

A wealth of research envisions using their solutions to foster collaboration among health professionals [79] or in the patient’s home environment, including studies that use smartphones as a communication component in a remote rehabilitation platform [18,25,33,35,70,76,80,81]. Although we designed Postural SmartVest primarily to assist patients during rehabilitation sessions, we are actively extending our research to the home environment by integrating Postural SmartVest with our experience sampling and programmed intervention platform [68].

XSL•FO
RenderX

Conclusions

Our goal was to create an affordable wearable, Postural SmartVest, aiding poststroke patients in posture maintenance during rehabilitation. Its simplicity and cost-effectiveness, using widely available devices with efficient sensors, enable multisensory feedback, enhancing accessibility. Patients and therapists expressed utility and satisfaction, emphasizing its clinical potential.

A comparison of patient session data using Postural SmartVest, both with and without feedback, informed essential insights. The significant difference in maintaining correct positions underscored the value of feedback. Although there was no disparity in the number of movements, the increased adjustments during sessions with feedback suggested active patient engagement. These findings emphasize the potential of Postural SmartVest as a valuable assistive technology for poststroke patients and highlight the importance of integrating feedback into rehabilitation tools.

Our study indicates several future research directions, including studying diverse stroke survivor populations, assessing long-term effects, exploring user engagement factors, and evaluating therapist training impact. Additionally, we should prioritize functional outcomes, technology adoption, cost-effectiveness, and addressing unmet needs like enhancing posture visualization and aiding long-term posture monitoring.

Acknowledgments

OJMN received support from the São Paulo Research Foundation (FAPESP; grant 14/01831-6). MDGCP was funded by FAPESP (grants 16/00351-6 and 16/50489-4) and CNPq (National Council for Scientific and Technological Development; grant 312058/2015-2). The authors also acknowledge the support of FAPESP for funding the publication of this manuscript (grant no. 24/15881-7). The Coordination for the Improvement of Higher Education Personnel (CAPES) contributed to this study (finance code 001). We deeply thank all participants in the interviews, feasibility study, and within-subject experiment for their participation. We thank the support from the Innovation and Entrepreneurship Laboratory in Assistive Technology, Sports, and Health at the Federal University of São Carlos for the Vicon evaluation.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional material.
[DOCX File , 499 KB - aging_v8i1e55455_app1.docx]

References

1. GBD 2019 Stroke Collaborators. Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Neurol 2021 Oct;20(10):795-820 [FREE Full text] [doi: 10.1016/S1474-4422(21)00252-0] [Medline: 34487721]
2. Rajati F, Rajati M, Rasulehvandi R, Kazeminia M. Prevalence of stroke in the elderly: a systematic review and meta-analysis. Interdiscip Neurosurg 2023;32:101746. [doi: 10.1016/j.inat.2023.101746]
3. Miranda M. Números do AVC. URL: https://avc.org.br/numeros-do-avc/ [accessed 2024-12-07]
4. Hsieh CL, Sheu CF, Hsueh IP, Wang CH. Trunk control as an early predictor of comprehensive activities of daily living function in stroke patients. Stroke 2002;33(11):2626-2630. [doi: 10.1161/01.str.0000033930.05931.93] [Medline: 12411652]
5. Lin SH, Dionne TP. Interventions to improve movement and functional outcomes in adult stroke rehabilitation: review and evidence summary. J Particip Med 2018;10(1):e3 [FREE Full text] [doi: 10.2196/jopm.8929] [Medline: 33052128]
6. Ryerson S, Byl NN, Brown DA, Wong RA, Hidler JM. Altered trunk position sense and its relation to balance functions in people post-stroke. J Neurol Phys Ther 2008;32(1):14-20. [doi: 10.1097/NPT.0b013e3181660f0c] [Medline: 18463551]
7. Pellicciari L, Sodero A, Campagnini S, Guolo E, Basagni B, Castagnoli C, et al. Factors influencing trunk control recovery after intensive rehabilitation in post-stroke patients: a multicentre prospective study. Top Stroke Rehabil 2023;30(2):109-118. [doi: 10.1080/10749357.2021.2016099] [Medline: 34994302]
8. Wang C, Chen Y, Wang C. Postural maintenance is associated with walking ability in people receiving acute rehabilitation after a stroke. Phys Ther 2022;102(4):pzab309. [doi: 10.1093/ptj/pzab309] [Medline: 35079798]
9. Yang S, Chung E, Lee J, Lee S, Lee B. The effect of trunk stability training based on visual feedback on trunk stability, balance, and upper limb function in stroke patients: a randomized control trial. Healthcare (Basel) 2021;9(5):532 [FREE Full text] [doi: 10.3390/healthcare9050532] [Medline: 34063280]
10. Valdés BA, Schneider AN, Van der Loos HFM. Reducing trunk compensation in stroke survivors: a randomized crossover trial comparing visual and force feedback modalities. Arch Phys Med Rehabil 2017;98(10):1932-1940. [doi: 10.1016/j.apmr.2017.03.034] [Medline: 28526482]

XSL•FO
RenderX

11. Appleby E, Gill ST, Hayes LK, Walker TL, Walsh M, Kumar S. Effectiveness of telerehabilitation in the management of adults with stroke: a systematic review. PLoS One 2019;14(11):e0225150. [doi: 10.1371/journal.pone.0225150] [Medline: 31714924]
12. Chan A, Cohen R, Robinson K, Bhardwaj D, Gregson G, Jutai JW, et al. Evidence and user considerations of home health monitoring for older adults: scoping review. JMIR Aging 2022;5(4):e40079. [doi: 10.2196/40079] [Medline: 36441572]
13. Muroi D, Saito Y, Koyake A, Hiroi Y, Higuchi T. Training for walking through an opening improves collision avoidance behavior in subacute patients with stroke: a randomized controlled trial. Disabil Rehabil 2023;46(5):1-9. [doi: 10.1080/09638288.2023.2181412] [Medline: 36815267]
14. Sánchez-Sánchez ML, Ruescas-Nicolau M, Arnal-Gómez A, Iosa M, Pérez-Alenda S, Cortés-Amador S. Validity of an android device for assessing mobility in people with chronic stroke and hemiparesis: a cross-sectional study. J Neuroeng Rehabil 2024;21(1):54 [FREE Full text] [doi: 10.1186/s12984-024-01346-5] [Medline: 38616288]
15. Van Criekinge T, Truijen S, Schröder J, Maebe Z, Blanckaert K, van der Waal C, et al. The effectiveness of trunk training on trunk control, sitting and standing balance and mobility post-stroke: a systematic review and meta-analysis. Clin Rehabil 2019;33(6):992-1002. [doi: 10.1177/0269215519830159] [Medline: 30791703]
16. Bonura A, Motolese F, Capone F, Iaccarino G, Alessiani M, Ferrante M, et al. Smartphone app in stroke management: a narrative updated review. J Stroke 2023;25(2):320-324 [FREE Full text] [doi: 10.5853/jos.2022.01410.e1] [Medline: 37282380]
17. Chiew TK, Khoo S, Ansari P, Kiruthika U. Mobile and wearable technologies for persons with disabilities: a bibliometric analysis (2000-2021). Disabil Rehabil Assist Technol 2024;19(3):994-1002. [doi: 10.1080/17483107.2022.2138996] [Medline: 36306172]
18. Lee BC, An J, Kim J, Lai EC. Performing dynamic weight-shifting balance exercises with a smartphone-based wearable telerehabilitation system for home use by individuals with parkinson's disease: a proof-of-concept study. IEEE Trans Neural Syst Rehabil Eng 2023;31:456-463 [FREE Full text] [doi: 10.1109/TNSRE.2022.3226368] [Medline: 36455080]
19. Nascimento LMSD, Bonfati LV, Freitas MLB, Mendes Junior JJA, Siqueira HV, Stevan SL. Sensors and systems for physical rehabilitation and health monitoring-a review. Sensors (Basel) 2020;20(15):4063 [FREE Full text] [doi: 10.3390/s20154063] [Medline: 32707749]
20. Rodgers MM, Alon G, Pai VM, Conroy RS. Wearable technologies for active living and rehabilitation: current research challenges and future opportunities. J Rehabil Assist Technol Eng 2019;6:2055668319839607 [FREE Full text] [doi: 10.1177/2055668319839607] [Medline: 31245033]
21. Ding Z, Luo Z, Causo A, Chen I, Yue K, Yeo S, et al. Inertia sensor-based guidance system for upperlimb posture correction. Med Eng Phys 2013;35(2):269-276. [doi: 10.1016/j.medengphy.2011.09.002] [Medline: 21978912]
22. Lu L, Zhang J, Xie Y, Gao F, Xu S, Wu X, et al. Wearable health devices in health care: narrative systematic review. JMIR Mhealth Uhealth 2020;8(11):e18907 [FREE Full text] [doi: 10.2196/18907] [Medline: 33164904]
23. Kim GJ, Parnandi A, Eva S, Schambra H. The use of wearable sensors to assess and treat the upper extremity after stroke: a scoping review. Disabil Rehabil 2022;44(20):6119-6138 [FREE Full text] [doi: 10.1080/09638288.2021.1957027] [Medline: 34328803]
24. Wang Q, Markopoulos P, Yu B, Chen W, Timmermans A. Interactive wearable systems for upper body rehabilitation: a systematic review. J Neuroeng Rehabil 2017;14(1):20 [FREE Full text] [doi: 10.1186/s12984-017-0229-y] [Medline: 28284228]
25. Aphiphaksakul P, Siriphorn A. Home-based exercise using balance disc and smartphone inclinometer application improves balance and activity of daily living in individuals with stroke: a randomized controlled trial. PLoS One 2022;17(11):e0277870 [FREE Full text] [doi: 10.1371/journal.pone.0277870] [Medline: 36409753]
26. Hou Y, Chiu Y, Chiang S, Chen H, Sung W. Development of a smartphone-based balance assessment system for subjects with stroke. Sensors (Basel) 2019;20(1):88 [FREE Full text] [doi: 10.3390/s20010088] [Medline: 31877843]
27. Afzal MR, Byun H, Oh M, Yoon J. Effects of kinesthetic haptic feedback on standing stability of young healthy subjects and stroke patients. J Neuroeng Rehabil 2015;12:27 [FREE Full text] [doi: 10.1186/s12984-015-0020-x] [Medline: 25889581]
28. Shirvani H, Salesi M, Samadi M, Shamsoddini A. Design and development of a 3-axis accelerometer biofeedback system for real-time correction of neck posture for long-time computer usersdesign and development of a 3-axis accelerometer biofeedback system for real-time correction of neck posture for long-time computer users. J Med Signals Sens 2021;11(4):269-273 [FREE Full text] [doi: 10.4103/jmss.JMSS_56_20] [Medline: 34820299]
29. Zampogna A, Mileti I, Palermo E, Celletti C, Paoloni M, Manoni A, et al. Fifteen years of wireless sensors for balance assessment in neurological disorders. Sensors (Basel) 2020;20(11):3247 [FREE Full text] [doi: 10.3390/s20113247] [Medline: 32517315]
30. Peters J, Abou L, Wong E, Dossou MS, Sosnoff JJ, Rice LA. Smartphone-based gait and balance assessment in survivors of stroke: a systematic review. Disabil Rehabil Assist Technol 2024;19(1):177-187. [doi: 10.1080/17483107.2022.2072527] [Medline: 35584288]
31. Hou YR, Chiu YL, Chiang SL, Chen HY, Sung WH. Feasibility of a smartphone-based balance assessment system for subjects with chronic stroke. Comput Methods Programs Biomed 2018;161:191-195. [doi: 10.1016/j.cmpb.2018.04.027] [Medline: 29852961]

XSL•FO
RenderX

32. Katzan I, Schuster A, Kinzy T. Physical activity monitoring using a fitbit device in ischemic stroke patients: prospective cohort feasibility study. JMIR Mhealth Uhealth 2021;9(1):e14494 [FREE Full text] [doi: 10.2196/14494] [Medline: 33464213]
33. Wu JM, Chen HS, Chen HH, Cheng BW, Huang CW, Chung MH. Enhancing patient self-management after a first stroke: an application of the wearable devices and the health management platform. Disabil Health J 2023;16(1):101392 [FREE Full text] [doi: 10.1016/j.dhjo.2022.101392] [Medline: 36333265]
34. Seo M, Shin M, Park TS, Park J. Clinometric gait analysis using smart insoles in patients with hemiplegia after stroke: pilot study. JMIR Mhealth Uhealth 2020;8(9):e22208 [FREE Full text] [doi: 10.2196/22208] [Medline: 32909949]
35. Chae SH, Kim Y, Lee K, Park H. Development and clinical evaluation of a web-based upper limb home rehabilitation system using a smartwatch and machine learning model for chronic stroke survivors: prospective comparative study. JMIR Mhealth Uhealth 2020;8(7):e17216 [FREE Full text] [doi: 10.2196/17216] [Medline: 32480361]
36. Signal NEJ, McLaren R, Rashid U, Vandal A, King M, Almesfer F, et al. Haptic nudges increase affected upper limb movement during inpatient stroke rehabilitation: multiple-period randomized crossover study. JMIR Mhealth Uhealth 2020;8(7):e17036 [FREE Full text] [doi: 10.2196/17036] [Medline: 32723718]
37. Georgiou T, Holland S, van der Linden J. Rhythmic haptic cueing for gait rehabilitation of people with hemiparesis: quantitative gait study. JMIR Biomed Eng 2020;5(1):e18649. [doi: 10.2196/18649]
38. Shim E, Lee S, Lee H, Choi YS. Smart pose: mobile posture-aware system for lowering physical health risk of smartphone users. : Association for Computing Machinery; 2013 Presented at: CHI '13 Extended Abstracts on Human Factors in Computing Systems; 2013 April 27; Paris, France p. 2257-2266 URL: https://doi.org/10.1145/2468356.2468747
39. Wang Q. Designing posture monitoring garments to support rehabilitation. 2016 Presented at: TEI '16: Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction; 2016 February 14; Eindhoven, Netherlands p. 709-712 URL: https://dl.acm.org/doi/10.1145/2839462.2854106
40. Ribeiro P, Soares AR, Girão R, Neto M, Cardoso S. Spine cop: posture correction monitor and assistant. Sensors (Basel) 2020;20(18):5376 [FREE Full text] [doi: 10.3390/s20185376] [Medline: 32961772]
41. Muşat EC, Borz SA. Learning from acceleration data to differentiate the posture, dynamic and static work of the back: an experimental setup. Healthcare (Basel) 2022;10(5):916 [FREE Full text] [doi: 10.3390/healthcare10050916] [Medline: 35628053]
42. Yang CL, Chui R, Mortenson WB, Servati P, Servati A, Tashakori A, et al. Perspectives of users for a future interactive wearable system for upper extremity rehabilitation following stroke: a qualitative study. J Neuroeng Rehabil 2023;20(1):77 [FREE Full text] [doi: 10.1186/s12984-023-01197-6] [Medline: 37312189]
43. Soares CL, Magalhães JP, Faria-Fortini I, Batista LR, Andrea Oliveira Lima L, Faria CD. Barriers and facilitators to access post-stroke rehabilitation services in the first six months of recovery in Brazil. Disabil Rehabil 2024;46(23):5601-5607. [doi: 10.1080/09638288.2024.2310756] [Medline: 38299553]
44. Nepal G, Yadav JK, Bhandari S, Gautam J, Gajurel BP. Low-cost alternatives for the management of acute ischemic stroke in low and middle-income countries. Ann Med Surg (Lond) 2021;72:102969 [FREE Full text] [doi: 10.1016/j.amsu.2021.102969] [Medline: 34992776]
45. Ariza J, Pearce JM. Low-cost assistive technologies for disabled people using open-source hardware and software: a systematic literature review. IEEE Access 2022;10:124894-124927. [doi: 10.1109/access.2022.3221449]
46. Negrini F, Gasperini G, Guanziroli E, Vitale JA, Banfi G, Molteni F. Using an accelerometer-based step counter in post-stroke patients: validation of a low-cost tool. Int J Environ Res Public Health 2020;17(9):3177 [FREE Full text] [doi: 10.3390/ijerph17093177] [Medline: 32370210]
47. Burns SP, Terblanche M, Perea J, Lillard H, DeLaPena C, Grinage N, et al. mHealth intervention applications for adults living with the effects of stroke: a scoping review. Arch Rehabil Res Clin Transl 2021;3(1):100095 [FREE Full text] [doi: 10.1016/j.arrct.2020.100095] [Medline: 33778470]
48. Shin DC, Song CH. Smartphone-based visual feedback trunk control training using a gyroscope and mirroring technology for stroke patients: single-blinded, randomized clinical trial of efficacy and feasibility. Am J Phys Med Rehabil 2016;95(5):319-329. [doi: 10.1097/PHM.0000000000000447] [Medline: 26829087]
49. Chung BPH, Chiang WKH, Lau H, Lau TFO, Lai CWK, Sit CSY, et al. Pilot study on comparisons between the effectiveness of mobile video-guided and paper-based home exercise programs on improving exercise adherence, self-efficacy for exercise and functional outcomes of patients with stroke with 3-month follow-up: a single-blind randomized controlled trial. Hong Kong Physiother J 2020;40(1):63-73 [FREE Full text] [doi: 10.1142/S1013702520500079] [Medline: 32489241]
50. van Vliet PM, Wulf G. Extrinsic feedback for motor learning after stroke: what is the evidence? Disabil Rehabil 2006;28(13-14):831-840. [doi: 10.1080/09638280500534937] [Medline: 16777770]
51. Soares M, Bonuzzi GMG, Coelho DB, Torriani-Pasin C. Effects of extrinsic feedback on the motor learning after stroke. Motriz rev educ fis 2019;25(1). [doi: 10.1590/s1980-6574201900010013]
52. Carr K, Zachariah N, Weir P, McNevin N. An examination of feedback use in rehabilitation settings. Crit Rev Phys Rehabil Med 2011;23(1-4):147-160. [doi: 10.1615/critrevphysrehabilmed.v23.i1-4.110]
53. Stanton R, Ada L, Dean CM, Preston E. Feedback received while practicing everyday activities during rehabilitation after stroke: an observational study. Physiother Res Int 2015;20(3):166-173. [doi: 10.1002/pri.1612] [Medline: 25451759]

XSL•FO
RenderX

54. Johnson L, Burridge J, Demain S. Internal and external focus of attention during gait re-education: an observational study of physical therapist practice in stroke rehabilitation. Phys Ther 2013;93(7):957-966. [doi: 10.2522/ptj.20120300] [Medline: 23559523]
55. Figueira V, Silva S, Costa I, Campos B, Salgado J, Pinho L, et al. Wearables for monitoring and postural feedback in the work context: a scoping review. Sensors (Basel) 2024;24(4):1341 [FREE Full text] [doi: 10.3390/s24041341] [Medline: 38400498]
56. Machado Neto OJ, Peracini APP, Cunha BCR, Rocha ALC, Elui VMC, Castro PC, et al. An architecture for creating mobile assistive products: a case study on posture monitoring of people with motor disability. 2016 Presented at: 2nd IET International Conference on Technologies for Active and Assisted Living (TechAAL 2016); 2016 October 24-25; London, UK URL: https://ieeexplore.ieee.org/document/7801339
57. Ometov A, Shubina V, Klus L, Skibińska J, Saafi S, Pascacio P, et al. A Survey on Wearable Technology: History, State-of-the-Art and Current Challenges. Comput Networks 2021 Jul;193:108074. [doi: 10.1016/j.comnet.2021.108074]
58. Heebner NR, Akins JS, Lephart SM, Sell TC. Reliability and validity of an accelerometry based measure of static and dynamic postural stability in healthy and active individuals. Gait Posture 2015;41(2):535-539. [doi: 10.1016/j.gaitpost.2014.12.009] [Medline: 25544692]
59. Guest G, MacQueen K, Namey E. Applied Thematic Analysis. Washington, DC: Sage Publications; 2014.
60. Demers L, Weiss-Lambrou R, Ska B. The Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST 2.0): an overview and recent progress. Technol Disabil 2002;14(3):101-105. [doi: 10.3233/tad-2002-14304]
61. de Carvalho KEC, Gois Júnior MB, Sá KN. Translation and validation of the Quebec User Evaluation of Satisfaction With Assistive Technology (QUEST 2.0) into Portuguese. Rev Bras Reumatol 2014;54(4):260-267 [FREE Full text] [doi: 10.1016/j.rbr.2014.04.003] [Medline: 25627220]
62. Jayaraman C, Mummidisetty CK, Mannix-Slobig A, McGee Koch L, Jayaraman A. Variables influencing wearable sensor outcome estimates in individuals with stroke and incomplete spinal cord injury: a pilot investigation validating two research grade sensors. J Neuroeng Rehabil 2018;15(1):19 [FREE Full text] [doi: 10.1186/s12984-018-0358-y] [Medline: 29534737]
63. Grouios G, Ziagkas E, Loukovitis A, Chatzinikolaou K, Koidou E. Accelerometers in our pocket: does smartphone accelerometer technology provide accurate data? Sensors (Basel) 2022;23(1):192 [FREE Full text] [doi: 10.3390/s23010192] [Medline: 36616798]
64. Riberto M, Miyazaki MH, Jucá SSH, Sakamoto H, Potiguara P. Validação da Versão Brasileira da Medida de Independência Funcional Validation of the Brazilian version of Functional Independence Measure. Acta Fisiatrica 2004;11 [FREE Full text]
65. Kidd D, Stewart G, Baldry J, Johnson J, Rossiter D, Petruckevitch A, et al. The functional independence measure: a comparative validity and reliability study. Disabil Rehabil 1995;17(1):10-14. [doi: 10.3109/09638289509166622] [Medline: 7858276]
66. Pedrosa D, Pimentel MDGC, Wright A, Truong KN. Filteryedping: design challenges and user performance of dwell-free eye typing. ACM Trans Access Comput 2015;6(1):1-37. [doi: 10.1145/2724728]
67. Open mHealth. 2019. URL: https://www.openmhealth.org/ [accessed 2024-11-30]
68. Cunha BCR, Rodrigues KRDH, Zaine I, da Silva EAN, Viel CC, Pimentel MDGC. Experience sampling and programmed intervention method and system for planning, authoring, and deploying mobile health interventions: design and case reports. J Med Internet Res 2021;23(7):e24278 [FREE Full text] [doi: 10.2196/24278] [Medline: 34255652]
69. Dantas L, Pimentel MDGC, Melo R, De Lima AP. Analysis of a digital instrument for multidimensional assessment of the older adults by undergraduate gerontology students. Cuad Ed Desar 2023;15(9):9168-9185.
70. Cogollor JM, Rojo-Lacal J, Hermsdörfer J, Ferre M, Arredondo Waldmeyer MT, Giachritsis C, et al. Evolution of cognitive rehabilitation after stroke from traditional techniques to smart and personalized home-based information and communication technology systems: literature review. JMIR Rehabil Assist Technol 2018;5(1):e4 [FREE Full text] [doi: 10.2196/rehab.8548] [Medline: 29581093]
71. Sorrentino G, Sale P, Solaro C, Rabini A, Cerri CG, Ferriero G. Clinical measurement tools to assess trunk performance after stroke: a systematic review. Eur J Phys Rehabil Med 2018;54(5):772-784 [FREE Full text] [doi: 10.23736/S1973-9087.18.05178-X] [Medline: 29684980]
72. Elnady A, Mortenson WB, Menon C. Perceptions of existing wearable robotic devices for upper extremity and suggestions for their development: findings from therapists and people with stroke. JMIR Rehabil Assist Technol 2018;5(1):e12 [FREE Full text] [doi: 10.2196/rehab.9535] [Medline: 29764799]
73. Hazelton C, McGill K, Campbell P, Todhunter-Brown A, Thomson K, Nicolson DJ, et al. Perceptual disorders after stroke: a scoping review of interventions. Stroke 2022;53(5):1772-1787 [FREE Full text] [doi: 10.1161/STROKEAHA.121.035671] [Medline: 35468001]
74. Huo C, Zheng Y, Lu W, Zhang T, Wang D, Xu D, et al. Prospects for intelligent rehabilitation techniques to treat motor dysfunction. Neural Regen Res 2021;16(2):264-269 [FREE Full text] [doi: 10.4103/1673-5374.290884] [Medline: 32859773]
75. Guo L, Wang J, Wu Q, Li X, Zhang B, Zhou L, et al. Clinical study of a wearable remote rehabilitation training system for patients with stroke: randomized controlled pilot trial. JMIR Mhealth Uhealth 2023;11:e40416 [FREE Full text] [doi: 10.2196/40416] [Medline: 36821348]

XSL•FO
RenderX

76. Toh SFM, Fong KNK, Gonzalez PC, Tang YM. Application of home-based wearable technologies in physical rehabilitation for stroke: a scoping review. IEEE Trans Neural Syst Rehabil Eng 2023. [doi: 10.1109/TNSRE.2023.3252880] [Medline: 37028029]
77. Davies RJ, Parker J, McCullagh P, Zheng H, Nugent C, Black ND, et al. A personalized self-management rehabilitation system for stroke survivors: a quantitative gait analysis using a smart insole. JMIR Rehabil Assist Technol 2016;3(2):e11 [FREE Full text] [doi: 10.2196/rehab.5449] [Medline: 28582260]
78. Fonte G, Schreiber C, Areno G, Masson X, Chantraine F, Schütz G, et al. Metabolic energy expenditure and accelerometer-determined physical activity levels in post-stroke hemiparetic patients. J Stroke Cerebrovasc Dis 2022;31(5):106397. [doi: 10.1016/j.jstrokecerebrovasdis.2022.106397] [Medline: 35228022]
79. Ploderer B, Fong J, Klaic M, Nair S, Vetere F, Cofré Lizama LE, et al. How therapists use visualizations of upper limb movement information from stroke patients: a qualitative study with simulated information. JMIR Rehabil Assist Technol 2016;3(2):e9 [FREE Full text] [doi: 10.2196/rehab.6182] [Medline: 28582257]
80. Toh SFM, Gonzalez PC, Fong KNK. Usability of a wearable device for home-based upper limb telerehabilitation in persons with stroke: a mixed-methods study. Digit Health 2023;9:20552076231153737 [FREE Full text] [doi: 10.1177/20552076231153737] [Medline: 36776407]
81. Lin WTM, Lin BS, Lee IJ, Lee SH. Development of a smartphone-based mHealth platform for telerehabilitation. IEEE Trans Neural Syst Rehabil Eng 2022;30:2682-2691. [doi: 10.1109/TNSRE.2022.3204148] [Medline: 36063516]

Abbreviations

QUEST: Quebec User Evaluation of Satisfaction with Assistive Technology

Edited by Q Fan; submitted 13.12.23; peer-reviewed by C Abbatantuono; comments to author 01.04.24; revised version received 29.08.24; accepted 06.09.24; published 22.01.25.

Please cite as:
Pereira AP, Machado Neto OJ, Elui VMC, Pimentel MDGC
Wearable Smartphone-Based Multisensory Feedback System for Torso Posture Correction: Iterative Design and Within-Subjects Study
JMIR Aging 2025;8:e55455
URL: https://aging.jmir.org/2025/1/e55455
doi:10.2196/55455
PMID:

©Amanda Polin Pereira, Olibario Jose Machado Neto, Valeria Meirelles Carril Elui, Maria da Graca Campos Pimentel. Originally published in JMIR Aging (https://aging.jmir.org), 22.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Estimation of Machine Learning–Based Models to Predict Dementia Risk in Patients With Atherosclerotic Cardiovascular Diseases: UK Biobank Study

Zhengsheng Gu[1*], MD; Shuang Liu[2*], MM; Huijuan Ma[2*], PhD; Yifan Long[2], MM; Xuehao Jiao[1], MD; Xin Gao[1], MM; Bingying Du[1,3*], PhD; Xiaoying Bi[1*], PhD; Xingjie Shi[2*], PhD

[1]Department of Neurology, First Affiliated Hospital of Naval Medical University, Shanghai, China

[2]KLATASDS-MOE, Academy of Statistics and Interdisciplinary Sciences, School of Statistics, East China Normal University, Shanghai, China

[3]Institute for Translational Brain Research, Fudan University, Shanghai, China

[*]these authors contributed equally

Corresponding Author:
Xingjie Shi, PhD
KLATASDS-MOE
Academy of Statistics and Interdisciplinary Sciences, School of Statistics
East China Normal University
No 3663 North Zhongshan Road, Putuo District
Shanghai, 200062
China
Phone: 86 21 622 332 23
Email: xjshi@fem.ecnu.edu.cn

Abstract

Background: The atherosclerotic cardiovascular disease (ASCVD) is associated with dementia. However, the risk factors of dementia in patients with ASCVD remain unclear, necessitating the development of accurate prediction models.

Objective: The aim of the study is to develop a machine learning model for use in patients with ASCVD to predict dementia risk using available clinical and sociodemographic data.

Methods: This prognostic study included patients with ASCVD between 2006 and 2010, with registration of follow-up data ending on April 2023 based on the UK Biobank. We implemented a data-driven strategy, identifying predictors from 316 variables and developing a machine learning model to predict the risk of incident dementia, Alzheimer disease, and vascular dementia within 5, 10, and longer-term follow-up in patients with ASCVD.

Results: A total of 29,561 patients with ASCVD were included, and 1334 (4.51%) developed dementia during a median follow-up time of 10.3 (IQR 7.6-12.4) years. The best prediction model (UK Biobank ASCVD risk prediction model) was light gradient boosting machine, comprising 10 predictors including age, time to complete pairs matching tasks, mean time to correctly identify matches, mean sphered cell volume, glucose levels, forced expiratory volume in 1 second *z* score, C-reactive protein, forced vital capacity, time engaging in activities, and age first had sexual intercourse. This model achieved the following performance metrics for all incident dementia: area under the receiver operating characteristic curve: mean 0.866 (SD 0.027), accuracy: mean 0.883 (SD 0.010), sensitivity: mean 0.637 (SD 0.084), specificity: mean 0.914 (SD 0.012), precision: mean 0.479 (SD 0.031), and F_1-score: mean 0.546 (SD 0.043). Meanwhile, this model was well-calibrated (Kolmogorov-Smirnov test showed goodness-of-fit *P* value>.99) and maintained robust performance across different temporal cohorts. Besides, the model had a beneficial potential in clinical practice with a decision curve analysis.

Conclusions: The findings of this study suggest that predictive modeling could inform patients and clinicians about ASCVD at risk for dementia.

(JMIR Aging 2025;8:e64148) doi:10.2196/64148

KEYWORDS

atherosclerotic cardiovascular disease; dementia; Alzheimer disease; vascular dementia; machine learning; UK Biobank

XSL•FO
RenderX

Introduction

Cardiovascular disease (CVD) is the leading cause of noncommunicable disease and mortality worldwide [1]. Meanwhile, the epidemiology of the atherosclerotic cardiovascular disease (ASCVD), which encompasses coronary heart disease and cerebrovascular disease (CeVD), has experienced substantial and rapid growth [2]. It is reported that in 2016, ASCVD was responsible for approximately 2.4 million deaths, representing 25% of all deaths and 61% of CVD-related deaths in China [3].

Dementia is another devastating disease affecting more than 50 million individuals worldwide [4]. Given the high costs and heavy burdens it imposes on families and society, scientists and scholars around the world are dedicated to identifying preventable interventions and reducing the incidence of dementia. Recently, a growing body of evidence indicates that lifestyle interventions early in life with a focus on reducing cardiovascular risk factors are a promising strategy for preventing dementia [5-9]. In particular, shared risk factors between dementia and ASCVD have been identified [10]. According to the Lancet Commission, it is estimated that approximately 40% of dementia cases can be prevented by targeting modifiable, primarily cardiovascular risk factors [4]. However, these studies were restricted by their use of classical statistical analyses (such as Cox or logistic regressions) and by considering only widely studied prespecified CVD risks. Therefore, the results were not sufficient in accuracy.

Machine learning (ML) is an emerging technical foundation of artificial intelligence, which enables the leverage of information from large and complex datasets [11]. Several studies have applied ML-based models to dementia diagnosis and risk prediction [12-15]. Nevertheless, the long-term risk of dementia progression (5 or 10 years) in patients with ASCVD remains uncertain.

In this study, we used comprehensive phenotypic and follow-up data from a cohort of over 500,000 UK Biobank participants to develop an ML-based model capable of predicting the 5- or 10-year risk of incident dementia in specific patients with ASCVD. We anticipate that this ML-derived early warning system will enhance clinician-patient counseling, enable targeted follow-up, and facilitate the development of personalized prevention strategies. This, ultimately, can optimize the health and care of individuals with ASCVD.

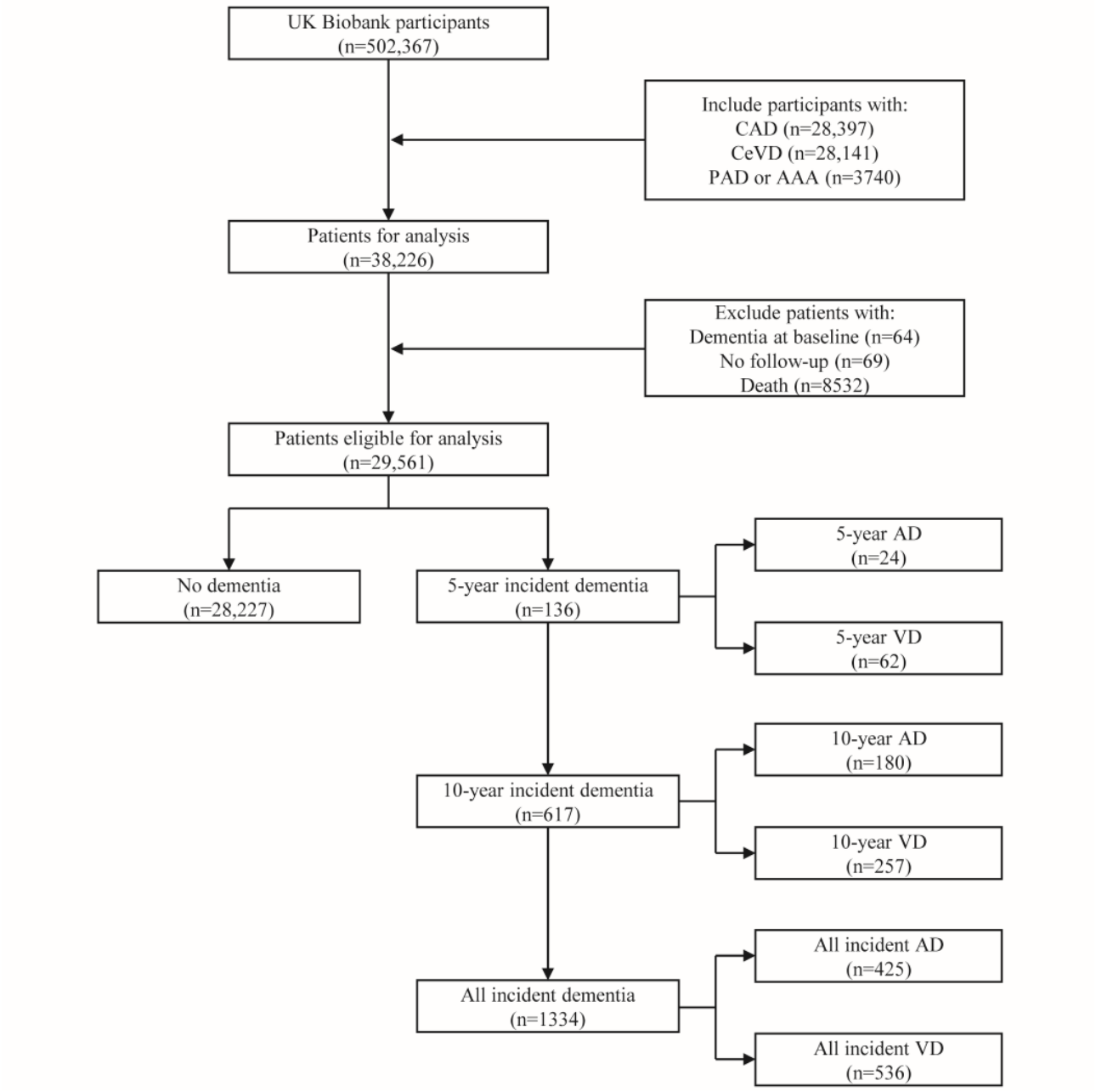
Methods

Data Source and Study Population

We analyzed data from the UK Biobank, a longitudinal prospective study that recruited over 500,000 participants between 2006 and 2010 [16]. The participants were enrolled from 22 recruitment centers across the United Kingdom and were aged between 40 and 69 years at the baseline assessment. Multiple data were collected from the participants, including questionnaires, physical measurements, biological sample assays, genotyping, imaging data, and ongoing hospital records. Figure 1 illustrated the enrollment process, where we included individuals with a prior history of any established ASCVD, such as coronary artery disease (n=28,397), CeVD (n=28,141), peripheral artery disease, or abdominal aortic aneurysm (n=3740) [17]. Participants were excluded at the baseline assessment if they met the following criteria: (1) had dementia at baseline (n=64), (2) had no follow-up records (n=69), and (3) death (n=8532). Ultimately, we included 29,561 participants with ASCVD who had at least 10 years of follow-up until April 2023.

XSL•FO
RenderX

Figure 1. Participant selection flowchart. UK Biobank participants were excluded if baseline dementia was self-reported or follow-up data were absent. The remaining participants were categorized according to their first reported years of dementia, AD, or VD after the baseline. AAA: abdominal aortic aneurysm; AD: Alzheimer disease; CAD: coronary artery disease; CeVD: cerebrovascular disease; PAD: peripheral artery disease; VD: vascular dementia.



Ethical Considerations

Ethical approval was obtained from the North West Multi-Centre Research Ethics Committee (11/NW/0382, 16/NW/0274, and 21/NW/0157). Written informed consent was provided by all participants during the collection of primary data. The UK Biobank data used were deidentified, and all personally identifiable information of participants has been removed to ensure privacy and confidentiality. Besides, the UK Biobank offered nonfinancial compensation in the form of travel reimbursements for attending the assessment centers and other incidental expenses related to participation. Additionally, participants were given feedback on their individual health data upon request, which provided valuable insights into their health status. This study adhered to the reporting guidelines of Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis [18].

Outcome

The primary end point of this study was the occurrence of all incident dementia, including Alzheimer disease (AD), vascular dementia (VD), frontotemporal dementia, and dementia associated with other neurodegenerative or specified diseases. Due to the high rate of incidence worldwide, AD and VD were examined as the secondary outcomes. To conduct a comprehensive survey on the incidence time, we categorized the patients into 5-year, 10-year, and all incident dementia, AD, and VD. The outcomes were ascertained and categorized based on the International Classification of Diseases and Read codes (Table S1 in Multimedia Appendix 1), which were obtained

XSL•FO
RenderX

from the "first occurrence" category in the UK Biobank including the primary care data, the hospital inpatient data, the death register records, and subsequent UK Biobank assessment center visits. Follow-up visits continued until the earliest of the following events: a dementia diagnosis, death, or the most recent available data from either the hospital or the general practitioner, whichever occurred first. What is noteworthy was that the imaging data and lumbar puncture results were not available to doctors to achieve the detailed diagnostic data.

Data Preparation

In this study, we included all clinically correlated variables during the participants' baseline visits. The assessment procedure involved a manual examination of each variable to determine its relevance to comprehensively understanding a participant's overall status. Variables not pertinent to these key domains or lacking in additional insights were excluded. Data screening was processed to exclude noninformative variables with missing values exceeding 40% among all participants. To prevent potential overfitting from oversampling, we applied random undersampling to the majority class, balancing the dataset more effectively. We also adjusted the class weights in our ML algorithms to give more importance to the minority class during model training. Overall, a total of 316 features were adopted, including the participants' demographic characteristics (n=2), touchscreen-recorded questionnaires (n=151), physical measures (n=66), cognitive function tests (n=22), and biological sample assays (n=59). Furthermore, to improve the informative value of the dataset, we used the available data to generate several variables (n=16) that were not directly extracted from the UK Biobank (Table S2 in Multimedia Appendix 1). Considering the significant impact of ASCVD on mortality, we identified and coded deaths as competing events to ensure accurate modeling of the primary outcome.

In this study, we used different missing data handling strategies tailored to each ML algorithm to ensure the accuracy and robustness of the models. Specifically, for the logistic regression model, we used mean imputation to handle missing values. For each variable with missing data, we calculated its mean in the training dataset and replaced the missing data points with this mean. This method is simple and efficient, making it suitable for models like logistic regression that require complete datasets. For other ML methods, including light gradient boosting machine (LightGBM), extreme gradient boosting machine, random forest, k-nearest neighbor, and artificial neural network, we adopted automatic imputation techniques. These automatic imputation methods leverage the inherent mechanisms of the algorithms or advanced imputation strategies within the preprocessing pipeline to dynamically estimate and replace missing values. To evaluate the robustness of our results, we conducted a sensitivity analysis using multiple imputation by chained equations (Table S4 in Multimedia Appendix 1). This approach generates several imputed datasets by modeling each missing value conditionally based on other variables, thereby accounting for the uncertainty associated with the imputations. By comparing the outcomes across different imputation methods, we assessed the stability and reliability of our predictive models.

To evaluate the model's stability and generalization across different time periods, this study used a time validation approach to partition dementia data from the UK Biobank database. First, the recruitment date was selected as the primary temporal variable, and all samples were sorted in ascending order based on this date to ensure chronological arrangement and prevent future data from leaking into the training process. Considering previous research and data volume, the dataset was divided into 2 periods: the training and validation sets comprised samples recruited from 2006 to 2009, while the test set included samples diagnosed in 2010. This division ensures that only past data were used for model training, and the model's predictive performance on future data was assessed during the validation and testing phases. To further guarantee temporal independence, feature selection and standardization were performed exclusively on the training set, with identical transformations applied to the validation and test sets, thereby avoiding the use of information from these sets during training. Given the typically low number of dementia cases, healthy samples with more than 5% missing variables were excluded from the training set to balance the class distribution and enhance model learning. Additionally, multiple imputation methods were used to handle missing data, ensuring data integrity. Through these steps, a time validation framework was established, maintaining the temporal independence and appropriate distribution of the training, validation, and test sets, thereby improving the model's predictive performance across different time periods and the credibility of the study's findings.

Predictor Selection

The predictors for model development were identified through a 2-step process: variable importance ranking and sequential forward selection (SFS) [19,20]. First of all, the importance of each variable was determined using a preliminary trained LightGBM classifier. Gradient boosting machine is a type of boosting that builds these simple models step-by-step, improving the model at each step to better fit the data. LightGBM, developed by Microsoft, is a faster and more efficient version of gradient boosting machine designed to handle large-scale data effectively. The "light" in LightGBM refers to its lightweight nature, meaning it uses less memory and runs faster. The top 50 variables were selected by LightGBM. Next, they were inputted into a hierarchical clustering algorithm, which used Spearman rank-order correlations to further identify and eliminate redundant variables with multicollinearity. We established a correlation threshold of 0.75, considering variables with pairwise correlations above this value as highly redundant. Within each cluster of such variables, we retained only the most predictive variable for the model, effectively reducing multicollinearity while preserving essential information. To avoid overfitting and enhance the robustness of feature selection, a nested cross-validation approach was used. Specifically, in the outer loop, we divided the dataset into multiple folds, selecting 1 fold as the test set and using the remaining folds for feature selection and model training. Within the inner loop, the training set was further split into inner training and validation sets, where features were selected based on performance in the inner validation sets. Finally, model performance was evaluated on the outer test set to ensure fair feature selection and robust

XSL•FO
RenderX

predictive capability. Then, an SFS approach was used, wherein the features within the preselected subset underwent reranking according to a newly developed classifier. Afterward, preselected variables were reranked, and multiple ML classifiers were used to sequentially add predictors one at each time. Finally, the classifier was selected based on achieving the best performance of area under the receiver operating characteristic curve (AUC), and we selected the top 10 variables according to the importance of each variable calculated by the LightGBM model. After selecting these 10 variables, adding any other variables did not significantly improve the model. The top 25 predictors are shown in Table 1. More details of predictor selection could be obtained in the part of the Methods section in Multimedia Appendix 1.

Table 1. Top 25 predictors for all incident dementia with light gradient boosting machine.

Number	Variables	Importance rating	Ranking
1	Forced vital capacity	0.133	1
2	Summed MET[a] minutes per week for all activity	0.12	2
3	Age	0.09	3
4	Pairs matching time	0.05	4
5	Mean sphered cell volume	0.038	5
6	Glucose	0.037	6
7	Mean time to correctly identify matches	0.037	7
8	FEV1[b] *z* score	0.035	8
9	Age first had sexual intercourse	0.033	9
10	C-reactive protein	0.031	10
11	Average weekly red wine intake	0.029	11
12	Calcium	0.029	12
13	Vitamin D	0.028	13
14	Pulse rate automated reading	0.026	14
15	Father age at death	0.025	15
16	Systolic blood pressure automated reading array	0.025	16
17	FEV1/FVC[c] ratio *z* score	0.024	17
18	Neuroticism score	0.023	18
19	Red blood cell erythrocyte distribution width	0.023	19
20	Apolipoprotein B	0.023	20
21	Total bilirubin	0.023	21
22	Cystatin C	0.022	22
23	Alanine aminotransferase	0.018	23
24	Average weekly beer plus cider intake	0.017	24
25	Result ranking	0.016	25

[a]MET: metabolic equivalent.

[b]FEV1: forced expiratory volume in 1 second.

[c]FVC: forced vital capacity.

Model Development

We implemented a range of ML techniques, including LightGBM, extreme gradient boosting machine, random forest, logistic regression, k-nearest neighbor, support vector machine, and artificial neural network to classify participants into 2 classes: 0 (predicted to remain no dementia) or class 1 (to develop all incident dementia, AD, or VD). The proposed model was developed using patients with ASCVD without dementia (n=28,227) and with all incident dementia (n=1334) from the UK Biobank dataset. In total, 10 identified predictors were incorporated into the model. We expanded our performance evaluation metrics to include receiver operating characteristic-AUC, precision, recall, and F_1-score, ensuring a comprehensive assessment of the models' performance on imbalanced data. Subsequently, LightGBM, the best-performing method, was used to develop a dementia risk prediction model of ASCVD, named the UK Biobank ASCVD risk prediction model. The hyperparameter tuning was performed through exhaustive selection from 10,000 sets of candidate parameters,

XSL•FO
RenderX

and the optimal set was chosen based on the performance measurement of AUC. Please refer to Table S8 in Multimedia Appendix 1 for detailed information on the search space and final adopted parameters. To evaluate the predictive performance of the models, we constructed and compared the traditional Cox proportional hazards model with the LightGBM model. Both models used identical predictor variables to ensure fairness and consistency in the comparison. The Cox model assessed hazard ratios for each variable through multivariate regression analysis, while the LightGBM model leveraged its robust ability to handle nonlinear relationships and variable interactions for risk prediction. Subsequently, the performance of both models was systematically compared using consistent evaluation metrics (such as AUC) to determine their predictive effectiveness within the study dataset. This comparison aims to validate the potential advantages of the LightGBM model in risk prediction and to provide a reference for the application of the traditional Cox model. The Cox proportional hazards model was also used to account for competing risks, which ensured that the risk of death did not bias the estimation of dementia event probabilities. To enhance the model's stability and applicability, a time validation approach was used for data analysis. Additionally, we calibrated the raw predicted probabilities into actual dementia risks (Figures S12-S14 in Multimedia Appendix 1). Finally, to assess the clinical utility of the prediction model, decision curve analysis (DCA) was conducted. First, the model's net benefit was calculated across various threshold probabilities and then compared with the baseline strategies of "treat all" and "treat none." The DCA curves were plotted using the *rmda* package in R (R Foundation for Statistical Computing) to illustrate the model's potential value in clinical decision-making. The ML algorithm was implemented using LightGBM library (version 3.3.2) and scikit-learn library (version 1.0.2) in Python (version 3.9; Python Software Foundation).

We also performed a 5-fold cross-validation to assess the stability of feature importance, randomly dividing the dataset into 5 equal parts. In each iteration, 4 folds were used for training, and 1 fold for validation. The training involved 2 stages: model development and calibration. The 4 training folds were split 3:1, with 3 folds for development and 1 fold for calibration. Validation sets were exclusively for performance evaluation. Results were averaged across folds with corresponding SDs.

Statistical Analysis

In an analysis of the variables of interest, continuous variables were summarized using the median and IQR, while discrete variables were summarized using frequency and percentage. Group comparisons (no dementia vs incident dementia or AD or VD) were conducted using chi-square tests for discrete variables and 2-tailed Student *t* tests for continuous variables. Multivariate analysis was used to calculate odds ratios based on normalized data.

The model's performance was evaluated using 2 accuracy metrics: discrimination and calibration. Discrimination was assessed using the AUC, which ranges from 0.5 for a noninformative model to 1 for a perfectly discriminating model. Calibration measures the agreement between predicted probabilities and observed event proportions. It was evaluated using the Kolmogorov-Smirnov test with 10 subgroups and visually represented in calibration plots. A *P* value greater than .05 signified an adequate goodness of fit.

Furthermore, we reported accuracy, sensitivity, specificity, precision, and the F_1-score, which were determined using the cutoff that maximized the Youden index. Additionally, we used Shapley Additive Explanations (SHAP) plots to visualize the individual contributions of each predictor to the target variable. All data analysis and visualizations were performed using Python (version 3.9) with packages from the scikit-learn library (version 1.0.2) and the SHAP library (version 0.40.0).

Results

Population Characteristics

After quality control, a total of 29,561 participants with ASCVD were included in this study. The median age of the participants was 62.0 (IQR 58.0-66.0) years. Among the participants, 36.63% (10,829/29,561) were women, and 94.12% (27,822/29,561) were White. During a median follow-up time of 10.3 (IQR 7.6-12.4) years, a subset of 1334 participants developed dementia after their baseline visits. Specifically, 617 participants had incidents within 10 years, and 136 had incidents within 5 years. Besides, the prevalence of all-cause dementia was 4.51% (1334/29,561), AD was 1.44% (425/29,561), and VD was 1.81% (536/29,561) in this study. The critical baseline predictors are presented by incident dementia, AD, and VD status in Table 2, and the percentage of missing values for the predictors is shown in Table S3 in Multimedia Appendix 1.

XSL•FO
RenderX

Table 2. The baseline characteristics of UK Biobank participants included in the study by dementia, Alzheimer disease (AD), and vascular dementia (VD) status.

Participants characteristics	Overall (n=29,561)	No dementia (n=28,227)	All incident dementia (n=1334)	All incident AD (n=425)	All incident VD (n=536)
Age (years), median (IQR)	62.0 (58.0-66.0)	62.0 (57.0-66.0)	66.0 (63.0-68.0)	66.0 (64.0-68.0)	66.0 (63.0-68.0)
Sex (female), n (%)	10,829 (36.63)	10,353 (36.68)	476 (35.68)	170 (40.00)	163 (30.41)
Ethnicity (White), n (%)	27,822 (94.12)	26,565 (94.11)	1257 (94.23)	404 (95.06)	504 (94.03)
Education (years), median (IQR)	10.0 (9.0-11.0)	10.0 (9.0-11.0)	10.0 (9.0-11.0)	9.0 (9.0-10.0)	9.0 (9.0-10.0)
Forced vital capacity (L), median (IQR)	3.5 (2.9-4.2)	3.5 (2.9-4.2)	3.2 (2.6-3.9)	3.2 (2.7-4.0)	3.1 (2.6-3.9)
Summed MET[a] minutes per week for all activity (minutes per week), median (IQR)	1662.0 (693.0-3546.0)	1671.0 (698.0-3546.0)	1398.0 (510.0-3288.0)	1653.0 (660.0-3546.0)	1308.0 (408.8-2942.6)
Pairs matching time (seconds), median (IQR)	411.0 (325.0-534.0)	408.0 (324.0-529.0)	487.5 (368.0-678.8)	487.0 (372.0-673.5)	496.0 (370.5-701.0)
Mean sphered cell volume (fL), median (IQR)	82.4 (79.2-85.9)	82.4 (79.2-85.8)	83.1 (79.5-86.7)	82.5 (79.6-86.1)	83.2 (79.2-87.3)
Glucose (mmol/L), median (IQR)	5.0 (4.7-5.5)	5.0 (4.7-5.5)	5.1 (4.7-5.9)	5.1 (4.7-5.6)	5.2 (4.7-6.2)
Mean time to correctly identify matches (seconds), median (IQR)	563.0 (500.0-644.0)	562.0 (500.0-641.0)	594.0 (531.0-699.8)	586.0 (527.0-684.0)	602.0 (532.0-707.5)
FEV1[b]z score (L), median (IQR)	0.6 (–0.1 to 1.3)	0.6 (–0.1 to 1.3)	0.7 (0.0-1.5)	0.5 (–0.0 to 1.4)	0.8 (0.2-1.6)
Age first had sexual intercourse (years), median (IQR)	18.0 (16.0-21.0)	18.0 (16.0-21.0)	18.0 (17.0-21.0)	19.0 (17.0-21.0)	18.0 (16.0-21.0)
C-reactive protein (mg/L), median (IQR)	1.5 (0.8-3.1)	1.5 (0.8-3.1)	1.5 (0.7-3.2)	1.5 (0.7-3.3)	1.7 (0.8-3.4)
CAD[c], n (%)	21,735 (73.53)	20,763 (73.56)	972 (72.86)	325 (76.47)	357 (66.60)
CeVD[d], n (%)	21,707 (73.43)	20,738 (73.47)	969 (72.64)	324 (76.24)	356 (66.42)
PAD[e] or AAA[f], n (%)	1345 (4.55)	1270 (4.50)	75 (5.62)	23 (5.41)	34 (6.34)

[a]MET: metabolic equivalent.

[b]FEV1: forced expiratory volume in 1 second.

[c]CAD: coronary artery disease.

[d]CeVD: cerebrovascular disease.

[e]PAD: peripheral artery disease.

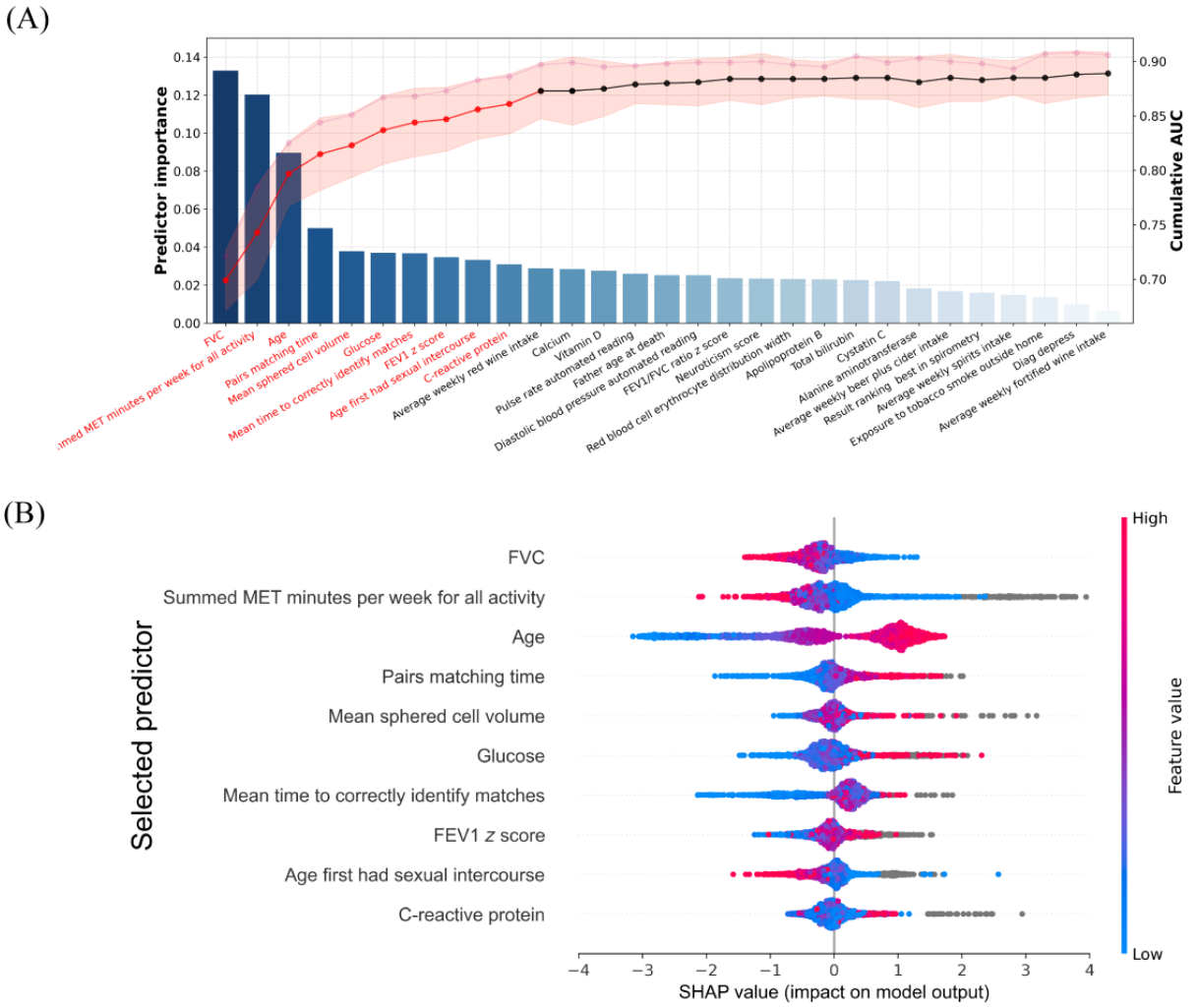
[f]AAA: abdominal aortic aneurysm.

Data-Driven Predictors Selection

Among the 316 candidate variables, we initially selected the top 50 variables based on the LightGBM classifier and performed the hierarchical clustering to eliminate the multicollinearity [21]. As shown in the bar chart of Figure 2A, a total of 29 variables were sorted according to their importance in the prediction task. The SFS strategy was used to strike a balance between model performance (AUC on the right axis) and the number of variables selected, as depicted in the line chart. The line chart showed that the model's performance experienced a sharp increase when incorporating the first few variables and eventually reached a plateau with the inclusion of additional variables. Ultimately, the top 10 variables were chosen as the final predictors for ML model development. Their summary statistics are displayed in Table 2.

XSL•FO
RenderX

Figure 2. Predictive variable selection and interpretation on all incident dementia. (A) Sequential forward selection from a preselected predictor pool. A bar chart ranked predictor importance by their contribution to classification, while a line chart depicted cumulative AUCs with each iterative predictor inclusion. The top 10 predictors, marked red, were selected for machine learning model construction. (B) SHAP-based visualization of salient predictors. Horizontal bar widths correspond to predictor impact on model predictions, with wider ranges indicating greater influence. Predictor intensity was color-coded, graduating from blue (low) to red (high), as per the color bar on the right. The x-axis orientation signified the probability of either dementia (right) or health (left). AUC: area under the receiver operating characteristic curve; FEV1: forced expiratory volume in 1 second; FVC: forced vital capacity; MET: metabolic equivalent; SHAP: Shapley Additive Explanations.

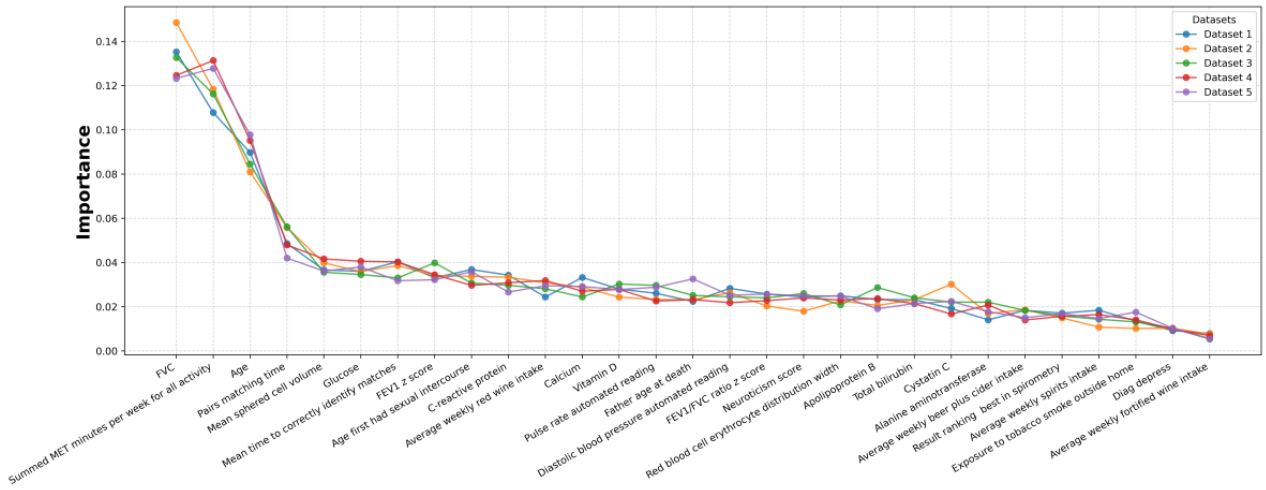


Model Interpretation of Selected Predictors

To interpret the influence of each selected predictor, we used SHAP values and visualized them in Figure 2B. The predictors were interpreted based on value magnitude (coded by gradient colors) and tendency on the horizontal axis (indicting the likelihood of developing dementia). Take the predictor forced vital capacity (FVC) as an example. Patients with ASCVD with lower FVC values (colored blue) were more likely to develop dementia (right side) compared to those with higher FVC values (colored red). Similarly, for the remaining predictors, patients with ASCVD who spend less time engaging in activities, being older, who take longer time to complete pair matching tasks, and who have higher mean sphered cell volume (MSCV), forced expiratory volume in 1 second *z* score, C-reactive protein, and glucose levels tend to have an increased risk of developing dementia. Interestingly, we found that patients who engaged in sexual intercourse at an earlier age were more likely to develop dementia. Moreover, a 5-fold cross-validation stability analysis of feature importance was conducted. The results indicated that most features exhibited high stability across different data subsets. Figure 3 shows the distribution of key feature importance across all folds.

XSL•FO
RenderX

Figure 3. Stability analysis of feature importance across different data subsets in the model of all incident dementia. The stability of feature importance in the model of all incident dementia was assessed across different data subsets, providing insights into the robustness of identified predictors. Each line represented the variability of feature importance in a specific subset, illustrating how consistent the predictive factors are across varying conditions. FEV1: forced expiratory volume in 1 second; FVC: forced vital capacity; MET: metabolic equivalent.



Model Performance Across Different Populations and Algorithms

Compared with other ML algorithms, it can be seen in Table 3 that LightGBM demonstrated superior performance across various metrics. We used the AUC metric to evaluate the discrimination performance of the UK Biobank ASCVD risk prediction model. As depicted in Table 4, the model achieved a mean AUC of 0.866 (SD 0.027) for all incident dementia cases. Furthermore, the model demonstrated promising results for the prediction of 10-year and 5-year incident dementia, with mean AUCs of 0.876 (SD 0.024) and 0.903 (SD 0.076), respectively. The model for all incident dementia exhibited a mean accuracy of 0.883 (SD 0.01), mean sensitivity of 0.637 (SD 0.084), mean specificity of 0.914 (SD 0.012), mean precision of 0.479 (SD 0.031), and mean F_1-score of 0.546 (SD 0.043). Apart from the 5-year AD and VD predictions, the model also displayed valuable discrimination abilities for different AD and VD population groups. Specifically, the mean AUCs for all and 10-year incident AD were 0.836 (SD 0.043) and 0.828 (SD 0.112), respectively, while the mean AUCs for all and 10-year incident VD achieved 0.870 (SD 0.029) and 0.881 (SD 0.031), respectively. For specific metrics of all, 10-year, and 5-year dementia, AD, and VD predictions, please refer to Table 4 and Figures S3-S11 and S15 in Multimedia Appendix 1. We also compared the performance of the traditional Cox proportional hazards model and the LightGBM-based model in risk prediction. The AUC of all incident dementia, AD, and VD for the Cox model were 0.67, 0.67, and 0.71, respectively (Figure S1 in Multimedia Appendix 1). After competing risk analysis with death, the prediction power of the Cox model did not show a significant difference in predicting AD and VD (Tables S5-S7 and Figure S2 in Multimedia Appendix 1).

Table 3. Model performance metrics for different machine learning classifiers on all incident dementiaa.

	Accuracy, mean (SD)	Sensitivity, mean (SD)	Specificity, mean (SD)	Precision, mean (SD)	F_1-score, mean (SD)	AUC[b], mean (SD)
LightGBM[c]	0.883 (0.010)	0.637 (0.084)	0.914 (0.012)	0.479 (0.031)	0.546 (0.043)	0.866 (0.027)
XGBoost[d]	0.814 (0.01)	0.709 (0.045)	0.826 (0.01)	0.323 (0.018)	0.444 (0.024)	0.853 (0.02)
Random forest	0.829 (0.01)	0.703 (0.047)	0.844 (0.012)	0.345 (0.02)	0.463 (0.024)	0.859 (0.02)
KNN[e]	0.829 (0.014)	0.561 (0.059)	0.86 (0.012)	0.32 (0.033)	0.407 (0.041)	0.765 (0.024)
Logistic regression	0.783 (0.014)	0.627 (0.071)	0.802 (0.016)	0.27 (0.022)	0.378 (0.031)	0.795 (0.035)
ANN[f]	0.836 (0.1)	0.638 (0.145)	0.859 (0.125)	0.377 (0.174)	0.46 (0.102)	0.833 (0.025)

[a]The cutoff for binarization was established by maximizing the Youden index (YI=sensitivity+specificity–1).

[b]AUC: area under the receiver operating characteristic curve.

[c]LightGBM: light gradient boosting machine.

[d]XGBoost: extreme gradient boosting machine.

[e]KNN: k-nearest neighbor.

[f]ANN: artificial neural network.

XSL•FO
RenderX

Table 4. Model performance metrics for the prediction on different types of dementia[a].

	Accuracy, mean (SD)	Sensitivity, mean (SD)	Specificity, mean (SD)	Precision, mean (SD)	F_1-score, mean (SD)	AUC[b], mean (SD)
All incident dementia	0.883 (0.01)	0.637 (0.084)	0.914 (0.012)	0.479 (0.031)	0.546 (0.043)	0.866 (0.027)
All incident AD[c]	0.847 (0.029)	0.656 (0.125)	0.855 (0.035)	0.148 (0.01)	0.241 (0.012)	0.836 (0.043)
All incident VD[d]	0.859 (0.039)	0.683 (0.045)	0.868 (0.043)	0.206 (0.045)	0.315 (0.053)	0.870 (0.029)
10-Year incident dementia	0.877 (0.033)	0.709 (0.059)	0.886 (0.037)	0.255 (0.057)	0.374 (0.063)	0.876 (0.024)
10-Year incident AD	0.668 (0.061)	0.814 (0.214)	0.666 (0.065)	0.036 (0.006)	0.07 (0.012)	0.828 (0.112)
10-Year incident VD	0.839 (0.051)	0.757 (0.1)	0.841 (0.055)	0.098 (0.02)	0.173 (0.029)	0.881 (0.031)
5-Year incident dementia	0.939 (0.031)	0.694 (0.155)	0.942 (0.033)	0.125 (0.045)	0.211 (0.069)	0.903 (0.076)
5-Year incident AD	0.979 (0.006)	0.4 (0.5)	0.98 (0.006)	0.039 (0.049)	0.061 (0.047)	0.775 (0.243)
5-Year incident VD	0.952 (0.012)	0.471 (0.288)	0.957 (0.012)	0.106 (0.063)	0.172 (0.102)	0.803 (0.11)

[a]Cutoffs were established by maximizing the Youden index (YI=sensitivity+specificity–1).

[b]AUC: area under the receiver operating characteristic curve.

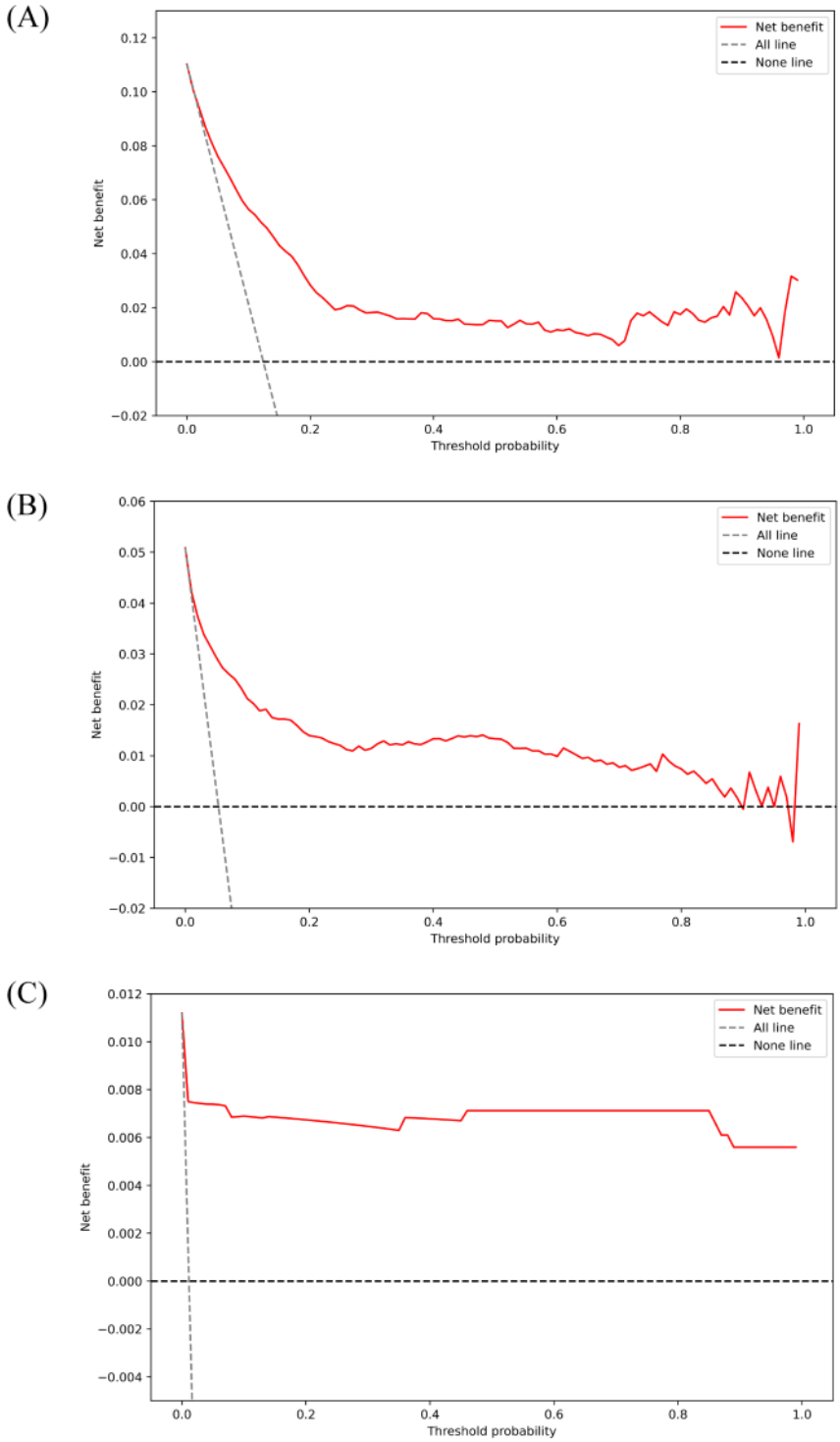
[c]AD: Alzheimer disease.

[d]VD: vascular dementia.

DCA demonstrated that our prediction model exhibited a higher net benefit within the threshold probability range of 0.04 to 0.97 across different time periods, significantly outperforming both the “treat all” and “treat none” baseline strategies (Figure 4). The Kolmogorov-Smirnov test was conducted to assess the calibration of the model. A *P* value greater than .05 indicates sufficient goodness of fit. Satisfactory calibrations for the development of all population groups, including 5-year or 10-year or all incident dementia, AD, and VD, were observed (Table 5 and Figures S12-S14 in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 4. Clinical applicability of dementia risk prediction with a decision curve analysis. (A-C) The decision curve analysis of the UK Biobank atherosclerotic cardiovascular disease risk prediction model on all incident and 10-year incident and 5-year incident times.



XSL•FO
RenderX

Table 5. Calibration data of the UK Biobank atherosclerotic cardiovascular disease risk prediction model[a].

Decile groups (10% quantile each)	All incident dementia[b]		All incident AD[b,c]		All incident VD[b,d]	
	Observed probability (‰)	Preserved probability (‰)	Observed probability (‰)	Preserved probability (‰)	Observed probability (‰)	Preserved probability (‰)
1	4.71	3.53	2.83	2.10	1.79	1.45
2	9.40	11.72	2.17	5.46	0.98	4.60
3	24.78	21.98	15.58	8.93	8.03	8.37
4	35.91	33.55	12.88	12.13	10.98	12.72
5	41.65	45.96	14.84	15.23	17.91	17.39
6	53.92	59.90	18.47	18.97	22.79	22.36
7	77.65	77.23	19.95	23.72	30.75	28.81
8	96.51	101.12	28.01	30.43	35.79	37.58
9	132.95	141.49	42.09	43.31	47.42	52.26
10	627.36	606.99	213.22	209.58	293.58	283.14

[a]Calibration data of the UK Biobank atherosclerotic cardiovascular disease risk prediction model on different types of dementia at all incident times. The 5-fold cross-validation strategy was performed to calculate the results. A *P* value less than .05 indicated the statistical significance of the results.

[b]Goodness-of-fit *P* value >.99.

[c]AD: Alzheimer disease.

[d]VD: vascular dementia.

A Temporal Validation of the Constructed Model

To assess the stability and generalizability of our dementia prediction model, we used a time validation approach using data from the UK Biobank. The model was trained and validated on samples recruited between 2006 and 2009 and tested on the 2010 cohort. In the training and validation set, the model achieved a mean AUC of 0.866 (SD 0.027) in predicting all incident dementia, indicating strong discriminatory ability. When applied to the test set from 2010, the model maintained robust performance with an AUC of 0.819, suggesting good generalization to future data. Other performance metrics in the test set: accuracy was 0.851, sensitivity was 0.691, specificity was 0.866, precision was 0.315, and the F_1-score was 0.433 (Table 6).

Table 6. Model performance metrics for the prediction on the test data divided by time period[a].

	Accuracy	Sensitivity	Specificity	Precision	F_1-score	AUC[b]
All incident dementia	0.851	0.691	0.866	0.315	0.433	0.819
All incident AD[c]	0.921	0.500	0.933	0.182	0.267	0.718
All incident VD[d]	0.916	0.619	0.926	0.218	0.322	0.838
10-Year incident dementia	0.925	0.576	0.943	0.350	0.435	0.824
10-Year incident AD	0.916	0.517	0.922	0.093	0.157	0.804
10-Year incident VD	0.946	0.590	0.953	0.207	0.307	0.834
5-Year incident dementia	0.984	0.480	0.991	0.4	0.436	0.882
5-Year incident AD	0.181	1	0.181	0.001	0.002	0.605
5-Year incident VD	0.002	1	0	0.002	0.004	0.595

[a]Cutoffs were established by maximizing the Youden index (YI=sensitivity+specificity–1).

[b]AUC: area under the receiver operating characteristic curve.

[c]AD: Alzheimer disease.

[d]VD: vascular dementia.

XSL•FO
RenderX

Discussion

Principal Findings and Comparisons With Prior Work

In this study, we developed a predictive model using the LightGBM algorithm and leveraging big data from the UK Biobank to assess the risk of dementia in patients with ASCVD. To the best of our knowledge, this is the first model that uses big data to predict the risk of dementia specifically in patients with ASCVD. Our model incorporates 10 clinical predictive factors, selected based on their importance, to accurately estimate the risk of dementia. Notably, our model demonstrates particularly strong performance in predicting all-cause dementia and VD, with AUC values exceeding 0.85. Furthermore, the model effectively calibrates the predicted probabilities and aligns well with the observed event ratios, indicating its reliability and accuracy in estimating the risk of dementia in patients with ASCVD.

The ASCVD has long been recognized as one of the most significant risk factors for dementia, especially VD [22]. Although previous studies have primarily concentrated on the influence of atherosclerotic CeVDs on dementia, emerging research suggests that systemic atherosclerotic diseases beyond CeVDs also significantly contribute to the development of dementia [23]. To the best of our knowledge, no prior studies have used big data to predict the risk of dementia in patients with ASCVD. Unlike models based on variables obtained from intricate neuroimaging and neuropsychological tests, the predictors in this model are more accessible and can be applied in various clinical settings and medical institutions.

Recent studies have indicated the importance of vascular risk factors in the development of dementia, which should be taken into consideration by clinical practitioners. During the establishment of a dementia risk model in the population with ASCVD, we identified several key and distinctive risk factors that differ from other longitudinal studies. Age was identified as one of the most significant influencing factors in this study. Other factors include low exercise time, high fasting blood glucose, and reaction time including pair matching time and mean time to correctly identify matches. These risk factors have been shown to have a close relationship with the overall health of the vascular system and are critical in the development of ASCVD events [24,25]. Besides, high plasma levels of C-reactive protein at baseline were associated with a high risk of all incident dementia in this study, which is corresponded with the result of the latest research [26]. Furthermore, we found that MSCV should be considered as a new and significant factor in assessing the risk of dementia in patients with ASCVD. MSCV is primarily a parameter in hematology used to assess changes in the volume of spherocytes. Currently, there is no evidence to suggest a direct relationship between MSCV and dementia. However, overall blood health can indirectly affect cognitive function, especially in the presence of chronic anemia or other systemic diseases [27]. If an abnormal MSCV is observed in clinical practice or research, it is important to consider the patient's overall health status comprehensively, including but not limited to neurological function, to fully evaluate the patient's dementia risk. Recent studies have also gradually found that lung function may play a role in the onset of dementia by influencing brain structure [28]. Our model indicated FVC and forced expiratory volume in 1 second *z* score as protective factors in decreasing dementia risk. Finally, age first had sexual intercourse, also known as age at first intercourse (AFS), was first identified as one of the significant risk factors for dementia in populations with ASCVD, which is interesting and worth attention. Compared to other biological traits, reproductive behaviors, especially sexual factors, have long been neglected when it comes to the study of CVDs and neurological disorders. Recent studies show that the earlier the age of first sexual intercourse, the higher the likelihood of developing hypertension and CVDs [29,30]. The specific processes driving this relationship are not yet fully understood, but they may include a mix of environmental and genetic influences. For instance, early sexual activities are often accompanied by adverse environmental factors, including lower educational attainment, increased smoking and alcohol consumption, and the use of illicit drugs [31], which are all closely associated with CVDs. Recent studies have identified a causal relationship between AFS and CVDs at the genetic level [30,32]. Although AFS has no direct impact on dementia, it might be induced from our study that AFS is indirectly related to dementia through intriguing CVDs especially in patients with ASCVD.

In our study, we performed a stability analysis of feature importance using 5-fold cross-validation to ensure the robustness of the identified predictors. The analysis revealed that several key features consistently ranked highly across all folds, indicating their strong and reliable association with the outcome. Specifically, FVC, summed metabolic equivalent minutes per week, age, and pair matching time maintained relatively high importance scores in every subset, underscoring their role as robust predictors. Although these 4 variables exhibited significant variation compared to other variables, the impact of this variation is minimal relative to their importance.

We acknowledge that the Cox proportional hazards model, as a mature and interpretable method, holds a significant position in survival analysis. However, our study results show that the LightGBM-based ML model significantly outperforms the Cox model in predictive performance metrics such as AUC, demonstrating its advantages in handling nonlinear relationships and complex interactions between variables. LightGBM effectively captures patterns in high-dimensional data, thereby enhancing the accuracy of risk prediction. Although ML models face certain challenges in terms of computational resources and interpretability, their substantial improvement in predictive performance illustrates their added value in practical applications. Future research could explore combining traditional Cox models with ML methods to balance predictive performance and model interpretability, thereby meeting diverse clinical application needs. Besides, our study indicated that mortality had a minimal impact on the primary outcomes, and the overall conclusions of the study remained largely unchanged. This suggested that the prediction power of our model remained robust even after accounting for competing risks.

The time validation results demonstrate that our dementia prediction model maintains robust performance across different

XSL•FO
RenderX

temporal cohorts. Specifically, in the population of all incident dementia, the AUC remained consistently high, with a slight decrease from 0.866 in the training set to 0.819 in the test set. Similarly, other performance metrics such as accuracy, precision, recall, and F_1-score showed only minor declines over time. This stability suggests that the model effectively captures underlying patterns associated with dementia risk that are persistent across the studied time periods. The consistency of performance metrics across the training and test sets indicates that the model's predictive capabilities are not significantly affected by temporal shifts in the data. The stable performance of the model over different time periods enhances its long-term applicability in clinical and public health settings. A model that maintains its predictive accuracy over time is invaluable for ongoing and future dementia screening programs, enabling early identification of at-risk individuals with confidence in its sustained reliability. However, it is also essential to acknowledge the significant performance decline observed in the test set when predicting the risk of 5-year AD and VD, which may be attributed to the low prevalence in patients with ASCVD during a relative short period. To ensure continued efficacy, periodic retraining and validation of the model with new data may be necessary. This approach would allow the model to adapt to any emerging trends or shifts in risk factors that may influence dementia incidence over time.

In clinical practice, the LightGBM model can be applied during the initial diagnosis or follow-up stages to early identify individuals with high dementia risk among patients with ASCVD, thus promoting timely intervention and treatment. For instance, using this predictive model during a patient's initial visit can assist physicians in swiftly identifying patients at high risk and arranging further diagnostic tests or interventions. Moreover, integrating the model's predictions into electronic health record systems can generate alerts and recommend further evaluations, thereby enhancing diagnostic accuracy and personalized treatment plans. To effectively communicate the predicted risk, doctors should use easily understandable language to explain the model results and their implications to the patient while also providing clear next steps and support resources to alleviate patient anxiety. Consider a hypothetical example: a 55-year-old male patient with hypertension and high cholesterol who recently experienced a heart attack. After using the LightGBM model for assessment, the results indicate a high dementia risk. Based on this, the doctor decides to schedule detailed cognitive function tests and recommends a comprehensive plan that includes cognitive training, a healthy diet, and regular exercise. Through these interventions, the patient can better manage his cardiovascular health while taking steps to reduce the likelihood of developing dementia.

However, the application of such predictive models raises potential ethical issues. First, there may be prediction bias due to training data, leading to unequal care, so continuous monitoring and validation are needed to ensure fairness. Second, doctors need to carefully communicate the model's predictions to avoid causing unnecessary anxiety for patients. Additionally, patient data use should require clear consent and ensure privacy protection. Finally, caution against overreliance on model predictions is necessary, with doctors maintaining primary responsibility for care decisions.

Limitations

Several limitations should be considered when interpreting the results. First, our study focused on a specific population of patients with ASCVD. Due to the relatively small sample size, we observed a lower AUC value when predicting the 5-year incidence rates of dementia, especially for AD and VD incidence over a 5-year period. This issue can be addressed by further expanding the sample size. Additionally, this study primarily used samples of European descent, which may restrict the generalizability of our findings to other populations. Genetic, environmental, and lifestyle differences across diverse ethnic groups could influence the model's performance and predictive accuracy. The limited diversity of the sample may affect the model's applicability to non-European populations. To ensure broader relevance and robustness, future research should include diverse ethnic backgrounds to validate and potentially refine the model for varied demographic groups. While the time validation results are promising, the model currently relies on static features collected at baseline. Integrating longitudinal data and time-varying covariates could potentially improve predictive performance and adaptability over extended periods. Despite incorporating death as a primary competing risk, there might still be other unrecognized or unadjusted competing factors, such as other chronic diseases or lifestyle changes, that could influence the results to some extent. The application of competing risk models relied on the correct specification of models and assumptions; any biases in model setup might affect the accuracy of the analysis. Therefore, future research should further explore additional potential competing risk factors and use more sophisticated statistical methods to comprehensively assess the prediction power. Regarding the ML algorithms chosen for our study, the use of LightGBM may lead to data overfitting because it generates deep decision trees. To mitigate overfitting, a maximum depth limit should be imposed during the use of LightGBM. Furthermore, it is important to acknowledge that LightGBM is a bias-based algorithm and can be sensitive to noise in data processing, which may potentially affect the final data analysis results. Additionally, it should be noted that the predictive variables identified in this study were derived from data-driven analytical models, which may induce some bias compared to actual clinical diagnostic and treatment experiences. While advanced predictive models and results have been obtained, their applicability to clinical practice remains uncertain. Therefore, future research should focus on validating the analysis results using other independent cohorts with larger sample sizes and extending the study methodology to populations from different countries, regions, and ethnicities. The integration of clinical practice experiences will contribute to the development of more universally applicable and practical models.

Conclusions

This study has identified several practical and novel predictors for dementia screening in patients with ASCVD. It is worthy of testing and evaluating the applicability of these factors in clinical practice. Future studies should focus on investigating

XSL•FO
RenderX

whether intervening in these factors can help prevent the incidence of dementia in patients with ASCVD. By exploring these possibilities, we can potentially improve the management and outcomes of patients with ASCVD and reduce the burden of dementia in this population.

Acknowledgments

This study was funded by grants from the Natural Science Foundation of Shanghai (23JS1400500 and 22ZR1420500), the Shanghai ShenKang Hospital Development Center (SHDC2020CR1038B), and 234 discipline peak climbing program of Shanghai Changhai Hospital (2019YXK013). The funding sources were not involved in the study's design and execution; data collection, management, analysis, and interpretation; manuscript drafting, review, or approval; nor the decision to submit the manuscript for publication. This study was conducted by the use of the UK Biobank Resource (application 97366). The authors appreciate all the participants and researchers from the UK Biobank.

Data Availability

All data used during this study are available from the UK Biobank Resource. As the datasets are the property of UK Biobank, they are not available for direct request. Researchers interested in the data can apply for access through the UK Biobank Access Management System.

Authors' Contributions

XS, XB, and BD designed the project and were responsible for the integrity of data and accuracy of analysis. HM, SL, and YL played a major role in the model's development and validation as well as statistical analysis. ZG, BD, XJ, and XG contributed to the discussion of results and drafted the manuscript. All authors contributed to the paper and took responsibility for submitting it for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
UK Biobank atherosclerotic cardiovascular disease study supplementary tables, figures, and methodological details.
[DOCX File , 5010 KB - aging_v8i1e64148_app1.docx]

References

1. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, et al. Global burden of cardiovascular diseases and risk factors, 1990-2019: update from the GBD 2019 study. J Am Coll Cardiol 2020;76(25):2982-3021 [FREE Full text] [doi: 10.1016/j.jacc.2020.11.010] [Medline: 33309175]
2. GBD 2017 Causes of Death Collaborators. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018;392(10159):1736-1788 [FREE Full text] [doi: 10.1016/S0140-6736(18)32203-7] [Medline: 30496103]
3. Zhao D, Liu J, Wang M, Zhang X, Zhou M. Epidemiology of cardiovascular disease in China: current features and implications. Nat Rev Cardiol 2019;16(4):203-212. [doi: 10.1038/s41569-018-0119-4] [Medline: 30467329]
4. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet 2020;396(10248):413-446 [FREE Full text] [doi: 10.1016/S0140-6736(20)30367-6] [Medline: 32738937]
5. Winstein CJ, Stein J, Arena R, Bates B, Cherney LR, Cramer SC, et al. Guidelines for adult stroke rehabilitation and recovery: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2016;47(6):e98-e169. [doi: 10.1161/STR.0000000000000098] [Medline: 27145936]
6. Sabia S, Dugravot A, Dartigues JF, Abell J, Elbaz A, Kivimäki M, et al. Physical activity, cognitive decline, and risk of dementia: 28 year follow-up of Whitehall II cohort study. BMJ 2017;357:j2709 [FREE Full text] [doi: 10.1136/bmj.j2709] [Medline: 28642251]
7. Sabia S, Fayosse A, Dumurgier J, Schnitzler A, Empana J, Ebmeier KP, et al. Association of ideal cardiovascular health at age 50 with incidence of dementia: 25 year follow-up of Whitehall II cohort study. BMJ 2019;366:l4414 [FREE Full text] [doi: 10.1136/bmj.l4414] [Medline: 31391187]
8. Schwarzinger M, Pollock BG, Hasan OSM, Dufouil C, Rehm J, QalyDays Study Group. Contribution of alcohol use disorders to the burden of dementia in France 2008-13: a nationwide retrospective cohort study. Lancet Public Health 2018;3(3):e124-e132 [FREE Full text] [doi: 10.1016/S2468-2667(18)30022-7] [Medline: 29475810]

XSL•FO
RenderX

9. Veronese N, Facchini S, Stubbs B, Luchini C, Solmi M, Manzato E, et al. Weight loss is associated with improvements in cognitive function among overweight and obese people: a systematic review and meta-analysis. Neurosci Biobehav Rev 2017;72:87-94 [FREE Full text] [doi: 10.1016/j.neubiorev.2016.11.017] [Medline: 27890688]
10. Nordestgaard LT, Christoffersen M, Frikke-Schmidt R. Shared risk factors between dementia and atherosclerotic cardiovascular disease. Int J Mol Sci 2022;23(17):9777 [FREE Full text] [doi: 10.3390/ijms23179777] [Medline: 36077172]
11. Deo RC. Machine learning in medicine. Circulation 2015;132(20):1920-1930 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.115.001593] [Medline: 26572668]
12. James C, Ranson JM, Everson R, Llewellyn DJ. Performance of machine learning algorithms for predicting progression to dementia in memory clinic patients. JAMA Netw Open 2021;4(12):e2136553 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.36553] [Medline: 34913981]
13. Zhan Y, Chen K, Wu X, Zhang D, Zhang J, Yao L, et al. Identification of conversion from normal elderly cognition to Alzheimer's disease using multimodal support vector machine. J Alzheimers Dis 2015;47(4):1057-1067 [FREE Full text] [doi: 10.3233/JAD-142820] [Medline: 26401783]
14. Park JH, Cho HE, Kim JH, Wall MM, Stern Y, Lim H, et al. Machine learning prediction of incidence of Alzheimer's disease using large-scale administrative health data. NPJ Digit Med 2020;3:46 [FREE Full text] [doi: 10.1038/s41746-020-0256-0] [Medline: 32258428]
15. Cui Y, Liu B, Luo S, Zhen X, Fan M, Liu T, et al. Identification of conversion from mild cognitive impairment to Alzheimer's disease using multivariate predictors. PLoS One 2011;6(7):e21896 [FREE Full text] [doi: 10.1371/journal.pone.0021896] [Medline: 21814561]
16. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015;12(3):e1001779 [FREE Full text] [doi: 10.1371/journal.pmed.1001779] [Medline: 25826379]
17. Hageman SHJ, McKay AJ, Ueda P, Gunn LH, Jernberg T, Hagström E, et al. Estimation of recurrent atherosclerotic cardiovascular event risk in patients with established cardiovascular disease: the updated SMART2 algorithm. Eur Heart J 2022;43(18):1715-1727 [FREE Full text] [doi: 10.1093/eurheartj/ehac056] [Medline: 35165703]
18. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD statement. BMJ 2015;350:g7594 [FREE Full text] [doi: 10.1136/bmj.g7594] [Medline: 25569120]
19. Wang J, Qiu J, Zhu T, Zeng Y, Yang H, Shang Y, et al. Prediction of suicidal behaviors in the middle-aged population: machine learning analyses of UK Biobank. JMIR Public Health Surveill 2023;9:e43419 [FREE Full text] [doi: 10.2196/43419] [Medline: 36805366]
20. Li Q, Yang X, Xu J, Guo Y, He X, Hu H, et al. Early prediction of Alzheimer's disease and related dementias using real-world electronic health records. Alzheimers Dement 2023;19(8):3506-3518 [FREE Full text] [doi: 10.1002/alz.12967] [Medline: 36815661]
21. Martin SA, Townend FJ, Barkhof F, Cole JH. Interpretable machine learning for dementia: a systematic review. Alzheimers Dement 2023;19(5):2135-2149 [FREE Full text] [doi: 10.1002/alz.12948] [Medline: 36735865]
22. Iadecola C. Revisiting atherosclerosis and dementia. Nat Neurosci 2020;23(6):691-692 [FREE Full text] [doi: 10.1038/s41593-020-0626-6] [Medline: 32424283]
23. Zlokovic BV, Gottesman RF, Bernstein KE, Seshadri S, McKee A, Snyder H, et al. Vascular contributions to cognitive impairment and dementia (VCID): a report from the 2018 National Heart, Lung, and Blood Institute and National Institute of Neurological Disorders and Stroke Workshop. Alzheimers Dement 2020;16(12):1714-1733. [doi: 10.1002/alz.12157] [Medline: 33030307]
24. Meyer-Lindemann U, Moggio A, Dutsch A, Kessler T, Sager HB. The impact of exercise on immunity, metabolism, and atherosclerosis. Int J Mol Sci 2023;24(4):3394 [FREE Full text] [doi: 10.3390/ijms24043394] [Medline: 36834808]
25. Stone NJ, Smith SC, Orringer CE, Rigotti NA, Navar AM, Khan SS, et al. Managing atherosclerotic cardiovascular risk in young adults: JACC state-of-the-art review. J Am Coll Cardiol 2022;79(8):819-836 [FREE Full text] [doi: 10.1016/j.jacc.2021.12.016] [Medline: 35210038]
26. Tachibana A, Iga JI, Ozaki T, Yoshida T, Yoshino Y, Shimizu H, et al. Serum high-sensitivity C-reactive protein and dementia in a community-dwelling Japanese older population (JPSC-AD). Sci Rep 2024;14(1):7374 [FREE Full text] [doi: 10.1038/s41598-024-57922-1] [Medline: 38548879]
27. Qiang YX, Deng YT, Zhang YR, Wang H, Zhang W, Dong Q, et al. Associations of blood cell indices and anemia with risk of incident dementia: a prospective cohort study of 313,448 participants. Alzheimers Dement 2023;19(9):3965-3976. [doi: 10.1002/alz.13088] [Medline: 37102212]
28. Ma YH, Shen LX, Li YZ, Leng Y, Yang L, Chen S, et al. Lung function and risk of incident dementia: a prospective cohort study of 431,834 individuals. Brain Behav Immun 2023;109:321-330. [doi: 10.1016/j.bbi.2023.02.009] [Medline: 36796705]
29. Ngueta G, Ndjaboue R. Early sexual experience and hypertension in US adults: results from the National Health and Nutrition Examination Survey 2001-2016. J Hypertens 2018;36(12):2414-2419. [doi: 10.1097/HJH.0000000000001821] [Medline: 29957720]

XSL•FO
RenderX

30. Zhuo C, Chen L, Wang Q, Cai H, Lin Z, Pan H, et al. Association of age at first sexual intercourse and lifetime number of sexual partners with cardiovascular diseases: a bi-directional Mendelian randomization study. Front Cardiovasc Med 2023;10:1267906 [FREE Full text] [doi: 10.3389/fcvm.2023.1267906] [Medline: 38146444]
31. Boisvert I, Boislard MA, Poulin F. Early sexual onset and alcohol use and misuse from adolescence into young adulthood. J Adolesc Health 2017;61(4):514-520. [doi: 10.1016/j.jadohealth.2017.04.013] [Medline: 28732719]
32. Mills MC, Tropf FC, Brazel DM, van Zuydam N, Vaez A, Pers TH, et al. Identification of 371 genetic variants for age at first sex and birth linked to externalising behaviour. Nat Hum Behav 2021;5(12):1717-1730 [FREE Full text] [doi: 10.1038/s41562-021-01135-3] [Medline: 34211149]

Abbreviations

AD: Alzheimer disease
AFS: age at first intercourse
ASCVD: atherosclerotic cardiovascular disease
AUC: area under the receiver operating characteristic curve
CeVD: cerebrovascular disease
CVD: cardiovascular disease
DCA: decision curve analysis
FVC: forced vital capacity
LightGBM: light gradient boosting machine
ML: machine learning
MSCV: mean sphered cell volume
SFS: sequential forward selection
SHAP: Shapley Additive Explanations
VD: vascular dementia

Edited by G Seçkin; submitted 10.07.24; peer-reviewed by T Yang, L Mercado Diaz; comments to author 22.08.24; revised version received 31.10.24; accepted 20.01.25; published 26.02.25.

Please cite as:
Gu Z, Liu S, Ma H, Long Y, Jiao X, Gao X, Du B, Bi X, Shi X
Estimation of Machine Learning–Based Models to Predict Dementia Risk in Patients With Atherosclerotic Cardiovascular Diseases: UK Biobank Study
JMIR Aging 2025;8:e64148
URL: https://aging.jmir.org/2025/1/e64148
doi:10.2196/64148
PMID:

©Zhengsheng Gu, Shuang Liu, Huijuan Ma, Yifan Long, Xuehao Jiao, Xin Gao, Bingying Du, Xiaoying Bi, Xingjie Shi. Originally published in JMIR Aging (https://aging.jmir.org), 26.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Development and Validation of a Brain Aging Biomarker in Middle-Aged and Older Adults: Deep Learning Approach

Zihan Li[1*], MS; Jun Li[2*], MD; Jiahui Li[1*], BM; Mengying Wang[1], MM; Andi Xu[1], MM; Yushu Huang[1], MM; Qi Yu[1], MM; Lingzhi Zhang[1], MM; Yingjun Li[3], PhD; Zilin Li[4], PhD; Xifeng Wu[1,5], MD, PhD; Jiajun Bu[6], PhD; Wenyuan Li[1,5], ScD

[1]Center for Clinical Big Data and Analytics, The Second Affiliated Hospital and Department of Big Data in Health Science, School of Public Health, Zhejiang University School of Medicine, Zhejiang University, Hangzhou, China

[2]Department of Radiology, The First Affiliated Hospital, Zhengzhou University, First Affiliated Hospital of Zhengzhou University, Zhengzhou, China

[3]Department of Epidemiology and Health Statistics, School of Public Health, Hangzhou Medical College, Hangzhou Medical College, Hangzhou, China

[4]School of Mathematics and Statistics, Northeast Normal University, Northeast Normal University, Changchun, China

[5]Zhejiang Provincial Key Laboratory of Intelligent Preventive Medicine, Hangzhou, China

[6]Zhejiang Key Laboratory of Accessible Perception and Intelligent Systems, College of Computer Science and Technology, Zhejiang University, Hangzhou, China

[*]these authors contributed equally

Corresponding Author:
Wenyuan Li, ScD
Center for Clinical Big Data and Analytics, The Second Affiliated Hospital and Department of Big Data in Health Science, School of Public Health
Zhejiang University School of Medicine
Zhejiang University
Zhejiang Campus
866 Yuhangtang Road
Hangzhou, 310058
China
Phone: 86 19521623869
Email: wenyuanli@zju.edu.cn

Abstract

Background: Precise assessment of brain aging is crucial for early detection of neurodegenerative disorders and aiding clinical practice. Existing magnetic resonance imaging (MRI)–based methods excel in this task, but they still have room for improvement in capturing local morphological variations across brain regions and preserving the inherent neurobiological topological structures.

Objective: To develop and validate a deep learning framework incorporating both connectivity and complexity for accurate brain aging estimation, facilitating early identification of neurodegenerative diseases.

Methods: We used 5889 T1-weighted MRI scans from the Alzheimer's Disease Neuroimaging Initiative dataset. We proposed a novel brain vision graph neural network (BVGN), incorporating neurobiologically informed feature extraction modules and global association mechanisms to provide a sensitive deep learning–based imaging biomarker. Model performance was evaluated using mean absolute error (MAE) against benchmark models, while generalization capability was further validated on an external UK Biobank dataset. We calculated the brain age gap across distinct cognitive states and conducted multiple logistic regressions to compare its discriminative capacity against conventional cognitive-related variables in distinguishing cognitively normal (CN) and mild cognitive impairment (MCI) states. Longitudinal track, Cox regression, and Kaplan-Meier plots were used to investigate the longitudinal performance of the brain age gap.

Results: The BVGN model achieved an MAE of 2.39 years, surpassing current state-of-the-art approaches while obtaining an interpretable saliency map and graph theory supported by medical evidence. Furthermore, its performance was validated on the UK Biobank cohort (N=34,352) with an MAE of 2.49 years. The brain age gap derived from BVGN exhibited significant difference across cognitive states (CN vs MCI vs Alzheimer disease; $P<.001$), and demonstrated the highest discriminative capacity between CN and MCI than general cognitive assessments, brain volume features, and apolipoprotein E4 carriage (area under the receiver operating characteristic curve [AUC] of 0.885 vs AUC ranging from 0.646 to 0.815). Brain age gap exhibited clinical feasibility

XSL•FO
RenderX

combined with Functional Activities Questionnaire, with improved discriminative capacity in models achieving lower MAEs (AUC of 0.945 vs 0.923 and 0.911; AUC of 0.935 vs 0.900 and 0.881). An increasing brain age gap identified by BVGN may indicate underlying pathological changes in the CN to MCI progression, with each unit increase linked to a 55% (hazard ratio=1.55, 95% CI 1.13-2.13; *P*=.006) higher risk of cognitive decline in individuals who are CN and a 29% (hazard ratio=1.29, 95% CI 1.09-1.51; *P*=.002) increase in individuals with MCI.

Conclusions: BVGN offers a precise framework for brain aging assessment, demonstrates strong generalization on an external large-scale dataset, and proposes novel interpretability strategies to elucidate multiregional cooperative aging patterns. The brain age gap derived from BVGN is validated as a sensitive biomarker for early identification of MCI and predicting cognitive decline, offering substantial potential for clinical applications.

(JMIR Aging 2025;8:e73004) doi:10.2196/73004

KEYWORDS

brain aging; deep learning; magnetic resonance imaging; MRI; imaging biomarker

Introduction

Background

As human life expectancy steadily rises, the global population aged ≥60 years is projected to nearly double by 2050, reaching an estimated 2.1 billion [1]. This demographic shift presents substantial burdens to society, including rising health care costs and caregiving demands [2]. In the area of aging, dementia accounts for 11.2% of disability years, surpassing that caused by strokes (9.5%), musculoskeletal disorders (8.9%), cardiovascular diseases (5%), and all types of cancer (2.4%) [3]. Scientific investigations have illuminated the association between an aging brain and neurodegenerative conditions, such as Alzheimer disease (AD), by revealing the molecular and cellular mechanisms, particularly mitochondrial impairment [4,5]. There is an urgent demand for accurate assessment of brain aging to enable early detection of age-related neurodegenerative diseases, with neuroimaging-based brain age estimation methods gaining popularity in recent years [6].

Given its exceptional soft tissue contrast, noninvasive nature, and multimodal capabilities, magnetic resonance imaging (MRI) is the preferred method for timely visualizing brain structure and subtle lesions [7,8]. Brain age gap estimation [9], a quantified MRI-based score, has been considered a reliable brain age scale [10,11], with numerous studies associating a positive brain age gap with brain abnormalities and mortality [12,13]. The typical approach in this field involves constructing a standardized model using structural MRI data from healthy individuals. This model is then used to evaluate neuroanatomical deviations from the norm in new participants, providing insight into underlying brain abnormalities [14]. Recent publications have demonstrated the efficacy of traditional methods for brain age gap estimation, such as support vector regression and relevance vector regression, with mean absolute errors (MAEs) ranging from 2.6 years to 7.7 years [15] and 3.7 years to 4.7 years [16], respectively. However, these approaches often rely on supervised learning, with inherent limitations related to manual labeling and model performance influenced by dataset bias [14].

In recent years, the emergence and widespread adoption of deep learning (DL) techniques, which require no manual annotation and excel in handling high-dimensional data [17], have propelled model precision to unprecedented heights, achieving MAEs of <3 years [18-20]. For instance, BrainAgeNeXt [21] achieved an MAE of 2.78 years. Convolutional neural networks (CNNs) are lauded for their adeptness in automatic feature extraction from MRI images but are limited by the gradient explosion or vanishing issue in deeper layers. Residual neural networks (ResNets) have made significant strides by introducing skip connections to mitigate these training challenges, albeit at the cost of increased computational complexity. Transformer-based models represent a paradigm shift with their self-attention mechanisms adept at capturing long-range dependencies, but they demand substantial data for optimal training. Ensemble DL approaches have garnered attention for their ability to synthesize the predictions of multiple models, thereby enhancing the robustness and accuracy of brain age estimation. This prediction strategy, while beneficial for improving generalization, incurs the surcharge of increased training complexity. However, most DL-based approaches regard brain MRI as a cube constructed of voxels and then transfer models that perform well in natural images or videos directly to this task. This approach neglects the unique biological characteristics of the brain, which distinguish it from natural images [22], particularly in morphology [23]. Moreover, the human connectome exhibits intricate higher-order connectivity, encompassing spatial, functional, and sex-specific distinctions [24] among different brain voxel regions. Common CNN architectures designed for natural images lack the capacity to modify their topologies [25]. Currently, there is a notable absence of a DL framework that simultaneously considers both brain morphology and voxel regions connectivity for assessing brain aging. In addition, there is insufficient evidence to compare and quantify the clinical application value of the brain age gap and the relationship between narrowing MAEs with improving clinical application performance.

This Study

Here we proposed a novel graph-based DL framework, a brain vision graph neural network (BVGN), which uses deformable kernels to extract brain morphological features and incorporated a graph neural network (GNN) to characterize voxel regions connectivity. We pioneered the adaptation of vision GNN [26], a backbone widely researched in natural image processing, to the domain of brain aging by extending its framework to 3D MRI. The BVGN leveraged an individual's standard

XSL•FO
RenderX

T1-weighted MRI as input and generated the corresponding neuroimaging biomarker-brain age gap to evaluate brain aging. Subsequently, we constructed multiple logistics regression models, each incorporating different types of input variables, to classify data samples labeled as cognitively normal (CN) and with mild cognitive impairment (MCI). We aimed to compare and validate the effects of brain age gap against conventional variables, including demographic factors, brain volume features, apolipoprotein E4 (APOE4) carriage, and general cognitive scales, in distinguishing CN and MCI states. In addition, we assessed the combined utility of the brain age gap with the Functional Activities Questionnaire (FAQ) cognitive assessment and the impact of lower MAEs. Cox proportional hazards regressions were performed separately using data from participants classified as CN and those with MCI at baseline to evaluate the potential of brain age gap as a risk biomarker for cognitive decline over a 2-year follow-up period.

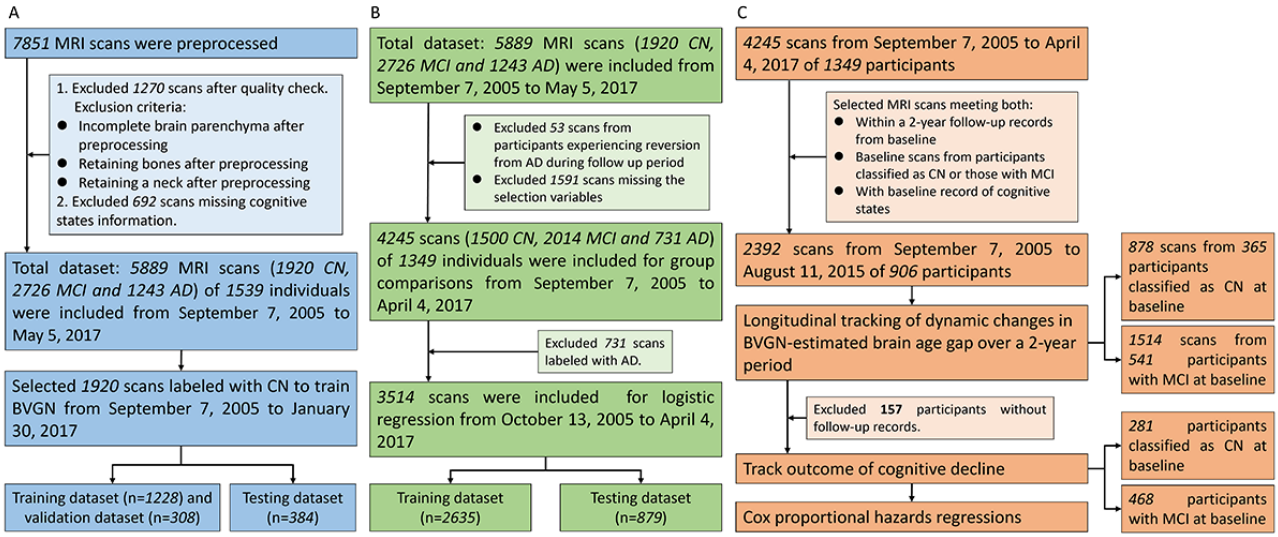
Methods

Dataset, Study Design, and Inclusion Criteria

Data used in the preparation of this paper were obtained from the database of Alzheimer's Disease Neuroimaging Initiative (ADNI) studies [27-29]. The ADNI was launched in 2003 as a public-private partnership, led by principal investigator Michael W Weiner, MD. The primary goal of ADNI has been to test whether serial MRI, positron emission tomography, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD.

This study gathered 7851 T1-weighted MRI scans from ADNI-1, ADNI-2, and ADNI-GO (Grand Opportunities) studies. The research design comprised 2 sequential phases: development and validation of the BVGN model for accurate brain age estimation, and clinical effect validation of the BVGN-estimated brain age gap metric by integrated cross-sectional and longitudinal analysis. For BVGN model development, we specifically chose 5889 scans from 1539 individuals that contained cognitive states labels and met both the ADNI MRI scanner protocols [30,31] and our strict training criteria (ie, presenting complete structural information after preprocessing). Data regarding the patient's cognitive states were retrieved from the ADNI web portal [32]. Only 1920 MRI scans of participants classified as CN were used for training the BVGN model (Figure 1A). To ensure comprehensive model development and evaluation, we divided the dataset consisting of participants classified as CN into training set with validation set and test set in a ratio of 8:2. The training set and validation set were used for model training and hyperparameter setting, while the test set was used to validate the performance of the model. A subset of the UK Biobank (UKB) dataset [33] was used for external validation to evaluate model generalizability. We curated 34,352 MRI scans, partitioned into training set (n=21,985, 64%), validation set (n=5497, 16%), and test set (n=6870, 20%).

Figure 1. Flow diagram to show study design and inclusion criteria. (A) Brain vision graph neural network (BVGN) training and testing part; (B) cross-sectional validation part; (C) longitudinal validation part. AD: Alzheimer disease; CN: cognitively normal; MCI: mild cognitive impairment; MRI: magnetic resonance imaging.



Given the progressive nature of dementia caused by neurodegenerative disorders, we excluded scans from participants who experienced reversion from AD during the follow-up period (Figure 1B). Following feature selection, 4245 MRI scans from 1349 participants were retained for cross-sectional analysis. All scans were used for intergroup comparisons, while only those from participants classified as CN and those with MCI were used for logistic regression analysis aiming at binary classification between CN and MCI. To validate the performance of logistic regression models, the dataset was split into training and test sets at an 8:2 ratio.

We combined various MRI scans from the same individual for longitudinal analysis (Figure 1C). Specifically, we used a dual-component longitudinal design to investigate the trajectories of BVGN-estimated brain age gap across cognitive progression groups and quantify cognitive decline risks through Cox proportional hazard models. For the brain age gap trajectory analysis, cognitive progression categories (CN-CN, CN-MCI or CN-AD, MCI-CN or MCI-MCI, and MCI-AD) were stratified

XSL•FO
RenderX

based on recorded cognitive states from baseline to a 2-year follow-up. The inclusion criteria required participants to be classified as either CN or having MCI at the baseline, yielding 1514 scans from 541 participants with MCI and 878 scans from 365 participants classified as CN.

To evaluate time-dependent cognitive decline, we conducted Cox proportional hazards regression stratified by baseline groups. We excluded 157 participants without follow-up data and performed separate analyses for the baseline CN group (tracking MCI conversion) and MCI group (tracking AD conversion). The final analysis included 749 participants, comprising 281 (37.5%) individuals classified as CN and 468 (62.5%) individuals with MCI, with the time-to-event defined as the interval between baseline MRI and first observed cognitive transition.

Quality Control and Data Preprocessing

For the ADNI dataset, each series within every examination was subjected to a rigorous quality control process at the Mayo Clinic. This involved 2 distinct levels of scrutiny—compliance with protocol-specific parameters and an assessment of the series-specific quality, such as the participant's movement and the extent of anatomical coverage. The quality of the scans was evaluated by trained analysts who assigned a subjective grade—scores of 1 to 3 were considered satisfactory, while a score of 4 indicated a failure, rendering the scans unusable. While the UKB dataset was only used to test the model's generalizability, all data also underwent standardized quality control procedures [34].

The acquisition protocols of different datasets required data preprocessing to ensure compatibility. The entire MRI preprocessing pipeline could be divided into 4 steps. First, the raw MRI data in neuroimaging informatics technology initiative format underwent minimum-maximum normalization. Second, nonbrain tissue, such as the skull and neck, was removed. In the third step, each MRI was registered from its native space to the standard Montreal Neurological Institute 152 1-mm^3 template [35] chosen for registration in our study. Finally, central cropping was performed on the complete MRI volume, resulting in a final output size of 160×196×160 mm^3. In addition, further preprocessing steps were conducted to exclude MRIs with incomplete brain structures, ensuring dataset quality. The preprocessing pipeline was implemented using SimpleITK [36] and FMRIB software library [37].

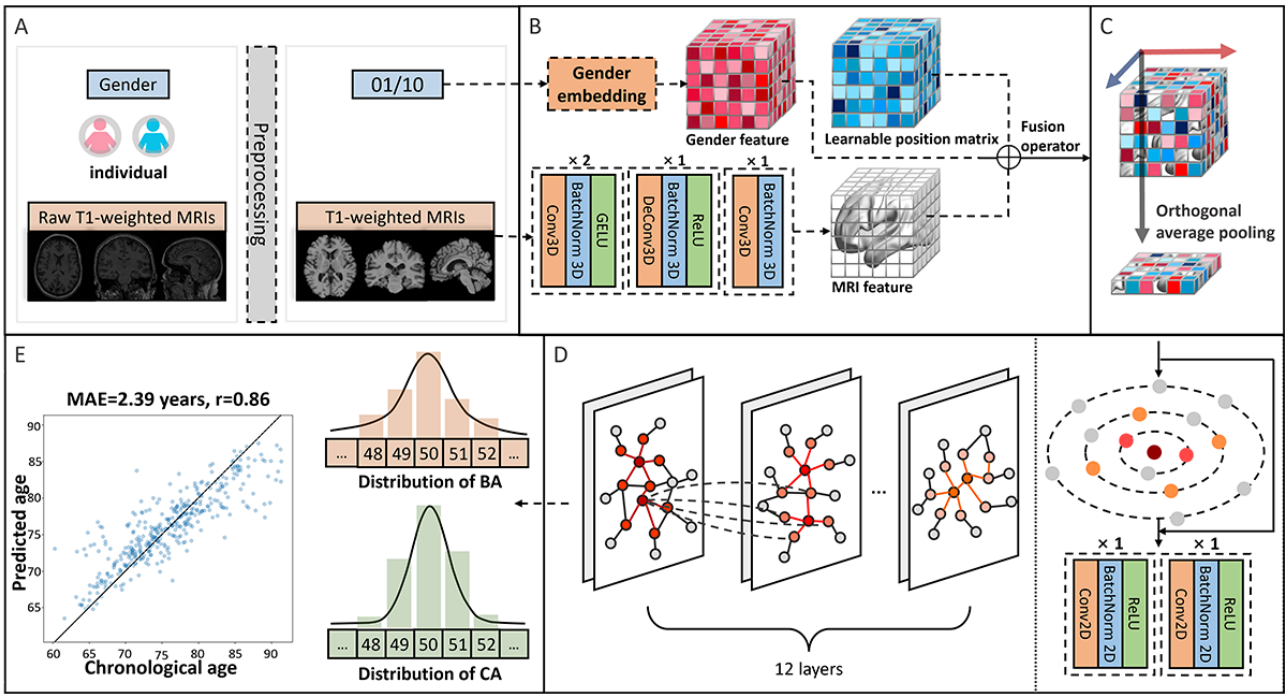
During the logistic and Cox regression modeling phase, *z* score normalization was applied to numerical variables, while categorical variables were transformed using one-hot encoding.

BVGN Model Development

Owing to the need for a large number of high-quality handcrafted features in traditional methods [14] and because inductive bias to brain structures is not effectively transferable from natural images when applying classical DL architectures, there are limitations in extracting features from irregular objects, such as cortical sulcus and gyrus, as well as parenchyma using grid-like convolution kernels that perform weighted summation of pixel values at the centroid and its neighbors [25,38,39]. Furthermore, CNN architectures lack the ability to modify topological relationships between voxel regions [22]. To address these issues, we proposed an end-to-end framework called BVGN (Figure 2), which used standardized MRIs to estimate brain aging of individuals. To be more specific, BVGN makes it easier to extract irregularly shaped structures in brain MRI by superimposing deformable convolution kernels. A graph CNN was used to model the relationship between different regions.

XSL•FO
RenderX

Figure 2. Overview of the brain vision graph neural network model. (A) A preprocessing pipeline was implemented, which involved converting gender into 1-hot encoding and transforming raw magnetic resonance imaging (MRI) data into brain tissue located in a standard space. (B) Discretization layer, the voxel space of brain is divided by the discretization layer, which uses a combination of multiple convolution modules and deformable convolution modules. Subsequently, positional features and gender features are integrated into cubes. (C) Orthogonal average pooling module (coronal axis, sagittal axis, transverse axis, the 3D feature cube is transformed into a 2D feature map). (D) Stacked graph neural network modules. (E) During training, the image on the right illustrates that the supervised signal of back-propagation is computed by evaluating the Kullback-Leibler divergence between estimated brain age and chronological age; whereas the image on the left depicts the model's performance. BA: brain age; CA: chronological age.



The BVGN framework consists of 3 components: discretization layer, stacked GNN layer, and aging distribution prediction header. Specifically, standardized MRIs were discretized into multiple cubes using nonoverlapping multiple deformable convolutional modules [38], where each cube's feature vectors contained all information within the receptive field range of deformable convolutional modules. In addition, an individual's sex and position information for each cube was incorporated into the feature map by passing through a projection module, modal-type embedding [40] and initializing a learnable parameter matrix [41]. An orthogonal average pooling module was implemented to project the 3D volumetric feature tensor along 3 mutually perpendicular anatomical axes (coronal, sagittal, and transverse). This dimensionality reduction operation transforms the 3D spatial representation into a 2D feature map while preserving critical spatial information across orthogonal planes. Detailed explanations of Figure 2 are provided in Multimedia Appendix 1.

The feature map was subsequently transformed into a graph using Euclidean distance and nearest neighbors as the basis, which was then input into a stacked GNN [26,42]. Importantly, the dynamic construction of the graph during forward processing relied on assessing similarity between feature maps at different depths. Ultimately, aging probability distributions [43,44] were obtained in the prediction header by integrating the feature map through bottlenecked convolutional kernels [45], based on predefined age ranges. Inspired by vision GNN [26], we implemented 2 distinct neural network architectures—the pyramid-shaped BVGN and the isotropic BVGN. The isotropic BVGN was used for interpretability analyses, while the pyramid-shaped BVGN was used to enhance overall performance. Both architectures were initiated with an identical discretization module featuring the same number of nodes. However, as the network deepened, the isotropic BVGN maintained a constant node count, ensuring uniform information flow akin to the transformer model's approach. In contrast, the pyramid-shaped BVGN underwent a process of node consolidation, progressively reducing the number of nodes to form a pyramidal structure, echoing the hierarchical feature extraction of CNNs in computer vision.

BVGN Model Training

The BVGN model was trained using multiple graphics processing units to ensure efficient computation. For optimization purposes, we implemented the stochastic gradient descent optimizer with an initial learning rate of 0.01, aiming to minimize the Kullback-Leibler divergence between the predicted age distribution and label age distribution. To prevent overfitting, an L2 weight decay coefficient value of 0.001 was applied during training sessions lasting up to 160 epochs. Furthermore, a batch size of 32 was adopted for improved performance, while progressively decreasing the learning rate by multiplying it with a factor of 0.7 after every consecutive span of 30 epochs.

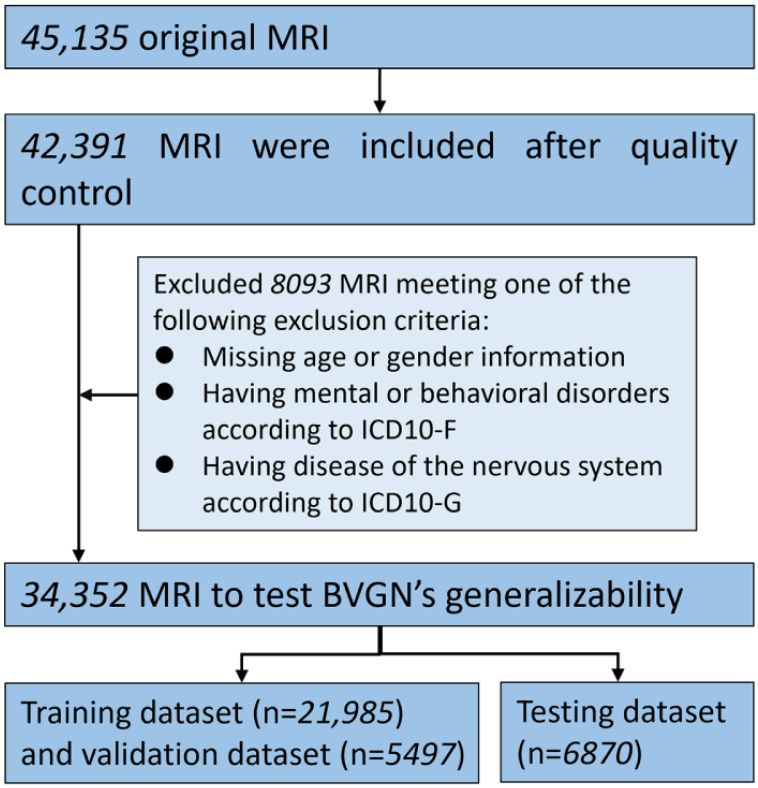
BVGN Model Performance and Generalizability Evaluation

To objectively assess the performance of our framework, we used MAE and Pearson correlation as validation metrics to evaluate the efficacy of BVGN as a prediction model. Furthermore, we simultaneously trained and tested the simple

XSL•FO
RenderX

fully convolutional network (SFCN) and ResNet models on the same dataset, ensuring a fair comparison. To validate the model's generalizability, we retrained it using the UKB dataset, which is large scale and not restricted to disease-specific populations. The detailed inclusion process and demographic characteristics of the included UKB dataset are presented in Figure 3 and Table S1 in Multimedia Appendix 1.

Figure 3. Flow diagram to show inclusion criteria for brain vision graph neural network's (BVGN's) generalizability. ICD: International Classification of Diseases; MRI: magnetic resonance imaging.



BVGN Model Interpretation

In contrast to models exclusively relying on CNNs, BVGN incorporates graph convolutional kernels to modify the relationships among local regions. This enables BVGN to offer interpretability not only through saliency maps but also by considering pertinent properties of the graph structure. More specifically, we first visualized a gradience-based saliency map [45] for participants classified as CN and those with MCI and AD. Subsequently, certified radiologists used the automated anatomical labelling atlas to ascertain the specific brain regions to which the saliency map corresponded. Furthermore, we conceptualized the brain as a graph-based complex system and characterized the interregional cooperative aging patterns by analyzing the connectivity of critical nodes across different hierarchical levels through data-driven visualization techniques.

Cross-Sectional Analysis: Intergroup Comparisons and Logistic Regression

During cross-sectional analysis, each MRI scan was treated as an individual sample, and all statistical analysis was conducted accordingly. Numerical variables in demographics and selection features were presented as mean (SD), while categorical variables were presented as counts (n) and percentages (%). Statistical comparisons between training and testing sets were performed with 2-tailed significance. Specifically, the Student *t* tests and a 1-way ANOVA were used for numerical variables, and chi-square tests were used for categorical variables. A $P<.05$ was considered statistically significant.

For intergroup comparisons, we first characterized the distribution of the brain age gap using the median and IQRs across CN, MCI, and AD classifications. Subsequently, we used the Kruskal-Wallis test followed by 1-tailed Dunn test to evaluate potential significant differences in the brain age gap among these groups, with a significance level set at $P<.05$. Potential confounding effects were addressed through a multiple linear regression model incorporating age and gender as covariates.

We used multiple logistic regression models aiming at binary classification between CN and MCI samples. Our objective was to validate the diagnostic ability of the brain age gap compared to other conventional variables and to assess the potential clinical value of achieving lower MAEs in brain age gap estimation. Inspired by previous studies, we considered the brain age gap along with other relevant cognitive variables, including demographic characteristics, general cognitive assessments, brain volume features, and APOE4 carriage. Given the simplicity and high discriminative and predictive power of FAQ as reported by previous studies, we specifically investigated the impact of combining the brain age gap with FAQ to determine if their integration improves diagnostic classification.

XSL•FO
RenderX

Model performance was evaluated using metrics, including accuracy, precision, recall, F_1-score, the receiver operating characteristic curves, and the area under the receiver operating characteristic curves (AUCs).

We initially established a benchmark model incorporating only demographic variables. Subsequently, brain age gap, various cognitive assessment scores, brain volume characteristics, and APOE4 carriage were added to the benchmark model to construct various logistic regression models. A comprehensive model integrating all the abovementioned variables was then developed, and their coefficients were compared. These analyses aimed to provide a holistic view in comparing the brain age gap with conventionally recognized cognition-related variables. Furthermore, we included both the FAQ and the brain age gap in the benchmark model to explore their combined utility. We also investigated the potential clinical value of lower MAEs by including diverse brain age gaps estimated from models with different precision, both in conjunction with the benchmark and comprehensive models, and compared their performances.

Longitudinal Analysis: Brain Age Gap Trajectories and Risk Quantification

We conducted a longitudinal analysis to investigate changes in brain age gap across follow-up scans of the same participants and assess the risk of brain age gap for cognitive decline. We used BVGN, SFCN, and ResNet models to estimate brain age gap across the total 2392 scans and depicted how the estimated brain age gap changed over follow-up time in different progression groups.

After excluding participants without follow-up records, we performed Cox regression using the baseline brain age gap for both the CN and MCI groups. This allowed us to evaluate the association between brain age gap and cognitive decline, as well as to quantify its hazard ratio (HR). We initially estimated the HR of brain age gap using a univariate Cox proportional hazards model, followed by adjustment for demographic features to control for potential confounding bias. Subsequently, a multivariable Cox proportional hazards model that included all selected variables was used to assess the HR of the brain age gap from a holistic view. In addition, we plotted Kaplan-Meier survival curves stratified by the upper and lower 50% levels of brain age gap in both CN and MCI groups to visualize the performance of the brain age gap in predicting cognitive decline.

Ethical Considerations

Data used in the preparation of this paper were obtained from the ADNI database. Consequently, the investigators within the ADNI contributed to the design and implementation of ADNI and provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found [46] on the internet.

Data collection and sharing for this project were funded by the ADNI (National Institutes of Health grant U01 AG024904) and DOD ADNI (Department of Defense grant W81XWH-12-2-0012).

The data used in the preparation of this paper were obtained from the UKB database. UKB has generic ethical approval from the North West Multi-centre Research Ethics Committee as a Research Tissue Bank (91486), and therefore researchers do not require separate ethical clearance to use the resource.

Results

Demographics and Selection Variable Characteristics

During the BVGN development process, the sample sizes for training, validating, and testing sets were 1228, 308, and 384, respectively. As shown in Table S2 in Multimedia Appendix 1, the mean ages for the training with validating sets and testing sets were 75.65 (SD 6.31) and 75.89 (SD 6.60) years, respectively. The percentages of male individuals in the training and validating set and testing set were 47.1% (724/1536) and 43.5% (167/384), respectively. No significant difference (P=.51) was observed.

Table 1 details the basic information of the dataset for intergroup comparison, which included 4245 MRI scans from 1349 participants. The discrepancy in participant count (1585 in Table 1) reflects instances where participants exhibited multiple cognitive manifestations during the follow-up period, leading to some being counted multiple times across groups. Specifically, the dataset comprised 1500 MRI scans from 499 participants classified as CN, 2014 scans from 709 participants with MCI, and 731 scans from 377 participants experiencing AD. The mean age was 75.39 years for participants classified as CN, 72.96 years for participants with MCI, and 74.76 years for participants with AD. The percentage of male participants was 46.1% (691/1500) for those classified as CN, 55.2% (1112/2014) for those with MCI, and 49.3% (360/731) for those with AD. Significant differences were observed between the CN, MCI, and AD groups for both age and gender characteristics.

XSL•FO
RenderX

Table 1. Demographic characteristics for the dataset of group comparisons.

	CN[a]	MCI[b]	AD[c]	*P* value
MRI[d] scans, n	1500	2014	731	—[e]
Participants[f], n	499	709	377	—
Age (y), mean (SD)	75.39 (6.06)	72.96 (7.45)	74.76 (7.53)	<.001
Gender, n (%)				<.001
Male	691 (46.1)	1112 (55.2)	360 (49.3)	
Female	809 (53.9)	902 (44.8)	371 (50.7)	

[a]CN: cognitively normal.

[b]MCI: mild cognitive impairment.

[c]AD: Alzheimer disease.

[d]MRI: magnetic resonance imaging.

[e]Not applicable.

[f]The total number of participants (n=1585) in Table 1 included instances where some participants had multiple cognitive manifestations during follow-up, leading to duplicates between groups. This total differed from the 1349 unique participants reported in Figure 1. Statistical analysis was performed by ANOVA test for age and chi-square test for gender. Two-tailed significance was used, with $P<.05$ indicating the presence of significance.

For the logistic regression modeling process, 3514 scans were included, with 2635 (74.99%) in the training set and 879 (25.01%) in the testing set. The characteristics of all variables used in the logistic regression are presented in Table 2. No significant differences were observed between training and testing sets for all the selected variables.

XSL•FO
RenderX

Table 2. Characteristics of selected variables in the training and testing datasets for logistic regression.

Variables	Total dataset (n=3514)	Training dataset (n=2635)	Testing dataset (n=879)	*P* value
Education (y), mean (SD)	16.10 (2.81)	16.09 (2.82)	16.12 (2.79)	.77
ADAS13[a] (points), mean (SD)	12.50 (7.07)	12.56 (7.12)	12.32 (6.94)	.37
MMSE[b] (points), mean (SD)	28.22 (1.96)	28.23 (1.97)	28.19 (1.93)	.69
RAVLT_Imm[c] (points), mean (SD)	39.48 (11.99)	39.46 (12.04)	39.54 (11.85)	.86
RAVLT_Learn[d], (points), mean (SD)	4.89 (2.66)	4.87 (2.67)	4.97 (2.63)	.34
RAVLT_Forget[e] (points), mean (SD)	4.17 (2.69)	4.17 (2.67)	4.18 (2.75)	.96
RAVLT_VF%[f] (points), mean (SD)	48.97 (35.42)	49.09 (35.78)	48.60 (34.33)	.72
TRABSCOR[g] (points), mean (SD)	95.92 (56.17)	95.31 (55.33)	97.72 (58.60)	.27
FAQ[h] (points), mean (SD)	1.90 (3.50)	1.87 (3.45)	1.98 (3.65)	.41
Ventricles (cm^3), mean (SD)	36.39 (19.47)	36.48 (19.38)	36.10 (19.75)	.62
Hippocampus (cm^3), mean (SD)	6.99 (1.10)	6.98 (1.10)	7.02 (1.09)	.36
Entorhinal (cm^3), mean (SD)	3.62 (0.73)	3.60 (0.73)	3.65 (0.75)	.11
Fusiform (cm^3), mean (SD)	17.60 (2.59)	17.58 (2.56)	17.69 (2.69)	.28
MidTemp (cm^3), mean (SD)	19.70 (2.79)	19.71 (2.76)	19.68 (2.88)	.77
ICV[i] (cm^3), mean (SD)	1524.19 (158.67)	1523.49 (157.98)	1526.28 (160.80)	.65
BVGN_BAG[j]	3.74 (5.22)	3.79 (5.39)	3.58 (4.68)	.30
SFCN_BAG[k]	3.71 (5.10)	3.73 (5.24)	3.66 (4.67)	.70
ResNer_BAG[l]	4.24 (5.23)	4.30 (5.37)	4.08 (4.76)	.30
Gender, n (%)				.51
Male	1803 (51.3)	1343 (51.0)	460 (52.3)	
Female	1711 (48.7)	1292 (49.0)	419 (47.7)	
Marriage, n (%)				.71
Married	2628 (74.8)	1980 (75.1)	648 (73.7)	
Widowed	400 (11.4)	289 (11.0)	111 (12.6)	
Divorced	338 (9.6)	256 (9.7)	82 (9.3)	
Never married	128 (3.6)	96 (3.6)	32 (3.6)	
Unknown	20 (0.6)	14 (0.5)	6 (0.7)	
APOE4_ε4[m], n (%)				.28
0	2100 (59.8)	1594 (60.5)	506 (57.6)	
1	1181 (33.6)	867 (32.9)	314 (35.7)	
2	233 (6.6)	174 (6.6)	59 (6.7)	
Cognitive states, n (%)				.99
CN[n]	1500 (42.7)	1125 (42.7)	375 (42.7)	
MCI[o]	2014 (57.3)	1510 (57.3)	504 (57.3)	

[a]ADAS13: Alzheimer's Disease Assessment—Cognitive Subscale 13-item version.

[b]MMSE: Mini-Mental State Examination.

[c]RAVLT_Imm: Rey Auditory Verbal Learning Test_immediate.

[d]RAVLT_Learn: Rey Auditory Verbal Learning Test_learning.

[e]RAVLT_Forget: Rey Auditory Verbal Learning Test_forgetting.

XSL•FO
RenderX

[f]RAVLT_VF%: Rey Auditory Verbal Learning Test_percentage_forgetting.

[g]TRABSCOR: trail making test part B, time (to complete in Neuropsychological Battery assessment).

[h]FAQ: Functional Activities Questionnaire.

[i]ICV: intracranial volume.

[j]BVGN_BAG: brain age gap estimated by the brain vision graph neural network model.

[k]SFCN_BAG: brain age gap estimated by the simple fully convolutional network model.

[l]ResNet_BAG: brain age gap estimated by the residual neural network model.

[m]APOE4_ε4: number of apolipoprotein E ε4 alleles.

[n]CN: cognitively normal.

[o]MCI: mild cognitive impairment.

Performance and Generalizability of BVGN

We have simultaneously trained various strong baseline models alongside our proposed model using the same training set and explored their performance on the same test set. The results are shown in Table 3.

Table 3. Performance of benchmark models and the brain vision graph neural network (BVGN) model.

Models	Performance	
	MAE[a] (y)	Pearson correlation coefficient (*r*)
3D ResNet15[b]	3.03	0.76
SFCN[c]	2.41	0.87
Isotropic BVGN	3.08	0.79
Pyramid BVGN	2.39	0.86
Pyramid BVGN (external validation)	2.49	0.92

[a]MAE: mean absolute error.

[b]ResNet15: residual neural network with 15 layers.

[c]SFCN: simple fully convolutional network.

Our isotropic BVGN attained an MAE of 3.08 years, which was slightly worse than the currently optimal model, SFCN, specifically designed for the brain age gap estimation (with an MAE of 2.41 years). After applying deformable convolutions, our pyramid BVGN achieved the best performance among all the tested models with an MAE of 2.39 years and a Pearson correlation coefficient of 0.86 between the predicted and chronological age.

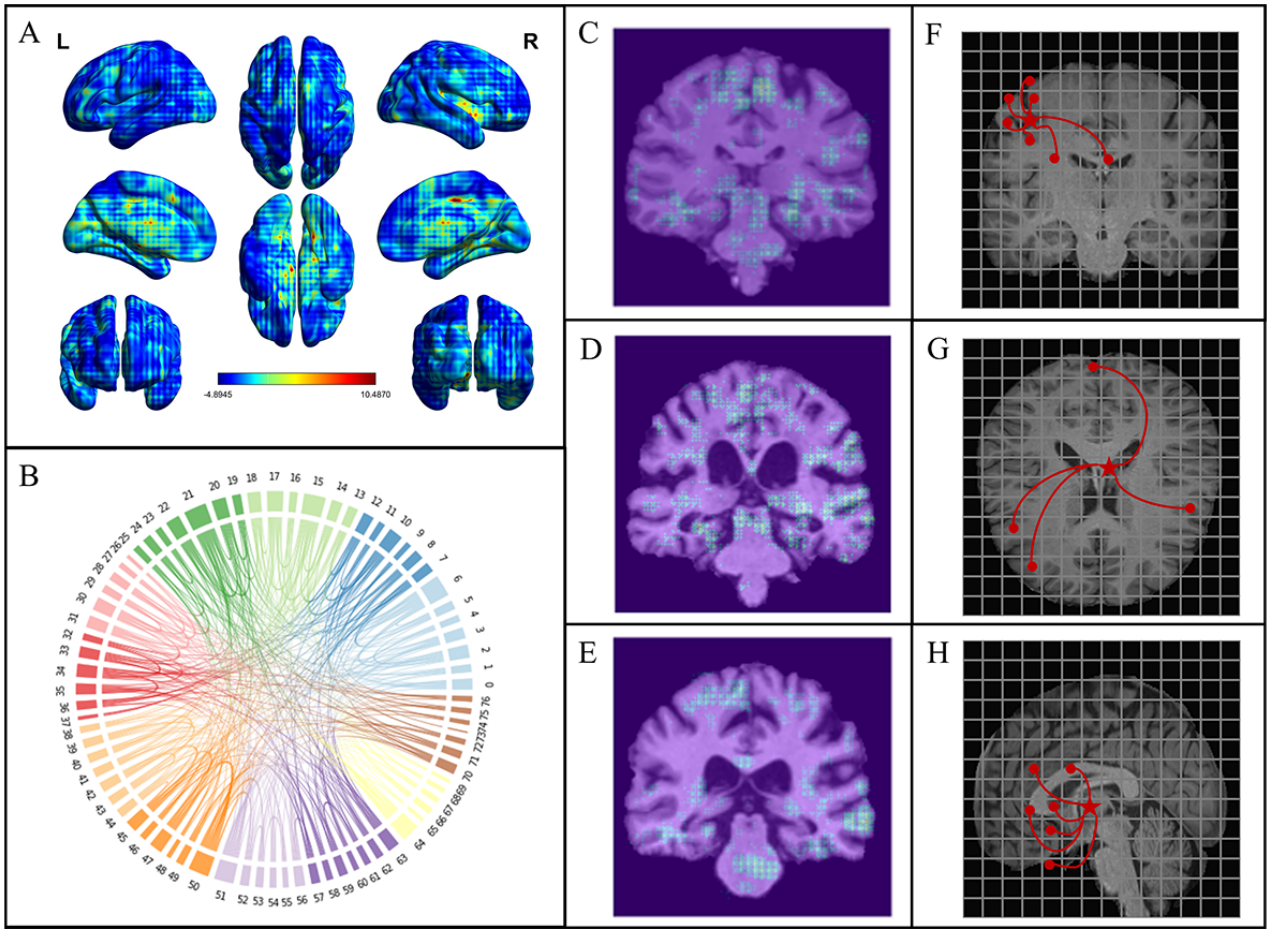
The BVGN model, when retrained on the UKB dataset, exhibited robust performance with an MAE of 2.49 years and a Pearson correlation coefficient of 0.92 (Figure S1 in Multimedia Appendix 1), underscoring its strong generalization capacity.

Interpretability Analysis of BVGN

As shown in Figures 4A and 4B, the saliency map indicated that our model exhibited a moderate focus on various dispersed regions within the cortical layer, while notably emphasizing relatively concentrated parenchymal regions. The regions of concern of the model are different in the population identified as CN (Figure 4C) and those with MCI (Figure 4D) and AD (Figure 4E).

XSL•FO
RenderX

Figure 4. Results for interpretability analysis.



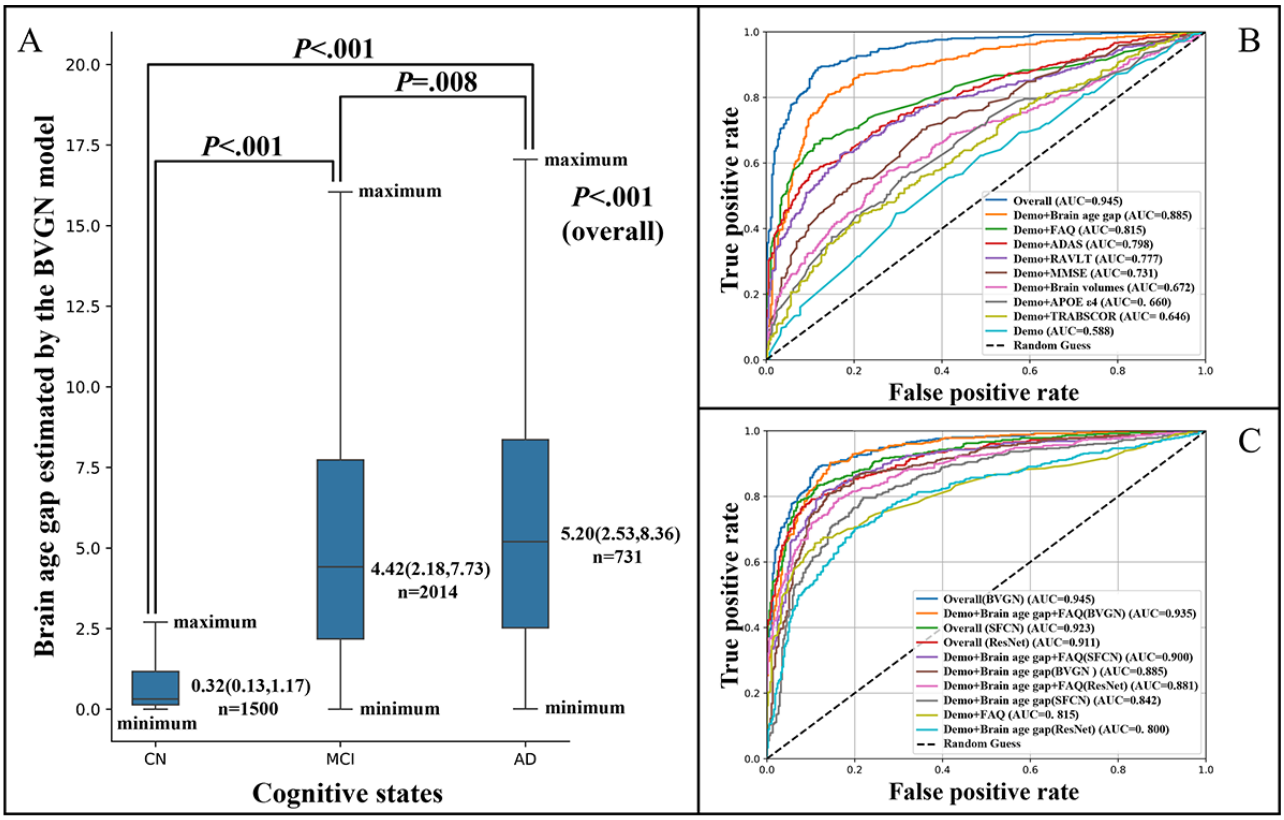
In addition, we described the brain as a complex system by extracting edge and node features from the last layer of an isotropic BVGN backbone. In Figure 4B, the top nodes with the highest degree centrality in the final layer of the isotropic BVGN are depicted, showcasing both these central nodes and their immediate neighboring nodes. This visualization provides insights into the most influential regions within the brain's voxel network as determined by the BVGN's analysis. Specifically, key cortical areas involved the frontal lobe and temporo-parieto-occipital regions, and parenchymal regions included the corpus callosum, cingulate gyrus, and parahippocampal gyrus, covering the striatum, thalamus, and adjacent hippocampus. In Figures 4F-4H, we leveraged a distinctive visualization approach inherent to the vision GNN backbone to illustrate the spatiotemporal dynamics of interregional aging patterns. The red star marks a specific voxel block, and the most relevant regions are delineated through red line connections.

Comparison of Brain Age Gap Among Different Cognitive State Groups

The median of brain age gap was 0.32 (IQR 0.13-1.17) among 1500 CN samples, 4.42 (IQR 2.18-7.73) across 2014 samples with MCI, and 5.20 (IQR 2.53-8.36) in 731 samples with AD (Figure 5A). A statistically significant difference in brain age gap was observed between AD and MCI (P=.008), with MCI demonstrating a significantly higher gap than CN (P<.001), indicating the potential of our model to better distinguish between individuals classified as CN and those with MCI. The results adjusted for age and gender are provided in Table S3 in Multimedia Appendix 1, which were similar to the results mentioned earlier.

XSL•FO
RenderX

Figure 5. Results or group comparisons and receiver operating characteristic curve of multiple logistic regressions. ADAS: Alzheimer's Disease Assessment Scale; APOE4: apolipoprotein E 4 alleles; AUC: area under the receiver operating characteristic curve; BVGN: brain vision graph neural network; CN: cognitively normal; FAQ: Functional Activities Questionnaire; MCI: mild cognitive impairment; MMSE: Mini-Mental State Examination; RAVLT: Rey Auditory Verbal Learning Test; ResNet: residual neural network; SFCN: simple fully convolutional network; TRABSCOR: trail making test part B, time.



Multiple Logistic Regression Models

Brain Age Gap Compared to Other Cognition-Related Variables

Detailed metrics of various logistic regression analyses aiming to compare brain age gap and other cognition-related variables are provided in Table 4, while multiple receiver operating characteristic curves and corresponding AUC scores are illustrated in Figure 5B. Initially, the benchmark model, including only demographic features, yielded an AUC of 0.588. Subsequent incorporation of additional variables, such as FAQ [47], Alzheimer's Disease Assessment Scale-Cognitive Subscale13-item version [48], Rey Auditory Verbal Learning Test [49], Mini-Mental State Examination [50], trail making test part B, time [51], brain volume features, APOE4 carriage, and brain age gap led to improved AUC scores, reaching 0.815, 0.798, 0.777, 0.731, 0.646, 0.672, 0.660, and 0.885, respectively. Notably, the comprehensive model encompassing all variables achieved an AUC of 0.945, with the brain age gap exhibiting the highest coefficient (Figure 6B). These results underscore the robust and superior discriminative ability of the brain age gap in distinguishing between CN and MCI samples.

XSL•FO
RenderX

Table 4. Performance of the logistics regression models in classifying cognitive states: cognitively normal and mild cognitive impairment (brain age gap versus conventional cognitive variables).

Input features	Accuracy	Precision	Recall	F_1-score	AUC[a]
Overall	0.882	0.915	0.875	0.895	0.945
Demographic+BVGN_BAG[b]	0.820	0.890	0.784	0.833	0.885
Demographic+SFCN_BAG[c]	0.780	0.839	0.764	0.800	0.842
Demographic+ResNet_BAG[d]	0.743	0.823	0.702	0.758	0.800
Demographic+FAQ[e]	0.758	0.880	0.669	0.760	0.815
Demographic+ADAS[f]	0.721	0.774	0.726	0.749	0.798
Demographic+RAVLT[g]	0.714	0.759	0.736	0.747	0.777
Demographic+MMSE[h]	0.678	0.723	0.710	0.717	0.731
Demographic+brain volumes	0.617	0.646	0.734	0.687	0.672
Demographic+APOE4[i]	0.631	0.645	0.794	0.712	0.660
Demographic+TRABSCOR[j]	0.611	0.644	0.720	0.680	0.646
Demographic	0.577	0.583	0.919	0.713	0.588

[a]AUC: area under the receiver operating characteristic curve.

[b]BVGN_BAG: brain age gap estimated by the brain vision graph neural network model.

[c]SFCN_BAG: brain age gap estimated by the simple fully convolutional network model.

[d]ResNet_BAG: brain age gap estimated by the residual neural network model.

[e]FAQ: Functional Activities Questionnaire.

[f]ADAS: Alzheimer's Disease Assessment Scale.

[g]RAVLT: Rey Auditory Verbal Learning Test.

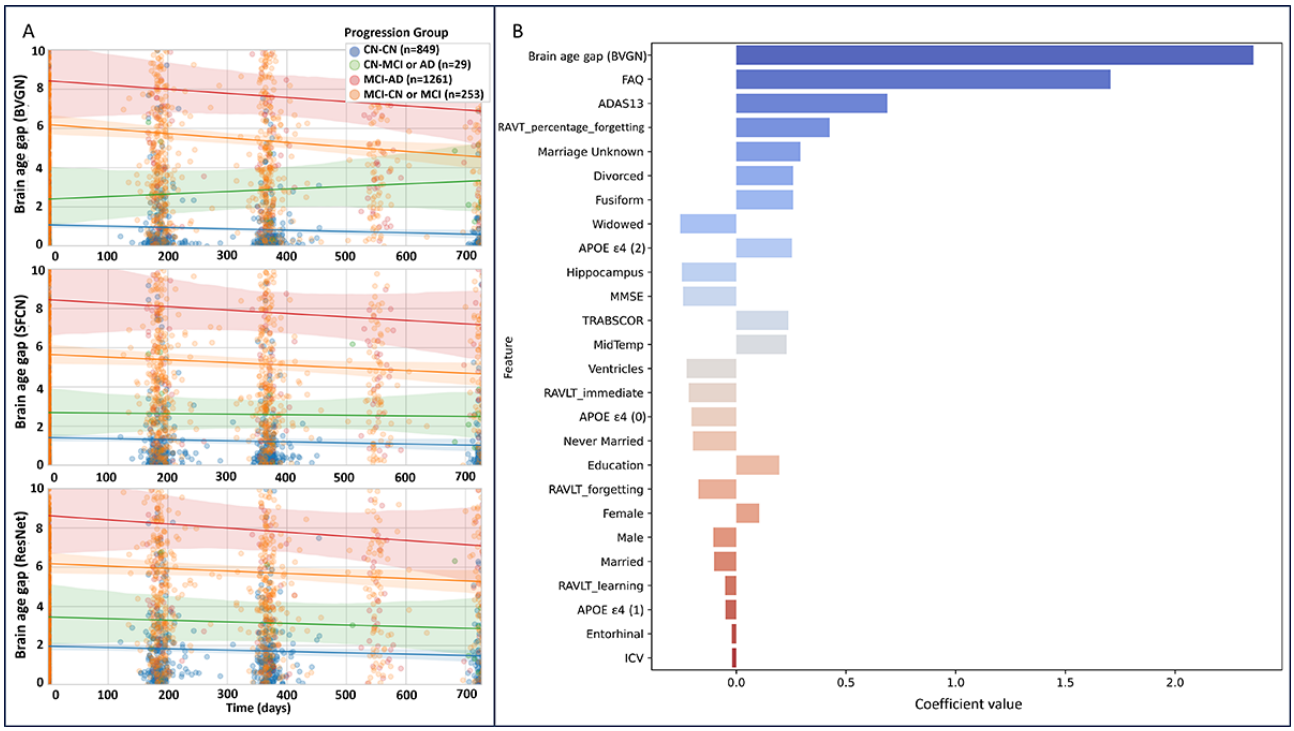
[h]MMSE: Mini-Mental State Examination.

[i]APOE4: apolipoprotein E ε4 alleles.

[j]TRABSCOR: Trail Making Test Part B Time.

XSL•FO
RenderX

Figure 6. (A) Visualization of longitudinal analysis for brain age gap from various models by progression groups over 2 years and (B) coefficients of various variables in the brain vision graph neural network (BVGN) comprehensive logistics regression model discriminating between cognitively normal (CN) and mild cognitive impairment (MCI). ADAS: Alzheimer's Disease Assessment Scale; APOE ε4: apolipoprotein E 4 alleles; AUC: area under the receiver operating characteristic curve; FAQ: Functional Activities Questionnaire; ICV: intracranial volume; MMSE: Mini-Mental State Examination; RAVLT: Rey Auditory Verbal Learning Test; ResNet: residual neural network; SFCN: simple fully convolutional network; TRABSCOR: trail making test part B, time.



Combined Performance of Brain Age Gap With FAQ

In the comprehensive logistics regression model of BVGN, FAQ demonstrated the second-largest coefficient following brain age gap (Figure 6B). When comparing the comprehensive models that included all selective variables with another model incorporating demographic variables, brain age gap and FAQ, minimal declines in AUCs were observed—BVGN decreased by 0.011 (from 0.945 to 0.935), SFCN by 0.024 (from 0.923 to 0.900), and ResNet by 0.031 (from 0.911 to 0.881; Table 5; Figure 5C). Notably, BVGN exhibited the smallest AUC reduction when all variables except FAQ were missing, outperforming other frameworks. In addition, with only demographic variables, brain age gap, and FAQ, BVGN could provide a higher AUC (0.935) compared to other frameworks, even when all selective features were incorporated (SFCN: AUC=0.923; ResNet: AUC=0.911). This underscored the combined efficiency of FAQ and brain age gap, as well as the robustness of BVGN in scenarios with missing variables.

XSL•FO
RenderX

Table 5. Performance of logistics regression models: brain age gaps evaluated by models with different precisions (mean absolute errors) and combined utility of brain age gap with Functional Activity Questionnaire (FAQ).

Input features and model	Accuracy	Precision	Recall	F_1-score	AUC[a]
Overall					
BVGN[b]	0.882	0.915	0.875	0.895	0.945
SFCN[c]	0.851	0.889	0.845	0.867	0.923
ResNet[d]	0.830	0.877	0.819	0.847	0.911
Demographic+brain age gap+FAQ					
BVGN	0.856	0.907	0.833	0.869	0.935
SFCN	0.835	0.895	0.808	0.849	0.900
ResNet	0.807	0.873	0.776	0.821	0.881
Demographic+brain age gap					
BVGN	0.820	0.890	0.784	0.833	0.885
SFCN	0.780	0.839	0.764	0.800	0.842
ResNet	0.743	0.823	0.702	0.758	0.800

[a]AUC: area under the receiver operating characteristic curve.

[b]BVGN: brain vision graph neural network.

[c]SFCN: simple fully convolutional network.

[d]ResNet: residual neural network.

Potential Clinical Value of Lower MAEs in Brain Age Gap Estimation

In assessing the potential clinical utility of models with lower MAEs, our BVGN model consistently outperformed SFCN and ResNet models (Table 5), both in the simplified model comprising only demographic features and brain age gap and in the comprehensive model incorporating all variables. Notably, brain age gap estimated by BVGN demonstrated the highest AUCs (Figure 5C). Specifically, in the simplified model, BVGN achieved AUCs of 0.885, while SFCN and ResNet yielded AUCs of 0.842 and 0.800, respectively. Similarly, in the comprehensive model, BVGN achieved AUCs of 0.945, surpassing SFCN and ResNet with AUCs of 0.923 and 0.911, respectively. These findings were consistent with the ranking of MAEs obtained from these 3 frameworks, indicating that models with lower MAEs might have the potential to enhance the model's discriminatory ability.

Longitudinal Value of BVGN

Brain Age Gap Trajectories in 2 Years

The number of MRI scans for CN-CN, CN-MCI or CN-AD, MCI-CN or MCI-MCI, and MCI-AD were 849, 29, 1261, and 253, respectively.

As illustrated in Figure 6A, we observed consistent ordering of brain age gaps across progression groups in all the frameworks over the 2-year period (MCI-AD>MCI-CN or MCI-MCI>CN-MCI or CN-AD>CN-CN). Moreover, brain age gaps tended to decrease over the 2-year follow-up period in each progression group in both the SFCN and ResNet models, as well as in all groups in the BVGN model except the CN-MCI or CN-AD group. However, in the BVGN model, an interesting increase in the brain age gap was observed in the CN-MCI or CN-AD group.

Risk Evaluation by Brain Age Gap

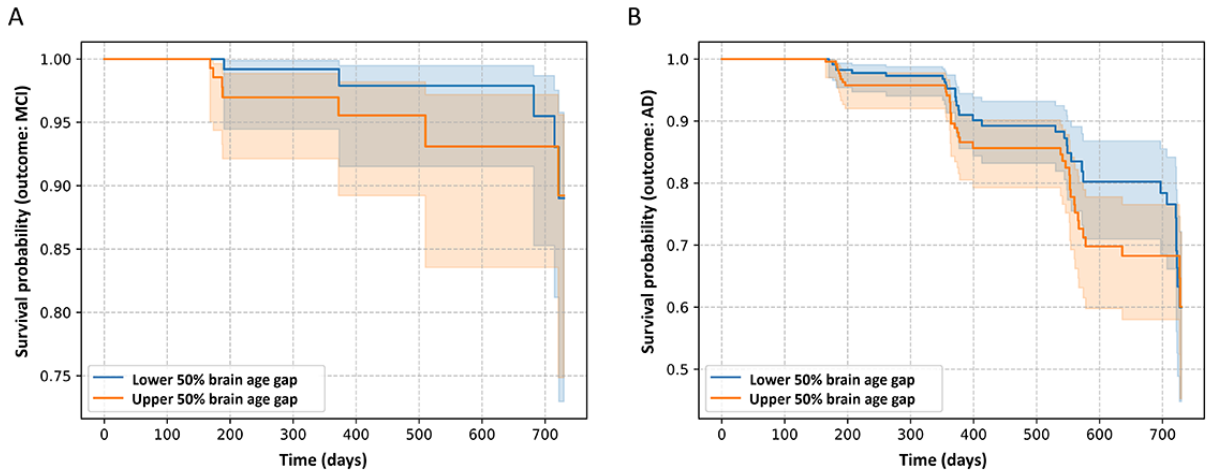
Among the 281 individuals classified as CN at the baseline, 12 (4.3%) declined to MCI, while 269 (95.7%) remained in the CN state. Among the 468 participants with MCI at the baseline, 74 (15.8%) declined to AD while 394 (84.2%) remained in the MCI state or improved to the CN state.

Both univariable Cox regression models and demographic-adjusted as well as fully adjusted multivariable models demonstrated that the brain age gap was a significant risk factor for cognitive decline in both the CN and MCI groups (Table S4 in Multimedia Appendix 1). In the univariable model, the HR for brain age gap was 1.55 (95% CI 1.13-2.13) in the CN group and 1.29 (95% CI 1.09-1.51) in the MCI group. For the CN group, the HR for the brain age gap was 1.47 (95% CI 1.06-2.05) in the demographic-adjusted model and increased to 2.13 (95% CI 1.01-4.47) in the fully adjusted model. In the MCI group, the HR was 1.30 (95% CI 1.10-1.53) for the demographic-adjusted model and 1.24 (95% CI 1.03-1.48) for the fully adjusted model.

The Kaplan-Meier curves indicated that in both CN (Figure 7A) and MCI (Figure 7B) groups, the survival curve for the lower 50% of the brain age gap group was higher than that for the upper 50% group. This finding aligned with the HRs, suggesting that a higher level of brain age gap was associated with an increased risk of cognitive decline.

XSL•FO
RenderX

Figure 7. Visualization of Kaplan-Meier curves stratified by the upper and lower 50% levels of brain age gap in both (A) cognitively normal and (B) mild cognitive impairment (MCI) groups. AD; Alzheimer disease.



Discussion

Principal Findings

In summary, this study proposed a novel framework, BVGN, for brain age estimation. This model not only introduced an innovative modeling approach but also achieved lower MAEs (2.39 years) compared to current state-of-the-art frameworks for brain age gap estimation. It also demonstrated a robust performance on a large-scale external dataset, maintaining strong generalizability (MAE=2.49 years; *r*=0.92; Figure S1 in Multimedia Appendix 1). In addition, BVGN offers an interpretable analysis approach not found in other DL models, which reveals patterns of coaging between brain regions by visualizing the locations of neighboring nodes associated with high-importance nodes. In our study, BVGN provided interpretable attention regions corroborated by medical evidence. Our validation experiments demonstrated its robust performance in MCI diagnosis, its capacity as a risk biomarker for cognitive decline, and its potential to aid in screening and clinical applications.

BVGN integrates deformable convolutional kernels and a dynamic graph for brain age gap estimation. To the best of our knowledge, this work represents the first successful application of vision GNN to T1-weighted MRI data without requiring additional brain segmentation, achieving competitive performance. Unlike 3D CNNs, which process brain images as regular voxel grids and struggle with the inherently irregular morphological characteristics of cerebral regions, often resulting in redundant computations and information loss, BVGN leverages graph-structured representations to preserve complex anatomical features. The integration of deformable convolutional kernels enables the model to focus on specific pathological patterns while suppressing extraneous noise, thereby enhancing local feature modeling [34]. Furthermore, the topological properties of graph structures facilitate multiregional feature aggregation across distant brain regions [38,39], a capability unattainable by grid-based CNNs. To strengthen age prediction accuracy, we transformed age from a discrete numerical variable into a smoothed probability distribution, preserving its continuity and capturing interval similarity [40,41,49,50]. Comparative evaluations demonstrate superior performance of BVGN over existing CNN-based models such as BrainAgeNeXt, achieving an MAE of 2.39 versus 2.78 years, and maintaining robust generalization capabilities (MAE=2.49 years; *r*=0.92) on the UKB dataset after retraining. Structure-functional coupling has garnered increasing attention in neuroscience, with graph-based modeling serving as a critical analytical framework [52]. For instance, structural connectivity derived from diffusion-weighted imaging via white matter fiber tractography and functional connectivity computed from functional MRI are commonly represented as adjacency matrices in graph structures. The framework, being graph-based in design, is inherently more scalable and adaptable to the aforementioned MRI modalities compared to other DL approaches, providing an interface with modal-type embedding for early feature fusion [40]. This capability not only enhances multimodal integration but also helps mitigate the semantic gap across different MRI modalities [53].

We found substantial evidence validating the BVGN's focused brain regions associated with cognitive decline, which provided strong interpretability for our model. The frontal lobes exhibited particularly prominent changes within an annual 0.87% reduction in overall cortical volume [54], while the temporo-parieto-occipital junction was intricately involved in high-level human neurological functions, such as language, memory, calculation, and writing [55]. Shape characteristics in the corpus callosum were demonstrated to have potential for distinguishing cognitive deterioration [56,57], and the cingulate gyrus was implicated in the onset of neurodegenerative and psychiatric disorders [58]. In addition, the parahippocampal gyrus integrates signals from the limbic and neocortex, predicting later memory [59], and the fractional anisotropy alterations in the parahippocampal white matter have been observed across AD stages [60,61].

In recent years, despite the substantial academic focus on the brain age gap, questions have persisted regarding its reliability as a clinical tool compared to cognitive assessments for diagnosing cognitive decline. After multiple experiments that incorporated various kinds of input variables to distinguish CN and MCI, our findings revealed that except for the trail making

XSL•FO
RenderX

test part B, time (AUC=0.646), cognitive assessments exhibited AUCs from 0.731 to 0.815, which surpassed both the AOPE4 carriage (AUC=0.660) and brain volume characteristics (AUC=0.672), and were inferior to brain age gap (AUC=0.885). While cognitive assessments provided affordability, the subjective bias among clinical practitioners might impact diagnosis [62], with limitations in sensitivity observed in certain cases, such as mild levels of impairment [32], subjective deficits, and healthy cognitive aging [63]. According to our findings, our model could offer alternative and more objective evaluations for clinical settings, potentially reducing physicians' workload and enhancing convenience.

The FAQ is a 10-item collateral-report scale with 4-point ordinal responses per item, totaling 0 to 30 points, with higher scores indicating greater functional impairment [64]. This makes it highly accessible in clinical settings and convenient for clinical practitioners. Scientific publications have demonstrated its good internal consistency and high discriminative validity for differentiating between CN and MCI, as well as between MCI and very mild AD [65,66]. Its high predictive validity in detecting individuals at risk of progression from MCI to AD and from CN to MCI has also been reported [47]. In our study, we observed minimal reductions in AUCs when comparing models that incorporated all selective variables and models with only demographic variables, brain age gap, and FAQ. Our selected demographic variables, education years, gender, and marriage status, were deliberately straightforward. This simplicity would enhance the feasibility of our proposed model for clinical applications, especially in cases of uncertain MCI diagnosis. By integrating basic demographics, FAQ responses, and T1-weighted MRI scans, our BVGN model could serve as a reliable assistant tool for clinicians in MCI diagnosis.

In our comprehensive model that integrated brain age gap, cognitive assessments, and other relevant variables, we observed the highest AUC of 0.945 in classification tasks. The integration of variables from diverse perspectives might contribute to improving performance. In addition, we observed brain age gap possessed the highest variable coefficient, with each scalarized unit increase associated with a 10.6 times higher risk of MCI. This further supported the brain age gap's capacity in MCI diagnosis.

Furthermore, our analysis revealed a relationship between model performance, as indicated by lower MAEs, and superior performance in discriminating between CN and MCI. Specifically, BVGN, SFCN, and ResNet achieved MAEs of 2.39, 2.41, and 2.65 years, respectively. Their corresponding AUCs in the simplified and comprehensive logistics models for classifying CN and MCI were 0.885, 0.842, and 0.800 and 0.945, 0.923, and 0.911. This finding underscored the importance of striving for minimal MAE in model development for enhanced predictive accuracy in subsequent cognitive decline tasks.

During longitudinal analysis, we observed an increase in the brain age gap derived from BVGN in CN-MCI or CN-AD progression group within 2 years after baseline. This increase was not observed in either SFCN or ResNet models. This group of participants classified as "CN" might be some individuals at high risk who have already experienced early pathological changes related to dementia yet remain classified as "CN" due to the absence of noticeable cognitive impairments. This finding supported BVGN as a sensitive marker for identifying participants at high risk who, despite being classified as "CN," were likely to progress to MCI in a few years. Existing scientific research suggests that MCI can be reversible to a certain extent, whereas AD, which is associated with significant cognitive degeneration, is challenging to reverse once it has progressed [67-69]. Therefore, early identification of these individuals at high risk for MCI is of great significance for better patient prognosis.

Multiple Cox proportional hazard models and Kaplan-Meier curves revealed a clear association between a higher brain age gap and an increased risk of cognitive decline. Specifically, each 1-unit increase in the brain age gap was associated with a 55% higher risk of cognitive decline in the CN group and a 29% increased risk in the MCI group. This finding underscored the brain age gap derived from BVGN as a sensitive risk biomarker for cognitive decline. Notably, the brain age gap appeared more sensitive in the CN groups. This observation aligned with existing research, which indicated the ability of the brain age gap to detect subtle preclinical or early neurodegenerative changes in populations classified as CN [70,71].

In clinical settings, the BVGN-estimated brain age gap serves as a tool for detecting cognitive decline and assessing the efficacy of innovative therapeutic interventions. For example, in the United States, 2 antiamyloid monoclonal antibodies, lecanemab (brand name Leqembi) and aducanumab (brand name Aduhelm) [72,73], have been approved for the treatment of AD. Despite the potential therapeutic benefits, these interventions are not devoid of inherent risks. Patients require vigilant monitoring for potential side effects, such as amyloid-related imaging abnormalities and infusion reactions, particularly at the onset of treatment. Currently, imaging provides limited insight into a patient's therapeutic response. However, BVGN could theoretically offer a novel approach by capturing changes in brain age pre- and posttreatment with monoclonal antibodies. This could provide an objective measure of a patient's response to the medication, enhancing our ability to assess the efficacy of these treatments.

Limitations and Future Directions

The limitation of BVGN lies in its inability to evaluate the performance and interpretability analysis results on larger datasets. In addition, the use of 3D deformable convolutional kernels in BVGN leads to increased computational costs during the training process when compared to other DL models. Our proposed DL-based brain aging framework is specifically designed for the early screening and prediction of neurodegenerative diseases. Epidemiological and neuroimaging studies indicate that structural alterations linked to neurodegenerative processes begin to manifest or accelerate after the age of 50 years [74]. To better capture brain age deviations closely associated with these pathological conditions, we selected individuals aged >45 years as our study population. This age threshold ensures the inclusion of individuals at the critical transition phase where neurodegeneration-related biomarkers become detectable. However, extending this

XSL•FO
RenderX

framework to the entire life span remains essential to comprehensively understand age-related neurodegenerative processes and identify potential early intervention strategies across all life stages.

Despite its limitations, BVGN remains a valuable tool in the study of brain aging, as it surpasses cognitive assessments, brain volume features, and susceptibility genes in classifying participants as CN and those with MCI, and has been demonstrated to be a risk factor for cognitive decline. This underscores the effectiveness of the brain age gap, estimated through individual MRI using BVGN, as an impactful neuroimaging biomarker. Furthermore, BVGN introduces graph theory analysis as a novel component of interpretability for the first time in brain age gap estimation. In the future, we will leverage the advantages of GNNs in brain analysis, and integrate them with other modalities of medical imaging to further enhance our framework and provide a more robust interpretability analysis based on graph theory. Future work will focus on 3 key directions to enhance the clinical utility and scientific depth of our framework. First, we will explore diverse graph convolutional feature aggregation mechanisms to optimize computational efficiency, enabling broader deployment in real-world clinical environments. Second, we aim to integrate multimodal MRI data by converting different imaging sequences (eg, T1-weighted, T2-weighted, and fluid attenuated inversion recovery) into graph-based representations within the BVGN framework. This multimodal integration will not only enhance the accuracy of brain aging quantification but also enable systematic investigation of cross-modal interactions underlying neurodegenerative processes. Finally, we plan to extend our framework across the entire life span to uncover age-specific aging patterns and their associations with disease progression.

Conclusions

This study proposed a novel DL framework, BVGN, for accurate brain age estimation. Our BVGN model used deformable convolutional kernel to capture brain complex morphology and graph theory to model brain topology, which was applicable to the biological nature of brain MRI. BVGN achieved a superior MAE of 2.39 years in the same testing set compared to other methods, with its attention region aligning with medical evidence of cognitive decline. The robust capacity of BVGN-derived brain age gap was validated through both cross-sectional and longitudinal analysis, which outperformed all conventional tools in aiding the diagnosis of MCI and was sensitive for identifying individuals at high risk for future MCI. The BVGN model can precisely evaluate brain aging and predict cognitive decline, offering substantial potential for improving early identification of neurodegenerative disorders and enhancing clinical applications.

Acknowledgments

This work was supported by the National Natural Science Foundation for Young Scientists of China (grant 82203984), the Healthy Zhejiang One Million People Cohort (grant K-20230085), the Zhejiang Provincial Key Laboratory of Intelligent Preventive Medicine (grant 2020E10004), the Leading Innovative and Entrepreneur Team Introduction Program of Zhejiang (grant 2019R01007), and the Key Research and Development Program of Zhejiang Province (grant 2020C03002).

The Alzheimer's Disease Neuroimaging Initiative (ADNI) is funded by the National Institute on Aging and the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc; Cogstate; Eisai Inc; Elan Pharmaceuticals, Inc; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc; Fujirebio; GE Healthcare; IXICO Ltd; Janssen Alzheimer Immunotherapy Research & Development, LLC; Johnson & Johnson Pharmaceutical Research & Development LLC; Lumosity; Lundbeck; Merck & Co, Inc; Meso Scale Diagnostics, LLC; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health [75]. The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

The data for external validation was obtained from the UK Biobank database [76].

Authors' Contributions

WL, XW, and ZL contributed to conceptualization, data curation, methodology, funding acquisition, project administration, resources, supervision, Writing, review, and editing. ZL, JL, and JL contributed to data curation, formal analysis, investigation, methodology, validation, visualization, software support, writing the original draft, writing, review, and editing. MW, AX, YH, QY, LZ, and YL contributed to investigation, validation, visualization, writing, review, and editing. MW contributed to data curation and methodology. JB contributed to writing, review, editing, methodology, validation, and resources.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Supplementary tables, figures, and detailed brain vision graph neural network (BVGN) architecture: exhaustive demographics for BVGN development and generalizability validation, Dunn-corrected intergroup comparisons, Cox regression results for brain-age-gap risk, and scatter plot performance across external cohorts.
[DOCX File , 59 KB - aging_v8i1e73004_app1.docx]

References

1. Kanasi E, Ayilavarapu S, Jones J. The aging population: demographics and the biology of aging. Periodontol 2000 2016 Oct;72(1):13-18. [doi: 10.1111/prd.12126] [Medline: 27501488]
2. Kehler DS. Age-related disease burden as a measure of population ageing. Lancet Public Health 2019 Mar;4(3):e123-e124 [FREE Full text] [doi: 10.1016/S2468-2667(19)30026-X] [Medline: 30851865]
3. Ferri CP, Prince M, Brayne C, Brodaty H, Fratiglioni L, Ganguli M, et al. Global prevalence of dementia: a Delphi consensus study. Lancet 2005 Dec 17;366(9503):2112-2117 [FREE Full text] [doi: 10.1016/S0140-6736(05)67889-0] [Medline: 16360788]
4. Grimm A, Eckert A. Brain aging and neurodegeneration: from a mitochondrial point of view. J Neurochem 2017 Nov;143(4):418-431 [FREE Full text] [doi: 10.1111/jnc.14037] [Medline: 28397282]
5. Azam S, Haque ME, Balakrishnan R, Kim IS, Choi DK. The ageing brain: molecular and cellular basis of neurodegeneration. Front Cell Dev Biol 2021;9:683459 [FREE Full text] [doi: 10.3389/fcell.2021.683459] [Medline: 34485280]
6. Ran C, Yang Y, Ye C, Lv H, Ma T. Brain age vector: a measure of brain aging with enhanced neurodegenerative disorder specificity. Hum Brain Mapp 2022 Nov;43(16):5017-5031 [FREE Full text] [doi: 10.1002/hbm.26066] [Medline: 36094058]
7. Muñoz-Ramírez V, Kmetzsch V, Forbes F, Meoni S, Moro E, Dojat M. Subtle anomaly detection: application to brain MRI analysis of de novo Parkinsonian patients. Artif Intell Med 2022 Mar;125:102251 [FREE Full text] [doi: 10.1016/j.artmed.2022.102251] [Medline: 35241258]
8. Lerch JP, van der Kouwe AJ, Raznahan A, Paus T, Johansen-Berg H, Miller KL, et al. Studying neuroanatomy using MRI. Nat Neurosci 2017 Feb 23;20(3):314-326 [FREE Full text] [doi: 10.1038/nn.4501] [Medline: 28230838]
9. Cole JH, Marioni RE, Harris SE, Deary IJ. Brain age and other bodily 'ages': implications for neuropsychiatry. Mol Psychiatry 2019 Feb;24(2):266-281 [FREE Full text] [doi: 10.1038/s41380-018-0098-1] [Medline: 29892055]
10. Cole JH, Franke K. Predicting age using neuroimaging: innovative brain ageing biomarkers. Trends Neurosci 2017 Dec;40(12):681-690 [FREE Full text] [doi: 10.1016/j.tins.2017.10.001] [Medline: 29074032]
11. Rokicki J, Wolfers T, Nordhøy W, Tesli N, Quintana DS, Alnaes D, et al. Multimodal imaging improves brain age prediction and reveals distinct abnormalities in patients with psychiatric and neurological disorders. Hum Brain Mapp 2021 Apr 15;42(6):1714-1726 [FREE Full text] [doi: 10.1002/hbm.25323] [Medline: 33340180]
12. Mishra S, Beheshti I, Khanna P. A review of neuroimaging-driven brain age estimation for identification of brain disorders and health conditions. IEEE Rev Biomed Eng 2023;16:371-385. [doi: 10.1109/RBME.2021.3107372] [Medline: 34428153]
13. Cearns M, Hahn T, Baune BT. Recommendations and future directions for supervised machine learning in psychiatry. Transl Psychiatry 2019 Oct 22;9(1):271 [FREE Full text] [doi: 10.1038/s41398-019-0607-2] [Medline: 31641106]
14. Baecker L, Garcia-Dias R, Vieira S, Scarpazza C, Mechelli A. Machine learning for brain age prediction: introduction to methods and clinical applications. EBioMedicine 2021 Oct;72:103600 [FREE Full text] [doi: 10.1016/j.ebiom.2021.103600] [Medline: 34614461]
15. Lee WH, Antoniades M, Schnack HG, Kahn RS, Frangou S. Brain age prediction in schizophrenia: does the choice of machine learning algorithm matter? Psychiatry Res Neuroimaging 2021 Apr 30;310:111270 [FREE Full text] [doi: 10.1016/j.pscychresns.2021.111270] [Medline: 33714090]
16. Baecker L, Dafflon J, da Costa PF, Garcia-Dias R, Vieira S, Scarpazza C, et al. Brain age prediction: a comparison between machine learning models using region- and voxel-based morphometric data. Hum Brain Mapp 2021 Jun 01;42(8):2332-2346 [FREE Full text] [doi: 10.1002/hbm.25368] [Medline: 33738883]
17. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015 May 28;521(7553):436-444. [doi: 10.1038/nature14539] [Medline: 26017442]
18. Kolbeinsson A, Filippi S, Panagakis Y, Matthews PM, Elliott P, Dehghan A, et al. Accelerated MRI-predicted brain ageing and its associations with cardiometabolic and brain disorders. Sci Rep 2020 Nov 17;10(1):19940 [FREE Full text] [doi: 10.1038/s41598-020-76518-z] [Medline: 33203906]
19. Wood DA, Kafiabadi S, Busaidi AA, Guilhem E, Montvila A, Lynch J, et al. Accurate brain-age models for routine clinical MRI examinations. Neuroimage 2022 Apr 01;249:118871 [FREE Full text] [doi: 10.1016/j.neuroimage.2022.118871] [Medline: 34995797]
20. He S, Feng Y, Grant PE, Ou Y. Deep relation learning for regression and its application to brain age estimation. IEEE Trans Med Imaging 2022 Sep;41(9):2304-2317 [FREE Full text] [doi: 10.1109/TMI.2022.3161739] [Medline: 35320092]
21. F A, La Rosa RF, Sather A, Rechtman E, Jalees M, Nabeel I, et al. MRI signature of brain age underlying post- traumatic stress disorder in World Trade Center responders. medRxiv. Preprint posted online on October 19, 2024 [FREE Full text] [doi: 10.1101/2024.10.18.24315761]

XSL•FO
RenderX

22. Zhang Z, Cui P, Zhu W. Deep learning on graphs: a survey. IEEE Trans Knowl Data Eng 2022 Jan 1;34(1):249-270. [doi: 10.1109/TKDE.2020.2981333]
23. Peng H, Xie P, Liu L, Kuang X, Wang Y, Qu L, et al. Morphological diversity of single neurons in molecularly defined cell types. Nature 2021 Oct 06;598(7879):174-181 [FREE Full text] [doi: 10.1038/s41586-021-03941-1] [Medline: 34616072]
24. Tadić B, Andjelković M, Melnik R. Functional geometry of human connectomes. Sci Rep 2019 Aug 19;9(1):12060 [FREE Full text] [doi: 10.1038/s41598-019-48568-5] [Medline: 31427676]
25. Despotović I, Goossens B, Philips W. MRI segmentation of the human brain: challenges, methods, and applications. Comput Math Methods Med 2015;2015:450341 [FREE Full text] [doi: 10.1155/2015/450341] [Medline: 25945121]
26. Han K, Wang Y, Guo J, Tang Y, Wu E. Vision GNN: an image is worth graph of nodes. In: Proceedings of the 36th International Conference on Neural Information Processing Systems. 2022 Presented at: NIPS'22; November 28-December 9, 2022; New Orleans, LA. [doi: 10.48550/arXiv.2206.00272]
27. Weiner MW, Veitch DP. Introduction to special issue: overview of Alzheimer's disease neuroimaging initiative. Alzheimers Dement 2015 Jul;11(7):730-733 [FREE Full text] [doi: 10.1016/j.jalz.2015.05.007] [Medline: 26194308]
28. Toga AW, Crawford KL. The Alzheimer's disease neuroimaging initiative informatics core: a decade in review. Alzheimers Dement 2015 Jul;11(7):832-839 [FREE Full text] [doi: 10.1016/j.jalz.2015.04.004] [Medline: 26194316]
29. Weiner MW, Kanoria S, Miller MJ, Aisen PS, Beckett LA, Conti C, et al. Overview of Alzheimer's disease neuroimaging initiative and future clinical trials. Alzheimers Dement 2025 Jan;21(1):e14321 [FREE Full text] [doi: 10.1002/alz.14321] [Medline: 39711072]
30. Esteban J, Starr A, Willetts R, Hannah P, Bryanston-Cross P. A Review of data fusion models and architectures: towards engineering guidelines. Neural Comput Appl 2005 Jun 21;14(4):273-281. [doi: 10.1007/s00521-004-0463-7]
31. Jack CRJ, Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D, et al. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. J Magn Reson Imaging 2008 Apr;27(4):685-691 [FREE Full text] [doi: 10.1002/jmri.21049] [Medline: 18302232]
32. Cullen B, O'Neill B, Evans JJ, Coen RF, Lawlor BA. A review of screening tests for cognitive impairment. J Neurol Neurosurg Psychiatry 2007 Aug;78(8):790-799 [FREE Full text] [doi: 10.1136/jnnp.2006.095414] [Medline: 17178826]
33. Miller KL, Alfaro-Almagro F, Bangerter NK, Thomas DL, Yacoub E, Xu J, et al. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. Nat Neurosci 2016 Nov;19(11):1523-1536 [FREE Full text] [doi: 10.1038/nn.4393] [Medline: 27643430]
34. Mansour L S, Di Biase MA, Smith RE, Zalesky A, Seguin C. Connectomes for 40,000 UK Biobank participants: a multi-modal, multi-scale brain network resource. Neuroimage 2023 Dec 01;283:120407 [FREE Full text] [doi: 10.1016/j.neuroimage.2023.120407] [Medline: 37839728]
35. Fonov V, Evans AC, Botteron K, Almli CR, McKinstry RC, Collins DL. Unbiased average age-appropriate atlases for pediatric studies. Neuroimage 2011 Jan 01;54(1):313-327 [FREE Full text] [doi: 10.1016/j.neuroimage.2010.07.033] [Medline: 20656036]
36. Yaniv Z, Lowekamp BC, Johnson HJ, Beare R. SimpleITK image-analysis notebooks: a collaborative environment for education and reproducible research. J Digit Imaging 2018 Jun 27;31(3):290-303 [FREE Full text] [doi: 10.1007/s10278-017-0037-8] [Medline: 29181613]
37. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. FSL. Neuroimage 2012 Aug 15;62(2):782-790. [doi: 10.1016/j.neuroimage.2011.09.015] [Medline: 21979382]
38. Dai J, Qi H, Xiong Y, Li Y, Zhang G, Hu H. Deformable convolutional networks. In: Proceedings of the IEEE International Conference on Computer Vision. 2017 Presented at: ICCV 2017; October 22-29, 2017; Venice, Italy. [doi: 10.1109/iccv.2017.89]
39. Zhu X, Hu H, Lin S, Dai J. Deformable ConvNets V2: more deformable, better results. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019 Presented at: CVPR 2019; June 15-20, 2019; Long Beach, CA. [doi: 10.1109/cvpr.2019.00953]
40. Kim W, Son B, Kim I. ViLT: vision-and-language transformer without convolution or region supervision. In: Proceedings of the 38th International Conference on Machine Learning. 2021 Presented at: ICML 2021; July 18-21, 2021; Online. [doi: 10.48550/arXiv.2102.03334]
41. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, et al. An image is worth 16x16 words: transformers for image recognition at scale. In: Proceedings of the Ninth International Conference on Learning Representations. 2021 Presented at: ICLR 2021; May 3-7, 2021; Online. [doi: 10.48550/arXiv.2010.11929]
42. Li G, Müller M, Thabet A, Ghanem B. DeepGCNs: can GCNs go as deep as CNNs? In: Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019 Presented at: ICCV 2019; October 27-November 02, 2019; Seoul, Korea (South). [doi: 10.1109/iccv.2019.00936]
43. Huo Z, Yang X, Xing C, Zhou Y, Hou P, Lv J. Deep age distribution learning for apparent age estimation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2016 Presented at: CVPRW 2016; June 26-July 01, 2016; Las Vegas, NV. [doi: 10.1109/cvprw.2016.95]

XSL•FO
RenderX

44. Gao BB, Zhou HY, Wu J, Geng X. Age estimation using expectation of label distribution learning. In: Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence. 2018 Presented at: IJCAI 2018; July 13-19, 2018; Stockholm, Sweden. [doi: 10.24963/ijcai.2018/99]
45. Simonyan K, Vedaldi A, Zisserman A. Deep inside convolutional networks: visualising image classification models and saliency maps. In: Proceedings of the International Conference on Learning Representations. 2014 Presented at: ICLR 2014; April 14-16, 2014; Banff, AB. [doi: 10.48550/arXiv.1312.6034]
46. Acknowledgement list for ADNI Publications. Alzheimer's Disease Neuroimaging Initiative. URL: https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf [accessed 2025-07-18]
47. Marshall GA, Zoller AS, Lorius N, Amariglio RE, Locascio JJ, Johnson KA, et al. Functional activities questionnaire items that best discriminate and predict progression from clinically normal to mild cognitive impairment. Curr Alzheimer Res 2015;12(5):493-502 [FREE Full text] [doi: 10.2174/1567205012051505261150003] [Medline: 26017560]
48. Skinner J, Carvalho JO, Potter GG, Thames A, Zelinski E, Crane PK, et al. The Alzheimer's Disease Assessment Scale-Cognitive-Plus (ADAS-Cog-Plus): an expansion of the ADAS-Cog to improve responsiveness in MCI. Brain Imaging Behav 2012 Dec 22;6(4):489-501 [FREE Full text] [doi: 10.1007/s11682-012-9166-3] [Medline: 22614326]
49. Schoenberg MR, Dawson KA, Duff K, Patton D, Scott JG, Adams RL. Test performance and classification statistics for the Rey Auditory Verbal Learning Test in selected clinical samples. Arch Clin Neuropsychol 2006 Oct;21(7):693-703. [doi: 10.1016/j.acn.2006.06.010] [Medline: 16987634]
50. Chun CT, Seward K, Patterson A, Melton A, MacDonald-Wicks L. Evaluation of available cognitive tools used to measure mild cognitive decline: a scoping review. Nutrients 2021 Nov 08;13(11):3974 [FREE Full text] [doi: 10.3390/nu13113974] [Medline: 34836228]
51. Reitan RM. Validity of the trail making test as an indicator of organic brain damage. Percept Mot Ski 1958 Dec;8(7):271-276. [doi: 10.2466/PMS.8.7.271-276]
52. Fotiadis P, Parkes L, Davis KA, Satterthwaite TD, Shinohara RT, Bassett DS. Structure-function coupling in macroscale human brain networks. Nat Rev Neurosci 2024 Oct;25(10):688-704. [doi: 10.1038/s41583-024-00846-6] [Medline: 39103609]
53. Bao H, Wang W, Dong L, Liu Q, Mohammed OK, Aggarwal K, et al. VLMO: unified vision-language pre-training with mixture-of-modality-experts. In: Proceedings of the 36th Conference on Neural Information Processing Systems. 2022 Presented at: NeurIPS 2022; November 28-December 9, 2022; New Orleans, LA. [doi: 10.48550/arXiv.2111.02358]
54. Cox SR, Harris MA, Ritchie SJ, Buchanan CR, Valdés Hernández MC, Corley J, et al. Three major dimensions of human brain cortical ageing in relation to cognitive decline across the eighth decade of life. Mol Psychiatry 2021 Jun 04;26(6):2651-2662 [FREE Full text] [doi: 10.1038/s41380-020-00975-1] [Medline: 33398085]
55. De Benedictis A, Duffau H, Paradiso B, Grandi E, Balbi S, Granieri E, et al. Anatomo-functional study of the temporo-parieto-occipital region: dissection, tractographic and brain mapping evidence from a neurosurgical perspective. J Anat 2014 Aug;225(2):132-151 [FREE Full text] [doi: 10.1111/joa.12204] [Medline: 24975421]
56. Mangalore S, Mukku SS, Vankayalapati S, Sivakumar PT, Varghese M. Shape profile of corpus callosum as a signature to phenotype different dementia. J Neurosci Rural Pract 2021 Jan;12(1):185-192 [FREE Full text] [doi: 10.1055/s-0040-1716805] [Medline: 33531781]
57. Kamal S, Park I, Kim YJ, Kim YJ, Lee U. Alteration of the corpus callosum in patients with Alzheimer's disease: deep learning-based assessment. PLoS One 2021;16(12):e0259051 [FREE Full text] [doi: 10.1371/journal.pone.0259051] [Medline: 34941878]
58. Foster BL, Koslov SR, Aponik-Gremillion L, Monko ME, Hayden BY, Heilbronner SR. A tripartite view of the posterior cingulate cortex. Nat Rev Neurosci 2023 Mar 01;24(3):173-189 [FREE Full text] [doi: 10.1038/s41583-022-00661-x] [Medline: 36456807]
59. Husain M. The parahippocampal region: organization and role in cognitive function. J Neurol Neurosurg Psychiatry 2003;74:829 [FREE Full text] [doi: 10.1136/jnnp.74.6.829-a]
60. Dalboni da Rocha JL, Bramati I, Coutinho G, Tovar Moll F, Sitaram R. Fractional anisotropy changes in parahippocampal cingulum due to Alzheimer's disease. Sci Rep 2020 Feb 14;10(1):2660 [FREE Full text] [doi: 10.1038/s41598-020-59327-2] [Medline: 32060334]
61. Salat DH, Tuch DS, van der Kouwe AJ, Greve DN, Pappu V, Lee SY, et al. White matter pathology isolates the hippocampal formation in Alzheimer's disease. Neurobiol Aging 2010 Feb;31(2):244-256 [FREE Full text] [doi: 10.1016/j.neurobiolaging.2008.03.013] [Medline: 18455835]
62. McClintock AH, Fainstad T, Jauregui J, Yarris LM. Countering bias in assessment. J Grad Med Educ 2021 Oct;13(5):725-726 [FREE Full text] [doi: 10.4300/JGME-D-21-00722.1] [Medline: 34721804]
63. Hess C, Levy B, Hashmi AZ, Hogan J, Greenspan S, Elber A, et al. Subjective versus objective assessment of cognitive functioning in primary care. J Am Board Fam Med 2020;33(3):417-425 [FREE Full text] [doi: 10.3122/jabfm.2020.03.190265] [Medline: 32430373]
64. Pfeffer RI, Kurosaki TT, Harrah CHJ, Chance JM, Filos S. Measurement of functional activities in older adults in the community. J Gerontol 1982 May;37(3):323-329. [doi: 10.1093/geronj/37.3.323] [Medline: 7069156]

XSL•FO
RenderX

65. González DA, Gonzales MM, Resch ZJ, Sullivan AC, Soble JR. Comprehensive evaluation of the functional activities questionnaire (FAQ) and its reliability and validity. Assessment 2022 Jun;29(4):748-763 [FREE Full text] [doi: 10.1177/1073191121991215] [Medline: 33543638]
66. Teng E, Becker BW, Woo E, Knopman DS, Cummings JL, Lu PH. Utility of the functional activities questionnaire for distinguishing mild cognitive impairment from very mild Alzheimer disease. Alzheimer Dis Assoc Disord 2010;24(4):348-353 [FREE Full text] [doi: 10.1097/WAD.0b013e3181e2fc84] [Medline: 20592580]
67. Ahmadzadeh M, Christie GJ, Cosco TD, Arab A, Mansouri M, Wagner KR, et al. Neuroimaging and machine learning for studying the pathways from mild cognitive impairment to Alzheimer's disease: a systematic review. BMC Neurol 2023 Aug 22;23(1):309 [FREE Full text] [doi: 10.1186/s12883-023-03323-2] [Medline: 37608251]
68. Öksüz N, Ghouri R, Taşdelen B, Uludüz D, Özge A. Mild cognitive impairment progression and Alzheimer's disease risk: a comprehensive analysis of 3553 cases over 203 months. J Clin Med 2024 Jan 17;13(2):518 [FREE Full text] [doi: 10.3390/jcm13020518] [Medline: 38256652]
69. Long JM, Holtzman DM. Alzheimer disease: an update on pathobiology and treatment strategies. Cell 2019 Oct 03;179(2):312-339 [FREE Full text] [doi: 10.1016/j.cell.2019.09.001] [Medline: 31564456]
70. Cole JH, Ritchie SJ, Bastin ME, Valdés Hernández MC, Muñoz Maniega S, Royle N, et al. Brain age predicts mortality. Mol Psychiatry 2018 May 25;23(5):1385-1392 [FREE Full text] [doi: 10.1038/mp.2017.62] [Medline: 28439103]
71. Franke K, Gaser C. Ten years of BrainAGE as a neuroimaging biomarker of brain aging: what insights have we gained? Front Neurol 2019;10:789 [FREE Full text] [doi: 10.3389/fneur.2019.00789] [Medline: 31474922]
72. Cummings J. Anti-amyloid monoclonal antibodies are transformative treatments that redefine Alzheimer's disease therapeutics. Drugs 2023 May 15;83(7):569-576 [FREE Full text] [doi: 10.1007/s40265-023-01858-9] [Medline: 37060386]
73. Agarwal A, Gupta V, Brahmbhatt P, Desai A, Vibhute P, Joseph-Mathurin N, et al. Amyloid-related imaging abnormalities in Alzheimer disease treated with anti-amyloid-β therapy. Radiographics 2023 Sep;43(9):e230009 [FREE Full text] [doi: 10.1148/rg.230009] [Medline: 37651273]
74. Yang Z, Wen J, Erus G, Govindarajan ST, Melhem R, Mamourian E, et al. Brain aging patterns in a large and diverse cohort of 49,482 individuals. Nat Med 2024 Oct;30(10):3015-3026 [FREE Full text] [doi: 10.1038/s41591-024-03144-x] [Medline: 39147830]
75. Raising private funds to expand NIH impact. The Foundation for the National Institutes of Health. URL: https://fnih.org/ [accessed 2025-05-29]
76. Health research data for the world. UK Biobank. URL: https://www.ukbiobank.ac.uk/ [accessed 2025-07-18]

Abbreviations

AD: Alzheimer disease
ADNI: Alzheimer’s Disease Neuroimaging Initiative
AUC: area under the receiver operating characteristic curve
BVGN: brain vision graph neural network
CN: cognitively normal
CNN: convolutional neural network
DL: deep learning
FAQ: Functional Activities Questionnaire
GNN: graph neural network
HR: hazard ratio
MAE: mean absolute error
MCI: mild cognitive impairment
MRI: magnetic resonance imaging
ResNet: residual neural network
SFCN: simple fully convolutional network
UKB: UK Biobank

Edited by N Schütz; submitted 23.02.25; peer-reviewed by C Yao, A Invernizzi; comments to author 15.05.25; revised version received 10.06.25; accepted 27.06.25; published 01.08.25.

Please cite as:
Li Z, Li J, Li J, Wang M, Xu A, Huang Y, Yu Q, Zhang L, Li Y, Li Z, Wu X, Bu J, Li W
Development and Validation of a Brain Aging Biomarker in Middle-Aged and Older Adults: Deep Learning Approach
JMIR Aging 2025;8:e73004
URL: https://aging.jmir.org/2025/1/e73004
doi:10.2196/73004
PMID:

XSL•FO
RenderX

©Zihan Li, Jun Li, Jiahui Li, Mengying Wang, Andi Xu, Yushu Huang, Qi Yu, Lingzhi Zhang, Yingjun Li, Zilin Li, Xifeng Wu, Jiajun Bu, Wenyuan Li. Originally published in JMIR Aging (https://aging.jmir.org), 01.08.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Changes in Sarcopenia Status and Subsequent Cardiovascular Outcomes: Prospective Cohort Study

Yuanyue Zhu[1,2*], MD; Kan Wang[3,4*], PhD; Zuolin Lu[5], PhD; Feika Li[1,2], MD, PhD; Yu Xu[3,4], MD, PhD; Linhui Shen[1,2], MD, PhD; Yufang Bi[3,4], Prof Dr; Weiguo Hu[1,2], Prof Dr

[1]Department of Geriatrics, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[2]Medical Center on Aging, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[3]Department of Endocrine and Metabolic Diseases, Shanghai Institute of Endocrine and Metabolic Diseases, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[4]National Clinical Research Center for Metabolic Diseases (Shanghai), Key Laboratory for Endocrine and Metabolic Diseases of the National Health Commission, National Research Center for Translational Medicine, State Key Laboratory of Medical Genomics, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[5]School of Population Medicine and Public Health, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Weiguo Hu, Prof Dr
Department of Geriatrics
Ruijin Hospital, Shanghai Jiao Tong University School of Medicine
No. 197 Ruijin Second Road, Shanghai
Shanghai, 200025
China
Phone: 86 02164370045
Email: wghu@rjh.com.cn

Related Article:

This is a corrected version. See correction statement: https://aging.jmir.org/2025/1/e84642

Abstract

Background: Sarcopenia is associated with cardiovascular diseases (CVDs). However, whether changes in sarcopenia status affect CVD risk remains unclear. In addition, how indoor fuel use impacts the sarcopenia transition process is less well studied.

Objective: This study prospectively examined the association of sarcopenia transitions with CVD risk, while exploring the effect of indoor fuel on these transitions.

Methods: In this prospective observational study, we used data from the China Health and Retirement Longitudinal Study waves 1 to 4 (2011 to 2018). In total, 8739 participants with complete data on sarcopenia and indoor fuel use were included for the indoor fuel use and sarcopenia transition analysis, and 6385 participants without previous CVDs were included for the sarcopenia transition and CVD risk analysis. Sarcopenia transition was defined according to the sarcopenia status at wave 1 (2011) and wave 2 (2013). Incident CVDs included heart diseases, stroke, and composite CVDs. Information on indoor fuel use was obtained at wave 1. Cox proportional hazards models were used to examine the effect of sarcopenia transition on incident CVDs. Logistic regression models were used to investigate the impact of indoor fuel use on these transitions.

Results: During a median of 7.0 years of follow-up, 1233 incident CVDs were documented. Compared to stably normal participants, progressing from a normal state to possible or confirmed sarcopenia brought increased risk of incident CVD (hazard ratio 1.42, 95% CI 1.15-1.77). Conversely, recovering to a normal state was associated with decreased risk (hazard ratio 0.72, 95% CI 0.55-0.95) for baseline participants with possible sarcopenia. In addition, clean fuel use increased the odds of achieving a possible-to-normal transformation (odds ratio 1.32, 95% CI 1.06-1.64), while both solid cooking and heating fuel use were associated with a higher risk of deterioration in sarcopenia status.

XSL•FO
RenderX

Conclusions: An unfavorable transition in sarcopenia status is associated with higher CVD risk, while reversion from possible sarcopenia to a normal state could reduce the risk. Therefore, early intervention for sarcopenia is imperative for CVD prevention, and promoting clean indoor fuel use is recommended.

(JMIR Aging 2025;8:e69860) doi:10.2196/69860

KEYWORDS

sarcopenia; transition; air pollution; cardiovascular disease

Introduction

Background

Sarcopenia is a common age-related condition, characterized by the loss of skeletal muscle mass, strength, and decline in physical function [1]. The prevalence of sarcopenia has been rising worldwide, with an estimated prevalence of 10% to 16% among older adults [2]. Globally, the epidemic of sarcopenia continuously imposes considerable health care and economic burdens as populations age.

Recent studies have shown that the clinical implications of sarcopenia extend well beyond the musculoskeletal system. Instead, it is also closely associated with several other adverse outcomes, including a higher risk of hospitalization and premature mortality [3,4]. Furthermore, emerging evidence suggests that sarcopenia is significantly associated with cardiovascular disease (CVD) [5,6]. For example, Gao et al [6] found that both possible and diagnosed sarcopenia were associated with higher CVD risk. This finding is further supported by an analysis of National Health and Nutrition Examination Survey, which demonstrated that sarcopenia served as an independent risk factor for CVD [7]. In addition, Mendelian randomization studies demonstrated a genetic correlation between sarcopenia and CVD development [8]. Indeed, sarcopenia and CVD have many shared risk factors, including insulin resistance, poor nutrition, smoking, and sedentary lifestyle [9,10]. Mechanistically, the loss of muscle mass can lead to chronic inflammation and insulin resistance, both established contributors to cardiovascular pathogenesis [11]. Moreover, imbalance of anabolic and catabolic homeostasis with reduction of muscle loss might lead to cardiovascular deconditioning as well [12].

However, much of the existing literature focused on the effect of sarcopenia status on CVD at single time points, with less attention paid to the changes in sarcopenia stages during follow-up [7,13-15]. In fact, sarcopenia is not a static condition; rather, it follows a dynamic progression. Starting from a subtle decline of muscle mass, it will either subsequently progress to clinical sarcopenia with noticeable reductions in muscle strength and physical performance, be reversed to a more normal state, or be stabilized with consistent muscle mass and quality [16]. Revealing the longitudinal association of transition in sarcopenia status with CVD can provide a more comprehensive understanding of the effect of sarcopenia on cardiovascular health; more importantly, such insights may identify opportunities to reverse or delay the progression of CVD brought on by muscle disorders [17].

Since the change in sarcopenia status might have long-term effects on CVD, identifying predictors of certain transition stereotypes is clinically important for developing targeted interventions. Various environmental, lifestyle, and socioeconomic factors may impact the progression of sarcopenia, thereby indirectly affecting cardiovascular outcomes [18-20]. Among these potential determinants, indoor air pollution represents a particularly relevant exposure given that older individuals spend a majority of time indoors [21]. Considering the cross-sectionally reported association between indoor air pollution and sarcopenia [22,23], we hypothesize that it may also affect the transition of sarcopenia stages. Despite the potential connections, limited research has examined how indoor air pollution influences sarcopenia transitions and their cardiovascular consequences, representing a critical knowledge gap for clinical practice.

Objectives

To fill these knowledge gaps, we aimed to explore the effect of sarcopenia transitions on the risk of subsequent CVDs and to figure out the role of indoor air pollution in specific changing patterns. Given the tremendous disease burden caused by sarcopenia and CVD, identifying modifiable environmental risk factors for sarcopenia progression and understanding their cardiovascular implications may contribute to more targeted intervention strategies.

Methods

Ethical Considerations

All study procedures have been approved by the Peking University Institutional Review Board (IRB00001052-11015). Participants provided written consent after being fully informed of the design and potential consequences of the study. To ensure privacy and confidentiality, no identifiable information has been included.

Study Population

The China Health and Retirement Longitudinal Study (CHARLS) is a community-based, prospective survey of the Chinese population aged 45 years or above. The detailed study design has been described elsewhere [24]. Briefly, the baseline survey was conducted between June 2011 and March 2012 with individuals selected through multistage probability sampling. Follow-up is performed every 2 years with physical measurements and blood testing. Three follow-up waves (2013, 2015, and 2018) were completed after the baseline survey.

A detailed flowchart for participants' selection in this study is shown in Figure S1 in Multimedia Appendix 1. Among 17,708 participants, 8969 (50.64 %) were excluded due to being

XSL•FO
RenderX

younger than 45 years (n=777, 4.4 %) or having missing information on sarcopenia status at wave 1 in 2011 (n=3992, 22.54 %) or wave 2 in 2013 (n=4200, 23.72 %), leaving 8739 participants for the indoor fuel use and sarcopenia transition analysis. Then, 2354 (13.29 %) participants were further excluded due to prevalent CVDs before wave 2 (2013) or loss to follow-up afterward, leaving 6385 (36.06 %) participants included for sarcopenia transition and CVD risk analysis.

Measurements and Definitions of Covariates

On the basis of a face-to-face computer-assisted personal interview system, structured questionnaires were administered and on-site physical measurements were conducted by trained field workers [25]. Information on socioeconomic factors (education level, marital status, household income, and urban and rural residence), lifestyle factors (drinking and smoking habits, sleep duration, and physical activity), and medical history (prevalent heart disease, stroke, diabetes, hypertension, chronic lung disease, asthma, and cancer, as well as antihypertensive and antidiabetic drugs) was collected through questionnaires. BMI was calculated as body weight divided by the square of height (kg/m^2). Ideal physical activity level was defined as participating in moderate or vigorous activity more than once a week [26]. Ideal sleep duration was defined as having a sleep duration of 6 to 8 hours per night [27]. Mini-Mental State Examination (MMSE) scores were used to assess the cognitive functions [28].

Measurement of Indoor Fuel Use

The CHARLS separately collects data on the energy source for cooking and heating purposes. Types of indoor fuel use were categorized according to a previous study [22]. Specifically, solid fuel was defined as the primary use of coal, crop residue, and wood. Clean fuel was defined as the primary use of solar energy, natural gas, liquefied petroleum gas, and electricity. This binary classification (solid vs clean) was consistently applied to both cooking fuel and heating fuel [29].

Assessment of Sarcopenia Status

Sarcopenia status at wave 1 (2011) and wave 2 (2013) was assessed based on the criteria recommended by the Asian Working Group for Sarcopenia 2019 [30], including muscle strength, physical performance, and appendicular skeletal muscle mass (ASM). Handgrip strength (in kg) was measured by a standardized dynamometer (Yuejian WL-1000); the cut-off point for low handgrip strength was defined as less than 28 kg for men and less than 18 kg for women. ASM was estimated by a validated anthropometric equation, which showed strong agreement with dual x-ray absorptiometry [31,32], and the cut-off value for defining low muscle mass was based on the sex-specific lowest 20% of the height-adjusted muscle mass (ASM/height^2) among the study population. Physical performance was estimated using the gait speed and the chair stand test. For gait speed, the time taken by each participant in walking a 2.5-m distance was recorded. The 5-time chair stand test measured the amount of time needed for the participants to rise continuously 5 times, keeping their arms folded across their chest, from the height of the 47-cm chair, and low physical performance was defined as a gait speed <1.0 m/s or a 5-time chair stand test ≥12 seconds [7].

According to the Asian Working Group for Sarcopenia 2019, possible sarcopenia is defined as low muscle strength or low physical performance, and confirmed sarcopenia is defined as low muscle mass plus low muscle strength or low physical performance. Therefore, all participants in this study were divided into 3 groups at each wave: normal, possible sarcopenia, and confirmed sarcopenia.

Categorization of Sarcopenia Transitions

Transition in sarcopenia status was defined as that between waves 2011 and 2013. In total, 9 distinct transition patterns in sarcopenia were identified based on participants' conditions in 2011 and 2013. To ensure clinical relevance, certain categories were combined. Specifically, the transitions from "normal" to "possible sarcopenia" and from "normal" to "confirmed sarcopenia" were grouped, as both reflect a decline from a healthy baseline and indicate any deterioration in muscle status. Similarly, the transition from "confirmed sarcopenia" to "normal or possible sarcopenia" was combined, as it represents an improvement from the most severe baseline state. Other categories were retained separately due to the distinctive clinical trajectories, despite smaller sample sizes.

Ascertainment of CVDs

The primary outcome was composite CVDs, including heart disease and stroke. Similar to previous studies, CVDs were ascertained by the following questions in each wave: "Have you been told by a physician that you have been diagnosed with a heart attack, angina, coronary heart disease, heart failure, or other heart problems?" "Have you been told by a doctor that you have been diagnosed with a stroke?" Participants who reported having heart disease or stroke were considered to have CVD [7,33]. In subsequent survey waves, participants were asked to confirm the existence of heart disease or stroke reported in the past wave. Once they disputed their earlier reports, the diagnosis would be retrospectively corrected [34]. Participants were followed until the first incident of CVD, death, or the end of the study period (the wave in 2018), whichever came first.

Statistical Analysis

Baseline characteristics are presented as mean (SD) for normally distributed continuous variables, median (IQR) for nonnormally distributed continuous variables, and number (percentage) for categorical variables. For continuous variables, 1-way ANOVA (for normally distributed variables) or the Kruskal-Wallis test (for nonnormally distributed variables) was used to test for overall differences across groups. For categorical variables, chi-square tests were used. The Sankey diagram was constructed to indicate the transitions of sarcopenia between 2011 and 2013.

Cox proportional hazard models were used to estimate the hazard ratios (HRs) and 95% CIs for incident CVDs associated with sarcopenia transitions. Adjustments included baseline age; sex; BMI; marital status (yes or no); high school education (yes or no); rural residence (yes or no); household income; smoking status (never, former, or current smoker); drinking status (never, former, or current drinker); optimal physical activity (yes or

XSL•FO
RenderX

no); ideal sleep duration (yes or no); antihypertensive and antidiabetic medication; chronic diseases (including diabetes, hypertension, chronic lung disease, asthma, and cancer); MMSE scores; and types of fuel use (solid or clean). The proportional hazard assumption was verified through the visual inspection of scaled Schoenfeld residual plots, with no significant deviations observed. In addition, the main analyses were repeated using specific cardiovascular outcomes (heart disease and stroke). To examine the association between baseline fuel use and sarcopenia status transitions between waves 2011 and 2013, multivariable logistic regression models were used. All models were adjusted for the aforementioned covariates except for the types of fuel use.

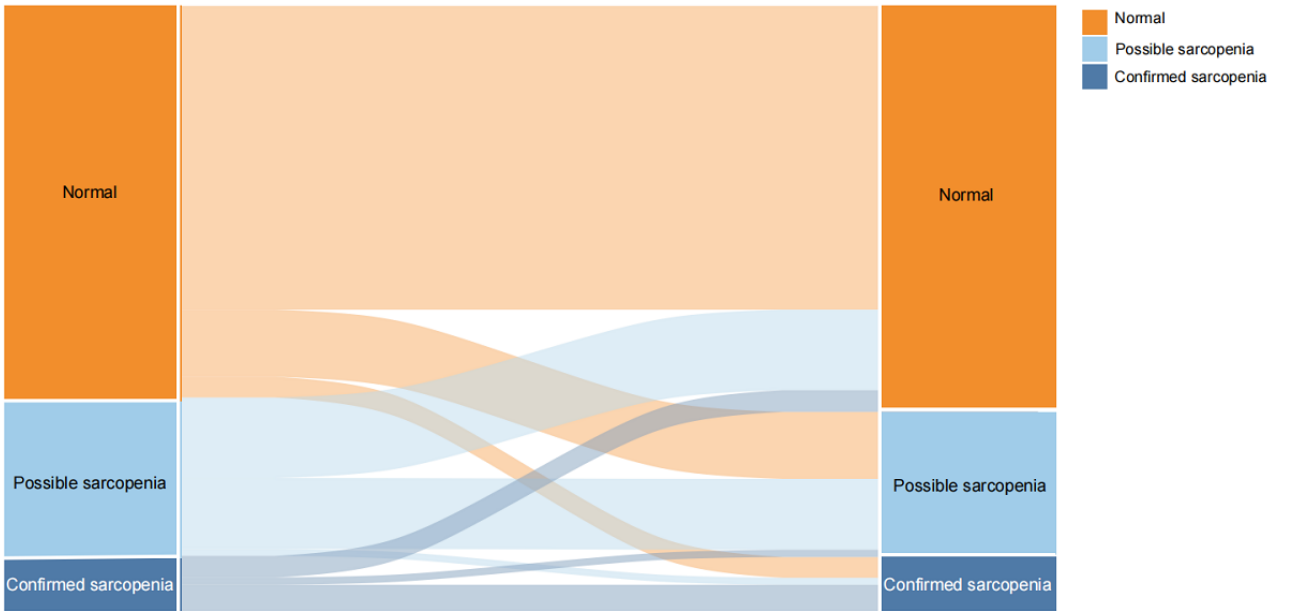
In addition, we performed stratified analyses according to age group (aged ≥65 y vs <65 y), sex, smoking status, physical activity level, and diabetes status to assess potential effect modification. These subgroup variables were selected based on their established associations with both sarcopenia and cardiovascular risk [5]. Statistical significance of multiplicative interaction was evaluated by the likelihood-ratio test [35]. The missing data of covariates were imputed using the multiple imputation with chained equations method [36]. On the basis of a previous study, only those with missing rates less than 20% would be imputed, multiple imputations are set to 10 times, not at least 10 times [37]. Effect estimates were computed separately for these datasets and then combined according to the Rubin rules [38]. To assess the robustness of our findings, we performed sensitivity analyses restricted to observations with complete covariate information. Data were handled and analyzed with SPSS statistics (version 25.0.0.1; IBM Corp) [39] and R CRAN (version 4.3.2; R Foundation) [40]. Specifically, we used the *networkD3* package of R software to generate the Sankey visualizations. All analyses were performed at a significance level of.05 (2 tailed).

Results

Baseline Characteristics

For the analysis of sarcopenia transition and CVD risk, the distribution and the exact numbers of various transition patterns between wave 1 (2011) and wave 2 (2013) is shown in Figure 1.

Figure 1. Transition of sarcopenia status between wave 1 (2011) and wave 2 (2013) in the China Health and Retirement Longitudinal Study (CHARLS).



Baseline characteristics varied significantly across sarcopenia transition groups (Table 1). The stable normal group demonstrated the most advantaged socioeconomic profiles, being the youngest (mean age 56.2, SD 7.7 years), with the highest proportion having high school education (n=397, 11.8%), the greatest likelihood of being married (n=3134, 93.00%), and the highest income levels (median CN ¥14,770, IQR ¥3367.5-¥37,082.5; The exchange rate of US dollars to Chinese yuan is US $1=CN ¥7.15), along with the highest MMSE scores (median 16.0, IQR 13.0-19.0) and greatest representation in rural areas (n=577, 17.1%). In contrast, the stable confirmed sarcopenia group comprised the oldest participants (mean age 70.9, SD 8.4 years) of whom the rates of high school education, marriage, and income levels were also the lowest. Meanwhile, the stable possible sarcopenia group displayed the most concerning lifestyle and health profiles, with the highest prevalence of being a current smoker (n=419, 66.1%), current drinker (n=428, 67.3%), prevalent diabetes (n=476, 74.6%), prevalent hypertension (n=345, 54.1%), and the incidence of heart disease (n=139, 21.9%) and stroke (n=73, 11.6%). Notably, individuals transitioning from possible to confirmed sarcopenia presented with the poorest cognitive and physical function, evidenced by the lowest MMSE scores (median 11.0, IQR 7.0-15.5) and the lowest rate of ideal physical activity levels (n=32, 49%).

Regarding indoor fuel use and sarcopenia transition analysis, baseline characteristics between different transition paradigms were shown in Table S1 in Multimedia Appendix 1, with similar patterns observed.

XSL•FO
RenderX

Table 1. Baseline characteristics of the participants for sarcopenia transition and cardiovascular disease analysis.

	Stable normal (n=3370)	Normal to possible or confirmed sarcopenia (n=890)	Possible sarcopenia to normal (n=827)	Stable possible sarcopenia (n=637)	Possible to confirmed sarcopenia (n=66)	Confirmed sarcopenia to normal or possible (n=297)	Stable confirmed sarcopenia (n=298)	*P* value
Age (y), mean (SD)	56.2 (7.7)	59. 7 (8.7)	57.3 (7.5)	62.3 (9.1)	67.4 (9.8)	64.7 (8.6)	70.9 (8.4)	<.001
Men, n (%)	1788 (53.1)	405 (45.5)	359 (43.4)	251 (39.4)	37 (56.1)	134 (45.1)	112 (37.6)	<.001
High school or above, n (%)	397 (11.8)	50 (5.6)	61 (7.4)	23 (3.6)	1 (1.5)	12 (4)	7 (2.3)	<.001
Married, n (%)	3134 (93.0)	774 (87)	746 (90.2)	533 (83.7)	46 (69.7)	241 (81.1)	196 (65.8)	<.001
Income (CN ¥[a])	14,770.0 (3367.5-37,082.5)	9900.0 (2460.0-28,950.0)	9100.0 (2360.0-30,000.0)	6200.0 (1325.0-26,875.0)	6240.0 (725.0-21,700.0)	4630.0 (1494.8-18,150.0)	3400.0 (900.0-18,122.5)	<.001
Rural, n (%)	577 (17.1)	90 (10.1)	121 (14.6)	81 (12.7)	4 (6.1)	19 (6.4)	23 (7.7)	<.001
BMI (kg/m^2), mean (SD)	23.6 (3.8)	22.9 (3.7)	24.2 (3.0)	24.9 (3.5)	22.3 (3.6)	19.1 (1.6)	19.1 (1.9)	<.001
MMSE[b]	16.0 (13.0-19.0)	14.0 (11.0-17.0)	15.0 (12.0-18.0)	13.0 (10.0-17.0)	11.0 (7.0-15.5)	14.0 (11.0-17.0)	12.0 (9.0-14.0)	<.001
Smoking status, n (%)								<.001
Never	1165 (34.6)	269 (30.2)	242 (29.3)	165 (25.9)	24 (36.4)	94 (31.6)	91 (30.5)	
Former	267 (7.9)	71 (8)	62 (7.5)	50 (7.8)	3 (4.5)	16 (5.4)	15 (5)	
Current	1935 (57.4)	548 (61.6)	518 (62.6)	419 (65.8)	38 (57.6)	186 (62.6)	191 (64.1)	
Drinking status, n (%)								<.001
Never	1292 (38.3)	291 (32.7)	256 (31)	151 (23.7)	17 (25.8)	97 (32.7)	69 (23.2)	
Former	247 (7.3)	55 (6.2)	68 (8.2)	57 (8.9)	7 (10.6)	25 (8.4)	36 (12.1)	
Current	1829 (54.3)	544 (61.1)	502 (60.7)	428 (67.2)	41 (62.1)	174 (58.6)	194 (65.1)	
Ideal physical activity, n (%)	2693 (79.9)	667 (74.9)	589 (71.2)	400 (62.8)	32 (48.5)	220 (74.1)	170 (57)	<.001
Optimal sleep duration, n (%)	2234 (66.3)	527 (59.2)	527 (63.7)	371 (58.2)	34 (51.5)	134 (45.1)	124 (41.6)	<.001
Chronic diseases, n (%)								
Hypertension	1076 (31.9)	342 (38.4)	287 (34.7)	345 (54.2)	22 (33.3)	90 (30.3)	128 (43)	<.001
Diabetes	2405 (71.4)	626 (70.3)	600 (72.6)	476 (74.7)	45 (68.2)	211 (71)	210 (70.5)	0.006
Chronic lung disease	239 (7.1)	81 (9.1)	66 (8)	66 (10.4)	7 (10.6)	32 (10.8)	40 (13.4)	0.001
Asthma	96 (2.8)	41 (4.6)	27 (3.3)	33 (5.2)	4 (6.1)	15 (5.1)	20 (6.7)	0.001
Cancer	22 (0.7)	5 (0.6)	7 (0.8)	8 (1.3)	0 (0)	1 (0.3)	0 (0)	0.345
Incident stroke	146 (4.3)	48 (5.4)	50 (6)	73 (11.5)	5 (7.6)	19 (6.4)	20 (6.7)	
Incident heart disease	369 (10.9)	144 (16.2)	112 (13.5)	139 (21.8)	13 (19.7)	40 (13.5)	55 (18.5)	
Antihypertensive drugs, n (%)	402 (11.9)	126 (14.2)	128 (15.5)	182 (28.6)	8 (12.1)	26 (8.8)	43 (14.4)	<.001
Antidiabetic drugs, n (%)	83 (2.5)	22 (2.5)	26 (3.1)	31 (4.9)	2 (3)	5 (1.7)	3 (1)	0.008
Solid fuel for cooking, n (%)	1850 (54.9)	587 (66)	533 (64.4)	443 (69.5)	51 (77.3)	216 (72.7)	234 (78.5)	<.001
Solid fuel for heating, n (%)	2598 (77.1)	745 (83.7)	691 (83.6)	541 (84.9)	62 (93.9)	268 (90.2)	265 (88.9)	<.001

XSL•FO
RenderX

	Stable normal (n=3370)	Normal to possible or confirmed sarcopenia (n=890)	Possible sarcopenia to normal (n=827)	Stable possible sarcopenia (n=637)	Possible to confirmed sarcopenia (n=66)	Confirmed sarcopenia to normal or possible (n=297)	Stable confirmed sarcopenia (n=298)	*P* value
Handgrip strength, n (%)	35.0 (29.0-42.8)	31.5 (26.5-38.0)	30.0 (24.0-37.3)	27.0 (21.0-34.5)	25.5 (21.6-30.3)	26.0 (20.5-33.0)	23.0 (17.1-27.0)	<.001
Appendicular skeletal muscle (kg)	17.8 (14.3-20.8)	16.4 (13.4-19.6)	16.9 (14.2-20.2)	16.7 (13.9-19.9)	17.1 (12.2-18.6)	12.3 (10.5-17.3)	11.4 (10.1-16.1)	<.001

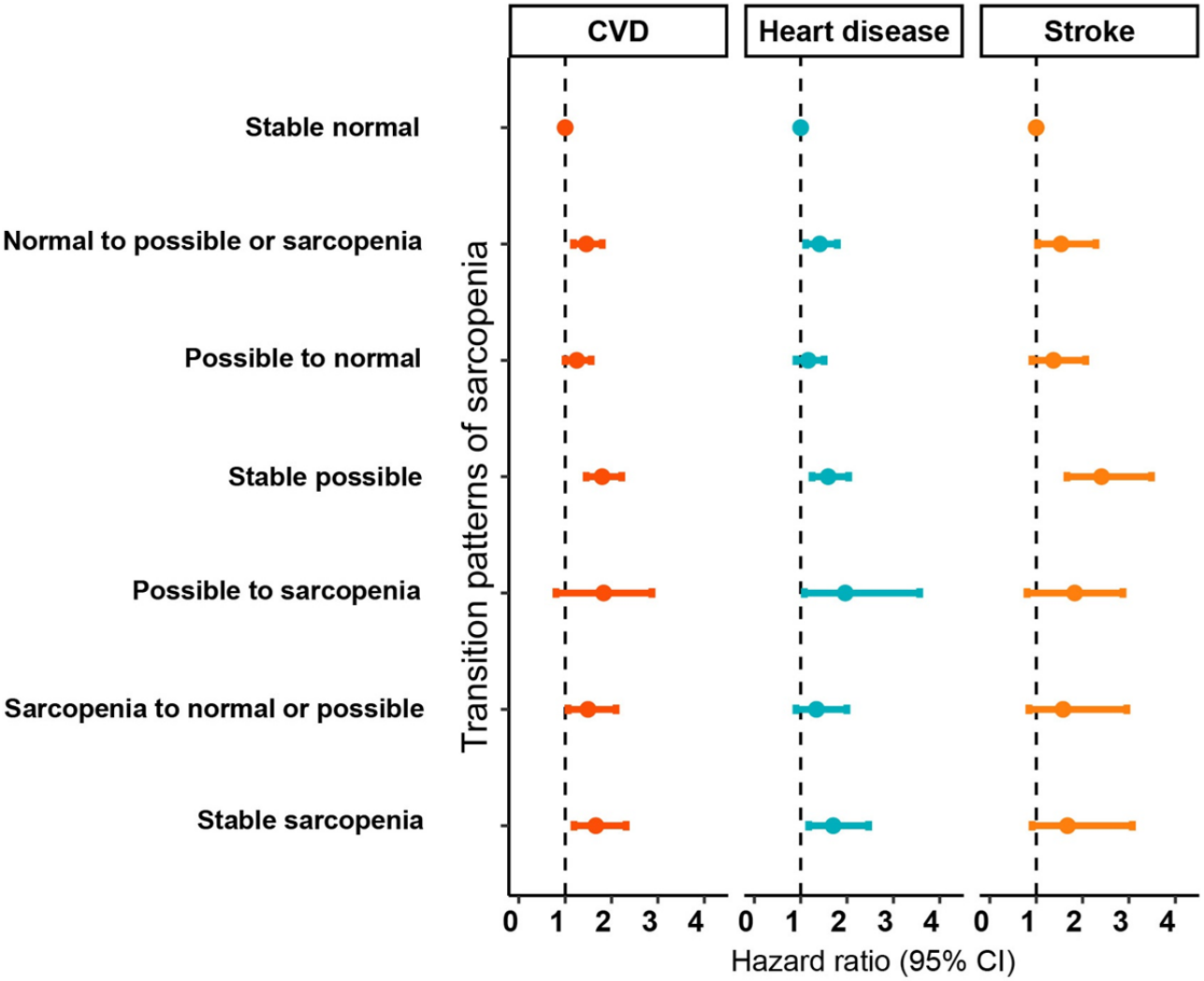
[a]US $1=CN ¥7.15.

[b]MMSE: Mini-Mental State Examination.

Association Between Transition in Sarcopenia Status and CVD

The effect of transition in sarcopenia status on CVD risk during follow-up is shown in Figure 2. Participants maintaining stably normal muscle function harbored the lowest risk for composite CVD and also its subtypes and were thus used as the reference.

Figure 2. Associations between sarcopenia transitions and the risk of cardiovascular diseases.



Compared with the reference group, HRs (95% CIs) for those who stayed possible sarcopenia and confirmed sarcopenia were 1.71 (1.43-2.05) and 1.55 (1.16-2.06), respectively. For participants undergoing an unfavorable shift to a more severe state, significantly higher risks of CVD were also observed. Specifically, progression from normal to possible or confirmed sarcopenia and possible to confirmed sarcopenia were associated with a 40% and 73% higher risk of CVD (Table S2 in Multimedia Appendix 1) compared with those maintaining a normal state, respectively. Importantly, the risk of CVD was still elevated among patients recovering from confirmed to possible sarcopenia and a normal state (HR 1.42, 95% CI 1.05-1.90). In contrast, no significant increase in CVD risk was observed in the “possible sarcopenia to normal” group (HR

XSL•FO
RenderX

1.13, 95% CI 0.94-1.36). Results were generally consistent for heart diseases. However, the risk of stroke was only significantly elevated among stable possible (HR 2.04, 95% CI 1.51-2.76). Further stratification analysis revealed that the association between sarcopenia transition and CVD was more prominent among male, physically active, and nondiabetic individuals (Tables S3-S7 in Multimedia Appendix 1).

Because most participants maintained their baseline status during follow-up, we subsequently used 3 stable groups as the corresponding references (eg, those who remained in a normal state, those with possible sarcopenia, and those with confirmed sarcopenia throughout the study period). Associations between sarcopenia transition and CVD of different baseline sarcopenia statuses are shown in Table 2. Among participants with normal baseline status, those who progressed to possible or confirmed sarcopenia showed elevated risk of CVD (HR 1.42, 95% CI 1.15-1.77) compared to stable normal participants. For those with possible sarcopenia at baseline, participants who improved to normal status demonstrated reduced risk (HR 0.72, 95% CI 0.55-0.95), while those who progressed to confirmed sarcopenia showed potentially higher risk, although without statistical significance. Moreover, compared with participants maintaining confirmed sarcopenia at baseline, those who recovered to possible sarcopenia or normal status had an HR of 0.80 (95% CI 0.51-1.23). Regarding specific outcomes, deterioration from normal status at baseline was associated with a 40% increased risk of heart disease, while the remission from possible sarcopenia to a normal state was not associated with significantly elevated risk. Moreover, recovery from possible sarcopenia to a normal state was associated with a lower risk of stroke (HR 0.54, 95% CI 0.34-0.86), compared with their stable counterparts. Results of stratification analysis within different baseline sarcopenia statuses were presented in Tables S8 to S12 in Multimedia Appendix 1, and significant interaction effects of sex and physical activity were observed.

Table 2. Associations between sarcopenia transition and cardiovascular diseases of different baseline sarcopenia statuses.

Sarcopenia transition	Cardiovascular disease, HR[a] (95% CI)	Heart disease, HR (95% CI)	Stroke, HR (95% CI)
Stable normal	1.00 (reference)	1.00 (reference)	1.00 (reference)
Normal to possible or confirmed sarcopenia	1.42 (1.15-1.77)	1.40 (1.10-1.79)	1.42 (0.94-2.14)
Stable possible sarcopenia	1.00 (reference)	1.00 (reference)	1.00 (reference)
Possible sarcopenia to normal	0.72 (0.55-0.95)	0.80 (0.58-1.00)	0.54 (0.34-0.86)
Possible to confirmed sarcopenia	1.09 (0.62-1.93)	1.16 (0.61-2.22)	1.05 (0.29-2.38)
Stable confirmed sarcopenia	1.00 (reference)	1.00 (reference)	1.00 (reference)
Confirmed sarcopenia to normal or possible	0.80 (0.51-1.23)	0.68 (0.41-1.12)	0.84 (0.37-1.90)

[a]HR: hazard ratio. Adjusted for baseline age, sex, BMI, married (yes or no); high school education (yes or no); rural residence (yes or no); household income; smoking status (never, former, or current smoker); drinking status (never, former, or current drinker); optimal physical activity (yes or no); ideal sleep duration (yes or no); antihypertensive and antidiabetic medication; prevalent diseases (including diabetes, hypertension, chronic lung disease, asthma, and cancer); and MMSE scores and types of fuel use (solid or clean).

Associations of Types of Cooking and Heating Fuel Use With Sarcopenia Transitions

The effect of types of fuel use on transition is shown in Table 3. Briefly, clean fuel use is a prediction of improvement in sarcopenia status. Compared with solid fuel, clean fuel use increased the odds of achieving the possible-to-normal transformation by 32% (*P*=.01). Meanwhile, both solid heating fuel (odds ratio [OR] 1.32, 95% CI 1.14-1.53) and solid cooking fuel (OR 1.20, 95% CI 1.01-1.44) were associated with a higher risk of the progression from normal to possible or confirmed sarcopenia. The association between fuel use and confirmed sarcopenia to normal or possible sarcopenia was not significant. In addition, significant effect modification was observed for sex, physical activity, and diabetes status (Tables S13-S17 in Multimedia Appendix 1).

Table 3. Associations of the use of cooking and heating fuel and sarcopenia transition.

Fuel type	Normal to possible or confirmed sarcopenia			Possible sarcopenia to normal			Confirmed sarcopenia to normal or possible		
	Clean fuel, OR[a] (95% CI)	Solid fuel, OR (95% CI)	*P* value	Clean fuel, OR (95% CI)	Solid fuel, OR (95% CI)	*P* value	Clean fuel, OR (95% CI)	Solid fuel, OR (95% CI)	*P* value
Cooking	1.00 (reference)	1.32 (1.14-1.53)	<.001	1.00 (reference)	1.32 (1.06-1.64)	.01	1.00 (reference)	1.16 (0.79-1.70)	.44
Heating	1.00 (reference)	1.20 (1.01-1.44)	.04	1.00 (reference)	0.92 (0.70-1.21)	.55	1.00 (reference)	0.98 (0.58-1.66)	.95

[a]OR: odds ratio.

Adjusted for baseline age, sex, BMI, married (yes or no); high school education (yes or no); rural residence (yes or no); household income; smoking status (never, former, or current smoker); drinking status (never, former, or current drinker); optimal physical activity (yes or no); ideal sleep duration (yes or no); antihypertensive and antidiabetic medication; prevalent

XSL•FO
RenderX

diseases (including diabetes, hypertension, chronic lung disease, asthma, and cancer); and MMSE scores.

Sensitivity Analysis

All missing rates of the covariates were <5% (Table S18 in Multimedia Appendix 1). Therefore, we repeated the analysis using the complete dataset without any missing covariates (Tables S19-S21 in Multimedia Appendix 1). These sensitivity analyses produced results consistent with the primary findings.

Discussion

Principal Findings

In this community-based cohort study conducted in China, we found that progression to worse sarcopenia status was associated with increased cardiovascular risk. Conversely, participants with possible sarcopenia at baseline but who returned to normal status demonstrated considerably reduced risk. Moreover, clean fuel use was associated with favorable transitions in sarcopenia status, whereas solid fuel use was a risk factor for negative shifts toward a deteriorated sarcopenia status.

Previous studies have well illustrated the association between different sarcopenia stages and CVD [5,7,14,15,41]. Despite these important observations, it is noteworthy that sarcopenia is a dynamic process rather than a stable state, and divergent changing paths of sarcopenia are associated with significantly distinct health outcomes [16]. For example, the transition toward deteriorated muscle mass was associated with metabolic abnormalities [42]. Unfortunately, little research has examined the association between sarcopenia transition and CVD risk.

In this study, we comprehensively reported that different trajectories of sarcopenia transitions were associated with markedly different risks of composite CVD. It is expected that a stably pathological state (ie, stable possible and confirmed sarcopenia) and the worsening in sarcopenia status were both associated with a higher risk of CVD. As shown in Table 1, the prevalence of hypertension and diabetes and the suboptimal lifestyle factor are higher among these categories. These negative influencers are also common risk factors of CVD. Moreover, reversion from possible sarcopenia to a normal state, but not from confirmed to possible sarcopenia, was associated with reduced risk. It is indicated that advanced sarcopenia involves increasingly irreversible metabolic alterations that persist despite apparent muscle mass recovery [43]. These findings suggest that intervention at the relatively early stages of sarcopenia may be more clinically beneficial. The weaker associations with stroke across all transition patterns may also reflect the mechanistic complexity of cerebrovascular disease, suggesting that the cardiovascular effects of sarcopenia may be more specific to cardiac rather than cerebrovascular pathways [7,9]. Methodological considerations also significantly contribute to our null findings. For example, the transition from possible to confirmed sarcopenia showed consistently nonsignificant associations across all cardiovascular outcomes, likely reflecting insufficient statistical power given the small sample size (n=66, 1 %), although point estimates generally favored increased risk. Similarly, the substantially reduced statistical power due to fewer stroke events (n=361, 5.7 %) compared to heart disease cases (n=872, 13.6 %) fundamentally limits our ability to detect modest but potentially clinically meaningful associations, emphasizing the need for larger studies with extended follow-up periods. We also noticed that the associations between sarcopenia transition and CVD were particularly pronounced in men, physically active individuals, and those without a history of diabetes. Given that both diabetes and physical inactivity are important risk factors for CVD, it may obscure or diminish the incremental risk associated with sarcopenia transition. Moreover, the gender difference in the effect of sarcopenia has been reported previously [44], which indicated that a male participant could exhibit a more pronounced susceptibility to the detrimental effects of sarcopenia.

Given that the transition and trajectory of sarcopenia have substantially varied health effects, understanding the risk factors is essential for developing effective interventions. Traditional determinants included aging, malnutrition, and physical inactivity [18,45]. However, little attention has been paid toward the environmental drivers. Previous studies indicated that fuel use was significantly associated with a higher risk of developing sarcopenia among the general population [46,47]. Our findings further demonstrated that solid fuel use for both cooking and heating contributed to declining muscle performance.

The mass and function of skeletal muscle progressively decline with aging, and sarcopenia is thus advancing [48]. This process is insidious and often accompanied by physical disability and chronic disease, including CVD. Sarcopenia is associated with cardiovascular remodeling and dysfunction through several pathways: insulin resistance, endothelial inflammation, oxidative stress, and neurohormonal abnormalities [49]. Efforts aimed at these pathways may help improve cardiovascular outcomes in patients with sarcopenia. Furthermore, explanations of how fuel use impacts the transition process remained uncertain. The hypothesis included harmful materials and heavy metals released by solid fuel, which induce neuromuscular degeneration [50,51]. Nonetheless, the exact mechanisms are yet to be explored.

Strengths and Limitations

This study has several strengths. This prospective study provides important evidence on the effect of sarcopenia transition on CVD risk, with the longitudinal design minimizing potential reverse causality. In addition, most attention in sarcopenia research has been paid toward the older adults, despite that the decline in muscle mass and quality already begins in the early 40s [52]. By including participants aged 45 years and above, this provided insights into a more comprehensive scope. However, some limitations are to be addressed. First, CVDs were identified based on self-reported physician diagnosis. However, the validity of self-reported CVDs has been confirmed by previous studies using the CHARLS database [53]. Second, ASM values were estimated by an equation rather than dual x-ray absorptiometry or magnetic resonance, despite the validity of this method having been established in previous studies [7,16]. Third, the lack of detailed temporal fuel exposure data constrains our ability to examine the dose-response relationships and critical exposure windows for sarcopenia development. Finally, it is important to note that the CHARLS only consisted

XSL•FO
RenderX

of Chinese participants, restricting the generalizability of our findings.

Our study suggested that the unfavorable transition of sarcopenia status is associated with a higher CVD risk, and reversion from possible sarcopenia may help reduce the related risk. Therefore, intervention at an earlier stage is imperative, and promoting the use of clean fuel may be helpful in averting muscle loss.

Acknowledgments

This work was supported by the Ruijin Hospital Youth Incubation Project (KY20240805), the China Postdoctoral Science Foundation (2024M752061), and the Noncommunicable Chronic Diseases-National Science and Technology Major Project (2024ZD0536700). The funders had no role in study design, data collection and analysis, the decision to publish, or the preparation of the manuscript. The authors would like to thank the staff and the participants in the China Health and Retirement Longitudinal Study.

Data Availability

The original China Health and Retirement Longitudinal Study dataset is available on the web. The full dataset used in this analysis is available from the corresponding author upon reasonable request.

Authors' Contributions

YZ, KW, YB, and WH are responsible for the study concept and design. YZ and KW compiled the statistical dataset, performed the statistical analyses, and wrote the manuscript. All authors contributed to interpreting the data and critically revising the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sensitivity analysis of the association between sarcopenia transition, indoor fuel use, and cardiovascular diseases.
[DOCX File , 83 KB - aging_v8i1e69860_app1.docx]

References

1. Rosenberg IH. Sarcopenia: origins and clinical relevance. J Nutr 1997 May;127(5 Suppl):990S-991S. [doi: 10.1093/jn/127.5.990S] [Medline: 9164280]
2. Yuan S, Larsson SC. Epidemiology of sarcopenia: prevalence, risk factors, and consequences. Metabolism 2023 Jul;144:155533 [FREE Full text] [doi: 10.1016/j.metabol.2023.155533] [Medline: 36907247]
3. Sayer AA, Cooper R, Arai H, Cawthon PM, Ntsama Essomba M, Fielding RA, et al. Sarcopenia. Nat Rev Dis Primers 2024 Sep 19;10(1):68. [doi: 10.1038/s41572-024-00550-w] [Medline: 39300120]
4. Damluji AA, Alfaraidhy M, AlHajri N, Rohant NN, Kumar M, Al Malouf C, et al. Sarcopenia and cardiovascular diseases. Circulation 2023 May 16;147(20):1534-1553 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.123.064071] [Medline: 37186680]
5. Zuo X, Li X, Tang K, Zhao R, Wu M, Wang Y, et al. Sarcopenia and cardiovascular diseases: a systematic review and meta-analysis. J Cachexia Sarcopenia Muscle 2023 Jun;14(3):1183-1198 [FREE Full text] [doi: 10.1002/jcsm.13221] [Medline: 37002802]
6. Gao K, Cao LF, Ma WZ, Gao YJ, Luo MS, Zhu J, et al. Association between sarcopenia and cardiovascular disease among middle-aged and older adults: findings from the China health and retirement longitudinal study. EClinicalMedicine 2022 Feb;44:101264 [FREE Full text] [doi: 10.1016/j.eclinm.2021.101264] [Medline: 35059617]
7. Yu Z, Zhao Z, Ding C, Liu Q, Ma T, Han X, et al. The association between sarcopenia and cardiovascular disease: an investigative analysis from the NHANES. Nutr Metab Cardiovasc Dis 2025 Jul;35(7):103864 [FREE Full text] [doi: 10.1016/j.numecd.2025.103864] [Medline: 40016026]
8. Liu X, Wang Y, Wang Z, Li L, Yang H, Liu J, et al. Association between sarcopenia-related traits and cardiovascular diseases: a bi-directional Mendelian randomization study. Front Endocrinol (Lausanne) 2023 Oct 13;14:1237971 [FREE Full text] [doi: 10.3389/fendo.2023.1237971] [Medline: 37900136]
9. Haines MS, Dichtel LE, Santoso K, Torriani M, Miller KK, Bredella MA. Association between muscle mass and insulin sensitivity independent of detrimental adipose depots in young adults with overweight/obesity. Int J Obes (Lond) 2020 Sep;44(9):1851-1858 [FREE Full text] [doi: 10.1038/s41366-020-0590-y] [Medline: 32404950]
10. Sartori R, Romanello V, Sandri M. Mechanisms of muscle atrophy and hypertrophy: implications in health and disease. Nat Commun 2021 Jan 12;12(1):330 [FREE Full text] [doi: 10.1038/s41467-020-20123-1] [Medline: 33436614]

XSL•FO
RenderX

11. Boonpor J, Pell JP, Ho FK, Celis-Morales C, Gray SR. In people with type 2 diabetes, sarcopenia is associated with the incidence of cardiovascular disease: A prospective cohort study from the UK Biobank. Diabetes Obes Metab 2024 Feb;26(2):524-531 [FREE Full text] [doi: 10.1111/dom.15338] [Medline: 37881162]
12. Jauffret C, Périchon R, Lamer A, Cortet B, Chazard E, Paccou J. Association between sarcopenia and risk of major adverse cardiac and cerebrovascular events-UK Biobank database. J Am Geriatr Soc 2024 Mar;72(3):693-706. [doi: 10.1111/jgs.18664] [Medline: 37945290]
13. Xia MF, Chen LY, Wu L, Ma H, Li XM, Li Q, et al. Sarcopenia, sarcopenic overweight/obesity and risk of cardiovascular disease and cardiac arrhythmia: a cross-sectional study. Clin Nutr 2021 Feb;40(2):571-580. [doi: 10.1016/j.clnu.2020.06.003] [Medline: 32593523]
14. Luo YX, Zhou XH, Heng T, Yang LL, Zhu YH, Hu P, et al. Bidirectional transitions of sarcopenia states in older adults: the longitudinal evidence from CHARLS. J Cachexia Sarcopenia Muscle 2024 Oct;15(5):1915-1929 [FREE Full text] [doi: 10.1002/jcsm.13541] [Medline: 39001569]
15. Zheng Y, Feng J, Yu Y, Ling M, Wang X. Advances in sarcopenia: mechanisms, therapeutic targets, and intervention strategies. Arch Pharm Res 2024 Apr 09;47(4):301-324. [doi: 10.1007/S12272-024-01493-2]
16. Sun B, Li S, Wang Y, Xiao W, Zhao H, Liu X, et al. Sarcopenia transitions and influencing factors among Chinese older adults with multistate Markov model. Innov Aging 2023;7(8):igad105 [FREE Full text] [doi: 10.1093/geroni/igad105] [Medline: 37954524]
17. Murphy RA, Ip EH, Zhang Q, Boudreau RM, Cawthon PM, Newman AB, Health, Aging,Body Composition Study. Transition to sarcopenia and determinants of transitions in older adults: a population-based study. J Gerontol A Biol Sci Med Sci 2014 Jun;69(6):751-758 [FREE Full text] [doi: 10.1093/gerona/glt131] [Medline: 24013673]
18. Li S, Wang Y, Li L, Wu H. The relationship between sleep disorders and combination of diabetes and sarcopenia in adults aged 45 years or older: 10-year nationwide prospective cohort study. JMIR Aging 2025 Jun 02;8:e66372 [FREE Full text] [doi: 10.2196/66372] [Medline: 40456103]
19. Jiang Q, Wang S, Zhang H, Guo Y, Lou Y, Huang S, et al. The association between solid fuel use and visual impairment among middle-aged and older Chinese adults: nationwide population-based cohort study. JMIR Public Health Surveill 2023 Jul 26;9:e43914 [FREE Full text] [doi: 10.2196/43914] [Medline: 37494091]
20. Hu Z, Tian Y, Song X, Zeng F, Yang A. Associations between indoor air pollution for cooking and heating with muscle and sarcopenia in Chinese older population. J Cachexia Sarcopenia Muscle 2023 Oct;14(5):2029-2043 [FREE Full text] [doi: 10.1002/jcsm.13281] [Medline: 37448255]
21. Shi W, Zhang T, Yu Y, Luo L. Association of indoor solid fuel use and long-term exposure to ambient PM (2.5) with sarcopenia in China: a nationwide cohort study. Chemosphere 2023 Dec;344:140356. [doi: 10.1016/j.chemosphere.2023.140356] [Medline: 37802484]
22. Chen LK, Woo J, Assantachai P, Auyeung TW, Chou MY, Iijima K, et al. Asian working group for sarcopenia: 2019 consensus update on sarcopenia diagnosis and treatment. J Am Med Dir Assoc 2020 Mar;21(3):300-7.e2. [doi: 10.1016/j.jamda.2019.12.012] [Medline: 32033882]
23. Bovet P, Banatvala N, Khaw KT, Reddy KS. Cardiovascular disease: burden, epidemiology and risk factors. In: Banatvala N, Bovet P, editors. Noncommunicable Diseases: A Compendium. New York, NY: Routledge; 2023:45-51.
24. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). Int J Epidemiol 2014 Feb 12;43(1):61-68 [FREE Full text] [doi: 10.1093/ije/dys203] [Medline: 23243115]
25. Chen X, Crimmins E, Hu PP, Kim JK, Meng Q, Strauss J, et al. Venous blood-based biomarkers in the china health and retirement longitudinal study: rationale, design, and results from the 2015 wave. Am J Epidemiol 2019 Nov 01;188(11):1871-1877 [FREE Full text] [doi: 10.1093/aje/kwz170] [Medline: 31364691]
26. Zhou M, Kuang L, Hu N. The association between physical activity and intrinsic capacity in Chinese older adults and its connection to primary care: China health and retirement longitudinal study (CHARLS). Int J Environ Res Public Health 2023 Mar 31;20(7):5361 [FREE Full text] [doi: 10.3390/ijerph20075361] [Medline: 37047975]
27. Chaput JP, Dutil C, Sampasa-Kanyinga H. Sleeping hours: what is the ideal number and how does age impact this? Nat Sci Sleep 2018;10:421-430 [FREE Full text] [doi: 10.2147/NSS.S163071] [Medline: 30568521]
28. Andersson B, Luo H. The mini-mental state examination in a Chinese population: reliability, validity, and measurement invariance. Innov Aging 2023;7(Supplement_1):385. [doi: 10.1093/geroni/igad104.1275]
29. Lin L, Wang HH, Liu Y, Lu C, Chen W, Guo VY. Indoor solid fuel use for heating and cooking with blood pressure and hypertension: a cross-sectional study among middle-aged and older adults in China. Indoor Air 2021 Nov 12;31(6):2158-2166 [FREE Full text] [doi: 10.1111/ina.12872] [Medline: 34118166]
30. Wen X, Wang M, Jiang CM, Zhang YM. Anthropometric equation for estimation of appendicular skeletal muscle mass in Chinese adults. Asia Pac J Clin Nutr 2011;20(4):551-556 [FREE Full text] [Medline: 22094840]
31. Yang M, Hu X, Wang H, Zhang L, Hao Q, Dong B. Sarcopenia predicts readmission and mortality in elderly patients in acute care wards: a prospective study. J Cachexia Sarcopenia Muscle 2017 Apr;8(2):251-258 [FREE Full text] [doi: 10.1002/jcsm.12163] [Medline: 27896949]
32. He D, Wang Z, Li J, Yu K, He Y, He X, et al. Changes in frailty and incident cardiovascular disease in three prospective cohorts. Eur Heart J 2024 Mar 27;45(12):1058-1068. [doi: 10.1093/eurheartj/ehad885] [Medline: 38241094]

XSL•FO
RenderX

33. Zhang Z, Xu H, Zhang R, Yan Y, Ling X, Meng Y, et al. Frailty and depressive symptoms in relation to cardiovascular disease risk in middle-aged and older adults. Nat Commun 2025 Jul 01;16(1):6008 [FREE Full text] [doi: 10.1038/s41467-025-61089-2] [Medline: 40595747]
34. Altman DG, Bland JM. Interaction revisited: the difference between two estimates. BMJ 2003 Jan 25;326(7382):219 [FREE Full text] [doi: 10.1136/bmj.326.7382.219] [Medline: 12543843]
35. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. Stat Med 2011 Feb 20;30(4):377-399. [doi: 10.1002/sim.4067] [Medline: 21225900]
36. Madley-Dowd P, Hughes R, Tilling K, Heron J. The proportion of missing data should not be used to guide decisions on multiple imputation. J Clin Epidemiol 2019 Jun;110:63-73 [FREE Full text] [doi: 10.1016/j.jclinepi.2019.02.016] [Medline: 30878639]
37. Campos AM, Moura FA, Santos SN, Freitas WM, Sposito AC, Brasilia Study on Healthy Aging and Brasilia Heart Study. Sarcopenia, but not excess weight or increased caloric intake, is associated with coronary subclinical atherosclerosis in the very elderly. Atherosclerosis 2017 Mar;258:138-144. [doi: 10.1016/j.atherosclerosis.2017.01.005] [Medline: 28129889]
38. Giganti MJ, Shepherd BE. Multiple-imputation variance estimation in studies with missing or misclassified inclusion criteria. Am J Epidemiol 2020 Dec 01;189(12):1628-1632 [FREE Full text] [doi: 10.1093/aje/kwaa153] [Medline: 32685964]
39. IBM SPSS Statistics 25. IBM Corp. URL: https://www.ibm.com/support/pages/downloading-ibm-spss-statistics-25 [accessed 2025-05-29]
40. R Core Team. R-4.5.1 for Windows. Cran R. URL: https://cran.r-project.org/bin/windows/base/ [accessed 2025-05-29]
41. Wu SE, Chen WL. Transitional states of sarcopenia: the trajectory of fat accumulation and glucose fluctuation on risk of metabolic syndrome. Ann N Y Acad Sci 2021 Nov;1503(1):60-71. [doi: 10.1111/nyas.14607] [Medline: 33942910]
42. Duarte MP, Pereira MS, Baião VM, Vieira FA, Silva MZ, Krug RR, SARC-HD Study Group. Design and methodology of the SARCopenia trajectories and associations with adverse clinical outcomes in patients on HemoDialysis: the SARC-HD study. BMC Nephrol 2023 Aug 15;24(1):239 [FREE Full text] [doi: 10.1186/s12882-023-03168-4] [Medline: 37582699]
43. Verstraeten LM, van Wijngaarden JP, Pacifico J, Reijnierse EM, Meskers CG, Maier A. Association between malnutrition and stages of sarcopenia in geriatric rehabilitation inpatients: RESORT. Clin Nutr 2021 Jun;40(6):4090-4096 [FREE Full text] [doi: 10.1016/j.clnu.2021.02.007] [Medline: 33622573]
44. Zhang B, Huang L, Zhu X, Ran L, Zhao H, Zhu Z, et al. Impact of household solid fuel use on sarcopenia in China: a nationwide analysis. Sci Total Environ 2023 Jun 15;877:162814. [doi: 10.1016/j.scitotenv.2023.162814] [Medline: 36933714]
45. Chen CH, Huang LY, Lee KY, Wu CD, Chiang HC, Chen BY, et al. Effects of PM2.5 on skeletal muscle mass and body fat mass of the elderly in Taipei, Taiwan. Sci Rep 2019 Aug 01;9(1):11176 [FREE Full text] [doi: 10.1038/s41598-019-47576-9] [Medline: 31371747]
46. Zengin A, Ó Breasail M, Parsons CM, Jarjou LM, Janha RE, Jobe M, et al. Sex-specific associations between cardiovascular risk factors and physical function: the Gambian Bone and Muscle Ageing Study. J Cachexia Sarcopenia Muscle 2023 Feb;14(1):84-92 [FREE Full text] [doi: 10.1002/jcsm.13069] [Medline: 36346161]
47. Larsson L, Degens H, Li M, Salviati L, Lee YI, Thompson W, et al. Sarcopenia: aging-related loss of muscle mass and function. Physiol Rev 2019 Jan 01;99(1):427-511 [FREE Full text] [doi: 10.1152/physrev.00061.2017] [Medline: 30427277]
48. Kinugasa Y, Yamamoto K. The challenge of frailty and sarcopenia in heart failure with preserved ejection fraction. Heart 2017 Feb;103(3):184-189. [doi: 10.1136/heartjnl-2016-309995] [Medline: 27940967]
49. Ali MU, Yu Y, Yousaf B, Munir MA, Ullah S, Zheng C, et al. Health impacts of indoor air pollution from household solid fuel on children and women. J Hazard Mater 2021 Aug 15;416:126127. [doi: 10.1016/j.jhazmat.2021.126127] [Medline: 34492921]
50. Hepburn C, Qi Y, Stern N, Ward B, Xie C, Zenghelis D. Towards carbon neutrality and China's 14th Five-Year Plan: clean energy transition, sustainable urban development, and investment priorities. Environ Sci Ecotechnol 2021 Oct;8:100130 [FREE Full text] [doi: 10.1016/j.ese.2021.100130] [Medline: 36156997]
51. Foley RN, Wang C, Ishani A, Collins AJ, Murray AM. Kidney function and sarcopenia in the United States general population: NHANES III. Am J Nephrol 2007;27(3):279-286. [doi: 10.1159/000101827] [Medline: 17440263]
52. Jiang M, Ren X, Han L, Zheng X. Associations between sarcopenic obesity and risk of cardiovascular disease: A population-based cohort study among middle-aged and older adults using the CHARLS. Clin Nutr 2024 Mar;43(3):796-802. [doi: 10.1016/j.clnu.2024.02.002] [Medline: 38350287]
53. Liu Y, Ning N, Sun T, Guan H, Liu Z, Yang W, et al. Association between solid fuel use and nonfatal cardiovascular disease among middle-aged and older adults: findings from the China Health and Retirement Longitudinal Study (CHARLS). Sci Total Environ 2023 Jan 15;856(Pt 2):159035. [doi: 10.1016/j.scitotenv.2022.159035] [Medline: 36191716]

Abbreviations

ASM: appendicular skeletal muscle mass
CHARLS: China Health and Retirement Longitudinal Study
CVD: cardiovascular disease
HR: hazard ratio

XSL•FO
RenderX

MMSE: Mini-Mental State Examination
OR: odds ratio

Edited by A Quialheiro; submitted 10.12.24; peer-reviewed by W Liu, Y Ou; comments to author 21.05.25; revised version received 12.07.25; accepted 01.08.25; published 08.09.25.

Please cite as:
Zhu Y, Wang K, Lu Z, Li F, Xu Y, Shen L, Bi Y, Hu W
Changes in Sarcopenia Status and Subsequent Cardiovascular Outcomes: Prospective Cohort Study
JMIR Aging 2025;8:e69860
URL: https://aging.jmir.org/2025/1/e69860
doi:10.2196/69860
PMID:

©Yuanyue Zhu, Kan Wang, Zuolin Lu, Feika Li, Yu Xu, Linhui Shen, Yufang Bi, Weiguo Hu. Originally published in JMIR Aging (https://aging.jmir.org), 08.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Dual Impact of Time and Content Exposure of Social Media on Diabetes Self-Management in Older Adults: Cross-Sectional Study

Qingyuan Ye[1], MSN; Ruiyang Xu[1], MSN; Li Li[1], PhD; Meng Zhao[1], PhD; Shan Wang[1], PhD; Sijing Peng[2], PhD; Si Chen[1], PhD; Fatema Ahmed[1], PhD; Chen Wu[1*], PhD; Kefang Wang[1*], PhD

[1]School of Nursing and Rehabilitation, Shandong University, Jinan, China

[2]School of Nursing, Anhui University of Chinese Medicine, Hefei, China

[*]these authors contributed equally

Corresponding Author:
Chen Wu, PhD
School of Nursing and Rehabilitation
Shandong University
NO.44 Wenhuaxi Road
Jinan, 250000
China
Phone: 86 17753573086
Email: wuchen@sdu.edu.cn

Abstract

Background: Self-management is critical for older adults with type 2 diabetes mellitus (T2DM); however, its practice remains suboptimal. Social media has become an accessible and effective stimulus source for the public, which has the potential to promote health behaviors, but its effect on the self-management of older adults with T2DM remains unknown.

Objective: We aimed to investigate the relationship between social media exposure, specifically time exposure and content exposure, and the self-management of older adults with T2DM.

Methods: In this cross-sectional study, we enrolled 257 older adults with T2DM who used short-form video apps from community health care centers. We assessed subjective and objective time and content exposure. We transformed text-based content exposure into diabetes-related content exposure encompassing irrelevant, harmful, hypobeneficial, and hyperbeneficial categories using Q-methodology. Self-management was assessed through a validated questionnaire. We used restricted cubic splines and linear regression models to model the relationships between time exposure and content exposure and self-management, respectively.

Results: Of 257 older adults with T2DM, the median age was 69 (IQR 65-72) years, 53.3% (n=137) were women, the mean sum score of self-management was 35.7 (SD 10.4), the median subjective time exposure was 120 (IQR 60-120) minutes, and 61.1% (n=157) of them were exposed to hyperbeneficial content. There was an approximate L-shaped dextrorotatory relationship between time exposure and self-management, with a decline in self-management when time exposure surpassed 139.8 minutes daily. Exposure to hyperbeneficial content was positively associated with the overall self-management (B=3.46, 95% CI 0.71-6.21). For participants exposed for more than 139.8 minutes daily, this positive association remained robust (B=7.27, 95% CI 1.54-13.00). In subdimensional analyses, hyperbeneficial content exposure was positively associated with general diet (B=1.51, 95% CI 0.54-2.49) and blood glucose testing (B=1.31, 95% CI 0.25-2.38).

Conclusions: Social media exposure presented a double-edged sword for self-management of older adults with T2DM. Self-management declined when the daily time spent on social media exceeded 139.8 minutes. However, exposure to hyperbeneficial content was associated with better self-management of individuals, regardless of excessive time spent on social media. Future longitudinal and experimental studies that validate the multifaceted association between social media exposure and health behaviors are needed. If confirmed, these findings would support the implementation of media prescription programs by health care providers in communities.

(JMIR Aging 2025;8:e67312) doi:10.2196/67312

XSL•FO
RenderX

KEYWORDS

social media; media exposure; type 2 diabetes mellitus; self-management; older adults

Introduction

Background

Type 2 diabetes mellitus (T2DM) imposes a substantial burden on health care systems worldwide [1], with older adults comprising over half of the adults diagnosed with T2DM worldwide [2]. Satisfactory self-management is crucial for older adults with T2DM, without which individuals are more likely to experience diabetes-associated complications, functional disability, and premature mortality [3]. Self-management is influenced by output information from external stimuli [4]. The lifestyle interventions are taken as the focal stimulus and function through providing reliable knowledge, regulated guidance, and interpersonal support theoretically [5], but face challenges of participants' low compliance and lack of perseverance in practice [6,7]. Therefore, the search for other optional and effective stimulus sources among older adults with T2DM is warranted.

Social media, which is abundant in disease-related information, can serve as a stimulus source for older adults with T2DM who frequently use these platforms to seek information about their condition [8]. Efforts from the government and businesses to bridge the digital divide have caused social media to explode in popularity among the older population. In the United States, 45% of individuals aged >65 years were active users of social media [9], while in China, older users make up 11.7% of all short-form video app users [10]. Instead of using text- or image-based platforms such as Twitter, most older adults prefer short-form video apps [11]. Social media is a powerful medium to promote health behaviors [12]. Supported by the cultivation theory [13] and empirical evidence, individuals' behavior may be influenced by the content they are exposed to on social media. Research demonstrated that exposure to COVID-19–related information on social media was associated with individuals' better practice of hand sanitation [14]. Ngqangashe and Backer [15] found that exposure to culinary videos on social media was associated with the food choice behaviors of middle school children. In addition, as an important dimension of social media exposure, time exposure on social media would also impact individuals' health behaviors. Zhang et al [16] found that women who reported excessive media use (>3 h/d) were more likely to have unhealthy lifestyles, including reduced physical activity, inadequate dietary diversity, and poorer sleep quality. Wu et al [17] found that longer-time exposure to short-form video was associated with decreased steps and time spent in physical activity in older women. However, multiple pitfalls in the literature have compromised the soundness of the existing findings. First, previous studies have reported significant but inconsistent associations between time and content exposure and health behaviors, highlighting that social media exposure has distinct characteristics that differentially influence self-management. Therefore, it is essential to examine both characteristics concurrently. Second, research often assumes a linear relationship between time exposure and the outcome, which has not been empirically confirmed. Third, the content that individuals are exposed to on social media is abundant, whereas researchers often collect specific content through yes or no questions. Such a practice fails to capture the full range of content that individuals are exposed to on social media [18]. Fourth, content exposure and time exposure have often been evaluated in studies with self-reported questions without considering the potential recall bias. Finally, the potential of social media to elicit self-management of older adults with T2DM has not been clarified.

Objectives

This study aimed to explore the role of social media as a source of stimuli in promoting self-management among older adults with T2DM and to investigate the associations between social media exposure (both content exposure and time exposure) on short-form video platforms and self-management of older adults with T2DM using both subjective and objective measures.

Methods

Ethical Considerations

The ethical oversight of this cross-sectional study was provided by the Shandong University (2023-R-017). All eligible adults provided permission to access their short-form video app use records, including time spent on each app, following, liking, bookmarking, and browsing history; agreed to provide deidentified individual information for research; and signed a written consent form before participating. This study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guidelines. A standardized monetary incentive (gift) was provided to encourage participation.

Recruitment

From February to June 2023, we launched a cross-sectional study named Media on Diabetes. A convenience sampling approach was used in this study, and upon the director's approval of the 5 community health centers, data were collected at the community health centers in Jinan, Shandong Province, China. Older people with T2DM were recruited during their routine physical examination visits. The inclusion criteria of the Media on Diabetes study were set as being aged ≥60 years, having been diagnosed with T2DM for at least 1 year, and using the short-form video apps Douyin or TikTok, Kuaishou, Haokan, and Toutiao on smartphones. Adults were excluded if they had been diagnosed with acute diabetes complications or physical disorders that could additionally affect their self-management, including cerebral ischemia, New York Health Association classes III and IV, otoliths, gout, severe asthma, chronic kidney disease at stage III or above, chronic obstructive pulmonary disease (Global Initiative for Chronic Obstructive Lung Disease III and IV), or moderate or severe swallowing disorders. We also excluded adults who were unable to walk independently even with walking aids, diagnosed with mental or cognitive disorders, diagnosed with hearing impairment and not wearing a hearing aid, or without diagnosed hearing impairment but

XSL•FO
RenderX

failing to hear when called aloud. The principal investigator and 2 trained research assistants collected data in the field. Adults filled in the questionnaires by themselves, and if they had inadequate reading capacity, the investigators assisted in reading questions. A total of 268 participants were recruited for this study, among whom 11 (4.1%) declined to complete the survey. Consequently, 257 valid questionnaires were included, yielding a response rate of 95.9%.

Sample Size

We used the formula to calculate the sample size for this study [19]. The formula is used for cross-sectional studies, where $Z_{\alpha/2}$ denotes the critical value for a 2-sided test at a 95% CI, *s* represents the SD derived from a previous study, and *d* represents the desired margin of error for estimating the population mean [19]. The mean (SD) of the sum score of self-management in the previous study was 37.5 (14.0) [20]; therefore, we took the values of $Z_{\alpha/2}$=1.96, *s*=14.0, and *d*=mean×5%=1.875. By adding 10% of adults to account for any withdrawal (eg, time conflicts, loss of interest, or privacy concerns), the minimal sample size for this study was calculated to be 235.

Measures

Social Media Exposure Assessment

Variables on the use of social media included social media exposure duration, exposure to multiplatforms, the feature of exposure, subjective time exposure, objective time exposure, and diabetes-related content exposure. Subjective time exposure was assessed through a self-reported question, and objective time exposure was recorded using data from the short-form video apps to account for potential recall bias. Participants were asked to indicate the types of content they watched from a list of 47 options (eg, health care, current affairs, and sport events) derived from a previous investigation on older adults [17], with space for additional comments. In addition, participants' interactions, such as liking, bookmarking, and browsing history on the short-form video apps, were gathered to supplement their self-reported data. Details of these variables are depicted in Table S1 in Multimedia Appendix 1 [21-24].

Self-Management Assessment

Diabetes self-management was assessed with the Chinese version of the Summary of Diabetes Self-care Activities, which was well adapted to Chinese culture [25]. It assessed 6 domains of self-management, including general diet, specific diet, exercise, blood glucose testing, foot care, and medication use of adults with T2DM. Each item was graded from 0 to 7, except for the item of high-fat food eating within the specific diet dimension, which was scored in reverse. The sum score of the Summary of Diabetes Self-care Activities ranged from 0 to 77, with higher scores demonstrating better self-management.

Covariates Assessment

By reviewing the literature on factors influencing diabetes self-management [26,27], covariates considered in the analyses included social demographic variables of sex, age, educational background, living status, marital status, working status, and monthly household income per capita; health-related variables of BMI, pain, diabetes duration, medication related to diabetes, complications of diabetes, multimorbidity, smoking history, and sleep duration; and psychological variables of self-efficacy, perceived stress, and social support. On the basis of the proposition of cultivation theory, which posits that traditional media exposure shapes health behaviors [13], we also considered variables of daily smartphone use time and traditional media use variables of self-reported traditional media exposure and health-related content exposure on traditional media. Coding and grading of the variables are shown in Tables S2 and S3 in Multimedia Appendix 1.

Statistical Analysis

The mean imputation approach was used to handle the missing data issue for the continuous variables of social media exposure duration (0.8%), self-efficacy (1.6%), perceived stress (1.2%), and social support (1.2%). Q-methodology is a robust approach that converts qualitative data into quantitative form and is widely applied in empirical studies [28,29]. In this study, we used Q-methodology [30] to assign scores to each social media content based on its relevance to T2DM and categorized participants according to their exposure to diabetes-related content. The procedures are outlined in Figure S1 in Multimedia Appendix 1. Participants who were exposed to irrelevant content only were coded as irrelevant content exposure, those exposed to harmful content with or without irrelevant content were coded as harmful content exposure, those exposed to beneficial content with a relevant score between 1 and 3 were coded as hypobeneficial content exposure, and those exposed to beneficial content with a relevant score of 4 or 5 were coded as hyperbeneficial content exposure.

The Kolmogorov-Smirnov goodness-of-fit test was used to assess the normality of continuous variables, which were described as mean and SD or median (IQR), as appropriate. Categorical variables were described as frequencies and percentages. The sum score of self-management was normally distributed, and as appropriate, the associations were calculated with Pearson or Spearman correlation for continuous variables and 2-sample *t* tests or ANOVA for categorical variables. Scores of self-management dimensions were not normally distributed, and we then used Spearman correlation, the Mann-Whitney *U* test, and the Kruskal-Wallis H test to test their associations with continuous and categorical variables. Consequently, covariates of the sum score of self-management and each score of self-management dimensions were selected a priori.

We used Spearman correlation to test the association between subjective time exposure and objective time exposure, and the correlation coefficient was 0.91 (P<.001), indicating that the recall bias was marginal. Therefore, we used subjective time exposure instead of objective time exposure for subsequent statistical inference due to its greater practical feasibility in community settings and comparability with the existing literature.

A scatter plot diagram was used to depict the relationship between time exposure and the sum score of self-management, and a nonlinear relationship was found. Therefore, the nonlinear relationship was estimated using a restricted cubic spline with

XSL•FO
RenderX

the estimator of ordinary least squares [31], while controlling for significant covariates. We repeated the steps to test the relationships between time exposure and each score of the self-management dimensions. A *P* value for an overall <.05 indicates the relationship between time exposure and self-management is statistically significant, and a *P* value for a nonlinear <.05 indicates a significant nonlinear relationship.

Multiple linear regression models were used to test the relationships between diabetes-related content exposure and the sum score of self-management; generalized linear models were used to evaluate the correlations between diabetes-related content exposure and scores of the self-management dimensions. If the nonlinear relationships between time exposure and self-management scores were found, we used the same modeling strategies to evaluate the associations between diabetes-related content exposure and self-management scores in subgroups based on time exposure, as appropriate.

All the analyses were performed with R (version 4.3.1; R Program for Statistical Computing), and 95% CIs for model estimates were reported. A 2-sided *P*<.05 was considered statistically significant. For diabetes-related content exposure as an independent variable, Bonferroni adjustment was applied due to multiple comparisons, and the threshold for statistical significance was adjusted to *P*<.017 (0.05/3=.017). Data were analyzed from July 2023 to May 2024.

Results

Participant Characteristics

Of the 257 participants, 137 (53.3%) were women, the median age was 69 (IQR 65-72) years, and the median diabetes duration was 7 (IQR 2-13) years (Table 1). The mean sum score of self-management of participants was 35.7 (SD 10.4). The median subjective time exposure of the participants was 120 (IQR 60-180) minutes. per day, the median social media exposure duration was 3 (IQR 1-5) years, and 173 (67.3%) were exposed to multiplatforms (Table 2).

XSL•FO
RenderX

Table 1. Participants' characteristics and their associations with self-management (N=257).

Characteristics	Participants	Statistics	*P* value[a]
Social demographic variables			
Sex, n (%)		1.239 (*df*=1)[b]	.22
Male	120 (46.7)		
Female	137 (53.3)		
Age (y), median (IQR)	69 (65-72)	–0.045[c]	.47
Educational background, n (%)		1.208 (*df*=3)[d]	.31
Elementary school or below	37 (14.4)		
Middle school	79 (30.7)		
High school	95 (37.0)		
College or above	46 (17.9)		
Living status, n (%)		0.397 (*df*=1)[b]	.69
Living alone	21 (8.2)		
Living with family, relatives, or friends	236 (91.8)		
Marital status, n (%)		–0.435 (*df*=1)[b]	.66
Married	232 (90.3)		
Unmarried, divorced, or widowed	25 (9.7)		
Working status, n (%)		–2.705 (*df*=1)[b]	.007[e]
Retired	234 (91.1)		
Still working	23 (8.9)		
Monthly household income per capita (CNY)[f], n (%)		2.473 (*df*=1)[d]	.09
<2000	30 (11.7)		
2000-4000	118 (45.9)		
>4000	109 (42.4)		
Health-related variables			
BMI (kg/m^2), n (%)		2.469 (*df*=2)[d]	.09
<24.0	103 (40.1)		
24.0-28.0	109 (42.4)		
≥28.0	45 (17.5)		
Pain, n (%)		2.939 (*df*=1)[b]	.004[e]
No	129 (50.2)		
Yes	128 (49.8)		
Diabetes duration (y), median (IQR)	7 (2-13)	0.258[c]	<.001[e]
Medication related to diabetes, n (%)		22.693 (*df*=2)[d]	<.001[e]
None	56 (21.8)		
Oral medicine	162 (63.0)		
Insulin	39 (15.1)		
Complications of diabetes, n (%)		0.158 (*df*=1)[b]	.88
No	187 (72.8)		
Yes	70 (27.2)		

XSL•FO
RenderX

Characteristics	Participants	Statistics	*P* value[a]
Multimorbidity, n (%)		–0.017 (*df*=1)[b]	.99
No	27 (10.5)		
Yes	230 (89.5)		
Smoking history, n (%)		2.043 (*df*=2)[d]	.13
Be smoking	31 (12.1)		
Quit smoking	19 (7.4)		
Nonsmoker	207 (80.5)		
Sleep duration, median (IQR), h	6 (6-7)	0.024[c]	.70
Psychological variables			
Self-efficacy, median (IQR)	29 (25-34)	0.249[c]	<.001[e]
Perceived stress, median (IQR)	15 (9-20.5)	–0.157[c]	.01[e]
Social support, mean (SD)	39.54 (7.16)	0.138[g]	.03[e]

[a]*P* values from univariate analyses for the sum score of self-management.

[b]Two-sample *t* tests were performed for binary categorical variables, including sex, living status, marital status, working status, pain, complications of diabetes, and multimorbidity.

[c]Spearman correlations were performed for continuous variables that do not conform to a normal distribution, including age, diabetes duration, sleep duration, self-efficacy, and perceived stress.

[d]ANOVA was performed for multicategorical variables, including educational background, monthly household income per capita, BMI, medication related to diabetes, and smoking history.

[e]Statistically significant (*P*<.05).

[f]1 CNY=US $0.14.

[g]Pearson correlation was performed for social support (a continuous variable that conforms to a normal distribution).

XSL•FO
RenderX

Table 2. Characteristics of media exposure and their associations with self-management in 257 older adults with type 2 diabetes mellitus.

Characteristics	Participants	Statistics	*P* value[a]
Traditional media			
Traditional media exposure, n (%)		0.980 (*df*=1)[b]	.33
No	75 (29.2)		
Yes	182 (70.8)		
Health-related content exposure on traditional media, n (%)		–1.210 (*df*=1)[b]	.23
No	220 (85.6)		
Yes	37 (14.4)		
Daily smartphone use (min), median (IQR)	120 (60-240)	0.041[c]	.57
Social media			
Social media exposure duration (y), median (IQR)	3 (1-5)	0.174[c]	.005[d]
Exposure to multiplatforms, n (%)		–0.135 (*df*=1)[b]	.89
No	84 (32.7)		
Yes	173 (67.3)		
The feature of exposure, n (%)		–0.906 (*df*=1)[b]	.37
Video-based	113 (44.0)		
Video- and text-based	144 (56.0)		

[a]*P* values from univariate analyses for the sum score of self-management.

[b]Two-sample *t* tests were performed for binary categorical variables, including traditional media exposure, health-related content exposure on traditional media, exposure to multiplatforms, and the feature of exposure.

[c]Spearman correlations were performed for continuous variables that do not conform to a normal distribution, including daily smartphone time use and social media exposure duration.

[d]Statistically significant (*P*<.05).

Of the diabetes-related content exposure variable, hyperbeneficial content comprises 5 core diabetes self-management components: dietary control, physical activity, medication adherence, blood glucose monitoring, and foot care. In contrast, hypobeneficial content includes information about diabetes-related comorbidities, for example, hypertension, dyslipidemia, and coronary heart disease. While clinically relevant, such content offers generalized health benefits rather than diabetes-specific advantages. The distribution of participants across the different categories of diabetes-related content exposure is presented in Table S4 in Multimedia Appendix 1. There were 64 (24.9%) participants who fell into the category of irrelevant content exposure, 1 (0.4%) who fell into the category of harmful content exposure, 35 (13.6%) who fell into the category of hypobeneficial content exposure, and 157 (61.1%) who fell into the category of hyperbeneficial content exposure.

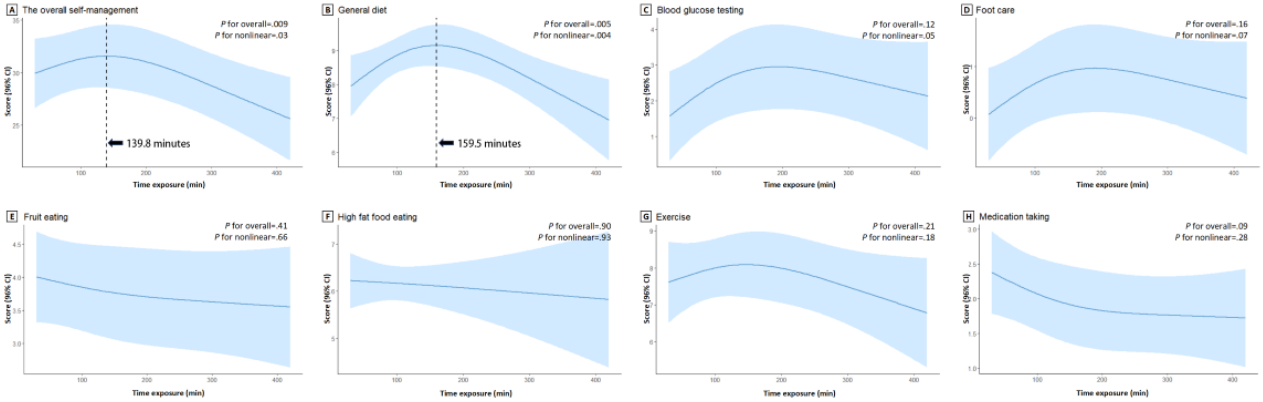
For the univariate analyses presented in Table 1, significant associations with the sum score of self-management were identified for working status (t_1=–2.705; *P*=.007), pain (t_1=2.939; *P*=.004), diabetes duration (*r*=0.258, *P*<.001), medication related to diabetes (F_2=22.693, *P*<.001), self-efficacy (*r*=0.249, *P*<.001), perceived stress (*r*=–0.157; *P*=.01), social support (*r*=0.138; *P*=.03), and social media exposure duration (*r*=–0.174; *P*=.005). The univariate analyses between covariates and each score of self-management dimensions are presented in Table S5 in Multimedia Appendix 1.

Associations Between Time Exposure and Self-Management

As shown in Figure 1, time exposure and the sum score of self-management followed an approximate L-shaped dextrorotatory trend (*P* for nonlinear=.03). The trend started with a marginal and almost linear increase between 10 minutes and 139.8 minutes, followed by a sharp and almost linear decline until 600 minutes. The sample was then divided into the ≤139.8 minutes group and the >139.8 minutes group for further analyses. The nonlinear relationship between time exposure and the general diet score was statistically significant (*P* for nonlinear=.004).

XSL•FO
RenderX

Figure 1. Associations between time exposure and the self-management dimensions. (A) The overall self-management adjustments include working status, diabetes duration, pain, medication related to diabetes, self-efficacy, perceived stress, social support, and social media exposure duration. (B) General diet adjustments include marital status, health-related content exposure to traditional media, and self-efficacy. (C) Blood glucose testing adjustments include diabetes duration, medication related to diabetes, and self-efficacy. (D) Foot care adjustments include diabetes duration, medication related to diabetes, complications of diabetes, and self-efficacy. (E) Fruit-eating adjustments include educational background, working status, monthly household income per capita, exposure to multiplatforms, and social support. (F) High-fat food eating adjustments include daily smartphone use time and traditional media exposure. (G) Exercise adjustments include sex, pain, complications of diabetes, self-efficacy, social support, perceived stress, and social media exposure duration. (H) Medication-taking adjustments include sex, diabetes duration, medication related to diabetes, and complications of diabetes.



Associations Between Diabetes-Related Content Exposure and Self-Management

One participant who fell into the category of harmful content exposure was not representative and was therefore excluded from the subsequent analyses. In the full sample, compared with irrelevant content exposure, hyperbeneficial content exposure was associated with the higher sum score of self-management (B=3.46, 95% CI 0.71 to 6.21); while the association between hypobeneficial content exposure and the sum score of self-management was positive but not statistically significant (B=0.17, 95% CI –3.54 to 3.88; Table 3). In subdimensional analyses, hyperbeneficial content exposure was associated with the higher score of general diet (B=1.51, 95% CI 0.54 to 2.49) and blood glucose testing (B=1.31, 95% CI 0.25 to 2.38). The association between hyperbeneficial content exposure and the foot care score was not statistically significant (B=0.75, 95% CI –0.10 to 1.52; Table S6 in Multimedia Appendix 1).

Table 3. Associations between diabetes-related content exposure and the sum score of self-management[a].

Diabetes-related content exposure (reference irrelevant)	B[b] (SE; 95% CI)	*P* value
Full sample (N=256[c]),		
Hypobeneficial	0.49 (1.91; –3.26 to 4.25)	.80
Hyperbeneficial	3.46 (1.40; 0.71 to 6.21)	.01[d]
The ≤139.8 min group (n=147)		
Hypobeneficial	–0.62 (2.26; –5.09 to 3.85)	.77
Hyperbeneficial	2.22 (1.73; –1.20 to 5.64)	.18
The >139.8 min group (n=109)		
Hypobeneficial	5.38 (3.76; –2.10 to 12.85)	.16
Hyperbeneficial	7.27 (2.89; 1.54 to 13.00)	.01[d]

[a]Covariates adjusted include working status, diabetes duration, pain, medication related to diabetes, self-efficacy, perceived stress, social support, and social media exposure duration.

[b]B: unstandardized regression coefficient.

[c]One case with harmful exposure was excluded.

[d]*P* values less than the Bonferroni-corrected level of significance of <.017.

In subgroup analyses, among participants with excessive time spent on social media, compared with irrelevant content exposure, the association between hyperbeneficial content exposure and the sum score of self-management remained robust (B=7.27, 95% CI 1.54-13.00; Table 3); while hyperbeneficial content exposure was not significantly associated with the score of general diet at the Bonferroni-adjusted level (B=2.38, 95% CI 0.36-4.40; Table S7 in Multimedia Appendix 1).

XSL•FO
RenderX

Discussion

Principal Findings

In this study, we reported that time exposure and content exposure on social media were associated with self-management in older adults with T2DM. Specifically, spending less than 139.8 minutes on social media per day was beneficial for self-management, but beyond that, increased time exposure would decrease self-management scores. Exposure to hyperbeneficial content was positively associated with overall self-management in both the full sample and among those using social media for more than 139.8 minutes per day.

Previous studies relied solely on subjective assessments [32,33]. It is often assumed that older adults are more sensitive to time, yet this assumption is not empirically confirmed. In this study, we did not take this assumption as a given and instead assessed both objective and subjective time simultaneously. Through comparison, we found these 2 measures to be highly correlated. In addition, considering that subjective time offers greater practical feasibility in community settings and better comparability with existing literature, we chose to use subjective time throughout our analyses. By identifying the strong associations between these 2 variables, we determined to use subjective time exposure in this study because it is feasible to assess. Instead of following the traditional paradigm, that is, assuming the linear relationships between time exposure and health behaviors [34], we used the restricted cubic spline and found an approximate L-shaped dextrorotatory trend between time exposure and the overall self-management. This suggests that excessive exposure to social media may compromise the self-management of adults with T2DM. We also found that excessive time exposure would significantly associate with a decline in participants' adherence to a diabetes diet. Understanding the mechanism underlying these findings falls outside the scope of this study, but some information in the literature would facilitate unraveling the hidden story in the black box. For example, excessive exposure to social media would reflect individuals' suboptimal executive function [35], and the latter would result in unsatisfactory diabetes self-management [36]. As for the general diet, social media is a popular platform for unhealthy food advertising, and excessive exposure might increase the likelihood of individuals becoming its audience and then being driven away from healthy diets [37,38]. Future studies may corroborate these findings and advance science by adding robust evidence on the mechanisms between time spent on social media and self-management of older adults with T2DM.

By sorting out content exposure gathered through a semiclosed question and records documented on social media, we found that 61.1% (157/257) had hyperbeneficial content exposure, including information about diabetes diet, complications, and treatment. With the high-resolution data, we found that compared with exposure to irrelevant content, participants exposed to hyperbeneficial content exhibited significantly better performance on overall self-management and subdimensions of general diet and blood glucose testing; the positive effects of hyperbeneficial content exposure on self-management remained robust among those with excessive time exposure. Cultivation theory posits that media exposure would reshape individuals' awareness and behaviors through information dissemination [13]. Consistent with this proposition, we found that exposure to hyperbeneficial content was significantly associated with better overall self-management, which validates the role of social media in shaping health behaviors. Furthermore, this study extends the proposition of cultivation theory by revealing the nonsignificant effects of hypobeneficial content and the negative effect of prolonged time exposure. If high-quality cohort studies or interventions confirm the findings from our study, this evidence may support media prescription in community settings. For example, health care providers may implement content dissemination programs through partnerships with social media platforms or curated content libraries for patients. It may also help social media servers optimize user profile capture, for example, by expanding information collection to include illness information and enhancing the recommender systems.

Subdimensional analyses revealed a positive trend between hyperbeneficial content exposure and foot care outcome, although this result was not statistically significant (P=.05). Foot care is often regarded as one of the most challenging practices encountered by health care providers [39,40], and this emerging trend nonetheless merits further investigation. Our results showed that only 2% (5/257) of the participants reported exposure to foot care content on social media. Empirical evidence showed that foot care information on social media was limited, but the available content was useful and highly reliable [41]. We failed to capture the association between hyperbeneficial content exposure and the practice of foot care, but this may be attributed to the small sample of individuals with foot care information exposure in our study. Gaps between diabetic foot care and other dimensions of self-management should be addressed, and we hope that health care providers may take their roles in posting high-quality foot care information on social media, getting familiar with and updating the information, and recommending it to older adults with T2DM.

Limitations

Although the findings of this study are strengthened by addressing the methodological pitfalls and content gaps in the literature, some limitations should be considered. First, several novel hypotheses were proposed regarding the relationships between social media exposure and self-management among older adults with T2DM. However, this study is cross-sectional, and future longitudinal or experimental research is required to determine the causality of these relationships. Second, most of the adults in our study had higher educational backgrounds, better economic status, and lived in urban areas. While individuals with these characteristics are the primary users of social media [42], this demographic skew may limit the extrapolation of our findings to populations with different characteristics, such as those living in rural areas and individuals with lower socioeconomic status. Third, we operationalized social media exposure as time exposure and content exposure, as often assessed in the literature. However, social media exposure may include other parameters, such as online social engagement, which reflects the *social* feature of social media

XSL•FO
RenderX

exposure. Future research could measure variables such as chatting frequency on chatting, posting, and gaming, and investigate their influences on older adults' self-management. A concept analysis of social media exposure should be conducted to lay the methodological foundation for research in this area.

Conclusions

This study reveals the dual effects of social media exposure on the self-management of older adults with T2DM. Self-management was aggravated when daily social media use time exceeded 139.8 minutes. However, exposure to hyperbeneficial content facilitated self-management, even under excessive time exposure. These findings provide new evidence suggesting that social media has the potential to serve as a valuable source for enhancing the self-management of older adults with T2DM. Further high-quality cohort studies are needed to confirm these findings, which may contribute to the development of strategies to guide appropriate social media exposure in older adults with T2DM to enhance their self-management.

Acknowledgments

The authors express their gratitude to the participants involved in data collection. RX was affiliated with the School of Nursing and Rehabilitation at Shandong University at the time of the investigation and is currently affiliated with the Hangzhou Zhongce Vocational School Qiantang. This study was supported by the National Natural Science Foundation of China (grant 82504534).

Authors' Contributions

Conceptualization: RX (lead), QY (equal), CW (equal), KW (supporting)
Data curation: RX (lead), QY (equal), LL (equal)
Formal analysis: RX (lead), QY (equal)
Funding acquisition: CW (lead)
Investigation: RX (lead), SW (supporting), SP (supporting)
Methodology: RX (lead), QY (equal), LL (equal), CW (supporting)
Project administration: CW (lead)
Supervision: KW (lead)
Validation: QY (lead)
Visualization: QY (lead)
Writing—original draft: RX (lead), QY (equal), CW (supporting)
Writing—review and editing: QY (lead), LL (equal), CW (equal), MZ (supporting), SW (supporting), SP (supporting), SC (supporting), FA (supporting), KW (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplemental tables and figures.
[DOCX File , 600 KB - aging_v8i1e67312_app1.docx]

References

1. IDF diabetes atlas 2021. International Diabetes Federation. URL: https://diabetesatlas.org/atlas/tenth-edition/ [accessed 2024-07-20]
2. National Diabetes Audit (NDA) 2023-24 quarterly report for England, Integrated Care Board (ICB), Primary Care Network (PCN) and GP practice. NHS England Digital. URL: https://tinyurl.com/45wju2k4 [accessed 2024-07-20]
3. Saelee R, Bullard KM, Hora IA, Pavkov ME, Pasquel FJ, Holliday C, et al. Trends and inequalities in diabetes-related complications among U.S. adults, 2000-2020. Diabetes Care 2025 Jan 01;48(1):18-28 [FREE Full text] [doi: 10.2337/dci24-0022] [Medline: 38905540]
4. Kempermann G. Environmental enrichment, new neurons and the neurobiology of individuality. Nat Rev Neurosci 2019 Apr 5;20(4):235-245. [doi: 10.1038/s41583-019-0120-x] [Medline: 30723309]
5. Ranjbar F, Karimi M, Zare E, Ghahremani L. The effect of educational intervention based on the behavioral reasoning theory on self-management behaviors in type 2 diabetes patients: a randomized controlled trial. BMC Public Health 2024 Jul 02;24(1):1761 [FREE Full text] [doi: 10.1186/s12889-024-19207-0] [Medline: 38956554]
6. Paudel G, Vandelanotte C, Dahal PK, Biswas T, Yadav UN, Sugishita T, et al. Self-care behaviours among people with type 2 diabetes mellitus in South Asia: a systematic review and meta-analysis. J Glob Health 2022 Aug 03;12:04056 [FREE Full text] [doi: 10.7189/jogh.12.04056] [Medline: 35916498]

XSL•FO
RenderX

7. Qian Y, He Y, Zhu W, Chen Q. Relationship between health beliefs, chronic illness resource utilization and self-management behavior in elderly diabetic patients. Mil Nurs 2022;39(12):58-61. [doi: 10.3969/j.issn.2097-1826.2022.12.015]
8. Wu C, Xu R, Cao J, Wang S, Peng S, Wang C, et al. Barriers and facilitators of self-management for older adults living with type 2 diabetes mellitus: a qualitative study in China. Sci Diabetes Self Manag Care 2024 Feb 19;50(1):44-55. [doi: 10.1177/26350106231221454] [Medline: 38240230]
9. Social network usage and growth statistics: how many people use social media in 2024? Backlinko. URL: https://backlinko.com/social-media-users [accessed 2024-04-15]
10. Zhang T, Wang L. China short-form video users value research report 2022. China Media Industry Development Report. 2023. URL: https://www.pishu.com.cn/skwx_ps/initDatabaseDetail?siteId=14&contentId=14647611&contentType=literature [accessed 2024-07-20]
11. The 49th statistical report on China's internet development. China Internet Network Information Center. URL: https://www3.cnnic.cn/n4/2022/0401/c88-1131.html [accessed 2024-07-14]
12. Global action plan on physical activity 2018-2030: more active people for a healthier world. World Health Organization. URL: https://www.who.int/publications/i/item/9789241514187 [accessed 2024-07-23]
13. Hermann E, Morgan M, Shanahan J. Cultivation and social media: a meta-analysis. New Media Soc 2023 Jul 08;25(9):2492-2511. [doi: 10.1177/14614448231180257]
14. Pan Y, Xin M, Zhang C, Dong W, Fang Y, Wu W, et al. Associations of mental health and personal preventive measure compliance with exposure to COVID-19 information during work resumption following the COVID-19 outbreak in China: cross-sectional survey study. J Med Internet Res 2020 Oct 08;22(10):e22596 [FREE Full text] [doi: 10.2196/22596] [Medline: 32936776]
15. Ngqangashe Y, Backer CJ. The differential effects of viewing short-form online culinary videos of fruits and vegetables versus sweet snacks on adolescents' appetites. Appetite 2021 Nov 01;166:105436. [doi: 10.1016/j.appet.2021.105436] [Medline: 34119561]
16. Zhang X, Liu J, Han N, Yin J. Social media use, unhealthy lifestyles, and the risk of miscarriage among pregnant women during the COVID-19 pandemic: prospective observational study. JMIR Public Health Surveill 2021 Jan 05;7(1):e25241 [FREE Full text] [doi: 10.2196/25241] [Medline: 33293263]
17. Wu C, Chen S, Wang S, Peng S, Cao J. Short-form video exposure and its two-sided effect on the physical activity of older community women in China: secondary data analysis. JMIR Mhealth Uhealth 2023 Sep 13;11:e45091 [FREE Full text] [doi: 10.2196/45091] [Medline: 37707321]
18. Griffioen N, Van Rooij MM, Lichtwarck-Aschoff A, Granic I. A stimulated recall method for the improved assessment of quantity and quality of social media use. J Med Internet Res 2020 Jan 28;22(1):e15529 [FREE Full text] [doi: 10.2196/15529] [Medline: 32012075]
19. Serdar CC, Cihan M, Yücel D, Serdar MA. Sample size, power and effect size revisited: simplified and practical approaches in pre-clinical, clinical and laboratory studies. Biochem Med (Online) 2021 Feb 16;31(1):27-53. [doi: 10.11613/bm.2021.010502]
20. Gimbel RW, Rennert LM, Crawford P, Little JR, Truong K, Williams JE, et al. Enhancing patient activation and self-management activities in patients with type 2 diabetes using the US Department of Defense Mobile Health Care Environment: feasibility study. J Med Internet Res 2020 May 26;22(5):e17968 [FREE Full text] [doi: 10.2196/17968] [Medline: 32329438]
21. Zhang JX, Schwarzer R. Measuring optimistic self-beliefs: a Chinese adaptation of the general self-efficacy scale. Psychol Int J Psychol Orient 1995;38(3):174-181 [FREE Full text]
22. Ma M, Ai Z, Shi Z. Reliability and validity of General Self-Efficacy Scale (Chinese version) in middle-aged and elderly patients with type 2 diabetes. J Tongji Univ Sci 2022;43(4):515-520. [doi: 10.12289/j.issn.1008-0392.21411]
23. Yang TZ, Huang HT. [An epidemiological study on stress among urban residents in social transition period]. Zhonghua Liu Xing Bing Xue Za Zhi 2003 Sep;24(9):760-764. [Medline: 14521764]
24. Xiao S. Theoretical basis and research applications of the Social Support Rating Scale (SSRS). J Clin Psychiatry 1994;4(2):98-100 [FREE Full text]
25. Qiaoqing W, Shaomei S, Xiaobin L, Jie P. Study on the reliability and validity of summary of diabetes self-care activities for type 2 diabetes patients. Chin J Pract Nurs 2008;24(07):26-27. [doi: 10.3760/cma.j.issn.1672-7088.2008.07.009]
26. Yang N, Masingboon K, Samartkit N. Factors influencing diabetes self-management among adults with type 2 diabetes mellitus in China. Belitung Nurs J 2022 Oct 21;8(5):389-395 [FREE Full text] [doi: 10.33546/bnj.2199] [Medline: 37554488]
27. Alexandre K, Campbell J, Bugnon M, Henry C, Schaub C, Serex M, et al. Factors influencing diabetes self-management in adults: an umbrella review of systematic reviews. JBI Evid Synth 2021 May;19(5):1003-1118. [doi: 10.11124/JBIES-20-00020] [Medline: 33741836]
28. Metla SG, Akhtar-Danesh N, Saini J, Bayer I, Mezil Y, Brewer-Deluce D, et al. Anatomy beyond the pandemic: a Q-methodology study exploring student perceptions toward a hybrid curriculum. Anat Sci Educ 2024 Jun 02;17(4):831-843. [doi: 10.1002/ase.2415] [Medline: 38563458]
29. Yim E, Yun M, Sok S. Subjectivity study on health conservation of elderly hemodialysis patients. BMC Geriatr 2024 Feb 27;24(1):196 [FREE Full text] [doi: 10.1186/s12877-024-04819-3] [Medline: 38413859]

XSL•FO
RenderX

30. Churruca K, Ludlow K, Wu W, Gibbons K, Nguyen HM, Ellis LA, et al. A scoping review of Q-methodology in healthcare research. BMC Med Res Methodol 2021 Jun 21;21(1):125 [FREE Full text] [doi: 10.1186/s12874-021-01309-7] [Medline: 34154566]
31. Marrie RA, Dawson NV, Garland A. Quantile regression and restricted cubic splines are useful for exploring relationships between continuous variables. J Clin Epidemiol 2009 May;62(5):511-7.e1. [doi: 10.1016/j.jclinepi.2008.05.015] [Medline: 19135859]
32. Niu Z, Willoughby J, Zhou R. Associations of health literacy, social media use, and self-efficacy with health information-seeking intentions among social media users in China: cross-sectional survey. J Med Internet Res 2021 Feb 25;23(2):e19134 [FREE Full text] [doi: 10.2196/19134] [Medline: 33629955]
33. Schuster AM, Cotten SR, Meshi D. Established adults, who self-identify as smartphone and/or social media overusers, struggle to balance smartphone use for personal and work purposes. J Adult Dev 2023 Sep 08;30(1):78-89 [FREE Full text] [doi: 10.1007/s10804-022-09426-3] [Medline: 36101741]
34. Al-Shoaibi AA, Shao IY, Ganson KT, Lavender JM, Testa A, Kiss O, et al. Prospective association of screen time with binge-eating disorder among adolescents in the United States: the mediating role of depression. Int J Eat Disord 2024 May 15;57(5):1192-1201 [FREE Full text] [doi: 10.1002/eat.24169] [Medline: 38358046]
35. Su C, Zhou H, Gong L, Teng B, Geng F, Hu Y. Viewing personalized video clips recommended by TikTok activates default mode network and ventral tegmental area. Neuroimage 2021 Aug 15;237:118136 [FREE Full text] [doi: 10.1016/j.neuroimage.2021.118136] [Medline: 33951514]
36. Ho HT, Lin SI, Guo NW, Yang YC, Lin MH, Wang CS. Executive function predict the quality of life and negative emotion in older adults with diabetes: a longitudinal study. Prim Care Diabetes 2022 Aug;16(4):537-542. [doi: 10.1016/j.pcd.2022.05.002] [Medline: 35659729]
37. Ali-Alsaadi AA, Cabeza-Ramírez LJ, Sántos-Roldán L, Loor-Zambrano HY. Digital marketing and fast-food intake in the UAE: the role of firm-generated content among adult consumers. Foods 2023 Nov 10;12(22):4089 [FREE Full text] [doi: 10.3390/foods12224089] [Medline: 38002147]
38. Mc Carthy CM, de Vries R, Mackenbach JD. The influence of unhealthy food and beverage marketing through social media and advergaming on diet-related outcomes in children-a systematic review. Obes Rev 2022 Jun 17;23(6):e13441 [FREE Full text] [doi: 10.1111/obr.13441] [Medline: 35301815]
39. Fitridge R, Chuter V, Mills J, Hinchliffe R, Azuma N, Behrendt C, et al. The intersocietal IWGDF, ESVS, SVS guidelines on peripheral artery disease in people with diabetes and a foot ulcer. Diabetes Metab Res Rev 2024 Mar 19;40(3):e3686 [FREE Full text] [doi: 10.1002/dmrr.3686] [Medline: 37726988]
40. Armstrong DG, Boulton AJ, Bus SA. Diabetic foot ulcers and their recurrence. N Engl J Med 2017 Jun 15;376(24):2367-2375. [doi: 10.1056/NEJMra1615439] [Medline: 28614678]
41. Akça Doğan D, Pehlivan S. Diabetic foot care training and the presence of nurses in Turkish YouTube videos. Prim Care Diabetes 2022 Jun;16(3):430-434. [doi: 10.1016/j.pcd.2022.03.010] [Medline: 35346614]
42. Kemp S. Digital 2025: global overview report. Digital Portal. URL: https://datareportal.com/reports/digital-2025-global-overview-report [accessed 2025-07-19]

Abbreviations

STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
T2DM: type 2 diabetes mellitus

Edited by Y Jiang; submitted 08.10.24; peer-reviewed by I Milani, M Peeples, H Wang, XM Yang; comments to author 14.07.25; revised version received 20.07.25; accepted 22.08.25; published 18.09.25.

Please cite as:
Ye Q, Xu R, Li L, Zhao M, Wang S, Peng S, Chen S, Ahmed F, Wu C, Wang K
The Dual Impact of Time and Content Exposure of Social Media on Diabetes Self-Management in Older Adults: Cross-Sectional Study
JMIR Aging 2025;8:e67312
URL: https://aging.jmir.org/2025/1/e67312
doi:10.2196/67312
PMID:

©Qingyuan Ye, Ruiyang Xu, Li Li, Meng Zhao, Shan Wang, Sijing Peng, Si Chen, Fatema Ahmed, Chen Wu, Kefang Wang. Originally published in JMIR Aging (https://aging.jmir.org), 18.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited.

XSL•FO
RenderX

The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Impact of a Consumer e-Learning Course on Beliefs, Treatment Choices, and Outcomes Among People With Hip and Knee Osteoarthritis: Qualitative Interview Study

Rachel K Nelligan[1], PhD; Libby Spiers[1], BPhysio (Hons); Rana S Hinman[1], PhD; Kim L Bennell[1], PhD

Centre for Health, Exercise and Sports Medicine, Department of Physiotherapy, School of Health Sciences, The University of Melbourne, Parkville, Australia

Corresponding Author:
Rachel K Nelligan, PhD
Centre for Health, Exercise and Sports Medicine, Department of Physiotherapy, School of Health Sciences, The University of Melbourne
161 Barry Street
Parkville, 3010
Australia
Phone: 61 3 8344 90
Email: rachel.nelligan@unimelb.edu.au

Abstract

Background: First-line management for hip and knee osteoarthritis includes lifestyle treatments, such as exercise and weight loss (if appropriate), whereas joint replacement surgery is recommended only for severe symptoms after these options have been exhausted. However, many people with osteoarthritis hold misconceptions about the condition, leading to lower acceptance of nonsurgical treatments, such as exercise, and the mistaken belief that surgery is their only option. Novel patient education approaches that address these misconceptions are recommended to improve uptake of lifestyle treatments, reduce unnecessary surgery, and improve outcomes for people with osteoarthritis. We developed a 4-week self-directed consumer e-learning course on osteoarthritis management. In a randomized controlled trial, using the course led to immediate and sustained improvements in osteoarthritis knowledge. However, participants' perspectives on the course and an understanding of how it impacted osteoarthritis beliefs, treatment choices, and outcomes were unknown.

Objective: This study aims to explore how an e-learning course for people with hip and knee osteoarthritis may have impacted their osteoarthritis beliefs, treatment choices, and outcomes.

Methods: In this qualitative study, we conducted semistructured individual interviews (N=20) with randomized controlled trial participants with hip or knee osteoarthritis who accessed a 4-week consumer e-learning course on osteoarthritis and its management. Interviews were audio recorded, transcribed verbatim, and thematically analyzed following a framework approach, which was guided by the common sense model of self-regulation.

Results: Four themes were developed from the interviews: (1) participants reshaped their beliefs and attitudes toward osteoarthritis and its management, (2) participants adopted a proactive approach to management, (3) participants developed a more positive mindset, and (4) the course supported learning and shifts in beliefs.

Conclusions: The e-learning course resulted in shifts in participants' beliefs and attitudes toward osteoarthritis and its management, increasing their confidence in living with osteoarthritis and resulting in a more optimistic outlook on the future. The e-learning course is freely available and could be a useful resource for people with osteoarthritis to enhance their understanding of the condition and its management.

(JMIR Aging 2025;8:e80282) doi:10.2196/80282

KEYWORDS

osteoarthritis; knee; hip; education; qualitative; self-management; e-learning

XSL•FO
RenderX

Introduction

Background

Osteoarthritis commonly affects the knee and hip joints, causing persistent pain, impaired function, and reduced quality of life [1]. First-line management, recommended for all people with osteoarthritis, includes lifestyle approaches, such as exercise and weight loss (if overweight) [2-5]. Despite joint replacement surgery being recommended only for those with severe symptoms who have exhausted these first-line approaches, many people believe that joint replacement surgery is the only option to manage their symptoms [6,7].

The common sense model of self-regulation has been used to better understand the influence of beliefs about osteoarthritis on the treatment choices and outcomes of people with hip and knee osteoarthritis [6-8]. According to this model, a person's beliefs about their health condition, including how they label it; their perceptions of its duration, causes, and consequences; and their belief in their ability to control it, can significantly influence treatment decisions and subsequently affect outcomes [9]. For instance, qualitative studies have shown that people with osteoarthritis commonly believe that their condition is caused by "wear and tear," resulting in "bone on bone." These misconceptions are associated with lower acceptance of recommended nonsurgical lifestyle treatments, such as exercise, and an inaccurate belief that surgical management is the only viable option [6-8,10]. As a result, novel consumer education methods that address such misconceptions have been identified as key strategies to increase the use of recommended lifestyle treatments and reduce unnecessary surgery [11]. However, the effectiveness of self-directed consumer education interventions in achieving these goals remains unknown; to the best of our knowledge, no qualitative studies have explored the impacts of a digital consumer education intervention for osteoarthritis.

This Study

To address this gap, our research center developed a 4-week self-directed consumer e-learning course about osteoarthritis and its recommended management [12]. We subsequently evaluated this course in a randomized controlled trial (RCT) of 124 people with hip or knee osteoarthritis and found that using the course led to immediate and sustained improvements in osteoarthritis knowledge when compared to a standard osteoarthritis education intervention, that is, provision of an online osteoarthritis pamphlet available from a reputable consumer organization [13]. Guided by the common sense model, the aim of this study was to gain a richer understanding of the perspectives of people who used the e-learning course within the RCT and explore how it may have impacted their osteoarthritis beliefs, treatment choices, and outcomes.

Methods

Ethical Considerations

The institutional human research ethics committee approved this study (responsible committee: STEMM 2; ethics ID 26676). All participants provided informed consent via an online form within REDCap (Research Electronic Data Capture; Vanderbilt University) software hosted at the University of Melbourne [14,15]. Participants received an Aus $50 (US $32.40) gift voucher as a token of appreciation upon completion of their interview.

Study Design

This study was approached through an interpretivist paradigm, focusing on the subjective perspectives of people within their specific context [16]. It was nested within an RCT that evaluated the effectiveness of a consumer e-learning course on osteoarthritis knowledge and pain self-efficacy in people with hip or knee osteoarthritis; the trial was registered on the Australian New Zealand Clinical Trials Registry (ACTRN12622001490763). Reporting complied with the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines [17].

Participants and Recruitment

Participants in this study were a subsample of those who participated in the RCT, were allocated to the e-learning group, and had completed their involvement in the trial (ie, completed 13-week assessments) within the past 4 weeks. Participants were purposively sampled to ensure variation across gender, age, duration of symptoms, and perceived usefulness of the course (as rated by an outcome measure collected as part of the RCT's 5-week assessment), with sampling monitored to include participants across the spectrum. The sample size was dictated by theoretical saturation [18]. This was monitored during the coding process by 2 independent researchers, with recruitment ceasing once both agreed there was mounting evidence that new data were providing minimal new insights to address the research question.

Initial recruitment for the RCT was conducted nationwide in Australia via internet sources (eg, social media and online newspapers) and our center's volunteer database. To be included in the RCT, participants needed to live in Australia and have (1) a native (unreplaced) hip or knee joint meeting osteoarthritis clinical criteria (aged ≥45 years, activity-related pain, and morning joint stiffness ≤30 min) [5], (2) pain in the joint for 3 months or more and on most days in the past month, (3) access to a computer with internet and email, and (4) the ability to provide consent and complete assessments. Participants were excluded from the RCT if they self-reported systemic arthritis (eg, rheumatoid arthritis or gout), were scheduled for lower limb joint surgery in the next 13 weeks, had completed an online education course about osteoarthritis that involved at least 2 hours of learning in the past 12 months, or were unable to easily read and understand English.

e-Learning Course Intervention

The course is accessible for free on the online education platform FutureLearn (Global University Systems) [19]. During the RCT, the course was available via FutureLearn's standard tiered access options. This included a free "'limited access" option, where each of the 4 modules was progressively unlocked weekly, and access ceased 4 weeks after registration, with a "buy this course" option, providing immediate unrestricted access. After randomization, participants in the RCT's e-learning group were sent an email containing the course link, instructions

XSL•FO
RenderX

on how to register for the free "limited access" option, and the recommendation to complete module 1 within 7 days, followed by subsequent modules weekly.

The full description of the e-learning course content and its development has been published elsewhere [12]. In summary, its development was informed by behavior and learning theory, osteoarthritis clinical practice guidelines, and consumer consultation. The course includes four modules: (1) learning about osteoarthritis, (2) physical activity and exercise for osteoarthritis, (3) body weight and osteoarthritis, and (4) additional management strategies and conclusion. Content is presented using a variety of formats, including written text, videos, infographics, learning activities (eg, quizzes and polls), and nonmoderated discussion boards. Each module contains a variety of recommended resources, including external links (eg, to evidence-informed self-directed online exercise programs, podcasts, and dietary guidelines), downloadable resources (eg, physical activity or exercise logbooks and action plan templates), and "finding out more" information (eg, information about the range of health care professionals who can provide further help with each topic and references with links to scientific papers that support content). The time required to complete all 4 modules and learning activities is approximately 4 hours (1 h per module). However, this is highly dependent on the user's level of interaction with the content and recommended resources.

Data Collection

Participants' demographic data were extracted from data collected within the RCT. Semistructured, one-on-one telephone interviews were conducted by one researcher (LS) who had no involvement in the development of the e-learning course or its RCT evaluation. Interviews followed an interview guide (Textbox 1), which was developed by members of the research team (RKN, KLB, and RSH) and aimed to explore the potential impact of the e-learning on osteoarthritis beliefs, treatment choices, and outcomes as guided by the common sense model of self-regulation. Interviews were audio recorded and transcribed verbatim by an external transcription service. Deidentified audio files and transcripts were stored in secure password-protected computer files on the university server, and only deidentified data were used for analysis.

XSL•FO
RenderX

Textbox 1. Semistructured interview guide.

Introduction

As part of your involvement in the Osteoarthritis Education study, you were asked to access the online course, Taking Control of your Osteoarthritis. This interview aims to get your feedback on the course and help us understand how the course might have influenced how you think about your knee/hip condition and how to manage it.

Before we start, I'd like to remind you that this interview will be audio recorded and then transcribed. Your responses during this interview will be kept confidential. Any data generated from this interview (e.g., the interview transcript) will not contain your name or any information that could identify you.

At any point, if you would like to skip a question or take breaks during the interview, please let me know. You are also able to cease the interview or withdraw from this study at any point with no adverse effects.

Do you have any questions before we start the interview and I start recording?

Overview

1. What were you hoping to gain when you enrolled into the study?

2. Before you joined the study, where had you received most of your information about OA and how to manage your symptoms?

What information did you receive about:

- exercise
- weight management
- surgery?

3. Overall, what did you think of the course you received as part of this study?

- What did you like?
- What didn't you like?

Common Sense Model of Self-Regulation domain: OA beliefs/illness representation

4. When you joined the study, what things did you want to learn about OA and its management?

- How well did the course meet your expectations? How/How not?

5. Can you describe if the course provided you with any new information about OA?

Any new information about:

i. what causes it

ii. its symptoms

iii. how it progresses

- Did any information surprise you?

6. Can you describe how the course influenced or changed your thoughts about treatment options for OA?

- Physical activity/Exercise?
- Weight management?
- Medications?
- Surgery—What about your thoughts about needing a Total Joint Replacement in the future?

Common Sense Model of Self-Regulation domain: Treatments/Coping strategies

7. Can you describe if the course has changed how you manage or plan to manage your joint symptoms?

- Can you describe if you are doing anything differently?
- Do you plan to do anything differently in the future?

Example prompts: seek health professional support/advice; changes in medication, physical activity, diet?

8. How could the course have better supported you in making choices about how to manage your osteoarthritis?

Common Sense Model of Self-Regulation domain: Outcomes

9. What has changed for you because of the information provided in the course? Why?

XSL•FO
RenderX

- Has the information provided in the course lead to any impact on you physically?
- Has the information provided in the course lead to any impact on you emotionally?

Wrap-up

10. After completing the Osteoarthritis Education Study, do you have any unanswered questions about OA and its management?

11. Do you have anything else you would like to add about the online course?

Data Analysis

Participants' demographic data were summarized using descriptive statistics. Thematic analysis followed a framework approach [20], which has previously been used in qualitative research applying the common sense model in hip or knee osteoarthritis [6,7]. Data were coded by 2 independent researchers with training and experience in qualitative methods (RKN and LS). Data collection and analysis occurred concurrently following the steps mentioned subsequently. First, the 2 independent researchers familiarized themselves with the interview data by reading transcripts and listening to sections of audio recordings, as required. Second, they separately conducted open coding using NVivo software (Lumivero) by rereading the transcript line by line and applying a "code" to summarize passages of data interpreted as important. Third, after the first several transcripts were coded, both researchers met to develop a working analytical framework. This involved the researchers comparing their independently applied codes and agreeing on codes to apply to subsequent transcripts. During this stage, inductively generated codes were also grouped into a priori categories guided by the common sense model: osteoarthritis beliefs, treatment choices, and outcomes. Osteoarthritis beliefs were further divided into 5 subcategories: identity, causes, timeline or progression, consequences or impacts on life, and perceived controllability. Fourth, both researchers continued to independently code subsequent transcripts according to the agreed framework and met two more times to modify the framework to include newly identified codes and discuss whether data saturation had been reached. Fifth, once data collection ceased, one researcher (RKN) organized codes and illustrative quotations into a matrix template within Microsoft Excel, with a priori categories as row headings, codes as row subheadings, and participant identifiers as column headings. This was done to facilitate the identification of patterns within and across participants and categories. The same researcher (RKN) then formulated potential overarching themes, which were reviewed by the interviewer and coder (LS). Finally, the independent coders met with the other members of the research team (KLB and RSH) to discuss the potential themes (with reference to the matrix template), which were refined until all members agreed on the final themes. To minimize bias, reflexivity was maintained by acknowledging the positions of team members and discussing how previous involvement in the development of the e-learning course might influence interpretation.

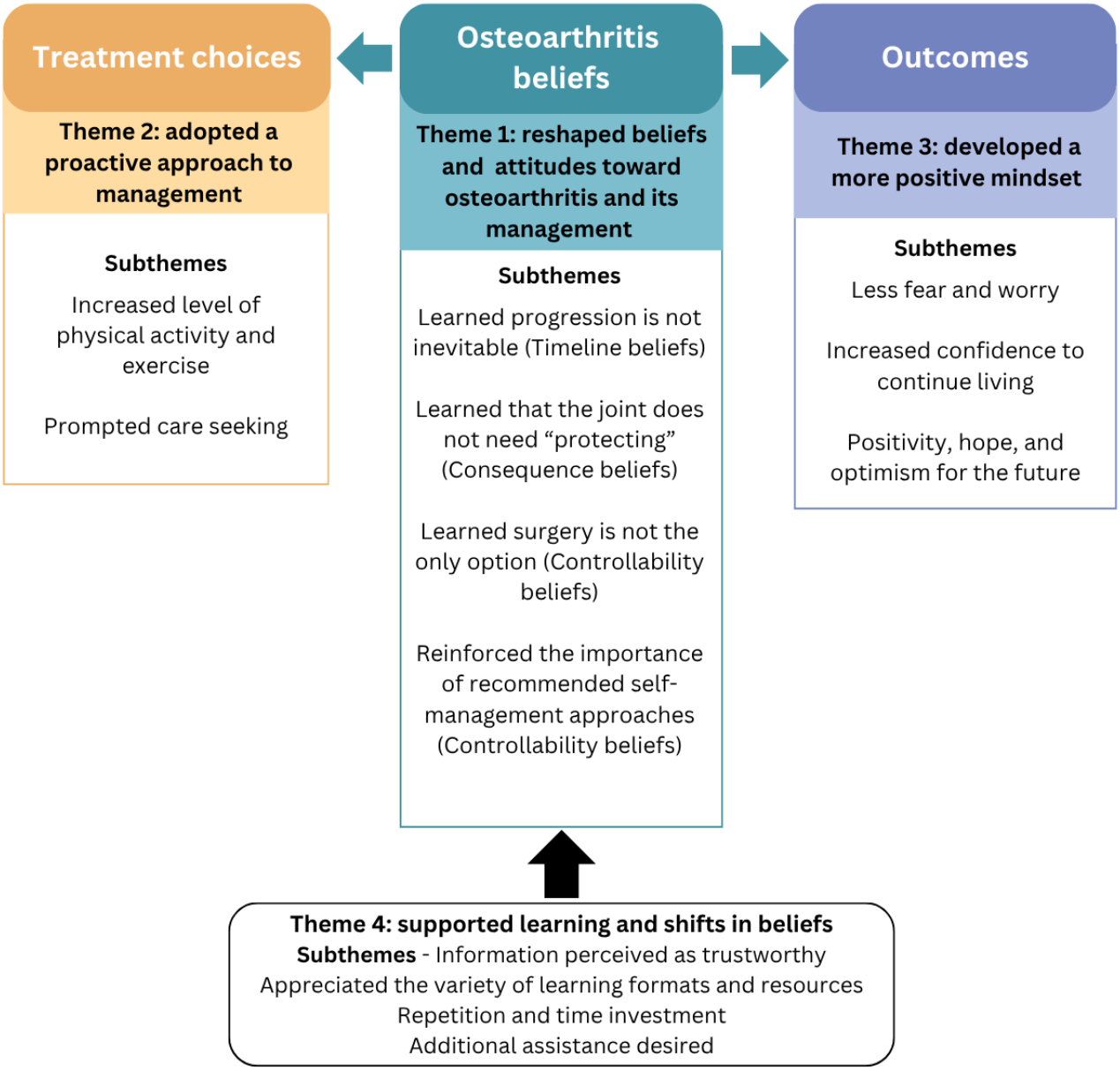
Results

Overview

A total of 20 participants were interviewed (Multimedia Appendix 1). The mean age of participants was 70 (SD 73) years, 16 (80%) were female, and 14 (70%) lived in metropolitan areas. Four themes were developed with up to 4 subthemes within each. Figure 1 presents these themes and their interactions within the common sense model. Multimedia Appendix 2 provides a summary of themes, subthemes, and exemplary quotes.

XSL•FO
RenderX

Figure 1. Themes (with subthemes) and their interactions within the common sense model.



Theme 1: Reshaped Beliefs and Attitudes Toward Osteoarthritis and Its Management

Learned Progression Is Not Inevitable

Most participants reflected that before accessing the course, they thought that osteoarthritis was an expected part of aging, and as one gets older, osteoarthritis inevitably worsens. Participants described how the course challenged these beliefs and provided them with a new understanding that osteoarthritis symptoms can be managed, stabilized, or may even improve over time.

Learned That the Joint Does Not Need "Protecting"

Participants reflected that, before accessing the course, they had been limiting their physical activities (eg, walking less, stopping running, or reducing work hours) to prevent further joint damage or deterioration to what they believed to be a vulnerable joint. Through the course, they learned that their joint affected by osteoarthritis did not need to be protected in this way.

Learned Surgery Is Not the Only Option

For all participants, the course challenged their belief that joint replacement surgery is the sole and inevitable solution—beliefs that were typically based on health professional advice or experience of family or friends undergoing the surgery. Participants felt encouraged and reported a sense of relief in having learned that surgical intervention was not necessarily needed and demonstrated an increased awareness of when surgery may be appropriate.

Reinforced the Importance of Recommended Self-Management Approaches

For all participants, the importance of staying active, engaging in regular exercise, and maintaining a healthy weight for joint

XSL•FO
RenderX

health was reinforced. Participants described a greater understanding of the mechanisms underlying why exercise and weight management were recommended, which further reinforced the importance of these approaches.

Theme 2: Adopted a Proactive Approach to Management

Increased Level of Physical Activity and Exercise

Participants reported an increased level of exercise and physical activity, which they linked to their changed beliefs (ie, the new understanding that they could be more active without causing damage), and many realized that the activities they wanted to do (but may have stopped) may be beneficial rather than exacerbating their symptoms and damaging their joints. Participants reported increasing their exercise and physical activity levels in a variety of ways, including starting new exercise routines (such as the exercise programs referred to in the course), returning to previous activities that they had ceased (eg, running), adjusting the frequency or intensity of the physical activities they were already doing (eg, spending more time in the garden or walking more), or using movement to alleviate pain (eg, breaking up sedentary time to alleviate joint pain and stiffness). For one participant, adoption of exercise as treatment led to disappointment, as despite their efforts, they did not experience the anticipated improvements that they believed were suggested as possible within the course.

Prompted Care Seeking

Many participants discussed that the course motivated them to seek guidance from a physiotherapist for osteoarthritis advice and tailored exercise plans. In addition, 3 participants indicated that the course validated their decision to consult with an orthopedic surgeon to explore joint replacement surgery as the next step in managing their osteoarthritis. They felt reassured in their decision, having tried the nonsurgical approaches recommended in the course without achieving their desired improvements and when the course did not offer further insights into new or alternative treatments.

Theme 3: Developed a More Positive Mindset

Less Fear and Worry

All participants described a reduced sense of worry and fear about having osteoarthritis because of shifted beliefs. Many described a change in how they viewed joint pain or discomfort; they were no longer fearful that it was a definite sign of "damage" and progression. Participants reported increased confidence to push through some discomfort to engage (or re-engage) in various physical activities. Many participants also reflected on the relief they felt at no longer being worried about the inevitability of a surgery.

Increased Confidence to Continue Living

Participants described a boost in confidence and feelings of empowerment and no longer feeling restricted by their previous beliefs about osteoarthritis (ie, that the joint needed to be protected and that surgery was inevitable). They described a sense of being granted permission to continue living their lives fully despite having osteoarthritis. All participants described how the course reinforced that they were on the right track and that they could keep doing the things they enjoyed (or start doing them again). This included restarting, continuing, or increasing activities such as walking, jogging, dancing, and camping and a reduced stigma associated with using a walking cane.

Positivity, Hope, and Optimism for the Future

Overall, participants described feeling happier and more positive about life, and most described a newfound optimism for the future, no longer viewing osteoarthritis as a "death sentence" and no longer believing that surgery was something that was inevitable.

Theme 4: Supported Learning and Shifts in Beliefs

Information Perceived as Trustworthy

Participants perceived that the course was created by a credible source (ie, researchers at a university), which increased their trust in the information provided.

Appreciated the Variety of Learning Formats and Resources

Overall, participants valued the variety of learning materials, such as linked resources, video stories, discussion boards, and quizzes. These options allowed them to tailor their learning experience and "get as much or as little" out of it as suited their needs. However, individual preferences varied; some found the discussion boards beneficial for learning from others' experiences, while others viewed them as platforms for personal complaints. Similarly, opinions diverged on video content, polls, quizzes, and the links provided to external resources, with some finding them valuable and others preferring not to engage with them.

Repetition and Time Investment

Similar information presented across various formats (text, images, and videos) resulted in many participants reporting that they found the course repetitive and time consuming. While some viewed this repetition positively, as a reinforcement of learning, others were frustrated and reported "skimming" various sections. The repetition became particularly bothersome when the content did not align with individual needs. For instance, several participants felt the weight management section was irrelevant to their circumstances, some did not like being asked for feedback "repeatedly" (eg, multiple prompts to engage with discussion boards, polls, and quizzes), and some found the external links hard to navigate.

Additional Assistance Desired

Most described that additional support to reinforce, supplement, or personalize course content would have been helpful. However, there were diverse opinions on the preferred form of this support. Suggestions ranged from having the discussion boards moderated to provide personalized responses to posed questions and having initial contact with a health professional to discuss individual concerns to having follow-up contact with a health professional to ask any outstanding questions or to support putting learning into action. Many also desired ongoing access to the content to enable them to revisit it as needed.

XSL•FO
RenderX

Discussion

Principal Findings

In this study, we interviewed people with hip or knee osteoarthritis who used an e-learning course about osteoarthritis and its recommended management within an RCT. This study explored the potential impact of the course on participants' beliefs, treatment choices, and outcomes related to osteoarthritis. Overall, participants valued the course, describing the ways it reshaped their beliefs and attitudes toward osteoarthritis and its management, which resulted in a more positive and optimistic mindset and increased engagement in physical activity. The perceived credibility of the course and the variety of learning formats and resources provided within it facilitated a reshaping of participants' beliefs. However, the variety of learning resources was also perceived as repetitive and time consuming by most participants.

To the best of our knowledge, this is the first study to qualitatively explore how a self-directed e-learning course for consumers can influence health condition beliefs, treatment choices, and outcomes of osteoarthritis or any other chronic condition. One study has explored, more generally, the perceptions of 20 people newly diagnosed with rheumatoid arthritis toward an e-learning course designed to support self-management [21]. Similar to our findings, the aforementioned qualitative study revealed overall positive participant perceptions toward e-learning. The participants with rheumatoid arthritis trusted the content developed by the program creators (ie, rheumatoid arthritis researchers and health professionals), who were perceived as credible, and found the course effective in supporting rheumatoid arthritis knowledge acquisition. However, some had unmet support needs (eg, emotional, social, or interpersonal support) and desired contact with a health care provider to support course content. As in our study, participants also expressed a desire for additional support to reinforce course content. Taken together, these findings suggest that for some people, an e-learning course alone may be insufficient to address their emotional and informational needs related to their health condition. Hence, future research assessing e-learning, when delivered in combination with health professional oversight, is warranted and is currently underway [22].

While participants in our study reported various physical benefits associated with completing the course (such as reduced joint pain, increased physical abilities, and weight loss), its true value appeared to lie in enhancing psychological well-being. Participants experienced reduced fear and worry about osteoarthritis and its prognosis, accompanied by newfound optimism for the future. Notably, participants expressed relief in discovering that they could lead fulfilling lives despite having osteoarthritis without worsening the condition or hastening the need for joint replacement surgery. These findings are particularly important given the substantial psychosocial impact of living with osteoarthritis. For instance, a recent systematic review comprising 21 qualitative studies (n=665) highlighted the impact that a perceived poor prognosis (such as the belief that osteoarthritis is a "progressive degenerative disease"), the perceived necessity for surgery, and one's preoccupation with joint damage have on the psychological well-being of people with knee osteoarthritis [23]. These beliefs negatively impacted quality of life and led to feelings of hopelessness, loss of control, grief, and inadequacy.

Despite the overall positive views regarding e-learning observed in our study, all participants described the course as time consuming and repetitive at times. This led some to "skim" various sections of content, particularly when they felt it did not align with their individual needs (eg, the body weight module). Many also wanted continuous access to revisit content rather than time-limited access. To reduce time pressure and better enable participants to learn at their own pace and more easily tailor how they engage with the course (eg, skip sections that feel repetitive or revisit sections at a later date to reinforce learning), we have since made the course available on the FutureLearn platform with free, unrestricted access. Future work should focus on developing strategies for large-scale implementation and dissemination.

Strengths and Limitations

Several strengths and limitations should be acknowledged. We used purposive sampling to ensure a variety of participants were included. To facilitate accurate recall, all interviews were conducted within a month of completing involvement in the RCT. To enhance rigor and validity, 2 researchers (RKN and LS) independently coded the data, with 1 (LS) having no involvement in course development. All researchers contributed to theme generation using the matrix of codes and quotes, and these processes also provided opportunities for reflexive discussion and critically challenging interpretations. There were several limitations. Despite the use of purposive sampling, selection bias may be present, as participants with more favorable opinions about the e-learning course may have been more inclined to provide consent to be interviewed. We did attempt to overcome this by deliberately recruiting people with lower perceived course usefulness; however, we purposely did not recruit those who reported that they found the course "not useful" at all in the RCT. This was because the 6 participants in the RCT who reported this also reported not having accessed the course (one because they believed the course required payment and the others for unknown reasons). While their exclusion was deemed appropriate to address our research question, it may nevertheless bias findings toward more favorable interpretations, as the perspectives of those who disengaged entirely were not captured. Finally, researcher perspectives could have affected findings, as most authors were involved in the development of the e-learning course. We attempted to address this by ensuring that the interviewer (who was also one of the independent coders) had no previous involvement in either the e-learning course development or its RCT evaluation.

Conclusions

In summary, participants in this qualitative study generally found the e-learning course beneficial, noting changes in their beliefs and attitudes toward osteoarthritis and its management. These shifts included understanding that osteoarthritis progression is not inevitable, the joint does not require constant

XSL•FO
RenderX

protection, and joint replacement is not an inevitable outcome. For this sample, such changes appeared to foster a more positive mindset, instilling newfound confidence in living with osteoarthritis and a more optimistic outlook for the future. The e-learning course is now available with free, unrestricted access and may serve as a useful resource to support people with osteoarthritis in improving their understanding of the condition and its management.

Acknowledgments

This study was funded by a University of Melbourne Early Career Researcher Grant 2023. RKN is supported by a University of Melbourne Sir Randal Heymanson Fellowship. RSH is supported by a National Health and Medical Research Council Fellowship (1154217). KLB is supported by a National Health and Medical Research Council Investigator Grant (1174431). For the purposes of open access, the authors have applied a Creative Commons Attribution public copyright license to any author-accepted manuscript version arising from this submission.

Data Availability

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

Authors' Contributions

RKN, KLB, and RSH conceived the idea for the study and developed the interview guide. LS conducted the interviews and, along with RKN, conducted the analyses. RKN wrote the manuscript, and all authors contributed to its development and finalization.

Conflicts of Interest

KLB receives personal fees from Wolters Kluwer for contributions to UpToDate knee osteoarthritis guidelines. All other authors declare no conflict of interest.

Multimedia Appendix 1
Participant details.
[PDF File (Adobe PDF File), 137 KB - aging_v8i1e80282_app1.pdf]

Multimedia Appendix 2
Themes, subthemes, and example quotes.
[PDF File (Adobe PDF File), 105 KB - aging_v8i1e80282_app2.pdf]

References

1. Hunter DJ, Bierma-Zeinstra S. Osteoarthritis. Lancet 2019 Apr 27;393(10182):1745-1759. [doi: 10.1016/S0140-6736(19)30417-9] [Medline: 31034380]
2. Guideline for the management of knee and hip osteoarthritis. The Royal Australian College of General Practitioners. URL: https://www.racgp.org.au/clinical-resources/clinical-guidelines/key-racgp-guidelines/view-all-racgp-guidelines/knee-and-hip-osteoarthritis [accessed 2025-05-29]
3. Bannuru RR, Osani M, Vaysbrot E, Arden N, Bennell K, Bierma-Zeinstra S, et al. OARSI guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis. Osteoarthritis Cartilage 2019 Nov;27(11):1578-1589 [FREE Full text] [doi: 10.1016/j.joca.2019.06.011] [Medline: 31278997]
4. Moseng T, Vliet Vlieland TP, Battista S, Beckwée D, Boyadzhieva V, Conaghan PG, et al. EULAR recommendations for the non-pharmacological core management of hip and knee osteoarthritis: 2023 update. Ann Rheum Dis 2024 May 15;83(6):730-740 [FREE Full text] [doi: 10.1136/ard-2023-225041] [Medline: 38212040]
5. Osteoarthritis in over 16s: diagnosis and management NICE guideline reference number:NG226. National Institute for Health and Care Excellence. URL: https://www.nice.org.uk/guidance/ng226 [accessed 2025-05-29]
6. de Oliveira BI, Smith AJ, O'Sullivan PP, Haebich S, Fick D, Khan R, et al. 'My hip is damaged': a qualitative investigation of people seeking care for persistent hip pain. Br J Sports Med 2020 Jul 24;54(14):858-865. [doi: 10.1136/bjsports-2019-101281] [Medline: 31980419]
7. Bunzli S, O'Brien P, Ayton D, Dowsey M, Gunn J, Choong P, et al. Misconceptions and the acceptance of evidence-based nonsurgical interventions for knee osteoarthritis. A qualitative study. Clin Orthop Relat Res 2019 Sep;477(9):1975-1983 [FREE Full text] [doi: 10.1097/CORR.0000000000000784] [Medline: 31192807]
8. Haber T, Hinman RS, Dobson F, Bunzli S, Hilton A, Hall M. Peoples' beliefs about their chronic hip pain and its care: a systematic review of qualitative studies. "I'm just getting old and breaking down". Pain 2023 May 01;164(5):926-947. [doi: 10.1097/j.pain.0000000000002792] [Medline: 36156531]

XSL•FO
RenderX

9. Leventhal H, Phillips LA, Burns E. The Common-Sense Model of Self-Regulation (CSM): a dynamic framework for understanding illness self-management. J Behav Med 2016 Dec 11;39(6):935-946. [doi: 10.1007/s10865-016-9782-2] [Medline: 27515801]
10. O'Brien P, Bunzli S, Ayton D, Dowsey MM, Gunn J, Manski-Nankervis J. What are the patient factors that impact on decisions to progress to total knee replacement? A qualitative study involving patients with knee osteoarthritis. BMJ Open 2019 Sep 24;9(9):e031310 [FREE Full text] [doi: 10.1136/bmjopen-2019-031310] [Medline: 31551388]
11. The National Osteoarthritis Strategy Project Group. National osteoarthritis strategy. The University of Sydney. URL: https://www.acsep.org.au/content/Document/National_OA_Strategy_PostSummit.pdf [accessed 2025-05-29]
12. Nelligan RK, Hinman RS, Egerton T, Gregory M, Bidgood N, McManus MF, et al. Effects of a massive open online course on osteoarthritis knowledge and pain self-efficacy in people with hip and/or knee osteoarthritis: protocol for the MOOC-OA randomised controlled trial. BMC Musculoskelet Disord 2023 May 15;24(1):381 [FREE Full text] [doi: 10.1186/s12891-023-06467-x] [Medline: 37189094]
13. Nelligan RK, Hinman RS, McManus F, De Silva AP, Gregory M, Bidgood N, et al. Effects of an eLearning course for patients on osteoarthritis knowledge and pain self-efficacy in people with hip and/or knee osteoarthritis: a randomised controlled trial. Patient Educ Couns 2025 Aug;137:108792 [FREE Full text] [doi: 10.1016/j.pec.2025.108792] [Medline: 40300349]
14. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
15. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
16. Groenewald T. A phenomenological research design illustrated. Int J Qual Methods 2004 Mar 01;3(1):42-55. [doi: 10.1177/160940690400300104]
17. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec 16;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
18. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018 Sep 14;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
19. Taking control of your hip and knee osteoarthritis. Future Learn. URL: https://www.futurelearn.com/courses/taking-control-hip-and-knee-osteoarthritis [accessed 2025-05-29]
20. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Sep 18;13(1):117 [FREE Full text] [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
21. Knudsen LR, Lomborg K, Hauge E, Zangi HA, de Thurah A. The WebRA study: opportunities and challenges in digital patient education from the perspective of patients with rheumatoid arthritis: a qualitative study. Patient Educ Couns 2023 Nov;116:107969 [FREE Full text] [doi: 10.1016/j.pec.2023.107969] [Medline: 37672918]
22. Nelligan RK, Hinman RS, Metcalf B, McManus F, Lamb KE, Heywood S, et al. Consumer-focused osteoarthritis e-learning to complement physiotherapy care in people with knee osteoarthritis: protocol for the PEEKO randomised controlled trial. Musculoskeletal Care 2025 Sep 03;23(3):e70156. [doi: 10.1002/msc.70156] [Medline: 40610701]
23. Wallis JA, Taylor NF, Bunzli S, Shields N. Experience of living with knee osteoarthritis: a systematic review of qualitative studies. BMJ Open 2019 Sep 24;9(9):e030060 [FREE Full text] [doi: 10.1136/bmjopen-2019-030060] [Medline: 31551381]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture

XSL•FO
RenderX

Edited by P Jansons; submitted 07.07.25; peer-reviewed by JJ Beilby, V Reis; comments to author 21.09.25; revised version received 01.10.25; accepted 06.10.25; published 05.11.25.

Please cite as:
Nelligan RK, Spiers L, Hinman RS, Bennell KL
Impact of a Consumer e-Learning Course on Beliefs, Treatment Choices, and Outcomes Among People With Hip and Knee Osteoarthritis: Qualitative Interview Study
JMIR Aging 2025;8:e80282
URL: https://aging.jmir.org/2025/1/e80282
doi:10.2196/80282
PMID:

©Rachel K Nelligan, Libby Spiers, Rana S Hinman, Kim L Bennell. Originally published in JMIR Aging (https://aging.jmir.org), 05.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Implementation of the German Clinical Practice Guideline for Multimorbidity Using a Digital Tool in Primary Care: Pilot Cluster Randomized Clinical Trial

Julia Nothacker[1], MSc; Valentina Paucke[1], MSc; Susanne Lezius[2], Dipl Biomath; Antonia Zapf[2], Prof Dr; Dagmar Lühmann[1], Dr med; Martin Scherer[1], Prof Dr Med; Ingmar Schäfer[1], PD, Dr phil

[1]Institute and Outpatients Clinic of General Practice and Primary Care, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[2]Institute of Medical Biometry and Epidemiology, University Medical Center Hamburg-Eppendorf, Hamburg, null, Germany

Corresponding Author:
Ingmar Schäfer, PD, Dr phil
Institute and Outpatients Clinic of General Practice and Primary Care
University Medical Center Hamburg-Eppendorf
Martinistr. 52
Hamburg, 20246
Germany
Phone: 49 40 7410 52760
Email: in.schaefer@uke.de

Abstract

Background: Clinical practice guidelines (CPGs) summarize the best available evidence in a specific field. To improve patient-centered outcomes, guidelines have to be implemented, using, for example, information and communications technology. Although there are CPGs addressing multimorbidity, there is still a lack of studies investigating their implementation.

Objective: This study aimed to evaluate whether the implementation of a CPG for multimorbidity using a digital tool is feasible and explore possible effects of this intervention.

Methods: A pilot cluster randomized clinical trial based on telephone interviews was conducted from October 25, 2023, to September 8, 2024. Patients enrolled in any disease management program who were aged ≥65 years and had at least 2 additional chronic conditions were randomly selected from 20 general practitioner (GP) practices and contacted for informed consent. Each practice was randomized after baseline interviews of all participating patients in the practice were finished. The use of a web application facilitating collection and documentation of treatment-relevant data in accordance with the German CPG for multimorbidity was compared with treatment as usual. The primary outcome was time spent in hospital. As a secondary outcome, the number of patients with at least one hospital admission was calculated. Further secondary outcomes included outpatient health care use, quality of life, patient satisfaction, and quality of care. Feasibility assessment included examination of sample size, participation rate, and compliance with the study protocol. Outcome measures were analyzed using linear, logistic, and negative binomial regressions with random intercepts for practices.

Results: Of 384 patients who were contacted, 123 (32%) agreed to participate, and 120 (31.3%, including 54/120, 45% in the intervention group and 66/120, 55% in the control group) completed baseline and follow-up assessments. Mean age was 75.4 (SD 6.6) years, and 51.7% (62/120) were women. The compliance rate, or the proportion of patients who were treated per protocol, was 89% (48/54). In our data, the incidence rate of hospital days was comparable in both groups (incidence rate ratio [IRR] 0.94, 95% CI 0.09-9.42; *P*=.96), but the odds of hospital admission were almost reduced by half in the intervention group (odds ratio 0.51, 95% CI 0.17-1.54; *P*=.23). Our data also suggest a higher incidence rate of GP contacts (IRR 1.13, 95% CI 0.83-1.53; *P*=.43) and lower incidence rate of contacts with outpatient specialists (IRR 0.79, 95% CI 0.54-1.15; *P*=.24) in the intervention group compared to usual care. Moreover, patients and GPs reported a better quality of care (mean difference 0.51, 95% CI −0.12 to 1.14; *P*=.12 and mean difference 1.19, 95% CI 0.13-2.25; *P*=.03, respectively) in the intervention group.

Conclusions: Implementation of the CPG using a digital tool was feasible. Our data suggest that the probability of hospital admissions and contacts with outpatient specialists might be reduced and quality of care might be improved.

Trial Registration: ClinicalTrials.gov NCT06061172; https://clinicaltrials.gov/study/NCT06061172

XSL•FO
RenderX

(JMIR Aging 2025;8:e79767) doi:10.2196/79767

KEYWORDS

multimorbidity; implementation; hospitalizations; clinical practice guidelines

Introduction

Clinical practice guidelines (CPGs) identify and summarize the best available evidence in a specific field [1]. To improve patient-centered outcomes, guidelines have to be implemented in regular health care. In contrast to development and dissemination, the implementation of guidelines is usually the responsibility of health care professionals [2]. There are various barriers to implementing guidelines, including time constraints and lack of resources [3,4]. One approach in optimizing guideline implementation is to provide structural changes by using information and communications technology [5,6]. A strength of this approach is facilitation of complex care [7] (eg, for patients with multimorbidity).

Multimorbidity is a common health problem in the primary care of older patients [8] that is associated with adverse outcomes such as decline in health-related quality of life [9] and increased risk of hospitalizations [10]. Patients with multimorbidity usually need tailored, person-centered care based on a prioritization of treatment goals [11]. In a recent systematic review, 20 guidelines specifically focusing on multimorbidity and polypharmacy were identified [12], but none have been implemented using a digital tool yet.

In Germany, the S3-level CPG for multimorbidity of the German Society of General Practitioners and Family Physicians does not focus on specific diseases but rather helps the general practitioner (GP) manage the patients' individual situation [13]. A field test emphasized the high complexity of the treatment algorithm and the need for specific assessment tools [14]. Therefore, we developed a digital tool that facilitates the implementation of this guideline [15] and conducted a cluster randomized clinical pilot study that examined the feasibility of the full-scale evaluation of an intervention based on this tool and explored the possible effects of this intervention.

Specifically, the primary objective of this study was to assess whether participant enrollment, data acquisition, and the implementation of the intervention in the intended way are feasible. In addition, our pilot study aimed to consolidate our hypotheses by analyzing whether the intervention is capable of reducing the time patients spend in hospital and their outpatient health care use and whether it might improve quality of care, patients' health-related quality of life, and patient satisfaction.

Methods

Study Oversight

Reporting followed the CONSORT (Consolidated Standards of Reporting Trials) extension for cluster randomized trials [16]. A project advisory board represented patients, caregivers, primary care physicians, self-governing bodies in the German health care system, the German professional society for primary care, and private companies developing digital therapeutics. The board informed the research team about the perspectives of health care providers and patients and provided guidance on how their interests and concerns can be considered in the design of our intervention and our pilot study.

Ethical Considerations

The protocol for this pilot study was approved on June 27, 2022, and September 5, 2022, by the ethics committee of the Medical Association of Hamburg (2022-100786-BO-ff) and registered on September 24, 2023, in ClinicalTrials.gov (NCT06061172).

Trial Design

We conducted a 2-group, parallel pilot cluster randomized clinical trial based on telephone interviews of older patients with multimorbidity and their GPs. On average, patient interviews lasted approximately 20 minutes, and GP interviews lasted approximately 10 minutes per patient. Answers were directly entered in deidentified form into Microsoft Excel spreadsheets. We conducted automated plausibility checks during data entry and retrospective evaluation of data quality before data preparation and analysis.

Participants

A total of 20 GP practices in the German federal states of Bremen, Hamburg, Lower Saxony, and Schleswig-Holstein were recruited via mail as a convenience sample. We included practices who were known from other research projects and formerly unknown practices identified on the official websites of the regional associations of statutory health insurance physicians. GPs were eligible if (1) they were registered as statutory health insurance physicians in the metropolitan area of Hamburg or its rural surroundings; (2) they participated in a disease management program for asthma, coronary heart disease, diabetes mellitus, or chronic obstructive pulmonary disease; and (3) they could provide a list of patients from their practice who were aged ≥65 years and participated in such a program.

GPs created a numbered list of all patients who fulfilled these criteria. A member of the study team then assigned them random numbers using a mobile app. On the basis of these numbers, GPs selected patients from the list and screened them for multimorbidity and exclusion criteria until 20 eligible patients in each practice were identified. Multimorbidity was operationalized as 2 additional diagnoses considering the 46 most prevalent chronic diseases [17]. Exclusion criteria included having no capacity to consent, functional limitations precluding participation in the intervention (eg, loss of vision or no capacity to use a mobile device) or telephone interviews (eg, loss of hearing), limited German language skills, and participation in other trials during the pilot study.

Eligible patients received an invitation letter from their GP that included written patient information, 2 consent forms, and a prepaid return envelope. Patients who wanted to participate signed one consent form, gave their contact data, and returned both to the study center. Participating patients were then

XSL•FO
RenderX

contacted and interviewed by the study team. In addition, attending GPs were interviewed about their patients and practice. Interviews were conducted at baseline and after the intervention period was finished (follow-up).

Sample Size

As this was a pilot trial, we did not conduct a formal sample size calculation. For confirmation of feasibility, we wanted to include 10 practices for each treatment arm with, on average, 5 patients per practice. Allowing for a reasonable dropout, we aimed at recruiting 6 patients in each practice. In summary, our intended sample size was 20 practices with 120 patients altogether. This is suitable according to general recommendations for minimum sample sizes [18,19].

Randomization

Each practice was randomized after the baseline interviews of all participating patients in the practice were finished. We used block randomization with computer-generated random numbers and variable block sizes stratified for region type (ie, rural and urban), which was generated by an independent statistician. The statistician received the practice number and region type via email but did not have access to any other patient or practice data. After randomization, the study team was informed about practice allocation via email and communicated this information to participating GPs via telephone. Patients were not actively informed about practice allocation but could not effectively be blinded.

Intervention and Comparator

The intervention was based on the web application MultiTool (version 1.0.0; University Medical Center Hamburg-Eppendorf), which implements all recommendations of the German Society of General Practitioners and Family Physicians S3-level CPG for multimorbidity [13]. Guideline recommendations, specific functions, and other contents of the tool are described in detail in another publication [15].

The main functions of the tool were digital questionnaires in the domains of (1) patient preferences (ie, prioritization of treatment goals, control preferences, and involvement of other health care professionals), (2) social life (ie, social contacts, activities, and participation), (3) treatment situation (ie, medication adherence and treatment burden), and (4) health concerns (ie, pain, anxiety and depression, and other health concerns).

GPs could give patients access to the questionnaires via email, and patients could fill out the questionnaires at home or in the GP practice using any digital device with a browser and access to the internet. The tool also facilitated “brown bag” medication reviews [20] by providing a standardized invitation letter clarifying which drugs patients should bring to the consultations and a guide on how the reviews should be structured by the GPs. Results from discussions between GPs and patients could be documented within the tool.

GPs in the intervention group received access to MultiTool, a short training on the digital tool and the intervention, a written manual, and contact data for telephone support. Approximately 3 months after randomization, the practices allocated to the intervention group were contacted via telephone by the study team to monitor the implementation of the intervention. During this call, possible adverse events were assessed by asking GPs in unstandardized form whether any undesirable or harmful event had occurred during the observation period, which may or may not be related to the intervention [21].

GPs in the intervention group had to use MultiTool for every patient participating in the trial at least once in every 3-month accounting period. They were further instructed, first, to use the questionnaire for prioritization of treatment goals and at least one questionnaire in each of the domains of social life, treatment situation, and health concerns. Second, the clinical consequences of the patients’ answers had to be discussed between GPs and patients, and third, at least one medication review during the trial was obligatory. Compliance with the study protocol was reported by GPs and defined as carrying out these 3 instructions with their patients.

GPs in the control group had no access to MultiTool and provided care as usual. Exposure time in both groups was 6 months starting from the randomization of the practice.

Clinical Outcomes and Other Variables

The primary outcome measure was time spent in hospital during the preceding 6 months. It was assessed in patient interviews by asking the patients how often they had been in hospital and how many days each hospital stay had lasted. The total time in hospital was calculated by adding the duration of all stays. As a secondary outcome, the number of patients with at least one hospital admission during the previous 6 months was also calculated.

Further secondary outcome measures included the number of contacts with the GP, outpatient specialists, and home care services during the preceding 6 months. Self-rated health was assessed on a scale from 0 to 100. Health-related quality of life was measured using the EQ-5D-5L, which comprises the domains of mobility, self-care, usual activities, pain or discomfort, and anxiety or depression [22], and the corresponding German value set, which assumes a value of 1 for perfect health and subtracts values between 0.026 and 0.612 for limitations in each of the 5 domains [23]. Patient satisfaction with clinical performance (17 items) and organization of care (6 items) were assessed on a 5-point Likert scale (0=bad; 4=excellent) using the European Task Force on Patient Evaluation of General Practice instrument [24]. For both subscales, a summary score was calculated by adding the values of each nonmissing item and dividing by the number of available answers. All these data were collected in patient interviews. In addition, quality of care was assessed using patient-reported (7 items) and GP-reported (12 items) summary scores of validated quality indicators for the primary care of patients with multimorbidity [25]. These scores were calculated by adding 1 point for each fulfilled indicator.

Patient interviews also included data on age, sex, marital status, living arrangement, educational level, and migration background. The educational level was coded using the Comparative Analysis of Social Mobility in Industrial Nations classification of tertiary, secondary, and primary education and

XSL•FO
RenderX

lower [26]. Migration background was defined by the country of birth of the study participants and their parents. Moreover, depressiveness was assessed in patient interviews using a summary score of the 15 items of the Geriatric Depression Scale [27], which gives 1 point for each positive answer. Additionally, the number of chronic conditions was collected in GP interviews.

Feasibility Assessment

To analyze the feasibility of recruitment of practices and patients, sample size and participation rate were analyzed. A priori, we expected that 400 eligible individuals could be identified and contacted and that 120 participants could be enrolled, which corresponds to an anticipated participation rate of 30%. To assess the quality of data acquisition, we analyzed the number of missing values for each variable. To examine the implementation of the intervention, we analyzed how many participants in each practice were compliant with the study protocol.

Statistical Analyses

Descriptive comparisons between the intervention and control groups were reported using means and SDs for continuous data and numbers and percentages for categorical data. Potential differences between patients in the intervention and control groups regarding their baseline characteristics and the primary and secondary outcomes were assessed using 2-tailed *t* tests and chi-square tests. Distinct baseline imbalances were defined by *P* values of <.05. Rates of change were calculated using the following formula:



(1)

According to their type, primary and secondary outcomes were analyzed using multilevel mixed-effects negative binomial, logistic, or linear regression, and the respective effect measures with corresponding 95% CIs were reported. All statistical models were adjusted for random effects at the level of GP practices, and negative binomial and linear regression models were also adjusted at the level of patients within practices. The secondary outcome of "proportion of patients who stayed in hospitals" was compared at follow-up and controlled for the number of hospital stays at baseline. Intraclass correlation coefficients were reported to quantify the correlation among observation within the clusters. For all analyses, the available dataset was used.

The capability of our intervention to influence outcomes was determined by the effect sizes of these analyses. As the power of our pilot study was not adequate, hypotheses were consolidated solely based on regression coefficients. In addition, the reported *P* values are to be understood as descriptive measures and not used for confirmatory conclusions. All statistical analyses were conducted using Stata (version 15.1; StataCorp) [28].

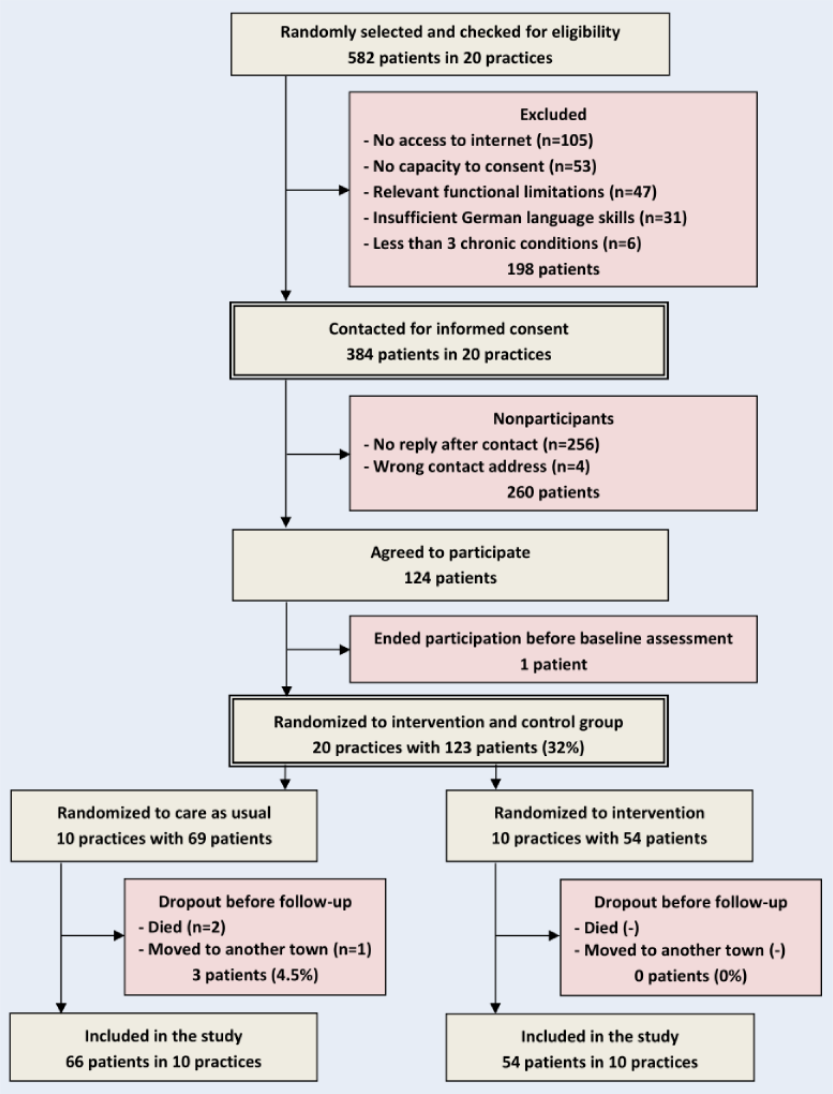
Results

Participants

The participant flow can be found in Figure 1. A total of 582 patients from 20 practices were screened for eligibility, and 384 (66%) were contacted by their GPs for informed consent. Baseline interviews were conducted between October 25, 2023, and January 31, 2024. Subsequently, 20 practices with 32% (123/384) of the contacted patients were randomized to the intervention and control groups. A total of 1.6% (2/123) of the patients died during the observation period, and 0.8% (1/123) moved to another town, both in the control group. The dropout rate was 2.4% (3/123) altogether. The final sample size was 66 patients in the control group and 54 patients in the intervention group. Follow-up interviews were conducted between June 3, 2024, and September 8, 2024.

XSL•FO
RenderX

Figure 1. Sampling and response rates.



Clinical Outcomes and Other Variables

On average, GPs in the control group were aged 52.0 (SD 8.4) years, and GPs in the intervention group were aged 54.1 (SD 8.4) years. In both groups, 60% (6/10) of the GPs were men, and 40% (4/10) were women. Both groups had a comparable mean number of GPs working in the practices (3.0, SD 1.5 in the intervention group vs 3.0, SD 2.0 in the control group) and a comparable number of treated patients per 3-month accounting period (mean 2300, SD 558 in the intervention group vs mean 2560, SD 1285 in the control group). The sociodemographic data of patients at baseline are reported in Table 1. There were no distinct baseline imbalances related to age (mean 75.8, SD 6.5 vs mean 74.9, SD 6.8 years; *P*=.48), sex (35/66, 53% men in the intervention group vs 23/54, 43% men in the control group; *P*=.26), or other sociodemographic data, but the mean number of chronic conditions in the control group was distinctly higher than that in the intervention group (10.4, SD 4.8 vs 8.8, SD 3.1, respectively; *P*=.03). The summary score on the Geriatric Depression Scale was comparable between both groups (mean 6.0, SD 1.5 vs mean 6.1, SD 1.7; *P*=.59).

There were no distinct imbalances in primary and secondary outcome measures at baseline between the control and intervention groups (Table 2). However, in the control group, fewer days in hospital during the 6 months before the baseline assessment were reported than in the intervention group (mean 0.9, SD 3.5 vs mean 2.4, SD 6.9, respectively; *P*=.14), and the probability of at least one hospital admission was also lower (11/66, 17% in the intervention group vs 12/54, 22% in the control group; *P*=.44). At follow-up, the probability of at least one hospital admission increased in the control group and decreased in the intervention group (13/66, 20% vs 8/54, 15%, respectively).

Home care services were only used by 6.7% (8/120) of the participants at baseline and 4.2% (5/120) of the participants at follow-up, and therefore, the effect of the intervention on the number of contacts with these services could not be analyzed. There was one missing value (1/66, 2%) for GP-reported quality of care in the control group and another missing value (1/54, 2%) for health-related quality of life scores in the intervention group. All other variables did not have any missing values.

XSL•FO
RenderX

Table 1. Sociodemographic data of the patients at baseline.

Characteristics	Control group (n=66)	Intervention group (n=54)	*P* value
Age (y), mean (SD)	75.8 (6.5)	74.9 (6.8)	.48
Sex (male), n (%)	35 (53)	23 (43)	.26
Marital status, n (%)			.38
Married	48 (73)	31 (57)	
Widowed	12 (18)	15 (28)	
Divorced	3 (5)	4 (7)	
Single	3 (5)	4 (7)	
Living arrangement (alone), n (%)	13 (20)	19 (35)	.06
Educational level, n (%)			.49
Primary or lower	42 (64)	29 (54)	
Secondary	16 (24)	15 (28)	
Tertiary	8 (12)	10 (19)	
Migration status, n (%)			.64
Patient and parents born in Germany	49 (74)	38 (70)	
Patient or at least one parent born abroad	17 (26)	16 (30)	

XSL•FO
RenderX

Table 2. Differences in primary and secondary outcomes at baseline and follow-up between the control and intervention groups.

	Control group	Intervention group	*P* value
Primary outcome			
Days in hospital (previous 6 mo)			.14
Baseline, mean (SD)	0.9 (3.5)	2.4 (6.9)	
Follow-up, mean (SD)	1.3 (4.1)	3.1 (9.8)	
Change, %	+38	+31	
Secondary outcomes			
Number of contacts with GP[a] (previous 6 mo)			.37
Baseline, mean (SD)	3.7 (2.7)	3.3 (2.7)	
Follow-up, mean (SD)	2.8 (2.1)	2.7 (1.8)	
Change, %	−25	−16	
Number of contacts with outpatient specialists (previous 6 mo)			.35
Baseline, mean (SD)	1.6 (1.5)	1.9 (2.1)	
Follow-up, mean (SD)	2.0 (2.2)	1.9 (1.7)	
Change, %	+24	−2	
GP-reported quality of care (previous 6 mo, 0-12)			.92
Baseline, mean (SD)	8.0 (2.6)	7.9 (2.8)	
Follow-up, mean (SD)	9.0 (3.0)	10 (1.6)	
Change, %	+13	+28	
Patient-reported quality of care (previous 6 mo, 0-7)			.17
Baseline, mean (SD)	3.0 (2.2)	3.6 (1.9)	
Follow-up, mean (SD)	2.9 (2.1)	3.9 (2.2)	
Change, %	−5	+10	
Self-rated health—EQ-5D-5L visual analog scale (0-100)			.94
Baseline, mean (SD)	72 (18)	72 (19)	
Follow-up, mean (SD)	73 (21)	75 (17)	
Change, %	+1	+4	
Health-related quality of life—EQ-5D-5L German value set (–0.661 to 1)			.14
Baseline, mean (SD)	0.8 (0.2)	0.8 (0.2)	
Follow-up, mean (SD)	0.8 (0.2)	0.9 (0.2)	
Change, %	+3	+1	
Patient satisfaction—EUROPEP[b] clinical performance subscale (0-4)			.11
Baseline, mean (SD)	3.2 (0.7)	3.4 (0.6)	
Follow-up, mean (SD)	3.1 (0.9)	3.4 (0.6)	
Change, %	−4	−2	
Patient satisfaction—EUROPEP organization of care subscale (0-4)			.71
Baseline, mean (SD)	3.2 (0.7)	3.1 (0.7)	
Follow-up, mean (SD)	3.1 (0.7)	3.1 (0.7)	
Change, %	−3	−3	

[a]GP: general practitioner.

[b]EUROPEP: European Task Force on Patient Evaluation of General Practice.

XSL•FO
RenderX

Feasibility Assessment

The intervention started in the first practice on November 27, 2023, and ended in the last practice on August 1, 2024. Table 3 provides an overview of compliance with the study protocol in each practice. In 70% (7/10) of the practices, all elements of the intervention were implemented per protocol with every participating patient, and in 30% (3/10) of the practices, the protocol was only partially followed. In summary, 98% (53/54) of the patients conducted assessments, 89% (48/54) participated in the intended discussions about the assessment results, and 98% (53/54) received a structured medication review. The compliance rate of patients who were treated per protocol was 89% (48/54). During the intervention, no adverse effects were reported by the GPs.

Table 3. Compliance with the study protocol by practice.

	Conducted assessments, n (%)	Discussion of assessment results, n (%)	Structured medication review, n (%)
Practice 1 (n=8 patients)	8 (100)	8 (100)	8 (100)
Practice 2 (n=5 patients)	4 (80)	4 (80)	4 (80)
Practice 3 (n=6 patients)	6 (100)	6 (100)	6 (100)
Practice 4 (n=2 patients)	2 (100)	0 (0)	2 (100)
Practice 5 (n=2 patients)	2 (100)	2 (100)	2 (100)
Practice 6 (n=5 patients)	5 (100)	5 (100)	5 (100)
Practice 7 (n=4 patients)	4 (100)	4 (100)	4 (100)
Practice 8 (n=10 patients)	10 (100)	10 (100)	10 (100)
Practice 9 (n=8 patients)	8 (100)	8 (100)	8 (100)
Practice 10 (n=4 patients)	4 (100)	1 (25)	4 (100)

Statistical Analyses

Associations between the intervention and primary and secondary outcome measures are shown in Table 4. In the analysis of the primary end point, in both groups, a comparable incidence rate of hospital days was detected (incidence rate ratio [IRR] 0.94, 95% CI 0.09-9.42; *P*=.96). In contrast, the alternative operationalization of the primary end point indicated almost half the odds of hospitalization in the intervention group (odds ratio 0.51, 95% CI 0.17-1.54; *P*=.23). Our data also suggest a higher incidence rate of GP contacts in the intervention group (IRR 1.13, 95% CI 0.83-1.53; *P*=.43) and a lower incidence rate of contacts with outpatient specialists (IRR 0.79, 95% CI 0.54-1.15; *P*=.24) compared to care as usual. Patients reported a higher quality of care in the intervention group (mean difference 0.51, 95% CI −0.12 to 1.14; *P*=.12), and the intervention was also associated with higher GP-reported quality of care (mean difference 1.19, 95% CI 0.13-2.25; *P*=.03). We observed no relevant association with the other secondary outcome measures.

XSL•FO
RenderX

Table 4. Ratios and differences between the control and intervention groups in the primary and secondary outcomes.

Outcome	Measure (95% CI)	*P* value	ICC_1[a]	ICC_2[b]
Days in hospital (primary): incidence rate ratio[c]	0.94 (0.09 to 9.42)	.96	—[d]	—[d]
At least one hospital admission (secondary): odds ratio[e]	0.51 (0.17 to 1.54)	.23	0.013	—[d]
Number of contacts with GP[f] (secondary): incidence rate ratio[c]	1.13 (0.83 to 1.53)	.43	—[d]	—[d]
Number of contacts with outpatient specialists (secondary): incidence rate ratio[c]	0.79 (0.54 to 1.15)	.24	—[d]	—[d]
GP-reported quality of care (secondary): mean difference[g]	1.19 (0.13 to 2.25)	.03	0.223	0.348
Patient-reported quality of care (secondary): mean difference[g]	0.51 (−0.12 to 1.14)	.12	0.091	0.649
Self-rated health—EQ-5D-5L visual analog scale (secondary): mean difference[g]	1.98 (−4.36 to 8.33)	.54	0.017	0.553
Health-related quality of life—EQ-5D-5L German value set (secondary): mean difference[g]	−0.02 (−0.08 to 0.05)	.58	0.002	0.661
Patient satisfaction—EUROPEP[h] clinical performance subscale (secondary): mean difference[g]	0.07 (−0.12 to 0.27)	.45	0.066	0.703
Patient satisfaction—EUROPEP organization of care subscale (secondary): mean difference[g]	0.04 (−0.18 to 0.26)	.70	0.132	0.626

[a]ICC_1: intraclass correlation coefficient at the practice level.

[b]ICC_2: intraclass correlation coefficient at the patient in practice level.

[c]Analyzed using multilevel mixed-effects negative binomial regression.

[d]Not applicable.

[e]Analyzed using multilevel mixed-effects logistic regression.

[f]GP: general practitioner.

[g]Analyzed using multilevel mixed-effects linear regression.

[h]EUROPEP: European Task Force on Patient Evaluation of General Practice.

Discussion

Principal Findings

In this pilot study, we were able to enroll the intended number of patients and collect almost all the intended data, and for 89% (48/54) of the patients, the intervention was implemented per protocol. Our results suggest that fewer hospital admissions, a better quality of care [25], and a reduced number of contacts with outpatient specialists could be effects of the intervention. Similarly, an increase in the incidence rate of contacts with the GP is a possible result of the intervention.

Comparison With the Literature

Guideline implementation using only written and oral dissemination is known to be ineffective [29]. Instead, proactive strategies tailored to each setting and target group are recommended to improve the management of chronic conditions [30,31]. However, to date, many implementation strategies have not been researched sufficiently [32], and few digital health interventions for the management of chronic diseases have demonstrated benefits in patient care [33].

For example, a clinical decision support system comprising recommendations from several single-disease guidelines led to a reduction in cardiovascular risk factors in patients with diabetes and a reduced risk of hospitalization [34]. Although a scoping review identified 3 other studies based on clinical guidelines [35], one did not report any benefits of the intervention [36], and the others did not assess any clinical outcomes [37,38]. Clinical results of an ongoing study using evidence-based, computerized decision-making for older patients with multimorbidity are still pending [39-41].

The high level of complexity (eg, due to heterogeneity of the population [17,42], fragmented care [43-45], and polypharmacy [8,46]) leads to difficulties in testing interventions designed to improve health outcomes for patients with multimorbidity [47]. A Cochrane systematic review comprising 17 randomized clinical trials [47] found little differences in clinical outcomes (eg, small improvements in patient-reported outcomes and medication adherence). Most of the studies included in the aforementioned Cochrane review and other reviews [48,49] did not find significant differences in hospitalizations.

Implications

Our pilot study demonstrated that implementation of a CPG for multimorbidity in GP practices using a digital tool is feasible, and our data suggest that patient-oriented outcomes can possibly be improved. Therefore, policymakers and clinicians should facilitate digital implementation of existing guidelines for multimorbidity, but due to the lack of robust data on their effectiveness, these approaches need to be thoroughly evaluated.

With regard to our target group, other studies have pointed to possible barriers, including functional impairment and limited adherence [50-52]. In our study, 98% (53/54) of the older patients with multiple chronic conditions shared their data using

XSL•FO
RenderX

the digital tool and participated in digitally assisted medication reviews. One facilitator in our study might be that the digital tool was developed and customized with participation of the target groups [15]. Therefore, similar studies should also consider participatory approaches.

Hospitalization is associated with multimorbidity [10], and it is a highly relevant outcome for this patient group [53]. However, there are various ways to operationalize hospitalization, and few interventions have reduced the number of hospitalizations of patients with multimorbidity [49]. In this context, our data suggest that the probability of at least one hospitalization might be a better operationalization than the number of hospital days.

Another implication for research is that health care use should be monitored when similar interventions are conducted. In line with other studies strengthening primary care [54,55], our data suggest that, on the one hand, the number of contacts with outpatient specialists could be reduced but, on the other hand, the number of contacts with GPs might increase due to this intervention.

Strengths and Limitations

The strengths of this study include the experimental design with pretest-posttest comparison, 2 treatment groups, cluster randomization, and prospective study registration, which all contribute to the robustness of our analyses. The sample size was based on recommendations from the literature. Interviewers received standardized training in advance, and data quality was further improved through numerous automated and retrospective plausibility checks. In-person interviews by telephone were chosen to increase the response rate and ensure that study participants were able to ask questions during the interviews, which made wrong answers due to misunderstandings less likely. With 1078 of 1080 values in outcome variables being available, we only had 0.19% (2/1080) of missing values in our data. Multilevel mixed-effects regression analyses considered similarities between patients and between treatment approaches at the practice level. In addition, participation of the target groups in the development of the intervention [15] and a project advisory board representing patients and the public improved the clinical relevance and applicability of our intervention.

However, due to the pilot study design and missing adjustment for multiplicity, the results of our analyses have to be interpreted as only exploratory and not confirmatory. Other limitations of this study include different sample sizes in the intervention and control group, which reduce the power of the study. We also found mild but mostly indistinct imbalances at baseline [56]. Therefore, we will include the number of recruited patients as an additional stratum for randomization, and sensitivity analyses [57] will be adjusted for possible confounding. However, as the full-scale evaluation will have a sample size 5 to 6 times higher, we expect that differences between both groups will generally be less pronounced.

Another limitation is the compliance rate of 89% (48/54) of patients being treated per protocol [58], which is still acceptable in comparison with those of other studies [59,60]. In the evaluation study, we will improve compliance with the trial protocol through a stricter monitoring of practices. In addition, we will equip intervention practices with tablet computers to facilitate digital assessments for patients without their own mobile devices.

To optimize outcome measurement in the evaluation study, we tested 2 operationalizations of the primary outcome, one of which was not defined in the registered protocol. Therefore, results regarding this measure should be interpreted with caution. Moreover, outcomes were assessed via self-reports, which are prone to social desirability bias, recall problems, and errors. In contrast to this pilot study, it is also planned to blind assessors in the full-scale evaluation.

Conclusions

This pilot study demonstrated that participant enrollment and data acquisition for evaluation and implementation of the intervention were feasible. Our data suggest that the probability of hospital admissions and contacts with outpatient specialists might be reduced and quality of care might be improved by our intervention. A full-scale cluster randomized clinical trial is now warranted to confirm these effects.

Acknowledgments

This study was funded by the German Federal Ministry of Education and Research under grant 01GY2109. The authors also acknowledge financial support from the Open Access Publication Fund of University Medical Center Hamburg-Eppendorf. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' Contributions

JN, SL, AZ, DL, MS, and IS conceived and design the study. SL, AZ, and IS conducted the statistical analyses. JN, VP, SL, AZ, DL, and IS interpreted the data. JN and IS drafted the manuscript, and the other authors reviewed it critically for important intellectual content. All authors approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-EHEALTH checklist (V 1.6.x).

XSL•FO
RenderX

[PDF File (Adobe PDF File), 1332 KB - aging_v8i1e79767_app1.pdf]

References

1. Lenzer J, Hoffman JR, Furberg CD, Ioannidis JP, Guideline Panel Review Working Group. Ensuring the integrity of clinical practice guidelines: a tool for protecting patients. BMJ 2013 Sep 17;347:f5535 [FREE Full text] [doi: 10.1136/bmj.f5535] [Medline: 24046286]
2. Shekelle P, Woolf S, Grimshaw JM, Schünemann HJ, Eccles MP. Developing clinical practice guidelines: reviewing, reporting, and publishing guidelines; updating guidelines; and the emerging issues of enhancing guideline implementability and accounting for comorbid conditions in guideline development. Implement Sci 2012 Jul 04;7:62 [FREE Full text] [doi: 10.1186/1748-5908-7-62] [Medline: 22762242]
3. McArthur C, Bai Y, Hewston P, Giangregorio L, Straus S, Papaioannou A. Barriers and facilitators to implementing evidence-based guidelines in long-term care: a qualitative evidence synthesis. Implement Sci 2021 Jul 09;16(1):70 [FREE Full text] [doi: 10.1186/s13012-021-01140-0] [Medline: 34243789]
4. Wang T, Tan JY, Liu XL, Zhao I. Barriers and enablers to implementing clinical practice guidelines in primary care: an overview of systematic reviews. BMJ Open 2023 Jan 06;13(1):e062158 [FREE Full text] [doi: 10.1136/bmjopen-2022-062158] [Medline: 36609329]
5. Peters S, Sukumar K, Blanchard S, Ramasamy A, Malinowski J, Ginex P, et al. Trends in guideline implementation: an updated scoping review. Implement Sci 2022 Jul 23;17(1):50 [FREE Full text] [doi: 10.1186/s13012-022-01223-6] [Medline: 35870974]
6. De Hert S, Paula-Garcia WN. Implementation of guidelines in clinical practice; barriers and strategies. Curr Opin Anaesthesiol 2024 Apr 01;37(2):155-162. [doi: 10.1097/ACO.0000000000001344] [Medline: 38390877]
7. Sucher JF, Moore FA, Todd SR, Sailors RM, McKinley BA. Computerized clinical decision support: a technology to implement and validate evidence based guidelines. J Trauma 2008 Mar;64(2):520-537. [doi: 10.1097/TA.0b013e3181601812] [Medline: 18301226]
8. Nicholson K, Liu W, Fitzpatrick D, Hardacre KA, Roberts S, Salerno J, et al. Prevalence of multimorbidity and polypharmacy among adults and older adults: a systematic review. Lancet Healthy Longev 2024 Apr;5(4):e287-e296 [FREE Full text] [doi: 10.1016/S2666-7568(24)00007-2] [Medline: 38452787]
9. Makovski TT, Schmitz S, Zeegers MP, Stranges S, van den Akker M. Multimorbidity and quality of life: systematic literature review and meta-analysis. Ageing Res Rev 2019 Aug;53:100903. [doi: 10.1016/j.arr.2019.04.005] [Medline: 31048032]
10. Rodrigues LP, de Oliveira Rezende AT, Delpino FM, Mendonça CR, Noll M, Nunes BP, et al. Association between multimorbidity and hospitalization in older adults: systematic review and meta-analysis. Age Ageing 2022 Jul 01;51(7):afac155 [FREE Full text] [doi: 10.1093/ageing/afac155] [Medline: 35871422]
11. Skou ST, Mair FS, Fortin M, Guthrie B, Nunes BP, Miranda JJ, et al. Multimorbidity. Nat Rev Dis Primers 2022 Jul 14;8(1):48 [FREE Full text] [doi: 10.1038/s41572-022-00376-4] [Medline: 35835758]
12. Wang Z, Zhu D, Zhang H, Wang L, He H, Li Z, et al. Recommendations and quality of multimorbidity guidelines: a systematic review. Ageing Res Rev 2024 Dec;102:102559 [FREE Full text] [doi: 10.1016/j.arr.2024.102559] [Medline: 39532235]
13. Scherer M, Wagner HO, Lühmann D, Muche-Borowski C, Schäfer I, Dubben HH, et al. Multimorbidität. S3-Leitlinie. AWMF-Register-Nr. 053-047. DEGAM-Leitlinie Nr. 20. Berlin: Deutsche Gesellschaft für Allgemeinmedizin und Familienmedizin (DEGAM); 2017.
14. Mühlhäuser U, Götz K, Götz K, Weinmayr LM, Steinhäuser S. DEGAM-Leitlinie „Multimorbidität“ im Praxistest. Z Allg Med 2018;94:64-69 [FREE Full text] [doi: 10.3238/zfa.2018.0064-0069]
15. Schäfer I, Jahns V, Paucke V, Lühmann D, Scherer M, Nothacker J. Integrating general practitioners' and patients' perspectives in the development of a digital tool supporting primary care for older patients with multimorbidity: a focus group study. Front Digit Health 2025 Jan 21;7:1499333 [FREE Full text] [doi: 10.3389/fdgth.2025.1499333] [Medline: 39906064]
16. Campbell MK, Piaggio G, Elbourne DR, Altman DG, CONSORT Group. Consort 2010 statement: extension to cluster randomised trials. BMJ 2012 Sep 04;345(sep04 1):e5661. [doi: 10.1136/bmj.e5661] [Medline: 22951546]
17. Schäfer I, Kaduszkiewicz H, Wagner HO, Schön G, Scherer M, van den Bussche H. Reducing complexity: a visualisation of multimorbidity by combining disease clusters and triads. BMC Public Health 2014 Dec 16;14(1):1285 [FREE Full text] [doi: 10.1186/1471-2458-14-1285] [Medline: 25516155]
18. Lancaster GA, Dodd S, Williamson PR. Design and analysis of pilot studies: recommendations for good practice. J Eval Clin Pract 2004 May 04;10(2):307-312. [doi: 10.1111/j..2002.384.doc.x] [Medline: 15189396]
19. Teare MD, Dimairo M, Shephard N, Hayman A, Whitehead A, Walters SJ. Sample size requirements to estimate key design parameters from external pilot randomised controlled trials: a simulation study. Trials 2014 Jul 03;15(1):264 [FREE Full text] [doi: 10.1186/1745-6215-15-264] [Medline: 24993581]
20. Steinman MA, Hanlon JT. Managing medications in clinically complex elders: "there's got to be a happy medium". JAMA 2010 Oct 13;304(14):1592-1601 [FREE Full text] [doi: 10.1001/jama.2010.1482] [Medline: 20940385]

XSL•FO
RenderX

21. Chou R, Aronson N, Atkins D, Ismaila AS, Santaguida P, Smith DH, et al. Assessing harms when comparing medical interventions. In: Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Rockville, MD: Agency for Healthcare Research and Quality (US); 2008.
22. Hinz A, Kohlmann T, Stöbel-Richter Y, Zenger M, Brähler E. The quality of life questionnaire EQ-5D-5L: psychometric properties and normative values for the general German population. Qual Life Res 2014 Mar;23(2):443-447. [doi: 10.1007/s11136-013-0498-2] [Medline: 23921597]
23. Ludwig K, Graf von der Schulenburg JM, Greiner W. German Value Set for the EQ-5D-5L. Pharmacoeconomics 2018 Jun;36(6):663-674 [FREE Full text] [doi: 10.1007/s40273-018-0615-8] [Medline: 29460066]
24. Wensing M, Mainz J, Grol R. A standardised instrument for patient evaluations of general practice care in Europe. Eur J Gen Pract 2009 Jul 11;6(3):82-87. [doi: 10.3109/13814780009069953]
25. Schäfer I, Schulze J, Glassen K, Breckner A, Hansen H, Rakebrandt A, et al. Validation of patient- and GP-reported core sets of quality indicators for older adults with multimorbidity in primary care: results of the cross-sectional observational MULTIqual validation study. BMC Med 2023 Apr 17;21(1):148 [FREE Full text] [doi: 10.1186/s12916-023-02856-0] [Medline: 37069536]
26. Brauns H, Steinmann S. Educational reform in France, West-Germany and the United Kingdom: updating the CASMIN educational classification. ZUMA Nachrichten 1999;44:7-44.
27. Yesavage JA, Sheikh JI. 9/Geriatric Depression Scale (GDS). Clin Gerontol 2008 Oct 25;5(1-2):165-173. [doi: 10.1300/j018v05n01_09]
28. Stata statistical software: release 15.1. StataCorp LLC. College Station, TX; 2017. URL: https://www.stata.com/ [accessed 2025-10-30]
29. Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. Health Technol Assess 2004 Feb;8(6):iii-iv, 1-72 [FREE Full text] [doi: 10.3310/hta8060] [Medline: 14960256]
30. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. Lancet 2003 Oct 11;362(9391):1225-1230. [doi: 10.1016/S0140-6736(03)14546-1] [Medline: 14568747]
31. NHS Centre for Reviews and Dissemination. Getting evidence into practice. Effect Health Care 1999 Feb;5(1):1-16 [FREE Full text] [doi: 10.7748/ns.13.32.31.s58]
32. Crawshaw J, Meyer C, Antonopoulou V, Antony J, Grimshaw JM, Ivers N, et al. Identifying behaviour change techniques in 287 randomized controlled trials of audit and feedback interventions targeting practice change among healthcare professionals. Implement Sci 2023 Nov 21;18(1):63 [FREE Full text] [doi: 10.1186/s13012-023-01318-8] [Medline: 37990269]
33. Ambrosi E, Mezzalira E, Canzan F, Leardini C, Vita G, Marini G, et al. Effectiveness of digital health interventions for chronic conditions management in European primary care settings: systematic review and meta-analysis. Int J Med Inform 2025 Apr;196:105820 [FREE Full text] [doi: 10.1016/j.ijmedinf.2025.105820] [Medline: 39922064]
34. Shi X, He J, Lin M, Liu C, Yan B, Song H, et al. Comparative effectiveness of team-based care with and without a clinical decision support system for diabetes management. Ann Intern Med 2023 Jan;176(1):49-58. [doi: 10.7326/M22-1950]
35. Wiwatkunupakarn N, Aramrat C, Pliannuom S, Buawangpong N, Pinyopornpanish K, Nantsupawat N, et al. The integration of clinical decision support systems into telemedicine for patients with multimorbidity in primary care settings: scoping review. J Med Internet Res 2023 Jun 28;25:e45944 [FREE Full text] [doi: 10.2196/45944] [Medline: 37379066]
36. Prabhakaran D, Jha D, Prieto-Merino D, Roy A, Singh K, Ajay VS, et al. Effectiveness of an mHealth-based electronic decision support system for integrated management of chronic conditions in primary care: the mWellcare cluster-randomized controlled trial. Circulation 2019 Jan 15;139(3):380-391 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.118.038192] [Medline: 30586732]
37. Marcolino MS, Oliveira JA, Cimini CC, Maia JX, Pinto VS, Sá TQ, et al. Development and implementation of a decision support system to improve control of hypertension and diabetes in a resource-constrained area in Brazil: mixed methods study. J Med Internet Res 2021 Jan 11;23(1):e18872 [FREE Full text] [doi: 10.2196/18872] [Medline: 33427686]
38. Peleg M, Shahar Y, Quaglini S, Broens T, Budasu R, Fung N, et al. Assessment of a personalized and distributed patient guidance system. Int J Med Inform 2017 May;101:108-130. [doi: 10.1016/j.ijmedinf.2017.02.010] [Medline: 28347441]
39. Gencturk M, Laleci Erturkmen GB, Akpinar AE, Pournik O, Ahmad B, Arvanitis TN, et al. Transforming evidence-based clinical guidelines into implementable clinical decision support services: the CAREPATH study for multimorbidity management. Front Med (Lausanne) 2024 May 27;11:1386689 [FREE Full text] [doi: 10.3389/fmed.2024.1386689] [Medline: 38860204]
40. Pournik O, Ahmad B, Lim Choi Keung SN, Khan O, Despotou G, Consoli A, et al. CAREPATH: developing digital integrated care solutions for multimorbid patients with dementia. Stud Health Technol Inform 2022 Jun 29;295:487-490. [doi: 10.3233/SHTI220771] [Medline: 35773917]
41. Erturkmen GB, Gencturk M, Akpinar AE, Pournik O, Arvanitis TN, Steinhoff A, et al. The CAREPATH study for personalized management of multimorbidity in elderly with mild cognitive impairment and mild dementia. Med Res Arch 2024 Aug;12(8):1-9 [FREE Full text] [doi: 10.18103/mra.v12i8.5626]

XSL•FO
RenderX

42. Gregg EW, Holman N, Sophiea M, Misra S, Pearson-Stuttard J, Valabhji J, et al. Multiple long-term conditions as the next transition in the global diabetes epidemic. Commun Med (Lond) 2025 Mar 14;5(1):42 [FREE Full text] [doi: 10.1038/s43856-025-00742-9] [Medline: 39953177]
43. Kappelin C, Sandlund C, Westman J, Wachtler C. Dancing with the patient: a qualitative study of general practitioners' experiences of managing patients with multimorbidity and common mental health problems. BMC Prim Care 2023 Apr 20;24(1):104 [FREE Full text] [doi: 10.1186/s12875-023-02056-y] [Medline: 37081385]
44. Nissen NK, Aarhus R, Ørtenblad L. Dynamics of a specialized and complex health care system: exploring general practitioners' management of multimorbidity. Chronic Illn 2022 Mar;18(1):155-168. [doi: 10.1177/1742395320928403] [Medline: 32498609]
45. Sinnott C, Mc Hugh S, Browne J, Bradley C. GPs' perspectives on the management of patients with multimorbidity: systematic review and synthesis of qualitative research. BMJ Open 2013 Sep 13;3(9):e003610 [FREE Full text] [doi: 10.1136/bmjopen-2013-003610] [Medline: 24038011]
46. Zhou X, Han L, Farmer A, Yao M, Xia Y, Yan M, et al. Challenges and barriers to physician decision-making for prescribing and deprescribing among patients with multimorbidity in eastern China's primary care settings: a qualitative study. BMJ Open 2025 Mar 08;15(2):e095063 [FREE Full text] [doi: 10.1136/bmjopen-2024-095063] [Medline: 39922587]
47. Smith SM, Wallace E, O'Dowd T, Fortin M. Interventions for improving outcomes in patients with multimorbidity in primary care and community settings. Cochrane Database Syst Rev 2021 Jan 15;1(1):CD006560 [FREE Full text] [doi: 10.1002/14651858.CD006560.pub4] [Medline: 33448337]
48. Zhang Y, Stokes J, Anselmi L, Bower P, Xu J. Can integrated care interventions strengthen primary care and improve outcomes for patients with chronic diseases? A systematic review and meta-analysis. Health Res Policy Syst 2025 Jan 06;23(1):5 [FREE Full text] [doi: 10.1186/s12961-024-01260-1] [Medline: 39762867]
49. Nothacker J, Butz S, Lühmann D, Duwe P, van den Akker M, Thiem U, et al. General practitioner-based interventions to reduce hospital admissions in patients with multimorbidity living at home - a rapid review. Z Evid Fortbild Qual Gesundhwes 2025 May;194:74-85 [FREE Full text] [doi: 10.1016/j.zefq.2025.01.005] [Medline: 40021381]
50. Mangin D, Parascandalo J, Khudoyarova O, Agarwal G, Bismah V, Orr S. Multimorbidity, eHealth and implications for equity: a cross-sectional survey of patient perspectives on eHealth. BMJ Open 2019 Mar 12;9(2):e023731 [FREE Full text] [doi: 10.1136/bmjopen-2018-023731] [Medline: 30760515]
51. Buawangpong N, Pinyopornpanish K, Pliannuom S, Nantsupawat N, Wiwatkunupakarn N, Angkurawaranon C, et al. Designing telemedicine for older adults with multimorbidity: content analysis study. JMIR Aging 2024 Jan 10;7:e52031 [FREE Full text] [doi: 10.2196/52031] [Medline: 38198201]
52. Endalamaw A, Zewdie A, Wolka E, Assefa Y. A scoping review of digital health technologies in multimorbidity management: mechanisms, outcomes, challenges, and strategies. BMC Health Serv Res 2025 Mar 15;25(1):382. [doi: 10.1186/s12913-025-12548-5] [Medline: 40089752]
53. Smith SM, Wallace E, Salisbury C, Sasseville M, Bayliss E, Fortin M. A core outcome set for multimorbidity research (COSmm). Ann Fam Med 2018 Mar;16(2):132-138 [FREE Full text] [doi: 10.1370/afm.2178] [Medline: 29531104]
54. Schnitzer S, Balke K, Walter A, Litschel A, Kuhlmey A. [Do gatekeeping programs increase equality of health care in Germany? A comparison of the health care situation of participants and nonparticipants]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2011 Aug 29;54(8):942-950. [doi: 10.1007/s00103-011-1317-y] [Medline: 21800242]
55. Laux G, Wensing M, Altiner A, Leutgeb R. Continuity of care in a pandemic: an observational study in GP-centred healthcare in Germany. BMJ Open Qual 2025 Mar 03;14(1):e002944 [FREE Full text] [doi: 10.1136/bmjoq-2024-002944] [Medline: 40032595]
56. Barker T, Stone JC, Sears K, Klugar M, Tufanaru C, Leonardi-Bee J, et al. The revised JBI critical appraisal tool for the assessment of risk of bias for randomized controlled trials. JBI Evid Synth 2023 Mar 01;21(3):494-506. [doi: 10.11124/JBIES-22-00430] [Medline: 36727247]
57. Vanderweele TJ, Arah OA. Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments, and confounders. Epidemiology 2011 Jan;22(1):42-52 [FREE Full text] [doi: 10.1097/EDE.0b013e3181f74493] [Medline: 21052008]
58. Andrade C. Intent-to-treat (ITT) vs completer or per-protocol analysis in randomized controlled trials. Indian J Psychol Med 2022 Jul 21;44(4):416-418 [FREE Full text] [doi: 10.1177/02537176221101996] [Medline: 35949630]
59. DiMatteo MR. Variations in patients' adherence to medical recommendations: a quantitative review of 50 years of research. Med Care 2004 Mar;42(3):200-209. [doi: 10.1097/01.mlr.0000114908.90348.f9] [Medline: 15076819]
60. Agbla SC, DiazOrdaz K. Reporting non-adherence in cluster randomised trials: a systematic review. Clin Trials 2018 Jun 02;15(3):294-304 [FREE Full text] [doi: 10.1177/1740774518761666] [Medline: 29608096]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
CPG: clinical practice guideline

XSL•FO
RenderX

GP: general practitioner
IRR: incidence rate ratio

Edited by J Wang; submitted 27.06.25; peer-reviewed by AA Klein, J Fernandez-Saez; comments to author 11.08.25; revised version received 25.08.25; accepted 26.09.25; published 20.11.25.

Please cite as:
Nothacker J, Paucke V, Lezius S, Zapf A, Lühmann D, Scherer M, Schäfer I
Implementation of the German Clinical Practice Guideline for Multimorbidity Using a Digital Tool in Primary Care: Pilot Cluster Randomized Clinical Trial
JMIR Aging 2025;8:e79767
URL: https://aging.jmir.org/2025/1/e79767
doi:10.2196/79767
PMID:

©Julia Nothacker, Valentina Paucke, Susanne Lezius, Antonia Zapf, Dagmar Lühmann, Martin Scherer, Ingmar Schäfer. Originally published in JMIR Aging (https://aging.jmir.org), 20.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Associations Between Social Media Use and Anxiety and Depression Among Older Adults : Cross-Sectional Study

Jiaoling Huang[1*], PhD; Zhenxing Ge[1*], MS; Yijing Chu[1*], BS; Yuge Yan[1], BS; Wei Zhang[1], BS; Hong Liang[2,3], PhD; Yuqi Yang[3], MS; Hui Wang[1], PhD

[1]School of Public Health, Shanghai Jiao Tong University School of Medicine, 227 Chongqing South Road, Shanghai, China

[2]School of Social Development and Public Policy, Fudan University, Shanghai, China

[3]Shanghai Institute of Infectious Disease and Biosecurity, Shanghai, China

[*]these authors contributed equally

Corresponding Author:
Hui Wang, PhD
School of Public Health, Shanghai Jiao Tong University School of Medicine, 227 Chongqing South Road, Shanghai, China

Abstract

Background: Social media engagement among older adults has surged worldwide, with China's older users exceeding 120 million in 2023. However, research remains disproportionately focused on youth. Critically, the dose-response relationship between use intensity and mental health in this population is poorly quantified, especially in rapidly aging societies such as China, where 23% of the population will be aged ≥65 years by 2035.

Objective: This study aimed to outline the social media use status among retired older adults and explore the association between social media use, including time spent on social media and social media addiction, and mental health status.

Methods: A cross-sectional survey was conducted in Shanghai, China, in 2024. A total of 15,986 retired participants were recruited via universities for older adults and primary health care institutions. Short versions of anxiety (the 2-item Generalized Anxiety Disorder scale) and depression (the 2-item Patient Health Questionnaire) scales were used to minimize the required time to complete the questionnaires for older adults. Logistic regressions were used to examine the associations between social media use and mental health after controlling for covariates. Subgroup analysis was conducted considering sex, age, marital status, urbanicity, and socioeconomic status.

Results: The participants had an average age of 68.49 (SD 7.6) years, with most (13,854/15,986, 86.7%) being married and living with their spouse and approximately half (8155/15,986, 51.0%) being male. Our research indicated that over 98% of retired older individuals (15,807/15,986, 98.88%) had used social media, with WeChat, Douyin, and Kuaishou being the most common platforms. Among them, 52.3% (8361/15,986) spent 2 to 3 hours a day on social media, 32.29% (5162/15,986) spent >4 hours a day, and 20.34% (3253/15,986) were addicted to social media. Older adults with ≥6 hours of daily social media use time exhibited higher rates of anxiety (odds ratio [OR] 1.44, 95% CI 1.20-1.72; *P*<.001) and depression (OR 1.50, 95% CI 1.25-1.79; *P*<.001) compared with those who used social media for ≤1 hour per day. Older adults addicted to social media had higher odds of anxiety (OR 2.81, 95% CI 2.57-3.08; *P*<.001) and depression (OR 2.51, 95% CI 2.30-2.75; *P*<.001). Subgroup analyses revealed stronger associations for women, people aged 49-75 years, those with a lower educational level and income, urban residents, and non–solo dwellers.

Conclusions: Retired older adults in Shanghai are an active group of social media users. Using social media for over 6 hours a day and social media addiction were significantly associated with anxiety and depression. Future social media research should pay more attention to older adults and explore these longitudinal relationships.

(JMIR Aging 2025;8:e71712) doi:10.2196/71712

KEYWORDS

social media use; addiction; older adults; mental health; anxiety; depression

Introduction

Mental disorders are increasingly recognized as leading causes of disease burden, with anxiety and depression being the most widespread [1]. The Global Burden of Disease Study (GBD) 2019 Mental Disorders Collaborators reported that, between 1990 and 2019, the global number of disability-adjusted life years due to mental disorders increased from 80.8 million to 125.3 million [2]. In addition to focusing on children and adolescents [3], current studies have focused extensively on older adults due to the recent COVID-19 pandemic. Existing studies suggest that long-term isolation, loneliness, fear of death,

XSL•FO
RenderX

and other factors have exacerbated the incidence of mental disorders among older adults [4,5]. At present, we do not have a clear picture of the global prevalence of mental disorders among older adults after the COVID-19 pandemic. However, the latest GBD data show that approximately 14% of adults aged ≥60 years lived with a mental disorder in 2021 [6], which might be underestimated as there has been an ongoing debate about the classification of mental disorders in the GBD and documented diagnostic gaps in older populations [7-9].

Due to the widespread use of electronic devices such as smartphones and tablets, social media is being used in China on an unprecedented scale. By the end of 2023, the number of internet users had reached 1.092 billion, whereas the number of smartphone owners had risen to 974.6 million [10,11]. Critically, older adults represent China's fastest-growing internet cohort, with users aged ≥60 years increasing by 300% since 2019 [12]. In China, social media platforms such as WeChat, Weibo, Douyin, and Xiaohongshu have gained widespread popularity across the country. The 53rd Statistical Report on the Development of the Internet in China by the China Internet Network Information Center shows that, as of the end of 2023, the size of the internet user group in China aged ≥60 years has reached 170 million (accounting for 57% of the older adult population), with the number of older adults spending more than 10 hours online per day exceeding 100,000 [13]. With the ongoing aging of China's population, the number of older individuals using social media is increasing annually. At present, although we lack specific data on internet use differences among various older adult age groups, current studies suggest that active older adults, particularly those who have recently retired, are likely to be the primary users of social media [14].

It would be quite interesting to observe how social media use affects the mental health of older adults in China. However, most published studies have focused on social media use among children, students, adolescents, and young adults, and older populations remain relatively understudied compared to these younger groups [15-17]. Although recent years have seen growing research interest in older adults' digital engagement, several research gaps remain. First, much of the current literature still assumes that older adults are largely averse to using the internet, thereby potentially underestimating their actual engagement with social media today [12]. However, up-to-date surveys on older adults' social media use are lacking, leaving us without solid evidence of their current patterns of engagement, particularly among active retirees. Second, current research suggests that online social networking has a positive impact on older adults, but recent systematic reviews highlight contradictory findings due to heterogeneous methodologies [18,19]. Reviewing these studies, we found that their measurement of social media use often relied on a simple binary variable—whether the respondent used social media at all. In measuring social media use among older adults, aspects such as frequency, intensity, and even potential addiction should not be overlooked. Third, other research findings in this area are inconclusive, especially because of a relative lack of research conducted on how sex, socioeconomic status (SES), and digital literacy mediate these effects [20]. Thus, this study focused on the vibrant older adult demographic that is most engaged with social media. We aimed to scrutinize the correlation between their social media use (use duration and addiction status) and mental well-being while also examining how this relationship varies among different sexes, age brackets, and SESs.

Methods

Participants and Study Design

This cross-sectional study was conducted in Shanghai, a typical Chinese mega city, in 2024, aiming to assess social media use and its association with mental well-being among retired older adults. This study was conducted in 2 phases using convenience sampling. In the first phase, the questionnaire was distributed through universities for older adults in Shanghai. In the second phase, after conducting a preliminary analysis of the questionnaires collected in the first stage, the areas with a smaller number of collected questionnaires were identified. Questionnaires were distributed through primary health care institutions, namely, community health service centers (CHSCs), to areas that either were not initially covered or had a smaller number of responses. All the questionnaires were distributed online. The inclusion criteria were as follows: (1) being retired, (2) no terminal illnesses, (3) ability to speak and read Chinese, and (4) no severe concurrent psychological or psychiatric diseases or other physical diseases that impact physical activities or mental health (self-report diagnosis). The exclusion criteria were as follows: (1) severe cognitive impairment, (2) active psychiatric disorders, and (3) inability to use a smartphone independently. A total of 15,986 valid online questionnaires were collected, and the questionnaire response rate was 99.95% (15,986/15,994).

Criteria for the Anxiety and Depression Scales

Short versions of anxiety and depression scales were used to minimize the time required to complete the questionnaires for older adults. The 2-item Generalized Anxiety Disorder scale (GAD-2) and the 2-item Patient Health Questionnaire (PHQ-2) were administered. The GAD-2 and PHQ-2 each consist of 2 questions, with each question scored from 0 to 3. The total possible score for each scale is 6 points, with a score of 2 or higher indicating clinically significant symptoms warranting further diagnostic evaluation [21,22]. While we acknowledge that they have a lower sensitivity than that of the full scales (the 9-item Patient Health Questionnaire and 7-item Generalized Anxiety Disorder scale), these instruments demonstrate adequate validity for population-level screening in older adults (area under the curve=0.78-0.84), particularly given the optimized acceptability in older cohorts with cognitive fatigue concerns [23-25].

Social Media Use

Social media use was represented using 2 variables: time spent on social media and social media addiction. Time spent on social media was measured using the self-reported amount of time spent on social media each day, categorized into 4 groups: ≤1 hour, 2 to 3 hours, 4 to 5 hours, and ≥6 hours. Social media addiction was assessed using the Bergen Social Media Addiction Scale, with a cutoff score of 19. A score above 19 indicates addiction to social media [26].

XSL•FO
RenderX

Covariates

The covariates in this study included sex (male=1; female=2), age, educational level (senior high school and lower=1; junior college=2; university and higher=3), marital status (married and living with spouse=1; married but not living with spouse=2; divorced or widowed=3; unmarried=4), urban or rural residence, living status (living alone=1; not living alone=2), and monthly income (≤¥6000 [US $842.05]=1; ¥6001‐¥9000 [US $842.19-$1263.07]=2; ≥¥9001 [US $1263.21]=3) of the active older individuals. Health behavior–related variables encompassed physical activity (regular exercise=1; no regular exercise=2), sedentary hours (<0.5=1; 0.5‐4=2; >4=3), sleep duration (≤6 hours=1; 7‐8 hours=2; ≥9 hours=3), and social engagement levels (hardly ever socialize=1; sometimes socialize=2; often socialize=3). All the covariates were based on the self-reports of the research subjects.

Statistical Analyses

For the descriptive statistics, we used the mean (with SE) for continuous variables and the number (with weighted percentage) for categorical variables. To identify association factors for anxiety and depression in older adults, we conducted logistic regression to calculate odds ratios (ORs) with 95% CIs. Adjusted ORs were obtained through multivariate regression analyses, using selection methods to account for potential confounding variables. Subgroup analyses were conducted to explore the associations between social media use and mental health among individuals of different sexes, ages, educational levels, marital statuses, monthly incomes, urban and rural residences, and living conditions. Database management and all statistical analyses were conducted using the *Survey* package in R (version 4.3.1; R Foundation for Statistical Computing) and ArcMap (version 10.8; Esri), which are software for statistical computing and plotting, respectively. A 2-sided *P* value of <.05 was considered statistically significant.

Ethical Considerations

This study was approved by the Ethics Committee of Public Health and Nursing Research, Shanghai Jiao Tong University School of Medicine (approval: SJUPN-2024-041-KS1). Written informed consent was obtained from all participants, with procedures for illiterate subjects involving witnessed verbal consent documented using institutional review board–approved forms. For secondary data analyses, original consent permitted anonymized reuse. All data were deidentified through replacement of personal identifiers with unique codes and aggregated for analysis to ensure confidentiality. There was no charge for participating in the study, and no compensation was provided to the participants. No identifiable images or personal information appear in this manuscript or the supplementary materials. This study was conducted in accordance with the Declaration of Helsinki.

Results

Descriptive Analysis

Table 1 shows the descriptive statistics of the sample. The average age of the respondents was 68.49 (SD 7.6) years. Men constituted 51.0% (8155/15,986) of the sample. A significant portion (5633/15,986, 35.2%) had received education at the university level or higher. Most respondents (13,854/15,986, 86.7%) were married and living with their spouse. A substantial number (9669/15,986, 60.5%) had a monthly income exceeding ¥9000 (US $1263.07), and 67.8% (10,833/15,986) resided in urban areas. The prevalence rates of anxiety and depression were 20.1% (3218/15,986) and 21.7% (3467/15,986), respectively. Regarding daily social media use time, 15.4% (2463/15,986) spent no more than 1 hour, 52.3% (8361/15,986) spent between 2 and 3 hours, 25.9% (4134/15,986) spent between 4 and 5 hours, and 6.4% (1028/15,986) spent more than 6 hours. The prevalence of social media addiction was 20.3% (3253/15,986). The social media platforms most used by older adults were WeChat, Douyin, and Kuaishou (Multimedia Appendix 1), and the most viewed type of content was current news and health-related information (Multimedia Appendix 1).

XSL•FO
RenderX

Table . Baseline sociodemographic data (N=15,986).

Characteristic	Values
Sex, n (%)	
Male	8155 (51.0)
Female	7831 (49.0)
Age (y), mean (SD)	68.49 (7.6)
Educational level, n (%)	
Senior high school and lower	3948 (24.7)
Junior college	6405 (40.1)
University and higher	5633 (35.2)
Marital status, n (%)	
Married and living with spouse	13,854 (86.7)
Married but not living with spouse	356 (2.2)
Divorced or widowed	1675 (10.5)
Unmarried	101 (0.6)
Monthly income, n (%)	
≤¥6000 (US $842.05)	2922 (18.3)
¥6001-¥9000 (US $842.19-$1263.07)	3395 (21.2)
≥¥9001 (US $1263.21)	9669 (60.5)
Urban or rural residence, n (%)	
Urban	10,833 (67.8)
Rural	5153 (32.2)
Social media use (hours per day), n (%)	
≤1	2463 (15.4)
2-3	8361 (52.3)
4-5	4134 (25.9)
≥6	1028 (6.4)
Social media addiction, n (%)	
No	12,733 (79.7)
Yes	3253 (20.3)
Anxiety, n (%)	
No	12,768 (79.9)
Yes	3218 (20.1)
Depression, n (%)	
No	12,519 (78.3)
Yes	3467 (21.7)

Logistic Analyses

Among the older adults with longer social media use times, a higher proportion exhibited higher scores for anxiety and depression (284/1028, 27.6% and 304/1028, 29.6%, respectively). Compared to those who used social media for 1 hour or less per day, those who used social media for 6 hours or more per day had 1.44 times the odds of anxiety (P<.001) and 1.50 times the odds of depression (P<.001). Among the older adults with social media addiction, a higher proportion had higher scores for anxiety and depression (1140/3,253, 35.0% and 1146/3253, 35.2%, respectively). The odds of anxiety were 2.81 times higher in those with addiction versus those without (P<.001), whereas their odds of depression were 2.51 times higher (P<.001; Table 2)

XSL•FO
RenderX

Table . Associations between social media use and mental health (N=15,986).

Characteristic	Participants, n (%)		Unadjusted analysis		Adjusted analysis[a]	
	No anxiety	Anxiety	OR[b] (95% CI)	*P* value	OR (95% CI)	*P* value
Anxiety						
Social media use time (h per d)						
≤1 (n=2463)	2014 (81.8)	449 (18.2)	—[c]	—	—	—
2-3 (n=8361)	6728 (80.5)	1633 (19.5)	1.09 (0.97-1.22)	.15	1.16 (1.03-1.31)	.01
4-5 (n=4134)	3282 (79.4)	852 (20.6)	1.16 (1.03-1.32)	.02	1.17 (1.02-1.34)	.02
≥6 (n=1028)	744 (72.4)	284 (27.6)	1.71 (1.44-2.03)	<.001	1.44 (1.20-1.72)	<.001
Social media addiction						
No (n=12,733)	10,655 (83.7)	2078 (16.3)	—	—	—	—
Yes (n=3253)	2113 (65.0)	1140 (35.0)	2.77 (2.54-3.01)	<.001	2.81 (2.57-3.08)	<.001
Depression						
Social media use time (hours per day)						
≤1 (n=2463)	1988 (80.7)	475 (19.3)	—	—	—	—
2-3 (n=8361)	6624 (79.2)	1737 (20.8)	1.10 (0.98-1.23)	.11	1.20 (1.07-1.35)	.002
4-5 (n=4134)	3183 (77.0)	951 (23.0)	1.25 (1.11-1.41)	<.001	1.29 (1.14-1.48)	<.001
≥6 (n=1028)	724 (70.4)	304 (29.6)	1.76 (1.49-2.08)	<.001	1.50 (1.25-1.79)	<.001
Social media addiction						
No (n=12,733)	10,412 (81.8)	2321 (18.2)	—	—	—	—
Yes (n=3253)	2107 (64.8)	1146 (35.2)	2.44 (2.24-2.75)	<.001	2.51 (2.30-2.75)	<.001

[a]Adjusted model: age, educational level, marital status, urban or rural residence, living status, monthly income, physical activity, sedentary time, sleep duration, and social engagement levels.

[b]OR: odds ratio.

[c]Reference group.

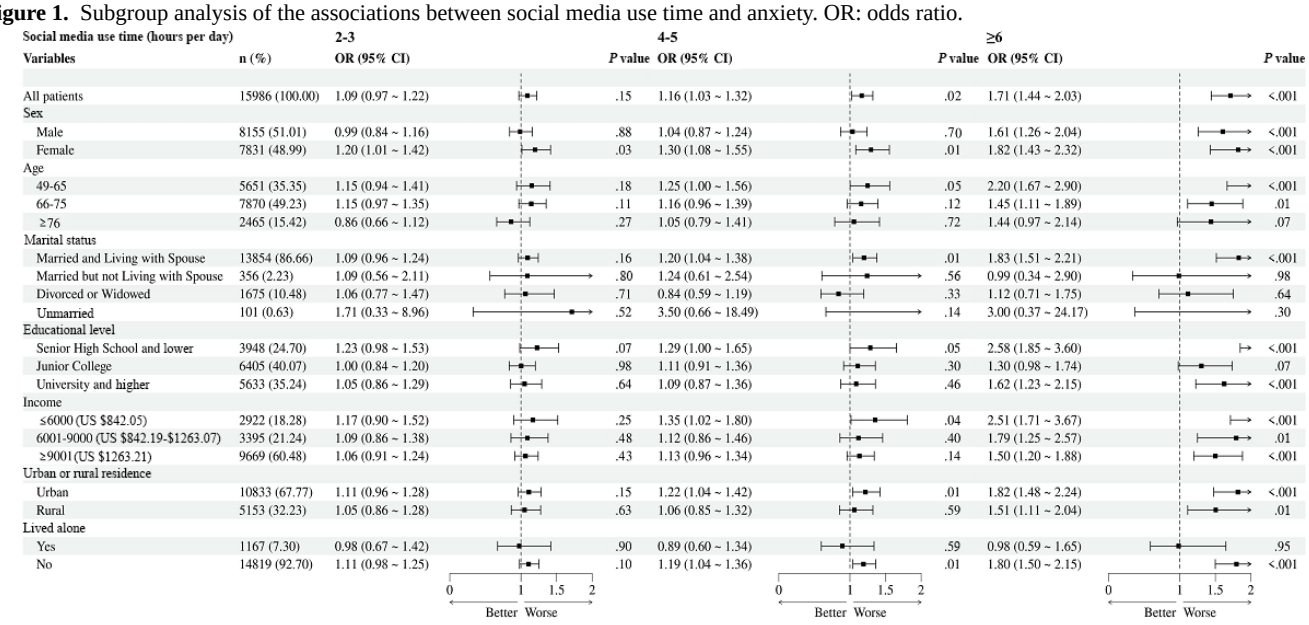
Subgroup Analyses

Associations Between Time Spent on Social Media and Mental Health

In the subgroup analysis of social media use time and anxiety (Figure 1), among the older adults who used social media for 6 or more hours per day, the odds of anxiety were higher for women (OR 1.82, 95% CI 1.43-2.32 vs OR 1.61, 95% CI 1.26-2.04; Table S1 in Multimedia Appendix 1), those aged 49 to 65 years (OR 2.20, 95% CI 1.67-2.90 vs OR 1.45, 95% CI 1.11-1.89; Table S2 in Multimedia Appendix 1), individuals who were married and living with their spouse (OR 1.83, 95% CI 1.51-2.21 vs OR 0.99, 95% CI 0.34-2.90; Table S3 in Multimedia Appendix 1), individuals with lower levels of education (OR 2.58, 95% CI 1.85-3.60 vs OR 1.62, 95% CI 1.23-2.15; Table S4 in Multimedia Appendix 1), those with a lower monthly income (OR 2.51, 95% CI 1.781 3.67 vs OR 1.79, 95% CI 1.25-2.57; Table S5 in Multimedia Appendix 1), those residing in urban areas (OR 1.82, 95% CI 1.48-2.24 vs OR 1.51, 1.11-2.04; Table S6 in Multimedia Appendix 1), and those not living alone (OR 1.80, 95% CI 1.50-2.15; Table S7 in Multimedia Appendix 1).

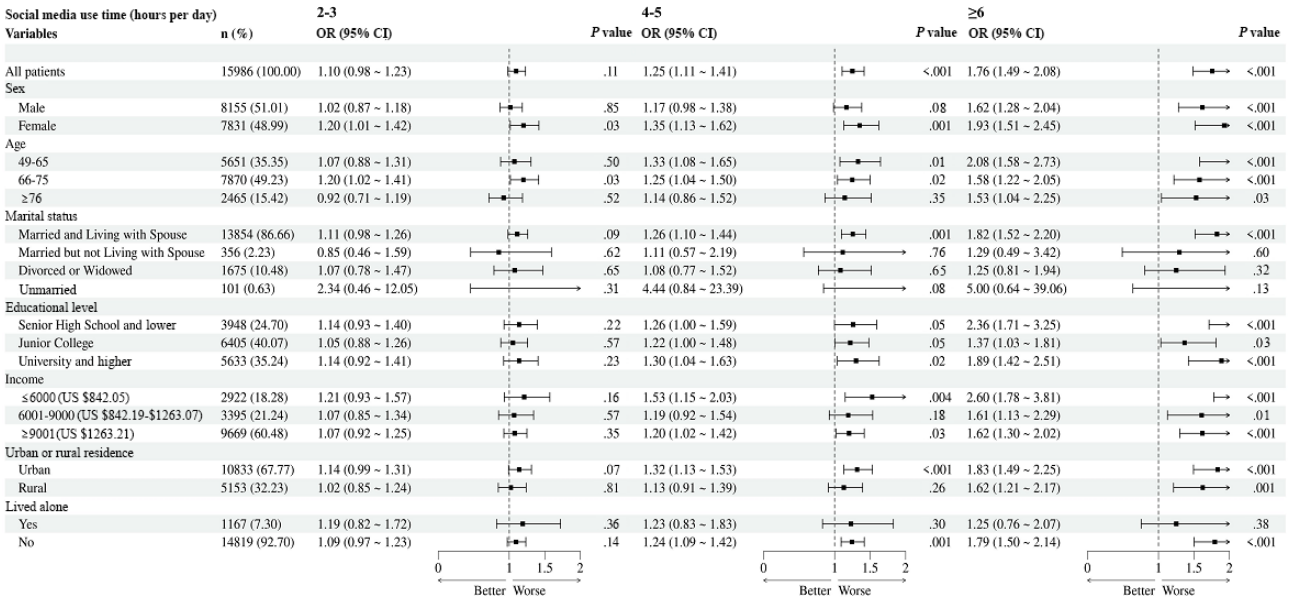
XSL•FO
RenderX

Figure 1. Subgroup analysis of the associations between social media use time and anxiety. OR: odds ratio.



In the subgroup analysis of social media use time and depression (Figure 2), among the older adults who used social media for 6 or more hours per day, the odds of depression were higher for women (OR 1.93, 95% CI 1.51-2.45 vs OR 1.62, 95% CI 1.28-2.04; Table S1 in Multimedia Appendix 1), those aged 49 to 65 years (OR 2.08, 95% CI 1.58-2.73 vs OR 1.58, 95% CI 1.22-2.05; Table S2 in Multimedia Appendix 1), individuals who were married and living with their spouse (OR 1.82, 95% CI 1.52-2.20 vs OR 1.25, 95% CI 0.81-1.94; Table S3 in Multimedia Appendix 1), individuals with lower levels of education (OR 2.36, 95% CI 1.71-3.25 vs OR 1.89, 95% CI 1.42-2.51; Table S4 in Multimedia Appendix 1), those with a lower monthly income (OR 2.60, 95% CI 1.78-3.81 vs OR 1.62, 95% CI 1.30-2.02; Table S5 in Multimedia Appendix 1), those residing in urban areas (OR 1.83, 95% CI 1.49-2.25 vs OR 1.62, 95% CI 1.21-2.17; Table S6 in Multimedia Appendix 1), and those not living alone (OR 1.79, 95% CI 1.50-2.14; Table S7 in Multimedia Appendix 1).

Figure 2. Subgroup analysis of the associations between social media use time and depression. OR: odds ratio.



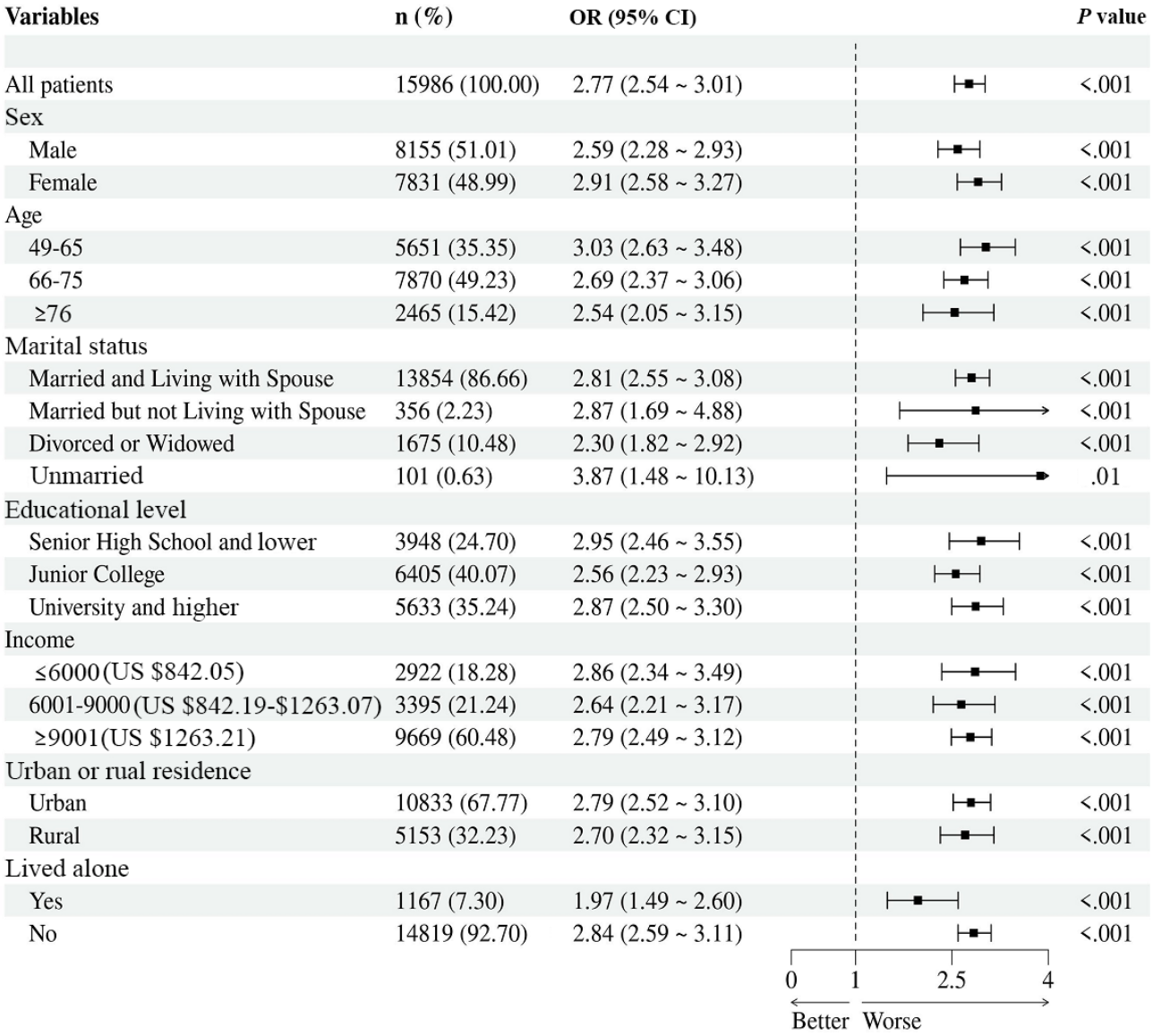
Associations Between Social Media Addiction and Mental Health

In the subgroup analysis of social media addiction and anxiety (Figure 3), among the older adults who were addicted to social media, the odds of anxiety were higher for women (OR 2.91, 95% CI 2.58-3.27 vs OR 2.59, 95% CI 2.28-2.93; Table S1 in Multimedia Appendix 1), those aged 49 to 65 years (OR 3.03, 95% CI 2.63-3.28 vs OR 2.69, 95% CI 2.37-3.06; Table S2 in Multimedia Appendix 1), individuals who were single (unmarried; OR 3.87, 95% CI 1.48-10.13 vs OR 2.87, 95% CI 1.69-4.88; Table S3 in Multimedia Appendix 1), those with lower levels of education (OR 2.95, 95% CI 2.46-3.55 vs OR 2.87, 95% CI 2.50-3.30; Table S4 in Multimedia Appendix 1), those with a lower monthly income (OR 2.86, 95% CI 2.34-3.49 vs OR 2.79, 95% CI 2.49-3.12; Table S5 in Multimedia Appendix 1), those residing in urban areas (OR 2.79, 95% CI 2.52-3.10 vs OR 2.70, 95% CI 2.32-3.15; Table S6 in

XSL•FO
RenderX

Multimedia Appendix 1), and those not living alone (OR 2.84, 95% CI 2.59-3.11 vs OR 1.97, 95% CI 1.49-2.60; Table S7 in Multimedia Appendix 1).

Figure 3. Subgroup analysis of the associations between social media addiction and anxiety. OR: odds ratio.

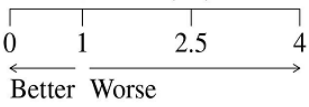


Social media addiction posed slightly higher odds of anxiety. In the subgroup analysis of depression and social media addiction (Figure 4), among the older adults who were addicted to social media, the odds of anxiety were higher for women (OR 2.61, 5%CI 2.32-2.94 vs OR 2.27, 95% CI 2.01-2.57; Table S1 in Multimedia Appendix 1), those aged 49 to 65 years (OR 2.68, 95% CI 2.33-3.08 vs OR 2.45, 95% CI 1.99-3.02; Table S2 in Multimedia Appendix 1), individuals who were single (unmarried; OR 2.61, 95% CI 1.02-6.66 vs OR 2.50, 95% CI 2.28-2.75; Table S3 in Multimedia Appendix 1), those with higher levels of education (OR 2.65, 95% CI 2.30-3.04 vs OR 2.41, 95% CI 2.11-2.76; Table S4 in Multimedia Appendix 1), those with a higher monthly income (OR 2.49, 95% CI 2.23-2.79 vs OR 2.46, 95% CI 2.02-3.00; Table S5 in Multimedia Appendix 1), those residing in urban areas (OR 2.46, 95% CI 2.22-2.72 vs OR 2.40, 95% CI 2.06-2.79; Table S6 in Multimedia Appendix 1), and those not living alone (OR 2.48, 95% CI 2.26-2.71 vs OR 1.90, 95% CI 1.45-2.49; Table S7 in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 4. Subgroup analysis of the associations between social media addiction and depression. OR: odds ratio.

Variables	n (%)	OR (95% CI)	*P* value
All patients	15986 (100.00)	2.44 (2.24 ~ 2.66)	<.001
Sex			
Male	8155 (51.01)	2.27 (2.01 ~ 2.57)	<.001
Female	7831 (48.99)	2.61 (2.32 ~ 2.94)	<.001
Age			
49-65	5651 (35.35)	2.68 (2.33 ~ 3.08)	<.001
66-75	7870 (49.23)	2.32 (2.05 ~ 2.64)	<.001
≥76	2465 (15.42)	2.45 (1.99 ~ 3.02)	<.001
Marital status			
Married and Living with Spouse	13854 (86.66)	2.50 (2.28 ~ 2.75)	<.001
Married but not Living with Spouse	356 (2.23)	2.21 (1.31 ~ 3.73)	.003
Divorced or Widowed	1675 (10.48)	1.97 (1.56 ~ 2.48)	<.001
Unmarried	101 (0.63)	2.61 (1.02 ~ 6.66)	.05
Educational level			
Senior High School and lower	3948 (24.70)	2.23 (1.87 ~ 2.68)	<.001
Junior College	6405 (40.07)	2.41 (2.11 ~ 2.76)	<.001
University and higher	5633 (35.24)	2.65 (2.30 ~ 3.04)	<.001
Income			
≤6000 (US $842.05)	2922 (18.28)	2.46 (2.02 ~ 3.00)	<.001
6001-9000 (US $842.19-$1263.07)	3395 (21.24)	2.29 (1.92 ~ 2.74)	<.001
≥9001(US $1263.21)	9669 (60.48)	2.49 (2.23 ~ 2.79)	<.001
Urban or rural residence			
Urban	10833 (67.77)	2.46 (2.22 ~ 2.72)	<.001
Rural	5153 (32.23)	2.40 (2.06 ~ 2.79)	<.001
Lived alone			
Yes	1167 (7.30)	1.90 (1.45 ~ 2.49)	<.001
No	14819 (92.70)	2.48 (2.26 ~ 2.71)	<.001



Discussion

Principal Findings

We observed that, in Shanghai, a typical Chinese megacity, older adults seem to have entered a new digital era. Our research found that over 98% of retired older adults in Shanghai (15,807/15,986, 98.8%) had ever used social media, particularly platforms such as WeChat, Douyin, and Kuaishou, with over 30% of them (5162/15,986, 32.3%) using it for more than 4 hours a day and over 20% (3253/15,986, 20.4%) being addicted to it. The frequency of social media use and the addiction rate among older adults in Shanghai have reached or even surpassed those of young people. Riehm et al [27] reported that approximately 80% of adolescents use social media for no more than 3 hours. In a study conducted by Eichenberg et al [28], approximately 20% of students were identified as being addicted to social media. On buses or subways in China, an increasing number of older adults being “phubbers” (people who are constantly looking at their phones) can be observed, mainly browsing short videos and social media feeds such as Moments on WeChat. Current social media platforms offer a variety of social features for older adults, including adding new friends and creating social circles, having voice and video conversations, sharing and commenting on daily life, and browsing and interacting with short videos. However, Jin et al [29] pointed out that more older adults are using social media such as WeChat and short videos, which can affect their daily life and mental health, possibly leading to addiction.

In our study, a positive correlation was found indicating that longer durations of social media use were associated with increased anxiety. Specifically, we found that social media use duration became highly associated with anxiety once it exceeded 6 hours, and such difference was not observed among those who used social media for less than 1 hour, 2 to 3 hours, or 4 to 5 hours. Similarly, the relationship between time spent on social media and depression was also positive and more prominent than that for anxiety. Existing research, which is limited, has provided contradictory evidence. Yan et al [30] investigated the association between internet exclusion and depressive symptoms among older adults in 32 countries and found that internet use was negatively associated with depressive

XSL•FO
RenderX

symptoms among older adults. Due to limitations in public databases, the aforementioned study only examined the relationship between internet use (yes or no) and depression among older adults, not the duration of social media use or social media addiction. In contrast, our analytic framework aligns with a growing body of evidence that documents a positive association between social media use frequency and intensity and depressive symptoms in later life [31,32]. Two factors appear to underlie this positive association among "young-old" retirees. First, active, healthy older adults increasingly replace offline social, physical, and community engagements with prolonged online activity, eroding face-to-face support networks and diminishing the mood-regulating benefits of in-person interaction. Second, older users predominantly engage in passive content consumption rather than active creation, resulting in frequent upward social comparisons and algorithm-amplified exposure to negative information—such as health-related rumors and sensationalized news [33]. These passive, comparison-driven experiences exacerbate feelings of inadequacy, anxiety, and hopelessness, thereby translating extended daily social media use into heightened depressive and anxiety symptomatology.

We attempted to explain the negative impact of social media on the mental health of older adults, and we infer that it may be related to their use habits. When older adults use social media, they often scroll and browse, such as watching short videos on platforms such as Douyin or scrolling through their Moments feed on the WeChat app, which is classified as passive social media use. According to Verduyn et al [34], passive social media use is defined as monitoring or observing content and behavior on social media platforms without direct exchanges with others (eg, scrolling and browsing), whereas active use involves more direct interactions such as posting, commenting, and sending private messages [35]. A substantial body of research has revealed a positive relationship between passive social media use and anxiety or depression among younger individuals [36]. Passive social media use is considered a cause of stress and anxiety, largely due to upward social comparison triggered by prevailing positive self-presentations [37,34]. In addition, the older adult group involved in this study was in the process of entering old age, with many having just started retirement or experiencing the empty nest stage. This cohort of older adults is unique in that they are part of the internet era, being familiar with smartphones and internet browsing. Spending a long time on social media tends to decrease their opportunities for face-to-face communication with family and friends, and passive use of social media further limits their chances to express themselves and engage in interactive communication. Further research should be conducted to delve into the pathways and mechanisms through which prolonged social media use among these younger members of the older population leads to anxiety and depression.

Additionally, our study provides clear evidence of a strong association between social media addiction and mental health issues among older adults. The odds of anxiety and depression for older individuals with social media addiction were 2.80 and 2.51 times higher, respectively, than for those without addiction (P<.001). This positive correlation has been consistently found in existing studies targeting adolescent populations [38,39]. Relatively few studies have focused on social media addiction among older adults; however, the results consistently show that social media addiction is highly correlated with mental health problems [40]. Some studies also suggest that mental health problems may lead to social media addiction among older adults [41]. Further exploration is needed to understand the causal relationship between the two.

Our study further identified that female, younger, married, and non–solo-dwelling older individuals are more likely to experience anxiety and depression when using social media. Sex and age factors have been partially explained in existing literature involving adolescents, such as female and younger individuals being more sensitive and active [40]. Özbek and Karaş [41] pointed out that sex roles may restrict women socially, so women may prefer to use social media to meet their social needs at home. Older individuals who are married and living with family or partners have also been identified (this was derived from 2 questions in the questionnaire. One was about marital status, and the other was about who one currently lives with, suggesting that, even when surrounded by others, they may still be at risk because social media can supplant direct human interaction, which in turn can be detrimental to their mental well-being [41]. In addition, we found that the association between social media use and mental well-being varied significantly by SES, with anxiety and depression being more prevalent among individuals with lower SES (lower income and educational level). Wang et al [42] identified area-level SES as a moderator between internet use and depressive symptoms, with an explanation that individuals residing in lower-SES regions were constrained by limited resources and social support, which could exacerbate the association between internet use and depressive symptoms. However, we also noticed that highly educated older adults who were addicted to social media were more likely to experience depression. We believe that the causal relationship requires further investigation. In addition, the mental health of older people in cities is more susceptible to the influence of social media, which is consistent with previous studies. Xu and Zhang [43] believe that although urban older people have easier access to the internet, information overload, online social comparison, and substitution of offline social interaction might be the underlying causes of higher levels of depression.

Drawing on our findings, we propose the following evidence-based policy recommendations. First, regulatory authorities should prioritize the development of age-friendly digital ecosystems by mandating algorithmic reforms that promote health-related content and enforce periodic use break reminders. This addresses our critical finding that older adults' social media engagement has reached levels comparable to those of younger users. Second, community health services should implement offline alternative programs such as subsidized "screen-free hour" activities and intergenerational interaction hubs to mitigate the mental health risks associated with prolonged screen time. Finally, resource allocation should be precisely targeted toward the high-risk subgroups identified in our stratified analyses—specifically, urban women aged 60 to 74 years, individuals with low SES, and non–solo

XSL•FO
RenderX

dwellers—through neighborhood-based digital literacy workshops, mental health screenings at CHSCs, and family-mediated device management initiatives. These measures aim to foster a healthier integration of the digital and physical aspects of life for China's aging population.

Limitations

Some limitations remain. First, although we targeted the older population, we mainly focused on the relatively younger and more vibrant segment of this demographic as they are the primary group of older social media users. However, we are not clear on the situation of more advanced-age older adults. Critically, we excluded working older people—omitting 8.9% of China's population aged >65 years, who may experience distinct digital stressors. Meanwhile, despite efforts to balance district-level representation through supplementary CHSC sampling, participants from these 2 sources differed in age and education profiles. Caution is warranted when generalizing the findings to isolated older adult populations beyond educational institutions. Second, self-reported data may incur recall bias, and the GAD-2 and PHQ-2 are screening tools, not clinical diagnoses. Our cutoff scores (≥3) may not generalize across subpopulations. Third, we merely identified the relationship between social media use and mental health, but the causality between the two remains unclear. A cohort study is recommended in the future, which could further clarify this causal relationship. Fourth, the links between social media use and mental health issues are not direct, and various factors such as prolonged sitting, sleep, and social support might play a mediating role, which needs further exploration. Fifth, despite our screening procedures (self-report diagnosis), residual confounding from undiagnosed clinical disorders may persist. Individuals with subclinical anxiety or depression might overuse social media and report poorer mental health, inflating the observed associations. Future studies should incorporate structured clinical interviews to better control for this confounder. Finally, exploratory subgroup analyses did not include formal interaction testing due to sample constraints.

Conclusions

We focused on retired older adults instead of on adolescents and explored the impact of social media use on their mental well-being, yielding some interesting findings. We found that older adults in Shanghai have a high frequency of social media use, with over 30% of them (15,807/15,986, 32.3%) spending more than 4 hours per day on social media and over 20% (3253/15,986, 20.4%) exhibiting signs of social media addiction. Significant associations between social media use, including time spent on and addiction to social media, and mental health, including anxiety and depression, were identified. Furthermore, subgroup analysis revealed that older individuals with lower SES who were female, younger, married, and not living alone were more prone to experiencing anxiety and depression through their engagement with social media.

Acknowledgments

This study was supported by grants from the National Natural Science Foundation of China (82030099) and the Shanghai Three-Year Action Plan for Public Health System Construction (GWVI-11.1-43 and GWVI-11.2-YQ54). The authors express their gratitude to the participants for their valuable contributions to the study, as well as to the current and former researchers and staff involved.

Data Availability

The datasets generated or analyzed during this study are not publicly available to protect participants' privacy but are available from the corresponding author on reasonable request.

Authors' Contributions

JH and HW conceptualized and designed the study. ZG, Y Yan, Y Yang, and WZ collected the data. YC conducted the statistical analyses. JH and YC drafted the manuscript. HL and HW contributed to methodology considerations. JH and HW reviewed the manuscript for intellectual content. JH and HW are the co–corresponding authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Word clouds of the most used social media platforms and most viewed content, plus full logistic regression and subgroup analysis tables for social media use versus anxiety and depression.
[DOCX File, 143 KB - aging_v8i1e71712_app1.docx]

References

1. Patel V, Saxena S, Lund C, et al. The Lancet Commission on Global Mental Health and Sustainable Development. The Lancet 2018 Oct;392(10157):1553-1598. [doi: 10.1016/S0140-6736(18)31612-X]

XSL•FO
RenderX

2. GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Psychiatry 2022 Feb;9(2):137-150. [doi: 10.1016/S2215-0366(21)00395-3]
3. Kieling C, Buchweitz C, Caye A, et al. Worldwide prevalence and disability from mental disorders across childhood and adolescence: evidence from the global burden of disease study. JAMA Psychiatry 2024 Apr 1;81(4):347-356. [doi: 10.1001/jamapsychiatry.2023.5051] [Medline: 38294785]
4. Grolli RE, Mingoti MED, Bertollo AG, et al. Impact of COVID-19 in the mental health in elderly: psychological and biological updates. Mol Neurobiol 2021 May;58(5):1905-1916. [doi: 10.1007/s12035-020-02249-x] [Medline: 33404981]
5. Mehra A, Rani S, Sahoo S, et al. A crisis for elderly with mental disorders: relapse of symptoms due to heightened anxiety due to COVID-19. Asian J Psychiatr 2020 Jun;51:102114. [doi: 10.1016/j.ajp.2020.102114] [Medline: 32334406]
6. Mental health of older adults. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/mental-health-of-older-adults [accessed 2024-09-04]
7. Vigo D, Thornicroft G, Atun R. Estimating the true global burden of mental illness. Lancet Psychiatry 2016 Feb;3(2):171-178. [doi: 10.1016/S2215-0366(15)00505-2] [Medline: 26851330]
8. Arias D, Saxena S, Verguet S. Quantifying the global burden of mental disorders and their economic value. EClinicalMedicine 2022 Dec;54:101675. [doi: 10.1016/j.eclinm.2022.101675] [Medline: 36193171]
9. Lin YJ, Huang MF, Yeh YC, Chen CS. Predicting risk of dementia among the elderly with major depressive disorder in remission: a prospective study. Int J Geriatr Psychiatry 2024 Feb;39(2):e6065. [doi: 10.1002/gps.6065] [Medline: 38319233]
10. Gill S. How many people own smartphones in the world? (2024-2029). Priori Data. 2024 Aug 16. URL: https://prioridata.com/data/smartphone-stats/ [accessed 2024-09-05]
11. The 53rd statistical report on the development of china's internet was released, highlighting how the internet stimulates the economy and society to embrace "new" dynamics (big data observation). Office of the Central Cyberspace Affairs Commission. URL: https://www.cnnic.com.cn/IDR/ReportDownloads/202405/P020240509518443205347.pdf [accessed 2024-09-05]
12. Lu X, Yao Y, Jin Y. Digital exclusion and functional dependence in older people: findings from five longitudinal cohort studies. EClinicalMedicine 2022 Dec;54:101708. [doi: 10.1016/j.eclinm.2022.101708] [Medline: 36353265]
13. The elderly legion breaks through "invisible loneliness". Yangtze Evening Post. URL: http://www.zgnt.net/jhwbszb/pc/c/202408/13/content_181506.html [accessed 2025-09-05]
14. Ke Y. Ageing on WeChat: the impact of social media on elders in urban China. J Vis Media Anthropol 2015;1(1):8-21 [FREE Full text]
15. Seabrook EM, Kern ML, Rickard NS. Social networking sites, depression, and anxiety: a systematic review. JMIR Ment Health 2016 Nov 23;3(4):e50. [doi: 10.2196/mental.5842] [Medline: 27881357]
16. O'Keeffe GS, Clarke-Pearson K, Council on Communications and Media. The impact of social media on children, adolescents, and families. Pediatrics 2011 Apr;127(4):800-804. [doi: 10.1542/peds.2011-0054] [Medline: 21444588]
17. Appel H, Gerlach AL, Crusius J. The interplay between Facebook use, social comparison, envy, and depression. Curr Opin Psychol 2016 Jun;9:44-49. [doi: 10.1016/j.copsyc.2015.10.006]
18. Chen E, Wood D, Ysseldyk R. Online social networking and mental health among older adults: a scoping review. Can J Aging 2022 Mar;41(1):26-39. [doi: 10.1017/S0714980821000040] [Medline: 33736725]
19. Wegener EK, Bergschöld JM, Whitmore C, Winters M, Kayser L. Involving older people with frailty or impairment in the design process of digital health technologies to enable aging in place: scoping review. JMIR Hum Factors 2023 Jan 27;10:e37785. [doi: 10.2196/37785] [Medline: 36705959]
20. Peng R, Li X, Guo Y, et al. Barriers and facilitators to acceptance and implementation of eMental-health intervention among older adults: a qualitative systematic review. Digit Health 2024;10:20552076241234628. [doi: 10.1177/20552076241234628] [Medline: 38444518]
21. Levis B, Sun Y, He C, et al. Accuracy of the PHQ-2 alone and in combination with the PHQ-9 for screening to detect major depression: systematic review and meta-analysis. JAMA 2020 Jun 9;323(22):2290-2300. [doi: 10.1001/jama.2020.6504] [Medline: 32515813]
22. Micoulaud-Franchi JA, Bartolomei F, McGonigal A. Ultra-short screening instruments for major depressive episode and generalized anxiety disorder in epilepsy: the NDDIE-2 and the GAD-SI. J Affect Disord 2017 Mar 1;210:237-240. [doi: 10.1016/j.jad.2016.12.038] [Medline: 28064112]
23. Li C, Friedman B, Conwell Y, Fiscella K. Validity of the Patient Health Questionnaire 2 (PHQ-2) in identifying major depression in older people. J Am Geriatr Soc 2007 Apr;55(4):596-602. [doi: 10.1111/j.1532-5415.2007.01103.x] [Medline: 17397440]
24. Richardson TM, Friedman B, Podgorski C, et al. Depression and its correlates among older adults accessing aging services. Am J Geriatr Psychiatry 2012 Apr;20(4):346-354. [doi: 10.1097/JGP.0b013e3182107e50] [Medline: 22434017]
25. Creighton AS, Davison TE, Kissane DW. The psychometric properties, sensitivity and specificity of the geriatric anxiety inventory, hospital anxiety and depression scale, and rating anxiety in dementia scale in aged care residents. Aging Ment Health 2019 May;23(5):633-642. [doi: 10.1080/13607863.2018.1439882] [Medline: 29470096]

XSL•FO
RenderX

26. Bányai F, Zsila Á, Király O, et al. Problematic social media use: results from a large-scale nationally representative adolescent sample. PLoS ONE 2017;12(1):e0169839. [doi: 10.1371/journal.pone.0169839] [Medline: 28068404]
27. Riehm KE, Feder KA, Tormohlen KN, et al. Associations between time spent using social media and internalizing and externalizing problems among US youth. JAMA Psychiatry 2019 Dec 1;76(12):1266-1273. [doi: 10.1001/jamapsychiatry.2019.2325] [Medline: 31509167]
28. Eichenberg C, Schneider R, Rumpl H. Social media addiction: associations with attachment style, mental distress, and personality. BMC Psychiatry 2024 Apr 15;24(1):278. [doi: 10.1186/s12888-024-05709-z] [Medline: 38622677]
29. Jin Y, Liu W, Zhao M, Wang D, Hu W. Short video app use and the life of mid-age and older adults: an exploratory study based on a social survey. Popul Res 2021;45(3):31 [FREE Full text]
30. Yan R, Liu X, Xue R, et al. Association between internet exclusion and depressive symptoms among older adults: panel data analysis of five longitudinal cohort studies. EClinicalMedicine 2024 Sep;75:102767. [doi: 10.1016/j.eclinm.2024.102767] [Medline: 39188711]
31. Lei X, Matovic D, Leung WY, Viju A, Wuthrich VM. The relationship between social media use and psychosocial outcomes in older adults: a systematic review. Int Psychogeriatr 2024 Sep;36(9):714-746. [doi: 10.1017/S1041610223004519] [Medline: 38287785]
32. Guzman AA, Brecht ML, Doering LV, Macey PM, Mentes JC. Social media use and depression in older adults: a systematic review. Res Gerontol Nurs 2023;16(2):97-104. [doi: 10.3928/19404921-20230220-05] [Medline: 36944173]
33. Cotten SR, Schuster AM, Seifert A. Social media use and well-being among older adults. Curr Opin Psychol 2022 Jun;45:101293. [doi: 10.1016/j.copsyc.2021.12.005]
34. Verduyn P, Ybarra O, Résibois M, Jonides J, Kross E. Do social network sites enhance or undermine subjective well‐being? A critical review. Soc Issues Policy Rev 2017 Jan;11(1):274-302. [doi: 10.1111/sipr.12033]
35. Gao J, Zheng P, Jia Y, et al. Mental health problems and social media exposure during COVID-19 outbreak. PLoS ONE 2020;15(4):e0231924. [doi: 10.1371/journal.pone.0231924] [Medline: 32298385]
36. Yue Z, Zhang R, Xiao J. Passive social media use and psychological well-being during the COVID-19 pandemic: the role of social comparison and emotion regulation. Comput Human Behav 2022 Feb;127:107050. [doi: 10.1016/j.chb.2021.107050] [Medline: 34646057]
37. Schmuck D, Karsay K, Matthes J, Stevic A. "Looking up and feeling down". The influence of mobile social networking site use on upward social comparison, self-esteem, and well-being of adult smartphone users. Telematics and Informatics 2019 Sep;42:101240. [doi: 10.1016/j.tele.2019.101240]
38. Keles B, McCrae N, Grealish A. A systematic review: the influence of social media on depression, anxiety and psychological distress in adolescents. Int J Adolesc Youth 2020 Dec 31;25(1):79-93. [doi: 10.1080/02673843.2019.1590851]
39. Lopes LS, Valentini JP, Monteiro TH, et al. Problematic social media use and its relationship with depression or anxiety: a systematic review. Cyberpsychol Behav Soc Netw 2022 Nov;25(11):691-702. [doi: 10.1089/cyber.2021.0300] [Medline: 36219756]
40. Karaş H, Küçükparlak İ, Özbek MG, Yılmaz T. Addictive smartphone use in the elderly: relationship with depression, anxiety and sleep quality. Psychogeriatrics 2023 Jan;23(1):116-125. [doi: 10.1111/psyg.12910] [Medline: 36366976]
41. Özbek MG, Karaş H. Associations of depressive symptoms and perceived social support with addictive use of social media among elderly people in Turkey. Psychogeriatrics 2022 Jan;22(1):29-37. [doi: 10.1111/psyg.12770] [Medline: 34605123]
42. Wang T, Zhang S, Li Q, et al. Socioeconomic inequalities in the relationship between internet usage patterns and depressive symptoms: evidence from a Chinese longitudinal study. J Glob Health 2024 Aug 16;14:04127. [doi: 10.7189/jogh.14.04127] [Medline: 39149826]
43. Xu J, Zhang Q. The relationship between internet use and mental health of the elderly: analysis of the differences between urban and rural. PLoS ONE 2023 Jan 26;18(1):e0280318. [doi: 10.1371/journal.pone.0280318]

Abbreviations

CHSC: community health service center
GAD-2: 2-item Generalized Anxiety Disorder scale
GBD: Global Burden of Disease Study
OR: odds ratio
PHQ-2: 2-item Patient Health Questionnaire
SES: socioeconomic status

XSL•FO
RenderX

Edited by J Wang, Y Jiang; submitted 02.02.25; peer-reviewed by J Zhang, W Vuong; revised version received 10.10.25; accepted 13.10.25; published 10.11.25.

Please cite as:
Huang J, Ge Z, Chu Y, Yan Y, Zhang W, Liang H, Yang Y, Wang H
Associations Between Social Media Use and Anxiety and Depression Among Older Adults : Cross-Sectional Study
JMIR Aging 2025;8:e71712
URL: https://aging.jmir.org/2025/1/e71712
doi:10.2196/71712

© Jiaoling Huang, Zhenxing Ge, Yijing Chu, Yuge Yan, Wei Zhang, Hong Liang, Yuqi Yang, Hui Wang. Originally published in JMIR Aging (https://aging.jmir.org), 10.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Online Communities as a Support System for Alzheimer Disease and Dementia Care: Large-Scale Exploratory Study

Sidharth Kaliappan[1], MS; Chunyu Liu[2], MS; Yoshee Jain[3]; Ravi Karkar[1], PhD; Koustuv Saha[3], PhD

[1]Manning College of Information & Computer Sciences, University of Massachusetts Amherst, Amherst, MA, United States

[2]School of Information Sciences, University of Illinois Urbana-Champaign, Champaign, United States

[3]Siebel School of Computing and Data Science, The Grainger College of Engineering, University of Illinois Urbana-Champaign, Urbana, IL, United States

Corresponding Author:
Koustuv Saha, PhD
Siebel School of Computing and Data Science
The Grainger College of Engineering
University of Illinois Urbana-Champaign
201 N. Goodwin Ave
Siebel Center for Computer Science, 4212
Urbana, IL, 61801
United States
Phone: 1 2172443824
Email: ksaha2@illinois.edu

Abstract

Background: Alzheimer disease (AD) is the leading type of dementia, demanding comprehensive understanding and intervention strategies. In the United States, where over 6 million people are impacted, the prevalence of AD and related dementias (AD/ADRD) presents a growing public health challenge. However, individuals living with AD/ADRD and their caregivers frequently express feelings of marginalization, describing interactions characterized by perceptions of patient infantilization and a lack of respect.

Objective: This study aimed to address 2 key research questions (RQs). For RQ1, we investigated the needs and concerns expressed by participants in online social communities focused on AD/ADRD, specifically on 2 platforms–Reddit's r/Alzheimers and ALZConnected. For RQ2, we examined the prevalence and distribution of social support corresponding to these needs and concerns, and the association between these needs and received support.

Methods: We collected 13,429 posts and comments from the r/Alzheimers subreddit spanning July 2014 to November 2023, and 90,113 posts and comments from ALZConnected between December 2020 (the community's earliest post) and November 2023. We conducted topic modeling using latent Dirichlet allocation (LDA), followed by labeling to identify the major topical themes of discussions. We used transfer learning classifiers to identify the occurrences of emotional support (ES) and informational support (IS) in the comments (or responses) in the discussions. We built regression models to examine how various topical themes are associated with the kinds of support received.

Results: Our analysis revealed a diverse range of topics reflecting community members' varying needs and concerns of individuals affected by AD/ADRD. These themes encapsulate the primary discussions within the online communities: memory care, nursing and caregiving, gratitude and acknowledgment, and legal and financial considerations. Our findings indicated a higher prevalence of IS compared to ES. Regression models revealed that ES primarily occurs in posts relating to nursing and caring, and IS primarily occurs in posts concerning medical conditions and diagnosis, legal and financial, and caregiving at home.

Conclusions: This study reveals that online communities dedicated to AD/ADRD support engage in discussions on a wide range of topics, such as memory care, nursing, caregiving, and legal and financial challenges. The findings shed light on the key pain points and concerns faced by individuals managing AD/ADRD in their households, revealing how they leverage online platforms for guidance and support. These insights underscore the need for targeted institutional and social interventions to address the specific needs of AD/ADRD patients, caregivers, and other family members.

(JMIR Aging 2025;8:e68890) doi:10.2196/68890

XSL•FO
RenderX

KEYWORDS

social media; natural language; Alzheimer disease; social support; online communities; machine learning

Introduction

Alzheimer disease and related dementias (AD/ADRD) consist of neurodegenerative conditions characterized by cognitive impairment, memory loss, and executive function decline [1]. Over 6 million individuals in the United States grapple with the challenges imposed by AD/ADRD [2], signaling a substantial and escalating public health concern, and underscoring the urgent need for comprehensive understanding and intervention. However, despite the existence of available resources, the long-term services and support for people with AD/ADRD exhibit inconsistencies within the existing health care system [3,4]. People with AD/ADRD and their caregivers often perceive interactions with others as reflective of patient infantilization and a lack of sensitivity and respect [5]. They frequently encounter added challenges, encompassing mental health issues, social exclusion, and instances of discrimination [6,7]. In fact, if caregivers' concerns are overlooked, it may accelerate cognitive decline in individuals with AD/ADRD [8].

As technology gets more integrated into our lives and society, previous research has also noted the use of technology by AD/ADRD individuals and their caregivers [9]. Research highlights the significance of acknowledging and addressing the technology requirements and acceptance within AD/ADRD care [10]. However, more emphasis is placed on incorporating technology into AD/ADRD care, with the objective of improving the user experience to assist both patients and caregivers in managing daily life and monitoring their health status on a regular basis. A significant concern among AD/ADRD caregivers, however, remains the limited availability of adequate support systems at both institutional and social levels [11,12]. With the ubiquity of the internet, more individuals are turning to online spaces to seek information, discuss challenges, and share personal experiences. In particular, online communities are spaces that enhance a sense of belonging by connecting individuals with similar experiences. These online communities enable individuals to candidly self-disclose their concerns and seek support from others [13-20]. Previous work also noted how these online platforms could play a role in shaping the experiences of those living with progressive AD/ADRD [21-23]. These online support communities are very much in line with the social support behavioral theory (SSBC) that identifies the interactions intended to provide social support toward helping individuals cope with stress, enhance well-being, and foster positive relationships [24]. Nonetheless, social expectations and requirements in these online communities are largely unknown. Such knowledge would not only enhance our understanding of how the needs and concerns of AD/ADRD individuals and caregivers evolve over time but also help inform the design of targeted online interventions to address these needs.

This study aims to explore the needs and concerns of members within online communities on AD/ADRD, as well as the social support they actively provide and seek. To avoid any biases embedded within specific communities, we examined 2 distinct online communities—r/Alzheimers on Reddit and AlzConnected—dedicated discussion forums for AD/ADRD. Our study is guided by the following research questions (RQs): (1) RQ1: What are the needs and concerns people discuss in online AD/ADRD communities? (2) RQ2: What is the prevalence and distribution of social support across various needs and concerns, and how are different types of needs associated with the types of support received?

Methods

Data

We sourced our data from 2 distinct online AD/ADRD communities—r/Alzheimers on Reddit and AlzConnected, which we describe below:

Reddit

Reddit is a widely used semianonymous online platform that houses a network of over 52 million active users and 100 thousand active online communities. These communities are called "subreddits,"—each offering demographic, topical, or interest-specific discussion boards. The semianonymity of Reddit is known to enable candid self-disclosure and seek social support on stigmatized and sensitive topics [13,25,26], including discussions related to AD/ADRD [22,23,27].

We used Reddit's search feature to look up subreddits related to AD/ADRD, and 2 of the largest subreddits returned were r/AlzheimersGroup (over 117 thousand members) and r/Alzheimers (over 13 thousand members). We then conducted sanity checks to understand the content of these subreddits. These investigations revealed that r/AlzheimersGroup was a community centered on humor and sarcasm, where people predominantly pretended to have forgetfulness and shared memes and cartoon references to Garfield [28]. r/AlzheimersGroup also self-describes that "Serious discussion belongs in r/Alzheimers," Therefore, we omitted r/AlzheimersGroup from our ensuing analyses. The r/Alzheimers subreddit consisted of discussions where people shared troubles and provided social support on AD/ADRD. We collected all the posts and comments data from this subreddit—1000 posts and 12,429 comments between July 2014 and November 2023.

AlzConnected

AlzConnected is another freely accessible website (alzconnected.org) consisting of discussion boards dedicated to AD/ADRD patients and their caregivers. Although a more recent online community, it has already grown a significant user base and is supported by the Alzheimer Association (alz.org). Here, community members express AD/ADRD-related concerns and provide and seek support from one another. This community has multiple thematic discussion boards—(1) living with AD/ADRD, further broken down into "I am living with Alzheimer or other dementia and I am living with younger-onset Alzheimer," and (2) Supporting someone living with AD/ADRD, further broken down into, "I am a caregiver (general topics), caring for a spouse or partner, caring for a parent, caring long distance, and supporting those who have lost someone."

XSL•FO
RenderX

We adopted web scraping approaches to collect the entire data from ALZConnected amounting to 10,328 posts and 79,785 comments from the community between December 2020 (earliest post in the community) and November 2023.

Table 1 summarizes our dataset from the 2 online communities.

Table 1. Summary of dataset from online communities catering to discussions on AD/ADRD[a].

Community	Number of posts	Number of comments
r/Alzheimers	1000	12,429
ALZConnected	10,328	79,785

[a]AD/ADRD: Alzheimer disease and related dementias.

Analyzing the Needs and Concerns

Toward answering RQ1, we conducted unsupervised topic modeling followed by human annotation to obtain the needs and concerns expressed by members of AD/ADRD online communities.

Topic Modeling and Thematic Analysis

We started by examining the content of posts and comments in r/Alzheimers and AlzConnected to gain insights into the prevalent needs and concerns discussed within AD/ADRD online communities.

Building a Topic Model

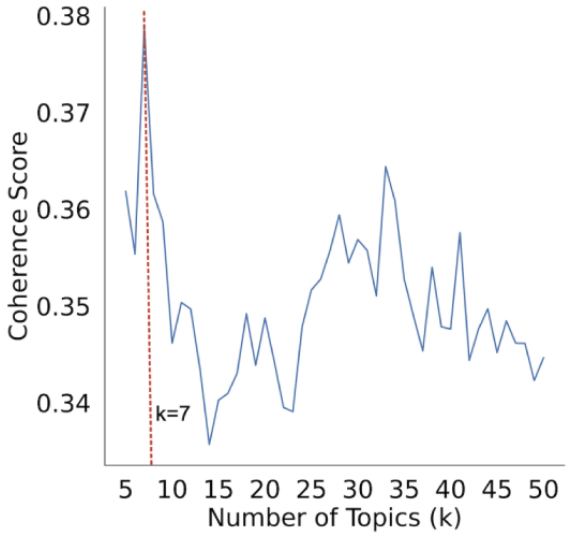
Topic modeling techniques have been adopted to analyze and identify key themes within social media discourse [29-32] including AD/ADRD-related discussions [22,23]. We built topic models on our entire dataset (posts and comments) by using latent Dirichlet allocation (LDA) [33], an unsupervised machine learning algorithm widely used for analyzing corpora. LDA generates latent topic distributions, making it a valuable tool for analyzing large, unlabeled document collections and their clustering into distinct groups based on common themes. LDA has been heavily leveraged in previous work on social media language analysis [29,30]. After tokenization, stop word removal, and adding n-grams (n=2,3), we created bag-of-words representations of the dataset.

Determining the Optimal Number of Topics

Given that LDA does not automatically determine the optimal number of topics, we adopted a semiautomated approach, motivated by previous work [22,30,31]. For this, we first used the coherence score as a metric to assess the model fit, varying the number of topics. The coherence score evaluates the tendency of words within a topic to co-occur and correlates well with topic modeling quality [34]. Then, we hand-selected a few topic models that occurred in the zone of highest coherence scores. Two coauthors manually evaluated the quality of these different topic models in terms of assessing the between-topic heterogeneity and within-topic homogeneity of words to narrow down the best topic model for our ensuing analyses.

We computed the coherence score for our 3 preprocessed corpora, varying the number of topics (k) from 5 to 50 (refer to Figure 1). We found the highest coherence at k=7. Consequently, we manually evaluated the quality of the topic modeling for k=5, k=7, and k=10.

Figure 1. Coherence values over varying number of topics in the LDA (latent Dirichlet allocation) topic model.



Labeling the Topics

The above steps yielded topical clusters, each of which contained clusters of keywords. We adopted a thematic analysis approach to label each topical cluster in a meaningful and interpretable way. The thematic analysis was conducted by 2 authors, familiar with social media analysis and AD/ADRD literature. They adopted an inductive coding approach to first independently code the topics based on the cluster of keywords, referencing sample posts corresponding to each topic. Then,

XSL•FO
RenderX

the authors engaged in collaborative discussions, comparing their codes, and gradually converging on cohesive thematic labels. This meticulous process ensured a nuanced and comprehensive analysis of the data. We found that the topic modeling with k=7 yielded the best results in terms of coherency within topics and boundaries between topics.

Classifying Support and Examining Supportive Expressions

Toward answering RQ2, we analyzed the responses within the online communities to examine the nature and forms of social support provided. This section details our methodology for categorizing these responses into distinct subcategories and contrasts the patterns of support provision within the Alzheimer community against those observed in a broader user demographic.

Classifying Support Expressions

Overview

The role of social support is essential in assisting individuals in managing psychological distress and various life adversities [35]. The emergence of social media platforms and virtual communities has led to the digital transposition of traditional social support mechanisms [13,36]. Specifically, online forums such as Reddit have increasingly mirrored the characteristics of virtual support groups, providing a platform for communal interaction and aid. The SSBC schema [37] offers a structured approach for classifying various types of support. Within this framework, informational support (IS) and emotional support (ES) emerge as 2 pivotal categories. Notably, these forms of support have garnered significant theoretical and empirical interest within social computing, as documented in various studies [20,38-43]. ES encompasses actions that provide encouragement, empathy, or care, whereas IS consists of assistance involving advice, information, or knowledge [37]. These 2 forms of support are particularly dominant in online interactions, as they can be effectively conveyed through text, making them well-suited for digital environments where individuals seek both guidance and emotional reassurance.

We categorized support offered in comments into ES and IS across both Reddit and ALZConnected datasets. To achieve this, we used a transfer learning methodology, a process wherein a machine learning classifier is developed by transferring insights from one labeled dataset to another unlabeled dataset with similar characteristics [44]. Specifically, we used a labeled Reddit dataset, previously expert-annotated with ES and IS labels on 396 comments collected from 55 mental health-related subreddits on Reddit [40]. This served as the seed dataset to train our classifiers. In addition, we manually labeled the ES and IS of 100 comments from AlzConnected and 100 comments from r/Alzheimers datasets and added these to the training datasets of the support classifiers. Our initial classifier was a supervised learning classifier for ES and IS, using n-grams (with n=1,2) as features in alignment with methodologies outlined in previous research [39-41]. We adopted a k-fold cross-validation (k=10) approach for evaluating our models.

We used a range of classification algorithms, including naive Bayes, logistic regression, support vector machine (SVM), random forest, and neural network. Among these, the Linear SVM model demonstrated the best performance, achieving a mean area under the curve of 0.84 for ES and 0.87 for IS (refer to Table 2). Figure 2 provides the receiver-operating-characteristic area-under-curve plot for the classifiers used in identifying support types. We use the linear SVM model for our ensuing analyses. For the ES classifier, the predominant features include: “good,” “sorry,” “right,” “dont,” “going,” “people,” “get,” “life,” “im,” and “feel.” In contrast, the IS classifier highlighted features such as: “help,” “might,” “thing,” “go,” “think,” “try,” “get,” “see,” “probably,” and “doesnt.”

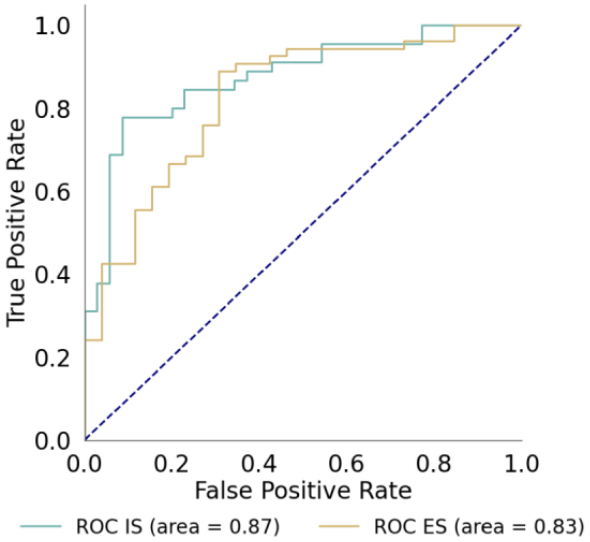
Table 2. Performance metrics of machine learning models of social support classifiers.

Model	Emotional support			Informational support		
	Precision	Recall	Area under the curve	Precision	Recall	Area under the curve
Naive Bayes	0.70	0.58	0.80	0.67	0.66	0.76
Logistic regression	0.73	0.69	0.76	0.75	0.76	0.81
Support vector machine	0.74	0.74	0.80	0.73	0.73	0.83
Random forest	0.68	0.54	0.75	0.67	0.66	0.75
Neural network	0.71	0.68	0.79	0.70	0.65	0.77
After adding augmented labeled data from AD/ADRD[a] communities						
Logistic regression	0.74	0.72	0.77	0.74	0.76	0.80
Support vector machine	0.75	0.76	0.83	0.77	0.78	0.87

[a]AD/ADRD: Alzheimer disease and related dementias.

XSL•FO
RenderX

Figure 2. Receiver-operating-characteristic area-under-curve plot of the emotional and informational support classifiers. ES: emotional support; IS: informational support.



To validate the applicability and relevance of the model, we applied transfer-based learning to the classifiers within our dataset. We conducted a manual validation on a randomly selected subset of 100 comments. This process involved 2 independent annotators classifying each reply for ES and IS. Discrepancies in annotations were resolved through the intervention of a third adjudicator. Further, we compared these manual annotations with the predictions made by the machine learning classifiers. The comparison resulted in an F_1-score of 0.76 for ES, characterized by a precision of 0.74 and accuracy of 0.76, and an F_1-score of 0.77 for IS, precision of 0.77, and accuracy of 0.77, which is an increase from the initial classifier. Given these results, we consider the classifier's performance satisfactory for further analysis. It is important to note that in our classification, ES and IS are not mutually exclusive categories; a single comment may be either both, or neither forms of support. Table 3 shows some example comments in our dataset with labels of ES and IS.

Table 3. Example comments and support labels with emotional and informational support (each comment may be either both, or neither forms of support).

Example comment	Emotional support	Informational support
Just as a note, when my Uncle passed away, I still remember him as the strong figure versus any thoughts of his Alzheimer. Man, my condolences and it's more of a blessing for him and your family for the rapid decline.	✓	x
Lumbar punctures are great and the data is very solid in the ability to diagnose AD[a]. However, LPs can be lengthy (and therefore expensive) procedures and are generally viewed as invasive by the general population. A blood test such as this would highly impact our ability to screen people more quickly for drug trials :)	x	✓
Sometimes you just have to chuckle. I just received yet another call from my person with dementia telling me they needed someone to fix their phone. The same phone they made a call with. And we live over three hours away.	✓	✓
This feels like a lecture, which is not what I need right now. No reply is necessary.	x	x

[a]AD: Alzheimer disease.

Associating the Relationship Between Topics of Concern and Support Received

To assess the support associated with each topic, we used linear regression models to predict the likelihood of receiving different types of support. We built separate models for ES and IS. Each model was designed to predict the probability of receiving a specific type of support based on the topic being discussed. So, we used the topics in the posts' content as independent variables, and the received support as the dependent variable. Essentially, by examining ordinary least squares regression models, we aimed to understand how different topics are associated with the probability of receiving a type of support. This approach allowed us to quantify the relationship between topics and the support they generate, providing valuable insights into the dynamics of support within the community.

Ethical Considerations

This project used publicly available data from online communities, and did not involve any direct interaction with human subject participants. Therefore, this research did not require ethics board approval. However, we are committed to the ethics of the research and followed practices to secure the privacy of the individuals in our dataset. We recognize the sensitivities of our study in terms of revealing the identities of the individuals, and thus deidentified all collected data. This paper only presents paraphrased quotes to reduce traceability while providing necessary context.

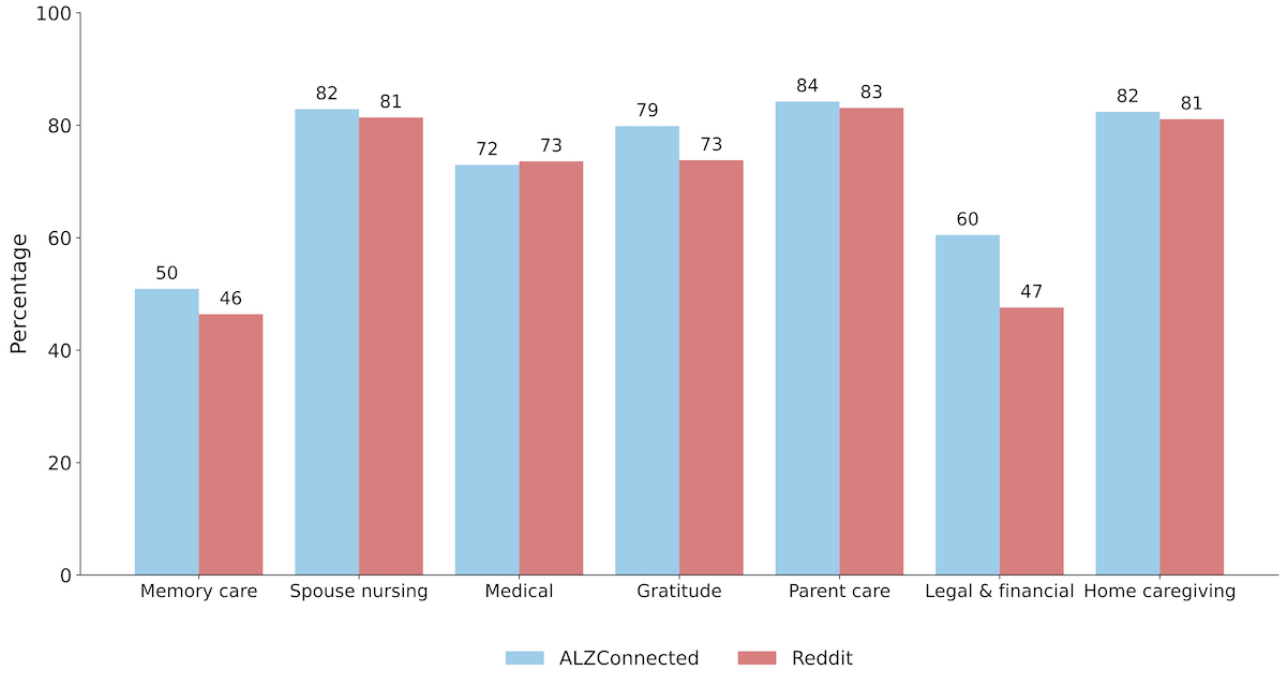
XSL•FO
RenderX

Results

RQ1: Themes of Needs and Concerns in Online Communities

We now describe the findings from our topic modeling followed by thematic analyses. Table 4 shows the 7 topics and relevant thematic analysis derived through LDA topic modeling. These topics serve as the primary themes, summarizing the needs and concerns of members within the AD/ADRD online communities. Figure 3 shows the overall distribution of topical occurrences in the 2 datasets, showing that the topics appear similarly on both platforms.

Figure 3. Distribution of topics as a percentage of posts on r/Alzheimers and ALZConnected.



We describe these topical themes as: First, the topics reveal diversity in community members' needs based on relationships with patients experiencing AD/ADRD. We find that several topics, such as memory care support, nursing for a spouse, caring for patients, and caregiving at home are linked to informal caregivers, such as patients' spouses, children, and close family members, along with professional caregivers providing in-home assistance. This reveals that community members with different roles may encounter various concerns and needs.

We find a substantial amount of content focusing on memory care, particularly about early onset patients with AD/ADRD, with words like "memory care" and "assisted living facility" (Topic 1, 5, and 7). In general, memory loss is a common symptom of AD/ADRD [45]. Memory care is frequently observed as dedicated and free-standing memory care living spaces within assisted living communities or integrated into life care communities [46]. Individuals also discuss how to choose an assisted living home with a memory care unit for patients and issues encountered in memory care units. One individual posted in r/Alzheimers to seek suggestions:

> *Hello everyone, could we discuss the options for finding an assisted living facility with memory care? My wife has Early-onset Alzheimer, and she is 57.*

We also find members who are early-stage patients looking for advice to plan the future and are gaining insight into experience within the memory care unit. An early onset patient posted:

> *Next week, I plan to visit an assisted living home with a memory care unit. My doctor has advised me to begin preparing for my future no matter what I can.*

AD/ADRD leads to a progressive decline in the capacity to perform daily living activities [47]. We found extensive discussions related to nursing and care, with the majority originating from caregivers and patients' family members (Topic 2, 5, and 7). Some content involves sharing AD/ADRD-related books and information. Within these posts, members recommend or summarize useful knowledge in a comprehensive way, and they are enthusiastic about discussing it with others. For instance, one caregiver posted their thoughts on the recommended book, "I recommend giving this book a read. Feel free to share your thoughts with me if you explore it." Also, a significant amount of content revolves around the sharing of experiences and challenges. Some members shared experiences when they felt happy in a largely challenging journey of caregiving for their family members with Alzheimer. For instance, a member wrote;

> *He didn't know who I was, but he got very excited that I helped him make handprint art to take to his 2 daughters and 3 granddaughters. He may not recognize me, but things like this show me the love is still there. Feeling blessed.*

XSL•FO
RenderX

Table 4. Topic themes with representative keywords and example posts.

	Topic keywords	Example post
Topic 1: Memory care and assisted living	care, facility, memory, memory care, wife, home, living, darling, dear darling, dear, darling wife, dear darling wife, assisted, assisted living, nursing, staff, significant, living facility, assisted living facility, and room	*Is there anyone willing to share their experience of placing their loved one into a memory care unit?*
Topic 2: Nursing for spouse	significant, director, director nursing, nursing, just, husband, know, dear, time, like, day, good, dear husband, think, sorry, hope, things, did, going, and got	*If I get COVID, would my wife have to be placed in a nursing home? What actions should I do in such a situation?*
Topic 3: Medical condition and diagnosis	years, alzheimer, disease, husband, ago, alzheimer disease, significant, diagnosed, year, just, stage, diagnosis, early, time, years agodementia, months, know, care, and memory	*What support or resources are available for people in the MCI stage who may be unaware of it? I'm asking because my younger sister has been diagnosed with early-onset Alzheimer.*
Topic 4: Acknowledgment and gratitude	significant, thank, dementia, good, hospice, help, loved, support, meds, know, doctor, thanks, best, sorry, disease, helpful, forum, new, and information	*I'd like to thank those who recommended getting an electronic pet for a loved one with Alzheimer. I bought one for my grandma. She absolutely loves it!*
Topic 5: Caring for parents	mom, just, significant, like, time, help, know, mother, things, home, going, day, really, think, doesn, house, need, feel, dad, and tell	*I am an only daughter and I have taken on the responsibility of caring for my mother who received a diagnosis of Alzheimer and vascular dementia for 2 years.*
Topic 6: Legal and financial matters	attorney, husband, dear husband, care, dear, power, power of attorney, need, law, Medicaid, significant, help, elder law, money, long, state, law attorney, elder law attorney, and pay	*My husband was just diagnosed with younger onset Alzheimer. We are in search of an elder care lawyer in the Naperville or surrounding area to assist us with financial planning.*
Topic 7: Caregiving at home	dementia, dad, person, mom, family, care, make, person dementia, people, like, time, life, significant, need, help, things, think, caregiver, nursing, and home	*I have taken on the role of a "mom" for my mom; she relies on me for all her necessities, and needs, or just for chat. I'm only 20. I wish to live my own life. However, guilt keeps me at home, and I'm concerned that moving out might intensify it. Please tell me what to do.*

Some members place individuals with AD/ADRD in skilled nursing homes, while others serve as caregivers at home, leading to varying sets of issues. Some common concerns in nursing homes are the occurrence of living issues faced by patients residing in facilities, inappropriate interactions among patients, and instances where patients get expelled from the nursing facility and sent back home. We also found an instance where the patient required hearing aids, but the care provided by nursing facilities was not satisfactory. Consequently, a member posted:

> *Regrettably, there are many instances where the thing is either not carried out properly or not done at all. During my visits, I've observed that there are times when she has the hearing aids in, but they are non-functional.*

In-home caregivers share the fatigue and strain arising from family relationships, as well as the distress of balancing their own lives with caregiving responsibilities. One member posted about being the sole caregiver without other siblings' assistance. They wrote, seeking advice:

> *I have a full-time job and my own family. I find myself in need of a break, and I'm hesitant to consider placing her in a nursing home. I'm concerned that the unfamiliar environment might exacerbate her dementia.*

Community members also expressed gratitude, encouragement, and acknowledgment. Members within online AD/ADRD communities can communicate about specific health-related topics, often finding online support a welcoming and comfortable venue for discussing sensitive issues [48]. Part of this involves mutual support and appreciation among community members, often occurring when someone shares their story or receives help and support from other members. Notably, posts and comments of this nature are typically longer and more detailed, prompting many individuals to express gratitude at the end, thanking others for taking the time to read their detailed posts. For instance, a member on r/Alzheimers, after venting about the challenges of being a caregiver, expressed:

> *Thanks for the opportunity to share my thoughts right here. I'll read your many posts before mine and offer my opinions and encouragement.*

XSL•FO
RenderX

Individuals also express gratitude not only directed toward other community members but also between caregivers and patients. Following the death of a loved one, members grieved recalling their shared past, the challenges of life, and caregiving after the diagnosis, such as, "I am thankful that his suffering finally ends. He is with my mom now in heaven." Furthermore, when reminiscing about early shared experiences, members often express gratitude and love for the individuals with AD/ADRD, such as, "Reflecting on the cherished moments we shared brings me joy, and I am grateful for the knowledge he imparted to me."

Finally, we also observed concerns related to legal and financial matters. Many of these concerns come from family members of individuals with AD/ADRD. With the progression of AD/ADRD, patients require increasing levels of care and assistance, leading to legal issues such as establishing guardianship or power of attorney (POA). Besides, financial concerns are a widely discussed topic in the communities, covering aspects such as handling patient's assets, as well as the costs associated with caregiving and treatment. For example, a member posted to seek advice from the community:

> *Her funds is running down. We still hope that we can place her in a nursing facility after her funds are used up by the memory care (MC) facility. I need to know if there is any assistance available to support her continued stay in MC.*

RQ2: Social Support in Online AD/ADRD Communities

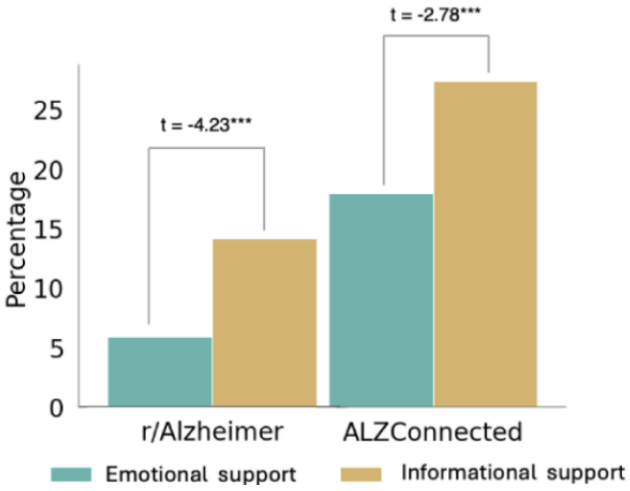
Distribution of ES and IS

We first examined how ES and IS are distributed in our datasets. We observe similar trends in the distribution of support in both communities—on average, IS is 6% more prevalent in r/Alzheimers and 10% more prevalent in AlzConnected than ES in respective communities. Figure 4 shows the overall distribution of ES and IS in the 2 datasets.

As previously noted, we obtained the topics on the posts and the support received in the comments to the posts. We break this down into the topics of the posts, and Figure 5 shows the distribution of ES and IS in the comments to the posts of varying topics in the 2 datasets.

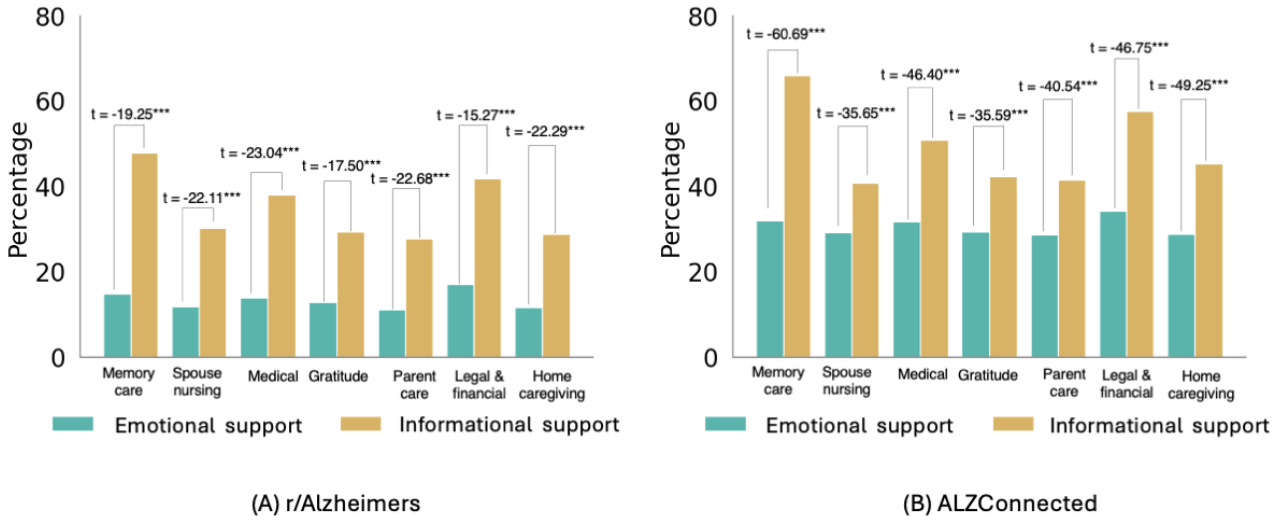
Both Figures 4 and 5 also show independent sample *t* tests between the occurrences of ES and IS. For each topical theme, we note statistically significant differences in occurrences of the 2 types of support.

Figure 4. Distribution of ES and IS to posts on r/Alzheimers and ALZConnected.



XSL•FO
RenderX

Figure 5. Distribution of emotional and IS in comments to posts with varying topics on (A) r/Alzheimers and (B) ALZConnected.



Regression Models for Predicting Support

To further analyze how support was provided corresponding to each topic, we built regression models. For each of our datasets (r/Alzheimers and AlzConnected), we built 2 models, 1 with IS as the dependent variable and another with ES as the dependent variable. Both of the models had all the topic occurrences as independent variables. Essentially, these models examined the likelihood of receiving ES and IS for each topic.

Table 5 shows the coefficients of the independent variables (topics) for ES and IS across the 2 datasets of r/Alzheimers and AlzConnected. All of these models show a significant goodness-of-fit (R^2). These coefficients signal the likelihood of receiving the respective support in the community. We describe our observations below.

Table 5. Coefficients of linear regression models with the topics of posts as independent variables and support (emotional and informational) responses as dependent variables.

Independent variables	Dependent variables, coefficient			
	Emotional support		Informational support	
	r/Alzheimers	ALZConnected	r/Alzheimers	ALZConnected
Intercept	–3.00E-04	3.00E-04	–0.008[a]	0.027[a]
Memory care	0.014	0.012[b]	0.191[a]	0.258[a]
Nursing for spouse	0.016[b]	0.079[a]	0.066[a]	0.013[a]
Medical condition	0.048[a]	0.067[a]	0.182[a]	0.164[a]
Acknowledgment and gratitude	0.041[a]	0.058[a]	0.038[a]	0.034[a]
Caring for parents	0.014[c]	0.089[a]	0.031[a]	0.048[a]
Legal and financial	0.056[a]	0.075[a]	0.088[a]	0.13[a]
Caregiving at home	0.035[a]	0.038[a]	0.055[a]	0.129[a]
R^2	0.777[a]	0.144[a]	0.260[a]	0.294[a]

[a]$P<.001$.
[b]$P<.01$.
[c]$P<.05$.

ES Posts and Topics

For both r/Alzheimers and AlzConnected, we find that topics relating to nursing and caring, such as “Nursing for Spouse,” “Caring for Parents,” and “Caregiving at Home” show high positive coefficients with statistical significance, that is, posts with these topics are likely to receive ES. Further, topics that are seemingly about informational content, such as “Legal and Financial” and “Medical Conditions and Diagnosis” also receive ES.

For instance, an individual self-disclosed their stress about the fact that their mom changed a POA and was in complete denial of her symptoms:

XSL•FO
RenderX

> *I believe my mom just changed her POA and took me completely off. Her husband is covering for her and is in complete denial about her diagnosis. [..] She is clearly paranoid and delusional. I will consult with an elder care attorney soon. Any other advice? She blames me for her driver's license getting taken away. I am the only family member in the state. Neither she nor her husband are capable of overseeing her medical care. [..] I am very stressed about this situation. Thank you for any thoughts.*

IS Posts and Topics

For both r/Alzheimers and AlzConnected, posts relating to "Memory Care Support," "Medical Condition and Diagnosis," "Legal and Financial," and "Caregiving at Home" show high positive coefficients with statistical significance. Interestingly, posts relating to "Nursing for Spouse" and "Caring for Parents," which occurred high for ES, are also likely to receive IS. For instance, an individual who described their situation, also asked, "How are folks keeping their careers while helping to care for a parent?"

Discussion

Principal Results

This study sheds light on the complex needs of individuals dealing with AD/ADRD as patients, caregivers, family members, and loved ones. Our analyses revealed seven primary topical themes—(1) memory care and assisted living, (2) nursing for a spouse, (3) medical condition and diagnosis, (4) acknowledgment and gratitude, (5) caring for parents, (6) legal and financial aspects, and (7) caregiving at home. We found that IS was more prevalently provided than ES across both platforms. In particular, ES was predominantly found in posts related to nursing and caregiving, suggesting that caregivers turn to these communities to cope with the emotional burdens associated with their roles. In addition, IS was more frequently associated with concerns about medical conditions, legal and financial issues, and caregiving at home, indicating a demand for practical advice and resources in managing the complexities of AD/ADRD.

These findings highlight the critical role that online communities play in supplementing the support networks of AD/ADRD caregivers and family members. Many caregivers face challenges in accessing adequate in-person support due to stigma, social isolation, or lack of resources, making online platforms an important alternative for both emotional validation and practical advice—as also noted in the case of other stigmatized mental health [13,38-41,49] and health conditions (such as cancer [17,19,20] and HIV [50]). The predominance of IS suggests that individuals seek concrete advice to manage the daily realities of caregiving, including navigating health care systems, legal documentation, and financial planning. This aligns with previous research indicating that caregivers often experience a steep learning curve when managing AD/ADRD, requiring access to timely and accurate information [51,52]. Furthermore, the need for ES around nursing and caregiving emphasizes the significant emotional and psychological toll on caregivers [53]. These findings underscore the importance of community support in mitigating mental health issues such as feelings of burnout, stress, and isolation—some of the major concerns experienced by AD/ADRD caregivers [54]. We also noted expressions of gratitude and acknowledgment within these communities, which may indicate a sense of belonging and solidarity among the members of online communities, especially among caregivers to help boost their resilience and sustain their caregiving roles over the long term. Therefore, fostering these online communities can be a low-cost and scalable way to address caregiver burnout, which is a critical issue given the growing aging population and the increasing prevalence of AD/ADRD.

Our study also bears important implications for health care practitioners, policy makers, and support organizations. First, online communities provide empirical insights into the plausible gaps in accessible resources for caregivers. For instance, institutional support from governmental policies, health care systems, and AD/ADRD care organizations can leverage online platforms to identify the needs. Further, these platforms can be used to disseminate accurate and specific information tailored to the needs of the caregivers, potentially reducing the burden on healthcare professionals and improving caregiver confidence in managing AD/ADRD-related challenges. In addition, these platforms can be integrated into existing health care support structures and information.

Health care systems and AD care organizations could leverage online platforms to disseminate accurate, easily digestible information tailored to the needs of caregivers, potentially reducing the burden on health care professionals and improving caregiver confidence in managing AD/ADRD-related challenges. In addition, these platforms could be integrated into existing health care support structures, allowing caregivers to receive timely advice from health care professionals, thus improving patient outcomes. From a policy perspective, there is a need to enhance public health initiatives that support caregivers, especially in terms of financial planning and legal guidance, which was one of the major topics of discussion.

Policy makers could consider developing targeted programs that address these specific needs, such as offering workshops on financial planning for families affected by AD/ADRD or increasing access to legal assistance. Moreover, expanding caregiver support programs, including respite care and mental health services, could alleviate the emotional burden as prevalently discussed in the online communities. For technology designers and developers, our study suggests opportunities to enhance digital tools aimed at supporting caregivers. In particular, with the increasing prevalence of using artificial intelligence (AI) in different sectors, platforms can also incorporate AI-driven recommendation systems to provide personalized resources based on each individual's specific question or concern. In addition, integrating mental health support tools, such as access to counseling services or peer support groups within these online communities, could further enhance the mental well-being of caregivers.

Limitations

Our study has limitations which also suggest interesting future research directions. Although adopting computational

XSL•FO
RenderX

approaches enabled our exploration of the depth and breadth of support dynamics across various topics, our findings are not necessarily generalizable to the broader population or other contexts beyond the online communities analyzed. As noted in Figure 3, although both communities follow the trend of having more IS posts than ES, the overall level of engagement across the 2 communities is not the same. In particular, our study is inherently limited by self-selection bias, as we only study individuals who actively chose to post and engage in these online communities. This means that our findings may disproportionately reflect the experiences of individuals who are more active on social media and are motivated and willing to share publicly on these platforms. In addition, this study does not capture the perspectives of passive users who browse forums for information without directly participating, which may exclude a significant portion of the online communities. Furthermore, digital inequity presents another challenge. Not all individuals have equal access to digital platforms, reliable internet, or the necessary skills to engage effectively online. This disparity can exclude voices from socioeconomically disadvantaged or digitally marginalized groups, leading to a biased understanding of support dynamics in online communities.

Accordingly, future research should explore strategies to mitigate these limitations by examining the representativeness of the user base and incorporating voices from underrepresented groups. For example, comparing findings from active participants with those of passive users or supplementing online data with offline sources could provide a more comprehensive view of support mechanisms. In addition, understanding how users of online support communities differ from those who might use AI-driven tools for AD/ADRD management is critical. Exploring the overlap and divergence between these groups will help inform the design of inclusive and effective digital interventions. Finally, the emergence of generative AI and large language models has enabled the development of personalized assistants and chatbots [55,56]. It would be interesting to examine how these AI chatbots respond to people's queries on AD/ADRD and how effective they are compared to online communities.

Comparison With Previous Work

Previous research has highlighted the significant burden imposed by AD/ADRD on patients, caregivers, and health care systems [54,57,58]. These caregivers often face stigma due to societal misconceptions about caregiving, which can result in judgment and disapproval, further intensifying their feelings of isolation and stress [59,60]. The growing demand for caregiving support due to the neurodegenerative progression of AD/ADRD necessitates increased levels of care over time that by 2030, an additional 1.2 million direct care workers will be required in the United States to address the rising needs of this population, pointing to a substantial gap in the current AD/ADRD caregiving infrastructure [2]. Our findings offer empirical evidence on the specific needs and concerns of individuals affected by AD/ADRD as they seek and exchange support within online communities.

Prior research has also underscored the psychosocial burdens faced by AD/ADRD caregivers, including feelings of marginalization and neglect of well-being [60,61]. A total of 69% of caregivers experience moderate to severe levels of caregiving burden, with the severity of the patient's condition significantly impacting their stress levels [62]. Our study expands on this by underlining the specific concerns expressed in online communities and revealing that discussions frequently center on memory care, nursing, caregiving challenges, and legal and financial issues. This suggests that caregivers seek both practical advice as well as ES to manage the complexities of their roles.

A notable contribution of our research is its focus on the role of online platforms as a source of support for AD/ADRD caregivers. Previous studies have acknowledged the use of digital tools among caregivers [63] and highlighted a disconnect between the availability of digital support and the recognition of caregiving efforts within social circles [64]. Our findings on the major concerns and needs—as identified by topic modeling—also align with recent computational and qualitative analyses of Reddit and other online communities for AD/ADRD caregivers, which have identified similar themes, including ES-seeking, decision-making about care facilities, and management of behavioral symptoms [23,27]. Therefore, the integration of individuals with AD/ADRD and caregivers' needs into the development of health and wellbeing technologies is necessary for future design [9,64]. Our study contributes to the recent body of work on the role of technologies in supporting caregivers—Bhat et al [65] emphasized the crucial role of caregiver-focused technological supports, and Kim et al [66] highlighted the need for tailored support strategies caregivers need during different phases of challenging behavioral episodes, and Meyerhoff et al [67] stressed the importance of user-centered digital mental health tools that adapt to individual needs.

The high prevalence of IS identified in our study aligns with existing literature that suggests caregivers often turn to digital platforms to fill gaps in health care services, particularly due to inconsistent coordination of long-term care [4]. The distribution of ES and IS across topics highlights a nuanced understanding of how caregivers use online communities to address specific facets of the caregiving experience. Aligning with how previous work focused on the use of technology for caregiving support [3,4], our study revealed the unique dynamics of social media as a tool for fostering a sense of belonging, solidarity, and social support among AD/ADRD caregivers. These observations align with findings from studies by Rains et al [68] and Naslund et al [69], who noted that individuals often leverage online spaces to combat loneliness and build supportive relationships. Together, our study builds upon the foundation laid by the SSBC, which categorizes support into 5 types—informational, emotional, esteem, tangible, and social network support [24]. Although this study focused specifically on ES and IS, we observed overlaps with other forms of support as well.

XSL•FO
RenderX

Conclusions

This study reveals that online communities centered on AD/ADRD engage in extensive discussions on a diverse array of topics. These topics include memory care strategies, nursing and caregiving practices, as well as legal, financial, and emotional challenges associated with managing the disease. The findings highlight not only the primary concerns and pain points experienced by caregivers and families who manage AD/ADRD in their homes but also the critical role online platforms play in providing them with a sense of community, practical advice, and ES. By analyzing these discussions, the study identifies how individuals affected by AD/ADRD turn to online communities to share personal experiences, seek solutions to caregiving dilemmas, and obtain validation and reassurance from others who are navigating similar challenges. These online interactions can alleviate feelings of isolation, foster a collaborative environment for exchanging best practices, and serve as an essential resource for information and guidance. Our findings underscore the urgent need for targeted institutional and social interventions that address the nuanced and evolving needs of AD/ADRD patients, their caregivers, and their families. These interventions could include enhanced support services, caregiver education programs, and accessible mental health resources that are tailored to the specific challenges highlighted in these online discussions. By recognizing and addressing these needs, stakeholders can improve the quality of life for both patients and caregivers, ultimately contributing to a more supportive and inclusive environment for those affected by AD/ADRD.

Conflicts of Interest

None declared.

References

1. Montenegro JMF. Alzheimer's disease diagnosis based on cognitive methods in virtual environments and emotions analysis. arXiv:1810.10941 2018.
2. NA. 2023 alzheimer's disease facts and figures. Alzheimers Dement 2023;19(4):1598-1695. [doi: 10.1002/alz.13016] [Medline: 36918389]
3. Rosebush CE, Stabler H, Nkimbeng M, Louwagie K, Fields NL, Jutkowitz E, et al. The porchlight project: A pilot study to adapt the senior companion program to enhance memory care services and supports. Gerontol Geriatr Med 2021;7:23337214211017651 [FREE Full text] [doi: 10.1177/23337214211017651] [Medline: 34036120]
4. Borson S, Chodosh J. Developing dementia-capable health care systems: a 12-step program. Clin Geriatr Med 2014;30(3):395-420. [doi: 10.1016/j.cger.2014.05.001] [Medline: 25037288]
5. Farhana N, Peckham A, Marani H, Roerig M, Marchildon G. The social construction of dementia: implications for healthcare experiences of caregivers and people living with dementia. J Patient Exp 2023;10:23743735231211066 [FREE Full text] [doi: 10.1177/23743735231211066] [Medline: 38026064]
6. Klimova B, Maresova P, Valis M, Hort J, Kuca K. Alzheimer's disease and language impairments: social intervention and medical treatment. Clin Interv Aging 2015;10:1401-1407 [FREE Full text] [doi: 10.2147/CIA.S89714] [Medline: 26346123]
7. El Haj M, Jardri R, Larøi F, Antoine P. Hallucinations, loneliness, and social isolation in alzheimer's disease. Cogn Neuropsychiatry 2016;21(1):1-13. [doi: 10.1080/13546805.2015.1121139] [Medline: 26740416]
8. Sun Y, Ji M, Leng M, Li X, Zhang X, Wang Z. Comparative efficacy of 11 non-pharmacological interventions on depression, anxiety, quality of life, and caregiver burden for informal caregivers of people with dementia: A systematic review and network meta-analysis. Int J Nurs Stud 2022;129:104204. [doi: 10.1016/j.ijnurstu.2022.104204] [Medline: 35247788]
9. Meiland F, Innes A, Mountain G, Robinson L, van der Roest H, García-Casal JA, et al. Technologies to support community-dwelling persons with dementia: a position paper on issues regarding development, usability, effectiveness and cost-effectiveness, deployment, and ethics. JMIR Rehabil Assist Technol 2017;4(1):e1 [FREE Full text] [doi: 10.2196/rehab.6376] [Medline: 28582262]
10. Berridge C, Turner NR, Liu L, Karras SW, Chen A, Fredriksen-Goldsen K, et al. Advance planning for technology use in dementia care: development, design, and feasibility of a novel self-administered decision-making tool. JMIR Aging 2022;5(3):e39335 [FREE Full text] [doi: 10.2196/39335] [Medline: 35896014]
11. Rettinger L, Zeuner L, Werner K, Ritschl V, Mosor E, Stamm T, et al. A mixed-methods evaluation of a supporting app for informal caregivers of people with dementia. 2020 Presented at: PETRA '20: The 13th PErvasive Technologies Related to Assistive Environments Conference; 2020 June 30-3 July; Corfu Greece p. 1-8. [doi: 10.1145/3389189.3397981]
12. Queluz FNFR, Kervin E, Wozney L, Fancey P, McGrath PJ, Keefe J. Understanding the needs of caregivers of persons with dementia: a scoping review. Int Psychogeriatr 2020;32(1):35-52 [FREE Full text] [doi: 10.1017/S1041610219000243] [Medline: 30967164]
13. De Choudhury M, De S. Mental health discourse on reddit: self-disclosure, social support, and anonymity. ICWSM 2014;8(1):71-80. [doi: 10.1609/icwsm.v8i1.14526]
14. Ernala SK, Rizvi AF, Birnbaum ML, Kane JM, De Choudhury M. Linguistic markers indicating therapeutic outcomes of social media disclosures of schizophrenia. Proc. ACM Hum.-Comput. Interact 2017;1(CSCW):1-27. [doi: 10.1145/3134678]

XSL•FO
RenderX

15. Jin Y, Chandra M, Verma G, Hu Y, Choudhury MD, Kumar S. Better to ask in english: cross-lingual evaluation of large language models for healthcare queries. 2024 Presented at: WWW '24: Proceedings of the ACM Web Conference 2024; 2024 May 13 - 17; Singapore p. 2627-2638. [doi: 10.1145/3589334.3645643]
16. Yang D, Yao Z, Kraut R. Self-disclosure and channel difference in online health support groups. ICWSM 2017;11(1):704-707. [doi: 10.1609/icwsm.v11i1.14935]
17. Levonian Z, Erikson DR, Luo W, Narayanan S, Rubya S, Vachher P, et al. Bridging qualitative and quantitative methods for user modeling: tracing cancer patient behavior in an online health community. ICWSM 2020;14:405-416. [doi: 10.1609/icwsm.v14i1.7310]
18. Saha K, Ernala SK, Dutta S, Sharma E, De Choudhury M. Understanding moderation in online mental health communities. : Springer; 2020 Presented at: 12th International Conference, SCSM 2020, Held as Part of the 22nd HCI International Conference, HCII 2020; 2020 July 19–24; Copenhagen, Denmark. [doi: 10.1007/978-3-030-49576-3_7]
19. Olteanu A, Varol A, Kiciman E. Distilling the outcomes of personal experiences: A propensity-scored analysis of social media. 2017 Presented at: CSCW '17: Computer Supported Cooperative Work and Social Computing; 2017 25 February; Portland Oregon USA p. 370-386. [doi: 10.1145/2998181.2998353]
20. Yang D, Yao Z, Seering J, Kraut R. The channel matters: self-disclosure, reciprocity and social spport in online cancer spport groups. 2019 Presented at: CHI '19: CHI Conference on Human Factors in Computing Systems; 2019 May 4 - 9; Glasgow Scotland Uk p. 1-15. [doi: 10.1145/3290605.3300261]
21. Lazar A, Brewer RN, Kacorri H, Hong J, Punzalan MND, Mahathir M, et al. How content authored by people with dementia affects attitudes towards dementia. Proc ACM Hum Comput Interact 2021;5 [FREE Full text] [doi: 10.1145/3479542] [Medline: 35187410]
22. Yoon S, Lucero R, Mittelman MS, Luchsinger JA, Bakken S. Mining twitter to inform the design of online interventions for hispanic alzheimer's disease and related dementias caregivers. Hisp Health Care Int 2020;18(3):138-143. [doi: 10.1177/1540415319882777] [Medline: 31646904]
23. Pickett AC, Valdez D, Sinclair KL, Kochell WJ, Fowler B, Werner NE. Social media discourse related to caregiving for older adults living with alzheimer disease and related dementias: computational and qualitative study. JMIR Aging 2024;7:e59294 [FREE Full text] [doi: 10.2196/59294] [Medline: 38896462]
24. Suhr JA, Cutrona CE, Krebs KK, Jensen SL. The social support behavior code (ssbc). In: Couple Observational Coding Systems. United Kingdom: Routledge; 2004:307-318.
25. Andalibi N, Haimson OL, Choudhury MD, Forte A. Social support, reciprocity, and anonymity in responses to sexual abuse disclosures on social media. ACM Trans. Comput.-Hum. Interact 2018;25(5):1-35. [doi: 10.1145/3234942]
26. Saha K, Kim SC, Reddy MD, Carter AJ, Sharma E, Haimson OL, et al. The language of LGBTQ+ minority stress experiences on social media. Proc ACM Hum Comput Interact 2019;3(CSCW):1-22 [FREE Full text] [doi: 10.1145/3361108] [Medline: 32935081]
27. Pickett AC, Valdez D, White LA, Loganathar P, Linden A, Boutilier JJ, et al. The careVirtue digital journal for family and friend caregivers of people living with alzheimer disease and related dementias: exploratory topic modeling and user engagement study. JMIR Aging 2024;7:e67992 [FREE Full text] [doi: 10.2196/67992] [Medline: 39719081]
28. r/alzheimersgroup. 2023. URL: https://knowyourmeme.com/memes/sites/ralzheimersgroup [accessed 2023-12-10]
29. Chancellor S, Lin Z, Goodman EL, Zerwas S, Choudhury MD. Quantifying and predicting mental illness severity in online pro-eating disorder communities. 2016 Presented at: CSCW '16: Computer Supported Cooperative Work and Social Computing; 2016 27 February-2 March; San Francisco California USA p. 1171-1184. [doi: 10.1145/2818048.2819973]
30. Zhou J, Saha K, Lopez Carron IM, Yoo DW, Deeter CR, De Choudhury M, et al. Veteran critical theory as a lens to understand veterans' needs and support on social media. Proc. ACM Hum.-Comput. Interact 2022;6(CSCW1):1-28. [doi: 10.1145/3512980]
31. Gong C, Saha K, Chancellor S. "The Smartest Decision for My Future": social media reveals challenges and stress during post-college life transition. Proc. ACM Hum.-Comput. Interact 2021;5(CSCW2):1-29. [doi: 10.1145/3476039]
32. Mehrotra R, Sanner S, Buntine W, Xie L. Improving LDA topic models for microblogs via tweet pooling and automatic labeling. 2013 Presented at: SIGIR '13: The 36th International ACM SIGIR conference on research and development in Information Retrieval; 2013 28 July-1 August; Dublin Ireland p. 889-892. [doi: 10.1145/2484028.2484166]
33. Blei DM, Ng AY, Jordan MI. Latent Dirichlet Allocation. Cambridge, Massachusetts: MIT Press; 2003:993-1022.
34. Wallach H, Murray I, Salakhutdinov R, Mimno D. Evaluation methods for topic models. 2009 Presented at: ICML '09: The 26th Annual International Conference on Machine Learning held in conjunction with the 2007 International Conference on Inductive Logic Programming; 2009 June 14 - 18; Montreal Quebec Canada p. 1105-1112. [doi: 10.1145/1553374.1553515]
35. Kummervold PE, Gammon D, Bergvik S, Johnsen JK, Hasvold T, Rosenvinge JH. Social support in a wired world: use of online mental health forums in Norway. Nord J Psychiatry 2002;56(1):59-65. [doi: 10.1080/08039480252803945] [Medline: 11869468]
36. Rains S, Young V. A meta-analysis of research on formal computer-mediated support groups: examining group characteristics and health outcomes. Hum. Communication Research 2009;35(1):309-336. [doi: 10.1111/j.1468-2958.2009.01353.x]

XSL•FO
RenderX

37. Cutrona CE, Suhr JA. Controllability of stressful events and satisfaction with spouse support behaviors. Communication Research 1992;19(2):154-174. [doi: 10.1177/009365092019002002]
38. Andalibi N, Haimson OL, Choudhury MD, Forte A. Understanding social media disclosures of sexual abuse through the lenses of support seeking and anonymity. 2016 Presented at: CHI'16: CHI Conference on Human Factors in Computing Systems; 2006 May 7 - 12; San Jose California USA p. 3908-3916. [doi: 10.1145/2858036.2858096]
39. De Choudhury M, Kiciman E. The language of social support in social media and its effect on suicidal ideation risk. ICWSM 2017;11(1):32-41. [doi: 10.1609/icwsm.v11i1.14891]
40. Sharma E, Choudhury MD. Mental health support and its relationship to linguistic accommodation in online communities. : ACM; 2018 Presented at: CHI '18: CHI Conference on Human Factors in Computing Systems; 2018 April 21 - 26; Montreal QC Canada p. 1-13. [doi: 10.1145/3173574.3174215]
41. Saha K, Sharma A. Causal factors of effective psychosocial outcomes in online mental health communities. ICWSM 2020;14:590-601. [doi: 10.1609/icwsm.v14i1.7326]
42. Saha K, Torous J, Caine ED, De Choudhury M. Psychosocial effects of the COVID-19 pandemic: large-scale quasi-experimental study on social media. J Med Internet Res 2020;22(11):e22600 [FREE Full text] [doi: 10.2196/22600] [Medline: 33156805]
43. Wang Y, Kraut R, Levine JM. To stay or leave?: the relationship of emotional and informational support to commitment in online health support groups. 2012 Presented at: CSCW '12: Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work; 2012 February 11 - 15; Seattle Washington USA p. 833-842. [doi: 10.1145/2145204.2145329]
44. Torrey L, Shavlik J. Transfer learning. In: Handbook of research on machine learning applications and trends: algorithms, methods, and techniques. Hershey, Pennsylvania, USA: IGI global; 2010:242-264.
45. Jahn H. Memory loss in alzheimer's disease. Dialogues Clin Neurosci 2013;15(4):445-454 [FREE Full text] [doi: 10.31887/DCNS.2013.15.4/hjahn] [Medline: 24459411]
46. Calkins MP. Memory care and alzheimer's units. In: Environmental Psychology and Human Well-Being. Amsterdam, Netherlands: Elsevier; 2018:365-386.
47. Rolland Y, Pillard F, Klapouszczak A, Reynish E, Thomas D, Andrieu S, et al. Exercise program for nursing home residents with alzheimer's disease: a 1-year randomized, controlled trial. J Am Geriatr Soc 2007;55(2):158-165. [doi: 10.1111/j.1532-5415.2007.01035.x] [Medline: 17302650]
48. White M, Dorman SM. Receiving social support online: implications for health education. Health Educ Res 2001;16(6):693-707. [doi: 10.1093/her/16.6.693] [Medline: 11780708]
49. Moore D, Drey N, Ayers S. Use of online forums for perinatal mental illness, stigma, and disclosure: an exploratory model. JMIR Ment Health 2017;4(1):e6 [FREE Full text] [doi: 10.2196/mental.5926] [Medline: 28219879]
50. Taggart T, Grewe ME, Conserve DF, Gliwa C, Roman Isler M. Social media and HIV: a aystematic review of uses of social media in HIV communication. J Med Internet Res 2015;17(11):e248 [FREE Full text] [doi: 10.2196/jmir.4387] [Medline: 26525289]
51. Soong A, Au ST, Kyaw BM, Theng YL, Tudor Car L. Information needs and information seeking behaviour of people with dementia and their non-professional caregivers: a scoping review. BMC Geriatr 2020;20(1):61 [FREE Full text] [doi: 10.1186/s12877-020-1454-y] [Medline: 32059648]
52. Kernisan LP, Sudore RL, Knight SJ. Information-seeking at a caregiving website: a qualitative analysis. J Med Internet Res 2010;12(3):e31 [FREE Full text] [doi: 10.2196/jmir.1548] [Medline: 20675292]
53. Blum K. Family caregivers. Palliat. care nursing. In: Qual. Care to End Life. Cham: Springer; 2010:187-198.
54. Schulz R, Sherwood PR. Physical and mental health effects of family caregiving. Journal of Social Work Education 2017;44(sup3):105-113. [doi: 10.5175/jswe.2008.773247702]
55. Xu L, Sanders L, Li K, Chow JCL. Chatbot for health care and oncology applications using artificial intelligence and machine learning: systematic review. JMIR Cancer 2021;7(4):e27850 [FREE Full text] [doi: 10.2196/27850] [Medline: 34847056]
56. Goodman S, Buehler E, Clary P, Coenen A, Donsbach A, Horne TN, et al. Lampost: design and evaluation of an ai-assisted email writing prototype for adults with dyslexia. 2022 Presented at: ASSETS '22: Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility; 2022 October 23 - 26; Athens Greece p. 1-8. [doi: 10.1145/3517428.3544819]
57. Gustavsson A, Norton N, Fast T, Frölich L, Georges J, Holzapfel D, et al. Global estimates on the number of persons across the alzheimer's disease continuum. Alzheimers Dement 2023;19(2):658-670. [doi: 10.1002/alz.12694] [Medline: 35652476]
58. Prince M. World Alzheimer report 2015. In: The Global Impact of Dementia: An Analysis of Prevalence, Incidence, Cost and Trends. London: Alzheimer's Disease International; 2015.
59. GBD 2019 Dementia Forecasting Collaborators. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the global burden of disease study 2019. Lancet Public Health 2022;7(2):e105-e125 [FREE Full text] [doi: 10.1016/S2468-2667(21)00249-8] [Medline: 34998485]
60. Corrigan PW, Watson AC, Miller FE. Blame, shame, and contamination: the impact of mental illness and drug dependence stigma on family members. J Fam Psychol 2006;20(2):239-246. [doi: 10.1037/0893-3200.20.2.239] [Medline: 16756399]

XSL•FO
RenderX

61. Werner P, Goldstein D, Buchbinder E. Subjective experience of family stigma as reported by children of alzheimer's disease patients. Qual Health Res 2010;20(2):159-169. [doi: 10.1177/1049732309358330] [Medline: 20065304]
62. Pudelewicz A, Talarska D, Bączyk G. Burden of caregivers of patients with alzheimer's disease. Scand J Caring Sci 2019;33(2):336-341. [doi: 10.1111/scs.12626] [Medline: 30378698]
63. Siddiqui F, Varghese D, Singh P, Bayyavarapu SB, Lindsay S, Chandrasekara D, et al. Exploring the digital support needs of caregivers of people with serious mental illness. 2023 Presented at: CHI '23: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems; 2023 April 23 - 28; Hamburg Germany p. 1-16. [doi: 10.1145/3544548.3580674]
64. Chen Y, Ngo V, Park SY. Caring for caregivers: designing for integrality. 2013 Presented at: CSCW '13: Computer Supported Cooperative Work; February 23 - 27 2013; San Antonio Texas USA p. 91-102. [doi: 10.1145/2441776.2441789]
65. Bhat KS, Hall AK, Kuo T, Kumar N. "We are half-doctors": family caregivers as boundary actors in chronic disease management. Proc. ACM Hum.-Comput. Interact 2023;7(CSCW1):1-29. [doi: 10.1145/3579545]
66. Kim T, Kim H, Roberts AC, Jacobs M, Kay M. Opportunities in mental health support for informal dementia caregivers suffering from verbal agitation. Proc. ACM Hum.-Comput. Interact 2024;8(CSCW1):1-26. [doi: 10.1145/3637381]
67. Meyerhoff J, Kornfield R, Mohr DC, Reddy M. Meeting yung adults' social support needs across the health behavior change journey: implications for digital mental health tools. Proc ACM Hum Comput Interact 2022;6(CSCW2):1-33 [FREE Full text] [doi: 10.1145/3555203] [Medline: 36387059]
68. Rains SA. Coping with Illness Digitally. Cambridge, Massachusetts: MIT Press; 2018.
69. Naslund JA, Grande SW, Aschbrenner KA, Elwyn G. Naturally occurring peer support through social media: the experiences of individuals with severe mental illness using YouTube. PLoS One 2014;9(10):e110171 [FREE Full text] [doi: 10.1371/journal.pone.0110171] [Medline: 25333470]

Abbreviations

AD/ADRD: Alzheimer Disease and Related Dementias
AD: Alzheimer Disease
AI: artificial intelligence
ES: emotional support
IS: informational support
LDA: latent Dirichlet allocation
POA: power of attorney
RQ: research question
SSBC: social support behavioral code
SVM: support vector machine

Edited by A Garcia; submitted 18.11.24; peer-reviewed by AC Pickett; comments to author 11.02.25; revised version received 04.03.25; accepted 07.03.25; published 05.05.25.

Please cite as:
Kaliappan S, Liu C, Jain Y, Karkar R, Saha K
Online Communities as a Support System for Alzheimer Disease and Dementia Care: Large-Scale Exploratory Study
JMIR Aging 2025;8:e68890
URL: https://aging.jmir.org/2025/1/e68890
doi:10.2196/68890
PMID:

©Sidharth Kaliappan, Chunyu Liu, Yoshee Jain, Ravi Karkar, Koustuv Saha. Originally published in JMIR Aging (https://aging.jmir.org), 05.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Analyzing Disparity in Geographical Accessibility to Home Medical Care Using a Claims Database and Geographical Information System: Simulation Study

Yasuhiro Morii[1], PhD; Yasuhiro Nakanishi[2], PhD; Yuichi Nishioka[3], PhD; Yukio Tsugihashi[3,4], PhD; Tatsuya Noda[3], PhD; Tomoya Myojin[3,5], PhD; Tomoaki Imamura[3], PhD; Manabu Akahane[2], PhD

[1]Center for Outcomes Research and Economic Evaluation for Health, National Institute of Public Health, Wako, Japan

[2]Department of Health and Welfare Services, National Institute of Public Health, 2-3-6 Minami, Wako, Japan

[3]Department of Public Health, Health Management and Policy, Nara Medical University, Kashihara, Japan

[4]Medical home care center, Tenri Hospital Shirakawa Branch, Tenri, Japan

[5]Department of Community Health and Preventive Medicine, Hamamatsu University School of Medicine, Hamamatsu, Japan

Corresponding Author:
Manabu Akahane, PhD
Department of Health and Welfare Services, National Institute of Public Health, 2-3-6 Minami, Wako, Japan

Abstract

Background: The demand for home medical care services has increased in aging societies. Therefore, allocating health care resources optimally to meet the needs of each community is essential. Geographical accessibility is an important factor affecting access to home medical care services; however, little research has been conducted on regional disparities in geographical accessibility.

Objective: This study aims to analyze the regional disparities in geographical accessibility to home medical care services using the Kokuho database (KDB), a comprehensive medical claims database for a prefecture in Japan.

Methods: This study included 39 municipalities in Nara Prefecture, Japan. Using a geographical information system, accessibility to home medical care services, that is, travel distance and time from hospitals and clinics to hypothetical patients, was analyzed in two scenarios: (1) an ideal scenario, where we assumed that all hospitals or clinics in Nara Prefecture provided those services and (2) an actual scenario, where hospitals or clinics in Nara Prefecture that actually provided home medical care services, identified from KDB data analysis, were used in the analysis. Hypothetical patients were randomly distributed on the geographical information system in accordance with the usage rates of home medical care services and with the distributions of the population aged ≥75 years. The usage rate by municipalities was aggregated from the analysis of KDB data of Nara Prefecture in FY2019.

Results: The median travel distance was longer than 16 km, the reference limit value specified in the Japanese fee table, and the median travel time exceeded 30 min in certain rural municipalities in the southern part of Nara Prefecture, in the actual scenario, whereas the travel distance and time were improved in the ideal scenario. The differences in travel time between the ideal and actual scenarios were the largest in the depopulated municipalities in the southern part, such as Totsukawa (32.6 vs 5.8 min), Kawakami (30.1 vs 11.8 min), Kurotaki (21.3 vs 5.2 min), and Kamikitayama (20.7 vs 3.5 min). The usage rates were also lower in rural municipalities in the southern part.

Conclusions: The results revealed that geographical accessibility was lower in depopulated municipalities in the southern part, and the disparity could be partly solved in the ideal scenario, especially in that area, highlighting the necessity of increasing supply in the southern areas. KDB is a comprehensive database that includes medical claims information for home medical care patients and details of the provision of medical institutions, enabling geographical analysis that reflects actual health care usage.

(JMIR Aging 2025;8:e70040) doi:10.2196/70040

KEYWORDS

home medical care; geographical accessibility; geographical information system; medical claims database; simulation

Introduction

Most limited-income countries have aging societies, including Japan, which is the most aging country in the world with an aging rate (rate of population aged ≥65 y) of 28.1% in 2018, followed by Italy (23.3%), Portugal (21.9%), and Germany (21.7%) [1]. Medical demand is estimated to increase with demographic changes [2]. The Japanese government has been promoting a structure in which home medical care and long-term care services are provided comprehensively to older adults,

XSL•FO
RenderX

ensuring they live their lives in the areas they are used to in an aging society [3]. Therefore, increased promotion of the use of home medical care services is of great societal importance.

The demand for home medical care services has also increased in recent years, with a peak at around 2040 [3]. Each prefecture in Japan is responsible for creating a Regional Medical Plan to ensure that health care services are effectively provided to meet the needs of the local population [4]. The Regional Medical Plan includes a plan to provide home medical care services, and each prefecture must specify the problems faced and measures to resolve them. Prefectures must consider optimal allocations and make the most of their limited resources to meet the needs.

Geographical accessibility is among the factors that could affect the provision of home medical services. The relationship between geographical factors and the usage of medical services has been previously studied [5-7]. For example, Schwarz et al [5] analyzed the relationship between the distance from hospitals to patients and the psychiatric treatments that patients received. Herein, people living in peripheral communities were mainly treated in inpatient settings. In contrast, those living in residential areas were more likely to receive intensive psychiatric home treatment. Similar to their study findings, geographical accessibility is important for home medical care services since health care providers need to travel to patients' homes.

Kuwayama et al [8] conducted a literature review of studies on home care nursing in Japan using a geographical information system (GIS) and discovered a few related studies [9,10]. In those studies, Naruse et al [7] performed a multilevel logistic analysis of the association between the use of home care nursing and the proportion of older adults reachable by facilities providing home-visit nursing services (calculated by buffer analysis using GIS). They found that people living in less reachable municipalities were less likely to use these services. However, few studies have been conducted on the geographical accessibility of home-based medical care services. Moreover, these previous studies were limited to visualizing the current situation.

Extensive databases containing medical claims and national health insurance information, such as the National Database of Health Insurance Claims and Specific Health Checkups of Japan (NDB) and National Health Insurance Database (Kokuho database: KDB), have recently become essential tools for medical and health care policy research. These real-world data sources provide valuable insights into health care usage, treatment patterns, and outcomes. The main advantages of using such databases include the ability to study targeted populations and understand the actual clinical practice comprehensively. Using such databases, a geographical analysis that reflects the actual health care usage and trends will be possible.

Several studies have conducted geographical analysis using a large real-world database. Cheng et al [11] analyzed the prescribing patterns of cardiovascular drugs across townships in Taiwan using a national claim database covering 97.5% of the total population. Kwak et al [12] analyzed the relationship between the incidence of obstetric complications for pregnant women and access time to delivery units in South Korea, using the national claim database.

However, there are few previous studies analyzing geographical accessibility to home medical care services using an extensive database.

Therefore, this study aimed to develop a methodology for analyzing regional disparities in geographical accessibility to home medical care services using the KDB database to help consider achieving the provision of more efficient and equitable service. Given the global trend of aging, the findings of this study in Japan may contribute to improving the efficiency and equity of service delivery and policymaking of home medical care services in other countries facing similar demographic challenges.

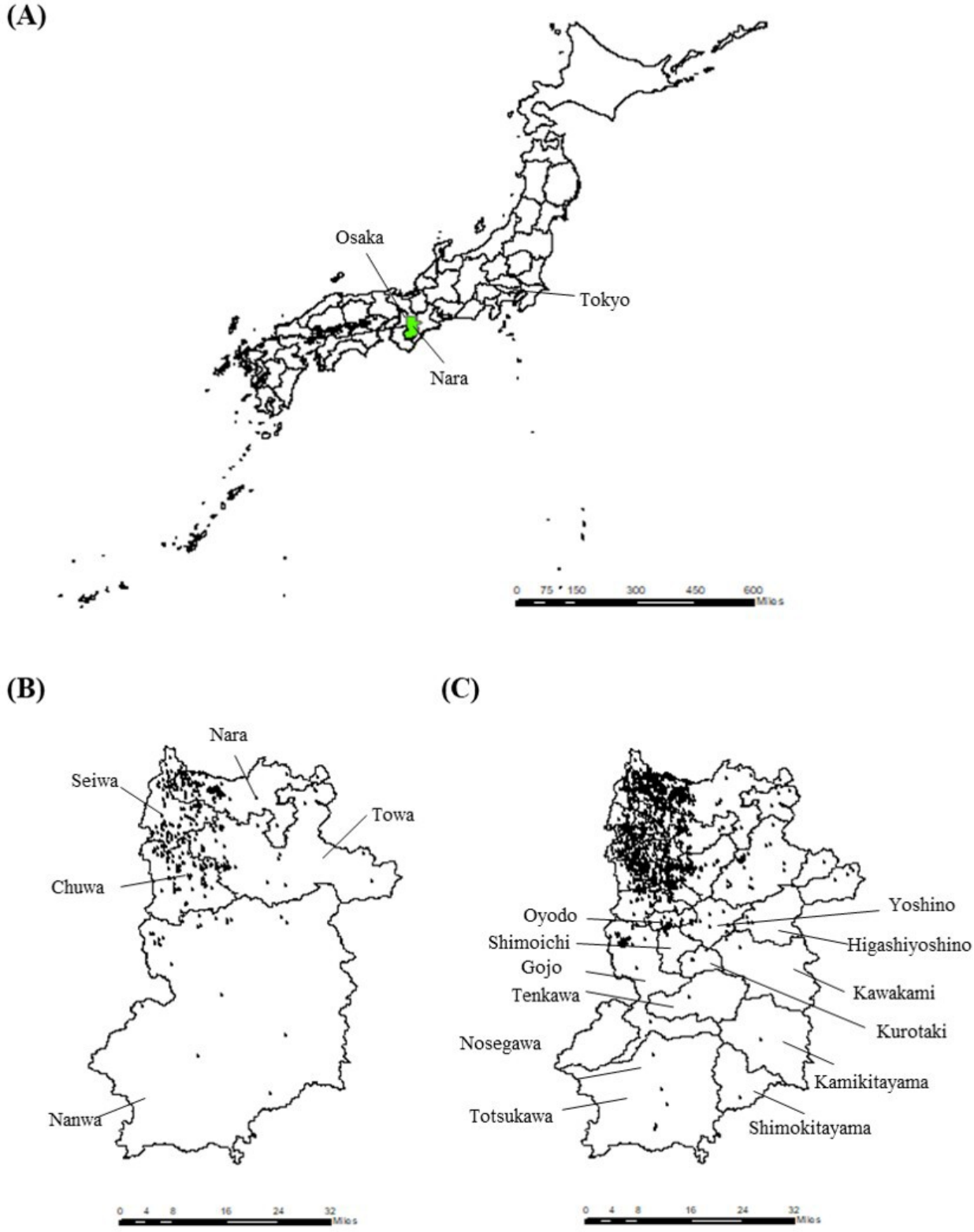
Methods

Subjects and Scheme

This study was conducted in the 39 municipalities in Nara Prefecture, Japan, as a model region. Nara Prefecture is located in the center of Japan's main island of Honshu (Figure 1A) and has both urban and rural (mountainous) areas. Nara Prefecture has a population of 1,286,651 (approximately 1/100th of the national total), a land area of 3691 km^2 (approximately 1/100th of the national total) [13,14], and 32% (0.4/1.25) of adults aged ≥65 years (compared to 36.4/125.52, 29% nationally; Table 1). The characteristics of municipalities, such as population and aging rate, are shown in Table 1. Figures 1B and C illustrate the geographical locations of the 5 secondary medical areas and 39 municipalities. Nanwa secondary area, in the southern part, is the most rural area with mountains. Since there seems to be a geographical difference within Nara Prefecture, it would be suitable as a model region for geographical analysis. Secondary medical areas provide general inpatient medical services, whereas municipalities or smaller regional units provide more basic medical services.

XSL•FO
RenderX

Figure 1. The geographical locations of Nara Prefecture in Japan, secondary medical areas in Nara, and municipalities in Nara show the geographical location of (A) Nara Prefecture in Japan and also illustrate the distributions of hospitals or clinics (B) in the actual scenario and (C) in the ideal scenario.



XSL•FO
RenderX

Table . Characteristics of each municipality and the result of usage rate.

Municipality	Population [14]	Aging rate (%) [14]	Land area (km^2) [13]	Population density ($/km^2$)	Number of claims	Usage rate (%)
Nara	347,947	32.5	276.9	1256.40	34,335	6.8
Kashihara	118,344	30	39.6	2991.50	5370	3.9
Ikoma	114,508	30	53.2	2154.40	7755	5.8
Yamatokoriyama	80,760	34.4	42.7	1891.80	5135	4.6
Kashiba	77,212	24.8	24.3	3182.70	3931	5.4
Tenri	60,717	27.9	86.4	702.6	4160	5.8
Yamatotakada	59,797	32.9	16.5	3628.50	3838	4.4
Sakurai	52,975	33.1	98.9	535.6	3812	5.4
Katsuragi	37,199	28.1	33.7	1103.20	1424	3.4
Koryo	33,910	27.3	16.3	2080.40	1154	3.6
Tawaramoto	30,840	32.6	21.1	1462.30	1497	4.9
Ikaruga	27,408	30.7	14.3	1920.70	2108	6.0
Uda	25,732	44.5	247.5	104	2103	4.7
Gojo	25,487	42.4	292	87.3	1769	4.2
Oji	23,667	28.9	7	3376.20	1518	5.4
Sango	22,688	33	8.8	2581.10	2018	7.2
Gose	22,259	43.9	60.6	367.4	1096	2.6
Kanmaki	20,815	37	6.1	3390.10	1586	5.3
Heguri	17,545	39.9	23.9	734.1	2484	7.9
Kawai	16,314	40.2	8.2	1982.30	1523	5.3
Oyodo	15,530	37.5	38.1	407.6	381	2.2
Kawanishi	7716	36	5.9	1301.20	457	3.6
Ando	6930	37.2	4.3	1607.90	439	4.9
Takatori	6235	43.6	25.8	241.8	274	2.5
Miyake	6097	37.2	4.1	1501.70	457	4.7
Yoshino	5457	54	95.7	57.1	407	2.7
Asuka	4767	43.2	24.1	197.8	238	3.1
Shimoichi	4356	49.7	62	70.3	154	2.0
Yamazoe	2909	51.6	66.5	43.7	289	4.4
Totsukawa	2695	43.4	672.4	4	254	3.6
Higashiyoshino	1337	60.5	131.7	10.2	224	5.1
Mitsue	1298	61.2	79.6	16.3	50	1.6
Soni	1175	53.5	47.8	24.6	140	5.0
Tenkawa	1048	53.3	175.7	6	60	1.5
Kawakami	1039	56.1	269.3	3.9	112	2.7
Shimokitayama	690	47.9	133.4	5.2	39	2.1
Kurotaki	536	54	47.7	11.2	24	0.9
Kamikitayama	384	49.6	274.2	1.4	37	2.6
Nosegawa	328	51.6	154.9	2.1	62	10.5
Total	1,286,651	32.7	3690.9	348.6	92,714	5.3

XSL•FO
RenderX

For each municipality, geographical accessibility to home medical care services was analyzed using GIS (ArcGIS Desktop version 10.8.1; ESRI Japan Corporation) in the following two scenarios: (1) actual scenario: geographical accessibility to patients from hospitals and clinics in Nara Prefecture that actually provided home medical care services, and (2) ideal scenario: geographical accessibility to patients from all hospitals or clinics in Nara Prefecture was used. The latter simulates geographical accessibility in an ideal scenario in which the number of hospitals or clinics that provide home medical care services would increase. These scenarios were different only in terms of the number of hospitals and clinics used in the analysis, as described hereafter. Hypothetical patients were generated on a GIS by using the methodology described in a subsequent section "Geographical Analysis."

Data

For the actual scenario, hospitals and clinics that provided home medical care services were defined using data from the KDB database of Nara Prefecture (Nara KDB) in FY2019. This is a prefecture-wide database of medical and long-term care insurance claims. To be defined as hospitals and clinics that actually provided home medical care services, hospitals and clinics were required to have at least one claim with home medical care service fees (medical service code: 114001110, 114030310, 114042110, 114042210, 114042810, 114046310, 114027710, and 114027810 in the national fee schedule) for patients aged ≥75 years in Nara KDB in FY2019. Consequently, 346 hospitals and clinics were identified in the actual scenario.

The ideal scenario used all the hospitals and clinics in Nara prefecture, that is, those providing or not providing home medical care services. The geographical data on the 1308 hospitals and clinics were obtained from the Geospatial Information Authority of Japan [15]. Dental clinics were excluded from the study. The distribution of hospitals and clinics in the 2 scenarios is shown in Figure 1.

Population data at each 500 m^2 mesh (ESRI Japan Corporation) were used to distribute the locations of the hypothetical patients.

Geographical Analysis

First, hypothetical patients were generated on GIS. The number of patients in each municipality was that of patients aged ≥75 years who used home medical care services in each municipality, identified from the Nara KDB data analysis. The usage of home medical care was defined in the same manner as identifying the hospitals and clinics in the actual scenario. That number of patients was randomly allocated to 500 m^2 meshes on GIS, reflecting the distribution of the population aged ≥75 years based on the national census 2020. In other words, the hypothetical patients were allocated to each mesh in a manner that the number of patients in each mesh is proportional to the mesh population aged ≥75 years using R version 4.1.2 (The R Foundation) [16,17]. This methodology enables the simulation of accessibility, reflecting the actual geographical distribution of potential patients. Thereafter, based on the patient allocations, hypothetical patient points were created on randomly determined points within the allocated meshes on ArcGIS using the "Create Random Points" function.

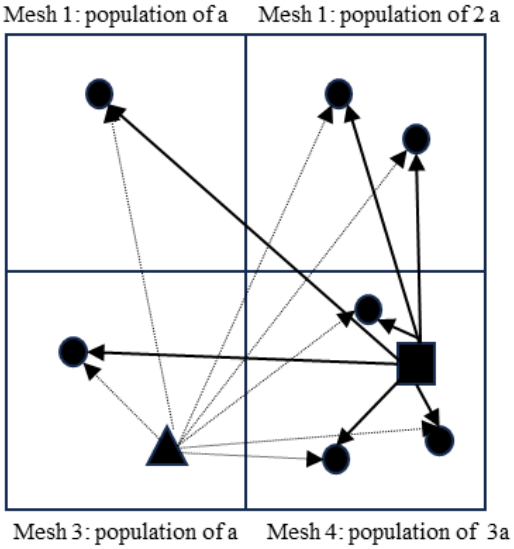
Subsequently, travel time from the patients to the hospitals or clinics was analyzed using the "Find Closest Facilities" function of ArcGIS for the 2 scenarios, respectively. We assumed that the travel to the patients was from the closest hospital or clinic. However, the closest hospitals or clinics can be located outside the municipality where the patients live.

The median and 25 - 75 percentiles of travel distance and time were aggregated for the patients in each municipality. Moreover, the rates of patients for whom travel distances were over 16 km were calculated since, in the Japanese universal fee table, the travel distance to provide home medical care services needs to be 16 km or less, unless an exceptional reason is present [18].

The usage rate was calculated by dividing the number of patients aged ≥75 years using home medical care services by the total population aged ≥75 years. The entire process of geographical analysis was repeated 10 times to consider the uncertainty of patient locations in each trial. The scheme of geographical analysis is visualized in Figure 2.

XSL•FO
RenderX

Figure 2. Scheme of the geographical analysis illustrating the scheme of the geographical analysis. GIS: geographical information system.



1. The hospitals or clinics defined to provide home medical care services (■) were identified from claims database analysis
2. All hospitals and clinics (both ▲ and ■) were identified from an open source
3. Locations of hypothetical patients were randomly decided in a manner to reflect actual distributions of population over 75 or more and usage rates for each municipality (using R 4.1.2) (eg. in a mesh with "2a" population likely has twice more patients than that with "a" population on average)
4. The hypothetical patients were mapped on GIS
5. Travel distance and time were analyzed using GIS
   - in ideal scenario: both ▲ and ■ were used
   - in actual scenario: only ■ was used (solid arrows)

Ethical Considerations

This study was approved by the ethics committee of the National Institute of Public Health, Japan (NIPH-IBRA number 12324 - 2). The requirement for informed consent was waived because all data were anonymized.

Results

The number of claims and usage rates are presented in Table 1. The distributions of hospitals or clinics used in each scenario were visualized in Figure 1. The travel time and distance results are shown in Figures 3-6 and Multimedia Appendices 1 and 2. In the actual scenario, the median travel time for all the patients in Nara Prefecture was 3.0 (IQR 2.0 - 4.3) minutes, and the travel distance was 0.7 (IQR 0.5 - 1.2) km. In the ideal scenario, the median travel time for all patients was 2.0 (IQR 1.2 - 2.9) minutes, and the travel distance was 0.5 (IQR 0.3 - 0.9) km.

XSL•FO
RenderX

Figure 3. The median travel time in the two scenarios. This figure illustrates the median travel time in (A) the ideal scenario and (B) the actual scenario.

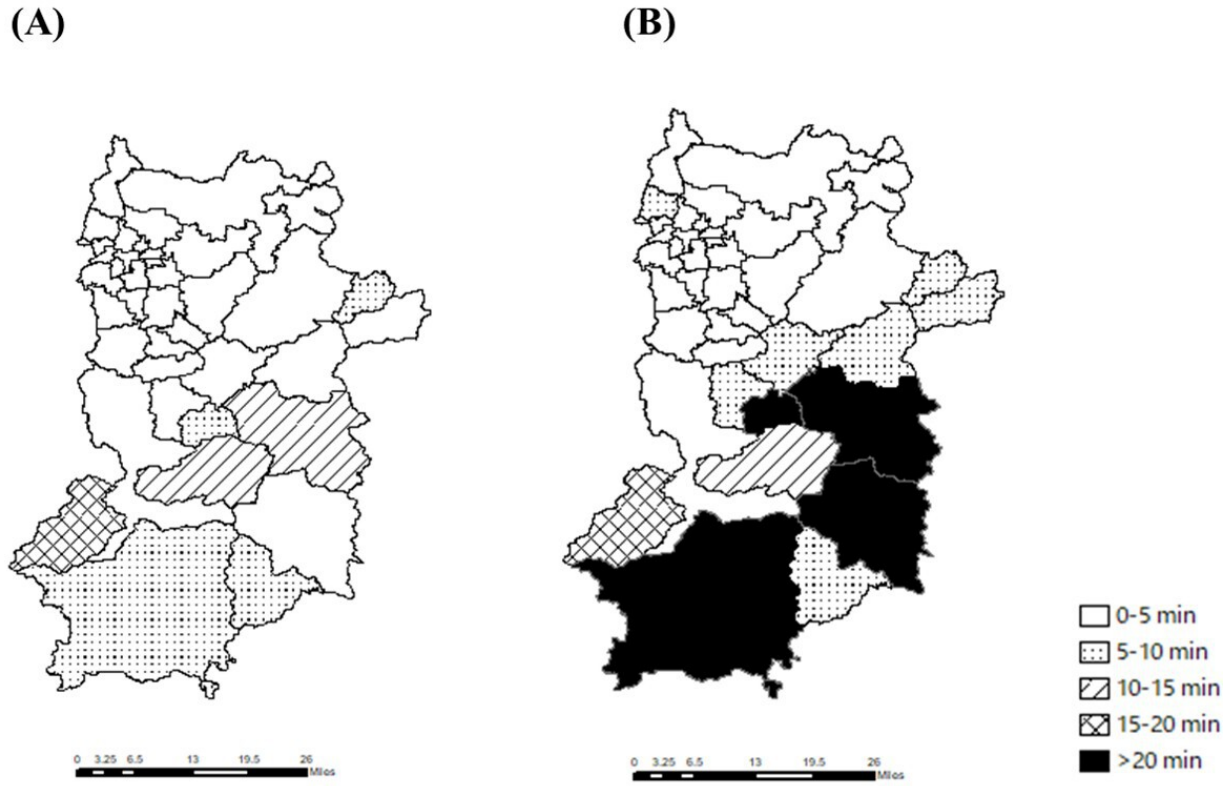
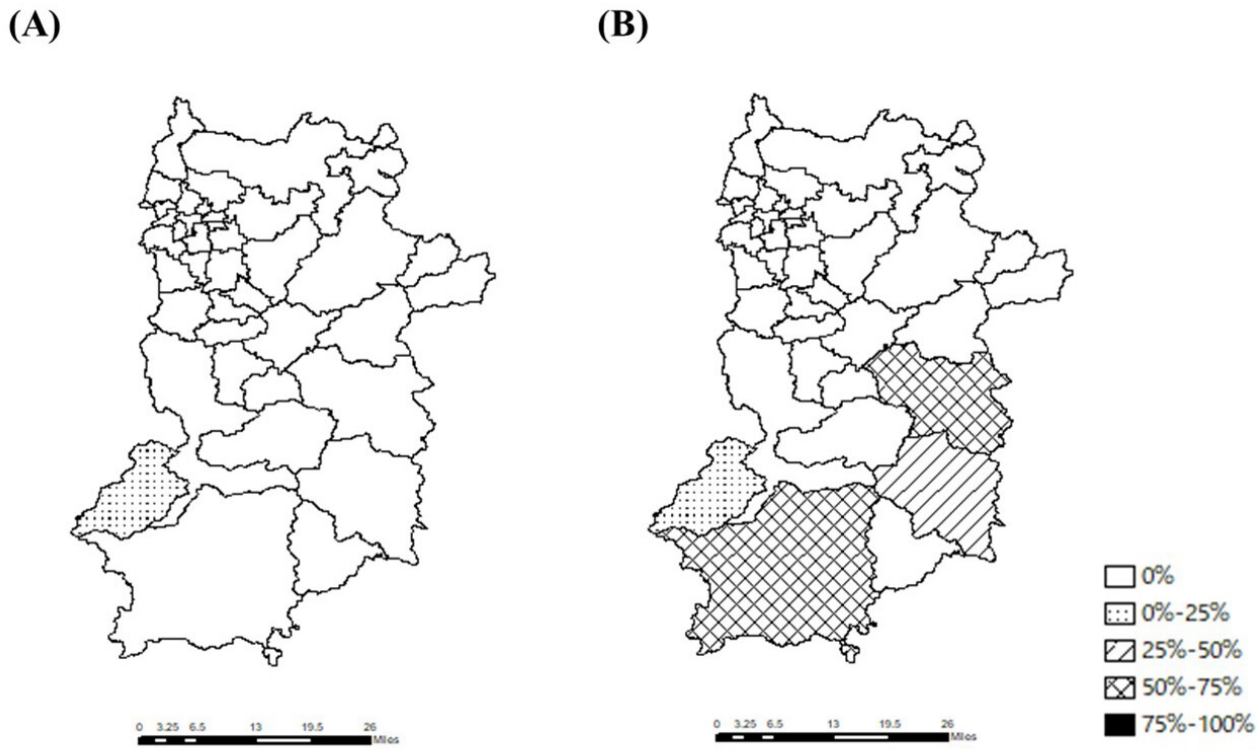


Figure 4. The rates of patients for whom the travel time was longer than 30 minutes. This figure illustrates the rates of patients for whom the travel time was longer than 30 minutes in (A) the ideal scenario and (B) the actual scenario.



XSL•FO
RenderX

Figure 5. The median travel distance for each municipality in the two scenarios illustrates the median travel distance in (A) the ideal scenario and (B) the actual scenario.

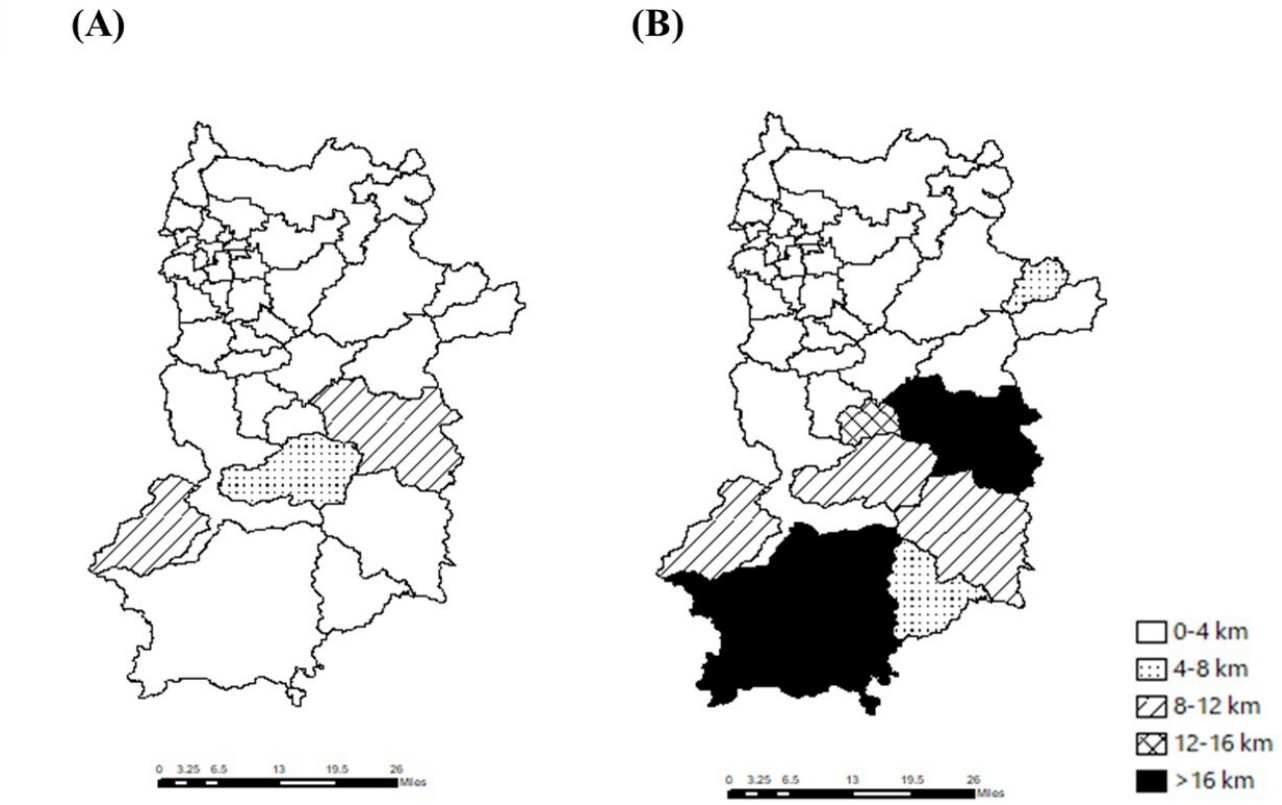
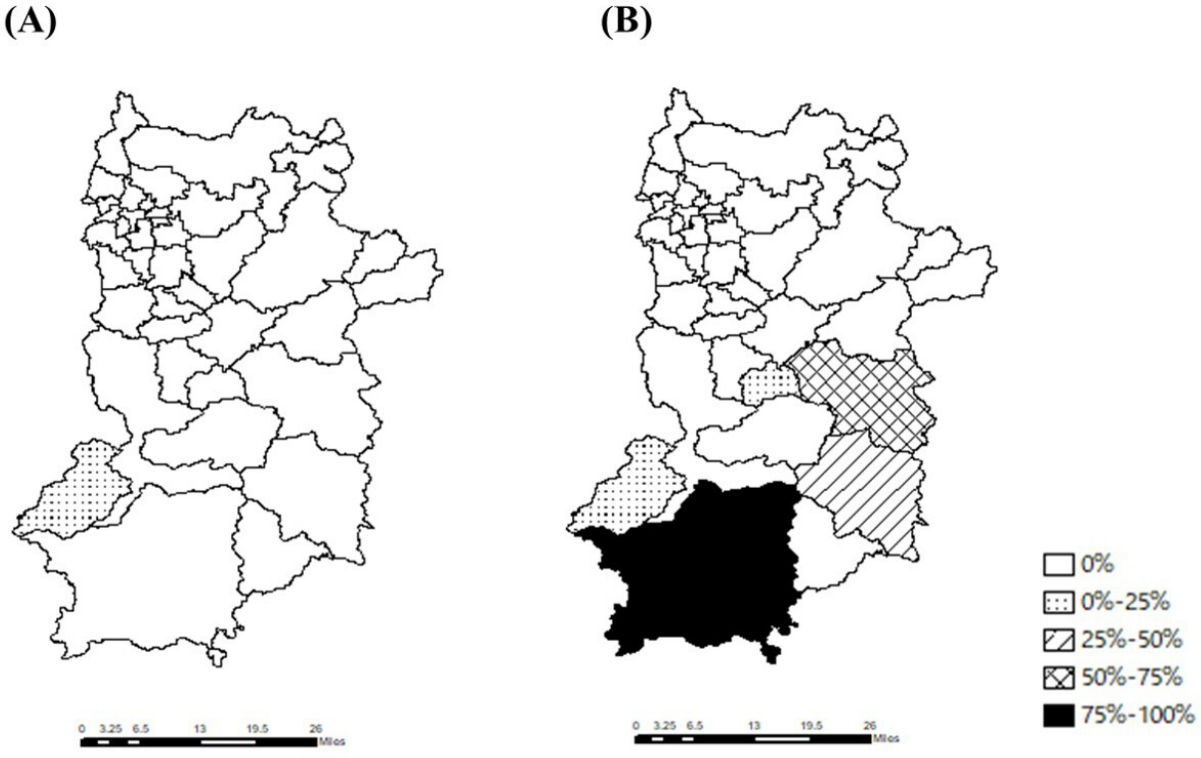


Figure 6. The rates of patients for whom the travel distance was longer than 16 km. This illustrates the rates of patients for whom the travel distance was longer than 16 km in (A) the ideal scenario and (B) the actual scenario.



XSL•FO
RenderX

The median travel times for each municipality in the 2 scenarios are shown in Figures 3 and 4. In the ideal scenario, the median travel time was less than 15 minutes, except for Nosegawa (18.1, IQR 2.2 - 25.5 min), while in the actual scenario, the median travel time was longer in the municipalities in the Nanwa area, the southern part (in Totsukawa (32.6, IQR 21.2 - 40.5 min), Kawakami (30.1, IQR 16.5 - 32.4 min), Kurotaki (21.3, IQR 20.2 - 24.1 min), and Kamikitayama (20.7, IQR 19.7 - 26.1 min; Figure 3).

When comparing the 2 scenarios, the differences in travel time between the scenarios were the largest in Totsukawa (26.8 min), Kawakami (18.3 min), Kamikitayama (17.2 min), and Kurotaki (16.1 min). All of these 4 municipalities are located in a depopulated secondary medical area of Nanwa.

Figure 4 shows the number of patients for whom the travel time was longer than 30 minutes. In the actual scenario, the rates were highest in Totsukawa (58.0%), Kawakami (50.0%), Kamikitayama (30.0%), and Nosegawa (25.0%).

The median travel distances for each municipality in the 2 scenarios are shown in Figure 5. In the ideal scenario, the median travel distances were shorter than 12 km for all municipalities. However, in the actual scenario, the median travel distances were longer than 16 km in Totsukawa (28.4, IQR 16.7 - 32.3 km) and Kawakami (17.2, IQR 10.3 - 19.3 km). Kurotaki, Nosegawa, and Kamikitayama showed similar findings (13.9, IQR 12.6 - 14.4; 11.5, IQR 1.3 - 16.1; and 10.8, IQR 10.3 - 16.4 km, respectively). The differences in travel distance between the scenarios were the largest in Totuskawa (24.5 km), Kurotaki (11.8 km), Kamikitayama (9.3 km), and Kawakami (8.6 km).

Figure 6 shows the number of patients whose travel distance was >16 km. In the actual scenario, the rates were 77.6% for Totsukawa, 56.2% for Kawakami, 30.0% for Kamikitayama, 25% for Nosegawa, 20% for Kurotaki, and 0% for the other municipalities. In the ideal scenario, the rate was zero, except in Nosegawa (25.4%).

The usage rates by municipalities are shown in Table 1. This table shows the usage rates of home medical care services in each municipality. As can be seen, the usage rates tended to be lower in the municipalities in the Nanwa area (colored black), except for Nosegawa, which had the highest rate.

Discussion

Principal Findings

In this study, an accessibility analysis of home medical care services in Nara Prefecture was conducted as a model region, using the KDB data to investigate the equitable distribution of home medical care services. Using the KDB, we accurately identified medical institutions that provide home medical care, estimated the usage rates of home medical care services for each municipality, and demonstrated the feasibility of analyzing geographical accessibility to home medical care services that reflect actual health care usage and clinical trends.

The findings of this study revealed that in certain municipalities within the Nanwa area, a rural secondary medical area with a declining population, the travel distance to a health care patient's home exceeded 16 km, and the travel time exceeded 30 minutes in some cases. These results highlight disparities among municipalities in terms of geographic accessibility to home medical care services. Its supply in the Nanwa area needs to be increased from the perspective of geographic accessibility.

In addition, the usage rates of these services were lower in rural municipalities in the Nanwa area. To promote the use of home medical care services, improving accessibility in these municipalities is necessary. Geographical accessibility—that is, the distance from service providers to patients—could affect service usage. Naruse et al [7] reported that travel distance can affect the usage of home nursing care services. Although geographical accessibility would also likely affect the usage of home medical care services, the extent of the effect has not been clear since this study did not quantify the effect by controlling other factors, such as socioeconomic status, care quality, and regional characteristics. Further studies with additional information beyond medical claim data are required to clarify this point.

In recent years, the government has promoted the use of home medical care to meet the increasing medical demands within each area, enabling local people to live their lives in areas they have always lived in [3]. Our results showed that a disparity occurred in geographical accessibility to home medical care based on the current practice in the actual scenario; however, in the ideal scenario, no municipality was observed in which the travel distance to any patient was longer than 16 km, except for Nosegawa. Geographical accessibility is much better in this scenario. For example, the average travel time was less than 15 minutes, except for Nosegawa. These results indicate that the disparity in terms of geographical accessibility would have been mostly resolved if all hospitals and clinics had provided these services. Furthermore, our results indicated that the potential benefit would be the largest in the Nanwa area, the most depopulated area. Therefore, if the number of hospitals or clinics providing these services is increased by a policy (eg, through dispatching health care professionals or funding), this will be effective in ensuring sufficient geographical accessibility for patients in the Nanwa area (specifically, Kawakami, Totsukawa, Kurotaki, and Kamikitayama). That is especially necessary in Kawakami and Kurotaki, since no hospitals or clinics were providing services, and the usage rates were low in these municipalities. However, it is neither feasible nor efficient for all hospitals and clinics to provide services because of limited resources. Further analysis using GIS could clarify how residents' geographical accessibility will change when a specific hospital or clinic starts to provide home medical care services.

Notably, while the travel distance was longer in Nosegawa than in the other municipalities in both scenarios, the usage rate was the highest. The high usage rate could have been achieved by factors other than geographical accessibility. Further study, including factors other than geographical accessibility, is required to clarify the reason.

Our study is among the few that have analyzed geographical accessibility to home medical care services using GIS and KDB. The Nara KDB used in this study is an integrated database of

XSL•FO
RenderX

medical and long-term care insurance claim information covering all residents aged 75 years and older, which matches the primary population of home medical care patients. This methodology enabled us to conduct a geographical analysis that reflected actual health care usage. Furthermore, our results are important when considering the optimal allocation of medical resources related to home medical care services and equal accessibility to these services. Japan is the most aging country in the world [1]. Home medical care services will be more important in other countries, and ensuring accessibility to these services will be of greater importance in the future. Geographical analyses using nationwide health care claims databases have been conducted in other countries for health care fields, such as drug prescription and pregnancy outcomes [11,12]. Likewise, our methodology can be applied to other countries for home medical care services, as those services will be more important as the aging of the population progresses.

Limitations

The KDB is a comprehensive database that includes medical claims information for home medical care patients and details the provision of medical institutions, enabling a geographical analysis that reflects actual health care usage. However, this study had some limitations. First, it did not consider the capacity of each hospital or clinic because it focused on geographical accessibility. Further studies are required to consider its capacity to be more informative for optimal resource allocation. Second, a hospital or clinic was regarded as providing home medical care services if at least one claim regarding services from a hospital or clinic was identified. Some hospitals and clinics included in the actual scenario might not routinely provide home medical care services. Third, the disparity in geographical accessibility within each municipality is beyond the scope of this study. Such disparities may exist particularly in municipalities with lower geographical accessibility. Fourth, the KDB was used to analyze home medical care patients aged ≥75 years; therefore, home health care patients aged <75 years were not included. In the future, using the NDB or KDB and applying the developed methods, the analysis could be expanded to include pediatric home medical care patients and those under the age of 75 years. Finally, this study only used KDB data from FY 2019 to conduct geographical analysis while excluding the impact of the COVID-19 pandemic. The usage of home medical care has been increasing over the years, and the usage patterns could also change over the years. By applying the methodology established in this study, a nationwide analysis using the most recent data, free from the influence of the COVID-19 pandemic, can provide valuable insights into the future of home medical care.

Conclusions

This study conducted a geographical accessibility analysis of home medical care services using Nara Prefecture as a model region with GIS and KDB. The results revealed that geographical accessibility was lower in depopulated municipalities in the southern part, such as Totsukawa, Kawakami, Kamikitayama, and Nosegawa, which have lower accessibility than urban areas. However, the results also showed that the disparity could be partly solved in the ideal scenario. The differences in travel time between the scenarios were especially larger in the depopulated areas in the southern part, such as Totsukawa (32.6 vs 5.8 min), Kawakami (30.1 vs 11.8 min), Kurotaki (21.3 vs 5.2 min), and Kamikitayama (20.7 vs 3.5 min). These results highlight the necessity of increasing supply in the southern areas.

Acknowledgments

This study was partially supported by a Grant-in-Aid for Scientific Research from the Ministry of Health, Labour and Welfare, Japan.

Data Availability

The datasets used in the current study are not publicly available owing to the publication policy of the Nara Prefecture but are available from the corresponding author on reasonable request subject to approval from the Nara Prefecture.

Authors' Contributions

YM, YN, and MA designed the research plan. YT, TN, TM, and TI contributed significantly to the development of the research. YM performed the geographical analysis and drafted the manuscript. YN performed the claims database analysis. AM led the research project.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Median and 25-75 percentiles of travel distance in the ideal and actual scenarios (km).
[DOCX File, 24 KB - aging_v8i1e70040_app1.docx]

Multimedia Appendix 2
Median and 25-75 percentiles of travel time in the ideal and actual scenarios (min).

XSL•FO
RenderX

[DOCX File, 23 KB - aging_v8i1e70040_app2.docx]

References

1. International comparison of elderly population (in japanese). Statistics Bureau of Japan. URL: https://www.stat.go.jp/data/topics/topi1135.html [accessed 2025-07-01]
2. Constructing the structure of providing home medical care services (in japanese). The Ministry of Health, Labour and Welfare. 2023. URL: https://www.mhlw.go.jp/content/10800000/001146149.pdf [accessed 2024-09-07]
3. Home medical care services -toward 8th regional medical plan (in japanese). The Ministry of Health, Labour and Welfare. 2021. URL: https://www.mhlw.go.jp/content/10800000/000911334.pdf [accessed 2024-09-07]
4. Regional Medical Plan (In Japanese). The Ministry of Health, Labour and Welfare. URL: https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryou/iryou/iryou_keikaku/index.html [accessed 2024-09-07]
5. Schwarz J, Hemmerling J, Kabisch N, et al. Equal access to outreach mental health care? Exploring how the place of residence influences the use of intensive home treatment in a rural catchment area in Germany. BMC Psychiatry 2022 Dec 26;22(1):826. [doi: 10.1186/s12888-022-04477-y] [Medline: 36572855]
6. Wang Y, Zhang Q, Spatz ES, et al. Persistent geographic variations in availability and quality of nursing home care in the United States: 1996 to 2016. BMC Geriatr 2019 Apr 11;19(1):103. [doi: 10.1186/s12877-019-1117-z] [Medline: 30975076]
7. Naruse T, Matsumoto H, Fujisaki-Sakai M, Nagata S. Measurement of special access to home visit nursing services among Japanese disabled elderly people: using GIS and claim data. BMC Health Serv Res 2017 May 30;17(1):377. [doi: 10.1186/s12913-017-2322-0] [Medline: 28558677]
8. Kuwayama T, Hamabata K, Kamesaki T, Koike S, Kotani K. Research on home care nursing in Japan using geographic information systems: a literature review. Jpn Clin Med 2018;9:1179670718814539. [doi: 10.1177/1179670718814539] [Medline: 30515029]
9. Shimizu J, Osanai S. Feasibility of home visit by nurses with higher palliative care specialty with visiting nurses in Japan—analysis of nationwide distribution and geographical associations [Abstract in English]. J Jpn Health Sci 2014;16:177-183.
10. Iizaka S, Sanada H. Distribution of Tokyo treatment centers with wound, ostomy, and continence nurses: a spatial analysis of hospitals and municipalities using a geographic information system. Jpn J PU 2016;18:118-126.
11. Cheng CL, Chen YC, Liu TM, Yang YHK. Using spatial analysis to demonstrate the heterogeneity of the cardiovascular drug-prescribing pattern in Taiwan. BMC Public Health 2011 May 24;11:380. [doi: 10.1186/1471-2458-11-380] [Medline: 21609462]
12. Kwak MY, Lee SM, Kim HJ, et al. How far is too far? A nationwide cross-sectional study for establishing optimal hospital access time for Korean pregnant women. BMJ Open 2019 Sep 20;9(9):e031882. [doi: 10.1136/bmjopen-2019-031882] [Medline: 31542767]
13. Geospatial information authority of japan. Land area (In Japanese). URL: https://www.gsi.go.jp/KOKUJYOHO/MENCHO/backnumber/GSI-menseki20250401.pdf [accessed 2024-09-19]
14. Nara prefectural government estimated population. Estimated population. 2024. URL: https://www.pref.nara.jp/6265.htm [accessed 2024-09-07]
15. Geospatial information authority of Japan. Digital Land Information Download Website (In Japanese). URL: https://nlftp.mlit.go.jp/ksj/index.html [accessed 2024-09-07]
16. Morii Y, Osanai T, Fujiwara K, et al. Analyzing Cost-Effectiveness of Allocating Neurointerventionist for Drive and Retrieve System for Patients with Acute Ischemic Stroke. J Stroke Cerebrovasc Dis 2021 Aug;30(8):105843. [doi: 10.1016/j.jstrokecerebrovasdis.2021.105843] [Medline: 34000607]
17. Morii Y, Osanai T, Ishikawa T, et al. Cost Effectiveness of Drive and Retrieve System in Hokkaido for Acute Ischemic Stroke Patient Treatment Using Geographic Information System. J Stroke Cerebrovasc Dis 2019 Aug;28(8):2292-2301. [doi: 10.1016/j.jstrokecerebrovasdis.2019.05.020] [Medline: 31200963]
18. A requirement on distance for home medical care services. The Ministry of Health, Labour and Welfare. 2023. URL: https://www.mhlw.go.jp/stf/shingi2/0000212500_00236.html [accessed 2024-09-27]

Abbreviations

GIS: geographical information system
KDB: Kokuho database
NDB: National Database of Health Insurance Claims and Specific Health Checkups of Japan

XSL•FO
RenderX

Edited by D Liu; submitted 13.12.24; peer-reviewed by F Zicker, L Jordan, RBL Lim, R Inokuchi; revised version received 16.04.25; accepted 07.06.25; published 06.08.25.

Please cite as:
Morii Y, Nakanishi Y, Nishioka Y, Tsugihashi Y, Noda T, Myojin T, Imamura T, Akahane M
Analyzing Disparity in Geographical Accessibility to Home Medical Care Using a Claims Database and Geographical Information System: Simulation Study
JMIR Aging 2025;8:e70040
URL: https://aging.jmir.org/2025/1/e70040
doi:10.2196/70040

© Yasuhiro Morii, Yasuhiro Nakanishi, Yuichi Nishioka, Yukio Tsugihashi, Tatsuya Noda, Tomoya Myojin, Tomoaki Imamura, Manabu Akahane. Originally published in JMIR Aging (https://aging.jmir.org), 6.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Determinants of Having Online Health Consultations During the COVID-19 Pandemic Among Middle-Aged and Older Adults in Germany: Representative Longitudinal Survey Study

Ariana Neumann, MSc; Hans-Helmut König, MPH, Prof Dr; André Hajek, Prof Dr

Department of Health Economics and Health Services Research, University Medical Center Hamburg-Eppendorf, Martinistr. 52, Hamburg, Germany

Corresponding Author:
Ariana Neumann, MSc
Department of Health Economics and Health Services Research, University Medical Center Hamburg-Eppendorf, Martinistr. 52, Hamburg, Germany

Abstract

Background: During the COVID-19 pandemic, telemedicine services represented a widely implemented alternative to in-person doctor and therapist appointments. Consequently, rates of telemedicine use rapidly increased worldwide, also in Germany. Research regarding longitudinal determinants of telemedicine use is needed, particularly from nationally representative German samples, to improve understanding of the use behavior of major target groups such as middle-aged and older adults.

Objective: This study aimed to longitudinally investigate determinants of online health consultation use among middle-aged and older individuals during the COVID-19 pandemic in Germany.

Methods: Nationally representative longitudinal data of German middle-aged and older adults (≥46 years old) were taken from the German Ageing Survey (DEAS). Data from the Compact Survey (conducted between June and July 2020) and wave 7 (conducted between November 2020 and March 2021) of the DEAS were observed (pooled analytic sample N=5456). Having experienced consultations with doctors or therapists on online platforms served as the outcome measure. Associations with socioeconomic, health- and health behavior–related, psychological, and COVID-19–related determinants were tested using random effects logistic regressions.

Results: In our sample, 49% (2673/5456) of participants were female and the mean age of the participants was 67.8 (SD 9.4) years. Past experience with online health consultations was reported by 10.3% (561/5456) of the sample. Online health consultation use was associated with high education (OR 1.43, 95% CI 1.06 - 1.93; *P*=.02), poor self-rated health (OR 0.60, 95% CI 0.49 - 0.75; *P*<.001), and higher frequency of physical activity (reference: low frequency; medium frequency: OR 1.58, 95% CI 1.15 - 2.17; *P*=.005; high frequency: OR 1.73, 95% CI 1.09 - 2.76; *P*=.02). Moreover, greater levels of loneliness (OR 1.43, 95% CI 1.06 - 1.93; *P*=.04) and life satisfaction (OR 1.33, 95% CI 1.02 - 1.73; *P*=.04) as well as perceiving the COVID-19 crisis as a greater personal threat (OR 1.08, 95% CI 1.01 - 1.15; *P*=.02) were associated with having online health consultations during the COVID-19 pandemic.

Conclusions: Online health consultation use does not seem to be exclusively associated with the health of middle-aged and older patients. Study findings emphasize the longitudinal association of education and psychosocial factors as well as health factors with telemedicine use during the COVID-19 pandemic in Germany. This knowledge may help to improve and adapt services to this patient group, which could contribute to higher utilization rates in the future. Future studies are needed to verify these initial findings under postpandemic circumstances and across different countries.

(JMIR Aging 2025;8:e60311) doi:10.2196/60311

KEYWORDS

telemedicine; digital health; remote consultation; health services for the aged; patient acceptance of health care; COVID-19

Introduction

Worldwide, health care systems are facing multiple challenges in the future related to the delivery of services (eg, access and continuity of care), human resources (eg, staff distribution and sufficiency), as well as leadership and governance (eg, strategic policies) [1]. In particular, the increasing prevalence of noncommunicable diseases, disability, and multimorbidity in connection with population aging calls for new solutions in health care (eg, [2]). According to predictions of the World Health Organization (WHO) [3], the number of global deaths due to noncommunicable diseases may increase from 36 million (2008) up to 52 million in 2030.

Promising methods for the delivery of health care in the future include digital solutions such as telemedicine [4]. The WHO Global Observatory for eHealth [5] defines telemedicine as

XSL•FO
RenderX

follows: (1) its purpose is to provide clinical support, (2) it is intended to overcome geographical barriers, connecting users who are not in the same physical location, (3) it involves the use of various types of information and communication technology, and (4) its goal is to improve health outcomes. Major strengths of telemedicine include improved access to care and information, time and cost savings, as well as convenience and flexibility [6,7]. Multiple studies evaluated the effectiveness and cost-effectiveness of telemedicine services. The services seem to be effective and can produce at least comparable effects to in-person services [8,9]. Moreover, the current literature suggests that telemedicine can be a cost-effective service delivery method [8-11]. In addition, providers as well as patients seem to be highly satisfied with the services [8,9,12].

Despite its benefits, telemedicine implementations were only limited and remained at a low level up until the occurrence of the COVID-19 pandemic [13-15]. To prevent further spreading of the virus and relieve the health care system, major changes in the delivery of health care services had to be made in response to the pandemic [15]. Telemedicine represented a valuable tool to avoid personal contact during health consultations. Consequently, the global telemedicine utilization increased tremendously [13-16]. For example, Koonin et al [17] reported a 154% increase in telehealth visits in the United States in March 2020 compared with March 2019. Also in Germany, which is the focus of our study, the proportion of contract physicians and psychotherapists who offered and billed for telemedicine services increased from 6.1% (2019) to 24.6% (2021) [18]. Particularly among the psychotherapeutic care sector, utilization rates were high in Germany [18]. Since many nonessential in-person medical appointments were canceled or postponed due to the pandemic in Germany [19,20], remote health consultations represented a valuable alternative format to assure the continuity of care in spite of pandemic circumstances. This digital transformation of the German health care system was facilitated by the introduction of laws including the Digital Health Care Act [21] and the Digital Health Application Ordinance [22], which supported the prescription of health care apps, provision of video consultations, and integration of digital provider networks. In the German health care system, health insurance is compulsory with about 90% of the population having statutory and 10% having private health insurance [23]. The implementation of the new laws enabled the coverage of digital health care by statutory insurances, ensuring that telemedicine users incur no additional costs.

Considering population aging, telemedicine is particularly relevant for the older patient groups. Telemedicine was found to be effective in older adults [24-26]. For instance, van den Berg et al [25] reviewed 68 interventional telemedical studies with a controlled design examining older adults and found that none of the included studies reported better outcomes for the control group (eg, randomized or matched control groups that received usual in-person care). Moreover, older adults seem to accept and are satisfied with the services [24,27]. Taking older adults' higher need for health care and probable mobility restrictions into account, older adults can particularly benefit from the remote format. Nevertheless, older patients use telemedicine services less often than younger age groups and report multiple barriers to using the services (eg, [28-32]). For example, Wilson et al [30] identified factors such as physical difficulties (eg, visual or hearing impairments), privacy concerns, lack of experience, training, or support as barriers to the use of eHealth by older adults.

In Germany, older patients also used telemedicine services less frequently than younger patients did during the pandemic [18,33-36]. However, increasing telemedicine use in older age groups can be helpful for the future delivery of health care services in Germany to deal with problems such as physician shortages or increased demand for (long-term) care caused by demographic change [37]. Therefore, it is important to examine factors that are associated with telemedicine use in older adults to increase utilization in Germany. Previous international reviews highlighted older patients' characteristics associated with higher telemedicine use [38-40]. This included younger age, higher education, higher self-efficacy, greater experience or skills in using electronic devices, access to technology and internet, greater social support or influence, higher (health-related) motivation, and greater openness to experience as well as fewer privacy concerns and less severe health impairments.

So far, only limited evidence from Germany regarding the determinants of telemedicine use during the pandemic exists. The previous studies stressed the association of socioeconomic (male or female sex, younger age, high or low education, social status, living in an urban area, and having children under 18 years), psychosocial (loneliness, digital literacy), health (mental or physical health problems), and COVID-19–related factors (higher perceived severity of COVID-19 infection, having had COVID-19 infection, subjective COVID-19–related challenges, COVID-19–related cognitive preoccupation, anxiety, and worries) with telemedicine use [18,34-36,41-43]. Very few studies exclusively looked at middle-aged and older adults in Germany during the pandemic. These studies highlighted the positive association of telemedicine use with education, living with a partner in the same household, mental (ie, anxiety and depression) or physical health problems, loneliness, life satisfaction as well as forgoing medical treatment due to the fear of being infected by the coronavirus [44,45]. A large variety of determinants was observed in the existing studies and more research is needed to further explore and verify the findings.

In addition to the small number of studies that observe determinants of telemedicine use in middle-aged and older adults during the pandemic in Germany, hardly any studies examined the determinants longitudinally. Solely, von der Groeben et al [35] used a quasi-longitudinal design (ie, they observed cross-sectional data from 3 different time points) to detect determinants of patient use and attitude toward using video and telephone conferences in a population-representative sample of adults (18 - 69 y) affected by depression during 3 different pandemic time points in Germany. Since telemedicine use and acceptance varied over the course of the pandemic (eg, [31,35,46]), it may be beneficial to consider more than just one time point when observing telemedicine use behavior during the pandemic. Moreover, the longitudinal approach gives further insight into the directionality of the relationships. Therefore,

XSL•FO
RenderX

our study aimed to longitudinally investigate determinants of online health consultation use in a large representative sample of middle-aged and older individuals during the COVID-19 pandemic in Germany.

Expanding the knowledge regarding determinants of online health consultation use in middle-aged and older adults could help to identify target groups for telemedicine services, as well as groups that would benefit from additional support for using the services. Furthermore, this knowledge may help to adapt telemedicine services to the needs and preferences of middle-aged and older adults. Consequently, important practical and theoretical implications may be derived from our findings, which could foster greater use of telemedicine services among middle-aged and older individuals, ultimately helping to deal with future health care challenges posed by population aging.

Methods

Sample

Nationally representative cross-sectional and longitudinal data were taken from the German Ageing Survey (DEAS [47]). The DEAS focuses on the German middle-aged and older population (starting at 40 y) and aims to describe living conditions and diversity among this population as well as aging and social change processes that are related to this life stage [47]. The first wave of the DEAS was conducted in 1996, followed by further waves in 2002, 2008, 2011, 2014, 2017, 2020/2021. The survey has a cohort-sequential design in which new baseline samples were added in 2002, 2008, and 2014. The baseline samples of the DEAS were disproportionally stratified into age groups, gender, and region [47]. In response to the COVID-19 pandemic, an additional Compact Survey to measure the pandemic impact on middle- and older adults' lives in Germany was implemented in 2020.

For the purpose of our study, data from the Compact Survey [48] and wave 7 [49] of the DEAS were observed. The Compact Survey was conducted from June until July 2020 and consisted of paper-and-pencil questionnaires that were sent to individuals who had taken part in the DEAS at least once in the past. The response rate for the Compact Survey was 57% (4823/8533) [50]. Due to the ongoing pandemic during wave 7, the usual computer-assisted personal interviews were replaced by telephone-administered interviews as well as an additional paper-and-pencil questionnaire. The data collection for wave 7 took place from November 2020 until March 2021 and the response rate was 66% (5402/8207) [51]. Overall, 4103 individuals participated in both surveys [51]. Since the last DEAS baseline sample was added in 2014, participants of the Compact Survey and wave 7 of the DEAS were at least 46 years old.

The DEAS is funded by the German Federal Ministry for Family Affairs, Senior Citizens, Women, and Youth and was conducted and developed by the German Centre of Gerontology. The fieldwork was carried out by the INFAS Institute for Applied Social Sciences.

Ethical Considerations

Written informed consent was obtained from all individuals who participated in the DEAS, and opting out of the survey was possible at all times [47]. DEAS participants received an incentive (eg, €10 in DEAS wave 7; conversion rate US $1=€0.951 in 2022 [51]). Due to data protection guidelines, a data distribution contract needs to be signed prior to using the anonymized DEAS data [47]. The DEAS study complies with the Declaration of Helsinki and did not require further ethical examination since the criteria for the need of ethical approval were not met for this survey (eg, missing information regarding the study or aim of the study, examination of vulnerable groups or patients, high risk or burden for participants due to participation).

Dependent Variable

For the sake of our study, we exclusively included DEAS participants, who indicated having access to the internet (Compact survey: 3858/4676, 83%; wave 7: 3676/4276, 86%). In both DEAS waves, participants were asked "How often do you use the Internet for the following purposes?." Among the listed items was "Consultations with doctors and therapists via an online platform" in the Compact survey and "Providing consultations with doctors or therapists on online platforms" in wave 7 of the DEAS. The response format consisted of a 6-point Likert scale indicating use frequency as "never," "less often or seldom," "1 to 3 times a month," "once a week," "several times a week," or "daily." Since only a small number of participants had multiple consultations with doctors or therapists on online platforms, the outcome was dichotomized for our analysis (0="never"; 1="less often or seldom," "1 to 3 times a month," "once a week," "several times a week," or "daily").

Independent Variables

Based on theoretical considerations and previous research [38-40,44,45], different groups of determinants were considered. This included socioeconomic, health- and health behavior–related, psychological, and COVID-19–related determinants. Regarding socioeconomic characteristics, we examined sex, age, educational level (International Standard Classification of Education 97 [52]: low, medium, or high education), employment status (used, retired, and other or unemployed), household income, migration background (no, yes), area lived in (metropolitan districts, urban districts, [partially] densely populated rural districts, sparsely populated rural districts), residential form of partnership (no partner, partner in the same household, and partner not in the same household), and presence of children (no, yes).

Health-related factors included self-rated health (ranging from 0=very bad to 4=very good) as well as the frequency of physical activity and walks (ranging from 0=never to 5=daily) as measures of health-related behavior. For the sake of our analysis, the values for the frequency of physical activity and walks were divided into tertiles and were included as categorical variables (low, medium, and high frequency).

Psychological determinants that were considered were depressive symptoms, loneliness, attitude toward own aging, and life satisfaction. Depressive symptoms were measured using

XSL•FO
RenderX

a 10-item German short form of the Center for Epidemiologic Studies - Depression scale [53] (CES-D; scores ranging from 0 to 30, higher values indicate more severe depressive symptoms) in the Compact survey. In wave 7 of the DEAS, depressive symptoms were measured using the German version of the 15-item CES-D [54] (scores ranging from 0 to 45, higher values indicate more severe depressive symptoms). The values for both scales were standardized to assure comparability between both surveys in our analysis. Both of these well-established instruments were evaluated in the past and have good psychometric properties [53,55,56]. Cronbach α values for both scales were 0.83 (Compact survey) and 0.84 (wave 7) and McDonald Omega was 0.85 (Compact survey) and 0.86 (wave 7) in our sample. In addition, loneliness was measured with the 6-item De Jong Gierveld Loneliness Scale [57] (scores ranging from 1 to 4, higher values indicate higher levels of loneliness). The scale has favorable psychometric properties [57] (Compact survey: Cronbach α=0.78, McDonald's Omega =0.79; wave 7: Cronbach α=0.80, McDonald Omega =0.81). The self-perception of one's own aging was examined using the German version of the 5-item Attitude Toward Own Aging subscale of the Philadelphia Geriatric Center Morale Scale [58,59] (scores ranging from 1 to 4, higher values indicate a more positive perception of own aging). This widely used scale was evaluated in different age groups in the past (eg, [60]). In our sample, Cronbach α was 0.77 (Compact Survey and wave 7) and McDonald Omega was 0.77 (Compact Survey and wave 7). Life satisfaction was measured with the German version of the Satisfaction with Life Scale [61,62] (scores ranging from 1 to 5, higher values indicate greater life satisfaction). The German version of the scale was evaluated in the past and showed good psychometric properties [63]. For this scale, the Cronbach α was 0.86 (Compact survey) and 0.84 (wave 7) and the McDonald Omega was 0.87 (Compact survey) and 0.86 (wave 7) in our sample.

Finally, we controlled for COVID-19–related determinants, which included perceiving the Corona crisis as a personal threat (scores ranging from 1=not at all a threat for me to 10=extreme threat for me), past infection with the Coronavirus by oneself (no, yes, and unknown), or by people from one's personal environment (no, yes, and unknown) as well as the feeling of being able to influence the infection with the Coronavirus (scores ranging from 1=not at all to 7=entirely).

Statistical Analysis

In the first step, sample characteristics of our pooled analytic sample were computed. The analytic sample consisted of individuals who participated in at least one of the 2 surveys (5456 observations corresponding to 3222 individuals). Second, random effects logistic regressions were calculated to test the associations of the determinants with online health consultation use. The random effects regression model considers the panel structure of the data and allows the inclusion of not only time-varying but also time-constant predictors in our model, under the assumption that unobserved unit-specific heterogeneity is not correlated with the independent variables [64]. When this independence assumption is fulfilled, the random effects model may be more efficient than the fixed effects model as it considers both between and within variation [64]. Our choice was supported by the Hausman test. The null hypothesis of the Hausman test states that both models (fixed and random effects model) are consistent while the random effects model is more efficient [64,65]. Therefore, the random effects model is preferred when the null hypothesis cannot be rejected. Since the Hausman test statistic was nonsignificant for our sample (P=.72), we used random effects models for our analysis. Stata (version 16.0, StataCorp) was used for the statistical analyses and the random effects logistic regression was calculated using the "xtlogit" command with the "re" option. The sample was stratified by sex and age groups (≤64 and ≥65 years) in additional analyses. Statistical significance was defined as an alpha level of P<.05. Missing data were handled using listwise deletion.

Results

Sample Characteristics

The pooled analytic sample characteristics for all included variables are presented in Table 1. In the pooled sample of the Compact survey and wave 7 of the DEAS, 49% (2673/5456) were female and the mean age of the participants was 67.8 (SD 9.4) years. When examining past consultations with doctors or therapists on online platforms, 10.3% (561/5456) reported past experience with online consultations.

XSL•FO
RenderX

Table . Analytic pooled sample characteristics (N=5456).

Characteristics[a]	Values
Consultations with doctors or therapists on online platforms, n (%)	
No	4895 (89.7)
Yes	561 (10.3)
Sex, n (%)	
Male	2783 (51)
Female	2673 (49)
Age (years), mean (SD)	67.8 (9.4)
Educational level, n (%)	
Low (ISCED[b] 0 - 2) or medium (ISCED 3 - 4)	2455 (45)
High (ISCED 5 - 6)	3001 (55.0)
Employment status, n (%)	
Employed	1745 (32.0)
Retired	3475 (63.7)
Other or unemployed	236 (4.3)
Monthly household income (€), mean (SD)[c]	4051.2 (11,806.1)
Migration background, n (%)	
No	5223 (95.7)
Yes	233 (4.3)
Area lived in, n (%)	
Metropolitan districts	1524 (27.9)
Urban districts	2014 (36.9)
(Partially) densely populated rural districts	1135 (20.8)
Sparsely populated rural districts	783 (14.4)
Residential form of partnership, n (%)	
No partner	1069 (19.6)
Partner in the same household	4152 (76.1)
Partner not in the same household	235 (4.3)
Having children, n (%)	
None	602 (11)
One or more	4854 (89)
Self-rated health, mean (SD)[d]	2.6 (0.8)
Frequency of physical activity, n (%)	
Low frequency	1589 (29.1)
Medium frequency	3234 (59.3)
High frequency	633 (11.6)
Frequency of walks, n (%)	
Low frequency	2135 (39.1)
Medium frequency	2147 (39.4)
High frequency	1174 (21.5)
Depressive symptoms, mean (SD)[e]	−0.1 (0.9)
Loneliness, mean (SD)[f]	1.8 (0.5)

XSL•FO
RenderX

Characteristics[a]	Values
Life satisfaction, mean (SD)[g]	3.9 (0.7)
Attitude toward own aging, mean (SD)[h]	3.0 (0.5)
Perceiving the Corona crisis as a personal threat, mean (SD)[i]	4.3 (2.1)
Oneself infected with the Coronavirus, n (%)	
No	5244 (96.1)
Yes	57 (1)
Unknown	155 (2.8)
People from personal environment infected with the Coronavirus, n (%)	
No	4517 (82.8)
Yes	853 (15.6)
Unknown	86 (1.6)
Feeling that one can influence the infection with the Coronavirus, mean (SD)[j]	4.7 (1.4)

[a]Due to differences in measurement tools, values for depressive symptoms had to be standardized. The standardized values should be interpreted as number of SDs by which the original values lay above or below their mean. For example, a value of 1 or -1 indicates that the reported overall score in the CES-D lays one SD above/below the mean CES-D score.

[b]ISCED: International Standard Classification of Education.

[c] Range (0-500,000); the conversion rate of USD to Euro was US $1=€0.846 in 2021 and US $1=€0.951 in 2022.

[d]Range 0-4. Higher values indicate better self-rated health.

[e]Range for standardized values –1.8 to 5.9. Higher values indicate more depressive symptoms.

[f]Range 1-4. Higher values indicate higher loneliness levels.

[g]Range 1-5. Higher values indicate greater life satisfaction.

[h]Range 1-4. Higher values indicate a more positive perception of own aging.

[i]Range 1=not at all a threat for me to 10=extreme threat for me.

[j]Range 1=not at all to 7=entirely.

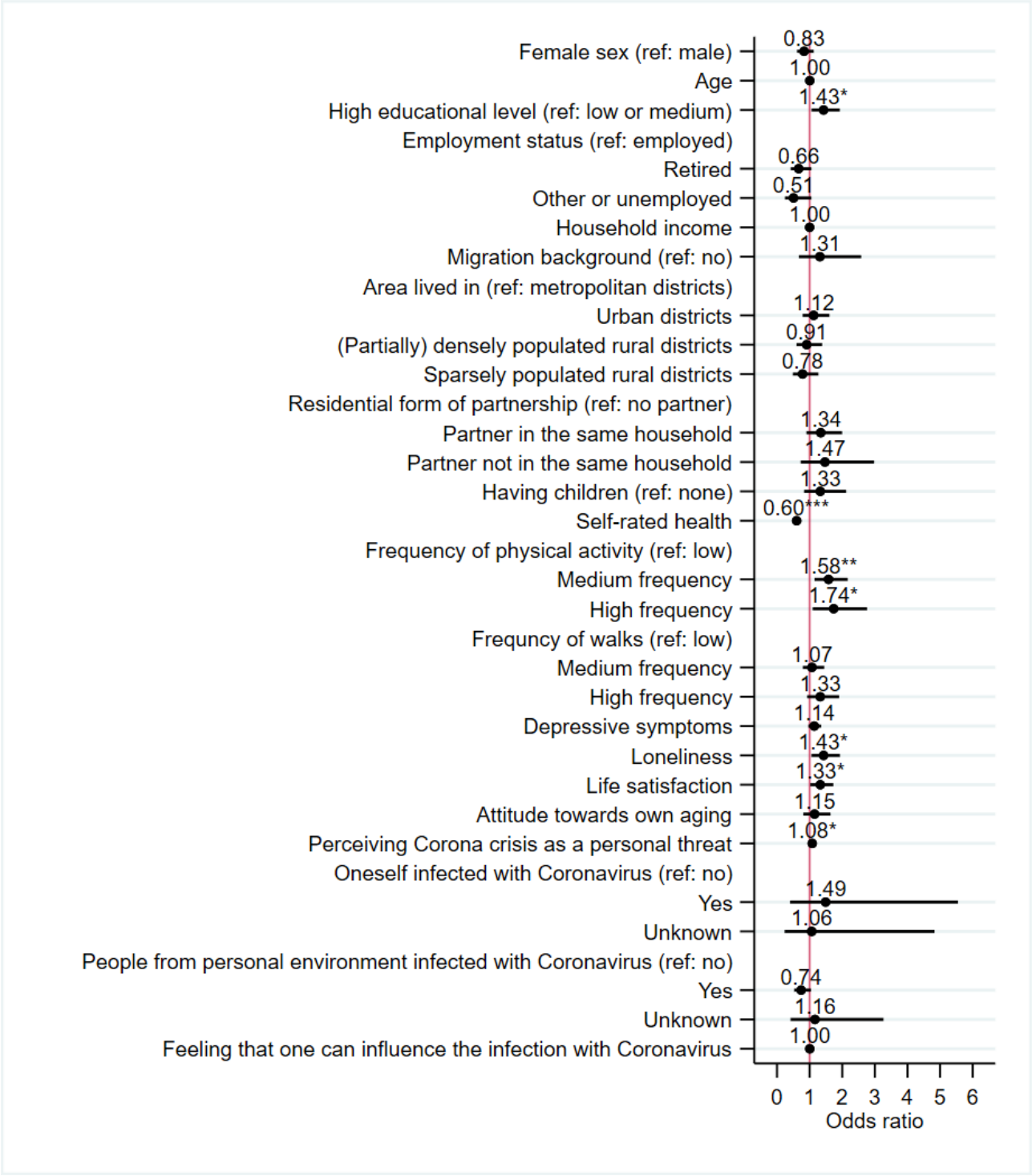
Regression Analysis

The results of the random effects logistic regression are presented in Figure 1 (Table S1 in Multimedia Appendix 1). The majority of the determinants were not significantly associated with online health consultation use during the COVID-19 pandemic. Nevertheless, we found a significant longitudinal association of the outcome with high education (OR 1.43, 95% CI 1.06 - 1.93; *P*=.02), poor self-rated health (OR 0.60, 95% CI 0.49 - 0.75; *P*<.001), higher frequency of physical activity (medium frequency: OR 1.58, 95% CI 1.15 - 2.17; *P*=.005; high frequency: OR 1.73, 95% CI 1.09 - 2.76; *P*=.02), higher loneliness (OR 1.43, 95% CI 1.06 - 1.93; *P*=.04), greater life satisfaction (OR 1.33, 95% CI 1.02 - 1.73; *P*=.04), and perceiving the Corona crisis as a greater personal threat (OR 1.08, 95% CI 1.01 - 1.15; *P*=.02).

In additional analyses, we stratified the sample by sex and age group (Tables S2 and S3 in Multimedia Appendix 1 for more details). While online health consultation use in female participants was only associated with poor self-rated health (OR 0.62, 95% CI 0.45 - 0.85; *P*=.004), in male individuals the outcome was associated with poor self-rated health (OR 0.59, 95% CI 0.44 - 0.79; *P*<.001) as well as high education (OR 1.70, 95% CI 1.12 - 2.56; *P*=.01), living with a partner in the same household (OR 2.27, 95% CI 1.13 - 4.59; *P*=.02), living with a partner not in the same household (OR 2.86, 95% CI 1.04 - 7.87; *P*=.04), higher loneliness (OR 1.94, 95% CI 1.25 - 2.99; *P*=.003), higher frequency of physical activity (medium frequency: OR 1.88, 95% CI 1.22 - 2.90; *P*=.004; high frequency: OR 2.17, 95% CI 1.17 - 4.03; *P*=.01) and walks (high frequency: OR 1.99, 95% CI 1.24 - 3.20; *P*=.004). In participants aged ≤64 years, online health consultation use was associated with poor self-rated health (OR 0.59, 95% CI 0.42 - 0.83; *P*=.003) and higher frequency of physical activity (medium frequency: OR 1.91, 95% CI 1.14 - 3.21; *P*=.01). In older participants (≥65 years), poor self-rated health (OR 0.61, 95% CI 0.46 - 0.81; *P*=.001) as well as older age (OR 1.03, 95% CI 1.00 - 1.07; *P*=.05), high education (OR 1.58, 95% CI 1.06 - 2.37; *P*=.026), and higher loneliness (OR 1.52, 95% CI 1.01 - 2.29; *P*=.04) were associated with online health consultation use.

XSL•FO
RenderX

Figure 1. Results of random effects logistic regression for determinants of online health consultation use during the COVID-19 pandemic (N=5456). Odds ratios with 95% CI are reported. Unless stated otherwise, the reference category is always zero or absence of the characteristic. ref: reference category; *** *P*<.001; ** *P*<.01; * *P*<.05.



Discussion

Principal Findings

Nationally representative longitudinal data from Germany were used to observe the longitudinal association of various determinants with the use of online consultations with doctors or therapists during the COVID-19 pandemic in a large sample of middle-aged and older individuals with access to the internet. Random effects logistic regressions revealed associations of education, health, and psychosocial factors with online health consultation use during the pandemic. In additional analyses stratified by sex and age, self-rated health was negatively associated with the outcome in all groups. Additional relationships with age, education, relationship status, and loneliness were only observed among the male and older (≥65 years) subgroups. Considering the limited evidence regarding determinants of telemedicine use (particularly based on longitudinal data), our longitudinal study considerably extend current knowledge on socioeconomic, health- and health

XSL•FO
RenderX

behavior–related, psychological, and COVID-19-related determinants.

Relation to Previous Research

In contrast to findings from other German samples [18,33,35,36,41,42,66], most of the socioeconomic determinants, such as sex or age, were not associated with online health consultation use in our sample. Nevertheless, a study that exclusively observed an older German sample also found no associations with socioeconomic characteristics [45]. This could imply that socioeconomic characteristics are less relevant for telemedicine use in middle-aged and older patients in Germany. However, the individual's educational level was positively associated with the outcome in our analysis. This is in line with findings from other large German [33,34,36,42] and international samples [38,40,67,68] and could suggest that high education is linked to higher digital literacy as well as access to necessary technical equipment for using telemedicine. A systematic review by Estrela et al [69] examined 36 international articles on digital health literacy and highlighted its positive relationship with education. Correspondingly, in a large US sample of old-aged Medicare beneficiaries (≥65 years), Choi et al [70] found associations of younger age and higher income with telemedicine use during the COVID-19 pandemic, which disappeared when controling for technology-enabling factors (eg, information and communication technology device ownership or use experience). Consequently, technology-enabling factors could be particularly considered when trying to enhance telemedicine use among older patients. For instance, implementing supporting material (eg, leaflets or video instructions), technical support (eg, via telephone) or hybrid care combinations in clinical practice may be helpful to strengthen competence and digital health literacy in older patients.

Health and health behavior were associated with online health consultation use in our sample. Individuals who reported a poorer health status were more likely to use online health consultations during the pandemic. Likewise, severe health limitations or poor health were associated with telemedicine use in the German Survey of Health, Ageing and Retirement in Europe (SHARE) sample of middle-aged and older adults [45] and in international samples [68,71]. This association could have been caused by greater health needs in unhealthy individuals or may be connected to precautions due to the COVID-19 pandemic. Likewise, perceiving the Corona crisis as a greater personal threat was associated with online health consultation use. Telemedicine might have presented a treatment option for patients who were scared of becoming infected with the Coronavirus and wanted to avoid personal contact (eg, [72]). Fear of the Coronavirus and pandemic-related challenges were also associated with telemedicine use in other German samples during the pandemic [36,41,42,45]. Other COVID-19–related factors such as the infection of oneself or close others with the virus were not associated with the outcome in our sample. A reason for that could be that only a small proportion of our pooled sample was infected with the virus (1%) or knew someone in their close personal environment who was infected (15.6%). Furthermore, the frequency of physical activity was positively associated with online consultation use. Physical activity is an important determinant of health and was associated with higher utilization of preventive or office-based health services and lower use of inpatient or emergency care among adults in previous international research (eg, [73,74]). Therefore, physically active individuals seem to show greater levels of health awareness, which might have been connected to the higher telemedicine use in this group.

Some psychological factors were associated with the outcome in our sample. Whereas depressive symptoms and the attitude towards one's own aging did not show an association, psychosocial factors including loneliness and life satisfaction had a significant relationship with online health consultation use. In the German SHARE sample, a positive association of depressive symptoms as well as loneliness with telemedicine use in middle-aged and older individuals was observed [45]. The mixed evidence concerning the relationship of depressive symptoms with telemedicine utilization might be explained by the different pandemic periods that were considered in the studies (Summer and Winter 2020 vs Summer 2021). In fact, depressive symptoms [75] and telemedicine acceptance [35] were found to have increased over the course of the pandemic. Regarding loneliness, Robbins et al [76] also observed higher loneliness rates among telemedicine users (telephone contacts), while in-person visits were associated with fewer feelings of loneliness among older adults (≥65 years) residing in the United States during the pandemic. It might be the case that older adults who indicated higher levels of loneliness were more open to using telemedicine services to satisfy their unmet social needs during pandemic times. Telemedicine services may be more accessible for older individuals (eg, no traveling for mobility-restricted individuals needed), which might encourage lonely individuals to take the initiative to foster social interaction through telemedicine appointments. Regarding life satisfaction, König et al [77] recently observed a positive correlation between life satisfaction and digital health literacy in a nationally representative survey of the population in Germany, which could explain the observed relationship between higher life satisfaction and online health consultation use in our sample. In addition, higher life satisfaction was associated with health-promoting behaviors in previous studies [78-80], which could have contributed to higher online health consultation use.

Additional analyses stratified by sex and age further highlighted the relationships with the determinants, especially in male and older participants. While online consultation use in females was only associated with health needs, male individual's use behavior was additionally associated with psychosocial factors (ie, relationship status and loneliness), education, and health behavior (ie, frequency of walks and physical activity). The decision of male patients to use telemedicine services seems to be connected to additional factors and therefore more complex compared with female patients. Male patients may be facing additional barriers to telemedicine utilization. Therefore, future research is needed to explore gender-specific determinants and barriers to the use of telemedicine. Furthermore, particularly among individuals aged 65 years and older, loneliness, education, and age were associated with online health consultation use. Consequently, compared with individuals aged

XSL•FO
RenderX

40 to 64 years, older patients might be especially affected by disparities in education or loneliness.

When considering the international context, multiple reviews mainly based on quantitative cross-sectional, randomized controlled or qualitative studies observed determinants of telemedicine use in older age groups and found positive associations with educational level, health needs, and health-related motivation to use the services [38-40], which is in line with our findings. Moreover, they observed additional relationships (eg, with age, sex, and social support or influence), which we did not find in our analysis. Age or sex were not associated with overall telemedicine use in our sample. The reviews [38-40] mostly included studies from the United States or Europe. Telemedicine regulations differ substantially between Germany and the United States, but also in the European context. Whereas Germany has implemented national telemedicine coverage rules during the pandemic, large state-specific variation in insurance coverage and regulations exist in the United States [71,81], which may have caused additional barriers for older telemedicine users. Nevertheless, our stratified analyses indicated differences in determinants of use in the different sex and age groups. Furthermore, we observed a positive relationship of loneliness with telemedicine use, which is in contrast to the observed negative relationship of social isolation or lack of social support with telemedicine use in international studies [39]. Raja et al [82] reviewed studies of older adults in European countries and found that social support and lack of social support were both associated with using new technologies, including telemedicine. Social support may be crucial when older adults are faced with problems when learning new technologies, which was also observed in US samples [68]. Nevertheless, lack of social support or loneliness might also motivate older adults to try out new technologies to address their social needs, which we observed in our sample. Moreover, the discrepancies might also be explained by differences in regulations or access to telemedicine care (eg, variations in out-of-pocket payments, supply, or complexity of use). For instance, the requirement for additional out-of-pocket payments or high barriers to use may lower the probability of using telemedicine services for the primary purpose of social interaction. Future research concerning differences in psychosocial determinants across different countries is needed.

Strengths and Limitations

The nationally representative, large DEAS sample of middle-aged and older individuals in Germany represents a key strength of our study. Middle-aged and older adults are a major target group for future telemedicine services in Germany, thus it is of particular importance to explore the telemedicine use behavior of this age group. In addition, longitudinal data were exploited, which enabled us to consider two different pandemic stages and account for the exceptional circumstances during that time. Since only few studies observed determinants of online health consultation use in German middle-aged and older adults in the past, our study adds valuable knowledge to the existing literature.

Nevertheless, some limitations should be noted. Telemedicine use was represented by having online consultations with doctors or therapists in our study. We neither examined specific patient groups nor focused on a certain telemedicine format (eg, video conferences or mobile apps). Therefore, we observed a potentially heterogeneous user group. Since telemedicine acceptance during the pandemic was found to vary across medical specialties and telemedicine formats [31], future research that tests for differences in use among patient groups or different telemedicine formats is needed to tailor future services to major user groups. However, our study provides initial insights into telemedicine use in middle and old age. In addition, the DEAS panel holds a slight selection bias. Young, highly educated, healthier, and female individuals were somewhat more likely to participate in the DEAS [83]. However, selection bias in the DEAS sample was found to be small and the distribution of major sociodemographic characteristics closely mirrors the distribution within the overall population of Germany [47]. Moreover, only individuals with access to the internet were included in our study. Therefore, generalization of the results might be slightly limited for some groups of the German middle-aged and older population.

Conclusion

Telemedicine services represent a valuable tool to deal with the increasing demand for health care caused by population aging. Knowledge about telemedicine use and its determinants, particularly in middle-aged and older individuals, is essential to promote widespread implementations in the future. Our study highlights the relationship of education, psychosocial, and health factors with telemedicine use of community-dwelling middle-aged and older individuals in Germany during the COVID-19 pandemic. Therefore, telemedicine use does not only depend on health needs of middle-aged and older patients. The finding that particularly highly educated individuals used online health consultations may point toward social inequality among telemedicine users. Consequently, efforts should be made to enable access to telemedicine for all patient groups and individual support should be provided (eg, for patients with low [digital] health literacy) to remove barriers to telemedicine use. Moreover, special attention should be paid to individuals with low life satisfaction and an unhealthy lifestyle since they seem harder to reach through telemedicine services. Finally, future research is needed to test the relevance of the observed relationships in the postpandemic context and identify potential reasons for use or nonuse of telemedicine services in middle-aged and older adults in Germany (eg, based on qualitative data). Moreover, cross-country comparisons regarding the determinants of telemedicine use remain to be explored.

Acknowledgments

We acknowledge financial support from the Open Access Publication Fund of UKE - Universitätsklinikum Hamburg-Eppendorf.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Results of random effects logistic regression for determinants of online health consultation use during the COVID-19 pandemic and additional analyses stratified by sex and age.
[DOCX File, 34 KB - aging_v8i1e60311_app1.docx]

References

1. Lehoux P, Roncarolo F, Silva HP, Boivin A, Denis JL, Hébert R. What health system challenges should responsible innovation in health address? Insights from an international scoping review. Int J Health Policy Manag 2019 Feb 1;8(2):63-75. [doi: 10.15171/ijhpm.2018.110] [Medline: 30980619]
2. Mahishale V. Ageing world: health care challenges. J Sci Soc 2015;42(3):138. [doi: 10.4103/0974-5009.165540]
3. Global status report on noncommunicable diseases 2010. World Health Organization. 2011. URL: https://iris.who.int/bitstream/handle/10665/44579/9789240686458_eng.pdf?sequence=1&isAllowed=y [accessed 2025-05-20]
4. Khemapech T. Telemedicine – meaning, challenges and opportunities. Siriraj Med Bull 2019 May 1;71(3):246-252. [doi: 10.33192/Smj.2019.38]
5. Telemedicine: opportunities and developments in member states: report on the second global survey on ehealth. WHO Global Observatory for eHealth. 2010. URL: https://iris.who.int/bitstream/handle/10665/44497/9789241564144_eng.pdf?sequence=1&isAllowed=y [accessed 2025-05-22]
6. Hjelm NM. Benefits and drawbacks of telemedicine. J Telemed Telecare 2005;11(2):60-70. [doi: 10.1258/1357633053499886] [Medline: 15829049]
7. Meneses ADF, Pimentel FF, da Cruz JPF, Candido Dos Reis FJ. Experiences of women with breast cancer using telehealth: a qualitative systematic review. Clin Breast Cancer 2023 Feb;23(2):101-107. [doi: 10.1016/j.clbc.2022.11.001] [Medline: 36464603]
8. Goharinejad S, Hajesmaeel-Gohari S, Jannati N, Goharinejad S, Bahaadinbeigy K. Review of systematic reviews in the field of telemedicine. Med J Islam Repub Iran 2021;35(1):184. [doi: 10.47176/mjiri.35.184] [Medline: 36042824]
9. Eze ND, Mateus C, Cravo Oliveira Hashiguchi T. Telemedicine in the OECD: An umbrella review of clinical and cost-effectiveness, patient experience and implementation. PLoS ONE 2020;15(8):e0237585. [doi: 10.1371/journal.pone.0237585] [Medline: 32790752]
10. Kichloo A, Albosta M, Dettloff K, et al. Telemedicine, the current COVID-19 pandemic and the future: a narrative review and perspectives moving forward in the USA. Fam Med Community Health 2020 Aug;8(3):e000530. [doi: 10.1136/fmch-2020-000530] [Medline: 32816942]
11. Atmojo JT, Sudaryanto WT, Widiyanto A, et al. Telemedicine, cost effectiveness, and patients satisfaction: a systematic review. J Health Policy Manage 2020;5(2):103-107 [FREE Full text] [doi: 10.26911/thejhpm.2020.05.02.02]
12. Nguyen M, Waller M, Pandya A, Portnoy J. A review of patient and provider satisfaction with telemedicine. Curr Allergy Asthma Rep 2020 Sep 22;20(11):72. [doi: 10.1007/s11882-020-00969-7] [Medline: 32959158]
13. Shaver J. The state of telehealth before and after the COVID-19 pandemic. Prim Care 2022 Dec;49(4):517-530. [doi: 10.1016/j.pop.2022.04.002] [Medline: 36357058]
14. Friedman AB, Gervasi S, Song H, et al. Telemedicine catches on: changes in the utilization of telemedicine services during the COVID-19 pandemic. Am J Manag Care 2022 Jan 1;28(1):e1-e6. [doi: 10.37765/ajmc.2022.88771] [Medline: 35049260]
15. Seivert S, Badowski ME. The rise of telemedicine: lessons from a global pandemic. EMJ Innov 2021;5(1):64-69. [doi: 10.33590/emjinnov/20-00239]
16. Nittari G, Savva D, Tomassoni D, Tayebati SK, Amenta F. Telemedicine in the COVID-19 era: a narrative review based on current evidence. Int J Environ Res Public Health 2022 Apr 22;19(9):5101. [doi: 10.3390/ijerph19095101] [Medline: 35564494]
17. Koonin LM, Hoots B, Tsang CA, et al. Trends in the use of telehealth during the emergence of the COVID-19 pandemic - United States, January-March 2020. MMWR Morb Mortal Wkly Rep 2020 Oct 30;69(43):1595-1599. [doi: 10.15585/mmwr.mm6943a3] [Medline: 33119561]
18. Heuer J, Osterwald A, Akmatov MK, Zentralinstitut für die kassenärztliche Versorgung in Deutschland (Zi), et al. Telemedizin als alternativer Zugang zu vertragsärztlicher ambulanter Versorgung–Trends im Zeitraum 2017 bis 2021 [Article in German]. Versorgungsatlas-Bericht Nr 2023;23(6):1-21. [doi: 10.20364/VA-23.06]
19. Heidemann C, Reitzle L, Schmidt C, Fuchs J, Prütz F, Scheidt-Nave C. Non-utilisation of health care services during the COVID-19 pandemic: results of the CoMoLo study. J Health Monit 2022 Mar;7(Suppl 1):2-17. [doi: 10.25646/9564] [Medline: 35356069]
20. Schäfer I, Haack A, Neumann M, Koch-Gromus U, Scherer M, Petersen E. Non-utilisation of medical services during the COVID-19 pandemic among persons with chronic diseases. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2023 Mar;66(3):275-282. [doi: 10.1007/s00103-023-03665-9] [Medline: 36723636]

XSL•FO
RenderX

21. Gesetz für eine bessere Versorgung durch Digitalisierung und Innovation (Digitale-Versorgung-Gesetz—DVG) [digital healthcare act]. BGBI I 2562. 2019. URL: https://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&start=%2F%2F%2A%5B%40attr_id=%27bgbl119s2562.pdf%27%5D#/switch/tocPane?_ts=1745927445559 [accessed 2025-05-22]
22. Verordnung über das Verfahren und die Anforderungen zur Prüfung der Erstattungsfähigkeit digitaler Gesundheitsanwendungen in der gesetzlichen Krankenversicherung (Digitale Gesundheitsanwendungen-Verordnung - DIGAV) [digital health applications ordinance]. BGBl I. 2020. URL: https://www.gesetze-im-internet.de/digav/BJNR076800020.html [accessed 2025-05-20]
23. Passon A, Lüngen M, Gerber A, Redaelli M, Stock S, In: Lauterbach K, Stock S, Brunner H, editors. Gesundheitsökonomie. Das Krankenversicherungssystem in Deutschland: Verlag Hans Huber, Hogrefe AG; 2009:105-136.
24. Şahin E, Yavuz Veizi BG, Naharci MI. Telemedicine interventions for older adults: a systematic review. J Telemed Telecare 2024 Feb;30(2):305-319. [doi: 10.1177/1357633X211058340] [Medline: 34825609]
25. van den Berg N, Schumann M, Kraft K, Hoffmann W. Telemedicine and telecare for older patients--a systematic review. Maturitas 2012 Oct;73(2):94-114. [doi: 10.1016/j.maturitas.2012.06.010] [Medline: 22809497]
26. Soh YY, Zhang H, Toh JJY, Li X, Wu XV. The effectiveness of tele-transitions of care interventions in high-risk older adults: a systematic review and meta-analysis. Int J Nurs Stud 2023 Mar;139:104428. [doi: 10.1016/j.ijnurstu.2022.104428] [Medline: 36682322]
27. Alsabeeha NHM, Atieh MA, Balakrishnan MS. Older adults' satisfaction with telemedicine during the COVID-19 pandemic: a systematic review. Telemed J E Health 2023 Jan;29(1):38-49. [doi: 10.1089/tmj.2022.0045] [Medline: 35613380]
28. Reinhardt G, Schwarz PEH, Harst L. Non-use of telemedicine: a scoping review. Health Informatics J 2021;27(4):14604582211043147. [doi: 10.1177/14604582211043147] [Medline: 34696613]
29. Saigí-Rubió F, Borges do Nascimento IJ, Robles N, et al. The current status of telemedicine technology use across the world health organization european region: an overview of systematic reviews. J Med Internet Res 2022 Oct 27;24(10):e40877. [doi: 10.2196/40877] [Medline: 36301602]
30. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556. [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
31. Drake C, Lian T, Cameron B, Medynskaya K, Bosworth HB, Shah K. Understanding telemedicine's "New Normal": variations in telemedicine use by specialty line and patient demographics. Telemed J E Health 2022 Jan;28(1):51-59. [doi: 10.1089/tmj.2021.0041] [Medline: 33769092]
32. Ternes S, Lavin L, Vakkalanka JP, et al. The role of increasing synchronous telehealth use during the COVID-19 pandemic on disparities in access to healthcare: A systematic review. J Telemed Telecare 2024 Apr 22;2024:1357633X241245459. [doi: 10.1177/1357633X241245459] [Medline: 38646804]
33. Reitzle L, Schmidt C, Färber F, et al. Perceived access to health care services and relevance of telemedicine during the COVID-19 pandemic in Germany. Int J Environ Res Public Health 2021 Jul 19;18(14):7661. [doi: 10.3390/ijerph18147661] [Medline: 34300110]
34. Uncovska M, Freitag B, Meister S, Fehring L. Patient acceptance of prescribed and fully reimbursed mHealth apps in Germany: an UTAUT2-based online survey study. J Med Syst 2023 Jan 27;47(1):14. [doi: 10.1007/s10916-023-01910-x] [Medline: 36705853]
35. von der Groeben S, Czaplicki A, Hegerl U, Reich H. Telemedicine during the COVID-19 pandemic in Germany: Results from three nationally representative surveys on use, attitudes and barriers among adults affected by depression. Internet Interv 2023 Apr;32:100622. [doi: 10.1016/j.invent.2023.100622] [Medline: 37091132]
36. Zöllner JP, Noda AH, McCoy J, et al. Use of health-related apps and telehealth in adults with epilepsy in Germany: a multicenter cohort study. Telemed J E Health 2023 Apr;29(4):540-550. [doi: 10.1089/tmj.2022.0238] [Medline: 35984859]
37. Kis A, Augustin M, Augustin J. Regional healthcare delivery and demographic change in Germany – scenarios for dermatological care in 2035. J Deutsche Derma Gesell 2017 Dec;15(12):1199-1209. [doi: 10.1111/ddg.13379]
38. Kampmeijer R, Pavlova M, Tambor M, Golinowska S, Groot W. The use of e-health and m-health tools in health promotion and primary prevention among older adults: a systematic literature review. BMC Health Serv Res 2016 Sep 5;16 Suppl 5(Suppl 5):290. [doi: 10.1186/s12913-016-1522-3] [Medline: 27608677]
39. Turcotte S, Bouchard C, Rousseau J, et al. Factors influencing older adults' participation in telehealth interventions for primary prevention and health promotion: A rapid review. Australas J Ageing 2024 Mar;43(1):11-30. [doi: 10.1111/ajag.13244] [Medline: 38014903]
40. Wang J, Fu Y, Lou V, Tan SY, Chui E. A systematic review of factors influencing attitudes towards and intention to use the long-distance caregiving technologies for older adults. Int J Med Inform 2021 Sep;153:104536. [doi: 10.1016/j.ijmedinf.2021.104536] [Medline: 34325206]
41. Hajek A, De Bock F, Merkel C, Kretzler B, König HH. Use of and satisfaction with telemedicine services during the pandemic: findings from the COVID-19 snapshot monitoring in Germany (COSMO). Healthcare (Basel) 2022 Jan 4;10(1):92. [doi: 10.3390/healthcare10010092] [Medline: 35052257]

XSL•FO
RenderX

42. Hannemann N, Götz NA, Schmidt L, Hübner U, Babitsch B. Patient connectivity with healthcare professionals and health insurer using digital health technologies during the COVID-19 pandemic: a German cross-sectional study. BMC Med Inform Decis Mak 2021 Aug 25;21(1):250. [doi: 10.1186/s12911-021-01605-8] [Medline: 34433452]
43. Rauschenberg C, Schick A, Goetzl C, et al. Social isolation, mental health, and use of digital interventions in youth during the COVID-19 pandemic: a nationally representative survey. Eur Psychiatry 2021 Mar 9;64(1):e20. [doi: 10.1192/j.eurpsy.2021.17] [Medline: 33686930]
44. Hajek A, König HH. Frequency and correlates of online consultations with doctors or therapists in middle-aged and older adults: nationally representative cross-sectional study. J Med Internet Res 2022 Apr 7;24(4):e29781. [doi: 10.2196/29781] [Medline: 35389360]
45. Neumann A, König HH, Hajek A. Determinants of telemedicine service use among middle-aged and older adults in Germany during the COVID-19 pandemic: cross-sectional survey study. JMIR Aging 2024 Apr 23;7:e50938. [doi: 10.2196/50938] [Medline: 38654578]
46. Mandal S, Wiesenfeld BM, Mann D, et al. Evidence for telemedicine's ongoing transformation of health care delivery since the onset of COVID-19: retrospective observational study. JMIR Form Res 2022 Oct 14;6(10):e38661. [doi: 10.2196/38661] [Medline: 36103553]
47. Klaus D, Engstler H, Mahne K, et al. Cohort profile: the German ageing survey (DEAS). Int J Epidemiol 2017 Aug 1;46(4):1105-1105g. [doi: 10.1093/ije/dyw326] [Medline: 28180273]
48. Vogel C, Wettstein M, Klaus D, Spuling S, Kortmann L, Alcántara AL, et al. German Ageing Survey: Instruments of the Compact Survey 2020: German centre of gerontology; 2021. [doi: 10.1007/978-3-319-69892-2_1115-1]
49. Vogel C, Wettstein M, Klaus D, Spuling S, Kortmann L, Alcántara AL, et al. German ageing survey, instruments of the survey 2020/2021: German centre of gerontology; 2021. [doi: 10.1007/978-3-030-22009-9_1115]
50. Engstler H. Deutscher Alterssurvey (DEAS): Kurzbeschreibung des SUF DEAS2020, version 1.0 (märz 2021) [Website in German]. : Deutsches Zentrum für Altersfragen; 2021. URL: https://nbn-resolving.org/urn:nbn:de:0168-ssoar-72029-0 [accessed 2025-05-20]
51. Stuth S. Deutscher Alterssurvey (DEAS): Kurzbeschreibung des SUF DEAS2020/21, version 11 [Website in German]. : Deutsches Zentrum für Altersfragen; 2022. URL: https://nbn-resolving.org/urn:nbn:de:0168-ssoar-79741-0 [accessed 2025-05-20]
52. United nations educational scientific and cultural organization. International Standard Classification of Education: ISCED 1997: UNESCO Institute for Statistics 2006. URL: https://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-1997-en_0.pdf [accessed 2025-05-23]
53. Andresen EM, Malmgren JA, Carter WB, Patrick DL. Screening for depression in well older adults: evaluation of a short form of the CES-D. Am J Prev Med 1994 Mar;10(2):77-84. [doi: 10.1016/S0749-3797(18)30622-6]
54. Hautzinger M, Bailer M. Allgemeine Depressions-Skala (ADS)[General Depression Scale]. Beltz 1993.
55. Mohebbi M, Nguyen V, McNeil JJ, et al. Psychometric properties of a short form of the Center for Epidemiologic Studies Depression (CES-D-10) scale for screening depressive symptoms in healthy community dwelling older adults. Gen Hosp Psychiatry 2018;51:118-125. [doi: 10.1016/j.genhosppsych.2017.08.002] [Medline: 28890280]
56. Hautzinger M. Die CES-D Skala: Ein Depressionsmesinstrument für Untersuchungen in der Allgemeinbevölkerung. [The CES-D scale: a depression-rating scale for research in the general population.]. Diagnostica 1988;34(2):167-173.
57. Gierveld JDJ, Tilburg TV. A 6-item ccale for overall, emotional, and social loneliness: confirmatory tests on survey data. Res Aging 2006;28(5):582-598. [doi: 10.1177/0164027506289723]
58. Lawton MP. The Philadelphia Geriatric Center Morale Scale: a revision. J Gerontol 1975 Jan;30(1):85-89. [doi: 10.1093/geronj/30.1.85] [Medline: 1109399]
59. Liang J, Bollen KA. The structure of the Philadelphia Geriatric Center Morale scale: a reinterpretation. J Gerontol 1983 Mar;38(2):181-189. [doi: 10.1093/geronj/38.2.181] [Medline: 6827034]
60. Jung S, Siedlecki KL. Attitude toward own aging: age invariance and construct validity across middle-aged, young-old, and old-old adults. J Adult Dev 2018 Jun;25(2):141-149. [doi: 10.1007/s10804-018-9283-3]
61. Diener E, Emmons RA, Larsen RJ, Griffin S. The satisfaction with life scale. J Pers Assess 1985 Feb;49(1):71-75. [doi: 10.1207/s15327752jpa4901_13] [Medline: 16367493]
62. Schumacher J. S. Satisfaction with life scale. In: Klaiberg A, Brähler E, editors. Diagnostische Verfahren Zu Lebensqualität Und Wohlbefinden2003.
63. Glaesmer H, Grande G, Braehler E, Roth M. The German version of the satisfaction with life scale (SWLS). Eur J Psychol Assess 2011 Jan;27(2):127-132. [doi: 10.1027/1015-5759/a000058]
64. Andreß HJ, Golsch K, Schmidt AW. Panel analysis of categorical variables. In: Andreß HJ, Golsch K, Schmidt AW, editors. Applied Panel Data Analysis for Economic and Social Surveys: Springer Berlin Heidelberg; 2013:203-285. [doi: 10.1007/978-3-642-32914-2_5]
65. Hausman JA. Specification tests in econometrics. Econometrica 1978 Nov;46(6):1251. [doi: 10.2307/1913827]
66. Muehlensiepen F, Petit P, Knitza J, Welcker M, Vuillerme N. Factors associated with telemedicine use among patients with rheumatic and musculoskeletal disease: secondary analysis of data from a German nationwide survey. J Med Internet Res 2023 Jan 27;25:e40912. [doi: 10.2196/40912] [Medline: 36705950]

XSL•FO
RenderX

67. Frydman JL, Li W, Gelfman LP, Liu B. Telemedicine uptake among older adults during the COVID-19 pandemic. Ann Intern Med 2022 Jan;175(1):145-148. [doi: 10.7326/M21-2972] [Medline: 34748380]
68. Li W, Ornstein KA, Li Y, Liu B. Barriers to learning a new technology to go online among older adults during the COVID-19 pandemic. J Am Geriatr Soc 2021 Nov;69(11):3051-3057. [doi: 10.1111/jgs.17433] [Medline: 34409589]
69. Estrela M, Semedo G, Roque F, Ferreira PL, Herdeiro MT. Sociodemographic determinants of digital health literacy: A systematic review and meta-analysis. Int J Med Inform 2023 Sep;177:105124. [doi: 10.1016/j.ijmedinf.2023.105124] [Medline: 37329766]
70. Choi NG, DiNitto DM, Marti CN, Choi BY. Telehealth use among older adults during COVID-19: associations with sociodemographic and health characteristics, technology device ownership, and technology learning. J Appl Gerontol 2022 Mar;41(3):600-609. [doi: 10.1177/07334648211047347] [Medline: 34608821]
71. Harju A, Neufeld J. Telehealth utilization during the COVID-19 pandemic: a preliminary selective review. Telemed Rep 2022;3(1):38-47. [doi: 10.1089/tmr.2021.0040] [Medline: 35720447]
72. Zubair Elahi M, Liang G, Jawad Malik M, Dilawar S, Ilyas B. Fear of Covid-19 and intentions towards adopting E-Health services: exploring the technology acceptance model in the scenario of pandemic. International Journal of Business, Economics and Management 2021;8(4):270-291 [FREE Full text] [doi: 10.18488/journal.62.2021.84.270.291]
73. Jacobs JM, Rottenberg Y, Cohen A, Stessman J. Physical activity and health service utilization among older people. J Am Med Dir Assoc 2013 Feb;14(2):125-129. [doi: 10.1016/j.jamda.2012.10.023] [Medline: 23218744]
74. Kang SW, Xiang X. Physical activity and health services utilization and costs among U.S. adults. Prev Med 2017 Mar;96:101-105. [doi: 10.1016/j.ypmed.2016.12.043]
75. Gobbi S, Płomecka MB, Ashraf Z, et al. Worsening of preexisting psychiatric conditions during the COVID-19 pandemic. Front Psychiatry 2020;11:581426. [doi: 10.3389/fpsyt.2020.581426] [Medline: 33391049]
76. Robbins R, DiClemente RJ, Baig N, Johnson A, Chou A, Van den Bulck J. Digital communications technology use and feelings of anxiety, depression, and loneliness among older adults during the COVID-19 pandemic. J Appl Gerontol 2023 Sep;42(9):1911-1920. [doi: 10.1177/07334648231169086] [Medline: 37041750]
77. König L, Kuhlmey A, Suhr R. Digital health literacy of the population in Germany and its association with physical health, mental health, life satisfaction, and health behaviors: nationally representative survey study. JMIR Public Health Surveill 2024 Feb 21;10:e48685. [doi: 10.2196/48685] [Medline: 38381497]
78. Grant N, Wardle J, Steptoe A. The relationship between life satisfaction and health behavior: A cross-cultural analysis of young adults. Int J Behav Med 2009;16(3):259-268. [doi: 10.1007/s12529-009-9032-x] [Medline: 19319695]
79. Martín-María N, Caballero FF, Moreno-Agostino D, et al. Relationship between subjective well-being and healthy lifestyle behaviours in older adults: a longitudinal study. Aging Ment Health 2020 Apr 2;24(4):611-619. [doi: 10.1080/13607863.2018.1548567]
80. Bai A, Li H, Huang Y, et al. A survey of overall life satisfaction and its association with breast diseases in Chinese women. Cancer Med 2016 Jan;5(1):111-119. [doi: 10.1002/cam4.565] [Medline: 26640035]
81. Gerke S, Stern AD, Minssen T. Germany's digital health reforms in the COVID-19 era: lessons and opportunities for other countries. NPJ Digit Med 2020;3(1):94. [doi: 10.1038/s41746-020-0306-7] [Medline: 32685700]
82. Raja M, Bjerkan J, Kymre IG, Galvin KT, Uhrenfeldt L. Telehealth and digital developments in society that persons 75 years and older in European countries have been part of: a scoping review. BMC Health Serv Res 2021 Oct 26;21(1):1157. [doi: 10.1186/s12913-021-07154-0] [Medline: 34696789]
83. Schiel S, Marcos J, Schulz S, Dickmann C, Alterssurvey F. Methodenbericht deutscher alterssurvey (DEAS): durchführung der 7. erhebungswelle 2020/2021 [article in german]. : INFAS Institut für angewandte Sozialwissenschaft GmbH; 2021 URL: https://www.ssoar.info/ssoar/bitstream/handle/document/79743/ssoar-2021-schiel_et_al-Methodenbericht_Deutscher_Alterssurvey_DEAS_Durchfuhrung.pdf?sequence=1&isAllowed=y&lnkname=ssoar-2021-schiel_et_al-Methodenbericht_Deutscher_Alterssurvey_DEAS_Durchfuhrung.pdf [accessed 2025-05-23]

Abbreviations

CES-D: Center of Epidemiologic Studies - Depression scale
DEAS: Deutsches Alterssurvey (German Ageing Survey)
ISCED: International Standard Classification of Education
OR: odds ratio
SHARE: Survey of Health, Ageing and Retirement in Europe
WHO: World Health Organization

XSL•FO
RenderX

Edited by G Seçkin; submitted 07.05.24; peer-reviewed by D Singh, L Mao, W Li; revised version received 06.02.25; accepted 23.02.25; published 26.05.25.

Please cite as:
Neumann A, König HH, Hajek A
Determinants of Having Online Health Consultations During the COVID-19 Pandemic Among Middle-Aged and Older Adults in Germany: Representative Longitudinal Survey Study
JMIR Aging 2025;8:e60311
URL: https://aging.jmir.org/2025/1/e60311
doi:10.2196/60311

© Ariana Neumann, Hans-Helmut König, André Hajek. Originally published in JMIR Aging (https://aging.jmir.org), 26.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Revisits, Readmission, and Mortality From Emergency Department Admissions for Older Adults With Vague Presentations: Longitudinal Observational Study

Sebastian Alejandro Alvarez Avendano[1], PhD; Amy Cochran[2], PhD; Valerie Odeh Couvertier[1], PhD; Brian Patterson[3], MPH, MD; Manish Shah[3], MPH, MD; Gabriel Zayas-Caban[1], PhD

[1]Department of Industrial and Systems Engineering, College of Engineering, University of Wisconsin-Madison, 3107 Mechanical Engineering Building, 1513 University Avenue, Madison, WI, United States

[2]Population Health Sciences, School of Medicine and Public Health, University of Wisconsin–Madison, Madison, WI, United States

[3]BerbeeWalsh Department of Emergency Medicine, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI, United States

Corresponding Author:
Gabriel Zayas-Caban, PhD
Department of Industrial and Systems Engineering, College of Engineering, University of Wisconsin-Madison, 3107 Mechanical Engineering Building, 1513 University Avenue, Madison, WI, United States

Abstract

Background: Older adults (65 years and older) often present to the emergency department (ED) with an unclear need for hospitalization, leading to potentially harmful and costly care. This underscores the importance of measuring the trade-off between admission and discharge for these patients in terms of patient outcomes.

Objective: This study aimed to measure the relationship between disposition decisions and 3-day, 9-day, and 30-day revisits, readmission, and mortality, using causal inference methods that adjust for potential measured and unmeasured confounding.

Methods: A longitudinal observational study (n=3591) was conducted using electronic health records from a large tertiary teaching hospital with an ED between January 1, 2014 and September 27, 2018. The sample consisted of older adult patients with 1 of 6 presentations with significant variability in admission: falls, weakness, syncope, urinary tract infection, pneumonia, and cellulitis. The exposure under consideration was the ED disposition decision (admission to the hospital or discharge). Nine outcome variables were considered: ED revisits, hospital readmission, and mortality within 3, 9, and 30 days of being discharged from either the hospital for admitted patients or the ED for discharged patients.

Results: Admission was estimated to significantly decrease the risk of an ED revisit after discharge (30-day window: −6.4%, 95% CI −7.8 to −5.0), while significantly increasing the risk of hospital readmission (30-day window: 5.8%, 95% CI 5.0 to 6.5) and mortality (30-day window: 1.0%, 95% CI 0.4 to 1.6). Admission was found to be especially adverse for patients with weakness and pneumonia, and relatively less adverse for older adult patients with falls and syncope.

Conclusions: Admission may not be the safe option for older adults with gray area presentations, and while revisits and readmissions are commonly used to evaluate the quality of care in the ED, their divergence suggests that caution should be used when interpreting either in isolation.

(JMIR Aging 2025;8:e55929) doi:10.2196/55929

KEYWORDS

gerontology; geriatric; older adults; elderly; older people; aging; emergency department; emergency room; ED; disposition decision; disposition; discharge; admission; revisit; readmission; observational study; health; hospital

Introduction

Care for acute illnesses has shifted from outpatient offices to emergency departments (EDs), leading to an increase in ED use that has outpaced population growth [1]. EDs diagnose and treat acute illnesses [2]. Therefore, emergency providers diagnose patients, initiate treatment, and predict the disease trajectory to decide whether to admit the patient to an inpatient unit or discharge the patient home [3,4]. Any change to admission decisions can impact outcomes and costs since two-thirds of ED health care costs in the United States come from visits that end in admission [5,6].

Whether to admit or discharge a patient is a weighty decision. This point is best exemplified with older adults (≥65 years of age). Discharging an older adult carries a high risk of adverse health outcomes, especially when compared with younger

XSL•FO
RenderX

patients [7-12]. Hence, increasing interest has been placed on identifying patients at risk for adverse outcomes after ED discharge [7,11,13,14] or on developing strategies for following up on discharged patients [15]. Admitting a patient who can be discharged carries its own risks. Older adults are vulnerable to deconditioning and hospital acquired infection, as well as developing delirium and accelerated cognitive decline [16-19]. These issues underscore the importance of measuring the trade-off between admitting and discharging a patient.

Admission decisions are usually based on well-defined clinical factors and practice guidelines, but often, patients fall into a gray area in which the need for hospital admission is unclear based on objective information or even local standards of care. A large group of such gray-area patients is those presenting with syndromic diagnoses such as falls [20] or weakness [21], or patients presenting with more definite diagnoses that are associated with wide variability in admission decision, such as syncope [22], chest pain [5,23,24], urinary tract infection (UTI) [25], pneumonia [26], and cellulitis [27]. Moreover, other factors influence disposition decisions including triage [28], crowding [29], patient home environment [30], diagnostic testing [31], patient ethnicity [32], and hospital capacity [32]. Consequently, admission rates vary widely for these patients between providers and between hospitals, leading to potentially harmful practice variation [33]. Reducing this variation may avoid harmful admissions while safeguarding patient safety, but in order to do so evidence must be used to guide decision-making.

In this study, we focused on older adults in the ED with gray-area diagnoses as follows: diagnoses associated with clinical ambiguity or high rates of potentially preventable hospitalizations and variability in admissions [16,17]. The goal of this work was to identify how the decision to admit drives subsequent revisit, readmission, and mortality among older adult patients with diagnoses that are syndromic (falls and weakness) or lacking a clear standard of practice (syncope, UTI, pneumonia, or cellulitis). This question was difficult to answer since patients with different disposition decisions differ in their clinical severity, complexity, and needs. Without adjusting for these differences, unfair conclusions can be drawn if patients with different disposition decisions are directly compared. We thus used a causal inference methodology for observational data to measure the relationship between disposition decisions and outcomes for these older adult patients [34,35].

Methods

Data

ED visits were analyzed using electronic health records (EHRs) from a large Midwestern academic health system. The dataset used was from a common EHR system, and the analyzed population consisted of encounters with a specific large ED between January 1, 2014 and September 27, 2018. ED visits were included for older adult patients (65 years of age or older) with presentations for falls (n=1581), weakness (n=564), syncope (n=468), UTI (n=456), pneumonia (n=299), and cellulitis (n=223). For all 6 presentations, inclusion and exclusion criteria were specified to capture patients who, based on objective criteria present in the EHR occupied a “gray area” with regards to criteria for admission. In general, the strategy was to include patients who met diagnostic criteria for each presentation but did not have further abnormalities indicating a clear indication for admission present within discrete fields of the EHR. General inclusion criteria were an acuity level (emergency severity index; ESI) of 2 or 3, excluding visits assigned the most and least severe ESI levels; and treated and either admitted to the hospital or discharged to their residence, excluding patients who eloped, left against medical advice, or were transferred to another facility. General exclusion criteria were a missing disposition decision, a visit in the last 45 days of the sample period (as these patients had inadequate follow-up for outcomes), indicators of acute coronary syndromes (troponin levels >0.10 ng/mL), severe vital sign abnormalities at any point during the ED visit (systolic blood pressure <80 mmHg, respiration rate >30 BPM, pulse oximetry <88%, or heart rate >120 BPM), or a specific diagnosis that unambiguously required admission (ie, stroke, myocardial infarction, or femur fractures). In addition, patients presenting to the ED with a combination of cellulitis and a higher temperature than 100.3 F were eliminated from the analysis since these encounters had a high admission probability.

Variables

Seven variables represented baseline characteristics as follows: age, sex, insurance, history of diabetes, history of congestive heart failure, history of hypertension, and Centers for Medicare and Medicaid Services Hierarchical Condition Category (HCC) score [36]. HCC scores predict health care costs and outcomes by assigning risk weights to comorbidities based on ICD codes; higher scores indicate a greater predicted risk and resource use. These baseline characteristics were selected because they are available prior to the ED visit and can capture key clinical variations that could influence a patient’s ED visit and outcomes.

Five variables represented initial observations serving as a proxy for underlying latent health state at the start of the admission process: acuity, temperature, blood pressure, respiratory rate, and heart rate. Acuity was measured using the ESI level, a 5-level triage system where lower ESI levels (eg, 1 or 2) indicate higher acuity, meaning the patient requires more urgent care and significant resources. These proxy variables were included because they are available before the admission process begins and can be considered indirect (or “noisy”) measurements of the patient’s latent health state.

Two variables represented the admission process: treatment time (duration between when a patient is placed in an ED room after triage for treatment and when their admission decision is made) and admission decision. Patients discharged to a Skilled Nursing Facility or Inpatient Rehabilitation Facility were treated as a discharge.

Three primary outcome variables were considered: ED revisits, hospital readmission, and mortality within 30 days of being discharged from either the hospital for admitted patients or the ED for discharged patients. Binary variables for revisits indicated whether an individual returned to the study ED (based on EHR data) within 30 days of discharge. Similarly, binary variables for readmission and mortality indicated whether an individual was readmitted or died within 30 days of discharge,

XSL•FO
RenderX

respectively. Alternative windows (3 d and 9 d instead of 30 d) were also considered as secondary outcomes, since a short window is likely to capture only subsequent events related to the original ED visit, whereas a long window is likely to capture all subsequent events related to the original ED visit. To optimize this trade-off, a 9-day window has previously been recommended [37].

Missing continuous and categorical variables were imputed with the median and most frequent category, respectively. Continuous variables were standardized and these transformed variables were then used for estimation.

Analysis

Admission decisions were evaluated using 3 types of analyses: unadjusted, adjusted, and subgroup. Each type of analysis was repeated for each of the 3 primary outcomes, and for the overall sample and each presentation group. Unadjusted analyses involved estimating the observed (ie, unadjusted) difference in risk of an outcome between admitted patients and discharged patients. Wald 95% CIs were recovered.

Adjusted analyses, constituting our primary analyses, involved estimating these same risk differences (RDs) but adjusted for the unmeasured or latent health state in addition to measured variables (ie, baseline characteristics and proxy variables). These estimates were recovered by implementing a latent-variable approach for evaluating admission decisions (see Cochran et al [34] for details and applied in Alvarez Avendaño et al [35] to chest pain patients). Briefly, this approach involves modeling both the measured variables and a latent "health state" variable, and then fitting this model to data using expectation maximization, from which an estimate of the average treatment effect (ie, average difference in potential outcomes were we to admit vs discharge a patient) and Wald CIs can be recovered. The latent health state is included to account for potential unmeasured confounding between the admission decision and each outcome. More specifically, it accounts for confounding by indication, whereby admission decisions are based on factors not captured in the data. The inclusion of proxy variables, such as acuity and vitals, is crucial as they provide indirect measures of the latent health state, thereby strengthening the model's ability to mitigate confounding by indication.

The model comprised several regression components: logistic regression to model the latent health state as a function of baseline characteristics; logistic and linear regression to model the proxy variables (eg, acuity and vitals) as a function of latent health and baseline factors; threshold regression to model the admission process (decision and timing) as a function of latent health, baseline characteristics, and proxy variables; and linear regression to model each outcome based on latent health. Once model parameters are estimated, we calculate average treatment effects by comparing outcomes for each visit under hypothetical admission and discharge scenarios while holding all other variables at their original value. We then calculate the difference between these hypothetical outcomes and average this difference across all visits to obtain the overall treatment effect. For subgroup analyses, we recover subgroup-specific average treatment effects by averaging these hypothetical outcomes across only those visits with certain baseline characteristics (eg, female patients).

To check the sensitivity of our conclusions to various factors, several additional analyses were conducted. First, we checked sensitivity to 2 key assumptions for the latent-variable approach, which is that potential outcomes are independent of admission decisions with similar latent health needs and baseline characteristics and that events related to different ED visits are independent. Second, we analyzed the secondary outcomes, which use alternative time windows (3 days and 9 days instead of 30 days). Third, we estimated the same RDs as the main analyses, but only adjusted for the measured variables as opposed to both the measured variables and the latent health state. These estimates were recovered using the causal inference methods known as inverse probability weighting (IPW) and g-estimation.

Due to space considerations, sensitivity analyses are detailed in Multimedia Appendix 1. Importantly, estimates from checking violations to our 2 key assumptions were generally consistent in terms of direction and magnitude with the adjusted estimates reported in the main text, and estimates from IPW and g-estimation were similarly consistent with unadjusted estimates, with a few exceptions detailed in Multimedia Appendix 1. In addition, cellulitis had a small sample size (n=223) relative to other diagnostic groups, leading to imprecise estimates. Therefore, results for cellulitis are presented in Multimedia Appendix 1. Finally, subgroup analyses for individual diagnoses and technical details of all our adjusted analyses can also be found in Multimedia Appendix 1.

Significance was considered at an α level of 0.05. Hypothesis tests were 2-tailed Wald tests. Multiple comparisons were not adjusted for, and as such, nominal CIs and *P* values are reported.

Ethical Considerations

This study was reviewed by the UW Minimal Risk Research Institutional Review Board (ID 2024-0106-CP001) and was deemed to meet the federal criteria for exemption.

Results

Baseline Characteristics

Sample characteristics are summarized in Table 1. Patients were predominantly female (2102/3591, 58.5%) with Medicare insurance (1460/1589, 91.9%) and had an average age of 79 years. Patients were diagnosed with falls (1581/3591, 44%), weakness (564/3591, 16%), syncope (468/3591, 13%), UTI (456/3591, 13%), pneumonia (299/3591, 8%), and cellulitis (223/3591, 6%).

XSL•FO
RenderX

Table . Descriptive statistics of sample by complaint.

	All (n=3591)	Falls (n=1581)	Weakness (n=564)	Syncope (n=468)	UTI[a] (n=456)	Pneumonia (n=299)	Cellulitis (n=223)
Age (years), mean (SD)	79.23 (8.93)	80.56 (8.94)	78.66 (8.4)	76.8 (8.73)	79.22 (8.91)	78.9 (9.27)	76.758 (8.36)
Comorbidity (HCC[b]), mean (SD)	1.65 (1.41)	1.49 (1.22)	1.82 (1.46)	1.32 (1.36)	1.879 (1.6)	2.107 (1.57)	1.965 (1.72)
Heart rate (BPM), mean (SD)	78.07 (14.66)	76.41 (13.54)	78.39 (15.06)	73.11 (13.52)	80.79 (14.82)	87.32 (15.74)	81.41 (14.59)
Temperature (°F), mean (SD)	97.61 (1.4)	97.44 (1.69)	97.56 (0.85)	97.34 (0.72)	97.88 (1.18)	98.55 (1.56)	97.73 (0.8)
Blood pressure (mmHg), mean (SD)	74.2 (13.7)	76.23 (13.92)	74.36 (12.9)	71.91 (12.19)	73.49 (14.72)	69.47 (13.57)	71.99 (12.41)
Respiration rate (BPM), mean (SD)	18.07 (3.18)	17.82 (2.92)	17.86 (3.23)	17.72 (3.57)	18.28 (3.07)	19.8 (3.72)	18.41 (2.58)
Treatment time (hours), mean (SD)	1.82 (0.12)	0.18 (0.1)	0.2 (0.12)	0.22 (0.2)	0.16 (0.07)	0.13 (0.06)	0.15 (0.08)
Female, n (%)	2102 (58.5)	994 (62.9)	299 (53)	255 (54.5)	290 (63.6)	149 (49.8)	115 (51.6)
Insurance, n (%)							
Medicaid/Badger Care	15 (0.4)	3 (0.2)	3 (0.5)	5 (1.1)	2 (0.4)	1 (0.3)	1 (0.4)
Medicare	3278 (91.9)	1460 (92.9)	526 (93.6)	415 (90.8)	409 (90.1)	273 (91.3)	195 (87.4)
Commercial/Worker's Compensation	271 (7.6)	109 (6.9)	33 (5.9)	35 (7.7)	42 (9.3)	25 (8.4)	27 (12.1)
Self-pay	3 (0.1)	0 (0)	0 (0)	2 (0.4)	1 (0.2)	0 (0)	0 (0)
Diabetes, n (%)	744 (20.7)	293 (18.5)	138 (24.5)	81 (17.3)	99 (21.7)	72 (24.1)	61 (27.4)
Congestive Heart failure, n (%)	413 (11.5)	173 (10.9)	62 (11)	33 (7.1)	56 (12.3)	53 (17.1)	36 (16.1)
Hypertension, n (%)	2081 (58.0)	923 (58.4)	350 (62.1)	244 (52.1)	255 (55.9)	170 (56.9)	139 (62.3)
Acuity (=2), n (%)	993 (27.7)	425 (26.9)	139 (24.6)	195 (41.7)	103 (22.6)	106 (35.5)	25 (11.2)
Admitted, n (%)	1401 (39.0)	333 (21.1)	252 (44.7)	179 (38.2)	240 (52.6)	255 (85.3)	142 (63.7)
30-day revisits, n (%)	644 (17.9)	293 (18.5)	108 (19.1)	53 (11.3)	105 (23)	40 (13.4)	45 (20.2)
30-day readmission, n (%)	207 (5.8)	50 (3.2)	36 (6.4)	23 (4.9)	43 (9.4)	33 (11)	22 (9.9)
30-day mortality, n (%)	127 (3.5)	50 (3.2)	30 (5.3)	6 (1.3)	11 (2.4)	27 (9)	3 (1.3)

[a]UTI: urinary tract infection.

[b]HCC: hierarchical condition category.

Revisits

Table 2 summarizes the unadjusted RD between admission and discharge for 30-day revisits across all presentations and for individual presentations. Across all presentations, admission carried a significantly lower unadjusted risk than a discharge of 30-day revisits (RD=−5.2%, 95% CI −6.3 to −4.1). Individual presentations yielded unadjusted estimates that generally agreed with that of the entire sample, with 1 exception. For syncope

XSL•FO
RenderX

patients, admission carried a greater unadjusted risk than a discharge of 30-day revisits (RD=2.5%, 95% CI 0.0 to 5.0).

Table . Unadjusted and adjusted estimates (95% CI) of risk differences (RD) for 30-day revisits, comparing admission to discharge (reference). Adjusted estimates account for latent health state and measured variables.

	Unadjusted RD (95% CI)	Adjusted RD (95% CI)
All	−5.2 (−6.3 to −4.1)	−6.4 (−7.8 to −5.1)
Falls	−4.5 (−6.5 to −2.5)	−33.2 (−35.5 to −30.8)
Weakness	−8.8 (−11.6 to −6.0)	4.4 (−2.8 to 11.6)
Syncope	2.5 (0 to 5)	−38.6 (−43.2 to −34.1)
UTI[a]	−10.8 (−14.1 to −7.5)	−56.7 (−61.3 to −52.2)
Pneumonia	−3.0 (−7.6 to 1.7)	23.4 (18.7 to 28.1)

[a]UTI: urinary tract infection.

Table 2 also summarizes the estimated RDs, adjusted for latent health state and measured patient variables, between admission and discharge for 30-day revisits. Across all presentations, admission carried a significantly lower adjusted risk than a discharge of 30-day revisits (RD=−6.4%, 95% CI −7.8 to −5.0). Individual presentations yielded adjusted estimates that generally agreed with that of the entire sample, with the following exceptions. For patients with weakness, admission carried a numerically greater adjusted risk than a discharge of 30-day revisits (RD=4.4%, 95% CI −2.8 to 11.6). For patients with pneumonia, admission carried a significantly greater adjusted risk than a discharge of 30-day revisits (RD=23.4%, 95% CI 18.7 to 28.1).

Readmission

Table 3 summarizes the unadjusted RD between admission and discharge for 30-day readmissions across all diagnoses and for individual diagnoses. Across all diagnoses, admission carried a significantly greater unadjusted risk than a discharge of 30-day readmission (RD =14.8%, 95% CI 14.1 to 15.4). Individual diagnoses yielded unadjusted estimates that agreed in terms of direction, significance, and magnitude (within 3%) with that of the entire sample.

Table . Unadjusted and adjusted estimates (95% CI) of risk differences (RDs), in percentage points, for 30-day readmissions, comparing admission to discharge (reference). Adjusted estimates account for latent health state and measured variables.

	Unadjusted RD (95% CI)	Adjusted RD (95% CI)
All	14.8 (14.1 to 15.4)	5.8 (5.0 to 6.5)
Falls	15.0 (14.1 to 15.9)	3.2 (2.3 to 4.0)
Weakness	14.3 (12.6 to 16.0)	61.6 (57.7 to 65.5)
Syncope	12.8 (11.1 to 14.6)	4.9 (3.0 to 6.9)
UTI[a]	17.9 (15.6 to 20.2)	9.4 (6.8 to 12.1)
Pneumonia	12.9 (8.7 to 17.2)	25.8 (20.9 to 30.6)

[a]UTI: urinary tract infection.

Table 3 also summarizes estimated RDs, adjusted for latent health state and measured patient variables, between admission and discharge for 30-day readmissions. Across all diagnoses, admission carried a significantly greater adjusted risk than a discharge of 30-day readmission (RD=5.8%, 95% CI 5.0 to 6.5). Individual diagnoses yielded adjusted estimates in the same direction as that of the entire sample, but the magnitude was notably larger in a few cases. For patients with weakness, admission carried a greater adjusted risk than a discharge 30-day readmission (RD=61.6%, 95% CI 57.7 to 65.5). For patients with pneumonia, admission carried a significantly greater adjusted risk than a discharge of 30-day readmission (RD=25.8%, 95% CI [20.9, 30.6]).

Mortality

Table 4 summarizes the unadjusted RD between admission and discharge for 30-day mortality across all diagnoses and for individual diagnoses. Across all diagnoses, admission carried a significantly greater unadjusted risk than a discharge of 30-day mortality (RD=3.8%, 95% CI 3.3 to 4.3). Individual diagnoses yielded unadjusted estimates that agreed in terms of direction, significance, and magnitude (within 3%) with those of the entire sample.

XSL•FO
RenderX

Table . Unadjusted and adjusted estimates (95% CI) of risk differences (RDs), in percentage points, for 30-day mortality, comparing admission to discharge (reference). Adjusted estimates account for latent health state and measured patient variables.

	Unadjusted RD (95% CI)	Adjusted RD (95% CI)
All	3.8 (3.3 to 4.3)	1.0 (0.4 to 1.6)
Falls	2.8 (1.9 to 3.7)	−0.9 (−1.8 to 0.0)
Weakness	6.9 (5.3 to 8.5)	58.1 (54.1 to 62.0)
Syncope	2.4 (1.6 to 3.3)	−21.9 (−53.6 to 9.7)
UTI[a]	1.9 (0.7 to 3.1)	1.1 (−0.3 to 2.5)
Pneumonia	5.3 (1.4 to 9.2)	7.7 (4.4 to 11.0)

[a]UTI: urinary tract infection.

Table 4 also summarizes estimated RDs, adjusted for latent health state and measured patient variables, between admission and discharge for 30-day mortality. Across all diagnoses, admission carried a significantly greater adjusted risk than a discharge of 30-day mortality (RD=1.0%, 95% CI 0.4 to 1.6). Individual diagnoses yielded adjusted estimates differing from that of the entire sample in several meaningful ways.

Reporting from high to low, admission for patients with weakness carried a significantly greater adjusted risk than a discharge of 30-day mortality (RD=58.1%, 95% CI 54.1 to 62.0). For patients with pneumonia, admission also carried a significantly adjusted risk than a discharge of 30-day mortality (RD=7.7%, 95% CI 4.4 to 11.0). For patients with UTI, admission carried a numerically greater adjusted risk than a discharge of 30-day mortality (RD=1.1%, 95% CI −0.3 to 2.5). For patients with syncope, admission carried a numerically lower adjusted risk of 30-day mortality (RD=−21.9%, 95% CI −53.6 to 9.7). For patients with falls, admission carried a lower adjusted risk than a discharge of 30-day mortality (RD=−0.9%, 95% CI −1.8 to 0.0).

Subgroup Analyses

Table 5 summarizes subgroup-specific estimates of RDs, adjusted for latent health state and measured patient variables, between admission and discharge for 30-day outcomes. All subgroup estimates had the same direction as that of the entire sample: negative for 30-day revisits and positive for 30-day readmission and mortality. For all 30-day outcomes, 5 subgroups had numerically greater adjusted estimates than that of the entire sample. These subgroups (in decreasing order of prominence) were individuals with congenital heart disease, individuals with diabetes, individuals with HCC scores greater than the mean score of 1.65, individuals greater than the mean age of 79.2 years, and individuals with hypertension. Of note, all these subgroups tend to have worse baseline health than their counterpart. The subgroup with the numerically lowest adjusted estimate for all 30-day outcomes was individuals with insurance other than Medicare.

XSL•FO
RenderX

Table . Subgroup-specific adjusted estimates (95% CI) of risk differences, in percentage points, for 30-day outcomes, comparing admission to discharge (reference). Adjusted estimates account for latent health state and measured patient variables.

Variable and subgroup	Revisit	Readmission	Mortality
Age (years)			
65 - 79.2	−7.2 (−8.7 to −5.6)	5.4 (4.7 to 6.2)	0.9 (0.2 to 1.5)
>79.2	−5.6 (−7.0 to −4.1)	6.1 (5.3 to 7.0)	1.2 (0.6 to 1.9)
Sex			
Male	−6.5 (−8.0 to −4.9)	5.8 (5.0 to 6.6)	1.0 (0.4 to 1.7)
Female	−6.4 (−7.9 to −5.0)	5.8 (5.0 to 6.6)	1.0 (0.4 to 1.7)
Insurance			
Medicare	−6.2 (−7.6 to −4.9)	5.8 (5.1 to 6.6)	1.1 (0.5 to 1.7)
Other	−8.5 (−10.8 to −6.3)	4.9 (3.9 to 5.9)	0.6 (−0.2 to 1.3)
Diabetes			
No	−6.8 (−8.2 to −5.4)	5.6 (4.9 to 6.4)	0.9 (0.3 to 1.6)
Yes	−5.0 (−6.7 to −3.2)	6.4 (5.4 to 7.4)	1.4 (0.7 to 2.1)
Congenital heart failure			
No	−7.0 (−8.4 to −5.6)	5.5 (4.8 to 6.2)	0.9 (0.3 to 1.5)
Yes	−1.8 (−3.9 to 0.3)	7.8 (6.6 to 9.0)	2.1 (1.3 to 2.9)
Hypertension			
No	−6.9 (−8.4 to −5.4)	5.6 (4.8 to 6.3)	0.9 (0.3 to 1.5)
Yes	−6.1 (−7.5 to −4.6)	5.9 (5.1 to 6.7)	1.1 (0.5 to 1.7)
Hierarchical condition category			
≤1.65	−7.4 (−8.8 to −6.0)	5.3 (4.6 to 6.1)	0.8 (0.2 to 1.4)
>1.65	−4.7 (−6.2 to −3.2)	6.5 (5.6 to 7.4)	1.4 (0.8 to 2.1)

Discussion

Principal Findings

Our goal was to investigate the impact of admitting older adult patients with 6 common but variably managed ED presentations: falls, weakness, syncope, UTI, pneumonia, and cellulitis. In particular, these conditions show considerable variability in admission practices, and understanding their specific outcomes can help clinicians better determine when admission is beneficial or potentially harmful. Using causal inference methods, we analyzed EHR data to compare outcomes such as ED revisits, hospital readmissions, and mortality, measuring these outcomes at 3, 9, and 30 days after discharge. By focusing on this “gray area” where no standard practice exists, we wanted to provide insights that support more consistent and evidence-based admission decisions for these presentations.

Our primary finding is that admission was associated with a significant decrease in the risk of an ED revisit, but with a significant increase in the risk of hospital readmission and mortality. The association with an increased readmission risk was consistent for all individual diagnoses, for all time frames, and for all estimation approaches (unadjusted, IPW, g-estimation, and latent variable as the primary analysis). The associations with a decreased revisit risk and with an increased mortality risk were generally consistent for individual diagnoses, all time frames, and all approaches, with a few exceptions discussed below. These results would suggest that while admission may help prevent immediate returns to the ED, it could lead to worse long-term outcomes for certain patients. This balance of risks highlights the importance of carefully weighing the short-term benefits of admission against the potential for readmission and mortality, especially in older adults with these conditions.

In our primary analysis, admission was associated with an increased revisit risk for patients with weakness and pneumonia, along with a notably higher readmission and mortality risk for these conditions compared with others. This suggests that patients with weakness or pneumonia may be especially vulnerable to poor outcomes from admission. In contrast, admission was associated with a numerically lower 30-day mortality risk for patients with falls or syncope, indicating that these individuals may benefit more from being admitted. These findings can help guide clinicians in prioritizing which patients may require closer monitoring, alternative care options, or potential admission.

Our subgroup analysis identified patient factors that may influence outcomes from hospital admission. Overall, admission appears less favorable for older individuals (>79 years) and those with comorbidities (ie, congenital heart disease, diabetes, high HCC scores, or hypertension). Important factors for

XSL•FO
RenderX

admission, however, may vary by specific presentation, as shown in our subgroup analysis in Multimedia Appendix 1. Despite these differences, we found consistent trends across subgroups, suggesting that standardized admission recommendations could potentially be based on the type of presentation alone, rather than on patient-specific factors like age, sex, insurance, or comorbidity.

For comparison, another study examined the impact of admitting older adult patients presenting to the ED with chest pain, another presentation with variable admission practices, and found that admission had a lower readmission risk but an increased mortality risk [35]. This study similarly estimated an increased mortality risk, yet it also found an increase in readmission risk for almost all presentations. The association of admission with both increased and decreased risks for revisits, readmission, and mortality, underscores why standardized care pathways are lacking for these presentations [21]. While our results suggest that, in general, discharge may be associated with lower readmission and mortality risks, discharging an older adult can still carry significant risks. These findings support the need for safe alternatives to hospitalization, such as outpatient care coordination or at-home ambulatory providers [38]. Such alternatives have been shown to effectively reduce hospital readmission rates on older adult patients discharged from the ED [39,40].

Limitations

A primary limitation of this study is the potential for unmeasured confounding, especially confounding by indication, where treatment decisions are influenced by clinical indicators not fully captured in the EHR. This limitation is significant because it directly impacts our ability to interpret the findings causally. Although we used multiple methods, including a latent variable approach to account for unobserved health status [34], no method can fully guarantee that all confounding is addressed. Consequently, our findings remain subject to potential bias if residual confounding exists. This is particularly relevant for interpreting the effect of admission on outcomes, as patients who are “sicker” based on unmeasured factors are more likely to be admitted, and these same factors could drive poor outcomes. While the latent variable approach yielded estimates generally more favorable to admission—suggesting some adjustment for underlying severity—and results were consistent across different time frames and presentations, caution is warranted in interpreting these effects as causal. This limitation highlights the need for further research to strengthen the causal inference of admission effects on outcomes.

In addition, this was a retrospective, observational study conducted using EHR data from an academic ED in the US Midwest, thereby limiting the generalizability of our findings. Moreover, the resulting patient population for specific presentations after applying our filters was relatively small, leading to nonsignificant estimates. Replicating this analysis in different EDs with larger samples is needed.

Conclusions

This study is one of the few to quantify the risks and benefits associated with admission decisions for ED presentations associated with unclear need for hospitalization. Existing studies have mainly focused on identifying factors that influence hospital admission of noncritically ill patients, often through physician surveys [41] or identifying risk patterns [33,42,43]. This study is intended to quantify the trade-offs of admission at a population level rather than directly influencing bedside decisions. Our findings indicate that admission is not always the “safe” choice, underscoring the need for further research to establish specific criteria that can better guide admission decisions for older adults in the ED.

Acknowledgments

This research was supported by the Clinical and Translational Science Award (CTSA) program, through the NIH National Center for Advancing Translational Sciences (NCATS) (GZC, grant numbers UL1TR002373 and KL2TR002374); and by the University of Wisconsin-Madison School of Medicine and Public Health's Wisconsin Partnership Program (WPP) (VOC, grant number 4358). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or WPP. The authors also wish to thank the Health Innovation Program at the University of Wisconsin-Madison School of Medicine and Public Health for their technical expertise and support in preparing the final curated dataset.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Accompanying appendix reporting (1) IPW and g-estimation results, (2) description of the latent variable approach, (3) 3-day outcomes results, (4) sensitivity analyses, (5) subgroup analyses for individual diagnoses, (6) analysis for cellulitis diagnosis, and (7) analysis for syndromic and no clear standard practice diagnoses.
[DOCX File, 430 KB - aging_v8i1e55929_app1.docx]

References

1. Pitts SR, Carrier ER, Rich EC, Kellermann AL. Where Americans get acute care: increasingly, it’s not at their doctor’s office. Health Aff (Millwood) 2010 Sep;29(9):1620-1629. [doi: 10.1377/hlthaff.2009.1026]

XSL•FO
RenderX

2. Kellermann AL, Martinez R. The ER, 50 years on. N Engl J Med 2011 Jun 16;364(24):2278-2279. [doi: 10.1056/NEJMp1101544]
3. Tang N, Stein J, Hsia RY, Maselli JH, Gonzales R. Trends and characteristics of US emergency department visits, 1997-2007. JAMA 2010 Aug 11;304(6):664-670. [doi: 10.1001/jama.2010.1112] [Medline: 20699458]
4. Morganti KG, Bauhoff S, Blanchard JC, et al. The Evolving Role of Emergency Departments in the United States: Rand health quarterly The RAND Corporation; 2013, Vol. 3.
5. Sabbatini AK, Nallamothu BK, Kocher KE. Reducing variation in hospital admissions from the emergency department for low-mortality conditions may produce savings. Health Aff (Millwood) 2014 Sep;33(9):1655-1663. [doi: 10.1377/hlthaff.2013.1318]
6. Galarraga JE, Pines JM. Costs of ED episodes of care in the United States. Am J Emerg Med 2016 Mar;34(3):357-365. [doi: 10.1016/j.ajem.2015.06.001]
7. Aminzadeh F, Dalziel WB. Older adults in the emergency department: a systematic review of patterns of use, adverse outcomes, and effectiveness of interventions. Ann Emerg Med 2002 Mar;39(3):238-247. [doi: 10.1067/mem.2002.121523]
8. Lowenstein SR, Crescenzi CA, Kern DC, Steel K. Care of the elderly in the emergency department. Ann Emerg Med 1986 May;15(5):528-535. [doi: 10.1016/S0196-0644(86)80987-8]
9. Hedges JR, Singal BM, Rousseau EW, et al. Geriatric patient emergency visits part II: perceptions of visits by geriatric and younger patients. Ann Emerg Med 1992 Jul;21(7):808-813. [doi: 10.1016/S0196-0644(05)81026-1]
10. Denman SJ, Ettinger WH, Zarkin BA, Coon PJ, Casani JA. Short‐term outcomes of elderly patients discharged from an emergency department. J Am Geriatr Soc 1989 Oct;37(10):937-943. [doi: 10.1111/j.1532-5415.1989.tb07278.x]
11. Caplan GA, Brown A, Croker WD, Doolan J. Risk of admission within 4 weeks of discharge of elderly patients from the emergency department—the DEED study. Age Ageing 1998;27(6):697-702. [doi: 10.1093/ageing/27.6.697]
12. Rowland K, Maitra AK, Richardson DA, Hudson K, Woodhouse KW. The discharge of elderly patients from an accident and emergency department: functional changes and risk of readmission. Age Ageing 1990;19(6):415-418. [doi: 10.1093/ageing/19.6.415]
13. Martin-Gill C, Reiser RC. Risk factors for 72-hour admission to the ED. Am J Emerg Med 2004 Oct;22(6):448-453. [doi: 10.1016/j.ajem.2004.07.023]
14. Gerson LW, Rousseau EW, Hogan TM, Bernstein E, Kalbfleisch N. Multicenter study of case finding in elderly emergency department patients. Acad Emerg Med 1995 Aug;2(8):729-734. [doi: 10.1111/j.1553-2712.1995.tb03626.x]
15. Stuck AR, Crowley C, Killeen J, Castillo EM. National survey of emergency physicians concerning home-based care options as alternatives to emergency department–based hospital admissions. J Emerg Med 2017 Nov;53(5):623-628. [doi: 10.1016/j.jemermed.2017.05.036]
16. Kozak LJ, Hall MJ, Owings MF. Trends in avoidable hospitalizations, 1980–1998. Health Aff (Millwood) 2001 Mar;20(2):225-232. [doi: 10.1377/hlthaff.20.2.225]
17. Nyweide DJ, Anthony DL, Bynum JPW, et al. Continuity of care and the risk of preventable hospitalization in older adults. JAMA Intern Med 2013 Nov 11;173(20):1879-1885. [doi: 10.1001/jamainternmed.2013.10059] [Medline: 24043127]
18. Hussain M, Oppenheim BA, O'Neill P, Trembath C, Morris J, Horan MA. Prospective survey of the incidence, risk factors and outcome of hospital-acquired infections in the elderly. J Hosp Infect 1996 Feb;32(2):117-126. [doi: 10.1016/S0195-6701(96)90053-5]
19. Taylor ME, Oppenheim BA. Hospital-acquired infection in elderly patients. J Hosp Infect 1998 Apr;38(4):245-260. [doi: 10.1016/S0195-6701(98)90074-3]
20. Hartholt KA, van der Velde N, Looman CWN, et al. Trends in fall-related hospital admissions in older persons in the Netherlands. Arch Intern Med 2010 May 24;170(10):905-911. [doi: 10.1001/archinternmed.2010.106] [Medline: 20498419]
21. Eagles D, Stiell IG, Clement CM, et al. International survey of emergency physicians' priorities for clinical decision rules. Acad Emerg Med 2008 Feb;15(2):177-182. [doi: 10.1111/j.1553-2712.2008.00035.x]
22. Daccarett M, Jetter TL, Wasmund SL, Brignole M, Hamdan MH. Syncope in the emergency department: comparison of standardized admission criteria with clinical practice. Europace 2011 Nov 1;13(11):1632-1638. [doi: 10.1093/europace/eur201]
23. Kocher KE, Meurer WJ, Fazel R, Scott PA, Krumholz HM, Nallamothu BK. National trends in use of computed tomography in the emergency department. Ann Emerg Med 2011 Nov;58(5):452-462. [doi: 10.1016/j.annemergmed.2011.05.020]
24. Kocher KE. Achieving the holy grail of emergency department evaluation for chest pain. Circ Cardiovasc Qual Outcomes 2017 Oct;10(10). [doi: 10.1161/CIRCOUTCOMES.117.004026]
25. Sammon JD, Sharma P, Rahbar H, et al. Predictors of admission in patients presenting to the emergency department with urinary tract infection. World J Urol 2014 Jun;32(3):813-819. [doi: 10.1007/s00345-013-1167-3]
26. Venkatesh AK, Chou SC, Li SX, et al. Association between insurance status and access to hospital care in emergency department disposition. JAMA Intern Med 2019 May 1;179(5):686-693. [doi: 10.1001/jamainternmed.2019.0037] [Medline: 30933243]
27. Volz KA, Canham L, Kaplan E, Sanchez LD, Shapiro NI, Grossman SA. Identifying patients with cellulitis who are likely to require inpatient admission after a stay in an ED observation unit. Am J Emerg Med 2013 Feb;31(2):360-364. [doi: 10.1016/j.ajem.2012.09.005]

XSL•FO
RenderX

28. Calder LA, Forster AJ, Stiell IG, et al. Mapping out the emergency department disposition decision for high-acuity patients. Ann Emerg Med 2012 Nov;60(5):567-576. [doi: 10.1016/j.annemergmed.2012.04.013]
29. Gorski JK, Batt RJ, Otles E, Shah MN, Hamedani AG, Patterson BW. The impact of emergency department census on the decision to admit. Acad Emerg Med 2017 Jan;24(1):13-21. [doi: 10.1111/acem.13103] [Medline: 27641060]
30. Nugus P, Carroll K, Hewett DG, Short A, Forero R, Braithwaite J. Integrated care in the emergency department: a complex adaptive systems perspective. Soc Sci Med 2010 Dec;71(11):1997-2004. [doi: 10.1016/j.socscimed.2010.08.013]
31. Florin TA, French B, Zorc JJ, Alpern ER, Shah SS. Variation in emergency department diagnostic testing and disposition outcomes in pneumonia. Pediatrics 2013 Aug;132(2):237-244. [doi: 10.1542/peds.2013-0179] [Medline: 23878049]
32. Studnicki J, Platonova EA, Fisher JW. Hospital-level variation in the percentage of admissions originating in the emergency department. Am J Emerg Med 2012 Oct;30(8):1441-1446. [doi: 10.1016/j.ajem.2011.11.008]
33. Pines JM, Mutter RL, Zocchi MS. Variation in emergency department admission rates across the United States. Med Care Res Rev 2013 Apr;70(2):218-231. [doi: 10.1177/1077558712470565]
34. Cochran AL, Rathouz PJ, Kocher KE, Zayas‐Cabán G. A latent variable approach to potential outcomes for emergency department admission decisions. Stat Med 2019 Sep 10;38(20):3911-3935. [doi: 10.1002/sim.8210]
35. Alvarez Avendaño SA, Cochran AL, Patterson B, Shah M, Smith M, Zayas-Cabán G. The average effect of emergency department admission on readmission and mortality for older adults with chest pain. Med Care 2020;58(10):881-888. [doi: 10.1097/MLR.0000000000001375]
36. Pope GC, Ellis RP, Ash AS, et al. Diagnostic Cost Group Hierarchical Condition Category Models for Medicare Risk Adjustment: Health Economics Research; 2000.
37. Rising KL, Victor TW, Hollander JE, Carr BG. Patient returns to the emergency department: the time-to-return curve. Acad Emerg Med 2014 Aug;21(8):864-871. [doi: 10.1111/acem.12442]
38. Kilaru AS, Resnick D, Flynn D, et al. Practical Alternative to Hospitalization for Emergency Department Patients (PATH): A Feasibility Study: Healthcare Elsevier; 2021:100545. [doi: 10.1016/j.hjdsi.2021.100545]
39. Guttman A, Afilalo M, Guttman R, et al. An emergency department–based nurse discharge coordinator for elder patients: does it make a difference? Acad Emerg Med 2004 Dec;11(12):1318-1327. [doi: 10.1197/j.aem.2004.07.006]
40. Caplan GA, Williams AJ, Daly B, Abraham K. A randomized, controlled trial of comprehensive geriatric assessment and multidisciplinary intervention after discharge of elderly from the emergency department—the DEED II study. J Am Geriatr Soc 2004 Sep;52(9):1417-1423. [doi: 10.1111/j.1532-5415.2004.52401.x]
41. Lewis Hunter AE, Spatz ES, Bernstein SL, Rosenthal MS. Factors influencing hospital admission of non-critically Ill patients presenting to the emergency department: a cross-sectional study. J Gen Intern Med 2016 Jan;31(1):37-44. [doi: 10.1007/s11606-015-3438-8] [Medline: 26084975]
42. Oster A, Bindman AB. Emergency department visits for ambulatory care sensitive conditions: insights into preventable hospitalizations. Med Care 2003;41(1):198-207. [doi: 10.1097/01.MLR.0000045021.70297.9F]
43. Pappas G, Hadden WC, Kozak LJ, Fisher GF. Potentially avoidable hospitalizations: inequalities in rates between US socioeconomic groups. Am J Public Health 1997 May;87(5):811-816. [doi: 10.2105/AJPH.87.5.811]

Abbreviations

ED: emergency department
EHR: electronic health record
ESI: emergency severity index
HCC: hierarchical condition category
IPW: inverse probability weighting
RD: risk difference
UTI: urinary tract infection

Edited by Y Jiang; submitted 29.12.23; peer-reviewed by A Hassan, K Musikatavorn, V Winter; revised version received 04.11.24; accepted 05.12.24; published 06.02.25.

Please cite as:
Alvarez Avendano SA, Cochran A, Odeh Couvertier V, Patterson B, Shah M, Zayas-Caban G
Revisits, Readmission, and Mortality From Emergency Department Admissions for Older Adults With Vague Presentations: Longitudinal Observational Study
JMIR Aging 2025;8:e55929
URL: https://aging.jmir.org/2025/1/e55929
doi:10.2196/55929

XSL•FO
RenderX

© Sebastian Alejandro Alvarez Avendano, Amy Cochran, Valerie Odeh Couvertier, Brian Patterson, Manish Shah, Gabriel Zayas-Caban. Originally published in JMIR Aging (https://aging.jmir.org), 6.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Passive Remote Monitoring Technologies' Influence on Home Care Clients' Ability to Stay Home: Multiprovincial Randomized Controlled Trial

Lorie Donelle[1], PhD; Bradley Hiebert[2,3], PhD; Grace Warner[4], PhD; Michael Reid[5], MSc, MMM; Jennifer Reid[6], MSc; Salimah Shariff[2,6], PhD; Emily Richard[7], PhD; Sandra Regan[2], PhD; Lori Weeks[8], PhD; Kathleen Ledoux[2], PhD

[1]College of Nursing, University of South Carolina, Columbia, SC, United States
[2]Arthur Labatt Family School of Nursing, Western University, London, ON, Canada
[3]St Joseph's Health Care London, London, ON, Canada
[4]School of Occupational Therapy, Dalhousie University, Halifax, NS, Canada
[5]Department of Health and Epidemiology, Dalhousie University, Halifax, NS, Canada
[6]ICES Western, London Health Sciences Centre Research Institute, London, ON, Canada
[7]Faculty of Nursing, University of New Brunswick, Moncton, NB, Canada
[8]School of Health Administration, Dalhousie University, Halifax, NS, Canada

Corresponding Author:
Lorie Donelle, PhD
College of Nursing
University of South Carolina
1601 Greene St
Columbia, SC, 29208
United States
Phone: 1 803 777 6528
Email: ldonelle@mailbox.sc.edu

Abstract

Background: Researchers in Nova Scotia and Ontario, Canada, implemented a passive remote monitoring (PRM) model of home care unique to their health system contexts. Each PRM model integrated tailored PRM devices (eg, motion sensors, cameras, and door alarms) into home care patients' residences with the aim of linking patients, family and friend caregivers, and health care providers to support older adults' aging in place.

Objective: The purpose of this study was to examine the use of PRM technologies in the home to support older adults' safe aging in place and avoidance or delay of higher levels of care.

Methods: This multiprovincial pragmatic randomized controlled trial examined how PRM technologies support older adults to safely remain in their home and avoid or delay admission to higher levels of care. Pairs of home care patients and their family and friend caregivers were recruited in Ontario and Nova Scotia. Participant pairs were randomly assigned to one of two conditions: (1) standard home care (ie, control) or (2) standard home care plus study-provided PRM (ie, intervention). Participants provided their provincial health insurance numbers to link with provincial health administrative databases and identify if patients were admitted to higher levels of care after 1 year. Cox proportional hazards models were used to evaluate the primary outcome in each province.

Results: In total, 313 patient-caregiver pairs were recruited: 174 pairs in Ontario (intervention: n=60; control: n=114) and 139 pairs in Nova Scotia (intervention: n=45; control: n=94). Results indicate PRM was associated with a nonsignificant 30% reduction in risk of patients being admitted to higher levels of care in Ontario (hazard ratio 0.7, 95% CI 0.3-1.4) and no reduction in risk in Nova Scotia (hazard ratio 1.1, 95% CI 0.3-3.7). Adjusting for patient sex had no impact on model estimates for either province.

Conclusions: Limitations related, in part, to the impact of the COVID-19 pandemic may have contributed to the effectiveness of the intervention. While our study did not yield statistically significant results (P=.30 and P=.90) regarding the effectiveness of the PRM model in prolonging home stays, the observed trends suggest that technology-assisted aging in place may be a valuable

XSL•FO
RenderX

goal for older adults. Further study is required to understand if longer follow-up time allows more effects of PRM on patients' avoidance of higher levels of care to be detected.

Trial Registration: ISRCTN ISRCTN79884651; https://www.isrctn.com/ISRCTN79884651

International Registered Report Identifier (IRRID): RR2-10.2196/15027

(JMIR Aging 2025;8:e69107) doi:10.2196/69107

KEYWORDS

remote monitoring technology; home care; health service use; aging in place

Introduction

Creative and effective ways to meet the health care needs of an aging population, often related to the high prevalence of chronic disease, are complicated by a strained health care work force, competing economic priorities, and evolving health information. Yet, these challenges create opportunities to reimagine how health care systems can provide Canadians with the right care, at the right time, in the right place. Although older adults may experience significant chronic disease that may challenge their ability to live independently, most older adults want to remain at home [1,2]. This sentiment was underscored by the findings of a recent survey where 85% of Canadian adults and 96% of Canadians aged ≥65 years reported that they would do everything they can to avoid institutionalization in a long-term care (LTC) setting [2,3]. Home care services that are appropriately managed and integrated into the health care system support older adults to successfully age at home, can improve the health and well-being of older adults and their families, and reduce costs of care in hospitals and LTC facilities [2,4]. However, in Canada, beyond the challenges previously noted, the provision of publicly funded home care services is limited and home care providers are increasingly unable to meet existing demands for home care service needs [2,4-6]. In fact, approximately 75% of home care support for older Canadians is provided *supposedly for free* by an unpaid family and friend caregiver [2,7], creating enhanced personal risk of physical health problems, stress, burnout, and depression [2,8].

A noted challenge of caring for older adults at home is providing suitable home care services and ensuring that older adults can safely follow established treatment plans. Common adverse events among older adults at home (eg, medication administration errors and falls) have been associated with increased use of health services, disability, and death [9,10]. Furthermore, many older adults have complex health needs, and the home care service gaps for older adults with complex care needs are further magnified by limited health human resources, access to partial home care services, a lack of direct support for or involvement of caregivers, and a lack of innovative strategies to expand home care [2,5]. These challenges highlight the urgency to develop more effective and respectful home care options for older adults.

A plausible intervention to address these home care challenges is the implementation of active and passive remote monitoring (PRM) technologies to support older adults to remain in their homes [11]. Active monitoring applications require individual participation, such as pushing a button on a wearable device (eg, pendant or bracelet), while passive monitoring technologies, such as motion sensors, do not require any action by the individual for the system to work. Remote monitoring is useful for tracking the behaviors of older adults with cognitive decline (eg, forgetting to take medications) and intervening quickly in the case of adverse events, such as falls. It also increases individuals' confidence in their ability to care for themselves and live independently at home [11,12]. These technologies can also benefit older adults living at home and their caregivers by increasing communication and collaboration among people in their circle of care and by enabling big data analytics that can contribute to improving health care delivery practices [11,12]. While there is emerging evidence around the value of monitoring technologies to support older adults at home, there is limited insight into the outcomes of PRM, the preferences of older adults and their caregivers, and the costs and benefits associated with the wide variety of remote monitoring technologies available.

The purpose of this study was to examine the use of PRM technologies in the home to support older adults with complex care needs to safely remain in their home and avoid or delay admission to higher levels of care, such as long-term hospitalization and LTC. We evaluated a technology-enabled remote monitoring model of home care conceptually informed by the principles of person-centered care, family as a client, supported self-management, and stakeholder collaboration, to address gaps in home care for older adults. Supportive self-management speaks to the extended health care responsibilities taken up by older adults and their caregivers that create care work. Supportive self-management within home care would include both clinical (eg, medication administration, symptom monitoring, and care coordination) and nonclinical (eg, homemaking, meal preparation, supportive housing, daily check-ins, transportation, and a 24/7 helpline) funded services [13,14]. In this model of care, family was conceptualized as the patient or client receiving home care services *and* their caregiver. Typically, caregivers play an integral role as part of the care team to support older adults in the home; health care assessments remain largely focused on the needs of the older adult with relatively marginal inquiry into caregiver health and well-being [13,14]. The primary focus of this study was to investigate the impact of the PRM model of home care on older adult home care recipients with a secondary focus on assessing the impact of PRM on caregivers' health and well-being. Reported here are our findings as to whether PRM along with usual home care supported older adults with complex care needs to safely remain in their home longer and delay or avoid admission to higher levels of care (ie, hospitalization and LTC)

XSL•FO
RenderX

when compared with older adults receiving usual home care alone.

Methods

Overview

This study was informed by the findings of previous pilot studies conducted in the Canadian provinces of Alberta, British Columbia, New Brunswick, and Nova Scotia [15,16]. The primary outcome of this study was to support older adults with complex care needs to safely remain in their home and avoid or delay admission to higher levels of care, such as long-term hospitalization and LTC. This outcome was operationalized among complex care older adult home care recipients as the occurrence of and time to the following events: terminal admission to the hospital awaiting admission to LTC and direct admission to LTC. The secondary outcomes were an assessment of health service use and mortality rates of home care recipients within 1 year of trial enrollment. A complete description of the study's methodology is reported elsewhere [17] and is summarized in the subsequent sections.

Study Design

This study was an unblinded pragmatic randomized controlled trial (PRCT) [18,19] with 2 parallel arms in two Canadian provincial sites: (1) Central, Western, and Northern Zones in Nova Scotia and (2) the South West Region in Ontario. PRCTs are well-suited to supporting decision-making around complex interventions tested in the *real world* in comparison to usual forms of care [20,21]. As PRCTs are meant to capture real-world situations and the experiences of individuals, broad inclusion criteria developed in collaboration with regional health care partners were purposely selected for this study (see Textbox 1 for the full criteria). For this study, participants were recruited in home care client-caregiver dyads because the PRM requires that older adults have a family caregiver who is willing to receive system notifications. This PRCT is registered with the ISRCTN (registration 79884651).

Textbox 1. Participant eligibility.

Participant group and inclusion criteria (eligible if all criteria are met)

- Home care clients
- Adult who is aged ≥65 years
- Requires home care and is at risk for higher levels of care as determined by the home care provider who makes these decisions
- Has a caregiver who is willing and able to receive the remote monitoring sensor notifications using a cell phone or a regular phone (ie, landline)
- Able to read and write in English or French
- Has the decisional capacity to consent or have a substitute decision-maker to consent for participation
- Caregivers
- Adult who is aged ≥18 years
- Is a caregiver to an adult who is aged ≥65 years that requires home care services and is assessed by home care service providers to be at risk for higher levels of care
- Can be a spouse, partner, child, sibling, other family relation, or friend who helps care for the patient at home
- Willing and able to receive the remote monitoring sensor notifications using a cell phone or a regular phone (ie, landline)
- Able to read and write in either English or French
- Has the decisional capacity to consent for participation

Recruitment

The participant sample size was determined using information from one health region regarding the time to higher levels of care for older adults who previously required complex home care. The sample size calculation was based on the following criteria: total institutionalization proportion among controls being 0.41, and the proportion for the experimental participants being 0.27 (a 34% reduction compared with controls); a 10% attrition rate due to dropout or loss to follow-up; a power of 80%; and a statistical significance level of α=.05. This resulted in a total target sample size of 160 home care clients (and paired caregivers) for the intervention group and 320 home care clients (and paired caregivers) in the control group for a total study sample size of 480 home care clients and 480 caregivers across the Ontario and Nova Scotia study sites. The required sample was estimated based on a prospective test of independence (continuity-corrected chi-square statistic test) between the experimental and control groups using 2 controls per intervention case to increase the study's power due to the cost of the intervention [22].

Participant recruitment occurred from April 2017 through January 2020 and was supported by the respective provincial regional authorities. Assessors and care coordinators employed at the health service provider that assesses and facilitates home care services in each region identified potential study participants (ie, clients) based on the study inclusion criteria (Textbox 1). The home care clients may have been on the assessors and care coordinators' case load for up to 6 months before being contacted about the study. Care coordinators provided information about the study to potential study participants and their family and friend caregivers and requested permission to provide their names and contact information to

XSL•FO
RenderX

the study research coordinator. Research staff contacted potential participants who agreed to have their contact information disclosed to arrange an appointment (in person or by telephone) to provide more information about the study. If both the home care client and family and friend caregiver met the eligibility criteria and verbally agreed to participate, the research staff scheduled an in-home meeting with the pair to obtain written consent and complete the baseline data collection forms (detailed in the *Data Collection* section).

Randomization

With participants' consent, baseline data were collected before random allocation into control or intervention groups. A block randomization process was used after every 3 pairs of participants were recruited [23] to randomly assign pairs in a 2:1 ratio to usual home care (ie, control group) or to usual home care plus study-provided PRM (ie, intervention group). This process was carried out independently in each province by members of the research team. Allocation bias was addressed through allocation concealment, and neither the home care case coordinator nor the researchers knew which group the participant would be assigned before baseline data collection. Following randomization research staff contacted each participant pair to inform them of their study arm. For intervention participants, research staff also forwarded their contact information to the technology provider to schedule an installation date.

Control Group

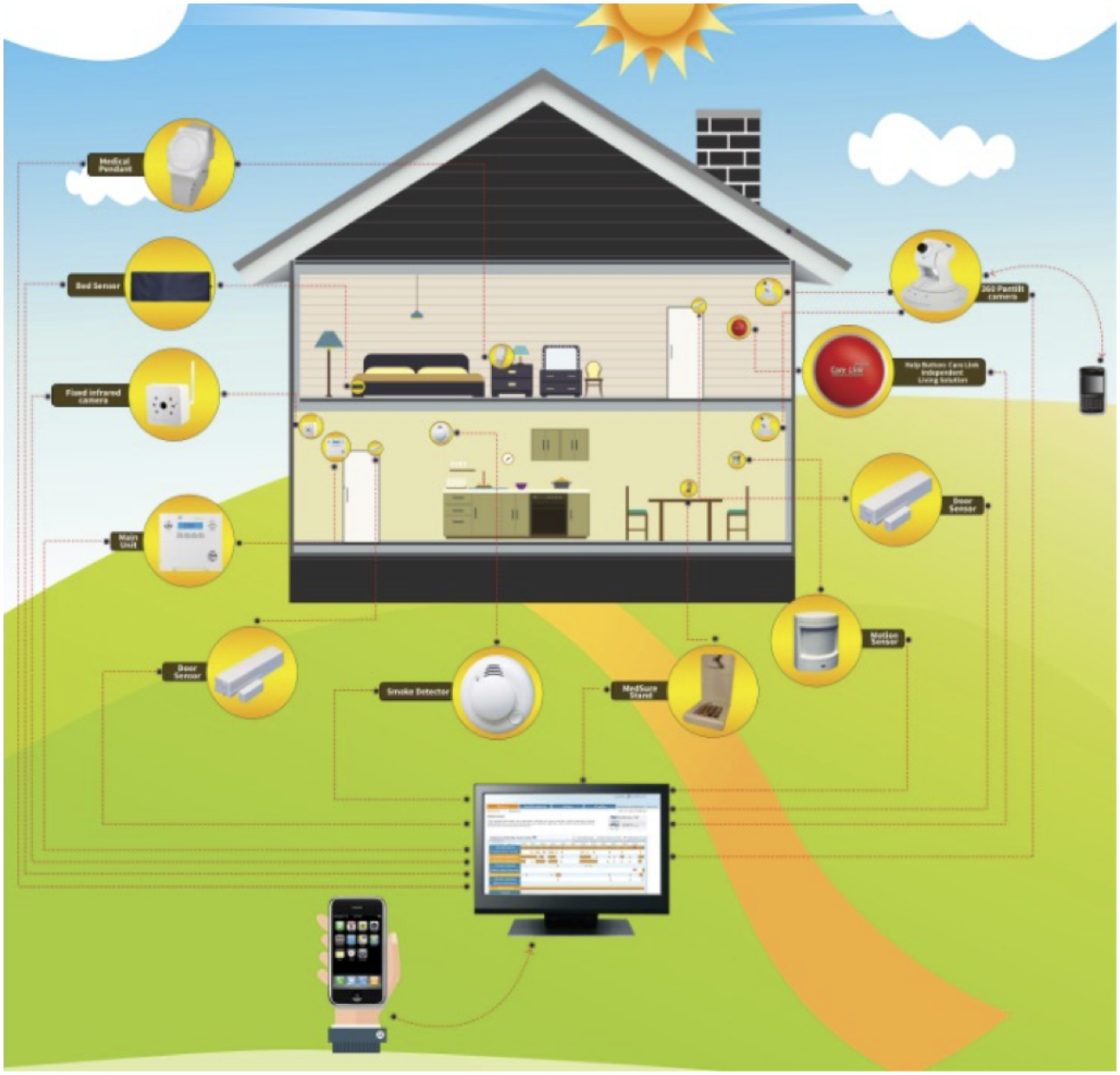
Participants received their usual publicly funded home care services provided by their provincial or regional health care authority; some participants also purchased privately funded home care services to supplement the publicly funded services received. These services included home visits by assistive personnel for activities of daily living, nursing care, and other supports deemed necessary by the home care case coordinator or case manager. Home care assessments were conducted as usual by the regional authority case coordinator or case manager. Once the assessment was completed and care services were decided, services were provided by a contracted home care agency.

Intervention Group

Participants received their usual home care services (described earlier) and PRM systems provided and funded by the study technology partner. Once enrolled in the study, the technology partner visited the clients and their caregiver. Participants received written and verbal overviews of the PRM options. Each PRM system installed was tailored to meet the needs and preferences of both the home care client and their caregiver. PRM system options included a combination of long-life battery–powered sensors including motion sensors; Wi-Fi–enabled cameras; fall alert pendants or bracelets; contact sensors for internal or external doors, cupboards, and refrigerators; pressure mats that could be placed at the base of a toilet, under a mattress, or on a chair; and a medication sensor system. Each PRM system had a main panel that was connected to a phone jack installed in the home care client's residence. The main panel received the signals from each sensor and sent them to a secured server to be transmitted back to the authorized recipient (eg, a caregiver) via a cellular signal or broadband module (Global System for Mobile Communications radio). The PRM systems only required internet access to transmit Wi-Fi-camera data. See Figure 1 for an example of how a PRM system used in this study could be set up.

XSL•FO
RenderX

Figure 1. Example of a passive remote monitoring system setup.



The role of the caregiver was key to the successful implementation of the PRM intervention. In addition to an array of sensors, the technology partner, in consultation with the home care client and caregiver, configured each PRM system to provide home care clients and their caregivers with a range of alerts for atypical events tailored to their situation. Examples of alerts include reminders for clients to take their medication; the ability for caregivers to monitor medication use; home care client movement patterns within the home or when leaving the home; fall emergencies; abnormal eating patterns; and abnormal length of time in bed, on a chair, or in a toilet. Notifications of atypical events were sent to the caregiver via email, SMS text message, or phone call. Notably, notifications of home care clients' atypical behaviors were not directly transmitted to a health care provider; rather, health care providers were notified at the discretion of the caregiver and in collaboration with the home care client, as appropriate. Possible caregiver actions based on notifications could include a telephone call to check on the client's safety, deploying assistive home care supports, checking the video (if cameras are present), taking the client to seek medical attention, or emergency action such as calling an ambulance. The technology partner provided education (written and verbal) to the home care clients and caregivers about PRM sensors and the types of alerts; the technology provider technician communicated any changes to the sensor system within the home with the research staff.

The intervention was provided to participants for 12 months during the study. After this time, the home care client was transitioned to usual care if they did not wish to retain their PRM system at a reduced cost.

Data Collection

Data Collection Window

Data collection aligned with the rolling recruitment of participants (see the *Recruitment* section) and lasted from April 2017 to February 2021. Paper format surveys comprised of standardized and researcher-developed tools were used to collect

XSL•FO
RenderX

data from participant pairs three times over the course of a 12-month period: (1) at the point of recruitment when consent was obtained (ie, baseline), (2) 6 months following recruitment, and (3) 12 months following recruitment. Data were collected in person at the home care client's home at baseline and by telephone or in person at the client's home for the 6-month and 12-month follow-ups. Home care clients and caregivers were interviewed separately where possible.

Survey

Both home care clients and caregivers completed a survey consisting of validated questionnaires and researcher-developed questions to provide demographic information (eg, date of birth, education, sex, marital status, and household income); provincial health insurance numbers (used for data linkage to provincial administrative health databases); feelings of home care clients' safety in their home; appraisals of home care quality; and satisfaction with and impact of PRM on both client and caregiver well-being (asked to the intervention group only). Additional researcher-developed items were asked of caregivers to understand their life satisfaction, self-rated mental health, and daily stress.

Home care clients completed standardized tools to assess their quality of life [24], Hospital Admission Risk Profile score, cognitive capacity (Mini-Mental State Exam) [25], and independent activities of daily living [26]. Caregivers also responded to standardized tools to assess their caregiving context, experiences, and information needs [27,28]; caregiver burden [29]; the home care client's abilities to complete independent activities of daily living (a part of the Hospital Admission Risk Profile) [26]; positive aspects of caregiving [30]; presenteeism at work [31]; and the financial impact of caregiving [32].

Impact of the COVID-19 Pandemic on Data Collection

The onset of the COVID-19 pandemic response in March 2020 occurred during the data collection period. Research staff adhered to all public health guidelines for risk mitigation and collected data entirely by phone during this period. Changes to health care service delivery models and stay-at-home orders in Ontario and Nova Scotia altered how these participants (1) could access and use health care services and (2) interacted with their participant partner (ie, the other half of the client-caregiver dyad). Participants' adherence to public health stay-at-home orders resulted in behaviors (eg, hesitancy to engage with acute and LTC services) that completely conflicted with and was contrary to our research outcome creating significant bias within the proportion of participants' data collected after March 2020. In consultation with epidemiologists and statisticians, a decision was made to control for the effects of the pandemic response, and participant pairs were only included for analysis if all study activities were completed before March 2020.

Administrative Data Linkage

Overview

Participants were enrolled between April 11, 2017, and January 15, 2020, in Ontario and between November 29, 2017, and December 16, 2019, in Nova Scotia. To determine the impact of the intervention on the primary study outcomes, the data collected by the research team (herein referred to as the PRCT database) was linked to health administrative databases held at the Institute for Clinical Evaluative Sciences (ICES; Ontario) and Health Data Nova Scotia (HDNS; Nova Scotia).

ICES is an independent, nonprofit research institute funded by an annual grant from the Ontario Ministry of Health and the Ministry of Long-Term Care. As a prescribed entity under Ontario's privacy legislation, ICES is authorized to collect and use health care data for the purposes of health system analysis, evaluation, and decision support. Secure access to these data is governed by policies and procedures that are approved by the Information and Privacy Commissioner of Ontario.

HDNS is a similar institution to ICES. It is a data repository based in the Faculty of Medicine, Department of Community Health and Epidemiology at Dalhousie University, that is focused on supporting data-driven research for researchers in Nova Scotia. HDNS facilitates research and innovation in Nova Scotia by providing access to linkable administrative health data and analysis for research and health service assessment purposes in a secure, controlled environment, while respecting the privacy and confidentiality of Nova Scotians.

Data Sources

We used the PRCT database to identify patient and caregiver participants, including measures of risk that the patient had to alternate levels of care and hospital admission and measures of burden on caregivers. Health administrative databases in both Ontario and Nova Scotia were used to determine other baseline characteristics and outcomes and identify participant data, as follows:

1. Demographic and geographic information as recorded in administrative data
2. Preexisting comorbidities (using validated algorithms)
3. Previous emergency department visits
4. Previous surgeries
5. Previous outpatient, primary care, and specialist visits and follow-up LTC claims
6. Previous prescription use and follow-up LTC
7. Previous home care services
8. Previous and follow-up hospital admissions
9. International classification of diseases diagnosis codes—version 10 and Ontario health insurance plan diagnosis codes.

In Ontario, datasets were linked using unique encoded identifiers and analyzed at ICES. In Nova Scotia, the databases were matched whenever possible. When databases could not be matched, HDNS used alternate databases with similar data to approximate what was done in Ontario. For a complete listing of all databases used, see Table 1.

XSL•FO
RenderX

Table 1. Databases used by the Institute for Clinical Evaluative Sciences (ICES) and Health Data Nova Scotia (HDNS) to support analysis.

Characteristics	ICES databases (Ontario)	HDNS databases (Nova Scotia)
Demographic and geographic information	The registered persons database and PCCF[a]	• Insured patient registry (MASTER) • Postal code conversion file (PCCF+)
Preexisting comorbidities using validated algorithms	Ontario Diabetes Dataset, Chronic Obstructive Pulmonary Disease, congestive heart failure, Ontario hypertension dataset (HYPER), Ontario asthma dataset (ASTHMA), Ontario dementia database (DEMENTIA)	• MSI physician's billings (MED)
Emergency department visits	NACRS[b]	• NACRS
Surgeries	Same Day Surgery database	—[c]
Outpatient, primary care, and specialist visits and follow-up long-term care claims	OHIP[d] claims	• MSI physician's billings (MED) • Eligibility group (EGROUP)
Prescription use and follow-up long-term care	Ontario Drug Benefit Claims and Drugs list	• Nova Scotia Drug Information System
Home care services	Home Care Database	• Nova Scotia Health Department of Seniors and Long Term Care
Hospital admissions	DAD[e]	• DAD
Descriptions for international classification of diseases diagnosis codes—version 10 and OHIP diagnosis codes	REF[f]	• REF

[a]PCCF: postal code conversion file.

[b]NACRS: National Ambulatory Care Reporting System.

[c]Not available.

[d]OHIP: Ontario health insurance plan.

[e]DAD: Discharge Abstract Database.

[f]REF: reference files.

Primary Outcome

The primary outcome of interest was to support older adults with complex care needs to safely remain in their home and avoid or delay admission to higher levels of care, such as long-term hospitalization and LTC. The primary outcome was defined as time to a client's inability to return home given the need for LTC or hospital admission under an alternative level of care designation within 1 year of enrollment in the trial. This was ultimately determined by the most responsible provider. The secondary outcome of health service use and mortality was assessed to confirm similar mortality rates within 1 year of enrollment in the trial. Clients were followed up to 1 year until February 28, 2020, to control for the impact of the COVID-19 pandemic. Participants enrolled before February 28, 2019, were included to ensure 1 year of follow-up before the COVID-19 pandemic.

Data Analysis

Baseline characteristics were summarized using descriptive statistics: continuous variables as means and SDs and categorical variables as proportions. Baseline characteristics between the control and intervention participants were examined to verify successful randomization using standardized differences that when greater than 0.1, indicates a potentially meaningful difference [33]. In accordance with ICES and HDNS privacy policies, cell sizes less than or equal to 5 were not reported.

A Cox proportional hazards model was used to evaluate our primary outcome of unable to return home, censoring on mortality and 1 year after enrollment, and our secondary outcome of mortality, only censoring on 1 year after enrollment. Hazard ratios (HRs) with 95% CI were reported. Proportional hazards were assessed using interaction with time terms, and no violations were observed.

Subgroup analyses compared health care use 1 year before and after enrollment for home care clients to determine what effect the intervention had on participants' use of the health care system. For the subgroup analysis, a paired analysis was conducted using the McNemar test for proportions and the Wilcoxon signed rank test for means to compare differences in health care use 1 year before and after enrollment in both patients (Ontario: n=108; Nova Scotia: n=105) and caregivers (Ontario: n=104; Nova Scotia: n=101). All analyses were conducted using SAS (version 9.4; SAS Institute). A statistical significance was defined as a 2-tailed α value less than .05.

Ethical Considerations

This study was approved by the Research Ethics Board at Western University (registration IRB 00000940) and at Nova Scotia Health (registration IRB 102203). All participants

XSL•FO
RenderX

provided written informed consent and could withdraw from the study at any time without penalty. All data were deidentified and assigned a study ID number to support data linkages throughout data collection and analyses. Participants received no compensation to complete the study; intervention group participants received the PRM systems at no personal expense.

Results

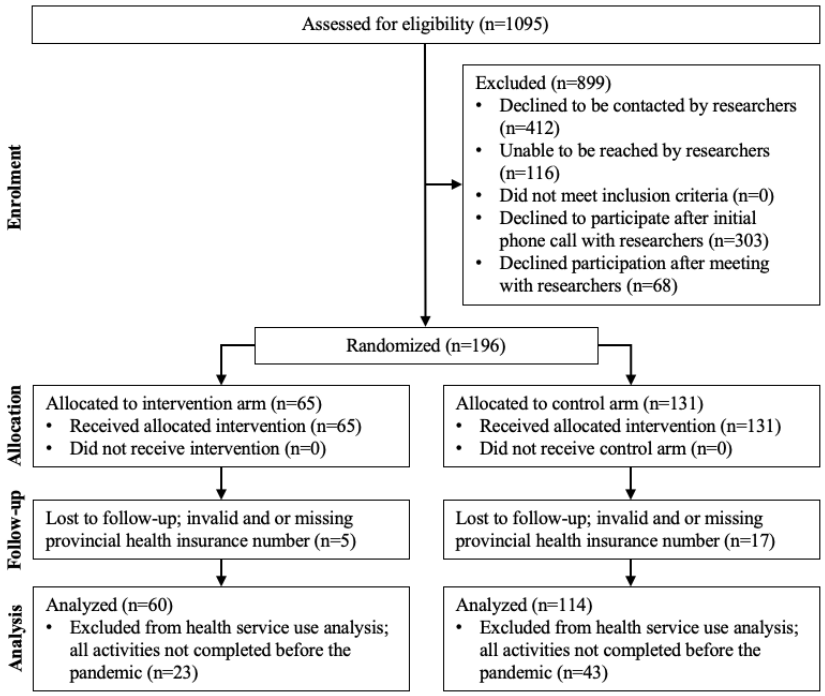
Due to challenges joining administrative patient data—such as data collected as part of this study—across provincial jurisdictions, a single analysis was not possible. The following sections present the results of the separate analyses conducted by ICES in Ontario and HDNS in Nova Scotia to address this study's research question.

Participant Characteristics

Ontario Cohort

The total Ontario cohort comprised 339 participants (home care clients: n=174; caregivers: n=165). In total, 196 home care clients were enrolled in Ontario from April 11, 2017, to January 15, 2020, of which 22 were excluded due to invalid or missing health card numbers, leaving 174 home care clients to be included in analysis. We were able to link 165 caregivers to the home care clients. In Ontario, 116 participants (home care clients: n=60; caregivers: n=56) included in this analysis received the study-provided PRM intervention plus their normal home care and 223 participants (home care clients: n=114; caregivers: n=109) received their normal home care without PRM. Figure 2 provides the randomization and analysis study flow diagram for the Ontario home care client cohort.

Figure 2. Ontario home care client cohort randomization and analysis flow diagram.

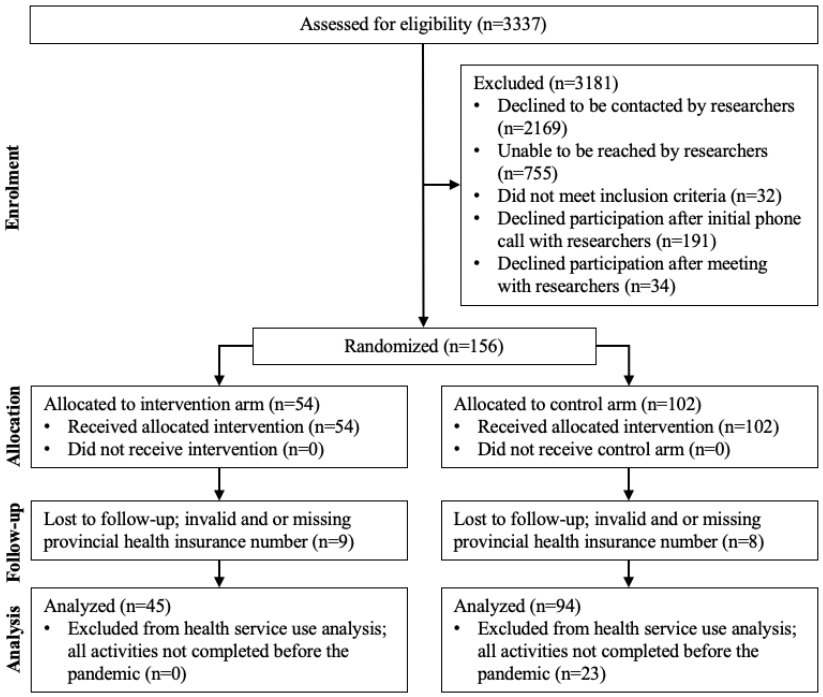


Nova Scotia Cohort

The total Nova Scotia cohort comprised 273 participants (home care clients: n=139; caregivers: n=134). In total, 156 home care clients were enrolled in Nova Scotia from November 29, 2017, to December 16, 2019, of which 17 were excluded due to invalid or missing health card numbers, leaving 139 home care clients to be included in the analysis. We were able to link 134 caregivers to the home care clients. In Nova Scotia, 92 participants (home care clients: n=45; caregivers: n=47) included in this analysis received the study-provided PRM intervention plus their normal home care and 181 participants (home care clients: n=94; caregivers: n=87) received their normal home care without PRM. Figure 3 provides the randomization and analysis study flow diagram for the Nova Scotia home care client cohort.

XSL•FO
RenderX

Figure 3. Nova Scotia home care client cohort randomization and analysis flow diagram.



Home Care Clients

Ontario Home Care Clients

In Ontario, there were no significant differences between home care clients in intervention and control groups at baseline on any of the demographic characteristics evaluated (all $P>.05$). At baseline, Ontario home care clients were on average aged 78 (SD 6.9) years. There was a higher proportion of male participants in the intervention group (40/60, 67%) compared with the control group (57/114, 50%). Most home care clients (125/174, 73%) resided in midsize urban centers. Approximately 20% (35/174) of home care clients resided in the lowest-income neighborhoods (quintile 1), while about 15% (26/174) resided in the highest-income neighborhoods. Most Ontario home care clients (109/174, 63%) lived with their spouse. Nearly all Ontario home care clients (>106/174, >90%) were at medium or high risk of hospitalization. Ontario home care clients had a variety of chronic disease comorbidities, most notably hypertension (144/174, approximately 83%), diabetes (95/174, approximately 55%), and dementia (67/174, approximately 40%). Complete home care client baseline characteristics are presented in Table 2.

XSL•FO
RenderX

Table 2. Ontario and Nova Scotia home care client baseline characteristics.

Characteristics	Ontario		Nova Scotia	
	Control (n=114)	Intervention (n=60)	Control (n=94)	Intervention (n=45)
Age (y), mean (SD)	77.9 (6.9)	78.1 (7.0)	80.1 (7.9)	81.5 (7.3)
Sex, n (%)				
Female	57 (50)	20 (33.3)	62 (66)	33 (73.3)
Male	57 (50)	40 (66.7)	32 (34)	12 (26.7)
Level of rurality, n (%)				
Large CMA[a] (≥500,000)	0 (0)	0 (0)	0 (0)	£5[b]
Midsize CMA (100,000-500,000)	80 (70.2)	45 (75)	66 (70.2)	35 (77.8)
Small CMA (10,000-100,000)	24 (21.1)	9 (15)	13 (13.8)	6 (13.3)
Non-CMA	10 (8.8)	6 (10)	15 (16)	£5
Neighborhood-level income quintile, n (%)				
Quintile 1 or missing	24 (21.1)	14 (23.3)	16 (17)	9 (20)
Quintile 2	20 (17.5)	15 (25)	29 (30.9)	17 (37.8)
Quintile 3	28 (24.6)	9 (15)	19 (20.2)	6 (13.3)
Quintile 4	23 (20.2)	10 (16.7)	13 (13.8)	7 (15.6)
Quintile 5	14 (12.3)	12 (20)	17 (18.1)	6 (13.3)
Living arrangements, n (%)				
Alone	≤5	0 (0)	56 (59.6)	29 (64.4)
With spouse	71 (62.3)	38 (63.3)	12 (12.8)	≤5
With children	0 (0)	0 (0)	0 (0)	0 (0)
Other	≤45	22 (36.7)	0 (0)	≤5
Missing	0 (0)	0 (0)	26 (27.6)	8 (18)
HARP[c], n (%)				
Low risk	≤5	≤5	8 (8.5)	≤5
Medium risk	73 (64)	33 (55)	43 (45.7)	≤20
High risk	≤40	≤30	40 (42.6)	23 (51.1)
Comorbidities (prevalence in sample), n (%)				
Hypertension	95 (83.3)	49 (81.7)	31 (33)	21 (46.7)
Diabetes	61 (53.5)	34 (56.7)	29 (30.9)	12 (26.7)
Dementia	42 (36.8)	25 (41.7)	31 (33)	12 (26.7)
COPD[d]	40 (35.1)	18 (30)	15 (16)	10 (22.2)
CHF[e]	30 (26.3)	16 (26.7)	7 (7.4)	6 (13.3)
Asthma	25 (21.9)	10 (16.7)	≤5	≤5
AMI[f]	6 (5.3)	≤5	≤5	≤5

[a]CMA: census metropolitan areas.

[b]Small cell sizes (ie, n£5) suppressed following the Institute for Clinical Evaluative Sciences and Health Data Nova Scotia reporting guidelines for maintaining participant confidentiality.

[c]HARP: Hospital Admission Risk Profile.

[d]COPD: chronic obstructive pulmonary disease.

[e]CHF: congestive heart failure.

[f]AMI: acute myocardial infarction.

XSL•FO
RenderX

Nova Scotia Home Care Clients

In Nova Scotia, there were no significant differences between home care clients in intervention and control groups at baseline on any of the demographic characteristics evaluated (all *P*>.05). In Nova Scotia, home care clients were on average 81 (SD 7.5) years of age. At baseline, there was a higher proportion of male participants in the control group (32/94, 34%) compared with the intervention group (12/45, 27%). Most Nova Scotia home care clients (101/139, 75%) resided in midsize urban centers. Approximately 20% (25/139) resided in the lowest-income neighborhoods, while approximately 15% (23/139) resided in the highest-income neighborhoods. The majority of Nova Scotia home care clients (85/139, approximately 60%) lived alone. Nearly all home care clients (>114/139, >90%) were at a medium or high risk of hospitalization. The Nova Scotia home care clients also had a variety of chronic disease comorbidities, most notably hypertension (52/139, approximately 40%), dementia (43/139, approximately 30%), and diabetes (41/139, approximately 30%). Complete home care client baseline characteristics are presented in Table 2.

Family and Friend Caregivers

Ontario Caregivers

In Ontario, there were no significant differences between intervention and control caregivers at baseline for any of the characteristics evaluated (all *P*>.05). Ontario caregivers were on average 67 (SD 12.4) years of age. There was a higher proportion of male participants in the control group (39/109, 36%) compared with the intervention group (13/56, 23%). Most caregivers (116/165, 70%) resided in midsize urban centers. Approximately 20% (32/165) resided in the lowest-income neighborhoods, while about 17.6% (29/165) resided in the highest-income neighborhoods. About 40% (<55/165) of Ontario caregivers attained a postsecondary education. Just over half of the caregivers in both groups (99/165, 60%) were retired, with 10% (12/109) of control and 25% (15/56) of intervention caregivers reporting full-time employment. Approximately 66% (114/165) of caregivers were participating in the study with their spouse or partner, and over 80% (144/165) of caregivers lived with the home care client they supported. Approximately 33% (58/165) indicated they would consider moving the home care client to a LTC facility. Ontario caregivers also had a variety of chronic disease comorbidities, most notably hypertension (control: 56/109, 51%; intervention: 37/56, 66%), diabetes (32/165, 20%), and chronic obstructive pulmonary disease (29/165, 18%). Ontario caregivers provided an average of 109 to 119 hours of care or support each week. Scores on the Zarit Burden Interview suggest that caregivers in both groups were similarly burdened by caregiving responsibilities and that such levels of caregiver burden approached high levels of burden. Complete caregiver baseline characteristics are presented in Table 3.

XSL•FO
RenderX

Table 3. Ontario and Nova Scotia family and friend caregiver baseline characteristics.

Characteristics	Ontario		Nova Scotia	
	Control (n=109)	Intervention (n=56)	Control (n=87)	Intervention (n=47)
Age (y), mean (SD)	66.8 (12.2)	67.3 (12.6)	65.4 (10.9)	60.6 (12.1)
Hours spent providing care or support for home care client per week, mean (SD)	109.0 (68.8)	119.0 (67.9)	80.9 (69.3)	80.6 (71.4)
Caregiver burden[a], mean (SD)	17.8 (9.7)	16.8 (8.4)	17.0 (9.3)	18.3 (11.7)
Sex, n (%)[b]				
Female	70 (64.2)	43 (76.8)	55 (63.2)	33 (70.2)
Male	39 (35.8)	13 (23.2)	32 (36.8)	14 (29.8)
Level of rurality, n (%)[b]				
Large CMA[c] (≥500,000)	≤5	0 (0)	0 (0)	0 (0)
Midsize CMA (100,000-500,000)	75 (68.8)	41 (73.2)	61 (70.1)	38 (80.9)
Small CMA (10,000-100,000)	23 (21.1)	8 (14.3)	13 (14.9)	6 (12.8)
Non-CMA	≤17	7 (12.5)	13 (14.9)	≤5
Missing	≤5	0 (0)	0 (0)	≤5
Neighborhood-level income quintile, n (%)[b]				
Quintile 1	22 (20.2)	10 (17.9)	15 (17.2)	6 (12.8)
Quintile 2	18 (16.5)	19 (33.9)	20 (23)	12 (25.5)
Quintile 3	≤40	7 (12.5)	16 (18.4)	13 (27.7)
Quintile 4	24 (22)	7 (12.5)	15 (17.2)	7 (14.9)
Quintile 5	16 (14.7)	13 (23.2)	21 (24.1)	9 (19.1)
Missing	≤5	0 (0)	0 (0)	0 (0)
Self-reported highest level of education, n (%)[b]				
Elementary school	21 (18.4)	9 (15)	10 (11.5)	7 (14.9)
High school	38 (33.3)	17 (28.3)	25 (28.7)	18 (38.3)
College, CEGEP[d], or university	40 (35.1)	28 (46.7)	42 (48.3)	15 (31.9)
Postgraduate degree	12 (10.5)	≤5	10 (11.5)	≤5
Other, declined, or N/A[e]	≤5	≤5	0 (0)	≤5
Missing	≤5	≤5	0 (0)	0 (0)
Self-reported employment status, n (%)[b]				
Full time	12 (10.5)	15 (25)	23 (26.4)	13 (27.7)
Part time	20 (17.5)	≤5	10 (11.5)	9 (19.1)
Leave of absence	≤5	0 (0)	≤5	≤5
Retired	65 (57)	34 (56.7)	48 (55.2)	14 (29.8)
Not employed	14 (12.3)	6 (10)	≤5	10 (21.3)
Declined or N/A	0 (0)	0 (0)	0 (0)	≤5
Missing	≤5	≤5	≤5	≤5
Relationship to home care client, n (%)[b]				
Spouse or partner	73 (64)	41 (68.3)	34 (39.1)	14 (29.8)
Adult child	25 (21.9)	16 (26.7)	35 (40.2)	24 (51.1)
Parent	6 (5.3)	0 (0)	≤5	≤5

XSL•FO
RenderX

Characteristics	Ontario		Nova Scotia	
	Control (n=109)	Intervention (n=56)	Control (n=87)	Intervention (n=47)
Other, declined, or N/A	≤15	≤5	17 (19.5)	8 (17)
Missing	≤5	≤5	≤5	≤5
Living with home care client, n (%[b])				
Yes	95 (83.3)	49 (81.7)	59 (67.8)	30 (63.8)
Would consider moving home care client to long-term care facility, n (%)[b]				
No	≤80	≤50	64 (73.6)	32 (68.1)
Yes	38 (33.3)	20 (33.3)	22 (25.3)	14 (29.8)
Declined or N/A	≤5	0 (0)	≤5	≤5
Missing	≤5	≤5	≤5	≤5
Comorbidities, n (%)[b]				
Hypertension	56 (51.4)	37 (66.1)	34 (39.1)	13 (27.7)
Diabetes	19 (17.4)	13 (23.2)	22 (25.3)	7 (14.9)
COPD[f]	17 (15.6)	12 (21.4)	≤5	≤5
CHF[g]	6 (5.5)	≤5	0 (0)	0 (0)
Asthma	10 (9.2)	7 (12.5)	≤5	≤5
Dementia	≤5	0 (0)	≤5	0 (0)
AMI[h]	≤5	≤5	≤5	0 (0)

[a]Caregiver burden assessed using the 12-item Zarit Burden Interview [29]. Scores range from 0 to 48, with higher scores reflecting higher levels of caregiver burden and scores of 17 and higher indicating a high level of caregiver burden.

[b]Small cell sizes (ie, n£5) suppressed following the Institute for Clinical Evaluative Sciences and Health Data Nova Scotia reporting guidelines for maintaining participant confidentiality.

[c]CMA: census metropolitan area.

[d]CGEP: collège d'enseignement général et professionnel.

[e]N/A: not applicable.

[f]COPD: Chronic Obstructive Pulmonary Disease.

[g]CHF: congestive heart failure.

[h]AMI: acute myocardial infarction.

Nova Scotia Caregivers

In Nova Scotia, there were no significant differences between intervention and control caregivers at baseline for any of the characteristics evaluated (all *P*>.05). On average, caregivers were aged 63 (SD 11.6) years. Approximately 35% (46/134) of Nova Scotia caregivers were male. Most caregivers (99/134, 75%) resided in midsize urban centers. Approximately 15% (21/134) of Nova Scotia caregivers resided in the lowest-income neighborhoods, while about 20% (30/134) resided in the highest-income neighborhoods. Postsecondary school was the highest level of education for 48% (42/87) of control caregivers and 32% (15/47) of intervention caregivers. Less than 30% (36/134) of Nova Scotia caregivers were employed full time; approximately 50% (48/87) of control caregivers and 30% (14/47) of intervention caregivers were retired. Approximately 35% (48/134) were participating with their spouse or partner, while approximately 45% (59/134) were participating with their parent. Approximately 66% (99/134) of Nova Scotia caregivers lived with the home care client they supported. Approximately 25% (36/134) of caregivers indicated they would consider moving the home care client to an LTC facility. Nova Scotia caregivers also had a variety of chronic disease comorbidities, most notably hypertension (control: 34/87, 39%; intervention: 13/47, 28%) and diabetes (control: 22/87, 25%; intervention: 7/47, 15%). Nova Scotia caregivers provided an average of approximately 81 hours of care or support each week. Scores on the Zarit Burden Interview suggests that caregivers in both groups were similarly burdened by caregiving responsibilities and that such levels of caregiver burden constituted high levels of burden. Complete caregiver baseline characteristics are presented in Table 3.

Home Care Clients' Admission to Higher Levels of Care

For the primary outcome in Ontario, there was a median follow-up time of 365 days, and the intervention group had an event rate of 5.5 per 10,000 person-days compared with the event rate of 8.0 per 10,000 person-days for the control group (Table 4). The intervention group was associated with a nonsignificant 30% reduction in risk of being unable to return home (HR 0.7, 95% CI 0.3-1.4). For the secondary outcome in

XSL•FO
RenderX

Ontario, home care client mortality, there was a median follow-up time of 365 days, and the intervention group had an event rate of 2.9 per 10,000 person-days vs the control group that had an event rate of 2.5 per 10,000 person-days (Table 4). Both groups had similar risk of mortality (HR 1.2, 95% CI 0.4-3.2).

Table 4. Unadjusted hazard ratio of primary and secondary outcomes and 95% CIs.

	Ontario			Nova Scotia		
	Control (n=114)	Intervention (n=60)	*P* value	Control (n=94)	Intervention (n=34)	*P* value
Primary outcome: unable to return home						
Events, n (%)	27 (24)	10 (17)	—[a]	8 (9)	≤5	—
Follow-up in days, median (IQR)	365 (257-365)	365 (334-365)	—	365 (204-365)	365 (200-365)	—
Event rate per 10,000 person-days	8.0	5.5	—	2.9	3.2	—
Values, hazard ratio (95% CI)	Reference	0.7 (0.3-1.4)	.30	Reference	1.1 (0.3-3.7)	.90
Secondary outcome: mortality						
Events, n (%)	10 (9)	6 (10)	—	16 (17)	8 (18)	—
Follow-up in days, median (IQR)	365 (365-365)	365 (365-365)	—	365 (211-365)	365 (214-365)	—
Event rate per 10,000 person-days	2.5	2.9	—	5.8	6.2	—
Values, hazard ratio (95% CI)	Reference	1.2 (0.4-3.2)	.80	Reference	1.1 (0.5-2.5)	.90

[a]Not applicable.

For the primary outcome in Nova Scotia, there was a median follow-up of 365 days, and the intervention group had an event rate of 3.2 per 10,000 person-days compared with the event rate of 2.9 per 10,000 person-days for the control group (Table 4). The intervention group was associated with a nonsignificant 10% increase in risk of being unable to return home (HR 1.1, 95% CI 0.3-3.7). For the secondary outcome in Nova Scotia, home care client mortality, there was a median follow-up time of 365 days, and the intervention group had an event rate of 6.2 per 10,000 person-days versus the control group that had an event rate of 5.8 per 10,000 person-days (Table 4). Both groups had similar risk of mortality (HR 1.1, 95% CI 0.5-2.5).

Health Care Use

Ontario Home Care Clients

There was no significant difference in health care use at the 1-year poststudy enrollment for Ontario control home care clients for all services included in the analysis (all $P>.05$). For Ontario intervention home care clients, significantly fewer participants used emergency departments ($P=.01$) at the 1-year poststudy enrollment (24/37, 65%) compared with prestudy enrollment (31/37, 84%). In addition, on average, Ontario intervention home care clients made significantly fewer visits to outpatient services (enrollment: mean 42.2, SD 31.0; after enrollment: mean 32.9, SD 26.0; $P=.04$) and specialist services (enrollment: mean 20.7, SD 25.3; after enrollment: mean 14.6 SD 18.4; $P=.02$). There were no differences in hospital admissions, primary care visits, or prescriptions at the 1-year poststudy enrollment for Ontario intervention home care clients. Complete health care use outcomes for both groups of Ontario home care clients are presented in Table 5.

XSL•FO
RenderX

Table 5. Ontario home care client health care use at the prestudy and 1-year poststudy enrollment.

Health care use	Control			Intervention		
	Before enrollment (n=71)	After enrollment (n=71)	*P* value	Before enrollment (n=37)	After enrollment (n=37)	*P* value
Hospital admissions						
Values, n (%)	26 (36.6)	31 (43.7)	.35	21 (56.8)	14 (37.8)	.07
Values, mean per client (SD)	0.7 (1.2)	0.7 (1.1)	.72	0.9 (1.1)	0.7 (1.0)	.38
Emergency department visits						
Values, n (%)	50 (70.4)	47 (66.2)	.56	31 (83.8)	24 (64.9)	.01
Values, mean per client (SD)	2.1 (2.4)	2.3 (2.8)	.99	2.5 (2.6)	2.0 (2.2)	.39
Outpatient physician visits						
Values, n (%)	71 (100)	71 (100)	—[a]	37 (100)	37 (100)	—
Values, mean per client (SD)	37.5 (35.3)	32.1 (28.2)	.35	42.2 (31.0)	32.9 (26.0)	.04
Primary care visits						
Values, n (%)	71 (100)	66-71 (92.9-100)	NR[b]	37 (100)	37 (100)	NR
Values, mean per client (SD)	14.6 (16.9)	13.0 (11.6)	.86	13.3 (9.5)	12.6 (10.8)	.45
Specialist visits						
Values, n (%)	66-71 (92.9-100)	64 (90.1)	NR	32-37 (86.5-100)	32-37 (86.5-100)	NR
Values, mean per client (SD)	16.0 (23.4)	13.0 (16.5)	.25	20.7 (25.3)	14.6 (18.4)	.02
Prescriptions						
Values, n (%)	66-71 (92.9-100)	66-71 (92.9-100)	NR	32-37 (86.5-100)	32-37 (86.5-100)	NR
Values, mean per client (SD)	6.5 (5.3)	6.7 (5.2)	.27	6.4 (4.5)	5.9 (4.4)	.24

[a]Not applicable.

[b]NR: not reported.

Nova Scotia Home Care Clients

On average, Nova Scotia home care clients in the control group made significantly fewer visits to primary care at the 1-year poststudy enrollment (enrollment: mean 9.2, SD 5.7; after enrollment: mean 7.8, SD 5.7; *P*=.004). There was no significant difference in health care use at the 1-year poststudy enrollment for Nova Scotia control home care clients for all other services included for the analysis (all *P*>.05). At the 1-year poststudy enrollment, Nova Scotia intervention home care clients averaged significantly more hospitalization visits (enrollment: mean 0.3, SD 0.5; after enrollment: mean 0.6, SD 0.9; *P*=.04) and emergency department visits (enrollment: mean 0.9, SD 1.3; after enrollment: mean 1.5, SD 1.8; *P*=.02) and significantly fewer outpatient services (enrollment: mean 16.6, SD 13.2); after enrollment: mean 13.1, SD 11.4; *P*=.005) and primary care visits (enrollment: 8.5, SD 6.8; after enrollment: mean 6.7, SD 5.5; *P*=.02). There were no differences in prescriptions at the 1-year poststudy enrollment for Nova Scotia intervention home care clients (all *P*>.05). Complete health care use outcomes for both groups of Nova Scotia home care clients are presented in Table 6.

XSL•FO
RenderX

Table 6. Nova Scotia home care client health care use at the prestudy and 1-year poststudy enrollment.

Health care use	Control			Intervention		
	Before enrollment (n=71)	After enrollment (=71)	*P* value	Before enrollment (n=34)	After enrollment (n=34)	*P* value
Hospital admissions						
Values, n (%)	19 (26.8)	27 (38)	.10	9 (26.5)	15 (44.1)	.08
Values, mean per client (SD)	0.3 (0.6)	0.5 (0.8)	.07	0.3 (0.5)	0.6 (0.9)	.04
Emergency department visits						
Values, n (%)	34 (47.9)	39 (54.9)	.28	18 (52.9)	22 (64.7)	.21
Values, mean per client (SD)	1.2 (1.7)	1.1 (1.5)	.82	0.9 (1.3)	1.5 (1.8)	.02
Outpatient physician visits						
Values, n (%)	71 (100)	71 (100)	—[a]	32 (94.1)	34 (100)	—
Values, mean per client (SD)	17.4 (9.9)	16.9 (17.2)	.03	16.6 (13.2)	13.1 (11.4)	.005
Primary care visits						
Values, n (%)	71 (100)	69 (97.2)	—	32 (94.1)	33 (97.1)	.56
Values, mean per client (SD)	9.2 (5.7)	7.8 (5.7)	.004	8.5 (6.8)	6.7 (5.5)	.02
Specialist visits						
Values, n (%)	65 (91.5)	58 (81.7)	.07	26 (76.5)	27 (79.4)	.65
Values, mean per client (SD)	5.0 (5.7)	5.6 (14.1)	.54	4.4 (5.3)	3.8 (4.1)	.39
Prescriptions						
Values, n (%)	63 (88.7)	65 (91.5)	.53	29 (85.3)	32 (94.1)	.18
Values, mean per client (SD)	5.3 (3.8)	5.3 (3.8)	.63	5.3 (4.3)	5.3 (3.9)	.62

[a]Not available.

Discussion

Primary Outcome of PRM Within Home Care

While our study did not yield statistically significant results regarding the effectiveness of the PRM model in prolonging home stays, the observed trends suggest that technology-assisted aging in place may be a valuable goal for older adults. Despite the underpowered design of our study, the use of the PRM model of care was associated with meaningful, albeit nonstatistically significant, differences between control and intervention groups on the primary outcome (ie, home care clients' ability to stay at home) in Ontario and Nova Scotia. Specifically, our findings with the Ontario intervention participants trended toward a greater number of home care clients staying home and staying home longer compared with those without PRM in their home. This is similar to the findings of others who have examined the impact of remote monitoring technologies on patient and caregiver well-being [34,35]. The Canadian Expert Panel on Aging highlights the importance of even insignificant delays to admission to LTC [2]. They report that retaining and appropriately supporting older adults at home—even for a month—positively impacts the overall LTC system incapacity or capacity and growing waitlist for care. Furthermore, Canadian older adults' overwhelmingly prefer to age at home and 11% to 30% of Canadians admitted to LTC could have remained at home and in their communities if adequate home care and community supports were available [36,37].

It is important to highlight several participant characteristics that contextualize the nature of this study's patient-caregiver dyads when considering participants' ability to stay at home. The PRM intervention in this study purposefully targeted a select cohort of older adult home care clients (mean age range 77-81 years), many (125/314, 40%) with limited personal resources (ie, education or income) but all who required significant supportive care to remain in their home environment. Home care clients in Ontario were mostly cared for by a spouse living in the same home residence, whereas in Nova Scotia, most older adults lived alone. All older adult home care clients in this study were dealing with significant comorbidities; the most prevalent were hypertension, diabetes, and dementia. In Ontario, about 2 in 5 home care clients were managing dementia, as were approximately 1 in 3 of those in Nova Scotia. This is significant, as previous research demonstrates that older adults (ie, home care clients in this study), especially those with

XSL•FO
RenderX

dementia, express a greater quality of life and social connections, less loneliness, and fewer depressive symptoms when living at home compared with institutionalized LTC [38-41]. The health status of home care clients in this study was characteristic of older adults at risk for higher levels of care, with dementia being the greatest predictor of LTC home admissions [2].

Secondary Outcomes of PRM Within Home Care

What was notable among Ontario intervention home care clients was a significant reduction in their health service use over the course of the study. We found that compared with those receiving usual care, a smaller proportion of Ontario older adult home care clients in the intervention group used emergency departments after 1 year compared with baseline; no such difference was observed in home care clients in the control arm. In addition, Ontario intervention home care clients experienced a significant decrease in the average number of visits to outpatient and specialist services after 1 year, while there were no changes in average visits to these services for home care clients in the control arm. This is an important finding given the *at-risk* health status of older adult home care recipients in the study. Finch et al [35] reported a reduction in health care costs reflective of fewer hospitalizations and emergency visits among "nursing home–eligible" older adults who were monitored by home-based PRM sensors to track their safety (eg, falls) and activities of daily living (eg, time in bed, toilet use, and opening and closing the refrigerator). Consistent with the findings of this study, Finch et al [35] concluded that PRM systems may contribute to a reduction in older adult health service use and associated reductions in health care costs.

However, this was not the case among home care clients in Nova Scotia who received the intervention. In contrast to their Ontario counterparts, Nova Scotia intervention home care clients averaged significantly greater hospital admissions and emergency department visits but significantly fewer outpatient and primary care visits. Discussion with Nova Scotia project partners (health care decision makers) provided some insight into the unanticipated health service use trends of Nova Scotia home care clients in our study. Consistent with our findings, Nova Scotia health system partners reported an observed trend of increased health system use among the Nova Scotia older adult population. The provincial health care decision makers conveyed that emergency services and hospitalization may be a work-around strategy people use to expedite access to LTC services ("skip the line") and circumvent existing policies governing LTC access. Current policies elevate the LTC priority status of hospitalized at-risk older adults who are on the LTC waitlist and given priority consideration for LTC placement (personal communication, June 2024). Home surveillance of older adults' behaviors, activities, and habits through PRM technologies by caregivers may also have contributed to increased health service use. Caregivers' surveillance of older adults' activities in their home may have illuminated older adults' need for greater care. In addition, PRM may generate greater insight into older adults' care needs that may influence (ie, increase) their health service use. This is consistent with pilot study findings of PRM home care services conducted in Nova Scotia (Report of CareLink Evaluation: The Nova Scotia experience, unpublished data, 2015).

Our findings of increased hospital health service use patterns may also be consequential to the 12% to 17% shortfall of primary care service provision within Nova Scotia [42]. With underresourced primary health care services, it may be that Nova Scotians, including complex care older adults needing health care services, would be inclined to access tertiary care services in the absence of primary care services. The Canadian Expert Panel on Aging cautioned that the lack of adequate home care and community-based health care services creates the conditions for premature institutionalization of older adults into LTC. Concernedly, current estimates suggest that up to 1 in 3 Canadians admitted to an LTC home could have remained home with adequate home care services [2]. Avoiding or delaying institutionalized care eases the health system burden particularly in LTC health settings that are perpetually underresourced and ill-equipped to manage the current and pending need for older adult care [43]. An estimated 52,000 Canadian adults are on waiting lists for a placement into an LTC home, while more than 430,000 are estimated to have unmet home care needs, with about 167,000 of them being aged ≥65 years [6,43].

Qualitative research evaluating active remote monitoring and PRM technologies has supported the use of PRM technology as a way to keep at-risk older adults in their home or extend their time at home before requiring institutionalized care [44]. The study, conducted in Nova Scotia, engaged expert key informants including provincial health care policy makers, home care providers, and remote monitoring technology experts to provide their insight into the use of home-based remote monitoring technologies. Key informant participants supported the use of active and PRM technologies as a way to keep complex care older adults at home or extend their time at home. The mitigated risk of adverse events, greater responsiveness to health emergencies (eg, falls), enhanced independence among older adults, and enhanced oversight and greater safety were reported with home monitoring compared with the oversight of care within LTC facilities [44].

Limitations and Strengths

This study was limited by a restricted sample size included in the primary outcomes analysis. Our study was profoundly affected by the COVID-19 pandemic response; public health measures to mitigate the spread of COVID-19 changed the nature of home care and expectations of family and friend caregivers in the middle of this study's sampling frame. This impacted the nature of the family caregiving relationships and the nature of support that caregivers were able to provide. The pandemic response also impacted what was considered normal standard of care for home care services due to resources being redirected to other sectors within the publicly funded health care systems in Ontario and Nova Scotia. Despite this, we did observe promising trends in the data related to PRM as a potentially supportive home care strategy among a cohort of older adults with significant and complex care needs. In addition, we acknowledge the unique opportunity to collect data via PRM systems (ie, information from the sensor notifications) regarding older adults' behaviors within the home setting and the ability to leverage PRM data into large datasets to be analyzed and used to better understand the home care needs of complex care older adults.

XSL•FO
RenderX

Perhaps a strength and limitation of the PRM model of home care for health care systems is the reliance on unpaid caregivers to receive and act on PRM system notifications and not the health care system. Alternatively, the PRM model of home care enhances unpaid caregivers' responsibility and accountability to the care of their family member.

There was also limited representation of marginalized communities within our study, which limits our understanding of the impact of PRM home care models among culturally diverse populations. Finally, the fact that most caregivers were spouses to the patient participants limited this study's understanding of how PRM technologies may impact caregivers in the sandwich generation (ie, those individuals caring for older adult parents and children dependents).

Future Research

There are several opportunities for future research. An economic analysis of the PRM model of care will provide greater insight regarding the implementation of the model of care among health system decision makers. Additional research is needed to fully explore concerns related to home care client, caregiver, and home care provider privacy. Research is needed into the contributions of the PRM model of home care intervention to resolve the challenges of limited health human resources within home care while balancing the health care (physical, emotional, and financial) services given to family caregivers. We also propose an opportunity for data science research in analyzing the aggregated data on older adults' behaviors to leverage historical tracking of data to support trend analysis and the development of best practices within the home care setting and, importantly, to generate a plan of home care services targeted to the unique needs of complex care older adults. A consideration of this PRM model of care against available DIY or off-the-shelf sensor systems requires investigation. The use of off-the-shelf systems has implications for enhanced digital skills with caregivers and older adults that is not required to the same extent with the PRM model of care in this study. Investigation into older adults' and caregiver digital health literacy skills will further inform uptake and use of the PRM model of home care. Research into system data visualization—dashboard components for older adult clients and caregivers—would also support informed decision-making and provide insight into user-friendly dashboards. Longitudinal studies with populations with different chronic diseases (eg, cancer and dementia) is also needed to determine PRM model impact.

Conclusions

The combination of growing waitlists for LTC, limited home care resources, and older adults' preference to remain in their home as long as possible creates the care context for PRM technologies. These technologies hold promise to support complex care older adults to safely remain in their home and avoid or delay admission to higher levels of care. Home is preferred by most older adults even if their health conditions challenge their ability to live independently; almost all Canadians aged ≥65 years reported their intention to take extraordinary measures to avoid institutionalization in an LTC setting. The PRM model of care in this study relied on unpaid caregivers, not the health care system, to receive and act on PRM notifications. While not statistically significant, the findings of this study demonstrate a trend in favor of PRM to support patients' aging in place and positive changes in health service use among older adult home care recipients. To appreciate the provincial differences reported in this study, the PRM model of care, as with any health care system intervention, must be considered within the larger health, social, and political context. Further study is required to understand if longer follow-up time allows more effects of PRM on patients' avoidance of higher levels of care to be detected.

Acknowledgments

This study was supported by the Institute for Clinical Evaluative Sciences, which is funded by an annual grant from the Ontario Ministry of Health and the Ministry of Long-Term Care. Portions of the data used in this report were made available by Health Data Nova Scotia of Dalhousie University. Although this research analysis is based on data obtained from the Nova Scotia Department of Health and Wellness, the observations and opinions expressed are those of the authors and do not represent those of either Health Data Nova Scotia or the Department of Health and Wellness. This study also received funding from Canadian Institutes of Health Research. This document used data adapted from the Statistics Canada Postal Code Conversion File, which is based on data licensed from Canada Post Corporation, and data adapted from the Ontario Ministry of Health Postal Code Conversion File, which contains data copied under license from Canada Post Corporation and Statistics Canada. The authors thank IQVIA Solutions Canada Inc for use of their Drug Information File. Parts of this material are based on data and information compiled and provided by the Canadian Institute for Health Information, Ontario Health, and the Ontario Ministry of Health. The analyses, conclusions, opinions, and statements expressed herein are solely those of the authors and do not reflect those of the funding or data sources; no endorsement is intended or should be inferred.

Authors' Contributions

The work for this manuscript was led by LD and BH. This study was conceptualized by LD (co-lead), SR (co-lead), GW, LW, and SS. Data curation was done by BH and JR. Formal analysis was completed by JR (lead), MR, and SS. Funding acquired by LD (co-lead), SR (co-lead), GW, LW, and SS. Investigation was coordinated by KL and BH. Data analysis methodology was done by JR, SS, LD, BH, MR, and GW. Project administration was done by KL, LD, GW, ER, and BH. Resources were secured by LD, SR, GW, and LW. Overall study supervision was provided by LD, SR, and BH. Results validation was done by LD, BH,

XSL•FO
RenderX

JR, MR, GW, ER, and SR. Visualization of manuscript was done by BH and LD. Writing of the original draft was done by LD, BH, GW, JR, SS, and MR. Reviewing and editing was done by LD (co-lead), BH (co-lead), GW, MR, JR, SS, ER, SR, LW, and KL.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 472 KB - aging_v8i1e69107_app1.pdf]

References

1. Turcotte M. Insights on Canadian society: Canadians with unmet home care needs. Statistics Canada. 2014 Sep 09. URL: https://www150.statcan.gc.ca/n1/en/pub/75-006-x/2014001/article/14042-eng.pdf?st=c76z_elg [accessed 2024-04-29]
2. National Institute on Ageing. Ageing in the right place: supporting older Canadians to live where they want. National Institute on Ageing. 2022. URL: https://static1.squarespace.com/static/5c2fa7b03917eed9b5a436d8/t/638e0857c959d1 [accessed 2024-08-01]
3. Pandemic perspectives on long-term care: insights from Canadians in light of COVID-19. National Institute on Ageing. 2021. URL: https://digitallibrary.cma.ca/link/digitallibrary5 [accessed 2022-08-01]
4. Seniors in need, caregivers in distress : what are the home care priorities for seniors in Canada? Health Council of Canada. 2012. URL: https://publications.gc.ca/site/eng/423913/publication.html [accessed 2024-04-29]
5. Brining care home. Expert Group on Home and Community Care. 2015. URL: https://www.onecaresupport.ca/Files/Bringing%20Care%20Home.pdf [accessed 2024-04-29]
6. Gilmour H. Unmet home care needs in Canada. Health Reports 2018 Nov 21;29(11):3-11 [FREE Full text] [Medline: 30485384]
7. MacDonald BJ, Wolfson M, Hirdes J. The future cost of long-term care in Canada. National Institute of Ageing. 2019 Oct. URL: https://tinyurl.com/yycuz3fd [accessed 2022-08-01]
8. Williams AP, Peckham A, Watkins J, Warrick N, Tam T, Rudoler D, et al. Caring for caregivers: facing up to tough challenges. Healthc Q 2014 Oct 30;17(3):20-23. [doi: 10.12927/hcq.2014.24015] [Medline: 25735060]
9. Blais R, Sears NA, Doran D, Baker GR, Macdonald M, Mitchell L, et al. Assessing adverse events among home care clients in three Canadian provinces using chart review. BMJ Qual Saf 2013 Dec 04;22(12):989-997 [FREE Full text] [doi: 10.1136/bmjqs-2013-002039] [Medline: 23828878]
10. Doran D, Blais R, Baker GR, Harrison M, Lang A, Macdonald M, et al. The safety at home study: an evidence base for policy and practice change. Healthc Q 2014 Oct 30;17(3):42-47. [doi: 10.12927/hcq.2014.24019] [Medline: 25591609]
11. Read E, Woolsey C, Donelle L, Weeks L, Chinho N. Passive remote monitoring and aging in place: a scoping review. Can J Aging 2023 Mar 01;42(1):20-32. [doi: 10.1017/S0714980822000198] [Medline: 35912590]
12. Tian YJ, Felber NA, Pageau F, Schwab DR, Wangmo T. Benefits and barriers associated with the use of smart home health technologies in the care of older persons: a systematic review. BMC Geriatr 2024 Feb 14;24(1):152 [FREE Full text] [doi: 10.1186/s12877-024-04702-1] [Medline: 38355464]
13. Swartz K, Collins L. Caregiver care. Am Fam Physician 2019 Jun 01;99(11):699-706 [FREE Full text] [Medline: 31150177]
14. Barnard D, Yaffe MJ. What is the physician's responsibility to a patient's family caregiver? Virtual Mentor 2014 May 01;16(5):330-338 [FREE Full text] [doi: 10.1001/virtualmentor.2014.16.05.ecas1-1405] [Medline: 24847702]
15. Atlantic Institute on Aging Carelink Advantage Pilot. Fredricton, NB: Orion Marketing Research; 2011.
16. Deis K, Gardner L, Ohinmaa A, Wanke M, Wilson P. Summative Evaluation of the Continuing Care Technology Innovation Pilot Project (CCTI). Edmonton, AB: Charis Management Consulting Inc; 2012.
17. Donelle L, Regan S, Kerr M, Zwarenstein M, Bauer M, Warner G, et al. Caring near and far by connecting community-based clients and family member/friend caregivers using passive remote monitoring: protocol for a pragmatic randomized controlled trial. JMIR Res Protoc 2020 Jan 10;9(1):e15027 [FREE Full text] [doi: 10.2196/15027] [Medline: 31922492]
18. Treweek S, Zwarenstein M. Making trials matter: pragmatic and explanatory trials and the problem of applicability. Trials 2009 Jun 03;10(1):37 [FREE Full text] [doi: 10.1186/1745-6215-10-37] [Medline: 19493350]
19. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008 Sep 29;337(sep29 1):a1655 [FREE Full text] [doi: 10.1136/bmj.a1655] [Medline: 18824488]
20. Husereau D, Jacobs P, Manns B, Hoomans T, Marshall D, Tamblyn R. Economic evaluation of complex health system interventions: a discussion paper. Institute of Health Economics. 2014. URL: https://www.ihe.ca/publications/economic-evaluation-of-complex-health-system-interventions-a-discussion-paper [accessed 2024-04-29]
21. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. BMJ 2015 May 08;350:h2147 [FREE Full text] [doi: 10.1136/bmj.h2147] [Medline: 25956159]

XSL•FO
RenderX

22. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods 2009 Nov;41(4):1149-1160. [doi: 10.3758/BRM.41.4.1149] [Medline: 19897823]
23. Kim J, Shin W. How to do random allocation (randomization). Clin Orthop Surg 2014 Mar;6(1):103-109 [FREE Full text] [doi: 10.4055/cios.2014.6.1.103] [Medline: 24605197]
24. Bowling A, Hankins M, Windle G, Bilotta C, Grant R. A short measure of quality of life in older age: the performance of the brief Older People's Quality of Life questionnaire (OPQOL-brief). Arch Gerontol Geriatr 2013 Jan;56(1):181-187 [FREE Full text] [doi: 10.1016/j.archger.2012.08.012] [Medline: 22999305]
25. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
26. Sager MA, Rudberg MA, Muhammad Jalaluddin MS, Franke T, Inouye SK, Landefeld CS, et al. Hospital admission risk profile (HARP): identifying older patients at risk for functional decline following acute medical illness and hospitalization. J Am Geriatr Soc 1996 Mar 27;44(3):251-257. [doi: 10.1111/j.1532-5415.1996.tb00910.x] [Medline: 8600192]
27. Selected caregiver assessment measures: a resource inventory for practitioners. Family Caregiver Alliance & Benjamin Rose Institute on Aging. 2012. URL: https://www.carealliance.ie/userfiles/file/SelCGAssmtMeas_ResInv_FINAL_12_10_12.pdf [accessed 2024-04-29]
28. California caregiver resource centers uniform assessment tool. California Caregiver Resource Centers. 2003. URL: https://www.caregiver.org/caregivers-count-too-s4-assessment-tool-examples [accessed 2024-04-29]
29. Bédard M, Molloy D, Squire L, Dubois S, Lever JA, O'Donnell M. The Zarit Burden Interview: a new short version and screening version. Gerontologist 2001 Oct;41(5):652-657. [doi: 10.1093/geront/41.5.652] [Medline: 11574710]
30. Tarlow BJ, Wisniewski SR, Belle SH, Rubert M, Ory MG, Gallagher-Thompson D. Positive aspects of caregiving: contributions of the REACH project to the development of new measures for Alzheimer's caregiving. Res Aging 2004 Jul 01;26(4):429-453. [doi: 10.1177/0164027504264493]
31. Koopman C, Pelletier KR, Murray JF, Sharda CE, Berger ML, Turpin RS, et al. Stanford presenteeism scale: health status and employee productivity. J Occup Environ Med 2002 Jan;44(1):14-20. [doi: 10.1097/00043764-200201000-00004] [Medline: 11802460]
32. Todtman K, Gustafson AW. The financial impact scale: an instrument for assessing long-term caregivers. J Gerontol Soc Work 2008 Oct 25;18(1-2):135-150. [doi: 10.1300/j083v18n03_10]
33. Faraone SV. Interpreting estimates of treatment effects: implications for managed care. P T 2008 Dec;33(12):700-711 [FREE Full text] [Medline: 19750051]
34. Gaugler JE, Zmora R, Mitchell LL, Finlay J, Rosebush CE, Nkimbeng M, et al. Remote activity monitoring for family caregivers of persons living with dementia: a mixed methods, randomized controlled evaluation. BMC Geriatr 2021 Dec 18;21(1):715 [FREE Full text] [doi: 10.1186/s12877-021-02634-8] [Medline: 34922475]
35. Finch M, Griffin K, Pacala JT. Reduced healthcare use and apparent savings with passive home monitoring technology: a pilot study. J Am Geriatr Soc 2017 Jun;65(6):1301-1305. [doi: 10.1111/jgs.14892] [Medline: 28407212]
36. 1 in 9 new long-term care residents potentially could have been cared for at home. Canadian Institute for Health Information. 2020. URL: https://www.cihi.ca/en/1-in-9-new-long-term-care-residents-potentially-could-have-been-cared-for-at-home [accessed 2022-08-01]
37. Seniors in transition: exploring pathways across the care continuum. Canadian Institute for Health Information. 2017. URL: https://www.cihi.ca/en/seniors-in-transition-exploring-pathways-across-the-care-continuum [accessed 2022-08-01]
38. Nikmat AW, Hawthorne G, Al-Mashoor SH. The comparison of quality of life among people with mild dementia in nursing home and home care--a preliminary report. Dementia (London) 2015 Jan 08;14(1):114-125. [doi: 10.1177/1471301213494509] [Medline: 24339093]
39. Olsen C, Pedersen I, Bergland A, Enders-Slegers MJ, Jøranson N, Calogiuri G, et al. Differences in quality of life in home-dwelling persons and nursing home residents with dementia - a cross-sectional study. BMC Geriatr 2016 Jul 11;16(1):137 [FREE Full text] [doi: 10.1186/s12877-016-0312-4] [Medline: 27400744]
40. Ramocha LM, Louw QA, Tshabalala MD. Quality of life and physical activity among older adults living in institutions compared to the community. S Afr J Physiother 2017;73(1):342 [FREE Full text] [doi: 10.4102/sajp.v73i1.342] [Medline: 30135899]
41. Mahmood A, Yamamoto T, Lee M, Steggell C. Perceptions and use of gerotechnology: implications for aging in place. J Hous Elderly 2008 Jun 16;22(1-2):104-126. [doi: 10.1080/02763890802097144]
42. Li K, Frumkin A, Bi WG, Magrill J, Newton C. Biopsy of Canada's family physician shortage. Fam Med Community Health 2023 May 12;11(2):e002236 [FREE Full text] [doi: 10.1136/fmch-2023-002236] [Medline: 37173094]
43. Cost estimate for Motion 77: improvements to longterm care. Office of the Parliamentary Budget Officer. URL: https://publications.gc.ca/site/eng/9.903146/publication.html [accessed 2022-08-01]
44. Warner G, Weeks LE, Chen YT, Hiebert B, Ledoux K, Donelle L. Key informant perceptions of challenges and facilitators to implementing passive remote monitoring technology for home care clients. Gerontechnology 2023;22(1):1-13 [FREE Full text] [doi: 10.4017/gt.2023.22.1.834.09]

XSL•FO
RenderX

Abbreviations

HDNS: Health Data Nova Scotia
HR: hazard ratio
ICES: Institute for Clinical Evaluative Sciences
LTC: long-term care
PRCT: pragmatic randomized controlled trial
PRM: passive remote monitoring

Edited by H Köttl; submitted 23.11.24; peer-reviewed by S Ofori, S Mukherjee; comments to author 22.12.24; revised version received 28.01.25; accepted 14.02.25; published 19.03.25.

Please cite as:
Donelle L, Hiebert B, Warner G, Reid M, Reid J, Shariff S, Richard E, Regan S, Weeks L, Ledoux K
Passive Remote Monitoring Technologies' Influence on Home Care Clients' Ability to Stay Home: Multiprovincial Randomized Controlled Trial
JMIR Aging 2025;8:e69107
URL: https://aging.jmir.org/2025/1/e69107
doi:10.2196/69107
PMID:

©Lorie Donelle, Bradley Hiebert, Grace Warner, Michael Reid, Jennifer Reid, Salimah Shariff, Emily Richard, Sandra Regan, Lori Weeks, Kathleen Ledoux. Originally published in JMIR Aging (https://aging.jmir.org), 19.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Artificial Intelligence–Enhanced Multi-Algorithm R Shiny Application for Predictive Modeling and Analytics: Case Study of Alzheimer Disease Diagnostics

Han Wenzheng[1*], MS; Edmund F Agyemang[1*], MS; Sudesh K Srivastav[1], PhD; Jeffrey G Shaffer[1], PhD; Samuel Kakraba[1,2*], PhD

[1]Department of Biostatistics and Data Science, Celia Scott Weatherhead School of Public Health and Tropical Medicine at Tulane University, 1440 Canal St, New Orleans, LA, United States

[2]Tulane Center for Aging, School of Medicine, Tulane University, New Orleans, LA, United States

[*]these authors contributed equally

Corresponding Author:
Samuel Kakraba, PhD
Department of Biostatistics and Data Science, Celia Scott Weatherhead School of Public Health and Tropical Medicine at Tulane University, 1440 Canal St, New Orleans, LA, United States

Abstract

Background: Artificial intelligence (AI) has demonstrated superior diagnostic accuracy compared with medical practitioners, highlighting its growing importance in health care. SMART-Pred (Shiny Multi-Algorithm R Tool for Predictive Modeling) is an innovative AI-based application for Alzheimer disease (AD) prediction using handwriting analysis.

Objective: This study aimed to develop and evaluate a noninvasive, cost-effective AI tool for early AD detection, addressing the need for accessible and accurate screening methods.

Methods: The study used principal component analysis for dimensionality reduction of handwriting data, followed by training and evaluation of 10 diverse AI models, including logistic regression, naïve Bayes, random forest, adaptive boosting, support vector machine, and neural network. Model performance was assessed using accuracy, sensitivity, precision, specificity, F_1-score, and area under the curve (AUC) metrics. The DARWIN (Diagnosis Alzheimer With Handwriting) dataset, comprising handwriting samples from 174 participants (89 patients with AD and 85 healthy controls), was used for validation and testing.

Results: The neural network classifier achieved an accuracy of 91% (95% CI 0.79 - 0.97) and an AUC of 94% on the test set after identifying the most significant features for AD prediction. These performance results surpass those of current clinical diagnostic tools, which typically achieve around 81% accuracy. SMART-Pred's performance aligns with recent AI advancements in AD prediction, such as Cambridge scientists' AI tool achieving 82% accuracy in identifying AD progression within 3 years, using cognitive tests and magnetic resonance imaging scans. The variables "air_time" and "paper_time" consistently emerged as critical predictors for AD across all 10 AI models, highlighting their potential importance in early detection and risk assessment. To augment transparency and interpretability, we incorporated the principles of explainable AI, specifically using Shapley Additive Explanations, a state-of-the-art method to emphasize the features responsible for our model's efficacy.

Conclusions: SMART-Pred offers noninvasive, cost-effective, and efficient AD prediction, demonstrating the transformative potential of AI in health care. While clinical validation is necessary to confirm the practical applicability of the identified key variables, the findings of this study contribute to the growing body of research on AI-assisted AD diagnosis and may lead to improved patient outcomes through early detection and intervention.

(JMIR Aging 2025;8:e70272) doi:10.2196/70272

KEYWORDS

artificial intelligence; AI; Shiny Multi-Algorithm R Tool for Predictive Modeling; SMART-Pred; machine learning; Alzheimer disease; predictive modeling; classification algorithms; disease diagnostics and surveillance

Introduction

Artificial intelligence (AI) is revolutionizing medical diagnosis, offering unprecedented accuracy and efficiency in health care [1]. AI algorithms can analyze vast amounts of complex medical data, including images, patient records, and genetic information, to detect diseases earlier and more precisely than traditional methods [1-3]. In some cases, AI outperforms human practitioners in diagnostic accuracy, particularly in fields like radiology and pathology [4-8]. For instance, AI systems can

XSL•FO
RenderX

identify subtle patterns in electrocardiograms to detect genetic heart conditions and analyze retinal images to diagnose diabetic retinopathy in its early stages [9-13]. By augmenting human expertise, AI not only improves diagnostic speed and accuracy but also has the potential to extend quality health care to remote areas, ultimately leading to better patient outcomes and more personalized treatment plans [1,3,4,14-17].

Alzheimer disease (AD) is a progressive brain disorder that gradually impairs memory and cognitive abilities, ultimately affecting the capacity to perform even the most basic tasks [18-22]. In most cases, especially those involving late-onset AD, symptoms usually appear in individuals during their mid-60s. Alzheimer disease and related dementias (ADRD) represent the primary cause of dementia among older adults [23-26]. In contrast, early-onset ADRD, which is relatively uncommon, can develop in people aged 30-65 years [27]. Aging is common to all cells and organelles in living organisms [28-31]. In recent times, there has been a disproportionate increase in the occurrence of ADRD, Huntington disease, Parkinson disease, amyotrophic lateral sclerosis, and other neurodegenerative diseases [32-35]. It is estimated that about 50 million people across the globe are living with ADRD [36,37]. ADRD are associated with a large family and caregiver burden. In 2018, the United States spent as much as US $277 billion on care for patients with ADRD. It is estimated that over 100 million people will be living with ADRD by the year 2050, for which the associated cost of treatment will be US $1200 billion [38-40]. Even though the prevalences of most diseases, such as breast cancer, prostate cancer, heart disease, stroke, and HIV, have decreased from 2000 to 2015, the prevalence of ADRD has increased by 123% within the same period [19].

Although many interventions, including novel nonsteroidal anti-inflammatory drugs and small molecules, have been explored for the treatment of ADRD [41-44], there is currently no cure for ADRD. Consequently, early diagnosis and the identification of ADRD biomarkers are critical to tailoring appropriate interventions for ADRD, reinforcing the need to design optimal methods for the prediction of ADRD and identify individuals at higher risk of ADRD in a timely manner [14,45-49].

The rapid growth of big data has necessitated the development of robust AI-driven algorithms for predictive modeling [50-58]. These algorithms are essential for efficiently handling large datasets and extracting relevant information through machine learning (ML) techniques [59]. Illustratively, ML algorithms, specifically hierarchical clustering, have been applied to visualize and examine the effects of point mutations that are associated with different phenotypes of diseases, like COVID-19, cystic fibrosis, and sickle cell anemia, resulting in an improved understanding of how point mutations lead to various variants of diseases [60-62]. AI and ML models have also found their place in disease prediction and risk analysis [54,63,64] for various diseases (neurodegenerative disorders, cardiovascular diseases, renal diseases, infectious diseases, diabetes, chronic diseases, neurological and psychiatric disorders, obesity, COVID-19, skin diseases, brain tumors, thyroid diseases, and liver diseases) [64-67]. However, no single algorithm can optimally solve every problem, highlighting the need for diverse ML approaches in predictive modeling and analytics. R, a widely adopted programming language for statistical data analysis, has become a powerful tool in data science [68]. The introduction of the Shiny package by RStudio revolutionized web development within the R ecosystem, making it accessible to both experienced data scientists and those with limited programming experience [69,70].

Methods

Development and Framework of the Shiny Multi-Algorithm R Tool for Predictive Modeling Application

SMART-Pred (Shiny Multi-Algorithm R Tool for Predictive Modeling) is an innovative AI-based R Shiny application developed to address the increasing demand for accessible data science tools and to simplify the complexity of ML workflows [67,71,72]. This user-friendly platform automates essential steps, such as data preprocessing, model training, hyperparameter tuning, model validation, and performance evaluation. It features a web interface that allows users to upload and preview datasets, configure variables, and compute descriptive statistics. SMART-Pred incorporates 10 classic ML algorithms, including logistic regression, random forest, adaptive boosting (AdaBoost), support vector machine (SVM), and neural network, for robust classification tasks. The tool also offers interactive data frames for comparing model evaluation metrics, alongside visualization capabilities like multimodel receiver operating characteristic (ROC) curve comparisons, feature importance ranking (by each selected algorithm saved as a CSV file in the same working directory in addition to the graphical display), and the Shapley Additive Explanations (SHAP) framework, which can provide the direction of influence of each feature. With one-click server deployment, SMART-Pred ensures accessibility anytime and anywhere, creating a convenient and intuitive environment for ML modeling that bridges the gap between complex AI algorithms and practical applications, potentially accelerating research and innovation across data-driven fields.

The DARWIN (Diagnosis Alzheimer With Handwriting) dataset [73] was used to validate SMART-Pred's classification capabilities and benchmark its performance against prior studies. Specifically curated for early AD detection, the DARWIN dataset comprises handwriting data from 174 participants (89 patients with AD and 85 healthy controls) across 452 columns, including unique identifiers, classification labels ("P" for patient and "H" for healthy), and 451 features extracted from 25 structured tasks (graphic, copy, memory, and dictation exercises) with 18 features per task. Collected using standardized protocols (eg, Wacom tablets at 200 Hz), this dataset captures fine motor and cognitive variations linked to neurodegeneration, offering a balanced sample to minimize algorithmic bias and richer multidimensional data compared to single-task datasets. Unlike neuroimaging-focused alternatives like Alzheimer's Disease Neuroimaging Initiative (ADNI), the DARWIN dataset supports noninvasive, cost-effective, and immediately deployable screening in clinical settings, aligning with SMART-Pred's goal of accessible early AD detection [73,74]. Its feature richness

XSL•FO
RenderX

and task diversity mitigate overfitting risks and provide direct insights into cognitive-motor integration, a hallmark of AD progression, while prioritizing actionable, low-resource biomarkers over complex biological data, thus enabling it to offer pragmatic solutions for early diagnosis in diverse health care environments. The SMART-Pred application and data are provided in Multimedia Appendix 1.

The ultimate objective of this study is to leverage handwriting characteristics from the DARWIN dataset to distinguish between patients with AD and healthy individuals, contributing to earlier, more accessible diagnostic methods that can inform tailored interventions. We hypothesized that the integration of user-configurable and automated ML methods within the SMART-Pred workflow will enable the reliable classification of patients with AD versus healthy individuals based on handwriting features from the DARWIN dataset, demonstrating the potential of handwriting-based digital biomarkers for early, scalable AD detection. SMART-Pred's end-to-end ML workflow for classification tasks is structured across 4 core components: data input and preview, data preprocessing, model training and validation, and model evaluation. This framework integrates user-configurable options with automated processing to ensure flexibility and methodological rigor. The process starts with data upload and preprocessing, advances through model training with multiple algorithmic choices, and culminates in comprehensive performance evaluation and visualization, as detailed in the flowchart presented in Figure 1. By democratizing access to advanced ML techniques through an intuitive interface, SMART-Pred aims to enhance research capabilities and clinical applicability in AD detection and beyond.

XSL•FO
RenderX

Figure 1. Schematic workflow of the SMART-Pred (Shiny Multi-Algorithm R Tool for Predictive Modeling) application. The comprehensive workflow of the SMART-Pred application is organized into four principal stages: (1) uploading and previewing the dataset, (2) preprocessing data, (3) model training and validation, and (4) model evaluation. Graphical elements indicate process types: dotted boxes represent user-configurable steps (eg, data selection, feature handling, and model settings), while solid boxes denote system-fixed operations. The color scheme clarifies each stage's nature: pink for data-related processes, green for machine learning operations, and blue for model outputs. This visual summary highlights both the modular structure and the flexible customization available within SMART-Pred for streamlined, transparent predictive analytics. AdaBoost: adaptive boosting; PCA: principal component analysis; ROC: receiver operating characteristic; SHAP: Shapley Additive Explanations; XGBoost: extreme gradient boosting.

XSL•FO
RenderX

SHAP-Based Interpretability Method

SHAP is a model-agnostic interpretability technique grounded in cooperative game theory, which quantifies how much each feature contributes to a model's prediction by computing Shapley values, ensuring consistent and locally accurate explanations for individual outcomes. In its general, model-agnostic form, SHAP approximates these values using a weighted linear regression approach (Kernel SHAP), though specialized algorithms exist for specific model types. SHAP is applied to interpret predictions from all 10 ML classifiers trained on AD data, revealing the relative impact of each input feature on classification outcomes [75,76]. The explanation model g approximates the original complex model f by representing the prediction as a linear combination of feature coalitions given by:

$$g(z') = \varphi_0 + \sum_{i=1}^{N} \varphi_i z_i', \quad z_i' \in \{0,1\}^N$$

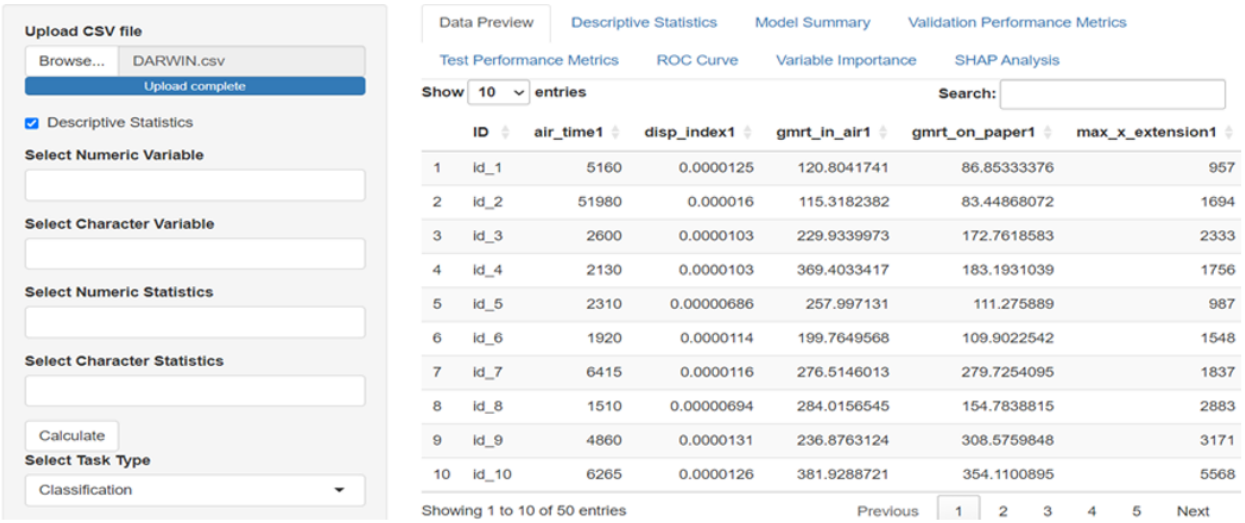
where z` denotes a binary vector indicating the presence or absence of features in the coalition, N is the total number of AD features, and φi corresponds to the Shapley value, reflecting the individual feature's contribution [77].

Through optimization, the Shapley value φi provides a theoretically grounded and consistent measure of feature importance, enhancing the interpretability of the ML classifier's predictions.

Data Input and Preview

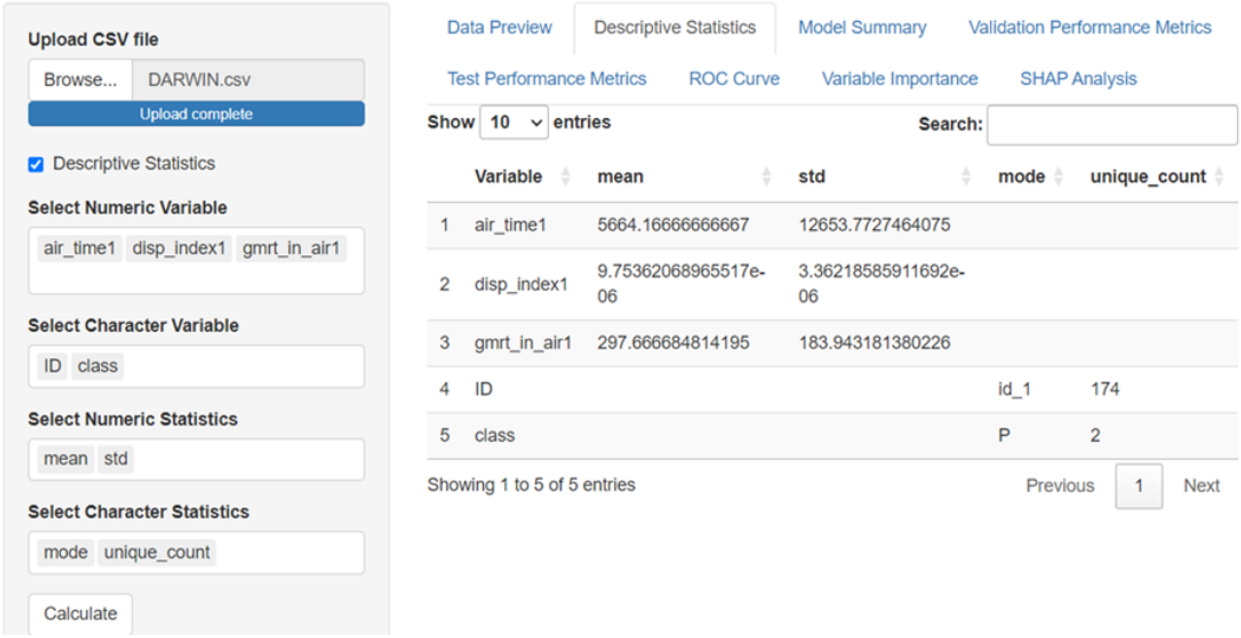
SMART-Pred allows users to upload local data files through a web interface. Upon uploading, users can preview the dataset in an interactive table format (Figure 2) and access basic descriptive statistics (Figure 3). For descriptive statistics, the interface provides options for variable and measure configuration, where users can choose specific statistical measures: mean, SD, minimum, median, and maximum for numerical variables, and mode and unique count for character variables. After data upload, users proceed to configure the analysis parameters through a structured variable selection process (Figure 4). First, users specify the analysis type by selecting either "Classification" or "Regression" from the task type dropdown menu. For the target variable selection, users choose the dependent variable that represents the outcome to be predicted. The interface then presents a comprehensive list of available independent variables in a scrollable selection box, including various measurement features and other kinematic parameters. Users can either select individual variables or use the "Select all independent variables" checkbox for comprehensive analysis. This structured approach ensures proper specification of the prediction task and relevant features for subsequent analysis. The interface also provides checkboxes for selecting from 10 ML models (Figure 5).

Figure 2. Data input and preview interface of SMART-Pred (Shiny Multi-Algorithm R Tool for Predictive Modeling). The left panel shows data input controls, including CSV file upload. The right panel displays an interactive data preview table with pagination controls (10 entries per page) and search functionality. The interface enables users to configure analytic parameters while maintaining an overview of the dataset structure. ROC: receiver operating characteristic; SHAP: Shapley Additive Explanations.



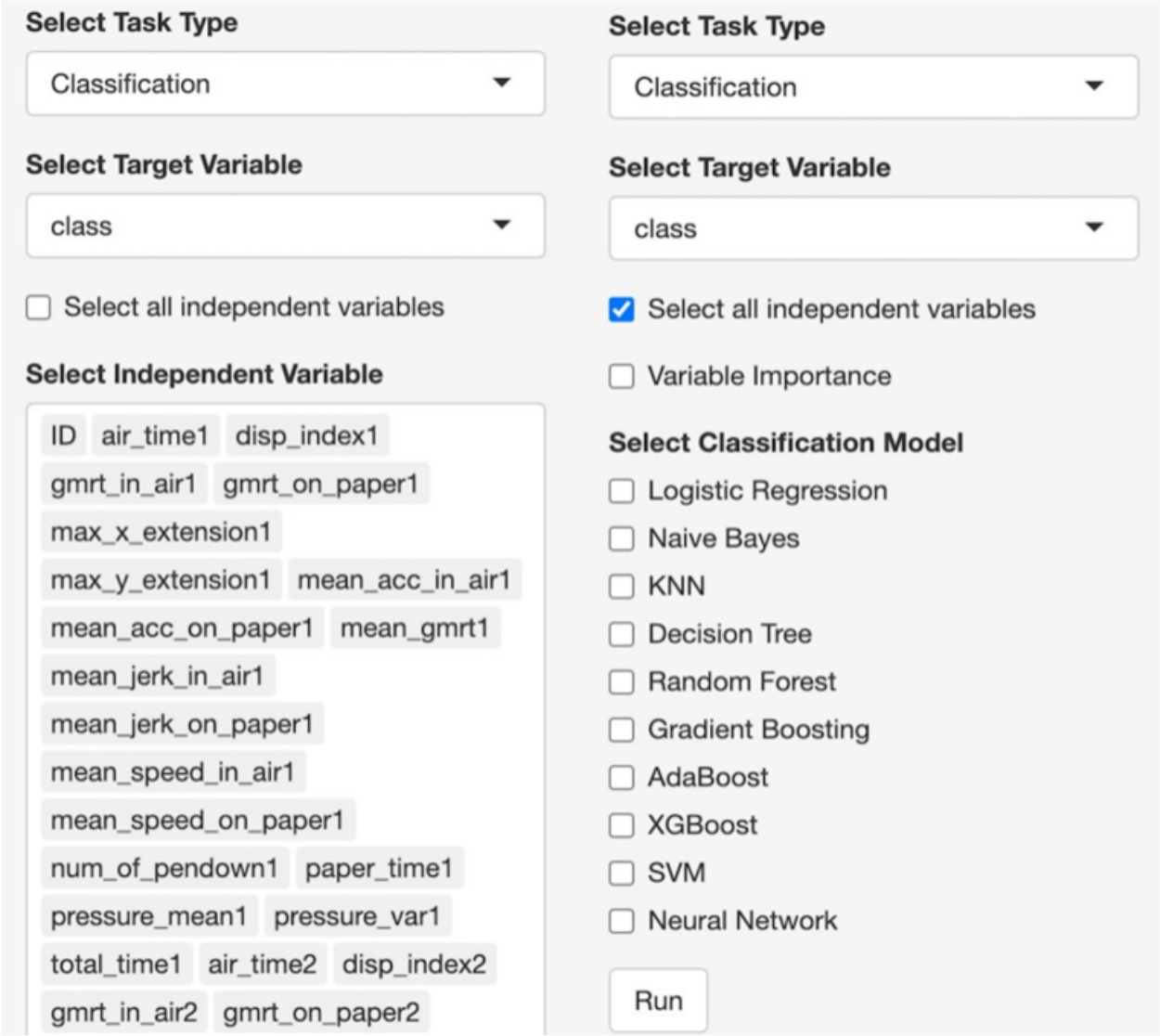
XSL•FO
RenderX

Figure 3. Descriptive statistics interface of SMART-Pred (Shiny Multi-Algorithm R Tool for Predictive Modeling). The left panel includes variable and statistics selection controls: separate boxes for numeric variables (eg, air_time1 and disp_index1) and character variables (eg, ID and class), along with corresponding statistics selection boxes (mean and std for numerical variables; mode and unique count for character variables) and a "Calculate" button to run the process. The right panel displays the calculated statistics in a paginated table format with sortable columns, showing descriptive measures for all selected variables. The interface includes navigation tabs at the top for accessing different analysis views and a search functionality for filtering results. ROC: receiver operating characteristic; SHAP: Shapley Additive Explanations.



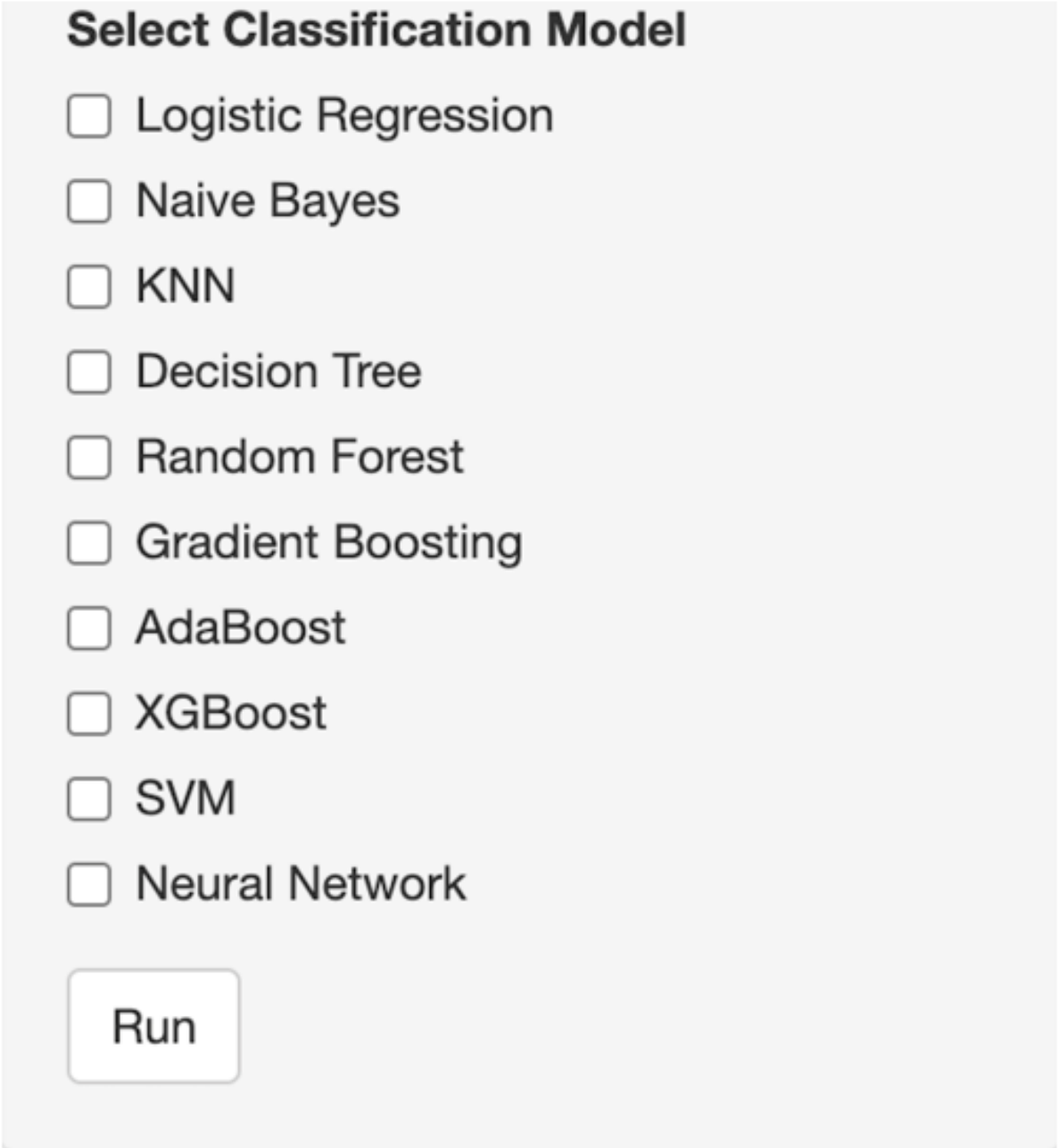
XSL•FO
RenderX

Figure 4. Task and variable selection interface of SMART-Pred (Shiny Multi-Algorithm R Tool for Predictive Modeling). The interface consists of 3 main components: a task type dropdown menu for selecting between classification and regression analysis, a target variable dropdown menu for specifying the dependent variable (eg, “class”), and a scrollable selection box displaying all available independent variables. Users can either select individual variables or use the “Select all independent variables” checkbox for batch selection. The interface displays variable names in a structured format, allowing users to view and select from all the features (eg, air_time1, disp_index1, and gmrt_in_air1). AdaBoost: adaptive boosting; KNN: k-nearest neighbors; SVM: support vector machine; XGBoost: extreme gradient boosting.



XSL•FO
RenderX

Figure 5. Model selection panel of SMART-Pred (Shiny Multi-Algorithm R Tool for Predictive Modeling). The interface provides checkboxes for selecting from 10 machine learning models: logistic regression, naïve Bayes, k-nearest neighbors (KNN), decision tree, random forest, gradient boosting, adaptive boosting (AdaBoost), extreme gradient boosting (XGBoost), support vector machine (SVM), and neural network. A "variable importance" option enables feature significance analysis, a "SHAP analysis" option estimates Shapley Additive Explanations (SHAP) values to emphasize the direction of the features responsible for the model's efficacy, and a "Run" button initiates the model training process. The panel allows multiple model selection for comparative analysis.



Data Preprocessing

The preprocessing pipeline (Figure 1) implements a systematic approach to data preparation. The preprocessing pipeline in SMART-Pred begins with a stratified 70 - 30 split to preserve class balance between patients with AD and controls, ensuring representative distributions in the training (70%) and test (30%) sets. For missing values, mean/mode imputation is systematically applied: numerical features use subgroup-specific means (eg, imputing based on AD stage or age cohort) to retain biological context, while categorical variables use mode replacement or an "Unknown" category to avoid information loss. This approach balances simplicity with clinical relevance, as arbitrary deletion would disproportionately affect smaller AD subgroups.

Principal component analysis (PCA) was strategically implemented to address the DARWIN dataset's high dimensionality (451 features), focusing on spatiotemporal handwriting variables like *air_time* and *paper_time*. PCA condensed these into orthogonal components that captured 95% cumulative variance, eliminating redundant kinematic features (eg, overlapping stroke velocity metrics) while preserving

XSL•FO
RenderX

discriminative power for AD-related motor-cognitive patterns. The threshold of 95% variance retention was chosen to minimize noise from minor handwriting variations unrelated to neurodegeneration. Prior to PCA, feature filtering removed near-zero variance metrics (<5% unique values) and highly correlated variables (Pearson r>0.9), reducing multicollinearity that could distort component interpretation.

For non-DARWIN datasets with missing values, SMART-Pred uses context-aware imputation: numerical features use median values stratified by diagnostic status to avoid dilution of AD-specific patterns, while categorical variables adopt a hybrid approach-mode replacement for features with <10% missingness and "Missing" flagging for higher rates. This dual strategy ensures robustness across heterogeneous data sources while maintaining compatibility with the tool's multi-algorithm framework. The preprocessing workflow optimizes computational efficiency without sacrificing clinical interpretability, enabling focused analysis of neurodegeneration-sensitive handwriting biomarkers.

Model Training and Validation

The model training process uses a robust validation strategy involving 10-fold cross-validation repeated 5 times. The system has 10 classic ML algorithms (Table 1). The models range from traditional approaches (eg, logistic regression and naïve Bayes) to advanced ensemble methods (eg, random forest and gradient boosting) and deep learning (eg, neural network), and the selection has been driven by the need to balance algorithmic diversity, clinical utility, and technical robustness. These models have been chosen to cover a broader spectrum of approaches: linear models (eg, logistic regression) provide interpretability for clinical stakeholders, ensemble methods (eg, random forest and AdaBoost) address nonlinear relationships while reducing overfitting, and deep learning models (eg, neural network) capture complex interactions. Criteria, such as accuracy, computational efficiency, and explainability, have guided the selection. Domain-specific validation ensures that the models excel at analyzing temporal handwriting dynamics and mitigates class imbalance. The inclusion of both high-precision neural networks and interpretable models optimizes the tradeoff between complexity and clinical transparency, while cross-validation confirms consistent performance. This multi-algorithm strategy ensures adaptability to diverse patient profiles and aligns with SMART-Pred's goal of delivering a robust, deployable tool for early AD detection. Detailed information about each algorithm's implementation and theoretical foundation has been provided previously [25-33]. Hyperparameter optimization is performed through a grid search process to identify the best model configuration.

Table . Machine learning models implemented in the SMART-Pred[a] application.

Algorithm	R function	Reference
Logistic regression	glm	[78]
Naïve Bayes	Naïve_bayes	[79]
K-nearest neighbor	knn	[80]
Decision tree	rpart	[81]
Random forest	rf	[82]
Gradient boosting	gbm	[83]
Adaptive boosting	ada	[84]
Extreme gradient boosting	xgbTree	[85]
Support vector machine	svmRadial	[80]
Neural network	nnet	[86]

[a]SMART-Pred: Shiny Multi-Algorithm R Tool for Predictive Modeling.

Evaluation Output

The evaluation module provides a comprehensive assessment through multiple metrics and visualizations. The system generates validation model performance metrics (Figures 6 and 7), including accuracy [87], sensitivity [88], specificity [88], and F_1-score [89], for model selection; test model performance metrics (Figure 8), including accuracy [34], sensitivity [35], specificity [35], F_1-score [36], and ROC [37], for final evaluation; and ROC curves (Figure 9) for comparing model performance. Additionally, variable importance analysis (Figure 10) is conducted to identify key predictive features across different models. Lastly, SHAP analysis is conducted to interpret the predictions made by the 10 ML classifiers. All evaluation results are presented through interactive visualizations and data frames, allowing the examination of model performance from multiple perspectives.

XSL•FO
RenderX

Figure 6. Model summary interface of SMART-Pred (Shiny Multi-Algorithm R Tool for Predictive Modeling) displaying detailed model training results. The interface presents comprehensive information for each trained model and an error message if a problem occurs during model training, including model type (eg, generalized linear model and naïve Bayes), dataset characteristics (eg, 121 samples and 451 predictors), preprocessing steps (eg, impute_mean, impute_mode, nzv, corr, center, scale, and pca), and cross-validation details (10-fold, repeated 5 times). Performance metrics are displayed in a structured format, showing receiver operating characteristic (ROC), sensitivity, and specificity values for each model's cross-validation results. SHAP: Shapley Additive Explanations.

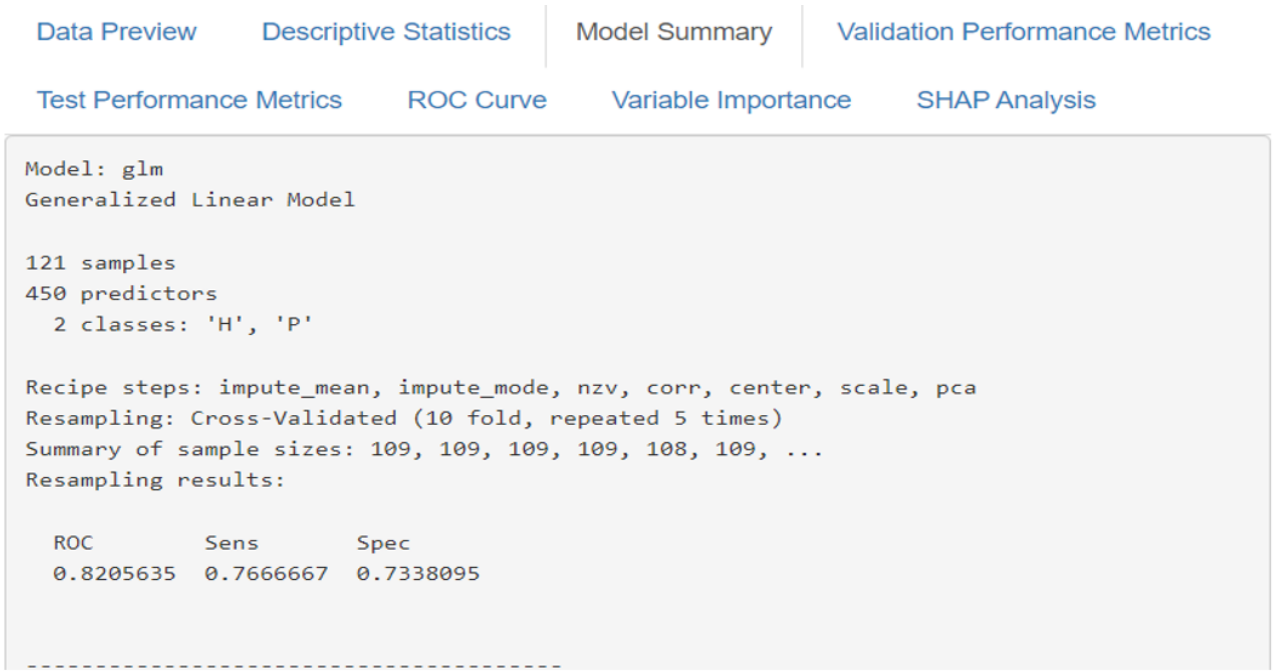
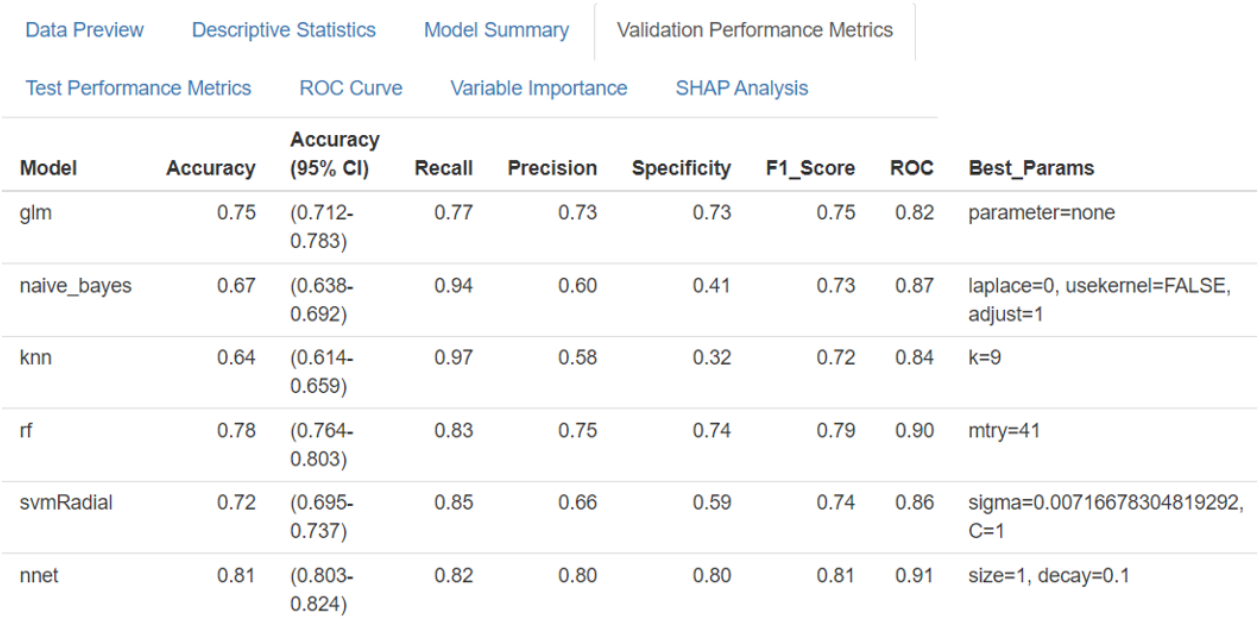


Figure 7. Validation performance metrics comparison interface of SMART-Pred (Shiny Multi-Algorithm R Tool for Predictive Modeling). The table displays comprehensive evaluation metrics for 6 selected machine learning models (logistic regression, naïve Bayes, k-nearest neighbors, random forest, support vector machine, and neural network), including accuracy, recall, precision, specificity, F_1-score, and receiver operating characteristic (ROC) value. Each model's optimized hyperparameters are shown in the Best_Params column. The interface is accessible through the "Validation Performance Metrics" tab, allowing users to compare model performance across multiple evaluation criteria. SHAP: Shapley Additive Explanations.



Model	Accuracy	Accuracy (95% CI)	Recall	Precision	Specificity	F1_Score	ROC	Best_Params
glm	0.75	(0.712-0.783)	0.77	0.73	0.73	0.75	0.82	parameter=none
naive_bayes	0.67	(0.638-0.692)	0.94	0.60	0.41	0.73	0.87	laplace=0, usekernel=FALSE, adjust=1
knn	0.64	(0.614-0.659)	0.97	0.58	0.32	0.72	0.84	k=9
rf	0.78	(0.764-0.803)	0.83	0.75	0.74	0.79	0.90	mtry=41
svmRadial	0.72	(0.695-0.737)	0.85	0.66	0.59	0.74	0.86	sigma=0.00716678304819292, C=1
nnet	0.81	(0.803-0.824)	0.82	0.80	0.80	0.81	0.91	size=1, decay=0.1

XSL•FO
RenderX

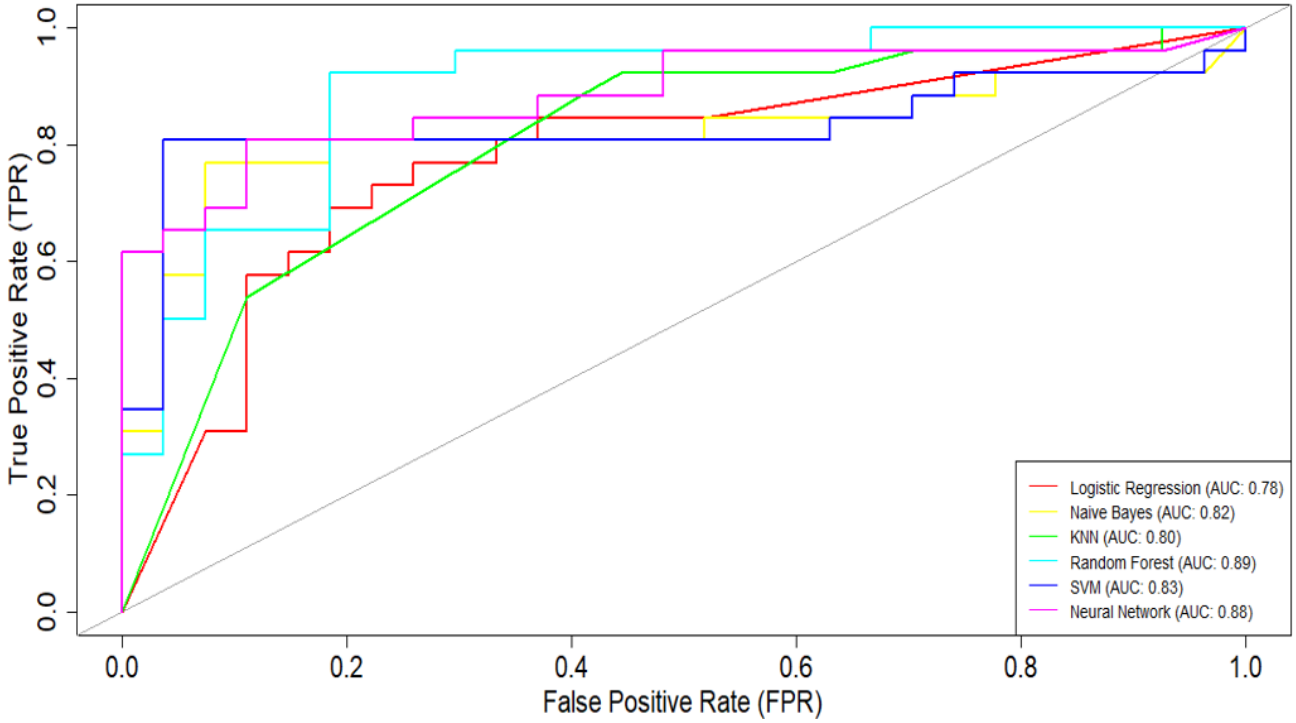
Figure 8. Test performance metrics comparison interface of SMART-Pred (Shiny Multi-Algorithm R Tool for Predictive Modeling). The table displays comprehensive evaluation metrics for 6 machine learning models (logistic regression, naïve Bayes, k-nearest neighbors, random forest, support vector machine, and neural network), including accuracy, recall, precision, specificity, F_1-score, and receiver operating characteristic (ROC) values. Each model's optimized hyperparameters are shown in the Best_Params column. The interface is accessible through the "Test Performance Metrics" tab, allowing users to compare model performance across multiple evaluation criteria. SHAP: Shapley Additive Explanations.

Data Preview | Descriptive Statistics | Model Summary | Validation Performance Metrics

Test Performance Metrics | ROC Curve | Variable Importance | SHAP Analysis

Model	Accuracy	Accuracy (95% CI)	Recall	Precision	Specificity	F1_Score	ROC	Best_Params
glm	0.74	(0.597-0.847)	0.65	0.77	0.81	0.71	0.78	parameter=none
naive_bayes	0.64	(0.498-0.769)	0.85	0.59	0.44	0.70	0.82	laplace=0, usekernel=FALSE, adjust=1
knn	0.62	(0.479-0.752)	0.96	0.57	0.30	0.71	0.80	k=9
rf	0.79	(0.659-0.892)	0.77	0.80	0.81	0.78	0.89	mtry=41
svmRadial	0.83	(0.702-0.919)	0.81	0.84	0.85	0.82	0.83	sigma=0.00716678304819292, C=1
nnet	0.79	(0.659-0.892)	0.69	0.86	0.89	0.77	0.88	size=1, decay=0.1

Figure 9. Receiver operating characteristic (ROC) curve comparison interface of SMART-Pred (Shiny Multi-Algorithm R Tool for Predictive Modeling). The plot displays the true positive rate against the false positive rate for 6 classification models, with each model represented by a different colored line. The diagonal reference line represents random prediction (area under the curve [AUC]=0.5). The interface is accessible through the "ROC Curve" tab, allowing users to compare model performance across multiple evaluation criteria. KNN: k-nearest neighbors; SVM: support vector machine.



XSL•FO
RenderX

Figure 10. Top 10 variable importance visualization interface in SMART-Pred (Shiny Multi-Algorithm R Tool for Predictive Modeling). The display consists of panels representing different models (eg, logistic regression, naïve Bayes, k-nearest neighbors [KNN], random forest, support vector machine [SVM], and neural network), with each showing bar charts of variable importance scores (y-axis, 0-100 scale) for the most influential predictors (x-axis). Each model type is distinguished by a unique color scheme, allowing for cross-model comparison of feature importance patterns. The interface is accessible through the "Variable Importance" tab, allowing users to compare variable importance across multiple evaluation criteria. ROC: receiver operating characteristic; SHAP: Shapley Additive Explanations.



Ethical Considerations

This study used secondary, deidentified data and therefore did not require Institutional Review Board approval. All data used in this study have been anonymized to protect privacy.

Results

Preliminary Findings

In SMART-Pred, the validation results (Table 2) demonstrated the diverse array of models analyzed, revealing significant variations in performance metrics, such as accuracy, sensitivity, specificity, F_1-score, and area under the curve (AUC), for the DARWIN dataset among the models. The neural network and AdaBoost classifiers emerged as top performers, exhibiting high accuracy scores of 81% and 80%, respectively. Moreover, AdaBoost and neural network achieved commendable AUC scores (92% and 91%, respectively), suggesting a robust capability to distinguish between positive and negative AD cases. Conversely, models, such as k-nearest neighbors (KNN) and naïve Bayes, were noted for lower overall performance, revealing potential limitations in handling the DARWIN dataset's intricacies.

XSL•FO
RenderX

Table . Validation performance metric comparison across the 10 machine learning models in the SMART-Pred[a] application for the full feature model.

Model name	Accuracy, value (95% CI)	Recall	Precision	Specificity	F_1-score	AUC[b]
Logistic regression	0.75 (0.71 - 0.78)	0.77	0.73	0.73	0.75	0.82
Naïve Bayes	0.67 (0.64 - 0.69)	0.94	0.60	0.41	0.73	0.87
K-nearest neighbors	0.64 (0.61 - 0.66)	0.97	0.58	0.32	0.72	0.84
Decision tree	0.69 (0.67 - 0.71)	0.58	0.73	0.79	0.65	0.80
Random forest	0.78 (0.76 - 0.80)	0.83	0.75	0.74	0.79	0.90
Gradient boosting	0.80 (0.79 - 0.82)	0.81	0.80	0.80	0.80	0.91
Adaptive boosting	0.80 (0.79 - 0.81)	0.83	0.78	0.77	0.80	0.92
Extreme gradient boosting	0.80 (0.79 - 0.80)	0.80	0.79	0.79	0.79	0.91
Support vector machine	0.72 (0.70 - 0.74)	0.85	0.66	0.59	0.74	0.86
Neural network	0.81 (0.80 - 0.82)	0.82	0.80	0.80	0.81	0.91

[a]SMART-Pred: Shiny Multi-Algorithm R Tool for Predictive Modeling.

[b]AUC: area under the curve.

Full Model With All Features

When the model was used for prediction in the test set (Table 3), the AdaBoost classifier stood out with a remarkable accuracy of 85%, a balanced sensitivity of 88%, an F_1-score of 85%, and an AUC of 92%, illustrating its ability to generalize well beyond the validation dataset. Such consistency between the validation and test phases highlights the robustness of the AdaBoost model for practical applications. The extreme gradient boosting (XGBoost) classifier also demonstrated reliable performance across the dataset, maintaining stable accuracy and specificity, further reinforcing its candidacy for real-world deployment. Many other algorithms in our pipeline, such as the gradient boosting and random forest classifiers, achieved commendable evaluation scores across the evaluation metrics, reaffirming the robustness of the pipeline for classification-based predictive modeling.

Table . Test performance metric comparison across the 10 machine learning models in the SMART-Pred[a] application for the full feature model.

Model name	Accuracy, value (95% CI)	Recall	Precision	Specificity	F_1-score	AUC[b]
Logistic regression	0.74 (0.60 - 0.85)	0.65	0.77	0.81	0.71	0.78
Naïve bayes	0.64 (0.50 - 0.77)	0.85	0.59	0.44	0.70	0.82
K-nearest neighbors	0.62 (0.48 - 0.75)	0.96	0.57	0.30	0.71	0.80
Decision tree	0.72 (0.58 - 0.83)	0.77	0.69	0.67	0.73	0.73
Random forest	0.79 (0.66 - 0.89)	0.77	0.80	0.81	0.78	0.89
Gradient boosting	0.85 (0.72 - 0.93)	0.81	0.88	0.89	0.84	0.89
Adaptive boosting	0.85 (0.72 - 0.93)	0.88	0.82	0.81	0.85	0.92
Extreme gradient boosting	0.81 (0.68 - 0.91)	0.73	0.86	0.89	0.79	0.88
Support vector machine	0.83 (0.70 - 0.92)	0.81	0.84	0.85	0.82	0.83
Neural network	0.79 (0.66 - 0.89)	0.69	0.86	0.89	0.77	0.88

[a]SMART-Pred: Shiny Multi-Algorithm R Tool for Predictive Modeling.

[b]AUC: area under the curve.

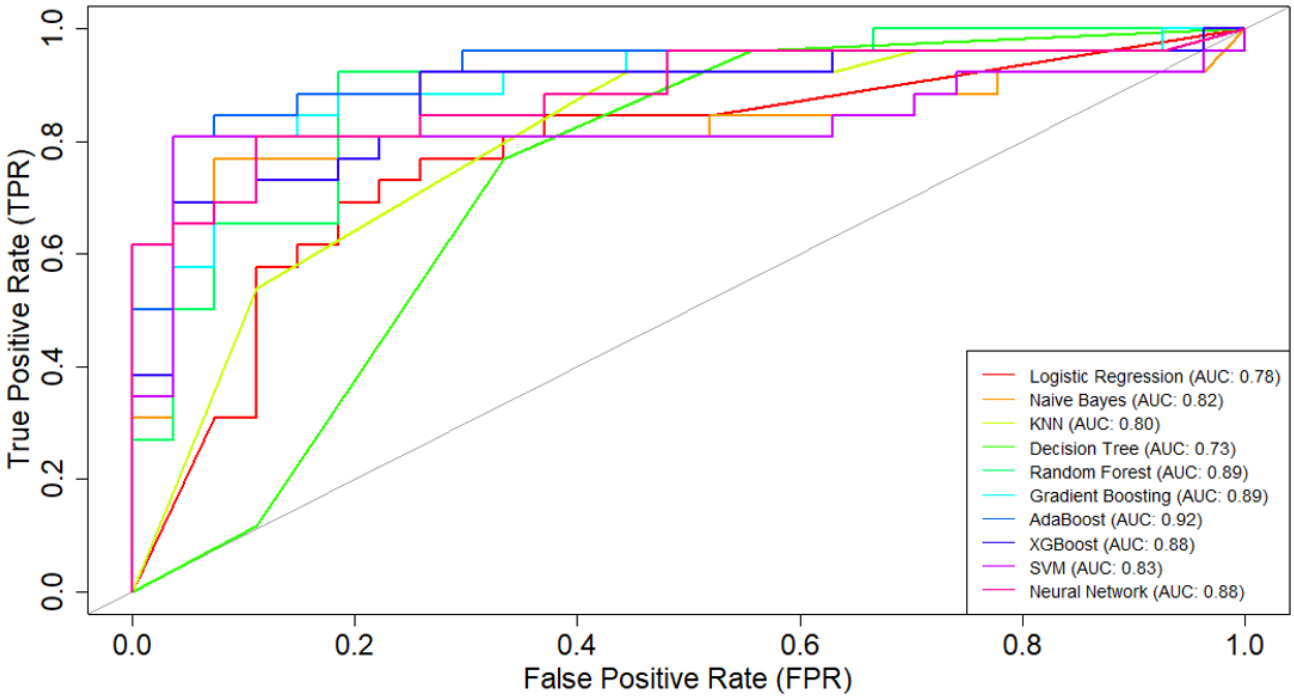
XSL•FO
RenderX

On integrating findings from both evaluation phases, it becomes evident that ensemble models (random forest, gradient boosting, AdaBoost, and XGBoost) hold significant promise due to their balanced performance metrics and strong generalization capabilities [90-94]. These models provide a compelling combination of sensitivity and specificity, which is crucial for the diagnostics required in AD detection. The observed fluctuations in the performance of the neural network classifier from validation to testing (especially for recall) suggest that while it possesses inherent strengths, there is room for enhancement.

Our comparative analysis underscores ensemble models (random forest, gradient boosting, AdaBoost, and XGBoost) as the most robust candidates for AD diagnostics in the scope of this work, given their superior performance across both the validation and test datasets [95-98]. These ensemble models have good, balanced accuracy in this regard, making them particularly suited for clinical settings where precise prediction and identification are critical. The insights garnered from this study pave the way for continued exploration into model optimization, potentially enhancing the applicability and accuracy of ML techniques in terms of advancing AD research and the early diagnosis of other diseases, which can be considered crucial in tailoring interventions to target diseases.

Based on the ROC curves (Figure 11), the ensemble models (AdaBoost: AUC=0.92; random forest: AUC=0.89; gradient boosting: AUC=0.89; XGBoost: AUC=0.88) had the strongest performance among the 10 ML models considered in this study. The findings suggest that the AdaBoost ensemble model is highly effective in correctly identifying positive and negative instances for the given classification task. AdaBoost was closely followed by random forest and gradient boosting, both of which demonstrated impressive AUC values of 0.89. The XGBoost and neural network classifiers both achieved an impressive AUC value of 0.88.

Figure 11. Receiver operating characteristic (ROC) curves across different machine learning (ML) models obtained from SMART-Pred (Shiny Multi-Algorithm R Tool for Predictive Modeling). The plot illustrates the relationship between the true positive rate (y-axis) and false positive rate (x-axis) across 10 different ML models. Each colored curve represents a distinct model's performance, with the area under the curve (AUC) values shown in the legend. The diagonal gray line represents the random classifier baseline (AUC=0.5). The closer the ROC curve is to the upper left corner of the graph, the higher is the accuracy of the test, as in the upper left corner, sensitivity is 1 and the false positive rate is 0 (specificity=1). The ideal ROC curve thus has an AUC of 1.0. AdaBoost: adaptive boosting; KNN: k-nearest neighbors; SVM: support vector machine; XGBoost: extreme gradient boosting.



In contrast, the decision tree model exhibited the weakest performance, with an AUC of 0.73. While this is still better than random guessing, it suggests that the decision tree algorithm may not be the most suitable choice. The remaining models, including logistic regression, naïve Bayes, SVM, and KNN, had AUC values ranging from 0.78 to 0.83, indicating reasonably good predictive capabilities. The intersecting nature of the ROC curves highlights the better tradeoffs between different models [99-102]. Certain algorithms may excel at specific levels of the false positive rate, which can be an important consideration when selecting the most appropriate model for the task at hand [63].

Building upon our variable importance analysis of SMART-Pred (Figure 12), the partial variable set was constructed by selecting the most influential features among all the algorithms identified through our importance ranking analysis, including "air_time" and "paper_time" variables. Results for the reduced feature model analysis via SMART-Pred are shown in Tables 4 and 5.

XSL•FO
RenderX

Table . Validation performance metric comparison across the 10 machine learning models in the SMART-Pred[a] application for the reduced feature model.

Model name	Accuracy, value (95% CI)	Recall	Precision	Specificity	F_1-score	AUC[b]
Logistic regression	0.85 (0.82 - 0.88)	0.85	0.85	0.86	0.85	0.89
Naïve Bayes	0.85 (0.83 - 0.87)	0.87	0.83	0.84	0.85	0.93
K-nearest neighbors	0.83 (0.81 - 0.84)	0.95	0.76	0.71	0.84	0.93
Decision tree	0.82 (0.80 - 0.84)	0.77	0.85	0.87	0.81	0.89
Random forest	0.88 (0.87 - 0.90)	0.88	0.89	0.89	0.88	0.95
Gradient boosting	0.88 (0.87 - 0.89)	0.88	0.88	0.88	0.88	0.94
Adaptive boosting	0.89 (0.88 - 0.90)	0.89	0.88	0.88	0.88	0.96
Extreme gradient boosting	0.87 (0.87 - 0.88)	0.87	0.87	0.88	0.87	0.94
Support vector machine	0.86 (0.85 - 0.88)	0.85	0.87	0.87	0.86	0.93
Neural network	0.86 (0.85 - 0.87)	0.87	0.85	0.85	0.86	0.95

[a]SMART-Pred: Shiny Multi-Algorithm R Tool for Predictive Modeling.

[b]AUC: area under the curve.

Table . Test performance metric comparison across the 10 machine learning models in the SMART-Pred[a] application for the reduced feature model.

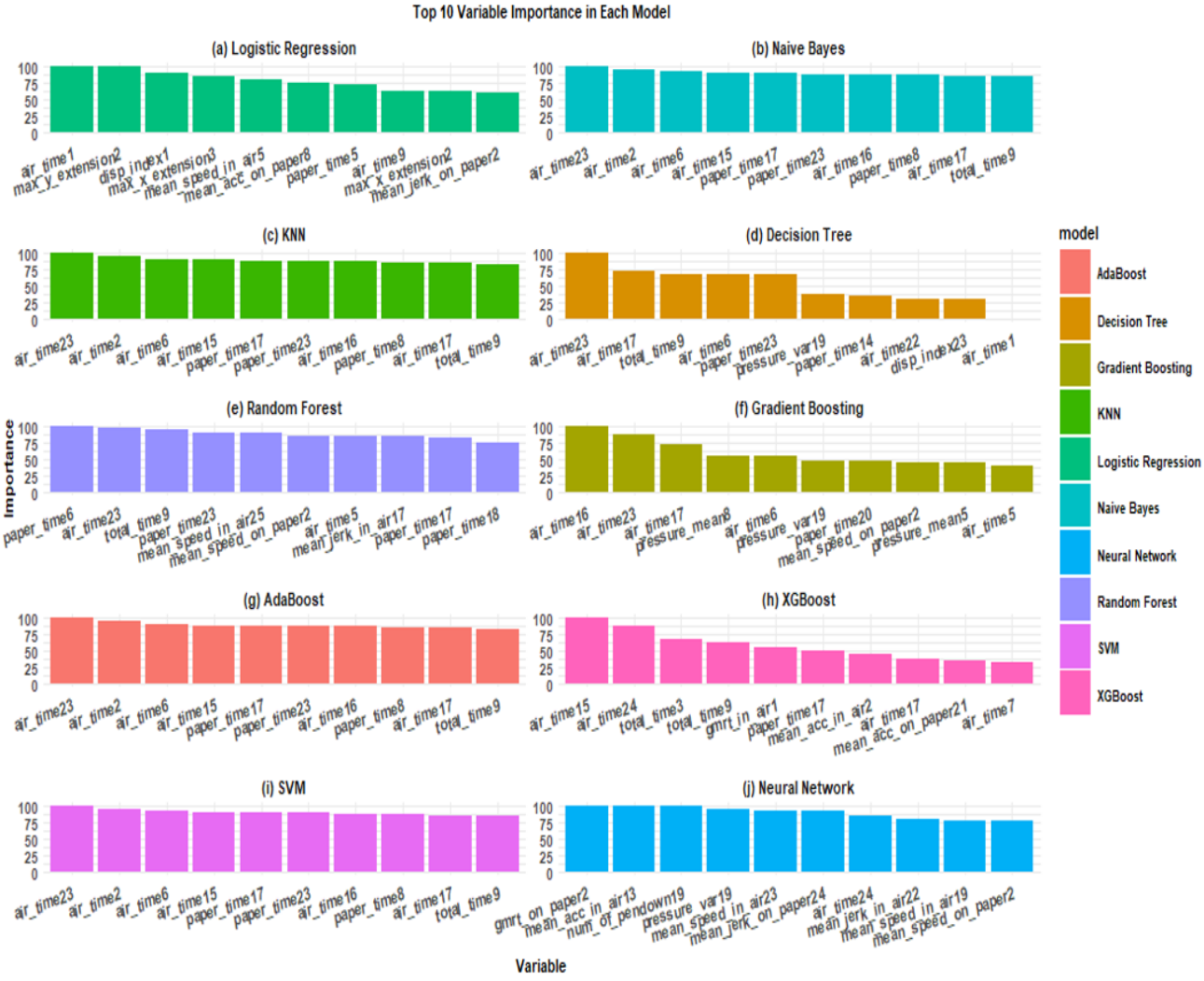
Model name	Accuracy, value (95% CI)	Recall	Precision	Specificity	F_1-score	AUC[b]
Logistic regression	0.87 (0.74 - 0.95)	0.88	0.85	0.85	0.87	0.88
Naïve Bayes	0.81 (0.68 - 0.91)	0.85	0.79	0.78	0.81	0.84
K-nearest neighbors	0.81 (0.68 - 0.91)	0.92	0.75	0.70	0.83	0.90
Decision tree	0.81 (0.68 - 0.91)	0.88	0.77	0.74	0.82	0.81
Random forest	0.87 (0.75 - 0.95)	0.85	0.88	0.89	0.86	0.91
Gradient boosting	0.87 (0.75 - 0.95)	0.85	0.88	0.89	0.86	0.90
Adaptive boosting	0.85 (0.72 - 0.93)	0.81	0.88	0.89	0.84	0.91
Extreme gradient boosting	0.85 (0.72 - 0.93)	0.81	0.88	0.89	0.84	0.87
Support vector machine	0.89 (0.77 - 0.96)	0.85	0.92	0.93	0.88	0.92
Neural network	0.91 (0.79 - 0.97)	0.88	0.92	0.93	0.90	0.94

[a]SMART-Pred: Shiny Multi-Algorithm R Tool for Predictive Modeling.

[b]AUC: area under the curve.

XSL•FO
RenderX

Figure 12. The top 10 variable importance findings across different models obtained from SMART-Pred (Shiny Multi-Algorithm R Tool for Predictive Modeling). The subplots represent the following models: logistic regression, a basic classification algorithm; naïve Bayes, a probabilistic classifier based on Bayes' theorem; k-nearest neighbors (KNN), an instance-based learning algorithm; decision tree, a tree-structured classification model; random forest, an ensemble learning method based on multiple decision trees; gradient boosting, a boosting-based ensemble method; adaptive boosting (AdaBoost), an adaptive boosting algorithm; extreme gradient boosting (XGBoost), an efficient implementation of gradient boosting; support vector machine (SVM), a kernel-based classification algorithm; and neural network, a deep learning architecture. Each subplot represents a specific model's top 10 most important features, with normalized importance scores (y-axis, 0-100) plotted against feature names (x-axis). Different colors distinguish between models, and features are arranged in descending order of importance within each subplot. The importance scores are normalized within each model, with the most important feature scaled to 100.



Reduced Model With Important Features

The comparative analysis of the reduced feature model and full feature model showed that models trained on the selected partial variables (reduced feature model) consistently demonstrated equal or superior performance compared to those trained on all variables (full feature model), as seen in Tables 4 and 5. The comparison between the full and partial variable sets (reduced feature model) revealed patterns in both validation (Figure 13) and test (Figure 14) metrics.

XSL•FO
RenderX

Figure 13. Comparison of validation metrics between the full feature model (all variables included) and the reduced feature model (only the top 2 most important variables across all models in the full feature model) across machine learning models. Each subplot represents a different model, with 6 standardized performance metrics (accuracy, recall, precision, specificity, F_1-score, and area under the curve [AUC]) on the x-axis and their corresponding values (0-1.0) on the y-axis. Red bars represent models trained with all variables, while blue bars indicate models trained with the top 2 most important variables ("air_time" and "paper_time"). This visualization enables a direct comparison of model performance between the full and reduced feature sets across multiple evaluation metrics. AdaBoost: adaptive boosting; KNN: k-nearest neighbors; SVM: support vector machine; XGBoost: extreme gradient boosting.



XSL•FO
RenderX

Figure 14. Comparison of test metrics between the full feature model (all variables included) and the reduced feature model (only the top 2 most important variables across all models in the full feature model) across machine learning models. Each subplot represents a different model, with 6 standardized performance metrics (accuracy, recall, precision, specificity, F_1-score, and area under the curve [AUC]) on the x-axis and their corresponding values (0-1.0) on the y-axis. Red bars represent models trained with all variables, while blue bars indicate models trained with the top 2 most important variables ("air_time" and "paper_time"). This visualization enables a direct comparison of model performance between the full and reduced feature sets across multiple evaluation metrics. AdaBoost: adaptive boosting; KNN: k-nearest neighbors; SVM: support vector machine; XGBoost: extreme gradient boosting.



The pattern was particularly profound for models, such as SVM, naïve Bayes, KNN, and neural network, which showed substantial improvements across all metrics when using the selected important variables (most notable was the neural

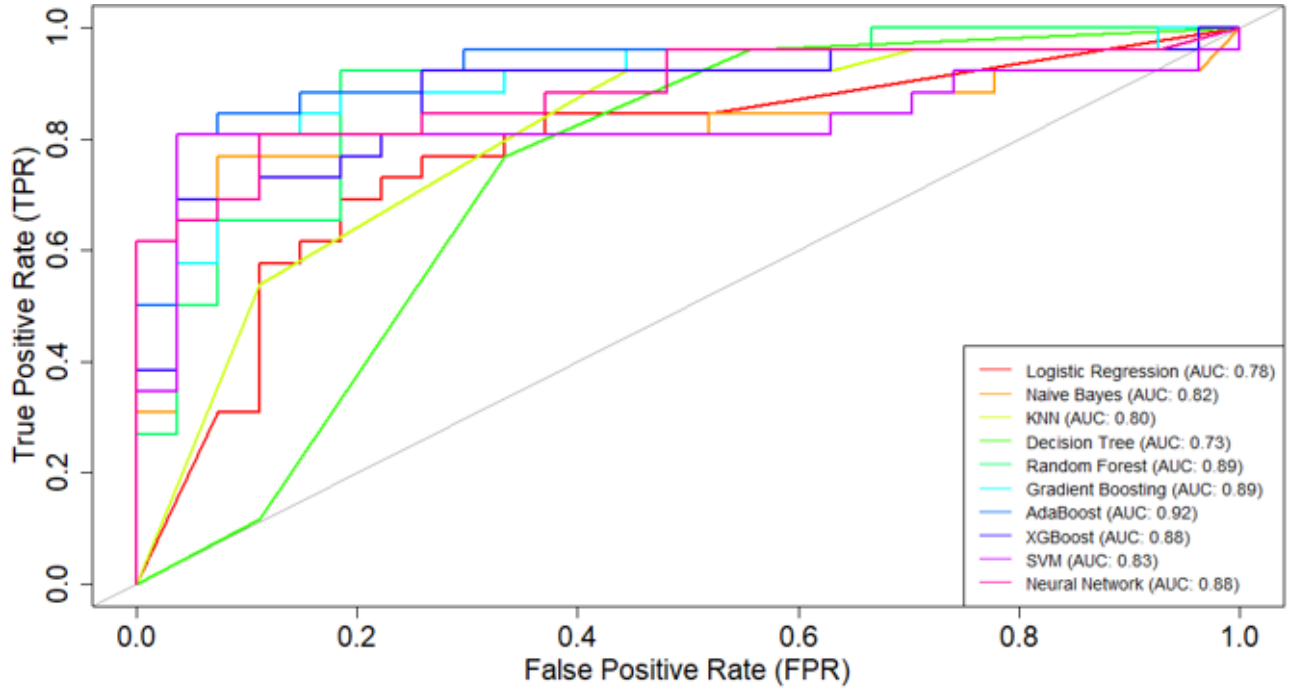
XSL•FO
RenderX

network's approximately 20% increment in the reduced feature model's test recall, from 69% to 88%). Ensemble methods like random forest, gradient boosting, AdaBoost, and XGBoost demonstrated slight improvements in performance across both variable sets, highlighting their inherent robustness regardless of feature selection [103,104]. Notably, the improvements observed in the validation metrics were largely mirrored in the test metrics, indicating that the benefits of using carefully selected important variables generalize well to unseen data.

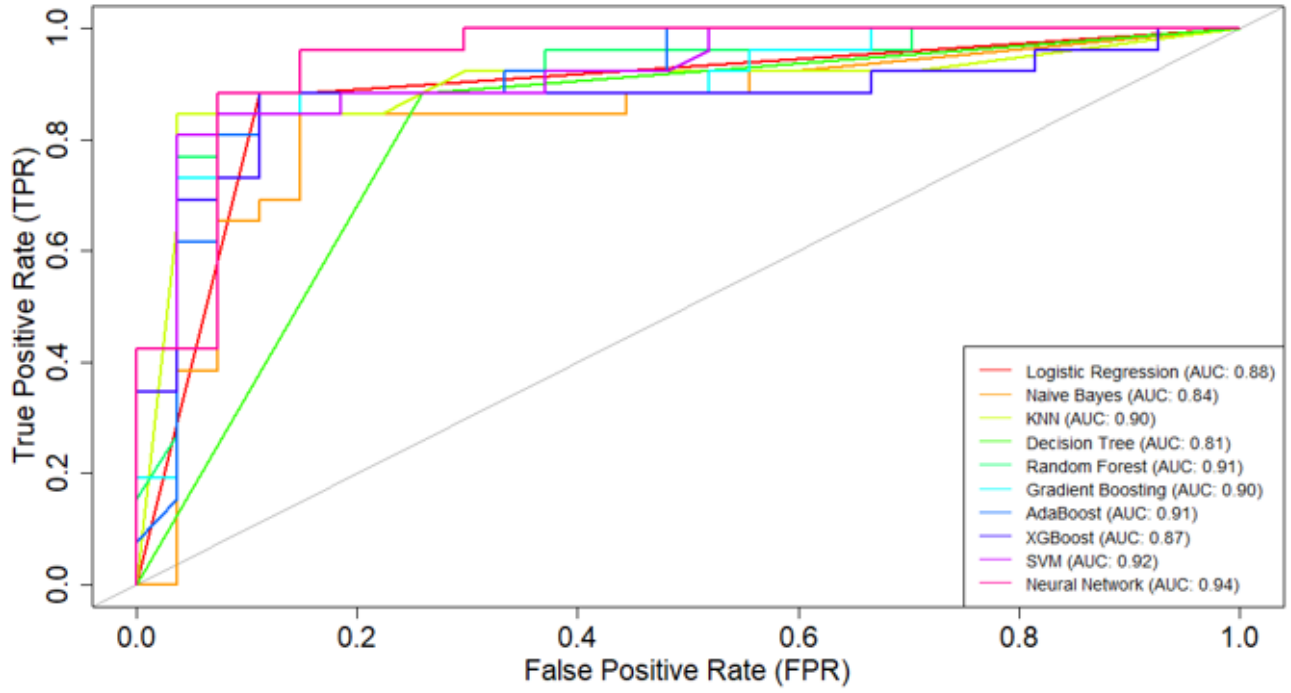
The consistent enhancement across various performance indicators, including accuracy, recall, precision, F_1-score, and ROC values, in both the validation and test scenarios reinforces the effectiveness of our variable importance–based feature selection approach. These findings strongly validate our strategy of using variable importance ranking for feature selection, demonstrating that focusing on the most influential variables not only maintains or improves model performance but also potentially reduces computational complexity and simplifies the model optimization process [105-108]. In the reduced feature model, the SVM classifier achieved an accuracy of 89% (95% CI 0.77 - 0.96) with an AUC of 92%, while the neural network algorithm reached an accuracy of 91% (95% CI 0.79 - 0.97) with an AUC of 94%, after identifying the most significant features for AD prediction in the test set. The comparison of ROC curves is presented in Figure 15.

XSL•FO
RenderX

Figure 15. Comparison of receiver operating characteristic (ROC) curves obtained directly from SMART-Pred (Shiny Multi-Algorithm R Tool for Predictive Modeling) between models trained with the full feature set (A) and selected important features (B). Each subplot shows the true positive rate against the false positive rate for 10 machine learning models, with corresponding area under the curve (AUC) values in the legend. Different colors represent different models, and the diagonal gray line indicates random prediction (AUC=0.5). This side-by-side comparison demonstrates the impact of feature selection on model discrimination ability. The closer the ROC curve is to the upper left corner of the graph, the higher is the accuracy of the test, as in the upper left corner, sensitivity is 1 and the false positive rate is 0 (specificity=1). The ideal ROC curve thus has an AUC of 1.0. AdaBoost: adaptive boosting; KNN: k-nearest neighbors; SVM: support vector machine; XGBoost: extreme gradient boosting.



A



B

SHAP Analysis on the DARWIN Dataset

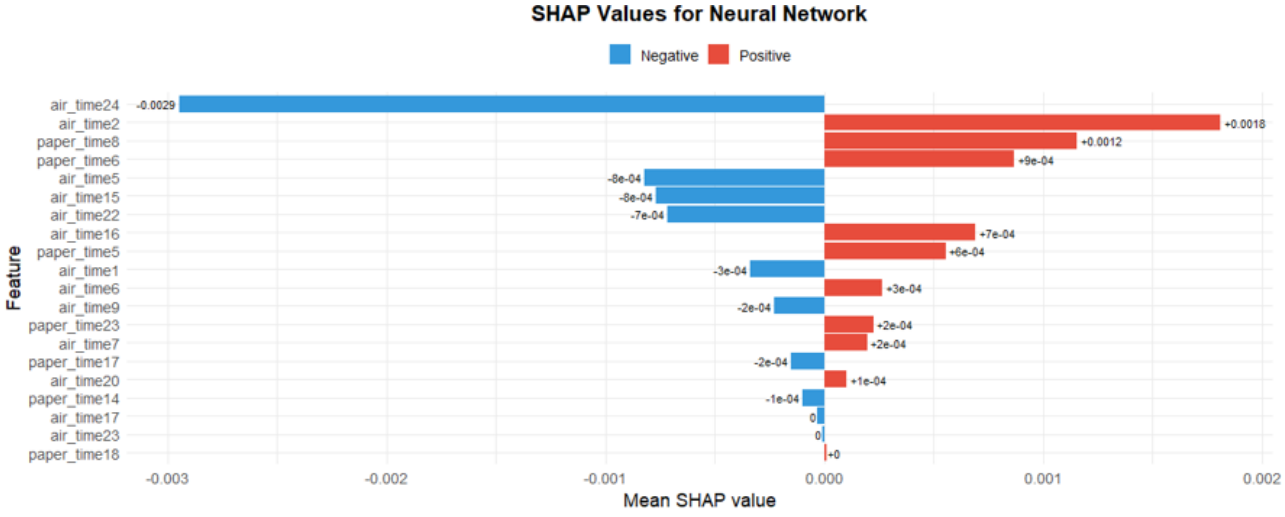
Interpreting the results and understanding the decision-making process of ML models has become paramount, especially in critical fields like medical diagnostics. In the context of AD detection, understanding which features or biological markers most influence a model's prediction (specifically direction of influence) can provide invaluable insights for medical professionals. In lieu of this, this study performed SHAP analysis. However, for brevity and ease of presentation, the

XSL•FO
RenderX

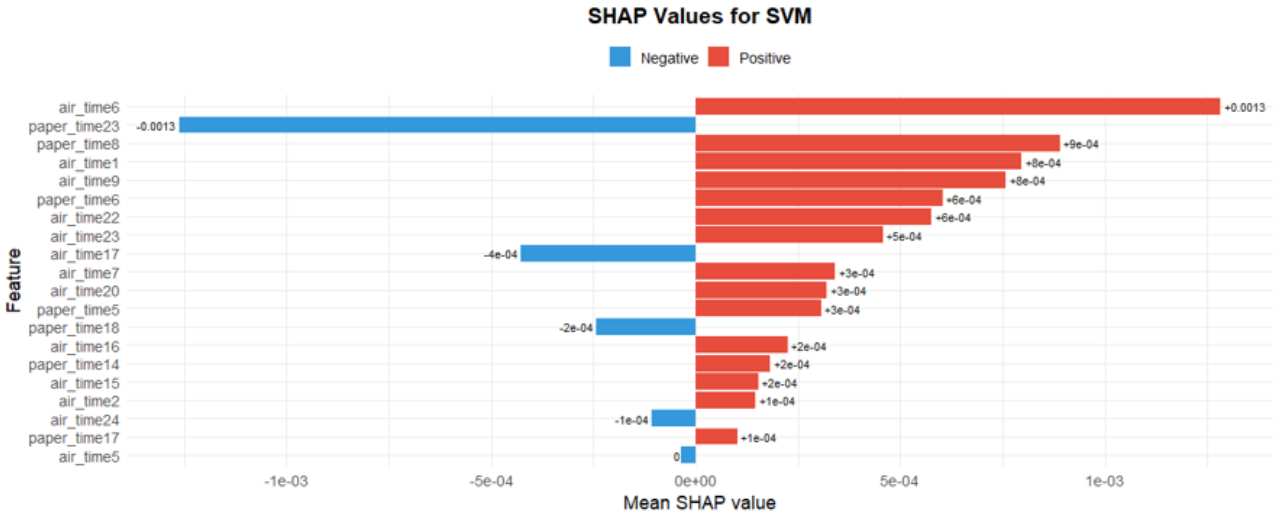
SHAP analysis of the 2 best models (neural network and SVM classifiers) is presented in this section, while the SHAP analysis of the remaining 8 models has been presented in Multimedia Appendix 2.

The SHAP analysis of the neural network classifier indicated that *air_time24* and *air_time5* (mean SHAP values of −0.0029 and −0.0008, respectively) had the greatest negative influence on AD prediction, while *air_time2* and *paper_time8* (mean SHAP values of +0.0018 and +0.0012, respectively) had the greatest positive influence on AD prediction (Figure 16A).

Figure 16. Shapley Additive Explanations (SHAP) summary plots for (A) neural network and (B) support vector machine (SVM) classifiers. Each plot visualizes the impact of individual features on model output using SHAP values, which quantify both the magnitude and direction of each feature's contribution to the prediction. In these summary plots, features are ranked from top to bottom by their overall importance in the model's decision-making process. Each point represents a feature value for an instance in the dataset, with color gradients indicating whether the feature value is low (blue) or high (red). The horizontal axis shows the SHAP value, reflecting whether an increased feature value drives the model output higher (positive SHAP value) or lower (negative SHAP value). Comparisons between neural network and support vector classifier results in the figure highlight both consensus and differences in feature importance and decision logic across models. These SHAP visualizations provide insights into which features most influence model predictions, supporting the interpretability and transparency of the machine learning models used in this research.



A



B

For the support vector classifier, *paper_time23* and *air_time17* (mean SHAP values of −0.0013 and −0.0004, respectively) had the greatest negative influence on AD prediction, while *air_time6* and *paper_time8* (mean SHAP values of +0.0013 and +0.0009, respectively) had the greatest positive influence on AD prediction (Figure 16B).

Discussion

Robustness of the SMART-Pred Application

AI is transforming medical practice and research by enhancing disease diagnostics, accelerating biomarker discovery, and revolutionizing disease surveillance [109-115]. In diagnostics, AI algorithms can examine complex medical data, such as imaging findings, genetic information, and clinical records, to identify subtle patterns that might be missed by human observers

XSL•FO
RenderX

[116-120]. This capability leads to earlier and more accurate disease detection, potentially improving patient outcomes. In biomarker identification, AI's ability to process vast amounts of multiomic data is uncovering novel indicators of disease presence, progression, and treatment response [121,122]. These AI-discovered biomarkers are paving the way for personalized medicine approaches and more targeted therapeutic interventions. For disease surveillance, AI systems are proving to be invaluable in monitoring and predicting disease outbreaks on a global scale [123]. By analyzing diverse data sources, including social media, environmental factors, and health records, AI can identify emerging health threats in real-time, enabling rapid response and containment strategies [124,125]. Moreover, AI's impact extends to drug discovery, treatment optimization, and health care resource allocation, promising a future of more efficient, effective, and accessible health care delivery [126,127].

The SMART-Pred application demonstrates a robust, user-friendly, and customizable approach to variable importance analysis across multiple ML models. This innovative tool offers valuable insights into the predictive power of various features in disease diagnostics, biomarker identification, and disease surveillance, particularly in the context of AD, as demonstrated in this study. By using the "varImp()" function and SHP analysis in R in SMART-Pred, our study identified the top-most influential variables for each model, providing a comprehensive view of feature relevance in AD prediction.

Discussion of the Findings

A thorough analysis of variable importance across all models revealed consistent patterns and model-specific characteristics. The "air_time" features, particularly air_time23, air_time2, and air_time6, emerged as consistently important variables across multiple models, including naïve Bayes, KNN, AdaBoost, and SVM. Additionally, "paper_time" features (paper_time17 and paper_time23) showed significant importance across several models, indicating the crucial role of temporal features in prediction. The selection of air_time (pen movement duration above paper) and paper_time (pen contact duration) in SMART-Pred's reduced feature model is rooted in their direct biological ties to AD pathology. Air_time reflects dysfunction of the prefrontal cortex, a region critical for motor planning and task initiation [128,129]. AD-related neurodegeneration is reflected in handwriting performance, with patients showing longer overall writing times and substantially increased air times between pen strokes across everyday copying tasks, indicating hesitation and disrupted sequencing of movements as they struggle to translate cognitive intentions into fluent motor output [130-132]. Paper_time, conversely, correlates with motor execution deficits caused by degeneration of the basal ganglia and cerebellar region, which impairs fine motor control. Quantitative tablet-based studies have reported that kinematic and pressure features, such as time in air, speed, and complexity, can correctly classify healthy controls, patients with mild cognitive impairment, and patients with AD, with roughly 70%-80% accuracy in 3-group models and above 90% accuracy in several pairwise contrasts. However, it was observed that the accuracy of classification was significantly influenced by the characteristics of the groups being categorized and the particular task, with rates varying from 63.5% to 100%. Moreover, the diagnostic accuracy derived from kinematic measurements exhibited higher specificity in differentiating normal cognition from impaired cognition (mild cognitive impairment and AD), whereas enhanced sensitivity was seen in discriminating between the degrees of impaired cognition (mild cognitive impairment vs AD) [133-135]. More recent digital pen work has shown that handwriting characteristics distinguish patients with AD from controls, with the task accuracy of the XGBoost model achieving a maximum value of 96.55% and the AUC reaching up to 99.10%, proving that handwriting characteristic analysis is promising in auxiliary AD screening or AD diagnosis [136]. On the other hand, the spatiotemporal dynamics of simple loop handwriting separate patients with early-stage AD from healthy older adults, with about 74% classification accuracy [137]. These findings support total duration, time in air, and related temporal variability as sensitive quantitative markers that differentiate healthy aging, mild cognitive impairment, and AD and that can contribute to early staging and screening.

The validity of the variables is further supported by multimodal evidence. Neuroimaging and digital clock drawing studies have linked temporal handwriting features, including "time in air" and "time on paper," to global brain and gray matter volumes, while time-based handwriting metrics have shown significant negative correlations with global cognitive scores, such as the MoCA (Montreal Cognitive Assessment), supporting the interpretation of prolonged air_time and paper_time variability as markers of cognitive decline [138-142]. Compared with invasive biomarkers, such as cerebrospinal fluid Aβ42, which require lumbar puncture and typically achieve diagnostic sensitivities of around 80% at a cost of approximately US $1000 per test [143,144], and FDG-PET (fluorodeoxyglucose positron emission tomography) scans, which show similar median sensitivity but are substantially more expensive (often around US $3000 per scan) [145,146], digital clock drawing metrics, such as air_time and paper_time, provide noninvasive, low-cost screening with high diagnostic accuracy for early cognitive impairment and AD [147], with our SMART-Pred protocol achieving 88% sensitivity at a cost of approximately US $5 per test (our study). Their ability to capture both executive dysfunction (via planning delays) and visuospatial deficits (via erratic stroke execution) provides a holistic view of AD's cognitive-motor integration breakdown, aligning with National Institutes of Health priorities for functional digital biomarkers [148-150]. This dual focus on mechanistic relevance and clinical practicality makes them ideal for scalable, early AD detection in routine care settings.

Consistent with the literature, our findings showed that random forest had a more balanced feature utilization [151,152], with paper_time6 showing maximum importance (100.00), followed by air_time23 (98.27) and total_time9 (94.82). The model also highlighted the significance of diverse features, such as paper_time23 (90.20), mean_speed_in_air25 (89.03), and mean_jerk_in_air17 (84.03), suggesting a more comprehensive feature integration approach [153-156]. Gradient boosting and XGBoost showed similar patterns but with distinct emphases [156]. Gradient boosting prioritized air_time16 (100.00), while XGBoost highlighted air_time15 (100.00) as the most important

XSL•FO
RenderX

feature. The gradient boosting model incorporated additional variables like pressure_mean8 and mean_speed_on_paper2 in the predictions, while the XGBoost model incorporated additional variables like gmrt_in_air1 and total_time3 in the predictions. The neural network model exhibited a unique feature importance pattern, emphasizing variables that were less prominent in other models. While including air_time features, it notably prioritized gmrt_on_paper2 (100.00), mean_acc_in_air13 (98.61), and num_of_pendown19 (98.58), suggesting its ability to capture better patterns in the data, particularly those related to acceleration and movement characteristics. These findings regarding variable importance support the notion that interpretability is crucial in AI systems, especially in medical applications [157-161]. The ability to identify and understand underlying diagnostic biomarkers is not only beneficial for enhancing the diagnostic process but also plays a vital role in tailoring treatment plans for patients [162-168].

ROC curve analysis provided a comparative evaluation of various ML models when assessed with all available variables (full feature model) versus a carefully selected subset of the most important variables ("air_time" and "paper_time" variables; reduced feature model). In the full feature model, AdaBoost demonstrated exceptional predictive capability with an AUC of 0.92, while decision tree showed relatively lower performance with an AUC of 0.73. Strong performance was also noted for random forest and gradient boosting (both with an AUC of 0.89) and XGBoost and neural network (both with an AUC of 0.89). In the reduced feature model, neural network and SVM classifiers outperformed all other algorithms (AUC values of 0.94 and 0.92, respectively), revealing their effectiveness even with fewer inputs. Ensemble models, including random forest, gradient boosting, XGBoost, and AdaBoost, continued to perform well, achieving AUC values ranging from 0.87 to 0.91, which reflect their robustness [169-172]. Interestingly, the decision tree's AUC improved to 0.81, indicating that it benefited significantly from targeted variable selection in the reduced feature model. Consistent with the literature, our results from the SMART-Pred analysis highlight the consistent strength of ensemble models, such as random forest, gradient boosting, XGBoost, and AdaBoost, across varying variable sets and support the strategic approach of leveraging variable importance ranking to enhance model performance [173-179]. By focusing on critical features, models can maintain or even improve accuracy, as demonstrated in this work. However, it should be emphasized that nothing in this work demonstrated causality between these variables and AD.

Comparison of SMART-Pred With Other AI Tools for AD Diagnostics

SMART-Pred offers a unique approach to AD diagnostics by employing handwriting analysis as a noninvasive biomarker, distinguishing itself from other AI tools in the field. Unlike most existing tools that use cognitive tests, magnetic resonance imaging scans, or blood-based markers, SMART-Pred leverages the DARWIN dataset, which includes handwriting samples from 174 participants (89 patients with AD and 85 healthy controls) across 25 tasks, identifying critical predictors like *air_time* and *paper_time*. This focus provides a novel, cost-effective method for early AD detection. The tool integrates 10 diverse ML models, including random forest, AdaBoost, and neural network, within a user-friendly R Shiny application, democratizing access to advanced AI for nonspecialists and enabling robust performance comparison and customization. With neural network achieving 91% accuracy and SVM achieving 89% accuracy on the test set of the reduced feature model, SMART-Pred surpasses many clinical diagnostic tools (typically having around 81% accuracy) using noninvasive data, and its emphasis on interpretability through variable importance ranking (eg, air_time and paper_time) via the "varImp()" function in R aids physician acceptance and hypothesis formulation.

SMART-Pred stands out due to its reliance on accessible handwriting data compared to resource-intensive methods like magnetic resonance imaging scans or genetic testing used by tools, such as the USC model (approximately 85% accuracy) [180] and the Cambridge AI tool (82% accuracy for progression prediction) [45]. While a quantum ML approach has reported an exceptional 99.89% accuracy, its practical application is limited by computational complexity [181]. On the other hand, SMART-Pred prioritizes ease of use and scalability. Unlike the UCSF model, which predicts AD risk up to 7 years in advance using clinical records (72% accuracy), SMART-Pred focuses on immediate diagnostic classification for direct clinical application. Its ability to bridge complex AI algorithms with user-friendly interfaces addresses a critical gap in clinical adoption, and this is further enhanced by its focus on interpretability, a challenge for many AI models. By pinpointing specific handwriting features as predictors, SMART-Pred not only facilitates early detection but also contributes to understanding potential risk factors, paving the way for tailored interventions.

The interplay between AI interpretability and physician acceptance is a key factor in integrating AI systems into health care. SMART-Pred, a multifaceted R Shiny AI-based predictive modeling application, demonstrated robust predictive power in AD diagnosis using the DARWIN dataset. Incorporating 10 ML algorithms with tunable parameters, it achieved optimal results, with ensemble models like random forest and AdaBoost excelling across evaluation metrics. This provides an efficient, customizable approach to AI modeling, particularly for complex datasets and disease diagnosis, which is extendable to other classification-based scenarios. Its ability to identify potential risk factors makes it valuable for guiding hypothesis formulation and experimentation across health care domains. Research indicates that physicians are more likely to adopt AI technologies when the reasoning behind insights is clear, especially in complex fields like AD diagnosis and management [182-184].

Future studies could focus on clinical assessments to evaluate whether top variables like *air_time* and *paper_time* are sufficient for AD diagnosis, strengthening the synergy between AI outcomes and experimental findings. Additionally, SMART-Pred could be integrated into electronic health record systems to identify biomarkers of successful aging, enhance disease diagnosis and surveillance, and analyze medical records, addressing public health concerns through data-driven insights

XSL•FO
RenderX

and personalized medicine strategies. The use of smart home data for predicting AD symptoms has also shown promise, highlighting the potential of diverse data sources in AD research, with SMART-Pred's handwriting analysis offering a complementary approach.

SMART-Pred demonstrated competitive performance compared to recent AI models for AD prediction. Its unique use of handwriting analysis and accessibility as an R Shiny application contribute significantly to the field. Future work could also explore integrating multiple data modalities to enhance its predictive accuracy and robustness. Additional clinical research is needed to evaluate the practical applicability of the top-ranked variables (*air_time* and *paper_time*) in real-world AD risk assessment and diagnosis.

A comprehensive and comparative analysis of the performance of AI tools for AD detection and prediction is provided in Table 6.

XSL•FO
RenderX

Table . Comparative performance of AI[a] tools for AD[b] detection and prediction.

AI tool	Primary data source	Accuracy	Key features/differentiators
SMART-Pred[c] (our study)	Handwriting analysis (DARWIN[d] dataset)	91.00% (neural network) with 95% CI: 0.79 - 0.97), AUC=94%	Multi-algorithm R Shiny application (10 ML models) with automated feature selection, explainable AI (SHAP), and cross-validation; non-invasive, cost-effective, user-friendly, generalizable framework applicable to diverse biomarkers and deployable beyond handwriting analysis
Cambridge AI tool [45]	Cognitive tests, MRI[e] scans	82.00% (progression prediction)	Predicts AD progression within 3 years; 3 times more accurate than standard clinical markers
USC model [180]	MRI scans, genetic data, biomarkers	85.00% (CNN[f]) with control and MCI[g]/AD groups	Uses deep learning for high accuracy in brain scan analysis; integrates diverse data types
Quantum ML[h] approach [181]	Unspecified (advanced computational methods)	99.89%	Exceptional accuracy but limited by complexity and resource needs; less practical for immediate use
Deep learning CNN [185]	Handwriting analysis (DARWIN dataset)	90.40%	Proposes a novel CNN as a cheap, fast, and accurate solution with a high accuracy, but with no XAI[i]
Particle swarm optimization (PSO) algorithm [186]	Handwriting analysis (DARWIN dataset)	90.57% (LR[j], AdaBoost[k], and RF[l], with RF dominating the accuracy battle with the best feature dimension of 20)	Uses a hybrid ML model for predicting AD using PSO-based feature selection, but with no XAI.
ML [73]	Handwriting analysis (DARWIN dataset)	85.29% (RF)	Proposes a single classifier merging the features from the best tasks, but with no XAI.
ML [187]	DARWIN dataset analysis	82.43% (RF)	Uses several ML techniques with selected tasks of different types and complexities to evaluate the healthy status of patients and a combination of different tasks, but with no XAI
ML [188]	Handwriting analysis (DARWIN dataset)	85% (XGBoost[m])	Uses several ensemble techniques by implementing a 5-fold cross-validation that ensured robust and unbiased results, but with no XAI
ML [189]	Handwriting analysis (DARWIN dataset)	85.86% (LR)	Uses several ML techniques involving extensive hyperparameter tuning, but with no XAI

[a]AI: artificial intelligence.

[b]AD: Alzheimer disease.

[c]SMART-Pred: Shiny Multi-Algorithm R Tool for Predictive Modeling.

[d]DARWIN: Diagnosis Alzheimer WIth Handwriting.

[e]MRI: magnetic resonance imaging.

[f]CNN: convolutional neural network.

[g]MCI: mild cognitive impairment.

[h]ML: machine learning.

[i]XAI: explainable artificial intelligence.

[j]LR: linear regression.

[k]AdaBoost: adaptive boosting.

[l]RF: random forest.

[m]XGBoost: extreme gradient boosting.

XSL•FO
RenderX

Practical Design Implications of SMART-Pred

The implementation of SMART-Pred, an AI-based tool for early AD detection via handwriting analysis, holds transformative practical design implications across clinical diagnostics, health care technology, and neurodegenerative research. In clinical settings, SMART-Pred provides a noninvasive, cost-effective screening option, achieving up to 91% accuracy with the neural network model (with 95% CI: 0.79 - 0.97 and AUC=94%), which can be seamlessly integrated into primary care where access to advanced neuroimaging or cerebrospinal fluid testing is limited. Using simple handwriting tasks from the DARWIN dataset, the application enables early identification of motor-cognitive decline, potentially reducing diagnostic delays, especially in underserved populations. This supports value-based care by enhancing efficiency and patient outcomes, aligning with broader trends of AI adoption in health care, and prioritizing patients for further evaluation or early interventions.

Beyond clinical use, SMART-Pred influences health care technology by showcasing a user-friendly R Shiny interface that automates complex ML workflows (data preprocessing, model training, and evaluation) for nonspecialist clinical researchers and providers while addressing challenges like infrastructure compatibility and algorithmic bias. Its focus on explainability through feature importance ranking (*air_time* and *paper_time*) promotes transparency, building trust among clinicians who are wary of black-box systems. In neurodegenerative research, SMART-Pred inspires exploration of noninvasive biomarkers for other disorders like Parkinson disease and Huntington disease, and future validation on datasets like ADNI could lead to hybrid models combining handwriting with imaging or genetic data for enhanced precision. By prioritizing clinical utility, SMART-Pred supports future AI development in radiology and beyond, encouraging interdisciplinary collaboration and clinician involvement in design to address real-world health care needs and drive innovation in diagnostics.

Limitations and Future Work

The SMART-Pred tool has demonstrated promising capabilities for AD prediction through handwriting analysis, but it faces a few critical limitations in its basic version. Computationally, its exclusive reliance on CPU processing within the R environment restricts parallel computing capabilities, leading to inefficiencies when handling large datasets or complex ML algorithms like neural network [190]. This architectural constraint directly impacts scalability, particularly for resource-intensive tasks such as hyperparameter optimization and multimodal data integration. A fundamental reproducibility challenge stems from the inability to save key outputs (model metrics, ROC curves, and trained model states), hindering longitudinal studies and comparative analyses [191-194]. These limitations collectively restrict the tool's utility in real-world clinical workflows.

To address these issues, the authors propose an advanced version integrating Python-based frameworks like TensorFlow for GPU-accelerated neural network models and enhanced computational efficiency. Planned upgrades include flexible hyperparameter optimization, automated parameter distribution analysis, and comprehensive result-saving functionalities. Future validation will expand beyond the DARWIN dataset to incorporate multimodal data from the ADNI, encompassing neuroimaging, genetic, and biochemical biomarkers. This broader validation aims to assess generalizability across diverse populations and clinical settings while enabling cross-modal biomarker comparisons, a critical step toward establishing SMART-Pred as a clinically actionable tool for early AD detection.

Comparison With Prior Work

Our study on SMART-Pred aligns with the growing trend of using AI for the early detection and prediction of AD. The performance of our model (91% accuracy with neural network, 95% CI: 0.79 - 0.97; AUC=94%) compares favorably with recent advancements in the field. Table 7 presents a concise comparison of studies that used ML, explainable AI, and the DARWIN dataset for AD detection.

XSL•FO
RenderX

Table . Comparative analysis of ML[a] and XAI[b] studies using the DARWIN[c] dataset for Alzheimer disease detection.

Dataset	ML techniques used	XAI methods applied	Key insights and features identified	Performance metrics
DARWIN dataset analysis [195]	LR[d], RF[e], SVM[f], and EBM[g]	SHAP[h] and EBM interpretability	"air_time" (in-air movement duration) identified as the most decisive feature group	89% (SD 5%) accuracy with EBM on the reduced dimensionality dataset
DARWIN dataset analysis [196]	ML and stacking	SHAP XAI features importance and direction of influence	"air_time" and "paper_time" (pen contact duration) features dominate the feature importance analysis using SHAP	88.57% accuracy with the stacking ensemble model
DARWIN dataset analysis [197]	LR, AdaBoost[i], XGBoost[j], KNN[k], RF, SVM, etc	LIME[l] and ELI5[m]	"air_time" and "paper_time" features were the most important features using ELI5	91.00% accuracy with RF
SMART-Pred[n] (current study)	LR, AdaBoost, XGBoost, KNN, RF, SVM, neural network, etc	SHAP features importance and direction of influence	"air_time" and "paper_time" as primary biomarkers using SHAP XAI.	91.00% accuracy with neural network; 95% CI 0.79 - 0.97; AUC=94%

[a]ML: machine learning.

[b]XAI: explainable artificial intelligence.

[c]DARWIN: Diagnosis Alzheimer WIth Handwriting.

[d]LR: logistic regression.

[e]RF: random forest.

[f]SVM: support vector machine.

[g]EBM: explainable boosting machine.

[h]SHAP: Shapley Additive Explanations.

[i]AdaBoost: adaptive boosting.

[j]XGBoost: extreme gradient boosting.

[k]KNN: k-nearest neighbors.

[l]LIME: Local Interpretable Model-Agnostic Explanations.

[m]ELI5: Explain Like I'm 5.

[n]SMART-Pred: Shiny Multi-Algorithm R Tool for Predictive Modeling.

The comparative analysis revealed 3 critical insights. First, the studies demonstrate feature consensus, identifying temporal handwriting features like "air_time" as pivotal predictors, thereby validating their biological relevance to AD motor-cognitive decline through their association with prefrontal cortex dysfunction and motor planning deficits. Second, SMART-Pred represents a methodological evolution from prior work by expanding model diversity (10 algorithms vs up to a maximum of 7 algorithms in earlier studies), incorporating neural networks to capture complex nonlinear patterns in handwriting dynamics, and advancing biomarker identification through dual focus on "air_time" (in-air movement duration) and "paper_time" (pen contact duration) rather than single-feature analysis. Third, despite architectural differences, SMART-Pred classifiers performed very well (with neural network achieving an accuracy of 91% [95% CI 0.79-0.97]), highlighting the DARWIN dataset's robustness for handwriting-based diagnosis while indicating that model selection impacts biomarker discovery more than absolute accuracy in this domain. This progression from single-feature interpretation to multi-algorithm, multibiomarker analysis reflects an intentional shift toward clinical translatability, balancing explainability with predictive power.

Conclusion of the Study

The SMART-Pred R Shiny application represents a significant advancement in AI-based predictive modeling for medical diagnostics, offering a user-friendly interface and comprehensive algorithm selection. This tool holds value for researchers and clinicians in personalized medicine and public health surveillance, demonstrating potential to enhance early detection and screening strategies for conditions such as AD and Parkinson disease. By bridging the gap between computational modeling and clinical application, SMART-Pred exemplifies the growing trend of leveraging interactive data visualization and AI-powered analysis in health care settings.

Future development of SMART-Pred will focus on clinical validation and address the challenges associated with imbalanced datasets. The roadmap includes expanding algorithmic support to encompass advanced ML techniques, image recognition architectures, and analysis capabilities for diverse data types such as text, time series, and graph data. Additionally, plans to improve visualization capabilities and user interface design aim to increase accessibility and power. These ongoing enhancements are designed to create a more flexible and comprehensive platform for data analysis across various scientific domains, solidifying SMART-Pred's role in advancing predictive modeling in health care and beyond.

XSL•FO
RenderX

Acknowledgments

This work was supported by start-up funds provided to SK by Celia Scott Weatherhead School of Public Health and Tropical Medicine at Tulane University. The article processing charges for this manuscript were supported by the Weatherhead Foundation through a gift to the Tulane University Celia Scott Weatherhead School of Public Health & Tropical Medicine in support of the School's AI and Data Science Strategic Initiative. This generous support has enabled us to advance work that has strong potential to contribute to the development and application of AI in public health, in alignment with the Dean's strategic AI initiative.

Funding

This work was supported by start-up funds provided to SK by Celia Scott Weatherhead School of Public Health and Tropical Medicine at Tulane University.

Data Availability

The DARWIN (Diagnosis Alzheimer With Handwriting) dataset [72,73] was used to validate the classification capabilities of the SMART-Pred (Shiny Multi-Algorithm R Tool for Predictive Modeling) application. A fully reproducible workflow using the provided R codes, including all necessary packages and code snippets for SMART-Pred, is available in our GitHub repository [198]. These scripts are designed to facilitate reproducibility and allow users to replicate the machine learning classification tasks outlined in our study.

Authors' Contributions

Conceptualization: SK
Data curation: HW, EFA, SK
Formal analysis: HW, EFA, SK, SKS
Funding acquisition: SK
Investigation: HW, EFA, SK, SKS
Methodology: HW, EFA, SK
Project administration: SK
Resources: SK
Software: HW, EFA, SK
Supervision: SK, SKS, JGS
Validation: HW, EFA, SK, SKS, JGS
Visualization: HW, EFA, SK
Writing – original draft: SK, HW
Writing – review & editing: SK, EFA, JGS, SKS

Conflicts of Interest

None declared.

Multimedia Appendix 1
SMART-Pred (Shiny Multi-Algorithm R Tool for Predictive Modeling) application and data.
[ZIP File, 356 KB - aging_v8i1e70272_app1.zip]

Multimedia Appendix 2
Shapley Additive Explanations analysis results for the additional machine learning classifiers.
[DOCX File, 125 KB - aging_v8i1e70272_app2.docx]

References

1. Alowais SA, Alghamdi SS, Alsuhebany N, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC Med Educ 2023 Sep 22;23(1):689. [doi: 10.1186/s12909-023-04698-z] [Medline: 37740191]
2. Choubey A, Choubey SB, K P, Daulatabad VS, John N. Healthcare transformation: artificial intelligence is the dire imperative of the day. Cureus 2024 Jun;16(6):e62652. [doi: 10.7759/cureus.62652] [Medline: 39036139]
3. Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthc J 2021 Jul;8(2):e188-e194. [doi: 10.7861/fhj.2021-0095] [Medline: 34286183]
4. Khalifa M, Albadawy M. AI in diagnostic imaging: revolutionising accuracy and efficiency. Computer Methods and Programs in Biomedicine Update 2024;5:100146. [doi: 10.1016/j.cmpbup.2024.100146]

XSL•FO
RenderX

5. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts H. Artificial intelligence in radiology. Nat Rev Cancer 2018 Aug;18(8):500-510. [doi: 10.1038/s41568-018-0016-5] [Medline: 29777175]
6. Yu F, Moehring A, Banerjee O, Salz T, Agarwal N, Rajpurkar P. Heterogeneity and predictors of the effects of AI assistance on radiologists. Nat Med 2024 Mar;30(3):837-849. [doi: 10.1038/s41591-024-02850-w] [Medline: 38504016]
7. Shafi S, Parwani AV. Artificial intelligence in diagnostic pathology. Diagn Pathol 2023 Oct 3;18(1):109. [doi: 10.1186/s13000-023-01375-z] [Medline: 37784122]
8. Sarwar S, Dent A, Faust K, et al. Physician perspectives on integration of artificial intelligence into diagnostic pathology. NPJ Digit Med 2019;2(1):28. [doi: 10.1038/s41746-019-0106-0] [Medline: 31304375]
9. Ghenciu LA, Dima M, Stoicescu ER, Iacob R, Boru C, Ha egan OA. Retinal imaging-based oculomics: artificial intelligence as a tool in the diagnosis of cardiovascular and metabolic diseases. Biomedicines 2024 Sep 23;12(9):2150. [doi: 10.3390/biomedicines12092150] [Medline: 39335664]
10. Kong M, Song SJ. Artificial intelligence applications in diabetic retinopathy: what we have now and what to expect in the future. Endocrinol Metab (Seoul) 2024 Jun;39(3):416-424. [doi: 10.3803/EnM.2023.1913] [Medline: 38853435]
11. Lopez-Jimenez F, Attia Z, Arruda-Olson AM, et al. Artificial intelligence in cardiology: present and future. Mayo Clin Proc 2020 May;95(5):1015-1039. [doi: 10.1016/j.mayocp.2020.01.038] [Medline: 32370835]
12. Li Z, Yin S, Wang S, Wang Y, Qiang W, Jiang J. Transformative applications of oculomics-based AI approaches in the management of systemic diseases: a systematic review. J Adv Res 2025 Sep;75:679-696. [doi: 10.1016/j.jare.2024.11.018]
13. Barry T, Farina JM, Chao CJ, et al. The role of artificial intelligence in echocardiography. J Imaging 2023 Feb 20;9(2):9. [doi: 10.3390/jimaging9020050] [Medline: 36826969]
14. Maleki Varnosfaderani S, Forouzanfar M. The role of AI in hospitals and clinics: transforming healthcare in the 21st century. Bioengineering (Basel) 2024 Mar 29;11(4):337. [doi: 10.3390/bioengineering11040337] [Medline: 38671759]
15. Olawade DB, Wada OZ, Odetayo A, David-Olawade AC, Asaolu F, Eberhardt J. Enhancing mental health with artificial intelligence: current trends and future prospects. Journal of Medicine, Surgery, and Public Health 2024 Aug;3:100099. [doi: 10.1016/j.glmedi.2024.100099]
16. Bekbolatova M, Mayer J, Ong CW, Toma M. Transformative potential of AI in healthcare: definitions, applications, and navigating the ethical landscape and public perspectives. Healthcare (Basel) 2024 Jan 5;12(2):125. [doi: 10.3390/healthcare12020125] [Medline: 38255014]
17. Quer G, Arnaout R, Henne M, Arnaout R. Machine learning and the future of cardiovascular care: JACC state-of-the-art review. J Am Coll Cardiol 2021 Jan 26;77(3):300-313. [doi: 10.1016/j.jacc.2020.11.030] [Medline: 33478654]
18. Drachman DA. Aging of the brain, entropy, and Alzheimer disease. Neurology (ECronicon) 2006 Oct 24;67(8):1340-1352. [doi: 10.1212/01.wnl.0000240127.89601.83] [Medline: 17060558]
19. Alzheimer's Association. 2018 Alzheimer's disease facts and figures. Alzheimer's & Dementia 2018 Mar;14(3):367-429. [doi: 10.1016/j.jalz.2018.02.001]
20. Breijyeh Z, Karaman R. Comprehensive review on Alzheimer's disease: causes and treatment. Molecules 2020 Dec 8;25(24):5789. [doi: 10.3390/molecules25245789] [Medline: 33302541]
21. Kumar A, Sidhu J, Lui F, Tsao JW. Alzheimer Disease: StatPearls Publishing; 2023. URL: https://www.ncbi.nlm.nih.gov/books/NBK499922/ [accessed 2025-11-29]
22. Jahn H. Memory loss in Alzheimer's disease. Dialogues Clin Neurosci 2013 Dec;15(4):445-454. [doi: 10.31887/DCNS.2013.15.4/hjahn] [Medline: 24459411]
23. Crecelius C. Diagnosis and treatment of non-Alzheimer's dementias. J Am Med Dir Assoc 2003;4(4 Suppl):H25-H29. [doi: 10.1097/01.JAM.0000078013.50275.FF] [Medline: 12857366]
24. Giau VV, Senanarong V, Bagyinszky E, An SSA, Kim S. Analysis of 50 neurodegenerative genes in clinically diagnosed early-onset Alzheimer's disease. Int J Mol Sci 2019 Mar 26;20(6):1514. [doi: 10.3390/ijms20061514] [Medline: 30917570]
25. Kennison M, Long E. The long journey of Alzheimer's disease. J Christ Nurs 2018;35(4):218-227. [doi: 10.1097/CNJ.0000000000000529] [Medline: 30198992]
26. Cummings JL. Current perspectives in Alzheimer's disease. Neurology (ECronicon) 1998 Jul;51(1_suppl_1):S2-S17. [doi: 10.1212/WNL.51.1_Suppl_1.S1]
27. Mendez MF. Early-onset Alzheimer disease. Neurol Clin 2017 May;35(2):263-281. [doi: 10.1016/j.ncl.2017.01.005] [Medline: 28410659]
28. DiLoreto R, Murphy CT. The cell biology of aging. Mol Biol Cell 2015 Dec 15;26(25):4524-4531. [doi: 10.1091/mbc.E14-06-1084] [Medline: 26668170]
29. Bouska M, Huang K, Kang P, Bai H. Organelle aging: lessons from model organisms. J Genet Genomics 2019 Apr 20;46(4):171-185. [doi: 10.1016/j.jgg.2019.03.011] [Medline: 31080045]
30. Niccoli T, Partridge L. Ageing as a risk factor for disease. Curr Biol 2012 Sep 11;22(17):R741-R752. [doi: 10.1016/j.cub.2012.07.024] [Medline: 22975005]
31. Morawe T, Hiebel C, Kern A, Behl C. Protein homeostasis, aging and Alzheimer's disease. Mol Neurobiol 2012 Aug;46(1):41-54. [doi: 10.1007/s12035-012-8246-0] [Medline: 22361852]
32. Cummings JL, Gonzalez MI, Pritchard MC, May PC, Toledo-Sherman LM, Harris GA. The therapeutic landscape of tauopathies: challenges and prospects. Alz Res Therapy 2023;15(1):168. [doi: 10.1186/s13195-023-01321-7]

XSL•FO
RenderX

33. Pais M, Martinez L, Ribeiro O, et al. Early diagnosis and treatment of Alzheimer's disease: new definitions and challenges. Braz J Psychiatry 2020 Aug;42(4):431-441. [doi: 10.1590/1516-4446-2019-0735] [Medline: 31994640]
34. Nayab DE, Din FU, Ali H, et al. Nano biomaterials based strategies for enhanced brain targeting in the treatment of neurodegenerative diseases: an up-to-date perspective. J Nanobiotechnology 2023 Dec 13;21(1):477. [doi: 10.1186/s12951-023-02250-1] [Medline: 38087359]
35. Singh A, Dawson TM, Kulkarni S. Neurodegenerative disorders and gut-brain interactions. J Clin Invest 2021 Jul 1;131(13):13. [doi: 10.1172/JCI143775] [Medline: 34196307]
36. GBD 2016 Dementia Collaborators. Global, regional, and national burden of Alzheimer's disease and other dementias, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol 2019 Jan;18(1):88-106. [doi: 10.1016/S1474-4422(18)30403-4] [Medline: 30497964]
37. Maresova P, Hruska J, Klimova B, Barakovic S, Krejcar O. Activities of daily living and associated costs in the most widespread neurodegenerative diseases: a systematic review. Clin Interv Aging 2020;15:1841-1862. [doi: 10.2147/CIA.S264688] [Medline: 33061334]
38. Alzheimer's Association. 2016 Alzheimer's disease facts and figures. Alzheimer's & Dementia 2016 Apr;12(4):459-509. [doi: 10.1016/j.jalz.2016.03.001]
39. Rizzi L, Rosset I, Roriz-Cruz M. Global epidemiology of dementia: Alzheimer's and vascular types. Biomed Res Int 2014;2014:908915. [doi: 10.1155/2014/908915] [Medline: 25089278]
40. Aisen PS, Cummings J, Jack CR Jr, et al. On the path to 2025: understanding the Alzheimer's disease continuum. Alzheimers Res Ther 2017 Aug 9;9(1):60. [doi: 10.1186/s13195-017-0283-5] [Medline: 28793924]
41. Bowroju SK, Mainali N, Ayyadevara S, et al. Design and synthesis of novel hybrid 8-hydroxy quinoline-indole derivatives as inhibitors of Aβ self-aggregation and metal chelation-induced Aβ aggregation. Molecules 2020 Aug 8;25(16):16. [doi: 10.3390/molecules25163610] [Medline: 32784464]
42. Kakraba S, Ayyadevara S, Penthala NR, et al. A novel microtubule-binding drug attenuates and reverses protein aggregation in animal models of Alzheimer's disease. Front Mol Neurosci 2019;12:310. [doi: 10.3389/fnmol.2019.00310] [Medline: 31920540]
43. Kakraba S, Ayyadevara S, Mainali N, et al. Thiadiazolidinone (TDZD) analogs inhibit aggregation-mediated pathology in diverse neurodegeneration models, and extend C. elegans life- and healthspan. Pharmaceuticals (Basel) 2023 Oct 20;16(10):10. [doi: 10.3390/ph16101498] [Medline: 37895969]
44. Ayyadevara S, Balasubramaniam M, Kakraba S, Alla R, Mehta JL, Shmookler Reis RJ. Aspirin-mediated acetylation protects against multiple neurodegenerative pathologies by impeding protein aggregation. Antioxid Redox Signal 2017 Dec 10;27(17):1383-1396. [doi: 10.1089/ars.2016.6978] [Medline: 28537433]
45. Lee LY, Vaghari D, Burkhart MC, et al. Robust and interpretable AI-guided marker for early dementia prediction in real-world clinical settings. EClinicalMedicine 2024 Aug;74:102725. [doi: 10.1016/j.eclinm.2024.102725] [Medline: 39764178]
46. Schöll M, Verberk IMW, Del Campo M, et al. Challenges in the practical implementation of blood biomarkers for Alzheimer's disease. Lancet Healthy Longev 2024 Oct;5(10):100630. [doi: 10.1016/j.lanhl.2024.07.013] [Medline: 39369727]
47. Agostinho D, Simões M, Castelo-Branco M. Predicting conversion from mild cognitive impairment to Alzheimer's disease: a multimodal approach. Brain Commun 2024;6(4):fcae208. [doi: 10.1093/braincomms/fcae208] [Medline: 38961871]
48. Chudzik A, Śledzianowski A, Przybyszewski AW. Machine learning and digital biomarkers can detect early stages of neurodegenerative diseases. Sensors (Basel) 2024 Feb 29;24(5):1572. [doi: 10.3390/s24051572] [Medline: 38475108]
49. Olawade DB, David-Olawade AC, Wada OZ, Asaolu AJ, Adereni T, Ling J. Artificial intelligence in healthcare delivery: prospects and pitfalls. Journal of Medicine, Surgery, and Public Health 2024 Aug;3:100108. [doi: 10.1016/j.glmedi.2024.100108]
50. Chakraborty C, Bhattacharya M, Pal S, Lee SS. From machine learning to deep learning: advances of the recent data-driven paradigm shift in medicine and healthcare. Current Research in Biotechnology 2024;7:100164. [doi: 10.1016/j.crbiot.2023.100164]
51. Jamarani A, Haddadi S, Sarvizadeh R, Haghi Kashani M, Akbari M, Moradi S. Big data and predictive analytics: a systematic review of applications. Artif Intell Rev 2024;57(7):176. [doi: 10.1007/s10462-024-10811-5]
52. Soori M, Arezoo B, Dastres R. Artificial intelligence, machine learning and deep learning in advanced robotics, a review. Cognitive Robotics 2023;3:54-70. [doi: 10.1016/j.cogr.2023.04.001]
53. Kakraba S, Yadem AC, Abraham KE. Unraveling protein secrets: machine learning unveils novel biologically significant associations among amino acids. Preprints. Preprint posted online on May 6, 2025. [doi: 10.20944/preprints202505.0139.v1]
54. Yang Z, Zhou H, Srivastav S, et al. Optimizing Parkinson's disease prediction: a comparative analysis of data aggregation methods using multiple voice recordings via an automated artificial intelligence pipeline. Preprints. Preprint posted online on Dec 5, 2024. [doi: 10.20944/preprints202412.0366.v1]
55. Parkinson-s-disease-prediction-code-and-data-repository-main. ResearchGate. 2025. URL: https://www.researchgate.net/publication/387796710_Parkinson-s-Disease-Prediction-Code-and-Data-Repository-main [accessed 2025-11-29]

XSL•FO
RenderX

56. Netsey EK, Naandam SM, Asante Jnr J, et al. Structural and functional impacts of SARS-CoV-2 spike protein mutations: insights from predictive modeling and analytics (Preprint). JMIR Bioinform Biotechnol 2025 Dec 8;6:e73637. [doi: 10.2196/73637] [Medline: 41359941]
57. Kakraba S, Ayyadevara S, Yadem Clement A, Abraham KE, Compadre CM, Shmookler Reis RJ. Machine Learning–Enhanced Quantitative Structure-Activity Relationship Modeling for DNA Polymerase Inhibitor Discovery: Algorithm Development and Validation. JMIR AI 2025 Dec 3;4:e77890. [doi: 10.2196/77890] [Medline: 41340396]
58. Kakraba S, Agyemang EF, Shmookler Reis RJ. Accelerating Discovery of Leukemia Inhibitors with AI-Driven QSAR Modeling (forthcoming). JMIR AI 2025 Dec 5;05/12/2025:81552 [FREE Full text] [doi: 10.2196/81552] [Medline: 41358925]
59. Mahesh B. Machine learning algorithms - a review. International Journal of Science and Research 2020 Jan 5;9(1):381-386. [doi: 10.21275/ART20203995]
60. Netsey EK, Kakraba S, Naandam SM, Yadem AC. A mathematical graph-theoretic model of single point mutations associated with sickle cell anemia disease. JBT 2021;9:1-14. [doi: 10.24297/jbt.v9i.9109]
61. Kakraba S, Knisley D. A graph-theoretic model of single point mutations in the cystic fibrosis transmembrane conductance regulator. JBT 2016;6(1):780-786. [doi: 10.24297/jbt.v6i1.4013]
62. Kakraba S. A hierarchical graph for nucleotide binding domain 2. : East Tennessee State University; 2015 URL: https://dc.etsu.edu/etd/2517/ [accessed 2025-11-29]
63. Ahsan MM, Luna SA, Siddique Z. Machine-learning-based disease diagnosis: a comprehensive review. Healthcare (Basel) 2022 Mar 15;10(3):541. [doi: 10.3390/healthcare10030541] [Medline: 35327018]
64. Shivahare BD, Singh J, Ravi V, et al. Delving into machine learning's influence on disease diagnosis and prediction. The Open Public Health Journal 2024 Apr 26;17(1). [doi: 10.2174/0118749445297804240401061128]
65. Xu D, Xu Z. Machine learning applications in preventive healthcare: a systematic literature review on predictive analytics of disease comorbidity from multiple perspectives. Artif Intell Med 2024 Oct;156:102950. [doi: 10.1016/j.artmed.2024.102950] [Medline: 39163727]
66. Park DJ, Park MW, Lee H, Kim YJ, Kim Y, Park YH. Development of machine learning model for diagnostic disease prediction based on laboratory tests. Sci Rep 2021 Apr 7;11(1):7567. [doi: 10.1038/s41598-021-87171-5] [Medline: 33828178]
67. Dixon D, Sattar H, Moros N, et al. Unveiling the influence of AI predictive analytics on patient outcomes: a comprehensive narrative review. Cureus 2024 May;16(5):e59954. [doi: 10.7759/cureus.59954] [Medline: 38854327]
68. Jia L, Yao W, Jiang Y, et al. Development of interactive biological web applications with R/Shiny. Brief Bioinformatics 2022 Jan 17;23(1). [doi: 10.1093/bib/bbab415]
69. Giorgi FM, Ceraolo C, Mercatelli D. The R language: an engine for bioinformatics and data science. Life (Basel) 2022 Apr 27;12(5):648. [doi: 10.3390/life12050648] [Medline: 35629316]
70. Perkins MA, Carrier JW. Using R Shiny to develop a dashboard using IPEDS, U.S. Census, and bureau of labor statistics data. PLOS ONE 2023;18(1):e0278573. [doi: 10.1371/journal.pone.0278573] [Medline: 36630356]
71. Bastian G, Baker GH, Limon A. Bridging the divide between data scientists and clinicians. Intelligence-Based Medicine 2022;6:100066. [doi: 10.1016/j.ibmed.2022.100066]
72. Singh P. Systematic review of data-centric approaches in artificial intelligence and machine learning. Data Science and Management 2023 Sep;6(3):144-157. [doi: 10.1016/j.dsm.2023.06.001]
73. Cilia ND, De Gregorio G, De Stefano C, Fontanella F, Marcelli A, Parziale A. Diagnosing Alzheimer's disease from on-line handwriting: a novel dataset and performance benchmarking. Eng Appl Artif Intell 2022 May;111:104822. [doi: 10.1016/j.engappai.2022.104822]
74. DARWIN. UCI Machine Learning Repository. URL: https://archive.ics.uci.edu/dataset/732/darwin [accessed 2025-11-29]
75. Nohara Y, Matsumoto K, Soejima H, Nakashima N. Explanation of machine learning models using shapley additive explanation and application for real data in hospital. Comput Methods Programs Biomed 2022 Feb;214:106584. [doi: 10.1016/j.cmpb.2021.106584] [Medline: 34942412]
76. Ponce‐Bobadilla AV, Schmitt V, Maier CS, Mensing S, Stodtmann S. Practical guide to SHAP analysis: explaining supervised machine learning model predictions in drug development. Clinical Translational Sci 2024 Nov;17(11):e70056. [doi: 10.1111/cts.70056]
77. Feng J, Liang J, Qiang Z, et al. A hybrid stacked ensemble and Kernel SHAP-based model for intelligent cardiotocography classification and interpretability. BMC Med Inform Decis Mak 2023 Nov 28;23(1):273. [doi: 10.1186/s12911-023-02378-y] [Medline: 38017460]
78. The R Core Team. R: a language and environment for statistical computing. : R Project; 2024 URL: https://cran.r-project.org/doc/manuals/r-release/fullrefman.pdf [accessed 2025-11-29]
79. Weihs C, Ligges U, Luebke K, Raabe N. KlaR analyzing German business cycles. In: Baier D, Decker R, Schmidt-Thieme L, editors. Data Analysis and Decision Support Studies in Classification, Data Analysis, and Knowledge Organization: Springer; 2005:335-343. [doi: 10.1007/3-540-28397-8_36]
80. Kuhn M. Building predictive models in R using the caret package. J Stat Soft 2008;28(5):1-26. [doi: 10.18637/jss.v028.i05]

XSL•FO
RenderX

81. Therneau T, Atkinson B, Ripley B. rpart: recursive partitioning and regression trees. : R Project; 2023 URL: https://cran.r-project.org/web/packages/rpart/rpart.pdf [accessed 2025-11-29]
82. Liaw A, Wiener M. Classification and regression by randomForest. R News 2002;2(3):18-22 [FREE Full text]
83. Ridgeway G, Edwards D, Kriegler B, et al. gbm: generalized boosted regression models. : R Project; 2024 URL: https://cran.r-project.org/web/packages/gbm/gbm.pdf [accessed 2025-11-29]
84. Culp M, Johnson K, Michailidis G. The R package ada for stochastic boosting. : R Project; 2016 URL: https://cran.r-project.org/web/packages/ada/ada.pdf [accessed 2025-11-29]
85. Chen T, He T. xgboost: extreme gradient boosting. : R Project; 2024 URL: https://cran.r-project.org/web/packages/xgboost/vignettes/xgboost.pdf [accessed 2025-11-29]
86. Venables WN, Ripley BD. Modern Applied Statistics with S, 4th edition: Springer; 2002. [doi: 10.1007/978-0-387-21706-2]
87. Uddin MF. Addressing accuracy paradox using enhanched weighted performance metric in machine learning. Presented at: Sixth HCT Information Technology Trends (ITT); Nov 20-21, 2019. [doi: 10.1109/ITT48889.2019.9075071]
88. Salihoğlu YS, Uslu Erdemir R, Aydur Püren B, et al. Diagnostic performance of machine learning models based on 18F-FDG PET/CT radiomic features in the classification of solitary pulmonary nodules. Mol Imaging Radionucl Ther 2022 Jun 27;31(2):82-88. [doi: 10.4274/mirt.galenos.2021.43760] [Medline: 35770958]
89. Ávila-Jiménez JL, Cantón-Habas V, Carrera-González MDP, Rich-Ruiz M, Ventura S. A deep learning model for Alzheimer's disease diagnosis based on patient clinical records. Comput Biol Med 2024 Feb;169:107814. [doi: 10.1016/j.compbiomed.2023.107814] [Medline: 38113682]
90. Sun Y, Zhang J, Zhang Y. Adaboost algorithm combined multiple random forest models (Adaboost-RF) is employed for fluid prediction using well logging data. Phys Fluids (1994) 2024 Jan 1;36(1). [doi: 10.1063/5.0179422]
91. Hossain S, Hasan MK, Faruk MO, Aktar N, Hossain R, Hossain K. Machine learning approach for predicting cardiovascular disease in Bangladesh: evidence from a cross-sectional study in 2023. BMC Cardiovasc Disord 2024 Apr 18;24(1):214. [doi: 10.1186/s12872-024-03883-2] [Medline: 38632519]
92. Ghasemieh A, Lloyed A, Bahrami P, Vajar P, Kashef R. A novel machine learning model with Stacking Ensemble Learner for predicting emergency readmission of heart-disease patients. Decision Analytics Journal 2023 Jun;7:100242. [doi: 10.1016/j.dajour.2023.100242]
93. Hussain SS, Zaidi SSH. AdaBoost ensemble approach with weak classifiers for gear fault diagnosis and prognosis in DC motors. Appl Sci (Basel) 2024;14(7):3105. [doi: 10.3390/app14073105]
94. Dostmohammadi M, Pedram MZ, Hoseinzadeh S, Garcia DA. A GA-stacking ensemble approach for forecasting energy consumption in a smart household: a comparative study of ensemble methods. J Environ Manage 2024 Jul;364:121264. [doi: 10.1016/j.jenvman.2024.121264] [Medline: 38870783]
95. Sarica A, Cerasa A, Quattrone A. Random forest algorithm for the classification of neuroimaging data in Alzheimer's disease: a systematic review. Front Aging Neurosci 2017;9:329. [doi: 10.3389/fnagi.2017.00329] [Medline: 29056906]
96. Moore PJ, Lyons TJ, Gallacher J, Alzheimer's Disease Neuroimaging Initiative. Random forest prediction of Alzheimer's disease using pairwise selection from time series data. PLoS ONE 2019;14(2):e0211558. [doi: 10.1371/journal.pone.0211558] [Medline: 30763336]
97. Tripoliti EE, Fotiadis DI, Argyropoulou M, Manis G. A six stage approach for the diagnosis of the Alzheimer's disease based on fMRI data. J Biomed Inform 2010 Apr;43(2):307-320. [doi: 10.1016/j.jbi.2009.10.004] [Medline: 19883796]
98. Song M, Jung H, Lee S, Kim D, Ahn M. Diagnostic classification and biomarker identification of Alzheimer's disease with random forest algorithm. Brain Sci 2021 Apr 2;11(4):453. [doi: 10.3390/brainsci11040453] [Medline: 33918453]
99. Metcalf CJE, Tate AT, Graham AL. Demographically framing trade-offs between sensitivity and specificity illuminates selection on immunity. Nat Ecol Evol 2017 Nov;1(11):1766-1772. [doi: 10.1038/s41559-017-0315-3] [Medline: 28963489]
100. Dimitriadis SI, Liparas D, Tsolaki MN. Random forest feature selection, fusion and ensemble strategy: combining multiple morphological MRI measures to discriminate among healhy elderly, MCI, cMCI and Alzheimer's disease patients: from the Alzheimer's disease neuroimaging initiative (ADNI) database. J Neurosci Methods 2018 May 15;302:14-23. [doi: 10.1016/j.jneumeth.2017.12.010] [Medline: 29269320]
101. Shree HV, Mouleeswaran SK. A comprehensive study of predictive models for mortality risk assessment in medical data. Presented at: 4th International Conference on Data Engineering and Communication Systems (ICDECS); Mar 22-23, 2024. [doi: 10.1109/ICDECS59733.2023.10502599]
102. Borkan D, Dixon L, Sorensen J, Thain N, Vasserman L. Nuanced metrics for measuring unintended bias with real data for text classification. Presented at: WWW '19: Companion Proceedings of The 2019 World Wide Web Conference; May 13-17, 2019. [doi: 10.1145/3308560.3317593]
103. Dumitrescu E, Hué S, Hurlin C, Tokpavi S. Machine learning for credit scoring: improving logistic regression with non-linear decision-tree effects. Eur J Oper Res 2022 Mar;297(3):1178-1192. [doi: 10.1016/j.ejor.2021.06.053]
104. Zhang YP, Zhang XY, Cheng YT, et al. Artificial intelligence-driven radiomics study in cancer: the role of feature engineering and modeling. Military Med Res 2023;10(1):22. [doi: 10.1186/s40779-023-00458-8]
105. Zhang S, Pan M. An AutoML-powered analysis framework for forest fire forecasting: adapting to climate change dynamics. Atmosphere (Basel) 2024;15(12):1481. [doi: 10.3390/atmos15121481]

XSL•FO
RenderX

106. Sui Q, Ghosh SK. Active learning for stacking and AdaBoost-related models. Stats 2024;7(1):110-137. [doi: 10.3390/stats7010008]
107. Ahmed SF, Alam M, Hassan M, et al. Deep learning modelling techniques: current progress, applications, advantages, and challenges. Artif Intell Rev 2023 Nov;56(11):13521-13617. [doi: 10.1007/s10462-023-10466-8] [Medline: 37362885]
108. Khalil M, AlSayed A, Liu Y, Vanrolleghem PA. An integrated feature selection and hyperparameter optimization algorithm for balanced machine learning models predicting N2O emissions from wastewater treatment plants. Journal of Water Process Engineering 2024 Jun;63:105512. [doi: 10.1016/j.jwpe.2024.105512]
109. Xu J, Luo Y, Peng K, et al. Supplementary motor area driving changes of structural brain network in blepharospasm. Brain (Bacau) 2023 Apr 19;146(4):1542-1553. [doi: 10.1093/brain/awac341] [Medline: 36130317]
110. Wang L, Zhao J. Artificial intelligence. In: Strategic Blueprint for Enterprise Analytics Studies in Big Data, Vol 150: Springer; 2024:111-138. [doi: 10.1007/978-3-031-55885-6_6]
111. Yadav S, Singh A, Singhal R, Yadav JP. Revolutionizing drug discovery: the impact of artificial intelligence on advancements in pharmacology and the pharmaceutical industry. Intelligent Pharmacy 2024 Jun;2(3):367-380. [doi: 10.1016/j.ipha.2024.02.009]
112. ZainEldin H, Gamel SA, Talaat FM, et al. Silent no more: a comprehensive review of artificial intelligence, deep learning, and machine learning in facilitating deaf and mute communication. Artif Intell Rev 2024;57(7):188. [doi: 10.1007/s10462-024-10816-0]
113. Pinto-Coelho L. How artificial intelligence is shaping medical imaging technology: a survey of innovations and applications. Bioengineering (Basel) 2023 Dec 18;10(12):1435. [doi: 10.3390/bioengineering10121435] [Medline: 38136026]
114. Moore I, Magnante C, Embry E, et al. Doctor AI? A pilot study examining responses of artificial intelligence to common questions asked by geriatric patients. Front Artif Intell 2024;7:1438012. [doi: 10.3389/frai.2024.1438012] [Medline: 39118788]
115. Li R, Wang X, Lawler K, Garg S, Bai Q, Alty J. Applications of artificial intelligence to aid early detection of dementia: a scoping review on current capabilities and future directions. J Biomed Inform 2022 Mar;127:104030. [doi: 10.1016/j.jbi.2022.104030] [Medline: 35183766]
116. Ahuja S, Zaheer S. Advancements in pathology: digital transformation, precision medicine, and beyond. J Pathol Inform 2025 Jan;16:100408. [doi: 10.1016/j.jpi.2024.100408] [Medline: 40094037]
117. Bohr A, Memarzadeh K. Chapter 2 - the rise of artificial intelligence in healthcare applications. In: Bohr A, Memarzadeh K, editors. Artificial Intelligence in Healthcare: Academic Press; 2020:25-60. [doi: 10.1016/B978-0-12-818438-7.00002-2]
118. Hirani R, Noruzi K, Khuram H, et al. Artificial intelligence and healthcare: a journey through history, present innovations, and future possibilities. Life (Basel) 2024 Apr 26;14(5):557. [doi: 10.3390/life14050557] [Medline: 38792579]
119. Iqbal J, Cortés Jaimes DC, Makineni P, et al. Reimagining healthcare: unleashing the power of artificial intelligence in medicine. Cureus 2023 Sep;15(9):e44658. [doi: 10.7759/cureus.44658] [Medline: 37799217]
120. Arya SS, Dias SB, Jelinek HF, Hadjileontiadis LJ, Pappa AM. The convergence of traditional and digital biomarkers through AI-assisted biosensing: a new era in translational diagnostics? Biosens Bioelectron 2023 Sep 1;235:115387. [doi: 10.1016/j.bios.2023.115387] [Medline: 37229842]
121. Singh M, Kumar A, Khanna NN, et al. Artificial intelligence for cardiovascular disease risk assessment in personalised framework: a scoping review. EClinicalMedicine 2024 Jul;73:102660. [doi: 10.1016/j.eclinm.2024.102660] [Medline: 38846068]
122. DeGroat W, Abdelhalim H, Peker E, et al. Multimodal AI/ML for discovering novel biomarkers and predicting disease using multi-omics profiles of patients with cardiovascular diseases. Sci Rep 2024 Nov 3;14(1):26503. [doi: 10.1038/s41598-024-78553-6] [Medline: 39489837]
123. Kapustina O, Burmakina P, Gubina N, Serov N, Vinogradov V. User-friendly and industry-integrated AI for medicinal chemists and pharmaceuticals. Artificial Intelligence Chemistry 2024 Dec;2(2):100072. [doi: 10.1016/j.aichem.2024.100072]
124. Olawade DB, Wada OZ, Ige AO, Egbewole BI, Olojo A, Oladapo BI. Artificial intelligence in environmental monitoring: advancements, challenges, and future directions. Hygiene and Environmental Health Advances 2024 Dec;12:100114. [doi: 10.1016/j.heha.2024.100114]
125. Yelne S, Chaudhary M, Dod K, Sayyad A, Sharma R. Harnessing the power of AI: a comprehensive review of its impact and challenges in nursing science and healthcare. Cureus 2023 Nov;15(11):e49252. [doi: 10.7759/cureus.49252] [Medline: 38143615]
126. Blanco-González A, Cabezón A, Seco-González A, et al. The role of AI in drug discovery: challenges, opportunities, and strategies. Pharmaceuticals (Basel) 2023 Jun 18;16(6):891. [doi: 10.3390/ph16060891] [Medline: 37375838]
127. Kakraba S, Ayyadevara S, Yadem Clement A, Abraham KE, Compadre CM, Reis RJS. Advanced QSAR modeling with machine learning for drug discovery: targeting DNA polymerase inhibitors (Preprint). JMIR AI. Preprint posted online on Aug 16, 2025. [doi: 10.2196/preprints.77890]
128. Ho NTN, Gonzalez P, Gogovi GK. Writing the signs: an explainable machine learning approach for Alzheimer's disease classification from handwriting. Healthc Technol Lett 2025;12(1):e70006. [doi: 10.1049/htl2.70006] [Medline: 39949642]
129. Onofri E, Mercuri M, Donato G, Ricci S. Cognitive fluctuations in connection to dysgraphia: a comparison of Alzheimer's disease with dementia Lewy bodies. Clin Interv Aging 2015;10:625-633. [doi: 10.2147/CIA.S79679] [Medline: 25848239]

XSL•FO
RenderX

130. Fernandes CP, Montalvo G, Caligiuri M, Pertsinakis M, Guimarães J. Handwriting changes in Alzheimer's disease: a systematic review. J Alzheimers Dis 2023;96(1):1-11. [doi: 10.3233/JAD-230438] [Medline: 37718808]
131. Schröter A, Mergl R, Bürger K, Hampel H, Möller HJ, Hegerl U. Kinematic analysis of handwriting movements in patients with Alzheimer's disease, mild cognitive impairment, depression and healthy subjects. Dement Geriatr Cogn Disord 2003;15(3):132-142. [doi: 10.1159/000068484] [Medline: 12584428]
132. Werner P, Rosenblum S, Bar-On G, Heinik J, Korczyn A. Handwriting process variables discriminating mild Alzheimer's disease and mild cognitive impairment. J Gerontol B Psychol Sci Soc Sci 2006 Jul;61(4):228-236. [doi: 10.1093/geronb/61.4.p228] [Medline: 16855035]
133. Lanctôt KL, Boada M, Tariot PN, et al. Association between clinical dementia rating and clinical outcomes in Alzheimer's disease. Alzheimers Dement (Amst) 2024;16(1):e12522. [doi: 10.1002/dad2.12522] [Medline: 38239329]
134. Aggarwal NT, Wilson RS, Beck TL, Bienias JL, Bennett DA. Motor dysfunction in mild cognitive impairment and the risk of incident Alzheimer disease. Arch Neurol 2006 Dec;63(12):1763-1769. [doi: 10.1001/archneur.63.12.1763] [Medline: 17172617]
135. Garre-Olmo J, Faúndez-Zanuy M, López-de-Ipiña K, Calvó-Perxas L, Turró-Garriga O. Kinematic and pressure features of handwriting and drawing: preliminary results between patients with mild cognitive impairment, Alzheimer disease and healthy controls. Curr Alzheimer Res 2017;14(9):960-968. [doi: 10.2174/1567205014666170309120708] [Medline: 28290244]
136. Qi H, Zhang R, Wei Z, et al. A study of auxiliary screening for Alzheimer's disease based on handwriting characteristics. Front Aging Neurosci 2023;15:1117250. [doi: 10.3389/fnagi.2023.1117250] [Medline: 37009455]
137. Kahindo C, El-Yacoubi MA, Garcia-Salicetti S, Rigaud AS, Cristancho-Lacroix V. Characterizing early-stage Alzheimer through spatiotemporal dynamics of handwriting. IEEE Signal Process Lett ;25(8):1136-1140. [doi: 10.1109/LSP.2018.2794500]
138. Zheng R, Zhang Y, Yang Z, Han S, Cheng J. Reduced brain gray matter volume in patients with first-episode major depressive disorder: a quantitative meta-analysis. Front Psychiatry 2021;12:671348. [doi: 10.3389/fpsyt.2021.671348] [Medline: 34276443]
139. Peng X, Srivastava S, Sutton F, Zhang Y, Badran BW, Kautz SA. Compensatory increase in ipsilesional supplementary motor area and premotor connectivity is associated with greater gait impairments: a personalized fMRI analysis in chronic stroke. Front Hum Neurosci 2024;18:1340374. [doi: 10.3389/fnhum.2024.1340374] [Medline: 38487103]
140. Koppelmans V, Silvester B, Duff K. Neural mechanisms of motor dysfunction in mild cognitive impairment and Alzheimer's disease: a systematic review. J Alzheimers Dis Rep 2022;6(1):307-344. [doi: 10.3233/ADR-210065] [Medline: 35891638]
141. Yang J, Yuan C, Chai Y, et al. Dynamic handwriting features for cognitive assessment in inflammatory demyelinating diseases: a machine learning study. Appl Sci (Basel) ;15(11):6257. [doi: 10.3390/app15116257]
142. Yuan J, Au R, Karjadi C, et al. Associations between the digital clock drawing test and brain volume: large community-based prospective cohort (Framingham Heart Study). J Med Internet Res 2022 Apr 15;24(4):e34513. [doi: 10.2196/34513] [Medline: 35436225]
143. Mattsson N, Zetterberg H, Hansson O, et al. CSF biomarkers and incipient Alzheimer disease in patients with mild cognitive impairment. JAMA 2009 Jul 22;302(4):385-393. [doi: 10.1001/jama.2009.1064] [Medline: 19622817]
144. Papaliagkas V, Kalinderi K, Vareltzis P, Moraitou D, Papamitsou T, Chatzidimitriou M. CSF biomarkers in the early diagnosis of mild cognitive impairment and Alzheimer's disease. Int J Mol Sci 2023 May 19;24(10):8976. [doi: 10.3390/ijms24108976] [Medline: 37240322]
145. Morbelli S, Garibotto V, Van De Giessen E, et al. A Cochrane review on brain [¹ F]FDG PET in dementia: limitations and future perspectives. Eur J Nucl Med Mol Imaging 2015 Sep;42(10):1487-1491. [doi: 10.1007/s00259-015-3098-2] [Medline: 26067090]
146. Wittenberg R, Knapp M, Karagiannidou M, Dickson J, Schott JM. Economic impacts of introducing diagnostics for mild cognitive impairment Alzheimer's disease patients. A&D Transl Res & Clin Interv 2019 Jan;5(1):382-387. [doi: 10.1016/j.trci.2019.06.001]
147. Müller S, Preische O, Heymann P, Elbing U, Laske C. Increased diagnostic accuracy of digital vs. conventional clock drawing test for discrimination of patients in the early course of Alzheimer's disease from cognitively healthy individuals. Front Aging Neurosci 2017;9:101. [doi: 10.3389/fnagi.2017.00101] [Medline: 28443019]
148. Palmqvist S, Zetterberg H, Mattsson N, et al. Detailed comparison of amyloid PET and CSF biomarkers for identifying early Alzheimer disease. Neurology (ECronicon) 2015 Oct 6;85(14):1240-1249. [doi: 10.1212/WNL.0000000000001991] [Medline: 26354982]
149. Andreasen N, Minthon L, Davidsson P, et al. Evaluation of CSF-tau and CSF-Abeta42 as diagnostic markers for Alzheimer disease in clinical practice. Arch Neurol 2001 Mar;58(3):373-379. [doi: 10.1001/archneur.58.3.373] [Medline: 11255440]
150. Krishnan K, Rossetti H, Hynan LS, et al. Changes in montreal cognitive assessment scores over time. Assessment 2017 Sep;24(6):772-777. [doi: 10.1177/1073191116654217] [Medline: 27318033]
151. Speiser JL. A random forest method with feature selection for developing medical prediction models with clustered and longitudinal data. J Biomed Inform 2021 May;117:103763. [doi: 10.1016/j.jbi.2021.103763] [Medline: 33781921]

XSL•FO
RenderX

152. Huang BFF, Boutros PC. The parameter sensitivity of random forests. BMC Bioinformatics 2016 Sep 1;17(1):331. [doi: 10.1186/s12859-016-1228-x] [Medline: 27586051]
153. Islam KI, Elias E, Carroll KC, Brown C. Exploring random forest machine learning and remote sensing data for streamflow prediction: an alternative approach to a process-based hydrologic modeling in a snowmelt-driven watershed. Remote Sens (Basel) 2023;15(16):3999. [doi: 10.3390/rs15163999]
154. Acharjee A, Larkman J, Xu Y, Cardoso VR, Gkoutos GV. A random forest based biomarker discovery and power analysis framework for diagnostics research. BMC Med Genomics 2020 Nov 23;13(1):178. [doi: 10.1186/s12920-020-00826-6] [Medline: 33228632]
155. Chen C, Liang J, Sun W, Yang G, Meng X. An automatically recursive feature elimination method based on threshold decision in random forest classification. Geo-spatial Information Science 2025 Jul 4;28(4):1494-1519. [doi: 10.1080/10095020.2024.2387457]
156. Shao Z, Ahmad MN, Javed A. Comparison of random forest and XGBoost classifiers using integrated optical and SAR features for mapping urban impervious surface. Remote Sens (Basel) 2024;16(4):665. [doi: 10.3390/rs16040665]
157. Ali S, Akhlaq F, Imran AS, Kastrati Z, Daudpota SM, Moosa M. The enlightening role of explainable artificial intelligence in medical & healthcare domains: a systematic literature review. Comput Biol Med 2023 Nov;166:107555. [doi: 10.1016/j.compbiomed.2023.107555] [Medline: 37806061]
158. Sadeghi Z, Alizadehsani R, Cifci MA, et al. A review of explainable artificial intelligence in healthcare. Computers and Electrical Engineering 2024 Aug;118:109370. [doi: 10.1016/j.compeleceng.2024.109370]
159. Frasca M, La Torre D, Pravettoni G, Cutica I. Explainable and interpretable artificial intelligence in medicine: a systematic bibliometric review. Discov Artif Intell 2024;4(1):15. [doi: 10.1007/s44163-024-00114-7]
160. Farah L, Murris JM, Borget I, Guilloux A, Martelli NM, Katsahian SIM. Assessment of performance, interpretability, and explainability in artificial intelligence-based health technologies: what healthcare stakeholders need to know. Mayo Clin Proc Digit Health 2023 Jun;1(2):120-138. [doi: 10.1016/j.mcpdig.2023.02.004] [Medline: 40206724]
161. Abgrall G, Holder AL, Chelly Dagdia Z, Zeitouni K, Monnet X. Should AI models be explainable to clinicians? Crit Care 2024 Sep 12;28(1):301. [doi: 10.1186/s13054-024-05005-y] [Medline: 39267172]
162. Ahmad A, Imran M, Ahsan H. Biomarkers as biomedical bioindicators: approaches and techniques for the detection, analysis, and validation of novel biomarkers of diseases. Pharmaceutics 2023 May 31;15(6):1630. [doi: 10.3390/pharmaceutics15061630] [Medline: 37376078]
163. Zhou Y, Tao L, Qiu J, et al. Tumor biomarkers for diagnosis, prognosis and targeted therapy. Signal Transduct Target Ther 2024 May 20;9(1):132. [doi: 10.1038/s41392-024-01823-2] [Medline: 38763973]
164. Khalifa M, Albadawy M. Artificial intelligence for clinical prediction: exploring key domains and essential functions. Computer Methods and Programs in Biomedicine Update 2024;5:100148. [doi: 10.1016/j.cmpbup.2024.100148]
165. Lawler M, Keeling P, Kholmanskikh O, et al. Empowering effective biomarker-driven precision oncology: a call to action. Eur J Cancer 2024 Sep;209:114225. [doi: 10.1016/j.ejca.2024.114225] [Medline: 39053288]
166. Passaro A, Al Bakir M, Hamilton EG, et al. Cancer biomarkers: emerging trends and clinical implications for personalized treatment. Cell 2024 Mar 28;187(7):1617-1635. [doi: 10.1016/j.cell.2024.02.041] [Medline: 38552610]
167. Llitjos JF, Carrol ED, Osuchowski MF, et al. Enhancing sepsis biomarker development: key considerations from public and private perspectives. Crit Care 2024 Jul 13;28(1):238. [doi: 10.1186/s13054-024-05032-9] [Medline: 39003476]
168. Marques L, Costa B, Pereira M, et al. Advancing precision medicine: a review of innovative in silico approaches for drug development, clinical pharmacology and personalized healthcare. Pharmaceutics 2024 Feb 27;16(3):332. [doi: 10.3390/pharmaceutics16030332] [Medline: 38543226]
169. Nadkarni SB, Vijay GS, Kamath RC. Comparative study of random forest and gradient boosting algorithms to predict airfoil self-noise. Presented at: RAiSE-2023 International Conference on Recent Advances in Science and Engineering; Oct 4-5, 2023. [doi: 10.3390/engproc2023059024]
170. Wei Y, Ding J, Yang S, Yang X, Wang F. Comparisons of random forest and stochastic gradient treeboost algorithms for mapping soil electrical conductivity with multiple subsets using Landsat OLI and DEM/GIS-based data at a type oasis in Xinjiang, China. European Journal of Remote Sensing 2021 Jan 1;54(1):158-181. [doi: 10.1080/22797254.2021.1888657]
171. Patri A, Patnaik Y. Random forest and stochastic gradient tree boosting based approach for the prediction of airfoil self-noise. Procedia Comput Sci 2015;46:109-121. [doi: 10.1016/j.procs.2015.02.001]
172. Sahin EK. Assessing the predictive capability of ensemble tree methods for landslide susceptibility mapping using XGBoost, gradient boosting machine, and random forest. SN Appl Sci 2020 Jul;2(7):1308. [doi: 10.1007/s42452-020-3060-1]
173. Naderi Yaghouti AR, Zamanian H, Shalbaf A. Machine learning approaches for early detection of non-alcoholic steatohepatitis based on clinical and blood parameters. Sci Rep 2024 Jan 30;14(1):2442. [doi: 10.1038/s41598-024-51741-0] [Medline: 38287043]
174. Perera LRD, Ganegoda GU. Alzheimer's disease prediction using clinical data approach. Presented at: 8th International Conference on Information Technology Research (ICITR); Dec 7-8, 2023. [doi: 10.1109/ICITR61062.2023.10382737]
175. Asadi B, Hajj R. Prediction of asphalt binder elastic recovery using tree-based ensemble bagging and boosting models. Constr Build Mater 2024 Jan;410:134154. [doi: 10.1016/j.conbuildmat.2023.134154]

XSL•FO
RenderX

176. Nguyen KL, Trinh HT, Pham TM. Prediction of punching shear strength in flat slabs: ensemble learning models and practical implementation. Neural Comput & Applic 2024 Mar;36(8):4207-4228. [doi: 10.1007/s00521-023-09296-0]
177. Nazat S, Abdallah M. XAI-based feature ensemble for enhanced anomaly detection in autonomous driving systems. ArXiv. Preprint posted online on Oct 20, 2024 URL: https://arxiv.org/abs/2410.15405 [accessed 2025-11-29]
178. Mahdizadeh Gharakhanlou N, Perez L. From data to harvest: leveraging ensemble machine learning for enhanced crop yield predictions across Canada amidst climate change. Sci Total Environ 2024 Nov 15;951:175764. [doi: 10.1016/j.scitotenv.2024.175764] [Medline: 39182775]
179. Ascher S, Wang X, Watson I, Sloan W, You S. Interpretable machine learning to model biomass and waste gasification. Bioresour Technol 2022 Nov;364:128062. [doi: 10.1016/j.biortech.2022.128062] [Medline: 36202285]
180. Chattopadhyay T, Ozarkar SS, Buwa K, et al. Comparison of deep learning architectures for predicting amyloid positivity in Alzheimer's disease, mild cognitive impairment, and healthy aging, from T1-weighted brain structural MRI. Front Neurosci 2024;18:1387196. [doi: 10.3389/fnins.2024.1387196] [Medline: 39015378]
181. Jenber Belay A, Walle YM, Haile MB. Deep ensemble learning and quantum machine learning approach for Alzheimer's disease detection. Sci Rep 2024 Jun 20;14(1):14196. [doi: 10.1038/s41598-024-61452-1] [Medline: 38902368]
182. Rosenbacke R, Melhus Å, McKee M, Stuckler D. How explainable artificial intelligence can increase or decrease clinicians' trust in AI applications in health care: systematic review. JMIR AI 2024 Oct 30;3:e53207. [doi: 10.2196/53207] [Medline: 39476365]
183. Markus AF, Kors JA, Rijnbeek PR. The role of explainability in creating trustworthy artificial intelligence for health care: a comprehensive survey of the terminology, design choices, and evaluation strategies. J Biomed Inform 2021 Jan;113:103655. [doi: 10.1016/j.jbi.2020.103655] [Medline: 33309898]
184. Gaube S, Suresh H, Raue M, et al. Non-task expert physicians benefit from correct explainable AI advice when reviewing X-rays. Sci Rep 2023 Jan 25;13(1):1383. [doi: 10.1038/s41598-023-28633-w] [Medline: 36697450]
185. Erdogmus P, Kabakus AT. The promise of convolutional neural networks for the early diagnosis of the Alzheimer's disease. Eng Appl Artif Intell 2023 Aug;123:106254. [doi: 10.1016/j.engappai.2023.106254]
186. Subha R, Nayana B, Selvadass M. Hybrid machine learning model using particle swarm optimization for effectual diagnosis of Alzheimer's disease from handwriting. Presented at: 4th International Conference on Circuits, Control, Communication and Computing; Dec 21-23, 2022. [doi: 10.1109/I4C57141.2022.10057948]
187. Gattulli V, Impedovo D, Pirlo G, Semeraro G. Handwriting task-selection based on the analysis of patterns in classification results on Alzheimer dataset. Presented at: IEEESDS'23: Data Science Techniques for Datasets on Mental and Neurodegenerative Disorders; Jun 22, 2023 URL: https://ceur-ws.org/Vol-3521/paper2.pdf [accessed 2025-11-29]
188. Önder M, Şentürk Ü, Polat K, Paulraj D. Diagnosis of Alzheimer's disease using boosting classification algorithms. Presented at: 2023 International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE); Nov 1-2, 2023. [doi: 10.1109/RMKMATE59243.2023.10369418]
189. Anacur CA, Yilmaz AG, Aykut M. Handwriting analysis for Alzheimer's diagnosis: feature enrichment and machine learning classification. Presented at: 2024 8th International Artificial Intelligence and Data Processing Symposium (IDAP); Sep 21-22, 2024. [doi: 10.1109/IDAP64064.2024.10711121]
190. Humphreys D, Kupresanin A, Boyer MD, et al. Advancing fusion with machine learning research needs workshop report. J Fusion Energ 2020 Aug;39(4):123-155. [doi: 10.1007/s10894-020-00258-1]
191. Olorunnimbe K, Viktor H. Deep learning in the stock market-a systematic survey of practice, backtesting, and applications. Artif Intell Rev 2023;56(3):2057-2109. [doi: 10.1007/s10462-022-10226-0] [Medline: 35791405]
192. Scheidt FV, Medinová H, Ludwig N, Richter B, Staudt P, Weinhardt C. Data analytics in the electricity sector – a quantitative and qualitative literature review. Energy and AI 2020 Aug;1:100009. [doi: 10.1016/j.egyai.2020.100009]
193. Sedlakova J, Daniore P, Horn Wintsch A, et al. Challenges and best practices for digital unstructured data enrichment in health research: a systematic narrative review. PLOS Digit Health 2023 Oct;2(10):e0000347. [doi: 10.1371/journal.pdig.0000347] [Medline: 37819910]
194. Kazemi P, Ghisi A, Entezami A. Artificial intelligence-powered computational strategies in selecting and augmenting data for early design of tall buildings with outer diagrids. Buildings 2024;14(4):1118. [doi: 10.3390/buildings14041118]
195. Moreira A, Ferreira A, Leite N. Prediction of Alzheimer disease on the DARWIN dataset with dimensionality reduction and explainability techniques. Presented at: 1st International Conference on Explainable AI for Neural and Symbolic Methods; Nov 20-22, 2024. [doi: 10.5220/0013017400003886]
196. Singh SK, Chaturvedi A. Leveraging handwriting dynamics, explainable AI and machine learning for Alzheimer prediction. In: Singh JP, Singh MP, Singh AK, Mukhopadhyay S, Mandal JK, Dutta P, editors. Computational Intelligence in Communications and Business Analytics CICBA 2024 Communications in Computer and Information Science, Vol 2366: Springer; 2025:338-351. [doi: 10.1007/978-3-031-81342-9_27]
197. Samrat N, Islam M, Sharmi S, et al. Explainable AI-driven handwriting analysis for early Alzheimer's disease detection using machine learning and ensemble models. Presented at: 2024 27th International Conference on Computer and Information Technology (ICCIT); Dec 20-22, 2024. [doi: 10.1109/ICCIT64611.2024.11021728]
198. SMART-Pred. GitHub. URL: https://github.com/Agyemang1z/SMART-Pred-Shiny-Multi-Algorithm-R-Tool-for-Predictive-Modeling-Manuscript [accessed 2025-12-10]

XSL•FO
RenderX

Abbreviations

AD: Alzheimer disease
AdaBoost: adaptive boosting
ADNI: Alzheimer's Disease Neuroimaging Initiative
ADRD: Alzheimer disease and related dementias
AI: artificial intelligence
AUC: area under the curve
DARWIN: Diagnosis Alzheimer With Handwriting
KNN: k-nearest neighbors
PCA: principal component analysis
ROC: receiver operating characteristic
SHAP: Shapley Additive Explanations
SMART-Pred: Shiny Multi-Algorithm R Tool for Predictive Modeling
SVM: support vector machine
XGBoost: extreme gradient boosting

Edited by J Wang; submitted 18.12.24; peer-reviewed by F Shojaei, G Gogovi, S Mohanadas, X Ding, Y Li; revised version received 29.09.25; accepted 05.11.25; published 30.12.25.

Please cite as:
Wenzheng H, Agyemang EF, Srivastav SK, Shaffer JG, Kakraba S
Artificial Intelligence–Enhanced Multi-Algorithm R Shiny Application for Predictive Modeling and Analytics: Case Study of Alzheimer Disease Diagnostics
JMIR Aging 2025;8:e70272
URL: https://aging.jmir.org/2025/1/e70272
doi:10.2196/70272

© Han Wenzheng, Edmund F Agyemang, Sudesh K Srivastav, Jeffrey G Shaffer, Samuel Kakraba. Originally published in JMIR Aging (https://aging.jmir.org), 30.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Daily-Life Walking Characteristics of Older Adults in Relation to Age, Sex, and Physical Function: the HUNT4 Trondheim 70+ Observational Study

Karoline Blix Grønvik[1], MSc; Anisoara Paraschiv-Ionescu[2], PhD; Gro Gujord Tangen[3,4], PhD; Øyvind Salvesen[5], PhD; Jorunn Lægdheim Helbostad[1], PhD; Nina Skjæret-Maroni[1], PhD; Beatrix Vereijken[1], PhD

[1]Department of Neuromedicine and Movement Science, Norwegian University of Science and Technology, Edvard Griegs Gate 8, Trondheim, Norway

[2]Signal Processing Laboratory 5, Department of Electrical and Electronic Engineering, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

[3]Department of Geriatric Medicine, Oslo University Hospital, Oslo, Norway

[4]Norwegian National Centre for Ageing and Health, Vestfold Hospital Trust, Tønsberg, Norway

[5]Department of Clinical and Molecular Medicine, Norwegian University of Science and Technology, Trondheim, Norway

Corresponding Author:
Karoline Blix Grønvik, MSc
Department of Neuromedicine and Movement Science, Norwegian University of Science and Technology, Edvard Griegs Gate 8, Trondheim, Norway

Abstract

Background: Knowledge about how older adults walk is crucial for the effective prevention and treatment of various mobility issues as well as treatment evaluation, but it is currently largely limited to laboratory-based measurements. Although laboratory-based data provide relevant information about what older adults can do under standardized conditions, they do not provide insight into how they actually walk in their daily life, a gap that needs to be addressed urgently.

Objective: The objective of this study was to describe how older adults walk in daily life, in relation to age, sex, and level of physical function, using wearable sensor data from a large sample of older adults with a wide range of age and function from the HUNT4 Trondheim 70+ study.

Methods: The current study is based on 1-week accelerometer data (Axivity AX3) from 1289 older adults (mean age 77.41, SD 6.06 years; age range 70 - 105 years; n=705, 54.7% women). Physical function was assessed using the Short Physical Performance Battery (SPPB). To investigate the effect of age and SPPB score on gait metrics (daily number of steps, 95th percentile speed, mode speed, 95th percentile cadence, mode cadence, and maximum walking bout [WB] distance) for women and men, univariate gamma regression models with log link were used for each outcome measure, with age and SPPB score in separate models. Sex differences were investigated using Mann-Whitney *U* tests.

Results: Older adults showed a large variation in how and how much they walked in daily life across age, sex, and physical function, particularly younger participants and those with better physical function. Most gait metrics decreased at an increasing rate with higher age, with men maintaining their levels up to higher ages than women. Poorer physical function led to an exponential or close-to-linear decrease in all gait metrics apart from habitual cadence, which remained stable up to a high age. Women had a lower daily number of steps, gait speed, and maximum distance but higher cadence than men (P<.001 for all). On average, 63% of all WBs lasted <10 seconds, corresponding to a median accumulated time of 99 (IQR 66 - 128) minutes. For WBs lasting 10 to 30 seconds, 30 to 60 seconds, and >60 seconds, the median accumulated time was 105 (IQR 65 - 154) minutes, 31 (IQR 18 - 47) minutes, and 113 (IQR 37 - 219) minutes, respectively.

Conclusions: Daily-life walking performance was affected more by functional ability than by age itself, except for the highest ages, and differed significantly between sexes. Although most WBs were very short, the total accumulated walking time in WBs shorter than 30 seconds was longer than that in longer WBs. Future research can build upon our findings by considering both the impact of short WBs and relevant group and sex differences when implementing daily-life mobility assessment in both clinical studies and patient follow-up.

(JMIR Aging 2025;8:e75835) doi:10.2196/75835

KEYWORDS

gait analysis; walking; daily life; real-world; older adults; Trøndelag Health Study; HUNT; Short Physical Performance Battery; SPPB; physical function

XSL•FO
RenderX

Introduction

Mobility, the ability to move around physically, is one of the key features of independence for older adults, contributing to sustained participation and overall quality of life [1,2]. With increasing age, the prevalence of mobility problems increases, with almost a third of people over the age of 75 years having moderate to severe walking difficulties [3], leading to increased dependency on caregivers and assistive equipment [4]. Therefore, an in-depth understanding of how older adults walk can provide valuable insights into factors supporting mobility and independence in an aging population. A better understanding is pivotal for preventing reduced walking ability, thereby sustaining independence, lowering societal costs, and maximizing available resources [5], as well as maintaining their well-being [6].

Most of our current knowledge about how older adults walk is derived from laboratory-based gait assessments, which have provided us with detailed information about how older adults can perform under controlled, supervised conditions. As a result, gait speed has been identified as a predictor of survival [7] and disability [8] among older adults. Older adults who have a preferred gait speed below 1 m/s but are otherwise well functioning have a higher risk of adverse health-related outcomes than those with a higher habitual gait speed [9]. Furthermore, higher age, female sex, and lower physical function are associated with a lower gait speed [10,11]. Additionally, older adults who walk at a cadence of at least 100 steps per minute have a 21% lower risk of all-cause mortality than those with a lower step frequency [12].

However, laboratory-based gait assessments of what older adults can do under standardized conditions are not representative of what they actually do in daily life, and hence gait assessments in the laboratory and in daily life represent 2 different constructs of mobility [13-16]. Laboratory assessments may be biased by factors such as the “white coat effect” or the “Hawthorne effect,” which can lead to overperformance or underperformance or a general change in behavior due to the awareness of being observed [17]. In addition, laboratory assessments are limited to snapshots of a person’s gait, and gait speed measured in the laboratory differs significantly from gait speed in the real world [14,18]. Moreover, real-world cadence is significantly associated with number of falls per year, which has not been shown for cadence, gait speed, or distance measured in a laboratory setting [19]. Therefore, assessment outside the laboratory is needed to gain detailed knowledge of how older adults walk in a setting that reflects the complexity of their daily environments [1]. Such knowledge can provide essential complementary insights into critical factors to prevent functional decline and mobility reduction.

The last two decades have seen a rise and rapid improvement in wearable sensor technology that allows for continuous monitoring outside of a laboratory or clinical setting. These technological advances can provide researchers with fine-grained real-world data in a wide variety of mobility contexts. One of the first studies to use this technology was conducted by Orendurff and colleagues [20], who described daily-life walking patterns of healthy, employed adults. Since then, additional studies have investigated real-world walking using body-worn accelerometers [18,21-23]. Although these earlier studies provided the first insights into real-world gait, most of them had a small sample size [24], did not include the frailest and/or oldest adults [18,21,22], or used algorithms that were not validated specifically for an older population [22,23,25].

Recent developments in algorithms and analysis pipelines specifically developed for older adults and people experiencing mobility problems, along with improved sensor technology, make it possible to account for atypical walking patterns that can occur in these populations. These state-of-the-art algorithms have been shown to reliably estimate various digital mobility outcomes in a real-world environment, both for healthy older adults in general and for patients with mobility problems [26,27]. This paves the way for investigating how older adults walk in daily life in large study samples with large variations in age and level of function. Therefore, the objective of this study was to describe how older adults walk in daily life, in relation to age, sex, and level of physical function, using wearable sensor data from a large sample of older adults with a wide range of age and function from the HUNT4 Trondheim 70+ study.

Methods

Study Population

A total of 1745 adults aged 70 years or older participated in HUNT4 Trondheim 70+, a substudy in the fourth wave of the population-based Trøndelag Health Study (HUNT) in Norway [28]. Data collection took place from October 2018 to June 2019. To be able to recruit the oldest and frailest older adults as well, we offered a home visit to those who could not come to the field station. The majority of participants (1289/1745, 73.9%) were assessed at a field station, with the remaining participants assessed in their own homes (198/1745, 11.3%) or nursing homes (258/1745, 14.8%). Of the 1745 participants included in the Trondheim 70+ substudy, 1363 older adults (aged 70 - 105 years) participated in activity monitoring.

Ethical Considerations

This study was approved by the Regional Committee for Medical and Health Research Ethics (reference number 51105). The participants provided written informed consent. In cases where assessors or health care personnel in nursing homes considered a participant unable to consent, written informed consent was provided by the closest proxy. Participants could withdraw from the study at any moment without providing any reason. All collected data were deidentified to protect participant privacy and confidentiality. No compensation was provided apart from an offer of transportation to the field station if needed.

Procedures

All participants in HUNT4 Trondheim 70+ underwent a standardized study protocol, including clinical tests, structured interviews, and self-reported questionnaires, which has been described in detail previously [29,30]. Height and body weight were measured as part of the clinical tests, and BMI

XSL•FO
RenderX

(weight/height²) was calculated retrospectively. Cognitive function was assessed using the Montreal Cognitive Assessment (MoCA) [31], generating a score of 0 to 30, where higher scores indicated better cognitive performance. Physical function was assessed using the Short Physical Performance Battery (SPPB), consisting of 3 subtests targeting static balance (feet together and in semi-tandem and tandem stance), gait speed (4 m walk test), and leg strength (5 times sit-to-stand test). Each subtest was scored with a value between 0 and 4, adding up to a total score of 0 to 12 points, where higher scores indicated better performance [32].

After the clinical tests, participants were asked to wear 2 triaxial accelerometer devices (Axivity AX3), one on their thigh (centrally on the thigh, approximately 10 cm above the upper border of the patella) and one on the lower back (on L3, the third lumbar segment), for 7 days. Devices were configured using the OmGui software (version 1.0.0.43; Open Movement) and fixed on the skin using adhesive film (Opsite Flexifix; Smith & Nephew). Acceleration was measured at a frequency of 50 Hz and a range of ±8*g*. After the measurement period, devices were delivered back to the field station by the participant or a team member, where the data were downloaded using the OmGui software. Data from the lower back sensor were used in the analyses in this study.

Data Processing

Of the 1363 participants who contributed activity monitoring data, 16 (1.2%) had a data collection registration error, 20 (1.5%) were missing data from the lower back sensor, and 31 (2.3%) had a data processing error. In addition, age and/or SPPB score were missing for 7 participants (0.5%). Data from the remaining 1289 participants (94.6%) were included in further analyses. First, the raw acceleration signal was resampled from 50 Hz to 100 Hz using interpolation [33], for compatibility with the subsequent algorithms. To distinguish between wearing and non-wearing of the sensor, we used an event-based algorithm that detects changes in the temperature signal from the Axivity AX3 device [34]. The number of days with accelerometry data ranged from 1 to 7. For 88.3% (1138/1289) of the participants, this value was 6 or 7 days; for 7.0% (90/1289), 4 or 5 days; and for less than 4.7% (61/1289), 3 days or less. Data processing was performed using MATLAB (version R2022a; MathWorks, Inc). The 456 participants excluded from the analyses (either because they never wore a body-fixed sensor or for the reasons mentioned above) were older (mean 82.2, SD 8.2 years versus mean 77.4, SD 6.1 years), had a lower MoCA score (mean 19.8, SD 6.9 versus 23.5, SD 4.2), and had a higher proportion of women (66.9% versus 54.7%) than those in the analyzed sample.

Real-World Walking Analysis

The data processing pipeline included validated, top-ranked algorithms [26,27] associated with the following processing blocks: foot initial contact detection; walking bout (WB) detection; and estimation of cadence, step length, speed, and distance. The validity criteria were a 95% CI intraclass correlation coefficient threshold of ≥0.7 for performance metrics (eg, sensitivity, positive predictive value, and accuracy) and a relative error of <20%, as described by Micó-Amigo et al [27] and Kirk et al [26]. The *foot initial contact detection* algorithm used the preprocessed vertical acceleration signal. Preprocessing included detrending, low-pass filtering, numerical integration, and continuous wavelet transformation. The initial contact events were identified as the positive peaks between zero-crossings [27]. For *WB detection*, the norm of the triaxial acceleration signal was first detrended and low-pass filtered (finite impulse response, cutoff frequency=3.2 Hz); then, steps-related peaks in the acceleration signal were selected using adaptive thresholds for amplitude and peak-to-peak duration [35,36]. A WB was identified as a sequence of 2 or more consecutive steps (potentially including steps taken while turning), allowing 1.5 seconds plus the average step duration between consecutive steps [35], to ensure accurate estimation of gait metrics (eg, speed and cadence) and reduce bias. *WB duration* (s) was defined as the time interval from the first to the last foot initial contact of the WB. *Cadence* (steps/min) estimation was based on the foot initial contact detection algorithm and was calculated for each detected WB as the step frequency during the respective WB [27]. The *step length* (m) estimation algorithm was based on the inverted pendulum biomechanical model proposed by Zijlstra et al [27,37,38], using a correction factor in the model for impaired gait patterns. *Speed* (m/s) was estimated as the product of cadence and step length. For each WB, step length and speed were obtained as averaged values across all steps within the WB. Finally, *WB distance* (m) was estimated as the product of WB duration and WB speed.

Data Aggregation and Gait Outcome Metrics

As the parameters characterizing real-world walking were estimated for each detected WB throughout the monitoring period, a data aggregation procedure was required to summarize these WB parameters as gait outcome variables for each participant before subsequent statistical analysis. To this end, for each participant, the WBs and the respective parameters were aggregated from (1) all WBs across valid wear days (daily aggregation) or (2) all WBs across the entire measurement period (weekly aggregation). *Daily number of steps* was obtained by summing the steps within all detected WBs across all valid days and then averaging across all valid days. Valid days were defined as days having at least 18 hours of detected wear time. The weekly aggregation procedure was based on various statistical metrics that describe the distribution of walking parameters across pooled WBs. The *within-subject mode* values of cadence and speed were used to represent habitual walking. The *within-subject 95th percentile* values of cadence and speed, referred to as fast cadence and fast walking speed, respectively, were used to represent walking capacity. *Maximum WB distance* was defined as the longest distance covered in a single WB across all measurement days per participant. Additional gait outcome variables derived for weekly aggregated WBs were as follows: (1) *WB duration distribution*, defined as the proportion of WBs in each of the 4 duration categories (≤10 s, >10 - 30 s, >30 - 60 s, >60 s), and (2) *time in WB duration categories*, defined as the accumulated time in each of the 4 WB duration categories.

Cohort Stratification Criteria

Participants were categorized on the basis of age and SPPB score to facilitate the graphical representation of group variance.

XSL•FO
RenderX

They were divided into 6 groups according to age (70.0 - 74.9, 75.0 - 79.9, 80.0 - 84.9, 85.0 - 89.9, 90.0 - 94.9, and ≥95.0 years) and 3 groups according to SPPB score (range=0 - 12 points; high=12 - 10 points, moderate=9 - 7 points, and low=6 - 0 points) [39].

Statistical Analysis

Statistical analyses were performed using R software (version 4.4.3; R Foundation for Statistical Computing) and RStudio (version 2024.04.1+748 Chocolate Cosmos; Posit Software, PBC). Participant characteristics were summarized using mean and SD for continuous variables and frequencies and percentages for categorical variables. The normality distribution of dependent variables was assessed visually using histograms and QQ plots and statistically using the Shapiro-Wilk test and skewness and kurtosis tests. To assess whether participant characteristics and gait metrics differed significantly by sex, we applied the nonparametric Mann-Whitney *U* test. A χ^2 test of independence was used to test whether the distribution of SPPB score differed between sexes.

To investigate the effect of age and SPPB score on gait metrics (daily number of steps, 95th percentile speed, mode speed, 95th percentile cadence, mode cadence, and maximum WB distance), univariate gamma regression models with log link were used for each outcome measure (one aggregated value per participant), with age and SPPB score in separate models. Models using gamma family and log link were chosen after inspecting residual distributions and model fit to account for the data distribution, and age and SPPB score were treated as continuous variables. The SPPB score was reversed to use the maximum score of 12 points as the reference. To prevent disproportionate influence on the results, single representatives of a specific age (2 women aged 97 and 99 years, respectively, and 2 men aged 93 and 100 years, respectively) were excluded from the gamma regression models. A power law relationship was observed between several log gait metrics and age or SPPB score. Therefore, the exponent providing the best fit was determined using maximum likelihood estimation. When the Akaike information criterion did not improve, the exponent was fixed at 1. Regression analyses were performed separately for women and men.

The relationship between each gait metric and age was assessed using gamma regression models assuming

$$\text{gait metric}=\exp(\beta 0+\beta 1(\text{age}-70)\beta p)$$

with a gamma distribution with mean 1 and variance φ for the error term .

The relationship between each gait metric and the SPPB score was assessed using gamma regression models assuming

$$\text{gait metric}=\exp(\beta 0+\beta 1(12-\text{score})\beta p)$$

with a gamma distribution with mean 1 and variance φ for the error term .

Results

A total of 1289 participants (n=705, 54.7% women) were included in the analyses. The mean age of the participants was 77.41 (SD 6.06) years, ranging from 70 to 105 years. Significant differences were found between sexes in age, height, weight, BMI, and overall SPPB score (see Table 1 for the respective *P* values). Furthermore, the distribution of SPPB categories differed significantly between women and men (χ^2_2=8.34, *P*=.02), with a higher proportion of men in the high SPPB score group. Additional participant characteristics and group frequencies are presented in Table 1.

XSL•FO
RenderX

Table . Participant characteristics and group frequencies.

	Total	Women	Men	*P* value[a]
Participants, n (%)	1289 (100)	705 (54.7)	584 (45.3)	—[b]
Age (years), mean (SD) [range]	77.41 (6.06) [70.1 - 105.4]	77.96 (6.45)	76.74 (5.49)	.005
Height (cm), mean (SD) [range]	168.80 (9.26) [141-199]	162.70 (6.36)	176.10 (6.49)	<.001
Weight (kg), mean (SD) [range]	76.20 (14.81) [36.4 - 195.7]	69.90 (12.72)	83.74 (13.56)	<.001
BMI (kg/m²), mean (SD) [range]	26.71 (4.43) [13.4 - 65.4]	26.45 (4.68)	27.02 (4.09)	.002
MoCA[c] score, mean (SD) [range]	23.50 (4.19) [3-30]	23.56 (4.45)	23.43 (3.87)	.10
SPPB[d] score, mean (SD) [range]	9.96 (2.94) [0 - 12]	9.71 (3.09)	10.27 (2.71)	<.001
Group frequency, n (%)				
SPPB[d] category				.02
High	923 (71.6)	482 (68.4)	441 (75.5)	
Moderate	203 (15.8)	121 (17.2)	82 (14)	
Low	163 (12.7)	102 (14.5)	61 (10.6)	
Age category (years)				—
70.0 - 74.9	588 (45.6)	310 (44)	278 (47.6)	
75.0 - 79.9	341 (26.4)	174 (24.7)	167 (28.6)	
80.0 - 84.9	188 (14.6)	110 (15.6)	78 (13.4)	
85.0 - 89.9	104 (8.1)	65 (9.2)	39 (6.7)	
90.0 - 94.9	54 (4.2)	35 (5)	19 (3.2)	
≥95.0	14 (1.1)	11 (1.6)	3 (0.5)	

[a]Derived from Mann-Whitney *U* tests and χ^2 test of independence.

[b]Not applicable.

[c]MoCA: the Montreal Cognitive Assessment. Generates a score of 0-30, where higher scores indicate better cognitive performance.

[d]SPPB: Short Physical Performance Battery. Generates a score of 0-12, where higher scores indicate better physical performance. high=12-10 points, moderate=9-7 points, low=6-0 points.

Number and Characteristics of WBs

The mean number of daily WBs was 253.46 (SD 123.63), ranging from 2.25 to 865.5 WBs. There was no significant difference between women and men in number of WBs (*P*=.47). Across participants, on average, 63% of all WBs had a duration of ≤10 seconds, while only 6% of all WBs lasted >30 seconds, independent of age, sex, and physical function. The median total accumulated time spent walking in WBs of <10 seconds was 99 (IQR 66 - 128) minutes, and it was 105 (IQR 65 - 154) minutes for WBs lasting between 10 and 30 seconds. Furthermore, WBs lasting between 30 and 60 seconds and >60 seconds had a median accumulated time of 31 (IQR 18 - 47) minutes and 113 (IQR 37 - 219) minutes, respectively. Of the 1289 participants, only 1 participant (0.1%) did not have WBs of >10 seconds, 35 participants (2.7%) did not have WBs of >30 seconds, and 83 participants (6.4%) did not have WBs of >60 seconds.

For women, WBs lasting <10 seconds accounted for 66.3% of all WBs, adding up to a median total accumulated time of 101 (IQR 65 - 133) minutes. In contrast, men had 59.2% WBs of <10 seconds, with a median total accumulated time of 96 (IQR 66 - 125) minutes. Bouts lasting between 10 and 30 seconds accounted for 28% of WBs for women and 33.6% for men, with a median total accumulated time of 92 (IQR 58 - 137) minutes and 120 (IQR 78 - 178) minutes, respectively. Furthermore, 2.7% of women's and 4.0% of men's WBs lasted between 30 and 60 seconds, with a median total accumulated time of 27 (IQR 14 - 39) minutes for women and 37 (IQR 24 - 55) minutes for men. Only 2.3% of women's WBs and 2.8% of men's WBs lasted longer than 60 seconds, with a median total accumulated time of 96 (IQR 29 - 210) minutes and 134 (IQR 54 - 227) minutes, respectively.

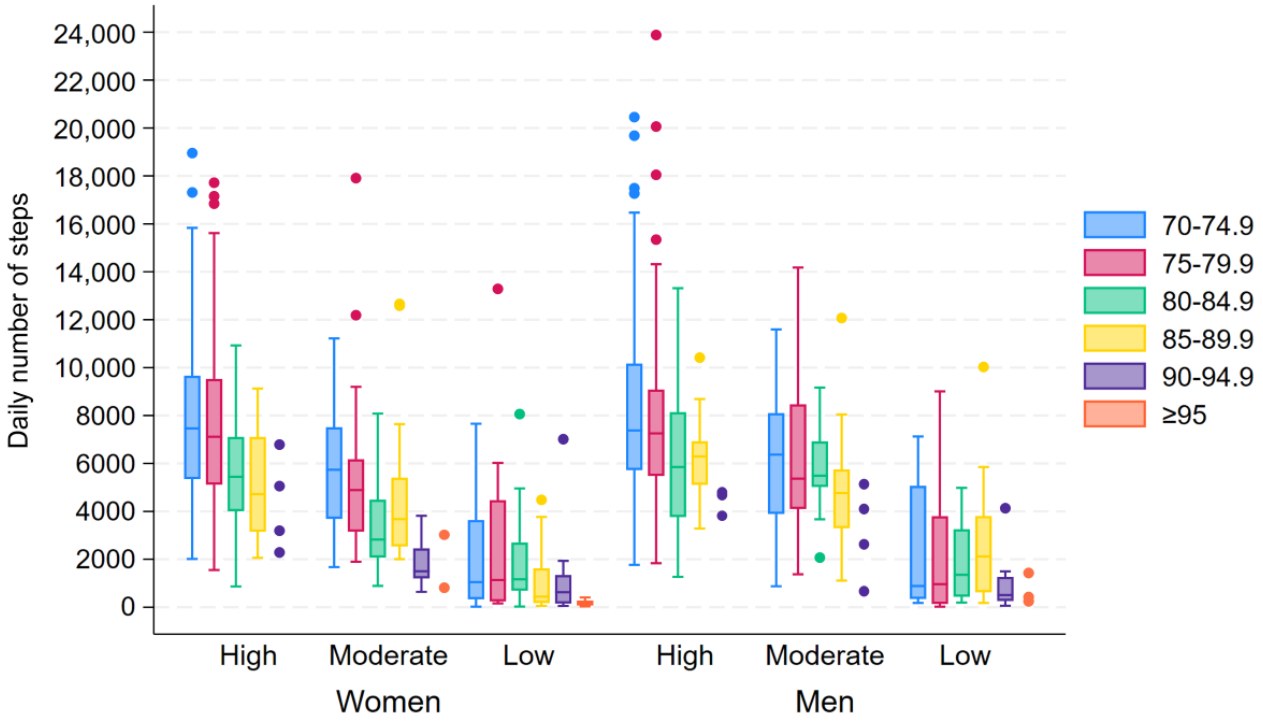
Daily Number of Steps

The average number of steps per day varied widely, particularly in younger age groups and higher SPPB score groups (Figure

XSL•FO
RenderX

1). Overall, women took significantly fewer daily steps than men, with the medians being 5704.5 (IQR 3441.6 - 8264.8) and 6504.5 (IQR 4442.3 - 8837.0), respectively (*P*<.001). Furthermore, participants with higher age and lower SPPB scores had fewer daily steps.

Figure 1. Box plots of daily number of steps, showing the median and IQR (whiskers) of daily steps across all valid days, stratified by age group (years), sex, and SPPB score group (high=12 - 10 points, moderate=9 - 7 points, low=6 - 0 points). For groups with fewer than 5 participants, individual data points are displayed. SPPB: Short Physical Performance Battery.



The daily number of steps decreased with higher age and lower SPPB scores for both women and men. For women, results from the gamma regression analyses examining the effect of age on the daily number of steps predicted a gradual decrease until approximately 77 years of age and then a decrease at an accelerating rate with advancing age (detailed results from all gamma regression models are presented in Tables S1-S4 in Multimedia Appendix 1). In comparison, men were predicted to show a slower decrease in the daily number of steps up to approximately 80 years of age, after which the rate of decrease became more prominent with increasing age, eventually exceeding the rate of decrease for women. The predicted decrease in the daily number of steps as age increased from 70 to 71 years was 0.06% for women and 0.006% for men, with the rate of decrease accelerating to 5.6% and 3.6% with an increase in age from 80 to 81 years and 14.8% and 15.2% with an increase in age from 90 to 91 years for women and men, respectively.

Likewise, the gamma regression analyses predicted an accelerating rate of decrease in the daily number of steps with poorer SPPB scores, although the acceleration was less prominent than that observed for age. Women showed a larger decrease with decreasing SPPB scores already in the highest score range, while men showed a slower decrease in the daily number of steps in the higher SPPB score range before the rate of decrease exceeded women's rate of decrease in the lowest SPPB score range. The predicted decrease in the daily number of steps with a reduction in SPPB score from 12 to 11 points was 5.45% for women and 1.19% for men, with the decrease escalating to 23.4% and 20.9% with a score reduction from 6 to 5 points and 34.6% and 44.9% with a score reduction from 1 to 0 points, respectively.

Walking Speed

Similar to the findings regarding the daily step count, there were large variations in both fast and habitual walking speeds, particularly in the younger age groups and higher SPPB score groups (see Figure 2 for fast speed and Figure 3 for habitual speed). Furthermore, both fast and habitual speeds decreased with increasing age and poorer SPPB scores, and women generally walked slower than men. For women and men, the median fast walking speed was 1.08 (IQR 0.91 - 1.24) m/s and 1.19 (IQR 1.03 - 1.35) m/s, respectively (*P*<.001), and the median habitual walking speed was 0.80 (IQR 0.68 - 1.01) m/s and 0.91 (IQR 0.76 - 1.12) m/s, respectively (*P*<.001).

XSL•FO
RenderX

Figure 2. Box plots of fast walking speed, showing the between-subject median and IQR of the within-subject 95th percentiles of gait speed (m/s), stratified by age group (years), sex, and SPPB score group (high=12 - 10 points, moderate=9 - 7 points, low=6 - 0 points). For groups with fewer than 5 participants, individual data points are displayed. SPPB: Short Physical Performance Battery.

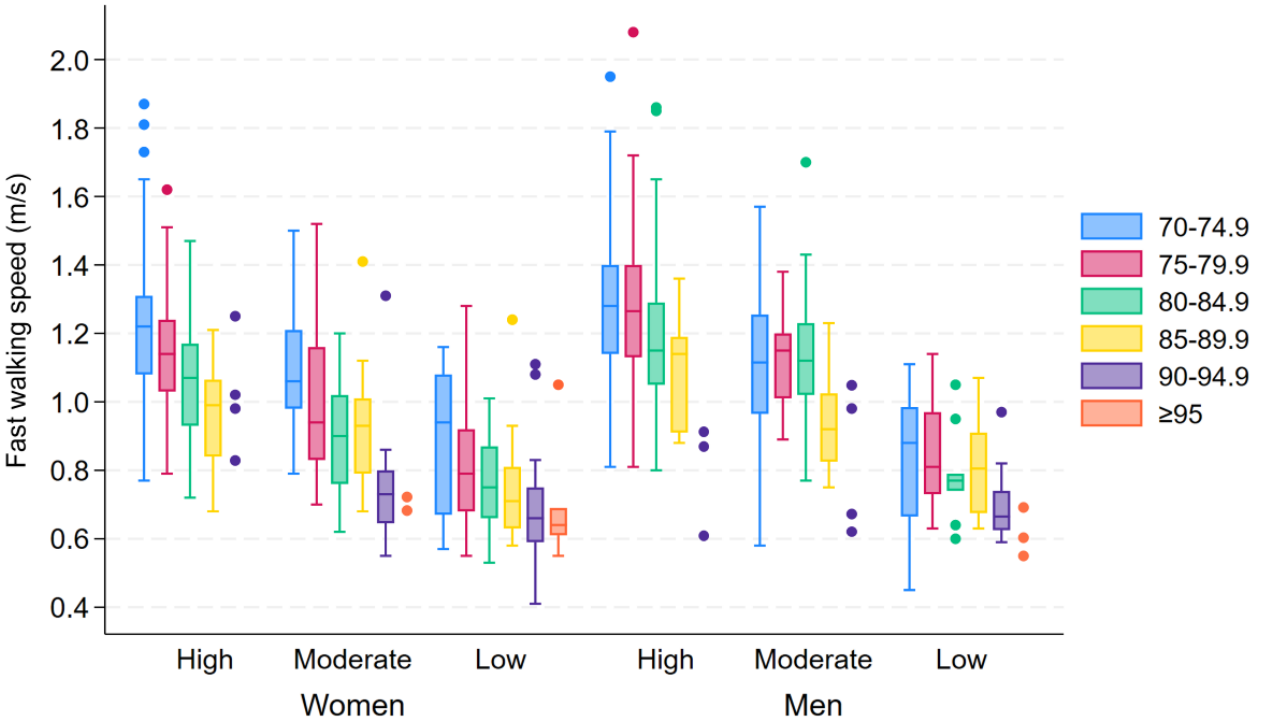
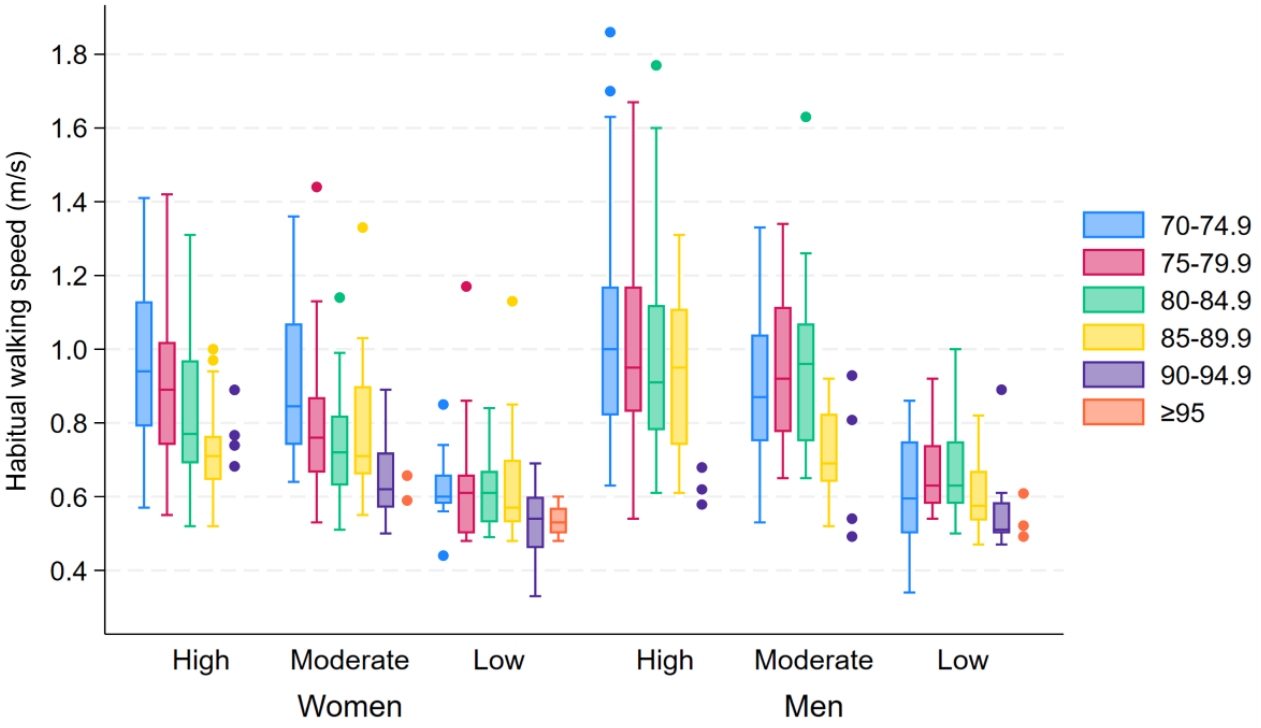


Figure 3. Box plots of habitual walking speed, showing the between-subject median and IQR of the within-subject modes of gait speed (m/s), stratified by age group (years), sex, and SPPB score group (high=12 - 10 points, moderate=9 - 7 points, low=6 - 0 points). For groups with fewer than 5 participants, individual data points are displayed. SPPB: Short Physical Performance Battery.



Overall, both fast and habitual walking speeds were predicted to decrease at an accelerating rate with age. However, both fast and habitual speeds were predicted to decrease evenly with age for women but to have a slow initial decline and steeper decline with higher age for men, similar to the predicted decline in daily number of steps with age. In other words, women experienced a more prominent decline at an earlier age than men, the latter showing a slower rate of decline until approximately 80 years

XSL•FO
RenderX

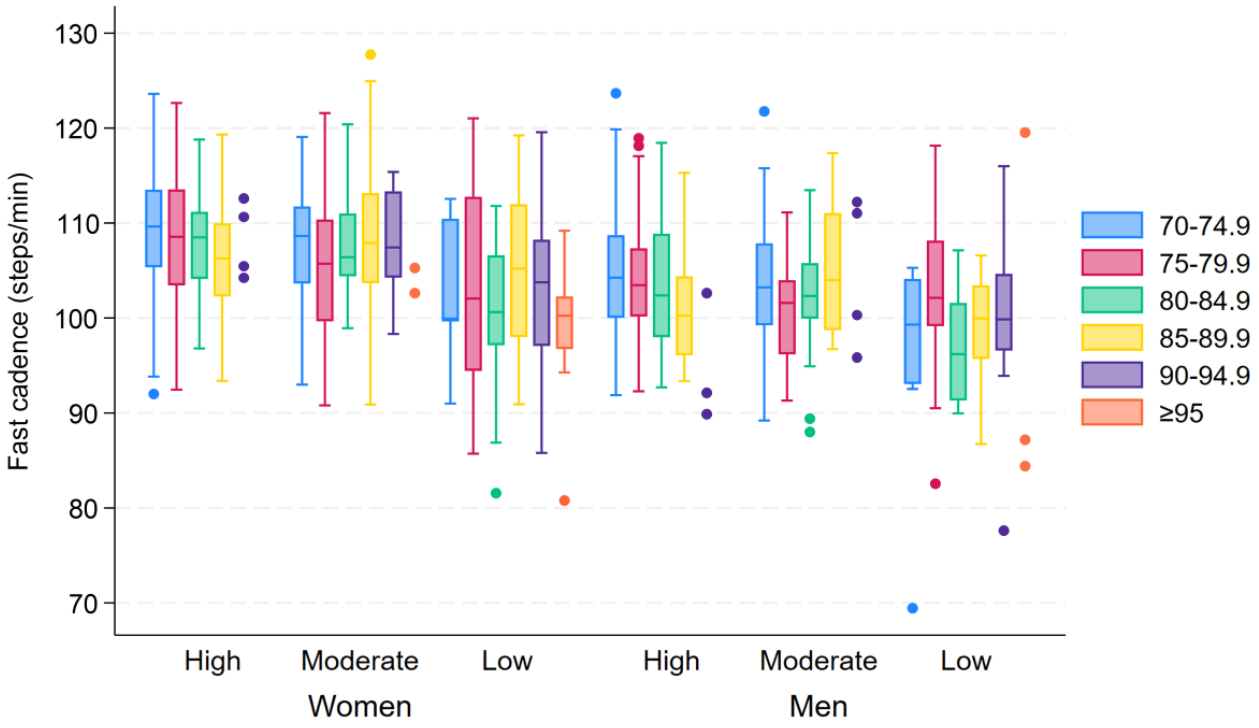
of age. With an increase in age from 70 to 71 years, women's fast and habitual speeds were predicted to decrease by 0.97% and 1.13%, respectively, with the decrease in fast and habitual speeds accelerating to 2.16% and 2.20% with an increase from 80 to 81 years and 2.55% and 2.51% with an increase from 90 to 91 years, respectively. In contrast, with an increase in age from 70 to 71 years, men were predicted to have an initial decrease of only 0.02% for fast speed and 0.01% for habitual speeds, with the decline accelerating to 1.88% and 1.69% with an increase from 80 to 81 years and 5.12% and 5.42% with an increase from 90 to 91 years, respectively.

With a decrease in SPPB score, an exponential decrease was predicted in both fast and habitual speeds for both sexes. In other words, results indicated a rapid initial decrease in gait speed as the score decreased from 12 points, with a diminishing rate of decrease in speed further down the score scale. Per 1-point decrement in SPPB score, women's fast speed and habitual speed were predicted to decrease by 5.06% and 5.15%, respectively. For men, the predicted decrease per 1-point decrement in SPPB score was 5.44% for fast speed and 5.29% for habitual speed.

Cadence

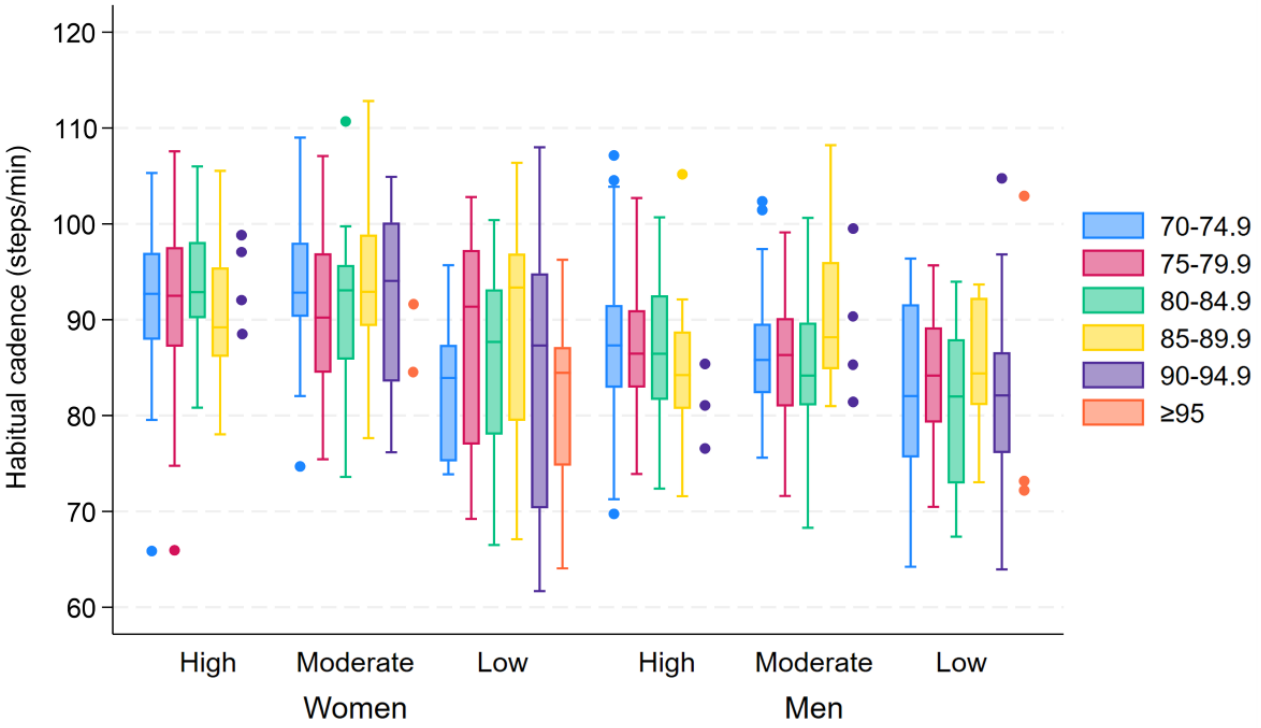
Although there was an overall large variation in cadence within all age and SPPB groups, both fast and habitual cadence remained stable up to higher ages than daily number of steps and speed. Furthermore, women had an overall higher cadence compared to men: median fast cadence 107.9 (IQR 103.3 - 112.2) steps/min and 103.3 (IQR 99.1 - 107.6) steps/min, respectively (P<.001; Figure 4), and median habitual cadence 92.1 (IQR 86.8 - 97) steps/min and 86.4 (IQR 82.4 - 91.1) steps/min, respectively (P<.001; Figure 5).

Figure 4. Box plots of fast cadence, showing the between-subject median and IQR of the within-subject 95th percentile of cadence (steps/min), stratified by age group (years), sex, and SPPB score group (high=12 - 10 points, moderate=9 - 7 points, low=6 - 0 points). For groups with fewer than 5 participants, individual data points are displayed. SPPB: Short Physical Performance Battery.



XSL•FO
RenderX

Figure 5. Box plots of habitual cadence, showing the between-subject median and IQR of the within-subject mode of cadence (steps/min), stratified by age group (years), sex, and SPPB score group (high=12 - 10 points, moderate=9 - 7 points, low=6 - 0 points). For groups with fewer than 5 participants, individual data points are displayed. SPPB: Short Physical Performance Battery.



With increasing age, fast cadence was predicted to decrease slowly for both women and men, with a decrease of 0.20% and 0.25% per 1-year increase in age, respectively. Results indicated virtually no decrease in habitual cadence until the age of approximately 85 to 90 years for both women and men, followed by a steep decline after the age of 90 years. With an increase in age from 70 to 71 years, habitual cadence decreased by <0.001% for both women and men, with the rate of decrease accelerating to 0.04% and 0.003% with an increase in age from 80 to 81 years and 0.88% and 1.18% with an increase in age from 90 to 91 years, respectively.

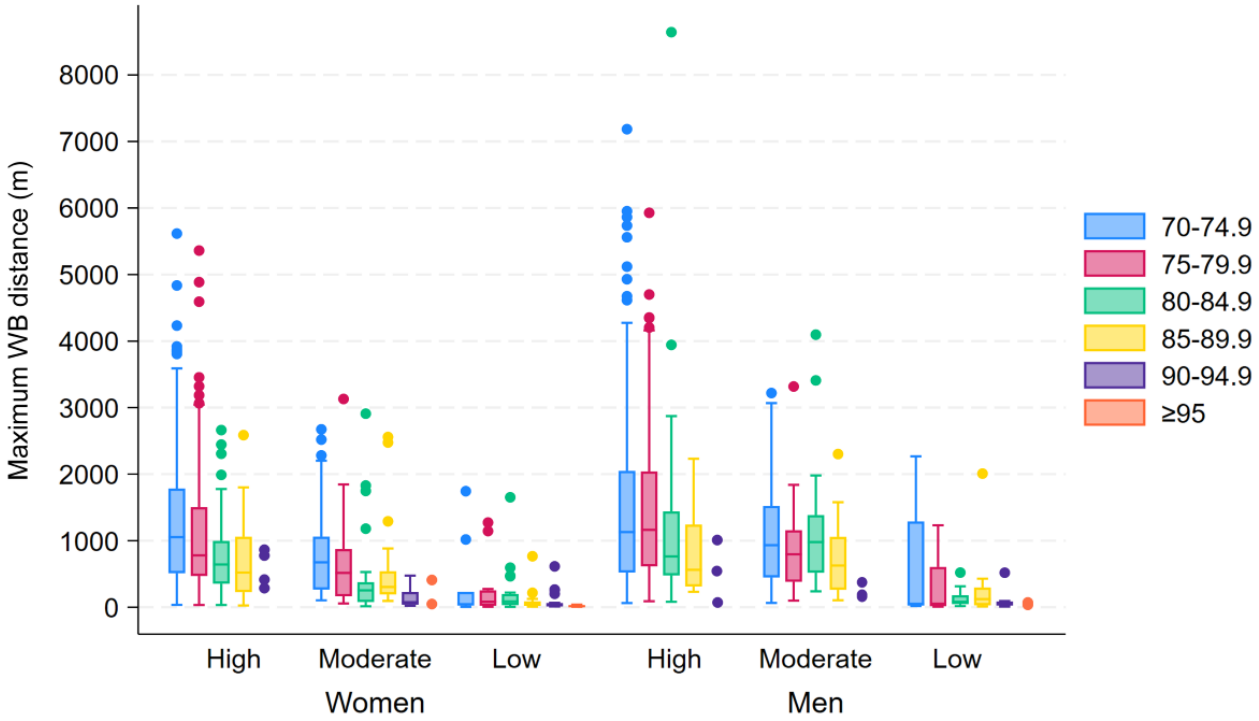
With a decrease in SPPB score, fast cadence was predicted to decrease exponentially for both women and men, with an overall larger effect in women than in men. For every 1-point decrement in the SPPB score, fast cadence was predicted to decrease by 0.69% for women and 0.50% for men. However, habitual cadence was predicted to decrease at an accelerating rate for every 1-point decrement in SPPB score for both sexes. In other words, the habitual cadence remained more stable at higher SPPB scores before decreasing at an accelerating rate toward the lowest SPPB scores, particularly for men. The rate of decrease in habitual cadence with a reduction in SPPB score from 12 to 11 points was 0.18% for women and 0.02% for men, with the rate of decrease accelerating to 0.98% and 0.62% for a score reduction from 6 to 5 points and 1.64% and 1.77% for a reduction from 1 to 0 points, respectively.

Maximum WB Distance

The longest distance covered in a single WB showed a pattern similar to that observed for daily number of steps and speed, with large variations, particularly in the younger age groups and in the higher SPPB score groups (Figure 6). Additionally, individuals with higher SPPB scores and younger age showed an overall longer maximum WB distance, and women generally covered a shorter maximum WB distance than men, with the median (IQR) values being 615.4 (229.6 - 1332.6) m and 869.8 (365.6 - 1724.1) m, respectively (P<.001).

XSL•FO
RenderX

Figure 6. Box plots of maximum walking bout (WB) distance, showing the between-subject median and IQR of the within-subject maximum WB distance (m), stratified by age group (years), sex, and SPPB score group (high=12 - 10 points, moderate=9 - 7 points, low=6 - 0 points). For groups with fewer than 5 participants, individual data points are displayed. SPPB: Short Physical Performance Battery.



In women, the maximum WB distance was predicted to initially decrease evenly, before stabilizing at a median between 10 and 40 m after the age of 90 years. Interestingly, although men's maximum WB distance initially decreased slowly, the rate of change subsequently accelerated from the age of around 80 years, ultimately leading to men having a steeper decrease at older ages than women. However, men's maximum WB distance did not fall below a median of 40 m until the age of 93 years. The predicted decrease in maximum WB distance with an increase in age from 70 to 71 years was 0.28% for women and 0.01% for men, with the decline accelerating to 8.14% and 5.13% with an increase in age from 80 to 81 years and 16.47% and 20.29% with an increase in age from 90 to 91 years, respectively.

As Figure 6 shows, both women and men had lower maximum WB distances with lower SPPB scores. For women, a relatively even decrease was predicted with lower SPPB scores, while men were predicted to maintain a higher maximum WB distance to a larger degree despite lower SPPB scores, dropping to low WB distances only in the highest age groups. The gamma regression analysis predicted a decrease in maximum WB distance of 11.74% for women and 2.97% for men when the SPPB score decreased from 12 to 11 points, with the rate of decrease accelerating to 28.56% and 26.26% with a decrease in score from 6 to 5 points and 36.11% and 46.16% with a decrease in score from 1 to 0 points, respectively.

Discussion

Principal Findings

This study investigated the daily-life walking performance of older adults, focusing on age, sex, and level of physical function. With comprehensive data from a large sample covering a wide range of age and function, our results bring valuable new insights into how older adults walk outside a laboratory or clinical setting. Overall, there was a large variation between older adults in how and how much they walked in daily life, particularly among younger participants and those with better physical function. Despite this variability, the results revealed clear patterns in daily-life walking performance and mobility among older adults, with notable differences between men and women. While women generally had a higher cadence and shorter WBs, men had a longer maximum WB distance. Across both sexes, higher age and lower physical function were consistently associated with reductions in the number of daily steps, walking speed, and maximum WB distance. These declines intensified at an advanced age, highlighting the nonlinear decline in mobility in later life. Surprisingly, older adults maintained their habitual cadence up to high ages. Finally, the total walking time in WBs shorter than 30 seconds exceeded that in bouts longer than 30 seconds, indicating that, on average, older adults accumulate the majority of their walking time in short WBs.

Daily-Life Walking in Relation to Age, Sex, and Physical Function

Consistent with earlier research, daily-life walking predominantly consisted of very short WBs [20,40]. Our results

XSL•FO
RenderX

built further on this finding by highlighting interesting sex differences in walking behavior. Women tended to engage in shorter WBs and accumulated more time in these brief episodes, while men generally walked more and accumulated more time in longer bouts. Although women generally had fewer daily steps than men, this was not the case in the youngest age group, which aligns with previous findings in the age groups 45 to 79 years [21] and 60 to 79 years [22]. Across all ages, women walked at slower speeds, consistent with previous findings from both in-laboratory [11,41,42] and daily life measurements [21,22,43]. Furthermore, the maximum WB distance was shorter for women, which stands in contrast to the findings of Kiselev et al [22], who reported longer maximum WB distances for women. This discrepancy may be partly due to the younger cohort (60 - 79 years) in Kiselev et al [22] than in this study (70-105 years). Despite walking slower, women maintained a higher cadence than men, a trend also observed in earlier research [42,43].

Interestingly, age-related changes in daily-life walking behavior were not linear. Daily steps, walking speed, and maximum WB distance declined gradually at first, followed by a more pronounced reduction with advancing age. Although women showed signs of decline earlier, men experienced a steeper drop in later years, eventually resulting in a higher rate of change at the highest ages. These findings go beyond patterns found in earlier studies that used linear models, such as Kiselev et al [22]. Our findings align with those of Ko et al [44], who observed increasingly pronounced declines in gait speed with age in laboratory settings, and of Obuchi et al [43], who reported similar sex-specific trends in daily-life walking in Japanese adults. Notably, habitual cadence remained stable up to a high age, with both sexes largely maintaining their habitual cadence until 85 to 90 years of age, after which the decrease accelerated drastically. Women simultaneously showed a decline in walking speed, suggesting a reduction in their stride length.

Physical function, assessed using the SPPB, emerged as a stronger influence than age itself on daily-life walking behavior. Except for habitual cadence, which remained relatively stable across most levels of physical function, declines in the remaining gait metrics were more pronounced with lower SPPB scores, indicating that functional ability plays a central role in maintaining mobility. This supports previous claims that gait changes cannot be fully explained by age, sex, height, or BMI alone [22] and highlights the importance of including an assessment of physical function. In contrast to our findings, van Gameren et al [13] found weaker associations between SPPB and walking metrics, potentially due to their exclusion of short WBs, which comprised over 60% of our dataset.

Although the SPPB is frequently used as a measure of physical function in older adults [10,16,30], it has known ceiling effects among those with high-level functioning [45,46], which may conceal early, more subtle declines in physical function. Hence, future research should strive to incorporate more sensitive assessment tools capable of distinguishing performance within this well-functioning subgroup.

Methodological Considerations

This study has 2 key methodological strengths: a robust sample and validated algorithms to estimate daily-life gait metrics. The HUNT4 Trondheim 70+ cohort comprises a large and diverse group of older adults, spanning a broad range in age and function, with near-equal sex distribution. Furthermore, the vast majority of participants contributed 6 or 7 days of accelerometry data, likely due to the use of body-fixed sensors secured with adhesive tape, minimizing participant interference and ensuring consistent data collection. The digital mobility outcomes in our study were derived using algorithms previously validated in older adults and various clinical populations [26,27], including frail individuals recovering from hip fractures [47]. Although these algorithms were developed to be device agnostic, we used a different yet similar device (the AX3 instead of the AX6) as well as a different sampling frequency, requiring resampling of our data. In the original validation work, algorithm performance improved further with the exclusion of WBs shorter than 10 seconds and among participants with better physical function [27,47]. However, as these short bouts represented the majority of walking behavior in our cohort, they were retained in the current analyses. The exclusion of breaks within WBs strengthened the estimation of speed and cadence but may have led to an underestimation of walking duration and distance when short breaks segmented a walking period into multiple bouts. However, the inclusion of short WBs and exclusion of breaks offers a close description of walking behavior in older adults as it occurs in their daily lives, which is frequently interrupted by environmental factors and activity transitions. Future research should explore whether and how the frequency and duration of breaks influence the estimations of walking duration and distance in naturalistic settings. Furthermore, future research should investigate how older adults walk in relation to other potentially influencing factors such as cognitive function, comorbidities, environmental conditions, and psychosocial factors.

Conclusion

Based on a large study sample with a wide range of age and function, this study provides new insights into how older adults walk in daily life. First of all, there was a persistent large variation in how older adults walk both across and within age groups up to the highest ages. Furthermore, although gait metrics generally decreased with increased age and decreased physical function, changes were largely nonlinear and showed relevant differences between men and women, with men maintaining how they walked in daily life up to a higher age than women but with accelerated change at the highest ages. Surprisingly, habitual cadence remained remarkably stable compared to other gait characteristics in both men and women throughout most ages and levels of physical function. Moreover, although most WBs were very short, the total accumulated walking time in bouts shorter than 30 seconds exceeded that in longer WBs. Finally, daily-life walking performance was affected more by changes in functional ability than by increased age per se. Future research can build upon these findings by implementing daily-life mobility assessment in both clinical studies and patient follow-up, to monitor changes in gait that may point to underlying changes in health condition and function, while

XSL•FO
RenderX

considering the impact of short WBs and relevant group and sex differences in how older adults walk in daily life.

Acknowledgments

The Trøndelag Health Study (HUNT) is a collaboration between HUNT Research Centre (Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology [NTNU]), Trøndelag County Council, Central Norway Regional Health Authority, and the Norwegian Institute of Public Health. NTNU funded the article processing fee for this publication.

Data Availability

Data can be obtained from the HUNT database upon application. Requests to access the datasets should be directed to HUNT Research Centre (HUNT - Trøndelag Health Study - NTNU).

Authors' Contributions

Conceptualization: KBG, AP-I, BV
Formal analysis: KBG, AP-I, ØS
Funding acquisition: BV
Project administration: BV
Software: AP-I
Supervision: AP-I, NS-M, BV
Visualization: KBG
Writing – original draft: KBG, AP-I, BV
Writing – review & editing: KBG, AP-I, GGT, ØS, JLH, NS-M, BV

Conflicts of Interest

None declared.

Multimedia Appendix 1
Results of gamma regression models with log link for each gait metric with age or Short Physical Performance Battery score
[DOCX File, 24 KB - aging_v8i1e75835_app1.docx]

References

1. Viceconti M, Tome M, Dartee W, et al. On the use of wearable sensors as mobility biomarkers in the marketing authorization of new drugs: a regulatory perspective. Front Med (Lausanne) 2022;9:996903. [doi: 10.3389/fmed.2022.996903] [Medline: 36213641]
2. Delgado-Ortiz L, Polhemus A, Keogh A, et al. Listening to the patients' voice: a conceptual framework of the walking experience. Age Ageing 2023 Jan 8;52(1):afac233. [doi: 10.1093/ageing/afac233] [Medline: 36729471]
3. Ageing Europe - statistics on health and disability. Eurostat. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Ageing_Europe_-_statistics_on_health_and_disability [accessed 2024-11-22]
4. Vall-Garcia E, Delgado-Ortiz L, Alcock L, et al. Aging towards walkable futures: insights from a multidisciplinary workshop held in Barcelona, Spain. BMC Public Health 2025 Jan 17;25(1):195. [doi: 10.1186/s12889-024-21223-z] [Medline: 39825288]
5. Wohlrab M, Klenk J, Delgado-Ortiz L, et al. The value of walking: a systematic review on mobility and healthcare costs. Eur Rev Aging Phys Act 2022 Dec 29;19(1):31. [doi: 10.1186/s11556-022-00310-3] [Medline: 36581809]
6. Ungvari Z, Fazekas-Pongor V, Csiszar A, Kunutsor SK. The multifaceted benefits of walking for healthy aging: from Blue Zones to molecular mechanisms. Geroscience 2023 Dec;45(6):3211-3239. [doi: 10.1007/s11357-023-00873-8] [Medline: 37495893]
7. Studenski S, Perera S, Patel K, et al. Gait speed and survival in older adults. JAMA 2011 Jan 5;305(1):50-58. [doi: 10.1001/jama.2010.1923] [Medline: 21205966]
8. Artaud F, Singh-Manoux A, Dugravot A, Tzourio C, Elbaz A. Decline in fast gait speed as a predictor of disability in older adults. J Am Geriatr Soc 2015 Jun;63(6):1129-1136. [doi: 10.1111/jgs.13442] [Medline: 26096387]
9. Cesari M, Kritchevsky SB, Penninx B, et al. Prognostic value of usual gait speed in well-functioning older people--results from the Health, Aging and Body Composition Study. J Am Geriatr Soc 2005 Oct;53(10):1675-1680. [doi: 10.1111/j.1532-5415.2005.53501.x] [Medline: 16181165]
10. Dapp U, Vinyard D, Golgert S, Krumpoch S, Freiberger E. Reference values of gait characteristics in community-dwelling older persons with different physical functional levels. BMC Geriatr 2022 Aug 29;22(1):713. [doi: 10.1186/s12877-022-03373-0] [Medline: 36038832]

XSL•FO
RenderX

11. Elshehabi M, Del Din S, Hobert MA, et al. Walking parameters of older adults from a lower back inertial measurement unit, a 6-year longitudinal observational study. Front Aging Neurosci 2022;14:789220. [doi: 10.3389/fnagi.2022.789220] [Medline: 36172482]
12. Brown JC, Harhay MO, Harhay MN. Walking cadence and mortality among community-dwelling older adults. J Gen Intern Med 2014 Sep;29(9):1263-1269. [doi: 10.1007/s11606-014-2926-6] [Medline: 24934147]
13. van Gameren M, Voorn PB, Bossen D, et al. The Short Physical Performance Battery does not correlate with daily life gait quality and quantity in community-dwelling older adults with an increased fall risk. Gait Posture 2024 Oct;114:78-83. [doi: 10.1016/j.gaitpost.2024.09.005] [Medline: 39276702]
14. Van Ancum JM, van Schooten KS, Jonkman NH, et al. Gait speed assessed by a 4-m walk test is not representative of daily-life gait speed in community-dwelling adults. Maturitas 2019 Mar;121:28-34. [doi: 10.1016/j.maturitas.2018.12.008] [Medline: 30704562]
15. Reijnierse EM, Geelen SJG, van der Schaaf M, et al. Towards a core-set of mobility measures in ageing research: the need to define mobility and its constructs. BMC Geriatr 2023 Apr 6;23(1):220. [doi: 10.1186/s12877-023-03859-5] [Medline: 37024827]
16. Urbanek JK, Zipunnikov V, Harris T, Crainiceanu C, Harezlak J, Glynn NW. Validation of gait characteristics extracted from raw accelerometry during walking against measures of physical function, mobility, fatigability, and fitness. J Gerontol A Biol Sci Med Sci 2018 Apr 17;73(5):676-681. [doi: 10.1093/gerona/glx174] [Medline: 28958000]
17. Warmerdam E, Hausdorff JM, Atrsaei A, et al. Long-term unsupervised mobility assessment in movement disorders. Lancet Neurol 2020 May;19(5):462-470. [doi: 10.1016/S1474-4422(19)30397-7] [Medline: 32059811]
18. Takayanagi N, Sudo M, Yamashiro Y, et al. Relationship between daily and in-laboratory gait speed among healthy community-dwelling older adults. Sci Rep 2019 Mar 5;9(1):3496. [doi: 10.1038/s41598-019-39695-0] [Medline: 30837520]
19. Urbanek JK, Roth DL, Karas M, et al. Free-living gait cadence measured by wearable accelerometer: a promising alternative to traditional measures of mobility for assessing fall risk. J Gerontol A Biol Sci Med Sci 2023 May 11;78(5):802-810. [doi: 10.1093/gerona/glac013] [Medline: 35029661]
20. Orendurff MS, Schoen JA, Bernatz GC, Segal AD, Klute GK. How humans walk: bout duration, steps per bout, and rest duration. J Rehabil Res Dev 2008;45(7):1077-1089. [doi: 10.1682/jrrd.2007.11.0197] [Medline: 19165696]
21. Chan LLY, Lord SR, Brodie MA. Daily-life walking speed, quality and quantity derived from a wrist motion sensor: large-scale normative data for middle-aged and older adults. Sensors (Basel) 2024 Aug 10;24(16):5159. [doi: 10.3390/s24165159] [Medline: 39204870]
22. Kiselev J, Nuritdinow T, Spira D, et al. Long-term gait measurements in daily life: results from the Berlin Aging Study II (BASE-II). PLoS One 2019;14(12):e0225026. [doi: 10.1371/journal.pone.0225026] [Medline: 31825966]
23. Yin S, Du L, He Q, et al. Walking more, not faster, is associated with bone health in China of community-dwelling older women: a cross-sectional study. Prev Med 2023 Oct;175:107722. [doi: 10.1016/j.ypmed.2023.107722] [Medline: 37783314]
24. Ando S, Higuchi Y, Kitagawa T, Murakami T, Todo E, Ueda T. Instrumental activities of daily living and number of daily steps in frail older females. J Phys Ther Sci 2019 Oct;31(10):780-784. [doi: 10.1589/jpts.31.780] [Medline: 31645806]
25. Weiss A, Brozgol M, Dorfman M, et al. Does the evaluation of gait quality during daily life provide insight into fall risk? A novel approach using 3-day accelerometer recordings. Neurorehabil Neural Repair 2013 Oct;27(8):742-752. [doi: 10.1177/1545968313491004] [Medline: 23774124]
26. Kirk C, Küderle A, Micó-Amigo ME, et al. Mobilise-D insights to estimate real-world walking speed in multiple conditions with a wearable device. Sci Rep 2024 Jan 19;14(1):1754. [doi: 10.1038/s41598-024-51766-5] [Medline: 38243008]
27. Micó-Amigo ME, Bonci T, Paraschiv-Ionescu A, et al. Assessing real-world gait with digital technology? validation, insights and recommendations from the Mobilise-D consortium. J Neuroeng Rehabil 2023 Jun 14;20(1):78. [doi: 10.1186/s12984-023-01198-5] [Medline: 37316858]
28. Åsvold BO, Langhammer A, Rehn TA, et al. Cohort profile update: the HUNT study, Norway. Int J Epidemiol 2023 Feb 8;52(1):e80-e91. [doi: 10.1093/ije/dyac095] [Medline: 35578897]
29. GjØra L, Strand BH, Bergh S, et al. Current and future prevalence estimates of mild cognitive impairment, dementia, and its subtypes in a population-based sample of people 70 years and older in Norway: the HUNT study. J Alzheimers Dis 2021;79(3):1213-1226. [doi: 10.3233/JAD-201275] [Medline: 33427745]
30. Melsæter KN, Tangen GG, Skjellegrind HK, Vereijken B, Strand BH, Thingstad P. Physical performance in older age by sex and educational level: the HUNT study. BMC Geriatr 2022 Oct 26;22(1):821. [doi: 10.1186/s12877-022-03528-z] [Medline: 36289472]
31. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
32. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994 Mar;49(2):M85-M94. [doi: 10.1093/geronj/49.2.m85] [Medline: 8126356]
33. Jackson DG. ResampleACC.m. GitHub: OpenMovement. 2019. URL: https://github.com/openmovementproject/openmovement/blob/master/Software/Analysis/Matlab/resampleACC.m [accessed 2019-11-27]

XSL•FO
RenderX

34. Pagnamenta S, Grønvik KB, Aminian K, Vereijken B, Paraschiv-Ionescu A. Putting temperature into the equation: development and validation of algorithms to distinguish non-wearing from inactivity and sleep in wearable sensors. Sensors (Basel) 2022 Feb 1;22(3):1117. [doi: 10.3390/s22031117] [Medline: 35161862]
35. Paraschiv-Ionescu A, Newman CJ, Carcreff L, Gerber CN, Armand S, Aminian K. Locomotion and cadence detection using a single trunk-fixed accelerometer: validity for children with cerebral palsy in daily life-like conditions. J Neuroeng Rehabil 2019 Feb 4;16(1):24. [doi: 10.1186/s12984-019-0494-z] [Medline: 30717753]
36. Paraschiv-Ionescu A, Soltani A, Aminian K. Real-world speed estimation using single trunk IMU: methodological challenges for impaired gait patterns. In: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) in Conjunction with the 43rd Annual Conference of the Canadian Medical and Biological Engineering Society 2020:4596-4599. [doi: 10.1109/EMBC44109.2020.9176281]
37. Soltani A, Aminian K, Mazza C, et al. Algorithms for walking speed estimation using a lower-back-worn inertial sensor: a cross-validation on speed ranges. IEEE Trans Neural Syst Rehabil Eng 2021;29:1955-1964. [doi: 10.1109/TNSRE.2021.3111681] [Medline: 34506286]
38. Zijlstra W, Hof AL. Assessment of spatio-temporal gait parameters from trunk accelerations during human walking. Gait Posture 2003 Oct;18(2):1-10. [doi: 10.1016/s0966-6362(02)00190-x] [Medline: 14654202]
39. Guralnik JM, Ferrucci L, Pieper CF, et al. Lower extremity function and subsequent disability: consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery. J Gerontol A Biol Sci Med Sci 2000 Apr;55(4):M221-M231. [doi: 10.1093/gerona/55.4.m221] [Medline: 10811152]
40. Ustad A, Sverdrup K, Tangen GG, et al. Daily physical activity in older adults across levels of care: the HUNT Trondheim 70 + study. Eur Rev Aging Phys Act 2024 Jul 17;21(1):20. [doi: 10.1186/s11556-024-00355-6] [Medline: 39014310]
41. Mobbs L, Fernando V, Fonseka RD, Natarajan P, Maharaj M, Mobbs RJ. Normative database of spatiotemporal gait metrics across age groups: an observational case-control study. Sensors (Basel) 2025 Jan 20;25(2):581. [doi: 10.3390/s25020581] [Medline: 39860951]
42. Auvinet B, Berrut G, Touzard C, et al. Reference data for normal subjects obtained with an accelerometric device. Gait Posture 2002 Oct;16(2):124-134. [doi: 10.1016/s0966-6362(01)00203-x] [Medline: 12297254]
43. Obuchi SP, Kawai H, Murakawa K. Reference value on daily living walking parameters among Japanese adults. Geriatr Gerontol Int 2020 Jul;20(7):664-669. [doi: 10.1111/ggi.13931] [Medline: 32378318]
44. Ko SU, Stenholm S, Metter EJ, Ferrucci L. Age-associated gait patterns and the role of lower extremity strength - results from the Baltimore Longitudinal Study of Aging. Arch Gerontol Geriatr 2012;55(2):474-479. [doi: 10.1016/j.archger.2012.04.004] [Medline: 22564361]
45. Bergland A, Strand BH. Norwegian reference values for the Short Physical Performance Battery (SPPB): the Tromsø Study. BMC Geriatr 2019 Aug 8;19(1):216. [doi: 10.1186/s12877-019-1234-8] [Medline: 31395008]
46. Tangen GG, Robinson HS. Measuring physical performance in highly active older adults: associations with age and gender? Aging Clin Exp Res 2020 Feb;32(2):229-237. [doi: 10.1007/s40520-019-01190-x] [Medline: 30977080]
47. Berge MA, Paraschiv-Ionescu A, Kirk C, et al. Evaluating the accuracy and reliability of real-world digital mobility outcomes in older adults after hip fracture: cross-sectional observational study. JMIR Form Res 2025 May 20;9(1):e67792. [doi: 10.2196/67792] [Medline: 40392580]

Abbreviations

HUNT: Trøndelag Health Study
MoCA: Montreal Cognitive Assessment
SPPB: Short Physical Performance Battery
WB: walking bout

Edited by P Jansons; submitted 11.04.25; peer-reviewed by C Becker, C Sakal, L Taylor, MA Amalbert-Birriel; accepted 02.11.25; published 04.12.25.

Please cite as:
Grønvik KB, Paraschiv-Ionescu A, Tangen GG, Salvesen Ø, Helbostad JL, Skjæret-Maroni N, Vereijken B
Daily-Life Walking Characteristics of Older Adults in Relation to Age, Sex, and Physical Function: the HUNT4 Trondheim 70+ Observational Study
JMIR Aging 2025;8:e75835
URL: https://aging.jmir.org/2025/1/e75835
doi:10.2196/75835

XSL•FO
RenderX

© Karoline Blix Grønvik, Anisoara Paraschiv-Ionescu, Gro Gujord Tangen, Øyvind Salvesen, Jorunn Lægdheim Helbostad, Nina Skjæret-Maroni, Beatrix Vereijken. Originally published in JMIR Aging (https://aging.jmir.org), 4.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Implementing a Digital Physical Activity Intervention for Older Adults: Qualitative Study

Cherish Boxall[1,2*], MRES, MSc; Laura Dennison[2*], PhD; Sascha Miller[2], PhD; Judith Joseph[2], PhD; Kate Morton[2], PhD; Jenny Corser[3,4,5], MSc; Joanna Kesten[2], PhD; Asgar Electicwala[2], PhD; Max J Western[6], PhD; Stephen Lim[2], PhD; Chloe Grimmett[2], PhD; Lucy Yardley[2,3], PhD; Katherine Bradbury[2], PhD

[1]Southampton Clinical Trials Unit, University of Southampton, Southampton, United Kingdom

[2]School of Psychology, University of Southampton, Building 44, Highfield Campus, Southampton, United Kingdom

[3]School of Psychological Science, University of Bristol, Bristol, United Kingdom

[4]The National Institute for Health and Care Research Applied Research Collaboration West (NIHR ARC West), University Hospitals Bristol and Weston NHS Foundation Trust, Bristol, United Kingdom

[5]The National Institute for Health and Care Research (NIHR) Health Protection Research Unit (HPRU) in Behavioural Science and Evaluation, University of Bristol, Bristol, United Kingdom

[6]Centre for Motivation and Health Behaviour Change, Department for Health, University of Bath, Bath, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Katherine Bradbury, PhD
School of Psychology, University of Southampton, Building 44, Highfield Campus, Southampton, United Kingdom

Abstract

Background: Physical activity (PA) in older adults can prevent, treat, or offset symptoms and deterioration from various health conditions and help maintain independence. However, most older adults are insufficiently active. Digital interventions have the potential for high reach at low cost.

Objective: This paper reports on the implementation of "Active Lives," a digital intervention developed specifically for older adults.

Methods: This study had a qualitative design. The implementation team approached a range of National Health Service, public health, community, and third-sector organizations in the United Kingdom to offer Active Lives to as large and diverse groups of older adults as possible. Alongside real-world implementation activities, research was conducted to explore what supports and inhibits the implementation of a digital intervention for PA in older adults. Data collection involved interviews with implementation partners (n=15) and the implementation team (n=3) plus extensive field notes from stakeholder communications. Inductive thematic analysis was used to analyze the data.

Results: Five broad themes were developed, capturing implementation barriers and facilitators. These were (1) complex and opaque networks and influencers, (2) forming an understanding of Active Lives and its fit, (3) a landscape of competition and conflicting interests, (4) navigating unclear approval processes, and (5) shifting strategies: small and effortful to high reach and passive. Identifying key decision makers proved arduous, consuming significant time and resources, and proposals from enthusiastic implementation partners often proved impractical or overly burdensome. Health care professionals demonstrated a comprehensive understanding of the potential benefits of digital interventions in alleviating operational burdens and improving patient care. However, stakeholders from disparate sectors held reservations about digital intervention and had different views on the best approaches to supporting PA among older adults. This discord was exacerbated by conflicts with existing local initiatives, such as group exercise programs, which occasionally hindered the implementation of Active Lives. Furthermore, bureaucratic hurdles within National Health Service trust approval processes acted as formidable obstacles, dampening progress and resolve, highlighting the need for guidance in identifying sustainable and scalable practices.

Conclusions: The findings highlight important implementation challenges to digital PA interventions for older adults such as bureaucratic barriers and alignment with ongoing initiatives. This research emphasizes the necessity for strategic direction and multilevel guidance to efficiently implement digital interventions for PA among community-dwelling older adults.

(JMIR Aging 2025;8:e64953) doi:10.2196/64953

KEYWORDS

implementation; digital technology; physical activity; physical exercise; exercising; older adult; older people; aging; mobile app; smartphone; digital health; digital intervention

XSL•FO
RenderX

Introduction

Background

Physical activity (PA) in older adults aged 65 years and older can help prevent, treat, or offset symptoms and deterioration from a range of common health conditions such as frailty, falls, depression, cardiovascular disease, Parkinson disease, and cognitive decline [1-3]. However, in the United Kingdom, 79% of adults aged 65 years and older are insufficiently active (ie, do not undertake 150 minutes of moderate-intensity PA per week), and 32% are completely inactive (ie, undertake no moderate-intensity activity per week), with similar patterns noted in US and Dutch samples [4]. Inactivity is detrimental to older adults' quality of life and ability to live independently, which is costly to health and social care services [3,5].

Digital behavior change interventions hold the potential to reach large portions of the older adult population. There is concern that older adults might be disadvantaged by the digital divide, where a lack of familiarity with technologies throughout their life course makes them less likely to accept, engage with, and benefit from digital solutions. In particular, older adults who are insufficiently active are more likely to experience socioeconomic disadvantage and have previously benefited less from digital PA interventions [6]. Encouragingly, the percentage of older adults using the internet has been rising, increasing the viability of digital solutions in this age group. In the United Kingdom, 83% of 65‐ to 74-year-old people were internet users, and 47% of 75-year-old people were internet users in 2019. This rose to 53% of 75-year-old people in 2020 [7,8].

Prior Work

The "Active Lives" digital intervention is a freely available website suited to computers, tablets, and smartphones. It was developed by our research group as an inclusive, equitable, and scalable solution to help older adults increase and maintain PA [9,10]. Active Lives supports users to reduce sedentary time, increase moderate activity, and undertake strength and balance exercises. It was developed using behavior change theory in conjunction with qualitative research to ensure that the intervention was acceptable and engaging to its intended end users [10-12]. The website uses simple language and navigation to improve accessibility for people with mild cognitive impairment and lower digital or health literacy. To adapt to the heterogeneous PA abilities observed within older adults, Active Lives starts by assessing the user's strengths and weaknesses and uses this information to tailor its recommendations, suiting it to a supportive patient-led intervention. The website addresses psychological barriers to beginning or increasing activity (eg, low motivation, low self-efficacy, concerns about injury, appropriateness, and safety), using evidence-based behavior change strategies such as self-monitoring and goal-setting [13]. These features set it apart from other available PA websites and apps [14-16]. Active Lives is currently being evaluated as a part of a 5-year multicomponent intervention (PA, diet, and cognitive exercises) to prevent cognitive impairment in older adults. To use Active Lives, users click on a website link to register with an email address and password and subsequently use the link to log in at any time. The study is registered on the United Kingdom's clinical study registry (ISRCTN17349359), which aims to promote transparency and reduce selective reporting in interventional and noninterventional clinical studies. Preliminary findings from 2 feasibility randomized controlled trials (RCTs) conducted within primary care indicated that people randomized to Active Lives engaged well with the Active Lives intervention, increased their walking, moderate activity, vigorous activity, and strength or balance exercises compared to baseline, with no adverse events reported [9]. The intervention was co-designed with 52 older adults with and without cognitive decline to ensure acceptability. Qualitative interviews showed that the intervention was acceptable, with key facilitators of engagement identified as maintaining independence and enjoyment, while managing ill health was found to be a barrier to engagement [12].

Digital interventions such as Active Lives can potentially have huge reach at low cost, so scaling up effective interventions is a priority, but there is a research-to-practice gap, where behavioral interventions are not implemented beyond their formal evaluation in RCTs [17], and there is limited implementation-focused research to guide how this gap could be shortened [18,19]. Specifically, for navigating the unique challenges associated with implementing digital behavior change interventions [17,20,21].

Implementation research spans various stages of intervention development, including planning, feasibility testing, evaluation, and the implementation phase itself [13]. Numerous studies have been conducted during digital intervention development and integrated within feasibility testing and evaluation, providing insights into factors that may influence future scale-up [22-26]. However, there is a paucity of published research investigating implementation following successful development and evaluation.

Existing studies in this domain have primarily focused on aspects of the digital intervention itself or end-user uptake, use, and engagement [27,28]. Some research has explored implementation barriers and facilitators within the context of large-scale digital service transformation programs involving extensive consortiums [20,29].

Nevertheless, there remains a gap in the literature regarding the implementation research on scaling up a single, fully developed intervention, particularly concerning the process of engaging partner organizations to reach patients or end users. Recent commentaries have highlighted various challenges encountered when attempting to make evidence-based digital interventions freely and widely available after the RCT, such as in eczema management [30] and asthma management [21].

The Goal of This Study

Digital interventions may be an effective solution for older adults to increase PA, but little is understood about the process of engaging with individuals and organizations that have the potential to enable reach to patients and end users. In this study, we explore the influencing factors to the implementation of a digital PA intervention aimed at older adults. We were interested in what supports or inhibits implementation and how challenges were perceived, approached, and managed from multiple perspectives, including the academic team and individuals

XSL•FO
RenderX

involved in the implementation process, an approach recommended by Murray et al [31].

Methods

Implementation Strategy and Context

The implementation activities took place in the United Kingdom between September 2020 and January 2023. In this paper, the terms “implementation team” and “implementation partners” are used to describe 2 broad sets of actors.

The implementation team was academics trying to get Active Lives to end users (via implementation partners). The implementation team was led by coauthor KB, a health psychologist involved in developing and evaluating Active Lives and supported by 2 postdoctoral researchers (co-authors JJ and SM).

Implementation partners were people, including clinical and academic grant coapplicants, who were approached about facilitating the implementation of Active Lives either at a strategic or practical level. The term is used in this paper to capture all such individuals, regardless of whether they eventually implemented Active Lives. Due to the location of the implementation team and their professional network, most of the implementation partners were based in the south of the United Kingdom. We began implementing Active Lives in the most socioeconomically deprived areas within Southampton and Portsmouth. These regions were selected due to their high prevalence of low education levels and established connections with local charities that serve vulnerable populations. Following these initial areas, Active Lives experienced organic expansion in a diverse range of health care and community organizations based in Devon, Berkshire, Sussex, Reading, Isle of Wight, and Dorset.

Active Lives is designed to be used without specialist support or follow-up. Therefore, the key action necessary to roll out Active Lives was that implementation partners shared a link (ie, a URL) with older adults and a brief verbal or written explanation of Active Lives, ideally framed as an endorsement or recommendation to encourage self-management of PA. The implementation team was open to exploring implementation opportunities that emerged naturally during the project. Several routes were anticipated as starting points through insights from pilot work and planning discussions at the grant proposal stage, including NHS falls and physiotherapy services, public health, social prescribers, charities, and community organizations.

The implementation team planned a theoretically informed implementation strategy. A thorough account of this is beyond the aims and scope of this paper. Table 1 outlines how 3 key theories and frameworks [32-34] were referred to for implementation planning and data collection activities. As implementation progressed, a Plan Do Study Act approach [35] was used to review and adapt the implementation strategy and activities.

XSL•FO
RenderX

Table . Theories and models used to plan the Active Lives implementation strategy.

Theory or model or framework	Overview of how the theory or model or framework was used	Examples
RE-AIM[a] [32]	RE-AIM was used as a broad overarching framework to consider important aspects of intervention success. It helped us to consider key implementation activities and to plan data collection (logging implementation activities and digital intervention use data).	• RE-AIM guided decisions on prioritizing routes that were likely to have high Reach. For example, more of the implementation team's limited time was invested in attempts to get implementation partners to set up routine digital or paper-based communication from NHS[b] services that serve thousands of people rather than ad-hoc face-to-face referrals via volunteers in small local community organizations that serve hundreds of people. • RE-AIM also helped to identify types of data that needed to be collected for evaluation purposes (where feasible). For example, the decision was made to create bespoke Active Lives URLs for different organizations and routes, so that users signing up via different routes could be tracked separately (Reach).
Normalization Process Theory (NPT) [33]	NPT was used to anticipate the work that implementation partners would need to undertake in order to make Active Lives become sustainable, normal practice within their organization. It led us to consider how implementation plans and activities could facilitate these processes. It also influenced data collection (topic guides).	• When considering the coherence concept, the implementation team developed a set of resources (videos, PDFs, and presentations or pitches) to facilitate implementation partners "sense-making" around Active Lives, particularly drawing out its unique features and making clear how it would be used. This was iteratively tailored to different settings based on the implementation team's evolving understandings of the context and priorities of each implementation partner.
The Consolidated Framework for Implementation Research (CFIR) [34]	The CFIR was used to identify, pre-empt, and plan for how to respond to a wide range of barriers and facilitators to implementation. It also influenced data collection (topic guides).	• The outer setting domain: External pressure construct focused the implementation team on the need to be aware of contemporary pressures on implementation partners and how this would link to their interest in and ability to implement Active Lives. For example, within the NHS, relevant factors included post–COVID-19 hospital waiting lists (and the "Waiting Well" agenda), deconditioning in older adults, various new strategies and policies, and restructuring of teams and services. Some of these "external pressures" aligned well with Active Lives, while others indicated competing pressures for potential implementation. • The inner setting domain: Internal pressure construct highlighted the need to understand (in advance) the organizations and teams that were potential implementation partners. Constructs such as "culture," "tension for change," and "relative priority" helped to plan and tailor communications and sometimes aided decision-making about spending time pursuing routes that were not working.

[a]RE-AIM: Reach, Effectiveness, Adoption, Implementation, Maintenance.

[b]NHS: National Health Service.

XSL•FO
RenderX

Participants

Health care professionals (HCPs; n=approximately 200), health service managers, public health specialists, sports and PA experts, and community and charity workers were involved in emails, meetings, or discussions about implementing Active Lives. We sent an invitation email to those who represented or had influential power in a group who were approached regarding the implementation of Active Lives, irrespective of their progress. The academic members of the implementation team conducting much of the work with implementation partners were also interviewed.

Data Collection

Overview

After completion of consent forms, guided semistructured interviews (SuppMat3 and 4) were completed using videoconferencing software (Microsoft Teams) and were audio-recorded and transcribed verbatim. The personnel who conducted the interviews were not directly involved in implementation activities and consisted of a Predoctoral Fellow who completed most implementation partner interviews and 2 experienced qualitative researchers, one of whom conducted the interviews with the academic team.

Implementation Team

Semistructured interviews captured the experiences and perceptions of the academic team involved in the implementation activities, including perceived barriers and facilitators (Multimedia Appendix 1).

Implementation Partners

Our sample size was determined pragmatically by the willingness of our partners to participate. Assessment of information power revealed a high degree, supported by the narrow aim, dense specificity, application of theory, and strong dialogue [36]. The topic guides (Multimedia Appendix 2) focused on their involvement with and perspective on Active Lives, including obstacles, facilitators, and learning points relating to implementation.

Field Notes

The implementation team made field notes. Typically, these were reflections on meetings or other communications with implementation partners. These included perceptions of power dynamics and behavior within meetings, issues that had been mentioned informally by implementation partners, personal feelings about the work involved with implementation, as well as problem-solving and suggestions for next steps.

Data Analysis

Analysis was led by LD and was broadly aligned with a reflexive thematic analysis approach [37]. It involved inductively developing themes that capture shared meaning and creating interpretations and narratives around observations and patterns recognized within the dataset. Analysis began with familiarization (listening to audio recordings and reading transcripts and field notes). NVivo (QSR International; version 12) was used to code the data related to the research question. Codes were organized and clustered into preliminary themes, which captured important features and patterns within the data. Themes were iteratively reworked and relabeled, with ongoing discussion from coauthors, to add additional analytic perspectives and ensure that the final themes were insightful, coherent, and well-supported by the data.

Reflexivity

Consistent with reflexive thematic analysis, we consider our analysis and the themes we generated to be inextricably linked to our own experiences, assumptions, and positionalities [37]. An important issue to draw attention to is the multiple and overlapping roles of coauthors KB, JJ, and SM who were involved in the development of Active Lives and formed the implementation team (and as such were also interview participants and producers of field notes). Although LD did not participate in implementation meetings and activities, her academic background, experiences, loyalties, and goals are naturally more aligned with the implementation team than the implementation partners.

Research Design and Reporting

The study is reported using the StaRI (Standard for Reporting Implementation Studies) statement for implementation studies [38] and the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist for qualitative studies [39] (Checklists 1 and 2).

Ethical Considerations

Ethics approval for this research was obtained via Health Research Authority and Health and Care Research Wales (ref 22/HRA/0130) and from the University of Southampton (ref 57536.A2). All participants were given an information sheet outlining the study's objectives and provided informed written consent. Participants were informed prior to and at the time of interview that their involvement was voluntary. To ensure anonymity and confidentiality, participant data were replaced with a numerical code (eg, 1001001), and identifying information within the transcript was anonymized prior to data analysis.

Results

In total, 17 interviews were carried out.

Implementation Team

Three interviews were with members of the implementation team. Interview durations varied between 76 and 97 minutes.

Implementation Partners

A total of 22 implementation partners were invited for interviews from 12 different organizations. In total, 8 people declined the invitation to interview, and 1 did not reply. In total, 14 interviews were completed with 13 implementation partners. This included NHS HCPs, most of whom were in senior or strategic leadership roles, partners from an Academic Health Sciences Network (AHSN), employees from sports or PA organizations, and from charities or community groups serving older adults.

XSL•FO
RenderX

Due to a change in roles and organization during the study, 1 interview was repeated because the participant was able to share implementation perspectives from both a strategic (regional) and organizational (health care department) level.

Interviews with implementation partners lasted between 27 and 57 minutes (median 33, IQR 30.5-47 minutes). Detailed demographic information is withheld from reporting in the paper to ensure the anonymity and confidentiality of implementation partners' participants.

Overview

Five broad themes were developed from the data, as depicted in Table 2. Below, each theme is explained, and supporting quotations are presented. To preserve anonymity, names and organizations are replaced with participant ID numbers and tagged as either implementation team (eg, ImpTeam_1) or implementation partner (eg, ImpPartner_2). For the latter, a broad indication of their professional role or sector is given (eg, NHS HCP).

Table . Themes and subthemes.

Theme	Subthemes
Complex and opaque networks and influencers: This theme is about the complexities of identifying people with sufficient interest and power to help implement Active Lives.	• Identifying open doors and gatekeepers: This subtheme is about the difficulties in identifying suitable implementation partners. • Champions and "antichampions": This subtheme is about the people whose actions facilitated and blocked implementation.
Forming an understanding of Active Lives and its fit: This theme is about how implementation partners understood (or misunderstood), labeled, and appraised Active Lives and grasped how it might fit into their organizations and alongside their priorities. It also covers how the implementation team facilitated sense-making about Active Lives.	• Labeling Active Lives as a good thing: This subtheme is about the positive aspects of Active Lives that implementation partners recognized. • Having doubts and concerns about Active Lives: This subtheme relates to objections to and worries about Active Lives that Implementation Partners had.
A landscape of competition and conflicting interests: This theme covers how Active Lives implementation occurred in a context of competition with other interventions and initiatives and how this inhibited or blocked implementation	• —[b]
Navigating unclear approval processes: This theme is about how a lack of clear process for large organizations (particularly NHS[a]) to approve Active Lives or sign off on decisions created burden and uncertainty, slowing or blocking implementation.	• —
Shifting strategies: small and effortful to high reach and passive: This theme is about how implementation partners were initially attracted to small-scale, effortful routes but could be persuaded to find more efficient, scalable routes.	• —

[a]NHS: National Health Service

[b]Not applicable

Complex and Opaque Networks and Influencers

Identifying Open Doors and Gatekeepers

The implementation team's professional networks made connections and introductions to potential implementation partners and helped them to understand "the landscape [...], the groups that are influential" (ImpTeam1). Implementation partners in both NHS and community settings considered supporting older adults to become more active to be relevant to their job, but it was rarely a key responsibility.

> *[It's] a "side thing" for most people.* [ImpTeam3]

Language reflecting the need for openness was used frequently; pursuing "open doors" (ImpPartner1_AHSN) rather than "pushing against closed doors" (ImpPartner2_AHSN) or "trying to knock [the door] down" (ImpPartner9_sports/PA sector). However, determining the openness of "doors" was unpredictable despite steers from a wide network. Many initial meetings were unproductive. Others seemed positive; yet, further contact or action never materialized. "Gatekeepers" (ImpTeam1) were gradually uncovered; people were enthusiastic but also, crucially, held decision-making power.

> *if they're not the right decision-maker then it's quite-it's a dead end.* [ImpTeam1]

Difficulties (highlighted below) in recognizing open doors and gatekeepers at the project outset meant "traction" (ImpTeam1) was slow, and progress only began "snowballing" (ImpTeam2), as project funding was running out.

Champions and Antichampions

At multiple levels (eg, regional to local department), champions were seen as crucial for spreading Active Lives and were recognized as "adding massive credibility" (ImpTeam1) and garnering and sustaining interest, commitment, and action.

> *[I'm] kind of banging away at the drum.* [ImpPartner8_NHS-HCP-leadership]
>
> *I have encouraged others to use it in their clinical work. I have just made sure that there weren't any*

XSL•FO
RenderX

barriers locally. [ImpPartner10_NHS-HCP-leadership]

People who championed Active Lives or facilitated introductions noted professional benefits, including prestige, improved networks, and better ability to perform their jobs. However, the work needed to support implementation could also be burdensome (see Forming an Understanding of Active Lives and Its Fit section).

Besides champions, participants discussed people who deliberately blocked or inhibited implementation: "Anti-Champions" (ImpTeam1). Sometimes blocking was overt, directly speaking out against Active Lives or behaving in rude and hostile ways. Participants also suspected that antichampions who were hostile to Active Lives were obstructing progress from behind the scenes. Considering this, ImpPartner10_NHS HCP-leadership advised of the need to "go around and above" those with a reputation for "territory defending" (Fieldnote-ImpTeam1). Although participants expressed some sense of mystery about the dynamics at play, antichampionship tended to be interpreted as arising from misunderstanding Active Lives or having competing interests (see Forming an Understanding of Active Lives and Its Fit section).

Forming an Understanding of Active Lives and Its Fit

Labeling Active Lives as a Good Thing

All implementation partners recognized PA in older people as a magic medicine (ImpPartner1_AHSN). They appreciated the need to better support PA in community and clinical settings.

An absolute no-brainer in having a benefit. [ImpPartner15_NHS HCP]

Most viewed Active Lives enthusiastically, valuing its coproduced development with older adults and its tailored activity recommendations designed to accommodate the population's diverse abilities, including those with chronic health conditions and lower levels of digital literacy. Coverage of psychological barriers to PA, support with different types of PA (including strength or balance), and emphasis on activity within daily life were appreciated. The absence of cost was important. The trial evidence base was valued by HCPs, whereas other partners found case studies and testimonials more persuasive.

Active Lives was perceived as having potential for extensive reach and broad relevance. One HCP declared that the patient group she would like to use it with was "basically the world!" (Fieldnote-ImpTeam2), another felt that "it could fit for anybody" (ImpPartner14_NHS HCP). Nevertheless, some participants described how PA interventions struggle to fit within the NHS, which is organized around conditions, not prevention.

There's just no agenda for prevention within the NHS or proactive interventions. [ImpPartner10_NHS HCP-leadership]

[the broad appeal] is partly its downfall [...] if you can be quite specific about where it fits and what pathway and what condition it fits with, you can almost be quite prescriptive in its implementation. [ImpPartner7_NHS HCP]

The implementation team, mindful of this, developed skills in understanding services, recognizing unmet needs, and helping HCPs to identify exactly how Active Lives fit: "the value to the organisation, the individual" (ImpPartner2_AHSN). They gradually honed the way they articulated the offer, finding the right "hook" (ImpPartner1_AHSN). When approached in this way, HCPs were particularly positive, easily grasping how Active Lives would benefit their patients and their working practices.

I immediately knew that it would fit in to my current work. [ImpPartner10_NHS HCP-leadership]

Somewhere easy to signpost people to. [ImpPartnerP12_NHS HCP]

Active Lives' value was readily recognized when services were undergoing new challenges or change. For example, when new regional or department-level plans were introduced. Its value was readily seen in the context of excessive waiting lists, deconditioning of older people, and hospital pressures resulting from long hospital stays.

[It] ticked all the boxes etc ... particularly during and since COVID. [ImpPartner2_AHSN]

Having Doubts and Concerns About Active Lives

Most implementation partners, particularly within the NHS, considered Active Lives' digital format suitable for many older adults. However, the suitability was occasionally rejected or questioned, particularly by those in community organizations serving deprived areas. Digital exclusion seemed to drive concerns. ImpPartner11_Charity/community sector described "grinding poverty" and "language barriers" in her deprived urban area. There were "just–too many barriers ... it's not to say that it's not a fantastic thing." Another described how

Around 60% of [the older adults we serve] are either not digitally included at all or their activity would be very restricted. [ImpPartner4_Charity/community sector]

Several implementation partners were highly concerned that digital interventions could exacerbate isolation and loneliness.

I would never ever support anything that would completely remove human contact. [ImpPartner11_Charity/community sector]

There was also a broader concern about digital services threatening the existence of face-to-face services. One implementation partner described older adults being reluctant to engage with (non-PA) digital services "because they feared that if they started engaging with us digitally, we'd decrease their access to face-to-face stuff" (ImpPartner4_charity/community sector).

Implementation partners consistently communicated that Active Lives was only suitable as part of a broader PA offer and repeated the importance of ensuring choice for patients or users.

As part of a suite of interventions. [ImpPartner10_NHS HCP-leadership]

XSL•FO
RenderX

> *A menu of solutions [...] a pick and mix.* [ImpPartner1_AHSN]

> *Beneficial to add to a selection of tools.* [ImpPartner6_NHS HCP]

The need to offer older adults a variety of PA approaches was acknowledged by the implementation team. Nonetheless, some implementation partners believed that Active Lives sought to replace rather than complement other services (see A Landscape of Competition and Conflicting Interests section).

Occasionally, concern was expressed over the safety of digital interventions for PA in older adults, including about dizziness or falling if PA was unsupervised and potential litigation if someone was injured.

> *[The public health team thought] services that are not provided by a certified exercise specialist might be risky.* [ImpTeam1]

Another unfavorable perception of Active Lives was linked to a distrust of academics and a view that researchers are detached from practical realities, thereby rendering evidence-based interventions inappropriate to meet the local population's needs. Importantly, this disconnect with academia seemed limited to community organizations who prioritized local insights to develop and commission PA services that got people out of the home and socializing, a singular approach that was privately challenged by the implementation team.

> *They trust each other and they don't necessarily trust academics.* [ImpTeam1]

One implementation partner described disinclination to implement Active Lives, which they considered "an academic research piece" rather than a "practical intervention that is really looking to make sustainable change in local communities" (ImpPartner3_Sports/PA sector).

A Landscape of Competition and Conflicting Interests

Most data underpinning this theme came from the implementation team interviews and their field notes about meetings and conversations. A key barrier to initial buy-in and sustained action on implementing Active Lives was a dedication to existing PA initiatives adopted by services and organizations. Specifically, Active Lives was perceived as posing a threat to and being in competition with other ventures.

A common pattern of concerns raised by implementation partners from the PA sector related to how they believed digital provision threatened or competed with doing PA outside of the home. A specific concern was that Active Lives would stop people from attending group exercise programs that they had commissioned. This appeared to be based on a misunderstanding that Active Lives motivated the use of exercise videos to follow alone at home rather than being a motivational and supportive behavioral intervention to encourage older people to adopt whatever form of PA they preferred (including group activities outside the home). The implementation team tried to address this, articulating how Active Lives could complement group exercise provision, serving different types of people, or potentially providing a "stepping stone" (ImpTeam1) to support currently inactive people to increase confidence and motivation to attend classes.

> *[we learned to] position Active Lives so that it didn't seem a threat ... So they could see that this was complementary, rather than alternative.* [ImpTeam1]

Perceived conflict and competition were also apparent with condition-focused self-management digital interventions with a PA component (eg, diabetes and chronic obstructive pulmonary disease), informational websites about PA, and national PA promotion campaigns.

> *[physical activity sector partners] felt a little threatened by it, because they felt that that was the area that they were trying to involve themselves in as well. I think there was a bit of a perceived conflict there.* [ImpPartner5_NHS HCP]

> *They're already paying for [disease self-management DI]; they don't want to confuse it by getting something else even though he thinks [Active Lives] is better.* [ImpTeam2]

Some of the commitment to other approaches and unwillingness to consider Active Lives seemed to stem from personal involvement with development or commissioning (eg, where a service had created its digital resources).

> *He's already got his own thing that he's doing and he doesn't want to promote [Active Lives].* [ImpTeam2]

The implementation team described experiencing territorial and defensive behavior, "[an Implementation Partner said] I was stepping on toes" (ImpTeam1). A sense of territory was reflected in the language of implementation partners, such as "hunting grounds" (ImpPart1_AHSN).

Other reasons for loyalty to existing interventions also became clear. With one charity, it emerged that a group PA program was generating income, and a successful adoption of Active Lives was perceived as threatening this income.

> *[that the income stream] was more of a motivator for them than us saying "here's something that can meet the need."* [ImpPartner1_AHSN]

Furthermore, informal conversations with some implementation partners highlighted fear of redundancy; digital interventions threatened the viability of in-person services and job security. The perception of an overcrowded and competitive space was frustrating to the implementation team who could differentiate Active Lives' unique scope and features and relevance to different populations.

> *It's something completely different!* [ImpTeam1]

Conflicting interests were often undeclared in formal meetings and discussed off-record. Sometimes it only became apparent to implementation partners when they consulted more widely.

The concerns expressed within this theme meant that several charity sector PA partners chose not to take on Active Lives. While there were clear financial concerns from these partners, relating to the perception that Active Lives could reduce attendance at their in-person exercise classes, we also examined

XSL•FO
RenderX

the funding sources of these services to understand whether funding might be related to interest in rolling out Active Lives. Charities that were approached to implement Active Lives were usually funded by a mix of funding streams, which included donations, county council support, and funding from the national lottery and Sport England. Those most reluctant to participate were more likely to be part-funded by Sport England and provided their own in-person PA classes. However, once we had understood the concerns of these organizations, we were able to refine our pitch to more clearly demonstrate how Active Lives could help motivate more people to attend their in-person classes rather than competing as an alternative to in-person support. After this, when we approached other similar organizations who were part-funded by Sport England (in some cases, part of the same organization in a different area), those services often did decide to roll out Active Lives.

Shifting Strategies: Small and Effortful to High Reach and Passive

The people who most recognized the value of Active Lives tended to be busy NHS HCPs who had limited "headspace" (ImpPartner5_NHS HCP) and expressed frustration with their ever-expanding workload and responsibilities. Typically, implementation partners initially suggested ways they wanted to spread Active Lives within their service organizations. They often proposed routes that were likely to be low reach, extremely effortful, or both.

> *People can often come up with ideas that aren't going to work.* [ImpTeam1]

Low-reach suggestions included waiting room posters or mentioning Active Lives on a long list of unrelated resources or a low-traffic website, strategies the implementation team was not eager to pursue.

> *If we're in a list, we just get buried.* [ImpTeam3]

Enthusiastic implementation partners usually felt that in-person signposting to patients or service users was the ideal way to provide access to Active Lives.

> *Signposting from kind of healthcare professionals is the kind of optimum route.* [ImpPartner2_AHSN]

HCPs took the suitability of Active Lives for their patients very seriously and wanted to ensure that Active Lives was discussed with the right patient at the right time to ensure safety and appropriateness. HCPs described needing to be both familiar with and in agreement with the content of Active Lives before discussing with a patient.

> *We've basically both just signed up to it and had a look through ourselves.* [ImpPartner13_NHS HCP]

> *If it goes wrong, and someone doesn't like the outcome of it, or doesn't enjoy it, or thinks that it's going to cause them an issue, it's us that the blame will fall back on.* [ImpPartner6_NHS HCP]

Some implementation partners anticipated it being necessary to verbally "onboard" users and even demonstrate how to use Active Lives.

> *Click on this, this is what you'll see. This is how you'll interact with it, this is what you'll do.* [ImpPartner7_NHS HCP]

However, implementation partners also highlighted how unfeasible this would be in the context of being overloaded and unable to cope with the "creep" (ImpPartner5_NHS HCP) of requirements for each clinical consultation. Having multiple HCPs within services absorb and retain information about Active Lives and finding time to discuss it with patients also proved unrealistic to organize and sustain.

> *To try and remind [colleagues], encourage them, to offer it as an option [...]. It can become quite overwhelming.* [ImpPartner6_NHS HCP]

The implementation team was clear that this was not a sustainable way to fulfill "the promise of digital—that you can reach lots and lots of people" (ImpTeam1).

Recognizing the tendency for implementation partners to identify overly effortful or low-reach routes, the implementation team began to offer keen implementation partners a consultation meeting and follow-up support to pursue routes "that are more passive, where people are not relied on" (ImpTeam1). Successful strategies included embedding information about Active Lives within routine letters or digital communications to patients (eg, SMS text messages) or otherwise signposting Active Lives into an existing process or protocol. Once suggested, these routes were enthusiastically supported, especially when examples of other organizations that were already doing this were given.

Navigating Unclear, Burdensome Approval Processes

Implementation partners from various settings expressed caution around implementing digital interventions, uncertainty about how to assess them, and a lack of time to learn. Some highlighted how they and their colleagues were not confident with digital technologies: "we're all a bit out of date" (ImpPartner15_NHS HCP).

In this context, NHS HCPs faced extra work if they tried to gain organizational approval to implement Active Lives.

> *You think, "this might be hard work." It almost puts a break on before you go any further.* [ImpPartner15_NHS HCP]

The work usually involved multiple meetings as discussions moved "up the food chain" (ImpTeam1). The process was extremely slow and frustrating to navigate for both implementation partners and the implementation team.

> *Feels like wading through mud.* [ImpPartner8_NHS HCP-leadership]

HCPs found themselves burdened with complicated approval paperwork and were expected to attend meetings about information governance. One HCP (too busy to be interviewed) described it as "overwhelming and intimidating" and "heavy handed" (Field notes-ImpTeam2). Uncertainty characterized the approval process, as different NHS trusts and services were inconsistent, and the process was not publicized. The implementation team felt unable to help implementation partners when discussions were ongoing within trusts because they could

XSL•FO
RenderX

neither see what was happening nor anticipate what objections might arise.

> *[it is like] a corridor with ten doors on it and you don't even know what those doors are, who they lead to.* [ImpTeam2]

During the implementation project, implementation partners encountered crises including the COVID-19 pandemic, industrial action, and NHS trust "critical incidents." An easy, successful implementation route came about during the chaos of the COVID-19 pandemic when an HCP added the Active Lives URL and a sentence explaining its benefits to a routine patient letter, without the typical wider consultation.

> *We perhaps could have gone further up and asked for permission at a higher level ... they might have said no, and I think the reason that I say that is they might've said no simply because they don't have time to give it due consideration.* [ImpPartner5_NHS HCP]

Discussion

Key Findings and Implications

This study captures the process of engaging with individuals and public sectors who have the potential to implement "Active Lives," a digital PA intervention for older adults. The emergence of previously unanticipated concerns not documented in existing literature offers novel insights into the challenges and opportunities associated with implementing digital PA interventions for older adults.

One prominent challenge highlighted in our study was the difficulty in identifying suitable implementation partners. Despite an extensive network available to facilitate introductions, finding "open doors," gaining access to decision makers, and advancing discussions about implementation needed considerable time investment spanning several months each time a new person was introduced. Although the academic team was able to communicate the underpinning principles and unique aspects of the intervention and sense-making about how the intervention matched patient or user needs, the time required to engage with multiple partners was incompatible with workloads and incentivization within academia, as noted in previous reports, where academics have attempted to implement digital interventions [21,30].

Furthermore, the divergence in perspectives between HCPs and stakeholders from the community, public health, and PA sectors accentuates the need to accommodate diverse organizational and individual needs and preferences. While HCPs valued the evidence base and viewed Active Lives as a valuable tool for alleviating operational pressures and improving patient health, other sectors relied on local insights and were resistant to shifting beliefs about getting people out of the house and socializing as an optimal strategy for older adults.

Additionally, a competitive landscape for PA interventions added further complexity to bridging gaps in perspectives of Active Lives. The presence of a crowded market in PA support provision was also unanticipated and undocumented in the literature, beyond an overall awareness that people and organizations may have conflicting priorities, agendas, and goals [26]. This study identified how revenue streams from in-person group classes could tie organizations to in-person approaches and how job insecurity and personal involvement with developing and commissioning other PA services might inhibit the uptake of digital interventions. This study also showed how certain individuals and groups seemed to act as significant blocks to implementation (antichampions) because their links to multiple organizations made their influence far-reaching.

A recent review of implementation studies of PA interventions in community settings revealed multiple barriers and facilitators relating to intervention characteristics, setting, individual characteristics, and processes of implementation [18]. However, none of the included studies were digital interventions, and only 1 targeted older adults.

Implementation theories and frameworks highlight the importance of the people involved with the implementation and their influence, attitudes, and behaviors [19,33,34]. Reviews of implementation studies of PA interventions in community settings recommended paying more attention to how individuals affect implementation [18]. However, at the project outset, it was unclear who would be involved, where champions and antichampions would be found, and what their objections to Active Lives would be.

Academic teams leading implementation efforts may not be the best route to scaling digital interventions. Commercialization, with a dedicated team to support initial and sustained implementation work, could be a more feasible route to widespread implementation and scaling up. This is problematic, however, as the intervention being cost-free to both services and end users was an important implementation facilitator.

The principle of considering the implementor's perspectives on suitable local ways of implementing interventions is highlighted within various implementation frameworks [18,34]. However, this study suggests the need to anticipate implementation partners' inclination toward low-reach or overly complex routes and discourage such approaches to maximize the benefits of digital interventions reaching many users at a low cost. Working collaboratively while guiding toward alternatives that are scalable and realistic is important.

However, despite ostensibly straightforward routes, the unpredictable and bureaucratic approval systems for digital interventions within the NHS were a substantial burden, potentially limiting adoption to only the most committed implementation partners. A consistent, proportionate approach is needed within NHS trusts to support HCPs in adopting evidence-based digital interventions that align with clinical and organizational needs. We did, however, identify that champions of Active Lives were motivated by professional benefits from involvement such as prestige and improved networks, a finding also noted elsewhere [18]. Emphasizing these benefits when engaging implementation partners in future projects may help offset their frustrations with the work involved.

Finally, future implementation strategies might need to consider how digital PA interventions can be used as stand-alone or

XSL•FO
RenderX

complementary interventions that can be used alongside existing in-person services. This approach may also overcome concerns regarding equitable access and engagement with PA, specifically in deprived areas.

Strengths and Limitations

This study identified several critical challenges in implementing a digital PA intervention for older adults. Key findings include the time-intensive process of identifying suitable implementation partners, divergent perspectives between HCPs and community sector stakeholders, an unexpectedly competitive landscape for PA interventions, the presence of influential "antichampions," a tendency among partners to suggest low-reach implementation strategies, burdensome approval processes within health care systems, and the incompatibility of engagement timelines with academic workloads. The study also revealed potential facilitators, such as the professional benefits for implementation champions and the incentive of the intervention being cost-free. These findings underscore the complex interplay of organizational, individual, and systemic factors influencing the implementation of digital health interventions, highlighting the need for tailored, multifaceted implementation strategies that address diverse stakeholder concerns and leverage existing organizational structures.

This study's strengths lie in its in-depth qualitative approach, using reflexive thematic analysis of field notes and interviews to explore real-world implementation dynamics. Our research contributes to the relatively underexplored area of implementing digital PA interventions for older adults.

However, several limitations warrant consideration. The brief duration of interviews with implementation partners and potential recruitment bias toward supportive individuals may have constrained the breadth and depth of insights. Furthermore, the focus on early implementation stages and prominent stakeholders limits the generalizability of findings to sustained implementation efforts.

Project funding timescales necessitated conducting interviews early in the implementation process, primarily focusing on key introducers, gatekeepers, and champions. Consequently, the analysis primarily offers insights into implementation partners' decision-making processes regarding Active Lives adoption rather than the ongoing implementation "work" [33]. Additionally, quantitative analysis of uptake, use, engagement, and changes in PA was beyond this study's scope.

Future research could address these limitations by incorporating longitudinal perspectives and quantitative analysis to further elucidate uptake, use, and engagement with digital PA interventions among older adults. Additionally, exploring the perspectives of implementation partners in greater depth and considering the role of contextual factors in shaping implementation outcomes could provide valuable insights for future intervention design and implementation efforts.

Conclusions

This paper examined factors influencing the implementation of the digital PA intervention "Active Lives" from the perspectives of the core implementation team and partners. Key issues include complexities with the identification of influencers and decision makers, navigating approval processes within tight time frames and competing demands. Misunderstandings of Active Lives and differing views on suitable interventions for older adults alongside loyalty to existing interventions emerged as barriers. Even partners eager to implement Active Lives tended to choose low-reach and effortful approaches, but a more directive strategy with anecdotal evidence helped promote passive, high-impact routes.

Acknowledgments

The authors express their gratitude to all implementation partners involved in discussions or activities relating to Active Lives. The authors would also like to acknowledge their patient and public involvement group for their invaluable time, insights, and contributions. The implementation activities and nested research reported here were funded by a grant from the National Institute of Health and Care Research (NIHR) Applied Research Collaboration (ARC) Wessex (IDA: Implementing a Digital Physical Activity intervention for older adults). LY is an NIHR senior investigator, and her research program is partly supported by NIHR ARC West and NIHR Health Protection Research Unit for Behavioural Science and Evaluation. JK is partly funded by NIHR ARC West and NIHR Health Protection Research Unit in Behavioural Science and Evaluation. The views expressed are those of the authors and not necessarily those of the NIHR ARCs or the Department of Health and Social Care.

Conflicts of Interest

KB, SM, and JJ were involved in the development of the Active Lives website and were part of the implementation team in this study.

Multimedia Appendix 1
Implementation team topic guide.
[DOCX File, 18 KB - aging_v8i1e64953_app1.docx]

Multimedia Appendix 2
Implementation partner topic guide.
[DOCX File, 19 KB - aging_v8i1e64953_app2.docx]

XSL•FO
RenderX

Checklist 1
StaRI (Standards for Reporting Implementation Studies) checklist.
[DOCX File, 41 KB - aging_v8i1e64953_app3.docx]

Checklist 2
COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.
[DOCX File, 24 KB - aging_v8i1e64953_app4.docx]

References

1. Northey JM, Cherbuin N, Pumpa KL, Smee DJ, Rattray B. Exercise interventions for cognitive function in adults older than 50: a systematic review with meta-analysis. Br J Sports Med 2018 Feb;52(3):154-160. [doi: 10.1136/bjsports-2016-096587] [Medline: 28438770]
2. Bauman A, Merom D, Bull FC, Buchner DM, Fiatarone Singh MA. Updating the evidence for physical activity: summative reviews of the epidemiological evidence, prevalence, and interventions to promote "Active Aging". Gerontologist 2016 Apr;56 Suppl 2(Suppl 2):S268-S280. [doi: 10.1093/geront/gnw031] [Medline: 26994266]
3. Cunningham C, O' Sullivan R, Caserotti P, Tully MA. Consequences of physical inactivity in older adults: a systematic review of reviews and meta‐analyses. Scandinavian Med Sci Sports 2020 May;30(5):816-827. [doi: 10.1111/sms.13616]
4. Kapteyn A, Banks J, Hamer M, et al. What they say and what they do: comparing physical activity across the USA, England and the Netherlands. J Epidemiol Community Health 2018 Jun;72(6):471-476. [doi: 10.1136/jech-2017-209703] [Medline: 29643112]
5. Ding D, Lawson KD, Kolbe-Alexander TL, et al. The economic burden of physical inactivity: a global analysis of major non-communicable diseases. Lancet 2016 Sep;388(10051):1311-1324. [doi: 10.1016/S0140-6736(16)30383-X]
6. Western MJ, Armstrong MEG, Islam I, Morgan K, Jones UF, Kelson MJ. The effectiveness of digital interventions for increasing physical activity in individuals of low socioeconomic status: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2021 Nov 9;18(1):148. [doi: 10.1186/s12966-021-01218-4] [Medline: 34753490]
7. Internet users, UK: 2020. : Office for National Statistics; 2021 URL: https://www.ons.gov.uk/businessindustryandtrade/itandinternetindustry/bulletins/internetusers/2020 [accessed 2025-08-18]
8. Internet users, UK: 2019. : Office for National Statistics; 2019 URL: https://www.ons.gov.uk/businessindustryandtrade/itandinternetindustry/bulletins/internetusers/2019 [accessed 2025-08-18]
9. Essery R, Pollet S, Bradbury K, et al. Parallel randomized controlled feasibility trials of the "Active Brains" digital intervention to protect cognitive health in adults aged 60-85. Front Public Health 2022;10:962873. [doi: 10.3389/fpubh.2022.962873] [Medline: 36203694]
10. Pollet S, Denison-Day J, Bradbury K, et al. A qualitative exploration of perceptions of a digital intervention to promote physical activity in older adults. J Aging Phys Act 2021 Jun 1;29(3):442-454. [doi: 10.1123/japa.2019-0484] [Medline: 33348316]
11. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan 30;17(1):e30. [doi: 10.2196/jmir.4055] [Medline: 25639757]
12. Essery R, Pollet S, Smith KA, et al. Planning and optimising a digital intervention to protect older adults' cognitive health. Pilot Feasibility Stud 2021 Aug 18;7(1):158. [doi: 10.1186/s40814-021-00884-2] [Medline: 34407886]
13. Skivington K, Matthews L, Simpson SA, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061. [doi: 10.1136/bmj.n2061] [Medline: 34593508]
14. Aromatario O, Van Hoye A, Vuillemin A, et al. How do mobile health applications support behaviour changes? A scoping review of mobile health applications relating to physical activity and eating behaviours. Public Health 2019 Oct;175:8-18. [doi: 10.1016/j.puhe.2019.06.011] [Medline: 31374453]
15. Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. J Med Internet Res 2014 Aug 15;16(8):e192. [doi: 10.2196/jmir.3469] [Medline: 25131661]
16. Middelweerd A, Mollee JS, van der Wal CN, Brug J, Te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. Int J Behav Nutr Phys Act 2014 Jul 25;11(1):1-9. [doi: 10.1186/s12966-014-0097-9] [Medline: 25059981]
17. Presseau J, Byrne-Davis LMT, Hotham S, et al. Enhancing the translation of health behaviour change research into practice: a selective conceptual review of the synergy between implementation science and health psychology. Health Psychol Rev 2022 Mar;16(1):22-49. [doi: 10.1080/17437199.2020.1866638] [Medline: 33446062]
18. Cooper J, Murphy J, Woods C, et al. Barriers and facilitators to implementing community-based physical activity interventions: a qualitative systematic review. Int J Behav Nutr Phys Act 2021 Sep 7;18(1):118. [doi: 10.1186/s12966-021-01177-w] [Medline: 34493306]

XSL•FO
RenderX

19. Koorts H, Eakin E, Estabrooks P, Timperio A, Salmon J, Bauman A. Implementation and scale up of population physical activity interventions for clinical and community settings: the PRACTIS guide. Int J Behav Nutr Phys Act 2018 Jun 8;15(1):51. [doi: 10.1186/s12966-018-0678-0] [Medline: 29884236]
20. Lennon MR, Bouamrane MM, Devlin AM, et al. Readiness for delivering digital health at scale: lessons from a longitudinal qualitative evaluation of a national digital health innovation program in the United Kingdom. J Med Internet Res 2017 Feb 16;19(2):e42. [doi: 10.2196/jmir.6900] [Medline: 28209558]
21. Ainsworth B, Bruton A, Thomas M, Yardley L. One year later: Highlighting the challenges and opportunities in disseminating a breathing-retraining digital behaviour change intervention. Digit Health 2020;6:2055207620936441. [doi: 10.1177/2055207620936441] [Medline: 32728474]
22. Bradbury K, Steele M, Corbett T, et al. Developing a digital intervention for cancer survivors: an evidence-, theory- and person-based approach. NPJ Digit Med 2019;2(1):85. [doi: 10.1038/s41746-019-0163-4] [Medline: 31508496]
23. Seljelid B, Varsi C, Solberg Nes L, Øystese KA, Børøsund E. A digital patient-provider communication intervention (InvolveMe): qualitative study on the implementation preparation based on identified facilitators and barriers. J Med Internet Res 2021 Apr 8;23(4):e22399. [doi: 10.2196/22399] [Medline: 33830063]
24. Biran N, Anthony Kouyaté R, Yucel E, et al. Adaptation and evaluation of a symptom-monitoring digital health intervention for patients with relapsed and refractory multiple myeloma: pilot mixed-methods implementation study. JMIR Form Res 2020 Nov 17;4(11):e18982. [doi: 10.2196/18982] [Medline: 33200997]
25. Morton K, Dennison L, Band R, et al. Implementing a digital intervention for managing uncontrolled hypertension in primary care: a mixed methods process evaluation. Implement Sci 2021 May 26;16(1):57. [doi: 10.1186/s13012-021-01123-1] [Medline: 34039390]
26. Smith J, Essery R, Yardley L, et al. Implementing a health care professional-supported digital intervention for survivors of cancer in primary care: qualitative process evaluation of the renewed intervention. JMIR Cancer 2022 Apr 1;8(2):e36364. [doi: 10.2196/36364] [Medline: 35363143]
27. Benton JS, Cotterill S, Hawkes RE, Miles LM, French DP. Changes in a digital type 2 diabetes self-management intervention during national rollout: mixed methods study of fidelity. J Med Internet Res 2022 Dec 7;24(12):e39483. [doi: 10.2196/39483] [Medline: 36476723]
28. Miller S, Ainsworth B, Weal M, et al. A web-based intervention (Germ Defence) to increase handwashing during a pandemic: process evaluations of a randomized controlled trial and public dissemination. J Med Internet Res 2021 Oct 5;23(10):e26104. [doi: 10.2196/26104] [Medline: 34519661]
29. Agbakoba R, McGee-Lennon M, Bouamrane MM, Watson N, Mair FS. Implementation factors affecting the large-scale deployment of digital health and well-being technologies: a qualitative study of the initial phases of the “Living-It-Up” programme. Health Informatics J 2016 Dec;22(4):867-877. [doi: 10.1177/1460458215594651] [Medline: 26276795]
30. Santer M, Yardley L, Muller I, Roberts A, Thomas KS. How to make Eczema Care Online freely available. BMJ 2022 Dec 8;379:o2973. [doi: 10.1136/bmj.o2973] [Medline: 36740876]
31. Murray E, Hekler EB, Andersson G, et al. Evaluating digital health interventions: key questions and approaches. Am J Prev Med 2016 Nov;51(5):843-851. [doi: 10.1016/j.amepre.2016.06.008]
32. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999 Sep;89(9):1322-1327. [doi: 10.2105/ajph.89.9.1322] [Medline: 10474547]
33. May C, Finch T. Implementing, embedding, and integrating practices: an outline of Normalization Process Theory. Sociology 2009 Jun;43(3):535-554. [doi: 10.1177/0038038509103208]
34. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009 Aug 7;4(1):50. [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
35. Plan-Do-Study-Act (PDSA) directions and examples. Agency for Healthcare Research and Quality. URL: https://www.ahrq.gov/health-literacy/improve/precautions/tool2b.html [accessed 2025-07-17]
36. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
37. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health 2019 Aug 8;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
38. Pinnock H, Barwick M, Carpenter CR, et al. Standards for Reporting Implementation Studies (StaRI) statement. BMJ 2017 Mar 6;356:i6795. [doi: 10.1136/bmj.i6795] [Medline: 28264797]
39. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]

Abbreviations

AHSN: Academic Health Sciences Network
COREQ: Consolidated Criteria for Reporting Qualitative Research

XSL•FO
RenderX

HCP: health care professional
NHS: National Health Service
PA: physical activity
RCT: randomized controlled trial
StaRI: Standard for Reporting Implementation Studies

Edited by F Puga; submitted 31.07.24; peer-reviewed by B Breznen, SA Tomaz; revised version received 12.05.25; accepted 03.06.25; published 21.08.25.

Please cite as:
Boxall C, Dennison L, Miller S, Joseph J, Morton K, Corser J, Kesten J, Electicwala A, Western MJ, Lim S, Grimmett C, Yardley L, Bradbury K
Implementing a Digital Physical Activity Intervention for Older Adults: Qualitative Study
JMIR Aging 2025;8:e64953
URL: https://aging.jmir.org/2025/1/e64953
doi:10.2196/64953

© Cherish Boxall, Laura Dennison, Sascha Miller, Judith Joseph, Kate Morton, Jenny Corser, Joanna Kesten, Asgar Electicwala, Max J Western, Stephen Lim, Chloe Grimmett, Lucy Yardley, Katherine Bradbury. Originally published in JMIR Aging (https://aging.jmir.org), 21.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Smart Health Wearable Adoption Among Singaporean Older Adults Based on Self-Determination Theory: Web-Based Survey Study

Hyunjin Kang[1], PhD; Tingting Yang[1], MA; Nazira Banu[1], MA; Sheryl Wei Ting Ng[1], MA; Jeong Kyu Lee[2], PhD

[1]Wee Kim Wee School of Communication and Information, Nanyang Technological University, 31 Nanyang Link, Singapore, Singapore

[2]Department of Health and Exercise Science, University of Oklahoma, Norman, OK, United States

Corresponding Author:
Hyunjin Kang, PhD
Wee Kim Wee School of Communication and Information, Nanyang Technological University, 31 Nanyang Link, Singapore, Singapore

Abstract

Background: Smart health wearables offer significant benefits for older adults, enabling seamless health monitoring and personalized suggestions based on real-time data. Promoting adoption and sustained use among older adults is essential to empower autonomous health management, leading to better health outcomes, improved quality of life, and reduced strain on health care systems.

Objective: This study investigates how autonomy-related contextual factors, including artificial intelligence (AI) anxiety, perceived privacy risks, and health consciousness, are related to older adults' psychological needs of competence, autonomy, and relatedness (RQ1). We then examined whether the fulfillment of these needs positively predicts older adults' intentions to adopt these devices (H1), and how they mediate the relationship between these factors and older adults' intentions to use smart health wearables (RQ2). Additionally, it compares experienced and nonexperienced older adult users regarding the influence of these psychological needs on use intentions (RQ3).

Methods: A web-based survey was conducted with individuals aged 60 years and above in Singapore, using a Qualtrics survey panel. A total of 306 participants (177 male; mean age of 65.47 years, age range 60 - 85 years) completed the survey. A structural equation model was used to analyze associations among AI anxiety, perceived privacy risks, and health consciousness, and the mediating factors of competence, autonomy, and relatedness, as well as their relationship to smart health wearable use intention.

Results: Health consciousness positively influenced all intrinsic motivation factors—competence, autonomy, and relatedness—while perceived privacy risks negatively affected all three. AI anxiety was negatively associated with competence only. Both privacy risk perceptions and health consciousness were indirectly linked to older adults' intentions to use smart health wearables through competence and relatedness. No significant differences were found in motivational structures between older adults with prior experience and those without.

Conclusions: This study contributes to the application of self-determination theory in promoting the use of smart technology for health management among older adults. The results highlight the critical role of intrinsic motivation—particularly competence—in older adults' adoption of smart health wearables. While privacy concerns diminish motivation, health consciousness fosters it. The study results offer valuable implications for designing technologies that align with older adults' motivations, potentially benefiting aging populations in other technologically advanced societies. Developers should focus on intuitive design, transparent privacy practices, and social features to encourage adoption, empowering older adults to use smart wearables for proactive health management.

(JMIR Aging 2025;8:e69008) doi:10.2196/69008

KEYWORDS

smart health wearables; self-determination theory; AI anxiety; perceived privacy risk; health consciousness

Introduction

Background

Smart health wearables are devices powered by artificial intelligence (AI) designed to be worn on the body, enhancing health management. This technology holds significant promise for the older population, offering a nonintrusive way to continuously monitor health information and provide valuable assistance [1,2]. The use of smart health wearables has been verified in clinical intervention research and proven to assist in self-health management and prevention of chronic diseases,

XSL•FO
RenderX

such as diabetes, hypertension, and dementia [3]. With technological advancements today, some of these technologies have already been integrated into personal devices, such as smartphones and other mobile devices, making it easier for older adults to track their health conveniently [4]. To fully leverage the potential of health wearables, it is essential to promote the adoption and sustained use among older adults, empowering them in autonomous health management. This can lead to better health outcomes, an improved quality of life, and reduced strain on the health care system [5].

The promotion of adoption and sustained use of smart health wearables can be understood through the lens of self-determination theory (SDT) [6,7], which posits that fostering intrinsic motivation, rather than extrinsic, is essential for encouraging prolonged active engagement in a behavior. SDT suggests that fostering an environment which supports an individual's self-determination in engaging in a specific behavior is crucial for satisfying fundamental human needs that drive intrinsic motivation [6]. We examine AI anxiety, privacy risk perceptions, and health consciousness as the critical contextual factors shaping self-determination of using smart health wearables. These 3 contextual factors are closely tied to key smart health wearable aspects. (1) AI anxiety [8] stems from the automated, AI-driven recommendations in smart health wearables, which can make users feel they are losing control over health decisions. (2) Privacy risk perceptions [9] relate to the continuous data collection by smart health wearables; if inadequately protected, users' sense of control over their personal information can be reduced. (3) Health consciousness aligns with smart health wearables' purpose of empowering users to take charge of their health, reinforcing autonomy for proactive self-management [10].

Another key proposition of SDT is that fulfilling the 3 psychological needs of competence, relatedness, and autonomy is essential for intrinsic motivation [6]. When applied to smart health wearables for older adults, the extent to which these devices help older adults feel capable, enable independent choices, and foster a sense of connection with others can shape their intrinsic motivation to engage with the devices and promote sustained usage.

Taken together, this study explores how 3 contextual factors related to self-determination in using smart health wearables—AI anxiety, privacy risk perceptions, and health consciousness—are associated with fulfilling the needs for competence, relatedness, and autonomy, thereby promoting use intentions. Additionally, the study examines whether differences exist between experienced and nonexperienced wearable users in terms of the psychological needs influencing their intentions. Understanding these distinctions provides insight into the motivators for each group at different adoption stages, offering evidence for targeted strategies to sustain user engagement among older adults and thereby promote continuous, effective self-health management [11].

Through this study, we aim to contribute to the literature on self-determination in health promotion through smart technology by identifying key factors that intrinsically motivate older adults to use smart health wearables, which ultimately enhances their long-term well-being. A web-based survey has been conducted in Singapore, one of the world's most connected nations with high information and communication technology penetration and a robust AI ecosystem [12]. Furthermore, Singapore has recently become a "super-aged society," with nearly one-fifth of its population over 65 years old [13]. As such, the insights gained from this research can offer valuable, generalizable implications for designing technologies that align with older adults' motivations, potentially benefiting aging populations in other technologically advanced societies.

Literature Review: Smart Health Wearables for Older Adults' Self-Health Management

Wearables refer to "intelligent computing devices integrated into various accessories, including clothing, fashion accessories, and other everyday items worn by consumers" [14]. Smart health wearables are wearable devices designed to enhance users' health management, offering users a seamless and integrated user experience. Smart health wearables have many benefits, with their main benefit being empowering older adults to take charge of their health [15]. By integrating sensors, (eg, bio, motion, and environmental sensors), internet connection, as well as AI, smart health wearables enable individual users to track and exchange data, and even make smart and personalized decisions for their health management [1]. Some examples of smart health wearables include biosensors, GPS, and radio frequency identification (RFID) technologies that can monitor health conditions by tracking physical information such as blood pressure, oxygen level, and sleep patterns to provide personalized health feedback and medical suggestions. This feedback is valuable for those managing chronic conditions, such as diabetes or cardiovascular diseases [16]. New and improved wearable designs (eg, smartwatches) incorporate advanced algorithms and machine learning to analyze users' data and provide personalized workout programs and nutrition recommendations [17].

The potential of health wearables to improve self-health management is increasingly recognized. However, adoption rates among older adults remain low, largely due to skepticism about the tangible benefits of using wearables [18,19]. Despite a growing body of research exploring enablers and barriers to adopting smart health wearables, the literature remains fragmented [20]. Further studies are needed to understand how these devices are perceived and what motivates older adults to adopt them long-term. For example, recent studies have reported that older adults are more likely to adopt technology if they perceive a clear and immediate benefit from its usage. If the advantages of wearables are not clearly communicated or experienced, older adults may not see them as worth the investment [21]. However, such insights only scratch the surface. More in-depth exploration is necessary to uncover intrinsically motivated factors that can help older adults perceive a clear benefit from adopting wearables. Understanding these factors is crucial to fostering long-term engagement with these wearable devices. This study draws upon SDT and the motivational technology model to identify intrinsic motivators that can encourage older adults to adopt smart health wearable devices for their self-health management.

XSL•FO
RenderX

Three Contextual Factors Related to Autonomous Motivation

Overview

One of the key propositions of SDT is that creating an environment that supports one's self-determination in engaging in a given behavior is essential for fulfilling the key human needs required for fostering intrinsic motivation. In environments that respect and encourage individuals' autonomy, they are more likely to develop a sense of their true selves and achieve self-determination for their actions [6]. In social conditions that pressure people's autonomous behavior, autonomy is deprived. In such situations, the fulfillment of the 3 needs for intrinsic motivation will be hampered. However, the key human needs for intrinsic motivation will be nurtured in autonomy-supportive contexts, where the external environment supports one's self-initiation and choice.

The role of autonomous motivation in encouraging behavior is particularly significant in health-specific contexts. Being in an autonomy-supportive health care climate, for instance, and having autonomous motivation are linked to successful physical activity and dietary management, which, in turn, leads to better health outcomes for patients with type 2 diabetes [22]. Likewise, feelings of autonomous motivation positively impacted the behavioral intentions and usage behaviors of sports apps among students, contrary to feelings of controlled motivation [23]. However, solely using smart health wearables is insufficient to encourage self-health management. A study that sought to promote physical activity among adolescents using health wearables found that, contrary to expectations, feelings of autonomous motivation decreased over an 8-week period because of feelings of peer comparison that created a social environment where participants' autonomy was undermined [24]. This highlights the importance of contextual factors in shaping autonomous motivation to use smart health wearables for health management. Moreover, the technology acceptance model (TAM) [25] and the unified theory of acceptance and use of technology (UTAUT) [26] explain how users accept a given technology. These theories suggest that external factors, such as social influence, should be considered to understand people's attitudes and intentions to use technology.

This study examines the roles of 3 social-contextual factors—AI anxiety, perceived privacy risks, and health consciousness—that are crucial for creating a context that fosters the autonomy of older adults in engaging in self-health management using smart health wearables.

AI Anxiety

There is growing tension between human agency and machine agency in the context of the rise of AI, especially when user experience relies on the algorithms of smart technology [27]. Deci and Ryan [28] defined human agency as motivated behaviors that emanate from one's integrated self. In the context of communication with smart technology, user agency is defined as "the degree to which the self feels like that he or she is a relevant actor" and a feeling that the technology provides "manipulability" for the user to exercise their influence throughout the interaction [29, p. 61]. The feeling of being in control is essential as it bolsters people's intrinsic motivation.

However, as AI techniques advance at an unprecedented speed, people may begin to experience negative psychological tension driven by their natural desire to limit AI agency [27]. AI anxiety describes the psychological status of fear and trepidation when people are concerned that they will lose autonomy and be controlled by AI [8]. It is caused by the constantly evolving technological advancement of AI, along with the challenges and uncertainty that AI brings [30]. For example, Airbnb hosts experienced this algorithm-related AI anxiety as they navigated a complex ecosystem filled with uncertainty and frustration, since the algorithms comprise both "known and unknown factors" [31, p. 9].

While existing research indicates the prevalence of AI anxiety across various populations [32], AI anxiety may be particularly prominent among older adult users. Research shows that age is positively associated with greater anxiety over computers [33]. Although there are ontological differences between traditional computing technologies and AI, AI technologies may generate a wider range of anxieties because of their complexity and perceived autonomy [34]. Furthermore, low levels of technological literacy among older adults may worsen AI anxiety since a key dimension of AI literacy involves the ability to learn AI, which can be particularly challenging for older adults [34,35].

AI anxiety becomes even more pronounced when AI is used to deliver health care services, like in the case of smart health wearables. The decision-making processes of AI models that directly affect the user's well-being may be beyond the understanding of users [36]. Studies have highlighted the fear and frustration expressed by older adults when using smart health wearables to monitor their blood oxygen saturation levels [37], which illustrates how smart health wearables can deepen pre-existing anxieties. Since studies indicate that technological anxiety can negatively affect attitudes and intentions to use technology [3,38,39], we investigate AI anxiety as a significant factor shaping intrinsic motivation for using smart health wearables.

Perceived Privacy Risks

Privacy risks, throughout the various stages of designing, distributing, and using smart health wearables, can cause privacy violations. Technically, (1) most wearable devices are equipped with various sensors (eg, motion, location, and physiological sensors) that continuously collect data, (2) these data are highly personal and thus sensitive, and (3) opportunistic data usage can lead to privacy breaches [40]. Hence, using smart health wearables for health management inevitably triggers people's concerns about smart health wearables' ability to protect privacy. Complicated terms regarding data collection and sharing further exacerbate this concern, with many users unaware of how their data are managed [18,41]. Additionally, incidents involving unauthorized access and cyberattacks compound these concerns [40]. Particularly for older adult users, concerns about privacy and data security are heightened [9].

XSL•FO
RenderX

Moreover, privacy is essential for developing an autonomous self. It is commonly accepted that the human body and health data are "the central mediator of autonomy and individual privacy" [42, p. xvi]. Therefore, individuals' controllability over their information [43] or "selective control of access to the self or to one's group" [44, p. 18] is the core element of privacy. Through such processes of boundary regulation, individuals decide how open they want to be to others. However, smart health wearables need health-related personal information to function (eg, customize exercise plans and make smart recommendations). Hence, there exists this tension between (1) the need to grant controlling power of private information to smart health wearables and (2) people's need for autonomous decision-making about information disclosure. For example, when using smart home applications integrating sensors and other network technologies, older adult users expressed their concern about their controllability over their privacy; however, this sense of tension can be ameliorated if they have "full control of the sensor," meaning they would feel more in control if they have a say about when to activate or deactivate the device [45, p. 4750]. Thus, perceived privacy risks are closely tied to feelings of autonomy, and the control of personal information is key to intrinsic motivation for using health care wearables [46].

Health Consciousness

Health consciousness refers to how individuals care about their health and the level at which they are motivated to engage in preventive health practices such as health information seeking, home-based exercise, and healthy food choices [47]. Specifically, health consciousness reflects people's orientation toward three health-related dimensions: (1) personal health awareness, (2) self-responsibility, and (3) health motivation [48]. Health consciousness has repeatedly been shown to have a positive effect on health-related behaviors like using the internet to search for health information [10], choosing foods based on their health benefits [47], or using dietary and fitness apps [49].

There are several reasons to explain why health consciousness motivates older adults to engage in self-health management. Health-conscious people tend to have a stronger belief in their ability to control their health, which influences their decisions to engage in health-promoting activities [50]. Furthermore, health consciousness can motivate people to be proactive in looking for ways to evaluate their conditions, like taking up information and communication technology to mitigate loneliness [51]. Proactiveness is a distinct characteristic of health consciousness and differs from simply reacting to negative health outcomes. This aligns with research using the health belief model, which suggests that health-related internet use was more strongly proactively motivated by health consciousness, rather than being influenced by perceptions of one's health risk in a reactive manner [10]. Smart health wearables that are designed to enable self-directed health management activities [52] precisely meet the needs of health-conscious individuals. Thus, it would be fruitful to assess health consciousness as a potential factor encouraging the adoption of smart health wearables among older adults.

Three Basic Needs for Self-Determination

We used the SDT [6,7] to examine the factors related to the motivation for adopting smart health wearables among older adults. The core idea of SDT is that intrinsic motivation is necessary for sustaining engagement in a particular behavior. The 3 basic needs—competence, relatedness, and autonomy—are fundamental sources of intrinsic motivation and the natural tendency for self-development.

Competence refers to the ability to achieve personal goals. Autonomy is the freedom to make self-initiated choices, and relatedness reflects feeling connected and cared for by social groups or communities. When individuals perceive that a particular behavior fulfills these 3 basic needs, they tend to find it enjoyable (ie, intrinsically motivating) and will engage in it without relying on external rewards. Hence, from the lens of SDT, this study examines how the 3 psychological needs shape intrinsic motivation and ensures the long-term use of smart health wearables for older adults' self-health management.

As with other interactive technologies, design elements in smart health wearables can influence the sense of competence, autonomy, and relatedness through their intelligent, user-focused features [53]. For example, features such as goal setting and reward badges can enhance users' competence by providing feedback on their achievements and progress. Yet, older adults often struggle with setting up or understanding the data generated by these devices, which can undermine their sense of competence [54]. Additionally, while personalized nudges and recommendations from these devices may sometimes compromise users' sense of autonomy by making decisions based on their data [27], customizable interfaces can empower users to feel more in control. Furthermore, social features—such as communities, friend invitations, and participation in challenges—can enhance a sense of relatedness among users [24,53]. Wu and Lim [21] found that older adults are more likely to use wearables when family members or friends in their close social circle also adopt the devices.

Overall, this study explores 3 social-contextual factors influencing the autonomous use of smart health wearables for health management—AI anxiety, perceived privacy risks, and health consciousness—and their potential associations with the psychological needs related to intrinsic motivation among older adults. Accordingly, we pose the following research question:

RQ1: How are autonomy-related contextual factors regarding smart health wearable use (ie, AI anxiety, perceived privacy risks, and health consciousness) associated with the sense of competence, autonomy, and relatedness among older adults?

In addition, we aim to unpack the relationship between the psychological needs of older adults and their intention to use smart health wearables. Self-determined motivation shaped by the fulfillment of these 3 needs plays a crucial role in people's intention to use technology [55]. For instance, a study found that intentions to use health technology, such as wellness clouds for health tracking, were positively associated with intrinsic motivation, driven by expectations of performance, playfulness, and ease of use [56]. Thus, we hypothesize:

XSL•FO
RenderX

H1: Perceived competence, autonomy, and relatedness in using smart health wearables will be positively associated with older adults' intentions to use these devices.

We expect that the needs of competence, autonomy, and relatedness fulfilled by health wearables use would mediate the relationships between the autonomy-related contextual factors and use intention. This study will explore these indirect associations by proposing the following research question:

RQ2: How do perceived competence, autonomy, and relatedness in using smart health wearables mediate the relationships between the autonomous-related contextual factors (ie, AI anxiety, perceived privacy risks, and health consciousness) and older adults' intentions to use these devices?

Experienced Versus Nonexperienced Users

We also examine potential differences in motivation between older adults with and without prior experience using smart health wearables. According to the UTAUT and TAM models, as discussed earlier, users' prior experience with using the technology is also one of the key variables, as prior experience can work as a conditional factor in changing the degree of effect of other antecedent factors on technology acceptance (eg, intention to use and attitudes toward a technology) [26]. Specifically, whether a user has experience using a technology may affect their evaluation of motivations, effort expectancy, social influence, and perceived ease of use, among other factors [26,57]. For example, one study about smart health wearables found that the relation between behavior intention and actual use is stronger for more experienced users, compared with the new users; in addition, the experienced users rated the facilitating conditions (eg, resource and knowledge they received) as less important [58]. Interestingly, among older adults, whether users have prior experience also showed a difference in what types of smart health wearables they choose: compared with experienced users whose choice focuses mostly on smartwatches and wristbands, nonexperienced users' choices were more dispersed, showing a wide variety of preferences such as smart rings and clips among others. This indicates that prior experience might affect users' overall expectations among many technology acceptance indicators [59], supporting the proposition that prior experience may be a conditional factor that differentiates users' motivational structure predicting their technology adoption intentions [60], as experienced users tend to exhibit greater confidence and skill in navigating wearable features [61,62]. Experienced older adults may more readily recognize benefits such as improvements in health metrics or physical activity [16]. Conversely, older adults with limited or no prior experience may have different expectations regarding wearables, often accompanied by uncertainty. While such uncertainty can foster positive expectations about the benefits of devices, it may also provoke feelings of intimidation. Older adults without prior experience may struggle to fully grasp the benefits of smart health wearables early on. Moreover, without clear evidence of the device's impact on their well-being, these older adults might question the use of the technology, resulting in lower intrinsic motivation to adopt such devices [63].

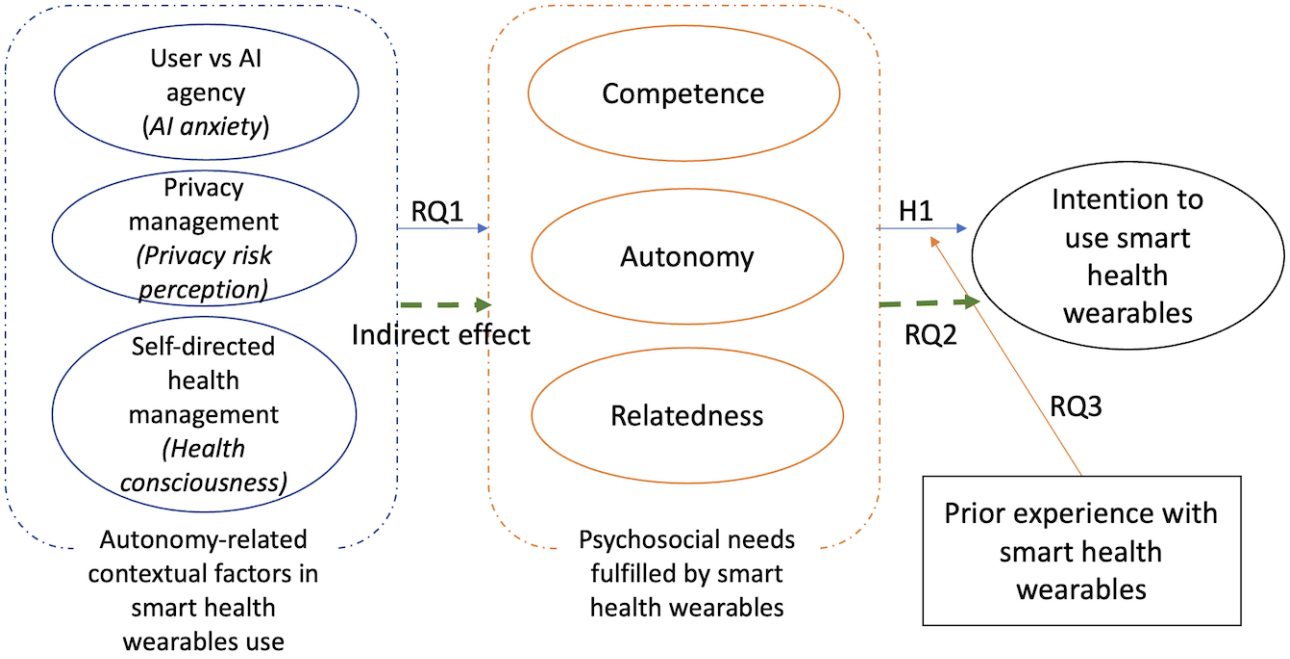
However, the moderating effect of prior experience from existing literature shows inconsistent effects. For example, one study examined whether users' prior experience of using robots predicts their approval of autonomous delivery robots, and the correlation was not significant [60]. In light of these discrepancies, we will examine whether and how these 2 groups—those with and without prior experience—differ in their motivational structures for using smart health wearables in health management. Understanding the motivational differences between those who are experienced versus not experienced users of smart health wearables offers valuable insights into the distinct needs of users in different adoption stages. These insights can guide the development of more user-friendly designs and tailored educational strategies, ultimately encouraging wider adoption among hesitant or inexperienced individuals. Accordingly, we propose the following research question:

RQ3: How are experienced older adult users different from the nonexperienced older adult group regarding the relationship between the 3 psychological needs and use intention?

Figure 1 summarizes the model tested in this study.

XSL•FO
RenderX

Figure 1. Hypothesized model summary. AI: artificial intelligence; H: hypothesis; RQ: research question.



Methods

Data Collection

A web-based survey was conducted with individuals over the age of 60 years in Singapore from March to April 2022. For data collection, we used a survey panel from Qualtrics, one of the major global survey companies, which provides panels across various age groups in different countries. In total, 306 participants (177 male; mean age of 65.47 years, age range 60 - 85 years; ethnicity: 84.6% [259/306] Chinese, 3.3% [10/306] Malay, 6.5% [20/306] Indian, and 5.6% [17/306] others) completed the survey. At the beginning of the study, they were given descriptions and examples of smart health wearables. Then, they were asked if they had used such devices before. Of the 306 participants, 163 (53.3%) responded "yes" indicating their prior experiences with smart health wearables, and fitness trackers were the most common type of wearable used (135/306, 44.1%). Meanwhile, 143 (46.7%) responded "no" to the question asking about their prior experience.

Ethical Considerations

The survey procedure and materials were reviewed and approved by the institutional review board at Nanyang Technological University (IRB-2022-163). Informed consent was obtained before participants began the survey, and they had the option to opt out. All collected data were deidentified. Compensation was provided to participants by the survey company in the form of credit points in accordance with its policy.

Measurement

AI Anxiety

To measure the anxiety caused by uncertainty about how AI works, 3 items were adapted from Meuter et al [64], originally developed for measuring user anxiety about service technology. A 7-point Likert scale was used (1=strongly disagree to 7=strongly agree; eg, "I feel apprehensive about using AI technology").

Perceived Privacy Risk

Perceived privacy risk regarding wearable health technology was measured using 3 items [65], each rated on a 7-point Likert scale (1=strongly disagree to 7=strongly agree; eg, "It would be risky to disclose my personal health information to wearable device vendors").

Health Consciousness

Health consciousness was measured using a 3-item scale, with participants indicating their agreement with each statement on a 7-point Likert scale (1=strongly disagree to 7=strongly agree; eg, "I take responsibility for the state of my health") [66].

Self-Determination Factors

Adapted from Ryan et al [67], 3 factors of self-determination—competence, autonomy, and relatedness—were measured using a 7-point Likert scale (1=strongly disagree to 7=strongly agree). Competence in using wearable health devices was measured using a 3-item scale (eg, "I would feel competent in using a smart wearable health care device"). Autonomy was assessed with a 4-item scale (eg, "I would feel free to decide for myself how to do things on a smart wearable health care device"). Relatedness to other users of wearable health care devices was measured using 3 items (eg, "I would find the relationships with other users of smart wearable health care devices fulfilling").

Future Use Intention

Intention to adopt smart wearable health care technology was measured using a 3-item semantic differential scale. Participants rated their likelihood, probability, and willingness to use smart wearable health care technology in the next 3 months, with responses ranging from 1 to 7 (eg, likelihood: 1=unlikely to 7=likely).

XSL•FO
RenderX

Refer to Table 1 for the measurement items, internal reliability (Cronbach α), and basic statistics for each factor's composite score.

XSL•FO
RenderX

Table . Measurement items, internal reliability, and descriptive statistics.

Measures and items			Cronbach α	Mean (SD)
AI[a] anxiety			0.78	3.95 (1.28)
	I feel apprehensive about using AI technology.			
	I hesitate to use AI technology for fear of making mistakes I cannot correct.			
	I have difficulty understanding AI-related technological matters.			
Perceived privacy risk			0.92	4.40 (1.44)
	It would be risky to disclose my personal health information to wearable device vendors.			
	There would be a high potential for loss associated with disclosing my personal health information to vendors providing wearable devices.			
	There would be too much uncertainty associated with giving my personal health information to vendors providing wearable devices.			
Health consciousness			0.63	5.37 (1.06)
	I am concerned about my health all the time.			
	I notice how I feel physically as I go through the day.			
	I take responsibility for the state of my health.			
Competence			0.95	4.83 (1.31)
	I would feel competent in using a smart wearable health care device.			
	I would feel capable when using a smart wearable health care device.			
	I would feel like I am effective when using a smart wearable health care device.			
Autonomy			0.91	4.70 (1.17)
	I would feel free to decide for myself how to do things on a smart wearable health care device.			
	I would generally feel free to express my ideas and opinions on a smart wearable health care device.			
	I would feel like I can pretty much be myself when I use a smart wearable health care device.			
	I would experience a lot of freedom when I use a smart wearable health care device.			
Relatedness			0.70	4.03 (.97)
	I would find the relationships with other users of smart wearable health care devices fulfilling.			
	I would find the relationships with other users of smart wearable health care devices important.			
	I would not feel close to other smart wearable health care device users (reverse-coded).			
Intention			0.94	4.78 (1.93)
	How likely are you to use smart wearable health care devices in the next 3 months?			
		Unlikely or likely (1 to 7 scale)		
		Not probable or probable (1 to 7 scale)		

XSL•FO
RenderX

Measures and items	Cronbach α	Mean (SD)
Unwilling or willing (1 to 7 scale)		

[a]AI: artificial intelligence.

Measurement Model

The measurement model was tested using comparative fit index (CFI) analysis. Covariances for the first 3 autonomy items and the first 2 competence items were allowed to improve model fit. After this modification, the model showed good fit: χ^2_{185}=365.9, *P*<.001, root mean square error of approximation (RMSEA)=0.057 (90% CI 0.048-0.065), CFI=0.966, Tucker-Lewis index (TLI)=0.957, and standardized root mean square residual (SRMR)=0.048, all within acceptable ranges.

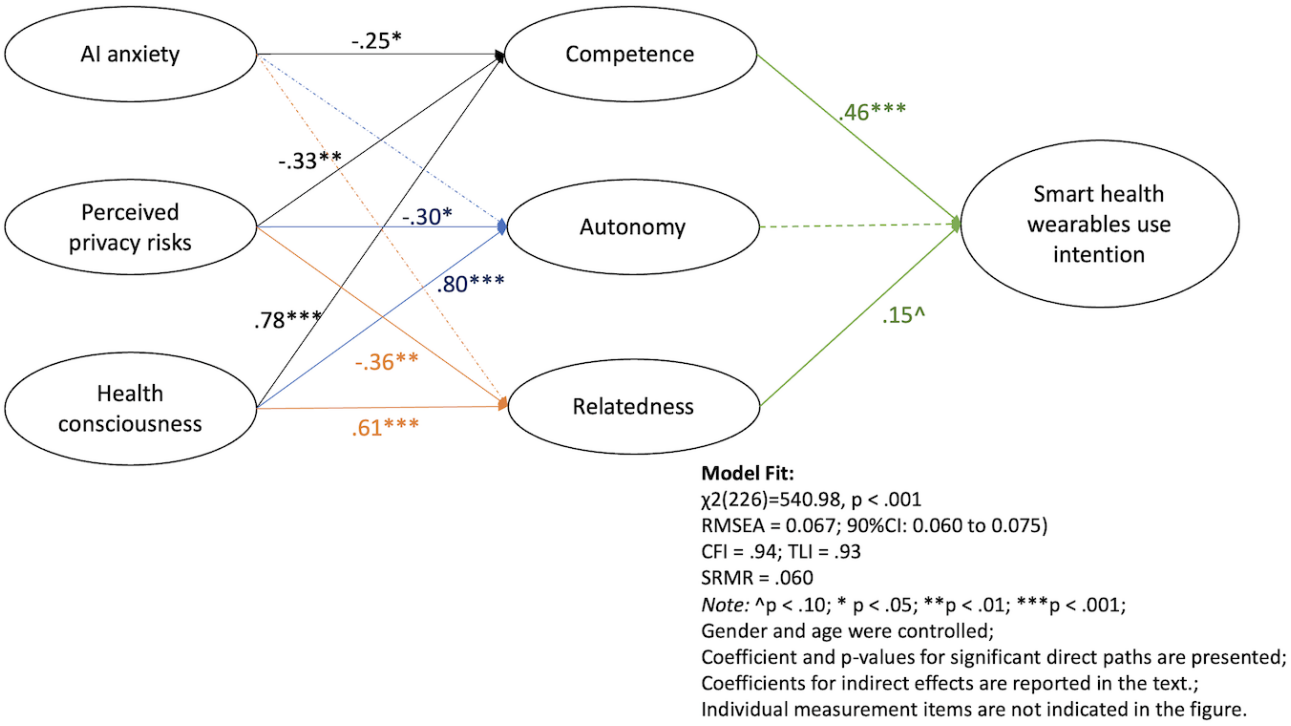
Results

The structural equation model (SEM) results indicate significant associations among AI anxiety, perceived privacy risks, and health consciousness, and the mediating factors of competence, autonomy, and relatedness, which are related to smart health wearable use intention. Model fit indices suggested a good fit: χ^2_{226}=541, *P*<.001; RMSEA=0.067 (90% CI 0.06-0.075); CFI=0.94; TLI=0.93; SRMR=0.06.

For RQ1, AI anxiety was negatively associated only with competence (β=−.25, *P*=.039); perceived privacy risks showed a significant negative association with competence (β=−.33, *P*=.002), autonomy (β=−.3, *P*=.012), and relatedness (β=−.36, *P*=.001). Health consciousness was positively associated with competence (β=.78, *P*<.001), autonomy (β=.8, *P*<.001), and relatedness (β=.61, *P*<.001).

H1 examined the associations between the 3 self-determination model dimensions and the intention to use smart health wearables. Competence was significantly associated with use intention (β=.46, *P*<.001), and relatedness also showed a marginally significant association (β=.15, *P*=.05). Figure 2 summarizes the SEM results.

Figure 2. Structural equation model results for RQ1 and H1. AI: artificial intelligence; CFI: comparative fit index; H: hypothesis; RMSEA: root mean square error of approximation; RQ: research question; SRMR: standardized root mean square residual.



To answer RQ2, we examined the indirect effects of AI anxiety, perceived privacy risks, and health consciousness on use intention through competence, autonomy, and relatedness as mediators, using 5000 bootstrapped samples. The total indirect effect of AI anxiety on use intention was not statistically significant (*B*=−0.31, 95% CI −0.78 to 0.05) although the indirect pathway through competence was significant in a negative direction (*B*=−0.27, 95% CI −0.71 to −0.05). For perceived privacy risks, the total indirect effect was significant in a negative direction (*B*=−0.36, 95% CI −0.59 to −0.14), via competence (*B*=−0.24, 95% CI −0.52 to −0.08) and relatedness (*B*=−0.09, 95% CI −0.23 to −0.004). Lastly, health consciousness showed a significant total indirect effect on use intention in a positive direction (*B*=1.37, 95% CI 1.01-2.08), with a strong positive mediation through competence (*B*=0.97, 95% CI 0.44-1.78) and relatedness (*B*=0.25, 95% CI 0.02-0.58). We found no significant indirect pathways through autonomy for any predictors.

Finally, RQ3 asked if experienced versus not experienced older adult users would have different motivation structures shaping the intention to use smart health wearables. First, measurement

XSL•FO
RenderX

invariance between experienced versus nonexperienced users for the 3 self-determination factors and use intention were tested. The models allowing for free estimation of all parameters across groups showed a good fit to the data: χ^2_{110}=183.8, *P*<.001; RMSEA=0.066, 90% CI 0.049-0.083; CFI=0.979; TLI=0.971; and SRMR=0.045. Additionally, the chi-square difference test comparing the model constraining the factor loadings to be equal across groups, against the unconstrained model, was not significant ($\Delta\chi^2_9$=6.8, *P*=.66), indicating that the factor loadings are invariant across groups. To assess the relationship between the 3 motivation factors and behavioral intention across experienced and nonexperienced user groups, 2 SEMs were estimated: an unconstrained path coefficient model (χ^2_{18}=38.5, *P*=.003, RMSEA=0.086, 90% CI 0.048-0.124, CFI=0.968) and a constrained model (χ^2_{30}=50.5, *P*=.011, RMSEA=0.067, 90% CI 0.032-0.098, CFI=0.968), both showing acceptable model fits. However, the chi-square difference test indicated that the relationships between competence, autonomy, and relatedness, and behavioral intention between the 2 groups do not statistically differ ($\Delta\chi^2_{12}$=12, *P*>.05).

Discussion

Theoretical Implications

Drawing on SDT, this study sheds light on the motivational factors influencing older adults' intention to use smart wearables for health care. By examining external contextual factors that may shape older adults' autonomous use of smart wearables for health management, we provide more nuanced insights into the self-determination model, enhancing the understanding of how to promote self-determined, proactive health management through technology. Moreover, this age-specific insight contributes to a deeper understanding of how SDT can be adapted to different user populations in the smart health wearables context.

Our study examined how external contextual factors influence intrinsic motivation, particularly in the adoption of smart health wearables among older adults (RQ1). The results reveal that perceived privacy risks erode all motivation factors: competence, autonomy, and relatedness. The fear of data breaches or the misuse of personal health information can make older adults feel that they lack control over their personal data, directly undermining their sense of autonomy. Perceived privacy risks also affect competence, as older adults may feel uncertain about their ability to navigate complex privacy settings or how their data are collected and used. This uncertainty creates a barrier to engaging with smart health wearables, as users who feel less competent in managing their personal information may be less likely to adopt such technologies. The negative association between privacy risk perception and relatedness may imply that fears of data misuse or unauthorized access can undermine users' willingness to engage in social features of smart health wearables, such as sharing health information with peers or participating in health communities. We also found a significant association between AI anxiety and competence, among 3 intrinsic motivation factors. AI anxiety is characterized by fear and apprehension about AI's decision-making processes [8]. Older adults, who may already feel less confident in their technological abilities, may find AI's opaque algorithms intimidating, leading them to doubt their capacity to effectively use smart health wearables, consequently reducing their intrinsic motivation to use them.

However, health consciousness strongly supports the fulfillment of all 3 psychological needs among older adult users; health-conscious older adults are more likely to feel competent, autonomous, and socially connected through the use of smart health wearables. This result indicates that the tendency of being proactive in self-health management among health-conscious individuals [47,50] can be largely translated to self-determined smart health wearables adoption. Especially, the significant association between health consciousness and relatedness suggests that older adults who care about their health are more likely to engage with smart health wearables and connect with others through health communities and shared goals, not just for self-health management. These findings suggest that while SDT offers a strong framework for understanding motivation across various domains, it may require adjustments for technology adoption among older adults.

We then examined the associations between the 3 psychological needs that foster intrinsic motivation and the intention to use smart health wearables (H1). The results indicate that, for older adults, feeling capable of using smart health wearables (competence) is a significant driver of wearable use, and experiencing social connection (relatedness) showed a marginally significant association. However, independent decision-making (autonomy) was not a significant factor. This result might be because older adults recognize the benefits of delegating monitoring and decision-making for health management to technology, which provides personalized recommendations for users. This result suggests that older adults may prioritize supported autonomy in using smart health wearables. Studies [29,68] have suggested that users seek to benefit from the autonomous capacity of AI-based technology while maintaining control over it.

It is also worth noting that our result contradicts Jung and Kang [52], who found that only autonomy, among the 3 intrinsic motivation factors, significantly predicted enjoyment in using smart fitness wearables, based on a survey conducted with a general population in the United States. This also indicates a possibility that autonomy might be experienced or expected differently among the older population in Singapore. Future research should explore how older adults balance autonomy and supported autonomy through health technologies and how smart health wearables can foster this balance without diminishing their sense of agency, especially through comparison with young users.

The mediating roles of the 3 psychological needs between contextual factors and the intention to use smart health wearables (RQ2) offer a deeper understanding of the factors influencing self-determination in wearable health technology use. The findings underscore the importance of fostering feelings of competence to encourage smart wearable adoption among older adults. Competence, as a key driver of intrinsic motivation, is particularly crucial for this demographic, who may feel

XSL•FO
RenderX

intimidated or uncertain about using such technology [54]. The strong indirect associations between privacy risk perceptions, health consciousness, and usage intentions through competence suggest that merely addressing privacy concerns and promoting health awareness may be insufficient. Instead, efforts should also focus on clearly communicating privacy safeguards and emphasizing the value of self-directed health management in a way that reinforces older adults' confidence and competence, ultimately helping them integrate wearables into their health routines successfully.

While relatedness played a secondary role, it nonetheless contributes to older adults' intrinsic motivation, aligning with previous research identifying relatedness as an important intrinsic motivator in social media use [69]. Our study indicates that the feelings of being connected to peer users, family, and caregivers remain a significant need for older adult users even when engaging with technology primarily intended for purposes other than social interaction or relationship building.

Finally, the lack of significant differences between experienced and nonexperienced users regarding the associations between motivational factors and use intention (RQ3) suggests that the need for competence and relatedness is important in shaping the intention to adopt smart health wearables, regardless of prior experience with such devices. Previous research found that individuals with prior experience using smart devices (eg, smartphones and smart televisions) perceived them as less complex and were more willing to try new technologies, as compared to those without experience [19]. However, the absence of differences between the 2 groups in our study could be interpreted in 2 ways. First, advancements in user-centered design and AI may have made smart health wearables more intuitive and less cognitively demanding [18], reducing the potential barriers or concerns around competence in using these devices. Alternatively, it is possible that smart health wearables remain complex and intimidating, even for older adults who have had prior experience with similar technology. Future research could use a qualitative approach, such as in-depth interviews with older adult users, to explore the underlying reasons behind this nonsignificant result. In addition, for older adult users, connecting with family and friends was identified as a major gratification in using new media technology [69-71]. Our study demonstrates that, for both experienced and nonexperienced groups, fostering relatedness with others remains an important reason for adopting health wearables among older adults, which often did not hold truth in other populations [24,52].

Practical Implications

Our findings indicate that perceived privacy risks can negatively impact older adults' sense of competence and autonomy when adopting smart health wearables. To address perceived privacy risks, smart health wearables should offer easy-to-use privacy settings that give older adults control over how their data are shared. Transparent communication about data security can enhance users' sense of control and trust, encouraging wider adoption and sustained use.

Competence emerged as the strongest predictor of smart health wearables adoption. Therefore, developers should focus on creating intuitive, easy-to-use devices tailored to older adults, with guides, tutorials, and support resources that enhance users' confidence and ability. Features such as personalized feedback, goal-setting, and progress tracking can further reinforce older adults' sense of accomplishment in managing their health, fostering continued engagement. By helping older adults feel capable and effective, developers can promote long-term use of wearables.

The feeling of social connectedness (relatedness) also plays a significant role in driving smart health wearables adoption. Developers should incorporate social features that allow older adults to connect with family, friends, or health communities—such as shared health data, group challenges, and support groups. These features foster a sense of social support, enhancing motivation. Health care providers and caregivers can encourage older adults to use smart health wearables in social settings, which can strengthen their motivation for ongoing use.

The nonsignificant difference between experienced and nonexperienced users in terms of the motivational structure for smart health wearables use suggests that both groups rely on feeling competent and socially connected when deciding to use smart health wearables. From a practical standpoint, this indicates that interventions aimed at enhancing these motivational factors—such as simplifying the technology to improve competence and fostering social connections—are equally important for both groups. For example, social workers and government organizations can provide training courses for older adults focusing on these aspects. A recent study showed that even short training courses (eg, teaching older adults how to use smartphones for social interaction and medical use) increased the positive effect of older adult users' effort expectancy (how understandable and easy to learn to use the technology) on behavior intention [72]. Therefore, strategies to increase smart health wearables adoption should focus on addressing these core needs universally, rather than differentiating between experienced users and nonexperienced users.

Limitations and Future Research

We acknowledge several limitations in this study. First, our cross-sectional design limits the ability to establish causal relationships between autonomy-related factors and wearables adoption. As a result, the associations observed in this study are correlational. Longitudinal research would be valuable in examining how variables such as AI anxiety, privacy concerns, and health consciousness impact both short- and long-term use of wearables over time, providing a clearer understanding of causality.

Second, this study was conducted via a web-based survey, which inherently requires participants to have a certain level of digital literacy and internet access. Consequently, our sample may not fully represent older adults with limited internet access or low digital literacy. To capture a more inclusive profile of older adult participants, future studies should consider using offline methods, such as paper-based surveys, to ensure a broader representation of older adults with varying digital skills.

XSL•FO
RenderX

Third, while our research primarily focuses on intrinsic motivation among older adult users for adopting smart health wearables, actual behavioral intention (ie, intention to use) is also influenced by external factors, such as the usability designs of the devices. Studies indicate that older adults have specific performance expectations regarding wearable functionality and ease of use [73]. Therefore, future studies should consider both intrinsic and external factors when exploring predictors of older adults' adoption and sustained use of smart health wearables.

Fourth, the high ownership rate of fitness trackers among our participants (135/306, 44.1%) can be attributed to Singapore's National Steps Challenge, initiated in 2015, which provided free fitness trackers to encourage physical activity. This unique context may have influenced our results, highlighting the need for further cross-cultural studies in other countries to gain complementary insights and verify the generalizability of these findings across different cultural and health care contexts.

Finally, according to the National Population and Talent Division, 81.8% of those over 60 years old are ethnic Chinese, 11% are Malay, 8% are Indian, and 1% belong to other ethnicities as of 2024 [74]. While our sample largely reflects the overall ethnic group composition, Chinese (259/306, 84.6%) and other ethnicities (17/306, 5.6%) are slightly overrepresented, and Malay (10/306, 3.3%) and Indian participants (20/306, 6.5%) are slightly underrepresented. Given that cultural and ethnic differences could significantly influence health-related [75] as well as technology and privacy-related factors [76,77], future research should aim for a more representative sample to capture the diverse experiences and preferences of all ethnic groups, ensuring broader applicability of the findings.

Conclusions

Through the lens of SDT, this study advances our understanding of older adults' adoption of smart health wearables. The results highlighted the critical roles of competence, relatedness, and autonomy-supporting contexts in shaping their intrinsic motivation. While perceived privacy risks and AI anxiety negatively impact motivation, health consciousness emerges as a strong enabler of self-determined adoption of health wearables. Our findings emphasize the need for user-friendly designs, robust privacy safeguards, and social features that foster connection and confidence among older adults. By addressing these factors, developers and stakeholders can better support older adults in integrating smart wearables into their proactive health management routines.

Acknowledgments

This study was supported by the Ministry of Education in Singapore Tier 1 Grant (to HK) and the University of Oklahoma Libraries' Open Access Fund (to JKL).

Conflicts of Interest

None declared.

References

1. Hiremath S, Yang G, Mankodiya K. Wearable Internet of Things: concept, architectural components and promises for person-centered healthcare. Presented at: 2014 4th International Conference on Wireless Mobile Communication and Healthcare – Transforming Healthcare Through Innovations in Mobile and Wireless Technologies (MOBIHEALTH); Nov 3-5, 2014; Athens, Greece p. 304-307. [doi: 10.4108/Mobihealth33544.2014.7015971]
2. Peng C, Zhao H, Zhang S. Determinants and cross-national moderators of wearable health tracker adoption: a meta-analysis. Sustainability 2021;13(23):13328. [doi: 10.3390/su132313328]
3. Jeng MY, Pai FY, Yeh TM. Antecedents for older adults' intention to use smart health wearable devices—technology anxiety as a moderator. Behav Sci (Basel) 2022 Apr 18;12(4):114. [doi: 10.3390/bs12040114] [Medline: 35447686]
4. Liu F, Ngai E, Ju X. Understanding mobile health service use: an investigation of routine and emergency use intentions. Int J Inf Manage 2019 Apr;45:107-117. [doi: 10.1016/j.ijinfomgt.2018.09.004]
5. Metcalf D, Milliard STJ, Gomez M, Schwartz M. Wearables and the Internet of Things for health: wearable, interconnected devices promise more efficient and comprehensive health care. IEEE Pulse 2016;7(5):35-39. [doi: 10.1109/MPUL.2016.2592260] [Medline: 28113167]
6. Deci EL, Ryan RM. Intrinsic Motivation and Self-Determination in Human Behavior: Springer US; 1985. [doi: 10.1007/978-1-4899-2271-7]
7. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000;55(1):68-78. [doi: 10.1037/0003-066X.55.1.68]
8. Johnson DG, Verdicchio M. AI anxiety. J Assoc Inf Sci Technol 2017 Sep;68(9):2267-2270. [doi: 10.1002/asi.23867]
9. Xing F, Peng G, Zhang B, Li S, Liang X. Socio-technical barriers affecting large-scale deployment of AI-enabled wearable medical devices among the ageing population in China. Technol Forecast Soc Change 2021 May;166:120609. [doi: 10.1016/j.techfore.2021.120609]
10. Ahadzadeh AS, Pahlevan Sharif S, Ong FS, Khong KW. Integrating health belief model and technology acceptance model: an investigation of health-related internet use. J Med Internet Res 2015 Feb 19;17(2):e45. [doi: 10.2196/jmir.3564] [Medline: 25700481]

XSL•FO
RenderX

11. Windasari NA, Lin FR, Kato-Lin YC. Continued use of wearable fitness technology: a value co-creation perspective. Int J Inf Manage 2021 Apr;57:102292. [doi: 10.1016/j.ijinfomgt.2020.102292]
12. AI Singapore. AI Singapore. 2022. URL: https://aisingapore.org [accessed 2024-11-16]
13. Lau D. How Singapore is preparing for a super aged society come 2026. CNA. 2024. URL: https://www.channelnewsasia.com/today/big-read/super-aged-2026-singapore-ready-4656756 [accessed 2024-11-16]
14. Wright R, Keith L. Wearable technology: if the tech fits, wear it. J Electron Resour Med Libr 2014 Oct 2;11(4):204-216. [doi: 10.1080/15424065.2014.969051]
15. Li J, Ma Q, Chan AH, Man SS. Health monitoring through wearable technologies for older adults: smart wearables acceptance model. Appl Ergon 2019 Feb;75:162-169. [doi: 10.1016/j.apergo.2018.10.006] [Medline: 30509522]
16. Cheung ML, Chau KY, Lam MHS, et al. Examining consumers' adoption of wearable healthcare technology: the role of health attributes. Int J Environ Res Public Health 2019 Jun 26;16(13):2257. [doi: 10.3390/ijerph16132257] [Medline: 31247962]
17. Yadav SS, Agarwal R, Bharath K, Rao S, Thakur CS. tinyRadar for fitness: a contactless framework for edge computing. IEEE Trans Biomed Circuits Syst 2023 Apr;17(2):192-201. [doi: 10.1109/TBCAS.2023.3244240] [Medline: 37022890]
18. Garcia Reyes EP, Kelly R, Buchanan G, Waycott J. Understanding older adults' experiences with technologies for health self-management: interview study. JMIR Aging 2023 Mar 21;6:e43197. [doi: 10.2196/43197] [Medline: 36943333]
19. Farivar S, Abouzahra M, Ghasemaghaei M. Wearable device adoption among older adults: a mixed-methods study. Int J Inf Manage 2020 Dec;55:102209. [doi: 10.1016/j.ijinfomgt.2020.102209]
20. Ferreira JJ, Fernandes CI, Rammal HG, Veiga PM. Wearable technology and consumer interaction: a systematic review and research agenda. Comput Human Behav 2021 May;118:106710. [doi: 10.1016/j.chb.2021.106710]
21. Wu C, Lim GG. Investigating older adults users' willingness to adopt wearable devices by integrating the technology acceptance model (UTAUT2) and the Technology Readiness Index theory. Front Public Health 2024;12:1449594. [doi: 10.3389/fpubh.2024.1449594] [Medline: 39421816]
22. Koponen AM, Simonsen N, Suominen S. Determinants of physical activity among patients with type 2 diabetes: the role of perceived autonomy support, autonomous motivation and self-care competence. Psychol Health Med 2017 Mar;22(3):332-344. [doi: 10.1080/13548506.2016.1154179] [Medline: 26952696]
23. Guo J. Influencing factors of college students' use of sports apps in mandatory situations: based on UTAUT and SDT. Biomed Res Int 2022;2022(1):9378860. [doi: 10.1155/2022/9378860] [Medline: 36193303]
24. Kerner C, Goodyear VA. The motivational impact of wearable healthy lifestyle technologies: a self-determination perspective on Fitbits with adolescents. Am J Health Educ 2017 Sep 3;48(5):287-297. [doi: 10.1080/19325037.2017.1343161]
25. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
26. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
27. Sundar SS. Rise of machine agency: a framework for studying the psychology of human–AI interaction (HAII). J Comput Mediat Commun 2020 Mar 23;25(1):74-88. [doi: 10.1093/jcmc/zmz026]
28. Deci EL, Ryan RM. Human autonomy. In: Kernis MH, editor. Efficacy, Agency, and Self-Esteem: Springer US; 1995:31-49. [doi: 10.1007/978-1-4899-1280-0_3]
29. Sundar SS. Self as source: agency and customization in interactive media. In: Mediated Interpersonal Communication: Routledge; 2008:72-88. [doi: 10.4324/9780203926864-12]
30. Lemay DJ, Basnet RB, Doleck T. Fearing the robot apocalypse: correlates of AI anxiety. Int J Learn Anal Artif Intell Educ IJAI 2020;2(2):24. [doi: 10.3991/ijai.v2i2.16759]
31. Jhaver S, Karpfen Y, Antin J. Algorithmic anxiety and coping strategies of Airbnb hosts. Presented at: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18); Apr 21-26, 2018; Montreal QC, Canada p. 1-12. [doi: 10.1145/3173574.3173995]
32. Neyazi TA, Ng SWT, Hobbs M, Yue A. Understanding user interactions and perceptions of AI risk in Singapore. Big Data Soc 2023 Jul;10(2). [doi: 10.1177/20539517231213823]
33. Fernández-Ardèvol M, Ivan L. Why age is not that important? An ageing perspective on computer anxiety. In: Zhou J, Salvendy G, editors. Human Aspects of IT for the Aged Population Design for Aging: Springer International Publishing; 2015:189-200. [doi: 10.1007/978-3-319-20892-3_19]
34. Li J, Huang JS. Dimensions of artificial intelligence anxiety based on the integrated fear acquisition theory. Technol Soc 2020 Nov;63:101410. [doi: 10.1016/j.techsoc.2020.101410]
35. Finkelstein R, Wu Y, Brennan-Ing M. Older adults' experiences with using information and communication technology and tech support services in New York City: findings and recommendations for post-pandemic digital pedagogy for older adults. Front Psychol 2023;14:1129512. [doi: 10.3389/fpsyg.2023.1129512] [Medline: 37138998]
36. Esmaeilzadeh P, Mirzaei T, Dharanikota S. Patients' perceptions toward human-artificial intelligence interaction in health care: experimental study. J Med Internet Res 2021 Nov 25;23(11):e25856. [doi: 10.2196/25856] [Medline: 34842535]

XSL•FO
RenderX

37. Arnaert A, Sumbly P, da Costa D, Liu Y, Debe Z, Charbonneau S. Acceptance of the Apple Watch Series 6 for telemonitoring of older adults with chronic obstructive pulmonary disease: qualitative descriptive study part 1. JMIR Aging 2023 Dec 26;6:e41549. [doi: 10.2196/41549] [Medline: 38147371]
38. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: an integrative systematic review. BMC Public Health 2024 Feb 16;24(1):506. [doi: 10.1186/s12889-024-18036-5] [Medline: 38365698]
39. Tsai TH, Lin WY, Chang YS, Chang PC, Lee MY. Technology anxiety and resistance to change behavioral study of a wearable cardiac warming system using an extended TAM for older adults. PLoS One 2020;15(1):e0227270. [doi: 10.1371/journal.pone.0227270] [Medline: 31929560]
40. Marakhimov A, Joo J. Consumer adaptation and infusion of wearable devices for healthcare. Comput Human Behav 2017 Nov;76:135-148. [doi: 10.1016/j.chb.2017.07.016]
41. Cilliers L. Wearable devices in healthcare: privacy and information security issues. Health Inf Manag J 2020 May;49(2-3):150-156. [doi: 10.1177/1833358319851684]
42. Akrivopoulou CM. Protecting the Genetic Self From Biometric Threats: Autonomy, Identity, and Genetic Privacy: IGI Global; 2015. [doi: 10.4018/978-1-4666-8153-8]
43. Kupfer J. Privacy, autonomy, and self-concept. Am Philos Q 1987;24:81-89 [FREE Full text]
44. Altman I. The Environment and Social Behavior: Privacy, Personal Space, Territory, Crowding: Brooks/Cole Publishing Company; 1975.
45. Townsend D, Knoefel F, Goubran R. Privacy versus autonomy: a tradeoff model for smart home monitoring technologies. Presented at: 2011 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society; Aug 30 to Sep 3, 2011; Boston, MA. [doi: 10.1109/IEMBS.2011.6091176]
46. Schönau A, Dasgupta I, Brown T, Versalovic E, Klein E, Goering S. Mapping the dimensions of agency. AJOB Neurosci 2021;12(2-3):172-186. [doi: 10.1080/21507740.2021.1896599] [Medline: 33764258]
47. Mai R, Hoffmann S. How to combat the unhealthy = tasty intuition: the influencing role of health consciousness. J Public Policy Mark 2015 Apr;34(1):63-83. [doi: 10.1509/jppm.14.006]
48. Gould SJ. Health consciousness and health behavior: the application of a new health consciousness scale. Am J Prev Med 1990;6(4):228-237. [doi: 10.1016/S0749-3797(18)31009-2] [Medline: 2223170]
49. Chen MF, Lin NP. Incorporation of health consciousness into the technology readiness and acceptance model to predict app download and usage intentions. Internet Res 2018 Apr 4;28(2):351-373. [doi: 10.1108/IntR-03-2017-0099]
50. Pu B, Zhang L, Tang Z, Qiu Y. The relationship between health consciousness and home-based exercise in China during the COVID-19 pandemic. Int J Environ Res Public Health 2020 Aug 6;17(16):5693. [doi: 10.3390/ijerph17165693] [Medline: 32781751]
51. Zhang Q, Guo X, Vogel D. Addressing elderly loneliness with ICT use: the role of ICT self-efficacy and health consciousness. Psychol Health Med 2022 May 28;27(5):1063-1071. [doi: 10.1080/13548506.2020.1847302]
52. Jung EH, Kang H. Self-determination in wearable fitness technology: the moderating effect of age. Int J Human Comput Interact 2022 Sep 14;38(15):1399-1409. [doi: 10.1080/10447318.2021.2002048]
53. Sundar SS, Bellur S, Jia H. Motivational technologies: a theoretical framework for designing preventive health applications. In: Bang M, Ragnemalm EL, editors. Persuasive Technology Design for Health and Safety: Springer; 2012:112-122. [doi: 10.1007/978-3-642-31037-9_10]
54. Lee C, Coughlin JF. PERSPECTIVE: Older adults' adoption of technology: an integrated approach to identifying determinants and barriers. J Prod Innov Manag 2015 Sep;32(5):747-759. [doi: 10.1111/jpim.12176]
55. Rupp MA, Michaelis JR, McConnell DS, Smither JA. The impact of technological trust and self-determined motivation on intentions to use wearable fitness technology. Proc Hum Factors Ergon Soc Annu Meet 2016 Sep;60(1):1434-1438. [doi: 10.1177/1541931213601329]
56. Hsieh PJ. The impact of motivations, health beliefs, and basic human needs on mobile self-management: an extension of the self-determination theory perspective. Behav Inf Technol 2023 Jun 11;42(8):1045-1063. [doi: 10.1080/0144929X.2022.2059007]
57. Lazaro MJS, Lim J, Kim SH, Yun MH. Wearable Technologies: Acceptance Model for Smartwatch Adoption Among Older Adults: Springer International Publishing; 2020:303-315. [doi: 10.1007/978-3-030-50252-2_23]
58. Ozkan-Yildirim S, Pancar T. Smart wearable technology for health tracking: what are the factors that affect their use? In: Marques G, Bhoi AK, Albuquerque V, KS H, editors. IoT in Healthcare and Ambient Assisted Living: Springer; 2021:165-199.
59. Rani N, Chu SL. Wearables can help me learn: a survey of user perception of wearable technologies for learning in everyday life. Educ Inf Technol 2022 Apr;27(3):3381-3401. [doi: 10.1007/s10639-021-10726-6]
60. Xu Z, Li Y, Hao L. An empirical examination of UTAUT model and social network analysis. Library Hi Tech 2022 Feb 14;40(1):18-32. [doi: 10.1108/LHT-11-2018-0175]
61. Mercer K, Giangregorio L, Schneider E, Chilana P, Li M, Grindrod K. Acceptance of commercially available wearable activity trackers among adults aged over 50 and with chronic illness: a mixed-methods evaluation. JMIR Mhealth Uhealth 2016 Jan 27;4(1):e7. [doi: 10.2196/mhealth.4225] [Medline: 26818775]

XSL•FO
RenderX

62. Kang HS, Exworthy M. Wearing the future-wearables to empower users to take greater responsibility for their health and care: scoping review. JMIR Mhealth Uhealth 2022 Jul 13;10(7):e35684. [doi: 10.2196/35684] [Medline: 35830222]
63. Sun S, Jiang L, Zhou Y. Associations between perceived usefulness and willingness to use smart healthcare devices among Chinese older adults: the multiple mediating effect of technology interactivity and technology anxiety. Digit Health 2024;10. [doi: 10.1177/20552076241254194] [Medline: 38812850]
64. Meuter ML, Ostrom AL, Bitner MJ, Roundtree R. The influence of technology anxiety on consumer use and experiences with self-service technologies. J Bus Res 2003 Nov;56(11):899-906. [doi: 10.1016/S0148-2963(01)00276-4]
65. Li H, Wu J, Gao Y, Shi Y. Examining individuals' adoption of healthcare wearable devices: an empirical study from privacy calculus perspective. Int J Med Inform 2016 Apr;88:8-17. [doi: 10.1016/j.ijmedinf.2015.12.010]
66. Shin DH, Biocca F. Health experience model of personal informatics: the case of a quantified self. Comput Human Behav 2017 Apr;69:62-74. [doi: 10.1016/j.chb.2016.12.019]
67. Ryan RM, Rigby CS, Przybylski A. The motivational pull of video games: a self-determination theory approach. Motiv Emot 2006 Dec 12;30(4):344-360. [doi: 10.1007/s11031-006-9051-8]
68. Kang H, Lou C. AI agency vs. human agency: understanding human–AI interactions on TikTok and their implications for user engagement. J Comput-Mediat Commun 2022 Aug 18;27(5). [doi: 10.1093/jcmc/zmac014]
69. Jung EH, Sundar SS. Senior citizens on Facebook: How do they interact and why? Comput Human Behav 2016 Aug;61:27-35. [doi: 10.1016/j.chb.2016.02.080]
70. Wilson G, Gates JR, Vijaykumar S, Morgan DJ. Understanding older adults' use of social technology and the factors influencing use. Ageing Soc 2023 Jan;43(1):222-245. [doi: 10.1017/S0144686X21000490]
71. Zhao W, Kelly RM, Rogerson MJ, Waycott J. Older adults using technology for meaningful activities during COVID-19: an analysis through the lens of self-determination theory. Presented at: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23); Apr 23-28, 2023; Hamburg, Germany p. 1-17. [doi: 10.1145/3544548.3580839]
72. Yang CC, Liu C, Wang YS. The acceptance and use of smartphones among older adults: differences in UTAUT determinants before and after training. Library Hi Tech 2023 Oct 31;41(5):1357-1375. [doi: 10.1108/LHT-12-2021-0432]
73. Chen J, Wang T, Fang Z, Wang H. Research on elderly users' intentions to accept wearable devices based on the improved UTAUT model. Front Public Health 2022;10:1035398. [doi: 10.3389/fpubh.2022.1035398] [Medline: 36699866]
74. Population in brief. National Population and Talent Division. 2024 Sep 24. URL: https://www.population.gov.sg/media-centre/publications/population-in-brief [accessed 2025-02-25]
75. Kagawa Singer M. Applying the concept of culture to reduce health disparities through health behavior research. Prev Med 2012 Nov;55(5):356-361. [doi: 10.1016/j.ypmed.2012.02.011]
76. Lee SG, Trimi S, Kim C. The impact of cultural differences on technology adoption. J World Bus 2013 Jan;48(1):20-29. [doi: 10.1016/j.jwb.2012.06.003]
77. Li Y, Rho EHR, Kobsa A. Cultural differences in the effects of contextual factors and privacy concerns on users' privacy decision on social networking sites. Behav Inf Technol 2022 Feb 17;41(3):655-677. [doi: 10.1080/0144929X.2020.1831608]

Abbreviations

AI: artificial intelligence
CFI: comparative fit index
H: hypothesis
RFID: radio frequency identification
RMSEA: root mean square error of approximation
RQ: research question
SDT: self-determination theory
SEM: structural equation model
SRMR: standardized root mean square residual
TAM: technology acceptance model
TLI: Tucker-Lewis index
UTAUT: Unified Theory of Acceptance and Use of Technology

XSL•FO
RenderX

Edited by Q Fan; submitted 19.11.24; peer-reviewed by J Farias, MA Azmi Murad; revised version received 28.01.25; accepted 30.01.25; published 19.03.25.

Please cite as:
Kang H, Yang T, Banu N, Ng SWT, Lee JK
Exploring Smart Health Wearable Adoption Among Singaporean Older Adults Based on Self-Determination Theory: Web-Based Survey Study
JMIR Aging 2025;8:e69008
URL: https://aging.jmir.org/2025/1/e69008
doi:10.2196/69008

© Hyunjin Kang, Tingting Yang, Nazira Banu, Sheryl Wei Ting Ng, Jeong Kyu Lee. Originally published in JMIR Aging (https://aging.jmir.org), 19.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Smartphone-Based Timed Up and Go Test Self-Assessment for Older Adults: Validity and Reliability Study

Melissa Johanna Böttinger[1,2], MSc; Sabato Mellone[3], PhD; Jochen Klenk[4,5], Prof Dr; Carl-Philipp Jansen[1], PhD; Marios Stefanakis[6], MSc; Elena Litz[1,2], MSc; Anastasia Bredenbrock[2], BSc; Jan-Philipp Fischer[7], MA; Jürgen M Bauer[1,2], MD; Clemens Becker[1], MD; Katharina Gordt-Oesterwind[1,2,8], PhD

[1]Geriatric Center, Medical Faculty Heidelberg, Heidelberg University, Bergheimer Str. 20, Heidelberg, Germany

[2]Network Aging Research, Heidelberg University, Heidelberg, Germany

[3]Department of Electrical, Electronic and Information Engineering, University of Bologna, Bologna, Italy

[4]Institute of Epidemiology and Medical Biometry, Ulm University, Ulm, Germany

[5]Study Center Stuttgart, IB University of Health and Social Sciences, Stuttgart, Germany

[6]Department of Clinical Gerontology and Geriatric Rehabilitation, Robert Bosch Hospital, Stuttgart, Germany

[7]TSG ResearchLab gGmbH, Zuzenhausen, Germany

[8]Institute of Sports and Sports Sciences, Heidelberg University, Heidelberg, Germany

Corresponding Author:
Melissa Johanna Böttinger, MSc
Geriatric Center, Medical Faculty Heidelberg, Heidelberg University, Bergheimer Str. 20, Heidelberg, Germany

Abstract

Background: The Timed Up and Go test (TUG) is recommended as an evidence-based tool for measuring physical capacity. Instrumented TUG (iTUG) approaches expand classical supervised clinical applications offering the potential of self-assessment for older adults.

Objective: This study aimed to evaluate the concurrent validity and test-retest reliability of a smartphone-based TUG self-assessment "up&go app."

Methods: A total of 52 community-dwelling older adults (>67 years old) were recruited. A validated and medically certified system attached with a belt at the lower back was used as a reference system to validate the "up&go app" algorithm. The participants repeated the TUG 5 times wearing, a smartphone with the "up&go app" in their front trouser pocket and an inertial sensor to test the concurrent validity. A subsample of 37 participants repeated the "up&go app" measurement 2 weeks later to examine the test-retest reliability.

Results: The correlation between the "up&go app" and the reference measurement was r=0.99 for the total test duration and r=0.97 for the 5 single repetitions. Agreement between the 5 repetitions was intraclass correlation coefficient (ICC)=0.9 (0.84 - 0.94). Leaving out the first repetition, the agreement was ICC=0.95 (0.92 - 0.97). Test-retest agreement had an ICC=0.79 (0.53 - 0.9).

Conclusions: The duration of 5 repetitions of the TUG test, measured with the pocket-worn "up&go app," was very consistent with the results of a lower-back sensor system, indicating excellent concurrent validity. Participants walked slower in the first round than in the other 4 repetitions within a test run. Test-retest reliability was also excellent. The "up&go app" provides a useful smartphone-based approach to measure 5 repetitions of the TUG. The app could be used by older adults as a self-screening and monitoring tool of physical capacity at home and thereby help to early identify functional limitations and take interventions when necessary.

(JMIR Aging 2025;8:e67322) doi:10.2196/67322

KEYWORDS

timed up and go test; self-assessment; instrumented assessment; technology-based assess-ment; physical capacity; mobility; aged; mobile applications; smartphone; diagnostic self evaluation

Introduction

The baby boomer generation embraces smartphone technology [1]. This particular cohort is willing and capable of using mobile health applications especially when they are designed in a user-centered way and increase the likelihood of a self-determined future [2-4]. Good examples are the use of fitness trackers, heart rate monitoring, and glucose monitoring [5].

XSL•FO
RenderX

The World Health Organization (WHO) and United Nations (UN) member states have declared the current decade as the UN "Decade of Healthy Ageing" (2021 - 2030) to improve the lives of older people, their families, and the communities in which they live [6-8]. Early identification of health risks and functional limitations in older adults is crucial to enable timely preventive measures and promote the maintenance of independence [9]. A recent systematic review was conducted to guide WHO recommendations and to establish a "standard" for evidence-based assessment of physical capacity [10]. It recommends the Timed Up and Go Test (TUG), in addition to the 30-Second Chair Rise Test, due to the high quality evidence for its sufficiently valid and reliable measurement of physical capacity in community-dwelling older adults [10,11]. The standard TUG captures the time recorded by a trained test administrator using a stopwatch. The test person rises from a chair, walks 3 meters at their usual pace, turns around, walks back, and sits down again. Longer test durations indicate limitations in physical capacity, particularly in strength and balance [10-12] and predict hospitalization and associated functional decline in older adults [13]. The need for and lack of a digital version of the TUG has been addressed during the meetings of the WHO Locomotor Capacity Working Group. This is even more relevant for persons living remotely or in low and middle-income countries where smartphone technology is often available but health care professionals are often lacking [2].

Wearable technologies offer opportunities to lower the barrier for integrating functional tests like the TUG into the assessment and clinical management of older people [14,15]. Such an instrumented Timed Up and Go Test (iTUG) not only enables standardized and digitized measurement in supervised settings [16-18] but also opens an option for low-cost self-screening and monitoring in people's everyday lives [19,20]. In clinical studies, camera-based systems, inertial sensors or smartphones that are attached to the back (eg, with a belt) are most commonly used to perform the iTUG [21]. For self-assessment purposes, however, an approach using common technologies and a convenient placement on the body would be more feasible, without the need for additional materials or training. Therefore, the "up&go" iTUG was developed as a self-test for older adults. The prototype of the "up&go" app emerged from several EU projects [17-19,22,23] and a cocreation process with older adults [24]. It incorporates an innovative approach: placing a smartphone in the user's front trouser pocket during 5 repetitions of the test. An unsupervised smartphone-based iTUG could fill a gap in the landscape of mobility assessment methods that is still dominated by patient-reported outcome measures and supervised assessments conducted in laboratory or clinical settings [25]. Digital self-assessments could support older adults to take up an active role in risk screening or monitoring, and to early identify changes in their mobility. If necessary, this would allow them to initiate further medical diagnostics, and to take primary or secondary preventive measures in a timely manner. Thus, the motivation for this study is to examine if the "up&go" self-assessment app fulfils the criteria of validity and reliability to be considered trustworthy and to serve as a credible tool for implementation in future iTUG studies and in the self-management of older people.

Hence, this study aims to examine the concurrent validity and test-retest reliability of the "up&go" smartphone-based TUG self-assessment for older adults. The measures considered are the total TUG test duration and the duration of the 5 individual repetitions. This study corresponds to an analytical validation according to the V3 framework [26].

The objectives of this study are to analyze the concurrent validity of the algorithm used in the "up&go app" against a validated and medically certified sensor-based system and to determine the test-retest reliability of the "up&go app" algorithm in the home environment of older adults.

Methods

Participant Recruitment

The findings are reported following the guidelines for reporting reliability and agreement studies [27]. Participant were recruited from the SMART-AGE randomized controlled trial conducted at Heidelberg University, Germany [28]. Participants were included in the SMART-AGE study if they were 67 years or older, lived in the community, had basic knowledge of using PCs or tablets and of the German language. Exclusion criteria were severe medical conditions (ie, heart failure with shortness of breath at rest, cardiac arrhythmia with dizziness, Parkinson disease with use of a walker or wheelchair, cancer with chemotherapy or radiotherapy, chronic lung disease with oxygen therapy, and planned major medical procedure with inpatient hospitalization within the next 3 months), severe visual or hearing impairment, and severe cognitive impairment.

Data for the sample description were extracted from the SMART-AGE initial assessment data: age, sex, BMI, citizenship, living situation, employment, WHO Quality of Life Scale score [29], Satisfaction With Life Scale score [30], fall history, Groll Functional Comorbidity Index [31], Trail Making Test A and B score [32], Six-item Cognitive Impairment Test score [33], Short Falls Efficacy Scale International score [34], 4-Meter Walk Test score [35], 30-second Chair Rise Test score [36], and stopwatch-measured Timed-Up and Go Test score [11].

Ethical Considerations

The SMART-AGE study and the amended protocol for this sub study were approved by Heidelberg University Medical Faculty's Ethical Committee (S-672/2022). All participants provided informed signed consent before participating.

Data Collection and Processing

The data collection was performed during a home visit 3 to 4 months after the SMART-AGE study's baseline assessment. A total of 3 pretrained assessors with a professional background in psychology (AB and EL) or physiotherapy (MJB) conducted the home visits following a standardized manual.

"Up&go App": Smartphone-Based Timed Up and Go Test

Previous versions of the "up&go app" algorithm have been developed and validated within the 2 European Commission funded projects, Farseeing [17,22] and PreventIT [18,19,23].

XSL•FO
RenderX

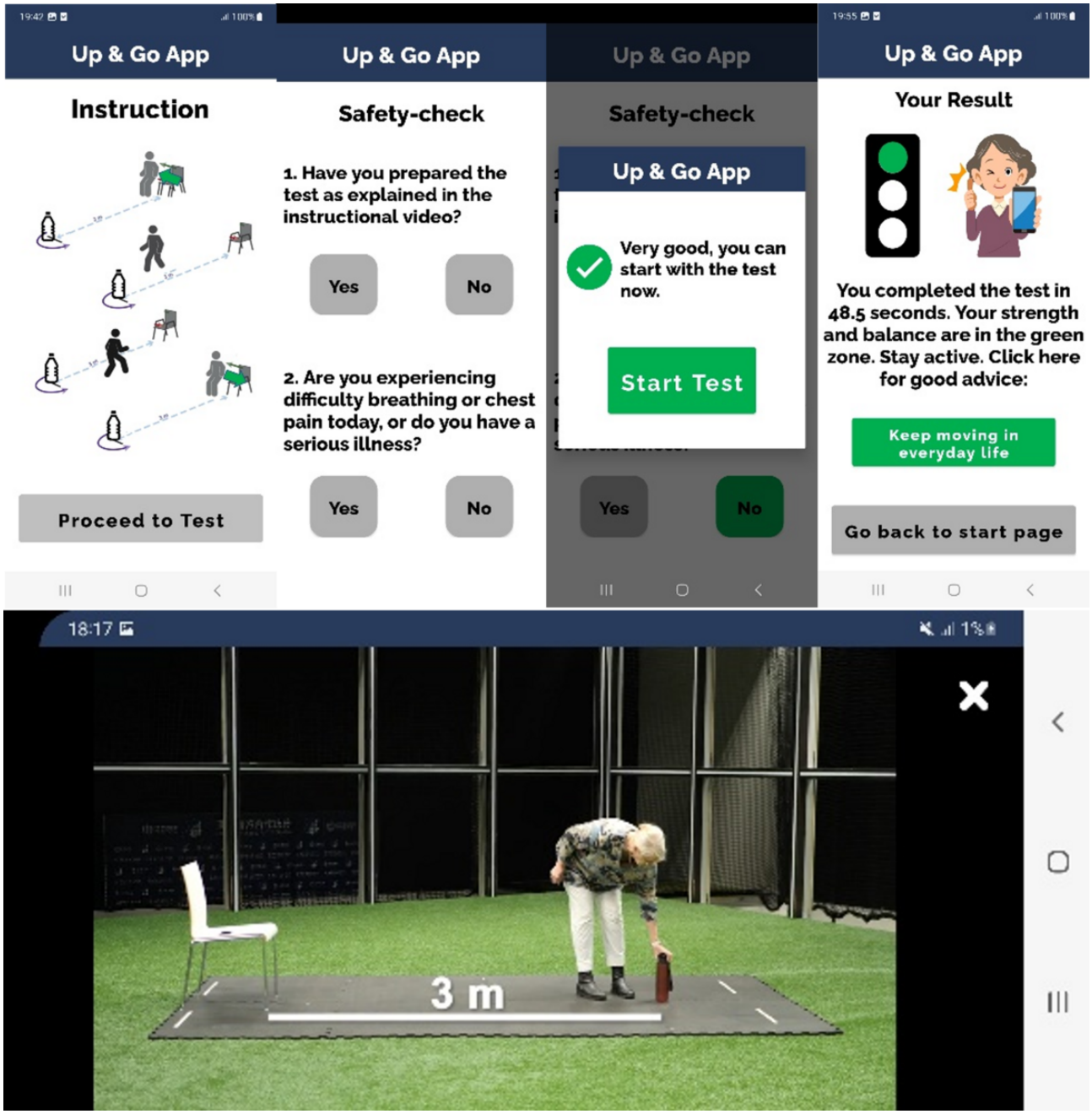
The “up&go app” investigated in this study [37] is a further development of the previous versions. It is based on the results of a study examining the usability of the previous Norwegian version of the PreventIT project [19] and a cocreation study with older adults [24].

In order to run the TUG test within the “up&go app,” the user is invited to watch and read the test instructions, prepare the test setting (chair, 3 m distance, marker), answer 2 safety questions, press the “start test” button, place the smartphone in one of the front trouser pockets and sit down to begin with the first round of the test at usual gait speed (Figure 1). The app includes an algorithm processing real-time accelerometer data to detect the body movements to guide the user through 5 consecutive repetitions of the TUG via audio prompts (eg, “Please put the phone in your front trouser pocket and sit down”). After each run there is a pause of about 3 seconds in which the algorithm detects the position change, and the new audio announcement is played (eg, “This was the third run. Please wait for the next start signal”). The embedded sensors in the smartphone collect data during the test.

After the test, the algorithm automatically postprocesses accelerometer data and computes the total time needed to complete 5 repetitions, excluding pauses between runs. The total time is shown on the smartphone screen in seconds together with a traffic light (Figure 1, green=<60 s, yellow 60 - 90 s, red>90 s) and an according recommendation for action (eg, green light: recommendation to stay active, yellow light: recommendation for supervised training, and red light: recommendation to seek medical advice). The color coding thresholds are based on the cut-off value of 12 seconds for 1 repetition, respectively 60 seconds for 5 repetitions, to indicate normal versus below normal mobility in community-dwelling older adults [13,38].

XSL•FO
RenderX

Figure 1. Screenshots from the “up&go app".



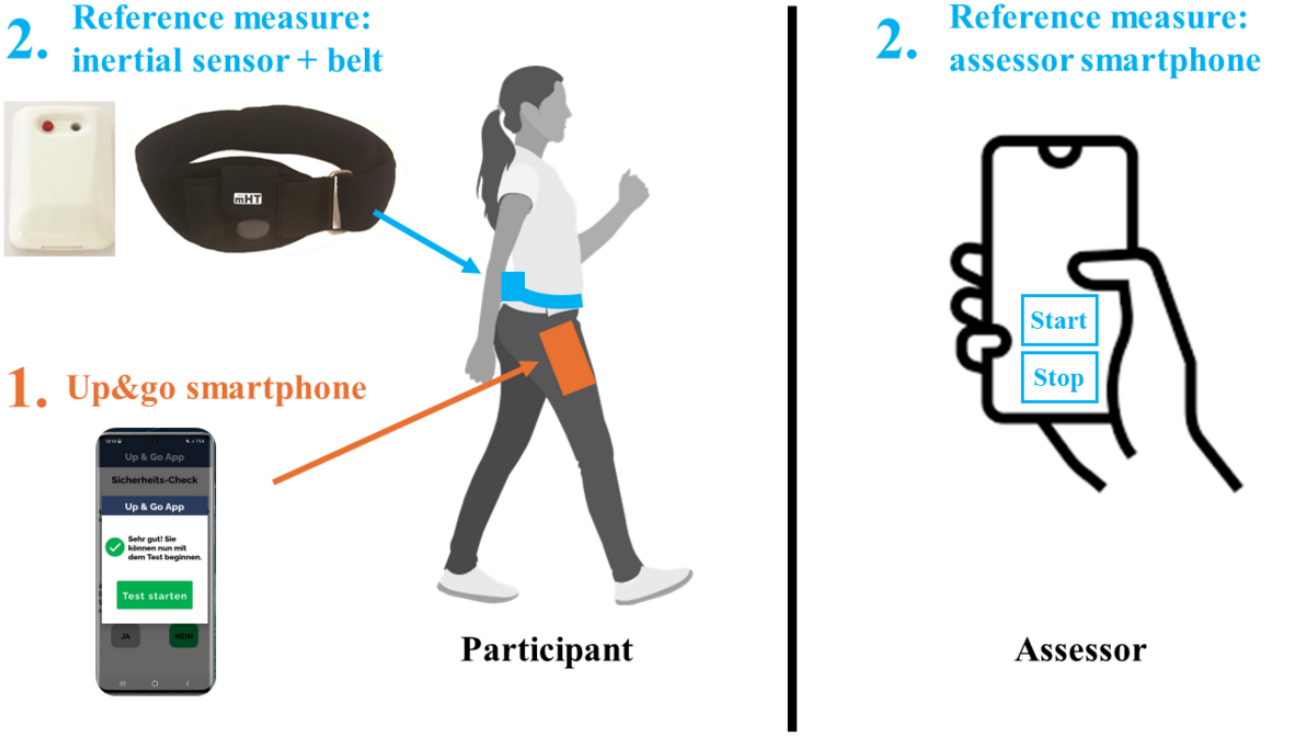
Concurrent Validity

In this study, the “mTUG” system (“mTUG” medical device, mHealth Technologies) was used as a reference measure to validate the “up&go app” algorithm. “mTUG” is a sensor-based system that has been validated in a population of older adults [17] and is certified as a medical device. The system consists of an inertial sensor (mHealth Technologies triaxial accelerometer range±2 g, triaxial gyroscope range±250°/sec, sampling rate: 100 Hz) connected by Bluetooth to an assessor smartphone (Samsung Galaxy S10e, Android 5.0.1) including a customized app (mHealth Technologies) allowing the assessor to manually start and stop the sensor measurement.

The TUG was repeated 5 times with participants wearing the 2 measurement systems simultaneously (Figure 2), smartphone with “up&go app” in their right front trouser pocket and an inertial sensor (mHealth Technologies) attached to an elastic belt worn on the lower back. A previous study with hospitalized hip fracture patients found that the performance of three TUG trials is recommended to achieve performance stability [39]. In this study, 5 repetitions were performed to take into account more variability in the performance of older adults conducting the test at home.

XSL•FO
RenderX

Figure 2. : Simultaneous measurement with (1) the "up&go app" and (2) the reference measure.



Participants were familiar with the TUG procedure as the conventional stopwatch TUG [11] had been performed beforehand as part of the baseline motor assessments within the SMART-AGE study. As the feasibility of the "up&go app" self-assessment was not part of this study, test preparation, viewing of instructional videos by participants were omitted for data collection. The test procedure was explained to the study partner (5 repetitions, usual walking pace, and follow audio instructions). The assessor prepared the TUG test setting, mounted the sensor in the belt on the participant's lower back and started the test on the "mTUG" assessor smartphone. Then, the assessor started the test on the "up&go" smartphone and handed it over to the participant. The participant then placed it in the right front trouser pocket and performed the 5 repetitions of the TUG test, following the audio instructions of the "up&go app."

The "up&go app" does not store any data on the user's smartphone to guarantee data confidentiality. We therefore developed an adapted version for this study, allowing the app to store all raw data collected from the accelerometer and gyroscope embedded in a Samsung Galaxy S21 smartphone (Android 5.0.1, triaxial accelerometer range±8 g, triaxial gyroscope range±1000°/sec, sampling rate 500 Hz) on the smartphone memory. This allowed postprocessing of the raw data and analysis of not only the total test duration, but also the duration of each of the 5 repetitions (rep1, rep2, rep3, rep4, and rep5).

Within the "mTUG" system, raw data was automatically stored on the assessor smartphone's memory and processed by a validated algorithm [17].

Test-Retest Reliability

To investigate the test-retest reliability of the "up&go app" algorithm, we repeated the "up&go" measurement at an additional home visit within 2 weeks after the first measurement, following the same procedure as described above.

Participants were asked to participate in the retest if they lived <20 km from the study center to reduce the burden of assessor time.

Statistical Analysis

The sample size for the validation part of the study was calculated using R (version 2023.09.1, R Foundation for Statistical Computing, package "pwr"). A strong correlation of ≥0.75 was assumed. With a power of 95% and a statistical significance level of α=.05, ≥45 participants are required. The sample size for the test-retest part of the study was calculated using R (version 2023.09.1, package "ICC.sample.size"). An intraclass correlation coefficient (ICC) of ≥0.75 was assumed. With power=95% and α=.05, ≥37 participants were required for 2 measurements (first measurement vs retest). Assuming a 15% dropout rate, 52 participants were recruited for the study. From this sample, we consecutively recruited participants for the retest measurement until a subsample size of at least 37 participants was reached to examine the test-retest reliability of the "up&go app" algorithm. The distribution of the data was checked before analyses (Shapiro-Wilk test). All data fulfilled the normal distribution assumption. For all statistical tests, α=.05 was used as the threshold for statistical significance. Pearson correlation coefficients were calculated to examine the association between the systems ("up&go app" and reference measure) regarding the total duration needed to complete the test as well as the duration of the 5 single repetitions. Correlation coefficients of 0 - 0.19 are interpreted as very weak, 0.2 - 0.39

XSL•FO
RenderX

as weak, 0.4 - 0.59 as moderate, 0.6 - 0.79 as strong, and 0.8 - 1 as very strong [40]. For graphical description of the agreement between the 2 systems ("up&go app" and reference measure), Bland-Altman plots were used [41], including the lower and upper limits of agreement (LLoA and ULoA). From the data collected during the first measurement ("up&go app" and reference measure), ICC 3,1 were calculated to analyze the agreement between the 5 repetitions. Test-retest reliability of the "up&go app" was evaluated using ICC [1,2,42]. An ICC value<0.4 indicates poor reliability, an ICC value between ≥0.4 and <0.75 fair to good reliability and an ICC value ≥0.75 indicates excellent reliability [43]. It was assumed that the association between "up&go app" and the reference measure is very strong (r>0.8) and the agreement between "up&go" and the reference measure is good (Bland-Altman plot). Furthermore, we assumed that the test-retest reliability of the "up&go" test meets excellent levels (ICC≥0.75). Statistical analyses were computed using statistical software R (version 2023.09.1) and MATLAB (version R2022b, MathWorks).

Results

Participant Data

A total of 52 community-dwelling older adults aged between 66 and 88 years (mean 73.6, SD 5.4) were measured, with a similar distribution of men (25/52, 48%) and women (27/52, 52%) and a mean BMI of 27.1 kg/m^2 (SD 4.1). Of those who provided the data, all were retired, almost all were German citizens (49/52, 98%) and one third (16/52, 33%) lived alone. Participants indicated an average of 2.3 (SD 1.7) comorbidities according to the Groll Index. Of these, degenerative disc disease (17/52), visual impairment (13/52), depression (11/52), and osteoarthritis (11/52) were reported most frequently. Cognition was normal (<7 errors on the Six-item Cognitive Impairment Test) and executive functions (Trail Making Test A and B) were in accordance with normative data of this age group [44]. The average duration of the TUG measured with the stopwatch in the SMART-AGE initial assessment was 9.7 (SD 1.7) seconds. None of the study participants used a walking aid during the TUG. Table 1 provides a detailed description of the sample.

XSL•FO
RenderX

Table . Characteristics of the study sample.

Data from the SMART-AGE initial assessment[a]		Values	Available sample
Age (years), mean (SD)		73.6 (5.4)	52
Female sex, n (%)		27 (52)	52
BMI (kg/m^2), mean (SD)		27.1 (4.1)	52
German citizenship, n (%)		49 (98)	50
Living alone, n (%)		16 (33)	49
Retired, n (%)		50 (100)	50
Stopwatch-measured TUG (s), mean (SD)		9.7 (1.7)	52
4-meter walk test, mean (SD)			
	Duration (s)	4.2 (0.9)	52
	Gait speed (m/s)	1.0 (0.2)	52
30-Second chair rise test, mean (SD)			
	Number of repetitions	12.3 (3.3)	52
Short Falls Efficacy Scale International, mean (SD)			
	Score	8.2 (2.2)	49
Groll Functional Comorbidity Index (score), mean (SD)		2.3 (1.7)	48
Fall history ("Yes, I fell in the last 12 months")[b], n (%)		19 (37)	51
Trail making test, mean (SD)			
	Test duration A (s),	45.2 (16.6)	52
	Test duration B (s)	99 (50.3)	52
	Ratio (B/A)	2.3 (1.1)	52
Six-item cognitive impairment test [c] , mean (SD)			
	Score	1.0 (1.5)	52
Satisfaction with life scale, mean (SD)			
	Score	5.0 (1.2)	49
WHO Quality of Life Scale, mean (SD)			
	Physical domain (score)	77.3 (13.1)	49
	Psychological domain (score)	72.4 (10.3)	49
	Social relationships domain (score)	65.3 (13.6)	49
	Environment domain (score)	79.8 (10.1)	49

[a]The initial SMART-AGE assessment took place 3-4 months before the data collection for this study.

[b]Fall history data was obtained one month after the initial SMART-AGE assessment.

[c]Participants with >7 error points were excluded from the SMART-AGE study due to a suspected dementia disorder.

Concurrent Validity

In total, 35 (67.3%) participants completed the test successfully on the first attempt. Due to technical problems (eg, smartphone shifted in trouser pocket), 12 (23.1%) participants had to repeat the test a second time, 4 (7.7%) participants a third time, and 1 (1.9%) participant a fourth time. After an unsuccessful attempt, participants always took a break of at least 5 minutes. As the app only saves data from complete test runs, failed attempts were not included.

On average, participants needed 52.4 (SD 7.8) seconds to complete 5 repetitions of the TUG, as measured by the "up&go app". Table 2 shows the results for the total duration of the test and the duration of the 5 repetitions as measured by the "up&go app" and the reference measure.

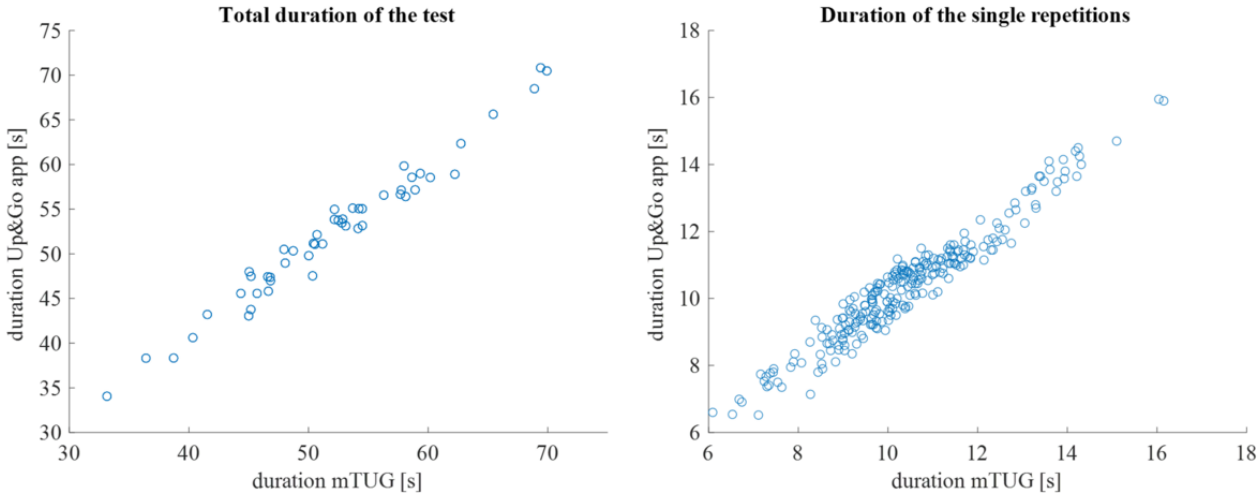
Data from all 52 participants were available for the concurrent validity analysis ("up&go app" vs reference measure). The Pearson correlation coefficient (r) for the total duration was 0.99 (95% CI 0.98 - 0.99) and that for the durations of the 5 single repetitions was 0.97 (95% CI 0.96-.97; Figure 3).

XSL•FO
RenderX

Table . Total test duration and duration of the 5 repetitions measured by the "up&go app" and reference measure.

Duration	"up&go app", mean (SD)	Reference measure, mean (SD)
Total duration of the test (s)	52.38 (7.8)	52.31 (8.11)
First repetition (s)	10.99 (2.1)	11.11 (2.17)
Second repetition (s)	10.33 (1.54)	10.39 (1.61)
Third repetition (s)	10.25 (1.57)	10.34 (1.64)
Fourth repetition (s)	10.21 (1.52)	10.2 (1.52)
Fifth repetition (s)	10.11 (1.46)	10.26 (1.6)

Figure 3. Correlation analysis between "up&go app" and reference measure: total duration and duration of the single repetitions. TUG: Timed Up and Go Test.

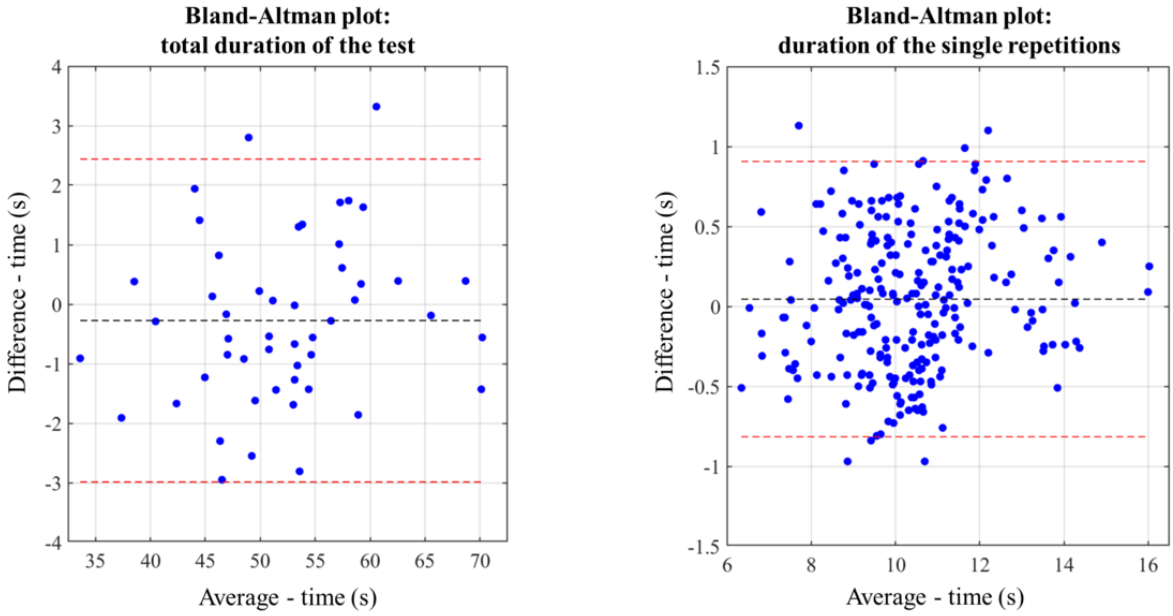


The agreement between the total duration measured by the "up&go app" and the reference measure is shown in Figure 4 (Bland-Altman plots). According to Bland-Altman analysis, the mean difference was −0.27 for the total duration of the test and −0.05 for the duration of the single repetitions. The difference in the total duration of the test was between −2.99 seconds (LLoA) and −2.44 seconds (ULoA) and the difference in the duration of a single repetition was within −0.82 seconds (LLoA) and −0.91 seconds (ULoA) with respect to the estimate of the reference measure.

Agreement between all "up&go app" repetitions was ICC(3,1)=0.9 (95% CI 0.84 - 0.94) and ICC(3,1)=0.87 (95% CI 0.80 - 0.93) between all repetitions measured by the reference measure. Leaving out the first repetition, the agreement of repetitions 2 - 5 was ICC(3,1)=0.95 (95% CI 0.92 - 0.97) for the "up&go app" and ICC(3,1)=0.93 (95% CI 0.89 - 0.96) for the reference measure.

XSL•FO
RenderX

Figure 4. Statistical agreement between the "up&go app" and the reference measure—Bland-Altman plot.



Test-Retest Reliability

Of the 52 participants, 37 were measured twice, and 1 participant was excluded due to noncompliance with the retest instruction to walk at habitual gait speed. Retests were performed on average 6 (SD 2.6) days after the first measurement. At the "up&go app" retest measurement, participants needed on average 48.6 (SD 8.5) seconds to complete the 5 repetitions. The ICC between the total duration during the first home visit and during the retest was ICC(2,1)=0.79 (95% CI 0.53 - 0.9).

Discussion

Principal Findings

To enable self-assessment for older people in their own homes, the "up&go app" was developed as a pocket-worn approach with 5 repetitions of the TUG. This study aimed to investigate the concurrent validity and test-retest reliability of the "up&go app," which records and processes data from sensors embedded in a smartphone that is placed in the front pocket of the trousers. The results show excellent agreement between the "up&go" data and the comparator system, a previously validated medically certified lower-back inertial sensor system. Overall, the correlation of the 5 single repetitions was very strong. The results indicate that participants did not seem to walk at the same speed in each of the 5 test rounds, that is, they tended to walk slower during the first round. This trend probably shows a learning or accustoming effect. Similar observations have been made in previous studies [39,45] where participants started rather slowly and more carefully during the first attempt of motor performance tests to follow the instructions precisely and to avoid errors [45]. This first round may therefore be comparable to a "trial run," which is often performed in assessments before the actual measurement. Performing several repetitions of the test provides the opportunity to observe any changes in walking speed during the test and, if necessary, to interpret it clinically regarding physical capacity. For example, in more frail populations, a significant reduction in walking speed during the test could indicate fatigue. Furthermore, this approach might achieve a better approximation toward the "real" usual walking speed compared with measuring only 1 or 2 rounds. Using the app to monitor differences in the total test duration across several measurements would be a clinically relevant use case. But what would be considered a "relevant change"? Minimal important differences (MID) of >2 seconds for 1 TUG repetition were reported in studies observing populations with age-associated disorders and TUG baseline values around 20 seconds [46,47]. In another study observing older adults with hip osteoarthritis and a TUG baseline duration of 7 seconds, a sample that is comparable with this study, the MID is significantly lower with, 0.8 seconds [48]. This suggests that ceiling effects make it more difficult to screen for relevant changes in more robust individuals. The "up&go app," however, should not only aim to screen for subjectively noticeable changes, but also for changes falling below the abovementioned MID thresholds. Looking at the Bland-Altman plots, a clinically relevant change could be assumed if the measured change is outside the limits of agreement. The limits identified in this study were about ±3 seconds (LLoA, ULoA) for the total duration of 5 repetitions and about ±1 second (LLoA, ULoA) for a single round. Especially in relatively fit, prefrail older adults, a 5-repetition approach could be more sensitive to capture small changes and clinically more useful for early identification of declining physical capacity.

The test-retest reliability analysis showed excellent results. Looking at the average total test duration, we observed a shorter total test duration in the retest. On average, retests took place 6 days after the initial measurement, which makes an improved TUG test performance through training effects unlikely. A similar observation was described elsewhere [45]. A learning effect could account for the shorter duration of the retest. Since participants were already familiar with the test from the first measurement, they may have been more confident and faster during the retest. Other iTUG systems measuring at the height of the navel, (ICC=0.97) [49] and the lower back (ICC=0.9 - 0.96) [50] show slightly higher test-retest reliability

XSL•FO
RenderX

values. This is to be expected, as the “up&go app” pocket position, specifically selected for the self-assessment purpose, is located more distant from the center of mass. Another reason could be that individuals wore different clothes with different types of pockets; these pockets may be tight or loose, which aggravates individual differences.

For future monitoring purposes, machine learning could be implemented to enable an in-app plausibility check. This might be a helpful feature to detect implausible changes in test duration (eg, >20% compared with previous test performances very recently).

Limitations

The sample was predominantly White, and participants were relatively fit, as demonstrated by the TUG performance. The average time needed to complete one repetition (9.7, SD 1.7 seconds; Table 1) is below the threshold of 12 seconds to distinguish normal versus below normal mobility [38]. The results are therefore limited in their transferability to more diverse cultural populations with greater mobility restrictions. For pragmatic reasons, we did not include a gold-standard reference measure, that is, optoelectronic measurement, limiting the results of this study accordingly. A segmental analysis of the “up&go app” data (sit-to-stand, walking, turning, and stand-to-sit) was not considered to be useful for self-monitoring. Furthermore, signal noise caused by the pocket-worn approach would be expected, as the smartphone moves back and forth in the trouser pocket in addition to the body movement during the test. Thus, the pocket-worn approach should be considered as a consumer-centered approach and not as a measurement system for segmental or kinematic analysis. For an accurate segmentation of the test, the use of technology that is placed close to the center of mass would be recommended. The instruction to walk at “usual walking speed” leaves room for individual interpretation, hampering standardization. The extent to which the gait speed during the TUG test corresponds to an individual’s actual real-world gait speed was not assessed in this study. It is assumed that performing 5 repetitions and the familiar home setting enabled a greater approximation to the normal walking speed. However, an activity measurement of several days would be required for comparison [51].

Future Perspectives

The smartphone app used in this study could be suitable as a physical capacity screening and monitoring tool incorporating an instrumented approach for the widely used TUG test. It is designed to be used by older adults, but it could also be implemented by health care professionals to measure TUG total time in clinical settings.

As this study was aimed at validating the test algorithm, it is currently not possible to draw conclusions on the feasibility and user experience of older adults when operating the app, setting up and conducting the test independently. The app should therefore be examined and further developed in future cocreation studies with the target group.

The investigated app is designed to screen physical capacity, representing one of several health domains. In the future, it could be used as part of a comprehensive digital self-assessment, which should include other relevant risk factors such as physical activity, cognition, and vision in addition to physical capacity [24,52]. Self-guided early detection of risks in these domains could enable timely, more specific clinical diagnostics and the initiation of appropriate care interventions. Implementing the app as an upstream assessment within a digital training platform would allow adopting the dosage and supporting the selection of target-specific exercises to be tailored to the user’s individual level of physical capacity.

Conclusion

The results show excellent concurrent validity and test-retest reliability of the pocket-worn iTUG approach with 5 repetitions. The “up&go app” could be suitable as a self-screening and self-monitoring of physical capacity for older adults at home. It provides a smartphone-based approach to accurately measure the total duration of TUG. This novel approach offers the potential for older adults to take an active role in their health management and preclinical risk detection.

Acknowledgments

This publication is based on data from the study "Smart Aging in Community Contexts: Testing Intelligent Assistive Systems for Self-regulation and Co-regulation under Real-Life Conditions" (SMART-AGE). The study is funded by the Carl-Zeiss-Foundation. We thank all project members from the SMART-AGE project and the TSG Research Lab for their contribution. In particular, the authors extend thanks to Carlo Tacconi, Jan Spielmann, Jan Mayer, Ronny Bergquist, Sarah Zimmermann, Tobias Eckert, Claudia Hellmund, Franziska Kramer-Gmeiner, Meike Steinhilber, Anna-Lena Schubert and Alica Mertens.

Authors' Contributions

MJB, CB, and KGO conceptualized the study. MJB, AB, EL, and MS took part in data collection. MJB and SM took part in the data analysis. MJB drafted the manuscript, JK, SM, CB, and KGO took part in a preliminary process of reviewing and editing the manuscript. All authors participated in the subsequent editing process, read the entire manuscript, and agreed to the published version.

XSL•FO
RenderX

Conflicts of Interest

SM holds a share of the mHealth Technologies srl company. None of the other authors declared conflicts of interest. The “up&go app” is freely available in the Google Play and Apple App Store, in German speaking countries.

References

1. Pew Research Center. Share of those 65 and older who are tech users has grown in the past decade. 2022. URL: https://www.pewresearch.org/short-reads/2022/01/13/share-of-those-65-and-older-who-are-tech-users-has-grown-in-the-past-decade/ [accessed 2024-10-09]
2. Chen C, Ding S, Wang J. Digital health for aging populations. Nat Med 2023 Jul;29(7):1623-1630. [doi: 10.1038/s41591-023-02391-8] [Medline: 37464029]
3. Fotteler M, Risch B, Gaugisch P, et al. Obstacles to using assistive technology for older adults - results from a focus group analysis. Stud Health Technol Inform 2021 May 27;281:994-998. [doi: 10.3233/SHTI210327] [Medline: 34042822]
4. Wilson SA, Byrne P, Rodgers SE, Maden M. A systematic review of smartphone and tablet use by older adults with and without cognitive impairment. Innov Aging 2022;6(2):igac002. [doi: 10.1093/geroni/igac002] [Medline: 35243008]
5. Malani P, Kullgren J, Solway E, et al. National Poll on Healthy Aging: Mobile Health App Use Among Older Adults. 2022.
6. United Nations, Department of Economic and Social Affairs, Population Division. World population ageing 2019: highlights. [(ST/ESA/SER.A/430)]. 2019.
7. World Health Organization. Decade of healthy ageing: plan of action. 2020. URL: https://www.who.int/publications/m/item/decade-of-healthy-ageing-plan-of-action [accessed 2024-10-09]
8. Amuthavalli Thiyagarajan J, Mikton C, Harwood RH, et al. The UN Decade of healthy ageing: strengthening measurement for monitoring health and wellbeing of older people. Age Ageing 2022 Jul 1;51(7):afac147. [doi: 10.1093/ageing/afac147] [Medline: 35776669]
9. Beard JR, Officer A, de Carvalho IA, et al. The World report on ageing and health: a policy framework for healthy ageing. Lancet 2016 May;387(10033):2145-2154. [doi: 10.1016/S0140-6736(15)00516-4]
10. Honvo G, Sabico S, Veronese N, et al. Measures of attributes of locomotor capacity in older people: a systematic literature review following the COSMIN methodology. Age Ageing 2023 Oct 28;52(Suppl 4):iv44-iv66. [doi: 10.1093/ageing/afad139] [Medline: 37902521]
11. Podsiadlo D, Richardson S. The timed “Up & Go”: a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991 Feb;39(2):142-148. [doi: 10.1111/j.1532-5415.1991.tb01616.x] [Medline: 1991946]
12. Coelho-Junior HJ, Rodrigues B, Gonçalves IDO, Asano RY, Uchida MC, Marzetti E. The physical capabilities underlying timed “Up and Go” test are time-dependent in community-dwelling older women. Exp Gerontol 2018 Apr;104:138-146. [doi: 10.1016/j.exger.2018.01.025] [Medline: 29410234]
13. Gatenio-Hefling O, Tzemah-Shahar R, Asraf K, Dilian O, Gil E, Agmon M. Revisiting the “Timed Up and Go” test: a 12-s cut-off can predict Hospitalization Associated Functional Decline in older adults. Geroscience 2024 Jul 16. [doi: 10.1007/s11357-024-01280-3] [Medline: 39014130]
14. Ponciano V, Pires IM, Ribeiro FR, et al. Is the Timed-Up and Go Test feasible in mobile devices? A systematic review. Electronics (Basel) 2020;9(3):528. [doi: 10.3390/electronics9030528]
15. Böttinger MJ, Labudek S, Schoene D, et al. “TiC-TUG”: technology in clinical practice using the instrumented timed up and go test-a scoping review. Aging Clin Exp Res 2024 Apr 27;36(1):100. [doi: 10.1007/s40520-024-02733-7] [Medline: 38676844]
16. Milosevic M, Jovanov E, Milenkovic A. Quantifying timed-up-and-go test: A smartphone implementation. Presented at: 2013 IEEE International Conference on Body Sensor Networks (BSN 2013): Cambridge, MA, USA; May 6-9, 2013. [doi: 10.1109/BSN.2013.6575478]
17. Mellone S, Tacconi C, Chiari L. Validity of a Smartphone-based instrumented Timed Up and Go. Gait Posture 2012 May;36(1):163-165. [doi: 10.1016/j.gaitpost.2012.02.006] [Medline: 22421189]
18. Bergquist R, Nerz C, Taraldsen K, et al. Predicting advanced balance ability and mobility with an instrumented Timed Up and Go Test. Sensors (Basel) 2020 Sep 3;20(17):4987. [doi: 10.3390/s20174987] [Medline: 32899143]
19. Bergquist R, Vereijken B, Mellone S, Corzani M, Helbostad JL, Taraldsen K. App-based self-administrable clinical tests of physical function: development and usability study. JMIR Mhealth Uhealth 2020 Apr 27;8(4):e16507. [doi: 10.2196/16507] [Medline: 32338616]
20. Fudickar S, Pauls A, Lau S, et al. Measurement system for unsupervised standardized assessments of Timed Up and Go Test and 5 Times Chair Rise Test in community settings-a usability study. Sensors (Basel) 2022 Jan 19;22(3):731. [doi: 10.3390/s22030731] [Medline: 35161478]
21. Ortega-Bastidas P, Gómez B, Aqueveque P, Luarte-Martínez S, Cano-de-la-Cuerda R. Instrumented Timed Up and Go Test (iTUG)-more than assessing time to predict falls: a systematic review. Sensors (Basel) 2023 Mar 24;23(7):3426. [doi: 10.3390/s23073426] [Medline: 37050485]
22. Mellone S, Tacconi C, Schwickert L, Klenk J, Becker C, Chiari L. Smartphone-based solutions for fall detection and prevention: the FARSEEING approach. Z Gerontol Geriat 2012 Dec;45(8):722-727. [doi: 10.1007/s00391-012-0404-5]

XSL•FO
RenderX

23. Coni A, Ancum JMV, Bergquist R, et al. Comparison of standard clinical and instrumented physical performance tests in discriminating functional status of high-functioning people aged 61 70 years old. Sensors (Basel) 2019 Jan 22;19(3):449. [doi: 10.3390/s19030449] [Medline: 30678268]
24. Böttinger MJ, Litz E, Gordt-Oesterwind K, et al. Co-creating a digital life-integrated self-assessment for older adults: user experience study. JMIR Aging 2023 Sep 26;6:e46738. [doi: 10.2196/46738] [Medline: 37751274]
25. Soubra R, Chkeir A, Novella JL. A systematic review of thirty-one assessment tests to evaluate mobility in older adults. Biomed Res Int 2019;2019:1354362. [doi: 10.1155/2019/1354362] [Medline: 31321227]
26. Goldsack JC, Coravos A, Bakker JP, et al. Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs). NPJ Digit Med 2020;3:55. [doi: 10.1038/s41746-020-0260-4] [Medline: 32337371]
27. Kottner J, Audigé L, Brorson S, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. J Clin Epidemiol 2011 Jan;64(1):96-106. [doi: 10.1016/j.jclinepi.2010.03.002] [Medline: 21130355]
28. Federal Institute for Drugs and Medical Devices. German clinical trials register: SMART aging in community contexts: testing intelligent assistive systems for self-regulation and co-regulation under real-life conditions (SMART-AGE). 2024. URL: https://drks.de/search/en/trial/DRKS00034316 [accessed 2024-12-13]
29. Skevington SM, Lotfy M, O'Connell KA, WHOQOL Group. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. Qual Life Res 2004 Mar;13(2):299-310. [doi: 10.1023/B:QURE.0000018486.91360.00] [Medline: 15085902]
30. Diener E, Emmons RA, Larsen RJ, Griffin S. The Satisfaction With Life Scale. J Pers Assess 1985 Feb;49(1):71-75. [doi: 10.1207/s15327752jpa4901_13] [Medline: 16367493]
31. Groll DL, To T, Bombardier C, Wright JG. The development of a comorbidity index with physical function as the outcome. J Clin Epidemiol 2005 Jun;58(6):595-602. [doi: 10.1016/j.jclinepi.2004.10.018] [Medline: 15878473]
32. Reitan RM. Trail Making Test: Reitan Neuropsychology Laboratory; 1992.
33. Gale TM, Larner AJ. Six-item cognitive impairment test (6CIT). In: Larner AJ, editor. Cognitive Screening Instruments: Springer; 2017:241-253. [doi: 10.1007/978-3-319-44775-9_11]
34. Kempen GIJM, Yardley L, van Haastregt JCM, et al. The Short FES-I: a shortened version of the falls efficacy scale-international to assess fear of falling. Age Ageing 2008 Jan;37(1):45-50. [doi: 10.1093/ageing/afm157] [Medline: 18032400]
35. Kim HJ, Park I, Lee HJ, Lee O. The reliability and validity of gait speed with different walking pace and distances against general health, physical function, and chronic disease in aged adults. J Exerc Nutrition Biochem 2016 Sep;20(3):46-50. [doi: 10.20463/jenb.2016.09.20.3.7] [Medline: 27757387]
36. Jones CJ, Rikli RE, Beam WC. A 30-s chair-stand test as A measure of lower body strength in community-residing older adults. Res Q Exerc Sport 1999 Jun;70(2):113-119. [doi: 10.1080/02701367.1999.10608028] [Medline: 10380242]
37. Google Ireland Limited. Up&go (google play). 2024. URL: https://play.google.com/store/apps/details?id=mht.tsg.upandgo&pli=1 [accessed 2024-10-09]
38. Bischoff HA, Stähelin HB, Monsch AU, et al. Identifying a cut-off point for normal mobility: a comparison of the timed "up and go" test in community-dwelling and institutionalised elderly women. Age Ageing 2003 May;32(3):315-320. [doi: 10.1093/ageing/32.3.315] [Medline: 12720619]
39. Kristensen MT, Ekdahl C, Kehlet H, Bandholm T. How many trials are needed to achieve performance stability of the Timed Up & Go test in patients with hip fracture? Arch Phys Med Rehabil 2010 Jun;91(6):885-889. [doi: 10.1016/j.apmr.2010.01.021] [Medline: 20510979]
40. Evans JD. Straightforward Statistics for the Behavioral Sciences: Brooks/Cole Pub. Co; 1996.
41. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986 Feb 8;1(8476):307-310. [Medline: 2868172]
42. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979 Mar;86(2):420-428. [doi: 10.1037//0033-2909.86.2.420] [Medline: 18839484]
43. Rosner B. Fundamentals of Biostatistics, 8th edition: Cengage Learning; 2016.
44. Tombaugh TN. Trail Making Test A and B: normative data stratified by age and education. Arch Clin Neuropsychol 2004 Mar;19(2):203-214. [doi: 10.1016/S0887-6177(03)00039-8] [Medline: 15010086]
45. Vartiainen MV, Rinne MB, Lehto TM, Pasanen ME, Sarajuuri JM, Alaranta HT. The test–retest reliability of motor performance measures after traumatic brain injury. Adv Physiother 2006 Jan;8(2):50-59. [doi: 10.1080/14038190600700195]
46. Gallagher R, Marquez J, Osmotherly P. Clinimetric properties and minimal clinically important differences for a battery of gait, balance, and cognitive examinations for the Tap Test in idiopathic mormal pressure hydrocephalus. Neurosurgery 2019 Jun 1;84(6):E378-E384. [doi: 10.1093/neuros/nyy286] [Medline: 30010977]
47. Braun T, Thiel C, Schulz RJ, Grüneberg C. Responsiveness and interpretability of commonly used outcome assessments of mobility capacity in older hospital patients with cognitive spectrum disorders. Health Qual Life Outcomes 2021 Mar 1;19(1):68. [doi: 10.1186/s12955-021-01690-3] [Medline: 33648508]

XSL•FO
RenderX

48. Wright AA, Cook CE, Baxter GD, Dockerty JD, Abbott JH. A comparison of 3 methodological approaches to defining major clinically important improvement of 4 performance measures in patients with hip osteoarthritis. J Orthop Sports Phys Ther 2011 May;41(5):319-327. [doi: 10.2519/jospt.2011.3515] [Medline: 21335930]
49. Yamada S, Aoyagi Y, Yamamoto K, Ishikawa M. Quantitative evaluation of gait disturbance on an instrumented Timed Up-and-go Test. Aging Dis 2019 Feb;10(1):23-36. [doi: 10.14336/AD.2018.0426] [Medline: 30705765]
50. Miller Koop M, Ozinga SJ, Rosenfeldt AB, Alberts JL. Quantifying turning behavior and gait in Parkinson's disease using mobile technology. IBRO Rep 2018 Dec;5:10-16. [doi: 10.1016/j.ibror.2018.06.002] [Medline: 30135951]
51. Warmerdam E, Hausdorff JM, Atrsaei A, et al. Long-term unsupervised mobility assessment in movement disorders. Lancet Neurol 2020 May;19(5):462-470. [doi: 10.1016/S1474-4422(19)30397-7] [Medline: 32059811]
52. Barry E, Galvin R, Keogh C, Horgan F, Fahey T. Is the Timed Up and Go test a useful predictor of risk of falls in community dwelling older adults: a systematic review and meta-analysis. BMC Geriatr 2014 Feb 1;14:14. [doi: 10.1186/1471-2318-14-14] [Medline: 24484314]

Abbreviations

ICC: intraclass correlation coefficient
iTUG: instrumented Timed Up and Go Test
LLoA: lower level of agreement
MID: minimal important difference
TUG: Timed Up and Go Test
ULoA: upper level of agreement
UN: United Nations
WHO: World Health Organization

Edited by AF Abdulai; submitted 18.10.24; peer-reviewed by H Kocaman, S Rego; revised version received 13.12.24; accepted 02.01.25; published 21.03.25.

Please cite as:
Böttinger MJ, Mellone S, Klenk J, Jansen CP, Stefanakis M, Litz E, Bredenbrock A, Fischer JP, Bauer JM, Becker C, Gordt-Oesterwind K
A Smartphone-Based Timed Up and Go Test Self-Assessment for Older Adults: Validity and Reliability Study
JMIR Aging 2025;8:e67322
URL: https://aging.jmir.org/2025/1/e67322
doi:10.2196/67322

© Melissa Johanna Böttinger, Sabato Mellone, Jochen Klenk, Carl-Philipp Jansen, Marios Stefanakis, Elena Litz, Anastasia Bredenbrock, Jan-Philipp Fischer, Jürgen M Bauer, Clemens Becker, Katharina Gordt-Oesterwind. Originally published in JMIR Aging (https://aging.jmir.org), 21.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Perceptions of the Use of Mobile Apps to Assess Sleep-Dependent Memory in Older Adults With Subjective and Objective Cognitive Impairment: Focus Group Approach

Aaron Lam[1,2,3], PhD; Simone Simonetti[1,2], PhD; Angela D'Rozario[3], PhD; David Ireland[4], PhD; DanaKai Bradford[4], PhD; Jurgen Fripp[4], PhD; Sharon L Naismith[1,2], PhD

[1]Healthy Brain Ageing Program, The Brain and Mind Centre, The University of Sydney, 96 Mallett Street, Camperdown, Australia

[2]School of Psychology, The University of Sydney, Camperdown, Australia

[3]Woolcock Institute of Medical Research, Macquarie University, Macquarie Park, Australia

[4]Australian eHealth Research Centre, Commonwealth Scientific and Industrial Research Organisation, Herston, Australia

Corresponding Author:
Aaron Lam, PhD
Healthy Brain Ageing Program, The Brain and Mind Centre, The University of Sydney, 96 Mallett Street, Camperdown, Australia

Abstract

Background: Sleep-dependent memory (SDM) is the phenomenon where newly obtained memory traces are consolidated from short-term memory stores to long-term memory, underpinning memory for daily life. Administering SDM tasks presents considerable challenges, particularly for older adults with memory concerns, due to the need for sleep laboratories and research staff being present to administer the task. In response, we have developed a prototype mobile app aimed at automating the data collection process.

Objective: This study investigates the perspectives of older adults, with subjective or objective cognitive impairment, regarding barriers and facilitators to using a new mobile app for at-home assessment of SDM.

Methods: In total, 11 participants aged 50 years and older were recruited from the Healthy Brain Ageing memory clinic, a specialized research memory clinic that focuses on the assessment and early intervention of cognitive decline. Two focus groups were conducted and thematically analyzed using NVivo (version 13; Lumivero).

Results: On average, participants were aged 68.5 (SD 5.1) years, and 4/11 were male. Eight participants had subjective cognitive impairment, and 3 participants had mild cognitive (objective) impairment. Two main themes emerged from the focus groups, shedding light on participants' use of mobile phones and the challenges and facilitators associated with transitioning from traditional laboratory-based assessments to home assessments. These challenges include maintaining accurate data, engaging with humans versus robots, and ensuring accessibility and task compliance. Additionally, potential solutions to these challenges were identified.

Conclusions: Our findings underscore the importance of app flexibility in accommodating diverse user needs and preferences as well as in overcoming barriers. While some individuals required high-level assistance, others expressed the ability to navigate the app independently or with minimal support. In conclusion, older adults provided valuable insights into the app modifications, user needs, and accessibility requirements enabling home-based SDM assessment.

(JMIR Aging 2025;8:e68147) doi:10.2196/68147

KEYWORDS

aging; mild cognitive impairment; subjective cognitive impairment; digital health; cognition; neuropsychology; sleep

Introduction

There are over 55 million people living with dementia [1]. Dementia is an umbrella term for a diverse number of substantial cognitive impairments that significantly interfere with an individual's daily living. The most common form of dementia is Alzheimer disease, in which memory impairment is a critical early feature [2] that is intricately linked to diminished well-being and quality of life [3]. Some facets of memory can be assessed by standardized neuropsychological tests that examine a person's ability to recall information presented to them after a short delay (eg, 20 - 30 minutes). An aspect of memory that cannot be assessed during a standardized neuropsychological test is sleep-dependent memory (SDM). This is a process by which memories encoded during the day are consolidated and strengthened during sleep [4]. This mechanism greatly influences overall memory capabilities by bolstering long-term memory storage, integration, and retrieval.

To date, most studies have explored SDM in younger adults [5]. Thus, tasks are tailored to suit the cognitive capacities of

XSL•FO
RenderX

younger individuals and often involve memorizing many items. This poses significant challenges for older adults, particularly those with cognitive impairments, leading to a notable research gap in this demographic. Given this gap and the importance of SDM in memory for daily life, our team developed an SDM task [6] tailored explicitly for older adults with cognitive concerns and mild cognitive impairment (MCI). MCI refers to a transitional state between normal aging and dementia, where cognitive impairment is apparent, but daily functioning is largely intact. Our prior work demonstrated that individuals with multiple-domain MCI had significantly more compromised SDM than healthy controls and those with single-domain MCI [6]. In this study, poorer SDM was linked to having greater sleep apnea severity for older adults without MCI. In contrast, for those with MCI, poorer performance was associated with decreased sleep spindle duration and smaller hippocampal subfield size. As various age-related sleep changes occur both naturally [7] and with neurodegenerative diseases [8,9], it is crucial to understand how these alterations impact SDM. Notably, preliminary evidence suggests that SDM may be improved through sleep apnea management [10], transcranial electrical stimulation [11], and acoustic stimulation [12]. These findings underscore the importance of identifying clinical correlates and predictors of poor SDM performance to inform targeted interventions.

Unfortunately, our understanding of SDM is limited to small sample case-control studies, and it is not routinely assessed in clinical evaluations. This may be partly attributable to the type of task used and the setting within which most research has been conducted. In addition to being predominantly in younger healthy samples, studies to date have largely been conducted in sleep laboratories where participants are asked to do memory testing before and after sleep [13]. This requires participants to sleep in an unfamiliar environment, meaning their sleep that night may not reflect their usual sleep [14], and they are often asked to learn significant amounts of information before sleep using tasks that are not feasible or suitable for clinical settings. Another barrier is the high demand for staff to conduct these studies (Naismith, personal communication, 2024). Staff must be present to administer SDM tasks before sleep (often within 3 hours before sleep) and after sleep (often after 1-hour postwaking) and to stay overnight to monitor the participant. Together, the requirement for sleep laboratories and high staff demand result in high costs associated with conducting SDM tasks. These barriers could be addressed by adapting SDM tasks for older clinical samples and delivering the tasks in the home environment.

The adoption of digital health technology in older adults has been increasing [15], including by individuals with subjective concerns about their cognition and those with MCI [16]. However, when considering new app-based approaches, barriers and facilitators that influence adoption and effective use must be identified [17]. A prior systematic review showed that barriers preventing the use of digital technology in older adults (including those with cognitive impairment) can include motor, sensory, cognitive, lack of familiarity, and device-specific challenges, though these factors can be mitigated through design modifications [18]. In contrast, digital health technology adoption can be facilitated by the perceived usefulness of the app, pre-existing knowledge, and ease of use. For SDM tasks, additional considerations are required, given that they are optimally tested under strict timings (ie, before and after sleep) and may require participants' sustained attention and interaction with a mobile app for extended periods (eg, potentially up to an hour).

This study sought to explore the perspectives of older adults with subjective or objective cognitive impairment regarding barriers to and facilitators of an at-home assessment of SDM using a new "chatbot" (mobile app), approach to assessment, named "Sleep Memories." A chatbot is designed to simulate human-like conversation with users through text or voice interactions and, in our scenario, to deliver a memory task and record a participant's responses. Previous work has shown the use of chatbots in various health populations, including chronic pain populations to deliver pain education [19], and monitoring of chemotherapy outcomes in older adults with cancer [20]. The objective was to use co-design methodologies to assess an initial app prototype, which could be incorporated into a revised version suitable for clinical testing. We were specifically interested in potential barriers to and facilitators of use.

Methods

Digitalization of the SDM Task

The SDM task, as described in Lam et al [6], was initially adapted to be suited for an app-based platform (Figure 1). This task is a verbal memory task that comprises 32 word pairs, half of which are semantically related and the other half unrelated. In step 1, participants are presented with all 32 pairs. In step 2, individual words are presented, for which participants are required to identify the corresponding pair word by speech or text. As described in the initial study [6], participants performed this task 4 times in the evening within 3 hours of bedtime to facilitate learning. In the morning, participants completed the recall component (step 2) and a multiple-choice test, comprising an individual word and 4 possible answers (1 correct corresponding pair word and 3 incorrect pair words), approximately 1 hour after waking. This task was digitalized for use in a mobile app, using a chatbot for automatic delivery and data collection. To humanize the chatbot for prototype testing, we named it "Aurora." The Sleep Memories app's development, in-house testing, and evaluation are detailed in Ireland et al [21].

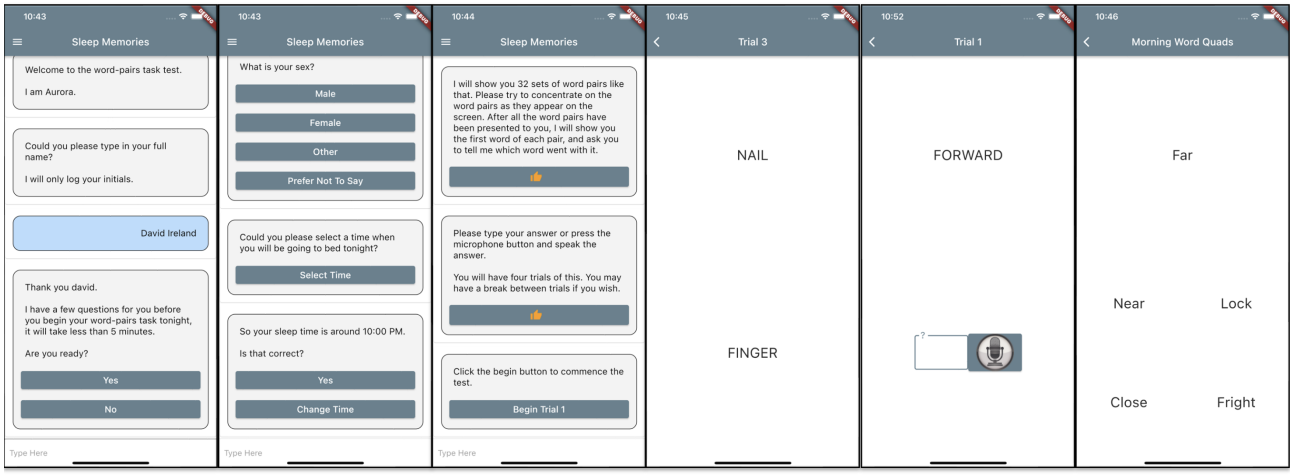
Sleep Memories was designed to collect information about a participant's habitual sleep patterns, through a short survey on opening the app, and deliver the SDM task within 3 hours of the participant's regular sleep time. After the participants respond to the questions about their sleep, the app provides a trial run, enabling participants to become comfortable with its functionalities. In line with Lam et al [6], each word pair is presented to the participants for 10 seconds. For the recall component, the app offers a 10-second window, allowing participants to either type or vocally record their answers, ensuring a user-friendly response period. In the morning, participants are prompted to confirm their bedtime and wake-up

XSL•FO
RenderX

time, which is followed by recall and multiple-choice testing to assess their memory retention.

Qualitative methods were selected for this study, as they are ideally suited to elicit participants' perspectives of Sleep Memories and their experiences with current digital health technologies relative to other methodologies. An interpretative philosophical perspective [22] was selected, as this method allowed us to situate the participants' perspectives within their broader life contexts and allowed for flexibility and adaptability regarding the emergent themes and meanings arising from the data.

Figure 1. Screen captures of the Sleep Memories app. (**A**) Aurora (chatbot) introducing themselves to the participant. (**B**) Aurora asking demographic and habitual sleep pattern questions. (**C**) Aurora giving instructions for the word-pairs task. (**D**) Learning component of the word-pairs task being delivered by the app (**E**) Recall component of the word-pairs task being delivered by the app. (**F**) Multiple-choice testing component of the word-pairs task being delivered by the app.



(A) (B) (C) (D) (E) (F)

Participant Recruitment

Participants were recruited via convenience sampling from the Healthy Brain Ageing (HBA) memory clinic. As described previously [23], the HBA clinic is a specialized memory clinic that accepts referrals from general practitioners and medical specialists and provides comprehensive neuropsychological, medical, and mood assessments for older adults aged >50 years with recent subjective cognitive or mood decline. Exclusion criteria are a Mini-Mental State Examination score<20 [24], intellectual disability, insufficient English proficiency for standardized neuropsychological assessment, history of nonaffective psychiatric disorder (eg, schizophrenia), history of stroke, history of head injury (loss of consciousness>30 minutes) or other neurological disorder (eg, epilepsy), or current substance dependence or abuse. Basic demographic (eg, age, sex, and years of education) and clinical data (eg, Mini-Mental State Examination score and MCI status) collected at the HBA clinic are reported here for descriptive purposes. MCI status was categorized into 3 distinct groups: subjective cognitive impairment (absence of objective cognitive impairment), nonamnestic MCI (objective cognitive impairment in nonmemory domains), and amnestic MCI (defined by objective cognitive impairment in memory domain). Participants were invited to participate in the study via email or phone call. Interested participants were provided with a participant information statement identifying the research team, funding, research aims, and risks and explaining confidentiality, consent, and withdrawal. Once informed consent was obtained, participants were invited to attend 1 of 2 digital focus groups via Zoom (Zoom Video Communications).

Data Collection

Before the focus groups, participants were required to complete a digital survey consisting of an Abbreviated version of the previously validated Mobile Device Proficiency Questionnaire (AMDPQ) [25]. In general, the abbreviated version of the questionnaire captured 7 items on basic proficiency in mobile phone use (eg, charging the phone and typing with the keyboard), 8 items on proficiency in communication (eg, sending emails and messaging), 3 items on data and file storage (eg, transferring files from mobile to computer), 8 items on internet use, and 6 items on troubleshooting and software management (eg, updating and deleting apps). Higher scores on each item indicated greater proficiency in each aspect.

A semistructured focus group guide was developed in collaboration with AL (male, program manager), Peta Mills (female, clinical project officer, PhD), SS (female, postdoctoral research associate), and SLN (female, professor and clinical neuropsychologist with extensive experience in focus groups, workgroups, interviews, and working parties for a variety of purposes such as to develop guidelines or models). The focus groups and questions were structured around three core themes (see Multimedia Appendix 1 for the question guidelines used):

1. Current use of mobile phone apps: informed by prior literature [17,18], this theme focuses on participants' current use of mobile phone apps, particularly for health and sleep-related functions. The theme includes use habits, types of health apps used, and experience with speech recognition features on mobile devices. Related questions were designed to assess participants' familiarity with mobile phone technology, health-related apps, as well as their interactions and experiences with speech recognition features.

XSL•FO
RenderX

2. Interest in completing cognitive tasks through a mobile app: informed by previous studies [18,26], this theme was selected to explore participants' willingness and comfort in using a mobile app for memory assessments. This theme includes the potential benefits and challenges of this approach compared to traditional methods. The aim is to understand the acceptability, motivation, practicality, and barriers or facilitators of adopting mobile apps for cognitive assessments.
3. Feedback on the proposed mobile app: this theme seeks to gather participants' perspectives on various aspects of the proposed mobile app, including the clarity of instructions and tasks, visual appeal, potential distractions, and privacy concerns. Participants were also invited to provide suggestions for enhancing the usability and engagement of the mobile app. The questions specifically targeted feedback on the app's features, such as the chatbot and task length or design, and were developed based on internal testing of the mobile app.

A Microsoft PowerPoint presentation (Multimedia Appendix 2) was used to guide participants through the focus group material. The presentation introduced the concept and significance of SDM, highlighting its potential impairment in aging and cognitive decline and its understudied nature due to current limitations in traditional data collection methods (ie, in a sleep laboratory setting). We showed how our mobile app may be a novel solution to these barriers to facilitate home-based assessments. The presentation also illustrated how participants would interact with Aurora, including the delivery of the word-pairs task, in which the process and structure of the SDM assessment involving learning and recall word pairs were explained.

Data were collected via 2 focus groups between October and November 2022. Each focus group was facilitated by separate researchers: SS, a postdoctoral research associate with 10 years of applied research experience and 4 years of qualitative research experience (interviews and focus groups), and Peta Mills, a clinical project officer with 11 years of applied research experience and 18 months of qualitative research experience (focus groups). At the time of the focus groups, both facilitators were not involved in this research project in any capacity and thus provided an impartial perspective. AL and DI (male, software engineer) attended the focus group sessions to answer participants' logistical and technical questions. Focus groups lasted 60 - 90 minutes and were audio-recorded, transcribed, and de-identified.

Ethical Considerations

This study was approved by The University of Sydney Human Ethics Committee (2022/HE000563) and conducted in accordance with the World Medical Association Declaration of Helsinki. Written informed consent was obtained from all participants prior to any study procedures and given the option to opt out of the study at any given time. Participants were given a A$20 (US $12.75) gift voucher as compensation for their time. All data collected were deidentified to ensure participant privacy and confidentiality.

Data Analysis

For clinical and survey data, frequency and descriptive statistics were computed using SPSS Statistics (version 24.0; IBM Corp). Focus group transcripts were analyzed thematically in NVivo (version 13; Lumivero), a qualitative data analysis computer software package, using the techniques described by Braun and Clarke [27]. The thematic analysis approach described by Braun and Clarke [27] is a systematic, iterative, and inductive process comprising 6 phases. Initially, researchers immerse themselves in the data to discern patterns and potential coding schemes. Subsequently, they generate initial codes to categorize salient features systematically. These codes are then organized into potential themes. Themes undergo rigorous validation to ensure coherence and alignment with the dataset. Once themes are finalized, they are defined, and any hierarchical relationships are identified. Finally, thematic insights are woven into a narrative to address relevant research questions.

As we had identified our 3 core themes during the development of the focus group guide, our thematic analysis was focused on a deeper exploration of these themes. Following Braun and Clarke [27] phases, AL and SS became familiar with the data via repeated reading and individually created initial codes for all transcripts by noting evocative phrases, ideas, and perceptions. Codes were then organized into meaningful groups and sorted into subthemes. Where there was disagreement in coding or theme development, the codes or themes were discussed and refined. To ensure the study conformed to best practice guidelines, we used the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist and met 27 of 32 items [28].

Results

Participants

In total, 19 people from the HBA memory clinic were invited to participate in the focus groups, and 11 of these agreed to participate. Their demographics and clinical information are reported in Table 1. Briefly, the sample on average was aged 68.5 (SD 5.1) years and predominately female (n=7, 64%). The sample also comprised 8 subjective cognitive impairment, 2 nonamnestic MCI, and 1 amnestic MCI. Of these participants, 10 completed the AMDPQ, with 1 participant declined, as they did not have sufficient proficiency to access the digital questionnaire. On average, participants demonstrated relatively high mobile proficiency (see Multimedia Appendix 3 for a full breakdown of participant responses to each individual item).

XSL•FO
RenderX

Table . Participant demographic and clinical information.

	Values
Age (years), mean (SD)	68.5 (5.1)
Sex (male), n (%)	4 (36)
Education (years), mean (SD)	15.1 (2.9)
MMSE[a] score (out of 30), mean (SD)	29.2 (0.9)
Subjective cognitive impairment, n (%)	8 (73)
Nonamnestic mild cognitive impairment, n (%)	2 (18)
Amnestic mild cognitive impairment, n (%)	1 (9)
AMDPQ[b] total score (out of 25), mean (SD)	20.5 (4.6)
AMDPQ—mobile device basic subscale (out of 5), mean (SD)	4.4 (0.6)
AMDPQ—communication subscale (out of 5), mean (SD)	4.0 (1.0)
AMDPQ—data and file storage subscale (out of 5), mean (SD)	3.7 (1.3)
AMDPQ—internet subscale (out of 5), mean (SD)	4.2 (1.1)
AMDPQ—troubleshooting and software management subscale (out of 5), mean (SD)	4.1 (1.1)

[a]MMSE: Mini-Mental State Examination.

[b]AMDPQ: Abbreviated Mobile Device Proficiency Questionnaire.

Thematic Analysis

Overview

The thematic analysis revealed two broad themes relevant to the 3 core focus group themes: (1) context of mobile phone use and (2) shifting from traditional laboratory-based assessments to home assessments. The latter theme should be considered within the framework of the former theme, as shown in Textbox 1. That is, theme 2 considers the challenges and facilitators identified by participants when shifting from laboratory settings to home environments within 2 specific contexts: (1a) participants' existing familiarity and proficiency with mobile phones and apps and (1b) the prevalent stereotypes and preconceived notions concerning aging and technology.

Textbox 1. Breakdown of the overarching themes and subthemes identified in our study by 2 primary domains: "context of mobile phone use" and "shifting from traditional laboratory-based assessments to home assessments."

Context of mobile phone use

- Current phone use
- Aging

Shifting from traditional laboratory-based assessments to home assessments

- Accuracy of data
- Human versus robot
- Accessibility
- Compliance

Context of Mobile Phone Use

Current Phone Use

Participants were asked to share their mobile phone use patterns and specify the types of apps they used. Across participants, mobile phone use varied substantially from the use of basic phone functions, such as phone calls, text messaging, and camera functions, to more complex use involving apps, such as emails, social media, search engines, maps, news, weather, and music. Although minimally used, health and sleep apps were accessed for relaxation and to monitor sleep, heart rate, and physical activity. One participant felt highly proficient in phone use and that the phone was integral to the organization of their day-to-day life: "I would be lost without my mobile phone." Speech-to-text functions were used as needed, particularly when driving and composing text messages, using navigation maps, or playing music. However, it is worth noting that not all participants chose to use this feature.

Aging

Individuals who reported limited phone use perceived their age as a barrier to gaining proficiency with new technologies.

XSL•FO
RenderX

> *I don't think I am alone in this, but it is a real problem when you're a bit older trying to pick up the technology that seems to be intuitive to other people.*

Participants expressed a sense of not keeping pace with technology, which stemmed from their preference for relying on what they were familiar with or their self-perception as old-fashioned. Even when participants successfully used a new technology, they felt the need to use it continuously to avoid losing the knowledge of how to operate it effectively.

> *There is so much learning involved and if you don't use these things more than once in 6 months it's gone.*

Irrespective of the reasons for their limited access to, and use of, technology, a prevailing sentiment among these individuals was a sense of missing out.

> *I have a Samsung phone, one of the latest, and I know there are lots of lovely things for me and I'm not using them. I would use them if they were set up.*

Younger adults were viewed as having an innate familiarity with and competence in using digital tools and, thus, were flagged as potential facilitators of accessing and using technology. Subsequently, participants who did not have a younger adult available to facilitate their use of new technologies felt at a disadvantage. These observations reinforced the prevailing belief that older generations faced a variety of disadvantages with technology.

The insights gathered suggest a spectrum of technology assistance needs among participants. Some individuals may require comprehensive setup support, including step-by-step instructions and human interaction throughout their tech journey, while others might only require minimal assistance. This underscores the importance of adopting 1 of 2 design approaches for apps: either a straightforward and user-friendly design that prioritizes ease of use or, in the case of complex apps, the provision of substantial human support.

These findings carry significant implications for the design of apps intended for home use as opposed to controlled laboratory settings. They serve as a foundation for understanding the crucial factors and context that should inform the app design process when moving apps from the laboratory to the home.

Shifting From Traditional Laboratory-Based Assessments to Home Assessments

Overview

There were several considerations for moving the SDM task from laboratory-based assessments to home assessments raised during focus groups. Participants identified facilitators as well as challenges and proposed potential solutions to overcome these obstacles (Table 2).

Table . Moving sleep-dependent memory from laboratory to home: challenges, solutions, and changes.

Challenge	Potential solution	App updates
Unexpected events during task completion (eg, phone ringing and someone knocking on the door).	Function to pause and recommence task.	The task was not implemented with a pause option because it could impact the results. Instead, Aurora advised participants to find a quiet, distraction-free place at the start of the task. If they could not finish because of unforeseen circumstances, they were offered another chance on a different night.
Difficulty with setting up apps and need help with troubleshooting.	Set up the app for participants and having someone to contact.	Through email, participants can reach a contact person who will provide technical support.
Participants did not want to enter identifying information (eg, names or date of birth).	Using a participant ID rather than identifying information.	We have given the participants an option to enter their participant ID. We now only collect participant initials and not name or date of birth.
Participants flagged that sometimes they ignore push notifications on their phones in the evening.	Option for the phone to ring or a text message.	The app now sends additional notifications to remind the participants to complete the task. Furthermore, the phone now sends short audio notification reminders to participants.

Accuracy of Data

Participants generally regarded the home environment as conducive to achieving higher accuracy levels relative to the laboratory environment, which was described as "ghastly" and distressing. Indeed, because of this perception, participants questioned the reliability of sleep laboratories, primarily because of the unconventional and unfamiliar sleep conditions they imposed. These doubts extended to the execution of the SDM task in the laboratory. Participants posited that conducting the SDM task in the laboratory might cause reduced accuracy compared to conducting the same tasks within the comfort of their homes, where they felt more relaxed as they did not have a researcher watching them and where they could expect a more restful night's sleep. Similarly, there was a recognition that the peculiar sleep laboratory environment might introduce sleep disturbance that may disrupt SDM processes. Therefore, the results of the SDM task completed in a laboratory may not truly reflect an individual's performance. To further illustrate these insights, one participant shared their personal experience with a sleep laboratory:

> *I recently, did a sleep study. It's just not what I consider to be an indicator of anything that's going on. I had a two-and-a-half-hour sleep if that through the whole night strapped into all these other things*

XSL•FO
RenderX

> *as well ... So definitely at home I think you'll get more relaxed, and you get a true picture of what's going on.*

Conversely, having a researcher present and demanding attention could lead to better performance.

> *In the clinic, I was being put on the spot ... "we're not just being casual here, we're doing a trial and I'm seeking your recall results, so you need to pay attention." So being in the clinic I was really on purpose for fear of failure ... At home I'd be far more relaxed and maybe that's a good thing too.*

In the laboratory setting, participants interact with the SDM task by saying their responses aloud. When considering the at-home mobile app version of the SDM task, participants were given a choice between typing their responses into an open text field or vocalizing their answers. Preferences varied, as participants felt that while using verbal responses may be more convenient than typing, the accuracy of verbal responses could be reduced. Nonetheless, being able to choose to use a text or a verbal response resonated positively with participants.

Human Versus Robot

During thematic analysis, some initial disagreements arose among the 2 scorers regarding the classification of human versus robot. AL suggested that the theme "human versus robot" may encompass accessibility as well, while SS suggested creating a separate theme for "accessibility." To resolve this, the team engaged in iterative discussions, referring back to the original transcripts to ensure alignment with participants' narratives. Through this process, a consensus was reached to have "human versus robot" and "accessibility" as 2 separate themes.

The facilitators explained to participants that they would primarily interact with Aurora, a chatbot designed to guide them through the SDM task via written text or verbal instructions. Participants overwhelmingly agreed that any kind of computerized intelligence should be as humanlike and natural as possible, suggesting that a computerized voice might be difficult to engage with.

> *I found it very easy to do it in the lab and the interaction with [researcher name] it was simple to do and it was a softer voice and it is relaxed and it is encouraging. All that subliminal processing of getting the task done.*

Relying solely on a computerized and automated system without human support raised several challenges, and the lack of human interaction was primarily viewed as a drawback. Indeed, participants flagged the importance of human support, both during setup and for ongoing assistance. Participants also identified that the app may struggle to accommodate diverse accents or may lack the mechanism to clarify potential response errors, tasks that were easily accomplished by a human. To assist with these challenges, participants recommended that the app should be capable of detecting potential errors and should offer corrections or alternative suggestions when necessary.

Additionally, participants expressed apprehension around the security of data, given personal information was to be entered directly into the app. There was a hesitation to provide personal information, such as names or dates of birth. Instead, participants proposed using participant IDs only for entry into the app and that other information, such as date of birth, is gathered by the researcher in another way and is linked externally to this ID. However, other participants displayed indifference to this issue, expressing a belief that personal information is already widely accessible: "Everyone knows everything about us anyway."

Accessibility

Participants identified 2 primary barriers that could impact the app's accessibility. These were physical and economic barriers. Several participants expressed concerns regarding age-related vision problems, which can hinder their ability to read or view small text. Participants proposed using larger screens, like tablets, and increasing font sizes might address this barrier. Participant responses from the AMDPQ suggested that certain participants exhibited greater proficiency when using personal computers or tablets than smartphones, indicating the potential value of designing apps with flexibility and accessibility across different platforms.

> *I do I get a bit irritable with the smallest of the screen prefer to do things on the laptop.*

Another physical barrier that was discussed was physical dexterity, which refers to fine motor skills and the ability to move their fingers. Age-related factors, such as tremors and arthritis, may be challenging for older adults to respond to Aurora or complete the SDM task.

> *The dexterity on the fingers may be important on how well someone clicks on the screen ... I might have trouble because I haven't done it on a phone and only do it on a PC.*

One participant raised concerns about whether the app required high bandwidth to use. High bandwidth apps can increase difficulty in accessing and using mobile apps as well as increase costs associated with using the app. Therefore, it is important to optimize apps for low bandwidth.

> *Is your application bandwidth heavy that it's just going to slow things down for those of us who actually don't have a proper wireless connection or a slow data connection and it's just going to make it after a while just unusable nothing.*

Participants noted that the initial setup phase of new technologies (eg, installing and setting up of apps) was either too challenging or time-consuming. Participants recognized that they required assistance or guidance to set up, navigate, and use technology effectively. This support could encompass various aspects, such as learning to use new devices, understanding complex software apps, or troubleshooting technical issues. However, they found that, in general, such support was lacking or insufficient. The consensus among participants strongly supports the idea that offering setup assistance, particularly during the initial stages of technology adoption, would be immensely beneficial. This will help familiarize participants with the mobile app, save time, and alleviate stress for older adults. Any kind of automation was also supported to avoid repeated data entry:

XSL•FO
RenderX

> *I don't need to put that same information in every time that I'm doing this test because this would be 5 - 10 minutes every day just putting in the same information ... when you are actually setting up the app on your phone is when you actually key that data in and then you don't have to do it again*

Compliance

The facilitators explained to the participants that the home-based completion of the chatbot would require dedicating approximately 40 minutes to the learning component, followed by a subsequent delay of 20 - 30 minutes and another 10 minutes to complete the task. While participants expressed comfort with the overall duration of the task, the true challenge was quarantining the time to do this task.

> *I think anything more than 40 minutes it may be burdensome. I do think 40 minutes is the top. The challenge is disruptions and to really quarantine the time and be strong about that.*

Participants proposed a solution, emphasizing the necessity of proactively earmarking dedicated time in advance to ensure the completion of the task. Having a large time window during which participants could complete the task would provide the flexibility needed for accommodating various schedules and preferences of participants.

> *I would quarantine in the time then for the next day and prepare myself because I don't have any other commitments or pressures on my time.*

Alternatively, participants suggested introducing a pause feature to mitigate any potential disruptions caused by unforeseen events, such as incoming phone calls or unexpected visitors. However, it is crucial to acknowledge that the introduction of a pause feature could potentially influence the integrity of the test. While these features may be convenient for the participants, this will introduce confounds to the results. Subsequently, a pause button was not implemented in Aurora.

> *I always find a pause button is important. For example: if I am starting the memory test or memory exercise right now and someone knock the door ... I have the opportunity that I can come back to that when I need to do something.*

Another challenge identified was participants occasionally forgetting to complete the task. The facilitators explained to the participants that there are existing notifications within the app to remind the participants to do the task in the evening and in the morning. It was suggested that increasing the frequency of notifications could be a viable solution. Moreover, participants highlighted a tendency to ignore push notifications, particularly in the evening. To address this, participants proposed implementing alternative reminder methods, such as phone calls or text messages. Participants identified that their interest and motivation to complete the memory task were closely tied to the app's clarity. The consensus was that instructions and information presented in an easily digestible manner would keep participants' engagement. It was evident that it is important to prioritize clear and straightforward instructions that would lead to sustained participant engagement, especially when participants are asked to complete the task on multiple occasions.

> *Okay, if you wanted me to use it repeatedly then it's got to be clear and simple and easy to access if you're talking about once every three months and that's ok I'll come back anytime.*

Participants nearly unanimously highlighted the critical role of feedback on their task performance. They underscored the importance of receiving real-time feedback to gauge their progress and enable a sense of accomplishment and motivation to improve and continue to complete the task for multiple assessments.

> *I don't want money but I do want feedback. I am involved in a lot of medical research I am surprised I don't find out about the results.*

Even participants who did not think that feedback was necessary thought it would be a welcomed supplementary feature, as it would motivate them to continue using the app. Subsequently, Sleep Memories integrated informative feedback to enhance user engagement and promote continued motivation. For others, the motivation was intrinsic (ie, feedback not needed), and people would remain motivated to contribute to science. Participants trusted that scientists knew what they were doing and would follow their guidance because they are experts in the field.

> *... The work that this unit does is so important. Just thanking them for their time and delivering, any material or rewards I find it degrading. I just want to donate my data to be used for the greater good. There is nothing you can give me to make it better.*

Discussion

Principal Findings

This study aimed to investigate the perspectives of older adults both with subjective or objective cognitive impairment regarding barriers and facilitators to using an at-home assessment of SDM through a novel app named Sleep Memories. Through co-design methodologies, we assessed the initial prototype of the mobile app to incorporate user feedback into a revised version suitable for clinical testing.

We found that digitalization and technological advancements can both facilitate and hinder everyday life for older adults. This phenomenon is reflected in the term Janus-faced technology [29]. According to this concept, the successful integration of technology can substantially improve daily activities for older adults. However, barriers to use can evoke feelings of alienation and disconnection from the digital world. Understanding these barriers is essential to minimize these negative feelings and to, instead, develop accessible and user-friendly technology solutions for older adults, particularly those with cognitive impairments.

One of the key findings of our study was the importance of app flexibility in accommodating users' diverse needs and preferences and in challenging barriers. While some individuals reported requiring high-level assistance, others indicated that

XSL•FO
RenderX

they could navigate the app independently with minimal support. A recent systematic review highlighted the critical importance of tailoring technology to meet the specific needs and preferences of older adults with MCI and dementia [30]. Below we examine each of the facilitators and barriers we identified and explore corresponding, flexible solutions in light of existing research.

Comparison to Prior Work

The "first night effect" is a confounding variable in sleep research where participants' sleep is disrupted due to being in an unfamiliar environment. This effect has been studied in prior studies. For instance, one study [31] demonstrated its potential to impact sleep disorder diagnoses. The benefit of home-based testing is that participants can sleep in their usual environment, removing the need for multiple night sleep laboratory testing. Furthermore, the SDM data collected are not confounded by the first night effect and more accurately represent a participant's true abilities than traditional SDM testing.

Within the discussion of voice responses, our findings align with prior work, with concerns from older adults regarding the accuracy of audio (eg, the chatbot's ability to accurately capture verbal responses) and text (eg, spelling mistakes) inputs [32]. In response to these concerns, Sleep Memories has implemented a dual-layered validation process. First, an autocorrect function is used to rectify minor spelling inaccuracies automatically. Second, when the system is uncertain of the response, a request for confirmation by a researcher is made. The implementation and availability of a voice response option increase the accessibility of the mobile app for those who may have age-related declines or disabilities [32].

To increase the accessibility of mobile apps for older adults, it is important to navigate both technological and socioeconomic challenges. Our focus groups highlighted concerns like high bandwidth and data consumption, which could impede the functionality and availability of health apps such as "Sleep Memories" for older adults. Furthermore, older adults may have age-related physical limitations, such as arthritis or hand pain, which may impact their interaction with mobile devices. To address the high bandwidth and data consumption, we allow the app to be used without an internet connection, and data are uploaded when connectivity is available. As previously mentioned, the implementation of voice responses can be a solution to enable those with physical limitations to use the app.

Participants raised concerns about forgetting to complete the task, particularly for those who might overlook setting personal reminders. This aligns with previous research in the context of medication compliance, where it was observed that a single reminder was often insufficient to prompt action, suggesting that a backup notification could serve as a viable solution [33]. A solution in our context that was proposed was to increase the frequency and elevate the distinctiveness of reminder notifications (eg, for Sleep Memories, we implemented a unique auditory notification, specifically a vocal prompt stating "it is now time to complete your memory task").

Similar to a previous study [34], we found that incorporating feedback mechanisms, like scores, within eHealth tools is imperative to foster user engagement and empowerment. These features motivate users by providing immediate, tangible feedback on their performance. However, it is essential for app developers and researchers to communicate the context and limitations of their scores, particularly when there is a lack of established normative data. Without proper context, participants may misinterpret their scores, leading to unnecessary anxiety or concerns.

Our study revealed that conducting tasks within a home setting may introduce various distractions. This was aligned with prior research on the environmental interference of unsupervised home cognitive assessments [26]. To mitigate this challenge, 2 solutions emerged during our focus group discussions. The first involves incentivizing participants' engagement through targeted motivation strategies, such as feedback, to maintain focus despite potential distractions. The second suggests the need for flexibility in task design, allowing participants to pause and resume activities, accommodating the unpredictability of the home environment. However, this approach must be balanced against the risk of compromising data integrity. For instance, in "Sleep Memories," participants can pause and take short breaks between trials without permitting interruptions during the trials themselves to safeguard data continuity and accuracy.

Limitations

Our study comprised older adults who had relatively high mobile proficiency at least in the subdomains that we examined, so our findings may not be generalizable to older adults who have a low mobile proficiency. Another limitation is that our sample had less representation from individuals with MCI (n=3), which limited our ability to detect or uncover differences in responses related to cognitive impairment. Future research is necessary to explore this important area by recruiting a larger cohort of individuals with MCI. Similarly, our participants generally scored highly on the AMDPQ, meaning that they showed decent proficiency in mobile phone use. Whether these findings apply to older adults with low proficiency in mobile phones needs further investigation. Nonetheless, our study is one of the first to examine the potential barriers and facilitators to the use of mobile device apps, in particular to complete memory tasks, in those with cognitive impairment.

Future Directions

In addition to the app modifications mentioned earlier, our findings provide valuable insights to meet the needs and preferences of our older adult users. When deploying Sleep Memories, we will first conduct a needs assessment to assess the ability to use existing and new technology as well as the user's circumstances (eg, Do they have time to complete the task in the app without distraction?) and environment (eg, Do they have access to a technology-savvy person in their network?). Based on this assessment, we will offer users various levels of support such as the ability for a phone call by a researcher or various notification options.

Conclusions

Our study demonstrates that while there are barriers related to accessibility, usability, and data integrity to using health-related phone apps in this population, there are important facilitators

XSL•FO
RenderX

that can be implemented in phone apps to create a flexible and inclusive digital health tool. Furthermore, we have shown evidence on the potential feasibility of the "Sleep Memories" app to enable the collection of SDM data within a user's natural environment.

Acknowledgments

The authors would like to extend their deepest gratitude to all the participants who dedicated their time and shared their experiences with them. The authors also wish to acknowledge Peta Mills for her assistance with their focus groups. Finally, the authors thank their clinical and research team at the Healthy Brain Ageing clinic. This study was supported by the National Health and Medical Research Council (NHMRC) Centre of Research Excellence to Optimise Sleep in Brain Ageing and Neurodegeneration (CogSleep), the Commonwealth Scientific and Industrial Research Organisation, and an NHMRC Fellowship awarded to SLN (GNT1135639) and AD (GNT2008001).

Data Availability

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

Conflicts of Interest

SLN has received consulting fees from Eisai and Roche pharmaceuticals. All other authors have no conflicts of interest to declare.

Multimedia Appendix 1
Discussion guide for focus groups.
[DOCX File, 19 KB - aging_v8i1e68147_app1.docx]

Multimedia Appendix 2
Chatbot for word pairs.
[DOCX File, 4460 KB - aging_v8i1e68147_app2.docx]

Multimedia Appendix 3
Abbreviated Mobile Device Proficiency Questionnaire.
[DOCX File, 26 KB - aging_v8i1e68147_app3.docx]

References

1. Dementia. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2023-11-01]
2. Förstl H, Kurz A. Clinical features of Alzheimer's disease. Eur Arch Psychiatry Clin Neurosci 1999;249(6):288-290. [doi: 10.1007/s004060050101] [Medline: 10653284]
3. Zuniga KE, Mackenzie MJ, Kramer A, McAuley E. Subjective memory impairment and well-being in community-dwelling older adults. Psychogeriatrics 2016 Jan;16(1):20-26. [doi: 10.1111/psyg.12112] [Medline: 25737426]
4. Stickgold R, Walker MP. Sleep-dependent memory triage: evolving generalization through selective processing. Nat Neurosci 2013 Feb;16(2):139-145. [doi: 10.1038/nn.3303] [Medline: 23354387]
5. Feld GB, Weis PP, Born J. The limited capacity of sleep-dependent memory consolidation. Front Psychol 2016;7:1368. [doi: 10.3389/fpsyg.2016.01368] [Medline: 27679589]
6. Lam A, Haroutonian C, Grummitt L, et al. Sleep-dependent memory in older people with and without MCI: the relevance of sleep microarchitecture, OSA, hippocampal subfields, and episodic memory. Cereb Cortex 2021 May 10;31(6):2993-3005. [doi: 10.1093/cercor/bhaa406] [Medline: 33565576]
7. Lam A, Kong S, Naismith SL. Recent advances in understanding of sleep disorders and disturbances for dementia risk and prevention. Curr Opin Psychiatry 2024 Mar 1;37(2):94-100. [doi: 10.1097/YCO.0000000000000921] [Medline: 38226546]
8. Orlando IF, O'Callaghan C, Lam A, et al. Sleep spindle architecture associated with distinct clinical phenotypes in older adults at risk for dementia. Mol Psychiatry 2024 Feb;29(2):402-411. [doi: 10.1038/s41380-023-02335-1] [Medline: 38052981]
9. Lam AKF, Carrick J, Kao CH, et al. Electroencephalographic slowing during REM sleep in older adults with subjective cognitive impairment and mild cognitive impairment. Sleep 2024 Jun 13;47(6):zsae051. [doi: 10.1093/sleep/zsae051] [Medline: 38394454]
10. Djonlagic IE, Guo M, Igue M, Kishore D, Stickgold R, Malhotra A. Continuous positive airway pressure restores declarative memory deficit in obstructive sleep apnea. Am J Respir Crit Care Med 2021 May 1;203(9):1188-1190. [doi: 10.1164/rccm.202011-4253LE] [Medline: 33347378]

XSL•FO
RenderX

11. Ketz N, Jones AP, Bryant NB, Clark VP, Pilly PK. Closed-loop slow-wave tACS improves sleep-dependent long-term memory generalization by modulating endogenous oscillations. J Neurosci 2018 Aug 15;38(33):7314-7326. [doi: 10.1523/JNEUROSCI.0273-18.2018] [Medline: 30037830]
12. Papalambros NA, Santostasi G, Malkani RG, et al. Acoustic enhancement of sleep slow oscillations and concomitant memory improvement in older adults. Front Hum Neurosci 2017;11:109. [doi: 10.3389/fnhum.2017.00109] [Medline: 28337134]
13. Gui WJ, Li HJ, Guo YH, Peng P, Lei X, Yu J. Age-related differences in sleep-based memory consolidation: a meta-analysis. Neuropsychologia 2017 Mar;97:46-55. [doi: 10.1016/j.neuropsychologia.2017.02.001] [Medline: 28161367]
14. Harrison EI, Roth RH, Lobo JM, et al. Sleep time and efficiency in patients undergoing laboratory-based polysomnography. J Clin Sleep Med 2021 Aug 1;17(8):1591-1598. [doi: 10.5664/jcsm.9252] [Medline: 33739259]
15. Mace RA, Mattos MK, Vranceanu AM. Older adults can use technology: why healthcare professionals must overcome ageism in digital health. Transl Behav Med 2022 Dec 30;12(12):1102-1105. [doi: 10.1093/tbm/ibac070] [Medline: 36073770]
16. LaMonica HM, Milton A, Braunstein K, et al. Technology-enabled solutions for Australian mental health services reform: impact evaluation. JMIR Form Res 2020 Nov 19;4(11):e18759. [doi: 10.2196/18759] [Medline: 33211024]
17. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556. [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
18. Wilson SA, Byrne P, Rodgers SE, Maden M. A systematic review of smartphone and tablet use by older adults with and without cognitive impairment. Innov Aging 2022;6(2):igac002. [doi: 10.1093/geroni/igac002] [Medline: 35243008]
19. Andrews NE, Ireland D, Vijayakumar P, et al. Acceptability of a pain history assessment and education chatbot (Dolores) across age groups in populations with chronic pain: development and pilot testing. JMIR Form Res 2023 Oct 6;7:e47267. [doi: 10.2196/47267] [Medline: 37801342]
20. Piau A, Crissey R, Brechemier D, Balardy L, Nourhashemi F. A smartphone chatbot application to optimize monitoring of older patients with cancer. Int J Med Inform 2019 Aug;128:18-23. [doi: 10.1016/j.ijmedinf.2019.05.013] [Medline: 31160007]
21. Ireland D, Lam A, Bradford D, Naismith S. A new role for chatbots: automation of a sleep-dependent memory task. Stud Health Technol Inform 2024 Jan 25;310:519-523. [doi: 10.3233/SHTI231019] [Medline: 38269863]
22. Spencer R, Pryce JM, Walsh J. Philosophical approaches to qualitative research. In: The Oxford Handbook of Qualitative Research: Oxford University Press; 2014:81-98.
23. Duffy SL, Lagopoulos J, Hickie IB, et al. Glutathione relates to neuropsychological functioning in mild cognitive impairment. Alzheimers Dement 2014 Jan;10(1):67-75. [doi: 10.1016/j.jalz.2013.01.005] [Medline: 23688577]
24. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
25. Roque NA, Boot WR. A new tool for assessing mobile device proficiency in older adults: the Mobile Device Proficiency Questionnaire. J Appl Gerontol 2018 Feb;37(2):131-156. [doi: 10.1177/0733464816642582] [Medline: 27255686]
26. Madero EN, Anderson J, Bott NT, et al. Environmental distractions during unsupervised remote digital cognitive assessment. J Prev Alzheimers Dis 2021;8(3):263-266. [doi: 10.14283/jpad.2021.9] [Medline: 34101782]
27. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
28. Tong A, Sainsbury P, Craig J. COnsolidated criteria for REporting Qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19(6):349-357. [doi: 10.1093/intqhc/mzm042]
29. Pirhonen J, Lolich L, Tuominen K, Jolanki O, Timonen V. "These devices have not been made for older people's needs"—older adults' perceptions of digital technologies in Finland and Ireland. Technol Soc 2020 Aug;62:101287. [doi: 10.1016/j.techsoc.2020.101287]
30. Holthe T, Halvorsrud L, Karterud D, Hoel KA, Lund A. Usability and acceptability of technology for community-dwelling older adults with mild cognitive impairment and dementia: a systematic literature review. Clin Interv Aging 2018;13:863-886. [doi: 10.2147/CIA.S154717] [Medline: 29765211]
31. Newell J, Mairesse O, Verbanck P, Neu D. Is a one-night stay in the lab really enough to conclude? First-night effect and night-to-night variability in polysomnographic recordings among different clinical population samples. Psychiatry Res 2012 Dec 30;200(2-3):795-801. [doi: 10.1016/j.psychres.2012.07.045] [Medline: 22901399]
32. Werner L, Huang G, Pitts BJ. Smart speech systems: a focus group study on older adult user and non-user perceptions of speech interfaces. Int J Hum-Comput Interact 2023 Mar 16;39(5):1149-1161. [doi: 10.1080/10447318.2022.2050541]
33. Stawarz K, Cox A, Blandford A. Don't forget your pill!: designing effective medication reminder apps that support users' daily routines. Presented at: CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Apr 26 to May 1, 2014; Toronto, Ontario, Canada. [doi: 10.1145/2556288.2557079]
34. Thompson HJ, Demiris G, Rue T, et al. A holistic approach to assess older adults' wellness using e-health technologies. Telemed J E Health 2011 Dec;17(10):794-800. [doi: 10.1089/tmj.2011.0059] [Medline: 22011052]

XSL•FO
RenderX

Abbreviations

AMDPQ: Abbreviated Mobile Device Proficiency Questionnaire
COREQ: Consolidated Criteria for Reporting Qualitative Research
HBA: Healthy Brain Ageing
MCI: mild cognitive impairment
SDM : sleep-dependent memory

Edited by M Mardini; submitted 29.10.24; peer-reviewed by E Okoyeocha, M Velazquez, RS Gomaa Mahmoud; revised version received 12.02.25; accepted 18.03.25; published 28.04.25.

Please cite as:
Lam A, Simonetti S, D'Rozario A, Ireland D, Bradford D, Fripp J, Naismith SL
Perceptions of the Use of Mobile Apps to Assess Sleep-Dependent Memory in Older Adults With Subjective and Objective Cognitive Impairment: Focus Group Approach
JMIR Aging 2025;8:e68147
URL: https://aging.jmir.org/2025/1/e68147
doi:10.2196/68147

© Aaron Lam, Simone Simonetti, Angela D'Rozario, David Ireland, DanaKai Bradford, Jurgen Fripp, Sharon L Naismith. Originally published in JMIR Aging (https://aging.jmir.org), 28.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Relationship Between Within-Session Digital Motor Skill Acquisition and Alzheimer Disease Risk Factors Among the MindCrowd Cohort: Cross-Sectional Descriptive Study

Andrew Hooyman[1], PhD; Matt J Huentelman[2], PhD; Matt De Both[2], BS; Lee Ryan[3], PhD; Kevin Duff[4], PhD; Sydney Y Schaefer[5], PhD

[1]Department of Physical Therapy, Chapman University, 9401 Jeronimo Rd, Irvine, CA, United States

[2]Division of Early Detection and Prevention, The Translational Genomics Research Institute, Phoenix, AZ, United States

[3]Department of Psychology, The University of Arizona, Tucson, AZ, United States

[4]Department of Neurology, Oregon Health and Science University, Portland, OR, United States

[5]School of Biological Health Systems Engineering, Arizona State University, Tempe, AZ, United States

Corresponding Author:
Andrew Hooyman, PhD
Department of Physical Therapy, Chapman University, 9401 Jeronimo Rd, Irvine, CA, United States

Abstract

Background: Previous research has shown that in-lab motor skill acquisition (supervised by an experimenter) is sensitive to biomarkers of Alzheimer disease (AD). However, remote unsupervised screening of AD risk through a skill-based task via the web has the potential to sample a wider and more diverse pool of individuals at scale.

Objective: The purpose of this study was to examine a web-based motor skill game ("Super G") and its sensitivity to risk factors of AD (eg, age, sex, *APOE* ε4 carrier status, and verbal learning deficits).

Methods: Emails were sent to 662 previous MindCrowd participants who had agreed to be contacted for future research and have their *APOE* ε4 carrier status recorded and those who were at least 45 years of age or older. Participants who chose to participate were redirected to the Super G site where they completed the Super G task using their personal computer remotely and unsupervised. Once completed, different Super G variables were derived. Linear and logistic multivariable regression was used to examine the relationship between available AD risk factors (age, sex, *APOE* ε4 carrier status, and verbal learning) and distinct Super G performance metrics.

Results: Fifty-four participants (~8% response rate) from the MindCrowd web-based cohort (mean age of 62.39 years; 39 females; and 23 *APOE* ε4 carriers) completed 75 trials of Super G. Results show that Super G performance was significantly associated with each of the targeted risk factors. Specifically, slower Super G response time was associated with being an *APOE* ε4 carrier (odds ratio 0.12, 95% CI 0.02-0.44; *P*=.006), greater Super G time in target (TinT) was associated with being male (odds ratio 32.03, 95% CI 3.74-1192,61; *P*=.01), and lower Super G TinT was associated with greater age (β −3.97, 95% CI −6.64 to −1.30; *P*=.005). Furthermore, a sex-by-TinT interaction demonstrated a differential relationship between Super G TinT and verbal learning depending on sex ($\beta_{male:TinT}$ 6.77, 95% CI 0.34-13.19; *P*=.04).

Conclusions: This experiment demonstrated that this web-based game, Super G, has the potential to be a skill-based digital biomarker for screening of AD risk on a large scale with relatively limited resources.

(JMIR Aging 2025;8:e67298) doi:10.2196/67298

KEYWORDS

digital health technology; web-based assessment; aging; APOE; motor skills; sensitivity; risk factors; adults; older adults

Introduction

Since the number of cases of Alzheimer disease (AD) is expected to double in the next 2 decades [1], there is an urgent need for widespread screening of older adults for their individual AD risk profile, which has implications for clinical care and research. Current options, such as positron emission tomography for measuring tau and beta amyloid pathology, tend to be expensive and invasive and require advanced imaging facilities [2]. Blood-based biomarkers may make screening for AD more affordable and less invasive, but their validity and standardization are still being established at this time [3,4]. Cognitive screening tests can indicate deficits, but these largely rely upon in-person administration from a trained clinician and may have reduced sensitivity in identifying individuals at high risk of AD in the earliest stages [5]. Thus, it is important to

XSL•FO
RenderX

identify measures that are sensitive to AD risk and yet can also be delivered accurately, easily, and directly to patients and prospective participants in clinical trials and research studies.

Digital biomarkers have great potential to meet the need for accessible and remote testing of AD risk. Broadly speaking, digital biomarkers are measurable indicators of health or disease collected from a digital device or through digital means. Some digital biomarkers of AD include finger tapping [5,6], repeated cognitive assessment on a Wi-Fi–enabled device [7], recorded speech [8], and digital clock drawing [9], which have all been shown to be sensitive to cognitive impairment. These measures have also been associated with disease status, differentiating between cognitively intact versus mild cognitive impairment (MCI) and MCI versus dementia to some degree. However, these examples have overall low sensitivity to risk factors of AD, such as age [10], sex [11], apolipoprotein-E (*APOE*) ε4 carrier status [12,13], and brain amyloid [9]. Considering this, another potential digital biomarker of AD could be motor skill. Motor skill acquisition is the within-session improvement in a motor skill as a function of practice [14]. Previous in-lab studies have associated motor skill deficits with *APOE* carrier status, hippocampal atrophy, functional decline, and amyloid deposition among people diagnosed with amnestic MCI [15-18]. This task can also be collected remotely [19,20], which would allow for a wider and more diverse sample of individuals at potential risk for AD and can more easily facilitate longitudinal testing as desired.

This study developed a web-based tool for assessing motor skill performance called Super G [21], which can be reliably played unsupervised on the web [20] regardless of device type and without downloading any app. Specifically, the objective of this study was to examine whether Super G performance was individually related to known risk factors of AD (eg, age, sex, *APOE* ε4 carrier status, and verbal learning). Based on prior motor skill studies in AD, it was hypothesized that Super G performance would be negatively associated with each risk factor.

Methods

Study Design

This was a cross-sectional descriptive study that examined within-session performance characteristics from a web-based motor skill task (Super G) and their association with AD risk factors among adults recruited from the MindCrowd study.

Participants

Participants were recruited in May of 2021 through MindCrowd, a web-based research study launched in 2013 to crowdsource demographic, medical history, lifestyle, and cognitive data to identify risk factors of AD [22,23]. Emails were sent through MindCrowd to a subset of 662 individuals who met our inclusion or exclusion criteria and were older than 45 years, who had previously provided a dried blood spot or saliva sample for *APOE* genotyping (see section “AD Risk Factors and Other Participant Characteristics From MindCrowd” for details) and had provided consent to be contacted for future studies. A hyperlink was included in the email that directed individuals to the Super G game website, on which participants digitally provided consent (approved by the Arizona State University institutional review board study no. 13081). Of the 662 individuals emailed, 54 participants (age: mean 62.39, SD 7.4 years; female: n=39) appropriately registered and completed all 75 trials of the game, equating to an 8.1% response rate. The mean (SD) time between MindCrowd data collection (specifically verbal learning, see section “AD Risk Factors and Other Participant Characteristics From MindCrowd” for details) and Super G data collection was 5.9 (1.4) years.

Super G

The Super G game was developed in Unity 5.3.1 and is hosted on Hostinger. Thus, participants were not required to download an app or program to their device. Super G was developed as a gamified version of a seminal motor skill paradigm [24] and has been validated against the original version [21]. The goal of Super G is to help an astronaut explore as many planets as possible within the game’s solar system. There are 16 planets to visit but only 75 attempts to reach them all. Participants use the left and right arrow keys on their keyboard to move the astronaut onto a planet. However, the game uses a rate control mechanism that may not be immediately apparent to participants [21]. Specifically, pressing the right arrow key applies a constant positive force to move the astronaut toward the planet, while pressing the left arrow key applies an equal negative force away from the planet. Since the virtual environment lacks gravity or drag, any force applied will result in a constant velocity until a negative force is applied to slow it down. Thus, participants must learn to apply negative force at the right time and for the right duration to land the astronaut on each planet.

Each trial in Super G begins with the astronaut positioned on the left side of the screen on the initial start planet. The target planet is located on the opposite side of the screen to the right (Figure 1A). The trial lasts 4.5 seconds, but the astronaut cannot leave the start planet until 1.5 seconds have elapsed, as indicated by the disappearance of the blue atmosphere around the start planet (Figure 1B). If the astronaut leaves too early, the trial resets and the astronaut is returned to the initial position on the start planet. Once the blue atmosphere disappears, the participant has 3 seconds to land the astronaut on the target planet. To achieve a successful landing, the astronaut must stay within the boundary of the target planet for 1 continuous second. If successful, a reward tone plays, and fireworks erupt from the target planet (Figure 1C). After a successful landing, Super G repositions at the start position, and the previously landed planet becomes the start planet, with a new target planet appearing in its place (Figure 1D). In the event of a failed landing, the planets for the next trial remain the same, and astronaut reappears at the start position. All 75 trials are completed within a single session for a total session length of approximately 6 minutes.

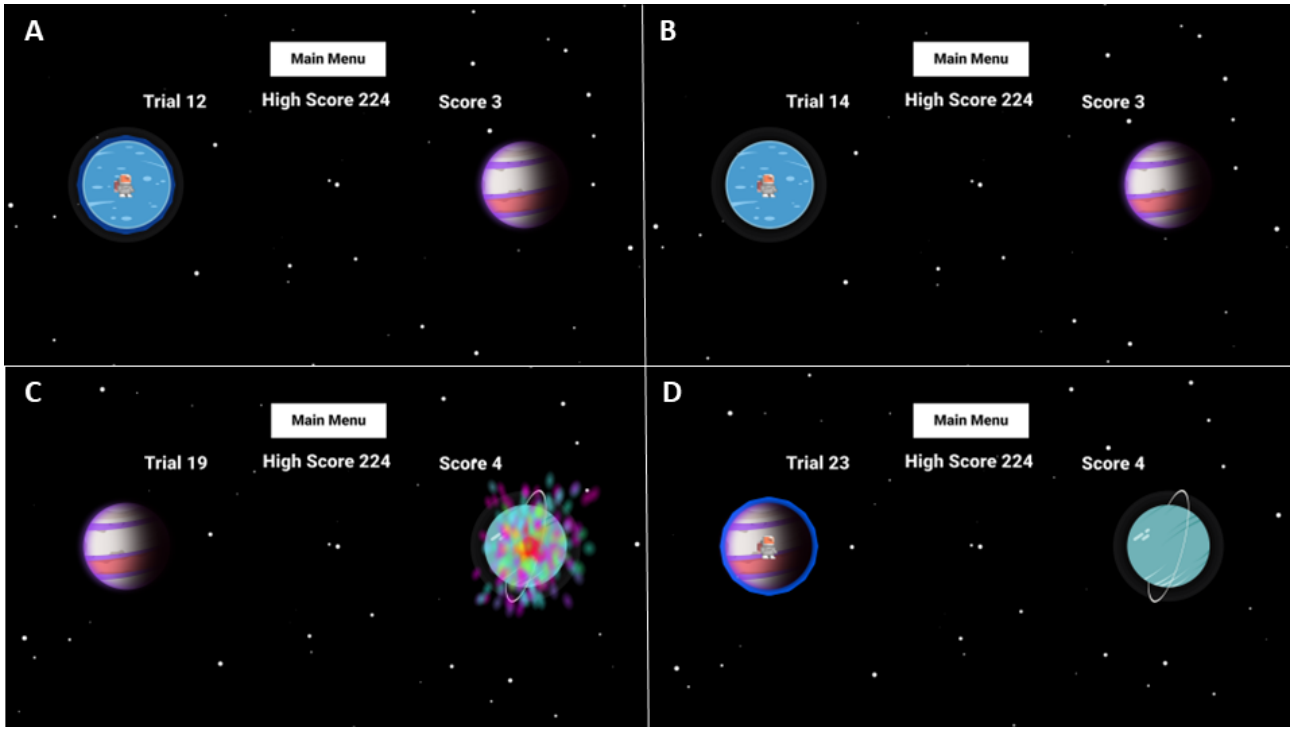
Cursor position and key press data from each trial of Super G were collected at 100 Hz, along with high scores, time and date of each trial. From these data, four performance variables were calculated. First, scaling ratio (SR) represents the ratio of negative force (duration of left button press) applied to Super G over positive force (duration of right button press), where values equal to 1 indicate equal scaling of forces. Second, time

XSL•FO
RenderX

of reversal (TR) is a measure of how well participants execute the timing of their movements by identifying when the left arrow key press occurs during the trial, where higher values are reversals that occur later in the trial. Successful trials require equal scaling of forces (ie, SR close to or equal to 1) and a reversal timed at the midway point of the trial, although we note that there can be trade-offs between the SR and TR to still allow for some degree of success in the task. Third, total time in the target (TinT) planet is the time that the cursor stayed within the target planet, with higher values representing better performance. Fourth, response time (RT) is a measure of how soon the cursor left the home planet after the disappearance of the start planet's atmosphere, and lower values equate to a faster RT.

Although these performance metrics do strongly correlate with one another, each represents distinct phases of individual motor skill acquisition, as it is possible to have a fast RT with a low TinT if either timing of cursor reversal or scaling of forces is not well executed. Overall, average TinT across all 75 trials is the primary measure of performance since it directly represents the task goal and performance of the task as it relates to execution of SR and TR. Average RT across all 75 trials was considered as a secondary measure of performance because it represents the earlier stage of skill acquisition as participants first need to anticipate when to exit the start planet prior to execution of their acquired movement strategy. Average TinT and RT across the 75 trials are used as the Super G metrics of individual performance due to our prior work [21], which demonstrated that average TinT and RT well describe individual within-session change in performance and delayed retention.

Figure 1. Each panel represents the different phases a participant may experience during Super G play. (A) The astronaut, Super G, spawns on the start planet at the beginning of each trial. There is a blue atmosphere around the planet that signals to the participant that Super G cannot leave yet. (B) Once the blue atmosphere disappears, then Super G can leave the start planet and attempt to land on the target planet. (C) If Super G successfully stays within the target planet boundary for 1 continuous second, then a reward tone plays and fireworks erupt from the planet. (D) Subsequent trials then render the previous target planet as the new start planet and a new target planet is put in place.



AD Risk Factors and Other Participant Characteristics From MindCrowd

Super G data were harmonized with MindCrowd through merging of hashed email addresses linked to MindCrowd data and used as the login for the Super G game. Super G variables were then merged with participant age, sex, level of education, verbal learning (measured via paired associates learning [PAL]) score [25], simple visual reaction time (svRT), and *APOE* genotype data, which were available in MindCrowd. The PAL is a verbal learning task that measures the ability to remember the associations between different word pairs. Specifically, participants are visually presented with 12 word pairs, with each word pair presented separately and at 2-second intervals. Participants were then presented with the first word of each pair and then used their keyboard to type the missing word. This procedure was repeated for 2 additional trials. The maximum score on the PAL is 36 (12 words across 3 trials). The svRT task measured the median reaction time across 6 trials, in which participants had to press any keyboard button as soon as a target stimulus appeared on their screen, with reaction time for each trial recorded in milliseconds. Median svRT was used due to the skewness of the distribution of reaction time in each participant's individual data (skewness svRT median 1.2 vs skewness svRT mean 4.1). The svRT test was used in this study as a control variable only. Participants were classified as *APOE* ε4 carriers or noncarriers. Carriers were defined as individuals who had either 1 or 2 copies of the ε4 allele, and noncarriers were defined as individuals with 0 copies of the ε4 allele. This was based on prior genotyping from biospecimen collection via

XSL•FO
RenderX

self-administered saliva or dried blood spot kits that were mailed to the participants by MindCrowd and then processed by the Translational Genomics Research Institute. Details regarding biospecimen collection and genotyping can be found here [26]. Complete visualization of *APOE* ε4 carrier status with svRT, PAL, and age can be viewed in Figure S1 in Multimedia Appendix 1.

Statistical Analysis

Wilcoxon rank sum tests were used due to a difference in sample sizes between *APOE* ε4 carriers (n=23) and noncarriers (n=31), since uneven sample sizes may lead to inaccurate and disproportionate estimates of variance of each group and violate assumptions of parametric testing. This determined whether there were differences in age, PAL score, svRT, and Super G performance between the 2 groups. Chi-square tests were used to determine whether there was a difference in proportion based on sex, race, education, and ethnicity between carriers and noncarriers. To provide broader context of the verbal learning (PAL) and reaction time (svRT) of this sample, percentile scores were calculated for each participant adjusted for their age, sex, and level of education. This allowed for better contextualization of the relative performance on each measure between carriers and noncarriers relative to the entire MindCrowd cohort (ie, if participants in this study are under- or overperformers compared with what would be expected to a random sample across the entire cohort).

Separate linear and logistic multivariable regression analyses, depending on outcome variable type, were used to test the relationship between Super G performance and the dependent variables, that is, AD risk factors. Specifically, different models were constructed to identify which Super performance metrics are related to specific AD risk factors, while controlling for possible confounding effects of the other AD risk factors and participant characteristics. Multivariable logistic regression was used for the dependent variable of *APOE* ε4 carrier status (where carriers were coded as “true” and noncarriers as “false”), along with the primary and secondary measures of Super G performance (TinT and RT), and control variables of age, sex, PAL score, hour of day Super G played, level of education, and svRT. This approach controlled for the other factors and was repeated with sex as the dependent variable (whereby male was coded as “true” and female as “false”) while switching *APOE* ε4 carrier status to a control variable. Multivariable regression was used when the dependent variable was PAL score, with a primary independent variable of mean Super G performance (TinT, RT, SR, or TR), and control variables of age, sex, *APOE* ε4 carrier status, hour of day Super G played, level of education, and svRT. The same approach was repeated with age as the dependent variable while switching PAL score to a control variable. In addition, to control for the potential delay between initial PAL scores and Super G measurement, PAL-adjusted scores were also generated and analyzed. Individual PAL scores were adjusted based on the amount of time between their PAL and Super G measurement, given previously reported expected decline in PAL based on age reported by Talboom and colleagues [26]. This resulted in an expected 0.2-point decline in PAL for every year between the initial PAL and current Super G measurement. The formulation of each model can be viewed in the following equations:

1. *APOE* ε4 carrier status ~ Super G Performance (TinT or RT) + age +sex + hour played + education +svRT + PAL (Logistic regression)
2. Sex ~ Super G Performance (TinT or RT) + age + hour played + education + svRT + PAL + *APOE* ε4 carrier status (Logistic regression)
3. Age ~ Super G Performance (TinT or RT) + sex + hour played + education + svRT + PAL + *APOE* ε4 carrier status (Linear regression)
4. PAL ~ Super G Performance (TinT or RT) + age + sex + hour played + education + svRT + *APOE* ε4 carrier status (Linear regression)
5. PAL adjusted ~ Super G Performance (TinT or RT) + age + sex + hour played + education + svRT + *APOE* ε4 carrier status (Linear regression)

Participants could play Super G at any time throughout the day, the variable of hour played (measured with a 24-hour clock rather than a 12-hour clock) was transformed using a cosine function to ensure that adjacent hours 0 and 23 were close together. This is visualized in Figure S2 in Multimedia Appendix 1. All numeric variables were standardized (age, svRT, PAL, hour played, and Super G Performance) to be centered at their respective mean and divided by their respective SD. Thus, a 1-unit change in the results of these variables with respect to all reported odds ratios (ORs) and beta coefficients represents a 1 SD change in either the outcome or the performance metric. This allowed for better relative comparison between all independent variables within each model. To detect multicollinearity, the variance inflation factor was calculated, and any variable with a variance inflation factor of >5 was removed. Outliers were identified using Cook’s distance and removed if their distance was >1.

Ethical Considerations

The study protocol was reviewed and approved by the office of Research Integrity and Assurance at Arizona State University (approval number STUDY00013081). To participate, participants needed to provide digital informed consent. All data collected from this experiment were deidentified for privacy and confidentiality. Participants were not compensated for their time in this study.

Results

Participant Characteristics

Overall, there were 23 *APOE* ε4 carriers (20 heterozygotes and 3 homozygotes) and 31 noncarriers. Between-group comparisons (carriers vs noncarriers) using the Wilcoxon rank sum test demonstrated that groups did differ by age (W=471; *P*=.04), with noncarriers being an average of 4.3 years older than carriers (carriers=59.9 years, noncarriers=64.2 years). For all other control variables, there were no observed group differences (Table 1).

XSL•FO
RenderX

Table . Genetic, demographic, cognitive, motor, and Super G performance data between *APOE* ε4 carries and noncarriers.

	APOE ε4 carriers	*APOE* ε4 noncarriers	Test statistic (*df*)	*P* value
Participants, n	23	31	—[a]	—
APOE alleles (X/X), n			—	—
2/2	—	0		
2/3	—	2		
3/3	—	29		
2/4	4	—		
3/4	16	—		
4/4	3	—		
Age (years), mean (SD)	59.9 (7.4)	64.2 (6.9)	*W*=471 (—)	.045[b]
Sex (male/female)	6/17	9/22	χ^2=0 (1)	>.99
Race			χ^2=0 (1)	>.99
White	23	31		
Ethnicity			χ^2=0 (1)	>.99
Not Latinx	23	31		
Education, n			χ^2=2.5 (3)	.47
High school diploma	1	0		
Some college	5	4		
Four-year degree	9	12		
Postgraduate degree	8	15		
PAL[c] score, mean (SD)	23.4 (8.4), 66th percentile	19.2 (8.5), 54th percentile	*W*=254 (—)	.07
median svRT[d] (ms), mean (SD)	389.7 (75), 71st percentile	418 (79), 65th percentile	*W*=451 (—)	.10
Hour of day played, mean (SD)[e]	13.1 (4.6)	14.9 (4.1)	*W*=441 (—)	.14
Hour of day played (cosine), mean (SD)	0.52 (0.6)	.45 (6)	*W*=394 (—)	.49
Super G time in target (ms), mean (SD)	653.9 (436.6)	469.4 (287.9)	*W*=279 (—)	.18
Super G response time (ms), mean (SD)	1614.8 (305.9)	1874.5 (307.4)	*W*=502 (—)	.01[b]
Super G time of reversal (ms), mean (SD)	2000.6 (592.4)	2060.3 (836.4)	*W*=432 (—)	.19
Super G scaling ratio, mean (SD)	0.66 (0.3)	0.64 (0.3)	*W*=343 (—)	.82

[a]N/A: not applicable.

[b]Statistical significance (*P*<.05).

[c]PAL: paired associates learning.

[d]svRT: simple visual reaction time.

[e]Based on 24-hour clock.

Relationship Between Super G Performance and AD Risk Factors

Within the *APOE* ε4 carrier status logistic regression (model 1) there was a significant association between *APOE* ε4 carrier status with Super G RT (OR 0.12, 95% CI 0.02-0.44; *P*=.006) (Figure 2A). Thus, a participant with a Super G RT that was 1 SD below the mean would have an 88% increase in odds of being an *APOE* ε4 noncarrier than carrier. Visual comparison between all Super G performance metrics and *APOE* ε4 can be visualized in Figure S3 in Multimedia Appendix 1. All other control variables, including PAL (*P*=.62) and svRT (*P*=.29),

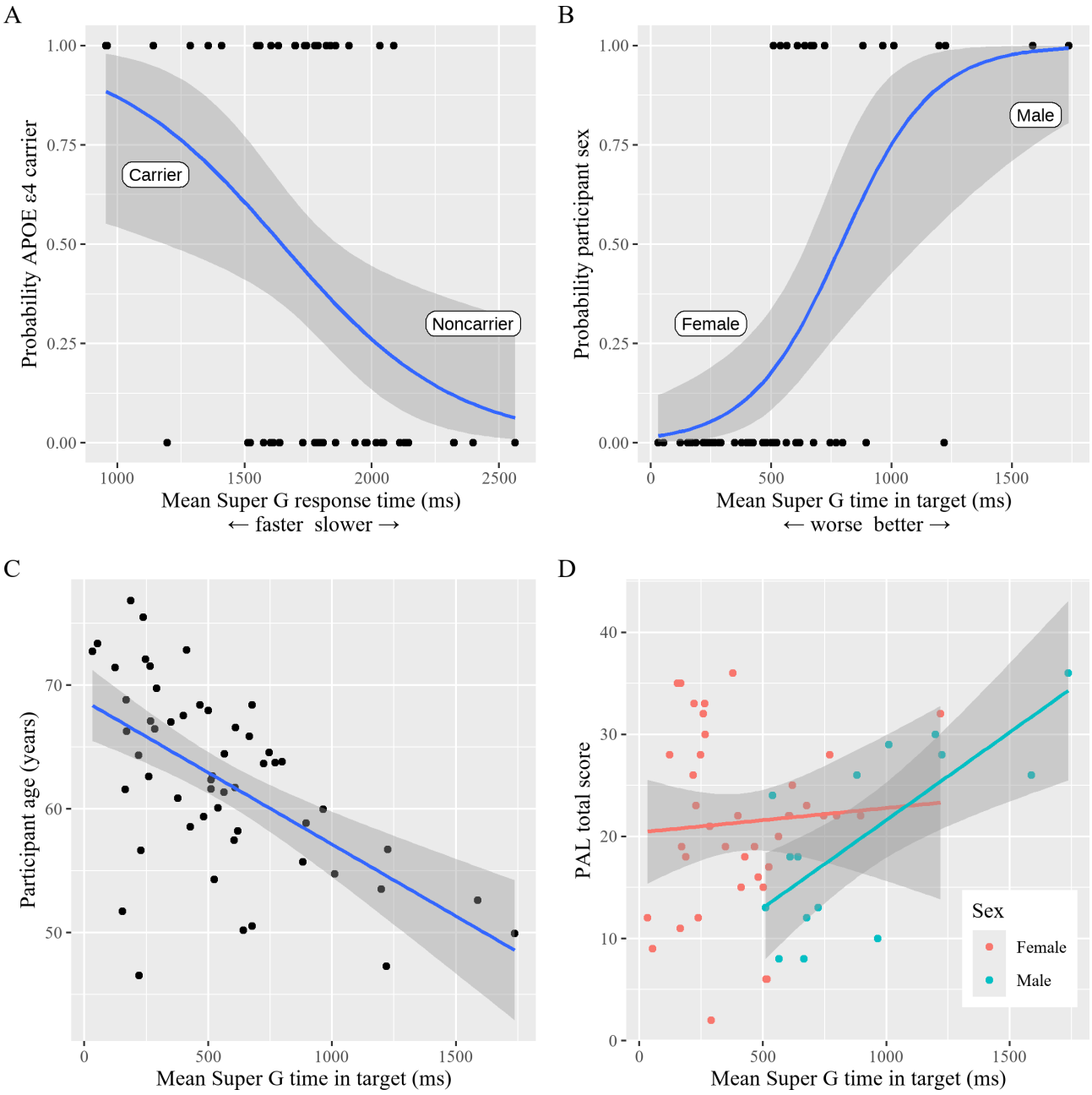
XSL•FO
RenderX

were not significantly related to *APOE* ε4 carrier status. Within the participant age linear regression (model 3), Super G TinT was significantly associated with age (β −3.97, 95% CI −6.64 to −1.30; *P*=.005) (Figure 2C) and within the participant sex logistic regression (model 2), TinT was also associated with sex (OR 32.03, 95% CI 3.74-1192.61; *P*=.01) (Figure 2B), with lower TinT values (poorer performance) associated with being older and being female, respectively. This is consistent with our earlier work in other cohorts [27]. Visual comparison between all Super G performance metrics and sex can be visualized in Figure S4 in Multimedia Appendix 1.

Given that previous research in MindCrowd has demonstrated a main effect of sex on verbal learning (as measured with the PAL test) [26,28], and that TinT is also strongly linked with participant sex (Figure 2B), within the PAL linear regression (model 4), a sex-by-TinT interaction was also included to best model PAL [26]. Results from the PAL model demonstrated that age was associated with PAL (β −3.37, 95% CI −6.34 to −0.4; *P*=.03), indicating that older age was associated with lower PAL scores. There was also a main effect of participant sex (β_{Male} −7.73, 95% CI −14.98 to −0.45; *P*=.04). This result indicates that male participants scored an average of 7 points lower on the PAL than females. There was a significant sex-by-TinT interaction ($\beta_{Male:TinT}$ 6.77, 95% CI 0.34-13.19; *P*=.04). This result indicates that males with a mean TinT 1 SD greater than the group mean are associated with an increase of 6.77 points on the PAL compared with females with the same TinT performance (Figure 2D). Furthermore, in the PAL adjusted regression (model 5), where PAL scores are modified based on the time between PAL and Super G measures, the results are nearly identical to the raw scores (β_{age} −3.41, 95% CI −6.38 to −0.44; *P*=.03; β_{Male} −7.57, 95% CI −14.81 to −0.33; *P*=.04; $\beta_{Male:TinT}$ 6.72, 95% CI 0.31-13.14; *P*=.04; Figure S5 in Multimedia Appendix 1). The results of this adjusted analysis provide evidence that the delay between PAL and Super G measures does not significantly impact the observed relationship between PAL and TinT given participant sex.

XSL•FO
RenderX

Figure 2. (A) Mean Super G response time (ie, time at which Super G exits the start planet once the blue atmosphere disappears) between *APOE* ε4 noncarriers and carriers. Y-axis represents the probability a participant is classified as a carrier or noncarrier based on their mean response time. The blue line represents the binomial relationship between response time and carrier status with faster response time more associated with being a carrier and slower response time more associated with being a noncarrier. The gray ribbon represents the 95% CI of the estimated probability of binomial relationship. (B) Mean Super G time in target between female and male participants. Y-axis represents the probability a participant is classified as a male or female based on their mean time in target. The blue line represents the binomial relationship between time in target and sex with better time in target associated with being a male and worse time in target associated with being a female. The gray ribbon represents the 95% CI of the estimated probability of binomial relationship. (C) Linear relationship between participant age in years and their mean Super G time in target. The blue line represents the least squares line fitted between the variable and the gray ribbon about the line represents the 95% CI. (D)The relationship between mean Super G time in target to individual PAL total score stratified by sex (male in blue and female in pink). Given known sex differences between males and females on PAL and observed sex differences on mean Super time in target, such a stratification by sex was necessary to control for potential confounding of sex on the Super G to PAL relationship. The blue line represents the line of least squares for males and the pink line represents the line of least squares for females. The gray ribbon about each line represents the 95% CI. PAL: paired associates learning.



Discussion

This study tested whether motor skill acquisition, as assessed by the Super G task, was sensitive to known risk factors of AD, namely age, sex, *APOE* ε4 carrier status, and verbal learning. Results showed that better Super G performance was directly and independently associated with each of these risk factors, suggesting that Super G may be a sensitive digital biomarker of AD risk that can be remotely collected. Interestingly, TinT and RT were related to different AD risk factors (sex, age, and verbal learning vs *APOE* ε4 carrier status, respectively), likely reflecting how different aspects of motor skill map onto different AD risk factors. Our prior work showed that TinT is the product of optimal execution of SR and TR and best characterizes overall

XSL•FO
RenderX

skill compared with the other Super G variables [21], and it was related to age, sex, and verbal learning. However, RT is a performance characteristic of early skill acquisition, as it measures an individual's ability to anticipate when to initialize movement at the beginning of the trial, which was related to *APOE* ε4 carrier status. Although *APOE* ε4 carriers unexpectedly outperformed noncarriers in terms of early skill acquisition (ie, better RT performance), this opens up the possibility of using Super G as a prognostic enrichment strategy for enriching AD-focused cohorts with ε4 carriers [29].

The observed *APOE* ε4 benefit in this study is contrary to previous work where *APOE* ε4 carrier status leads to worse performance on memory tests in carriers compared with noncarriers [30-32]. However, the role *APOE* ε4 in aging independent of AD pathology may be significant [33], as a small but growing body of evidence in both cognitively unimpaired humans and rodents shows that visual working memory and learning is better among *APOE* ε4 carriers than noncarriers [34-36]. Although *APOE* ε4 is the primary genetic risk factor for AD [13], evidence suggests a possible benefit, or compensatory behavior [37], of learning at an earlier age while leading to impairments in later life [38,39]. Previous research has shown that cognitively unimpaired individuals who are *APOE* ε4 carriers have greater gray matter volume in frontal regions, can better allocate cognitive control, and possess better visual working memory and learning than *APOE* ε4 noncarriers [34,40,41]. Furthermore, task-based neuroimaging studies have associated better performance on visual working memory tasks among carriers with greater activation in frontal and parietal regions [34,36,42]. This leads to the hypothesis that increased frontal activation is a compensatory mechanism that modifies behavior in AD, given that frontal brain regions are relatively spared in AD [43,44]. Although these data do not provide direct support for the *APOE* ε4 compensatory mechanisms, it may be plausible that Super G may be able to measure a suspected *APOE* ε4 benefit. This interpretation is further supported by the fact that in-lab motor skill studies have associated greater skill with higher white matter integrity of frontoparietal tracts [45,46] in cognitively unimpaired older adults, which could serve as a candidate neural substrate for such a compensatory mechanism as proposed previously, but further study would be needed to confirm.

Prior research that has investigated sex differences in motor skill learning and performance has demonstrated a preferential advantage for males compared with females [47]. Similarly, this study revealed that males tended to perform better than females on Super G. The estimated effect size of mean time in target between males and females was very large, with Cohen *d*=1.66. This result is consistent with previous research using Super G, which also found large sex differences in performance [27]. Several factors may explain the observed sex differences in Super G performance. One possible biological explanation [48] may be due to early tau deposition which has been shown to be elevated in females compared with males, consistent with the higher risk for developing AD among females. Thus, a remote and unsupervised motor skill task that is sensitive to sex differences may aid in the detection of sex-specific changes in behavior due to disease in a scalable way. Moreover, there was an interaction between sex, verbal learning, and Super G performance, which suggests that motor skill in males (ie, Super G performance) may be linked to their verbal learning to a greater degree than in females. Larger sample sizes are needed to examine the interactions between sex and other behavioral variables in the context of AD [49,50].

Several limitations to this study should be acknowledged. First, the mean time between when the PAL and Super G was 5 years. Although this is a substantial delay between 2 variables of interest, there was no correlation between PAL score and this intertest interval. Furthermore, when PAL was adjusted for this delay, there was no change in the reported relationship between Super G and PAL regardless of whether raw or adjusted PAL scores were used as the dependent variable (Figure S5 in Multimedia Appendix 1). Thus, the reported result between PAL and Super G appears robust, even with such a delay between measures. Second, the study sample was all non-Hispanic White and highly educated (>80% with at least a college degree), preventing any analysis of interactions between race or ethnicity and *APOE* ε4. Analyses that consider the interaction between race and carrier status are important, given that the link between the *APOE* ε4 allele and the AD is weaker among Black/African American and Hispanic/Latino individuals despite their increased risk of developing AD overall [12,51,52]. In addition, the study sample size was relatively limited due to our required inclusion criteria, particularly the existence of *APOE* genotyping data. As such, future Super G research will recruit larger and more diverse samples. We plan to expand the size of this work with future collaborations, and it is important to note that the MindCrowd cohort itself has increased in size, in racial and ethnic diversity, and in the number of individuals in the cohort who have *APOE* genotyping data, suggesting a path to addressing this limitation in the future. Third, we did not collect participant data on previous video game experience, which may be a contributing factor to the observed relationship between Super G performance and sex [53]. Fourth, there is no defined minimal clinically important difference or clinical cutoff score for the PAL, making it difficult to determine the meaningfulness of the observed relationship between PAL and Super G performance. For example, it cannot be determined whether the predicted change in PAL of 6.5 points among males, in relation to TinT performance, is representative of a meaningful increase or decrease in verbal learning. Finally, with only an 8.1% response rate in this study (54 participants completed all trials of Super G out of the 662 contacted), there may be limited generalizability to the broader MindCrowd cohort or the general public. However, this response rate is similar to other web-based AD-focused cohorts such as the Alzheimer's Prevention Trials Webstudy [54], although the participants in this study were not paid to participate (in contrast to the Alzheimer's Prevention Trials Webstudy), were not actively seeking care, and were emailed only once during the recruitment process. It is noted that these are the factors that could influence participation rates. Despite these limitations, this study establishes the proof of concept that Super G may be a feasible skill-based digital biomarker of individual AD risk.

XSL•FO
RenderX

Acknowledgments

The authors would like to thank the participants of this study without whom this research would not be possible. This work is supported by the National Institute on Aging F32 AG071110-01A1 (principal investigator [PI]: AH), Mueller Family Charitable Trust (PI: MJH), Arizona DHS in support of the Arizona Alzheimer's Consortium (PI: Reiman), Flinn Foundation (PI: MJH), and the National Institute on Aging U19AG065169 (PI: Barnes).

Data Availability

The data, analysis, and visualizations for this paper can be found at Open Science Framework [55].

Authors' Contributions

AH participated in the original draft (lead), formal analysis (lead), writing—review and editing (equal), software (lead), methodology (equal), conceptualization (equal), and data management (equal). MJH participated in writing—review and editing (equal), conceptualization (equal), data management (equal), and methodology (equal). MDB participated in data management (equal), software (supporting), and writing—review and editing (equal). LR contributed to writing—review and editing (equal). KD contributed to writing—review and editing (equal). SYS participated in writing—review and editing (equal), conceptualization (equal), data management (equal), and methodology (equal).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplemental figures for between-group comparison across covariates and outcome measures contrasted with different Super G outcomes.
[PDF File, 670 KB - aging_v8i1e67298_app1.pdf]

References

1. Matthews KA, Xu W, Gaglioti AH, et al. Racial and ethnic estimates of Alzheimer's disease and related dementias in the United States (2015-2060) in adults aged ≥65 years. Alzheimers Dement 2019 Jan;15(1):17-24. [doi: 10.1016/j.jalz.2018.06.3063] [Medline: 30243772]
2. Klunk WE. Amyloid imaging as a biomarker for cerebral β-amyloidosis and risk prediction for Alzheimer dementia. Neurobiol Aging 2011 Dec;32(Suppl 1):S20-S36. [doi: 10.1016/j.neurobiolaging.2011.09.006] [Medline: 22078170]
3. Nakamura A, Kaneko N, Villemagne VL, et al. High performance plasma amyloid-β biomarkers for Alzheimer's disease. Nature New Biol 2018 Feb 8;554(7691):249-254. [doi: 10.1038/nature25456] [Medline: 29420472]
4. Chong MS, Sahadevan S. Preclinical Alzheimer's disease: diagnosis and prediction of progression. Lancet Neurol 2005 Sep;4(9):576-579. [doi: 10.1016/S1474-4422(05)70168-X] [Medline: 16109364]
5. Dagum P. Digital biomarkers of cognitive function. NPJ Digit Med 2018;1:10. [doi: 10.1038/s41746-018-0018-4] [Medline: 31304295]
6. Omberg L, Chaibub Neto E, Perumal TM, et al. Remote smartphone monitoring of Parkinson's disease and individual response to therapy. Nat Biotechnol 2022 Apr;40(4):480-487. [doi: 10.1038/s41587-021-00974-9] [Medline: 34373643]
7. Papp KV, Jutten RJ, Soberanes D, et al. Early detection of amyloid-related changes in memory among cognitively unimpaired older adults with daily digital testing. Ann Neurol 2024 Mar;95(3):507-517. [doi: 10.1002/ana.26833] [Medline: 37991080]
8. Mueller KD, Van Hulle CA, Koscik RL, et al. Amyloid beta associations with connected speech in cognitively unimpaired adults. Alzheimers Dement (Amst) 2021;13(1):e12203. [doi: 10.1002/dad2.12203] [Medline: 34095435]
9. Rentz DM, Papp KV, Mayblyum DV, et al. Association of digital clock drawing with PET amyloid and tau pathology in normal older adults. Neurology (ECronicon) 2021 Apr 6;96(14):e1844-e1854. [doi: 10.1212/WNL.0000000000011697] [Medline: 33589537]
10. Gurland BJ, Wilder DE, Lantigua R, et al. Rates of dementia in three ethnoracial groups. Int J Geriatr Psychiatry 1999 Jun;14(6):481-493. [Medline: 10398359]
11. Podcasy JL, Epperson CN. Considering sex and gender in Alzheimer disease and other dementias. Dialogues Clin Neurosci 2016 Dec;18(4):437-446. [doi: 10.31887/DCNS.2016.18.4/cepperson] [Medline: 28179815]
12. Tang MX, Stern Y, Marder K, et al. The APOE-epsilon4 allele and the risk of Alzheimer disease among African Americans, whites, and Hispanics. JAMA 1998 Mar 11;279(10):751-755. [doi: 10.1001/jama.279.10.751] [Medline: 9508150]
13. Liu CC, Liu CC, Kanekiyo T, Xu H, Bu G. Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. Nat Rev Neurol 2013 Feb;9(2):106-118. [doi: 10.1038/nrneurol.2012.263] [Medline: 23296339]
14. Schmidt R, Lee T, Winstein C, Wulf G, Zelaznik H. Motor Control and Learning: A Behavioral Emphasis, 6th edition: Human Kinetics; 2019.

XSL•FO
RenderX

15. Rogojin A, Gorbet DJ, Hawkins KM, Sergio LE. Cognitive-motor integration performance is affected by sex, APOE status, and family history of dementia. J Alzheimers Dis 2019;71(2):685-701. [doi: 10.3233/JAD-190403]
16. Schaefer SY, Malek-Ahmadi M, Hooyman A, King JB, Duff K. Association between motor task performance and hippocampal atrophy across cognitively unimpaired, amnestic mild cognitive impairment, and Alzheimer's disease individuals. J Alzheimers Dis 2022;85(4):1411-1417. [doi: 10.3233/JAD-210665] [Medline: 34958015]
17. Schaefer SY, Hooyman A, Duff K. Using a timed motor task to predict one-year functional decline in amnestic mild cognitive impairment. J Alzheimers Dis 2020;77(1):53-58. [doi: 10.3233/JAD-200518] [Medline: 32651327]
18. Schaefer SY, Duff K, Hooyman A, Hoffman JM. Improving prediction of amyloid deposition in mild cognitive impairment with a timed motor task. Am J Alzheimers Dis Other Demen 2022;37:15333175211048262. [doi: 10.1177/15333175211048262] [Medline: 35200059]
19. Tsay JS, Lee A, Ivry RB, Avraham G. Moving outside the lab: the viability of conducting sensorimotor learning studies online. Neuron Behav Data Anal Theory 2021;5(3). [doi: 10.51628/001c.26985]
20. Hooyman A, Huentelman MJ, De Both M, Ryan L, Schaefer SY. Establishing the validity and reliability of an online motor learning game: applications for Alzheimer's disease research within MindCrowd. Games Health J 2023 Apr;12(2):132-139. [doi: 10.1089/g4h.2022.0042] [Medline: 36745382]
21. Hooyman A, Gordon J, Winstein C. Unique behavioral strategies in visuomotor learning: hope for the non-learner. Hum Mov Sci 2021 Oct;79:102858. [doi: 10.1016/j.humov.2021.102858] [Medline: 34392189]
22. Talboom JS, De Both MD, Naymik MA, et al. Two separate, large cohorts reveal potential modifiers of age-associated variation in visual reaction time performance. NPJ Aging Mech Dis 2021 Jul 1;7(1):14. [doi: 10.1038/s41514-021-00067-6] [Medline: 34210964]
23. Lewis CR, Talboom JS, De Both MD, et al. Smoking is associated with impaired verbal learning and memory performance in women more than men. Sci Rep 2021 May 13;11(1):10248. [doi: 10.1038/s41598-021-88923-z] [Medline: 33986309]
24. Brooks V, Hilperath F, Brooks M, Ross HG, Freund HJ. Learning "what" and "how" in a human motor task. Learn Mem 1995;2(5):225-242. [doi: 10.1101/lm.2.5.225]
25. Pike KE, Rowe CC, Moss SA, Savage G. Memory profiling with paired associate learning in Alzheimer's disease, mild cognitive impairment, and healthy aging. Neuropsychology 2008 Nov;22(6):718-728. [doi: 10.1037/a0013050] [Medline: 18999345]
26. Talboom JS, Håberg A, De Both MD, et al. Family history of Alzheimer's disease alters cognition and is modified by medical and genetic factors. Elife 2019 Jun 18;8:e46179. [doi: 10.7554/eLife.46179] [Medline: 31210642]
27. Hooyman A, Schaefer SY. Age and sex effects on Super G performance are consistent across internet devices. Int J Serious Games 2023 Jun 7;10(2):25-36. [doi: 10.17083/ijsg.v10i2.598] [Medline: 37846217]
28. Kaushanskaya M, Marian V, Yoo J. Gender differences in adult word learning. Acta Psychol (Amst) 2011 May;137(1):24-35. [doi: 10.1016/j.actpsy.2011.02.002] [Medline: 21392726]
29. Cummings J, Apostolova L, Rabinovici GD, et al. Lecanemab: appropriate use recommendations. J Prev Alzheimers Dis 2023;10(3):362-377. [doi: 10.14283/jpad.2023.30] [Medline: 37357276]
30. Caselli RJ, Reiman EM, Osborne D, et al. Longitudinal changes in cognition and behavior in asymptomatic carriers of the APOE e4 allele. Neurology (ECronicon) 2004 Jun 8;62(11):1990-1995. [doi: 10.1212/01.wnl.0000129533.26544.bf] [Medline: 15184602]
31. Marioni RE, Campbell A, Scotland G, Hayward C, Porteous DJ, Deary IJ. Differential effects of the APOE e4 allele on different domains of cognitive ability across the life-course. Eur J Hum Genet 2016 Jun;24(6):919-923. [doi: 10.1038/ejhg.2015.210] [Medline: 26395552]
32. Evans SL, Dowell NG, Prowse F, Tabet N, King SL, Rusted JM. Mid age APOE ε4 carriers show memory-related functional differences and disrupted structure-function relationships in hippocampal regions. Sci Rep 2020 Feb 20;10(1):3110. [doi: 10.1038/s41598-020-59272-0] [Medline: 32080211]
33. Palmer JM, Huentelman M, Ryan L. More than just risk for Alzheimer's disease: APOE ε4's impact on the aging brain. Trends Neurosci 2023 Sep;46(9):750-763. [doi: 10.1016/j.tins.2023.06.003] [Medline: 37460334]
34. Lu K, Nicholas JM, Pertzov Y, et al. Dissociable effects of APOE-ε4 and β-amyloid pathology on visual working memory. Nat Aging 2021 Nov;1(11):1002-1009. [doi: 10.1038/s43587-021-00117-4] [Medline: 34806027]
35. McLean JW, Bhattrai A, Vitali F, Raikes AC, Wiegand JP, Brinton RD. Contributions of sex and genotype to exploratory behavior differences in an aged humanized APOE mouse model of late-onset Alzheimer's disease. Learn Mem 2022 Sep;29(9):321-331. [doi: 10.1101/lm.053588.122]
36. Scheller E, Peter J, Schumacher LV, et al. APOE moderates compensatory recruitment of neuronal resources during working memory processing in healthy older adults. Neurobiol Aging 2017 Aug;56:127-137. [doi: 10.1016/j.neurobiolaging.2017.04.015] [Medline: 28528773]
37. Filbey FM, Chen G, Sunderland T, Cohen RM. Failing compensatory mechanisms during working memory in older apolipoprotein E-epsilon4 healthy adults. Brain Imaging Behav 2010 Jun;4(2):177-188. [doi: 10.1007/s11682-010-9097-9] [Medline: 20502990]
38. Iacono D, Feltis GC. Impact of apolipoprotein E gene polymorphism during normal and pathological conditions of the brain across the lifespan. Aging (Albany NY) 2019 Jan 24;11(2):787-816. [doi: 10.18632/aging.101757] [Medline: 30677746]

XSL•FO
RenderX

39. Palmer JM, Grilli MD, Lawrence AV, Ryan L. The impact of context on pattern separation for objects among younger and older apolipoprotein 4 carriers and noncarriers. J Int Neuropsychol Soc 2023 Jun;29(5):439-449. [doi: 10.1017/S1355617722000455] [Medline: 36416211]
40. Dean DC III, Jerskey BA, Chen K, et al. Brain differences in infants at differential genetic risk for late-onset Alzheimer disease. JAMA Neurol 2014 Jan 1;71(1):11. [doi: 10.1001/jamaneurol.2013.4544]
41. Zink N, Bensmann W, Arning L, Beste C, Stock AK. Apolipoprotein ε4 is associated with better cognitive control allocation in healthy young adults. Neuroimage 2019 Jan 15;185:274-285. [doi: 10.1016/j.neuroimage.2018.10.046] [Medline: 30342978]
42. Wishart HA, Saykin AJ, Rabin LA, et al. Increased brain activation during working memory in cognitively intact adults with the APOE epsilon4 allele. Am J Psychiatry 2006 Sep;163(9):1603-1610. [doi: 10.1176/ajp.2006.163.9.1603] [Medline: 16946187]
43. Tuminello ER, Han SD. The apolipoprotein E antagonistic pleiotropy hypothesis: review and recommendations. Int J Alzheimers Dis 2011 Feb 24;2011:726197. [doi: 10.4061/2011/726197] [Medline: 21423560]
44. Han SD, Bondi MW. Revision of the apolipoprotein E compensatory mechanism recruitment hypothesis. Alzheimers Dement 2008 Jul;4(4):251-254. [doi: 10.1016/j.jalz.2008.02.006] [Medline: 18631975]
45. Lingo VanGilder J, Bergamino M, Hooyman A, et al. Using whole-brain diffusion tensor analysis to evaluate white matter structural correlates of delayed visuospatial memory and one-week motor skill retention in nondemented older adults: a preliminary study. PLoS One 2022;17(9):e0274955. [doi: 10.1371/journal.pone.0274955] [Medline: 36137126]
46. Regan E, Fridriksson J, Schaefer SY, et al. Neural correlates of within-session practice effects in mild motor impairment after stroke: a preliminary investigation. Exp Brain Res 2021 Jan;239(1):151-160. [doi: 10.1007/s00221-020-05964-y] [Medline: 33130906]
47. Moreno-Briseño P, Díaz R, Campos-Romo A, Fernandez-Ruiz J. Sex-related differences in motor learning and performance. Behav Brain Funct 2010 Dec 23;6(1):74. [doi: 10.1186/1744-9081-6-74] [Medline: 21182785]
48. Buckley RF, Mormino EC, Rabin JS, et al. Sex differences in the association of global amyloid and regional tau deposition measured by positron emission tomography in clinically normal older adults. JAMA Neurol 2019 May 1;76(5):542-551. [doi: 10.1001/jamaneurol.2018.4693] [Medline: 30715078]
49. Caselli RJ, Langlais BT, Dueck AC, et al. Neuropsychological decline up to 20 years before incident mild cognitive impairment. Alzheimers Dement 2020 Mar;16(3):512-523. [doi: 10.1016/j.jalz.2019.09.085] [Medline: 31787561]
50. Savignac C, Villeneuve S, Badhwar A, et al. APOE alleles are associated with sex-specific structural differences in brain regions affected in Alzheimer's disease and related dementia. PLoS Biol 2022 Dec;20(12):e3001863. [doi: 10.1371/journal.pbio.3001863] [Medline: 36512526]
51. Farrer LA, Cupples LA, Haines JL, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. JAMA 1997;278(16):1349-1356. [Medline: 9343467]
52. Harwood DG, Ownby RL. Ethnicity and dementia. Curr Psychiatry Rep 2000 Feb;2(1):40-45. [doi: 10.1007/s11920-000-0040-4] [Medline: 11122930]
53. Terlecki M, Brown J, Harner-Steciw L, et al. Sex differences and similarities in video game experience, preferences, and self-efficacy: implications for the gaming industry. Curr Psychol 2011 Mar;30(1):22-33. [doi: 10.1007/s12144-010-9095-5]
54. Walter S, Clanton TB, Langford OG, et al. Recruitment into the Alzheimer Prevention Trials (APT) Webstudy for a Trial-Ready Cohort for Preclinical and Prodromal Alzheimer's Disease (TRC-PAD). J Prev Alzheimers Dis 2020;7(4):219-225. [doi: 10.14283/jpad.2020.46] [Medline: 32920623]
55. Hooyman A. Super G and AD risk factors in Mindcrowd. OSF. URL: https://osf.io/rnzg6 [accessed 2025-03-13]

Abbreviations

AD: Alzheimer disease
MCI: mild cognitive impairment
OR: odds ratio
PAL: paired associates learning
RT: response time
SR: scaling ratio
svRT: simple visual reaction time
TinT: time in the target
TR: time of reversal

XSL•FO
RenderX

Edited by A Garcia; submitted 08.10.24; peer-reviewed by A Lewis, AC Cavalcanti Maciel; revised version received 24.02.25; accepted 11.03.25; published 24.04.25.

Please cite as:
Hooyman A, Huentelman MJ, De Both M, Ryan L, Duff K, Schaefer SY
Relationship Between Within-Session Digital Motor Skill Acquisition and Alzheimer Disease Risk Factors Among the MindCrowd Cohort: Cross-Sectional Descriptive Study
JMIR Aging 2025;8:e67298
URL: https://aging.jmir.org/2025/1/e67298
doi:10.2196/67298

© Andrew Hooyman, Matt J Huentelman, Matt De Both, Lee Ryan, Kevin Duff, Sydney Y Schaefer. Originally published in JMIR Aging (https://aging.jmir.org), 24.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Implications of Public Disclosure of Personal Information in a Mobile Alert App for People Living With Dementia Who Go Missing: Qualitative Descriptive Study

Adebusola Adekoya[1], MN; Christine Daum[1], PhD; Noelannah Neubauer[1], PhD; Antonio Miguel-Cruz[2], PhD; Lili Liu[1], PhD

[1]School of Public Health Sciences, Faculty of Health, University of Waterloo, Waterloo, ON, Canada

[2]Department of Occupational Therapy, Faculty of Rehabilitation Medicine, University of Alberta, Edmonton, AB, Canada

Corresponding Author:
Lili Liu, PhD
School of Public Health Sciences, Faculty of Health
University of Waterloo
200 University Avenue West
Waterloo, ON, N2L 3G1
Canada
Phone: 1 519 888 4567
Email: lili.liu@uwaterloo.ca

Abstract

Background: People living with dementia are at risk of getting lost and going missing due to memory loss, confusion, and disorientation. Missing person incidents involving people living with dementia are increasing. Alert systems such as Community ASAP can promote community engagement in locating missing persons with dementia and aid in search and rescue efforts. However, the implications of public disclosure of personal information such as name, age, sex, and physical description within such alert systems have yet to be explored.

Objective: This study aimed to identify and discuss the implications of public disclosure of personal information in Community ASAP for people living with dementia at risk of going missing.

Methods: This study used a qualitative descriptive research design drawing from naturalistic inquiry. A total of 19 participants including people living with dementia, care partners, first responders, and service providers were recruited from Ontario, Alberta, and British Columbia, Canada. Semistructured interviews were used to explore participants' perspectives on the perceived implications of the release of personal information when using Community ASAP. NVivo (version 12) was used to manage data, and conventional content analysis was conducted to identify key themes of the implications of public disclosure of personal information in Community ASAP.

Results: In total, 10/19 (53%) of the participants were women and 9/19 (47%) were men. Of the 19 participants, 3 (16%) were people living with dementia, 5 (26%) were care partners, 4 (21%) were first responders, and 7 (37%) were service providers. In total, 4 key themes were identified as implications of public disclosure of personal information in Community ASAP: *right to autonomy*, *safety versus privacy*, *informed and knowledgeable consent*, and *stigmatization.* Participants discussed how the public disclosure of personal information in Community ASAP could undermine a person's choice not to be found and contribute to stigmatization. Participants emphasized a need to balance safety and privacy concerns. Informed and knowledgeable consent is important when using an alert system to locate missing persons with dementia.

Conclusions: Community ASAP can promote community engagement in locating missing persons with dementia. However, the public disclosure of personal information in alert systems has implications. Users' right to autonomy, a balance between safety and privacy, informed and knowledgeable consent, and risks of stigmatization are perceived impacts of disclosure of personal information in alert systems.

(JMIR Aging 2025;8:e64847) doi:10.2196/64847

KEYWORDS

alert systems; technology; missing persons; dementia; autonomy; privacy; stigmatization; consent

XSL•FO
RenderX

Introduction

Background

Dementia is an increasing health problem affecting >55 million people worldwide [1]. People living with dementia are at risk of getting lost and going missing due to memory loss, confusion, and disorientation [2]. Missing person incidents involving people living with dementia have been increasing [3]. Missing persons with dementia are at risk of being exposed to injuries and death. If not found within 24 hours, half of missing persons with dementia will sustain serious injuries or be found deceased [4]. It is crucial to engage the community in locating missing persons and not view missing person incidents solely as a family issue [5]. This inclusive approach provides extra eyes on the ground. Strategies such as alert systems promote community engagement by encouraging community members to be on the lookout for missing persons. This, in turn, aids in search and rescue efforts and minimizes the risks associated with going missing [6].

Alert systems such as the United States' publicly funded Silver Alert program disseminate information about missing persons with dementia or other cognitive disabilities to the public through media broadcasts (eg, commercial radio and television stations) and electronic billboards [6]. The Silver Alert system has been implemented in all but 5 states [6]. However, there are limitations to these programs. The process of issuing Silver Alerts and the duration differ across states, and media broadcasts may reach fewer people due to an increase in the use of mobile devices to access information. Furthermore, there are concerns about declining media sensitivity due to alert fatigue [7].

In Canada, when a person living with dementia goes missing, police services alert the public about a missing person using social media such as Facebook and X (formerly Twitter) and not an emergency alert system such as an Amber alert [8]. Amber alerts notify the public about missing abducted children using Alert Ready. Alert Ready is Canada's emergency alert system that sends critical alerts about hazards or impending dangers (eg, floods, tornados, chemical spills, and fires) to Canadians through radio, satellite television, and compatible wireless devices [9]. As missing person incidents involving people living with dementia are more frequent than missing children and not typically related to a crime, the use of the Alert Ready system would not be appropriate [8]. Furthermore, the use of the Alert Ready system can desensitize the public to alerts and cause alert fatigue, which already exists with Amber alerts.

Ideally, programs that mitigate risks of going missing among people living with dementia would engage local communities but minimize alert fatigue and be location specific. For example, the community-led British Columbia Silver Alert program notifies community subscribers about missing persons with dementia, cognitive impairment, and autism using channels such as social media (eg, Facebook and X) and via email or SMS text message [10]. Interest in community alert systems in other Canadian provinces is growing, as evidenced by a recent national petition; amendments to Missing Persons (Silver Alert) Acts in Alberta, Manitoba, and Ontario [11-13]; and the launch of Silver Alert pilot projects in Quebec [14].

To date, a localized alert system based on volunteer subscription does not exist. To this end, our team developed a localized area alert app called Community ASAP, the first of its kind in Canada, with representatives of end users, including people living with dementia, care partners, and community organizations [3]. Community ASAP allows community volunteers, local businesses, and police services to work together to locate missing persons with cognitive impairment and is available as a mobile app. As a requirement in Canada, a Community ASAP coordinator from police services would initiate an alert when a person living with dementia is reported missing. Community volunteers who register with Community ASAP receive alerts about missing persons based on their geographic preferences. The alert includes a link to the missing individual's personal information, such as their first and last names, nickname, age, physical description (eg, height, weight, eye color, and hair color), mobile number (if any), and locations they typically visit [3].

Community ASAP and other public alert systems such as Silver Alerts and Amber alerts are designed to enhance public awareness of vulnerable missing persons, but they differ in their scope and approaches. Community ASAP is a subscription-based alert system that requires community engagement and provides real-time updates and geofencing features. For example, volunteers receive missing person alerts based on their geographic preferences, allowing them to choose up to 5 locations, such as home and work addresses, and set a radius (1, 3, 6, 12, or 25 km) for each location [3]. However, its effectiveness depends on widespread app adoption and access to mobile devices. In contrast, Silver Alerts and Amber alerts are government funded or community funded and broadcast information about missing persons through large-scale public channels such as highway signs, radio, and television, reaching a wider audience [6,15]. The activation of Community ASAP in Canada and public alert systems in the United States rely on police verification, which means that alerts are not triggered immediately when a missing person incident occurs [3,6,16].

This study was a part of the development and evaluation of the accuracy and usability of Community ASAP [3]. During the development of Community ASAP, we identified the need to understand how the rights of people living with dementia to self-determination and privacy are respected while using the alert system. The use of alert systems has raised ethical concerns about the privacy of people living with dementia [6,15]. Furthermore, there is limited knowledge about concerns related to using alert systems to locate missing persons with dementia [15], particularly how the release of personal information to the public may impinge on the human rights of vulnerable older adults [17].

Objectives

The purpose of this study was to identify and discuss the implications of public disclosure of personal information in an alert system called Community ASAP for people living with dementia at risk of going missing.

XSL•FO
RenderX

Methods

Research Design

This study used a qualitative descriptive research design drawing from naturalistic inquiry, which allows a phenomenon to be studied in its natural state [18]. This method is appropriate when the intent of the study is to provide a basic description and summary of people's experiences [19,20].

Recruitment

Purposeful sampling [21] was used to select participants who had professional and personal experiences on the topic to best provide insights into concerns that could arise from the release of information via Community ASAP. Specific strategies included intensity [21] and snowball [22] sampling. Intensity sampling was used to sample information-rich cases. In snowball sampling, participants helped share study information with other prospective participants whom they believed had knowledge or experience with the topic of interest. These strategies were chosen for practical reasons, including the challenges of selecting potentially difficult-to-access or vulnerable groups, such as people living with dementia. We aimed to include the perspectives of people living with dementia, who are often underrepresented in research. Studies have shown the importance of involving people living with dementia in research, highlighting their ability to provide reliable self-report information, personal perspectives, and valuable insights [23-25]. Inclusion of people living with dementia in this study was crucial to understanding their individual experiences with an alert system and enhancing their safety and well-being while using this program.

Participants (people living with dementia, care partners, first responders, and service providers) were recruited and identified via email or face-to-face conversations through our research team's existing professional networks serving people living with dementia, such as Alzheimer societies, dementia advocacy organizations, and first responders. Thus, some of the participants were known to the research team, and others were not. None of the participants were family members or friends of or had personal relationships with the researchers. All prospective participants who were invited to take part in the study enrolled. Participants had lived or professional experience with dementia and the use of technologies to manage the risk of going missing or responding to missing person incidents. During recruitment, participants were asked to self-identify their experience and knowledge of technologies. All participants had some level of familiarity or experience with technology. We did not inquire about the extent or specific type of their experience.

We selected participants from across 3 Canadian provinces (Ontario, Alberta, and British Columbia) where our professional networks were located and from 4 stakeholder groups: people living with dementia, care partners, service providers (eg, social workers, dementia educators, representatives from vulnerable person registries, and support workers), and first responders (eg, police and search and rescue managers). Inclusion criteria were to (1) be aged ≥18 years, (2) speak English, and (3) have lived or professional experience with or knowledge of the use of technologies to manage the risk of going missing or responding to missing person incidents. Exclusion criteria were having no understanding about technologies to manage the risk of going missing, inability to articulate perspectives due to impaired cognitive abilities, and severe visual or hearing limitations that could not be corrected with the use of an assistive device.

Ethical Considerations

Ethics approval was received from the University of Alberta Research Ethics Board (Pro00078537). All participants gave written informed consent prior to the study, were told that their participation is voluntary, and were informed of their right to withdraw at any time. People living with mild cognitive impairment or mild dementia who had the ability to consent as identified by our professional networks were invited to participate in this study. We used the teach-back method [26] to ensure that these participants had cognitive ability to engage in one-on-one interviews with research team members. The person living with dementia was asked an open-ended question about what they had read in the information letter to ensure that they understood the procedure, risks, and what to do should they wish to withdraw from the study. Transcripts were cleaned for accuracy and deidentified to maintain anonymity. Participants did not receive an honorarium.

Data Collection

Semistructured interviews [27] were conducted with each participant in person, via phone, or through Zoom (Zoom Video Communications) videoconference. The purpose of the interviews was to understand participant perspectives on the implications of public disclosure of personal information in alert systems. Each person selected a preferred mode of participation based on their geographical location and other time commitments. In total, we interviewed 19 stakeholders (n=16, 84% individual interviews and 1 group interview with n=3, 16% of the participants at their request). No other individuals (ie, care recipients or care partners) were present during the interviews with people living with dementia as these individuals had the ability to consent and respond to questions. Before the interview began, the researcher introduced herself, the overarching study [3] in which the current project was embedded, and the purpose and procedures, and written informed consent was obtained. An interview guide (Textbox 1) that contained 5 open-ended questions was used to elicit participants' perspectives on the possible implications of the release of information through Community ASAP. Probes were used to elaborate on participant responses and clarify meaning [22]. The interviews were conducted by female research team members (NN and CD) with backgrounds in occupational therapy, experience with qualitative research methods, and PhDs in rehabilitation science. Each participant was interviewed once, and the interviews were approximately 30 to 90 minutes in length. Interviewers documented their observations and any interruptions in field notes, and the interviews were digitally recorded and transcribed verbatim. Transcripts were not returned to participants as we did not use a member-checking approach to review the transcripts. NVivo (version 12; Lumivero) was used to manage the data.

XSL•FO
RenderX

Textbox 1. Interview guide questions.

Questions

- From your perspective as a (person with dementia or caregiver of a person with dementia, service provider, first responder, expert in ethics, or the law), what ethical concerns are associated with release of personal information in the C-ASAP system?
- From your perspective as a (person with dementia or caregiver of a person with dementia, service provider, first responder, expert in ethics, or the law), what legal concerns are associated with release of personal information in the C-ASAP system?
- How does the release of personal information in the proposed C-ASAP system compare with other registries such as MedicAlert Safely Home and various Vulnerable Persons Registries?
- What are the implications of having a missing person's name released to the public (and thus a part of the public record)? What is the balance (or tipping point) between privacy and safety?
- [For group interviews involving representatives from agencies that collect information about vulnerable persons and store them on registries]: What agreements are in place between the vulnerable person or their family and the agency upon signing on with a particular registry?

Data Analysis

Conventional content analysis [28] guided our analytic approach. A thematic approach is preferred when there is limited existing theory or research literature on a phenomenon, allowing for the identification of specific meanings and determination of appropriate categories and themes. A female doctoral candidate in public health sciences (AA) with a background in nursing and qualitative research methods analyzed the data. She immersed herself in the data by listening to the recordings while reading the transcripts and identifying initial reflections. Next, she coded the transcripts using keywords. Codes were described, and similar codes were grouped and refined to create categories. As new codes were generated, they were added to the framework, and transcripts that had been previously coded were updated to reflect hierarchy. Category descriptions were generated to describe their contents, and themes that represented the categories were created inductively. Analysis continued until saturation of the data was reached.

Data analysis was an iterative process. The coding hierarchy, categories, and themes were reviewed, scrutinized, and confirmed by the analyst and the 2 team members who conducted the interviews; that is, we used peer debriefing as a trustworthiness strategy [29]. This helped us improve clarity and the internal and external homogeneity of codes and categories [30] and consider alternative interpretations and explanations. Thus, the coding hierarchy and main themes were refined repeatedly, thereby enhancing credibility. Our process is consistent with the approach by Morse et al [31] to verification in which transcripts, codes, categories, and themes are rechecked.

Results

Overview

A total of 19 participants (n=10, 53% women and n=9, 47% men) were interviewed. Of these 19 participants, 3 (16%) were people living with dementia, 5 (26%) were care partners, 4 (21%) were first responders (search and rescue and police officers), and 7 (37%) were service providers. Participants were predominantly White individuals (17/19, 89%), and the remainder were Asian individuals (Korean and Filipino; 2/19, 11%; Table 1). All participants had some level of familiarity or experience with technology. In total, 58% (11/19) of the participants (first responders and service providers) had professional experience with technology to manage or respond to missing person incidents, such as alert systems, MedicAlert, and locator devices (GPS devices and Project Lifesaver). The remaining participants (people living with dementia and care partners; 8/19, 42%) had lived experience or familiarity these technologies.

XSL•FO
RenderX

Table 1. Study participants.

	Care partners (n=5), n (%)	First responders (n=4), n (%)	People living with dementia (n=3), n (%)	Service providers (n=7), n (%)
Sex				
Male	4 (80)	4 (100)	1 (33)	0 (0)
Female	1 (20)	0 (0)	2 (67)	7 (100)
Ethnicity				
Filipino	0 (0)	0 (0)	0 (0)	1 (14)
Korean	1 (20)	0 (0)	0 (0)	0 (0)
White	4 (80)	4 (100)	3 (100)	6 (86)
Province				
Alberta	1 (20)	1 (25)	0 (0)	0 (0)
British Columbia	1 (20)	1 (25)	1 (33)	2 (29)
Ontario	3 (60)	2 (50)	2 (67)	5 (71)

Thematic Findings

Overview

Four key themes represent the implications of public disclosure of personal information in Community ASAP for people living with dementia who have gone missing: (1) right to autonomy, (2) safety versus privacy, (3) informed and knowledgeable consent, and (4) stigmatization (Textbox 2).

XSL•FO
RenderX

Textbox 2. Themes with participant quotes.

Right to autonomy

- "Depending on where the person is in their journey, what if they just wanted to be alone for a few days and then got this whole thing going on because someone decided that they've disappeared. That's a legal issue for me." [Person living with dementia 3, an experienced user of GPS devices who had previously been lost]
- "We locate a missing person, we let them know you've been reported missing, here's the person who's reported you missing...However, they do have a right to be missing so to speak." [First responder 4, an experienced police officer with expertise in search and rescue]
- "Is that what they wanted when they were in the real self before dementia, there is an ethical issue there because you're not honoring what the person wants in the moment." [Service provider 3, a service provider who worked directly with people living with dementia providing education and counseling]

Safety versus privacy

- "The more eyes the better. I think if you're missing, I don't think privacy should be an issue." [Person living with dementia 1, who had experience with MedicAlert]
- "I think that the tipping point is safety, and so that might be it is 30 to 20 degrees Celsius outside. It's not as extreme situation as it is 30 or minus 25. So, all of these things need to be taken into account." [Care partner 2, a family member who cared for a person living with dementia]
- "Some people aren't as open to having other people or the general public know a person with dementia has been missing more around, that they have the condition, they have dementia." [Service provider 6, a person with expertise in supporting the safety of vulnerable persons and who had experience with GPS locator devices for this population]
- "If they are making this phone call it's urgent enough, it's important enough that yes, we are willing to give up that right to this in order to find this person quickly." [First responder 1, an experienced search and rescue member who also provided education about locator devices]

Stigmatization

- "For a lot of people, I worry about their faces and the local papers, this stuff. And that still makes them very vulnerable, because someone recognizes them after the fact." [Person living with dementia 3, an experienced user of GPS devices who had previously been lost]
- "So, with my dad, we didn't want anybody to know that he had Alzheimer's. I think it's because there may be a stigma of Alzheimer's." [Care partner 3, a family member of a person living with dementia who got lost]
- "So, it's such a stigma, it's such a huge issue that sometimes you might think that having an alert system would be really useful for the community to know how to interact or to keep their eye out for somebody with dementia in the community. But is a stigma sort of perpetuating." [Service provider 7, a service provider who works directly with people living with dementia providing education and counseling]
- "The one thing we are always cognizant of in our media releases, is our management and our upper management are always cognizant of the...stigma." [First responder 3, a police officer with experience with missing persons]

Informed and knowledgeable consent

- "If the person is like me and can give consent, that would be OK...if I've already deteriorated, then I think it would be up to my family or my power of attorney because they know what my wishes are." [Person living with dementia 2, a person who was previously lost]
- "Whenever possible, that person who is actually living with dementia, I certainly would advocate and support their involvement from the get-go. No matter what the timeframe, I think you need to be respectful and try to include their voice as much as possible, knowing that cognitively that could change at any time, and if it does then I would hope that the person has the best interests for that individual." [Care partner 5, a family member of a person living with dementia who had experience with locator devices]
- "I think the idea of capacity and the ability to give informed consent is not an on/off button. And to always have that conversation as much as possible to the extent that it is possible for you and whether that's through care partners. I think it's really at the heart of it, really enabling that person to have the choice. But the challenge with that is that on one day the person might say 'yeah I'm completely fine with it' and then the next day they're like 'ugh, I'm not fine with it at all.'" [Service provider 7, a service provider who works directly with people living with dementia providing education and counseling]
- "Informed consent is the only consent that would certainly help alleviate some problems down the road, and I would think if the person can't give consent, because of their mental capacity, hopefully the caregiver would have that Power of Attorney to be able to make that decision for them, on their behalf." [First responder 2, an experienced police officer and search and rescue expert]

Right to Autonomy

This theme relates to concerns about how the release of personal information can compromise a person's choice not to be found. Participants discussed the importance of respecting people's autonomy and right to go missing without wanting to be found. People living with dementia also possess the rights of adults and have the right to be alone or go missing, whereas this does not apply to children. A participant (first responder 4) with many years of experience in search and rescue emphasized adults' rights to go missing. He expressed the following:

> *When we locate a missing person, we let them know you've been reported missing, here's the person who's reported you missing...However, they do have a right to be missing so to speak.*

XSL•FO
RenderX

Labeling intentionally missing persons as missing, especially those capable of decision-making, may raise ethical and legal concerns, as noted by some participants. This concern was further expanded by a participant living with dementia:

> *Depending on where the person is in their journey, what if they just wanted to be alone for a few days and then got this whole thing going on because someone decided that they've disappeared. That's a legal issue for me. Like where in the journey it is decided that someone...I don't even know how to explain that because like for me, it's not uncommon for me, I mean now I always tell somebody just because that's my life. I quite often go "Oh well I'm going to take off and go here for a couple of days."* [Person living with dementia 3, an experienced user of GPS devices who had previously been lost]

Participants pointed out that individuals' views on risk and outlook on life may change as their dementia progresses. Hence, it is crucial that the use of Community ASAP respects the preferences of those living with dementia. In addition, there might be a need to re-evaluate personal information disclosure preferences as dementia progresses. A service provider highlighted the importance of honoring the current wishes of people living with dementia when using Community ASAP:

> *Somebody's idea about what they consent to and what they want when they're first diagnosed might be different than what they want later on, and what's the record of truth. Is that what they wanted when they were in the real self before dementia, there is an ethical issue there because you're not honoring what the person wants in the moment.* [Service provider 3, a service provider who worked directly with people living with dementia providing education and counseling]

Safety Versus Privacy

This theme highlights concerns related to the balance between safety and privacy, acknowledging the diverse perspectives on how to maintain this balance for missing persons. While Community ASAP serves as a safety net, offering peace of mind to people living with dementia and their care partners, it also facilitates a sense of independence. However, some participants expressed concerns about privacy when it comes to sharing their personal information publicly. For many, safety was the tipping point as finding the missing person alive and minimizing the risks of harm is paramount. A care partner emphasized the urgency of finding missing persons promptly, especially in severe weather conditions in Canada:

> *I think that the tipping point is safety, and so that might be it is 30 to 20 degrees Celsius outside. It's not as extreme situation as it is 30 or minus 25. So, all of these things need to be taken into account.* [Care partner 2, a family member who cared for a person living with dementia]

Privacy concerns are influenced by culture as cultural beliefs play a key role in how individuals view privacy, particularly regarding the disclosure of certain information to the public. For instance, a Korean participant (care partner 3, a family member of a person living with dementia who got lost) shared cultural beliefs about having a "conservative background" and keeping his father's dementia diagnosis confidential within the family and refraining from sharing it with outsiders.

There are concerns regarding the permanence of personal information on the internet. Details such as a missing person's name, age, photo, and physical description shared with the public might linger on the internet even after the person is located. Some participants expressed unease about their personal information being public posthumously.

Participants also emphasized the importance of only sharing essential information, such as the person's name, age, photo, and physical description. While some information is necessary for community members (volunteers) to identify a missing individual, disclosing medical information (eg, cognitive impairment) could heighten the risks of fraud, identity theft, and abuse. One participant raised concerns about potential exploitation of people living with dementia when using Community ASAP:

> *The risk is that the volunteers would prey upon this person if he/she is found and returned home, identified to have dementia. They are already vulnerable to scammers, to people who would be able to redirect them and take advantage of them.* [Service provider 1, a service provider who worked directly with people living with dementia providing education and counseling]

Some participants expressed that a dementia diagnosis could make individuals vulnerable to exploitation regardless of whether their personal information is shared via an alert system such as Community ASAP. One participant (person living with dementia 1) stressed the significance of not disclosing banking details to the public:

> *You're not going to give out any information regarding bank accounts. If people want to take advantage of you, they're gonna do it anyways.*

Stigmatization

This theme discusses concerns about how the use of Community ASAP could lead to stigma for people living with dementia. People living with dementia may experience stigma due to their dementia diagnosis and assumptions about their abilities and behaviors. People may assume that individuals living with dementia do not have the capacity to make decisions for themselves. Therefore, sharing personal information such as a cognitive impairment or dementia diagnosis publicly could heighten the existing risk of stigmatization for people living with dementia. A police officer (first responder 3) with experience with missing persons expressed the following:

The one thing we are always cognizant of in our media releases, is our management and our upper management are always cognizant of the...stigma.

A service provider also explained the stigma associated with using alert systems:

XSL•FO
RenderX

> *So, it's such a stigma, it's such a huge issue that sometimes you might think that having an alert system would be really useful for the community to know how to interact or to keep their eye out for somebody with dementia in the community. But is a stigma sort of perpetuating.* [Service provider 7, a service provider who works directly with people living with dementia providing education and counseling]

Participants pointed out that the public may not understand how to approach people living with dementia or communicate with them, hence the need to address dementia-related stigma and misconceptions through public education and awareness campaigns. It was also noted that service providers and first responders should receive adequate training on recognizing, communicating with, and responding to missing persons with dementia. A service provider discussed the importance of using nonstigmatizing language when interacting with missing persons with dementia:

> *I think one of the issues we may see is how responsive behaviors are presented. So, if there are triggers, because I know it is important to communicate that to the community, like you know do not touch this person because they might respond in this way. How can we train coordinators to use language that is going to be non-stigmatizing but still express the nature of that behavior.* [Service provider 1, a service provider who worked directly with people living with dementia providing education and counseling]

Informed and Knowledgeable Consent

This theme focuses on concerns related to ensuring individuals' informed and knowledgeable consent when using Community ASAP. Participants highlighted the importance of people living with dementia having the capacity to understand the implications of an alert system, its objectives, and the reasons behind collecting and sharing their personal data. They emphasized the necessity of providing sufficient information about Community ASAP to those living with dementia and respecting their autonomy in decision-making. However, dementia can affect an individual's ability to process and understand information adequately, leading to challenges in decision-making regarding alert systems, which does not always align with a simple *present or absent* perspective. This concern was elaborated on by a service provider:

> *I think the idea of capacity and the ability to give informed consent is not an on/off button. And to always have that conversation as much as possible to the extent that it is possible for you and whether that's through care partners. I think it's really at the heart of it, really enabling that person to have the choice. But the challenge with that is that on one day the person might say "yeah I'm completely fine with it" and then the next day they're like "ugh, I'm not fine with it at all."* [Service provider 7, a service provider who worked directly with people living with dementia providing education and counseling]

For individuals with limited decision-making capacity, service providers should seek consent from their proxies or substitute decision makers (eg, relatives) and make decisions in their best interests. Participants emphasized the importance of a respectful decision-making process that considers the evolving wishes and preferences of individuals with dementia as the condition advances. A first responder discussed the significance of informed consent and ensuring that decisions made on behalf of an individual living with dementia align with their preferences:

> *Informed consent is the only consent and that would certainly help alleviate some problems down the road, and I would think if the person can't give consent, because of their mental capacity, hopefully the caregiver would have that Power of Attorney to be able to make that decision for them, on their behalf. What's the end goal? The end goal is to try and help that person the best you can, and if they are in a position that is maybe not good for them, at least hopefully we put enough checkmarks in place that they'll come out okay, and that's what you have to look at.* [First responder 2, an experienced police officer and search and rescue expert]

The need to involve people living with dementia in the decision-making process regarding the use of Community ASAP was discussed by participants. It was emphasized that involving the person with dementia in decisions at an early stage is crucial as dementia progression can greatly impact decision-making capacity. Care partner 5, a family member of a person living with dementia who had experience with locator devices, expressed the following:

> *You need to be respectful and try to include their voice as much as possible, knowing that cognitively that could change at any time, and if it does...I would hope that the person has the best interests for that individual.*

Discussion

Principal Findings

Our study aimed to identify the implications of public disclosure of personal information in an alert system called Community ASAP for people living with dementia who are at risk of going missing. The implications were discussed under 4 key themes: *right to autonomy*, *safety versus privacy*, *stigmatization*, and *informed and knowledgeable consent.*

Community ASAP differs from public alert systems for missing persons with dementia such as Silver Alerts. While Community ASAP is a subscription-based, localized alert system that provides real-time updates and geofencing features, Silver Alerts can be government funded or community funded and use large-scale channels such as highway signs and media to reach a wider audience [3,6]. Given the limited research on subscription-based, localized alert systems for missing persons with dementia, our references drew primarily from studies on Silver Alerts.

Participants raised concerns about respecting individuals' right to autonomy when using Community ASAP. The literature emphasizes the importance of honoring the autonomy of people

XSL•FO
RenderX

living with dementia in alert systems [17,32]. Autonomy, or self-determination, is a fundamental human right tied to a person's capacity to assess options, weigh risks, communicate choices, and act accordingly [33,34]. Unlike children requiring guardianship, adults with dementia, even those with appointed guardians, retain some decision-making capacity [33].

People living with dementia may go missing intentionally to avoid disclosing their location or unintentionally due to disorientation or difficulty recognizing familiar places [16,35]. Respecting the right to go missing intentionally is controversial as risk is often viewed negatively and people living with dementia are considered a vulnerable group needing protection [36,37]. The use of technology to locate missing persons raises ethical concerns about balancing autonomy and privacy with safety [6,17,34]. For example, dementia-related wandering is typically seen as "risky" due to potential harm, but it can also represent autonomy and agency and even provide health benefits as a form of exercise [38]. In care settings, health care providers may monitor "wandering" due to ethical responsibilities and potential litigation fears [39]. However, disclosing personal information can violate a person's autonomy [6]. It is crucial to balance safety with respect for autonomy, necessitating open discussions between people living with dementia and their care partners or health care professionals [39]. These conversations should focus on the person's values, care preferences, and ways to support their choices while living with risks.

The public disclosure of personal information in Community ASAP also raises concerns about balancing safety and privacy. While safety focuses on locating missing persons and minimizing risks of harm, privacy emphasizes the right to choose what information is shared with others [33]. Participants expressed that disclosing details such as dementia diagnoses could violate privacy rights. Similarly, Wasser and Fox [32] found in their analysis of Silver Alerts that such systems can unintentionally infringe on the privacy rights of individuals with dementia. In addition, cultural beliefs influence perceptions of dementia [40] and privacy as some cultures view such diagnoses as private. Understanding these cultural factors is essential when using alert systems while ensuring the privacy of people living with dementia.

For most participants, the safety of missing persons outweighs privacy rights. The literature shows that care partners often prioritize safety over privacy concerns [15,16]. Alert systems can serve as a safety net for people living with dementia, providing comfort and independence while easing care partners' worries. Balancing safety and privacy requires understanding individual preferences and the perceptions of risk. Disclosure of personal information may be necessary for safety, but technology can also create a false sense of security if it fails [8,16,33]. To reduce the risk of individuals going missing, additional measures such as GPS devices and return home interviews should be used alongside alert systems [8]. Informed dialogue among people living with dementia, care partners, and health care professionals is essential, along with effective policies to address these concerns [6,16,33].

Public disclosure of personal information in Community ASAP can contribute to the stigmatization and abuse of people living with dementia. Stigma can lead to delays in seeking help, exclusion from social activities, and reduced quality of life [41,42]. Participants in this study emphasized the need for dementia education and staff training to address stigma and improve communication regarding missing persons. Community education is also crucial to raise awareness about the risks associated with alert systems [6,8]. Furthermore, a participatory approach in the design, development, and evaluation of alert systems is necessary [3].

People living with dementia are at heightened risk of abuse due to their diagnosis [43,44]. Publicly sharing information such as dementia diagnoses and home addresses can lead to financial exploitation and other crimes. For example, the United States' Silver Alert may broadcast home addresses, license plates, and photographs of missing persons, unintentionally increasing risks of victimization or identity theft [7,16,17]. It is vital to limit the information shared in alert systems to what is necessary and implement safety measures such as using general locations instead of specific addresses and ensuring that sensitive data are controlled and removed promptly [3].

Another concern regarding the public disclosure of personal information in Community ASAP is informed and knowledgeable consent. Consent is considered knowledgeable if "the individual knows the purposes of the collection, use or disclosure and knows that they have the right to give, withhold or withdraw consent" [45]. Participants noted that people living with dementia may face challenges in providing informed consent for using alert systems, often due to assumptions about their capacity [6,17]. While dementia can impact decision-making capacity in some cases, individuals can still make decisions with support from family, friends, or health care professionals [33,46]. Assessment of decision-making capacity can be difficult as dementia-related cognitive fluctuations affect attention and memory [46,47]. A dementia diagnosis does not automatically imply incapacity [46], and capacity assessments should account for the person's stage of dementia and be conducted during moments when they are experiencing minimal confusion [47]. Continuous consent should be sought as the condition progresses, and early discussions are encouraged to address future decision-making needs. For those requiring substitute decision makers, involving them in decisions respects autonomy and dignity. Regardless of dementia progression, individuals should be included in conversations about their health and the use of alert systems, ensuring that their voices are considered in the design and implementation of these technologies.

Strengths and Limitations

While there is research on public alert systems such as Silver Alerts, little is known about the implications of public disclosure of personal information in subscription-based, localized alert systems. Our study addressed this gap by examining how Community ASAP may infringe on the rights of people living with dementia. Specifically, it highlights the need to understand and address end users' concerns about disclosure of personal information, the right to autonomy, the balance between safety and privacy, and informed and knowledgeable consent. Including participants with lived and professional experience

XSL•FO
RenderX

with dementia and technologies for managing risks of going missing provided diverse perspectives on a topic with limited existing research. The qualitative descriptive approach used in this study enabled an in-depth exploration of concerns associated with public disclosure of personal information in subscription-based, localized alert systems.

Community ASAP and this study have limitations. The subscription-based, localized design of Community ASAP and the small participant sample size for each category limit its generalizability. The subscription model may exclude lower-income users and create a self-selection bias, whereas the app's reliance on a high density of community users may reduce its effectiveness in underengaged or rural areas. These factors may have influenced participant responses and perspectives. To enhance broader applicability, addressing barriers such as cost, accessibility, and adoption in diverse communities would be essential in future research.

In addition, it was unclear how much experience participants had with localized alert systems. Thus, the concept of the release of personal information in such alert systems was hypothetical. Participants' perspectives may differ with real-world experience, and the abstract nature of the topic may have led to the generation of data that are somewhat limited compared to those that a more concrete topic would yield. Recruitment methods also presented a limitation, potentially privileging certain voices. Participants were predominantly White individuals, which does not represent the demographics of Canada. In addition, participants were recruited through the researchers' networks, likely including individuals who were more vocal, confident, or willing to participate. This may have excluded underrepresented perspectives such as those from ethnic communities where alert systems might be viewed differently or where stigma around dementia is more significant.

Conclusions

Localized alert systems such as Community ASAP differ from existing Silver Alert programs that push alerts to the public through the media. Instead, Community ASAP pushes alerts to targeted community members who volunteer to be extra eyes and ears on the lookout for missing persons with dementia. This can promote community engagement in locating missing persons with dementia. However, the public disclosure of personal information in alert systems has significant implications. Such disclosure can compromise an individual's right to autonomy and privacy and lead to stigmatization. The balance of privacy rights with safety concerns presents a challenge. Informed and knowledgeable consent should be a fundamental part of using alert systems. Regardless of the stage of dementia, individuals living with the condition have a right to be included in conversations about their health, including the use of alert systems. An understanding of these implications paves the way to respecting the rights of people living with dementia at risk of going missing and improving their safety and well-being.

Acknowledgments

This work was supported by the Centre for Aging + Brain Health Innovation and Aging Gracefully Across Environments Using Technology to Support Wellness, Engagement, and Long Life from the Networks of Centres of Excellence of Canada. Community ASAP was conceived by Ron Beleno, a caregiver for his father who lived with dementia. The authors thank the study participants for contributing their time and feedback, including the Calgary Missing Older Adult Resource Network, Alzheimer Society of Ontario, Alzheimer Society of Durham Region, and Coquitlam Search and Rescue, as well as Canadian police services in Durham, Ontario; Toronto, Ontario; and Calgary, Alberta, Canada.

Conflicts of Interest

None declared.

References

1. Dementia. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2024-06-22]
2. Ward R, Rummery K, Odzakovic E, Manji K, Kullberg A, Clark A, et al. Getting lost with dementia: encounters with the time-space of not knowing. Health Place 2022 Nov;78:102940 [FREE Full text] [doi: 10.1016/j.healthplace.2022.102940] [Medline: 36403420]
3. Neubauer N, Daum C, Miguel-Cruz A, Liu L. Mobile alert app to engage community volunteers to help locate missing persons with dementia. PLoS One 2021;16(7):e0254952. [doi: 10.1371/journal.pone.0254952] [Medline: 34280219]
4. Alzheimer's Association. 2016 Alzheimer's disease facts and figures. Alzheimers Dement 2016 Apr;12(4):459-509. [doi: 10.1016/j.jalz.2016.03.001] [Medline: 27570871]
5. Buckley DM. Towards more effective missing women investigations: police relationships with victims' families, the community and the media a policy discussion. Missing Women Commission of Inquiry. 2012. URL: https://missingwomen.library.uvic.ca/wp-content/uploads/2010/10/POL-2-March-2012-MB-Towards-More-Effective-Missing-Women-Investigations.pdf [accessed 2023-09-18]
6. Gergerich E, Davis L. Silver alerts: a notification system for communities with missing adults. J Gerontol Soc Work 2017 Apr;60(3):232-244. [doi: 10.1080/01634372.2017.1293757] [Medline: 28409711]
7. Gier VS, Kreiner DS, Lampinen JM. Factors affecting recognition of senior citizens in a silver alert. J Police Crim Psych 2016 Oct 18;32(3):185-196. [doi: 10.1007/S11896-016-9210-0]

XSL•FO
RenderX

8. Adekoya A, Daum C, Neubauer N, McLennan L, Liu L. Policy brief Community alert systems for missing vulnerable adults: a call to action for policy makers. University of Waterloo. URL: https://uwaterloo.ca/aging-innovation-research-program/sites/default/files/uploads/files/policy_brief_community_alert_systems_finalcitation.pdf [accessed 2024-04-29]
9. Emergency alerts and the national public alerting system. Government of Canada. URL: https://crtc.gc.ca/eng/television/services/alert.htm [accessed 2023-09-18]
10. Alert criteria. BC Silver Alert. URL: https://bcsilveralert.ca/ [accessed 2023-09-18]
11. Bill 204: missing persons (silver alert) amendment act. Legislative Assembly of Alberta. URL: https://www.assembly.ab.ca/assembly-business/bills/bill?billinfoid=11993&from=bills [accessed 2023-09-18]
12. the missing persons amendment act (silver alert): 2nd session, 41st legislature. The Legislative Assembly of Manitoba. 2017. URL: https://web2.gov.mb.ca/bills/41-2/b214e.php [accessed 2023-09-18]
13. Bill 74, missing persons amendment act. Legislative Assembly of Ontario. URL: https://www.ola.org/en/legislative-business/bills/parliament-43/session-1/bill-74#:~:text=A%20vulnerable%20persons%20alert%20can,by%20a%20person%20in%20a [accessed 2023-09-18]
14. Quebec to launch 'Silver Alert' pilot project to locate missing seniors. Canadian Broadcasting Corporation (CBC). URL: https://www.cbc.ca/news/canada/montreal/silver-alert-quebec-1.6464902 [accessed 2023-09-18]
15. Carr D, Muschert GW, Kinney J, Robbins E, Petonito G, Manning L, et al. Silver alerts and the problem of missing adults with dementia. Gerontologist 2010 Apr 25;50(2):149-157. [doi: 10.1093/geront/gnp102] [Medline: 19556393]
16. Griffin T, Wiecko F. “Worth it if it saves just one life”? Ethical reservations regarding AMBER alert. Crim Justice Ethics 2015 Aug 14;34(2):156-176. [doi: 10.1080/0731129X.2015.1059572]
17. Petonito G, Muschert GW, Carr DC, Kinney JM, Robbins EJ, Brown JS. Programs to locate missing and critically wandering elders: a critical review and a call for multiphasic evaluation. Gerontologist 2013 Feb;53(1):17-25 [FREE Full text] [doi: 10.1093/geront/gns060] [Medline: 22565495]
18. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g] [Medline: 10940958]
19. Kim H, Sefcik JS, Bradway C. Characteristics of qualitative descriptive studies: a systematic review. Res Nurs Health 2017 Feb 30;40(1):23-42 [FREE Full text] [doi: 10.1002/nur.21768] [Medline: 27686751]
20. Neergaard MA, Olesen F, Andersen RS, Sondergaard J. Qualitative description - the poor cousin of health research? BMC Med Res Methodol 2009 Jul 16;9:52 [FREE Full text] [doi: 10.1186/1471-2288-9-52] [Medline: 19607668]
21. Emmel N. Purposeful sampling. In: Emmel N, editor. Sampling and Choosing Cases in Qualitative Research: A Realist Approach. Thousand Oaks, CA: Sage Publications; 2013.
22. Patton MQ. Qualitative Research & Evaluation Methods. Thousand Oaks, CA: Sage Publications; 2002.
23. Adekoya AA, Guse L. Walking interviews and wandering behavior: ethical insights and methodological outcomes while exploring the perspectives of older adults living with dementia. Int J Qual Methods 2020 Apr 20;19(3 Suppl. 2):89-95. [doi: 10.1177/1609406920920135]
24. Brooks J, Savitch N, Gridley K. Removing the 'gag': Involving people with dementia in research as advisers and participants. Soc Res Pract 2017;3:14 [FREE Full text]
25. O'Connor CM, Liddle J, O'Reilly M, Meyer C, Cartwright J, Chisholm M, et al. Advocating the rights of people with dementia to contribute to research: considerations for researchers and ethics committees. Australas J Ageing 2022 Jun 25;41(2):309-313. [doi: 10.1111/ajag.13023] [Medline: 34821448]
26. Holtz LR. Applying ethical principles to enrolling older adults with cognitive impairment in research. In: Byrdsong Q, editor. Sage Research Methods Cases: Medicine and Health. Thousand Oaks, CA: Sage Publications; 2023.
27. Richards L, Morse JM. Readme First for a User's Guide to Qualitative Methods. Thousand Oaks, CA: Sage Publications; 2007.
28. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
29. Barber JP, Walczak KK. Conscience and critic: peer debriefing strategies in grounded theory research. In: Proceedings of the 2009 Conference on Annual Meeting of the American Educational Research Association. 2009 Presented at: AERA '09; April 13-17, 2009; San Diego, CA p. 1-19 URL: https://www.researchgate.net/publication/242479874_Conscience_and_Critic_Peer_Debriefing_Strategies_in_Grounded_Theory_Research [doi: 10.4135/9781412986281.n252]
30. Mayan MJ. Essentials of Qualitative Inquiry. Walnut Creek, CA: Routledge; 2009.
31. Morse JM, Barrett M, Mayan M, Olson K, Spiers J. Verification strategies for establishing reliability and validity in qualitative research. Int J Qual Methods 2002 Jun 01;1(2):13-22. [doi: 10.1177/160940690200100202]
32. Wasser TD, Fox PK. For whom the bell tolls: Silver Alerts raise concerns regarding individual rights and governmental interests. J Am Acad Psychiatry Law 2013;41(3):421-429. [Medline: 24051596]
33. Liu L, Daum C, Neubauer N, Miguel Cruz A, Rios Rincón A. Autonomy and Independence: Aging in an Era of Technology. New York, NY: Morgan & Claypool Publishers; 2022.

XSL•FO
RenderX

34. Bantry-White E. Supporting ethical use of electronic monitoring for people living with dementia: social work's role in assessment, decision-making, and review. J Gerontol Soc Work 2018 Apr 16;61(3):261-279. [doi: 10.1080/01634372.2018.1433738] [Medline: 29381128]
35. Greene KS, Clarke CL, Pakes F, Holmes L. People with dementia who go missing: a qualitative study of family caregivers decision to report incidents to the police. Policing 2019;13(2):241-253. [doi: 10.1093/police/paz007]
36. Brannelly T. Citizenship and people living with dementia: a case for the ethics of care. Dementia (London) 2016 May 10;15(3):304-314. [doi: 10.1177/1471301216639463] [Medline: 27170583]
37. Robinson L, Hutchings D, Dickinson HO, Corner L, Beyer F, Finch T, et al. Effectiveness and acceptability of non-pharmacological interventions to reduce wandering in dementia: a systematic review. Int J Geriatr Psychiatry 2007 Jan;22(1):9-22. [doi: 10.1002/gps.1643] [Medline: 17096455]
38. Adekoya AA, Guse L. Wandering behavior from the perspectives of older adults with mild to moderate dementia in long-term care. Res Gerontol Nurs 2019 Sep 01;12(5):239-247 [FREE Full text] [doi: 10.3928/19404921-20190522-01] [Medline: 31158296]
39. Adekoya A. Tensions between risk, autonomy, and health care settings. In: Liu L, Daum C, Neubauer N, Miguel Cruz A, Rios Rincón A, editors. Autonomy and Independence: Aging in an Era of Technology. New York, NY: Morgan & Claypool Publishers; 2022:69-70.
40. Hillman A, Latimer J. Cultural representations of dementia. PLoS Med 2017 Mar 28;14(3):e1002274 [FREE Full text] [doi: 10.1371/journal.pmed.1002274] [Medline: 28350889]
41. Benbow SM, Jolley D. Dementia: stigma and its effects. Neurodegen Dis Manage 2012 Apr 16;2(2):165-172. [doi: 10.2217/nmt.12.7]
42. Rewerska-Juśko M, Rejdak K. Social stigma of people with dementia. J Alzheimers Dis 2020;78(4):1339-1343. [doi: 10.3233/JAD-201004] [Medline: 33185610]
43. Dong X, Chen R, Simon MA. Elder abuse and dementia: a review of the research and health policy. Health Aff (Millwood) 2014 Apr;33(4):642-649 [FREE Full text] [doi: 10.1377/hlthaff.2013.1261] [Medline: 24711326]
44. Storey JE. Risk factors for elder abuse and neglect: a review of the literature. Aggress Violent Behav 2020 Jan;50:101339. [doi: 10.1016/j.avb.2019.101339]
45. Elements of consent. Information and Privacy Commissioner of Ontario. URL: https://www.ipc.on.ca/part-x-cyfsa/consent-and-capacity/elements-of-consent/#_edn2 [accessed 2023-09-18]
46. Poppe C, Elger BS, Wangmo T, Trachsel M. Evaluation of decision-making capacity in patients with dementia: challenges and recommendations from a secondary analysis of qualitative interviews. BMC Med Ethics 2020 Jul 06;21(1):55 [FREE Full text] [doi: 10.1186/s12910-020-00498-y] [Medline: 32631304]
47. Trachsel M, Hermann H, Biller-Andorno N. Cognitive fluctuations as a challenge for the assessment of decision-making capacity in patients with dementia. Am J Alzheimers Dis Other Demen 2015 Jun 18;30(4):360-363 [FREE Full text] [doi: 10.1177/1533317514539377] [Medline: 24946784]

Edited by X Xiang; submitted 29.07.24; peer-reviewed by L Haslam-Larmer; comments to author 03.12.24; revised version received 16.12.24; accepted 20.12.24; published 07.02.25.

Please cite as:
Adekoya A, Daum C, Neubauer N, Miguel-Cruz A, Liu L
Implications of Public Disclosure of Personal Information in a Mobile Alert App for People Living With Dementia Who Go Missing: Qualitative Descriptive Study
JMIR Aging 2025;8:e64847
URL: https://aging.jmir.org/2025/1/e64847
doi:10.2196/64847
PMID:39918846

©Adebusola Adekoya, Christine Daum, Noelannah Neubauer, Antonio Miguel-Cruz, Lili Liu. Originally published in JMIR Aging (https://aging.jmir.org), 07.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Transcultural Adaptation, Validation, Psychometric Analysis, and Interpretation of the 22-Item Thai Senior Technology Acceptance Model for Mobile Health Apps: Cross-Sectional Study

Nida Buawangpong[1], MD; Penprapa Siviroj[2,3], DrPH; Kanokporn Pinyopornpanish[1], MD, PhD; Wachiranun Sirikul[2,4,5], MSc, MD

[1]Department of Family Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[2]Department of Community Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[3]Environmental and Occupational Medicine Excellence Center, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[4]Center of Data Analytics and Knowledge Synthesis for Health Care, Chiang Mai University, Chiang Mai, Thailand

[5]Department of Biomedical Informatics and Clinical Epidemiology, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

Corresponding Author:
Wachiranun Sirikul, MSc, MD
Department of Community Medicine
Faculty of Medicine
Chiang Mai University
110 Intrawarorot road
Sriphum, Mueng
Chiang Mai, 50200
Thailand
Phone: 66 53935472
Email: wachiranun.sir@cmu.ac.th

Abstract

Background: The rapid advancement of technology has made mobile health (mHealth) a promising tool to mitigate health problems, particularly among older adults. Despite the numerous benefits of mHealth, assessing individual acceptance is required to address the specific needs of older people and promote their intention to use mHealth.

Objective: This study aims to adapt and validate the senior technology acceptance model (STAM) questionnaire for assessing mHealth acceptance in the Thai context.

Methods: In this cross-sectional study, we adapted the original, 38-item, English version of the STAM using a 10-point Likert scale for mHealth acceptability among the Thai population. We translated the mHealth STAM into Thai using forward and backward translation. A total of 15 older adults and experts completed the pilot questionnaire and were interviewed to assess its validity. The pilot items of the Thai mHealth STAM were then reworded and revised for better comprehension and cross-cultural compatibility. The construct validity of the Thai mHealth STAM was evaluated by a multidimensional approach, including exploratory and confirmatory factor analysis and nonparametric item response theory analysis. Discriminative indices consisting of sensitivity, specificity, and area under the receiver operating characteristic (AUROC) were used to determine appropriate banding and discriminant validity for the intention to use mHealth. Internal consistency was assessed using Cronbach α and McDonald ω coefficients.

Results: Out of the 1100 participants with a mean age of 62.3 (SD 8.8) years, 360 (32.7%) were adults aged 45-59 years, and 740 (67.3%) were older adults aged 60 years and older. Of the 40-item pilot questionnaire, exploratory factor analysis identified 22 items with factor loadings >0.4 across 7 principal components, explaining 91.45% of the variance. Confirmatory factor analysis confirmed that 9-dimensional sets of 22 items had satisfactory fit indices (comparative fit index=0.976, Tucker-Lewis index=0.968, root mean square error of approximation=0.043, standardized root mean squared residual=0.044, and R^2 for each item>0.30). The score banding D (low≤151, moderate 152-180, and high≥181) was preferred as the optimal 22-item Thai mHealth STAM cutoff score based on the highest sensitivity of 89% (95% CI 86.1%-91.5%) and AUROC of 72.4% (95% CI 70%-74.8%) for predicting the intention to use mHealth. The final Thai mHealth STAM, consisting of 22 items, exhibited remarkable internal consistency, as evidenced by a Cronbach α of 0.88 (95% CI 0.87-0.89) and a McDonald ω of 0.85 (95% CI 0.83-0.87). For all 22 items, the corrected item-total correlations ranged between 0.26 and 0.71.

XSL•FO
RenderX

Conclusions: The 22-item Thai mHealth STAM demonstrated satisfactory psychometric properties in both validity and reliability. The questionnaire has the potential to serve as a practical questionnaire in assessing the acceptance and intention to use mHealth among pre-older and older adults.

(JMIR Aging 2025;8:e60156) doi:10.2196/60156

KEYWORDS

STAM; senior technology acceptance model; validity; reliability; mHealth; older adult; technology acceptance; mobile health; app; transcultural adaptation; psychometric analysis; geriatrics; cross-sectional study; Thai; theory analysis; Cronbach α; McDonald ω; quality of life; well-being; social media; telehealth; health informatics; eHealth; mobile phone

Introduction

As the global population ages, the integration of technology into the lives of older adults becomes increasingly crucial for enhancing their quality of life, independence, and well-being [1]. An emerging technology that promotes healthy aging is mobile health (mHealth). mHealth refers to medical and public health services facilitated by mobile devices [2]. It can provide individualized care plans for older adults to sustain functional ability and enhance quality of life [3]. Examples of mHealth innovations for older adults include supporting services for age-friendly health and facilitating the establishment of behavioral changes [3,4]. However, the adoption of technology, for example, mHealth, among older adults remains a complex and multifaceted issue, influenced by various factors such as individual perception and experience, ease of use, technological support, and sociocultural contexts [5,6]. To address this challenge, numerous theoretical frameworks have been proposed to understand and predict older adults' acceptance of technology.

Assessing technology acceptance is essential for the successful implementation and use of mHealth technologies, as it directly influences user engagement, health outcomes, and health care delivery efficiency. Understanding acceptance helps developers create user-friendly applications [7,8], improves health outcomes through better adherence to interventions [9,10], and guides implementation strategies to address barriers effectively [11,12]. It also informs policy makers and administrators, enabling evidence-based decisions on mHealth investments [13,14]. Therefore, the lack of validated questionnaires for assessing technology acceptance could lead to a limited understanding of user needs and missed opportunities for improvement. Addressing this gap by developing and validating robust assessment is critical for maximizing the benefits of mHealth technologies and ensuring their effective adoption across diverse populations.

In the field of mHealth, various instruments and frameworks have been developed to assess adoption, intention to use, and acceptance. Established instruments like the Health Information Technology Usability Evaluation Scale (Health-ITUES) [15,16], System Usability Scale (SUS) [17,18], and mHealth App Usability Questionnaire (MAUQ) [19,20] provided insights focusing on user experiences and satisfaction. Broader frameworks include the unified theory of acceptance and use of technology (UTAUT) [21,22], which was extended to include additional factors relevant to mHealth, such as trust and perceived reliability, and was used in various studies to predict mHealth acceptance; the Fit between Individuals, Tasks, and Technology (FITT) Framework [23] is another, which was introduced to measure acceptance in clinical environments, emphasizing the alignment between user needs and technology capabilities. Despite their usability, these instruments and frameworks often lack specificity when addressing the unique needs of older adults.

The senior technology acceptance model (STAM) [24] stands out due to its tailored approach for older adults, which addresses their unique challenges and enhances the relevance of mHealth technologies for this population, making it more relevant than general models like the technology acceptance model (TAM) [25] or the UTAUT [26]. Furthermore, it emphasizes the role of social influence and support, which are critical for older adults who may rely on family and caregivers for technological adoption and addressing common health conditions in older adults, such as cognitive load and physical limitations.

The STAM was first proposed by Chen and Chan [24] in 2014 and has gained prominence for its focus on the unique needs and characteristics of older adults. This model was developed based on a study of 1012 older adults aged 55 years and older in Hong Kong, and it specifically targets older adults as its primary population of interest. The STAM integrates concepts from established technology acceptance frameworks, such as the TAM [25] and the UTAUT [26], tailored to address the specific considerations of older adults and provides a thorough framework for studying the factors that influence technological adoption in this age group. The study indicated 8 dimensions associated with technology acceptance in older adults, which included gerontechnology self-efficacy, gerontechnology anxiety, facilitating conditions, self-reported health conditions, cognitive ability, social relationships, attitude toward life and satisfaction, and physical functioning. Sociodemographic factors such as age, gender, education, and economic status are taken into account [24].

While the STAM has been used in different cultural contexts in other Asian countries, including Hong Kong [24] and South Korea [27], its applicability to the Thai population has not been validated. Thailand, like many other countries, is experiencing rapid population aging, emphasizing the urgency of understanding and promoting health technology acceptance among older adults [28]. However, cultural background, social norms, and technological infrastructures specific to Thailand may influence older adults' perceptions and behaviors toward technology differently than in other contexts. Therefore, this study aimed to adapt, validate, and define the interpretation of the STAM questionnaire for evaluating the acceptance and intent to use mHealth in Thailand.

XSL•FO
RenderX

Methods

Study Design and Study Population

The cross-sectional study was conducted from August 2022 to July 2023 through a nationwide, web-based survey and a community survey. Eligible criteria for the study were Thai citizens aged 45 years and older on the date of the survey who could read and communicate in the Thai language and had no underlying conditions or diseases that limited their ability to complete the survey or use mHealth apps (eg, dementia, active psychological problems, or severe visual problems). The web-based survey was disseminated through an assortment of social media platforms, including the department websites, Facebook, Line, Twitter (rebranded as X in 2023), and Instagram. The information on community survey setting and recruitment is described in the section below. For the survey data collection, the respondents to both the web-based and community surveys used the Research Electronic Data Capture (REDCap; Vanderbilt University) survey platform to self-complete the questionnaires. REDCap [21,22] is a secure, web-based software platform designed to support data capture for research studies, providing (1) an intuitive interface for validated data capture, (2) audit trails for tracking data manipulation and export procedures, (3) automated export procedures for seamless data downloads to common statistical packages, and (4) procedures for data integration and interoperability with external sources. All study data were collected and managed using REDCap tools hosted at the Faculty of Medicine, Chiang Mai University. All respondents provided their informed consent, which was included in the screening questionnaire and study information sheet, before participating in this survey. The study excluded incomplete respondents or participants who spent less than 2 minutes or more than 60 minutes on the survey. This study was reported in accordance with COSMIN (Consensus-Based Standards for the Selection of Health Status Measurement Instruments)

reporting guidelines for studies on measurement properties of patient-reported outcome measures [29].

Community Survey Setting and Recruitment

The community survey was distributed by the investigator team, consisting of medical students and health care personnel at primary care units from 10 subdistricts in Chiang Mai province. To identify eligible participants in the target area, officers from the subdistrict primary care units reviewed periodic health survey data for community-dwelling adults aged 45 years and older. Subsequently, patients' information was verified with the health-promoting hospital databases to exclude individuals with health conditions that impeded survey participation or mHealth use as described above. The subdistrict primary health care teams invited eligible individuals to participate in the study through individual contact by community health care volunteers, community radio announcements, and posters displayed at primary care units.

Ethical Considerations

The ethical consideration of the human subject research was approved by the Institutional Review Board of the Faculty of Medicine, Chiang Mai University (COM-2565-09079). All respondents provided their informed consent, as outlined in the screening questionnaire and study information page, before participating in this survey. For the web-based survey, respondents remained anonymous, and no identification data were recorded. In the case of the community survey, identification data of eligible participants were used solely for recruitment purposes within each target area and were not recorded in either the survey form or the study database. Participants received 100 Thai Baht (US $3) as compensation for answering the questionnaires.

Translation and Adaptation of the Thai mHealth STAM

The original, English, 38-item STAM is a 10-point Likert scale consisting of 10 subscales and 38 items that capture the acceptance of general technology use for the older adult population. The total ranges from 38 to 380 points, with a higher score indicating greater acceptance of technology. The validity and reliability of 38-item STAM have been established on a satisfactory scale in 1012 older adults aged 55 years and older in Hong Kong [24]. The construct validity of the STAM was also evaluated with the confirmatory factor analysis (CFA) and revealed a satisfactory model fit with the proposed structure (comparative fit index [CFI]=0.938, root mean square error of approximation [RMSEA]=0.054, and standardized root mean square residual [SRMR]=0.075). The reliability of each subscale with Cronbach α coefficients ranged from 0.67 to 0.95.

Translation and adaptation of the Thai mHealth STAM was performed in accordance with the second edition of the International Test Commission (ITC) Guidelines for Translating and Adapting Tests [30]. In accordance with the ITC precondition guidelines, permission from the holder of the intellectual property rights relating to the 38-item STAM was obtained before performing any translation and adaptation of the STAM. The forward and backward translation with an expert reconciliation design was performed as recommended by the ITC test development guidelines. Before beginning the forward translation process, we decided to include a new subscale, perceived barriers, in the Thai STAM version due to the findings from the previous scoping review [31] on adopting mobile apps for health-related interventions among older adults. It revealed that barriers to adopting mHealth apps among older adults were the most common topics identified in the included studies. Insufficient technological skills, perceived lack of capability and time, concerns regarding personal data privacy, and trust in mHealth providers were the four items comprising the perceived barriers subscale. Following the translation protocol, the original, English, 38-item STAM was adapted to specify mHealth apps in all items and then forward translated into Thai by a professional translator to ensure accuracy for the target audience. The expert panel, which included a digital health expert (family physician and epidemiologist), 2 gerontology physicians, and a public health expert in community medicine, reviewed the forward translation of the Thai STAM questionnaire to ensure readability and transcultural adaptation. The backward translation was done by another professional translator into English. Then, the expert panel reconciled the backward translation version with the original STAM version.

XSL•FO
RenderX

The investigator's team resolved any discrepancies by reaching a final consensus through discussions with the expert panel. To ensure the face and content validity of the proposed questionnaire, a literature review, an expert review, and public interviews were incorporated into the adaptation of the Thai mHealth STAM. In total, 15 older adults participated in this phase to complete the pilot 40-item Thai STAM. Participants were subsequently interviewed to assess the following: overall questionnaire readability, clarity of instructions and items/response options, comprehension of the questionnaire, and other feedback on each item. Then, the pilot 40-item Thai STAM was reworded and revised as recommended on input from both participants and expert interviews. Finally, the pilot 40-item Thai mHealth STAM was given to a group of 40 older adults to verify its reliability and scale usability.

Sample Size Estimation

The sample size was estimated based on three parameters, which are as follows: (1) a stable structure for an exploratory factor analysis (EFA) based on the rule of thumb, which is 10 cases per question; (2) expected CFI for a CFA based on the structural equation modeling; and (3) expected Cronbach α for the internal consistency of the questionnaire. For the first parameter, according to the rule of thumb, at least 440 respondents, accounting for 10% of the dropout rate, were required for an EFA. To achieve the expected CFI of 0.95 for a CFA, at least 459 respondents, accounting for 10% of the dropout rate, were required based on an average factor loading of 0.60 and an average factor correlation of .30 to ensure a .05 α (type I) error and power of 90% [32]. For testing overall reliability, at least 146 total respondents were required based on expected Cronbach α=0.80 (SD .05), a confidence level of 95%, and a dropout rate of 10% [33]. All sample size estimation was performed by the web-based sample size calculator [34]. Finally, the minimal required sample size for this study was 920, which was divided into 460 each for the EFA and CFA, respectively.

Statistical Analysis

Descriptive Analysis

All statistical analyses were conducted using Stata (version 17.0; StataCorp). A *P* value below .05 indicated statistical significance. Categorical data were presented as frequency and percentage, while continuous data were described using mean (SD). Univariable analysis for comparison was performed as appropriate. The Thai mHealth STAM item scores were summarized with central estimations, measures of variability, floor and ceiling effect, skewness, and kurtosis tests. The overall psychometric properties of the Thai mHealth STAM were evaluated for validity and reliability as follows:

Dimensionality

To explore and reduce the dimensionality of the proposed questionnaire, an EFA was performed using a principal component analysis (PCA). The selection of PCA over common factor analysis was based on its ability to enhance parsimony and aid in the selection of factors for CFA [35]. Communalities were initially evaluated, and then orthogonal rotation with the varimax criteria and oblique rotation with promax criteria of the component was conducted. The Kaiser-Meyer-Olkin (KMO) measure and the Bartlett test of sphericity were conducted to verify the appropriateness of using factor analysis. A KMO value greater than 0.8 [36] and a Bartlett test with a *P* value less than .05 [37] are suggested for assessing sample adequacy and the suitability of the data for factor analysis, respectively. Eigenvalues greater than 1, the cumulative percentage of variance, and the scree plot with the number of factors that explained more than 5% of the variance were used to determine the number of factors to be retained [38,39]. A parallel study was conducted to validate the optimal threshold for the number of included factors [40]. Then, we used the following criteria to evaluate the adequacy of the EFA results. First, each should be saliently loaded with at least three items to ensure reliability and stability. In case a factor contains only 2 items, the expert panel consensus will be reached to ensure that the factor is meaningful based on the context and theoretical basis. Second, each item should load saliently on only 1 factor without complex or cross-loadings. Third, each factor should demonstrate internal consistency reliability ≥0.70. Fourth, all factors should be theoretically meaningful [35,41,42].

Construct Validity

For a CFA, structural equation modeling using a maximum likelihood estimation was performed to assure the factor structure based on the exploratory factor, as described previously. To determine the appropriateness of the proposed model, the specific fit indices were evaluated as follows: RMSEA<0.100, SRMR<0.100, CFI>0.900, and Tucker-Lewis Index (TLI)>0.900 [43-45]. To establish acceptance of the final structure of the final model, the coefficient of determination (R^2) and item-scale correlation (standardized factor loading) should be at least 0.30 and 0.40, respectively. Finally, a nonparametric item response theory (IRT) analysis was done to confirm that the final Thai mHealth STAM had the unidimensional set for the relationship between the latent trait and the responses to the items [46]. The IRT analysis was assessed based on fundamental assumptions, including unidimensionality, local independence, and monotonicity. Loevinger H coefficients (H^s) less than 0.3, between 0.3 and 0.4, and greater than 0.4, as determined by the item traces, correspond to poor, medium, and strong scalability properties, respectively. The monotonicity assumption criterion was determined by a critical value of less than 80.

Discriminant Validity

To determine the discriminant validity of the final questionnaire, the intention to use mHealth, as indicated in the external question, "If there are available mHealth applications for you, do you want to use them? (yes/no)," was used as the anchor-based question. The discriminative indices, including sensitivity, specificity, and area under the receiver operating characteristic (AUROC), were used with the intention of determining the appropriate cutoff scores. The 6 proposed bandings for the Thai mHealth STAM scores are categorized into low, moderate, and high acceptance based on score tertiles. Associations between these bandings and the intention to use mHealth are presented by adjusted odds ratios (aORs) with 95% CI from a multivariable logistic regression adjusted for potential

XSL•FO
RenderX

confounders such as age, gender, education, income, and living alone.

Reliability

To estimate the correlation statistics for reliability, 95% CI using 1000 bootstrap resampling was presented alongside the reported correlation statistics. An internal consistency consisting of Cronbach α and McDonald ω coefficients was calculated for each item of the final questionnaire, as well as the entirety of the final questionnaire, to determine internal consistency, reliability, and the degree to which every item on a scale measures the same construct. The values of at least .70 indicated acceptable reliability of the questionnaire [47]. In addition, the item-total correlations and the corrected item-total correlations between .20 and .80 were considerably acceptable. A subgroup analysis of adults aged 45-59 years and adults aged 60 years and older was also performed, recognizing the importance of understanding the unique health needs and challenges faced by both current older populations and those who will age into this group in the future.

Results

Findings From the Translation and Adaptation of the Thai mHealth STAM

After reviewing the forward translation, the panel of experts decided to remove 2 items from the gerontechnology self-efficacy subscale, as they were redundant with the facilitating condition (FC) subscale (FC1 and FC2). Independent back-translation provided an additional check of the semantic equivalence of the translation. A total of 4 items, including PU2, PEOU2, P4, and P8, were modified based on the backward translation. For face and content validity, we conducted interviews with 15 older adults similar to the target population. Based on participants' feedback, 4 items (FC1, FC2, C4, and P2) were slightly modified for clarity. In addition, 2 gerontology experts suggested rephrasing 2 items (A1 and A2) regarding attitude to aging and life satisfaction due to the sensitive wording. Finally, the 40-item Thai mHealth STAM in the pilot group of 40 older adults indicated acceptable internal consistency (Cronbach α =0.91). The details of the full 40 items (10 dimensions) of the Thai mHealth STAM are presented in Table S1 in Multimedia Appendix 1.

Participant Characteristics

From the total of 1100 participants, the mean age was 62.3 (SD 8.8) years. The majority of participants were female (776/1100, 70.5%). Among the 1100 participants, 360 (32.7%) were adults aged 45-59 years, and 740 (67.3%) were older adults aged 60 years and older. Statistically significant differences in the characteristics between adults and older adults were observed in marital status (P=.003), education levels (P<.001), income (P<.001), underlying diseases (P<.001), and technology experience (P<.001). The characteristics of the participants of the study population are presented in Table 1. The derived data were randomly divided in a 1:1 ratio into 2 datasets in preparation for the EFA and CFA. The characteristics of the participants involved in the EFA and CFA are described in Table S2 in Multimedia Appendix 1.

XSL•FO
RenderX

Table 1. Participant characteristics of the study population.

Characteristics	Total (N=1100)	Adults (n=360)	Older adults (n=740)	*P* value
Age (year), mean (SD)	62.3 (8.8)	52.4 (5.5)	67 (5.5)	<.001
Male, n (%)	324 (29.5)	99 (27.5)	225 (30.4)	.32
Marital status, n (%)				.003
Single	96 (8.7)	43 (11.9)	53 (7.2)	
Married	747 (67.9)	250 (69.4)	497 (67.2)	
Separated, divorced, or widowed	257 (23.4)	67 (18.6)	190 (25.7)	
Education levels, n (%)				<.001
No education	18 (1.6)	1 (0.3)	17 (2.3)	
Primary school	725 (65.9)	160 (44.4)	565 (76.4)	
Secondary school	97 (8.8)	50 (13.9)	47 (6.4)	
High school and vocational training	162 (14.7)	99 (27.5)	63 (8.5)	
Pre-university	11 (1)	6 (1.7)	5 (0.7)	
Bachelor’s degree	79 (7.2)	41 (11.4)	38 (5.1)	
Master’s degree	8 (0.7)	3 (0.8)	5 (0.7)	
Income (THB[a]), n (%)				
<10,000	948 (86.2)	275 (76.4)	673 (90.9)	<.001
10,001-30,000	138 (12.5)	79 (21.9)	59 (8)	
>30,001	14 (1.3)	6 (1.7)	8 (1.1)	
Living status, n (%)				.91
Alone	108 (9.8)	34 (9.4)	74 (10)	
With family	988 (89.8)	325 (90.3)	663 (89.6)	
With others	4 (0.4)	1 (0.3)	3 (0.4)	
Living area, n (%)				.49
Urban	220 (20)	78 (21.7)	142 (19.2)	
Sub-urban	377 (34.3)	116 (32.2)	261 (35.3)	
Rural	503 (45.7)	166 (46.1)	337 (45.5)	
Had any underlying disease, n (%)	726 (66)	189 (52.5)	537 (72.6)	<.001
Hypertension, n (%)	495 (45)	115 (31.9)	380 (51.4)	<.001
Dyslipidemia, n (%)	375 (34.1)	87 (24.2)	288 (38.9)	<.001
Diabetes mellitus, n (%)	184 (16.7)	55 (15.3)	129 (17.4)	.37
Chronic kidney disease, n (%)	17 (1.5)	3 (0.8)	14 (1.9)	.18
Vision problems, n (%)	612 (55.6)	208 (57.8)	404 (54.6)	.32
Wore glasses or contact lens, n (%)	399 (65.2)	144 (69.2)	255 (63.1)	.13
Hearing problems, n (%)	120 (10.9)	15 (4.2)	105 (14.2)	<.001
Used hearing aids, n (%)	4 (3.3)	0 (0)	4 (3.8)	.44
Had experience using a smartphone or tablet, n (%)	873 (79.4)	332 (92.2)	541 (73.1)	<.001
Had own smartphone, n (%)	843 (76.6)	317 (88.1)	526 (71.1)	<.001
Had own tablet, n (%)	20 (1.8)	12 (3.3)	8 (1.1)	.009
Had experience using the internet, n (%)	784 (71.3)	323 (89.7)	461 (62.3)	<.001
Had experience in using mHealth[b] apps, n (%)	439 (50.3)	205 (61.7)	234 (43.3)	<.001
Intention to use mHealth apps, n (%)	537 (48.8)	220 (61.1)	317 (42.8)	<.001

XSL•FO
RenderX

[a]THB 1=US $0.0296195.

[b]mHealth: mobile health.

Dimensionality

According to the item analysis, we excluded 6 items from the physical function subscale (P2, P3, P4, P5, P7, and P8) due to a floor effect or ceiling effect of >80% (Table S3 in Multimedia Appendix 1). An EFA was conducted using PCA with 34 remaining items. The Bartlett test of sphericity obtained $P<.001$, indicating that the correlation matrix was not random [37]. The KMO statistic was 0.875, well above the minimum standard for conducting factor analysis [36]. Therefore, we determined that the input data were appropriate for EFA. Subsequently, the rotation of principal components was performed using both orthogonal rotation (varimax) and oblique rotation (promax) in an attempt to achieve a simple structure. Given the fact that an oblique rotation is generally recommended by measurement specialists to facilitate the emergence of factor intercorrelations [48-50], almost all social sciences measurements exhibit some degree of correlation [51]. In addition, the correlation matrix for the factors with oblique (promax) rotation indicated that the highest correlation was 0.445 (Table S3 in Multimedia Appendix 1); we thereby determined that the factors were correlated, and hence, oblique rotation was an appropriate approach. The results of parallel analysis (Table S4 in Multimedia Appendix 1) and PCA with or without oblique (promax) rotation all recommended the retention of 7 factors. According to the previous criteria, 2-item factors were identified, including factor 4 (PBR1 and PBR2) and factor 6 (S1 and S2). The internal consistency of seven factors demonstrated Cronbach α of 0.884 with 95% CI (0.875-0.894), which met acceptable thresholds. Within the context and theoretical framework of the STAM [24] and the UTAUT [11,21,26,52], social factors significantly influence behavioral intentions to use technology, particularly in the use of mHealth. Perceived barriers also play a role in determining intentions to use mHealth, as demonstrated by the aforementioned scoping review [31]. The inclusion of factor 4 and factor 6, which represented perceived barriers and social relationships, was considered appropriate. Based on the priori criteria and consensus of the panel experts, the EFA identified 22 candidate items (ATT1, ATT2, PU1, PU2, PU3, PEOU1, PEOU2, PB1, PB2, ANX1, ANX2, FC2, FC4, FC5, H1, H2, H5, C2, C3, C4, S1, and S2) with factor loadings greater than 0.4 that encompassed the 7 factors. The final EFA result is presented in Table 2.

XSL•FO
RenderX

Table 2. Exploratory factor analysis of the final 22-item Thai mobile health (mHealth) senior technology acceptance model (STAM).

Items	Factor loadings[a]							Communality value
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	
ATT1	0.639[b]	0.096	0.057	0.027	0.040	0.057	0.060	0.824
ATT2	0.631[b]	0.059	0.060	0.036	0.058	0.071	0.057	0.822
PU1	0.899[b]	0.061	0.030	0.065	0.053	0.040	0.032	0.825
PU2	0.921[b]	0.031	0.027	0.081	0.036	0.054	0.099	0.878
PU3	0.919[b]	0.046	0.015	0.079	0.046	0.056	0.072	0.873
PEOU1	0.473[b]	0.188	0.074	0.298	0.055	0.007	0.175	0.677
PEOU2	0.513[b]	0.169	0.108	0.281	0.080	0.019	0.175	0.746
PBR1	0.115	0.203	0.071	0.854[b]	0.025	0.000	0.038	0.795
PBR2	0.141	0.200	0.061	0.855[b]	0.018	0.009	0.069	0.811
PBR3	0.031	0.880[b]	0.025	0.120	–0.017	0.003	0.032	0.806
PBR4	0.073	0.894[b]	0.012	0.109	–0.016	0.023	0.058	0.831
ANX1	0.130	0.720[b]	0.061	0.183	0.015	0.019	0.044	0.734
ANX2	0.107	0.643[b]	0.081	0.231	0.030	0.031	0.055	0.688
FC2	0.324	0.088	0.032	0.103	0.034	0.054	0.579[b]	0.486
FC4	0.380	0.171	0.147	0.237	0.064	–0.009	0.449[b]	0.606
FC5	0.330	0.201	0.020	0.124	0.032	0.079	0.563[b]	0.504
H1	0.074	–0.020	0.172	0.007	0.723[b]	0.100	0.017	0.573
H2	0.092	–0.009	0.129	0.042	0.713[b]	0.097	0.009	0.556
H5	0.127	–0.037	0.297	0.098	0.505[b]	0.112	0.072	0.498
C1	0.009	0.040	0.614[b]	0.081	0.221	0.066	0.026	0.497
C2	0.163	0.086	0.611[b]	0.164	0.177	0.117	0.073	0.582
C3	0.019	0.054	0.693[b]	0.067	0.144	0.129	0.024	0.547
C4	0.043	0.065	0.624[b]	0.068	0.106	0.237	0.024	0.497
S1	0.126	0.040	0.226	0.021	0.147	0.669[b]	0.024	0.548
S2	0.142	0.034	0.162	–0.012	0.116	0.678[b]	0.040	0.532
% of variance[b]	26.41	15.99	12.37	11.77	10.76	7.32	6.83	—[c]
Cumulative % of variance[b]	26.41	42.40	54.77	66.54	77.30	84.60	91.45	—

[a]The extraction method was principal component analysis, with the rotation method by oblique, promax rotation.

[b]Items load on the assigned factor loadings >0.4 are highlighted.

[c]Not applicable.

Construct Validity

From the EFA, the 22 items of the 7-factor Thai mHealth STAM explained 91.45% of the variance. The unidimensionality of each factor (subscale) and the overall models were assessed by analyzing modification indices in the CFA. Of the 7 factors from the EFA, the CFA of each factor (subscale) showed that only 5 factors, consisting of cognitive ability (C2, C3, C4), perceived barriers (PB1 and PB2), facilitating conditions (FC2, FC4, and FC5), self-reported health conditions (H1, H2, and H5), and social relationships (S1 and S2), showed satisfactory information criteria indices of the CFA, as presented Table 3. Factor 1, which included items from the attitude toward using (ATT1 and ATT2), perceived usefulness (PU1, PU2, and PU3),

XSL•FO
RenderX

and perceived ease of use (PEOU1 and PEOU2), did not meet the CFA criteria due to over-factoring issues (Table S5 in Multimedia Appendix 1). Factor 2, combining perceived barriers (PB3 and PB4) and gerontechnology anxiety (ANX1 and ANX2), was unfit according to CFA criteria, with a low CFI (0.799), low TLI (0.698), and high RMSEA (0.306, 90% CI 0.293-0.319). Attempts to combine subscales also did not meet CFA criteria (Table S5 in Multimedia Appendix 1). However, when items were separated as in the original STAM, including attitude toward using (ATT1 and ATT2), perceived usefulness (PU1, PU2, and PU3), perceived ease of use (PEOU1 and PEOU2), and gerontechnology anxiety (ANX1 and ANX2), these separated factors showed a good fit with CFA criteria (Table S5 in Multimedia Appendix 1). Out of the 9 factors from the single latent factor analysis, 5 were 2-item factors. These were kept in the final CFA model because 3 factors (attitude toward using, perceived ease of use, and gerontechnology anxiety) were originally designed as 2-item factors, similar to the original STAM. The perceived barriers and social relationships were also retained because of their contextual relevance, as described above. Finally, the CFA confirmed 9-dimensional sets of 22 items with satisfactory fit indices, as shown in Table 3. The details of the CFAs of evaluated and reevaluated models are described in Table S4 in Multimedia Appendix 1.

A nonparametric IRT analysis also affirmed the unidimensionality, local independence, and monotonicity of the 22-item model with 8 factors (Table S6 in Multimedia Appendix 1). For the scalability, all 22 items of the Thai mHealth STAM had H^s coefficients over 0.4, which indicates medium to strong scalability properties (Table S6 in Multimedia Appendix 1). The correlation among the final 22-item Thai mHealth STAM subscales ranged from 0.040 to 0.685 (Table S7 in Multimedia Appendix 1). The final 22-item Thai mHealth STAM questions, along with the English version and modeling indices, are described in Table 4.

Each item is scored on a 10-point Likert scale from 1 (very unsatisfied or strongly disagree) to 10 (very satisfied or strongly agree), with reverse scaling for perceived barriers and gerontechnology anxiety.

Table 3. Confirmatory factor analysis of the final Thai mobile health (mHealth) senior technology acceptance model (STAM).

Factor	Number of items	Threshold for acceptable fit					Model fit
		CFI[a] (>0.90)	TLI[b] (>0.90)	RMSEA[c] (<0.10 [90% CI])	SRMR[d] (<0.10)	R^2 (>0.30)	
Attitude toward using	2 items (ATT1 and ATT2)	1.000	1.000	<0.001 (<0.001 to <0.001)	<0.001	All >0.30	Acceptable
Perceived of benefits	3 items (PU1, PU2, and PU3)	1.000	1.000	<0.001 (<0.001 to <0.001	<0.001	All >0.30	Acceptable
Perceived ease of use	2 items (PEOU1 and PEOU2)	1.000	1.000	<0.001 (<0.001 to <0.001)	<0.001	All >0.30	Acceptable
Perceived of barriers	2 items (PB1 and PB2)	1.000	1.000	<0.001 (<0.001 to <0.001)	<0.001	All >0.30	Acceptable
Gerontechnology anxiety	2 items (ANX1 and ANX2)	1.000	1.000	<0.001 (<0.001 to <0.001)	<0.001	All >0.30	Acceptable
Facilitating conditions	3 items (FC2, FC4, and FC5)	1.000	1.000	<0.001 (<0.001 to <0.001)	<0.001	All >0.30	Acceptable
Self-reported health conditions	3 items (H1, H2, and H5)	1.000	1.000	<0.001 (<0.001 to <0.001)	<0.001	All >0.30	Acceptable
Cognitive ability	3 items (C2, C3, and C4)	1.000	1.000	<0.001 (<0.001 to <0.001)	<0.001	All >0.30	Acceptable
Social relationships	2 items (S1 and S2)	1.000	1.000	<0.001 (<0.001 to <0.001)	<0.001	All >0.30	Acceptable
Final Thai mHealth STAM 9-dimensional model	22 items	0.976	0.968	0.043 (0.039 to 0.047)	0.044	All >0.30	Acceptable

[a]CFI: comparative-fit index.

[b]TLI: Tucker-Lewis Index.

[c]RMSEA: root mean square error of approximation.

[d]SRMR: standardized root mean squared residual.

XSL•FO
RenderX

Table 4. The final 22-item Thai mobile health (mHealth) senior technology acceptance model (STAM).

Items and questions		Mean (SD)	Ceiling, %	Floor, %	Skewness	Kurtosis	Standardized factor loading (95% CI)	R^2
Attitude toward use								
ATT1	Using mobile health applications is a good idea.	8.18 (2.30)	48.45	2.50	–1.11	3.45	0.94 (0.92-0.95)	0.875
ATT2	You like the idea of using mobile health applications.	8.02 (2.39)	45.85	3.11	0.34	2.25	0.93 (0.91-0.95)	0.862
Perceived usefulness								
PU1	Using mobile health applications would enhance your effectiveness in life.	7.66 (2.59)	40.91	4.09	–0.92	2.93	0.91 (0.90-0.92)	0.831
PU2	Using mobile health applications would make your life more convenient.	7.72 (2.61)	43.08	4.09	–0.95	2.93	0.94 (0.93-0.95)	0.897
PU3	You would find mobile health applications useful in your life.	7.82 (2.62)	45.42	3.92	–1.03	3.03	0.94 (0.93-0.95)	0.891
Perceived ease of use								
PEOU1	You would find mobile health applications are easy to use.	6.37 (3.18)	28	11.64	–0.34	1.76	0.83 (0.81-0.86)	0.693
PEOU2	You could be skillful at using mobile health applications.	6.93 (3.06)	34	9	–0.62	2.10	0.89 (0.87-0.92)	0.803
Gerontechnology anxiety								
ANX1	You feel apprehensive about using mobile health applications.	5.89 (3.16)	23.56	13.2	–0.09	1.70	0.89 (0.86-0.92)	0.799
ANX2	You hesitate to use the technology for fear of making mistakes you cannot correct.	5.77 (3.11)	20.66	13.92	–0.38	1.75	0.95 (0.93-0.98)	0.918
Perceived barriers								
PB1	You need to put in a lot of effort to use mobile health applications?	4.81 (3.14)	13.27	21.42	0.34	1.79	0.89 (0.86-0.92)	0.704
PB2	You need to spend a lot of time to use mobile health applications?	5.00 (3.19)	14.78	20.35	0.26	1.70	0.96 (0.93-0.98)	0.819
Facilitating conditions								
FC2	A specific person (or group) is available for assistance with difficulties using mobile health applications.	7.45 (3.09)	44.79	10.06	–0.96	2.57	0.63 (0.59-0.68)	0.407
FC4	When you want or need to use mobile health applications, they are accessible to you.	7.20 (2.93)	36.92	7.83	–0.75	2.38	0.79 (0.76-0.83)	0.631
FC5	Your family and friends think/support that you should use mobile health applications.	6.72 (3.33)	36.62	14.04	–0.55	1.83	0.68 (0.64-0.72)	0.461
Self-reported health conditions								
H1	How are your general health conditions?	7.73 (1.75)	21.78	0.27	–0.52	2.79	0.80 (0.77-0.84)	0.651
H2	How are your health conditions compared with the same-age groups?	7.91 (2.03)	33.69	0.18	–0.69	2.54	0.77 (0.73-0.81)	0.595
H5	How well are you able to move around?	8.56 (1.96)	52.85	0.36	–1.37	4.24	0.61 (0.56-0.65)	0.370
Cognitive ability								
C2	How satisfied are you with your ability to learn new information?	7.89 (2.25)	38.38	1.16	–0.98	3.27	0.65 (0.61-0.70)	0.432
C3	How well are you able to concentrate?	8.75 (1.72)	54.23	0.18	–1.51	4.96	0.75 (0.71-0.79)	0.563

XSL•FO
RenderX

Items and questions		Mean (SD)	Ceiling, %	Floor, %	Skewness	Kurtosis	Standardized factor loading (95% CI)	R^2
C4	How satisfied are you with your ability to make decisions?	8.91 (1.58)	58.15	0.09	–1.48	4.64	0.76 (0.72-0.80)	0.578
Social relationships								
S1	How satisfied are you with your personal relationships?	9.34 (1.29)	70.99	0.09	–2.36	8.95	0.84 (0.80-0.89)	0.712
S2	How satisfied are you with the support you get from your friends and family?	9.39 (1.27)	74.01	0.09	–2.59	10.47	0.75 (0.71-0.80)	0.568
Overall (possible range 22-220)		164.16 (30.55)	—[a]	—	—	—	—	0.999

[a]Not applicable.

Discriminant Validity

Considering the absence of a reference standard, it is theoretically reasonable that more participants with higher STAM scores will result in greater acceptance and adoption of technology. The discriminative indices, including sensitivity, specificity, and AUROC, were used to determine the cutoff scores for the proposed questionnaire, considering the intention to use mHealth from the external question. The 6 proposed sets of the final 22-item Thai mHealth STAM bands were classified into low, moderate, and high acceptance, as presented in Table 5. The set D of the possible banding was preferred as the optimal 22-item Thai mHealth STAM cutoff score based on the highest sensitivity of 89% (95% CI 86.1%-91.5%) and AUROC of 72.4% (95% CI 70%-74.8%). This finding also confirmed the discrimination performance of the 22-item Thai mHealth STAM in identifying persons with and without the intention to use mHealth. For set D, low, moderate, and high scores are defined as ≤151, 152-180, and ≥181, respectively. In addition, we conducted a subgroup analysis based on age groups: pre-older adults (aged 45-59 years) and older adults (aged 60 years and older). The result revealed that the set D banding had robust discriminant validity in older adults (AUROC 73%, 95% CI 70%-76%), but the discriminant validity decreased in the pre-older adult group (AUROC 67.7%, 95% CI 63.3%-71.9%). The discriminant validity of the 22-item Thai mHealth STAM by the subpopulation cohorts is shown in Table S8 in Multimedia Appendix 1.

XSL•FO
RenderX

Table 5. Proposed sets of the final 22-item Thai mobile health (mHealth) senior technology acceptance model (STAM) bands.

Possible bandings[a]		Discriminant validity (intention to use mHealth)			
Set and band	Score	aOR[b] (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	AUROC[c] (95% CI)
Set A					
Low	≤121	Ref[d]	Ref	Ref	Ref
Moderate	122-150	2.69[e] (1.21-5.95)	98.5 (97.1-99.4)	16.0 (13.1-19.3)	57.3 (55.6-58.8)
High	≥151	15.53[e] (7.31-5.95)	89.0 (86.1-91.5)	53.5 (49.2-57.6)	71.2 (68.8-73.7)
Set B					
Low	≤131	Ref	Ref	Ref	Ref
Moderate	132- 160	4.37[e] (2.42-7.88)	97.2 (95.4-98.4)	26.1 (22.5-29.9)	61.7 (59.7-63.6)
High	≥161	15.53[e] (8.75-27.52)	78.0 (74.3-81.5)	64.8 (60.7-68.8)	71.4 (68.8-74.1)
Set C					
Low	≤141	Ref	Ref	Ref	Ref
Moderate	142- 170	4.33[e] (2.82-6.66)	93.7 (91.3-95.6)	40.9 (36.8-45.0)	67.3 (65.0-69.5)
High	≥171	13.18[e] (8.-20.26)	66.1 (61.9- 70.1)	75.8 (72.1-79.3)	71.0 (68.3-73.6)
Set D					
Low	≤151	Ref	Ref	Ref	Ref
Moderate	152-180	5.73[e] (4.01–8.19)	89.0 (86.1-91.5)	55.8 (51.6-59.9)	72.4 (70.0-74.8)
High	≥181	12.49[e] (8.45-18.47)	49.9 (45.6-54.2)	84.7 (81.5–87.6)	67.3 (64.7-69.9)
Set E					
Low	≤161	Ref	Ref	Ref	Ref
Moderate	162-190	3.55[e] (2.62-4.83)	76.7 (72.9-80.2)	65.5 (61.5-69.5)	71.1 (68.5-73.4)
High	≥191	8.46[e] (5.74-12.47)	37.4 (33.3-41.7)	90.9 (88.3-93.2)	64.2 (61.8-66.6)
Set F					
Low	≤171	Ref	Ref	Ref	Ref
Moderate	172-200	4.24[e] (3.14-5.74)	65.2 (61.0-69.2)	77.1 (73.4-80.5)	71.1 (68.4-73.8)
High	≥201	7.59[e] (4.68-12.29)	21.8 (18.4-25.5)	95.6 (93.5-97.1)	58.7 (56.7-60.6)

[a]The final 22-item Thai mHealth STAM is highlighted.

[b]aOR: adjusted odds ratio.

[c]AUROC: area under receiver operating characteristic curve.

[d]Ref: reference.

[e]All reported aORs were statistically significant with *P* value<.05. aORs were estimated using a multivariable logistic regression with adjustment for age, gender, education levels (no, primary, secondary, and university education), income levels (low: <10,000 baht, moderate: 10,000-30,000 baht, and high: >30,000 bath), and living alone.

Scale Reliability

Out of 1100 overall participants, the final 22-item Thai mHealth STAM demonstrated an excellent internal consistency in both the Cronbach α (0.88, 95% CI 0.87-0.89) and the McDonald ω coefficients (0.85, 95% CI 0.83-0.87), as shown in Table 6. By subpopulation, the Cronbach α and the McDonald ω coefficients were 0.88 (95% CI 0.86-0.90) and 0.84 (95% CI 0.81- 0.89) for adults aged 45-59 years and 0.88 (95% CI 0.86- 0.89) and 0.83 (95% CI 0.81-0.86) for older adults. All 22 items revealed the corrected item-total correlations ranging from 0.26 to 0.71, achieving a level of acceptance between 0.20 and 0.80 (Table 6).

XSL•FO
RenderX

Table 6. Reliability of the final 22-Item Thai mobile health (mHealth) senior technology acceptance model (STAM).

Items	n	Item-total correlations	Corrected item-total correlations	Average interitem correlation	Cronbach α	McDonald ω
ATT1	1158	0.652	0.605	1.682	0.877	0.879
ATT2	1156	0.652	0.603	1.675	0.877	0.879
PU1	1149	0.684	0.634	1.645	0.875	0.878
PU2	1149	0.720	0.673	1.631	0.874	0.876
PU3	1147	0.713	0.667	1.632	0.874	0.877
PEOU1	1143	0.721	0.664	1.585	0.873	0.878
PEOU2	1144	0.761	0.713	1.575	0.871	0.876
PB1	1130	0.514	0.431	1.682	0.881	0.885
PB2	1130	0.543	0.462	1.666	0.880	0.884
ANX1	1129	0.472	0.385	1.701	0.883	0.886
ANX2	1128	0.481	0.396	1.698	0.883	0.886
FC2	1123	0.546	0.469	1.671	0.880	0.884
FC4	1124	0.703	0.649	1.613	0.874	0.879
FC5	1125	0.586	0.507	1.640	0.879	0.883
H1	1125	0.316	0.263	1.805	0.885	0.888
H2	1125	0.342	0.281	1.790	0.884	0.887
H5	1124	0.401	0.344	1.775	0.883	0.886
C2	1123	0.522	0.465	1.725	0.880	0.883
C3	1123	0.340	0.288	1.800	0.884	0.887
C4	1123	0.366	0.320	1.799	0.883	0.886
S1	1117	0.343	0.305	1.814	0.884	0.886
S2	1116	0.329	0.291	1.817	0.884	0.887
Test scale, Cronbach α (95% CI)	—[a]	—	—	1.701	0.884 (0.875-0.894)	0.85 (0.83-0.87)

[a]Not applicable.

Discussion

Principal Findings

The study aimed to adapt and validate the STAM questionnaire for assessing mHealth technology acceptance among pre-older and older populations regarding the use of health support. The results confirmed the scale's factor structure, supported an 8-factor model with 22 items, and showed good discriminant validity in predicting mHealth intention. The optimal version was a 22-item Thai mHealth STAM using the scoring cutoff (≥152). Subgroup analysis indicated no significant difference in discriminant validity between pre-older and older adults. The scale demonstrated strong internal consistency and stability, with reliability confirmed by Cronbach α and McDonald ω coefficients. This adapted 22-item version is more relevant for assessing mHealth intention among older adults and is suitable for public surveys and routine practice, which take less than 15 minutes to complete.

Our findings are consistent with the previous study conducted by the owner of the original STAM [53], which was subsequently developed into a brief form to save administration time and reduce the burden on respondents. The 14-item brief version of the STAM questionnaire consisted of a 4-factor structure: attitudinal beliefs, control beliefs, gerontechnology anxiety, and health. These findings are consistent with ours, reflecting the original STAM model constructs and the age-related health characteristics of older adults. We observed a decrease in discriminant validity within the pre-older adult group, indicating a need for additional factors to explain their behavioral intentions. For example, older adults with different genders, education levels, income, marital status, and ethnicity may have different intentions and purposes to use mHealth for their health [54].

Strengths and Limitations

On the strength side, this is the first Thai version of the STAM questionnaire suitable for evaluating technology acceptance in Thai older adults. The 22-item Thai STAM version demonstrates structural balance, reliability, and validity in assessing technology acceptance among older individuals. The evaluation process is time-efficient. In addition, this tool can be used with both pre-older adults and older adults to prepare them for engaging with technology in their future lives. Furthermore, the

XSL•FO
RenderX

evaluation takes into account the influence of Thai cultural norms on the adoption and acceptance of mHealth.

However, there are some limitations to consider. Although the psychometric properties of the 22-item Thai mHealth STAM are satisfied through transcultural adaptation in terms of validity and reliability in both the pre-older and older populations, this scale can be applied for use in a broad. However, our study participants may not be representative of the overall Thai pre-older and older populations, as almost all of the participants lived in the northern part of Thailand, particularly in Chiang Mai province. In order to address this concern, future studies, including those based on different regions of Thailand and other specific populations (eg, teenagers, vulnerable groups, minorities, and specific groups of patients) that could potentially derive advantages from mHealth usage, are recommended to expand the generalizability and usability of this scale. Finally, the 22-item Thai mHealth STAM was evaluated based on the board's definition of mHealth. It is possible that the proposed questionnaire may not be compatible with all of the existing mHealth technologies due to the diverse range of mHealth technologies in health care. The patient's choice may vary depending on several factors, such as health care providers, types of services, or the specific application. Hence, we suggest using this questionnaire to assess their acceptance and intention to use it in conjunction with the designated mHealth technology.

Practical Implications of the 22-Item Thai mHealth STAM

The 22-item Thai mHealth STAM offers a practical assessment of patients' acceptability—a crucial factor often overlooked, as evidenced by a recent systematic review of technology acceptability in health care, which revealed that only 10% (142/1219) of the reviewed studies examined patient acceptance [55]. This publicly available questionnaire has the potential to support health care professionals, policy makers, and developers in making informed decisions [56,57], particularly regarding the adoption and acceptance of mHealth within Thai cultural norms. This questionnaire can be incorporated into the research and development (R&D) processes of mHealth and used as a questionnaire to define the target population based on levels of acceptability, as well as ascertain the factors that encourage or hinder the adoption of their mHealth technologies [58,59]. This information is important for informing stakeholders and developers in advance of the mHealth R&D and implementation stages, which necessitate user data for resource allocation and planning in consideration of user requirements and experiences [60,61].

Conclusion

The increasing number of older people, along with their growing adoption of technology, indicates that mHealth technologies might offer a new approach to enhancing the health of older adults with lower health care expenses. Although there are many advantages to using mHealth apps, it is important to consider their acceptance and intention to use them for health-related objectives. We proposed the 22-item Thai mHealth STAM as the questionnaire to evaluate the levels of acceptability and intention to use mHealth in the Thai community of pre-older and older adults. The 22-item Thai mHealth STAM has demonstrated satisfactory psychometric properties in terms of validity and reliability. As a result, it is now feasible to use this questionnaire in a public survey to support stakeholders in making informed decisions. Nevertheless, to improve generalizability and long-term use, further study is needed to investigate the various demographic groups with the specific mHealth interventions.

Acknowledgments

This work was supported by the Faculty of Medicine, Chiang Mai University (grant no. 124-2565), and Chiang Mai University. We would like to express our gratitude to Ms Asapaporn Nimmannit from the Research Institute for Health Sciences at Chiang Mai University for her support in translating the questionnaire. We used QuillBot [62] to review grammar and edit language. We critically reviewed the feedback generated by QuillBot and, based on this, revised the writing using our own words and expressions.

Data Availability

The data sets generated and analyzed during this study are not publicly available due to the institutional policies for protecting participant confidentiality but are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables.
[DOCX File , 72 KB - aging_v8i1e60156_app1.docx]

References

1. Lee C. Technology and aging: the jigsaw puzzle of design, development and distribution. Nat Aging 2022;2(12):1077-1079. [doi: 10.1038/s43587-022-00325-6] [Medline: 37118534]

XSL•FO
RenderX

2. Wang Q, Liu J, Zhou L, Tian J, Chen X, Zhang W, et al. Usability evaluation of mHealth apps for elderly individuals: a scoping review. BMC Med Inform Decis Mak 2022;22(1):317 [FREE Full text] [doi: 10.1186/s12911-022-02064-5] [Medline: 36461017]
3. Liaw ST, Ansari S, Jonnagaddala J, Narasimhan P, Ashraf MM, Harris-Roxas B, et al. Use of mHealth for promoting healthy ageing and supporting delivery of age-friendly care services: a systematic review. Int J Integr Care 2019;19(4):147. [doi: 10.5334/ijic.s3147]
4. Changizi M, Kaveh MH. Effectiveness of the mHealth technology in improvement of healthy behaviors in an elderly population-a systematic review. Mhealth 2017;3:51 [FREE Full text] [doi: 10.21037/mhealth.2017.08.06] [Medline: 29430455]
5. Lee C, Coughlin JF. PERSPECTIVE: older adults' adoption of technology: an integrated approach to identifying determinants and barriers. J of Product Innov Manag 2014;32(5):747-759 [FREE Full text] [doi: 10.1111/jpim.12176]
6. Barnard Y, Bradley MD, Hodgson F, Lloyd AD. Learning to use new technologies by older adults: Perceived difficulties, experimentation behaviour and usability. Computers in Human Behavior 2013;29(4):1715-1724 [FREE Full text]
7. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017;17(1):88 [FREE Full text] [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
8. Miyamoto SW, Henderson S, Young HM, Pande A, Han JJ. Tracking health data Is not enough: a qualitative exploration of the role of healthcare partnerships and mHealth technology to promote physical activity and to sustain behavior change. JMIR Mhealth Uhealth 2016;4(1):e5 [FREE Full text] [doi: 10.2196/mhealth.4814] [Medline: 26792225]
9. Bokolo AJ. Application of telemedicine and eHealth technology for clinical services in response to COVID-19 pandemic. Health Technol (Berl) 2021;11(2):359-366 [FREE Full text] [doi: 10.1007/s12553-020-00516-4] [Medline: 33469474]
10. Stover A, Hayes A, Barrera L, Golbeck E, Malloy C. A comparison of use of the mobile electronic health record by medical providers based on clinical setting. J Emerg Invest 2023. [doi: 10.59720/22-196]
11. Garavand A, Samadbeik M, Nadri H, Rahimi B, Asadi H. Effective factors in adoption of mobile health applications between medical sciences students using the UTAUT model. Methods Inf Med 2019;58(4-05):131-139. [doi: 10.1055/s-0040-1701607] [Medline: 32170717]
12. Koivumäki T, Pekkarinen S, Lappi M, Väisänen J, Juntunen J, Pikkarainen M. Consumer adoption of future myData-based preventive eHealth services: an acceptance model and survey study. J Med Internet Res 2017;19(12):e429 [FREE Full text] [doi: 10.2196/jmir.7821] [Medline: 29273574]
13. Agarwal S, Perry HB, Long LA, Labrique AB. Evidence on feasibility and effective use of mHealth strategies by frontline health workers in developing countries: systematic review. Trop Med Int Health 2015;20(8):1003-1014 [FREE Full text] [doi: 10.1111/tmi.12525] [Medline: 25881735]
14. Kim S, Lee KH, Hwang H, Yoo SY. Analysis of the factors influencing healthcare professionals' adoption of mobile electronic medical record (EMR) using the unified theory of acceptance and use of technology (UTAUT) in a tertiary hospital. BMC Med Inform Decis Mak 2016;16(1):12 [FREE Full text] [doi: 10.1186/s12911-016-0249-8] [Medline: 26831123]
15. Muro-Culebras A, Escriche-Escuder A, Martin-Martin J, Roldán-Jiménez C, De-Torres I, Ruiz-Muñoz M, et al. Tools for evaluating the content, efficacy, and usability of mobile health apps according to the consensus-based standards for the selection of health measurement instruments: systematic review. JMIR Mhealth Uhealth 2021;9(12):e15433 [FREE Full text] [doi: 10.2196/15433] [Medline: 34855618]
16. Schnall R, Cho H, Liu J. Health information technology usability evaluation scale (Health-ITUES) for usability assessment of mobile health technology: validation study. JMIR Mhealth Uhealth 2018;6(1):e4 [FREE Full text] [doi: 10.2196/mhealth.8851] [Medline: 29305343]
17. Hajesmaeel-Gohari S, Khordastan F, Fatehi F, Samzadeh H, Bahaadinbeigy K. The most used questionnaires for evaluating satisfaction, usability, acceptance, and quality outcomes of mobile health. BMC Med Inform Decis Mak 2022;22(1):22 [FREE Full text] [doi: 10.1186/s12911-022-01764-2] [Medline: 35081953]
18. Patel B, Thind A. Usability of mobile health apps for postoperative care: systematic review. JMIR Perioper Med 2020;3(2):e19099 [FREE Full text] [doi: 10.2196/19099] [Medline: 33393925]
19. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth app usability questionnaire (MAUQ): development and validation study. JMIR Mhealth Uhealth 2019;7(4):e11500 [FREE Full text] [doi: 10.2196/11500] [Medline: 30973342]
20. Mustafa N, Safii NS, Jaffar A, Sani NS, Mohamad MI, Abd Rahman AHA, et al. Malay version of the mHealth app usability questionnaire (M-MAUQ): translation, adaptation, and validation study. JMIR Mhealth Uhealth 2021;9(2):e24457 [FREE Full text] [doi: 10.2196/24457] [Medline: 33538704]
21. Schomakers EM, Lidynia C, Vervier LS, Valdez AC, Ziefle M. Applying an extended UTAUT2 model to explain user acceptance of lifestyle and therapy mobile health apps: survey study. JMIR Mhealth Uhealth 2022;10(1):e27095. [doi: 10.2196/27095] [Medline: 35040801]
22. Schretzlmaier P, Hecker A, Ammenwerth E. Extension of the unified theory of acceptance and use of technology 2 model for predicting mHealth acceptance using diabetes as an example: a cross-sectional validation study. BMJ Health Care Inform 2022;29(1):e100640. [doi: 10.1136/bmjhci-2022-100640] [Medline: 36379608]

XSL•FO
RenderX

23. Noblin A, Shettian M, Cortelyou-Ward K, Dugre JS. Exploring physical therapists' perceptions of mobile application usage utilizing the FITT framework. Informatics for Health and Social Care 2016;42(2):180-193. [doi: 10.1080/17538157.2016.1178118]
24. Chen K, Chan AHS. Gerontechnology acceptance by elderly Hong Kong Chinese: a senior technology acceptance model (STAM). Ergonomics 2014;57(5):635-652. [doi: 10.1080/00140139.2014.895855] [Medline: 24655221]
25. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989;13(3):319 [FREE Full text] [doi: 10.2307/249008]
26. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Quarterly 2003;27(3):425. [doi: 10.2307/30036540]
27. Shin HR, Um SR, Yoon HJ, Choi EY, Shin WC, Lee HY, et al. Comprehensive senior technology acceptance model of daily living assistive technology for older adults with frailty: cross-sectional study. J Med Internet Res 2023;25:e41935 [FREE Full text] [doi: 10.2196/41935] [Medline: 37036760]
28. Jantavongso S. Toward global digital literate citizens: a case of 's aging generation. E J Info Sys Dev Countries 2021;88(2):e12207 [FREE Full text] [doi: 10.1002/isd2.12207]
29. Gagnier JJ, Lai J, Mokkink LB, Terwee CB. COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. Qual Life Res 2021;30(8):2197-2218. [doi: 10.1007/s11136-021-02822-4] [Medline: 33818733]
30. The ITC guidelines for translating and adapting tests (Second edition). International Test Commission. 2017. URL: https://www.intestcom.org/ [accessed 2025-12-10]
31. Ahmad NA, Mat Ludin AF, Shahar S, Mohd Noah SA, Mohd Tohit N. Willingness, perceived barriers and motivators in adopting mobile applications for health-related interventions among older adults: a scoping review. BMJ Open 2022;12(3):e054561 [FREE Full text] [doi: 10.1136/bmjopen-2021-054561] [Medline: 35264349]
32. Kim KH. The relation among fit indexes, power, and sample size in structural equation modeling. Structural Equation Modeling: A Multidisciplinary Journal 2005;12(3):368-390. [doi: 10.1207/s15328007sem1203_2]
33. Bonett DG. Sample size requirements for testing and estimating coefficient alpha. J Educat Behav Stat 2002;27(4):335-340. [doi: 10.3102/10769986027004335]
34. Arifin WN. Sample size calculator (web). 2017-2024 [cited 2022]. 2017. URL: http://wnarifin.github.io [accessed 2025-02-25]
35. Watkins MW. Exploratory factor analysis: a guide to best practice. J Black Psychol 2018;44(3):219-246. [doi: 10.1177/0095798418771807]
36. Kaiser HF. An index of factorial simplicity. Psychometrika 2025;39(1):31-36. [doi: 10.1007/BF02291575]
37. Bartlett MS. A note on the multiplying factors for various χ approximations. J Royal Stat Soc: Series B (Methodological) 2018;16(2):296-298. [doi: 10.1111/j.2517-6161.1954.tb00174.x]
38. Norris M, Lecavalier L. Evaluating the use of exploratory factor analysis in developmental disability psychological research. J Autism Dev Disord 2010;40(1):8-20. [doi: 10.1007/s10803-009-0816-2] [Medline: 19609833]
39. Roberson RB, Elliott TR, Chang JE, Hill JN. Exploratory factor analysis in rehabilitation psychology: a content analysis. Rehabil Psychol 2014;59(4):429-438. [doi: 10.1037/a0037899] [Medline: 25221958]
40. Dinno A. Implementing horn's parallel analysis for principal component analysis and factor analysis. The Stata Journal: Promoting communications on statistics and Stata 2009;9(2):291-298. [doi: 10.1177/1536867x0900900207]
41. Schweizer K, DiStefano C. editors. Principles and Methods of Test Construction: Standards and Recent Advances. Boston, MA: Hogrefe; 2016.
42. Fabrigar LR, Wegener DT. Exploratory Factor Analysis. Oxford, United Kingdom: Oxford University Press; 2011.
43. Bentler PM. Comparative fit indexes in structural models. Psychol Bull 1990;107(2):238-246 [FREE Full text] [doi: 10.1037/0033-2909.107.2.238] [Medline: 2320703]
44. Yuan KH, Chan WS, Marcoulides GA, Bentler PM. Assessing structural equation models by equivalence testing with adjusted fit indexes. Structural Equation Modeling: A Multidisciplinary Journal 2015;23(3):319-330. [doi: 10.1080/10705511.2015.1065414]
45. Hooper D, Coughlan J, Mullen M. Structural equation modeling: guidelines for determining model fit. EJBRM 2007:6 [FREE Full text]
46. Hardouin JB, Bonnaud-Antignac A, Sébille V. Nonparametric item response theory using stata. The Stata Journal: Promoting communications on statistics and Stata 2011;11(1):30-51. [doi: 10.1177/1536867x1101100102]
47. Streiner DL, Norman GR, Cairney J. Health Measurement Scales: A practical guide to their development and use. Oxford, United Kingdom: Oxford University Press; 2014.
48. Flora DB, Labrish C, Chalmers RP. Old and new ideas for data screening and assumption testing for exploratory and confirmatory factor analysis. Front Psychol 2012;3:55 [FREE Full text] [doi: 10.3389/fpsyg.2012.00055] [Medline: 22403561]
49. Price LR. Psychometric methods: Theory into practice. New York, NY: Guilford Press; 2017.
50. Gaskin CJ, Happell B. On exploratory factor analysis: a review of recent evidence, an assessment of current practice, and recommendations for future use. Int J Nurs Stud 2014;51(3):511-521. [doi: 10.1016/j.ijnurstu.2013.10.005] [Medline: 24183474]

XSL•FO
RenderX

51. Meehl PE. Why summaries of research on psychological theories are often uninterpretable. PR 1990;66(1):195. [doi: 10.2466/pr0.66.1.195-244]
52. Liu Y, Lu X, Zhao G, Li C, Shi J. Adoption of mobile health services using the unified theory of acceptance and use of technology model: self-efficacy and privacy concerns. Front Psychol 2022;13:944976 [FREE Full text] [doi: 10.3389/fpsyg.2022.944976] [Medline: 36033004]
53. Chen K, Lou VWQ. Measuring senior technology acceptance: development of a brief, 14-Item scale. Innov Aging 2020;4(3):igaa016 [FREE Full text] [doi: 10.1093/geroni/igaa016] [Medline: 32617418]
54. AshaRani PV, Lai D, Koh J, Subramaniam M. Purpose in life in older adults: a systematic review on conceptualization, measures, and determinants. Int J Environ Res Public Health 2022;19(10):5860 [FREE Full text] [doi: 10.3390/ijerph19105860] [Medline: 35627396]
55. AlQudah AA, Al-Emran M, Shaalan K. Technology acceptance in healthcare: a systematic review. Applied Sciences 2021;11(22):10537. [doi: 10.3390/app112210537]
56. Alsyouf A, Lutfi A, Alsubahi N, Alhazmi FN, Al-Mugheed K, Anshasi RJ, et al. The use of a technology acceptance model (TAM) to predict patients' usage of a personal health record system: the role of security, privacy, and usability. Int J Environ Res Public Health 2023;20(2):1347 [FREE Full text] [doi: 10.3390/ijerph20021347] [Medline: 36674105]
57. Hsiao CH, Tang KY. Examining a model of mobile healthcare technology acceptance by the elderly in taiwan. Journal of Global Information Technology Management 2015;18(4):292-311. [doi: 10.1080/1097198x.2015.1108099]
58. Nadal C, Sas C, Doherty G. Technology acceptance in mobile health: scoping review of definitions, models, and measurement. J Med Internet Res 2020;22(7):e17256 [FREE Full text] [doi: 10.2196/17256] [Medline: 32628122]
59. Yau Y, Hsiao CH. The technology acceptance model and older adults' exercise intentions-a systematic literature review. Geriatrics (Basel) 2022;7(6):124 [FREE Full text] [doi: 10.3390/geriatrics7060124] [Medline: 36412613]
60. Noorbergen TJ, Adam MTP, Teubner T, Collins CE. Using co-design in mobile health system development: a qualitative study with experts in co-design and mobile health system development. JMIR Mhealth Uhealth 2021;9(11):e27896. [doi: 10.2196/27896] [Medline: 34757323]
61. Taherdoost H. Importance of technology acceptance assessment for successful implementation and development of new technologies. GJES 2019;1(3). [doi: 10.33552/gjes.2019.01.000511]
62. QuillBot. URL: https://quillbot.com/ [accessed 2025-03-03]

Abbreviations

aOR: adjusted odds ratio
AUROC: area under the receiver operating characteristic
CFA: confirmatory factor analysis
CFI: comparative fit index
COSMIN: Consensus-Based Standards for the Selection of Health Status Measurement Instruments
EFA: exploratory factor analysis
FC: facilitating condition
FITT: Fit between Individuals, Tasks, and Technology
Health-ITUES: Health Information Technology Usability Evaluation Scale
IRT: item response theory
ITC: International Test Commission
KMO: Kaiser-Meyer-Olkin
MAUQ: Mobile Health App Usability Questionnaire
mHealth: mobile health
PCA: principal component analysis
R&D: research and development
REDCap: Research Electronic Data Capture
RMSEA: root mean square error of approximation
SRMR: standardized root mean square residual
STAM: senior technology acceptance model
SUS: System Usability Scale
TAM: technology acceptance model
TLI: Tucker-Lewis Index
UTAUT: unified theory of acceptance and use of technology

XSL•FO
RenderX

Edited by Y Jiang; submitted 03.05.24; peer-reviewed by S Hejazi, M Lee; comments to author 02.12.24; revised version received 17.12.24; accepted 19.02.25; published 11.03.25.

Please cite as:
Buawangpong N, Siviroj P, Pinyopornpanish K, Sirikul W
Transcultural Adaptation, Validation, Psychometric Analysis, and Interpretation of the 22-Item Thai Senior Technology Acceptance Model for Mobile Health Apps: Cross-Sectional Study
JMIR Aging 2025;8:e60156
URL: https://aging.jmir.org/2025/1/e60156
doi:10.2196/60156
PMID:

©Nida Buawangpong, Penprapa Siviroj, Kanokporn Pinyopornpanish, Wachiranun Sirikul. Originally published in JMIR Aging (https://aging.jmir.org), 11.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Developing a Dyadic Immersive Virtual Environment Technology Intervention for Persons Living With Dementia and Their Caregivers: Multiphasic User-Centered Design Study

Elizabeth A Rochon[1], BA; Ayush Thacker[2], BS; Mirelle Phillips[3], BA; Christine Ritchie[2,4*], MSPH, MD; Ana-Maria Vranceanu[1,4*], PhD; Evan Plys[1,4], PhD

[1]Center for Health Outcomes and Interdisciplinary Research, Department of Psychiatry, Massachusetts General Hospital, Boston, MA, United States

[2]Mongan Institute Center for Aging and Serious Illness and the Division of Palliative Care and Geriatric Medicine, Department of Medicine, Massachusetts General Hospital, Boston, MA, United States

[3]Studio Elsewhere, Brooklyn, NY, United States

[4]Harvard Medical School, Boston, MA, United States

[*]these authors contributed equally

Corresponding Author:
Evan Plys, PhD
Center for Health Outcomes and Interdisciplinary Research
Department of Psychiatry
Massachusetts General Hospital
1 Bowdoin Square
Boston, MA, 02114
United States
Phone: 1 617 724 0200
Email: eplys@mgh.harvard.edu

Abstract

Background: Persons living with dementia and their caregivers experience frequent emotional health challenges. Across the illness spectrum, engaging in shared pleasant activities is an important feature of well-being for persons living with dementia–caregiver dyads. Under the umbrella of virtual reality, immersive virtual environment technology (IVET) offers artificial sensory experiences and shows promise in this population. IVET development benefits from a user-centered design approach, and as an emerging field, preliminary testing of safety, usability, and engagement for person living with dementia–caregiver dyads is required.

Objective: We aimed to develop a preliminary IVET intervention for psychosocial health among person living with dementia–caregiver dyads. In doing so, we highlight design considerations and user preferences to ensure the safety and usability of technology-based interventions in the context of dementia.

Methods: We engaged 10 clinicians, 8 caregivers, and 3 persons living with dementia in 5 rounds of focus groups to evaluate the safety and usability of preliminary intervention features. Following prototype development, we engaged caregivers and persons living with dementia (n=9 dyads) in beta testing workshops to observe real-time user interaction with the intervention and guide refinements. Rapid data analysis was used to extract themes relevant to intervention development.

Results: The following themes emerged from focus groups to inform prototype development: (1) designing flexibly to allow users to tailor the intervention experience to their own environmental context and circumstance, (2) designing with the dyad's clinical and relational needs in mind, and (3) accounting for illness and aging-related challenges in design. The following themes emerged from workshops to inform prototype refinements: (1) increasing user support through more feedback and (2) increasing variety of visual and auditory feedback.

Conclusions: Using user feedback throughout the development process, we developed a prototype of an IVET intervention, Toolkit for Experiential Well-Being in Dementia (the Isle of TEND), tailored to the needs of persons living with dementia and their caregivers. Our prototype uses specific design features to promote safety, usability, and engagement in the context of dementia. Future feasibility testing of the intervention is warranted.

International Registered Report Identifier (IRRID): RR2-10.2196/52799

XSL•FO
RenderX

(JMIR Aging 2025;8:e66212) doi:10.2196/66212

KEYWORDS

dementia; caregivers; dyads; technology; design; immersive

Introduction

Background

Throughout the course of dementia, emotional health challenges (eg, depression, negative affect, and worry) are common among both persons living with a dementia diagnosis and their informal caregivers (eg, family members and friends) [1,2]. Persons living with dementia and their caregivers present with different emotional needs throughout the course of dementia [3,4]; however, maintaining relationship quality within the context of changing interpersonal dynamics remains an important component of emotional health for both members of the person living with dementia–caregiver dyad [5].

Dyadic research is gaining popularity within the fields of emotional health and dementia care [6,7]. Dyadic research highlights the impact of interpersonal relationships and the transmission of emotional states between members of the person living with dementia–caregiver dyad, which influence the emotional health of each member [8]. For example, perceived relationship strain by the persons living with dementia is associated with increased emotional distress in their caregivers and vice versa [9,10]. Alternatively, positive affective states can be shared among dyad members (eg, laughter), increasing emotional health for both dyad members [11]. Indeed, relationships and the emotional health needs of persons living with dementia and caregivers often change throughout the course of dementia, especially in advanced stages, as cognitive and communication limitations are more common [4,12]. However, despite changes, the need for positive experiences shared between persons living with dementia and caregivers remains common and relates to fewer behavioral and psychiatric symptoms among persons living with dementia and less stress among caregivers [13]. Thus, shared pleasant activities are an important component of emotional health among person living with dementia–caregiver dyads that can be targeted at any stage of illness.

As dementia advances and cognitive impairments become more pronounced, identifying shared activities that can promote emotional health for both persons living with dementia and caregivers may become more challenging [14-16]. Indeed, many dyadic interventions in the context of dementia are concerned with systematic approaches to identifying and engaging in shared activities [17]. For example, the tailored activity program intervention helps dyads identify and engage in personalized shared activities to reduce agitation among persons living with dementia and improve well-being among caregivers [18,19]. However, many of these shared activity interventions rely heavily on caregivers to plan and initiate activities, which may reduce shared positive affective states due to the caregiver-care recipient role, rather than equitable activity engagement [11]. For example, caregivers may find themselves in the role of an “instructor” or “supervisor,” rather than fully engaging in and receiving the emotional benefits of an activity. Therefore, designing interventions that allow both persons living with dementia and caregivers to fully engage in shared activities may create a positive emotional experience for each dyad member, contributing to their emotional health. Technology may aid this endeavor by creating interactive experiences that minimize the need for caregiver initiation (ie, using a technology platform that engages and guides both members of the dyad through an activity, rather than relying on the caregiver to lead the activity). Immersive virtual environment technology (IVET) is one such option for engaging person living with dementia–caregiver dyads in an intervention together.

IVET exists under the umbrella of virtual and augmented reality [20] and is defined as an artificial sensory experience that users actively engage with as if it were a real experience [21]. IVET promotes active engagement between users and computer-simulated environments through multisensory user-driven experiences [13,21,22]. These methods align well with existing approaches to promoting positive affect among persons living with dementia (ie, multisensory stimulation and personalized activities) [23]. IVETs often follow the 3 I’s design framework [24], which suggests that user engagement is achieved through (1) immersion (ie, creating a feeling of presence within the virtual environment), (2) interaction (ie, creating a sense of agency and control over the virtual environment), and (3) imagination (ie, the potential to create new experiences within the virtual environment). In nondementia populations, IVET interventions have demonstrated positive effects on emotional health outcomes, such as positive affect and emotional health (eg, depression and stress) [25]. Importantly, IVET interventions vary in terms of the level of immersion, categorized as low (ie, screen-based intervention), medium (ie, cave automatic virtual environments and surrounding projection screens), or high (ie, head-mounted displays for isolation from physical reality) [25]. In nondementia populations, medium immersion produces the greatest results on emotional health, with little difference observed between low and high immersion [25]. However, given the potential safety risks, such as cybersickness, falls, and agitation associated with higher levels of immersion in the context of dementia [21,26], IVET with low immersion may be suitable for this population. Indeed, most IVET interventions tested within the context of dementia use lower levels of immersion [27].

IVET interventions allow for multiple users to engage jointly (eg, persons living with dementia and their caregivers participate in shared activities) [21]. However, existing interventions have been tested primarily among persons living with dementia and caregivers individually [27-30]. For persons living with dementia, IVET interventions have demonstrated feasibility and usability, primarily targeting cognitive rehabilitation, exercise, or spatial navigation training [27]. For caregivers, IVET interventions have demonstrated feasibility and usability, often focused on caregiver education and skills [31,32]. Fewer IVET interventions have been developed for person living with

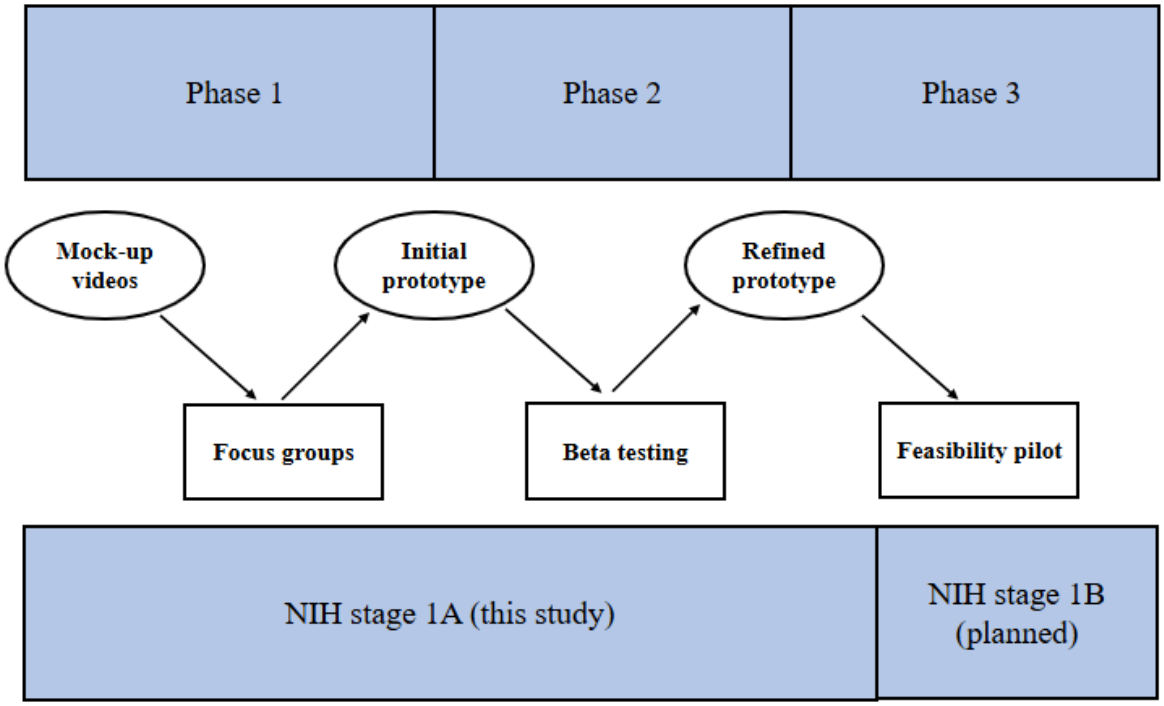
XSL•FO
RenderX

dementia–caregiver dyads to participate in jointly (ie, dyadically), with a focus on shared positive experiences for emotional health. SENSE-GARDEN is an example of a dyadic IVET intervention, which creates an immersive environment based on preferred autobiographical experiences, such as music, images, videos, and personal media (eg, family photos), for persons living with dementia who have moderate to severe cognitive limitations, allowing them to navigate along with their caregivers. The intervention is designed to support reminiscence between the persons living with dementia and caregivers in the present moment. Notably, caregivers have an active role in personalizing the information in SENSE-GARDEN but receive a lower "active dose" of the intervention (ie, fewer caregiver-focused experiences). Rather, caregiver benefit is conceptualized as arising from positive interactions with persons living with dementia. A clinical trial is currently underway for this intervention [33].

A major challenge to designing dyadic IVET interventions in the context of dementia is addressing cognitively mismatched dyads, which is more common as dementia advances. Promoting shared engagement among person living with dementia–caregiver dyads across the illness spectrum must embrace the 3 I's of design (ie, immersion, interaction, and imagination) for dyads with potentially different cognitive ability levels. While challenging from a design perspective, if achieved, this type of intervention would be an invaluable opportunity for dyad members to truly engage in shared activities together (ie, not initiated by the caregiver), helping them to briefly abandon the caregiver-care recipient dynamic and, in doing so, derive relational and emotional benefits. These types of more equitable interactions may promote positive affect and interpersonal connection among both persons living with dementia and caregivers, dignity for the persons living with dementia, and positive aspects of caregiving for the caregivers [34]. However, currently, guidelines for designing dyadic IVET interventions to promote shared engagement among person living with dementia–caregiver dyads are lacking, especially for dyads with mismatched cognitive abilities [6,35].

Additional work is needed to guide the development of dyadic IVET interventions throughout the course of dementia, especially in more advanced stages. For this endeavor, user-centered design approaches may be particularly important to ensure that interventions have the potential to be acceptable and feasible for both dyad members at various stages of illness. User-centered design is an iterative approach that centers the user experience early and often throughout the intervention development process [36]. The aim of user-centered design is to maximize the feasibility and acceptability of an intervention [37]. This study followed a multiphasic conceptual model of user-centered design rooted in human factors research, as outlined by Witteman et al [38]; refer to Figure 1. In phase 1, design teams engaged users and other knowledge holders with the goal of understanding user needs, goals, and strengths. In phase 2, design teams iteratively developed and refined an initial intervention prototype with user feedback. In phase 3, design teams observed users' naturalistic interactions with the initial prototype to inform additional refinements [38]. On the basis of user-centered design principles in the context of dementia [39], we prioritized the following factors across each design phase: understandability (ie, the user understands the intervention functions), learnability (ie, the user is able to easily learn the intervention), usability (ie, the user is able to participate in a task with minimal assistance), and safety (ie, tasks do not inadvertently lead to negative affect or other adverse outcomes).

Figure 1. User-centered iterative design to develop immersive virtual environment technology. Circles denote activities initiated by the design team, and rectangles denote activities designed to elicit user feedback. NIH: National Institutes of Health.



XSL•FO
RenderX

Purpose

This study aimed to develop a dyadic IVET intervention to support the emotional health of person living with dementia–caregiver dyads. Our design team represents an academic-industry partnership. In this study, we highlight our multiphasic user-centered design process [38], focused on the early stages of intervention development of a dyadic IVET intervention (phases 1 and 2 of user-centered design). First, we present data from focus groups with clinicians, caregivers, and person living with dementia–caregiver dyads to understand user needs, goals, strengths, and usability and safety concerns (phase 1). Then, we discuss the process of iteratively developing and refining an initial intervention prototype based on user feedback from multiple user workshops with person living with dementia–caregiver dyads (phase 2). The resulting dyadic IVET intervention prototype will be ready for pilot testing in a pilot feasibility trial (phase 3). In this study, by describing phases 1 and 2 of user-centered design, we highlight early design considerations for future teams focused on developing IVET or other novel technologies with person living with dementia–caregiver dyads across stages of dementia.

Methods

Study Design

This study followed a multiphasic user-centered design framework for the early stages of intervention development, as outlined by Witteman et al [38] and with the following dementia-specific considerations incorporated [39]: (1) understanding user needs, goals, strengths, limitations, and safety concerns; (2) iteratively developing and refining an initial prototype with user feedback; and (3) observing users' naturalistic interactions with the initial prototype. In phase 1, we conducted 5 focus groups with clinicians, caregivers, and person living with dementia–caregiver dyads to identify preferred activity features as well as design considerations for usability (ie, preferred intervention activities and accommodations) and safety in the context of dementia. The study team determined readiness to move to phase 2 after we believed that we understood major safety risks and could generate enough content for mock-ups of an intervention prototype. In phase 2, we conducted 8 beta testing workshops with person living with dementia–caregiver dyads to inform the development and refinement of an initial prototype. We determined readiness to move to phase 3 when usability concerns were sufficiently addressed, as evidenced by minimal usability issues reported by workshop participants. At the conclusion of phase 2, we had a finalized intervention prototype ready for feasibility testing in phase 3, which is not reported in this study. Phases 1 and 2 map onto stage 1A of the National Institutes of Health (NIH) Stage Model of Behavioral Intervention Development (ie, intervention generation, refinement, modification, and adaptation), and phase 3 maps onto stage 1B (ie, pilot testing) [40,41].

Phases 1 and 2 of intervention development were informed by our intervention conceptual model, as suggested by Rochon et al [42]. Specifically, we aimed to develop an IVET intervention that could promote sustained attention, positive emotions, and active engagement (ie, mechanistic targets) among both persons living with dementia and caregivers, leading to increased relationship satisfaction, decreased distress, and reduced agitation among persons living with dementia, as well as decreased burden among caregivers (ie, outcomes).

Ethical Considerations

Procedures were approved by the Massachusetts General Hospital Institutional Review Board (2022P001401). The information reported in this manuscript follows the Guidance for the Reporting of Intervention Development standards [43]. Potential clinician participants provided verbal consent before participating in focus group interviews. Caregivers provided verbal consent or e-consent during screening phone calls. A research assistant administered the University of California, San Diego Brief Assessment of Capacity to Consent Questionnaire (UBACC [44]) to persons living with dementia to determine their ability to provide consent. For those unable to provide consent, the persons living with dementia provided assent, and their caregivers provided surrogate consent to participate. All data were deidentified to protect anonymity. In phase 1, clinicians were compensated US $30, and persons living with dementia and caregivers were compensated US $50 each. In phase 2, persons living with dementia and caregivers were compensated US $50 each for their time.

Participant Recruitment

During phases 1 and 2, this study enrolled three categories of participants: (1) clinicians involved in dementia care, (2) caregivers of persons living with dementia, and (3) person living with dementia–caregiver dyads.

Clinicians were recruited via emails to dementia care clinics within the affiliated health care system. Clinicians were eligible if they (1) self-identified as dementia care providers and (2) were able to participate via Zoom (Zoom Communications). Of the 11 clinicians screened for focus groups, 1 (9%) did not participate due to a scheduling conflict.

Caregivers and dyads were both recruited via research registries (eg, study fliers). Recruitment materials were designed to target caregivers. Interested caregivers responded to the research team to obtain additional information about the study and review eligibility. Caregivers were eligible if they (1) aged >18 years, (2) self-identified as informal caregivers for a person living with dementia (eg, a family member or friend providing support), (3) could participate in English, and (4) had access to a video-enabled device for Zoom participation. Of the 31 caregivers screened for focus groups, 5 (16%) were ineligible due to not identifying as a caregiver of a person living with dementia, 13 (42%) had scheduling conflicts, and 5 (16%) could not be contacted.

For person living with dementia–caregiver dyads, caregivers were targeted using similar recruitment strategies discussed above. Interested caregivers identified and contacted their care recipient with dementia to participate in the study. Then, caregivers and persons living with dementia participated in a screening call together. Using the same eligibility criteria stated in the Participant Recruitment section, caregivers provided verbal consent or e-consent during screening calls. Eligibility

XSL•FO
RenderX

criteria for persons living with dementia included (1) having an informal caregiver participating in the study, (2) a self- or caregiver-reported dementia diagnosis, and (3) the ability to participate in an interview conducted in English. The UBACC uses 2 open-ended questions to assess the understanding of informed consent information (eg, "Can you briefly tell me what the purpose of the study is based on what I just read?"); scores range from 0 to 2, with scores <2 indicating lack of capacity to provide consent. The UBACC scores show a strong correlation with global cognition [45], suggesting that lower scores (ie, the inability to provide consent) are indicative of more advanced stages of dementia. Across stages, of the 12 persons living with dementia screened for this study, 8 (67%) scored a 0, a total of 2 (17%) scored 1, and 2 (17%) scored 2 (ie, able to consent independently). Of the 26 caregivers screened for workshops, 1 (4%) was ineligible due to not identifying as a caregiver, 4 (15%) had scheduling conflicts, 3 (12%) declined due to the care partner's worsening medical status, and 9 (35%) could not be contacted.

Phase 1: Focus Groups

Participants

There were 21 participants across 5 focus groups. A total of 10 clinicians participated across 2 clinician-only focus groups, representing the following professions: geriatric medicine (n=2, 20%), neurology (n=3, 30%), social work (n=2, 20%), occupational therapy (n=1, 10%), psychology (n=1, 10%), and psychiatry (n=1, 10%). Clinician degrees included Occupational Therapist (1/10, 10%), Licensed Psychologist (1/10, 10%), Master of Social Work (1/10, 10%), Licensed Clinical Social Worker (1/10, 10%), Licensed Independent Clinical Social Worker (1/10, 10%), Doctor of Medicine (6/10, 60%), Doctor of Science (1/10, 10%), and Doctor of Philosophy (PhD; 3/10, 30%). Participants reported an average of 10.95 (SD 10.20; range 2-20) years of experience in dementia care. The majority were women (6/10, 60%) and White (9/10, 90%) individuals.

A total of 4 caregivers participated in a caregiver-only group. Participants were primarily women (4/4, 100%), and half identified as White (2/4, 50%) individuals.

A total of 3 person living with dementia–caregiver dyads (n=6) and 1 solo caregiver (ie, the corresponding person living with dementia was unable to attend) participated across 2 dyad focus groups. Persons living with dementia were primarily men (2/3, 67%) and White (2/3, 67%) individuals. Caregivers were primarily women (4/4, 100%), and half identified as White (2/2, 50%) individuals.

Procedure

The sequence of focus groups was clinicians (n=2), person living with dementia–caregiver dyads (n=2), and caregiver-only (n=1) group. A PhD-level clinical psychologist with experience in dementia care and qualitative research led 4 focus groups with clinicians and person living with dementia–caregiver dyads. During dyadic focus groups, the facilitator made every effort to ensure active participation from persons living with dementia, including seeking collateral information from dyad members (eg, confirming caregiver statements with persons living with dementia), asking binary questions (eg, yes or no), rephrasing questions when needed, and making behavioral observations with verbal confirmations (eg, "I noticed you perked up when I mentioned music, do you enjoy listening to music?"). Two trained bachelor-level research assistants coled the caregiver-only group. None of the interviewers had previous relationships with the participants. Group sessions lasted an average of 54 (range 51–57) minutes. All focus groups were completed virtually over secure Zoom, audio recorded, and transcribed.

During each focus group, participants responded to semistructured interview questions targeting activity preferences, technology preferences, usability considerations, and safety concerns; refer to Multimedia Appendix 1 for example focus group questions. In addition, participants provided feedback on a series of mock-up videos (ie, disjointed visuals designed to elicit conceptual feedback and ideas for intervention modules), including passive nature views (ie, waves crashing and Northern lights), guided relaxation with nature background (ie, night sky with voice-over), and abstract color flow (ie, colorful viscous swirl). Mock-up videos were previously used by the design team to inform the development of other IVET platforms targeting nondementia populations. Videos were chosen to provide insight into design features that aligned with our proposed mechanistic targets of sustained attention, positive emotions, and active engagement [42]. Participant feedback was designed to elicit preferences, usability considerations, and safety concerns (eg, "What problems do you see with this experience?" and "Do you foresee any concerns for safety or otherwise?"). Consistent with an iterative process of design, videos and interview questions were updated between each round of focus groups, such that videos or topics that elicited concerns for usability (eg, querying about the utility of sensory kits of household items, such as spices, to accompany videos, which elicited concern for feasibility) or safety (eg, abstract color flows, which elicited concern about increasing agitation or confusion) were omitted, and others were added (eg, new nature scenes with multiple points of view).

During each focus group, 1 or 2 bachelor-level research assistants observed the interviews and completed live field notes using a standardized rapid data analysis (RDA) template created for use in this study; refer to Multimedia Appendix 1 for the RDA template. At the conclusion of each focus group, all facilitators and observers completed reflective field notes using a standardized form with categories similar to those in the RDA template.

Phase 2: Beta Testing Workshops

Participants

A total of 9 person living with dementia–caregiver dyad (n=18) participants were recruited across 8 beta testing workshops. Dyads were recruited separately for phases 1 and 2; none of the dyads who participated in phase 1 participated in phase 2. However, dyads enrolled in phase 2 were allowed to participate in multiple workshops if desired. Of the 9 dyads, most (6, 67%) participated in 1 workshop and 3 (33%) participated in 2 workshops; however, none participated in >2 workshops. Participants ranged from 1 to 4 dyads per workshop, with the majority consisting of 1 dyad (6/8, 75% workshops).

XSL•FO
RenderX

Across workshops, 100% (18/18) of participants identified as White individuals. Most persons living with dementia (7/9, 78%) identified as male, and most caregivers (7/9, 78%) identified as female. Caregivers provided proxy reports of dementia diagnoses, per their understanding of the medical records of the persons living with dementia, which included Alzheimer disease (2/9, 22%), unspecified dementia (4/9, 44%), Lewy body dementia (1/9, 11%), primary progressive aphasia (1/9, 11%), and frontotemporal dementia (1/9, 11%). Given that this study did not access medical records, caregiver reports were not confirmed, and the stage of illness was not reported. Multiple dyad relationship types were noted, including spousal (4/9, 44%), sibling (1/9, 11%), and adult child (1/9, 11%) caregivers.

Procedure

Beta testing workshops were led by members of the industry partner design team, with 1 to 2 trained research assistants in attendance to assist dyads in navigating the platform and to complete live field notes for RDA. All workshops were completed virtually over secure Zoom and recorded for the research team to review after session completion (eg, detailed feedback and behavioral observations). If multiple dyads engaged in 1 session, breakout rooms were used to enable private viewing experiences. Beta testing workshops lasted an average of 63 (range 34-103) minutes.

Workshops were completed via Zoom. At the start of each workshop, the design team shared a website link to the initial IVET intervention prototype with participants. Dyad participants were instructed to share their screen after opening the website. A workshop facilitator then prompted dyads to navigate the intervention both freely (eg, “As you explore the area, please share your thoughts out loud”) and with specific instructions (eg, “Can you figure out how to navigate back to the home screen?”). Participants started by freely navigating the prototype, providing live feedback on usability and acceptability (eg, confusing instructions and enjoyable features). Then, participants were presented with 1 to 3 modules to navigate, guided by facilitator prompts that primarily focused on usability and acceptability (eg, “Are there specific features on the island that drew your interest?” “What are your initial thoughts on the visuals of the activity?” and “Was navigating around the island intuitive or challenging?”). Dyads could choose the module they were most interested in navigating. In some situations, the second or third module was chosen for the dyad to ensure that feedback was gathered for each module. At the conclusion of each workshop, caregivers and persons living with dementia independently responded to quantitative measures assessing the perceived feasibility and usability of the platform. As workshops progressed, feedback became more targeted (eg, preference for specific images) than in earlier workshops (eg, preferred level of feedback from the platform), moving toward an initial prototype.

During workshops, research assistants completed observational field notes using a standardized template created for use in this study; refer to Multimedia Appendix 1 for the template. Field notes were used as part of RDA to inform intervention refinements between each workshop.

Measures

Usability was assessed using the System Usability Scale (SUS) [46]. The SUS consists of 10 items evaluating the perceived usability of technological platforms, including websites. The SUS is the most widely used scale to assess usability for technology-based interventions in the context of dementia, but self-report challenges may exist with more advanced cognitive limitations [47]. However, the SUS is primarily applied to persons living with dementia with mild cognitive impairments and lacks psychometric data across the full spectrum of dementia (ie, limited data available in advanced stages of dementia) [47]. Furthermore, the SUS uses 5-point Likert scale for response options [48,49]. Therefore, following best practices to reduce cognitive load and ensure that persons living with dementia could meaningfully participate, we adapted this scale to reflect yes or no response categories [47,50]. This decision increased the likelihood that persons living with dementia in this study could provide feedback independently.

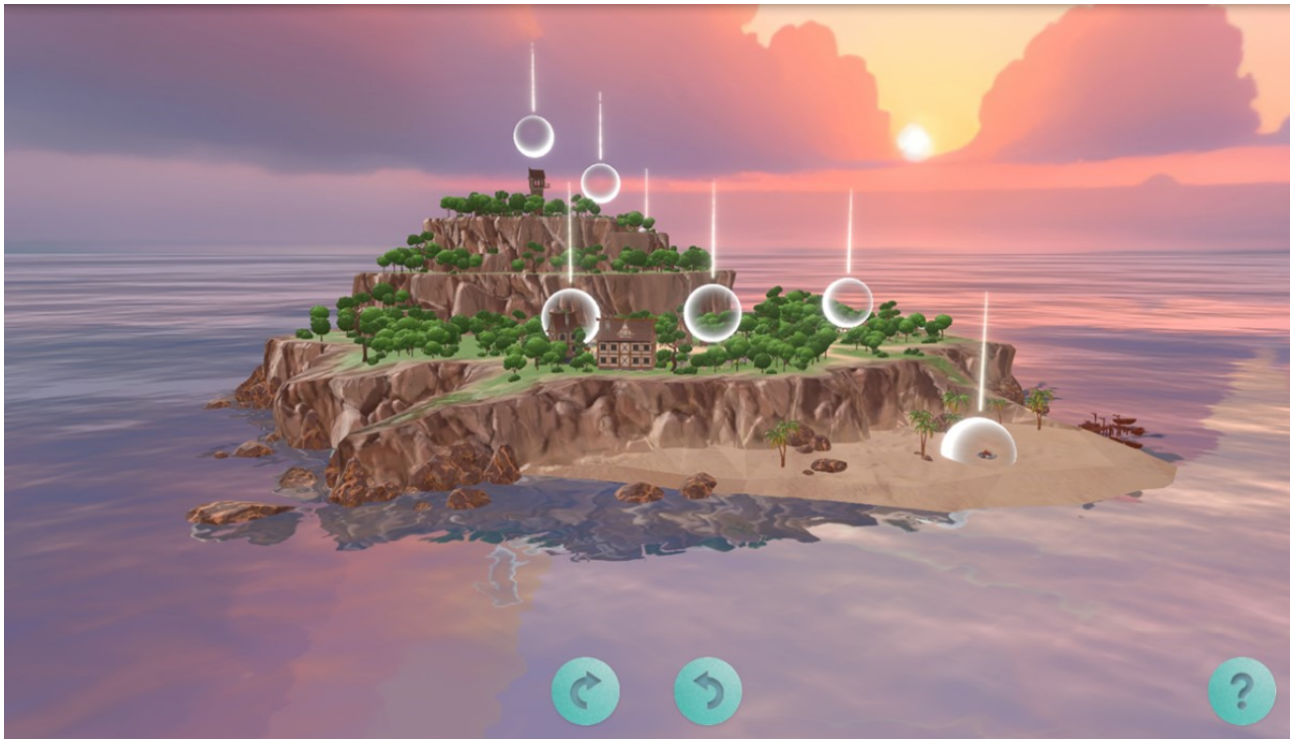
A total of 3 additional study-specific feasibility questions targeted entertainment, safety, and control. These items were adapted from previous technology-based intervention studies with participants with mild cognitive impairment and their caregivers [48] and provided feedback salient to our conceptual model (eg, engagement and positive affect) and design aims (eg, safety). These items were also asked in a yes or no format to increase inclusivity for persons living with dementia.

Intervention Prototype

Activities outlined in Phase 1: Focus Groups and Phase 2: Beta Testing Workshops sections culminated in the development of an initial prototype of a web-based dyadic IVET intervention, referred to as the Toolkit for Experiential Well-Being in Dementia (Isle of TEND). The Isle of TEND website contained a home page (Figure 2) with the ability to navigate to 8 modules, which contained different shared activities; refer to Multimedia Appendix 2 for an overview of modules. Consistent with our conceptual model [42], each module in the Isle of TEND featured a shared activity intended to promote sustained attention, positive emotions, and active engagement with the intervention. The intervention was designed to be used by both persons living with dementia and caregivers together in the home.

XSL•FO
RenderX

Figure 2. The Toolkit for Experiential Well-Being in Dementia (Isle of TEND) home screen.



Data Analysis

This study used RDA as an analytic framework to identify themes from focus groups and workshops. RDA is used to quickly identify themes in qualitative data [51]. Rather than written transcriptions, this technique relies on live coding directly from interviews, such as taking observational notes during interviews or while listening to an audio recording [52-54]. For this study, RDA was chosen to identify specific feedback from participants related to preferences, usability, and safety that could guide prototype development and refinement. By quickly capturing this information, our research team was able to communicate findings back to our design team to iteratively respond to feedback and update the intervention prototype to gather more feedback from users.

Consistent with previous studies using RDA [55], we compiled RDA template notes from 2 to 3 team members (ie, 1 PhD-level clinical psychologist and 2 bachelor-level research assistants) into a single matrix for each focus group. Data from this matrix were used to guide discussions with the design team and informed design decisions toward initial prototype development.

At the conclusion of the study, matrices were combined and used to facilitate thematic analysis of focus groups and workshops separately. Focus group data were analyzed as a single dataset (ie, clinician, caregiver, and dyad focus groups analyzed together), such that we identified themes that emerged across groups. We did not compare themes across populations, as this was not consistent with the aims of this study. Two research assistants (ER and AT) trained in qualitative RDA methods independently reviewed matrices to identify codes. Coders met on multiple occasions during the coding process to resolve discrepancies. Codes were grouped by content to form themes. The thematic structure was iteratively developed and confirmed with a meeting between coders and a clinical psychologist on the study team with experience in dementia care (EP). After a thematic structure was agreed upon, the research assistants (ER and AT) reviewed transcripts to identify salient quotes for each thematic category. Both focus groups and workshops followed the same analytic steps. However, transcripts were only reviewed for focus groups because the data collected during workshops produced fewer exemplar quotes (eg, workshops contained more “thinking aloud” methods or behavioral observation, making data less relevant for exemplar quotations).

Results

Phase 1: Focus Groups

Three themes emerged from focus groups with clinicians, caregivers, and dyads: (1) designing flexibly to allow users to tailor the intervention experience to their own environmental context and circumstance, (2) designing with the dyad’s clinical and relational needs in mind, and (3) accounting for illness and aging-related challenges in the design.

Theme 1: Designing Flexibly to Allow Users to Tailor the Intervention Experience to Their Own Environmental Context and Circumstance

The ability for users to individualize their experience emerged as an important design feature to increase acceptability and usability as well as to empower choices among users. Participants noted that the option to individualize or tailor intervention content could help accommodate a larger range of users based on their access to and comfort navigating technology. In addition, participants recommended that the IVET intervention be accessible via various platforms. For example, a clinician noted the following:

XSL•FO
RenderX

> *[Consider] an app that is free, that can download the materials into [users] cell phone or their device so it doesn't depend on the Internet connection...and maybe having the option of audio only or audio plus visual.*

Participants also noted the importance of developing stimuli that were familiar to and pleasant for participants of various backgrounds, cohorts, and identities to promote individualization. Furthermore, participants noted that familiar stimuli may especially increase comfort, enjoyment, engagement, and safety for persons living with dementia. A caregiver highlighted the potential benefits of familiar stimuli:

> *Having an inspirational waterfall isn't necessarily going to put me at ease, but maybe hearing my family's familiar voice, or the clattering of my own kitchen, or having something that individualizes this to me would.*

Similarly, a clinician highlighted the need for inclusivity as follows:

> *I just wonder if the salience of some of the images might be different depending on the cultural groups. The aurora borealis is beautiful, but for maybe people who [are] not familiar with that part of the of the world...what are the scenes, the sounds that are most salient and comforting to that individual.*

Of note, related to the environmental context of an at-home IVET intervention, caregivers discouraged the use of a proposed sensory kit of household materials, citing feasibility issues with the ability for a range of users to locate these materials in their home (eg, spices).

Theme 2: Designing With the Dyad's Clinical and Relational Needs in Mind

Participants noted that designing an intervention with a specific purpose and orienting users to a rationale for using the intervention may increase usability. Specifically, participants suggested recommendations for appropriate times and intended outcomes of use (eg, using when calm and avoiding when agitated). For example, a clinician noted the potential clinical utility of an IVET intervention:

> *I think that [the calming sounds are] good, and it appeals to an inherent positive emotional experience which could also aid caregivers when they're trying to prep for redirection or transition. Basically, if we're trying to get somebody to do something that they may not necessarily want to do we're far more likely to be able to if a person is in a positive emotional state versus a negative one and so these types of things could aid in that process of prepping a patient for a transition.*

Caregivers believed that they would likely need to be present to initiate the intervention as well as to guide the persons living with dementia through some of the modules. For example, a caregiver noted the following:

> *[Caregivers] would have to be there to sort of guide [Person living with dementia] or kinda help them figure out what [the intervention] is.*

Participants were skeptical that an IVET intervention could be engaging for both dyad members, noting challenges of achieving sustained attention and immersion for both persons living with dementia and caregivers. This is highlighted by a caregiver who stated the following:

> *[Person living with dementia] loses interest pretty quick, and I think sometimes he has trouble following the storyline. So, it would be hard to find something that would be suitable for both of us, cause we're kind of in very different places now. And that's one of the things that's difficult is we don't have as many shared experiences as we used to.*

Participants recommended structured stimuli to quickly capture and sustain the attention of users, which they noted could be challenging in the context of dementia. For example, a caregiver stated the following:

> *Without having a clear beginning or end, [the person living with dementia] might get bored or discouraged, it needs to have more structure.*

Furthermore, participants emphasized that the IVET intervention would need to be designed to sustain attention and immersion without overstimulating or confusing users with cognitive challenges.

Theme 3: Accounting for Illness and Aging-Related Challenges in the Design

Participants noted the importance of ensuring the reduction, rather than an unintentional exacerbation, of behavioral and psychiatric symptoms of dementia, particularly agitation and hallucinations. This concern is highlighted by a clinician, who stated the following:

> *Dementias with any psychiatric component or psychotic features, paranoia, delusional thinking...you always want to be careful about providing stimulation when we don't know where that may lead the patient's train of thought.*

There were specific stimuli that participants recommended avoiding, such as 1 video with a moving (ie, swirling) bright color scheme against a black background (Figure 3), which participants suggested could potentially exacerbate psychiatric symptoms. For example, a clinician stated the following:

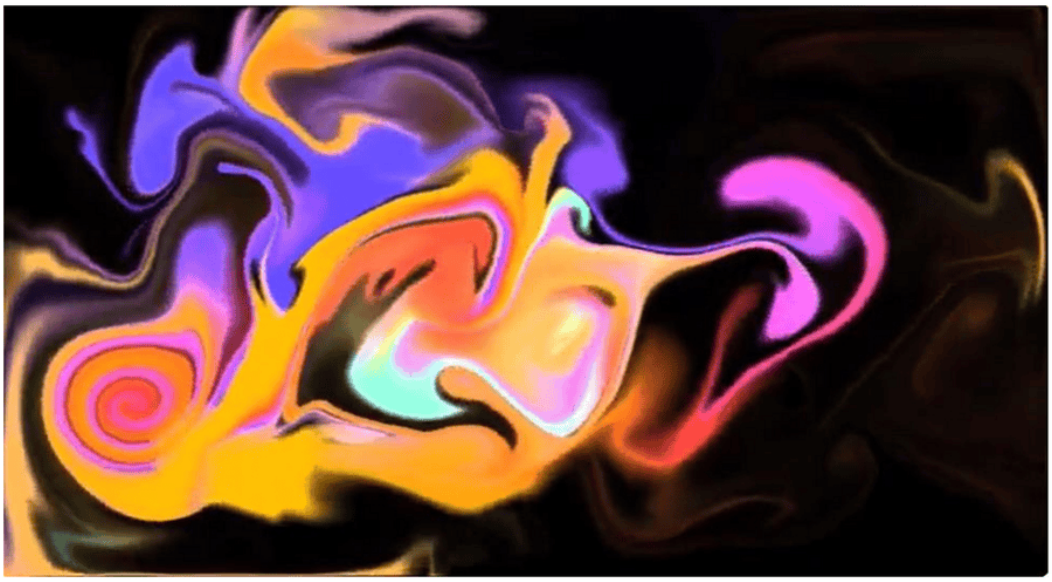
> *I think patients would not like [the swirling color video], they have a lot of visual spatial misperceptions.*

Furthermore, participants cautioned against dark backgrounds, which might increase psychiatric symptoms and confusion or disorientation. For example, a caregiver noted the following:

> *My husband is really paranoid, he thinks people are out to get him. I was thinking of the lighting, if there is darkness that might increase his paranoia.*

XSL•FO
RenderX

Figure 3. Swirling color scheme video.



Another caregiver noted the following:

> *The edges of my video were very dark. So, it gives you almost like that movie theater feel to it, which I worry, if you're immersed in something too much, do you have less grasp on here and now where you are.*

In addition, as noted across multiple themes, participants relayed that certain cognitive symptoms, especially attentional deficits and apathy, could pose a challenge to engagement and usability. One clinician stated the following:

> *I wonder if in some instances this [mock-up videos] may be a little too unstimulating to capture the patient's attention, unless they are really already very calm and wanting to be calm.*

Another clinician stated the following:

> *for images...it might be really, really simple ones, the less complicated you make it, the better received because [person living with dementia] cognitive abilities and processual abilities are impaired.*

Participants also noted that visual clarity was an important design feature and recommended that the IVET intervention accommodate for age-related changes and pathologies more common with older age. For example, 1 caregiver stated the following:

> *There's something about lighting that I would pay attention to for elders...I have another elder in my life who has glaucoma and she can't see things in depth of perception and other things.*

Phase 2: Beta Testing Workshops

Qualitative Results

Two main themes emerged from beta testing workshops with dyads that informed intervention refinement: (1) more feedback to guide using the intervention and (2) more variety of visual and auditory experiences.

Theme 1: Increasing User Support Through More Feedback

Participants noted that more purposeful or goal-directed actions within intervention modules could reduce confusion during use. When engaging with the intervention, caregivers and persons living with dementia frequently inquired about the end goal of modules, especially when feeling "lost" or "stuck." Relatedly, participants recommended more instructions, which, per observation, appeared to be related to participant concerns about making an error or doing the experience "wrong." Together, an open-ended module without clear directions, which was initially developed to increase exploration and reduce the need for caregiver instructions, actually increased confusion and perceived misuse of the intervention by users.

Furthermore, participants voiced the need for continuous instructions throughout use to understand next steps, which was particularly relevant in modules that included multistep tasks. For example, a module that developed a poem based on typed responses to multiple prompts caused some confusion among users because dyads were unsure of the task they were being prompted to complete (eg, not knowing that prompts would generate a poem) or the end goal of the multistep process (eg, the need to press next and respond to another prompt to develop the full poem). In addition to more prompts, participants suggested having multiple methods of engagement, such as a speech-to-text function, to promote accessibility and usability.

Theme 2: Increasing the Variety of Visual and Auditory Feedback

Participants recommended increasing the presence and variability of responsive sounds, background music, and interactive visuals. Caregivers and persons living with dementia both noted that alternating music or background colors might reduce redundancy after multiple uses and make the intervention more engaging. In addition to increasing engagement, dyads noted that sound effects, especially calming nature sounds, improved their experience by creating a sense of calm while using the intervention.

XSL•FO
RenderX

Quantitative Analysis

In general, users rated the intervention positively (Table 1). On the study-specific feasibility scale, 100% (11/11) of caregivers and 67% (6/9) of persons living with dementia (n=2, 22% were unsure or did not respond) felt that the program was safe to use, 82% (9/11) of caregivers and 100% (9/9) of persons living with dementia felt that the program was entertaining, and 82% (9/11) of caregivers and 89% (8/9) of persons living with dementia felt that the program helped them "feel in charge of their own experience." On the SUS, 82% (9/11) of caregivers and 89% (8/9) of persons living with dementia thought that the program was "easy to use," and 73% (8/11) of caregivers and 78% (7/9) of persons living with dementia "felt very confident using the program" (n=1, 11% were unsure or did not respond).

Table 1. Feasibility and usability results of beta testing workshops.

	Caregivers, (n=11), n (%)			Persons living with dementia (n=9), n (%)		
	Yes	No	Unsure or unable to answer	Yes	No	Unsure or unable to answer
"I think that I would like to use this program."	7 (64)	2 (18)	2 (18)	6 (67)	2 (22)	1 (11)
"I found the program unnecessarily complex."	1 (9)	10 (91)	0 (0)	1 (11)	8 (89)	0 (0)
"I thought the program was easy to use."	9 (82)	2 (18)	0 (0)	8 (89)	1 (11)	0 (0)
"I think that I would need the support of a technical person to be able to use this program."	4 (36)	7 (64)	0 (0)	3 (33)	4 (44)	2 (22)
"I found the various functions in this program were well integrated."	9 (82)	2 (18)	0 (0)	6 (67)	3 (33)	0 (0)
"I thought there was too much inconsistency in this program."	2 (18)	9 (82)	0 (0)	2 (22)	5 (56)	2 (22)
"I would imagine that most people would learn to use this program very quickly."	7 (64)	3 (27)	1 (9)	4 (44)	0 (0)	5 (56)
"I found the program very cumbersome to use."	0 (0)	11 (100)	0 (0)	1 (11)	5 (56)	3 (33)
"I felt very confident using the program."	8 (73)	2 (18)	1 (9)	7 (78)	1 (11)	1 (11)
"I needed to learn a lot of things before I could get going with this program."	1 (9)	10 (91)	0 (0)	2 (22)	4 (44)	3 (33)
"Was this program able to entertain your relative and you?"	9 (82)	2 (18)	0 (0)	9 (100)	0 (0)	0 (0)
"Did you feel that the program was safe to use for your relative and you?"	11 (100)	0 (0)	0 (0)	6 (67)	1 (11)	2 (22)
"The program helped my relative feel in charge of their own experience."	9 (82)	2 (18)	0 (0)	8 (89)	1 (11)	0 (0)

Feedback-Design Integration

Throughout intervention development, our academic-industry team met weekly to discuss results from focus groups and workshops and translate findings into pragmatic design elements. Often, the research team helped provide the context of findings within the illness context of dementia, as well as guide a systematic approach to intervention design that would prepare the team for a future, scientifically rigorous clinical trial. The industry design team often discussed the possibilities and limitations of an IVET intervention delivered with minimal equipment (ie, via a computer or tablet). These perspectives led to productive discussions about incorporating user feedback into intervention design. Together, this iterative process led to the end product of an IVET intervention ready for initial feasibility testing (NIH stage 1B) that was deemed safe and usable by end users (ie, person living with dementia–caregiver dyads). Feedback and design decisions are provided in Table 2.

XSL•FO
RenderX

Table 2. User feedback and design team responses.

User feedback	Design feature
Focus groups: initial prototyping	
Implement choice-based activities when possible.	The team decided on a web-based IVET[a] format so that users could guide themselves through the platform. The intervention was user driven, and users chose from various modules in the intervention, deciding the type and length of activity within each module.
Establish the ability to adjust audio or visuals depending on preferences.	Developers enabled access to the intervention on the user's own tablet or computer so that users could adjust the volume and brightness to their preference.
Include visuals or audio that are appealing and familiar to diverse populations.	The design team developed 8 different modules, including modules set in a forest, an ocean, a beach, and a lighthouse.
Integrate instructions for the intervention.	The design team included text instructions on the home screen to orient users when they logged into the intervention. The instructions reviewed how to navigate through the home screen and the modules.
Intervention can be used among people with and without cognitive impairment.	Modules were "error-free" so that users were unable to "fail" or choose the "wrong" option, halting the experience; regardless of user choice, the experience continued.
Include stimuli that are reactive to touch or hand or arm movement.	Developers created two activities, which accessed the device's camera to track and react to hand movement: (1) Sound Garden, which used hand movement to control the music pitch and volume in a garden scene, and (2) Art Studio, which used hand movement to create brush strokes on the screen to reveal an art piece.
Design modules that are grounded in reality (familiar structure and clear beginning and ending).	The team set the intervention in recognizable locations, such as a waterfall, a lighthouse, and a beach. The game had a clear beginning with text instructions when users logged on.
Avoid abstract images and darkness in visuals.	The team removed abstract "swirling colors" stimuli. The team discarded a video set in a dark forest environment and brightened the colors in the remaining modules.
Beta testing workshop: prototype refinements	
Clarify the end goal or purpose of the intervention. Include more goal-directed aspects to the intervention.	The design team added an introduction slide clarifying that the intervention was designed to create "shared adventures" between the dyad and "engage the senses and spark imagination." The team included a "help" button across all modules to guide users.
Implement different methods of engagement, such as a speech-to-text option instead of typing.	Developers included a speech-to-text function in modules that required users to type. These activities accessed the device microphone to capture spoken responses from the dyad and convert them into text. To increase engagement, developers created custom soundscapes for scenes and changed them based on the visual stimuli presented.
Add more instructions and audio-to-text instructions.	Developers added a module description when users hovered their mouse over each module location. The team added step-by-step tutorials, which isolated different interactive aspects of the experience when users clicked on each module. In addition, the team added voice-over to text as an accessibility option to all introductions and guides.
Increase the visual stimuli throughout use to reduce redundancy.	The team increased the visual options in modules throughout the experience by creating more artwork and audio clips. For example, in the Storytelling module Campfire Cove, the team increased the number of available pictures to choose from and created an accompanying story.
Include visual and audio stimuli, such as animation or music, in modules that lacked sounds or featured static images. Add calming music or nature audio when possible.	The design team added nature-based animation and sound on the home screen, including auditory and visual waves, so that when logging on, users were greeted with these stimuli. In addition, the design team added nature-based sounds to the background of module instructions, such as birds chirping and "calming" music.

[a]IVET: immersive virtual environment technology.

XSL•FO
RenderX

Discussion

Principal Findings

Following a multiphasic user-centered design framework with dementia-specific considerations [38], this study used focus groups and user workshops to develop and refine the initial prototype of a dyadic IVET intervention to support the emotional health of person living with dementia–caregiver dyads (ie, the Isle of TEND). The intervention was rooted in our conceptual model of shared activities that can promote sustained attention, positive emotions, and active engagement among both dyad members simultaneously [42]. The Isle of TEND is now ready for feasibility testing, consistent with NIH stage 1B.

Design Considerations

Overview

Across focus groups and workshops, we identified several key considerations for designing an IVET intervention for person living with dementia–caregiver dyads. Importantly, we did not constrict our intervention to a specific stage of illness. Rather, we engaged users, including those with cognitive limitations as evidenced by UBACC scores, and other knowledge holders who could provide feedback relevant to users with varying degrees of cognitive ability. This approach allowed us to capture multiple voices and perspectives to ensure that we are designing for all persons living with dementia, including those often not considered during the design stage due to significant cognitive impairment.

Personalization and Variety of Experiences

Consistent with previous research on technology-based interventions with persons living with dementia [56,57] and activity engagement in them [58], personalization and variety of experiences emerged as key design features for our intervention. This feedback is not surprising, given the heterogeneity of identities, preferences, and lived experiences of dyads managing dementia. Increasing the number of novel or rotating visuals and sounds in each module allowed for personalization and helped to reduce boredom with repeated use. Our design team was cautious to only make slight variations in activities (eg, sounds or storytelling prompts) to reduce boredom without increasing confusion. Consistently, participants noted that nature-based sounds and animation, such as waves or chirping birds, evoked positive emotions. Therefore, our design team focused on nature scenes for many of our modules. Previous research with persons living with dementia supports that nature sounds introduced through technology, referred to as a soundscape, are perceived positively with minimal safety concerns [59,60].

Optimizing Autonomy and Choice While Sustaining Ease of Use

Feedback related to personalization highlighted an important component of designing in dementia: optimizing autonomy and choice while sustaining ease of use. In addition to increasing acceptability, the ability to tailor interventions to personal preferences may create an opportunity for persons living with dementia to have more control over their user experience. This was important for both members of the dyad to promote a sense of accomplishment, autonomy, and confidence and, ideally, allow the dyad to break the care recipient-caregiver script, enabling both members to feel as though they have ownership of their own experiences with the intervention. Previous technology-based interventions developed for persons living with dementia with mild cognitive impairments also note the importance of prioritizing user control [48]. Furthermore, in the context of dyads, increasing autonomy and choice for persons living with dementia means reducing the need for the caregiver to "deliver" the intervention. While refinements to our prototype are still needed to fully achieve this outcome, it is an important endeavor for dyadic interventionists in the context of dementia, as allowing for a break in the care recipient-caregiver script may produce positive emotional benefits for both dyad members across all stages of illness.

Users consistently provided feedback that some user-driven choices in our platform were too open ended (eg, coming up with a poem without prompts), which could lead to confusion and produce the opposite emotional experience for dyad members than intended (ie, frustration with being "wrong" and the perceived need for caregiver-directed behaviors). To address this balance, our design team attempted to provide ample instructions and feedback for use while not imposing "rules" (ie, "error-free" design) and promoting exploration. For example, several of our modules initially present users with directions for navigating the activity and then provide participants with open-ended (ie, pictures of different scenes) and forced-choice (ie, multiple choices for storytelling) decisions to guide their experience based on their needs and preferences. In addition to providing directions throughout the module, as recommended by users, we developed responsive features to reduce cognitive load and avoid frustration. For example, following effective communication strategies in the context of dementia [61], if a dyad did not respond to a multiple-choice prompt within a period of time, a forced-choice option appeared (ie, this or that). Most users with dementia rated that they felt in control of their own experience and confident in using the Isle of TEND, which may be a product of these design features.

Maintaining Engagement Among Cognitively Mismatched Users

We set out to design an IVET intervention that can maintain sustained attention and engagement among cognitively mismatched dyad users while at the same time not overwhelming the persons living with dementia or underwhelming the caregivers. Participants recommended that to achieve this goal, we needed to have a clear structure, continuous feedback, and novel stimuli, which are consistent with theory (eg, flow) [62]. In response to these recommendations, our design team developed a home screen with an introduction to the IVET intervention, tutorials and instructions for each module, and a navigation guide that appears constantly. Furthermore, our design team developed several modules that are reactive to touch or movement. For example, 1 module, the Art Garden, accesses the device camera and uses hand movement as "brushstrokes" to paint a picture. Despite these features designed to promote sustained attention and engagement, user feedback in the workshops was still mixed regarding the ability to capture and

XSL•FO
RenderX

sustain engagement across cognitively mismatched dyads and reduce caregiver-directed behaviors (eg, caregiver reiterating directions). However, it is important to note that 82% (9/11) of caregivers and 100% (9/9) of persons living with dementia still found the intervention to be entertaining. The field still needs to address the important topic of designing for mismatched cognitive abilities, as there are potential clinical benefits associated with shared engagement for person living with dementia–caregiver dyads managing dementia throughout the course of illness (eg, emotional closeness and relaxation). It is possible that design teams need to attend to more nuanced interactions between users and platforms to evaluate common elements of sustained attention and engagement across users with varying cognitive abilities (eg, movement).

Not Exacerbating Behavioral and Psychiatric Symptoms of Dementia

Regarding safety concerns, participants flagged specific stimuli that could potentially lead to distress or exacerbate behavioral or psychiatric symptoms that accompany certain types of dementia. In particular, participants recommended avoiding abstract stimuli, bright colors, or moving objects against a dark background, as these could potentially increase adverse events (eg, agitation) in more advanced stages of dementia. These images may relate to overstimulation or confusion, which increases behavioral and psychiatric disturbances in the context of advanced dementia when persons living with dementia have increased vulnerability to environmental stressors [63]. Users suggested that, in general, simpler and easily recognizable images may be better received by both persons living with dementia and caregivers. Participants also identified dark or shaded visuals as potentially problematic (eg, blurriness and glare) for users with age-related or other visual impairments. In response to the feedback, the design team removed visuals and modules that relied on limited lighting and highly abstract images. Of note, there were no observed exacerbations of dementia-related behavioral or psychiatric symptoms during workshops.

Limitations

This initial intervention development study has several limitations to consider. First, our convenience sample of caregivers, dyads, and clinicians consisted of a majority of White participants. In future pilot testing, we will ensure that participants are more representative of the US population. Second, we did not administer cognitive assessments to describe the level of cognitive impairment of our sample. Therefore, we cannot empirically comment on the cognitive profile of participants with objective measurement, including the severity of cognitive impairment. This study used self-reports from persons living with dementia and caregivers to capture the presence of dementia, and the UBACC provides additional insight into the presence of cognitive limitations inhibiting the ability to provide autonomous consent. Indeed, dementia research must balance participant burden, the reality of difficulties obtaining a dementia diagnosis, and study goals. In line with our broader goal of offering TEND as a free intervention for individuals in community settings, we included participants with varying levels of cognitive ability, as reflected by the fact that only 17% (2/12) of our sample was able to provide independent consent. To circumvent challenges with reporting on quantitative surveys for those with more advanced dementia, our team took multiple steps to support persons living with dementia in reporting their own outcomes, including using binary scales and reading questions aloud. However, these methods may have reduced the validity of our measures, and we still cannot be sure that all items were fully comprehended. Finally, we chose RDA over other forms of qualitative analysis. For this study, prioritizing a quick summary of information and relaying this to our design team consistent with RDA was necessary, but this analytic approach may also be more susceptible to researcher bias.

Conclusions

The result of this multiphasic user-centered design study with an academic-industry design team was the development of an initial dyadic IVET intervention for persons living with dementia and their caregivers to use together at home. On the basis of our conceptual model, we aimed to promote sustained attention, positive emotions, and active engagement in both dyad members to increase relationship satisfaction and reduce psychosocial distress. In this study, we highlight early intervention development with a focus on designing for usability and safety. While some of our design considerations may be helpful for future investigators and design teams focusing on dyads managing dementia, we also recognize that there are gaps that still need to be filled, such as promoting mutual engagement among cognitively mismatched dyads.

Conflicts of Interest

MP is the founder and chief executive officer of Studio Elsewhere and receives financial compensation for her role. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
Example questions from phases 1 and 2 and rapid data analysis template.
[DOCX File , 27 KB - aging_v8i1e66212_app1.docx]

Multimedia Appendix 2
Toolkit for Experiential Well-Being in Dementia (Isle of TEND) modules.
[DOCX File , 12522 KB - aging_v8i1e66212_app2.docx]

XSL•FO
RenderX

References

1. Sallim AB, Sayampanathan AA, Cuttilan A, Ho R. Prevalence of mental health disorders among caregivers of patients with Alzheimer disease. J Am Med Dir Assoc 2015 Dec;16(12):1034-1041. [doi: 10.1016/j.jamda.2015.09.007] [Medline: 26593303]
2. Leung DK, Chan WC, Spector A, Wong GH. Prevalence of depression, anxiety, and apathy symptoms across dementia stages: a systematic review and meta-analysis. Int J Geriatr Psychiatry 2021 Sep 04;36(9):1330-1344. [doi: 10.1002/gps.5556] [Medline: 33905138]
3. Blandin K, Pepin R. Dementia grief: a theoretical model of a unique grief experience. Dementia (London) 2017 Jan;16(1):67-78 [FREE Full text] [doi: 10.1177/1471301215581081] [Medline: 25883036]
4. Gitlin LN, Jutkowitz E, Gaugler JE. Caregiver intervention research: now and into the future. College of Nursing and Health Professions. 2019. URL: https://sites.nationalacademies.org/cs/groups/dbassesite/documents/webpage/dbasse_196527.pdf [accessed 2025-05-13]
5. Conway ER, Watson B, Tatangelo G, McCabe M. Is it all bleak? A systematic review of factors contributing to relationship change in dementia. Int Psychogeriatr 2018 Nov;30(11):1619-1637 [FREE Full text] [doi: 10.1017/S1041610218000303] [Medline: 29667571]
6. Plys E, Szapary CL, Vranceanu AM. Flexibly incorporating the dyad into psychosocial intervention: considerations for neuropalliative care. Semin Neurol 2024 Oct;44(5):559-566. [doi: 10.1055/s-0044-1791568] [Medline: 39362315]
7. Wilson SJ, Novak JR, Yorgason JB, Martire LM, Lyons KS. New opportunities for advancing dyadic health science in gerontology. Gerontologist 2024 Jan 01;64(1):gnac187 [FREE Full text] [doi: 10.1093/geront/gnac187] [Medline: 36534908]
8. Guterman EL, Allen IE, Josephson SA, Merrilees JJ, Dulaney S, Chiong W, et al. Association between caregiver depression and emergency department use among patients with dementia. JAMA Neurol 2019 Oct 01;76(10):1166-1173 [FREE Full text] [doi: 10.1001/jamaneurol.2019.1820] [Medline: 31282955]
9. Campbell JL, Rowe MA, Marsiske M. Behavioral symptoms of dementia: a dyadic effect of caregivers' stress process? Res Gerontol Nurs 2011 Jul;4(3):168-184 [FREE Full text] [doi: 10.3928/19404921-20100901-02] [Medline: 20873693]
10. Miller LM, Kaye JA, Lyons KS, Lee CS, Whitlatch CJ, Caserta MS. Well-being in dementia: a cross-sectional dyadic study of the impact of multiple dimensions of strain on persons living with dementia and their family care partners. Int Psychogeriatr 2019 May;31(5):617-626 [FREE Full text] [doi: 10.1017/S104161021800203X] [Medline: 30722800]
11. Chen KH, Merrilees J, Brown CL, Yee C, Sapozhnikova A, Wells JL, et al. Interpersonal linkage in positive and negative emotional behaviors, emotional well-being, and physical functioning in dementia caregivers. Clin Psychol Sci 2024 Jan 02;12(6):1039-1055 [FREE Full text] [doi: 10.1177/21677026231214530]
12. Hoel V, Koh WQ, Sezgin D. Enrichment of dementia caregiving relationships through psychosocial interventions: a scoping review. Front Med (Lausanne) 2022;9:1069846 [FREE Full text] [doi: 10.3389/fmed.2022.1069846] [Medline: 36687423]
13. Petrovsky DV, Yildiz M, Yefimova M, Sefcik JS, Baker ZG, Ma KP, et al. Shared activities as a protective factor against behavioral and psychological symptoms of dementia and caregiver stress. Innov Aging 2024;8(5):igae034 [FREE Full text] [doi: 10.1093/geroni/igae034] [Medline: 38660116]
14. Roland KP, Chappell NL. Meaningful activity for persons with dementia: family caregiver perspectives. Am J Alzheimers Dis Other Demen 2015 Sep;30(6):559-568 [FREE Full text] [doi: 10.1177/1533317515576389] [Medline: 25788432]
15. Van't Leven N, de Lange J, Prick AE, Pot AM. How do activating interventions fit the personal needs, characteristics and preferences of people with dementia living in the community and their informal caregivers? Dementia (London) 2019 Jan;18(1):157-177. [doi: 10.1177/1471301216662378] [Medline: 27509919]
16. Lai FH, Yan EW, Tsui WS, Yu KK. A randomized control trial of activity scheduling for caring for older adults with dementia and its impact on their spouse care-givers. Arch Gerontol Geriatr 2020;90:104167. [doi: 10.1016/j.archger.2020.104167] [Medline: 32663699]
17. Balvert SC, Del Sordo GC, Milders MV. The efficacy of dyadic interventions for community-dwelling people with dementia and their caregivers: a systematic review and meta-analysis. Ageing Res Rev 2024 Apr;96:102258. [doi: 10.1016/j.arr.2024.102258] [Medline: 38479479]
18. Gitlin LN, Marx K, Piersol CV, Hodgson NA, Huang J, Roth DL, et al. Effects of the tailored activity program (TAP) on dementia-related symptoms, health events and caregiver wellbeing: a randomized controlled trial. BMC Geriatr 2021 Oct 20;21(1):581 [FREE Full text] [doi: 10.1186/s12877-021-02511-4] [Medline: 34670502]
19. Gitlin LN, Winter L, Vause Earland T, Adel Herge E, Chernett NL, Piersol CV, et al. The Tailored Activity Program to reduce behavioral symptoms in individuals with dementia: feasibility, acceptability, and replication potential. Gerontologist 2009 Jun;49(3):428-439 [FREE Full text] [doi: 10.1093/geront/gnp087] [Medline: 19420314]
20. Motejlek J, Alpay E. Taxonomy of virtual and augmented reality applications in education. IEEE Trans Learn Technol 2021 Jun 1;14(3):415-429. [doi: 10.1109/tlt.2021.3092964]
21. Smith JW. Immersive virtual environment technology to supplement environmental perception, preference and behavior research: a review with applications. Int J Environ Res Public Health 2015 Sep 11;12(9):11486-11505 [FREE Full text] [doi: 10.3390/ijerph120911486] [Medline: 26378565]

XSL•FO
RenderX

22. Blascovich J, Loomis J, Beall AC, Swinth KR, Hoyt CL, Bailenson JN. Immersive virtual environment technology as a methodological tool for social psychology. Psychol Inquiry 2002 Aug;13(2):103-124 [FREE Full text] [doi: 10.1207/s15327965pli1302_03]
23. Cheng C, Baker G, Dursun S. Use of multisensory stimulation interventions in the treatment of major neurocognitive disorders. Psychiatry Clin Psychopharmacol 2019 Dec 23;29(4):916-921. [doi: 10.1080/24750573.2019.1699738]
24. Burdea GC, Coiffet P. Virtual Reality Technology. Hoboken, NJ: John Wiley & Sons; 2003.
25. Li H, Ding Y, Zhao B, Xu Y, Wei W. Effects of immersion in a simulated natural environment on stress reduction and emotional arousal: a systematic review and meta-analysis. Front Psychol 2022 Jan 9;13:1058177 [FREE Full text] [doi: 10.3389/fpsyg.2022.1058177] [Medline: 36698558]
26. Shu S, Woo BK. Use of technology and social media in dementia care: current and future directions. World J Psychiatry 2021 Apr 19;11(4):109-123 [FREE Full text] [doi: 10.5498/wjp.v11.i4.109] [Medline: 33889536]
27. Strong J. Immersive virtual reality and persons with dementia: a literature review. J Gerontol Soc Work 2020 Apr;63(3):209-226. [doi: 10.1080/01634372.2020.1733726] [Medline: 32091323]
28. Hirt J, Beer T. Use and impact of virtual reality simulation in dementia care education: a scoping review. Nurse Educ Today 2020 Jan;84:104207. [doi: 10.1016/j.nedt.2019.104207] [Medline: 31669968]
29. Hodge J, Balaam M, Hastings S, Morrissey K. Exploring the design of tailored virtual reality experiences for people with dementia. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI '18; April 21-26, 2018; Montreal, QC. [doi: 10.1145/3173574.3174088]
30. Kokorelias KM, Chiu M, Paul S, Zhu L, Choudhury N, Craven CG, et al. Use of virtual reality and augmented reality technologies to support resilience and skill-building in caregivers of persons with dementia: a scoping review. Cureus 2024 Jul;16(7):e64082. [doi: 10.7759/cureus.64082] [Medline: 39114214]
31. Huang Y, Ho KH, Christensen M, Wong DW, Wang S, Su JJ, et al. Virtual reality-based simulation intervention for enhancing the empathy of informal caregivers of people with dementia: a mixed-methods systematic review. Int J Ment Health Nurs 2024 Apr;33(2):241-258. [doi: 10.1111/inm.13240] [Medline: 37817470]
32. Zhi S, Gao S, Sun J, Ma D, Zhao Y, He M, et al. Virtual reality-based dementia educational programmes for formal and informal caregivers of people with dementia: a scoping review. J Clin Nurs 2024 Apr 29. [doi: 10.1111/jocn.17198] [Medline: 38685766]
33. Goodall G, Ciobanu I, Taraldsen K, Sørgaard J, Marin A, Draghici R, et al. The use of virtual and immersive technology in creating personalized multisensory spaces for people living with dementia (SENSE-GARDEN): protocol for a multisite before-after trial. JMIR Res Protoc 2019 Aug 19;8(9):e14096 [FREE Full text] [doi: 10.2196/14096] [Medline: 31538942]
34. Sideman AB, Merrilees J, Dulaney S, Kiekhofer R, Braley T, Lee K, et al. "Out of the clear blue sky she tells me she loves me": connection experiences between caregivers and people with dementia. J Am Geriatr Soc 2023 Jul;71(7):2172-2183 [FREE Full text] [doi: 10.1111/jgs.18297] [Medline: 36917135]
35. Paliokas I, Tzallas A, Katertsidis N, Votis K, Tzovaras D. Gamification in social networking: a platform for people living with dementia and their caregivers. In: Proceedings of the 17th International Conference on Bioinformatics and Bioengineering. 2017 Presented at: BIBE 2017; October 23-25, 2017; Washington, DC. [doi: 10.1109/bibe.2017.00015]
36. Lyon AR, Koerner K. User-centered design for psychosocial intervention development and implementation. Clin Psychol (New York) 2016 Jun;23(2):180-200 [FREE Full text] [doi: 10.1111/cpsp.12154] [Medline: 29456295]
37. Gould JD, Lewis C. Designing for usability: key principles and what designers think. Commun ACM 1985 Mar;28(3):300-311. [doi: 10.1145/3166.3170]
38. Witteman HO, Dansokho SC, Colquhoun H, Coulter A, Dugas M, Fagerlin A, et al. User-centered design and the development of patient decision aids: protocol for a systematic review. Syst Rev 2015 Jan 26;4(1):11 [FREE Full text] [doi: 10.1186/2046-4053-4-11] [Medline: 25623074]
39. Ceccacci S, Generosi A, Giraldi L, Mengoni M. An user-centered approach to design smart systems for people with dementia. In: Proceedings of the 2017 IEEE 7th International Conference on Consumer Electronics - Berlin. 2017 Presented at: ICCE-Berlin 2017; September 3-6, 2017; Berlin, Germany. [doi: 10.1109/icce-berlin.2017.8210650]
40. National Institutes of Health (NIH) homepage. National Institutes of Health (NIH). URL: https://www.nih.gov/ [accessed 2024-08-07]
41. Onken L. Implementation science at the national institute on aging: the principles of it. Public Policy Aging Report 2022;32(1):39-41 [FREE Full text] [doi: 10.1093/ppar/prab034]
42. Rochon EA, Sy M, Phillips M, Anderson E, Plys E, Ritchie C, et al. Bio-experiential technology to support persons with dementia and care partners at home (TEND): protocol for an intervention development study. JMIR Res Protoc 2023 Dec 29;12:e52799 [FREE Full text] [doi: 10.2196/52799] [Medline: 38157239]
43. Duncan E, O'Cathain A, Rousseau N, Croot L, Sworn K, Turner KM, et al. Guidance for reporting intervention development studies in health research (GUIDED): an evidence-based consensus study. BMJ Open 2020 Apr 08;10(4):e033516 [FREE Full text] [doi: 10.1136/bmjopen-2019-033516] [Medline: 32273313]
44. Jeste DV, Palmer BW, Appelbaum PS, Golshan S, Glorioso D, Dunn LB, et al. A new brief instrument for assessing decisional capacity for clinical research. Arch Gen Psychiatry 2007 Aug;64(8):966-974. [doi: 10.1001/archpsyc.64.8.966] [Medline: 17679641]

XSL•FO
RenderX

45. Seaman JB, Terhorst L, Gentry A, Hunsaker A, Parker LS, Lingler JH. Psychometric properties of a decisional capacity screening tool for individuals contemplating participation in Alzheimer's disease research. J Alzheimers Dis 2015;46(1):1-9 [FREE Full text] [doi: 10.3233/JAD-142559] [Medline: 25765917]
46. Brooke J. SUS: a 'quick and dirty' usability scale. In: Usability Evaluation In Industry. Boca Raton, FL: CRC Press; 1996.
47. Miguel Cruz A, Daum C, Comeau A, Salamanca JD, McLennan L, Neubauer N, et al. Acceptance, adoption, and usability of information and communication technologies for people living with dementia and their care partners: a systematic review. Disabil Rehabil Assist Technol 2023 May;18(4):443-457. [doi: 10.1080/17483107.2020.1864671] [Medline: 33378627]
48. Quintana M, Anderberg P, Sanmartin Berglund J, Frögren J, Cano N, Cellek S, et al. Feasibility-usability study of a tablet app adapted specifically for persons with cognitive impairment-SMART4MD (support monitoring and reminder technology for mild dementia). Int J Environ Res Public Health 2020 Sep 18;17(18):6816 [FREE Full text] [doi: 10.3390/ijerph17186816] [Medline: 32961993]
49. Lewis JR. The system usability scale: past, present, and future. Int J Hum Comput Interact 2018 Mar 30;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
50. Kolanowski A, Gilmore-Bykovskyi A, Hill N, Massimo L, Mogle J. Measurement challenges in research with individuals with cognitive impairment. Res Gerontol Nurs 2019 Jan 01;12(1):7-15 [FREE Full text] [doi: 10.3928/19404921-20181212-06] [Medline: 30653646]
51. Schexnayder J, Perry KR, Sheahan K, Majette Elliott N, Subramaniam S, Strawbridge E, et al. Team-based qualitative rapid analysis: approach and considerations for conducting developmental formative evaluation for intervention design. Qual Health Res 2023 Jul 06;33(8-9):778-789. [doi: 10.1177/10497323231167348] [Medline: 37278662]
52. Halcomb EJ, Davidson PM. Is verbatim transcription of interview data always necessary? Appl Nurs Res 2006 Feb;19(1):38-42. [doi: 10.1016/j.apnr.2005.06.001] [Medline: 16455440]
53. Suchman L, Omoluabi E, Kramer J, Vallin J, Sedlander E, Gitome S, et al. Analyzing fast and slow: combining traditional and rapid qualitative analysis to meet multiple objectives of a complex transnational study. Front Sociol 2023;8:961202 [FREE Full text] [doi: 10.3389/fsoc.2023.961202] [Medline: 36818663]
54. Vindrola-Padros C, Johnson GA. Rapid techniques in qualitative research: a critical review of the literature. Qual Health Res 2020 Aug;30(10):1596-1604. [doi: 10.1177/1049732320921835] [Medline: 32667277]
55. Lewinski AA, Crowley MJ, Miller C, Bosworth HB, Jackson GL, Steinhauser K, et al. Applied rapid qualitative analysis to develop a contextually appropriate intervention and increase the likelihood of uptake. Med Care 2021 Jun 01;59(Suppl 3):S242-S251 [FREE Full text] [doi: 10.1097/MLR.0000000000001553] [Medline: 33976073]
56. Goodall G, Taraldsen K, Serrano JA. The use of technology in creating individualized, meaningful activities for people living with dementia: a systematic review. Dementia (London) 2021 May;20(4):1442-1469 [FREE Full text] [doi: 10.1177/1471301220928168] [Medline: 32475166]
57. Eun SJ, Kim EJ, Kim J. Artificial intelligence-based personalized serious game for enhancing the physical and cognitive abilities of the elderly. Future Gener Comput Syst 2023 Apr;141:713-722. [doi: 10.1016/j.future.2022.12.017]
58. Trahan MA, Kuo J, Carlson MC, Gitlin LN. A systematic review of strategies to foster activity engagement in persons with dementia. Health Educ Behav 2014 Oct;41(1 Suppl):70S-83S [FREE Full text] [doi: 10.1177/1090198114531782] [Medline: 25274714]
59. Houben M, Brankaert R, Bakker S, Kenning G, Bongers I, Eggen B. The role of everyday sounds in advanced dementia care. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020 Presented at: CHI '20; April 25-30, 2020; Honolulu, HI. [doi: 10.1145/3313831.3376577]
60. Devos P, Aletta F, Thomas P, Petrovic M, Vander Mynsbrugge T, Van de Velde D, et al. Designing supportive soundscapes for nursing home residents with dementia. Int J Environ Res Public Health 2019 Dec 04;16(24):4904 [FREE Full text] [doi: 10.3390/ijerph16244904] [Medline: 31817300]
61. Small JA, Gutman G, Makela S, Hillhouse B. Effectiveness of communication strategies used by caregivers of persons with Alzheimer's disease during activities of daily living. J Speech Lang Hear Res 2003 Apr;46(2):353-367. [doi: 10.1044/1092-4388(2003/028)] [Medline: 14700377]
62. Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi. Dordrecht, The Netherlands: Springer; 2014.
63. Kales HC, Gitlin LN, Lyketsos CG. Assessment and management of behavioral and psychological symptoms of dementia. BMJ 2015 Mar 02;350:h369 [FREE Full text] [doi: 10.1136/bmj.h369] [Medline: 25731881]

Abbreviations

Isle of TEND: Toolkit for Experiential Well-Being in Dementia
IVET: immersive virtual environment technology
NIH: National Institutes of Health
RDA: rapid data analysis
SUS: System Usability Scale
UBACC: University of California, San Diego Brief Assessment of Capacity to Consent Questionnaire

XSL•FO
RenderX

Edited by M Antoniou; submitted 24.09.24; peer-reviewed by E Tonkin, D Urbanski, R Bratches; comments to author 06.01.25; revised version received 17.03.25; accepted 30.03.25; published 21.05.25.

Please cite as:
Rochon EA, Thacker A, Phillips M, Ritchie C, Vranceanu AM, Plys E
Developing a Dyadic Immersive Virtual Environment Technology Intervention for Persons Living With Dementia and Their Caregivers: Multiphasic User-Centered Design Study
JMIR Aging 2025;8:e66212
URL: https://aging.jmir.org/2025/1/e66212
doi:10.2196/66212
PMID:

©Elizabeth A Rochon, Ayush Thacker, Mirelle Phillips, Christine Ritchie, Ana-Maria Vranceanu, Evan Plys. Originally published in JMIR Aging (https://aging.jmir.org), 21.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Design of a Mobile App and a Clinical Trial Management System for Cognitive Health and Dementia Risk Reduction: User-Centered Design Approach

Hannes Hilberger[1,2], MSc[‡]; Bianca Buchgraber-Schnalzer[1], MSc; Simone Huber[1], MSc; Theresa Weitlaner[1], MSc; Markus Bödenler[1], PhD; Alara Abaci[3], MSc; Jeroen Bruinsma[4], PhD; Ana Diaz[5], PhD; Anna Giulia Guazzarini[6,7], MSc; Jenni Lehtisalo[8], PhD; Seungjune Lee[9,10]; Vasileios Loukas[11,12], MSc; Francesca Mangialasche[3,13], PhD; Patrizia Mecocci[6,7], PhD; Tiia Ngandu[3,8,14], PhD; Anna Rosenberg[3,8], PhD; Elisabeth Stögmann[9,10], MD; Konsta Valkonen[8,15], MSc; Elena Uhlik[9,10]; Helena Untersteiner[9,10], MD; Laura Kneß[1], MSc; Helmut Ahammer[2], PhD; Sten Hanke[1], PhD

[1]Institute of eHealth, University of Applied Sciences - FH JOANNEUM, Graz, Austria

[2]Gottfried Schatz Research Center, Division of Medical Physics and Biophysics, Medical University of Graz, Graz, Austria

[3]Division of Clinical Geriatrics, Center for Alzheimer Research, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden

[4]Department of Health Promotion, Care and Public Health Research Institute, Maastricht University, Maastricht, The Netherlands

[5]Alzheimer Europe, Luxembourg, Luxembourg

[6]Division of Gerontology and Geriatrics, Department of Medicine and Surgery, University of Perugia, Perugia, Italy

[7]Division of Clinical Geriatrics, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden

[8]Lifestyles and Living Environments Unit, Department of Public Health, Finnish Institute for Health and Welfare, Helsinki, Finland

[9]Department of Neurology, Medical University of Vienna, Vienna, Austria

[10]Comprehensive Center for Clinical Neurosciences and Mental Health, Medical University of Vienna, Vienna, Austria

[11]Department of Materials Science and Engineering, Unit of Medical Technology and Intelligent Information Systems, University of Ioannina, Ioannina, Greece

[12]Biomedical Research Institute, Foundation for Research and Technology–Hellas, Ioannina, Greece

[13]Medical Unit Aging, Theme Inflammation and Aging, Karolinska Institutet, Stockholm, Sweden

[14]Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio, Finland

[15]Institute of Clinical Medicine, University of Eastern Finland, Kuopio, Finland

[‡]LETHE Consortium

Corresponding Author:
Hannes Hilberger, MSc
Institute of eHealth
University of Applied Sciences - FH JOANNEUM
Institute of eHealth
Eckertstraße 30i
Graz, 8020
Austria
Phone: 43 316 5453 6529
Email: hannes.hilberger@fh-joanneum.at

Abstract

Background: The rising prevalence of dementia is a major concern, with approximately 45% of cases linked to 14 modifiable risk factors. The European project LETHE aims to develop a personalized digital intervention model to delay or prevent cognitive decline through risk factor management.

Objective: The objective of our study was to design a clinical trial platform for older individuals at risk of cognitive decline, including a mobile app for study participants and a clinical trial management system (CTMS) for health professionals.

Methods: Using a user-centered design approach, workshops and feedback rounds involved potential participants representing the target group and professionals. The LETHE app's usability was assessed among 156 older adults enrolled in a 2-year

XSL•FO
RenderX

multinational randomized controlled trial evaluating the feasibility of a digitally supported lifestyle program for dementia risk reduction. The randomized controlled trial is currently ongoing; the System Usability Scale (SUS) was administered 1 month after baseline to map first user experiences. Feedback on the LETHE CTMS was collected from 21 users.

Results: Of the 78 participants in the trial intervention group, 66 (85%) provided responses for the mobile app, with a median SUS score of 70 (IQR 55-82). Within the control group, 73% (57/78) of responses were received, with a median SUS score of 73 (IQR 63-90). For the CTMS, we received 71% (15/21) of responses, and the feedback was mostly positive. A ranking of the features that could be considered beyond state of the art showed that the integration of personalized activities (mean 2.23, SD 1.17) and real-time appointments (mean 2.46, SD 1.51) were considered the most novel ones.

Conclusions: The LETHE app and CTMS were developed to support a personalized digital intervention method within a study involving 156 participants. Limitations include participants having digital literacy and internet access, potentially impacting the generalizability of the findings. Despite these limitations, positive feedback and high usability scores suggest promising potential for the LETHE app and CTMS in supporting personalized interventions to prevent cognitive decline in older adults.

(JMIR Aging 2025;8:e66660) doi:10.2196/66660

KEYWORDS

clinical trial; medical informatics; health information systems; eHealth; mobile health; mHealth; mobile apps; dementia; cognitive decline; prevention; multidomain interventions; artificial intelligence; AI

Introduction

Background

Overview

The rising prevalence of dementia, driven by sociodemographic changes, constitutes a significant global health challenge. Projections indicate a notable rise from 57 million affected individuals in 2019 to an anticipated 153 million by 2050 [1]. Notably, the Lancet Commission's report underscores that approximately 45% of all cases of dementia are associated with 14 potentially modifiable risk factors at different phases of the life span [2], and diet has also been suggested as an additional factor [3]. Previous multidomain intervention studies have demonstrated their positive impact on cognitive performance while simultaneously targeting various lifestyle domains, including diet, physical activity, cognitive training, management of cardiovascular risk factors, or social interaction [4-6].

Information and communications technology (ICT) solutions such as apps running on phones and tablets, as well as wearable devices that collect user-generated health data (digital biomarkers) in an automated way, could potentially help provide digitally supported interventions for a broader audience and groups at risk of cognitive decline. At the same time, these technologies can collect data to monitor progress and adherence to lifestyle interventions. Other research projects have shown that digital interventions can be effective and feasible for older adults, but this depends on several factors such as digital literacy, usability of the technology, and the design of the study with regard to human support (hybrid intervention) [7-9].

The Horizon 2020 LETHE project [10] has been initiated to provide such a personalized digital intervention model and components for reducing risk of developing cognitive decline by evolving the successful Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER) multidomain intervention model [5]. Leveraging ICT-based methods and guided in-person as well as remote sessions, LETHE aims to prevent cognitive decline in at-risk older individuals by providing personalized feedback and intervention methods.

The LETHE project focuses on the aforementioned modifiable risk factors grouped into the following lifestyle domains: "Physical Activity," "Nutrition," "Cognitive Training," "Management of vascular/metabolic risk factors," "Sleep/Relaxation," and "Social Activity."

The LETHE study encompasses a currently ongoing 2-year randomized controlled multicenter, parallel-group feasibility trial (ClinicalTrials.gov; NCT05565170) involving 156 participants at risk of cognitive decline distributed across clinical centers in Austria (Medical University of Vienna), Finland (Finnish Institute for Health and Welfare), Sweden (Karolinska Institute), and Italy (University of Perugia). Participants were equally randomized to one of the study groups (intervention and control), both supported by digital tools (mobile app and smartwatch). At baseline, participants had a mean age of 68.8 (SD 4.5) years, with an average of 14.9 (SD 3.1) years of education. A total of 64.7% of the participants were women. Notably, most participants (84%) used their smartphones at least 6 times per day. Two-thirds of the participants had previous experience using their smartphones for health tracking, and approximately 48% had used lifestyle-related apps. In addition, 63% of the participants reported using the internet for eHealth-related tasks such as booking physician appointments or checking test results [11].

To provide intervention components and continuously collect data from the trial participants, a LETHE smartphone app was designed and implemented. Clinical data, including blood test results, are entered into the LETHE clinical trial management system (CTMS), which was also developed during the course of the project, by health professionals at each visit. The CTMS was developed to ensure seamless integration with the app, allowing all participant-entered data to be displayed for health professionals to review and respond accordingly. Furthermore, data and services from third parties have been integrated, including cognitive training data and data from Fitbit devices, which track steps, sleep, and physical activity. The feasibility

XSL•FO
RenderX

of other novel technologies (audio glasses and a robot that reflect the content of the smartphone app) is being investigated in a substudy.

This study aimed to describe the design and implementation of the LETHE technical components, which were carefully developed through a joint process involving potential users, clinical experts, and designers and technicians.

Hybrid Intervention Approach

LETHE adopts a hybrid intervention (blended therapies) approach, seamlessly integrating digital intervention sessions with in-person and group sessions led by a trained coach (eg, physiotherapist or nutritionist). This comprehensive strategy, meticulously managed by health professionals, is in harmony with the study's mobile app and CTMS concept.

The literature on hybrid interventions for older adults at risk of cognitive decline, incorporating technology with coaching and peer sessions in comparison to stand-alone technology interventions, is still developing and has so far mainly focused on already symptomatic individuals. Nevertheless, emerging studies indicate that fostering social connection and interaction can be pivotal for success in dementia interventions [12]. Digital health platforms, easily accessible via mobile technology, can be efficient and cost-effective tools for dementia prevention. They not only offer individualized cognitive training but can also provide an engaging user experience [13,14]. The currently ongoing prevention of dementia using mobile phone applications (PRODEMOS) study creates an evidence-based dementia prevention strategy using mobile health (mHealth) accessible to at-risk individuals [15]. In contrast, Maintain Your Brain was a completely online lifestyle intervention targeting modifiable risk factors for dementia that was delivered to the participants via a web interface [16]. Wesselman et al [17] conducted a comprehensive overview of web-based multidomain lifestyle programs, combining it with a meta-analysis to evaluate their effectiveness. Through a systematic literature research, they collected a wide range of web-based programs.

Other studies have provided app-supported self-regulation for older adults based on self-determination theory, with a tablet facilitating tailored exercise programs and playing a key role in action planning and execution of behaviors [18]. The involvement of a personal coach was pivotal, adapting exercises to individual preferences and providing motivating remote monitoring to empower older adults to enhance physical activity levels at home. The experience conveyed in the study by Mehra et al [18] emphasizes the value of the personal coach, particularly during the initial phases of goal setting, behavior execution, and evaluation in self-regulation to achieve specific goals by guiding one's own behavior. Their findings indicate that the availability of a personal coach remains crucial even when technology supports the intervention. Nevertheless, concerns about adherence and safety arose in the absence of instructor guidance [18].

The success of this hybrid intervention approach could be further supported by the seamless interaction of a mobile app and a web-based CTMS for professionals. This integrated system allows professionals to interact with and monitor patients' development, enabling timely interventions when needed.

Digital Intervention Applications

mHealth apps are increasingly used by individuals to engage in health behaviors, aid in the self-management of chronic conditions, or enact preventive measures [19]. These apps hold promise in personalizing and tailoring behavior change interventions based on real-time data, thereby enhancing health outcomes [20]. Another crucial consideration is the necessity for co-designing apps in collaboration with the end users and health professionals from the relevant sector [21]. This is particularly vital for older adults as their participation in the technological design process is pivotal for their acceptance and adoption of the technology [22].

Digital interventions, whether delivered through mHealth apps or computers, are subject to evaluation in randomized controlled trials (RCTs) targeting various conditions such as psychotic disorders [23,24], mental health [25-28], eating disorders [29], diabetes [30], chronic pain [31], insomnia [32], or speech disorders [33]. These interventions often incorporate features such as audio or video content for physical activity or relaxation [23,24,26,32]; mood tracking [23,25,26]; personal tasks [23,25,31]; questionnaires [25,27,29,31]; educational materials [23-26,32]; self-tracking of, for example, vital signs in the form of a diary [24,32]; and habit libraries for behavior changes [34].

An umbrella review comprising 48 systematic reviews concentrating on mHealth apps in RCTs across various health conditions such as diabetes or hypertension indicates the potential effectiveness of app-based health interventions [35]. This further strengthens the approach of delivering a mobile-based app for study participants.

CTMSs for Intervention Projects

Equally important when intervention apps are used in clinical trials or research settings is central monitoring and management by the study coordinator or a corresponding physician. Using a CTMS and the integration of electronic data capture has become commonplace to serve this purpose [36]. Such a platform should facilitate and effectively support the multidimensional data management process in clinical trials [37] throughout their phases, from participant onboarding to completion [38].

Electronic systems allow for streamlined data transfer from clinics to the CTMS, remote enrollment capabilities, greater transparency of trial conduct, timeline monitoring such as specific tasks, tracking of participant visits, enhanced research documentation, and robust reporting [36,38,39]. A CTMS presents a variety of advantages; for example, it enables research teams to access up-to-date study information and simplifies collaboration as all project members can work efficiently together on the same task [38]. Moving beyond traditional methods of data collection in clinical trials, which may involve manual completion of paper case report forms (CRFs), leads to increased accuracy in findings, enhanced productivity as all necessary elements for managing the trial are consolidated in a single location, and higher data quality and compliance while reducing the risk of bias in clinical outcomes [36]. Important

XSL•FO
RenderX

considerations when selecting a CTMS include the feature set, usability, customization, and cost [38].

A systematic review of 19 research papers examining the technical features of clinical data management systems revealed that most of these systems were developed on a web-based platform to meet the individual needs of specific clinical trials within a short time frame [40]. Reportedly, such systems used in research centers showed limitations and inability to fully support the automation of all dimensions of the clinical data and workflow management process. In addition, the review found that most of the systems lacked flexibility and extensibility for further system development.

Health professionals such as study nurses, who are key members of the clinical research team and one of the targeted user groups for CTMSs, play a critical role in achieving accurate outcomes for clinical research studies. Furthermore, the effective implementation of a CTMS in RCT studies can have multiple benefits, including improved completion rates and increased fidelity while ensuring the safety of individual research participants [36].

The studies mentioned in this section collectively underscore the value of CTMSs in enhancing the efficiency and effectiveness of complex multisite intervention projects.

Objectives

This paper describes the iterative design process and reports on the technical implementation of the LETHE app and the LETHE CTMS within the context of a multifaceted and hybrid intervention to prevent cognitive decline in older adults. By sharing the results, we aim to contribute valuable insights to the field of dementia prevention and risk factor management through personalized ICT-supported hybrid intervention methods. We collaborated intensively with end users and health professionals to design the clinical trial platform. In addition, in this study, we evaluated how the trial participants and clinicians accepted the developed applications by using subjective usability assessments. This helped us to understand whether such digital tools keep the participants more engaged and identify directions for further improvement.

Methods

Design of a Mobile App and Clinical Trial Platform

Overview

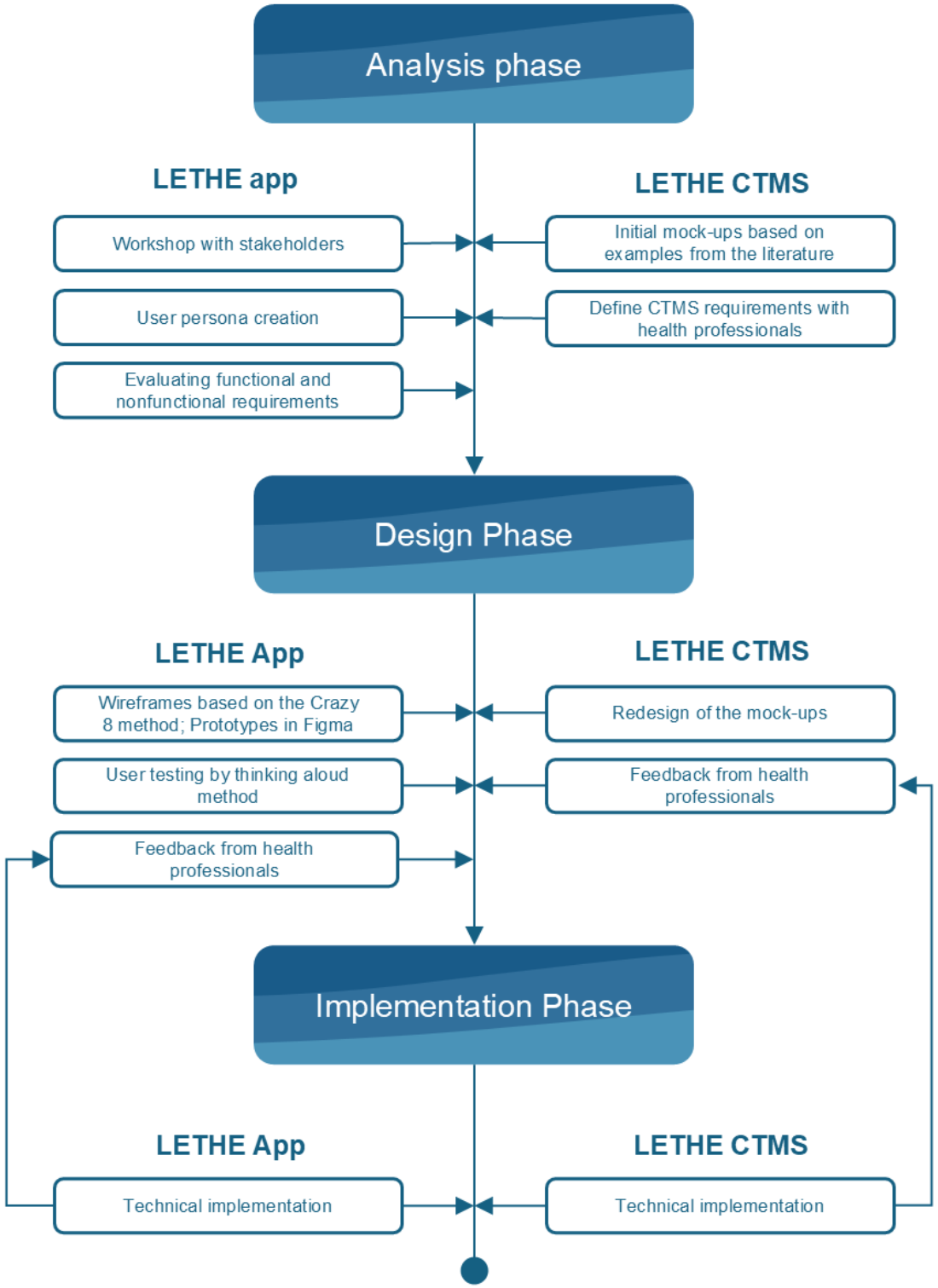
To ensure high usability, the user-centered design (UCD) approach was chosen for the LETHE app as well as for the CTMS design process. The UCD process involves the participation of end users (in our case, older individuals at risk of cognitive decline and health professionals supporting them during a preventive lifestyle intervention using the CTMS) to define key user requirements. The design approach consists of 3 main phases: analysis, design, and implementation [41].

The UCD approach encompasses 5 key principles: defining user requirements, gathering and considering feedback from users to specify requirements, involving users from the outset to evaluate design iterations, consistently adhering to UCD, and using an iterative design process [42].

Figure 1 illustrates the different design phases of the LETHE app and the LETHE CTMS.

XSL•FO
RenderX

Figure 1. Different phases for the design and development of the LETHE app and LETHE clinical trial management system (CTMS).



Requirement Analysis and Design of the LETHE App

First, a requirement workshop, which was held online, was carried out with the goal of listing requirements for the entire LETHE ecosystem and especially for the LETHE app. A total of 21 individuals from the LETHE project, including clinicians, behavioral experts, and representatives of public involvement in dementia, were invited to the workshop.

During the workshop, 3 user personas [43] were created, which represented typical end users of the LETHE app. General requirements were evaluated via open-ended questions from a clinical view. Specific requirements such as the design or the scientific use were ranked according to importance. The functional requirements of the LETHE app included, on the one hand, the content features according to the FINGER lifestyle domains; on the other hand, general digital requirements were collected from professionals during the workshop. The design

XSL•FO
RenderX

of the digital intervention was based on established behavior change theories to ensure that the app's features were effective. We focused on understanding relevant psychological factors that influence behaviors related to dementia risk [44,45] by relying on multiple behavior change theories to achieve a comprehensive understanding [46]. We used intervention mapping elements [47] and relied on a behavior change taxonomy [48] to select strategies that would best target relevant psychological factors to promote behavior change. Methods such as self-monitoring, setting clear goals, and planning for challenges were built into the app's features. Throughout the design process, a collaboration with health professionals and an advisory board of older adults at risk of or living with dementia was established. The results of this workshop were collected via online collaborating tools such as Mentimeter [49] and Padlet [50].

Afterward, wireframes based on the requirements and identified personas were created. During the user experience design, the method "The crazy 8" for wireframing was used to create as many screens as possible [51]. This method is a fast-sketching method to engage designers in sketching 8 distinct ideas in 8 minutes. In the next step, the wireframes were built using Figma (Figma, Inc) [52] as clickable prototypes to obtain continuous feedback from the involved health professionals.

In addition, a first user test with 4 German-speaking citizens aged between 65 and 85 years was carried out at the Department of Neurology at the Medical University of Vienna, and results were integrated during the implementation phase. One participant had a diagnosis of dementia, whereas the other 3 had no diagnosis. The user tests were conducted using the wireframes as clickable prototypes. First, participants explored the prototype using the thinking-aloud method [53]. Afterward, they were asked to perform 2 specific tasks (ie, choosing among a prepopulated list of pieces of advice for lifestyle improvement and setting them as goals for personal change and entering self-measured blood pressure values). Finally, participants provided verbal feedback and shared ideas and recommendations to further improve the prototypes.

When an implementation-ready design was obtained after the workshop and the user test, the implementation phase (including translation tasks) involved continuous feedback from professionals. Textbox 1 summarizes the phases for creating the design of the LETHE app.

Textbox 1. Phases for creating the design of the LETHE app, with steps 4 and 5 being recurring.

- Workshop for creation of user personas and functional and nonfunctional requirements (step 1)
- Creation of wireframes and clickable prototypes based on the workshop (step 2)
- User testing with potential end users (step 3)
- Adaption of wireframes and implementation phase (step 4)
- Continuous improvement, modification, and development after feedback from professionals and suggestions by study participants and advisory board members (step 5)

Requirement Analysis and Design of the LETHE CTMS

In the initial phase of our UCD methodology, an interactive workshop dedicated to the analysis of user requirements for the LETHE CTMS was conducted. The workshop was held online via Microsoft Teams (Microsoft Corp) with 13 health professionals as participants from different study centers located in Finland, Italy, Sweden, and Austria, as well as technical partners involved in the LETHE project. The professionals of the study centers included neurologists, gerontologists, public health experts, and professionals in the field of quality of life and dementia who represented the end users of the CTMS. The aim was to collaboratively identify the essential features and functionalities required for the LETHE CTMS. The workshop used the Mural tool [54], facilitating the creation of digital whiteboards using sticky notes for brainstorming in smaller breakout rooms consisting of 3 to 4 persons. In those breakout rooms, initial design considerations were discussed. Those considerations included general thoughts on a CTMS, an overview of the participants, data entering, and an artificial intelligence (AI) risk simulation. The process behind the design of the CTMS is described in Textbox 2 [55-57].

The results of the workshop influenced the subsequent design phase, during which multiple design proposals were iteratively developed incorporating feedback from clinical professionals. Once the design phase concluded, we transitioned into the implementation phase.

Textbox 2. Process to design the LETHE clinical trial management system, with steps 3 and 4 being recurring steps.

- Creation of initial mock-ups based on examples from the literature [55-57] and the study protocol of the LETHE trial for workshops together with clinical and technical partners (step 1)
- Conduction of the workshops to gather feedback and potentially missing features (step 2)
- Redesign of the mock-ups and technical implementation (step 3)
- Feedback on the redesign (step 4)

XSL•FO
RenderX

Evaluation of the Mobile App and Clinical Trial Platform

Overview

This section provides a detailed overview of the methodology used to analyze user perspectives within the currently ongoing 2-year clinical feasibility trial, focusing on both the LETHE app and the LETHE CTMS (or LETHE dashboard). The aim was to gather insights on basic usability assessments, encompassing considerations such as satisfaction with the applications.

The analysis was conducted using R (version 4.3.2; R Foundation for Statistical Computing) [58]. The normality test for the data was carried out using the Shapiro-Wilk test (α=.05). To assess the difference in the System Usability Scale (SUS) scores [59] between the study groups, the Mann-Whitney *U* test was used. For pairwise comparisons of the various countries, the Kruskal-Wallis test followed by the Dunn post hoc test with Bonferroni correction was used.

Evaluation of the LETHE App

Older adults who participated in the LETHE RCT provided feedback within the LETHE app through a structured questionnaire, "Experiences with LETHE app," consisting of 10 closed-ended questions based on an adapted version of the SUS, which was simplified for older adults and adults with cognitive impairment [60]. Using a Likert Scale [61] with 5 points ranging from "Strongly Disagree" to "Strongly Agree," participants expressed subjective feelings about the frequency of LETHE app use, perceptions of app complexity, and their need for assistance or confidence in navigating the mobile app. The questionnaire can be found in Multimedia Appendix 1.

This survey will be administered at 3 distinct time points throughout the study: 1 month, 6 months, and 24 months after randomization. Participants have 4 weeks to complete the survey at each time point. This timeline allows for the observation of participant sentiments over an extended duration, offering insights into the subjective aspects of use patterns and the perceived complexity of a mobile app specifically designed for older individuals. As the trial is still ongoing, this study focused on data from the first time point. As a result, we share participants' initial user experiences at this stage.

The LETHE app exists in 2 versions: one for the intervention group and one for the control group. Due to the limited functionality of the version for the control group, which lacks more sophisticated features such as personalized or intervention activities, the analysis was conducted separately for the intervention and control groups.

Within the LETHE trial, 156 study participants were invited to complete the questionnaire, with 78 (50%) in the intervention group and 78 (50%) in the control group. The calculated SUS was assigned a grade [62] as shown in Table 1, with "C" as the average grade and 68 as the center of the range.

Table 1. Interpretation of the System Usability Scale (SUS) score based on a grading scale by Sauro and Lewis [62].

SUS range	Grade	Percentile range
84.1-100	A+	96-100
80.8-84	A	90-95
78.9-80.7	A–	85-89
77.2-78.8	B+	80-84
74.1-77.1	B	70-79
72.6-74	B–	65-69
71.1-72.5	C+	60-64
65-71	C	41-59
62.7-64.9	C–	35-40
51.7-62.6	D	15-34

Evaluation of the LETHE CTMS

To obtain a comprehensive understanding of perspectives from relevant health professionals who used the LETHE CTMS, a 41-question online survey was conducted including both open- and closed-ended questions via Microsoft Forms (Microsoft Corp). The survey was distributed 10 months after the start of the study, and users were given 3 weeks to complete it. The primary objective was to gather feedback on the functionalities of the LETHE CTMS and assess user satisfaction with its implementation. The survey aimed to provide insights essential for mitigating common bottlenecks in the future development of similar tools. Notably, the survey was self-administered and structured into different parts, starting with a brief introduction outlining its objective before participants provided information about their role in the project. Subsequently, general inquiries about the CTMS, such as the onboarding process and overall experience, were posed. Participants then provided feedback on each functionality. Finally, recommendations, challenges, and additional comments were solicited.

All users of the LETHE CTMS were invited to participate and were informed about the survey duration.

The mixing of open- and closed-ended questions served a dual purpose. Closed-ended questions, including Likert scaling and ranking formats, enabled the quantification of participant sentiments, whereas parallel open-ended inquiries facilitated the collection of qualitative feedback. Noteworthy is the

XSL•FO
RenderX

inclusion of a single ranking question designed to identify the beyond-state-of-the-art features of the LETHE CTMS.

Moreover, a 5-point Likert scale [61] was used for closed-ended questions, encompassing the following response options: “Strongly Agree,” “Agree,” “Neither Agree nor Disagree,” “Disagree,” and “Strongly Disagree.”

The structure of the survey can be found in Multimedia Appendix 2.

Ethical Considerations

The trial has been approved by ethical committees in Austria (Ethics Committee of the Medical University of Vienna; 1392/2022), Finland (Hospital District of Helsinki and Uusimaa Ethical Committee; HUS/13675/2022), Italy (Regional Ethics Committee Umbria; 25723/22/AV), and Sweden (Swedish Ethical Review Authority; 2022-03961-01). All participants provided written informed consent before enrollment. The data used in this study for analysis were anonymized. Study participants received a wearable device, and a phone if they chose to have one. Some of those in the intervention group were additionally provided with a tablet for use during the trial.

Results

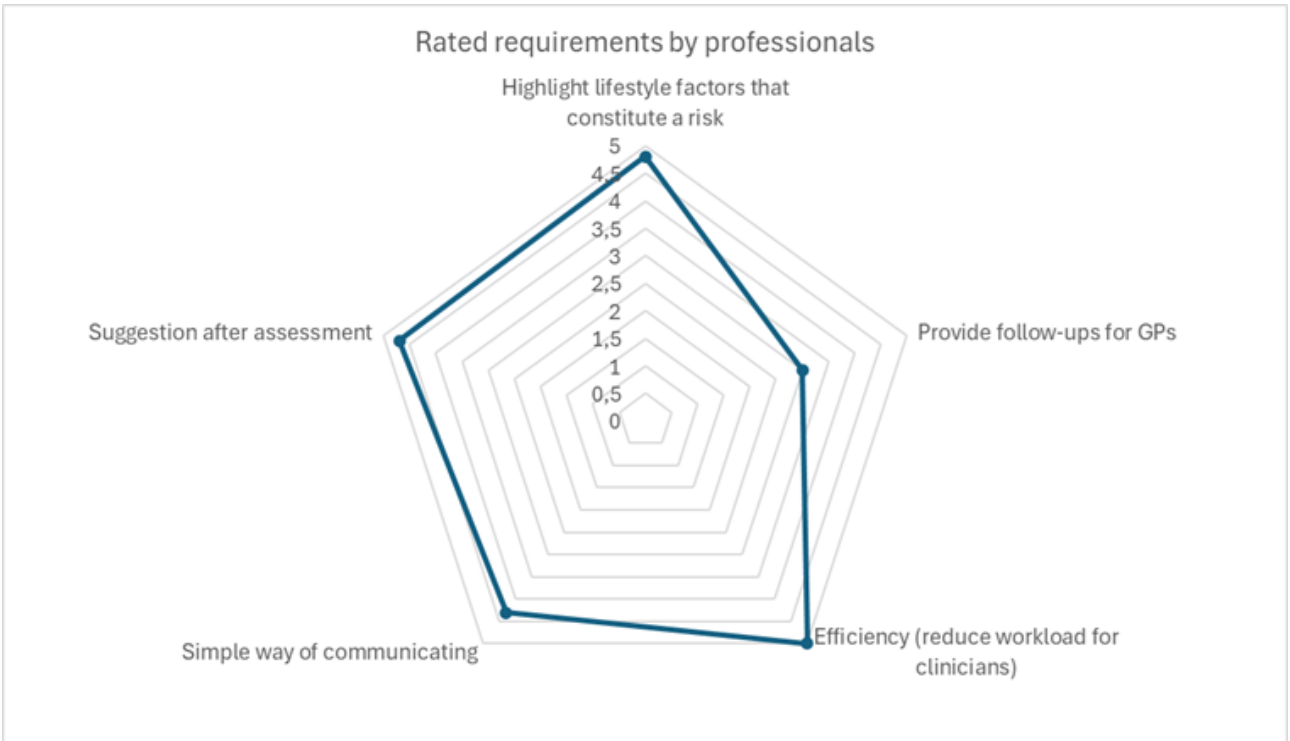
Design, Functionalities, and Evaluation of the LETHE App

Results of the Requirement Workshop and End-User Testing

The workshop together with the clinical and technical partners identified several requirements regarding the LETHE app, including the highlighting of lifestyle factors that constitute a risk or where the participant could improve. Additional observations include the LETHE app’s provision of guidelines to identify and flag study participants who are not using it. The communication between the specialists and the study participants should be simple. There was also an emphasis on data transfer and data sharing, highlighting that study participants can independently input data without professional assistance, such as for questionnaires.

The identified requirements were categorized and ranked via a survey. The main outcomes for the LETHE app are summarized in Figure 2. The highest-ranked requirements were to reduce the workload for professionals and highlight the lifestyle factors that constitute a risk. The least important requirement was to provide follow-ups for general practitioners. Those main findings were considered when designing the prototype.

Figure 2. Categorized and ranked requirements for the LETHE app. GP: general practitioner.



While conducting the user evaluation with the 4 participants at the Department of Neurology at the Medical University of Vienna, several key findings were obtained. The size of the user interface elements was criticized as being too small, and the tiles were not understood as large buttons. A section about the summary of the previous week as well as the share function of third-party apps (eg, for cognitive training or videos), which allows users to view content on another device via a QR code or by emailing it to themselves, was not considered intuitive to understand. The given tasks were missing additional information and were not clear without further explanation. Suggestions from the user testing included adding a notification for remembering to eat vegetables and fruits, as well as the marking of special input data values.

XSL•FO
RenderX

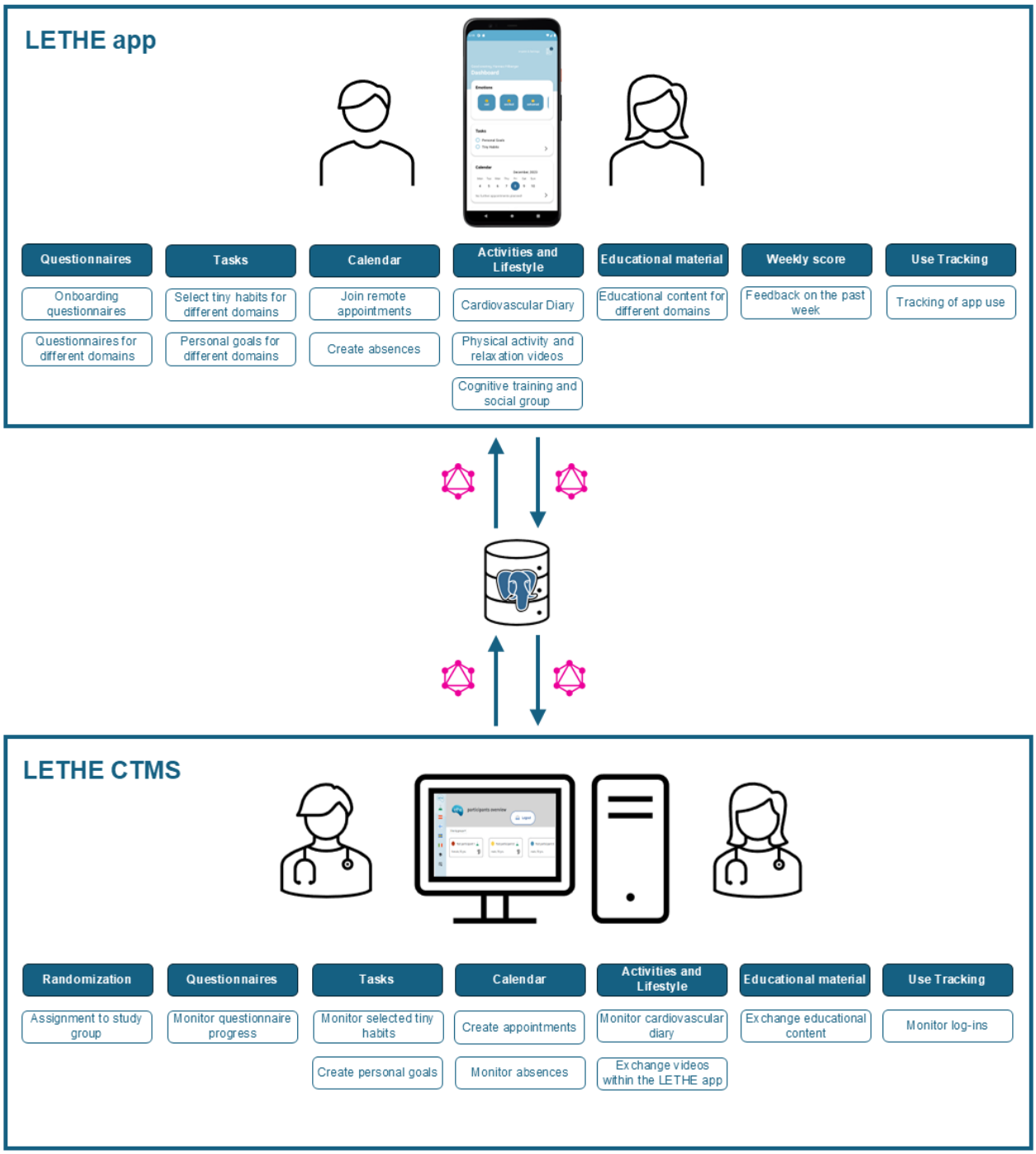
Functionalities of the LETHE App

Overview

The trial incorporates the LETHE app, an Android-native app using Java, to facilitate digital aspects and streamline data collection from study participants. The LETHE app has been released through the Google Play Store and is publicly available for download. However, access to the app's content is restricted to authorized study participants. Data storage and retrieval are managed through a .NET 7 (Core; .NET Foundation) application programming interface, which exposes data via GraphQL (Meta Platforms) protocols for external interaction with a PostgreSQL database (PostgreSQL Global Development Group). Figure 3 illustrates the modules of the LETHE app and the interaction with the LETHE CTMS.

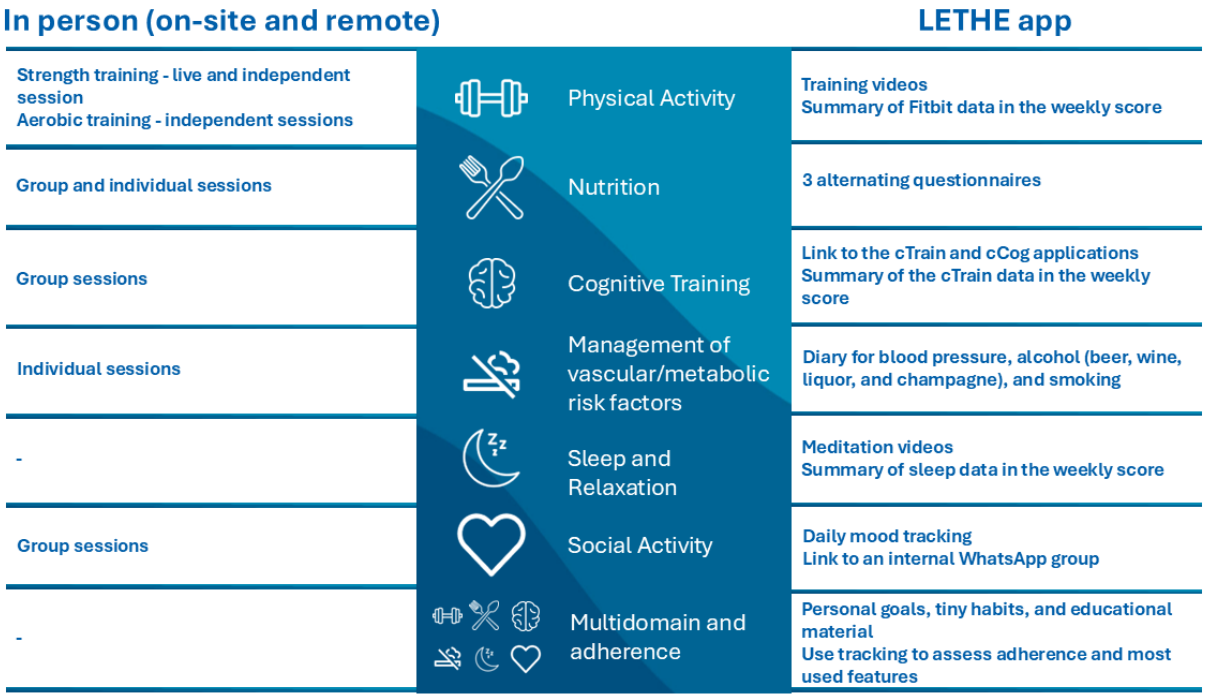
The LETHE app correspondingly aligned the lifestyle domains with a specific digital feature based on the FINGER study [5] and included the additional domains of *sleep/relaxation* and *social activity*. Figure 4 illustrates a comparison between the in-person and digital components of the lifestyle intervention trial.

Figure 3. Overview of the modules, functionalities, and interactions between the LETHE app and the LETHE clinical trial management system (CTMS).



XSL•FO
RenderX

Figure 4. Comparison between in-person and digital components for each lifestyle domain.



After completing the baseline study visit and initial questionnaires through the LETHE app, participants are randomly assigned to either the intervention or control group through the LETHE CTMS. Subsequently, the LETHE app dynamically adjusts the content based on group assignment. Both groups have access to the calendar, settings, questionnaires, mood tracking, and educational content. However, the intervention group has additional functionalities, such as personalized activities and the LETHE lifestyle program.

The LETHE app supports 5 languages (English, Finnish, Italian, Swedish, and German), with translations provided by the study centers. To enhance flexibility for features such as tiny habits and questionnaires, most of the translated content is dynamically retrieved, facilitating faster translation adjustments and content extensions. The translation process involved preparing an initial draft in English, which was then translated by native speakers at the clinical centers. The translations were reviewed to ensure accuracy for each region.

Designed with an older population in mind, the LETHE app uses a tile-based approach with large tiles as entry points for each functionality. This design allows individual elements to be hidden, shown, or reused, creating a modular structure. Following this approach, each component could be packaged independently, facilitating potential reuse in other research projects with similar requirements. Screenshots of each functionality of the LETHE app can be found in Multimedia Appendix 3.

Questionnaires

The questionnaires may be completed at single or multiple time points, with variations based on study group or participant gender. Each questionnaire consists of a title, due date, progress bar, and background color divided into green, yellow, and red, indicating the time remaining to complete the questionnaire. Participants can answer questions across multiple sessions, with answers cached on the device and sent only upon completion. The technical structure adheres to the Fast Healthcare Interoperability Resources (FHIR) standard [63], ensuring interoperability with other systems and enabling the questionnaire module's reuse by incorporating different FHIR questionnaires while maintaining the same structure.

LETHE App Dashboard

Once assigned to either the intervention or control group, participants unlock access to the LETHE app dashboard, which consists of different tiles to access various functionalities. Participants can choose here how they feel daily. Each possible mood item is accompanied by a short description and an emoticon for visualization. The list of the available mood items is dynamically adaptable and could expand in future development based on participant feedback.

Tasks

The *Tasks* section encompasses the personalized intervention module, comprising 2 distinct areas: personal goals and tiny habits. Personal goals are collaboratively established with clinical professionals during visits, targeting specific lifestyle domains for personalized interventions. Personal goals follow the specific, measurable, achievable, realistic, and time-bound principle [64]. Unlike tiny habits, personal goals are not predefined, but they can be packaged as a future library of personalized objectives. These goals can be set daily, weekly, or monthly, offering frequency flexibility, including options such as bidaily or biweekly. Study participants can mark goals as completed or incomplete. Each personal goal is bound to a lifestyle domain, providing participants with an indication of their personal domains for lifestyle improvement.

Tiny habits provide practical everyday tips and behavioral suggestions to help individuals implement manageable healthy

XSL•FO
RenderX

habits [65,66] and is adapted from the StopDia library [34], which is available in Finnish under a CC BY 4.0 license [67]. Tiny habits can be individually set by participants and are available for all lifestyle domains. Each tiny habit includes a description of the activity, a health fact about its benefits, and a place where it should be performed. There are >500 tiny habits in total to choose from, and they are assigned to a lifestyle domain. Given the subjective nature of tiny habit completion, participants are asked weekly about their success based on perceived completion in a questionnaire.

Calendar

The calendar displays all study-related appointments, both remote and in person, with a monthly and daily view. Remote visits can be joined directly from the daily view, which features larger tiles for individual appointments tailored to the target group. Participants can add vacation or unavailability periods, marking them as absences in the calendar. The absence menu is distinct in contrast to all other functionalities, providing a holiday like feel for participants.

Activity and Lifestyle

This section encapsulates the comprehensive LETHE app lifestyle program, featuring various activities tailored to distinct lifestyle domains. In the "Diary" section, study participants can log different health metrics such as blood pressure, alcohol consumption, and daily cigarette intake, directly contributing to the "Management of vascular/metabolic risk factors" lifestyle domain.

A range of activities aligned with different lifestyle areas unfold in the horizontal list at the top of the screen. It starts with a variety of videos dedicated to the areas of "Physical Activity" and "Relaxation." For the "Cognitive Training" domain, there is a link to an external application called cTrain, a cognitive training game package encompassing games similar to those used in the original FINGER study [5,68].

For the "Social Activity" domain, participants from the intervention group can engage with a country-specific and internal WhatsApp group facilitated by clinical staff, ensuring exclusivity for trial participants. The last tile in the list serves as an educational hub, presenting web resources including web feeds provided by professionals via the LETHE CTMS for each lifestyle domain.

Weekly Scores

The weekly score with motivational feedback messages offers participants feedback on the previous week. The feedback message adjusts if the score increased or decreased. The score displays 3 rings representing "App Data/Lifestyle," "Fitness," and "Brain Training." The "App Data/Lifestyle" category includes how often the LETHE app is used. The "Fitness" section displays Fitbit data, including step count, sleep duration, and individual fitness sessions, whereas "Brain Training" presents an overview of cognitive training data.

Use Tracking

To obtain insight into participants' adherence and identify features of particular interest to older individuals in a lifestyle intervention app, a mechanism for use tracking was developed. Each time a participant opens the LETHE app, a unique ID is assigned for the session. Every screen is assigned to an event, and each event is saved with new time stamps when navigating through screens until the participant closes the app. This monitoring technique aims to understand useful features and incorporate measures for adherence analysis [69].

Evaluation of the LETHE App

Overview

Overall, of the 156 participants, 123 (78.8%) provided responses. The distribution of SUS scores in both the intervention group (P=.03) and the control group (P=.02) deviated significantly from normality. However, there was no significant difference in SUS scores between the groups (P=.18).

Intervention Group

A total of 78 participants from the intervention group were invited for the SUS, and 66 (85%) responses were received. Of these 66 participants, 12 (18%) did not answer all the questions. The analysis revealed that the highest satisfaction levels were associated with learning how to use the app, whereas the lowest satisfaction levels were linked to the integration of various components, as shown in Figure 5 and Multimedia Appendix 4. The integration of the various components also received the most neutral responses among the questions. The question regarding the need for assistance when using the app generated the most disagreement, whereas learning about the app before use had the smallest response rate of 85% (56/66) for all questions.

Multimedia Appendix 4 provides an additional analysis for all 4 countries. Finland achieved the highest response rate (19/20, 95% of participant responses), and Sweden achieved the lowest response rate (12/18, 67%).

As 18% (12/66) of the participants did not provide complete answers, missing values were replaced with the neutral value of 3 following the approach outlined by Lewis [70] to calculate the SUS. Figure 6 highlights the presence of 1 outlier with a total score of 0, which was not excluded.

The median SUS score was 70 (IQR 55-82), which is above the average score and qualifies the app for a grade C, as shown in Table 2.

XSL•FO
RenderX

Figure 5. System Usability Scale (SUS) responses for the LETHE app in the intervention group. Panel (A) shows positively formulated questions, whereas panel (B) highlights negatively formulated questions. Green indicates high user satisfaction, and red indicates low user satisfaction. Q: question.

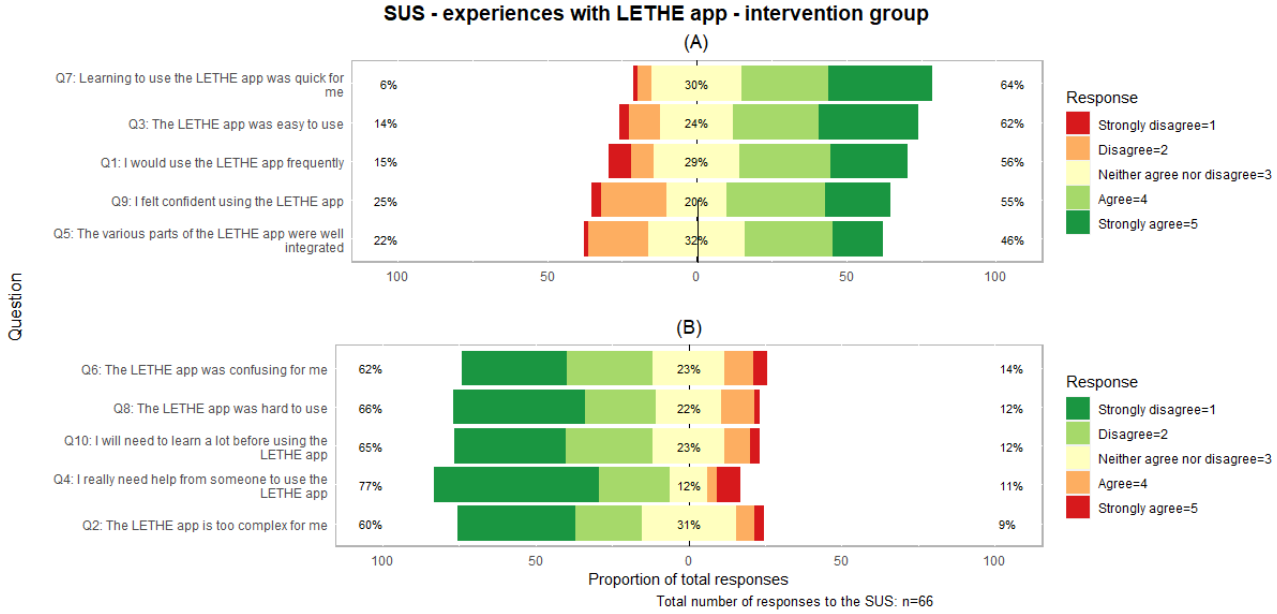


Figure 6. Box plot of the System Usability Scale (SUS) scores for each country in the intervention group.

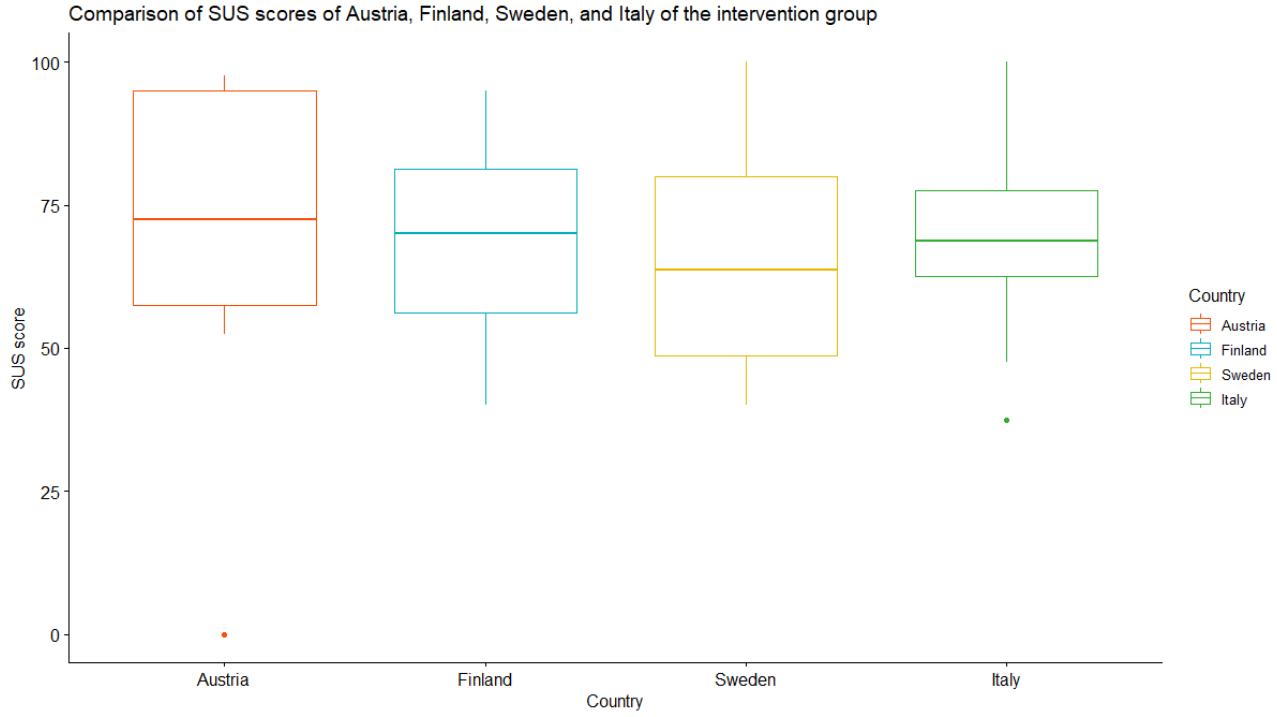


Table 2. Scores on the System Usability Scale (SUS) and its subscales for study participants in the intervention group.

	Measure[a]	Values, median (IQR)	Grade
Total (n=66)	SUS	70 (55-82)	C
Austria (n=17)	SUS	73 (58-95)	C+
Finland (n=19)	SUS	70 (57-82)	C
Sweden (n=12)	SUS	64 (49-80)	C–
Italy (n=18)	SUS	69 (63-78)	C

[a]The SUS ranges from 0 to 100, with 100 as the highest score.

XSL•FO
RenderX

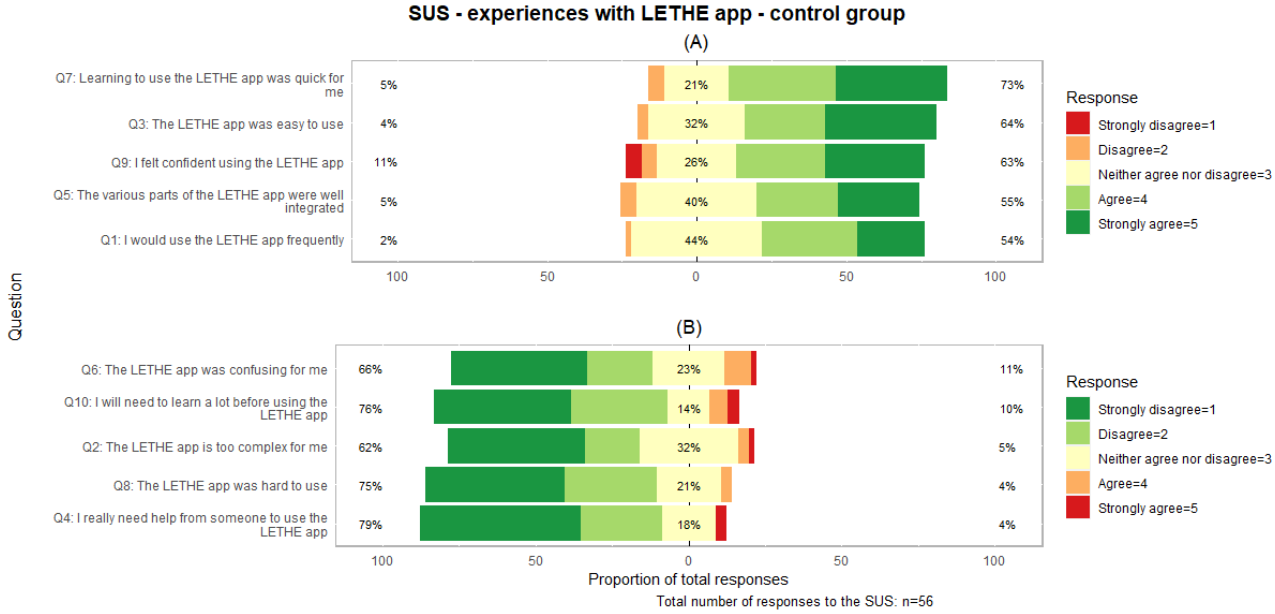
Austria achieved the highest median SUS score at 73 (IQR 58-95), whereas Sweden recorded the lowest median score at 64 (IQR 49-80). Austria was the only country where the app received a C+ score, distinguishing it from the others. Austria had 1 participant who was classified as an outlier with a total score of 0 but was still included in the analysis.

The distribution of SUS scores in Austria departed significantly from normality (P=.01). The distribution of SUS scores in Finland (P=.66), Sweden (P=.19), and Italy (P=.94) did not show evidence of nonnormality. The median SUS score did not differ significantly between the countries in the intervention group (P=.74).

Control Group

For the limited version of the LETHE app, 73% (57/78) of the study control group participants responded. Notably, 18% (10/57) of the respondents did not answer all questions. Figure 7 and Multimedia Appendix 5 indicate that the highest satisfaction was achieved in terms of how quickly participants could learn to use the LETHE app, followed by its ease of use. The only instance of a "Strongly disagree" response was regarding the confidence in using the LETHE app. The question about the LETHE app's frequent use gathered the most neutral responses, possibly due to its limited functionalities compared to the intervention version. A total of 79% (44/56) of the participants disagreed with the idea of needing help from others while using the LETHE app. Similarly to the intervention group, the question about learning about the LETHE app before using it received the smallest response rate of 98% (56/57) of all questions.

Figure 7. System Usability Scale (SUS) responses for the LETHE app in the control group. Panel (A) shows positively formulated questions, whereas panel (B) highlights negatively formulated questions. Green indicates high user satisfaction, and red indicates low user satisfaction. Q: question.



In the control group, Italy exhibited the lowest response rate, with 60% (12/20) of the participants responding, whereas Austria had the highest response rate, with 85% (17/20) of the participants responding. A detailed breakdown of responses per country is available in Multimedia Appendix 5.

To calculate the SUS, all missing values were substituted with the neutral value of 3, mirroring the approach taken with the intervention group. Table 3 illustrates that the median SUS score was 73 (IQR 63-90), corresponding to a B– grade.

Table 3. Scores on the System Usability Scale (SUS) and its subscales for study participants in the control group.

	Measure[a]	Values, median (IQR)	Grade
Total (n=57)	SUS	73 (63-90)	B–
Austria (n=17)	SUS	90 (73-95)	A+
Finland (n=16)	SUS	69 (65-76)	C
Sweden (n=12)	SUS	64 (50-84)	C–
Italy (n=12)	SUS	78 (54-90)	B+

[a]The SUS ranges from 0 to 100, with 100 as the highest score.

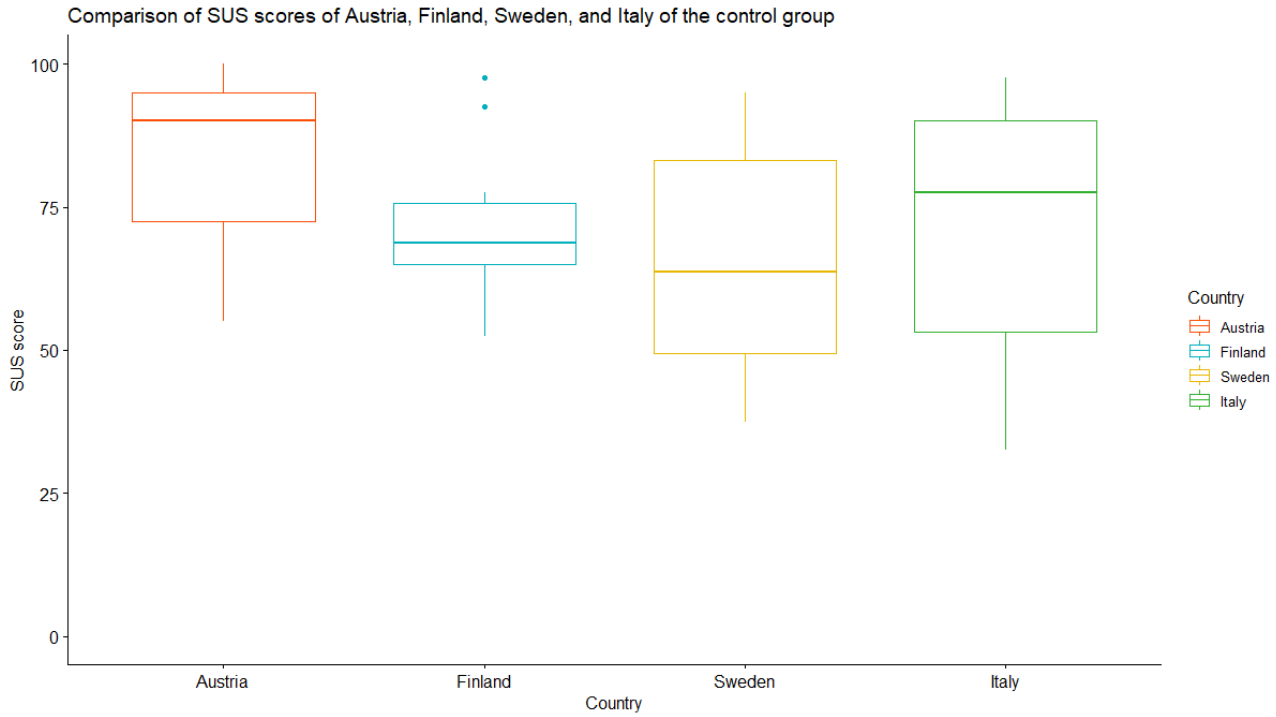
Austria had the highest median SUS score at 90 (IQR 73-95), giving the app an A+ grade, whereas Sweden had the lowest median score at 64 (IQR 50-84), giving the app a C– grade. The app obtained a C grade in Finland and a B+ grade in Italy, with median scores of 69 (IQR 65-76) and 78 (IQR 54-90), respectively. Notably, in contrast to the intervention group, the

XSL•FO
RenderX

median SUS score was higher in Austria and Italy and lower in Finland, and it stayed the same in Sweden. Figure 8 visualizes the box plot for each country.

The SUS scores of Austria (P=.03) and Finland (P=.047) departed significantly for normality. The distribution of the SUS scores of Sweden (P=.41) and Italy (P=.11) did not show evidence of nonnormality. The SUS scores did differ between the countries in the control group (P=.03). Pairwise comparisons showed that the median SUS score of Austria was significantly different than the median SUS score of Sweden (P=.02). For the other pairwise comparisons, no significant difference was found. Specifically, the comparisons between Austria and Finland (P=.33), Austria and Italy (P=.36), Finland and Sweden (P=.99), Finland and Italy (P=.99), and Sweden and Italy (P=.99) all yielded P values greater than the threshold for statistical significance.

Figure 8. Box plot of the System Usability Scale (SUS) scores for each country in the control group.



Design, Functionalities, and Evaluation of the LETHE CTMS

Feedback From the Workshop

After an examination of the LETHE study protocol and the examples from the literature [55-57] for a design perspective, initial mock-ups for the LETHE CTMS, a web application, were created. The mock-ups were kept very general and built the foundation for the discussions in the workshop with health professionals. These mock-ups encompassed (1) an overview page featuring details on all study participants, (2) a dedicated view for an individual study participant, (3) the conceptualization of an AI simulation, and (4) a data entry page aligned with the study protocol.

The feedback on the mock-ups included the exclusion of real names given that all study participants would be collectively visible on one page in the mock-up. Emphasis was placed on the immediate visibility of adherence using colors (eg, green, yellow, and red) and dropouts. In addition, diverse roles were implemented to restrict access to sensitive information, such as ensuring that health professionals could only view data related to their own country as well as a blinded role to not see the group of the study participants. Workshop participants also expressed the desire for a descriptive overview of the dataset variables.

Subsequently, the discussion turned to the overview page, where all study participants are listed. Key considerations included the inclusion of adherence information and the need for clear differentiation between participants from the intervention and control groups. Given the sizable participant list of 156 individuals, there was a request for sorting and search functionalities.

During discussions regarding the page presenting information for an individual study participant, workshop participants advocated for distinct perspectives tailored to individuals in the intervention and control groups, emphasizing adherence and dropout summaries. Further dialogues included role-specific views, aiming for the visibility of only relevant data. The workshop participants also expressed a preference for a comprehensive presentation of lifestyle categories, including detailed information and distinct sections for different clinical professionals as well as additional risk factors such as the consumption of alcohol or cigarettes. Requests for additional functionalities included the presentation of the Clinical Dementia Rating [71], allergies, questionnaire responses and scores, notes on participant interactions, details of adverse events such as muscle pain after a prescribed aerobic workout, and automated calculations such as BMI. Another necessity identified was the inclusion of a data export functionality.

XSL•FO
RenderX

When considering the AI dementia risk simulation, there was a discussion about whether it could be used during the study intervention or solely for research purposes. The AI risk simulation is currently available only to researchers due to ethical considerations. As the models used in the simulation have not yet been fully tested or validated, they were not implemented with real participants at this stage. Moreover, the risk information generated by the simulation is not disclosed to participants. Should the disclosure of this risk information be considered in the future, it would require careful evaluation. This process would involve consulting with both the participants and experts in ethics to assess the potential psychological impacts and the broader ethical implications of sharing risk-related information with participants.

The workshop discussed the insertion of all CRF data into the LETHE CTMS, encompassing visit-related data and scores such as the Clinical Dementia Rating and Neuropsychological Test Battery as in the FINGER study. Workshop participants underscored the inclusion of medications, validation checks, and prefillable fields. The study protocol served as a reference for defining the data entry page fields.

Following the conclusion of the workshop, participants were given a clickable prototype. Subsequently, a follow-up workshop was conducted to incorporate adaptations based on the initial workshop insights. The most significant modification was made to the data entry page, necessitating a complete restructuring due to a higher-than-anticipated number of data fields. After several iterations, a design for the beginning of the study was proposed.

Functionalities of the LETHE CTMS

Overview

Following the completion of the LETHE CTMS design, pivotal components were identified to commence the clinical trial. This included an overview of all study participants, a detailed view of a study participant, and the data input portal for clinical professionals. The LETHE CTMS was developed as a web application using React and uses the same backend infrastructure as the LETHE app. The LETHE CTMS is deployed as a Docker container (Docker, Inc). Once a new version is pushed to the main branch of the GitLab repository, a Continuous Integration/Continuous Delivery or Deployment (CI/CD) pipeline is automatically triggered, handling the build, release, and deployment process. After the container is pushed to the GitLab container registry, deployment is carried out using an Ansible playbook (Ansible Inc).

The LETHE CTMS has a consistent design throughout the application, achieved by using a fundamental page structure featuring headers, card-based content representation, a uniform color scheme inspired by the LETHE project colors, and font selection. Ensuring that the data are relevant only to the relevant users was achieved through ongoing communication with end users, and various roles were defined. There are roles that have access to all participants from all study centers; roles that are blinded and, therefore, can only enter visit forms; and roles that are not blinded and can have access to all participant details from their study center.

The LETHE CTMS underwent iterative modifications over time driven by continuous user feedback and evolving requirements. In particular, each section has been given a standardized nomenclature. At the moment, the following sections are included: overview page, detail page, clinician data entry page (containing the electronic CRF [eCRF]), and the configuration pages to adapt the content in the LETHE app. For now, the AI risk simulation is only available for research purposes and not visible to the end users, so a further description is not provided. Screenshots of the different pages can be found in Multimedia Appendix 6.

Overview Page

The overview page serves as a visual representation of all study participants within the selected country. Upon initializing this page, the study participants are loaded based on the study site affiliation of the clinical professional. The header features a legend outlining variables encompassing distinct adherence pathways and the respective participant groups. Color-coded adherence pathways range from “low” (red) to “medium” (yellow) to “high” (green) and include a category of “not calculated.” These pathways are grouped into distinct types, with more details provided on the detail page.

The central section of the page displays all participants from the current site within individual rounded boxes, which encompasses variables such as the participant’s ID, and other essential participant information, including adherence, group affiliation, age, and gender, is also presented.

There is also the possibility to access the configuration pages, allowing for the modification of content related to educational material and videos on the LETHE app.

Detail Page

Upon selecting an individual on the overview page, a more detailed view is presented, encompassing comprehensive data on that particular participant.

The detailed view interface is organized into 3 columns, each with functional components. In the first column, baseline visit data are featured, encompassing demographics, cardiovascular metrics, comorbidities, medications, allergies, and blood values. Users can review completed and pending questionnaires as well as access the responses and corresponding scores. This column concludes with a messaging feature that allows professionals to schedule notifications directly sent to study participants’ smartphones.

The second column commences with notes about adverse events, the entering and tracking of personal goals, and the monitoring of the selected tiny habits and cardiovascular risk factor data entered into the LETHE app. Professionals can also view participant absences to adjust appointment scheduling and reschedule personal goals accordingly.

In the last column, professionals can schedule recurring remote or in-person meetings with study participants together with reminders. Each meeting can be marked as completed to help track the adherence. Participant-professional contacts are recorded, and additional information is displayed, including last

XSL•FO
RenderX

LETHE app log-in, participant consent status, and participation in substudies.

At the bottom of the page, adherence to various categories, such as app use, cognitive activity, diet, and physical activity, is visually displayed, providing an overview of the participant's engagement with key aspects of the study protocol. Table 4 shows the thresholds used to assign each pathway to one of the adherence levels. On the basis of adherence and lifestyle domain, participants receive tailored messages on specific weekdays: physical activity on Mondays, app use on Tuesdays, diet on Wednesdays, cognitive activity on Thursdays, cardiovascular risk factors and social interaction on Fridays, and relaxation and sleep on Saturdays. Participants consistently in the red pathway receive domain-specific or holistic messages. In addition, tiny habit messages are sent on Mondays, Wednesdays, and Thursdays based on reported engagement.

Table 4. Adherence pathways and thresholds for different categories.

	Data type	Instrument used to measure the data and frequency of measurement	Green threshold	Yellow threshold	Red threshold
Physical activity	Activity min and frequency (moderate to vigorous intensity)	Fitbit; weekly (Mondays); passive	Months 1-3: 30-45 min (1 time per wk); months 4-6: 30-45 min (2-3 times per wk); months 7-9: 30-60 min (3-4 times per wk); months 10-24: 45-60 min (3-5 times per wk)	Months 1-3: 15-29 min (1 time per wk); months 4-6: 15-29 min (1 time per wk); months 7-9: 15-29 min (1-2 times per wk); months 10-24: 15-44 min (1-2 times per wk)	1 wk with 0 min or 3 weeks in a row in the yellow path
App use	Changes between the screens within the LETHE app	App; weekly (Tuesdays); passive	≥3 screens (4-7 d)	≥3 screens (0-3 d)	0 screens for 14 d
Diet patterns	Completing food-monitoring item questionnaire within the LETHE app (3 alternating questionnaires)	App; weekly (Wednesdays); active	2/3 or 3/3 blocks completed (in 3 wks)	1/3 blocks completed (in 3 wks)	0% completed (in 3 wks)
Cognitive training	Use of cognitive training program	cTrain; weekly (Thursdays); active	2-3 times per wk	0-1 times per wk	0 times for 3 wk

Data Entry Page

Diverging from the workshop, a series of design iterations led to the decision that the data entry page would act as a repository for information gathered from all study visits, whereas tasks such as contacting study participants and scheduling appointments were moved to the detailed view of an individual participant.

Within the eCRF, each primary section, covering screening, baseline, and subsequent visits, contains subforms that address different documentation areas (eg, forms to report results for blood tests, medication use, and neurological assessments such as the Mini-Mental State Examination [72]).

Each subform within the eCRF incorporates features such as automatic calculations and validation checks, whereby abnormal values trigger a visual highlighted in yellow. This ensures data accuracy and enhances the efficiency of data entry. The eCRF design, embedded with these functionalities, serves as a comprehensive and streamlined platform for recording and managing diverse aspects of study visit documentation.

Evaluation of the LETHE CTMS

Overview

This section analyzes the feedback from the CTMS survey completed by health professionals. First, user-specific details, such as their role in the project and within the application are presented. Second, the quantitative and qualitative feedback for each of the functionalities and the design is analyzed. Finally, a ranking of the state-of-the-art features is presented. The survey refers to the LETHE CTMS as *LETHE dashboard*, but both terms are interchangeable.

Users of the LETHE CTMS

The survey targeted all users of the LETHE CTMS, gathering general information such as their role in the project, their function within the application, and the frequency of their application access. When the survey was conducted, 21 users had access to the LETHE CTMS, and 15 (71%) responded to the survey. The summarized results are presented in Table 5. Most respondents held roles as coordinators or in unblinded positions, with access to almost all features. Of the 15 respondents, 12 (80%) accessed it at least once a week.

XSL•FO
RenderX

Table 5. Role of professionals in the LETHE project and in the LETHE clinical trial management system (CTMS) and frequency of CTMS use (N=15).

Answer option	Professionals, n (%)
"Could you please describe your role in the project?"	
Coordinator	1 (7)
Digital coach	2 (13)
Neuropsychologist	3 (20)
Nutritionist	2 (13)
Physiotherapist	1 (7)
Principal investigator	1 (7)
Study nurse	4 (27)
Study physician	1 (7)
"What dashboard role do you have?"	
Blinded role	1 (7)
Coordinator role	6 (40)
Unblinded role	8 (53)
"How frequently do you access the dashboard?"	
Daily	1 (7)
More than once a week	6 (40)
Once a week	5 (33)
More than once a month	1 (7)
Once a month	2 (13)

Feedback on the Design and Functionalities of the LETHE CTMS

Figure 9 and Table 6 indicate that the highest level of user satisfaction was related to the comprehensive display of relevant participant information in an easily understandable way, with a score of 100% above neutral. A detailed breakdown of all quantitative responses can be found in Multimedia Appendix 7. Overall, feedback was positive, reflecting successful integration of design approaches from workshops and feedback. Users found the LETHE CTMS intuitive, with a visually appealing and clear layout featuring consistent colors, logically divided sections, and pictograms. Pop-ups during item saving aided navigation and provided feedback on actions. Suggested improvements included reducing the steps to move between participants, addressing inconsistencies in single- and double-click requirements, and ensuring visibility of study participant phone information on the participant overview page.

XSL•FO
RenderX

Figure 9. Survey responses for the LETHE clinical trial management system from health professionals. Green indicates high user satisfaction, and red indicates low user satisfaction.

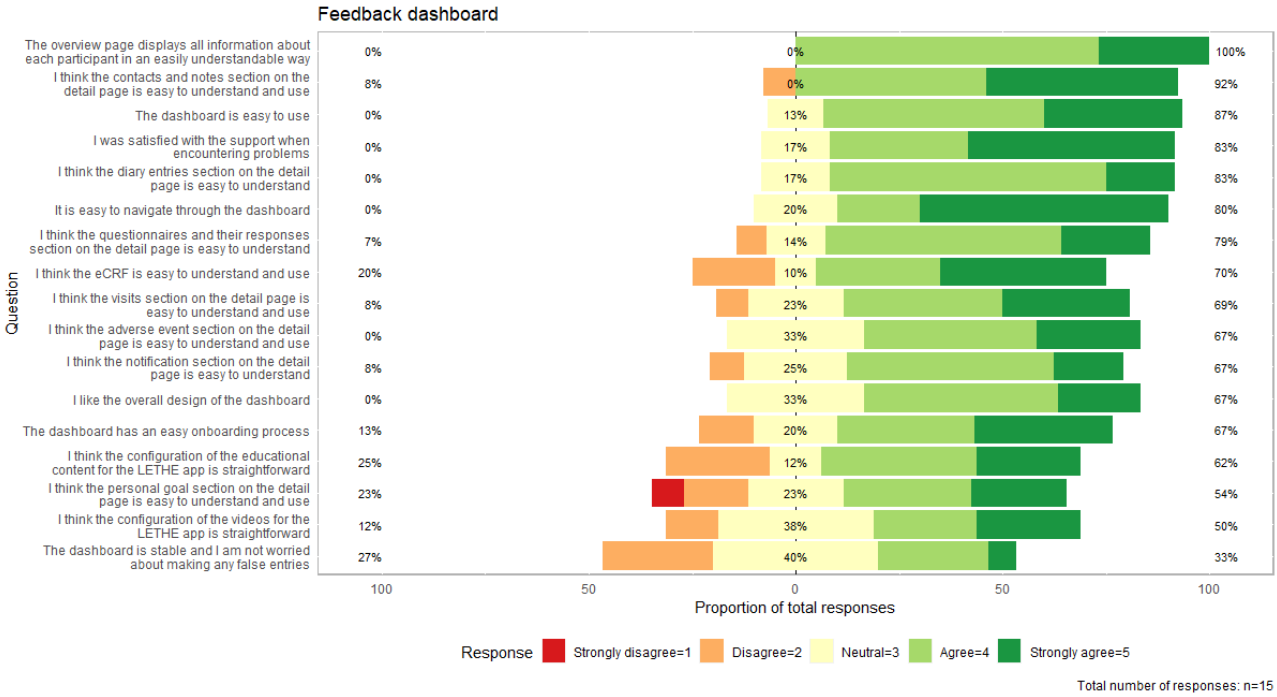


Table 6. Means and SDs of the LETHE clinical trial management system feedback survey targeting the different functionalities.

	Values, mean (SD)[a]
“The overview page displays all relevant information about each participant in an easily understandable way” (n=15)	4.27 (0.46)
“I think the Contacts/Notes Section on the Detail Page is easy to understand use” (n=13)	4.31 (0.85)
“The Dashboard is easy to use” (n=15)	4.20 (0.68)
“I was satisfied with the support when encountering problems” (n=12)	4.33 (0.78)
“I think the Diary Entries Section on the Detail Page is easy to understand” (n=12)	4.00 (0.60)
“It is easy to navigate through the Dashboard” (n=15)	4.40 (0.83)
“I think the Questionnaires and their Responses Section on the Detail Page is easy to understand” (n=14)	3.93 (0.83)
“I think the eCRF is easy to understand and use” (n=10)	3.90 (1.20)
“I think the Visits Section on the Detail Page is easy to understand and use” (n=13)	3.92 (0.95)
“I think the Adverse Event Section on the Detail Page is easy to understand and use” (n=12)	3.92 (0.79)
“The Dashboard has an easy onboarding process” (n=15)	3.87 (1.06)
“I like the overall design of the Dashboard” (n=15)	3.87 (0.74)
“I think the Notification Section on the Detail Page is easy to understand” (n=12)	3.75 (0.87)
“I think the configuration of the Educational Content for the LETHE App is straightforward” (n=8)	3.63 (1.19)
“I think the Personal Goal Section on the Detail Page is easy to understand and use” (n=13)	3.46 (1.27)
“I think the configuration of the Videos for the LETHE App is straightforward” (n=8)	3.63 (1.06)
“The Dashboard is stable and I am not worried about making any false entries” (n=15)	3.13 (0.92)

[a]The scale ranges from 1 (*Strongly disagree*) to 5 (*Strongly agree*). A higher level of agreement indicates greater satisfaction.

Concerning the onboarding process, users expressed that it involved multiple steps and should be streamlined further. However, once granted access, the general opinion was that the process was straightforward. The lowest satisfaction regarding functionalities was observed in the educational content and video configuration, where users missed options for uploading videos and materials instead of web links. The stability and prevention of false entries, particularly in the eCRF, received the lowest satisfaction rating. Users reported issues with scrolling affecting eCRF values and occasional problems with

XSL•FO
RenderX

saving values. Suggestions included implementing a log to track data edits for added transparency, a measure already in place but not visible to users. Another notable disagreement in the *Personal Goals* section was that users found it difficult to set goals at regular intervals, and the lifestyle domain associated with each personal goals was not visible in the detailed view of a study participant.

Feature suggestions included displaying follow-up data, editing values in different sections, providing additional Fitbit data insights, and summarizing clinical test names and scores at the bottom of the eCRF page. Positive aspects highlighted satisfaction with support, emphasizing quick response times, correction of implementations, and clear explanations.

Other feedback suggested the ability to send messages to groups of study participants as well as create group meetings. Overall, those surveyed commented on the effectiveness of the LETHE CTMS for study staff and participants, noting a continuous improvement since its first use.

Ranking the Most Useful and Beyond-State-of-the-Art Features

One of the survey questions asked respondents to rank various features integrated into the LETHE CTMS based on their perceived usefulness and how they surpassed state-of-the-art capabilities. As shown in Table 7, the integration of personalized activities (mean 2.23, SD 1.17) and real-time appointment planning (mean 2.46, SD 1.51) stood out as the most impactful features. An additional comment from a respondent highlighted the novelty of the interplay between the LETHE CTMS and the LETHE app, with a note that further optimization was needed. In addition, the LETHE CTMS holds the potential to serve as both an eCRF and a tool to assist in intervention delivery. This dual functionality was seen as a valuable development for the future to meet the evolving needs of professionals conducting clinical trials.

Table 7. Ranking of different features based on surpassing state-of-the-art capabilities, with R1 being the best-ranked feature and R7 being the worst-ranked feature.

	Choices (n=13), n (%)							Total (n=91), n (%)
	C1[a]	C2[b]	C3[c]	C4[d]	C5[e]	C6[f]	C7[g]	
R1	2 (15)	2 (15)	1 (8)	0 (0)	1 (8)	4 (31)	3 (23)	13 (14)
R2	1 (8)	0 (0)	0 (0)	0 (0)	1 (8)	4 (31)	7 (54)	13 (14)
R3	1 (8)	0 (0)	2 (15)	4 (31)	3 (23)	2 (15)	1 (8)	13 (14)
R4	0 (0)	3 (23)	3 (23)	3 (23)	1 (8)	2 (15)	1 (8)	13 (14)
R5	1 (8)	1 (8)	3 (23)	2 (15)	5 (38)	0 (0)	1 (8)	13 (14)
R6	2 (15)	5 (38)	1 (8)	3 (23)	1 (8)	1 (8)	0 (0)	13 (14)
R7	6 (46)	2 (15)	3 (23)	1 (8)	1 (8)	0 (0)	0 (0)	13 (14)
Total (n=91)	13 (14)	13 (14)	13 (14)	13 (14)	13 (14)	13 (14)	13 (14)	91 (100)

[a]Overview page of different study participants in different countries, including their digital intervention adherence pathways (mean 5.08, SD 2.43).

[b]Direct notifications to study participants via the dashboard (mean 4.84, SD 1.99).

[c]Electronic case report form (mean 4.69, SD 1.80).

[d]Immediate results of onboarding and questionnaires on the dashboard (mean 4.54, SD 1.40).

[e]Configuration of the content of the app via the dashboard (mean 4.15, SD 1.68).

[f]Real-time planning of appointments within the dashboard for the app (mean 2.46, SD 1.51).

[g]Integration of personalized intervention tasks (*Personal Goals*; mean 2.23, SD 1.17).

Discussion

Principal Findings

This paper outlines the design process, functionalities, and evaluation of the digital intervention study components within the LETHE project, focusing on older individuals at risk of cognitive decline and clinical professionals. The design process involved multiple sessions with potential end users and clinical experts. Furthermore, the setup is currently being evaluated in a 2-year intervention study, and the first results have been presented [11,69].

The original FINGER multidomain trial, along with related studies, demonstrated that cognitive benefits can be achieved through a 2-year intervention. As a result, this duration was established as the timeline for the LETHE trial. However, there is only limited information available on longer studies supported by ICT components, such as digital apps requiring daily interactions and wearable devices.

These requirements made it necessary to carefully design the study in terms of human support and create user-friendly ICT components for both participants and clinical professionals. The learnings regarding ICT use in clinical studies, as well as data collected in intervention studies (clinical data and digital biomarkers), are essential and can influence future studies and clinical trial setups using ICT.

XSL•FO
RenderX

To emphasize the importance of user-friendly design, this paper presents the process from requirement gathering to the design of the components. The findings will be evaluated and will inform further improvements to the setup.

Overall, the median SUS score of the intervention group of 70 (IQR 55-82) was comparable to that of the control group of 73 (IQR 63-90). The slightly higher median score of the control group might be attributed to the less complex functionalities compared to those of the intervention group version, specifically the absence of personalized and intervention activities. However, both median scores surpassed 68, indicating above-average user acceptance and satisfaction according to SUS guidelines [62]. Given that the initial user experience for the control group was quite positive even with fewer features, there is hope that they will remain engaged throughout the trial. It is anticipated that these values will improve as participants gain more experience with the LETHE app during the 2-year RCT [73]. Evaluating the other time points will give us more information about the longer-term usability and user engagement.

Noteworthy is the variation in median SUS scores between countries, with Austrian participants providing the highest scores (intervention group: median 73, IQR 58-95; control group: median 90, IQR 73-95) and Sweden providing the lowest scores (intervention group: median 64, IQR 49-80; control group: median 64, IQR 50-84). Performing a pairwise comparison using the Dunn test, a significant difference in median SUS scores between Austria and Sweden was found (P=.02).

While the LETHE lifestyle intervention program is centrally coordinated, and the activities in the 4 countries are harmonized to ensure comparable content, certain local adaptations are allowed to optimize feasibility (eg, in the detailed intervention delivery [balance between in-person and digital sessions] and how the digital components are leveraged). A qualitative study is currently underway to collect additional data on the barriers to and facilitators of the use of digital tools among participants in the LETHE trial.

Overall, users expressed satisfaction with the LETHE CTMS, and workshops on design, usability, and functionality proved beneficial. However, clarity regarding the purpose of an AI risk simulation and addressing user concerns about inaccurate data entry emerged as key challenges during the design process. Further exploration into user satisfaction with the intervention's introduction by clinical professionals across different countries is warranted for future research.

Research should investigate reasons behind decreased personal engagement in long-term ICT-supported studies. The LETHE protocol, with its 2-year ICT-supported design, offers a unique contribution to support cognitive health in older individuals.

In terms of future research, several key directions should be explored. First, a thorough evaluation of the platform's use and usability after the completion of the full 2-year trial would provide valuable insights into its long-term effectiveness. Second, further investigations could focus on the AI-driven risk simulation within the CTMS, with a particular emphasis on its potential role in guiding future interventions. Finally, exploring the adaptability of the LETHE platform among populations with low digital literacy and assessing the feasibility of a hybrid approach would be an important area for future research.

We are confident that LETHE app's diverse features, together with high retention, positive feedback and frequent use [11,69], provide promising directions for future hybrid multimodal interventions.

Strengths and Limitations

This study has several strengths, particularly in how the LETHE app and CTMS incorporate perspectives from both northern and southern Europe. This allows for a broader range of views to be considered. Moreover, this study is regularly improved through feedback from professionals, suggestions from study participants, and input from advisory board members, ensuring ongoing refinement and relevance. One limitation lies in assuming that all participants have sufficient digital skills and internet access, which was an inclusion criterion for the RCT. However, according to a recent study, the gap in digital health trends between younger and older people may vanish in 10 years as today's individuals aged <65 years are highly adapted to digital solutions [74].

Furthermore, all participants were provided with introductory materials and manuals for the LETHE app to facilitate use. Those factors could impact their ratings. Moreover, the subjective nature of the SUS, which relies on participants' self-reported perceptions of usability, might not fully grasp the nuanced usability challenges, especially among older users who may have varying cognitive and physical abilities. Another limitation is that the LETHE app content (eg, educational materials or videos) varies between countries and may influence the ratings.

Another consideration is that the initial user testing involved a small sample size, with only 4 German-speaking participants for the first user testing. However, the evaluation of the LETHE app is part of a study in which health professionals and 156 study participants are included.

The LETHE app was developed exclusively for the Android platform. This decision was based on practical considerations, including resource constraints, time limitations, and the need for a streamlined pilot phase. Focusing on a single operating system allowed for a more efficient development, testing, and release process, ensuring that the feasibility trial could be completed within the available time frame. Android was chosen due to its widespread use and compatibility with various devices, facilitating broader accessibility. To mitigate potential accessibility barriers, participants who did not have an Android device were given the option to receive one for the study. This approach enabled a standardized evaluation of the app's functionality and user experience. Future iterations of the app may consider cross-platform availability, including an iOS or web version.

Regarding the LETHE CTMS, a limitation of the survey is that most users held roles with the highest level of privileges in terms of functionalities, with frequent role changes occurring throughout the project. Improved planning of roles is advisable for future applications to mitigate this issue.

XSL•FO
RenderX

Comparison With Prior Work

Previous studies indicate declining user engagement with mHealth interventions over time due to multiple factors, such as unengaging content, complex protocols, or poor design. For instance, studies such as predictive model-based decision support for diabetes patient empowerment (POWER2DM) [75] have investigated ICT-supported health systems integrating features such as shared decision-making, personal goal setting, mood tracking, and exercise components, which is similar to the interaction between the LETHE app and LETHE CTMS but within the context of diabetes self-management.

Compared to existing digital interventions for dementia prevention, such as PRODEMOS [15] and Maintain Your Brain [16], LETHE introduces several additions and refinements. LETHE builds on standard goal setting but also includes the tiny habits method as an add-on. It also covers additional lifestyle domains, including social engagement and sleep, which are not included in PRODEMOS but are, meanwhile, part of the Lancet Commission's report [2].

In terms of data collection, LETHE uses Fitbit devices to passively monitor physical activity and sleep, whereas PRODEMOS relies on self-reported step counts. Both interventions use a hybrid approach, but LETHE's study duration is extended to 24 months (compared to 18 months in PRODEMOS) and builds on the successful FINGER protocol [5].

LETHE differs from Maintain Your Brain by offering immediate access to all modules rather than introducing them sequentially. Although LETHE does not include a separate mental health module, it collects related information through questionnaires. Maintain Your Brain, compared to LETHE, is entirely internet based and spans 3 years, with only quarterly booster activities after the first year and yearly follow-up assessments for the remaining 2 years.

Finally, unlike both PRODEMOS and Maintain Your Brain, a primary outcome of LETHE is the feasibility of a digitally supported multimodal lifestyle intervention, assessed through participant engagement with the LETHE app and a Fitbit device, attendance to study activities, and implementation of a hybrid approach that combines active and passive data collection across 4 countries.

One significant factor contributing to the unsatisfactory use of ICT components in ICT intervention studies is the usability of the functions provided. This paper provides a comprehensive account of the design process for both the app and the CTMS, aiming to maximize their usability. Subsequent publications will detail the outcomes concerning use and user satisfaction. Identified gaps in involving older adults in the design process of digital health technologies shed light on issues such as the exclusion of individuals with low digital skills during recruitment, emphasizing the importance of using diverse methods such as focus groups, interviews, or workshops [76]. While both the LETHE app and LETHE CTMS design processes incorporated various methods to solicit continuous feedback from diverse user populations, they did not explicitly include persons with low digital skills.

Conclusions

In summary, this paper presented and evaluated the design and functionalities of a comprehensive clinical trial system involving a mobile app tailored for older individuals at risk of cognitive decline. The LETHE app and a web-based monitoring and configuration system for clinical professionals, the LETHE CTMS, were thoroughly evaluated through survey-based assessments. Overall, our approach facilitates real-time interaction, providing 2 distinct applications for professionals and study participants. The results elucidated the design principles, stakeholder involvement, and essential functionalities of such an eHealth system. Subsequent research will delve into the posttrial use of the LETHE app over the 2 years, shedding further light on its effectiveness and user engagement.

Acknowledgments

The LETHE project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement 101017405 [77] (accessed on January 31, 2024). FM is also supported by an ALF (Region Stockholm) grant. This research is supported by a larger consortium, including FH JOANNEUM GmbH, Medical University of Vienna, University of Perugia, Karolinska Institute, Finnish Institute for Health and Welfare, Alzheimer Europe, Foundation for Research and Technology–Hellas, Maastricht University, Kaasa solution GmbH, Innovation2Grow, the EGI Foundation, Extra Red SRL, Infotrend Innovations Company Limited, Combinostics Oy, and the Lisbon Council for Economic Competitiveness asbl.

Conflicts of Interest

None declared.

Multimedia Appendix 1
LETHE app survey.
[PDF File (Adobe PDF File), 200 KB - aging_v8i1e66660_app1.pdf]

Multimedia Appendix 2
LETHE clinical trial management system survey.
[PDF File (Adobe PDF File), 366 KB - aging_v8i1e66660_app2.pdf]

XSL•FO
RenderX

Multimedia Appendix 3
LETHE app screenshots.
[PPTX File , 25265 KB - aging_v8i1e66660_app3.pptx]

Multimedia Appendix 4
LETHE app evaluation—intervention group.
[XLSX File (Microsoft Excel File), 15 KB - aging_v8i1e66660_app4.xlsx]

Multimedia Appendix 5
LETHE app evaluation—control group.
[XLSX File (Microsoft Excel File), 15 KB - aging_v8i1e66660_app5.xlsx]

Multimedia Appendix 6
LETHE clinical trial management system screenshots.
[PPTX File , 549 KB - aging_v8i1e66660_app6.pptx]

Multimedia Appendix 7
LETHE clinical trial management system cross-table.
[XLSX File (Microsoft Excel File), 12 KB - aging_v8i1e66660_app7.xlsx]

References

1. GBD 2019 Dementia Forecasting Collaborators. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. Lancet Public Health 2022 Feb;7(2):e105-e125 [FREE Full text] [doi: 10.1016/S2468-2667(21)00249-8] [Medline: 34998485]
2. Livingston G, Huntley J, Liu KY, Costafreda SG, Selbæk G, Alladi S, et al. Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. Lancet 2024 Aug 10;404(10452):572-628. [doi: 10.1016/S0140-6736(24)01296-0] [Medline: 39096926]
3. Kivipelto M, Mangialasche F, Anstey KJ. Pivotal points in the science of dementia risk reduction. Lancet 2024 Aug 10;404(10452):501-503. [doi: 10.1016/S0140-6736(24)01546-0] [Medline: 39096928]
4. Noach S, Witteman B, Boss HM, Janse A. Effects of multidomain lifestyle interventions on cognitive decline and Alzheimer's disease prevention: a literature review and future recommendations. Cereb Circ Cogn Behav 2023;4:100166 [FREE Full text] [doi: 10.1016/j.cccb.2023.100166] [Medline: 37215433]
5. Ngandu T, Lehtisalo J, Solomon A, Levälahti E, Ahtiluoto S, Antikainen R, et al. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. Lancet 2015 Jun 06;385(9984):2255-2263. [doi: 10.1016/S0140-6736(15)60461-5] [Medline: 25771249]
6. Thunborg C, Wang R, Rosenberg A, Sindi S, Andersen P, Andrieu S, et al. Integrating a multimodal lifestyle intervention with medical food in prodromal Alzheimer's disease: the MIND-AD randomized controlled trial. Alzheimers Res Ther 2024 May 30;16(1):118 [FREE Full text] [doi: 10.1186/s13195-024-01468-x] [Medline: 38812047]
7. Whitfield T, McConnell B, Renouf P, Mansour H, Zabihi S, Aguirre E, et al. The effect of remotely delivered lifestyle interventions on cognition in older adults without dementia: a systematic review and meta-analysis. Ageing Res Rev 2021 Dec;72:101505 [FREE Full text] [doi: 10.1016/j.arr.2021.101505] [Medline: 34757173]
8. Richard E, Moll van Charante EP, Hoevenaar-Blom MP, Coley N, Barbera M, van der Groep A, et al. Healthy ageing through internet counselling in the elderly (HATICE): a multinational, randomised controlled trial. Lancet Digit Health 2019 Dec;1(8):e424-e434 [FREE Full text] [doi: 10.1016/S2589-7500(19)30153-0] [Medline: 33323224]
9. Kourtis LC, Regele OB, Wright JM, Jones GB. Digital biomarkers for Alzheimer's disease: the mobile/ wearable devices opportunity. NPJ Digit Med 2019;2:9 [FREE Full text] [doi: 10.1038/s41746-019-0084-2] [Medline: 31119198]
10. Hanke S, Mangialasche F, Bödenler M, Neumayer B, Ngandu T, Mecocci P, et al. AI-based predictive modelling of the onset and progression of dementia. Smart Cities 2022 May 20;5(2):700-714. [doi: 10.3390/smartcities5020036]
11. Rosenberg A, Untersteiner H, Guazzarini AG, Bödenler M, Bruinsma J, Buchgraber-Schnalzer B, on behaf of the LETHE Consortium. A digitally supported multimodal lifestyle program to promote brain health among older adults (the LETHE randomized controlled feasibility trial): study design, progress, and first results. Alzheimers Res Ther 2024 Nov 21;16(1):252 [FREE Full text] [doi: 10.1186/s13195-024-01615-4] [Medline: 39574193]
12. Bruinsma J, Visser LN, Abaci A, Rosenberg A, Diaz A, Hanke S, et al. Social activities in multidomain dementia prevention interventions: insights from practice and a blueprint for the future. Front Psychiatry 2024 May 20;15:1386688 [FREE Full text] [doi: 10.3389/fpsyt.2024.1386688] [Medline: 38832328]

XSL•FO
RenderX

13. Di Lorito C, Bosco A, Rai H, Craven M, McNally D, Todd C, et al. A systematic literature review and meta-analysis on digital health interventions for people living with dementia and mild cognitive impairment. Int J Geriatr Psychiatry 2022 Jun 19;37(6):5730 [FREE Full text] [doi: 10.1002/gps.5730] [Medline: 35588315]
14. Sohn M, Yang J, Sohn J, Lee JH. Digital healthcare for dementia and cognitive impairment: a scoping review. Int J Nurs Stud 2023 Apr;140:104413. [doi: 10.1016/j.ijnurstu.2022.104413] [Medline: 36821951]
15. Eggink E, Hafdi M, Hoevenaar-Blom MP, Song M, Andrieu S, Barnes LE, et al. Prevention of dementia using mobile phone applications (PRODEMOS): protocol for an international randomised controlled trial. BMJ Open 2021 Jun 09;11(6):e049762 [FREE Full text] [doi: 10.1136/bmjopen-2021-049762] [Medline: 34108173]
16. Heffernan M, Andrews G, Fiatarone Singh MA, Valenzuela M, Anstey KJ, Maeder A, Maintain Your Brain Collaborative Team, et al. Maintain your brain: protocol of a 3-year randomized controlled trial of a personalized multi-modal digital health intervention to prevent cognitive decline among community dwelling 55 to 77 year olds. J Alzheimers Dis 2019 Nov 23;70(s1):S221-S237 [FREE Full text] [doi: 10.3233/JAD-180572] [Medline: 30475762]
17. Wesselman LM, Hooghiemstra AM, Schoonmade LJ, de Wit MC, van der Flier WM, Sikkes SA. Web-based multidomain lifestyle programs for brain health: comprehensive overview and meta-analysis. JMIR Ment Health 2019 Apr 09;6(4):e12104 [FREE Full text] [doi: 10.2196/12104] [Medline: 30964438]
18. Mehra S, van den Helder J, Visser B, Engelbert RH, Weijs PJ, Kröse BJ. Evaluation of a blended physical activity intervention for older adults: mixed methods study. J Med Internet Res 2020 Jul 23;22(7):e16380 [FREE Full text] [doi: 10.2196/16380] [Medline: 32459652]
19. Moorthy P, Weinert L, Schüttler C, Svensson L, Sedlmayr B, Müller J, et al. Attributes, methods, and frameworks used to evaluate wearables and their companion mHealth apps: scoping review. JMIR Mhealth Uhealth 2024 Apr 05;12:e52179 [FREE Full text] [doi: 10.2196/52179] [Medline: 38578671]
20. Arigo D, Jake-Schoffman DE, Wolin K, Beckjord E, Hekler EB, Pagoto SL. The history and future of digital health in the field of behavioral medicine. J Behav Med 2019 Feb;42(1):67-83 [FREE Full text] [doi: 10.1007/s10865-018-9966-z] [Medline: 30825090]
21. Fox S, Brown LJ, Antrobus S, Brough D, Drake RJ, Jury F, et al. Co-design of a smartphone app for people living with dementia by applying agile, iterative co-design principles: development and usability study. JMIR Mhealth Uhealth 2022 Jan 14;10(1):e24483 [FREE Full text] [doi: 10.2196/24483] [Medline: 35029539]
22. Harrington CN, Wilcox L, Connelly K, Rogers W, Sanford J. Designing health and fitness apps with older adults: examining the value of experience-based co-design. In: Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare. 2018 Presented at: PervasiveHealth '18; May 21-24, 2018; New York, NY p. 15-24 URL: https://dl.acm.org/doi/10.1145/3240925.3240929 [doi: 10.1145/3240925.3240929]
23. Lim MH, Gleeson JF, Rodebaugh TL, Eres R, Long KM, Casey K, et al. A pilot digital intervention targeting loneliness in young people with psychosis. Soc Psychiatry Psychiatr Epidemiol 2020 Jul 14;55(7):877-889. [doi: 10.1007/s00127-019-01681-2] [Medline: 30874828]
24. Audibert CE, de Moura Fereli Reis A, Zazula R, Machado RC, Guariente SM, Nunes SO. Development of digital intervention through a mobile phone application as an adjunctive treatment for bipolar disorder: MyBee project. Clin eHealth 2022 Dec;5:72-78. [doi: 10.1016/j.ceh.2022.09.001]
25. Peake E, Miller I, Flannery J, Chen L, Lake J, Padmanabhan A. Preliminary efficacy of a digital intervention for adolescent depression: randomized controlled trial. J Med Internet Res 2024 Feb 07;26:e48467 [FREE Full text] [doi: 10.2196/48467] [Medline: 38324367]
26. Ponzo S, Morelli D, Kawadler JM, Hemmings NR, Bird G, Plans D. Efficacy of the digital therapeutic mobile app BioBase to reduce stress and improve mental well-being among university students: randomized controlled trial. JMIR Mhealth Uhealth 2020 Apr 06;8(4):e17767 [FREE Full text] [doi: 10.2196/17767] [Medline: 31926063]
27. Gräfe V, Moritz S, Greiner W. Correction to: health economic evaluation of an internet intervention for depression (deprexis), a randomized controlled trial. Health Econ Rev 2020 Jul 30;10(1):24 [FREE Full text] [doi: 10.1186/s13561-020-00280-1] [Medline: 32734480]
28. Atik E, Stricker J, Schückes M, Pittig A. Efficacy of a brief blended cognitive behavioral therapy program for the treatment of depression and anxiety in university students: uncontrolled intervention study. JMIR Ment Health 2023 Aug 25;10:e44742 [FREE Full text] [doi: 10.2196/44742] [Medline: 37624631]
29. Maramis C, Moulos I, Ioakimidis I, Papapanagiotou V, Langlet B, Lekka I, et al. A smartphone application for semi-controlled collection of objective eating behavior data from multiple subjects. Comput Methods Programs Biomed 2020 Oct;194:105485. [doi: 10.1016/j.cmpb.2020.105485] [Medline: 32464588]
30. Pihlajamäki J, Männikkö R, Tilles-Tirkkonen T, Karhunen L, Kolehmainen M, Schwab U, StopDia study group. Digitally supported program for type 2 diabetes risk identification and risk reduction in real-world setting: protocol for the StopDia model and randomized controlled trial. BMC Public Health 2019 Mar 01;19(1):255 [FREE Full text] [doi: 10.1186/s12889-019-6574-y] [Medline: 30823909]
31. Rabbi M, Aung MS, Gay G, Reid MC, Choudhury T. Feasibility and acceptability of mobile phone-based auto-personalized physical activity recommendations for chronic pain self-management: pilot study on adults. J Med Internet Res 2018 Oct 26;20(10):e10147 [FREE Full text] [doi: 10.2196/10147] [Medline: 30368433]

XSL•FO
RenderX

32. Schuffelen J, Maurer LF, Lorenz N, Rötger A, Pietrowsky R, Gieselmann A. The clinical effects of digital cognitive behavioral therapy for insomnia in a heterogenous study sample: results from a randomized controlled trial. Sleep 2023 Nov 08;46(11):zsad184 [FREE Full text] [doi: 10.1093/sleep/zsad184] [Medline: 37428712]
33. Upton E, Doogan C, Fleming V, Leyton PQ, Barbera D, Zeidman P, et al. Efficacy of a gamified digital therapy for speech production in people with chronic aphasia (iTalkBetter): behavioural and imaging outcomes of a phase II item-randomised clinical trial. EClinicalMedicine 2024 Apr;70:102483 [FREE Full text] [doi: 10.1016/j.eclinm.2024.102483] [Medline: 38685927]
34. Harjumaa M, Absetz P, Ermes M, Mattila E, Männikkö R, Tilles-Tirkkonen T, et al. Internet-based lifestyle intervention to prevent type 2 diabetes through healthy habits: design and 6-month usage results of randomized controlled trial. JMIR Diabetes 2020 Aug 11;5(3):e15219 [FREE Full text] [doi: 10.2196/15219] [Medline: 32779571]
35. Chong SO, Pedron S, Abdelmalak N, Laxy M, Stephan A. An umbrella review of effectiveness and efficacy trials for app-based health interventions. NPJ Digit Med 2023 Dec 16;6(1):233 [FREE Full text] [doi: 10.1038/s41746-023-00981-x] [Medline: 38104213]
36. Barlow C. Oncology research: clinical trial management systems, electronic medical record, and artificial intelligence. Semin Oncol Nurs 2020 Apr;36(2):151005. [doi: 10.1016/j.soncn.2020.151005] [Medline: 32253050]
37. Nourani A, Ayatollahi H, Dodaran MS. A review of clinical data management systems used in clinical trials. Rev Recent Clin Trials 2019;14(1):10-23. [doi: 10.2174/1574887113666180924165230] [Medline: 30251611]
38. Clinical Trial Management System (CTMS). Definitive Healthcare Limited. URL: https://www.definitivehc.com/resources/glossary/clinical-trial-management-system [accessed 2025-06-11]
39. Drew CJ, Poile V, Trubey R, Watson G, Kelson M, Townson J, et al. Integrating technology into complex intervention trial processes: a case study. Trials 2016 Nov 17;17(1):551 [FREE Full text] [doi: 10.1186/s13063-016-1674-9] [Medline: 27855710]
40. Nourani A, Ayatollahi H, Dodaran MS. Clinical trial data management software: a review of the technical features. Rev Recent Clin Trials 2019;14(3):160-172. [doi: 10.2174/1574887114666190207151500] [Medline: 30734683]
41. Ghazali M, Ariffin NA, Omar R. User centered design practices in healthcare: a systematic review. In: Proceedings of the 3rd International Conference on User Science and Engineering. 2014 Presented at: i-USEr '15; September 2-5, 2014; Shah Alam, Malaysia p. 91-96 URL: https://ieeexplore.ieee.org/document/7002683 [doi: 10.1109/iuser.2014.7002683]
42. ISO 9241-210:2019: ergonomics of human-system interaction: part 210: human-centred design for interactive systems. International Organization for Standardization. URL: https://www.iso.org/standard/77520.html [accessed 2025-05-29]
43. Matthews T, Judge T, Whittaker S. How do designers and user experience professionals actually perceive and use personas? In: Proceedings of the 2012 SIGCHI Conference on Human Factors in Computing Systems. 2012 Presented at: CHI '12; May 5-10, 2012; Austin, TX p. 1219-1208 URL: https://dl.acm.org/doi/10.1145/2207676.2208573 [doi: 10.1145/2207676.2208573]
44. Bruinsma J, Heger I, Loukas VS, Kassiotis T, Karanasiou G, Fotiadis DI, et al. Public perspectives on lifestyle-related behavior change for dementia risk reduction: an exploratory qualitative study in the Netherlands. J Alzheimers Dis 2023;95(4):1635-1642 [FREE Full text] [doi: 10.3233/JAD-230217] [Medline: 37718799]
45. Bruinsma J, Loukas VS, Kassiotis T, Heger I, Rosenberg A, Visser LN, et al. Socio-cognitive determinants of lifestyle behavior in the context of dementia risk reduction: a population-based study in the Netherlands. J Alzheimers Dis 2024;99(3):941-952 [FREE Full text] [doi: 10.3233/JAD-231369] [Medline: 38759007]
46. Peters GJ, Crutzen R. Pragmatic nihilism: how a theory of nothing can help health psychology progress. Health Psychol Rev 2017 Jun 10;11(2):103-121 [FREE Full text] [doi: 10.1080/17437199.2017.1284015] [Medline: 28110627]
47. Bartholomew Eldredge LK, Markham CM, Ruiter RA, Fernández ME, Kok G, Parcel GS. Planning Health Promotion Programs: An Intervention Mapping Approach. 4th edition. New York, NY: Jossey-Bass Inc; 2016.
48. Kok G, Gottlieb NH, Peters GJ, Mullen PD, Parcel GS, Ruiter RA, et al. A taxonomy of behaviour change methods: an intervention mapping approach. Health Psychol Rev 2016 Sep;10(3):297-312 [FREE Full text] [doi: 10.1080/17437199.2015.1077155] [Medline: 26262912]
49. Interactive presentation software. Mentimeter. URL: https://www.mentimeter.com/ [accessed 2025-05-29]
50. Padlet. URL: https://de.padlet.com/ [accessed 2025-05-29]
51. Li SH, Achilles MR, Spanos S, Habak S, Werner-Seidler A, O'Dea B. A cognitive behavioural therapy smartphone app for adolescent depression and anxiety: co-design of ClearlyMe. Cogn Behav Ther 2022 Mar 16;15:e13. [doi: 10.1017/S1754470X22000095]
52. The collaborative interface design tool. Figma GmbH. URL: https://www.figma.com/ [accessed 2025-05-29]
53. Jaspers MW, Steen T, van den Bos C, Geenen M. The think aloud method: a guide to user interface design. Int J Med Inform 2004 Nov;73(11-12):781-795. [doi: 10.1016/j.ijmedinf.2004.08.003] [Medline: 15491929]
54. Work better together with Mural's visual work platform. Mural. URL: https://mural.co/ [accessed 2023-11-06]
55. Christen OM, Mösching Y, Müller P, Denecke K, Nüssli S. Dashboard visualization of information for emergency medical services. Stud Health Technol Inform 2020 Nov 23;275:27-31. [doi: 10.3233/SHTI200688] [Medline: 33227734]

XSL•FO
RenderX

56. Bach K, Marling C, Mork PJ, Aamodt A, Mair FS, Nicholl BI. Design of a clinician dashboard to facilitate co-decision making in the management of non-specific low back pain. J Intell Inf Syst 2018 Dec 3;52(2):269-284. [doi: 10.1007/s10844-018-0539-y]
57. Hensley NB, Grant MC, Cho BC, Suffredini G, Abernathy JA. How do we use dashboards to enhance quality in cardiac anesthesia? J Cardiothorac Vasc Anesth 2021 Oct;35(10):2969-2976. [doi: 10.1053/j.jvca.2021.04.034] [Medline: 34059439]
58. R Core Team. R: a language and environment for statistical computing. The R Foundation. URL: https://www.R-project.org/ [accessed 2025-05-29]
59. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry. Boca Raton, FL: CRC Press; 1996:189-194.
60. Holden RJ. A simplified System Usability Scale (SUS) for cognitively impaired and older adults. Proc Int Symp Hum Factors Ergon Healthc 2020 Sep 16;9(1):180-182. [doi: 10.1177/2327857920091021]
61. Taherdoost H. Designing a questionnaire for a research paper: a comprehensive guide to design and develop an effective questionnaire. Asian J Manag Sci 2022 Apr 25;11(1):8-16. [doi: 10.51983/ajms-2022.11.1.3087]
62. Sauro J, Lewis JR. Quantifying the User Experience: Practical Statistics for User Research. Waltham, MA: Elsevier Inc; 2012.
63. Welcome to FHIR®. HL7 FHIR. URL: https://www.hl7.org/fhir/? [accessed 2024-01-03]
64. Bjerke MB, Renger R. Being smart about writing SMART objectives. Eval Program Plann 2017 Apr;61:125-127. [doi: 10.1016/j.evalprogplan.2016.12.009] [Medline: 28056403]
65. Fogg BJ, Hreha J. Behavior wizard: a method for matching target behaviors with solutions. In: Proceedings of the 5th International Conference on Persuasive Technology. 2010 Presented at: PERSUASIVE '10; June 7-10, 2010; Copenhagen, Denmark p. 117-131 URL: https://link.springer.com/chapter/10.1007/978-3-642-13226-1_13 [doi: 10.1007/978-3-642-13226-1_13]
66. Wood W, Neal DT. Healthy through habit: interventions for initiating and maintaining health behavior change. Behav Sci Policy 2016 Apr 01;2(1):71-83. [doi: 10.1177/237946151600200109]
67. Stop DC. Pienet teot -sovellus. University of Eastern Finland. URL: https://sites.uef.fi/stopdia/t2d-data-pienet-teot/ [accessed 2025-06-11]
68. Kivipelto M, Solomon A, Ahtiluoto S, Ngandu T, Lehtisalo J, Antikainen R, et al. The Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER): study design and progress. Alzheimers Dement 2013 Nov;9(6):657-665. [doi: 10.1016/j.jalz.2012.09.012] [Medline: 23332672]
69. Hilberger H, Bödenler M, Hanke S. Tracking usage patterns of a mobile application for dementia risk assessment in elderly individuals. Stud Health Technol Inform 2024 Aug 22;316:478-482. [doi: 10.3233/SHTI240452] [Medline: 39176781]
70. Lewis JR. The system usability scale: past, present, and future. Int J Hum Comput Interact 2018 Mar 30;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
71. Hughes CP, Berg L, Danziger W, Coben LA, Martin RL. A new clinical scale for the staging of dementia. Br J Psychiatry 1982 Jun 29;140(6):566-572. [doi: 10.1192/bjp.140.6.566] [Medline: 7104545]
72. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
73. Mclellan S, Muddimer A, Peres SC. The effect of experience on system usability scale ratings. J Usability Stud 2012;7(2):56-67 [FREE Full text]
74. Győrffy Z, Boros J, Döbrössy B, Girasek E. Older adults in the digital health era: insights on the digital health related knowledge, habits and attitudes of the 65 year and older population. BMC Geriatr 2023 Nov 27;23(1):779 [FREE Full text] [doi: 10.1186/s12877-023-04437-5] [Medline: 38012565]
75. Ruissen MM, Torres-Peña JD, Uitbeijerse BS, Arenas de Larriva AP, Huisman SD, Namli T, POWER2DM Consortium. Clinical impact of an integrated e-health system for diabetes self-management support and shared decision making (POWER2DM): a randomised controlled trial. Diabetologia 2023 Dec 29;66(12):2213-2225 [FREE Full text] [doi: 10.1007/s00125-023-06006-2] [Medline: 37775611]
76. Wegener EK, Bergschöld JM, Whitmore C, Winters M, Kayser L. Involving older people with frailty or impairment in the design process of digital health technologies to enable aging in place: scoping review. JMIR Hum Factors 2023 Jan 27;10:e37785 [FREE Full text] [doi: 10.2196/37785] [Medline: 36705959]
77. LETHE (λήθη) – a personalized prediction and intervention model for early detection and reduction of risk factors causing dementia, based on AI and distributed Machine Learning. European Commission. URL: https://cordis.europa.eu/project/id/101017405 [accessed 2025-05-29]

Abbreviations

AI: artificial intelligence
CI/CD: Continuous Integration/Continuous Delivery or Deployment
CRF: case report form
CTMS: clinical trial management system

XSL•FO
RenderX

eCRF: electronic case report form
FHIR: Fast Healthcare Interoperability Resources
FINGER: Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability
ICT: information and communications technology
mHealth: mobile health
POWER2DM: predictive model-based decision support for diabetes patient empowerment
PRODEMOS: prevention of dementia using mobile phone applications
RCT: randomized controlled trial
SUS: System Usability Scale
UCD: user-centered design

Edited by P Jansons; submitted 19.09.24; peer-reviewed by K Kawaguchi, R Pryss; comments to author 30.03.25; revised version received 06.05.25; accepted 26.05.25; published 02.07.25.

Please cite as:
Hilberger H, Buchgraber-Schnalzer B, Huber S, Weitlaner T, Bödenler M, Abaci A, Bruinsma J, Diaz A, Guazzarini AG, Lehtisalo J, Lee S, Loukas V, Mangialasche F, Mecocci P, Ngandu T, Rosenberg A, Stögmann E, Valkonen K, Uhlik E, Untersteiner H, Kneß L, Ahammer H, Hanke S
Design of a Mobile App and a Clinical Trial Management System for Cognitive Health and Dementia Risk Reduction: User-Centered Design Approach
JMIR Aging 2025;8:e66660
URL: https://aging.jmir.org/2025/1/e66660
doi:10.2196/66660
PMID:

©Hannes Hilberger, Bianca Buchgraber-Schnalzer, Simone Huber, Theresa Weitlaner, Markus Bödenler, Alara Abaci, Jeroen Bruinsma, Ana Diaz, Anna Giulia Guazzarini, Jenni Lehtisalo, Seungjune Lee, Vasileios Loukas, Francesca Mangialasche, Patrizia Mecocci, Tiia Ngandu, Anna Rosenberg, Elisabeth Stögmann, Konsta Valkonen, Elena Uhlik, Helena Untersteiner, Laura Kneß, Helmut Ahammer, Sten Hanke. Originally published in JMIR Aging (https://aging.jmir.org), 02.07.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Digital Serious Game (Anticip'action) to Support Advance Care Planning Discussions in the General Population: Usability Study

Dafne Campioni[1*], MA; Frederic Ehrler[2*], PhD; Antoine Berger[2], MA; Christine Clavien[1], PhD

[1]Institute for Ethics, History, and the Humanities, Faculty of Medicine, University of Geneva, Geneva, Switzerland

[2]Department of Medical Information Sciences, University Hospital of Geneva, Geneva, Switzerland

[*]these authors contributed equally

Corresponding Author:
Christine Clavien, PhD
Institute for Ethics, History, and the Humanities
Faculty of Medicine
University of Geneva
CMU
1 rue Michel Servet
Geneva, 1211
Switzerland
Phone: 41 22 379 4609
Fax: 41 22 379 4619
Email: christine.clavien@unige.ch

Abstract

Background: In the context of an aging population and increasingly medicalized end-of-life practices, it is crucial to promote early discussions to help patients express their view on what is essential in their life as well as articulate their preferences regarding future medical treatments and end-of-life issues. An interprofessional research team at Geneva University and the Geneva University Hospitals has developed Anticip'action, a card game designed to help initiate and conduct advance care planning and end-of-life discussions. It is available for free in paper format in diverse languages and in a digital version in French.

Objective: This study aims to assess the ergonomic quality of the digital version of the game with primary users.

Methods: Overall, 10 users (women: n=5, 50%; men: n=5, 50%; mean age 41, SD 13.4 years; range 25-65 years; education: upper level; comfortable with smartphones) completed an online usability test. The test began with a rapid desirability test to capture initial impressions of the game's main screen without knowing what it is about. This was followed by a think-aloud procedure, including 26 tasks to perform all steps of the game. Posttest questionnaires were administered to collect participants' subjective perceptions of the usability (System Usability Scale), attractiveness (AttrakDiff), and relevance as well as overall endorsement of the game (Mobile Application Rating Scale). Open-ended questions were used to further explore usability issues. Usability problems were categorized and evaluated using standard evaluation grids. Content readability was assessed with Scolarius.

Results: The rapid desirability test revealed an overall good or average impression. In 83.2% (208/250) of the cases, participants successfully completed the think-aloud tasks without assistance. Some of the tasks (4/25, 16%) revealed multiple usability issues requiring assistance. Analysis of the 23 failures and difficulties encountered revealed that 3 (13%) issues were due to suboptimal wording of the task instructions and that there were 9 (39%) major usability problems. All could be addressed through minor modifications. The Scolarius test indicated that the card titles were understandable at an elementary school level, while the explanations on the back of the cards required a high school reading level. Participants rated Anticip'action as good or excellent in usability (79 out of 100 on the System Usability Scale), attractiveness (1.57 on the −3 to +3 AttrakDiff scale), and relevance (4.1 out of 5 on section F of the Mobile Application Rating Scale). Participants provided overall positive qualitative feedback.

Conclusions: The usability testing of the digital French version of Anticip'action produced positive results, with some areas for improvement identified. It can be recommended as a valuable resource for patients, families, and caregivers to prompt reflection, raise awareness, and support advance care planning conversations. Further tests should be conducted on wider population groups, including older patients and individuals less comfortable with digital solutions.

(JMIR Aging 2025;8:e73378) doi:10.2196/73378

XSL•FO
RenderX

KEYWORDS

advance care planning; mobile health; mHealth; mobile apps; mobile phone; palliative care; usability; serious game; conversation tool; end of life

Introduction

Background

In the context of an aging population and increasingly medicalized end-of-life practices, it is crucial to promote early discussions to help patients express their view on what is essential in their life as well as articulate their preferences regarding future medical treatments and end-of-life issues [1,2]. Advance care planning and end-of-life discussions may include clarifying one's own values and priorities in life; preferences regarding situations or medical treatments that are expected or not desired; and other important matters that should be addressed in case of loss of decision-making capacity or death, such as the management of financial matters, funeral planning, organ donation, the designation of a surrogate medical decision maker, and so on. These discussions are mostly relevant for patients facing a life-threatening illness or for older people with increased fragility. They help increase patient empowerment and decrease distress among loved ones and health professionals when important decisions have to be made [3]. However, these anticipatory discussions are often delayed or avoided due to a multitude of barriers such as health care professionals' lack of trust in their communication abilities, time constraints, emotional burden, and fear of generating misunderstandings or conflicts [4-7]. To overcome these obstacles, it is crucial to develop strategies and communication tools that facilitate the initiation and conduction of end-of-life discussions.

Studies indicate that free online digital tools (websites, apps, online games, etc) could enhance interest, communication, and decision-making related to advance care planning [8-12]. These communication tools may serve as icebreakers, provide a supportive structure for conducting a discussion, and save clinicians' time. Indeed, digital tools are easy to share with patients and can provide relevant information and input to conduct part of the reflection in a private and familial context ahead of the medical consultation. However, with some exceptions [13], existing solutions are mostly in English and not necessarily adapted to the European cultural context.

An interprofessional research team at Geneva University and the Geneva University Hospitals (HUG), composed of ethicists, physicians, nurses, trained patients as partners, and IT professionals, has developed information and communication tools to support the culture of advance care planning. In particular, the digital tool Accordons-nous ("Let's agree") is available for free as a module integrated into the larger Concerto application available on the HUG website [14]. The Concerto app can also be downloaded from the Google Play Store and Apple App Store. Accordons-nous [15] is currently only available in French. It includes information on advance care planning, discussion prompts, an advance care directive form, and a digital version of the card game Anticip'action.

Anticip'action, the Card Game

Anticip'action is a card game designed to help users clarify their values and priorities in life, discuss them with their loved ones and health care professionals, and plan concrete actions regarding care and end-of-life decisions. It is inspired by the game Go Wish [16] but contains its own set of cards adapted to the Swiss cultural context and additional rules and content. It can be used with caregivers during structured advance care planning discussions, or it can be proposed as an icebreaker to be played independently at home with loved ones. A printable version of Anticip'action is available in different languages (English, French, German, and Italian) for people who are not at ease with digital tools. The digital version is currently only available in French in Accordons-nous (Textbox 1; Figures 1-6). In this paper, we present the results of the usability test of the digital version of the game.

Textbox 1. Short description of the game rules of Anticip'action.

The game is divided into three phases

1. "Prioritize"—in this phase, the task is to sort 32 cards into 4 categories ("Very important," "Important," "Not important," and "Unsure"). Examples of card statements are "Looking after my appearance" and "Expressing what I couldn't stand." Each card includes explanatory text on the back.
2. "Specify"—in this phase, the task is to add personal details to the cards that have been marked as "Very important" and decide whether the topic has already been addressed or still requires action.
3. "Act"—in this phase, the task is to write down actions that could be taken to address the important issues highlighted in the previous phase. To assist in this task, each card suggests possible actions (eg, scheduling an appointment with a health care provider or arranging a meeting with a loved one). The output of the game is a PDF summary of cards ranking, specifications provided during the game, and to-do-list of actions to be done.

XSL•FO
RenderX

Figure 1. Welcome page of Anticip'action included in the Accordons-nous module. HUG: Geneva University Hospitals.

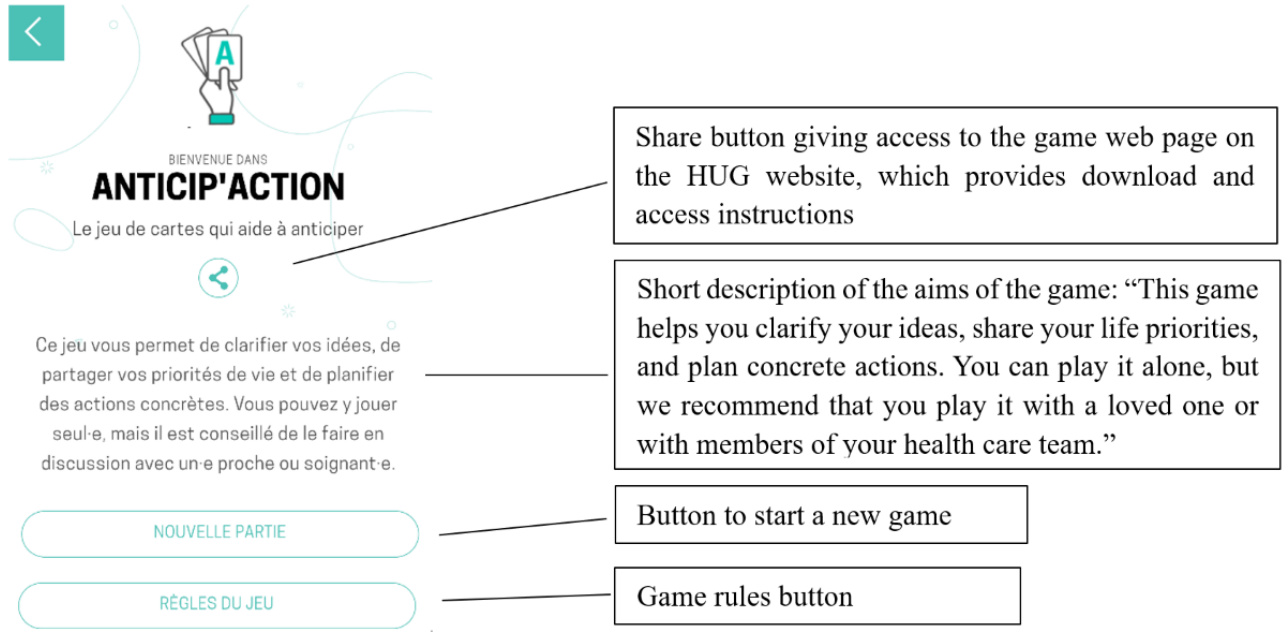
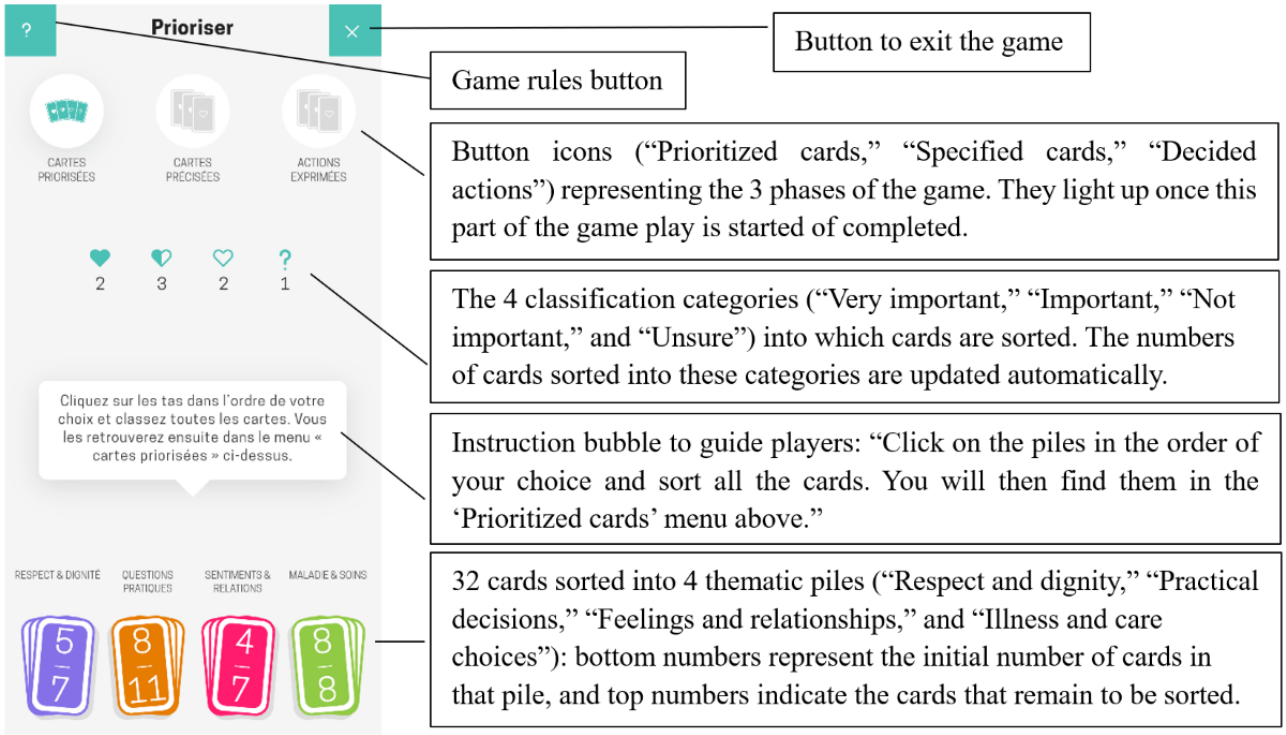


Figure 2. Main screen of the game.



XSL•FO
RenderX

Figure 3. Game phase 1—"Prioritize": screen for sorting cards allocated to the theme "Practical decisions.".

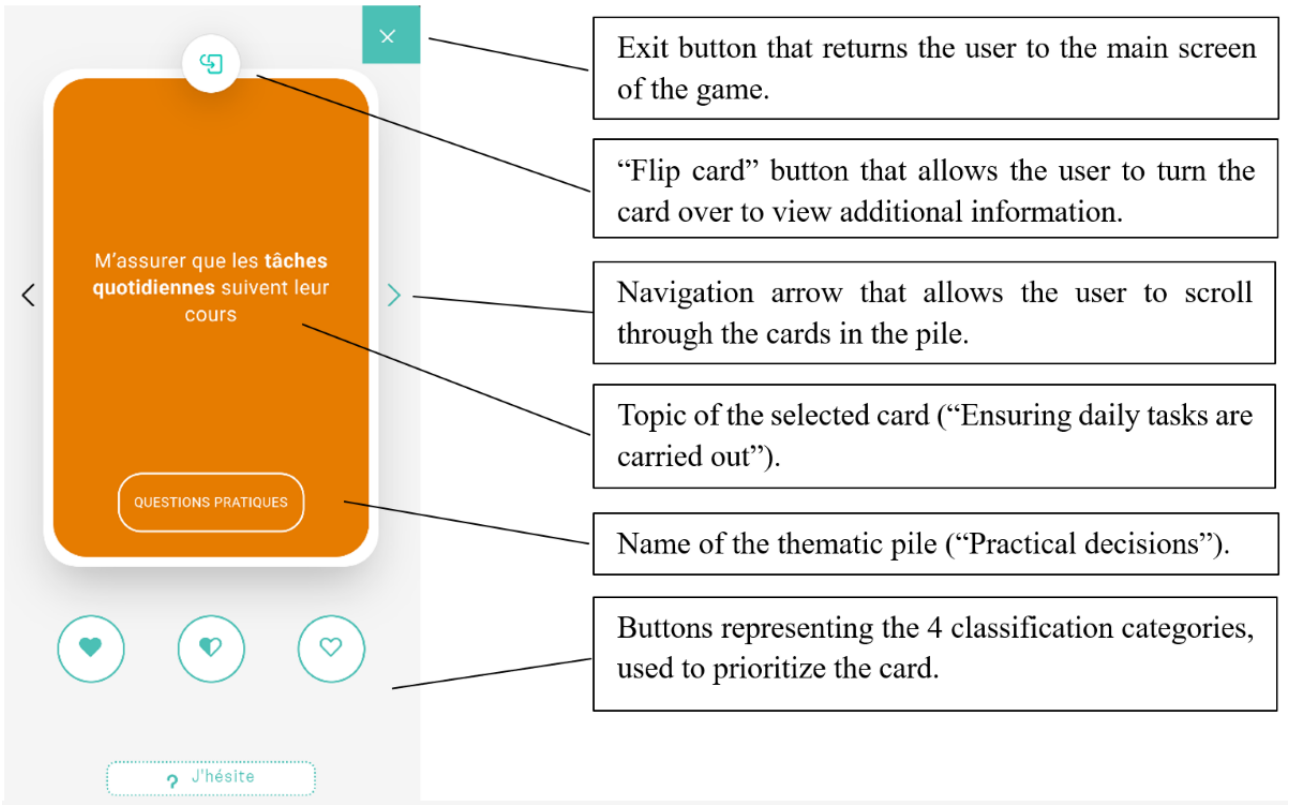
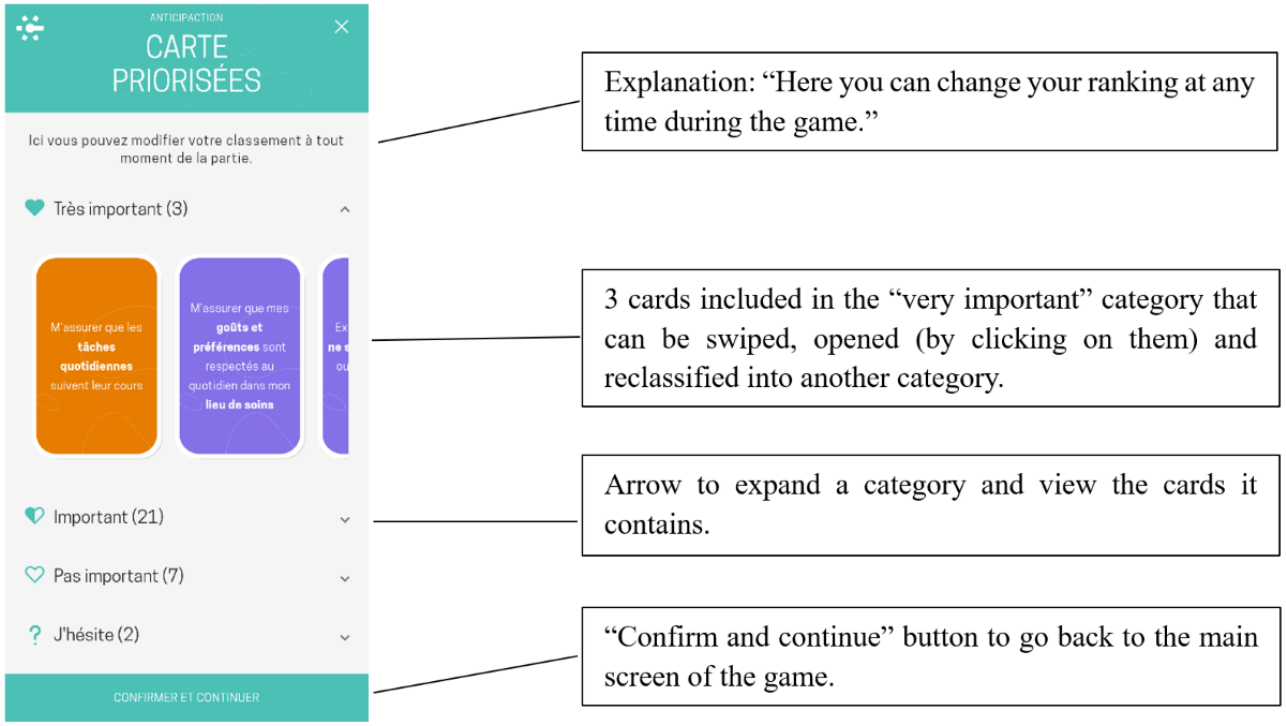


Figure 4. "Prioritized cards" screen, which allows users to re-sort cards into different categories at any time during the game. This screen can be accessed from the main screen by clicking on the "Prioritized cards" button icon.



XSL•FO
RenderX

Figure 5. Game phase 2: “Specify”—screen that allows users to specify reasons for having marked a card as “Very important.”.

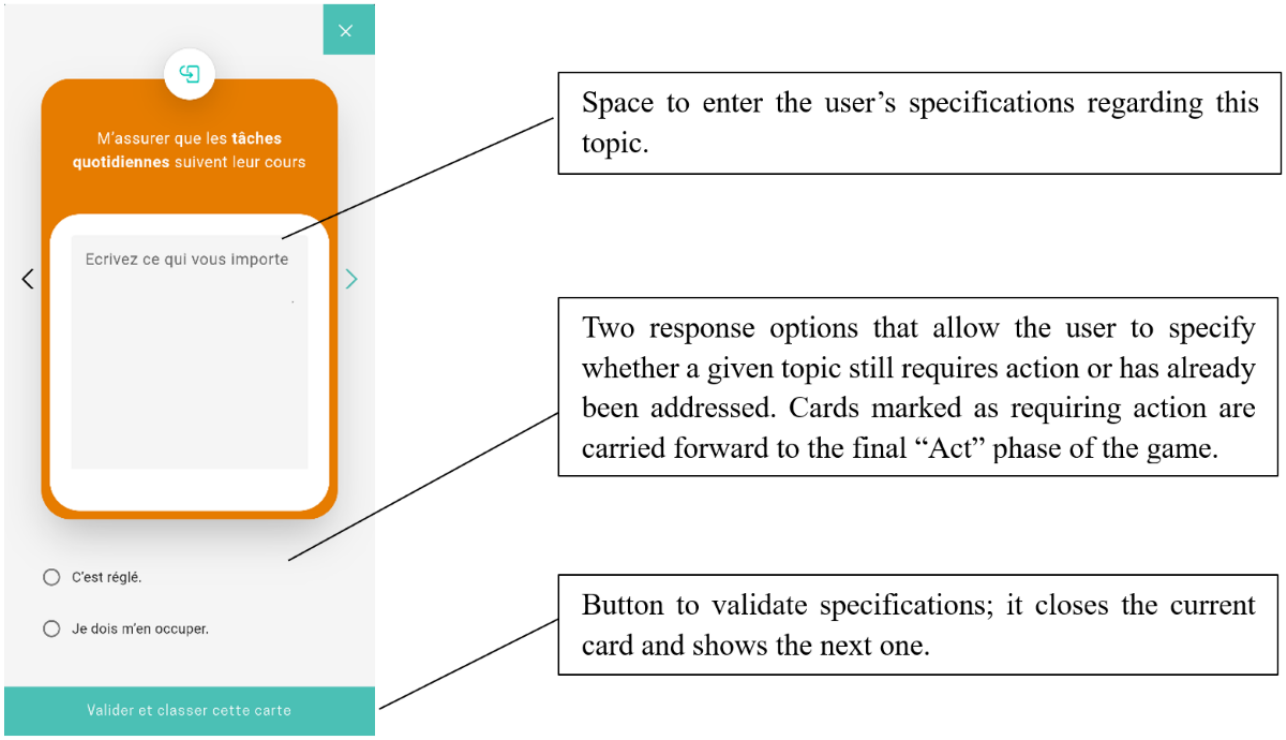
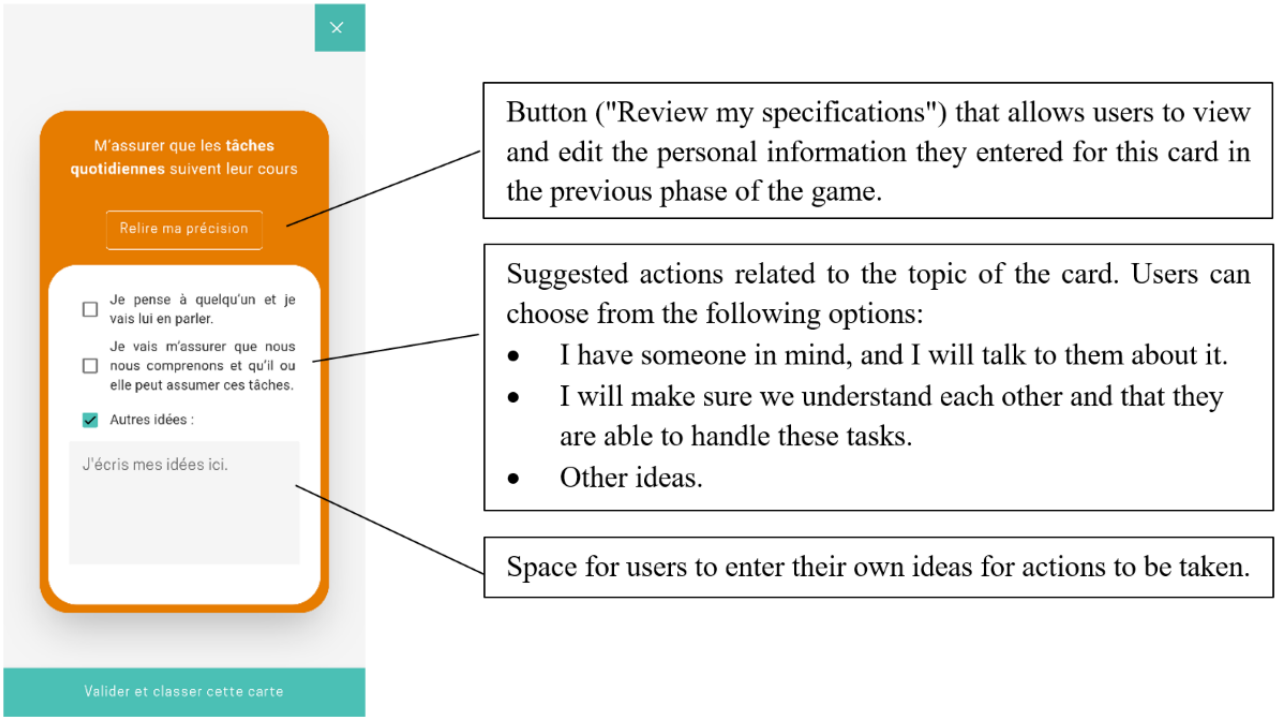


Figure 6. Game phase 3: “Act”—screen that allows users to choose actions to complete in relation to the topic of the card.



Methods

Recruitment

The aim was to collect data from 10 participants, as Nielsen and Landauer [17] have demonstrated that a usability test conducted with 10 users can identify >98% of usability issues. The inclusion criteria were as follows: being an adult, having oral and written proficiency in French, and feeling comfortable using a smartphone. A single exclusion criterion was applied: participants should not be familiar with or have previously navigated the Accordons-nous module in the Concerto app. The exclusion criterion was enforced to prevent testers from being biased by prior knowledge that could facilitate ease of use. Overall, the aim was to ensure diversity in gender and age and to include individuals with at least minimal comfort using digital tools. Participants were recruited through the patients-as-partners platform at the HUG as well as via snowball sampling and networking. Participation was voluntary, and participants

XSL•FO
RenderX

received an information sheet and consent form in advance. The study was preregistered on the OSF platform [18].

Ethical Considerations

The research protocol was submitted to the Geneva Commission Cantonale d'Ethique de la Recherche (Req-ID 2024-00994), which determined that the project did not fall within the scope of the Human Research Ordinance and therefore waived the requirement for ethics approval (decision date: September 17, 2024). This study was conducted in accordance with the Declaration of Helsinki. Willing participants were informed of the study aims and procedures and signed an informed consent form. During the data collection and analysis phase, only 3 authors (DC, FE, and CC) had access to the recorded interviews. The recordings were destroyed after analysis. No identifying information was retained in the remaining study files.

Study Intervention

The entire evaluation was conducted online (via email and videoconference using Zoom [Zoom Video Communications, Inc]). Participants received in advance via email the information sheet and consent form along with instructions on how to join the videoconference with the experimenter and how to install the Concerto app. They were asked not to use the Concerto app before the test and were not informed about the name of the game or where to find it in the app. On the day of the test, the experimenter answered any questions, made sure that informed consent had been signed and returned, and assisted participants in connecting their smartphone to the Zoom videoconference to observe and record their navigation behavior.

The test began with a few demographic and technical questions (age, gender, level of education, type of smartphone used, and smartphone use habits), after which participants were shown a screenshot of the game to gather their first impressions. They then completed a series of think-aloud tasks [19] during which they used the main functions of the game. The 26 tasks were designed to evaluate the usability of the game and to ensure that an average user could navigate effectively through it. Participants were asked to imaging being an individual aged 50 years concerned about illness and advance care directives, who had been advised to use the game Anticip'action. The experimenter then presented the scripted task instructions one by one (Table 1). While completing the tasks, participants were encouraged to verbalize their thoughts, actions, and decisions.

XSL•FO
RenderX

Table 1. List and description of the tasks completed during the think-aloud procedure.

Tasks	Expected achievement	Exact wording of the instruction	When the task was considered complete
1. Find and open the game.	Find the game nested in the Accordons-nous module in the Concerto app.	Your friend told you about the game. He told you that the game is in Concerto. You are now in Concerto. More precisely, the game is in the Accordons-nous module. Your first task is to find the game and open it. Let's go!	The participant opens the Concerto app, scrolls to and clicks on the Accordons-nous module, clicks on the "Talk about it" menu at the bottom of the screen to access the game entry point, and then clicks on the "Play" button.
2. Read rules and start.	Understand the logic of the welcome page, which includes 2 buttons: one for accessing the rules of the game and one for starting a new game (Figure 1).	First, read the rules, and start the game.	The participant clicks on the "Game rules" button, reads the rules, closes the rules page by clicking on the X icon, clicks on the "New game" button, reads the pop-up instructions when they appear, and then closes the pop-up.
3. Select "Feelings and relationships."	Understand that there are 4 thematic card piles that can be opened (Figure 2).	Select the "Feelings and relationships" pile, and open it.	The participant clicks on the third (pink) icon at the bottom of the screen, which represents the "Feelings and relationships" pile.
4. Flip the first card.	Understand that the card can be flipped to obtain more information on the topic (Figure 3).	Flip the first card.	The participant clicks on the "flip card" button at the top of the screen.
5. Scroll through cards.	Understand that the content of the opened pile can be explored by scrolling (Figure 3).	Scroll once to the left and once to the right to see the other cards in this pile.	The participant uses the navigation arrows to scroll through the cards in the pile.
6. Classify as "Unsure."	Understand the classification system, including the "Unsure" category (Figure 3).	Classify the first card as "Unsure."	The participant clicks on the rectangular "Unsure" button.
7. Classify as "Very important."	Understand that the first filled heart icon represents the "Very important" category (Figure 3).	Classify the next 3 cards as "Very important."	The participant clicks on the filled heart icon.
8. Exit the pile.	Understand how to return to the main screen of the game when a pile is opened (Figure 3).	Exit this pile.	The participant clicks on the X icon on the top right of the screen.
9. Open the "Practical questions" pile.	Identify and open another thematic card pile (Figure 2).	Open the "Practical questions" pile.	The participant clicks on the second (orange) icon at the bottom of the screen.
10. Classify 9 cards as "Very important."	Learn the rule limiting the number of cards classified as "Very important" to a maximum of 10 (Figure 3)	Classify the first 9 cards as "Very important."	The participant clicks on the filled heart icon until a pop-up message appears.
11. Read and close pop-up.	Understand the content of an instruction pop-up and close it.	[Pop-up appears] Read this message, and close it.	The participant reads the pop-up message and clicks on either the X icon or the "Understood" button.
12. Find and scroll through surplus cards.	Discover where to find already classified cards (Figures 2 and 4).	Find the surplus cards, and scroll through them.	The participant clicks on the X icon on the top right of the screen to exit the pile, clicks on the "Prioritized cards" button icon on the main screen (top left) to view and update already classified cards, clicks on "Very important," and scrolls through the cards.
13. Classify surplus cards as "Important."	Understand how to reclassify cards in another category (Figure 4).	Classify the surplus cards as "Important" and then confirm.	The participant clicks on a card classified as "Very important" and clicks on the half-filled heart icon, repeating the process until task completion.
14. Classify all cards.	Understand that all cards must be classified in the first "Prioritize" phase of the game (Figure 2).	Classify the following cards as "Not important."	The participant clicks on card piles and clicks on the empty heart icon below each card, repeating the process until no more cards are available.

XSL•FO
RenderX

Tasks	Expected achievement	Exact wording of the instruction	When the task was considered complete
15. Check card count in each category.	Understand the logic of the 4 classification categories and their associated numbers (located on the main screen; Figure 2).	Do you know how many cards are present in each classification category ("Very important," "Important," "Not important," "Unsure")?	The participant reads and interprets the classification numbers located under the category's icons.
16. Go to next phase of the game.	Understand how to go to the second phase of the game.	You can proceed to the next step.	The participant clicks on the "Next step" button (it appears at the bottom of the main screen when all cards have been classified).
17. Read instructions and continue.	Discover the game rules for the next phase ("Specify") of the game.	[Pop-up appears] Read the instructions, and continue the game.	The participant reads the instructions and closes the pop-up by clicking on either the X icon or the "Understood" button.
18. Add text to a card.	Understand how to add text to a card (Figure 5).	Write comments on the first card by adding text in the designated space (eg, a letter).	The participant clicks on the card pile (which opens the first "Very important" card) and enters text in the designated space on the card.
19. Validate and define whether to act.	Understand that the specifications provided on a card can only be validated after it is additionally specified whether the topic still requires action or has already been addressed (Figure 5).	Validate the card to move on to the next one, knowing that action is still required.	The participant clicks on the second radio button ("I have to take care of this") and then clicks on the "Validate and classify this card" button.
20. Locate the rules.	Identify how to access the game rules at any time during the game (Figure 2).	Imagine that you suddenly have doubts about some game rules and would like to consult them. Where do you find the rules of the game?	The participant clicks on the X icon to exit the pile and, on the main screen, clicks on the "?" icon (top left); alternatively, the participant quits the game by clicking on the X icon on the main screen and then clicking on the "Game rules" button on the welcome page.
21. Validate as "Resolved."	Understand that for all cards in the "Specify" phase, a decision has to be made about whether or not it calls for an action (Figure 5).	Validate the following cards (without writing text on them) as "Resolved" and continue the game.	The participant clicks on the pile of cards not yet specified (located at the bottom of the main screen); for each card, the participant clicks on the first radio button ("This has been taken care of") and then clicks on the button "Validate and classify this card."
22. Review and check all actions.	Understand how to handle the cards remaining in the final phase ("Act") of the game: open each card, review the specifications, and choose related actions (Figure 6).	Review your specifications, and check all the actions.	The participant clicks on the pile of cards requiring action (located at the bottom of the main screen), clicks on the "Review my specifications" button, clicks on the "flip card" button again, and clicks on all actions.
23. Add more text to a card.	Understand how to enter one's own ideas about actions to be taken (Figure 6).	Add text in the designated space (eg, a letter).	The participant clicks in the "Other ideas" textbox and enters text.
24. Validate the card.	Understand how to validate action choices related to a card (Figure 6).	Validate the card.	The participant clicks on the "Validate and classify this card" button.
25. Export results and return to the game.	Understand how to export a summary of the game.	Export your results, and return to the game.	The participant clicks on the "Export game summary" button (it appears when all actions have been defined), reads the data confidentiality pop-up, and clicks on the "Understood" button.
26. Erase data and exit the game.	Understand that a decision can be made to erase or retain the data when exiting the game.	Exit the game, and delete the data.	The participant clicks on the "Quit" button and, when the pop-up appears, clicks on the "Quit and delete my data" gray band.

Finally, posttest questionnaires were administered to gather participants' subjective evaluation of the usability, attractiveness, and relevance of the game as well as their overall endorsement. The questionnaires included open-ended questions to further explore usability issues (Multimedia Appendix 1). While usability was the primary focus of the study, these additional measures helped provide a broader understanding of users' perceptions and expectations.

XSL•FO
RenderX

Measures Used and Data Analysis

First Impressions

The 5-second test [20] was used to gather participants' spontaneous first impressions of the game. In this test, the main screen of the game (Figure 2) was shown for exactly 5 seconds before being replaced by a black screen. Participants were then asked four questions: (1) "What do you remember about the visual elements of the interface?" (2) "What do you think are the objectives of the game?" (3) "What is your overall impression of the game (on a 5-point scale ranging from *very bad* to *very good*)?" (4) "How do you evaluate the aesthetics of the game (on a 5-point scale ranging from *very unsightly* to *very attractive*)?" The 2 open-ended questions were presented first because this has been shown to better support memory recall [21].

Think-Aloud Procedure

Task Success

Performance was assessed based on whether participants succeeded or failed in completing each task, with or without assistance from the experimenter. The following scoring system was applied: 0=the participant failed, and the experimenter ultimately provided the answer; 1=the participant succeeded but only with the experimenter's help; 2= the participant succeeded after the experimenter repeated the task instructions (in cases where the participant had forgotten the task; this was not considered assistance); 3=the participant succeeded independently but explored the interface before arriving at the correct result (correct result and path but after some exploration); 4=the participant succeeded without help but did not use the shortest path (correct result via an alternative route); and 5=the participant succeeded without assistance and easily found the shortest path.

Clicks to Complete Tasks

The number of clicks required to complete each task was recorded. As shown in Table 1, some tasks required more clicks than others, whereas for certain tasks, multiple response paths were possible.

Time Spent on Tasks

The time (in seconds) that participants took to complete each task was recorded. Timing began when the experimenter finished giving the task instruction for the first time and ended when the participant successfully completed the task. Any time spent digressing during the task (eg, making critical comments about the game's design or sharing a personal memory) was recorded as "digression time" and subtracted from the total time.

Analysis of Errors and Problems Encountered

The errors and problems encountered by participants while completing the 26 tasks were recorded and analyzed. Each issue was documented in a Microsoft Excel file; categorized according to the heuristics (such as information density, consistency, and code meaning) of Bastien and Scapin [22]; and assigned a severity score using a 0 to 4 scale based on the recommendations of Nielsen and Landauer [17]: 0=disagreement on whether it constitutes a usability problem; 1=cosmetic issue—need not be fixed unless extra time is available; 2=minor usability issue—fixing this should be given low priority; 3=major usability issue—important to fix and should be given high priority; and 4=usability catastrophe—imperative to fix this before the product can be launched.

Data Analysis

To ensure reliability, all authors independently coded 2 recordings. The results were then compared and discussed to resolve discrepancies or doubts about specific interpretations. The first author coded the remaining data independently.

Questionnaires

Perceived Usability

The French version of the standard System Usability Scale (SUS) was used to evaluate participants' perceptions of the game's usability [23]. The SUS provides an indication of the effectiveness, efficiency, and overall ease of use of a digital product. It consists of 10 items rated on a 0 to 4 Likert scale, yielding a score ranging from 0 to 100 (each item contributing up to 10 points). A score of approximately 85 is considered excellent, approximately 72 good, approximately 53 average, and approximately 38 poor.

As discussing end-of-life and advance care planning issues is not an everyday activity, Anticip'action is not expected to be played repetitively. Consequently, the first item of the SUS—"I think I would like to use this system frequently"—is only partially applicable to evaluating the game. To address this, a slightly modified version was added to the original 10 items: "I think that if I need to consider what is important to me at the end of life, I would give priority to this game." This adjustment allowed for the calculation of 2 SUS scores: the standard version and an adapted version that included the modified item.

Perceived Attractiveness

The French version of the standard 28-item AttrakDiff questionnaire [24] was used to evaluate participants' experience across four dimensions of user experience, each comprising 7 items: (1) pragmatic quality, which measures the product's usability and the perceived ease with which users can achieve their goals; (2) hedonic quality–stimulation, which assesses the extent to which the product satisfies the need for stimulation by offering new, interesting, and engaging content, features, and interaction styles; (3) hedonic quality–identity, which examines how well the product supports a social function and communicates a sense of user identity; and (4) overall attractiveness, which represents the perceived value of the product based on its pragmatic and hedonic qualities. This questionnaire includes pairs of contrasting adjectives rated on a 7-point scale ranging from −3 to +3. Scores close to the mean (between 0 and 1) indicate that the product has achieved its objectives without any significant negative impact, although improvements could be made to create a more positive user experience or to enhance attractiveness. Scores outside the neutral zone should be interpreted as either positive (from 1 to 3) or negative (from −1 to −3).

XSL•FO
RenderX

Perceived Relevance and Endorsement

To assess participants' perceptions of the app's impact on user awareness, knowledge, attitudes, intention to change, and likelihood of actual change in advance care planning behavior, the 6 "perceived impact" items (section F) of the validated Mobile Application Rating Scale (MARS) were used [25]. These items yield an overall mean score on a 5-point Likert scale, interpreted as follows: 1=inadequate, 2=poor, 3=acceptable, 4=good, and 5=excellent. To evaluate endorsement, the first 5-point Likert item of the "subjective quality" section (section E) of the MARS was used. This item asks participants whether they would recommend the digital game to others who could benefit from it. This questionnaire was translated into French in accordance with the best practices described by Wild et al [26].

Participants' Qualitative Feedback

At the end of the study, participants were asked 5 open-ended questions to gather their final feedback on the game. These questions aimed to assess (1) what participants appreciated about the game, (2) what aspects they disliked, (3) which tasks they found particularly challenging, (4) what they felt was missing in the game, and (5) whether they had any additional comments to share. To analyze the responses, the guidelines set by Paillé and Mucchielli [27] for qualitative data analysis were followed. Specifically, Excel was used for organizing and working with the material, with integrated in-text notes that were differentiated by color to facilitate categorization. A continuous thematic analysis approach was applied to systematically identify emerging themes and patterns throughout the feedback.

Content Readability

To assess the readability of the game's textual content, all text was compiled into a Microsoft Word document and evaluated using the Scolarius test [28], which was developed by Influence Communication, a Canadian media analysis organization. The readability score is calculated similarly to the Flesch Reading Ease score (commonly used for English-language materials). The test measures word and paragraph length to provide a score ranging from 50 to 250, interpreted as follows: a score between 50 and 89 corresponds to an elementary school reading level, 90 to 119 indicates a high school reading level, 120 to 149 reflects a college reading level, and 150 to >189 represents a university reading level.

Results

Recruitment

The procedure was pilot-tested with 2 participants to master the technical aspects of conducting the test online with multiple devices and to refine the multiple task instructions. These data are not included in the results. During September and October 2024, we recruited 12 participants for the online usability test. However, data from 2 participants could not be analyzed due to major technical issues: a failure of the screen-sharing functionality on Zoom in 1 case and bad quality recording in the other. The final dataset includes complete responses from 10 participants.

The mean age of the participants included in the final dataset was 41 (range 25-65) years, and the sample was diverse in terms of gender (men: 5/10, 50%; women: 5/10, 50%) and type of smartphone (operating system) used (Table 2). All participants reported using their smartphones daily, demonstrating familiarity with digital interfaces and technological tools, and they had a rather high level of education. Overall, the sample was skewed toward younger adults, with only 1 (10%) of the 10 participants being aged >60 years. All participants reported at least moderate digital literacy.

Table 2. Participants' characteristics.

Participant	Gender	Age (y)	Education	Smartphone operating system
P1	Woman	54	Professional higher education	iOS
P2	Woman	25	University	iOS
P3	Woman	35	Professional higher education	iOS
P4	Man	53	High school	iOS
P5[a]	Woman	36	High school	iOS
P6	Man	34	High school	Android
P7[a]	Man	59	University	iOS
P8	Man	35	University	Android
P9	Woman	32	University	iOS
P10	Man	27	University	iOS
P11	Man	65	University	iOS
P12	Woman	49	University	iOS

[a]Data from these participants could not be used because of technical issues.

XSL•FO
RenderX

First Impressions

In the visual recall task, 80% (8/10) of the participants noticed the presence of colored cards. Of these 8 participants, 3 (38%) observed that the cards had titles, and 2 (25%) noted the numbers displayed in the center of each card. The second most frequently recalled element (5/10, 50%) was the instruction bubble positioned in the center of the screen. Finally, the 3 circular button icons representing the game's 3 phases were remembered by 5 (50%) of the 10 participants; within this group, 3 (60%) also noticed the grayed-out card images, and another 3 (60%) noticed the text beneath the icons. Some elements remained unnoticed, including the gray background of the game; the title of the game phase at the top of the screen; the 2 square buttons on the top left (question mark icon) and top right (X icon) to access the rules or exit the game, respectively; the classification symbols (heart icons and question mark); and the numbers below these symbols. Some participants had false memories. Of the 10 participants, 1 (10%) remembered the presence of a question, possibly confusing it with the instruction bubble in the center of the screen; 1 (10%) reported 4 button icons at the top of the screen (instead of 3); and 1 (10%) mentioned "choices with several options," possibly referring to the instructions and heart icons. In addition, there was evidence of imprecise recall and potential misinterpretation of certain elements; for example, of the 10 participants, 1 (10%) described a white background for the game, 1 (10%) reported that the colored cards were placed in the center of the screen (instead of at the bottom), and 1 (10%) misremembered the position of the text as being below the cards rather than above them.

Regarding participants' understanding of the game's objectives, half (5/10, 50%) were unable to provide an answer, while the other half (5/10, 50%) identified it as a card game. Of those who recognized it as a card game, 60% (3/5) suggested that the objective might involve making choices. Each of these 5 participants provided at least 1 interpretation of the game: asking questions about various hospital-related topics, completing a multiple-choice quiz, associating each card pile with a specific value, obtaining information with the option to change one's mind, sorting cards by theme, and clicking on heart icons to indicate preferences (this participant seemed to recall the presence of the heart icons although this was not explicitly expressed during the visual recall task).

Participants' general impression of the game was mixed: of the 10 participants, 4 (40%) rated it as good, 3 (30%) as average, 2 (20%) as bad, and 1 (10%) as very bad.

Their evaluation of the game's aesthetics was also mixed: of the 10 participants, 1 (10%) found it very attractive, 2 (20%) rated it as attractive, 5 (50%) as neither attractive nor unsightly, 1 (10%) as unsightly, and 1 (10%) as very unsightly.

Think-Aloud Procedure

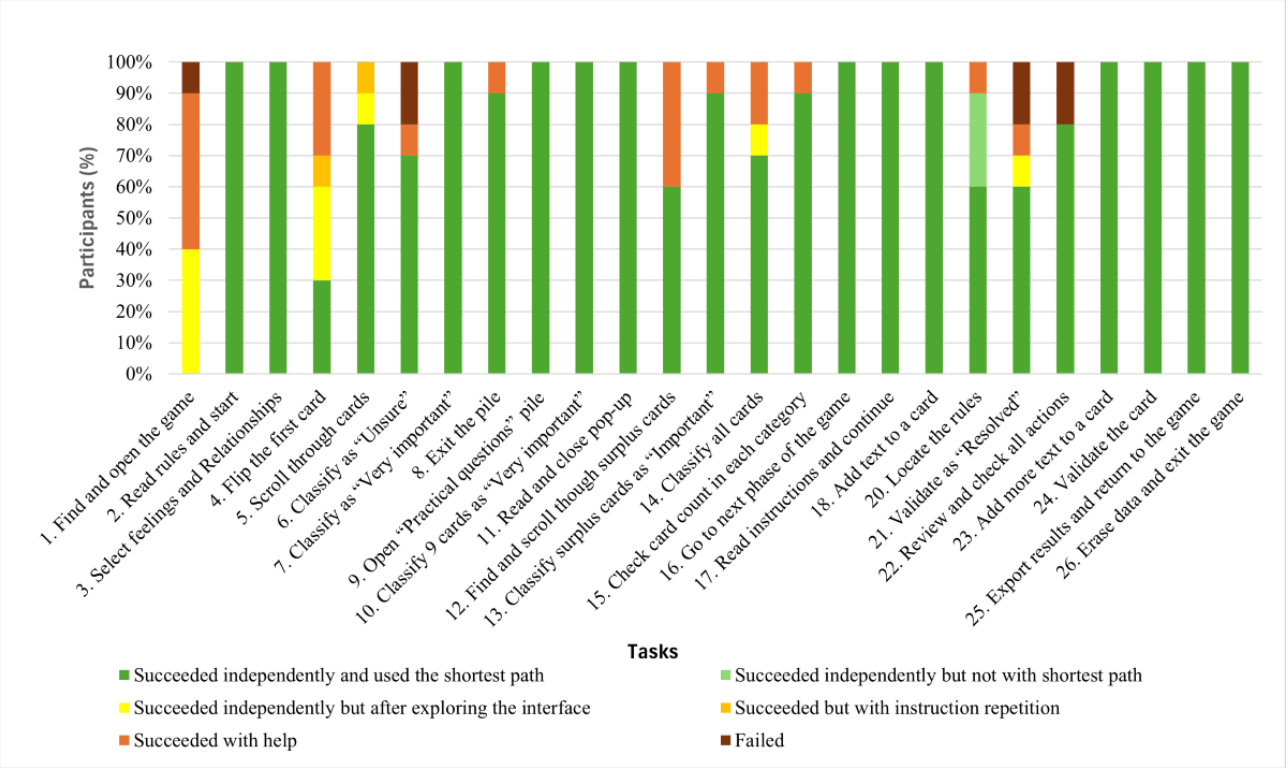
Task Success

At the analysis stage, task 19 (which overlapped partially with task 21) was excluded from the results because it seemed that most participants failed to understand the instruction due to suboptimal wording. The instruction began with "Validate the card" before unclearly alluding to the constraint that a card can only be validated once a decision has been made about whether one should do something about the card topic. Consequently, participants attempted to validate the card prematurely, before ticking the appropriate box. Unfortunately, this methodological flaw was not identified during the pilot testing phase.

As illustrated in Figure 7 and Table S1 in Multimedia Appendix 1, most of the tasks (21/25, 84%) were completed successfully by all participants, while some (4/25, 16%) required assistance. In 83.2% (208/250) of the cases, participants successfully completed tasks without assistance and found the shortest path. In 1.2% (3/250) of the cases, participants succeeded independently but did not follow the most direct route. In 4% (10/250) of the cases, participants explored the interface before finding the correct solution. Meanwhile, 8.8% (22/250) of the cases required intervention from the experimenter. Finally, 2.8% (7/250) of the cases resulted in failure, as participants were unable to complete the tasks even with assistance.

XSL•FO
RenderX

Figure 7. Proportion of participants who succeeded or failed in completing study tasks during the think-aloud procedure, with or without input from the experimenter, the type of input received, and whether the shortest path was used.



Some tasks required only minimal assistance (eg, tasks 4 and 14), while other tasks presented notable usability issues (Table 3). In particular, for completing task 1—finding the game in the Concerto app—6 (60%) of the 10 participants required assistance, such as instruction repetition, hints, additional information, or the correct answer. For task 4, which involved flipping a card, 6 (60%) of the 10 participants experienced difficulties locating the dedicated button (Figure 3). For task 6—classifying a card in the "Unsure" category—3 (30%) of the 10 participants needed substantial intervention (such as additional information, guiding questions, or answers to questions) to understand that the square "Unsure" button was clickable (the other clickable categories were represented by round buttons; Figure 3). For task 12, which involved finding the surplus cards classified as "Very important," 6 (60%) of the 10 participants faced difficulties, necessitating multiple forms of assistance, such as instruction repetition, additional information, hints, answers to questions, or guiding questions. The nature of these difficulties indicates that the wording of the task instruction was not optimal: specifically, the phrase "cartes en surplus" ("surplus cards") used in the instruction could be misleadingly interpreted as the existence of some "extra cards" in the game, although the aim was to convey the information that too many cards had been allocated to the "Very important" category, and some would need to be removed from the category before moving on to the next phase of the game. Task 21 was challenging for 3 (30%) of the 10 participants. It required understanding that in the "Specify" phase of the game, each card needed to be evaluated to determine whether it required further action before it could be marked as "resolved," after which the participant could proceed to the "Act" phase of the game (Figure 5). However, the task instruction was short and allusive, which could explain the participants' need for more guidance. In task 22, of the 10 participants, 1 (10%) failed to understand that the pile in the center needed to be "resolved," and 1 (10%) failed to click on the "Review my specifications" button to check the specifications related to the card (Figure 6). In both cases, the experimenter provided the answer.

XSL•FO
RenderX

Table 3. Description of the problems encountered by participants (n=10) during the think-aloud tasks.

Tasks and problems encountered by participants	Frequency, n (%)	Category (according to the heuristics of Bastien and Scapin [22])	Severity (0 to 4 scale based on the recommendations of Nielsen and Landauer [17])
1. Find and open the game.			
1.1. Could not find the game and clicked on the wrong tab	3 (30)	Guidance: grouping and distinction	4
1.2. Could not find the game: did not see the menu at the bottom	2 (20)	Guidance: grouping and distinction; consistency	4
1.3. Thought they were already in the game when they were on the wrong page in Accordons-nous	2 (20)	Guidance: prompting; workload: information density	3
1.4. Complained about too much text	1 (10)	Workload: information density	1
1.5. Did not see the “play” button, which allows entry into the game	1 (10)	Guidance: legibility	3
1.6. Misunderstood the “Anticip’action” tab as a way to enter the game, although it is actually a tab to navigate within sections in the Accordons-nous module	2 (20)	Guidance: grouping and distinction	3
2. Read rules and start.			
2.1. Felt overwhelmed by the text	2 (20)	Workload: information density	1
4. Flip the first card.			
4.1. Did not see the button to flip the card	4 (40)	Guidance: grouping and distinction; compatibility	3
4.2. Interpreted the return icon as a door to return to the menu	1 (10)	Consistency	1
4.3. Interpreted the card title as a clickable button	1 (10)	Consistency	0
6. Classify as “Unsure.”			
6.1. Used drag and drop to classify a card (which is not a function in the game)	2 (20)	Consistency and compatibility	1
6.2. Clicked on the wrong button	2 (20)	Consistency and significance of codes	2
8. Exit the pile.			
8.1. Did not see the X icon (located on the top right of the screen) to exit the pile	1 (10)	Guidance: prompting	0
12. Find and scroll through surplus cards.			
12.1. Tried to find the extra cards by clicking on the heart icons on the main page (which is not a function in the game)	3 (30)	Guidance: prompting and grouping and distinction; significance of codes	3
12.2. Could not find the extra cards after exploring the screen	6 (60)	Guidance: grouping and distinction; consistency	3
13. Classify surplus cards as “Important.”			
13.1. Did not see the confirmation button after completing the task	1 (10)	Workload: minimal actions; adaptability: flexibility	2
18. Add text to a card.			
18.1. Did not understand that to validate the written text, one needs to click on the X icon	1 (10)	Guidance: prompting; compatibility	2
18.2. Clicked in the text area when only the written area was active	1 (10)	Compatibility	1
20. Locate the rules.			

XSL•FO
RenderX

Tasks and problems encountered by participants	Frequency, n (%)	Category (according to the heuristics of Bastien and Scapin [22])	Severity (0 to 4 scale based on the recommendations of Nielsen and Landauer [17])
20.1. Did not see the button to find the rules	3 (30)	Guidance: legibility	1
20.2. Did not want to exit the game for fear of losing data	1 (10)	Error management: error protection; workload: information density	2
21. Validate as "Resolved."			
21.1. Did not click on the card pile in phase 2 to continue the game	1 (10)	Guidance: grouping and distinction; workload: information density	0
22. Review and check all actions.			
22.1. Clicked on the specified cards button instead of the blue pile at the bottom	1 (10)	Guidance: prompting; significance of codes	1
22.2. Did not see the button to review the details	2 (20)	Guidance: immediate feedback and legibility; significance of codes	3

Clicks to Complete Tasks

As shown in Table 4, a comparison between participants' actual number of clicks and the expected optimal number reveals varying levels of efficiency across tasks. For most of the tasks, participants generally stayed close to the expected optimal number of clicks, indicating a solid understanding of the interface and task requirements. However, several tasks stand out with higher click counts, suggesting areas where participants may have encountered challenges or required additional exploration before completing the task. In particular, task 1, which involved finding the game in the Concerto app, was completed with a mean of 10.30 (SD 4.47) clicks compared to the expected 3.

XSL•FO
RenderX

Table 4. Participants' overall click performance per task: mean number of clicks used, expected optimal number of clicks, and deviation from the optimal number.

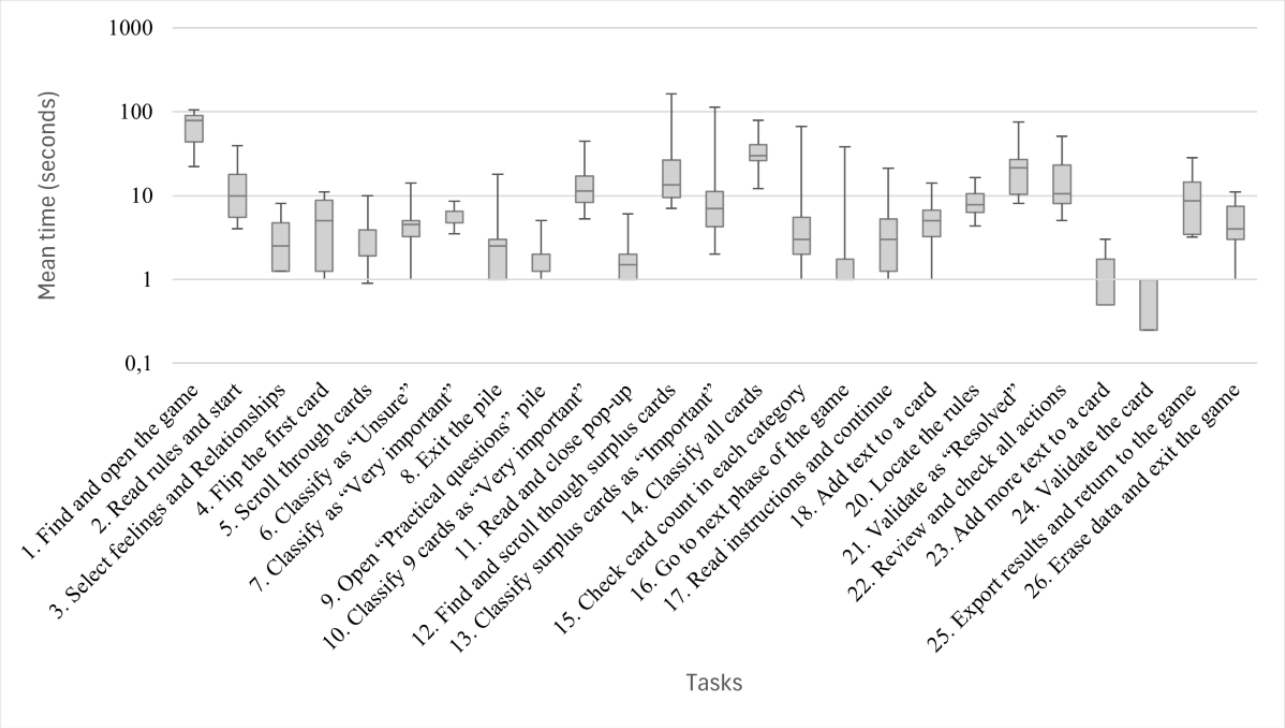
Tasks	Number of clicks, mean (SD)	Optimal number of clicks, n	Difference
1. Find and open the game.	10.30 (4.47)	3.00	7.30
2. Read rules and start.	4.00 (0.00)	4.00	0.00
3. Select "Feelings and relationships."	1.00 (0.00)	1.00	0.00
4. Flip the first card.	2.50 (1.43)	1.00	1.50
5. Scroll through cards.	3.00 (1.70)	2.00	1.00
6. Classify as "Unsure."	2.40 (2.84)	1.00	1.40
7. Classify as "Very important."	3.20 (0.42)	3.00	0.20
8. Exit the pile.	1.20 (0.63)	1.00	0.20
9. Open "Practical questions" pile.	1.00 (0.00)	1.00	0.00
10. Classify 9 cards as "Very important."	8.10 (0.32)	8.00	0.10
11. Read and close pop-up.	1.00 (0.00)	1.00	0.00
12. Find and scroll through surplus cards.	5.20 (2.30)	4.00	1.20
13. Classify surplus cards as "Important."	3.40 (0.97)	3.00	0.40
14. Classify all cards.	25.40 (0.97)	25.00	0.40
15. Check card count in each category.	0.00 (0.00)	0.00	0.00
16. Go to next phase of the game.	1.00 (0.00)	1.00	0.00
17. Read instructions and continue.	1.00 (0.00)	1.00	0.00
18. Add text to a card.	2.70 (0.95)	2.00	0.70
20. Locate the rules.	2.40 (0.52)	2.00	0.40
21. Validate as "Resolved."	20.30 (2.26)	19.00	1.30
22. Review and check all actions.	6.30 (3.13)	5.00	1.30
23. Add more text to a card.	2.40 (0.70)	2.00	0.40
24. Validate the card.	1.00 (0.00)	1.00	0.00
25. Export results and return to the game.	3.00 (0.00)	3.00	0.00
26. Erase data and exit the game.	2.00 (0.00)	2.00	0.00

Time Spent on Tasks

As illustrated in Figure 8, the analysis of time allocation per task reveals substantial variability among participants. Nevertheless, some patterns and deviations from precalculated optimal times provide insight into specific usability challenges. Notably, task 1 ("Find and open the game") required considerably more time than anticipated, indicating that participants faced initial difficulties with the game's discoverability. Participants 6 and 8 displayed markedly extended times for task 1, suggesting that the entry point for locating the game lacked intuitive visibility or clarity. Divergence from the optimal could also be observed in tasks 10, 14, and 21, which required participants to categorize or "resolve" numerous cards. By design, these tasks were lengthy rather than difficult to navigate: while some participants elected to read each card before sorting it, others proceeded with the task more quickly without detailed review, leading to substantial individual variations in completion time. A similar pattern could be observed with task 2, which asked participants to read the rules of the game and start: some participants took much more time to read the rules than others. The time taken to complete tasks 12 and 21 (discussed previously) also diverged notably from optimal times, partly due to issues stemming from unclear instructions and partly due to suboptimal guidance in the game.

XSL•FO
RenderX

Figure 8. Mean time (in seconds) spent by participants to complete each of the 25 tasks during the think-aloud procedure, displayed on a logarithmic scale.



Analysis of Errors and Problems Encountered

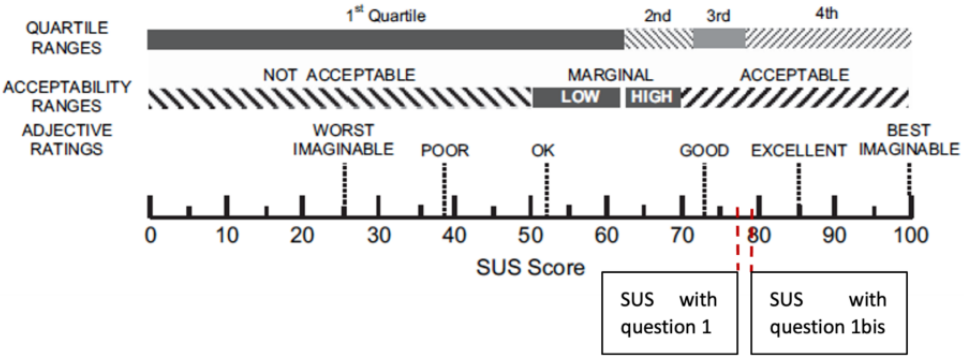
Following the criteria formulated by Bastien and Scapin [22], we recorded 23 problems encountered by the participants while completing the 26 tasks (Table 3). Of these 23 problems, 2 (9%) were classified as usability catastrophes, 7 (30%) as major usability problems, 5 (22%) as minor usability problems, 6 (26%) as cosmetic issues, and 3 (13%) as not constituting a usability problem. A majority of the problems (14/23, 61%) were related to guidance difficulties, such as the visibility of visual items or the way they were displayed in relation to one another.

Questionnaires

Perceived Usability

As illustrated in Figure 9, the mean score on the standard SUS was 77.75 (SD 5.1) out of 100. The adapted version of the SUS, which included a more fitting variant of the first question, yielded a slightly higher mean score of 79.25 (SD 5.6) out of 100. These results indicate that participants rated the game's usability between good and excellent.

Figure 9. Scores of Anticip'action on the standard and adapted versions of the System Usability Scale (SUS).

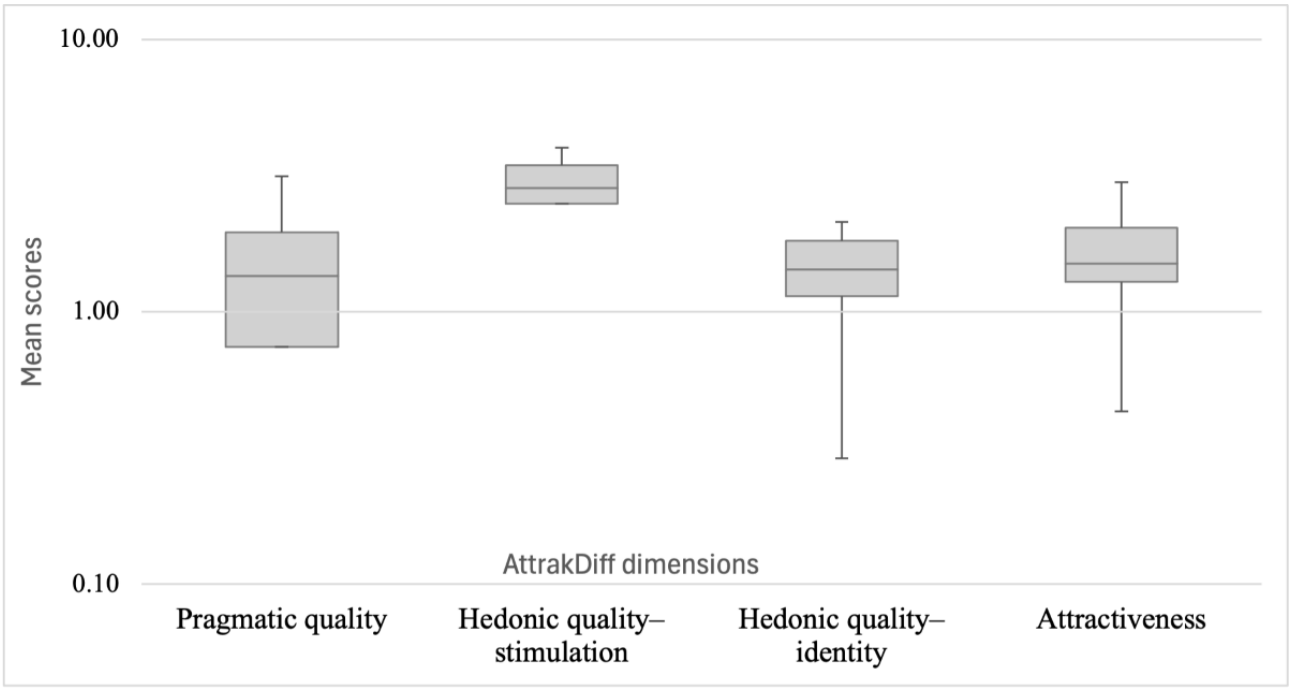


Perceived Attractiveness

The game obtained an overall attractiveness score of 1.57, indicating a positive assessment. As illustrated in Figure 10, positive results were obtained on all 4 dimensions of the test, with scores ranging from 1.16 to 1.57.

XSL•FO
RenderX

Figure 10. Mean scores of Anticip'action on the 4 dimensions of the AttrakDiff questionnaire, displayed on a logarithmic scale.



Perceived Relevance and Endorsement

The mean score on the MARS questionnaire was 4.1 (SD 2.1), indicating that participants evaluated the relevance of the game as good. As illustrated in Table 5, participants were mostly convinced that the game can increase awareness, knowledge, and intention to change regarding the topic of anticipating difficult situations that may arise in the future.

The game also ranked positively on the "I would recommend" question, with a mean score of 3.3 (SD 1.1). Of the 10 participants, 2 (20%) indicated that they would highly recommend the game to everyone who could benefit from it, 1 (10%) would recommend it to many people, 5 (50%) to several people, and 2 (20%) to very few people. No participant stated that they would not recommend the game.

Table 5. Average responses of the 10 participants to the Mobile Application Rating Scale (MARS) questionnaire, section F. All questions ended with "[...] the anticipation of difficult situations that may arise in the future."

	Scores, mean (SD)
Awareness: this game is likely to increase awareness of the importance of...	4.5 (2.4)
Knowledge: this game is likely to increase understanding of...	4.2 (1.9)
Attitude: this game is likely to change attitudes toward improving...	3.8 (1.9)
Intention to change: this game is likely to increase motivation to address...	4.1 (1.4)
Help seeking: use of this game is likely to encourage further help seeking...	4.0 (2.6)
Behavior change: use of this game is likely to increase...	4.0 (2.7)

Participants' Qualitative Feedback

Participants generally provided positive feedback about the game. Most found it user-friendly, with clear visuals and easy navigation. Of the 10 participants, 7 (70%) highlighted that the game addressed the sensitive topic of advanced care directives and end-of-life care in an approachable and meaningful way, while 6 (60%) stated that the game encouraged reflection on complex health-related issues and promoted awareness of often-overlooked topics.

Areas for improvement were also suggested. Several participants reported feeling overwhelmed by the amount of information and recommended adjustments to reduce cognitive load. They suggested adding an onboarding tutorial to help users understand the game's mechanics. Some participants proposed that offering personalized content options (eg, an option to filter cards based on user profile) may help to better tailor the game to individual users' needs. Navigation issues were also mentioned: participants indicated that the "Unsure" and the "flip card" buttons were not visually intuitive as clickable elements, and navigation could be improved through clearer visual cues such as more prominent buttons. Moreover, some participants recommended incorporating dynamic features such as animations, which could increase accessibility and maintain user engagement. Finally, several participants noted that the term "jeu" ("game") does not fit with the seriousness of the content and suggested some rephrasing or providing a content advisory to prepare users for the sensitive nature of the topics discussed.

XSL•FO
RenderX

Content Readability

The readability analysis of the game cards, assessed using the Scolarius program, yielded distinct scores for the rectos (card titles) and versos (explanatory text). The rectos had an average score of 50.5, indicating content accessible to individuals with an elementary school education. By contrast, the versos scored 119.8, which corresponds to a high school reading level. Specifically, in terms of readability, 61% (20/33) of the verso cards were at the elementary school level, 18% (6/33) at the high school level, and 21% (7/33) at the university level.

Modifications to Be Made Based on the Results

Analysis of the 23 usability issues and the proposed solutions (Table 6) showed that most problems were related to users' difficulty in identifying which visual elements were actionable. Some buttons were not consistently interpreted as clickable, while certain titles were occasionally mistaken for buttons. No substantial structural modifications seem to be required to address these usability issues, with the exception of the location of the game: it is currently nested within a section of a module (Accordons-nous) inside an app (Concerto), making it difficult for users to access Anticip'action.

Table 6. Changes to be made to the game in response to the identified usability issues.

Items	Problems encountered by participants	Changes to be made
1.1, 1.2, and 1.3	Could not find the game nested in the Accordons-nous module	Create a video explaining where the game is located
1.4	Complained about too much text	No changes to be made
1.5 and 1.6	Did not see the "play" button, which allows entry into the game	Highlight the button by bolding the text
2.1	Felt overwhelmed by the text	No changes to be made
4.1	Did not see the button to flip the card	Move the button to the corner or bottom of the card
4.2	Interpreted the return icon as a door to return to the menu	No changes to be made
4.3	Interpreted the card title as a clickable button	Redesign the title by, for example, removing the border
6.1	Used drag and drop to classify a card (which is not a function in the game)	No changes to be made
6.2	Clicked on the wrong button	Align the 4 buttons on the same level, enlarge the question mark, ensure that it uses the same visual weight and style as the heart buttons (eg, same height, rounded shape, and icon structure). Alternatively, add a heart icon with a question mark in the center and place the phrase "I am unsure" below it.
8.1	Did not see the X icon (located on the top right of the screen) to exit the pile	No changes to be made
12.1 and 12.2	Could not find the extra cards	Gray out the hearts or make them clickable to access the menu
13.1	Did not see the confirmation button after completing the task	Change the button design, round the edges, and ensure consistency
18.1	Did not understand that to validate the written text, one needs to click on the X icon	Add a button to validate the text
18.2	Clicked in the text area when only the written area was active	No changes to be made
20.1	Did not see the button to find the rules	No changes to be made
20.2	Did not want to exit the game for fear of losing data	Specify more clearly that exiting the game will not result in data loss
21.1	Did not click on the card pile in phase 2 to continue the game	No changes to be made
22.1	Clicked on the specified cards button instead of the blue pile at the bottom	No changes to be made
22.2	Did not see the button to review the details	Redesign the button, which resembles the card titles too closely, to make it more visible

XSL•FO
RenderX

Discussion

Principal Findings

This study evaluated the usability, readability, and user experience of the digital version of the game Anticip'action, designed to facilitate end-of-life discussions and advance care planning decisions. The findings indicate overall satisfactory results.

After 5 seconds of visual exposure to the main screen of the game, participants were able to recall key visual elements, such as the colored cards and accompanying text. However, they struggled to notice other important interface components, including classification symbols and navigation buttons. Some misinterpretations of visual elements were observed, which was confirmed by participants' fragmented and diverse understanding of the game's objectives. Moreover, the aesthetics of the game left participants mostly indifferent. These results indicate that there is room for improvement in the visual appearance of the game to better engage users. However, these findings should be interpreted in light of the fact that in a real-world scenario, before reaching the main game screen used for the 5-second test, users would always receive preliminary information about the game's content, purpose, and rules. Specifically, the welcome page included a pictogram with cards and introductory text explaining the game's purpose (Figure 1). Moreover, a pop-up explained the card sorting task and classification symbols at the start of a new game.

The think-aloud procedure yielded a high overall success rate. In 83.2% (208/250) of the cases, participants completed the tasks without assistance. Participants' performance also closely matched the expected number of clicks, indicating that the game offered clear navigation paths. However, for some tasks (4/25, 16%), participants required assistance, revealing serious usability problems. Analysis of the 23 failures and difficulties encountered revealed that 3 (13%) issues were partly due to a suboptimal wording of task instructions (tasks 12 and 21), that 14 (61%) problems were related to guidance, and that there were 9 (39%) major usability issues that needed to be addressed (Table 3). Specifically, the following tasks highlighted important difficulties related to navigation and discoverability: task 1, which involved finding the game within the Concerto app; and task 12, which involved finding and reclassifying the surplus cards classified in the "Very important" category. About half of the participants needed guidance and took longer than the expected optimal time to complete these tasks (Figure 7; Table 7). These issues need to be addressed with clearer instructions and improved visual guidance. Moreover, the test revealed occasional design inconsistencies (eg, some buttons were not visible enough in tasks 1, 4, and 13), excessive cognitive load (eg, some pages had numerous elements and text), and inadequate error prevention (eg, once players opened a pile of cards, they expected to navigate by scrolling, whereas "resolving" cards required clicking on a button; task 21). All identified usability issues could be addressed through minor modifications, as highlighted in Table 6.

Table 7. Summary of relevant usability problems encountered during the think-aloud test.

Tasks	Difference between the average time spent by participants per task and the expected optimal time (seconds)	Problems (n=23), n (%)	Participants affected (n=10), n (%)
1. Find and open the game.	29	6 (26)	7 (70)
2. Read rules and start.	7	1 (4)	2 (20)
4. Flip the first card.	2	3 (13)	5 (50)
6. Classify as "Unsure."	2	2 (7)	3 (30)
8. Exit the pile.	3	1 (4)	1 (10)
12. Find and scroll through surplus cards.	16	2 (7)	5 (50)
13. Classify surplus cards as "Important."	16	1 (4)	1 (10)
18. Add text to a card.	4	2 (7)	2 (20)
20. Locate the rules.	4	2 (7)	4 (40)
21. Validate as "Resolved."	8	1 (4)	1 (10)
22. Review and check all actions.	5	2 (7)	3 (30)

At the end of the test, Anticip'action was rated between good and excellent on the SUS. It obtained a mean score of 77.75 out of 100 on the standard SUS and a mean score of 79.25 out of 100 on the adapted version. The increase in score resulting from the adaptation of the first question highlights the importance of context-specific adjustments to usability assessment tools.

The game also received an overall positive assessment of its attractiveness, although there remains room for improvement. The global score of 1.57 on the AttrakDiff questionnaire reflects a generally positive perception. Scores on the pragmatic and hedonic scales ranged from 1.16 to 1.57, all above zero, indicating a baseline level of user satisfaction. Hedonic quality received the lowest score, suggesting that improvements are needed to better engage users and sustain their interest. Enhancements to the design and animated and visually attractive content, such as the inclusion of more visually dynamic elements in the game, could increase its attractiveness while addressing some navigation or cognitive load issues. Reducing the number

XSL•FO
RenderX

of actions required in certain steps, notably for re-sorting the cards or "resolving" them in the "Specify" phase of the game, could help maintain user engagement. At a minimum, greater clarity and streamlining of these task processes would improve user experience.

Although this test included a small number of participants and no quantitative conclusions should be drawn, we noticed that older participants and those recruited via the patients-as-partners program at the hospital rated the game more positively. This may be due to a higher baseline receptivity to the topic of advance care planning discussions.

Anticip'action was also evaluated as relevant, with an overall score of 4.1 out of 5 on the MARS. Specifically, the game was judged as particularly impactful in raising awareness, improving knowledge, and motivating behavior change regarding advance care planning. Participants also endorsed the game with a mean score of 3.3 out of 5 on the "I would recommend" question. Given the specificity and sensitivity of the topic, this result is promising.

The readability analysis showed that the card titles are understandable for individuals with an elementary school education, which is an excellent result. However, the longer explanations on the versos of the cards require a high school reading level, indicating room for improvement. Part of the difficulty may be due to the use of gender-neutral language (which has the effect of increasing the lengths of sentences in French) as well as to the use of unavoidable technical terms such as *surrogate decision maker*, *digital legacy*, or *lasting power of attorney*. Efforts are underway to streamline and simplify the card content while preserving its informational value.

Participants provided overall positive qualitative feedback while pointing out aspects that could be improved. Most described the game as user-friendly, with clear visuals and easy navigation. They highlighted its relevance in fostering engagement and encouraging reflection on complex topics such as advance care directives.

However, some participants identified areas that needed improvement. Their concerns included the cognitive overload and some unclarity in the navigation process related to specific tasks. Regarding cognitive overload, some parts of the game contained dense textual information. Although the game achieved good usability scores overall, this suggests that some refinement may be necessary to improve accessibility, for instance, through progressive disclosure, pictorial supports, or guided onboarding elements. Some participants emphasized the need for an onboarding tutorial to improve accessibility as well as more animations, intuitive interactive elements, clearer buttons, and personalized content pathways to improve visual clarity and user engagement. Some participants also disagreed with the label "game" for such a conversation tool and expressed the need for providing a content advisory to prepare users for the sensitive nature of the topics discussed. This highlights a semantic tension between the playful connotation of the term "game" and the seriousness of the topic, which may require careful framing to ensure user acceptance.

To address the identified issues while taking participants' recommendations into account, several improvements are envisaged. Easily implementable solutions are detailed in Table 6. Furthermore, modifications may include the introduction of an onboarding tutorial to enhance initial navigation and reduce cognitive load. Options to personalize some features of the game (eg, relaxing the 10-card limit) or to filter subgroups of cards may also help further align the game with diverse user needs and preferences. In addition, enhancing the game's visual appeal by incorporating more dynamic features (eg, animated movement of cards between piles) could increase engagement and satisfaction, especially among younger users.

Limitations

Although this study reports positive usability results, it does not demonstrate that Anticip'action is an effective tool to promote end-of-life discussions. Moreover, the quantitative results from the SUS, AttrakDiff, and MARS questionnaires should be interpreted with caution, given the limited number of participants (n=10). Complementary studies with appropriate designs and more diverse populations are necessary to confirm the game's perceived value and effectiveness.

One important limitation of this study is the demographic profile of the participants. The sample included 10 individuals of different ages (ranging from 25 to 65 y), of whom 6 (60%) were in the younger-to–middle adult range (25-45 y). Prior research has shown that age can influence usability metrics, including task performance and perceived ease of use [29]. The limited sample size in this study does not allow for an evaluation of this effect. Moreover, the sample included a participant aged >60 years and none with low digital literacy or lower levels of education. As such, the findings may not be generalizable to these groups. Usability outcomes may differ significantly among older adults or individuals with low digital literacy, who may face greater barriers in navigating the tool. This possibility should be considered when interpreting our findings. Future work should explicitly include a more diverse participant pool, encompassing a wider range of health statuses and cultural backgrounds, to evaluate the accessibility and acceptability of the tool in broader populations.

Participants were asked to imagine being an individual aged 50 years reflecting on end-of-life preferences. This scenario was used to standardize the testing context, but it does not replace real-life engagement with the topic. Future studies should include participants who are directly involved in advance care planning processes.

The test was conducted online with the game displayed on participants' devices. Therefore, their facial expressions could not be observed and used as nonverbal cues to assess the gravity of usability issues.

Part of the participants' failure to complete some tasks (in particular, task 19, which was removed from the analysis, as well as tasks 12 and 21) may be due to suboptimal wording of the instructions.

Some participants took more time than others to read the text on the cards before completing the tasks. This difference in participants' behavioral style could have influenced the time

XSL•FO
RenderX

spent on tasks, although it is not an indication of failure to complete them. This was particularly obvious on task 14 ("Classify all cards") and may have also affected tasks 10 and 21. Therefore, time-based measures are difficult to interpret as usability factors in this study.

Conclusions

Anticip'action is a conversation tool aimed at initiating end-of-life and advance care planning discussions and increasing patients' empowerment during these discussions. It offers accessible content, simple navigation, and thoughtful discussion prompts to support individuals, families, and health care professionals in exploring these important and often challenging topics. It is also designed to motivate related action planning, such as writing advance directives.

This study evaluated the usability, readability, and user experience of the digital version of the game with participants of different ages (ranging from 25 to 65 y) who were already familiar with digital interfaces and technological tools. The results demonstrated that participants found the game generally user-friendly, with clear visuals and intuitive navigation. Qualitative feedback highlighted its relevance. Specific challenges were identified, including cognitive overload and some unclear navigation elements. Addressing these issues through improved design, clearer instructions, and better user guidance would further enhance user experience and engagement.

Overall, the usability findings validated Anticip'action as an accessible conversation game for patients who are minimally at ease with digital tools and the French language. While additional testing with health care professionals and broader user groups (including older patients and individuals less comfortable with digital solutions) would strengthen its evidence base, our findings indicate that Anticip'action could be a valuable resource for some patients, families, and caregivers navigating end-of-life decisions and advance care planning. It would be interesting to translate it into other languages, as is the case for the printed version of Anticip'action.

Acknowledgments

The authors would like to thank everyone who contributed to the development of the game, notably Catherine Bollondi, Marta Fadda, Samia Hurst, Pascale Lefuel, Sophie Pautex, Arnaud Ricci, Marianne Rubio, Céline Schöpfer, Inès Serre, Pierre Sutter, and Aida Yamani. The authors also wish to thank the 2 anonymous reviewers who helped improve the draft during the review process. The project received financial support from the Geneva University Hospital and the Fondation Privée des Hôpitaux Universitaires de Genève, a nonprofit organization that had no involvement in any stage of the study.

Data Availability

The datasets generated and analyzed during this study are available in the Open Science Foundation (OSF) repository [30].

Authors' Contributions

CC, DC, and FE conceptualized the study. AB and FE were responsible for software. All authors contributed to data curation. CC and FE were responsible for resources, supervision, validation, and project administration. DC conducted the formal analysis and was responsible for investigation. CC was responsible for acquiring funding. AB and DC were responsible for visualization. CC, DC, and FE were responsible for methodology. CC and DC wrote the original draft. All authors reviewed and edited the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Table showing task completion success rates and related details, along with the questionnaires and the discussion grid used for the final qualitative feedback, provided in both the original French and English translation.
[DOCX File , 87 KB - aging_v8i1e73378_app1.docx]

References

1. Rietjens JA, Sudore RL, Connolly M, van Delden JJ, Drickamer MA, Droger M, et al. Definition and recommendations for advance care planning: an international consensus supported by the European Association for Palliative Care. Lancet Oncol 2017 Sep;18(9):e543-e551 [FREE Full text] [doi: 10.1016/S1470-2045(17)30582-X] [Medline: 28884703]
2. McMahan RD, Tellez I, Sudore RL. Deconstructing the complexities of advance care planning outcomes: what do we know and where do we go? A scoping review. J Am Geriatr Soc 2021 Jan 07;69(1):234-244 [FREE Full text] [doi: 10.1111/jgs.16801] [Medline: 32894787]
3. Anderson RM, Funnell MM. Patient empowerment: myths and misconceptions. Patient Educ Couns 2010 Jun;79(3):277-282 [FREE Full text] [doi: 10.1016/j.pec.2009.07.025] [Medline: 19682830]

XSL•FO
RenderX

4. Trankle SA, Shanmugam S, Lewis E, Nicholson M, Hillman K, Cardona M. Are we making progress on communication with people who are near the end of life in the Australian health system? A thematic analysis. Health Commun 2018 Nov 26;35(2):158-167. [doi: 10.1080/10410236.2018.1548335]
5. Periyakoil V, Neri E, Kraemer H. No easy task: a mixed methods study of barriers to conducting effective end-of-life conversations reported by multi-specialty doctors (S776). J Pain Symptom Manage 2015 Feb;49(2):446-447. [doi: 10.1016/j.jpainsymman.2014.11.255]
6. Serey K, Cambriel A, Pollina-Bachellerie A, Lotz JP, Philippart F. Advance directives in oncology and haematology: a long way to go-a narrative review. J Clin Med 2022 Feb 23;11(5):1195 [FREE Full text] [doi: 10.3390/jcm11051195] [Medline: 35268299]
7. Momen NC, Barclay SI. Addressing 'the elephant on the table': barriers to end of life care conversations in heart failure - a literature review and narrative synthesis. Curr Opin Support Palliat Care 2011 Dec;5(4):312-316. [doi: 10.1097/SPC.0b013e32834b8c4d] [Medline: 21897257]
8. Daruwalla Z, Thakkar V, Aggarwal M, Kiasatdolatabadi A, Guergachi A, Keshavjee K. Patient empowerment: the role of technology. Stud Health Technol Inform 2019;257:70-74. [Medline: 30741175]
9. Riley SR, Voisin C, Stevens EE, Bose-Brill S, Moss KO. Tools for tomorrow: a scoping review of patient-facing tools for advance care planning. Palliat Care Soc Pract 2024 Jun 24;18:26323524241263108 [FREE Full text] [doi: 10.1177/26323524241263108] [Medline: 39045292]
10. Liu L, Zhao YY, Yang C, Chan HY. Gamification for promoting advance care planning: a mixed-method systematic review and meta-analysis. Palliat Med 2021 Jun 29;35(6):1005-1019. [doi: 10.1177/02692163211005343] [Medline: 33775174]
11. Van Scoy LJ, Green MJ, Reading JM, Scott AM, Chuang CH, Levi BH. Can playing an end-of-life conversation game motivate people to engage in advance care planning? Am J Hosp Palliat Care 2017 Sep 12;34(8):754-761 [FREE Full text] [doi: 10.1177/1049909116656353] [Medline: 27406696]
12. McDarby M, Llaneza D, George L, Kozlov E. Mobile applications for advance care planning: a comprehensive review of features, quality, content, and readability. Am J Hosp Palliat Care 2021 Aug 17;38(8):983-994 [FREE Full text] [doi: 10.1177/1049909120959057] [Medline: 32938207]
13. Dupont C, Smets T, Monnet F, Eneslätt M, Tishelman C, Van den Block L. The cultural adaptation of the go wish card game for use in Flanders, Belgium: a public health tool to identify and discuss end-of-life preferences. BMC Public Health 2022 Nov 17;22(1):2110 [FREE Full text] [doi: 10.1186/s12889-022-14523-9] [Medline: 36397020]
14. Concerto. Centre de l'Innovation, Hôpitaux Universitaires de Genève. URL: https://www.hug.ch/concerto [accessed 2025-07-07]
15. Anticip'action. Hôpitaux Universitaires de Genève. URL: https://www.hug.ch/centre-linnovation/anticipaction [accessed 2025-07-07]
16. Menkin ES. Go Wish: a tool for end-of-life care conversations. J Palliat Med 2007 Apr;10(2):297-303. [doi: 10.1089/jpm.2006.9983] [Medline: 17472498]
17. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. In: Proceedings of the INTERACT' 93 and CHI' 93 Conference on Human Factors in Computing Systems. 1993 Presented at: INTERCHI93; April 24-29, 1993; Amsterdam, The Netherlands.
18. Anticip'action - Usability test. OSF Registries. URL: https://osf.io/8gzsf [accessed 2025-07-16]
19. Jaspers MW, Steen T, van den Bos C, Geenen M. The think aloud method: a guide to user interface design. Int J Med Inform 2004 Nov;73(11-12):781-795. [doi: 10.1016/j.ijmedinf.2004.08.003] [Medline: 15491929]
20. Gronier G. Measuring the first impression: testing the validity of the 5 second test. J User Exp 2016;12(1):8-25 [FREE Full text]
21. Doncaster P. The UX Five-Second Rules: Guidelines for User Experience Design's Simplest Testing Technique. Amsterdam, The Netherlands: Elsevier Science; 2014.
22. Bastien JM, Scapin DL. Ergonomic criteria for the evaluation of human-computer interfaces. Inria. 1993 May. URL: https://www.usabilis.com/wp-content/uploads/2017/09/Criteres-Ergonomiques-pour-l-Evaluation-d-Interfaces-Utilisateur-Scapin-Bastien.pdf [accessed 2025-07-14]
23. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation in Industry. London, UK: CRC Press; 1996.
24. Lallemand C, Koenig V, Gronier G, Martin R. Création et validation d'une version française du questionnaire AttrakDiff pour l'évaluation de l'expérience utilisateur des systèmes interactifs. Eur Rev Appl Psychol 2015 Sep;65(5):239-252. [doi: 10.1016/j.erap.2015.08.002]
25. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
26. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR task force for translation and cultural adaptation. Value Health 2005;8(2):94-104 [FREE Full text] [doi: 10.1111/j.1524-4733.2005.04054.x] [Medline: 15804318]

XSL•FO
RenderX

27. Paillé P, Mucchielli A. L'analyse Qualitative en Sciences Humaines et Sociales Fifth Edition. Paris, France: Armand Colin; 2021.
28. Scolarius homepage. Scolarius. URL: https://www.scolarius.com/ [accessed 2024-09-01]
29. Sonderegger A, Schmutz S, Sauer J. The influence of age in usability testing. Appl Ergon 2016 Jan 01;52:291-300. [doi: 10.1016/j.apergo.2015.06.012] [Medline: 26360221]
30. Campioni D, Ehrler F, Clavien C. Anticip'action - Usability test. OSF Home. 2024. URL: https://osf.io/n3kbc/?view_only=e5a4f79e1a7b4384bc86d9d180963e9e [accessed 2025-07-16]

Abbreviations

HUG: Geneva University Hospitals
MARS: Mobile Application Rating Scale
SUS: System Usability Scale

Edited by H LaMonica; submitted 04.03.25; peer-reviewed by HF Hsieh, JH Song; comments to author 21.04.25; revised version received 17.05.25; accepted 15.06.25; published 21.08.25.

Please cite as:
Campioni D, Ehrler F, Berger A, Clavien C
A Digital Serious Game (Anticip'action) to Support Advance Care Planning Discussions in the General Population: Usability Study
JMIR Aging 2025;8:e73378
URL: https://aging.jmir.org/2025/1/e73378
doi:10.2196/73378
PMID:

©Dafne Campioni, Frederic Ehrler, Antoine Berger, Christine Clavien. Originally published in JMIR Aging (https://aging.jmir.org), 21.08.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluating the Feasibility of a Dyadic, Touch-Based Multimedia Tablet Intervention and Its Effects on the Caregiver-Patient Relationship Among Individuals With Mild Cognitive Impairment: Qualitative Triangulation Study

Alexa von Bosse[1,2], BSc, MSc; Anna Thum[3], BSc; Sonja Mahler[3], BSc; Henriette Schulz[3], BSc; Paulina Weuthen[3], BA; Christophe Kunze[1], Prof Dr

[1]Care & Technology Lab, Faculty 3: Health, Medical & Life Sciences, Furtwangen University, Furtwangen, Germany

[2]Competence Center Dementia, Department Health, Ostschweizer Fachhochschule OST, St Gallen, Switzerland

[3]Faculty 3: Health, Medical & Life Sciences, Furtwangen University, Furtwangen, Germany

Corresponding Author:
Alexa von Bosse, BSc, MSc
Care & Technology Lab
Faculty 3: Health, Medical & Life Sciences
Furtwangen University
Robert-Gerwig-Platz 1
Furtwangen, 78120
Germany
Phone: 49 7723 920 2583
Email: alexa.vonbosse@ost.ch

Abstract

Background: Approximately 20% of the global population is affected by mild cognitive impairment (MCI), with around 15% progressing to dementia within 2 years. Touch-based multimedia applications can support cognitive, social, and physical functioning, potentially enhancing daily life and strengthening caregiver-patient relationships through shared engagement. Although interest in dyadic, technology-assisted interventions is increasing, empirical evidence on their feasibility and acceptability in home-based settings remains scarce. In particular, little is known about their impact on caregiver–care recipient dynamics and the factors that facilitate or hinder their use.

Objective: We aimed to evaluate the feasibility of a dyadic, tablet-based multimedia intervention for individuals with MCI and their caregivers in a home setting, focusing on user experiences, use barriers and facilitators, and the intervention's impact on the caregiver-patient relationship.

Methods: We applied qualitative triangulation, combining naturalistic observations and semistructured interviews. Data were analyzed using qualitative content analysis. The intervention was codeveloped by a multidisciplinary team and implemented as a user-centered, tablet-based modular platform with customizable cognitive, physical, and interactive exercises.

Results: We recruited a total of 40 participants, comprising 20 (50%) individuals with MCI and their 20 (50%) caregivers. Our study confirmed the feasibility of a touch-based multimedia intervention for both groups. Despite initial challenges with navigation and touch interfaces, most participants demonstrated increased confidence and competence, particularly with tailored caregiver support. Instances of enhanced communication and emotional connection were described by care partners and then became visible during their interaction with the intervention. Shared moments of laughter, mutual encouragement, and coordinated task execution indicated that the intervention could create opportunities for relational closeness. Biography-related tasks proved particularly effective, stimulating meaningful conversation and storytelling that facilitated the sharing of personal memories and experiences. While some dyads experienced occasional tensions due to impatience, differing expectations, or dominant behavior by one partner, the overall atmosphere was one of cooperation, support, and adaptability. The intervention demonstrated how technology can serve as a tool to facilitate shared experiences, promote cognitive engagement, and enrich interpersonal relationships in everyday caregiving contexts.

Conclusions: Our findings show that the home-based implementation of a dyadic, touch-based multimedia intervention can be feasible for individuals with MCI and their caregivers. A key finding is that biographically oriented content fosters caregiver-patient

XSL•FO
RenderX

relationships by acting as a catalyst for personal dialogues and collective reminiscence. These interactions enhance emotional intimacy and mutual understanding, highlighting the potential of technology-driven interventions in dementia care. The intervention incorporated multiple domains, including cognitive stimulation, physical activation, and communicative-social interaction, all of which proved highly promising. In addition, the regular implementation of the intervention in home settings appears to be a realistic and achievable approach.

(JMIR Aging 2025;8:e75189) doi:10.2196/75189

KEYWORDS

mild cognitive impairment; caregivers; relationship; feasibility; tablet-based intervention; multimedia intervention; dyadic intervention; artificial intelligence

Introduction

Epidemiology of Dementia and Mild Cognitive Impairment

Approximately 50 million people are living with dementia, and the prevalence of dementia is projected to increase to 152 million by 2050 [1]. Mild cognitive impairment (MCI) is characterized by a decline in cognitive function that is more pronounced than normal aging but not as severe as dementia, often manifesting as memory impairment [2]. Approximately 15% of individuals with MCI progress to dementia within 2 years [3]. While MCI does not typically lead to significant impairments in activities of daily living [2], individuals with dementia often experience substantial limitations in performing everyday tasks [4].

Tablet-Based Interventions and Dyadic Interaction

Given the potential progression to dementia and the increasing prevalence of MCI with advancing age, early interventions, such as cognitive training applications, are particularly important in an aging society [3]. Maintaining cognitive, social, and physical functioning is essential for preserving quality of life. Assistive technologies, including touch-based multimedia applications, have been shown to support daily activities, enhance quality of life, and reduce the financial burden of inpatient care [4]. Cognitive stimulation, such as through touch-based multimedia applications, may contribute to the management of MCI and mitigate the cognitive decline of affected individuals [5].

While tablet-based interventions for individuals with MCI may offer cognitive and physical benefits, research suggests that these interventions, particularly those incorporating reminiscence activities and fostering dyadic interactions between caregivers and individuals with MCI in both inpatient and home settings, can also enhance communication and social engagement [6]. Moreover, tablet-based interventions provide an engaging and enjoyable means of interaction for individuals with MCI, potentially strengthening social relationships [7].

Some interventions incorporate activity content such as photos, audio elements, and videos, which have been shown to facilitate memory activation and increase social engagement. However, research on the effects of technology-based interventions on caregiver-patient relationships and social interaction in individuals with cognitive impairments remains limited [6]. Nevertheless, existing evidence indicates a reciprocal relationship between the perceived burden of caregiving and the caregiver-patient relationship. While the ongoing cognitive decline of care recipients can heighten role strain for caregivers, dyadic relational resources, such as mutuality, have been shown to alleviate this burden [8].

To strengthen the dyadic relationship, it is essential to continuously adapt to changes in cognitive health and relational dynamics as the condition progresses [9]. This also applies to changes in communication and social interaction, which can be impaired in individuals with MCI and may affect their ability to participate socially and maintain relationships [10]. Given that cognitive decline in individuals with MCI can negatively impact their relationships with caregivers and potentially lead to interpersonal tensions [11], further research is needed to explore the positive effects of dyadic tablet-based interventions on individuals with MCI and their caregivers. In addition, involving both patients and caregivers in the design of digital applications is crucial for the early identification of design and usability challenges in technology intended for individuals with cognitive impairments [12].

Study Objectives and Research Questions

We designed this study based on the findings from an earlier investigation by Klein et al [13], in which a prototype of the intervention for cognitive and motor activation, as well as stimulation of social interaction, was tested with individuals with MCI. Insights from Klein et al [13] informed the further development of the touch-based multimedia tablet intervention, incorporating user feedback and refinements to enhance its usability and effectiveness.

Building on previous research, we introduced a dyadic approach involving individuals with MCI and their caregivers to examine how shared tablet-based interventions affect their interaction. By integrating elements of physical, social, and cognitive activation within the touch-based demonstrator, we sought to explore the feasibility of the intervention and the effects experienced by individuals with MCI and their caregivers with regard to their interaction and relationship.

In this study, we addressed the following research questions:

1. Caregiver-patient relationship—How does the use of the intervention influence the relationship between individuals with MCI and their caregivers from their perspectives? What relational dynamics become visible during its use, and how can the intervention be leveraged to foster positive interactions and strengthen the caregiver-patient bond?

XSL•FO
RenderX

2. Feasibility—To what extent is the implementation of a touch-based multimedia tablet intervention in a home environment practicable for individuals with MCI and their caregivers? What challenges and facilitators do they experience during its use?

Methods

Ethical Considerations

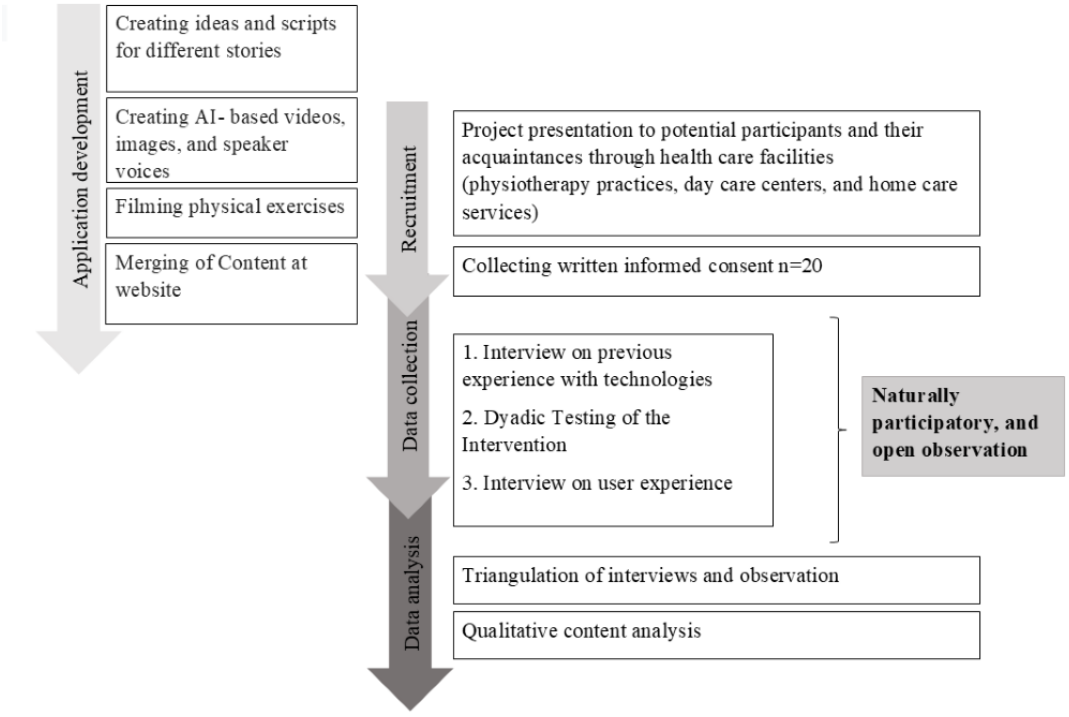
Following a thorough review process, this study received approval from the Ethics Committee of the University of Furtwangen (application number 24-052). It complies with the tenets of the Declaration of Helsinki. In accordance with institutional guidelines and ethical standards, all participants were assessed for their capacity to provide informed consent before enrollment. A letter containing pertinent information in an easily understandable form and a consent form were distributed to the participants in advance. Only individuals with MCI were included in the study; those with moderate or advanced dementia were excluded to ensure the ability of participants to provide informed and voluntary participation. The informed consent process was conducted with both the person with MCI and their caregiver present. This arrangement allowed for clarification of the study aims and procedures, and provided an additional opportunity to confirm the participant's comprehension and willingness to participate. Throughout the testing sessions, both members of the dyad were present, ensuring that the person with MCI was always accompanied by a familiar individual to promote comfort and emotional support during the intervention.

All participants were anonymized. No financial compensation was given to the participants.

Overview of This Study's Design

We used a qualitative triangulation approach, combining 20 naturalistic observations of participants and 20 interviews, with data from each method systematically aligned. The research was conducted in Germany and the reporting adheres to the Standards for Reporting Qualitative Research [14]. A summary of the application development and the study procedure is provided in Figure 1.

Figure 1. Application development and study procedure. AI: artificial intelligence.



Intervention Development Process

We drew upon the TiDIER (Template for Intervention Description and Replication) [15] checklist as a guiding framework to structure the description of the intervention (Table 1). This checklist provides a structured approach to outlining how interventions are designed, delivered, and evaluated, supporting transparency, reproducibility, and adaptability. Using these guidelines as a reference, we aimed to enhance the clarity and rigor of the intervention's description while allowing for flexibility in its adaptation to the specific research context. The intervention was collaboratively designed and developed by a multidisciplinary research team with expertise in health, computer science, and professional health care (including occupational therapists and physical therapists). This approach ensured that both technical and therapeutic aspects were thoroughly considered, resulting in a user-centered, holistic solution.

XSL•FO
RenderX

Table 1. Template for Intervention Description and Replication (TiDIER)–based description of the intervention.

Item	Description of item
Why (rationale)	• The intervention aimed to improve cognitive, physical, and social functioning in older adults with MCI[a] and their caregivers. This may include family members, friends, or professional care staff with whom these people with MCI interact on a regular basis.

What (materials and procedures)	• A tablet-based digital intervention was developed using a modular design for ease of use. The overarching design principle of simplicity in implementation was adopted to ensure accessibility and ease of use. To this end, a dedicated web platform was developed to support the design and delivery of the tablet intervention. This platform functioned as a modular toolkit, empowering researchers to assemble the stories by combining various media elements, including self-created recordings (images, sounds, and videos); royalty-free images; and sound files. In addition, AI[b]-generated avatars were facilitated through HeyGen and AI-generated speech through ElevenLabs. The interventions were administered through the website as well. The intervention involved the use of a tablet computer (Galaxy Tab A9+), which was provided by the study team. The tablet contained the demonstration tool. An introductory module was included to explain the website's operation and the nature of the tasks (cognitive, physical, and interactive), ensuring caregivers could effectively facilitate intervention. Participants then selected 1 of 4 stories to engage with. These stories included a city trip to Berlin, a visit to a farm, a trip to Spain, and a day in the forest. The explicit design of these stories aimed to address three core task areas:
	  • cognitive exercises with a focus on ADLs[c]—these exercises are designed to stimulate memory, attention, and problem-solving skills while maintaining a practical focus on ADL [5]. Examples include planning a travel day, shopping for groceries, and managing the process of paying (examples of the task types can be found in Multimedia Appendix 1).
	  • physical exercises related to ADL [6]—these exercises are carefully designed to promote movement and enhance functional independence. Examples include activities such as simulated stair climbing, exercises involving hand movement exercises, and relaxation exercises. These exercises are adapted to the participant's fitness level and performed either while seated or while standing.
	  • interaction exercises [7]—the objective of these exercises is to promote social engagement and the exchange of personal experiences, thereby fostering the relationship between people with MCI and their caregivers. These interactions are structured to incorporate biographical references, enabling participants to share meaningful memories and build connections during the intervention.
	• The selection of exercises and task areas was based on expert consensus [16] and interdisciplinary team input. The chosen story was played through once. Following the story, feedback questions with predefined answer options appeared on the screen, which the participants responded to.

Who provided (intervention delivery)	• The intervention was designed to be facilitated jointly by a person with MCI and their caregiver at home. The digital application guided them through the story-based sessions and interactive elements. Before implementation, participants received a structured introduction and training session to familiarize them with the application and its components, ensuring that they felt confident in using the intervention autonomously at home.
How (mode of delivery)	• The intervention was delivered through a tablet-based application, designed for independent, dyadic use at home. The tasks and stories were experienced via touchscreen interaction and audiovisual playback.
Where, when, and how much (setting and dose)	• The intervention was designed to be used in a home environment or any familiar setting of the participant's choice. Each session involved playing one story and engaging in associated tasks, with a total intervention duration of approximately 15 to 20 minutes. The intervention was applied in a single session, as this was a demonstration version. For future implementation, regular joint use on a weekly basis is recommended to support continuity, strengthen dyadic interaction, and enhance potential therapeutic effects over time.
Tailoring (customization)	• . Participants could choose between 4 stories and 2 levels of difficulty. Within each story, chapters could be skipped or selected freely to adjust the experience to individual preferences and needs. In addition, each story included movement exercises that could be performed at different levels of difficulty, either in a seated or standing position, to accommodate individual physical abilities and safety considerations.
Modifications	• No modifications were made during the delivery of the intervention.

[a]MCI: mild cognitive impairment.

[b]AI: artificial intelligence.

[c]ADL: activities of daily living.

Participant Selection and Recruitment Process

We conducted the intervention in German to ensure linguistic and cultural relevance, as the study took place in rural areas of southern Germany. Recruitment was conducted using a multipronged, community-based approach designed to ensure low-threshold access and preserve voluntariness throughout the process. We initially contacted local institutions (eg, day care

XSL•FO
RenderX

centers and physiotherapy practices) via email to introduce the study. After establishing contact, we visited selected facilities in person to present the project to staff and—in the case of day care programs—directly to older adults and their caregivers during regular group sessions. Recruitment in these settings always included an open invitation format, allowing participants to approach us voluntarily after the presentations for further questions or to express interest.

In addition, we actively recruited in dementia cafés and volunteer services. In the cafés, short presentations with a few Microsoft PowerPoint slides were integrated as program points into the regular schedule, usually in informal settings, such as coffee and cake gatherings. These events often attracted individuals with MCI and their family members together, creating a familiar and trusting atmosphere. We demonstrated small elements of the intervention to make it more tangible and understandable, encouraging attendees to decide freely whether they would like to participate. Flyers were distributed to supplement the verbal information.

Interested individuals could then approach us directly and schedule an individual appointment for testing the intervention. The timing and location of these sessions were arranged entirely based on the preferences and availability of both the person with MCI and the caregiver. This flexible and person-centered approach helped build trust and facilitated participation.

The study included participants aged 65 to 95 years with MCI, along with their caregivers. The term “caregiver” was defined as an individual who maintains a close and ongoing relationship with the person with MCI and provides assistance in daily activities, independent of their participation in the study. This could include spouses, family members, caregivers from day care facilities, friends, volunteers, or other relevant individuals who are regularly involved in the care and support of the person with MCI.

Individuals were excluded from participation if they (1) had severe cognitive impairment and were unable to provide informed consent; (2) were diagnosed with delirium, schizophrenia, or depression; (3) had a high risk of falling or significant mobility limitations; and (4) were living in residential care facilities.

Participants were recruited via local health care providers, home care services, and day care centers to ensure sample diversity. Dyadic appointments were arranged by personally contacting potential participants.

Data Collection

The data collection process was executed during the period spanning from September 2024 to November 2024. A total of 20 dyads participated in testing the tablet-based demonstrator and providing semistructured dyadic interviews [17].

The interviews were conducted either at participants’ homes or during their time in day care and in accordance with a pre-established, semistructured interview guideline (Multimedia Appendix 2), which provided a framework of deductive categories. The guideline was developed by the interdisciplinary research team and was informed by the overarching research questions, previous research on dyadic technology use in dementia care, and the team’s experience with related qualitative studies. It aimed to explore participants’ previous digital experiences; their handling of the intervention; perceived usability and usefulness of the different task types (cognitive, physical, and interactive); and any observed or experienced relational effects. This structure allowed for open-ended responses and follow-up questions based on participant input.

The first part of the interview focused on collecting demographic and personal data, including gender; age; ownership; use; and knowledge of devices such as tablets, PCs, and smartphones. After a brief introduction to the intervention, the caregiver and the respondent tested the tablet-based intervention together. As the last step, a second short interview took place to evaluate and discuss the feasibility of the tablet-based intervention. In addition, participants were asked about their perceptions of one another during the use of the intervention. The interviews were conducted in accordance with a pre-established interview guideline (Multimedia Appendix 2), which provided a structured framework of deductive categories. This process was recorded by audio recordings and naturalistic, participatory, and open observations, and took from 30 to 60 minutes for each dyad. Relevant interview topics were as follows: previous experience with digital technology, user skills, and personal characteristics of dyads; evaluation of tablet use; overall evaluation of the content; evaluation of different task types (physical, cognitive, and social); and shared use and influence on the relationship.

Furthermore, a naturalistic, participatory, and open observation was accomplished by 2 researchers, following an observation protocol that was created based on the deductive categories [18]. During the intervention sessions, one researcher assumed a dual role: one researcher facilitated the session by providing support to the dyads as needed (eg, with technical questions regarding the tablet application) and simultaneously took part in observing the interactions. The second researcher adopted a nonparticipant observer role, sitting slightly apart from the dyad and focusing exclusively on the documentation of field notes using a structured observation protocol.

The observation protocol was developed by the interdisciplinary study team based on previous experience in qualitative feasibility studies and informed by the relevant literature on dyadic interaction and intervention research in dementia and MCI contexts. The guiding questions served to focus the observations on key domains of interest, such as task feasibility, motivation, and dyadic communication, without applying standardized rating scales.

In the observation protocol, the place, time, and sequence of the observed processes were recorded. Also, it was distinguished between observations, contextual information, methodological and role reflections, and theoretical reflections. This distinction ensures that the observations could later be interpreted as such without being prematurely shaped by hasty theorizing [18]. The protocol supported a systematic but open-ended documentation of relevant behaviors, reactions, and interaction patterns, which were later analyzed qualitatively. Specifically, it consisted of 3 parts (Multimedia Appendix 3). One part of the observation protocol was dedicated to the observation of the care dyad. A

XSL•FO
RenderX

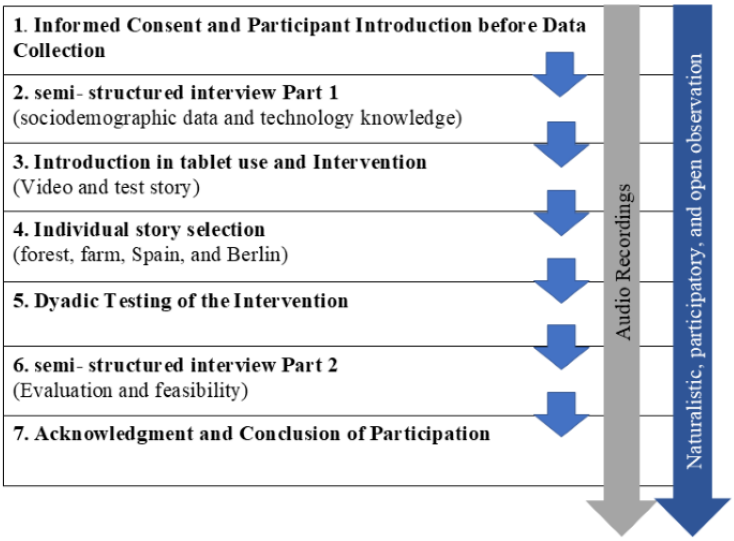
second part focused on the observation of the caregiver, while the third part entailed the observation of the communication dynamics within the dyad. For better comparability and assignability, guiding questions for each observation category were created. Parts 1 and 2 contained identical topics and guiding questions, whereas part 3 contained other topics that aimed to provide a comprehensive description of the communication dynamics within the dyad. The relevant topics and observation criteria are listed in Textbox 1.

The data collection process is summarized in Figure 2.

Textbox 1. Relevant categories and guiding questions of the observation protocol.

1. Person with MCI (overall feasibility, cognitive tasks, physical tasks, and interaction tasks). For example, how was the motivation during the exercise?
2. Caregiver (overall feasibility, cognitive tasks, physical tasks, and interaction tasks). For example, how was the reaction to the exercise? (positive or negative)
3. Care dyad: communication dynamics and relationship within the dyad (communication, mood, and change in mood; support; attitude of caregivers toward tablet computer; shared use; stimulus for conversation; and overall impression). For example, did the mood undergo any changes? To what extent did the caregiver's support meet the needs of the patient, and how well was it accepted?

Figure 2. Intervention and data collection process.



Data Analysis

The methodological approach followed a systematic and rigorous process to ensure the reliability and validity of the qualitative analysis. The first step involved the creation of a deductive category system based on the overarching themes of the research questions, which provided a structured framework for the analysis, along with a detailed coding guide to ensure consistency during the coding process. Audio recordings of interviews and observations were transcribed. The structured field notes during the observation underwent a double-validation process, whereby 2 researchers independently reviewed the observation data from each other and resolved any discrepancies through discussion. Following this, triangulation of the transcribed interview data and the validated observation protocols was carried out to align insights from multiple sources and enhance the depth of the analysis. The combination of methods allowed us to create results from different perspectives, particularly regarding relationship dynamics. The observations made it possible to include nonverbal communication patterns, power symmetries or asymmetries, and subtle interactions into the analysis.

The qualitative content analysis was conducted based on the methodological approach described by Kuckartz [19] and followed a structured, multistep process. After transcription, both coders independently conducted the initial coding of the interview and observation data using the MAXQDA software (Verbi GmbH), using a combination of deductive categories based on the research questions and inductively emerging subcategories from the material. Theme development was conducted collaboratively. The 2 coders jointly identified and refined preliminary themes through iterative discussions, comparing and synthesizing their interpretations. To ensure analytical rigor and increase trustworthiness, these themes were subsequently reviewed and discussed with additional members of the research team. Discrepancies were resolved through consensus, and categories were adjusted where necessary to better reflect the data. This reflexive and team-based approach aimed to ensure consistency and credibility throughout the analysis process.

Anchor quotations representative of key themes and findings were also identified to illustrate the findings.

XSL•FO
RenderX

Researcher Characteristics and Reflexivity

In line with the COREQ (Consolidated Criteria for Reporting Qualitative Research) criteria [20], the research team's backgrounds and positions in the field were considered throughout the research process to enhance transparency and reflexivity. The interdisciplinary study team consisted of health and social scientists with academic training and professional experience in gerontology, rehabilitation sciences, nursing, and qualitative research methods. All researchers had previous experience conducting studies with older adults, including individuals with cognitive impairments and their caregivers. They were trained in qualitative interviewing and observational techniques and were familiar with the ethical and communicative considerations required when working with cognitively vulnerable populations. To ensure reflexivity and reduce potential bias, the research team engaged in regular debriefing sessions in which preliminary impressions and interpretations were discussed throughout the data collection and analysis process. Throughout all interactions, researchers maintained a respectful, nondirective stance to minimize influence on participants' behavior and responses.

Results

Overview

Overall, we recruited 40 individuals, including persons with MCI (n=20, 50%) and their caregivers (n=20, 50%), to participate in our study. The level of experience in using technical devices was grouped into 4 categories: minimal experience, basic experience, moderate experience, and advanced experience. The description of the variables is provided in Table 2.

Table 2. Description of the participants' characteristics.

Variables	People with MCI[a] (n=20)	Caregiver (n=20)
Age (y), mean (SD)	80.95 (8.27)	60.63 (18.1)
Sex (female), n (%)	15 (75)	16 (80)
Previous technical experience, n (%)		
Minimal experience[b]	7 (35)	2 (10)
Basic experience[c]	7 (35)	2 (10)
Moderate experience[d]	2 (10)	1 (5)
Advanced experience[e]	4 (20)	15 (75)
Relationship, n (%)		
Spouse	—[f]	5 (25)
Child	—	4 (20)
Caregiver of a person attending a day care center	—	4 (20)
Friend	—	3 (15)
Volunteer	—	3 (15)
Other	—	1 (5)

[a]MCI: mild cognitive impairment.

[b]Minimal experience: use of simple, traditional technologies (eg, radio, television, and telephone) and have little or no experience with modern technology (eg, tablets and smartphones).

[c]Basic experience: limited access to or occasional use of modern technology; may own devices (eg, cell phone and tablet) but rarely use them or have had difficulty using them.

[d]Moderate experience: regular use of modern devices, but mainly for basic applications or with support from others.

[e]Advanced experience: in-depth experience and independent, either through intensive previous use or current skills; varied use of modern technology.

[f]Not applicable.

The results were organized along the 2 overarching research questions. On the basis of the qualitative data analysis, 5 main categories were developed deductively and supplemented by inductively derived subcategories. To enhance clarity and traceability, these categories were assigned to the respective research questions according to their thematic relevance:

1. Caregiver-patient relationship—relational dynamics and the influence of the intervention

1.1 Dyadic interaction and support

1.2 Impact of the intervention on communication and engagement

2. Feasibility—practicability, challenges, and conditions for successful implementation

2.1 Personal characteristics and experiences with digital technology

2.2 Feasibility, media, and task engagement

2.3 Future use and recommendations for the intervention

XSL•FO
RenderX

A visualization of this structure is provided in Table 3 to support the reader in contextualizing the categories within the scope of the research questions.

Table 3. Deductively developed main categories in relation to superordinate research questions.

Categories	Research question 1: caregiver-patient relationship	Research question 2: feasibility
1.1 Dyadic interaction and support	✓	—[a]
1.2. Impact of the intervention on communication and engagement,	✓	—
2.1 Personal characteristics and experiences with digital technology	Indirectly relevant	✓
2.2 Feasibility, media, and task engagement	—	✓
2.3 Future use and recommendations for the intervention	Implicit, relationship perspective	✓

[a]Not applicable.

Research Question 1: Caregiver-Patient Relationship—Relational Dynamics and the Influence of the Intervention

Dyadic Interaction and Support

The dynamics within dyads were characterized by mutual trust, encouragement, and humor, fostering a positive and engaged atmosphere. Praise and nonverbal reassurance, such as eye contact, helped maintain concentration:

> *Very good! Where did you learn that? So on with the story.* [Caregiver 4]

While most dyads communicated harmoniously, differences in communication styles emerged, particularly among married couples, where disagreements over pacing were observed. Some dyads adapted over time, shifting from skepticism to more relaxed interactions. Decision-making varied, with some caregivers assuming control (eg, “don’t keep pressing, I have to read first” [person with MCI 4]), while others facilitated joint decision-making, promoting autonomy. A key aspect was “getting to know each other better” through biographical storytelling (eg, “talking about Berlin or the farm—I think it’s great that you can talk about things you’ve experienced yourself” [caregiver 13]), shared memories, and collaborative tasks. Many dyads perceived the joint implementation as enriching (eg, “yes, that’s good. Two is always better than one [...] and one person can help the other a bit” [caregiver 4]). The medium itself was secondary to the interaction it facilitated, whether through a tablet or other activities. However, challenges such as impatience, paternalism, physical differences, or differing expectations could lead to tension.

Caregivers played a crucial role in technical and content-related support, adapting their assistance based on the person with MCI’s abilities. Some caregivers encouraged independence:

> *Or do you want to press? With your finger on it.* [Caregiver 14]

Other caregivers guided tablet use through instruction:

> *Exactly. Most of the time you have to use the fingertip.* [Caregiver 1]

As familiarity increased, caregiver intervention decreased, allowing for greater autonomy. However, excessive assistance risked creating dependency, limiting engagement:

Caregiver 5: “If you could choose now, where would you rather go?”

Person with MCI 5: “Spain.”

Caregiver 5: “Yeah. If you want, you can get up. But I don’t care. It’s up to you.”

Content-related support varied from cooperative problem-solving to task takeover. Some caregivers dominated tasks, limiting opportunities for independent learning (eg, “that brings us to 42” [caregiver 13]), while others facilitated responses through guidance and rephrasing. When caregivers provided encouragement and structured support, participation and self-confidence improved. Joint problem-solving fostered a sense of security and trust, allowing individuals with MCI to engage more actively:

> *You see, it’s very important that you’re there.* [Person with MCI 2]

Impact of the Intervention on Communication and Engagement

The intervention stimulated conversation, particularly through biographical topics and open-ended questions. References to personal experiences (eg, travel, markets, and gardens) encouraged storytelling, deep discussions, and caregiver engagement:

> *She [Avatar] asked what your best trip was? Do you want to tell?* [Caregiver 13]

Collaborative exercises enriched interactions, alternating between independent efforts and joint decision-making. Participants valued this approach, stating the following:

> *Oh yes, but rather together than alone.* [Caregiver 12]

> *It gives more motivation, and when you do it as a pair, where both can give answers, that’s nice.* [Person with MCI 10]

Laughter and shared experiences enhanced motivation, yet caregiver dominance sometimes led to answer anticipation,

XSL•FO
RenderX

impacting engagement. Differences of opinion occasionally arose but were mitigated by the structured research setting. While knowledge-based questions occasionally led to uncertainty and insecurity, the intervention proved effective in fostering dynamic exchanges, particularly when topics were personally relevant. In addition to self-reported experiences, observations during the joint use of the intervention revealed moments of closeness and shared engagement. These appeared socially meaningful, particularly when task complexity, autonomy, and role distribution were attuned to the dyad's needs.

Research Question 2: Feasibility—Practicability, Challenges, and Conditions for Successful Implementation

Personal Characteristics and Experiences With Digital Technology

Overview

Interview and observation analyses revealed that the personal backgrounds of individuals with MCI—shaped by their life experiences, social environments, and daily routines—strongly influenced their perspectives and engagement. Key themes emerged, including regional origins, cultural influences, former professions, and family roles. For example, one participant described how her career as a seamstress and raising her children structured her daily life. Health limitations, such as visual or hearing impairments, as well as technological barriers, also impacted their independence and activities.

Family support played a crucial role in overcoming daily technological challenges. One participant shared how her daughter introduced her to mobile phones:

> *Yes, and when my daughter got a new cell phone, she always gave me her old one and always said, do it, and I've been doing it for a few years. My grandson always helped me at the beginning when he was there, yes, then we write emails to each other and try to do something like that.* [Person with MCI 17]

Participants' experiences with technology varied widely, from complete unfamiliarity to routine use. Personal attitudes and biographical contexts influenced these differences. Some participants expressed skepticism, rejecting digital devices as unnecessary:

> *Well, I'm definitely not a computer freak.* [Person with MCI 8]

> *I have my TV, that's enough.* [Person with MCI 11]

However, others had extensive experience and integrated technology into their daily lives:

> *I did all my written work online. And then I have six grandchildren of all ages. And they keep me fit.* [Person with MCI 9]

Professional and family contexts played a crucial role in the development of technical skills.

Acceptance and Rejection of the Technology

Survey responses showed a complex picture of both acceptance and skepticism regarding the intervention. Some participants were enthusiastic, viewing the technology as stimulating and beneficial (eg, one participant stated that they "absolutely wanted to take part in the study" [person with MCI 17]). Key factors for acceptance included feelings of success, motivation, and interactive elements. Two participants were motivated by the desire to promote their own health through the movement exercises (eg, "it's healthy" [person with MCI 5] and "movement promotes blood circulation, and blood circulation in turn promotes the ability to think" [person with MCI 20]). Participants also highlighted the value of memory activation and cognitive stimulation:

> *It was fun to name the things that you know. Yes, then you're happy that you already know something and that you're somehow making progress.* [Person with MCI 17]

> *I'm interested in something like that, so that it sticks in my memory. I have a problem with my memory because the doctor said it wasn't very good.* [Person with MCI 18]

The social aspect contributed positively to acceptance, with some participants emphasizing that technology facilitated shared experiences:

> *It was nice to do it together.* [Caregiver 6]

However, technical challenges and design limitations posed barriers to acceptance. Some participants were skeptical about the relevance of the technology in their daily lives:

> *Do you really need it?* [Caregiver 1]

The success of the intervention depended on intuitive usability, cognitive stimulation, and interactive design, while obstacles included technical difficulties, content adaptation, and perceived relevance. In summary, acceptance was driven by positive user experiences, while rejection stemmed from usability concerns and doubts about the technology's practical value.

Content Aspects and Choice of Stories

Participants' topic preferences were strongly influenced by personal interests and past experiences. The "forest story" was the most frequently chosen, as the forest is a familiar and meaningful place for many. Familiarity with a topic often led participants to engage more, as they associated it with positive past experiences, such as family vacations. Conversely, a lack of personal connection acted as a barrier, leading some to reject certain topics:

> *Spain, that's not for us, we've never been there.* [Person with MCI 3]

However, for others, unfamiliar topics sparked curiosity and provided an opportunity to learn something new. The forest theme received broad acceptance and positive responses due to personal connections with nature:

> *We lived in the Black Forest for many years and spent a lot of time in the forest, and we still talk about those beautiful walks today.* [Person with MCI 3]

XSL•FO
RenderX

> *I am a nature lover through and through.* [Person with MCI 17]

Many recalled special memories, such as hiking, animal-watching, or past experiences in the military, particularly from a time when they were not yet physically restricted. The farm story also resonated well with participants, as it was both biographically relevant and generally familiar. Many had childhood memories or past visits to farms, while others found the rural lifestyle intriguing:

> *We have four sons, and when they were relatively young, we always spent our vacations on a farm.* [Caregiver 9]

The Spain story provided various points of reference, particularly for those with travel experiences or a general interest in the region:

> *I have already driven through Spain. As I said, above Coimbra, to Portugal, to Fatima, and then back via Madrid and back to Germany.* [Person with MCI 20]

The Spain story was often described as informative and entertaining, although some found it more challenging than other stories. The Berlin story triggered reflections on the city's history, especially regarding the Berlin Wall and the divided past of the city:

> *Berlin is beautiful, big, loud, all sorts of things, but um, the wall, it makes you thoughtful.* [Person with MCI 13]

Participants' interaction with the buttons varied widely. While some tapped independently, others were hesitant, required assistance, or sought clarification. The floating "Next" button caused occasional confusion, as its function was not immediately clear. Similarly, uncertainty arose with selection buttons under images, as some participants attempted to tap the images directly, leading to frustration when no response occurred. In addition, prolonged pauses between slides without interactive elements resulted in uncertainty about the next steps, with one participant explicitly asking for guidance:

> *So where do I press? It's a bit of uncharted territory, isn't it? You have to think about it.* [Person with MCI 3]

At the beginning of cognitive questions, answer options were initially hidden, leading to confusion that diminished once they became visible.

Initial difficulties in using the tablet were common, often due to unfamiliarity with the technology. Typing proved challenging for some, especially those with long fingernails or those who exerted excessive pressure on the touchscreen. While some participants remained hesitant at first, many gained confidence over time and eventually typed independently. Notably, participants with no previous tablet experience were initially skeptical but adapted quickly. One participant described typing as "new territory" but ultimately found it enjoyable. Despite early uncertainties, many reported increasing speed and confidence with continued use.

Feasibility, Media, and Task Engagement

The use of rewards, such as animal pictures after task completion, proved to be an effective emotional motivator, fostering joy and enthusiasm in both individuals with MCI and caregivers. Statements such as "once a lamb and once the dog. That builds you up" (person with MCI 2) illustrate how such elements encouraged engagement. The reward system not only increased motivation and attention but also enhanced social interaction within dyads:

Tom (avatar): "Well done. That was correct."

Caregiver 2: "The contestant has 100 points" (laughs).

Person with MCI 2: "Yes, yes. Now he's continuing. Look there the donkey said 'Bravo' That's funny."

Visual media, particularly nature-related images (eg, forest paths, flower meadows, streams, and animals), were consistently well received and described as visually appealing and emotionally uplifting:

> *They were beautiful, especially the forest and everything.* [Person with MCI 1]

Videos featuring landscapes and moving animals provided an immersive experience:

> *Seeing the landscape again, from a different perspective from above.* [Caregiver 1]

Narrator: "Ella watches as an orange butterfly lands on a blade of grass. To remember the moment, Ella takes another."

Person with MCI 1: "Oh, that's a beautiful fellow. I think I've barely seen any this year. Oh wait, a yellow one did fly by the other day."

Artificial intelligence (AI)–generated content (eg, images, speaker voices, and avatars) was widely accepted, seamlessly integrating into participants' experiences without compromising authenticity:

Interviewer: "Tom—this person—we created him using artificial intelligence. So it's not a real person, but someone we generated with a computer."

Caregiver 15: "Oh, but at the time, it wasn't really noticeable that it wasn't a real person."

Computer-generated voices and sound effects were generally perceived as clear and natural, although minor issues such as choppy audio or low volume were noted, but did not significantly affect comprehension. Background sounds, such as birdsong and flowing water, were praised for enhancing the atmosphere. A well-balanced combination of voices and ambient sounds was found to improve the overall user experience. Text-based and interactive elements received mixed feedback. While most found the font legible, some suggested increasing the font size, stating, "slightly larger" (dyad 17). Button placement and touch recognition issues were noted, with recommendations to space buttons further apart for better usability.

XSL•FO
RenderX

The range of tasks was generally well received, particularly movement-based activities, which were seen as beneficial for both physical health and cognitive activation:

> *Movement promotes blood circulation, which supports thinking.* [Person with MCI 20]

Personal anecdotes often accompanied exercises, for example, "feeding chickens" (farm story) or "running up the cathedral tower" (Spain story), enhancing engagement. Dyadic interaction influenced task acceptance, while relaxation and breathing exercises elicited mixed reactions—some found them calming, others reacted humorously. Interaction exercises encouraged quick responses and biographical storytelling (eg, farm visits, nature, and markets), effectively activating memories and fostering dialogue. Cognitive tasks were generally perceived as appropriate, although some participants, particularly those with relevant backgrounds, desired greater challenges. Many participants enjoyed problem-solving tasks, expressing enthusiasm through clapping, smiling, or laughing. Open-ended knowledge questions encouraged discussions and thought sharing, making them especially popular. The variety of question types was widely appreciated.

Future Use and Recommendations for the Intervention

The tablet-based intervention received positive feedback. One participant noted the following:

> *Time passed quickly, and it was a nice change.* [Caregiver 16]

Many participants were open to future use, even those with limited previous tablet experience:

> *Very good, it was very good. I'm surprised. She [Person with MCI] read and listened to everything, responded, and with such great interest. Better than usual.* [Caregiver 18]

One participant with MCI 3 expressed confidence:

> *I think I can do it now...What do you call it? Tablet? That's easy.*

In response, caregiver 5 added, "You can still learn to use a tablet at 92, right?" However, some saw limited personal use, mainly as an activity to share with grandchildren:

> *If my grandson suggested it, I'd join in.* [Caregiver 8]

Others viewed it as a one-time experience:

> *That was nice, but just once.* [Person with MCI 15]

Some rejected it outright:

> *I don't feel like it. Why should I learn this?* [Caregiver 5]
>
> *I have my TV, that's enough for me.* [Person with MCI 1]

Overall, willingness to use the intervention was mostly positive, but some participants preferred alternative health-promoting activities over learning new technology. Participants suggested adding more animals and matching sounds to the forest story, making background noise quieter to improve focus, and using more relevant feedback images. They recommended a larger touchscreen with bigger, clearly highlighted buttons and a more intuitive introduction with arrows pointing to where to press. Zooming options for images would help people with vision impairments. Shorter pauses between slides and longer stories (eg, the Berlin story was "too short") were also preferred. Participants favored focusing on one type of exercise (cognition, interaction, or movement) and wanted more feedback on quiz answers, as well as opportunities for independent writing tasks. A sing-along song for more interactivity was suggested. In terms of content, participants expressed interest in travel stories set in places such as Spain, the Netherlands, or local cities with landmarks, castles, and regional foods. They also wanted stories about trips to the sea, the mountains, the Camino de Santiago, or farm visits. Other ideas included traditional crafts, iconic routes such as Route 66, boat tours, market visits, and exploring local heritage and cuisine.

Discussion

Principal Findings

This study explored the feasibility of a dyadic, touch-based multimedia intervention for individuals with MCI and their caregivers in a home setting, focusing on its impact on caregiver-patient relationships and dyadic communication. The intervention, incorporating physical, cognitive, and interactive exercises, aimed to address cognitive and relational challenges through shared activities.

A key finding was the positive influence on caregiver-patient relationships. Shared activities facilitated connection, mutual support, and emotional intimacy, particularly through biographical elements, which sparked meaningful dialogue and strengthened bonds. Humor, encouragement, and collaborative problem-solving enhanced communication. Caregivers played an essential role, with varying levels of involvement. Some encouraged autonomy, while others were more controlling, occasionally limiting independence. Striking a balance in caregiver support is key to maintaining a collaborative dynamic.

The intervention was feasible for home use, with participants adapting well despite initial unfamiliarity. The integration of AI-driven visual and auditory elements was particularly promising, as most participants did not perceive these features as unnatural. However, challenges such as confusion with interactive buttons, slide transition delays, and difficulties for those with limited previous experience pointed to areas for improvement, such as an intuitive interface, larger buttons, and faster system responsiveness. Mixed reactions on future use highlighted a preference for customization. While many participants were enthusiastic about continuing, others preferred traditional activities, emphasizing the need for relevant and varied content. Feedback mechanisms, adaptive difficulty, and offline accessibility could enhance its appeal.

Comparison With Previous Work

According to our study, a dyadic, touch-based multimedia intervention for individuals with MCI and their caregivers in a home setting is feasible. Participants' difficulties mainly depended on their previous technical experience, with those having moderate or advanced experience finding the tablet easier

XSL•FO
RenderX

to use. Despite initial uncertainties, most individuals with MCI gained confidence over time. An introductory video facilitated familiarization with the tablet, supporting findings by Quintana et al [21]. However, unfamiliarity with touch surfaces and interactive fields created challenges, such as pressing too hard or not knowing where to press. These issues were exacerbated by inconsistent button positioning, suggesting the need for uniformity in future designs. The dyadic concept proved helpful, as caregivers provided support, enhancing interaction and helping overcome difficulties, as noted by Klein et al [13].

User acceptance was high, with participants describing the intervention as both enjoyable and cognitively engaging. Two levels of difficulty were implemented to accommodate varying conditions and daily performance, which the participants appreciated. Our findings support the importance of adapting multimedia applications to users' needs [16]. However, technical barriers, such as limited Wi-Fi access, posed challenges, leading to disruptions such as audio interruptions. These issues could be mitigated with an offline version, now available, which should be considered for future interventions. The perceived relevance of the technology played a crucial role in acceptance, as users who found the tablet relevant demonstrated higher levels of engagement. This point has been emphasized by Chen et al [7] and Thordardottir et al [22]. These findings can be better understood through the lens of the technology acceptance model [23], which posits that 2 key factors—perceived usefulness and perceived ease of use—determine users' acceptance of a technology. In our study, both aspects emerged as highly relevant: while many participants appreciated the intervention's cognitive and relational value (usefulness), others were discouraged by technical hurdles or limited familiarity with touch-based interfaces (ease of use). This aligns with previous findings [24,25] that emphasize the importance of intuitive design and contextual relevance in enhancing user engagement. Integrating the technology acceptance model in future evaluations may provide a more structured framework for understanding acceptance dynamics and guide further refinement of the intervention.

Content relevance also played a role, with participants preferring stories connected to their own experiences, such as those involving forests and farms. City stories were less engaging, which aligns with the findings of Klein et al [13] that content relevance impacts user acceptance. A participatory approach in design could further ensure that the content meets users' needs, as suggested by Holthe et al [26]. Our results indicate that targeted adaptations, involving both people with MCI and caregivers in the design process, are key for long-term acceptance and success.

The study highlights both positive impacts and challenges for the relationship between individuals with MCI and their caregivers. A key finding is that biographically oriented content can strengthen the caregiver-patient relationship by facilitating personal dialogues and shared reminiscence, which enhances emotional intimacy and mutual understanding. This finding aligns with that of Ryan et al [27], who showed that biographical content strengthens emotional bonds in tablet-based interventions. Similarly, Hoel et al [6] highlighted how technology can enhance social interaction and alleviate communication pressures. Klein et al [13] also noted that caregiver support is crucial for the success of tablet-based interventions.

Caregivers played a central role in guiding and encouraging active participation of individuals with MCI, reinforcing the importance of adaptive support as seen in the study by Ingebrand et al [28]. However, challenges arose when caregivers became overly controlling, limiting the autonomy of individuals with MCI. This finding aligns with that of Ingebrand et al [28], who emphasized the need for a balanced approach to caregiver involvement. Gilson et al [29] similarly pointed out that while technology-based interventions can strengthen emotional bonds, they also risk inducing frustration if caregivers overstep their supportive roles or if the technology is perceived as too complex.

Moreover, our study suggests that the home environment positively influenced the intervention's success, supporting the findings of Gilson et al [29], who showed that home-based interventions can yield favorable outcomes for individuals with dementia. Future research should focus on investigating long-term relational dynamics, optimizing caregiver involvement, and assessing the scalability of such interventions. Tailoring content to cultural and biographical contexts will be crucial for maintaining relevance and increasing acceptance. Future studies should focus on optimizing caregiver involvement, tailoring content to cultural and biographical contexts, and assessing scalability.

Strengths and Limitations

We used a qualitative research approach, which is particularly well suited for exploring lived experiences and subjective meanings in complex social contexts [30]. By implementing methodological triangulation, we combined different data sources and perspectives, increasing the robustness and validity of our findings. In addition, the study explicitly addressed the dyad as a unit, acknowledging the dynamic interaction and relational dynamics between individuals with MCI and their caregivers. This holistic approach reflected the real-life caregiving context, ensuring that both social and cognitive aspects were considered in evaluating the intervention's impact. The study also benefited from a diverse recruitment strategy, including participants from various caregiving settings, such as home care, daycare facilities, and physiotherapy practices, which enhanced the practical relevance of the findings. Furthermore, the integration of observational data alongside interviews provided deeper insights into participants' actual behavior, reducing the risk of self-report bias.

Despite these strengths, several limitations must be acknowledged. One limitation was the lack of internet access among a substantial portion of the target population, posing challenges as the intervention was exclusively online. However, the dependence on technology may have posed a barrier for individuals less familiar with digital tools. Another potential limitation was social desirability bias, as participants may have provided overly positive feedback during interviews. This risk was mitigated by observations and researcher cross-validation. Furthermore, the study provides a short-term perspective on the intervention, meaning that insights into long-term engagement

XSL•FO
RenderX

and effects remain limited. Future research should explore how sustained use influences outcomes over time.

Overall, this study offers valuable insights into dyadic interactions in a digital intervention while highlighting both opportunities and challenges in its implementation. While this feasibility study was not designed to evaluate relational outcomes, the data suggest that shared use of the intervention may open opportunities for meaningful dyadic interaction. Future research should further explore the potential relational impact in a targeted design.

Conclusions

This study evaluated the feasibility of a dyadic, tablet-based multimedia intervention for individuals with MCI and their caregivers in a home setting, with a focus on its impact on caregiver-patient relationships. The findings indicate that the intervention was feasible, with initial uncertainties mitigated by thematic introductions and caregiver support, even for participants with limited technical experience.

A key factor in user acceptance and feasibility was the relevance of the intervention to participants' daily lives. One of the study's strengths was the implementation of 2 difficulty levels, allowing for adaptability to the fluctuating cognitive performance of individuals with MCI. Participants generally provided positive feedback, partly attributing benefits to cognitive and physical engagement. Our findings also underscore both the potential and challenges of using technology to support caregiver-patient relationships. Biographically oriented content appeared to facilitate social interaction and strengthen bonds. Furthermore, the integration of AI-generated visual and auditory elements was well received, as most participants did not perceive them as unnatural. This suggests that AI-based storytelling can enhance engagement and provide more personalized content. In future interventions, AI could be used to generate individualized, person-centered narratives that spark meaningful conversations within the dyad, further strengthening relationships and improving communication.

However, as this was a cross-sectional study, no conclusions about long-term effects can be drawn. Future research should prioritize longitudinal studies to assess sustained relational impacts and explore strategies to optimize caregiver involvement.

Acknowledgments

This study was funded by the German Federal Ministry of Education and Research (BMBF) as part of the KoMoS (Kognitiv-motorisches Spiel [cognitive-movement game]) project (reference 16SV9255) within the startInteraktiv (START-interaktiv [START-interactive]) research program.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Examples of the task types within the demonstrator.
[DOCX File , 10438 KB - aging_v8i1e75189_app1.docx]

Multimedia Appendix 2
Guidelines for the dyadic interviews—English version.
[DOCX File , 19 KB - aging_v8i1e75189_app2.docx]

Multimedia Appendix 3
Observation protocol—English version.
[DOCX File , 21 KB - aging_v8i1e75189_app3.docx]

References

1. World Alzheimer Report 2018: the state of the art of dementia research: new frontiers. Alzheimer's Disease International. 2018. URL: https://www.alzint.org/u/WorldAlzheimerReport2018.pdf [accessed 2025-04-03]
2. Seidl U, Ahlsdorf E, Toro P. Die leichte kognitive Beeinträchtigung: epidemiologische, neuropsychologische und neurobiologische Aspekte. J Prev Med 2007;3:286-293 [FREE Full text]
3. Petersen RC, Lopez O, Armstrong MJ, Getchius TS, Ganguli M, Gloss D, et al. Practice guideline update summary: mild cognitive impairment [RETIRED]: report of the guideline development, dissemination, and implementation subcommittee of the American Academy of Neurology. Neurology 2018 Jan 16;90(3):126-135 [FREE Full text] [doi: 10.1212/WNL.0000000000004826] [Medline: 29282327]
4. Lee-Cheong S, Amanullah S, Jardine M. New assistive technologies in dementia and mild cognitive impairment care: a PubMed review. Asian J Psychiatr 2022 Jul;73:103135. [doi: 10.1016/j.ajp.2022.103135] [Medline: 35569363]

XSL•FO
RenderX

5. Ortega Morán JF, Pagador JB, Gilete Preciado V, Moyano-Cuevas JL, Rodríguez Domínguez T, Santurino Muñoz M, et al. A serious game for cognitive stimulation of older people with mild cognitive impairment: design and pilot usability study. JMIR Aging 2024 Apr 04;7:e41437 [FREE Full text] [doi: 10.2196/41437] [Medline: 38596860]
6. Hoel V, Feunou CM, Wolf-Ostermann K. Technology-driven solutions to prompt conversation, aid communication and support interaction for people with dementia and their caregivers: a systematic literature review. BMC Geriatr 2021 Mar 04;21(1):157 [FREE Full text] [doi: 10.1186/s12877-021-02105-0] [Medline: 33663401]
7. Chen K, Lou VW, Lo SS. Exploring the acceptance of tablets usage for cognitive training among older people with cognitive impairments: a mixed-methods study. Appl Ergon 2021 May;93:103381. [doi: 10.1016/j.apergo.2021.103381] [Medline: 33578065]
8. Yang CT, Liu HY, Shyu YI. Dyadic relational resources and role strain in family caregivers of persons living with dementia at home: a cross-sectional survey. Int J Nurs Stud 2014 Apr;51(4):593-602. [doi: 10.1016/j.ijnurstu.2013.09.001] [Medline: 24083977]
9. Conway ER, Watson B, Tatangelo G, McCabe M. Is it all bleak? A systematic review of factors contributing to relationship change in dementia. Int Psychogeriatr 2018 Nov;30(11):1619-1637 [FREE Full text] [doi: 10.1017/S1041610218000303] [Medline: 29667571]
10. Sobral AI, de Araújo CM, Sobral MF. Mild cognitive impairment in the elderly: relationship between communication and functional capacity. Dement Neuropsychol 2018 Jun;12(2):165-172 [FREE Full text] [doi: 10.1590/1980-57642018dn12-020009] [Medline: 29988360]
11. Connors MH, Seeher K, Teixeira-Pinto A, Woodward M, Ames D, Brodaty H. Mild cognitive impairment and caregiver burden: a 3-year-longitudinal study. Am J Geriatr Psychiatry 2019 Nov;27(11):1206-1215. [doi: 10.1016/j.jagp.2019.05.012] [Medline: 31230914]
12. Fox S, Brown LJ, Antrobus S, Brough D, Drake RJ, Jury F, et al. Co-design of a smartphone app for people living with dementia by applying agile, iterative co-design principles: development and usability study. JMIR Mhealth Uhealth 2022 Jan 14;10(1):e24483 [FREE Full text] [doi: 10.2196/24483] [Medline: 35029539]
13. Klein M, von Bosse A, Kunze C. The needs and experiences of people with early-stage dementia using an application for cognitive and physical activation in Germany: qualitative study. JMIR Aging 2024 Dec 10;7:e62689 [FREE Full text] [doi: 10.2196/62689] [Medline: 39656992]
14. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251 [FREE Full text] [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
15. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014 Mar 07;348(mar07 3):g1687 [FREE Full text] [doi: 10.1136/bmj.g1687] [Medline: 24609605]
16. von Bosse A, Bejan A, Wessel M, Kunze C. Gestaltungs- und Evaluationsmöglichkeiten von Multimediaanwendungen zur Unterstützung der Betreuung von Menschen mit leichter Demenz [Design and evaluation options of multimedia applications to support the care of people with mild dementia]. Z Gerontol Geriatr 2024 Aug;57(5):382-388. [doi: 10.1007/s00391-023-02280-2] [Medline: 38265455]
17. DeJonckheere M, Vaughn LM. Semistructured interviewing in primary care research: a balance of relationship and rigour. Fam Med Community Health 2019 Mar 08;7(2):e000057 [FREE Full text] [doi: 10.1136/fmch-2018-000057] [Medline: 32148704]
18. Przyborski A, Wohlrab-Sahr M. Qualitative Sozialforschung: Ein Arbeitsbuch. Berlin, Germany: De Gruyter Brill; 2021.
19. Kuckartz U. Qualitative Text Analysis: A Guide to Methods, Practice and Using Software. Thousand Oaks, CA: SAGE Publications; 2014.
20. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
21. Quintana M, Anderberg P, Sanmartin Berglund J, Frögren J, Cano N, Cellek S, et al. Feasibility-usability study of a tablet app adapted specifically for persons with cognitive impairment-SMART4MD (support monitoring and reminder technology for mild dementia). Int J Environ Res Public Health 2020 Sep 18;17(18):6816 [FREE Full text] [doi: 10.3390/ijerph17186816] [Medline: 32961993]
22. Thordardottir B, Malmgren Fänge A, Lethin C, Rodriguez Gatta D, Chiatti C. Acceptance and use of innovative assistive technologies among people with cognitive impairment and their caregivers: a systematic review. Biomed Res Int 2019 Mar 06;2019:9196729 [FREE Full text] [doi: 10.1155/2019/9196729] [Medline: 30956989]
23. Davis FD. A technology acceptance model for empirically testing new end-user information systems: theory and results. Massachusetts Institute of Technology, Sloan School of Management. 1986. URL: https://dspace.mit.edu/handle/1721.1/15192 [accessed 2025-08-03]
24. Paneru CP, Tarigan AK, To a C. Exploring user engagement and perceptions towards smart energy applications in Norway: insights from a text mining analysis of user reviews. Energy Effic 2025 Mar 17;18(4):25. [doi: 10.1007/s12053-025-10309-1]

XSL•FO
RenderX

25. Omaghomi TT, Elufioye OA, Akomolafe O, Anyanwu EC, Daraojimba AI. Health apps and patient engagement: a review of effectiveness and user experience. World J Adv Res Rev 2024;21(2):432-440 [FREE Full text] [doi: 10.30574/wjarr.2024.21.2.0476]
26. Holthe T, Jentoft R, Arntzen C, Thorsen K. Benefits and burdens: family caregivers' experiences of assistive technology (AT) in everyday life with persons with young-onset dementia (YOD). Disabil Rehabil Assist Technol 2018 Nov 11;13(8):754-762. [doi: 10.1080/17483107.2017.1373151] [Medline: 28891356]
27. Ryan AA, McCauley CO, Laird EA, Gibson A, Mulvenna MD, Bond R, et al. 'There is still so much inside': the impact of personalised reminiscence, facilitated by a tablet device, on people living with mild to moderate dementia and their family carers. Dementia (London) 2020 May;19(4):1131-1150. [doi: 10.1177/1471301218795242] [Medline: 30193535]
28. Ingebrand E, Samuelsson C, Hydén LC. Supporting people living with dementia in novel joint activities: managing tablet computers. J Aging Stud 2023 Jun;65:101116 [FREE Full text] [doi: 10.1016/j.jaging.2023.101116] [Medline: 37268389]
29. Gilson A, Dodds D, Kaur A, Potteiger M, Ford Ii JH. Using computer tablets to improve moods for older adults with dementia and interactions with their caregivers: pilot intervention study. JMIR Form Res 2019 Sep 03;3(3):e14530 [FREE Full text] [doi: 10.2196/14530] [Medline: 31482847]
30. Flick U. An Introduction to Qualitative Research. London, UK: SAGE Publications; 2014.

Abbreviations

ADL: activities of daily living
AI: artificial intelligence
COREQ: Consolidated Criteria for Reporting Qualitative Research
MCI: mild cognitive impairment
TiDIER: Template for Intervention Description and Replication

Edited by M Gray; submitted 01.04.25; peer-reviewed by K Seibert, K Meyer; comments to author 24.06.25; revised version received 07.07.25; accepted 08.07.25; published 28.08.25.

Please cite as:
von Bosse A, Thum A, Mahler S, Schulz H, Weuthen P, Kunze C
Evaluating the Feasibility of a Dyadic, Touch-Based Multimedia Tablet Intervention and Its Effects on the Caregiver-Patient Relationship Among Individuals With Mild Cognitive Impairment: Qualitative Triangulation Study
JMIR Aging 2025;8:e75189
URL: https://aging.jmir.org/2025/1/e75189
doi:10.2196/75189
PMID:

©Alexa von Bosse, Anna Thum, Sonja Mahler, Henriette Schulz, Paulina Weuthen, Christophe Kunze. Originally published in JMIR Aging (https://aging.jmir.org), 28.08.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Design and Evaluation of a Mobile App for Intergenerational Communication: User-Centered Participatory Design and Experimental Mixed Methods Study

Soondool Chung[1*], PhD; Hannah Lee[2*], PhD; Jeehye Jung[1], MA

[1]Department of Social Welfare, Ewha Womans University, Seoul, Republic of Korea

[2]Ewha Institute of Intergeneration for Age Integration Research, Ewha Womans University, Seoul, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Hannah Lee, PhD
Ewha Institute of Intergeneration for Age Integration Research
Ewha Womans University
SK Telecom Building, Room 504-1, 5th Floor, 52, Ewhayeodae-gil, Seodaemun-gu
Seoul, 03760
Republic of Korea
Phone: 82 10 3038 0831
Email: hannah21@ewha.ac.kr

Abstract

Background: Social isolation and weakened intergenerational ties pose significant threats to the emotional well-being and social support networks of older adults. Although structured intergenerational programs can reduce age-related stereotypes and promote connectedness, their accessibility is often hindered by physical and logistical constraints. The increasing digital literacy among older populations presents new opportunities for technology-based interventions to support meaningful cross-generational engagement.

Objective: This study aimed to design and evaluate a mobile app that fosters intergenerational communication and enhances perceived social support in older adults using a user-centered design framework grounded in the double diamond model.

Methods: The development process followed the 4 phases of the double diamond model. In the discover phase, surveys with older and younger adults identified distinct usability preferences. The define phase synthesized these insights into key design principles. In the develop phase, a prototype was created and iteratively refined through usability testing. Finally, in the deliver phase, a 2-week experimental study involving 39 participants (20 older adults aged 68-82 years and 19 younger adults aged 22-39 years) assessed changes in intergenerational interaction, perceived social support, and user satisfaction.

Results: The app appeared to enhance intergenerational communication and perceived social support, particularly among older participants. Users reported increased comfort and emotional connection in cross-generational conversations. Accessibility features and engaging content were noted as contributing to positive user experiences across age groups.

Conclusions: This study suggests the potential of user-centered digital platforms to promote social well-being among older adults. By addressing the unique needs of multiple generations, such interventions may help foster inclusive digital environments and contribute to age-friendly, connected societies. Despite limitations related to sample size, duration, and cultural context, the study provides preliminary evidence for the potential of co-designed digital tools in supporting intergenerational communication and aging-in-place.

(JMIR Aging 2025;8:e75950) doi:10.2196/75950

KEYWORDS

intergenerational communication; digital inclusion; user-centered design; mobile app development; technology acceptance model; intergenerational programs; social connectedness

XSL•FO
RenderX

Introduction

Intergenerational conflict has emerged as a critical social issue, undermining societal cohesion beyond mere differences in opinions between older and younger generations. Such conflicts manifest as misunderstandings, stereotypes, discriminatory attitudes, and competition for resources, resulting in tensions that hinder effective intergenerational collaboration and societal harmony [1,2]. This phenomenon is prevalent across various cultural contexts and is often exacerbated by the increasing generational divide, which limits opportunities for mutual understanding and interaction [2,3]. According to social identity theory [4], individuals tend to categorize themselves into in-groups and out-groups, fostering biases and discriminatory attitudes toward those outside their group. The generational divide intensifies such perceptions, leading to heightened tensions and societal fragmentation.

In preindustrial societies, intergenerational communication was inherently facilitated by extended family structures and community-based living arrangements, where multiple generations coexisted and engaged in shared responsibilities. However, the shift toward nuclear family living arrangements and the advent of digital media have significantly reduced these opportunities, reinforcing psychological and physical distances between age groups. This decline in direct contact between age groups has contributed to social alienation and hindered the formation of intergenerational trust. According to the contact hypothesis [5], sustained intergroup contact reduces prejudice and promotes positive relationships. This has been empirically demonstrated in structured intergenerational programs, which consistently reduce stereotypes and foster mutual respect [6,7].

To address the widening generational divide, intergenerational programs have been widely implemented to facilitate direct engagement between older and younger generations. Such programs yield mutual benefits; older adults experience enhanced emotional well-being, cognitive stimulation, and increased social participation, while younger participants benefit from identity formation and skill acquisition [8]. Despite their well-documented advantages [9], face-to-face engagement remains constrained by physical distance, mobility limitations, and logistical barriers [10,11].

Digital communication tools offer a scalable and accessible alternative, particularly for populations with restricted opportunities for in-person interaction. The COVID-19 pandemic further accelerated digital adoption among older adults; in Germany, for instance, internet use among older individuals increased from 58% to 81% during the pandemic [12]. Moreover, empirical evidence suggests that even indirect or extended intergenerational contact can reduce age-based prejudice among younger people [7], underscoring the potential of digital platforms to foster intergenerational understanding. These findings highlight the value of digital environments as a means of promoting meaningful engagement across generations, extending the reach and impact of conventional face-to-face programs.

In line with this shift, various digital interventions have emerged globally to enhance intergenerational engagement. For example, Müller et al [13] developed a web-based platform to improve communication between older adults and their families through shared modules and collaborative learning. Similarly, Phang et al [14] identified diverse programs using accessible technologies—such as smartphones, video calls, and messaging apps—to strengthen emotional connection and intergenerational bonding. However, most existing platforms lack mechanisms for sustained, bidirectional interaction. Although older adults' digital literacy has increased [14], many platforms still fall short of addressing the nuanced needs and co-design preferences of both age groups. This gap underscores the need for inclusive, user-centered digital solutions that enable emotionally reciprocal and goal-oriented communication across generations.

In response to these limitations, this study introduces CoGen (Ewha Institute for Age Integration Research), a mobile app designed to promote intergenerational communication. Unlike prior platforms that rely on structured prompts or 1-way content delivery, CoGen supports user-generated, bidirectional interaction designed to accommodate the nuanced needs and co-design preferences of both younger and older users. To achieve this, CoGen's development is anchored in user-centered design (UCD) principles [15,16] and operationalized through the double diamond model [17], a 4-phase process (discover, define, develop, and deliver) that facilitates a structured, iterative approach for identifying user needs and refining design solutions in response to ongoing feedback. Implicit in this design philosophy is the contact hypothesis [5], which posits that sustained interaction under conditions, such as equal status, common goals, intergroup cooperation, and institutional support, can reduce prejudice. While this study does not directly test the hypothesis, its framework serves as a useful theoretical lens for understanding how intergenerational communication—especially in structured digital settings—might contribute to prejudice reduction.

Rather than simply developing a communication tool, this study aims to create CoGen as a user-centered, age-inclusive digital environment that enables users to exchange experiences, share cultural knowledge, and engage in collaborative activities. By fostering sustained and meaningful digital interaction, the project seeks to bridge the intergenerational divide, enhance older adults' social connectedness, and support global efforts toward digitally inclusive aging. In doing so, the development and evaluation of CoGen align with international movements advocating for scalable, human-centered technologies that empower older populations and promote equitable participation.

Methods

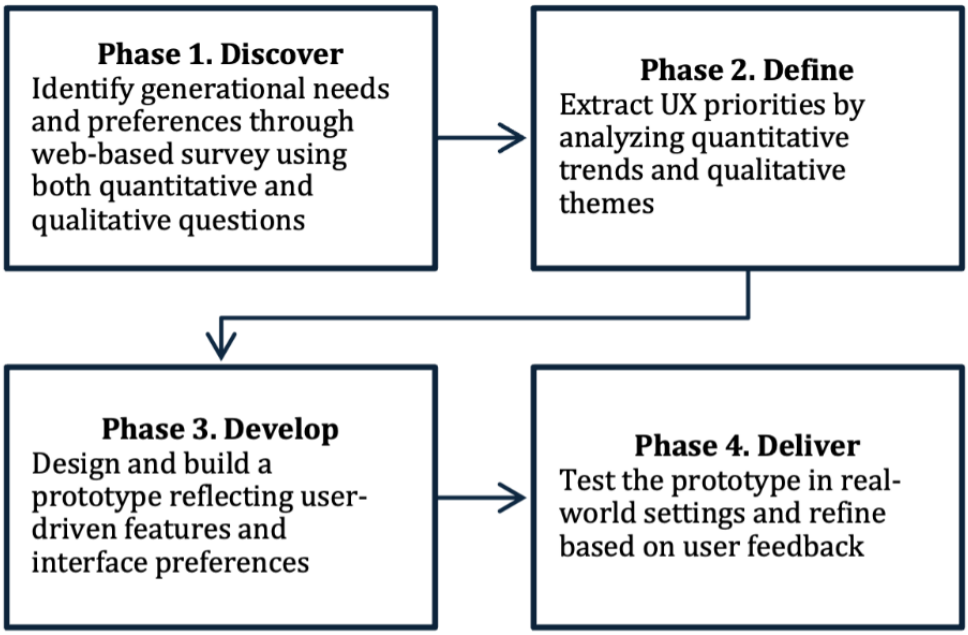
Double Diamond Model and Research Process

To guide the design and evaluation of the proposed mobile app, this study adopts the double diamond model [17], a 4-phase framework comprising discover, define, develop, and deliver. Alternating between divergent and convergent thinking, this model enabled a user-centered, iterative approach to address challenges in intergenerational digital communication. Each phase incorporated continuous input from both younger and older users to ensure that the final design was both accessible and responsive to real-world needs. Figure 1 presents a visual

XSL•FO
RenderX

summary of the study design, showing the sequential flow of the 4 phases along with key methods and objectives associated with each phase.

Figure 1. User-centered app development for intergenerational communication using the double diamond model. UX: user experience.



Phase 1: Discover: User Research

Research Focus

In the discover phase, the technology acceptance model [18] guided the identification of key factors influencing user attitudes and adoption intentions across age groups. A cohort-specific survey approach was used to capture generational differences in communication needs, motivations, and feature preferences. By incorporating both usability concerns of older adults [19] and engagement drivers for younger users [20], this phase laid the groundwork for a user-centered app design.

Participants and Research Procedure

Recognizing the distinct usability barriers faced by older adults and the engagement drivers influencing younger adults, this study recruited 2 age groups: younger adults (20-39 years) and older adults (60-79 years). A nationwide web-based survey was conducted in South Korea, using proportional stratified random sampling from an existing panel database managed by a professional research firm. Stratification criteria included age, gender, and prior digital communication experience to ensure demographic diversity within each group. Email invitations outlining the study's purpose were sent, and a total of 413 participants—202 younger adults and 211 older adults—voluntarily completed the survey.

Measures and Data Collection

Measurement Approach

To assess participants' attitudes, expectations, and key adoption factors for an intergenerational communication app, we used a mixed methods approach with quantitative survey items and qualitative open-ended responses. Our measurement strategy primarily used single-item indicators, carefully adapted from established multi-item scales for their high face validity and relevance [21-27]. This approach prioritized feature-level granularity over composite scoring, aligning with our exploratory UCD goals. Furthermore, we followed prior methodological support for single-item measures when constructs are narrow, concrete, and easily understood [28-30]. As these items were not intended to form reflective scales, internal consistency metrics like Cronbach α were not computed.

Attitudes Toward Intergenerational Communication

Participants' overall attitudes toward intergenerational communication were assessed using an adapted version of the scale by Lee and Lee [21]. Two conceptually distinct items measured interest and perceived importance (eg, "How interested are you in communicating with other generations?" and "How important do you think it is to have communication between generations?"), rated on a 5-point Likert scale (1=not at all and 5=very much).

Perceived Necessity and Willingness to Use an Intergenerational Communication App

To assess both the perceived necessity of an intergenerational communication app and participants' willingness to use it, two key items were included: (1) "How necessary do you think an app for intergenerational communication is?" (1=not necessary at all and 5=very necessary). (2) "If an app for intergenerational communication were developed, how often do you think you would use it?" (1=not at all and 5=very often). These items aimed to evaluate both the demand for such an app and participants' likelihood of adoption.

Motivational Factors for App Use

This section examined the anticipated use cases of the app, including social communication (eg, maintaining relationships across generations), entertainment (eg, engaging in shared recreational activities), information exchange (eg, learning about generational perspectives), and emotional support (eg,

XSL•FO
RenderX

alleviating social isolation). Participants rated 8 motivational factors on a 5-point Likert scale (1=not relevant at all and 5=very relevant) to assess their perceived significance. These factors included social communication, passing time (relieving boredom), entertainment and enjoyment, exchanging necessary information or opinions, reducing loneliness, health management, discussing concerns not easily shared with same-generation peers, and cultural exchange. The items were adapted from Lee and Na [22] and Jun et al [23] to align with prior research on intergenerational communication and digital engagement.

Preferred App Features

Guided by Davis's [18] technology acceptance model, participants evaluated both content and design attributes likely to influence perceived ease of use and perceived usefulness of the app. For perceived ease of use, 7 items were assessed (eg, ease of access to content, clarity and accuracy of information, and simplicity of design). For perceived usefulness, 9 items were included (eg, convenience, sense of belonging, emotional satisfaction, and health relevance). These items were adapted from prior studies on digital content evaluation and intergenerational engagement [24,25].

Aesthetic and Gamification Components

Six aesthetic and gamification components were assessed to evaluate their perceived importance in enhancing user engagement. These included rewards (eg, points, badges, and levels), customization options (eg, avatars and profiles), friend integration (eg, KakaoTalk and Facebook), immersive meeting spaces (eg, augmented reality or virtual reality), visual appeal (eg, icons and backgrounds), and auditory appeal (eg, sounds and music). All items were rated on a 5-point importance scale (1=not important at all and 5=very important). The items were adapted from Choi et al [26] and Kim and Kim [27] to ensure consistency with prior research on digital engagement and user experience (UX) design.

Qualitative Responses on Cross-Generational Information and Conversation Topics

In addition to structured survey items, participants provided open-ended responses about specific topics and types of information they believed would best facilitate meaningful intergenerational interactions. These qualitative insights supplemented the quantitative findings, ensuring that the app's design aligns with real-world conversational preferences.

Statistical and Thematic Analysis

Quantitative data were analyzed using descriptive statistics and independent-samples 2-tailed *t* tests to examine generational differences across key variables, including perceived necessity, willingness to use the app, motivational factors, and preferences for content and design features. Mean comparisons allowed for the identification of statistically significant age-related patterns in user needs and expectations.

In parallel, qualitative data from open-ended survey responses were analyzed using inductive thematic analysis. One researcher (JJ) conducted an initial review of all responses to identify semantically similar ideas and organized them into preliminary categories. These were grouped into broader conceptual themes through a data-driven approach—for example, responses referencing real estate, stock investments, and financial planning were grouped under "finance," while those related to exercise, dieting, and eating habits were categorized under "health." To enhance the credibility of the analysis, a second researcher (HL) independently reviewed the categorized data and cross-checked the thematic structure. Discrepancies in categorization were discussed and resolved collaboratively.

Phase 2: Define: UX Design Extraction Through Data Analytics

In the define phase, findings from the quantitative and qualitative analyses conducted in the discover phase were synthesized to establish key UX objectives and design requirements. This phase focused on translating user insights into actionable design principles, aligning core functionalities and content strategies with generational communication patterns and technological expectations. No new analyses were conducted; instead, this phase served as a bridge between empirical findings and UCD decisions.

Phase 3: Develop: Prototype Development and UX Optimization

Design Objectives

This phase focused on translating key insights from the define phase into a functional prototype that supports intergenerational communication through an intuitive, engaging, and accessible design. The primary goal was to develop an interface that balances structured discussions with active engagement while ensuring seamless navigation for both younger and older users. The prototype underwent preliminary usability testing to validate design decisions and inform further refinement.

User Interface or UX Design and Wireframing Strategy

In the develop phase, findings from the define stage were translated into concrete interface components through low- and high-fidelity prototyping. Design decisions were guided by age-specific user preferences identified in previous phases. For older adults, design emphasis was placed on simplicity, clarity, and visual accessibility, including large font sizes and high-contrast modes. For younger adults, vibrant color palettes and dynamic layouts were incorporated to enhance visual engagement and usability.

Wireframing began with a streamlined home screen comprising 5 core functions to reduce cognitive load while maintaining structural clarity. To support long-term engagement, visual elements such as level indicators and progress bars were introduced, emphasizing intuitive feedback mechanisms.

Prototype Feature Integration Plan

Based on the motivational and usability patterns revealed in the discover and define phases, a set of core features was developed. These included theme-based discussion boards, digital conversation prompts, proprietary 1:1 friend matching, and progression-based gamification. Design mapping was used to align each feature with specific user needs—for example, older adults' emphasis on meaningful social bonding informed the

XSL•FO
RenderX

inclusion of noncompetitive reward systems, while younger adults' preference for entertainment-driven experiences led to the development of dynamic prompts and visual engagement indicators.

Preliminary Usability Testing Procedure

Before broader evaluation, a pilot usability test was conducted to evaluate the prototype's interface, features, and overall usability. In total, 8 participants—5 younger adults (aged 20-39 years) and 3 older adults (aged 60-79 years)—were recruited to test core functions including onboarding, feature navigation, and emotional appeal. Feedback was collected through structured observation and posttask interviews. Key points of friction and areas of positive engagement were identified to guide iterative improvements before moving into the final testing phase.

Phase 4: Deliver: Usability Testing and User-Centered Refinement of CoGen

Phase Overview

The deliver phase aimed to evaluate the CoGen prototype's effectiveness in fostering intergenerational engagement in real-world settings. A 2-week experimental study using a mixed methods approach was conducted to assess UX, communication outcomes, and app usability.

Study Design and Procedures

A total of 39 participants (17 men and 22 women) aged 20 to 82 (mean 48.57, SD 24.56) years took part in a 2-week experimental study conducted from October 28 to November 12, 2024. The sample included 20 older adults (aged 68-82 years) recruited from 2 senior welfare centers in Seoul and 19 younger adults (aged 22-39 years) recruited via web-based advertisements. To minimize usability barriers, only older adults who owned and regularly used smartphones were eligible to participate.

Prior to the intervention, all participants attended an orientation session held at a welfare center in Seoul, with simultaneous Zoom (Zoom Video Communications) access for those unable to attend in person. The orientation lasted approximately 1 hour and included an explanation of the study purpose and procedures, followed by the administration of the preintervention survey. Over the subsequent 2 weeks, participants were instructed to log in daily to the CoGen app, post or respond to discussion prompts, and engage with others through comments and reactions.

Following the 2-week use period, participants completed a postintervention survey. Subsequently, focus group interviews (FGIs) were conducted to gather qualitative feedback on UX, perceived value, and emotional engagement. All 39 participants were divided into 7 small groups (fewer than 6 participants per group), and each group participated in a single FGI session, resulting in 7 sessions in total. Interviews took place in November 2024, either in person at senior welfare centers or on the web via Zoom, depending on participant accessibility. All younger adults participated through web-based sessions, whereas older adults primarily attended in person unless remote participation was required. Each session lasted approximately 1 hour and was moderated by a lead researcher (HL), with a second researcher (JJ) observing and taking notes. Participants provided both verbal and written consent for audio recording and transcription. A structured open-ended question guide was used to explore experiences of intergenerational communication, app usability, emotional responses, and perceived facilitators and barriers.

Measures

Postuse App Evaluation and Behavioral Intentions

Participants' postuse evaluation of the app and their behavioral intentions were assessed using a multidimensional measurement strategy comprising both multi-item and single-item indicators, all rated on 5-point Likert scales. Items were adapted from instruments used in the discover phase and modified to capture users' emotional, cognitive, and behavioral responses after 2 weeks of app use. Single-item measures were selected for their clarity and conceptual precision, consistent with prior methodological recommendations [28-30].

Affective and Overall Evaluation

To evaluate affective aspects of intergenerational communication facilitated through the CoGen app, participants were asked to rate their comfort during cross-generational conversations ("How comfortable did you feel while having conversations with individuals from other generations using the CoGen app?") and their satisfaction with those conversations ("How satisfied were you with the intergenerational conversations through the app?"). These items captured users' emotional responses and overall impressions of the app-mediated interaction.

Emotional Engagement

Emotional engagement was measured through 4 items: "enjoyment," "happiness," "interest," and "a sense of similarity or affinity with others." These were presented under the prompt, "To what extent did you experience the emotions or values you had expected from using the intergenerational communication app?" and were designed to evaluate participants' subjective emotional experiences during app use.

Functional Usability

Functional usability was measured using 2 items: "It was easy to find the information I needed in the CoGen app" and "The information and content provided in the app were easy to understand." These were also rated on a 5-point Likert scale (1=strongly disagree and 5=strongly agree). All items were analyzed individually, and descriptive statistics were used to examine users' perceptions of the app's usability, clarity, and interface design.

Perceived Necessity, Use Intentions, and Recommendation Willingness

Participants' attitudes toward the CoGen app were assessed using 3 single-item measures, each rated on a 5-point Likert scale. Two items (eg, "Do you think the CoGen app is necessary?" and "Do you intend to continue using the CoGen app in the future?") were administered both before and after the intervention to assess perceived necessity and continued use intention, respectively. An additional item (eg, "Would you

XSL•FO
RenderX

recommend the CoGen app to others?") was included in the postintervention survey to assess participants' willingness to recommend the app to others.

Intergenerational Communication

Changes in intergenerational communication were measured using 6 items administered before and after the 2-week intervention, rated on a 5-point Likert scale (1=strongly disagree and 5=strongly agree). Items were adapted from the European Values Study and the revised Solidarity Scale [31,32], modified for app-based cross-generational contexts. Intergenerational interaction was assessed with 2 items (pre: α=.76; post: α=.79), such as "I now talk with members of other generations more frequently than before." Intergenerational understanding was measured with 4 items (pre: α=.87; post: α=.91), including "I have come to respect the culture of other generations" and "I have come to understand the realities faced by other generations."

Psychosocial Outcomes

Perceived social support was assessed using 8 items from the Medical Outcomes Study Social Support Survey [33], rated on a 5-point Likert scale (1=strongly disagree and 5=strongly agree) before and after the intervention (pre: α=.91; post: α=.93). Example items include "There is someone who makes me feel cared for" and "I have someone to talk to about work or household problems."

Depressive symptoms were measured using the 10-item short form of the Center for Epidemiologic Studies Depression Scale, validated in Korean [34], with responses rated on a 5-point scale (1=not at all and 5=almost always; pre: α=.83; post: α=.85). Sample items include "I felt depressed" and "I had trouble concentrating on what I was doing."

Data Analysis

A mixed methods approach was used to analyze both quantitative and qualitative data. For the quantitative component, paired-sample 2-tailed *t* tests were conducted to compare participants' pre- and postintervention responses on key outcome variables, including intergenerational interaction, perceived understanding, social support, and depressive symptoms. Descriptive statistics were also used to summarize postuse evaluations, usability perceptions, and emotional engagement indicators.

For the qualitative component, transcripts were analyzed using thematic analysis [35]. Three members of the research team—including 2 senior researchers with expertise in aging and 1 qualitative methods specialist—independently reviewed transcripts and field notes. An inductive coding process was used to identify recurring patterns and refine overarching themes such as perceived benefits, design-related frustrations, and recommendations for improvement. Triangulation and consensus meetings were used to enhance the trustworthiness and analytical rigor of the findings.

Ethical Considerations

This study received institutional review board approval from Ewha Womans University (IRB ewha-202203-0022-01). All participants provided written informed consent prior to participation, and data were anonymized to ensure confidentiality. Identifiable information was removed during transcription and analysis to protect participant privacy. As compensation, each participant received a gift voucher worth approximately US $15.

Results

Phase 1: Discover; Key Findings From User Research

Attitudes Toward Intergenerational Communication

Older adults exhibited significantly higher interest in intergenerational communication (mean 3.87, SD 0.63) compared to younger adults (mean 3.61, SD 0.81; t_{411}=–3.63; *P*<.001; *d*=0.36). They also placed greater importance on exchanging necessary information (mean 4.05, SD 0.66) than younger adults (mean 3.86, SD 0.82; t_{411}=–2.59; *P*=.01; *d*=0.26).

Perceived Necessity and Willingness to Use an Intergenerational Communication App

Older adults rated the perceived necessity of an intergenerational communication app significantly higher (mean 3.82, SD 0.73) than younger adults (mean 3.47, SD 0.83; t_{411}=–4.54; *P*<.001; *d*=0.45). They also demonstrated a greater willingness to use the app (mean 3.81, SD 0.71) compared to younger adults (mean 3.17, SD 0.87; t_{411}=–8.17; *P*<.001; *d*=0.81).

Motivational Factors for App Use

Distinct differences emerged in how and why users expected to use the app. Older adults were more likely to use the app for social communication (mean 3.87, SD 0.63) than younger adults (mean 3.61, SD 0.81; t_{411}=–3.63; *P*<.001; *d*=0.36), indicating a stronger interest in maintaining intergenerational relationships. They also placed greater importance on passing time (mean 2.97, SD 0.86) than younger adults (mean 3.19, SD 1.01; t_{411}=2.38; *P*=.018; *d*=–0.23), although the difference favored younger adults. In addition, older adults expressed significantly greater interest in entertainment and enjoyment (mean 3.50, SD 0.73) than younger adults (mean 3.19, SD 1.03; t_{411}=–3.52; *P*<.001; *d*=0.35).

Regarding exchanging necessary information or opinions, older adults again rated this higher (mean 4.05, SD 0.66) than younger adults (mean 3.86, SD 0.82; t_{411}=–2.59; *P*=.01; *d*=0.26). They also viewed the app as more useful for reducing loneliness (mean 3.10, SD 0.89) than younger adults (mean 2.89, SD 1.12; t_{411}=–2.10; *P*=.04; *d*=0.21) and placed greater emphasis on health management (mean 3.77, SD 0.88) compared to younger adults (mean 3.31, SD 1.02; t_{411}=–4.90; *P*<.001; *d*=0.47).

Younger adults, on the other hand, reported significantly greater interest in discussing concerns not easily shared with same-generation peers (mean 3.59, SD 0.94) than older adults (mean 3.06, SD 0.94; t_{411}=5.73; *P*<.001; *d*=–0.56). Finally, older adults rated cultural exchange significantly higher (mean 3.72, SD 0.72) than younger adults (mean 3.44, SD 0.94; t_{411}=–3.39; *P*=.001; *d*=0.33), suggesting a stronger openness to learning about different generational perspectives.

XSL•FO
RenderX

Preferred App Features

Perceived Ease of Use

Both younger and older adults highly prioritized the reliability of provided information (younger adults: mean 4.26, SD 0.86 and older adults: mean 4.41, SD 0.66), with older adults showing a marginally higher preference (t_{411}=–1.98; *P*=.048; *d*=0.20). Similarly, clarity and accuracy of information received high, comparable ratings from both groups (younger adults: mean 4.16, SD 0.83 and older adults: mean 4.26, SD 0.68; t_{411}=–1.34; *P*=.18; *d*=0.13).

A significant difference emerged in the ease of access to information, where older adults placed notably higher importance (mean 4.14, SD 0.60) than younger adults (mean 4.00, SD 0.73; t_{411}=–2.12; *P*=.03; *d*=0.21), underscoring the need for an intuitive and navigable interface for this demographic. While simplicity and readability of content were also more valued by older adults, this difference was not statistically significant.

Both groups rated the relevance, appropriateness, and immediate accessibility of content similarly, with no significant differences. Interestingly, younger adults placed slightly more emphasis on engagement and entertainment value, though this difference was not significant.

Perceived Usefulness

Older adults consistently reported significantly higher expectations across several usefulness-related values. They notably emphasized health as a critical factor (older adults: mean 4.02, SD 0.79 vs younger adults: mean 3.60, SD 1.03; t_{411}=–4.64; *P*<.001; *d*=0.46), highlighting their preference for health-related features. Additionally, novelty was rated significantly higher by older adults (mean 3.97, SD 0.74) than by younger adults (mean 3.45, SD 0.89; t_{411}=–6.44; *P*<.001; *d*=0.62), suggesting an attraction to fresh and innovative experiences.

In terms of emotional and social benefits, older adults consistently placed greater importance on convenience (mean 3.91, SD 0.62), fun (mean 3.85, SD 0.70), and friendship (mean 3.37, SD 0.72) compared to younger adults (convenience: mean 3.59, SD 0.95; fun: mean 3.46, SD 0.96; and friendship: mean 3.03, SD 0.97). They also rated confidence (mean 3.66, SD 0.81) and a sense of belonging (mean 3.54, SD 0.79) significantly higher than younger adults (confidence: mean 3.21, SD 0.91 and belonging: mean 3.25, SD 1.01). While not statistically significant, both age groups similarly recognized the emotional benefits related to memories, though these were less differentiated between them.

Aesthetic and Gamification Components

In evaluating aesthetic and gamification components, older adults generally expressed a greater interest in several features compared to younger adults. Specifically, older adults showed a significantly stronger interest in customization options like avatars and personalized profiles (older adults: mean 3.28, SD 0.89 vs younger adults: mean 2.92, SD 1.08; t_{411}=–3.69; *P*<.001; *d*=0.36), suggesting a desire for more personalized UXs.

A notable difference also emerged in friend integration, with older adults showing a much higher preference for linking the app with external platforms like KakaoTalk or Facebook (mean 3.64, SD 0.88 vs mean 3.04, SD 1.17; t_{411}=–5.87; *P*<.001; *d*=0.57). This highlights the importance of seamless social connectivity for older users. Furthermore, older adults reported greater interest in immersive meeting spaces (eg, augmented reality or virtual reality features) compared to younger adults (mean 3.12, SD 0.87 vs mean 2.91, SD 1.12; t_{411}=–2.12; *P*=.03; *d*=0.20), suggesting that immersive environments may enhance their digital engagement.

While older adults appreciated rewards like points and badges more than younger adults (mean 3.49, SD 0.84 vs mean 3.32, SD 1.12), this difference was not statistically significant (t_{411}=–1.74; *P*=.08; *d*=0.17). Finally, visual appeal (eg, icons and backgrounds) was rated similarly across both age groups (older adults: mean 3.50, SD 0.81 and younger adults: mean 3.46, SD 0.93; t_{411}=–0.47; *P*=.64; *d*=0.04), indicating a shared preference for aesthetically pleasing user interface (UI).

Table 1 summarizes generational differences in attitudes toward intergenerational communication, perceptions of app necessity and willingness to use, motivational factors, and preferred app features.

XSL•FO
RenderX

Table 1. Comparison of younger and older adults' attitudes, perceptions, motivations, and preferred features for an intergenerational communication app (N=413).

	Younger (n=202), mean (SD)	Older (n=211), mean (SD)	*t* test (*df*=411)	*P* value
Attitudes toward intergenerational communication				
Interest in cross-generational communication	3.61 (0.81)	*3.87 (0.63)* [a]	*–3.63*	*<.001*
Perceived importance of cross-generational communication	3.86 (0.82)	*4.05 (0.66)*	*–2.59*	*.01*
Perceived necessity and willingness to use an intergenerational communication app				
Perceived necessity of an intergenerational communication app	3.47 (0.83)	*3.82 (0.73)*	*–4.54*	*<.001*
Willingness to use the app	3.17 (0.87)	*3.81 (0.71)*	*–8.17*	*<.001*
Motivational factors for app use				
Social communication	3.61 (0.81)	*3.87 (0.63)*	*–3.63*	*<.001*
Passing time (relieving boredom)	3.19 (1.01)	*2.97 (0.86)*	*2.38*	*.01*
Entertainment and enjoyment	3.19 (1.03)	*3.50 (0.73)*	*–3.52*	*<.001*
Exchanging necessary information or opinions	3.86 (0.82)	*4.05 (0.66)*	*–2.59*	*.01*
Reducing loneliness	2.89 (1.12)	*3.10 (0.89)*	*–2.10*	*.04*
Health management	3.31 (1.02)	*3.77 (0.88)*	*–4.90*	*<.001*
Discussing concerns not easily shared with same-generation peers	3.59 (0.94)	*3.06 (0.94)*	*5.73*	*<.001*
Cultural exchange	3.44 (0.94)	*3.72 (0.72)*	*–3.39*	*.001*
Preferred app features				
Perceived ease of use				
Ease of access to information or content	4.00 (0.73)	*4.14 (0.60)*	*–2.12*	*.03*
Immediate accessibility	3.84 (0.85)	3.86 (0.70)	–0.26	.79
Reliability of provided information	4.26 (0.86)	*4.41 (0.66)*	*–1.98*	*.048*
Clarity and accuracy of information	4.16 (0.83)	4.26 (0.68)	–1.34	.18
Relevance and appropriateness of content	4.05 (0.75)	4.12 (0.69)	–0.99	.32
Simplicity and readability of content	3.91 (0.93)	4.03 (0.78)	–1.42	.16
Engagement and entertainment value	3.87 (0.79)	3.81 (0.75)	0.79	.43
Perceived usefulness				
Convenience	3.59 (0.95)	*3.91 (0.62)*	*–4.03*	*<.001*
Fun	3.46 (0.96)	*3.85 (0.70)*	*–4.70*	*<.001*
Friendship	3.03 (0.97)	*3.37 (0.72)*	*–4.03*	*<.001*
Calmness	3.44 (1.01)	3.58 (0.74)	–1.60	.11
Confidence	3.21 (0.91)	*3.66 (0.81)*	*–5.30*	*<.001*
Sense of belonging	3.25 (1.01)	*3.54 (0.79)*	*–3.24*	*.001*
Memories	3.48 (1.01)	3.46 (0.86)	0.22	.83
Satisfaction	3.43 (0.95)	*3.63 (0.76)*	*–2.36*	*.019*
Health	3.60 (1.03)	*4.02 (0.79)*	*–4.64*	*<.001*
Novelty	3.45 (0.89)	*3.97 (0.74)*	*–6.44*	*<.001*
Aesthetics and gamification components				
Rewards (points, badges, levels)	3.32 (1.12)	3.49 (0.84)	–1.74	.08
Customization (avatars, profiles)	2.92 (1.08)	*3.28 (0.89)*	*–3.69*	*<.001*
Friend integration (KakaoTalk, Facebook)	3.04 (1.17)	*3.64 (0.88)*	*–5.87*	*<.001*
Immersive meeting spaces (AR[b] or VR[c])	2.91 (1.12)	*3.12 (0.87)*	*–2.12*	*.03*

XSL•FO
RenderX

	Younger (n=202), mean (SD)	Older (n=211), mean (SD)	*t* test (*df*=411)	*P* value
Visual appeal (icons, backgrounds)	3.46 (0.93)	3.5 (0.81)	–0.47	.64

[a]Values in italics format indicate statistically significant differences at $P<.05$.

[b]AR: augmented reality.

[c]VR: virtual reality.

Qualitative Responses on Intergenerational Information and Conversation Topics

Participants expressed strong interest in sharing hobbies, health tips, financial advice, travel experiences, and pet care—topics that may serve as effective entry points for intergenerational dialogue. These findings support the use of thematically structured discussion boards to guide user-generated content, providing a flexible yet focused environment for sustained and reciprocal interaction based on shared interests.

Phase 2: Define; UX Priorities and Design Implications

This phase translated the findings from phase 1 into concrete design priorities for the CoGen prototype. Quantitative and qualitative analyses revealed distinct generational differences in usability preferences, interaction styles, and motivational drivers.

Generational Differences in UX Preferences

Older adults emphasized clarity, structure, and accessibility. For instance, 73.4% (n=155) of older users selected "easy navigation" as a top priority compared to 52.6% (n=106) of younger users. They also rated the need for structured information layouts as significantly higher (mean 4.26, SD 0.62) than younger adults (mean 3.68, SD 0.81; t_{411}=–4.57; $P<.001$).

In contrast, younger adults favored engaging and dynamic elements. A total of 68.3% (n=138) of younger users prioritized "fun and dynamic design" compared to 41.2% (n=87) of older adults. They also showed greater interest in gamified features such as conversation streaks and badges (mean 3.92, SD 0.77) than older users (mean 3.41, SD 0.84; t_{411}=–3.21; P=.002).

Both groups expressed strong interest in structured discussion themes, though topic preferences varied. Older adults preferred themes such as health, pets, and daily life tips, while younger users leaned toward travel, hobbies, and financial advice. Open-ended responses supported these findings, with one older adult stating, "If I could talk about gardening or health tips, I would join more often," and a younger participant noting, "It's fun when it feels like social media, not a lecture."

Design Implications for the CoGen App

These generational patterns directly informed the prototype development of CoGen. The interface was designed to balance simplicity with visual appeal, offering a minimalist layout suited to older users while providing customizable elements to satisfy younger users' preferences. Accessibility features such as optional auditory cues (eg, notification sounds and voice-based guidance) were included to accommodate diverse needs without overwhelming the user. Thematic discussion boards organized around popular topics such as health, hobbies, and finance were implemented to support guided yet flexible interactions. Additionally, participatory components like conversation streaks, shared milestones, and feedback prompts were incorporated to foster social engagement and maintain user interest over time, especially among younger participants.

Phase 3: Develop; Finalized Features and User-Driven Adjustments

This phase focused on translating insights from earlier phases into a functional CoGen prototype through iterative development. Drawing on findings from phases 1 and 2, the design process integrated key generational preferences into the app's structure and functionality, followed by small-scale usability testing and refinement.

UI or UX Design and Wireframing Outcomes

Based on generational needs identified in earlier phases, the CoGen prototype was wireframed with a focus on accessibility for older adults and visual engagement for younger users. Older participants responded positively to features such as enlarged font sizes and high-contrast modes, frequently describing them as "easier to read" and "less tiring on the eyes." In contrast, younger users favored vibrant color schemes and modern layouts, stating that the design "felt more like a social space." The home screen—structured around 5 core functions—was perceived as intuitive and manageable across age groups. Additionally, the inclusion of a numeric level indicator with a horizontal progress bar was initially well received as a simple motivator, though later feedback suggested a need for more emotionally resonant visualizations.

Prototype Development and Feature Integration Outcomes

Prototype Features and User Feedback

The functional prototype incorporated features designed to foster sustained intergenerational engagement, informed by user needs identified in earlier phases. Feedback from older and younger participants highlighted differing priorities—older adults emphasized meaningful conversation and ease of use, while younger users favored interactivity and entertainment-driven engagement. Four key features were evaluated and refined based on user reactions.

Discussion Boards

Theme-based forums were designed around topics such as hobbies, travel, and health. This structure aimed to support older users' desire for organized, topic-specific discussions while offering younger users spaces for more casual and spontaneous interaction.

XSL•FO
RenderX

Conversation Triggers

Participatory elements, including prompts and quizzes, were introduced to ease the initiation of conversations. These features were intended to reduce entry barriers and encourage mutual storytelling, particularly among younger users who prefer playful engagement.

1:1 Friend Matching

A proprietary friend-matching feature was included to support more personalized interactions, especially for older adults who seek deeper, one-on-one conversations. This function was designed as an in-app alternative to social media integration, minimizing complexity and maintaining user privacy.

Generational Identity Icons

To promote age-aware interaction, generation-specific icons (eg, youth and older adult symbols) were added next to each user ID. This visual cue enabled users to identify the generational background of their conversation partners, fostering mutual understanding and respectful communication across age groups (Figure 2).

Figure 2. Generational identity icons: youth (left) and older adult (right).



Progression-Based Gamification

To encourage consistent app engagement, a light-touch gamification system was implemented. Users could earn points through participation (eg, posting or commenting), which contributed to levels and badges. The system was intentionally noncompetitive, reflecting older users' preference for collaborative over competitive dynamics.

Preliminary Usability Testing and Prototype Refinement

Preliminary Testing Outcomes

Initial usability testing with 8 participants (5 younger adults and 3 older adults) revealed critical insights that informed refinements to the CoGen prototype. Participants evaluated the app's overall appeal, navigability, and clarity of features. Several features were refined based on this feedback.

Branding and Identity

The app was named "CoGen" to reflect its focus on cogenerational engagement. The elephant logo—symbolizing wisdom, memory, and social bonds—resonated with participants' desire for meaningful and emotionally grounded experiences.

Weekly Question Feature

Participants favored prompts that evoked shared memories and generational perspectives. Based on this input, the "Weekly Question" was redesigned to include open-ended questions that invite storytelling and nostalgic reflection.

Engagement Indicator

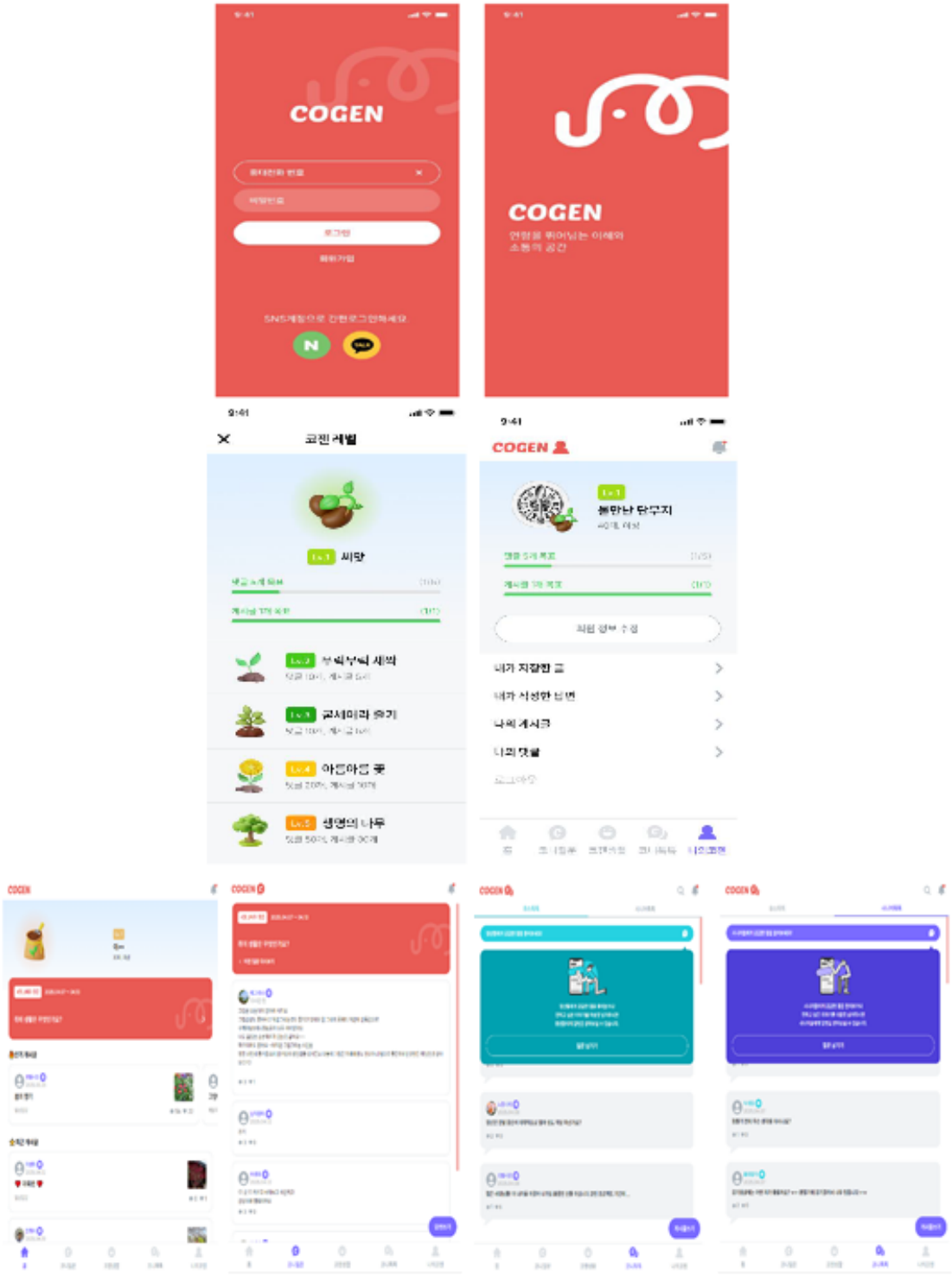
The original numeric level and progress bar were perceived as impersonal. In response, the engagement indicator was redesigned into a metaphorical growth journey—a seed blossoming into a flower—offering a more intuitive and emotionally resonant representation of user progress, especially appreciated by older participants.

Tone of Feature Labels

Younger participants emphasized the importance of a conversational tone. In response, feature names were modified to be more inviting and playful. For example, the discussion board was renamed "CoGen Life," and the weekly prompt became "Coni's Question." These refinements finalized the CoGen prototype by incorporating user-driven feedback into its core features and design, setting the stage for real-world testing in the deliver phase (Figure 3).

XSL•FO
RenderX

Figure 3. Final version of the CoGen app interface, displaying (from top to bottom): the login screen, level indicator, home screen with "Coni's Question," and "Cogen TalkTalk" features.



Phase 4: Deliver; Real-World Testing Outcomes and Final Revisions

User Engagement Patterns Based on Activity Log Data

To assess the feasibility and user engagement of the CoGen platform during the 2-week intervention, activity log data were analyzed. Participants logged into the app an average of 31.43 (SD 10.56) times, with a total of 2975 posts and comments generated—equivalent to an average of 76.28 (SD 20.13) entries per user. This indicates both frequent access and active participation in intergenerational discussions.

Engagement with in-app features—such as quizzes and feedback prompts—was also high, with 80.3% (n=31) of users completing them at least once. The dropout rate remained low at 5.3% (n=2), and activity levels were stable throughout the period, suggesting sustained engagement.

XSL•FO
RenderX

UX: Functional Satisfaction

Participants reported moderate to high levels of comfort in cross-generational conversations (mean 3.62, SD 0.81) and relatively high overall satisfaction (mean 3.82, SD 0.76) with intergenerational conversations via the app.

Regarding emotional engagement, users reported high levels of enjoyment (mean 3.74, SD 0.85) and happiness (mean 3.79, SD 0.77), with perceived interest being the highest measured variable (mean 4.05, SD 0.79). While the perceived sense of similarity or affinity was moderate (mean 3.46, SD 0.94), it still indicated a positive trend in user closeness.

Conversely, ease of access received a comparatively lower rating (mean 2.59, SD 0.88), indicating potential navigation difficulties. Similarly, perceived understanding of content was moderate (mean 3.23, SD 0.71), suggesting opportunities for optimizing content presentation and guidance.

Impact on Perceived Necessity, Use Intentions, and Recommendation Willingness

No statistically significant changes were observed in participants' perceived necessity (t_{37}=1.58; *P*=.13; *d*=0.45) or intention to continue using the app (t_{37}=0.81; *P*=.43; *d*=0.25; Table 2). However, perceived necessity was already high before the intervention (mean 3.70, SD 0.73; mean 4.05, SD 0.83), suggesting a potential ceiling effect. Intention to use remained moderate but stable (mean 3.30, SD 0.86; mean 3.50, SD 0.69). Participants also showed a moderate willingness to recommend the app (mean 3.55, SD 1.01), reflecting its perceived value in fostering cross-generational communication.

Table 2. Changes in key variables before and after the intervention (n=39).

Variable	Pre, mean (SD)	Post, mean (SD)	*t* test (*df*=37)	*P* value
Intergenerational interaction	2.85 (1.04)	3.85 (0.75)	–3.82	.001
Intergenerational understanding	1.42 (0.40)	3.02 (0.72)	–10.26	<.001
Perceived social support	3.03 (0.53)	3.85 (0.78)	–5.53	<.001
Depressive symptoms	1.85 (0.36)	1.82 (0.37)	0.4	.69
Perceived necessity of the app	3.70 (0.73)	4.05 (0.83)	1.58	.13
Intention to continue using the app	3.30 (0.86)	3.50 (0.69)	0.81	.43
Willingness to recommend the app (post)	—[a]	3.55 (1.01)	—	—

[a]Not available.

Changes in Intergenerational Communication: Interaction and Understanding

Paired-sample 2-tailed *t* tests revealed statistically significant improvements in both intergenerational interaction and understanding after the 2-week intervention (Table 2). Intergenerational interaction increased significantly (t_{37}=–3.82; *P*=.001; *d*=0.52), signifying more frequent and natural conversations. Intergenerational understanding also showed a substantial and statistically significant improvement (t_{37}=–10.26; *P*<.001; *d*=1.16), indicating that the app effectively enhanced empathy toward different generational perspectives.

Impact on Psychosocial Outcomes: Social Support and Depressive Symptoms

Perceived social support significantly increased following 2 weeks of CoGen app use (t_{37}=–5.53; *P*<.001; *d*=0.66), suggesting that the app strengthened emotional bonds (Table 2). In contrast, depressive symptoms showed no statistically significant change (t_{37}=0.40; *P*=.69; *d*=–0.06), implying that while social connectedness improved, immediate effects on emotional well-being were not observed within the short time frame.

Qualitative Insights From the FGI

To supplement the survey findings, an FGI was conducted, revealing that participants from both age groups found the app emotionally enriching and socially meaningful. They especially appreciated the structured yet open environment that supported respectful and thoughtful dialogue.

Mutual Adaptation and Respectful Communication

Participants reported that their attitudes toward intergenerational communication shifted positively during the 2-week period. Younger adults initially expressed concern about potential generational conflict but noted that the platform encouraged the use of respectful and considerate language. In turn, older adults also exercised caution to avoid appearing condescending or outdated, demonstrating a shared effort to communicate with mutual respect. This digital form of etiquette was recognized by both groups as an important cultural adjustment, especially when engaging with unfamiliar age cohorts.

Psychosocial Gains and Shifting Perceptions

The app served as a space where younger users could seek emotional support, practical advice, and broader perspectives from older participants. Many reported gaining comfort, validation, and a renewed appreciation for the insights of older generations. Meanwhile, older adults described feeling revitalized by the interactions, often expressing surprise at the empathy and openness of younger users. They also viewed the app as a rare opportunity to reflect on their own life stories while learning about new cultural trends and digital practices.

Common Recognition of Generational Diversity

Participants across age groups highlighted how the experience disrupted stereotypical perceptions. Younger users came to recognize the diversity among older adults, moving beyond

XSL•FO
RenderX

generalized assumptions. Conversely, older participants reported gaining a more nuanced understanding of younger people's struggles and viewpoints, fostering empathy and solidarity.

Conditions for Sustained Use

Despite the positive feedback, participants noted several functional improvements necessary for long-term engagement. Many called for real-time chat features, simplified login processes, and improved content navigation to support more spontaneous and efficient communication. Additionally, participants expressed that future iterations of the app should focus on enabling users to "gain something meaningful" through dialogue—such as wisdom, emotional encouragement, or practical advice. Both groups emphasized the need for structured prompts and shared interests (eg, health and hobbies) to sustain motivation.

Design Suggestions

Users also offered specific UI or UX recommendations, including expressive elements like emojis or stickers, and personalized notification systems. These enhancements were seen as important for maintaining emotional resonance and playful interaction across generations, supporting more engaging and expressive communication for all users.

Discussion

Principal Findings

This study provides preliminary evidence that the CoGen app may contribute to improvements in interaction, perceived social support, and comfort in cross-generational conversations. These results align with previous findings, indicating that intentional and structured opportunities for age-diverse contact can reduce prejudice and strengthen social bonds [5,7]. Specifically, the observed improvements in dialogue across generational lines are consistent with the principles of the contact hypothesis, suggesting that digital platforms hold potential spaces for meaningful engagement when designed to facilitate regular and constructive exchanges among different age groups.

Improvements in perceived social support align with previous research, indicating that older adults benefit from increased digital communication by maintaining social ties and mitigating feelings of isolation [36,37]. Younger adults also reported gaining fresh perspectives and insights from older generations, suggesting that positive, bidirectional exchanges can foster empathy and reduce generational stereotyping. This finding is aligned with prior studies, highlighting the significance of qualitatively rich interactions over mere contact frequency in achieving prejudice reduction and mutual understanding [8].

Furthermore, the study's results suggest that participants experienced an increased sense of comfort and satisfaction in conversations with individuals from different generations. These findings tentatively support the view that intergenerational communication can be nurtured beyond physical copresence, especially as digital literacy expands among older adults. Although ease of access scored relatively lower, it underlines the importance of continued efforts to design age-friendly digital interfaces that reduce navigation barriers and enhance older adults' confidence in adopting new technologies [38]. Participants specifically emphasized the need for auto-login features, clearer content categorization, and easier navigation to locate new or popular posts. Older adults in particular faced challenges in understanding the interface layout and expressed a desire for simplified, intuitive pathways to access discussions. To enhance emotional engagement, suggestions included incorporating expressive elements such as emojis and stickers.

A key emergent finding from the FGIs was the positive reception of the generational badge feature, which visually distinguished older and younger users. Despite initial concerns about reinforcing age-based divisions, many participants noted that it surprisingly encouraged respectful and age-sensitive communication, even in anonymous interactions. Younger users, in particular, reported feeling a greater sense of obligation to maintain courtesy, while older users appreciated being addressed with politeness. These findings suggest that CoGen's features organically aligned with Korean cultural norms rooted in Confucian values, such as intergenerational respect and age-based etiquette. Rather than merely enabling conversations, the app fostered a communicative atmosphere marked by mutual consideration—an emergent property shaped by both interface design and cultural expectations.

Moreover, younger participants shared in the FGIs that the app served as a psychologically safe space to seek advice or express concerns they might hesitate to bring up in face-to-face settings (eg, with parents). In contrast to offline environments where social expectations and *noonchi*—the culturally embedded sensitivity to others' feelings—can act as barriers, and the app allowed for open and honest exchanges without fear of breaching generational decorum. This illustrates how CoGen's design facilitated not just communication, but culturally meaningful connection, by providing users with both symbolic cues for respectful interaction and a space that eased the emotional burden of navigating age hierarchies.

In contrast to many existing digital interventions that often emphasize structured prompts or 1-way content delivery, CoGen was specifically designed to support user-generated content and dynamic, reciprocal communication between generations. Most prior platforms still lack robust mechanisms for sustained, bidirectional interaction, and despite increasing digital literacy among older adults [13,14], many fall short in addressing the nuanced interaction needs and co-design preferences of both younger and older users. Addressing this critical gap, CoGen positions itself as a hybrid model bridging formal programs and informal social interaction—an area still underdeveloped in the digital aging landscape. The app integrates both synchronous and asynchronous communication, encourages user-generated content, and offers structured prompts to support long-term engagement. This approach is strongly reinforced by CoGen's adherence to principles of inclusive co-design and iterative user-centered development [15]. As literature increasingly highlights the digital literacy of older adults [39] and the importance of inclusive co-design [40], CoGen specifically reflects these principles through its iterative development and participatory UX testing. These practices not only improve usability but also enhance users' sense of ownership and

XSL•FO
RenderX

emotional attachment to the platform, further supporting long-term adoption.

Despite its promising outcomes, the relatively lower usability scores underscore the need for improved accessibility. Participants emphasized visual clarity, reduced click-depth, and features like auto-login and richer emotional expression tools. These findings reinforce the importance of UCD and inclusive codevelopment, ensuring that intergenerational digital tools remain not only functional but also emotionally engaging and culturally responsive. CoGen's iterative design process and hybrid communication structure offer a promising model—particularly in societies like Korea, where age-based etiquette shapes interaction norms—by fostering respectful and psychologically safe exchanges across generations.

Theoretical and Practical Implications

The structured features of CoGen—such as discussion boards, quizzes, and shared achievements—appeared to reduce social distance and foster mutual understanding across generations. These affordances reflect core principles of empathy-building and cooperative engagement found in offline intergenerational programs [7,41], while offering scalable digital alternatives. Although the contact hypothesis was not directly tested, several of its key conditions—such as equal status, shared goals, and structured interaction—were naturally embedded in the design of CoGen. These elements may have contributed to the observed improvements in mutual understanding and reduction of age-based bias. This highlights the value of grounding digital intervention development not only in UCD principles but also in theoretically informed frameworks, which together can enhance both usability and psychosocial outcomes.

As digital connectivity expands globally, platforms like CoGen hold promise for supporting intergenerational relationships, despite physical separation. This function aligns with models of digital age-friendly communities [38] and aging-in-place strategies [42], both of which emphasize the role of sustained social engagement in promoting healthy aging. By enabling remote yet meaningful interactions, CoGen represents a digital extension of the World Health Organization's vision of inclusive, communicative, and socially connected environments for older adults [43].

The observed improvement in perceived social support reinforces previous findings that emotionally resonant communication enhances older adults' well-being and cognitive engagement [44,45]. CoGen's focus on sustained dialogue may thus contribute to improved life satisfaction and emotional resilience. In line with World Health Organization guidelines [42], promoting regular digital interactions through accessible platforms can also support independent living and active aging.

In the Korean context, where Confucian values such as filial piety and age-based respect remain culturally salient, CoGen's emphasis on structured, respectful dialogue aligns well with traditional norms of intergenerational harmony. Participant feedback indicated that the platform created a safe space for expressing concerns—especially for younger users—while maintaining cultural expectations around age hierarchy. This suggests the culturally adaptive potential of digital tools to bridge generational gaps by balancing modern communication and traditional values.

Finally, design feedback points to future opportunities for enhancing cross-generational engagement. Participants emphasized the value of algorithm-based matching systems to facilitate interest-based dialogue, as well as the need for simplified UI or UX for older adults. Incorporating expressive features such as emojis and stickers may further support engagement among younger users. These insights highlight the importance of participatory, multigenerational design frameworks in creating emotionally inclusive and functionally accessible digital environments.

Limitations and Future Research

Despite its promising findings, this study has several limitations that offer clear directions for future research. First, the sample size was relatively small, and the 2-week intervention period was short. These factors may limit the generalizability of our findings and the understanding of long-term engagement. Future research should involve large-scale, longitudinal studies to track sustained engagement and relational outcomes over an extended period. Such studies would allow us to verify if the observed benefits persist and evolve across more diverse cultural and demographic contexts.

Second, our design could be enhanced by incorporating real-time interaction and deeper personalization. While the app effectively supported asynchronous communication, participants expressed a strong desire for synchronous features like direct messaging. Moreover, to move beyond basic usability and better capture the emotionally sensitive nature of intergenerational dialogue, future design cycles will integrate formal tools from design thinking. Specifically, we plan to use empathy maps and user journey visualizations to inform the development of more sophisticated features, such as adaptive dialogue prompts and sentiment-aware interaction flows, thereby creating a more responsive and emotionally resonant UX.

Third, our reliance on self-reported measures could be complemented by objective data. While valuable, self-reports on comfort or connection may not capture the full picture of user behavior. Future studies could create a more holistic understanding by integrating use analytics—such as the frequency, duration, and content analysis of interactions—with self-report data to triangulate findings on how digital communication truly evolves.

Finally, several key constructs were measured using single-item indicators. This was a deliberate methodological choice grounded in both practical and theoretical considerations. Rather than combining conceptually distinct items into composite scores, we retained item-level granularity to capture feature-specific responses that could directly inform app development. This approach aligns with prior research indicating that single-item measures can be valid and useful, especially when constructs are concrete, unidimensional, and behaviorally specific [28]. Although the items were carefully adapted from previously validated multi-item instruments, we acknowledge that multi-item scales generally offer higher reliability. Future studies should incorporate established multi-item measures to

XSL•FO
RenderX

validate the findings and enable more robust analyses, such as latent variable modeling, to advance theoretical understanding. By addressing these limitations, future research can build upon our work to design inclusive digital interventions that foster meaningful connections and contribute to the broader goals of healthy aging and creating age-friendly societies.

Acknowledgments

This work was supported by the National Research Foundation of Korea (grant NRF-2020S1A5C2A03092919).

Conflicts of Interest

None declared.

References

1. Francioli SP, Danbold F, North MS. Millennials versus boomers: an asymmetric pattern of realistic and symbolic threats drives intergenerational tensions in the United States. Pers Soc Psychol Bull 2024;50(11):1546-1562 [FREE Full text] [doi: 10.1177/01461672231164203] [Medline: 37133238]
2. Chung S, Lee S, Park M, Shin B, Kim JY. Areas and causes of intergenerational conflict between age groups based on expert Delphi survey. Korean J Soc Welf 2021;73(3):229-253. [doi: 10.20970/kasw.2021.73.3.009]
3. Park M, Chung S. The effect of social isolation of young generation on perception and behavior of discrimination against the older people: mediating effect of perception of generational conflict. Korea Gerontol Soc 2020;40(6):1383-1401. [doi: 10.31888/jkgs.2020.40.6.1383]
4. Tajfel H, Turner J, Austin W, Worchel S. An integrative theory of intergroup conflict. In: Organizational Identity: A Reader. London, England: Psychology Press; 1979:94-109.
5. Allport GW. The Nature of Prejudice. 1st Edition. Reading, MA: Addison-Wesley; 1954:1-537.
6. Teater B. How stereotypes and attitudes toward older adults are challenged through intergenerational contact: young people explain the process. Gerontol Geriatr Educ 2018;39(1):104-116. [doi: 10.1080/02701960.2015.1115985] [Medline: 26886686]
7. Drury L, Hutchison P, Abrams D. Direct and extended intergenerational contact and young people's attitudes towards older adults. Br J Soc Psychol 2016;55(3):522-543 [FREE Full text] [doi: 10.1111/bjso.12146] [Medline: 27256485]
8. Webster M, Norwood K, Waterworth J, Leavey G. Effectiveness of intergenerational exchange programs between adolescents and older adults: a systematic review. J Intergener Relatsh 2023;22(4):603-644. [doi: 10.1080/15350770.2023.2267532]
9. Zhong S, Lee C, Foster MJ, Bian J. Intergenerational communities: a systematic literature review of intergenerational interactions and older adults' health-related outcomes. Soc Sci Med 2020;264:113374. [doi: 10.1016/j.socscimed.2020.113374] [Medline: 33017736]
10. Döring N, Mikhailova V, Brandenburg K, Broll W, Gross H, Werner S, et al. Digital media in intergenerational communication: status quo and future scenarios for the grandparent-grandchild relationship. Univers Access Inf Soc 2022;23:1-16 [FREE Full text] [doi: 10.1007/s10209-022-00957-w] [Medline: 36530861]
11. ElGibreen H, Al Ali G, AlMegren R, AlEid R, AlQahtani S. Telepresence robot system for people with speech or mobility disabilities. Sensors (Basel) 2022;22(22):8746 [FREE Full text] [doi: 10.3390/s22228746] [Medline: 36433341]
12. Conroy KE, Kochoska A, Hey D, Pablo H, Hambleton KM, Jones D, et al. Physics of eclipsing binaries. V. General framework for solving the inverse problem. Astrophys J Suppl Ser 2020;250(2):34. [doi: 10.3847/1538-4365/abb4e2]
13. Müller M, Otero N, Milrad M. Guiding the design and implementation of interactive public displays in educational settings. J Comput Educ 2023;11(3):823-854. [doi: 10.1007/s40692-023-00280-0]
14. Phang JK, Kwan YH, Yoon S, Goh H, Yee WQ, Tan CS, et al. Digital intergenerational program to reduce loneliness and social isolation among older adults: realist review. JMIR Aging 2023;6:e39848 [FREE Full text] [doi: 10.2196/39848] [Medline: 36598801]
15. Norman D, Draper S. User Centered System Design: New Perspectives on Human-Computer Interaction. Hillsdale, NJ: Lawrence Erlbaum Associates; 1986.
16. Cornet VP, Toscos T, Bolchini D, Rohani Ghahari R, Ahmed R, Daley C, et al. Untold stories in user-centered design of mobile health: practical challenges and strategies learned from the design and evaluation of an app for older adults with heart failure. JMIR Mhealth Uhealth 2020;8(7):e17703 [FREE Full text] [doi: 10.2196/17703] [Medline: 32706745]
17. A study of the design process: the Double Diamond. Design Council. 2005. URL: https://www.designcouncil.org.uk/our-resources/the-double-diamond/ [accessed 2025-02-07]
18. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989;13(3):319-340. [doi: 10.2307/249008]
19. Gomez-Hernandez M, Ferre X, Moral C, Villalba-Mora E. Design guidelines of mobile apps for older adults: systematic review and thematic analysis. JMIR Mhealth Uhealth 2023;11:e43186 [FREE Full text] [doi: 10.2196/43186] [Medline: 37733401]
20. Florenthal B. J Res Interact Mark 2019;13(3):351-391. [doi: 10.1108/jrim-05-2018-0064]

XSL•FO
RenderX

21. Hyeon-Woo L, Jeong-jin L. Differences of ideological conflict between generations. National Strategy 2013;19(1):57-81. [doi: 10.35390/sejong.19.1.201303.003]
22. Lee JA, Na EY. The influence of motivation for using smart phone applications and innovativeness/compatibility on application use and substitution for other media/living equipments. Korean J Journal Commun Stud 2012;56(5):5-35 [FREE Full text]
23. Jun D, Kim D, Loni H. The impact of mobile social media use on user's emotional status. J Korean Assoc Reg Inf Soc 2017;20(2):65-88. [doi: 10.22896/karis.2017.20.2.004]
24. Kim Y, Kim D. Factors of buying defined on smart phone app. Using AHP. J Digit Interact Des 2012;11:23-33 [FREE Full text] [doi: 10.21195/jidr.2012.11.1.002]
25. Kim SS, Han KS, Kim BS, Park SK, Ahn SK. An empirical study on users' intention to use mobile applications. J Korean Inst Inf Technol 2011;9(8):213-228 [FREE Full text]
26. Choi J, Lee JS, Choi SH. Factors influencing mobile messenger interface design quality: with a focus on KakaoTalk application. J Digit Des 2012;12(3):211-220. [doi: 10.17280/jdd.2012.12.3.021]
27. Kim H, Kim J. A study on the graphic user interface for emotional communications in the mobile messenger—focused on 'Kakao Talk' and ‘WhatsApp' by case analysis. J Korea Des Knowl 2011;18:142-155. [doi: 10.17246/jkdk.2011..18.015]
28. Allen MS, Iliescu D, Greiff S. Single item measures in psychological science. Eur J Psychol Assess 2022;38(1):1-5. [doi: 10.1027/1015-5759/a000699]
29. Bergkvist L, Rossiter JR. The predictive validity of multiple-item versus single-item measures of the same constructs. J Market Res 2007;44(2):175-184. [doi: 10.1509/jmkr.44.2.175]
30. Wanous JP, Reichers AE, Hudy MJ. Overall job satisfaction: how good are single-item measures? J Appl Psychol 1997;82(2):247-252. [doi: 10.1037/0021-9010.82.2.247] [Medline: 9109282]
31. Wheeless LR. Self-disclosure and interpersonal solidarity: measurement, validation, and relationships. Human Comm Res 1976;3(1):47-61. [doi: 10.1111/j.1468-2958.1976.tb00503.x]
32. Voicu B, Bartolome Peral E, Rusu H, Rosta G, Comşa M, Vasile O, et al. COVID-19 and orientations towards solidarity: the cases of Spain, Hungary, and Romania. Eur Soc 2020;23(Suppl 1):S887-S904. [doi: 10.1080/14616696.2020.1852439]
33. Sherbourne CD, Stewart AL. The MOS social support survey. Soc Sci Med 1991;32(6):705-714. [doi: 10.1016/0277-9536(91)90150-b] [Medline: 2035047]
34. Shin SY. Validity study of short forms of the Korean version Center for Epidemiologic Studies Depression Scale (CES-D) [Unpublished master's thesis]. Seoul National University. 2011 Feb. URL: http://dcollection.snu.ac.kr:80/jsp/common/DcLoOrgPer.jsp?sItemId=000000030438 [accessed 2024-05-16]
35. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
36. Nguyen MH, Gruber J, Marler W, Hunsaker A, Fuchs J, Hargittai E. Staying connected while physically apart: digital communication when face-to-face interactions are limited. New Media Soc 2021;24(9):2046-2067. [doi: 10.1177/1461444820985442]
37. Thangavel G, Memedi M, Hedström K. Customized information and communication technology for reducing social isolation and loneliness among older adults: scoping review. JMIR Ment Health 2022;9(3):e34221 [FREE Full text] [doi: 10.2196/34221] [Medline: 35254273]
38. Ahn C, Ham Y, Kim J, Kim J. A digital twin city model for age-friendly communities: capturing environmental distress from multimodal sensory data. 2020 Presented at: Proceedings of the 53rd Hawaii International Conference on System Sciences; January 7-10, 2020; Hawaii, United States p. 7-10. [doi: 10.24251/hicss.2020.206]
39. Vaportzis E, Clausen MG, Gow AJ. Older adults perceptions of technology and barriers to interacting with tablet computers: a focus group study. Front Psychol 2017;8:1687 [FREE Full text] [doi: 10.3389/fpsyg.2017.01687] [Medline: 29071004]
40. Waycott J, Vetere F, Pedell S. Designing communication technologies for socially isolated older people: fieldwork reflections. 2013 Presented at: Proceedings of the Changing Perspectives, the 2013 ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2013); April 27-May 2, 2013; Paris, France.
41. Geng S, Gai L, Bu W. Interactive product design from the perspective of intergenerational integration. 2024 Presented at: Proceedings of the 3rd International Conference on Art Design and Digital Technology (ADDT 2024); May 24-26, 2024; Luoyang, China. [doi: 10.4108/eai.24-5-2024.2350165]
42. Global age-friendly cities: a guide. World Health Organization. 2007. URL: https://www.who.int/publications/i/item/9789241547307 [accessed 2007-10-05]
43. Baldassar L, Wilding R, Boccagni P, Merla L. Aging in place in a mobile world: new media and older people’s support networks. Transnational Soc Rev 2017;7(1):2-9. [doi: 10.1080/21931674.2016.1277864]
44. Huxhold O, Miche M, Schüz B. Benefits of having friends in older ages: differential effects of informal social activities on well-being in middle-aged and older adults. J Gerontol B Psychol Sci Soc Sci 2014;69(3):366-375. [doi: 10.1093/geronb/gbt029] [Medline: 23682165]
45. Shin H, Park C. Social support and psychological well-being in younger and older adults: the mediating effects of basic psychological need satisfaction. Front Psychol 2022;13:1051968 [FREE Full text] [doi: 10.3389/fpsyg.2022.1051968] [Medline: 36507030]

XSL•FO
RenderX

Abbreviations

FGI: focus group interview
UCD: user-centered design
UI: user interface
UX: user experience

Edited by N Schütz; submitted 14.04.25; peer-reviewed by JJ Yoo, JT Tseng; comments to author 29.05.25; revised version received 01.07.25; accepted 20.07.25; published 17.09.25.

Please cite as:
Chung S, Lee H, Jung J
Design and Evaluation of a Mobile App for Intergenerational Communication: User-Centered Participatory Design and Experimental Mixed Methods Study
JMIR Aging 2025;8:e75950
URL: https://aging.jmir.org/2025/1/e75950
doi:10.2196/75950
PMID:

©Soondool Chung, Hannah Lee, Jeehye Jung. Originally published in JMIR Aging (https://aging.jmir.org), 17.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

An mHealth App (NeoMayor) to Promote Healthy Lifestyles and Brain Health in Older Adults: Design and Validation Study

Myriam Gutiérrez[1,2,3,4*], MS, PhD; Victoria Cabello[1,5*], MS, PhD; Carol D SanMartín[6,7], MS, PhD; Jorge Mauro[4], PT, MS; Gada Musa[1], MS; Melissa Torres[1], BS, PT; Maria Consuelo San Martin[1,8], MS, PhD; Carlos Márquez[1,7,9], PhD, MS; Rodrigo Beltrán[1], BS; Nicole Rogers[2,10,11], MD, PhD; Felipe Salech[2,6], MD, PhD; Daniela Ponce[6], BEng; Christian Ampuero[12], BEng; Pia Varas[12], BA; Bastián Gamboa-Labbé[12], BEng; Víctor David Cortés[12], BEng; Carlos Vega[12], BEng; Rocío Ruiz[12], BEng, MBA; Juan Velásquez[12,13], BEng, PhD, MS; Rodrigo Vergara[14,15], PhD; Maria Isabel Behrens[6], MD, PhD; Jamileth More[1,6], PhD, MS; Carolina Delgado Derio[1,2], MD, MS

[1]Unidad de Cerebro Saludable, Hospital Clínico Universidad de Chile, Santiago, Chile

[2]Facultad de Medicina, Departamento de Neurociencia, Universidad de Chile, Santiago, Chile

[3]Núcleo Magíster en Salud de la Mujer, Facultad de Medicina y Ciencias de la Salud, Universidad Mayor, Santiago, Chile

[4]Facultad de Salud y Odontología, Escuela de Kinesiología, Universidad Diego Portales, Santiago, Chile

[5]Escuela de Psicología, Universidad de Humanismo Cristiano, Universidad de Humanismo Cristiano, Santiago, Chile

[6]Centro de Investigación Clínica Avanzada (CICA), Hospital Clínico Universidad de Chile, Santiago, Chile

[7]Instituto de Nutrición y Tecnología de los Alimentos (INTA), Universidad de Chile, Santiago, Chile

[8]Escuela de Psicología, Universidad de Los Andes, Santiago, Chile

[9]Departamento Medicina Interna, Facultad de Medicina, Universidad de La Frontera, Temuco, Chile

[10]Instituto de Neurocirugía Dr. Alfonso Asenjo, Santiago, Chile

[11]Clínica Universidad de Los Andes, Santiago, Chile

[12]Web Intelligence Centre (WIC), Departamento de Ingeniería Industrial, Universidad de Chile, Santiago, Chile

[13]Complex Engineering System Institute, Santiago, Chile

[14]Departamento de Kinesiología, Facultad de Artes y Educación Física, Universidad Metropolitana de Ciencias de la Educación, Santiago, Chile

[15]Centro Nacional de Inteligencia Artificial (CENIA), Santiago, Chile

[*]these authors contributed equally

Corresponding Author:
Carolina Delgado Derio, MD, MS
Unidad de Cerebro Saludable
Hospital Clínico Universidad de Chile
Dr. Carlos Lorca Tobar 999, Independencia
Santiago, 8380456
Chile
Phone: 56 229788260
Email: cdelgado@uchile.cl

Abstract

Background: In Latin American countries, the prevalence of noncommunicable diseases has increased rapidly in recent decades. Mobile health (mHealth) apps are now widely available at low cost and are easy to implement, offering an opportunity to encourage healthy lifestyles in older adults. However, at present, there are no mHealth apps that integrate multidomain healthy lifestyle interventions specifically adapted for older adults in Chile.

Objective: This study aims to describe the development and validation of NeoMayor, an mHealth app designed to promote healthy lifestyles as well as cardiovascular and brain health in older adults in Chile.

Methods: NeoMayor was developed iteratively with feedback from users and input from a multidisciplinary team of clinicians, researchers, and software developers. Using lean user experience methodology, we ensured user involvement throughout the design and validation process. The research was conducted in 2 phases. In the design and development phase, we created and adapted evidence-based recommendations. In the validation phase, we conducted a pilot study to assess usability, adherence, and cardiovascular health (CVH). A total of 60 functionally independent and cognitively healthy participants used the NeoMayor app

XSL•FO
RenderX

for 2 months. Clinical and cognitive assessments were conducted before and after app use. We held 26 cocreation sessions with users, consulted experts, performed a literature review, and collaborated with a team of app developers to create a functional prototype.

Results: The mean age of the participants was 71 (SD 15) years, and 85% (51/60) were female. Participants had an average of 9.8 (SD 3.6) years of education. At the end of the 2-month intervention, usability testing indicated high engagement and adherence, with participants using the app for an average of 6.6 (SD 11.85) minutes per day twice a week. Improvements were observed in global CVH, with the mean Life's Essential 8 CVH index score increasing from 64 (SD 10) to 68 (SD 11; $P<.001$). Reductions were noted in systolic blood pressure (10 mm Hg) and waist circumference (7 cm), along with increased high-density lipoprotein cholesterol levels. Participants also showed improvements in self-reported physical activity and diet, higher scores on the Short Physical Performance Battery, and faster performance times on the sit-to-stand and gait speed tests. The app was optimized for broad compatibility with Android devices, safe data collection and storage, and compliance with data privacy regulations following good clinical practices. The final product is ready for testing in a randomized controlled trial.

Conclusions: This study represents the first initiative in Chile to develop and validate an mHealth app to promote healthy lifestyles as well as cardiovascular and brain health in older adults, offering an effective, accessible, and affordable solution for promoting healthy aging in Latin American countries.

(JMIR Aging 2025;8:e71936) doi:10.2196/71936

KEYWORDS

mobile health; mHealth; healthy aging; noncommunicable diseases; brain health; dementia prevention

Introduction

Background

Accelerated population aging in recent decades, particularly in low- and middle-income countries, has led to an increase in the prevalence of noncommunicable diseases (NCDs) in adults [1,2]. Among people aged ≥70 years, cardiovascular diseases and dementia are now the leading causes of disability [3,4]. In Latin America and the Caribbean, dementia prevalence is projected to rise by 200% to 300% over the next 30 years [5]. Globally, an estimated 80% of cardiovascular disease risk factors are potentially modifiable [5,6]; in Latin America and the Caribbean, approximately 54% of dementia risk factors are also potentially modifiable [7]. This highlights the urgent need for preventive initiatives aimed at reducing the burden of cardiovascular diseases and dementia [8].

Increasing evidence supports the effectiveness of multidomain healthy lifestyle interventions—including physical activity, healthy diet, cognitive stimulation, social engagement, and health control—for treating and preventing NCDs, reducing stroke risk, and enhancing cognitive reserve [9-11]. In Chile, despite the existence of public primary prevention programs for NCDs, there has been a significant increase in the prevalence of diabetes and obesity in the past decade [12,13]. These unfavorable outcomes of Chilean prevention campaigns could be attributed to their limited outreach; low population adherence; and multiple implementation challenges [14], including health care inequities [15].

To address this implementation gap in health promotion, mobile health (mHealth) platforms have emerged as promising tools. mHealth encompasses medical practices supported by mobile devices such as smartphones, tablets, and wearable sensors. These technologies are easily accessible, widely available, have affordable costs, and are easily implementable [16-18]. Recent evidence supports the effectiveness of web-based and smartphone interventions in promoting cardiovascular health (CVH) in adults [16,19], providing insights for the design of user-friendly and accessible health technologies for older adults [20,21]. In high-income countries, numerous initiatives have been developed to promote healthy lifestyles through the use of mHealth [22]. Although many of these mHealth solutions can be easily adopted by young adults, older adults face several barriers to using mHealth, reducing its effectiveness in this demographic group [21,23]. These barriers include lower digital literacy and technological skills [24], physical and cognitive limitations [21,25], lack of trust and privacy concerns [26], and cost and technology access [23]. Regarding technology access, in Chile in 2022, 65% of older adults owned a smartphone (rising to 82% among people aged 60-69 y), and 60% had home internet access (rising to 90% among those living with younger people) [27-30]. Educational level is the most important factor related to technology use in Chile's older adults, possibly associated with technology use skills and technology access [29]. Therefore, the design of mHealth tools targeting older adults should incorporate a user-centered approach that accounts for their physical and cognitive capacities as well as their technological skills [31]. Prior research recommends solutions that are intuitive, easy to navigate, supported by health care professionals, and accessible, for example, apps with large interfaces and simplified interactions [32,33]. However, in Chile, despite relatively good smartphone access, there are currently no mHealth solutions for promoting healthy lifestyles that are adapted to older adults. Our main hypothesis was that health care content delivered through a mobile app—cocreated using participatory design methodologies and based on a multidomain intervention approach—would be easy to use and would improve CVH in older adults.

Objectives

We aim to describe the development and validation of NeoMayor, a mobile app cocreated with older adults to promote healthy lifestyle habits, support NCD management, and enhance cardiovascular and brain health in this population.

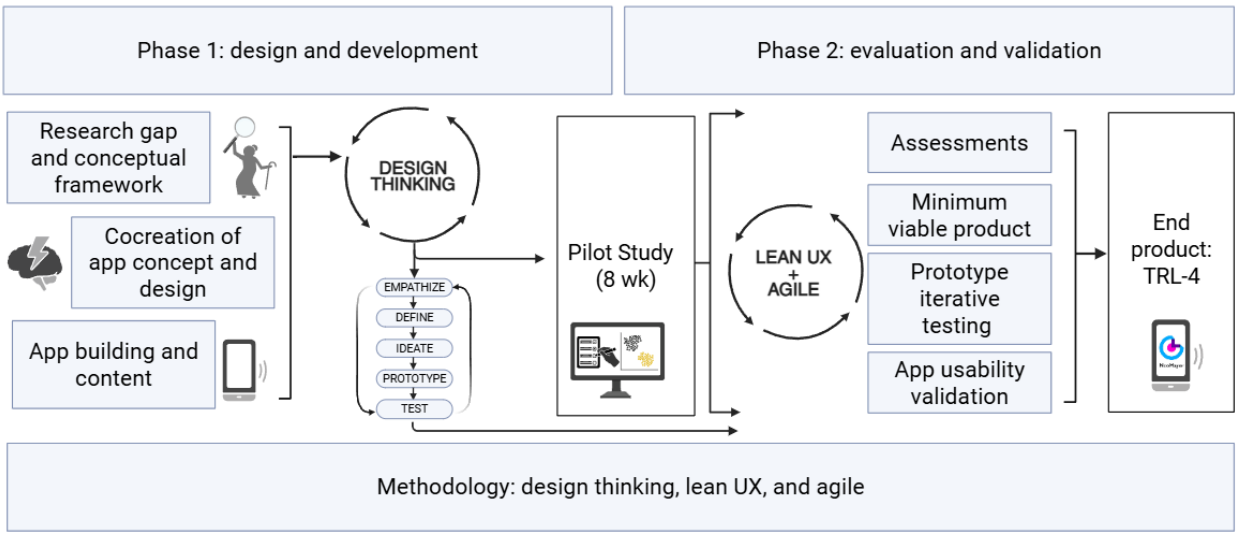
XSL•FO
RenderX

Methods

Study Design

The development and validation of NeoMayor were planned using a user-centered design approach [31]. The study was conducted in 2 phases: phase 1 (design and development with older adults, guided by lean user experience [UX] methodology) and phase 2 (prototype validation through a 2-mo pilot study in which participants used NeoMayor on their smartphones). The entire process lasted 26 months, from February 2022 to April 2024 (Figure 1 [34]).

Figure 1. The 2 phases of the NeoMayor app study. RCT: randomized controlled trial; TRL: technology readiness level; UX: user experience. Created with BioRender [34].



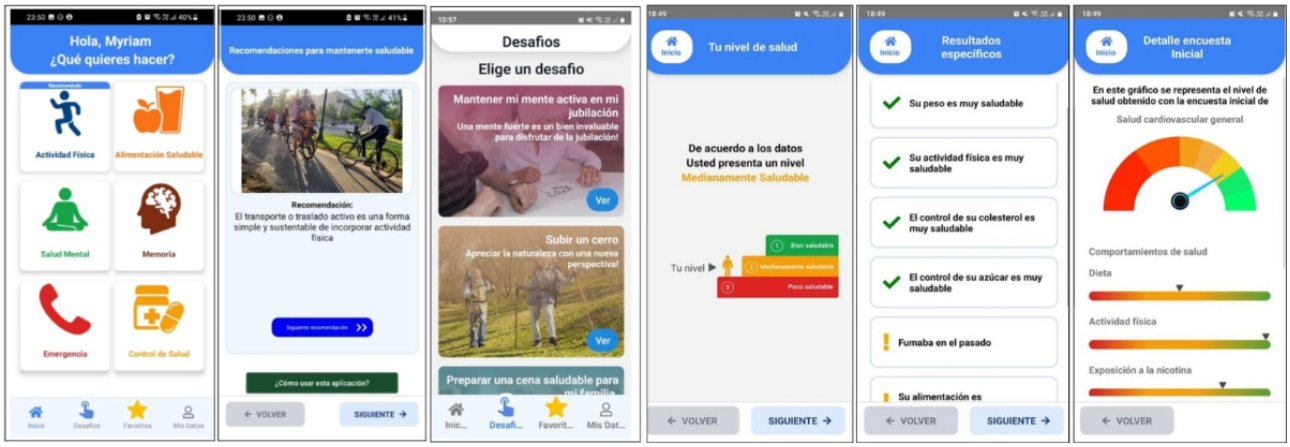
Phase 1: Design and Development

The NeoMayor app was developed to promote healthy lifestyles as well as cardiovascular and brain health in older adults in Chile. From the outset, the app was designed to provide personalized lifestyle recommendations based on each user's baseline health status. Its architecture includes a secure electronic database compliant with clinical data privacy standards [35].

The clinical multidisciplinary team, which included experts in neurology, geriatrics, physical therapy, nutrition, psychology, and biomedical sciences, led the creation of the NeoMayor app and collaborated closely with software developers to design and program the software. Using a design thinking process, we first identified user needs (empathy approach) and analyzed existing health technologies (definition approach). After gathering theoretical information, hypotheses were formulated, and proposed content from the clinical teams was validated by expert peers. Content creation was based on multiple guidelines, such as the World Health Organization (WHO) Integrated Care for Older People (ICOPE) handbook [36], the WHO guidelines for risk reduction of cognitive decline and dementia [36,37], the Chilean Cardiovascular Health Program [38], and the American Heart Association's Life's Essential 8 [39]. Life's Essential 8 provides self-care recommendations to manage 8 modifiable risk factors for cardiovascular disease by adopting healthy lifestyle behaviors: eat better, be more active, quit tobacco, get healthy sleep, manage weight, control cholesterol, manage blood glucose, and manage blood pressure. Each risk factor can be measured on a semiquantitative scale from 0 (poorest) to 100 (healthiest), with the average corresponding to the global CVH index [25]. We selected the CVH index as our primary outcome because it has been validated as a predictor of cardiovascular events [40], brain health, and dementia risk [41]. In addition, because it comprises only modifiable risk factors, the CVH index can change over time, making it suitable for tracking progress in CVH [39]. We adapted the CVH index into a mobile questionnaire that provides feedback on health status and tailors recommendations accordingly. Participants are categorized as optimal (score: 79-100), moderate (score: 50-79), or poor (score: <50), using graphic images (Figure 2A). The app recommends periodic assessments to monitor changes in CVH over time.

XSL•FO
RenderX

Figure 2. Screens of the NeoMayor app. (A) Categorization of cardiovascular health. (B) Initial menu and goal-directed navigation.



The lifestyle content is structured into 6 modules: physical activity, healthy eating, mental health [42,43], memory, health control [6,38,39] (Figure 2B; based on guidelines from the WHO [44,45], the Chilean Ministry of Health [38], and Life's Essential 8 [39]), and "emergency" (incorporated at the request of participants, who wanted quick access to emergency numbers and official resources on legal and health benefits). Content is delivered through video, audio, and educational images. Currently, the app features a repository of 33 videos, each approximately 5 minutes long; 30 audio recordings, each lasting 15 minutes; and 50 educational images. The app includes 2 navigation options: the first is free navigation, in which users can choose from 6 modules on the menu screen; and the second is challenge-based navigation, in which users set a specific goal (eg, "being able to climb a hill"), prompting customized, targeted training suggestions that combine content from multiple modules over a defined period (Figure 2B).

The next step was NeoMayor's iterative refinement and development, guided by a combination of lean UX [46], agile [47], and design thinking methodologies [48]. The selection of lean UX and design thinking was intentional, given their strong alignment with the specific challenges and needs of older adults. These user-centered frameworks facilitate continuous feedback integration, empathy-based exploration of UX, and rapid prototyping, all of which are essential for optimizing usability and ensuring the app's relevance for older adults with medium to high digital literacy [49]. Lean UX enabled rapid iteration to refine accessibility features such as color contrast, touch target size, and interface clarity [50]. In parallel, design thinking fostered cocreation through structured empathy and ideation stages, ensuring that the final product resonated with the lived experiences of older users [51]. Low-fidelity prototypes of the NeoMayor app were iteratively refined to emphasize accessibility, intuitive navigation, and content clarity.

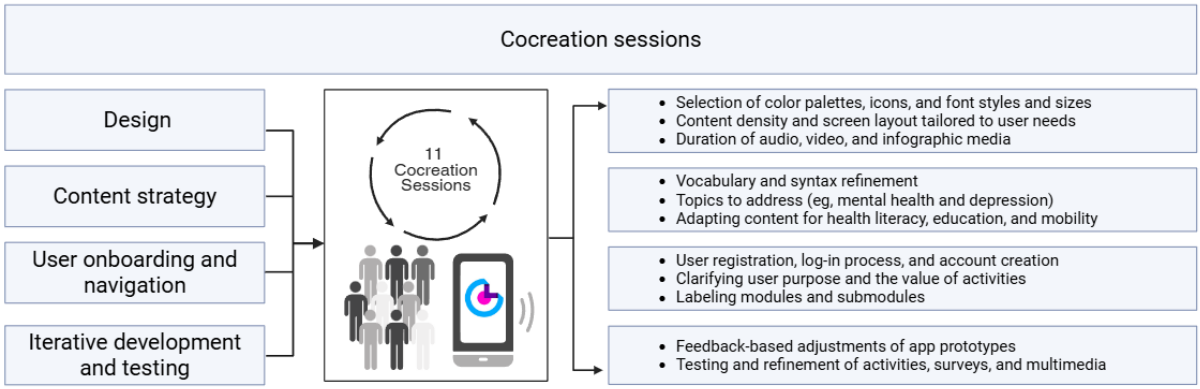
The development team followed agile Scrum practices, with biweekly sprints and sprint retrospectives involving clinicians and software engineers. Tasks were managed using Trello, and code repositories were version-controlled using Git and GitHub. Continuous integration was achieved using GitHub Actions, which facilitated automated testing and deployment to staging environments before release to the Play Store.

As part of this cocreative and coconstructive process [52], we implemented several interface adjustments, including high-contrast colors, large touch targets, simplified gestures, and accommodations for motor and cognitive needs. This combined approach enabled the development of a tool that is not only clinically robust but also practical, intuitive, and acceptable to older adults across a spectrum of technological familiarity. Classification algorithms and user surveys further guided customization of the UX.

We held 26 cocreation sessions with 30 older adults (female: n=21, 70%; male: n=9, 30%; age: range 60-75 y) to validate the language, content, and user interface elements (Figure 3 [53]). These sessions provided critical feedback on layout, navigation, and instructional clarity, which shaped the app's design.

XSL•FO
RenderX

Figure 3. The NeoMayor iterative participative design process. Created with BioRender [53].



After the content was finalized, the software development team generated both front-end and back-end components. The app was built using the open-source React Native framework for cross-platform compatibility on Android devices and a Django backend on servers maintained by the Web Intelligence Center for secure data handling. Once compiled with Gradle, the NeoMayor app was validated through the Google Play Console and released on the Google Play Store as a downloadable app.

The app deployment was managed through the Google Play Console, with version tracking and staged rollouts to monitor performance and crash reports. Updates were issued using semantic versioning (eg, v1.0.0) and were preceded by internal QA cycles. A dedicated beta testing group of older adults provided feedback on each release candidate.

Throughout the iterative design process, the team (users, clinicians, designers, and engineers) validated the app's content, navigation flow, and overall UX. As suggested by pivotal software development literature [54], we conducted micro-usability tests to evaluate ease of navigation, icon comprehension, and text readability. These tests provided insights on interface improvements, leading to high-contrast designs, labeled icons, and enlarged touch targets (Figure 3). Thus, the NeoMayor app was evaluated at the prototype level in terms of usability, functionality, and efficacy, achieving positive results across all evaluated domains. After 12 months, we obtained a minimum viable product: a mobile app that classifies users according to CVH and mobility and offers personalized content organized into 6 modules.

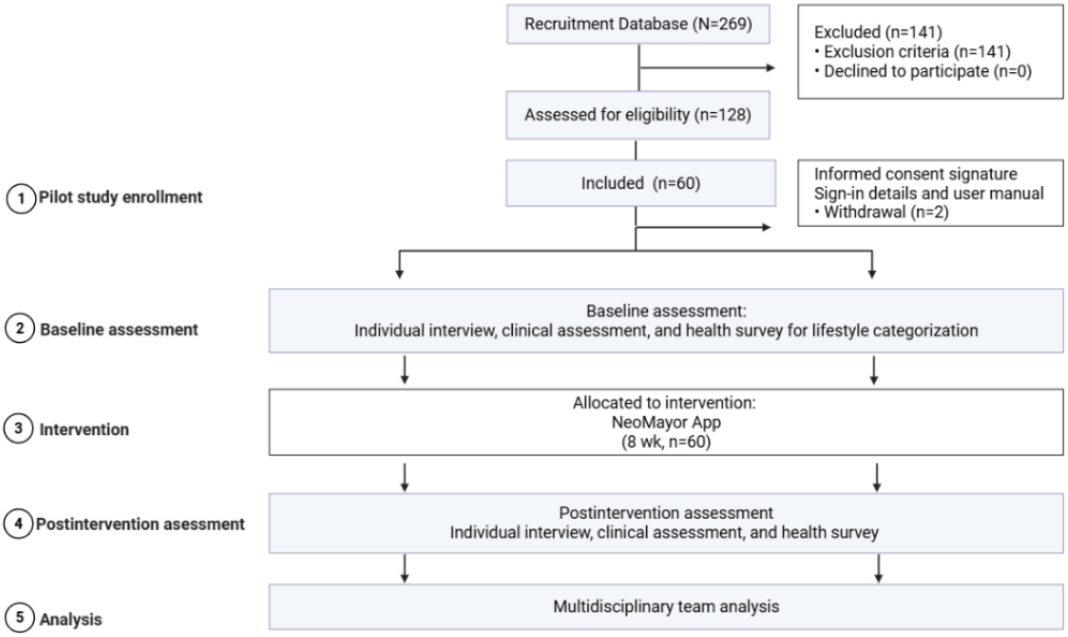
Phase 2: Evaluation and Validation (Pilot Study)

Participants

We recruited functionally independent older adults aged 60 to 77 years living in Santiago, Chile, who were able to read, write, and use technology (digital skills sufficient to independently use a mobile smartphone; Figure 4 [55]). Participants were recruited through voluntary sign-up notices from ambulatory programs at the Hospital Clínico de la Universidad de Chile, civil society representatives of the Servicio Nacional de Personas Mayores, and the advisory committee of the Instituto Nacional de Geriatría. Printed invitation posters were also distributed in various community and public spaces across Santiago. To qualify for inclusion, participants were required to have a moderate to high cardiovascular risk, defined as the presence of at least three of the following cardiovascular risk factors: (1) hypertension (blood pressure >140/90 mm HG or receiving pharmacological treatment), (2) diabetes or fasting glycemia level of >100 mg/dL, (3) obesity (BMI ≥30 kg/m^2), (4) current smoking status, and (5) elevated total cholesterol level (>240 mg/dL). In addition, participants had to be free of dementia, defined as a total score of ≥8 on the memory, fluency, and orientation test, which is a 5-minute screening test for cognitive decline validated for the Chilean population [56]. The exclusion criteria included significant impairments in independent mobility, substantial auditory or visual impairments, neurodegenerative diseases, untreated psychiatric disorders or psychological diagnoses, and the presence of a severe or terminal illness. Participant recruitment began 1 month before the initiation of the pilot study. Different participants were involved in the 2 phases of the NeoMayor design: those engaged in cocreation sessions, design, and development (n=30), described in the previous section; and those who participated in the pilot study (n=60).

XSL•FO
RenderX

Figure 4. The NeoMayor pilot study flowchart. Created with BioRender [55].



Pilot Methodology

A database of 269 individuals who expressed interest in participating in the pilot study was established, and we contacted 128 (47.6%) by telephone to assess their eligibility based on the inclusion and exclusion criteria. Of these 128 individuals, 62 (48.4%) met the inclusion criteria. Subsequently, 2 (3%) of these 62 participants withdrew voluntarily for personal reasons, leaving a final sample of 60 (97%) participants. Participants then underwent face-to-face interviews and received a complete clinical, physical, and cognitive assessment at baseline and after the 2-month NeoMayor app use period. Trained professionals (physicians, physical therapists, and psychologists) performed the assessments. For consistency, the same evaluator performed the pre- and postpilot assessments, blinded to adherence results. The clinical assessment lasted approximately 1 hour and included a structured interview to collect demographic and sociocultural data, current health conditions, and medication use; measurement of biometric data, including anthropometry; assessment of blood pressure; blood tests (glucose and lipid levels); and structured questionnaires to assess lifestyle habits adapted for Latin America [11]: the Pittsburgh Sleep Quality Index [57], the Mediterranean-DASH Diet Intervention for Neurodegenerative Delay Diet Scale [58], and the International Physical Activity Questionnaire [59]. By using these data, we calculated the CVH index, including the total score and 8 subscores [39]. The cognitive evaluation included the administration of the Mini-Mental State Examination [60]; the Geriatric Depression Scale [61]; self-perception of memory scales (MCA) [62]; and the memory, fluency, and orientation test, which is a 5-minute screening test for cognitive impairment [56]. The physical performance assessment included the use of the handgrip test; the 5-repetition chair stand test, as recommended by the WHO ICOPE handbook [36]; the timed up and go test (both simple and dual tasks); and the Short Physical Performance Battery (SPPB) [63].

The NeoMayor app was installed on participants' smartphones, and they received both training in its use and a printed user manual. For those who did not have a smartphone, the project provided one. All participants used the NeoMayor app for 2 months. In case of significant problems with the app or smartphones, assistance was provided. After 3 weeks, we contacted half of the users (30/60, 50%) via telephone to address any problems in using the app. Finally, to evaluate the pilot implementation, we used the RE-AIM (Reach, Effectiveness, Adoption, Implementation, and Maintenance) [64] methodology. At the conclusion of the pilot study, all participants completed surveys to provide qualitative data about usability, user satisfaction, and content preferences. Further assessments will be included in a randomized controlled trial (RCT).

Data Management

The data were collected in a REDCap (Research Electronic Data Capture; Vanderbilt University) database. All data analyses were conducted using R (version 4.0.0, 64-bit; R Foundation for Statistical Computing) and SPSS software (version 27.0; IBM Corp). After checking the distribution of the data, we performed descriptive statistics. Continuous variables were expressed as mean (SD) and 95% CI, and categorical variables were reported as percentages and 95% CIs. The potential effect of NeoMayor app use on participants' CVH was measured using a paired sample *t* test comparing the CVH index before and after app use (within-group analysis) [65]. We calculated the effect size of within-group changes by using Cohen *d*.

Finally, we categorized the participants according to whether they showed improvement. Improvement was defined as significant if the difference between the CVH index before and after app use was >1 SD from the baseline CVH index. We then performed a binary logistic regression to establish which variables were associated with the improvement, using improvement (yes or no) as the dependent variable, with age, educational level, sex, baseline depressive symptoms, and baseline CVH index as independent variables.

XSL•FO
RenderX

Ethical Considerations

The study was approved by the scientific and research ethics committee of the Hospital Clínico de la Universidad de Chile (Acta 36, OAIC 1220/21, FONDEF ID21I10096). All study participants provided signed informed consent.

Results

Demographic and Biomedical Data

A total of 60 older adults were included in the pilot study. The participants' mean age was 71 (SD 15) years, and they had an average of 9.8 (SD 3.6) years of education. Of the 60 participants, 51 (85%) were female. Most of the participants used the internet at home (57/60, 95%) and owned an Android smartphone (49/60, 82%). We provided a smartphone to 11 (18%) of the 60 participants (n=7, 64% owned an iPhone; n=4, 36% did not have their own smartphone, or their smartphone was too old). The baseline characteristics of participants are detailed in Table 1.

Table 1. Participants' baseline characteristics (n=60).

Characteristics	Values
Age (y), mean (SD)	71 (15)
Sex, n (%)	
Female	51 (85)
Male	9 (15)
Education: ≥12 y, n (%)	37 (62)
High blood pressure, n (%)	35 (58)
Obesity, n (%)	28 (47)
Dyslipidemia, n (%)	47 (78)
Tobacco use, n (%)	14 (23)
Hypothyroidism, n (%)	14 (23)
Mild vision problems, n (%)	60 (100)
Mild hearing problems, n (%)	1 (2)
Heart failure, n (%)	4 (7)
Chronic respiratory disease, n (%)	12 (20)
Diabetes, n (%)	14 (23)
Lives alone, n (%)	13 (22)
Has internet access at home, n (%)	57 (95)
Knows how to use a smartphone, n (%)	56 (93)
Has Android smartphone, n (%)	49 (82)

Validation of the App's Assessment

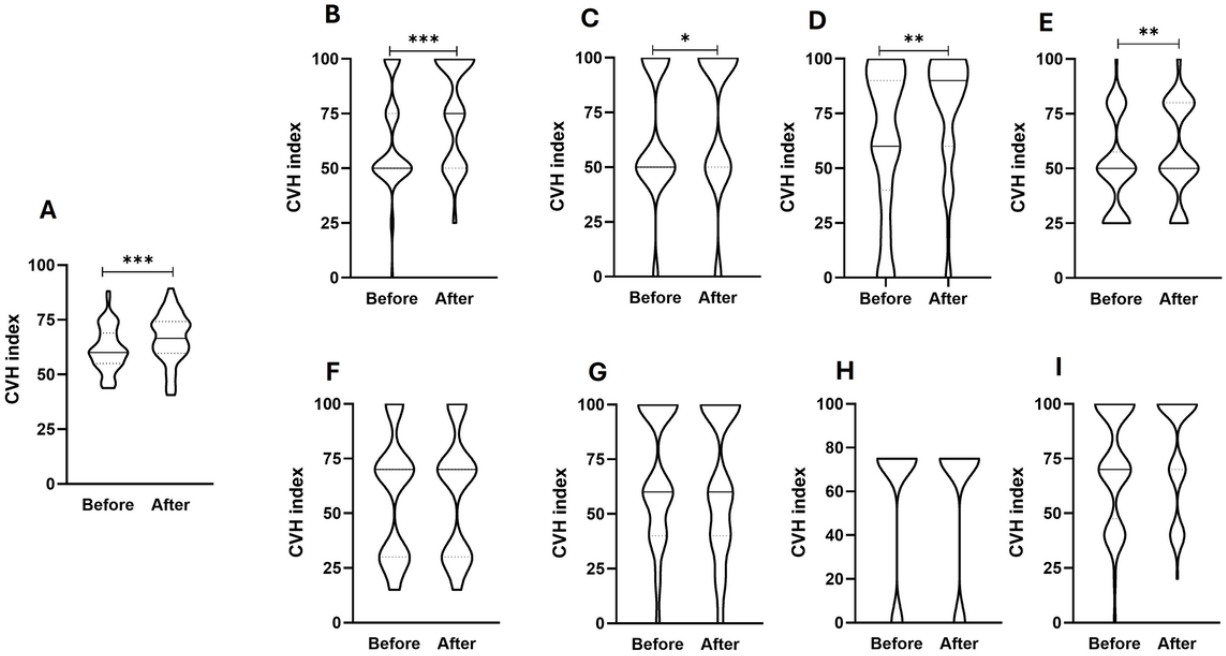
To validate the CVH categorization, we compared the NeoMayor app categorization and the face-to-face clinical assessment, which yielded a κ index of 0.51 (P<.001), indicating a moderate level of agreement between the 2 evaluations. In addition, the Pearson correlation between the 2 assessments (r=0.692; P=.001) indicated a good correlation for the total CVH index score. At baseline, our results indicate that 54 (90%) of the 60 participants had intermediate CVH, 4 (7%) had low CVH, and 4 (7%) had optimal CVH.

Potential Effectiveness

The average CVH index score at baseline was 64.3 (SD 10.8), which improved significantly to 69.0 (SD 11.03) after 2 months of NeoMayor app use (P<.001; Figure 5A), resulting in a moderate effect size (Cohen d=−0.548, −2.55 to 1.46). Participants showed significant improvements in 4 of the 8 subscores of the CVH index: blood pressure (Figure 5B), blood glucose level (Figure 5C), physical activity (Figure 5D), and diet (Figure 5E). However, no significant changes were observed in BMI, blood lipids, tobacco use, and sleep (Figures 5F-5I; Table 2). Indeed, the number of participants who smoked remained unchanged. Next, we classified the participants into those who showed improvement and those who did not: 32% (19/60) showed improvement. There were no significant predictors of improvement after testing the effects of age, sex, educational level, baseline depressive symptoms, and baseline CVH index, but there was a tendency for a negative effect of higher baseline depressive symptoms (odds ratio 0.69, 95% CI 0.48–1.00; P=.05) and healthier baseline CVH status (odds ratio 0.94, 95% CI 0.88–1.00; P=.06) on the possibility of improvement.

XSL•FO
RenderX

Figure 5. Changes in the cardiovascular health (CVH) index after NeoMayor use. (A) Total Life's Essential 8. (B) Blood pressure. (C) Blood glucose level. (D) Physical activity. (E) Diet. (F) BMI. (G) Blood lipids. (H) Tobacco use. (I) Sleep. Violin plots compare the global CVH index and its subscores before and after NeoMayor use. The solid lines represent the median, and the dashed lines represent the quartiles. A paired sample t test compared CVH index differences. *P<.05, **P<.005, ***P<.001.



Regarding other results, we observed changes in anthropometric measurements, with an average weight reduction of 1 kg and a decrease in waist circumference of 7 cm, both indicative of improved body fat distribution. Physical performance showed improvement, with faster gait speed (an average reduction of 2 s), enhanced upper limb strength (measured by the handgrip test), increased lower limb strength (measured by the 5-repetition chair stand test), and improved SPPB scores. In addition, a significant reduction of 10 mm Hg in systolic blood pressure was noted. Blood tests revealed a decrease in high-density lipoprotein cholesterol levels, although low-density lipoprotein cholesterol and triglyceride levels showed nonsignificant increases. There were no differences in cognitive assessment or depressive symptoms, but there was an improvement in self-perceived memory after app use (Table 2).

XSL•FO
RenderX

Table 2. Comparison of variables assessed before and after the NeoMayor pilot study (n=60).

Variables	A: before pilot, mean (SD)	B: after pilot, mean (SD)	*P* value: A versus B[a]
Physical examination			
Weight (kg)	72.0 (14.0)	71.0 (17.0)	.53
Waist circumference (cm)	99.0 (11.0)	92.0 (12.0)	*<.001* [b]
Systolic blood pressure (mm Hg)	131.0 (17.0)	121.0 (14.0)	*<.001*
Diastolic blood pressure (mm Hg)	72.0 (14.0)	72.0 (14.0)	.57
Handgrip test (kg)	25.0 (6.0)	26.0 (6.0)	.48
5-repetition chair stand test time (s)	10.0 (2.0)	8.0 (2.0)	*<.001*
SPPB[c] (score: 0-12)	11.3 (1.1)	11.9 (0.5)	*<.001*
Timed up and go simple task (s)	10.0 (2.0)	8.0 (2.0)	*<.001*
Timed up and go dual task (s)	11.0 (2.0)	9.0 (2.0)	*<.001*
Blood tests			
Fasting blood glucose level	107.0 (26.0)	104.0 (24.0)	.85
Fasting blood total cholesterol level	192.0 (36.0)	198.0 (43.0)	.17
Fasting blood HDL[d] cholesterol level	57.0 (15.0)	52.0 (14.0)	*.005*
Fasting blood LDL[e] cholesterol level	109.0 (34.0)	117.0 (39.0)	.45
Fasting blood triglycerides level	134.0 (52.0)	146.0 (67.0)	.43
Life's Essential 8			
CVH[f] index	63.0 (10.8)	69.0 (11.3)	*<.001*
Blood pressure	63.7 (23.2)	77.5 (23.3)	*<.001*
Blood lipids	69.0 (29.5)	65.0 (32.4)	.18
Blood glucose	65.8 (29.8)	73.3 (32.5)	*.04*
BMI	58.2 (28.2)	59.6 (28.1)	.38
Physical activity	62.2 (34.4)	76.3 (27.4)	*.003*
Diet	49.8 (21.1)	57.3 (21.6)	*.006*
Tobacco use	58.7 (31.1)	58.7 (31.1)	>.99
Sleep	73.8 (26.1)	79.6 (25.1)	.07
Questionnaires and cognitive test scores			
MIND[g] food score	8.0 (2.0)	8.0 (6.0)	.49
IPAQ[h] score	1.8 (0.6)	2.4 (0.6)	*<.001*
PSQI[i] (self-reported sleep duration)	6.4 (1.0)	6.4 (1.0)	.74
PSQI score	8.0 (4.0)	8.0 (4.0)	.75
MCA[j]	5.0 (3.0)	3.0 (2.0)	*.02*
GDS[k]	2.0 (2.0)	2.0 (3.0)	.08
MEFO[l]	12.0 (1.0)	12.0 (1.0)	.14

[a]Paired sample *t* test with equal variances.

[b]Italicization indicates values that met the significance threshold ($P<.05$).

[c]SPPB: Short Physical Performance Battery.

[d]HDL: high-density lipoprotein.

[e]LDL: low-density lipoprotein.

[f]CVH: cardiovascular health.

[g]MIND: Mediterranean-DASH Diet Intervention for Neurodegenerative Delay.

XSL•FO
RenderX

[h]IPAQ: International Physical Activity Questionnaire.

[i]PSQI: Pittsburgh Sleep Quality Index.

[j]MCA: self-perception of memory scales.

[k]GDS: Geriatric Depression Scale.

[l]MEFO: memory, fluency, and orientation test.

Qualitative Results

Overview

The usability as well as barriers and facilitators to using the NeoMayor app were evaluated through an iterative process involving the clinical team and cocreation session participants. Among the participants, 88% (53/60) stated that their primary motivation for using the app was "to improve overall health," followed by improving physical fitness and enhancing memory. In addition, 79% (47/60) of the participants considered improving their lifestyle especially important, and the same proportion believed that the app would help them achieve this goal. Regarding usability, 85% (51/60) of the participants found the app easy and intuitive to use. For the initial health survey, 90% (54/60) considered the personalized results delivered by the survey useful and easy to understand. In terms of user satisfaction, 66% (39/60) found the app highly engaging, and 64% (38/60) were very satisfied with their UX. Moreover, 63% (38/60) rated the app as highly motivating and found the app intuitive to use.

Adherence

To assess study adherence, we monitored app use during the first month. Participants spent an average of 12,242.6 (SD 2359.0) seconds using the app, equivalent to approximately 204.0 (SD 5.9) minutes. This corresponds to approximately 400 seconds (6.7 min) of app use per day.

Discussion

Principal Findings

In this study, we described the design, development, and validation of NeoMayor, the first mHealth app aimed at promoting healthy lifestyles as well as cardiovascular and brain health in older adults in Chile through multidomain healthy lifestyle interventions. The NeoMayor app categorizes users' CVH profiles based on a survey inspired by Life's Essential 8, which was adapted and validated in this pilot study. The app then delivers personalized recommendations and content covering physical activity, nutrition, mental health, memory, and general health management for healthy brain promotion [66], providing a personalized guide to healthier living. This structured and systematic development process, combined with the validation of the content and design, allows for potential replication in other mHealth apps, particularly for populations with limited digital literacy or access to health technologies [52]. These findings also align with previously described barriers and facilitators to adopting mHealth-mediated strategies [21,26,27].

Although several international initiatives exist to promote health care in older adults, most have been developed in high-income countries and are not designed for populations with lower educational levels, who require special adaptations and cultural validation to meet their needs [67]. In this context, NeoMayor is the first app specifically adapted for Chilean older adults to support integral health care promotion.

Regarding adherence, the average daily use time of the NeoMayor app was notably higher (6.7 min/d) than that reported in other studies of older adults using mobile apps (2 min/d), indicating a high level of engagement [68], which could be associated with the participants' high technological skills and motivation. One important factor to consider in the adoption of mHealth technology is technology-related anxiety [69] in people with limited digital literacy, which may require positive management to encourage app use [17]. Although this study included participants with functional independence, cognitive ability, and basic digital skills, broader implementation for older people with low digital literacy or cognitive issues [70] may require alternative strategies. These could include health care professional– or coach-assisted mHealth interventions to overcome moderate mHealth literacy or social digital barriers to engagement and empower health literacy [71].

The findings of the pilot study suggest that using the NeoMayor app led to improvements in both metabolic health and lifestyle habits, which are key indicators of the app's potential effectiveness [72]. The improvements observed in blood pressure (a reduction of 10 mm Hg in systolic blood pressure), blood glucose levels, physical activity, and diet—key elements in the prevention of cardiovascular disease and other NCDs—represent a relevant finding that should be confirmed in a larger-scale RCT [65]. Although we did not find significant changes in BMI, participants experienced a significant reduction in waist circumference, which is more closely associated with cardiovascular risk than BMI [73]. This is a very important result, considering that 47% (28/60) of the participants were obese.

The significant enhancement in functional physical performance, measured by the SPPB [74], gait speed alone, and the 5-repetition chair stand test alone, is relevant, considering its critical importance in reducing the risk of physical fragility and falls in older adults [75]. Moreover, the 2-second reduction in the 5-repetition chair stand test represents a statistically significant and clinically relevant change among older Chileans, which is associated with higher survival rates [76]. These improvements in physical performance may be attributed to the physical activity module, which includes instructional videos, consistent with findings from previous research on the impact of video-guided physical activity interventions [52,77]. In addition, these improvements contribute to enhanced metabolic health. Furthermore, we found an increase in self-perceived memory, which could be associated with both physical activity and the global cognitive activity promoted by the use of the app. Previous studies report that physical activity combined with cognitive training can improve self-perceived memory in

XSL•FO
RenderX

older adults [78], and exercise is a main component of cognitive health [79].

As an additional strength, usability testing demonstrated easy user navigation, with a main menu that provided intuitive access to the app's modules. Participants reported a preference for challenge-based navigation, which allowed them to set specific health goals and stay motivated. However, minor modifications should be addressed in future iterations of the app, including more engaging navigation elements, considering the importance of structural measures to support lifestyle changes [14,80].

Limitations

This study included only participants without sensory or cognitive impairments and with good technological skills. Future adaptations could expand the intervention's applicability by considering the co-design of an inclusive version of the app tailored to varying levels of digital literacy. In addition, because most of the participants were female—a common proportion in similar studies [81]—the generalizability validity of our findings may be limited. This is a frequent issue in pilot studies, where feasibility is the main objective, and could be overcome in future studies using the NeoMayor app in population-based samples. As a pilot study, the methodology used cannot establish the efficacy of NeoMayor in CVH independent of confounding factors, but it provides a solid foundation for a subsequent RCT with strategies to overcome sample bias. Indeed, with the implementation of the design improvements, the NeoMayor app has reached technology readiness level 4 (Figure 2). The next phase will involve evaluating the effectiveness of NeoMayor on CVH and lifestyle in an RCT, which should include at least 200 participants with different digital literacy and technological skills, based on sample size estimation considering the moderate effect size with regard to CVH observed in this pilot.

Conclusions

This study described the design, development, and validation of a comprehensive mHealth app that aims to promote healthy lifestyles as well as cardiovascular and brain health in older adults in Chile. Using a person-centered, multidisciplinary approach grounded in lean UX and agile methodology for co-design and codevelopment, NeoMayor demonstrated good feasibility and potential effectiveness in improving CVH and physical performance, which will be evaluated soon in a larger RCT. These findings indicate that the NeoMayor app can provide an accessible, affordable, and implementable strategy to support healthy aging.

Acknowledgments

The authors thank the NeoMayor team, the Healthy Brain Unit personnel at Hospital Clínico de la Universidad de Chile in Santiago, the Web Intelligence Centre team at the Universidad de Chile in Santiago, and all participants in the cocreation and testing sessions. This research was funded by a grant from the Chilean Fund for the Promotion of Scientific and Technological Development (ID2110096 [CDD]) and the Chilean National Agency for Research and Development grant for doctoral studies (21221932 [MG]) and Centro Nacional de Inteligencia Artificial CENIA (FB210017, BASAL, ANID [RV]). The article processing charge was funded by Universidad Mayor (MG). Figures 1, 3, and 4 were created using BioRender (BioRender Inc).

Authors' Contributions

CDD conceptualized the study. CDD, J More, and CDS were responsible for methodology. VDC, CA, CV, JV, V Cortes, and PV were responsible for software. CDD, J More, FS, and MIB were responsible for validation. CDD, MG, V Cabello, and CDSM were responsible for formal analysis. MT, CDD, V Cabello, MG, CM, GM, CDSM, CDS, RB, DP, and NR were responsible for investigation. CDD and MIB were responsible for resources. CDD, J Mauro, CDS, and MG were responsible for data curation. MG wrote the original draft. MG, V Cabello, J More, J Mauro, and CDD reviewed and edited the manuscript. CDD supervised the project. CDD and MT were responsible for project administration. CDD, MG, and RV were responsible for funding acquisition. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Moreno X, Lera L, Albala C. Disability-free life expectancy and life expectancy in good self-rated health in Chile: gender differences and compression of morbidity between 2009 and 2016. PLoS One 2020 Apr 30;15(4):e0232445 [FREE Full text] [doi: 10.1371/journal.pone.0232445] [Medline: 32353089]
2. Albala C, Sánchez H, Lera L, Angel B, Cea X. Socioeconomic inequalities in active life expectancy and disability related to obesity among older people. Rev Med Chil 2011 Oct;139(10):1276-1285 [FREE Full text] [Medline: 22286726]
3. Global burden of disease 2021: findings from the GBD 2021 Study. Institute for Health Metrics and Evaluation. 2024. URL: https://www.healthdata.org/research-analysis/library/global-burden-disease-2021-findings-gbd-2021-study [accessed 2025-08-27]
4. Sousa RM, Ferri CP, Acosta D, Albanese E, Guerra M, Huang Y, et al. Contribution of chronic diseases to disability in elderly people in countries with low and middle incomes: a 10/66 Dementia Research Group population-based survey. The Lancet 2009 Nov;374(9704):1821-1830. [doi: 10.1016/s0140-6736(09)61829-8]

XSL•FO
RenderX

5. Paradela RS, Calandri I, Castro NP, Garat E, Delgado C, Crivelli L, et al. Population attributable fractions for risk factors for dementia in seven Latin American countries: an analysis using cross-sectional survey data. The Lancet Glob Health 2024 Oct;12(10):e1600-e1610. [doi: 10.1016/s2214-109x(24)00275-4]
6. Roth GA, Mensah GA, Fuster V. The global burden of cardiovascular diseases and risks: a compass for global action. J Am Coll Cardiol 2020 Dec 22;76(25):2980-2981 [FREE Full text] [doi: 10.1016/j.jacc.2020.11.021] [Medline: 33309174]
7. Baker LD, Snyder HM, Espeland MA, Whitmer RA, Kivipelto M, Woolard N, et al. Study design and methods: U.S. study to protect brain health through lifestyle intervention to reduce risk (U.S. POINTER). Alzheimers Dement 2024 Feb 30;20(2):769-782 [FREE Full text] [doi: 10.1002/alz.13365] [Medline: 37776210]
8. Mukadam N, Wolters FJ, Walsh S, Wallace L, Brayne C, Matthews FE, et al. Changes in prevalence and incidence of dementia and risk factors for dementia: an analysis from cohort studies. The Lancet Public Health 2024 Jul;9(7):e443-e460. [doi: 10.1016/s2468-2667(24)00120-8]
9. Ngandu T, Lehtisalo J, Solomon A, Levälahti E, Ahtiluoto S, Antikainen R, et al. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. The Lancet 2015 Jun;385(9984):2255-2263. [doi: 10.1016/s0140-6736(15)60461-5]
10. Kivipelto M. World-wide fingers network: a global approach to risk reduction and prevention of dementia. J Neurol Sci 2019 Oct;405:18-19. [doi: 10.1016/j.jns.2019.10.050]
11. Crivelli L, Calandri IL, Suemoto CK, Salinas RM, Velilla LM, Yassuda MS, et al. Latin American Initiative for Lifestyle Intervention to Prevent Cognitive Decline (LatAm-FINGERS): study design and harmonization. Alzheimers Dement 2023 Sep;19(9):4046-4060. [doi: 10.1002/alz.13101] [Medline: 37204054]
12. Parra-Soto S, Leiva-Ordoñez AM, Troncoso-Pantoja C, Matus-Castillo C, Petermann-Rocha F, Martínez-Sanguinetti MA, et al. Association of adiposity and diabetes mellitus type 2 by education level in the Chilean population. Rev Med Chil 2021 Jun;149(6):819-828 [FREE Full text] [doi: 10.4067/s0034-98872021000600819] [Medline: 34751340]
13. Encuesta Nacional de Salud 2016-2017: Primeros resultados. Ministerio de Salud (Minsal), Subsecretaría de Salud Pública, División de Planificación Sanitaria, Departamento de Epidemiología. 2018. URL: https://redsalud.ssmso.cl/wp-content/uploads/2018/02/ENS-2016-17_PRIMEROS-RESULTADOS-ilovepdf-compressed.pdf [accessed 2025-09-15]
14. Valentino G, Vio F, Rodriguez-Osiac L. Analysis of the Chilean health promotion policy "choose a healthy lifestyle". Rev Med Chil 2023 Feb;151(1):42-51 [FREE Full text] [doi: 10.4067/s0034-98872023000100042] [Medline: 37906745]
15. Legaz A, Altschuler F, Gonzalez-Gomez R, Hernández H, Baez S, Migeot J, et al. Structural inequality linked to brain volume and network dynamics in aging and dementia across the Americas. Nat Aging 2025 Feb 27;5(2):259-274. [doi: 10.1038/s43587-024-00781-2] [Medline: 39730822]
16. Richard E, Moll van Charante EP, Hoevenaar-Blom MP, Coley N, Barbera M, van der Groep A, et al. Healthy ageing through internet counselling in the elderly (HATICE): a multinational, randomised controlled trial. The Lancet Digital Health 2019 Dec;1(8):e424-e434. [doi: 10.1016/s2589-7500(19)30153-0]
17. Hafdi M, Eggink E, Hoevenaar-Blom MP, Witvliet MP, Andrieu S, Barnes L, et al. Design and development of a mobile health (mHealth) platform for dementia prevention in the prevention of dementia by mobile phone applications (PRODEMOS) project. Front Neurol 2021 Dec 16;12:733878 [FREE Full text] [doi: 10.3389/fneur.2021.733878] [Medline: 34975710]
18. mHealth: new horizons for health through mobile technologies. World Health Organization. URL: https://www.afro.who.int/publications/mhealth-new-horizons-health-through-mobile-technologie [accessed 2025-08-27]
19. Van Charante EM, Hoevenaar-Blom MP, Song M, Andrieu S, Coley N, Georges J, et al. Prevention of dementia using mobile phone applications (PRODEMOS): a multinational, randomised, controlled effectiveness-implementation trial. Lancet Healthy Longev 2024 Jun;5(6):e431-e442. [doi: 10.1093/eurheartj/ehae666.3555] [Medline: 38763155]
20. Kim S, Gajos KZ, Muller M, Grosz BJ. Acceptance of mobile technology by older adults: a preliminary study. In: Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services. 2016 Presented at: MobileHCI '16; September 6-9, 2016; Florence, Italy. [doi: 10.1145/2935334.2935380]
21. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556 [FREE Full text] [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
22. Eggink E, Hafdi M, Hoevenaar-Blom MP, Song M, Andrieu S, Barnes LE, et al. Prevention of dementia using mobile phone applications (PRODEMOS): protocol for an international randomised controlled trial. BMJ Open 2021 Jun 09;11(6):e049762 [FREE Full text] [doi: 10.1136/bmjopen-2021-049762] [Medline: 34108173]
23. Barreda Gutiérrez M, Cantarero-Prieto D, Pascual Sáez M. Age, technology, and the digital divide: are they directly related to mental health problems? Healthcare (Basel) 2024 Dec 05;12(23):2454 [FREE Full text] [doi: 10.3390/healthcare12232454] [Medline: 39685075]
24. Cajita MI, Hodgson NA, Lam KW, Yoo S, Han HR. Facilitators of and barriers to mHealth adoption in older adults with heart failure. Comput Inform Nurs 2018 Aug;36(8):376-382 [FREE Full text] [doi: 10.1097/CIN.0000000000000442] [Medline: 29742549]

XSL•FO
RenderX

25. Mitzner TL, Boron JB, Fausset CB, Adams AE, Charness N, Czaja SJ, et al. Older adults talk technology: technology usage and attitudes. Comput Human Behav 2010 Nov 01;26(6):1710-1721 [FREE Full text] [doi: 10.1016/j.chb.2010.06.020] [Medline: 20967133]
26. Wildenbos GA, Peute L, Jaspers M. Aging barriers influencing mobile health usability for older adults: a literature based framework (MOLD-US). Int J Med Inform 2018 Jun;114:66-75. [doi: 10.1016/j.ijmedinf.2018.03.012] [Medline: 29673606]
27. Mace RA, Mattos MK, Vranceanu AM. Older adults can use technology: why healthcare professionals must overcome ageism in digital health. Transl Behav Med 2022 Dec 30;12(12):1102-1105 [FREE Full text] [doi: 10.1093/tbm/ibac070] [Medline: 36073770]
28. Jongstra S, Beishuizen C, Andrieu S, Barbera M, van Dorp M, van de Groep B, et al. Development and validation of an interactive internet platform for older people: the healthy ageing through internet counselling in the elderly study. Telemed J E Health 2017 Feb;23(2):96-104. [doi: 10.1089/tmj.2016.0066] [Medline: 27463120]
29. Observatorio del Envejecimiento. Uso de internet y tecnologías de la información y comunicación en las personas mayores. Centro de Estudios de Vejez y Envejecimiento de la Pontificia Universidad Católica de Chile y Compañía de Seguros Confuturo. 2022. URL: https://observatorioenvejecimiento.uc.cl/wp-content/uploads/2022/06/Observatorio-Reporte-TICS.pdf [accessed 2025-08-27]
30. Chile y Sus Mayores: Sexta Encuesta Nacional de Calidad de Vida en la Vejez 2022 UC-Caja Los Andes. Pontificia Universidad Católica de Chile. 2025. URL: https://encuestacalidaddevidaenlavejez.uc.cl/wp-content/uploads/2023/08/Libro-completo-VI-Encuesta_compressed.pdf [accessed 2025-09-15]
31. Or CK, Holden RJ, Valdez RS. Human factors engineering and user-centered design for mobile health technology: enhancing effectiveness, efficiency, and satisfaction. In: Duffy VG, Ziefle M, Rau PL, Tseng MM, editors. Human-Automation Interaction. Cham, Switzerland: Springer; Dec 15, 2022.
32. Alruwaili MM, Shaban M, Elsayed Ramadan OM. Digital health interventions for promoting healthy aging: a systematic review of adoption patterns, efficacy, and user experience. Sustainability 2023 Dec 02;15(23):16503. [doi: 10.3390/su152316503]
33. Borges do Nascimento IJ, Abdulazeem H, Vasanthan LT, Martinez EZ, Zucoloto ML, Østengaard L, et al. Barriers and facilitators to utilizing digital health technologies by healthcare professionals. NPJ Digit Med 2023 Sep 18;6(1):161 [FREE Full text] [doi: 10.1038/s41746-023-00899-4] [Medline: 37723240]
34. Gutiérrez M. BioRender: Scientific image and illustration software. 2025. URL: https://app.biorender.com/citation/6782ce4b55c5ed2b903ded2e [accessed 2025-09-12]
35. LaMonica HM, Roberts AE, Lee GY, Davenport TA, Hickie IB. Privacy practices of health information technologies: privacy policy risk assessment study and proposed guidelines. J Med Internet Res 2021 Sep 16;23(9):e26317 [FREE Full text] [doi: 10.2196/26317] [Medline: 34528895]
36. Integrated care for older people (ICOPE) : guidance for person-centered assessment and pathways in primary care. Pan American Health Organization. 2020. URL: https://www.paho.org/en/documents/integrated-care-older-people-icope-guidance-person-centered-assessment-and-pathways [accessed 2025-08-27]
37. Risk reduction of cognitive decline and dementia: WHO guidelines. World Health Organization. 2019 Jan 1. URL: https://www.who.int/publications/i/item/risk-reduction-of-cognitive-decline-and-dementia [accessed 2025-08-27]
38. Primary or essential arterial hypertension in people 15 years of age and older. Superintendencia de Salud. URL: http://superdesalud.gob.cl/orientacion-en-salud/hipertension-arterial-primaria-o-esencial-en-personas-de-15-anos-y-mas/ [accessed 2025-08-27]
39. Lloyd-Jones DM, Allen NB, Anderson CA, Black T, Brewer LC, Foraker RE, et al. Life's essential 8: updating and enhancing the American Heart Association's construct of cardiovascular health: a presidential advisory from the American Heart Association. Circulation 2022 Aug 02;146(5):e18-e43 [FREE Full text] [doi: 10.1161/CIR.0000000000001078] [Medline: 35766027]
40. Howard G, Cushman M, Blair J, Wilson NR, Yuan Y, Safford MM, et al. Comparative discrimination of life's simple 7 and life's essential 8 to stratify cardiovascular risk: is the added complexity worth it? Circulation 2024 Mar 19;149(12):905-913. [doi: 10.1161/circulationaha.123.065472]
41. Clocchiatti-Tuozzo S, Rivier CA, Renedo D, Huo S, Hawkes MA, de Havenon A, et al. Life's essential 8 and poor brain health outcomes in middle-aged adults. Neurology 2024 Nov 26;103(10):e209990. [doi: 10.1212/WNL.0000000000209990] [Medline: 39442069]
42. Abad Andrades MS, Gálvez Pérez MJ, Jana Hoebel D, Vergara Andueza O. Guía práctica salud mental y bienestar para personas mayores. SENAMA, Ministerio de Desarrollo Social y Familia, Gobierno de Chile. 2021. URL: https://www.senama.gob.cl/storage/docs/SD_OK_-2.pdf [accessed 2025-08-27]
43. Plan nacional de Salud mental 2017-2025. Biblioteca del Congreso Nacional de Chile. 2024 Feb. URL: https://obtienearchivo.bcn.cl/obtienearchivo?id=repositorio/10221/35814/1/BCN_programas_nacionales_salud_mental_FINAL.pdf [accessed 2025-08-27]
44. WHO guidelines on physical activity and sedentary behaviour. World Health Organization. 2020 Nov 25. URL: https://www.who.int/publications/i/item/9789240015128 [accessed 2025-08-27]

XSL•FO
RenderX

45. Fiscal policies to promote healthy diets: WHO guideline. World Health Organization. 2024 Jun 14. URL: https://www.who.int/publications/i/item/9789240091016 [accessed 2025-08-27]
46. Mess EV, Kramer F, Krumme J, Kanelakis N, Teynor A. Use of creative frameworks in health care to solve data and information problems: scoping review. JMIR Hum Factors 2024 Sep 13;11:e55182 [FREE Full text] [doi: 10.2196/55182] [Medline: 39269739]
47. Hodson N, Woods P, Sobolev M, Giacco D. A digital microintervention supporting evidence-based parenting skills: development study using the agile scrum methodology. JMIR Form Res 2024 Jun 28;8:e54892 [FREE Full text] [doi: 10.2196/54892] [Medline: 38941594]
48. Riede M, Breitschwerdt R, Liebe JD. Inclusive design thinking for the development of digital health applications: a methodology review. Stud Health Technol Inform 2024 Aug 30;317:324-334. [doi: 10.3233/shti240874]
49. Vercruyssen A, Schirmer W, Geerts N, Mortelmans D. How “basic” is basic digital literacy for older adults? Insights from digital skills instructors. Front Educ 2023 Sep 13;8. [doi: 10.3389/feduc.2023.1231701]
50. Gothelf J, Seiden J. Lean UX: Applying Lean Principles to Improve User Experience. Sebastopol, CA: O'Reilly Media; 2013.
51. Gonen E. Tim Brown, change by design: how design thinking transforms organizations and inspires innovation (2009). Mark Global Dev Rev 2019;4(2). [doi: 10.23860/mgdr-2019-04-02-08]
52. Daniels K, Lemmens R, Knippenberg E, Marinus N, Vonck S, Baerts J, et al. Promoting physical activity and a healthy active lifestyle in community-dwelling older adults: a design thinking approach for the development of a mobile health application. Front Public Health 2023 Nov 29;11:1280941 [FREE Full text] [doi: 10.3389/fpubh.2023.1280941] [Medline: 38106904]
53. Gutiérrez M. BioRender: Scientific image and illustration software. 2025. URL: https://app.biorender.com/citation/6783251fbf1713cce012e75c [accessed 2025-09-12]
54. Or C, Tao D. Usability study of a computer-based self-management system for older adults with chronic diseases. JMIR Res Protoc 2012 Nov 08;1(2):e13 [FREE Full text] [doi: 10.2196/resprot.2184] [Medline: 23612015]
55. Gutiérrez M. BioRender: Scientific image and illustration software. 2025. URL: https://app.biorender.com/citation/677ed37933288e508b1fb1a1 [accessed 2025-09-12]
56. Delgado Derio C, Guerrero Bonnet S, Troncoso Ponce M, Araneda Yañez A, Slachevsky Chonchol A, Behrens Pellegrino MI. Memory, fluency, and orientation (MEFO): a five-minute screening test for cognitive decline. Neurología (English Edition) 2013 Sep;28(7):400-407. [doi: 10.1016/j.nrleng.2012.10.001]
57. Buysse DJ, Reynolds CF3, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
58. Morris MC, Tangney CC, Wang Y, Sacks FM, Barnes LL, Bennett DA, et al. MIND diet slows cognitive decline with aging. Alzheimers Dement 2015 Sep 15;11(9):1015-1022 [FREE Full text] [doi: 10.1016/j.jalz.2015.04.011] [Medline: 26086182]
59. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
60. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
61. Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, et al. Development and validation of a geriatric depression screening scale: a preliminary report. J Psychiatr Res 1982 Jan;17(1):37-49. [doi: 10.1016/0022-3956(82)90033-4] [Medline: 7183759]
62. Zelinski EM, Gilewski MJ. A 10-item Rasch modeled memory self-efficacy scale. Aging Ment Health 2004 Jul 19;8(4):293-306. [doi: 10.1080/13607860410001709665] [Medline: 15370046]
63. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994 Mar 01;49(2):M85-M94. [doi: 10.1093/geronj/49.2.m85] [Medline: 8126356]
64. Ory MG, Altpeter M, Belza B, Helduser J, Zhang C, Smith ML. Perceived utility of the RE-AIM framework for health promotion/disease prevention initiatives for older adults: a case study from the U.S. evidence-based disease prevention initiative. Front Public Health 2014 Apr 27;2:143 [FREE Full text] [doi: 10.3389/fpubh.2014.00143] [Medline: 25964897]
65. Thabane L, Ma J, Chu R, Cheng J, Ismaila A, Rios LP, et al. A tutorial on pilot studies: the what, why and how. BMC Med Res Methodol 2010 Jan 06;10(1):1 [FREE Full text] [doi: 10.1186/1471-2288-10-1] [Medline: 20053272]
66. Rost NS, Salinas J, Jordan JT, Banwell B, Correa DJ, Said RR, et al. The brain health imperative in the 21st century—a call to action. Neurology 2023 Sep 26;101(13):570-579. [doi: 10.1212/wnl.0000000000207739]
67. Liu N, Yin J, Tan SS, Ngiam KY, Teo HH. Mobile health applications for older adults: a systematic review of interface and persuasive feature design. J Am Med Inform Assoc 2021 Oct 12;28(11):2483-2501 [FREE Full text] [doi: 10.1093/jamia/ocab151] [Medline: 34472601]

XSL•FO
RenderX

68. Sweeney M, Barton W, Nebeker C. Evaluating mobile apps targeting older adults: descriptive study. JMIR Form Res 2023 Apr 27;7:e37329 [FREE Full text] [doi: 10.2196/37329] [Medline: 37103995]
69. Hoque R, Sorwar G. Understanding factors influencing the adoption of mHealth by the elderly: an extension of the UTAUT model. Int J Med Inform 2017 May;101:75-84. [doi: 10.1016/j.ijmedinf.2017.02.002] [Medline: 28347450]
70. Gutiérrez M, Mauro J, Asecio J, Acevedo F, Herrada J, Torres C, et al. Implementation of a virtual telehealth course for Chilean older adults and health students. Rev Med Chil 2023 Feb;151(1):61-71 [FREE Full text] [doi: 10.4067/s0034-98872023000100061] [Medline: 37906747]
71. Kim M, Kim B, Park S. Social support, eHealth literacy, and mhealth use in older adults with diabetes: moderated mediating effect of the perceived importance of app design. Comput Inform Nurs 2024 Feb 01;42(2):136-143. [doi: 10.1097/CIN.0000000000001081] [Medline: 38129323]
72. Lehtisalo J, Rusanen M, Solomon A, Antikainen R, Laatikainen T, Peltonen M, et al. Effect of a multi-domain lifestyle intervention on cardiovascular risk in older people: the FINGER trial. Eur Heart J 2022 Jun 01;43(21):2054-2061 [FREE Full text] [doi: 10.1093/eurheartj/ehab922] [Medline: 35051281]
73. Ross R, Neeland IJ, Yamashita S, Shai I, Seidell J, Magni P, et al. Waist circumference as a vital sign in clinical practice: a Consensus Statement from the IAS and ICCR Working Group on Visceral Obesity. Nat Rev Endocrinol 2020 Mar 04;16(3):177-189 [FREE Full text] [doi: 10.1038/s41574-019-0310-7] [Medline: 32020062]
74. Guralnik JM, Ferrucci L, Pieper CF, Leveille SG, Markides KS, Ostir GV, et al. Lower extremity function and subsequent disability: consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery. J Gerontol A Biol Sci Med Sci 2000 Apr;55(4):M221-M231. [doi: 10.1093/gerona/55.4.m221] [Medline: 10811152]
75. Dent E, Morley JE, Cruz-Jentoft AJ, Woodhouse L, Rodríguez-Mañas L, Fried LP, et al. Physical frailty: ICFSR international clinical practice guidelines for identification and management. J Nutr Health Aging 2019 Nov;23(9):771-787 [FREE Full text] [doi: 10.1007/s12603-019-1273-z] [Medline: 31641726]
76. Albala C, Lera L, Sanchez H, Angel B, Márquez C, Arroyo P, et al. Frequency of frailty and its association with cognitive status and survival in older Chileans. Clin Interv Aging 2017 Jun;12:995-1001. [doi: 10.2147/cia.s136906]
77. Blair CK, Harding E, Wiggins C, Kang H, Schwartz M, Tarnower A, et al. A home-based mobile health intervention to replace sedentary time with light physical activity in older cancer survivors: randomized controlled pilot trial. JMIR Cancer 2021 Apr 13;7(2):e18819 [FREE Full text] [doi: 10.2196/18819] [Medline: 33847588]
78. Miller KJ, Siddarth P, Gaines JM, Parrish JM, Ercoli LM, Marx K, et al. The memory fitness program: cognitive effects of a healthy aging intervention. Am J Geriatr Psychiatry 2012 Jun;20(6):514-523. [doi: 10.1097/jgp.0b013e318227f821]
79. Suzuki T, Shimada H, Makizako H, Doi T, Yoshida D, Ito K, et al. A randomized controlled trial of multicomponent exercise in older adults with mild cognitive impairment. PLoS One 2013;8(4):e61483 [FREE Full text] [doi: 10.1371/journal.pone.0061483] [Medline: 23585901]
80. Recio-Rodríguez JI, Lugones-Sanchez C, Agudo-Conde C, González-Sánchez J, Tamayo-Morales O, Gonzalez-Sanchez S, et al. Combined use of smartphone and smartband technology in the improvement of lifestyles in the adult population over 65 years: study protocol for a randomized clinical trial (EVIDENT-Age study). BMC Geriatr 2019 Jan 23;19(1):19 [FREE Full text] [doi: 10.1186/s12877-019-1037-y] [Medline: 30674284]
81. Lyons EJ, Swartz MC, Lewis ZH, Martinez E, Jennings K. Feasibility and acceptability of a wearable technology physical activity intervention with telephone counseling for mid-aged and older adults: a randomized controlled pilot trial. JMIR Mhealth Uhealth 2017 Mar 06;5(3):e28 [FREE Full text] [doi: 10.2196/mhealth.6967] [Medline: 28264796]

Abbreviations

CVH: cardiovascular health
ICOPE: Integrated Care for Older People
mHealth: mobile health
NCD: noncommunicable disease
RCT: randomized controlled trial
RE-AIM: Reach, Effectiveness, Adoption, Implementation, and Maintenance
REDCap: Research Electronic Data Capture
SPPB: Short Physical Performance Battery
UX: user experience
WHO: World Health Organization

XSL•FO
RenderX

Edited by A Garcia; submitted 30.01.25; peer-reviewed by C Or, E Spanakis; comments to author 29.03.25; revised version received 19.04.25; accepted 20.04.25; published 01.10.25.

Please cite as:
Gutiérrez M, Cabello V, SanMartín CD, Mauro J, Musa G, Torres M, San Martin MC, Márquez C, Beltrán R, Rogers N, Salech F, Ponce D, Ampuero C, Varas P, Gamboa-Labbé B, Cortés VD, Vega C, Ruiz R, Velásquez J, Vergara R, Behrens MI, More J, Delgado Derio C
An mHealth App (NeoMayor) to Promote Healthy Lifestyles and Brain Health in Older Adults: Design and Validation Study
JMIR Aging 2025;8:e71936
URL: https://aging.jmir.org/2025/1/e71936
doi:10.2196/71936
PMID:

©Myriam Gutiérrez, Victoria Cabello, Carol D SanMartín, Jorge Mauro, Gada Musa, Melissa Torres, Maria Consuelo San Martin, Carlos Márquez, Rodrigo Beltrán, Nicole Rogers, Felipe Salech, Daniela Ponce, Christian Ampuero, Pia Varas, Bastián Gamboa-Labbé, Víctor David Cortés, Carlos Vega, Rocío Ruiz, Juan Velásquez, Rodrigo Vergara, Maria Isabel Behrens, Jamileth More, Carolina Delgado Derio. Originally published in JMIR Aging (https://aging.jmir.org), 01.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

End User and Primary Care Physicians' Perspectives on Digital Innovations in Dementia Risk Detection: Focus on a Digital Sleep Biomarker

Ríona Mc Ardle[1,2], PhD; Marie Poole[3], PhD; Sophie Horrocks[4,5], MA, MSc; Josh King-Robson[6], BMBS; Jonathan M Schott[6,7], MD, FMedSci; David Sharp[4,8], PhD; Matthew Harrison[4,5], MA, MEng; Louise Robinson[3], MD

[1]Translational and Clinical Research Institute, Newcastle University, Newcastle Upon Tyne, United Kingdom

[2]NIHR Newcastle Biomedical Research Centre, Newcastle Upon Tyne, United Kingdom

[3]Population Health Sciences Institute, Newcastle University, Newcastle Upon Tyne, United Kingdom

[4]Care Research and Technology Centre, UK Dementia Research Institute, London, United Kingdom

[5]Institute of Global Health Innovation, Helix Centre, Imperial College London, London, United Kingdom

[6]Queen Square Institute of Neurology, Dementia Research Centre, University College London, London, United Kingdom

[7]University College London, UK Dementia Research Institute, London, England, United Kingdom

[8]Department of Brain Sciences, Faculty of Medicine, Imperial College London, London, United Kingdom

Corresponding Author:
Matthew Harrison, MA, MEng
Institute of Global Health Innovation
Helix Centre
Imperial College London
2, Paterson Wing, St Mary's Campus
London, W2 1NY
United Kingdom
Phone: 1 02033125525
Email: matthew.jc.harrison@imperial.ac.uk

Abstract

Background: Dementia is a global health priority. Early identification in asymptomatic or mildly symptomatic individuals (ie, dementia risk detection) is proposed as a clinical solution for early intervention and could support researchers to identify novel neuropathological targets and recruit to clinical trials. Digital biomarkers of behavioral or physiological markers, including sleep, are cited as a potential low-cost, noninvasive, and objective method for dementia risk detection. Understanding perspectives on digital biomarkers, particularly acceptability, from potential end users and clinical staff is required when considering implementation within any clinical service. With emerging evidence of sleep as a risk marker for dementia, the efficacy of the Dementia Research Institute Sleep Index (DRI-SI), based on continuous remote monitoring of sleep patterns detected by a digital sleep mat, for dementia risk detection, is currently being explored by the InSleep46 study.

Objective: This qualitative substudy aimed to explore perspectives of potential end users and primary care physicians regarding the use of a digital sleep mat to measure the DRI-SI and its application towards dementia risk detection.

Methods: Thirty-one potential end users (age: 31-82 years, 11 female and 20 male) from Newcastle and London, United Kingdom, with personal or caregiving experience related to dementia, participated in qualitative focus group workshops. They shared opinions on integrating the sleep mat into their homes, the DRI-SI's potential for identifying dementia risk, and the necessary information for engagement with related clinical services. Seven primary care physicians from across England participated in semistructured interviews regarding the potential application of the DRI-SI in dementia risk detection and its integration into current clinical practice. Inductive thematic analysis was conducted to identify key themes.

Results: Four key themes emerged from end user focus groups: (1) practical use, (2) prospective acceptability, (3) clinical management, and (4) data concerns. Three main themes came from the semistructured interviews with physicians: (1) prospective acceptability, (2) health care provision, and (3) practical considerations. Common themes were identified in both groups but held differing perspectives. End users were focused on practical aspects of integrating the digital sleep mat within their daily life, the effect of the DRI-SI on clinical care, and privacy concerns regarding data use. Primary care physicians were concerned more

XSL•FO
RenderX

broadly with how the DRI-SI and dementia risk detection service would integrate into current clinical practice, the impact on clinical resources and patient well-being, and the need for clinical actionability and guidance on discussing results with patients.

Conclusions: End users would find the DRI-SI acceptable as part of their clinical care, but primary care physicians require a more robust evidence base. Future research should explore the integration of the DRI-SI into clinical care/research pathways to enhance clinical acceptability. Five key recommendations have been made for further development of digital biomarkers for dementia risk populations.

(JMIR Aging 2025;8:e74307) doi:10.2196/74307

KEYWORDS

dementia; digital technology; digital health; mass screening; qualitative research; community-based participatory research

Introduction

Dementia is considered a global health priority by the World Health Organization and the leading cause of disability and dependency amongst older adults [1]. Dementia is a syndrome characterized by significant cognitive impairment leading to impairments in functional abilities. Fifty-five million people are currently living with the condition worldwide, with numbers expected to rise to 139million by 2050 [2]. Dementia is not an inevitability; the recent Lancet Commission suggested that ≈45% of all dementia cases are preventable, reporting 14 known modifiable risk factors [3]. Further risk factors, such as sleep, are under investigation [4]. In the absence of a globally accessible cure or treatment, the World Alzheimer's Report 2023 declared dementia risk reduction to be the most proactive and feasible way to combat dementia [2].

Dementia risk detection, which involves screening for dementia-causing diseases in those who are apparently well, has been proposed to enhance identification of high-risk individuals in asymptomatic populations [5-8]. For patients, this would provide access to earlier treatments, such as personalized risk reduction interventions [6], or disease-modifying therapies for Alzheimer disease; shown to be most efficacious when given early in the disease course [9]. From a scientific perspective, dementia risk detection may support identification of neurodegenerative changes in asymptomatic individuals, providing novel disease-modifying targets, and aiding recruitment of high-risk individuals into clinical trials [10]. There is a growing interest in "digital biomarkers," defined as objective, quantifiable physiological and behavioral data measured using digital technologies, to act as low-cost, noninvasive methods to detect dementia risk [5,10]. Digital sleep biomarkers may be one such method [5,11].

Sleep disturbances are a common symptom of dementia, presenting as difficulties falling or remaining asleep, excessive daytime napping, "sundowning" whereby people become confused and agitated during evenings, and rapid eye movement sleep behavior disorder [12-15]. Although the cause-and-effect process is unclear, sleep disturbances may occur prior to cognitive symptom onset as they are associated with an increased risk of developing dementia in both cognitively normal older adults and people with Parkinson disease [16,17]. Early evidence indicates that differences in sleep patterns, quantified by a digital sleep biomarker, known as the Dementia Research Institute Sleep Index (DRI-SI), may be able to distinguish dementia from normal ageing [11]. The DRI-SI can be objectively measured through continuous remote monitoring of sleep patterns using an under-mattress digital sleep mat (Withings Sleep analyzer) and is now being explored in asymptomatic individuals, with the view of assessing its utility to detect dementia risk prior to cognitive symptom onset.

This research is timely, given the rising interest in a "digital first" health care system in England. Most recently, the UK's Digital Health and Social Care Plan set out government-level goals to improve patient access to health care services via digital means and expand remote monitoring to manage long-term conditions [18]. National and international policies, coupled with significant international efforts to technically and clinically validate digital biomarkers and endpoints, indicate that digital dementia risk detection tools may soon be a feasible and useful avenue to support dementia risk reduction by identifying high-risk individuals requiring intervention [4,5,19-24]. However, such tools and resulting digital risk scores must be acceptable to both the potential end users (ie, the general public) and clinicians in order to be commissioned within health care services [8,25], especially primary care, where most screening occurs [26]. Recent surveys of the public, including underserved groups, have indicated that dementia risk detection is generally acceptable to the public [27]. However, there has been limited exploration regarding their perspectives on the use of a continuous remote monitoring digital device, such as a digital sleep mat, for this purpose [27,28]. A recent systematic review has indicated that primary care physicians have mixed views on the implementation of dementia risk screening but are interested in engaging in dementia risk reduction approaches [29]. Negative perspectives on dementia risk screening were largely attributed to limited time, resources, and clinicians' lack of education and knowledge in the area. Conversely, the review highlighted positive perceptions of a dedicated dementia screening service that would provide protected time and resources to conduct screening and communicate results with patients in a nonstigmatizing manner. Regarding digital biomarkers, one study has reported that primary care physicians recognize the potential of wearable technology to detect dementia but hold concerns regarding health inequities caused by direct-to-consumer products [30]. As such, there is currently not enough literature on this topic to infer how potential end users and primary care physicians would perceive a digital sleep mat and the DRI-SI, and their use within a potential dementia screening service. Further research is therefore required to identify prospective barriers to acceptability and to proactively consider solutions prior to resource-intensive development of

XSL•FO
RenderX

the DRI-SI, ensuring that issues around appropriateness and uptake are addressed from the outset. The aim of this study is to understand the perspectives of (1) potential end users and (2) primary care physicians regarding the use of a digital sleep mat for dementia risk detection.

Methods

Overview

This study is part of a larger study (InSleep46: Detecting and monitoring dementia using dynamic digital biomarkers of night-time behavior and sleep) with the overall aim to produce a digital biomarker, the DRI-SI, for use in screening for neurodegenerative pathology and progression towards dementia [11]. The purpose of the qualitative substudy was to consult the public and relevant professionals on the acceptability and practicality of using digital sleep data as part of a hypothetical dementia screening service, whereby a digital sleep mat would be supplied to participants' homes for 3 months, and the DRI-SI would be derived and communicated. An exploratory, health services research approach was adopted, using qualitative methods, to explore the perspectives of potential end users, including those who were worried about their dementia risk, with a family history or caregiver experience of dementia, and physicians regarding the potential implementation of the digital sleep mat for dementia risk detection [31]. This approach studies how health care services are organized, delivered, and experienced by multiple stakeholders, and seeks to understand how these interactions affect access and outcomes of care.

Ethical Considerations

The study was approved by the Imperial College Ethics Committee (reference number: 6554509) on May 15, 2023. Participants were provided with written information about the study, and their written informed consent was obtained prior to participation. Each participant received remuneration in the form of either shopping vouchers or financial payment in line with National Institute for Health Research (NIHR) guidance for their involvement after participation [32]. No identifiable information about participants is provided.

Recruitment

Potential end users were recruited via local research and community networks in the North East of England and London using purposeful sampling, particularly advertising through organizations inclusive to rural and low socioeconomic communities [33]. The inclusion criteria for potential end users were as follows: experience of dementia diagnosis from perspective of a family caregiver, professional role, or recent experience of own diagnosis; or at risk of developing dementia, or aged 65 years, or worried about their own memory or cognitive abilities; able to communicate effectively in English; resident in the United Kingdom; physically able to attend and participate in an in-person workshop or having the digital skills and equipment to join a web-based group meeting (Zoom/Teams Call). Potential participants were excluded from the study if they had moderate/advanced dementia or if they lacked capacity to understand the project and consent to involvement in line with the principles set out in the Mental Capacity Act (2005) Code of Practice [34].

This study was also designed to capture clinical perspectives; during end user interviews (ie, general public), participants highlighted that the most appropriate clinical context for implementation would be primary care. We therefore focused our recruitment on primary care physicians to ensure that the clinical perspectives reflected where the tool might be applied in practice. Primary care physicians were recruited using convenience and snowball sampling from across England [35]. All physicians worked in the National Health Service (NHS) in England as general practitioners, with experience of providing health care to people with suspected cognitive impairment and dementia. Physicians were included if they had experience in dementia diagnosis in a professional role, were residents in the United Kingdom, and could communicate effectively in English.

Data Collection

Data collection was conducted with potential end service users prior to exploring physician views to allow us to structure physician interviews based on the emergent findings from focus groups with end service users.

Four in-person focus groups with potential end users took place between September 2023 and January 2024, two in Newcastle and 2 in London, facilitated by 2 experienced (MP and LR) and 3 novice qualitative researchers (SH, MH, and RMA; 4 female and 1 male).

Before beginning the focus group discussion, the moderators outlined the InSleep46 study to participants, including the associations between sleep and dementia risk, the practical use of the sleep mat, an example of the data received from the sleep mat, and a brief explanation of the DRI-SI. A semistructured topic guide was developed to explore initial perceptions of the sleep mat, probing on practicalities of use and trust in the data; willingness to use the proposed screening service and DRI-SI to identify dementia risk, probing on using the sleep mat for 3 months and receiving feedback following this, and the information provided when receiving the DRI-SI score; and where participants envisioned this service sitting within health care or as the consumer-based landscape (Multimedia Appendix 1). The topic guide evolved iteratively during data gathering through discussions within the research team. All focus groups were audio recorded and transcribed verbatim using a university-approved transcription company, then checked and anonymized by the research team (RMA).

A semistructured interview approach was used with physicians (Multimedia Appendix 2 for interview schedule). Key themes and considerations identified in the end user focus groups were embedded into an interview schedule, which guided questioning, with ad hoc follow-up questions used to explore salient points. A total of 6 of 7 interviews were conducted by one of 2 researchers (MP and RMA) via one-to-one teleconferencing call, while one interview was in person. Informed consent was obtained in writing prior to the interview and verbally recorded at the beginning of the interview. All interviews were audio recorded, transcribed verbatim, checked, and anonymized by the research team (RMA) prior to analysis.

XSL•FO
RenderX

Public Patient Involvement

Two contributors with lived experience of dementia were involved in this substudy. One contributor supported the development of the aims and methodology at the grant proposal stage, before withdrawing due to changes in their condition. The second contributor has been involved in the conduct of the research, including but not limited to reviewing interview schedules, supporting recruitment efforts, attending project meetings, and contributing to the thematic analysis process and interpretation of data, by providing insights from their lived experience.

Data Analysis

All transcripts were analyzed using inductive thematic analysis, aligned with a critical realism epistemology [36]. NVivo (version 14; Lumivero) software was used for the management of transcripts [37]. This systematic, "bottom-up," reflexive approach involved multiple phases, and data collection and analysis ran in parallel. A member of the research team (RMA) familiarized herself with the data through repeated reading of transcripts. She generated initial codes and discussed these with the research team, including all focus group moderators (MP, MH, SH, and LR) and a Public Patient Contributor with lived experience of dementia, using triangulation to ensure a unified interpretation and relevance to study aims. First-order themes were constructed through researchers' interpretation of the coded data; these were reviewed regularly in discussions with the research team and refined further, ensuring credibility and consistency in analysis and minimizing bias in interpretation. First-order themes were clustered into a second-order series of themes based on commonality of meaning and patterns in the data. This was further refined, with second-order themes expanding to encompass others or shrinking for specificity. Themes were reviewed against the coded data and the entire dataset (either end users or primary care physicians, where relevant) to ensure participants' perspectives were accurately captured.

Reflexivity

The research team was multidisciplinary, including expertise in dementia research, qualitative methods, human-centered design, digital health, and primary care practice. While the team collaborated with the wider InSleep46 team, this work package operated independently from the team developing the DRI-SI. The development of the end user topic guide was developed prior to the first author (RMA) joining the team, although she observed all end user focus groups and kept field notes, using these to input into the interview schedule for primary care physicians, which she conducted. She led the analysis on both sets of data. The first author's expertise is in digital health and dementia research, with prior involvement in dementia risk detection research; this may have initially shaped expectations around the perceived acceptability and value of the DRI-SI. Reflexive engagement was maintained throughout analysis through memoing, maintaining, and iteratively refining a codebook, and regular team discussions. The codebook evolved from an initially broad set of descriptive codes to a refined structure that captured the conceptual relationships between themes. Discussions with other team members were used to challenge assumptions, refine theme definitions, and ensure interpretations remained grounded in participants' accounts rather than disciplinary biases.

Results

Participants

Thirty-one potential end user participants (age: 31-82 years; 11 females and 20 males; 25/31, 81% White) from Newcastle and London, United Kingdom, with personal or caring experience related to dementia, took part in focus group workshops. Seven primary care physicians were recruited from across England and had a range of clinical experience and related interests, outlined in Table 1.

XSL•FO
RenderX

Table 1. Demographic information on primary care physicians.

Participant	Sex	Practice area	Common types of patients	Additional expertise
GP1	Male	Two practices: Urban, suburban; Newcastle, North East	• University students • Older patients • Families • Care home residents	Academic researcher in dementia risk
GP2	Female	Two practices: Urban, suburban; Newcastle, North East	• Care home residents • Dementia/memory problems	Technology Lead for Clinical Commissioning Group (CCG) previously
GP3	Female	Two practices: Urban, suburban; Newcastle, North East	• University students • Wide age range • Dementia risk cases	—[a]
GP4	Male	Semirural, Northumberland, North East	• Older patients • Care home residents • Lower risk profile than other areas	Researcher focused on chronic illness in care homes
GP5	Male	Urban, London, South East	• Older patients with cognitive issues/dementia • Frailty management • Nursing home residents	Researcher in technology and dementia
GP6	Male	Semirural, Bristol, South West	• Patients with dementia • Farming communities	Researcher in dementia diagnostic tests
GP7	Female	Urban, Manchester, North West	• Patients with memory concerns and dementia • Health disparities	Researcher on health inequity in dementia and primary care

[a]Not available.

Briefly, four main themes were constructed through research interpretation from the end user focus groups: (1) practical use, (2) prospective acceptability, (3) clinical management, and (4) data concerns. Three main themes were generated from analysis of the semistructured interviews with primary care physicians: (1) prospective acceptability, (2) health care provision, and (3) practical considerations. Although common themes were identified for both groups concerning prospective acceptability, practical considerations, and health care management, the groups approached these from different perspectives. End users were primarily concerned about the practicalities of using the digital sleep mat and had privacy concerns about personal data. Primary care physicians wanted to know how the device and corresponding DRI-SI would fit into broader health care delivery and decision-making processes. Tables 2 and 3 provide definitions of themes and subthemes, with examples of common topics, for each stakeholder group, respectively. Thematic analysis results are presented below for end users, followed by the primary care physicians.

XSL•FO
RenderX

Table 2. Themes and subthemes which were generated from focus groups with potential end users.

Theme and subtheme	Definition	Common topics
Practical use: considers every day, tangible aspects of using the sleep mat		
Considerations for practical use	Refers to the practicalities of using the sleep mat, which an end user may consider an issue or barrier for use.	• Bed features • Design features • Digital connectivity • Out-of-home activities • Reminders/prompts • Safety considerations • Technical support • Time period for device use
Cost implications	Refers to the costs associated with the sleep mat and proposed dementia screening services, whether experienced by end users, clinicians, or the National Health Service.	• End user costs • Cost-benefit analysis • Supplementation of costs
Prospective acceptability: focuses on perceptions and attitudes of end users regarding the sleep mat		
General acceptability of device and digital technology	Perception of acceptable components or uses of the sleep mat and digital technology.	• Positive perceptions • Negative perceptions • Use of technology • Data monitoring • Trust in commercial or third-party organizations • Trust in data access • Trust in technology
Comparison to other digital products	Perceptions of the sleep mat to other digital devices, such as smart watches, or experiences of using similar devices to the sleep mat, that is, devices which assess sleep.	• Competition with existing devices • Information provision
Comparison to other clinical markers	Perceptions of how the sleep mat compares to other clinical markers for dementia or health (eg, blood biomarkers).	• Part of the clinical toolkit • Preferential perception of the sleep mat • Time period
Clinical management: considers the integration of the sleep mat into the end users' health care journey		
Service provision	Considers who the dementia risk detection service should be provided to and at what clinical timepoint and stage.	• Early detection • Memory impairments • Prevention tool • Sleep problems • Population of interest • Concerns with sleep • Safety device • General health and well-being • Monitoring disease progression
Reciprocal communication	Two-way communication of information relating to the sleep mat/service (eg, results) from both the device itself or from clinical services.	• Clinical service • Communication of data to the patient • Communication of results • Patient to clinician communication
Postscreening actions	Actions that would be taken following results from the sleep mat/clinical service.	• Clinical support • Self-management
Data concerns:considers data handling, access, and accuracy		

XSL•FO
RenderX

Theme and subtheme	Definition	Common topics
Data accuracy	Experiences or confounding factors that may impact the accuracy of the sleep data.	• Accessibility • Alcohol • Bed features • Bed partners • Erratic sleep patterns • Medication • Mental Health • Microphone disturbances • Nonsleep bed behaviors • Sleeping outside the bed • Dementia subtype • Cultural differences • Toilet behaviors • Dementia subtype • Sleep apnea • Hawthorne effect
Data management	Management of data from a processing or governance perspective.	• Patient access • Societal benefit • Data sharing across clinical services • Data processing

XSL•FO
RenderX

Table 3. Themes and subthemes generated from the semistructured interviews with primary care physicians.

Theme and subtheme	Definition	Common topics
Prospective acceptability: considers what makes the sleep mat acceptable to clinicians, from functionality to clinical feasibility and consequences		
Feasibility, reliability, and validity	Reflects concerns regarding practical implementation and trustworthiness of using the sleep mat for dementia risk detection.	• Complexity of sleep behavior • External influences affecting data accuracy • Feasibility of use in end users' homes • Behavioral changes due to Hawthorne effect • Patient acceptance of the mat. • Comparison with existing digital products.
Dementia risk detection	Refers to both patient and health care providers' challenges in interpreting, communicating, and understanding dementia risk scores, and concerns regarding the potential outcomes of dementia risk screening.	• Perception of risk • Purpose of the risk score • Communication of risk • Desire for diagnosis • Patient Choice • Actionability • Negative consequences for patients
Health care provision: c onsiders broader infrastructure and logistical elements needed for risk score successful deployment		
Current clinical practice	Refers to the existing landscape of how sleep and related digital data are managed in primary care.	• Sleep concerns • Sleep management • Consideration of consumer digital products in decision-making
Service provision	Refers to logistical aspects of implementation, such as who, where, and when the dementia sleep risk score should be offered.	• Service delivery • Clinical timepoint • Target population
Health care readiness and actionability	Refers to how equipped the current health care system is to act on the dementia sleep risk score.	• Actionability of risk score • Resource readiness • Integration into current pathways
Requirements for implementation	Factors required to be met to ensure successful deployment of the dementia risk score.	• Cost and value • Trust and credibility
Practical considerations: focuses on how results will be presented and understood by both patients and clinicians		
Data visualization	How the report is presented and shared with end users.	• Report preferences • Comparison with other reports for chronic condition screening • Sharing the report with patients
Clinical interpretation and decision-making	Requirements for the clinician to interpret the data and make informed decisions.	• Need for training and guidance • Meaningful metrics • Identification of reasons for poor sleep or anomalies • Standardized decision-making tool

Thematic Analysis: Potential End Users

Theme 1: Practical Use by End Users

Overview

The first theme was practical use of the device and its use in a dementia screening service, which has 2 subthemes: considerations for practical use and cost implications.

Considerations for Practical Use

Participants identified practical considerations for using the sleep mat in an optimal way. Commonly, they considered bed (eg, mattress type, electric blanket usage, and bed size) and design features (eg, size and robustness of the device), safety issues such as fire hazards, and the need for digital connectivity, such as access to Wi-Fi. They highlighted the need for technical support to set up the device, and reminders and prompts to ensure the device was switched on and collecting data during the required period.

> *I was just going to say, when you wrote 'a patient would install it,' in some ways it may be better to explicitly refer to 'or family or friend' because even though they probably can install it, it's a very simple thing, but I know my grandma, with dementia, would have a lot of insecurity. She would think she's done it wrong. She wouldn't know where to put it.* [G3P02, female, London]

XSL•FO
RenderX

Participants felt that a longer time period than the recommended 3 months would be required and suggested multiple follow-up assessments (eg, once a year) to assess progression or identify new risk signs. Additionally, there was some interest in keeping the device as a permanent monitoring feature. These perspectives may reflect a shift towards personal responsibility and autonomy in managing one's own health, alongside growing expectations of continuous, data-driven engagement with health and well-being data in the emerging digital landscape.

Cost Implications

Cost implications of the device were discussed from the perspective of the end user, such as the price of purchasing the device personally and the additional expense of running the device in one's own home. There was some advocacy for receiving the device on prescription and having running costs supplemented by the health care services, which may reflect an expectation that such costs are the responsibility of a publicly funded health care system rather than the individual. Other participants expressed willingness to purchase the device for their own use, reflecting a broader interest in investing in their personal health. Differing views may be shaped by individuals' socioeconomic status, influencing whether participants regarded cost as a barrier to participation.

> *It would be the same as if you had to pay to have a medical test done. In this country, we don't. It may take that bit longer, or whatever, but everyone has the chance to get treatment. Whereas, if you're having to pay for something, people put it off.* [G2P01, male, Newcastle]

There was also some consideration as to whether the benefit of the device and dementia screening service would benefit the health care service by reducing medication and care costs.

> *I suppose as well, if you've got doctors that realize that early intervention is going to be better for them, they're going to save money on their practice and they're going to save on medication in the long term as well. So, if they're sold on the fact that it's going to make a difference, well that would be a big difference for the practice.* [G1P05, female, Newcastle]

Theme 2: Prospective Acceptability by End Users

Overview

The second theme was prospective acceptability of the device and its use in a dementia screening service, with 3 subthemes: general acceptability of the device, comparison to other digital products, and comparison to other clinical markers.

General Acceptability of the Device and Digital Technology

Overall, the device was viewed positively, considered "unobtrusive," "easy," and beneficial to wider society. Possible negative perceptions related mainly to the microphone feature due to perceived intrusion of privacy, and anxiety over the possible results.

> *Knowing the mat is underneath you, recording this data, might cause you anxiety in itself. Thinking, "If I move-"* [G3P06, female, London]

Familiarity with technology was cited as an important factor for the general acceptability of using the device. Participants highlighted willingness to have their sleep data recorded for use in health care services or to support health care research, but expressed distrust in having the data stored, accessed, and used by commercial organizations or third parties. Similarly, some participants cited distrust in technology generally as a barrier to use.

Comparison to Other Digital Products

The sleep mat was compared with other digital products such as smart watches and phone apps, particularly considering the comparative accuracy to determine sleep, the ability to capture napping behaviors outside bed, and the incentive to use the sleep mat rather than an existing product. Data about sleep can be provided from existing devices, and this information was of interest to participants. Participants considered the sleep mat an additional tool to use to observe their sleep behaviors, alongside existing devices, suggesting an inherent trust in familiar technologies already used to monitor health and well-being This trust appeared linked to perceptions of visibility and access to data, with devices providing continuous real-time feedback viewed as more reliable or desirable because one could actively observe and interpret their own data. Contrastingly, the sleep mat's limited scope of only capturing in-bed activity reduced its perceived value, with concerns that it would miss napping behaviors in other locations (eg, an armchair) and not provide an accurate report to clinicians.

> *There are lots of apps that track, for example, when you went to bed, when you woke up, and then they can do your REM cycles. How would you incentivize someone to see the value of putting a device under their bed when they could actually change nothing and still get that data for themselves?* [G3P02, female, London]

Comparison to Other Clinical Markers

Some participants suggested a preference for the sleep mat over other clinical tests they were aware of or had experience of, such as cognitive interviews, stool samples, blood tests, and sleep diaries, as it is less obtrusive and stigmatizing. They felt it could be paired with other tests to support screening efforts, suggesting a desire for a holistic screening service.

> *I can see the benefits of the mat if it's used in conjunction with other tests, rather than being a sole determinant.* [G3P01, male, London]

Theme 3: Clinical Management

Overview

The third theme was clinical management, largely focusing on the potential dementia screening service proposed, encompassing 3 subthemes: service provision, reciprocal communication, and postscreening actions.

XSL•FO
RenderX

Service Provision

Participants discussed who the dementia risk detection service would be provided to in terms of age, family history, and existing health conditions, and at which clinical stage. They suggested that the service might be provided when people present with sleep problems, and that it should be considered an early detection or preventative service. There were mixed feelings regarding the provision of the service to those presenting with memory problems, generally feeling this was too late, although one participant thought it could be provided to people with mild cognitive impairment.

> *One of my initial thoughts was, by the time you've got to the memory clinic, very often that almost seems too late.* [G1P01, female, Newcastle]

Participants also commonly advocated the provision of the sleep mat and service to those with a family history of dementia. They felt that younger people may be more inclined to use the service and make risk-reducing lifestyle changes accordingly, although this may be deterred by the perception that dementia is "for old people" (G1P03, male, Newcastle). Additionally, provisions for middle-aged people and sports players were mentioned as these groups were perceived to have an increased risk of dementia.

Participants also suggested that the sleep mat could be used for additional purposes, such as monitoring sleep in those with concerns, acting as a safety device for people with dementia (eg, detecting falls), as a general well-being device, or as a method to monitor disease progression in dementia.

Reciprocal Communication

The communication of data and results was discussed by participants. Participants wanted to be able to see their own data, while expressing the need to have an expert relay what that data means to them.

> *Well, really, I mean, the common lay person doesn't know what information they're looking for, really. I mean, at the end of the day, it seems, sleep, they don't know the in depths of it, what's good and what's bad. They'll be looking for the GP or the professional to let them know the information that they thought was important.* [G2P02, male, Newcastle]

Participants also expressed the desire to be able to bring results from the sleep mat to the doctor if they were worried before the 3 months were up, or to be able to use the data to communicate abnormalities with clinicians themselves. This reflects the interest in actively engaging with health care services through control of one's own data, rather than passively waiting for the clinician's decision.

Postscreening Actions

Participants expressed the need for clinical support or self-management avenues following sleep-based dementia screening, particularly regarding improving their sleep. Referrals to sleep specialists, talking therapies, and cognitive behavioral therapies were proposed. Some participants expressed willingness to adopt lifestyle changes around sleep if they received a high-risk score or demonstrated poor sleep.

> *"We don't say you're going to have dementia, but we think you're not sleeping very well." Whether I'm going to have dementia or not, I'm going to try and start looking at things that will help me to sleep better.* [G4P02, female, London]

Theme 4: Data Concerns by End Users

Overview

The final theme identified was data concerns, composed of 2 subthemes: data accuracy and management.

Data Accuracy

Participants highlighted multiple factors that could potentially impact sleep and the DRI-SI results, including alcohol, erratic sleep patterns, medications, mental health, daytime sleep outside of the bed, dementia subtypes, and sleep apnea. Additionally, participants outlined concerns regarding characteristics or experiences which might inaccurately be perceived as sleep behaviors relating to dementia risk, such as the influence of bed partners (eg, waking up when partner gets up for the toilet) or incorrectly detecting a bed partner's heart rate as their own including spouses, partners, children and pets; behaviors in the bed besides sleeping such as watching TV or intimate moments; and frequent nighttime toilet use. They were concerned that the microphone features may inaccurately classify white noise or talking as sleep-related noises. Finally, they highlighted that results may be impacted by the accessibility of the device; for example, only collecting and comparing data from those who could afford to use the device. Multiple participants also suggested that end users may change their sleeping behaviors due to a heightened awareness of sleep monitoring.

> *What would be disastrous is if you had this thing and then you stopped doing whatever it is you're doing or trying to change whatever it is, the way you normally do things because you've got this mat.* [G4P02, female, Newcastle]

Data Management

Participants generally expressed the desire to have access to their own sleep data as collected by the device. Some participants considered wider access to their data, such as by researchers, as acceptable if it was positively benefiting wider society, and advocated for the data to be shared across clinical services.

> *So long as it's anonymized, I don't have a problem with the collection of data. Also, provided that it's doing useful research, but also that the research is open, again, I wouldn't mind. I wouldn't want it being sold to the highest bidder… I wouldn't want them being sold to a private company to make a profit off. I'd want any medicine that comes out of it to be open.* [G2P01, male, Newcastle]

Participants questioned who would be processing the sleep data and suggested that it should be processed by the health care services rather than third-party organizations. These perceptions likely reflect wider social anxieties about data misuse and commercial exploitation, and limited awareness of how health

XSL•FO
RenderX

data is often managed through the commissioning of external companies to process tests and support digital infrastructure.

> *But I'm saying that, if you're saying, from a patient point of view, "Where's my data being held and who's processing it?" I would want it to be processed- I don't see why, in the medium to long term- In the short term it's different. But in the medium to long term, why can't that be processed within the auspices of the NHS?* [G4P01, male, London]

Thematic Analysis: Primary Care Physicians

Theme 1: Prospective Acceptability

Overview

Physicians considered what makes the digital sleep mat acceptable to health care professionals, from functionality and validity of the device to clinical feasibility and consequences of dementia risk detection. Two subthemes were identified as follows: (1) feasibility, reliability, and validity and (2) dementia risk detection.

Feasibility, Reliability, and Validity

Regarding the use of the device, physicians considered how feasible and acceptable it would be for use in potential end users. They generally considered it noninvasive and simple, although one mentioned that there may be concerns regarding data privacy. They wanted to understand the benefit of the digital sleep mat in comparison to other familiar digital devices, such as smart watches, particularly as the digital sleep mat may not detect napping behaviors. There were some practical concerns about end users not using it correctly or unplugging the device.

> *I suppose the cables dangling around and could that get in the way? Are people just going to say, "Oh that's in the way," or trip over it and remove the sensor.* [GP4]

Physicians highlighted sleep as a complex, multifactorial behavior, with many different confounding factors for poor sleep quality, including nighttime toileting, partners' sleep behaviors, low mood, chronic pain, heart problems, and sleep apnea. They felt it would be difficult to disentangle these factors from the DRI-SI and required more information on how these confounders would be accounted for. Most physicians agreed that there may be a Hawthorne effect early in the sleep monitoring process, whereby end users would change their sleeping behaviors due to a heightened awareness of being monitored, but that the 3-month period would alleviate this effect.

> *There's so much more about sleep than just… yeah, and I think that would be harder to sell to patients and clinicians. They'd be like, "Well, actually there's so much more to sleep than just memory and the picking up dementia."* [GP2]

> *Whenever a patient knows they're being monitored for something, I suspect what they would do is they would try to be a good, model patient. They'll go to sleep on time and they'll try to make sure they're sleeping above the sensor. But I think the fact that it's over three months means that your initial data might seem quite pure, let's say. But then, later on, maybe after the first few weeks, when the patient has got used to it, they may have forgotten about the sensor. So, then, it becomes a much more natural version of their routine.* [GP1]

Additionally, physicians wanted information about the scientific validity of the device and DRI-SI, including the validity of sleep as a risk marker of dementia, the sensitivity and specificity of the score to detect dementia risk, at which clinical stage or disease subtype, and the measurement error and reliability of the device.

> *And that kind of comes into, is poor sleep a symptom of you're going to develop dementia, or is it a risk factor that causes you to develop dementia?* [GP7]

> *I just want to be knowing how it's working and how accurate it is.* [GP4]

Dementia Risk Detection

Physicians described being familiar with communicating risk scores for other clinical conditions, such as diabetes and heart disease, and generally felt confident that patients understand these concepts. However, they noted that fear and stigma of dementia and ageing stereotypes may reduce patients' desire to know their dementia risk or to understand the role sleep may play in it, although this appears to be shifting for the younger older generation. They noted that some patients may struggle to separate risk from diagnosis and may consider the condition as an inevitability. This may be due to ineffective communication, as explaining risk requires skill and experience.

> *Sometimes people will have experience of dementia and older parents or older relatives, and they'll have almost decided… Well, both my parents had this, you know. So, I'm going to get this. So, I don't know if the score per se would be but in terms of whether they would see that as a modifiable thing, in the same way that they would look at their blood pressure, their cholesterol. I'm not sure.* [GP3]

> *What they understand is the feelings that those numbers trigger off in them. So, in my experience, when I explain to someone that, "You might have a 20% or 30% risk of having a stroke," or whatever, the patient doesn't hear 20% or 30%. What they hear is, "Stroke," so that they hear the thing that they can latch onto, which is the thing which then triggers off an emotional response in them.* [GP6]

To support discussions regarding dementia risk, physicians highlighted that they would need to understand the relationship between sleep and dementia, the components of the risk score, and have guidance on how to explain it to patients, as is the protocol with other common screening processes.

> *I think the one thing that they would want to know is, what is the risk of a sleep disturbance? Or what is it that makes sleep something that puts you at risk of dementia?* [GP1]

Finally, physicians considered the provision of the DRI-SI as acceptable only if there was a consequential positive action to

XSL•FO
RenderX

improve health or reduce and modify risk. Without this, physicians expressed concerns about frightening patients and causing unintended harm by sharing a risk score without a pathway to modification.

> *I wouldn't share it with a patient and frighten them and say that "You're moderate risk of dementia," because I'm sorry until somebody can explain to me the clear evidence for that then I would cause harm to someone by doing that so…* [GP2]

Potential interventions need to be carefully considered and tailored to a patient's individualized response, as this information may support some to make lifestyle changes, but negatively impact others due to emotional connotations and comorbid conditions. Some physicians questioned the need for the DRI-SI based on the current risk reduction evidence, citing primary population-level prevention as a simpler, effective strategy. However, risk modification was not the only considered benefit. Risk detection could also support advanced care planning and trigger safety assessments, such as reviewing driving abilities. Additionally, physicians wanted to know if the DRI-SI could assess intervention effectiveness, with a reduced score showing modified risk.

> *There's a possibility and there's a plausible scenario, given that, in my opinion, a lot of it is driven by emotion triggers, that the people that are least well equipped to be able to actually change their behavior, because their response gets driven by an overwhelming emotional response, then end up not being able to change their behavior, even if they are, in fact, at the greatest risk of getting dementia.* [GP6]

Theme 2: Health Care Provision

Overview

Physicians considered the broader infrastructure and logistical elements required to successfully deploy the digital sleep mat and subsequent DRI-SI within health care, leading to four subthemes: (1) current clinical practice, (2) health care readiness and actionability, (3) service provision, and (4) requirements for implementation.

Current Clinical Practice

Physicians indicated that they often speak to people about sleep problems, usually part of a larger consultation. However, sleep concerns do not usually occur in the context of help-seeking for dementia-related symptoms. Physicians talked about a variety of sleep management techniques they might use, often involving identifying the root cause and trying to make small changes to alleviate the impact. They indicated that they would generally refer people to social prescribers to speak about sleep hygiene, or to websites and apps to encourage lifestyle modifications rather than medical treatment. However, physicians felt that sleep management is variable depending on the clinician and their experience, and generally indicated that sleep is not an area where they have great clinical expertise.

> *Well, I don't have any treatments for poor-quality sleep. And I have asked the powers that be, and I'm not convinced that there's anything that I can really do. So, we talk about sleep hygiene, and exercise, and not eating too late at night, and limiting caffeine intake, particularly in the latter part of the day. Of course, we do all that. We have a speculative idea of what melatonin might do. And many people just go buy it. We avoid benzodiazepines, in the form of hypnotics, very firmly, but we do give them. If there's complete sleeplessness, usually, on a short-term basis. But what's a treatment for sleep disturbance?* [GP5]

> *I don't have any formal training in sleep stuff I suppose like this so I'm coming back to myself.* [GP4]

Physicians also indicated that an increasing number of patients were bringing their own direct-to-consumer devices and outputs relating to sleep, physical activity, and heart rate to a consultation, but there were mixed views on what to do with this data. Some physicians appreciated the role these devices can play in empowering patients and promoting self-advocacy, while others felt that they would not know how to interpret or use the data, and do not allow it to impact clinical decision-making.

> *I think it's a great thing. I think, if people are interested and engaged their health, then that's fantastic, isn't it? But I also think, yeah, it's quite interesting to look through someone's app or their resting heart rate.* [GP6]

> *It wouldn't be like the main point of a consultation if anything… And actually, I might note it down and stuff but it doesn't add much to the history of the condition because I don't then go, "Okay. So, he's had this much REM sleep or this much non-REM sleep."* [GP1]

Health Care Readiness and Actionability

A common area of discussion amongst physicians regarding the implementation of a dementia risk detection service based on the digital sleep mat was a lack of resources and overburden on primary care. Physicians were generally concerned with their lack of time and "the never-ending list of things that the GP does" (GP7). They felt that conducting the proposed dementia risk detection service would remove resources from other critical tasks, such as annual reviews, cognitive assessments, and in-depth conversations about memory concerns. They also discussed how detecting dementia risk might lead to an influx of people being screened, leaving fewer resources for those who require postdiagnostic support and care.

> *Because you haven't got… you've got 10 minutes with someone and I'm looking to see if you've got a risk of dementia. I'm not going to be going on about sleep for 10 minutes, I'll be looking at function and cardiovascular risk and all this else.* [GP2]

> *We already know that many people, when they are diagnosed, experience a high degree of abandonment and not really getting very much post-diagnostic support at all. So, the problem is, quite understandably, many people are really interested in trying to widen the funnel and get more people in at an earlier stage and get something done about it. But*

XSL•FO
RenderX

> *the problem is that that means, unless you have an extremely effective filter and funnel at a relatively early stage in the funnel, you have more people progressing through.* [GP6]

Physicians wanted to know how the service would integrate into current care pathways and have clear guidance on where to refer high-risk individuals.

> *"Here, by the way, here's the risk for dementia, and it's really high," there has to be a, "So what?" a, "So what? It's really high. What can I do that is going to change that?" and in a way which is personalized to the individual, because otherwise referring back to the GP will mean diddly-squat.* [GP6]

Service Provision

Physicians discussed where this sleep-based dementia risk detection service would best fit and which population might benefit the most from it. Generally, they felt it should be commissioned as part of the NHS as opposed to a direct-to-consumer service, citing clinicians' "conscious incompetence" with unfamiliar data, exacerbating health inequalities, and over-investigation due to the uncontrolled provision of the digital sleep mat.

> *So, I think that that kind of digital health revolution is going to create significant problems going forward. The sleep mat isn't going to contribute to that if it's done in a controlled way and it's monitored and it's used in the right people and the right population.* [GP7]

Within the NHS, physicians advocated for the integration of the digital sleep mat into either established health check services or annual reviews, or to consider the DRI-SI as one marker in a clinical toolkit for a dedicated dementia screening service. Regarding the target population, patients with existing risk factors such as diabetes and heart disease were commonly mentioned, alongside people with a family history of dementia, genetic vulnerability, or middle-aged people who chose to be screened due to the need for future planning, such as caregivers for disabled children or partners. Some physicians also suggested using the DRI-SI at a later clinical time point, such as in people with suspected dementia or mild cognitive impairment, and to increase rates of diagnosis within the care home population.

> *You know, everybody who has heart disease, diabetes, that kind of thing they get, you know, anywhere between… sort of… one and thee review appointments a year with the practice nurse. And again, I think the opportunity to have a sleep mat would fit quite nicely into that.* [GP3]

> *If this was part of a wider NHS Dementia Health Check, I think it could certainly fit in and I think that service is massively needed and I think it would be a hugely beneficial use of primary care's time.* [GP7]

Requirements for Implementation

To have this service integrated into the NHS, physicians highlighted the device and DRI-SI to be trustworthy and credible, as evidenced by empirical evidence published in peer-reviewed journals, guidelines from credible sources, such as the National Institute of Clinical Excellence, and appropriate licensing, such as acquiring a CE mark for this use. Additionally, they felt that the cost-benefit must be apparent and that the service would need to be funded with financial incentives for service providers.

> *I think you just find that the reality for any practice would be that unless there was some funding tagged onto it that they could use, they wouldn't.* [GP3]

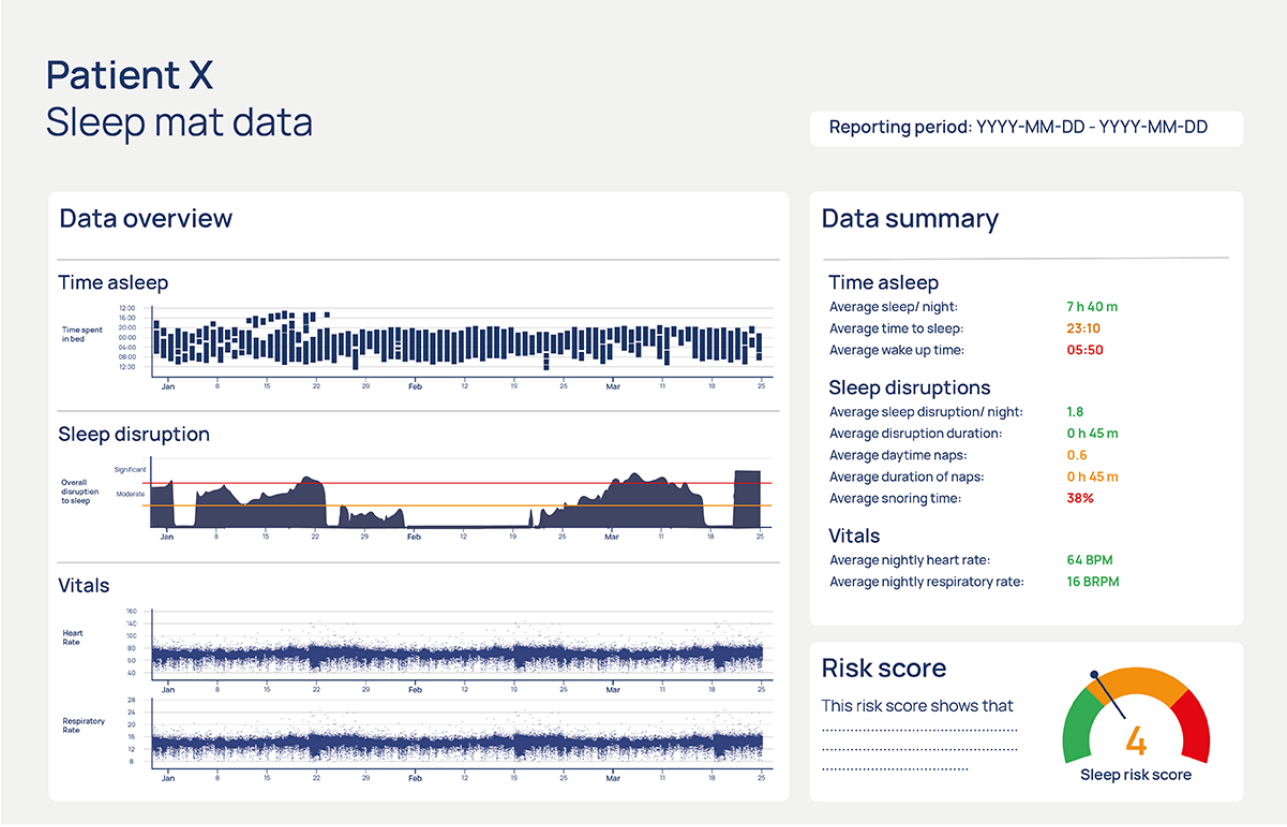
Theme 3: Practical Considerations

Overview

The final theme focuses on how results will be presented and understood by both patients and physicians, involving two subthemes as follows: (1) digital sleep data report, and (2) clinical interpretation and decision-making. These data were largely derived from discussions with the physicians after presenting them with the prototype clinical report (Figure 1).

XSL•FO
RenderX

Figure 1. Example prototype of the data visualization of the Dementia Research Institute Sleep Index risk score.



Digital Sleep Data Report

Generally, physicians felt the digital sleep data report was "easy to read," "presented nice and clear," "helpfully color coded," and "simpler" than other reports. They tended to be drawn towards the graphs. They compared the report to other clinical reports that they were familiar with, such as echocardiograms, DEXA scans (ie, a bone density test), or QRISK results (ie, a risk score for developing cardiovascular disease), highlighting that these outcomes generally have a standardized summary to aid interpretation and support clinical decision-making. Physicians felt more comfortable with this standardized reporting approach as "you know what is going to come next and you know which bits to focus on" (GP1). They mostly felt comfortable with sharing this report with patients but highlighted that they would need guidance and a "patient decision aid document."

> *But in terms of the presentation, it's very clear in terms of the science for me to understand, but also for me to communicate that to a patient; it's, it's very, very good, I think.* [GP7]

Clinical Interpretation and Decision-Making

Regarding their use of the report to make clinical decisions, physicians highlighted the need for meaningful metrics. They commonly focused on the vital signs in the report, due to their greater familiarity with interpreting and communicating these results. Some tried to interpret the sleep scores, such as sleep disruption, but highlighted the need for training to use such metrics clinically. Physicians also felt that some patient feedback or self-reporting would support their decision-making when considering anomalies in the data. Finally, they wanted a standardized decision-making section within the report to guide them on what actions to take next with the patient.

> *In clinical practice we're less attuned to looking at sleep, the sleep graph, so my immediate attention would be to the vital signs because then that's data that we're more familiar in interpreting.* [GP4]

> *I think a summary by the technician who's reporting this, to say, "These are the things it would be helpful to counsel the patient on," would make our job a lot easier. Because, otherwise, if this is going to be used nationally, you're going to have to train a lot of GPs to be able to determine that. It would take a lot of effort to do that.* [GP1]

Discussion

Principal Results

This is the first study to capture perceptions and attitudes of potential end users and primary care physicians regarding the use of a digital sleep mat to derive a digital sleep biomarker (DRI-SI) for dementia risk detection. Potential end users find both the device and the proposal of dementia risk detection generally acceptable; their concerns focused on the practicality of the device, their trust in technology, and the personal impact of sleep monitoring and dementia risk detection. In contrast,

XSL•FO
RenderX

primary care physicians were reluctant to provide the device or engage in dementia risk detection without clear guidelines, empirical evidence of benefit, and clinical actionability. Findings from this study are informative for the development and implementation of other noninvasive biomarkers for dementia, particularly those using digital methods.

Comparison With Prior Work

End users and primary care physicians differed in their views regarding the acceptability of the use and implementation of the DRI-SI. Potential end users largely focused on the use of the digital sleep mat and the information it could provide, considering how it would fit into their daily lives and how they felt about continuous data monitoring. Regarding dementia risk detection, they believed that dementia risk reduction techniques could be implemented early and effectively, reducing overall health care costs and burden. These findings reflect recent systematic reviews and surveys showing positive attitudes of the general public around dementia risk detection due to beliefs of positive outcomes, such as slowing the disease via lifestyle changes and providing more time to plan for the future [27,28]. Comparatively, primary care physicians were more concerned with the clinical validation and health care system integration of the DRI-SI, requiring a more robust evidence base for implementation of a digital sleep biomarker for dementia. This echoes a recent systematic review, which reported that primary care physicians were reluctant to consider the provision of dementia risk detection, due to a lack of knowledge, skills, and training on dementia, resources to provide the service, and concerns regarding the consequences to patients [29]. While physicians held concerns regarding the accuracy of the DRI-SI to detect dementia risk, they also highlighted the more general issues of limited evidence regarding benefits versus harms of dementia risk detection and for accessible treatments that can slow or stop disease progression [25,38-41]. The complexity of the issue moves beyond the development and validation of the digital biomarker and highlights the need for good-quality evidence on clinical actionability of dementia risk detection and benefit to patients and the wider health care system.

Strengths and Limitations

Strengths of this study include the recruitment of potential end users with a range of ages and experiences related to dementia (ie, care partner, family history, subjective memory concerns), a high representation of males, with ethnic diversity representative of the United Kingdom's 2021 census [42]. Physicians were recruited from across different clinical practices and English regions, providing a greater diversity of professional experiences. Physician-focused interview guides were informed by the end user perspectives, reflecting real-world concerns, and supporting alignment of stakeholder perspectives across the proposed clinical pathway. Finally, throughout this substudy, we have engaged meaningfully with people with lived experience of dementia through our focus groups with potential end users and our partnership with public patient involvement contributors. To address recently reported gaps in meaningful end user engagement in digital health research for neurological populations, we have reported these efforts in adherence to the recommended minimum reporting guidelines [43].

Clinical viewpoints in this study are limited to primary care physicians; this decision was based on the perspectives from end users of where the DRI-SI would best fit. However, this limits our understanding of the broader clinical perspective of those involved in the dementia diagnosis and care pathways, such as old age psychiatrists, geriatricians, and neurologists. Additionally, perspectives are limited to end users and physicians living in England and may differ internationally. For example, in the USA, the American Academy of Neurology and the Alzheimer's Association both recommend annual cognitive screening for older adults [44,45], which may lead to greater willingness to uptake the DRI-SI as a dementia risk detection tool. We also did not collect detailed demographic data on end users, such as education level or socioeconomic status, so it is unclear whether the participants included reflected the established dementia risk factors [3].

Implications for Policy and Practice

Digital biomarker efforts often propose value as population-wide dementia risk detection tools, as they are unobtrusive, objective, usually passive, may reduce health care burden by being conducted remotely, and are speculated to detect dementia risk in "real time" as they continuously monitor changes in physiology and behavior [4,10,22,23,26,46]. However, the reluctance from physicians in this study to use the DRI-SI without its integration in clinical guidelines reflects the wider controversy around dementia risk detection [26,40]. No detection tool, digital or otherwise, has yet to meet the internationally-recognized criteria for population screening, which require an adequate understanding of the latent stages of the condition, an accurate and acceptable tool for both patients and physicians, a target population, provision of an accepted treatment, reasonable facilities for diagnosis and treatment, and economic cost-benefit [8,10]. As such, the UK's National Screening Committee currently recommends against such programs [26]. Many of these criteria go beyond the scope of digital biomarker developers' ambitions and require a comprehensive multidisciplinary team science approach, which would include physicians, health care commissioners, and patient voices. Despite these significant barriers, the physicians in this study felt strongly that, if implemented, the DRI-SI and resulting dementia risk detection service should be integrated into existing clinical pathways with guidance for referral in the national health care services.

To meet this challenge, the emergence of "brain health" clinics in the United Kingdom could offer a potential bridge to continue to gather the required evidence while working with health care providers. Brain health clinics are joint clinical-research services that provide access to high-quality assessments not routinely available to patients under investigation for cognitive impairment and physicians, including biomarkers under investigation [47,48]. Recent reports show that 93.5% of patients consent to the use of their data collected as part of the brain health clinic to be used for research [48]. Brain health clinics are considered a safe and acceptable care model for the conduct of more novel biomarker assessments, and have the resources to implement secondary prevention interventions for high-risk individuals [49]. They also meet the suggestion by both end users and physicians to implement the DRI-SI as part of a wider

XSL•FO
RenderX

clinical toolkit and dedicated service to alleviate burden from primary care, while allowing integration into an already established clinical pathway with actionable endpoints, that is, dissemination of information to clinicians in the memory clinic [29,48]. The term "brain health" can also be less stigmatizing than "dementia" and is preferred by primary care physicians when making referrals [50]. Further work should consider capturing the attitudes and perspectives of brain health clinic staff and associated memory clinics regarding the use of the DRI-SI to support dementia risk detection.

Looking to the Future: Wider Impact for Digital Biomarker Research

Overview

While the optimal application of the DRI-SI, like most digital biomarkers, remains unclear, further work could develop and validate it for a clinical service, as previously proposed, or as a tool to support identification of participants for clinical trials and disease-modifying targets [10,26]. From our findings, we have identified 5 key recommendations that should be considered for all proposed applications for the DRI-SI and are relevant to digital biomarkers more widely, prior to any integration into health care systems.

Clinical Staff Require Training on the Proposed Physiology or Behavior

In both this study and the wider literature, physicians indicated limited knowledge and education on dementia-related risks, particularly the association between sleep and cognitive impairment [29]. This leads to mistrust in the DRI-SI, a lack of confidence in how to explain this link to patients, and heterogeneous intervention strategies. Any digital biomarker will require clear and concise clinical information on the underlying association with dementia, which can be shared with end users.

Digital Technologies Need to Be Usable for All Stakeholders

Both end users and physicians held concerns about the usability of the device. End users were primarily concerned about setting up the device and ensuring its continued maintenance, while physicians felt that setup and maintenance would be beyond their resource abilities. Digital biomarker developers need to develop sustainable models of technical support, should these devices be integrated into clinical services or trials. This should include tailored support for individuals most at risk of digital exclusion, to reduce further exacerbation of health inequities in these groups [30,51-53].

Technical and Clinical Validity Need to Be Clearly Communicated

Both end users and physicians worried about confounding factors influencing data accuracy. Previous research regarding clinical use of artificial intelligence models reports that physicians are more likely to trust these tools if they can readily explain and justify them; this requirement was highlighted by both our stakeholder groups [54]. Digital biomarker developers need to develop adaptive algorithms that account for such factors and communicate clearly with physicians how these algorithms work so that they can answer concerns from end users. Additionally, physicians need access to evidence of clinical validation via credible sources (eg, National Institute of Clinical Excellence guidelines and peer-reviewed journals) to improve confidence in the digital biomarker. Additionally, both stakeholder groups indicated a rise in the desire of patients to manage and advocate for their health and well-being, with an increased ability to continuously monitor relevant outcomes through "wellness" consumer-based products such as smart watches. However, both groups also highlighted a lack of "expert" understanding of these data and the continued need for clinical interpretation to make sense of digital health metrics. While self-management supported through digital means is a key principle of the NHS' 10 Year Health Plan for England [55], there is a disparity between the standards of evidence required for commercial "wellness" devices and those approved for clinical use. Clearer communication of these differences to the general public and clinical bodies is essential to prevent unnecessary anxiety among the "worried well" and to support clinicians with reliable, evidence-based guidance on trusted digital tools.

Outcomes Need to Be Meaningful and Interpretable

End users and physicians wanted to understand and see easily interpretable data with meaningful outcomes, that is, aspects of health that have personal importance for end users [56]. This is not just important to our stakeholders; evidence of meaningfulness of outcomes is now required by regulatory bodies such as the European Medical Association and the Food and Drug Administration for qualification of clinical endpoints [57]. Additionally, patient decision aids and clinical results with standardized advice need to be provided through effective data visualization tools, codeveloped with physicians and end users to ensure effective communication and decision-making, as seen with common screening tools such as the QRISK [58].

Transparency in Data Management and Governance

Aligned with common perspectives on digital technologies, end users indicated that they were concerned with their data management and governance, preferring their data to be held and processed by trusted health services rather than third-party providers [26]. To improve trust in these processes and meet recent ethical guidance, digital biomarker developers should be transparent about who has their data and for what use, with guidance on how physicians can discuss this with end users [59,60].

Conclusions

Overall, this study found contrasting perspectives regarding the acceptability of using a digital sleep biomarker as a dementia risk detection tool, based on the current evidence base and clinical practice. Further work to integrate the DRI-SI in clinical-research pathways may support continued development and improve clinical acceptability and adoptability. Findings from this study support broader recommendations for the development of digital biomarkers for dementia populations, including requirements of further training for clinical staff, development of usable technology and meaningful digital

XSL•FO
RenderX

outcomes, and transparency regarding technical and clinical validation and data governance procedures.

Acknowledgments

The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care. We would like to acknowledge all participants for giving their time to this research and the Public Patient Involvement contributors for supporting the study design and data interpretation.

JMS is a National Institute for Health Research (NIHR) Senior Investigator and acknowledges the support of the NIHR University College London Hospitals Biomedical Research Centre and the UCL Centre of Research Excellence, an initiative funded by British Heart Foundation (RE/24/130013). This work is supported by the UK Dementia Research Institute through UK DRI Ltd, principally funded by the Medical Research Council. He acknowledges grant funding from Alzheimer's Research UK, LifeArc, Brain Research UK, Weston Brain Institute, Medical Research Council, British Heart Foundation, Wolfson Foundation, Brain Research UK and Alzheimer's Association (SG-666374-UK BIRTH COHORT). This work uses data provided by study participants or patients and collected through clinical studies or as part of their care and support.

Data Availability

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

Funding

This study was funded by a National Institutes of Health Research (NIHR) Invention for Innovation (i4i) grant (NIHR204286) and supported by the NIHR Imperial Biomedical Research Centre. RMA is funded by the NIHR (NIHR 301677) and NIHR Three Schools: Dementia Career Development Award (NIHR-SSCR-DP-CDA30) and supported by the NIHR Newcastle Biomedical Research Centre based at The Newcastle upon Tyne Hospital National Health Service (NHS) Foundation Trust; Newcastle University; and the Cumbria, Northumberland and Tyne and Wear (CNTW) NHS Foundation Trust.
Insight 46 is funded by grants from Alzheimer's Research UK (ARUK-PG2014-1946 and ARUK-PG2017-1946), Alzheimer's Association (SG-666374-UK BIRTH COHORT), the Medical Research Council Dementias Platform UK (CSUB19166), The Wolfson Foundation (PR/ylr/18575), The Medical Research Council (MC_UU_10019/1 and MC_UU_10019/3) and Brain Research Trust (UCC14191). Florbetapir amyloid tracer was provided in kind by AVID Radiopharmaceuticals (a wholly owned subsidiary of Eli Lilly), who had no part in the design of the study. The funders of the study had no role in study design, data collection, analysis, interpretation, report writing or in the decision to submit the article for publication.

Authors' Contributions

RMA contributed to the investigation, formal analysis, data curation, writing of the original draft, and review and editing of the manuscript. MP contributed to conceptualization, methodology, investigation, formal analysis, review and editing of the manuscript, and supervision. JKR contributed to the review and editing of the manuscript. SH contributed to conceptualization, methodology, investigation, visualization, and review and editing of the manuscript. JMS and DS contributed to conceptualization, funding acquisition, and review and editing of the manuscript. MH contributed to conceptualization, methodology, investigation, visualization, and review and editing of the manuscript. LR contributed to conceptualization, methodology, investigation, supervision, funding acquisition, and review and editing of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Topic guide for focus groups.
[DOCX File , 19 KB - aging_v8i1e74307_app1.docx]

Multimedia Appendix 2
Interview Schedule for primary care physicians.
[DOCX File , 15 KB - aging_v8i1e74307_app2.docx]

References

1. Global action plan on the public response to dementia 2017 - 2025. World Health Organisation. 2017. URL: https://www.who.int/publications/i/item/global-action-plan-on-the-public-health-response-to-dementia-2017---2025 [accessed 2025-11-07]
2. World alzheimer report 2023: reducing dementia risk: never too early, never too late. Alzheimer's Disease International. 2023. URL: https://www.alzint.org/resource/world-alzheimer-report-2023/ [accessed 2025-11-07]

XSL•FO
RenderX

3. Livingston G, Huntley J, Liu KY, Costafreda SG, Selbæk G, Alladi S, et al. Dementia prevention, intervention, and care: 2024 report of the lancet standing commission. Lancet 2024;404(10452):572-628. [doi: 10.1016/S0140-6736(24)01296-0] [Medline: 39096926]
4. Frey A, Karran M, Jimenez R, Baxter J, Adeogun M, Chan D, et al. Harnessing the potential of digital technologies for the early detection of neurodegenerative diseases. Preprint 2019:5. [doi: 10.31219/osf.io/u49z5]
5. Lyall DM, Kormilitzin A, Lancaster C, Sousa J, Petermann-Rocha F, Buckley C, Deep Dementia Phenotyping (DEMON) Network, et al. Artificial intelligence for dementia-applied models and digital health. Alzheimers Dement 2023;19(12):5872-5884 [FREE Full text] [doi: 10.1002/alz.13391] [Medline: 37496259]
6. Kivipelto M, Mangialasche F, Snyder HM, Allegri R, Andrieu S, Arai H, et al. World-wide FINGERS network: a global approach to risk reduction and prevention of dementia. Alzheimers Dement 2020;16(7):1078-1094 [FREE Full text] [doi: 10.1002/alz.12123] [Medline: 32627328]
7. Deckers K, Köhler S, Ngandu T, Antikainen R, Laatikainen T, Soininen H, et al. Quantifying dementia prevention potential in the FINGER randomized controlled trial using the LIBRA prevention index. Alzheimers Dement 2021;17(7):1205-1212 [FREE Full text] [doi: 10.1002/alz.12281] [Medline: 33403822]
8. Wilson JM, Jungner YG. [Principles and practice of mass screening for disease]. Bol Oficina Sanit Panam 1968;65(4):281-393. [Medline: 4234760]
9. Sims JR, Zimmer JA, Evans CD, Lu M, Ardayfio P, Sparks J, TRAILBLAZER-ALZ 2 Investigators. Donanemab in early symptomatic alzheimer disease: the TRAILBLAZER-ALZ 2 randomized clinical trial. JAMA 2023;330(6):512-527 [FREE Full text] [doi: 10.1001/jama.2023.13239] [Medline: 37459141]
10. Ford E, Milne R, Curlewis K. Ethical issues when using digital biomarkers and artificial intelligence for the early detection of dementia. Wiley Interdiscip Rev Data Min Knowl Discov 2023;13(3):e1492 [FREE Full text] [doi: 10.1002/widm.1492] [Medline: 38439952]
11. Soreq E, Kolanko M, Ravindran K, Monica C, Revell V, Daniels S. Contactless longitudinal monitoring of ageing and dementia-related sleep trajectories in the home. Res Square 2024. [doi: 10.21203/rs.3.rs-4523167/v1]
12. Koren T, Fisher E, Webster L, Livingston G, Rapaport P. Prevalence of sleep disturbances in people with dementia living in the community: a systematic review and meta-analysis. Ageing Res Rev 2023;83:101782 [FREE Full text] [doi: 10.1016/j.arr.2022.101782] [Medline: 36356799]
13. Benca R, Teodorescu M. Sleep physiology and disorders in aging and dementia. Handb Clin Neurol 2019;167:477-493. [doi: 10.1016/B978-0-12-804766-8.00026-1] [Medline: 31753150]
14. Ferini-Strambi L, Liguori C, Lucey BP, Mander BA, Spira AP, Videnovic A, et al. Role of sleep in neurodegeneration: the consensus report of the 5th think tank world sleep forum. Neurol Sci 2024;45(2):749-767. [doi: 10.1007/s10072-023-07232-7] [Medline: 38087143]
15. Zhang Y, Ren R, Yang L, Zhang H, Shi Y, Okhravi HR, et al. Sleep in alzheimer's disease: a systematic review and meta-analysis of polysomnographic findings. Transl Psychiatry 2022;12(1):136 [FREE Full text] [doi: 10.1038/s41398-022-01897-y] [Medline: 35365609]
16. Shi L, Chen S, Ma M, Bao Y, Han Y, Wang Y, et al. Sleep disturbances increase the risk of dementia: a systematic review and meta-analysis. Sleep Med Rev 2018;40:4-16. [doi: 10.1016/j.smrv.2017.06.010] [Medline: 28890168]
17. Iranzo A, Fernández-Arcos A, Tolosa E, Serradell M, Molinuevo JL, Valldeoriola F, et al. Neurodegenerative disorder risk in idiopathic REM sleep behavior disorder: study in 174 patients. PLoS One 2014;9(2):e89741 [FREE Full text] [doi: 10.1371/journal.pone.0089741] [Medline: 24587002]
18. A plan for digital health and social care. Department of Health and Social Care. 2022. URL: https://www.gov.uk/government/publications/a-plan-for-digital-health-and-social-care/a-plan-for-digital-health-and-social-care [accessed 2025-11-07]
19. Muurling M, de Boer C, Kozak R, Religa D, Koychev I, Verheij H, RADAR-AD Consortium. Remote monitoring technologies in alzheimer's disease: design of the RADAR-AD study. Alzheimers Res Ther 2021;13(1):89 [FREE Full text] [doi: 10.1186/s13195-021-00825-4] [Medline: 33892789]
20. Rochester L, Mazzà C, Mueller A, Caulfield B, McCarthy M, Becker C, et al. A roadmap to inform development, validation and approval of digital mobility outcomes: the mobilise-D approach. Digit Biomark 2020;4(Suppl 1):13-27 [FREE Full text] [doi: 10.1159/000512513] [Medline: 33442578]
21. Micó-Amigo ME, Bonci T, Paraschiv-Ionescu A, Ullrich M, Kirk C, Soltani A, Mobilise-D consortium. Assessing real-world gait with digital technology? validation, insights and recommendations from the mobilise-D consortium. J Neuroeng Rehabil 2023;20(1):78 [FREE Full text] [doi: 10.1186/s12984-023-01198-5] [Medline: 37316858]
22. Kourtis LC, Regele OB, Wright JM, Jones GB. Digital biomarkers for alzheimer's disease: the mobile/ wearable devices opportunity. NPJ Digit Med 2019;2:9 [FREE Full text] [doi: 10.1038/s41746-019-0084-2] [Medline: 31119198]
23. Buegler M, Harms RL, Balasa M, Meier IB, Exarchos T, Rai L, et al. Digital biomarker-based individualized prognosis for people at risk of dementia. Alzheimers Dement (Amst) 2020;12(1):e12073 [FREE Full text] [doi: 10.1002/dad2.12073] [Medline: 32832589]
24. Koychev I, Lawson J, Chessell T, Mackay C, Gunn R, Sahakian B, et al. Deep and frequent phenotyping study protocol: an observational study in prodromal alzheimer's disease. BMJ Open 2019;9(3):e024498 [FREE Full text] [doi: 10.1136/bmjopen-2018-024498] [Medline: 30904851]

XSL•FO
RenderX

25. Adult screening programme: dementia. UK National Screening Committee. 2019. URL: https://view-health-screening-recommendations.service.gov.uk/dementia/ [accessed 2025-11-07]
26. Costa A, Milne R. Detecting value(s): digital biomarkers for alzheimer's disease and the valuation of new diagnostic technologies. Sociol Health Illn 2024;46(S1):261-278. [doi: 10.1111/1467-9566.13713] [Medline: 37740673]
27. Alzheimer's Research UK. Detectingdiagnosing Alzheimer's disease. In: Enhancing Our Understanding of Public Attitudes to Improving Early Detection and Diagnosis. Cambridge: Alzheimer's Research UK; 2019.
28. Brar M, Mc Ardle R, Hagan A, Al-Oraibi A, Hanjari M, Stephan B, et al. Attitudes and preferences towards screening for dementia with a focus on ethnic minority and low socio-economic groups: a systematic review of research studies written in the english language. J Alzheimers Dis 2024;100(4):1315-1331 [FREE Full text] [doi: 10.3233/JAD-240315] [Medline: 39031361]
29. Mc Ardle R, Dsouza A, Hagan A, Al-Oraibi A, Hanjari M, Stephan BCM, et al. Attitudes and preferences towards screening for dementia from the perspectives of healthcare professionals: an updated systematic review. Int J Geriatr Psychiatry 2025;40(2):e70057. [doi: 10.1002/gps.70057] [Medline: 39978789]
30. Wilson S, Tolley C, Mc Ardle R, Beswick E, Slight SP. Key considerations when developing and implementing digital technology for early detection of dementia-causing diseases among health care professionals: qualitative study. J Med Internet Res 2023;25:e46711 [FREE Full text] [doi: 10.2196/46711] [Medline: 37606986]
31. Bowling A. Research Methods in Health: Investigating Health and Health Services. Berkshire: McGraw-hill education (UK); 2014.
32. Payment guidance for researchers and professionals. National Institute of Health and Care Research. 2024. URL: https://www.learningforinvolvement.org.uk/content/resource/nihr-payment-guidance-for-researchers-and-professionals/ [accessed 2025-11-07]
33. Emmel N. Sampling and Choosing Cases in Qualitative Research: A Realist Approach. London: SAGE Publications Ltd; 2013.
34. Mental capacity act 2005. Legislation.gov.uk. 2005. URL: https://www.legislation.gov.uk/ukpga/2005/9/contents [accessed 2025-11-07]
35. Green J, Thorogood N. Qualitative Methods for Health Research. Thousand Oaks: Sage Publications; 2018.
36. Braun V, Clarke V. Thematic Analysis: A Practical Guide. London: sage; 2022.
37. Dhakal K. NVivo. J Med Libr Assoc 2022;110(2):270-272 [FREE Full text] [doi: 10.5195/jmla.2022.1271] [Medline: 35440911]
38. Cognitive impairment (2016). Canadian Task Force on Preventative Health Care. 2016. URL: https://canadiantaskforce.ca/guidelines/published-guidelines/cognitive-impairment/ [accessed 2025-11-07]
39. US Preventive Services Task Force, Owens DK, Davidson KW, Krist AH, Barry MJ, Cabana M, et al. Screening for cognitive impairment in older adults: US preventive services task force recommendation statement. JAMA 2020;323(8):757-763. [doi: 10.1001/jama.2020.0435] [Medline: 32096858]
40. Brayne C. Evidence still lacking for recommendation of screening for cognitive impairment in older adults. JAMA Intern Med 2020;180(4):483-484. [doi: 10.1001/jamainternmed.2019.7522] [Medline: 32096825]
41. Fowler NR, Perkins AJ, Gao S, Sachs GA, Boustani MA. Risks and benefits of screening for dementia in primary care: the Indiana university cognitive health outcomes investigation of the comparative effectiveness of dementia screening (IU CHOICE) trial. J Am Geriatr Soc 2020;68(3):535-543 [FREE Full text] [doi: 10.1111/jgs.16247] [Medline: 31792940]
42. Ethnic group, England and Wales: census 2021. Office for National Statistics (ONS). 2023. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/ethnicity/bulletins/ethnicgroupenglandandwales/census2021 [accessed 2025-11-07]
43. Hanrahan M, Wilson C, Keogh A, Barker S, Rochester L, Brittain K, et al. How can patients shape digital medicine? a rapid review of patient and public involvement and engagement in the development of digital health technologies for neurological conditions. Expert Rev Pharmacoecon Outcomes Res 2025;25(2):137-154 [FREE Full text] [doi: 10.1080/14737167.2024.2410245] [Medline: 39376020]
44. Cordell CB, Borson S, Boustani M, Chodosh J, Reuben D, Verghese J, Medicare Detection of Cognitive Impairment Workgroup. Alzheimer's association recommendations for operationalizing the detection of cognitive impairment during the medicare annual wellness visit in a primary care setting. Alzheimers Dement 2013;9(2):141-150 [FREE Full text] [doi: 10.1016/j.jalz.2012.09.011] [Medline: 23265826]
45. AAN recommends people 65+ be screened yearly for memory problems [news release]. American Academy of Neurology. URL: https://www.aan.com/PressRoom/Home/PressRelease/27482019 [accessed 2025-11-07]
46. Butler J, Watermeyer TJ, Matterson E, Harper EG, Parra-Rodriguez M. The development and validation of a digital biomarker for remote assessment of alzheimer's diseases risk. Digit Health 2024;10:20552076241228416 [FREE Full text] [doi: 10.1177/20552076241228416] [Medline: 38269369]
47. Griffanti L, Gillis G, O'Donoghue MC, Blane J, Pretorius PM, Mitchell R, et al. Adapting UK biobank imaging for use in a routine memory clinic setting: the Oxford brain health clinic. Neuroimage Clin 2022;36:103273 [FREE Full text] [doi: 10.1016/j.nicl.2022.103273] [Medline: 36451375]

XSL•FO
RenderX

48. O'Donoghue MC, Blane J, Gillis G, Mitchell R, Lindsay K, Semple J, et al. Oxford brain health clinic: protocol and research database. BMJ Open 2023;13(8):e067808 [FREE Full text] [doi: 10.1136/bmjopen-2022-067808] [Medline: 37541753]
49. Venkataraman AV, Kandangwa P, Lemmen R, Savla R, Beigi M, Hammond D. The SLaM Brain Health Clinic: a remote biomarker enhanced memory clinic for patients with mild cognitive impairment within an NHS mental health trust. medRxiv 2024. [doi: 10.1101/2024.04.23.24303268]
50. Jones D, Drewery R, Windle K, Humphrey S, de Paiva AF. Dementia prevention and the GP's role: a qualitative interview study. Br J Gen Pract 2024;74(741):e242-e249 [FREE Full text] [doi: 10.3399/BJGP.2023.0103] [Medline: 37549993]
51. Honeyman M, Maguire D, Evans H, Davies A. Digital technology and health inequalities: a scoping review. EuroHealthNet. 2020. URL: https://health-inequalities.eu/jwddb/digital-technology-and-health-inequalities-a-scoping-review-wales/ [accessed 2025-11-07]
52. 2023 consumer digital index. Lloyds Bank PLC. 2023. URL: https://www.lloydsbankinggroup.com/assets/pdfs/media/consumer-digital-index/2023-consumer-digital-index-report.pdf [accessed 2025-11-07]
53. Wilson S, Tolley C, Mc Ardle R, Lawson L, Beswick E, Hassan N, et al. Recommendations to advance digital health equity: a systematic review of qualitative studies. NPJ Digit Med 2024;7(1):173 [FREE Full text] [doi: 10.1038/s41746-024-01177-7] [Medline: 38951666]
54. Reddy S, Allan S, Coghlan S, Cooper P. A governance model for the application of AI in health care. J Am Med Inform Assoc 2020;27(3):491-497 [FREE Full text] [doi: 10.1093/jamia/ocz192] [Medline: 31682262]
55. NHS England. Fit for the future: 10 year health plan for England. NHS England. 2025. URL: https://www.england.nhs.uk/long-term-plan/
56. Mammen JR. Mapping Meaningful Symptoms and Impacts of Early Parkinson's Disease to Digital Outcome Measures. Island: University of Rhode; 2022.
57. Letter to Dr. martin daumer regarding DDT COA #000106. U.S. Food and Drug Administration. 2018. URL: https://www.fda.gov/drugs/clinical-outcome-assessment-coa-qualification-program/ddt-coa-000106-actibeltr-multiple-sclerosis [accessed 2025-11-07]
58. Patient decision aid: should I take a statin? National Institute for Health and Care Excellence. 2023. URL: https://www.nice.org.uk/guidance/ng238/resources/patient-decision-aid-on-should-i-take-a-statin-pdf-243780159 [accessed 2025-11-07]
59. Ethics and governance of artificial intelligence for health: guidance on large multi-modal models. World Health Organization. 2024. URL: https://www.who.int/publications/i/item/9789240084759 [accessed 2025-11-07]
60. Lekadir K, Feragen A, Fofanah AJ, Frangi AF, Buyx A, Emelie A, et al. FUTURE-AI: international consensus guideline for trustworthy and deployable artificial intelligence in healthcare. arXiv:2309.12325 2025. [doi: 10.1136/bmj.r340]

Abbreviations

DRI-SI: Dementia Research Institute Sleep Index
NHS: National Health Service
NIHR: National Institute for Health Research

Edited by J Wang; submitted 21.03.25; peer-reviewed by A Keogh, C Newton, S van der Landen; comments to author 26.09.25; revised version received 13.10.25; accepted 05.11.25; published 01.12.25.

Please cite as:
Mc Ardle R, Poole M, Horrocks S, King-Robson J, Schott JM, Sharp D, Harrison M, Robinson L
End User and Primary Care Physicians' Perspectives on Digital Innovations in Dementia Risk Detection: Focus on a Digital Sleep Biomarker
JMIR Aging 2025;8:e74307
URL: https://aging.jmir.org/2025/1/e74307
doi:10.2196/74307
PMID:

©Ríona Mc Ardle, Marie Poole, Sophie Horrocks, Josh King-Robson, Jonathan M Schott, David Sharp, Matthew Harrison, Louise Robinson. Originally published in JMIR Aging (https://aging.jmir.org), 01.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Associations of Wearable-Measured Sleep and Physical Activity With Memory Performance in Older Adults: Cross-Sectional Study With Actigraphy and MRI

Geng-Hao Liu[1,2,3,4], MD, MMS, PhD; Yueh-Hsiang Huang[5,1], MD, PhD; Tzu-Chiao Yuan[6], MD; Yueh-Peng Chen[7,8,9], PhD; Jung-Lung Hsu[10,11], MD; Shwu-Hua Lee[12,13], MD; Chih-Ming Lin[14,13,15], MD; Cheng-Hong Toh[16,1], MD; Ji-Tseng Fang[17,13], MD; Shih-Wei Lin[18,4,19], MD; Li-Pang Chuang[18,4,13,19], MD, PhD; Ning-Hung Chen[18,1,4,19], MD

[1]School of Traditional Chinese Medicine, College of Medicine, Chang Gung University, Taoyuan, Taiwan

[2]Graduate Institute of Traditional Chinese Medicine, School of Traditional Chinese Medicine, College of Medicine, Chang Gung University, Taoyuan, Taiwan

[3]Division of Acupuncture and Moxibustion, Center for Traditional Chinese Medicine, Linkou Chang Gung Memorial Hospital, Taoyuan, Taiwan

[4]Sleep Center, Taoyuan Chang Gung Memorial Hospital, Taoyuan, Taiwan

[5]Division of Chinese Internal Medicine, Center for Traditional Chinese Medicine, Chang Gung Memorial Hospital, Taipei, Taiwan

[6]Taoying Fengze Chinese Medicine Clinic, Fengze Chinese Medicine Clinic, Taoyuan, Taiwan

[7]Master of Science Degree Program in Innovation for Smart Medicine, Chang Gung University, Taoyuan, Taiwan

[8]Center for Artificial Intelligence in Medicine, Linkou Chang Gung Memorial Hospital, Taoyuan, Taiwan

[9]Division of Rheumatology, Allergy and Immunology, Linkou Chang Gung Memorial Hospital, Taoyuan, Taiwan

[10]Department of Neurology, New Taipei Municipal TuCheng Hospital (Built and Operated by Chang Gung Medical Foundation), New Taipei, Taiwan

[11]Graduate Institute of Mind, Brain and Consciousness, Taipei Medical University, Taipei, Taiwan

[12]Department of Psychiatry, Linkou Chang Gung Memorial Hospital, Taoyuan, Taiwan

[13]College of Medicine, Chang Gung University, Taoyuan, Taiwan

[14]Division of Internal Medicine, Chang Gung Memorial Hospital, Taipei, Taiwan

[15]Department of Health Management, Chang Gung Health and Culture Village, Taoyuan, Taiwan

[16]Department of Medical Imaging and Intervention, Linkou Chang Gung Memorial Hospital, Taoyuan, Taiwan

[17]Kidney Research Center, Department of Nephrology, Linkou Chang Gung Memorial Hospital, Taoyuan, Taiwan

[18]Department of Pulmonary and Critical Care Medicine, Linkou Chang Gung Memorial Hospital, Taoyuan, Taiwan

[19]Department of Respiratory Therapy, Linkou Chang Gung Memorial Hospital, Taoyuan, Taiwan

Corresponding Author:
Ning-Hung Chen, MD
Department of Pulmonary and Critical Care Medicine
Linkou Chang Gung Memorial Hospital
5 Fuxing St
Taoyuan, 333423
Taiwan
Phone: 886 3 328 1200 ext 2281
Email: ninghung@yahoo.com.tw

Abstract

Background: Cognitive decline is a common aspect of aging, and identifying modifiable lifestyle factors, such as physical activity and sleep, is crucial for promoting healthy brain aging. While both are individually linked to cognition, few studies have simultaneously assessed their independent and combined effects using objective wearable-based data, particularly in older Asian populations.

Objective: This study aimed to examine the independent and interactive effects of wearable-assessed sleep and physical activity parameters on memory performance in healthy older adults. We also explored whether age and hippocampal volume moderated these associations.

Methods: This prospective cross-sectional analysis included 88 cognitively healthy community-dwelling adults (≥60 years of age) from the *Integrating Systematic Data of Geriatric Medicine to Explore the Solution for Healthy Aging* cohort in Taiwan.

XSL•FO
RenderX

Participants underwent 12-day wrist-worn actigraphy, brain magnetic resonance imaging, and neuropsychological assessments. Light-intensity physical activity (LPA) and wake after sleep onset (WASO) were selected based on age-adjusted partial correlations with Consortium to Establish a Registry for Alzheimer's Disease Neuropsychological Battery memory scores. Multivariate regressions, age-stratified models (cutoff=72 years), and *PROCESS* moderation and mediation analyses were conducted, adjusting for age, education, daytime sleepiness, and hippocampal volume.

Results: Partial correlation analyses adjusting for age showed that higher LPA (r=0.260; P=.02) and lower WASO (r=–0.251; P=.02) were significantly associated with better memory scores. Age significantly moderated both effects: LPA was beneficial beyond 73.8 years of age, and WASO was detrimental beyond 71.1 years of age. Multivariate regression models confirmed that both WASO (β=–.044; P=.04) and LPA (β=.042; P=.01) were significant predictors of memory. In subgroup analyses (age ≥72 years), both LPA (β=.054; P=.04) and WASO (β=–.111; P=.01) remained significant predictors. Moderated mediation analyses showed that WASO was associated with reduced LPA (β=–.325; P=.03), but the indirect effect on memory via LPA was not significant. Instead, WASO exerted a direct and age-moderated effect on memory performance. Hippocampal volume moderated both associations, supporting the brain reserve hypothesis.

Conclusions: Our findings highlight WASO and LPA, as measured by wearable devices, as modifiable behavioral factors linked to memory function in older adults. The impact of these factors intensifies with advancing age and may be influenced by hippocampal reserve. Promoting daily light physical activity and maintaining sleep continuity may serve as accessible, age-tailored strategies for preserving cognitive health in aging populations.

Trial Registration: ClinicalTrials.gov NCT04207502; https://classic.clinicaltrials.gov/ct2/show/NCT04207502

(JMIR Aging 2025;8:e80584) doi:10.2196/80584

KEYWORDS

CERAD-NB; cognitive aging; hippocampus; light physical activity; memory function; sleep fragmentation; wake after sleep onset; wearable devices

Introduction

Cognitive decline is a well-recognized physiological aspect of aging. Even in the absence of dementia or mild cognitive impairment, older adults may experience subtle yet measurable decreases in memory performance and other cognitive domains [1]. These cognitive changes often occur alongside physical changes such as reductions in mobility, muscle strength, and endurance, which collectively affect quality of life and functional independence in late life. Among the most common behavioral alterations in aging populations are disrupted sleep patterns—such as changes in sleep timing and reduced sleep quality [2]—and reduced physical activity levels [3]. Growing evidence suggests that lifestyle factors play a critical role in both physical and cognitive health in older adults [4-6]. Given the lack of well-established pharmacological treatments for early-stage cognitive decline, modifiable lifestyle behaviors such as sleep and physical activity are increasingly recognized as vital contributors to cognitive reserve and healthy brain aging.

Many studies have investigated the association between self-reported sleep quality and cognitive function, although the findings have been inconsistent. Some reports have linked longer sleep latency and longer or shorter total sleep duration with poorer cognitive outcomes [7-10], including U-shaped relationships between sleep duration and cognitive performance. Studies using objective sleep measures have further identified associations between poor sleep quality—characterized by increased wake after sleep onset (WASO), prolonged sleep latency, and reduced sleep efficiency—and higher risk of cognitive impairment [11,12].

Physical activity is another important modifiable factor associated with cognitive aging. Observational studies have shown that higher levels of physical activity are linked to better cognitive function in older adults [13,14], whereas sedentary behavior (SB) is associated with poorer outcomes [15]. Interventional trials have also demonstrated cognitive benefits from regular physical exercise [16,17]. Mechanistically, physical activity may promote brain health through the release of muscle-derived myokines, which support neuroplasticity and anti-inflammatory processes [18]. Additionally, sarcopenia has been associated with cognitive decline, highlighting the potential interplay between muscle and brain health [19].

Although a growing body of literature has explored the impact of sleep or physical activity on cognitive functioning, studies that simultaneously examine both factors using objective, device-based methods remain limited, particularly among older adults in Asian populations. Furthermore, little is known about how these associations vary across different age thresholds in older people, despite evidence that the aging process may alter the relative influence of these lifestyle behaviors. In this study, we aimed to investigate the associations between sleep characteristics, physical activity levels, and memory function in healthy, community-dwelling older adults in Taiwan, using actigraphy-based monitoring and age-stratified as well as moderation analyses to explore potential age-related differences.

Methods

Study Population

This study was part of a prospective observational cohort established to create a comprehensive geriatric database for community-dwelling older adults in Taiwan without major physical or psychological disabilities. Community-dwelling adults aged 60 years and older were invited to participate in the *Integrating Systematic Data of Geriatric Medicine to Explore*

XSL•FO
RenderX

the Solution for Healthy Aging (ISDHA) study between September 2019 and October 2020 during annual routine health examinations at Chang Gung Memorial Hospital. Participants were residents of either the Chang Gung Health and Culture Village in Taoyuan or the Songshan District in Taipei, Taiwan. The inclusion criteria were as follows: (1) aged ≥60 years, (2) had visited Chang Gung Memorial Hospital at least once within the past year, and (3) had resided in Taiwan for more than 180 days within the year prior to enrollment. Exclusion criteria included (1) clinical evidence of major organ dysfunction; (2) a history of severe autoimmune disease; (3) ongoing cancer treatment at the time of recruitment; (4) antibiotic use within 1 month prior to enrollment; (5) cognitive impairment, defined as an Ascertain Dementia 8 (AD8) questionnaire score of ≥2 or Mini-Mental State Examination (MMSE) score of ≤26; (6) current follow-up at a cognitive disorder outpatient clinic; (7) depressive symptoms, defined as a Geriatric Depression Scale (GDS) score of ≥5; (8) a prior diagnosis of dementia or major depressive disorder; (9) severe hearing, visual, or cognitive impairment, preventing meaningful participation in interviews; or (10) physical frailty precluding the ability to stand or walk independently.

For this analysis, only participants who completed brain magnetic resonance imaging (MRI) and objective assessments of sleep quality and physical activity using actigraphy were included.

Demographic and Clinical Data Collection

Demographic information, including age, educational attainment, smoking status, and habitual exercise, was obtained through structured interviews. Medical history, such as the presence of hypertension, hyperlipidemia, diabetes mellitus, insomnia, obstructive sleep apnea, peptic ulcer disease, and anxiety, was collected from both electronic medical records and participant self-reports. Vital signs, including systolic and diastolic blood pressure and resting pulse rate, were measured using an automated sphygmomanometer. Body height and weight were measured using standardized procedures, and BMI was calculated as weight in kilograms divided by the square of height in meters (kg/m^2).

Blood Sample Collection and Laboratory Tests

Fasting venous blood samples were collected from all participants following an overnight fast of 8-12 hours. Laboratory tests included a complete blood count, liver and renal function panels, metabolic markers, serum vitamin D and albumin levels, as well as proinflammatory cytokines. All analyses were performed in the central laboratory of Chang Gung Memorial Hospital using standardized protocols.

Body Composition Assessments

Body composition was assessed using bioelectrical impedance analysis, a noninvasive and reliable method based on the electrical conductivity properties of different tissues [20]. Measurements were performed using a multifrequency segmental body composition analyzer (Tanita MC-780 MA; Tanita Corp), following the manufacturer's standard operating procedures. The parameters obtained included fat mass (kg), fat-free mass (kg), percentage of body fat, visceral fat rating, muscle mass (kg), bone mass (kg), total body water (TBW; kg), intracellular water (kg), extracellular water (ECW; kg), and basal metabolic rate (kcal).

Cognitive Function Assessments

Global cognitive status was assessed using the MMSE [21] and the AD8 questionnaire [22]. To specifically evaluate memory function, the Consortium to Establish a Registry for Alzheimer's Disease Neuropsychological Battery (CERAD-NB) was administered following standardized procedures [23]. The CERAD-NB is a widely validated tool for assessing cognitive performance in normal aging and Alzheimer disease [24] and includes several subtests: Verbal Fluency, the Modified Boston Naming Test, MMSE, Word List Memory, and Constructional Praxis [25]. The memory total score of the CERAD-NB was calculated by summing the scores from the 3 trials of the Word List Learning test, Delayed Recall, and Recognition subtests, with a maximum total score of 50 [25].

Subjective Sleep Assessments

Subjective sleep quality and daytime sleepiness were assessed using 2 standardized self-report questionnaires. The Pittsburgh Sleep Quality Index (PSQI) was used to evaluate overall sleep quality over the past month, while the Epworth Sleepiness Scale (ESS) assessed the general level of daytime sleepiness [26]. Both instruments have been widely validated in older adult populations and are commonly used in aging and sleep research.

Actigraphy-Based Monitoring of Physical Activity and Sleep

Objective monitoring of physical activity and sleep was performed using a wrist-worn accelerometer (Actiwatch 2; Philips Respironics), which incorporates a piezoelectric accelerometer to detect movement and a photodiode sensor to measure ambient light (illuminance in lux). The device recorded activity as counts per minute and light exposure at 60-second epochs. Participants were instructed to wear the device on the nondominant wrist for 14 consecutive days. Data from the first and last days were excluded from analysis to avoid partial-day bias. Sleep and wake periods were scored by a trained sleep specialist using Philips Actiware software (version 6.0.9) in conjunction with participants' sleep diaries. Epochs with activity counts below the wake threshold were classified as sleep, while those above the threshold were classified as wake. Circadian rhythm parameters were derived using nonparametric analysis. M10 and L5 were defined as the mean activity counts per minute during the most active 10 consecutive hours and the least active 5 consecutive hours of the day, respectively, reflecting rest-activity patterns [27].

Actigraphy recordings were considered valid if participants wore the device for at least 7 valid days (≥10 hours of daytime wear per day) without prolonged removal. Participants with fewer than 7 valid days (n=1) or who did not complete the actigraphy assessment (n=7) were excluded from the final analysis. The average number of valid actigraphy days before excluding the first and last partial days was 13.6 (SD 1.3; range 9-14) days. After trimming these 2 boundary days to avoid partial-day bias, a continuous 12-day dataset was used for the final analysis. Participants wore the Actiwatch 2 for an average

XSL•FO
RenderX

of 1412 (SD 12.5) minutes per day (~23.5 hours), with all valid recordings exceeding 870 minutes per day (~14.5 hours), indicating high adherence and minimal device removal during monitoring. Most participants (84/88, 95.5%) provided 13-14 days of valid recordings, demonstrating excellent compliance and data completeness.

Physical activity parameters were categorized using validated counts-per-minute thresholds obtained from Actiwatch 2 actigraphy data [28]. SB was defined as activity of <145 counts/minute (corresponding to <1.5 metabolic equivalents of task [METs]), light-intensity physical activity (LPA) as 145-274 counts/minute (~1.5-3 METs), moderate-intensity physical activity (MPA) as 274-597 counts/minute (~3-6 METs), and vigorous-intensity physical activity (VPA) as >597 counts/minute (>6 METs). Wake activity time was calculated as the total time spent in any activity category (SB, LPA, MPA, or VPA) during waking hours. Total activity counts during the wake period were also recorded as an indicator of overall activity level.

Light exposure during the wake period was evaluated using the following measures: wake light illuminance, defined as the average light intensity (lux); wake light time, representing the total duration (minutes) exposed to light; and wake light exposure, calculated as the cumulative product of light intensity and time (lux × minutes).

Sleep-related parameters were also derived from the actigraphy data. These included time in bed (TIB), defined as the duration between bedtime and final wake time; total sleep time (TST), referring to the total number of minutes scored as sleep; sleep onset latency, defined as the time from attempting to sleep to sleep onset; sleep efficiency, calculated as the percentage of time asleep while in bed; and WASO, representing the total number of minutes awake between sleep onset and the final awakening. In addition, light exposure during the sleep period was evaluated, including sleep light illuminance (average light intensity), sleep light time (duration of light exposure during sleep), sleep light exposure (lux × minutes), and sleep total activity (cumulative movement counts during sleep).

Summary statistics of subjective questionnaires and actigraphy-derived parameters, including sleep, physical activity, and light exposure, are presented in Textbox 1.

Textbox 1. Clinical assessments included subjective questionnaires (Pittsburgh Sleep Quality Index [PSQI] and Epworth Sleepiness Scale [ESS]) and actigraphy-derived parameters for sleep, physical activity, and light exposure. The actigraphy-derived parameters were obtained from 12 consecutive days of wrist-worn recordings.

Subjective assessments

- Sleep parameters include (1) PSQI (past 1 month) and (2) ESS (past 1 month).

Objective assessments

- Actigraphy-derived sleep parameters include (1) time in bed (min), (2) total sleep time (min), (3) onset latency (min), (4) sleep efficiency (%), and (5) wake after sleep onset (min).
- Actigraphy-derived physical activity parameters include (1) M10 (counts/min; indicates the mean activity during the most active 10 consecutive hours), (2) L5 (counts/min; indicates the mean activity during the least active 5 consecutive hours), (3) awaking total activity (counts), (4) sleep total activity (counts), and (5) physical activity time (min), which is further categorized into (i) sedentary behavior time (min), (ii) light-intensity physical activity time (min), (iii) moderate-intensity physical activity time (min), and (iv) vigorous-intensity physical activity time (min).
- Actigraphy-derived light exposure parameters include (1) wake light exposure (lux-min) and (2) sleep light exposure (lux-min).

Brain MRI Acquisition and Analysis

Brain MRI was performed using 3-Tesla scanners (Discovery MR750w and Discovery MR750; GE Healthcare). For each participant, 3 imaging sequences were acquired: a 3D T1-weighted anatomical scan, a 3D T2 fluid-attenuated inversion recovery scan, and diffusion tensor imaging. Gray matter and white matter volumes were estimated from the high-resolution T1-weighted images using the Computational Anatomy Toolbox (version 12.7), implemented within the Statistical Parametric Mapping software (version 12; Wellcome Trust Centre for Neuroimaging).

Statistical Analysis

All statistical analyses were performed using SPSS Statistics software (version 25.0; IBM Corp). Descriptive statistics were used to summarize participant characteristics, including age, education level, sleep and physical activity parameters, and cognitive performance. The normality of continuous variables was assessed using the Kolmogorov-Smirnov test. As most variables did not meet the assumption of normality, nonparametric methods were applied, and between-group comparisons were conducted using the Mann-Whitney *U* test.

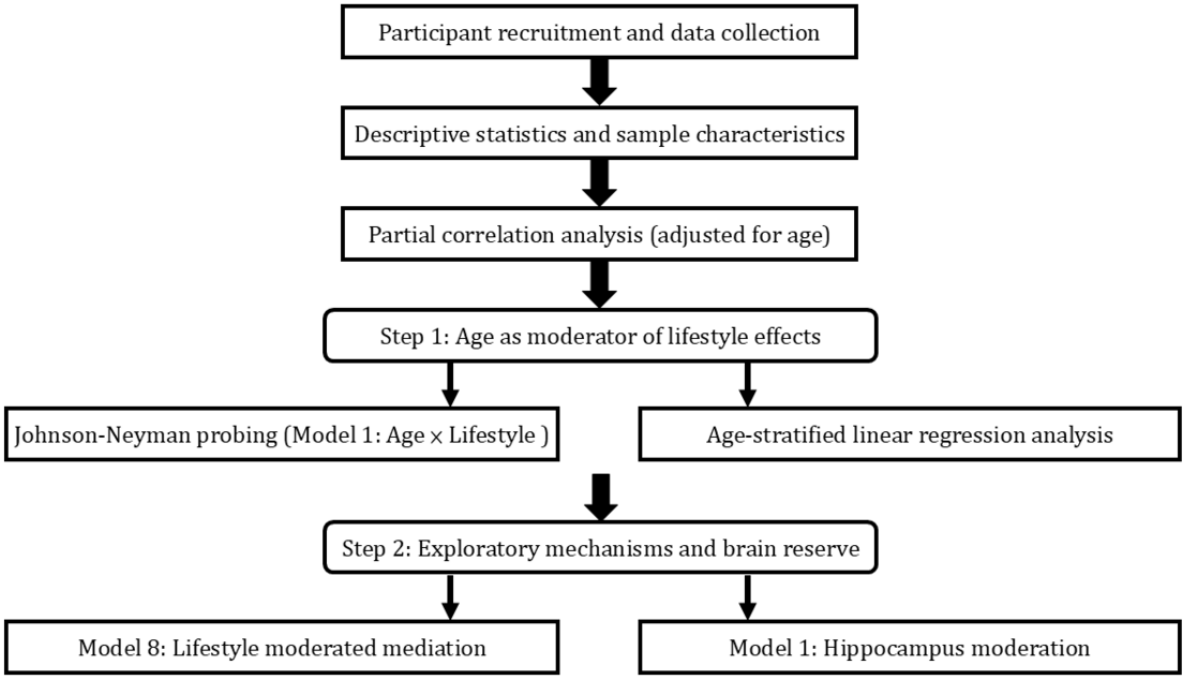
To examine the associations between actigraphy-derived sleep and physical activity variables and memory performance, partial correlation analyses adjusted for age were conducted. Based on these results, WASO (r=–0.251; P=.02) and LPA time (r=0.26; P=.02) were identified as candidate modifiable lifestyle factors and selected for further analysis. Subsequently, multivariate linear regression analyses were performed to evaluate the independent contributions of WASO and LPA time to CERAD-NB memory scores, adjusting for relevant covariates, including age, education level, ESS score, and mean hippocampal volume. To investigate whether the effects of lifestyle factors on memory performance were moderated by age, moderation analyses were conducted using the *PROCESS* macro (version 4.2) for SPSS, developed by Andrew F Hayes. Model 1 was used to test interaction effects between age and each lifestyle factor (eg, age × WASO), and the

XSL•FO
RenderX

Johnson-Neyman technique was applied to identify the age threshold at which the interaction became statistically significant. Based on these results, a cutoff age of ≥72 years was established and used for subsequent age-stratified regression analyses.

Finally, exploratory analyses were conducted to examine potential mechanisms linking sleep and physical activity to memory performance. This included moderated mediation analyses using Model 8 of the *PROCESS* macro to assess indirect effects and additional moderation analyses (Model 1) evaluating the role of hippocampal volume as a potential moderator, in alignment with the brain reserve hypothesis. All statistical tests were 2-tailed, with *P* values of <.05 considered statistically significant. For moderation and mediation models, bias-corrected bootstrap 95% CIs based on 5000 resamples were used to assess the robustness of indirect and interaction effects. The overall analytic framework is illustrated in Figure 1.

Figure 1. Analysis flowchart of the study design.



This diagram illustrates the sequential analytic framework used to investigate the associations between lifestyle factors and memory performance in older adults. After participant recruitment and descriptive analysis, partial correlation analysis (adjusted for age) was conducted to identify candidate actigraphy-derived lifestyle parameters. Step 1 examined whether age moderated the relationship between these lifestyle variables and memory performance, using Johnson-Neyman probing (*PROCESS* Model 1) and age-stratified linear regression. Step 2 further explored potential mechanisms and the brain reserve hypothesis through lifestyle-moderated mediation analysis (Model 8) and mean hippocampal volume moderation (Model 1).

Ethical Considerations

The study was approved by the Institutional Review Board of Chang Gung Medical Foundation, Taiwan (approval number 201900702A3). Written informed consent was obtained from all participants. To ensure privacy and confidentiality, all data were deidentified using unique study codes, with no personally identifiable information stored with research data. Participants received NT $100 (US $3.10) for nutritional supplementation and NT $200 (US $6.20) for transportation for each visit involving additional blood sampling or questionnaire-based assessments.

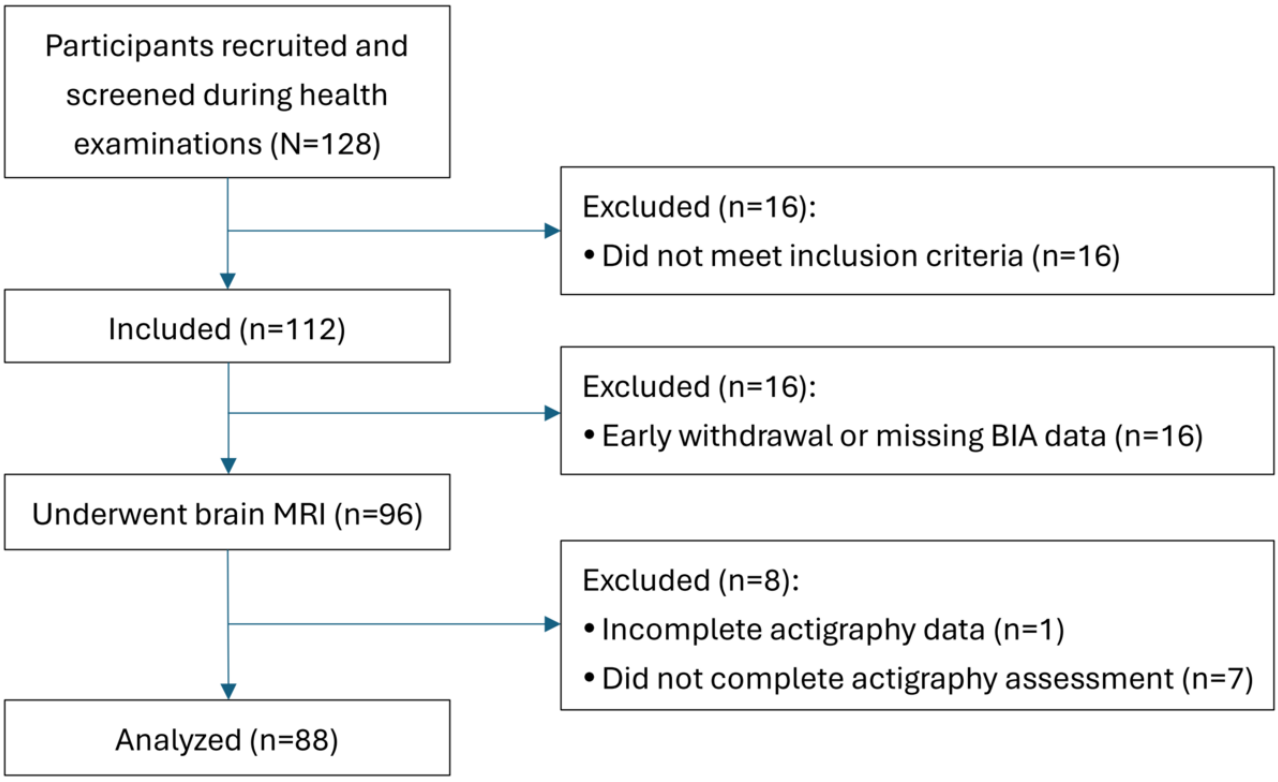
Results

Participants' Characteristics

Of the 128 individuals initially enrolled in the study, 112 met the inclusion criteria. Participants who did not complete the body composition assessment (bioelectrical impedance analysis), brain MRI, or valid actigraphy recordings were excluded. After these exclusions, a total of 88 participants were included in the final analysis (flow diagram in Figure 2). Their demographic characteristics, cognitive performance, subjective sleep assessments, and actigraphy-derived sleep and physical activity parameters are summarized in the total column of Table 1.

XSL•FO
RenderX

Figure 2. Flow diagram of the inclusion process. BIA: bioelectrical impedance analysis; MRI: magnetic resonance imaging.



The mean age of participants was 73.5 (SD 7) years, and 55.7% (49/88) were female. Educational attainment was distributed as follows: 8% (7/88) had completed primary school, 23.9% (21/88) had completed junior or senior high school (including vocational programs), 52.3% (46/88) had attended college or university, and 15.9% (14/88) held a graduate degree. The mean BMI was 23.7 (SD 3.4) kg/m^2, and the average systolic and diastolic blood pressures were 138.4 (SD 17) mm Hg and 75 (SD 9.6) mm Hg, respectively. Regarding habitual exercise, 75% (66/88) of participants reported engaging in regular physical activity exceeding 150 minutes per week, 19.3% (17/88) exercised less than 150 minutes per week, and only 5.7% (5/88) reported no regular exercise habit.

The average CERAD-NB memory score was 37 (SD 5.7). In terms of subjective sleep assessments, the mean scores were 6.4 (SD 3.8) for the PSQI and 3.7 (SD 3.7) for the ESS.

Actigraphy-derived sleep parameters revealed an average TIB of 468 (SD 81.2) minutes and TST of 391.5 (SD 80.7) minutes. The mean sleep onset latency was 11.6 (SD 13.4) minutes, sleep efficiency was 83.6% (SD 8.2%), and WASO averaged 56.2 (SD 25.6) minutes. In terms of physical activity, participants showed an average M10 activity of 267 (SD 97.2) counts/minute and L5 activity of 21.7 (SD 15.6) counts/minute. Daily activity distributions included a mean LPA time of 141.5 (SD 33.2) minutes, MPA of 236.2 (SD 67.6) minutes, and VPA of 141.3 (SD 88.5) minutes. Average wake light exposure was 136576.4 (SD 129264.3) lux-minute, and sleep light exposure was 497.9 (SD 1103.9) lux-minute. A comprehensive summary of subjective and actigraphy-derived sleep and activity parameters is provided in Textbox 1.

XSL•FO
RenderX

Table 1. Characteristics of 88 participants analyzed in the *Integrating Systematic Data of Geriatric Medicine to Explore the Solution for Healthy Aging* study.

Demographics	Total (n=88)	Age		*P* value
		<72 years (n=44)	≥72 years (n=44)	
Age (years), mean (SD)	73.5 (7)	67.9 (2.4)	79.1 (5.3)	<.001
Sex (female), n (%)	49 (55.7)	25 (56.8)	24 (54.5)	.83
Systolic blood pressure (mm Hg), mean (SD)	138.4 (17)	135 (14.7)	141.7 (18.5)	.08
Diastolic blood pressure (mm Hg), mean (SD)	75 (9.6)	74.9 (10.3)	75.1 (8.8)	.95
BMI (kg/m^2), mean (SD)	23.7 (3.4)	23.2 (2.8)	24.2 (3.8)	.33
Education, n (%)				.34
Primary school	7 (8)	3 (6.8)	4 (9.1)	
Junior or senior high (including vocational)	21 (23.9)	9 (20.5)	12 (27.3)	
University or college	46 (52.3)	24 (54.5)	22 (50)	
Graduate school	14 (15.9)	8 (18.2)	6 (13.6)	
Exercise (in 1 week), n (%)				.52
No exercise habit	5 (5.7)	1 (2.3)	4 (9.1)	
<150 minutes	17 (19.3)	9 (20.5)	8 (18.2)	
≥150 minutes	66 (75)	34 (77.3)	32 (72.7)	
Questionnaires, mean (SD)				
CERAD-NB[a] memory score	37 (5.7)	38.7 (4.2)	35.4 (6.5)	.02
Pittsburgh Sleep Quality Index	6.4 (3.8)	6.5 (3.9)	6.3 (3.7)	.90
Epworth Sleepiness Scale	3.7 (3.7)	3.7 (3.7)	3.8 (3.7)	.96
Actigraphy-derived sleep parameters, mean (SD)				
Time in bed (min)	468 (81.2)	444.3 (67.6)	491.8 (87.2)	.006
Total sleep time (min)	391.5 (80.7)	370.5 (73.1)	412.5 (83.2)	.02
Onset latency (min)	11.6 (13.4)	11.2 (17)	12 (8.6)	.17
Sleep efficiency (%)	83.6 (8.2)	83.3 (10.5)	83.9 (5.2)	.62
Wake after sleep onset (min)	56.2 (25.6)	52.9 (27.3)	59.4 (23.8)	.12
Actigraphy-derived physical activity parameters, mean (SD)				
M10[b] (counts/min)	267 (97.2)	296.9 (99.9)	237.1 (85.4)	.006
L5[c] (counts/min)	21.7 (15.6)	21.4 (20)	22.1 (9.8)	.06
Awaking total activity (counts)	280508.1 (107005.9)	291065.6 (107660.9)	230846.9 (80836.7)	.007
Sleep total activity (counts)	9119.8 (4950.1)	7796.7 (5378.5)	9549.3 (3944.9)	.005
Physical activity time (min)	972.6 (81.9)	996.7 (66.2)	948.5 (89.4)	.01
SB[d] time (min)	453.7 (117.5)	441.8 (121.7)	465.5 (113.2)	.34
LPA[e] time (min)	141.5 (33.2)	140.1 (30.2)	143 (36.2)	.55
MPA[f] time (min)	236.2 (67.6)	240.5 (58.2)	231.9 (76.2)	.81
VPA[g] time (min)	141.3 (88.5)	174.4 (99.1)	108.2 (61.6)	<.001
Actigraphy-derived light exposure parameters (lux-min), mean (SD)				
Wake light exposure	136576.4 (129264.3)	175062.9 (149387.8)	98089.9 (91973.2)	<.001
Sleep light exposure	497.9 (1103.9)	247.7 (362.3)	748 (1485.4)	.32

[a]CERAD-NB: Consortium to Establish a Registry for Alzheimer's Disease Neuropsychological Battery.

[b]M10 indicates the mean activity (counts/min) during the most active 10 consecutive hours.

XSL•FO
RenderX

[c]L5 indicates the mean activity (counts/min) during the least active 5 consecutive hours.

[d]SB: Sedentary behavior.

[e]LPA: Light-intensity physical activity.

[f]MPA: moderate-intensity physical activity.

[g]VPA: vigorous-intensity physical activity.

Correlation Between Sleep, Physical Activity, and Memory Performance

Partial correlation analyses adjusted for age were conducted to examine the associations between actigraphy-derived parameters and memory performance, as measured by the CERAD-NB memory score (Table 2).

Among sleep-related variables, WASO showed a significant negative correlation with memory performance (r=–0.251; P=.02), indicating that greater sleep fragmentation was associated with lower memory scores. Other sleep parameters, such as TIB (r=–0.199; P=.07), sleep efficiency (r=0.152; P=.16), and onset latency (r=–0.119; P=.28), did not reach statistical significance.

Table 2. Partial correlations between actigraphy parameters and Consortium to Establish a Registry for Alzheimer's Disease Neuropsychological Battery memory score.

Actigraphy variables	Values (adjusted for age)	
	r	*P* value
Sleep		
Time in bed (min)	–0.199	.07
Total sleep time (min)	–0.082	.45
Onset latency (min)	–0.119	.28
Sleep efficiency (%)	0.152	.16
Wake after sleep onset (min)	–0.251	.02
Physical activity		
M10[a] (counts/min)	0.216	.046
L5[b] (counts/min)	–0.098	.37
Awaking total activity (counts)	0.166	.13
Sleep total activity (counts)	–0.160	.14
Physical activity time (min)	0.196	.07
SB[c] time (min)	–0.132	.23
LPA[d] time (min)	0.260	.02
MPA[e] time (min)	0.170	.12
VPA[f] time (min)	0.120	.27
Light exposure		
Wake light exposure (lux-min)	–0.144	.19
Sleep light exposure (lux-min)	–0.167	.12

[a]M10 indicates the mean activity (counts/min) during the most active 10 consecutive hours.

[b]L5 indicates the mean activity (counts/min) during the least active 5 consecutive hours.

[c]SB: sedentary behavior.

[d]LPA: light-intensity physical activity.

[e]MPA: moderate-intensity physical activity.

[f]VPA: vigorous-intensity physical activity.

Regarding physical activity, several metrics showed modest positive correlations with memory. Notably, LPA time was significantly associated with better memory scores (r=0.260; P=.02). Additionally, M10—the mean activity (counts/min) during the most active 10 consecutive hours—was also significantly correlated with memory scores (r=0.216; P=.046). However, LPA was prioritized for subsequent multivariate and moderation analyses because it represents a clearly defined, low-intensity, and modifiable behavioral category, whereas M10 reflects a composite index of total daytime movement combining multiple activity intensities. Thus, LPA provided a more interpretable and intervention-relevant indicator for

XSL•FO
RenderX

modeling behavioral effects on memory. Total physical activity time (r=0.196; P=.07) and MPA (r=0.170; P=.12) showed nonsignificant trends in the expected direction. No significant associations were observed for sedentary time, VPA, or nighttime sleep activity.

Light exposure variables, including wake light exposure and sleep light exposure, also did not show significant correlations with memory performance (P>.10). Taken together, WASO and LPA time emerged as the 2 behavioral parameters showing statistically significant and physiologically meaningful associations with memory and were therefore selected for subsequent regression and moderation models to examine their independent and age-dependent effects on cognitive performance.

Age as a Moderator of the Association Between Lifestyle Factors and Memory

To determine whether age moderated the relationship between modifiable lifestyle factors and memory performance, moderation analyses were conducted using *PROCESS* Model 1 with the Johnson-Neyman technique. As illustrated in Figure 3, age significantly moderated both associations of LPA time and WASO with CERAD-NB memory scores. In contrast, no significant moderating effects of age were observed for MPA, VPA, or combined moderate to vigorous physical activity (MVPA) on memory performance (Figure S1 in Multimedia Appendix 1).

Figure 3. Interaction effects of age with light-intensity physical activity (LPA) time and wake after sleep onset (WASO) on Consortium to Establish a Registry for Alzheimer's Disease Neuropsychological Battery (CERAD-NB) memory performance. (A) Interaction between LPA time and age on CERAD-NB memory scores. (B) Johnson-Neyman plot showing the age-dependent significance threshold for the association between LPA time and memory. (C) Interaction between WASO and age on CERAD-NB memory scores. (D) Johnson-Neyman plot showing the age-dependent significance threshold for the association between WASO and memory.

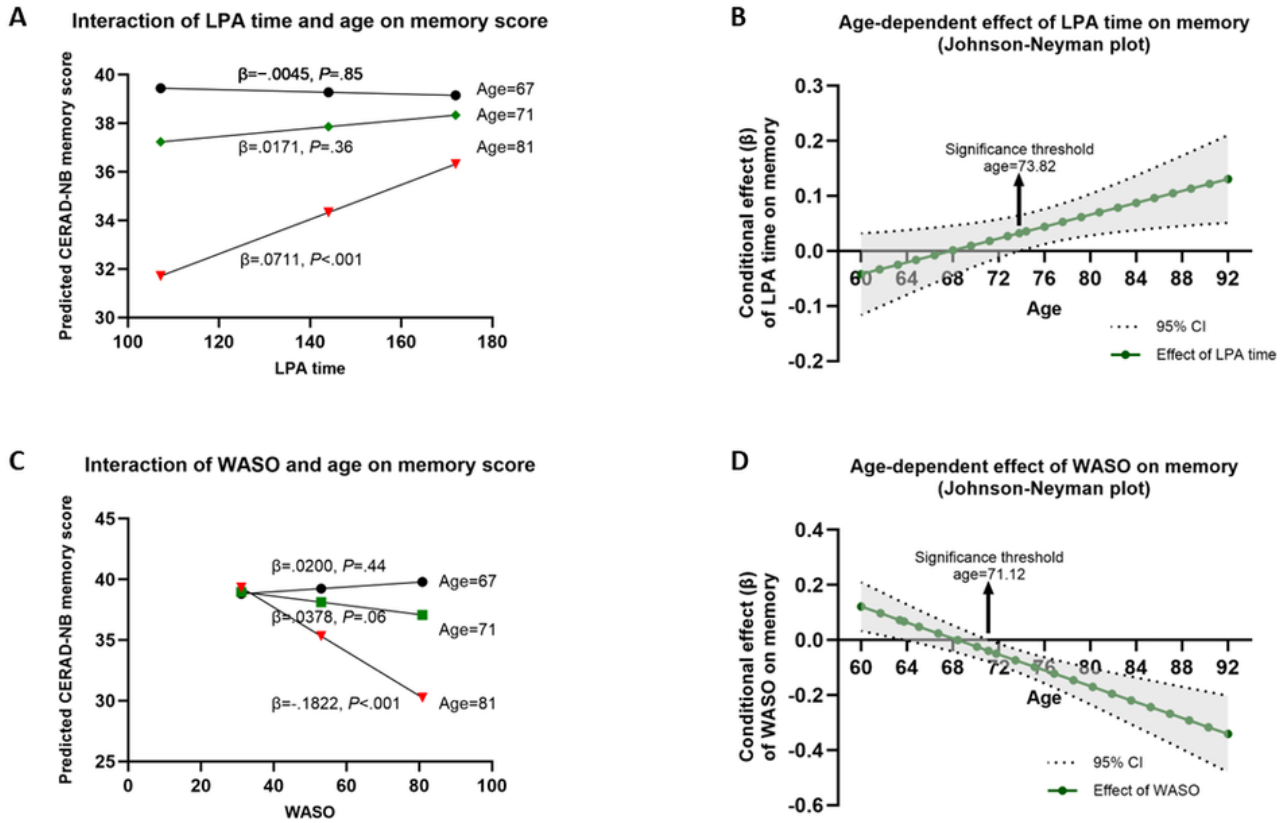


Figure 3A illustrates the interaction between LPA time and age on CERAD-NB memory scores. While no significant association was found in younger adults (eg, aged 67 or 71 years), higher LPA time was significantly associated with better memory performance at older ages (eg, 81 years; β=.0711; P<.001). Figure 3B shows the Johnson-Neyman plot, indicating that the association between LPA time and memory became statistically significant at age 73.82 years onward.

Figure 3C depicts the interaction between WASO and age. While WASO was not significantly associated with memory among younger individuals, it showed a significant negative relationship at older ages (eg, 81 years; β=–.1822; P<.001). The Johnson-Neyman plot in Figure 3D indicates a significance threshold at age 71.12 years, after which increased WASO was associated with worse memory performance.

All models were tested using the *PROCESS* macro Model 1 (moderation analysis), examining the interaction between age and each lifestyle factor.

For LPA time, the conditional effect on memory became statistically significant at age 73.82 years, above which greater LPA time was associated with higher memory scores. Conversely, for WASO, the association with memory became significantly negative starting from age 71.12 years, indicating that sleep fragmentation had a more detrimental impact on

XSL•FO
RenderX

memory among older participants. Based on these results, we stratified participants into 2 groups using an age cutoff of 72 years for subsequent subgroup analyses.

Compared to younger participants, older adults (aged ≥72 years), as shown in the age-stratified columns of Table 1, had significantly lower CERAD-NB memory scores (35.4, SD 6.5 minutes vs 38.7, SD 4.2; *P*=.02), longer TIB (491.8, SD 87.2 minutes vs 444.3, SD 67.6 minutes; *P*=.006), and greater TST (412.5, SD 83.2 minutes vs 370.5, SD 73.1 minutes; *P*=.02). However, no significant age differences were observed in sleep onset latency, sleep efficiency, or WASO.

In terms of physical activity, participants aged ≥72 years had significantly lower M10 activity (237.1, SD 85.4 counts/min vs 296.9, SD 99.9 counts/min; *P*=.006), shorter VPA time (108.2, SD 61.6 minutes vs 174.4, SD 99.1 minutes; *P*<.001), and reduced awake total activity and sleep total activity counts (*P*<.01). They also exhibited significantly lower wake light exposure (*P*<.001).

Further comparisons in Table S1 in Multimedia Appendix 1 revealed significant age-related physiological and structural differences. Older participants had slower gait speed, lower estimated glomerular filtration rate, and smaller hippocampal volume, including both right and left hippocampi (*P*<.001). Additionally, significant age-related differences were observed in visceral fat rating, ECW and TBW ratio, and regional brain structure volumes (eg, anterior cingulate cortex), consistent with expected patterns of biological aging. Moreover, among the inflammatory markers assessed, only interleukin-8 showed a significant increase with age (*P*=.02), whereas other cytokines (interleukin-1β, interleukin-6, interleukin-10, tumor necrosis factor-alpha) did not differ significantly between groups, suggesting mild low-grade inflammation in this otherwise healthy cohort.

Together, these findings support the use of age-stratified analyses to further explore how sleep and activity factors impact memory differently across the aging continuum.

Multivariate Models and Brain-Based Mechanisms

To assess the independent contributions of LPA time and WASO to memory performance, multivariate linear regression analyses were conducted, adjusting for age, education level, ESS scores, and mean hippocampal volume. As shown in Table 3, both LPA time (β=.042; *P*=.01) and WASO (β=–.044; *P*=.04) remained significant predictors of CERAD-NB memory scores in the overall sample. Increasing age was independently associated with lower memory scores (β=–.318; *P*=.003), whereas education level, ESS, and hippocampal volume did not show significant contributions in the total model. Among participants aged ≥72 years, the effects of LPA time (β=.054; *P*=.04) and WASO (β=–.111; *P*=.01) on memory remained statistically significant, suggesting that these modifiable lifestyle factors exert a stronger influence in later age.

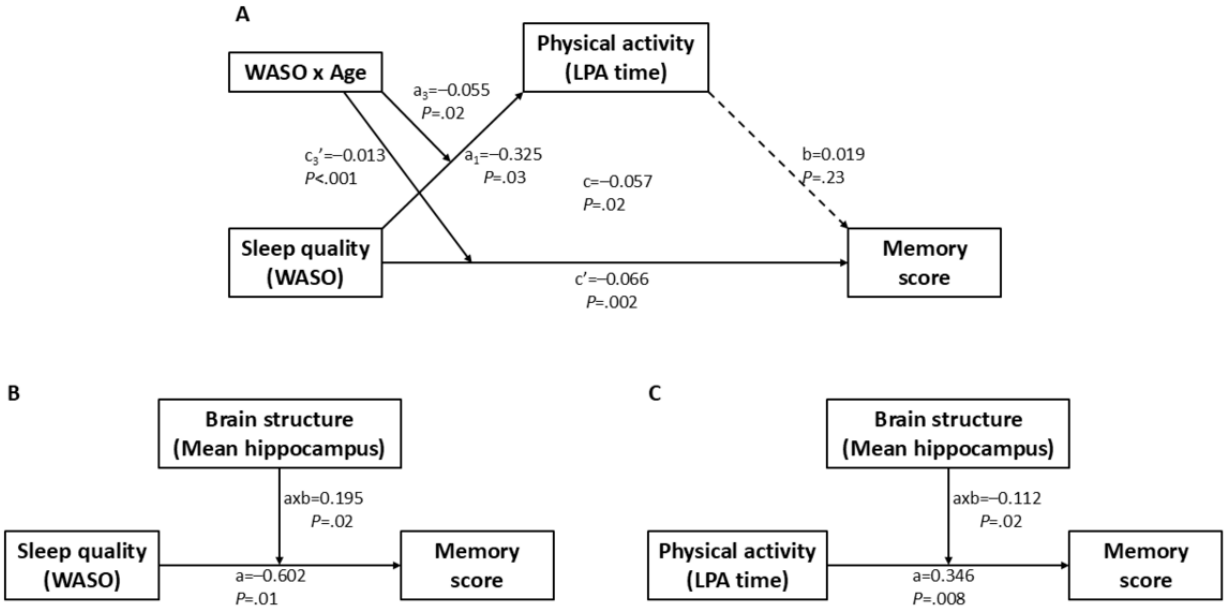
Table 3. Multivariate linear regression models examining associations of light-intensity physical activity (LPA) time and wake after sleep onset (WASO) with Consortium to Establish a Registry for Alzheimer's Disease Neuropsychological Battery memory scores. Two adjusted models are presented: the left column reports results for the full sample, and the right column reports results for participants aged ≥72 years. Both overall models were statistically significant (*P*<.001). All models were adjusted for age, education level, Epworth Sleepiness Scale score, and mean hippocampal volume. Regression coefficients (β) and corresponding *P* values are reported.

Factors	Models (total)		Models (age ≥72 years)	
	β	*P* value	β	*P* value
Age	–.318	.003	–.268	.15
Education	1.144	.09	.886	.39
Epworth Sleepiness Scale	–.207	.17	–.257	.25
Mean hippocampal volume	.054	.98	6.053	.09
LPA time	.042	.01	.054	.04
WASO	–.044	.04	–.111	.01

To further explore the potential mechanisms linking sleep fragmentation, physical activity, and memory decline, 2 exploratory models were tested using the *PROCESS* macro (Figure 4A). A moderated mediation model (Model 8) showed that higher WASO was associated with lower LPA time (a_1=–0.325; *P*=.03), and this relationship was significantly moderated by age (a_3=–0.055; *P*=.02), suggesting that sleep disruption more strongly reduced activity levels in older adults. However, LPA itself was not a significant predictor of memory performance (b=0.019; *P*=.23), and the indirect effect of WASO on memory via LPA was not statistically significant. Instead, WASO exerted a direct effect on memory performance (c'=–0.066; *P*=.002) that was also moderated by age (c_3'=–0.013; *P*<.001), supporting an age-dependent vulnerability to sleep fragmentation on cognitive function.

XSL•FO
RenderX

Figure 4. Exploratory models examining potential mechanisms linking sleep fragmentation, physical activity, and memory performance in older adults. (A) Moderated mediation model examining whether the association between wake after sleep onset (WASO) and memory performance is mediated by light-intensity physical activity (LPA) time and moderated by age. (B) Moderation model testing whether mean hippocampal volume moderates the association between WASO and memory performance. (C) Moderation model testing whether mean hippocampal volume moderates the association between LPA time and memory performance.



In Figure 4A, a moderated mediation model (*PROCESS* Model 8) was used to examine whether the relationship between sleep fragmentation (measured by WASO) and memory performance was mediated by LPA time and whether age moderated these behavioral pathways. Higher WASO was associated with reduced LPA (a_1=–0.325; *P*=.03), and this effect was moderated by age (a_3=–0.055; *P*=.02), indicating a stronger negative effect of sleep disruption on physical activity in older individuals. However, LPA was not a significant predictor of memory performance (b=0.019; *P*=.23), and the indirect path from WASO to memory via LPA was not statistically significant. In contrast, the direct effect of WASO on memory remained significant after accounting for LPA (c'=–0.066; *P*=.002), and this effect was also moderated by age (c_3'=–0.013; *P*<.001), indicating that the negative impact of WASO on memory increased with advancing age.

In addition, moderation analyses (*PROCESS* Model 1) were used to test the brain reserve hypothesis by examining hippocampal volume as a moderator of the LPA-memory and WASO-memory associations (Figures 4B and 4C). For WASO, hippocampal volume significantly moderated the association with memory (interaction effect: a×b=0.195; *P*=.02), suggesting that individuals with greater hippocampal volume were less affected by sleep fragmentation. Similarly, for LPA, the moderation effect was significant (a×b=–0.112; *P*=.02), indicating that low hippocampal reserve amplified the cognitive impact of reduced physical activity.

These findings provide evidence for both behavioral (activity-mediated) and structural (hippocampal-modulated) pathways linking modifiable lifestyle factors to memory outcomes in older adults.

In Figure 4B, moderation models (*PROCESS* Model 1) examined whether mean hippocampal volume moderated the effects of WASO and LPA on memory, consistent with the brain reserve hypothesis. These models showed that mean hippocampal volume significantly moderated the relationships between both WASO (a=–0.602; *P*=.01; a×b=0.195; *P*=.02) and LPA (a=0.346; *P*=.008; a×b=–0.112; *P*=.02) with memory scores. These findings support the idea that greater structural brain reserve may buffer the negative impact of sleep disruption and low activity on cognitive outcomes in aging.

These findings underscore WASO as a direct and age-sensitive predictor of memory performance, with its negative impact particularly evident in older individuals and those with lower hippocampal reserve. Although LPA did not mediate the sleep-memory relationship, it remained an independent modifiable factor, also moderated by brain reserve. Together, sleep continuity and light physical activity represent key behavioral targets for cognitive preservation in aging.

Discussion

Principal Findings

In this study, using wearable device–based monitoring, we identified 2 objectively measured lifestyle factors—LPA time and WASO—that were significantly associated with memory performance in community-dwelling older adults. Higher LPA time was linked to better memory scores, while greater WASO was associated with poor memory. These associations remained

XSL•FO
RenderX

robust after adjusting for age, education, daytime sleepiness, and hippocampal volume. Importantly, age moderated both effects, with LPA showing a protective role from age 73.8 years onward, and WASO exerting a detrimental influence from age 71.1 years onward. Subgroup regression analyses confirmed that both LPA (β=.054; *P*=.04) and WASO (β=–0.111; *P*=.01) were stronger predictors of memory among older participants (≥72 years of age). Moreover, hippocampal volume moderated these associations, consistent with the brain reserve hypothesis.

Beyond these primary findings, this study contributes novel insights in 3 ways. First, it simultaneously examined sleep and physical activity using objective wearable-based monitoring, rather than focusing on one lifestyle factor in isolation. Second, it integrated these measures with validated neuropsychological testing and structural brain imaging, providing a multimodal framework to explore behavioral and neural determinants of cognitive health. Third, by testing age and hippocampal reserve as moderators, it highlights how the effects of lifestyle factors intensify with advancing age and reduced neural reserve. Taken together, this work advances understanding of how modifiable lifestyle behaviors interact with biological aging processes to influence memory in older adults, particularly within an underrepresented Asian population.

This study highlights LPA and sleep continuity (measured as WASO) as key modifiable lifestyle factors associated with memory performance in community-dwelling older adults. Our findings on the correlation between sleep and memory function align with those of other studies. Actigraphy-based research has identified links between sleep disturbances, prolonged sleep duration, and poorer cognitive performance [6,12,20,29]. Recent evidence from an accelerometer-based study in *JMIR Aging* further demonstrated that greater night-to-night variability in sleep efficiency was associated with poorer cognitive performance in older adults [30], reinforcing the role of objective sleep fragmentation in cognitive aging. Consistently, wearable-derived rhythm metrics, such as interdaily stability and sleep-wake fragmentation measured by the myRhythmWatch platform, have been linked to cognitive impairment among community-dwelling elders [31], supporting the concept that circadian rhythm disruption contributes to cognitive vulnerability. Similarly, good sleep quality, as objectively measured by polysomnography, has been shown to correlate with enhanced cognitive performance in older adults [11]. Various potential mechanisms have been proposed to explain how sleep influences cognition in older people. Sleep disturbances might negatively impact synaptic plasticity by disrupting neuronal and glial signaling pathways [32]. Inadequate sleep could promote neuroinflammation and impair neurogenesis, potentially leading to neurodegeneration in the hippocampus [33]. Furthermore, increased sleep fragmentation, often due to sleep-disordered breathing, could induce neuron apoptosis and brain atrophy through inflammatory and oxidative pathways [34]. A polysomnography-based study also highlighted the relationship between sleep and glymphatic function, which, in turn, was associated with neuropsychological function [35] and gray matter volumes [36].

Numerous studies have identified physical activity as a crucial, modifiable lifestyle factor influencing cognitive and memory function in aging populations [14,37]. However, the optimal duration, frequency, and intensity of physical activity necessary to mitigate cognitive decline in older adults remain uncertain. Several mechanisms have been proposed to explain the cognitive benefits of physical activity. One key hypothesis centers on its impact on brain structure; that is, MRI-based studies have shown that regular activity is associated with preserved brain microstructure and reduced hippocampal atrophy [38,39]. At the molecular level, exercise stimulates the release of myokines from skeletal muscle, promoting the production of neurotrophic factors such as brain-derived neurotrophic factor and insulin-like growth factor-1 [18]. These neurotrophic mediators support synaptic plasticity, neurogenesis, and angiogenesis in regions critical to memory, including the hippocampus [40-44]. This interaction between muscle and brain could account for the observed benefits of physical activity on cognitive function in older people. Reports suggest that higher intensity and variability in physical activity are linked to a reduced risk of dementia [45,46]. While MVPA is often the focus of public health recommendations, our findings indicate that even low-intensity movement, particularly LPA, plays a meaningful role in supporting cognitive health. Interestingly, in our relatively healthy older adult sample, 75% (66/88 participants) reported having regular exercise habits exceeding 150 minutes per week, and accelerometry data showed that MPA time was substantially higher than LPA time (mean MPA 236.2, SD 67.6 minutes/day vs mean LPA: 141.5, SD 33.2 minutes/day). The absence of a significant association between MPA and memory in our data may partly reflect a ceiling effect, as most participants already met or exceeded recommended activity levels. However, this explanation remains speculative and was not directly tested. Future studies with broader variability in physical activity levels are warranted to verify this interpretation. In contrast, LPA, which includes incidental movement such as light walking, standing, or household activities, showed greater interindividual variation and a stronger correlation with cognitive function. Consistent with this interpretation, supplementary analyses (Figure S1 in Multimedia Appendix 1) demonstrated that neither MPA, VPA, nor their combined MVPA time showed significant age-moderated associations with memory performance, further supporting that low-intensity, sustained movement may be a more sensitive behavioral indicator of cognitive health in highly active older adults.

Our findings extend the emerging evidence supporting incidental light physical activity as a health-promoting behavior. A recent large-scale study [47] demonstrated that incidental light physical activity was associated with a dose-dependent reduction in cardiovascular events and all-cause mortality, even in populations not engaging in structured exercise. Building on these results, our study provides novel evidence that LPA is also beneficial for brain health, particularly memory performance in older adults. Additionally, physical activity may contribute uniquely to cognition through nonfatiguing, sustained movement that stabilizes cerebral perfusion, minimizes sleep disruption, and potentially supports glymphatic clearance. Recent evidence suggests that long-term exercise can enhance brain waste removal via glymphatic and meningeal lymphatic pathways [48], providing a novel mechanism linking physical activity with cognitive health. Taken together, these findings

XSL•FO
RenderX

suggest that while maintaining structured exercise remains important [49], encouraging older adults to stay consistently active throughout the day, even at low intensity, may be an equally vital strategy for preserving memory and cognitive function.

Our findings indicate that age significantly moderates the impact of lifestyle factors on memory performance. Using Johnson-Neyman analysis, we identified distinct age thresholds—73.8 years for LPA and 71.1 years for WASO—beyond which the cognitive effects of these modifiable behaviors became statistically significant. While both LPA and WASO were significant predictors of memory performance in the overall sample, subgroup regression analyses showed that their predictive effects were even stronger among participants aged ≥72 years. This observation aligns with emerging evidence that aging is not a uniform or linear process but one marked by biological inflection points across the lifespan. For instance, a large-scale proteomic study [50] revealed 3 major age-related transitions in the human plasma proteome—around ages 34, 60, and 78 years—indicating phase-like shifts in biological aging trajectories. Similarly, a recent multiomics analysis [51] showed cliff-like drops in functional integrity occurring around ages 44 and 60 years, supporting a model of nonlinear, stepwise aging rather than gradual decline. While our study does not directly assess molecular aging, our results extend this nonlinear perspective by showing that the cognitive impact of behavioral factors such as physical activity and sleep continuity intensifies with advancing age. These findings carry important clinical implications. They suggest that behavioral risks are not static across age but rather become more influential in late life, even in otherwise healthy older adults. This emphasizes the importance of early and consistent engagement in health-promoting behaviors throughout adulthood, while also recognizing that targeted interventions in older age may yield greater benefits. Age-stratified strategies that prioritize sleep consolidation and daily activity may help optimize cognitive preservation throughout aging.

Our exploratory analyses suggest that the relationship between sleep and memory is not only behavioral but also shaped by neurobiological and age-related factors. Although sleep fragmentation (WASO) was associated with reduced LPA in older adults, this behavioral pathway did not significantly mediate memory outcomes. Instead, WASO exerted a robust and direct negative influence on memory, with the effect strengthening significantly with increasing age. The absence of a significant indirect pathway suggests that sleep fragmentation may impair memory primarily through direct neurophysiological mechanisms—such as disrupted synaptic homeostasis or impaired glymphatic clearance—rather than solely through behavioral reductions in activity. This interpretation aligns with prior evidence linking sleep continuity to neural restoration and waste clearance processes. In parallel, hippocampal volume moderated the associations of both WASO and LPA with memory performance. Participants with larger hippocampi showed relative resilience to the adverse effects of poor sleep and low activity, whereas those with smaller hippocampal volume demonstrated stronger behavior-cognition associations. These findings support the brain reserve hypothesis, which posits that individuals with greater structural brain integrity are more capable of withstanding behavioral and pathological insults without manifesting cognitive decline. While hippocampal volume has previously been linked to cognitive resilience in the context of neuropathology, depression, and sleep disorders [52-54], our results extend this model to modifiable lifestyle factors measured objectively via wearable technologies in cognitively healthy older adults.

To explore the reasons behind the varying correlation patterns between sleep, physical activity, and memory function in participants older than 72 years, we compared several physical, laboratory, and imaging parameters between participants younger and older than this age threshold. Significant differences were observed in some variables, such as pulse rate, visceral fat rating, the ratio of ECW to TBW, indicators of muscle function, and hippocampal volume, but not in the appendicular skeletal muscle mass index between the 2 groups. Previous studies have indicated that higher levels of body fat negatively affect brain volume and contribute to memory decline [55]. Furthermore, higher ratios of ECW to TBW have been linked to poorer cognition [56]. Whether these differences in body composition are responsible for diminishing the effects of physical activity on memory function in individuals aged 72 years or older requires further investigation.

In addition to these physiological factors, systemic inflammation may also play a role in mediating age-related differences. Inflammatory markers, including interleukin-1β, interleukin-6, interleukin-10, and tumor necrosis factor-alpha, were assessed in this study (Table S1 in Multimedia Appendix 1). Among these, only interleukin-8 levels were significantly higher in older adults, while other cytokines showed no marked differences. This pattern suggests that low-grade inflammation may exist but remains relatively mild in this healthy, community-dwelling sample. The modest change in IL-8 could reflect early immune activation associated with aging and sleep disruption, consistent with prior evidence linking chronic sleep fragmentation to subtle increases in inflammatory signaling [57-59]. Future studies incorporating oxidative stress and glial activation markers could help clarify whether inflammatory processes mediate the observed age-specific behavioral and cognitive effects.

In this study, we used the Actiwatch 2 accelerometer to objectively assess physical activity and estimate sleep quality and duration. As a lightweight and easy-to-wear device, actigraphy enables continuous monitoring over multiple days or weeks [60], thereby capturing behavioral patterns that more closely reflect individuals' habitual, real-world lifestyles. While actigraphy is known to overestimate TST and underestimate nocturnal awakenings when compared to polysomnography [60], it remains a practical and scalable tool in population-based and longitudinal research [61]. In an era of increasingly accessible wearable technology, actigraphy offers valuable insight into daily lifestyle rhythms and provides meaningful, ecologically valid data to inform behavioral recommendations and health interventions, especially in community and aging populations where in-laboratory polysomnography is often not feasible.

XSL•FO
RenderX

Our findings contribute new insights by integrating wearable-derived behavioral data, validated neuropsychological memory outcomes, and structural brain imaging into a unified analytic model. The demonstration that sleep and activity behaviors interact with age and hippocampal reserve to influence memory performance underscores the complexity and heterogeneity of cognitive aging. Importantly, these results suggest that individuals with lower hippocampal volume—who may be at higher risk of future cognitive decline—could particularly benefit from targeted interventions that improve sleep continuity and promote consistent physical activity. This highlights the potential value of combining lifestyle-based strategies with brain-based risk stratification in the design of personalized cognitive health programs for older adults.

From a translational perspective, these findings underscore the potential of promoting incidental, low-intensity movement, such as walking, household activities, and standing breaks, as realistic behavioral goals for older adults who may find structured exercise challenging. Public health strategies emphasizing total daily movement, rather than only moderate to vigorous exercise, may therefore be particularly beneficial in aging populations.

This study has several notable strengths. First, we used objective, wearable-based monitoring to assess both sleep and physical activity over an extended period in a real-world setting, avoiding the biases associated with self-reported lifestyle data. Second, by focusing on memory-specific cognitive outcomes, rather than global cognition, we provide a more targeted understanding of functional decline in aging. Third, we incorporated structural neuroimaging data to test how brain reserve, operationalized as hippocampal volume, interacts with behavioral risk factors. To our knowledge, this is one of the few studies to examine the joint impact of objectively measured lifestyle factors and neural reserve on memory performance in healthy older adults, using advanced moderation and mediation models to capture complex interactions between age, behavior, and brain structure.

Despite these strengths, several limitations should be acknowledged. First, although this was a prospectively designed cross-sectional study, the absence of longitudinal follow-up limits our ability to infer causal relationships between physical activity, sleep, and memory function. Future longitudinal or interventional studies are warranted to verify the temporal and causal pathways suggested by our findings. Second, participants were recruited from a single hospital-based health examination program in northern Taiwan, which may introduce selection bias and limit external validity. Third, our sample consisted primarily of relatively healthy, community-dwelling older adults with generally high physical activity levels. While this ensured sample homogeneity and reduced confounding, it also excluded high-risk individuals, such as those with cognitive impairment or major chronic diseases, thus limiting the generalizability of our results to more clinically vulnerable populations. Fourth, participants were aware that their activity and sleep were being monitored, which may have temporarily increased activity levels and slightly reduced the accuracy of habitual behavior estimates. Fifth, although actigraphy is not the gold standard for evaluating sleep architecture or activity intensity, it remains a US Food and Drug Administration–approved, user-friendly tool that enables unobtrusive and ecologically valid data collection under free-living conditions. In an era of widespread wearable technology use [62], actigraphy offers scalable and meaningful behavioral data for aging research. Sixth, polysomnography data were collected as part of the broader ISDHA cohort but were not included in this analysis because this study focused on wearable-based lifestyle measures to capture real-world behavior; polysomnography findings will be reported separately. Seventh, we did not account for other potentially influential factors, such as diet, socioeconomic status, or social engagement, that may also modulate cognitive outcomes. Finally, although hippocampal volume was examined as a moderator, other neural or molecular mechanisms (eg, functional connectivity or serum neurotrophic and inflammatory markers) were not analyzed. While basic inflammatory cytokines were assessed, more specific biomarkers of oxidative stress or neuroinflammation (eg, glial or endothelial activation indices) should be incorporated in future studies to clarify underlying biological pathways. Additionally, the proposed ceiling effect for moderate physical activity should be interpreted cautiously, as it was not formally tested. Collectively, these limitations highlight the need for longitudinal, multisite studies integrating multimodal biomarkers to validate and expand our findings. Nonetheless, this study provides an essential foundation for identifying modifiable behavioral factors and developing age- and brain-based strategies to preserve cognitive health in older adults.

Conclusions

In this study of cognitively healthy older adults, we demonstrated that 2 modifiable lifestyle factors—sleep continuity and LPA, objectively measured via wearable devices—were significantly associated with memory performance. These associations were further shaped by age and hippocampal volume, emphasizing the interplay between behavioral and neurobiological factors in cognitive aging. Our findings provide empirical support for the brain reserve hypothesis and highlight that the cognitive impact of lifestyle behaviors becomes more pronounced in later life and among individuals with lower neural reserve. Encouraging consistent, light-intensity activity throughout the day, together with maintaining sleep continuity, may represent a practical, scalable, and age-tailored strategy to support cognitive health in older adults. Future longitudinal and intervention studies are warranted to validate these pathways and inform personalized behavioral approaches for preserving memory function across aging populations.

XSL•FO
RenderX

Acknowledgments

Data collection, data management, and IT services are based on the Structured Research and Medical Informatics Cloud (SRMIC) platform system, provided by the Clinical Trial Center, Chang Gung Memorial Hospital, Linkou, Taiwan, Republic of China, with support from the Ministry of Health and Welfare under grant MOHW110-TDU-B-212-124005. We thank the Genomic and Proteomic Core Laboratory of Chang Gung Memorial Hospital for the assistance with the experiment.

Funding

This research was supported by Chang Gung Memorial Hospital under grants CORPG3J0371, CLRPG3H0015, CMRPG3N1101, and CMRPG3N1102; by the Ministry of Health and Welfare, Taiwan (grant MOHW110-TDU-B-212-124005); and by the National Science and Technology Council, Taiwan, under grants MOST 111-2320-B-182-013 and NSTC 114-2320-B-182-009.

Data Availability

The data collected and analyzed for this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: GHL and NHC
Methodology: JLH, SHL, CHT, NHC, and JTF
Software: YPC
Validation: GHL and YHH
Formal analysis: GHL, YHH, and TCY
Investigation: GHL
Resources: CML
Data curation: JTF
Writing—original draft: YHH and TCY
Writing—review & editing: GHL and NHC
Visualization: YHH
Supervision: NHC and JTF
Project administration: JTF
Funding acquisition: JTF

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interactions of physical activity measures with age on memory performance, and age-group differences in physical, laboratory, and brain magnetic resonance imaging characteristics.
[DOCX File , 124 KB - aging_v8i1e80584_app1.docx]

References

1. Harada CN, Natelson Love MC, Triebel KL. Normal cognitive aging. Clin Geriatr Med 2013;29(4):737-752 [FREE Full text] [doi: 10.1016/j.cger.2013.07.002] [Medline: 24094294]
2. Feinsilver SH. Normal and abnormal sleep in the elderly. Clin Geriatr Med 2021;37(3):377-386. [doi: 10.1016/j.cger.2021.04.001] [Medline: 34210444]
3. DiPietro L. Physical activity in aging: changes in patterns and their relationship to health and function. J Gerontol A Biol Sci Med Sci 2001;56 Spec No 2:13-22. [doi: 10.1093/gerona/56.suppl_2.13] [Medline: 11730234]
4. Knoops KTB, de Groot LCPGM, Kromhout D, Perrin A, Moreiras-Varela O, Menotti A, et al. Mediterranean diet, lifestyle factors, and 10-year mortality in elderly European men and women: the HALE project. JAMA 2004;292(12):1433-1439. [doi: 10.1001/jama.292.12.1433] [Medline: 15383513]
5. Clare L, Wu YT, Teale JC, MacLeod C, Matthews F, Brayne C, CFAS-Wales study team. Potentially modifiable lifestyle factors, cognitive reserve, and cognitive function in later life: a cross-sectional study. PLoS Med 2017;14(3):e1002259 [FREE Full text] [doi: 10.1371/journal.pmed.1002259] [Medline: 28323829]
6. Kimura N, Aso Y, Yabuuchi K, Ishibashi M, Hori D, Sasaki Y, et al. Modifiable lifestyle factors and cognitive function in older people: a cross-sectional observational study. Front Neurol 2019;10:401 [FREE Full text] [doi: 10.3389/fneur.2019.00401] [Medline: 31068892]
7. Suh SW, Han JW, Lee JR, Byun S, Kwon SJ, Oh SH, et al. Sleep and cognitive decline: a prospective nondemented elderly cohort study. Ann Neurol 2018;83(3):472-482. [doi: 10.1002/ana.25166] [Medline: 29394505]

XSL•FO
RenderX

8. Faubel R, López-García E, Guallar-Castillón P, Graciani A, Banegas JR, Rodríguez-Artalejo F. Usual sleep duration and cognitive function in older adults in Spain. J Sleep Res 2009;18(4):427-435. [doi: 10.1111/j.1365-2869.2009.00759.x] [Medline: 19691473]
9. Kronholm E, Sallinen M, Suutama T, Sulkava R, Era P, Partonen T. Self-reported sleep duration and cognitive functioning in the general population. J Sleep Res 2009;18(4):436-446. [doi: 10.1111/j.1365-2869.2009.00765.x] [Medline: 19732318]
10. Tworoger SS, Lee S, Schernhammer ES, Grodstein F. The association of self-reported sleep duration, difficulty sleeping, and snoring with cognitive function in older women. Alzheimer Dis Assoc Disord 2006;20(1):41-48. [doi: 10.1097/01.wad.0000201850.52707.80] [Medline: 16493235]
11. Bastien CH, Fortier-Brochu E, Rioux I, LeBlanc M, Daley M, Morin CM. Cognitive performance and sleep quality in the elderly suffering from chronic insomnia. Relationship between objective and subjective measures. J Psychosom Res 2003;54(1):39-49. [doi: 10.1016/s0022-3999(02)00544-5] [Medline: 12505554]
12. Blackwell T, Yaffe K, Ancoli-Israel S, Schneider JL, Cauley JA, Hillier TA, Study of Osteoporotic Fractures Group. Poor sleep is associated with impaired cognitive function in older women: the study of osteoporotic fractures. J Gerontol A Biol Sci Med Sci 2006;61(4):405-410. [doi: 10.1093/gerona/61.4.405] [Medline: 16611709]
13. Hillman CH, Weiss EP, Hagberg JM, Hatfield BD. The relationship of age and cardiovascular fitness to cognitive and motor processes. Psychophysiology 2002;39(3):303-312. [doi: 10.1017/s0048577201393058] [Medline: 12212649]
14. Hayes SM, Alosco ML, Hayes JP, Cadden M, Peterson KM, Allsup K, et al. Physical activity is positively associated with episodic memory in aging. J Int Neuropsychol Soc 2015;21(10):780-790 [FREE Full text] [doi: 10.1017/S1355617715000910] [Medline: 26581790]
15. Falck RS, Davis JC, Liu-Ambrose T. What is the association between sedentary behaviour and cognitive function? A systematic review. Br J Sports Med 2017;51(10):800-811. [doi: 10.1136/bjsports-2015-095551] [Medline: 27153869]
16. Langlois F, Vu TTM, Chassé K, Dupuis G, Kergoat M, Bherer L. Benefits of physical exercise training on cognition and quality of life in frail older adults. J Gerontol B Psychol Sci Soc Sci 2013;68(3):400-404. [doi: 10.1093/geronb/gbs069] [Medline: 22929394]
17. Ruscheweyh R, Willemer C, Krüger K, Duning T, Warnecke T, Sommer J, et al. Physical activity and memory functions: an interventional study. Neurobiol Aging 2011;32(7):1304-1319. [doi: 10.1016/j.neurobiolaging.2009.08.001] [Medline: 19716631]
18. Pedersen BK. Physical activity and muscle-brain crosstalk. Nat Rev Endocrinol 2019;15(7):383-392. [doi: 10.1038/s41574-019-0174-x] [Medline: 30837717]
19. Scisciola L, Fontanella RA, Surina, Cataldo V, Paolisso G, Barbieri M. Sarcopenia and cognitive function: role of myokines in muscle brain cross-talk. Life (Basel) 2021;11(2):173 [FREE Full text] [doi: 10.3390/life11020173] [Medline: 33672427]
20. Falck RS, Best JR, Davis JC, Liu-Ambrose T. The independent associations of physical activity and sleep with cognitive function in older adults. J Alzheimers Dis 2018;63(4):1469-1484. [doi: 10.3233/JAD-170936] [Medline: 29782311]
21. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
22. Galvin JE, Roe CM, Xiong C, Morris JC. Validity and reliability of the AD8 informant interview in dementia. Neurology 2006;67(11):1942-1948. [doi: 10.1212/01.wnl.0000247042.15547.eb] [Medline: 17159098]
23. Morris JC, Heyman A, Mohs RC, Hughes JP, van Belle G, Fillenbaum G, et al. The consortium to establish a registry for Alzheimer's Disease (CERAD). Part I. clinical and neuropsychological assessment of Alzheimer's disease. Neurology 1989;39(9):1159-1165. [doi: 10.1212/wnl.39.9.1159] [Medline: 2771064]
24. Sotaniemi M, Pulliainen V, Hokkanen L, Pirttilä T, Hallikainen I, Soininen H, et al. CERAD-neuropsychological battery in screening mild Alzheimer's disease. Acta Neurol Scand 2012;125(1):16-23. [doi: 10.1111/j.1600-0404.2010.01459.x] [Medline: 21198445]
25. Chandler MJ, Lacritz LH, Hynan LS, Barnard HD, Allen G, Deschner M, et al. A total score for the CERAD neuropsychological battery. Neurology 2005;65(1):102-106. [doi: 10.1212/01.wnl.0000167607.63000.38] [Medline: 16009893]
26. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The pittsburgh sleep quality index: a new instrument for psychiatric practice and research. Psychiatry Res 1989;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
27. Gonçalves BSB, Cavalcanti PR, Tavares GR, Campos TF, Araujo JF. Nonparametric methods in actigraphy: an update. Sleep Sci 2014;7(3):158-164 [FREE Full text] [doi: 10.1016/j.slsci.2014.09.013] [Medline: 26483921]
28. Neil-Sztramko SE, Rafn BS, Gotay CC, Campbell KL. Determining activity count cut-points for measurement of physical activity using the Actiwatch2 accelerometer. Physiol Behav 2017;173:95-100. [doi: 10.1016/j.physbeh.2017.01.026] [Medline: 28108333]
29. Cavuoto MG, Ong B, Pike KE, Nicholas CL, Bei B, Kinsella GJ. Objective but not subjective sleep predicts memory in community-dwelling older adults. J Sleep Res 2016;25(4):475-485. [doi: 10.1111/jsr.12391] [Medline: 26868539]
30. Sakal C, Li T, Li J, Yang C, Li X. Association between sleep efficiency variability and cognition among older adults: cross-sectional accelerometer study. JMIR Aging 2024;7:e54353 [FREE Full text] [doi: 10.2196/54353] [Medline: 38596863]

XSL•FO
RenderX

31. Jones CD, Wasilko R, Zhang G, Stone KL, Gujral S, Rodakowski J, et al. Detecting sleep/wake rhythm disruption related to cognition in older adults with and without mild cognitive impairment using the myrhythmwatch platform: feasibility and correlation study. JMIR Aging 2025;8:e67294 [FREE Full text] [doi: 10.2196/67294] [Medline: 40194310]
32. Havekes R, Vecsey CG, Abel T. The impact of sleep deprivation on neuronal and glial signaling pathways important for memory and synaptic plasticity. Cell Signal 2012;24(6):1251-1260 [FREE Full text] [doi: 10.1016/j.cellsig.2012.02.010] [Medline: 22570866]
33. Zhu B, Dong Y, Xu Z, Gompf HS, Ward SAP, Xue Z, et al. Sleep disturbance induces neuroinflammation and impairment of learning and memory. Neurobiol Dis 2012;48(3):348-355 [FREE Full text] [doi: 10.1016/j.nbd.2012.06.022] [Medline: 22776332]
34. Nair D, Dayyat EA, Zhang SX, Wang Y, Gozal D. Intermittent hypoxia-induced cognitive deficits are mediated by NADPH oxidase activity in a murine model of sleep apnea. PLoS One 2011;6(5):e19847 [FREE Full text] [doi: 10.1371/journal.pone.0019847] [Medline: 21625437]
35. Ma J, Chen M, Liu G, Gao M, Chen N, Toh CH, et al. Effects of sleep on the glymphatic functioning and multimodal human brain network affecting memory in older adults. Mol Psychiatry 2025;30(5):1717-1729. [doi: 10.1038/s41380-024-02778-0] [Medline: 39397082]
36. Siow TY, Toh CH, Hsu J, Liu G, Lee S, Chen N, et al. Association of sleep, neuropsychological performance, and gray matter volume with glymphatic function in community-dwelling older adults. Neurology 2022;98(8):e829-e838. [doi: 10.1212/WNL.0000000000013215] [Medline: 34906982]
37. Bherer L, Erickson KI, Liu-Ambrose T. A review of the effects of physical activity and exercise on cognitive and brain functions in older adults. J Aging Res 2013;2013:657508 [FREE Full text] [doi: 10.1155/2013/657508] [Medline: 24102028]
38. Dawe RJ, Yu L, Leurgans SE, James BD, Poole VN, Arfanakis K, et al. Physical activity, brain tissue microstructure, and cognition in older adults. PLoS One 2021;16(7):e0253484 [FREE Full text] [doi: 10.1371/journal.pone.0253484] [Medline: 34232955]
39. Hashimoto M, Araki Y, Takashima Y, Nogami K, Uchino A, Yuzuriha T, et al. Hippocampal atrophy and memory dysfunction associated with physical inactivity in community-dwelling elderly subjects: the Sefuri study. Brain Behav 2017;7(2):e00620 [FREE Full text] [doi: 10.1002/brb3.620] [Medline: 28239530]
40. Erickson KI, Voss MW, Prakash RS, Basak C, Szabo A, Chaddock L, et al. Exercise training increases size of hippocampus and improves memory. Proc Natl Acad Sci USA 2011;108(7):3017-3022 [FREE Full text] [doi: 10.1073/pnas.1015950108] [Medline: 21282661]
41. Cassilhas RC, Antunes HKM, Tufik S, de Mello MT. Mood, anxiety, and serum IGF-1 in elderly men given 24 weeks of high resistance exercise. Percept Mot Skills 2010;110(1):265-276. [doi: 10.2466/PMS.110.1.265-276] [Medline: 20391891]
42. Cassilhas RC, Viana VAR, Grassmann V, Santos RT, Santos RF, Tufik S, et al. The impact of resistance exercise on the cognitive function of the elderly. Med Sci Sports Exerc 2007;39(8):1401-1407 [FREE Full text] [doi: 10.1249/mss.0b013e318060111f] [Medline: 17762374]
43. Lista I, Sorrentino G. Biological mechanisms of physical activity in preventing cognitive decline. Cell Mol Neurobiol 2010;30(4):493-503. [doi: 10.1007/s10571-009-9488-x] [Medline: 20041290]
44. Vaynman S, Gomez-Pinilla F. License to run: exercise impacts functional plasticity in the intact and injured central nervous system by using neurotrophins. Neurorehabil Neural Repair 2005;19(4):283-295. [doi: 10.1177/1545968305280753] [Medline: 16263961]
45. Angevaren M, Vanhees L, Wendel-Vos W, Verhaar HJJ, Aufdemkampe G, Aleman A, et al. Intensity, but not duration, of physical activities is related to cognitive function. Eur J Cardiovasc Prev Rehabil 2007;14(6):825-830. [doi: 10.1097/HJR.0b013e3282ef995b] [Medline: 18043306]
46. Podewils LJ, Guallar E, Kuller LH, Fried LP, Lopez OL, Carlson M, et al. Physical activity, APOE genotype, and dementia risk: findings from the cardiovascular health cognition study. Am J Epidemiol 2005;161(7):639-651. [doi: 10.1093/aje/kwi092] [Medline: 15781953]
47. Stamatakis E, Biswas RK, Koemel NA, Sabag A, Pulsford R, Atkin AJ, et al. Dose response of incidental physical activity against cardiovascular events and mortality. Circulation 2025;151(15):1063-1075 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.124.072253] [Medline: 40228066]
48. Yoo R, Kim J, Moon HY, Park JY, Cheon S, Shin H, et al. Long-term physical exercise facilitates putative glymphatic and meningeal lymphatic vessel flow in humans. Nat Commun 2025;16(1):3360 [FREE Full text] [doi: 10.1038/s41467-025-58726-1] [Medline: 40204790]
49. Dubbeldam R, Stemplewski R, Pavlova I, Cyma-Wejchenig M, Lee S, Esser P, et al. Technology-assisted physical activity interventions for older people in their home-based environment: scoping review. JMIR Aging 2025;8:e65746 [FREE Full text] [doi: 10.2196/65746] [Medline: 40952779]
50. Lehallier B, Gate D, Schaum N, Nanasi T, Lee SE, Yousef H, et al. Undulating changes in human plasma proteome profiles across the lifespan. Nat Med 2019;25(12):1843-1850 [FREE Full text] [doi: 10.1038/s41591-019-0673-2] [Medline: 31806903]
51. Shen X, Wang C, Zhou X, Zhou W, Hornburg D, Wu S, et al. Nonlinear dynamics of multi-omics profiles during human aging. Nat Aging 2024;4(11):1619-1634. [doi: 10.1038/s43587-024-00692-2] [Medline: 39143318]

XSL•FO
RenderX

52. Vuoksimaa E, Panizzon MS, Chen C, Fiecas M, Eyler LT, Fennema-Notestine C, et al. The genetic association between neocortical volume and general cognitive ability is driven by global surface area rather than thickness. Cereb Cortex 2015;25(8):2127-2137 [FREE Full text] [doi: 10.1093/cercor/bhu018] [Medline: 24554725]
53. Stern Y. Cognitive reserve. Neuropsychologia 2009;47(10):2015-2028 [FREE Full text] [doi: 10.1016/j.neuropsychologia.2009.03.004] [Medline: 19467352]
54. Jack CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, Contributors. NIA-AA research framework: toward a biological definition of Alzheimer's disease. Alzheimers Dement 2018;14(4):535-562 [FREE Full text] [doi: 10.1016/j.jalz.2018.02.018] [Medline: 29653606]
55. Nyberg CK, Fjell AM, Walhovd KB. Level of body fat relates to memory decline and interacts with age in its association with hippocampal and subcortical atrophy. Neurobiol Aging 2020;91:112-124. [doi: 10.1016/j.neurobiolaging.2019.10.005] [Medline: 32224068]
56. Low S, Ng TP, Lim CL, Ang SF, Moh A, Wang J, et al. Higher ratio of extracellular water to total body water was associated with reduced cognitive function in type 2 diabetes. J Diabetes 2021;13(3):222-231. [doi: 10.1111/1753-0407.13104] [Medline: 32786001]
57. Irwin MR, Olmstead R, Carroll JE. Sleep disturbance, sleep duration, and inflammation: a systematic review and meta-analysis of cohort studies and experimental sleep deprivation. Biol Psychiatry 2016;80(1):40-52 [FREE Full text] [doi: 10.1016/j.biopsych.2015.05.014] [Medline: 26140821]
58. Patel SR, Zhu X, Storfer-Isser A, Mehra R, Jenny NS, Tracy R, et al. Sleep duration and biomarkers of inflammation. Sleep 2009;32(2):200-204 [FREE Full text] [doi: 10.1093/sleep/32.2.200] [Medline: 19238807]
59. Wang J, Chen X, Men X, Chen M, Tao J, Lu Z. Chronic insomnia is associated with higher circulating interleukin-8 in patients with atherosclerotic cerebral small vessel disease. Nat Sci Sleep 2020;12:93-99 [FREE Full text] [doi: 10.2147/NSS.S239030] [Medline: 32104118]
60. Ancoli-Israel S, Cole R, Alessi C, Chambers M, Moorcroft W, Pollak CP. The role of actigraphy in the study of sleep and circadian rhythms. Sleep 2003;26(3):342-392. [doi: 10.1093/sleep/26.3.342] [Medline: 12749557]
61. Gabb VG, Blackman J, Morrison H, Li H, Kendrick A, Turner N, et al. Longitudinal remote sleep and cognitive research in older adults with mild cognitive impairment and dementia: prospective feasibility cohort studya. JMIR Aging 2025;8:e72824 [FREE Full text] [doi: 10.2196/72824] [Medline: 40435500]
62. Ding H, Brown S, Paquette DR, Orwig TA, Spartano NL, Lin H. Passive measures of physical activity and cadence as early indicators of cognitive impairment: observational study. J Med Internet Res 2025;27:e72946 [FREE Full text] [doi: 10.2196/72946] [Medline: 40939123]

Abbreviations

AD8: Ascertain Dementia 8
CERAD-NB: Consortium to Establish a Registry for Alzheimer’s Disease Neuropsychological Battery
ECW: extracellular water
ESS: Epworth Sleepiness Scale
GDS: Geriatric Depression Scale
ISDHA: Integrating Systematic Data of Geriatric Medicine to Explore the Solution for Healthy Aging
L5: mean activity counts per minute during the least active 5 consecutive hours
LPA: light-intensity physical activity
M10: mean activity counts per minute during the most active 10 consecutive hours
MET: metabolic equivalents of task
MMSE: Mini-Mental State Examination
MPA: moderate-intensity physical activity
MRI: magnetic resonance imaging
MVPA: moderate to vigorous physical activity
PSQI: Pittsburgh Sleep Quality Index
SB: sedentary behavior
TBW: total body water
TIB: time in bed
TST: total sleep time
VPA: vigorous-intensity physical activity
WASO: wake after sleep onset

XSL•FO
RenderX

Edited by J Wang; submitted 15.07.25; peer-reviewed by Q He, B Nairn, G Grove; comments to author 20.08.25; accepted 18.11.25; published 09.12.25.

Please cite as:
Liu GH, Huang YH, Yuan TC, Chen YP, Hsu JL, Lee SH, Lin CM, Toh CH, Fang JT, Lin SW, Chuang LP, Chen NH
Associations of Wearable-Measured Sleep and Physical Activity With Memory Performance in Older Adults: Cross-Sectional Study With Actigraphy and MRI
JMIR Aging 2025;8:e80584
URL: https://aging.jmir.org/2025/1/e80584
doi:10.2196/80584
PMID:41258815

©Geng-Hao Liu, Yueh-Hsiang Huang, Tzu-Chiao Yuan, Yueh-Peng Chen, Jung-Lung Hsu, Shwu-Hua Lee, Chih-Ming Lin, Cheng-Hong Toh, Ji-Tseng Fang, Shih-Wei Lin, Li-Pang Chuang, Ning-Hung Chen. Originally published in JMIR Aging (https://aging.jmir.org), 09.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Web-Based Assessment of Cognition and Dementia Risk Factors in Over 3000 Norwegian Adults Aged 50 Years and Older: Cross-Sectional PROTECT Norge Study

Ingelin Testad[1,2], PhD; Jon Arild Aakre[1,3], PhD; Martha Therese Gjestsen[1,3], PhD; Clive Ballard[2], MD; Anne Corbett[2], PhD; Dag Aarsland[1,4], MD, PhD; Ellie Pickering[2], MSc; Anastasia Ushakova[1], PhD

[1]Centre for Age-Related Medicine - SESAM, Stavanger University Hospital, Stavanger, Norway

[2]College of Medicine and Health, Faculty of Health and Life Sciences, University of Exeter, Exeter, United Kingdom

[3]Department of Clinical Medicine, University of Bergen, Haukeland Universitetssykehus, Laboratoriebygget, Bergen, Norway

[4]Centre for Healthy Brain Ageing, Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

Corresponding Author:
Jon Arild Aakre, PhD
Centre for Age-Related Medicine - SESAM, Stavanger University Hospital, Stavanger, Norway

Abstract

Background: With the growing number of older adults in the Norwegian population and the associated rapid rise in dementia and cognitive impairment, novel and more efficient methodologies are needed to facilitate research, improve diagnostic triage, and deliver effective brain health interventions in the community. Platform for Research Online to Investigate Genetics and Cognition in Ageing Norge (PROTECT Norge) is a web-based, remote research platform on the aging brain, culturally adapted from the UK PROTECT study, incorporating a Norwegian cohort of adults aged 50 years and older, where participants complete study activities via a dedicated website. Data were collected through study activities, which included self-reported questionnaires and a computerized neuropsychological test battery.

Objective: The study aimed to assess associations between dementia risk factors, including obesity, hypertension, smoking, and hearing loss, and cognition using baseline data from the PROTECT Norge study.

Methods: Data from the PROTECT Norge study were used to assess associations between dementia risk factors and cognition. These associations were assessed using independent-sample *t* tests for each of the cognitive tests, which included paired associate learning, self-ordered search, digit span, and verbal reasoning tasks.

Results: During the first 2 years of data collection, records from 3214 participants were obtained. Associations between established risk factors and cognitive performance were found, with significant detriments to cognitive performance on the computerized neuropsychological test battery. In the PROTECT Norge cohort, consisting of 74.5% females, the mean age was 64.1 (SD 7.7), and 94% of participants provided consent for contact regarding future research.

Conclusions: These data show the associations between dementia risk factors and cognition and attest to the excellent feasibility of the PROTECT Norge cohort, with over 3000 participants included over a period of 2 years and accessibility for people with early cognitive impairment. This, combined with the cohort's willingness to participate in future studies (94%), positions PROTECT Norge as a tremendous opportunity for cost-efficient, large-scale brain health research and potentially for clinical digital cognitive health programs.

(JMIR Aging 2025;8:e69867) doi:10.2196/69867

KEYWORDS

PROTECT Norge; cognitive function; dementia risk factors; online cohort study; brain health; epidemiology

Introduction

With an aging population, developing new approaches to improve brain health and prevent cognitive decline is essential to empower the community and mitigate against a potential "dementia epidemic." For example, there are already more than 50 million people worldwide living with dementia, and this is projected to rise rapidly over the next 30 years [1]. Aging is the single biggest risk factor for dementia [2]. In addition, genetic predispositions contribute significantly to dementia risk, both of which are factors with limited potential for mitigation. However, within the domains of lifestyle, mental health, and physical health, there exists a spectrum of potential modifiable risk factors for dementia [3], which are thought to account for 40% of the total dementia risk. With no cure for dementia,

XSL•FO
RenderX

addressing these potentially modifiable risk factors through proactive management and preventative interventions is proposed as a key strategy going forward. This approach aims to maintain and improve cognitive health in aging and slow the onset or progression of dementia, to address the expected increase of dementia cases in the population [3,4]. As populations worldwide grow older, the urgency of dementia prevention becomes increasingly evident. In Norway, this trend is reflected in a projected 30.1% increase in those aged 50 years and older from 2024 to 2050 [5], with rural municipalities seeing the greatest impact, where individuals aged 70 years and older could make up a third of the population [6]. Digital technologies show promise in supporting health promotion and disease prevention, with adults aged 50 years and older increasingly integrating them into their daily lives for this purpose [7].

In the United Kingdom, a web-based aging cohort was set up in 2014 to address these challenges. The Platform for Research Online to investigate Cognition and Genetics in Ageing- United Kingdom (PROTECT-UK [8-10]) study aimed to collate large-scale health and cognitive data from older adults to provide novel insights into cognitive trajectory in aging and the risk factors associated with decline. PROTECT-UK (University of Exeter) was built on emerging strengths in digital technologies and the rapid increase in engagement with the internet, computers, and smartphone devices by people in mid and later life. PROTECT-UK demonstrated the enormous potential for this approach to longitudinal research, enabling rapid recruitment of 20,000 participants within the first year, with strong year-on-year retention, for a fraction of the cost of traditional in-person cohort methodologies. Norway was the first country outside the United Kingdom to adopt the PROTECT model, contributing to its international expansion, with Canada later implementing a similar approach [11]. By providing large-scale, standardized cognitive data across multiple countries, PROTECT enables cross-national comparisons and enhances global understanding of dementia risk factors.

Platform for Research Online to Investigate Genetics and Cognition in Ageing Norge (PROTECT Norge) builds upon the PROTECT-UK infrastructure but is tailored to the Norwegian context. PROTECT Norge was launched in September 2020 and has already collected data on more than 5000 participants [12]. The PROTECT infrastructure has enabled large-scale collection of longitudinal data on a range of lifestyle, physical health, mental health, and well-being, in addition to robust, well-validated computerized neuropsychology data in the Norwegian population, providing an accessible, engaging, and cost-efficient approach to cognitive health research using a web-based, remote methodology, which combines computerized cognitive tests and health questionnaires on a dedicated web-based platform with remote genetic testing. The platform provides the opportunity to engage people at scale with epidemiological studies, genomic and biomarker research, and clinical trials. Upon registration, potential participants are given the option to voluntarily consent to being recontacted for future studies, which will enable us to capitalize on their participation in clinical trials relevant to cognition, health, and well-being. For example, this will be an important vehicle to increase the proportion of people who are offered participation in clinical studies, contributing to reach the Government's action plan and the national target of including 5% of the patient population in Norwegian hospitals by 2025.

The need for such digital initiatives like PROTECT Norge is underscored by the introduction of antiamyloid antibodies for dementia, whose effectiveness depends on early diagnosis. However, this remains a major challenge, as only 1 in 3 dementia cases in the Norwegian community are diagnosed, and diagnosis often occurs in late stages of the condition when irreversible cognitive decline has already led to pronounced functional impairments in daily life and an increased need for care [13].

PROTECT Norge is the only study of this scale incorporating remote data collection, including longitudinal cognitive testing in Norway. The platform is uniquely situated to optimize approaches for the early identification of people with progressive cognitive decline and provide a potential opportunity to evolve as a primary care and community health delivery pathway and resource. While PROTECT Norge is tailored to the Norwegian context, its findings contribute to the wider understanding of cognitive aging and dementia risk factors, supporting ongoing research efforts in dementia prevention and early detection.

This study presents information regarding the baseline characteristics, neuropsychological performance, and risk factors for dementia in the PROTECT Norge cohort with the objective of assessing associations between previously identified dementia risk factors and cognition.

Methods

Design

This is a cross-sectional analysis of data collected through PROTECT Norge, a web-based observational cohort study focusing on the aging brain of individuals aged 50 years and older living in the community in Norway. All data are collected remotely via the PROTECT Norge study website. The reporting of the study follows the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) [14] guidelines and the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) [15] (the CHERRIES checklist is provided in Checklist 1).

Cross-Cultural Adaptation of PROTECT Norge

All questionnaires without an existing Norwegian translation, along with cognitive tests, underwent a thorough translation process [16]. Two native Norwegian speakers independently translated each measure, 1 familiar with the content and 1 unfamiliar. They then collaborated to create a consensus version. This was back-translated by 2 bilingual translators, preferably native English speakers. Where this was not feasible, Norwegians fluent in English with relevant academic or professional experience were used. An expert committee, including translators from all stages and PROTECT Norge working group members, reviewed all versions and finalized the Norwegian adaptations. For medical and mental health-related questionnaires, additional domain experts and professional translators contributed to quality assurance.

XSL•FO
RenderX

The PROTECT Norge platform underwent multiple rounds of user acceptance testing (UAT). Initial alpha and beta testing of an unscripted manner was conducted by PROTECT-UK and PROTECT Norge personnel. A scripted UAT was then carried out with 30 participants, including end users, user representatives, researchers, and PROTECT Norge personnel. This allowed for simultaneous platform testing and translation validation, ensuring clarity and usability [17].

Participants

Participants fulfill the inclusion criteria for PROTECT Norge, with all participants aged 50 years or older, residing in Norway, without a diagnosis of dementia, and with a good understanding of the Norwegian language. Participants must also have access to a computer or tablet with an internet connection. The vast majority of people over the age of 50 years in Norway are eligible, thus providing an opportunity for the whole community within this age group to participate in research. Participation is entirely remote and can be completed at home.

Recruitment

Participants were recruited to PROTECT Norge through publicity in national and regional broadcasting channels, newspaper coverage, placement of written materials in health care facilities, and oral presentations. However, social media, particularly Facebook, served as a primary recruitment channel through both nontargeted posts and targeted advertisements, the latter applying demographic filters based on the inclusion criteria (age 50 years and older and residing in Norway). The participants register, confirm eligibility, and provide consent through an ethically approved, validated web-based procedure, which includes voluntary consents for contact for future research and consent for providing saliva for genetic sampling. Following registration, participants provide demographic information and then receive email reminders for assessments and study activities. Participants complete an annual assessment of validated questionnaires and assessments on the PROTECT Norge platform.

Cognitive Assessments

Participants complete an annual cognitive assessment using a validated computerized cognitive assessment system. This analysis presents data from 4 widely used and well-validated cognitive tests that were in use on the platform from 2020 to 2022, including tests of spatial working memory (Paired Associate Learning [18], Self-Ordered Search [19]), numerical working memory (Digit Span [20]), and executive function (Verbal Reasoning [21]). The Trail-making B test was also included to enable benchmarking against well-established cognitive norms [22]. Mild cognitive impairment (MCI) was identified based on established thresholds [23].

Although not presented in this paper, it is important to note that in 2022, the cognitive assessment protocol was updated to the FLAME (Factors of Longitudinal Attention, Memory, and Executive Function; University of Exeter) cognitive test battery which includes the 4 original tests in addition to 3 tests of reaction time and attention, and 1 of attention and episodic memory. The FLAME battery is well validated [23], showing good discrimination between different levels of cognitive impairment, good sensitivity to change, and good concurrent validity with functional change. Importantly, the battery is validated for self-test in their own home [10].

Health Questionnaire Assessments

Participants complete annual questionnaire assessments to capture key risk and cognitive association factors. Participants provide information regarding current and past medical conditions through a tick-box selection of key conditions, including hypertension and hearing loss, free-text capture of prescribed medications, and height and weight to enable BMI calculation. Participants also complete a comprehensive questionnaire set to capture lifestyle and modifiable risk factors, including questions on smoking status (current-, ever-, and nonsmoker). Where participants have voluntarily nominated an informant, the informant is invited to complete adapted versions of specific questionnaires pertaining to the participant. Although not evaluated as part of the current paper, the other assessments undertaken are summarized in Textbox 1 for information.

XSL•FO
RenderX

Textbox 1. Other assessments completed in PROTECT Norge (Platform for Research Online to Investigate Genetics and Cognition in Ageing Norge).

The text box lists other study assessments undertaken by participants in PROTECT Norge:

- Health Questionnaire: captures Medical conditions and prescribed medications, sensory impairments, pain, and history of traumatic brain injury
- Family history of neurodegenerative disease questionnaire
- The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): self and informant questionnaire
- Mild Behavior Impairment (MBI) scale: measures neuropsychiatric symptoms; available as self and informant questionnaire
- Instrumental Activities of Daily Living questionnaire
- Lifestyle questionnaire: includes items capturing use of technology, cognitive training, alcohol use, caffeine intake, dietary supplement use, use of cognitively stimulating activities, and the number of languages spoken, physical activity, including CHAMPS (Community Healthy Activities Model Program for Seniors), physical activity questionnaire for older adults. This contains adaptive questioning, whereby certain items are conditionally displayed based on previous responses
- Composite mental health questionnaire: Patient Health Questionnaire (PHQ-9) and General Anxiety Disorder (GAD-7) scales as assessments of depression and anxiety, history of mental health disorder, items from the World Health Organization (WHO) composite international diagnostic interview psychosis and manic modules, addictions, including Alcohol Use Disorder Identification Test (AUDIT), questions relating to self-harm, suicidality, well-being, and experiences in adult life and childhood. Items exploring neurodivergent symptoms from childhood. This questionnaire is voluntary in its entirety, and participants may choose “prefer not to answer” for any item. It also uses adaptive questioning
- Sleep: St. Mary’s Sleep Questionnaire, Insomnia Severity Index, Bergen Insomnia Scale, and Epworth Sleepiness Scale
- Awareness of age-related change-10 SF: assesses perceived gains and losses with aging
- University of California, Los Angeles Three-Item Loneliness Scale (UCLA-3)
- ASK (Aging, Social Dynamics, and Knowledge) questionnaire: captures information about living situations, social interactions, technology use, travel habits, future housing preferences, and views on future care models

For all questionnaires, respondents can review and modify their responses at any time before final submission.

Patient and Public Involvement

User involvement is paramount in all health research and aligns with the mission of the Ministry of Health and Care Services to regional health authorities and has been the core value in all activities at the Centre for Age-Related Medicine - SESAM since its establishment in 2010. PROTECT Norge is supported by 2 designated user representatives who have been actively engaged since the start of PROTECT Norge. They are part of the steering group, have been represented in seminars, and have been actively involved in the translation of material from PROTECT-UK.

Statistical Analysis

Categorical variables were described using count and percentage. Continuous variables were described using mean (SD) if symmetrically distributed and median with IQR if otherwise. For each test, the average value of all scores over the number of times it was completed (from 1 to 3 at each testing time point) was used as a final score.

For each cognitive test score and each potential medical risk factor (Hypertension–present or absent, smoking ever, or nonsmoker), BMI (<30 v>30), hearing loss—present or absent—identified as risk factors from the Lancet commission meta-analysis [3], we performed a 2-tailed independent sample *t* test for each of the cognitive tests (df=3212). Two-tailed *P* values <.05 were considered significant; no adjustments for confounders and no correction for multiple testing were done. All statistical analyses were undertaken using SPSS (version 29; IBM Corp).

Ethical Considerations

This study uses data from the PROTECT Norge Study [12], launched in September 2020 and approved by the Regional Committee for Medical and Health Research Ethics, West (2019/478), and the study has been approved by the Stavanger University hospital. Informed consent was obtained through a validated web-based procedure, and data are stored in compliance with the General Data Protection Regulation (GDPR) and Stavanger University Hospital’s data security policy [24]. To protect privacy, data were pseudonymized before analysis. No compensation was provided for participants, except for gaining access to brain training games on the PROTECT Norge platform.

Results

Cohort Characteristics

During the first 2 years of the data collection period, 3214 participants completed the baseline assessment. The average age of the cohort was 64 years, of whom 74.5% (n=2393) were female and nearly 80% (n=2503) had a higher education degree. The majority (n=2289, 71.7%) of the participants were either widowed, divorced, or single, and 50% reported that they were working either part- or full-time. Of all the participants (n=3214), 58 had MCI. The full cohort characteristics are shown in Table 1.

The majority of participants (74.4%) reported hearing about PROTECT Norge through social media. Additional sources included friends and acquaintances (10.3%), traditional media (TV, radio, and newspapers; 4.4%), other study participants (4.2%), other sources (5.8%), and courses or conferences (0.9%).

XSL•FO
RenderX

Consent for contact for future research was provided by 94% of the participants in PROTECT Norge.

Table . Characteristics of PROTECT Norge[a] cohort participants.

Characteristics (n=3214)[b]	Values
Age (years), mean (SD), min/max	64.1 (7.7), 50.0/89.2
Female, n (%)	2393 (74.5)
Higher education, n (%)	2503 (77.9)
Marital status, n (%)	
Married, civil partnership, or cohabiting	2289 (71.7)
Widowed, separated, divorced, or single	902 (28.3)
Working (full-time or part-time), n (%)	1608 (50.2)
High blood pressure, n (%)	656 (38.3)
Mild cognitive impairment, n (%)	58 (1.8)
Hearing loss, n (%)	818 (29.6)
Smoking, n (%)	
Current smoker	153 (5.6)
Stopped smoking	1465 (53.2)
BMI (reference: normal), n (%)	
Underweight	24 (0.9)
Overweight	1057 (38.3)
Obese	374 (13.6)
Very obese	26 (0.9)

[a]PROTECT Norge: Platform for Research Online to Investigate Genetics and Cognition in Ageing Norge.

[b]This table presents the demographic and health-related characteristics of the participants in the PROTECT Norge cohort study.

Cognitive Profile of the Cohort

The mean scores for each of the neuropsychological tests undertaken are described and compared between participants with and without MCI, showing that performance was significantly more impaired on the 4 PROTECT tasks in people with MCI compared with those with normal cognition (Table 2).

Table . Neuropsychological test scores and comparison of cognitive performance between people with MCI[a] and healthy cognition.

Test[b]	Mean (SD)	MCI vs normal cognition	
		t test (*df*)	*P* value
Digit Span Memory Test	6.3 (2.5)	6.7 (3212)	<.001
Paired Associate Learning	3.6 (1.0)	3.3 (3212)	<.001
Verbal Reasoning task	24.5 (10.1)	8.8 (3212)	<.001
Self-Ordered Search task	33.9 (14.4)	6.3 (3212)	<.001

[a]MCI: mild cognitive impairment.

[b]This table presents the mean scores (SD) on neuropsychological tests among individuals with mild cognitive impairment (MCI) compared to those with normal cognition. The statistical comparison between the MCI group and the normal cognition group is shown with corresponding *t* values (df=3212) and *P* values.

Association of Medical Risk Factors With Cognitive Performance

A consistent pattern of mainly significant associations was seen between the Lancet commission risk factors for dementia examined (hypertension, smoking, hearing loss, and BMI>30) [3] and the neuropsychological test battery scores. The full results are shown in Table 3.

XSL•FO
RenderX

Table . Association between medical risk factors for dementia and cognitive test scores in PROTECT Norge[a].

Medical risk factors[b]	Digit span		PAL[c]		Verbal reasoning		Self-ordered search	
	t test (*df*)	*P* value	*t* test (*df*)	*P* value	*t* test (*df*)	*P* value	*t* test (*df*)	*P* value
Hypertension	2.3 (3212)	.02	2.7 (3212)	.006	2.4 (3212)	.02	1.5 (3212)	.13
Hearing loss	3.6 (3212)	<.001	3.5 (3212)	<.001	2.9 (3212)	.004	1.7 (3212)	.09
Smoking	3.0 (3212)	.003	2.5 (3212)	.007	2.8 (3212)	.005	2.7 (3212)	.007
BMI >30	1.9 (3212)	.06	1.5 (3212)	.07	4.3 (3212)	<.001	3.1 (3212)	.002

[a]PROTECT Norge: Platform for Research Online to Investigate Genetics and Cognition in Ageing Norge.

[b]This table displays the association between various medical risk factors for dementia and cognitive test scores among participants in the PROTECT Norge study. Statistical significance is indicated by *t* values, including degrees of freedom and corresponding *P* values (P).

[c]PAL: paired associate learning.

Discussion

Principal Findings

This study reports the baseline characteristics of the first 3214 participants in the PROTECT Norge digital cohort study of people aged 50 years and older. We found a consistent pattern of mainly significant associations between the Lancet Commission risk factors for dementia [3], including hypertension, smoking, hearing loss, and BMI>30, and lower cognitive performance on the neuropsychological test battery scores.

Comparison to Previous Work

Established risk factors for dementia (BMI, smoking, hypertension, and hearing loss) were all significantly associated with lower cognitive performance on the computerized neuropsychological tests, which is important for several reasons. First, it acts as a validation of the PROTECT Norge cohort, demonstrating the same pattern of risk associations that have been identified in other international studies [3]. Second, the work builds on this by demonstrating that in addition to the longer-term risks for dementia identified in other studies, these risk factors are associated with subtle but measurable cognitive impairments in individuals who are cognitively healthy but with an increased future risk of dementia. This also highlights the opportunities to improve cognitive health through intervention studies in the Norwegian population. Hearing loss is a relatively newly identified risk factor for dementia, and it is therefore of particular note that a strong association was seen between hearing loss and cognitive test performance. Emerging work also suggests that the risk can be mitigated by wearing hearing aids [25,26], with important potential public health implications.

Strengths and Limitations

First, a strength of the PROTECT Norge study is its fully digitalized study design, which has enabled the inclusion of over 3000 Norwegian individuals over a period of 2 years, during which high-quality data, including computerized neuropsychology, validated questionnaires collecting health, mental health, and lifestyle data were successfully collected. In addition, 94% of the cohort has consented to be contacted for future research. Second, the computerized neuropsychology cognitive test battery is objectively measured, demonstrates good discrimination between different levels of cognitive impairment, and has been validated for self-testing in a home setting [10]. However, other health assessments used in this study rely on self-reported data, which may introduce potential biases such as recall errors and misestimation. Third, it should be acknowledged that the findings reported in this study are based on cross-sectional data, meaning causality cannot be inferred.

Fourth, the cohort is predominantly Caucasian, female, and well-educated, in line with other digital health studies, and the self-selection recruitment strategy and a heavy reliance on Facebook for outreach have likely contributed to this imbalance, as women in this age group are more engaged on this platform [24,27]. Finally, although it is encouraging that participants with MCI successfully completed study assessments, only 58 of 3214 participants had MCI, well below population estimates and the proportion seen in PROTECT UK.

Future Direction

In 2030, for the first time in Norway, there will be more people over 65 years than 19 years and younger [6] and the number of people with dementia will increase from 101,118 in 2020 to almost 235,000 in 2050 [28], along with a reduction in the ratio of people working to those who are retired. These numbers are even more dramatic in rural districts [6]. Better brain health and improved initiatives to diagnose early cognitive impairment and prevent dementia will be imperative. Lifestyle and cognitive training interventions have been found to improve brain health and prevent cognitive decline [29,30], and the PROTECT infrastructure is well placed to deliver large-scale interventions to the Norwegian population. Digitalized research and health care platforms will be an essential part of improving the brain health of older people in our society.

This study demonstrates the feasibility of the PROTECT model, highlighting the tremendous opportunity for this approach to contribute significantly to large-scale, cost-effective brain health research and service delivery as we move forward. In addition, despite the current demographic limitations seen in PROTECT Norge, this study still provides valuable insights into how these risk factors influence cognition within a large cohort, contributing to the understanding of cognitive health at a population level. However, future efforts are warranted, specifically to improve demographic diversity in the cohort and increase the representation of individuals with early cognitive

XSL•FO
RenderX

impairment. Moving this forward could be done by specifically targeting underrepresented demographic groups on social media platforms, whilst simultaneously expanding outreach beyond social media to relevant interest groups or organizations and collaboration with primary care providers.

Conclusion

In summary, our study using PROTECT Norge data reveals significant associations between established dementia risk factors, including obesity, hypertension, smoking, and hearing loss, and lower cognitive performance, even among individuals who are cognitively healthy but at increased future risk of dementia. These findings support the validity of the PROTECT model and highlight its potential to detect early cognitive changes in at-risk populations. The successful collection of large-scale, remotely administered data further demonstrates the feasibility of this digital approach. PROTECT Norge offers a unique and cost-effective opportunity to scale brain health research and interventions across Norway, particularly by targeting individuals with modifiable risk factors.

However, to improve cohort representativeness, more targeted recruitment strategies are needed to reduce selection bias and enhance demographic diversity. Future efforts should focus on expanding outreach channels and optimizing participant engagement to improve the inclusion of underrepresented demographic groups, including individuals with MCI.

Acknowledgments

The PROTECT Norge (Platform for Research Online to Investigate Genetics and Cognition in Ageing Norge) study received funding from the National Collaboration Group for Health Research in the Specialist Health Care Services (Nasjonal samarbeidsgruppe for helseforskning i spesialisthelsetjenesten), Stavanger Health Trust and from the Western Norway Regional Health Authority. We extend our gratitude to Adam Hampshire for providing the cognitive tests for the study. We also thank the study participants from PROTECT Norge, as well as the patient and public involvement representatives, Pamela Cranner and Tor Cranner, organized through WiseAge.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to data governance restrictions, but may be accessed on reasonable request. Applications for access to the PROTECT Norge data must be submitted, and this process is initiated by contacting the corresponding author.

Conflicts of Interest

None declared.

Checklist 1
Checklist for Reporting Results of Internet E-Surveys.
[PDF File, 209 KB - aging_v8i1e69867_app1.pdf]

References

1. Patterson C. World alzheimer report 2018: the state of the art of dementia research: new frontiers. Alzheimer's Disease International. 2018 URL: https://www.alzint.org/resource/world-alzheimer-report-2018/ [accessed 2025-06-10]
2. van der Flier WM, Scheltens P. Epidemiology and risk factors of dementia. J Neurol Neurosurg Psychiatry 2005 Dec;76 Suppl 5(Suppl 5):v2-v7. [doi: 10.1136/jnnp.2005.082867] [Medline: 16291918]
3. Livingston G, Huntley J, Sommerlad A, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet 2020 Aug 8;396(10248):413-446. [doi: 10.1016/S0140-6736(20)30367-6] [Medline: 32738937]
4. Risk reduction of cognitive decline and dementia: WHO guidelines. World Health Organization. 2019. URL: https://www.who.int/publications/i/item/9789241550543 [accessed 2025-08-08]
5. 14282: Population projections 1 January, by sex, age, immigration category and country background, in 15 alternatives 2024 – 2100. Statistics Norway. 2024. URL: https://www.ssb.no/en/statbank/table/14282 [accessed 2024-06-13]
6. Leknes S, Løkken SA. Municipal Population Projections 2024: Statistics Norway; 2024. URL: https://www.ssb.no/befolkning/befolkningsframskrivinger/artikler/befolkningsframskrivinger-for-kommunene-2024 [accessed 2025-06-10]
7. De Santis KK, Mergenthal L, Christianson L, Busskamp A, Vonstein C, Zeeb H. Digital technologies for health promotion and disease prevention in older people: scoping review. J Med Internet Res 2023 Mar 23;25:e43542. [doi: 10.2196/43542] [Medline: 36951896]
8. University of Exeter. PROTECT UK. 2014. URL: https://www.protectstudy.org.uk/ [accessed 2025-06-10]
9. Corbett A, Williams G, Creese B, et al. Cognitive decline in older adults in the UK during and after the COVID-19 pandemic: a longitudinal analysis of PROTECT study data. Lancet Healthy Longev 2023 Nov;4(11):e591-e599. [doi: 10.1016/S2666-7568(23)00187-3] [Medline: 37924840]

XSL•FO
RenderX

10. Huntley J, Corbett A, Wesnes K, et al. Online assessment of risk factors for dementia and cognitive function in healthy adults. Int J Geriatr Psychiatry 2018 Feb;33(2):e286-e293. [doi: 10.1002/gps.4790] [Medline: 28960500]
11. Ismail Z, Guan DX, Vellone D, et al. The Canadian platform for research online to investigate health, quality of life, cognition, behaviour, function, and caregiving in aging (CAN-PROTECT): Study protocol, platform description, and preliminary analyses. Aging and Health Research 2024 Dec;4(4):100207. [doi: 10.1016/j.ahr.2024.100207]
12. Centre for age-related medicine - SESAM, Stavanger University Hospital. PROTECT Norge. 2020. URL: https://www.protect-norge.no/ [accessed 2025-06-10]
13. Gjøra L, Strand BH, Bergh S, et al. Prevalence and determinants of diagnosed dementia: a registry linkage study linking diagnosis of dementia in the population-based HUNT study to registry diagnosis of dementia in primary care and hospitals in Norway. J Alzheimers Dis 2024;99(1):363-375. [doi: 10.3233/JAD-240037] [Medline: 38701153]
14. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008 Apr;61(4):344-349. [doi: 10.1016/j.jclinepi.2007.11.008] [Medline: 18313558]
15. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
16. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976) 2000 Dec 15;25(24):3186-3191. [doi: 10.1097/00007632-200012150-00014] [Medline: 11124735]
17. Gordon S, Crager J, Howry C, et al. Best practice recommendations: user acceptance testing for systems designed to collect clinical outcome assessment data electronically. Ther Innov Regul Sci 2022 May;56(3):442-453. [doi: 10.1007/s43441-021-00363-z] [Medline: 35233726]
18. Barnett JH, Blackwell AD, Sahakian BJ, Robbins TW. The Paired Associates Learning (PAL) Test: 30 Years of CANTAB Translational Neuroscience from Laboratory to Bedside in Dementia Research. In: Murray RM, Lewis DA, Weinberger DR, editors. Curr Top Behav Neurosci 2016;28:449-474. [doi: 10.1007/7854_2015_5001] [Medline: 27646012]
19. Gillett R. Assessment of working memory performance in self-ordered selection tests. Cortex 2007 Nov;43(8):1047-1056. [doi: 10.1016/s0010-9452(08)70702-0] [Medline: 18044665]
20. Schroeder RW, Twumasi-Ankrah P, Baade LE, Marshall PS. Reliable digit span: a systematic review and cross-validation study. Assessment 2012 Mar;19(1):21-30. [doi: 10.1177/1073191111428764] [Medline: 22156721]
21. Baddeley AD. A 3 min reasoning test based on grammatical transformation. Psychon Sci 1968 Oct;10(10):341-342. [doi: 10.3758/BF03331551] [Medline: 19717586]
22. Tombaugh TN. Trail making test A and B: normative data stratified by age and education. Arch Clin Neuropsychol 2004 Mar;19(2):203-214. [doi: 10.1016/S0887-6177(03)00039-8] [Medline: 15010086]
23. Brooker H, Williams G, Hampshire A, et al. FLAME: a computerized neuropsychological composite for trials in early dementia. Alzheimers Dement (Amst) 2020;12(1):e12098. [doi: 10.1002/dad2.12098] [Medline: 33088895]
24. Privacy Statement for Helse Stavanger HF. Helse Stavanger. 2024. URL: https://www.helse-stavanger.no/om-oss/personvernerklering-for-helse-stavanger-hf [accessed 2024-06-13]
25. Cantuaria ML, Pedersen ER, Waldorff FB, et al. Hearing loss, hearing aid use, and risk of dementia in older adults. JAMA Otolaryngol Head Neck Surg 2024 Feb 1;150(2):157-164. [doi: 10.1001/jamaoto.2023.3509] [Medline: 38175662]
26. Lin FR, Pike JR, Albert MS, et al. Hearing intervention versus health education control to reduce cognitive decline in older adults with hearing loss in the USA (ACHIEVE): a multicentre, randomised controlled trial. Lancet 2023 Sep 2;402(10404):786-797. [doi: 10.1016/S0140-6736(23)01406-X] [Medline: 37478886]
27. Chidiac M, Ross C, Marston HR, Freeman S. Age and gender perspectives on social media and technology practices during the COVID-19 pandemic. Int J Environ Res Public Health 2022 Oct 27;19(21):36360853. [doi: 10.3390/ijerph192113969] [Medline: 36360853]
28. Gjøra L, Kjelvik G, Strand BH, Kvello-Alme M, Selbæk G. Forekomst Av Demens i Norge: Aldring og helse; 2020.
29. Corbett A, Williams G, Creese B, et al. Impact of short-term computerized cognitive training on cognition in older adults with and without genetic risk of Alzheimer's disease: outcomes from the START randomized controlled trial. J Am Med Dir Assoc 2024 May;25(5):860-865. [doi: 10.1016/j.jamda.2024.03.008] [Medline: 38642588]
30. Ngandu T, Lehtisalo J, Solomon A, et al. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. Lancet 2015 Jun 6;385(9984):2255-2263. [doi: 10.1016/S0140-6736(15)60461-5] [Medline: 25771249]

Abbreviations

CHERRIES: Checklist for Reporting Results of Internet E-Surveys
FLAME: Factors of Longitudinal Attention, Memory, and Executive Function
GDPR: General Data Protection Regulation
MCI: mild cognitive impairment
PROTECT Norge: Platform for Research Online to Investigate Genetics and Cognition in Ageing Norge

XSL•FO
RenderX

PROTECT-UK: Platform for Research Online to Investigate Cognition and Genetics in Ageing- United Kingdom
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
UAT: user acceptance testing

Edited by Y Jiang; submitted 09.01.25; peer-reviewed by A Jamal, S Ogundepo; revised version received 12.06.25; accepted 12.06.25; published 25.08.25.

Please cite as:
Testad I, Aakre JA, Gjestsen MT, Ballard C, Corbett A, Aarsland D, Pickering E, Ushakova A
Web-Based Assessment of Cognition and Dementia Risk Factors in Over 3000 Norwegian Adults Aged 50 Years and Older: Cross-Sectional PROTECT Norge Study
JMIR Aging 2025;8:e69867
URL: https://aging.jmir.org/2025/1/e69867
doi:10.2196/69867

©Ingelin Testad, Jon Arild Aakre, Martha Therese Gjestsen, Clive Ballard, Anne Corbett, Dag Aarsland, Ellie Pickering, Anastasia Ushakova. Originally published in JMIR Aging (https://aging.jmir.org), 25.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Network Analysis of Key Instrumental Activities of Daily Living and Cognitive Domains for Targeted Intervention in US Older Adults Without Dementia: Cross-Sectional Study

Jiaying Li[1,2], PhD; Rendong He[1,3], MSN; Erh-Chi Hsu[1], MPH; Junxin Li[1], PhD

[1]School of Nursing, Johns Hopkins University, 525 N Wolfe St, Baltimore, MD, United States

[2]School of Nursing, Li Ka Shing Faculty of Medicine, University of Hong Kong, China (Hong Kong)

[3]School of Nursing, Jilin University, Changchun, China

Corresponding Author:
Junxin Li, PhD
School of Nursing, Johns Hopkins University, 525 N Wolfe St, Baltimore, MD, United States

Abstract

Background: Cognitive impairment in older adults reduces independence and raises health care costs but can be mitigated through stimulating activities. Based on network theory, intricate relationships within and between clusters of instrumental activities of daily living (IADLs) and cognitive domains suggest the existence of central IADLs and cognitive domains, as well as bridge IADLs. Modifying these can significantly enhance daily living activities and cognitive functions holistically.

Objective: This study aims to identify central IADLs (key activities within the IADL network), central cognitive domains (key domains within the cognitive network), and bridge IADLs (linking IADL and cognitive networks). These insights will inform targeted interventions to effectively improve IADL and cognitive well-being in older adults.

Methods: A cross-sectional analysis of adults aged 65 years and older in the United States focused on 5 IADLs and 6 cognitive domains from the National Health and Aging Trends Study (NHATS). Network analysis identified central and bridge variables. Nonparametric and case-dropping bootstrap methods checked network stability. Network comparison tests assessed sex differences with Benjamini-Hochberg adjustments.

Results: Of the 2239 participants, 56.4% were female (n=976). We computed and tested 3 networks: IADL, cognition, and bridge-with correlation stability coefficients of 0.67, 0.75, and 0.44, respectively (all>0.25). Meal preparation was identified as the central IADL, with a centrality index of 3.87, which was significantly higher than that of other IADLs (all P<.05). Visual attention emerged as the central cognition domain, with a centrality index of 0.86, which was significantly higher than that of other cognition domains (all P<.05). Shopping was determined to be the bridge IADL, with a centrality index of 0.41, which was significantly higher than that of other IADLs (all P<.05). Notably, gender differences emerged in the IADL network, with stronger associations between laundry and meal preparation in females (1.69 vs males: 0.74; P=.001) and higher centrality in meal preparation among females (difference=1.99; P=.007).

Conclusions: While broad enhancements in all IADL and cognitive domains are beneficial, targeting meal preparation, visual attention, and shopping may leverage their within-network influence to yield a more pronounced improvement in holistic IADL, holistic cognition, and holistic cognition function through IADL interventions among older adults. Notably, meal preparation interventions may be less effective in males, requiring tailored approaches.

(JMIR Aging 2025;8:e67632) doi:10.2196/67632

KEYWORDS

cognition function; older adults; intervention targets; elder; elderly; cognitive impairment; stimulating activity; instrumental activities of daily living; IADL; daily living activity; cognitive domain; non-demented; cognitive network; holistic cognition; holistic cognition function; network comparison; central variables; bridge variables; network analysis

Introduction

Cognitive function encompasses mental processes such as acquiring knowledge, manipulating information, and reasoning, including perception, memory, learning, attention, decision-making, and language abilities [1]. Globally, cognitive impairment in older adults has a prevalence ranging from 5.1% to 41%, with incidence rates around 53.97 per 1000 person-years [2]. This impairment not only predicts older adults' future incidence of dementia, but also significantly reduces functional independence and quality of life [2]. Beyond the individual level, economically, cognitive impairment incurs 44% higher direct medical costs and significantly increases the need for

XSL•FO
RenderX

informal care [3]. Given these consequences, early detection and proactive management are essential to mitigate impacts and prevent progression to more severe conditions and greater health care burdens.

Instrumental activities of daily living (IADLs) encompass complex tasks necessary for independent living, such as cooking, cleaning, transportation, laundry, and financial management [4]. These activities assessment sometime adjusted by country or age will include driving or medication management [5,6] and demand cognitive skills like planning, memory, and problem-solving, linking their performance closely to cognitive well-being [4]. In performing these tasks, individuals engage in practical cognitive training that builds cognitive reserve and supports overall brain health. Furthermore, regularly performing IADLs provides ongoing cognitive training that builds "cognitive reserve," a concept suggesting that mentally stimulating tasks strengthen neural connections, thereby enhancing brain resilience [7]. Although other pursuits, such as employment or volunteering, also require high-level cognition, IADLs foster daily independence in older adults, and their everyday, repetitive occurrence makes them an ideal approach for continuous cognitive training. However, current interventions usually target specific IADL domains, such as shopping or meal preparation, highlighting the challenges and resource demands of broad IADL interventions [8].

A significant research gap exists in identifying the most impactful IADL component that can efficiently enhance global cognitive function, particularly crucial in resource-limited settings. This need is supported by complex interactions between IADLs and cognition [9-12], correlations within individual IADL domains (eg, between laundry and meal preparation due to similar cognitive and physical demands) [13], and relationships within cognition domains (between psychomotor function and visual attention stemming from their joint role in tasks requiring quick coordination and responses) [14]. Notably, the cognitive demands of specific IADLs vary; for instance, shopping necessitates skills in navigation, selection, and financial transactions, whereas meal preparation involves planning, execution, and presentation. Such distinctions suggest the existence of certain IADLs that are more closely linked to multiple cognitive domains. Traditional analytic methods, which often isolate relationships or assume predictor independence, may overlook the nuanced, simultaneous interactions among nodes. In contrast, network analysis captures these complex dynamics by representing variables as nodes interconnected by edges [15]. This approach not only elucidates direct interactions but also reveals broader network structures, thereby identifying "central" nodes-those exerting significant within-group influence-and "bridge" nodes that connect disparate networks [15]. Specifically, within our framework, a central IADL is defined as the activity with the highest connectivity within the IADL network, while central cognition refers to the cognitive domain with the most extensive links. A bridge IADL, by linking the IADL and cognitive networks, may have an outsized impact on overall function when its performance changes. Such insights can inform target interventions and strategic resource allocation aimed at enhancing both daily living activities and cognitive function [15]. Furthermore, gender-specific differences in cognitive decline and IADL performance further complicate this landscape. Women generally demonstrate superior executive function and memory; however, their executive function appears to decline more rapidly than that of men, while memory trajectories remain similar between sexes. In contrast, difficulties in financial management and medication adherence are more predictive of dementia in men, whereas transportation challenges serve as stronger indicators in women [16,17]. Consequently, a nuanced exploration of these differences is critical for developing targeted intervention strategies.

This study aimed to identify the central IADL and cognition domains within their respective networks, pinpoint the bridge IADL most substantially linked to overall cognitive function, and examine sex differences in these variables. We hypothesized the existence of central, bridge variables and sex-based differences affecting them. The findings are expected to reveal network dynamics, pinpoint key intervention targets for effectively enhancing holistic IADL and cognitive functions in the elderly, and indicate the necessity of sex-specific interventions.

Methods

Study Design and Data Source

This cross-sectional analysis used waves 11 and 12 (2021 - 2022) of the National Health and Aging Trends Study (NHATS), a nationally representative longitudinal database of Medicare beneficiaries aged 65 years and older. Waves 11 and 12 of NHATS were chosen for their comprehensive 6-domain cognitive assessment, unlike earlier waves that measured only episodic memory, executive function, and orientation. Data were gathered during in-home interviews by NHATS interviewers. For those included in both waves, only data from wave 12 were retained to ensure the latest cognitive assessments were used. The reporting of this study followed the CHERRIES checklist [18].

Participants and Sample Size Calculation

Eligible participants were cognitively intact individuals aged 65 years and older, not residing in nursing homes, and without signs of cognitive impairment. Cognitive intactness was determined by absence of potential or probable dementia. According to previous NHATS literature, potential dementia was defined by scores≤1.5SDs below the mean in one cognitive domain, while probable dementia was indicated by similar scores in at least two domains, meeting AD8 criteria, or having a clinical dementia diagnosis [19]. To ensure adequate statistical power, our network analysis of 11 nodes and 55 edges required a minimum of 165 participants, adhering to the 3-participants-per-parameter rule [20].

Measures

Cognitive performance was assessed across six domains: (1) Episodic memory, scored 0 - 20 from immediate and delayed recall of 10 words; (2) executive function, scored 0 - 5 by the clock drawing test; (3) orientation, scored 0 - 8 from knowledge of the current date and names of the president or vice president; (4) psychomotor function, measured by reaction speed in log-transformed milliseconds from the cogstate detection task,

XSL•FO
RenderX

where participants respond when a card turns face up; (5) visual attention, assessed by reaction speed for correct responses in log-transformed milliseconds from the cogstate identification task, where participants decide if a card is red or black; and (6) working memory, evaluated by accuracy in the cogstate one card back task, asking participants to remember if they have seen a card before. Cogstate, a tablet-based test used since wave 11 in NHATS, includes card detection, identification, and one card back activities, expanding cognitive assessments beyond traditional tests. This computerized assessment has been validated against traditional paper-based cognitive tests, demonstrating adequate reliability and validity in differentiating adults with cognitive impairment from those without [21]. Further details are available in the NHATS Cogstate user guide [22].

IADLs were assessed via self-reports on managing medication, laundry, shopping, meal preparation, and banking. Participants reported performance over the past month using 5 options: “1” Did not do by self last month; “2” Did by self last month with no difficulty; “3” Did by self last month with difficulty; “4” -Don’t know or refuse, with no difficulty; and “5” Don’t know or refuse, with difficulty. We dichotomized responses rather than using a Likert scale because the options do not form a natural continuum, but instead distinguish independent performance from any difficulty. Participants with no difficulty (responses “2” or “4”) were classified as “no difficulty,” while those with difficulty or inability (responses “1,” “3,” or “5”) were classified as “difficulty.” This binary approach preserves the key distinction in functional independence and aligns with previous research [23].

Sociodemographic variables included age (70 - 74, 75 - 79, 80 - 84, 85 - 89, and 90+ years), sex (female or male), living arrangement (alone, with others), race (Hispanic, non-Hispanic Black, non-Hispanic White, and other non-Hispanic), income status (poverty, low income, and normal), and self-rated health (poor, fair, good, very good, and excellent). Income status was defined according to federal guidelines from the Office of the Assistant Secretary for Planning and Evaluation at the US Department of Health and Human Services (ASPE HHS, 2024 version) as follows:≤100% Federal Poverty Level (FPL), >100% to≤200% FPL, and>200% FPL, replacing the previous labels of “poverty,” “low income,” and “normal.” Self-rated health was measured using a single 5-point Likert-scale item (1=poor, 5=excellent): “Would you say that, in general, your health is poor, fair, good, very good, or excellent?”

Statistical Analysis

Data were organized in a Microsoft Excel database and subjected to rigorous quality control checks. The analysis was carried out using R statistical software (version 4.1.1; R Core Team). Descriptive statistics summarized participant demographics and performance in IADLs and cognitive functions. Continuous variables were checked for normality with P-P plots and described using mean and SD; categorical variables were presented as frequencies and percentages. To maintain consistency across all nodes within the IADL and cognition networks, necessary reverse coding adjustments were made to ensure that higher scores consistently indicate diminished capabilities. Network analysis proceeded through 5 phases: evaluating topological overlap, estimating the network, assessing network stability, calculating centrality and bridge centrality indices, and conducting network comparison tests.

Checking Topological Overlap

We used the R package *network tools*’ goldbricker function to identify unique variables and avoid artificial relationships from symptom similarity in our network analysis. A significance threshold of 0.25 and a *P* value<.01 determined statistical significance [24].

Network Estimation

We developed 3 networks for our study: an IADL network for all 5 IADL domains, a cognitive network for all 6 cognitive domains, and a bridge network linking both. Each network consisted of nodes (items within each domain) and edges (relationships between items). For the cognition network with continuous data, we applied the EBICglasso method, which used the Extended Bayesian Information Criterion (EBIC) with the least absolute shrinkage and selection operator (LASSO) for partial correlation analysis, reducing confounding by shrinking coefficients and zeroing smaller correlations. The IADL network, based on binary data, was analyzed using the IsingFit method, which used logistic regression to adjust node states and determine conditional probabilities. The bridge network was assessed using the “mgm” method, designed for mixed data types, using conditional independence tests tailored to heterogeneous data. Network visualization was performed using R packages *bootnet* and *qgraph*, where edge thickness represented association strength—blue for positive and red for negative associations [25].

Network Stability

The *bootnet* package was used to assess edge and centrality stability in each network [25]. Edge stability was evaluated through nonparametric bootstrap, with 95% CIs reflecting accuracy; narrower CIs indicate higher network reliability [25]. Centrality stability was measured with a case-dropping subset bootstrap, as reflected by the correlation stability coefficient (CS-C); values above 0.25, ideally over 0.5, denote optimal stability [25].

Central Node, Centrality, Bridge Node, and Bridge Centrality

A central node in a network has substantial influence due to its extensive connections with other nodes [26]. Bridge nodes connect different communities or clusters within a network, facilitating interactions that would otherwise not occur [27]. Centrality measures in network analysis typically include strength, betweenness, closeness, and expected influence; however, due to the instability of betweenness and closeness, and because strength ignores negative edges (summing only absolute values), we exclusively used expected influence for central nodes and bridge expected influence for bridge nodes [28]. The expected influence index, which accounts for both positive and negative edge values, was calculated using the *qgraph* package in R [26,27]. Similarly, the top bridge node was identified through the highest bridge expected influence (1-step) index, which sums the edge values connecting the node

XSL•FO
RenderX

to those outside its community, determined by the *networktools* package in R [26,27]. Differences in node centrality were analyzed using Wilcoxon tests with 1000 bootstrapped indices from the bootnet package in R, applying Benjamini-Hochberg corrections for multiple comparisons [29].

Network Comparison Test

To analyze gender differences across 3 networks, we used the *network comparison test* package in R. This involved performing both a network invariance test, which identified significant variations in edges among subgroup networks, and a global strength invariance test, which evaluated the total weighted sum of all edges to measure the connection strength among network variables. Should the network invariance test yield significant results, we then conducted an edge invariance test to pinpoint specific edge pairs that varied between subgroups. In addition, we compared node centrality between men and women. To adjust for multiple comparisons at the level of individual edges and centralities, the Benjamini-Hochberg correction method was used.

Ethical Considerations

This secondary analysis of the NHATS dataset relies on publicly available data. The original data collection, which obtained informed consent from all participants, was approved by the Johns Hopkins University IRB. As no restricted data were used, further IRB review was not required.

Results

Sample Characteristics and Descriptions of IADL and Cognitive Domains

Of the 2239 participants (1194 from wave 12 and 245 from wave 11), 1263 were female (56.41%), and 1720 (76.82%) were White. The predominant age group was 75‐79 years, representing 748 participants (33.41%). Detailed sociodemographic data, as well as descriptions and abbreviations of the IADL items and cognitive domains, are provided in Table 1.

XSL•FO
RenderX

Table . Demographics, descriptions, and abbreviations of instrumental activities of daily living (IADL) and cognitive domain items (N=2239).

Variables	Results	Range
Demographics, n (%)		
Sex		
Female	1263 (56.41)	—[a]
Male	976 (43.59)	—
Age group		
70 - 74 years	214 (9.56)	—
75 - 79 years	748 (33.41)	—
80 - 84 years	620 (27.69)	—
85 - 89 years	395 (17.64)	—
90+ years	262 (11.70)	—
Self-rated health		
Poor	69 (3.08)	—
Fair	431 (19.25)	—
Good	893 (39.88)	—
Very good	676 (30.19)	—
Excellent	169 (7.55)	—
Missing value	1 (0.04)	—
Race		
Non-Hispanic others	41 (1.83)	—
Non-Hispanic Black	375 (16.75)	—
Hispanic	81 (3.62)	—
Non-Hispanic White	1720 (76.82)	—
Missing	22 (0.98)	—
Living arrangement		
Alone	809 (36.13)	—
Living with someone	1430 (63.87)	—
Income status		
Poverty	859 (38.37)	—
Low income	413 (18.45)	—
Normal	967 (43.19)	—
Description of items (abbreviations in networks), n (%)		
Difficulty in managing medication (I1)		
No	1798 (80.30)	—
Yes	441 (19.70)	—
Difficulty in managing laundry (I2)		
No	1321 (59)	—
Yes	918 (41)	—
Difficulty in managing shopping (I3)		
No	1214 (54.22)	—
Yes	1025 (45.78)	—
Difficulty in managing meal (I4)		
No	1373 (61.32)	—

XSL•FO
RenderX

Variables	Results	Range
Yes	866 (38.68)	—
Difficulty in managing banking (I5)		
No	1505 (67.22)	—
Yes	734 (32.78)	—
Episodic memory (C1), mean (SD)	10.15 (3.21)	0-9
Executive function (C2), mean (SD)	0.84 (0.89)	0-5
Orientation (C3), mean (SD)	0.96 (1.33)	0-8
Psychomotor function (C4), mean (SD)	2.66 (0.13)	2.36-3.26
Visual attention (C5), mean (SD)	2.82 (0.10)	2.57-3.41
Working memory (C6)	0.34 (0.25)	0-1.57

[a]Not applicable.

Items Remained After Redundancy Check

The Goldbricker analysis confirmed no redundancy in the IADL and cognitive domains, thus all items were retained.

Stability of IADLs, Cognition, and the Bridge Networks

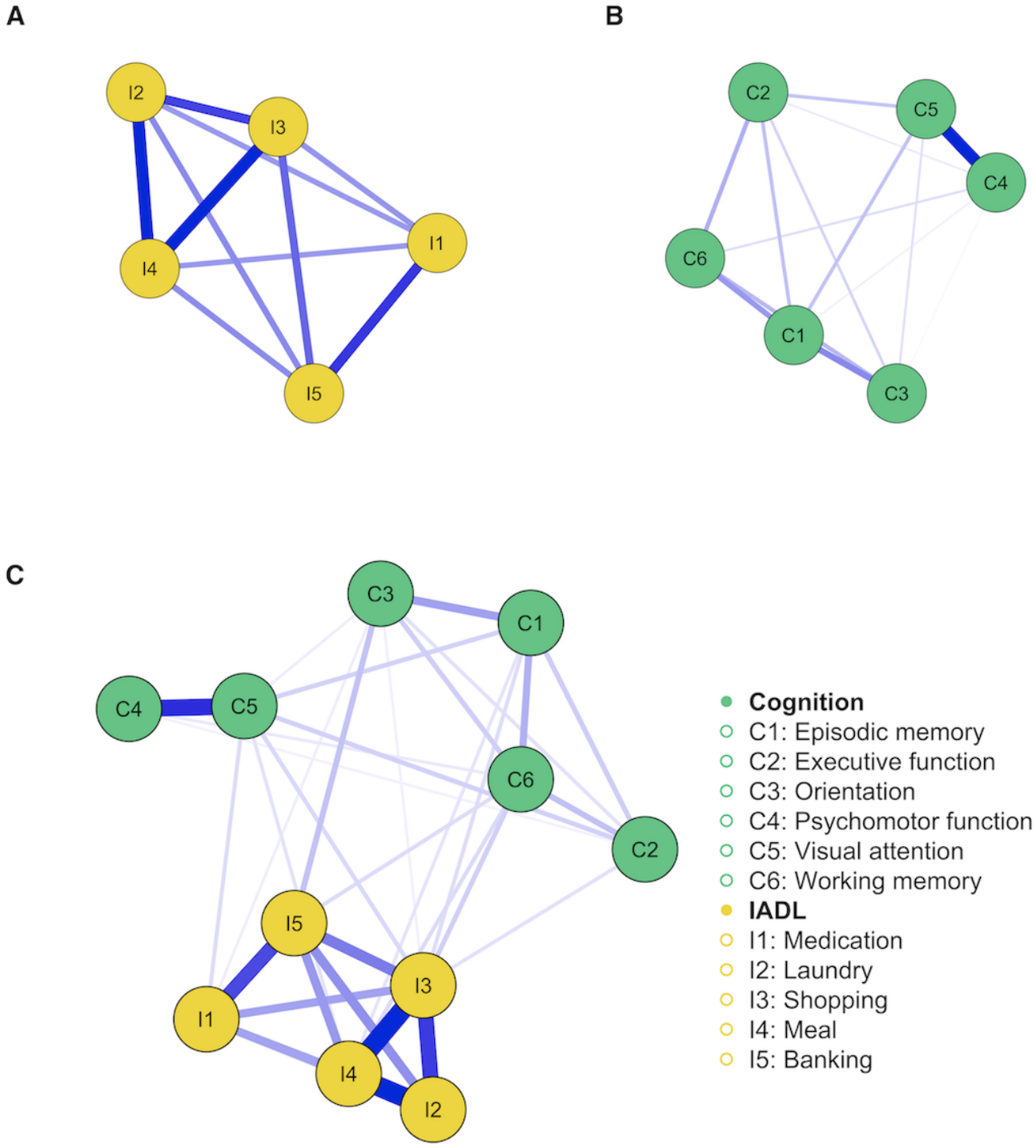
The bootstrapped 95% CI analysis revealed narrow CIs across all 3 networks (IADL, cognition, and bridge), indicating precise edge-weight estimates (Figures S1, S3, and S5 in Multimedia Appendix 1). In addition, CS-C values for the IADL, cognition, and bridge networks were 0.67, 0.75, and 0.44, respectively, all surpassing the recommended threshold of 0.25, confirming the networks' interpretability and reliability (Figures S2, S4, and S6 in Multimedia Appendix 1).

IADL Network

Figure 1A illustrates the IADL network, with all edges (10/10, 100%) nonzero, indicating strong connectivity among nodes. The most robust connections were between I3 and I4 (I3: shopping-I4:meal, edge weight 1.38), I2 and I4 (I2: laundry-I4:meal, 1.33), and I1 and I5 (I1: medication- I5: banking, 1.08). Logistic regression coefficients for other edges are detailed in Table S1 in Multimedia Appendix 1. Figure 2A shows that node I4 (meal) had the highest expected influence of 3.87. Figure 2B's centrality bootstrapped difference test highlights I4's (meal's) significantly higher influence (all $P<.05$ after Benjamini-Hochberg corrections), underscoring its central role in the network.

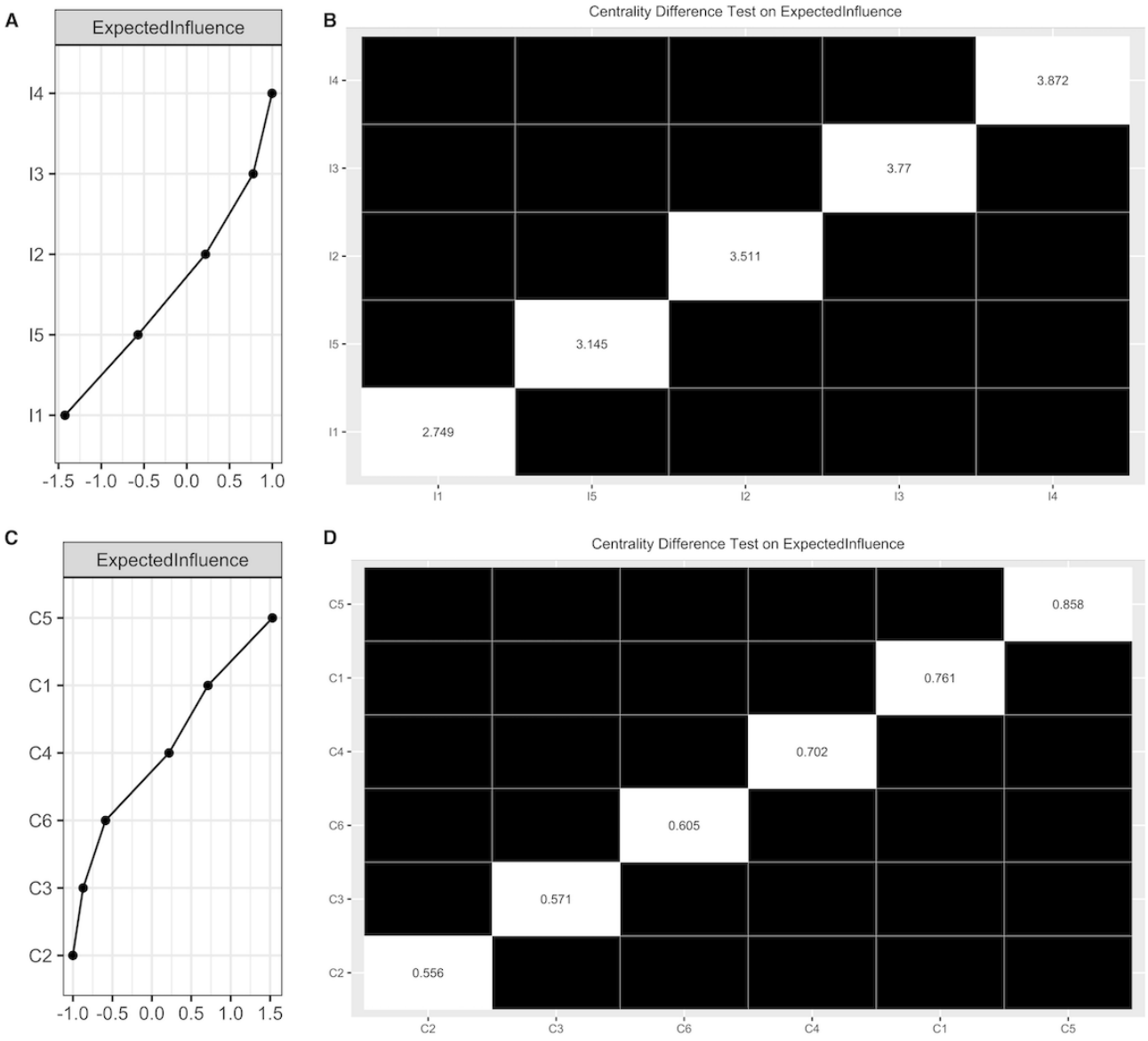
XSL•FO
RenderX

Figure 1. Network structure of instrumental activities of daily living (IADL) network, cognition network, and the bridge network.



XSL•FO
RenderX

Figure 2. Expected influence centrality index and centrality bootstrapped difference tests for variables in the IADL network (Pane A and Panel B) and the cognition network (Panel C and Panel D). Gray boxes indicate variables that do not significantly differ from one-another. Black boxes represent variables that differ significantly from one another (α=.05). White boxes show the values of bridge expected influence. I1: difficulty in managing medication; I2: difficulty in managing laundry; I3: difficulty in managing shopping; I4: difficulty in managing meals; I5: difficulty in managing banking; C1: episodic memory; C2: executive function; C3: orientation; C4: psychomotor function; C5: visual attention; C6: working memory.



Cognition Network

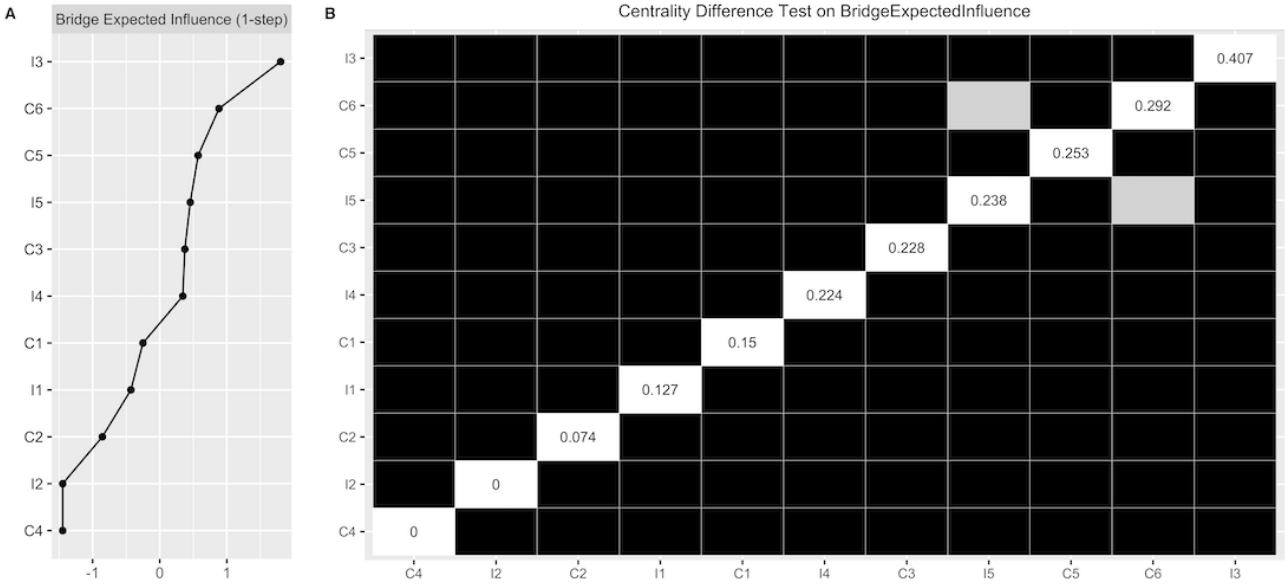
Figure 1B shows the cognition network structure, with 14/15 edges (93.33%) nonzero, reflecting strong connectivity among nodes. The strongest connections were C4-C5 (C4: psychomotor function-C5: visual attention; edge weight 0.54), C1-C3 (C1: episodic memory-C3: orientation, 0.25), and C1-C6 (C1: episodic memory-C6: working memory, 0.22). The partial correlation matrix for other edges is in Table S2 in Multimedia Appendix 1. Figure 2C displays the expected influence index for all nodes, with C5 (visual attention) having the highest at 0.86. Figure 2D's centrality bootstrapped difference test underscores C5's (visual attention's) significant influence within the cognitive domains (all P<.05 after Benjamini-Hochberg corrections).

Bridge Network

Figure 1C shows the bridge network between IADL and cognition, with 34/55 edges (61.82%) nonzero, indicating strong connectivity. Details on all edges are in Table 3 in Multimedia Appendix 1. Figure 3A highlights the bridge expected influence index, with I3 (shopping) recording the highest at 0.41, followed by I5 (banking, 0.24) and I4 (meal, 0.22). Significant bridge edges include I3-C6 (I3: shopping-C6: working memory, edge weight: 0.12), I5-C3 (I5: banking-C3: orientation, edge weight: 0.15), and I4-C6 (I4: meal-C6: working memory, edge weight: 0.09). Figure 3B's centrality bootstrapped difference test confirms I3's (shopping's) prominent role in connecting IADL and cognitive domains (all P<.05 after Benjamini-Hochberg corrections).

XSL•FO
RenderX

Figure 3. Bridge expected influence centrality index (Panel A) and centrality bootstrapped difference tests (Panel B) for variables in the bridge network. Gray boxes indicate variables that do not significantly differ from one-another. Black boxes represent variables that differ significantly from one another (α=.05). White boxes show the values of bridge expected influence. I1: difficulty in managing medication; I2: difficulty in managing laundry; I3: difficulty in managing shopping; I4: difficulty in managing meals; I5: difficulty in managing banking; C1: episodic memory; C2: executive function; C3: orientation; C4: psychomotor function; C5: visual attention; C6: working memory.



Sex Differences in Networks

Network invariance and global strength tests revealed no significant sex differences in the cognition network (*M*=0.07, *P*=.64; *S*=0.10, *P*=.21). Conversely, significant sex differences were evident in the IADL and bridge networks, confirmed by network invariance tests (IADL: *M*=0.95; *P*=.005 and bridge: *M*=0.49; *P*=.002) and global strength tests (IADL: *S*=1.94; *P*=.001 and bridge: *S*=1.79; *P*=.002). Following these findings, further edge and centrality invariance tests were conducted for the IADL and bridge networks between sexes. In the IADL network, edge invariance tests identified significant sex disparities between I2 (laundry) and I4 (meal), with females demonstrating stronger associations (female 1.69, male 0.74; *P*=.001) and greater centrality in I4 (meal: difference=1.99; *P*=.007). However, in the bridge network, no significant sex differences were detected through further edge and centrality invariance tests (all *P*>.05 after Benjamini–Hochberg corrections).

Discussion

Principal Findings

Using network analysis, the main findings of this study elucidates the detailed interactions within and between IADL and cognitive domains in older adults, identifying 3 key variables as predictive markers and potential intervention targets for enhancing global IADL and cognitive function, while also noting sex differences. First, meal preparation difficulty and visual attention are central nodes within their respective IADL and cognition networks, with higher levels predictive of better functionality; targeted modifications could significantly improve overall functionality. Secondly, shopping difficulty within the IADL network has the strongest association with global cognition. Given the cause-and-effect relationship between IADL and cognitive function, early interventions targeting shopping difficulties could effectively boost global cognition. Finally, a sex difference was observed, with meal preparation exerting a greater influence in the IADL network among females than males, suggesting its higher predictive relevance and intervention efficacy for females.

Within the IADL network, the biggest 3 connections include those between shopping and meal preparation, laundry and meal preparation, and medication management and banking. The strong link between shopping and meal preparation is due to their shared planning, organization, and physical demands necessary for food tasks [30-32]. Likewise, laundry and meal preparation share organizational and physical demands [13]. The link between medication management and banking stems from their reliance on the same executive functions and working memory [33,34]. Importantly, meal preparation is central in the IADL network, perhaps due to impacts on other domains like medication, shopping, and banking owning to necessary executive skills such as multitasking and planning [35,36]. Data from our cross-sectional analysis indicate that meal preparation is the most central node within the IADL network. This finding suggests that same unite improvement in meal preparation ability can yield the largest overall enhancement in IADL performance relative to other tasks. Therefore, targeted interventions—such as the integration of assistive kitchen technologies or cognitive orthoses—may be especially effective in promoting independent cooking and, by extension, broader functional independence [37,38].

In the cognition network, the strongest connections are between psychomotor function and visual attention, and between episodic memory, orientation, and working memory. The link between psychomotor and visual attention is due to their roles in coordinated, rapid response tasks [14]. Significant correlations also exist between episodic memory and gray matter volume in the bilateral hippocampus and parahippocampal gyrus, which are key for orientation [39]. Furthermore, the association

XSL•FO
RenderX

between episodic memory and working memory is supported by evidence that working memory capacity and prefrontal cortex executive functions are essential for episodic memory formation and retrieval [40]. Importantly, visual attention, the most central node in the cognitive network, is crucial for selective focus necessary for advanced cognitive processes like reasoning and problem-solving [41], and it enhances complex task execution through interactions with working memory [42]. Studies show that enhancing visual attention can lead to sustained improvements in cognitive performance [43,44], supporting its pivotal role in overall cognition and highlighting the need for targeted interventions to enhance focus and processing of visual stimuli. Potential interventions include computer-based training programs, video games, mindfulness exercises, and virtual reality applications specifically designed for visual tracking tasks, all of which may contribute to enhanced cognitive function [45-48].

In the bridge network, which integrates both IADLs and cognition, connectivity is notably high, with 61.82% (34/55) of edges being nonzero. This finding indicates a robust interconnection among multiple cognitive domains and IADLs, suggesting that daily tasks rely on a synergy of cognitive processes rather than on any single discrete skill [49,50]. Such an observation aligns with mounting evidence that real-life cognition operates as an integrated set of processes-often termed functional cognition [51,52]. While traditional neuropsychological models emphasize isolated cognitive constructs, functional cognition highlights how domains such as attention, executive function, and memory converge to support everyday activities [53]. By adopting this framework, our findings on the specific links ("edges") between each cognitive domain and each IADL can offer theoretical guidance for real-world functional cognition rehabilitation or training aimed at improving IADL performance. Importantly, the prominent bridge edge was identified between shopping and working memory. When considering the overall impact on global cognition, shopping ranked first as the bridge IADL, followed by banking and meal preparation. These latter activities are also highly demanding cognitively and should be targeted in interventions. However, if resources are limited, prioritizing shopping may yield the greatest benefits in enhancing overall cognitive function. Notably, shopping functions as a bridge IADL because it draws on a broad range of cognitive skills: episodic memory for recalling past purchases and layouts; executive function for planning and budgeting; spatial orientation for navigation; psychomotor skills for handling products and carts; visual attention for identifying items; and working memory for tracking purchases and costs [33,54-57]. To leverage its bridge position, shopping tasks could be integrated into routine cognitive assessments, and regular shopping activities could be encouraged to maintain cognitive function. Virtual reality simulations of shopping tasks, artificial intelligence-powered service robotics, and the use of audio recorders as assistive technology to enhance shopping independence among older adults [58,59], can leverage the bridge IADL role of shopping to effectively improve global cognitive function.

Sex differences within the IADL network show females with a stronger association between laundry and meal preparation and a higher centrality of meal preparation. These patterns likely result from societal norms assigning women more IADL responsibilities, especially laundry and meal preparation [17,60,61]. The elevated centrality of meal preparation in women's IADL networks suggests a ripple effect where challenges in meal preparation deplete time management, mental energy, and physical resources, reducing efficiency in other tasks. This ripple effect also explains the stronger association between meal preparation and laundry among women compared with men. Previous studies indicate that IADLs do not measure equivalently for men and women [61]; our study finds that meal preparation has a lower predictive value for overall IADL function in males than in females. Interventions to improve IADL performance through meal preparation should be tailored with an awareness that these activities vary in importance and difficulty between sexes.

Limitations

Several limitations are worth mentioning. First, the assessment of IADL relied on self-reported data, which, while expedient, may compromise reliability and necessitate cautious interpretation of the findings. Future research should use more objective measures to validate these results. Second, the cross-sectional design precludes the establishment of causality and does not capture the temporal dynamics between IADL capabilities and cognitive function, underscoring the need for longitudinal approaches. Third, we excluded older adults with dementia, as dementia-related deficits may mask subtle variations in both cognition and IADL performance, which is critical for identifying central or bridge nodes in network analysis. Consequently, our findings should be interpreted primarily for community-dwelling older adults with relatively preserved cognitive function. Future research should incorporate participants with more severe cognitive impairments to validate whether our findings remain consistent. Fourth, although we set eligibility at ≥65 years, all final participants were aged 70+ years, likely because those aged 65 - 69 years in earlier waves either dropped out or turned 70 years old by Waves 11 - 12. This may limit the generalizability of our findings to younger older adults. Fifth, our study relies on the NHATS dataset, which limits the assessment to a select set of IADL and cognitive domains. Sixth, one limitation of our study is that the network analysis did not adjust for external confounders (eg, age and cultural background). Network models focus solely on the relationships among the included nodes and do not account for factors outside the network. Future research might address this limitation by residualizing each node on confounders before constructing the network, or by using subgroup or multigroup analyses to examine how these factors influence network structure. While these measures capture key aspects of functioning, we acknowledge that not including additional, more nuanced domains may affect the generalizability of our findings. Finally, the generalizability of our findings is limited to the American population studied; it remains unclear if these results can be generalized to populations with differing cultural, economic, or health system backgrounds. Further studies should

XSL•FO
RenderX

expand the demographic scope to determine if these findings hold across diverse populations.

Conclusions

The central IADL, central cognitive domain, and bridge IADL connecting global cognition were meal preparation, visual attention, and shopping, respectively, underscoring the need for targeted interventions to maximize resource efficiency and effectiveness. Specifically, enhancing meal preparation in older adults may significantly boost holistic IADL capabilities through interventions such as cooking classes, nutritional education, and tailored tools, along with support services such as interdisciplinary collaboration, caregiver training, and smart appliances. Similarly, focusing on visual attention training through methods such as computer-based programs, neurofeedback, and mindfulness exercises may substantially improve global cognitive function. Given the link between IADL performance and cognitive function, interventions centered on shopping are expected to be highly effective. This can be achieved by integrating shopping tasks into cognitive assessments and promoting regular shopping activities. Technological aids such as GPS, virtual reality simulations, and caregiver education on the cognitive benefits of shopping can further support elderly care and quality of life. In addition, observed sex differences suggest that meal preparation interventions may vary in effectiveness, with potentially lower efficacy among males, highlighting the need for tailored strategies to maximize outcomes.

Acknowledgments

This study was supported by funding from the National Institute on Aging (grant R01AG080613).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Tables S1-S3 and Figures S1-S6.
[DOCX File, 6100 KB - aging_v8i1e67632_app1.docx]

References

1. Kiely KM. Cognitive function. In: Michalos AC, editor. Encyclopedia of Quality of Life and Well-Being Research: Dordrecht: Springer Netherlands; 2014:974-978. [doi: 10.1007/978-94-007-0753-5_426]
2. Pais R, Ruano L, P Carvalho O, Barros H. Global cognitive impairment prevalence and incidence in community dwelling older adults—a systematic review. Geriatrics (Basel) 2020 Oct 27;5(4):84. [doi: 10.3390/geriatrics5040084] [Medline: 33121002]
3. Zhu CW, Sano M, Ferris SH, Whitehouse PJ, Patterson MB, Aisen PS. Health-related resource use and costs in elderly adults with and without mild cognitive impairment. J Am Geriatr Soc 2013 Mar;61(3):396-402. [doi: 10.1111/jgs.12132] [Medline: 23414481]
4. Guo HJ, Sapra A. Instrumental Activity of Daily Living: StatPearls Publishing; 2023.
5. Toups R, Chirles TJ, Ehsani JP, et al. Driving performance in older adults: current measures, findings, and implications for roadway safety. Innov Aging 2022;6(1):igab051. [doi: 10.1093/geroni/igab051] [Medline: 35028434]
6. Sino CGM, Sietzema M, Egberts TCG, Schuurmans MJ. Medication management capacity in relation to cognition and self-management skills in older people on polypharmacy. J Nutr Health Aging 2014 Jan;18(1):44-49. [doi: 10.1007/s12603-013-0359-2] [Medline: 24402388]
7. Li X, Song R, Qi X, et al. Influence of cognitive reserve on cognitive trajectories. Neurology (ECronicon) 2021 Oct 26;97(17):e1695-e1706. [doi: 10.1212/WNL.0000000000012728] [Medline: 34493618]
8. Buele J, Varela-Aldás JL, Palacios-Navarro G. Virtual reality applications based on instrumental activities of daily living (iADLs) for cognitive intervention in older adults: a systematic review. J Neuroeng Rehabil 2023 Dec 19;20(1):168. [doi: 10.1186/s12984-023-01292-8] [Medline: 38110970]
9. Huang S, Zhong W, Cheng Q, Shuai Y, Zhu J, Diao J. Instrumental activities of daily living function and cognitive status among Chinese older adults: a serial multiple mediation model. Front Public Health 2024;12:1378979. [doi: 10.3389/fpubh.2024.1378979] [Medline: 38756886]
10. Toth C, Tulliani N, Bissett M, Liu KP. The relationship between cognitive function and performance in instrumental activities of daily living in older adults. British Journal of Occupational Therapy 2022 Feb;85(2):120-129. [doi: 10.1177/03080226211008722]
11. Romero-Ayuso D, Castillero-Perea Á, González P, et al. Assessment of cognitive instrumental activities of daily living: a systematic review. Disabil Rehabil 2021 May;43(10):1342-1358. [doi: 10.1080/09638288.2019.1665720] [Medline: 31549907]
12. Lau SCL, Connor LT, Heinemann AW, Baum CM. Cognition and daily life activities in stroke: a network analysis. OTJR (Thorofare N J) 2022 Oct;42(4):260-268. [doi: 10.1177/15394492221111730] [Medline: 35838366]

XSL•FO
RenderX

13. Jefferson AL, Paul RH, Ozonoff A, Cohen RA. Evaluating elements of executive functioning as predictors of instrumental activities of daily living (IADLs). Arch Clin Neuropsychol 2006 May;21(4):311-320. [doi: 10.1016/j.acn.2006.03.007] [Medline: 16814980]
14. Hemager N, Vangkilde S, Thorup A, et al. Visual attention in 7-year-old children at familial high risk of schizophrenia or bipolar disorder: The Danish high risk and resilience study VIA 7. J Affect Disord 2019 Nov 1;258:56-65. [doi: 10.1016/j.jad.2019.07.079] [Medline: 31394459]
15. Hevey D. Network analysis: a brief overview and tutorial. Health Psychol Behav Med 2018 Sep 25;6(1):301-328. [doi: 10.1080/21642850.2018.1521283] [Medline: 34040834]
16. Levine DA, Gross AL, Briceño EM, et al. Sex differences in cognitive decline among US adults. JAMA Netw Open 2021 Feb 1;4(2):e210169. [doi: 10.1001/jamanetworkopen.2021.0169] [Medline: 33630089]
17. Ryu HJ, Moon Y. Gender differences in items of the instrumental activities of daily living in mild cognitive impairment and alzheimer's disease dementia. Dement Neurocogn Disord 2024 Apr;23(2):107-114. [doi: 10.12779/dnd.2024.23.2.107] [Medline: 38720826]
18. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
19. Kasper JD, Freedman VA, Spillman BC. Classification of persons by dementia status in the national health and aging trends study. Technical paper.: National Health and Aging Trends Study (NHATS); 2013.
20. Li J, Fong DYT, Lok KYW, et al. Key lifestyles and interim health outcomes for effective interventions in general populations: A network analysis of A large international observational study. J Glob Health 2023 Oct 20;13:04125. [doi: 10.7189/jogh.13.04125] [Medline: 37861130]
21. Lim YY, Ellis KA, Harrington K, et al. Use of the CogState Brief Battery in the assessment of Alzheimer's disease related cognitive impairment in the Australian Imaging, Biomarkers and Lifestyle (AIBL) study. J Clin Exp Neuropsychol 2012;34(4):345-358. [doi: 10.1080/13803395.2011.643227] [Medline: 22248010]
22. Hu M, Freedman VA. National hHealth and Aging Trends Study (NHATS) cogstate user guide. : Johns Hopkins University School of Public Health; 2022 URL: https://nhats.org/sites/default/files/2022-07/NHATS%20Cogstate%20User%20Guide.pdf [accessed 2025-03-11]
23. Edwards RD, Brenowitz WD, Portacolone E, et al. Difficulty and help with activities of daily living among older adults living alone with cognitive impairment. Alzheimers Dement 2020 Aug;16(8):1125-1133. [doi: 10.1002/alz.12102] [Medline: 32588985]
24. Hittner JB, May K, Silver NC. A Monte Carlo evaluation of tests for comparing dependent correlations. J Gen Psychol 2003 Apr;130(2):149-168. [doi: 10.1080/00221300309601282] [Medline: 12773018]
25. Epskamp S, Borsboom D, Fried EI. Estimating psychological networks and their accuracy: A tutorial paper. Behav Res Methods 2018 Feb;50(1):195-212. [doi: 10.3758/s13428-017-0862-1] [Medline: 28342071]
26. Fonseca-Pedrero E. Network analysis in psychology. Papeles del Psicólogo 2018;39(1):1-12. [doi: 10.23923/pap.psicol2018.2852]
27. Jones PJ, Ma R, McNally RJ. Bridge centrality: a network approach to understanding comorbidity. Multivariate Behav Res 2021;56(2):353-367. [doi: 10.1080/00273171.2019.1614898] [Medline: 31179765]
28. Li J, Fong DYT, Lok KYW, et al. Country-specific key lifestyle factors and health outcomes for resource allocation in the general population: a network analysis across 29 countries. J Glob Health 2025 Jan 10;15:04011. [doi: 10.7189/jogh.15.04011] [Medline: 39791329]
29. Li J, Fong DYT, Lok KYW, et al. Key lifestyles and health outcomes across 16 prevalent chronic diseases: A network analysis of an international observational study. J Glob Health 2024;14. [doi: 10.7189/jogh-14-04068]
30. Lawton M, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Nurs Res 1970;19(3):278. [doi: 10.1097/00006199-197005000-00029]
31. Giere S, Rudrud E, McKay S. A functional approach to meal planning, preparation, and shopping skills. Australia and New Zealand Journal of Developmental Disabilities 1989 Jan;15(2):81-97. [doi: 10.1080/07263868900033551]
32. Davies G. Time, food shopping and food preparation: some attitudinal linkages. British Food Journal 1997 Apr 1;99(3):80-88. [doi: 10.1108/00070709710168914]
33. Engle RW. Working memory capacity as executive attention. Curr Dir Psychol Sci 2002 Feb;11(1):19-23. [doi: 10.1111/1467-8721.00160]
34. Carpenter PA, Just MA, Reichle ED. Working memory and executive function: evidence from neuroimaging. Curr Opin Neurobiol 2000 Apr;10(2):195-199. [doi: 10.1016/s0959-4388(00)00074-x] [Medline: 10753796]
35. Craik FIM, Bialystok E. Planning and task management in older adults: Cooking breakfast. Mem Cognit 2006 Sep;34(6):1236-1249. [doi: 10.3758/BF03193268]
36. Doherty TA, Barker LA, Denniss R, Jalil A, Beer MD. The cooking task: making a meal of executive functions. Front Behav Neurosci 2015;9:22. [doi: 10.3389/fnbeh.2015.00022] [Medline: 25717294]
37. Kuoppamäki S, Tuncer S, Eriksson S, McMillan D. Designing kitchen technologies for ageing in place: a video study of older adults' cooking at home. Proc ACM Interact Mob Wearable Ubiquitous Technol 2021;5(2):1-19. [doi: 10.1145/3463516]

XSL•FO
RenderX

38. Yaddaden A, Couture M, Gagnon-Roy M, et al. Using a cognitive orthosis to support older adults during meal preparation: Clinicians' perspective on COOK technology. J Rehabil Assist Technol Eng 2020;7:2055668320909074. [doi: 10.1177/2055668320909074] [Medline: 32435504]
39. Maillet D, Rajah MN. Association between prefrontal activity and volume change in prefrontal and medial temporal lobes in aging and dementia: a review. Ageing Res Rev 2013 Mar;12(2):479-489. [doi: 10.1016/j.arr.2012.11.001] [Medline: 23183352]
40. Guarnieri RV, Alves MV, Nogueira AML, et al. Haloperidol-induced impairment on working memory capacity affecting long term memory performance: the binding hypothesis. Neuroscience. . [doi: 10.1101/2019.12.24.888073]
41. Moore T, Zirnsak M. Neural mechanisms of selective visual attention. Annu Rev Psychol 2017 Jan 3;68(1):47-72. [doi: 10.1146/annurev-psych-122414-033400] [Medline: 28051934]
42. Giuffrida V, Marc IB, Ramawat S, et al. Reward prospect affects strategic adjustments in stop signal task. Front Psychol 2023;14:1125066. [doi: 10.3389/fpsyg.2023.1125066] [Medline: 37008850]
43. Willis SL, Tennstedt SL, Marsiske M, et al. Long-term effects of cognitive training on everyday functional outcomes in older adults. JAMA 2006 Dec 20;296(23):2805-2814. [doi: 10.1001/jama.296.23.2805] [Medline: 17179457]
44. Ball K, Berch DB, Helmers KF, et al. Effects of cognitive training interventions with older adults: a randomized controlled trial. JAMA 2002 Nov 13;288(18):2271-2281. [doi: 10.1001/jama.288.18.2271] [Medline: 12425704]
45. Belchior P, Marsiske M, Sisco SM, et al. Video game training to improve selective visual attention in older adults. Comput Human Behav 2013 Jul 1;29(4):1318-1324. [doi: 10.1016/j.chb.2013.01.034] [Medline: 24003265]
46. Ferreira S, Marmeleira J, Del Pozo-Cruz J, et al. Acute effects of augmented reality exergames versus cycle ergometer on reaction time, visual attention, and verbal fluency in community older adults. Int J Environ Res Public Health 2022 Nov 8;19(22):14667. [doi: 10.3390/ijerph192214667] [Medline: 36429385]
47. Chambon C, Herrera C, Romaiguere P, Paban V, Alescio-Lautier B. Benefits of computer-based memory and attention training in healthy older adults. Psychol Aging 2014 Sep;29(3):731-743. [doi: 10.1037/a0037477] [Medline: 25244490]
48. Isbel B, Weber J, Lagopoulos J, Stefanidis K, Anderson H, Summers MJ. Neural changes in early visual processing after 6 months of mindfulness training in older adults. Sci Rep 2020 Dec 3;10(1):21163. [doi: 10.1038/s41598-020-78343-w] [Medline: 33273707]
49. Giles GM, Radomski MV, Wolf TJ. Cognition, cognitive rehabilitation, and occupational performance. Am J Occup Ther 2019 Nov 1;73(Supplement_2):7312410010p1-7312410010p25. [doi: 10.5014/ajot.2019.73S201]
50. Giles GM, Edwards DF, Morrison MT, Baum C, Wolf TJ. Screening for functional cognition in postacute care and the Improving Medicare Post-Acute Care Transformation (IMPACT) Act of 2014. Am J Occup Ther 2017;71(5):7105090010p1-7105090010p6. [doi: 10.5014/ajot.2017.715001] [Medline: 28809645]
51. Giles GM, Edwards DF, Baum C, et al. Making functional cognition a professional priority. Am J Occup Ther 2020;74(1):7401090010p1-7401090010p6. [doi: 10.5014/ajot.2020.741002] [Medline: 32078504]
52. Giles GM, Goverover Y, Foster ER, et al. Functional cognition: moving the field forward. Occup Ther J Res ;2024:15394492241296811. [doi: 10.1177/15394492241296811]
53. Wesson J, Clemson L, Brodaty H, Reppermund S. Estimating functional cognition in older adults using observational assessments of task performance in complex everyday activities: A systematic review and evaluation of measurement properties. Neuroscience & Biobehavioral Reviews 2016 Sep;68:335-360. [doi: 10.1016/j.neubiorev.2016.05.024]
54. Greenwood PM, Lambert C, Sunderland T, Parasuraman R. Effects of apolipoprotein E genotype on spatial attention, working memory, and their interaction in healthy, middle-aged adults: results From the National Institute of Mental Health's BIOCARD study. Neuropsychology 2005 Mar;19(2):199-211. [doi: 10.1037/0894-4105.19.2.199] [Medline: 15769204]
55. Kim YS, Park JH, Lee SA. Is a program to improve grocery-shopping skills clinically effective in improving executive function and instrumental activities of daily living of patients with schizophrenia? Asian J Psychiatr 2020 Feb;48:101896. [doi: 10.1016/j.ajp.2019.101896]
56. Mouri N, Ohta R, Sano C. Effects of shopping rehabilitation on older people's daily activities. Int J Environ Res Public Health 2022 Jan 5;19(1):569. [doi: 10.3390/ijerph19010569] [Medline: 35010839]
57. Okahashi S, Seki K, Nagano A, Luo Z, Kojima M, Futaki T. A virtual shopping test for realistic assessment of cognitive function. J Neuroeng Rehabil 2013 Jun 18;10(1):1-13. [doi: 10.1186/1743-0003-10-59] [Medline: 23777412]
58. Barnett MD, Hardesty DR, Griffin RA, Parsons TD. Performance on a virtual environment shopping task and adaptive functioning among older adults. J Clin Exp Neuropsychol 2023 Aug;45(5):464-472. [doi: 10.1080/13803395.2023.2249175] [Medline: 37638858]
59. Ghazal M, Yaghi M, Gad A, et al. AI-powered service robotics for independent shopping experiences by elderly and disabled people. Appl Sci (Basel) 2021;11(19):9007. [doi: 10.3390/app11199007]
60. Merritt BK, Fisher AG. Gender differences in the performance of activities of daily living. Arch Phys Med Rehabil 2003 Dec;84(12):1872-1877. [doi: 10.1016/s0003-9993(03)00483-0] [Medline: 14669197]
61. Sheehan CM, Tucker-Drob EM. Gendered expectations distort male–female differences in instrumental activities of daily living in later adulthood. J Gerontol B Psychol Sci Soc Sci 2019 Apr 12;74(4):715-723. [doi: 10.1093/geronb/gbw209] [Medline: 28158847]

XSL•FO
RenderX

Abbreviations

ASPE HHS: Office of the Assistant Secretary for Planning and Evaluation at the US Department of Health and Human Services
CS-C: correlation stability coefficient
EBIC: Extended Bayesian Information Criterion
FPL: Federal Poverty Level
IADL: instrumental activity of daily living
LASSO: least absolute shrinkage and selection operator
NHATS: National Health and Aging Trends Study

Edited by Y Jiang; submitted 16.10.24; peer-reviewed by D Fong, K Nanjo, Anonymous, Y Lee; revised version received 17.02.25; accepted 19.02.25; published 19.03.25.

Please cite as:
Li J, He R, Hsu EC, Li J
Network Analysis of Key Instrumental Activities of Daily Living and Cognitive Domains for Targeted Intervention in US Older Adults Without Dementia: Cross-Sectional Study
JMIR Aging 2025;8:e67632
URL: https://aging.jmir.org/2025/1/e67632
doi:10.2196/67632

©Jiaying Li, Rendong He, Erh-Chi Hsu, Junxin Li. Originally published in JMIR Aging (https://aging.jmir.org), 19.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Online Group–Based Dual-Task Training to Improve Cognitive Function of Community-Dwelling Older Adults: Randomized Controlled Feasibility Study

Pui Hing Chau[1], PhD; Denise Shuk Ting Cheung[1], PhD; Jojo Yan Yan Kwok[1], PhD; Wai Chi Chan[2], MBChB; Doris Sau Fung Yu[1], PhD

[1]School of Nursing, The University of Hong Kong, Hong Kong, China (Hong Kong)

[2]Department of Psychiatry, The University of Hong Kong, Hong Kong, China (Hong Kong)

Corresponding Author:
Pui Hing Chau, PhD
School of Nursing
The University of Hong Kong
5/F, 3 Sassoon Road
Pokfulam
Hong Kong
China (Hong Kong)
Phone: 852 39176626
Email: phchau@graduate.hku.hk

Abstract

Background: Cognitive training for older adults is crucial before cognitive impairment emerges. During periods of social distancing like the COVID-19 pandemic, cognitive stimuli are lacking. Online dual-task training is proposed as a solution to address these needs.

Objective: We aimed to explore the feasibility, acceptance, and potential effects of online group-based dual-task training as an intervention for enhancing cognitive function among community-dwelling older adults.

Methods: A randomized controlled feasibility study was conducted with 76 participants in Hong Kong, randomly assigned to the intervention and attention control groups in a ratio of 2:1 (n=50, 66% and n=26, 34%, respectively). The intervention group underwent 60-minute online dual-task training sessions twice a week for 12 weeks, incorporating cognitive components (upper limb and finger movement, arithmetic operation, and verbal fluency) and physical components (chair-based exercises) developed through a co-design approach. The attention control group received online health talks. Outcomes related to feasibility and acceptance included class attendance and self-reported satisfaction. Main outcomes related to potential effects included the Memory Inventory in Chinese and the Montreal Cognitive Assessment 5 Minutes (Hong Kong Version) at baseline, 6 weeks (midintervention), 12 weeks (postintervention) and 18 weeks (follow-up). Descriptive statistics and linear mixed effects models were used. Effect size was described with Cohen *d*. Qualitative feedback was collected from 12 informants and analyzed by thematic analysis.

Results: About 72% (36/50) of the participants in the intervention group and 62% (16/26) in the control group attended over 75% of the classes. In total, 44 (88%) participants from the intervention group provided acceptance feedback; 82% (36/44) were satisfied and 84% (37/44) would recommend the training to others. Improvement in the Memory Inventory in Chinese score in the intervention group was observed at midintervention, postintervention, and follow-up, with a medium-to-large effect size (d=0.65, 0.43 and 0.85, respectively). Adjusting for baseline values, the between-group differences in the Montreal Cognitive Assessment 5 Minutes (Hong Kong Version) score attained a small-to-medium effect size at midintervention (d=0.34) and postintervention (d=0.23). Qualitative feedback highlighted the timesaving and convenient aspects of online dual-task training, with participants finding the sessions challenging and enjoyable, and reporting benefits across cognitive, physical, and psychosocial domains. However, a preference for traditional in-person training was noted among the older adults despite the advantages of online training.

Conclusions: Online dual-task training is a feasible intervention accepted by the older adults, with potential benefits in cognitive abilities. Online training may complement in-person sessions. Further investigation in a full-scale randomized controlled trial is warranted to comprehensively explore its effects and address areas for improvement.

XSL•FO
RenderX

Trial Registration: ClinicalTrials.gov NCT05573646; https://clinicaltrials.gov/study/NCT05573646

(JMIR Aging 2025;8:e67267) doi:10.2196/67267

KEYWORDS

cognitive training; dual-task; co-design; online; older adults; Hong Kong

Introduction

Cognitive Decline

The cognitive health of older people is critical to aging in place. Even among healthy older adults, about one-third may experience cognitive decline within 18 months [1], impacting decision-making and daily activities like managing finances, medications, transportation, and meal preparation [2]. Cognitive decline is also linked to decreased physical performance and psychosocial issues, such as loneliness and depression [3]. The health care burden may thus increase in the long run [3].

Existing Dual-Task Interventions

Interventions to enhance cognitive function or to prevent decline are essential for healthy aging. When older adults have limited capacity in performing both functional tasks and cognitive tasks, they tend to prioritize the former over the latter because of the lower chance of being injured [4]. Dual-task training, which involves combining cognitive and physical training components, has been shown to improve physical and cognitive performance, including, but not limited to, gait, balance, and memory and cognitive function [5]. Compared with training with single components, dual-task training tends to have potentially greater benefits [6,7]. Dual-task training has been widely applied among populations with clinical conditions, such as stroke, Parkinson disease, and cognitive impairment. A systematic review of 13 randomized controlled trials (RCTs) on stroke survivors reported that dual-task training could improve walking and balance function [8]. A recent meta-analysis of 14 RCTs estimated a small-to-moderate effect in improving cognitive function among stroke survivors [9]. Meanwhile, a meta-analysis of 17 RCTs on people with Parkinson disease revealed positive effects in physical performance, such as gait and balance, with effect size varying from small to large [10], and a narrative review of 3 studies reported potential improvements in cognitive function among this group, although statistical significance was not always achieved [11]. A 2022 network meta-analysis concluded that dual-task training was promising for addressing the cognitive and motor symptoms of patients with Parkinson disease [12]. The benefits of dual-task training for cognitive function were commonly investigated among cognitively impaired people. A review of 18 RCTs on those with dementia or mild cognitive impairment reported improvements in overall cognitive function, attention, and functional mobility [7]. Another meta-analysis of 21 RCTs reported small-to-moderate effects in global cognitive function, memory, executive function, and attention in cognitively impaired people [13]. A recent meta-analysis of 20 RCTs also reported benefits for global cognition, executive function, and working memory with effect sizes that were moderate, small to moderate, and moderate to large, respectively [14]. There is growing literature on studies targeting cognitively healthy older people. A meta-analysis of 8 controlled trials on cognitively healthy older adults reported improvement in cognitive functions, such as global cognition, working memory, and executive function [15]. The same review also reported no statistically significant difference in efficacy between cognitively healthy older adults and those with mild cognitive impairment [15].

Service Gaps

In practice, motivating healthy older adults to engage in cognitive training can be challenging, especially when they perceive no immediate risk of decline. Presenting interventions as recreational and stimulating can enhance participation. Compared with solely cognitive training, dual-task training may address this need. Meanwhile, when there was social distancing, such as during the COVID-19 pandemic, older adults lost their opportunity to exercise and interact with others. Reductions in physical activity and social interaction not only challenged their physical and psychological health but also challenged their cognitive performance. Although group-based dual-task training could be beneficial, existing dual-task training programs have been conducted face-to-face, which limits their application during times of social distancing, prompting the exploration of online delivery methods. As the participants could join the training in their own homes, the training components had to be specially designed to simultaneously ensure safety, enjoyment, and efficacy. Hence, as an initial step, the feasibility, acceptability, and potential effects of the proposed intervention had to be investigated through a feasibility study.

Aim and Objectives

This study aimed to explore the feasibility, acceptance, and potential effects of online group-based dual-task training as an intervention for enhancing cognitive function among community-dwelling older adults. To accomplish this aim, we had three objectives: (1) to develop the intervention through a co-design approach, (2) to explore its feasibility and acceptance among older adults, and (3) to examine its potential effects in terms of the cognitive, physical, and psychosocial status of the intervention group before and after the intervention and in comparison with the control group. We hypothesized that online based dual-task training would be feasible for healthy older adults.

Methods

Study Design

This was a randomized controlled feasibility study with 2 parallel arms. The trial was registered at ClinicalTrials.gov (NCT05573646).

XSL•FO
RenderX

Ethical Considerations

Ethics approval was obtained from The University of Hong Kong/Hospital Authority Hong Kong West Cluster Institutional Review Board (UW22-038).

Setting

This study was conducted in both online and community settings across Hong Kong. Data collection was conducted at older adults community centers, an older adult educational center, and a university campus, and the intervention was delivered online.

Target Population and Sample

The target population comprised community-dwelling older adults. The inclusion criteria were (1) age ≥65 years, (2) no communication problems, and (3) ability to use an online meeting platform. The exclusion criteria were (1) contraindications to chair-based exercises and (2) engagement in any kind of cognitive training 3 months before the study or during the study period.

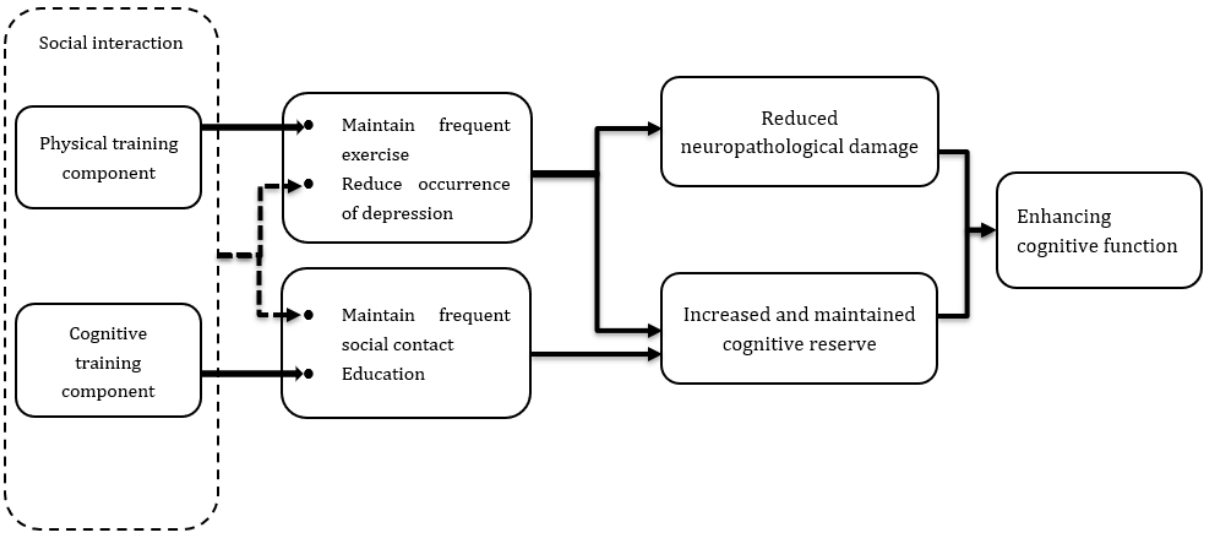
According to the literature, for a main trial designed with 80% power and 2-sided 5% significance to detect a small-to-medium effect (Cohen *d* between 0.1 and 0.3), a sample size of at least 20 per treatment arm is needed for a pilot study [16]. Meanwhile, to allow the sample to detect a medium within-group effect size in the intervention group (d=0.5) with 80% power and 2-sided 5% significance, a sample size of 34 was required for the intervention group. Taking both scenarios and 20% attrition into account, 50 participants in the intervention group and 25 participants in the control group were required.

Intervention: Online Group-Based Dual-Task Training

Participants in the intervention group received a 1-hour online group-based dual-task training session twice a week for 12 weeks, resulting in a total intervention time of 24 hours. Such frequency and duration were within the range of 4 to 25 weeks and 30 to 240 minutes per week, as used in previous dual-task training programs for healthy older adults [6].

Figure 1 shows the conceptual framework underpinning the development of the intervention based on the 2020 Report of the Lancet Commission on dementia prevention, intervention, and care [17]. The framework proposed that modifiable risk factors prevent dementia through reduction of neuropathological damage and increase and maintenance of a cognitive reserve. Among these factors, maintenance of frequent exercise and reduction of depression act on both paths, while maintenance of frequent social contact and education act on the cognitive reserve path. In our intervention, the physical training components helped to maintain frequent exercise (at least twice weekly) and the cognitive training component mimicked the education effect. The social interaction via group Zoom (Zoom Communications, Inc) classes maintained frequent social contact for at least twice a week and thus could help to relieve depression symptoms.

Figure 1. Conceptual framework.



In the development of the cognitive training component, various factors were considered, including the lower education level of the local older adults, the online delivery format of the intervention, the cultural appeal to the local older population, and the effectiveness and safety of the training. A diverse set of training components were incorporated, such as upper limb and finger movements, arithmetic operations, and verbal fluency exercises, based on their proven efficacy in previous dual-task training studies [18,19]. A co-design approach was adopted to develop the dual-task training components. In total, 9 participants from the target population were invited to a trial run to provide feedback on their interests, challenges, and suggestions for tasks they preferred. The suggested tasks were then reviewed by the project team to assess their intensity and potential impact on training outcomes. Considering what would be acceptable and effective, 18 upper limb and finger movement games, 22 arithmetic operation games, and 6 verbal fluency training components were finalized (Textbox 1). During the 1-hour training session, participants engaged in different tasks according to a predefined schedule (Table 1). Specifically, participants spent 12 minutes on upper limb and finger movement exercises, 15 minutes on arithmetic operations, and 15 minutes on verbal fluency tasks. Each game started at the easiest level, with the facilitator adjusting the cognitive demand

XSL•FO
RenderX

as participants progressed. For each game, the facilitator would start with the easiest level first. To sustain motivation and enjoyment, the difficulty level of each component increased when participants demonstrated mastery, typically indicated by 3 consecutive rounds without losing a game.

Textbox 1. Examples of cognitive training components.

Upper limb and finger movements

- Goal-directed finger movements, such as counting and reverse counting with the fingers and making motions in different directions with different hands
- Forming patterns in the air

Arithmetic operation

- Counting
- Reverse counting
- Counting or reverse counting with subtraction or addition
- Clapping at specific numbers or multiples

Verbal fluency

- "Jie Long" (a well-known Chinese word chain game)
- Naming games (particularly on Chinese cultural concepts, such as Chinese dim sum and greetings in the Chinese new year)
- Memory games (participants needed to recap the previous terms presented by others and then add their own)

Table 1. Dual-task training schedule for each 60-minute session.

Time (minutes)	Dual-task training	
	Physical components (chair-based)	Cognitive components
5	Warm-up exercise	Safety reminder or casual chit-chat
3	Toes raises	Upper limb and finger movement 1
3	Heel raises	Upper limb and finger movement 2
3	Toes raises	Upper limb and finger movement 3
3	Heel raises	Upper limb and finger movement 4
3	Break	Break
3	Stepping	Arithmetic operation 1
3	Rest	Arithmetic operation 2
3	Stepping	Arithmetic operation 3
3	Rest	Arithmetic operation 4
3	Stepping	Arithmetic operation 5
3	Break	Break
3	Stepping	Verbal fluency 1
3	Clam exercise	Verbal fluency 2
3	Stepping	Verbal fluency 3
3	Clam exercise	Verbal fluency 4
3	Stepping	Verbal fluency 5
3	Clam exercise	Verbal fluency 6
4	Cool-down exercise	Safety reminder

For the physical training component, we selected chair-based exercises, which were chosen due to their suitability for the limited space available in most local households. Chair-based exercises offer various benefits, such as improving mood and well-being, enhancing certain activities of daily living, promoting social interaction, and increasing muscle strength [20]. Alongside the cognitive training games, participants engaged in exercises like toe or heel raises on alternate sides,

XSL•FO
RenderX

stepping, and clam exercises (Table 1). The exercises were performed at a low frequency, such as 60 steps per minute. To ensure safety, participants were permitted to use their hands for balance support during stepping or clam exercises. Therefore, when upper limb and finger movement tasks were involved, only toe or heels raise were performed.

A group-based training approach was adopted as group-based cognitive training might yield greater interaction and cognitive benefits than individual-based training [21]. Group-based exercise was also reported as having the additional advantage of creating a social support environment that could enhance psychosocial well-being [22]. Moreover, group-based interventions tend to have higher adherence rates than general exercise programs due to the motivational support provided by group dynamics [23]. To encourage interaction among participants and enable close monitoring by the facilitator, a group size of less than 10 people was used.

Zoom was selected as the training platform due to its widespread use and familiarity among local older adults. Throughout the session, participants were instructed to sit on a sturdy chair with a back rest in front of a desktop computer, laptop, tablet, or mobile phone securely placed on a table. Mobile device users were provided with stands for better posture (all participants received the stand as a souvenir). To allow enough space for movement, the chair and table needed to be positioned 60 cm apart. Participants were encouraged to turn on their cameras during the session to facilitate interaction with one another. They were also asked to keep their audio on to actively participate in the cognitive tasks by verbalizing their answers.

The sessions were led by a trained facilitator. The facilitator did not need to be health professional, but they were required to undergo necessary training to ensure they had the knowledge and skills to effectively work with older adults [20]. A written manual with all the rules and regulations of the training was provided. During the initial sessions of the intervention, the project lead provided on-site supervision of the facilitator to ensure that the intervention was delivered according to the protocol. In addition, random spot checks were conducted by the project lead to ensure strict adherence to the protocol.

Attention Control: Interactive Health Talks

In line with recommendations for attention control groups [24], participants in the control group underwent 8 one-hour online group-based interactive health talks on food label knowledge and its practical applications. These sessions were led by a facilitator with nutrition training.

Outcome Measures

The primary outcomes were related to feasibility and acceptance. Feasibility was reflected by class attendance. Reasons for absence were documented when available. There was no consensus on the definition of high attendance [25]. We took 75% attendance as high attendance, as a review reported 3 of 11 studies on exercise interventions adopted this definition [25]. The proportion of participants with high attendance was calculated using the number of participants randomized to the group as denominator. A review of group exercise interventions for older adults reported that in 4 of 6 studies, 65% to 67% of participants adhered to the program [22]. Hence, our study considered over 65% of participants completing at least 75% of the classes as feasible.

Acceptance was reflected by 5-point Likert scale questions on the level of satisfaction with the intervention (“Are you satisfied with this training?”) and the level of likeliness of recommending the intervention to other people (“Would you recommend this training to friends and family”), with answers collected at the end of the intervention via instant text messaging. The proportion of respondents providing a positive response was calculated. As a conservative estimate, another proportion was calculated by assuming those who did not provide feedback as not being satisfied and not likely to recommend the intervention, with the number of participants randomized to the group as the denominator.

The secondary outcomes were related to potential effects, which included subjective memory complaints, cognitive status, working memory, executive function, physical function, instrumental activities of daily living, happiness, and social networks.

Subjective memory complaints were assessed using the Memory Inventory in Chinese (MIC) [26]. This scale consists of 27 items on memory concerns related to daily activities in the past month. Responses are rated on a scale from 0 (none) to 4 (once or more per day or continuously), with scores totaling between 0 and 108. A higher score indicates a greater frequency of memory concerns.

Cognitive status was assessed using the validated Montreal Cognitive Assessment 5 Minutes (Hong Kong Version) (HK-MoCA 5-Min), which covers 4 domains, namely, attention, executive functions or language, orientation, and memory [27]. The total score ranges from 0 to 30, with a higher score indicating better cognitive status.

Working memory was assessed by Digit Span Test [28]. Participants were required to repeat progressively longer series of digits. The test began with 2 digits and increased in length, with a maximum of 8 digits for the forward test and 7 digits for the backward test. The participant’s ability to recall the longest digit series in forward and backward order was recorded as the forward and backward scores, respectively. A higher score indicated a better working memory capacity.

Executive function was assessed by the Chinese version of the Victoria Stroop Test [29]. The test consisted of 3 parts, each presenting different stimuli: colored dots, common words unrelated to color, and color words. Participants were required to name the color in which the stimuli were printed. Inference scores were calculated based on the difference in completion time between the word or color test and the dot test, with a longer completion time indicating a poorer condition.

Lower-limb muscle strength was assessed by the 5-time chair stand task [30]. Participants were timed on how long it took them to complete the task, with a longer time indicating poorer performance. If participants were unable to finish the task within 1 minute, the test was stopped, and a time of 60 seconds was imputed for analysis purposes.

XSL•FO
RenderX

Instrumental activities of daily living were assessed by Lawton's 8 selected tasks, which include telephone use, shopping, food preparation, housekeeping, laundry, transportation, medicine use, and handling finances [31]. Participants rated their level of dependence in performing each task on a 5-point Likert scale. The score was the sum of all items, ranging from 0 to 8, with a higher score indicating greater independence.

Happiness was assessed by the Chinese version of the validated 4-item Subjective Happiness Scale [32]. Participants rated their responses on a 7-point Likert scale. The score was the average of the 4 items, ranging from 1 to 7, with a higher score indicating a higher level of happiness.

Depressive symptoms were assessed by the Patient Health Questionnaire-9 (PHQ-9) [33]. PHQ-9 was validated as a reliable tool for evaluating depression among older adults. The scale ranged from 0 to 27, with a higher score indicating a higher level of depressive symptoms.

Social networks were assessed by the Chinese version of the 6-item Lubben Social Network Scale [34]. Participants rated their responses on a 5-point Likert scale. The total score ranged from 0 to 30, with a higher score indicating a stronger social network.

All effect-related outcomes were assessed at recruitment (T0, baseline), 6 weeks after baseline (T1, midintervention assessment), 12 weeks after baseline (T2, postintervention assessment), and 18 weeks after baseline (T3, follow-up assessment). Demographic, medical and lifestyle information, such as age, gender, education level, exercise habits (measured by the International Physical Activity Questionnaire-Short Form), frailty status (measured by the FRAIL scale, which consists of 5 items, namely, fatigue, resistance, ambulation, illnesses, and loss of weight) and comorbidity (coexistence of 2 or more chronic illnesses) were collected at baseline. For participants who could not undergo face-to-face assessment for reasons such as quarantine or sickness, the 5-time chair stand task was not administered and was regarded as missing data.

Qualitative feedback was collected from a purposive sample at the end of the follow-up period through phone interviews. A semistructured interview guide was used to explore reasons for participating in the intervention, perceived difficulties encountered, logistical aspects of the intervention, and strengths and limitations perceived by participants. Prompts were provided as appropriate. Participants who exhibited varying degrees of change in the MIC score were invited to ensure diverse perspectives. Data saturation was achieved with 12 informants.

Procedure

Participants were recruited through social media and nongovernmental organizations in Hong Kong. They were randomly assigned to the intervention or control groups at a ratio of 2:1. As a strategy to enhance recruitment, a ratio of 2:1 was proposed for the number of people in the intervention and control groups, such that less participants would be disappointed with the allocation to the control group. A face-to-face orientation session was conducted to guide participants on using the Zoom platform and provide them with mobile device stands. Furthermore, for intervention group participants, demonstrations of the chair-based exercises were provided as they were unable to observe lower-limb movements on Zoom.

Randomization and Concealment

Block randomization with varying block sizes was used to generate the allocation sequence. An independent research assistant generated the sequence using an online platform and prepared sequentially numbered opaque sealed envelopes. These envelopes were opened in front of the participants immediately after obtaining informed consent.

Blinding

This was a single-blinded study as participants and interventionists could not be blinded. Trained research assistants involved in assessing outcomes related to potential effects were blinded to the allocation.

Data Analysis

The participants' characteristics were summarized using descriptive statistics. Attendance statistics, regarding the percentage of scheduled classes attended, were calculated. The percentage of participants satisfied with the intervention and those who would recommend it to others were also calculated. Adopting intent-to-treat, linear mixed effects models were used to analyze effect-related outcomes, adjusting for sociodemographic characteristics, medical history, and lifestyle patterns. A significance level of 5% was used. However, given the small sample size, focus was placed on effect size with the corresponding 95% CI. A Cohen *d* of 0.2 was considered clinically meaningful [35]. Missing outcome measures were not imputed as the linear mixed effects models could handle missing data. Statistical analyses were performed using SPSS version 28.

As for the qualitative interview, the audio recordings were transcribed verbatim for thematic analysis [36]. First, 2 researchers read the transcripts several times and conducted the systematic coding independently. They proceeded to generate the initial codes and the recurrent pattern within the data and identify and name the themes and subthemes according to the underlying meaning. Then, a third researcher checked for the consensus on the themes and subthemes and discussed with the team accordingly. They also worked together to identify the common threads that extend across the interviews of the participants. Analyses were performed manually without the aid of software.

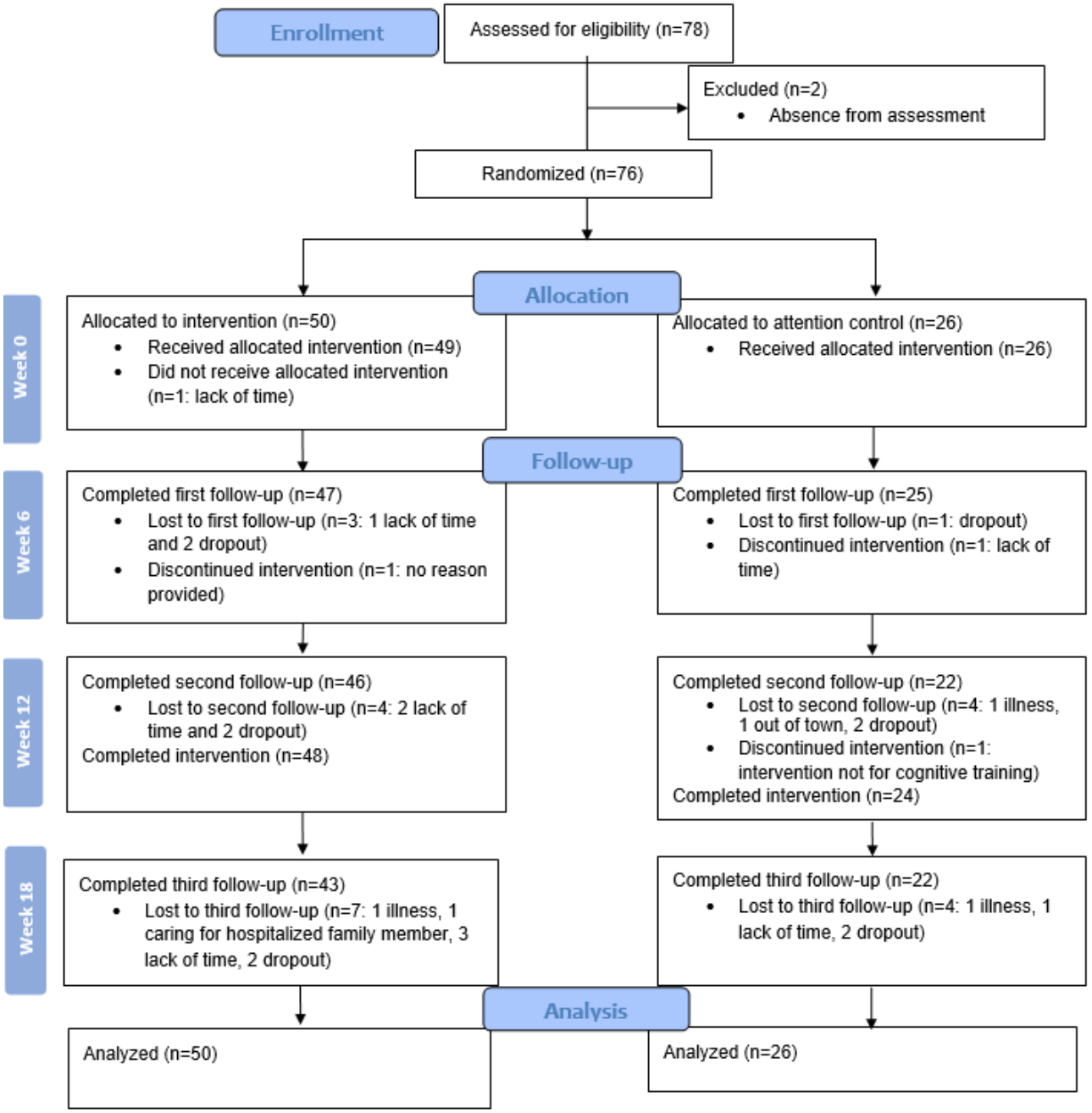
Results

Overview

In total, 76 participants were recruited from October 2022 to May 2023 (Figure 2). To manage class sizes, the 50 participants in the intervention group were divided into 8 subgroups, each with 24 classes, and the 26 participants in the attention control group were divided into 2 subgroups, each with 8 classes. Overall, 48 (96%) participants from the intervention group and 24 (92%) from the control group completed the assigned intervention.

XSL•FO
RenderX

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) diagram.



The mean ages of the participants in the intervention and the control groups were 71.5 (SD 5.2) and 71.1 (SD 6.0) years, respectively. There were 64% (32/50) female participants in the intervention group and 88% (23/26) in the control group. One-fifth (15/76, 20%) of the participants had a primary level of education or below. About 44% (22/50) of the participants in the intervention group and 54% (14/26) in the control group had multimorbidity. Most participants in both groups (30/50, 60% and 16/26, 61%) had low levels of physical activity. Prefrailty was observed in about 38% (19/50) of intervention-group participants and 50% (13/26) of control-group participants. Table 2 shows the characteristics of the participants.

XSL•FO
RenderX

Table 2. Baseline characteristics of participants (N=76).

Characteristics	Intervention (n=50)	Control (n=26)
Age (y), mean (SD)	71.5 (5.2)	71.1 (6.0)
Sex, n (%)		
Male	18 (36)	3 (11)
Female	32 (64)	23 (89)
Education level, n (%)		
Primary or below	10 (20)	5 (19)
Secondary	27 (54)	14 (54)
Tertiary or above	13 (26)	7 (27)
Multimorbidity[a], n (%)		
No	28 (56)	12 (46)
Yes	22 (44)	14 (54)
Physical activity level (measured by IPAQ[b]), n (%)		
Low	30 (60)	16 (61)
Moderate	15 (30)	9 (35)
High	5 (10)	1 (4)
Frailty status (measured by FRAIL[c] scale), n (%)		
Robust	31 (62)	13 (50)
Prefrail	19 (38)	13 (50)

[a]Multimorbidity is defined as the co-existence of at least 2 of the following long-term health conditions: hypertension, heart disease, high cholesterol, cancer, osteoporosis, and other chronic illnesses, as reported by the participants.

[b]IPAQ: International Physical Activity Questionnaire.

[c]FRAIL: Fatigue, Resistance, Ambulation, Illnesses, and Loss of weight.

Feasibility

In the intervention group, 36 (72%) participants attended over 75% of the 24 classes. For the attention control group, 16 participants (61%) attended over 75% of the 8 classes. Common reasons for absence in both groups included time clashes with medical appointments, illness, or other prior commitments.

Acceptability

Among 44 participants who provided acceptance feedback on the intervention, 82% (36/44) expressed being very satisfied or satisfied with the intervention, and 84% (37/44) indicated they would be very likely or likely to recommend the intervention to others. Assuming that those who did not provide feedback were not satisfied and were not likely to recommend the intervention, the satisfaction rate was 72% (36/50) and the recommendation rate was 74% (37/50).

Potential Effects

Regarding the effect-related outcomes, owing to the pilot nature of the trial, statistical significance was not the primary focus, instead, the effect size was noted. Within-group changes were explored and are presented in Multimedia Appendix 1. For the intervention group, the effect size of the MIC improvements at midintervention, postintervention, and follow-up compared to baseline were 0.65, 0.43 and 0.85, whereas the estimates for the control group were 0.51, 0.37 and 0.78, respectively. Similarly, both groups experienced an increase in HK-MoCA 5-Min score with moderate to large effect sizes. However, changes in working memory, as assessed by the Digit Span Test, did not consistently show the same direction in both groups. The intervention group showed a moderate improvement in executive function, as shown by the inference score for color from the Victoria Stroop Test, while the control group showed a small-to-large improvement in the inference score for words. Improvement was not observed for other outcomes in both groups.

Between-group differences, adjusting for the baseline differences, were explored and are presented in Multimedia Appendix 1. The intervention group did not outperform the control group in the MIC, but it showed better performance in the HK-MoCA 5-Min at the midintervention and the postintervention assessments (effect size 0.34 and 0.23, respectively). Meanwhile, a medium effect size was observed for between-group difference in the Digit Span Test at the follow-up. However, the findings for the Victoria Stroop Test were inconsistent. Details on other outcomes are presented in Multimedia Appendix 1.

Comparing the effect-related outcomes between participants with high attendance (≥75% of class) and low attendance in the intervention group (Multimedia Appendix 1), those with high attendance had fewer subjective memory complaints than those

XSL•FO
RenderX

with low attendance at midintervention and postintervention, with effect sizes of 1.18 and 1.53, respectively. At the follow-up assessment, the between-group difference was observed to be in favor of the high-attendance group in the Digit Span Test, inference score for color, and PHQ-9 score.

Qualitative Feedback

Table 3 shows the demographic characteristics of the 12 informants. Table 4 shows the 5 identified themes and the corresponding subthemes derived from the qualitative feedback, along with the significant quotes. The themes are summarized as presented in Textbox 2.

Table 3. Characteristics of the 12 informants providing qualitative feedback.

Informant	Sex	Age group (y)	Education level
A	Female	65-74	Primary
B	Female	65-74	Secondary
C	Male	≥75	Primary
D	Female	65-74	Secondary
E	Female	65-74	Secondary
F	Female	65-74	Tertiary
G	Female	≥75	Secondary
H	Female	65-74	Secondary
I	Female	65-74	Secondary
J	Male	65-74	Tertiary
K	Male	65-74	Primary
L	Male	65-74	Tertiary

XSL•FO
RenderX

Table 4. Five identified themes and the corresponding subthemes from the qualitative feedback and the significant quotes.

Themes and subthemes	Significant quotes by informants
Perceived strength of the training (before participation)	
Arousing curiosity and motivation to join	• "I heard an introduction at the elderly community center, and I want to learn more." [Informant A] • "Curious about online dual task training." [Informant B]
Giving them hope to slow down cognitive decline	• "I am taking precautions because I don't know if I will have cognitive impairment in the future. I just want to prevent it. Early prevention is better." [Informant D] • "I want to maintain, not improve, cognitive ability." [Informant L]
Offering opportunity for social interaction	• "I want to make new friends." [Informant C]
Save travel time	• "No need to spend time going back and forth; you can stay at home. Nowadays, we mostly go to centers to learn things. In-person is much better, but it takes more time to get there and back." [Informant E] • "But overall, I think online is good for participants, as it saves travel time." [Informant J]
Positive experience from the training	
Comprehensive training	• "I think the course is quite good and comprehensive; it does not just focus on those with cognitive impairment." [Informant I]
Challenging training	• "There is some difficulty; sometimes I forget, like moving a hand or getting distracted and forgetting to do the stepping. You need to stay focused to complete it." [Informant B] • "At the beginning, I feel a bit frustrated when playing. However, it has its benefits. If everything were easy, there would be no need for training." [Informant G] • "If you have to move both your foot and hand, you often end up neglecting one. Sometimes, you also have to keep track of numbers, so you need to pay attention to what the other participants are saying." [Informant K]
Engaging facilitator	• "The host guided us with patience." [Informant A] • "Overall, I think the team is very dedicated. Sometimes when we can't hear, she [the facilitator] reminds us. Or when we can't think of an answer, she gives us hints at the right time. I think the whole process is very smooth." [Informant I]
Peer support and encouragement during the training	• "We encourage each other during the training." [Informant L]
Enjoy the convenience at home environment	• "I can start the training right after breakfast. When we have breaks, we can go to the washroom and drink water at home. After the session ends, we can prepare lunch ourselves." [Informant A]
Challenges encountered with online training	
Constrained supervision and intensity of training	• "At least with Zoom, the camera cannot capture my feet, so you cannot tell if participants are doing stepping. I think in-person is ultimately more useful because you can see everyone's movements." [Informant B] • "Because in-person, you can immediately see everyone's performance." [Informant D]
Harder to follow the sequence if the training involved taking turns	• "Because it is online and most participants are women, it is hard for me to know that it is my turn unless there is a man preceding my turn. But after a few sessions, everyone starts to get familiar with each other, and we build up rapport about the sequence." [Informant J]
Less concentrated and less cordial	• "During training, some might not be able to concentrate, for example, if there are things happening at home." [Informant A] • "It's not as easy to get distracted if it is in-person." [Informant D]
Class atmosphere not comparable to in-person class	• "But in-person is more enjoyable, with more group interactions, which makes it more fun." [Informant G]
Technical issues encountered	• "The display is limited in size, and there might be some lag time." [Informant J] • "The problem is that some resource-limited families don't have Wi-Fi at home, so they have to go to the center to access it, which is less convenient." [Informant K]

XSL•FO
RenderX

Themes and subthemes	Significant quotes by informants
Perceived benefits of the training	
Improved cognitive ability (including attention and concentration)	• "Initially, I felt very unfamiliar with it, but after a few more sessions, I absorbed new knowledge and it made my mind more agile." [Informant G] • "I think I can use my brain more, as I usually don't use it much." [Informant H] • "My memory has been maintained without much decline." [Informant L]
Improved coordination and physical flexibility	• "After participating, my knowledge has broadened, and both my hands and mind have become more agile." [Informant C] • "I find that after training, the coordination of my hands and feet has become a bit more agile compared to before." [Informant L]
Improved communication skills	• "It can train an older person to handle different communication demands." [Informant J]
Promotion of positive mood and social interactions	• "Someone accompanies you. I also find it fun and enjoyable." [Informant F]
Learn new skills for self-practice in future	• "I have not experienced this kind of hand-eye coordination, but now I do. Sometimes, when I'm alone, I practise the movements." [Informant E] • "I'll go back and practise it myself." [Informant K]
Increased self-efficacy and self-confidence	• "My brain is now willing to exercise and let me do calculations." [Informant F] • "I actually feel quite happy and it gives a sense of accomplishment because you can achieve your goals and meet the training requirements, so there is a sense of success." [Informant J]
Self-explanation of less benefit from "cheating"	• "I do not feel like I have gained much. Maybe I prepare the answers in advance. For example, knowing that the next lesson will involve 'clapping on 7,' I will prepare the answers beforehand. But if it is in-person, it is harder to cheat." [Informant I]
Suggestions for improvements	
Continue the training (beyond 12 weeks)	• "I think I need to keep doing it consistently to see improvement. But I understand it is difficult if resources are lacking. Once the training stops, like now, I have forgotten a lot." [Informant K]
Booster training	• "Suggest to add online virtual courses. Each session would be half an hour, allowing students to choose their own study times." [Informant K] • "If possible, it would be best to have a booster session once a month to refresh what we have learned." [Informant L]
Consider in-person mode (but reduced frequency to accommodate time and venue constraints)	• "I think in-person sessions are ultimately more useful because you can see everyone's movements. In-person sessions could be held once a week, not too frequently, since not everyone can attend twice a week, and the center might not be able to provide the space for two days of training." [Informant B]

Textbox 2. Summary of the themes and subthemes derived from qualitative feedback.

- Perceived strength of the training (before participation): informants found dual-task training to be a novel approach, which interested them and motivated them to participate. They expected the training would help to slow down cognitive decline. The online format of the training saved them travel time but could still offer social interaction.
- Positive experience from the training: informants described the training comprehensive and challenging. The engaging facilitator increased their adherence. The group training enabled peer support, enhancing their training experience. Informants also found having the training at home very convenient.
- Challenges encountered with online training: supervision and the intensity of training was limited by the online format. Some informants found it harder to follow the sequence if the training involved taking turns. Some found the home environment made them less concentrated and the class atmosphere was not comparable to in-person class. The quality of the experience was also influenced by the display size of their device, Wi-Fi speed, and more importantly accessibility to a stable internet connection.
- Perceived benefits after the training: informants perceived benefits in cognitive ability, coordination, physical flexibility, communication skills, positive moods, and social interactions. Some valued acquiring new skills that could enable future self-practice. Some mentioned increased self-efficacy and self-confidence. Meanwhile, one mentioned the use of a cheat sheet as related to the negligible improvement.
- Suggestions for improvements: informants expressed a desire to continue the training by either extending the training period or offering regular booster sessions. However, informants generally preferred in-person training, although they acknowledged limitations regarding time and venue.

XSL•FO
RenderX

Discussion

Principal Findings

This study demonstrated feasibility and explored potential effects of an online dual-task training for community-dwelling older adults. The innovative aspects of the online dual-task training included (1) the delivery mode, allowing for uninterrupted training even during a pandemic, (2) chair-based exercises suitable for home-based practice, (3) cognitive training tailored to the needs of Chinese older adults with lower education levels, and (4) a co-design approach that considered the preferences and interests of the participants when designing the training components.

This study was launched at the end of 2022 when the COVID-19 pandemic was still ongoing. While there was no citywide quarantine policy in place, there was still a home quarantine policy for those who were infected and their close contacts. Older adults tended to limit their outings to reduce the risk of infection, as well as chance of becoming close contacts. The older population had experienced a lack of activities for a prolonged period, potentially leading to a faster decline in cognitive performance. The group-based online dual-task training addressed this gap by offering training that the participants could participate in at home.

The findings of our pilot study showed the intervention was highly feasible, as 72% (36/50) of the participants in the intervention group attended over three-quarters of the classes, which was above the target of 65%. The intervention was also well received by the participants; 82% (36/44) of the participants were satisfied with the intervention and 84% (37/44) would recommend the intervention to others. Even adopting the most conservative approach, such figures remained high at 72% (36/50) and 74 (37/50), respectively. The dual-task training successfully attracted some relatively cognitively healthy participants to join and could be continued irrespective of the pandemic situation. This initiative addressed the existing gap in services by providing online cognitive training for community-dwelling older adults.

Within-group improvement was observed in subjective memory complaints and cognitive function in the intervention group. Such improvements became evident at midintervention and persisted until the final follow-up, showing a medium-to-large effect size. Despite the qualitative feedback indicating self-perceived benefits, other outcomes did not consistently favor the intervention. Controlled for the baseline outcomes, small-to-medium effects ranging from 0.23 to 0.34 were observed in cognitive function for the between-group difference in midintervention and postintervention assessments. These findings are consistent with previous literature reporting small effects of 0.22 to 0.29 for dual-task training over control groups [37,38]. Nevertheless, inconsistent results were reported for other outcomes. Moreover, the fluctuation in the magnitude of the improvement was unexpected. Although possible self-practicing of the dual-task training after the intervention period might explain the increase in within-group improvement at the follow-up, the drop at the postintervention assessment was unexpected. There were factors that might have affected the evaluation. For example, engagement in the environment might influence achievement in relation to the effects of cognitive training [39], but the social context of the participants was not assessed. Another possibility might be COVID-19 infection, which might influence the performance of the participants not only during the infection period but afterward [40]. Moreover, some assessments could not be performed face-to-face but were conducted via Zoom; the different modes of data collection might have also influenced their performance. After the pandemic, a proper RCT should be conducted.

Meanwhile, within-group improvement was observed for the attention control group. The attention control condition consisted of a series of interactive health talks focusing on nutrition and food labeling. Participants memorized the contents delivered in each talk so they could answer the instructors' questions. Moreover, some participants had to learn to operate the Zoom platform as it was new to them. The learning process might also be beneficial to the participants, as some calculation was involved in comparing food labels using different units. In future trials, usual care control may be adopted.

Qualitative feedback highlighted the strength of the training regarding convenience, innovation, and comprehensiveness. In contrast, participants expressed a preference for in-person interactions. Hence, online training could be considered as a potential alternative or complement to traditional in-person training. Participants engaged in the training before receiving a cognitive impairment diagnosis, indicating that dual-task training could motivate early engagement in cognitive training. The facilitator leading the training does not necessarily be a health professional [20], and it is the engaging nature of the training, as revealed from the feedback, that the participants highly valued. This makes implementation of the intervention more feasible. Apart from the cognitive training, the participants also treasured social interactions and peer support during the class, and these further support the need to organize group-based training.

Implications

Dual-task training offers sufficient challenges to community-dwelling older adults; they found it interesting and reported that it boosted their motivation to participate in cognitive training. The contents of the training were comprehensive, covering motion, arithmetic operations, and verbal fluency. The training was conducted in a stimulating and recreational atmosphere, which was well received by the participants.

Online training offers the advantage of saving time and being convenient. However, it is not intended to replace traditional in-person training, as older adults generally prefer face-to-face interactions. Nonetheless, online training can be viewed as an alternative or complement to in-person sessions. Participants highlighted that organizing training twice a week in-person would be challenging, indicating a potential solution of combining one day of in-person training with one day of online training to maintain a twice-weekly training frequency. As older adults become more accustomed to online training, it can serve as a viable option during situations like a pandemic, enabling

XSL•FO
RenderX

training to continue even when older adult community centers are closed or during citywide quarantine measures.

When considering the implementation of in-person or hybrid mode training, it is essential to conduct research studies to assess their effectiveness. If in-person dual-task training is to be provided, incorporating more physical components could be beneficial. Similarly, any modified training approaches should undergo evaluation before being integrated into routine services. Future research studies could involve including a usual care control group or a waitlist control group to accurately determine the true impact of the training, building upon the improvement areas identified in this study. For instance, potential strategies for future interventions may involve implementing a 6-week intervention period, as some participants in the outcome evaluations exhibited positive effects as early as 6 weeks after the intervention; combining both in-person and online training modalities; and incorporating booster sessions. The exclusion criterion regarding engagement in any form of cognitive training 3 months before or during the study period may pose challenges once activities return to normal after the pandemic. Therefore, future trials should take a pragmatic approach, ensuring that any benefits derived from the intervention are in addition to participants' existing daily lifestyle routines.

Strengths and Limitations

The strength of this study included the integration of both quantitative and qualitative outcomes, which would reflect the feasibility and perceived benefits more holistically. The multiple outcome domains also helped to examine the potential effects more comprehensively. The multiple assessment points (4 time points) would facilitate the exploration of potential effects relative to the dosage of the intervention. The innovation components of the interventions were not restricted to the local context but could largely be applied to other regions.

Nevertheless, there were some limitations. The use of an attention control group might have limited the examination of potential effects from the intervention. Moreover, COVID-19 infection among the participants or their family members and the overall social environment during the study period were not documented, limiting the ability to control for the influences from these factors on the study outcomes. The cognitive domains assessed were limited to attention, verbal fluency, orientation, working memory, and executive function. Other cognitive domains, such as perceptual-motor control and social cognition, that might benefit from the intervention were not evaluated. User satisfaction was based on a self-developed question, and a validated scale could be considered in future. Owing to the nature of the feasibility study, the sample size was small, and a proper trial is needed in the future involving larger sample size.

Conclusions

Online dual-task training was shown to be a feasible intervention and was well received by older adults during the pandemic. Within-group improvement was observed in subjective memory complaints and cognitive function, with a medium-to-large effect size in the intervention group. Further investigation in a full-scale RCT is required to fully explore the intervention's potential effects.

Acknowledgments

This project was supported by the Health and Medical Research Fund (19200991) under the Food and Health Bureau of the government of the Hong Kong Special Administrative Region. The authors would like to thank the New Territories West Elder Academies Cluster and Hong Kong Sheng Kung Hui Welfare Council for their support and the older adults for their participation.

Authors' Contributions

All authors were responsible for conceptualization, funding acquisition, methodology, and writing−review and editing. PHC was responsible for formal analysis, project administration, and writing−original draft.

Conflicts of Interest

PHC is an Associate Editor for JMIR Aging.

Multimedia Appendix 1
Within-group and between-group differences in effect-related outcomes based on mixed effects model.
[PDF File (Adobe PDF File), 140 KB - aging_v8i1e67267_app1.pdf]

Multimedia Appendix 2
CONSORT eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 18050 KB - aging_v8i1e67267_app2.pdf]

References

1. Woodard JL, Seidenberg M, Nielson KA, Smith JC, Antuono P, Durgerian S, et al. Prediction of cognitive decline in healthy older adults using fMRI. J Alzheimers Dis 2010;21(3):871-885 [FREE Full text] [doi: 10.3233/JAD-2010-091693] [Medline: 20634590]

XSL•FO
RenderX

2. Salthouse T. Consequences of age-related cognitive declines. Annu Rev Psychol 2012;63:201-226 [FREE Full text] [doi: 10.1146/annurev-psych-120710-100328] [Medline: 21740223]
3. Comijs HC, Dik MG, Aartsen MJ, Deeg DJ, Jonker C. The impact of change in cognitive functioning and cognitive decline on disability, well-being, and the use of healthcare services in older persons. Results of Longitudinal Aging Study Amsterdam. Dement Geriatr Cogn Disord 2005;19(5-6):316-323. [doi: 10.1159/000084557] [Medline: 15785032]
4. Li KZ, Lindenberger U, Freund AM, Baltes PB. Walking while memorizing: age-related differences in compensatory behavior. Psychol Sci 2001 May;12(3):230-237. [doi: 10.1111/1467-9280.00341] [Medline: 11437306]
5. Pereira Oliva HN, Mansur Machado FS, Rodrigues VD, Leão LL, Monteiro-Júnior RS. The effect of dual-task training on cognition of people with different clinical conditions: an overview of systematic reviews. IBRO Rep 2020 Dec;9:24-31 [FREE Full text] [doi: 10.1016/j.ibror.2020.06.005] [Medline: 33336101]
6. Lauenroth A, Ioannidis AE, Teichmann B. Influence of combined physical and cognitive training on cognition: a systematic review. BMC Geriatr 2016 Jul 18;16:141 [FREE Full text] [doi: 10.1186/s12877-016-0315-1] [Medline: 27431673]
7. Yu D, Li X, He S, Zhu H, Lam FM, Pang MY. The effect of dual-task training on cognitive ability, physical function, and dual-task performance in people with dementia or mild cognitive impairment: a systematic review and meta-analysis. Clin Rehabil 2024 Apr;38(4):443-456. [doi: 10.1177/02692155231216621] [Medline: 38009067]
8. He Y, Yang L, Zhou J, Yao L, Pang MY. Dual-task training effects on motor and cognitive functional abilities in individuals with stroke: a systematic review. Clin Rehabil 2018 Jul 23;32(7):865-877. [doi: 10.1177/0269215518758482] [Medline: 29473480]
9. Feng QM, Chen JW. Effects of exercise combined with cognitive dual-task training on cognitive function and sustained effects in stroke patients: a systematic review and meta-analysis. Disabil Rehabil 2024 Dec 06:1-10. [doi: 10.1080/09638288.2024.2435523] [Medline: 39641346]
10. García-López H, de Los Ángeles Castillo-Pintor M, Castro-Sánchez AM, Lara-Palomo IC, Obrero-Gaitán E, Cortés-Pérez I. Efficacy of dual-task training in patients with Parkinson's disease: a systematic review with meta-analysis. Mov Disord Clin Pract 2023 Sep;10(9):1268-1284 [FREE Full text] [doi: 10.1002/mdc3.13823] [Medline: 37772294]
11. Xiao Y, Yang T, Shang H. The impact of motor-cognitive dual-task training on physical and cognitive functions in Parkinson's disease. Brain Sci 2023 Mar 03;13(3):437 [FREE Full text] [doi: 10.3390/brainsci13030437] [Medline: 36979247]
12. Kwok JY, Smith R, Chan LM, Lam LC, Fong DY, Choi EP, et al. Managing freezing of gait in Parkinson's disease: a systematic review and network meta-analysis. J Neurol 2022 Jun;269(6):3310-3324. [doi: 10.1007/s00415-022-11031-z] [Medline: 35244766]
13. Ali N, Tian H, Thabane L, Ma J, Wu H, Zhong Q, et al. The effects of dual-task training on cognitive and physical functions in older adults with cognitive impairment; a systematic review and meta-analysis. J Prev Alzheimers Dis 2022;9(2):359-370. [doi: 10.14283/jpad.2022.16] [Medline: 35543010]
14. Ye JY, Chen R, Chu H, Lin HC, Liu D, Jen HJ, et al. Dual-task training in older adults with cognitive impairment: a meta-analysis and trial sequential analysis of randomized controlled trials. Int J Nurs Stud 2024 Jul;155:104776. [doi: 10.1016/j.ijnurstu.2024.104776] [Medline: 38703695]
15. Han K, Tang Z, Bai Z, Su W, Zhang H. Effects of combined cognitive and physical intervention on enhancing cognition in older adults with and without mild cognitive impairment: a systematic review and meta-analysis. Front Aging Neurosci 2022;14:878025 [FREE Full text] [doi: 10.3389/fnagi.2022.878025] [Medline: 35928994]
16. Whitehead AL, Julious SA, Cooper CL, Campbell MJ. Estimating the sample size for a pilot randomised trial to minimise the overall trial sample size for the external pilot and main trial for a continuous outcome variable. Stat Methods Med Res 2016 Jun;25(3):1057-1073 [FREE Full text] [doi: 10.1177/0962280215588241] [Medline: 26092476]
17. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet 2020 Aug 08;396(10248):413-446 [FREE Full text] [doi: 10.1016/S0140-6736(20)30367-6] [Medline: 32738937]
18. Yamada M, Aoyama T, Tanaka B, Nagai K, Ichihashi N. Seated stepping exercise in a dual-task condition improves ambulatory function with a secondary task: a randomized controlled trial. Aging Clin Exp Res 2011;23(5-6):386-392. [doi: 10.1007/BF03337763] [Medline: 21042047]
19. Morita E, Yokoyama H, Imai D, Takeda R, Ota A, Kawai E, et al. Effects of 2-year cognitive motor dual-task training on cognitive function and motor ability in healthy elderly people: a pilot study. Brain Sci 2018 May 11;8(5):86 [FREE Full text] [doi: 10.3390/brainsci8050086] [Medline: 29751661]
20. Robinson KR, Leighton P, Logan P, Gordon AL, Anthony K, Harwood RH, et al. Developing the principles of chair based exercise for older people: a modified Delphi study. BMC Geriatr 2014 May 19;14:65 [FREE Full text] [doi: 10.1186/1471-2318-14-65] [Medline: 24884392]
21. Kelly ME, Loughrey D, Lawlor BA, Robertson IH, Walsh C, Brennan S. The impact of cognitive training and mental stimulation on cognitive and everyday functioning of healthy older adults: a systematic review and meta-analysis. Ageing Res Rev 2014 May;15:28-43. [doi: 10.1016/j.arr.2014.02.004] [Medline: 24607830]
22. Farrance C, Tsofliou F, Clark C. Adherence to community based group exercise interventions for older people: a mixed-methods systematic review. Prev Med 2016 Jun;87:155-166 [FREE Full text] [doi: 10.1016/j.ypmed.2016.02.037] [Medline: 26921655]

XSL•FO
RenderX

23. Ehlers DK, Daugherty AM, Burzynska AZ, Fanning J, Awick EA, Chaddock-Heyman L, et al. Regional brain volumes moderate, but do not mediate, the effects of group-based exercise training on reductions in loneliness in older adults. Front Aging Neurosci 2017;9:110 [FREE Full text] [doi: 10.3389/fnagi.2017.00110] [Medline: 28487648]
24. Aycock DM, Hayat MJ, Helvig A, Dunbar SB, Clark PC. Essential considerations in developing attention control groups in behavioral research. Res Nurs Health 2018 Jun;41(3):320-328. [doi: 10.1002/nur.21870] [Medline: 29906317]
25. Hawley-Hague H, Horne M, Skelton DA, Todd C. Review of how we should define (and measure) adherence in studies examining older adults' participation in exercise classes. BMJ Open 2016 Jun 23;6(6):e011560 [FREE Full text] [doi: 10.1136/bmjopen-2016-011560] [Medline: 27338884]
26. Lui VW, Lam LC, Chiu HF. Validation of a memory inventory for the assessment of awareness of memory deficits in Alzheimer's disease in Chinese elderly. Int J Geriatr Psychiatry 2006 Oct;21(10):917-923. [doi: 10.1002/gps.1580] [Medline: 16927409]
27. Yeung PY, Wong LL, Chan CC, Leung JL, Yung CY. A validation study of the Hong Kong version of Montreal Cognitive Assessment (HK-MoCA) in Chinese older adults in Hong Kong. Hong Kong Med J 2014 Dec;20(6):504-510 [FREE Full text] [doi: 10.12809/hkmj144219] [Medline: 25125421]
28. Leung JL, Lee GT, Lam YH, Chan RC, Wu JY. The use of the Digit Span Test in screening for cognitive impairment in acute medical inpatients. Int Psychogeriatr 2011 Dec;23(10):1569-1574 [FREE Full text] [doi: 10.1017/S1041610211000792] [Medline: 21729426]
29. Lee TM, Chan CC. Stroop interference in Chinese and English. J Clin Exp Neuropsychol 2000 Aug;22(4):465-471. [doi: 10.1076/1380-3395(200008)22:4;1-0;FT465] [Medline: 10923056]
30. Takai Y, Ohta M, Akagi R, Kanehisa H, Kawakami Y, Fukunaga T. Sit-to-stand test to evaluate knee extensor muscle size and strength in the elderly: a novel approach. J Physiol Anthropol 2009;28(3):123-128 [FREE Full text] [doi: 10.2114/jpa2.28.123] [Medline: 19483373]
31. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179-186. [Medline: 5349366]
32. Nan H, Ni MY, Lee PH, Tam WW, Lam TH, Leung GM, et al. Psychometric evaluation of the Chinese version of the Subjective Happiness Scale: evidence from the Hong Kong FAMILY Cohort. Int J Behav Med 2014 Aug;21(4):646-652 [FREE Full text] [doi: 10.1007/s12529-014-9389-3] [Medline: 24515396]
33. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
34. Chang Q, Sha F, Chan CH, Yip PS. Validation of an abbreviated version of the Lubben Social Network Scale ("LSNS-6") and its associations with suicidality among older adults in China. PLoS One 2018;13(8):e0201612 [FREE Full text] [doi: 10.1371/journal.pone.0201612] [Medline: 30071067]
35. Sedaghat AR. Understanding the minimal clinically important difference (MCID) of patient-reported outcome measures. Otolaryngol Head Neck Surg 2019 Oct;161(4):551-560. [doi: 10.1177/0194599819852604] [Medline: 31159641]
36. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: implications for conducting a qualitative descriptive study. Nurs Health Sci 2013 Sep;15(3):398-405. [doi: 10.1111/nhs.12048] [Medline: 23480423]
37. Zhu X, Yin S, Lang M, He R, Li J. The more the better? A meta-analysis on effects of combined cognitive and physical intervention on cognition in healthy older adults. Ageing Res Rev 2016 Nov;31:67-79. [doi: 10.1016/j.arr.2016.07.003] [Medline: 27423932]
38. Gavelin HM, Dong C, Minkov R, Bahar-Fuchs A, Ellis KA, Lautenschlager NT, et al. Combined physical and cognitive training for older adults with and without cognitive impairment: a systematic review and network meta-analysis of randomized controlled trials. Ageing Res Rev 2021 Mar;66:101232. [doi: 10.1016/j.arr.2020.101232] [Medline: 33249177]
39. Park DC, Bischof GN. The aging mind: neuroplasticity in response to cognitive training. Dialogues Clin Neurosci 2013 Mar;15(1):109-119 [FREE Full text] [doi: 10.31887/DCNS.2013.15.1/dpark] [Medline: 23576894]
40. Hampshire A, Azor A, Atchison C, Trender W, Hellyer PJ, Giunchiglia V, et al. Cognition and memory after Covid-19 in a large community sample. N Engl J Med 2024 Feb 29;390(9):806-818 [FREE Full text] [doi: 10.1056/NEJMoa2311330] [Medline: 38416429]

Abbreviations

MIC: Memory Inventory in Chinese
HK-MoCA 5-Min: Montreal Cognitive Assessment 5 Minutes (Hong Kong Version)
PHQ-9: Patient Health Questionnaire-9
RCT: randomized controlled trial

XSL•FO
RenderX

Edited by Y Luo; submitted 10.10.24; peer-reviewed by A Ogunniyi, A Iizuka; comments to author 07.01.25; revised version received 02.03.25; accepted 31.03.25; published 16.05.25.

Please cite as:
Chau PH, Cheung DST, Kwok JYY, Chan WC, Yu DSF
Online Group–Based Dual-Task Training to Improve Cognitive Function of Community-Dwelling Older Adults: Randomized Controlled Feasibility Study
JMIR Aging 2025;8:e67267
URL: https://aging.jmir.org/2025/1/e67267
doi:10.2196/67267
PMID:

©Pui Hing Chau, Denise Shuk Ting Cheung, Jojo Yan Yan Kwok, Wai Chi Chan, Doris Sau Fung Yu. Originally published in JMIR Aging (https://aging.jmir.org), 16.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Predicting Postoperative Delirium in Older Patients Before Elective Surgery: Multicenter Retrospective Cohort Study

Shun-Chin Jim Wu[1,2,3*], MD, MSc; Nitin Sharma[1,2*], MSc; Anne Bauch[1,2], PhD; Hao-Chun Yang[1,2], PhD; Jasmine L Hect[4], BSc; Christine Thomas[1,5], MD; Sören Wagner[6,7], MD; Bernd R Förstner[8], PhD; Christine A F von Arnim[9,10], MD; Tobias Kaufmann[1,2,11], PhD; Gerhard W Eschweiler[1,2,12], MD; Thomas Wolfers[1,2,11], PhD; PAWEL Study Group[1,2*]

[1]Department of Psychiatry and Psychotherapy, University of Tübingen, Tübingen, Germany

[10]German Center for Cardiovascular Research (DZHK), Göttingen, Germany

[11]Norwegian Centre for Mental Disorders Research, University of Oslo, Oslo, Norway

[12]Geriatric Center, Universitätsklinikum Tübingen, Tübingen, Germany

[2]German Center for Mental Health (DZPG), partner site Tübingen, Tübingen, Germany

[3]Department of Psychiatry, University of Cambridge, Cambridge, United Kingdom

[4]Department of Neurological Surgery, University of Pittsburgh School of Medicine, Pittsburgh, PA, United States

[5]Department of Geriatric Psychiatry and Psychotherapy, Klinikum Stuttgart, Stuttgart, Germany

[6]Department of Anesthesiology, Klinikum Stuttgart, Stuttgart, Germany

[7]Department of Anesthesiology and Intensive Care, School of Medicine and Health, Klinikum Rechts der Isar, Technical University of Munich, Munich, Germany

[8]Social and Preventive Medicine, Department of Sports and Health Sciences, University of Potsdam, Potsdam, Germany

[9]Department of Geriatrics, University Medical Center Göttingen, Göttingen, Germany

[*]these authors contributed equally

Corresponding Author:
Shun-Chin Jim Wu, MD, MSc
Department of Psychiatry and Psychotherapy, University of Tübingen, Tübingen, Germany

Abstract

Background: Elective surgeries for older adults are increasing. Machine learning could enhance risk assessment, influencing surgical planning and postoperative care. Preoperative cognitive assessment may facilitate early detection and management of postoperative delirium (POD).

Objective: This study aims to assess machine learning models' predictive ability for POD, focusing on the added predictive value of the neuropsychological assessments before elective surgery.

Methods: This retrospective cohort study analyzed data from the multicenter PAWEL (Patient safety, Efficiency and Life quality in elective surgery) and PAWEL-R (risk) studies, encompassing older patients (≥70 y) undergoing elective surgeries from July 2017 to April 2019. A total of 1624 patients (52.3% male, N=850; age: mean 77.9, SD 4.9 years) were included, with a POD diagnosis made before discharge. Sociodemographic, clinical, surgical, and neuropsychological features were collected pre- and intraoperatively by care providers. Machine learning models' performance was evaluated using the area under the receiver operating characteristic curve (AUC), with permutation testing for significance, and Shapley Additive Explanations to identify effective neuropsychological assessments.

Results: Predicting POD before surgery with a random forest model achieved an AUC of 0.760. Incorporating all pre- and intraoperative features into the model yielded a slightly higher AUC of 0.783, with no statistically significant difference observed (P=.24). While cognitive factors alone were not strong predictors (AUC=0.617), specific tests within neuropsychological assessments, such as the Montreal Cognitive Assessment and Trail Making Tests, showed high feature attribution and played a crucial role in further enhancing prediction before surgery.

Conclusions: Preoperative risk prediction for POD can increase risk awareness in presurgical assessment and improve perioperative management in older patients at a high risk for delirium.

Trial Registration: German Clinical Trials Register (Germany) DRKS12797; https://drks.de/search/en/trial/DRKS00012797/details and DRKS13311; https://drks.de/search/en/trial/DRKS00013311/details

International Registered Report Identifier (IRRID): RR2-10.1186/s13063-018-3148-8

(JMIR Aging 2025;8:e67958) doi:10.2196/67958

XSL•FO
RenderX

KEYWORDS

postoperative delirium; predictive model; preoperative cognition; geriatrics; elective surgery

Introduction

In an aging society, there is a rising demand for elective surgeries due to the changing health care needs of older people [1-3]. However, this increase in elective surgeries raises concerns about additional adverse outcomes, particularly given the unique challenges posed by aging, such as pre-existing health conditions, disease sequelae, and diminished physiological reserves [4,5]. Meeting the growing demand for elective surgeries among older adults requires a comprehensive strategy, including thorough preoperative assessments, personalized care plans, and continuous postoperative support [6-8].

Postoperative delirium (POD), characterized by acute and fluctuating inattention with alterations in thinking or consciousness after surgery, affects 12% to 51% of older patients, with incidence varying by surgical procedures and regions [9]. POD in older patients is linked to heightened rehospitalization rates, persistent postoperative cognitive dysfunction, increased incidence of dementia, and elevated mortality [10-12]. Some studies have shown significant associations between POD and factors collected before or during surgery (pre- or intraoperatively) [13,14]. These factors could aid in predicting the POD [15,16]. However, different pre- and intraoperative feature categories may vary in their importance for POD risk prediction. Sociodemographic factors, such as age and sex, are critical for assessing POD risk [17-24]. Clinical data, including blood samples and chronic disease medication, is also predictive [25-27]. The type and duration of surgery and anesthesia have been shown to be indispensable for POD prediction [18,28]. Although preoperative neuropsychological assessments, like the Mini-Mental State Examination (MMSE) and others [29-31], help identify at-risk patients for early risk mitigation [32-34], these evaluations are not yet incorporated into clinical routine despite experts' recommendations [35]. The early identification of POD predictors enables clinicians to proactively assess risks to mitigate the occurrence of POD [36] and might affect patients' decisions before nonemergent surgery. Therefore, identifying preoperative predictors for POD is crucial in facilitating personalized surgical risk assessment before surgery [37,38]. Moreover, examining POD risks remains challenging due to its multifactorial origins [18,22,39]. The precise categories of preoperative and intraoperative features with superior predictive capabilities for POD remain unknown and unvalidated [40].

In this study, we (1) predict POD by using machine learning approaches using diverse pre- and intraoperative features from a large multicenter cohort of older patients and (2) conduct a comparative analysis of independent models using various predictor categories, including sociodemographic, clinical, surgical, and neuropsychological features. By featuring one of the largest cohorts to date, this multicenter study provides an extensive assessment of POD predictors in both cardiac and noncardiac surgeries through machine learning techniques. This approach can enhance presurgical risk assessment and perioperative management in older patients.

Methods

Participants

Building on previous research [13,41,42], this study comprehensively investigates the predictive capabilities of various feature categories of the entire cohort from the PAWEL (Patientensicherheit, Wirtschaftlichkeit und Lebensqualität bei elektiven Operationen, English: Patient safety, Efficiency and Life quality in elective surgery) study, originally recruited between November 21, 2017, and April 12, 2019. In addition, it also includes additional participants from the add-on PAWEL-R study (R for risk estimation), which extended the recruitment period from July 11, 2017, to January 15, 2019. By incorporating both datasets, our study expanded the cohort size, strengthening the predictive model's reliability. Machine learning models were used to predict POD, departing from original statistical non–cross-validated methods in the previous reports [13]. Patients were recruited from 5 major medical institutions in Germany (3 university hospitals: Tübingen, Freiburg, and Ulm, and 2 tertiary medical centers: Stuttgart and Karlsruhe). The study adheres to CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) [43] and STROBE (Strengthening the Reporting of Observational studies in Epidemiology) [44] guidelines (the CONSORT-EHEALTH checklist is provided in Checklist 1). This study was approved by ethical votes under the PAWEL-R study, and ethical approval (Ethical Committee 233/2017BO1) was provided by the Ethical Committee of Tübingen University Hospitals, Tübingen, Germany on June 6, 2017.

Participants included patients aged 70 years and older undergoing elective surgery (joint, spine, vessels, heart, lung, abdomen, urogenital system, and other organs) with an expected surgical duration exceeding 60 minutes. Exclusion criteria covered patients unable to communicate effectively in German, those undergoing emergency surgery, severe dementia with MMSE <15 or Montreal Cognitive Assessment (MoCA) <8, or an estimated survival time of less than 15 months. A stepped-wedge cluster randomized design was used for equitable intervention group allocation [41,42]. The study, initially including 1631 patients, conducted thorough postoperative assessments within 1 week after surgery but before discharge. To ensure accuracy in predicting POD diagnosis, patients without sufficient information for POD diagnosis before discharge (N=7) were excluded, resulting in 1624 patients included in prediction models (Figure S1). The study reported 23.1% of patients diagnosed with POD, with detailed between-group comparisons available in Tables S1 and S2.

Measures

The PAWEL and PAWEL-R studies extensively evaluated elective surgery patients to identify risk factors and outcomes related to POD. Diagnosis involved the Confusion Assessment Method (I-CAM) algorithm and chart review within the first postoperative week or until discharge. Features were categorized

XSL•FO
RenderX

into 4 feature groups: sociodemographic, neuropsychological, clinical, and surgical information.

Sociodemographic data, including age, sex, education, alcohol and smoking habits, living arrangements, and hospital location, were collected preoperatively. Neuropsychological assessments such as MoCA, Trail Making Test (TMT) parts A and B, digit span backward, Subjective Memory Impairment (SMI), and Patient Health Questionnaire-4 (PHQ-4) were conducted during admission. Clinical profiles included blood samples, past medical histories (including pre-existing mild or moderate dementia and previous delirium history), preoperative and intraoperative medication dose (including benzodiazepine, neuroleptics, opiates, and propofol), polypharmacy, multimorbidity, the mininutritional assessment, American Society of Anesthesiologists (ASA) physical status classification system, Charlson comorbidity index, Barthel index, clinical frailty scale, and sensory impairments (auditory, visual, and others). Surgical information covered types and duration of surgery and anesthesia, cardiopulmonary bypass, and intraoperative physiological changes. Preoperative and intraoperative features were defined to assess predictive performance achieved at 2 distinct time points: before and during surgery. Additional details can also be found in Table S1.

To identify effective predictors for POD diagnosis, the specific and combined predictive capacity of sociodemographic, clinical, and surgical features before (preoperative) and during surgery (intraoperative) were analyzed. Models were developed and tested with both features and preoperative features alone. In addition, models were compared with and without neuropsychological assessments to examine their contribution to predictive performance.

Preprocessing, Imputation, and Features

Two features with excessive missingness were excluded, namely, albumin level (979/1624, 60.28%) and depth of anesthesia (1224/1624, 75.37%) were excluded [16]. Information about missing values is available in Table S3. Data imputation involved using multiple imputation (IterativeImputer) for continuous variables and random sampling from the original probability distribution for discrete and binary variables within cross-validation folds for all models, except for gradient boosting, which does not require imputation [45]. Continuous and discrete variables were scaled within cross-validation folds for all models. Potential outliers or data entry errors in clinical assessments were identified when a data point exceeded 5 SD from the mean and were subsequently imputed using the above method.

Blood samples, including hemoglobin, sodium levels, and C-reactive protein (CRP), were interpreted and discretized following clinical guidelines from Harrison's Principles of Internal Medicine, 20th version, in line with standard practice [46]. Specifically, hemoglobin levels of less than 12 g/dL were considered indicative of anemia, while values of 12 g/dL or higher were considered normal. Serum sodium levels were categorized as hyponatremia if below 135 mmol/L and hypernatremia if above 145 mmol/L, with values within this range considered normal. Similarly, CRP levels greater than 3 mg/L indicated elevated levels in adults aged 65 years and older [47], while values within the normal range were not considered clinically significant. Redundant features with a perfect correlation were excluded; therefore, use of heart-lung machine was excluded for its perfect correlation with cardiopulmonary bypass. Nonbinary categorical features, including location, SMI, types of anesthesia and surgery, and transformed blood samples were one-hot encoded because they had no natural ordinal relationship among their categories, and assigning numerical labels to them could introduce bias or incorrect assumptions in the model.

Fifty-nine features were used, divided into preoperative (48 features) and intraoperative (11 features) categories. Preoperative features encompassed sociodemographic (7 features), clinical (20 features), surgical categories (6 features), and neuropsychological assessments (15 features). Intraoperative features included clinical (4 features) and surgical (7 features) categories. Different combinations of features were compared, including preoperative only, preoperative and intraoperative, and each category of preoperative features. The study evaluated model performance, feature category effectiveness, and the additional benefits of preoperative neuropsychological assessments.

The study aimed to develop a prediction model in a naturalistic setting. Information regarding interventions was included only as a sensitivity analysis to demonstrate its potential impact on the prediction model. The potential imbalance of the dataset for all models was tested by oversampling with the Synthetic Minority Oversampling Technique [48]. Given that the overall sex distribution in the cohort was fairly balanced (774/1624, 47.66% female), we did not apply additional rebalancing for sex to avoid potential bias. In addition, no significant age differences were observed between patients with and without POD. Various sensitivity analyses described were performed.

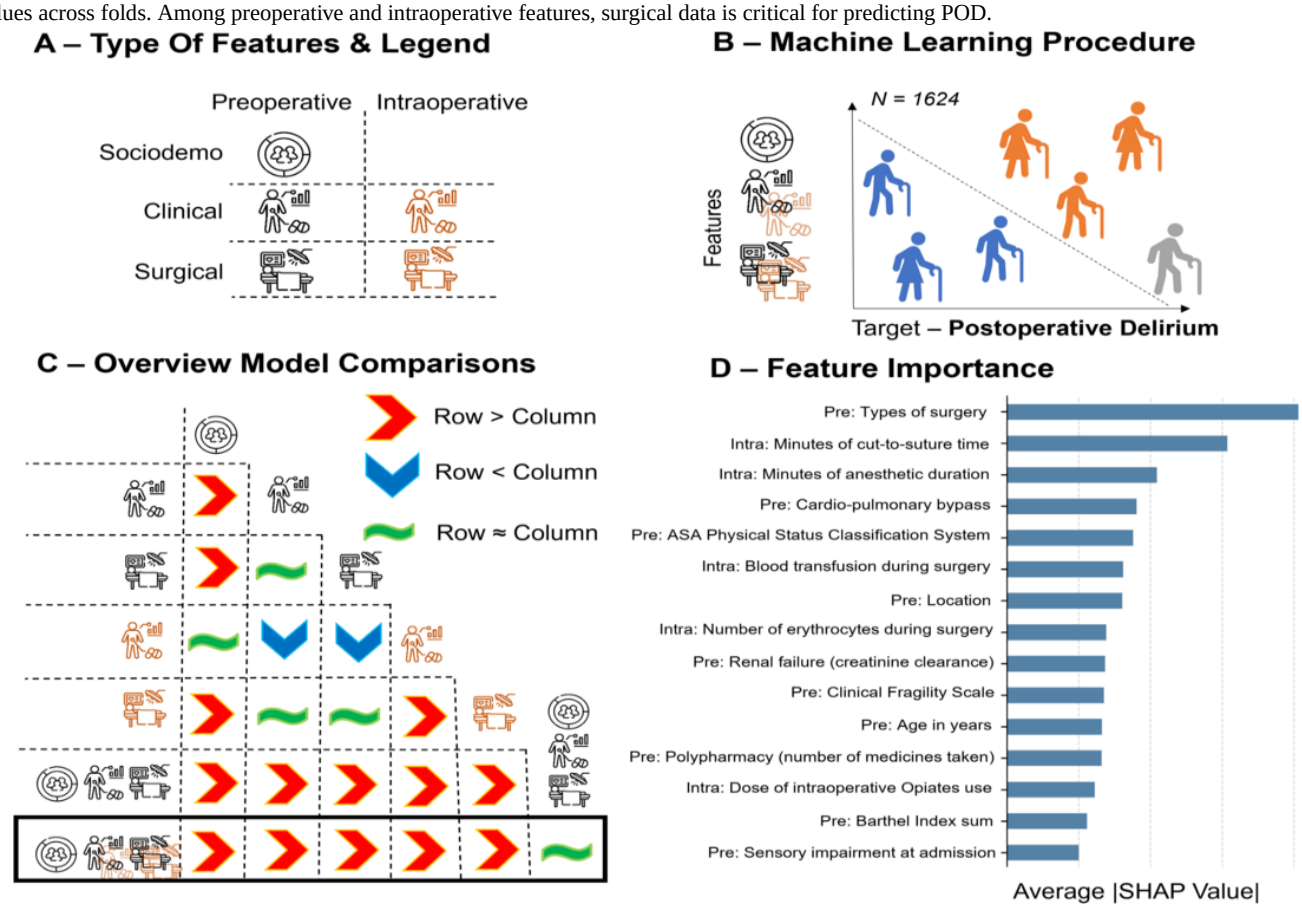
Machine Learning Models, Performance Evaluation, and Feature Importance

Machine learning models were used to predict POD, including logistic regression, support vector machines, random forest, and gradient boosting without hyperparameter tuning using the scikit-learn library version 1.2.2 [49] and the Xgboost library version (1.7.3) [50]. Independent variables were feature variables, while the dependent variable was POD diagnosis, as illustrated in Figure 1B. Feature selection was performed using the SelectFromModel function from scikit-learn, leveraging model-based feature selection, to retain the most predictive variables. Given the sample size relative to the complexity of the models, we did not anticipate substantial improvements from hyperparameter tuning. However, we conducted additional sensitivity analyses to assess its impact on model performance via nested cross-validation (Tables S12-S15 in Multimedia Appendix 1). To ensure model stability and interpretability, we used default parameters from the software for the primary analysis (Table S11 in Multimedia Appendix 1). 5-fold cross-validation, with balanced labels across folds, measured model performance at testing using the area under the receiver operating characteristic curve (AUC) as the primary metric. In order to robustly evaluate model performance against chance and between models, permutation testing was used to assess

XSL•FO
RenderX

whether the AUC of each model was significantly greater than expected by chance. Specifically, POD diagnostic labels were randomly shuffled, and AUC values were recalculated 1000 times to generate a null distribution. The *P* value was derived by comparing the observed AUC to this null distribution (Table S4 in Multimedia Appendix 1). Furthermore, to compare AUC differences between models, permutation testing was also used to generate a null distribution of AUC differences by randomly permuting the labels 1000 times while preserving data dependencies. The *P* value for model comparison was computed based on the observed difference relative to this null distribution [51,52]. Lastly, to obtain a robust measure of variability in model performance, 95% CI for AUC using bootstrapping with 1000 resamples was estimated. Additional metrics included precision, recall, sensitivity, specificity, balanced accuracy, and area under the precision-recall curve presented in Table S5 in Multimedia Appendix 1.

Figure 1. Individual feature categories perform worse than combined feature categories pre-, intraoperatively, and postoperative delirium (POD) can effectively be predicted prior to surgery (see back box). (**A**) The symbols are consistently used in this study and represent the different types of features. (**B**) A random forest classifier with 5-fold cross-validation was used to forecast POD onset in 1624 older individuals. Among them, 23.1% experienced POD. Orange indicates patients with POD, blue represents those without POD, and gray signifies patients predicted to have POD by the algorithm. (**C**) A comparative analysis of different feature combinations based on their respective area under the receiver operating characteristic curve (AUC) values, calculated over 1000 permutations was conducted. Combined feature sets consistently outperformed individual feature types. (**D**) Shapley Additive Explanations (SHAP) was used to assess feature importance. The feature importance was based on the average absolute value of the aggregated SHAP values across folds. Among preoperative and intraoperative features, surgical data is critical for predicting POD.

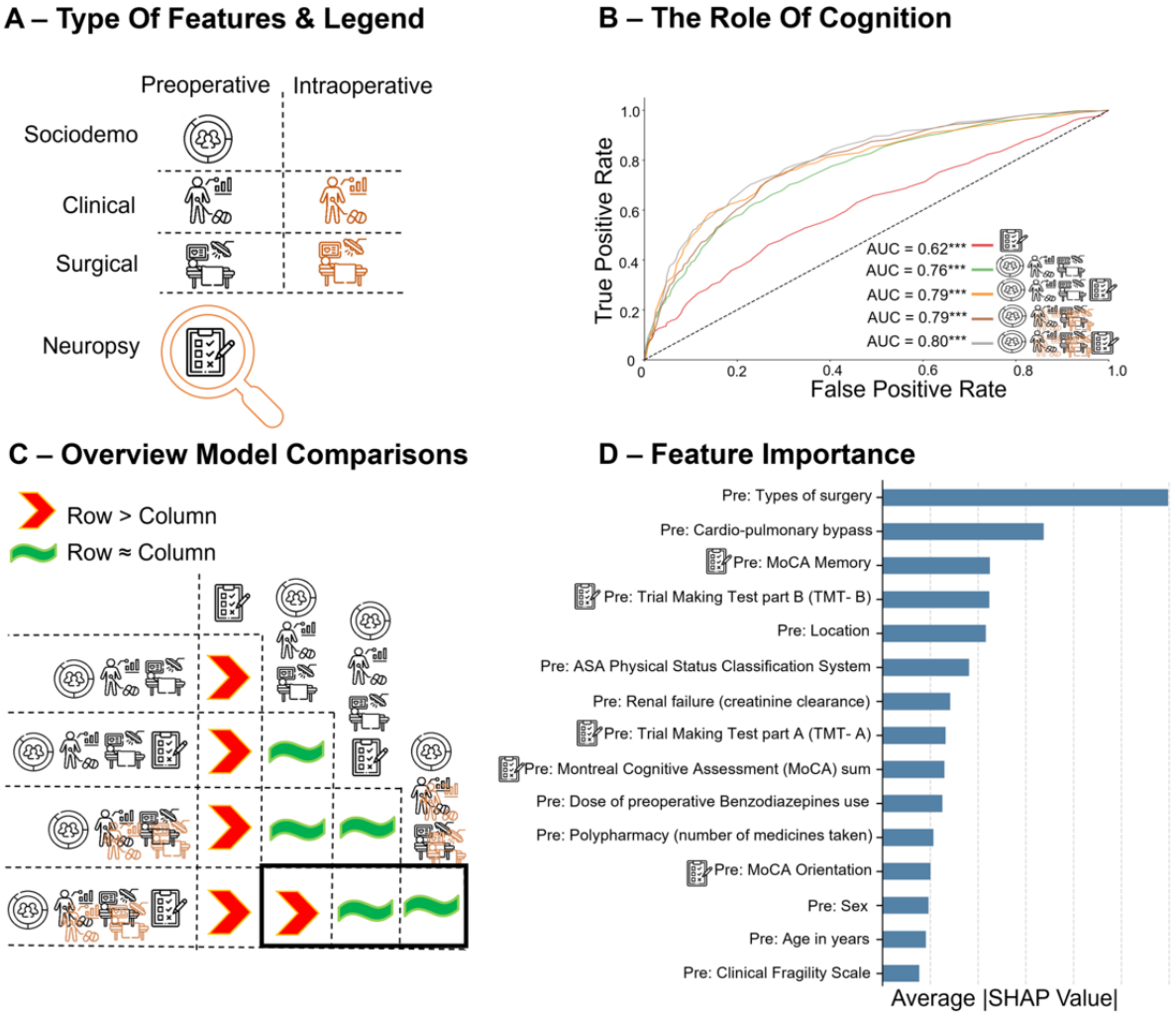


The Shapley Additive Explanations (SHAP) values were used to assess feature importance. Positive SHAP values increased the probability of POD, while negative values decreased it. SHAP values were computed across all cross-validation folds and aggregated over 5-fold cross-validation with 5 shuffles to ensure stability and robustness in feature importance rankings. To facilitate explanation of the main models, we presented the top 15 features contributing to model performance and complete feature attributions in major feature sets: preoperative features with and without neuropsychological assessments, and combined preoperative and intraoperative features with and without neuropsychological assessments (Figures 1D and 2D, Figure S2 in Multimedia Appendix 1, and Tables S7-S10 in Multimedia Appendix 1). The SHAP library version 0.41.0 in Python was used [53]. To evaluate the calibration of our models, calibration plots were generated (applying Platt scaling if needed), and the number of patients with high-confidence positive predictions (>0.9) and high-confidence negative predictions (<0.1) was calculated for each model. This analysis was conducted across different feature sets to assess whether the addition of features affected the model's probability estimates.

All preprocessing steps and analyses are available on GitHub upon publication from Sharma forked to our laboratory's page, and data can be requested by addressing the PAWEL consortium as indicated in the data availability section.

XSL•FO
RenderX

Figure 2. Cognition, in isolation, demonstrates limited predictive power. However, when combined with preoperative features, the performance matches that of all pre- and intraoperative features (see back box). (**A**) The predictive capability of preoperative neuropsychological tests was assessed for postoperative delirium (POD) individually and in combination with other features. (**B**) Our analysis, as depicted in the receiver operating characteristic curve, demonstrated the successful prediction of POD by preoperative neuropsychological tests. However, when combined with pre- and intraoperative data, the performance gains were only marginal. Asterisks (*) denote the Area Under the receiver operating characteristic Curve (AUC) values, which were assessed against chance permutation testing, revealing a significance (**$P \leq .001$). (**C**) The AUC values of neuropsychological tests were compared to those integrated into the pre- and intraoperative features. Adding preoperative neuropsychological tests to the preoperative model improves POD prediction to a level comparable to the best-performing model, similar to the effect of incorporating intraoperative features. In contrast, using preoperative features alone without neuropsychological tests results in lower performance. (**D**) Specific neuropsychological tests, such as the Montreal Cognitive Assessment (MoCA) and Trail Making Tests (TMTs), were particularly influential in POD prediction.



Ethical Considerations

This study was approved by the Ethics Commission of the Faculty of Medicine of the University of Tübingen (233/2017BO1, October 12, 2017) and the Ethics Commission of the University of Potsdam (38/2017, December 11, 2017), and was registered in the German Clinical Trials Register (DRKS00012797, July 2017).

All participants provided written informed consent for participation and for the publication of any potentially identifiable data. Data were pseudonymized, stored securely, and handled in accordance with applicable privacy and confidentiality regulations. Participants were not financially compensated but were reimbursed for travel expenses when applicable.

Results

Predicting POD With Combined, Pre-, and Intraoperative Features

The models incorporating combined and independent pre- and intraoperative features exhibited robust performance, as evidenced by AUC values surpassing chance levels, all with $P<.002$ (Table S4 in Multimedia Appendix 1) and Table S5 in Multimedia Appendix 1 displays the performance of these models through receiver operating characteristic curves with the random forest. The combined model using only preoperative features (Pre-Op) achieved an AUC of 0.760, comparable to a model incorporating both pre- and intraoperative features (Pre- and Intra-Op) with an AUC of 0.783, showing no statistically significant difference (Figure 1C and Table S6 in Multimedia

XSL•FO
RenderX

Appendix 1). Independent models exclusively using preoperative clinical, preoperative surgical, or intraoperative surgical features also demonstrated high AUC values (0.691, 0.664, and 0.670, respectively) without statistically significant difference. Notably, key predictors of POD included surgery type and cardiopulmonary bypass (preoperative surgical; recruiting hospital and age (preoperative sociodemographic); ASA status, Clinical Frailty Scale, Barthel Index, polypharmacy, and creatinine clearance (preoperative clinical); cut-to-suture time, anesthesia duration, and blood loss (intraoperative surgical) shown in Figure 1D and Figure S2-A in Multimedia Appendix 1. Further details, including all pairwise model comparisons and *P* values for differences in AUC values, can be found in Table S6 in Multimedia Appendix 1. A detailed list of the feature importance can be found in Tables S7-S8 in Multimedia Appendix 1.

Addition of Preoperative Neuropsychological Assessments

The model using preoperative neuropsychological assessments exclusively exhibited AUC values of 0.617. Integrating neuropsychological assessments into the model, using both pre- and intraoperative features (Pre- and Intra-Op+NeuroPsy), led to a slight improvement in the AUC, reaching 0.803 (Figure 2B and Table S6 in Multimedia Appendix 1). Adding neuropsychological assessments to the preoperative model (Pre-Op+ NeuroPsy) improved the AUC from 0.760 to 0.787, which resulted in a model that matched the above best-performing model (Figure 2C and Table S6 in Multimedia Appendix 1). Specifically, the MoCA scores and TMTs before surgery were important for predicting POD, as illustrated in Figure 2D and Figure S2-B in Multimedia Appendix 1. For detailed pairwise comparisons of AUC values and corresponding *P* values, please refer to Table S6 in Multimedia Appendix 1. A detailed list of the feature importance can be found in Tables S9-S10 in Multimedia Appendix 1.

Robustness of Predictive Models in Performance Evaluation

Evaluating the performance of 4 classifiers with tuned hyperparameters revealed comparable results (Tables S12-S15 in Multimedia Appendix 1). The random forest model, demonstrating marginally better performance in AUC values in models with combined features, was highlighted above. In addition, random forest models demonstrated good calibration, as indicated by a consistent distribution of high-confidence predictions across different feature sets. The inclusion of intraoperative and neuropsychological information did not compromise calibration, ensuring the model's reliability in distinguishing high- and low-risk patients (Figure S3 in Multimedia Appendix 1).

Including intervention allocation information had no discernible impact on predictive performance (Tables S5 and S16 in Multimedia Appendix 1). Finally, there was no statistically significant difference in AUC between models with and without oversampling (Table S17 in Multimedia Appendix 1).

Discussion

Principal Findings

Leveraging machine learning, we predicted the occurrence of POD after elective surgeries through a combination of preoperative and intraoperative features with a large multicenter cohort, achieving an AUC above 0.8. This performance exceeds that of traditional scoring systems such as the Delirium Risk Assessment Tool, Delirium Risk Assessment Score, and Delirium Elderly At-Risk, which have demonstrated AUC values between 0.5 and 0.7 in a large cohort [54-56]. Surgical information both before and during the surgery was critical in predicting POD. Integrating neuropsychological tests into preoperative features enhanced the AUC to a level comparable to the best-performing model, effectively replacing intraoperative features for predicting POD before surgery. This improvement was primarily driven by MoCA and TMTs A and B, as elucidated by model explanations. This integrated analysis improves conventional clinical risk profiling, furnishing superior predictive capacity with promising implications for surgical planning in the era of machine learning-assisted health care and empowering the prioritization of pivotal features in future work.

Critical Surgical Information

Surgical features, both preoperative and intraoperative, emerged as good solo predictors category for POD (Figure 1C). While baseline clinical profiles, including ASA physical status, renal function, frailty, and polypharmacy, are important, intraoperative surgical features such as individual surgery and anesthesia duration and factors related to blood loss were particularly important for predicting POD (Figure 1D and Figure S2-A in Multimedia Appendix 1). In addition, preoperative surgical information such as types of surgery and use of cardiopulmonary bypass were also associated with a higher risk of POD, aligning with a higher overall incidence of cardiac surgery in patients with POD relative to those without POD (206/375, 54.9% vs 264/1249, 21.1% as indicated in Table S1 in Multimedia Appendix 1). Consistent with our findings, a 30-minute increase in surgery duration corresponded to a 6% rise in POD risk [28], and this risk is further elevated in prolonged cardiac surgeries using cardiopulmonary bypass [57], potentially leading to hypoperfusion or microembolism [58]. This suggests a potential cumulative effect, particularly in patients with poor preoperative clinical profiles. Overall, our findings highlight the critical importance of cardiovascular surgical risk measures in POD prediction.

Looking ahead, the integration of real-time predictive technology into surgical workflows holds promise. This advancement could potentially facilitate on-the-fly predictions during surgery, enabling timely adjustments to medication or nonpharmacological intervention to mitigate potential adverse outcomes associated with surgical interventions. Our study emphasizes the substantial value of information gleaned from measures taken during surgery, shedding light on their crucial role in enhancing our understanding and prediction of POD.

XSL•FO
RenderX

Enhanced POD Prediction Before Elective Surgery Through Neuropsychological Assessments

With a larger cohort and a more comprehensive battery of preoperative neuropsychological assessments, this study demonstrated slightly improved predictive performance compared to a previous model that used a smaller sample and a limited neuropsychological assessment (MoCA) [37]. Consistent with prior research [16], our results indicate that preoperative models can achieve comparable performance to those incorporating intraoperative features in predicting POD. Although incorporating intraoperative data slightly improved POD predictions, they do not diminish the value of neuropsychological testing for early risk stratification, especially in elective surgeries where timing allows for actionable planning. A key limitation of intraoperative data prevents surgical planning and decision-making beforehand, allowing only adjustments in real time. These findings are particularly relevant to older patients undergoing elective surgery, as they have sufficient preoperative time for noninvasive neuropsychological assessments and for adjusting surgical strategies accordingly. Therefore, augmenting the predictive performance by incorporating data that can be gathered before a surgical procedure is important, as it allows for potential prehabilitation strategies, integrated surgical planning, and informed decision making prior to any invasive or surgical procedure.

In our study, preoperative neuropsychological assessments were predictive above chance (Figure 2B). To achieve performance comparable to the model combining all available features, adding neuropsychological tests to the preoperative model can effectively replace intraoperative features to predict POD before surgery (Figure 2C), which is critical given that surgical and postoperative management could be optimized. Preoperative models with neuropsychological tests effectively predicted POD before surgery substantiating previous observations [13,29,59]. This could be attributed to preoperative neuropsychological tests revealing subtle cognitive deficits that are not captured by dementia or delirium history. Fewer than 2% of patients in this cohort reported a diagnosis of mild or moderate dementia before elective surgery (Table S1 in Multimedia Appendix 1), and neither dementia nor delirium history was influential predictors with consistently almost zero feature attributions (Tables S7-S10 in Multimedia Appendix 1). These subtle deficits may progress into POD. In line with this explanation, timely preoperative cognitive interventions can mitigate the risk of POD and long-term cognitive dysfunction after cardiac surgeries [60,61]. In addition, patients with pre-existing cognitive decline face increased risks of other postoperative complications [62,63]. Consequently, baseline neuropsychological assessments are valuable for improving the prediction of POD beyond what clinical history alone can offer.

Selecting suitable neuropsychological assessments for clinical use is crucial. Our study identified the MoCA and TMTs as effectively indicated by their average absolute SHAP values (Figure 2D and Figure S2-B in Multimedia Appendix 1). Low scores on the MoCA and longer test times on the TMT indicate poor cognitive performance and executive dysfunction. These tests are crucial for predicting POD risk, as demonstrated in a previous prediction study [15]. Our findings, while focused on prediction [64], can complement previous etiological studies that show patients with mild cognitive impairment at baseline are more likely to develop POD [26], while good preoperative cognitive performance is protective against POD [18]. Previous studies have often used the MMSE [29,32,65], Clock-Drawing Test [20], or MoCA score as preoperative risk factors [66]. Critically, we replicated the strong association between baseline MoCA and POD risk in the previous theory-driven etiological PAWEL-R study with a larger and extended cohort [13]. These findings offer a thorough understanding of the efficacy of individual preoperative neuropsychological tests in predicting POD. By conducting comprehensive assessments of predictors for pre-existing risks, we may unlock new avenues for optimizing surgical planning and postoperative management. Our study underscores the added, albeit moderate, advantage of evaluating cognition, emphasizing its importance and advocating for its inclusion in future developments aimed at refining preoperative risk assessments.

Limitations and Recommendations

The study exhibits several limitations that require consideration. First, interpreting and using SHAP values warrants caution [67], as is generally the case for methods using model explanations in medicine [68,69]. Unstable explanations are not uncommon for complex models trained on large datasets [70,71]. While the ranking of importance may fluctuate, features with higher mean absolute SHAP values generally maintain consistent attributions. To enhance stability and reduce dependence on specific cohort segments, we aggregated SHAP values across multiple folds with shuffling, ensuring a more reliable assessment of feature importance (Tables S7-S10 in Multimedia Appendix 1). Second, although permutation is a robust statistical method, its conservative nature means nonsignificance does not always indicate the absence of a difference. Third, to minimize temporal bias and avoid causal leakage, we excluded perioperative features whose timing could not be clearly verified to precede the onset of delirium. While such features may have clinical relevance, their inclusion risks introducing postoutcome information. We recommend future research explore causal-learning approaches to better define temporal relationships, though limitations of observational data must be considered [72]. Fourth, we harnessed cognitive data derived from standardized assessments. However, it is crucial to note that these procedures, while standardized, often remain nondigitalized. This presents significant untapped potential for future advancements in the realm of risk prediction before surgical procedures. Fifth, while our study included a wide range of preoperative neuropsychological assessments, it is not an exhaustive list. As neuropsychological assessments are time-consuming and require training for assessors to conduct accurate tests and interpret results. Therefore, incorporating semiautomatic assessments of cognition [73] may prove advantageous and a relevant direction for future research.

Conclusion and Relevance

This study highlights the feasibility of predicting POD before elective surgery in adults aged 70 years and older by using a diverse set of features, including neuropsychological

XSL•FO
RenderX

assessments. Our research advances the understanding of POD predictors, enabling a more targeted approach to POD risk prediction in clinical practice. The findings offer crucial insights into predictive features for POD, underscoring the importance of integrating these predictors into the digital transformation of preoperative risk assessment.

Acknowledgments

This work was supported by grant VF1_2016- - 201 from the Innovationsfonds (fund of the Gemeinsamer Bundesausschuss, GBA) as well as the German Research Foundation (DFG) Emmy Noether with reference 513851350 (TW) and the BMBF/DLR Project FEDORA: 01EQ2403G (TW). This work was supported by the BMBF-funded de.NBI Cloud within the German Network for Bioinformatics Infrastructure (de.NBI) (031A532B, 031A533A, 031A533B, 031A534A, 031A535A, 031A537A, 031A537B, 031A537C, 031A537D, 031A538A).

We would like to thank the PAWEL (Patientensicherheit, Wirtschaftlichkeit und Lebensqualität bei elektiven Operationen) and PAWEL-R (Patientensicherheit, Wirtschaftlichkeit und Lebensqualität bei elektiven Operationen–Risk) Study group and the patients' participation in this study, making the work possible.

"PAWEL Study Group" includes members: Florian Metzger, Department of Psychiatry and Psychotherapy, University Hospital Tübingen; Andreas Straub, Department of Anesthesiology, University Hospital Tübingen; Tobias Krüger, Department of Heart Surgery, University Hospital TübingenFelix Bausenhart, Department of Orthopedics, University Hospital Tübingen; Petra Renz, Nursing Department, University Hospital Tübingen; Andreas Walther, Department of Anaesthesiology and Intensive Care, Katharinen Hospital, Klinikum Stuttgart; Carola Bruns, Department of Old Age Psychiatry and Psychotherapy, Klinikum Stuttgart; Juliane Spank, Department of Old Age Psychiatry and Psychotherapy, Klinikum Stuttgart; Patricia Sabbah, Department of Old Age Psychiatry and Psychotherapy, Klinikum Stuttgart; Andreas Häusler, Department of Social and Preventive Medicine, University of Potsdam; Bernd Förstner, Department of Social and Preventive Medicine, University of Potsdam; Susanne Schulze, Department of Social and Preventive Medicine, University of Potsdam; Markus Martin, Department of Neurology and Neurophysiology, University of Freiburg; Bernhard Heimbach, Sebastian Voigt-Radloff, Center for Geriatric Medicine and Gerontology (ZGGF), University of Freiburg; Heiko Reichel, Department of Orthopedics, University Hospital Ulm; Andreas Liebold, Department of Cardiothoracic and Vascular Surgery, University Hospital Ulm; Simone Brefka, Agaplesion Bethesda Clinic, Geriatric Medicine, Ulm University and Geriatric Center Ulm; Stephan Kirschner, Department of Orthopedics, ViDia Christian Clinics Karlsruhe; Nina Stober, Geriatric Center Karlsruhe, ViDia Christian Clinics Karlsruhe; Uwe Mehlhorn, Helios Clinic for Cardiac Surgery, Study Center in KarlsruheJürgen Wasem, Institute for Healthcare Management and Research, University Duisburg-Essen; Anja Neumann, Institute for Healthcare Management and Research, University Duisburg-Essen

Data Availability

All preprocessing steps and analyses are available on GitHub [74] upon publication. The dataset analyzed during this study is available from the PAWEL study consortium on reasonable request.

Conflicts of Interest

CT received honoraria from serving on the scientific advisory board of Roche, a research grant from the gemeinsame Bundesausschuß der Krankenkassen and has received funding for travel and speaker honoraria from several hospitals for scientific education. GWE has nothing to disclosure. CAFvA received honoraria from serving on the scientific advisory board of Biogen, Roche, Novo Nordisk, Biontech, MindAhead UG, Lilly and Dr. Willmar Schwabe GmbH &Co. KG and has received funding for travel and speaker honoraria from Lilly, Biogen, Roche diagnostics AG, Novartis, Medical Tribune Verlagsgesellschaft mbH, Landesvereinigung für Gesundheit und Akademie für Sozialmedizin Niedersachsen e. V., FomF GmbH | Forum für medizinische Fortbildung and Dr. Willmar Schwabe GmbH &Co. KG and has received research support from Roche diagnostics AG.

Multimedia Appendix 1
Additional figures and tables.
[DOCX File, 724 KB - aging_v8i1e67958_app1.docx]

Checklist 1
CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth; version 1.6.1) checklist.
[PDF File, 3180 KB - aging_v8i1e67958_app2.pdf]

References

XSL•FO
RenderX

1. Deng C, Mitchell S, Paine SJ, Kerse N. Retrospective analysis of the 13-year trend in acute and elective surgery for patients aged 60 years and over at Auckland City Hospital, New Zealand. J Epidemiol Community Health 2020 Jan;74(1):42-47. [doi: 10.1136/jech-2019-212283] [Medline: 31649040]
2. Lee YZ, Dharmawan A, Zhang X, Chua DYC, Low JK. The changing landscape of general surgery in the elderly - trends over a decade in a tertiary centre in Singapore. ANZ J Surg 2022 Sep;92(9):2018-2024. [doi: 10.1111/ans.17674] [Medline: 35373437]
3. Lutz W, Sanderson W, Scherbov S. The coming acceleration of global population ageing. Nature New Biol 2008 Feb 7;451(7179):716-719. [doi: 10.1038/nature06516] [Medline: 18204438]
4. Glance LG, Benesch CG, Holloway RG, et al. Association of time elapsed since ischemic stroke with risk of recurrent stroke in older patients undergoing elective nonneurologic, noncardiac surgery. JAMA Surg 2022 Aug 1;157(8):e222236. [doi: 10.1001/jamasurg.2022.2236] [Medline: 35767247]
5. Graham LA, Hawn MT. Managing competing risks for surgical patients with complex medical problems—considering confounding. JAMA Surg 2024 Feb 1;159(2):149. [doi: 10.1001/jamasurg.2023.5952]
6. Partridge JSL, Harari D, Martin FC, Dhesi JK. The impact of pre-operative comprehensive geriatric assessment on postoperative outcomes in older patients undergoing scheduled surgery: a systematic review. Anaesthesia 2014 Jan;69 Suppl 1(s1):8-16. [doi: 10.1111/anae.12494] [Medline: 24303856]
7. Kim KI, Park KH, Koo KH, Han HS, Kim CH. Comprehensive geriatric assessment can predict postoperative morbidity and mortality in elderly patients undergoing elective surgery. Arch Gerontol Geriatr 2013;56(3):507-512. [doi: 10.1016/j.archger.2012.09.002] [Medline: 23246499]
8. Aceto P, Antonelli Incalzi R, Bettelli G, et al. Perioperative Management of Elderly patients (PriME): recommendations from an Italian intersociety consensus. Aging Clin Exp Res 2020 Sep;32(9):1647-1673. [doi: 10.1007/s40520-020-01624-x] [Medline: 32651902]
9. Whitlock EL, Vannucci A, Avidan MS. Postoperative delirium. Minerva Anestesiol 2011 Apr;77(4):448-456. [Medline: 21483389]
10. Witlox J, Eurelings LSM, de Jonghe JFM, Kalisvaart KJ, Eikelenboom P, van Gool WA. Delirium in elderly patients and the risk of postdischarge mortality, institutionalization, and dementia: a meta-analysis. JAMA 2010 Jul 28;304(4):443-451. [doi: 10.1001/jama.2010.1013] [Medline: 20664045]
11. Vasilevskis EE, Han JH, Hughes CG, Ely EW. Epidemiology and risk factors for delirium across hospital settings. Best Practice & Research Clinical Anaesthesiology 2012 Sep;26(3):277-287. [doi: 10.1016/j.bpa.2012.07.003]
12. Gleason LJ, Schmitt EM, Kosar CM, et al. Effect of delirium and other major complications on outcomes after elective surgery in older adults. JAMA Surg 2015 Dec;150(12):1134-1140. [doi: 10.1001/jamasurg.2015.2606] [Medline: 26352694]
13. Eschweiler GW, Czornik M, Herrmann ML, et al. Presurgical screening improves RISK prediction for delirium in elective surgery of older patients: The PAWEL RISK study. Front Aging Neurosci 2021;13:679933. [doi: 10.3389/fnagi.2021.679933]
14. Susano MJ, Scheetz SD, Grasfield RH, et al. Retrospective analysis of perioperative variables associated with postoperative delirium and other adverse outcomes in older patients after spine surgery. J Neurosurg Anesthesiol 2019;31(4):385-391. [doi: 10.1097/ANA.0000000000000566]
15. Sadlonova M, Hansen N, Esselmann H, et al. Preoperative delirium risk screening in patients undergoing a cardiac surgery: Results from the prospective observational FINDERI study. Am J Geriatr Psychiatry 2024 Jul;32(7):835-851. [doi: 10.1016/j.jagp.2023.12.017]
16. Xue B, Li D, Lu C, et al. Use of machine learning to develop and evaluate models using preoperative and intraoperative data to identify risks of postoperative complications. JAMA Netw Open 2021 Mar 1;4(3):e212240. [doi: 10.1001/jamanetworkopen.2021.2240] [Medline: 33783520]
17. Chaiwat O, Chanidnuan M, Pancharoen W, et al. Postoperative delirium in critically ill surgical patients: incidence, risk factors, and predictive scores. BMC Anesthesiol 2019 Mar 20;19(1):39. [doi: 10.1186/s12871-019-0694-x] [Medline: 30894129]
18. Greaves D, Psaltis PJ, Davis DHJ, et al. Risk factors for delirium and cognitive decline following coronary artery bypass grafting surgery: A systematic review and meta-analysis. J Am Heart Assoc 2020 Nov 17;9(22):e017275. [doi: 10.1161/JAHA.120.017275] [Medline: 33164631]
19. Jankowski CJ, Trenerry MR, Cook DJ, et al. Cognitive and functional predictors and sequelae of postoperative delirium in elderly patients undergoing elective joint arthroplasty. Anesth Analg 2011 May;112(5):1186-1193. [doi: 10.1213/ANE.0b013e318211501b] [Medline: 21415433]
20. Vasilian CC, Tamasan SC, Lungeanu D, Poenaru DV. Clock-drawing test as a bedside assessment of post-operative delirium risk in elderly patients with accidental hip fracture. World J Surg 2018 May;42(5):1340-1345. [doi: 10.1007/s00268-017-4294-y] [Medline: 29030678]
21. Wang CG, Qin YF, Wan X, Song LC, Li ZJ, Li H. Incidence and risk factors of postoperative delirium in the elderly patients with hip fracture. J Orthop Surg Res 2018 Jul 27;13(1):186. [doi: 10.1186/s13018-018-0897-8] [Medline: 30049276]
22. Wu J, Yin Y, Jin M, Li B. The risk factors for postoperative delirium in adult patients after hip fracture surgery: a systematic review and meta-analysis. Int J Geriatr Psychiatry 2021 Jan;36(1):3-14. [doi: 10.1002/gps.5408] [Medline: 32833302]

XSL•FO
RenderX

23. Wang H, Guo X, Zhu X, et al. Gender differences and postoperative delirium in adult patients undergoing cardiac valve surgery. Front Cardiovasc Med 2021;8:751421. [doi: 10.3389/fcvm.2021.751421]
24. Ansaloni L, Catena F, Chattat R, et al. Risk factors and incidence of postoperative delirium in elderly patients after elective and emergency surgery. Br J Surg 2010 Feb;97(2):273-280. [doi: 10.1002/bjs.6843] [Medline: 20069607]
25. Ayob F, Lam E, Ho G, Chung F, El-Beheiry H, Wong J. Pre-operative biomarkers and imaging tests as predictors of post-operative delirium in non-cardiac surgical patients: a systematic review. BMC Anesthesiol 2019 Feb 23;19(1):25. [doi: 10.1186/s12871-019-0693-y] [Medline: 30797230]
26. Kazmierski J, Banys A, Latek J, et al. Mild cognitive impairment with associated inflammatory and cortisol alterations as independent risk factor for postoperative delirium. Dement Geriatr Cogn Disord 2014;38(1-2):65-78. [doi: 10.1159/000357454] [Medline: 24603477]
27. Kassie GM, Nguyen TA, Kalisch Ellett LM, Pratt NL, Roughead EE. Do risk prediction models for postoperative delirium consider patients' preoperative medication use? Drugs Aging 2018 Mar;35(3):213-222. [doi: 10.1007/s40266-018-0526-6] [Medline: 29423780]
28. Ravi B, Pincus D, Choi S, Jenkinson R, Wasserstein DN, Redelmeier DA. Association of duration of surgery with postoperative delirium among patients receiving hip fracture repair. JAMA Netw Open 2019 Feb 1;2(2):e190111. [doi: 10.1001/jamanetworkopen.2019.0111] [Medline: 30794305]
29. Liu J, Li J, He J, Zhang H, Liu M, Rong J. The age-adjusted Charlson Comorbidity Index predicts post-operative delirium in the elderly following thoracic and abdominal surgery: A prospective observational cohort study. Front Aging Neurosci 2022;14:979119. [doi: 10.3389/fnagi.2022.979119]
30. Lin X, Liu F, Wang B, et al. Subjective cognitive decline may be associated with post-operative delirium in patients undergoing total hip replacement: The PNDABLE study. Front Aging Neurosci 2021;13:680672. [doi: 10.3389/fnagi.2021.680672]
31. Segernäs A, Skoog J, Ahlgren Andersson E, Almerud Österberg S, Thulesius H, Zachrisson H. Prediction of postoperative delirium after cardiac surgery with a quick test of cognitive speed, Mini-Mental State Examination and Hospital Anxiety and Depression Scale. Clin Interv Aging 2022;17:359-368. [doi: 10.2147/CIA.S350195] [Medline: 35400995]
32. Veliz-Reissmüller G, Agüero Torres H, van der Linden J, Lindblom D, Eriksdotter Jönhagen M. Pre-operative mild cognitive dysfunction predicts risk for post-operative delirium after elective cardiac surgery. Aging Clin Exp Res 2007 Jun;19(3):172-177. [doi: 10.1007/BF03324686] [Medline: 17607083]
33. Cao SJ, Chen D, Yang L, Zhu T. Effects of an abnormal mini-mental state examination score on postoperative outcomes in geriatric surgical patients: a meta-analysis. BMC Anesthesiol 2019 May 15;19(1):74. [doi: 10.1186/s12871-019-0735-5] [Medline: 31092206]
34. Sadeghirad B, Dodsworth BT, Schmutz Gelsomino N, et al. Perioperative factors associated with postoperative delirium in patients undergoing noncardiac surgery: An individual patient data meta-analysis. JAMA Netw Open 2023 Oct 2;6(10):e2337239. [doi: 10.1001/jamanetworkopen.2023.37239] [Medline: 37819663]
35. Peden CJ, Miller TR, Deiner SG, Eckenhoff RG, Fleisher LA, Members of the Perioperative Brain Health Expert Panel. Improving perioperative brain health: an expert consensus review of key actions for the perioperative care team. Br J Anaesth 2021 Feb;126(2):423-432. [doi: 10.1016/j.bja.2020.10.037] [Medline: 33413977]
36. Jin Z, Hu J, Ma D. Postoperative delirium: perioperative assessment, risk reduction, and management. Br J Anaesth 2020 Oct;125(4):492-504. [doi: 10.1016/j.bja.2020.06.063] [Medline: 32798069]
37. Menzenbach J, Kirfel A, Guttenthaler V, et al. PRe-Operative Prediction of postoperative DElirium by appropriate SCreening (PROPDESC) development and validation of a pragmatic POD risk screening score based on routine preoperative data. J Clin Anesth 2022 Jun;78:110684. [doi: 10.1016/j.jclinane.2022.110684] [Medline: 35190344]
38. Bishara A, Chiu C, Whitlock EL, et al. Postoperative delirium prediction using machine learning models and preoperative electronic health record data. BMC Anesthesiol 2022 Jan 3;22(1):8. [doi: 10.1186/s12871-021-01543-y] [Medline: 34979919]
39. Chen H, Mo L, Hu H, Ou Y, Luo J. Risk factors of postoperative delirium after cardiac surgery: a meta-analysis. J Cardiothorac Surg 2021 Apr 26;16(1):113. [doi: 10.1186/s13019-021-01496-w] [Medline: 33902644]
40. Bramley P, McArthur K, Blayney A, McCullagh I. Risk factors for postoperative delirium: An umbrella review of systematic reviews. Int J Surg 2021 Sep;93:106063. [doi: 10.1016/j.ijsu.2021.106063] [Medline: 34411752]
41. Sánchez A, Thomas C, Deeken F, et al. Patient safety, cost-effectiveness, and quality of life: reduction of delirium risk and postoperative cognitive dysfunction after elective procedures in older adults-study protocol for a stepped-wedge cluster randomized trial (PAWEL Study). Trials 2019 Jan 21;20(1):71. [doi: 10.1186/s13063-018-3148-8] [Medline: 30665435]
42. Deeken F, Sánchez A, Rapp MA, et al. Outcomes of a delirium prevention program in older persons after elective surgery: A stepped-wedge cluster randomized clinical trial. JAMA Surg 2022 Feb 1;157(2):e216370. [doi: 10.1001/jamasurg.2021.6370] [Medline: 34910080]
43. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126. [doi: 10.2196/jmir.1923] [Medline: 22209829]
44. Cuschieri S. The STROBE guidelines. Saudi J Anaesth 2019 Apr;13(Suppl 1):S31-S34. [doi: 10.4103/sja.SJA_543_18] [Medline: 30930717]

XSL•FO
RenderX

45. Dong Y, Peng CYJ. Principled missing data methods for researchers. Springerplus 2013 Dec;2(1):222. [doi: 10.1186/2193-1801-2-222] [Medline: 23853744]
46. Jameson JL, Fauci AS, Kasper DL, Hauser SL, Longo DL, Loscalzo J, editors. Harrison's Principles of Internal Medicine: New York, NY: McGraw-Hill Education; 2018.
47. Wyczalkowska-Tomasik A, Czarkowska-Paczek B, Zielenkiewicz M, Paczek L. Inflammatory markers change with age, but do not fall beyond reported normal ranges. Arch Immunol Ther Exp (Warsz) 2016 Jun;64(3):249-254. [doi: 10.1007/s00005-015-0357-7] [Medline: 26283530]
48. Chawla N, Bowyer K, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. ArXiv. Preprint posted online on 1813. [doi: 10.1613/jair.953]
49. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in Python. J Mach Learn Res 2011;12:2825-2830 [FREE Full text]
50. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. . Preprint posted online on 2016.
51. Rutherford S, Barkema P, Tso IF, et al. Evidence for embracing normative modeling. Elife 2023 Mar 13;12:e85082. [doi: 10.7554/eLife.85082] [Medline: 36912775]
52. Ojala M, Garriga GC. Permutation tests for studying classifier performance. Presented at: 2009 Ninth IEEE International Conference on Data Mining; Dec 6-9, 2009; Miami Beach, FL, USA. [doi: 10.1109/ICDM.2009.108]
53. Lundberg S, Lee SI. A unified approach to interpreting model predictions. . Preprint posted online on 2017.
54. Singh M, Sahhar M, Nassar JE, et al. Analysis of delirium risk assessment tools for prediction of postoperative delirium following lumbar spinal fusion. Spine (Phila Pa 1976) 2025 Jan 30;9900. [doi: 10.1097/BRS.0000000000005271] [Medline: 39882654]
55. Vreeswijk R, Kalisvaart I, Maier AB, Kalisvaart KJ. Development and validation of the delirium risk assessment score (DRAS). Eur Geriatr Med 2020 Apr;11(2):307-314. [doi: 10.1007/s41999-019-00287-w] [Medline: 32297197]
56. Freter SH, Dunbar MJ, MacLeod H, Morrison M, MacKnight C, Rockwood K. Predicting post-operative delirium in elective orthopaedic patients: the Delirium Elderly At-Risk (DEAR) instrument. Age Ageing 2005 Mar;34(2):169-171. [doi: 10.1093/ageing/afh245] [Medline: 15713861]
57. O'Neal JB, Billings FT 4th, Liu X, et al. Risk factors for delirium after cardiac surgery: a historical cohort study outlining the influence of cardiopulmonary bypass. Can J Anaesth 2017 Nov;64(11):1129-1137. [doi: 10.1007/s12630-017-0938-5] [Medline: 28718100]
58. Levin P. Chapter 223 - postoperative delirium. In: Atlee JL, editor. Complications in Anesthesia: Philadelphia: W.B. Saunders; 2007:888-889.
59. Kaźmierski J, Miler P, Pawlak A, et al. Lower preoperative verbal memory performance is associated with delirium after coronary artery bypass graft surgery: A prospective cohort study. Arch Clin Neuropsychol 2023 Jan 21;38(1):49-56. [doi: 10.1093/arclin/acac064] [Medline: 35915987]
60. Butz M, Meyer R, Gerriets T, et al. Increasing preoperative cognitive reserve to prevent postoperative delirium and postoperative cognitive decline in cardiac surgical patients (INCORE): Study protocol for a randomized clinical trial on cognitive training. Front Neurol 2022;13:1040733. [doi: 10.3389/fneur.2022.1040733] [Medline: 36578306]
61. Moller J, Cluitmans P, Rasmussen L, et al. Long-term postoperative cognitive dysfunction in the elderly: ISPOCD1 study. The Lancet 1998 Mar;351(9106):857-861. [doi: 10.1016/S0140-6736(97)07382-0]
62. Weiss Y, Zac L, Refaeli E, et al. Preoperative Cognitive Impairment and Postoperative Delirium in Elderly Surgical Patients: A Retrospective Large Cohort Study (The CIPOD Study). Ann Surg 2023 Jul 1;278(1):59-64. [doi: 10.1097/SLA.0000000000005657] [Medline: 35913053]
63. Scholz AFM, Oldroyd C, McCarthy K, Quinn TJ, Hewitt J. Systematic review and meta-analysis of risk factors for postoperative delirium among older patients undergoing gastrointestinal surgery. Br J Surg 2016 Jan;103(2):e21-e28. [doi: 10.1002/bjs.10062] [Medline: 26676760]
64. Ramspek CL, Steyerberg EW, Riley RD, et al. Prediction or causality? A scoping review of their conflation within current observational research. Eur J Epidemiol 2021 Sep;36(9):889-898. [doi: 10.1007/s10654-021-00794-w] [Medline: 34392488]
65. Zhao J, Liang G, Hong K, et al. Risk factors for postoperative delirium following total hip or knee arthroplasty: A meta-analysis. Front Psychol 2022;13:993136. [doi: 10.3389/fpsyg.2022.993136]
66. Frei BW, Woodward KT, Zhang MY, et al. Considerations for clock drawing scoring systems in perioperative anesthesia settings. Anesth Analg 2019;128(5):e61-e64. [doi: 10.1213/ANE.0000000000004105]
67. Fryer D, Strumke I, Nguyen H. Shapley values for feature selection: The good, the bad, and the axioms. IEEE Access 2021;9:144352-144360. [doi: 10.1109/ACCESS.2021.3119110]
68. Petch J, Di S, Nelson W. Opening the black box: The promise and limitations of explainable machine learning in cardiology. Can J Cardiol 2022 Feb;38(2):204-213. [doi: 10.1016/j.cjca.2021.09.004] [Medline: 34534619]
69. Oh S, Park Y, Cho KJ, Kim SJ. Explainable machine learning model for glaucoma diagnosis and its interpretation. Diagnostics (Basel) 2021 Mar 13;11(3):510. [doi: 10.3390/diagnostics11030510] [Medline: 33805685]
70. Amparore E, Perotti A, Bajardi P. To trust or not to trust an explanation: using LEAF to evaluate local linear XAI methods. PeerJ Comput Sci 2021;7:e479. [doi: 10.7717/peerj-cs.479] [Medline: 33977131]

XSL•FO
RenderX

71. Bordt S, von Luxburg U. From shapley values to generalized additive models and back. arXiv. URL: https://arxiv.org/abs/2209.04012 [accessed 2025-08-08] [doi: 10.48550/arXiv.2209.04012]
72. Sokolova E, von Rhein D, Naaijen J, et al. Handling hybrid and missing data in constraint-based causal discovery to study the etiology of ADHD. Int J Data Sci Anal 2017;3(2):105-119. [doi: 10.1007/s41060-016-0034-x] [Medline: 28691055]
73. Moon KJ, Son CS, Lee JH, Park M. The development of a web-based app employing machine learning for delirium prevention in long-term care facilities in South Korea. BMC Med Inform Decis Mak 2022 Aug 17;22(1):220. [doi: 10.1186/s12911-022-01966-8] [Medline: 35978303]
74. MHM-lab/PAWEL-delirium-prediction. GitHub. URL: https://github.com/MHM-lab/PAWEL-Delirium-Prediction [accessed 2025-08-08]

Abbreviations

ASA: American Society of Anesthesiologists
AUC: area under the receiver operating characteristic curve
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
CRP: C-reactive protein
I-CAM: Confusion Assessment Method
MMSE: Mini-Mental State Examination
MoCA: Montreal Cognitive Assessment
PAWEL: Patientensicherheit, Wirtschaftlichkeit und Lebensqualität bei elektiven Operationen
PAWEL-R: Patientensicherheit, Wirtschaftlichkeit und Lebensqualität bei elektiven Operationen–Risk
PHQ-4: Patient Health Questionnaire-4
POD: postoperative delirium
SHAP: Shapley Additive Explanations
SMI: Subjective Memory Impairment
STROBE: Strengthening the Reporting of Observational studies in Epidemiology
TMT: Trail Making Test

Edited by J Job; submitted 24.10.24; peer-reviewed by CR King, H Frohnhofen; revised version received 02.06.25; accepted 26.06.25; published 19.08.25.

Please cite as:
Wu SCJ, Sharma N, Bauch A, Yang HC, Hect JL, Thomas C, Wagner S, Förstner BR, von Arnim CAF, Kaufmann T, Eschweiler GW, Wolfers T, PAWEL Study Group
Predicting Postoperative Delirium in Older Patients Before Elective Surgery: Multicenter Retrospective Cohort Study
JMIR Aging 2025;8:e67958
URL: https://aging.jmir.org/2025/1/e67958
doi:10.2196/67958

© Shun-Chin Jim Wu, Nitin Sharma, Anne Bauch, Hao-Chun Yang, Jasmine L Hect, Christine Thomas, Sören Wagner, Bernd R Förstner, Christine A F von Arnim, Tobias Kaufmann, Gerhard W Eschweiler, Thomas Wolfers, PAWEL Study Group. Originally published in JMIR Aging (https://aging.jmir.org), 19.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Cognitive Functioning Trajectories and Their Association With Mental Health in Older Adults With Hypertension: Secondary Analysis of the Chinese Longitudinal Healthy Longevity Survey Data

Peiyun Wu, MS; Cheng Lin, MS; Jing Yang, MS; Zhizhuo Wang, MS, OTD

Department of Rehabilitation Medicine, School of Health, Fujian Medical University, No.1 Xuefu North Road, University Town, Fuzhou, China

Corresponding Author:
Zhizhuo Wang, MS, OTD
Department of Rehabilitation Medicine, School of Health, Fujian Medical University, No.1 Xuefu North Road, University Town, Fuzhou, China

Abstract

Background: There is a lack of research on the trajectories of cognitive functioning in older people with hypertension, as well as possible contributing factors and correlations between cognitive functioning and mental health.

Objective: This study aimed to conduct a secondary longitudinal analysis to examine cognitive functioning trajectories and their associated factors in older people with hypertension.

Methods: The data used in our study were retrieved from the Chinese Longitudinal Healthy Longevity Survey. The Chinese version of the Mini-Mental State Examination was used to assess cognitive functioning. The trajectories of cognitive functioning of older individuals with hypertension were determined by using group-based trajectory modeling. The binary logistic regression analyses were performed to examine how participant factors affected the trajectories of cognitive functioning in older individuals with hypertension. The relationships between cognitive functioning and mental health were investigated using multivariable linear regression models. Data were analyzed using SPSS (version 20.0; SPSS Inc) and Stata (version 16.0; StataCorp LLC).

Results: A total of 642 older people with hypertension were included. Cognitive functioning was categorized into 2 trajectories according to group-based trajectory modeling: “rapid decline” (48/642, 7.9%) and “slow decline” (594/642, 92.1%). Binary logistic regression results showed that older adults with hypertension aged equal to or greater than 80 years had an elevated risk of rapid decline in cognitive functioning (odds ratio 5.484, 95% CI 2.365 - 12.719), and higher score in mental health was the protector of rapid decline in cognitive functioning during the following 13 years (odds ratio 0.918, 95% CI 0.852 - 0.988). In the unadjusted model, mental health was positively associated with cognitive functioning (β=.246, 95% CI 0.125 - 0.234, *P*<.001), and this association was maintained after partial or complete adjustment for covariates (β=.159, 95% CI 0.059 - 0.174, *P*<.001; β=.138, 95% CI 0.043 - 0.158, *P*=.001). Subgroup analyses by age showed that this positive correlation was only seen in the 60 - 69 years age group (β=.183, 95% CI 0.036 - 0.193, *P*=.004), while subgroup analyses by sex revealed that the association between these 2 indicators was no longer presented among males in the fully adjusted model (β=.082, 95% CI −0.024 to 0.119, *P*=.19), and BMI fully stratification demonstrated this association persisted in the healthy weight group (β=.125, 95% CI 0.039 - 0.210, *P*=.004).

Conclusions: Our research showed that the decline in cognitive functioning is associated with lower mental health and occurs more rapidly in older adults with hypertension who are older than 80 years.

(JMIR Aging 2025;8:e74916) doi:10.2196/74916

KEYWORDS

cognitive functioning; mental health; hypertension; older people; cognitive functioning trajectories; older people with hypertension

Introduction

According to the summary of the China Cardiovascular Health and Illness Report 2022, there will be 245 million individuals with hypertension in China, and the prevalence of cardiovascular illness is expected to continue rising. Currently, cardiovascular disease ranks first among all causes of death for Chinese citizens living in both urban and rural areas. Patients with hypertension account for 50% of deaths from coronary heart disease and stroke, and the disease has increased the burden on society and the populace [1]. One of the main chronic illnesses in China that has a significant impact on health care costs is hypertension, which also poses the greatest risk of stroke, heart disease, and renal disease [2].

XSL•FO
RenderX

Hypertension is well established as a key risk factor that raises the likelihood of vascular cognitive impairment and Alzheimer disease, which affects a considerable fraction of the adult and geriatric population [3]. Due to its disruption of the cerebral circulation's structural and functional integrity, hypertension primarily affects cerebral blood supply by causing vascular thinning and dysfunction as well as neurovascular uncoupling. Furthermore, because it compromises the blood-brain barrier, hypertension also increases the risk of neuroinflammation and amyloidosis. In addition, age worsens the body's ability to recover from cellular stress and maintain homeostasis, which exacerbates the negative consequences of hypertension on the cerebral vasculature [4]. Cognitive decline is closely associated with all of these factors.

There may also be a connection between poor mental health and cognitive deterioration in hypertension. According to recent research, there is evidence linking mental illnesses with hypertension, and chronic stress is acknowledged as a common but underappreciated risk factor for cardiovascular disease [5,6]. Another significant risk factor for cognitive deterioration that has been identified is depression. Studies have indicated a link between chronic depressed symptoms and decreased cognitive functioning in late middle age [7], and that depression is a risk factor for developing Alzheimer disease [8]. The link between the development of cardiovascular disease, cognitive decline, and poorer mental health is uncertain and needs to be researched in depth.

In light of the high incidence and unfavorable effects of hypertension, it is critical to investigate how the disease may affect cognitive trajectories and its connections to mental health, all of which may aid in the early detection and management of hypertension. By extracting pertinent data from the Chinese Longitudinal Healthy Longevity Survey (CLHLS), the study aimed to investigate the cognitive functioning trajectories, potential affecting factors, and link with mental health in older persons with hypertension.

Methods

Ethical Considerations

The CLHLS study was approved by the Research Ethics Committee of Peking University (IRB00001052-13074), and all written informed consent had been obtained from participants or their proxy respondents and the participants' privacy was protected. All data were fully anonymized to ensure participant confidentiality.

Study Design and Population

Data included in the study were retrieved from 5 waves of the CLHLS from 2005 to 2018. The CLHLS, initiated in 1998, is a prospective cohort designed to investigate factors influencing longevity among older adults. It uses a nationally representative sample drawn from 22 of China's 34 provinces. The study uses a multistage, nonequal target random sampling strategy: approximately half of all counties, county-level cities, and municipal districts within the participating provinces were randomly chosen as survey sites. These selected administrative areas were further stratified into "large sample" and "small sample" categories for data collection purposes. Qualified interviewers, centrally trained and assessed for competency, conducted the surveys. Data gathering involved direct interviews with surviving participants (older adults) and proxy interviews with the adult children of decedent participants. For more information on the sampling design and data quality assessment, readers can refer to this literature [9].

Participants in the wave of 2005 as baseline involved 15,638 older people, and we excluded those who lacked necessary information on cognitive functioning and blood pressure in the following CLHLS 2008, 2011, 2014, and 2018 waves (n=14,404). This study focused only on older adults with hypertension. Therefore, of the 1234 participants who remained, we further excluded those who did not satisfy the diagnostic criteria of hypertension previously established in the literature (n=531), who missed the information on covariates (n=12), and those with unreliable data (n=49). Ultimately, a total of 642 participants were included in the analysis. The adequacy of our sample size (N=642) was rigorously evaluated using Green's power-based framework for regression analyses with 12 predictors [10]: for multivariable linear regression examining mental health-cognition associations, detecting medium effect sizes (f^2=0.15) required ≥128 participants for overall model significance ($N \geq L/f^2$), while testing individual coefficients (partial correlations) needed ≥65 participants ($N \geq 8/f^2+(m-1)$). As for binary logistic regression analyzing cognitive trajectory predictors, the Peduzzi criterion (10 events per predictor) necessitated ≥600 participants ($N \geq 10 \times m/p$). Our sample size of 642 substantially surpassed all these thresholds, providing over 99% power to detect medium effects at α of .05 (calculated via Cohen tables). Our fullest model (k=12 covariates) required n ≥116, well below our analytic sample. Therefore, the current sample size provides adequate statistical power for the core objectives. Based on the literature [11], hypertension in this study was defined as systolic blood pressure ≥140 mmHg or diastolic blood pressure ≥90 mmHg or self-reported diagnosis of hypertension in the hospital prior to enrollment.

Cognitive Functioning

Cognitive functioning was measured by the Chinese version of the Mini-Mental State Examination (C-MMSE), which evaluated 4 aspects of cognitive functioning, including orientation, calculation, recall, and language [12]. The C-MMSE comprised 13 questions with total scores ranging from 0 to 30. A higher score indicated better cognitive functioning. The psychometrics of the C-MMSE have been verified in previous studies [13,14].

Covariates

The selection of covariates was guided by established gerontological and epidemiological frameworks, aiming to control for potential confounding factors and capture key dimensions influencing health outcomes in older adults, particularly cognitive functioning [15]. The covariates included in this study were as indicated below: age category (60-69 y, 70-79 y, and ≥80 y), sex (male or female), smoking status (current, previous, and never), drinking status (current, previous, and never), exercise (current, previous, and never), educational level (zero year of schooling, 1-6 y of schooling, and ≥7 y of

XSL•FO
RenderX

schooling), marital status (married and living with spouse, separated/divorced/widowed/never married), BMI (<18.5 as underweight, 18.5-23.9 as healthy weight, 24-27.9 as overweight, and ≥28 as obesity), and quality of sleep (bad, general, and good). Among these, smoking, drinking, and exercise were self-reported by participants, who could report their current, previous, or never status for each factor. Social participation was evaluated by 10 activities in the CLHLS questionnaire according to the previous study [16]. Activities of daily living (ADL) were measured by 6 items including bathing, dressing, toileting, indoor transferring, continence, and feeding, and instrumental activities of daily living (IADL) were assessed by 8 items including going out to visit neighbors, shopping, making food, washing clothes, walking 1 kilometer, carrying 5 kg weight, crouching and standing 3 times, and taking public transport [16]. Mental health was examined by 7 items, such as "look on the bright side of thing," "keep my belongings neat and clean," "feel fearful or anxious," "feel lonely and isolated," "make own decision," "feel useless with age," and "be happy as younger." The negative-oriented items were coded reversely. The maximum score is 28, and a higher score indicates a greater degree of mental health [17].

Statistical Analysis

The data analyses were administered by using the SPSS (version 20.0; SPSS Inc) for general statistical analyses. Stata (version 16.0; StataCorp LLC) was used for group-based trajectory modeling (GBTM) to identify the trajectories of cognitive functioning in older adults with hypertension from 2005 to 2018. The normally and abnormally distributed continuous data were manifested as the mean (SD) and median (IQR), respectively. The categorical data were expressed as a number (%). The *t* test was used to compare continuous variables with a normal distribution; otherwise, the Mann-Whitney *U* rank sum test was used. Pearson chi-square or Fisher exact test was applied to compare categorical variables. The best-fitting model was determined based on the following parameters and principles [18,19]: (1) the Bayesian Information Criterion closer to 0; (2) lower Akaike Information Criterion values; (3) the odds of correct classification higher than 5 for each group; (4) the entropy values closer to 1; (5) the average posterior probabilities of group membership greater than 0.7 for each group; and (6) the presence of sufficient number of individuals in each group (at least 5%). After confirming the optimal number of trajectory groups and appropriate trajectory shapes, we visually delineated the trajectories of cognitive functioning. Binary logistic regression analyses (Forward: LR) were conducted to analyze the influence of participants' characteristics on the trajectories of cognitive functioning of older adults with hypertension. Multivariable linear regression models were applied to examine the associations between mental health and cognitive functioning. Model 1 was crudely adjusted (unadjusted); model 2 was adjusted for age, sex, drinking status, smoking status, exercise, educational level, marital status, BMI, and quality of sleep; and model 3 was additionally adjusted for social participation, ADL, and IADL. Subgroup analyses stratified by either age or sex were performed with all covariates except those adjusted. $P<.05$ was considered statistically significant differences.

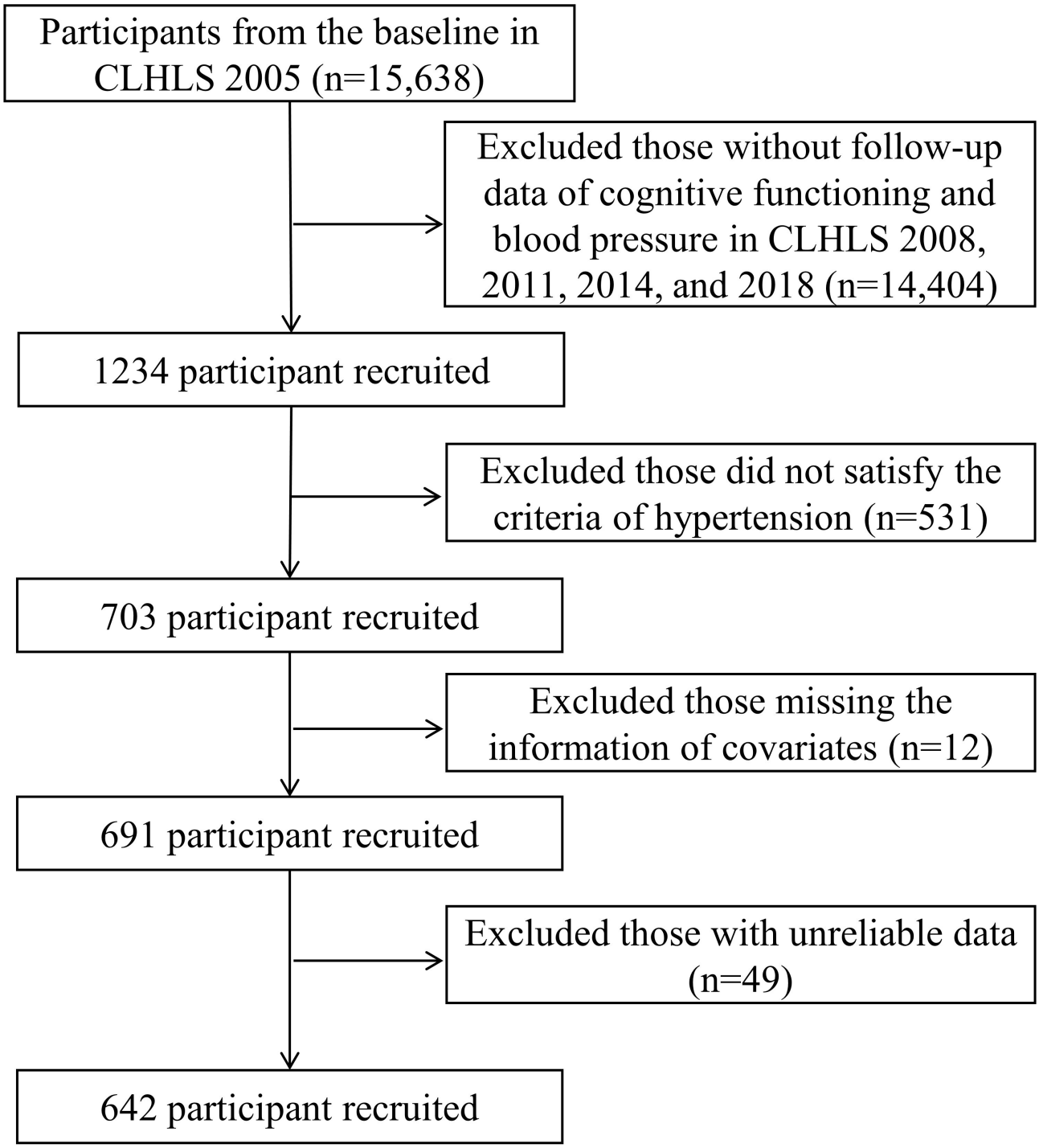
Results

Baseline Characteristics of Study Population

A total of 642 older adults with hypertension were included in this study (Figure 1). Of the included participants, 292 (45.48%) participants were men, and 350 (54.52%) participants were women (as seen in Table 1). The age distribution of older adults aged 60-69, 70-79, and ≥80 years was 41.12% (n=264), 45.95% (n=295), and 12.93% (n=83), respectively. For BMI, participants categorized as underweight, healthy weight, overweight, and obesity were 171 (26.64%), 346 (53.89%), 93 (14.49%), and 32 (4.98%), respectively. The median of social participation, ADL, IADL, and mental health was 17 (14-20), 12 (12-12), 16 (15-16), and 20 (17-23), respectively. In terms of blood pressure, the median of systolic blood pressure and diastolic blood pressure was 130 (120-130) and 80 (80-85) mmHg, respectively. From the cognitive perspective, participants rated a median score of 29 (27-30) in the baseline assessment (the cohort of 2005).

XSL•FO
RenderX

Figure 1. The flowchart of participants' inclusion and exclusion process. CLHLS: Chinese Longitudinal Healthy Longevity Survey.



XSL•FO
RenderX

Table . Comparison of baseline characteristics between cognitive functioning trajectory groups.

Variables	Total participants (N=642)	Trajectory group		*P* value
		Group 1 (Rapid decline, n=48)	Group 2 (Slow decline, n=594)	
Age category (years), n (%)				<.001
60-69	264 (41.12)	10 (20.83)	254 (42.76)	
70-79	295 (45.95)	22 (45.84)	273 (45.96)	
≥80	83 (12.93)	16 (33.33)	67 (11.28)	
Sex, n (%)				.51
Male	292 (45.48)	24 (50.00)	268 (45.12)	
Female	350 (54.52)	24 (50.00)	326 (54.88)	
Smoking status, n (%)				.86
Current	155 (24.14)	11 (22.92)	144 (24.24)	
Previous	78 (12.15)	7 (14.58)	71 (11.95)	
Never	409 (63.71)	30 (62.50)	379 (63.81)	
Drinking status, n (%)				.95
Current	168 (26.17)	13 (27.08)	155 (26.09)	
Previous	62 (9.66)	4 (8.34)	58 (9.77)	
Never	412 (64.17)	31 (64.58)	381 (64.14)	
Exercise, n (%)				.47
Current	253 (39.41)	15 (31.25)	238 (40.07)	
Previous	51 (7.94)	4 (8.33)	47 (7.91)	
Never	338 (52.65)	29 (60.42)	309 (52.02)	
Educational level, n (%)				.04
Zero year of schooling	308 (47.98)	29 (60.42)	279 (46.97)	
1-6 years of schooling	240 (37.38)	17 (35.42)	223 (37.54)	
≥7 years of schooling	94 (14.64)	2 (4.16)	92 (15.49)	
Marital status, n (%)				.13
Married and living with spouse	229 (35.67)	22 (45.83)	207 (34.85)	
Separated/divorced widowed/never married	413 (64.33)	26 (54.17)	387 (65.15)	
BMI (kg/m^2), n (%)				.22
<18.5 (underweight)	171 (26.64)	19 (39.58)	152 (25.59)	
18.5-23.9 (healthy weight)	346 (53.89)	22 (45.84)	324 (54.54)	
24-27.9 (overweight)	93 (14.49)	6 (12.50)	87 (14.65)	
≥28 (obesity)	32 (4.98)	1 (2.08)	31 (5.22)	
Quality of sleep, n (%)				.81
Bad	211 (32.87)	16 (33.34)	195 (32.83)	
General	316 (49.22)	25 (52.08)	291 (48.99)	
Good	115 (17.91)	7 (14.58)	108 (18.18)	
Social participation, median (IQR)	17 (14-20)	17.00 (13.25-20.00)	17.00 (14.00-20.25)	.17
ADL[a], median (IQR)	12 (12-12)	12.00 (12.00-12.00)	12.00 (12.00-12.00)	.27
IADL[b], median (IQR)	16 (15-16)	16.00 (13.00-16.00)	16.00 (15.00-16.00)	.01

XSL•FO
RenderX

Variables	Total participants (N=642)	Trajectory group		*P* value
		Group 1 (Rapid decline, n=48)	Group 2 (Slow decline, n=594)	
Mental health, median (IQR)	20 (17-23)	18.00 (15.25-21.00)	20 (17.00-23.00)	.007

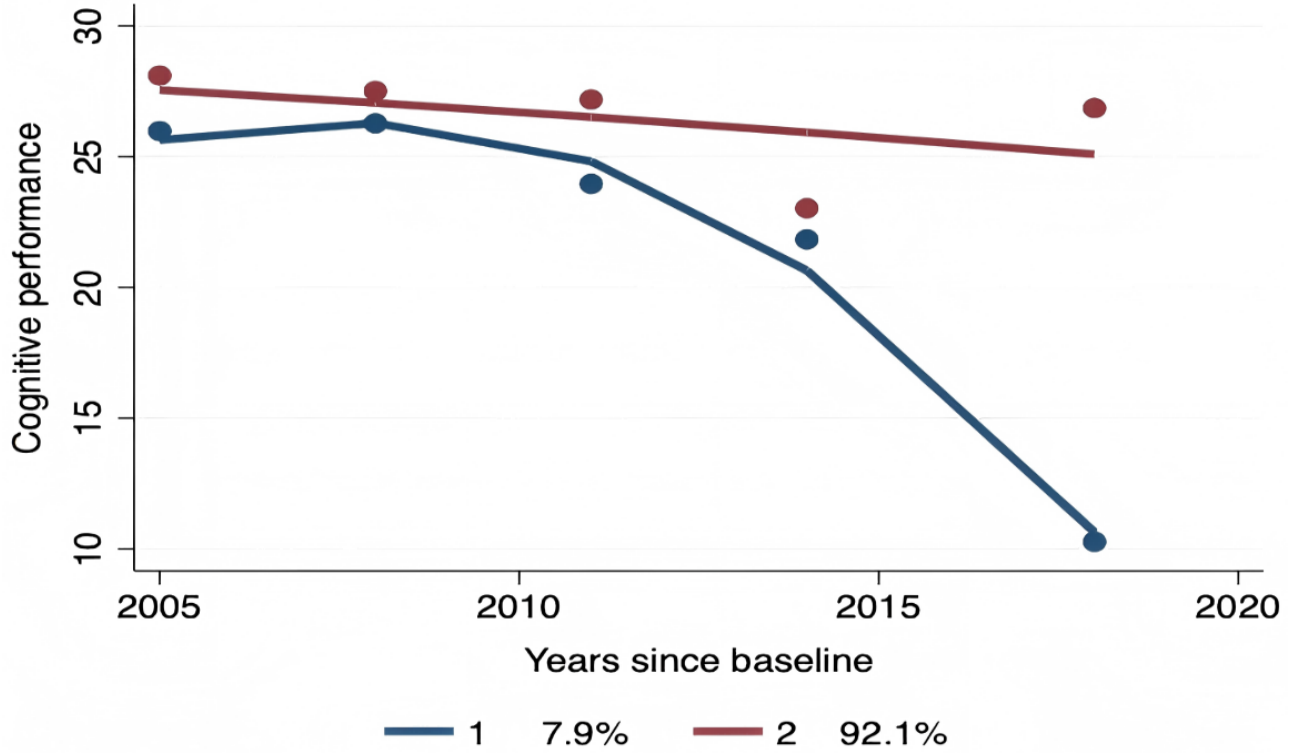
[a]ADL: activities of daily living.

[b]IADL: instrumental activities of daily living.

Trajectories of Cognitive Performance for Older People With Hypertension

According to GBTM, the cognitive functioning was divided into 2 trajectories. As shown in Figure 2, the trajectory of group 1 comprised 48 (7.9%) older adults with hypertension and was defined as “rapid decline”. The second group was estimated to include 594 (92.1%) older people with hypertension labeled as “slow decline.”

Figure 2. Trajectories of cognitive functioning for older people with hypertension.



The average posterior probabilities for these 2 trajectory groups were 0.91 and 0.99, respectively, which were higher than the recommended cut-off of 0.70. The odds of correct classification values were 114.58 for group 1 and 6.82 for group 2, with the lower value (6.82) surpassing the recommended threshold of 5. Model fit indices further supported this solution: the Bayesian Information Criterion was –8107.51, the Akaike Information Criterion was –8089.65, and entropy was 0.937.

Comparison of Cognitive Functioning Trajectory Groups

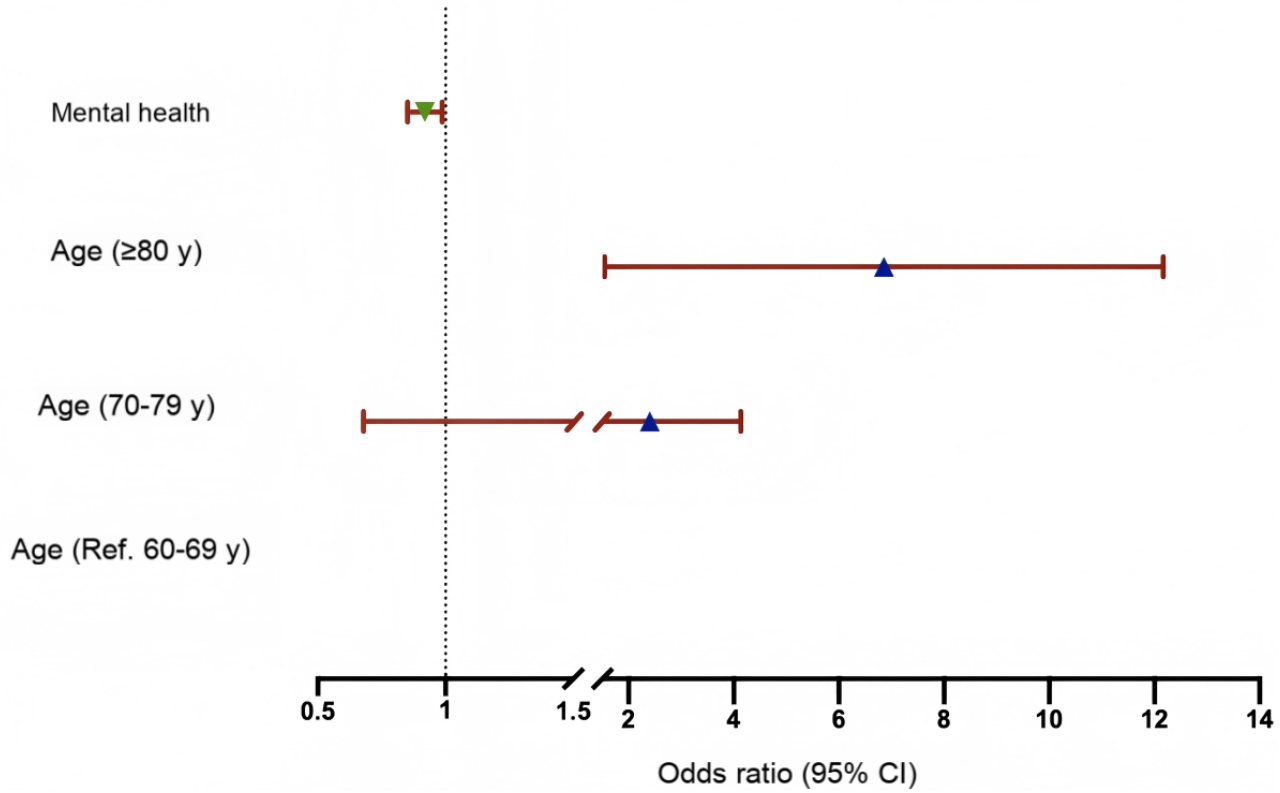
As demonstrated in Table 1, there were significant differences between the cognitive functioning trajectory groups in age (χ^2_2=21.901, *P*<.001), educational level (χ^2_2=5.979, *P*=.04), and IADL (*z*=−2.544, *P*=.01). Compared with group 1, the median of mental health was significantly higher (*z*=−2.691, *P*=.007). However, no statistically significant difference was found between groups in sex (χ^2_1=0.427, *P*=.51), BMI (χ^2_3=4.227, *P*=.22), smoking status (χ^2_2=0.297, *P*=.86), drinking status (χ^2_2=0.112, *P*=.95), exercise (χ^2_2=1.527, *P*=.47), marital status (χ^2_1=2.335, *P*=.13), quality of sleep (χ^2_2=0.411, *P*=.81), social participation (*z*=−1.363, *P*=.17), and ADL (*z*=−1.113, *P*=.27).

Factors Influencing the Cognitive Functioning Trajectories

Figure 3 presented the results of binary logistic regressions between group 1 and group 2. Older adults with hypertension aged 80 years or older had an elevated risk of rapid decline in cognitive functioning (odds ratio 5.484, 95% CI 2.365 - 12.719). However, a higher score in mental health was the protector of rapid decline in cognitive functioning during the following 13 years (odds ratio 0.918, 95% CI 0.852 - 0.988).

XSL•FO
RenderX

Figure 3. Logistic regression forest plot of the factors associated with the cognitive functioning trajectory groups.



Association Between Mental Health and Cognitive Functioning

As shown in Table 2, the multivariable linear regression model was used to examine the association between mental health and cognitive functioning in older people with hypertension. In the unadjusted model, mental health was positively associated with cognitive functioning (β=.246, 95% CI 0.125 - 0.234, *P*<.001). After adjusting the covariates partially (model 2) or fully (model 3), the positive relationship also remained unchanged (β=.159, 95% CI 0.059 - 0.174, *P*<.001; β=.138, 95% CI 0.043 - 0.158, *P*=.001).

XSL•FO
RenderX

Table . Association between mental health and cognitive functioning.

Mental health with cognitive functioning	Model 1		Model 2		Model 3	
	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value
Overall	0.246 (0.125 to 0.234)	<.001	0.159 (0.059 to 0.174)	<.001	0.138 (0.043 to 0.158)	.001
Stratified by age (years)						
60-69	0.255 (0.086 to 0.233)	<.001	0.199 (0.046 to 0.202)	.002	0.183 (0.036 to 0.193)	.004
70-79	0.197 (0.062 to 0.228)	.001	0.119 (−0.001 to 0.176)	.05	0.106 (−0.012 to 0.168)	.09
≥80	0.322 (0.102 to 0.480)	.003	0.239 (−0.020 to 0.452)	.07	0.107 (−0.127 to 0.319)	.39
Stratified by sex						
Male	0.215 (0.059 to 0.190)	<.001	0.128 (0.002 to 0.146)	.04	0.082 (−0.024 to 0.119)	.19
Female	0.238 (0.110 to 0.275)	<.001	0.179 (0.058 to 0.231)	.001	0.165 (0.047 to 0.219)	.002
Stratified by BMI (kg/m²)						
<18.5 as underweight	0.186 (0.081 to 0.292)	.001	0.126 (0.015 to 0.237)	.03	0.108 (−0.04 to 0.220)	.06
18.5-23.9 as healthy weight	0.184 (0.100 to 0.269)	<.001	0.142 (0.054 to 0.230)	.002	0.125 (0.039 to 0.210)	.004
24-27.9 as overweight	0.068 (−0.037 to 0.132)	.20	0.053 (−0.060 to 0.166)	.36	0.067 (−0.052 to 0.186)	.26
≥28 as obesity	0.052 (−0.127 to 0.231)	.56	0.061 (−0.180 to 0.303)	.60	−0.017 (−0.261 to 0.228)	.89

Subgroup analysis stratified by age, sex, and BMI was conducted to further evaluate the relationship between mental health and cognitive functioning. After adjusting for all confounders except for age, subgroup analysis by age showed that the positive association between these 2 variables was merely in the age group of 60-69 years (β=.183, 95% CI 0.036 - 0.193, *P*=.004). Meanwhile, subgroup analysis by sex revealed that although the positive relationship between these 2 indicators was shown in model 1 and model 2 for both males and females, the association was no longer presented in model 3 for males (β=.082, 95% CI −0.024 to 0.119, *P*=.19). Additionally, subgroup analysis by BMI demonstrated that in the partially adjusted model, mental health was significantly associated with cognitive functioning in underweight (β=.126, 95% CI 0.015 - 0.237, *P*=.03) and healthy weight groups (β=.142, 95% CI 0.054 - 0.230, *P*=.002). After full adjustment, this association persisted only in the healthy weight group (β=.125, 95% CI 0.039 - 0.210, *P*=.004), but not in underweight, overweight, or obesity groups.

Discussion

Principal Findings

This longitudinal study identified 2 distinct cognitive trajectories among Chinese older adults with hypertension using GBTM: a rapid decline trajectory (7.9%) and a predominant slow decline trajectory (92.1%). Key factors significantly differentiating these trajectories included age, education level, IADL, and mental health. Notably, age of 80 years or older emerged as a critical risk factor for rapid decline. Furthermore, the protective association between mental health and cognitive functioning was found to be significantly moderated by BMI, remaining robust only in normal-weight individuals after comprehensive adjustment for confounders.

Trajectory Classification and Risk Profiles

As far as research on the cognitive functioning trajectories of older persons with hypertension, possible contributing factors, and their connection to mental health is concerned, this is the first longitudinal survey of its kind. A total of 642 older adults with hypertension were included in our study. We distinguished between 2 categories of cognitive trajectories of patients with hypertension based on GBTM analysis: “rapid decline” (7.9%) and “slow decline” (92.1%). These 2 groups differed significantly in terms of age, education, and IADL, and the group with higher mental health scores showed a trend towards slow decline. According to one study, having hypertension was linked to worse cognitive trajectories, while treating and controlling the condition was linked to improved cognitive trajectories [20]. In a different study, Zhang et al [21] identified trajectory categories of executive functioning and situational memory in cognitive functioning. Their findings showed that patients with hypertension who had higher education levels,

XSL•FO
RenderX

moderate sleep at night, fewer depressive symptoms, and at least one alcoholic beverage per month were more likely to be in the optimal group for stabilized executive functioning; as time went on, participants with higher BMI, adequate daytime naps, and higher education levels were also more likely to be in the optimal group for stabilized executive functioning. Individuals with stable situational memory over time [21]. This is in keeping with our findings, which showed variations in age and education level between those who saw a quick and slow decrease.

Stratified Effects in Key Subgroups

We additionally propose possible causes for this interparticipant variability based on the binary logistic regression results. Stratified by age, we found that patients with hypertension aged 80 years or older had a higher risk of rapid cognitive decline compared with other age groups. Hypertension is known to affect a large portion of the adult and older population and is an important risk factor for vascular cognitive impairment and dementia in later life. Hypertension impairs cognitive functioning by altering or destroying the structural and functional integrity of the cerebral circulation, which in turn is exacerbated by aging [22]. This suggests that the age factor plays an important role in the risk of cognitive decline in patients with hypertension. As we can see from the logistic regression forest plot of the factors associated with cognitive functioning, higher mental health scores emerged as a protective factor against cognitive functioning decline. This illustrates the link between cognitive decline in hypertension and poor mental health. The study by Bai et al [23] confirmed that interventions targeting depression have the potential to alleviate the deterioration of cognitive symptoms. This corroborates the results obtained in our study and provides a theoretical basis for identifying risk factors and alleviating cognitive functioning decline.

We used multivariate linear regression models to examine the association between the trajectories of cognitive ability and mental health. According to our research, there is a positive correlation between mental health and cognitive capacity. This correlation persists even after controlling for confounders either fully (model 3) or partially (model 2). We performed age-, gender-, and BMI-stratified subgroup analyses to examine the connection between mental health and cognitive functioning in more detail. Age-specific subgroup analysis revealed that only the 60 - 69 years age group had a positive correlation between the 2, even after controlling for all other variables save age. Gender subgroup analysis revealed that in models 1 and 2, there were positive correlations between the 2 indicators for both males and females, but in model 3, there were no longer any for males. BMI subgroup analysis also revealed that mental health's cognitive protection was consistently significant only in normal-weight individuals. Effects attenuated in underweight and disappeared in overweight/obese groups, suggesting metabolic factors may override psychological benefits when BMI exceeds 23.9 kg/m^2. Clinically, mental health interventions should prioritize normal-weight elders with hypertension, while weight management precedes psychological therapies in higher BMI groups. This is in line with a study by Zhou et al [24] that found a negative correlation between serious depressive symptoms and lower cognitive function in an older Chinese community population. Guo et al [25] showed a substantial correlation between cognitive deterioration and older women with depression. Yeom et al [26] also revealed that the relationship between depressive symptoms and cognitive function is related to BMI. In summary, our findings indicate that the relationship between mental health and cognitive ability is significantly influenced by BMI and age, with normal-weight individuals and specific age groups showing stronger correlations even after controlling for confounders.

Conclusions

In summary, this population-based longitudinal study's data indicated that the cognitive trajectories of older Chinese persons with hypertension can be divided into 2 categories: rapid decline and moderate decline. Age was a significant factor, with hypertension patients older than 80 years having a larger chance of experiencing rapid cognitive decline than patients in other age groups; rapid cognitive decline was indicated by lower mental health scores. To determine the causal association between mental health and cognitive trajectory in older persons with hypertension, further longitudinal cohort studies should be carried out.

Strengths and Limitations

There are significant advantages to our study. As the first longitudinal study using trajectory analysis to examine older Chinese adults with hypertension, our work first clarifies changes in trajectories of cognitive ability, examines potential causes, and identifies correlations between trajectories of cognitive ability and mental health in this population. Second, we performed a thorough examination of cognitive functioning while keeping track of a sizable sample of middle-aged and older Chinese adults that was nationally representative. In order to better understand the fundamental causes of cognitive loss, we also fit the trajectory of cognitive aging using a cutting-edge statistical model (GBTM). It is crucial to acknowledge that this study has several limitations. First of all, it is challenging to explain causality in this observational study due to the cross-sectional methodology. Another limitation is the reliance on secondary data from the CLHLS. The variables and measurements were limited by the original survey design, potentially omitting unmeasured confounding factors. Moreover, we used a set of cognitive tests and questionnaires for cognitive assessment. These tools might not fully capture the complexity of cognitive functions or subtle variations in cognitive functioning. Finally, the reliance on self-reported measures may introduce recall bias and social desirability bias, particularly among older adults with cognitive impairment whose self-assessment accuracy could be compromised.

XSL•FO
RenderX

Acknowledgments

We sincerely express our gratitude to the colleagues and participants of CLHLS for their great effort. This study was funded by the Fujian Social Science Planning Project (grant number: FJ2024C094). The funding agencies had no role in the study design or conduct.

Data Availability

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

Authors' Contributions

PW was responsible for conceptualization, formal analysis, data curation, methodology, and writing the original draft. CL contributed to data curation and methodology, and participated in the review and editing of the manuscript. JY also contributed to data curation and methodology, and assisted with review and editing of the manuscript. ZW was involved in conceptualization and supervision, and contributed to the review and editing of the manuscript.

Conflicts of Interest

None declared.

References

1. China Cardiovascular Health and Disease Report Writing Group. Summary of the China cardiovascular health and disease report 2022. Chin J Circ 2023;38(6):583-612. [doi: 10.1097/CM9.0000000000002927] [Medline: 38018129]
2. The National Essential Public Health Service Program Office for Management of Hypertension in Primary Health Care, National Center for Cardiovascular Diseases, National Committee on Hypertension Management in Primary Health Care. National Guidelines for Prevention and Management of Hypertension at the Primary Level (2020 Edition) [Article in Chinese]. Chin J Circ 2021;36(3):209-220. [doi: 10.3969/j.issn.1000-3614.2021.03.001]
3. Scheltens P, Blennow K, Breteler MMB, et al. Alzheimer's disease. Lancet 2016 Jul 30;388(10043):505-517. [doi: 10.1016/S0140-6736(15)01124-1] [Medline: 26921134]
4. Santisteban MM, Iadecola C, Carnevale D. Hypertension, neurovascular dysfunction, and cognitive impairment. Hypertension 2023 Jan;80(1):22-34. [doi: 10.1161/HYPERTENSIONAHA.122.18085] [Medline: 36129176]
5. Osborne MT, Shin LM, Mehta NN, Pitman RK, Fayad ZA, Tawakol A. Disentangling the links between psychosocial stress and cardiovascular disease. Circ Cardiovasc Imaging 2020 Aug;13(8):e010931. [doi: 10.1161/CIRCIMAGING.120.010931] [Medline: 32791843]
6. Schaare HL, Blöchl M, Kumral D, et al. Associations between mental health, blood pressure and the development of hypertension. Nat Commun 2023 Apr 7;14(1):1953. [doi: 10.1038/s41467-023-37579-6] [Medline: 37029103]
7. Singh-Manoux A, Akbaraly TN, Marmot M, et al. Persistent depressive symptoms and cognitive function in late midlife: the Whitehall II study. J Clin Psychiatry 2010 Oct;71(10):1379-1385. [doi: 10.4088/JCP.09m05349gry] [Medline: 20584520]
8. Sinclair LI, Mohr A, Morisaki M, et al. Correction: is later-life depression a risk factor for Alzheimer's disease or a prodromal symptom: a study using post-mortem human brain tissue? Alzheimers Res Ther 2024 Feb 13;16(1):33. [doi: 10.1186/s13195-024-01404-z] [Medline: 38351042]
9. Zhang Q, Wu Y, Han T, Liu E. Changes in cognitive function and risk factors for cognitive impairment of the elderly in China: 2005–2014. IJERPH 2019;16(16):2847. [doi: 10.3390/ijerph16162847]
10. Green SB. How many subjects does it take to do a regression analysis. Multivariate Behav Res 1991 Jul 1;26(3):499-510. [doi: 10.1207/s15327906mbr2603_7] [Medline: 26776715]
11. Du J, Zhu G, Yue Y, Liu M, He Y. Blood pressure and hypertension prevalence among oldest-old in China for 16 year: based on CLHLS. BMC Geriatr 2019 Sep 9;19(1):248. [doi: 10.1186/s12877-019-1262-4] [Medline: 31500574]
12. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
13. Lv X, Li W, Ma Y, et al. Cognitive decline and mortality among community-dwelling Chinese older people. BMC Med 2019 Mar 15;17(1):63. [doi: 10.1186/s12916-019-1295-8] [Medline: 30871536]
14. Zeng Y, Feng Q, Hesketh T, Christensen K, Vaupel JW. Survival, disabilities in activities of daily living, and physical and cognitive functioning among the oldest-old in China: a cohort study. Lancet 2017 Apr 22;389(10079):1619-1629. [doi: 10.1016/S0140-6736(17)30548-2] [Medline: 28285816]
15. Li C, Zhu Y, Ma Y, Hua R, Zhong B, Xie W. Association of cumulative blood pressure with cognitive decline, dementia, and mortality. J Am Coll Cardiol 2022 Apr 12;79(14):1321-1335. [doi: 10.1016/j.jacc.2022.01.045] [Medline: 35393012]
16. Zhang C, Zhao Y, Chen X, et al. Trajectories of social participation and its predictors in older adults: based on the CLHLS cohorts from 2002 to 2018. Int J Environ Res Public Health 2023 Feb 27;20(5):4260. [doi: 10.3390/ijerph20054260] [Medline: 36901270]

XSL•FO
RenderX

17. Pan C, Kelifa MO, Liang J, Wang P. Joint trajectories of disability and related factors among older adults in China. Public Health (Fairfax) 2021 Oct;199:96-102. [doi: 10.1016/j.puhe.2021.08.018] [Medline: 34583202]
18. Yang J, Qu Y, Zhan Y, et al. Trajectories of antepartum depressive symptoms and birthweight: a multicenter and prospective cohort study. Psychiatry Clin Neurosci 2023 Dec;77(12):631-637. [doi: 10.1111/pcn.13590] [Medline: 37632723]
19. Bruce DG, Davis WA, Davis TME. Group-based trajectory modelling of changes in mobility over six years in type 2 diabetes: the fremantle diabetes study phase II. J Clin Med 2023 Jul 6;12(13):4528. [doi: 10.3390/jcm12134528] [Medline: 37445563]
20. Zhang B, Jiang S. Heterogeneity in longitudinal trajectories of cognitive performance among middle-aged and older individuals with hypertension: growth mixture modeling across an 8-year cohort study. Hypertens Res 2022 Jun;45(6):1037-1046. [doi: 10.1038/s41440-021-00829-5] [Medline: 34952952]
21. Zhang B, Hu M, Sun Y, Lin Y, Zhu C. Associations between the prevalence, treatment, control of hypertension and cognitive trajectories among Chinese middle-aged and older adults. Am J Geriatr Psychiatry 2022 Oct;30(10):1123-1134. [doi: 10.1016/j.jagp.2022.03.005] [Medline: 35466016]
22. Forouzanfar MH, Liu P, Roth GA, et al. Global burden of hypertension and systolic blood pressure of at least 110 to 115 mm Hg, 1990-2015. JAMA 2017 Jan 10;317(2):165-182. [doi: 10.1001/jama.2016.19043] [Medline: 28097354]
23. Bai W, Zhang J, Smith RD, et al. Inter-relationship between cognitive performance and depressive symptoms and their association with quality of life in older adults: a network analysis based on the 2017-2018 wave of Chinese Longitudinal Healthy Longevity Survey (CLHLS). J Affect Disord 2023 Jan 1;320:621-627. [doi: 10.1016/j.jad.2022.09.159] [Medline: 36206885]
24. Zhou L, Ma X, Wang W. Relationship between cognitive performance and depressive symptoms in Chinese older adults: the China Health and Retirement Longitudinal Study (CHARLS). J Affect Disord 2021 Feb;281:454-458. [doi: 10.1016/j.jad.2020.12.059]
25. Guo L, Luo F, Gao N, Yu B. Social isolation and cognitive decline among older adults with depressive symptoms: prospective findings from the China Health and Retirement Longitudinal Study. Arch Gerontol Geriatr 2021;95:104390. [doi: 10.1016/j.archger.2021.104390] [Medline: 33752099]
26. Yeom HE, Kim YJ. Age and sex-specific associations between depressive symptoms, body mass index and cognitive functioning among Korean middle-aged and older adults: a cross-sectional analysis. BMC Geriatr 2022 May 10;22(1):412. [doi: 10.1186/s12877-022-03079-3] [Medline: 35538446]

Abbreviations

ADL: activities of daily living
C-MMSE: Chinese version of the Mini-Mental State Examination
CLHLS: Chinese Longitudinal Healthy Longevity Survey
GBTM: group-based trajectory modeling
IADL: instrumental activities of daily living

Edited by PH Chau; submitted 25.03.25; peer-reviewed by NF Agbangla, X Xian; revised version received 04.07.25; accepted 04.07.25; published 21.08.25.

Please cite as:
Wu P, Lin C, Yang J, Wang Z
Cognitive Functioning Trajectories and Their Association With Mental Health in Older Adults With Hypertension: Secondary Analysis of the Chinese Longitudinal Healthy Longevity Survey Data
JMIR Aging 2025;8:e74916
URL: https://aging.jmir.org/2025/1/e74916
doi:10.2196/74916

© Peiyun Wu, Cheng Lin, Jing Yang, Zhizhuo Wang. Originally published in JMIR Aging (https://aging.jmir.org), 21.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Comparison of 3 Aging Metrics in Dual Declines to Capture All-Cause Dementia and Mortality Risk: Cohort Study

Anying Bai[1], PhD; Shan He[2], MD; Yu Jiang[3,4], PhD; Weihao Xu[5], MD; Zhanyi Lin[5], MD

[1]School of Population Medicine and Public Health, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[2]Haikou Cadre's Sanitarium of Hainan Military Region, Haikou, China

[3]School of Health Policy and Management, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[4]School of Population Medicine and Public Health, Peking Union Medical College, Beijing, China

[5]Department of Geriatrics, Guangdong Provincial Geriatrics Institute, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Southern Medical University, No. 106, Zhongshan 2nd Road, Yuexiu District, Guangzhou, China

Corresponding Author:
Weihao Xu, MD
Department of Geriatrics, Guangdong Provincial Geriatrics Institute, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Southern Medical University, No. 106, Zhongshan 2nd Road, Yuexiu District, Guangzhou, China

Abstract

Background: The utility of aging metrics that incorporate cognitive and physical function is not fully understood.

Objective: We aim to compare the predictive capacities of 3 distinct aging metrics—motoric cognitive risk syndrome (MCR), physio-cognitive decline syndrome (PCDS), and cognitive frailty (CF)—for incident dementia and all-cause mortality among community-dwelling older adults.

Methods: We used longitudinal data from waves 10-15 of the Health and Retirement Study. Cox proportional hazards regression analysis was employed to evaluate the effects of MCR, PCDS, and CF on incident all-cause dementia and mortality, controlling for socioeconomic and lifestyle factors, as well as medical comorbidities. Discrimination analysis was conducted to assess and compare the predictive accuracy of the 3 aging metrics.

Results: A total of 2367 older individuals aged 65 years and older, with no baseline prevalence of dementia or disability, were ultimately included. The prevalence rates of MCR, PCDS, and CF were 5.4%, 6.3%, and 1.3%, respectively. Over a decade-long follow-up period, 341 cases of dementia and 573 deaths were recorded. All 3 metrics were predictive of incident all-cause dementia and mortality when adjusting for multiple confounders, with variations in the strength of their associations (incident dementia: MCR odds ratio [OR] 1.90, 95% CI 1.30 - 2.78; CF 5.06, 95% CI 2.87 - 8.92; PCDS 3.35, 95% CI 2.44 - 4.58; mortality: MCR 1.60, 95% CI 1.17 - 2.19; CF 3.26, 95% CI 1.99 - 5.33; and PCDS 1.58, 95% CI 1.17 - 2.13). The C-index indicated that PCDS and MCR had the highest discriminatory accuracy for all-cause dementia and mortality, respectively.

Conclusions: Despite the inherent differences among the aging metrics that integrate cognitive and physical functions, they consistently identified risks of dementia and mortality. This underscores the importance of implementing targeted preventive strategies and intervention programs based on these metrics to enhance the overall quality of life and reduce premature deaths in aging populations.

(JMIR Aging 2025;8:e66104) doi:10.2196/66104

KEYWORDS

gerontology; geriatrics; older adults; older people; aging; motoric cognitive risk syndrome; MCR; physio-cognitive decline syndrome; PCDS; cognitive frailty; CF; frailty; discrimination; risk factors; prediction; dementia risk; mortality risk

Introduction

The aging process encompasses various physiological declines across multiple systems. As the global population ages, the prevalence of dementia is rising, expected to reach 131 million individuals by 2050. This increase poses a significant economic challenge, potentially equating to 1.1% of the global gross domestic product by 2030 [1]. The lack of curative treatments for dementia and its substantial public health impact underscore the necessity for early detection to mitigate or delay its onset. Given dementia's heterogeneous and lengthy preclinical phase, early screening and diagnosis in at-risk individuals are vital for disease management and caregiver preparedness [2].

Cognitive and physical declines with age often occur simultaneously, suggesting shared underlying mechanisms [3]. Researchers have developed metrics across molecular, phenotypic, and functional areas [4] to reflect the complex nature of aging accurately. Cognitive deterioration typically

XSL•FO
RenderX

precedes dementia by several years, with evidence indicating that motor decline, especially in walking speed, can precede cognitive decline by over a decade [5,6]. Thus, composite aging metrics, encompassing both physical and cognitive functions, offer a promising method to gauge the functional status of the aging population. A meta-analysis highlighted an increased dementia risk in individuals with both physical frailty and cognitive impairment compared to those with cognitive impairment alone [7]. Therefore, composite aging metrics may serve as focal points for interventions aimed at preventing or delaying disability onset and enhancing the healthy lifespan of the elderly [8,9].

An ideal dementia screening tool for primary care should be brief, easily administered, acceptable to older individuals, and exhibit high sensitivity and specificity. Research on aging metrics that incorporate both cognitive and physical functions is gaining traction in gerontological studies due to their strong predictive power for adverse health outcomes. Various metrics have been introduced, such as cognitive frailty (CF) [10], motoric cognitive risk syndrome (MCR) [11], and physio-cognitive decline syndrome (PCDS) [12]. Despite their conceptual similarities, detailed assessments of their definitions and attributes are scarce, hindering their application in research and clinical settings.

Research has shown an increased incidence of concurrent gait and cognitive impairments in older adults susceptible to dementia [5]; however, the specific clinical traits of those experiencing both declines are not well-defined, and direct comparisons between different aging metrics have not been made. Additionally, the effectiveness of these metrics in identifying at-risk individuals and assessing the risk of adverse health events within the same population is not well understood. This lack of knowledge is clinically important for effectively categorizing older adults and identifying potentially reversible conditions in individuals with concurrent declines, thereby informing targeted strategies to slow dementia progression or reduce mortality rates. Consequently, our study seeks to fill these voids by comparing the risk of future dementia and all-cause mortality across 3 aging metrics and examining the predictive abilities of MCR, PCDS, and CF for adverse events among community-dwelling older adults without dementia at baseline, employing data from a broad population-based cohort.

Methods

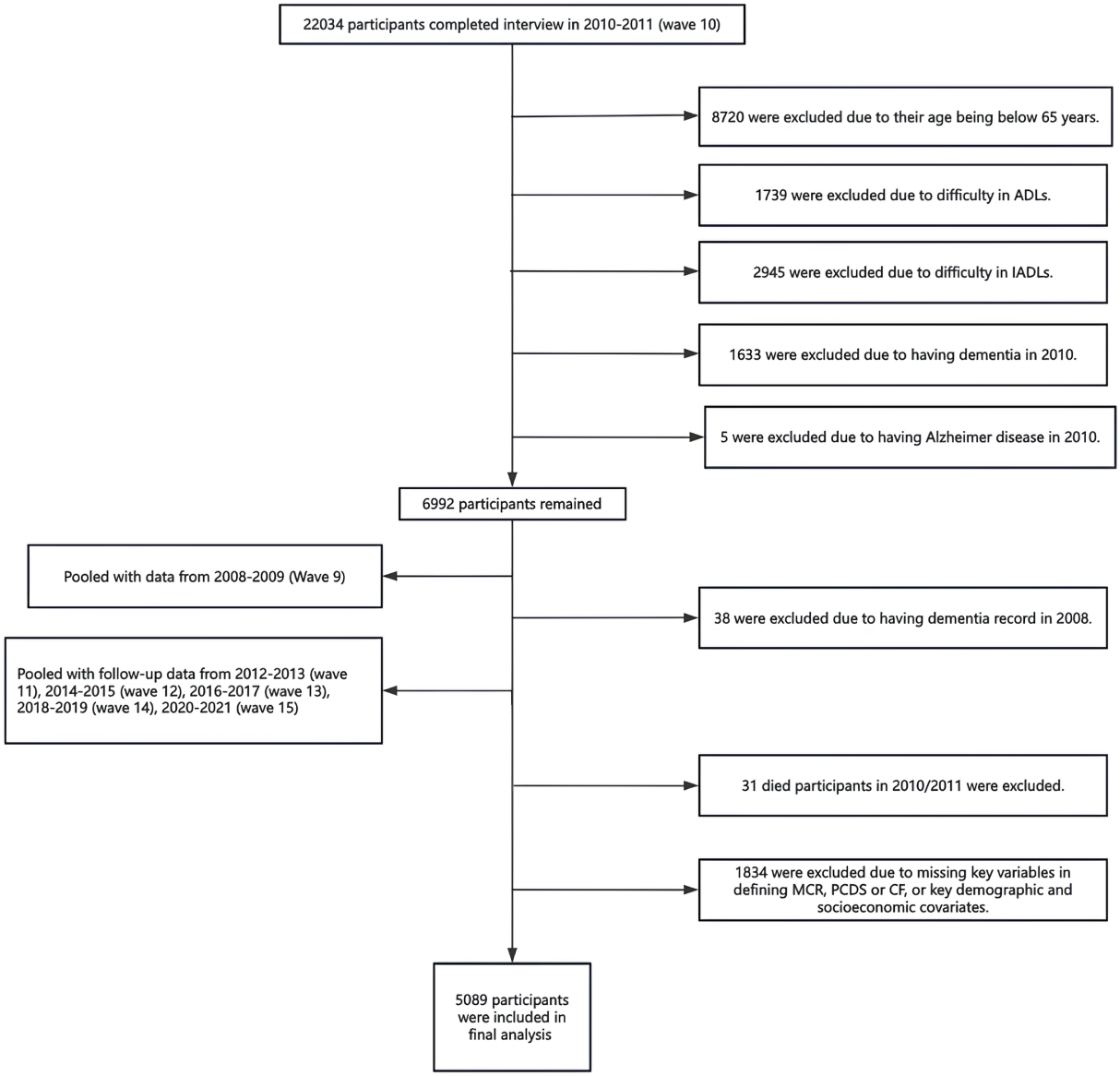
Sample

This study analyzes data from waves 10 - 15 of the Health and Retirement Study (HRS), the most extensive longitudinal study examining the aging experiences of Americans aged 51 and older. The HRS uses multi-stage probability sampling of U.S. households to obtain a nationally representative sample of adults in this age group [13]. It collects self-reported data on demographics, chronic health conditions, daily activities, disability status, and other health determinants at baseline and every 2 years thereafter. In 2006, the HRS began conducting enhanced face-to-face interviews that included physical performance assessments, biomarker collections, and a leave-behind questionnaire on psychosocial topics. Half of the households were chosen randomly for the enhanced interview in 2006, with the remainder selected in 2008, a process that continues in subsequent waves. Further information on the HRS's recruitment strategies and design is detailed in previous literature [13].

The baseline for this analysis combined data from the 2008 - 2009 (wave 9) and 2010 - 2011 (wave 10) waves, marking the initial occasion respondents were asked about diagnoses of Alzheimer disease or dementia, instead of a "memory-related disease." Mortality data has been available since 2011. A total of 22,034 individuals completed wave 10 and were followed biennially through to 2020 - 2021 (wave 15). This study is a secondary analysis of the de-identified HRS public data, and the original HRS was approved by the University of Michigan Institutional Review Board. All participants signed the informed consent at the time of participation. Our final sample consisted of 2372 individuals who (1) were 65 years or older, (2) had complete baseline data on MCR, PCDS, and CF measures, (3) reported no difficulty with any activities of daily living and instrumental activities of daily living at baseline, (4) did not have Alzheimer disease or dementia at baseline, and (5) were alive in 2010 and 2011. Figure 1 shows the participant flow through each selection stage according to the inclusion criteria.

XSL•FO
RenderX

Figure 1. Study flowchart of participant selection. ADL: activity of daily living; IADL: instrumental activity of daily living; CF: cognitive frailty; MCR: motoric cognitive risk syndrome; PCDS: physio-cognitive decline syndrome.



Measures

Cognitive Function

Biennial cognitive function tests were administered by trained HRS interviewers either in-person or via telephone using the Modified Telephone Interview for Cognitive Status (TICS-m), a global cognition test modeled on the Mini-Mental State Examination. The TICS-m comprises immediate and delayed 10-noun free recall tests (score range: 0 - 10 for each), a serial 7 subtraction test (score range: 0 - 5), and a test of counting backward from 20 (score range: 0 - 2). Higher scores denote better cognitive performance. During each assessment, HRS participants were categorized into normal cognition, mild cognitive impairment (MCI), or dementia based on established thresholds and comprehensive evaluations, including expert clinician adjudication from the Aging, Demographics, and Memory Study, a dementia sub-study within the HRS framework. Cognitive status was categorized into 3 distinct groups based on continuous scores [14], where scores from 12 to 27 indicated no impairment; scores between 7 and 11 signified cognitive impairment without dementia or MCI; and scores from 0 to 6 suggested dementia [15].

Motoric Cognitive Risk Syndrome

MCR was defined by subjective cognitive complaints coupled with slow gait in older adults who did not have a mobility disability or dementia [16]. In the HRS, gait speed (measured in meters per second) was determined from the time taken (in seconds) to walk a 2.5-meter course at a normal pace within participants' homes. A slow gait was defined as performance ≥1 SD below the mean for the participant's age and sex, a criterion previously used in HRS to identify MCR [17].

Subjective cognitive complaints were assessed using 2 questions: 1. "How would you rate your memory at the present time? Would you say it is excellent, very good, good, fair, or poor?" and 2. "Compared with the previous interview, would you say your memory is better now, about the same, or worse now than it was then?" Responses of 'fair' or 'poor' to the first question,

XSL•FO
RenderX

or 'worse' to the second, were considered indicative of cognitive complaints.

Cognitive Frailty

CF was defined as the co-occurrence of physical frailty and MCI [18]. The concept of frailty was identified using the 5 criteria outlined by Fried et al [19] in the Cardiovascular Health Study: unintended weight loss, physical inactivity, exhaustion, weakness, and reduced speed. Unintended weight loss is recognized as either a 10% or greater reduction in BMI since the last measurement in 2008 or a current BMI less than 18.5 kg/m^2. Physical activity levels were quantified by averaging the frequencies of activities at 3 levels of intensity—mild, moderate, and vigorous—weighted by the average metabolic equivalent of task (MET) scores for each intensity level: mild (1 - 3 MET), moderate (4 - 6 MET), and vigorous (7 - 10 MET). Participants were deemed physically inactive if their average physical activity fell within the lowest 20%. Exhaustion was determined based on responses to 2 questions from the Center for Epidemiologic Studies Depression scale [20], including, "I could not get going" and "I felt that everything I did was an effort." Muscle strength was assessed through the average of 2 grip strength measurements using a dynamometer on the dominant hand. Weakness was determined by grip strength falling below thresholds adjusted for BMI and gender, as established in the CHS. Reduced speed or slowness was defined as a speed <0.762 m/s for women taller than 159 cm or men taller than 173 cm and as <0.653 m/s for women 159 cm tall or less or men 173 cm tall or less, measured on a 2.5-m course [19,21]. Participants were considered to have missing data for physical measures if they were unable to perform the assessments due to the absence of suitable facilities or equipment or recent surgery. The diagnosis of frailty was based on the number of these criteria fulfilled, with those meeting 3-5 criteria classified as frail.

Physio-Cognitive Decline Syndrome

PCDS was defined as slowness or weakness (using cutoffs from the 2019 consensus update by the Asian Working Group for Sarcopenia), accompanied by cognitive performance that is at least 1.5 standard deviations below the mean for age-, sex-, and education-matched controls across all cognitive domains [12,22]. This assessment is based on comprehensive objective neuropsychological tests.

All-Cause Dementia and Mortality

The diagnosis of dementia was based on physician-diagnosed dementia and TICS scores ranging from 0 to 6. Mortality data were recorded, including the year and month of death, obtained from an exit interview or a spouse or partner's core interview.

Covariates

Covariates included sociodemographic factors, clinical characteristics, and health-related lifestyle behaviors, all of which were assessed at baseline. Sociodemographic characteristics included age (in years), sex (male or female), educational background (primary school or below, high school or equivalent, college and above), and marital status (married vs unmarried). Health-related lifestyle behaviors included excessive drinking, defined as more than 14 drinks per week for men and more than 7 drinks per week for women. Alcohol consumption was calculated by multiplying the number of days per week that alcohol (drink liquor or beer or wine or rice) was consumed by the number of drinks (liang or bottles or mugs) per day. Clinical characteristics included history of hypertension, diabetes, and heart disease. Hypertension was defined as systolic blood pressure ≥140 mm Hg, diastolic blood pressure ≥90 mm Hg, physical diagnosis, or antihypertensive medication use [23,24]. Diabetes was defined as having a diabetes diagnosis by a physician, being on treatment for diabetes, and having a fasting glucose level greater than or equal to 126 mg/dL and HbA_{1c} >6.5%. The presence of heart disease was determined via a physician's diagnosis obtained through an in-person visit with study personnel via a questionnaire.

Statistical Analysis

The incidence rates of all-cause dementia and mortality were calculated as the number of incident cases divided by the number of person-years of follow-up within the observation year (from 2008 to 2021). Differences between the MCR and non-MCR groups, PCDS and non-PCDS groups, and CF and non-CF groups were assessed using a 2-sided, independent *t*-test and the χ^2 test. To evaluate the impact of MCR, PCDS, and CF on the occurrence of all-cause dementia and mortality, we employed Cox proportional hazards regression analysis. The observation period extended from the index date to the earliest of the following: onset of dementia, death, or the conclusion of the observation period (December 31, 2018). Adjusted hazard ratios (AHRs) for MCR, PCDS, and CF in predicting the onset of dementia and all-cause mortality were calculated, accounting for covariates in an initially unadjusted model. Subsequent adjustments for covariates were made in 2 stages: Model 1 adjusted for age and gender; Model 2 further incorporated socioeconomic (education level and marital status), lifestyle (excessive drinking), and medical conditions (hypertension, diabetes, heart disease, and stroke). To assess and compare the predictive accuracy of all models, discrimination, defined by the model's ability to differentiate between individuals who develop dementia and those who do not, was quantified using Harrell C-statistic, taking survival into account.

Several sensitivity analyses were conducted to verify the stability of our findings. First, to focus on new cases and reduce reverse causation bias, individuals diagnosed with dementia or who died within 2 years of follow-up (sensitivity analysis I) were excluded. Second, to address the competing risk of death for dementia occurrences, Fine and Gray competing risk models were used [25], comparing these to the results from Cox proportional hazards regression models (sensitivity analysis II). The sub-distribution hazard function, defined at time *t*, represents the immediate risk of event *k* among individuals not previously experiencing event *k*. Third, to minimize selection bias, associations within individual samples with complete data on MCR, PCDS, and CF were analyzed separately (sensitivity analysis III). Statistical analyses were performed using 2-tailed tests with a significance level of $P<.05$ and 95% CIs, employing Stata (version 17) for all statistical procedures.

XSL•FO
RenderX

Ethical Considerations

This investigation has been conducted in accordance with the ethical standards of the Declaration of Helsinki, as well as national and international guidelines. The study has been approved by the Institutional Review Board at the University of Michigan (approval number: HUM00061128). All participants were provided with detailed information about the study, including its purpose, procedures, potential risks and benefits, and their rights to withdraw at any time. Written informed consent was obtained from all participants prior to their involvement in the study. To ensure privacy and confidentiality, all data collected were anonymized and deidentified. No identifying information was retained or published. Protective measures were in place to safeguard participant information, including secure storage of data and restricted access to study records. Participants were not compensated for their involvement in this study. The research was conducted on a voluntary basis, and no financial or other incentives were provided.

Results

Baseline Characteristics

The study participants' baseline characteristics are outlined in Table 1. The initial cohort comprised 2367 individuals, featuring prevalence rates for MCR, PCDS, and CF at 5.4% (n=121), 6.3% (n=140), and 1.3% (n=31), respectively. Among these, CF patients were the oldest on average (75.7, SD 6.1 years), with the distribution of men being 52.89% (64/121) in the MCR group, 47.14% (66/140) in the PCDS group, and 58.06% (10/31) in the CF group. During the follow-up, 573 (24.2%) patients died. The proportions of all-cause dementia for MCR, PCDS, and CF were 24.8% (30/121), 33.6% (47/140), and 41.9% (13/31), respectively.

Table . Characteristics of included patients at baseline according to 3 aging metrics.

Variable	Non-MCR[a] (n=2246)	MCR (n=121)	*P* value	Non-PCDS[b] (n=2227)	PCDS (n=140)	*P* value	Non-CF[c] (n=2336)	CF (n=31)	*P* value
Age (years), mean (SD)	73.78 (5.65)	73.29 (5.54)	.17	73.72 (5.62)	74.29 (6.02)	.08	73.73 (5.63)	75.65 (6.11)	.52
Male, n (%)	1086 (48.35)	64 (52.89)	.33	1084 (48.68）	66 (47.14)	.73	1132 (48.46)	18 (58.06)	.29
Educational background, n (%)			.001			.80			<.001
Illiterate	286 (12.73)	27 (22.31)		297 (13.34）	16 (11.43)		301 (12.89）	12 (38.71)	
Primary or above	1286 (57.26)	73 (60.33)		1276 (57.30%)	83 (59.29)		1343 (57.49)	16 (51.61)	
Secondary or above	674 (30.01)	21 (17.36)		654 (29.37)	41 (29.29)		692 (29.62)	3 (9.68)	
Medical history, n (%)									
Hypertension	1398 (62.24)	91 (75.21)	.004	1392 (62.51)	97 (69.29)	.11	1466 (62.76)	23 (74.19)	.19
Diabetes	431 (19.19)	25 (20.66)	.69	410 (18.41)	46 (32.86)	<.001	446 (19.09)	10 (32.26)	.07
Heart disease	605 (26.94)	47 (38.84)	.004	615 (27.62)	37 (26.43)	.76	638 (27.31)	14 (45.16)	.03
Excessive drink	1436 (63.94)	95 (78.51)	.001	1428 (64.12)	103 (73.57)	.02	1509 (64.60)	22 (70.97)	.46
Incident all-cause dementia	311 (13.85)	30 (24.79)	.001	294 (13.20)	47 (33.57)	<.001	328 (14.04)	13 (41.94)	<.001
Mortality	530 (23.60)	43 (35.54)	.003	526 (23.62)	47 (33.57)	.008	556 (23.80）	17 (54.84)	<.001

[a] MCR: motoric cognitive risk syndrome.

[b]PCDS: physio-cognitive decline syndrome.

[c]CF: cognitive frailty.

XSL•FO
RenderX

Relationships of MCR, PCDS, and CF With Incident Dementia and All-Cause Mortality

Overall, there were 341 incident dementia cases during follow-up, for an overall incidence rate of 19.52 (95% CI 17.56-21.71) per 1000 person-years. There were 573 cases that died during follow-up, for an overall incidence rate of 30.14 (95% CI 27.77-32.71) per 1000 person-years (Table 2). The incidence rates of all-cause dementia among MCR, CF, and PCDS patients were 38.04, 108.11, and 55.80 per 1000 person-years, respectively—significantly higher than those observed in the relatively healthy control group. Similarly, the incidence rates of all-cause mortality were 47.99, 100.79, and 44.70 per 1000 person-years for MCR, CF, and PCDS patients, respectively—again, markedly higher than in the healthy controls. Table 2 also demonstrates significant associations of the 3 conditions with increased risks of incident dementia and all-cause mortality in various models (all *P* values <.001). CF had the highest AHR for both outcomes (5.06; 95% CI 2.87-8.92 for dementia, and 3.26; 95% CI 1.99 - 5.33 for mortality). Participants with dual decline experienced a two to threefold increased risk of dementia progression (AHR: 1.90 - 2.22 in the MCR group; 3.21 - 3.35 in the PCDS group) compared with those without dual decline. The AHR for all-cause mortality ranged from 1.60 to 1.84 in the MCR group and from 1.56 to 1.63 in the PCDS group. The trend was consistent across models: AHRs for dementia and mortality increased from the unadjusted model to adjusted model 1, then decreased upon further adjustment for covariates, including socioeconomic status, lifestyle factors, and medical comorbidities. Sensitivity analyses I (Table S1 in Multimedia Appendix 1) and III (Table S2 in Multimedia Appendix 1) consistently showed an elevated risk of dementia and mortality for MCR, CF, and PCDS. However, in sensitivity analysis II, the standardized hazard ratios for MCR and CF were not statistically significant post-adjustment (Table S3 in Multimedia Appendix 1).

XSL•FO
RenderX

Table . Associations between 3 aging metrics, incident all-cause dementia, and all-cause mortality.

	Total sample	MCR[a]		CF[b]		PCDS[c]	
		No	Yes	No	Yes	No	Yes
All-cause dementia							
Events and sample size, n/N	341/2367	311/2246	30/121	328/2336	13/31	294/2227	47/140
Incidence (95% CI)[d]	19.52 (17.56 - 21.71)	18.65 (16.69 - 20.84)	38.04 (26.59 - 54.40)	18.91 (16.97 - 21.07)	108.11 (62.77 - 186.18)	17.68 (15.77 - 19.83)	55.80 (41.93 - 74.27)
Unadjusted HR[e] (95% CI)	—[f]	Ref.[g]	2.03 (1.40 - 2.95)	Ref.	6.23 (3.57 - 10.88)	Ref.	3.21 (2.36 - 4.37)
Model 1[h]: adjusted HR (95% CI)	—	Ref.	2.22 (1.53 - 3.24)	Ref.	6.62 (3.79 - 11.57)	Ref.	3.29 (2.41 - 4.48)
Model 2[i]: adjusted HR (95% CI)	—	Ref.	1.90 (1.30 - 2.78)	Ref.	5.06 (2.87 - 8.92)	Ref.	3.35 (2.44 - 4.58)
All-cause mortality							
Events and sample size, n/N	573/2367	530/2246	43/121	556/2336	17/31	526/2227	47/140
Incidence (95% CI)	30.14 (27.77 - 32.71)	29.26 (26.87 - 31.86)	47.99 (35.59 - 64.71)	29.51 (27.16 - 32.07)	100.79 (62.66 - 162.13)	29.29 (26.89 - 31.90)	44.70 (33.58 - 59.49)
Unadjusted HR (95% CI)	—	Ref.	1.68 (1.23 - 2.29)	Ref.	3.93 (2.42 - 6.38)	Ref.	1.56 (1.16 - 2.11)
Model 1: adjusted HR (95% CI)	—	Ref.	1.84 (1.35 - 2.51)	Ref.	4.04 (2.49 - 6.57)	Ref.	1.63 (1.21 - 2.19)
Model 2: adjusted HR (95% CI)	—	Ref.	1.60 (1.17 - 2.19)	Ref.	3.26 (1.99 - 5.33)	Ref.	1.58 (1.17 - 2.13)

[a]MCR: motoric cognitive risk syndrome.

[b]CF: cognitive frailty.

[c]PCDS: physio-cognitive decline syndrome.

[d]Incidence rates = events per 1000 person-years.

[e]HR: hazard ratio.

[f]Not applicable.

[g]Ref: reference.

[h]Model 1 adjusted for age and gender.

[i]Model 2 further adjusted for educational background, marital status, excessive drinking, hypertension, diabetes, and heart disease.

Discriminations of MCR, PCDS, and CF for All Outcomes

The discrimination abilities of MCR, PCDS, and CF regarding all-cause dementia and mortality are detailed in Table 3. The C-index for PCDS in identifying all-cause dementia was 0.732 (95% CI 0.703-0.760), outperforming the C-indices for MCR and CF (PCDS vs MCR: 0.012, 95% CI −0.001 to 0.025; *P*=.08; CF vs MCR: 0.005, 95% CI −0.004 to 0.013; *P*=.35). In contrast, MCR's C-index for identifying all-cause mortality was the highest among the 3, at 0.727 (95% CI 0.706-0.748). Similar to the case with incident dementia, the differences in discrimination ability among MCR, PCDS, and CF for all-cause mortality were not statistically significant.

XSL•FO
RenderX

Table . Harrell C-index for Cox regression models predicting incident all-cause dementia, and all-cause mortality.

	C-index[a]	95% CI	*P* value	Difference	*P* value
Incident all-cause dementia					
MCR[b]	0.7194	0.6915 to 0.7472	<.001	Reference	
PCDS[c]	0.7315	0.7028 to 0.7602	<.001	0.0121 (−0.0009 to 0.0252)	.08
CF[d]	0.7239	0.6959 to 0.7520	<.001	0.0046 (−0.0038 to 0.0129)	.35
All-cause mortality					
MCR	0.727	0.7061 to 0.7479	<.001	Reference	
PCDS	0.7259	0.7048 to 0.7469	<.001	−0.0012 (−0.0054 to 0.0030)	.59
CF	0.7254	0.7044 to 0.7465	<.001	−0.0016 (−0.0054 to 0.0022)	.41

[a]All indexes were estimated in models adjusted for age, gender, educational background, marital status, excessive drinking, hypertension, diabetes, and heart disease

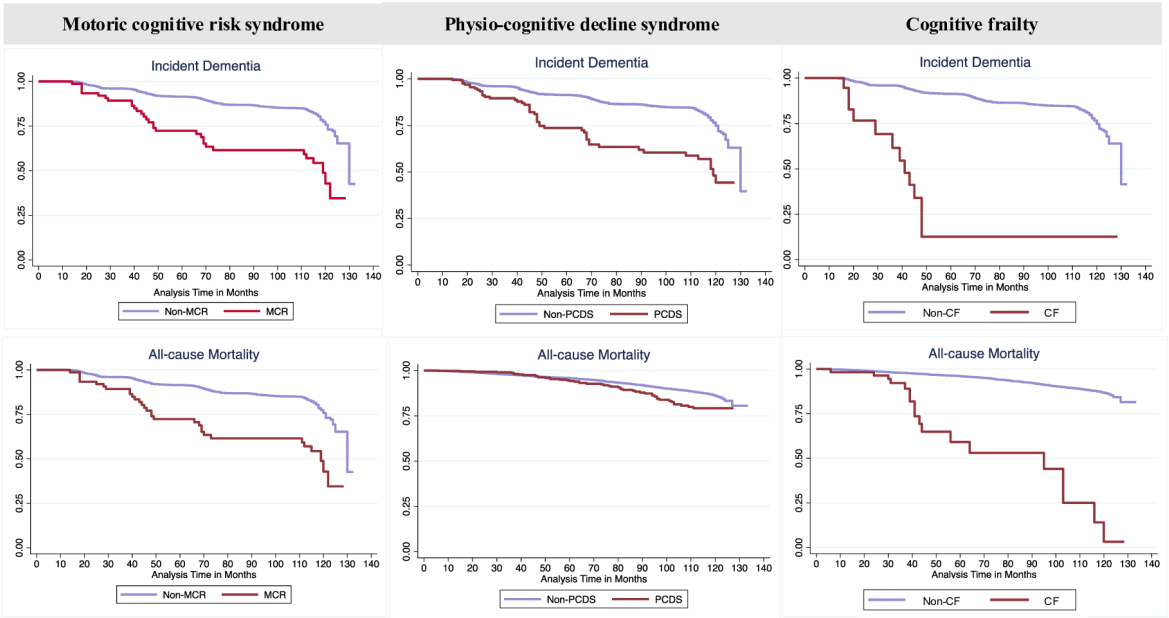
[b]MCR: motoric cognitive risk syndrome.

[c]PCDS: physio-cognitive decline syndrome.

[d]CF: cognitive frailty.

The Kaplan-Meier curve, depicted in Figure 2, shows the duration to incident dementia or all-cause mortality, stratified by MCR, PCDS, or CF, and healthy controls, with adjustments made for all covariates. Both curves demonstrate a decline over the follow-up period, with a notably pronounced decrease observed among patients (all *P* values for log rank <.001).

Figure 2. Kaplan-Meier survival curve depicting the proportion of participants remaining dementia-free or surviving during follow-up, comparing motoric cognitive risk syndrome (MCR), physio-cognitive decline syndrome (PCDS), and cognitive frailty (CF) patients with their respective healthy control groups.



XSL•FO
RenderX

Discussion

Principal Findings

Analyzing data from a large-scale cohort in the United States, this study is the first to illustrate that MCR, PCDS, and CF have significant associations with incident dementia and all-cause mortality within the same population, despite the inherent heterogeneity of these metrics. Among them, CF emerged as the highest-risk group for both dementia and mortality, with markedly elevated incidence rates and the highest AHRs (5.06 for dementia and 3.26 for mortality), identifying CF as the most vulnerable subgroup. These findings offer valuable insights into the future health trajectories of frail individuals with cognitive impairments and highlight the need for targeted interventions and early health education to mitigate dementia and mortality risks in later life at lower costs. While PCDS showed the strongest predictive ability for dementia and MCR for mortality, the differences in predictive accuracy were not statistically significant. Implementing appropriate management strategies can help alleviate the health care burden for individuals with varying cognitive and physical conditions.

To forestall unhealthy aging, health care systems need to identify individuals at risk of functional declines that are still preventable or reversible. Our results are in line with prior studies, indicating that MCR syndrome is linked with a 90% increased risk of incident dementia and a 60% higher risk of mortality, even after adjustments for demographics, SES, and cardiovascular comorbidities (hypertension, diabetes, and heart disease). Consistent with research from Chung et al, the prevalence of hypertension and heart disease was significantly higher in the MCR group, highlighting the importance of regular screening to manage cardiovascular risk factors and delay dementia onset [26,27]. These observations further emphasize the potential neurodegenerative nature of MCR pathologic changes [28].

CF has been associated with an increased risk of dementia and mortality, as demonstrated by a meta-analysis revealing that older adults with CF face higher mortality and dementia risks than their healthier counterparts [29]. However, the absence of a standardized assessment tool for CF hampers its use in widespread epidemiological studies. Using the frailty phenotype alongside the MMSE, although common, proves to be time-intensive and presents challenges for deployment in community and hospital settings [30]. Furthermore, the lower prevalence of CF compared to other measures (1.31% for CF vs 5.91% for PCDS vs 5.11% for MCR) in the community-dwelling older adult cohort indicates difficulties in identifying a substantial number of at-risk individuals for interventions in similar groups.

In contrast to CF, PCDS, with its precise definition encompassing physical decline and cognitive impairment, provides a more targeted approach. By focusing on the mobility aspects of frailty linked to worse cognitive outcomes and increased mortality risk, researchers can more accurately identify the target population, concentrating on specific causal mechanisms [31,32]. PCDS serves as an advantageous focus for multidomain interventions, integrating physical activity, cognitive training, nutritional counseling, and disease management education to improve the condition of vulnerable but potentially reversible older individuals [33]. A recent Singapore study that used a 12-week dual-task exercise program to examine the potential reversibility of CF, MCR, and PCDS also found clinical improvements in PCDS, but the longer-term effects remain uncertain [34]. Nonetheless, Lee et al [35] demonstrated that PCDS was related to a 6-year, but not 3-year, incidence of dementia, even when using inverse probability weighting analysis to account for bias from missing data. This finding implies a latency period for the development of dementia, highlighting this interval as potentially crucial for preventive interventions.

While PCDS demonstrated the best predictive performance for incident dementia, MCR emerged as a superior predictor for all-cause mortality. The choice between these 2 metrics may depend on practical considerations, with MCR offering several unique advantages for clinical use. MCR is simpler and more efficient to assess, unaffected by education levels or learning effects, which enhances its credibility and reliability as a screening tool. In contrast, the comprehensive cognitive assessments required for PCDS diagnosis can be labor-intensive and require specific skills, which may limit their use in community or primary care settings. Given MCR's higher predictive accuracy for all-cause mortality and its ease of implementation, it is well-suited for use in routine health care assessments, particularly in primary health care settings. Additionally, research suggests that while handgrip strength declines earlier in aging, walking speed in later life is a stronger predictor of mortality [36]. This further supports MCR's higher predictive accuracy for all-cause mortality.

Despite representing a prodromal phenotype of accelerated aging, MCR remains an important target for intervention to prevent poor outcomes in older adults. MCR offers incremental validity in predicting dementia beyond what is provided by MCI subtypes [16] and individual components such as subjective cognitive complaints or slow gait [16,37]. Our findings suggest that MCR could serve as a simpler, more efficient tool to identify individuals at higher risk for dementia, especially given its independence from educational background or learning effects. The motor components of MCR, such as the time-up-and-go test or one-leg-standing test, are reliable and valid methods for quantifying motor function, making MCR an easily implementable screening tool in various health care settings. However, one study has shown that the MCR subtype defined by the 5-times-sit-to-stand test, which also includes a balance component, was less effective in predicting cognitive decline compared to MCR defined by slow gait [38]. Further research is needed to determine how different gait and balance assessments influence the identification of individuals at risk for dementia, and whether integrating MCR into routine geriatric assessments could enhance early intervention strategies in aging populations.

Early and appropriate intervention might postpone the onset of dementia or reduce its risk, particularly given the lengthy incubation period associated with the development of dementia. Our results could aid clinicians in the early implementation of screening and prevention strategies and inform government decisions on community health prevention. The strength of this

XSL•FO
RenderX

study lies in its extensive, well-delineated cohort, featuring longitudinal assessments, uniform measures, and verified outcomes. To our knowledge, this is the first study to assess the predictive accuracy of MCR, PCDS, and CF regarding dementia incidence and all-cause mortality within the same cohort. The reliability and consistency of our findings, supported by sensitivity analyses, bolster the study's validity. Nonetheless, several limitations should be acknowledged. First, the study lacked objective neuropsychological testing [39] and did not account for some potential confounding factors, such as *APOE* genotype or imaging biomarkers [40,41]. However, our sensitivity analysis—excluding participants diagnosed with dementia or those who died within 2 years of the index date—strengthened the robustness of our conclusions. Second, while we used data from a large longitudinal cohort, the final sample size of 2372 individuals may limit the representativeness of our findings. Additionally, the prevalence rates of MCR (5.4%), PCDS (6.3%), and CF (1.3%) in our sample may be higher than those in the general older population, introducing potential selection bias. To address this, sensitivity analyses were conducted within individual subgroups with complete data on MCR, PCDS, and CF. These analyses included larger sample sizes and showed lower prevalence rates for the 3 conditions, aligning more closely with those observed in the general population. The robustness of these results further confirmed our conclusions, enhancing their validity and generalizability while mitigating potential selection bias. Future research could benefit from larger, more diverse, and representative samples to enhance generalizability. Third, the potential influence of mortality on dementia risk assessment must be considered. Individuals who died during follow-up could bias dementia outcome evaluations. To address this competing risk, we employed Fine and Gray competing risk models and compared the results with those from Cox proportional hazards regression models as part of our sensitivity analysis. This approach ensured that our findings remained robust and accounted for the impact of mortality on dementia risk associations. Finally, variations in syndrome definitions and differences in the prevalence of the 3 aging metrics within the population could affect their predictive power for adverse outcomes. This variability may limit the generalizability of our results to other populations. To mitigate this, sensitivity analyses within subgroups with complete data on MCR, PCDS, and CF were conducted, reinforcing the reliability of the associations between these aging metrics and adverse health outcomes.

Conclusions

The integration of cognitive and physical functions into aging metrics consistently indicates risks for incident dementia and all-cause mortality, despite their significant differences, underscoring their utility for effective risk stratification in research and clinical settings. Patients with CF represent the most vulnerable subgroup, highlighting the need for prioritized preventive strategies and interventions. Meanwhile, MCR emerged as a particularly efficient and accurate screening tool for both dementia and mortality. These findings emphasize the importance of implementing targeted prevention and intervention programs based on these metrics to enhance quality of life and reduce premature mortality among aging populations. Further research on the longitudinal dynamics of these aging metrics in relation to dementia, mortality, and other outcomes is essential for a deeper understanding of their long-term impact.

Acknowledgments

This study was supported by the Natural Science Foundation of Guangdong Province of China (2023A1515011259).

Data Availability

The datasets analyzed during this study are available from the Health and Retirement Study [42].

Authors' Contributions

AYB and WHX contributed to conception and design. AYB and SH contributed to writing the original draft and reviewing and editing the manuscript. AYB, YJ, and SH were involved in data curation, and data analysis. YJ and ZYL contributed to reviewing and editing. AYB and WHX contributed to the methodology, project administration, supervision, and verified the underlying data. All authors gave final approval of the version to be published.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary materials.
[DOCX File, 26 KB - aging_v8i1e66104_app1.docx]

References

1. Zeisel J, Bennett K, Fleming R. World Alzheimer report 2020: design, dignity, dementia: dementia-related design and the built environment. 2020.
2. Borson S, Frank L, Bayley PJ, et al. Improving dementia care: the role of screening and detection of cognitive impairment. Alzheimers Dement 2013 Mar;9(2):151-159. [doi: 10.1016/j.jalz.2012.08.008] [Medline: 23375564]

XSL•FO
RenderX

3. Robertson DA, Savva GM, Kenny RA. Frailty and cognitive impairment--a review of the evidence and causal mechanisms. Ageing Res Rev 2013 Sep;12(4):840-851. [doi: 10.1016/j.arr.2013.06.004] [Medline: 23831959]
4. Ferrucci L, Levine ME, Kuo PL, Simonsick EM. Time and the metrics of aging. Circ Res 2018 Sep 14;123(7):740-744. [doi: 10.1161/CIRCRESAHA.118.312816] [Medline: 30355074]
5. Montero-Odasso MM, Barnes B, Speechley M, et al. Disentangling cognitive-frailty: results from the gait and brain study. J Gerontol A Biol Sci Med Sci 2016 Nov;71(11):1476-1482. [doi: 10.1093/gerona/glw044] [Medline: 26984391]
6. Dumurgier J, Artaud F, Touraine C, et al. Gait speed and decline in gait speed as predictors of incident dementia. J Gerontol A Biol Sci Med Sci 2016;72(5). [doi: 10.1093/gerona/glw110]
7. Zheng L, Li G, Gao D, et al. Cognitive frailty as a predictor of dementia among older adults: a systematic review and meta-analysis. Arch Gerontol Geriatr 2020;87:103997. [doi: 10.1016/j.archger.2019.103997] [Medline: 31846833]
8. Cano A. Cognitive frailty, a new target for healthy ageing. Maturitas 2015 Oct;82(2):139-140. [doi: 10.1016/j.maturitas.2015.07.026] [Medline: 26277253]
9. Ruan Q, Yu Z, Chen M, Bao Z, Li J, He W. Cognitive frailty, a novel target for the prevention of elderly dependency. Ageing Res Rev 2015 Mar;20:1-10. [doi: 10.1016/j.arr.2014.12.004] [Medline: 25555677]
10. Kelaiditi E, Cesari M, Canevelli M, et al. Cognitive frailty: rational and definition from an (I.A.N.A./I.A.G.G.) International Consensus Group. J Nutr Health Aging 2013 Nov;17(9):726-734. [doi: 10.1007/s12603-013-0367-2]
11. Verghese J, Annweiler C, Ayers E, et al. Motoric cognitive risk syndrome: multicountry prevalence and dementia risk. Neurology (ECronicon) 2014 Aug 19;83(8):718-726. [doi: 10.1212/WNL.0000000000000717] [Medline: 25031288]
12. Chung CP, Lee WJ, Peng LN, et al. Physio-cognitive decline syndrome as the phenotype and treatment target of unhealthy aging. J Nutr Health Aging 2021;25(10):1179-1189. [doi: 10.1007/s12603-021-1693-4] [Medline: 34866145]
13. Heeringa SG, Connor JH. Technical Description of the Health and Retirement Survey Sample Design: University of Michigan; 1995.
14. Langa KM, Plassman BL, Wallace RB, et al. The Aging, Demographics, and Memory Study: study design and methods. Neuroepidemiology 2005;25(4):181-191. [doi: 10.1159/000087448] [Medline: 16103729]
15. Crimmins EM, Kim JK, Langa KM, Weir DR. Assessment of cognition using surveys and neuropsychological assessment: the Health and Retirement Study and the Aging, Demographics, and Memory Study. J Gerontol B Psychol Sci Soc Sci 2011 Jul;66(Suppl 1):i162-i171. [doi: 10.1093/geronb/gbr048] [Medline: 21743047]
16. Verghese J, Wang C, Lipton RB, Holtzer R. Motoric cognitive risk syndrome and the risk of dementia. J Gerontol A Biol Sci Med Sci 2013 Apr;68(4):412-418. [doi: 10.1093/gerona/gls191] [Medline: 22987797]
17. Ayers E, Verghese J. Motoric cognitive risk syndrome and risk of mortality in older adults. Alzheim Dementia 2016 May;12(5):556-564. [doi: 10.1016/j.jalz.2015.08.167]
18. Gaspar PM, Campos-Magdaleno M, Pereiro AX, Facal D, Juncos-Rabadán O. Cognitive reserve and mental health in cognitive frailty phenotypes: Insights from a study with a Portuguese sample. Front Psychol 2022;13:968343. [doi: 10.3389/fpsyg.2022.968343] [Medline: 36110284]
19. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
20. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977;1(3):385-401. [doi: 10.1177/014662167700100306]
21. Bandeen-Roche K, Xue QL, Ferrucci L, et al. Phenotype of frailty: characterization in the women's health and aging studies. J Gerontol A Biol Sci Med Sci 2006 Mar;61(3):262-266. [doi: 10.1093/gerona/61.3.262] [Medline: 16567375]
22. Chen LK, Arai H. Physio-cognitive decline as the accelerated aging phenotype. Arch Gerontol Geriatr 2020;88:104051. [doi: 10.1016/j.archger.2020.104051] [Medline: 32278485]
23. Chen S, Sudharsanan N, Huang F, Liu Y, Geldsetzer P, Bärnighausen T. Impact of community based screening for hypertension on blood pressure after two years: regression discontinuity analysis in a national cohort of older adults in China. Br Med J 2019 Jul 11;366:l4064. [doi: 10.1136/bmj.l4064] [Medline: 31296584]
24. Geldsetzer P, Manne-Goehler J, Marcus ME, et al. The state of hypertension care in 44 low-income and middle-income countries: a cross-sectional study of nationally representative individual-level data from 1·1 million adults. Lancet 2019 Aug 24;394(10199):652-662. [doi: 10.1016/S0140-6736(19)30955-9] [Medline: 31327566]
25. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 1999 Jun;94(446):496-509. [doi: 10.1080/01621459.1999.10474144]
26. Ngandu T, Lehtisalo J, Solomon A, et al. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. The Lancet 2015 Jun;385(9984):2255-2263. [doi: 10.1016/S0140-6736(15)60461-5]
27. Crous-Bou M, Minguillón C, Gramunt N, Molinuevo JL. Alzheimer's disease prevention: from risk factors to early intervention. Alz Res Therapy 2017 Dec;9(1):71. [doi: 10.1186/s13195-017-0297-z]
28. Sekhon H, Allali G, Launay CP, et al. Motoric cognitive risk syndrome, incident cognitive impairment and morphological brain abnormalities: systematic review and meta-analysis. Maturitas 2019 May;123:45-54. [doi: 10.1016/j.maturitas.2019.02.006] [Medline: 31027677]

XSL•FO
RenderX

29. Bu Z, Huang A, Xue M, Li Q, Bai Y, Xu G. Cognitive frailty as a predictor of adverse outcomes among older adults: a systematic review and meta-analysis. Brain Behav 2021 Jan;11(1):e01926. [doi: 10.1002/brb3.1926] [Medline: 33159430]
30. Zhang XM, Wu XJ, Cao J, Jiao J, Chen W. Association between cognitive frailty and adverse outcomes among older adults: a meta-analysis. J Nutr Health Aging 2022;26(9):817-825. [doi: 10.1007/s12603-022-1833-5] [Medline: 36156673]
31. Wu YH, Liu LK, Chen WT, et al. Cognitive function in individuals with physical frailty but without dementia or cognitive complaints: results from the I-Lan Longitudinal Aging Study. J Am Med Dir Assoc 2015 Oct 1;16(10):899. [doi: 10.1016/j.jamda.2015.07.013] [Medline: 26321467]
32. Liu LK, Chen CH, Lee WJ, et al. Cognitive frailty and its association with all-cause mortality among community-dwelling older adults in Taiwan: results from I-Lan Longitudinal Aging Study. Rejuvenation Res 2018 Dec;21(6):510-517. [doi: 10.1089/rej.2017.2038] [Medline: 29644921]
33. Chen LK, Hwang AC, Lee WJ, et al. Efficacy of multidomain interventions to improve physical frailty, depression and cognition: data from cluster-randomized controlled trials. J Cachexia Sarcopenia Muscle 2020 Jun;11(3):650-662. [doi: 10.1002/jcsm.12534] [Medline: 32134208]
34. Merchant RA, Chan YH, Hui RJY, et al. Motoric cognitive risk syndrome, physio-cognitive decline syndrome, cognitive frailty and reversibility with dual-task exercise. Exp Gerontol 2021 Jul 15;150:111362. [doi: 10.1016/j.exger.2021.111362] [Medline: 33887381]
35. Lee WJ, Peng LN, Lin MH, et al. Six-year transition of physio-cognitive decline syndrome: results from I-Lan Longitudinal Aging Study. Arch Gerontol Geriatr 2022;102:104743. [doi: 10.1016/j.archger.2022.104743] [Medline: 35687948]
36. Zaccardi F, Franks PW, Dudbridge F, Davies MJ, Khunti K, Yates T. Mortality risk comparing walking pace to handgrip strength and a healthy lifestyle: a UK Biobank study. Eur J Prev Cardiol 2021 Jul 10;28(7):704-712. [doi: 10.1177/2047487319885041] [Medline: 34247229]
37. Verghese J, Lipton RB, Hall CB, Kuslansky G, Katz MJ, Buschke H. Abnormality of gait as a predictor of non-Alzheimer's dementia. N Engl J Med 2002 Nov 28;347(22):1761-1768. [doi: 10.1056/NEJMoa020441] [Medline: 12456852]
38. Sekhon H, Launay CP, Chabot J, Allali G, Beauchet O. Motoric cognitive risk syndrome: could it be defined through increased five-times-sit-to-stand test time, rather than slow walking speed? Front Aging Neurosci 2018;10:434. [doi: 10.3389/fnagi.2018.00434] [Medline: 30766485]
39. Alexander LK, Lopes B, Ricchetti-Masterson K, Yeatts KB. Sources of systematic error or bias: Information bias. ERIC Notebook 2015;2(14):1-5.
40. Baumgart M, Snyder HM, Carrillo MC, Fazio S, Kim H, Johns H. Summary of the evidence on modifiable risk factors for cognitive decline and dementia: a population-based perspective. Alzheimers Dement 2015 Jun;11(6):718-726. [doi: 10.1016/j.jalz.2015.05.016] [Medline: 26045020]
41. Campbell NL, Unverzagt F, LaMantia MA, Khan BA, Boustani MA. Risk factors for the progression of mild cognitive impairment to dementia. Clin Geriatr Med 2013 Nov;29(4):873-893. [doi: 10.1016/j.cger.2013.07.009] [Medline: 24094301]
42. HRS data portal. Health and Retirement Study. URL: https://hrsdata.isr.umich.edu/ [accessed 2025-01-28]

Abbreviations

AHR: adjusted hazard ratio
CF: cognitive frailty
HRS: Health and Retirement Study
MCI: mild cognitive impairment
MCR: motoric cognitive risk syndrome
MET: metabolic equivalent of task
OR: odds ratio
PCDS: physio-cognitive decline syndrome
TICS-m: Modified Telephone Interview for Cognitive Status

Edited by Y Du; submitted 03.09.24; peer-reviewed by L Qin, M Colombo, S Wang; revised version received 27.11.24; accepted 05.12.24; published 30.01.25.

Please cite as:
Bai A, He S, Jiang Y, Xu W, Lin Z
Comparison of 3 Aging Metrics in Dual Declines to Capture All-Cause Dementia and Mortality Risk: Cohort Study
JMIR Aging 2025;8:e66104
URL: https://aging.jmir.org/2025/1/e66104
doi:10.2196/66104

XSL•FO
RenderX

© Anying Bai, Shan He, Yu Jiang, Weihao Xu, Zhanyi Lin. Originally published in JMIR Aging (https://aging.jmir.org), 30.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Building Consensus on the Relevant Criteria to Screen for Depressive Symptoms Among Near-Centenarians and Centenarians: Modified e-Delphi Study

Carla Gomes da Rocha[1,2,3], MScN; Armin von Gunten[2], MPhil, MD; Pierre Vandel[2], PhD; Daniela S Jopp[4,5], PhD; Olga Ribeiro[6], PhD; Henk Verloo[1], PhD

[1]School of Health Sciences, University of Applied Sciences and Arts Western Switzerland (HES-SO), Sion, Switzerland

[2]Service of Old Age Psychiatry, Lausanne University Hospital and University of Lausanne, Prilly-Lausanne, Switzerland

[3]Institute of Biomedical Sciences Abel Salazar, University of Porto, Porto, Portugal

[4]Institute of Psychology, University of Lausanne, Lausanne, Switzerland

[5]Swiss National Centre of Competence in Research LIVES, University of Lausanne, Lausanne, Switzerland

[6]Nursing School of Porto (ESEP) and RISE-Health, Porto, Portugal

Corresponding Author:
Carla Gomes da Rocha, MScN
School of Health Sciences
University of Applied Sciences and Arts Western Switzerland (HES-SO)
Chemin de l'Agasse 5
Sion, 1950
Switzerland
Phone: 41 58 606 84 73
Email: carla.gomesdarocha@hevs.ch

Abstract

Background: The number of centenarians worldwide is expected to increase dramatically, reaching 3.4 million by 2050 and >25 million by 2100. Despite these projections, depression remains a prevalent yet underdiagnosed and undertreated condition among this population that carries significant health risks.

Objective: This study aimed to identify and achieve consensus on the most representative signs and symptoms of depression in near-centenarians and centenarians (aged ≥95 years) through an e-Delphi study with an international and interdisciplinary panel of experts. Ultimately, the outcomes of this study might help create a screening instrument that is specifically designed for this unique population.

Methods: A modified e-Delphi study was carried out to achieve expert consensus on depressive symptoms in near-centenarians and centenarians. A panel of 28 international experts was recruited. Consensus was defined as 70% agreement on the relevance of each item. Data were collected through a web-based questionnaire over 3 rounds. Experts rated 104 items that were divided into 24 dimensions and 80 criteria to identify the most representative signs and symptoms of depression in this age group.

Results: The panel consisted of experts from various countries, including physicians with experience in old age psychiatry or geriatrics as well as nurses and psychologists. The response rate remained consistent over the rounds (20/28, 71% to 21/28, 75%). In total, 4 new dimensions and 8 new criteria were proposed by the experts, and consensus was reached on 86% (24/28) of the dimensions and 80% (70/88) of the criteria. The most consensual potentially relevant dimensions were *lack of hope* (21/21, 100%), *loss of interest* (27/28, 96%), *lack of reactivity to pleasant events* (27/28, 96%), *depressed mood* (26/28, 93%), and *previous episodes of depression or diagnosed depression* (19/21, 90%). In addition, the most consensual potentially relevant criteria were *despondency, gloom, and despair* (25/25, 100%); *depressed* (27/27, 100%); *lack of reactivity to pleasant events or circumstances* (28/28, 100%); *suicidal ideation* (28/28, 100%); *suicide attempt(s)* (28/28, 100%); *ruminations* (27/28, 96%); *recurrent thoughts of death or suicide* (27/28, 96%); *feelings of worthlessness* (25/26, 96%); *critical life events* (20/21, 95%); *anhedonia* (20/21, 95%); *loss of interest in activities* (26/28, 93%); *loss of pleasure in activities* (26/28, 93%); and *sadness* (24/26, 92%). Moreover, when assessing depression in very old age, the duration, number, frequency, and severity of signs and symptoms should also be considered, as evidenced by the high expert agreement.

XSL•FO
RenderX

Conclusions: The classification of most elements as *relevant* highlights the importance of a multidimensional approach for optimal depression screening among individuals of very old age. This study offers a first step toward improving depression assessment in near-centenarians and centenarians. The development of a more adapted screening tool could improve early detection and intervention, enhancing the quality of mental health care for this population.

(JMIR Aging 2025;8:e64352) doi:10.2196/64352

KEYWORDS

centenarians; near-centenarians; depressive symptoms; depression diagnosis; screening; assessment; e-Delphi technique; web-based survey

Introduction

Background

The centenarian population is growing significantly and is projected to increase sharply in the coming decades. As of 2024, there were approximately 722,000 centenarians worldwide, with Japan, the United States, and China having the highest numbers [1]. This figure is expected to rise to 3.4 million by 2050 and exceed 25 million by 2100 [2,3]. In Europe, notable examples include France, Italy, and Greece, each with >20 centenarians per 100,000 inhabitants registered already in 2011. Portugal also stands out with a density of 14.4 centenarians per 100,000 inhabitants that same year, ranking it among the top 15 in Europe [4]. In Switzerland, the centenarian population more than doubled from 787 in 2000 to 1948 in 2022, reaching a noteworthy prevalence of 22.1 per 100,000 inhabitants [5].

Amid extreme longevity, individuals' health experiences vary significantly, from robust health to substantial declines [6-8]. Depression, a prevalent mood disorder among older adults [9], emerges as a major public health concern [10]. The *International Classification of Diseases, 11th Revision* characterizes a depressive episode as "a period of depressed mood or diminished interest in activities occurring most of the day, nearly every day during a period lasting at least two weeks accompanied by other symptoms such as difficulty concentrating, feelings of worthlessness or excessive or inappropriate guilt, hopelessness, recurrent thoughts of death or suicide, changes in appetite or sleep, psychomotor agitation or retardation, and reduced energy or fatigue" [6]. However, this definition broadly encompasses the general population and may not fully address the possible difference and heterogeneity of depressive symptomatology in older adults, potentially leading to diagnostic imprecision.

Accurate diagnosis is crucial as depression should not be misconstrued as an inherent aspect of aging even in very advanced age [11,12]. A systematic literature review reported substantial variations in depression prevalence among near-centenarians and centenarians across countries, such as 12.8% in Italy, 13.5% in Australia, 20% in the United States, and 29% in Mexico [13]. These differences may reflect cultural, health care, and socioeconomic influences as well as methodological variations in study conduct [13,14]. Interestingly, comparisons between centenarian cohorts and younger groups have shown inconsistent findings, with half of the studies reporting higher rates of depressive symptoms among centenarians [13]. This highlights the diversity of aging experiences and suggests that mental health may not necessarily worsen with age. Particularly noteworthy are the findings from Sweden, where near-centenarians and centenarians exhibited almost double the depression rates of older adults aged 85 years at 32.3% versus 16.8%, respectively. These statistics highlight the need not only for careful mental health evaluations among individuals of very old age but also for prompt discussion on tailored approaches to support mental well-being in this population [13,15].

Depression can lead to severe physical, cognitive, and psychological effects. Physically, it can cause malnutrition, falls, delirium, functional decline, and increased mortality [13,16-20]. Cognitively, depression has been connected to cognitive decline and dementia [21-25]. Psychologically, it has been associated with anxiety and suicidal ideation [26,27]. Furthermore, depression places a significant burden on caregivers and health care systems [16-18,28,29], increasing costs and resource demands. These consequences underscore the critical need for early detection and intervention. Therefore, proactive measures such as regular mental health screenings are essential to prevent the severe impact of depression on individuals' quality of life and alleviate its wider societal effects [11]; however, these are not part of the current practice in most countries worldwide.

Addressing this gap in care is difficult as depression in older adults frequently remains underdiagnosed and undertreated, in part because its manifestations are often misinterpreted as natural aspects of aging [18,19,30,31]. Studies have estimated that approximately half of depression cases are not identified by frontline health care providers [16,18], and of those detected, approximately half do not receive adequate treatment [16]. The challenge in detecting late-life depression (LLD) stems from its sometimes subtle, hidden, or atypical signs and symptoms [30,32]. Physical manifestations such as weight loss, psychomotor retardation, and exhaustion, as well as emotional manifestations such as loss of interest in activities, heightened anxiety, or irritability, might overshadow the typical sadness often associated with depression [9,18]. This complexity requires greater clinical vigilance. Moreover, the possible overlap of symptoms between depression and dementia adds another layer of diagnostic difficulty as both conditions can present with symptoms such as memory impairment, psychomotor retardation, and reduced motivation [9]. Beyond dementia, several other medical conditions can mimic or exacerbate depressive symptoms. Chronic pain, thyroid dysfunction, Parkinson disease, stroke, and medication side effects—particularly in polymedicated individuals—can contribute to *depression-like* symptoms, including fatigue, apathy, functional dependence, memory impairment, and

XSL•FO
RenderX

disturbances in mood and sleep [33-38]. These overlapping symptoms underscore the need for a comprehensive assessment to distinguish LLD from other conditions, aiming for an accurate differential diagnosis. The wide range of signs and symptoms associated with LLD can extend beyond the conventional diagnostic criteria for depressive disorders. Symptom clusters not fully meeting the diagnostic criteria for depression may still carry significant clinical importance due to their link to reduced quality of life and increased disability [9,18].

Objectives

Given the prevalence of depression among individuals of very old age, better detection of early signs and symptoms is crucial. However, despite the availability of various screening tools, to the best of our knowledge, none have been specifically validated for near-centenarians and centenarians (aged ≥95 years). Therefore, our study aimed to compile a comprehensive range of potentially relevant depressive features that have been documented in scientific literature. These features were then submitted to an international panel of experts to achieve consensus on the most representative signs and symptoms of depression in individuals of very old age. The outcomes of this approach are 2-fold. First, by systematically examining and integrating diverse diagnostic criteria and screening tools, this study aimed to contribute to the ongoing debate about the complexities and challenges of effectively diagnosing depression in individuals of very old age. Second, achieving expert consensus is central to developing a new screening instrument specifically tailored to this unique population. This effort might be an important first step toward early detection and effective intervention, which might lead to better outcomes and enhanced general well-being for this age group, and seeks to contribute to promoting mental health and supporting optimal care for individuals of very old age.

Methods

Study Design

We conducted a modified e-Delphi study to systematically combine expert opinions and achieve an informed group consensus on the symptomatology of depression in very old age. The Delphi technique is an established approach in which a panel of experts is asked to provide their opinions on a given issue over the course of several rounds [39]. In each round, the questions are informed by the findings of the previous one, allowing the study to evolve over time based on the collected data [39].

Given the prevalence of depression among older adults and the complexities of its diagnosis, the Delphi technique is particularly useful for gathering expert consensus on this issue. In our study, we used an e-Delphi approach to involve experts from different locations and facilitate efficient data collection through electronic communication methods. To adapt the traditional Delphi approach to the specific requirements of our study, we implemented several modifications (referred to as a *modified e-Delphi study*). To keep the focus on the items rather than on the degree of agreement, experts were informed at the start of each new round about which items had reached consensus without being given the precise consensus rates. In addition, for items on which consensus had already been reached, experts were unable to review their previous responses. This decision was made to avoid revisiting settled issues and manage the large number of items efficiently, ensuring the stability of agreed-upon elements. An additional aspect of our modified methodology was a final evaluation of open issues by the steering committee. This last step took place to discuss the items that had not achieved consensus over the e-Delphi rounds, ensuring a thorough examination of these elements.

The materials presented to the e-Delphi panel were based on a comprehensive analysis of the existing literature on depression in very old age (personal communication by Gomes da Rocha et al, 2023). This ensured that the expert panel's input was based on a thorough understanding of the relevant literature, thereby enhancing the validity of the final consensus, which may be seen as a lever to encourage scientific debate and foster new developments toward an effective diagnosis of depression in very old age [40].

This study adhered to the Conducting and Reporting Delphi Studies guidelines [41,42].

Steering Committee

The steering committee for this e-Delphi study based in Western Switzerland comprised 4 members: 2 physicians who are old age psychiatrists and senior researchers (AvG and PV), 1 clinical nurse specialist in geriatric care and a senior researcher skilled in the e-Delphi technique (HV), and 1 clinical nurse specialist in geriatric care who is also a junior researcher (CGdR). The committee was responsible for overseeing the study design, ensuring ethical compliance, and guiding the iterative process of questionnaire refinement and consensus building. Meetings were held monthly, with methodological decisions made based on discussion. The committee played a crucial role in maintaining the study's rigor, managing data collection and analysis, and addressing any lack of consensus among panelists.

Selection of Experts

In the Delphi approach, nonprobability sampling methods are used through the purposive selection of an expert panel. However, clear selection criteria were applied to limit researcher bias based on a five-step procedure [43]:

1. *Identifying the most appropriate categories of experts for the panel:* we identified experts from 3 main professional categories—physicians, nurses, and psychologists—each with a minimum of 5 years of professional experience and competencies in LLD. These professionals were selected for their specific work experience or contact with individuals of very old age in clinical, research, or teaching domains within the fields of old age psychiatry or geriatrics. In addition, our recruitment strategy aimed for a broad geographical representation by including experts from various continents (ie, Asia, Australia, Europe, and North America) to enrich the panel with diverse perspectives and, thereby, enhance the global view on depression screening among individuals of very old age.
2. *Identifying experts:* names were compiled from various sources, including previous research involvement, publications on the subject, professional email lists,

XSL•FO
RenderX

professional associations or societies, and boards of professional organizations. This method aimed to create a list of potential panelists recognized for their expertise and contributions to the field that was as comprehensive as possible.

3. *Contacting some readily reachable experts from each professional category to nominate other experts:* initial contact was made with experts who were easily accessible. These experts were then asked to nominate additional professionals whom they believed could contribute valuable insights to the study, thereby using a snowball sampling technique to ensure a wide and diverse pool of expertise.
4. *Creating sublists for each professional category and ranking experts based on specific criteria:* experts were organized into sublists by professional category and ranked according to their representation of professional role or specialty and their field of practice. This was a strategic approach intended to foster a diverse and balanced panel that reflects a broad spectrum of experiences and viewpoints, contributing to the robustness and credibility of the consensus process.
5. *Inviting the experts.*

The Delphi method offers no consensus on an ideal sample size as the focus is on qualitative depth and the use of expert judgment for a comprehensive exploration of specific topics rather than on statistical representativeness [44]. Usually, a panel comprises 7 to 15 experts to facilitate effective information processing and decision-making; however, some studies may involve up to 30 experts for broader perspectives [44]. For our study, we aimed to assemble a convenience sample of at least 20 experts from various countries.

Recruitment Procedure

On the basis of the predefined selection of potential panelists, we emailed invitation letters to a total of 84 experts. These letters outlined the study's objectives, described the e-Delphi context, detailed the questionnaire (eg, language, type of questions, and estimated completion time), and clarified what was expected from the participants, ensuring that they understood the significance of their contributions to the study. Interested parties could access the questionnaire through a direct URL to the survey. Anticipating a 25% to 30% response rate for the first-round questionnaire, we expected approximately 20 to 25 experts to participate. We aimed to retain these experts across all subsequent rounds, planning for a maximum of 4 rounds.

Data Collection Procedures

The SurveyMonkey tool (SurveyMonkey Inc) was used as the survey software as it facilitates secure and efficient data collection. In accordance with literature recommendations [43], each e-Delphi round remained open for a minimum of 4 weeks, during which up to 3 automatic reminder emails were dispatched to nonrespondents. At the start of each new round, experts were provided access to the overall results from the previous round, including a list of items for which consensus was either *achieved* (≥70% agreement) or *not achieved* (<70% agreement) [43]. The items that did not reach consensus were integrated into the current round for further deliberation, along with any new items that emerged from panel suggestions through open-ended questions.

Measures

The questionnaire was divided into 3 sections.

e-Delphi Section

This section aimed to achieve consensus on the topic under study and consisted of 104 items. These were divided into 24 dimensions (eg, *impaired concentration* and *lack of hope*) and 80 criteria (eg, *impaired ability to think or concentrate* and *hopelessness*), which experts rated according to their relevance using a Likert scale from 1 (*Not relevant*) to 4 (*Very relevant*). Experts also had the possibility to comment on their answers and suggest additional items that they considered relevant to the topic for subsequent rounds.

In addition to rating the relevance of each item, experts were asked to indicate the most suitable type of assessment for each criterion: self-assessment, hetero-assessment, or both. This was intended to identify preferred methods of assessing potential depressive signs and symptoms. The aim was not to achieve a stringent consensus of 70% as with the relevance ratings but to capture a broad spectrum of expert opinions on assessment approaches.

The initial selection of 104 items was based on a comprehensive review of existing instruments widely used in research to screen for and evaluate depressive signs and symptoms among near-centenarians and centenarians [13] supplemented by additional insights from recent analysis (personal communication by Gomes da Rocha et al, 2023). We started with all items available in these measures to be as comprehensive as possible. These instruments included self-assessment tools such as the Geriatric Depression Scale–30 [45], the Center for Epidemiologic Studies Depression Scale [46], the depression-related items from the Brief Symptom Inventory [47], and the Hospital Anxiety and Depression Scale [48], as well as the Montgomery-Åsberg Depression Rating Scale [49], which is administered by a professional rater. Furthermore, given the significant prevalence of neurocognitive disorders in our aging population and the need for reliable assessment of depressive symptoms in individuals with cognitive impairments, we also considered instruments specifically designed for such assessments: the Cornell Scale for Depression in Dementia [50,51] and the depression section of the Neuropsychiatric Inventory [52,53]. In addition, we included the *International Classification of Diseases, 11th Revision* and the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* diagnostic criteria to anchor our selections in the gold standards for depression diagnosis [6,54]. To complete the pool of items, the Depression Rating Scale from the Resident Assessment Instrument was also retained given its widespread use among adults of old and very old age in diverse settings, such as private homes (Resident Assessment Instrument–Home Care) or long-term care facilities (Resident Assessment Instrument–Long-Term Care Facilities) [17,55-57]. After compiling all items from the selected instruments, we checked for overlap and excluded identical items, which led to the final selection of 104 items. For instance, items such as "Do you feel

XSL•FO
RenderX

that your situation is hopeless?" from the Geriatric Depression Scale–30 and "Feeling hopeless about the future" from the Brief Symptom Inventory necessitated thematic grouping under a unified dimension (*Lack of Hope* in this example). This process enabled us to construct more general items designated as *dimensions* that reflected the essence of the criteria within each group, thereby streamlining the questionnaire without compromising the depth of depressive symptom screening. The list of all selected instruments and original items can be found in Multimedia Appendix 1.

In summary, this strategic selection of items aimed to cover a broad spectrum of relevant criteria for the screening of depressive signs and symptoms, ensuring the questionnaire's comprehensive scope and its applicability to adults of very old age with varying levels of cognitive functioning.

Sociodemographic and Professional Characteristics

This section included questions designed to capture the sociodemographic and professional characteristics of the panelists in round 1, including age, gender, nationality, profession, years of professional experience, and their country and field of practice. In subsequent rounds, only essential information (age, gender, nationality, and profession) was collected to ensure continuity and to accurately link responses from the same panelists across all rounds.

Contact With Individuals of Very Old Age

Although this could be considered part of professional characterization, this section specifically gathered information on the frequency of the panelists' experiences or interactions with individuals of very old age (aged ≥85 and ≥95 years). It was specific to the first round, aiming to inform the context of their responses. While this study's target population was defined as near-centenarians and centenarians (≥95 years), there is only a limited number of experts focusing on this exact population. However, most experts working with individuals aged ≥85 years have regular contact with near-centenarians and centenarians. Thus, these experts were also invited provided that they felt confident in drawing on their experiences and perceptions relevant to the population aged ≥95 years.

Pretest

Before launching round 1, the questionnaire was pretested with a small purposive sample of 5 professionals (nurses, psychologists, and a physician) to verify its clarity and comprehensibility, with only a minor adjustment made based on 1 expert's suggestion to include specific examples for certain items. Despite the positive outcomes of the pretest, experts were invited to rate the questionnaire at the end of round 1 for comprehensiveness, readability, and clarity.

Overall, 93% (26/28) of the experts participating in round 1 rated the questionnaire as *Comprehensive* or *Very comprehensive*, and 86% (24/28) found it *Easy* or *Very easy* to read and understand. Due to this positive feedback, no further structural adjustments were implemented before proceeding to the subsequent rounds.

Data Analysis

Data collected via SurveyMonkey were transferred to a Microsoft Excel spreadsheet (Microsoft Corp) and examined for any potential errors or missing data. They were then imported into the SPSS Statistics software (version 29.0; IBM Corp) for analysis [58]. Descriptive statistics were computed to characterize the panel from a sociodemographic and professional perspective using means and SDs as well as medians and interquartile ranges for quantitative variables and frequency distributions and percentages for qualitative variables. The level of agreement among experts required to achieve consensus on a given item was set at 70% for each round. This threshold is commonly used in Delphi studies [59-61] as it allows for the inclusion of minority viewpoints in the early rounds, ensuring that diverse perspectives are considered before achieving final consensus. This is particularly relevant in the context of depressive symptom screening given the challenges associated with detecting LLD, as previously mentioned. Hence, this consensus rate is considered acceptable for preserving methodological rigor, as supported by the literature [43], which indicates that such a threshold aligns with norms for ensuring reliability and validity in the context of an iterative, consensus-building approach. The consensus rate for each item was calculated using a 4-level Likert scale (1=*Not relevant*, 2=*Somewhat relevant*, 3=*Relevant*, and 4=*Very relevant*). In the first step, these responses were recoded into 2 levels: values of 1 and 2 were recoded as "0" (*Overall not relevant*), and values of 3 and 4 were recoded as "1" (*Overall relevant*). Next, we aggregated the recoded responses for each item to determine consensus, which was achieved when the percentage of experts identifying an item as *Overall relevant* or *Overall not relevant* met or exceeded the predefined threshold of 70%.

After the first e-Delphi round, items that did not reach the consensus level were carried forward to the subsequent round. At the beginning of each subsequent round, experts were given the opportunity to review the general outcomes from the previous round (ie, *consensus achieved* or *consensus not achieved* for each evaluated item). Items that achieved consensus were not included again in subsequent rounds. After conducting 3 rounds with the participating panel and observing only limited changes in the consensus rate after the third round, we decided to proceed with the last round, which involved a final evaluation of unresolved aspects by the steering committee. This last step aimed to gain a deeper understanding of the reasoning behind the outcomes and further refine the consensus process. Items that still did not achieve consensus after the 3 e-Delphi rounds were thoroughly examined during this phase.

Ethical Considerations

The Ethics Commission on Human Research of the Canton of Vaud was consulted and confirmed that this project fell outside the scope of the Swiss Human Research Act (Req-2023-00844; Multimedia Appendix 2) as the study did not involve the collection of health-related personal data or biological material.

All participants received detailed information about the study, including its aim and relevance, what to expect, a description of the questionnaire, instructions for participation, and contact information for the main researcher, along with the invitation

XSL•FO
RenderX

email. The questionnaires for all rounds (1, 2, and 3) included an introductory section emphasizing the voluntary nature of participation in the study, as well as specific information regarding confidentiality, data protection, risks, and benefits and a statement on ethical considerations. Participants were explicitly informed that, by answering the questionnaire, they were implicitly providing consent to take part in the study. They did not receive any compensation for their participation. Each participant received a survey link for accessing the questionnaire, and their IP addresses were not recorded to ensure confidentiality. The collected data were deidentified, with no personal identifiers retained. Responses stored on SurveyMonkey were deleted after transferring the data to the Microsoft Excel spreadsheet. All data are securely stored on a password-protected institutional server to comply with Swiss data protection regulations and will be definitively deleted after 5 years.

Results

Expert Participation and Sociodemographic and Professional Characteristics

A total of 84 experts were invited to this modified e-Delphi study. Of these 84 experts, 28 (33%) participated in the first round, forming the baseline panel. This group (n=28) was consequently invited to the second and third rounds, for which the participation rates were 75% (21/28) and 71% (20/28), respectively.

The panel exhibited sociodemographic and professional diversity, including different nationalities, professions, and fields of practice. The average age ranged slightly between 44.1 (SD 10.6) and 46.7 (SD 11.4) years over the rounds. A majority of women participated in the first and second rounds (17/28, 61% and 13/21, 62%, respectively), whereas men were more prevalent in the third round (11/20, 55%). Swiss and Portuguese nationalities were the most represented (12/20, 60% to 18/28, 64%), with the nursing profession being predominant (9/20, 45% to 12/21, 57%), followed by the medical profession (5/21, 24% to 8/20, 40%). The average professional experience was 19.4 (SD 8-9) years, with nearly 80% (22/28, 79%) of the experts being engaged in purely clinical activities or a combination of clinical and academic activities, among others (eg, service organization). More than 80% (23/28, 82%) reported having regular or very regular contact (daily, weekly, or monthly) with individuals aged ≥85 years, and 75% (21/28) reported having regular or very regular contact with individuals aged ≥95 years (Table 1).

XSL•FO
RenderX

Table 1. Sociodemographic and professional characteristics of the experts from the panel.

	Round 1 (n=28)	Round 2 (n=21)	Round 3 (n=20)
Age (y)			
Values, mean (SD)	44.1 (10.6)	45.2 (11.6)	46.7 (11.4)
Values, median (IQR)	39.5 (18.8)	39.0 (22.0)	44.0 (22.8)
Gender, n (%)			
Men	11 (39)	8 (38)	11 (55)
Women	17 (61)	13 (62)	9 (45)
Nationality, n (%)			
Portuguese	9 (32)	6 (29)	4 (20)
Swiss	9 (32)	7 (33)	8 (40)
French	3 (11)	2 (10)	3 (15)
Swiss and Portuguese	2 (7)	2 (10)	1 (5)
Australian	1 (4)	1 (5)	1 (5)
Belgian	1 (4)	1 (5)	1 (5)
German	1 (4)	1 (5)	0 (0)
Italian	1 (4)	0 (0)	1 (5)
Norwegian	1 (4)	1 (5)	1 (5)
Profession, n (%)			
Nurse	15 (54)	12 (57)	9 (45)
Physician	8 (29)	5 (24)	8 (40)
Psychologist	4 (14)	4 (19)	3 (15)
Other: gerontologist	1 (4)	0 (0)	0 (0)
Professional experience (y; n=27)			
Values, mean (SD)	19.4 (8.9)	__[a]	—
Values, median (IQR)	17.0 (11.0	—	—
Country of professional practice, n (%)			
Switzerland	15 (54)	—	—
Portugal	7 (25)	—	—
France	3 (11)	—	—
Australia	1 (4)	—	—
Norway	1 (4)	—	—
Portugal, United Kingdom, Saudi Arabia, and UAE[b]	1 (4)	—	—
Field of professional activity, n (%)			
Clinical activity			
Inpatient care only	6 (21)	—	—
Outpatient care only	2 (7)	—	—
Long-term care only	2 (7)	—	—
Academic activity			
Teaching only	1 (4)	—	—
Research only	1 (4)	—	—
Teaching and research	3 (11)	—	—
Clinical activity[c] and academic activity[d]	9 (32)	—	—

XSL•FO
RenderX

	Round 1 (n=28)	Round 2 (n=21)	Round 3 (n=20)
Clinical activity[c], academic activity[d], and other (eg, quality improvement or service organization)	3 (11)	—	—
Management activity	1 (4)	—	—
Contact with individuals aged ≥85 years, n (%)			
Daily	10 (36)	—	—
Weekly	9 (32)	—	—
Monthly	4 (14)	—	—
Yearly	4 (14)	—	—
Previously but not currently	1 (4)	—	—
Contact with individuals aged ≥95 years, n (%)			
Daily	4 (14)	—	—
Weekly	8 (29)	—	—
Monthly	9 (32)	—	—
Yearly (or less[e])	6 (21)	—	—
Previously but not currently	1 (4)	—	—

[a]Not applicable.

[b]UAE: United Arab Emirates.

[c]*Clinical activity* includes inpatient, outpatient, or long-term care settings.

[d]*Academic activity* includes teaching or research roles.

[e]One expert commented the following: "Not yearly, but once in a while."

Questionnaire Completion Times

The completion dates and the mean time taken for participants to complete the questionnaires in each round are summarized in Table 2. Globally, the e-Delphi rounds took place between August 2023 and February 2024, with an overall weighted mean completion time of 16 minutes and 40 seconds (SD 11 min and 27 s).

Table 2. Summary of overall consensus achievements across e-Delphi rounds regarding potential dimensions and criteria for depression screening.

	Questionnaire completion dates	Questionnaire completion time, mean (SD)	Dimensions analyzed	Consensus achieved (dimensions), n (%)	Consensus not achieved (dimensions), n (%)	New dimensions proposed	Criteria analyzed	Consensus achieved (criteria), n (%)	Consensus not achieved (criteria), n (%)	New criteria proposed
Round 1	August 15, 2023, to October 1, 2023	24 min, 51 s (10 min, 51 s)	24	13 (54)[a]	11 (46)[a]	4	80	49 (61)[b]	31 (39)[b]	8
Round 2	October 26, 2023, to December 12, 2023	15 min, 31 s (8 min, 18 s)	15	10 (67)[c]	5 (33)[c]	0	39	12 (31)[d]	27 (69)[d]	0
Round 3	January 31, 2024, to February 29, 2024	6 min, 26 s (3 min, 55 s)	5	1 (20)[e]	4 (80)[e]	0	27	9 (33)[f]	18 (67)[f]	0

[a]n=24.

[b]n=80.

[c]n=15.

[d]n=39.

[e]n=5.

[f]n=27.

XSL•FO
RenderX

Overall Consensus Achievements

A consensus threshold of ≥70% was established for both dimensions and criteria, as previously described.

Depression Screening: Potentially Relevant Dimensions

In round 1, experts assessed 24 dimensions, reaching consensus on 13 (54%). In addition, 4 new dimensions were suggested by the experts and incorporated in round 2: *underlying cognitive disorder* (dimension 25), *previous episodes of depression or diagnosed depression* (dimension 26), *personal care (eg, poor hygiene)* (dimension 27), and *language/speech (eg, poor speech)* (dimension 28). Thus, 15 dimensions (n=11, 73% without consensus and n=4, 27% new proposals) were reassessed in round 2, with consensus achieved on 10 (67%). In round 3, a total of 5 dimensions were revisited, with consensus reached on only 1 (20%; Table 2).

Table 3 summarizes the consensus rates for the 28 dimensions assessed across the 3 e-Delphi rounds. A total of 14% (4/28) of the dimensions failed to reach the ≥70% consensus threshold: memory problems (dimension 21), mood-congruent delusions (dimension 22), reduced adaptation capacity (dimension 24), and language/speech (eg, poor speech) (dimension 28)—the only newly suggested dimension without consensus. Of the dimensions on which consensus was achieved, all were considered relevant for screening for depression in individuals aged ≥95 years except for increased appetite (dimension 8), which was deemed not relevant. Notably, several relevant dimensions reached ≥90% consensus: lack of hope (dimension 20; 21/21, 100%), loss of interest (dimension 4; 27/28, 96%), lack of reactivity to pleasant events (dimension 5; 27/28, 96%), depressed mood (dimension 1; 26/28, 93%), and previous episodes of depression or diagnosed depression (dimension 26; 19/21, 90%).

Figure S1 in Multimedia Appendix 3 provides an overview of the consensus rates for the potentially relevant dimensions for depression screening.

XSL•FO
RenderX

Table 3. Consensus level on potential dimensions for depression screening[a].

	Round 1 (n=28)			Round 2 (n=21)			Round 3 (n=20)		
	Level of consensus at ≥70%	Result	Participants, n (%)	Level of consensus at ≥70%	Result	Participants, n (%)	Level of consensus at ≥70%	Result	Participants, n (%)
Depressed mood	Achieved	Relevant	26 (93)[b]	—[c]	—	—	—	—	—
Anxiety	Achieved	Relevant	24 (86)[b]	—	—	—	—	—	—
Irritability	Not achieved	N/A[d]	N/A	Achieved	Relevant	17 (81)[e]	—	—	—
Loss of interest	Achieved	Relevant	27 (96)[b]	—	—	—	—	—	—
Lack of reactivity to pleasant events	Achieved	Relevant	27 (96)[b]	—	—	—	—	—	—
Psychomotor changes	Not achieved	N/A	N/A	Achieved	Relevant	16 (76)[e]	—	—	—
Appetite loss	Achieved	Relevant	19 (70)[f]	—	—	—	—	—	—
Increased appetite	Not achieved	N/A	N/A	Achieved	Not relevant	16 (76)[e]	—	—	—
Reduced sleep or insomnia	Achieved	Relevant	19 (73)[g]	—	—	—	—	—	—
Hypersomnia	Not achieved	N/A	N/A	Achieved	Relevant	16 (76)[e]	—	—	—
Lack of energy	Achieved	Relevant	21 (78)[f]	—	—	—	—	—	—
Death wishes	Achieved	Relevant	22 (79)[b]	—	—	—	—	—	—
Suicidal thoughts	Achieved	Relevant	24 (86)[b]	—	—	—	—	—	—
Poor self-esteem	Achieved	Relevant	23 (82)[b]	—	—	—	—	—	—
Pessimism	Achieved	Relevant	20 (74)[f]	—	—	—	—	—	—
Low life satisfaction	Achieved	Relevant	20 (71)[b]	—	—	—	—	—	—
Impaired concentration	Not achieved	N/A	N/A	Not achieved	N/A	N/A	Achieved	Relevant	14 (70)[h]
Impaired decision-making capacity	Not achieved	N/A	N/A	Achieved	Relevant	15 (71)[e]	—	—	—
Negative feelings	Achieved	Relevant	24 (86)[b]	—	—	—	—	—	—
Lack of hope	Not achieved	N/A	N/A	Achieved	Relevant	21 (100)[e]	—	—	—
Memory problems	Not achieved	N/A	N/A	Not achieved	N/A	N/A	Not achieved	N/A	N/A
Mood-congruent delusions	Not achieved	N/A	N/A	Not achieved	N/A	N/A	Not achieved	N/A	N/A
Repercussions of functional dependence	Not achieved	N/A	N/A	Achieved	Relevant	16 (76)[e]	—	—	—
Reduced adaptation capacity	Not achieved	N/A	N/A	Not achieved	N/A	N/A	Not achieved	N/A	N/A
Underlying cognitive disorder (added in round 2)	N/A	N/A	N/A	Achieved	Relevant	15 (71)[e]	—	—	—
Previous episodes of depression or diagnosed depression (added in round 2)	N/A	N/A	N/A	Achieved	Relevant	19 (90)[e]	—	—	—

XSL•FO
RenderX

	Round 1 (n=28)			Round 2 (n=21)			Round 3 (n=20)		
	Level of consensus at ≥70%	Result	Participants, n (%)	Level of consensus at ≥70%	Result	Participants, n (%)	Level of consensus at ≥70%	Result	Participants, n (%)
Personal care (eg, poor hygiene; added in round 2)	N/A	N/A	N/A	Achieved	Relevant	16 (76)[e]	—	—	—
Language/speech (eg, poor speech; added in round 2)	N/A	N/A	N/A	Not achieved	N/A	N/A	Not achieved	N/A	N/A

aThis list includes 4 dimensions suggested by the expert participants in round 1 and incorporated into the subsequent rounds: underlying cognitive disorder, previous episodes of depression or diagnosed depression, personal care (eg, poor hygiene), and language/speech (eg, poor speech).

[b]n=28.

[c]Data are not available for the corresponding entries.

[d]N/A: not applicable.

[e]n=21.

[f]n=27.

[g]n=26.

[h]n=20.

Depression Screening: Potentially Relevant Criteria

In round 1, a total of 80 criteria were assessed, with consensus reached on 49 (61%). In addition, 8 new criteria emerged from suggestions by the panel and were incorporated in round 2: *critical life events (eg, loss of a child or another loved one)* (criterion 81), *emotional indifference* (criterion 82), *anhedonia* (criterion 83), *afraid of being alone* (criterion 84), *resurfacing of "old wounds" (eg, long-standing conflicts with children)* (criterion 85), *feeling of accomplishment throughout life* (criterion 86), *closeness with family and significant others* (criterion 87), and *subjective health* (criterion 88). As a result, 39 criteria (n=31, 79% without consensus and n=8, 21% new proposals) were reassessed in round 2, with consensus achieved on 12 (31%). In round 3, a total of 27 criteria were revisited, with consensus reached on 9 (33%; Table 2).

The consensus rates for the 88 criteria assessed across the 3 e-Delphi rounds are presented in Multimedia Appendix 4. In total, 20% (18/88) of the criteria failed to reach the ≥70% consensus threshold: *easily annoyed* (criterion 20); *short-tempered* (criterion 21); *staying home instead of going out and doing new things* (criterion 26); *psychomotor agitation* (criterion 29); *significant unintentional weight gain (more than 5% in a month)* (criterion 33); *restless sleep* (criterion 35); *full of energy* (criterion 42); *feels as good as other people* (criterion 51); *pessimism* (criterion 52); *thinks most people are better off than him/her* (criterion 56); *finding life exciting, wonderful, and enjoyable* (criterion 58); *getting bored often* (criterion 62); *feeling fearful* (criterion 64); *feeling that people are unfriendly or dislike him/her* (criterion 66); *memory problems* (criterion 69); *the mind is as clear as it used to be* (criterion 70); *having a sense of direction and purpose in life* (criterion 79); and *afraid of being alone* (criterion 84)—the only newly suggested criterion without consensus. Of the criteria on which consensus was achieved, all were considered *relevant* for screening for depression in individuals aged ≥95 years except for *hard to get started on new projects* (criterion 28), *fear of dying* (criterion 75), and *sufficient financial resources* (criterion 77), which were considered *not relevant*. Notably, several *relevant* criteria reached a ≥90% consensus: *despondency, gloom, and despair* (criterion 5; 25/25, 100%); *depressed* (criterion 6; 27/27, 100%); *lack of reactivity to pleasant events or circumstances* (criterion 25; 28/28, 100%); *suicidal ideation* (criterion 45; 28/28, 100%); *suicide attempt(s)* (criterion 46; 28/28, 100%); *ruminations* (criterion 14; 27/28, 96%); *recurrent thoughts of death or suicide* (criterion 43; 27/28, 96%); *feelings of worthlessness* (criterion 47; 25/26, 96%); *critical life events* (criterion 81; 20/21, 95%); *anhedonia* (criterion 83; 20/21, 95%); *loss of interest in activities* (criterion 23; 26/28, 93%); *loss of pleasure in activities* (criterion 24; 26/28, 93%); and *sadness* (criterion 4; 24/26, 92%).

Figure S2 in Multimedia Appendix 3 provides an overview of the consensus rates for the potentially relevant criteria for depression screening.

Signs and Symptoms of Depression: Duration, Number, Frequency, and Severity

In round 1, experts (n=28) were asked about the importance of considering the duration of depressive signs or symptoms in addition to their presence. It is important to note that, for the purpose of this assessment, they were instructed not to consider *suicidal thoughts or ideation* as these require immediate attention and, thus, were deliberately excluded from the duration criteria. All participants (28/28, 100%) agreed on its importance. When specifically asked about the minimum duration for considering these signs or symptoms as clinically significant, 25 experts provided their insights—21 (84%) suggested a range from *a few days* (1 to 3) to *4 weeks*, with *2 weeks* being the most cited (n=9, 36%).

In addition, 7% (2/28) of the experts suggested differentiating the duration of signs or symptoms between recurrent episodes in chronic depression and new-onset cases. For this reason, a question addressing this specificity was introduced in round 2 (n=21), with 62% (13/21) supporting the distinction. For recurrent episodes, most experts favored a minimum duration

XSL•FO
RenderX

of *2 weeks* (9/13, 69%), whereas for a new onset, opinions varied, with "4 weeks" (5/13, 38%) being the most cited.

In round 2 (n=21), 62% (13/21) emphasized the importance of considering the number of signs or symptoms during screening. Of the 12 experts who specified a threshold, 7 (58%) favored a minimum of 3 symptoms.

Regarding sign or symptom frequency, all experts who provided an opinion in round 2 (20/20, 100%) considered it relevant. Of these 20 experts, 19 (95%) specified a particular frequency, of whom 9 (47%) suggested a frequency of *several times per day*, 3 (16%) suggested a frequency of *nearly daily*, and 6 (32%) suggested a frequency of *several times per week*.

For severity, 90% (18/20) considered it an important factor. Moreover, on a severity scale from 0 to 10 (with 10 being the most severe), the average threshold for clinical concern was 4.9 (SD 1.4), with a median of 5.0 (IQR 2.0).

Further details regarding the results described previously can be found in Multimedia Appendix 5.

Steering Committee's Decision on Dimensions and Criteria Lacking Consensus After 3 e-Delphi Rounds

The steering committee conducted 2 additional rounds of discussion to decide on the relevance of the 4 dimensions and 18 criteria that did not reach a consensus of ≥70% among the panelists during the 3 e-Delphi rounds.

In the first round of discussion, committee members reviewed the panelists' comments and agreement levels for each dimension and criterion, analyzing potential reasons for the lack of consensus and considering the clinical implications. Each member shared their perspectives drawing on their clinical expertise. In the second round, the committee revisited the dimensions and criteria after reflecting on the insights from the first round, aiming to resolve any remaining uncertainties and reach a collective decision. Consensus was prioritized, with majority voting used when necessary.

A summary of the committee's final determinations is provided in Multimedia Appendix 6, where the rationale behind the inclusion or exclusion of each dimension and criterion is detailed.

Table 4 presents an overview of the consensus levels reached in each e-Delphi round and the final decisions made by the steering committee.

XSL•FO
RenderX

Table 4. Steering committee's decision on the dimensions and criteria lacking consensus after 3 e-Delphi rounds.

	Round 1 (n=28), n (%)	Round 2 (n=21), n (%)	Round 3 (n=20), n (%)	Steering committee final decision
Dimensions				
Memory problems				Not relevant
Not relevant	15 (56)[a]	11 (52)[b]	9 (45)[c]	
Relevant	12 (44)[a]	10 (48)[b]	11 (55)[c]	
Mood-congruent delusions				Relevant
Not relevant	12 (44)[a]	9 (43)[b]	9 (45)[c]	
Relevant	15 (56)[a]	12 (57)[b]	11 (55)[c]	
Reduced adaptation capacity				Relevant
Not relevant	12 (43)[d]	7 (33)[b]	8 (40)[c]	
Relevant	16 (57)[d]	14 (67)[b]	12 (60)[c]	
Language or speech (eg, poor speech; added in round 2)				Not relevant
Not relevant	—[e]	8 (38)[b]	12 (60)[c]	
Relevant	—	13 (62)[b]	8 (40)[c]	
Criteria				
Easily annoyed				Not relevant
Not relevant	11 (39)[d]	12 (57)[b]	11 (55)[c]	
Relevant	17 (61)[d]	9 (43)[b]	9 (45)[c]	
Short-tempered				Relevant
Not relevant	11 (41)[a]	11 (58)[f]	10 (50)[c]	
Relevant	16 (59)[a]	8 (42)[f]	10 (50)[c]	
Staying home instead of going out and doing new things				Not relevant
Not relevant	11 (39)[d]	9 (43)[b]	12 (60)[c]	
Relevant	17 (61)[d]	12 (57)[b]	8 (40)[c]	
Psychomotor agitation				Not relevant
Not relevant	11 (41)[a]	8 (38)[b]	9 (45)[c]	
Relevant	16 (59)[a]	13 (62)[b]	11 (55)[c]	
Significant unintentional weight gain (more than 5% in a month)				Not relevant
Not relevant	18 (67)[a]	14 (67)[b]	11 (55)[c]	
Relevant	9 (33)[a]	7 (33)[b]	9 (45)[c]	
Restless sleep				Relevant
Not relevant	10 (38)[g]	8 (40)[c]	8 (40)[c]	
Relevant	16 (62)[g]	12 (60)[c]	12 (60)[c]	
Full of energy				Not relevant
Not relevant	18 (67)[a]	12 (57)[b]	12 (60)[c]	
Relevant	9 (33)[a]	9 (43)[b]	8 (40)[c]	
Feels as good as other people				Not relevant
Not relevant	14 (52)[a]	13 (65)[c]	8 (40)[c]	

XSL•FO
RenderX

	Round 1 (n=28), n (%)	Round 2 (n=21), n (%)	Round 3 (n=20), n (%)	Steering committee final decision
Relevant	13 (48)[a]	7 (35)[c]	12 (60)[c]	
Pessimism				Relevant
Not relevant	9 (32)[d]	9 (45)[c]	8 (40)[c]	
Relevant	19 (68)[d]	11 (55)[c]	12 (60)[c]	
Thinks most people are better off than him or her				Not relevant
Not relevant	10 (37)[a]	9 (43)[b]	7 (35)[c]	
Relevant	17 (63)[a]	12 (57)[b]	13 (65)[c]	
Finding life exciting, wonderful, and enjoyable				Not relevant
Not relevant	10 (37)[a]	9 (45)[c]	9 (45)[c]	
Relevant	17 (63)[a]	11 (55)[c]	11 (55)[c]	
Getting bored often				Not relevant
Not relevant	11 (41)[a]	12 (57)[b]	7 (37)[f]	
Relevant	16 (59)[a]	9 (43)[b]	12 (63)[f]	
Feeling fearful				Relevant
Not relevant	9 (33)[a]	11 (52)[b]	10 (50)[c]	
Relevant	18 (67)[a]	10 (48)[b]	10 (50)[c]	
Feeling that people are unfriendly or dislike him or her				Not relevant
Not relevant	15 (56)[a]	13 (62)[b]	7 (35)[c]	
Relevant	12 (44)[a]	8 (38)[b]	13 (65)[c]	
Memory problems				Not relevant
Not relevant	14 (50)[d]	12 (57)[b]	9 (45)[c]	
Relevant	14 (50)[d]	9 (43)[b]	11 (55)[c]	
The mind is as clear as it used to be				Not relevant
Not relevant	14 (52)[a]	11 (52)[b]	10 (50)[c]	
Relevant	13 (48)[a]	10 (48)[b]	10 (50)[c]	
Having a sense of direction and purpose in life				Relevant
Not relevant	11 (41)[a]	7 (33)[b]	9 (45)[c]	
Relevant	16 (59)[a]	14 (67)[b]	11 (55)[c]	
Afraid of being alone (added in round 2)				Relevant
Not relevant	—	9 (43)[b]	8 (40)[c]	
Relevant	—	12 (57)[b]	12 (60)[c]	

[a]n=27.
[b]n=21.
[c]n=20.
[d]n=28.
[e]Not applicable.
[f]n=19.
[g]n=26.

XSL•FO
RenderX

Types of Assessment for Potential Depression Criteria: Self-Assessment, Hetero-Assessment, or Both?

In addition to assessing the relevance of each criterion, the panel provided opinions on the preferred assessment method (self-assessment, hetero-assessment, or both). Rather than seeking strict consensus, the goal was to identify general trends. Using a 50% threshold as an example, 35% (28/80) of the criteria were deemed suitable for *self-assessment* only, 4% (3/80) of the criteria were deemed suitable for *hetero-assessment* only, and 40% (32/80) of the criteria were deemed suitable for both methods. Detailed information is available in Multimedia Appendix 7, whereas Figure S3 in Multimedia Appendix 3 provides an overview of the consensus rates for these assessment methods.

Discussion

Principal Findings

This modified e-Delphi study aimed to identify and achieve consensus on the most representative signs and symptoms of depression in very old age, focusing on near-centenarians and centenarians.

Our results revealed that most of the items proposed—86% (24/28) of the dimensions and 80% (70/88) of the criteria—which had been assembled based on a literature review, were considered *relevant* for depression screening in very old age. This highlights the importance of a multidimensional approach and acknowledges the complexity of accurately diagnosing depression in this population.

Most Consensual Potentially Relevant Depressive Signs and Symptoms

Focusing on items with the highest consensus rates (≥90%), some overlap between dimensions and criteria was observed, which was expected as dimensions grouped related criteria.

For example, the *depressed mood* dimension was widely considered *relevant*, aligning with criteria such as *depressed*; *sadness*; and *despondency, gloom, and despair*, which reflect the emotional core of depressive disorders [6,54]. Similarly, *loss of interest in activities*, *loss of pleasure in activities*, *lack of reactivity to pleasant events or circumstances*, and *anhedonia* also achieved broad consensus as critical indicators of depression [6,54,62-64]. Collectively, these signs and symptoms highlight the significant impact of depression on a person's capacity for joy and engagement with life, underscoring their relevance in depression screening.

Criteria such as *ruminations* and *feelings of worthlessness* equally achieved a strong consensus due to their role in the negative cognitive and emotional patterns of depression. Previous research has shown that *ruminations* exacerbate psychopathology, including depression, by intensifying and extending negative mood states, thus interfering with problem-solving and perpetuating physiological stress responses [65]. Similarly, *feelings of worthlessness* are strongly linked to self-blame, lack of hope, and an increased suicide risk even after depressive episodes remit [66,67]. Building on this, the *lack ofhope* dimension achieved unanimous consensus (21/21, 100%) as a relevant element, reflecting its importance as a significant predictor of psychological well-being [68]. Within this broader dimension, the *hopelessness* and *hopeful about the future* criteria also showed high consensus (24/27, 89% and 16/21, 76%, respectively), emphasizing their important role in depression screening in individuals of very old age whether as a negative indicator or a positive resource.

Furthermore, *previous episodes of depression or diagnosed depression* were also identified as relevant, highlighting the importance of considering an individual's psychiatric history as past experiences with depression increase the risk of recurrence [69]. In addition, *critical life events* perceived as negative were recognized for their significant association with higher levels of depressive symptoms in a recent meta-analysis [70].

Finally, the strong agreement on severe criteria such as *recurrent thoughts of death or suicide*, *suicidal ideation*, and past *suicide attempt(s)* underscores the critical need to address these factors due to their immediate risk to individuals' safety [26].

Most Consensual Potentially Relevant Depressive Signs and Symptoms in the Context of Old Age

Having established the most consensual depressive features, it is essential to explore how these characteristics manifest specifically in older adults as unique age-related factors may shape their expression and impact.

As individuals age, they may face different challenges, some of which are particularly relevant to LLD. These include agism; loss of income; the death of loved ones; and the risk of isolation, functional limitations, multimorbidity, cognitive decline, and institutionalization [71]. Such multidimensional challenges can impact well-being and mental health, contributing to reduced aging satisfaction and increased loneliness, including near-centenarians and centenarians [71,72]. While the cumulative effect of these difficulties can increase the risk of LLD, it is important to note that experiencing them does not necessarily lead to the development of a depressive state [72].

The dimensions and criteria that achieved high consensus, as previously discussed, clearly represent central elements in depression screening across all age groups, aligning with current diagnostic standards [6,54]. Our findings confirm their importance for accurate diagnosis in individuals of very old age, underscoring their transversality across generations and relevance even in the final stages of life. While depression in very old age may have unique aspects, some of the core signs and symptoms of the disorder remain consistent and should not be overlooked, including in near-centenarians and centenarians. However, these signs and symptoms may carry additional significance due to the unique psychosocial and physical challenges faced by these individuals [72].

For instance, *loss of interest in activities*, *loss of pleasure in activities*, *lack of reactivity to pleasant events or circumstances*, and *anhedonia* may highlight a deeper disengagement from life that can be difficult to detect in older adults. These symptoms are often masked by typical aging challenges such as physical decline, social isolation, or even cognitive issues, complexifying their identification. In addition, *feelings of worthlessness* and

XSL•FO
RenderX

the negative impact of *critical life events* may be particularly relevant in late-life contexts, where significant life changes and losses are common [67,73]. As highlighted in a meta-synthesis [74], older adults often experience declines in physical and cognitive abilities, leading to feelings of disconnection and a perception of being unneeded by their communities. This sense of being unneeded can foster feelings of isolation and insignificance, thereby triggering or exacerbating depressive symptoms. Moreover, societal changes and loss of social roles further deepen these feelings of worthlessness, underscoring the importance of addressing these issues in older populations [74].

While the high consensus on many classic criteria suggests their continued relevance across the life cycle, our findings also point to unique aspects of depression in very old age that may not be fully captured by current approaches. For example, the unanimous agreement on the relevance of *lack ofhope* highlights its particular significance in this age group. Unlike other factors typically associated with normal aging processes, low levels of hope have been identified as a significant risk factor in older adults, reflecting a deeper sense of despair with potentially serious consequences. A systematic review [75] underscored that lack of hope is closely linked to negative outcomes in older populations, including an increased risk of depression and poorer overall health outcomes.

Moreover, the strong agreement on severe symptoms such as *recurrent thoughts of death or suicide*, *suicidal ideation*, and past *suicide attempt(s)* underscores the importance of also addressing these factors in very old age, where such topics often remain taboo [76]. Reluctance to discuss or acknowledge these symptoms may lead to underdiagnosis and inadequate treatment, emphasizing the need for sensitive and proactive approaches in this age group. Recent qualitative research [76] has revealed that suicidal thoughts in later life are often subtly expressed and normalized in everyday conversations, making them difficult to detect. However, the extent to which this applies to centenarians is unclear. Previous research suggests that only a minority of centenarians express a wish to die [77,78], making explicit suicidal thoughts likely rare and not always linked to depressive states. Nonetheless, feelings of life being *completed and no longer worth living* reflect a deeper disconnection from life, characterized by aching loneliness, feelings of not mattering, and a physical and mental tiredness [79]. Creating safe spaces for informal—yet meaningful—discussions is essential, alongside strategies for timely recognition and management of these symptoms [76].

Potentially Not Relevant Depressive Signs and Symptoms in the Context of Old Age

The *increased appetite* dimension may have been considered *not relevant* as it can result from situational factors such as attending a day center or moving to a nursing home, where tailored meals and social dining often enhance the willingness to eat, contributing to an increased appetite. Similarly, the *significant unintentional weight gain* criterion was likely deemed *not relevant* for related reasons as previous research has found no association between weight gain and incident depressive symptoms in older adults [80].

Certain cognitive dimensions such as *memory problems* and *language/speech (eg, poor speech)* were also classified as *not relevant*, likely due to their overlap with nondepressive phenomena commonly observed in older adults. This highlights the need for tailored screening protocols that can effectively differentiate between depression and cognitive impairment, conditions that often coexist in individuals of very old age and require distinct therapeutic strategies [81,82]. The criterion *the mind is as clear as it used to be* was likely deemed *not relevant* for the same reasons.

Criteria such as *hard to get started on new projects*; *staying home instead of going out and doing new things*; *finding life exciting, wonderful, and enjoyable*; *easily annoyed*; and *getting bored often* may be more closely related to motivational factors, social involvement, social support, or underlying health conditions rather than directly indicating a depressive state [83,84]. Similarly, criteria such as *thinks most people are better off than him/her*, *feeling that people are unfriendly or dislike him/her*, and *feels as good as other people* may reflect personality traits, which are relatively stable over the life span and may not necessarily reflect a change in mental health status [85].

Symptoms such as *psychomotor agitation* illustrate the challenge in differentiating between depression-related symptoms and those arising from other geriatric conditions such as dementia or delirium [86], suggesting that certain depressive features may be less applicable or not sufficient for individuals of very old age [87]. In addition, the *full of energy* criterion, typically seen as a positive indicator, was not deemed relevant, possibly because it is rare for a near-centenarian or centenarian without depression to report feeling full of energy. This likely reflects their overall physical condition rather than a depressive state, and the absence of this criterion may often be confounded with physical exhaustion linked to frailty [88].

The *fear of dying* criterion was also classified as *not relevant*. Although it might be assumed that individuals in the final stage of life would fear death, previous research suggests that death anxiety is generally low among older adults, who are more often concerned with the dying process rather than death itself [89]. Similarly, *sufficient financial resources* was deemed *not relevant*. While financial difficulties can impact mental health, their link to psychological distress in older adults appears to be more influenced by underlying factors such as mastery and self-esteem rather than acting as a direct indicator of depression [90].

The Specificity of Near-Centenarians and Centenarians

Research on near-centenarians and centenarians has predominantly focused on physical, cognitive, and social health, with less attention paid to psychiatric aspects [13]. Studies on depression in this population often rely on overall assessment scores from standardized instruments rather than detailed analysis of individual items [72,87,91-94]. This gap, alongside the variety of existing tools for screening for depression in very old age, also motivated this study.

Notably, one study conducted a detailed item-level analysis using a 14-item version of the Geriatric Depression Scale in

XSL•FO
RenderX

Portuguese centenarians [88] and found that 51.4% reported *feelings of worthlessness*, supporting the relevance of this specific criterion in our e-Delphi study. In addition, only 12.3% reported having more *memory problems*, 43.3% expressed a preference for *staying home instead of going out and doing new things*, 39.8% reported *getting bored often*, and 33.3% agreed with the statement *thinks most people are better off than they are* [88]. Globally, these findings align with our results, where these criteria were considered not relevant in this age group.

While these findings suggest reduced presence of symptoms established for younger populations, it remains unclear whether depression in near-centenarians and centenarians significantly differs from that in other older individuals, such as octogenarians; the key features of depression may not vary meaningfully. Therefore, it is worth questioning whether the development of a tool specifically designed for very old age, including the near-centenarian and centenarian population, is truly necessary. Nevertheless, psychological resources such as resilience appear to be particularly strong in near-centenarians and centenarians [77,95]. These protective characteristics may lead traditional depression screening tools to overlook subtle signs of depression that are masked by these strengths. A tool specifically designed for this age group could better capture the subtle differences and specific needs, enhancing the accuracy and effectiveness of depression screening.

Duration, Number, Frequency, and Severity of Depressive Signs and Symptoms and Type of Assessment

Our findings emphasize the importance of considering the duration, number, frequency, and severity of depressive signs and symptoms in near-centenarians and centenarians. Notably, experts differentiated between chronic and new-onset depression, advocating for a nuanced approach that may significantly impact clinical assessments and interventions. This aligns with existing literature documenting distinct clinical profiles for persistent and nonchronic depressive disorders in adults aged 18 to 79 years [96,97], with persistent cases often involving earlier onset, more severe symptoms, and greater treatment resistance [96]. While these findings are based on younger individuals, they reinforce the necessity of a personalized life span approach that includes rigorous screening and comprehensive clinical strategies for the oldest age groups.

Furthermore, experts in our study highlighted the potential value of combining self- and hetero-assessment techniques to screen for depressive signs and symptoms in very old age. However, this approach may require further consideration. Cognitive limitations at this age may affect the reliability of self-reports, complicating the integration with caregiver observations. Developing strategies to reconcile these differing perspectives—such as guidelines for interpreting discrepancies—could ensure that assessments are accurate and sensitive to the unique challenges faced by individuals of very old age. Flexible, combined approaches remain relevant as research suggests that they enhance diagnostic accuracy and patient engagement [98].

Further Considerations

Given the findings of our study, a more comprehensive approach to depression screening in near-centenarians and centenarians may be advantageous. The high consensus among experts regarding certain depressive features highlights their relevance and significance for this particular age group. Importantly, the diagnosis of LLD in adults of very old age can be particularly complex due to overlapping symptoms with other mental or somatic conditions, as previously outlined. This overlap necessitates a differential diagnostic approach that integrates a broader perspective. This emphasizes the need for screening tools that are both comprehensive and sensitive to the unique experiences and needs of individuals of very old age. Implementing such tailored screening methods could contribute to improved mental health outcomes within this population.

Our work extends existing literature by providing a comprehensive overview of a wider range of depressive features, including considerations of their duration, number, frequency, and severity—factors that are less considered in standard depression screenings. This highlights the complexity of diagnosing depression in adults of very old age and confirms the necessity of incorporating these specific aspects into the screening process.

While the e-Delphi method was valuable for identifying key elements in detecting depressive signs and symptoms in those of very old age, a separate validation study is necessary to confirm whether the extended screening process would be more effective compared to current tools. Nonetheless, this study hopefully represents a meaningful step forward in improving the detection of depressive signs and symptoms in this population.

Limitations

We acknowledge several limitations. First, despite the international and professional diversity of the expert panel, it may not have captured the full spectrum of clinical practices or cultural perceptions related to aging and depression. The relative overrepresentation of nurses among the panelists in rounds 1 and 2 may have introduced a bias toward nursing-specific priorities, such as a stronger focus on the practical applicability of screening criteria in daily care. However, this also reflects their pivotal role in direct care for individuals of very old age, making their insights particularly relevant for developing pragmatic screening criteria. Given this, a *groupthink* effect in which the dominant profession influenced the consensus cannot be excluded. In addition, our panel lacked experts from Asian countries despite notable recruitment efforts even though these regions have some of the highest proportions of centenarians worldwide. This absence may have introduced a bias toward Western perspectives. At the same time, the concept of depression in very old age may be less clearly defined in some Asian cultures due to cultural background. Second, the selection of a 70% consensus threshold may have influenced the interpretation of the results as it reflects a compromise between inclusivity and stringency, subtly shaping the scope of the consensus achieved. Third, the structured nature of the e-Delphi process may have limited the exploration of nuanced symptoms critical to understanding depression in individuals of very old

XSL•FO
RenderX

age, a population whose cognitive decline may complicate symptom presentation. Fourth, some overlap between dimensions and criteria was observed, which may be perceived as a limitation to some extent. However, the steering committee was aware of this overlap and intentionally decided to retain it to confirm certain trends in consensus on specific items.

Despite these challenges, this study's insights have the potential to stimulate discussions on improving mental health strategies for older populations, especially those of very old age. The diverse and interdisciplinary composition of the panel further contributes to the generalizability of the findings, offering valuable insights to health care professionals and researchers worldwide.

Future research should aim to develop and validate a screening tool specifically designed for depressive signs and symptoms in very old age, including near-centenarians and centenarians. A validation study comparing the extended method with established tools or a *gold standard* diagnostic reference is crucial to determining its efficacy. In addition, it is important to assess the tool's effectiveness across different cultural contexts to ensure its broad applicability. The incorporation of Patient-Reported Experience Measures in future developments would also be beneficial in capturing the lived experiences and perceptions of this age group.

Conclusions

This expert consensus could significantly contribute to enhancing early detection and intervention for depression in individuals of very old age, notably by informing the creation of a new screening tool. The findings have the potential to improve diagnostic accuracy and, consequently, allow for the personalization of treatment plans tailored to the unique needs of individuals of very old age.

In conclusion, this study used expert consensus to establish a foundational step toward enhancing the detection and treatment of depression among individuals of very old age, including near-centenarians and centenarians. This study emphasizes the need for cautious, multidimensional, and multidisciplinary approaches to depression screening in this age group, advocating for the continuous refinement of screening tools to address the complex nature of very old age effectively.

Acknowledgments

This work was supported by the Swiss National Science Foundation through the Sinergia grant CRSII5_186239/1 (SWISS100). The authors gratefully acknowledge the contributions of all the experts who participated in this study. They recognize those who agreed to be publicly acknowledged: Alcina Matos Queirós, Department of Health and Social Welfare, Canton of Vaud, Switzerland; Carina Santos, Riviera-Chablais Hospital, Switzerland; Christophe Nakamura, Lausanne University Hospital, Switzerland; Eivind Aakhus, the Norwegian National Centre for Ageing and Health, Vestfold Hospital Trust, Norway; Filipa Pereira, School of Health Sciences, University of Applied Sciences and Arts Western Switzerland Valais-Wallis, Switzerland; Jean Roche, Lille University Hospital, France; Kim Uittenhove, Swiss Federal Institute of Technology in Lausanne, Switzerland; Nancy A Pachana, the University of Queensland, Australia; Oscar Ribeiro, University of Aveiro, Portugal; Raquel Neves, Almoosa Rehabilitation Hospital, Kingdom of Saudi Arabia; Sofia Fernandes, School of Health Sciences, University of Applied Sciences and Arts Western Switzerland Valais-Wallis, Switzerland; Sylvain Nguyen, Riviera-Chablais Hospital, Switzerland; Sylvie Bonin-Guillaume, Aix-Marseille University, University Hospital of Marseille, France; and Tiago Ferreira, Lausanne University Hospital, Switzerland. The authors also extend their sincere appreciation to the experts who preferred to remain anonymous. Their equally valuable contributions were central in shaping this study's outcomes.

Authors' Contributions

CGdR, AvG, and HV contributed to the conceptualization of the study. AvG, PV, and HV provided methodological input. AvG, PV, DSJ, OR, and HV contributed scientific input. CGdR was responsible for data collection, data analysis, and writing the original draft. CGdR, AvG, PV, DSJ, OR, and HV reviewed and edited the manuscript. AvG and HV supervised the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Selected instruments and their respective items.
[DOCX File , 27 KB - aging_v8i1e64352_app1.docx]

Multimedia Appendix 2
Decision of the ethics committee.
[DOCX File , 180 KB - aging_v8i1e64352_app2.docx]

Multimedia Appendix 3
Graphical overview of the consensus rates (dimensions, criteria, and type of assessment).

XSL•FO
RenderX

[DOCX File , 774 KB - aging_v8i1e64352_app3.docx]

Multimedia Appendix 4
Consensus level on potential criteria for depression screening.
[DOCX File , 42 KB - aging_v8i1e64352_app4.docx]

Multimedia Appendix 5
Aspects of signs and symptoms of depression (duration, number, frequency, and severity).
[DOCX File , 20 KB - aging_v8i1e64352_app5.docx]

Multimedia Appendix 6
The steering committee's final determinations.
[DOCX File , 17 KB - aging_v8i1e64352_app6.docx]

Multimedia Appendix 7
Findings on the preferred assessment methods for depression screening criteria.
[DOCX File , 25 KB - aging_v8i1e64352_app7.docx]

References

1. Schaeffer K. U.S. centenarian population is projected to quadruple over the next 30 years. Pew Research Center. URL: https://tinyurl.com/4z8tuu3d [accessed 2024-04-08]
2. World population ageing. United Nations (UN). URL: https://tinyurl.com/33chm345 [accessed 2024-04-29]
3. The world population prospects: 2015 revision. United Nations (UN). URL: https://www.un.org/en/development/desa/publications/world-population-prospects-2015-revision.html [accessed 2024-04-29]
4. Teixeira L, Araújo L, Jopp D, Ribeiro O. Centenarians in Europe. Maturitas 2017 Oct;104:90-95. [doi: 10.1016/j.maturitas.2017.08.005] [Medline: 28923181]
5. Living to 100 years and beyond. Federal Statistical Office. URL: https://tinyurl.com/bdzcvhzd [accessed 2024-04-08]
6. ICD-11 for mortality and morbidity statistics. World Health Organization. URL: https://icd.who.int/browse/2024-01/mms/en [accessed 2024-04-29]
7. Rapport mondial sur le vieillissement et la santé. World Health Organization. URL: https://tinyurl.com/mrwmk5bs [accessed 2024-04-29]
8. Jopp DS, Boerner K, Rott C. Health and disease at age 100: findings from the second Heidelberg Centenarian studies. Dtsch Arztebl Int 2016 Mar 25;113(12):203-210 [FREE Full text] [doi: 10.3238/arztebl.2016.0203] [Medline: 27118718]
9. Devita M, de Salvo R, Ravelli A, de Rui M, Coin A, Sergi G, et al. Recognizing depression in the elderly: practical guidance and challenges for clinical management. Neuropsychiatr Dis Treat 2022;18:2867-2880 [FREE Full text] [doi: 10.2147/NDT.S347356] [Medline: 36514493]
10. Lleshi V, Bizzozzero T. La dépression du sujet âgé. Rev Med Suisse 2009;5:1785-1789. [doi: 10.53738/revmed.2009.5.216.1785]
11. Cappeliez P. La dépression. In: Voyer P, editor. Soins infirmiers aux aînés en perte d'autonomie 3e. Montréal, Québec: Éditions du Renouveau Pédagogique Inc; 2021:185-198.
12. Gallarda T, Lôo H. Dépression et personnes âgées. Encephale 2009 Jun;35(3):269-280. [doi: 10.1016/j.encep.2008.10.013] [Medline: 19540414]
13. Cheng A, Leung Y, Harrison F, Brodaty H. The prevalence and predictors of anxiety and depression in near-centenarians and centenarians: a systematic review. Int Psychogeriatr 2019 Nov 29;31(11):1539-1558 [FREE Full text] [doi: 10.1017/S1041610219000802] [Medline: 31354113]
14. Eshun S, Gurung RA. Culture and Mental Health: Sociocultural Influences, Theory, and Practice. Halifax, NS: Wiley-Blackwell; 2009.
15. Bergdahl E, Gustavsson JM, Kallin K, von Heideken Wågert P, Lundman B, Bucht G, et al. Depression among the oldest old: the Umeå 85+ study. Int Psychogeriatr 2005 Dec;17(4):557-575 [FREE Full text] [doi: 10.1017/S1041610205002267] [Medline: 16185377]
16. Clément JP. La dépression à l'âge avancé. In: Giannakopoulos P, Gaillard M, editors. Abrégé de Psychiatrie de l'Âge Avancé. Chêne-Bourg, Switzerland: Médecine & Hygiène; 2010:171-183.
17. Hirdes JP, Ikegami N, Jónsson PV, Topinková E, Maxwell CJ, Yamauchi K. Cross-national comparisons of antidepressant use among institutionalized older persons based on the minimum data set (MDS). Can J Aging 2010 Nov 29;19(S2):18-37. [doi: 10.1017/S0714980800013878]
18. Hottin P, Trudel JF. Dépression et toubles affectifs. In: Arcand M, Hébert R, editors. Précis Pratique de Gériatrie. Montreal, QC: Edisem inc; 2007:305-322.

XSL•FO
RenderX

19. Mental health of older adults. World Health Organization (WHO). 2017. URL: https://www.who.int/news-room/fact-sheets/detail/mental-health-of-older-adults [accessed 2024-04-29]
20. Wei J, Hou R, Zhang X, Xu H, Xie L, Chandrasekar EK, et al. The association of late-life depression with all-cause and cardiovascular mortality among community-dwelling older adults: systematic review and meta-analysis. Br J Psychiatry 2019 Aug;215(2):449-455. [doi: 10.1192/bjp.2019.74] [Medline: 30968781]
21. Cherbuin N, Kim S, Anstey KJ. Dementia risk estimates associated with measures of depression: a systematic review and meta-analysis. BMJ Open 2015 Dec 21;5(12):e008853 [FREE Full text] [doi: 10.1136/bmjopen-2015-008853] [Medline: 26692556]
22. Diniz BS, Butters MA, Albert SM, Dew MA, Reynolds 3rd CF. Late-life depression and risk of vascular dementia and Alzheimer's disease: systematic review and meta-analysis of community-based cohort studies. Br J Psychiatry 2013 May;202(5):329-335 [FREE Full text] [doi: 10.1192/bjp.bp.112.118307] [Medline: 23637108]
23. John A, Patel U, Rusted J, Richards M, Gaysina D. Affective problems and decline in cognitive state in older adults: a systematic review and meta-analysis. Psychol Med 2019 Feb;49(3):353-365 [FREE Full text] [doi: 10.1017/S0033291718001137] [Medline: 29792244]
24. Santabarbara J, Sevil-Perez A, Olaya B, Gracia-Garcia P, Lopez-Anton R. [Clinically relevant late-life depression as risk factor of dementia: a systematic review and meta-analysis of prospective cohort studies]. Rev Neurol 2019 Jun 16;68(12):493-502 [FREE Full text] [doi: 10.33588/rn.6812.2018398] [Medline: 31173329]
25. Thomas A. Depression in older people. In: Dening T, Thomas A, editors. Oxford Textbook of Old Age Psychiatry. 2nd edition. New York, NY: Oxford University Press; 2013:545-569.
26. Gramaglia C, Calati R, Zeppegno P. Rational suicide in late life: a systematic review of the literature. Medicina (Kaunas) 2019 Sep 29;55(10):52 [FREE Full text] [doi: 10.3390/medicina55100656] [Medline: 31569542]
27. Bian Z, Xu R, Shang B, Lv F, Sun W, Li Q, et al. Associations between anxiety, depression, and personal mastery in community-dwelling older adults: a network-based analysis. BMC Psychiatry 2024 Mar 07;24(1):192 [FREE Full text] [doi: 10.1186/s12888-024-05644-z] [Medline: 38454373]
28. Balsamo M, Cataldi F, Carlucci L, Padulo C, Fairfield B. Assessment of late-life depression via self-report measures: a review. Clin Interv Aging 2018;13:2021-2044 [FREE Full text] [doi: 10.2147/CIA.S178943] [Medline: 30410319]
29. Luppa M, Sikorski C, Motzek T, Konnopka A, König HH, Riedel-Heller SG. Health service utilization and costs of depressive symptoms in late life - a systematic review. Curr Pharm Des 2012;18(36):5936-5957. [doi: 10.2174/138161212803523572] [Medline: 22681171]
30. British Geriatrics Society. Depression among older people living in care homes report. Royal College of Psychiatrists and British Geriatrics Society. URL: https://tinyurl.com/5e4tx7yj [accessed 2024-04-29]
31. Sudhir Kumar CT, George S, Augustine A. Treatment of late-life depression. J Geriatr Care Res 2020;7(1):9-16 [FREE Full text]
32. Fußer F, Karakaya T, Pantel J. Late-life depression. In: Carvalho AF, McIntyre RS, editors. Mental Disorders in Primary Care: A Guide to their Evaluation and Management. New York, NY: Oxford University Press; 2017:251-264.
33. Charoenngam N, Rittiphairoj T, Ponvilawan B, Prasongdee K. Thyroid dysfunction and risk of Parkinson's disease: a systematic review and meta-analysis. Front Endocrinol (Lausanne) 2022;13:863281 [FREE Full text] [doi: 10.3389/fendo.2022.863281] [Medline: 35600588]
34. Dagnino AP, Campos MM. Chronic pain in the elderly: mechanisms and perspectives. Front Hum Neurosci 2022;16:736688 [FREE Full text] [doi: 10.3389/fnhum.2022.736688] [Medline: 35308613]
35. Lee PH, Yeh TT, Yen HY, Hsu WL, Chiu VJ, Lee SC. Impacts of stroke and cognitive impairment on activities of daily living in the Taiwan longitudinal study on aging. Sci Rep 2021 Jun 09;11(1):12199 [FREE Full text] [doi: 10.1038/s41598-021-91838-4] [Medline: 34108582]
36. Liu M, Hou T, Nkimbeng M, Li Y, Taylor JL, Sun X, et al. Associations between symptoms of pain, insomnia and depression, and frailty in older adults: a cross-sectional analysis of a cohort study. Int J Nurs Stud 2021 May;117:103873 [FREE Full text] [doi: 10.1016/j.ijnurstu.2021.103873] [Medline: 33621722]
37. Pazan F, Wehling M. Polypharmacy in older adults: a narrative review of definitions, epidemiology and consequences. Eur Geriatr Med 2021 Jun 10;12(3):443-452 [FREE Full text] [doi: 10.1007/s41999-021-00479-3] [Medline: 33694123]
38. Wang EY, Meyer C, Graham GD, Whooley MA. Evaluating screening tests for depression in post-stroke older adults. J Geriatr Psychiatry Neurol 2018 May;31(3):129-135. [doi: 10.1177/0891988718778791] [Medline: 29793370]
39. Barrett D, Heale R. What are Delphi studies? Evid Based Nurs 2020 Jul;23(3):68-69. [doi: 10.1136/ebnurs-2020-103303] [Medline: 32430290]
40. Cole ZD, Donohoe HM, Stellefson ML. Internet-based Delphi research: case based discussion. Environ Manage 2013 Mar;51(3):511-523 [FREE Full text] [doi: 10.1007/s00267-012-0005-5] [Medline: 23288149]
41. Spranger J, Homberg A, Sonnberger M, Niederberger M. Reporting guidelines for Delphi techniques in health sciences: a methodological review. Z Evid Fortbild Qual Gesundhwes 2022 Aug;172:1-11 [FREE Full text] [doi: 10.1016/j.zefq.2022.04.025] [Medline: 35718726]

XSL•FO
RenderX

42. Jünger S, Payne SA, Brine J, Radbruch L, Brearley SG. Guidance on conducting and REporting DElphi studies (CREDES) in palliative care: recommendations based on a methodological systematic review. Palliat Med 2017 Sep 13;31(8):684-706. [doi: 10.1177/0269216317690685] [Medline: 28190381]
43. Gill FJ, Leslie GD, Grech C, Latour JM. Using a web-based survey tool to undertake a Delphi study: application for nurse education research. Nurse Educ Today 2013 Nov;33(11):1322-1328. [doi: 10.1016/j.nedt.2013.02.016] [Medline: 23510701]
44. Naisola-Ruiter V. The Delphi technique: a tutorial. Res Hosp Manag 2022 Jun 06;12(1):91-97. [doi: 10.1080/22243534.2022.2080942]
45. Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, et al. Development and validation of a geriatric depression screening scale: a preliminary report. J Psychiatr Res 1982;17(1):37-49. [doi: 10.1016/0022-3956(82)90033-4] [Medline: 7183759]
46. Radloff LS. The CES-D scale. Applied Psychol Manag 1977 Jun 01;1(3):385-401. [doi: 10.1177/014662167700100306]
47. Derogatis L. Brief Symptom Inventory (BSI): Administration, Scoring, and Procedures Manual. 4th edition. Minneapolis, MN: National Computer Systems; 1993.
48. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983 Jun;67(6):361-370. [doi: 10.1111/j.1600-0447.1983.tb09716.x] [Medline: 6880820]
49. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry 1979 Apr;134:382-389. [doi: 10.1192/bjp.134.4.382] [Medline: 444788]
50. Alexopoulos GS, Abrams RC, Young RC, Shamoian CA. Cornell scale for depression in Dementia. Biol Psychiatry 1988 Feb 01;23(3):271-284 [FREE Full text] [doi: 10.1016/0006-3223(88)90038-8] [Medline: 3337862]
51. Resnick B, van Haitsma K, Kolanowski A, Galik E, Boltz M, Ellis J, et al. Reliability and validity of the Cornell Scale for depression in dementia and invariance between black versus white residents in nursing homes. J Am Med Dir Assoc 2022 Jul;23(7):1236-41.e3 [FREE Full text] [doi: 10.1016/j.jamda.2021.11.016] [Medline: 34896057]
52. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. Neurology 1994 Dec;44(12):2308-2314. [doi: 10.1212/wnl.44.12.2308] [Medline: 7991117]
53. Kaufer DI, Cummings JL, Christine D, Bray T, Castellon S, Masterman D, et al. Assessing the impact of neuropsychiatric symptoms in Alzheimer's disease: the Neuropsychiatric Inventory Caregiver Distress Scale. J Am Geriatr Soc 1998 Feb;46(2):210-215. [doi: 10.1111/j.1532-5415.1998.tb02542.x] [Medline: 9475452]
54. American Psychiatric Association. Diagnostic and Statistical Manual Of Mental Disorders: DSM-5™, 5th edition. New York, NY: American Psychiatric Publishing; 2013.
55. Burrows AB, Morris JN, Simon SE, Hirdes JP, Phillips C. Development of a minimum data set-based depression rating scale for use in nursing homes. Age Ageing 2000 Mar;29(2):165-172. [doi: 10.1093/ageing/29.2.165] [Medline: 10791452]
56. Instruments overview. interRAI. URL: https://www.interrai.org/instruments/ [accessed 2024-04-29]
57. Monod S, Büla C, Hongler T, Castelli T, Clivaz-Luchez P. Le Resident Assessment Instrument-Home-Care (RAI-Domicile) : ce que le médecin de premier recours doit savoir [Article in French]. Rev Med Suisse 2011 Nov 09;7:2176-2183. [doi: 10.53738/REVMED.2011.7.316.2176]
58. IBM SPSS Statistics for Windows, Version 29.0. IBM Corp. URL: https://www.ibm.com/support/pages/downloading-ibm-spss-statistics-29 [accessed 2024-04-29]
59. Cascella M, Miceli L, Cutugno F, Di Lorenzo G, Morabito A, Oriente A, et al. A Delphi consensus approach for the management of chronic pain during and after the COVID-19 Era. Int J Environ Res Public Health 2021 Dec 19;18(24):22 [FREE Full text] [doi: 10.3390/ijerph182413372] [Medline: 34948983]
60. Lam K, Iqbal FM, Purkayastha S, Kinross JM. Investigating the ethical and data governance issues of artificial intelligence in surgery: protocol for a Delphi study. JMIR Res Protoc 2021 Feb 22;10(2):e26552 [FREE Full text] [doi: 10.2196/26552] [Medline: 33616543]
61. Vogel C, Zwolinsky S, Griffiths C, Hobbs M, Henderson E, Wilkins E. A Delphi study to build consensus on the definition and use of big data in obesity research. Int J Obes (Lond) 2019 Dec;43(12):2573-2586 [FREE Full text] [doi: 10.1038/s41366-018-0313-9] [Medline: 30655580]
62. Coryell W. Depressive disorders. MSD Manual Professional Version. URL: https://www.msdmanuals.com/professional/psychiatric-disorders/mood-disorders/depressive-disorders [accessed 2024-08-29]
63. de Fruyt J, Sabbe B, Demyttenaere K. Anhedonia in depressive disorder: a narrative review. Psychopathology 2020 Jul 15;53(5-6):274-281. [doi: 10.1159/000508773] [Medline: 32668436]
64. Serretti A. Anhedonia and depressive disorders. Clin Psychopharmacol Neurosci 2023 Aug 31;21(3):401-409 [FREE Full text] [doi: 10.9758/cpn.23.1086] [Medline: 37424409]
65. Watkins ER, Roberts H. Reflecting on rumination: consequences, causes, mechanisms and treatment of rumination. Behav Res Ther 2020 Apr;127:103573. [doi: 10.1016/j.brat.2020.103573] [Medline: 32087393]
66. Wakefield JC, Schmitz MF. Feelings of worthlessness during a single complicated major depressive episode predict postremission suicide attempt. Acta Psychiatr Scand 2016 Apr;133(4):257-265. [doi: 10.1111/acps.12521] [Medline: 26538107]

XSL•FO
RenderX

67. Harrison P, Lawrence AJ, Wang S, Liu S, Xie G, Yang X, et al. The psychopathology of worthlessness in depression. Front Psychiatry 2022;13:818542 [FREE Full text] [doi: 10.3389/fpsyt.2022.818542] [Medline: 35664464]
68. Long KN, Kim ES, Chen Y, Wilson MF, Worthington Jr EL, VanderWeele TJ. The role of Hope in subsequent health and well-being for older adults: an outcome-wide longitudinal approach. Glob Epidemiol 2020 Nov;2:100018. [doi: 10.1016/j.gloepi.2020.100018]
69. Saunders R, Cohen ZD, Ambler G, DeRubeis RJ, Wiles N, Kessler D, et al. A patient stratification approach to identifying the likelihood of continued chronic depression and relapse following treatment for depression. J Pers Med 2021 Dec 04;11(12):1295 [FREE Full text] [doi: 10.3390/jpm11121295] [Medline: 34945767]
70. Haehner P, Würtz F, Kritzler S, Kunna M, Luhmann M, Woud ML. The relationship between the perception of major life events and depression: a systematic scoping review and meta-analysis. J Affect Disord 2024 Mar 15;349:145-157. [doi: 10.1016/j.jad.2024.01.042] [Medline: 38190863]
71. Voyer P. Soins infirmiers aux aînés en perte d'autonomie, 3e édition. Montréal, QC: Éditions du Renouveau Pédagogique Inc; 2021.
72. Jopp DS, Boerner K, Cimarolli V, Hicks S, Mirpuri S, Paggi M, et al. Challenges experienced at age 100: findings from the Fordham centenarian study. J Aging Soc Policy 2016;28(3):187-207. [doi: 10.1080/08959420.2016.1163652] [Medline: 27010530]
73. Schladitz K, Löbner M, Stein J, Weyerer S, Werle J, Wagner M, et al. Grief and loss in old age: exploration of the association between grief and depression. J Affect Disord 2021 Mar 15;283:285-292. [doi: 10.1016/j.jad.2021.02.008] [Medline: 33578340]
74. Kitzmüller G, Clancy A, Vaismoradi M, Wegener C, Bondas T. "Trapped in an empty waiting room"-the existential human core of loneliness in old age: a meta-synthesis. Qual Health Res 2018 Jan;28(2):213-230. [doi: 10.1177/1049732317735079] [Medline: 29235943]
75. Hernandez SC, Overholser JC. A systematic review of interventions for hope/hopelessness in older adults. Clin Gerontol 2021;44(2):97-111. [doi: 10.1080/07317115.2019.1711281] [Medline: 31913808]
76. Hafford-Letchfield T, Hanna JR, Ellmers TJ, Rasmussen S, Cogan N, Gleeson H, et al. Talking really does matter: lay perspectives from older people on talking about suicide in later life. Front Psychol 2022;13:1009503 [FREE Full text] [doi: 10.3389/fpsyg.2022.1009503] [Medline: 36467190]
77. Jopp D, Rott C. Adaptation in very old age: exploring the role of resources, beliefs, and attitudes for centenarians' happiness. Psychol Aging 2006 Jun;21(2):266-280. [doi: 10.1037/0882-7974.21.2.266] [Medline: 16768574]
78. Rott C, Jopp D. Subjective well-being in centenarians. Glob Ageing 2006;4:52-62. [doi: 10.1007/978-981-287-080-3_187-1]
79. van Wijngaarden E, Leget C, Goossensen A. Ready to give up on life: the lived experience of elderly people who feel life is completed and no longer worth living. Soc Sci Med 2015 Aug;138:257-264 [FREE Full text] [doi: 10.1016/j.socscimed.2015.05.015] [Medline: 25982088]
80. Zhang L, Li JL, Guo LL, Xu G, Yang L, Wang C, et al. Weight changes and the incidence of depressive symptom in the middle-aged and older adults: findings from the Chinese nationwide cohort study. BMC Public Health 2022 Dec 06;22(1):2282. [doi: 10.1186/s12889-022-14624-5] [Medline: 36474202]
81. Weisenbach SL, Boore LA, Kales HC. Depression and cognitive impairment in older adults. Curr Psychiatry Rep 2012 Aug;14(4):280-288. [doi: 10.1007/s11920-012-0278-7] [Medline: 22622478]
82. Muhammad T, Meher T. Association of late-life depression with cognitive impairment: evidence from a cross-sectional study among older adults in India. BMC Geriatr 2021 Jun 15;21(1):364 [FREE Full text] [doi: 10.1186/s12877-021-02314-7] [Medline: 34130632]
83. Kelly ME, Duff H, Kelly S, McHugh Power JE, Brennan S, Lawlor BA, et al. The impact of social activities, social networks, social support and social relationships on the cognitive functioning of healthy older adults: a systematic review. Syst Rev 2017 Dec 19;6(1):259 [FREE Full text] [doi: 10.1186/s13643-017-0632-2] [Medline: 29258596]
84. Dawson-Townsend K. Social participation patterns and their associations with health and well-being for older adults. SSM Popul Health 2019 Aug;8:100424 [FREE Full text] [doi: 10.1016/j.ssmph.2019.100424] [Medline: 31312713]
85. Kandler C, Kornadt AE, Hagemeyer B, Neyer FJ. Patterns and sources of personality development in old age. J Pers Soc Psychol 2015 Jul;109(1):175-191. [doi: 10.1037/pspp0000028] [Medline: 25545842]
86. Dening KH. Differentiating between dementia, delirium and depression in older people. Nurs Stand 2019 Aug 31;35(1):43-50. [doi: 10.7748/ns.2019.e11361] [Medline: 31762251]
87. Jopp DS, Park MS, Lehrfeld J, Paggi ME. Physical, cognitive, social and mental health in near-centenarians and centenarians living in New York City: findings from the Fordham Centenarian Study. BMC Geriatr 2016 Jan 05;16:1 [FREE Full text] [doi: 10.1186/s12877-015-0167-0] [Medline: 26729190]
88. Ribeiro O, Duarte N, Teixeira L, Paúl C. Frailty and depression in centenarians. Int Psychogeriatr 2018 Jan;30(1):115-124 [FREE Full text] [doi: 10.1017/S1041610217001910] [Medline: 28931449]
89. Sinoff G. Thanatophobia (death anxiety) in the elderly: the problem of the child's inability to assess their own parent's death anxiety state. Front Med (Lausanne) 2017;4:11 [FREE Full text] [doi: 10.3389/fmed.2017.00011] [Medline: 28289681]

XSL•FO
RenderX

90. Bierman A, Upenieks L, Lee Y, Harmon M. Consequences of financial strain for psychological distress among older adults: examining the explanatory role of multiple components of the self-concept. Socius 2023 Oct 13;9(12):e1001940 [FREE Full text] [doi: 10.1177/23780231231197034] [Medline: 26701262]
91. Zaccaria D, Cavalli S, Masotti B, Gomes Da Rocha C, von Gunten A, Jopp DS. Social isolation and loneliness among near-centenarians and centenarians: results from the Fordham centenarian study. Int J Environ Res Public Health 2022 May 13;19(10):5940 [FREE Full text] [doi: 10.3390/ijerph19105940] [Medline: 35627476]
92. Cimarolli VR, Jopp DS, Boerner K, Minahan J. Depressive symptoms in the oldest-old: the role of sensory impairments. Arch Gerontol Geriatr 2018;78:249-254. [doi: 10.1016/j.archger.2018.07.009] [Medline: 30032074]
93. Li R, Zong ZY, Gu XX, Wang DN, Dong C, Sun C, et al. Higher dietary diversity as a protective factor against depression among older adults in China: a cross-sectional study. Ann Palliat Med 2022 Apr;11(4):1278-1289 [FREE Full text] [doi: 10.21037/apm-21-1449] [Medline: 34775772]
94. Madrigal-Leer F, Martìnez-Montandòn A, Solìs-Umaña M, Helo-Guzmàn F, Alfaro-Salas K, Barrientos-Calvo I, et al. Clinical, functional, mental and social profile of the Nicoya Peninsula centenarians, Costa Rica, 2017. Aging Clin Exp Res 2020 Feb;32(2):313-321. [doi: 10.1007/s40520-019-01176-9] [Medline: 30919261]
95. Merino MD, Sánchez-Ortega M, Elvira-Flores E, Mateo-Rodríguez I. Centenary personality: are there psychological resources that distinguish centenarians? J Happiness Stud 2023 Nov 08;24(8):2723-2745. [doi: 10.1007/s10902-023-00700-z]
96. Nübel J, Guhn A, Müllender S, Le HD, Cohrdes C, Köhler S. Persistent depressive disorder across the adult lifespan: results from clinical and population-based surveys in Germany. BMC Psychiatry 2020 Feb 10;20(1):58 [FREE Full text] [doi: 10.1186/s12888-020-2460-5] [Medline: 32041560]
97. Comijs HC, Nieuwesteeg J, Kok R, van Marwijk HW, van der Mast RC, Naarding P, et al. The two-year course of late-life depression; results from the Netherlands study of depression in older persons. BMC Psychiatry 2015 Feb 12;15:20 [FREE Full text] [doi: 10.1186/s12888-015-0401-5] [Medline: 25775143]
98. Pinho LG, Lopes MJ, Correia T, Sampaio F, Arco HR, Mendes A, et al. Patient-centered care for patients with depression or anxiety disorder: an integrative review. J Pers Med 2021 Aug 10;11(8):e510 [FREE Full text] [doi: 10.3390/jpm11080776] [Medline: 34442420]

Abbreviations

LLD: late-life depression

Edited by A Garcia; submitted 08.09.24; peer-reviewed by K Mikkonen; comments to author 24.01.25; revised version received 29.01.25; accepted 01.02.25; published 05.03.25.

Please cite as:
Gomes da Rocha C, von Gunten A, Vandel P, Jopp DS, Ribeiro O, Verloo H
Building Consensus on the Relevant Criteria to Screen for Depressive Symptoms Among Near-Centenarians and Centenarians: Modified e-Delphi Study
JMIR Aging 2025;8:e64352
URL: https://aging.jmir.org/2025/1/e64352
doi:10.2196/64352
PMID:40053803

©Carla Gomes da Rocha, Armin von Gunten, Pierre Vandel, Daniela S Jopp, Olga Ribeiro, Henk Verloo. Originally published in JMIR Aging (https://aging.jmir.org), 05.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Cyberchondria in Older Adults and Its Relationship With Cognitive Fusion, Health-Related Quality of Life, and Mental Well-Being: Mediation Analysis

Richard Huan Xu[1], PhD; Vladan Starcevic[2], Prof Dr Med

[1]Department of Rehabilitation Sciences, Faculty of Health and Social Sciences, Hong Kong Polytechnic University, Kowloon, China (Hong Kong)
[2]Faculty of Medicine and Health, Sydney Medical School, Nepean Clinical School, Brain and Mind Centre, University of Sydney, Sydney, Australia

Corresponding Author:
Richard Huan Xu, PhD
Department of Rehabilitation Sciences
Faculty of Health and Social Sciences
Hong Kong Polytechnic University
9 Hoi Ting Rd
Hung Hom
Kowloon
China (Hong Kong)
Phone: 852 27664199
Email: richard.xu@polyu.edu.hk

Abstract

Background: Cyberchondria is the compulsive searching for health information online that continues despite harmful effects. It leads to increased health anxiety and lower health-related quality of life (HRQOL). Older adults face higher risks of cyberchondria due to their limited digital literacy skills and more frequent health concerns. However, researchers have not thoroughly studied how cyberchondria affects this age group.

Objective: This study aimed to explore cyberchondria in the older population and investigate its relationship with cognitive fusion (ie, the tendency to become entangled with thoughts and perceive them as literal truths that dictate behavior), HRQOL, and mental well-being.

Methods: A web-based, cross-sectional survey was conducted in May 2024 with a sample of 638 participants from China aged ≥60 years recruited through the online panel of a survey company. The participants completed questionnaires assessing cyberchondria (using the Cyberchondria Severity Scale-12 [CSS-12]), cognitive fusion, HRQOL, and mental well-being. Structural equation modeling (SEM) was used to assess the hypothesized mediation model, and standardized estimates and their 95% CIs were calculated for all structural paths.

Results: Participants had a mean CSS-12 score of 40 (SD 8.5), suggesting a fairly high level of cyberchondria in this sample. Participants with a higher socioeconomic status tended to report lower levels of cyberchondria. The SEM showed that cyberchondria was positively associated with cognitive fusion (β=0.505, *P*<.001 for both models) and negatively associated with HRQOL (β=–0.221, *P*<.001) and mental well-being (β=–0.212, *P*<.001). The mediation model showed a good fit and demonstrated that cognitive fusion fully mediated the total effect of cyberchondria on HRQOL and mental well-being.

Conclusions: Cyberchondria may be more prominent in older Chinese adults, especially those residing in rural areas and with a lower socioeconomic status. Additionally, cyberchondria can enhance cognitive fusion, contributing to poor HRQOL and mental well-being. Interventions focused on “defusing” cyberchondria-relevant thoughts may help reduce maladaptive behaviors associated with cyberchondria and improve the overall well-being of older populations.

(JMIR Aging 2025;8:e70302) doi:10.2196/70302

KEYWORDS

cyberchondria; cognitive fusion; health-related quality of life; well-being; mediation

XSL•FO
RenderX

Introduction

Background

The internet can amplify health anxiety by creating an environment in which distinguishing credible online health information from that derived from unreliable sources may be difficult, often leading to confusion and fear [1,2]. Cyberchondria refers to excessive or repeated online health information seeking that persists despite negative consequences and is associated with increased health anxiety [3] and decreased health-related quality of life (HRQOL) [4,5], with its toll including diminished well-being and compromised daily functioning. Cyberchondria and health anxiety are closely intertwined, and studying this link is crucial to address the growing impact of cyberchondria. The rise of cyberchondria is intimately linked to the digital revolution, and online misinformation and worst-case scenarios often lead to catastrophic interpretations that exacerbate anxiety. The internet also facilitates confirmation bias, allowing individuals to seek information that aligns with their existing fears, reinforcing anxiety. Easy access to information facilitates compulsive checking behaviors, which provide temporary relief but ultimately perpetuate anxiety. This creates a vicious cycle, in which increased anxiety leads to more compulsive online health information seeking [4,6,7].

Cyberchondria in Older Adults

Older adults face heightened vulnerability to cyberchondria. Their vulnerability stems not only from lower digital literacy and more frequent health concerns [8,9] but also from their generation's unique beliefs and life experiences [10]. Older adults, particularly those from generations that emphasized deference to authority figures, may uncritically accept health-related information presented in an authoritative tone, even from unreliable online sources [11]. This attitude, coupled with passive health information-seeking habits developed in an era of limited access to medical knowledge, heightens susceptibility to misinformation and catastrophic health interpretations.

Moreover, older age brings an expectation of decline, a psychological state shaped by imminent health risks, accumulated losses, and direct experiences with chronic illness [12]. These experiences may foster hypervigilance toward bodily changes, prompting compulsive online searches that reinforce anxiety through confirmation bias. The interplay of factors, such as cohort beliefs, anticipatory health anxiety, and cumulative adversity, may create a unique pathway for cyberchondria in older adults, distinct from younger populations. Despite these risks, limited research has focused on cyberchondria in this age group, even as studies highlight their frequent use of the internet for health information, which may increase their anxiety and distress [12-14].

HRQOL and Cyberchondria

HRQOL is a multidimensional construct that encompasses an individual's physical health, mental health, social functioning, and role functioning [15]. HRQOL is particularly affected by cyberchondria, as the anxiety and stress from constant health-related searches have a negative impact on emotional well-being. This anxiety can also manifest as psychosomatic symptoms, affecting physical health and daily functioning.

Unravelling the mechanisms linking cyberchondria to HRQOL is essential, as individual differences in coping strategies, digital literacy, and pre-existing health conditions interact dynamically with the aforementioned specific vulnerabilities in older adults. As cyberchondria becomes more prevalent [16], targeted interventions must address not only the skill gaps but also the ingrained cognitive and emotional patterns that predispose older adults to its harms, ultimately safeguarding their well-being in an increasingly digital health landscape.

Mental Well-Being and Cyberchondria

Mental well-being and HRQOL are distinct concepts that address different dimensions of human experience. Unlike HRQOL, which quantifies the impact of health on physical, social, and role functioning, mental well-being focuses on emotional and psychological states and encompasses positive aspects such as happiness, life satisfaction, resilience, and a sense of purpose [17]. Rooted in frameworks like positive psychology, mental well-being relates to the presence of flourishing mental states (eg, optimism, self-acceptance) and absence of psychological distress (eg, anxiety, depression). Cyberchondria can jeopardize mental well-being by heightening anxiety about one's health [18]. It can also destabilize emotional health, trigger obsessive thought patterns, and impair resilience, ultimately eroding the core components of mental well-being and leaving individuals trapped in a cycle of distress and hypervigilance [19].

Elucidating the mechanisms connecting cyberchondria to mental well-being is vital, especially for older adults who face unique risks due to diminished digital literacy, heightened health concerns, and reduced adaptability to stress [20]. Factors such as pre-existing anxiety, limited ability to critically evaluate online content, and a propensity for catastrophic thinking can intensify cyberchondria's impact, making older adults particularly susceptible to its psychological toll [21]. With the growing prevalence of digital health-seeking behaviors, research must prioritize these interactions to inform robust, tailored interventions, combining enhanced digital education with cognitive-behavioral strategies, to break the cycle of anxiety and safeguard mental well-being in this vulnerable population [22].

Cognitive Fusion and Cyberchondria

Cognitive fusion refers to the tendency for behavior to be overly regulated and influenced by cognition. Cognitive fusion occurs when individuals get overly entangled with their thoughts, treating them as literal truths rather than just mental phenomena [23]. Although cognitive fusion is primarily a state phenomenon, some individuals may experience it more persistently, leading to patterns that resemble a trait. This means that certain people might be more prone to cognitive fusion across various situations, making it appear as a stable, trait-like characteristic [24]. Studies have shown that cognitive fusion is strongly associated with health care avoidance and weight stigma [25] and that higher levels of cognitive fusion are associated with

XSL•FO
RenderX

increased anxiety and depression and poorer HRQOL [26,27]. Cognitive fusion is also related to rumination, shame, and reduced self-compassion [28]. Cognitive fusion involving anxious thoughts can exacerbate anxiety-related maladaptive behaviors, such as avoidance, checking, and reassurance seeking. The inability to distance oneself from such thoughts perpetuates the anxiety cycle and strengthens the perceived validity of the thoughts [29].

Previous studies have shown an association between cyberchondria and metacognitive beliefs [30,31], a concept potentially linked to cognitive fusion. Although metacognitive beliefs refer to ways of appraising one's thoughts and other cognitive functions, cognitive fusion is about a strong attachment to one's thoughts and their perceived truth. Compared with cognitive fusion, metacognitive beliefs reflect a broader construct, can be either positive or negative, and lack an element of attachment to one's own thoughts, regardless of the nature of these thoughts. Both constructs influence how individuals experience and respond to their thoughts, but considering the potential role that cognitive fusion can play in psychopathology, it has not received sufficient attention from researchers in the context of cyberchondria.

Although cognitive fusion can affect people of all ages, older adults may be more severely impacted than their younger counterparts, given their thoughts about numerous life transitions, including retirement, loss of loved ones, high risks of social isolation, and declining health [32]. Cognitive fusion may amplify health-related anxiety of older adults by making it difficult for them to separate their concerning health-related thoughts from the reality about their health. However, limited empirical evidence exists on the relationship between cognitive fusion and health anxiety among older adults [33]. Studying cognitive fusion in this demographic may help develop tailored interventions to enhance coping mechanisms of older adults and improve their mental health and overall HRQOL [34].

The relationship between cognitive fusion and cyberchondria can be understood through the lens of cognitive processes and reinforcement. Cognitive fusion may contribute to negative perceptions of online health information [35]. In other words, the threatening thoughts about online health information, especially if this information is ambiguous or incongruent, are experienced as real and not a product of one's own interpretation or perception. This increases anxiety and makes a person search compulsively for health information in an effort to alleviate the sense of threat, thus leading to cyberchondria [29,36]. However, such behavior only reinforces the sense of threat, especially in the long run. For older adults whose health-related searches are often driven by anticipatory anxiety [12], seeking ambiguous or alarming information may deepen cognitive fusion ("I must be ill because I keep reading about symptoms"), creating a feedback loop. This aligns with studies showing that repetitive behaviors amplify cognitive rigidity in aging populations [33]. Thus, we hypothesized cyberchondria as a behavioral trigger that intensifies cognitive fusion, which in turn erodes HRQOL.

Research Questions and Hypotheses

Research Questions

Drawing from the theoretical and empirical considerations outlined earlier, this study proposed 2 research questions to guide the investigation. First, what are the levels of cyberchondria and cognitive fusion among older adults, and which sociodemographic factors correlate with these constructs? Second, how are cyberchondria, cognitive fusion, HRQOL, and mental well-being interrelated in older adults? To address the second question, 5 hypotheses were formulated. The conceptual framework is presented in Multimedia Appendix 1.

Hypothesis 1

Prior research has demonstrated that cyberchondria contributes to functional impairment and diminished quality of life [18]. In older adults, this effect may be amplified due to age-related vulnerabilities, such as reduced online health literacy and the presence of chronic illnesses, which heighten susceptibility to health anxiety and its consequences [37]. Therefore, we proposed, as hypothesis 1 (H1), that cyberchondria negatively impacts HRQOL in older adults.

Hypothesis 2

Studies have shown that excessive online health information seeking is linked to increased psychological distress, including heightened anxiety and reduced emotional well-being [38,39]. Older adults may be particularly vulnerable to these effects due to declining physical health, reduced resilience to stress, and limited digital literacy, which can exacerbate the psychological toll of cyberchondria. Thus, we hypothesized, as hypothesis 2 (H2), that cyberchondria negatively impacts mental well-being in older adults.

Hypothesis 3

Research has indicated that cognitive fusion, characterized by entanglement with thoughts and treating them as literal truths, is associated with anxiety-driven behaviors [40]. Cyberchondria, as a behavioral pattern driven by health anxiety, may similarly foster cognitive fusion by reinforcing rigid and catastrophic thought patterns about health. Accordingly, we proposed, as hypothesis 3 (H3), that cyberchondria positively impacts cognitive fusion in older adults.

Hypothesis 4

The theoretical framework of acceptance and commitment therapy, supported by studies, suggests that cognitive fusion amplifies the negative effects of maladaptive thoughts on well-being, including health-related outcomes [41]. In the context of cyberchondria, older adults who excessively seek online health information may develop rigid, catastrophic cognitions that impair HRQOL through cognitive fusion [42]. Therefore, we hypothesized, as hypothesis 4 (H4), that cognitive fusion mediates the negative impact of cyberchondria on HRQOL in older adults.

Hypothesis 5

Evidence from prior research indicates that cognitive fusion intensifies the emotional impact of anxiety-driven behaviors, mediating the association between repetitive thought patterns

XSL•FO
RenderX

and psychological distress [40]. For older adults, cyberchondria may similarly exacerbate emotional distress through cognitive fusion, undermining mental well-being [23]. Thus, we proposed, as hypothesis 5 (H5), that cognitive fusion mediates the negative impact of cyberchondria on mental well-being in older adults.

Methods

Data and Participants

This study used data from a web-based cross-sectional survey conducted between April 2024 and May 2024 to examine the health and social status of China's older population. The participants were recruited through Wenjuanxing, a Chinese survey company with an online panel of over 2.6 million members. The inclusion criteria were as follows: (1) age ≥60 years, (2) ability to read and speak Mandarin, (3) absence of cognitive impairments, and (4) ability to provide informed consent. Eligible participants were invited to complete a series of questionnaires starting with an informed consent form. Only those who consented to participate then completed the questionnaire. Several measures were implemented to ensure data quality. First, we used CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) to prevent bot submissions, ensuring that only genuine human respondents participated. We also conducted a time analysis to exclude implausibly rapid responses, which could indicate a lack of thoroughness. Additionally, we limited submissions to one per IP address within a set time frame to prevent duplicate entries. To further enhance data integrity, we identified and filtered parallel response patterns, including both consistent and repetitive patterns, which could suggest inattentive or automated responses.

Measures

The standard and validated Chinese versions of all instruments were administered. The English versions are included in Multimedia Appendix 2 for reference only.

Cyberchondria

The severity of cyberchondria was measured using the Cyberchondria Severity Scale-12 (CSS-12). It possesses good psychometric properties, comparable to those of the original version, and has been validated in the Chinese population [43]. The CSS-12 items were rated on a Likert-type scale ranging from 1 (never) to 5 (always). The total score ranges from 12 to 60, with higher scores indicating higher levels of cyberchondria.

Cognitive Fusion

The Cognitive Fusion Questionnaire (CFQ) was designed to measure the extent to which individuals are entangled in their thoughts [23]. It consists of 7 items, each rated on a 7-point Likert scale ranging from 1 (never true) to 7 (always true). Total scores range from 7 to 49, with higher scores reflecting a higher level of cognitive fusion. The psychometric properties of the CFQ have been reported to be satisfactory in the Chinese population [44].

HRQOL

The EQ-5D-5L was used to measure HRQOL in this study. It comprises 5 health-related items (mobility, self-care, usual activities, pain or discomfort, and anxiety or depression), each rated on a 5-point Likert scale ranging from 1 (no problems) to 5 (extreme problems) [45]. To reflect HRQOL, all health states described by the descriptive system can be converted into a single utility score using a scoring algorithm based on public preferences. This study used the EQ-5D-5L China value set and scoring algorithm [46].

Mental Well-Being

The World Health Organization-5 Well-Being Index (WHO-5) is a widely recognized and validated tool for assessing subjective psychological well-being [47]. It consists of 5 questions that measure positive mood, vitality, and general interest over the past 2 weeks. Each item is rated on a 6-point Likert scale ranging from 0 (not present) to 5 (constantly present), allowing for a maximum score of 25. Higher scores indicate better well-being. The psychometric properties of the WHO-5 in the Chinese population have been reported to be satisfactory [48].

Statistical Analysis

Descriptive statistics were used to analyze participants' sociodemographic characteristics. Continuous variables (eg, age) were analyzed using means and SDs, whereas categorical variables (eg, sex) were analyzed using frequencies and percentages. Pearson correlation coefficients (r) were calculated to determine associations between variables, with r values ≥0.3 and ≥0.5 indicating moderate and large effects, respectively [49]. ANOVA was conducted to assess the differences in CSS-12 scores across different socioeconomic groups.

We used structural equation modeling (SEM) with full-information likelihood estimation to assess the hypothesized mediation model. Latent variables for cyberchondria (CSS-12), HRQOL (EQ-5D-5L), mental well-being (WHO-5), and cognitive fusion (CFQ) were created based on the sum score of the 4 instruments. Respondents' characteristics (ie, sex, age, education level, and presence of a chronic condition) were included in the model for adjustment. To assess the direct, indirect, and total effects, we used 5000 bootstrapped samples, derived effect estimates, and bias-corrected 95% CIs [50]. We assessed model fit using a Tucker-Lewis index>0.90, comparative fit index>0.90, root mean square error of approximation<0.05, and standardized root mean square residual<0.05 [51]. We set significance at P<.05 (2-tailed) and conducted all statistical analyses using R software [52].

Ethical Considerations

This study was performed in line with the principles of the Declaration of Helsinki. The study protocol and informed consent were approved by the Institutional Review Board of Hong Kong Polytechnic University (reference ID: HSEARS20230502006). Informed consent was obtained from all the participants. All data were anonymized to protect participants' privacy and ensure confidentiality.

XSL•FO
RenderX

Results

Participants' Sociodemographic Characteristics

Table 1 presents the sociodemographic characteristics of the 638 survey participants. The majority were men, comprising 64.1% (409/637) of the sample. In terms of educational background, 51.4% (328/637) had a high school education or less, 25.9% (165/637) had attended college, and 22.7% (145/637) held a university degree. Regarding their household registration, 37.5% (239/637) were from rural areas, while 62.5% (399/637) were from urban areas. When asked about their perceived socioeconomic status, 11% (70/637) considered themselves below the local average, 79.1% (505/637) felt they were equal to it, and 9.9% (63/637) viewed themselves as above it. Moreover, 60.7% (387/637) were long-term caregivers, and 55.3% (353/537) had chronic conditions.

Table 1. Participants' demographics and socioeconomic status (n=637).

Characteristics	Results, n (%)
Sex	
Male	409 (64.1)
Female	229 (35.9)
Age (years)	
60-65	533 (81.6)
66-79	105 (18.4)
Educational level	
High school or less	328 (51.4)
College	165 (25.9)
University	145 (22.7)
Household registration	
Rural	239 (37.5)
Urban	399 (62.5)
Marital status	
Single	9 (1.4)
Married	578 (90.6)
Divorced/widow(er)	51 (8)
Perceived socioeconomic status	
Lower than local average	70 (11)
Equal to local average	505 (79.1)
Higher than local average	63 (9.9)
Long-term caregiver	
Yes	387 (60.7)
No	251 (39.3)
Chronic conditions	
Yes	353 (55.3)
No	285 (44.7)

Profiles of Cyberchondria, HRQOL, Mental Well-Being, and Cognitive Fusion and Associations Between Them

Table 2 presents the participants' scores on the 4 assessment instruments, while Table 3 shows the correlations among the instruments. The average CSS-12 (cyberchondria) score was 40 (SD 8.5) points. The mean WHO-5 (mental well-being) score was 20.4 (SD 5.0) points, the mean EQ-5D-5L (HRQOL) score was 0.8 (SD 0.19) points, and the mean CFQ (cognitive fusion) score was 36.8 (SD 12.3) points. The Cronbach α coefficients for all 4 instruments were above 0.8, indicating good internal consistency and reliability.

XSL•FO
RenderX

Table 2. Measure profiles.

Measures	Mean (SD)	Median (range)	Cronbach α
CSS-12[a]	40 (8.5)	42 (12-58)	0.87
WHO-5[b]	20.4 (5.0)	21 (5-30)	0.87
EQ-5D-5L	0.8 (0.19)	0.88 (0.08-1)	0.83
CFQ[c]	36.8 (12.3_	39 (9-63)	0.94

[a]CSS-12: Cyberchondria Severity Scale-12.

[b]WHO-5: World Health Organization-5 Well-Being Index.

[c]CFQ: Cognitive Fusion Questionnaire.

Table 3. Correlations between the measures.

Measures	CSS-12[a]	WHO-5[b]	EQ-5D-5L	CFQ[c]
CSS-12				
r	—[d]	–0.20	–0.30	0.50
P value	—	<.001	<.001	<.001
WHO-5				
r	–0.20	—	0.36	–0.42
P value	<.001	—	<.001	<.001
EQ-5D-5L				
r	–0.30	0.36	—	–0.48
P value	<.001	<.001	—	<.001
CFQ				
r	0.50	–0.42	–0.48	—
P value	<.001	<.001	<.001	—

[a]CSS-12: Cyberchondria Severity Scale-12.

[b]WHO-5: World Health Organization-5 Well-Being Index.

[c]CFQ: Cognitive Fusion Questionnaire.

[d]Not applicable.

Mental well-being (WHO-5) was positively correlated with HRQOL (EQ-5D-5L; r= 0.36, P<.001), suggesting that greater well-being aligns with better HRQOL. Conversely, mental well-being was negatively correlated with cognitive fusion (CFQ; r=–0.42, P<.001), indicating that higher well-being corresponds with reduced cognitive fusion. Similarly, HRQOL was negatively correlated with cognitive fusion (r=–0.48, P<.001), demonstrating that better HRQOL is linked to lower cognitive fusion. Cyberchondria (CSS-12) had negative associations with mental well-being (r=–0.2, P<.001) and HRQOL (r=–0.3, P<.001) and a positive association with cognitive fusion (r=0.5, P<.001). These results suggest that elevated cyberchondria is associated with poorer mental well-being, diminished HRQOL, and heightened cognitive fusion (Table 3).

Differences in Cyberchondria Across Sociodemographic Groups

Table 4 illustrates the differences in cyberchondria (CSS-12) scores across various sociodemographic groups. Although no statistically significant differences were observed between the sexes, substantial variations were evident in all other group comparisons. Notably, individuals with higher educational levels, who resided in urban areas, who were younger, who had a higher socioeconomic status, who did not need caregivers, and who had no chronic conditions had lower cyberchondria scores.

XSL•FO
RenderX

Table 4. Participants' responses on the Cyberchondria Severity Scale-12 (CSS-12) and Cognitive Fusion Questionnaire (CFQ), stratified by demographics.

Demographic characteristics	Cyberchondria (CSS-12)		Cognitive fusion (CFQ)	
	Mean (SD)	*P* value	Mean (SD)	*P* value
Sex		.36		.72
Male	39.8 (8.8)		37.0 (12.4)	
Female	40.4 (7.8)		36.6 (12.2)	
Educational level		<.001		<.001
High school	41.0 (7.8)		38.5 (12.2)	
College	39.5 (8.3)		35.8 (11.8)	
University	38.3 (9.8)		34.5 (12.7)	
Household registration		<.001		<.001
Rural	42.6 (6.4)		41.6 (10.0)	
Urban	38.5 (9.2)		34.0 (12.7)	
Age group (years)		.02		.02
60-65	38.3(9.3)		34.4(12.5)	
66-79	40.3(8.2)		37.4(12.2)	
Socioeconomic status		.01		<.001
Lower than average	42.1 (8.7)		43.4 (11.0)	
Average	39.9 (8.2)		36.3 (12.1)	
Higher than average	38.5 (9.6)		34.5 (13.6)	
Do you need a caregiver?		.001		<.001
Yes	39.1 (8.9)		35.6 (12.8)	
No	41.4 (7.5)		38.9 (11.3)	
Do you have chronic condition?		<.001		<.001
Yes	41.1 (7.7)		40.0 (11.0)	
No	38.6 (9.1)		33.0 (12.7)	

Differences in Cognitive Fusion Across Sociodemographic Groups

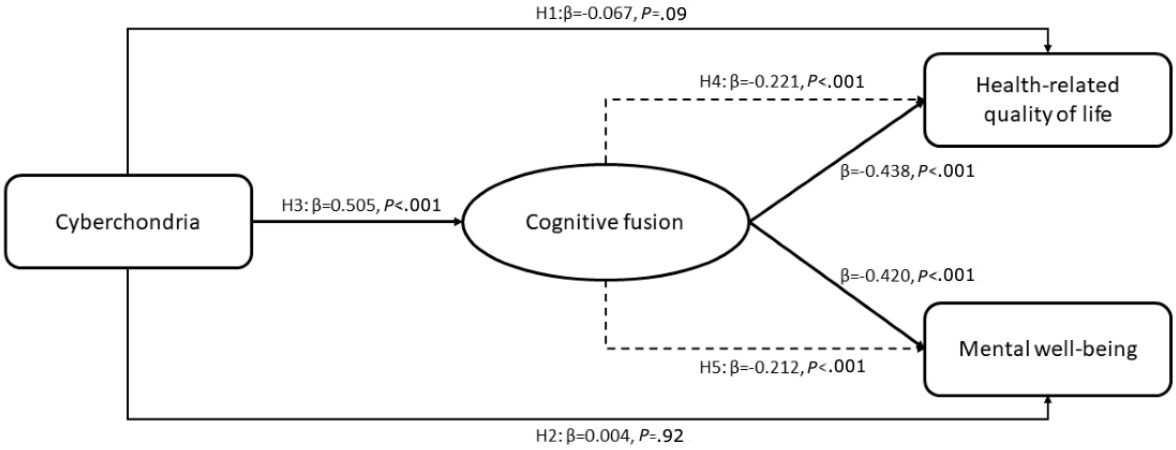
The differences in respondents' cognitive fusion across socioeconomic groups were similar to those found for levels of cyberchondria. Participants exhibited significantly higher cognitive fusion if they had lower educational levels, lived in rural areas, were older, had a lower socioeconomic status, and had caregivers or chronic conditions (Table 4).

Results of the Hypothesized Mediation Model

Figure 1 shows the results of the SEM analysis. The mediation model, adjusted for sex, age, and education level, demonstrated acceptable model indices (comparative fit index=0.997, Tucker-Lewis index=0.993, root mean square error of approximation=0.017), suggesting that unmeasured potential confounders were relatively unlikely to affect the results. Outcomes of the SEM analysis showed the lack of statistically significant direct effects between cyberchondria and HRQOL (β=–0.067, *P*=.09; H1 is not supported) and between cyberchondria and mental well-being (β=0.004, *P*=.92; H2 is not supported). The coefficients of the mediation paths were statistically significant. Cyberchondria was significantly positively associated with cognitive fusion (β=0.505, *P*<.001; H3 is supported) with significant negative indirect associations with HRQOL (β=–0.221, *P*<.001; H4 is supported) and mental well-being (β=–0.212, *P*<.001; H5 is supported).

XSL•FO
RenderX

Figure 1. Results of mediation analysis.



Discussion

Principal Findings

This study demonstrates that the levels of cyberchondria in the sample of older Chinese adults were rather high. Our participants had an average score of 40 (of the maximum score of 60) on the CSS-12, which surpassed the mean scores found in earlier research [53-56]. According to the CSS-12 cutoff value established by Xu [57], 60% of our participants exhibited significant features of cyberchondria. These results provide empirical evidence that internet-related mental health issues affect older adults, challenging the belief that such problems are mainly of relevance for younger individuals [58,59]. Additionally, the mediation analysis revealed that associations between cyberchondria and both HRQOL and mental well-being were complex, as no significant direct effects were observed in the model (H1 and H2 were not supported). However, cyberchondria appeared to have a negative and indirect effect on HRQOL and mental well-being by enhancing cognitive fusion among older adults (H3, H4, and H5 were supported). This highlights the importance of addressing cyberchondria not just as a mental health issue but also as a broader HRQOL concern. Interventions that aim to improve overall well-being, not just reduce anxiety, in this population should be encouraged. However, since the significant associations disappeared in the mediation models, the mechanisms relating cyberchondria to HRQOL and mental well-being might be complex and call for further research scrutiny.

The factors contributing to the high levels of cyberchondria are complex. Our study found that older adults with a lower socioeconomic status reported significantly higher cyberchondria levels than those with a higher socioeconomic status. This finding is consistent with results of previous studies conducted in other populations. For example, a Turkish study revealed a significant relationship between cyberchondria levels and family income among adolescents [60]. Several reasons may help explain this. Older individuals with a lower socioeconomic status often have limited access to health care resources, poor internet connectivity, and lower eHealth literacy [61,62]. These factors may compel them to rely heavily on potentially unreliable online health information. Moreover, older individuals with a lower socioeconomic status are more likely to experience social isolation and have smaller support networks [63,64], which may exacerbate their health anxiety and lead to higher levels of cyberchondria. Given these complexities, cyberchondria should be studied within the broader context of psychosocial determinants of health rather than as a simple health-related phenomenon.

A significant difference in CSS-12 scores emerged between urban and rural residents. Older urban residents exhibited lower levels of cyberchondria than their rural counterparts. Although a previous study found that health anxiety is more prevalent and severe in rural areas than in urban areas [65], the study empirically confirmed that the variation in cyberchondria levels was linked to household registration. The reasons for the striking difference in cyberchondria levels between urban and rural older adults largely overlap with the aforementioned reasons for the difference in cyberchondria levels between older individuals with higher and lower socioeconomic statuses. In addition, the urban migration of younger generations in China often leaves older family members in rural areas, which can hinder older people's access to direct support for using digital technologies [66]. Thus, a shift to digital health information, emphasized by the government, may become a challenge for rural residents [67,68].

The cognitive fusion scores in our sample of older adults were slightly higher than those reported in younger populations [32,69] but lower than those in populations with mental health problems, such as depression [70] and suicidal intention [71]. Gillanders et al [26] found that cognitive fusion is negatively associated with HRQOL, which is consistent with our findings. Additionally, we found a significant association between cognitive fusion and education; older respondents with higher educational levels had lower levels of cognitive fusion. Previous studies indicated that education often enhances cognitive flexibility [72], improves critical-thinking skills [73], and boosts problem-solving abilities [74]. These factors may help better-educated individuals avoid rigid thought patterns, thus reducing their proneness to cognitive fusion.

The mediation model revealed that cyberchondria may increase cognitive fusion, leading individuals to become overly entangled with their thoughts about online health information. Heightened

XSL•FO
RenderX

cognitive fusion can negatively impact HRQOL and mental well-being by making it challenging for older individuals to manage their distressing thoughts and emotions. This finding provides support to H2 of this study and partially aligns with the findings of previous studies showing that cognitive fusion significantly mediates the relationship between individual well-being and other health-related factors [75-78].

Building on previous research and clinical findings, our study extends the understanding of how cognitive fusion may serve as a potential mechanism by which cyberchondria causes negative health outcomes [79]. These findings provide valuable insights into how excessive online health information seeking may affect mental health and other functions. This study has clinical implications, as it suggests that interventions targeting cognitive fusion may be effective in mitigating the impact of cyberchondria on HRQOL and mental well-being.

As populations worldwide age, many countries are experiencing a surge in internet use among older adults, driven by increased access to smartphones and online platforms. The rising trend of cyberchondria among Chinese older adults found in this study is very relevant for other regions with aging populations, such as Japan, the United States, and much of Europe, where older adults increasingly turn to the web for health information [80]. The internet's role in amplifying anxiety through misinformation and compulsive online behavior is a universal concern [81]. Other countries can learn from China's experience to prepare for similar challenges, particularly where health care systems might struggle because of the increasing number of older patients with high levels of health anxiety and cyberchondria.

Strengths and Limitations

This is the first study conducted in China to specifically examine cyberchondria and its sociodemographic correlates in the older population. The novelty of this study lies in its examination of previously overlooked relationships between cyberchondria, cognitive fusion, HRQOL, and mental well-being in this population. The findings of this study enhance our understanding of cyberchondria by linking a specific cognitive construct (ie, cognitive fusion) with behavioral aspects (ie, excessive online health information seeking).

However, this study has several limitations that warrant consideration. First, our web-based sampling method, despite using various quality assurance techniques, may have introduced selection bias. Individuals unfamiliar with online surveys or those not part of the company's panel may have been excluded, potentially limiting the generalizability of our findings. Second, although national census data for the population aged 60 years and older was unavailable, our sample was representative only to some extent of the demographic characteristics of China's older population, as reported in previous studies. Specifically, our sample had higher proportions of educated individuals, men, and urban residents. These sampling differences may introduce bias in our findings. Third, our reliance on self-reported measures means that participants' subjective views may have influenced their responses. Fourth, the study did not assess some potentially important variables, such as digital literacy. This was due to the limited length of the study questionnaire and the need to reduce the cognitive burden on older respondents. Variables like digital literacy may affect cyberchondria, and their direct or moderating effects should be explored in future studies. Fifth, our decision to model HRQOL using the EQ-5D-5L single utility score as a manifest variable in the SEM analysis deviates from conventional SEM practices. An alternative approach, such as conceptualizing HRQOL as a latent construct with the five EQ-5D-5L dimensions as formative indicators, might better capture the multidimensional nature of HRQOL. Although this alternative does not align with the standard use of the EQ-5D-5L utility score, it might reflect the differential contributions of each dimension and adhere more closely to typical SEM frameworks. Our approach may have constrained the model's ability to account for the nuanced interplay of HRQOL dimensions, and future research should consider latent variable modeling to address this limitation. Last, our study used a cross-sectional design. Although SEM can offer some insights into potential causal relationships between variables, longitudinal data will be necessary in the future to establish definitive causal links.

Conclusion

This study shows that almost two-thirds of the older Chinese online population might exhibit features of cyberchondria. Higher levels of cyberchondria were observed among older adults living in rural areas and those with lower socioeconomic status. Targeted educational programs could be useful for empowering older adults to better navigate cyberspace and identify credible online health information. Cyberchondria can enhance cognitive fusion in older populations, resulting in poorer HRQOL and mental well-being. Interventions aimed at "defusing" cyberchondria-relevant thoughts and breaking the cycle of anxiety-fueled, excessive online health information seeking can help reduce cyberchondria and improve the overall well-being of older populations.

Data Availability

Derived data supporting the findings of this study are available from the corresponding author on request.

Authors' Contributions

Conceptualization: RHX
Data curation: RHX
Investigation: RHX
Software: RHX
Supervision: RHX, VS

XSL•FO
RenderX

Writing – original draft: RHX
Writing – review & editing:RHX, VS

Conflicts of Interest

None declared.

Multimedia Appendix 1
The conceptual framework of the study and study hypotheses. Hypothesis 1 (H1): Cyberchondria correlates negatively with HRQoL and mental well-being in older adults. Hypothesis 2 (H2): Cognitive fusion mediates the relationships between cyberchondria and reduced HRQoL and mental well-being in older adults.
[DOCX File , 30 KB - aging_v8i1e70302_app1.docx]

Multimedia Appendix 2
Survey questionnaire (English version).
[DOCX File , 27 KB - aging_v8i1e70302_app2.docx]

Multimedia Appendix 3
Factor loadings of the SEM analysis.
[DOCX File , 20 KB - aging_v8i1e70302_app3.docx]

References

1. Starcevic V. Cyberchondria: challenges of problematic online searches for health-related information. Psychother Psychosom 2017 May 11;86(3):129-133 [FREE Full text] [doi: 10.1159/000465525] [Medline: 28490037]
2. Starcevic V, Aboujaoude E. Cyberchondria, cyberbullying, cybersuicide, cybersex: "new" psychopathologies for the 21st century? World Psychiatry 2015 Feb;14(1):97-100 [FREE Full text] [doi: 10.1002/wps.20195] [Medline: 25655165]
3. Starcevic V, Berle D, Arnáez S. Recent insights into cyberchondria. Curr Psychiatry Rep 2020 Aug 27;22(11):56 [FREE Full text] [doi: 10.1007/s11920-020-01179-8] [Medline: 32852626]
4. Ambrosini F, Truzoli R, Vismara M, Vitella D, Biolcati R. The effect of cyberchondria on anxiety, depression and quality of life during COVID-19: the mediational role of obsessive-compulsive symptoms and Internet addiction. Heliyon 2022 May;8(5):e09437 [FREE Full text] [doi: 10.1016/j.heliyon.2022.e09437] [Medline: 35600442]
5. Rahme C, Akel M, Obeid S, Hallit S. Cyberchondria severity and quality of life among Lebanese adults: the mediating role of fear of COVID-19, depression, anxiety, stress and obsessive-compulsive behavior-a structural equation model approach. BMC Psychol 2021 Oct 29;9(1):169 [FREE Full text] [doi: 10.1186/s40359-021-00674-8] [Medline: 34715930]
6. Mathes BM, Norr AM, Allan NP, Albanese BJ, Schmidt NB. Cyberchondria: overlap with health anxiety and unique relations with impairment, quality of life, and service utilization. Psychiatry Res 2018 Mar;261:204-211. [doi: 10.1016/j.psychres.2018.01.002] [Medline: 29324396]
7. Xu RH, Chen C. Moderating effect of coping strategies on the association between the infodemic-driven overuse of health care services and cyberchondria and anxiety: partial least squares structural equation modeling study. J Med Internet Res 2024 Apr 09;26:e53417 [FREE Full text] [doi: 10.2196/53417] [Medline: 38593427]
8. Ding K, Wang S, Xu W, Lin L, Liao D, Chen H, et al. Low mental health literacy and its association with depression, anxiety and poor sleep quality in Chinese elderly. Asia Pac Psychiatry 2022 Dec 09;14(4):e12520. [doi: 10.1111/appy.12520] [Medline: 36210054]
9. Chesser AK, Keene Woods N, Smothers K, Rogers N. Health literacy and older adults: a systematic review. Gerontol Geriatr Med 2016;2:2333721416630492 [FREE Full text] [doi: 10.1177/2333721416630492] [Medline: 28138488]
10. Hill S, Watts D. Googling and health anxiety in older age-a case study of CBT incorporating a single case experimental design. Clin Case Rep 2024 Aug 10;12(8):e9316 [FREE Full text] [doi: 10.1002/ccr3.9316] [Medline: 39130804]
11. Laidlaw K, Thompson LW, Dick-Siskin L, Gallagher-Thompson D. Cognitive behaviour therapy with older people. Hoboken, NJ: John Wiley & Sons Ltd; 2008.
12. Magsamen-Conrad K, Dillon JM, Billotte Verhoff C, Faulkner SL. Online health-information seeking among older populations: family influences and the role of the medical professional. Health Commun 2019 Jul 23;34(8):859-871 [FREE Full text] [doi: 10.1080/10410236.2018.1439265] [Medline: 29474125]
13. Xu RH, Tian L, Zhu L, Cao Y, Chan SK, Dong D, et al. Age differences in electronic mental health literacy: qualitative study. J Med Internet Res 2024 Jul 16;26:e59131 [FREE Full text] [doi: 10.2196/59131] [Medline: 39012686]
14. Zhao YC, Zhao M, Song S. Online health information seeking behaviors among older adults: systematic scoping review. J Med Internet Res 2022 Feb 16;24(2):e34790 [FREE Full text] [doi: 10.2196/34790] [Medline: 35171099]
15. Cheung YB, Thumboo J. Developing health-related quality-of-life instruments for use in Asia: the issues. Pharmacoeconomics 2006;24(7):643-650. [doi: 10.2165/00019053-200624070-00003] [Medline: 16802840]

XSL•FO
RenderX

16. Mestre-Bach G, Potenza MN. Cyberchondria: a growing concern during the COVID-19 pandemic and a possible addictive disorder? Curr Addict Rep 2023 Feb 09;10(1):77-96 [FREE Full text] [doi: 10.1007/s40429-022-00462-3] [Medline: 36785757]
17. Dale H, Brassington L, King K. The impact of healthy lifestyle interventions on mental health and wellbeing: a systematic review. Mental Health Review Journal 2014;19(1):1-26. [doi: 10.1108/MHRJ-05-2013-0016]
18. Starcevic V, Berle D. Cyberchondria: towards a better understanding of excessive health-related Internet use. Expert Rev Neurother 2013 Feb 09;13(2):205-213. [doi: 10.1586/ern.12.162] [Medline: 23368807]
19. Fergus TA. Cyberchondria and intolerance of uncertainty: examining when individuals experience health anxiety in response to Internet searches for medical information. Cyberpsychol Behav Soc Netw 2013 Oct;16(10):735-739. [doi: 10.1089/cyber.2012.0671] [Medline: 23992476]
20. Zheng H, Sin SJ, Kim HK, Theng Y. Cyberchondria: a systematic review. INTR 2020 Oct 27;31(2):677-698. [doi: 10.1108/intr-03-2020-0148]
21. White RW, Horvitz E. Cyberchondria. ACM Trans. Inf. Syst 2009 Nov 30;27(4):1-37. [doi: 10.1145/1629096.1629101]
22. Jamal A, McKenzie K, Clark M. The impact of health information technology on the quality of medical and health care: a systematic review. Health Inf Manag 2009 Oct 01;38(3):26-37. [doi: 10.1177/183335830903800305] [Medline: 19875852]
23. Gillanders DT, Bolderston H, Bond FW, Dempster M, Flaxman PE, Campbell L, et al. The development and initial validation of the cognitive fusion questionnaire. Behav Ther 2014 Jan;45(1):83-101 [FREE Full text] [doi: 10.1016/j.beth.2013.09.001] [Medline: 24411117]
24. García-Gómez M, Guerra J, López-Ramos VM, Mestre JM. Cognitive fusion mediates the relationship between dispositional mindfulness and negative affects: a study in a sample of Spanish children and adolescent school students. Int J Environ Res Public Health 2019 Nov 25;16(23):4687 [FREE Full text] [doi: 10.3390/ijerph16234687] [Medline: 31775280]
25. Harkins A, Bernent M, Hunter S. Relation between Health, Wellness, and Academic Achievement in Adolescent Females. In: J Allied Health. 2016 Presented at: ASAHP 2016 Annual Conference; October 19-21, 2016; New Orleans, LA URL: https://www.jstor.org/stable/48721952
26. Gillanders DT, Sinclair AK, MacLean M, Jardine K. Illness cognitions, cognitive fusion, avoidance and self-compassion as predictors of distress and quality of life in a heterogeneous sample of adults, after cancer. Journal of Contextual Behavioral Science 2015 Oct;4(4):300-311. [doi: 10.1016/j.jcbs.2015.07.003]
27. Guzmán A, Gillanders D, Stevenson A, Ross K. Psychosocial adjustment to mild cognitive impairment: the role of illness perceptions, cognitive fusion and cognitive impairment. Dementia (London) 2021 Feb 16;20(2):464-484. [doi: 10.1177/1471301219893862] [Medline: 31948271]
28. O'Loughlin CM, Bennett DS, O'Hayer CV. The nomological network of cognitive fusion among people living with HIV: associations with rumination, shame, and depressive symptoms. Journal of Contextual Behavioral Science 2020 Jan;15:245-252. [doi: 10.1016/j.jcbs.2020.01.012]
29. Fergus TA. I really believe I suffer from a health problem: examining an association between cognitive fusion and healthy anxiety. J Clin Psychol 2015 Sep 20;71(9):920-934. [doi: 10.1002/jclp.22194] [Medline: 25994488]
30. Fergus TA, Spada MM. Cyberchondria: examining relations with problematic internet use and metacognitive beliefs. Clin Psychol Psychother 2017 Nov;24(6):1322-1330. [doi: 10.1002/cpp.2102] [Medline: 28621035]
31. Nadeem F, Malik N, Atta M, Ullah I, Martinotti G, Pettorruso M, et al. Relationship between health-anxiety and cyberchondria: role of metacognitive beliefs. J Clin Med 2022 May 05;11(9):2590 [FREE Full text] [doi: 10.3390/jcm11092590] [Medline: 35566713]
32. Bodenlos JS, Hawes ES, Burstein SM, Arroyo KM. Association of cognitive fusion with domains of health. Journal of Contextual Behavioral Science 2020 Oct;18:9-15. [doi: 10.1016/j.jcbs.2020.08.001]
33. Sanduvete-Chaves S, Chacón-Moscoso S, Cano-García FJ. Effectiveness of psychological interventions to decrease cognitive fusion in patients with chronic pain: a systematic review and meta-analysis. J Psychosom Res 2024 Nov;186:111888 [FREE Full text] [doi: 10.1016/j.jpsychores.2024.111888] [Medline: 39213942]
34. Kállai J, Rózsa S, Hupuczi E, Hargitai R, Birkás B, Hartung I, et al. Cognitive fusion and affective isolation: blurred self-concept and empathy deficits in schizotypy. Psychiatry Res 2019 Jan;271:178-186. [doi: 10.1016/j.psychres.2018.11.045] [Medline: 30481696]
35. Barrera-Caballero S, Romero-Moreno R, Márquez-González M, Jiménez-Gonzalo L, Huertas-Domingo C, Olazarán J, et al. Longitudinal effects of cognitive fusion in depressive and anxious symptoms of family caregivers of people with dementia. Journal of Contextual Behavioral Science 2024 Jul;33:100782. [doi: 10.1016/j.jcbs.2024.100782]
36. Ford N. Modeling cognitive processes in information seeking: from Popper to Pask. J. Am. Soc. Inf. Sci 2004 Feb 23;55(9):769-782. [doi: 10.1002/asi.20021]
37. McElroy E, Shevlin M. The development and initial validation of the cyberchondria severity scale (CSS). J Anxiety Disord 2014 Mar;28(2):259-265. [doi: 10.1016/j.janxdis.2013.12.007] [Medline: 24508033]
38. Fergus TA, Russell LH. Does cyberchondria overlap with health anxiety and obsessive-compulsive symptoms? An examination of latent structure and scale interrelations. J Anxiety Disord 2016 Mar;38:88-94. [doi: 10.1016/j.janxdis.2016.01.009] [Medline: 26851749]

XSL•FO
RenderX

39. Baumgartner SE, Hartmann T. The role of health anxiety in online health information search. Cyberpsychol Behav Soc Netw 2011 Oct;14(10):613-618. [doi: 10.1089/cyber.2010.0425] [Medline: 21548797]
40. Bardeen JR, Fergus TA. The interactive effect of cognitive fusion and experiential avoidance on anxiety, depression, stress and posttraumatic stress symptoms. Journal of Contextual Behavioral Science 2016 Jan;5(1):1-6. [doi: 10.1016/j.jcbs.2016.02.002]
41. Hayes SC, Levin ME, Plumb-Vilardaga J, Villatte JL, Pistorello J. Acceptance and commitment therapy and contextual behavioral science: examining the progress of a distinctive model of behavioral and cognitive therapy. Behav Ther 2013 Jun;44(2):180-198 [FREE Full text] [doi: 10.1016/j.beth.2009.08.002] [Medline: 23611068]
42. Kashdan TB, Rottenberg J. Psychological flexibility as a fundamental aspect of health. Clin Psychol Rev 2010 Nov;30(7):865-878 [FREE Full text] [doi: 10.1016/j.cpr.2010.03.001] [Medline: 21151705]
43. Wang D, Sun L, Shao Y, Zhang X, Maguire P, Hu Y. Research and evaluation of a cyberchondria severity scale in a Chinese context. PRBM 2023 Nov;Volume 16:4417-4429. [doi: 10.2147/prbm.s431470]
44. Zhang WC, Ji Y, Li X, Guo HN, Zhu ZH. Reliability and validity of the Chinese version of the Cognitive Fusion Questionnaire. Chinese Mental Health Journal 2014;28(1):40-44 [FREE Full text] [doi: 10.3969/j.issn.1000-6729.2014.01.007]
45. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736 [FREE Full text] [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
46. Luo N, Liu G, Li M, Guan H, Jin X, Rand-Hendriksen K. Estimating an EQ-5D-5L value set for China. Value Health 2017 Apr;20(4):662-669 [FREE Full text] [doi: 10.1016/j.jval.2016.11.016] [Medline: 28408009]
47. Wellbeing measures in primary health care/the DepCare Project: report on a WHO meeting: Stockholm, Sweden, 12–13 February 1998. World Health Organization. 1998. URL: https://iris.who.int/handle/10665/349766 [accessed 2025-05-14]
48. Fung S, Kong CYW, Liu Y, Huang Q, Xiong Z, Jiang Z, et al. Validity and psychometric evaluation of the Chinese version of the 5-Item WHO Well-Being Index. Front Public Health 2022 Mar 30;10:872436 [FREE Full text] [doi: 10.3389/fpubh.2022.872436] [Medline: 35433612]
49. Cohen J. A power primer. Psychological Bulletin 1992;112(1):155-159. [doi: 10.1037/0033-2909.112.1.155]
50. Bollen KA. Structural Equation Models. In: Armitage P, Colton T, editors. Encyclopedia of Biostatistics. Hoboken, NJ: John Wiley & Sons, Ltd; 2005:2005.
51. Kline RB. Principles and practices of structural equation modelling. New York, NY: Guilford Press; 2015.
52. R Core Team. R: A language and environment for statistical computing. R foundation for statistical computing. 2021. URL: https://www.R-project.org/ [accessed 2025-05-14]
53. Hallit S, Rogoza R, Abi Semaan C, Azzi V, Sawma T, Obeid S. Validation of the Arabic version of the cyberchondria severity scale 12 items (CSS-12-Ar) among a sample of Lebanese adults. BMC Psychiatry 2023 Aug 23;23(1):618. [doi: 10.1186/s12888-023-05123-x] [Medline: 37612590]
54. Robles-Mariños R, Alvarado GF, Maguiña JL, Bazo-Alvarez JC. The short-form of the Cyberchondria Severity Scale (CSS-12): adaptation and validation of the Spanish version in young Peruvian students. PLoS One 2023 Oct 5;18(10):e0292459 [FREE Full text] [doi: 10.1371/journal.pone.0292459] [Medline: 37796833]
55. Barke A, Bleichhardt G, Rief W, Doering BK. The Cyberchondria Severity Scale (CSS): German validation and development of a short form. Int J Behav Med 2016 Oct 1;23(5):595-605. [doi: 10.1007/s12529-016-9549-8] [Medline: 26931780]
56. McElroy E, Kearney M, Touhey J, Evans J, Cooke Y, Shevlin M. The CSS-12: development and validation of a short-form version of the Cyberchondria Severity Scale. Cyberpsychol Behav Soc Netw 2019 May;22(5):330-335. [doi: 10.1089/cyber.2018.0624] [Medline: 31013440]
57. Xu RH. Estimation of cut-off value for the Cyberchondria Severity Scale-12 in Chinese population. Research Square Preprint posted online on August 05, 2024. [doi: 10.21203/RS.3.RS-4721206/v1]
58. Wang Q, Zhang W, An S. A systematic review and meta-analysis of Internet-based self-help interventions for mental health among adolescents and college students. Internet Interv 2023 Dec;34:100690 [FREE Full text] [doi: 10.1016/j.invent.2023.100690] [Medline: 38023965]
59. Kwak Y, Kim H, Ahn J. Impact of internet usage time on mental health in adolescents: using the 14th Korea Youth Risk Behavior Web-Based Survey 2018. PLoS One 2022 Mar 23;17(3):e0264948 [FREE Full text] [doi: 10.1371/journal.pone.0264948] [Medline: 35320295]
60. Köse S, Murat M. Examination of the relationship between smartphone addiction and cyberchondria in adolescents. Arch Psychiatr Nurs 2021 Dec;35(6):563-570. [doi: 10.1016/j.apnu.2021.08.009] [Medline: 34861946]
61. Shi Y, Ma D, Zhang J, Chen B. In the digital age: a systematic literature review of the e-health literacy and influencing factors among Chinese older adults. Z Gesundh Wiss 2023 Jun 04;31(5):679-687 [FREE Full text] [doi: 10.1007/s10389-021-01604-z] [Medline: 34104627]
62. McMaughan DJ, Oloruntoba O, Smith ML. Socioeconomic status and access to healthcare: interrelated drivers for healthy aging. Front Public Health 2020 Jun 18;8:231 [FREE Full text] [doi: 10.3389/fpubh.2020.00231] [Medline: 32626678]
63. Golden J, Conroy RM, Lawlor BA. Social support network structure in older people: underlying dimensions and association with psychological and physical health. Psychol Health Med 2009 May 14;14(3):280-290. [doi: 10.1080/13548500902730135] [Medline: 19444706]

XSL•FO
RenderX

64. Donovan NJ, Blazer D. Social isolation and loneliness in older adults: review and commentary of a National Academies report. Am J Geriatr Psychiatry 2020 Dec;28(12):1233-1244 [FREE Full text] [doi: 10.1016/j.jagp.2020.08.005] [Medline: 32919873]
65. Wang M, Pan Q. The rural-urban differences and influencing factors in the anxiety symptoms of Chinese elderly people. Chinese General Practice 2021;24(31):3963-3970. [doi: 10.12114/j.issn.1007-9572.2021.00.294]
66. Fan S, Yang Y. How does internet use improve mental health among middle-aged and elderly people in rural areas in China? A quasi-natural experiment based on the China Health and Retirement Longitudinal Study (CHARLS). Int J Environ Res Public Health 2022 Oct 16;19(20):13332 [FREE Full text] [doi: 10.3390/ijerph192013332] [Medline: 36293913]
67. Zhou X, Chen L. Digital health care in China and access for older people. The Lancet Public Health 2021 Dec;6(12):e873-e874. [doi: 10.1016/s2468-2667(21)00051-7]
68. Dai B. The old age health security in rural China: where to go? Int J Equity Health 2015 Nov 04;14(1):119 [FREE Full text] [doi: 10.1186/s12939-015-0224-5] [Medline: 26530528]
69. Luque-Reca O, Gillanders D, Catala P, Peñacoba C. Psychometric properties of the Cognitive Fusion Questionnaire in females with fibromyalgia. Curr Psychol 2021 Aug 30;42(11):9485-9495. [doi: 10.1007/s12144-021-02214-4]
70. Chen N, Xi J, Fan X. Correlations among psychological resilience, cognitive fusion, and depressed emotions in patients with depression. Behav Sci (Basel) 2023 Jan 25;13(2):100 [FREE Full text] [doi: 10.3390/bs13020100] [Medline: 36829329]
71. Hu Z, Yu H, Zou J, Zhang Y, Lu Z, Hu M. Relationship among self-injury, experiential avoidance, cognitive fusion, anxiety, and depression in Chinese adolescent patients with nonsuicidal self-injury. Brain Behav 2021 Dec;11(12):e2419 [FREE Full text] [doi: 10.1002/brb3.2419] [Medline: 34816613]
72. Al-Zoubi EM. Quality of life and its relationship with cognitive flexibility among higher education students. Journal of Educational and Social Research 2020 Jul 10;10(4):156 [FREE Full text] [doi: 10.36941/jesr-2020-0075]
73. Hart C, Da Costa C, D'Souza D, Kimpton A, Ljbusic J. Exploring higher education students' critical thinking skills through content analysis. Thinking Skills and Creativity 2021 Sep;41:100877. [doi: 10.1016/j.tsc.2021.100877]
74. Kanbay Y, Okanlı A. The effect of critical thinking education on nursing students' problem-solving skills. Contemp Nurse 2017 Jun;53(3):313-321. [doi: 10.1080/10376178.2017.1339567] [Medline: 28589781]
75. Faustino B, Vasco AB, Farinha-Fernandes A, Delgado J. Psychological inflexibility as a transdiagnostic construct: relationships between cognitive fusion, psychological well-being and symptomatology. Curr Psychol 2021 Jun 04;42(8):6056-6061. [doi: 10.1007/s12144-021-01943-w]
76. Fernández-Rodríguez C, Coto-Lesmes R, Martínez-Loredo V, Cuesta-Izquierdo M. Psychological inflexibility, anxiety and depression: the moderating role of cognitive fusion, experiential avoidance and activation. Psicothema 2022 May 01;34(2):240-248. [doi: 10.7334/psicothema2021.344] [Medline: 35485537]
77. Ferreira C, Mendes A, Marta-Simões J, Trindade I. The impact of early affiliative memories on psychological health: the mediator role of external shame and cognitive fusion. Eur. psychiatr 2020 Mar 23;33(S1):S518-S518. [doi: 10.1016/j.eurpsy.2016.01.1916]
78. Soltani E, Bahrainian SA, Farhoudian A, Masjedi Arani A, Gachkar L. Effectiveness of acceptance commitment therapy in social anxiety disorder: application of a longitudinal method to evaluate the mediating role of acceptance, cognitive fusion, and values. Basic Clin Neurosci 2023 Jul 01;14(4):479-490 [FREE Full text] [doi: 10.32598/bcn.2021.2785.1] [Medline: 38050569]
79. McMullan RD, Berle D, Arnáez S, Starcevic V. The relationships between health anxiety, online health information seeking, and cyberchondria: systematic review and meta-analysis. J Affect Disord 2019 Feb 15;245:270-278. [doi: 10.1016/j.jad.2018.11.037] [Medline: 30419526]
80. Bansal N, Choudhary H. Growing old in the digital era: a qualitative study of internet use and outcomes among urban Indian older adults. WWOP 2023 Oct 02;28(4):317-335. [doi: 10.1108/wwop-08-2023-0032]
81. Maftei A, Holman A, Merlici I. Using fake news as means of cyber-bullying: the link with compulsive internet use and online moral disengagement. Computers in Human Behavior 2022 Feb;127:107032. [doi: 10.1016/j.chb.2021.107032]

Abbreviations

CAPTCHA: Completely Automated Public Turing test to tell Computers and Humans Apart
CFQ: Cognitive Fusion Questionnaire
CSS-12: Cyberchondria Severity Scale-12
H1: hypothesis 1
H2: hypothesis 2
H3: hypothesis 3
H4: hypothesis 4
H5: hypothesis 5
SEM: structural equation modeling
WHO-5: World Health Organization-5 Well-Being Index

XSL•FO
RenderX

Edited by J Wang; submitted 19.12.24; peer-reviewed by S Hill, S Bidmon; comments to author 06.02.25; revised version received 01.03.25; accepted 12.05.25; published 21.05.25.

Please cite as:
Xu RH, Starcevic V
Cyberchondria in Older Adults and Its Relationship With Cognitive Fusion, Health-Related Quality of Life, and Mental Well-Being: Mediation Analysis
JMIR Aging 2025;8:e70302
URL: https://aging.jmir.org/2025/1/e70302
doi:10.2196/70302
PMID:

©Richard Huan Xu, Vladan Starcevic. Originally published in JMIR Aging (https://aging.jmir.org), 21.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Technology-Assisted Physical Activity Interventions for Older People in Their Home-Based Environment: Scoping Review

Rosemary Dubbeldam[1*], PhD; Rafal Stemplewski[2*], PhD; Iuliia Pavlova[3*], Prof. PhD; Magdalena Cyma-Wejchenig[2*], PhD; Sunwoo Lee[4*], PhD; Patrick Esser[5*], PhD; Ellen Bentlage[1*], PhD; Veysel Alcan[6*], PhD; Özge Selin Çevik[7*], PhD; Eleni Epiphaniou[8], PhD; Francesca Gallè[9], PhD; Antoine Langeard[10], PhD; Simone Gafner[11], PhD; Mona Ahmed[1], PhD; Niharika Bandaru[12], MSc; Arzu Erden Güner[13], PhD; Evrim Göz[14], PhD; Ilke Kara[15], MSc; Ayşe Kabuk[16], PhD; Ilayda Türkoglu[17], PhD; Zada Pajalic[18], Prof. PhD; Jan Vindiš[4], PhD; Damjan Jaksic[19], PhD; U ur Verep[15], PhD; Ioanna Chouvarda[20], PhD; Vera Simovska[21], Prof. PhD; Yael Netz[22*], Prof. PhD; Jana Pelclova[4*], Prof. PhD

[1]Department of Movement Science, Institute for Sport and Exercise Sciences, University Münster, Horstmarer Landweg 62B, Münster, Germany

[10]Normandy University, UNICAEN, INSERM, COMETE, CYCERON, Caen, France

[11]School of Health Sciences, HES-SO Valais-Wallis, Leukerbad, Switzerland

[12]Institut of Sport Science, Otto-Von Guericke Universität of Magdeburg, Magdeburg, Germany

[13]Department of Physiotherapy and Rehabilitation, Faculty of Health Sciences, Karadeniz Technical University, Trabzon, Turkey

[14]Department of Physiotherapy and Rehabilitation, Faculty of Health Sciences, Tarsus University, Tarsus, Turkey

[15]Department of Physical Therapy and Rehabilitation, Institute of Health Sciences, Dokuz Eylul University, Izmir, Turkey

[16]Faculty of Health Sciences, Zonguldak Bülent Ecevit University, Zonguldak, Turkey

[17]Department of Fundamentals of Nursing, Hamidiye Faculty of Nursing, Health Sciences University, İstanbul, Turkey

[18]Faculty of Health and Social Sciences, University of South-Eastern Norway, Kongsberg, Norway

[19]Faculty of Sport and Physical Education, University of Novi Sad, Novi Sad, Serbia

[2]Department of Digital Technologies in Physical Activity, University of Physical Education, Poznan, Poland

[20]Lab of Computing, Medical Informatics and Biomedical-Imaging Technologies, School of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece

[21]HEPA Macedonia National Organisation for the Promotion of Health-Enhancing Physical Activity, Skopje, North Macedonia

[22]The Levinsky-Wingate Academic Center, Wingate Institute, Netanya, Israel

[3]Department of Theory and Methods of Physical Culture, Ivan Boberskyj Lviv State University of Physical Culture, Lviv, Ukraine

[4]Faculty of Physical Culture, Palacký University Olomouc, Olomouc, Czech Republic

[5]Centre for Movement, Occupation and Rehabilitation Sciences, Oxford Brookes University, Oxford, United Kingdom

[6]Department of Electrical and Electronics Engineering, Faculty of Engineering, Tarsus University, Tarsus, Turkey

[7]Department of Physiology, Faculty of Medicine, Mersin University, Mersin, Turkey

[8]School of Humanities, Social and Education Sciences, European University Cyprus, Engomi, Cyprus

[9]Department of Medical, Movement and Wellbeing Sciences, University of Naples Parthenope, Naples, Italy

[*]these authors contributed equally

Corresponding Author:
Rosemary Dubbeldam, PhD
Department of Movement Science, Institute for Sport and Exercise Sciences, University Münster, Horstmarer Landweg 62B, Münster, Germany

Abstract

Background: Technology-assisted physical activity interventions for older adults in their home-based environment have been used to promote physical activity. Previous research has reported that such interventions benefit body composition, aerobic fitness, cognitive abilities, and postural control, reducing the risk of falls and maintaining regular physical activity among the older population.

Objective: While previous reviews on technology-assisted physical activity interventions focused on health-related outcomes, this scoping review explores the characteristics of the technology in relation to the characteristics of populations, purpose of the activity, and usability in terms of adverse events, drop-outs, adherence, and user experience.

XSL•FO
RenderX

Methods: A full search was performed in Medline, Embase, CINAHL, SportDiscus, and Web of Science. Sources were considered for inclusion if the participants aged on average 60 years and older, if the physical activity intervention was assisted by technology, and if performed within home-based environments.

Results: We identified 8496 sources. After title and abstract screening, 455 full texts were assessed, and 148 were included, representing 12,717 participants aged 74 (SD 6) years. In total, 63% (93/148) of the sources reported on the population's health status. The main purpose of the interventions was balance (75/148, 51%), and strength and power (64/148, 43%) and the intervention purposes were not related to the embedded technology. In studies where the participant's health status was reported as healthy, 53% (78/148) implemented exergames compared to only 27% (40/148) in studies with participants with a clinical condition. Mobile apps (30/148, 20%) and trackers (16/148, 11%) were implemented likewise in both groups. The technology was embedded to provide continuous exercise information (40/148, 27%) and exercise feedback (40/148, 27%) or to record real-time movement data (38/148, 26%). Adverse events were reported in 46% (68/148) of the sources with three quarters (49/68) reporting no adverse events. Only two mild events were related to technology. Dropout rates were reported in 68% (100/148) of the sources, with no differences between intervention (16 SD 16%) and control (14 SD 12%) groups. Dropout reasons related to technology are only 3% (3/100). Adherence was reported in 53% (79/148) sources and was slightly higher in the intervention group (80 SD 18%) compared to the control group (71 SD 25%). A significantly higher adherence was found between interventions that were tailored (83 SD 15%) versus those that were not (75 SD 21%). General enjoyment of the technology was captured in 37% (55/148) of the sources. Within those sources, 91% rated positive (91/100), 7% neutral (7/100), and 2% negative (2/100). Occasionally reported wishes were related to goal setting, feedback, technical support, exercise variation, and social setting.

Conclusions: Various technologies were successfully used in healthy and clinical older populations, though evidence regarding their implementation in physical activity interventions in hospital settings remains limited. The embedded technology was not a reason for additional dropouts, led to slightly better adherence, and adverse events were rarely related to technology. When assessed, the technology was well accepted and positively enjoyed.

Trial Registration: OSF Registries OSF.IO/NVECQ; https://doi.org/10.17605/OSF.IO/B2NGD

(JMIR Aging 2025;8:e65746) doi:10.2196/65746

KEYWORDS

technology-assisted interventions; technological interventions; physical activity; physical exercise; exercise; gerontology; geriatric; older adult; elderly; older people; aging; home-based; dropout; adherence; adverse effects; scoping review; review

Introduction

The benefits of physical activity in advanced age are well-documented, yet paradoxically, sedentary behavior increases with age [1,2]. Even if individuals are motivated to engage in physical activity outside the home, environmental barriers such as weather conditions, infrastructure constraints, or specific constraints such as COVID-19 can pose challenges for older adults [3,4]. This paradox has prompted clinicians to seek technological solutions to encourage older adults to be more physically active within their home environment. It is important to distinguish between physical activity and exercise. Physical activity encompasses any bodily movement that expends energy, while exercise is a structured subset of physical activity specifically designed to improve fitness components [5].

Various electronic devices, software, or wearable technology such as smartphones [6], fitness trackers [7], mobile phone applications [8], tablets [9], virtual reality [10], exergames [11], E-coaching [12], and Robotics [13] have been reported to support beneficial changes in sedentary time and physical activity in the older adult population [14,15]. Such technological interventions benefit body composition, aerobic fitness, cognitive abilities, and postural control and reduce the risk of falls among the older population [9,16,17]. However, while there are indications that older adults tend to show high levels of adherence to technology-assisted physical activity interventions [15], quite a few barriers limiting the use of this technology have been reported, such as functional barriers related to the design of e-health programs and their interface with older end users, or lack of technological support [18].

With the continuously increasing number of technology-assisted healthcare tools, particularly those for home-based physical activity, and the difficulties older adults face in adapting to and operating these tools, it is not surprising that several reviews have been conducted in recent years to examine the usability of technology-assisted physical activity programs [15,19-25], more than half of them in recent years – 2021 - 2024 [19,20,22,23,25]. Most of these reviews focused on a specific population, egfor example, healthy, frail, or people with cardiovascular disease, focused on specific aspects of interventions, for example, balance or falls, or focused on a specific type of technology, for example, exergames [15,19-25]. Considering the boost of assistive-technology provision of physical activity in the recent five years, some are relatively dated [15,24]. A recent review by Costa-Brito et al elaborated on the usage of technology to improve physical function among older community-dwelling adults [23] and focused on the purpose and volume of an activity, types of technology, aspects of the usage of technology (adherence, acceptance, feasibility), and health outcomes. The review, however, did not go into the specific characteristics of the study participants, the technology's functionality, or its user interface. Furthermore, it did not explore or analyze key aspects of technology design, function, and interface in relation to specific populations or physical activity goals. Just as many other reviews [19-21,25],

XSL•FO
RenderX

Costa-Brito et al also neglected to examine safety concerns or adverse events associated with the use of technology.

Aligned with the recommendations of the PhysAgeNet, a European Cooperation in Science and Technology (COST) network devoted to evidence-based physical activity in old age [26] and with the proliferation of technology related to home-based physical activity, our aim in this study was to explore the characteristics of that technology in relation to the characteristics of older populations (eg, age group or health status) and the interventions. Furthermore, we aimed to examine the usability of that technology in terms of adverse events, drop-outs, adherence, and the user experience (acceptability and enjoyment). This may promote the optimal utilization of technology to provide efficient home-based physical activity programs for older adults. A scoping review was selected as the most suitable methodology due to its capacity to encompass a broad spectrum of study designs and methodologies. This approach facilitates the mapping of key concepts and the identification of research gaps, which is particularly critical in this emerging field characterized by diverse evidence sources, including randomized controlled trials, pre-post studies, feasibility studies, qualitative research, and mixed methods investigations.

Methods

Review Process

This scoping review was conducted in accordance with the Joanna Briggs Institute (JBI) methodology for scoping reviews [26] and the Preferred Reporting Items for Systematic Reviews and Meta-analyses extension for scoping reviews (PRISMA-ScR) checklist [27]. The protocol for the scoping review is available at the Open Science Framework Registries [28]. The review was conducted by an international interdisciplinary team of researchers with regular weekly meetings to discuss, distribute tasks, and ensure consistency during the different steps of the process.

Eligibility

Sources were considered for inclusion if they met the following criteria: (1) participants aged on average 60 years and older with and without health conditions; (2) studies that focused on technology-assisted physical activity interventions; and (3) the intervention was based within home-based environments. Table 1 gives a detailed definition of the participants, concept, and context.

Table . Details of the participants, concept, and context (PCC framework) which form the basis for source inclusion consideration.

PCC	Inclusion	Definitions and descriptions
Population	Older people	Older adults aged 60 years and older (the participant's group mean age should be ≥60 y), living independently or in residential facilities (aged care facilities), with or without health conditions.
Concept	Technology-assisted physical activity interventions	A physical intervention is a structured or unstructured measure taken to improve or maintain the user's physical activity and practiced individually. It can be tailored to specific health or fitness goals (structured) or it can be general and not specifically prescribed by a health professional (unstructured) [29]. The intervention includes any form of physical activity such as human daily activities, exercise, training, or physical fitness training. The technologies used to deliver the intervention are digital systems and include, among others, mobile applications, telemedicine, wearable electronic devices, exergaming, and virtual reality. The delivery of the intervention can take place supervised, facilitated (partly supervised), or unsupervised.
Context	Home-based environment	The home-based location and environment, ie, the immediate vicinity in the community environment, ie, a public open-access setting [29]. The participants may be living in a community setting on their own (community dwelling) or in an institution such as a care home or a hospital.
	Health status	People with or without mental or physical impairments.

Sources were excluded if (1) the study design was a review or study protocol, or if the publication type was an abstract or non-peer-reviewed conference proceedings, (2) participants were on average younger than 60 years of age, (3) no physical activity intervention was present or the intervention took place within clinical setting, (4) digital technology was not present or not an integral part of the intervention, (5) the intervention did not take place in home-based setting, or (6) the study language was other than English.

XSL•FO
RenderX

Search Strategy

According to JBI recommendation, an initial limited search of PubMed and Web of Science was conducted to identify potentially relevant key search terms for the PCC framework. In an iterative process, the keywords in the text of retrieved sources were analyzed. The search terms provided were consulted with experts within the team to capture all relevant keywords and then grouped into four main headings: population group, residence, technology, and intervention. The search string was further expanded by running searches with Medical Subject Heading (MeSH Terms) as well as with adequate non-MeSH terms for databases that do (eg, PubMed) or do not use a MeSH tree (eg, Web of Science), respectively and cross-checked by a subject librarian at Oxford Brookes University: ("aged" OR "elder*" OR "old" OR "older people" OR "geriatric" OR "senior#") AND ("home*" OR "resident*" OR "care facilit*" OR "nursing home*") AND ("digital technolog*" OR "mobile health" OR "telehealth" OR "eHealth" OR "mHealth" OR "exergam*" OR "video gam*" OR "serious gam*" OR "wearable technolog*" OR "wearable electronic device*" OR "application*" OR "mobile app*" OR "virtual reality") AND ("exercise#" OR "training" OR "intervention*" OR "physical activit*" OR "physical fitness") NOT ("review"[Title/Abstract] OR "rehabilitation" [Title/Abstract])

The created search strings were the basis for developing a full search strategy in Medline (PubMed), Embase (Ovid), CINAHL (Ebsco), SportDiscus (Ebsco), and Web of Science and converted as necessary (Multimedia Appendix 1). The search in the different databases was performed between September 26 and October 6, 2022, and all sources published in previous years were included. As a third step of the search strategy, an additional search was performed in the reference list of the identified sources selected from the full text search.

Source of Evidence Selection

Following the search, all identified sources were collated and uploaded into Rayyan [30], and duplicates were removed. Two reviewers performed a pre-screening pilot on 100 titles and abstracts to develop screening instructions (total agreement=95%, Cohen's Kappa=0.701). Any disagreements were resolved via consensus with the help of a third reviewer. After the whole team approved the final version of the screening instruction tool, the 2 reviewers screened the remaining titles and abstracts (total agreement=95.5%, Cohen's Kappa=0.515). Again, disagreements were resolved by a third reviewer.

Full texts of potentially eligible sources were uploaded into Rayyan. Ten different pairs of reviewers assessed these sources (total agreement =81%, Cohen's Kappa=0.696) within Rayyan [30]. Disagreements were resolved via consensus in a team of five other (independent) reviewers. The results of the search, the sources' inclusion process, and exclusion reasons have been reported in full and presented according to the PRISMA flow diagram [27]. All sources obtained after the search process were stored in Mendeley (Mendeley Ltd., Elsevier).

Data Extraction

A data extraction tool (Multimedia Appendix 2) was made in line with the JBI guidelines [26,31] and a template for intervention description and replication [32]. Five reviewers piloted the draft extraction tool, topic explanations were added to ensure the correct data was extracted into relevant columns and a final data extraction tool was prepared after which a group analysis and discussion was conducted, making sure that each extractor understood the tool. Finally, ten pairs of reviewers independently completed data extractions. When preparing the data for the analysis, a third reviewer cross-checked the extracted data and, where necessary, consulted the full text for details. Discrepancies and uncertainties were discussed within the team.

Data Analysis and Presentation

Four teams consisting of 3 - 5 reviewers analyzed the extracted data regarding the following topics: study design, participants, interventions and outcome measures, and technologies and key findings. An inductive thematic analysis was conducted. Codes or labels were defined to categorize the data and a coding framework with definitions of the categories and subcategories was developed. This coding framework allowed us to characterize the participants (age, health status, mobility, country of residence, etc), intervention (purpose, tailoring, temporal data, etc), the used technologies (design, function, interface), participation (feasibility, effectiveness, safety, etc), and health outcomes (physical and mental capacities, fall risk, quality of life, etc). A detailed description of the coding framework is presented in Multimedia Appendix 3. During the extracted data categorization process, the reviewers came together regularly in their small teams to discuss discrepancies until consensus was achieved and all data categorized. For this review, the focus lies with the participants, intervention, and especially the technology categories.

The quantitative and qualitative data were analyzed descriptively using absolute and relative frequency, percentage, average, and standard deviation. Pie charts and bar charts were made to visualize results.

Crosstabulation statistical analysis was used to get a better understanding of what kind of digital technologies were used to integrate specific physical activity intervention programs and for which populations. Observations between 2 or more nominal variables were explored, whereby the frequency of observations between variables was counted and reported accordingly and expressed as a percentage against all observations. Specifically, the cross-tabulation analysis was conducted between the following categorical variables: Technology (design, function, interface) versus Population (age group, health status, residency) and Intervention (aim of physical activity, tailoring).

Adherence (attendance and duration) and dropout rates were expressed as a percentage of the relevant study population and compared against population and intervention characteristics (2 or more categories) using differential tests. In the case of a binomial category, an independent *t* test was used. The following categories and linear outcomes were compared: Participation outcomes (dropouts, adherence) versus Population (age groups, health status) and Intervention (group, tailoring, supervision).

Relationships between linear parameters were regressed to extract the coefficient of determination (R^2) with associated *P*

XSL•FO
RenderX

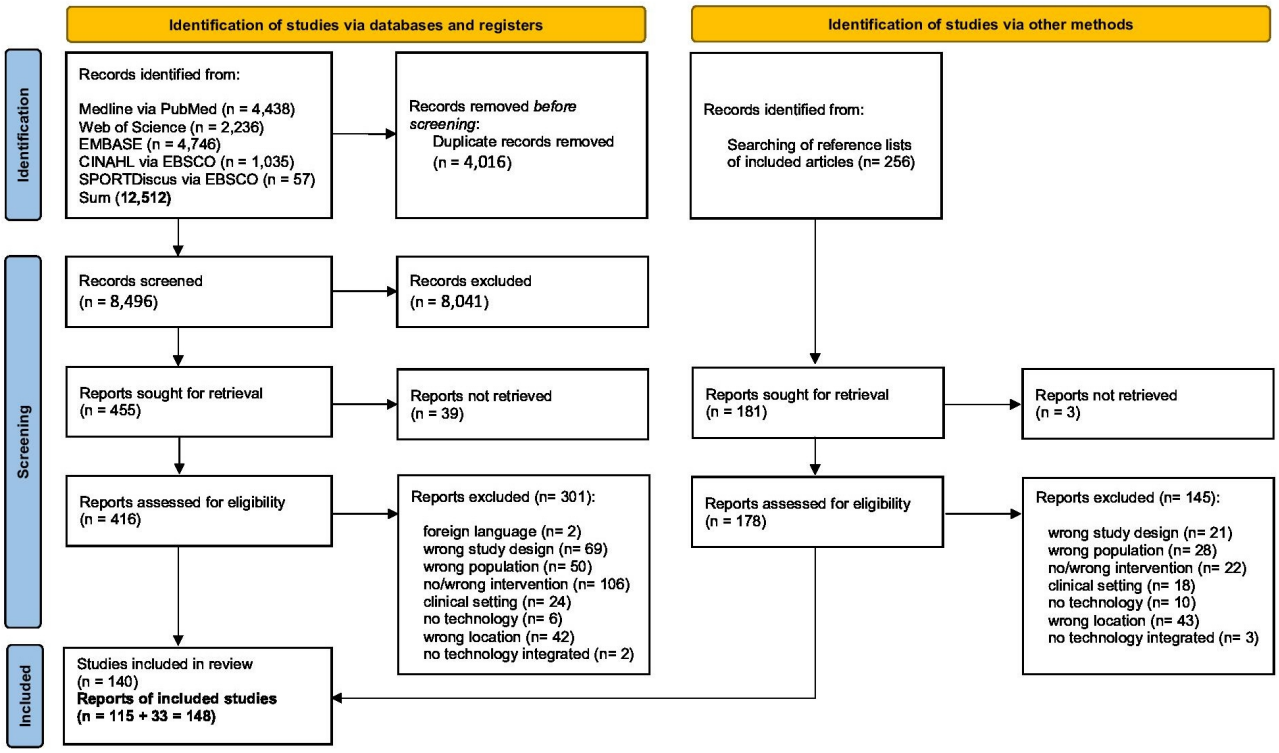
value to indicate its significance. These were tested for the dropout and adherence outcomes, against the total number of intervention sessions and session durations.

Results

Source Selection

A total of 12,512 sources were identified during searching in Medline (PubMed), Embase (Ovid), CINAHL (EBSCO), SportDiscus (EBSCO), and Web of Science databases. After screening, 148 sources remained. Inclusion and exclusion details can be found in the PRISMA flowchart (Figure 1). An overview of all excluded sources with reasons and all included sources and corresponding extracted data can be found in the supplementary material. Detailed results from the data analysis and the references to the 148 included sources can be found in Multimedia Appendix 4 [33-175]. The full data extraction table and reasons for source exclusions from this review can be found in Multimedia Appendices 5 and 6, respectively.

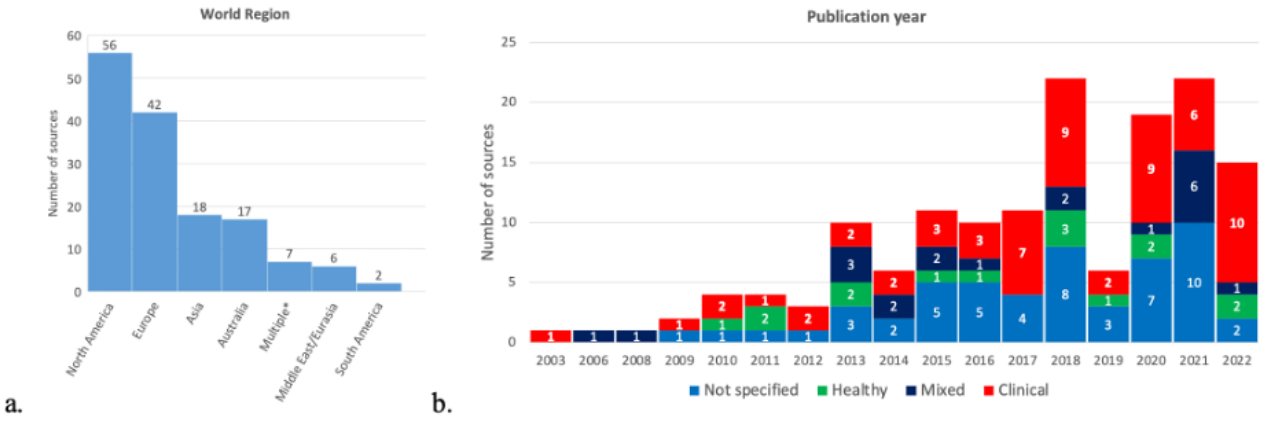
Figure 1. Flowchart search process.



Study Characteristics

Publication years ranged from 2003 to 2022, with most of the included sources published in the last five years (84/148, 57%). An overview of the number of sources performed by region and publication year is presented in Figure 2.

Figure 2. (**A**) Number of included sources by world region. (**B**) and by publication year and health status.



The sources included 51% (75/148) randomized controlled trials, 24% (35/148) pre-post-test, and 12% (17/148) feasibility studies and to a lesser extent qualitative studies (6/148, 4%), non-randomized controlled trials (5/148, 3%), experimental studies (5/148, 3%), mixed methods (4/148, 3%), and one case study (1/148, 1%).

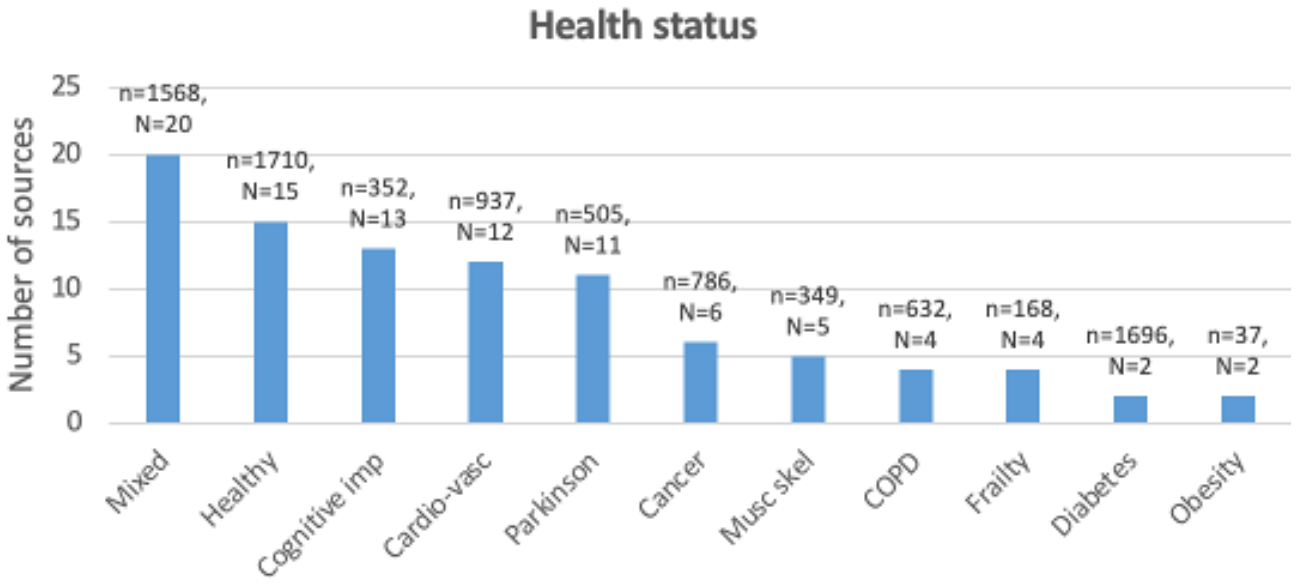
XSL•FO
RenderX

Participant Characteristics

The included sources contained a total sample of 12,717 participants, aged 74 (SD 6) years on average, with a range of 60 - 88 years. Most sources were devoted to the youngest-old group of 60 to 74 years (87/148, 59%), followed by the middle-old group of 75 to 84 years (51/148, 35%) and the oldest-old group of 85 years and older (6/148, 4%). A small sample of sources reported on mixed (2/148, 1%) age categories. In 2 studies, the age of the participants was not reported. With increasing age groups, the percentage of women increased from 53% (5160/9802) (youngest-old) to 70% (1750/2510) (middle-old) and 79% (119/150) (old-old).

A larger focus on the clinical population is visible since 2017, although the health status is not always properly reported (Figure 2B). In total, 63% (93/148) of the sources reported on the population's health status (Figure 3). In 37% (55/148) of the sources, information on comorbid conditions was indicated, the average number of comorbidities was 4.2 (SD 2.6; range 1 - 10).

Figure 3. Number of included sources (N) by reported health status. With n being the total number of participants with cardiovascular disease (Cardio-vasc); cognitive impairments (Cognitive imp); chronic obstructive pulmonary disease (COPD) and musculoskeletal disease (Musc skel).



Most sources (101/148, 68%) included participants living in community-dwelling residences, 26% of the sources included participants from care homes, while 13% (19/148) of the sources were not specific on participant living arrangements (Figure 3). No source reported an intervention in a hospital setting.

Prior experience with technology was stated in 16% (24/148) of sources, and 15% (22/148) of sources stated that participants had access to technology at home before the implementation of the intervention. Five percent of the sources (7/148) highlight information about the prior frequency of use of these services or devices.

Intervention Design Characteristics

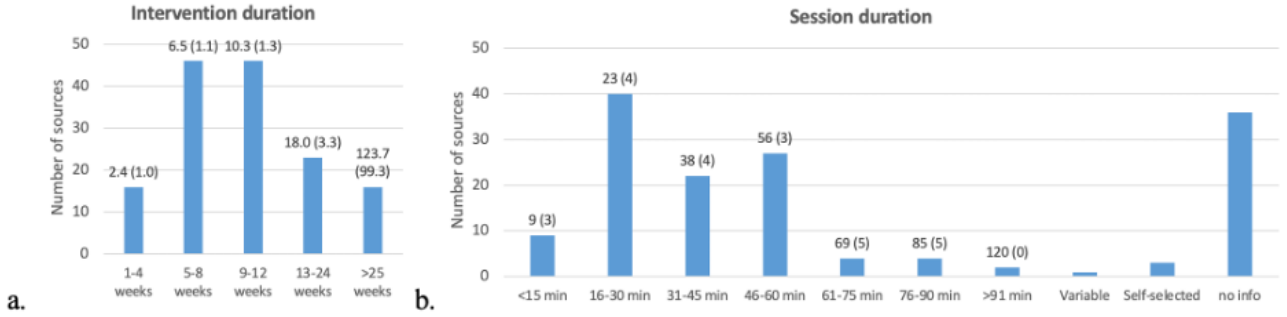
In 44% (65/148) of the sources, participants received supervised interventions (professional or non-professional) with a further 55% (82/148) of the sources including unsupervised interventions. Additionally, one source tested supervised versus unsupervised technology-based interventions.

In most of the sources (93/148, 63%), the intervention was tailored to individuals and/or groups of participants. The other 37% (55/148) applied a generic intervention approach to all participants. In the tailored interventions, the physical activity was adapted to the participant's specific needs, or the progression of the activity was personalized by adapting the difficulty, intensity, or frequency level of the activity. Furthermore, individual assessments and feedback systems allowed for continuous adjustment of activities based on performance.

Intervention durations varied from one week up to 288 weeks (Figure 4A). A total of 2 interventions (2/148, 1%) did not report intervention duration. Session duration varied by study from 5 minutes up to 120 minutes (Figure 4B). Session frequency was reported in 119 studies and ranged between one to 21 per week. In 2 studies, the weekly session frequency could be chosen independently. Most sources (87/148, 59%) ranged from one to three sessions per week.

XSL•FO
RenderX

Figure 4. Number of sources stratified by (A). intervention duration (average and SD) and (B) session duration (average and SD). For intervention duration in minutes (min), the average duration and corresponding standard deviation per group is also presented.



The activity intensity was described in only 14% (20/148) of the sources. Of these, 3% (5/148) used Rate of Perceived Exertion, 2% (3/148) the Borg scale, with a further 6% (9/148) of the sources reported intensity in a different manner, eg, the level of the game. Only one source (1/148, 1%) set the intensity at 60% of Heart Rate Reserve, with another 2 sources (2/148,1%) using self-pacing of intensity.

In total, 2 sources (2/148, 1%) did not report on the physical activity used as part of the intervention. The other 146 sources (146/148, 97%) reported 2 to 5 different purposes for the interventions resulting in a total of 296 different purposes. In these interventions, the top 5 purposes were balance (73/148, 49%), muscular strength and power (64/148, 43%), cardiorespiratory fitness (30/148, 20%), functional mobility (28/148, 19%) and general physical activity (25/148, 17%). In 40% (58/148) of the sources, the focus was on one purpose, whereas the majority (88/148, 59%) had on average 3 (min 2, max 5) purposes.

Technology Characteristics

An overview of the technology designs, functions, and interfaces used in the sources is given in Table 2. Below results of the cross-tabulation between technology and population or intervention characteristics are described. All detailed results can be found in Multimedia Appendix 4.

XSL•FO
RenderX

Table . Extracted counts for the Design, Function, and Interface categories. N equals the total reported for each category, and in brackets, the percentage of the total reported per Design, Function, and Interface, respectively.

Variable	Categories	N (%)
Design		
	Computer-based	20 (9)
	Sound and video recording systems	8 (4)
	Mobile applications (web-based, tablet, mobile phone)	48 (22)
	Web video/ phone call	30 (13)
	Exergames (incl. screen)	68 (30)
	Trackers	21 (9)
	Virtual reality	6 (3)
	Augmented reality	1 (0)
	Wearable devices	14 (6)
	Other	7 (3)
Function		
	Assessment of outcomes	21 (6)
	Providing exercise program information once	28 (8)
	Providing exercise information continuously	102 (27)
	Feedback of exercise	102 (27)
	Real-time health metrics	6 (2)
	Real-time movement data	98 (26)
	Real-time functionality ability	9 (2)
	Other	5 (1)
Interface		
	Graphical interface	94 (27)
	Command line with feedback	24 (7)
	Command line without feedback	13 (4)
	Menu-driven	40 (11)
	(Haptic) touch	6 (2)
	Auditory or visual	61 (17)
	Body movement	93 (27)
	Form-based	1 (0)
	Natural language processing	0 (0)
	Remote driven	14 (4)
	Other	3 (1)

Cross-Tabulation Between Technology and Population

The results of the cross-tabulation between participant characteristics and technology design are presented in Table 3. The cross-tabulation results between participant characteristics and technology function and interface can be found in Multimedia Appendix 4. Exergames were implemented in half of the sources with oldest-old adult participants who are aged 85 years and older. In the sources with middle-old age groups (75 - 85 y), exergames (25/75, 33%) and mobile apps (16/75, 21%) were both frequently used, and alternative designs such as web/video phone call (10/75, 13%) were implemented as well. In sources with youngest-old age groups (60 - 74 y), trackers (18/125, 14%) were more often implemented than in the middle-old or oldest-old age groups (2/75, 3% and 1/18, 6% respectively). In sources where the participant's health status was reported as healthy, 53% (10/19) implemented exergames compared to only 27% (34/127) in sources with participants with a clinical condition. For both health status groups, mobile apps (20%, ie., 4/19 and 24/127) and trackers (11%, ie., 2/19 and 14/127) were likewise implemented. However, sources with participants with a clinical condition also made use of other

XSL•FO
RenderX

technological designs such as computer-based (14/127, 11%), web video/phone call (19/127, 15%) or wearable devices (10/127, 8%). Age and health status did not influence the technological functions or user interface.

Table . Cross-tabulation results for participant characteristic age, health status, and residency versus technology design.

Design	Sum[a]	Computer based	Sound/ video recording systems	Mobile apps	Web/ video phone call	Exer-games	Trackers	Virtual reality	Augmented reality	Wearable devices	Other
AGE											
Youngest old (60 - 74 y)	125	7%	2%	24%	15%	25%	14%	2%	-	6%	4%
Middle old (75 - 85 y)	75	11%	7%	21%	13%	33%	3%	4%	-	7%	1%
Oldest old (>85 + y)	18	11%	6%	11%	6%	50%	6%	-	-	6%	6%
Mixed	3	-	-	-	-	67%	-	-	33%	-	-
HEALTH STATUS											
Healthy	19	-	-	21%	5%	53%	11%	5%	-	-	5%
Clinical condition	127	11%	5%	19%	15%	27%	11%	2%	-	8%	3%
RESIDENCY											
Care Homes	42	10%	5%	7%	2%	50%	7%	10%	-	2%	7%
Community Dwelling	153	9%	3%	24%	18%	26%	10%	1%	1%	5%	3%
Hospital	1	-	-	-	-	-	100%	-	-	-	-
Other	33	6%	3%	24%	6%	21%	3%	18%	-	18%	-

[a]The horizontal rows add up to 100% of the sum in the second column. The other columns represent the percentages of the sum, and the total of each row equals 100%. With abbreviations: Apps, applications; cmnd, command; record., recording; y, years.

In community setting studies, exergames (40/154, 26%) and apps (37/154, 24%) were equally deployed. However, in care homes, half of the sources used exergames. Wearable technology was used frequently in settings where residency was unfortunately not specified, whereas the use of web video/phone call designs was highest (relatively) in community setting studies. Making use of specific functions and interfaces was similarly distributed for community dwelling and care home residents. Only the body movement interface was 34% (21/61) in care homes compared to 25% (59/236) in community settings, probably related to the implemented exergames.

Cross-Tabulation Between Technology and Intervention

The results of the cross-tabulation analyses for purpose of the physical activity and technology are presented in Multimedia Appendix 4. Exergame and mobile app designs were used to achieve improvements for a large range of physiological parameters such as balance, muscle strength, functional mobility, or cardiovascular health. Similar for web video/phone calls, but these technologies were incorporated into the sources to a much lesser amount. Exergames were predominantly used to improve cardiorespiratory fitness (17/41, 41%), neuromotor function (14/29, 48%), mental and cognitive function (16/32, 50%), and fall risk (8/14, 57%). Only enhancement of general physical activity was trained most using trackers (13/39, 33%, eg, for step counting). Regardless of the purpose, the technologies provided continuous information on the activity, feedback, and assessed real-time movement data. These technological functions were achieved through graphical (25% - 30%), body movement (20% - 36%), or auditory and visual (13% - 23%) interfaces with the users.

Tailoring of the intervention's physical activity was observed across all technologies. Hence, incorporating tailoring of the intervention's physical activity did not seem to be related to the type of technology design, function of user interface, and vice-versa.

XSL•FO
RenderX

Outcome Measurements

Adverse Events

In total, 64% (68/148) of the sources reported information on adverse events. Out of these, 72% (49/68) of the sources reported no adverse events, with the other 28% (19/68) reporting a total of 78 adverse events (n=78, 100%) categorized in grade 1 (mild, 40/78, 51%), grade 2 (moderate, 10/78, 13%), grade 3 (severe, 28/78, 36%) events, with none reported in grade 4 (life-threatening) or grade 5 (death).

Of these 78 adverse events, 47 events could be related to the intervention and included 37 mild events, e.g., primarily musculoskeletal pain in joints, 8 moderate events, e.g., rib fracture or COPD-related, and 2 severe events, eg, severe injury due to (bicycle) fall accident. Only 2 adverse events were related to technology and included mild impairments from a skin rash due to Fitbit wearing and nausea due to the Virtual Reality environment.

Dropout and Adherence

Information on dropout and adherence for the intervention groups was reported in 100 sources (68%) and 79 sources (53%), respectively. Information on dropout and adherence for the control groups was reported more sparely in 45 sources (47%) and 16 sources (17%), respectively. While the number of participants for all sources is variable, dropout rates of control groups (14 SD 12%) and intervention groups (16 SD 16%) were similar (t_{34} = −0.648, P=.521). Average adherence tended to be lower in the control groups (71 SD 25%) compared to the intervention groups (80 SD 18%), though this could not be tested statistically due to underreporting in the control group.

In total, the dropout reasons for 815 participants were quantitatively reported. Reasons for dropout were mostly health-related (28%), though these were unrelated to the intervention. Furthermore, dropout participants mentioned time-related issues (17%), such as lack of time or traveling and burden or motivation-related issues (15%) such as lost interest or feeling burdened. Often another reason (17%) was indicated for the dropout, or no reason was mentioned at all (15%). About 5% of the dropout participants were lost due to follow-up issues ascribed to e.g., loss of contact or moving. Only 3% of the dropouts were related to technology: for example, broken computer, internet malfunctioning, lack of device or space.

No differences (t_{98}=−1.209, P=.115, d=−0.252) were found in dropout when interventions were tailored (15 SD 16%) or not tailored (19 SD 14%). A significantly higher adherence (t_{77}=1.995, P=0.025, d=0.455) was found between interventions that were tailored (83 SD 15%) versus those that were not (75 SD 21%). On the contrary, no differences were observed when comparing supervised versus unsupervised interventions, when assessing dropout (t_{98}=-1.374, P=.086, d=−0.275) and adherence (t_{77}=0.606, P=.273, d=0.136).

Sources including participants with reported clinical conditions as health status (18 SD 11% dropout) reported a similar level of dropout rate compared to those who used healthy populations (17 SD 18% dropout), which was nonsignificant (t_{62}=0.063, P=.475, d=0.022). There were no differences between adherence rates when comparing the sources with reported healthy participants versus participants with a reported clinical condition (t_{48}=0.027, P=.489, d=0.010). Differential testing between age groups was not possible, due to the low reported dropout ratios across three age categories, although no trends were observed of dropout rate differences across age categories.

No further relationships were found between dropout or adherence, when regressing against the total number of intervention sessions (R^2=0.012, P=.924) or session duration (R^2=0.001, P=.850).

User Experience

A total of 55 sources (37%) explored the general enjoyment of the technology part of their intervention. An overwhelming 91% reported a positive experience, with a vast minority reporting a neutral or negative experience (7% and 2% respectively).

Only 21 sources (14%) investigated the system usability. The system usability was reported using the System Usability Scale (SUS, N=12, 8%) or other questionnaires, mostly using Likert scales or percentages (N=9, 6%). The reported SUS values ranged from 62 to 90 with an average of 77 (SD 9) indicating okay, good, and excellent overall user friendliness [176]. In several sources (N=2), it was reported that the SUS was assessed, but then the result was not presented.

Regardless of the study design, wishes and needs of the users were seldom (systematically) assessed. When reported, the wishes were related to goal setting, feedback (regarding own or other participants' performance, messages), support (technical, medical), the physical activity (variation, choice of level), the technology (access, comfort, malfunctioning, touch screen), and social setting (in groups, partnering up, avatar look, timing).

Discussion

Principal Findings

This scoping review explores the application of digital technologies in relation to the characteristics of populations (eg, health status), interventions and usability in home-based physical activity interventions for older adults. Analyzing 148 sources, the review finds that technologies like exergames and mobile apps are widely applicable and easy to use across diverse older adult populations, including healthy and those with clinical conditions. These technologies offer continuous features like information about the physical activity programs, feedback, and real-time movement tracking, addressing various goals such as balance, strength, functional mobility, and cardiovascular fitness. Interestingly, technology use did not lead to higher dropout rates, and adherence was significantly higher in tailored interventions. Adverse events directly related to the technology were rare. User evaluations indicated positive feedback regarding the usability and enjoyment of the technology, suggesting its potential to increase physical activity engagement among older adults.

XSL•FO
RenderX

Participants

Our research underscores the growing evidence of technology-based physical activity interventions in home-based environments (community dwelling and care homes) since 2010, with a notable increase over the past five years [177]. Our review revealed that technology-assisted physical activity interventions have been explored among older adults in various age groups and with various health conditions. Given the increasing prevalence of age-related diseases and chronic health conditions among older adults, technology holds promise in promoting physical activity engagement, particularly among those with health challenges [19,25].

The use of technology-assisted physical activity interventions for older adults in hospital settings is limited [178-180]. Initially, this review identified no such studies (N=0), likely due to excluding abstracts mentioning “rehabilitation” in our search. While physical activity interventions in hospitals generally show positive effects on functional decline or hospital-associated disability [3,181,182], Sheerman et al report inconsistent findings [183]. Similarly, technology-assisted interventions in hospitals report mixed results: adherence was high for supervised interactive gaming [180] and activity trackers [178] but low for self-regulated exergames [179]. Furthermore, Laver et al reported improved Time-Up-and-Go test results with no serious adverse events, while Oesch et al noted no balance improvements and 2 adverse events related to pain [179,180].

Studies failed to report participants’ education (only 40 %) and socio-economic status (only 9 %), which are critical factors as lower education and lower socioeconomic status often correlate with lower digital health literacy [184,185]. Lower education influences the level of health literacy in general as well [184], which may influence judgment on the relevance of engaging with health technology [186]. Furthermore, lower socio-economic status has been related to suboptimal access to internet and digital tools [184], e.g., due to costs [187]. Unfortunately, only limited information was reported by the included sources (15%) on access to technology. Also, technology familiarity or e-literacy was poorly reported (15%). Lower education and the lack of familiarization with technology may introduce barriers to the capacity of users to understand and attend a technology-supported intervention and may lead to increased anxiety and dropouts when attending [186-188]. Hence, the limited reporting on such participant characteristics raises concerns about sample bias and the generalizability of results.

Interventions

This scoping review reveals significant diversity in technology-assisted intervention designs, supervision levels, and tailoring. While many interventions are appropriately tailored to the specific needs of older adults and target populations, this tailoring naturally results in variability in intervention durations, session lengths, and frequencies. This does not necessarily indicate a lack of structure but reflects the necessary flexibility in designing interventions that meet individual or group needs. However, the lack of consistent reporting frameworks for these interventions poses a challenge for comparing results across studies. Our findings are consistent with other reviews in the field (eg, [177,189-191]), which also identified significant variability in intervention designs and reporting frameworks. Those reviews, like ours, emphasize the importance of individualization and tailoring as the key components in ensuring that interventions are effective, but the absence of standardized reporting makes it difficult to evaluate the relative success of these programs. Particularly, the underreporting of intervention intensity poses a significant challenge, with only 14% of sources detailing intervention intensity using varied methods. This is unexpected since technology can easily track and report intensity, providing both motivation and precise data for older adults. For instance, studies using Kinect systems and wearables like Fitbit have shown potential in accurately tracking activity intensity, enhancing engagement and usability [192-194].

Technologies

In the sources we reviewed, exergaming emerged as the most used approach, featuring 53% of interventions for healthy populations and 27% for clinical groups. Despite most exergaming applications targeting younger populations [195,196], our scoping review identifies a significant shift towards older users, with exergames implemented in half of the sources involving participants aged 85 and above. The findings support and build upon previous research suggesting the utility of exergames in promoting physical activity, cognitive enhancement, and overall well-being in the older population [11]. Furthermore, the diversity of exergame intervention contents highlights a comprehensive strategy for addressing age-related challenges such as cardiorespiratory fitness (41%), neuromotor functions (48%), mental and cognitive abilities (50%), and reducing falls (57%).

Our findings also indicate that while exergames and mobile applications are the primary technologies used in these physical activity interventions, web video and phone call technologies or wearable devices also play a role, albeit less prominently. Especially in the clinical population, more often other designs than exergames were used compared to the interventions performed with healthy older adults. These other forms of technology design, offering a more passive interaction than the dynamic engagement found in exergames, may not be as prevalent but are nonetheless vital. Older, or more prone older adults may prefer and benefit from their easy set-up and simple usage [197-199]. Their integration highlights their significance within a comprehensive technological intervention framework, which aligns with effective telehealth components like text messaging for education and reminders, web-based content for goal-setting, and interactive platforms for health data exchange [200].

In advocating for the integration of technology in physical activity interventions, our review highlights the pivotal role of technology functions and user interfaces, as supported by Bentlage et al. (2023), who stress their essential contribution to improve the activation factors skills, knowledge, and motivation, yielding effective behavioral change [20]. We pointed out the primary technological functions as the continuous provision of physical activity information (28%), performance feedback (28%), and the recording of real-time

XSL•FO
RenderX

movement data (26%). Such continuous and synchronous activity information offers users an immersive, enjoyable, and engaging experience [201], fosters regular habits [202], can facilitate the visualization of their movements, and significantly increase their motivation [203]. Enhancing user engagement and overcoming engagement barriers can be achieved by using advanced technology to incorporate different user interfaces, preferably multiple types [20,204]. While graphical or body movement interfaces were prevalent in many sources (each 63%), alternative user interfaces were limited (ie, 41% visual/auditory, 27% menu driven, and 16% command line).

Notably, the tailoring of intervention exercises or activities did not significantly influence the choice of technology design, function, or user interface. This might indicate that the deployed technological solutions are versatile and adaptable across various intervention needs without requiring significant modifications [205,206]. Moreover, our review emphasizes a shift towards personalized interventions tailored to individual or group needs, with 63% of interventions adopting this approach (versus 37% with a generic approach). This shift aligns with emerging evidence suggesting the effectiveness of personalized strategies in promoting adherence and achieving positive health outcomes among older adults [207].

Technology-Related Outcomes

Technology engagement and adherence to protocols are tightly linked to the usability, enjoyment, and effectiveness of physical activity-based interventions [208]. Yet only a slight majority of studies reported on dropout and adherence rates, whereas an even smaller number reported on reasons for dropout or (reduced) adherence. Despite the Consolidated Standard of Reporting Trials (CONSORT) [209] statement and TIDieR (Template for Intervention Description and Replication) [32] guidelines dating back to 2010 and 2014, and the relevant Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) 2022 extension [210], there is still a lack of consistency of reporting.

Sources that reported the reasons for dropout among participants primarily cited health-related issues, with only a few instances attributed to technology usage, underscoring the older participants' willingness and engagement with home-based technology. Such instances of technology-related dropout (3%) were predominantly due to technological failures, connectivity issues, or general disinterest. These findings are consistent with recent studies, which also identified a lack of motivation and low familiarity with technology as key barriers to its implementation in older adults [15,23,211]. Furthermore, tailored interventions resulted in significantly higher adherence than control interventions. This underscores that it is imperative to adapt the technology to the preferences and needs of the target aging population [22,23,197-199,212-214]. To achieve suitable technological approaches, a generative co-design framework, could be used to overcome technological barriers in an aging population, as suggested by Bird et al [215].

Strengths and Limitations of This Scoping Review

Adopting a common vocabulary and definition of terms is challenging in research on technology-assisted physical activity interventions, which can be the intersection of medicine, physiotherapy, engineering and informatics, sensor and gaming technology, and even psychology and social sciences. Though challenging, we discussed and agreed upon definitions and vocabulary usage at each phase of the review. Furthermore, multiple protocols and double-blind scoring were incorporated in most phases of the review process to ensure consistency [216]. The relatively large number of screeners and extractors involved may be considered a potential limitation. However, the screening, data extraction, and data analysis tools were carefully prepared and discussed within the group, and conflicts were solved with a third independent reviewer or group discussions.

Regarding the included sources, most of the reviewed studies were from highly developed and advanced countries. In addition, the sources included in this scoping review often failed to describe the used technology in sufficient detail to allow for replication of the design, function, and/or user interface. For example, exergames rely on an often bilateral interaction between the user and the system. However, the inputs (eg, handheld devices, sensors, video cameras) are poorly described on how these were adapted for use. Valenzuela et al 217 explored older adults' experiences using an interactive cognitive-motor step training program. Yet, there is no clear description of how the interface has been adapted for their specific user group. Other sources assumed prior or even in-depth knowledge of commercial exergaming setups used within the intervention, avoiding interpretation of the data or replicability due to unknown use of commercial games (Wii-Fit), body movement recognition through three-dimensional camera setups (Xbox Kinect), or not clearly described engagement elements of online video calling for physical activity class delivery (eg, Yoga via Zoom calls) [218,219].

This review has been carried out in a systematic and transparent manner [220]. However, we have not conducted a quality assessment or risk of bias of the included sources. Therefore, we cannot use the results of the included sources to draw significant conclusions on intervention effects on health outcomes.

Recommendations

There is no compelling reason to exclude specific older populations from home-based, technology-assisted physical activity interventions, as factors such as age, health status, and residency do not appear to negatively impact dropout rates, adherence, or result in technology-related adverse events. However, evidence regarding the use of such technologies in hospital settings remains limited. Vulnerable older adults, in particular, could benefit significantly from these interventions, as reduced physical activity during hospitalization often leads to functional decline and a loss of independence in daily activities at discharge. Regardless of the setting, the design of technology-assisted interventions must account for the physical and cognitive impairments commonly observed among older adults, which may result in a mismatch between their abilities and the demands of the intervention or technology [186]. Careful tailoring of both the intervention and the technology is therefore essential. For example, participatory design processes that

XSL•FO
RenderX

engage older adults directly can ensure better alignment with their needs and preferences [215]. Our review highlights that tailored approaches are associated with significantly higher adherence rates. Furthermore, feedback from older adults highlights their strong preference for tailored approaches, indicating that tailoring is not only beneficial but also actively desired by older adults. To improve the accessibility and impact of these interventions, training to enhance health literacy (knowledge of the intervention's relevance) and digital literacy (competence in using the technology) should be incorporated during both the recruitment and intervention phases. Developing clear recruitment guidelines may further promote inclusivity. Digital literacy barriers, often linked to low education levels and socioeconomic status, remain significant challenges to technology adoption and engagement [184,186-188]. To address this, technical support—either private or as part of the intervention—should be provided, aligning with preferences expressed by older adults in user feedback. Lastly, behavioral change techniques that enhance skills, knowledge, and motivation should be integrated into the design of technology-assisted physical activity interventions [20]. Feedback mechanisms, such as performance-based metrics, functional ability assessments, and real-time health data (eg, movement intensity), are currently underutilized, appearing in less than one-third of interventions. Implementing such features could encourage behavioral change and fulfill the desires for actionable feedback, as reported by older adults.

Furthermore, future research should prioritize transparent reporting. The participant's health status, including co-morbidities, mental and physical impairments, and activity level, as well as technology experience, should be recorded. The intervention design, including purpose, duration, intensity, supervision, and tailoring, should be well described. The technology design, functions (including consideration of behavioral change techniques), and user interface should be described in detail. Lastly, intervention and technology-related outcomes such as drop-out, adherence, adverse events, and feedback on user experience should be reported. Guidelines for the development of technology-assisted physical activity interventions in older adults with various health conditions may lead to conformity and could contribute to better designs engaging older adults. Such guidelines may improve the effectiveness of the interventions and make studies better comparable by adequate reporting. The latter results in a deeper understanding of user preferences and advances the development of tailored technology-assisted interventions.

Conclusions

The review suggests that using technology in physical activity interventions is feasible for all older adults in community dwellings and care homes, without additional risks of adverse events or dropouts. Furthermore, generally higher adherence was reported in technology-assisted interventions, which was significantly higher where interventions were tailored. Interventions were delivered to all older age groups, as well as those with and without clinical conditions and in a wide range of physical activity topics. The technology approaches for home-based physical activity interventions were found to be safe and enjoyable for older adults with acceptable usability. Despite the large number of sources that have been identified in this article, the lack of standardized reporting limited a more in-depth analysis of the main research question. Therefore, there is a need to update reporting guidelines for technology-specific interventions in this area.

Acknowledgments

This publication is based upon work from EU COST Action CA20104 – Network on evidence-based physical activity in old age (PhysAgeNet), supported by COST (European Cooperation in Science and Technology). We acknowledge the following members of the COST action 'PhysAgeNet' Working group 3 for their support in the data acquisition: Bassam Bouaziz from University of Sfax, Juel Jarani from University of Tirana, and Bojan Masanovic from University of Montenegro.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search string development.
[DOCX File, 40 KB - aging_v8i1e65746_app1.docx]

Multimedia Appendix 2
Data extraction topics.
[DOCX File, 37 KB - aging_v8i1e65746_app2.docx]

Multimedia Appendix 3
Coding framework for data extraction.
[DOCX File, 69 KB - aging_v8i1e65746_app3.docx]

Multimedia Appendix 4

XSL•FO
RenderX

Result tables and sources.
[DOCX File, 137 KB - aging_v8i1e65746_app4.docx]

Multimedia Appendix 5
Full data extraction.
[XLSX File, 257 KB - aging_v8i1e65746_app5.xlsx]

Multimedia Appendix 6
Excluded sources reasons.
[XLSX File, 105 KB - aging_v8i1e65746_app6.xlsx]

Checklist 1
PRISMA-ScR checklist.
[DOCX File, 110 KB - aging_v8i1e65746_app7.docx]

References

1. Vogel T, Brechat PH, Leprêtre PM, Kaltenbach G, Berthel M, Lonsdorfer J. Health benefits of physical activity in older patients: a review. Int J Clin Pract 2009 Feb;63(2):303-320. [doi: 10.1111/j.1742-1241.2008.01957.x] [Medline: 19196369]
2. Bull FC, Al-Ansari SS, Biddle S, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med 2020 Dec;54(24):1451-1462. [doi: 10.1136/bjsports-2020-102955] [Medline: 33239350]
3. Aubertin-Leheudre M, Rolland Y. The importance of physical activity to care for frail older adults during the COVID-19 pandemic. J Am Med Dir Assoc 2020 Jul;21(7):973-976. [doi: 10.1016/j.jamda.2020.04.022] [Medline: 32513558]
4. Bosco A, McGarrigle L, Skelton DA, Laventure RME, Townley B, Todd C. Make movement your mission: evaluation of an online digital health initiative to increase physical activity in older people during the COVID-19 pandemic. Digit Health 2022;8:20552076221084468. [doi: 10.1177/20552076221084468] [Medline: 35295764]
5. Bangsbo J, Blackwell J, Boraxbekk CJ, et al. Copenhagen consensus statement 2019: physical activity and ageing. Br J Sports Med 2019 Jul;53(14):856-858. [doi: 10.1136/bjsports-2018-100451] [Medline: 30792257]
6. Sohaib Aslam A, van Luenen S, Aslam S, van Bodegom D, Chavannes NH. A systematic review on the use of mHealth to increase physical activity in older people. Clinical eHealth 2020;3:31-39. [doi: 10.1016/j.ceh.2020.04.002]
7. Shin G, Jarrahi MH, Fei Y, et al. Wearable activity trackers, accuracy, adoption, acceptance and health impact: a systematic literature review. J Biomed Inform 2019 May;93(October 2018):103153. [doi: 10.1016/j.jbi.2019.103153] [Medline: 30910623]
8. Yerrakalva D, Yerrakalva D, Hajna S, Griffin S. Effects of mobile health app interventions on sedentary time, physical activity, and fitness in older adults: systematic review and meta-analysis. J Med Internet Res 2019 Nov 28;21(11):e14343. [doi: 10.2196/14343] [Medline: 31778121]
9. Delbaere K, Valenzuela T, Woodbury A, et al. Evaluating the effectiveness of a home-based exercise programme delivered through a tablet computer for preventing falls in older community-dwelling people over 2 years: study protocol for the Standing Tall randomised controlled trial. BMJ Open 2015 Oct 22;5(10):e009173. [doi: 10.1136/bmjopen-2015-009173] [Medline: 26493461]
10. Dermody G, Whitehead L, Wilson G, Glass C. The role of virtual reality in improving health outcomes for community-dwelling older adults: systematic review. J Med Internet Res 2020 Jun 1;22(6):e17331. [doi: 10.2196/17331] [Medline: 32478662]
11. Kappen DL, Mirza-Babaei P, Nacke LE. Older adults' physical activity and exergames: a systematic review. International Journal of Human–Computer Interaction 2019 Jan 20;35(2):140-167. [doi: 10.1080/10447318.2018.1441253]
12. Chatterjee A, Gerdes M, Prinz A, Martinez S. Human coaching methodologies for automatic electronic coaching (eCoaching) as behavioral interventions with information and communication technology: systematic review. J Med Internet Res 2021 Mar 24;23(3):e23533. [doi: 10.2196/23533] [Medline: 33759793]
13. Lotfi A, Langensiepen C, Yahaya SW. Socially assistive robotics: robot exercise trainer for older adults. Technologies (Basel) 2018;6(1):32. [doi: 10.3390/technologies6010032]
14. Stockwell S, Schofield P, Fisher A, et al. Digital behavior change interventions to promote physical activity and/or reduce sedentary behavior in older adults: A systematic review and meta-analysis. Exp Gerontol 2019 Jun;120(January):68-87. [doi: 10.1016/j.exger.2019.02.020] [Medline: 30836130]
15. Valenzuela T, Okubo Y, Woodbury A, Lord SR, Delbaere K. Adherence to technology-based exercise programs in older adults: a systematic review. J Geriatr Phys Ther 2018;41(1):49-61. [doi: 10.1519/JPT.0000000000000095]
16. Langeard A, Bigot L, Maffiuletti NA, et al. Non-inferiority of a home-based videoconference physical training program in comparison with the same program administered face-to-face in healthy older adults: the MOTION randomised controlled trial. Age Ageing 2022 Mar 1;51(3):afac059. [doi: 10.1093/ageing/afac059]

XSL•FO
RenderX

17. Lee J, Jung D, Byun J, Lee M. Effects of a combined exercise program using an iPad for older adults. Healthc Inform Res 2016 Apr;22(2):65-72. [doi: 10.4258/hir.2016.22.2.65] [Medline: 27200215]
18. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556. [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
19. Batalik L, Filakova K, Sladeckova M, Dosbaba F, Su J, Pepera G. The cost-effectiveness of exercise-based cardiac telerehabilitation intervention: a systematic review. Eur J Phys Rehabil Med 2023 Apr;59(2):248-258. [doi: 10.23736/S1973-9087.23.07773-0] [Medline: 36692413]
20. Bentlage E, Nyamadi JJ, Dubbeldam R. The importance of activating factors in physical activity interventions for older adults using information and communication technologies: systematic review. JMIR Mhealth Uhealth 2023 Oct 24;11:e42968. [doi: 10.2196/42968] [Medline: 37933182]
21. McGarrigle L, Boulton E, Todd C. Map the apps: a rapid review of digital approaches to support the engagement of older adults in strength and balance exercises. BMC Geriatr 2020 Nov 18;20(1):483. [doi: 10.1186/s12877-020-01880-6] [Medline: 33208117]
22. Piech J, Czernicki K. Virtual reality rehabilitation and exergames—physical and psychological impact on fall prevention among the elderly—a literature review. Appl Sci (Basel) 2021;11(9):4098. [doi: 10.3390/app11094098]
23. Costa-Brito AR, Bovolini A, Rúa-Alonso M, et al. Home-based exercise interventions delivered by technology in older adults: A scoping review of technological tools usage. Int J Med Inform 2024 Jan;181:105287. [doi: 10.1016/j.ijmedinf.2023.105287] [Medline: 37972483]
24. Skjæret N, Nawaz A, Morat T, Schoene D, Helbostad JL, Vereijken B. Exercise and rehabilitation delivered through exergames in older adults: An integrative review of technologies, safety and efficacy. Int J Med Inform 2016 Jan;85(1):1-16. [doi: 10.1016/j.ijmedinf.2015.10.008] [Medline: 26559887]
25. Yang Y, Li S, Cai Y, et al. Effectiveness of telehealth-based exercise interventions on pain, physical function and quality of life in patients with knee osteoarthritis: A meta-analysis. J Clin Nurs 2023 Jun;32(11-12):2505-2520. [doi: 10.1111/jocn.16388] [Medline: 35872635]
26. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evidence Synthesis 2020;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167]
27. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
28. Dubbeldam R, Stemplewski R, Cyma-Wejchenig M, et al. Scoping review on technology-assisted physical interventions for older people in their home environment. OSF 2023 [FREE Full text]
29. Denton F, Power S, Waddell A, et al. Is It really home-based? A commentary on the necessity for accurate definitions across exercise and physical activity programmes. Int J Environ Res Public Health 2021 Sep 1;18(17):9244. [doi: 10.3390/ijerph18179244] [Medline: 34501833]
30. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
31. Pollock D, Peters MDJ, Khalil H, et al. Recommendations for the extraction, analysis, and presentation of results in scoping reviews. JBI Evid Synth 2023 Mar 1;21(3):520-532. [doi: 10.11124/JBIES-22-00123] [Medline: 36081365]
32. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014 Mar 7;348(3):g1687. [doi: 10.1136/bmj.g1687] [Medline: 24609605]
33. Sheehy L, Sveistrup H, Knoefel F, et al. The use of home-based nonimmersive virtual reality to encourage physical and cognitive exercise in people with mild cognitive impairment: a feasibility study. J Aging Phys Act 2022 Apr 1;30(2):297-307. [doi: 10.1123/japa.2021-0043] [Medline: 34453024]
34. Lee BJ, Park YH, Lee JY, Kim SJ, Jang Y, Lee JI. Smartphone application versus pedometer to promote physical activity in prostate cancer patients. Telemed J E Health 2019 Dec;25(12):1231-1236. [doi: 10.1089/tmj.2018.0233] [Medline: 30758247]
35. Taylor ME, Close JCT, Lord SR, et al. Pilot feasibility study of a home-based fall prevention exercise program (StandingTall) delivered through a tablet computer (iPad) in older people with dementia. Australas J Ageing 2020 Sep;39(3):e278-e287. [doi: 10.1111/ajag.12717] [Medline: 31538401]
36. Leskinen T, Suorsa K, Tuominen M, et al. The effect of consumer-based activity tracker intervention on physical activity among recent retirees-an RCT study. Med Sci Sports Exerc 2021 Aug 1;53(8):1756-1765. [doi: 10.1249/MSS.0000000000002627] [Medline: 34261997]
37. Lauzé M, Martel DD, Agnoux A, et al. Feasibility, acceptability and effects of a home-based exercise program using a gerontechnology on physical capacities after a minor injury in community-living older adults: a pilot study. J Nutr Health Aging 2018;22(1):16-25. [doi: 10.1007/s12603-017-0938-8] [Medline: 29300417]
38. Nikitina S, Didino D, Baez M, Casati F. Feasibility of virtual tablet-based group exercise among older adults in Siberia: findings from two pilot trials. JMIR Mhealth Uhealth 2018 Feb 27;6(2):e40. [doi: 10.2196/mhealth.7531] [Medline: 29487045]
39. Janssen S, Tange H, Arends R. A preliminary study on the effectiveness of Exergame Nintendo “Wii Fit Plus” on the balance of nursing home residents. Games Health J 2013 Apr;2(2):89-95. [doi: 10.1089/g4h.2012.0074] [Medline: 26192126]

XSL•FO
RenderX

40. Talbot LA, Gaines JM, Huynh TN, Metter EJ. A home-based pedometer-driven walking program to increase physical activity in older adults with osteoarthritis of the knee: a preliminary study. J Am Geriatr Soc 2003 Mar;51(3):387-392. [doi: 10.1046/j.1532-5415.2003.51113.x] [Medline: 12588583]
41. Albores J, Marolda C, Haggerty M, Gerstenhaber B, Zuwallack R. The use of a home exercise program based on a computer system in patients with chronic obstructive pulmonary disease. J Cardiopulm Rehabil Prev 2013;33(1):47-52. [doi: 10.1097/HCR.0b013e3182724091] [Medline: 23235321]
42. Hawley-Hague H, Tacconi C, Mellone S, et al. Smartphone apps to support falls rehabilitation exercise: app development and usability and acceptability study. JMIR Mhealth Uhealth 2020 Sep 28;8(9):e15460. [doi: 10.2196/15460] [Medline: 32985992]
43. Silveira P, van de Langenberg R, van Het Reve E, Daniel F, Casati F, de Bruin ED. Tablet-based strength-balance training to motivate and improve adherence to exercise in independently living older people: a phase II preclinical exploratory trial. J Med Internet Res 2013 Aug 12;15(8):e159. [doi: 10.2196/jmir.2579] [Medline: 23939401]
44. Gschwind YJ, Schoene D, Lord SR, et al. The effect of sensor-based exercise at home on functional performance associated with fall risk in older people - a comparison of two exergame interventions. Eur Rev Aging Phys Act 2015;12:11. [doi: 10.1186/s11556-015-0156-5] [Medline: 26865875]
45. Demeyer H, Louvaris Z, Frei A, et al. Physical activity is increased by a 12-week semiautomated telecoaching programme in patients with COPD: a multicentre randomised controlled trial. Thorax 2017 May;72(5):415-423. [doi: 10.1136/thoraxjnl-2016-209026] [Medline: 28137918]
46. Pettersson B, Janols R, Wiklund M, Lundin-Olsson L, Sandlund M. Older adults' experiences of behavior change support in a digital fall prevention exercise program: qualitative study framed by the self-determination theory. J Med Internet Res 2021 Jul 30;23(7):e26235. [doi: 10.2196/26235] [Medline: 34328438]
47. Muñoz J, Mehrabi S, Li Y, et al. Immersive virtual reality exergames for persons living with dementia: user-centered design study as a multistakeholder team during the COVID-19 pandemic. JMIR Serious Games 2022 Jan 19;10(1):e29987. [doi: 10.2196/29987] [Medline: 35044320]
48. Hong J, Kim J, Kim SW, Kong HJ. Effects of home-based tele-exercise on sarcopenia among community-dwelling elderly adults: Body composition and functional fitness. Exp Gerontol 2017 Jan;87(Pt A):33-39. [doi: 10.1016/j.exger.2016.11.002] [Medline: 27838369]
49. Neumann S, Meidert U, Barberà-Guillem R, Poveda-Puente R, Becker H. Effects of an exergame software for older adults on fitness, activities of daily living performance, and quality of life. Games Health J 2018 Oct;7(5):341-346. [doi: 10.1089/g4h.2017.0079] [Medline: 30325234]
50. Chao YY, Scherer YK, Wu YW, Lucke KT, Montgomery CA. The feasibility of an intervention combining self-efficacy theory and Wii Fit exergames in assisted living residents: a pilot study. Geriatr Nurs 2013;34(5):377-382. [doi: 10.1016/j.gerinurse.2013.05.006] [Medline: 23764366]
51. Golla A, Müller T, Wohlfarth K, Jahn P, Mattukat K, Mau W. Home-based balance training using Wii Fit: a pilot randomised controlled trial with mobile older stroke survivors. Pilot Feasibility Stud 2018;4:143. [doi: 10.1186/s40814-018-0334-0] [Medline: 30155268]
52. Espay AJ, Baram Y, Dwivedi AK, et al. At-home training with closed-loop augmented-reality cueing device for improving gait in patients with Parkinson disease. J Rehabil Res Dev 2010;47(6):573-581. [doi: 10.1682/jrrd.2009.10.0165] [Medline: 20848370]
53. Jaarsma T, Klompstra L, Ben Gal T, et al. Effects of exergaming on exercise capacity in patients with heart failure: results of an international multicentre randomized controlled trial. Eur J Heart Fail 2021 Jan;23(1):114-124. [doi: 10.1002/ejhf.1754] [Medline: 32167657]
54. Rosenberg D, Depp CA, Vahia IV, et al. Exergames for subsyndromal depression in older adults: a pilot study of a novel intervention. Am J Geriatr Psychiatry 2010 Mar;18(3):221-226. [doi: 10.1097/JGP.0b013e3181c534b5] [Medline: 20173423]
55. Vestergaard S, Kronborg C, Puggaard L. Home-based video exercise intervention for community-dwelling frail older women: a randomized controlled trial. Aging Clin Exp Res 2008 Oct;20(5):479-486. [doi: 10.1007/BF03325155] [Medline: 19039291]
56. Li F, Harmer P, Fitzgerald K, Winters-Stone K. A cognitively enhanced online Tai Ji Quan training intervention for community-dwelling older adults with mild cognitive impairment: a feasibility trial. BMC Geriatr 2022 Jan 25;22(1):76. [doi: 10.1186/s12877-021-02747-0] [Medline: 35078407]
57. Davis JC, Hsu CL, Cheung W, et al. Can the Otago falls prevention program be delivered by video? A feasibility study. BMJ Open Sport Exerc Med 2016;2(1):e000059. [doi: 10.1136/bmjsem-2015-000059] [Medline: 27900151]
58. Fillol F, Paris L, Pascal S, et al. Possible impact of a 12-month web- and smartphone-based program to improve long-term physical activity in patients attending spa therapy: randomized controlled trial. J Med Internet Res 2022 Jun 16;24(6):e29640. [doi: 10.2196/29640] [Medline: 35708743]
59. Duarte-Rojo A, Bloomer PM, Rogers RJ, et al. Introducing EL-FIT (exercise and liver fitness): a smartphone app to prehabilitate and monitor liver transplant candidates. Liver Transpl 2021 Apr;27(4):502-512. [doi: 10.1002/lt.25950] [Medline: 33232547]

XSL•FO
RenderX

60. Zadro JR, Shirley D, Nilsen TIL, Mork PJ, Ferreira PH. Family history influences the effectiveness of home exercise in older people with chronic low back pain: a secondary analysis of a randomized controlled trial. Arch Phys Med Rehabil 2020 Aug;101(8):1322-1331. [doi: 10.1016/j.apmr.2020.03.019] [Medline: 32376326]
61. Loh KP, Sanapala C, Watson EE, et al. A single-arm pilot study of a mobile health exercise intervention (GO-EXCAP) in older patients with myeloid neoplasms. Blood Adv 2022 Jul 12;6(13):3850-3860. [doi: 10.1182/bloodadvances.2022007056] [Medline: 35320340]
62. Wu YZ, Lin JY, Wu PL, Kuo YF. Effects of a hybrid intervention combining exergaming and physical therapy among older adults in a long-term care facility. Geriatr Gerontol Int 2019 Feb;19(2):147-152. [doi: 10.1111/ggi.13575] [Medline: 30523659]
63. Yuan RY, Chen SC, Peng CW, Lin YN, Chang YT, Lai CH. Effects of interactive video-game-based exercise on balance in older adults with mild-to-moderate Parkinson's disease. J Neuroeng Rehabil 2020 Jul 13;17(1):91. [doi: 10.1186/s12984-020-00725-y] [Medline: 32660512]
64. Duscha BD, Piner LW, Patel MP, et al. Effects of a 12-week mHealth program on functional capacity and physical activity in patients with peripheral artery disease. Am J Cardiol 2018 Sep 1;122(5):879-884. [doi: 10.1016/j.amjcard.2018.05.018] [Medline: 30049464]
65. Wang CH, Chou PC, Joa WC, et al. Mobile-phone-based home exercise training program decreases systemic inflammation in COPD: a pilot study. BMC Pulm Med 2014 Aug 30;14:142. [doi: 10.1186/1471-2466-14-142] [Medline: 25175787]
66. Auerswald T, Meyer J, von Holdt K, Voelcker-Rehage C. Application of activity trackers among nursing home residents-a pilot and feasibility study on physical activity behavior, usage behavior, acceptance, usability and motivational Impact. Int J Environ Res Public Health 2020 Sep 14;17(18):6683. [doi: 10.3390/ijerph17186683] [Medline: 32937840]
67. Goode AP, Taylor SS, Hastings SN, Stanwyck C, Coffman CJ, Allen KD. Effects of a home-based telephone-supported physical activity program for older adult veterans with chronic low back pain. Phys Ther 2018 May 1;98(5):369-380. [doi: 10.1093/ptj/pzy026] [Medline: 29669086]
68. Schwartz H, Har-Nir I, Wenhoda T, Halperin I. Staying physically active during the COVID-19 quarantine: exploring the feasibility of live, online, group training sessions among older adults. Transl Behav Med 2021 Mar 16;11(2):314-322. [doi: 10.1093/tbm/ibaa141] [Medline: 33447852]
69. van den Helder J, Mehra S, van Dronkelaar C, et al. Blended home-based exercise and dietary protein in community-dwelling older adults: a cluster randomized controlled trial. J Cachexia Sarcopenia Muscle 2020 Dec;11(6):1590-1602. [doi: 10.1002/jcsm.12634] [Medline: 33103379]
70. Barnason S, Zimmerman L, Schulz P, Tu C. Influence of an early recovery telehealth intervention on physical activity and functioning after coronary artery bypass surgery among older adults with high disease burden. Heart Lung 2009;38(6):459-468. [doi: 10.1016/j.hrtlng.2009.01.010] [Medline: 19944870]
71. McAuley E, Wójcicki TR, Gothe NP, et al. Effects of a DVD-delivered exercise intervention on physical function in older adults. J Gerontol A Biol Sci Med Sci 2013 Sep;68(9):1076-1082. [doi: 10.1093/gerona/glt014] [Medline: 23401566]
72. Bao T, Noohi F, Kinnaird C, et al. Retention effects of long-term balance training with vibrotactile sensory augmentation in healthy older adults. Sensors (Basel) 2022 Apr 14;22(8):3014. [doi: 10.3390/s22083014] [Medline: 35459000]
73. Agmon M, Perry CK, Phelan E, Demiris G, Nguyen HQ. A pilot study of Wii Fit exergames to improve balance in older adults. J Geriatr Phys Ther 2011;34(4):161-167. [doi: 10.1519/JPT.0b013e3182191d98] [Medline: 22124415]
74. Moy ML, Collins RJ, Martinez CH, et al. An internet-mediated pedometer-based program improves health-related quality-of-life domains and daily step counts in COPD: a randomized controlled trial. Chest 2015 Jul;148(1):128-137. [doi: 10.1378/chest.14-1466] [Medline: 25811395]
75. Anderson-Hanley C, Arciero PJ, Brickman AM, et al. Exergaming and older adult cognition: a cluster randomized clinical trial. Am J Prev Med 2012 Feb;42(2):109-119. [doi: 10.1016/j.amepre.2011.10.016] [Medline: 22261206]
76. Coultas DB, Jackson BE, Russo R, et al. Home-based physical activity coaching, physical activity, and health care utilization in chronic obstructive pulmonary disease. Chronic obstructive pulmonary disease self-management activation research trial secondary outcomes. Ann Am Thorac Soc 2018 Apr;15(4):470-478. [doi: 10.1513/AnnalsATS.201704-308OC] [Medline: 29283670]
77. Schwenk M, Grewal GS, Honarvar B, et al. Interactive balance training integrating sensor-based visual feedback of movement performance: a pilot study in older adults. J Neuroeng Rehabil 2014 Dec 13;11:164. [doi: 10.1186/1743-0003-11-164] [Medline: 25496052]
78. Esculier JF, Vaudrin J, Bériault P, Gagnon K, Tremblay LE. Home-based balance training programme using Wii Fit with balance board for Parkinsons's disease: a pilot study. J Rehabil Med 2012 Feb;44(2):144-150. [doi: 10.2340/16501977-0922] [Medline: 22266676]
79. Snyder A, Colvin B, Gammack JK. Pedometer use increases daily steps and functional status in older adults. J Am Med Dir Assoc 2011 Oct;12(8):590-594. [doi: 10.1016/j.jamda.2010.06.007] [Medline: 21450191]
80. Wu G, Keyes L, Callas P, Ren X, Bookchin B. Comparison of telecommunication, community, and home-based Tai Chi exercise programs on compliance and effectiveness in elders at risk for falls. Arch Phys Med Rehabil 2010 Jun;91(6):849-856. [doi: 10.1016/j.apmr.2010.01.024] [Medline: 20510973]

XSL•FO
RenderX

81. Swinnen N, Vandenbulcke M, de Bruin ED, Akkerman R, Stubbs B, Vancampfort D. Exergaming for people with major neurocognitive disorder: a qualitative study. Disabil Rehabil 2022 May;44(10):2044-2052. [doi: 10.1080/09638288.2020.1822934] [Medline: 32962436]
82. Garcia JA, Schoene D, Lord SR, Delbaere K, Valenzuela T, Navarro KF. A bespoke kinect stepping exergame for improving physical and cognitive function in older people: a pilot study. Games Health J 2016 Dec;5(6):382-388. [doi: 10.1089/g4h.2016.0070] [Medline: 27860515]
83. Olafsdottir SA, Jonsdottir H, Bjartmarz I, et al. Feasibility of ActivABLES to promote home-based exercise and physical activity of community-dwelling stroke survivors with support from caregivers: a mixed methods study. BMC Health Serv Res 2020 Jun 22;20(1):562. [doi: 10.1186/s12913-020-05432-x] [Medline: 32571316]
84. Villumsen BR, Jorgensen MG, Frystyk J, Hørdam B, Borre M. Home-based "exergaming" was safe and significantly improved 6-min walking distance in patients with prostate cancer: a single-blinded randomised controlled trial. BJU Int 2019 Oct;124(4):600-608. [doi: 10.1111/bju.14782] [Medline: 31012238]
85. Zahedian-Nasab N, Jaberi A, Shirazi F, Kavousipor S. Effect of virtual reality exercises on balance and fall in elderly people with fall risk: a randomized controlled trial. BMC Geriatr 2021 Sep 25;21(1):509. [doi: 10.1186/s12877-021-02462-w] [Medline: 34563120]
86. Klompstra L, Jaarsma T, Mårtensson J, Strömberg A. Exergaming through the eyes of patients with heart failure: a qualitative content analysis study. Games Health J 2017 Jun;6(3):152-158. [doi: 10.1089/g4h.2016.0087] [Medline: 28422516]
87. Schoene D, Lord SR, Delbaere K, Severino C, Davies TA, Smith ST. A randomized controlled pilot study of home-based step training in older people using videogame technology. PLoS ONE 2013;8(3):e57734. [doi: 10.1371/journal.pone.0057734] [Medline: 23472104]
88. O'Brien T, Troutman-Jordan M, Hathaway D, Armstrong S, Moore M. Acceptability of wristband activity trackers among community dwelling older adults. Geriatr Nurs 2015;36(2 Suppl):S21-S25. [doi: 10.1016/j.gerinurse.2015.02.019] [Medline: 25771957]
89. Wu G, Keyes LM. Group tele-exercise for improving balance in elders. Telemed J E Health 2006 Oct;12(5):561-570. [doi: 10.1089/tmj.2006.12.561] [Medline: 17042710]
90. Padala KP, Padala PR, Lensing SY, et al. Home-based exercise program improves balance and fear of falling in community-dwelling older adults with mild alzheimer's disease: a pilot study. J Alzheimers Dis 2017;59(2):565-574. [doi: 10.3233/JAD-170120] [Medline: 28655135]
91. Hosteng KR, Simmering JE, Polgreen LA, et al. Multilevel mHealth intervention increases physical activity of older adults living in retirement community. J Phys Act Health 2021 Jul 1;18(7):851-857. [doi: 10.1123/jpah.2020-0592] [Medline: 34039774]
92. Golsteijn RHJ, Bolman C, Volders E, Peels DA, de Vries H, Lechner L. Short-term efficacy of a computer-tailored physical activity intervention for prostate and colorectal cancer patients and survivors: a randomized controlled trial. Int J Behav Nutr Phys Act 2018 Oct 30;15(1):106. [doi: 10.1186/s12966-018-0734-9] [Medline: 30376857]
93. Yousefi Babadi S, Daneshmandi H. Effects of virtual reality versus conventional balance training on balance of the elderly. Exp Gerontol 2021 Oct 1;153:111498. [doi: 10.1016/j.exger.2021.111498] [Medline: 34311059]
94. Swinnen N, Vandenbulcke M, de Bruin ED, et al. The efficacy of exergaming in people with major neurocognitive disorder residing in long-term care facilities: a pilot randomized controlled trial. Alzheimers Res Ther 2021 Mar 30;13(1):70. [doi: 10.1186/s13195-021-00806-7] [Medline: 33785077]
95. Jung DI, Ko DS, Jeong MA. Kinematic effect of Nintendo Wii(TM) sports program exercise on obstacle gait in elderly women with falling risk. J Phys Ther Sci 2015 May;27(5):1397-1400. [doi: 10.1589/jpts.27.1397] [Medline: 26157228]
96. Yeşilyaprak SS, Yıldırım M, Tomruk M, Ertekin Ö, Algun ZC. Comparison of the effects of virtual reality-based balance exercises and conventional exercises on balance and fall risk in older adults living in nursing homes in Turkey. Physiother Theory Pract 2016;32(3):191-201. [doi: 10.3109/09593985.2015.1138009] [Medline: 27049879]
97. Padala KP, Padala PR, Malloy TR, et al. Wii-fit for improving gait and balance in an assisted living facility: a pilot study. J Aging Res 2012;2012:597573. [doi: 10.1155/2012/597573] [Medline: 22745909]
98. Maddison R, Pfaeffli L, Whittaker R, et al. A mobile phone intervention increases physical activity in people with cardiovascular disease: Results from the HEART randomized controlled trial. Eur J Prev Cardiol 2015 Jun;22(6):701-709. [doi: 10.1177/2047487314535076] [Medline: 24817694]
99. van Diest M, Stegenga J, Wörtche HJ, Verkerke GJ, Postema K, Lamoth CJC. Exergames for unsupervised balance training at home: A pilot study in healthy older adults. Gait Posture 2016 Feb;44:161-167. [doi: 10.1016/j.gaitpost.2015.11.019] [Medline: 27004651]
100. Silveira P, van het Reve E, Daniel F, Casati F, de Bruin ED. Motivating and assisting physical exercise in independently living older adults: a pilot study. Int J Med Inform 2013 May;82(5):325-334. [doi: 10.1016/j.ijmedinf.2012.11.015] [Medline: 23273420]
101. Li F, Harmer P, Voit J, Chou LS. Implementing an online virtual falls prevention intervention during a public health pandemic for older adults with mild cognitive impairment: a feasibility trial. Clin Interv Aging 2021;16:973-983. [doi: 10.2147/CIA.S306431] [Medline: 34079243]

XSL•FO
RenderX

102. Khaghani Far I, Ferron M, Ibarra F, et al. The interplay of physical and social wellbeing in older adults: investigating the relationship between physical training and social interactions with virtual social environments. PeerJ Comput Sci ;1:e30. [doi: 10.7717/peerj-cs.30]
103. Cinini A, Cutugno P, Ferraris C, et al. Final results of the NINFA project: impact of new technologies in the daily life of elderly people. Aging Clin Exp Res 2021 May;33(5):1213-1222. [doi: 10.1007/s40520-019-01357-6] [Medline: 31587153]
104. McDermott MM, Spring B, Berger JS, et al. Effect of a home-based exercise intervention of wearable technology and telephone coaching on walking performance in peripheral artery disease: the HONOR randomized clinical trial. JAMA 2018 Apr 24;319(16):1665-1676. [doi: 10.1001/jama.2018.3275] [Medline: 29710165]
105. Song J, Paul SS, Caetano MJD, et al. Home-based step training using videogame technology in people with Parkinson's disease: a single-blinded randomised controlled trial. Clin Rehabil 2018 Mar;32(3):299-311. [doi: 10.1177/0269215517721593] [Medline: 28745063]
106. Chao YY, Scherer YK, Montgomery CA, Wu YW, Lucke KT. Physical and psychosocial effects of Wii Fit exergames use in assisted living residents: a pilot study. Clin Nurs Res 2015 Dec;24(6):589-603. [doi: 10.1177/1054773814562880] [Medline: 25488422]
107. Geraedts HAE, Zijlstra W, Zhang W, et al. A home-based exercise program driven by tablet application and mobility monitoring for frail older adults: feasibility and practical implications. Prev Chronic Dis 2017 Feb 2;14:E12. [doi: 10.5888/pcd14.160227] [Medline: 28152361]
108. Pischke CR, Voelcker-Rehage C, Ratz T, et al. Web-based versus print-based physical activity intervention for community-dwelling older adults: crossover randomized trial. JMIR Mhealth Uhealth 2022 Mar 23;10(3):e32212. [doi: 10.2196/32212] [Medline: 35319484]
109. Batsis JA, Petersen CL, Clark MM, et al. Feasibility and acceptability of a technology-based, rural weight management intervention in older adults with obesity. BMC Geriatr 2021 Jan 12;21(1):44. [doi: 10.1186/s12877-020-01978-x] [Medline: 33435877]
110. Gschwind YJ, Eichberg S, Ejupi A, et al. ICT-based system to predict and prevent falls (iStoppFalls): results from an international multicenter randomized controlled trial. Eur Rev Aging Phys Act 2015;12:10. [doi: 10.1186/s11556-015-0155-6] [Medline: 26865874]
111. Jansons P, Dalla Via J, Daly RM, Fyfe JJ, Gvozdenko E, Scott D. Delivery of home-based exercise interventions in older adults facilitated by amazon alexa: a 12-week feasibility trial. J Nutr Health Aging 2022;26(1):96-102. [doi: 10.1007/s12603-021-1717-0] [Medline: 35067710]
112. Rodrigues EV, Gallo LH, Guimarães ATB, Melo Filho J, Luna BC, Gomes ARS. Effects of dance exergaming on depressive symptoms, fear of falling, and musculoskeletal function in fallers and nonfallers community-dwelling older women. Rejuvenation Res 2018 Dec;21(6):518-526. [doi: 10.1089/rej.2017.2041] [Medline: 29669458]
113. Morrison S, Simmons R, Colberg SR, Parson HK, Vinik AI. Supervised balance training and Wii Fit-based exercises lower falls risk in older adults with type 2 diabetes. J Am Med Dir Assoc 2018 Feb;19(2):185. [doi: 10.1016/j.jamda.2017.11.004] [Medline: 29289541]
114. Muñoz GF, Cardenas RAM, Pla F. A kinect-based interactive system for home-assisted active aging. Sensors (Basel) 2021 Jan 8;21(2):417. [doi: 10.3390/s21020417] [Medline: 33435557]
115. Brickwood KJ, Ahuja KDK, Watson G, O'Brien JA, Williams AD. Effects of activity tracker use with health professional support or telephone counseling on maintenance of physical activity and health outcomes in older adults: randomized controlled trial. JMIR Mhealth Uhealth 2021 Jan 5;9(1):e18686. [doi: 10.2196/18686] [Medline: 33399541]
116. Peng X, Su Y, Hu Z, et al. Home-based telehealth exercise training program in Chinese patients with heart failure: a randomized controlled trial. Medicine (Baltimore) 2018 Aug;97(35):e12069. [doi: 10.1097/MD.0000000000012069] [Medline: 30170422]
117. Mehra S, van den Helder J, Visser B, Engelbert RHH, Weijs PJM, Kröse BJA. Evaluation of a blended physical activity intervention for older adults: mixed methods study. J Med Internet Res 2020 Jul 23;22(7):e16380. [doi: 10.2196/16380] [Medline: 32459652]
118. Lewis ZH, Ottenbacher KJ, Fisher SR, et al. The feasibility and RE-AIM evaluation of the TAME health pilot study. Int J Behav Nutr Phys Act 2017 Aug 14;14(1):106. [doi: 10.1186/s12966-017-0560-5] [Medline: 28807041]
119. Yin Z, Martinez CE, Li S, et al. Adapting Chinese qigong mind-body exercise for healthy aging in older community-dwelling low-income Latino adults: pilot feasibility study. JMIR Aging 2021 Nov 1;4(4):e29188. [doi: 10.2196/29188] [Medline: 34723824]
120. Klompstra L, Jaarsma T, Strömberg A. Exergaming to increase the exercise capacity and daily physical activity in heart failure patients: a pilot study. BMC Geriatr 2014 Nov 18;14:119. [doi: 10.1186/1471-2318-14-119] [Medline: 25407612]
121. Irvine AB, Gelatt VA, Seeley JR, Macfarlane P, Gau JM. Web-based intervention to promote physical activity by sedentary older adults: randomized controlled trial. J Med Internet Res 2013 Feb 5;15(2):e19. [doi: 10.2196/jmir.2158] [Medline: 23470322]
122. Mansson L, Lundin-Olsson L, Skelton DA, et al. Older adults' preferences for, adherence to and experiences of two self-management falls prevention home exercise programmes: a comparison between a digital programme and a paper booklet. BMC Geriatr 2020 Jun 15;20(1):209. [doi: 10.1186/s12877-020-01592-x] [Medline: 32539711]

XSL•FO
RenderX

123. Anderson-Hanley C, Stark J, Wall KM, et al.): effects of a 3-month in-home pilot clinical trial for mild cognitive impairment and caregivers. Clin Interv Aging 2018;13:1565-1577. [doi: 10.2147/CIA.S160756] [Medline: 30233154]
124. Martel D, Lauzé M, Agnoux A, et al. Comparing the effects of a home-based exercise program using a gerontechnology to a community-based group exercise program on functional capacities in older adults after a minor injury. Exp Gerontol 2018 Jul 15;108:41-47. [doi: 10.1016/j.exger.2018.03.016] [Medline: 29577975]
125. Chao YY, Scherer YK, Montgomery CA, Lucke KT, Wu YW. Exergames-based intervention for assisted living residents: a pilot study. J Gerontol Nurs 2014 Nov;40(11):36-43. [doi: 10.3928/00989134-20140407-04] [Medline: 24716646]
126. Baez M, Khaghani Far I, Ibarra F, Ferron M, Didino D, Casati F. Effects of online group exercises for older adults on physical, psychological and social wellbeing: a randomized pilot trial. PeerJ 2017;5:e3150. [doi: 10.7717/peerj.3150] [Medline: 28392983]
127. Auerswald T, Hendker A, Ratz T, et al. Impact of activity tracker usage in combination with a physical activity intervention on physical and cognitive parameters in healthy adults aged 60+: a randomized controlled trial. Int J Environ Res Public Health 2022 Mar 22;19(7):3785. [doi: 10.3390/ijerph19073785] [Medline: 35409466]
128. Netz Y, Yekutieli Z, Arnon M, et al. Personalized exercise programs based upon remote assessment of motor fitness: a pilot study among healthy people aged 65 years and older. Gerontology 2022;68(4):465-479. [doi: 10.1159/000517918] [Medline: 34515118]
129. Gandolfi M, Geroin C, Dimitrova E, et al. Virtual reality telerehabilitation for postural instability in parkinson's disease: a multicenter, single-blind, randomized, controlled trial. Biomed Res Int 2017;2017:7962826. [doi: 10.1155/2017/7962826] [Medline: 29333454]
130. Ambrens M, van Schooten KS, Lung T, et al. Economic evaluation of the e-Health StandingTall balance exercise programme for fall prevention in people aged 70 years and over. Age Ageing 2022 Jun 1;51(6):afac130. [doi: 10.1093/ageing/afac130] [Medline: 35679193]
131. Baez M, Ibarra F, Far IK, Ferron M, Casati F. 2016 international conference on collaboration technologies and systems (CTS). Online group-exercises for older adults of different physical abilities . [doi: 10.1109/CTS.2016.0098]
132. Monteblanco Cavalcante M, Fraga I, Dalbosco B, et al. Exergame training-induced neuroplasticity and cognitive improvement in institutionalized older adults: a preliminary investigation. Physiol Behav 2021 Nov 1;241:113589. [doi: 10.1016/j.physbeh.2021.113589] [Medline: 34509470]
133. Wall K, Stark J, Schillaci A, et al. The enhanced interactive physical and cognitive exercise system (iPACESTM v2.0): pilot clinical trial of an in-home ipad-based neuro-exergame for mild cognitive impairment (MCI). J Clin Med 2018 Aug 30;7(9):249. [doi: 10.3390/jcm7090249] [Medline: 30200183]
134. Shubert TE, Chokshi A, Mendes VM, et al. Stand tall-a virtual translation of the otago exercise program. J Geriatr Phys Ther 2020;43(3):120-127. [doi: 10.1519/JPT.0000000000000203] [Medline: 29958232]
135. Lynch BM, Nguyen NH, Moore MM, et al. A randomized controlled trial of a wearable technology-based intervention for increasing moderate to vigorous physical activity and reducing sedentary behavior in breast cancer survivors: the ACTIVATE trial. Cancer 2019 Aug 15;125(16):2846-2855. [doi: 10.1002/cncr.32143] [Medline: 31012970]
136. Yang WC, Wang HK, Wu RM, Lo CS, Lin KH. Home-based virtual reality balance training and conventional balance training in Parkinson's disease: a randomized controlled trial. J Formos Med Assoc 2016 Sep;115(9):734-743. [doi: 10.1016/j.jfma.2015.07.012] [Medline: 26279172]
137. Kim T, Xiong S. Effectiveness and usability of a novel kinect-based tailored interactive fall intervention system for fall prevention in older people: a preliminary study. Front Public Health 2022;10:884551. [doi: 10.3389/fpubh.2022.884551] [Medline: 35712291]
138. Callisaya ML, Jayakody O, Vaidya A, Srikanth V, Farrow M, Delbaere K. A novel cognitive-motor exercise program delivered via a tablet to improve mobility in older people with cognitive impairment - standingtall cognition and mobility. Exp Gerontol 2021 Sep;152:111434. [doi: 10.1016/j.exger.2021.111434] [Medline: 34098009]
139. Sparrow D, Gottlieb DJ, Demolles D, Fielding RA. Increases in muscle strength and balance using a resistance training program administered via a telecommunications system in older adults. J Gerontol A Biol Sci Med Sci 2011 Nov;66(11):1251-1257. [doi: 10.1093/gerona/glr138] [Medline: 21852283]
140. Jennings SC, Manning KM, Bettger JP, et al. Rapid transition to telehealth group exercise and functional assessments in response to COVID-19. Gerontol Geriatr Med 2020;6:2333721420980313. [doi: 10.1177/2333721420980313] [Medline: 33403222]
141. Flynn A, Preston E, Dennis S, Canning CG, Allen NE. Home-based exercise monitored with telehealth is feasible and acceptable compared to centre-based exercise in Parkinson's disease: a randomised pilot study. Clin Rehabil 2021 May;35(5):728-739. [doi: 10.1177/0269215520976265] [Medline: 33272025]
142. Szanton SL, Walker RK, Lim JH, et al. Development of an exergame for urban-dwelling older adults with functional limitations: results and lessons learned. Prog Community Health Partnersh 2016;10(1):73-81. [doi: 10.1353/cpr.2016.0005] [Medline: 27018356]
143. Alagumoorthi G, D BJ, Thirunavukarasu S, V R, A K. Effectiveness of Wii sports- based strategy training in reducing risk of falling, falls and improving quality of life in adults with idiopathic Parkinson's disease- a randomized comparative trial. Clin Rehabil 2022 Aug;36(8):1097-1109. [doi: 10.1177/02692155221089030] [Medline: 35315706]

XSL•FO
RenderX

144. Hong J, Kong HJ, Yoon HJ. Web-based telepresence exercise program for community-dwelling elderly women with a high risk of falling: randomized controlled trial. JMIR Mhealth Uhealth 2018 May 28;6(5):e132. [doi: 10.2196/mhealth.9563] [Medline: 29807877]
145. Muellmann S, Buck C, Voelcker-Rehage C, et al. Effects of two web-based interventions promoting physical activity among older adults compared to a delayed intervention control group in Northwestern Germany: Results of the PROMOTE community-based intervention trial. Prev Med Rep 2019 Sep;15:100958. [doi: 10.1016/j.pmedr.2019.100958] [Medline: 31410347]
146. Ellmers TJ, Paraskevopoulos IT, Williams AM, Young WR. Recalibrating disparities in perceived and actual balance abilities in older adults: a mixed-methods evaluation of a novel exergaming intervention. J Neuroeng Rehabil 2018 Mar 22;15(1):26. [doi: 10.1186/s12984-018-0369-8] [Medline: 29566720]
147. Ogawa EF, Huang H, Yu LF, et al. Effects of exergaming on cognition and gait in older adults at risk for falling. Med Sci Sports Exerc 2020 Mar;52(3):754-761. [doi: 10.1249/MSS.0000000000002167] [Medline: 31652241]
148. Daly RM, Gianoudis J, Hall T, Mundell NL, Maddison R. Feasibility, usability, and enjoyment of a home-based exercise program delivered via an exercise app for musculoskeletal health in community-dwelling older adults: short-term prospective pilot study. JMIR Mhealth Uhealth 2021 Jan 13;9(1):e21094. [doi: 10.2196/21094] [Medline: 33439147]
149. Soancatl Aguilar V, Lamoth CJC, Maurits NM, Roerdink J. Assessing dynamic postural control during exergaming in older adults: A probabilistic approach. Gait Posture 2018 Feb;60:235-240. [doi: 10.1016/j.gaitpost.2017.12.015] [Medline: 29288962]
150. Batsis JA, Naslund JA, Gill LE, Masutani RK, Agarwal N, Bartels SJ. Use of a wearable activity device in rural older obese adults: a pilot study. Gerontol Geriatr Med 2016;2:2333721416678076. [doi: 10.1177/2333721416678076] [Medline: 28138502]
151. Keogh JWL, Power N, Wooller L, Lucas P, Whatman C. Physical and psychosocial function in residential aged-care elders: effect of Nintendo Wii Sports games. J Aging Phys Act 2014 Apr;22(2):235-244. [doi: 10.1123/japa.2012-0272] [Medline: 23752164]
152. Adcock M, Thalmann M, Schättin A, Gennaro F, de Bruin ED. A pilot study of an in-home multicomponent exergame training for older adults: feasibility, usability and pre-post evaluation. Front Aging Neurosci 2019;11:304. [doi: 10.3389/fnagi.2019.00304] [Medline: 31824295]
153. Jessen JD, Lund HH. Playful home training for falls prevention. : Institute of Electrical and Electronics Engineers Inc; 2015 Presented at: 2015 IEEE International Conference on Advanced Intelligent Mechatronics (AIM); Busan, South Korea p. 311-317. [doi: 10.1109/AIM.2015.7222550]
154. Saenz-de-Urturi Z, Garcia-Zapirain Soto B. Kinect-based virtual game for the elderly that detects incorrect body postures in real time. Sensors (Basel) 2016 May 16;16(5):704. [doi: 10.3390/s16050704] [Medline: 27196903]
155. Vaziri DD, Aal K, Ogonowski C, et al. Exploring user experience and technology acceptance for a fall prevention system: results from a randomized clinical trial and a living lab. Eur Rev Aging Phys Act 2016;13:6. [doi: 10.1186/s11556-016-0165-z] [Medline: 27293489]
156. Vallabhajosula S, McMillion AK, Freund JE. The effects of exergaming and treadmill training on gait, balance, and cognition in a person with Parkinson's disease: A case study. Physiother Theory Pract 2017 Dec;33(12):920-931. [doi: 10.1080/09593985.2017.1359867] [Medline: 28812419]
157. Blair CK, Harding E, Wiggins C, et al. A home-based mobile health intervention to replace sedentary time with light physical activity in older cancer survivors: randomized controlled pilot trial. JMIR Cancer 2021 Apr 13;7(2):e18819. [doi: 10.2196/18819] [Medline: 33847588]
158. Gooch HJ, Jarvis KA, Stockley RC. Behavior change approaches in digital technology-based physical rehabilitation interventions following stroke: scoping review. J Med Internet Res 2024 Apr 24;26:e48725. [doi: 10.2196/48725] [Medline: 38656777]
159. Fu AS, Gao KL, Tung AK, Tsang WW, Kwan MM. Effectiveness of exergaming training in reducing risk and incidence of falls in frail older adults with a history of falls. Arch Phys Med Rehabil 2015 Dec;96(12):2096-2102. [doi: 10.1016/j.apmr.2015.08.427] [Medline: 26360975]
160. Barisch-Fritz B, Bezold J, Scharpf A, Trautwein S, Krell-Roesch J, Woll A. ICT-based individualized training of institutionalized individuals with dementia. evaluation of usability and trends toward the effectiveness of the InCoPE-app. Front Physiol 2022;13:921105. [doi: 10.3389/fphys.2022.921105] [Medline: 35874545]
161. Uzor S, Baillie L. Investigating the Long-Term Use of Exergames in the Home with Elderly Fallers: CHI '14; 2014. URL: https://dl.acm.org/doi/proceedings/10.1145/2556288 [doi: 10.1145/2556288.2557160] [Medline: 25283179]
162. Lauzé M, Martel DD, Aubertin-Leheudre M. Feasibility and effects of a physical activity program using gerontechnology in assisted living communities for older adults. J Am Med Dir Assoc 2017 Dec 1;18(12):1069-1075. [doi: 10.1016/j.jamda.2017.06.030] [Medline: 28818422]
163. Bao T, Carender WJ, Kinnaird C, et al. Effects of long-term balance training with vibrotactile sensory augmentation among community-dwelling healthy older adults: a randomized preliminary study. J Neuroeng Rehabil 2018 Jan 18;15(1):5. [doi: 10.1186/s12984-017-0339-6] [Medline: 29347946]

XSL•FO
RenderX

164. Hsu YI, Chen YC, Lee CL, Chang NJ. Effects of diet control and telemedicine-based resistance exercise intervention on patients with obesity and knee osteoarthritis: a randomized control trial. Int J Environ Res Public Health 2021 Jul 21;18(15):7744. [doi: 10.3390/ijerph18157744] [Medline: 34360036]
165. Middleton A, Simpson KN, Bettger JP, Bowden MG. COVID-19 pandemic and beyond: considerations and costs of telehealth exercise programs for older adults with functional impairments living at home-lessons learned from a pilot case study. Phys Ther 2020 Aug 12;100(8):1278-1288. [doi: 10.1093/ptj/pzaa089] [Medline: 32372072]
166. Studenski S, Perera S, Hile E, Keller V, Spadola-Bogard J, Garcia J. Interactive video dance games for healthy older adults. J Nutr Health Aging 2010 Dec;14(10):850-852. [doi: 10.1007/s12603-010-0119-5] [Medline: 21125204]
167. Radhakrishnan K, Julien C, O'Hair M, et al. Usability testing of a sensor-controlled digital game to engage older adults with heart failure in physical activity and weight monitoring. Appl Clin Inform 2020 Oct;11(5):873-881. [doi: 10.1055/s-0040-1721399] [Medline: 33378780]
168. Adcock M, Fankhauser M, Post J, et al. Effects of an in-home multicomponent exergame training on physical functions, cognition, and brain volume of older adults: a randomized controlled trial. Front Med (Lausanne) 2019;6:321. [doi: 10.3389/fmed.2019.00321] [Medline: 32047751]
169. Jung Y, Li KJ, Janissa NS, Gladys WLC, Lee KM. Games for a better life. In: IE '09 2009. URL: https://dl.acm.org/doi/proceedings/10.1145/1746050 [doi: 10.1145/1746050.1746055]
170. Ibrahim F, Usman J, Hamzah N, Yazed Ahmad M, Teh SJ, editors. 2nd International Conference for Innovation in Biomedical Engineering and Life Sciences 2018. URL: http://link.springer.com/10.1007/978-981-10-7554-4 [doi: 10.1007/978-981-10-7554-4_6]
171. Smith ST, Davies TA, Lennox J. Step training system: an ICT solution to measure and reduce fall risk in older adults. Annu Int Conf IEEE Eng Med Biol Soc 2013;2013:7033-7035. [doi: 10.1109/EMBC.2013.6611177] [Medline: 24111364]
172. Weinstock RS, Brooks G, Palmas W, et al. Lessened decline in physical activity and impairment of older adults with diabetes with telemedicine and pedometer use: results from the IDEATel study. Age Ageing 2011 Jan;40(1):98-105. [doi: 10.1093/ageing/afq147] [Medline: 21081539]
173. Pettersson B, Wiklund M, Janols R, et al. "Managing pieces of a personal puzzle" - older people's experiences of self-management falls prevention exercise guided by a digital program or a booklet. BMC Geriatr 2019 Feb 18;19(1):43. [doi: 10.1186/s12877-019-1063-9] [Medline: 30777026]
174. Zalecki T, Gorecka-Mazur A, Pietraszko W, et al. Visual feedback training using WII Fit improves balance in Parkinson's disease. Folia Med Cracov 2013;53(1):65-78. [Medline: 24858332]
175. Dekker-van Weering M, Jansen-Kosterink S, Frazer S, Vollenbroek-Hutten M. User experience, actual use, and effectiveness of an information communication technology-supported home exercise program for pre-frail older adults. Front Med 2017;4. [doi: 10.3389/fmed.2017.00208]
176. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact 2008 Jul 29;24(6):574-594. [doi: 10.1080/10447310802205776]
177. Pajalic Z, de Sousa DA, Strøm BS, et al. Welfare technology interventions among older people living at home-a systematic review of RCT studies. PLOS Digit Health 2023 Jan;2(1):e0000184. [doi: 10.1371/journal.pdig.0000184] [Medline: 36812629]
178. Kim B, Hunt M, Muscedere J, Maslove DM, Lee J. Using consumer-grade physical activity trackers to measure frailty transitions in older critical care survivors: exploratory observational study. JMIR Aging 2021 Feb 23;4(1):e19859. [doi: 10.2196/19859] [Medline: 33620323]
179. Oesch P, Kool J, Fernandez-Luque L, et al. Exergames versus self-regulated exercises with instruction leaflets to improve adherence during geriatric rehabilitation: a randomized controlled trial. BMC Geriatr 2017 Mar 23;17(1):77. [doi: 10.1186/s12877-017-0467-7] [Medline: 28330455]
180. Laver K, George S, Ratcliffe J, et al. Use of an interactive video gaming program compared with conventional physiotherapy for hospitalised older adults: a feasibility trial. Disabil Rehabil 2012;34(21):1802-1808. [doi: 10.3109/09638288.2012.662570] [Medline: 22409245]
181. Ortiz-Alonso J, Bustamante-Ara N, Valenzuela PL, et al. Effect of a simple exercise program on hospitalization-associated disability in older patients: a randomized controlled trial. J Am Med Dir Assoc 2020 Apr;21(4):531-537. [doi: 10.1016/j.jamda.2019.11.027] [Medline: 31974063]
182. Juneau A, Bolduc A, Nguyen P, et al. Feasibility of implementing an exercise program in a geriatric assessment unit: the SPRINT program. Can Geriatr J 2018 Sep;21(3):284-289. [doi: 10.5770/cgj.21.311] [Medline: 30271514]
183. Scheerman K, Raaijmakers K, Otten RHJ, Meskers CGM, Maier AB. Effect of physical interventions on physical performance and physical activity in older patients during hospitalization: a systematic review. BMC Geriatr 2018 Nov 23;18(1):288. [doi: 10.1186/s12877-018-0965-2] [Medline: 30470205]
184. Estrela M, Semedo G, Roque F, Ferreira PL, Herdeiro MT. Sociodemographic determinants of digital health literacy: a systematic review and meta-analysis. Int J Med Inform 2023 Sep;177:105124. [doi: 10.1016/j.ijmedinf.2023.105124] [Medline: 37329766]
185. Wang X, Luan W. Research progress on digital health literacy of older adults: a scoping review. Front Public Health 2022;10:906089. [doi: 10.3389/fpubh.2022.906089]

XSL•FO
RenderX

186. Kebede AS, Ozolins LL, Holst H, Galvin K. Digital engagement of older adults: scoping review. J Med Internet Res 2022 Dec 7;24(12):e40192. [doi: 10.2196/40192] [Medline: 36477006]
187. Choi NG, Dinitto DM. The digital divide among low-income homebound older adults: Internet use patterns, eHealth literacy, and attitudes toward computer/Internet use. J Med Internet Res 2013 May 2;15(5):e93. [doi: 10.2196/jmir.2645] [Medline: 23639979]
188. Borghouts J, Eikey E, Mark G, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021 Mar 24;23(3):e24387. [doi: 10.2196/24387] [Medline: 33759801]
189. Manser P, Herold F, de Bruin ED. Components of effective exergame-based training to improve cognitive functioning in middle-aged to older adults – a systematic review and meta-analysis. Ageing Res Rev 2024 Aug;99(102385):102385. [doi: 10.1016/j.arr.2024.102385]
190. Torre MM, Temprado JJ. Effects of exergames on brain and cognition in older adults: a review based on a new categorization of combined training intervention. Front Aging Neurosci 2022 Mar 30. [doi: 10.3389/FNAGI.2022.859715/PDF]
191. Wollesen B, Herden M, Lamberti N, Giannaki CD. Defining and reporting exercise intensity in interventions for older adults: a modified Delphi process. Eur Rev Aging Phys Act 2024 Feb 2;21(1):3. [doi: 10.1186/s11556-024-00337-8] [Medline: 38302886]
192. VandeBunte A, Gontrum E, Goldberger L, et al. Physical activity measurement in older adults: wearables versus self-report. Front Digit Health 2022;4:869790. [doi: 10.3389/fdgth.2022.869790] [Medline: 36120711]
193. Matsuura H, Mukaino M, Otaka Y, et al. Validity of simplified, calibration-less exercise intensity measurement using resting heart rate during sleep: a method-comparison study with respiratory gas analysis. BMC Sports Sci Med Rehabil 2019;11(1):27. [doi: 10.1186/s13102-019-0140-x] [Medline: 31700643]
194. Knippenberg E, Timmermans A, Palmaers S, Spooren A. Use of a technology-based system to motivate older adults in performing physical activity: a feasibility study. BMC Geriatr 2021 Jan 28;21(1):81. [doi: 10.1186/s12877-021-02021-3] [Medline: 33509098]
195. Calcaterra V, Vandoni M, Marin L, et al. Exergames to limit weight gain and to fight sedentarism in children and adolescents with obesity. Children (Basel) 2023 May 24;10(6):928 [FREE Full text] [doi: 10.3390/children10060928] [Medline: 37371160]
196. Benzing V, Schmidt M. Exergaming for children and adolescents: strengths, weaknesses, opportunities and threats. J Clin Med 2018 Nov 8;7(11):422. [doi: 10.3390/jcm7110422] [Medline: 30413016]
197. Garcia Reyes EP, Kelly R, Buchanan G, Waycott J. Understanding older adults' experiences with technologies for health self-management: interview study. JMIR Aging 2023 Mar 21;6:e43197. [doi: 10.2196/43197] [Medline: 36943333]
198. Krishnan S, Mandala MA, Wolf SL, Howard A, Kesar TM. Perceptions of stroke survivors regarding factors affecting adoption of technology and exergames for rehabilitation. PM R 2023 Nov;15(11):1403-1410. [doi: 10.1002/pmrj.12963]
199. Yusif S, Soar J, Hafeez-Baig A. Older people, assistive technologies, and the barriers to adoption: a systematic review. Int J Med Inform 2016 Oct;94:112-116. [doi: 10.1016/j.ijmedinf.2016.07.004] [Medline: 27573318]
200. CDC. Telehealth interventions to improve chronic disease. 2022. URL: https://www.cdc.gov/cardiovascular-resources/php/data-research/telehealth.html [accessed 2025-08-28]
201. Tanaka K, Sheehan G, Holash D JR, Katz L. A comparison of exergaming interfaces for use in rehabilitation programs and research. Loading – The Journal of the Canadian Game Studies Association 2012;6(9) [FREE Full text]
202. Oda N, Kim JH, Yamamoto Y. Information presentation methods for setting achievable and meaningful goals on fitness apps. 2022 Sep 7 Presented at: GoodIT 2022; Sep 7-9, 2022; Limassol Cyprus p. 61-67 URL: https://dl.acm.org/doi/proceedings/10.1145/3524458 [doi: 10.1145/3524458.3547237]
203. Munoz JE, Cao S, Boger J. Kinematically adaptive exergames: personalizing exercise therapy through closed-loop systems. Presented at: 2019 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR); Dec 9-11, 2019; San Diego, CA, USA p. 118-1187 URL: https://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=8938596 [doi: 10.1109/AIVR46125.2019.00026]
204. Abedi A, Colella TJF, Pakosh M, Khan SS. Artificial intelligence-driven virtual rehabilitation for people living in the community: a scoping review. NPJ Digit Med 2024 Feb 3;7(1):25. [doi: 10.1038/s41746-024-00998-w] [Medline: 38310158]
205. Delbaere K, Valenzuela T, Lord SR, et al. E-health standingtall balance exercise for fall prevention in older people: results of a two year randomised controlled trial. BMJ 2021;373:n740. [doi: 10.1136/bmj.n740]
206. Fischl C, Blusi M, Lindgren H, Nilsson I. Tailoring to support digital technology-mediated occupational engagement for older adults - a multiple case study. Scand J Occup Ther 2020 Nov;27(8):577-590. [doi: 10.1080/11038128.2020.1760347] [Medline: 32396419]
207. Collado-Mateo D, Lavín-Pérez AM, Peñacoba C, et al. Key factors associated with adherence to physical exercise in patients with chronic diseases and older adults: an umbrella review. Int J Environ Res Public Health 2021 Feb 19;18(4):2023. [doi: 10.3390/ijerph18042023] [Medline: 33669679]
208. Tuena C, Pedroli E, Trimarchi PD, et al. Usability issues of clinical and research applications of virtual reality in older people: a systematic review. Front Hum Neurosci 2020;14:93. [doi: 10.3389/fnhum.2020.00093] [Medline: 32322194]
209. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. J Pharmacol Pharmacother 2010 Jul;1(2):100-107. [doi: 10.4103/0976-500X.72352] [Medline: 21350618]

XSL•FO
RenderX

210. Butcher NJ, Monsour A, Mew EJ, et al. Guidelines for reporting outcomes in trial protocols. JAMA 2022 Dec 20;328(23):2345. [doi: 10.1001/jama.2022.21243]
211. Iancu I, Iancu B. Designing mobile technology for elderly. a theoretical overview. Technol Forecast Soc Change 2020 Jun;155:119977. [doi: 10.1016/j.techfore.2020.119977]
212. Borghese NA, Murray D, Paraschiv-Ionescu A, et al. Rehabilitation at Home: A Comprehensive Technological Approach 2014:289-319. [doi: 10.1007/978-3-642-54816-1_16]
213. LaMonica HM, Davenport TA, Roberts AE, Hickie IB. Understanding technology preferences and requirements for health information technologies designed to improve and maintain the mental health and well-being of older adults: participatory design study. JMIR Aging 2021 Jan 6;4(1):e21461. [doi: 10.2196/21461] [Medline: 33404509]
214. Ringgenberg N, Mildner S, Hapig M, et al. ExerG: adapting an exergame training solution to the needs of older adults using focus group and expert interviews. J Neuroeng Rehabil 2022 Aug 16;19(1):89. [doi: 10.1186/s12984-022-01063-x] [Medline: 35974409]
215. Bird M, McGillion M, Chambers EM, et al. A generative co-design framework for healthcare innovation: development and application of an end-user engagement framework. Res Involv Engagem 2021 Mar 1;7(1):12. [doi: 10.1186/s40900-021-00252-7] [Medline: 33648588]
216. Peters MDJ, Marnie C, Colquhoun H, et al. Scoping reviews: reinforcing and advancing the methodology and application. Syst Rev 2021 Oct 8;10(1):263. [doi: 10.1186/s13643-021-01821-3] [Medline: 34625095]
217. Valenzuela T, Razee H, Schoene D, Lord SR, Delbaere K. An interactive home-based cognitive-motor step training program to reduce fall risk in older adults: qualitative descriptive study of older adults' experiences and requirements. JMIR Aging 2018 Nov 30;1(2):e11975. [doi: 10.2196/11975] [Medline: 31518251]
218. Hong J, Kim SW, Joo H, Kong HJ. Effects of smartphone mirroring-based telepresence exercise on body composition and physical function in obese older women. Aging Clin Exp Res 2022 May;34(5):1113-1121. [doi: 10.1007/s40520-021-02033-4] [Medline: 35028918]
219. Park J, Heilman KJ, Sullivan M, et al. Remotely supervised home-based online chair yoga intervention for older adults with dementia: feasibility study. Complement Ther Clin Pract 2022 Aug;48:101617. [doi: 10.1016/j.ctcp.2022.101617]
220. Pollock D, Davies EL, Peters MDJ, et al. Undertaking a scoping review: a practical guide for nursing and midwifery students, clinicians, researchers, and academics. J Adv Nurs 2021 Apr;77(4):2102-2113. [doi: 10.1111/jan.14743] [Medline: 33543511]

Abbreviations

CONSORT: Consolidated Standard of Reporting Trials
COST: Cooperation in Science and Technology
JBI: Joanna Briggs Institute
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-analyses extension for scoping reviews
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials

Edited by Q Fan; submitted 25.08.24; peer-reviewed by M Gasmi, P Manser, RS Gomaa Mahmoud; revised version received 19.12.24; accepted 25.12.24; published 15.09.25.

Please cite as:
Dubbeldam R, Stemplewski R, Pavlova I, Cyma-Wejchenig M, Lee S, Esser P, Bentlage E, Alcan V, Çevik ÖS, Epiphaniou E, Gallè F, Langeard A, Gafner S, Ahmed M, Bandaru N, Erden Güner A, Göz E, Kara I, Kabuk A, Türkoglu I, Pajalic Z, Vindiš J, Jaksic D, Verep U, Chouvarda I, Simovska V, Netz Y, Pelclova J
Technology-Assisted Physical Activity Interventions for Older People in Their Home-Based Environment: Scoping Review
JMIR Aging 2025;8:e65746
URL: https://aging.jmir.org/2025/1/e65746
doi:10.2196/65746

© Rosemary Dubbeldam, Rafal Stemplewski, Iuliia Pavlova, Magdalena Cyma-Wejchenig, Sunwoo Lee, Patrick Esser, Ellen Bentlage, Veysel Alcan, Özge Selin Çevik, Eleni Epiphaniou, Francesca Gallè, Antoine Langeard, Simone Gafner, Mona Ahmed, Niharika Bandaru, Arzu Erden Güner, Evrim Göz, Ilke Kara, Ayşe Kabuk, Ilayda Türkoglu, Zada Pajalic, Jan Vindiš, Damjan Jaksic, U ur Verep, Ioanna Chouvarda, Vera Simovska, Yael Netz, Jana Pelclova. Originally published in JMIR Aging (https://aging.jmir.org), 15.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using Artificial Intelligence–Based Technologies for the Early Detection of Behavioral and Psychological Symptoms of Dementia: Scoping Review

Sofia Fernandes[1,2,3], MScN; Joëlle Rosselet Amoussou[4], MSci; Carla Gomes da Rocha[1], PhD; Elodie Perruchoud[1], MScN; Armin von Gunten[5], MPhil; Cédric Mabire[3], PhD; Henk Verloo[1], PhD

[1]School of Health Sciences, University of Applied Sciences and Arts Western Switzerland (HES-SO) Valais-Wallis, Chemin de l'Agasse 5, Sion, Switzerland

[2]Les Maisons de la Providence Nursing Home, Montagnier, Switzerland

[3]Lausanne University Hospital and University of Lausanne Faculty of Biology and Medicine, Institute of Higher Education and Research in Healthcare, Lausanne, Switzerland

[4]Medical Library-Cery, Lausanne University Hospital and University of Lausanne, Site de Cery, Prilly, Switzerland

[5]Service of Old Age Psychiatry, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland

Corresponding Author:
Sofia Fernandes, MScN
School of Health Sciences, University of Applied Sciences and Arts Western Switzerland (HES-SO) Valais-Wallis, Chemin de l'Agasse 5, Sion, Switzerland

Abstract

Background: People with dementia commonly display behavioral and psychological symptoms, which have multiple negative consequences. Artificial intelligence–based technologies (AITs) have the potential to support earlier detection of the behavioral and psychological symptoms of dementia (BPSD). The recent surge of interest in this topic underscores the need to comprehensively examine the existing evidence.

Objective: This scoping review aimed to identify and summarize the types and uses of AITs currently used for the early detection of BPSD among people diagnosed with the disease. We also examined which health care professionals were involved, nursing involvement and experience, the care settings in which these technologies are used, and the characteristics of the BPSD that were assessed.

Methods: Our scoping review was conducted in accordance with the Joanna Briggs Institute manual for scoping reviews. Searches were conducted in March 2025 in the following bibliographic databases: MEDLINE ALL Ovid, Embase, APA PsycINFO Ovid, CINAHL EBSCO, Web of Science Core Collection, the Cochrane Library Wiley, and ProQuest Dissertations and Theses A&I. Additional searches were performed using citation tracking strategies and by consulting the Association for Computing Machinery Digital Library. Eligible studies included primary research involving people with dementia and examining the use of AITs for the detection of BPSD in real-world care settings.

Results: After screening 3670 articles for eligibility, the review includes 12 studies. The studies retained were conducted between 2012 and 2025 in 5 countries and encompassed a range of care settings. The AITs used were predominantly based on classic machine learning approaches and used information from environmental sensors, wearable devices, and data recording systems. These studies primarily assessed behavioral and physiological parameters and focused specifically on symptoms, such as agitation and aggression. None of the retained studies explored nurses' roles or their specific skills in using these technologies.

Conclusions: The use of AITs for managing BPSD represents an emerging field of research offering novel opportunities to enhance their detection in various health care contexts. We recommended that nurses be actively engaged in developing and assessing these technologies. Future research should prioritize investigations into how effective AITs are across diverse populations, whether they can have a long-term impact on managing BPSD, and whether they can improve the quality of life of patients and caregivers.

(JMIR Aging 2025;8:e76074) doi:10.2196/76074

KEYWORDS

artificial intelligence; behavioral and psychological symptoms of dementia; early detection; older adults; scoping review

XSL•FO
RenderX

Introduction

Background

The growing prevalence of dementia is a major global concern [1]. As the disease progresses, cognitive decline is often accompanied by the behavioral and psychological symptoms of dementia (BPSD) [2,3]. The International Psychogeriatric Association defines the BPSD as disturbances in perception, thought content, mood, or behavior [4]. These symptoms can be grouped into two subtypes: (1) behavioral symptoms, which are typically observed directly (agitation, verbal or physical aggression, wandering, sexual disinhibition, repetitive behavior, catastrophic reactions, sundowning, sleep problems, sleeplessness, and apathy), and (2) psychological symptoms, which are usually assessed through interviews with patients and relatives (psychosis, delusions, hallucinations, misidentifications, mood changes, anxiety, restlessness, and depression) [2,4]. BPSD has a deleterious impact on the quality of life of people with dementia, accelerating cognitive decline, reducing life expectancy, and precipitating hospitalizations and early placement in nursing homes [5,6]. Managing patients' BPSD often requires constant vigilance, leading to heightened stress and depressive symptoms among informal caregivers. Furthermore, indirect costs, including medical consultations, nonreimbursed medication, home adaptations, and reduced professional activity, can lead to substantial socioeconomic burdens [7-9]. For health care professionals, BPSD can increase workloads, stress, and the risk of professional burnout, which may reduce the quality of the care provided [10,11]. Regarding the health care system overall, BPSD leads to higher costs due to more consultations, hospitalizations, and prescriptions of psychotropic drugs [12].

The etiopathogeneses of BPSD are complex and multifactorial, influenced by biopsychosocial and environmental factors, and vary considerably in frequency and intensity between individuals [13,14]. Their early detection through the identification of subtle behavioral and psychological changes (eg, apathy, anxiety, sleep problems, irritability, and social withdrawal), before any formal clinical intervention, is of major clinical importance [15,16]. Early identification enables proactive management and personalized interventions, thereby helping to mitigate the symptoms' deleterious impact. However, early detection is challenging as early signs are often transient, underrecognized, and misinterpreted as normal aging or personality changes [15,16]. Furthermore, standard detection tools, such as the Neuropsychiatric Inventory, the Alzheimer's Disease Behavioral Pathology Assessment Scale, or the Cohen-Mansfield Agitation Inventory, can be time-consuming, and health care professionals do not always perceive them to be precise or practical [11,14,17]. These challenges are exacerbated by growing demands for care and persistent shortages of qualified health care staff [18,19]. New health care technologies offer promising solutions to these challenges [20,21].

Recent technological developments have focused on artificial intelligence–based technologies (AITs), which combine artificial intelligence (AI) with remote sensors, robotics, and decision support algorithms [21-24]. AI encompasses a wide range of techniques and methodologies designed to enable computer systems to mimic human cognitive functions, such as learning, problem-solving, and decision-making [24]. When combined with critical thinking and human judgment, AI can optimize clinical reasoning by accelerating and refining assessment, anticipation, the synthesis of multiple data sources, and knowledge generation [25]. In health care, AI is already applied in diagnosis and screening, treatment personalization, patient monitoring, robot-assisted surgery, and health system management [22]. In dementia care, current AITs include monitoring systems, social robots, and daily living assistance tools, primarily involving wearable and environmental sensors [26-28]. These technologies have shown positive impacts on depression, quality of life, social engagement, cognitive function, and patient monitoring [26-28]. They have also demonstrated good performance in health monitoring and dementia screening [26-28].

Management of BPSD requires a multidisciplinary team, with nurses playing a key role in the early detection of symptoms, thanks to their continuous proximity with patients [4]. However, due to the complexity of these symptoms and the diversity of beliefs and conceptualizations regarding BPSD, it is not uncommon for nurses' clinical judgments to be based on intuition and their standard practice rather than on a clear, standardized methodology [29,30]. This underscores the potential for introducing AITs to support clinical judgment and thus enhance the quality of nursing care for people with dementia. Registered and advanced practice nurses have been actively helping to develop and implement AITs in health care [31-33]. They engage in activities including problem-solving, defining contextual needs and formulating priorities, idea generation, finding concrete solutions, and determining end user satisfaction [31-33].

While previous reviews have examined the use of AITs in the care of patients with dementia and BPSD, their scope differed from this work [34-38]. Husebo et al [37] focused on treatment monitoring, and Jao et al [35] emphasized model performance without mapping contexts of use. AI has evolved rapidly since 2020, and it is likely that new solutions, not covered in these reviews, have now emerged. More recently, other authors have addressed the use of AITs in the general care of those with dementia and BPSD, combining prediction, assessment, and treatment, but without a specific focus on early detection [34,38]. Similarly, Shaik et al [36] reviewed remote sensing technologies and care frameworks without focusing on the application of AITs for the early detection of the BPSD. Our preliminary literature searches identified no completed or ongoing scoping reviews on the topic. Given the rapid evolution of AITs and their growing integration into dementia care, an updated and targeted synthesis of the literature is necessary.

Aim and Review Questions

This scoping review aimed to identify and summarize the different types of AITs currently available for the early detection of BPSD among older adults. It also sought to examine the functional uses of those technologies, the health care professionals involved, nursing involvement and experience with AITs, the care setting in which they are used, and the

XSL•FO
RenderX

characteristics of the BPSD they assess. Our research questions were developed following the population, concept, and context framework. First, what type of AITs are currently available for the early detection of BPSD among people with dementia? Second, what are the functional uses of the AITs involved in the early detection of BPSD? Third, which health care professionals are involved in the use and implementation of AITs for the early detection of BPSD? Fourth, in which care settings are AITs used for the early detection of BPSD? Which specific characteristics of BPSD do AITs assess? Finally, what are nurses' specific roles, responsibilities, and experiences regarding the use of AITs for the early detection of BPSD?

Methods

Protocol and Registration

This scoping review was conducted in accordance with the Joanna Briggs Institute manual for scoping reviews [39]. The PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews; checklist provided in Checklist 1) standards were followed for reporting [40]. The scoping review's protocol was registered at the Open Science Framework [41].

Eligibility Criteria

The eligibility criteria for studies with the potential to be assessed in this review were defined based on their population, concepts, and types of sources. The inclusion and exclusion criteria are detailed in Table 1.

XSL•FO
RenderX

Table . Scoping review inclusion and exclusion criteria.

Criteria	Inclusion criteria	Exclusion criteria
Participants	• Studies involving people aged 65 y and older with any dementia (eg, Alzheimer disease, frontotemporal dementia, vascular dementia, and Parkinson dementia).	• Studies involving individuals with mental disorders other than dementia (eg, mild cognitive impairment, depression, delirium, schizophrenia, and bipolar disorder).
Concept		
Artificial intelligence: techniques and methodologies designed to enable computer systems to mimic human cognitive functions, such as learning, problem-solving, and decision-making [24]. Other terms: "computational intelligence," "computer reasoning," "machine learning," and "machine intelligence."	• Studies involving all types of devices and software based on AITs[a] to screen, detect, and assess clinical BPSD[b].	• Studies focusing on AITs intended for the prediction or management of symptoms or diseases (eg, type and stage of dementia, psychiatric disorders, chronic illnesses, and acute conditions).
BPSD: Disturbances in perception, thought content, mood, or behavior [4].	• Studies involving all types of clinical BPSD: agitation, aggression, verbal aggression, physical aggression, psychosis, delusions, hallucinations, misidentifications, repetitive behavior, wandering, sundowning, restlessness, catastrophic reactions, sexual disinhibition, mood changes, depression, apathy, anxiety, sleep problems, and sleeplessness [4].	• Studies that did not use scales rating BPSD to report psychological symptoms.
Context		
(1) Acute care: Immediate and intensive treatment services provided within the health care system to address severe, sudden and often unforeseen health episodes: emergency medicine, trauma services, prehospital emergency response, acute surgical care, critical care units, urgent care, and short-term stabilization of patients in an inpatient setting [42]. (2) Community care: social support services and programs designed to empower individuals with specific needs—including those with intellectual disabilities, mental health challenges, and older adults—to lead independent lives and actively engage in their communities [43]. (3) Long-term care: support services aimed at addressing the health and personal care needs of individuals who are unable to manage daily activities independently: home-based care, community care and residential care [44].	• Studies involving acute care, community care, or long-term care from any country.	• Studies involving virtual environments.
Types of sources	• Experimental and quasi-experimental studies; • Analytical and descriptive observational studies; • Qualitative and mixed methods studies.	• All types of reviews: narrative reviews, integrative reviews, scoping reviews, systematic reviews, umbrella reviews, and bibliometric analyses; • Opinion letters; • Commentaries; • Books and book chapters.

[a]AIT: artificial intelligence–based technology.

[b]BPSD: behavioral and psychological symptoms of dementia.

Information Sources and Search Strategy

In March 2025, we conducted a literature search of the following bibliographic databases in collaboration with a medical librarian (JRA): MEDLINE ALL Ovid, Embase, APA PsycInfo Ovid, CINAHL EBSCO, Web of Science Core Collection, the Cochrane Database of Systematic Reviews Wiley, the Cochrane Central Register of Controlled Trials Wiley, and ProQuest Dissertations and Theses A&I. No language or date restrictions were applied. The final search strategies were peer-reviewed by a second information specialist and are documented in Multimedia Appendix 1, detailing the search syntax, keywords, and index terms used. Using the studies that were retained, we performed a backward and forward citation search using

XSL•FO
RenderX

Citationchaser [45] (SF). Further research was carried out in the Association for Computing Machinery Digital Library (SF).

Study or Source of Evidence Selection

After the search, a medical librarian (JRA) imported all the references identified into EndNote 20 (Clarivate Analytics), and duplicate references from the databases were removed using Deduklick (Risklick AG) [46]. The references were then imported into the Rayyan interface, and 2 reviewers (SF and CGdR) independently screened the titles and abstracts according to the established inclusion criteria. All the potentially relevant sources were retrieved in full and imported into a database. Two reviewers (SF and EP) then independently assessed the full texts of the references retained to ensure they met the inclusion criteria. The rationales for excluding references at the full-text review stage were documented. Disagreements between the reviewers during this selection process were discussed and resolved with the assistance of a third reviewer (HV). Due to resource constraints, the selection of references identified through citation tracking was conducted by a single reviewer (SF). The results are presented in a PRISMA-ScR flow diagram [47].

Data Extraction

Data from the references retained for the scoping review were extracted by a single reviewer (SF), and a random 10% subset of the full-text articles included was verified by a second reviewer (HV) to ensure accuracy and completeness [48]. The second reviewer cross-checked the completed data extraction tool's results against the corresponding full-text articles. Data were extracted using a data extraction tool developed in accordance with the study's objectives and search questions. The data extracted included the year of publication, country of origin, study characteristics (design, objectives, sample size, health care setting, and health care professionals), participant characteristics (age, sex, and dementia type and severity), specific types of the AITs (sensors, social robots, and decision support systems), delivery methods for these AITs (remote monitoring platforms and mobile apps), the functional purposes of these AITs, the AI techniques applied, characteristics of the BPSD assessed, the traditional instruments used for detecting BPSD, nurses' involvement and experience, and key findings relevant to the review questions (eg, effectiveness, usability, feasibility, users' satisfaction with AITs, users' perspectives about AITs, and barriers and facilitators to adopting AITs). When necessary, study authors were contacted to request missing or additional data.

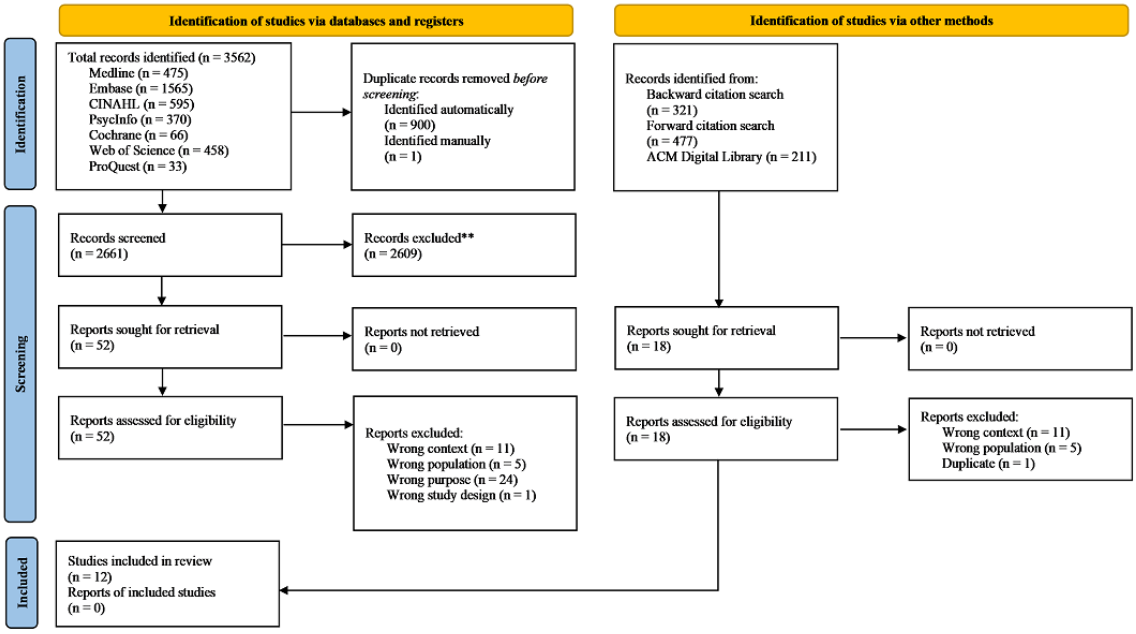
Data Analysis and Presentation

We use a narrative approach to present the extracted and analyzed evidence in the different sections below [39]. The sections include (1) the characteristics of the included studies, (2) the types and uses of AITs, (3) the characteristics of the BPSD detected, and (4) the involvement and experience of nurses. The types of AITs were analyzed and categorized according to the definitions proposed by Zhang and Lu [49], which distinguish the different drivers and technologies of AI. Tables, charts, and diagrams are used to summarize and make the evidence more readable.

Results

Our bibliographic database searches identified 3562 publications. After removing 901 duplicate records, 2661 publications underwent title and abstract screening. Of these, 2609 were excluded for not meeting our inclusion criteria, and 41 of the 52 full-text articles assessed for eligibility were excluded. Following the recommendation of an expert, 18 full-text articles were found using citation searching and a search of the Association for Computing Machinery Digital Library; however, after assessing them for eligibility, 17 were excluded. The scoping review finally retained 12 articles [50-61] (Figure 1).

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram (source: Page et al [47]).



XSL•FO
RenderX

Study Characteristics

The 12 studies reviewed came from 5 countries and were published from 2012 - 25. Over half of the studies were published in the last 5 years (2020 - 2025) [50-56]. They included 3 single case studies [53,57,58], 2 studies using quasi-experimental approaches [52,59], 1 randomized controlled trial [60], 1 correlational study [61], 1 exploratory observational study [54], 3 retrospective longitudinal studies [50,55,56], and 1 qualitative study [51]. Furthermore, 3 publications examined acute care settings [52,57,61], 6 examined community care settings [50,54,55,58-60], and 3 examined long-term care settings [51,53,56].

The study participants included groups of older adults with mean ages ranging from 78.8 to 84 years old and presenting with mild to advanced dementia. Although 1 study [51] did not involve participants with a confirmed diagnosis of dementia, it was retained due to its focus on the implementation of a behavioral disturbance detection system in long-term care settings, which predominantly serve people with dementia. As such, it provided valuable contextual insights relevant to integrating such technologies into dementia care. Most of the health care professionals included in the studies were nurses [51-54,56,57], followed by physicians [53,56,57], care workers [51,56], a rehabilitation assistant [51,56], an occupational therapist [51,56], and a physiotherapist [51,56]. Table 2 describes the main characteristics of the studies retained.

XSL•FO
RenderX

Table . Characteristics of the studies retained (n=12).

Study, country	Design (aim)	Sample size (N)	Health care setting	Health care professionals involved	Participants (age, sex, dementia type, and severity)
Bafaloukou et al [50], United Kingdom	Retrospective longitudinal study (develop a machine learning framework able to identify episodes of agitation).	63	Community care (home)	—[a]	Age: N/P[b], sex: female (n=22), male (n=41), dementia: Alzheimer, vascular frontotemporal, Parkinson
Bourennane et al [57], France	Single case study (describe experiments to detect deviations in a patient's behavioral observations using a multi-sensor).	1	Acute care (Alzheimer's care unit)	Physicians and nurses	Age: mean 84 y, sex: female, dementia: Alzheimer
Choukou et al [51], Canada	Qualitative case study (explore the perceptions of healthcare professionals and family members about the potential for introducing a behavioral disturbances detector in long-term care homes).	8 health care professionals and 1 proxy	Long-term care (tertiary care facility)	Nurses, care workers, rehabilitation assistants, occupational therapists, and a physiotherapist	Age: N/P, sex: female (n=9), dementia: N/P
Enshaeifar et al [59], United Kingdom	Quasi-experimental, interrupted-time series (detect patterns directly correlated with agitation, irritation, and aggression in people with dementia).	12	Community care (home)	—	Age: N/P, sex: N/P, dementia: mild to moderate
Iaboni et al [52], Canada	Quasi-experimental, interrupted-time series (develop personalized machine learning models capable of detecting individual patterns in BPSD[c]).	17	Acute care (specialist dementia unit)	Nurses	Age: mean 78.9 (SD 8.9) y, sex: female (n=10), dementia: advanced
Kim et al [53], South Korea	Single case study (detect a dementia patient's abnormal behavior using sensors and developing a machine-learning-based system).	1	Long-term care (nursing homes)	Physicians and nurses	—
König et al [61], France	Correlational (investigate whether automatic speech analysis could be used to characterize and detect apathy).	60	Acute care (hospital memory clinic)	—	Age female with apathy: mean 79.50 (SD 5.86) y, age male with apathy: mean 79.58 (SD 5.45) y, dementia: N/P
Lee et al [54], South Korea	Exploratory observational study (investigate a predictive model for sleep efficiency and its associated features among older adults living with dementia in their own homes).	69	Community care (home)	Nurses	Age sleep efficiency >85%: mean 80.15 (6.11) y, Age sleep efficiency <85%: mean 80.64 (4.63) y, sex: N/P, dementia: N/P

XSL•FO
RenderX

Study, country	Design (aim)	Sample size (N)	Health care setting	Health care professionals involved	Participants (age, sex, dementia type, and severity)
Lotfi et al [58], United Kingdom	Single case study (design unobtrusive activities in a daily living monitoring system for identifying and predicting abnormal behavior).	1	Community care (home)	—	Age: N/P, sex: female, dementia: N/P
Palermo et al [55], United Kingdom	Retrospective longitudinal study (identify episodes of agitation among people living with dementia using a deep recurrent model applied to continuous in-home monitoring data).	Total=46 (33 for training, 8 for validation, and 5 for testing)	Community care (home)	—	Age: mean 82.6 (SD 7.2) y, sex: N/P, dementia: Alzheimer, frontotemporal, vascular, Parkinson
Rostill et al [60], United Kingdom	Randomized control trial, parallel groups (test the technology integrated health management system in a real-world setting).	204 patients and 204 carers	Community care (home)	—	Age: N/P, sex: N/P, dementia: mild to moderate
Zhu et al [56], Australia	Retrospective longitudinal study (determine the prevalence of agitated behaviors among people with dementia using natural language processing).	3528	Long-term care (nursing homes)	Nurses, care workers, and other health care professionals (eg, physicians and therapists)	Age: N/P, sex: female (n=2267), male (n=1261), dementia: N/P

[a]Not applicable.

[b]N/P: not provided.

[c]BPSD: behavioral and psychological symptoms of dementia.

AITs and Their Use

Data Collection

Most of the studies adopted a multimodal input approach, combining 2 or more data types. In total, 6 studies reported on combinations of environmental sensors and wearable devices [50,53-55,57,59,60] and 1 study used wearable devices with video cameras [52]; 2 studies relied on signal data, referring to continuous, time-series measurements, captured through voice recorders or environmental sensors [58,61]. Voice recorders were not classified as wearable devices as they were set in fixed positions during structured interviews rather than being worn continuously for passive data collection. One study used structured data derived from an electronic health care records system [56]. The studies used a variety of environmental sensors, including passive infrared sensors or motion detectors, door and window entry point sensors, electricity power usage sensors, and bed and sofa pressure sensors. Wearable devices included wristbands, smartwatches, sweat patches, and GPS trackers integrated into textiles or attached to the body. These devices were equipped with sensors collecting motion data via accelerometers, monitored sleep patterns through actigraphy, and recorded other structured data, such as blood volume pulse, electrodermal activity, transpiration, and skin temperature. Smart thermometers and smart scales were used to collect physiological data, including body temperature, body composition, and heart rate. Furthermore, 4 studies reported using wireless technology systems to communicate between the various devices [53,57,58,60].

AI Learning Processes and Approaches

Supervised learning processes were used in 7 studies in which models were trained on labeled datasets to predict specific behavioral outcomes [50,52-55,58,61]; 2 studies used unsupervised learning approaches aiming to discover patterns or clusters within unlabeled data [56,59], while 1 study adopted a hybrid process combining both supervised and unsupervised methods [57]. Most of the studies relied on machine learning techniques, including algorithms such as random forest, logistic regression, k-nearest neighbors, decision trees, support vector machines, LightGBM, and CatBoost [50,52-54,56,57,59-61]. Only 2 studies applied deep learning methods, specifically recurrent neural networks and long short-term memory networks [55,58]. Table 3 provides an overview of the main data sources, AI techniques, learning processes, and AI approaches reported in the studies selected.

XSL•FO
RenderX

Table . Principal methods and data sources of the artificial intelligence–based technologies used.

Study	Data sources (input)	AI[a] technique	Learning process		AI approach	
			Supervised	Unsupervised	Machine learning[b]	Deep learning[c]
Bafaloukou et al [50]	Multimodal signal	Light gradient-boosting machine	✓	—[d]	✓	—
Bourennane et al [57]	Multimodal signal	K-means, k-nearest neighbors	✓	✓	✓	—
Enshaeifar et al [59]	Multimodal signal	Markov chains	—	✓	✓	—
Iaboni et al [52]	Multimodal signal	Random forest	✓	—	✓	—
Kim et al [53]	Multimodal signal	k-nearest neighbors, random forest, logistic regression, decision tree, support vector machine, multilayer perceptron	✓	—	✓	—
König et al [61]	Signal	Logistic regression	✓	—	✓	—
Lee et al [54]	Multimodal signal	CatBoost	✓	—	✓	—
Lotfi et al [58]	Signal	Recurrent neural networks	✓	—	—	✓
Palermo et al [55]	Multimodal signal	Long short-term memory	✓	—	—	✓
Rostill et al [60]	Multimodal signal	N/P[e]	N/P	N/P	✓	—
Zhu et al [56]	Structured data	Computerized natural language processing	—	✓	✓	—

[a]AI: artificial intelligence.

[b]Machine learning is a set of algorithms that learn patterns from data using manually engineered features, often applied to structured datasets [24].

[c]Deep learning uses artificial neural networks to automatically learn hierarchical feature representations directly from raw data [24].

[d]Not applicable.

[e]N/P: not provided.

Data Availability and Visualization

Of the total, 2 studies included used cloud-based systems to retrieve data from devices and transmit them to AI-driven analytical platforms [53,60], while 1 study used the Zigbee communication protocol to transmit data [57]. Furthermore, 5 of the studies retained made their data derived from AI available and visualized it on secure web portals [55,57-60]. In comparison, 1 study used a mobile health app [53]. In addition, 6 studies described the use of AITs with alert systems triggered by the detection of abnormal behavior [50,53,55,57,59,60]; 4 studies reported using AITs to identify abnormal behavior without an alert system [52,56,58,61]. The studies retained in our analysis defined abnormal behavior heterogeneously, and 4 studies defined it as a deviation from the established norm in each participant's profile [53,57-59]. These deviations were characterized by different behavioral features, and their nature varied across studies, with 2 focusing on movement in specific rooms or living areas [57,58], 2 focusing on agitation [57,59], and 1 focusing on repetitive questions, behaviors, and sleep profiles [53]. None of these 4 studies provided any specifications regarding the variance defined by the norm. The principal elements of the interventional studies are outlined in Table 4.

XSL•FO
RenderX

Table . Principal elements of the interventional studies.

Study	Intervention duration	Device used	Intervention description
Bafaloukou et al [50]	27 mo	• Passive infrared sensors and door and kitchen appliance sensors • Bed mattresses	• Weekly behavioral monitoring questionnaire to report the presence or absence of participants' behavioral symptoms. • Traffic light system (interactive interface), grouping estimates of the probability of agitation produced by the model.
Bourennane et al [57]	30 d	• Wearable tag with an inbuilt GPS tracker • Infrared and pressure detectors in bed mattresses	• Wireless communication technology system to relay data to a central computer in the nursing staff area. • Abnormal behavior: hyperactivity, deviations from the standard means[a] for patient movements through common living areas by time of day and for agitation profiles from the past 30 d. • Alerts sent via SMS text messages to caregivers' mobile phones.
Enshaeifar et al [59]	6 mo	• Passive sensors: 2 passive infrared sensors, 4 motion sensors, 2 pressure sensors, 1 door sensor, and 1 central energy consumption monitoring device • Bluetooth-enabled medical devices	• Anomaly detection of outlying values[a] for agitation, irritation, and aggression during daily activities. • Alerts were generated on a secure web portal and verified by the clinical team by contacting the caregiver.
Iaboni et al [52]	8 wk	• Wristband (Empatica E4) • Video cameras installed in common living areas	• Customized machine learning models to classify the presence of behavioral symptoms by type (aggression and agitation).
Kim et al [53]	30 d	• Smartwatch • Lifelog sensors attached to assorted pieces of furniture	• Wireless communication technology system interacting between a smartwatch and lifelog sensors. • Abnormal behavior: repetitive questions or behaviors and sleep disturbances. • Multilayer perceptron machine learning model determined the presence or absence of abnormal behavior. • Alerts were transmitted to the smartphones of physicians and nurses via an app.
Lee et al [54]	14 d	• Wristband (Actiwatch2 [Respironics, Inc]) • Sweat patches (PharmChek [SCRAM Systems])	• Sleep efficacy analysis: actigraphy for sleep and physical activity, sweat patch for cytokines. • Machine learning analysis to select the sleep efficiency prediction model and exploration of its top 10 associated features. • Categorical boosting machine learning algorithm based on decision trees.

XSL•FO
RenderX

Study	Intervention duration	Device used	Intervention description
Lotfi et al [58]	18 mo	• JustChecking wireless motion sensor in the kitchen, bedroom, bathroom, and living room • Door entry point sensors in the front and back doors	• Abnormal behavior: deviations from the standard mean[a] for the length of time spent in a specific area of the house. • Recurrent neural networks to analyze data with results transmitted to caregivers via a secure web interface.
Palermo et al [55]	27 mo	• Motion and tracking sensors (multiple doors, hallway, kitchen, and other rooms) • Home appliance use, smart plugs (kitchen appliances, television, and other commonly used devices) • Wearable device • Smart temporal thermometer • Smart scale	• Digital platform to integrate and analyze data. • Data labeled as true (episode of agitation validated by a clinical monitoring team) or false (not an episode of agitation when the monitoring team contacted participants). • Recurrent neural networks are used to analyze the risk of episodes of agitation.
Rostill et al [60]	6 mo	• Sensors • Vital signs monitors • GPS tracker • Gateway device	• Machine learning algorithm to analyze data. • Alerts generated and prioritized on a digital dashboard and followed up by a centralized monitoring team overseen by clinicians. • Alert follow-up process: contact caregivers and, if necessary, liaise with other services (social care teams, Alzheimer's Society, and emergency services).

[a]Defined threshold not available.

AI Model Performance

Of the total, 8 studies reported that their AI model performed well in identifying and classifying distinct types of BPSD [50,52-56,59,61]. The performance metrics used included accuracy, precision, recall, F_1-score, and area under the curve, as they are commonly defined in the literature [62]: accuracy refers to the proportion of correctly classified instances among all instances; precision measures the proportion of true positives among all positive predictions, while recall reflects the proportion of true positives correctly identified among all actual positives; the F_1-score is the harmonic mean of precision and recall, balancing both metrics, the area under the curve indicates the model's ability to distinguish between classes across all thresholds. The parameters used and their corresponding scores are described in Table 5.

XSL•FO
RenderX

Table . Performance of the artificial intelligence models used in the retained studies.

Study	Model or algorithm	Performance metrics, mean (SD)				
		Accuracy	Precision	Recall	F_1-score	AUC[a]
Bafaloukou et al [50]	Gradient boosting algorithm	mean 71.3 (SD 7.21)	mean 71.81 (SD 7.94)	N/P[b]	mean 71.12 (SD 7.28)	mean 77.63 (SD 6.59)
Enshaeifar et al [59]	Hierarchical fusion algorithm	N/P	0.81	0.83	0.80	N/P
Iaboni et al [52]	Interpretable models	N/P	N/P	N/P	N/P	0.87
Kim et al [53]	Multilayer perceptron model	0.99	1.00	1.00	1.00	N/P
König et al [61]	Logistic regression models	N/P	N/P	N/P	N/P	0.88 for men and 0.77 for women
Lee et al [54]	Categorical boosting algorithm	0.96	0.87	0.77	0.82	0.99
Palermo et al [55]	Recurrent neural networks model	N/P	0.27	0.79	0.37	N/P
Zhu et al [56]	Rule-based NLP[c] model	N/P	N/P	N/P	N/P	0.89

[a]AUC: area under the curve.

[b]N/P: not provided.

[c]NLP: natural language processing.

The Adoption of AITs

In total, 2 studies reported that the participating health care professionals found AITs useful in helping them understand patients' health statuses [51,60]; 1 study reported that its participants perceived the use of AITs to be feasible [54]. Another study documented how the majority of its participants, both caregivers and patients, accepted and would have given a favorable recommendation on using the AIT under evaluation [60]. These positive perceptions were primarily linked to the devices' ease of use and the involvement of a clinical monitoring team to interpret the data generated by the system. Indeed, the few nonendorsements of that AIT were attributed to technical issues, user-perceived difficulties with device use, and the perceived inconvenience associated with the routine monitoring of vital signs. One study identified forgetfulness, refusal, and a lack of cooperation as the primary factors contributing to patients not adhering to their wearable device protocol [54]. Technical problems, such as battery discharge, software errors, device malfunction, and detection failures were also reported as contributing factors [54]. One study referenced future potential AIT users' perceived requirements for successfully incorporating AITs into clinical environments [51]. These included the necessity for training, the standardized use of technical devices and giving priority access to them to nursing and care staff. Furthermore, that study indicated that those potential AIT users had concerns about data confidentiality and the costs associated with AITs [51].

Characteristics of the BPSD Detected

The studies retained for analysis reported their assessments of specific behavioral and physiological parameters acting as surrogate markers of BPSD. Regarding behavioral parameters, 6 studies assessed motion [50,52,53,55,57,60], 4 assessed activities of daily living [53,55,56,58], 1 assessed sleep efficiency [54], and 1 assessed speech quality [61]. Regarding physiological signs, 5 studies assessed vital and biological signs (blood pressure, heart rate, body temperature, blood oxygenation, weight, and hydration) [50,52,55,59,60]. One study measured electrodermal activity [52], while another study examined patients' levels of cytokines, such as interleukin-6 and interleukin-10 [54]. One study examined environmental factors associated with BPSD, such as indoor light levels and ambient temperature [50].

Furthermore, 2 studies focused on detecting the 3 specific BPSD, that is, agitation, irritation, and aggression [59,60]; 3 studies examined agitation alone [50,55,56], while another concentrated on detecting agitation and aggression [52]. In addition, 1 study investigated apathy as a distinct form of BPSD [61], and another concentrated on sleep efficacy [54].

Regarding the use of specific scales for detecting and measuring BPSD, 4 studies used the Neuropsychiatric Inventory [50,52,54,61], and another combined the Neuropsychiatric Inventory with the Apathy Inventory [61].

One study indicated that biological parameters, including blood pressure, blood oxygen, heart rate, body temperature, and hydration, were the most significant markers for detecting agitation, irritability, and aggression [60]. Another study identified heart rate and body temperature as the principal markers for verbal aggression, while an accelerometer reading was the most significant marker for physical aggression [52]. One study identified respiratory rate, number of nocturnal awakenings, and the indoor light level as the most important

XSL•FO
RenderX

characteristics for identifying agitation [50]. Table 6 summarizes this information.

Table . Characteristics of the behavioral and psychological symptoms of dementia found in the studies retained for analysis.

Sign of BPSD[a]	Types of BPSD		
	Agitation or aggression	Apathy	Sleep problems
Behavioral	• Sleep • Daily activity pattern • Sequence of daily activities • Movement activity patterns	• Paralinguistic markers	• Physical activity • Percentage of time awake • Total sleep duration • Daytime sleep minutes • Nocturnal sleep minutes • Sleep regularity
Physiological	• Blood pressure • Heart rate • Blood volume pulses • Oxygen saturation • Electrodermal activity • Body temperature • Weight • Hydration	• None	• Interleukin-6 • Interleukin-10

[a]BPSD: behavioral and psychological symptoms of dementia.

Nurses' Involvement and Experience

Only 3 of the studies retained clearly described nurses' roles in identifying BPSD [51,52,54]. One documented their participation in data collection, specifically documenting episodes of BPSD [52]. Another study described their involvement as participants in a group of health care professionals and the exploration of their perceptions of the potential use of AITs in clinical practice [51]. In research contexts, nurses were part of the research team in 1 study [52] and led the design and implementation of another [54]. None of the retained studies reported on the extent of nurses' knowledge, skills, and experience in the use of AITs.

Discussion

Principal Results

This scoping review highlighted how AITs are emerging and being used for the early detection of BPSD, as evidenced by the limited but growing number of publications over the past 5 years (between 2019 and 2025). Despite this limited body of literature, the studies retained were diverse in terms of design, care environment, and health care professional involvement, as well as in terms of their study populations' characteristics, and this provided a more comprehensive understanding of the various approaches to using AITs in the care of people with dementia.

The studies included in this review encompassed the domains of acute, community, and long-term care and included several types of health professionals, such as physicians, nurses, home care workers, and therapists. Only one of the studies included explored the integration of AITs from an interprofessional perspective, highlighting potential benefits such as improved communication and more proactive care [51]. These findings aligned with previous reviews highlighting AITs' roles in enabling real-time clinical decision-making and interdisciplinary data sharing [36,38]. However, health care professionals also expressed concerns about how the challenges of implementing AITs might hinder their successful adoption in daily practice, such as workflow disruptions, overreliance on technology, threats to professional identity, costs, and data privacy [51]. These concerns, also reported in previous literature [36,63,64], involve organizational, technological, professional, and cost issues. Identifying these barriers early on is therefore essential to guiding implementation strategies and systems design, thereby enhancing the usability, acceptability, and integration of AITs, particularly in complex interprofessional settings such as dementia care.

Regarding data collection systems and devices, many of the retained studies reported on the combined use of environmental sensors and wearable devices [52,53,55,57,60]. It should be noted that the innovation in these studies does not lie in the use of particular devices or the application of AI, as both have been used before in dementia care settings, particularly to monitor and enhance cognitive function, as well as to track daily activities and physiological parameters [26,28,36,64-66]. Distinctive innovation lies in integrating data collected by these devices with an AI-based analysis specifically aimed at identifying and generating alerts for episodes of BPSD. While similar integrations have been explored in broader smart health monitoring contexts, their targeted use for the early detection of BPSD remains largely underexplored [26,37,38]. These findings are consistent with those reported in previous reviews, which also mentioned mobile apps, text analysis systems, and robots [27,36-38]. For instance, Shaik et al [36] reported on the use of robots capable of monitoring patient behavior. However, none of the studies included in our review used robots specifically for the detection of BPSD, reflecting a broader trend where AITs form parts of psychosocial interventions rather than acting as early detection tools [34].

The variety of connectable devices is particularly interesting, offering health care professionals the flexibility to tailor technologies to care objectives, patients' preferences, and

XSL•FO
RenderX

context-specific needs. Concerning communication systems and data visualization, most of the studies retained used wireless technologies and a secure web portal, reflecting the growing integration of the Internet of Medical Things in health care [53,57-60]. This type of infrastructure enables real-time data transmission and storage, supporting remote monitoring and system interoperability [67]. Using wireless systems facilitates technological integration by reducing the physical constraints that can be particularly relevant for people with dementia, as restrictive or obtrusive devices could cause distress or resistance [68]. However, device acceptability varied across the studies retained, reflecting the mixed findings reported in previous literature [27,36,67]. Acceptability was primarily judged according to ease of use [51,54,60], consistent with previous reviews that also emphasized the importance of simple, intuitive interfaces [27,36]. Regarding device design, 1 study on wearable devices reported their overall feasibility [54]. However, this contrasted with the conclusion by Loveys et al [27], who identified several barriers to acceptance, including discomfort when wearing devices, intrusiveness due to continuous monitoring, and sleep disturbances caused by bed sensors. Another study involving wearable devices similarly reported such inconveniences, particularly a feeling of being constantly monitored [60], in alignment with the barriers described by Loveys et al [27]. Additional concerns included technical issues and forgetting to wear the device [51,60]. These findings are consistent with previous literature [27,36,67] and further reinforce the importance of anticipating such barriers early on in the system design process to improve the user experience and support successful implementation in dementia care settings.

None of the studies included used advanced data storage and sharing systems (such as edge computing) that can improve security and enable the offline functionalities that are important in settings with limited or unreliable internet access, such as in home care or remote regions. While data security remains a major concern in the Internet of Medical Things, AI offers promising support. As Messinis et al [69] stated, machine and deep learning techniques can enhance cybersecurity. However, in the studies reviewed, AI was primarily used for data analysis in research contexts rather than being fully integrated into clinical practice.

Most of the studies retained used classic machine learning to classify or detect BPSD, training models on labeled data to identify predefined behaviors [50,52-55,58,61]. This reflects a broader trend in early-stage research, where annotated datasets are key for model development and validation [35,36].

While some of the studies retained defined abnormal behaviors as deviations from individual baselines, alerts for the occurrence of BPSD were not systematically triggered at fixed thresholds [53,57-59]. Instead, detection often relied on adaptive methods, such as multivariate pattern recognition or unsupervised learning algorithms that adjusted to individual variations over time. These approaches are particularly well-suited to the heterogeneity and fluctuation of BPSD [13,14] and have the added advantage of requiring few or no labeled data, making them suitable for real-time and continuous monitoring [24,70]. In parallel, some studies used expert validation to verify the accuracy or clinical relevance of alerts, either through retrospective review by clinicians [55,57,60] or the integration of rule-based insights during model development [52,53,59]. One study combined supervised and unsupervised techniques to enhance adaptability to individual patterns and clinical relevance, thus supporting a flexible, person-centered approach [57].

The studies retained for review assessed behavioral and physiological parameters as signs of the BPSD, with most being behavioral parameters, including movement and the activities of daily living [50,52-58,60,61]. Regarding specific BPSD, agitation, irritability, aggression, and apathy were the most frequently investigated [50,52,55,56,59-61], with biological parameters and accelerometer data being the most significant markers for detecting them [52,60]. This particular focus may be explained by the clinical urgency and disruptive nature of these symptoms, which place a significant burden on care staff and often require rapid intervention [5,71,72]. Furthermore, these behaviors tend to generate observable and measurable physiological signals, making them more technically accessible to current AITs. These findings were consistent with previous reviews, one of which also highlighted wandering as a distinct type of BPSD commonly targeted by AITs [35,36]. However, this trend underscores the need to assess how effectively AITs can capture the broader spectrum of BPSD, particularly those that are less externally visible, lack distinct physiological markers, and may manifest more gradually over time, such as anxiety, mood-related symptoms, or apathy.

Overall, the algorithms developed across the retained studies demonstrated a good ability to detect the BPSD targeted. This observation aligned with the findings of other reviews, which report the promising detection or classification accuracy of AITs in dementia care contexts [35,36]. However, some limitations must be considered, as they constrain the generalizability of our findings. These include the small sample sizes used in certain studies and the limited occurrence of behavioral events, which may have affected the robustness of algorithm training. Furthermore, although the diversity of study settings and populations represents a strength of the current body of evidence, the heterogeneity in study contexts and their definitions of abnormal behaviors may have limited their comparability and the broader applicability of our findings.

Although nurses represented the majority of the health care professionals included in the studies retained, their knowledge, skills, and experiences of using AITs to detect the BPSD remain undocumented. Given their central role in detecting and managing BPSD, the lack of insight into how nurses interpret algorithmic outputs, respond to alerts, and integrate AITs into their clinical reasoning raises concerns about the clinical relevance, acceptability, and long-term sustainability of such technologies. Only 2 studies included nurses in their research teams, and only one of those was clearly led by nurse researchers [52,54]. The absence of nursing-specific expertise in the development process may result in technologies that are misaligned with real-world care dynamics, increase the risk of misinterpreting AIT-generated alerts, or inadvertently add to workloads rather than reducing them. These findings appear to contradict the existing literature, which emphasizes nurses' active involvement in developing and implementing AITs in

XSL•FO
RenderX

diverse settings and domains of care, including geriatric care [31,33].

Implications for Practice and Research

Given the lack of studies examining the efficacy of AITs for detecting the BPSD, recommendations regarding their use remain limited and cannot be generalized to clinical practice. Furthermore, the heterogeneity of devices, data types, and analytical approaches hinders the establishment of standards and limits the development of evidence-based recommendations.

Nevertheless, the proliferation of AITs in clinical and care environments is an emergent phenomenon that cannot be overlooked. To ensure that these technologies are properly understood, safely integrated, and clinically meaningful, it is essential to develop core competencies and critical awareness of AITs at an early stage. In parallel, assessing health care professionals' readiness to provide feedback for training is equally important as it allows educational initiatives to be tailored to existing knowledge, expectations, and perceived needs. This effort should be grounded in an interprofessional perspective, as the implementation and interpretation of AIT outputs typically involve diverse professional disciplines, including nurses, physicians, and allied health staff. Frameworks, such as the conceptual framework for digitally supported communication, coordination, cooperation, and collaboration model, which emphasizes communication, coordination, cooperation, and collaboration, may provide a useful basis for aligning technological solutions with clinical goals, care contexts, and each profession's roles and responsibilities [73].

Building on this foundation, several actions are needed to support the effective and ethical integration of AITs into dementia care. Educational institutions should reinforce curricula in digital health literacy, algorithmic reasoning, and AI ethics across health care professions, including training on how to interpret algorithmic outputs and assess system reliability. In clinical settings, interprofessional training (through workshops, simulations, or case reviews) can foster shared understanding and critical reflection on AIT use [63]. Health care institutions should identify clinical champions to liaise with developers, support day-to-day integration, and promote peer learning [74]. Beyond health care and academic institutions, the development of these competencies should be supported by professional and regulatory bodies, including organizations issuing evidence-based practice guidelines, as well as geriatric and psychogeriatric associations, which can help contextualize AIT within ethical and clinical standards. Experts in long-term care can also contribute to preparing position statements and evaluation frameworks. From a research perspective, promoting nurse-led, co-designed research is essential to ensure that technological solutions are aligned with the day-to-day realities of care and person-centered values. Evidence shows that nurse-led research, especially by advanced practice nurses, improves care quality, successful implementation, and health system efficiency [75,76]. These efforts are particularly needed in settings where AITs are already being piloted, such as in long-term care and psychiatric units, to ensure safe, supported, and meaningful deployment from the outset.

Most of the studies retained used classic machine learning techniques, with only 2 exploring advanced methods like deep learning, which offers a greater capacity to model complex, temporal, and heterogeneous data [24,77]. Although some studies combined behavioral and physiological data, only 1 [55] applied deep learning to such multimodal inputs, suggesting that this is an underexplored area of research. The use of deep learning to explore the interactions between internal physiological states and external behavioral manifestations of the BPSD holds promise for enhancing the sensitivity, robustness, and contextual adaptability of detection systems. This is particularly relevant given the fluctuating and multifactorial nature of the BPSD. For instance, a review by Wu et al [78] showed that deep learning–based multimodal emotion recognition models outperformed conventional models by integrating behavioral, vocal, and physiological inputs, enabling a more robust and accurate interpretation of emotional and behavioral states.

However, the application of deep learning in health care raises well-known challenges, including a lack of transparency about analysis and prediction process, the limited interpretability of outputs, and a loss of clinical trust [24,74,79]. One promising strategy to address these concerns is the adoption of explainable AI approaches. These rely on techniques such as local interpretable model-agnostic explanations, which help clarify prediction accuracy, and Deep Learning Important FeaTures, which enable traceability by comparing neuron activations to reference states [24]. Promoting such approaches is essential to ensure that AI-driven decisions are not only technically reliable but also aligned with human values and ethical standards.

Furthermore, the practical value of advanced models depends not only on technical performance but also on feasibility and user acceptance. However, limited attention has been given to clinical and contextual factors, such as comorbidities, medication use, functional status, care environments, or staff ratios, which can all affect the expression and detection of the BPSD [2,79-82]. This reflects the exploratory nature of early-stage studies, which prioritized technical development, algorithmic performance, and system functionality. As previously highlighted, several studies also reported barriers to device acceptability, further underscoring the need for future research to consider issues involving clinical complexity and contextual variability.

In this regard, it is important to explore whether and how advanced methods, such as deep learning, can adapt to such variability and enhance the detection of BPSD. Multisite experimental studies should be conducted to assess the robustness and generalizability of algorithms across diverse care settings. Furthermore, the diversity of technological approaches used to detect BPSD highlights the need for experimental or quasi-experimental comparative studies to determine which AITs offer the most effective and reliable outcomes. Such evidence is critical for guiding clinical decision-making and future technological development based on proven efficacy.

In addition, mixed methods studies incorporating user-centered evaluations (such as perceived usefulness, usability, workflow

XSL•FO
RenderX

integration, perceived technical performance, and interprofessional communication) and guided by implementation frameworks may offer valuable insights into how contextual factors shape technical outcomes and clinical integration while helping to anticipate potential barriers to adoption. Finally, longitudinal studies will be essential to assess AITs' long-term impact on the detection and management of the BPSD, on the quality of life of people with dementia and their informal caregivers, on health care professionals' workloads, and on overall cost efficiency.

Limitations

This review had some limitations. The number of relevant publications was relatively limited compared with other reviews on the use of AITs in dementia care [34,36-38], which is probably attributable to AI's relatively recent integration into the management of the BPSD. During our study screening process, a substantial number of publications referenced the use of AITs to detect dementia and evaluate functional status. However, some studies may have involved the BPSD without using this explicit terminology, which could have resulted in their exclusion during the selection of references. Also, no formal quality appraisal of the studies included was performed, as per standard scoping review methodology. However, eligibility criteria and dual screening helped ensure relevance. Despite addressing several dimensions of this topic, the overall lack of available data limited our exploration of other important contextual factors, such as costs, the availability of AITs, and health care infrastructure. This, too, restricted the scope of our recommendations.

Conclusion

The use of AITs for the early detection of BPSD represents a rapidly emerging field of research with the potential to significantly enhance the quality of dementia care. This scoping review highlighted the diversity of AITs explored across various care settings, reflecting promising new avenues for improving the detection of BPSD. However, their clinical efficacy remains to be demonstrated. Future research should evaluate whether AITs support earlier and more accurate detection of BPSD, reduce symptom severity or duration, and improve the quality of life of people with dementia and their informal caregivers. Studies should also examine AITs' impacts on clinical decision-making, care coordination, and workload management—across diverse populations and care contexts—to ensure relevance and generalizability.

To support their meaningful integration, nurses must be more actively involved in developing and evaluating these technologies. Their clinical expertise and proximity to patients make them key stakeholders in assessing the applicability of these tools in clinical settings. Their contribution should be strengthened through participatory design approaches, inclusion in pilot implementation studies, and interprofessional collaboration throughout the development cycles of new AITs.

The successful integration of AITs into dementia care will depend not only on technical innovation but also on inclusive, clinical, context-sensitive research strategies that can provide robust evidence of their impact in everyday practice.

Authors' Contributions

Conceptualization: SF, JRA, AvG, CM, HV
Data curation: SF
Formal analysis: SF, CGdR, EP, HV
Validation: JRA, AvG, CM, HV
Writing - original draft: SF
Writing - review & editing: SF, JRA, AvG, CM, HV, CGdR, EP

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategies.
[PDF File, 209 KB - aging_v8i1e76074_app1.pdf]

Checklist 1
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) fillable checklist.
[PDF File, 192 KB - aging_v8i1e76074_app2.pdf]

References

1. Dementia. World Health Organization. 2025 Mar 31. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2025-09-02]
2. Gerlach LB, Kales HC. Managing behavioral and psychological symptoms of dementia. Psychiatr Clin North Am 2018 Mar;41(1):127-139. [doi: 10.1016/j.psc.2017.10.010] [Medline: 29412841]

XSL•FO
RenderX

3. Schwertner E, Pereira JB, Xu H, et al. Behavioral and psychological symptoms of dementia in different dementia disorders: a large-scale study of 10,000 individuals. J Alzheimers Dis 2022;87(3):1307-1318. [doi: 10.3233/JAD-215198] [Medline: 35491774]
4. IPA complete guides to behavioral and psychological symptoms of dementia (BPSD). International Psychogeriatric Association. 2015. URL: https://www.ipa-online.org/publications/guides-to-bpsd [accessed 2023-02-14]
5. Bränsvik V, Granvik E, Minthon L, Nordström P, Nägga K. Mortality in patients with behavioural and psychological symptoms of dementia: a registry-based study. Aging Ment Health 2021 Jun;25(6):1101-1109. [doi: 10.1080/13607863.2020.1727848] [Medline: 32067466]
6. Cepoiu-Martin M, Tam-Tham H, Patten S, Maxwell CJ, Hogan DB. Predictors of long-term care placement in persons with dementia: a systematic review and meta-analysis. Int J Geriatr Psychiatry 2016 Nov;31(11):1151-1171. [doi: 10.1002/gps.4449] [Medline: 27045271]
7. Lu X, Ye R, Wu J, Rao D, Liao X. Comparing behavioral and psychological symptoms of dementia and caregiver distress caused between older adults with dementia living in the community and in nursing homes. Front Psychiatry 2022;13:881215. [doi: 10.3389/fpsyt.2022.881215] [Medline: 35651819]
8. Pinyopornpanish K, Soontornpun A, Wongpakaran T, et al. Impact of behavioral and psychological symptoms of Alzheimer's disease on caregiver outcomes. Sci Rep 2022 Aug 19;12(1):14138. [doi: 10.1038/s41598-022-18470-8] [Medline: 35986203]
9. Pinyopornpanish M, Pinyopornpanish K, Soontornpun A, et al. Perceived stress and depressive symptoms not neuropsychiatric symptoms predict caregiver burden in Alzheimer's disease: a cross-sectional study. BMC Geriatr 2021 Mar 12;21(1):180. [doi: 10.1186/s12877-021-02136-7] [Medline: 33711938]
10. Kang Y, Hur Y. Nurses' experience of nursing workload-related issues during caring patients with dementia: a qualitative meta-synthesis. Int J Environ Res Public Health 2021 Oct 4;18(19):10448. [doi: 10.3390/ijerph181910448] [Medline: 34639748]
11. Steele C, Berry K, Brown LJE. Healthcare professionals' experiences of using a biopsychosocial approach to understand behavioural and psychological symptoms of dementia: a qualitative interview study. Int J Older People Nurs 2022 Mar;17(2):e12427. [doi: 10.1111/opn.12427] [Medline: 34561970]
12. Butler M, Gaugler JE, Talley KMC, et al. Care interventions for people living with dementia and their caregivers. Agency for Healthcare Research and Quality (AHRQ). 2020. [doi: 10.23970/AHRQEPCCER231]
13. Tible OP, Riese F, Savaskan E, von Gunten A. Best practice in the management of behavioural and psychological symptoms of dementia. Ther Adv Neurol Disord 2017 Aug;10(8):297-309. [doi: 10.1177/1756285617712979] [Medline: 28781611]
14. Cloak N, Al Khalili Y. Behavioral and Psychological Symptoms in Dementia: StatPearls Publishing; 2023. URL: http://www.ncbi.nlm.nih.gov/books/NBK551552 [accessed 2025-10-16]
15. Ismail Z, Smith EE, Geda Y, et al. Neuropsychiatric symptoms as early manifestations of emergent dementia: provisional diagnostic criteria for mild behavioral impairment. Alzheimers Dement 2016 Feb;12(2):195-202. [doi: 10.1016/j.jalz.2015.05.017] [Medline: 26096665]
16. Creese B, Brooker H, Ismail Z, et al. Mild behavioral impairment as a marker of cognitive decline in cognitively normal older adults. Am J Geriatr Psychiatry 2019 Aug;27(8):823-834. [doi: 10.1016/j.jagp.2019.01.215] [Medline: 30902566]
17. Kales HC, Gitlin LN, Lyketsos CG. Assessment and management of behavioral and psychological symptoms of dementia. BMJ 2015 Mar 2;350:h369. [doi: 10.1136/bmj.h369] [Medline: 25731881]
18. Barreira LF, Paiva A, Araújo B, Campos MJ. Challenges to systems of long-term care: mapping of the central concepts from an umbrella review. Int J Environ Res Public Health 2023 Jan 17;20(3):1698. [doi: 10.3390/ijerph20031698] [Medline: 36767064]
19. ICN policy brief - the global nursing shortage and nurse retention. International Council of Nurses (ICN). 2021 Mar 11. URL: https://www.icn.ch/node/1297 [accessed 2023-09-06]
20. Boyne JJ, Ski CF, Fitzsimons D, Amin H, Hill L, Thompson DR. The changing role of patients, and nursing and medical professionals as a result of digitalization of health and heart failure care. J Nurs Manag 2022 Nov;30(8):3847-3852. [doi: 10.1111/jonm.13888] [Medline: 36329647]
21. Schoenborn NL, Chae K, Massare J, et al. Perspectives on AI and novel technologies among older adults, clinicians, payers, investors, and developers. JAMA Netw Open 2025 Apr 1;8(4):e253316. [doi: 10.1001/jamanetworkopen.2025.3316] [Medline: 40184066]
22. Diaz-Flores E, Meyer T, Giorkallos A. Evolution of artificial intelligence-powered technologies in biomedical research and healthcare. Adv Biochem Eng Biotechnol 2022;182:23-60. [doi: 10.1007/10_2021_189] [Medline: 35262750]
23. Maddox T, Babski D, Embi P, et al. Generative Artificial Intelligence in Health and Medicine: Opportunities and Responsibilities for Transformative Innovation: National Academies Press; 2025. [doi: 10.17226/28907]
24. Liu S, Russo C, Suero Molina E, Di Ieva A. Artificial intelligence methods. Adv Exp Med Biol 2024;1462:21-38. [doi: 10.1007/978-3-031-64892-2_3] [Medline: 39523258]
25. Artificial intelligence, critical thinking, and the nursing process. HIMSS. 2019. URL: https://legacy.himss.org/resources/artificial-intelligence-critical-thinking-and-nursing-process [accessed 2025-10-08]

XSL•FO
RenderX

26. Steijger D, Christie H, Aarts S, IJselsteijn W, Verbeek H, de Vugt M. Use of artificial intelligence to support quality of life of people with dementia: a scoping review. Ageing Res Rev 2025 Jun;108:102741. [doi: 10.1016/j.arr.2025.102741] [Medline: 40188991]
27. Loveys K, Prina M, Axford C, et al. Artificial intelligence for older people receiving long-term care: a systematic review of acceptability and effectiveness studies. Lancet Healthy Longev 2022 Apr;3(4):e286-e297. [doi: 10.1016/S2666-7568(22)00034-4] [Medline: 35515814]
28. Li R, Wang X, Lawler K, Garg S, Bai Q, Alty J. Applications of artificial intelligence to aid early detection of dementia: a scoping review on current capabilities and future directions. J Biomed Inform 2022 Mar;127:104030. [doi: 10.1016/j.jbi.2022.104030] [Medline: 35183766]
29. Grootscholten E, Poslawsky I, Bakker T. Nursing staff's observations of BPSD amongst older adults with dementia living in a nursing home: a qualitative study. Nurs Rep 2023 Jan 31;13(1):166-178. [doi: 10.3390/nursrep13010018] [Medline: 36810269]
30. Kim D, Choi YR, Lee YN, Chang SO. How do nursing home nurses conceptualize the management of behavioral and psychological symptoms of dementia? A phenomenographic study. Nurs Health Sci 2024 Jan 1. [doi: 10.1111/nhs.13079] [Medline: 38163765]
31. Raymond L, Castonguay A, Doyon O, Paré G. Nurse practitioners' involvement and experience with AI-based health technologies: a systematic review. Appl Nurs Res 2022 Aug;66:151604. [doi: 10.1016/j.apnr.2022.151604] [Medline: 35840270]
32. Sodeau A, Fox A. Influence of nurses in the implementation of artificial intelligence in health care: a scoping review. Aust Health Rev 2022;46(6):736-741. [doi: 10.1071/AH22164]
33. von Gerich H, Moen H, Block LJ, et al. Artificial Intelligence -based technologies in nursing: a scoping literature review of the evidence. Int J Nurs Stud 2022 Mar;127:104153. [doi: 10.1016/j.ijnurstu.2021.104153] [Medline: 35092870]
34. Fernandes S, von Gunten A, Verloo H. Using AI-based technologies to help nurses detect behavioral disorders: narrative literature review. JMIR Nurs 2024 May 28;7(NA):e54496. [doi: 10.2196/54496] [Medline: 38805252]
35. Jao YL, Liao YJ, Yuan F, et al. AI-assisted methods for assessing affect and behavioral symptoms in dementia: a systematic review. Innovation in Aging 2022 Dec 20;6(Supplement_1):765-765. [doi: 10.1093/geroni/igac059.2774]
36. Shaik MA, Anik FI, Hasan MM, et al. Advancing remote monitoring for patients with Alzheimer disease and related dementias: systematic review. JMIR Aging 2025 May 14;8(1):e69175. [doi: 10.2196/69175] [Medline: 40367504]
37. Husebo BS, Heintz HL, Berge LI, Owoyemi P, Rahman AT, Vahia IV. Sensing technology to monitor behavioral and psychological symptoms and to assess treatment response in people with dementia. a systematic review. Front Pharmacol 2019;10(NA):1699. [doi: 10.3389/fphar.2019.01699] [Medline: 32116687]
38. Ardelean A, Redolat R. Supporting behavioral and psychological challenges in Alzheimer using technology: a systematic review. Act Adapt Aging 2024 Jan 2;48(1):21-52. [doi: 10.1080/01924788.2023.2172900]
39. Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. JBI Manual for Evidence Synthesis: JBI; 2024. URL: https://jbi-global-wiki.refined.site/space/MANUAL [accessed 2025-10-16]
40. Peters MDJ, Godfrey C, McInerney P, Munn Z, Tricco AC, Khalil H. Scoping reviews (2020). In: Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. JBI Manual for Evidence Synthesis: JBI; 2020. URL: https://jbi-global-wiki.refined.site/space/MANUAL/4687342/Chapter+11%3A+Scoping+reviews [accessed 2024-09-06]
41. Fernandes S. Utilization of artificial intelligence-based technologies for the early detection of behavioral and psychological symptoms of dementia: a scoping review protocol. OSF 2024. [doi: 10.17605/OSF.IO/9WFPM]
42. Hirshon JM, Risko N, Calvello EJB, et al. Health systems and services: the role of acute care. Bull World Health Organ 2013 May 1;91(5):386-388. [doi: 10.2471/BLT.12.112664] [Medline: 23678202]
43. Chang HH, Chou YC, Chao CY. Community-based care. In: Maggino F, editor. Encyclopedia of Quality of Life and Well-Being Research: Springer International Publishing. [doi: 10.1007/978-3-319-69909-7_3655-2]
44. What is long-term care? National Institute on Aging. 2023 Oct 12. URL: https://www.nia.nih.gov/health/long-term-care/what-long-term-care [accessed 2024-02-16]
45. Haddaway NR, Grainger MJ, Gray CT. Citationchaser: a tool for transparent and efficient forward and backward citation chasing in systematic searching. Res Synth Methods 2022 Jul;13(4):533-545. [doi: 10.1002/jrsm.1563] [Medline: 35472127]
46. Borissov N, Haas Q, Minder B, et al. Reducing systematic review burden using Deduklick: a novel, automated, reliable, and explainable deduplication algorithm to foster medical research. Syst Rev 2022 Aug 17;11(1):172. [doi: 10.1186/s13643-022-02045-9] [Medline: 35978441]
47. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
48. Pollock D, Peters MDJ, Khalil H, et al. Recommendations for the extraction, analysis, and presentation of results in scoping reviews. JBI Evid Synth 2023 Mar 1;21(3):520-532. [doi: 10.11124/JBIES-22-00123] [Medline: 36081365]
49. Zhang C, Lu Y. Study on artificial intelligence: the state of the art and future prospects. Journal of Industrial Information Integration 2021 Sep;23:100224. [doi: 10.1016/j.jii.2021.100224]

XSL•FO
RenderX

50. Bafaloukou M, Schalkamp AK, Fletcher-Lloyd N, et al. An interpretable machine learning tool for in-home monitoring of agitation episodes in people living with dementia: a proof-of-concept study. EClinicalMedicine 2025 Feb;80(NA):103032. [doi: 10.1016/j.eclinm.2024.103032] [Medline: 39896868]
51. Choukou MA, Mbabaali S, East R. Healthcare professionals' perspective on implementing a detector of behavioural disturbances in long-term care homes. Int J Environ Res Public Health 2021 Mar 8;18(5):2720. [doi: 10.3390/ijerph18052720] [Medline: 33800257]
52. Iaboni A, Spasojevic S, Newman K, et al. Wearable multimodal sensors for the detection of behavioral and psychological symptoms of dementia using personalized machine learning models. Alz & Dem Diag Ass & Dis Mo 2022 Jan;14(1):e12305. [doi: 10.1002/dad2.12305]
53. Kim K, Jang J, Park H, Jeong J, Shin D, Shin D. Detecting abnormal behaviors in dementia patients using lifelog data: a machine learning approach. Information 2023;14(8):433. [doi: 10.3390/info14080433]
54. Lee JY, Yang E, Cho AY, Choi Y, Lee S, Lee KH. Sleep efficiency in community-dwelling persons living with dementia: exploratory analysis using machine learning. J Clin Sleep Med 2025 Feb 1;21(2):393-400. [doi: 10.5664/jcsm.11436] [Medline: 39484800]
55. Palermo F, Li H, Capstick A, et al. Designing a clinically applicable deep recurrent model to identify neuropsychiatric symptoms in people living with dementia using in-home monitoring data. . Preprint posted online on Dec 2, 2021. [doi: 10.48550/arXiv.2110.09868]
56. Zhu Y, Song T, Zhang Z, et al. Agitation prevalence in people with dementia in Australian residential aged care facilities: findings from machine learning of electronic health records. J Gerontol Nurs 2022 Apr;48(4):57-64. [doi: 10.3928/00989134-20220309-01] [Medline: 35343838]
57. Bourennane W, Charlon Y, Bettahar F, Campo E, Esteve D. Homecare monitoring system: a technical proposal for the safety of the elderly experimented in an Alzheimer's care unit. IRBM 2013 Apr;34(2):92-100. [doi: 10.1016/j.irbm.2013.02.002]
58. Lotfi A, Langensiepen C, Mahmoud SM, Akhlaghinia MJ. Smart homes for the elderly dementia sufferers: identification and prediction of abnormal behaviour. J Ambient Intell Human Comput 2012 Sep;3(3):205-218. [doi: 10.1007/s12652-010-0043-x]
59. Enshaeifar S, Zoha A, Markides A, et al. Health management and pattern analysis of daily living activities of people with dementia using in-home sensors and machine learning techniques. PLOS ONE 2018;13(5):e0195605. [doi: 10.1371/journal.pone.0195605] [Medline: 29723236]
60. Rostill H, Nilforooshan R, Barnaghi P, Morgan A. Technology-integrated dementia care: trial results. Nursing and Residential Care 2019 Sep 2;21(9):489-494. [doi: 10.12968/nrec.2019.21.9.489]
61. König A, Linz N, Zeghari R, et al. Detecting apathy in older adults with cognitive disorders using automatic speech analysis. J Alzheimers Dis 2019;69(4):1183-1193. [doi: 10.3233/JAD-181033] [Medline: 31127764]
62. Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks. Information Processing & Management 2009 Jul;45(4):427-437. [doi: 10.1016/j.ipm.2009.03.002]
63. Ta'an W, Damrah S, Al-Hammouri MM, Williams B. Professional identity and its relationships with AI readiness and interprofessional collaboration. PLOS ONE 2025;20(5):e0322794. [doi: 10.1371/journal.pone.0322794] [Medline: 40378129]
64. Li H, Zhu G, Zhong Y, Zhang Z, Li S, Liu J. Applications of artificial intelligence in psychiatric nursing: a scope review. Stud Health Technol Inform 2024 Jul 24;315:74-80. [doi: 10.3233/SHTI240109] [Medline: 39049229]
65. De Santis KK, Mergenthal L, Christianson L, Busskamp A, Vonstein C, Zeeb H. Digital technologies for health promotion and disease prevention in older people: scoping review. J Med Internet Res 2023 Mar 23;25(1):e43542. [doi: 10.2196/43542] [Medline: 36951896]
66. Sohn M, Yang J, Sohn J, Lee JH. Digital healthcare for dementia and cognitive impairment: a scoping review. Int J Nurs Stud 2023 Apr;140:104413. [doi: 10.1016/j.ijnurstu.2022.104413] [Medline: 36821951]
67. Tiersen F, Batey P, Harrison MJC, et al. Smart home sensing and monitoring in households with dementia: user-centered design approach. JMIR Aging 2021 Aug 11;4(3):e27047. [doi: 10.2196/27047] [Medline: 34383672]
68. Li C, Wang J, Wang S, Zhang Y. A review of IoT applications in healthcare. Neurocomputing 2024 Jan;565:127017. [doi: 10.1016/j.neucom.2023.127017]
69. Messinis S, Temenos N, Protonotarios NE, Rallis I, Kalogeras D, Doulamis N. Enhancing internet of medical things security with artificial intelligence: a comprehensive review. Comput Biol Med 2024 Mar;170:108036. [doi: 10.1016/j.compbiomed.2024.108036] [Medline: 38295478]
70. Manickam P, Mariappan SA, Murugesan SM, et al. Artificial intelligence (AI) and internet of medical things (IoMT) assisted biomedical systems for intelligent healthcare. Biosensors (Basel) 2022 Jul 25;12(8):562. [doi: 10.3390/bios12080562] [Medline: 35892459]
71. Zhao QF, Tan L, Wang HF, et al. The prevalence of neuropsychiatric symptoms in Alzheimer's disease: systematic review and meta-analysis. J Affect Disord 2016 Jan 15;190:264-271. [doi: 10.1016/j.jad.2015.09.069] [Medline: 26540080]
72. Maust DT, Kales HC, McCammon RJ, Blow FC, Leggett A, Langa KM. Distress associated with dementia-related psychosis and agitation in relation to healthcare utilization and costs. Am J Geriatr Psychiatry 2017 Oct;25(10):1074-1082. [doi: 10.1016/j.jagp.2017.02.025] [Medline: 28754586]

XSL•FO
RenderX

73. Nordmann K, Redlich MC, Schaller M, Sauter S, Fischer F. Toward a conceptual framework for digitally supported communication, coordination, cooperation, and collaboration in interprofessional health care: scoping review. J Med Internet Res 2025 May 26;27(1):e69276. [doi: 10.2196/69276] [Medline: 40418789]
74. Chen M, Decary M. Artificial intelligence in healthcare: an essential guide for health leaders. Healthc Manage Forum 2020 Jan;33(1):10-18. [doi: 10.1177/0840470419873123] [Medline: 31550922]
75. Kilpatrick K, Savard I, Audet LA, et al. A global perspective of advanced practice nursing research: a review of systematic reviews. PLOS ONE 2024;19(7):e0305008. [doi: 10.1371/journal.pone.0305008] [Medline: 38954675]
76. Eckert M, Kennedy K, Neylon K, et al. A scoping review of nurse-led randomised controlled trials. J Clin Nurs 2023 Sep;32(17-18):5550-5561. [doi: 10.1111/jocn.16632] [Medline: 36737840]
77. Li FF, Manning C, Reuel A. The standford emerging technology review 2025. : Standford University; 2025 URL: https://setr.stanford.edu/sites/default/files/2025-02/SETR2025_01-AI_web-240128.pdf [accessed 2025-07-01]
78. Wu Y, Mi Q, Gao T. A comprehensive review of multimodal emotion recognition: techniques, challenges, and future directions. Biomimetics (Basel) 2025 Jun 27;10(7):418. [doi: 10.3390/biomimetics10070418] [Medline: 40710231]
79. Bitkina OV, Park J, Kim HK. Application of artificial intelligence in medical technologies: a systematic review of main trends. Digit Health 2023;9:20552076231189331. [doi: 10.1177/20552076231189331] [Medline: 37485326]
80. Bessey LJ, Walaszek A. Management of behavioral and psychological symptoms of dementia. Curr Psychiatry Rep 2019 Jul 1;21(8):66. [doi: 10.1007/s11920-019-1049-5] [Medline: 31264056]
81. Altomari N, Bruno F, Laganà V, et al. A comparison of behavioral and psychological symptoms of dementia (BPSD) and BPSD sub-syndromes in early-onset and late-onset Alzheimer's disease. J Alzheimers Dis 2022;85(2):691-699. [doi: 10.3233/JAD-215061] [Medline: 34864668]
82. McCabe MP, Mellor D, Karantzas G, Von Treuer K, Davison TE, O'Connor D. Organizational factors related to the confidence of workers in working with residents with dementia or depression in aged care facilities. Aging Ment Health 2017 May;21(5):487-493. [doi: 10.1080/13607863.2015.1118011] [Medline: 26666515]

Abbreviations

AI: artificial intelligence
AIT: artificial intelligence–based technology
BPSD: behavioral and psychological symptoms of dementia
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by A Garcia; submitted 16.04.25; peer-reviewed by D Steijger, SQ Yoong; revised version received 03.09.25; accepted 05.09.25; published 21.10.25.

Please cite as:
Fernandes S, Rosselet Amoussou J, Gomes da Rocha C, Perruchoud E, von Gunten A, Mabire C, Verloo H
Using Artificial Intelligence–Based Technologies for the Early Detection of Behavioral and Psychological Symptoms of Dementia: Scoping Review
JMIR Aging 2025;8:e76074
URL: https://aging.jmir.org/2025/1/e76074
doi:10.2196/76074

© Sofia Fernandes, Joëlle Rosselet Amoussou, Carla Gomes da Rocha, Elodie Perruchoud, Armin von Gunten, Cédric Mabire, Henk Verloo. Originally published in JMIR Aging (https://aging.jmir.org), 21.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effectiveness of mHealth-Based Gamified Interventions on Physical Activity in Older Adults: Systematic Review

Lin Chen[1], MS; Fenglin Jang[1,2], MD; Min Li[3], MS; Wei Zong[4], MS; Huiqin Yu[5], MS

[1]Science and Education Training Department, Pengzhou City People's Hospital, No. 255, Section 2, South Third Ring Road, Zhihe Town, Pengzhou City, Chengdu City, China

[2]School of Nursing, National University of Singapore, Singapore, Singapore

[3]Endocrinology and Metabolism Department, First Hospital of China Medical University, Shenyang, China

[4]Intensive Care Unit, Affiliated Hospital of Xuzhou Medical College, Xuzhou, China

[5]Endocrinology and Metabolism Department, Changsha People's Hospital, Changsha, China

Corresponding Author:
Lin Chen, MS
Science and Education Training Department, Pengzhou City People's Hospital, No. 255, Section 2, South Third Ring Road, Zhihe Town, Pengzhou City, Chengdu City, China

Abstract

Background: Global aging presents significant socioeconomic and health challenges, particularly for older adults who face an increased risk of chronic diseases and reduced physical activity levels. Although physical activity is crucial for maintaining health, most older adults do not meet the recommended guidelines. Gamification and mobile health (mHealth) technologies offer innovative solutions to motivate physical activity; however, research focusing on older adults is limited, especially regarding the effectiveness and sustainability of such interventions.

Objective: This study aims to synthesize evidence on the effectiveness of mHealth-based gamified interventions for improving physical activity in older adults.

Methods: This systematic review followed Preferred Reporting Items for Systematic Reviews and Meta-Analyses and Meta-Analysis of Observational Studies in Epidemiology guidelines and analyzed studies from PubMed, Embase, Web of Science, CINAHL, Scopus, and Wiley Online Library, covering relevant literature from their inception up to May 2025. The inclusion criteria focused on gamified mHealth interventions for adults aged 60+ years, excluding serious games. Quality assessment was conducted according to the Joanna Briggs Institute standards, with data extracted on study design, gamification elements, and outcomes such as step counts and moderate-to-vigorous physical activity.

Results: Of 2944 studies identified from the database search, 1454 individuals from 8 trials were included. Gamified interventions significantly increased daily step counts and time spent in moderate-to-vigorous physical activity among older adults. Goal setting and rewards were the most frequently used components, and the combined use of mobile and wearable devices offered greater flexibility and accessibility. A classification framework indicated that interventions integrating multiple gamification elements with hybrid technology systems were most effective, particularly when guided by a theoretical basis. However, the heterogeneity in study designs, small sample sizes, and lack of long-term follow-up studies limited the generalizability of the findings.

Conclusions: mHealth-based gamification interventions demonstrate potential for increasing physical activity in older adults. Future interventions should consider employing multifaceted designs combining advanced gamification with hybrid technology systems, while also prioritizing theoretical integration, long-term sustainability, and caregiver involvement to improve sustainability and inclusivity. This review highlights the need for theory-driven, technology-mediated strategies that address the unique health needs of older adults.

(JMIR Aging 2025;8:e78686) doi:10.2196/78686

KEYWORDS

gamification; geriatric; mobile health; older adults; physical activity; Preferred Reporting Items for Systematic Reviews and Meta-Analyses; PRISMA; systematic review

Introduction

The issue of global aging is becoming increasingly pronounced. According to United Nations statistics, by 2050, the global population aged 60 years and older is projected to reach 2.1 billion, accounting for 22% of the total global population [1]. The aging of the population not only has a profound impact on the socioeconomic landscape but also presents substantial challenges in the health management of older adults. As

XSL•FO
RenderX

individuals age, they commonly experience rising rates of chronic diseases, declines in physical function, and reductions in cognitive abilities—factors that collectively affect their quality of life and capacity for self-care [2].

Research indicates that daily physical activity plays a crucial role in maintaining the physical and mental health of older adults. Moderate exercise can significantly reduce the risk of chronic diseases such as cardiovascular disease, diabetes, and osteoporosis, while also promoting cognitive function and mental well-being [3]. However, the reality is that the daily activity levels of most older adults fall far below the recommended health guidelines. A sedentary lifestyle has become a prevalent issue, and insufficient physical activity accelerates functional decline, significantly increasing the risk of falls, disability, and social isolation [4]. Thus, identifying effective strategies to enhance the physical activity levels of older adults is a critical issue that requires urgent attention in public health.

Gamification, as an innovative behavioral intervention strategy, has garnered increasing attention in health promotion in recent years. Gamification involves incorporating game design elements—such as points, leaderboards, progress bars, and badges—into nongame contexts to motivate and enhance positive user behavior [5]. Existing studies have shown that gamification-based interventions are highly effective in promoting health behaviors, rehabilitation training, and chronic disease management, significantly improving participants' motivation and adherence [6-8].

Mobile health (mHealth) refers to a comprehensive technological system that employs mobile devices, sensors, and digital media technologies for health monitoring, disease management, and health promotion [9]. Compared to traditional gamified interventions, mHealth gamified interventions provide remote health management and real-time feedback and integrate personalized, adaptable health content. This enables older adults to engage in health activities more flexibly within their daily routines. Moreover, the widespread use of mobile devices allows this intervention model to reach a broader population of older adults, particularly in home care and telemedicine settings, where its advantages are particularly pronounced [10].

Although gamified interventions combined with mHealth technology show considerable potential in enhancing the activity levels of older adults, challenges such as cognitive decline, lower education levels, and limited digital proficiency among older adults have hindered their broader adoption [11,12]. As a rapidly growing group of technology users, older adults have unique needs and face significant challenges, but they are expected to undergo a phase of rapid development in the near future. However, there is currently a lack of systematic quantitative research that comprehensively evaluates the actual effects of gamified interventions on the physical functioning and activity levels of older adults. Therefore, synthesizing existing empirical studies into a systematic review to clarify their effectiveness and applicability holds significant academic and practical value.

Specifically, this study focuses on the following core issues:

1. Does gamification based on mHealth offer advantages over traditional interventions in promoting physical activity in older adults?
2. What are the most frequently used gamification elements and mHealth types in the existing studies?
3. Is the process of gamification design integrated with behavioral or other related theories or principles?
4. What factors may affect the feasibility and effectiveness of gamified interventions for older adults?

Methods

Overview

The reviewed protocol was developed and approved by all authors prior to initiation. It complies with all recommendations from the systematic review, strictly adhering to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist and Meta-Analysis of Observational Studies in Epidemiology guidelines [13]. Literature searches were performed across the following databases: PubMed, Web of Science, Embase, CINAHL, Scopus, and Wiley Online Library. The search period spanned from each database's inception until May 20, 2025. This study was registered with the International Prospective Register of Systematic Reviews (PROSPERO; registration number CRD420251056689).

Search Strategy

In most studies on gamification, younger individuals are typically the primary target group for game-based interventions due to their generally higher sense of self-efficacy and greater experience with digital gaming. In this context, the study initially focused on the research subjects and intervention measures but overlooked other retrieval elements when designing the retrieval strategy. However, during the subsequent literature screening, these overlooked elements will be reassessed and evaluated. To maximize the breadth of the search, the study selected the keywords "gamification," "game elements," and "elderly" and refined the retrieval strategy by incorporating a combination of controlled vocabulary, natural language terms, and synonyms, in accordance with the requirements of different databases. An additional file shows the search strategy in more detail (Multimedia Appendix 1).

Literature Inclusion Criteria

To ensure that the analyzed literature aligns with the objectives and to mitigate the impact of conceptual confusion on the validity of the research results, this review clarifies the distinction between serious games and gamification, incorporating this differentiation throughout the literature selection process. Gamification refers to the integration of game elements, such as points, leaderboards, and progress bars, into nongame environments to enhance user engagement and motivation through game-like experiences, without altering the nature of the tasks. For example, in rehabilitation training for older adults, adding a reward system or virtual progress bar can extend training time [14].

Serious games, on the other hand, represent a comprehensive design approach that provides an immersive gaming experience, with the primary goal of education, training, or therapy, rather

XSL•FO
RenderX

than entertainment. They are sometimes referred to as purpose-driven games (eg, virtual reality rehabilitation training games for older adults) [15,16]. However, the distinction between these 2 concepts may be ambiguous and subjective. When these concepts cannot be clearly differentiated, the research team engages in discussion and makes a final decision.

The studies included in this systematic review must meet the following criteria: (1) studies targeting older adults with a mean age ≥60 years and minimum age threshold aligned with World Health Organization functional aging criteria (typically ≥55 y). When the exact minimum age was unreported, statistical estimation (mean − 2SD) confirmed ≥90% of the participants met age criteria. (2) The use of gamification or at least 1 game element is clearly specified. Serious games (eg, video games, immersive virtual reality games, and augmented reality games) or studies that merely mention gamification without analysis are excluded. (3) Gamification is implemented through digital devices such as tablets, smartphones, or wearables. (4) At least 1 change related to physical or social activities is described, which may be presented as subjective self-reports or objective measurement indicators. (5) The study is original empirical research (must be experimental in nature). Review articles, research proposals, guidelines, policy documents, and research programs are excluded. (6) Only peer-reviewed, published papers in English are included, including doctoral dissertations and conference papers.

Quality Assessment

The research team conducted quality assessments of the studies that met the inclusion criteria, following the quality assessment standards established by the Joanna Briggs Institute [17,18]. There are 13 criteria for evaluating randomized controlled trials (RCTs) and 9 criteria for evaluating quasi-experimental studies. Studies that meet at least 50% of the assessment criteria are accepted for review. For RCTs, the research team decided to exclude the blinding of researchers and participants from the quality assessment, as blinding is considered impractical in the context of gamified interventions. Any disputes are resolved through discussions with a third-party investigator to reach a final consensus.

Data Extraction and Analysis

Two researchers independently extracted data from the selected studies using standardized forms based on the content of the literature. The extracted data included the first author's name, publication year, country, study design, subject characteristics (sample size, mean age, and gender distribution), intervention characteristics (study setting, type of equipment, and duration), gamification characteristics (game name, game elements, and theories used), primary and secondary outcomes related to physical activity, and a summary of the research conclusions. Any disputes were resolved through discussions with a third-party investigator to reach a final consensus.

Due to substantial methodological and reporting differences that prevented quantitative data synthesis, we conducted a narrative synthesis organized by the PICO framework to systematically summarize the findings of the studies.

To systematically address heterogeneity in gamification elements and technological platforms, we developed a dual-axis classification framework. The first dimension categorizes interventions by gamification complexity, distinguishing basic designs employing 1 or 2 core game elements from intermediate configurations integrating 3 or 4 elements, and advanced systems incorporating 5 or more interactive components. The second dimension classifies technological integration into standalone mobile apps, hybrid systems combining mobile apps with wearable devices, and wearable-centric approaches relying primarily on sensor technology. This framework enables structured cross-study comparisons while acknowledging inherent methodological variations.

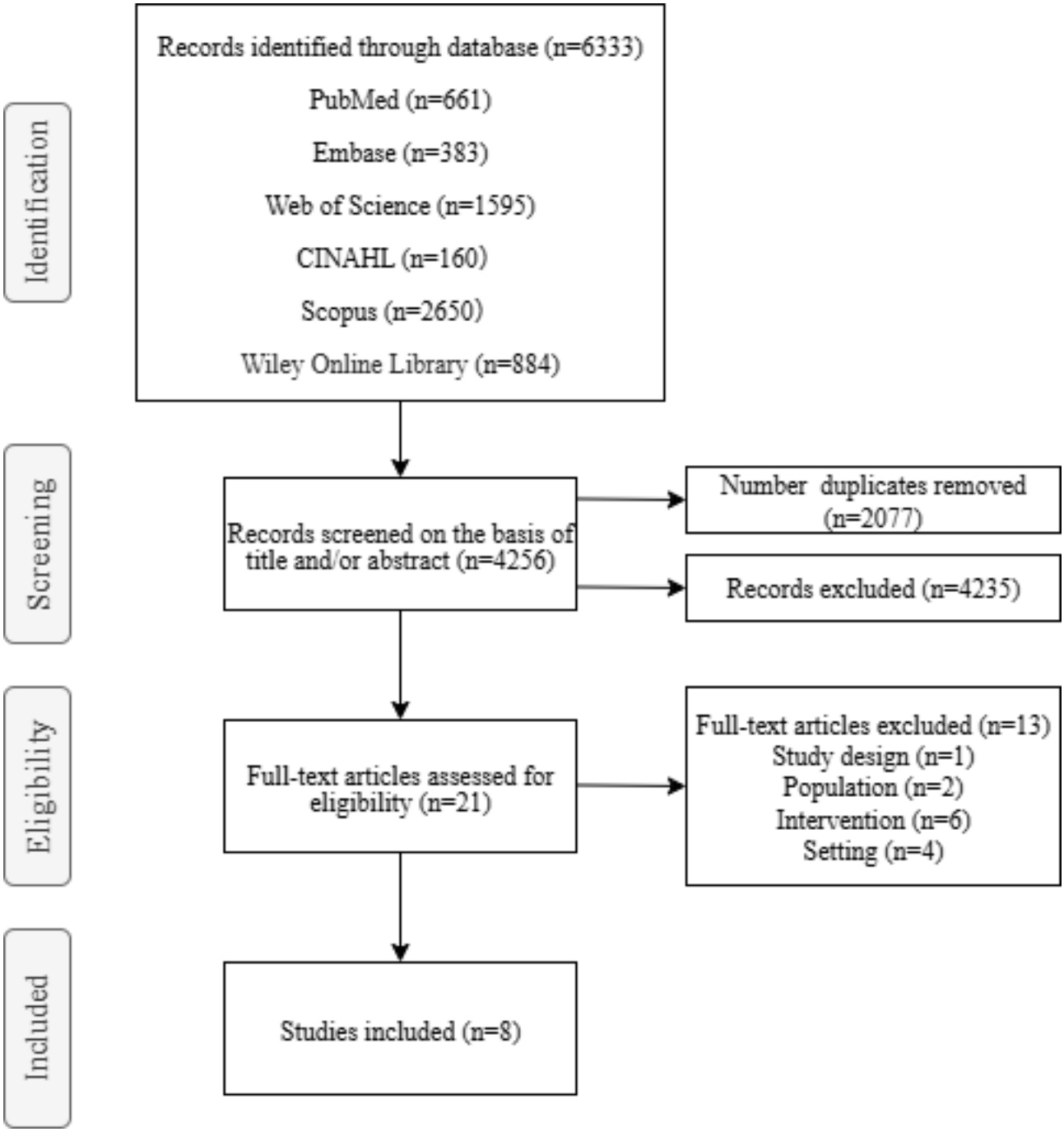
Results

Search Outcomes

A total of 6333 records were identified. After removing 2077 duplicate entries using EndNote, the remaining 4256 records were screened by title and abstract. Following rigorous assessment, a substantial number of records were excluded, primarily due to ineligible age, nonphysical activity focus, or the others (Multimedia Appendix 2). Subsequently, 21 articles were initially included. Reviews, clinical record comparisons, case reports, and other literature that did not meet the inclusion criteria were excluded based on their titles and abstracts. Following a more detailed examination of the full text, 13 articles were further excluded for not meeting the inclusion criteria. An additional file shows this in more detail (Multimedia Appendix 3). Ultimately, 8 studies [14,19-25] were included in this review. The search results are presented in Figure 1.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow chart.



Study Characteristics

The studies selected for this review were conducted between 2018 and 2024 in Sweden [14], Portugal [19], Japan [20,21,24], Brazil [21], and the United States [23,25]. At least 5 studies [14,22-25] were randomized controlled trials, while 3 studies [19-21] were quasi-experimental studies. Sample sizes ranged from 12 to 1062 participants, with a total of 1454 participants. The average age of participants ranged from 62 to 81.5 years, with a higher percentage of women. At least 4 studies [20-22,24] included community-dwelling older adults, 1 study [14] included participants with musculoskeletal conditions, 1 study [23] included participants at risk of cardiovascular disease, 1 study [25] included participants at risk of Alzheimer disease, and 1 study [19] did not specify participant characteristics or the gender ratio. A summary of the participants, measures, interventions, controls, and outcomes is presented in Table 1.

XSL•FO
RenderX

Table . Study characteristics.

Authors/year/country	Duration /study design	Participants	Gamification characteristics	Intervention(s)	Results	Conclusions
		Characteristics/age/sample size (female, n, %)	Game name/elements/theory used	Sample/program/mobile device		
Randriambelonoro et al [14] 2023 (Switzerland)	RCT[a] 3-week intervention +3-week follow-up period	Older adults with musculoskeletal problems Mean (SD) age: 81.5 (6.8) y Sample size: 57 (39, 68.4%)	Name: activLife Elements: challenges/customization/goals/progress bars Theory: none	IG[b] (35): 3 times/week gamified rehabilitation training using Stepwatch activity monitor combined with ActivLife device CG[c] (22): standard rehabilitation treatments	Steps were slightly higher but not significantly different in IG (IG:1839 vs CG:1504 steps, *P*=.25)	Game-based rehabilitation training is as effective as standard care in restoring functional ability and has a promoting effect on daily activity
Martinho et al [19] 2023 (Portugal)	Single-group (pre-post) study 4-week baseline +1-week intervention period	older adults aged 55 years and above Mean age: 62.0 y Sample size: 12 (no data)	Name: coaFEId Elements: levels/challenges/rewards/feedback/customization/progress bars Theory: circumplex model of affect (based on the dimensions of wakefulness and valence) [26]	IG(12): coaFEId application installed on smartphones	Significant difference in daily steps between pre-post intervention (2532 vs 4931 steps, Cohen *d*=0.65, *P*=.03)	Game-based agent system significantly improved daily steps of older adults in the short term and had high availability
Santos et al [20] 2018 (Japan)	Single-group (pre-post) study 1-week baseline +2-week intervention period	Older adults in communities Mean (SD) age: 75.0 (3.4) y Sample size: 12 (9, 75%)	Name: shinpo Elements: rewards/challenges/collaboration/goals/avatars Theory: none	IG(9): collecting cards by smartphones that are preinstalled with Shinpo	7.6% - 8.1% more steps than the baseline (mean: 22,567.2, all participants)	Gamification improves the activities of older adults, but the mechanism of social interaction needs to be optimized
Santos et al [21] 2019 (Brazil)	Quasi-experimental Study 4-week intervention period	Older adults in communities Mean (SD) age: 62.3 (5.8) y Sample size: 18 (14, 78%)	Name: trilhas Elements: rewards/challenges/goals/social support/collaboration Theory: none	IG(11): trilhas application installed on smartphones with social interaction function CG(7): trilhas application installed on smartphones without social interaction function	Step counts significantly increased in IG (absolute change: η^2=0.19, *P*=.01; percent change: η^2=0.27, *P*=.009)	Social interaction in gamification can effectively improve the level of physical activity in older adults
Santos et al [22] 2021 (Japan)	RCT 1-week baseline +3-week intervention period	Older adults in communities Mean (SD) age: 63.4 (6.8) y Sample size: 18 (14, 78%)	Name: shinpo Elements: rewards/goals/collaboration/social support Theory: None	IG(9): shinpo application installed on smartphones with social interaction function to collect cards CG(9): shinpo application installed on smartphones without social interaction function to collect cards	Step counts significantly increased in IG (absolute change: η^2=0.31, *P*=.04; percent change: η^2=0.30, *P*=.03)	Social interaction in gamification can effectively improve the level of physical activity in older adults

XSL•FO
RenderX

Authors/year/country	Duration /study design	Participants	Gamification characteristics	Intervention(s)	Results	Conclusions
		Characteristics/age/sample size (female, n, %)	Game name/elements/theory used	Sample/program/mobile device		
Fanaroff et al [23] 2024 (America)	RCT 12-month intervention +6-month follow-up period	Older adults with high risk of cardiovascular Mean (SD) age: 67 (8.0) y Sample size: 1062 (648, 61%)	Name: none Elements: levels/goals/points/social support/rewards Theory: prospect theory [27]	G1(304): gamified intervention based on Fitbit Charge (wearable device)/smartphone/tablet G2(302): financial incentive intervention based on Fitbit Charge (wearable device)/smartphone/tablet G3(305): gamification+financial incentives CG(151): feedback only	Significant difference in daily steps (538 steps, *P*=.003; and follow-up 459.8 steps, *P*=.017) and MVPA (min/wk) (15.4, *P*=.033) between G1 and CG	Gamification significantly improved the activity of older adults and was more suitable for clinical promotion because of its lower cost.
Kawaguchi et al [24] 2024 (Japan)	RCT 12-week intervention period	Older adults in communities Mean (SD) age: 70.0 (6.3) y Sample size: 181 (100, 55.3%)	Name: ESP Elements: levels/feedback Theory: behavior change [28]	IG(87): ESP application installed on smartphones combined with Google Fit CG(94): Google Fit only	Significant difference in activity frequency (times/2 mo) (3.03, *P*=.04) but not in weekly steps	ESP significantly improved the activity of older adults but was ineffective for step counts
Greysen et al [25] 2024 (America)	RCT 12-week intervention +3-week follow-up period	Older adults at risk of Alzheimer disease Mean age: 70.6 (3.1) y Sample size: 94 (73, 78%)	Name: none Elements: rewards/levels/goals/feedback/collaboration Theory: prospect theory [27]	IG(44): text messages–based gamification and social support combined with FitBit device/application CG(50): FitBit device/application only	Significant difference in daily steps (Δ=1699 steps, *P*<.0001) and MVPA (min/day) (Δ=6.6, *P*=.017 between 2 groups	Gamification combined with supportive partners can effectively increase physical activity in high-risk older adults, and the effect is sustainable

[a]RCT: randomized controlled trial.

[b]IG: intervention group.

[c]CG: control group.

Risk of Bias

The quality assessment scores for the RCT studies ranged from 7 to 10 points (Table 2). All 5 studies employed randomization; however, the concealment of treatment group allocation was unclear in 3 studies. Due to the nature of the intervention, blinding of participants or personnel was not possible. In 3 studies [23-25], outcome assessors were blinded to the treatment allocation. Control of confounding factors was unclear in 1 study [14]. The average score was 9/11, indicating that all the RCT studies demonstrated high methodological quality.

The 3 quasi-experimental studies each received a score of 6 points (Table 3) for methodological quality. In 1 study [21], the age of participants in the 2 groups was not similar, and the characteristics of health status and technical proficiency were not clearly presented. The criteria for this assessment were not explicit. None of the studies reported in detail whether participants in the intervention group received any other care or treatments besides the intervention. At least 2 studies [19,20] employed before-and-after designs without an independent control group. Another 2 studies [20,21] had incomplete follow-ups, and although the reasons for dropout were stated, the impact of dropout on the results was not analyzed. At least 1 study [20] only reported descriptive statistics (means, standard deviations) and did not perform significance testing or calculate effect sizes. The statistical methods used were relatively simple.

XSL•FO
RenderX

Table . Quality assessment scores of the selected RCT studies[a].

Authors (year)	1	2	3	4	5	6	7	8	9	10	11	12	13	Total
Randriambelonoro et al [14] 2023	Y	N	Y	N/A	N/A	?	?	Y	N	Y	Y	Y	Y	7/11
Santos et al [20] 2018	Y	?	Y	N/A	N/A	?	Y	Y	Y	Y	Y	Y	Y	9/11
Fanaroff et al [23] 2024	Y	?	Y	N/A	N/A	Y	Y	Y	Y	Y	Y	Y	Y	10/11
Kawaguchi et al [24] 2024	Y	Y	Y	N/A	N/A	Y	Y	Y	Y	Y	N	Y	Y	10/11
Greysen et al [25] 2024	Y	Y	Y	N/A	N/A	Y	Y	Y	N	Y	Y	Y	Y	10/11

[a]Rating scale: Yes (Y), No (N), Unclear (?), Not applicable (N/A); items for critical appraisal: 1. Were the participants' group assignments sufficiently randomized? 2. Was the control group's allocation concealed? 3. At baseline, were the groups comparable? 4. Were staff members and participants blinded to the intervention's allocation? 5. Were the researchers blinded to the allocation of the intervention? 6. Did the outcome assessors have blinded access to the allocation of the intervention? 7. Aside from the intervention, did the control group receive the same care as the intervention group? 8. Was the follow-up completed? If not, were group differences sufficiently explained and analyzed? 9. Were participants assessed with the goal of randomization in mind? 10. Did all intervention groups employ the same techniques for measuring outcomes? 11. Were the outcome measurements accurate? 12. Was the correct statistical analysis applied? 13. Was the trial design suitable? Did the analysis and interpretation account for any deviations from the conventional RCT design?

Table . Quality assessment scores of the selected quasi-experimental studies[a].

Authors (year)	1	2	3	4	5	6	7	8	9	Total
Martinho et al [19] 2023	Y	Y	?	N	Y	N	Y	Y	Y	6/9
Santos et al [21] 2019	Y	Y	?	N	Y	Y	Y	Y	N	6/9
Santos et al [22] 2021	Y	N	?	Y	Y	N	Y	Y	Y	6/9

[a]Rating scale: Yes (Y), No (N), Unclear (?), Not applicable (NA); items for critical appraisal: (1) Are the study's "cause" and "effect" (ie, the order of the variables) evident? (2) Are the individuals being compared similar? (3) Aside from the intervention, do the individuals in the comparison group receive the same care or treatment? (4) Is a control group present? (5) Are the results measured more than once before and after the intervention? (6) Has the follow-up been completed? If not, are the differences sufficiently explained and analyzed? (7) Do the individuals being compared have the same outcome measured in the same way? (8) Is the outcome measurement accurate? (9) Is the correct statistical analysis applied?

Characteristics of the Described Interventions

Table 1 and Multimedia Appendix 4 display the gamification features of the studies included in our systematic review. The number of game elements used in the gamification interventions for older adults' physical activity ranged from 2 to 6, with the majority incorporating 5 game elements. The most commonly used game elements were goals and rewards, followed by challenges and collaboration, while virtual identity and points were the least used.

Half of the 8 studies [19,23-25] employed theories or principles to design the gamification interventions. At least 2 studies [23,25] conducted in the United States used prospect theory from behavioral economics, while the other 2 studies employed the theory of behavioral change [24] and the emotional

XSL•FO
RenderX

circumplex model [19]. All of these studies used a single theory without combining them.

In 1 study, family members or friends were incorporated into the gamified intervention as support partners to provide encouragement and assist participants in achieving their activity goals.

There are 3 digital approaches to gamification for older adults. Most studies [19-22] use mobile apps only (4/8), while some studies [14,24,25] combine mobile apps with activity monitors (3/8), and only 1 study [23] uses activity monitors alone (eg, Stepwatch or Fitbit). The shortest intervention time was 1 week [19], and the longest was 12 months [23]. Only 3 studies [18,23,25] included a follow-up period, highlighting the need for further evaluation of gamification interventions to assess their long-term sustainability. Another 5 studies [14,21,22,24,25] compared a gamification intervention to usual care, while 2 studies [19,20] did not include a control group. The study by Fanaroff et al [23] assigned participants to 4 groups and compared the effects of gamification to financial incentives, a combined intervention of both, and explored the impact of gamification on physical activity in older adults.

Effectiveness of Interventions

The included studies all assessed the effect of gamification interventions on the number of steps taken by older adults in their daily lives, as measured by activity monitors or smartphone sensors such as Google Fit. The findings from both controlled and single-arm studies were largely consistent in confirming the positive impact of gamification interventions on step count, including results observed during the follow-up period. The only exception was the study by Kawaguchi et al [24], which reported a nonsignificant increase in the number of steps following the gamification intervention. However, this study did report that intervention participants had a higher frequency of participation in specific types of activities, particularly hobbies and culture-related activities, compared to the control group.

At least 2 RCTs [23,25] reported changes in the time spent in moderate-to-vigorous physical activity (MVPA) per day, as measured by activity monitors. After adjusting for factors such as baseline activity levels, proficiency with the devices, and genetic risk, between-group differences indicated that gamification can significantly increase the duration of MVPA in older adults, with these effects being sustainable.

Patterns From Intervention Classification

Application of the classification framework revealed consistent patterns across the 8 included studies. Advanced gamification designs consistently demonstrated superior outcomes [23,25], particularly when integrated with hybrid technology systems, showing statistically significant improvements in both step counts and MVPA duration [23,25]. Basic gamification approaches proved effective for step count enhancement [20,21] but exhibited limited capacity for tracking complex activity metrics [19]. Standalone mobile apps demonstrated strengths in facilitating social interaction mechanisms [20,22], while wearable-centric implementations may suffer from scalability limitations [14]. Crucially, all advanced interventions incorporated theoretical frameworks [19,23-25], whereas basic and intermediate designs operated predominantly without explicit theoretical foundations [20-22] (Table 4).

Effect strength was evaluated through standardized criteria: strong, >15% improvement from baseline OR Cohen d >0.8 OR η^2>0.25 with P<.05; moderate, 5% - 15% improvement OR Cohen d 0.5 - 0.8 OR η^2 0.10 - 0.25 with P<.05; weak, <5% improvement OR Cohen d <0.5 OR η^2 <0.10 with P<.05, NS, nonsignificant.

XSL•FO
RenderX

Table . Systematic categorization of interventions.

Authors (year)	Complexity	Tech integration	Primary outcome	Effect direction	Statistical significance	Theoretical framework
Randriambelonoro et al [14] 2023	Intermediate	Wearable-centric	Daily steps	NS[a]	*P*=.25	None
Martinho et al [19] 2023	Advanced	Standalone	Daily steps	Moderate	Cohen *d*=0.65 *P*=.03	Circumplex model
Santos et al [20] 2018	Basic	Standalone	% Steps	Moderate	Descriptive statistics only	None
Santos et al [21] 2019	Basic	Standalone	% Steps	Moderate	η^2=0.19, *P*=.01	None
Santos et al [22] 2021	Intermediate	Standalone	% Steps	Moderate	η^2=0.31, *P*=.04	None
Fanaroff et al [23] 2024	Advanced	Hybrid	Daily steps and MVPA[b]	Strong	*P*=.08[c] *P*=.03[d]	Prospect theory
Greysen et al [25] 2024	Advanced	Hybrid	Daily steps and MVPA	Strong	*P*<.001[c] *P*=.08[d]	Prospect theory
Kawaguchi et al [24] 2024	Intermediate	Hybrid	Weekly steps and activity frequency	Weak	*P*=.04 (steps NS)	Behavior change

[a]NS: nonsignificant.

[b]MVPA: moderate-to-vigorous physical activity.

[c]*P* value of daily steps.

[d]*P* value of MVPA.

Discussion

Principal Findings

The present review included only 8 full-text articles that met stringent inclusion criteria, reflecting both methodological rigor and domain-specific constraints. Strict adherence to gamification standards, mandatory mHealth components, and a focused eligibility criterion of adults aged 60 years and above resulted in the exclusion of a considerable proportion of initially identified records. This selectivity contrasts sharply with the broad application of gamification in general healthcare, which has largely targeted younger populations [11]. Nonetheless, this condensed evidence base captures a developing nexus among gerontechnology, behavioral engagement, and mHealth—an area that remains understudied yet highly relevant for physical activity promotion in older adults. Despite these constraints, the main findings indicate that mHealth-based gamified interventions offer meaningful benefits in improving physical activity levels within this demographic.

A larger sample size and longer intervention duration would enhance the generalizability of the data and provide more opportunities to determine usage patterns and preferences. However, most of the reviewed studies employed shorter intervention durations and did not include follow-up assessments, involving only a small number of participants. This phenomenon is also observed in other areas of gamified interventions for older adults. For example, Savulich et al [29] used a sample size of 21 participants in their intervention group for cognitive training using iPads for older adults diagnosed with amnestic mild cognitive impairment, with an intervention duration of 4 weeks. Costa et al [30] used a sample size of 20 participants in their intervention group for mental health interventions for community-dwelling older adults through a gamified learning platform, with an intervention duration of 6 weeks. Despite the increased opportunities to access new technologies in recent years, the application of these tools among the older adult population still faces significant challenges.

Choosing game elements that cater to the needs of older adults or developing corresponding gamified applications can significantly enhance the effectiveness, practicality, and sustainability of interventions, thereby increasing user participation and reducing sample attrition rates during follow-up processes [31]. The studies included in this review generally adopt goal-setting (eg, daily step counts) and instant rewards (eg, points and badges) as core elements. These elements are widely used due to their strong alignment with the core mechanisms of behavior change [32], while also addressing the research need for quick verification of the immediate effects of gamification through short-term interventions. This approach aligns with the conclusions of 2 previous systematic reviews, which did not specify the research subjects [33,34].

However, the adoption of complex gamification elements, such as virtual identities and social collaboration, requires older adults to construct abstract self-representations or coordinate rules for multilateral interactions. This may lead to cognitive load or operational obstacles for some older adults [29], resulting in fewer studies designed with such elements. Additionally, we note that the included studies all employed interventions with 2 or more combined elements, suggesting that combinations of multiple game elements may be more effective than single

XSL•FO
RenderX

elements. However, due to the lack of detailed information in the initial studies, no definitive conclusions can yet be drawn about the optimal number or composition of game elements in these interventions.

One point that needs clarification is that the most frequently used gamification elements do not necessarily indicate the optimal effectiveness of those elements on older adults or physical activity. As Dugas et al [35] pointed out, personality differences can significantly affect how certain gamification characteristics influence older adults' use of mobile health technologies. When gamification mechanisms are unsuitable, older adults often fail to understand and recognize the relevant benefits of these services. Therefore, we must approach the results of this systematic review with caution, as it could lead to cognitive biases. In the study by Kostopoulos et al [36], the authors established goal-setting and achievement features that allow users to receive different rewards after completing various goals and tasks. The purpose was to further explain how these goals and tasks are generated for each user, which may help in understanding the different motivations and preferences that drive users to perform various tasks. This is an innovative approach, as it provides feedback on user performance while validating the effects of gamification interventions.

We have also found that Greysen et al [25] systematically included primary caregivers (support partners) as a core component of the intervention. And although other studies did not independently analyze or evaluate the involvement of nursing staff, most of them indirectly addressed similar concepts through social functions or external support mechanisms. This suggests that incorporating nursing staff into the network could be a valuable development. Including family members, health care professionals, and social workers in mobile health gamification interventions may enhance older adults' perceptions of social and emotional support, thereby increasing their acceptance of these interventions [37]. Additionally, from a technical perspective, this model can improve older adults' understanding of the intervention content and their sense of security in using the interventions [38]. This approach is particularly beneficial for older adults who are interested in gamification but lack independent operational skills, as well as for those who prefer direct care from health care professionals, providing an effective pathway for the broader application of gamification.

Technological platforms that accompany gamification are another important factor attracting the interest of older adults. Currently, validated technological features used among the older adult population include self-management systems [9], physical robots [10], consoles [11], portable devices, and wearable technology [12]. Although studies by Kitakoshi et al [38] and Mocanu et al [39] have confirmed that older adults enjoy using systems and interacting with robots, mobile health offers more diverse gamification opportunities in real-world applications. This aligns with the viewpoint of Steinert et al [40], who pointed out that older adults are more motivated by tangible information provided by activity monitors, sensors, and smartphones, rather than more abstract game feedback. This is also the primary reason why our review is limited to gamification intervention research supported by mobile technology.

In this systematic review, the mHealth technologies used to provide physical activity gamification interventions include 1 activity monitor (1/8), 4 mobile apps (4/8), and 3 combinations of both (3/8). This distribution of technological application patterns reflects a shift in the current research field, from single-device use to multimodal and collaborative intervention plans. Tablet devices offer significant advantages in screen size and clarity, and previous research [30] has confirmed their ability to improve specific cognitive functions in older adults. However, the physical characteristics of the devices, such as weight, button design, and keyboard operation, may affect the user experience of older adults participating in activities. As a result, only 1 study in this review mentioned tablet devices [23].

The theoretical framework also warrants discussion, as it enables a deeper analysis of the behavioral mechanisms behind older adults' participation in physical activities, accurately addressing their physiological, psychological, and social needs. However, among the reviewed studies, only 4 [19,23-25] explicitly adopted a single theoretical framework or model to design gamified interventions for physical activities targeting older adults. The classification framework revealed that these theory-oriented designs demonstrated more effective outcomes while accommodating greater gamification complexity, suggesting that theoretically integrated gamification may play a positive role in reducing engagement burden among older adults.

At least 2 of these studies [23,25] utilized behavioral economics, incorporating design principles such as loss aversion, precommitment, and the fresh start effect. The significance of this theoretical application lies in viewing participation in physical activities as an investment in future health, guiding older adults' behavioral choices through subtle environmental adjustments rather than rigid directives. In another study [19], the circumplex model of affect was translated into operational gamification solutions, offering new insights for the design of digital therapeutics in geriatric health.

Self-Determination Theory, a comprehensive motivation theory, has become a key framework for health behavior interventions, but only 1 study [24] claims to have used it, differing from the 2 systematic reviews mentioned earlier [33,34]. Other commonly considered behavior change theories include the Health Belief Model [41], Theory of Planned Behavior [42], Transtheoretical Model [43], and Habit Formation Theory [44], but these theories appear to be rarely applied in gamified interventions for older adults, even in nonphysical activity areas. The reasons for this may partly stem from the relatively novel concept of introducing gamified design to the older adult population, where traditional theories may not align well. Alternatively, it could be that the behavior patterns of older adults are complex and diverse, making the application of theory more challenging. Therefore, a more systematic and appropriate gamification theory system for the older adult population could be developed in the future through qualitative exploration, theoretical integration, and other methods.

Finally, a study published by Bellotti et al [45] in the field of human-computer interaction suggested that game-based interventions could serve as a strategy for patient self-assessment, enriching the typical content of assessments

XSL•FO
RenderX

for older adults. Although no authors in the reviewed studies directly made such a claim, this idea may also apply to gamification for older adults. By continuously tracking the activity performance of older adults, research can gather multidimensional data, such as task completion rates, execution times, and failure counts. This information, when combined with commonly used assessment tools for older adults, such as the Activity of Daily Living scale [46] and the Morse fall risk assessment scale [47], could provide a more comprehensive reflection of functionality and psychological state, thereby guiding clinical decision-making and personalized intervention strategies. While this is a reasonable inference, further investigation is needed. Additionally, it is important to consider data privacy and the risk of false positives to avoid excessive monitoring, which could lead to psychological stress or self-report biases due to the Hawthorne effect.

Limitations

Although this study systematically reviews the effectiveness of mHealth gamification interventions on physical activity in older adults, several limitations should be noted. First, the included studies are relatively few (n=8), with generally small sample sizes, potentially limiting the generalizability and reliability of the findings. Second, intervention durations varied significantly (1 wk to 12 mo), and most studies lacked follow-up periods, complicating the assessment of long-term sustainability. Additionally, considerable heterogeneity existed across studies, including variations in gamification elements, technological platforms, theoretical frameworks, and outcome measurement methods, precluding meta-analysis and unified conclusions. Another limitation is the incomplete reporting of baseline participant characteristics (eg, digital literacy, health status), potentially biasing intervention effect interpretations. Finally, as most studies focused on high-income countries (eg, the United States, Japan, Switzerland), the findings may lack generalizability to older adults in other regions or cultural contexts. Furthermore, due to incomplete reporting of minimum participant ages, this review inferred population homogeneity using a functionally oriented aging definition and statistical estimation, which may overestimate or misrepresent intervention effectiveness.

Conclusions

This systematic review suggests that mHealth-based gamified interventions have potential benefits in increasing the physical activity levels of older adults, particularly by enhancing daily step counts and the duration of MVPA. Common gamification elements, such as goal-setting and immediate rewards, effectively stimulate the participation motivation of older adults, while the combination of mobile apps and wearable devices further enhances the flexibility and accessibility of these interventions. However, existing evidence still has limitations, including a limited number of studies, small sample sizes, and a lack of long-term follow-up data. Future studies should aim to expand sample sizes, extend intervention durations, and include more diverse older adult populations to validate the universality and sustainability of the intervention effects. Additionally, there is a need to further explore the optimal combination of gamification elements, the integration of theoretical frameworks, and the potential impact of caregiver network participation on the effectiveness of these interventions. Overall, mHealth gamified interventions offer an innovative and feasible strategy for promoting physical activity among older adults, but their design and implementation must carefully consider the specific needs and varying abilities of older adults to ensure the practicality and acceptance of these interventions.

Data Availability

All data generated or analyzed during this study are included in this published article (and its supplementary information files).

Authors' Contributions

Conceptualization: FJ, LC
Investigation: FJ, LC, ML
Resources: FJ, LC, ML, WZ
Visualization: FJ, HY, LC, WZ
Writing – original draft: LC, ML
Writing – review & editing: HY, LC

Conflicts of Interest

None declared.

Multimedia Appendix 1
Full search strategy.
[DOCX File, 12 KB - aging_v8i1e78686_app1.docx]

Multimedia Appendix 2
References screened by title and abstract.
[XLSX File, 519 KB - aging_v8i1e78686_app2.xlsx]

XSL•FO
RenderX

Multimedia Appendix 3
References excluded after reading the full text.
[DOCX File, 19 KB - aging_v8i1e78686_app3.docx]

Multimedia Appendix 4
Game elements observed in the studies.
[PNG File, 6 KB - aging_v8i1e78686_app4.png]

Checklist 1
PRISMA checklist.
[PDF File, 159 KB - aging_v8i1e78686_app5.pdf]

References

1. Ageing and health. World Health Organisation. 2022. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2025-05-26]
2. Beard JR, Officer A, de Carvalho IA, et al. The World report on ageing and health: a policy framework for healthy ageing. The Lancet 2016 May;387(10033):2145-2154. [doi: 10.1016/S0140-6736(15)00516-4]
3. Bull FC, Al-Ansari SS, Biddle S, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med 2020 Dec;54(24):1451-1462. [doi: 10.1136/bjsports-2020-102955] [Medline: 33239350]
4. Dunlop DD, Song J, Arnston EK, et al. Sedentary time in US older adults associated with disability in activities of daily living independent of physical activity. J Phys Act Health 2015 Jan;12(1):93-101. [doi: 10.1123/jpah.2013-0311] [Medline: 24510000]
5. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining "Gamification". Presented at: MindTrek '11: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments; Sep 28-30, 2011; Tampere Finland p. 9-15. [doi: 10.1145/2181037.2181040]
6. Gavelin HM, Dong C, Minkov R, et al. Combined physical and cognitive training for older adults with and without cognitive impairment: a systematic review and network meta-analysis of randomized controlled trials. Ageing Res Rev 2021 Mar;66:101232. [doi: 10.1016/j.arr.2020.101232] [Medline: 33249177]
7. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. JMIR Serious Games 2013 Dec 12;1(1):e3. [doi: 10.2196/games.3139] [Medline: 25658754]
8. Alahäivälä T, Oinas-Kukkonen H. Understanding persuasion contexts in health gamification: a systematic analysis of gamified health behavior change support systems literature. Int J Med Inform 2016 Dec;96:62-70. [doi: 10.1016/j.ijmedinf.2016.02.006] [Medline: 26944611]
9. Free C, Phillips G, Galli L, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013;10(1):e1001362. [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
10. Seko Y, Kidd S, Wiljer D, McKenzie K. Youth mental health interventions via mobile phones: a scoping review. Cyberpsychol Behav Soc Netw 2014 Sep;17(9):591-602. [doi: 10.1089/cyber.2014.0078] [Medline: 25007383]
11. Reiners F, Sturm J, Bouw LJW, Wouters EJM. Sociodemographic factors influencing the use of eHealth in people with chronic diseases. Int J Environ Res Public Health 2019 Feb 21;16(4):645. [doi: 10.3390/ijerph16040645] [Medline: 30795623]
12. Walker DM, Hefner JL, Fareed N, Huerta TR, McAlearney AS. Exploring the digital divide: age and race disparities in use of an inpatient portal. Telemed J E Health 2020 May;26(5):603-613. [doi: 10.1089/tmj.2019.0065] [Medline: 31313977]
13. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
14. Randriambelonoro M, Perrin Franck C, Herrmann F, et al. Gamified physical rehabilitation for older adults with musculoskeletal issues: pilot noninferiority randomized clinical trial. JMIR Rehabil Assist Technol 2023 Mar 6;10:e39543. [doi: 10.2196/39543] [Medline: 36877563]
15. Kleschnitzki JM, Grossmann I, Beyer R, Beyer L. Modification in the motor skills of seniors in care homes using serious games and the impact of COVID-19: field study. JMIR Serious Games 2022 May 10;10(2):e36768. [doi: 10.2196/36768] [Medline: 35536610]
16. Cuevas-Lara C, Sáez de Asteasu ML, Ramírez-Vélez R, et al. Effects of game-based interventions on functional capacity in acutely hospitalised older adults: results of an open-label non-randomised clinical trial. Age Ageing 2022 Jan 6;51(1):afab247. [doi: 10.1093/ageing/afab247] [Medline: 35077558]
17. Barker TH, Stone JC, Sears K, et al. The revised JBI critical appraisal tool for the assessment of risk of bias for randomized controlled trials. JBI Evid Synth 2023 Mar 1;21(3):494-506. [doi: 10.11124/JBIES-22-00430] [Medline: 36727247]
18. Barker TH, Habibi N, Aromataris E, et al. The revised JBI critical appraisal tool for the assessment of risk of bias for quasi-experimental studies. JBI Evid Synth 2024 Mar 1;22(3):378-388. [doi: 10.11124/JBIES-23-00268] [Medline: 38287725]

XSL•FO
RenderX

19. Martinho D, Crista V, Carneiro J, Matsui K, Corchado JM, Marreiros G. Effects of a gamified agent-based system for personalized elderly care: pilot usability study. JMIR Serious Games 2023 Nov 23;11:e48063. [doi: 10.2196/48063] [Medline: 37995116]
20. Santos LH, Okamoto K, Hiragi S, et al. Pervasive game design to evaluate social interaction effects on levels of physical activity among older adults. J Rehabil Assist Technol Eng 2019;6:2055668319844443. [doi: 10.1177/2055668319844443] [Medline: 31285836]
21. Santos L, Okamoto K, Funghetto SS, et al. Effects of social interaction mechanics in pervasive games on the physical activity levels of older adults: quasi-experimental study. JMIR Serious Games 2019;7(3):e13962. [doi: 10.2196/13962]
22. Santos L, Okamoto K, Otsuki R, et al. Promoting physical activity in Japanese older adults using a social pervasive game: randomized controlled trial. JMIR Serious Games 2021 Jan 6;9(1):e16458. [doi: 10.2196/16458] [Medline: 33404507]
23. Fanaroff AC, Patel MS, Chokshi N, et al. Effect of gamification, financial incentives, or both to increase physical activity among patients at high risk of cardiovascular events: the BE ACTIVE randomized controlled trial. Circulation 2024 May 21;149(21):1639-1649. [doi: 10.1161/CIRCULATIONAHA.124.069531] [Medline: 38583084]
24. Kawaguchi K, Nakagomi A, Ide K, Kondo K. Effects of a mobile app to promote social participation on older adults: randomized controlled trial. J Med Internet Res 2024 Sep 30;26:e64196. [doi: 10.2196/64196] [Medline: 39348180]
25. Greysen SR, Oon AL, Harkins K, et al. Effect of gamification with a support partner to increase physical activity in older adults at risk for Alzheimer's disease: the STEP 4Life randomized clinical trial. Alzheimer's & Dementia 2024 Aug;20(8):5450-5459. [doi: 10.1002/alz.14058]
26. Russell JA. A circumplex model of affect. J Pers Soc Psychol 1980;39(6):1161-1178. [doi: 10.1037/h0077714] [Medline: 7431205]
27. Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. Econometrica 1979 Mar;47(2):263. [doi: 10.2307/1914185]
28. Kwasnicka D, Dombrowski SU, White M, Sniehotta F. Theoretical explanations for maintenance of behaviour change: a systematic review of behaviour theories. Health Psychol Rev 2016 Sep;10(3):277-296. [doi: 10.1080/17437199.2016.1151372] [Medline: 26854092]
29. Savulich G, Piercy T, Fox C, et al. Cognitive training using a novel memory game on an iPad in patients with amnestic mild cognitive impairment (aMCI). Int J Neuropsychopharmacol 2017 Aug 1;20(8):624-633. [doi: 10.1093/ijnp/pyx040] [Medline: 28898959]
30. Costa LV, Veloso AI, Loizou M, Arnab S. Games for active ageing, well-being and quality of life: a pilot study. Behav Inf Technol 2018 Aug 3;37(8):842-854. [doi: 10.1080/0144929X.2018.1485744]
31. Devos P, Min Jou A, De Waele G, Petrovic M. Design for personalized mobile health applications for enhanced older people participation. Eur Geriatr Med 2015 Dec;6(6):593-597. [doi: 10.1016/j.eurger.2015.10.004]
32. Hicks K, Gerling K, Richardson G, Pike T, Burman O, Dickinson P. Understanding the effects of gamification and juiciness on players. Presented at: 2019 IEEE Conference on Games (CoG); Aug 20-23, 2019; London, United Kingdom p. 1-8. [doi: 10.1109/CIG.2019.8848105]
33. Vermeir JF, White MJ, Johnson D, Crombez G, Van Ryckeghem DML. The effects of gamification on computerized cognitive training: systematic review and meta-analysis. JMIR Serious Games 2020 Aug 10;8(3):e18644. [doi: 10.2196/18644] [Medline: 32773374]
34. Xu L, Shi H, Shen M, et al. The effects of mHealth-based gamification interventions on participation in physical activity: systematic review. JMIR Mhealth Uhealth 2022 Feb 3;10(2):e27794. [doi: 10.2196/27794] [Medline: 35113034]
35. Dugas M, Crowley K, Gao GG, et al. Individual differences in regulatory mode moderate the effectiveness of a pilot mHealth trial for diabetes management among older veterans. PLoS ONE 2018;13(3):e0192807. [doi: 10.1371/journal.pone.0192807] [Medline: 29513683]
36. Kostopoulos P, Kyritsis AI, Ricard V, Deriaz M, Konstantas D. Enhance daily live and health of elderly people. Procedia Comput Sci 2018;130:967-972. [doi: 10.1016/j.procs.2018.04.097]
37. Merkel S, Kucharski A. Participatory design in gerontechnology: a systematic literature review. Gerontologist 2019 Jan 9;59(1):e16-e25. [doi: 10.1093/geront/gny034] [Medline: 29788319]
38. Kitakoshi D, Okano T, Suzuki M. An empirical study on evaluating basic characteristics and adaptability to users of a preventive care system with learning communication robots. Soft Comput 2017 Jan;21(2):331-351. [doi: 10.1007/s00500-015-1631-7]
39. Mocanu I, Marian C, Rusu L, Arba R. A kinect based adaptive exergame. Presented at: 2016 IEEE 12th International Conference on Intelligent Computer Communication and Processing (ICCP); Sep 8-10, 2016; Cluj-Napoca, Romania p. 117-124. [doi: 10.1109/ICCP.2016.7737132]
40. Steinert A, Buchem I, Merceron A, Kreutel J, Haesner M. A wearable-enhanced fitness program for older adults, combining fitness trackers and gamification elements: the pilot study fMOOC@Home. Sport Sci Health 2018 Aug;14(2):275-282. [doi: 10.1007/s11332-017-0424-z]
41. Rosenstock IM. Historical origins of the health belief model. Health Educ Monogr 1974 Dec;2(4):328-335. [doi: 10.1177/109019817400200403]

XSL•FO
RenderX

42. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
43. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997 Sep;12(1):38-48. [doi: 10.4278/0890-1171-12.1.38]
44. Messinis G. Habit formation and the theory of addiction. J Econ Surv 1999 Sep;13(4):417-442. [doi: 10.1111/1467-6419.00089]
45. Bellotti F, Kapralos B, Lee K, Moreno-Ger P, Berta R. Assessment in and of serious games: an overview. Adv Hum-Comput Interact 2013;2013:1-11. [doi: 10.1155/2013/136864]
46. Mahoney FI, Barthel DW. Functional evaluation: the Barthel index. Md State Med J 1965;14:61-65. [Medline: 14258950]
47. Chu JJ, Chen XJ, Shen SS, et al. A poor performance in comprehensive geriatric assessment is associated with increased fall risk in elders with hypertension: a cross-sectional study. J Geriatr Cardiol 2015 Mar;12(2):113-118. [doi: 10.11909/j.issn.1671-5411.2015.02.006] [Medline: 25870613]

Abbreviations

mHealth: mobile health
MVPA: moderate-to-vigorous physical activity
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial

Edited by Y Du; submitted 10.06.25; peer-reviewed by C Gissel, ME Heidari; revised version received 03.09.25; accepted 13.09.25; published 31.10.25.

Please cite as:
Chen L, Jang F, Li M, Zong W, Yu H
Effectiveness of mHealth-Based Gamified Interventions on Physical Activity in Older Adults: Systematic Review
JMIR Aging 2025;8:e78686
URL: https://aging.jmir.org/2025/1/e78686
doi:10.2196/78686

© Lin Chen, Fenglin Jang, Min Li, Wei Zong, Huiqin Yu. Originally published in JMIR Aging (https://aging.jmir.org), 31.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Examining Technology Perspectives of Older Adults With Mild Cognitive Impairment: Scoping Review

Snezna Bizilj Schmidt[1], BSc (Hons); Stephen Isbel[2], HScD; Blooma John[1], PhD; Ramanathan Subramanian[1], PhD; Nathan Martin D'Cunha[3], PhD

[1]Faculty of Science and Technology, Centre for Intelligent Computing and Systems (CICS), School of Information Technology and Systems, University of Canberra, 11 Kirinari Street, Bruce, ACT, Australia

[2]Faculty of Health, Discipline of Occupational Therapy and Centre for Ageing Research and Translation, University of Canberra, Bruce, Australia

[3]Faculty of Health, Centre for Ageing Research and Translation, University of Canberra, Bruce, ACT, Australia

Corresponding Author:
Snezna Bizilj Schmidt, BSc (Hons)
Faculty of Science and Technology, Centre for Intelligent Computing and Systems (CICS), School of Information Technology and Systems, University of Canberra, 11 Kirinari Street, Bruce, ACT, Australia

Abstract

Background: Mild cognitive impairment (MCI) affects up to 20% of people older than the age of 65 years. The global incidence of MCI is increasing, and technology is being explored for early intervention. Theories of technology adoption predict that useful and easy-to-use solutions will have higher rates of adoption; however, these models do not specifically consider older adults with cognitive impairments or the unique human-computer interaction challenges posed by MCI. There are gaps in understanding the combined impacts of aging and cognitive impairment on factors affecting technology adoption for older adults with MCI, and it is not clear how MCI impacts human-computer interaction and device and interaction modality preferences in this population.

Objective: This study aimed to collate perspectives from older adults with MCI about technology solutions proposed for them, to understand whether solutions are perceived as useful, easy to use, and what changes are suggested. It also identifies which devices and interaction modalities are preferred, and other factors that may affect usage and adoption.

Methods: This scoping review was completed according to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines. A consistent search was performed across 9 electronic databases (ACM Digital Library, EBSCOhost CINAHL Plus with Full Text, EBSCOhost Computers and Applied Sciences Complete, Google Scholar, JMIR Publications, IEEE Xplore, EBSCOhost MEDLINE, Scopus, and Web of Science Core Collection) for studies published between January 1, 2014, and May 1, 2024. Extracted data were analyzed using inductive thematic analysis.

Results: We identified 4271 studies, and after the removal of duplicates and screening, 83 studies were included for data extraction. Inductive thematic analysis of feedback from older adults with MCI about technology solutions proposed for them identified five themes: (1) purpose and need, (2) solution design and ease of use, (3) self-impression, (4) lifestyle, and (5) interaction modality. Solutions were perceived as useful, even though gaps in functional support exist; however, they were not perceived as entirely easy to use due to issues related to usability and user experience. Devices that are lightweight, portable, familiar, and have large screens are preferred, as is multimodal interaction—particularly speech, visual or text, and touch.

Conclusions: Using technology can create feelings that positively or negatively affect a user's comfort, confidence, and overall well-being. Older adults with MCI value independence and autonomy, and solution designs should support these. Usefulness, ease of use, security, privacy, cost, physical comfort, and convenience are important considerations for technology use. Reliable technology creates trust, confidence, and feelings of empowerment. This review recommends future work to (1) improve usability and user experience, (2) enhance personalization, (3) better understand interaction preferences and effectiveness, (4) enable options for multimodal interaction, and (5) more seamlessly integrate solutions into users' lifestyles.

(JMIR Aging 2025;8:e78229) doi:10.2196/78229

KEYWORDS

mild cognitive impairment; MCI; technology adoption; multimodal; interaction; preference; usability; user experience; UX

XSL•FO
RenderX

Introduction

Background

Mild cognitive impairment (MCI) affects a person's memory or how they think, feel, or behave. It is not a normal part of aging, but up to 20% of adults older than the age of 65 years may be affected, and up to 15% of those may progress to dementia [1,2]. Historically, MCI was considered a decline in a person's memory, but the criteria and definition of MCI have expanded with more research to include other cognitive domain deficits and affective attributes such as language, executive functioning, or visuospatial ability [3-5]. A memory decline is only labeled single-domain amnestic MCI. A decline in 1 nonmemory cognitive domain is labeled single-domain nonamnestic MCI, and a decline in memory and one or more other cognitive domains is labeled multiple-domain amnestic MCI. People with MCI may forget things more frequently, lose items more often, struggle to remember words or have problems with language, have trouble making decisions or following instructions, lose their train of thought, have new difficulties regulating emotions, behave more impulsively, have trouble with visual perception, have trouble sleeping, experience motor coordination issues, or be less able to follow daily routines [2]. Given that visual, perceptual, cognitive, and behavioral functions are needed for human-computer interaction (HCI), the limitations associated with MCI create challenges for HCI, such as difficulties remembering process steps or commands, understanding instructions, interpreting feedback, correcting errors, or maintaining focus and attention. Changes to emotions, feelings, or mood may also make MCI more difficult and thereby create frustration, increase fear of making mistakes, and reduce self-confidence, all of which have been shown to contribute to factors influencing technology adoption (TA). These challenges can be managed through a considered design of technology solutions.

The global prevalence of MCI is increasing and placing additional stress on both health and aged care services [6]. The number of older Australians with MCI is estimated to grow from 884,000 in 2022 to 2.03 million in 2070 [7,8]. The Australian Royal Commission into Aged Care Quality and Safety [9] noted that aged care services are already under pressure and recommended exploring technology options to address the growing needs of older Australians. Australia's National Dementia Action Plan 2024 - 2034 [10] recommended that action be taken to (1) implement evidence-based early interventions to reduce cognitive decline, and (2) provide support and resources for people with MCI.

Research has shown that technology provides opportunities to increase health care service productivity [11], but only if solutions are adopted by the target users. Aging may result in slower cognitive speed, poorer memory, reduced concentration, and negatively impact vision, hearing, speech, dexterity, mobility, and learning abilities. These limitations have been found to discourage the use of technology [12], and studies have shown that the use of digital health solutions is lower among older adults than younger adults [13]. In addition, as people experience MCI uniquely and with different symptoms, older adults with MCI will have different and diverse needs for technology support.

Literature Review

TA has been studied from several perspectives. Early models, such as the technology acceptance model (TAM) [14], TAM2 [15], and TAM3 [16], take a behavioral perspective and consider TA in an organizational context. They describe TA as being dependent on a user's perception of the technology's usefulness, ease of use, and the determinants of these. The unified theory of acceptance and use of technology (UTAUT) [17] considers TA from a combined behavioral, psychological, and information management perspective, and UTAUT2 [18] extends UTAUT to consider TA outside of business settings, such as at home, and shows that the context of use has an impact on the relevance and contribution of determinants. However, these models were not developed based on data about older adults with cognitive impairment or the assessment of health care technologies. More recent models, such as the senior technology acceptance model (STAM) [19] and the model for the adoption of technology by older adults (MATOA) [20], focus on older users and the effects of physical and cognitive aging on TA. They note that ease of use becomes more important as users age, and MATOA identifies self-management (the extent to which a person can remain independent and maintain control of their life, emotions, and role) as a determinant for intention to use. The lifespan theory of control [21] describes the importance of control and states that people will attempt to retain control of how they behave and interact with their external environment for as long as possible because control contributes to a sense of well-being. It is therefore relevant to consider user empowerment as a valuable design feature in technology solutions designed for older adults, and this is supported by a literature review of older adults' intention to use technologies [22]. STAM and MATOA do not specifically consider TA by older adults with cognitive impairments or technology designed for health care or assistance with activities of daily living. The health care technology acceptance model (H-TAM) [23] describes acceptance of health care technologies by older adults with hypertension. It includes many factors that appear in TAM-based TA models and proposes compatibility (how well the technology integrates into a person's everyday life) as a determinant of facilitating conditions.

A study of people's intention to use service robots [24] suggests that TA models may need to be revisited to incorporate the influence of technology innovations such as smart technologies and robots, which may be adaptive and exhibit human-like features. These advances have enabled new uses and new modes of interaction, and users can now choose their device and interaction mode, which is empowering. However, it is not clear how this impacts TA for older users with cognitive impairments, in particular, perceptions of usefulness and ease of use, which are shown to be determinants of TA. Researchers have investigated the effectiveness of unimodal and multimodal interaction for older users and found that modality affects use [12,25-29]; however, these studies did not include older adults with cognitive impairments. Subtle speech differences can be observed in people with MCI [30], and literature reviews have reported on the use of voice-activated technology solutions by

XSL•FO
RenderX

older adults [27,31,32]. However, these did not always include older adults with cognitive impairment as participants.

There are gaps in understanding the combined impacts of aging and cognitive impairment on factors affecting TA for older adults with MCI. It is also unclear how MCI impacts HCI and device and interaction mode preferences. This scoping literature review examines published studies about technology solutions developed for older adults with MCI (amnestic or nonamnestic, and single- or multi-domain) to answer the following research questions:

- Research question 1 (RQ1): What are the opinions of older people with MCI about the technology solutions that have been proposed for them? (A) Are they useful? (B) Are they easy to use, or are changes and improvements suggested?
- RQ2: What feedback do older people with MCI provide about usage? (A) Preferred devices? (B) Interaction modality and ways of use?

To our knowledge, this is the first study that collates opinions and feedback from older adults with MCI about technology solutions proposed for them.

Methods

Overview

This scoping review design was guided by previous literature [33,34] and is reported according to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines [35]. It was preregistered on the Open Science Framework Storage (United States) [36].

Eligibility Criteria

An electronic database search was performed to identify studies published in English between January 1, 2014, and May 1, 2024. A 10-year period is considered sufficient given the fast evolution of technology and the potential for change in people's opinions.

Information Sources

The search was performed consistently across nine electronic databases (ACM Digital Library, EBSCOhost CINAHL Plus with Full Text, EBSCOhost Computers and Applied Sciences Complete, Google Scholar, JMIR Publications, IEEE Xplore, EBSCOhost MEDLINE, Scopus, and Web of Science Core Collection).

Search Strategy

The search strategy was developed with the assistance of a qualified librarian. The following search query was used in all databases:

> *((MCI OR "mild cognitive impairment" OR "mild cognitive disability" OR "mild neurocognitive")AND(technolog* OR ICT OR smart OR wearable* OR computer OR PC OR laptop OR tablet OR "touch screen" OR "touch-screen" OR "mobile phone" OR "mobile device" OR "personal device" OR robot OR reality OR VR OR "assistive technolog*" OR "embodied conversational agent" OR ECA OR multimedia OR "multi-media")AND(modalit* OR mode OR channel OR interact* OR engage* OR touch OR type OR voice OR gesture OR use OR usage)AND(preference* OR experience* OR perception* OR attitude* OR feeling OR practices OR "technology acceptance" OR qualitative OR "design science" OR "innovation resistance theory"))*

Selection Criteria

Studies that satisfied the search criteria were imported into the referencing software program EndNote (Clarivate Analytics) and then uploaded to Covidence (Veritas Health Innovation) to remove duplicates and for screening [37].

Three reviewers and the primary author screened the titles and abstracts of each study and excluded any that did not meet the following inclusion criteria:

- Participants have MCI and were living independently in the community (not in residential aged care)
- Participants used or evaluated the technology solutions or devices
- Participants provided feedback about a technology idea, paper prototype, or abstract concept
- The study assessed one or more of the participants' experiences, feelings, or attitudes regarding the technology, and preferences for use

Papers were excluded if:

- Participants lived in residential care or a nursing home
- The paper did not report on participant feedback.
- The paper reported only on the efficacy of the solution
- If the study included multiple participant cohorts, the feedback was not identifiable or reported by cohort

Where it was unclear if a study met the inclusion criteria, or there was a disagreement among reviewers, the study was carried forward into the full-text review.

The primary author independently reviewed all full-text studies extracted from the previous stage. The remaining authors shared the role of second reviewer. Full-text papers were included if they reported the following:

- Research about participants' use of technology and their experiences, feelings, attitudes, opinions, or preferences
- If multiple participant cohorts were included, the data were identifiable and reported by cohort
- Participants were assessed for MCI
- Participants had no comorbidities that cause neurological issues or affect cognition

Any disagreements were resolved by discussion among the authors.

Data Analysis Process

Data were extracted into a structured Microsoft Excel spreadsheet by the first author and evaluated by 2 other authors. Data were extracted about (1) study details (such as author, location of study, year published, journal name, study aim, study type, technology being evaluated, and participant details) and

XSL•FO
RenderX

(2) participant feedback. Participant feedback was recorded as any reference to (1) a participant's opinion about the technology (usefulness, ease of use, usage, and context of use), (2) how the solution made them feel, (3) preferences for device or interaction, (4) issues or challenges with use, and (5) recommendations for changes. Statements were captured as written, relating to the type of technology and solution.

All extracted data were reviewed for familiarization and to start the process of identifying similarities, relationships, and trends. Participant feedback was organized and analyzed using inductive, thematic analysis as described by Thomas and Harden [38], to ensure transparency and traceability from the data to the outcomes. This process has been used effectively to complete other thematic analyses of qualitative data that describe people's experiences, opinions, and feelings.

Inductive thematic analysis of participant feedback was guided by the 2 research questions and completed in 4 steps:

1. Keyword allocation: each feedback item was assigned one or more keyword phrases that succinctly described its meaning. The first author used existing keywords or created new ones as needed.
2. Consistency check: review of all feedback items with the same keyword to ensure consistent allocation.
3. Category grouping: related keyword phrases were grouped into higher-level categories, which were then named.
4. Theme creation: the categories were reviewed, and 5 abstract themes were created to provide meaning to the feedback and answer the research questions.

During the above 4-step process, there was a continuous feedback cycle with coauthors.

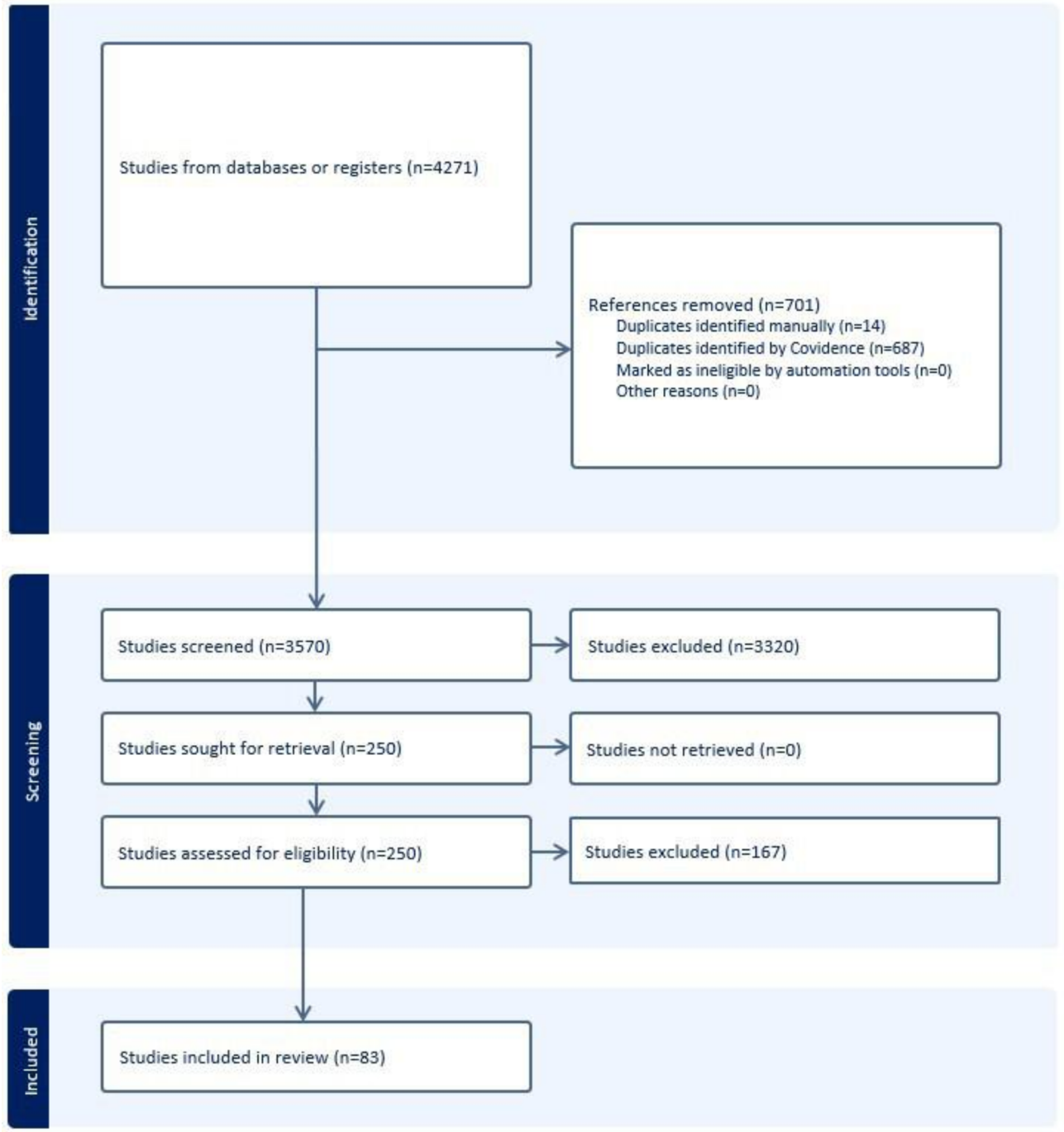
Results

Overview

A total of 4271 studies were imported for screening, 687 duplicates were identified and removed by Covidence, and 14 duplicates were identified and removed manually. A total of 3570 studies were screened by title and abstract, and 3320 studies were subsequently excluded. The full text of 250 studies was assessed for eligibility, and 167 studies were excluded, leaving 83 studies for data extraction. Figure 1 presents the flow of studies through the PRISMA-ScR process, and Table 1 provides a summary of the studies and extracted data, sorted alphabetically by the author.

XSL•FO
RenderX

Figure 1. Flow of studies through the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) process.



XSL•FO
RenderX

Table . Summary of studies, sorted alphabetically by author.

Study details	Study type	Type of technology	Purpose of the solution	Participant details
Afifi et al [39]	Quantitative	VR[a] headset	Social interaction, messaging, chat, or information provision	Female: 18; male: 3; location: community center; 9 participants had MCI[b]
Bartels et al [40]	Mixed methods	Mobile phone	Lifestyle (ADL[c]) support, social interaction, or companionship	Female: 5; male: 16; location: home
Beentjes et al [41]	Mixed methods	Touch-screen laptop or tablet	Lifestyle (ADL) support, social interaction, or companionship helps people find apps for self-management and meaningful activities	Female: 8 (3 in experimental group and 5 in control group); male: 12 (7 in experimental group and 5 in control group); location: home
Bernini et al [42]	Mixed methods	Touch-screen laptop or tablet	CBT[d]	Female: 5; male: 5; location: home
Bernini et al [43]	Quantitative	PC	CBT	Female: 17 telehealth and 13 in-person; male: 14 telehealth and 12 in-person; location: clinic (PC) or home (laptop)
Bogza et al [44]	Mixed methods	Web solution	Decision aid	Female: 6; male: 6; location: clinic, research center, or participant home
Bouzida et al [45]	Mixed methods	Robot with integrated touch screen	CBT	Female: 1; male: 2; location: home
Chang et al [46]	Quantitative	Screen and gesture sensitive device	Physical and cognitive training	Female: 8; male: 7; location: clinic
Chen et al [47]	Mixed methods	Touch-screen laptop or tablet	CBT	Female: 44 in evaluation group and 4 in focus group (same people); male: 13 in evaluation and 0 in focus group; location: clinic
Christiansen et al [48]	Qualitative	No device	Data collection (eg, UI[e] design feedback, technology use or acceptance, icon design preferences, changes to cognition, cognitive training or games design, and needs for robotic assistants)	Female: 6 (2 were 70 - 75 years and 4 were older than 81 years); male: 12 (4 were 70 - 75 years, 6 were 76 - 80 years, and 2 were older than 81 years); location: 15 at home and 3 in lab
Collette et al [49]	Qualitative	Integrated computer-based assistant	Health or medication management eg, BP[f], reminders to drink water	Female: 7; male: 3; location: clinic
Contreras-Somoza et al [50]	Qualitative	Touch-screen laptop or tablet	Lifestyle (ADL) support, social interaction, or companionship	Female: 11; male: 2; location: clinic
Contreras-Somoza et al [51]	Qualitative	Touch-screen laptop or tablet	Lifestyle (ADL) support, social interaction, or companionship	Female: 21; male: 14; survey: 35 participants (15 Spain, 4 Netherlands, 2 Italy, 2 France, 4 Israel, 5 Serbia, and 3 Slovenia)
Cunnah et al [52]	Qualitative	Web solution	Social interaction, messaging, chat, or information provision	Total: 100; location: home-100 dyads (49 control, 51 intervention); intervention group got 1:1 training followed by group training; 25% drop-out; 75 dyads completed the study (39 control and 36 intervention).

XSL•FO
RenderX

Study details	Study type	Type of technology	Purpose of the solution	Participant details
Demiris et al [53]	Mixed methods	Smart conversational assistant with touch screen (eg, Google Home and Alexa)	Companionship and reminders (computer-based pet)	Female: 10; male: 0; location: home
Dixon et al [54]	Qualitative	No device	Data collection (eg, UI design feedback, about technology use or acceptance, icon design preferences, changes to cognition, cognitive training or games design, and needs for robotic assistants)	Female: 1; male: 1; location: home
Franco-Martín et al [55]	Qualitative	PC	CBT	Not stated
Gasteiger et al [56]	Mixed methods	Robot with integrated touch screen with pen and magnetic blocks (tactile sensors) Used 3 robots (Bomy 1, Bomy, and Silbot) in the project	CBT	Not stated; location: location most convenient to participants, including the university, workplaces, clinic, home, or via Skype (for experts; developed by Janus Friis, Niklas Zennström, Ahti Heinla, Priit Kasesalu, and Jaan Tallinn)
Gelonch et al [57]	Mixed methods	Wearable camera	Memory aid-digital camera records daily activities	Female: 4; male: 5; location: adult day center
Givon Schaham et al [58]	Mixed methods	Touch-screen laptop or tablet	CBT	Female: 13; male: 15; location: clinic and home
Guzman-Parra et al [59]	Quantitative	No device	Data collection (eg, UI design feedback, about technology use or acceptance, icon design preferences, changes to cognition, cognitive training, games design, and needs for robotic assistants)	Female: 576; male: 510; location: used secondary data; 1086 dyads: 299 with dementia, 787 with MCI; data not split for the MCI and dementia groups
Haesner et al [60]	Qualitative	No device	Data collection (eg, UI design feedback, about technology use or acceptance, icon design preferences, changes to cognition, cognitive training, games design, and needs for robotic assistants)	Female: 3; male: 3; location: clinic
Hassandra et al [61]	Mixed methods	VR headset and controller, hand motion trackers, or wireless mouse	Physical and cognitive training	Female: 19; male: 8; location: clinic
Heatwole and Kendra [62]	Qualitative	No device	Data collection (eg, UI design feedback, about technology use or acceptance, icon design preferences, changes to cognition, cognitive training, games design, needs for robotic assistants)	Female: 6; male: 4; location: outing location
Hedman et al [63]	Qualitative	No device	Data collection (eg, UI design feedback, about technology use or acceptance, icon design preferences, changes to cognition, cognitive training, games design, and needs for robotic assistants)	Female: 2; male: 4; location: home

XSL•FO
RenderX

Study details	Study type	Type of technology	Purpose of the solution	Participant details
Hedman et al [64]	Mixed methods	No device	Data collection (eg, UI design feedback, about technology use or acceptance, icon design preferences, changes to cognition, cognitive training, games design, and needs for robotic assistants)	Female: 18; male: 19; location: home; only 21 participants left at year 5
Horn et al [65]	Qualitative	Mobile phone and smartwatch	Memory aid-facial recognition aid to identify people	Female: 6; male: 14; location: home
Hu et al [66]	Mixed methods	No device	Data collection (eg, UI design feedback, about technology use or acceptance, icon design preferences, changes to cognition, cognitive training, games design, and needs for robotic assistants)	Female: 18; male: 13; location: clinic
Infarinato et al [67]	Mixed methods	A robot and a separate touch screen	Physical and cognitive training	Female: 8; male: 7; location: home
Irazoki et al [68]	Qualitative	Touch screen laptop or tablet	Cognitive training or cognition assessment	Female: 11; male: 2; location: clinic
Korchut et al [69]	Mixed methods	No device	Data collection (eg, UI design feedback, about technology use or acceptance, icon design preferences, changes to cognition, cognitive training, games design, and needs for robotic assistants)	Female: 36; male: 21; survey
Kubota et al [70]	Qualitative	Robot with integrated touch screen	CBT	Female: 0; male: 3; location: online session
LaMonica et al [71]	Quantitative	No device	Data collection (eg, UI design feedback, about technology use or acceptance, icon design preferences, changes to cognition, cognitive training, games design, and needs for robotic assistants)	Female: 127; male: 94; survey: 137 had MCI, 61 had SCI[g], and 23 had dementia
Law et al [72]	Mixed methods	Robot with integrated touch screen	CBT	Female: 6; male: 4; location: lab
Lazarou et al [73]	Mixed methods	No device	Data collection (eg, UI design feedback, about technology use or acceptance, icon design preferences, changes to cognition, cognitive training, games design, and needs for robotic assistants)	Female: 9; male: 6; survey
Leese et al [74]	Mixed methods	Fitness tracker smartwatch	Collect daily movement and exercise data	Female: 1; male: 14; location: home
Lindqvist et al [75]	Qualitative	No device	Data collection (eg, UI design feedback, about technology use or acceptance, icon design preferences, changes to cognition, cognitive training, games design, and needs for robotic assistants)	Total: 5; location: clinic

XSL•FO
RenderX

Study details	Study type	Type of technology	Purpose of the solution	Participant details
Madjaroff and Mentis [76]	Qualitative	No device	Data collection (eg, UI design feedback, about technology use or acceptance, icon design preferences, changes to cognition, cognitive training, games design, and needs for robotic assistants)	Female: 3; male: 2; location: clinic
Maier et al [77]	Qualitative	Wearable watch, smartphone, and smartboard	Memory aid for daily activities or cognitive support or rehabilitation or share health information	Total: 6; location: clinic; 3 in focus group, 3 in usability evaluation of 2 prototypes; prototype 1 used a smartphone (not wearable technology) and prototype 2 used a smartphone and a smartboard (wall calendar)
Manca et al [78]	Mixed methods	Robot with integrated touch screen	CBT	Female: 9; male: 5; location: lab
Manera et al [79]	Quantitative	VR headset and controller or wireless mouse; screen and gesture-sensitive device; 3D glasses	CBT-train selective and sustained attention using image-based rendered environment	Female: 13; male: 15; location: lab
Mathur et al [80]	Mixed methods	Smart conversational assistant with touch screen (eg, Google Home and Alexa)	Health or medication management (eg, BP) reminders to drink water	Female: 4 total (phase 1 - 2, phase 2 - 2); male: 8 total (phase 1 - 5, phase 2 - 3); location: home
Matsangidou et al [81]	Mixed methods	VR headset	Regulate emotions and mood	Female: 19; male: 11; location: clinic
Mattos et al [82]	Mixed methods	PC and smartwatch	Insomnia intervention or sleep management	Female: 7; male: 5; location: home-only 10 completed the study
McCarron et al [83]	Mixed methods	Mobile phone and smartwatch	Memory aid-facial recognition aid to identify people	Female: 25; male: 23; location: home-29 participants had dementia
Mehrabian et al [84]	Qualitative	Touch screen laptop or tablet	Lifestyle (ADL) support, social interaction or companionship	Female: 19; male: 11; location: lab
Mondellini et al [85]	Mixed methods	VR headset and controller or hand motion trackers or wireless mouse	Cognitive training or cognition assessment	Female: 14; male: 1; location: lab
Moro et al [86]	Mixed methods	Humanoid robot, cartoon robot, and tablet	Lifestyle (ADL) support	Female: 6; male: 0; location: kitchen
Mrakic-Sposta et al [87]	Mixed methods	VR and stationary bike with controller on handlebars	Physical and cognitive training	Female: 6 (3 in experimental group, 3 in control group); male: 4 (2 in experimental group and 2 in control group); location: lab
Nie et al [88]	Mixed methods	Web solution and Webcam	Social interaction or messaging or chat or information provision	Total: 7 (5 in phase 1 and 2 in phase 2); location: clinic
Nieto-Vieites et al [89]	Mixed methods	Touch screen laptop or tablet	CBT	Female: 5 (study 2) and unknown (study 3); male: 3 (study 2) and unknown (study 3); location: clinic
Ortega Morán et al [90]	Qualitative	Touch screen laptop or tablet	CBT	Female: 16; male: 3; location: clinic
Park et al [91]	Quantitative	Touch screen tablet and wearable motion sensor device	Physical and cognitive training	Female: 12; male: 2; location: home

XSL•FO
RenderX

Study details	Study type	Type of technology	Purpose of the solution	Participant details
Park et al [92]	Quantitative	VR headset and controller or hand motion trackers or wireless mouse	CBT	Female: 7 control group and 7 VR group; male: 3 control group and 4 VR group; location: clinic
Piasek et al [93]	Mixed methods	Robot with integrated touch screen	Physical and cognitive training	Female: 3; male: 1; location: home
Pino et al [94]	Mixed methods	Robot with integrated touch screen	Lifestyle (ADL) support, social interaction or companionship	Female: 6; male: 4; location: clinic
Quintana et al [95]	Mixed methods	Touch screen laptop or tablet with tablet pen	Memory aid for daily activities, cognitive support, rehabilitation, or share health information	Female: 3 (Sweden) and 5 (Spain); male: Sweden-6 (Sweden) and Spain-5; location: clinic and home
Rossi et al [96]	Mixed methods	Robot with integrated touch screen	Lifestyle (ADL) support	Female: 2 (phase 1) and phase 3; male: 2 (phase 1) and phase 4; location: lab (phase 1) and home (phase 2)-
Saini et al [97]	Mixed methods	Videoconferencing (Zoom) and webcam	CBT	Total: 12; survey: 6 in each randomized group (face-to-face CBT or CBT via videoconferencing)
Scase et al [98]	Qualitative	Touch screen laptop or tablet	CBT	Female: development: 3 focus groups (group 1: 4; group 2: 4; group 3: 3), evaluation: 22; male: development: 3 focus groups (group 1: 5; group 2: 1; group 3: 1), evaluation: 3; location: clinic and home
Scheibe et al [99]	Qualitative	TV with remote control or tablet; sphygmomanometer (to measure blood pressure)	Telemonitoring medical app to communicate with multidisciplined health professionals	Female: 8; male: 4; location: home
Shamir et al [100]	Mixed methods	Touch screen laptop or tablet	CBT	Female: 6; male: 8; location: home
Shellington et al [101]	Mixed methods	Mobile phone	CBT	Female: 14; male: 5; location: home
Shin et al [102]	Qualitative	Remote controlled robot (robot is not with user, but controlled via app on PC, tablet, or mobile phone)	Lifestyle (ADL) support	Female: 0; male: 6; location: clinic or home
Stogl et al [103]	Quantitative	Robotic walker (user holds handles and pushes the device for exercise)	Physical and cognitive training	Female: 2; male: 8; location: clinic
Stramba-Badiale et al [104]	Mixed methods	VR (or TV), joystick, and foot pad or rudder	Assist with navigation and training of spatial memory	Female: 4; male: 3; location: lab
Tuena et al [105]	Mixed methods	VR (or TV), joystick, foot pad or rudder, and 3D glasses	Assist with navigation and training of spatial memory	Female: 2; male: 6; location: lab
Van Assche et al [106]	Qualitative	Robot with integrated touch screen	Data collection and design guidance for robot use cases	Female: 16; male: 14; location: home
Votis et al [107]	Qualitative	Touch screen laptop or tablet	CBT	Total: 17; location: lab
Wargnier et al [108]	Mixed methods	Smart conversational assistant with touch screen (eg, Google Home and Alexa)	Health or medication management (eg, BP and reminders to drink water)	Female: 11; male: 3; location: clinic (5 had Alzheimer disease)

XSL•FO
RenderX

Study details	Study type	Type of technology	Purpose of the solution	Participant details
Weering et al [109]	Qualitative	Web solution	Physical and cognitive training	Female: 43; male: 14; location: clinic
Wolf et al [110]	Mixed methods	VR headset	Lifestyle (ADL) support	Female: 6; male: 0; location: clinic kitchen
Wu et al [111]	Qualitative	No device	Data collection (eg, UI design feedback, about technology use or acceptance, icon design preferences, changes to cognition, cognitive training, games design, and needs for robotic assistants)	Female: focus group 4, interview 12; Male: focus group 1, interview 3; location: lab
Wu et al [112]	Mixed methods	Robot with integrated touch screen	Lifestyle (ADL) support	Total: 6; location: clinic
Yamazaki et al [113]	Qualitative	Robot (sitting on table, programmed to interact with user, sing, talk, and respond to spoken words)	Companionship	Female: 2; male: 0; location: home
Yun et al [114]	Quantitative	VR headset and controller or hand motion trackers or wireless mouse	CBT	Female: 6; male: 5; location: clinic
Yurkewich et al [115]	Mixed methods	Touch screen laptop or tablet	Social interaction, messaging, video chat, or information provision	Female: 6; male: 2; location: home
Zafeiridi et al [116]	Mixed methods	Web solution	Lifestyle (ADL) support, social interaction, or companionship	Female: 14; male: 10; location: home
Zedda et al [117]	Mixed methods	Robot with integrated touch screen	CBT	Female: 6; male: 10; location: clinic
Zhang and Gao [118]	Mixed methods	No device	Data collection (eg, UI design feedback, about technology use or acceptance, icon design preferences, changes to cognition, cognitive training, games design, and needs for robotic assistants)	Total: 5; location: clinic
Zhang and Liu [119]	Mixed methods	Touch screen laptop or tablet with foot stand with pressure sensors	Physical and cognitive training	Female: 79; male: 58; location: survey and usability evaluation in lab: 137 responses for phase 1 (requirements gathering) but only 30 responses were from people with MCI; 107 responses were from family of people with MCI
Zhu et al [120]	Mixed methods	Mobile phone	Memory aid-digital storytelling	Female: 9; male: 3; location: clinic
Zubatiy et al [121]	Mixed methods	Smart conversational assistant with touch screen (eg, Google Home and Alexa)	Lifestyle (ADL) support	Female: 4; male: 6; location: home (MCI and caregiver dyads included in study)

[a]VR: virtual reality.

[b]MCI: mild cognitive impairment.

[c]ADL: activity of daily living.

[d]CBT: cognitive behavioral therapy.

[e]UI: user interface.

[f]BP: blood pressure.

[g]SCI: subjective cognitive impairment.

XSL•FO
RenderX

Below, we first list 4 broad categorizations of the examined studies in terms of (1) study details and nature of participants, (2) purpose of technological solutions, (3) types of interaction devices, and (4) interaction modalities, and provide a short description of how studies evaluated usability. This is followed by the resulting themes derived from analysis of participant feedback.

Categorization of Examined Study Details

Characteristics of Included Studies and Study Participants

Over half of the included studies (n=45) were conducted in Europe, with almost a quarter (n=10) of these in Italy. Some of the European studies included participants from multiple countries [51,59,67,69,90,95,116]. Table 1 shows that 61 studies included fewer than 20 participants, 49 studies included fewer than 15 participants, 25 studies included fewer than 10 participants, and 8 studies included fewer than 5 participants. Many studies commented on the lack of cultural diversity among their participants.

Studies used a mixed methods (n=47), qualitative (n=26), or quantitative (n=10) approach. Studies that did not explicitly declare their design type are classified in this paper based on the type of data collected. Qualitative studies used semistructured interviews, focus groups, workshops, and observations to collect data. Quantitative studies used surveys or questionnaires with closed questions to gather data. Most studies, regardless of type, collected quantitative data about participant demographics. Quantitative data about usability, acceptability, and user experience (UX) were typically gathered using questionnaires with responses on a Likert-type scale.

Purpose of Technology Solutions in Studies

Over a third (n=32) of solutions provide cognitive training in the form of cognitive behavior therapy (CBT; n=21), a combination of CBT and physical training (n=9), or CBT and cognitive assessment (n=2). Other functions mentioned as needs by participants and supported by solutions were lifestyle or activities of daily living support (n=13), memory training or management (n=6), health and medication management (n=4), staying in touch with family and friends, social engagement, entertainment (n=4), companionship (n=2), spatial navigation or orientation (n=2), insomnia intervention (n=1), and daily exercise monitoring (n=1). Solutions also provided support for decision support (n=1) and mood regulation (n=1), but these were not identified as needs by participants. Further needs identified by participants but not supported by solutions were the management of finances and help to locate lost items. Sixteen studies were designed to collect data to (1) guide the design of future solutions, (2) inform amendments for existing solutions, or (3) collect information about the needs of people with MCI.

Interaction Devices

Studies included various technology platforms and devices as presented in Table 2 in the Multimedia Appendix 1. Almost half (n=37) of the proposed solutions include a touch screen as a component of a laptop or tablet (n=17), a robot (n=14), or a mobile phone (n=6). This reflects research by [25,122], which showed that touch screen devices are effective and easy to use by older adults. Some (n=4) robot solutions did not include sound, but the majority (n=12) included voice or sound. Eleven solutions involved the use of virtual reality (VR). Eight solutions included wearables (other than VR headsets or 3D glasses) in the form of a smartwatch (n=5), a wearable motion sensor (n=1), a sphygmomanometer to measure blood pressure (n=1), and a wearable camera (n=1). Six studies included a mobile phone, among which 3 paired the phone with a smartwatch. Five studies used a conversational agent such as Google Home (n=3), an integrated conversational agent on a laptop (n=1), and a computer-based pet avatar on a laptop or tablet (n=1).

Interaction Modalities

The variety of technology and devices made different options available for users to interact with the solutions and vice versa. Over half the studies (n=48) involved touch interaction either as unimodal (n=24) or multimodal (n=24). Seven of these touch solutions could also be used with a keyboard and mouse. This reflects research by [25,122], which demonstrates that touch screen interaction is an effective interaction mode for both old and young users, even with some usability issues. Nine studies combined a touch screen device with speech output, while 13 studies combined a touch screen device with speech input and output. Combining visual and speech modalities is supported by research findings that multimodal interaction is more effective and usable for older adults [123], as compared to unimodal touch or speech [124]. Voice or sound was associated with 29 solutions; 13 studies involved solutions that provided speech or sound outputs, and 16 studies involved solutions that afforded speech or sound input and output. Voice or sound outputs were associated with VR solutions (n=4), robots (n=6), tablet device solutions (n=2), and a web solution (n=1). Voice or sound input-plus-output was associated with VR solutions (n=2), robots (n=6), tablet device solutions (n=3), embodied conversational agent solutions (n=3), a web solution (n=1), and a mobile phone app (n=1). Only 2 studies used gesture as an interaction modality; 1 used a robot that gestured to the user as a demonstration of extraversion [117], and the other was a cognitive game based on Tetris, which included a user gesture capture device [46]. Four studies used haptic (vibration) interaction via a smartwatch and a touch screen device (n=2), a vibrating keyboard with a touch screen device (n=1), or a vibrating handset as part of a VR solution (n=1). Additional interaction modalities provided by VR solutions included joysticks and pedals or rudders (n=4), and robot solutions with lights, facial expressions, and movement (n=7). Table 2 in the Multimedia Appendix 1 lists the device and interaction modality combinations for each study.

How Studies Evaluated Usability or Acceptability

Studies used different methods to evaluate usability. Some studies used the System Usability Scale [42,44-46,61,80,81,88,95,104,105,109], which is a simple, 10-item Likert scale used to assess usability [125]. Research [126] has shown that the System Usability Scale is suitable for evaluating digital health applications. Some studies of robotic solutions used the Almere questionnaire [127], which is based

XSL•FO
RenderX

on UTAUT and developed to evaluate acceptance of assistive social agents by older adults [86,93,94,96,108,112]. Several studies designed their own questionnaires based on TAM, TAM3, STAM, MATOA, or UTAUT.

Themes Derived From Thematic Analysis of Participant Feedback

Inductive thematic analysis of extracted data identified five themes: (1) purpose and need, (2) solution design and ease of use, (3) self-impression (how the solution makes the user feel), (4) lifestyle (how well the solution fits with the user's life and routines), and (5) interaction modality.

Purpose and Need

Participants identified that improving or maintaining their existing cognition level is the primary need, and over a third of solutions focused on this function. The remaining studies reveal other needs identified by participants, and only 2 studies provided solutions for activities that were not identified as needs. Assistance with managing personal finances was identified as a need [75,76]; however, none of the studies provided a solution to this end. Being able to manage one's own finances decreases vulnerability and is an important enabler for independence. Further investigation is needed to determine how to support older adults with MCI to have control of and manage their finances in a safe and secure manner. Participants also identified help to find lost items as a need [93], which was again not supported by any studies. The answer to RQ1A is that proposed solutions for older adults with MCI are perceived as useful, but there are also gaps where functional support is missing. Given the demonstrated importance of perceived usefulness for TA, this is an area where further work is needed.

Solution Design and Ease of Use

Study participants reported a variety of usability and usage issues and noted that an ability to configure and personalize a solution provides flexibility and increases usability.

> *There's all grades of dementia as well isn't there. And some would need more help than others.* [56]

Participants said it would be useful if they could adjust (1) solution tasks to match user skills [42,49,52,60], (2) task complexity [45,55,56,81], (3) task workflow [115], (4) timing of reminders [65], (5) user interface design (colors of background, text, and widgets [56,98], size of text and widgets [47,66], and color contrast [56]), (6) speed, volume, and accent of speech [56,95], (7) language (formal or casual) [108], (8) robotic facial features and height [69,86,94,112], and (9) avatar image [50].

Consistent design, labeling, and widget placement were mentioned as important enablers of usability [88,120], and this is especially critical for users with cognitive impairments or memory concerns. Participants added that they may use multiple toolsets and that seamless integration and consistency across these toolsets would increase usefulness and ease of use.

Participants provided ideas about how to increase usability of solutions, such as (1) reduce or eliminate the need to remember passwords or specific trigger words because this increases complexity [121], (2) make messages specific and include sufficient detail so the message is actionable [80], (3) provide sufficient time to read text messages [42,68,105], (4) use language appropriate for older users [98], (5) use simple, familiar, and common words [88,95], (6) use sound as well as visual indicators to make it obvious that an action was performed successfully [95], (7) avoid loud noises and too many animations because these create a distraction [60], and (8) make task steps simple [119]. Participants prefer solutions that provide immediate feedback about task progress and task completion because they are engaging and provide motivation and incentives for use [60,63,74,80,90,64]. This illustrates how UX can influence TA.

Specific feedback points about user interfaces were (1) layout should be simple and not too busy or overwhelming so that controls are obvious and the interface does not overstimulate or place a burden on the user's cognition [44,68,88,120], (2) use visual aids [70], (3) use simple images so the item is easily recognizable [66,110], (4) use intuitive and simple navigation [41,107], (5) use bright and contrasting colors [89,98,107], (6) use large text, buttons, and widgets [40,88,107], (7) ensure all options are visible so information is transparent [44], and (8) list options in order of user priorities because users often select the first option listed [44]. These comments are consistent with design principles proposed by Nielsen [128-130].

Reliable technology creates trust, confidence, and feelings of empowerment [94,95]. Studies where participants experienced technical issues [56,67,72,108], or if the device had a low battery life or charged slowly [99], resulted in frustration and negative feedback.

Participants commented that age-related physical limitations may affect how they use technology, so accessibility is an important design feature of technology solutions for older adults with MCI [40,95,98,121]. Context of use and location were also raised as considerations in this regard (eg, rural vs metropolitan and speed of internet connection [55]).

Studies showed that older adults with MCI prioritize security and privacy [48,94]. Some participants said they prefer robots to have mechanical features and to be shorter than a person [94,112] because they felt less under surveillance. This preference has also been reported by other research [131]. Safety was identified as a priority, related to both (1) personal safety when using technology (especially robots or VR solutions) and (2) when using the internet or social media [48]. Participants said that even though privacy is a need, they are willing to compromise privacy and share personal data or their location with medical professionals or carers in an emergency [73]. Some participants said that it would be useful to receive regular reminders of what data is being shared [116].

Given the breadth of usability issues reported and the number of changes suggested, the answer to RQ1B is that proposed solutions for older adults with MCI are not perceived as completely easy to use, and this is an area where future technologies need to focus.

Self-Impression

Participants commented that the feelings they experience while using a solution can positively or negatively affect their comfort,

XSL•FO
RenderX

confidence, and overall well-being. The wearable camera solution made participants feel uncomfortable because it drew attention to them, and they felt conspicuous.

> *If people don't know you, they look at you as if they are thinking. Well, what is he doing?* [57]

Participants also felt vulnerable because the camera may capture private moments, such as when they are going to the bathroom.

> *The camera is black and people look at it a lot. I think that if it were a more natural color, people would look at it less.* [57]

Some participants said that technology would be useful for other people, but not themselves. Technology is often associated with negative aspects of aging, such as loneliness or ill health [111] because it advertises the user's limitations and lack of independence.

> *The telecare system would be useful for me if I had more deficits. But so far, I can manage by myself at home* [84]

Participants said they want solutions to reflect their identity and present a positive self-image. They suggested that solutions should (1) use affirming language that does not remind them about their cognitive limitations [40,68,72,76], (2) include nondirective task workflows (eg, participants prefer a system to check in rather than remind [80]), and (3) not draw attention to the user or the solution; this is especially true for robotic solutions [94,111,112] and the wearable camera [57]. Self-concept and social presence have been shown to be determinants of TA in the MATOA model, and a literature review [132] describes similar feedback.

Participant personality traits did not appear to affect acceptance of robots in the long term [96]; however, differences were observed over the study duration. Some participants felt comfortable engaging with a robot and interacted with it in a human-like manner [86,106]. They said a robot is novel and suggested that it should be human-like with autonomy, be able to perform simple tasks, and interact via 2-way communication [45,69]. This is consistent with existing research [133], which demonstrated that robots that take initiative, exhibit more activity, and are extraverted are better accepted by users.

> *I don't see much difference between [my doctors] and the robot except one is ran by electricity and the other one is a human being.* [45]

Other participants said they want to engage with real people rather than anthropomorphic technology. They value human connection [88,106,111], and there is a concern that the use of technology will reduce their existing human contact. These participants preferred more machine-like robots that did not imitate human traits.

> *A robot doesn't have a heart, It must be for people who are very handicapped. It's not for me. It makes me think that my life is terminated. I'd rather die than live with a robot.* [112]

Other concerns identified by participants relating to the use of robot solutions were that the robot would (1) be hard to use or experience technical problems, (2) not do what they needed or wanted it to do, (3) be expensive, (4) increase isolation, or (5) make them too dependent on the robot [94,111,112]. Nevertheless, many participants felt optimistic about anthropomorphic solutions and recognized their potential.

> *The human aspect, the "feeling," is hard to create with a robot. But there is an enormous progress in that world, and I do believe in it. Of course, still in combination with us, humans.* [106]

Participants who said they prefer more machine-like robots were less willing to create an emotional connection with the robot [111,112]. However, some participants got very attached to their robot solutions, treated the robot like a friend, and became sad when the study ended [53,69,78,113]. This raises important issues for designers about how to support users if technology solutions fail or become unavailable.

Lifestyle

Participants said that independence and autonomy are important, and some said they felt a responsibility to their relatives and carers to maintain their self-sufficiency [45]. Participants want to feel in control of their lives [45,51,60,75,76,80,88,103,111], and technology solutions that create a sense of user empowerment are preferred rather than solutions that impose mandatory interventions or training [60]. While participants appreciate the increased independence technology can provide [76], they are also concerned about receiving too much assistance (from technology or people) in case they become reliant on this and lose their autonomy [111] or independence [69,76]. The studies reinforced that IT literacy, training, and availability of support from carers and family are enablers for acceptance and continuing use of technology [47,48,63], which is consistent with TAMs. However, it is important that technology solutions do not create a dependence on support services, which in turn would erode independence.

Cost was identified by study participants as an important factor when considering adoption of a technology solution [63,84,94,106,112]. This is consistent with UTAUT2 and H-TAM, which include cost, price, and value or perceived benefit as determinants of adoption.

Physical comfort and convenience are important considerations for the use of technology. Most participants enjoyed the novelty of VR solutions; however, some experienced visual discomfort or nausea, found the headsets uncomfortable, or said the VR solutions were not practical [79,81,110, 39].

> *Well, I liked it, but it was too heavy and big for me. I had difficulty breathing because it wasn't staying in place and fell into my face blocking my nose.* [81]

The answer to RQ2A is that participants prefer devices that are light and portable (mobile phone or tablet), commonly owned (mobile phone), readily available, and have a reasonably large screen [74,84,119,120]. This review also identified that older adults with MCI do not like small wearable technology, such as watches, because the screens are too small, and many participants said that watches are not convenient because they are no longer part of their everyday toolset [74]. In addition, the physical size of the technology solution and its placement in the home affect ease of access and contribute to convenience

XSL•FO
RenderX

(eg, smaller robots take up less room and are less intrusive in a home setting [45,106]).

Studies where solutions were integrated into participants' existing lifestyles and routines made participants feel comfortable, confident, and more in control. This is consistent with the STAM, MATOA, and H-TAM models, which show that compatibility, ease of use, and maintaining a sense of control are important for older users of technology.

> *I like that Google kind of feels like a part of the house, and not something that I have to keep answering to all the time like my morning alarm.* [80]

Interaction Modality

Study participants provided positive feedback for voice interaction and said it was intuitive and easier than writing [45]. They also mentioned challenges such as (1) the solution's accent was hard to understand or the solution did not understand the user's accent or speech pattern [56,72], (2) the speech was too fast or the volume was too loud or too soft [68,78], (3) the solution spoke for too long [108], (4) the solution only recognized a limited vocabulary and participants had to remember specific "trigger" words to activate the solution or give it commands [53,110,121].

> *If it was more New Zealand English spoken it would be easier for me I think.* [56]

Similar feedback has been documented by other researchers studying voice interaction with older users [27,32]. Research has demonstrated that subtle differences in speech can be observed in older adults with MCI [30], and people with cognitive decline speak slower, in smaller chunks, pause more often, have longer silences, respond slower to questions, and provide shorter answers [134]. This highlights the importance of including older adults with cognitive impairments when designing and evaluating voice interaction solutions. Mozilla [135] is leading the world's largest open-source voice data project called "Common Voice." The project is building and refining a voice dataset that is being used to teach machines to speak, and so far, it includes 131 languages and 32,584 hours of recorded voice data. However, analysis [136] shows that people older than the age of 90 years are represented only in the Mongolian and Polish languages, and people in their 80s are only represented in the Esperanto, Turkish, Abkhazian, Toki Pona, and Hebrew languages. Of the English-contributing voices, there are 0.01% in their 70s, 0.04% in their 60s, 0.05% in their 50s, 0.09% in their 40s, 0.14% in their 30s, 0.25% in their 20s, 0.06% in their teens, and 0.36% of unspecified age. It is not known if any of the recorded voices are from older adults with MCI. Technology solutions using speech interaction increase their risk of error if they cannot recognize the speech of older users or the speech of people who are cognitively impaired. Users' current optimism creates an opportunity to invest in voice interaction to improve its usability and effectiveness.

Participants said that touch interaction is effective and easy, especially for less IT literate people [55,107], but they experience difficulties such as (1) trouble dragging items across the screen, (2) knowing how much pressure to use [100], (3) small screen or widget size making it hard to select the target accurately, and (4) using the touch screen if the user has dexterity problems [41,72,95,100]. Some participants chose to use an instrument to avoid these challenges [41,95].

> *I have a problem only with, first, with the gentleness of the finger. My finger isn't the gentlest thing.* [100]

These findings are consistent with outcomes from a literature review investigating the use of touch screens by older users [122], and an investigation of touch interaction, which identified usability issues associated with clicking, dragging, zooming, and rotating on a touch device for both older and younger users [25]. Based on feedback from study participants in this review, it appears that these challenges have not been fully resolved.

Participants said robotic assistants should be able to engage in both a verbal and nonverbal manner, and they preferred robots that can speak, listen, and answer questions rather than those with 1-way interaction via a touch screen [45,69]. If the solution is embodied, they want to be able to touch it physically and interact with it rather than simply view it on a screen [53]. Active robots that performed tasks were deemed more pleasant [96]. This is consistent with findings that robots with unimodal auditory 2-way feedback are highly usable and acceptable for older adults [27].

Participants were agnostic to haptic interaction, except for conveying that a smartwatch was not a preferred device, and suggested that it is useful to be able to control technology remotely [108,119], which supports the need for comfort and convenience.

The answer to RQ2B is that participants prefer multimodal interaction, in particular speech, visual or text, and touch [60,80,88,115,118,121] because it is flexible, provides options for use if people have physical or cognitive limitations, and makes solutions accessible to people with poor hearing or vision or speech limitations. This supports existing research about the effectiveness of different interaction modalities for older users, which showed that multimodal interaction is most effective [28,29,137] and that the effectiveness of interaction modalities and multimodalities is task- and user-dependent [26,29,138]. However, no studies included older adults with cognitive impairment, which creates a gap in understanding the effectiveness of different interaction modality combinations for this population.

Discussion

Principal Findings

Participants across studies reported similar needs but often in a different priority order, and this scoping review cannot (1) provide a definitive priority order for needs; (2) identify if needs are relatable to technologies or devices; or (3) attribute differences to the context of use, cultural differences, participant demographics, or the way MCI is affecting participants' lives. Mondellini et al [85] found that user needs appear to be location-specific, but further research is required to understand which needs are location-specific versus those that are generalized.

XSL•FO
RenderX

Fadzil et al [139] completed a review of existing technology solutions for older adults with cognitive impairments and found that current solutions are focused on (1) caregiver support, (2) strengthening cognition, (3) cognitive rehabilitation, and (4) using technology for disease prevention or self-management. Fadzil et al [139] do not comment on reported needs, technology preferences, or opinions about existing solutions for older adults with cognitive impairments. Neither does it describe how well existing technology supports user needs, nor how well technology is aligned with users' preferred ways of doing things. This gap contributes to the lack of understanding of TA by older adults with MCI.

Our review identified that security and privacy are important for older adults with MCI; nonetheless, they are prepared to compromise these to feel safe. However, they do not want to be continuously watched or feel like they are under surveillance. Nastjuk et al [140] used the term "techno-invasion" to describe the stress caused by constant observation and monitoring. Solutions must balance the need for users' privacy and security with their need for safety and well-being.

A positive UX is a significant enabler for technology acceptance, and ease of use contributes greatly to UX, which suggests that designers must give both priority. User attitude has been identified in TAMs as a determinant of use, and several studies commented that MCI has an impact on a person's attitude toward technology use [48,74]. This observation has also been made by other researchers [43], who noted that we need a greater understanding of how sociodemographic factors and cognition affect attitude.

Some older adults with MCI are not willing to connect with a robot because they prefer to interact with a human, while others are comfortable engaging with anthropomorphic solutions. Our review identified that participants who created an emotional connection with an anthropomorphic solution were vulnerable to becoming depressed once the solution was no longer available. This raises important issues for designers about how to support users if technology solutions fail, and ethical issues about how to manage emotional connection and attachment to more anthropomorphic solutions. It illustrates that technology represents part of the solution but not the entire solution [75], and that service providers must balance the need for human connection with the need to increase efficiency and reduce demand on health services.

Our review found that participants experience a variety of usability and usage issues, and that the flexibility of use and the ability to configure a solution increase ease of use. Older adults with MCI feel more comfortable, confident, and in control when solutions are consistent, predictable, and integrate into their existing lifestyles and routines. People often use multiple toolsets, and technology providers should prioritize consistency and seamless integration of their products. Overall, our analysis reveals that technology solutions should support how people behave and live, not require them to change and adapt to how the technology operates.

Some studies recommend that technology solutions for older adults with MCI should be designed together with this population [48,73,74,95,102,107,116] to ensure that all needs are identified and appropriately prioritized. This recommendation is supported by our findings, which show that while there is consistency in the top-priority need, there are differences in how participants in different locations prioritize the order of needs. Existing research shows that there is cultural and contextual influence on TA, and differences in the needs and priorities for independent living of older adults in Sweden and Poland, and the technologies they would consider using [141]. Given the evidence for the importance of perceived usefulness and ease of use as factors for TA, it is critical to understand the needs and characteristics of the target user base so we can design to maximize TA. User-centered design (UCD), first proposed by Kling [142], is a design methodology that focuses on usability and UX. UCD promotes the inclusion of end users in every stage of solution creation, starting with initial data gathering to understand users' characteristics, environment of use, and requirements. UCD is iterative and incorporates regular artifact evaluation and refinement based on user feedback. UCD processes and techniques have been applied successfully to design assistive technologies for people with dementia and shown to be an effective way to accurately capture requirements and ensure solution usability and acceptance [143].

Scope of Future Work

This review revealed that older adults with MCI have opinions about how and when they want to use technology, and it is imperative that future solutions are designed collaboratively with this population. Research is needed to accurately identify the needs of older adults with MCI so we understand what would be considered useful and how best to position technology in their lives. It would also be beneficial to identify needs that are common and which needs vary depending on demographics, location, or other factors.

Solutions designed for older adults with MCI must generate a positive experience because this affects a user's comfort, confidence, and overall well-being. This review collated suggestions from participants, but more research is required to fully understand the factors to be considered.

Participant feedback suggests that personalization and control, and enablement are important for older adults with MCI, but it is not clear how choice with respect to technology, device, and interaction modality affects TA. This review identified that older adults with MCI do have preferred devices and interaction modalities; however, further work is needed to fully understand the effectiveness of different combinations of interaction modalities and how these are influenced by the effects of MCI, context of use, device being used, and task being performed.

Based on the findings, older adults with MCI want to have autonomy, independence, and feel safe. They are accepting of technology if it positively contributes to their lifestyle and provides a sense of well-being. Future research should aim to provide clarity about how to effectively meet safety, privacy, and security needs and establish technology as an enabler for independence and autonomy in older age.

Finally, the collection of longitudinal data about technology use by older adults with MCI would capture changes to usage patterns, preferences, and adoption factors as MCI advances

XSL•FO
RenderX

and enable us to understand how solution designs should allow for adaptation to support MCI progression.

Study Strengths and Limitations

This review comprehensively searched 9 databases for studies over a 10-year period. It considered research about a broad set of technology solutions proposed for older adults with MCI in the pre– and post–COVID-19 environments.

However, the study is not without limitations. This scoping review only included studies in English, so it is possible that untranslated relevant studies were overlooked. Most studies published their outcomes promptly after data collection; however, [56] began data gathering in 2016 but did not publish the results until 2022. This is a long time in terms of technology evolution and potential for change in people's opinions and attitudes. Some studies did not analyze nonverbal feedback, so it is possible that important information was missed [50,68].

Most studies established an MCI diagnosis using the Montreal Cognitive Assessment [144], the Mini-Mental State Examination [145], or Petersen's criteria [146]. However, the cutoff criteria for each assessment tool were not applied consistently across studies.

The definition of "older adult" is not consistent across the 83 studies. Most studies considered older than 65 years as "older;" however, some studies included participants aged between 48 - 59 years. Some studies did not have an equal representation of genders.

Many studies listed small participant numbers or a lack of cultural diversity as limitations, and most studies included participants from only one location, meaning it may not be possible to generalize outcomes. Less than half of the studies were completed in a real-life setting, so outcomes may not be fully representative of real-life results.

Studies did not collect a consistent set of participant data (such as IT literacy, existing devices and technology being used, education level, or profession), so it is not possible to comment on how these may have impacted participants' opinions, although age, sex, education level, or experience are listed as moderators in existing TAMs such as TAM, UTAUT, and MATOA.

Some studies did not have people with MCI represented in all phases or excluded people with physical limitations, so the opinions of these people have not been considered. In addition, the perspectives of other important stakeholders, such as medical and allied health professionals who treat older adults with MCI, are not considered in this paper. A study exploring the perspectives of 20 medical and allied health professionals from France, Italy, Japan, and Germany, who treat older adults with cognitive impairments [147], identified that physical, social, and cognitive interventions are considered important, and technology is an effective way to deliver and support these. This scoping literature review is the first part of a multistage research project, and future studies will include other stakeholders, such as medical and health professionals, family, and caregivers.

Conclusions

This scoping review identified that improving or maintaining their current level of cognition is the top-priority need for older adults with MCI. Proposed technology solutions are perceived as useful by this population, even though gaps exist where functional support is missing. However, they are not perceived as completely easy to use, due to a variety of usability issues. In addition, technology solutions may have a positive or negative effect on the user's comfort, confidence, self-esteem, self-image, and overall well-being, depending on the feelings and emotions triggered by the solution design. Older adults with MCI have preferences for how they use and interact with technology, and which devices they use. They prefer multimodal interaction—especially speech, visual or text, and touch—because it is effective and provides options for use, and devices that are comfortable, convenient, lightweight, readily available, cost-effective, and have large screens. Future work should focus on (1) gathering information about the needs of older adults with MCI and clarifying how these are dependent on location, demographics, or other factors; (2) improving ease of use and UX; (3) enabling customization and personalization; (4) further exploring interaction preferences and the effectiveness of different interaction modes; (5) providing options for multimodal interaction; and (6) integrating technology solutions more seamlessly into people's existing lifestyle and routines.

Acknowledgments

The authors would like to thank the University of Canberra librarian, Murray Turner, for his advice regarding databases to search. The first author wishes to gratefully acknowledge the support of her research by an Australian Government Research Training Program Scholarship [148].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study extract data.
[XLSX File, 144 KB - aging_v8i1e78229_app1.xlsx]

Checklist 1

XSL•FO
RenderX

PRISMA-ScR checklist.
[DOCX File, 86 KB - aging_v8i1e78229_app2.docx]

References

1. Health Resource Directory. URL: https://healthresourcedirectory.org.au [accessed 2024-12-17]
2. Dementia Australia. URL: https://www.dementia.org.au [accessed 2024-12-17]
3. Polcher A, Wolfsgruber S, Peters O, et al. A comparison of operational definitions for mild cognitive impairment. J Alzheimers Dis 2022;88(4):1663-1678. [doi: 10.3233/JAD-215548] [Medline: 35811516]
4. Anderson ND. State of the science on mild cognitive impairment (MCI). CNS Spectr 2019 Feb;24(1):78-87. [doi: 10.1017/S1092852918001347] [Medline: 30651152]
5. Díaz-Mardomingo MDC, García-Herranz S, Rodríguez-Fernández R, Venero C, Peraita H. Problems in classifying mild cognitive impairment (MCI): one or multiple syndromes? Brain Sci 2017 Sep 1;7(9):111. [doi: 10.3390/brainsci7090111] [Medline: 28862676]
6. Gillis C, Mirzaei F, Potashman M, Ikram MA, Maserejian N. The incidence of mild cognitive impairment: a systematic review and data synthesis. Alzheimers Dement (Amst) 2019 Dec;11(1):248-256. [doi: 10.1016/j.dadm.2019.01.004]
7. Population projections, Australia. Australian Bureau of Statistics. 2023 Nov 23. URL: https://www.abs.gov.au/statistics/people/population/population-projections-australia/latest-release [accessed 2024-12-17]
8. Prevalence of dementia. Australian Institute of Health and Welfare. 2024 Sep 13. URL: https://www.aihw.gov.au/reports/dementia/dementia-in-aus/contents/population-health-impacts-of-dementia/prevalence-of-dementia [accessed 2024-12-18]
9. Final report. Royal Commissions. 2021 Mar 1. URL: https://www.royalcommission.gov.au/aged-care/final-report [accessed 2024-12-17]
10. National dementia action plan 2024-2034. Australian Government, Department of Health, Disability and Ageing. 2024. URL: https://www.health.gov.au/our-work/national-dementia-action-plan [accessed 2025-01-07]
11. Almalki ZS, Simsim DA. The role of health technology in transforming healthcare delivery and enhancing spending efficiency. Global J Med Ther 2020;2(3). [doi: 10.46982/gjmt.2020.107]
12. Emery VK, Edwards PJ, Jacko JA, et al. Toward achieving universal usability for older adults through multimodal feedback. Presented at: CUU '03: Proceedings of the 2003 Conference on Universal Usability; Nov 10-11, 2003; Vancouver, BC p. 46-53. [doi: 10.1145/957205.957214]
13. Wildebos G, Peute L, Jaspers M. Facilitators and barriers of electronic health record patient portal adoption by older adults: a literature study. Stud Health Technol Inform 2017;235:308-312. [doi: 10.3233/978-1-61499-753-5-308]
14. Davis FD. A technology acceptance model for empirically testing new end-user information systems: theory and results [Doctoral thesis]. : Massachusetts Institute of Technology; 1986 URL: http://dspace.mit.edu/handle/1721.1/15192
15. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
16. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. Decision Sci 2008 May;39(2):273-315. [doi: 10.1111/j.1540-5915.2008.00192.x]
17. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
18. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 2012;36(1):157. [doi: 10.2307/41410412]
19. Chen K, Chan AHS. Gerontechnology acceptance by elderly Hong Kong Chinese: a senior technology acceptance model (STAM). Ergonomics 2014;57(5):635-652. [doi: 10.1080/00140139.2014.895855] [Medline: 24655221]
20. Wang KH, Chen HG, Chen HG. A model of technology adoption by older adults. Soc Behav Pers 2017 May 7;45(4):563-572. [doi: 10.2224/sbp.5778]
21. Heckhausen J, Schulz R. A life-span theory of control. Psychol Rev 1995 Apr;102(2):284-304. [doi: 10.1037/0033-295x.102.2.284] [Medline: 7740091]
22. Yap YY, Tan SH, Choon SW. Elderly's intention to use technologies: a systematic literature review. Heliyon 2022 Jan;8(1):e08765. [doi: 10.1016/j.heliyon.2022.e08765] [Medline: 35128090]
23. Harris MT, Rogers WA. Developing a Healthcare Technology Acceptance Model (H-TAM) for older adults with hypertension. Ageing Soc 2023 Apr;43(4):814-834. [doi: 10.1017/s0144686x21001069] [Medline: 37007645]
24. Huang D, Chen Q, Huang S, Liu X, Huang S. Consumer intention to use service robots: a cognitive–affective–conative framework. Int J Contemp Hospitality Manage 2024 Apr 29;36(6):1893-1913. [doi: 10.1108/IJCHM-12-2022-1528]
25. Gao Q, Sun Q. Examining the usability of touch screen gestures for older and younger adults. Hum Factors 2015 Aug;57(5):835-863. [doi: 10.1177/0018720815581293] [Medline: 25957042]
26. Liu M, Wang C, Hu J. Older adults' intention to use voice assistants: usability and emotional needs. Heliyon 2023 Nov;9(11):e21932. [doi: 10.1016/j.heliyon.2023.e21932]
27. Nault E, Baillie L, Broz F. Investigating the usability of a socially assistive robotic cognitive training task with augmented sensory feedback modalities for older adults. Presented at: 2022 31st IEEE International Conference on Robot and Human

XSL•FO
RenderX

Interactive Communication (RO-MAN); Aug 29 to Sep 2, 2023; Napoli, Italy p. 735-742. [doi: 10.1109/RO-MAN53752.2022.9900825]
28. Warnock D, McGee-Lennon M, Brewster S. Multiple notification modalities and older users. Presented at: CHI '13: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Apr 27 to May 2, 2013; Paris, France p. 1091-1094. [doi: 10.1145/2470654.2466139]
29. Jacko J, Emery VK, Edwards PJ, et al. The effects of multimodal feedback on older adults' task performance given varying levels of computer experience. Behav Inf Technol 2004 Jul;23(4):247-264. [doi: 10.1080/01449290310001659213]
30. Mueller KD, Hermann B, Mecollari J, Turkstra LS. Connected speech and language in mild cognitive impairment and Alzheimer's disease: a review of picture description tasks. J Clin Exp Neuropsychol 2018 Nov;40(9):917-939. [doi: 10.1080/13803395.2018.1446513] [Medline: 29669461]
31. Jakob D. Voice controlled devices and older adults – a systematic literature review. In: Human Aspects of IT for the Aged Population Design, Interaction and Technology Acceptance: Springer; 2022:175-200. [doi: 10.1007/978-3-031-05581-2_14]
32. Pednekar S, Dhirawani P, Shah R, Shekokar N, Ghag K. Voice-based interaction for an aging population: a systematic review. Presented at: 2023 3rd International Conference on Intelligent Communication and Computational Techniques (ICCT); Jan 19-20, 2023; Jaipur, India p. 1-8. [doi: 10.1109/ICCT56969.2023.10075801]
33. Paré G, Tate M, Johnstone D, Kitsiou S. Contextualizing the twin concepts of systematicity and transparency in information systems literature reviews. Eur J Inf Syst 2016 Nov;25(6):493-508. [doi: 10.1057/s41303-016-0020-3]
34. Paré G, Trudel MC, Jaana M, Kitsiou S. Synthesizing information systems knowledge: a typology of literature reviews. Inf Manag 2015 Mar;52(2):183-199. [doi: 10.1016/j.im.2014.08.008]
35. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
36. Schmidt S, Isbel S, John B, Subramanian R, D'Cunha N. Opinions of older people with mild cognitive impairment (MCI) about technology solutions, preferred devices, uses and interaction modality. Open Science Framework. 2024 May 4.
37. Mitterfellner R, D'Cunha NM, Isbel S. Occupation-based interventions to improve occupational performance among older adults living in long-term care: a systematic review. Am J Occup Ther 2024 Jan 1;78(1):7801205140. [doi: 10.5014/ajot.2024.050441] [Medline: 38231082]
38. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol 2008 Jul 10;8(1):45. [doi: 10.1186/1471-2288-8-45] [Medline: 18616818]
39. Afifi T, Collins N, Rand K, et al. Using virtual reality to improve the quality of life of older adults with cognitive impairments and their family members who live at a distance. Health Commun 2023 Oct;38(9):1904-1915. [doi: 10.1080/10410236.2022.2040170] [Medline: 35253531]
40. Bartels SL, van Knippenberg RJM, Malinowsky C, Verhey FRJ, de Vugt ME. Smartphone-based experience sampling in people with mild cognitive impairment: feasibility and usability study. JMIR Aging 2020 Oct 16;3(2):e19852. [doi: 10.2196/19852] [Medline: 33064084]
41. Beentjes KM, Kerkhof YJF, Neal DP, et al. Process evaluation of the FindMyApps program trial among people with dementia or MCI and their caregivers based on the MRC guidance. Gerontechnology 2020;20(1):1-15. [doi: 10.4017/gt.2020.20.406.11]
42. Bernini S, Panzarasa S, Quaglini S, et al. HomeCoRe system for telerehabilitation in individuals at risk of dementia: a usability and user experience study. Front Med (Lausanne) 2023;10:1129914. [doi: 10.3389/fmed.2023.1129914] [Medline: 36873886]
43. Bernini S, Ballante E, Fassio F, et al. In person versus remote cognitive rehabilitation in patients with subjective cognitive decline or neurocognitive disorders: what factors drive patient's preference? Front Psychol 2023;14. [doi: 10.3389/fpsyg.2023.1266314]
44. Bogza LM, Patry-Lebeau C, Farmanova E, et al. User-centered design and evaluation of a web-based decision aid for older adults living with mild cognitive impairment and their health care providers: mixed methods study. J Med Internet Res 2020 Aug 19;22(8):e17406. [doi: 10.2196/17406] [Medline: 32442151]
45. Bouzida A, Kubota A, Cruz-Sandoval D, Twamley EW, Riek LD. CARMEN: a cognitively assistive robot for personalized neurorehabilitation at home. Presented at: HRI '24: Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction; Mar 11-15, 2024; Boulder, CO p. 55-64. [doi: 10.1145/3610977.3634971]
46. Chang CH, Yeh CH, Chang CC, Lin YC. Interactive somatosensory games in rehabilitation training for older adults with mild cognitive impairment: usability study. JMIR Serious Games 2022 Jul 14;10(3):e38465. [doi: 10.2196/38465] [Medline: 35834303]
47. Chen K, Lou VWQ, Lo SSC. Exploring the acceptance of tablets usage for cognitive training among older people with cognitive impairments: a mixed-methods study. Appl Ergon 2021 May;93:103381. [doi: 10.1016/j.apergo.2021.103381] [Medline: 33578065]
48. Christiansen L, Lindberg C, Sanmartin Berglund J, Anderberg P, Skär L. Using mobile health and the impact on health-related quality of life: perceptions of older adults with cognitive impairment. Int J Environ Res Public Health 2020 Apr 13;17(8):2650. [doi: 10.3390/ijerph17082650] [Medline: 32294911]

XSL•FO
RenderX

49. Collette B, Shotwell J, Shepherd A, Shepherd K, Renzi-Hammond LM, Beer JM. Developing an integrated virtual assistant (IVA): feasibility of a behavior tracking and reminder prototype solution to assist persons with dementia and their care partners. Proc Hum Factors Ergon Soc Annu Meet 2021 Sep;65(1):16-20. [doi: 10.1177/1071181321651279]
50. Contreras-Somoza LM, Toribio-Guzmán JM, Sánchez-Gómez MC, et al. Perceptions of older people with cognitive impairment, caregivers, and professionals about ehcoBUTLER (tablet health care platform): a qualitative focus group study. Int J Environ Res Public Health 2022 Jun 1;19(11):6761. [doi: 10.3390/ijerph19116761] [Medline: 35682344]
51. Contreras-Somoza LM, Irazoki E, Castilla D, et al. Study on the acceptability of an ICT platform for older adults with mild cognitive impairment. J Med Syst 2020 May 25;44(7):120. [doi: 10.1007/s10916-020-01566-x] [Medline: 32451740]
52. Cunnah K, Howe D, Thorpe J, et al. Training people with dementia/cognitive impairment and their carers in the use of web-based supportive technologies (Innovative practice). Dementia (London) 2021 Feb;20(2):796-806. [doi: 10.1177/1471301219887592]
53. Demiris G, Thompson HJ, Lazar A, Lin SY. Evaluation of a Digital Companion for Older Adults with Mild Cognitive Impairment. AMIA Annu Symp Proc 2016;2016:496-503. [Medline: 28269845]
54. Dixon E, Anderson J, Lazar A. Understanding how sensory changes experienced by individuals with a range of age-related cognitive changes can affect technology use. ACM Trans Access Comput 2022 Jun 30;15(2):1-33. [doi: 10.1145/3511906] [Medline: 35919105]
55. Franco-Martín MA, Diaz-Baquero AA, Bueno-Aguado Y, et al. Computer-based cognitive rehabilitation program GRADIOR for mild dementia and mild cognitive impairment: new features. BMC Med Inform Decis Mak 2020 Oct 22;20(1):274. [doi: 10.1186/s12911-020-01293-w] [Medline: 33092577]
56. Gasteiger N, Ahn HS, Lee C, et al. Participatory design, development, and testing of assistive health robots with older adults: an international four-year project. J Hum-Robot Interact 2022 Dec 31;11(4):1-19. [doi: 10.1145/3533726]
57. Gelonch O, Ribera M, Codern-Bové N, et al. Acceptability of a lifelogging wearable camera in older adults with mild cognitive impairment: a mixed-method study. BMC Geriatr 2019 Apr 16;19(1):110. [doi: 10.1186/s12877-019-1132-0] [Medline: 30991948]
58. Givon Schaham N, Vitek H, Donda N, Elbo Golan I, Buckman Z, Rand D. The development and feasibility of TECH: Tablet Enhancement of Cognition and Health, a novel cognitive intervention for people with mild cognitive impairment. Games Health J 2020 Oct 1;9(5):346-352. [doi: 10.1089/g4h.2019.0157] [Medline: 33054487]
59. Guzman-Parra J, Barnestein-Fonseca P, Guerrero-Pertiñez G, et al. Attitudes and use of information and communication technologies in older adults with mild cognitive impairment or early stages of dementia and their caregivers: cross-sectional study. J Med Internet Res 2020 Jun 1;22(6):e17253. [doi: 10.2196/17253] [Medline: 32442136]
60. Haesner M, O'Sullivan JL, Gövercin M, Steinhagen-Thiessen E. Requirements of older adults for a daily use of an internet-based cognitive training platform. Inform Health Soc Care 2015 Mar;40(2):139-153. [doi: 10.3109/17538157.2013.879149] [Medline: 24725153]
61. Hassandra M, Galanis E, Hatzigeorgiadis A, et al. A virtual reality app for physical and cognitive training of older people with mild cognitive impairment: mixed methods feasibility study. JMIR Serious Games 2021 Mar 24;9(1):e24170. [doi: 10.2196/24170] [Medline: 33759797]
62. Heatwole Shank KS. "You Know, I Swipe My Card and Hope for the Best": technology and cognition as dual landscapes of change. Gerontol Geriatr Med 2022 Jan;8:23337214221128402. [doi: 10.1177/23337214221128402] [Medline: 36203744]
63. Hedman A, Lindqvist E, Nygård L. How older adults with mild cognitive impairment relate to technology as part of present and future everyday life: a qualitative study. BMC Geriatr 2016 Mar 31;16(1):73. [doi: 10.1186/s12877-016-0245-y] [Medline: 27036538]
64. Hedman A, Kottorp A, Almkvist O, Nygård L. Challenge levels of everyday technologies as perceived over five years by older adults with mild cognitive impairment. Int Psychogeriatr 2018 Oct;30(10):1447-1454. [doi: 10.1017/S1041610218000285] [Medline: 29615145]
65. Horn BL, Albers EA, Mitchell LL, et al. Can technology-based social memory aids improve social engagement? Perceptions of a novel memory aid for persons with memory concerns. J Appl Gerontol 2023 Mar;42(3):399-408. [doi: 10.1177/07334648221134869] [Medline: 36274581]
66. Hu HJX, Chen LH, Liu YC. Simplicity is not always best. Presented at: ISS '18: Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces; Nov 25-28, 2018; Tokyo, Japan p. 375-380. [doi: 10.1145/3279778.3279910]
67. Infarinato F, Jansen-Kosterink S, Romano P, et al. Acceptance and potential impact of the eWALL platform for health monitoring and promotion in persons with a chronic disease or age-related impairment. Int J Environ Res Public Health 2020 Oct 28;17(21):7893. [doi: 10.3390/ijerph17217893] [Medline: 33126506]
68. Irazoki E, Sánchez-Gómez MC, Contreras-Somoza LM, et al. A qualitative study of the cognitive rehabilitation program GRADIOR for people with cognitive impairment: outcomes of the focus group methodology. J Clin Med 2021 Feb 19;10(4):859. [doi: 10.3390/jcm10040859] [Medline: 33669716]
69. Korchut A, Szklener S, Abdelnour C, et al. Challenges for service robots—requirements of elderly adults with cognitive impairments. Front Neurol 2017;8:228. [doi: 10.3389/fneur.2017.00228] [Medline: 28620342]
70. Kubota A, Cruz‐Sandoval D, Kim S, Twamley EW, Riek LD. Cognitively assistive robots at home: translating clinical interventions to robots. Alzheimers Dement 2022 Dec;18(S2). [doi: 10.1002/alz.062600]

XSL•FO
RenderX

71. LaMonica HM, English A, Hickie IB, et al. Examining internet and eHealth practices and preferences: survey study of Australian older adults with subjective memory complaints, mild cognitive impairment, or dementia. J Med Internet Res 2017 Oct 25;19(10):e358. [doi: 10.2196/jmir.7981] [Medline: 29070481]
72. Law M, Ahn HS, MacDonald B, et al. User testing of cognitive training games for people with mild cognitive impairment: design implications. In: Social Robotics: Springer; 2019, Vol. 2019:464-473. [doi: 10.1007/978-3-030-35888-4_43]
73. Lazarou I, Stavropoulos TG, Mpaltadoros L, et al. Human factors and requirements of people with cognitive impairment, their caregivers, and healthcare professionals for mHealth apps including reminders, games, and geolocation tracking: a survey-questionnaire study. J Alzheimers Dis Rep 2021;5(1):497-513. [doi: 10.3233/ADR-201001] [Medline: 34368634]
74. Leese MI, Dorociak KE, Noland M, Gaugler JE, Mattek N, Hughes A. Use of in-home activity monitoring technologies in older adult veterans with mild cognitive impairment: the impact of attitudes and cognition. Gerontechnology 2021 Jun;20(2):1-12. [doi: 10.4017/gt.2021.20.2.10.06] [Medline: 34305492]
75. Lindqvist E, PerssonVasiliou A, Hwang AS, et al. The contrasting role of technology as both supportive and hindering in the everyday lives of people with mild cognitive deficits: a focus group study. BMC Geriatr 2018 Aug 17;18(1):185. [doi: 10.1186/s12877-018-0879-z] [Medline: 30119653]
76. Madjaroff G, Mentis H. Narratives of older adults with mild cognitive impairment and their caregivers. Presented at: ASSETS '17: Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility; Oct 20 to Nov 1, 2017; Baltimore, MD p. 140-149. [doi: 10.1145/3132525.3132554]
77. Maier AM, Özkil AG, Bang MM, Forchhammer BH. Remember to remember: a feasibility study adapting wearable technology to the needs of people aged 65 and older with mild cognitive impairment (MCI) and Alzheimer's dementia. Presented at: Proceedings of the 20th International Conference on Engineering Design (ICED15); Jul 27-30, 2015; Milan, Italy URL: https://backend.orbit.dtu.dk/ws/portalfiles/portal/107110242/Remember_to_remember.pdf [accessed 2025-10-22]
78. Manca M, Paternò F, Santoro C, et al. The impact of serious games with humanoid robots on mild cognitive impairment older adults. Int J Hum Comput Stud 2021 Jan;145:102509. [doi: 10.1016/j.ijhcs.2020.102509]
79. Manera V, Chapoulie E, Bourgeois J, et al. A feasibility study with image-based rendered virtual reality in patients with mild cognitive impairment and dementia. PLoS ONE 2016;11(3):e0151487. [doi: 10.1371/journal.pone.0151487] [Medline: 26990298]
80. Mathur N, Dhodapkar K, Zubatiy T, Li J, Jones B, Mynatt E. A collaborative approach to support medication management in older adults with mild cognitive impairment using conversational assistants (CAs). Presented at: ASSETS '22: Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility; Oct 23-26, 2022; Athens, Greece p. 1-14. [doi: 10.1145/3517428.3544830]
81. Matsangidou M, Solomou T, Frangoudes F, et al. Affective out-world experience via virtual reality for older adults living with mild cognitive impairments or mild dementia. Int J Environ Res Public Health 2023 Feb 7;20(4):2919. [doi: 10.3390/ijerph20042919] [Medline: 36833616]
82. Mattos MK, Manning CA, Quigg M, et al. Feasibility and preliminary efficacy of an internet-delivered intervention for insomnia in individuals with mild cognitive impairment. J Alzheimers Dis 2021;84(4):1539-1550. [doi: 10.3233/JAD-210657] [Medline: 34690141]
83. McCarron HR, Zmora R, Gaugler JE. A web-based mobile app with a smartwatch to support social engagement in persons with memory loss: pilot randomized controlled trial. JMIR Aging 2019 Jun 18;2(1):e13378. [doi: 10.2196/13378]
84. Mehrabian S, Extra J, Wu YH, Pino M, Traykov L, Rigaud AS. The perceptions of cognitively impaired patients and their caregivers of a home telecare system. Med Devices (Auckl) 2015;8:21-29. [doi: 10.2147/MDER.S70520] [Medline: 25552909]
85. Mondellini M, Arlati S, Gapeyeva H, et al. User experience during an immersive virtual reality-based cognitive task: a comparison between Estonian and Italian older adults with MCI. Sensors (Basel) 2022 Oct 27;22(21):8249. [doi: 10.3390/s22218249] [Medline: 36365947]
86. Moro C, Lin S, Nejat G, Mihailidis A. Social robots and seniors: a comparative study on the influence of dynamic social features on human–robot interaction. Int J Soc Robotics 2019 Jan;11(1):5-24. [doi: 10.1007/s12369-018-0488-1]
87. Mrakic-Sposta S, Di Santo SG, Franchini F, et al. Effects of combined physical and cognitive virtual reality-based training on cognitive impairment and oxidative stress in MCI patients: a pilot study. Front Aging Neurosci 2018 Oct 1;10:282. [doi: 10.3389/fnagi.2018.00282] [Medline: 30327596]
88. Nie Q, Nguyen LT, Myers D, et al. Design guidance for video chat system to support social engagement for older adults with and without mild cognitive impairment. Gerontechnology 2020 Aug 29;20(1):3024. [Medline: 37904899]
89. Nieto-Vieites A, Mallo SC, Pérez-Blanco L, et al. A narrative video game for adults with subjective and objective cognitive impairment. Design and preliminary results on user-interaction and efficacy. Behav Inf Technol 2024 Jun 10;43(8):1518-1530. [doi: 10.1080/0144929X.2023.2220042]
90. Ortega Morán JF, Pagador JB, Gilete Preciado V, et al. A serious game for cognitive stimulation of older people with mild cognitive impairment: design and pilot usability study. JMIR Aging 2024 Apr 4;7:e41437. [doi: 10.2196/41437] [Medline: 38596860]
91. Park C, Mishra RK, York MK, et al. Tele-medicine based and self-administered interactive exercise program (Tele-Exergame) to improve cognition in older adults with mild cognitive impairment or dementia: a feasibility, acceptability, and

XSL•FO
RenderX

proof-of-concept study. Int J Environ Res Public Health 2022 Dec 6;19(23):16361. [doi: 10.3390/ijerph192316361] [Medline: 36498431]
92. Park JH, Liao Y, Kim DR, et al. Feasibility and tolerability of a culture-based virtual reality (VR) training program in patients with mild cognitive impairment: a randomized controlled pilot study. Int J Environ Res Public Health 2020 Apr 27;17(9):3030. [doi: 10.3390/ijerph17093030] [Medline: 32349413]
93. Piasek J, Wieczorowska-Tobis K. Acceptance and long-term use of a social robot by elderly users in a domestic environment. Presented at: 2018 11th International Conference on Human System Interaction (HSI); Jul 4-6, 2018; Gdansk, Poland p. 478-482. [doi: 10.1109/HSI.2018.8431348]
94. Pino M, Boulay M, Jouen F, Rigaud AS. "Are we ready for robots that care for us?" Attitudes and opinions of older adults toward socially assistive robots. Front Aging Neurosci 2015;7:141. [doi: 10.3389/fnagi.2015.00141] [Medline: 26257646]
95. Quintana M, Anderberg P, Sanmartin Berglund J, et al. Feasibility-usability study of a tablet app adapted specifically for persons with cognitive impairment-SMART4MD (Support Monitoring and Reminder Technology for Mild Dementia). Int J Environ Res Public Health 2020 Sep 18;17(18):6816. [doi: 10.3390/ijerph17186816] [Medline: 32961993]
96. Rossi S, Di Napoli C, Garramone F, Salvatore E, Santangelo G. Personality-based adaptation of robot behaviour: acceptability results on individuals with cognitive impairments. Int J Soc Robotics 2024 Jan;16(1):211-226. [doi: 10.1007/s12369-023-01074-1]
97. Saini J, Owen A, Hammer A, Brown W, Kulshreshtha A. Zooming towards better health: participant experience in a web-based group cognitive behavioral intervention. iProc 2018;4(2):e11773. [doi: 10.2196/11773]
98. Scase M, Kreiner K, Ascolese A. Development and evaluation of cognitive games to promote health and wellbeing in elderly people with mild cognitive impairment. Stud Health Technol Inform 2018;248:255-262. [Medline: 29726445]
99. Scheibe M, Lang C, Druschke D, et al. Independent use of a home-based telemonitoring app by older patients with multimorbidity and mild cognitive impairment: qualitative study. JMIR Hum Factors 2021 Jul 12;8(3):e27156. [doi: 10.2196/27156] [Medline: 34255664]
100. Shamir D, Loubani K, Schaham NG, Buckman Z, Rand D. Experiences of older adults with mild cognitive impairment from cognitive self-training using touchscreen tablets. Games Health J 2024 Feb;13(1):13-24. [doi: 10.1089/g4h.2023.0017] [Medline: 37768834]
101. Shellington EM, Felfeli T, Shigematsu R, Gill DP, Petrella RJ. HealtheBrain: an innovative smartphone application to improve cognitive function in older adults. Mhealth 2017 May 5;3:17. [doi: 10.21037/mhealth.2017.04.05] [Medline: 28567413]
102. Shin MH, McLaren J, Ramsey A, Sullivan JL, Moo L. Improving a mobile telepresence robot for people with Alzheimer disease and related dementias: semistructured interviews with stakeholders. JMIR Aging 2022 May 3;5(2):e32322. [doi: 10.2196/32322] [Medline: 35503518]
103. Stogl D, Armbruster O, Mende M, Hein B, Wang X, Meyer P. Robot-based training for people with mild cognitive impairment. IEEE Robot Autom Lett 2019;4(2):1916-1923. [doi: 10.1109/LRA.2019.2898470]
104. Stramba-Badiale C, Tuena C, Goulene KM, et al. Enhancing spatial navigation skills in mild cognitive impairment patients: a usability study of a new version of ANTaging software. Front Hum Neurosci 2024;17. [doi: 10.3389/fnhum.2023.1310375]
105. Tuena C, Serino S, Stramba-Badiale C, et al. Usability of an embodied CAVE system for spatial navigation training in mild cognitive impairment. J Clin Med 2023 Mar 1;12(5):1949. [doi: 10.3390/jcm12051949] [Medline: 36902733]
106. Van Assche M, Petrovic M, Cambier D, Calders P, Van Gelder P, Van de Velde D. The perspectives of older adults with mild cognitive impairment and their caregivers on the use of socially assistive robots in healthcare: exploring factors that influence attitude in a pre-implementation stage. Disabil Rehabil Assist Technol 2024 Jan;19(1):222-232. [doi: 10.1080/17483107.2022.2075477] [Medline: 35587020]
107. Votis K, Giakoumis D, Vasileiadis M, Doumpoulakis S, Segkoyli S, Tzovaras D. Mobile cognitive training games for older adults with mild cognitive impairment. Presented at: MobileHCI '15: Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct; Aug 24-27, 2015; Copenhagen, Denmark p. 932-935. [doi: 10.1145/2786567.2794307]
108. Wargnier P, Benveniste S, Jouvelot P, Rigaud AS. Usability assessment of interaction management support in LOUISE, an ECA-based user interface for elders with cognitive impairment. Technol Disabil 2018 Aug;30(3):105-126. [doi: 10.3233/TAD-180189]
109. Weering M, Jansen-Kosterink S, Tabak M. The implementation, actual use and user experience of an online home exercise program that fits the needs of older adults with mild cognitive impairments. Presented at: 5th International Conference on Information and Communication Technologies for Ageing Well and e-Health; May 2-4, 2019; Heraklion, Greece p. 189-196. [doi: 10.5220/0007696801890196]
110. Wolf D, Besserer D, Sejunaite K, Schuler A, Riepe M, Rukzio E. A demonstration of care. Presented at: MUM '19: Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia; Nov 26-29, 2019; Pisa, Italy p. 1-4. [doi: 10.1145/3365610.3368472]
111. Wu YH, Cristancho-Lacroix V, Fassert C, Faucounau V, de Rotrou J, Rigaud AS. The attitudes and perceptions of older adults with mild cognitive impairment toward an assistive robot. J Appl Gerontol 2016 Jan;35(1):3-17. [doi: 10.1177/0733464813515092] [Medline: 24652924]

XSL•FO
RenderX

112. Wu YH, Wrobel J, Cornuet M, Kerhervé H, Damnée S, Rigaud AS. Acceptance of an assistive robot in older adults: a mixed-method study of human-robot interaction over a 1-month period in the Living Lab setting. Clin Interv Aging 2014;9:801-811. [doi: 10.2147/CIA.S56435] [Medline: 24855349]
113. Yamazaki R, Nishio S, Nagata Y, et al. A preliminary study of robotic media effects on older adults with mild cognitive impairment in solitude. In: Social Robotics: Springer; 2021, Vol. 2021:453-463. [doi: 10.1007/978-3-030-90525-5_39]
114. Yun SJ, Kang MG, Yang D, et al. Cognitive training using fully immersive, enriched environment virtual reality for patients with mild cognitive impairment and mild dementia: feasibility and usability study. JMIR Serious Games 2020 Oct 14;8(4):e18127. [doi: 10.2196/18127] [Medline: 33052115]
115. Yurkewich A, Stern A, Alam R, Baecker R. A Field Study of Older Adults with Cognitive Impairment using Tablets for Communication at Home. Int J Mob Hum Comput Interact 2018;10(2):1-30. [doi: 10.4018/IJMHCI.2018040101]
116. Zafeiridi P, Paulson K, Dunn R, et al. A web-based platform for people with memory problems and their caregivers (CAREGIVERSPRO-MMD): mixed-methods evaluation of usability. JMIR Form Res 2018 Mar 12;2(1):e4. [doi: 10.2196/formative.9083] [Medline: 30684403]
117. Zedda E, Manca M, Paterno F, Santoro C. MCI older adults' user experience with introverted and extraverted humanoid robot personalities. Presented at: Proceedings of the 15th Biannual Conference of the Italian SIGCHI Chapter; Sep 20-23, 2023; Turin, Italy p. 1-12. [doi: 10.1145/3605390.3605405]
118. Zhang BY, Gao YM. Evaluation of older adults' preference factors for serious games. Presented at: 2023 9th International Conference on Control, Automation and Robotics (ICCAR); Apr 21-23, 2023; Beijing, China p. 324-329. [doi: 10.1109/ICCAR57134.2023.10151718]
119. Zhang Q, Liu Y. Home-based dual task training for older people with mild cognitive impairment. Presented at: Chinese CHI '22: Proceedings of the Tenth International Symposium of Chinese CHI; Oct 22-23, 2022; Guangzhou, China and online p. 196-202. [doi: 10.1145/3565698.3565784]
120. Zhu D, Al Mahmud A, Liu W. Digital storytelling intervention for enhancing the social participation of people with mild cognitive impairment: co-design and usability study. JMIR Aging 2024 Jan 17;7:e54138. [doi: 10.2196/54138] [Medline: 38231541]
121. Zubatiy T, Vickers KL, Mathur N, Mynatt ED. Empowering dyads of older adults with mild cognitive impairment and their care partners using conversational agents. Presented at: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems; May 8-13, 2021; Yokohama, Japan p. 1-15. [doi: 10.1145/3411764.3445124]
122. Motti L, Nadine V, Gorce P. Interaction techniques for older adults using touchscreen devices: a literature review from 2000 to 2013. Journal d'Interaction Personne-Système 2014;3(2). [doi: 10.46298/jips.3180]
123. Lee JH, Poliakoff E, Spence C. The effect of multimodal feedback presented via a touch screen on the performance of older adults. In: Haptic and Audio Interaction Design: Springer; 2009:128-135. [doi: 10.1007/978-3-642-04076-4_14]
124. Jian C, Shi H, Sasse N, et al. Modality preference in multimodal interaction for elderly persons. In: Biomedical Engineering Systems and Technologies: Springer; 2014:378-393. [doi: 10.1007/978-3-662-44485-6_26]
125. Brooke J. SUS: a quick and dirty usability scale. In: Usability Evaluation in Industry: CRC Press; 1996. [doi: 10.1201/9781498710411]
126. Hyzy M, Bond R, Mulvenna M, et al. System Usability Scale benchmarking for digital health apps: meta-analysis. JMIR Mhealth Uhealth 2022 Aug 18;10(8):e37290. [doi: 10.2196/37290] [Medline: 35980732]
127. Heerink M, Kröse B, Evers V, Wielinga B. Assessing acceptance of assistive social agent technology by older adults: the Almere model. Int J of Soc Robotics 2010 Dec;2(4):361-375. [doi: 10.1007/s12369-010-0068-5]
128. Nielsen J. 10 usability heuristics for user interface design. Nielsen Norman Group. 2024 Jan 30. URL: https://www.nngroup.com/articles/ten-usability-heuristics/ [accessed 2024-12-17]
129. Shneiderman B. Human-Centered AI: Oxford University Press; 2022. [doi: 10.1093/oso/9780192845290.001.0001]
130. Pitale A, Bhumgara A. Human computer interaction strategies — designing the user interface. Presented at: 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT); Nov 27-29, 2019; Tirunelveli, India p. 752-758. [doi: 10.1109/ICSSIT46314.2019.8987819]
131. Yuan F, Anderson JG, Wyatt TH, et al. Assessing the acceptability of a humanoid robot for Alzheimer's disease and related dementia care using an online survey. Int J Soc Robotics 2022 Jul;14(5):1223-1237. [doi: 10.1007/s12369-021-00862-x]
132. Holthe T, Halvorsrud L, Karterud D, Hoel KA, Lund A. Usability and acceptability of technology for community-dwelling older adults with mild cognitive impairment and dementia: a systematic literature review. Clin Interv Aging 2018;13:863-886. [doi: 10.2147/CIA.S154717] [Medline: 29765211]
133. Kan John P, Zhu X, Gedeon T, Zhu W. Evaluating human impressions of an initiative-taking robot. Presented at: CHI EA '22: Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems; Apr 29 to May 5, 2022; New Orleans, LA p. 1-7. [doi: 10.1145/3491101.3519633]
134. Müller S, Mychajliw C, Reichert C, Melcher T, Leyhe T. Autobiographical memory performance in Alzheimer's disease depends on retrieval frequency. J Alzheimer's Dis 2016 Jun 7;52(4):1215-1225. [doi: 10.3233/JAD-151071]
135. Common Voice. URL: https://commonvoice.mozilla.org/en [accessed 2024-12-18]
136. Reid K. Mozilla Common Voice v19 dataset metadata coverage. Observable. 2024. URL: https://observablehq.com/@kathyreid/mozilla-common-voice-v19-dataset-metadata-coverage [accessed 2025-10-17]

XSL•FO
RenderX

137. Proctor RW, Vu KPL. Human information processing: an overview for human-computer interaction. In: The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Application, 3rd edition: Taylor and Francis Group; 2012.
138. Liu H, Wang W, Liu Y, Song F, Wang S, Guo H. Study on the differences of icon cognition of graphical interface for age-friendly design. J Gerontol Soc Work 2023 Jul;66(5):662-679. [doi: 10.1080/01634372.2022.2139319] [Medline: 36309977]
139. Md Fadzil NH, Shahar S, Singh DKA, et al. Mapping the landscape: a bibliometric analysis of information and communication technology adoption by older adults with cognitive frailty or impairment. Geriatrics Gerontology Int 2024 Mar;24(3):251-262. [doi: 10.1111/ggi.14814]
140. Nastjuk I, Trang S, Grummeck-Braamt JV, Adam MTP, Tarafdar M. Integrating and synthesising technostress research: a meta-analysis on technostress creators, outcomes, and IS usage contexts. Eur J Inf Syst 2024 May 3;33(3):361-382. [doi: 10.1080/0960085X.2022.2154712]
141. Kolkowska E, Soja E. Attitudes towards ICT solutions for independent living among older adults in Sweden and Poland: a preliminary study. Presented at: International Conference on ICT Management for Global Competitiveness and Economic Growth in Emerging Economies; Oct 23-24, 2017; Wrocław, Poland p. 36-46 URL: https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=3140412#page=40 [accessed 2025-10-22]
142. Kling R. The organizational context of user-centered software designs. MIS Q 1977 Dec;1(4):41. [doi: 10.2307/249021]
143. Köhler S, Perry J, Biernetzky OA, Kirste T, Teipel SJ. Ethics, design, and implementation criteria of digital assistive technologies for people with dementia from a multiple stakeholder perspective: a qualitative study. BMC Med Ethics 2024 Jul 27;25(1):84. [doi: 10.1186/s12910-024-01080-6] [Medline: 39068472]
144. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
145. Folstein MF, Folstein SE, McHugh PR. “Mini-mental state”. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6]
146. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment. Arch Neurol 1999 Mar 1;56(3):303. [doi: 10.1001/archneur.56.3.303]
147. Stara V, Soraci L, Takano E, et al. Intrinsic capacity and active and healthy aging domains supported by personalized digital coaching: survey study among geriatricians in Europe and Japan on eHealth opportunities for older adults. J Med Internet Res 2023 Oct 12;25:e41035. [doi: 10.2196/41035] [Medline: 37824183]
148. Research Training Program. Australian Government: Department of Education. URL: https://www.education.gov.au/research-block-grants/research-training-program [accessed 2025-10-22]

Abbreviations

CBT: cognitive behavioral therapy
H-TAM: healthcare technology acceptance model
HCI: human-computer interaction
MATOA: model for the adoption of technology by older adults
MCI: mild cognitive impairment
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
STAM: senior technology acceptance model
TA: technology adoption
TAM: technology acceptance model
UCD: user-centered design
UTAUT: unified theory of acceptance and use of technology
UX: user experience
VR: virtual reality

Edited by H LaMonica; submitted 30.05.25; peer-reviewed by E Maranesi, J Farias, L Haslam-Larmer; revised version received 04.09.25; accepted 25.09.25; published 30.10.25.

Please cite as:
Schmidt SB, Isbel S, John B, Subramanian R, D'Cunha NM
Examining Technology Perspectives of Older Adults With Mild Cognitive Impairment: Scoping Review
JMIR Aging 2025;8:e78229
URL: https://aging.jmir.org/2025/1/e78229
doi:10.2196/78229

XSL•FO
RenderX

© Snezna Bizilj Schmidt, Stephen Isbel, Blooma John, Ramanathan Subramanian, Nathan Martin D'Cunha. Originally published in JMIR Aging (https://aging.jmir.org), 30.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Patient and Carer-Related Facilitators and Barriers to the Adoption of Assistive Technologies for the Care of Older Adults: Systematic Review

Stephen Malden[1], MPH, PhD; Kris McGill[2], PhD; Bruce Guthrie[2], PhD; Helen Frost[2], PhD; Susan D Shenkin[2,3], MD; Adanna Ezike[4], MBBS; Bethany Kate Bareham[2,5], PhD; Stewart W Mercer[2], PhD; Caroline Pearce[2], PhD; Cara Wilson[2], PhD; Ian Underwood[2,6], PhD; John Vines[2], PhD; Sue Lewis[2], PhD; Amy O'Donnell[2,5], PhD

[1]Scottish Collaboration for Public Health Research and Policy, School of Health in Social Science, University of Edinburgh, Old Medical School, Teviot Place, Edinburgh, United Kingdom

[2]Advanced Care Research Centre, Usher Institute, University of Edinburgh, Edinburgh, United Kingdom

[3]Ageing and Health, Usher Institute, University of Edinburgh, Edinburgh, United Kingdom

[4]Forth Valley Royal Hospital, NHS Forth Valley, Larbert, United Kingdom

[5]Population Health Sciences Institute, Newcastle University, Newcastle, United Kingdom

[6]School of Engineering, The University of Edinburgh, Edinburgh, United Kingdom

Corresponding Author:
Stephen Malden, MPH, PhD
Scottish Collaboration for Public Health Research and Policy, School of Health in Social Science, University of Edinburgh, Old Medical School, Teviot Place, Edinburgh, United Kingdom

Abstract

Background: Assistive technologies (ATs) are used increasingly in community settings to assist in the care of older adults. Despite a rapid increase in the capabilities and uptake of these technologies, gaps remain in understanding the main barriers to their usage.

Objective: This systematic review investigated the barriers and facilitators to the use of AT in the care of older adults.

Methods: Six electronic databases were searched from January 2011 to March 2024. Primary studies were included if they used qualitative methods reporting findings related to barriers or facilitators to the implementation of AT (eg, ambient and wearable sensors, alarms, telehealth or mobile health [mHealth]) for older adults (from the perspective of either carers or older adults) in community settings. All data were screened independently by two reviewers. Study quality was assessed using the Critical Appraisal Skills Program (CASP). Data from each included study were synthesized using thematic synthesis, before barriers were mapped against the domains of the Technology Acceptance Model (TAM).

Results: Ninety-five studies were included in the review. The number of studies published in the field of barriers to AT use has increased 3-fold post-COVID-19 in comparison to the previous decade. Ten barriers—privacy, cost, insufficient knowledge, fear of misuse, usability, poor functionality, perceived lack of need, stigma, and lack of human interaction—were identified, as well as three facilitators—awareness of health benefits, targeted training, and user-centered design. Persistent barriers relating to all domains of the TAM were identified, with the majority of these relating to the "behavioral intention to use" domain (cost, privacy, stigma, and fear of misuse). The majority of studies had a moderate/high risk of bias.

Conclusions: There remain distinct barriers to sustained usage of AT for the care of older adults, particularly concerning adoption as defined by the TAM. Further studies investigating the acceptability of ATs are needed to increase the understanding of optimization strategies.

Trial Registration: PROSPERO CRD42021266656; https://www.crd.york.ac.uk/PROSPERO/view/CRD42021266656

(JMIR Aging 2025;8:e73917) doi:10.2196/73917

KEYWORDS

older adult; assistive technology; care; aging in place; barriers; facilitators

Introduction

Increased life expectancy in many regions around the world has not been matched by improvements in health span (average length of disease-free healthy life) [1]. This has in turn increased the need for health and social care in older populations [2]. Assistive technologies (ATs) could potentially help enable older adults to live independently for longer, which is most people's

XSL•FO
RenderX

preference, and potentially reserve other kinds of more expensive care for those with higher needs. The term ATs covers a range of devices or products designed to support older people to perform activities of daily living such as sensors, wearables, robotics, and information communication technologies (ICT) such as telemedicine [3]. The challenges to in-person delivery of health care experienced during the COVID-19 pandemic led to rapid development and deployment of technology to remotely address gaps in care provision [4].

There are potential benefits of AT for older adults and their caregivers, such as reduced caregiver burden and increased perceived independence [5,6]. However, there is frequently a mismatch between the design, functionality, and usability of AT, on the one hand, and the needs and preferences of end users, on the other hand [7], leading to significant rates of abandonment [8] and barriers to optimal use [7,9-11].

The Technology Acceptance Model (TAM) offers a framework to assess how barriers to AT use may influence sustained adoption [12]. The TAM essentially proposes that an individual's motivation to adopt new technology can be explained by three factors: perceived ease of use, perceived usefulness, and attitude toward using or behavioral intention to use. Although other comprehensive technology frameworks exist (eg, Unified Theory of Acceptance and Use of Technology 2 [UTAUT2]), the advantage of TAM in relation to others is its relative simplicity, which is often preferred by practitioners and patients over complexity [13].

Systematic reviews of technology adoption in the care of older adults have shown that barriers such as cost, unfamiliarity with technology, perceived lack of need, and poor functionality can all impede the sustained adoption of AT [7,13-17]. However, these reviews either investigated just one technology type [15], only included specific populations with chronic conditions [7,13], or were older than 4 years [16,17]. Given the broad range of available AT, its adoption across diverse older adult care groups, and the pace at which it is developing, this review [18] aimed to synthesize the latest published qualitative literature, informed by the TAM. This then allowed for the identification of barriers and facilitators to implementing a wide range of AT to support paid and unpaid care for older adults.

Methods

Protocol Registration

The protocol for this systematic review was prospectively registered on PROSPERO systematic review registry (CRD42021266656) in July 2021 and was reported in accordance with the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) statement checklist [19] and the Transparency in Reporting the Synthesis of Qualitative Research guidelines [20].

Search Strategy and Selection Criteria

A comprehensive search strategy was developed in collaboration with a medical librarian prior to execution in 6 electronic databases (Medline, CINAHL, Embase, PsychINFO, Web of Science, and Cochrane Library) from January 2011 to March 5, 2024. Searches were limited by year to capture the most recent evidence and technological advances. Searches included appropriate keywords and medical subject headings. Database searches were supplemented with searches for gray literature in Google Scholar, OpenGray, and MedRXiv preprint server and forward and backward citation searching (to October 2024), including the reference lists of similar systematic reviews found from Google Scholar. The full search strategy used in Medline can be viewed in Table S9 in Multimedia Appendix 1.

Inclusion and Exclusion Criteria

Primary studies using qualitative methods to understand contextualized personal perspectives of the barriers and facilitators of ATs to support paid and unpaid care for older adults in the community setting were eligible for inclusion. We restricted the search to more recent studies, given rapid changes in AT [21-23]. Barriers and facilitators could be from the perspectives of either older adults or their carers. Studies exclusively containing participants with a single condition were excluded (such as heart failure or diabetes), as a review that only included participants with chronic conditions was published in 2024, which removed the need to synthesize studies in this area [13]. Therefore, this present review only included studies of the general population of older adults receiving care. Given the large volume of studies published in the field of AT for care, this also allowed the review to remain manageable in terms of the number of potentially relevant studies. Similarly, studies exclusively reporting on the use of technology to replace normal care practices during COVID-19 social distancing were also excluded, as these were often temporary and were not planned as a long-term care solution. A detailed description of further inclusion criteria used to select relevant studies is outlined in Textbox 1, while Table 1 describes the definitions applied throughout this review.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria

Inclusion criteria

- Participants were older adults (defined as aged 60 years or older).
- Participants were receiving some form of paid or unpaid care in the community setting (at home receiving domiciliary care or family care, assisted living, care homes, short-term care in the community, eg, rapid response teams to prevent emergency admission, long-term stays in community hospitals).
- The study explores some aspects of either barriers or facilitators of assistive technology usage, or both.
- Technologies explored must be in the context of assisting in the care or independence of older adults (eg, via remote monitoring).

Exclusion criteria

- Studies where participants' mean age was <60 years.
- Participants are in short-term or acute hospitals.
- A study population made up exclusively of older adults with a specific condition (with the exception of dementia).
- Studies using technology explicitly to replace usual care during social isolation periods of the COVID-19 pandemic.
- Quantitative research.
- Secondary research.
- Study focuses on the general use and acceptability of technologies without exploring barriers or facilitators to actual usage.

Table . Definitions used for inclusion or exclusion criteria application.

Criteria name	Descriptor
Paid and unpaid care	Any care provided to older adults by health care professionals, third sector organizations (paid/formal care), and friends and family (unpaid, informal care).
Assistive technologies (AT)	Any devices, products, systems, and services that were installed or utilized within a participant's living environment, community-based care environment, or on their person, with the primary aim to assist in activities of daily living or with the process of care provision through remote monitoring or data generation for care provision.
Facilitators	Factors that increase, improve, promote, facilitate, or enable the use of AT.
Barriers	Factors that prevent, obstruct, or decrease the use of AT.

Screening and Data Extraction

Consistent with the Joanna Briggs Institute Manual for Qualitative Evidence Synthesis [24], title and abstract and full-text screening were undertaken by two independent reviewers using Covidence® systematic review software, with conflicts discussed and resolved by consensus (SM screened all studies, with second screening by either BG, AOD, HF, SDS, SWM, BB, AE, KM, or CP). Data extraction was undertaken using a predeveloped data extraction template covering study identifiers, methods, population, technology, and facilitators and barriers. One reviewer extracted all studies, with a second reviewer checking extraction for accuracy, with any discrepancies discussed and resolved by consensus.

Quality Appraisal

Studies were assessed for quality using the Critical Appraisal Skills Program (CASP) tool for qualitative studies [25], extracting methodological issues that could be sources of bias, rather than assigning scores of quality.

Thematic Synthesis

We adopted the principles of framework analysis to conduct a deductive thematic synthesis [26]. Two researchers with qualitative experience (SM and AOD) independently coded facilitators and barriers from two primary studies to develop a working analytical framework, which was then tested on a further study. After modifications, the remaining studies were coded by the lead author. During the final stage of data interpretation, facilitator and barrier themes were developed by the lead author by looking at the data set as a whole and through discussions with the wider team, in addition to categorizing studies by participant type (ie, carer, older adult, or both). Illustrative quotes relating to each theme are presented in Table S6 in Multimedia Appendix 1 to provide context. Following thematic synthesis, barriers were mapped onto the domains of the TAM (by SM) to help identify aspects of technology adoption which may present greater need for support, or where particular interventions can facilitate usage. Specifically, identified barriers were aligned with the stage of the TAM at which they were determined to present themselves within the literature. For example, cost is a barrier that mainly prevented

XSL•FO
RenderX

participants from uptaking AT in the first place; therefore, this barrier would be mapped to the "behavioral intent to use" TAM domain. While a multitude of technology acceptance frameworks have been developed [27], the TAM was selected as the most appropriate framework for this review for two reasons. First, the TAM was developed using two behavioral theories which directly influence adoption behaviors; namely, theory of reasoned action [28] and theory of planned behavior [29]. Second, the TAM is a simple model consisting of five easily interpreted domains, rendering it a useful framework for a wider audience to comprehend or apply to their own work, thus potentially increasing transferability of the review findings.

Additionally, the proportion of barriers within each TAM domain for papers classified as pre-COVID (ie, primary studies indicate data collection was conducted prior to January 2020) or post-COVID (ie, primary studies indicate in methods that data collection was conducted post January 2020) was calculated. Results were summarized to observe both the volume of papers published in these two periods and any proportional change in TAM domains (represented by different colors) covered by identified barriers.

Results

Summary of Included Studies

Ninety-five studies met inclusion criteria (Figure 1). There was considerable study-to-study variation in both the type of technologies reviewed and the level of detail reported regarding the technologies. Studies related to AT facilitated remote monitoring. The most commonly investigated technologies were telehealth or mHealth systems, generic information communication technologies such as smartphones, and other communication devices such as tablets (n=44). Smart home technologies such as ambient sensors were the second most common technologies (n=23), followed by wearables such as sensors and alarms (n=13). The remaining studies investigated workforce support tools such as electronic health records or decision support tools (n=8) and robots (n=10); four studies reported on findings for more than one AT type. Further details on each included study can be viewed in Tables S1-S5 in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 1. PRISMA flow diagram of review. “Reports” indicates a full text document for screening. “Studies” indicates individual investigations. PRISMA: Preferred Reporting Items for Systematic Review and Meta-Analyses.

XSL•FO
RenderX

Study Quality

Overall, studies included in this review effectively specified their research question—target population group—and described their findings adequately in relation to the aims of the study. Recruitment strategies were often not reported in sufficient detail. Additionally, in general, studies did not adequately report steps taken to ensure rigor of findings, with only five studies adequately considering any potential influences the researcher(s) may have had over participants (eg, whether the researcher is known to the participants, is an external evaluator, or a member of the intervention implementation team), in turn meeting all criteria of CASP [30-34]. Some studies provided limited descriptions of the types of technologies tested in the studies, which in turn made defining specific barriers between technologies challenging. The majority of studies were small scale and of a moderate to high risk of bias overall. Table S7 in Multimedia Appendix 1 provides full CASP scores for all included studies.

Barriers to AT Use

Three of the 10 themes were related to technology design (intrusive/discomfort, poor functionality, and usability; see Table 2). The most common barrier across all technology types was usability, and lack of knowledge was also consistently identified (eg, lack of experience with technology, unfamiliarity with digital interfaces, etc). Stigma showed considerable variation, with no barriers related to stigma identified for telehealth or mHealth but 30% (n=4) of studies assessing wearables reporting stigma as a barrier. In particular, stigma was more common in wearable technologies and alarm systems, which were perceived to draw attention to users. Indeed, certain barriers appear to be unique to a specific AT. For example, discomfort was cited as a barrier to use in 36% (n=5) of wearable studies but not across any of the other AT groups. Similarly, lack of human interaction was more common in smart home technology, robots, and telehealth or mHealth but not cited for wearables or workforce support. A narrative summary of each identified barrier and facilitator is provided below, with supporting quotes available in Table S6 in Multimedia Appendix 1.

Table . Overarching barrier themes with descriptions.

Barrier and facilitator theme names	Descriptor
Barriers	
Privacy	Any reference to perceptions of the AT[a] invading personal privacy, such as data breaches, being watched, or having access to personal information
Cost	Any reference to the financial implication of using AT, such as the cost of the equipment, its upkeep, or subscription cost.
Usability	Any reference to how easy the AT is to use. For example, if buttons are hard to press or screens are small, making legibility difficult
Poor functionality	Distinct from usability. Any perception that the AT does not work as it is intended to, reducing confidence in its effectiveness, for example, false alarms, poor battery life
Stigma	Any perception that using AT draws negative attention to the user such as pity or perceptions of frailty
Intrusive/discomfort	Distinct from privacy, AT that is uncomfortable to wear or creates obstacles in home
Lack of need	Perception that the AT is not required for care or that the participants do not feel they require such devices for their health
Lack of knowledge	Perception that the participant does not know enough about the correct way to use the AT to use it effectively
Fear of misuse	Reluctant to use AT for fear of breaking it due to incorrect use or compromising safety/privacy through incorrect use
Lack of human interaction	Perceptions that AT is replacing human carers to the detriment of older adults, for example, reduced carer visits due to smart home monitoring, reducing social interactions
Facilitators	
Tailored training	Any form of structured training/education in the use of AT
Viewed positively by trusted relations	Carers, medical professionals, or family members encourage the use of the AT and advocate for it
Support	Any informal help from carers, friends, or family members to be more proficient at using AT

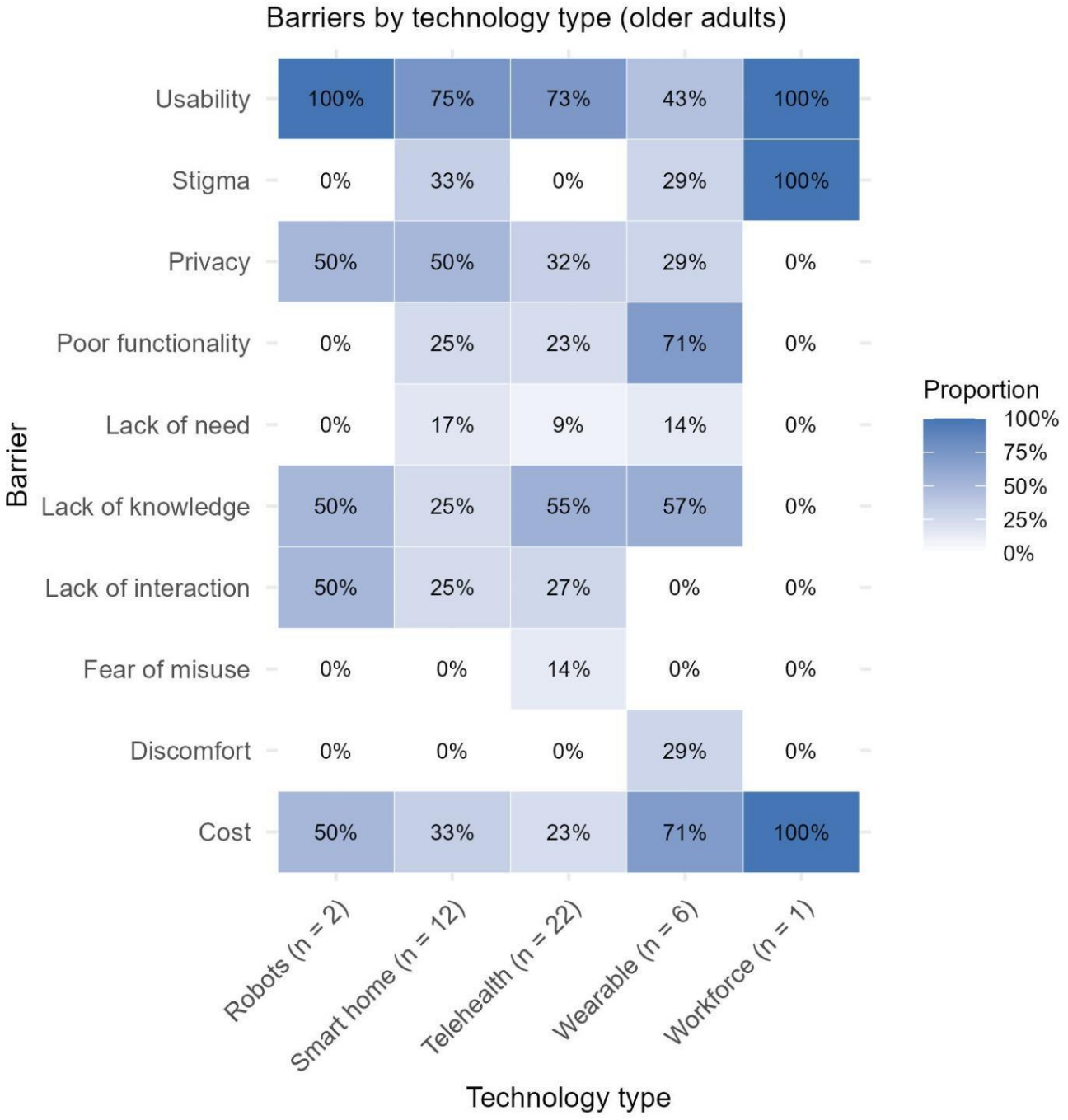
[a]AT: assistive technology.

XSL•FO
RenderX

Identified Barriers by Study Participant Group

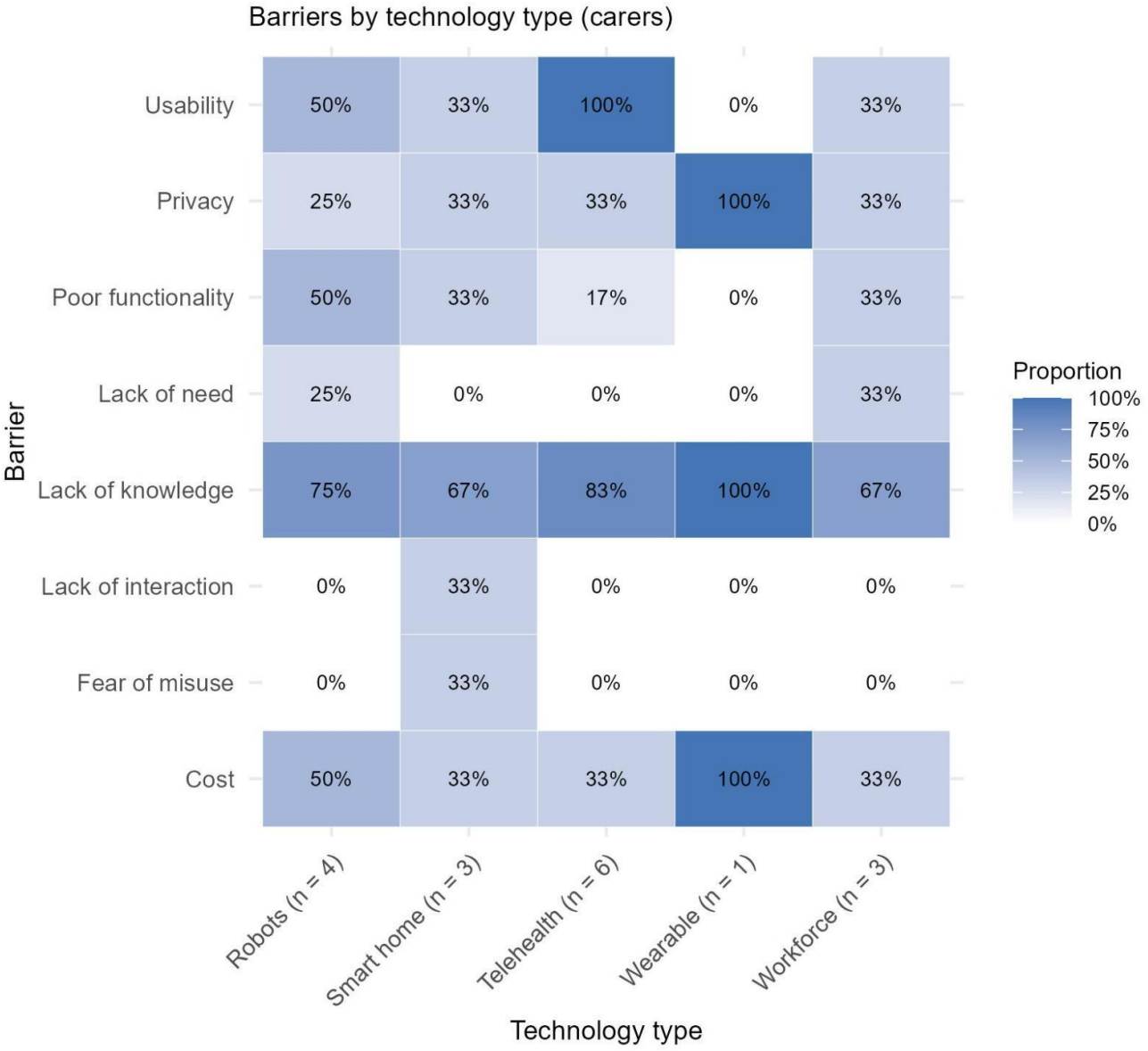
The majority of studies exclusively included older adults as the research participants, with no carer perspectives included (n=43). Conversely, a minority of studies exclusively recruited carers (n=17), with the remaining studies (n=36) including both older adults and carers in the study sample. Of these combined studies, presentation of distinct findings by each group was limited, making separation of barriers or facilitators experienced by older adults versus carers challenging. Figure 2 present the proportion of barriers for each technology type across studies of older adults, carers, and combined samples separately. When identified barriers between the carer group and older adult group are considered, there are clear differences in the proportional distribution of some barriers. For example, no barriers related to stigma were identified for any technology among carers, while this was a significant barrier for older adults, particularly regarding wearables and smart home technologies. Lack of human interaction also appeared to be more of an issue across technology types for older adults than for carers. Conversely, carers generally identified lack of knowledge as a barrier across all technologies to a greater extent than older adults.

Figure 2. Matrix plot of proportion of barriers identified across technology types for studies including older adults only.



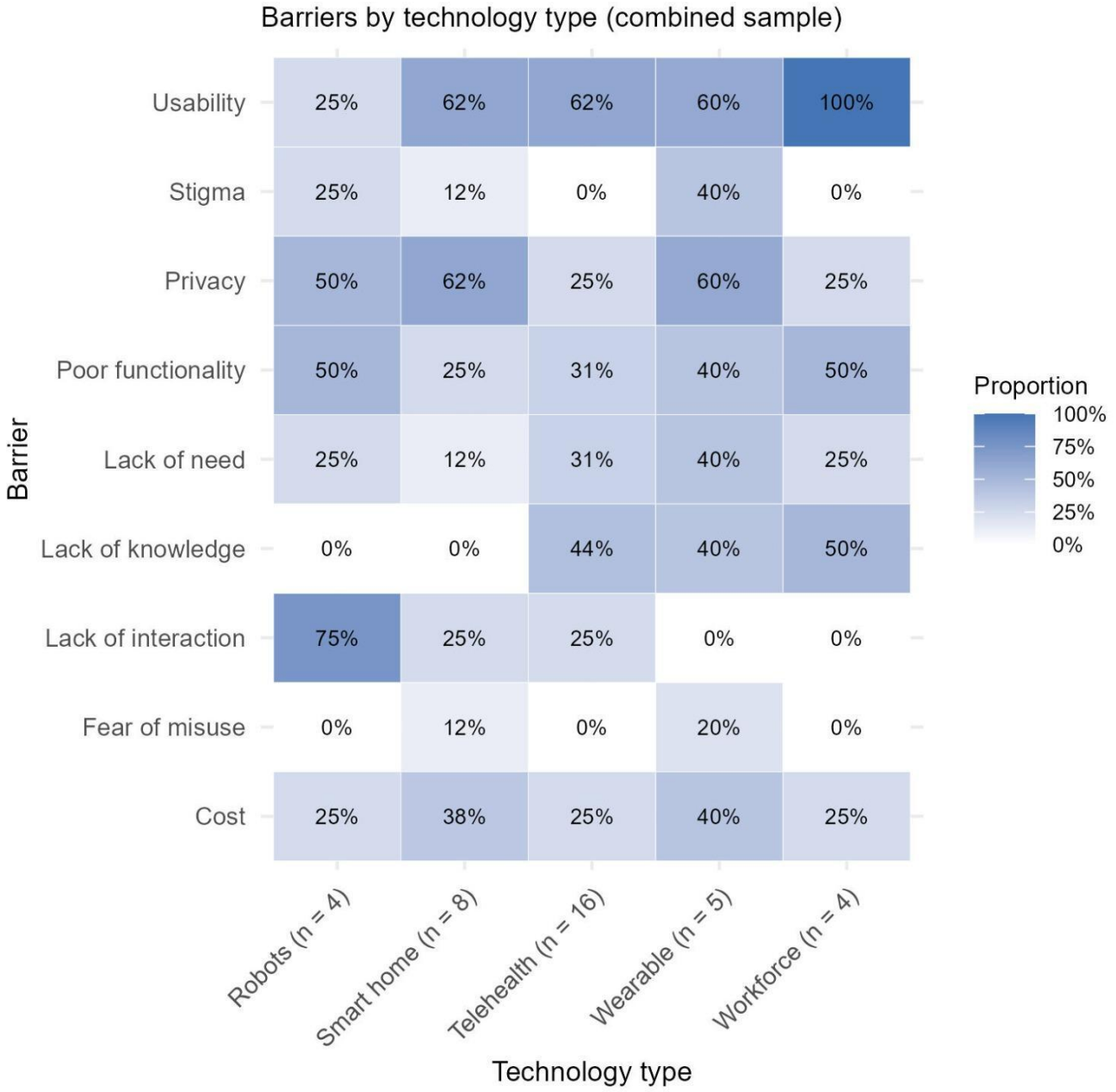
XSL•FO
RenderX

Figure 3. Matrix plot of proportion of barriers identified across technology types for studies including carers only.



XSL•FO
RenderX

Figure 4. Matrix plot of proportion of barriers identified across technology types for studies including carers only.



Technology Design

Usability

Usability was the most persistent barrier across all AT types (Figures 2-4) [5,30,31,34-87]. Complex interfaces, intricate buttons/controls, and small text/letters coupled with physical and cognitive limitations were all cited as barriers to effective use of these technologies.

Poor functionality

Poor functionality was a persistent barrier across all AT types [33,35,37-39,41,46,55,60,70,71,80,81,85,88-102]; this pertained to the perception that AT did not actually address the issue it was designed to assist with. This was particularly true where the technology failed or did not work as intended (eg, battery life of sensors was not fit for purpose, inaccurate alarms sounding when no adverse event had taken place, or software that crashed).

Discomfort

Some wearable devices (but not other AT types) were also viewed as uncomfortable to wear, which discouraged usage. Four studies [61,81,97] highlighted that participants did not like wearing the device as it felt uncomfortable and "in the way."

Other Barriers Related to Particular Technologies

Cost

Cost was a major barrier identified across AT types, with 28 studies citing concerns with the financial implications of installing or using AT [5,32-34,38,43,45,47,49,50,53,58,59,70,75,77,82,83,94,96,99,100,103-107]. There was a perception that there would be ongoing costs associated with owning these technologies.

One study of falls prevention technology highlighted that financial cost was only identified as a barrier for participants in their sample that were considered low income but not for

XSL•FO
RenderX

more affluent participants [92], highlighting the potential for AT to widen health inequalities if the barriers imposed by cost are not considered from the outset.

Out of all the barriers identified in this review, cost was one for which there could be potential contextual implications relating to geographical location. Specifically, as the United States is the only high-income country without universal health care, it could be theorized that cost would be a more pronounced barrier among US-based studies. However, of the 27 US-based studies included in this review, only 9 (33%) identified cost as a barrier, which does not considerably differ from the 24 out of 69 (34%) non US–based studies that found cost to be a barrier.

Lack of knowledge and fear of misuse

A lack of knowledge or general fear and anxiety about using technology incorrectly or breaking devices was common. This barrier was particularly evident in studies focused on technology that required participants to interact with software, such as smartphones, wearable devices, or computers [51,66,74,85,97,98,108,109]. In contrast, lack of knowledge was distinctly lower proportionally for smart home technologies, which tend to be more passive in their function (Figures 2-4).

This lack of knowledge and fear of incorrect use often led to feelings of anxiety among older adults that they were unknowingly using the technology wrong, which could lead to negative consequences (eg, inaccurate information being communicated to their providers regarding their health) [98].

Perceived lack of need

Perceived lack of need was a persistent barrier to AT usage, particularly for technologies designed to assist older adults to live independently (such as robots or fall prevention systems). This relates to individual self-perception that they were not yet of the age or physical condition to require such assistance, even when it was recommended by health care professionals or family members [44,48,57,62,66,78,92,94,98,104,106,110-112]. One study of older adults' usage of technology to assess frailty in the home demonstrated that although participants acknowledged that such technology could be useful in older age, they did not think that they personally needed such AT to live independently at this point in time, a sentiment shared by participants in other studies [88]. Similarly, some of the technologies reported in this review were perceived by participants (both older adults and caregivers) as simply not useful [88,109].

Concern About Adverse Consequences

Lack of interaction

Lack of interaction was a commonly identified barrier in studies relating to smart home tech, telehealth or mHealth, and robots, in the sense of ATs leading to a decrease in human interaction and opportunities to have informal discussions with health care providers [5,31,35,45,47,52,56,62,84,103,104,113-115]. For some individuals, this represented one of their only opportunities to get out of their home and meet another person (eg, a health care provider), or for someone to visit them (eg, a carer).

Privacy

Thirty-eight studies identified privacy concerns as a barrier to AT usage [5,30-32,34,36,44,45,47,50,56,58,63,67-70,72,75,77,85,89,96,99,100,105,111,112,115-117], either due to perceived intrusion on personal data or potential invasion of personal privacy by smart home technologies. Both older adults and carers in a number of studies perceived the installation of technologies such as ambient sensors or cameras as an invasion of personal space and privacy, which could make them feel uneasy in their own home.

Older adults also expressed concern that health professionals, insurance companies, or members of their family would be receiving private or delicate information regarding their health status, movements, or activities, which was perceived as an invasion of privacy.

Stigmatization

Stigmatization or general awareness of what others would think of them as a result of AT use was also a persistent barrier among older adults identified in ten studies [53,59,70,79,89,92,103] . This was evident in a study of ATs for older adults with age-related vision loss.

Stigma was most prevalent as a barrier to use of wearable technologies, while no studies relating to telehealth or mHealth cited stigma as a barrier. This is likely due to the attention some wearables can draw to the individual due to audible alarms or the visual design of the AT [118].

Facilitators to AT Use

Overview

As is common in research on barriers and facilitators in health care, many of the facilitators identified in this review directly reflect the removal of the barriers described above. For example, affordability, privacy measures, and ease of use were all identified as facilitators to AT use. Below we explicitly describe three identified facilitators that are distinct from simple removal of barriers.

Awareness of Benefits

Certain ATs were highlighted as allowing older adults to learn more about their own health behaviors and daily activities, which was viewed as a facilitator (particularly for wearables) [119,120]. If technology was perceived to increase safety and independence, participants were more likely to state that they would be willing to use it in a number of studies [40,119,120]. For example, across multiple studies of ambient sensors, participants highlighted that the AT offered real benefits by increasing safety and reducing caregiver burden.

Targeted Training

One of the most consistent facilitators to AT usage cited in the included studies was the use of training for older adults on how to properly use the technologies they were to be given [54]. This included strategies such as preimplementation of training materials on AT use to familiarize older adults with the technology [39], and one-to-one or group assistance with usage during initial implementation [53].

XSL•FO
RenderX

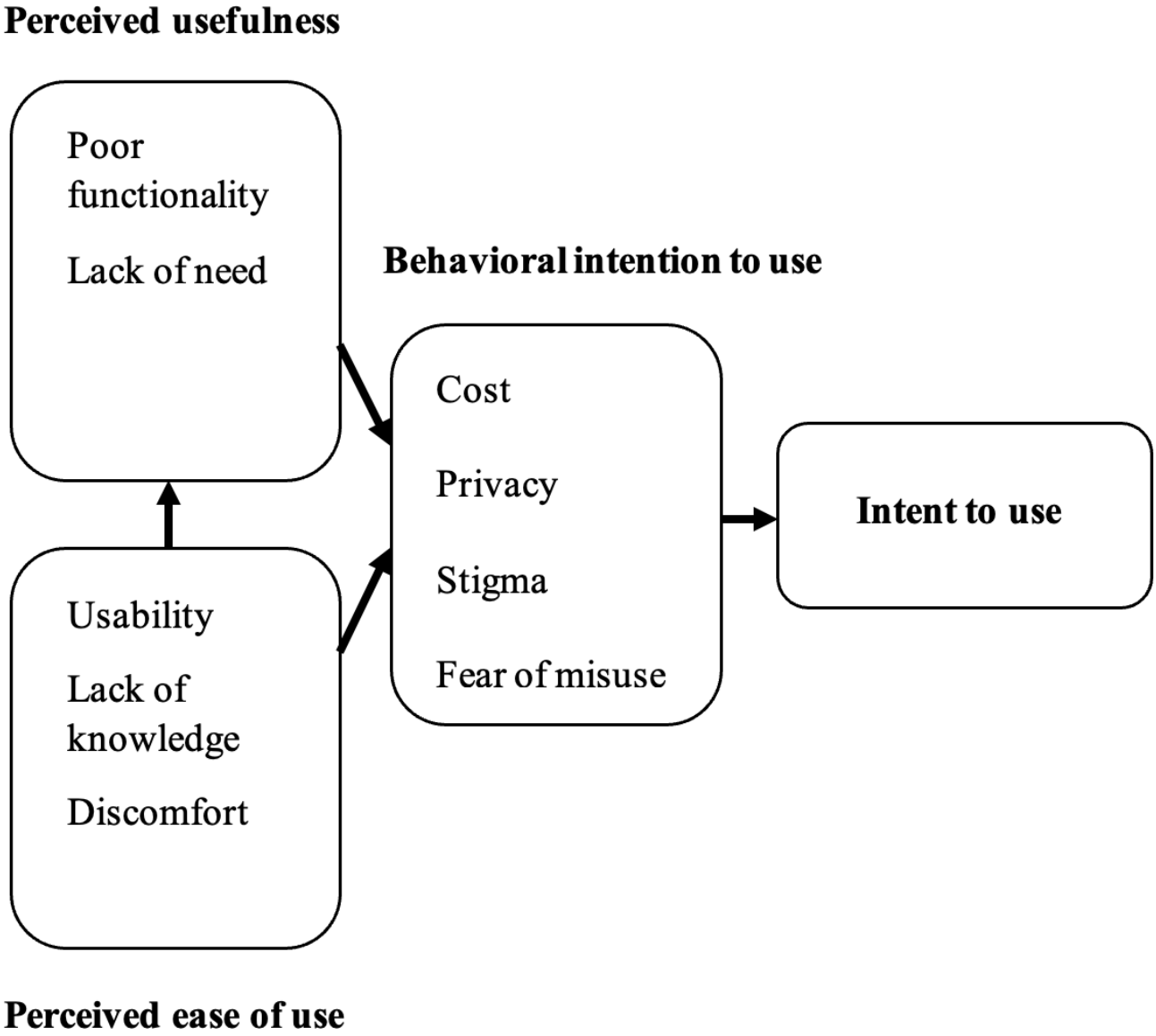
User-centered design

A user-centered design was viewed as a facilitator of more user-friendly technology development in 23 studies. Participants in a number of studies cited efforts to involve users in the design process as a benefit to usability. For example, Young and colleagues [87] employed an iterative design process (involving formal interviews and journey mapping) to develop a digital home cognitive screening tool. They found this person-centered design approach to be an asset in developing an AT that was acceptable across different user groups.

Barriers and Facilitators in Relation to the TAM

Figure 5 illustrates how all 10 of the identified barriers from the primary studies map onto the TAM, highlighting the points at which specific barriers present problems in sustained use of AT for care. While the majority of these aligned with the "behavioral intention to use" domain of TAM (cost, privacy, stigma, fear of misuse, and lack of human interaction), barriers such as poor functionality and lack of need impacted on "perceived usefulness," and three barriers (discomfort, usability, and lack of knowledge) mapped onto the "perceived ease of use" domain. Likewise, ATs that were found not to work as intended, or where no explicit care need was perceived to have been addressed, corresponded with perceived usefulness domain. All identified barriers were able to be mapped against a domain of the TAM (Figure 5).

Figure 5. Identified barriers to assistive technology use mapped against the technology acceptance model.



Barriers to AT by TAM Domain Pre- Versus Post-COVID

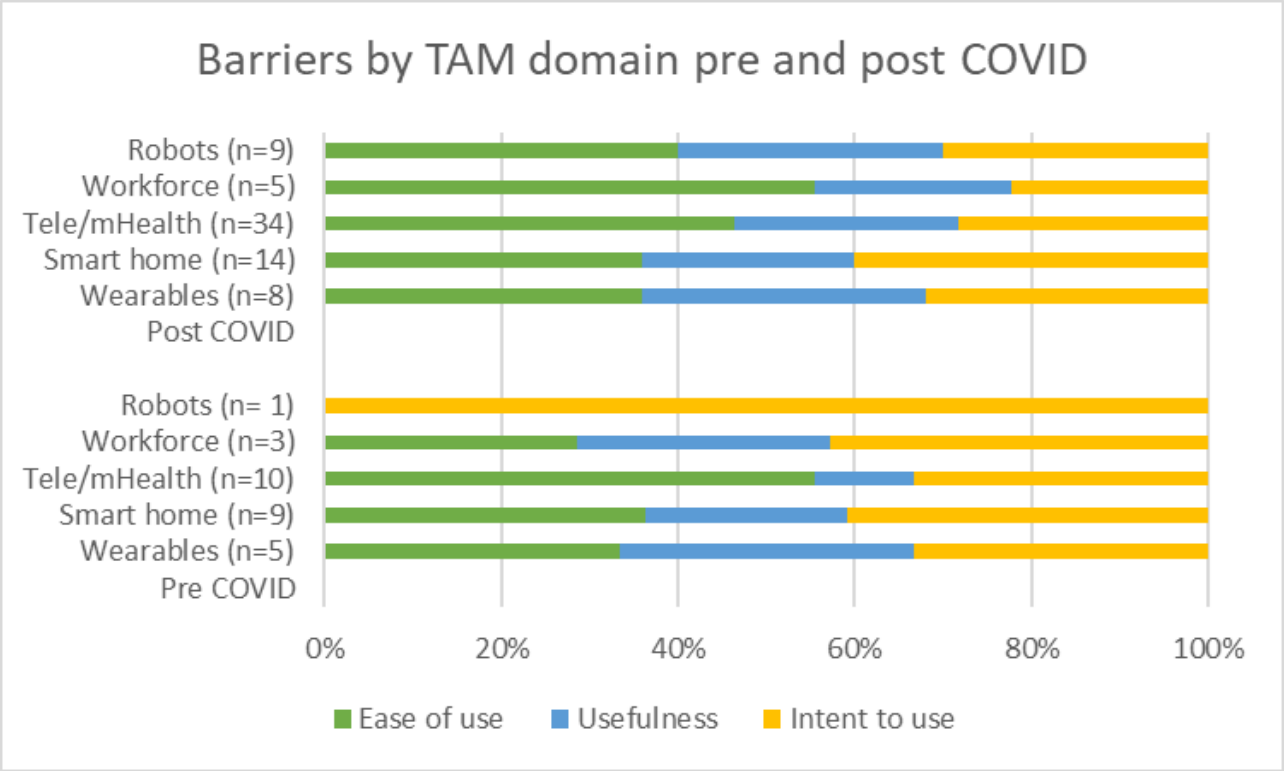
Figure 6 illustrates the potential influence that the COVID-19 pandemic may have had on both the quantity of studies assessing AT for care and their acceptance by users in relation to the domains of the TAM.

The type of barriers and their proportions in relation to the TAM is generally similar between pre- and post-COVID. However, for robots, studies identified further barriers in ease of use, usefulness, and intent to use. In telehealth or mHealth, there was a proportional increase in usefulness barriers, and for workforce support tools, there was a proportional increase in

XSL•FO
RenderX

usability barriers. There was no marked change in barriers relating to intention to use AT in any technology type.

Figure 6. Proportions of identified barriers by TAM domain for studies published pre (2011 - 2020) and post COVID (2021 - 2024). TAM: Technology Acceptance Model.



Discussion

Principal Findings

This systematic review identified 10 barriers related to the use of AT for the care of older adults in community settings. These mapped onto various domains across the TAM, highlighting the point at which such barriers may negatively influence adoption. The findings of this review are reinforced by previous reviews on this topic [13,15,16]. For example, privacy of AT was a consistent barrier, reflecting the findings of previous reviews [15,17]. Specifically, Yusif et al [17] identified privacy to be the most prevalent barrier identified in their review, being cited within 34% (n=15) of the 44 included studies. Similar findings were observed for usability, both for our review and existing reviews [14,17]. Also, in line with the findings of our review, Kruse et al's [14] review of 57 telehealth and mHealth acceptability studies found both technical illiteracy and a lack of desire to be two of the most prevalent barriers identified. This reflects findings in our review whereby perceived lack of need and lack of knowledge were persistent barriers identified across all AT, including studies of telehealth and mHealth. Stigma was also cited as another barrier to AT adoption across several reviews [14,15,17]. This appears to be more prevalent in ATs that are perceived to draw attention to the user, such as body-worn sensors or alarm systems [118]. Wang et al's [16] mixed methods review of long-distance care technologies including 41 studies highlighted the importance of considering personal factors which may mediate adoption of AT among older adults. Specifically, their review highlighted that factors such as individual personality and cohabiting circumstances can influence AT adoption, just as frequently as technology factors such as poor functionality and usability. These personal factors are often more challenging to capture in primary studies, highlighting the need for more robust study designs such as controlled studies with representative participant groups. Additionally, Vassli and Farshchian's [15] systematic review of qualitative studies investigating ICT adoption in older adults mapped identified themes to the UTAUT2. Barriers identified largely mirrored those identified in our review (cost, functionality, lack of knowledge, lack of human interaction, discomfort, and stigma). Mapping the UTAUT2 provided further nuance to these findings, highlighting where specific barriers may influence adoption in relation to each domain. For example, obtrusiveness or discomfort was theorized to negatively impact effort expectancy, thus in turn reducing behavioral intention to adopt intrusive ICT, while stigma could be classified as a barrier to adoption within the social influence domain of the UTAUT2 [15].

To our knowledge, this is the first review to also assess the barriers of a comprehensive range of AT types in conjunction with stages of the TAM. This exercise allowed for the identification of the barriers to AT use in relation to TAM domains. In conducting this exercise, it was important to acknowledge that the COVID-19 pandemic increased the volume of research output in this field exponentially and is likely to have expedited the pace at which a number of technologies have progressed from state of the art to well established and commonly used in the care of older adults [18]. What is less clear, however, is whether these advances in widespread use of ATs have also coincided with progress in addressing the

XSL•FO
RenderX

persistent barriers, which prevent older adults from optimally engaging with care technologies. This review has demonstrated an almost 3-fold increase in the volume of studies in this area, within a third of the timeframe post-COVID-19. It is important to note, however, that it was not possible to identify the mechanisms and drivers behind the increase in research output post-COVID-19. While the pandemic and the necessity to adhere to social distancing played a role, the natural increase in pace and progression of technological advances was also a factor.

Using the TAM domains to map identified barriers, our analysis demonstrates persistent challenges in addressing barriers, which affect older adults' behavioral intention to use technologies. Barriers relating to ease of use and perceived usefulness are largely similar in proportion for both pre and post COVID-19, with the exception of robots, for which a large increase is observed mainly due to more studies now being published versus the one study that existed pre-COVID. This indicates that despite the increased rate of publication in the field, barriers remain the same as those which impeded AT use before the pandemic.

This is unsurprising given the known complexities involved in influencing behavior change in clinical interventions [121]. In older studies of AT adoption, barriers relating to ease of use and perceived usefulness tended to be easier to rectify with alterations to AT hardware or software (albeit with unique challenges related to use by older adults, who often have additional needs such as frailty or vision loss that can require products to need further optimization) [122]. Conversely, addressing barriers such as perceived stigma, fear of misuse, or a lack of human interaction requires more than adaptations to the AT itself and will likely require multicomponent strategies to address the cultural and behavioral issues around the adoption of such technologies [16]. While both primary research and systematic reviews such as the present study can help identify barriers to AT use, adopting strategies to overcome such barriers will require careful planning from all stakeholders. Implementation science frameworks offer researchers and practitioners a viable avenue to achieve this [123]. One such framework that has been successfully applied to technological interventions is the Consolidated Framework for Implementation Research (CFIR) [124,125]. Briefly, CFIR consists of five domains (intervention characteristics, outer setting, inner setting, individual characteristics, and implementation processes). These domains contain further constructs upon which identified barriers of an intervention or ICT can be mapped in order to determine relevant implementation strategies [126]. If used in conjunction with the TAM, these implementation strategies can be further optimized to suit the stages at which barriers occur in the adoption process, as recent studies have achieved. For example, Gallant et al [127] conducted a mixed methods preimplementation study whereby the CFIR was used in conjunction with UTAUT to inform the implementation of automated pain behavior monitoring technology in a health care setting, leading to increased understanding of behavioral intentions to use technology and refinements to the UTAUT, which may help in making the framework more widely applicable to practice. Therefore, the results of this review, which has mapped commonly identified barriers in AT adoption to the TAM, can help researchers to expedite this process when adopting frameworks such as CFIR.

Strengths of this review include the comprehensive search strategy and the consideration of the potential influence of the COVID-19 pandemic to understand how this has influenced technology acceptance. There are a number of limitations that should be considered when interpreting the findings. For example, following the risk of bias assessment using CASP, a large portion of the included studies was deemed to be of moderate to high risk of bias, limiting the trustworthiness of the findings reported. Specifically, studies generally did not adequately address the relationship between the interviewer and the participants. This observation is not limited to this current review. A review of qualitative study methodologies found that none of the 19 included articles adequately described reflexive practice in data collection/analysis [128]. It was noted that most of the included studies appeared to be stand-alone, small-scale feasibility studies. While such studies are an important step in the refinement of ATs, their formative status can limit their generalizability when attempting to use such findings to inform practice at scale [129]. Additionally, previous reviews of the literature have highlighted that participants in studies of technology acceptance tend to be higher than average socioeconomic or educational status and at the younger end of older age [130]. Similar issues likely apply to the studies included in this review (Tables S1-S5 in Multimedia Appendix 1), which could limit generalizability. Finally, we applied a pragmatic approach to data synthesis, using double coding of a proportion of studies to develop a coding framework, before single coding was applied to the remaining studies as is common practice in reviews of a similar nature [131]. Consequently, there is an increased risk of misclassification bias of identified themes due to this. However, given the barriers typically identified in AT use among older adults are fairly distinct and straightforward to define, the true impact of this methodological limitation is likely minimal.

Conclusions

The scope and scale of ATs offering support to older adults and their carers continues to expand. However, ATs commonly are not optimized for use in this context, which could be addressed by better codesign or user-centered design (TAM domains of usefulness and ease of use). Other important barriers relate to the TAM domain of behavioral intent to use ATs, which requires a different set of interventions. Future research should take an implementation science approach by identifying solutions to the barriers to AT use, while conducting more rigorous evaluations of optimized ATs accompanied by robust process evaluations to help identify the impact of any persistent barriers on adoption and other outcomes of interest.

XSL•FO
RenderX

Acknowledgments

The authors thank Marshall Dozier for assistance with search strategy development. The views expressed are those of the authors and not necessarily those of Legal and General PLC.

Funding

SWM, BG, BB, SL, IU and HF are supported by the Advanced Care Research Center (ACRC), which is funded by Legal and General PLC (as part of their corporate social responsibility program), providing a research grant to establish the independent ACRC at University of Edinburgh. The funder had no role in the conduct of the study, interpretation, or the decision to submit for publication.

Authors' Contributions

SM, KM, BG, and AOD conceptualized the study and developed the review protocol. SM and KM conducted searches. SM, KM, HF, SWM, AE, BKB, SDS, BG, and AOD conducted screening, data extraction, and data analysis. SM and AOD drafted the manuscript. All authors revised the manuscript. All authors reviewed and approved the final submitted manuscript.

Conflicts of Interest

None declared

Multimedia Appendix 1
Full study description tables, full text exclusions with reasons, CASP quality assessment scoring, and MEDLINE search strategy.
[DOCX File, 144 KB - aging_v8i1e73917_app1.docx]

Checklist 1
PRISMA checklist.
[PDF File, 918 KB - aging_v8i1e73917_app2.pdf]

References

1. Kontis V, Bennett JE, Mathers CD, Li G, Foreman K, Ezzati M. Future life expectancy in 35 industrialised countries: projections with a Bayesian model ensemble. Lancet 2017 Apr;389(10076):1323-1335. [doi: 10.1016/S0140-6736(16)32381-9]
2. Beard JR, Officer A, de Carvalho IA, et al. The world report on ageing and health: a policy framework for healthy ageing. Lancet 2016 May;387(10033):2145-2154. [doi: 10.1016/S0140-6736(15)00516-4]
3. HM Government. Assistive technology: definition and safe use: Care DoHS; 2023. URL: https://www.gov.uk/government/publications/assistive-technology-definition-and-safe-use/assistive-technology-definition-and-safe-use#introduction [accessed 2025-11-18]
4. Balcombe L, De Leo D. An integrated blueprint for digital mental health services amidst COVID-19. JMIR Ment Health 2020 Jul 22;7(7):e21718. [doi: 10.2196/21718] [Medline: 32668402]
5. Peek STM, Luijkx KG, Rijnaard MD, et al. Older adults' reasons for using technology while aging in place. Gerontology 2016;62(2):226-237. [doi: 10.1159/000430949] [Medline: 26044243]
6. Klimova B, Valis M, Kuca K. Exploring assistive technology as a potential beneficial intervention tool for people with Alzheimer's disease - a systematic review. Neuropsychiatr Dis Treat 2018;14:3151-3158. [doi: 10.2147/NDT.S181849] [Medline: 30532546]
7. Howard J, Fisher Z, Kemp AH, Lindsay S, Tasker LH, Tree JJ. Exploring the barriers to using assistive technology for individuals with chronic conditions: a meta-synthesis review. Disabil Rehabil Assist Technol 2022 May;17(4):390-408. [doi: 10.1080/17483107.2020.1788181] [Medline: 32663110]
8. Sugawara AT, Ramos VD, Alfieri FM, Battistella LR. Abandonment of assistive products: assessing abandonment levels and factors that impact on it. Disabil Rehabil Assist Technol 2018 Oct;13(7):716-723. [doi: 10.1080/17483107.2018.1425748] [Medline: 29334475]
9. Abrilahij A, Boll T. A qualitative metasynthesis of reasons for the use or nonuse of assistive technologies in the aging population. GeroPsych (Bern) 2019 Jun;32(2):79-92. [doi: 10.1024/1662-9647/a000203]
10. Martin JK, Martin LG, Stumbo NJ, Morrill JH. The impact of consumer involvement on satisfaction with and use of assistive technology. Disabil Rehabil Assist Technol 2011;6(3):225-242. [doi: 10.3109/17483107.2010.522685] [Medline: 20929425]
11. O'Connor S, Hanlon P, O'Donnell CA, Garcia S, Glanville J, Mair FS. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. BMC Med Inform Decis Mak 2016 Sep 15;16(1):120. [doi: 10.1186/s12911-016-0359-3] [Medline: 27630020]

XSL•FO
RenderX

12. Davis FD. A technology acceptance model for empirically testing new end-user information systems: theory and results [phd thesis]. : Massachusetts Institute of Technology, Sloan School of Management; 1986 URL: http://hdl.handle.net/1721.1/15192 [accessed 2025-11-18]
13. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: an integrative systematic review. BMC Public Health 2024 Feb 16;24(1):506. [doi: 10.1186/s12889-024-18036-5] [Medline: 38365698]
14. Kruse C, Fohn J, Wilson N, Nunez Patlan E, Zipp S, Mileski M. Utilization barriers and medical outcomes commensurate with the use of telehealth among older adults: systematic review. JMIR Med Inform 2020 Aug 12;8(8):e20359. [doi: 10.2196/20359] [Medline: 32784177]
15. Vassli LT, Farshchian BA. Acceptance of health-related ICT among elderly people living in the community: a systematic review of qualitative evidence. Int J Human Comput Interact 2018 Feb;34(2):99-116. [doi: 10.1080/10447318.2017.1328024]
16. Wang J, Fu Y, Lou V, Tan SY, Chui E. A systematic review of factors influencing attitudes towards and intention to use the long-distance caregiving technologies for older adults. Int J Med Inform 2021 Sep;153:104536. [doi: 10.1016/j.ijmedinf.2021.104536] [Medline: 34325206]
17. Yusif S, Soar J, Hafeez-Baig A. Older people, assistive technologies, and the barriers to adoption: A systematic review. Int J Med Inform 2016 Oct;94:112-116. [doi: 10.1016/j.ijmedinf.2016.07.004] [Medline: 27573318]
18. Renu N. Technological advancement in the era of COVID-19. SAGE Open Med 2021;9:20503121211000912. [doi: 10.1177/20503121211000912] [Medline: 33786181]
19. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009 Aug 18;151(4):264-269. [doi: 10.7326/0003-4819-151-4-200908180-00135] [Medline: 19622511]
20. Tong A, Flemming K, McInnes E, Oliver S, Craig J. Enhancing transparency in reporting the synthesis of qualitative research: ENTREQ. BMC Med Res Methodol 2012 Nov 27;12(1):1-8. [doi: 10.1186/1471-2288-12-181] [Medline: 23185978]
21. de Joode E, van Heugten C, Verhey F, van Boxtel M. Efficacy and usability of assistive technology for patients with cognitive deficits: a systematic review. Clin Rehabil 2010 Aug;24(8):701-714. [doi: 10.1177/0269215510367551] [Medline: 20543021]
22. Blaschke CM, Freddolino PP, Mullen EE. Ageing and technology: a review of the research literature. Br J Soc Work 2009 Jun 1;39(4):641-656. [doi: 10.1093/bjsw/bcp025]
23. Magnusson L, Hanson E, Borg M. A literature review study of Information and communication technology as a support for frail older people living at home and their family carers. Technol Disabil 2004;16(4):223-235. [doi: 10.3233/TAD-2004-16404]
24. Vardell E, Malloy M. Joanna Briggs Institute: an evidence-based practice database. Med Ref Serv Q 2013;32(4):434-442. [doi: 10.1080/02763869.2013.837734] [Medline: 24180651]
25. Singh J. Critical appraisal skills programme. J Pharmacol Pharmacother 2013;4(1):76. [doi: 10.4103/0976-500X.107697]
26. Ritchie J, Spencer L, Bryman A, Burgess RG. Analysing Qualitative Data: Routledge; 1994. URL: https://www.taylorfrancis.com/books/edit/10.4324/9780203413081/analyzing-qualitative-data-alan-bryman-bob-burgess [accessed 2025-11-25]
27. Chang A. UTAUT and UTAUT2: a review and agenda for future research. The Winners 2012;13(2):10. [doi: 10.21512/tw.v13i2.656]
28. Hale JL, Householder BJ, Greene KL. The theory of reasoned action. In: The Persuasion Handbook: Developments in Theory and Practice: Sage; 2002, Vol. 14:259-286. [doi: 10.4135/9781412976046.n14]
29. Ajzen I. The theory of planned behaviour: reactions and reflections. Psychol Health 2011 Sep;26(9):1113-1127. [doi: 10.1080/08870446.2011.613995]
30. Chan-Nguyen S, O'Riordan A, Morin A, et al. Patient and caregiver perspectives on virtual care: a patient-oriented qualitative study. CMAJ Open 2022;10(1):E165-E172. [doi: 10.9778/cmajo.20210065] [Medline: 35232815]
31. Conway A, Ryan A, Harkin D, Mc Cauley C. "It's Another Feather in My Hat"-exploring factors influencing the adoption of apps with people living with dementia. Dementia (London) 2023 Oct;22(7):1487-1513. [doi: 10.1177/14713012231185283] [Medline: 37365816]
32. Dermody G, Fritz R, Glass C, Dunham M, Whitehead L. Family caregiver readiness to adopt smart home technology to monitor care-Dependent older adults: a qualitative exploratory study. J Adv Nurs 2024 Feb;80(2):628-643. [doi: 10.1111/jan.15826] [Medline: 37614010]
33. Hunter I, Lockhart C, Rao V, Tootell B, Wong S. Enabling rural telehealth for older adults in underserved rural communities: focus group study. JMIR Form Res 2022 Nov 4;6(11):e35864. [doi: 10.2196/35864] [Medline: 36331533]
34. Qi C, Wang Y, Qi X, Jiao Y, Que C, Chen Y. Barriers to providing internet-based home care services for urban older adults in China: a qualitative study of the service providers. BMC Geriatr 2023 May 23;23(1):320. [doi: 10.1186/s12877-023-04028-4] [Medline: 37221471]

XSL•FO
RenderX

35. Alexander NB, Phillips K, Wagner-Felkey J, et al. Team VA Video Connect (VVC) to optimize mobility and physical activity in post-hospital discharge older veterans: baseline assessment. BMC Geriatr 2021 Sep 22;21(1):502. [doi: 10.1186/s12877-021-02454-w] [Medline: 34551725]
36. Amiri P, Pirnejad H, Bahaadinbeigy K, Baghini MS, Khazaee PR, Niazkhani Z. A qualitative study of factors influencing ePHR adoption by caregivers and care providers of Alzheimer's patients: an extension of the unified theory of acceptance and use of technology model. Health Sci Rep 2023 Jul;6(7):e1394. [doi: 10.1002/hsr2.1394] [Medline: 37425233]
37. Arkkukangas M, Cederbom S, Tonkonogi M, Umb Carlsson Õ. Older adults' experiences with mHealth for fall prevention exercise: usability and promotion of behavior change strategies. Physiother Theory Pract 2021 Dec;37(12):1346-1352. [doi: 10.1080/09593985.2020.1712753] [Medline: 31910707]
38. Arthanat S, Rossignol H, Preble E, et al. Perspectives on a telepresence robot at an independent living facility: lessons learned and implications. JET 2024 Feb 21;18(1):1-12. [doi: 10.1108/JET-05-2023-0014]
39. Blocker KA, Kadylak T, Rogers WA. Understanding the needs of older adults learning to use digital home assistants: a demonstration study. Gerontechnology 2023 Aug 31;22(2):1-11 [FREE Full text] [doi: 10.4017/gt.2023.22.2.blo.08]
40. Blok M, van Ingen E, de Boer AH, Slootman M. The use of information and communication technologies by older people with cognitive impairments: from barriers to benefits. Comput Human Behav 2020 Mar;104:106173. [doi: 10.1016/j.chb.2019.106173]
41. Braspenning AM, Cranen K, Snaphaan L, Wouters EJM. A multiple stakeholder perspective on the drivers and barriers for the implementation of lifestyle monitoring using infrared sensors to record movements for vulnerable older adults living alone at home: a qualitative study. Int J Environ Res Public Health 2022 Jan 5;19(1):570. [doi: 10.3390/ijerph19010570] [Medline: 35010829]
42. Castro Rojas MD. Barriers and supportive factors for older adults learning about and using information and communication technologies for healthy aging in Costa Rica. Educ Gerontol 2021 Oct 3;47(10):433-446. [doi: 10.1080/03601277.2021.1989228]
43. Cha J, Peterson CM, Millenbah AN, et al. Delivering personalized recommendations to support caregivers of people living with dementia: mixed methods study. JMIR Aging 2022 May 3;5(2):e35847. [doi: 10.2196/35847] [Medline: 35503650]
44. Curtis K, Brooks S. Digital health technology: factors affecting implementation in nursing homes. Nurs Older People 2020 Mar 24;32(2):14-21. [doi: 10.7748/nop.2020.e1236] [Medline: 32159302]
45. Dequanter S, Steenhout I, Fobelets M, et al. Technology implementation in care practices for community-dwelling older adults with mild cognitive decline: perspectives of professional caregivers in Quebec and Brussels. Digit Health 2022;8:20552076221139693. [doi: 10.1177/20552076221139693] [Medline: 36420317]
46. Di Pelino S, Lamarche L, Carr T, et al. Lessons learned through two phases of developing and implementing a technology supporting integrated care: case study. JMIR Form Res 2022 Apr 11;6(4):e34899. [doi: 10.2196/34899] [Medline: 35404253]
47. Elers P, Hunter I, Whiddett D, Lockhart C, Guesgen H, Singh A. User requirements for technology to assist aging in place: qualitative study of older people and their informal support networks. JMIR Mhealth Uhealth 2018 Jun 6;6(6):e10741. [doi: 10.2196/10741] [Medline: 29875083]
48. Essery R, Pollet S, Smith KA, et al. Planning and optimising a digital intervention to protect older adults' cognitive health. Pilot Feasibility Stud 2021 Aug 18;7(1):158. [doi: 10.1186/s40814-021-00884-2] [Medline: 34407886]
49. Faisal S, Ivo J, McDougall A, Patel T. Stakeholder feedback of electronic medication adherence products: qualitative analysis. J Med Internet Res 2020 Dec 1;22(12):e18074. [doi: 10.2196/18074] [Medline: 33258778]
50. Garcia Reyes EP, Kelly R, Buchanan G, Waycott J. Understanding older adults' experiences with technologies for health self-management: interview study. JMIR Aging 2023 Mar 21;6:e43197. [doi: 10.2196/43197] [Medline: 36943333]
51. Gately ME, Tickle-Degnen L, McLaren JE, Ward N, Ladin K, Moo L. Factors influencing barriers and facilitators to in-home video telehealth for dementia management. Clin Gerontol 2022;45(4):1020-1033. [doi: 10.1080/07317115.2021.1930316] [Medline: 34096477]
52. Glomsås HS, Knutsen IR, Fossum M, Christiansen K, Halvorsen K. Family caregivers' involvement in caring for frail older family members using welfare technology: a qualitative study of home care in transition. BMC Geriatr 2022 Mar 18;22(1):223. [doi: 10.1186/s12877-022-02890-2] [Medline: 35303816]
53. Gordon S, Telford-Sharp F, Crowe W, Champion S. Effectiveness of a co‐designed technology package on perceptions of safety in community‐dwelling older adults. Australas J Ageing 2022 Sep;41(3):e257-e265. [doi: 10.1111/ajag.13095] [Medline: 35635481]
54. Gottwald S, Laatikainen TE, Kyttä M. Exploring the usability of PPGIS among older adults: challenges and opportunities. Int J Geogr Inf Sci 2016 Dec 1;30(12):2321-2338. [doi: 10.1080/13658816.2016.1170837]
55. Hafdi M, Eggink E, Hoevenaar-Blom MP, et al. Design and development of a mobile health (mHealth) platform for dementia prevention in the prevention of dementia by mobile phone applications (PRODEMOS) project. Front Neurol 2021;12:733878. [doi: 10.3389/fneur.2021.733878] [Medline: 34975710]
56. Halvorsrud L, Holthe T, Karterud D, Thorstensen E, Lund A. Perspectives on assistive technology among older Norwegian adults receiving community health services. Disabil Rehabil Assist Technol 2023 Jul;18(5):685-692. [doi: 10.1080/17483107.2021.1906962] [Medline: 33861681]

XSL•FO
RenderX

57. Hoel V, Ambugo EA, Wolf-Ostermann K. Sustaining our relationship: dyadic interactions supported by technology for people with dementia and their informal caregivers. Int J Environ Res Public Health 2022 Sep 2;19(17):10956. [doi: 10.3390/ijerph191710956] [Medline: 36078671]
58. Johansson-Pajala RM, Gustafsson C. Significant challenges when introducing care robots in Swedish elder care. Disabil Rehabil Assist Technol 2022 Feb;17(2):166-176. [doi: 10.1080/17483107.2020.1773549] [Medline: 32538206]
59. Kabacińska K, Vu K, Tam M, Edwards O, Miller WC, Robillard JM. "Functioning better is doing better": older adults' priorities for the evaluation of assistive technology. Assist Technol 2023 Jul 4;35(4):367-373. [doi: 10.1080/10400435.2022.2113180] [Medline: 35972791]
60. Kang HS, Koh IS, Makimoto K, Yamakawa M. Nurses' perception towards care robots and their work experience with socially assistive technology during COVID-19: A qualitative study. Geriatr Nurs (Lond) 2023 Mar;50:234-239. [doi: 10.1016/j.gerinurse.2023.01.025]
61. Kim B, Ghasemi P, Stolee P, Lee J. Clinicians and older adults' perceptions of the utility of patient-generated health data in caring for older adults: exploratory mixed methods study. JMIR Aging 2021 Nov 5;4(4):e29788. [doi: 10.2196/29788] [Medline: 34738913]
62. Ladekjær Larsen E, Waldorff FB, Hansen HP, la Cour K. Home-based training technology for persons with dementia: a qualitative study of barriers and facilitators for mobility-based training at home. BMC Geriatr 2022 Oct 14;22(1):800. [doi: 10.1186/s12877-022-03505-6] [Medline: 36241981]
63. Lee AR, McDermott O, Orrell M. Understanding barriers and facilitators to online and app activities for people living with dementia and their supporters. J Geriatr Psychiatry Neurol 2023 Sep;36(5):366-375. [doi: 10.1177/08919887221149139] [Medline: 36597870]
64. Mao A, Tam L, Xu A, et al. Barriers to telemedicine video visits for older adults in independent living facilities: mixed methods cross-sectional needs assessment. JMIR Aging 2022 Apr 19;5(2):e34326. [doi: 10.2196/34326] [Medline: 35438648]
65. Mishra RK, Park C, Momin AS, et al. Care4AD: a technology-driven platform for care coordination and management: acceptability study in dementia. Gerontology 2023;69(2):227-238. [doi: 10.1159/000526219] [Medline: 36096091]
66. Mishuris RG, Stewart M, Fix GM, et al. Barriers to patient portal access among veterans receiving home-based primary care: a qualitative study. Health Expect 2015 Dec;18(6):2296-2305. [doi: 10.1111/hex.12199] [Medline: 24816246]
67. Nahm ES, Son H, Yoon JM. Older adults' use of patient portals: experiences, challenges, and suggestions shared through discussion board forums. Geriatr Nurs (Lond) 2020 Jul;41(4):387-393. [doi: 10.1016/j.gerinurse.2019.12.001]
68. Nallam P, Bhandari S, Sanders J, Martin-Hammond A. A question of access: exploring the perceived benefits and barriers of intelligent voice assistants for improving access to consumer health resources among low-income older adults. Gerontol Geriatr Med 2020;6:2333721420985975. [doi: 10.1177/2333721420985975] [Medline: 33457459]
69. Naudé B, Rigaud AS, Kamali L, Pino M. Barriers and enablers for the use of digital interactive television in nursing home settings: an interview case study with older adults and professionals. Int J Environ Res Public Health 2023 Jan 18;20(3):1813. [doi: 10.3390/ijerph20031813] [Medline: 36767181]
70. Neubauer N, Spenrath C, Philip S, Daum C, Liu L, Miguel-Cruz A. Identifying adoption and usability factors of locator devices for persons living with dementia. Dementia (London) 2022 Apr;21(3):862-881. [doi: 10.1177/14713012211065381] [Medline: 34964391]
71. Oche O, Dobyns R, Lin C, Kennelty KA. Determining the factors that impede or facilitate the utilization of telemedicine (video visits) for geriatric patients. Telemed Rep 2022;3(1):156-165. [doi: 10.1089/tmr.2022.0018] [Medline: 36127949]
72. Patel Y, Vassilev I, du Toit SHJ. Implementing a digital tool to support meaningful engagement with socially isolated or lonely older adults. Health Social Care Comm 2022 Nov;30(6):e6456-e6466. [doi: 10.1111/hsc.14090]
73. Piculell E, Skär L, Sanmartin Berglund J, Anderberg P, Bohman D. Using a mobile application for health communication to facilitate a sense of coherence: experiences of older persons with cognitive impairment. Int J Environ Res Public Health 2021 Oct 28;18(21):11332. [doi: 10.3390/ijerph182111332] [Medline: 34769845]
74. Pirhonen J, Lolich L, Tuominen K, Jolanki O, Timonen V. "These devices have not been made for older people's needs" – older adults' perceptions of digital technologies in Finland and Ireland. Technol Soc 2020 Aug;62:101287. [doi: 10.1016/j.techsoc.2020.101287]
75. Pogorzelska-Maziarz M, Rising K, Gentsch A, et al. Home healthcare patient, caregiver and provider perspectives on use of unscheduled acute care and the usability and acceptability of on-demand telehealth solutions. Geriatr Nurs (Lond) 2021 Sep;42(5):1029-1034. [doi: 10.1016/j.gerinurse.2021.06.009]
76. Qi Tan AJ, Chua WL, McKenna L, Chin Tan LL, Lim YJ, Liaw SY. Enablers and barriers to nurse-facilitated geriatric teleconsultations in nursing homes: a qualitative descriptive multi-site study. Age Ageing 2022 Dec 5;51(12):afac268. [doi: 10.1093/ageing/afac268] [Medline: 36580553]
77. Sawik B, Tobis S, Baum E, et al. Robots for elderly care: review, multi-criteria optimization model and qualitative case study. In: Stachnik K, editor. Healthcare (Basel) 2023 Apr 30;11(9):1286. [doi: 10.3390/healthcare11091286] [Medline: 37174828]
78. Smith KM, Scerpella D, Guo A, et al. Perceived barriers and facilitators of implementing a multicomponent intervention to improve communication with older adults with and without dementia (SHARING Choices) in primary care: a qualitative

XSL•FO
RenderX

study. J Prim Care Community Health 2022;13:21501319221137251. [doi: 10.1177/21501319221137251] [Medline: 36398937]
79. Taylor A, Pizey H, Whittet C. Supporting safe and independent living: findings, themes and concept designs from a user-centred study on digital telecare. Interact Comput 2021 Jul 11;33(4):367-379. [doi: 10.1093/iwc/iwac006]
80. Thach KS, Lederman R, Waycott J. Key considerations for the design of technology for enrichment in residential aged care: an ethnographic study. In: Waycott J, editor. 2023 Apr 19 Presented at: CHI '23 URL: https://dl.acm.org/doi/proceedings/10.1145/3544548 [doi: 10.1145/3544548.3581176]
81. Thilo FJ, Bilger S, Halfens RJ, Schols JM, Hahn S. Involvement of the end user: exploration of older people's needs and preferences for a wearable fall detection device - a qualitative descriptive study. Patient Prefer Adherence 2017;11:11-22. [doi: 10.2147/PPA.S119177] [Medline: 28053509]
82. Tsertsidis A. Challenges in the provision of digital technologies to elderly with dementia to support ageing in place: a case study of a Swedish municipality. Disabil Rehabil Assist Technol 2021 Oct;16(7):758-768. [doi: 10.1080/17483107.2019.1710774] [Medline: 31913734]
83. Vaportzis E, Clausen MG, Gow AJ. Older adults perceptions of technology and barriers to interacting with tablet computers: a focus group study. Front Psychol 2017 Oct 4;8:1687. [doi: 10.3389/fpsyg.2017.01687] [Medline: 29071004]
84. Vergouw JW, Smits-Pelser H, Kars MC, et al. Needs, barriers and facilitators of older adults towards eHealth in general practice: a qualitative study. Prim Health Care Res Dev 2020 Dec 2;21:e54. [doi: 10.1017/S1463423620000547] [Medline: 33263272]
85. Wang S, Bolling K, Mao W, et al. Technology to Support Aging in Place: Older Adults' Perspectives Healthcare: MDPI; 2019. [doi: 10.3390/healthcare7020060]
86. Xiang X, Kayser J, Sun Y, Himle J. Internet-based psychotherapy intervention for depression among older adults receiving home care: qualitative study of participants' experiences. JMIR Aging 2021 Nov 22;4(4):e27630. [doi: 10.2196/27630] [Medline: 34813491]
87. Young SR, Lattie EG, Berry ABL, et al. Remote cognitive screening of healthy older adults for primary care with the MyCog mobile app: iterative design and usability evaluation. JMIR Form Res 2023 Jan 10;7:e42416. [doi: 10.2196/42416] [Medline: 36626223]
88. Cohen C, Kampel T, Verloo H. Acceptability of an intelligent wireless sensor system for the rapid detection of health issues: findings among home-dwelling older adults and their informal caregivers. Patient Prefer Adherence 2016;10:1687-1695. [doi: 10.2147/PPA.S113805] [Medline: 27660417]
89. Dosso JA, Kailley JN, Guerra GK, Robillard JM. Older adult perspectives on emotion and stigma in social robots. Front Psychiatry 2022;13:1051750. [doi: 10.3389/fpsyt.2022.1051750] [Medline: 36713914]
90. Dukhanin V, Wolff JL, Salmi L, et al. Co-designing an initiative to increase shared access to older adults' patient portals: stakeholder engagement. J Med Internet Res 2023 Nov 22;25:e46146. [doi: 10.2196/46146] [Medline: 37991827]
91. Gomez R, Habib A, Maidment DW, Ferguson MA. Smartphone-connected hearing aids enable and empower self-management of hearing loss: a qualitative interview study underpinned by the behavior change wheel. Ear Hear 2022;43(3):921-932. [doi: 10.1097/AUD.0000000000001143] [Medline: 34653030]
92. Hall A, Boulton E, Stanmore EJD. Older adults' perceptions of wearable technology hip protectors: implications for further research and development strategies. Disabil Rehabil Assist Technol 2019 Oct;14(7):663-668. [doi: 10.1080/17483107.2018.1491647] [Medline: 30371132]
93. Hawley CE, Wagner C, Venegas MD, et al. Connecting the disconnected: leveraging an in-home team member for video visits for older adults. J Am Geriatr Soc 2024 May;72(5):1408-1419. [doi: 10.1111/jgs.18663] [Medline: 37960887]
94. Hawley-Hague H, Tacconi C, Mellone S, et al. One-to-one and group-based teleconferencing for falls rehabilitation: usability, acceptability, and feasibility study. JMIR Rehabil Assist Technol 2021 Jan 12;8(1):e19690. [doi: 10.2196/19690] [Medline: 33433398]
95. Jahnke I, Riedel N, Popescu M, Skubic M, Rantz M. Social practices of nurse care coordination using sensor technologies - challenges with an alert system adoption in assisted living communities for older adults. Int J Nurs Sci 2021 Jul 10;8(3):289-297. [doi: 10.1016/j.ijnss.2021.05.011] [Medline: 34307777]
96. Kavčič M, Petrovčič A, Dolničar V. Model of acceptance and use of market-ready home-based e-care services: a qualitative study with care receivers and informal caregivers. Int J Human Comput Interact 2023 Mar 16;39(5):988-1002. [doi: 10.1080/10447318.2022.2041898]
97. Kononova A, Li L, Kamp K, et al. The use of wearable activity trackers among older adults: focus group study of tracker perceptions, motivators, and barriers in the maintenance stage of behavior change. JMIR Mhealth Uhealth 2019 Apr 5;7(4):e9832. [doi: 10.2196/mhealth.9832] [Medline: 30950807]
98. Munck B, Sandgren A, Fridlund B, Mårtensson J. Next-of-kin's conceptions of medical technology in palliative homecare. J Clin Nurs 2012 Jul;21(13-14):1868-1877. [doi: 10.1111/j.1365-2702.2012.04123.x] [Medline: 22582763]
99. Nymberg VM, Bolmsjö BB, Wolff M, Calling S, Gerward S, Sandberg M. "Having to learn this so late in our lives…" Swedish elderly patients' beliefs, experiences, attitudes and expectations of e-health in primary health care. Scand J Prim Health Care 2019 Mar;37(1):41-52. [doi: 10.1080/02813432.2019.1570612] [Medline: 30732519]

XSL•FO
RenderX

100. Parker SJ, Jessel S, Richardson JE, Reid MC. Older adults are mobile too!Identifying the barriers and facilitators to older adults' use of mHealth for pain management. BMC Geriatr 2013 May 6;13:1-8. [doi: 10.1186/1471-2318-13-43] [Medline: 23647949]
101. Shah MN, Morris D, Jones CMC, et al. A qualitative evaluation of a telemedicine-enhanced emergency care program for older adults. J Am Geriatr Soc 2013 Apr;61(4):571-576. [doi: 10.1111/jgs.12157] [Medline: 23452094]
102. Wong J, Young E, Hung L, Mann J, Jackson L. Beyond plan-do-study-act cycle - staff perceptions on facilitators and barriers to the implementation of telepresence robots in long-term care. BMC Health Serv Res 2023 Jul 19;23(1):772. [doi: 10.1186/s12913-023-09741-9] [Medline: 37468953]
103. Davenport RD, Mann W, Lutz B. How older adults make decisions regarding smart technology: an ethnographic approach. Assist Technol 2012;24(3):168-181. [doi: 10.1080/10400435.2012.659792] [Medline: 23033734]
104. Fothergill L, Holland C, Latham Y, Hayes N. Understanding the value of a proactive telecare system in supporting older adults' independence at home: qualitative interview study among key interest groups. J Med Internet Res 2023 Dec 14;25:e47997. [doi: 10.2196/47997] [Medline: 38096023]
105. Bin Noon G, Hanjahanja-Phiri T, Dave H, Fadrique L, Teague J, Morita PP. Exploring the role of active assisted living in the continuum of care for older adults: thematic analysis. JMIR Aging 2023 May 22;6(1):e40606. [doi: 10.2196/40606] [Medline: 37213201]
106. Waycott J, Zhao W, Kelly RM, Robertson E. Technology-mediated enrichment in aged care: survey and interview study. JMIR Aging 2022 Apr 12;5(2):e31162. [doi: 10.2196/31162] [Medline: 34975014]
107. Chaudhuri S, Kneale L, Le T, et al. Older adults' perceptions of fall detection devices. J Appl Gerontol 2017 Aug;36(8):915-930. [doi: 10.1177/0733464815591211] [Medline: 26112030]
108. Baker S, Waycott J, Robertson E, et al. Evaluating the use of interactive virtual reality technology with older adults living in residential aged care. Inf Process Manag 2020 May;57(3):102105. [doi: 10.1016/j.ipm.2019.102105]
109. Orellano-Colón EM, Mann WC, Rivero M, et al. Hispanic older adult's perceptions of personal, contextual and technology-related barriers for using assistive technology devices. J Racial Ethnic Health Disparities 2016 Dec;3(4):676-686. [doi: 10.1007/s40615-015-0186-8]
110. Mahmoudi Asl A, Toribio-Guzmán JM, Castro-González Á, Malfaz M, Salichs MA, Franco Martín M. Evaluating the user experience and usability of the MINI robot for elderly adults with mild dementia and mild cognitive impairment: insights and recommendations. Sensors (Basel) 2024 Nov 8;24(22):7180. [doi: 10.3390/s24227180] [Medline: 39598957]
111. Gris F, D'Amen B, Lamura G, Paciaroni L, Socci M, Melchiorre MG. Personalized technological support for informal caregivers of older people with dementia: a co-design approach involving potential end users and healthcare professionals in three focus groups in Italy. Healthcare (Basel) 2023 Sep 28;11(19):2640. [doi: 10.3390/healthcare11192640] [Medline: 37830677]
112. Read EA, Gagnon DA, Donelle L, et al. Stakeholder perspectives on in-home passive remote monitoring to support aging in place in the province of New Brunswick, Canada: rapid qualitative investigation. JMIR Aging 2022 May 11;5(2):e31486. [doi: 10.2196/31486] [Medline: 35544304]
113. Albers EA, Mikal J, Millenbah A, et al. The use of technology among persons with memory concerns and their caregivers in the United States during the COVID-19 pandemic: qualitative study. JMIR Aging 2022 Mar 17;5(1):e31552. [doi: 10.2196/31552] [Medline: 35134748]
114. Judson TJ, Subash M, Harrison JD, et al. Patient perceptions of e-visits: qualitative study of older adults to inform health system implementation. JMIR Aging 2023 May 26;6(1):e45641. [doi: 10.2196/45641] [Medline: 37234031]
115. Zsiga K, Edelmayer G, Rumeau P, Péter O, Tóth A, Fazekas G. Home care robot for socially supporting the elderly: focus group studies in three European countries to screen user attitudes and requirements. Int J Rehabil Res 2013 Dec;36(4):375-378. [doi: 10.1097/MRR.0b013e3283643d26] [Medline: 24189106]
116. Cimperman M, Brenčič MM, Trkman P, Stanonik MDL. Older adults' perceptions of home telehealth services. Telemed J E Health 2013 Oct;19(10):786-790. [doi: 10.1089/tmj.2012.0272] [Medline: 23931702]
117. Zafrani O, Nimrod G, Krakovski M, Kumar S, Bar-Haim S, Edan Y. Assimilation of socially assistive robots by older adults: an interplay of uses, constraints and outcomes. Front Robot AI 2024;11:1337380. [doi: 10.3389/frobt.2024.1337380] [Medline: 38646472]
118. Demiris G, Chaudhuri S, Thompson HJ. Older adults' experience with a novel fall detection device. Telemed J E Health 2016 Sep;22(9):726-732. [doi: 10.1089/tmj.2015.0218] [Medline: 26959299]
119. Bian C, Ye B, Hoonakker A, Mihailidis A. Attitudes and perspectives of older adults on technologies for assessing frailty in home settings: a focus group study. BMC Geriatr 2021 May 8;21(1):298. [doi: 10.1186/s12877-021-02252-4] [Medline: 33964887]
120. Pol M, van Nes F, van Hartingsveldt M, Buurman B, de Rooij S, Kröse B. Older people's perspectives regarding the use of sensor monitoring in their home: Table 1. GERONT 2016 Jun;56(3):485-493. [doi: 10.1093/geront/gnu104]
121. Conner M, Norman P. Health Behaviour: Current Issues and Challenges: Taylor & Francis; 2017:895-906. [doi: 10.1080/08870446.2017.1336240]
122. McCreadie C, Tinker A. The acceptability of assistive technology to older people. Ageing Soc 2005 May;25(1):91-110. [doi: 10.1017/S0144686X0400248X]

XSL•FO
RenderX

123. Bauer MS, Kirchner J. Implementation science: what is it and why should I care? Psychiatry Res 2020 Jan;283:112376. [doi: 10.1016/j.psychres.2019.04.025] [Medline: 31036287]
124. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. Implementation Sci 2022;17(1):75. [doi: 10.1186/s13012-022-01245-0]
125. Rangachari P, Mushiana SS, Herbert K. A scoping review of applications of the Consolidated Framework for Implementation Research (CFIR) to telehealth service implementation initiatives. BMC Health Serv Res 2022 Nov 30;22(1):1450. [doi: 10.1186/s12913-022-08871-w] [Medline: 36447279]
126. Aldridge G, Reupert A, Wu L, et al. Developing pre-implementation strategies for a co-designed, technology-assisted parenting intervention using the Consolidated Framework for Implementation Research (CFIR) and Expert Recommendations for Implementing Change (ERIC) approach. Int J Environ Res Public Health 2024 Nov 30;21(12):1599. [doi: 10.3390/ijerph21121599] [Medline: 39767440]
127. Gallant NL, Hadjistavropoulos T, Stopyn RJN, Feere EK. Integrating technology adoption models into implementation science methodologies: a mixed-methods preimplementation study. Gerontologist 2023 Mar 21;63(3):416-427. [doi: 10.1093/geront/gnac098] [Medline: 35810405]
128. Newton BJ, Rothlingova Z, Gutteridge R, LeMarchand K, Raphael JH. No room for reflexivity? Critical reflections following a systematic review of qualitative research. J Health Psychol 2012 Sep;17(6):866-885. [doi: 10.1177/1359105311427615] [Medline: 22147063]
129. Beets MW, von Klinggraeff L, Weaver RG, Armstrong B, Burkart S. Small studies, big decisions: the role of pilot/feasibility studies in incremental science and premature scale-up of behavioral interventions. Pilot Feasibility Stud 2021 Sep 10;7(1):173. [doi: 10.1186/s40814-021-00909-w] [Medline: 34507624]
130. Marangunić N, Granić A. Technology acceptance model: a literature review from 1986 to 2013. Univ Access Inf Soc 2015 Mar;14(1):81-95. [doi: 10.1007/s10209-014-0348-1]
131. Raskind IG, Shelton RC, Comeau DL, Cooper HLF, Griffith DM, Kegler MC. A review of qualitative data analysis practices in health education and health behavior research. Health Educ Behav 2019 Feb;46(1):32-39. [doi: 10.1177/1090198118795019] [Medline: 30227078]

Abbreviations

AT: assistive technologies
CASP: Critical Appraisal Skills Program
CFIR: Consolidated Framework for Implementation Research
PRISMA: Preferred Reporting Items for Systematic Review and Meta-Analyses
TAM: Technology Acceptance Model
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by P Jansons; submitted 13.03.25; peer-reviewed by P Sockolow, Y Wang; revised version received 14.09.25; accepted 16.09.25; published 27.11.25.

Please cite as:
Malden S, McGill K, Guthrie B, Frost H, Shenkin SD, Ezike A, Bareham BK, Mercer SW, Pearce C, Wilson C, Underwood I, Vines J, Lewis S, O'Donnell A
Patient and Carer-Related Facilitators and Barriers to the Adoption of Assistive Technologies for the Care of Older Adults: Systematic Review
JMIR Aging 2025;8:e73917
URL: https://aging.jmir.org/2025/1/e73917
doi:10.2196/73917

© Stephen Malden, Kris McGill, Bruce Guthrie, Helen Frost, Susan D Shenkin, Adanna Ezike, Beth Bareham, Stewart W Mercer, Caroline Pearce, Cara Wilson, Ian Underwood, John Vines, Sue Lewis, Amy O'Donnell. Originally published in JMIR Aging (https://aging.jmir.org), 27.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Applications of Machine Learning for Cognitive Health in Older Individuals With HIV: Rapid Systematic Review

Hwayoung Cho[1], PhD, RN; Jiyoun Song[2], PhD, AGACNP-BC, APRN; Hannah Cho[2], PhD, AGACNP-BC, APRN; Lin Li[3], MS; Renjie Liang[3], MS; Railton Miranda[1], MHS, RN; Qianqian Song[3], PhD; Jiang Bian[4,5], PhD

[1]College of Nursing, University of Florida, 1225 Center Dr., Gainesville, FL, United States

[2]School of Nursing, University of Pennsylvania, Philadelphia, PA, United States

[3]College of Medicine, University of Florida, Gainesville, FL, United States

[4]School of Medicine, Indiana University, Indianapolis, IN, United States

[5]Regenstrief Institute, Indianapolis, IN, United States

Corresponding Author:
Hwayoung Cho, PhD, RN
College of Nursing, University of Florida, 1225 Center Dr., Gainesville, FL, United States

Abstract

Background: More than half of people with HIV are now older than 50 years, and they face an approximately 60% higher risk of developing dementia compared with the general population. In recent years, the application of artificial intelligence, particularly machine learning, combined with the growing availability of large datasets, has opened new avenues for developing prediction models that improve dementia detection, monitoring, and management.

Objective: This systematic review aimed to synthesize the existing literature on the application of machine learning in dementia research among older people with HIV and identify directions for future research.

Methods: A comprehensive search was conducted in PubMed, CINAHL, and Embase in September 2024, limited to studies published within the past 10 years. Eligible articles included original research involving people with HIV applying at least 1 machine learning technique and reporting dementia-related outcomes.

Results: The search yielded 721 articles, of which 26 (3.6%) met the inclusion criteria. Most studies were retrospective and conducted in the United States (n=14, 53.8%), primarily focusing on neurocognitive impairment and HIV-associated neurocognitive disorders. Supervised machine learning techniques were most frequently used and demonstrated strong predictive performance. Common methodological challenges included small sample sizes, lack of external validation, limited participant diversity, and concerns about biological interpretability and generalizability.

Conclusions: Machine learning research on dementia among older people with HIV primarily targets HIV-associated neurocognitive disorders, with limited exploration of age-related neurodegenerative diseases such as Alzheimer disease and related dementias. The absence of longitudinal studies and external validation remains a key limitation. Future research should broaden the focus to all-cause dementia beyond HIV-specific conditions; apply advanced machine learning methods; and leverage large-scale longitudinal, multimodal datasets. Strengthening methodological rigor and enhancing real-world applications will be critical to improving early detection and effective management of cognitive health in this unique aging population.

(JMIR Aging 2025;8:e80433) doi:10.2196/80433

KEYWORDS

aging; HIV; dementia; cognitive impairment; machine learning; artificial intelligence; PRISMA; Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Introduction

Globally, there were approximately 39.9 million individuals living with HIV at the end of 2023 [1]. With advances in antiretroviral therapy, people with HIV are living longer, and more than 53% of them in the United States are now older than 50 years [2]. Although “older adults” commonly refers to individuals aged 65 years and above, research in the context of HIV often considers aging-related health concerns to emerge earlier due to accelerated and accentuated aging processes. There is a high frequency of neurocognitive decline reported in people with HIV, with approximately a 60% higher risk of developing dementia as they age compared with the general population [2]. Given that there is no cure for dementia, early detection and effective management are critical [3].

However, developing early diagnostic and effective management tools for dementia in this population presents unique challenges. People with HIV exhibit distinct neurological changes, which

XSL•FO
RenderX

include not only Alzheimer disease (AD) and AD-related dementias (ADRD; such as vascular dementia, Lewy body dementia, frontotemporal dementia, and mixed dementia) but also HIV-associated neurocognitive disorders and HIV-associated dementia. The complexity of HIV-related pathologies and comorbidities, along with inconsistent reporting, has posed significant challenges in HIV and aging research [4-6].

Despite this elevated risk, existing dementia detection and management strategies often fail to address the unique needs of people with HIV. In recent years, the application of artificial intelligence, particularly machine learning, combined with the increasing availability of large datasets, has opened new avenues for developing prediction models to improve dementia detection, progress monitoring, and management [3,7-11], offering promising alternatives for unraveling these complex relationships. Although the number of studies leveraging machine learning in predictive modeling of dementia outcomes in people with HIV has been growing, the scientific literature still lacks a synthesized review focused on this population. To address this gap, we conducted a rapid systematic review that followed the core principles of systematic review methodology but omitted a formal study quality appraisal to expedite evidence synthesis [12-14]. The objective of this review was to synthesize the existing literature that has applied machine learning in dementia research for people with HIV and to highlight directions for future research.

Methods

This systematic literature review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [15].

Data Sources and Search Strategy

A comprehensive search strategy was developed by the authors (Hwayoung Cho, LL, and RL) through a literature review, with assistance from a health sciences librarian at the academic institution where the study was conducted. Using 3 databases (ie, PubMed, CINAHL, and Embase), we searched the literature on September 24, 2024, with a 10-year publication limit to ensure inclusion of the most current advancements in artificial intelligence and machine learning applications in this area of research. To capture all relevant studies, additional hand searches were also conducted using reference lists obtained from the literature relevant to this review. The search terms used in the title, abstract, or keywords included the following themes: HIV infections; all-cause dementia including AD and ADRD, HIV-associated dementia, and HIV-associated neurocognitive disorders; machine learning approaches including artificial intelligence, machine learning, and deep learning. Medical Subject Headings (MeSH) terms for PubMed, CINAHL Headings for CINAHL, Emtree terms for Embase, and free-text terms were used with the Boolean searching technique. Details on the search strategies are presented in Multimedia Appendix 1.

Overview of Machine Learning Techniques

To provide context, we briefly summarize the most common machine learning techniques [11,16,17]. Supervised learning approaches, such as support vector machines (SVMs) [18,19], random forests, and logistic regression, use labeled data to classify or predict outcomes. Unsupervised methods, such as clustering, identify hidden patterns in unlabeled data. Semisupervised methods combine these two approaches, while deep learning techniques [20], including convolutional neural networks [21] and deep neural networks [22], use multilayered neural architectures to capture complex nonlinear patterns. These approaches are increasingly applied in health care research to improve diagnostic accuracy and risk prediction [11,16,17].

Study Selection and Eligibility Criteria

A web-based collaboration software platform that streamlines the production of systematic reviews, Covidence (Veritas Health Innovation), was used to facilitate the title and abstract screening, full-text review, data extraction, conflict resolution, and data verification. During each screening and review stage, titles and abstracts (LL, RL, RM, JS, and Hannah Cho) and full texts (JS and Hannah Cho) were independently assessed by two reviewers to determine study eligibility. To assess interrater reliability during the screening and review phases, the Cohen κ was calculated. Differences were resolved by consensus and, if necessary, by a sixth reviewer (Hwayoung Cho).

The inclusion criteria included an original research article that (1) focused on people with HIV as the target study population, (2) examined at least one dementia-related outcome, and (3) applied at least one machine learning technique and reported evaluation metrics (eg, area under the curve [AUC] of the receiver operating characteristic, sensitivity, specificity, precision, and F_1 -score). We excluded studies that examined only mild cognitive impairment without including any form of dementia, as our focus was on research specifically involving dementia-related conditions. We also excluded studies that were not available in the English language or full text, or gray literature (eg, clinical trials registries, conference abstracts, dissertations or theses, government reports, issues papers, letters, comments, editorials, correspondences, blogs, or newsletters).

Data Extraction and Synthesis

The following information was extracted from the studies included in the final review using a Microsoft Excel spreadsheet: last name of the first author, year of publication, location of the study conducted, study objective, study design, study sample (eg, demographics and sample size), data source, data period, and subtype of dementia focused. We also synthesized the following information: main aim of using machine learning, type of machine learning techniques, evaluation metrics used for machine learning models, key outcomes from machine learning models, and machine learning–related limitations reported in the article. Adopting machine learning breakdown frameworks [11,17], we subcategorized machine learning techniques by how they infer patterns from data into the following five types: (1) supervised, (2) unsupervised, (3) semisupervised (ie, a combination of supervised and unsupervised), (4) reinforcement learning, and (5) deep learning.

XSL•FO
RenderX

Consistent with rapid review methodology, we did not perform a formal study quality appraisal. Instead, we summarized the methodological challenges reported in the included studies. This approach aligns with the goal of rapid reviews to provide timely, systematic evidence synthesis using streamlined methods [12-14].

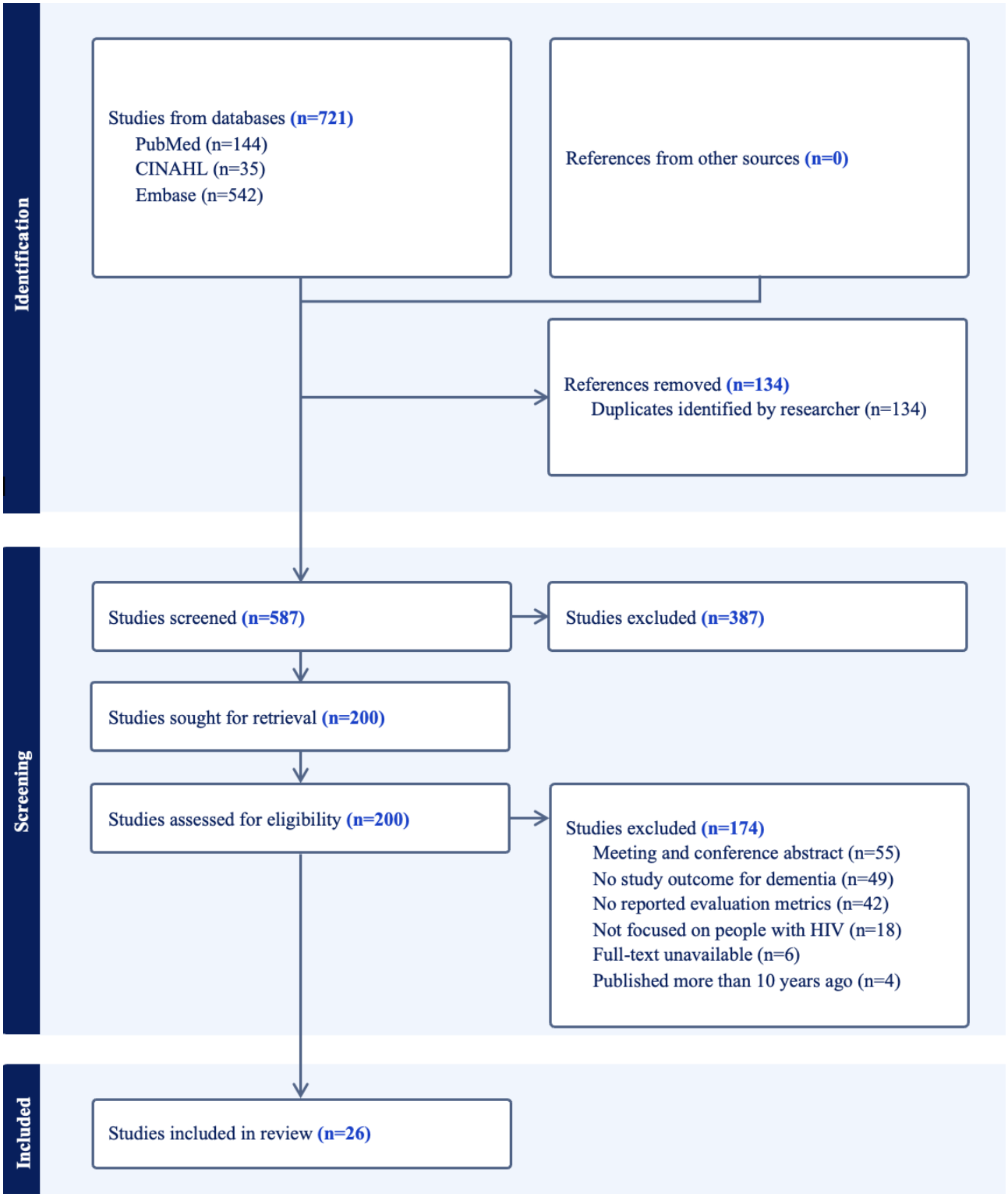
Results

Search Results

The initial searches resulted in 721 articles, including 144 (20%) from PubMed, 35 (4.9%) from CINAHL, and 542 (75.1%) from Embase. After removing duplicates, 587 (81.4%) articles were screened for titles and abstracts, and 200 (34.1%) articles were reviewed with full texts based on our eligibility criteria. We had high agreement between raters when making review decisions (Cohen $\kappa \geq 0.80$). A total of 26 (13%) studies that met the criteria were included in the review. The PRISMA flow diagram illustrating the review process is depicted in Figure 1.

XSL•FO
RenderX

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram of the study selection process.



Study Characteristics

An overview of study characteristics for a total of 26 studies is presented in Multimedia Appendix 2. Most studies were conducted in the United States (n=14, 53.9%) [23-36], followed by China (n=5, 19.2%) [37-41]; Canada (n=3, 11.5%) [42-44]; and other countries (n=4, 15.4%) including Colombia [45], Brazil [46], Ethiopia [47], and Japan [48].

The majority of studies used retrospective observational designs, including 3 (11.5%) cohort studies [25,27,42] and 2 (7.7%) cross-sectional [45,47] studies, while 21 (80.8%) did not specify their design. Data sources included neuroimaging (magnetic resonance imaging [MRI] and functional MRI [fMRI]) [26,28,29,31,32,34,39-41,44], neuropsychological tests, clinical and demographic variables from electronic health records and research cohorts [27,38,40,43,44,47], and other specialized data such as substance use questionnaires [25], the Center for Epidemiological Studies Depression Scale [25,46], and

XSL•FO
RenderX

biological samples from research programs [25,30,31]. Cognitive measures varied across studies and included the Mini-Mental State Examination [26,45], the International HIV Dementia Scale [47], self-reported memory assessments [25,31], and comprehensive neuropsychological test batteries [36,41,43].

The studies examined multiple forms of cognitive impairment among people with HIV, including neurocognitive impairment [26-28,38,42,46,47] and HIV-associated neurocognitive disorders [23-25,29,30,34,39,44,45,48], with particular attention to asymptomatic neurocognitive impairment as a subset of HIV-associated neurocognitive disorder [37,40,41]. Control groups included healthy individuals [26,28,35,38-40,48], HIV-negative cognitively normal participants [24,31,37,39], preclinical HIV-asymptomatic neurocognitive impairment cases [41], and neuropsychologically normal controls [23,30,33,43].

Sample sizes ranged from 9 to 8490 participants, with ages spanning 21 to 81 years. Almost half of the 26 studies (n=12, 46.2%) focused on middle-aged adults (40‐50 years) [24-27,32,33,36,42-46]. Most participants were male (>70%) [23,24,26,28,31-33,36,37,42,44,46]. Several studies (n=4, 15.4%) reported that participants were predominantly White, followed by African American or Black individuals [30-32,42]. Two (7.7%) US studies included primarily African American samples (82.6 % and 69.1%‐89.8%) [27,33]. Hispanic participants were represented in some studies (n=4, 15.4%) [23,30-32], typically comprising 4.8% to 24% of samples [23,31]. A few studies (n=3, 11.5%) reported Asian or other racial groups, but these consistently accounted for less than 5% of participants [23,31,32].

Key Findings on Machine Learning

An overview of main findings from the 26 studies is presented in Multimedia Appendix 3, summarizing machine learning applications in dementia research for people with HIV, the techniques and model performance used, and the limitations of these applications reported across studies.

Applications for Machine Learning in Dementia Research Among People With HIV

In this review, machine learning was applied to improve the diagnosis, prevention, and management of dementia among older people with HIV across studies. Of the 26 studies, some (n=12, 46.2%) applied machine learning algorithms to enhance diagnostic accuracy by analyzing neuroimaging data, cognitive assessments, and biomarkers (including 1, 3.8% study focusing on genetic features [48]) to identify early signs of dementia in people with HIV [23,28-30,33,35,36,41,43,46-48], while other studies (n=2, 7.7%) used machine learning to monitor disease progression, identify symptom clusters, and optimize tailored treatment plans for people with HIV [38,45].

Machine Learning Techniques and Model Performance

Various machine learning techniques were identified across the 26 included studies and classified into supervised, unsupervised, semisupervised (a combination of supervised and unsupervised), reinforcement learning, and deep learning [11,17]. Supervised machine learning was the most prevalent approach, applied in 20 (77%) studies [23,24,26,28-33,35-37,39-43,45-47]. Only 1 (3.8%) study applied an unsupervised machine learning approach [38], while a combination of supervised and unsupervised methods was used in 5 (19.2%) studies [25,27,34,44,48]. Deep learning was used in 2 (7.7%) studies [35,36], whereas reinforcement learning was not reported in any of the included studies.

Supervised machine learning techniques included SVMs, random forests, logistic regression, ensemble models (eg, Adaptive Boosting and Extreme Gradient Boosting), and feature selection methods such as least absolute shrinkage and selection operator and classification and regression trees [23,24,26,28-33,35-37,39-43,45-47]. SVMs were the most frequently applied algorithm, used in 7 (26.9%) studies [24,30,32,33,39,41,48], with reported AUC values up to 0.85 and classification accuracy as high as 82% [24]. Logistic regression was also implemented in 7 (26.9%) studies [25,26,28,37,43,46,47], achieving classification accuracy between 68% and 90% [43]. Random forests were used in 5 (19.2%) studies [31,42,43,46,48], demonstrating strong predictive performance with AUC values up to 0.87 and accuracy levels exceeding 80% [46]. Ensemble models, such as Adaptive Boosting, showed robust performance, with precision and recall scores of 0.80 and 0.77, respectively [46].

Unsupervised learning techniques, including k-means clustering, hierarchical clustering, and mutual connectivity analysis, were primarily used to stratify patients into subgroups and identify distinct cognitive profiles. For example, k-means clustering was applied to classify HIV-associated neurocognitive disorder subtypes and demonstrated strong performance in identifying connectivity profiles, achieving AUC values up to 0.89 [34].

Deep learning techniques were applied in only a few (n=2, 7.7%) studies [35,36]. Convolutional neural networks demonstrated classification accuracies surpassing 90% across various domains [35], while deep neural networks achieved an accuracy of 82% for cognitive impairment and 75% for frailty classification [36].

Internal validation methods, such as k-fold cross-validation, were implemented in 18 studies (69.2%) [24,26,28-36,39-41,43,45,46,48]. In contrast, 6 (23.1%) studies did not explicitly report validation methods [23,25,37,42,44,47], and none included external validation.

Limitations of Machine Learning in Dementia Research Among People With HIV

The researchers in the included studies acknowledged several methodological challenges. A predominant limitation across studies was small sample size [26,30-32,35-37,39-43,48]. Additionally, study cohorts were largely male, limiting generalizability [23,26,40,43]. Some (4/26, 15.4%) studies highlighted unaddressed confounding variables, such as recreational drug use [37,47], antiretroviral therapy resistance [42], and comorbidities, including hepatitis C virus coinfection [42] and depression [23]. Overlapping variables between HIV-associated neurocognitive disorders and other conditions were also reported [43], along with potential overdiagnosis due to scoring methods such as the Global Deficit Score [33].

The “black box” nature of machine learning models raised concerns about clinical relevance and whether predictive

XSL•FO
RenderX

accuracy should outweigh biological plausibility [24,33]. For example, brain connectivity patterns in fMRI data [24] and least absolute shrinkage and selection operator-selected MRI features [33] lacked clear neuropathological explanations, undermining interpretability. Finally, some (3/26, 11.5%) studies did not explicitly acknowledge limitations [28,34,46], potentially omitting critical considerations.

Discussion

Principal Findings

This is the first systematic review to examine how machine learning has been applied to dementia-related outcomes in people with HIV, addressing a critical gap in the literature at the intersection of aging, HIV, cognitive decline, and artificial intelligence. Historically, research on cognitive decline in HIV has centered on HIV-associated neurocognitive disorders, including HIV-associated dementia—a form of neurocognitive impairment resulting from the direct effects of HIV on the brain [4,49-56]. This emphasis largely reflected the shorter life expectancy in the early antiretroviral era, when age-related dementias were less prevalent.

We identified 26 studies, most of which focused on neurocognitive impairment related to HIV-associated neurocognitive disorders, with few investigating age-related dementias such as AD and ADRD. This imbalance underscores a major gap in the current literature and highlights the need for future research exploring the full spectrum of cognitive decline among older people with HIV. Although “older adults” is a common term in aging research, many included studies focused on middle-aged populations, reflecting the earlier onset of aging-related concerns in the context of HIV. As people with HIV continue to age, our findings highlight the need for future research to explore the intersection of HIV, aging, and neurodegenerative conditions beyond HIV-associated neurocognitive disorders.

A rigorous study design is essential when applying machine learning techniques in health care research. Most included studies were retrospective or cross-sectional, limiting understanding of disease progression. Because most studies used cross-sectional designs, they were unable to model change over time or to distinguish transient from persistent cognitive impairments. Such designs cannot capture how small errors in measurement or prediction may accumulate and influence dementia trajectories. Future research should use longitudinal designs and temporal modeling approaches to track cognitive changes and evaluate progression risk among older people with HIV.

In addition, retrospective and cross-sectional study designs restrict the ability of current machine learning models to account for real-world contextual factors such as environmental influences, day-to-day variability in cognitive performance, social interactions, and cultural factors that may affect how dementia symptoms manifest or are reported in people with HIV [57,58]. Without longitudinal or ecologically valid data, current machine learning models cannot fully capture the dynamic and multidimensional nature of cognitive changes over time [59,60]. Moreover, most studies did not incorporate contextual or social determinants of health, such as social support, treatment adherence, living conditions, or daily activity patterns, even though these factors are known to influence cognitive trajectories and dementia risk among older people with HIV. Integrating such variables into future machine learning frameworks could enhance the ecological validity and predictive performance of dementia models [61,62]. Longitudinal datasets are also critical for developing models that predict dementia onset or progression [63], as they enable temporal analyses and improve generalizability. Leveraging multimodal data, such as MRI and fMRI, laboratory tests, and neuropsychological assessments across time, can provide a comprehensive view of brain health and facilitate early detection of dementia among older people with HIV.

Integrating biomarker and genetic data with clinical and neuroimaging features offers promise for enhancing prediction of dementia outcomes [64-67]. Of the studies included in this systematic review, only 1 used machine learning to identify key genetic features predictive of HIV-associated neurocognitive disorder status and proposed a framework for biomarker development [48]. Future research should integrate biomarker and genomic information to improve predictive performance and support clinical applications.

Supervised learning methods, including SVMs, logistic regression, and random forests, were the most frequently used and generally demonstrated strong predictive performance across studies. However, despite their high accuracy, supervised machine learning models often require large datasets and substantial computing resources that may limit their scalability and generalizability, particularly in resource-constrained settings or smaller research cohorts [68]. Unsupervised and semisupervised methods (eg, k-means clustering and mutual connectivity analysis) were less common but useful for identifying subgroups and latent cognitive profiles (eg, brain connectivity profiles) [34]. Deep learning methods, although applied in only a few studies, showed promising performance in classifying cognitive impairment and related outcomes. No studies used reinforcement learning, representing an opportunity for future work. As datasets grow in size and complexity, advanced approaches such as deep and reinforcement learning could enhance early detection and personalized risk assessment for dementia in people with HIV [69]. Lessons from these machine learning–based dementia studies may also inform broader efforts to develop scalable, multimodal prediction tools for older populations with complex comorbidities. These approaches can extend beyond HIV-specific contexts to other aging populations facing overlapping challenges in cognitive health and care delivery.

A key limitation identified in this review was the lack of external validation across studies. Although most studies used internal validation methods (eg, k-fold cross-validation), few tested models on external datasets, limiting generalizability and real-world applicability [70]. Future research should prioritize external validation to enhance reproducibility and clinical relevance of machine learning–based dementia prediction in people with HIV.

XSL•FO
RenderX

Limitations

Our systematic review has several limitations. First, although a comprehensive search strategy was used across multiple databases, it is possible that relevant studies were missed due to publication bias or indexing issues (eg, the limited timeframe, exclusion of non-English publications, or exclusion of gray literature). Second, despite the systematic screening and review process conducted independently by multiple reviewers, the subjective nature of eligibility assessment might introduce potential reviewer bias. To mitigate this, we assessed interrater reliability using the Cohen κ and achieved strong agreement. We also did not perform a formal study quality appraisal; rather, we reported the methodological limitations noted by individual studies. While this approach aligns with rapid review guidance [12-14], it limits formal assessment of study quality. Several included studies relied on self-reported measures, such as subjective memory complaints, which are inherently susceptible to recall bias and response error. These sources of measurement bias were not explicitly addressed or adjusted for in the reviewed studies, limiting the interpretability of self-reported cognitive outcomes. Finally, while our review aimed to categorize and evaluate machine learning techniques used for dementia-related outcomes in people with HIV, the rapid advancements in machine learning may mean that recently published or ongoing studies were not captured at the time of our search; thus, continuous updating is needed in future systematic reviews.

Conclusions

This review examined how machine learning methods have been applied in dementia research for people with HIV, summarizing the techniques used, their strengths and limitations, and practical implications for future research. Despite promising predictive performance, most studies used supervised machine learning methods and lacked external validation, which limits their generalizability. Future dementia research in HIV could benefit from the adoption of advanced machine learning methods. The dominance of studies focused on HIV-associated neurocognitive disorders, with little attention to age-related neurodegenerative diseases such as AD and ADRD, underscores a critical gap in literature. As the population of people with HIV ages, there is a need to expand longitudinal cohort studies using large-scale real-world data that integrate multimodal information from multiple sources, including clinical covariates, neuroimaging, cognitive assessments, and genetic data, to capture disease progression, identify early biomarkers, and enable personalized risk assessment.

Future machine learning research in this area should ensure methodological rigor, the inclusion of diverse data sources, and external validation to enhance clinical applicability and ultimately improve dementia-related outcomes in older people with HIV. By highlighting current trends and research gaps, this review provides a foundation for advancing machine learning–driven cognitive health research to improve early detection and management of all-cause dementia in older people with HIV. These insights may also inform broader efforts to enhance neurodegenerative outcomes in aging populations through informatics and digital health tools.

Funding

This study was supported by the National Institute on Aging under grant K01AG090118 to the principal investigator Hwayoung Cho; the National Institute of Mental Health under grant R34MH135768 to multiple principal investigators Hwayoung Cho and Yiyang Liu, PhD; the National Heart, Lung, and Blood Institute under grant 1K99HL169940 to principal investigator JS; and the National Institute of General Medical Sciences under grant R35GM151089 to principal investigator QS of the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability

All data generated or analyzed during this study are included in this published paper and its multimedia appendices.

Authors' Contributions

Hwayoung Cho, LL, and RL were involved in the study concept and design. Hwayoung Cho, LL, RL, RM, JS, and Hannah Cho performed data acquisition and analysis through the systematic review process. All authors contributed equally to later manuscript drafts and agreed to the final submitted version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategies.
[DOCX File, 16 KB - aging_v8i1e80433_app1.docx]

Multimedia Appendix 2
Study characteristics.

XSL•FO
RenderX

[DOCX File, 68 KB - aging_v8i1e80433_app2.docx]

Multimedia Appendix 3
Machine learning application.
[DOCX File, 56 KB - aging_v8i1e80433_app3.docx]

Checklist 1
PRISMA checklist.
[PDF File, 166 KB - aging_v8i1e80433_app4.pdf]

References

1. HIV and AIDS. World Health Organization. 2025 Jul 15. URL: https://www.who.int/news-room/fact-sheets/detail/hiv-aids#:~:text=Key%20facts,in%20the%20WHO%20African%20Region [accessed 2025-12-11]
2. Lam JO, Hou CE, Hojilla JC, et al. Comparison of dementia risk after age 50 between individuals with and without HIV infection. AIDS 2021 Apr 1;35(5):821-828. [doi: 10.1097/QAD.0000000000002806] [Medline: 33394681]
3. Robinson L, Tang E, Taylor JP. Dementia: timely diagnosis and early intervention. BMJ 2015 Jun 16;350:h3029. [doi: 10.1136/bmj.h3029] [Medline: 26079686]
4. Heaton RK, Clifford DB, Franklin DR Jr, et al. HIV-associated neurocognitive disorders persist in the era of potent antiretroviral therapy: CHARTER Study. Neurology (ECronicon) 2010 Dec 7;75(23):2087-2096. [doi: 10.1212/WNL.0b013e318200d727] [Medline: 21135382]
5. Milanini B, Valcour V. Differentiating HIV-associated neurocognitive disorders from Alzheimer's disease: an emerging issue in geriatric NeuroHIV. Curr HIV/AIDS Rep 2017 Aug;14(4):123-132. [doi: 10.1007/s11904-017-0361-0] [Medline: 28779301]
6. Robertson K, Liner J, Meeker RB. Antiretroviral neurotoxicity. J Neurovirol 2012 Oct;18(5):388-399. [doi: 10.1007/s13365-012-0120-3] [Medline: 22811264]
7. Tomašev N, Glorot X, Rae JW, et al. A clinically applicable approach to continuous prediction of future acute kidney injury. Nature New Biol 2019 Aug;572(7767):116-119. [doi: 10.1038/s41586-019-1390-1] [Medline: 31367026]
8. Mitelpunkt A, Galili T, Kozlovski T, et al. Novel Alzheimer's disease subtypes identified using a data and knowledge driven strategy. Sci Rep 2020 Jan 28;10(1):1327. [doi: 10.1038/s41598-020-57785-2] [Medline: 31992745]
9. Kumar S, Oh I, Schindler S, Lai AM, Payne PRO, Gupta A. Machine learning for modeling the progression of Alzheimer disease dementia using clinical data: a systematic literature review. JAMIA Open 2021 Jul;4(3):ooab052. [doi: 10.1093/jamiaopen/ooab052] [Medline: 34350389]
10. Grueso S, Viejo-Sobera R. Machine learning methods for predicting progression from mild cognitive impairment to Alzheimer's disease dementia: a systematic review. Alzheimers Res Ther 2021 Sep 28;13(1):162. [doi: 10.1186/s13195-021-00900-w] [Medline: 34583745]
11. Bucholc M, James C, Al Khleifat A, et al. Artificial intelligence for dementia research methods optimization. ArXiv. 2023 Mar 2. [doi: 10.1002/alz.13441] [Medline: 36911275]
12. Tricco AC, Langlois EV, Straus SE. Rapid reviews to strengthen health policy and systems: a practical guide. World Health Organization. 2017. URL: https://wkc.who.int/resources/publications/i/item/2017-08-10-rapid-reviews-to-strengthen-health-policy-and-systems-a-practical-guide [accessed 2025-12-11]
13. Stevens A, Hersi M, Garritty C, et al. Rapid review method series: interim guidance for the reporting of rapid reviews. BMJ Evid Based Med 2025 Mar 21;30(2):118-123. [doi: 10.1136/bmjebm-2024-112899] [Medline: 39038926]
14. Devane D, Hamel C, Gartlehner G, et al. Key concepts in rapid reviews: an overview. J Clin Epidemiol 2024 Nov;175:111518. [doi: 10.1016/j.jclinepi.2024.111518] [Medline: 39245414]
15. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Rev Esp Cardiol (Engl Ed) 2021 Sep;74(9):790-799. [doi: 10.1016/j.rec.2021.07.010] [Medline: 34446261]
16. Sarker IH. Machine learning: algorithms, real-world applications and research directions. SN Comput Sci 2021;2(3):160. [doi: 10.1007/s42979-021-00592-x] [Medline: 33778771]
17. Matheny M, Thadaney Israni S, Ahmed M, Whicher D. Artificial Intelligence in Health Care: The Hope, the Hype, the Promise, the Peril: National Academy of Medicine; 2022. URL: https://nam.edu/wp-content/uploads/2019/12/AI-in-Health-Care-PREPUB-FINAL.pdf [accessed 2025-12-11]
18. Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995;20:273-297. [doi: 10.1007/BF00994018]
19. Guido R, Ferrisi S, Lofaro D, Conforti D. An overview on the advancements of support vector machine models in healthcare applications: a review. Information 2024;15(4):235. [doi: 10.3390/info15040235]
20. Miotto R, Wang F, Wang S, Jiang X, Dudley JT. Deep learning for healthcare: review, opportunities and challenges. Brief Bioinform 2018 Nov 27;19(6):1236-1246. [doi: 10.1093/bib/bbx044] [Medline: 28481991]
21. Alzubaidi L, Zhang J, Humaidi AJ, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data 2021;8(1):53. [doi: 10.1186/s40537-021-00444-8] [Medline: 33816053]

XSL•FO
RenderX

22. Mall PK, Singh PK, Srivastav S, et al. A comprehensive review of deep neural networks for medical image processing: recent developments and future opportunities. Healthc Anal 2023 Dec;4:100216. [doi: 10.1016/j.health.2023.100216]
23. Anderson AE, Jones JD, Thaler NS, Kuhn TP, Singer EJ, Hinkin CH. Intraindividual variability in neuropsychological performance predicts cognitive decline and death in HIV. Neuropsychology 2018 Nov;32(8):966-972. [doi: 10.1037/neu0000482] [Medline: 30211610]
24. Chockanathan U, DSouza AM, Abidin AZ, Schifitto G, Wismüller A. Automated diagnosis of HIV-associated neurocognitive disorders using large-scale Granger causality analysis of resting-state functional MRI. Comput Biol Med 2019 Mar;106:24-30. [doi: 10.1016/j.compbiomed.2019.01.006] [Medline: 30665138]
25. Levine AJ, Martin E, Sacktor N, Munro C, Becker J, Multicenter AIDS Cohort Study-Neuropsychology Working Group. Predictors and impact of self-reported suboptimal effort on estimates of prevalence of HIV-associated neurocognitive disorders. J Acquir Immune Defic Syndr 2017 Jun 1;75(2):203-210. [doi: 10.1097/QAI.0000000000001371] [Medline: 28328547]
26. Narvid J, McCoy D, Dupont SM, et al. Abnormal cerebral perfusion profile in older adults with HIV-associated neurocognitive disorder: discriminative power of arterial spin-labeling. AJNR Am J Neuroradiol 2018 Dec;39(12):2211-2217. [doi: 10.3174/ajnr.A5902] [Medline: 30467218]
27. Yang X, Zhang J, Chen S, Weissman S, Olatosi B, Li X. Comorbidity patterns among people living with HIV: a hierarchical clustering approach through integrated electronic health records data in South Carolina. AIDS Care 2021 May;33(5):594-606. [doi: 10.1080/09540121.2020.1844864] [Medline: 33172284]
28. Zhang Y, Kwon D, Esmaeili-Firidouni P, et al. Extracting patterns of morphometry distinguishing HIV associated neurodegeneration from mild cognitive impairment via group cardinality constrained classification. Hum Brain Mapp 2016 Dec;37(12):4523-4538. [doi: 10.1002/hbm.23326] [Medline: 27489003]
29. DSouza AM, Abidin AZ, Leistritz L, Wismüller A. Identifying HIV associated neurocognitive disorder using large-scale granger causality analysis on resting-state functional MRI. Proc SPIE Int Soc Opt Eng 2017 Feb;10133:101330M. [doi: 10.1117/12.2254690] [Medline: 29167591]
30. Pulliam L, Liston M, Sun B, Narvid J. Using neuronal extracellular vesicles and machine learning to predict cognitive deficits in HIV. J Neurovirol 2020 Dec;26(6):880-887. [doi: 10.1007/s13365-020-00877-6] [Medline: 32681213]
31. Finkelstein A, Faiyaz A, Weber MT, et al. Fixel-based analysis and free water corrected DTI evaluation of HIV-associated neurocognitive disorders. Front Neurol 2021 Nov 4;12:725059. [doi: 10.3389/fneur.2021.725059] [Medline: 34803875]
32. Underwood J, Cole JH, Leech R, Sharp DJ, Winston A, CHARTER group. Multivariate pattern analysis of volumetric neuroimaging data and its relationship with cognitive function in treated HIV disease. J Acquir Immune Defic Syndr 2018;78(4):429-436. [doi: 10.1097/QAI.0000000000001687] [Medline: 29608444]
33. Xu Y, Lin Y, Bell RP, et al. Machine learning prediction of neurocognitive impairment among people with HIV using clinical and multimodal magnetic resonance imaging data. J Neurovirol 2021 Feb;27(1):1-11. [doi: 10.1007/s13365-020-00930-4] [Medline: 33464541]
34. DSouza AM, Abidin AZ, Wismüller A. Investigating changes in resting-state connectivity from functional MRI data in patients with HIV associated neurocognitive disorder using MCA and machine learning. Proc SPIE Int Soc Opt Eng 2017 Mar 13;10137:101371C. [doi: 10.1117/12.2254189] [Medline: 29170578]
35. Zhang J, Zhao Q, Adeli E, et al. Multi-label, multi-domain learning identifies compounding effects of HIV and cognitive impairment. Med Image Anal 2022 Jan;75:102246. [doi: 10.1016/j.media.2021.102246] [Medline: 34706304]
36. Luckett P, Paul RH, Navid J, et al. Deep learning analysis of cerebral blood flow to identify cognitive impairment and frailty in persons living with HIV. J Acquir Immune Defic Syndr 2019 Dec 15;82(5):496-502. [doi: 10.1097/QAI.0000000000002181] [Medline: 31714429]
37. Zhan Y, Cai DC, Liu Y, et al. Altered metabolism in right basal ganglia associated with asymptomatic neurocognitive impairment in HIV-infected individuals. Heliyon 2023 Dec 6;10(1):e23342. [doi: 10.1016/j.heliyon.2023.e23342] [Medline: 38169709]
38. Zhu Z, Hu Y, Xing W, et al. Identifying symptom clusters among people living with HIV on antiretroviral therapy in China: a network analysis. J Pain Symptom Manage 2019 Mar;57(3):617-626. [doi: 10.1016/j.jpainsymman.2018.11.011] [Medline: 30465875]
39. Fu D, Mo K, Deng W, et al. Application value of machine learning method in measuring gray matter volume of AIDS patients. Dis Markers 2022 Jun 16;2022:1210002. [doi: 10.1155/2022/1210002] [Medline: 35756486]
40. Qi Y, Wang W, Rao B, et al. Value of radiomic analysis combined with diffusion tensor imaging in early diagnosis of HIV-associated neurocognitive disorders. J Magn Reson Imaging 2023 Dec;58(6):1882-1891. [doi: 10.1002/jmri.28741] [Medline: 37118972]
41. Qi Y, Xu M, Wang W, et al. Early prediction of putamen imaging features in HIV-associated neurocognitive impairment syndrome. BMC Neurol 2021 Mar 9;21(1):106. [doi: 10.1186/s12883-021-02114-x] [Medline: 33750319]
42. Amusan P, Power C, Gill MJ, et al. Lifetime antiretroviral exposure and neurocognitive impairment in HIV. J Neurovirol 2020 Oct;26(5):743-753. [doi: 10.1007/s13365-020-00870-z] [Medline: 32720232]

XSL•FO
RenderX

43. Tu W, Chen PA, Koenig N, et al. Machine learning models reveal neurocognitive impairment type and prevalence are associated with distinct variables in HIV/AIDS. J Neurovirol 2020 Feb;26(1):41-51. [doi: 10.1007/s13365-019-00791-6] [Medline: 31520320]
44. Gomez D, Power C, Gill MJ, Koenig N, Vega R, Fujiwara E. Empiric neurocognitive performance profile discovery and interpretation in HIV infection. J Neurovirol 2019 Feb;25(1):72-84. [doi: 10.1007/s13365-018-0685-6] [Medline: 30519968]
45. Martinez-Banfi M, Vélez JI, Mebarak Chams MR, et al. Utility of a short neuropsychological protocol for detecting HIV-associated neurocognitive disorders in patients with asymptomatic HIV-1 infection. Brain Sci 2021 Aug 4;11(8):1037. [doi: 10.3390/brainsci11081037] [Medline: 34439656]
46. Pinheiro LI, Pereira ML, Fernandez MP, Filho FM, de Abreu WJ, Pinheiro PG. Application of data mining algorithms for dementia in people with HIV/AIDS. Comput Math Methods Med 2021 Jul 9;2021:4602465. [doi: 10.1155/2021/4602465] [Medline: 34335861]
47. Salahuddin M, Manzar MD, Hassen HY, et al. Prevalence and predictors of neurocognitive impairment in Ethiopian population living with HIV. HIV AIDS (Auckl) 2020;12:559-572. [doi: 10.2147/HIV.S260831] [Medline: 33116918]
48. Ogishi M, Yotsuyanagi H. Prediction of HIV-associated neurocognitive disorder (HAND) from three genetic features of envelope gp120 glycoprotein. Retrovirology (Auckl) 2018 Jan 27;15(1):12. [doi: 10.1186/s12977-018-0401-x] [Medline: 29374475]
49. Yang HY, Beymer MR, Suen SC. Chronic disease onset among people living with HIV and AIDS in a large private insurance claims dataset. Sci Rep 2019 Dec 6;9(1):18514. [doi: 10.1038/s41598-019-54969-3] [Medline: 31811207]
50. Alford K, Vera JH. Cognitive impairment in people living with HIV in the ART era: a review. Br Med Bull 2018 Sep 1;127(1):55-68. [doi: 10.1093/bmb/ldy019] [Medline: 29868901]
51. Wei J, Hou J, Su B, et al. The prevalence of Frascati-criteria-based HIV-associated neurocognitive disorder (HAND) in HIV-infected adults: a systematic review and meta-analysis. Front Neurol 2020;11:581346. [doi: 10.3389/fneur.2020.581346] [Medline: 33335509]
52. Heaton RK, Franklin DR, Ellis RJ, et al. HIV-associated neurocognitive disorders before and during the era of combination antiretroviral therapy: differences in rates, nature, and predictors. J Neurovirol 2011 Feb;17(1):3-16. [doi: 10.1007/s13365-010-0006-1] [Medline: 21174240]
53. Robertson KR, Smurzynski M, Parsons TD, et al. The prevalence and incidence of neurocognitive impairment in the HAART era. AIDS 2007 Sep 12;21(14):1915-1921. [doi: 10.1097/QAD.0b013e32828e4e27] [Medline: 17721099]
54. Cysique LA, Heaton RK, Kamminga J, et al. HIV-associated neurocognitive disorder in Australia: a case of a high-functioning and optimally treated cohort and implications for international neuroHIV research. J Neurovirol 2014 Jun;20(3):258-268. [doi: 10.1007/s13365-014-0242-x] [Medline: 24696363]
55. Sacktor N, Skolasky RL, Seaberg E, et al. Prevalence of HIV-associated neurocognitive disorders in the Multicenter AIDS Cohort Study. Neurology 2016 Jan 26;86(4):334-340. [doi: 10.1212/WNL.0000000000002277] [Medline: 26718568]
56. Makinson A, Dubois J, Eymard-Duvernay S, et al. Increased prevalence of neurocognitive impairment in aging people living with human immunodeficiency virus: the ANRS EP58 HAND 55-70 study. Clin Infect Dis 2020 Jun 10;70(12):2641-2648. [doi: 10.1093/cid/ciz670] [Medline: 31755936]
57. Birbeck GL, Kvalsund MP, Byers PA, et al. Neuropsychiatric and socioeconomic status impact antiretroviral adherence and mortality in rural Zambia. Am J Trop Med Hyg 2011 Oct;85(4):782-789. [doi: 10.4269/ajtmh.2011.11-0187] [Medline: 21976587]
58. Eshun-Wilson I, Ford N, Mody A, et al. Strengthening implementation guidelines for HIV service delivery: considerations for future evidence generation and synthesis. PLoS Med 2023 Mar;20(3):e1004168. [doi: 10.1371/journal.pmed.1004168] [Medline: 36877738]
59. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd edition 2009. [doi: 10.1007/978-0-387-84858-7]
60. Vasilenko SA, Piper ME, Lanza ST, Liu X, Yang J, Li R. Time-varying processes involved in smoking lapse in a randomized trial of smoking cessation therapies. Nicotine Tob Res 2014 May;16 Suppl 2(Suppl 2):S135-S143. [doi: 10.1093/ntr/ntt185] [Medline: 24711627]
61. Webb EK, Weis CN, Huggins AA, et al. Neighborhood disadvantage is associated with stable deficits in neurocognitive functioning in traumatically-injured adults. Health Place 2021 Jan;67:102493. [doi: 10.1016/j.healthplace.2020.102493] [Medline: 33321457]
62. Reed RG, Hillmann AR. Neighborhood-level socioeconomic disadvantage is associated with multiple cognitive domains in a community sample of older adults. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2025 Jul;32(4):621-635. [doi: 10.1080/13825585.2025.2454517] [Medline: 39825636]
63. DSouza AM, Abidin AZ, Schifitto G, Wismüller A. A multivoxel pattern analysis framework with mutual connectivity analysis investigating changes in resting state connectivity in patients with HIV associated neurocognitve disorder. Magn Reson Imaging 2019 Oct;62:121-128. [doi: 10.1016/j.mri.2019.06.001] [Medline: 31189074]
64. Gatz M, Reynolds CA, Fratiglioni L, et al. Role of genes and environments for explaining Alzheimer disease. Arch Gen Psychiatry 2006 Feb;63(2):168-174. [doi: 10.1001/archpsyc.63.2.168] [Medline: 16461860]

XSL•FO
RenderX

65. Sims R, Hill M, Williams J. The multiplex model of the genetics of Alzheimer's disease. Nat Neurosci 2020 Mar;23(3):311-322. [doi: 10.1038/s41593-020-0599-5] [Medline: 32112059]
66. Karlsson IK, Escott-Price V, Gatz M, et al. Measuring heritable contributions to Alzheimer's disease: polygenic risk score analysis with twins. Brain Commun 2022;4(1):fcab308. [doi: 10.1093/braincomms/fcab308] [Medline: 35169705]
67. Rousset RZ, den Braber A, Verberk IM, et al. Heritability of Alzheimer's disease plasma biomarkers: a nuclear twin family design. Alzheimers Dement 2025 Jan;21(1):e14269. [doi: 10.1002/alz.14269] [Medline: 39588748]
68. Yang X, Chen A, PourNejatian N, et al. A large language model for electronic health records. NPJ Digit Med 2022 Dec 26;5(1):194. [doi: 10.1038/s41746-022-00742-2] [Medline: 36572766]
69. Monshi MMA, Poon J, Chung V. Deep learning in generating radiology reports: a survey. Artif Intell Med 2020 Jun;106:101878. [doi: 10.1016/j.artmed.2020.101878] [Medline: 32425358]
70. Balendran A, Beji C, Bouvier F, et al. A scoping review of robustness concepts for machine learning in healthcare. NPJ Digit Med 2025 Jan 17;8(1):38. [doi: 10.1038/s41746-024-01420-1] [Medline: 39824951]

Abbreviations

AD: Alzheimer disease
ADRD: Alzheimer disease–related dementias
AUC: area under the curve
fMRI: functional magnetic resonance imaging
MeSH: Medical Subject Headings
MRI: magnetic resonance imaging
SVM: support vector machine

Edited by N Schütz; submitted 10.07.25; peer-reviewed by P Ramaswamy, R Singh; revised version received 08.10.25; accepted 20.11.25; published 31.12.25.

Please cite as:
Cho H, Song J, Cho H, Li L, Liang R, Miranda R, Song Q, Bian J
Applications of Machine Learning for Cognitive Health in Older Individuals With HIV: Rapid Systematic Review
JMIR Aging 2025;8:e80433
URL: https://aging.jmir.org/2025/1/e80433
doi:10.2196/80433

© Hwayoung Cho, Jiyoun Song, Hannah Cho, Lin Li, Renjie Liang, Railton Miranda, Qianqian Song, Jiang Bian. Originally published in JMIR Aging (https://aging.jmir.org), 31.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Literature on Data Governance in the Health Care of Older Persons: Scoping Review

Angely Garcia[1,2,3*], MPH, RN; Ayra Mae Balingbing[1*], PTRP; Yves Palad[3*], EdD, PTRP

[1]Institute on Aging, National Institutes of Health, University of the Philippines Manila, 623 Pedro Gil St Ermita, Manila, Philippines

[2]College of Nursing, University of the Philippines Manila, Manila, Philippines

[3]College of Allied Medical Professions, University of the Philippines Manila, Manila, Philippines

[*]all authors contributed equally

Corresponding Author:
Angely Garcia, MPH, RN
Institute on Aging, National Institutes of Health, University of the Philippines Manila, 623 Pedro Gil St Ermita, Manila, Philippines

Abstract

Background: Health data are growing rapidly, and the processing of such data is evolving. Research on data governance in older persons' health care is unexplored, providing little guidance for practice and future studies.

Objective: This scoping review aimed to synthesize available information on data governance in the context of older persons' health based on evidence from literature.

Methods: The study followed the methodological framework of Arksey and O'Malley and PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews). Online databases, namely, PubMed, Cochrane, Ovid, ACM, IEEE Xplore, and Google Scholar were searched by 2 independent reviewers (AG and AMB) for studies on older persons' health data governance published from January 2000 to April 2024. The independent reviewers performed the search, screening, data extraction, and review of full-text papers. A third reviewer (YP) made the final decision for unresolved discrepancies between the first 2 reviewers. The framework by the World Health Organization Pan American Health Organization, a high-level framework for planning and implementing data governance in public health, was used in the data extraction and analysis. Descriptive statistics were used, and a descriptive approach was used to summarize the results of the scoping review.

Results: A total of 9840 titles were identified and 57 papers were included. Of these, 35 (61.4%) focused on technology, 19 (33.3%) on processes, and 3 (5.3%) on people. Data controller, processor, researchers, data subject or patient (including family or relatives), and relevant organizations were involved in older persons' data governance. Data governance frameworks were designed and implemented by reviewing the current evidence, involving the stakeholders throughout the process, implementing specific procedures (eg, collection and aggregating health data), and monitoring and evaluating them.

Conclusions: The review underscores the importance of the involvement of relevant stakeholders and the use of various innovative tools and approaches in governing data related to the health of older persons. Meanwhile, research specifically addressing data governance for older persons' health conditions is limited. To enhance health outcomes for older persons, effective data governance is essential, alongside further research on relevant policies and practices.

(JMIR Aging 2025;8:e73625) doi:10.2196/73625

KEYWORDS

older persons; health; data governance; scoping review; medical records

Introduction

Data governance is a management discipline and an emerging data management approach in health care. It is defined within the context of information technology as specifying the framework for decision rights and accountabilities to vitalize desirable behavior in the use of data [1]. Data governance highlights the responsibilities of those with authority in the organization [2], internal and external stakeholders [3], such as data stewards [4], civic society, public bodies [5], professional bodies [6], and the individuals who contributed their data [4], in managing data through its life cycle to generate quality information that can inform decision-making. Data life cycle begins with the capture or collection to the processing, use, storage, and disposal of data [7].

Data governance is also important in health systems as it is considered by many organizations as a promising method of maintaining health data [8]. Health data are considered an important asset to improve health through their use in public health, epidemiology, and health informatics [9]. Governance of such data allows health organizations to successfully manage, protect, maintain, and use data to generate information that

XSL•FO
RenderX

improves health care quality, health outcomes, and health system performance [10].

Real-time generation and efficiency in obtaining knowledge are possible when data governance principles are applied in health care [3]. In contrast, the absence of data governance could lead to failure in decision-making and addressing the individual needs of the public sector [11]. Literature underscores the impact of data governance on a nation's health care system, and the need to collect the right data, effectively process it to generate quality information for evaluating the health system, and identify where and when it is not functioning well [12]. Available evidence calls for the need to establish, streamline, and institutionalize a strong and comprehensive data governance process [5,13-19].

Data governance is essential in digital transformation initiatives of organizations as it improves data quality and accuracy and facilitates real-time data exchange [20]. One of the principles of the global health sector's digital transformation is accelerating progress toward inclusive digital health with emphasis on the most vulnerable populations [15]. These vulnerable populations include the older persons. The potential of data governance in the field of older persons' care has been highlighted by the World Economic Forum by presenting a new approach to data governance [21]. It emphasized that the combination of caregiver skills, for example, with the older persons' data such as their specific care needs, can result in better and more precisely tailored care for older persons.

The state of research on data governance in the health care of older persons is unexplored locally and internationally. This leaves little guidance for its application in practice and for future research on the topic. Thus, this scoping review aimed to map out the available literature on data governance in the context of older persons' health. Specifically, this review sought to answer the following questions: Who is involved in the planning and implementation of data governance ("people")? How are data governance frameworks designed and implemented? What are the governance processes that lead to the improvement of older persons' health ("process")? What tools and technologies are used to effectively govern data ("technology")?

Methods

Study Design and Framework

This scoping review followed the methodological framework of Arksey and O'Malley [22] and PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) [23]. The review questions were informed by the data governance framework in public health by Pan American Health Organization (PAHO), a high-level framework for planning, implementing, and continuously improving data governance [10]. The framework highlights the involvement of 3 components: people, processes, and technology. The people component includes establishing decision-making structures (eg, both the executive decision-making and the technical bodies); defining roles and responsibilities of those involved in the management of data throughout its life cycle; communicating these roles and responsibilities, data-related decisions, policies, and processes; and ensuring transparency on adherence to standards. Processes include managing data assets, enabling processes and standard operating procedures, and establishing processes for policy management and ensuring standards. The technology component pertains to identifying and implementing tools and technology required to manage data such as hardware and software; ensuring quality, availability, and security of information systems; and ensuring performance of the tools [10].

Aside from the framework for planning and implementing data governance, the data governance functions throughout the data life cycle also guided this scoping review. Data governance should address and include defining accountabilities, prioritizing investment requirements, establishing policies, implementing processes, setting standards, managing risks, and monitoring performance related to data throughout its life cycle. The data life cycle is as follows: data collection; data aggregation; data quality assurance and monitoring; data storage; data protection; data access, use, and disclosure; and data retention and destruction [10].

Sources of Information

Six databases, namely, PubMed, Cochrane, Ovid, IEEE Xplore, Association for Computing Machinery (ACM), and Google Scholar were searched. PubMed, Ovid, and Cochrane contained a vast collection of health literature. IEEE Xplore and ACM digital libraries contained resource materials from the fields of electrical engineering, information technology, computer science, and electronics, which capture the technology-related aspect of this review. Google Scholar indexes the full text or metadata of scholarly literature and contains a wide variety of disciplines and sources.

Search Strategy

The strategy was developed in consultation with an academic librarian and the search terms used were "data governance," "health data governance," "older person," "older people," "older adult," "elderly," "senior citizen," and "aged." Boolean operators were used to filter search results (Table S1 in Multimedia Appendix 1).

This review also included evidence from nonresearch sources [24] such as reports, projects, and economy papers from Google Scholar. Moreover, experts in data governance, eHealth, and geriatrics were consulted to explore additional literature sources. The experts were identified through the existing networks of the investigators and organizations relevant to digital health. Searching the reference list of the identified data governance literature was also conducted.

All the literature searched through the online databases, digital libraries, consultation with experts, and reference listing was uploaded in a shared Google Drive folder accessible only to the reviewers using their official university email addresses. The titles were encoded using Google Sheets to identify and remove duplicates.

Selection Process

Considering that data governance was introduced in 2000 and literature on the topic started to be published around the same

XSL•FO
RenderX

time [1,2,25], the search was filtered by publication dates between January 1, 2000, and April 22, 2024. The papers that meet the following criteria were included: (1) topic related to data governance including people, processes, or technology, (2) topic relevant to older persons' health care, (3) peer-reviewed publications, reports, policies, programs, and policy briefs, and (4) in the English language. Resources that were not related to data governance and older persons' health, conference abstracts, and published in languages other than English were excluded.

Following the eligibility criteria, 2 reviewers (AG and AMB) independently searched databases, screened papers for eligibility, reviewed the full texts, and extracted data from the included studies. If the reviewers are unable to screen papers based on the title alone, abstract screening was done to check whether the papers are related to health, data governance, and older persons. Initial data extraction from the abstracts was done to document the topic they covered using a data extraction form.

The reviewers independently documented the search yields and listed the titles obtained per online database using Google Sheets. The search yields of each reviewer per online database were compared as a form of initial validation. To ensure consistency among the reviewers, the procedures from searching to data extraction were pilot tested using 10 randomly selected samples. Reviewers proceeded with the next phases only upon reaching 75% agreement [24].

For each phase, the reviewers documented the reasons for exclusion and settled discrepancies by discussing them throughout the selection process. The reviewers met online and conferred about the discrepancies, which led to an amendment of the eligibility criteria. Based on the initial assessment of the abstracts and available full texts, there were various papers with mixed populations as their study samples (eg, 18 years and above). Considering the focus of this review, the reviewers decided to only include papers with older persons explicitly stated in their titles or abstracts or representing more than 50% of the study sample. Meetings between the reviewers were conducted to discuss which of the papers would be included in the review.

Data Extraction

Once the list of included studies was finalized, the 2 reviewers (AG and AMB) independently extracted the relevant information using an extraction tool in Google Sheets. Information such as authors, citation, publication year, type of publication, research design, setting of care, population, data governance function and component, and outcomes of the study and the technological intervention, among others, were extracted for further analysis. The reviewers conducted regular meetings to discuss and reconcile differences in the extracted data. Any disagreements that remained unresolved were discussed with a third reviewer (YP) to arrive at a consensus. The third reviewer made the final decision for the unresolved disagreements between the first 2 reviewers.

Data Management and Analysis

The search results and study selection process are documented in this scoping review report and presented using a PRISMA-ScR flow diagram. All data were documented in a Google Sheets document and exported into an Excel file after the entire process ended. For the unstructured texts in the data extraction form, a qualitative content analysis was conducted to describe how data governance is planned and implemented in the care of older persons. The data were coded according to the framework components involved, namely, person, processes, and technology, and data governance functions. A descriptive approach was used to summarize the result of the scoping review. Numerical or categorical data were presented using counts and proportions. Tables and a figure were used to present the extracted data for each extraction category.

Ethical Considerations

The study was monitored and classified as exempt from ethics review (UPMREB 2023 - 0864-EX) by the University of the Philippines Manila Research Ethics Board.

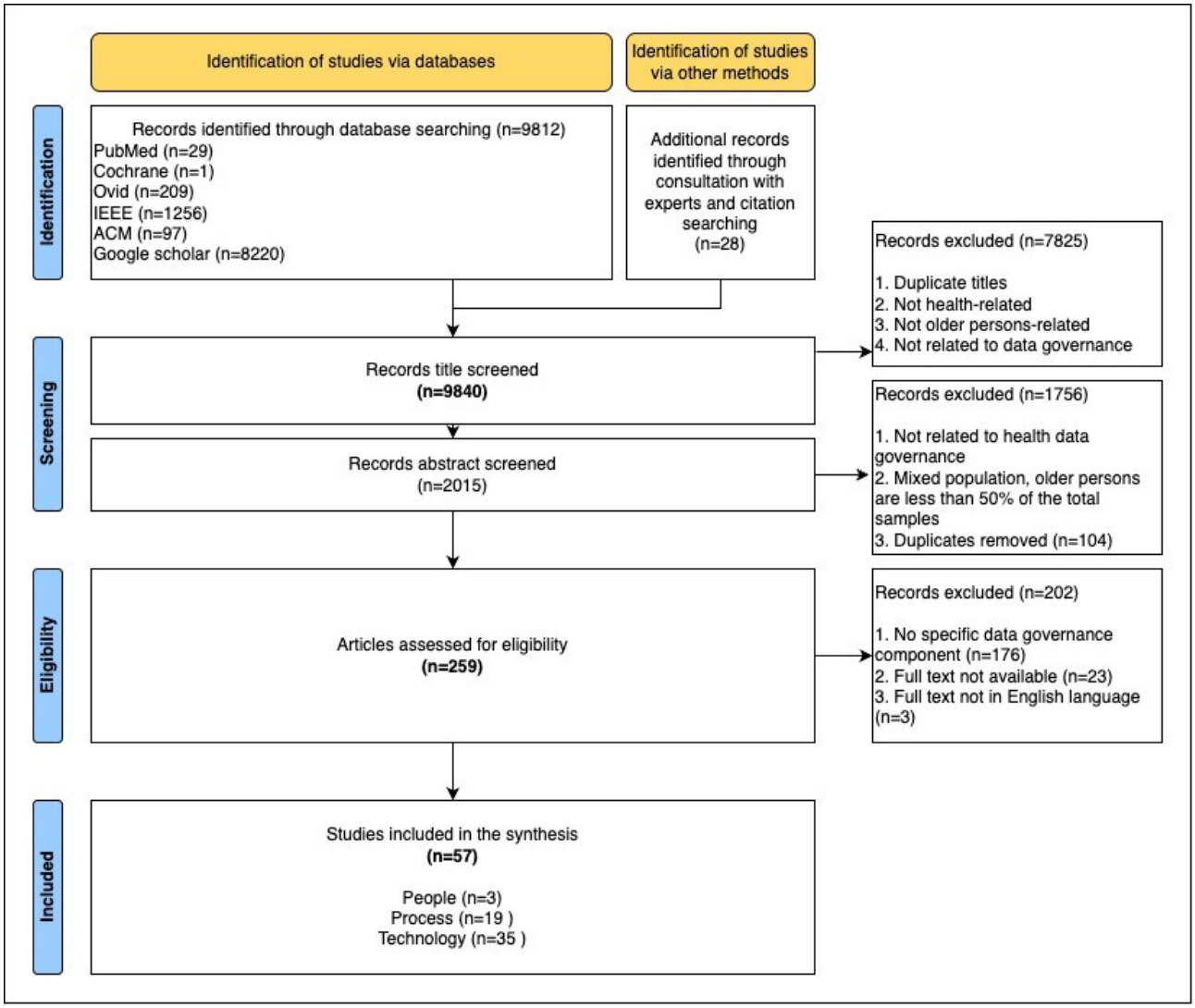
Results

Overview

A total of 9840 titles were identified from 6 databases, consultation with 3 experts, and reference listing (Figure 1). Duplicates and all papers not related to older persons, health, and data governance, and those with mixed populations were removed during the screening phase, leaving a total of 259 papers assessed for eligibility. Of these papers, 202 were removed due to the unavailability of full texts, non-English language, and lack of specific data governance components. A total of 57 (22%) papers met the criteria for inclusion in the synthesis.

XSL•FO
RenderX

Figure 1. PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) diagram.



Description of Included Studies

The included studies were published between 2003 and 2024; 56.1% (32/57) of these were published in the year 2020 onwards. The majority (32/57, 56.1%) of the included studies were from the IEEE digital library. Most (47/57, 82.5%) were journal papers, followed by conference proceedings (7/57, 12.3%). These conference proceedings were peer-reviewed full papers from IEEE and ACM (Table 1).

In terms of study design, 36.8% (21/57) were observational, followed by 29.8% (17/57) classified as experimental studies. These experimental studies are subdivided into 2 groups, namely, technology development and randomized controlled trials. About 30% (16/57) of the studies did not specify their design.

The studies were conducted mostly in home settings (11/57, 15.8%), followed by those conducted in multiple settings of care such as home and hospital, community and hospital, and hospital and long-term care facilities. More than 30% of the included studies did not specify their study setting. In terms of geographic coverage, most were conducted in Europe, followed by Asia and those with multicountry sites. There were no included studies that covered the following continents: Africa, Middle East, and South America (Table 1).

Sixty-seven percent of the included studies provided information about the study population, particularly, the age group, mean age, or age range they covered. Although 33.3% (19/57) did not indicate the specific age of their study population, the technology or the data governance component being relevant to older persons was explicitly stated in the title, abstract, and full text. Due to the variation of available information on age and unavailability of other information, the overall mean age cannot be computed.

About 70% of the included studies specified the sample size. A study had 1 older person sample, which was about a novel cloud-based framework for the older adult health care services [26]. The largest sample size among the included studies was 591,726 electronic health records of primary care patients aged 65 years and older [27].

XSL•FO
RenderX

Table . Profile of included studies.

Description	Included studies (n=57), n (%)
Data sources	
IEEE	32 (56.1)
Google Scholar	15 (26.3)
ACM[a]	4 (7.0)
PubMed	4 (7.0)
Ovid	2 (3.5)
Expert consultation	0 (0)
Publication year	
2000 - 2009	4 (7.0)
2010 - 2019	21 (36.8)
2020 and onward	32 (56.1)
Publication type	
Journal paper	47 (82.5)
Conference proceeding	7 (12.3)
Thesis	1 (1.8)
Review	1 (1.8)
Economy paper	1 (1.8)
Study design	
Observational	21 (36.8)
Experimental	17 (29.8)
Technology development	15 (26.3)
Randomized controlled trial	2 (3.5)
Review (eg, systematic review and scoping review)	3 (5.3)
Not specified	16 (28.1)
Study settings	
Home	11 (19.3)
Multiple settings	9 (15.8)
Hospital	5 (8.8)
Long-term care setting	5 (8.8)
Community	4 (7.0)
Others: research laboratory or special facility for experiment, university, office	4 (7.0)
Not specified	19 (33.3)
Geographic coverage	
Europe	12 (21.1)
Asia	6 (10.5)
Australia and Oceania	3 (5.3)
North America	2 (3.5)
Multicountry or continent sites	6 (10.5)
Not specified	28 (49.1)
Age (years) of older persons in study sample	

XSL•FO
RenderX

Description	Included studies (n=57), n (%)
Specified (age covered/mean/range)	38 (66.7)
Minimum age	18 (N/A[b])
Maximum age	104 (N/A)
Not specified	19 (33.3)
Sample size	
Specified	39 (68.4)
Minimum sample size	1 (N/A)
Maximum sample size (datasets)	591,726 (N/A)
Not specified	18 (31.6)

[a]ACM: Association for Computing Machinery.

[b]N/A: not applicable.

Data Governance Components and Functions

A majority (35/57, 61.4%) of the included studies were largely related to technology, tools and technology, in particular (Table 2). This is followed by processes (19/57, 33.3%), which are mostly centered on enabling processes and standard operating procedures. Only 3 papers were focused on people, particularly, roles and responsibilities and decision-making structure.

Interestingly, there were no included studies on communication and transparency under the people component, standards and policy management in the process component, availability and security, and performance alone in the technology component (Table 2).

Table 2 also shows that the data governance function of the included studies mostly focused on implementing processes, while only a few covered monitoring performances [27-30], defining accountability [31-33], setting standards [34,35], and managing risks [36].

XSL•FO
RenderX

Table . Data governance components and functions of included studies.

Data governance	Included studies (n=57), n (%)
Data governance component	
People	3 (5.3)
None	54 (94.7)
Roles and responsibilities	2 (3.5)
Decision-making structure	1 (1.8)
Communication and transparency	0 (0)
Process	19 (33.3)
None	37 (64.9)
Enabling processes and SOPs[a]	11 (19.3)
Data asset management	8 (15.8)
Standards and policy management	0 (0)
Technology	35 (61.4)
None	22 (38.6)
Tools and technology	33 (57.9)
Tools and technology and performance	2 (3.5)
Availability and security	0 (0)
Performance	0 (0)
Data governance function	
Implement processes	47 (82.5)
Monitor performance	4 (7.0)
Define accountability	3 (5.3)
Set standards	2 (3.5)
Manage risks	1 (1.8)
Prioritize investment	0 (0)
Establish policy	0 (0)

[a]SOP: standard operating procedure.

Designing and Implementing Data Governance

The identified actors involved in designing and implementing data governance related to the health of older persons include data controllers, processors, and subjects. Specific data controllers identified in this review include a hospital [32], researchers [33], and a steering committee [37]. The data subjects were patients [32], persons with dementia [33], nursing home residents [37], friends or relatives [32], legally authorized representatives [33], and hospital staff [32]. The data processors were a platform service provider [32], research team members [33], and privacy, scientific expert, and data access committees [37].

All included studies on the people component of data governance highlighted the crucial roles of relevant stakeholders (eg, controllers, processors, and subjects) in planning, designing, and implementing data governance [27,32,33]. Moreover, the composition, roles, and accountabilities of these stakeholders vary depending on the settings (eg, hospital, research, and nursing home).

An example specifying the roles and responsibilities of people involved in data governance was highlighted by Bernsmed [32] in the context of medical data collected from sensors that are exchanged between the older persons, their families and friends, and health personnel. The roles of data controller, processor, and subject were defined, and the accountability obligations were outlined. A data controller, which can be a person, public authority, agency, or other actors, determines the purposes and means of personal data processing. The data processor processes the personal data on behalf of the controller. Meanwhile, a data subject is a person who can be identified directly or indirectly through a reference identification number. In the study, the patients and relatives or friends, as well as the hospital staff, were the data subjects, while the hospital is the controller of its patients' and personnel's data. The data controller (hospital) is accountable to patients, relatives or friends, and hospital staff (data subjects) in informing and obtaining consent for the collection, processing, and management of their data [32].

A total of 10 papers provided information on how frameworks related to data governance of older persons' health data are

XSL•FO
RenderX

designed and implemented [26,32,33,35,37-42]. Specific examples of the use of data governance frameworks were related to addressing a particular issue such as fall detection [41], dementia care mapping [42], depression risk prediction [35], cardiac image processing [40], and digital twin health care for real-time supervision and crisis warning [43].

Conceptual content analysis of the methods using the PAHO framework emphasized the involvement of relevant stakeholders, implementation of specific processes, and conduct of literature review, demonstrating the feasibility or testing of the developed framework and defining the roles and accountabilities of the actors or stakeholders (Table S2 in Multimedia Appendix 1).

Data Governance Processes, Tools, and Outcomes

A total of 19 papers provided information on the specific processes or interventions relevant to older persons' health [26,30,34,35,38,39,41-53]. Majority focused on implementing processes pertaining to data collection, aggregation, and access, use, and disclosure. One study covered 6 data life cycles from data collection up to disclosure [44]. There were no papers that covered the process of defining accountability, managing risks, prioritizing investment, and establishing policy (Table 2).

Several tools and technologies were identified and used for data governance throughout the data life cycle. A total of 35 papers provided information on the specific technologies relevant to older persons' health [27-29,36,40,54-82]. Most of the identified technologies focused on implementing processes while only a few were related to use of technology in monitoring performance, defining accountability, and managing risks. Furthermore, these technologies are mostly involved in data collection.

In implementing data governance processes, various tools and technologies are used in collecting data for the assessment and monitoring of health needs; assessment of fall risk and detection of fall; identification of mental health–related issues such as mild cognitive impairment, dementia, stress, loneliness, and depression; and improvement of lifestyle and health (Table S2 in Multimedia Appendix 1).

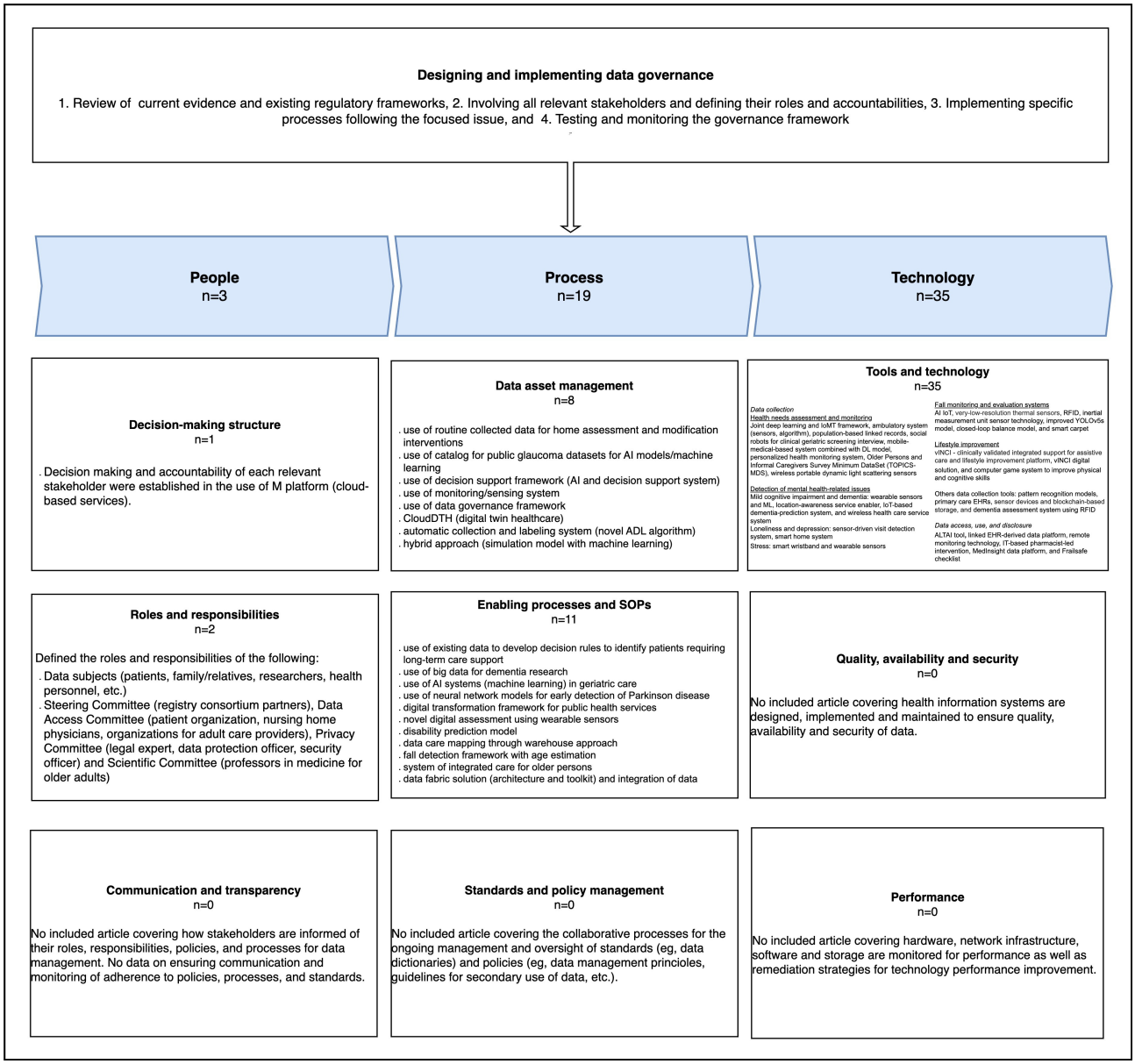
In terms of specific outcomes of the technology component of data governance, a total of 34 papers provided information on these. Most of the outcomes were process-related, such as efficient data access; quality control; communication and care support; recognition, detection, and prediction of various conditions; monitoring; and others. Only 2 included papers provided information on health outcomes among the older person samples, namely, improvement of quality of life [62] and cognitive skills and function [68] (Table S2 in Multimedia Appendix 1).

Older Persons' Health Care Data Governance

Figure 2 summarizes the results of this scoping review following the data governance framework in public health by PAHO [10].

XSL•FO
RenderX

Figure 2. Older person's health care data governance based on the available literature. ADL: activities of daily living; AI: artificial intelligence; EHR: electronic health record; IoT: Internet of Things; DL: deep learning; ML: machine learning; RFID: radio frequency identification; SOP: standard operating procedure.



Discussion

Principal Results

This scoping review aimed to synthesize available information on how data governance is planned and implemented in the context of older persons' health based on evidence from literature. Specifically, it aimed to answer the following questions: Who are involved in the planning and implementation of data governance (people)? How are data governance frameworks designed and implemented? What are the governance processes that lead to the improvement of older persons' health (process)? What tools and technologies are used to effectively govern data (technology)?

Findings show that available studies related to the health of older persons covered mostly data governance components related to technology and process. Moreover, data governance is designed and implemented by initially examining the current evidence and existing regulatory frameworks, involving all relevant stakeholders and defining their roles and accountabilities, implementing specific processes following the focused issue, and testing and monitoring the governance framework. Various data governance processes, tools, and technologies are used and contribute to the improvement of older persons' health. The use of these mostly leads to process outcomes.

This review emphasizes the involvement of stakeholders, particularly in technology development and testing process [40,41], sharing perspectives on digital transformation of public health services [39], and mapping of service users [42]. Moreover, the accountability of data controllers, processors, and subjects was outlined [32,33,37]. This underscores the data governance responsibilities of those with the authority in the organization [2] (data controller), the internal and external stakeholders (data processors and subjects) [3-6], and, ultimately, the data subjects who contributed their data [4] in

XSL•FO
RenderX

managing data through their life cycle to inform decision-making.

Limitations

This review included literature only on data governance relevant to older persons published from the past 2 decades and those peer-reviewed publications, reports, policies, programs, or policy briefs written in English language. Identified studies eligible for review but without full texts were excluded from the synthesis. Despite the efforts of the reviewers and consultant librarian in accessing the papers, 23 full texts were not retrieved. These papers were mostly from ACM and Google Scholar, published 2021 and beyond, included conference proceedings, and focused on the technology component of data governance. These exclusions might have affected the comprehensiveness of the reported findings on technology.

Many included studies did not specify the study design, setting, geographic coverage, age of the study population, and sample size. It would be helpful if this information were available to provide a complete description of the state of data governance literature for older persons. Moreover, the assessment or appraisal of the included studies for methodological rigor and quality was not performed as this review aimed to describe information only on data governance in the context of older persons' health.

Various processes and technologies aimed to contribute to the improvement of older persons' health based on their objectives and purpose were identified. However, since most of the included studies are related to technology, at their preclinical stage of technological development, and focused only on data collection, limited information on the actual health outcomes of data governance specific to older persons was collected.

Comparison With Prior Work

Studies related to health information systems included in this review [27,60,79] also support data governance as an important player in the health system, a promising method of maintaining health data [8], and recognize its potential in the field of older persons' care [83,84]. Various processes, tools, and innovative approaches in governing health data of older persons were explored in this review. Available evidence suggests that the use of these tools and technologies supports data governance, leading to positive health outcomes and better processes [27,31,61-63,68,75-77]. Although positive health outcomes were noted, further interventional studies with utilization of a data governance framework are essential to determine its direct effects on actual health outcomes among the older population.

In terms of processes, findings in this review are consistent with Cave [85] where strategies for implementing data governance were explored through a qualitative case study. These strategies include structured oversight with committees and boards, obtaining stakeholder buy-in, and benchmarking and standardization. Benchmarking and standardization through review of current evidence and existing regulatory frameworks were identified to be important steps in planning and implementing data governance [32,33,35,38,39,42]. These preliminary steps are crucial in ensuring a well-designed and effective governance framework.

Strategies for effective and strategic communications and compliance with regulations were not adequately covered in this review due to the lack of studies on this topic. Communication and transparency, as well as compliance with standards and policies, should be considered in implementing data governance. This is to ensure the protection of people, promotion of health value, and prioritization of equity, which are the key health data governance principles [86]. Moreover, although there were no papers that covered the processes of defining accountability and managing risks, these were discussed in the literature in terms of the role of persons involved in data governance and as integral components of designing and implementing data governance.

Various geriatric syndromes, such as falls and fall risk [54,55,65,69,71,73,74], mild cognitive impairment and dementia [33,42,53,64,66,75,82], frailty [29,80], loneliness, social isolation, and depression [56,57], were the focus of the studies included in this review. However, there were no included studies related to other geriatric syndromes such as functional decline, incontinence, delirium, pressure ulcers, polypharmacy, malnutrition, sleep problems, and others. These geriatric syndromes are common health conditions among older persons [87-90], often having multifactorial causes, and may have a major impact on their quality of life and disability [90]. Data governance covering these conditions can also generate information on strategies to address the needs of older persons and provide quality care to improve their health and quality of life.

Only 1 paper explicitly proposed a data governance framework specific for older persons, which was published a decade ago [38]. The proposed framework was drawn from the Data Management Body of Knowledge of the DAMA International, Inc [2,91] and the work of Cleven and Wortmann [92]. Dahlberg's [38] motive for proposing this specific governance framework was the fragmentation of older persons' data and the necessity to consolidate these data to make it more useful. The paper defined the types and sources of data that exist about older persons, and how these can be integrated into a comprehensive framework. It suggested defining the data categories, attributes, and sources of data first to improve data governance of older persons. There is no evidence that the framework has been validated. Hence, testing this framework is essential to determine its value in data governance for older persons.

It is notable that the included papers that provided information on designing and implementing data governance mostly came from high-income countries such as Australia, United States, Canada, United Kingdom, Japan, Singapore, Netherlands, and other countries in Europe. These countries are actively implementing digital health care solutions in the forms of telemedicine, electronic records, mobile health, and other digital tools for diagnosis and management, hence, the availability of data governance frameworks. On the other hand, no studies were included in the review from Africa, Middle East, and South America. This can be attributed to the challenges experienced in digital health, such as limited infrastructure, affordability issues, and data privacy concerns, especially in low-income and middle-income countries [93-95]. While others have established

XSL•FO
RenderX

and are implementing data governance, the rest are yet to address the challenges of digital health transformation.

The need for data governance in general [14] and in the care of older persons has been highlighted by the reviewed papers specifically on dementia research [42,53], health care applications for geriatric clinical care [49], and nursing homes for quality improvement [37]. New paradigms on the use of big data in dementia research and clinical care [42], as well as international dementia care mapping, require data governance [53]. Likewise, the standardization of machine learning approaches tailored to health care applications is required to evaluate whether these applications improve clinical care for older persons [49]. In nursing homes, literature suggests the need for enhancing transparency, specifically in presenting understandable information to the residents and their representatives on which data will be used, how they will be used, and for what purposes [37].

Findings agree with other papers not included in this review such as the call for institutionalizing a strong and comprehensive data governance process [5,10,12-19,96-98], a national health data governance framework to ensure availability and use of personal health data for public interest [99], and data governance for learning health systems to reduce concerns about privacy and trust in the system [100,101].The included studies were analyzed according to the data life cycle, data governance function, and components they covered to identify research gaps. The people data governance component covered only data access, use, and disclosure. Information on how people are involved, their roles and responsibilities, and communication and transparency during data collection, aggregation process, quality assurance, storage, data protection, retention, and destruction, is lacking. In terms of the process component, limited studies were related to data quality, storage, and protection, and none for defining accountability and managing risk. Meanwhile, no tools and technology were identified from the studies that covered data aggregation, data quality, data protection, and setting standards.

Studies on data governance in health care of older persons are largely focused on various forms of technology followed by process. The major research gap across all data governance components was the lack of included studies on data retention and destruction, prioritizing investments, and establishing policy (Table S3 in Multimedia Appendix 1). Findings and gaps that were identified from this review inform future research directions and practices of older persons' health data governance.

Conclusions

This review highlighted the importance of benchmarking, involvement of relevant stakeholders, and the use of various innovative tools and approaches in governing data related to the health of older persons. However, studies that are explicitly centered on data governance of older persons' health data, their common health conditions, or other geriatric syndromes, are limited. Likewise, identified research gaps include the lack of studies on data retention and destruction, establishment of policy, prioritizing investment, and communication and transparency.

Available studies on the health of older persons covered mostly technology and process data governance components. Only a few focused on people, which were identified to have crucial roles in designing governance frameworks, planning, and implementing data governance. To optimize the use of data in improving the health and quality of life of older persons, a well-designed process that considers the essential data governance components, functions, and principles should be developed and implemented.

Recommendations

Effective governance of older persons' health data requires a multisectoral approach. Benchmarking, utilization of innovative tools and approaches, and collaboration between relevant stakeholders, including policy makers, program planners, health facility administrators, health care providers, information and communication technology and digital health experts, data privacy and legal experts, and, ultimately, the older persons and their families, are essential in designing and implementing data governance.

In establishing health data governance policies or incorporating them into the national standards, legislations, and organizational policies, the key governance principles, components, functions, and data life cycle should be considered. In practice, health settings and personnel should ensure the involvement of the patients (including their families and relatives) in collecting and processing their data. The practical application of data governance frameworks and technologies in the care of older persons is essential to support favorable health outcomes. Further studies on the governance of data on common health conditions of older persons to inform clinical practice, health data retention schedules and destruction methods, and data governance policies and investments are recommended.

Acknowledgments

This scoping review was funded by the National Institutes of Health (NIH), University of the Philippines Manila through the NIH Commissioned Grants 2023 (project code: NIH 2023-027). The project team would like to thank and acknowledge the following: the NIH Executive Committee for funding this scoping review; Dr Shelley Ann de la Vega, director of the Institute on Aging, for her valuable support to the project especially during the funding application; Ms Vivian Biala for providing research grant administrative support; CAMP Research Committee for carefully reviewing and providing technical approval of the proposal; UP Manila Research Ethics Board for exempting the study from ethical review; experts consultants (Dr Alvin Marcelo, Dr Evelyn Hovenga, and Dr John Richard Hizon) for providing valuable insights based on their expertise; Mr Cristopher Naguit for helping the team process the salary and disbursement documents; and, finally, to the UP Manila channels for the processing of administrative,

XSL•FO
RenderX

legal, financial, and technical documents needed for this project. This project would not have been completed without the support of the above personalities and offices. All authors declared that they had insufficient or no funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided APF support for the publication of this paper.

Data Availability

All study data are present in this paper and the supplementary materials.

Authors' Contributions

AG performed protocol development, data collection, analysis, manuscript writing, review, and approval of the paper; AMB performed data collection, analysis, manuscript writing, review, and approval of the paper; and YP performed protocol development, data analysis, manuscript writing, review, and approval of the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search terms used by database, summary of extracted entries from the included studies, and research gaps on data governance in the health care of older persons.
[DOCX File, 42 KB - aging_v8i1e73625_app1.docx]

Checklist 1
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) checklist.
[PDF File, 660 KB - aging_v8i1e73625_app2.pdf]

References

1. Weill P, Ross JW. IT Governance: How Top Performers Manage IT Decision Rights for Superior Results: Harvard Business School Press; 2004.
2. Data Administration Management Association. In: Mosley M, Brackett M, Earley S, Henderson D, editors. The DAMA Guide to the Data Management Body of Knowledge: DAMA-DMBOK Guide, 1st edition: Technics Publications, LLC; 2010.
3. Hovenga EJS, Grain H. Health data and data governance. Stud Health Technol Inform 2013;193(193):67-92. [Medline: 24018511]
4. Rosenbaum S. Data governance and stewardship: designing data stewardship entities and advancing data access. Health Serv Res 2010 Oct;45(5 Pt 2):1442-1455. [doi: 10.1111/j.1475-6773.2010.01140.x] [Medline: 21054365]
5. Micheli M, Ponti M, Craglia M, Berti Suman A. Emerging models of data governance in the age of datafication. Big Data Soc 2020 Jul;7(2). [doi: 10.1177/2053951720948087]
6. Hovenga E, Grain H. Quality data, design, implementation, and governance. In: Roadmap to Successful Digital Health Ecosystems: A Global Perspective: Elsevier; 2022:209-237. [doi: 10.1016/B978-0-12-823413-6.00013-6]
7. Oachs PK, Watters A, editors. Health Information Management: Concepts, Principles, and Practice, 6th edition: American Health Information Management Association; 2020.
8. Ngesimani NL, Ruhode E, Harpur PA. Data governance in healthcare information systems: a systematic literature review. S Afr J Inf Manag 2022 Aug 18;24(1). [doi: 10.4102/sajim.v24i1.1475]
9. Dickens A. From information to valuable asset: the commercialization of health data as a human rights issue. Health Hum Rights 2020 Dec;22(2):67-69. [Medline: 33390695]
10. Data Governance in Public Health: Pan American Health Organization; 2022. URL: https://iris.paho.org/handle/10665.2/56576 [accessed 2025-05-21]
11. Benfeldt Nielsen O. A comprehensive review of data governance literature. Selected Papers of the IRIS 2017 Dec 31(8):3 [FREE Full text]
12. Hovenga EJS. Impact of data governance on a nation's healthcare system building blocks. Stud Health Technol Inform 2013;193(193):24-66. [Medline: 24018510]
13. Healthcare data governance. American Health Information Management Association (AHIMA). 2022. URL: https://www.ahima.org/media/pmcb0fr5/healthcare-data-governance-practice-brief-final.pdf [accessed 2025-05-31]
14. Recommendation of the council on health data governance. OECD. 2025. URL: https://legalinstruments.oecd.org/public/doc/348/348.en.pdf [accessed 2025-06-12]

XSL•FO
RenderX

15. From the evolution of information systems for health to the digital transformation of the health sector. : Pan American Health Organization; 2021 URL: https://iris.paho.org/bitstream/handle/10665.2/53364/PAHOEIHIS210006_eng.pdf [accessed 2025-05-21]
16. Paparova D, Aanestad M, Vassilakopoulou P, Bahus MK. Data governance spaces: the case of a national digital service for personal health data. Inf Organ 2023 Mar;33(1):100451. [doi: 10.1016/j.infoandorg.2023.100451]
17. Zayas-Cabán T, Chaney KJ, Rucker DW. National health information technology priorities for research: a policy and development agenda. J Am Med Inform Assoc 2020 Apr 1;27(4):652-657. [doi: 10.1093/jamia/ocaa008]
18. Meeting on strengthening health information systems for sustainable development goals and UHC monitoring in the Western Pacific region. : World Health Organization Regional Office for the Western Pacific (WHO WPRO); 2019 URL: https://iris.who.int/handle/10665/325941?locale-attribute=fr&show=full [accessed 2025-05-21]
19. Williamson K, Nimegeer A, Lean M. Navigating data governance approvals to use routine health and social care data to evidence the hidden population with severe obesity: a case study from a clinical academic's perspective. J Res Nurs 2022 Nov;27(7):623-636. [doi: 10.1177/17449871221122040] [Medline: 36405806]
20. Good data: data sharing, public trust and willingness. World Economic Forum. 2021. URL: https://www.weforum.org/publications/good-data-sharing-data-and-fostering-public-trust-and-willingness [accessed 2025-05-21]
21. Why we need to leverage data for healthy ageing. World Economic Forum. 2021. URL: https://www.weforum.org/agenda/2021/10/we-need-to-leverage-data-for-healthy-ageing/ [accessed 2025-05-21]
22. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
23. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
24. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
25. Nooteboom B. Learning by interaction: absorptive capacity, cognitive distance and governance. J Manage Governance 2000;4(1/2):69-92. [doi: 10.1023/A:1009941416749]
26. Liu Y, Zhang L, Yang Y, et al. A novel cloud-based framework for the elderly healthcare services using Digital Twin. IEEE Access 2019;7(7):49088-49101. [doi: 10.1109/ACCESS.2019.2909828]
27. Joseph RM, Knaggs RD, Coupland CAC, et al. Frequency and impact of medication reviews for people aged 65 years or above in UK primary care: an observational study using electronic health records. BMC Geriatr 2023 Jul 14;23(1):435. [doi: 10.1186/s12877-023-04143-2] [Medline: 37442984]
28. Pradhan A, Kuruganti U, Chester V. Biomechanical parameters and clinical assessment scores for identifying elderly fallers based on balance and dynamic tasks. IEEE Access 2020;8(8):193532-193543. [doi: 10.1109/ACCESS.2020.3033194]
29. Papoutsi C, Poots A, Clements J, Wyrko Z, Offord N, Reed JE. Improving patient safety for older people in acute admissions: implementation of the Frailsafe checklist in 12 hospitals across the UK. Age Ageing 2018 Mar 1;47(2):311-317. [doi: 10.1093/ageing/afx194] [Medline: 29315375]
30. Daniels H, Hollinghurst J, Fry R, et al. The value of routinely collected data in evaluating home assessment and modification interventions to prevent falls in older people: systematic literature review. JMIR Aging 2021 Apr 23;4(2):e24728. [doi: 10.2196/24728] [Medline: 33890864]
31. Fang L, Wu Y, Wu C, Yu Y. A nonintrusive elderly home monitoring system. IEEE Internet Things J 2021;8(4):2603-2614. [doi: 10.1109/JIOT.2020.3019270] [Medline: 37974935]
32. Bernsmed K. Accountable health care service provisioning in the cloud. 2014 Presented at: 2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing (UCC); Dec 8-11, 2014; London, England, United Kingdom p. 902-907. [doi: 10.1109/UCC.2014.147]
33. Thorogood A, Mäki-Petäjä-Leinonen A, Brodaty H, Global Alliance for Genomics and Health, Ageing and Dementia Task Team, et al. Consent recommendations for research and international data sharing involving persons with dementia. Alzheimers Dement 2018 Oct;14(10):1334-1343. [doi: 10.1016/j.jalz.2018.05.011] [Medline: 30293575]
34. Kiefer R, Abid M, Steen J, Ardali MR, Amjadian E. A catalog of public glaucoma datasets for machine learning applications: a detailed description and analysis of public glaucoma datasets available to machine learning engineers tackling glaucoma-related problems using retinal fundus images and OCT images. : ACM; 2023 Presented at: 2023 the 7th International Conference on Information System and Data Mining (ICISDM); May 10-12, 2023; Atlanta, United States p. 24-31. [doi: 10.1145/3603765.3603779]
35. Billis AS, Papageorgiou EI, Frantzidis CA, Tsatali MS, Tsolaki AC, Bamidis PD. A decision-support framework for promoting independent living and ageing well. IEEE J Biomed Health Inform 2015 Jan;19(1):199-209. [doi: 10.1109/JBHI.2014.2336757] [Medline: 25073180]
36. Chung-Chih L, Ping-Yeh L, Po-Kuan L, Guan-Yu H, Wei-Lun L, Ren-Guey L. A healthcare integration system for disease assessment and safety monitoring of dementia patients. IEEE Trans Inform Technol Biomed 2008;12(5):579-586. [doi: 10.1109/TITB.2008.917914]

XSL•FO
RenderX

37. Wieland-Jorna Y, Verheij RA, Francke AL, et al. Setting up a governance framework for secondary use of routine health data in nursing homes: development study using qualitative interviews. J Med Internet Res 2023 Jan 25;25(25):e38929. [doi: 10.2196/38929] [Medline: 36696162]
38. Dahlberg T. A framework for and issues in the management and governance of data on elderly citizens (ongoing res.). Presented at: 8th International Conference on Theory and Practice of Electronic Governance; Oct 27-30, 2014; Guimaraes, Portugal. [doi: 10.1145/2691195.2691295]
39. Zhou S, Ni Z, Ogihara A, Wang X. Behavioral patterns of supply and demand sides of health services for the elderly in sustainable digital transformation: a mixed methods study. Int J Environ Res Public Health 2022 Jul 5;19(13):8221. [doi: 10.3390/ijerph19138221] [Medline: 35805878]
40. Zhang T, Sodhro AH, Luo Z, et al. A joint deep learning and internet of medical things driven framework for elderly patients. IEEE Access 2020;8(8):75822-75832. [doi: 10.1109/ACCESS.2020.2989143]
41. Lin D, Yao C, Min W, et al. A novel fall detection framework with age estimation based on cloud-fog computing architecture. IEEE Sensors J 2024 Feb;24(3):3058-3071. [doi: 10.1109/JSEN.2023.3334555]
42. Khalid S. Towards data governance for international dementia care mapping (DCM). a study proposing DCM data management through a data warehousing approach [Master's thesis]. : The University of Bradford; 2011 Nov 18 URL: https://bradscholars.brad.ac.uk/handle/10454/5226 [accessed 2025-05-21]
43. Lin CH, Wang FC, Kuo TY, Huang PW, Chen SF, Fu LC. Early detection of Parkinson's disease by neural network models. IEEE Access 2022;10(10):19033-19044. [doi: 10.1109/ACCESS.2022.3150774]
44. Buleje I, Siu VS, Hsieh KY, et al. A versatile data fabric for advanced IoT-based remote health monitoring. Presented at: 2023 IEEE International Conference on Digital Health (ICDH); Jul 2-8, 2023; Chicago, IL, United States. [doi: 10.1109/ICDH60066.2023.00021]
45. Béland F, Bergman H, Lebel P, et al. A system of integrated care for older persons with disabilities in Canada: results from a randomized controlled trial. J Gerontol 2006 Apr 1;61(4):367-373. [doi: 10.1093/gerona/61.4.367]
46. Greene BR, McManus K, Redmond SJ, Caulfield B, Quinn CC. Digital assessment of falls risk, frailty, and mobility impairment using wearable sensors. NPJ Digit Med 2019;2(2):125. [doi: 10.1038/s41746-019-0204-z] [Medline: 31840096]
47. Shulman MA, Wallace S, Gilbert A, Reilly JR, Kasza J, Myles PS. Predicting death or disability after surgery in the older adult. Anesthesiology 2023;139(4):420-431. [doi: 10.1097/ALN.0000000000004683]
48. Kim JY, Liu N, Tan HX, Chu CH. Unobtrusive monitoring to detect depression for elderly with chronic illnesses. IEEE Sensors J 2017 Sep 1;17(17):5694-5704. [doi: 10.1109/JSEN.2017.2729594]
49. Choudhury A, Renjilian E, Asan O. Use of machine learning in geriatric clinical care for chronic diseases: a systematic literature review. JAMIA Open 2020 Oct;3(3):459-471. [doi: 10.1093/jamiaopen/ooaa034] [Medline: 33215079]
50. Elbattah M, Molloy O. Coupling simulation with machine learning: a hybrid approach for elderly discharge planning. : ACM; 2016 Presented at: Proceedings of the 2016 ACM SIGSIM Conference on Principles of Advanced Discrete Simulation; May 15-18, 2016; Banff, Alberta, Canada p. 47-56. [doi: 10.1145/2901378.2901381]
51. Hu X, Chu D, Li Z, et al. Coarse-to-fine activity annotation and recognition algorithm for solitary older adults. IEEE Access 2020;8(8):4051-4064. [doi: 10.1109/ACCESS.2019.2962843]
52. Subbe CP, Goulden N, Mawdsley K, Smith R. Anticipating care needs of patients after discharge from hospital: frail and elderly patients without physiological abnormality on day of admission are more likely to require social services input. Eur J Intern Med 2017 Nov;45(45):74-77. [doi: 10.1016/j.ejim.2017.09.029] [Medline: 28974330]
53. Unleashing the power of big data for Alzheimer's disease and dementia research: main points of the OECD expert consultation on unlocking global collaboration to accelerate innovation for Alzheimer's disease and dementia. : OECD Publishing; 2014 Mar URL: http://dx.doi.org/10.1787/5jz73kvmvbwb-en [accessed 2025-06-06]
54. Lin KC, Wai RJ. A feasible fall evaluation system via artificial intelligence gesture detection of gait and balance for sub-healthy community- dwelling older adults in Taiwan. IEEE Access 2021;9(9):146404-146413. [doi: 10.1109/ACCESS.2021.3123297]
55. Taramasco C, Rodenas T, Martinez F, et al. A novel monitoring system for fall detection in older people. IEEE Access 2018;6(6):43563-43574. [doi: 10.1109/ACCESS.2018.2861331]
56. Schutz N, Botros A, Hassen SB, et al. A sensor-driven visit detection system in older adults' homes: towards digital late-life depression marker extraction. IEEE J Biomed Health Inform 2022 Apr;26(4):1560-1569. [doi: 10.1109/JBHI.2021.3114595] [Medline: 34550895]
57. Austin J, Dodge HH, Riley T, Jacobs PG, Thielke S, Kaye J. A smart-home system to unobtrusively and continuously assess loneliness in older adults. IEEE J Transl Eng Health Med 2016;4(4):2800311. [doi: 10.1109/JTEHM.2016.2579638] [Medline: 27574577]
58. Rajamäki J, Gioulekas F, Rocha PAL, Garcia XDT, Ofem P, Tyni J. ALTAI tool for assessing ai-based technologies: lessons learned and recommendations from SHAPES pilots. Healthcare (Basel) 2023 May 17;11(10):1454. [doi: 10.3390/healthcare11101454]
59. Najafi B, Aminian K, Paraschiv-Ionescu A, Loew F, Büla CJ, Robert P. Ambulatory system for human motion analysis using a kinematic sensor: monitoring of daily physical activity in the elderly. IEEE Trans Biomed Eng 2003 Jun;50(6):711-723. [doi: 10.1109/TBME.2003.812189] [Medline: 12814238]

XSL•FO
RenderX

60. Elhussein L, Jödicke AM, He Y, et al. Characterising complex health needs and the use of preventive therapies in the older population: a population-based cohort analysis of UK primary care and hospital linked data. BMC Geriatr 2023 Jan 31;23(1):58. [doi: 10.1186/s12877-023-03770-z] [Medline: 36721104]
61. Manh Do H, Sheng W, Harrington EE, Bishop AJ. Clinical screening interview using a social robot for geriatric care. IEEE Trans Automat Sci Eng 2021;18(3):1229-1242. [doi: 10.1109/TASE.2020.2999203]
62. Spinsante S, Poli A, Mongay Batalla J, et al. Clinically-validated technologies for assisted living. J Ambient Intell Human Comput 2023 Mar;14(3):2095-2116. [doi: 10.1007/s12652-021-03419-y]
63. Bajenaru L, Marinescu IA, Dobre C, Tomescu M, Herghelegiu AM. Clinically-validated technologies for older adults' quality of life self-management: vINCI ecosystem. Presented at: 2022 14th International Conference on Electronics, Computers and Artificial Intelligence (ECAI); Jun 30 to Jul 1, 2022; Ploiesti, Romania. [doi: 10.1109/ECAI54874.2022.9847465]
64. Delmastro F, Martino FD, Dolciotti C. Cognitive training and stress detection in MCI frail older people through wearable sensors and machine learning. IEEE Access 2020;8(8):65573-65590. [doi: 10.1109/ACCESS.2020.2985301]
65. Yao L, Sheng QZ, Li X, et al. Compressive representation for device-free activity recognition with passive RFID signal strength. IEEE Trans Mobile Comput 2018 Feb;17(2):293-306. [doi: 10.1109/TMC.2017.2706282]
66. Moreno PA, Hernando ME, Gómez EJ. Design and technical evaluation of an enhanced location-awareness service enabler for spatial disorientation management of elderly with mild cognitive impairment. IEEE J Biomed Health Inform 2015 Jan;19(1):37-43. [doi: 10.1109/JBHI.2014.2327638] [Medline: 25486651]
67. Andrew NE, Beare R, Ravipati T, et al. Developing a linked electronic health record derived data platform to support research into healthy ageing. Int J Popul Data Sci 2023;8(1):2129. [doi: 10.23889/ijpds.v8i1.2129] [Medline: 37670961]
68. Lin YH, Mao HF, Lin KN, Tang YL, Yang CL, Chou JJ. Development and evaluation of a computer game combining physical and cognitive activities for the elderly. IEEE Access 2020;8(8):216822-216834. [doi: 10.1109/ACCESS.2020.3041017]
69. McManus K, Greene BR, Ader LGM, Caulfield B. Development of data-driven metrics for balance impairment and fall risk assessment in older adults. IEEE Trans Biomed Eng 2022 Jul;69(7):2324-2332. [doi: 10.1109/TBME.2022.3142617] [Medline: 35025734]
70. Wu J, Chang L, Yu G. Effective data decision-making and transmission system based on mobile health for chronic disease management in the elderly. IEEE Syst J 2021;15(4):5537-5548. [doi: 10.1109/JSYST.2020.3024816]
71. Chen T, Ding Z, Li B. Elderly fall detection based on improved YOLOv5s network. IEEE Access 2022;10(10):91273-91282. [doi: 10.1109/ACCESS.2022.3202293]
72. Doyle J, Caprani N, Kealy A, Bond R, Komaba Y, Inomata A. Healthcare professionals views on technology to support older adults transitioning from hospital to home. Presented at: Proceedings of the 30th International BCS Human Computer Interaction Conference; Jul 11-15, 2016; Poole, United Kingdom. [doi: 10.14236/ewic/HCI2016.34]
73. Ayena JC, Zaibi H, Otis MJD, Menelas BAJ. Home-based risk of falling assessment test using a closed-loop balance model. IEEE Trans Neural Syst Rehabil Eng 2016 Dec;24(12):1351-1362. [doi: 10.1109/TNSRE.2015.2508960] [Medline: 26685258]
74. Muheidat F, Tawalbeh LA. In-home floor based sensor system-smart carpet- to facilitate healthy aging in place (AIP). IEEE Access 2020;8(8):178627-178638. [doi: 10.1109/ACCESS.2020.3027535]
75. Kim J, Cheon S, Lim J. IoT-based unobtrusive physical activity monitoring system for predicting dementia. IEEE Access 2022;10(10):26078-26089. [doi: 10.1109/ACCESS.2022.3156607]
76. Yu L, Chan WM, Zhao Y, Tsui KL. Personalized health monitoring system of elderly wellness at the community level in Hong Kong. IEEE Access 2018;6(6):35558-35567. [doi: 10.1109/ACCESS.2018.2848936]
77. van der Meer HG, Wouters H, Teichert M, et al. Feasibility, acceptability and potential effectiveness of an information technology-based, pharmacist-led intervention to prevent an increase in anticholinergic and sedative load among older community-dwelling individuals. Ther Adv Drug Saf 2019;10(10):2042098618805881. [doi: 10.1177/2042098618805881] [Medline: 31019675]
78. Nath RK, Thapliyal H. Smart wristband-based stress detection framework for older adults with cortisol as stress biomarker. IEEE Trans Consumer Electron 2021;67(1):30-39. [doi: 10.1109/TCE.2021.3057806]
79. Lutomski JE, Baars MAE, Schalk BWM, et al. The development of the Older Persons and Informal Caregivers Survey Minimum DataSet (TOPICS-MDS): a large-scale data sharing initiative. PLoS One 2013;8(12):e81673. [doi: 10.1371/journal.pone.0081673] [Medline: 24324716]
80. Lewis ET, Williamson M, Lewis LP, et al. The feasibility of deriving the electronic frailty index from Australian general practice records. Clin Interv Aging 2022;17(17):1589-1598. [doi: 10.2147/CIA.S384691] [Medline: 36353269]
81. Zherebtsov EA, Zharkikh EV, Loktionova YI, et al. Wireless dynamic light scattering sensors detect microvascular changes associated with ageing and diabetes. IEEE Trans Biomed Eng 2023;70(11):3073-3081. [doi: 10.1109/TBME.2023.3275654]
82. Lin CC, Chiu MJ, Hsiao CC, Lee RG, Tsai YS. Wireless health care service system for elderly with dementia. IEEE Trans Inform Technol Biomed 2006;10(4):696-704. [doi: 10.1109/TITB.2006.874196]

XSL•FO
RenderX

83. Garcia AP, De La Vega SF, Mercado SP. Health information systems for older persons in select government tertiary hospitals and health centers in the Philippines: cross-sectional study. J Med Internet Res 2022 Feb 14;24(2):e29541. [doi: 10.2196/29541] [Medline: 35156927]
84. Monitoring the Building Blocks of Health Systems: A Handbook of Indicators and Their Measurement Strategies: World Health Organization; 2010. URL: https://iris.who.int/handle/10665/258734 [accessed 2025-05-21]
85. Cave A. Exploring strategies for implementing data governance practices [dissertation]. : Walden University; 2017 Jan 1 URL: https://scholarworks.waldenu.edu/dissertations/4206
86. Health Data Governance. 2025. URL: https://healthdatagovernance.org/ [accessed 2025-05-21]
87. Aïdoud A, Gana W, Poitau F, et al. High prevalence of geriatric conditions among older adults with cardiovascular disease. J Am Heart Assoc 2023 Jan 17;12(2):e026850. [doi: 10.1161/JAHA.122.026850] [Medline: 36628962]
88. Gómez-Ramos JJ, González-Guerra M, Dávalos-Rodríguez IP, Pérez-Ruíz ME, Peña-Durán E, Marín-Medina A. Geriatric syndromes and their relationship with mortality in a population of Mexican older adults aged 65 and over, admitted to the emergency department of a second-level care hospital. Healthcare (Basel) 2024 Jun 8;12(12):1166. [doi: 10.3390/healthcare12121166]
89. Wang LY, Hu ZY, Chen HX, Tang ML, Hu XY. Multiple geriatric syndromes in community-dwelling older adults in China. Sci Rep 2024 Feb 12;14(1):3504. [doi: 10.1038/s41598-024-54254-y] [Medline: 38347070]
90. Ward KT, Reuben DB, Schmader KE, Givens J. Comprehensive geriatric assessment. UpToDate. 2024 May 17. URL: https://www.uptodate.com/contents/comprehensive-geriatric-assessment [accessed 2025-06-12]
91. Earley S, Data Management Association. The DAMA Dictionary of Data Management: Technics Publications; 2011.
92. Cleven A, Wortmann F. Uncovering four strategies to approach master data management. Presented at: 2010 43rd Hawaii International Conference on System Sciences; Jan 5-8, 2010; Honolulu, HI, United States. [doi: 10.1109/HICSS.2010.488]
93. El-Jardali F, Bou-Karroum L, Jabbour M, et al. Digital health in fragile states in the Middle East and North Africa (MENA) region: a scoping review of the literature. PLoS One 2023;18(4):e0285226. [doi: 10.1371/journal.pone.0285226]
94. Giebel GD, Speckemeier C, Abels C, et al. Problems and barriers related to the use of digital health applications: scoping review. J Med Internet Res 2023 May 12;25:e43808. [doi: 10.2196/43808] [Medline: 37171838]
95. Monlezun DJ, Omutoko L, Oduor P, et al. Digitalization of health care in low- and middle-income countries. Bull World Health Organ 2025 Feb 1;103(2):148-154. [doi: 10.2471/BLT.24.291643] [Medline: 39882493]
96. A. Alampay E. Open data governance in the time of the pandemic. : ACM Presented at: 14th International Conference on Theory and Practice of Electronic Governance; Oct 6-8, 2021; Athens, Greece p. 455-462. [doi: 10.1145/3494193.3494255]
97. Koltay T. Data governance, data literacy and the management of data quality. IFLA J 2016 Dec;42(4):303-312. [doi: 10.1177/0340035216672238]
98. Delacroix S, Lawrence ND. Bottom-up data trusts: disturbing the ‘one size fits all’ approach to data governance. Int Data Privacy Law 2019 Oct 1:ipz014. [doi: 10.1093/idpl/ipz014]
99. Yukhno A. Digital transformation: exploring big data governance in public administration. Public Organ Rev 2024 Mar;24(1):335-349. [doi: 10.1007/s11115-022-00694-x]
100. Platt JE, Raj M, Wienroth M. An analysis of the learning health system in its first decade in practice: scoping review. J Med Internet Res 2020 Mar 19;22(3):e17026. [doi: 10.2196/17026] [Medline: 32191214]
101. Zurynski Y, Smith CL, Vedovi A, et al. Mapping the learning health system: a scoping review of current evidence. In: NHRMC Partnership Centre for Health System Sustainability: Australian Institute of Health Innovation; 2020.

Abbreviations

ACM: Association for Computing Machinery
PAHO: Pan American Health Organization
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews

Edited by PH Chau; submitted 07.03.25; peer-reviewed by MA Zoubi, S Mohanadas, T Adebambo; revised version received 26.04.25; accepted 30.04.25; published 27.06.25.

Please cite as:
Garcia A, Balingbing AM, Palad Y
Exploring Literature on Data Governance in the Health Care of Older Persons: Scoping Review
JMIR Aging 2025;8:e73625
URL: https://aging.jmir.org/2025/1/e73625
doi:10.2196/73625

XSL•FO
RenderX

© Angely Garcia, Ayra Mae Balingbing, Yves Palad. Originally published in JMIR Aging (https://aging.jmir.org), 27.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Nonpharmacological Multimodal Interventions for Cognitive Functions in Older Adults With Mild Cognitive Impairment: Scoping Review

Raffy Chi-Fung Chan[1*], MA; Joson Hao-Shen Zhou[1*], MMedSc; Yuan Cao[2], PhD; Kenneth Lo[3,4], PhD; Peter Hiu-Fung Ng[1], PhD; David Ho-Keung Shum[1,4], PhD; Arnold Yu-Lok Wong[1,4], PhD

[1]Department of Rehabilitation Sciences, Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[2]Department of Social Work and Social Administration, University of Hong Kong, Hong Kong, China (Hong Kong)

[3]Department of Food Science and Nutrition, Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[4]Research Institute for Smart Ageing, Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[*]these authors contributed equally

Corresponding Author:
Arnold Yu-Lok Wong, PhD
Department of Rehabilitation Sciences
Hong Kong Polytechnic University
QT522, 5/F, Core T
The Hong Kong Polytechnic University
Hong Kong
China (Hong Kong)
Phone: 852 2766 6741
Email: arnold.wong@polyu.edu.hk

Abstract

Background: As the global population ages, the prevalence of dementia is expected to rise significantly. To alleviate the burden on health care systems and the economy, it is essential to develop effective strategies to enhance cognitive function in older adults. Previous studies have shown that combined nonpharmacological interventions can improve cognition across various domains in older individuals. However, there is no established gold standard for the exact combination and duration of these interventions, which makes it challenging to assess their overall effectiveness.

Objective: Given the diversity of nonpharmacological multimodal interventions aimed at preventing cognitive decline in older adults with mild cognitive impairment (MCI), this scoping review sought to identify and summarize the characteristics and outcomes of these interventions.

Methods: We adhered to the Arksey and O'Malley methodological framework and the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) and searched 4 electronic databases (MEDLINE, PsycINFO, CINAHL, and Web of Science) systematically on July 6, 2023, and updated the search on April 17, 2024, using specific terms and keywords.

Results: This review included 45 studies from 18 countries with 4705 participants from 2014 to 2024 encompassing different combinations of physical training (PT), cognitive training (CT), nutrition intervention, psychosocial intervention, social activities, and electrical stimulation. There is a growing numbers of studies combining PT and CT for MCI treatment, with additional modalities often added to address various aspects of the condition. Compared to single-modal interventions and usual care, multimodal approaches demonstrated significantly better improvements in cognition domains such as attention, global cognition, executive function, memory, processing speed, and verbal fluency. Technology has been instrumental in delivering these interventions and enhancing the effects of PT and CT. Multimodal interventions also show promise in terms of acceptability and user experience, which can improve treatment adherence.

Conclusions: Research is limited regarding the cost-effectiveness and optimal dosage of these interventions, making it difficult to assess the additional benefits of incorporating more modalities. Future research should examine the long-term effects of incorporating multiple modalities, using standardized MCI criteria and outcome measures.

XSL•FO
RenderX

(JMIR Aging 2025;8:e70291) doi:10.2196/70291

KEYWORDS

mild cognitive impairment; multimodal intervention; prevention; randomized controlled trial; cognitive decline

Introduction

Overview

Mild cognitive impairment (MCI) represents a transitional stage between normal aging and dementia, characterized by cognitive decline greater than expected for age, impacting one or more domains such as attention, memory, orientation, executive functioning, language, and visuospatial skills [1,2]. However, the decline does not significantly interfere with daily activities [3]. The lifetime prevalence of MCI in individuals aged above 60 years is estimated to be 15% to 20% [4], with this prevalence increasing with age. The annual transition rate from MCI to various subtypes of dementia ranges from 10% to 15% and can reach up to 25.2% for adults aged 80 to 84 years [5,6]. Older adults with MCI are 46% more likely to develop dementia within 3 years compared to just 3% in the normal-aging population [7].

The trajectory for individuals with MCI may vary, leading to dementia, stable cognition, or a return to normal cognition function [6]. Untreated MCI may progress to dementia, a neurodegenerative disease that significantly impacts daily functions and affects one's physical, psychological, social, and economic aspects. Dementia also directly affects one's caregivers, families, and society [8]. With the global population of individuals aged 60 years and above projected to double from 1 billion in 2020 to 2 billion by 2050 [9], there is mounting concern regarding the rising prevalence of dementia and the urgent need for preventive measures to address the associated social and economic burdens. Factors, such as modifiable risk factors, genetics, and interventions, can affect MCI progression [10]. Therefore, understanding various interventions and their effectiveness in preventing progression from MCI to dementia is crucial given the rising prevalence of dementia among older adults.

Currently, there is no gold standard for treatments or interventions to manage MCI. The American Academy of Neurology guidelines indicate insufficient empirical evidence to support pharmacological treatments for MCI in older adults [11]. Conversely, numerous studies have advocated nonpharmacological interventions, such as physical training (PT) or cognitive training (CT), as effective strategies for managing MCI [11-13].

Nonpharmacological Interventions

PT, including aerobic, strengthening, and balance exercises, has been shown to stimulate norepinephrine release in the brain, promote brain plasticity, increase brain volume, and enhance cerebral blood flow [14]. These effects are crucial for improving cognition, mood, and physiological abilities [15]. Combining aerobic and strengthening exercises has been identified as particularly effective for cognitive improvement [16].

On the contrary, CT primarily uses cognitive stimulation and repetitive tasks to enhance various cognitive domains, particularly memory, attention, and executive function [17]. The effectiveness of CT in improving cognition lies in its capacity to strengthen the functioning and plasticity of neural networks and cognitive reserves [17]. Research suggested that memory-focused CT increased activation and connectivity in the frontal, temporal, and occipital brain regions [18]. These areas are crucial for memory, motor function, processing, attention, language, mood, and problem-solving. Therefore, CT shows promise for improving overall cognition in older adults with MCI.

Given that both PT and CT can enhance brain plasticity and stimulate brain regions responsible for various cognitive functions, their combined application, whether delivered separately or through dual tasking, is the most common approach for managing cognitive impairment in MCI. This multifaceted approach effectively targets different aspects of cognitive decline, countering cognitive decay and neurodegeneration.

Growing evidence suggests that engaging in more social activities (SA) can lower the risk of cognitive decline in individuals with MCI [5], making SA a potential key modality for MCI management. SA involves participation in activities that allow interactions or engagements with others [5]. Karp et al [19] found that older adults with MCI who participated in a broader range of activities, including mental stimulation, physical activities, and social recreation, had a lower risk of developing dementia than those who participated in fewer or no such activities. On the other hand, beyond PT, CT, and SA, emerging evidence suggests that modifiable lifestyle factors, including diet and nutritional intervention (NI), electrical stimulation, or psychosocial intervention (PI) may also improve cognitive functions in this population [10,20].

Multimodal Interventions

A multimodal intervention, integrating various methods such as PT, CT, SA, NI, electrical stimulation, and PI, either sequentially or simultaneously, addresses cognitive decline across different domains of MCI. This approach has been proven effective in managing MCI in older adults, enhancing cognitive abilities, mood, sleep, activities of daily living, functional capacities, and physical abilities, with benefits lasting up to 2 years [14]. A systematic review and meta-analysis [21] showed that combining 2 or more interventions had small to medium effects on global cognition, memory, executive function, and verbal fluency, demonstrating a synergistic effect. Studies also showed that multimodal interventions outperformed single-modal interventions in managing MCI [15,16]. Another systematic review and meta-analysis [22] also supported that combined PT and CT had a small to medium effect on global cognition than various types of cognitive-only interventions in older adults with MCI. However, these reviews solely focused on the effects of combined PT and CT [22] or compared the

XSL•FO
RenderX

effectiveness of multimodal interventions only with single-modal interventions [21], thus lacking direct comparisons between various types of multimodal interventions to inform researchers or clinicians regarding which combinations of multimodal interventions would yield better results. In addition, previous reviews did not consider the user experience of multimodal interventions, which limits the clinical applicability of their findings and important factors to determine the feasibility of multimodal interventions. Thus, a comprehensive evaluation of multimodal interventions should include both feasibility and user experience to optimize benefits.

Optimizing Multimodal Interventions: Technology, Dosage, and Cost-Effectiveness

The increasing adoption of digital health care technology in managing MCI [23] could enhance the delivery and reduce the costs of interventions, especially benefiting those in remote areas [24]. Given the high prevalence and economic burden of MCI on communities and health care systems, technology such as the use of computerized CT (CCT) can help mitigate costs associated with nonpharmacological interventions, addressing the shortage of trained professionals. While previous systematic reviews and meta-analyses have acknowledged the role of technology in MCI management, they mainly focused on CCT [24-26] without considering the broader potential of technology beyond CCT.

Addressing the impact of different dosages of nonpharmacological interventions for MCI is crucial for developing effective and sustainable therapeutic strategies. This involves determining the optimal frequency, intensity, type, and duration of interventions to maximize cognitive benefits while balancing time and financial costs [27]. A study [27] highlighted that CT, PT, NI, and the majority of combined PT and CT significantly improved cognition in individuals with MCI, with particularly effective doses being 1 to 2 sessions per week with 60 to 120 minutes per session and interventions lasting over 12 weeks. However, as this study focused on only one type of multimodal intervention (combined PT and CT), the dosage effects of other nonpharmacological multimodal interventions remain uncertain. Therefore, this scoping review aimed to summarize dosage effects from the literature on multimodal interventions.

In addition, managing cognitive decline imposes a significant global economic burden. The cost-effectiveness of multimodal interventions varies by region because of the differences in health care resources. Although previous research supported the effectiveness of multimodal interventions for cognitive decline, their adoption of multiple modalities often increases cost. However, no prior review has summarized the cost-effectiveness of these interventions, so this scoping review aimed to address this important gap to inform clinical decision-making.

Rationale for a Scoping Review

Considering the variety of multimodal MCI interventions, both with and without technology, and the diverse methodologies and research focus of the existing literature, a scoping review is warranted to identify the current research gaps to inform future research and clinical practice. The complexity and heterogeneity of the interventions, coupled with the rapidly evolving nature of this field, make a systematic review less feasible for comprehensively mapping the use of different types of multimodal interventions and investigating the current research trends. As such, this scoping review aimed to map and describe the latest development in MCI interventions and provide researchers and clinicians with insights into current trends and limitations of the existing approaches and studies. Specifically, this review mapped the current landscape of nonpharmacological multimodal interventions for older adults with MCI, identified and summarized the components of these interventions, research trends, and the use of different outcome measures. It aimed to enhance the understanding of MCI management and provide future research directions on multimodal interventions.

Methods

Overview

This review adhered to the Arksey and O'Malley methodological framework [28] and the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) [29]. The protocol was registered on the Open Science Framework platform [30]. The review process involved five key steps: (1) formulating the research questions; (2) devising the search strategy; (3) identifying and selecting relevant studies; (4) data charting; and (5) synthesizing and presenting findings.

Identified Research Questions

This review explored the following research questions:

1. What are the research trends in multimodal interventions for older adults with MCI?
2. What components were included in these multimodal interventions? What results have been reported?
3. How cost-effective were the identified multimodal interventions?
4. What role did technology play in these interventions for older adults with MCI?
5. What insights were available regarding the acceptability, user experiences, and dose responses of these interventions?

Identifying and Selecting Relevant Studies

We conducted a search across 4 databases—MEDLINE, PsycINFO, CINAHL, and Web of Science—on July 6, 2023, and updated the search on April 17, 2024, using specific Medical Subject Headings terms and keywords such as "combine," "multi," "dual," "mix interventions, and "mild cognitive impairment" (Multimedia Appendix 1). Although quality assessment is optional for scoping reviews [31], this review included only randomized controlled trials to enhance study quality. Two independent reviewers (JHSZ and RCFC) screened titles, abstracts, and full texts against the eligibility criteria (Textbox 1). Disagreements were resolved by consensus or by consulting a third reviewer (AYLW). The interrater reliability of the screening process as measured by the kappa coefficient was 0.87.

XSL•FO
RenderX

Textbox 1. Selection criteria for the scoping review.

Inclusion criteria

- Participants: diagnosed with mild cognitive impairment (MCI) by clinicians or psychologists using well-established criteria
- Intervention: at least one combination of a nonpharmacological multimodal intervention in managing older adults with MCI
- Control: received at least one or multiple forms of an intervention, a placebo or sham training, health education, or treatment as usual
- Outcome: must use at least one well-established measurement for testing cognitive outcomes
- Study design: must be an experimental study (randomized controlled trial or quasi-experimental study)
- Other: full-text and peer-reviewed study written in English

Exclusion criteria

- Participants: diagnosed with dementia or the cognitive impairment resulted from drug use or psychiatric or other neurological disorders (eg, bipolar disorder, schizophrenia, stroke, Parkinson disease, or epilepsy)
- Intervention: pharmacological interventions or nonpharmacological experimental studies with a single-modal intervention
- Study design: systematic reviews, scoping reviews, opinion letters, conference proceedings, dissertations, and research design protocols
- Other: gray literature

Definition of Types of Intervention and Control Groups

The included studies featured diverse intervention components and control groups. Detailed operational definitions for these interventions and control groups are provided in Textbox 2.

Textbox 2. Operational definition of the intervention and control groups in the included studies.

Interventions

- Physical training: activities, exercises, or training that required older adults to do physical activities with or without the guidance or supervision of a professional trainer or clinician
- Cognitive training: activities or training that used a standardized systematic cognitive stimulation, rehabilitation task, and training to improve cognitive function
- Nutrition intervention: the use of any type of dietary supplement, including herb extract, or any form of dietary counseling
- Psychosocial intervention: the use of activities, training, counseling, therapy, or education that aimed to improve psychological well-being, including music therapy and mindfulness
- Social activities: activities that encourage social engagement or facilitate social interaction between older adults
- Electric current stimulation: the use of current stimulation including transcranial alternating current stimulation and transcranial direct current stimulation

Control groups

- Active control: a group of participants who received at least one form of intervention
- Placebo control: a group of participants who received a placebo or sham training
- Health education control: a group of participants who received health education
- Inactive control: a group of participants who received no additional treatment, treatment as usual, or only health advice

Data Extraction

Relevant data, including authors, publication year, country, place of recruitment, diagnostic criteria, participants, intervention types, outcomes, treatment frequency and duration, follow-up time points, use of measurement tools, control group characteristics, results, and interpretations of findings, were extracted by 2 independent reviewers (JHSZ and RCFC). All extracted findings were compared. Discrepancies were reconciled by consensus or through consultation with a third reviewer (AYLW).

Results

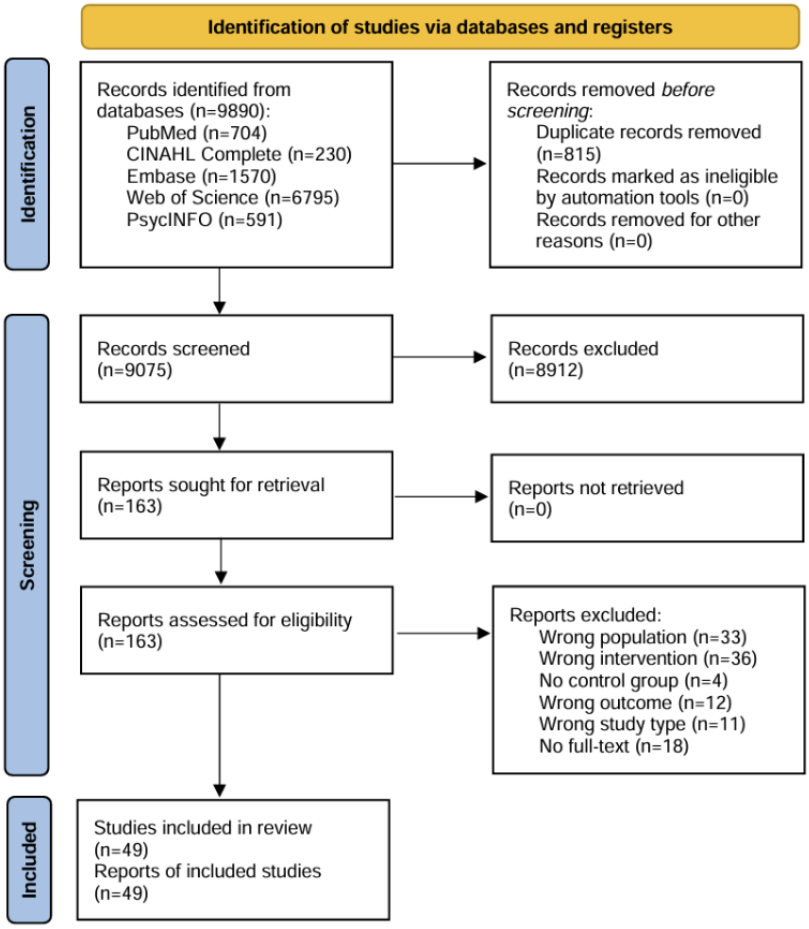
Study Selection and Characteristics

The initial search identified 9890 articles. After removing 815 duplicates, 9075 titles and abstracts were screened; 163 out of 9075 titles and abstracts were selected for full-text screening. Studies based on the same cohort were counted as a single study, including 4 studies by Hagovská and Nagyova [32], Hagovská and Olekszyova [33,34], and Hagovská et al [35] and 2 studies by Liao et al [36,37]. Exclusions were made for reasons

XSL•FO
RenderX

including incorrect target population (n=33), absence of multimodal intervention (n=36), lack of a control group (n=4), different study outcomes (n=12), absence of cognitive outcomes (n=11), and unavailability of full text (n=18). Ultimately, 45 studies from 49 articles published between 2014 and 2024, encompassing a total of 4705 participants, were included in this review. A detailed description of the study selection is presented in Figure 1.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram of literature search and screening.



The characteristics of the included studies, including participant characteristics, intervention tasks, treatment duration and frequency, and assessment time points, are presented in Multimedia Appendix 2. These studies used various criteria to diagnose MCI, such as the Peterson criteria (n=14) [15,16,38-49], the Albert criteria (n=5) [50-54], and cutoff scores from established screening tools like the Alzheimer Disease Assessment Scale–Cognitive subscale (cutoff score not provided) [38,40], Mini-Mental State Examination between 20 and 27 [55], Montreal Cognitive Assessment with scores up to 28 [15,23,36,37,52,56-65], Clinical Dementia Rating scores from 0.5 to less than 1 [45,60,66], or direct diagnoses from professional psychologists. Table 1 provides a detailed description of the diagnostic criteria used for a comprehensive reference.

Table 1. Diagnostic criteria used in the included studies.

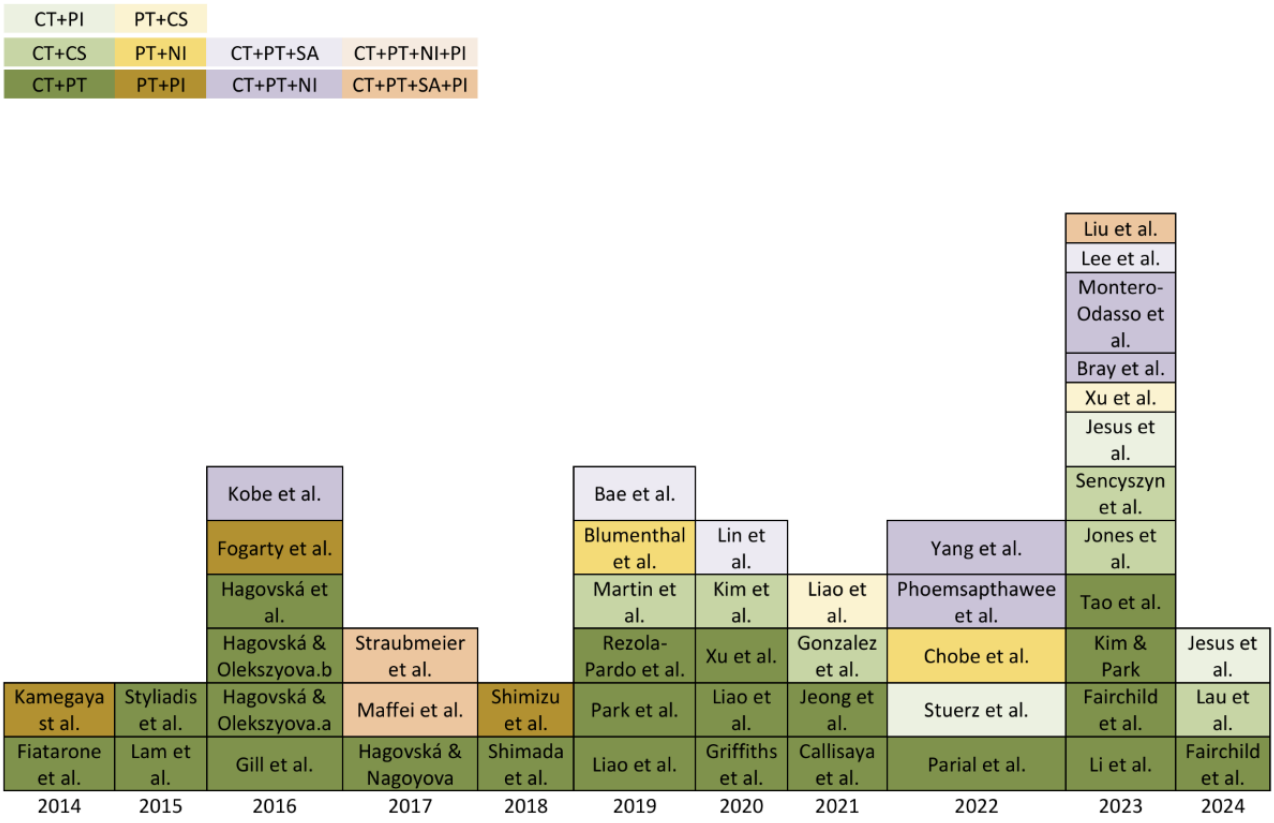
Criteria	Definition
Peterson criteria [11]	• Self-report cognitive decline • Objective cognitive impairment compared with age • No impact on daily functioning • No dementia
Albert criteria [3]	• Cognitive change reported by the patient, informant, or clinician • Objective evidence of cognitive impairment in one or more domains (typically 1 to 1.5 SDs below the mean when compared with their peers with matched age and education level) • Independent in functional abilities • No dementia

XSL•FO
RenderX

Publication or Study Trend

There has been a noticeable increase in both the number of studies and the variety of multimodal intervention combinations from 2014 to 2024 (Figure 2 [15,16,32-78]). Specifically, studies incorporating NIs as part of multimodal MCI interventions grew from 1 in 2016 to 3 in both 2020 and 2023. In addition, the use of technology in delivering multimodal interventions has gradually increased from 1 study in 2015 to 5 in 2023.

Figure 2. The number of studies that incorporated multimodal interventions for mild cognitive impairment throughout the years. CS: current stimulation; CT: cognitive training; NI: nutritional intervention; PI: psychosocial intervention; PT: physical training; SA: social activities.



Types of Multimodal Interventions

The included studies used combinations of 2, 3, and 4 types of interventions. Although most included studies primarily focused on combining PT and CT, there is an increasing number of studies integrating additional interventions to address the multifaceted nature of MCI. The following section details these specific combinations, the types of control groups used, and the overall outcomes across cognitive domains, such as attention, executive function, global cognition, memory, processing speed, and verbal fluency. Comprehensive outcomes, measurement tools, and targeted cognitive domains of the included studies are summarized in the following sections.

Bimodal Interventions

Overview

Most studies that used combined PT and current stimulation [45,67], or PT and NI [56,58] reported significantly greater improvements in various cognition domains compared to active control groups, while the use of combined CT and PI [61,68,69] reported significantly greater improvements compared to inactive control groups. However, studies with combined PT and CT showed mixed results. Conversely, most studies comparing combined PT and PI [39,41,48], CT and electroacupuncture [42], or CT and current stimulation [44,47,53,60,62] with an active or inactive control intervention found little to no significant advantage in the intervention groups (Table 2).

XSL•FO
RenderX

Table 2. Studies that incorporated a bimodal intervention and outcomes (n=35).

Study	Sample size (control or controls sample size)	Control or controls	Measured cognitive domains	Clinical outcomes
PT[a]+CT[b]				
Callisaya et al [57], 2021	44 (49)	• IN[c]	• GC[d] (COWAT[e], DSCT[f], HVLT[g], TMT-A[h], TMT-B[i], and SCWT[j])	• No significant difference between the two groups for all tests (P>.05)
Fairchild et al [51], 2024	36 (36)	• IN	• EF[k] (DST[l], TMT-B, and SCWT) • ME[m] (WMS-LM I[n] and II[o], and RAVLT[p]) • PS[q] (TMT-A and SCWT)	• No significant difference between the two groups for all tests (P>.05)
Fiatarone Singh et al [38], 2014	27 (22; 24; 27)	• CT only • PT only • PC[r]	• ATT[s] (SDMT[t]) • GC[u] (ADAS-Cog[v]) • EF (Matrices and Similarities subtests of the WAIS-III[w]) • ME (ADAS-Cog, BVRT[x], and LM[y] from WMS-III[z]) • PS (SDMT) • VF[aa] (COWAT and animal naming)	• The intervention group did not show significant differences compared to other groups across all tests (P>.05)
Gill et al [59], 2016	23 (21)	• PT only	• GC (composite score) • EF (TMT-A and TMT-B) • ME (AVLT[ab]) • PS (DSST[ac]) • VF (VFT[ad])	• GC (mean difference 0.2, P=.04[ae]) • EF (mean difference 0.11, P=.60) • ME (mean difference 0.3, P=.02) • PS (mean difference −0.06, P=.78) • VF (mean difference 0.62, P=.003)
Griffiths et al [52], 2020	35 (35)	• IN	• ATT (TMT-A and TMT-B) • EF (BD[af] of the WAIS-IV[ag]) • ME (DST-F[ah], DST-B[ai], DSS[aj] from WAIS-IV, VFT, and WLL[ak])	• ATT (TMT-A: mean difference −27.86, P=.36) • EF (mean difference 0.34, P=.58) • ME (DST-F and DST-B: mean difference 1.23, P=.60; DSS: mean difference 1.030, P=.31; VFT-Letter[al]: mean difference 4.59, P=.001; VFT-Category[am]: mean difference 2.81, P=.23)
Hagovská et al [35], 2016	40 (40)	• PT only	• ATT (SCWT) • GC (ACE[an] and MMSE[ao]) • ME (ACE and AVLT) • PS (DRT-II[ap]) • VF (ACE)	• ATT (SCWT: $\eta 2$=0.0001, P=.97) • GC (ACE: Cohen d=0.71, P=.002; MMSE: $\eta 2$=0.189, P=.001) • ME (ACE: Cohen d=0.64, P=.007; AVLT: $\eta 2$=0.173, P=.001) • PS (DRT-II: $\eta 2$=0.033, P=.11) • VF (Cohen d=0.73, P=.001)
Jeong et al [40], 2021	13 (13)	• HE[aq]	• ATT (TMT-A and TMT-B) • GC (ADAS-Cog and KMMSE[ar]) • PS (DSST)	• Group×time interaction • ATT—TMT-A: P<.05; TMT-B: P=.01 • GC (ADAS-Cog: P=.11); KMMSE (P=.72) • PS (P=.02)
Kim and Park [63], 2023	21 (21)	• CT only	• EF (EFPT-K[as] and FAB[at])	• EF (EFPT-K: $\eta 2$=0.132, P<.01; FAB: $\eta 2$=0.305, P<.001)
Lam et al [74], 2015	93 (115; 114; 101)	• CT only • PT only • SA only	• ATT (VFT-Category) • GC (ADAS-Cog, CDR-SOB[au], and CMMSE[av]) • ME (list learning delayed recall test)	• ATT (VFT-C: $\chi 2$=23.38, P<.001) • GC (ADAS-Cog: $\chi 2$=3.31, P=.2; CDR-SOB: $\chi 2$=1.82, P=.61; CMMSE: $\chi 2$=4.28, P=.23) • ME ($\chi 2$=3.31, P=.35)

XSL•FO
RenderX

Study	Sample size (control or controls sample size)	Control or controls	Measured cognitive domains	Clinical outcomes
Li et al [66], 2023	19 (35; 30)	• CT only • IN	• ATT, EF, ME, language, and spatial ability (AVLT, STT-A[aw] and STT-B[ax], CFT[ay], SCWT, and BNT-30[az]) • GC (ADAS-Cog and CMMSE)	• ATT, EF, and ME: (AVLT immediate recall: mean difference 0.46, *P*>.05; STT-A: mean difference −0.28, *P*>.05; STT-B: mean difference −0.15, *P*>.05; CFT: mean difference 0.34, *P*>.05; BNT-30: mean difference −1.14, *P*>.05) • GC: ADAS-Cog (mean difference 0.03, *P*>.05; MMSE: mean difference 0.22, *P*>.05)
Liao et al [36,37], 2019, 2020	16 (18)	• PT and CT	• GC (MoCA[ba]) • EF (EXIT-25[bb] and SCWT) • ME (CVVLT[bc])	• Group×time interaction • GC (*P*=.18) • EF (EXIT-25: *P*=.72; SCWT: number: *P*=.84; time: *P*=.32) • ME (CVVLT: immediate recall: *P*=.15; delayed recall: *P*=.12)
Parial et al [15], 2022	25 (26)	• HE	• GC (MoCA) • EF (TMT-B) • ME (DST)	• GC (Wald χ2=26.88, *P*<.001) • EF (Wald χ2=18.67, *P*<.001) • ME (immediate recall: Wald χ2=16.97, *P*<.001; delayed recall: Wald χ2=11.89, *P*<.003)
Park et al [72], 2019	25 (24)	• IN	• GC (ADAS-Cog) • EF (DSST) • ME (DST)	• GC: ADAS-Cog—*P*<.01; 95% CI −1.9 (−0.9 to −2.3) • EF: DSST—*P*<.01; 95% CI −6.3 (−4.2 to −8.4) • ME: DST—*P*=.02; 95% CI −0.6 (−0.3 to 1.6)
Rezola-Pardo et al [70], 2019	43 (42)	• PT only	• Overall cognitive performance (MoCA, symbol search and coding test from WAIS-IV, semantic fluency test, VFT, RAVLT, and TMT-A)	• MoCA (Cohen d=0.15; *P*=.90) • Symbol search (Cohen d=0.06; *P*=.85) • Semantic fluency (Cohen d=0.01; *P*=.28) • VFT (Cohen d=0.02; *P*=.67) • RAVLT (Cohen d=0.11; *P*=.70) • TMT-A (Cohen d=0.04; *P*=.65)
Shimada et al [71], 2018	137 (129)	• HE	• GC (MMSE) • EF (TMT-A and TMT-B) • ME (WMS-LM II and RAVLT) • VF (VFT)	• CG (mean difference 0.8, *P*=.01) • EF (mean difference −0.4, *P*=.35) • ME (WMS-LM II: mean difference 1.0, *P*=.004; RAVLT: mean difference 0.2, *P*=.35) • VF: (mean difference 2.2, *P*=.002)
Styliadis et al [16], 2015	14 (14; 14; 14; 14)	• PT only • CT only • HE • IN	• GC (MMSE)	• The intervention group did not show significant differences compared to other groups across all tests (*P*>.05)
Tao et al [73], 2023	51 (52)	• IN	• GC (MoCA and MMSE)	• GC (MoCA: η2=0.442, *P*=.001; F=39.550; MMSE: η2=0.33, *P*<.001; F=24.614)
Xu et al [65], 2020	5 (6; 6)	• NI only • IN	• GC (HK-MoCA[bd])	• GC: improvement only in the NI[be]-only group (*P*=.008)
PT+PI[bf]				

XSL•FO
RenderX

Study	Sample size (control or controls sample size)	Control or controls	Measured cognitive domains	Clinical outcomes
Fogarty et al [39], 2016	22 (18)	• PI only	• ATT (TEA[bg]) • ME (RBMT[bh]) • PS (digit symbol test)	• ATT: TEA (visual selective ATT: Cohen d=0.15, *P*=.82; F2,70=0.196; sustained ATT: Cohen d=0.127, *P*=.88; F2,62=0.125; ATT switching: Cohen d=0.0487, *P*=.13; F2,70=2.071) • ME (Cohen d=0.583, *P*=.06; F2,70=2.972) • PS (Cohen d=0.16, *P*=.80; F2,72=0.230)
Kamegaya et al [41], 2014	19 (24)	• IN	• GC (Five-Cog test) • EF (DSST and YKSSTbi)	• GC (F=2.999, *P*=.09; F1,38=2.999) • EF (DSST: F=1.165, *P*=.29; YKSST: F=0.096, *P*=.76; F1,38=1.165)
Shimizu et al [48], 2018	30 (9)	• PT only	• GC (FAB)	• GC: effect size (r=0.002, *P*=.81)
PT+NI				
Blumenthal et al [56], 2019	40 (41; 41; 38)	• PT only • NI only • HE	• GC (modified CDR-SOB) • EF (TMT-A, TMT-B, SCWT, DST-F, DST-B, DSST, Ruff 2 and 7 Test, and animal naming) • ME (HVLT-revised and CFT) • VF (COWAT and animal naming)	• GC (Cohen d=0.36, *P*=.03) • EF (Cohen d=0.4, *P*=.01) • ME (PT factor: Cohen d=0.19, *P*=.24; NI factor: Cohen d=0.15, *P*=.35) • VF (PT factor: Cohen d=0.12, *P*=.45; NI factor: Cohen d=0.23, *P*=.24)
Chobe et al [58], 2022	24 (25; 23)	• PT only • NI only	• ATT (DST-F and DST-B) • EF (TMT-A and TMT-B) • ME (RAVLT) • VF (COWAT)	• ATT (DST-F: F=2.921, *P*=.06; DST-B: F=5.766, *P*=.005) • EF (TMT-A: F=0.837, *P*=.44; TMT-B: F=0.677, *P*=.51) • ME (F=4.727, *P*=.01) • VF (F=3.028, *P*=.06)
CT+electroacupuncture				
Kim et al [42], 2020	16 (16)	• CT only	• GC (ADAS-Cog-K[bi] and MoCA-K[bj])	• GC: (ADAS-Cog-K: Z=−0.38, *P*=.70; MoCA-K: Z=−0.72, *P*=.47)
CT+PI				
Stuerz et al [68], 2022	26 (24)	• Reverse sequential control	• ATT (ACT) • GC (MMSE)	• ATT: ACT (η2<0.01, *P*=.21) • GC: MMSE (η2=0.05, *P*=.09)
Jesus et al [68], 2023a	27 (24)	• IN	• ATT (digit symbol coding of WAIS-III) • GC (ACE-R[bk]) • ME (ACE-R; WL[bl] and LM of WMS-III)	• ATT (F1,49=11.64; η2=0.19, *P*=.001) • GC (F1,49=4.70; η2=0.09, *P*=.04) • ME: ACE-R (F1,49=7.01; η2=0.13, *P*=.01); WL (F1,49=23.76; η2=0.33, *P*<.001); LM (F1,49=10.98; η2=0.18, *P*=.002)
Jesus et al [69], 2023b	98 (101)	• IN	• ATT (digit symbol coding of WAIS-III) • GC (ACE-R) • ME (LM of WMS-III)	• ATT (Cohen d=0.41, *P*<.001) • GC (Cohen d=0.33, *P*<.001) • ME (immediate recall: Cohen d=0.42, *P*<.001; delayed recall: Cohen d=0.35, *P*<.001)
CT+current stimulation				

XSL•FO
RenderX

Study	Sample size (control or controls sample size)	Control or controls	Measured cognitive domains	Clinical outcomes
Gonzalez et al [60], 2021	22 (24; 21)	• Sham tDCS[bm]+CT • CT only	• ATT (TMT-A) • GC (MoCA) • ME (DST and RBMT, 3rd edition)	• ATT: TMT-A (F=0.64, *P*=.62) • GC: MoCA (F=0.34, *P*=.85) • ME: DST (F=1.99, *P*=.09); RBMT-3 (F=0.13, *P*=.96)
Jones et al [62], 2023	13 (12)	• PC+CT	• ATT (ACE-X[bn]) • ME (ACE-X)	• ATT (Cohen d=0.94, *P*=.047) • ME (η2=0.05, *P*=.35)
Martin et al [53], 2019	33 (35)	• CT+PC	• ME (PALT[bo] and RVIP[bp]) • PS (SDMT) • Subjective cognitive functioning (CFQ[bq])	• ME: PALT (F=0.26, *P*=.77); RVIP (F=0.2, *P*=.82) • PS (F=1.39, *P*=.25) • Subjective cognitive functioning (F=1.25, *P*=.29)
Lau et al [44], 2024	11 (11)	• CT+PC	• GC (MoCA) • EF (TMT-A, TMT-B, and Tower of London) • ME (N-back task and CVVLT)	• GC: MoCA (η2=0.05, *P*=.35) • EF: TMT-A (η2=0.18, *P*=.06); TMT-B (η2=0.12, *P*=.13) • Tower of London (η2=0.03, *P*=.42); • ME: N-back (1-back: η2=0.17, *P*=.07; 2-back: η2=0.16, *P*=.07); CVVLT (verbal memory: η2=0.12, *P*=.07; delayed recall: η2=0.02, *P*=.59)
Senczyszyn et al [47], 2023	13 (13; 12)	• Current stimulation only • Sham current stimulation only	• GC (CANTAB[br]) • VF (Verbal Fluency FAS test)	• GC: CANTAB (Swms6: η2=0.042, *P*>.05; Palta4: η2=0.019, *P*=.03; Prmpci: η2=0.091, *P*=.02) • VF (η2=0.042, *P*>.05)
PT+current stimulation				

XSL•FO
RenderX

Study	Sample size (control or controls sample size)	Control or controls	Measured cognitive domains	Clinical outcomes
Liao et al [45], 2021	10 (10)	• PT+PC	• GC (MoCA) • EF (TMT-A and TMT-B) • ME (change detection task and CVVLT)	• GC: F1,18=0.246; η2=0.246, *P*=.63 • EF: TMT-A (F1,18=1.022; η2=0.054, *P*=.33); TMT-B (η2=0.271, *P*=.02) • ME: change detection task (F1,18=1.046; η2=0.058, *P*=.32); CVVLT (F1,18=0.024; η2=0.001, *P*=.88)
Xu et al [67], 2023	44 (49; 44; 43)	• PT+sham current stimulation • Sham PT+current stimulation; sham PT+sham current stimulation	• ATT (TAP[bs]) • GC (MoCA) • EF (SCWT) • ME (AVLT, ROCF, and Chinese Wechsler Memory Scale—Revised MQ[bt])	• ATT (auditory: F=2.66, *P*=.05; visual: F=4.536, *P*=.004; sustained reaction: F=3.609, *P*=.02) • GC (F=7.415, *P*<.001) • EF (F=.058, *P*=.98) • ME: AVLT (immediate recall: F=3.207, *P*=.03; delayed recall: F=1.13, *P*=.34); ROCF (copy: F=2.489, *P*=.062; recall: F=0.571, *P*=.64); MQ (F=3.584, *P*=.02)

[a]PT: physical training.
[b]CT: cognitive training.
[c]IN: inactive control.
[d]GC: global cognition.
[e]COWAT: Controlled Oral Word Association Test.
[f]DSCT: Digit Symbol Substitution Test.
[g]HVLT: Hopkins Verbal Learning Test.
[h]TMT-A: Trail Making Test–A.
[i]TMT-B: Trail Making Test–B.
[j]SCWT: Stroop Color-Word Test.
[k]EF: executive function.
[l]DST: digit span test.
[m]ME: memory.
[n]WMS-LM I: Wechsler Memory Scale-Revised–Logical Memory I.
[o]WMS-LM II: Wechsler Memory Scale-Revised–Logical Memory II.
[p]RAVLT: Rey Auditory Verbal Learning Test.
[q]PS: processing speed.
[r]PC: placebo control.
[s]ATT: attention.
[t]SDMT: Symbol Digit Modalities Test.
[u]GC: global cognition.
[v]ADAS-Cog: Alzheimer Disease Assessment Scale–Cognitive.
[w]WAIS-III: Wechsler Adult Intelligence Scale–III.
[x]BVRT: Benton Visual Retention Test.
[y]LM: logical memory.
[z]WMS-III: Wechsler Memory Scale–III.
[aa]VF: verbal fluency.
[ab]AVLT: Auditory Verbal Learning Test.
[ac]DSST: Digit Symbol Substitution Test.
[ad]VFT: Verbal Fluency Test.
[ae]Italicized values indicate statistical significance.
[af]BD: block design.
[ag]WAIS-IV: Wechsler Adult Intelligence Scale–IV.
[ah]DST-F: Digit Span Test–Forward.
[ai]DST-B: Digit Span Test–Backward.

XSL•FO
RenderX

[aj]DSS: digit span sequence.
[ak]WLL: Word-List Learning Test.
[al]VFT-Letter: Verbal Fluency Test–Letter.
[am]VFT-Category: Verbal Fluency Test–Category.
[an]ACE: Addenbrooke Cognitive Examination.
[ao]MMSE: Mini-Mental State Examination.
[ap]DRT-II: Disjunctive Reaction Time.
[aq]HE: health education.
[ar]KMMSE: Korean version of the Mini-Mental State Examination.
[as]EFPT-K: Executive Function Performance Test.
[at]FAB: Frontal Assessment Battery.
[au]CDR-SOB: Clinical Dementia Rating Sum of Boxes.
[av]CMMSE: Mini-Mental State Examination, Chinese version.
[aw]STT-A: Shape Trail Test–A.
[ax]STT-B: Shape Trail Test–B.
[ay]CFT: Complex Figure Test.
[az]BNT-30: Boston Naming Test.
[ba]MoCA: Montreal Cognitive Assessment.
[bb]EXIT-25: Executive Interview-25.
[bc]CVVLT: Chinese version of the Verbal Learning Test.
[bd]HK-MoCA: Montreal Cognitive Assessment, Hong Kong version.
[be]NI: nutritional intervention.
[bf]PI: psychosocial intervention.
[bg]TEA: Test of Everyday Attention.
[bh]RBMT: Rivermead Behavioral Memory Test.
[bi]ADAS-Cog-K: Alzheimer Disease Assessment Scale–Cognitive, Korean version.
[bj]MoCA-K: Montreal Cognitive Assessment, Korean version.
[bk]ACE-R: Addenbrooke Cognitive Examination–Revised.
[bl]WL: waitlist control.
[bm]tDCS: transcranial direct current stimulation.
[bn]ACE-X: Adaptive Cognitive Evaluation-Explorer
[bo]PALT: Paired Associative Learning Test.
[bp]RVIP: Rapid Visual Information Processing.
[bq]CFQ: Cognitive Failures Questionnaire.
[br]CANTAB: Cambridge Neuropsychological Test Automated Battery.
[bs]TAP: Test Of Attentional Performance.
[bt]MQ: memory quotient.

The Combination of PT and CT

Among these 18 studies, 4 used a single-modal active control group [34,59,63,70], 3 compared against a health education group [15,40,71], and 5 used a passive control group [51,52,57,72,73]. Two studies used both a single-modal active control and a passive control group [65,66]. One study compared results with 2 single-modal active controls and a placebo group [38], while another used 3 single-modal active control groups [74]. In addition, 1 study contrasted results with 2 single-modal active controls, a health education group, and a passive control group [16]. Another study used a bimodal active control group [36].

Among these studies, 7 reported significantly greater improvements in all measured cognitive functions within the bimodal intervention groups [15,35,40,52,59,63,72]. Three studies revealed improvements in both bimodal intervention and control groups [16,37,61,70]. One study found greater improvements in executive function, global cognition, and verbal fluency within the PT-only control group, whereas attention and processing speed improved significantly across all groups [38]. Another study reported no changes in cognitive function across any group [57]. In addition, 1 study observed improvements in attention, episodic memory, and working memory across all groups but no changes in global cognition [74]. Finally, 1 study reported greater improvement in global cognition within the NI-only control group [65].

Three Studies Investigated the Combination of PT and PI

Among these 3 studies, 2 studies compared results with a single-modal active control group [39,48], and one compared with a passive control group [41]. One study found that both the bimodal intervention and control groups enhanced attention, memory, and processing speed, but only the intervention group significantly improved executive function [39]. Another study

XSL•FO
RenderX

revealed improvements in global cognition only within the bimodal intervention group, although both groups improved in executive function [41]. In addition, one study reported significantly greater improvements in global cognition exclusively in the bimodal intervention group [48].

Two Studies Investigated the Combination of PT and NI

One study compared the intervention group with 2 single-modal active control groups and a health education control group [56]. It found that both the bimodal intervention and PT-only groups showed greater improvements in executive function and global cognition, although no gains were observed in memory or verbal fluency across groups [56]. Conversely, another study contrasted the intervention group with 2 single-modal active control groups [58] and revealed improvements in global cognition in all groups [58].

One Study Investigated the Combination of CT and Electroacupuncture

This study compared the intervention group with a single-modal active control group. It found that all groups demonstrated significant improvements in global cognition [42].

Three Studies Investigated the Combination of CT and PI

One study compared a reverse sequential control group [68], whereas 2 studies compared results with a passive control group [61,69]. Results from the latter 2 studies revealed significantly greater improvements in attention, episodic memory, and global cognition in the bimodal intervention group [61,69]. Conversely, the study with a reverse sequential control group found enhanced global cognition in both groups. Notably, the group starting with the PI showed more significant attention improvements after the initial training, while the group beginning with CT exhibited greater improvements following the second, combined training phase [68].

Five Studies Investigated the Combination of CT and Current Stimulation

Two of these studies compared the intervention group with a bimodal active control group [53,62]. One study contrasted the intervention group with a single-modal and a bimodal active control group [60]. Another study used a combination of CT and sham current stimulation as a control group [44]. One study included a current stimulation–only group, a sham current stimulation–only group, and a sham current stimulation control group [47].

One study observed improvements in everyday memory and global cognition across all groups, with attention improvements unique to the bimodal intervention group [60]. However, the CT-only control group showed greater working memory improvements than the bimodal intervention group [60]. Notably, transcranial direct current stimulation (tDCS) appeared to enhance the efficacy of CT in processing speed [60], despite not being a direct target. Another study showed that the bimodal intervention group demonstrated more significant attention improvements than both the placebo and CT groups [62]. In addition, 1 study found improvements in attention; processing speed; subjective cognitive functioning; and verbal, visual, and working memory across all groups [53]. Another study revealed improvements in episodic memory, executive function, and visual memory in both bimodal and active control groups [44].

Two Studies Investigated the Combination of PT and Transcranial Brain Stimulation

One study compared the intervention group with a bimodal active control group (PT and sham brain stimulation) [45], while another study used 2 bimodal active controls (PT and sham stimulation and sham PT with brain stimulation) along with a passive control group [67].

Results showed improvements in episodic memory, executive function, global cognition, and visual working memory in both the bimodal intervention and control groups [45]. However, the bimodal intervention group displayed greater enhancements across all cognitive functions compared to the control groups [67].

Trimodal Interventions

Overview

Studies comparing combined PT, CT, and SA with health education controls reported significantly better cognitive improvements, especially in the memory domain [55,75,76]. Similarly, 3 out of 5 studies using combined PT, CT, and NI [46,49,54] showed significantly greater improvements in global cognition compared to control groups (Table 3).

XSL•FO
RenderX

Table 3. Studies that incorporated trimodal intervention and outcomes (k=8).

Study	Sample size (control sample size)	Control	Measured cognitive domains	Clinical outcomes
PT[a]+CT[b]+NI[c]				
Bray et al [50], 2023	18 (15; 20; 15; 14)	• PT+CT+PC[d] • PT+NI+PC • PT+PC+PC • PC only	• GC[e] (ADAS-Cog[f]) • EF[g] and ME[h] (ADAS-Cog, TMT-A[i], and TMT-B[j])	• The intervention group did not show significant differences compared to other groups across all tests (*P*>.05)
Kobe et al [43], 2016	13 (9)	• NI[k]	• GC (AVLT[l], DST[m], VFT[n], TMT-A, TMT-B, and SCWT[o])	• No significant difference between the two groups for all tests (*P*>.05)
Montero-Odasso et al [54], 2023	34 (35; 37; 35)	• PT+CT+sham NI • PT+sham CT+NI • PT+sham CT+sham NI • Sham PT+sham CT+sham NI	• GC (ADAS-Cog) • EF (ADAS-Cog Plus variant)	• GC (mean difference −2.64, *P*=.005[p], Cohen d=0.71) • EF (ADAS-Cog Plus variant: no significant improvement; *P*>.05)
Phoemsapthawee et al [46], 2022	20 (20; 18)	• PT only • PC	• GC (MMSE[q]) • EF (TMT-A and TMT-B) • ME (DST-F[r] and DST-B[s])	• GC (F=13.158; η2=0.328, *P*<.001) • EF: TMT-A (F=30.142; η2=0.527, *P*<.001); TMT-B (η2=0.376, *P*<.001) • ME: DST-F (F=17.208; η2=0.389, *P*<.001); DST-B (η2=.495, *P*<.001)
Yang et al [49], 2022	55 (57)	• IN	• GC (MoCA[t])	• GC: group×time interaction ($\chi^2$3=303.928; *P*<.001)
PT+CT+SA				
Bae et al [55], 2019	41 (42)	• HE[u]	• GC (MMSE) • EF (TMT-A and TMT-B) • ME (spatial span task)	• GC (*P*=.14) • EF: TMT-A (*P*=.43); TMT-B (*P*=.68) • ME (*P*=.02)
Lee et al [75], 2023	140 (140)	• HE	• ATT[v] and EF (TMT-A and TMT-B) • ME (NCGG-FAT[w]) • PS[x] (DSST[y])	• ATT and EF: TMT-A (F=0.25); TMT-B (F=0.00) • ME (F=5.04, *P*<.05) • PS (F=0.67)
Lin et al [76], 2020	61 (61)	• HE	• ATT (SLUMS[z]) • GC (SLUMS) • EF (SLUMS)	• ATT (Wald χ2=50.84, *P*<.001) • GC (Wald χ2=252.81, *P*<.001) • EF (Wald χ2=115.99, *P*<.001)

[a]PT: physical training.

[b]CT: cognitive training.

[c]IN: inactive control.

[d]PC: placebo control.

[e]GC: global cognition.

[f]ADAS-Cog: Alzheimer Disease Assessment Scale–Cognitive.

[g]EF: executive function.

[h]ME: memory.

[i]TMT-A: Trail Making Test–A.

[j]TMT-B: Trail Making Test–B.

[k]NI: nutritional intervention.

[l]AVLT: Auditory Verbal Learning Test.

[m]DST: Digit Span Test.

[n]VFT: Verbal Fluency Test.

[o]SCWT: Stroop Color-Word Test.

[p]Italicized values indicate statistical significance.

[q]MMSE: Mini-Mental State Examination.

XSL•FO
RenderX

[r]DST-F: Digit Span Test–Forward.

[s]DST-B: Digit Span Test–Backward.

[t]MoCA: Montreal Cognitive Assessment.

[u]HE: health education.

[v]ATT: attention.

[w]NCGG-FAT: National Center for Geriatrics and Gerontology–Functional Assessment Tool.

[x]PS: processing speed.

[y]DSST: Digit Symbol Substitution Test.

[z]SLUMS: Saint Louis University Mental Status Examination.

Five Studies Investigated the Combination of PT, CT, and NI

One study compared the trimodal intervention group with an NI-only active control group [43], another compared it with both a PT-only active control and a placebo group [46], and a third study used an inactive control group [49]. One study contrasted outcomes with 4 control groups: combined PT and CT with placebo NI; combined PT and NI with sham CT; combined PT with sham CT and sham NI; and an inactive control group [54]. The fifth study compared the trimodal intervention with various combinations of PT, CT, and NI placebos [50].

In the first study, the trimodal intervention group study reported no substantial changes in cognitive functions across all groups [43]. Conversely, the second study found the intervention group showed significantly greater improvements in all cognitive functions compared with the placebo group, but not when compared to the PT-only control group [46] and the third study observed more significant enhancements in global cognition in the trimodal intervention group [49]. The fourth study also found greater improvements in global cognition in the trimodal intervention group, but no significant changes in other measures [54]. In the last study, the trimodal intervention group showed no significant difference compared to the controls [50].

Three Studies Investigated the Combination of PT, CT, and SA

All 3 studies contrasted the intervention group with a health education control group, revealing significant improvements in global cognition and memory for the trimodal intervention group [55,76]. In addition, 1 study found significant enhancements in memory compared to control groups [55], while another reported significant improvements in both attention and executive function during a 12-month follow-up in the intervention group [76].

Quadrimodal Interventions

Most studies incorporating quadrimodal interventions reported small to moderate effects across various cognition domains, especially in global cognition (Table 4).

XSL•FO
RenderX

Table 4. Studies that incorporated quadrimodal interventions and outcomes (n=3).

Study	Sample size (control sample size)	Control	Measured cognitive domains	Clinical outcomes
PT[a]+CT[b]+PI[c]+SA[d]				
Maffei et al [77], 2017	55 (58)	IN[e]	• GC[f] (ADAS-Cog[g])	• GC (Effect Size=−0.55, *P*<.001[h])
Straubmeier et al [64], 2017	208 (154)	IN	• GC (MMSE[i])	• GC (Cohen d=0.26, *P*=.01)
PT+CT+PI+NI[j]				
Liu et al [78], 2023	86 (106)	IN	• GC (MMSE and MoCA[k]) • ME (AVLT[l] and PALT[m])	• GC (Hedges g=0.40, *P*=.03) • ME (*P*=.05)

[a]PT: physical training.

[b]CT: cognitive training.

[c]PI: psychosocial intervention.

[d]SA: social activities.

[e]IN: inactive control.

[f]GC: global cognition.

[g]ADAS-Cog: Alzheimer Disease Assessment Scale–Cognitive.

[h]Italicized values indicate statistical significance.

[i]MMSE: Mini-Mental State Examination.

[j]NI: nutritional intervention.

[k]MoCA: Montreal Cognitive Assessment.

[l]AVLT: Auditory Verbal Learning Test.

[m]PALT: Paired Associative Learning Test.

Two Studies Investigated the Combination of PT, CT, PI, and SA

The 2 studies compared a quadrimodal intervention group with a passive control group [64,77]. One study found significantly greater improvements in global cognition in the quadrimodal intervention group [77]. In contrast, the other study reported no significant improvement in cognitive function within the intervention group, but the control group experienced deterioration [64].

One Study Investigated the Combination of PT, CT, PI, and NI

Compared to a passive control group, the quadrimodal intervention group showed significantly larger posttreatment improvements in global cognition and memory. However, these effects showed no difference between the two groups at the 12-month follow-up [78].

The Use of Technology in MCI Interventions

Between 2015 and 2024, a total of 16 studies incorporated technology to enhance intervention delivery. The Nintendo Wii (n=1) was first used as a hardware platform for delivering PT [16]. Subsequent studies used various technologies, including virtual reality (VR) technology (n=1) for combined PT and CT intervention [37]; Nintendo Switch (n=1) [44]; and iPad (n=2) [50,57] for sensorimotor, visuomotor [54], and CT [50], respectively.

In addition, 13 studies used a range of computerized programs for CT targeting functions such as attention, memory, and executive function. These included AKL-T01 (n=1) [62], Brain Fitness (n=1) [16], COGPACK (n=2) [53,68], CogniPlus (n=1) [32], NeuronUP (n=1) [60], Neuropeak (n=1) [54], and RehaCom (n=3) [42,47,63] and 2 studies with no specified programs [44,49]. In addition, 1 study used the FitForAll program [16] for PT. Studies that incorporated CCT alongside other modalities showed significantly better improvements, particularly in global cognition, than studies with either active and inactive control groups. Details on tasks, duration, and major findings of these programs are presented in Multimedia Appendix 3.

Five studies incorporated tDCS with CT [44,47,53,60,62], while 2 studies combined tDCS with PT [45,67]. Specifically, they applied 1.5mA to 2 mA of tDCS to the left dorsolateral prefrontal cortex (DLPFC) or 1.5 mA of transcranial alternating current stimulation (tACS) on the prefrontal cortex [62].

Acceptability and User Experience

User compliance is influenced by the acceptability and experience of treatments. Four studies indicated high acceptability and positive user experiences regarding multimodal interventions. Jesus et al [61] reported that 24 out of 27 participants demonstrated a 97% acceptability rate for the treatment effect and frequency of a combined CT and PI. Another study on a combined PT, CT, and PI intervention reported that 78% (39/50) of participants rated the program as “very good,” 20% (10/50) rated it as “good,” and 2% (1/50) rated it as “pleasurable” [68]. In addition, a study on a combined PT, CT, and SA program showed that 86% (35/41) of

XSL•FO
RenderX

participants were satisfied with the intervention duration and 81% (33/41 participants) expressed a desire to continue [55].

Similarly, a study on a combined PT and CT intervention reported that 39 out of 44 (89%) participants perceived subjective benefits. In terms of overall satisfaction, 14 out of 44 (32%) participants rated it as "very good," 24 out of 44 (54%) participants rated it as "good," and 6 out of 44 (13%) participants rated it as "neither good nor bad." In addition,12 out of 44 (27%) participants considered the user-friendliness of the intervention as "very good" [57].

Dosage Effect and Health Economics

While treatment dosage may affect clinical outcomes, only 1 study evaluated the dosage effects of a quadrimodal intervention (combined PT, CT, PI, and SA) [64]. It found that 1 to 2 sessions per week of the quadrimodal intervention in daycare centers yielded no significant difference in cognitive outcomes compared with 3 to 5 sessions per week [64]. As such, the study suggested that less-frequent sessions (1-2 wk) might be as effective as more frequent sessions, thus further studies are warranted to validate these findings and test with other multimodal interventions.

In addition, while the cost-effectiveness of multimodal cognitive interventions is important for clinical practice, none of the included studies examined the cost-effectiveness of the identified intervention.

Discussion

Overview

This scoping review examines the current research landscape on nonpharmacological multimodal interventions for MCI, highlighting publication trends, intervention types, technology use, user experience, and dosage effects. Over the past decade, the variety of multimodal intervention combinations has increased. The most common interventions combine PT and CT, with additional components, such as NI, electroacupuncture, PI, and current stimulation. Cognitive outcome measures are diverse, targeting various domains. Bimodal and trimodal interventions generally outperform single-modal ones in improving global cognition, attention, and executive function. Notably, PT with current stimulation and PT with NI often demonstrate better cognitive improvements compared to active controls, while CT with PI shows more improvements compared to inactive controls. However, most studies on combined PT and CT report mixed results. Quadrimodal interventions also show superior improvements, although their long-term effect remains uncertain. This review, following the Arksey and O'Malley framework, did not conduct a meta-analysis or risk of bias assessments. The included studies suggested that trimodal and quadrimodal interventions, especially those including SA, might offer better cognitive outcomes, but findings should be interpreted cautiously due to the limited number of studies. Technologies, such as VR, gaming, computerized programs, and transcranial stimulation, are increasingly adopted in MCI interventions. Although technology-assisted training showed significant improvements in various cognitive domains comparable to traditional trainings, results should be interpreted with care given the limited studies. Future research should investigate whether technology-assisted training, with or without additional modalities, offers significantly better cognitive improvements than traditional training. Although only 4 included studies assessed the acceptability of multimodal interventions, they were generally well-received by users. Only 1 included study evaluated the dosage effects of a quadrimodal intervention, suggesting that 3 or more sessions per week may not be beneficial, although this finding should be interpreted cautiously. As none of the included studies evaluated the cost-effectiveness of these interventions, future studies should address this gap.

Types of Modalities

Aside from traditional PT like aerobic and stretching exercises, recent research has been expanding to explore various PT types beyond traditional methods, including mind-body exercises such as tai chi [66], yoga [58], and dancing [15] as viable components in multimodal interventions for managing MCI.

Mind-body exercises, which engage the mind to influence bodily functions, are popular among older adults, partly due to cultural preferences in certain populations. The effectiveness of such exercises, such as tai chi, varies between Asian [74,79] and Western populations [39], highlighting cultural influences. tai chi, for instance, demands whole-body coordination, rhythmic movements, dynamic weight shifting, single-limb support, integrating movement recall (memorization and concentration), spatial orientation, and cognitive activities such as attention and executive control. This provides simultaneous PT and CT that could be beneficial for older adults with MCI [20,66]. In addition, these mind-body exercises offer relaxation and social support in group settings, potentially improving mood and motivation to participate in activities. Given their multifaceted benefits, mind-body physical activities may be incorporated into multimodal interventions for MCI.

With the growing interest in multimodal interventions, there is a noticeable increase in incorporating various interventions in addition to combined PT and CT. Modifiable lifestyle factors, including NI, are crucial for cognitive improvements among older adults with MCI. However, the effectiveness of NI as a standalone intervention has yielded mixed results for cognitive outcomes. Different dietary patterns and supplements have been suggested to promote cognitive health. For instance, the Mediterranean-Dietary Approaches to Stop Hypertension Intervention for Neurodegenerative Delay diet is popular. However, a large-scale study found that the Mediterranean-Dietary Approaches to Stop Hypertension Intervention for Neurodegenerative Delay diet was not significantly superior to a control diet with mild caloric restriction over 3 years [80]. In addition, systematic reviews suggest that certain single-nutrient supplements, such as folate, vitamin E, omega-3 fatty acids, and probiotics, show promising but preliminary results, often based on weak evidence or low-quality studies [81-83].

NI has drawn attention to its potential in multimodal interventions for MCI. All the included studies that incorporated NI demonstrated promising outcomes when combined with PT and CT compared to the control groups. However, some studies

XSL•FO
RenderX

lacked detailed descriptions of their use of dietary guidelines, limiting the understanding of whether the specific dietary changes could be contributed to cognitive improvement. The current findings suggest that while a single NI may offer limited results, a combination of different modalities can provide a synergistic protective effect, highlighting the effectiveness of a multimodal approach over a single intervention.

In addition, there is an increasing use of PI and SA as supplementary components. Previous studies suggested that engaging in SA could help delay the development of dementia and cognitive decline [84,85]. SA and leisure activities not only have therapeutic benefits but also allow participants to engage with peers, fostering long-term adherence to interventions. It is important to note that the number of studies that incorporated the use of PI and SA is still relatively small; more studies are warranted to strengthen whether the implementation of additional PI and SA might provide additional benefits in a multimodal intervention for MCI.

Collectively, all the aforementioned intervention components play a role in enhancing mental health and cognitive functions across various aspects of MCI, suggesting that a larger number of combined modalities could be beneficial. This aligns with a previous systematic review that suggested an increased number of modalities might yield greater improvement in cognition [21]. Indeed, most of the included studies incorporating 3 or 4 modalities reported notable improvements in various cognitive domains, especially when additional modalities (including SA, NI, and PI) were combined with PT and CT. However, there is no consensus regarding the optimal number or combinations of modalities for multimodal interventions. Furthermore, it is worth noting that adopting multiple lifestyle changes may pose challenges for older adults. As Schneider and Yvon [86] suggested, such changes are more feasible for older adults in better health and with higher education levels. In current MCI management, while health advice is important, it may not sufficiently inform or motivate individuals with MCI to adopt new management strategies. Therefore, health education emerges as a crucial component for the successful implementation of lifestyle changes because it provides vital information on lifestyle modifications and guides the execution of physical or cognitive exercises at home.

Nonetheless, it is noteworthy that most of the included studies only compared the effectiveness of 3 or more modal combinations with single-modal interventions, health education, or passive control groups. It remains unclear whether there are more modalities or specific combinations that would offer greater benefit to individuals with MCI.

The Current Use of Technology

Beyond traditional interventions, technology is increasingly embraced by researchers for its potential as a valuable asset in delivering multimodal interventions for clinicians and researchers from 2015 (n=1) [16] to 2023 (n=4) [47,54,62,63]. CCT comprises guided drill-and-practice on standardized tasks aimed at various cognitive domains that gradually progress with adjustable difficulty based on user performance. This approach is both cost-effective and safe for widespread use [87].

This review identified several computer programs that provide personalized CCT targeting various cognitive domains. Our findings concur with a previous systematic review, showing CCT and VR-CT as effective in enhancing attention, executive function, global cognition, memory, processing speed, verbal fluency, and visuospatial ability in older adults at risk of cognitive decline [25]. These outcomes are on par with traditional in-person PT and CT, suggesting CCT and VR-CT as promising ways for broad-scale cognitive intervention delivery.

Interactive video game–like programs offer training across multiple cognitive domains, including attention, executive function, processing speed, and visuomotor and visuospatial abilities, providing more engaging and motivating experiences for older adults with MCI than traditional face-to-face CT [26,88-90]. These programs also facilitate PT through diverse hardware. For instance, one included study used Nintendo Wii, Wii Remote, and Wii Balance Board for physical exercises (eg, aerobic exercises) [16]. Advanced technologies such as the latest Nintendo Switch, which tailored physical exercise games and hardware (eg, Ring Fit Adventure), adapt challenges to individual capacities, progressively improving cognitive function in older adults with MCI.

In addition to the computerized training program, the included studies that incorporated VR technology in delivering combined PT and CT demonstrated significant improvement in executive function and global cognition. These improvements were greater and consistent with the improvements from the combined PT and CT control group without the use of VR [36,37]. This positions VR as a promising tool for integrated therapy. As suggested by recent systematic reviews, VR could act as an assistive device to deliver training and help improve executive function and global cognition in individuals with MCI [25,88]. Compared to traditional PT and CT, VR-PT and VR-CT offer accessibility, cost-effectiveness, and immersive personalized experience to facilitate skill transfer to activities of daily living [36,37,90]. With advancements in technology, VR equipment has become more user-friendly, affordable, and suitable for home use and rehabilitation, benefiting older adults with mobility issues. However, potential side effects such as stimulator sickness, discomfort, tiredness, mood change induced by immersive experience, and some health-related issues from wearing the VR headgear [90] highlight the need for clinician oversight to adjust treatment as necessary.

Although the development of computerized training programs can be costly, the potential for widespread deployment can offset the expenses of hiring numerous qualified assessors or clinicians and the costs associated with traveling for face-to-face training. Future research should explore integrating these programs with multimodal interventions, such as NI, to enhance their effectiveness and improve the overall well-being of individuals with MCI.

MCI affects multiple cognitive functions. The DLPFC plays a crucial role in functional connectivity with other brain regions and controls various cognitive functions, including attention, decision-making, planning, and working memory. Research has found that individuals with MCI exhibit functional disconnection

XSL•FO
RenderX

in the left DLPFC, leading to attention and working memory deficits [60]. This review included studies that used noninvasive brain stimulation techniques (eg, tDCS [45,53,60] targeting the left DLPFC and tACS targeting the prefrontal cortex); results from these studies suggested that using tDCS or tACS alongside PT or CT could enhance the effectiveness of PT or CT by improving cognitive capacity and reducing the load needed to perform cognitive tasks, thereby increasing processing speed and potentially serving as a viable component in future studies.

Taken together, technology stands as a valuable adjunct in administrating PT and CT, offering standardized, cost-effective, and assessable personalized experiences. Such technological integration could help researchers develop comprehensive management strategies for the widespread treatment of MCI in older adults.

Acceptability and User Experience

Acceptability and user experience are pivotal for the success of interventions aimed at managing MCI, which directly influence its adherence rates. Adherence, in turn, significantly affects the effectiveness of multimodal interventions. Therefore, the interplay between user experience, acceptability, and adherence is vital in developing and implementing MCI interventions. However, most included studies have not investigated these factors, which may hinder clinicians or senior management in choosing or recommending treatment options. One included study suggested that the low adherence rate in older adults might be attributed to family responsibilities, health issues, and social commitments. Leveraging technology, particularly eHealth and mHealth tools, could improve adherence by alleviating barriers to adherence [65]. These tools could be used to facilitate easier implementation and monitoring of interventions, thereby reducing the time and cost required for broader deployment across various populations and regions [91]. Although none of the included studies have implemented such technological applications, future research should explore the potential of mobile technology in enhancing adherence in older adults with MCI.

Limitations and Future Directions

While this scoping review offers a comprehensive overview of current study directions in multimodal interventions for MCI management, it has several limitations. First, by focusing solely on experimental studies with control groups, it may have overlooked relevant insights from qualitative research and gray literature. Second, this scoping review only included English articles, which may introduce cultural bias. Third, this review concentrated exclusively on the cognitive outcomes, omitting physical and psychological outcomes that might be crucial considerations for policy or clinical decisions. Fourth, several included studies had relatively small sample sizes, ranging from 19 to 27 participants [16,43,45,62,65]. Therefore, results should be interpreted with caution. Fifth, adhering to the Arksey and O'Malley framework and PRISMA-ScR guidelines meant no risk of bias assessments or evidence syntheses were conducted. Future systematic reviews or network meta-analyses should aim to summarize evidence regarding the relative effectiveness of different combinations of various multimodal interventions for MCI management.

While the current American Academy of Neurology guidelines for MCI management only recommend PT or CT, recent studies highlight the potential benefits of NI, PI, and SI for MCI management. With the annual conversion rate of MCI to dementia ranging from 10% to 15% per year [7], further research should explore various multimodal intervention combinations, focusing on follow-up, treatment duration, and frequency to potentially reduce or delay conversion rates. Although some included studies used brain imaging, most relied on self-reported or assessment scores, which might not accurately reflect real-world improvements. Future research should incorporate brain imaging technologies to validate results more robustly and elucidate the mechanisms underlying clinical improvements in cognitive function.

Considering the limited research on the cost-effectiveness of multimodal interventions for older adults with MCI and the growing aging population, future studies should investigate the economic viability of these interventions. In addition, the unclear dosage effects of different multimodal intervention combinations warrant investigation. Only one included study revealed that attending 1 or 2 sessions per week of a quadrimodal intervention was not significantly different from 3 to 5 sessions per week, suggesting that fewer sessions of quadrimodal intervention (PT, CT, PI, and SA) might still significantly improve global cognition [64]. However, this finding from a single study is not conclusive, further research is needed to determine the optimal dosage and combination of modalities to optimize the effectiveness of multimodal interventions for MCI.

Previous studies highlight the potential of lifestyle modifications such as dietary patterns in managing cognitive decline in older adults [92-94]. However, these studies used diverse nutrition supplements or dietary patterns. Similarly, while social engagement could help manage MCI [5], there is a dearth of studies on its effectiveness in enhancing physical and emotional well-being in older adults with MCI.

Finally, the included studies used diverse diagnostic criteria and cognitive measurements with different cutoff scores, potentially resulting in heterogeneous participant cohorts that could affect the results. An international consortium should be formed to develop standardized outcome measures for consistent outcome comparisons.

Conclusions

This scoping review provides a comprehensive update on the use of multimodal interventions for improving cognitive functions in older adults with MCI. It presents study directions, multimodal intervention types, respective findings in the included studies, the role of technology in these interventions, and potential research directions. The most common multimodal intervention combines PT and CT. However, various types of interventions such as NI, PI, SI, and brain stimulation have also been incorporated into multimodal interventions more recently. Given the capacity of these interventions to stimulate multiple cognitive domains, the effectiveness of various combinations of modalities should be explored. The research gaps highlighted in this review pave the path for future large-scale clinical trials to help develop more effective management strategies for cognitive decline in older adults with MCI.

XSL•FO
RenderX

Acknowledgments

The authors would like to thank the authors and participants of the included and referenced studies in this scoping review. In addition, they would like to thank the Health and Social Science faculty librarian at the Hong Kong Polytechnic University for her help with our search terms. This scoping review was supported by funding from the Research Institute for Smart Ageing at The Hong Kong Polytechnic University (grant 1-CD4Z). David Ho-Keung Shum is supported by the Yeung Tsang Wing Yee and Tsang Wing Hing Endowed Professorship.

Authors' Contributions

Conceptualization: RCFC (lead), JHSZ (equal), AYLW (equal), YC (supporting), KL (supporting), DHKS (supporting).
Data Curation: RCFC (lead), JHSZ (equal).
Funding Acquisition: DHKS (lead), AYLW (equal).
Methodology: RCFC (lead), JHSZ (equal).
Supervision: AYLW (lead), YC (supporting), KL (supporting), PHFN (supporting), DHKS (supporting).
Visualization: PHFN.
Writing – original draft: RCFC (lead), JHSZ (equal).
Writing – review & editing: AYLC (lead), YC (supporting), KL (supporting), PHFC (supporting), DHKS (supporting).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search keywords and terms used in the current scoping review and results from each database.
[DOCX File , 19 KB - aging_v8i1e70291_app1.docx]

Multimedia Appendix 2
Detailed table of the included studies' and participants' demographic characteristics.
[DOCX File , 76 KB - aging_v8i1e70291_app2.docx]

Multimedia Appendix 3
Table of computerized training programs used in the included studies.
[DOCX File , 25 KB - aging_v8i1e70291_app3.docx]

Multimedia Appendix 4
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Review) Checklist.
[DOCX File , 27 KB - aging_v8i1e70291_app4.docx]

References

1. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011 May;7(3):270-279 [FREE Full text] [doi: 10.1016/j.jalz.2011.03.008] [Medline: 21514249]
2. Mrakic-Sposta S, Di Santo SG, Franchini F, Arlati S, Zangiacomi A, Greci L, et al. Effects of combined physical and cognitive virtual reality-based training on cognitive impairment and oxidative stress in MCI patients: a pilot study. Front Aging Neurosci 2018;10:282 [FREE Full text] [doi: 10.3389/fnagi.2018.00282] [Medline: 30327596]
3. Winblad B, Palmer K, Kivipelto M, Jelic V, Fratiglioni L, Wahlund L, et al. Mild cognitive impairment--beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. J Intern Med 2004 Sep;256(3):240-246 [FREE Full text] [doi: 10.1111/j.1365-2796.2004.01380.x] [Medline: 15324367]
4. Petersen RC. Mild cognitive impairment. Continuum (Minneap Minn) 2016 Apr;22(2 Dementia):404-418 [FREE Full text] [doi: 10.1212/CON.0000000000000313] [Medline: 27042901]
5. Lydon EA, Nguyen LT, Nie Q, Rogers WA, Mudar RA. An integrative framework to guide social engagement interventions and technology design for persons with mild cognitive impairment. Front Public Health 2022;9:750340 [FREE Full text] [doi: 10.3389/fpubh.2021.750340] [Medline: 35096730]
6. Sachdev PS, Lipnicki DM, Kochan NA, Crawford JD, Thalamuthu A, Andrews G, Cohort Studies of Memory in an International Consortium (COSMIC). The prevalence of mild cognitive impairment in diverse geographical and ethnocultural

XSL•FO
RenderX

regions: the COSMIC collaboration. PLoS One 2015;10(11):e0142388 [FREE Full text] [doi: 10.1371/journal.pone.0142388] [Medline: 26539987]
7. Pal K, Mukadam N, Petersen I, Cooper C. Mild cognitive impairment and progression to dementia in people with diabetes, prediabetes and metabolic syndrome: a systematic review and meta-analysis. Soc Psychiatry Psychiatr Epidemiol 2018 Nov;53(11):1149-1160 [FREE Full text] [doi: 10.1007/s00127-018-1581-3] [Medline: 30182156]
8. Global action plan on the public health response to dementia 2017-2025. World Health Organization. 2017. URL: https://www.who.int/publications/i/item/global-action-plan-on-the-public-health-response-to-dementia-2017?2025 [accessed 2024-04-29]
9. Ageing and health. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2024-04-29]
10. He CY, Zhou Z, Kan MM, Chan DH, Wong AC, Mok KH, et al. Modifiable risk factors for mild cognitive impairment among cognitively normal community-dwelling older adults: a systematic review and meta-analysis. Ageing Res Rev 2024 Aug;99:102350. [doi: 10.1016/j.arr.2024.102350] [Medline: 38942197]
11. Petersen RC, Lopez O, Armstrong MJ, Getchius TS, Ganguli M, Gloss D, et al. Practice guideline update summary: mild cognitive impairment: report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. Neurology 2018 Jan 16;90(3):126-135 [FREE Full text] [doi: 10.1212/WNL.0000000000004826] [Medline: 29282327]
12. Vasenina EE, Levin OS, Sonin AG. Sovremennye tendentsii v épidemiologii dementsii i vedenii patsientov s kognitivnymi narusheniiami. Zh Nevrol Psikhiatr Im S S Korsakova 2017;117(6. Vyp. 2):87-95. [doi: 10.17116/jnevro20171176287-95] [Medline: 28980619]
13. Schmitter-Edgecombe M, Dyck DG. Cognitive rehabilitation multi-family group intervention for individuals with mild cognitive impairment and their care-partners. J Int Neuropsychol Soc 2014 Oct;20(9):897-908. [doi: 10.1017/S1355617714000782] [Medline: 25222630]
14. Biazus-Sehn LF, Schuch FB, Firth J, Stigger FD. Effects of physical exercise on cognitive function of older adults with mild cognitive impairment: a systematic review and meta-analysis. Arch Gerontol Geriatr 2020;89:104048. [doi: 10.1016/j.archger.2020.104048] [Medline: 32460123]
15. Parial LL, Kor PP, Sumile EF, Leung AY. Dual-task Zumba gold for improving the cognition of people with mild cognitive impairment: a pilot randomized controlled trial. Gerontologist 2023 Aug 24;63(7):1248-1261. [doi: 10.1093/geront/gnac081] [Medline: 35679826]
16. Styliadis C, Kartsidis P, Paraskevopoulos E, Ioannides AA, Bamidis PD. Neuroplastic effects of combined computerized physical and cognitive training in elderly individuals at risk for dementia: an eLORETA controlled study on resting states. Neural Plast 2015;2015:172192 [FREE Full text] [doi: 10.1155/2015/172192] [Medline: 25945260]
17. Cheng ST. Cognitive reserve and the prevention of dementia: the role of physical and cognitive activities. Curr Psychiatry Rep 2016 Sep 1;18(9):85 [FREE Full text] [doi: 10.1007/s11920-016-0721-2] [Medline: 27481112]
18. Simon SS, Yokomizo JE, Bottino CM. Cognitive intervention in amnestic mild cognitive impairment: a systematic review. Neurosci Biobehav Rev 2012 Apr;36(4):1163-1178. [doi: 10.1016/j.neubiorev.2012.01.007] [Medline: 22322184]
19. Karp A, Paillard-Borg S, Wang HX, Silverstein M, Winblad B, Fratiglioni L. Mental, physical and social components in leisure activities equally contribute to decrease dementia risk. Dement Geriatr Cogn Disord 2006;21(2):65-73. [doi: 10.1159/000089919] [Medline: 16319455]
20. Lissek V, Suchan B. Preventing dementia? Interventional approaches in mild cognitive impairment. Neurosci Biobehav Rev 2021 Mar;122:143-164. [doi: 10.1016/j.neubiorev.2020.12.022] [Medline: 33440197]
21. Salzman T, Sarquis-Adamson Y, Son S, Montero-Odasso M, Fraser S. Associations of multidomain interventions with improvements in cognition in mild cognitive impairment: a systematic review and meta-analysis. JAMA Netw Open 2022 May 02;5(5):e226744 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.6744] [Medline: 35503222]
22. Yang C, Moore A, Mpofu E, Dorstyn D, Li Q, Yin C. Effectiveness of combined cognitive and physical interventions to enhance functioning in older adults with mild cognitive impairment: a systematic review of randomized controlled trials. Gerontologist 2020 Nov 23;60(8):633-642. [doi: 10.1093/geront/gnz149] [Medline: 31697831]
23. Sohn M, Yang J, Sohn J, Lee JH. Digital healthcare for dementia and cognitive impairment: a scoping review. Int J Nurs Stud 2023 Apr;140:104413. [doi: 10.1016/j.ijnurstu.2022.104413] [Medline: 36821951]
24. Di Lorito C, Bosco A, Rai H, Craven M, McNally D, Todd C, et al. A systematic literature review and meta-analysis on digital health interventions for people living with dementia and mild cognitive impairment. Int J Geriatr Psychiatry 2022 Jun;37(6):5730 [FREE Full text] [doi: 10.1002/gps.5730] [Medline: 35588315]
25. Coyle H, Traynor V, Solowij N. Computerized and virtual reality cognitive training for individuals at high risk of cognitive decline: systematic review of the literature. Am J Geriatr Psychiatry 2015 Apr;23(4):335-359. [doi: 10.1016/j.jagp.2014.04.009] [Medline: 24998488]
26. Zhang H, Huntley J, Bhome R, Holmes B, Cahill J, Gould RL, et al. Effect of computerised cognitive training on cognitive outcomes in mild cognitive impairment: a systematic review and meta-analysis. BMJ Open 2019 Aug 18;9(8):e027062 [FREE Full text] [doi: 10.1136/bmjopen-2018-027062] [Medline: 31427316]

XSL•FO
RenderX

27. Shao Z, Hu M, Zhang D, Zeng X, Shu X, Wu X, et al. Dose-response relationship in non-pharmacological interventions for individuals with mild cognitive impairment: a systematic review and meta-analysis of randomised controlled trials. J Clin Nurs 2022 Dec;31(23-24):3390-3401. [doi: 10.1111/jocn.16240] [Medline: 35098610]
28. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
29. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
30. A scoping review of non-pharmacological multi-modal interventions on multi-domain functions in older adults with mild cognitive impairment. Open Science Framework. URL: https://osf.io/y4hgk/ [accessed 2024-05-29]
31. Tricco AC, Lillie E, Zarin W, O'Brien K, Colquhoun H, Kastner M, et al. A scoping review on the conduct and reporting of scoping reviews. BMC Med Res Methodol 2016 Feb 09;16:15 [FREE Full text] [doi: 10.1186/s12874-016-0116-4] [Medline: 26857112]
32. Hagovska M, Nagyova I. The transfer of skills from cognitive and physical training to activities of daily living: a randomised controlled study. Eur J Ageing 2017 Jun;14(2):133-142 [FREE Full text] [doi: 10.1007/s10433-016-0395-y] [Medline: 28804398]
33. Hagovská M, Olekszyová Z. Relationships between balance control and cognitive functions, gait speed, and activities of daily living. Z Gerontol Geriatr 2016 Jul;49(5):379-385. [doi: 10.1007/s00391-015-0955-3] [Medline: 26458911]
34. Hagovská M, Olekszyová Z. Impact of the combination of cognitive and balance training on gait, fear and risk of falling and quality of life in seniors with mild cognitive impairment. Geriatr Gerontol Int 2016 Sep;16(9):1043-1050. [doi: 10.1111/ggi.12593] [Medline: 26338465]
35. Hagovská M, Takáč P, Dzvoník O. Effect of a combining cognitive and balanced training on the cognitive, postural and functional status of seniors with a mild cognitive deficit in a randomized, controlled trial. Eur J Phys Rehabil Med 2016 Feb;52(1):101-109 [FREE Full text] [Medline: 26325026]
36. Liao YY, Tseng HY, Lin YJ, Wang CJ, Hsu WC. Using virtual reality-based training to improve cognitive function, instrumental activities of daily living and neural efficiency in older adults with mild cognitive impairment. Eur J Phys Rehabil Med 2020 Feb;56(1):47-57 [FREE Full text] [doi: 10.23736/S1973-9087.19.05899-4] [Medline: 31615196]
37. Liao YY, Chen IH, Lin YJ, Chen Y, Hsu WC. Effects of virtual reality-based physical and cognitive training on executive function and dual-task gait performance in older adults with mild cognitive impairment: a randomized control trial. Front Aging Neurosci 2019 Jul 16;11:162 [FREE Full text] [doi: 10.3389/fnagi.2019.00162] [Medline: 31379553]
38. Fiatarone Singh MA, Gates N, Saigal N, Wilson GC, Meiklejohn J, Brodaty H, et al. The Study of Mental and Resistance Training (SMART) study—resistance training and/or cognitive training in mild cognitive impairment: a randomized, double-blind, double-sham controlled trial. J Am Med Dir Assoc 2014 Dec;15(12):873-880. [doi: 10.1016/j.jamda.2014.09.010] [Medline: 25444575]
39. Fogarty JN, Murphy KJ, McFarlane B, Montero-Odasso M, Wells J, Troyer AK, et al. Taoist Tai Chi® and memory intervention for individuals with mild cognitive impairment. J Aging Phys Act 2016 Apr;24(2):169-180 [FREE Full text] [doi: 10.1123/japa.2014-0062] [Medline: 25838271]
40. Jeong MK, Park KW, Ryu JK, Kim GM, Jung HH, Park H. Multi-component intervention program on habitual physical activity parameters and cognitive function in patients with mild cognitive impairment: a randomized controlled trial. Int J Environ Res Public Health 2021 Jun 09;18(12):6240 [FREE Full text] [doi: 10.3390/ijerph18126240] [Medline: 34207701]
41. Kamegaya T, Araki Y, Kigure H, Long-Term-Care Prevention Team of Maebashi City, Yamaguchi H. Twelve-week physical and leisure activity programme improved cognitive function in community-dwelling elderly subjects: a randomized controlled trial. Psychogeriatrics 2014 Mar 16;14(1):47-54. [doi: 10.1111/psyg.12038] [Medline: 24528600]
42. Kim JH, Han JY, Park GC, Lee JS. Cognitive improvement effects of electroacupuncture combined with computer-based cognitive rehabilitation in patients with mild cognitive impairment: a randomized controlled trial. Brain Sci 2020 Dec 14;10(12):984 [FREE Full text] [doi: 10.3390/brainsci10120984] [Medline: 33327627]
43. Köbe T, Witte AV, Schnelle A, Lesemann A, Fabian S, Tesky VA, et al. Combined omega-3 fatty acids, aerobic exercise and cognitive stimulation prevents decline in gray matter volume of the frontal, parietal and cingulate cortex in patients with mild cognitive impairment. Neuroimage 2016 May 01;131:226-238. [doi: 10.1016/j.neuroimage.2015.09.050] [Medline: 26433119]
44. Lau CI, Liu MN, Cheng FY, Wang HC, Walsh V, Liao YY. Can transcranial direct current stimulation combined with interactive computerized cognitive training boost cognition and gait performance in older adults with mild cognitive impairment?: a randomized controlled trial. J Neuroeng Rehabil 2024 Feb 16;21(1):26 [FREE Full text] [doi: 10.1186/s12984-024-01313-0] [Medline: 38365761]
45. Liao YY, Liu MN, Wang HC, Walsh V, Lau CI. Combining transcranial direct current stimulation with tai chi to improve dual-task gait performance in older adults with mild cognitive impairment: a randomized controlled trial. Front Aging Neurosci 2021;13:766649 [FREE Full text] [doi: 10.3389/fnagi.2021.766649] [Medline: 34966268]

XSL•FO
RenderX

46. Phoemsapthawee J, Ammawat W, Prasertsri P, Sathalalai P, Leelayuwat N. Does Gotu kola supplementation improve cognitive function, inflammation, and oxidative stress more than multicomponent exercise alone?: a randomized controlled study. J Exerc Rehabil 2022 Oct;18(5):330-342 [FREE Full text] [doi: 10.12965/jer.2244388.194] [Medline: 36420467]
47. Senczyszyn A, Szcześniak D, Wieczorek T, Maciaszek J, Małecka M, Bogudzińska B, et al. Improvement of working memory in older adults with mild cognitive impairment after repetitive transcranial magnetic stimulation - a randomized controlled pilot study. Front Psychiatry 2023;14:1196478 [FREE Full text] [doi: 10.3389/fpsyt.2023.1196478] [Medline: 38111617]
48. Shimizu N, Umemura T, Matsunaga M, Hirai T. Effects of movement music therapy with a percussion instrument on physical and frontal lobe function in older adults with mild cognitive impairment: a randomized controlled trial. Aging Ment Health 2018 Dec;22(12):1614-1626 [FREE Full text] [doi: 10.1080/13607863.2017.1379048] [Medline: 28937272]
49. Yang QH, Lyu X, Lin QR, Wang ZW, Tang L, Zhao Y, et al. Effects of a multicomponent intervention to slow mild cognitive impairment progression: a randomized controlled trial. Int J Nurs Stud 2022 Jan;125:104110. [doi: 10.1016/j.ijnurstu.2021.104110] [Medline: 34736073]
50. Bray NW, Pieruccini-Faria F, Witt ST, Bartha R, Doherty TJ, Nagamatsu LS, et al. Combining exercise with cognitive training and vitamin D 3 to improve functional brain connectivity (FBC) in older adults with mild cognitive impairment (MCI). Results from the SYNERGIC trial. Geroscience 2023 Jun;45(3):1967-1985 [FREE Full text] [doi: 10.1007/s11357-023-00805-6] [Medline: 37162700]
51. Fairchild JK, Myers J, Louras P, Jo B, McNerney MW, Hallmayer J, et al. Multimodal exercise and cognitive training program improves cognitive function in amnestic mild cognitive impairment. Am J Geriatr Psychiatry 2024 Apr;32(4):463-474. [doi: 10.1016/j.jagp.2023.12.002] [Medline: 38220592]
52. Griffiths J, Thaikruea L, Wongpakaran N, Munkhetvit P, Kittisares A, Varnado P. Effects of combined physical movement activity and multifaceted cognitive training in older people with mild neurocognitive disorder in a rural community: a randomized control trial. Dement Geriatr Cogn Disord 2020;49(2):194-201. [doi: 10.1159/000507922] [Medline: 32535601]
53. Martin DM, Mohan A, Alonzo A, Gates N, Gbadeyan O, Meinzer M, et al. A pilot double-blind randomized controlled trial of cognitive training combined with transcranial direct current stimulation for amnestic mild cognitive impairment. J Alzheimers Dis 2019 Aug 12;71(2):503-512. [doi: 10.3233/JAD-190306] [Medline: 31424410]
54. Montero-Odasso M, Zou G, Speechley M, Almeida QJ, Liu-Ambrose T, Middleton LE, Canadian GaitCognition Network. Effects of exercise alone or combined with cognitive training and vitamin D supplementation to improve cognition in adults with mild cognitive impairment: a randomized clinical trial. JAMA Netw Open 2023 Jul 03;6(7):e2324465 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.24465] [Medline: 37471089]
55. Bae S, Lee S, Lee S, Jung S, Makino K, Harada K, et al. The effect of a multicomponent intervention to promote community activity on cognitive function in older adults with mild cognitive impairment: a randomized controlled trial. Complement Ther Med 2019 Feb;42:164-169. [doi: 10.1016/j.ctim.2018.11.011] [Medline: 30670238]
56. Blumenthal JA, Smith PJ, Mabe S, Hinderliter A, Lin P, Liao L, et al. Lifestyle and neurocognition in older adults with cognitive impairments: a randomized trial. Neurology 2019 Jan 15;92(3):e212-e223 [FREE Full text] [doi: 10.1212/WNL.0000000000006784] [Medline: 30568005]
57. Callisaya ML, Jayakody O, Vaidya A, Srikanth V, Farrow M, Delbaere K. A novel cognitive-motor exercise program delivered via a tablet to improve mobility in older people with cognitive impairment - StandingTall cognition and mobility. Exp Gerontol 2021 Sep;152:111434. [doi: 10.1016/j.exger.2021.111434] [Medline: 34098009]
58. Chobe S, Patra SK, Chobe M, Metri K. Efficacy of integrated yoga and ayurveda rasayana on cognitive functions in elderly with mild cognitive impairment: non-randomized three-arm clinical trial. J Ayurveda Integr Med 2022;13(1):100373 [FREE Full text] [doi: 10.1016/j.jaim.2020.11.003] [Medline: 33349568]
59. Gill DP, Gregory MA, Zou G, Liu-Ambrose T, Shigematsu R, Hachinski V, et al. The healthy mind, healthy mobility trial: a novel exercise program for older adults. Med Sci Sports Exerc 2016 Feb;48(2):297-306. [doi: 10.1249/MSS.0000000000000758] [Medline: 26285025]
60. Gonzalez PC, Fong KN, Brown T. Transcranial direct current stimulation as an adjunct to cognitive training for older adults with mild cognitive impairment: a randomized controlled trial. Ann Phys Rehabil Med 2021 Sep;64(5):101536 [FREE Full text] [doi: 10.1016/j.rehab.2021.101536] [Medline: 33957292]
61. Jesus AG, de Lima MP, Vilar M, Pachana NA. Feasibility, acceptability, and preliminary efficacy of a cognitive-emotional intervention program in homebound older adults: a pilot randomized controlled trial. Clin Gerontol 2023;46(3):359-375. [doi: 10.1080/07317115.2022.2105670] [Medline: 35916398]
62. Jones KT, Ostrand AE, Gazzaley A, Zanto TP. Enhancing cognitive control in amnestic mild cognitive impairment via at-home non-invasive neuromodulation in a randomized trial. Sci Rep 2023 May 08;13(1):7435 [FREE Full text] [doi: 10.1038/s41598-023-34582-1] [Medline: 37156876]
63. Kim JH, Park JH. Does cognitive-physical dual-task training have better clinical outcomes than cognitive single-task training does? A single-blind, randomized controlled trial. Healthcare (Basel) 2023 May 25;11(11):1544 [FREE Full text] [doi: 10.3390/healthcare11111544] [Medline: 37297684]

XSL•FO
RenderX

64. Straubmeier M, Behrndt EM, Seidl H, Özbe D, Luttenberger K, Graessel E. Non-pharmacological treatment in people with cognitive impairment. Dtsch Arztebl Int 2017 Dec 01;114(48):815-821 [FREE Full text] [doi: 10.3238/arztebl.2017.0815] [Medline: 29249224]
65. Xu Z, Zhang D, Lee AT, Sit RW, Wong C, Lee EK, et al. A pilot feasibility randomized controlled trial on combining mind-body physical exercise, cognitive training, and nurse-led risk factor modification to reduce cognitive decline among older adults with mild cognitive impairment in primary care. PeerJ 2020;8:e9845 [FREE Full text] [doi: 10.7717/peerj.9845] [Medline: 33194354]
66. Li B, Tang H, He G, Jin Z, He Y, Huang P, et al. Tai Chi enhances cognitive training effects on delaying cognitive decline in mild cognitive impairment. Alzheimers Dement 2023 Jan 15;19(1):136-149. [doi: 10.1002/alz.12658] [Medline: 35290704]
67. Xu Y, Zhu J, Liu H, Qiu Z, Wu M, Liu J, et al. Effects of Tai Chi combined with tDCS on cognitive function in patients with MCI: a randomized controlled trial. Front Public Health 2023;11:1199246 [FREE Full text] [doi: 10.3389/fpubh.2023.1199246] [Medline: 37608981]
68. Stuerz K, Hartmann S, Holzner B, Bichler CS, Niedermeier M, Kopp M, et al. Effects of a two-step cognitive and relaxation training program in care home residents with mild cognitive impairment. Int J Environ Res Public Health 2022 Jul 07;19(14):8316 [FREE Full text] [doi: 10.3390/ijerph19148316] [Medline: 35886176]
69. Jesus AG, Lima MP, Vilar M, Pachana NA. HEPPI: a randomized controlled trial of the efficacy of a cognitive-emotional intervention for homebound older adults. Aging Ment Health 2024;28(3):491-501. [doi: 10.1080/13607863.2023.2260760] [Medline: 37747057]
70. Rezola-Pardo C, Arrieta H, Gil SM, Zarrazquin I, Yanguas JJ, López MA, et al. Comparison between multicomponent and simultaneous dual-task exercise interventions in long-term nursing home residents: the Ageing-ONDUAL-TASK randomized controlled study. Age Ageing 2019 Nov 01;48(6):817-823. [doi: 10.1093/ageing/afz105] [Medline: 31595289]
71. Shimada H, Makizako H, Doi T, Park H, Tsutsumimoto K, Verghese J, et al. Effects of combined physical and cognitive exercises on cognition and mobility in patients with mild cognitive impairment: a randomized clinical trial. J Am Med Dir Assoc 2018 Jul;19(7):584-591. [doi: 10.1016/j.jamda.2017.09.019] [Medline: 29153754]
72. Park H, Park JH, Na HR, Hiroyuki S, Kim GM, Jung MK, et al. Combined intervention of physical activity, aerobic exercise, and cognitive exercise intervention to prevent cognitive decline for patients with mild cognitive impairment: a randomized controlled clinical study. J Clin Med 2019 Jun 28;8(7):940 [FREE Full text] [doi: 10.3390/jcm8070940] [Medline: 31261796]
73. Tao M, Liu H, Cheng J, Yu C, Zhao L. Motor-cognitive interventions may effectively improve cognitive function in older adults with mild cognitive impairment: a randomized controlled trial. Behav Sci (Basel) 2023 Sep 04;13(9):737 [FREE Full text] [doi: 10.3390/bs13090737] [Medline: 37754015]
74. Lam LC, Chan WC, Leung T, Fung AW, Leung EM. Would older adults with mild cognitive impairment adhere to and benefit from a structured lifestyle activity intervention to enhance cognition?: a cluster randomized controlled trial. PLoS One 2015;10(3):e0118173 [FREE Full text] [doi: 10.1371/journal.pone.0118173] [Medline: 25826620]
75. Lee S, Harada K, Bae S, Harada K, Makino K, Anan Y, et al. A non-pharmacological multidomain intervention of dual-task exercise and social activity affects the cognitive function in community-dwelling older adults with mild to moderate cognitive decline: a randomized controlled trial. Front Aging Neurosci 2023;15:1005410 [FREE Full text] [doi: 10.3389/fnagi.2023.1005410] [Medline: 36993908]
76. Lin YC, Wang CJ, Chang YF, Wang JJ. Effects of the biopsychosocial functional activity program on cognitive function for community older adults with mild cognitive impairment: a cluster-randomized controlled trial. Nurs Health Sci 2020 Dec;22(4):1065-1075. [doi: 10.1111/nhs.12772] [Medline: 32865857]
77. Train the Brain Consortium. Randomized trial on the effects of a combined physical/cognitive training in aged MCI subjects: the Train the Brain study. Sci Rep 2017 Jan 03;7:39471 [FREE Full text] [doi: 10.1038/srep39471] [Medline: 28045051]
78. Liu X, Ma Z, Zhu X, Zheng Z, Li J, Fu J, et al. Cognitive benefit of a multidomain intervention for older adults at risk of cognitive decline: a cluster-randomized controlled trial. Am J Geriatr Psychiatry 2023 Mar;31(3):197-209. [doi: 10.1016/j.jagp.2022.10.006] [Medline: 36414488]
79. Sungkarat S, Boripuntakul S, Kumfu S, Lord SR, Chattipakorn N. Tai Chi improves cognition and plasma BDNF in older adults with mild cognitive impairment: a randomized controlled trial. Neurorehabil Neural Repair 2018 Feb;32(2):142-149. [doi: 10.1177/1545968317753682] [Medline: 29353543]
80. Barnes LL, Dhana K, Liu X, Carey VJ, Ventrelle J, Johnson K, et al. Trial of the MIND diet for prevention of cognitive decline in older persons. N Engl J Med 2023 Aug 17;389(7):602-611 [FREE Full text] [doi: 10.1056/NEJMoa2302368] [Medline: 37466280]
81. Andrews V, Zammit G, O'Leary F. Dietary pattern, food, and nutritional supplement effects on cognitive outcomes in mild cognitive impairment: a systematic review of previous reviews. Nutr Rev 2023 Oct 10;81(11):1462-1489 [FREE Full text] [doi: 10.1093/nutrit/nuad013] [Medline: 37027832]
82. McGrattan AM, McEvoy CT, McGuinness B, McKinley MC, Woodside JV. Effect of dietary interventions in mild cognitive impairment: a systematic review. Br J Nutr 2018 Dec;120(12):1388-1405 [FREE Full text] [doi: 10.1017/S0007114518002945] [Medline: 30409231]
83. Vlachos GS, Scarmeas N. Dietary interventions in mild cognitive impairment and dementia. Dialogues Clin Neurosci 2019 Mar;21(1):69-82 [FREE Full text] [doi: 10.31887/DCNS.2019.21.1/nscarmeas] [Medline: 31607782]

XSL•FO
RenderX

84. Kuiper JS, Zuidersma M, Zuidema SU, Burgerhof JG, Stolk RP, Oude Voshaar RC, et al. Social relationships and cognitive decline: a systematic review and meta-analysis of longitudinal cohort studies. Int J Epidemiol 2016 Aug 06;45(4):1169-1206. [doi: 10.1093/ije/dyw089] [Medline: 27272181]
85. Wang HX, Karp A, Winblad B, Fratiglioni L. Late-life engagement in social and leisure activities is associated with a decreased risk of dementia: a longitudinal study from the Kungsholmen project. Am J Epidemiol 2002 Jun 15;155(12):1081-1087. [doi: 10.1093/aje/155.12.1081] [Medline: 12048221]
86. Schneider N, Yvon C. A review of multidomain interventions to support healthy cognitive ageing. J Nutr Health Aging 2013 Mar;17(3):252-257 [FREE Full text] [doi: 10.1007/s12603-012-0402-8] [Medline: 23459978]
87. Hill NT, Mowszowski L, Naismith SL, Chadwick VL, Valenzuela M, Lampit A. Computerized cognitive training in older adults with mild cognitive impairment or dementia: a systematic review and meta-analysis. Am J Psychiatry 2017 Apr 01;174(4):329-340. [doi: 10.1176/appi.ajp.2016.16030360] [Medline: 27838936]
88. Wu J, Ma Y, Ren Z. Rehabilitative effects of virtual reality technology for mild cognitive impairment: a systematic review with meta-analysis. Front Psychol 2020 Sep 25;11:1811 [FREE Full text] [doi: 10.3389/fpsyg.2020.01811] [Medline: 33101098]
89. Thapa N, Park HJ, Yang JG, Son H, Jang M, Lee J, et al. The effect of a virtual reality-based intervention program on cognition in older adults with mild cognitive impairment: a randomized control trial. J Clin Med 2020 Apr 29;9(5):1283 [FREE Full text] [doi: 10.3390/jcm9051283] [Medline: 32365533]
90. D'Cunha NM, Nguyen D, Naumovski N, McKune AJ, Kellett J, Georgousopoulou EN, et al. A mini-review of virtual reality-based interventions to promote well-being for people living with dementia and mild cognitive impairment. Gerontology 2019;65(4):430-440. [doi: 10.1159/000500040] [Medline: 31108489]
91. Rosenberg A, Mangialasche F, Ngandu T, Solomon A, Kivipelto M. Multidomain interventions to prevent cognitive impairment, Alzheimer's disease, and dementia: from FINGER to world-wide FINGERS. J Prev Alzheimers Dis 2020;7(1):29-36 [FREE Full text] [doi: 10.14283/jpad.2019.41] [Medline: 32010923]
92. Hugo J, Ganguli M. Dementia and cognitive impairment: epidemiology, diagnosis, and treatment. Clin Geriatr Med 2014 Aug;30(3):421-442 [FREE Full text] [doi: 10.1016/j.cger.2014.04.001] [Medline: 25037289]
93. Lu Y, An Y, Guo J, Zhang X, Wang H, Rong H, et al. Dietary intake of nutrients and lifestyle affect the risk of mild cognitive impairment in the Chinese elderly population: a cross-sectional study. Front Behav Neurosci 2016;10:229 [FREE Full text] [doi: 10.3389/fnbeh.2016.00229] [Medline: 27965552]
94. Smith PJ, Blumenthal JA. Dietary factors and cognitive decline. J Prev Alzheimers Dis 2016 Mar;3(1):53-64 [FREE Full text] [doi: 10.14283/jpad.2015.71] [Medline: 26900574]

Abbreviations

CCT: computerized cognitive training
CT: cognitive training
DLPFC: dorsolateral prefrontal cortex
MCI: mild cognitive impairment
NI: nutrition intervention
PI: psychosocial intervention
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
PT: physical training
SA: social activities
tACS: transcranial alternating current stimulation
tDCS: transcranial direct current stimulation
VR: virtual reality

Edited by H LaMonica; submitted 19.12.24; peer-reviewed by F Shojaei, D Tao, S Moolsart; comments to author 16.02.25; revised version received 07.04.25; accepted 21.04.25; published 12.05.25.

Please cite as:
Chan RCF, Zhou JHS, Cao Y, Lo K, Ng PHF, Shum DHK, Wong AYL
Nonpharmacological Multimodal Interventions for Cognitive Functions in Older Adults With Mild Cognitive Impairment: Scoping Review
JMIR Aging 2025;8:e70291
URL: https://aging.jmir.org/2025/1/e70291
doi:10.2196/70291
PMID:

XSL•FO
RenderX

©Raffy Chi-Fung Chan, Joson Hao-Shen Zhou, Yuan Cao, Kenneth Lo, Peter Hiu-Fung Ng, David Ho-Keung Shum, Arnold Yu-Lok Wong. Originally published in JMIR Aging (https://aging.jmir.org), 12.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Advancing Remote Monitoring for Patients With Alzheimer Disease and Related Dementias: Systematic Review

Mohmmad Arif Shaik[1], MS; Fahim Islam Anik[2], MS; Md. Mehedi Hasan[1], MS; Sumit Chakravarty[1], PhD; Mary Dioise Ramos[3], RN, PhD; Mohammad Ashiqur Rahman[4], PhD; Sheikh Iqbal Ahamed[5], PhD; Nazmus Sakib[1], PhD

[1]Department of Electrical and Computer Engineering, Kennesaw State University, Kennesaw, GA, United States

[2]Department of Mechanical Engineering, Khulna University of Engineering & Technology, Khulna, Bangladesh

[3]Louisiana State University Health Sciences Center (LSUHSC) New Orleans, School of Nursing, Louisiana State University, New Orleans, United States

[4]Department of Electrical and Computer Engineering, Florida International University, Miami, FL, United States

[5]Department of Computer Science, Marquette University, Milwaukee, WI, United States

Corresponding Author:
Mohmmad Arif Shaik, MS
Department of Electrical and Computer Engineering
Kennesaw State University
Kennesaw, GA
United States
Phone: 1 4147975981
Email: mshaik14@students.kennesaw.edu

Abstract

Background: Using remote monitoring technology in the context of Alzheimer disease (AD) care presents exciting new opportunities to lessen caregiver stress and improve patient care quality. The application of wearables, environmental sensors, and smart home systems designed specifically for patients with AD represents a promising interdisciplinary approach that integrates advanced technology with health care to enhance patient safety, monitor health parameters in real time, and provide comprehensive support to caregivers.

Objective: The objectives of this study included evaluating the effectiveness of various remote sensing technologies in enhancing patient outcomes and identifying strategies to alleviate the burden on health care professionals and caregivers. Critical elements such as regulatory compliance, user-centered design, privacy and security considerations, and the overall efficacy of relevant technologies were comprehensively examined. Ultimately, this study aimed to propose a comprehensive remote monitoring framework tailored to the needs of patients with AD and related dementias.

Methods: Guided by the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework, we conducted a systematic review on remote monitoring for patients with AD and related dementias. Our search spanned 4 major electronic databases—Google Scholar, PubMed, IEEE Xplore, and DBLP on February 20, 2024, with an updated search on May 18, 2024.

Results: A total of 31 publications met the inclusion criteria, highlighting 4 key research areas: existing remote monitoring technologies, balancing practicality and empathy, security and privacy in monitoring, and technology design for AD care. The studies revealed a strong focus on various remote monitoring methods for capturing behavioral, physiological, and environmental data yet showed a gap in evaluating these methods for patient and caregiver needs, privacy, and usability. The findings also indicated that many studies lacked robust reference standards and did not consistently apply critical appraisal criteria, underlining the need for comprehensive frameworks that better integrate these essential considerations.

Conclusions: This comprehensive literature review of remote monitoring technologies for patients with AD provides an understanding of remote monitoring technologies, trends, and gaps in the current research and the significance of novel strategies for remote monitoring to enhance patient outcomes and reduce the burden among health professionals and caregivers. The proposed remote monitoring framework aims to inspire the development of new interdisciplinary research models that advance care for patients with AD.

(JMIR Aging 2025;8:e69175) doi:10.2196/69175

XSL•FO
RenderX

KEYWORDS

dementia; Alzheimer disease; remote monitoring; Alzheimer; caregiver; fall detection; artificial intelligence

Introduction

Background

Alzheimer disease (AD) is the most generic form of dementia, and significantly affects the brain cells, declines cognitive abilities, and makes it difficult to perform daily tasks [1]. The main effects of the disease are loss of memory, inability to use problem-solving and logical thinking skills, anxiety, depression, and confusion [2]. AD has remained the leading cause of death among people aged >60 years. The main cause of AD is the accumulation of abnormal proteins such as β-amyloid and microtubule-associated tau protein in the brain; the aggregation of these proteins causes synaptic disintegration and neural loss, which characterize AD [3,4]. Disturbingly, it is estimated that the number of AD cases could grow to 13.8 million by 2060. The pervasive nature of this disease underscores the pressing need for continued research, intervention strategies, and support systems to address the growing challenges posed by AD in an aging population [5,6].

Caring for individuals with AD is an enduring and demanding journey, often placing a significant burden on family members who become primary caregivers or on those who opt to hire professional caregivers [7]. The responsibilities encompass a spectrum of care, requiring a deep commitment to providing the essential support needed for patients with AD to navigate their daily lives. Caregivers play a crucial role in assisting individuals with AD in tasks integral to their routine while also implementing lifestyle adjustments aimed at alleviating stress, confusion, anxiety, and agitation. The caregiving journey involves fostering an environment that promotes mental, physical, and social well-being. This includes implementing tailored strategies to keep patients with AD mentally engaged, encouraging physical activities that align with their abilities, and facilitating social interactions to prevent isolation. However, continuous monitoring and provision of care for patients with AD causes a substantial burden on both the family members and caregivers [8,9]. This responsibility often causes a lot of stress and depression among those providing care. Family members and caregivers find themselves dealing with the progressive loss of individuals with AD. In addition to that, they must navigate between caregiver responsibilities and sustaining a work-life balance. Disputes between family members, financial burdens, and escalating medical expenses extend the challenges that those undertaking caregiver roles face [10]. While existing research has identified various scales to measure the burden on family members, there is a notable gap in the literature concerning effective strategies to alleviate this burden. Manzini and do Vale [11] and Liao et al [12] address this critical gap by exploring remote monitoring methods facilitated by ITs. By leveraging innovative approaches, the research seeks to enhance the quality of care provided to patients with AD and substantially reduce the burden experienced by family members and caregivers. Through the implementation of such technological solutions, the aforementioned studies [11,12] aspired to contribute valuable insights into improving the overall well-being of both patients with AD and those entrusted with their care.

Despite the growing body of research on remote monitoring for AD and related dementias (ADRD), significant gaps remain in the literature, particularly regarding interdisciplinary integration. Previous reviews have primarily focused on individual aspects of remote monitoring, such as wearable devices [13-16], smart home sensors [17-19], or privacy and security considerations [20,21], often in isolation. However, the effectiveness of these technologies depends on a holistic approach that integrates expertise from health care, computer science, cybersecurity, and user-centered design. The lack of interdisciplinary research has led to fragmented solutions that may not fully address the practical needs of patients and caregivers. This review sought to bridge these gaps by offering a comprehensive assessment of remote monitoring technologies while emphasizing their usability, security, and ethical implications. Therefore, we formulated four research questions (RQs) to explore existing remote monitoring technologies and their applications, assess their needs and advantages, examine privacy concerns related to remote monitoring, and investigate technology adoption among older adults:

- RQ 1: What are the primary remote monitoring technologies to capture different behavioral, physiological, and environmental data?
- RQ 2: To what extent do these existing remote monitoring technologies achieve an optimal balance of efficiency, practicality, and empathy in addressing the needs of both patients and caregivers?
- RQ 3: What are the critical security and privacy considerations and concerns involved in these remote monitoring technologies?
- RQ 4: What are the critical design considerations and technology components for a comprehensive remote monitoring framework specifically tailored to patients with AD?

While this review focused on remote monitoring for individuals with AD, it is important to acknowledge that many of the challenges, technologies, and solutions discussed are also relevant to other dementia subtypes. Conditions such as vascular dementia [22,23], Lewy body dementia [24], and frontotemporal dementia [25] share overlapping symptoms [26,27], including cognitive decline, behavior changes, and mobility issues, which can similarly benefit from remote monitoring interventions. However, our decision to center this review on AD was guided by its high prevalence and the significant body of research available on remote monitoring solutions tailored specifically to ADRD. The insights presented in this paper may be broadly applicable to dementia care, although future research should further explore technology adaptations for different dementia subtypes.

The remainder of this paper is divided into 3 key sections. Section 2 outlines the methodology, comprising 3 subsections that detail the scoping criteria, literature search strategy, and

XSL•FO
RenderX

data analysis procedures. Section 3 addresses the RQs in depth. Finally, section 4 concludes the literature review, summarizing the key findings and their implications.

Objectives

Our objective with this paper was to critically assess the effectiveness of various remote sensing technologies in enhancing patient outcomes, particularly for individuals with ADRD, while also identifying strategies to reduce the burden on health care professionals (HCPs) and caregivers. To perform this assessment, the research comprehensively examined key factors essential for the successful implementation and adoption of these technologies. These factors include patient and caregiver compliance, ensuring that the technologies are used consistently and effectively; user-centered design, which focuses on creating accessible and intuitive systems tailored to the specific needs of end users; privacy and security concerns, addressing the safeguarding of sensitive patient data and ensuring adherence to ethical and regulatory standards; and the overall efficacy of these technologies, evaluating their accuracy, reliability, and impact on patient care.

Methods

Overview

In the landscape of AD care, this review undertook a comprehensive examination of the complex dynamics introduced by remote monitoring technologies. We focused on different remote monitoring technologies, types of algorithms used to process the data, issues and concerns related to security and privacy, caregiver burdens, and specific monitoring frameworks for patients with AD. We conducted an extensive search to identify relevant research papers across multiple disciplines from 2011 to 2024. Our search encompassed a diverse array of digital libraries, including Google Scholar, IEEE Xplore, PubMed, and DBLP on February 20, 2024, with an updated search on May 18, 2024, to ensure coverage across multidisciplinary domains, including health care, engineering, and computer science. These databases were selected based on their relevance to remote monitoring technologies, security and privacy considerations, and AD care. While Scopus, Embase, and Cochrane are widely recognized for systematic reviews in clinical and biomedical research, our focus on technological, interdisciplinary, and engineering aspects made IEEE Xplore and DBLP particularly valuable sources. In addition, PubMed was included to ensure coverage of relevant medical and health care studies. Google Scholar was used to capture a broader range of studies, including conference proceedings and gray literature, which are crucial for emerging technologies. On Google Scholar, the search string used was *"Remote monitoring technologies" AND "Wearable sensors" AND Alzheimer AND "Privacy concerns,"* ensuring a broad examination of related literature. PubMed and IEEE Xplore were queried using the search string *(Remote monitoring technologies) AND Alzheimer AND ("Privacy concerns")* to capture articles with a focus on technical and medical perspectives. In addition, DBLP was queried using the search string *Alzheimer AND "Remote monitoring"* to include publications emphasizing computer science and engineering insights. In addition, we cross-referenced the citations of each article to uncover additional relevant articles or government reports aligning with our inclusion criteria. Notably, we imposed some constraints on the period during which the papers were published, and only English-language papers were considered. In total, we identified 31 articles that fulfilled our inclusion criteria.

Our search strategy included keywords such as "Remote monitoring technologies" AND "Wearable sensors" AND "Alzheimer" AND "Privacy concerns" to ensure the inclusion of studies addressing security and data protection—critical challenges in adopting remote monitoring solutions. However, we acknowledge that this keyword choice may have limited the retrieval of studies focusing on broader aspects, such as clinical effectiveness, usability, and caregiver perspectives. While our selection process incorporated additional screening to identify studies covering a wider range of topics, future research could benefit from refining the search terms to ensure a more balanced representation of different remote monitoring challenges.

Study Identification and Selection

A total of 79 papers were initially identified based on their titles and abstracts that were deemed to be potentially relevant to the research topic. Of the 61 records assessed for eligibility, 30 (49%) were excluded for various reasons: 7 (23%) were books, 6 (20%) were partial articles, and 17 (57%) were out-of-context studies. Finally, 31 papers were identified to extract information and insights related to the RQs. These papers were thoroughly analyzed to identify common themes, methodologies, findings, and gaps in the existing literature. This analysis provided a deeper understanding of the current research and valuable insights to guide this review. This systematic approach to literature review and analysis enhanced the validity and reliability of the research, contributing to its overall credibility.

In the following sections, the findings from the papers are presented and discussed. The results are organized based on the information identified in the literature review with respect to the RQs.

Results and Discussion

Primary Remote Monitoring Technologies (RQ 1)

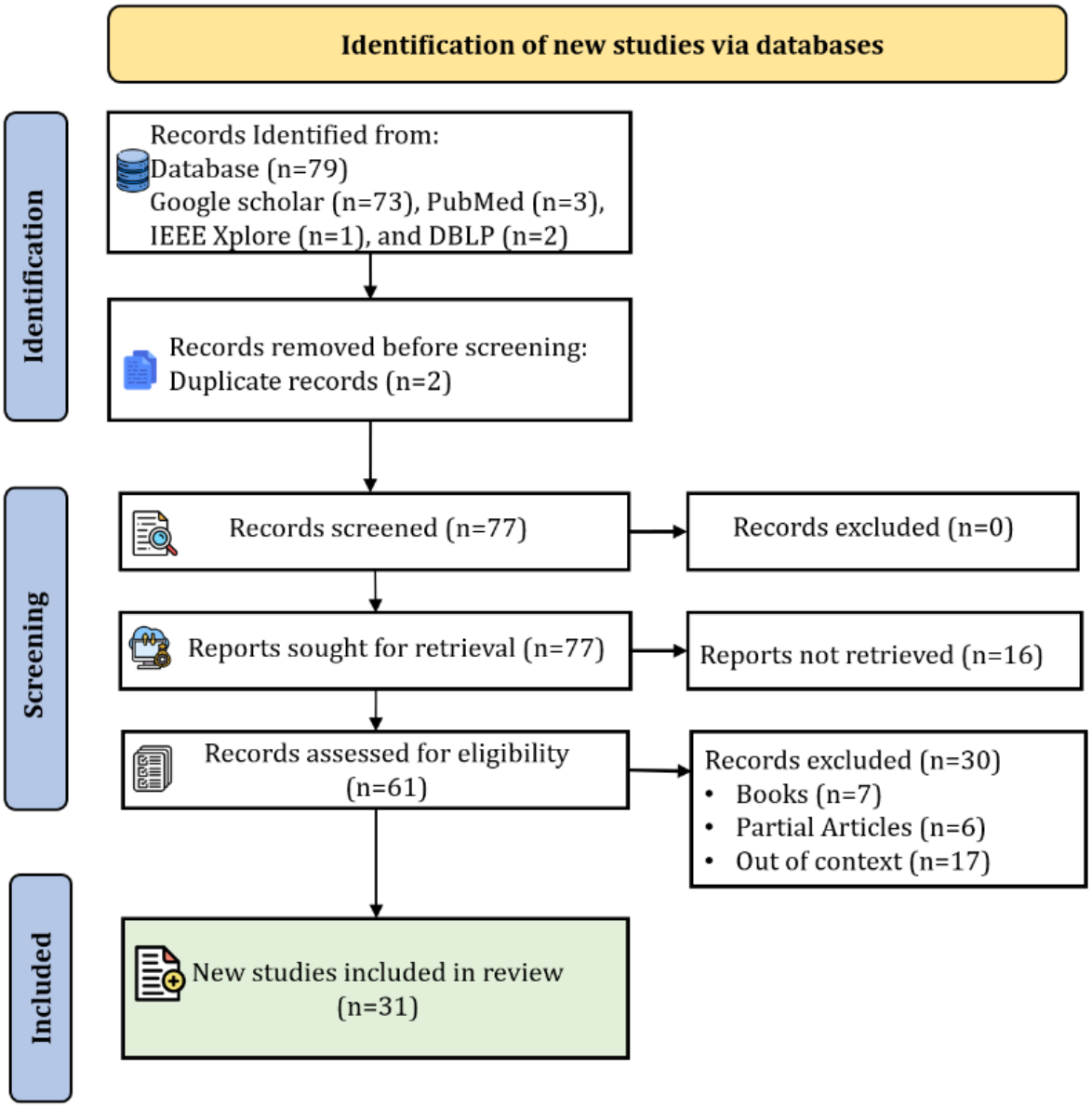
Overview

Remote monitoring technologies using wearable and portable device–based sensors have revolutionized health care by enabling continuous and real-time tracking of various physiological and environmental parameters. These sensors, embedded in devices such as smartwatches, fitness bands, and portable health monitors, can measure heart rate, blood pressure, glucose levels, temperature, and physical activity [28,29]. They offer significant advantages in managing chronic conditions, detecting early signs of health deterioration, and providing personalized care. Wearable sensors are particularly beneficial in monitoring older patients and individuals with chronic diseases, allowing for prompt interventions and reducing hospital visits. Furthermore, the integration of these devices with mobile health (mHealth) apps and cloud-based platforms facilitates seamless data collection, analysis, and sharing with

XSL•FO
RenderX

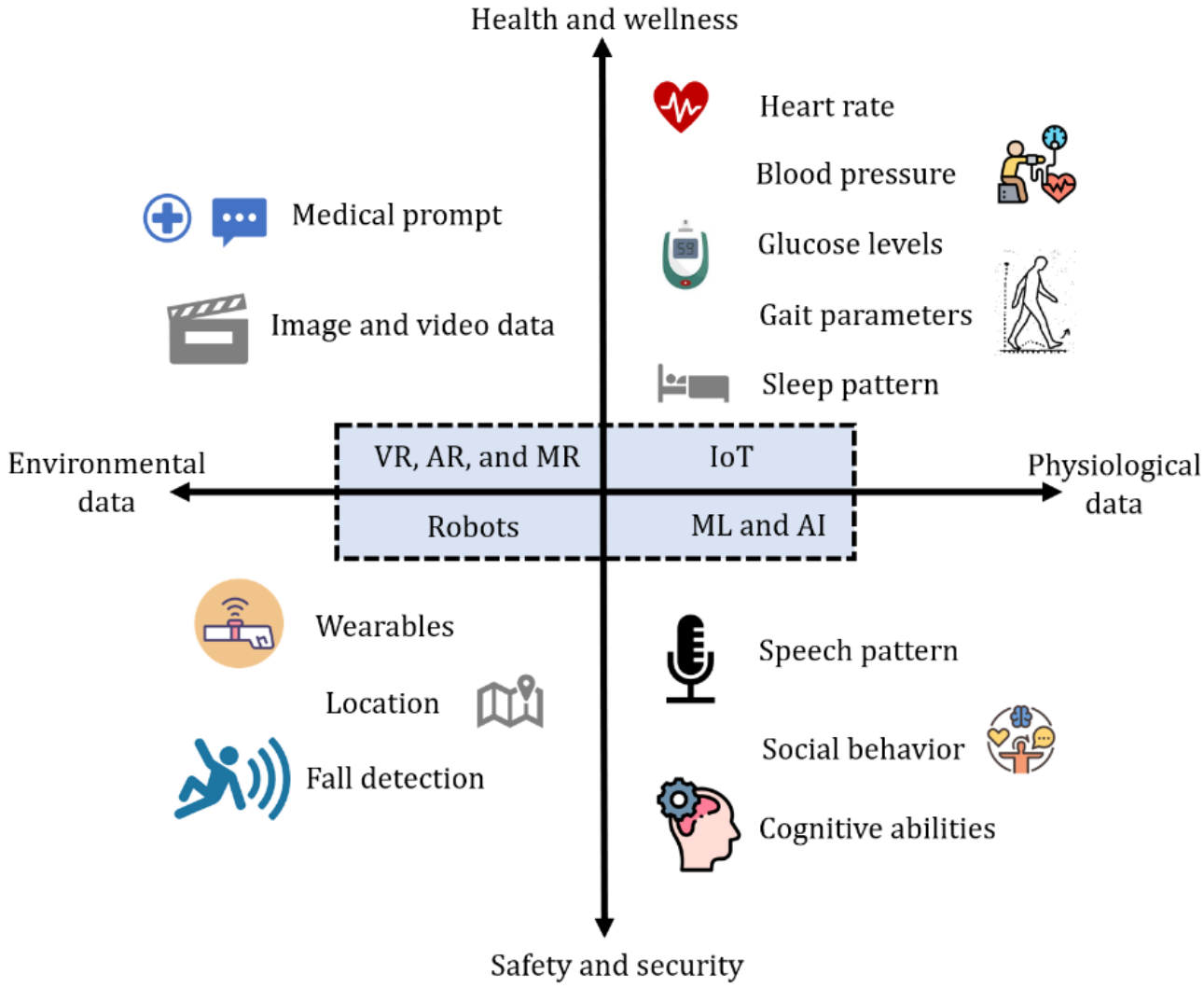
health care providers, enhancing patient engagement and improving health outcomes. The subsequent subsections provide a detailed description of the various aspects of remote monitoring technologies. To ensure a systematic approach, a thorough review of the literature was conducted following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework [30,31] (Multimedia Appendix 1), as shown in Figure 1. Figure 2 illustrates the integration and synergy of different remote monitoring technologies based on relevant parameters. These are further elaborated on in the following subsections and in Multimedia Appendices 2-5 [14,32-52].

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for the study selection in the literature review.



XSL•FO
RenderX

Figure 2. Integration and synergy of different remote monitoring technologies. AI: artificial intelligence; AR: augmented reality; IoT: Internet of Things; ML: machine learning; MR: mixed reality; VR: virtual reality.



Wearables

Wearable technologies refer to devices worn on the body either as an accessory or as clothing. Devices embedded with an accelerometer and gyroscope, such as wrist or ankle devices, are used for analyzing gait parameters such as speed, variability, and stride length [32]. There are also wearable devices that capture physical activity data, including steps taken, intensity, and duration, aiding in tracking physical activity, which helps the patient with AD be physically active and slow down the progression of the disease [32]. In addition, wrist devices and smart clothes are equipped with sensors, whereas various wearables, including wristbands, smartwatches, and chest straps, collect different data such as physiological signs, posture, gait, movement, and sleep data, and enable the continuous monitoring of blood pressure. Devices such as the Basis Health Tracker, Misfit Shine, Fitbit Flex, Withings Pulse O2, and Actiwatch Spectrum actigraph are used for sleep detection and monitoring daily activities [53-57]. These devices enable continuous monitoring of vulnerable patients, offering medical specialists a comprehensive view of the patients' health for more accurate diagnosis and treatment [33,34]. Moreover, dream headbands are used to record and store physiological data in real time during sleep, facilitating the monitoring of daily sleep patterns and enabling patients and caregivers to access their reports freely [35]. Further details are provided in Multimedia Appendix 2 [32-35,37-45] and Multimedia Appendix 3 [32-34,36-38,42,45-47].

Environment and Home Sensors

Remote monitoring through advanced environment and sensor technology provides a promising approach to enhancing care, monitoring daily activities, and detecting early signs of cognitive decline. In the context of home-based remote monitoring, various sensors are used to capture and analyze vital health and activity data. Infrared sensors strategically placed throughout the home facilitate gait analysis by detecting individual movements and providing insights into gait characteristics such as walking speed, rhythm, and variability [32]. Wireless mattress sensors and ambient sensors are used to monitor sleep time, movements, and patterns, aiding in the early detection of mild cognitive impairment (MCI) and AD. In addition, night monitoring systems equipped with motion detectors are used to monitor sleep patterns and alert caregivers to unusual nighttime behavior [32]. In the case of continuous physical activity monitoring, infrared sensors are used to collect internal

XSL•FO
RenderX

movement data, which allows for the assessment of activity level without the need for active participation of the individual [32,36]. A multisensory approach, including sound sensors, video cameras, radio-frequency identification readers, pressure sensors, light sensors, proximity sensors, temperature sensors, humidity sensors, and location trackers, has been applied to many activities and environmental conditions in the home [33]. This rich dataset facilitates the assessment of disease progression and provides valuable information about changes in daily activities and environmental factors. Behavior patterns are tracked using sensors installed in various areas of the home, including the kitchen, bedroom, and bathroom, along with fall detection sensors. These sensors help monitor mobility, stability, frequency, and time spent in different areas, contributing to a deeper understanding of daily routines and potential risks [32]. Microsoft Kinect cameras are specifically used for gait monitoring to detect and assess falls among older individuals, with the collected video data aiding in fall prevention efforts [37].

Finally, motion sensor–activated light paths and bed pressure sensors are used for fall prevention and sleep problem detection, respectively. These sensors provide an additional layer of safety and monitoring, enhancing the overall effectiveness of the remote monitoring system [34].

Personal Devices

Personal devices, specifically, smartphones equipped with advanced sensors such as accelerometers, gyroscopes, and magnetometers, play a pivotal role in gait analysis and physical activity monitoring. Apps installed on these smartphones analyze the sensor data to determine various gait characteristics, including stride length, walking speed, and rhythm, providing valuable insights into mobility patterns [32]. Moreover, waist-worn accelerometers facilitate continuous monitoring of physical activity parameters such as movements and acceleration, aiding individuals in maintaining optimal levels of physical activity and energy use [32]. In addition to physical activity monitoring, smartphones serve as effective tools for assessing social behavior. By monitoring calls, SMS text messages, and internet-browsing activities, these devices offer insights into an individual's social activity levels and patterns. This information is crucial for health care providers and caregivers to understand the social interactions of individuals, identify potential signs of isolation, and implement interventions to promote social engagement and well-being [37].

Machine Learning and Artificial Intelligence

Machine learning and artificial intelligence (AI) play a pivotal role in health care by leveraging diverse data sources to enhance diagnostic capabilities, predict disease progression, and support clinical decision-making. Driving patterns collected through GPSs are analyzed to detect early signs of preclinical AD. These patterns are compared with disease progression to identify behavior changes linked to AD [32]. Activity data from sensor-embedded devices such as wearables are used for data mining and predictive modeling, aiding in identifying potential health issues and informing health care decisions [46]. Sleep patterns collected from these devices offer valuable insights into an individual's health status, providing clinicians with early warnings of health deteriorations or relapses and enabling timely interventions to improve patient outcomes [47].

Speech pattern analysis using data from health records, including calls and SMS text messages, offers diagnostic capabilities for neurodegenerative diseases such as Parkinson disease and AD, enabling early detection and intervention [38]. In addition, AI-powered image and video processing techniques assist in diagnosing various conditions, such as cancer from x-ray and computed tomography scans and degenerative diseases from fundus photography, enhancing diagnostic accuracy and treatment planning [38]. Furthermore, data collected from wrist-worn actigraphy devices such as the ActiGraph wGT3X are analyzed using a variety of machine learning algorithms, including logistic regression, random forest, gradient-boosting machine, and support vector machine. These algorithms help predict the risks and progression of disease, providing valuable insights for personalized health care management and intervention strategies [45].

Robots

Robots such as Nao, Pepper, Paro, Stevie, Zoro, Mini, and Tangy are emerging as valuable tools in health care settings, particularly for supporting the needs of older adults. These robotic companions engage older individuals through interactive activities such as conversations, quizzes, tongue twisters, and arithmetic calculations, demonstrating promising results in enhancing social engagement and alleviating feelings of isolation [46]. AI-powered robots equipped with advanced sensors and algorithms can access environmental data and user-specific characteristics to deliver personalized medical prompts and emotional support. These robots play a crucial role in physical rehabilitation, reducing feelings of isolation and promoting overall well-being by providing companionship and interactive engagement [38,42]. Furthermore, specialized robots such as Nao, Pepper, and Paro are designed to cater to the unique needs of patients with dementia with cognitive impairment. These robots offer a comprehensive range of services, including behavioral monitoring, physical exercise tutoring, recreational activities, stress management, and medication reminders. By providing multifaceted support, these robotic companions significantly enhance the quality of life of older adults and contribute to improved health care outcomes [34]. Robots named Stevie and Zoro are specially tailored to support caregivers and older adults with memory loss. These robots offer care support, entertainment options, cognitive engagement activities, and social connectivity features. Their interactive nature stimulates older adults through engaging exercise and interaction, fostering a sense of community and connection [32]. Robots such as Mini and Tangy enhance cognitive abilities in older adults using educational games and imitation learning, aiming to improve memory, cognitive function, and mental agility through innovative stimulation and support [37].

Internet of Things

The Internet of Things (IoT) has revolutionized health care monitoring by using novel sensors and wireless communication methods such as Bluetooth, Wi-Fi, near-field communication, and Zigbee to access and securely store real-time data on vital signs, including temperature, blood pressure, heart rate, and

XSL•FO
RenderX

glucose levels, in the cloud. This advancement facilitates effective device management and real-time data monitoring and offers a secure storage solution, enhancing patient care and treatment outcomes [47]. Furthermore, the integration of multiple sensors into wearable devices, environment monitoring systems, and in-home setups provides pervasive and unobtrusive intelligence to support individuals proactively in their daily lives. These sensors enable continuous monitoring of various parameters, including temperature, humidity, and blood pressure, offering valuable insights for personalized health care interventions [39,42].

IoT-based tools such as smart pillboxes assist individuals in medication management by sending reminders according to prescribed medication plans, ensuring adherence to treatment regimens, and improving medication compliance [39]. Moreover, the implementation of semantic frameworks and advanced data processing techniques enables the collection and analysis of multi-sensor data to determine disease progression and provide timely alerts through real-time monitoring. This approach aids in the early detection of health deterioration, facilitating timely interventions and improved patient outcomes [40]. In addition, the DCARE model used a computational approach to process data from monitoring devices, incorporating content prediction, behavior recognition, and content recognition. This model serves as a comprehensive solution for general monitoring, data processing, and sending alerts to caregivers in case of potential risks or emergencies, enhancing patient safety and care coordination [41].

Virtual, Augmented, and Mixed Reality

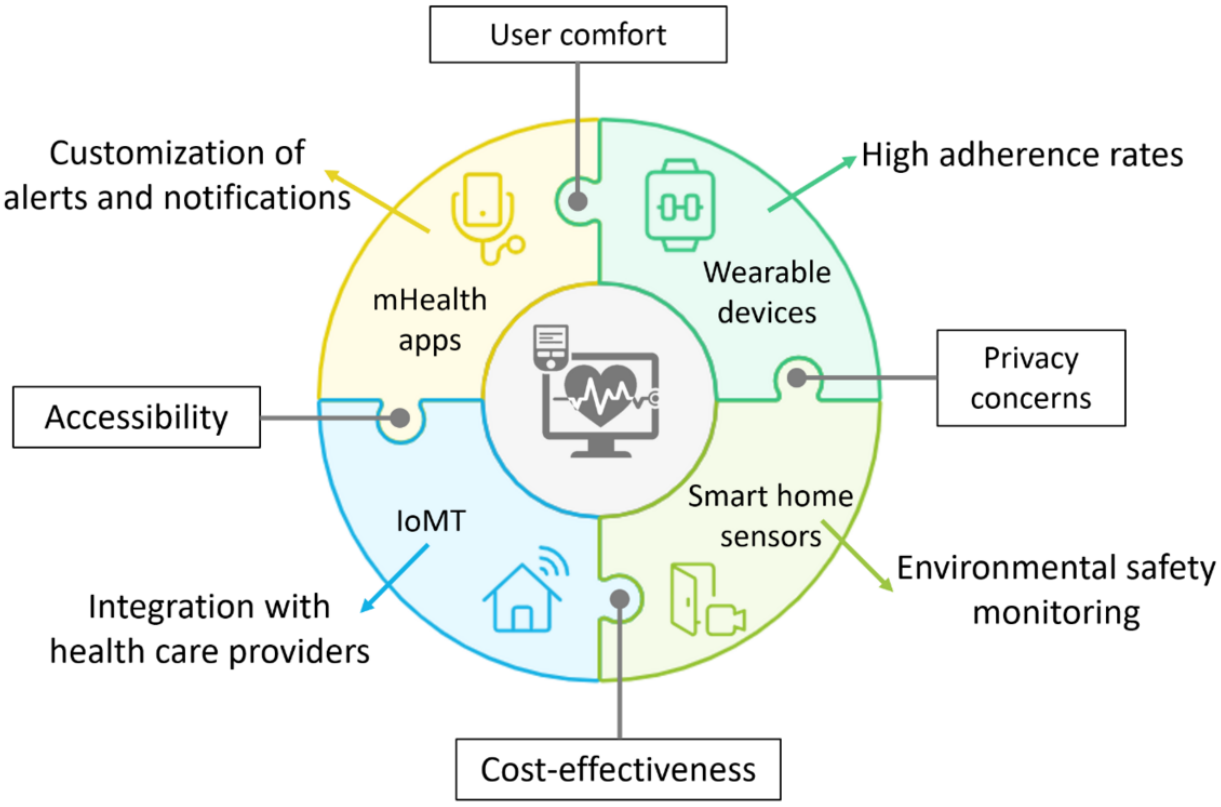
Virtual reality, augmented reality, and mixed reality technologies, facilitated by devices such as headsets, smart glasses, and next-generation smartphones, offer an immersive experience that extends beyond entertainment to include web-based home shopping, leisure activities, and communication, enriching both social interaction and physiological well-being [42]. Smart spectacles equipped with machine learning algorithms for face and object recognition enable personalized video and image playback, assisting patients with dementia and memory loss by stimulating memory recall and cognitive engagement [43]. In addition, GPS and device tracking functionalities, combined with medication reminders via Bluetooth connectivity, empower caregivers to monitor patient locations and medication adherence, ensuring timely interventions and enhancing patient safety [44].

Effectiveness of Remote Monitoring Technologies (RQ 2)

Overview

In addressing RQ 2, which investigated the effectiveness of remote monitoring technologies for individuals with ADRD, a thorough review of previous research papers was conducted. The findings are synthesized and summarized in Multimedia Appendix 4 [14,32,33,36-38,42,45,46,48-52] and Figure 3, providing a detailed comparison of various monitoring methods, their advantages, and their limitations. This comprehensive assessment helps achieve an optimal balance in addressing the needs of both patients and caregivers. The following subsections delineate different aspects pertaining to RQ 2.

Figure 3. Key features and considerations of different remote monitoring technologies. IoMT: Internet of Medical Things; mHealth: mobile health.



XSL•FO
RenderX

Wearable Devices

Wearable devices have emerged as valuable tools for remote monitoring of various health conditions, including cerebrovascular diseases, neurodegenerative diseases, and MCI. Godkin et al [48] investigated the adherence value of wearable devices focusing on 3 key aspects—adherence based on sensor location, time of day, and wear location—and participant and study partner acceptance of continuous multi-sensor wear for a week-long period. The outcomes of this study demonstrated the feasibility of a continuous multi-sensor approach to remote monitoring for individuals living with cerebrovascular diseases and neurodegenerative diseases, showing high adherence rates during both day- and nighttime wear throughout the study period. The aforementioned study also addressed concerns related to device discomfort, including itchiness, catching on clothing during dressing or on bed sheets during sleep, device size, and appearance.

Similarly, a comprehensive investigation conducted by Stavropoulos et al [14] focused on public involvement activities to capture the preferences, priorities, and concerns of people with AD and their caregivers regarding the use of monitoring wearables. Face-to-face discussions with the patient advisory board revealed varying levels of acceptance toward new technologies among older individuals, highlighting concerns regarding the cost of devices, adaption to modern technologies, and potential difficulties in operating technology. Furthermore, Stavropoulos et al [14] used a questionnaire to understand the specific requirements, preferences, barriers, and daily obstacles faced by individuals with MCI, their caregivers, and HCPs in using wearables. The Human Factors and Technology Requirements Questionnaire collected survey responses to various types of questions, such as open-ended, multiple-choice, and closed-ended questions, and the importance of user-centered design for assistive technologies using wearables and highlighted the implications for future research. The aforementioned study provided valuable insights into the perspective of individuals with MCI, caregivers, and HCPs regarding the use of assistive technologies with wearable trackers, emphasizing the need to consider privacy and ethical issues in the technology design and the disadvantages of traditional record-keeping methods.

Smart Homes (Multiple Sensors)

Smart home technologies equipped with multiple sensors have shown promising benefits in remote monitoring, particularly through video monitoring methods. However, there is a growing interest in exploring novel remote monitoring technologies beyond traditional video monitoring, emphasizing the need to diversify research in this area [36]. Despite the potential advantages, there is limited evidence to conclusively determine the acceptance of these remote sensing technologies, indicating a gap in understanding alternative methods [36]. In a comprehensive systematic review, it was highlighted that smart home remote monitoring has the potential to support older persons and caregivers by addressing individual family needs [49]. Nevertheless, critical barriers need to be addressed for successful adoption and implementation, including concerns about the time, training, and technical challenges associated with smart home technologies [49]. These challenges include issues with battery life, unstable wireless connectivity, and incompatibility, which have raised concerns about potential feelings of loneliness and isolation among older people with cognitive impairments [49]. Furthermore, there is a need for more research on alternative remote sensing technologies, such as vibration and acoustics technologies, to better understand their potential applications and effectiveness [36].

Addressing stigma and negative associations with disability through appropriate design and customization of sensing technologies, such as the use of smart fabric sensors, could help mitigate privacy concerns and enhance acceptability [36]. Future research should focus on understanding end-user conditions for acceptability and addressing various barriers, including usability, technical issues, cost concerns, and social implications, to facilitate the broader adoption of smart home technologies in health care settings [36,49].

Multimode Sensor (Machine Learning Approach)

In addressing the challenges of predicting and managing the behavioral and psychological symptoms of dementia (BPSD), multimode sensor technologies coupled with machine learning approaches have emerged as promising tools [45]. These approaches leverage diverse data sources to offer valuable insights into the relative contributions of various factors, such as caregiver-perceived triggers, cognitive and functional status, personality traits, and actigraphy-derived parameters, in predicting BPSD. However, the efficacy of these technologies can be compromised by challenges related to the improper wear of devices, leading to limited data availability for accurate prediction [45]. Despite these obstacles, machine learning models have demonstrated their capability to effectively predict different subsyndromes of BPSD. Such predictive capabilities hold significant potential to inform personalized interventions tailored to individual symptom profiles, thus facilitating timely and targeted care.

Nevertheless, it is essential to acknowledge the existing limitations and gaps in current research [45]. Cho et al [45] emphasize the need for future studies to address these challenges. This includes validating findings in larger and more diverse populations, integrating objective monitoring technologies, and developing more precise prediction algorithms to enhance the accuracy and reliability of these predictive models. Such advancements could significantly contribute to improving the quality of care and support provided to individuals living with dementia and their caregivers.

Patient and Public Involvement Survey

Patient and public involvement surveys emphasize the importance of prioritizing user-friendly design and personalized approaches when developing mHealth apps tailored for individuals with cognitive impairment [50]. Survey findings in the study by Lazarou et al [50] highlighted participants' keen interest in brain games aimed at memory enhancement, showcasing a strong demand for cognitive support tools. In addition, the participants expressed a desire for medication reminders and GPS tracking functionalities within mHealth apps, demonstrating the potential utility of such features in enhancing daily living and emergency response for individuals

XSL•FO
RenderX

with cognitive impairment [50]. Despite these valuable insights, it is important to acknowledge the survey's limitations. The research was conducted among a limited population, potentially introducing self-bias and not capturing the full spectrum of needs and preferences across diverse groups [50]. To address these challenges, future research should aim to explore more diverse populations and incorporate a broader range of user perspectives. This will help in developing more inclusive and effective mHealth interventions that cater to the diverse needs of individuals with cognitive impairment, caregivers, and HCPs.

Internet of Medical Things and Robots

The Internet of Medical Things (IoMT) is revolutionizing the health care landscape by integrating advanced technologies into smart health care systems. In the study by Mathkor et al [52], the authors provided a comprehensive overview of IoMT, highlighting its role in current and future biomedical systems. The study emphasized the integration of technologies such as deep learning, machine learning, blockchain, AI, radio-frequency identification, and Industry 5.0. These technologies hold immense potential in enhancing patient care, improving diagnostic accuracy, and streamlining health care operations. However, the implementation of IoMT systems presents several challenges that need to be addressed for successful adoption. High costs associated with device development, infrastructure setup, and connectivity solutions can pose financial barriers. In addition, many IoMT devices operate on battery power, leading to concerns about battery life and continuous operation. Data quality issues, including noise, biases, and inaccuracies, can also arise due to device limitations or environmental factors. To fully realize the benefits of IoMT, it is crucial to address these challenges and develop robust, scalable, and secure solutions. Companion robots have emerged as potential tools to address the social and emotional needs of older adults, particularly those with dementia.

In the study by Pou-Prom et al [51], the authors conducted an experiment involving 19 participants diagnosed with AD to explore the feasibility of using conversational companion robots. The study revealed that, while some participants enjoyed interacting with the robots, others expressed doubtfulness and reluctance. Technical limitations such as speech recognition challenges, lack of dialogue context, and keyword matching were identified as barriers to seamless communication. Furthermore, the absence of features such as miscommunication and acoustic features hindered the assessment of dementia-related symptoms through robot interactions. The mixed reactions from participants highlighted the importance of ongoing research to refine companion robot technologies. Future work should focus on addressing the identified limitations, exploring the trade-off between autonomous and semiautonomous approaches, and optimizing the design and functionality of these robots to better support the well-being of older adults. Collaborative efforts involving HCPs, engineers, and end users are essential to develop more effective and user-friendly companion robot solutions that can have a meaningful impact on the lives of older adults with dementia.

Cognitive Monitoring for Dementia Care

Advancements in remote monitoring have enabled the integration of cognitive assessment tools to track disease progression and evaluate treatment effectiveness in dementia care. Digital cognitive assessments such as smartphone-based memory and attention tests offer frequent, scalable evaluations that provide real-time insights into cognitive changes. In addition, AI-driven speech and typing analysis can detect subtle linguistic and motor impairments associated with cognitive decline, enabling early intervention and treatment adjustments. Gaze tracking and eye movement analysis further enhance cognitive monitoring by assessing reading patterns, recognition abilities, and attention shifts, which may indicate neurological changes over time. Meanwhile, wearable electroencephalography devices and neurofeedback tools provide continuous, objective measurements of brain activity, allowing researchers and clinicians to assess how therapeutic interventions impact cognitive function. These emerging technologies hold great promise for improving personalized treatment strategies, monitoring drug efficacy, and facilitating early diagnosis, ultimately enhancing the quality of care for individuals with dementia.

Privacy and Security in Remote Monitoring (RQ 3)

Overview

When examining the critical security and privacy considerations involved in remote monitoring technologies, particularly for patients with AD, it is essential to delve into several layers of complexity that encompass data security, privacy regulations, ethical issues, and the unique vulnerabilities of the patient population. Remote monitoring technologies, including smartphones, wearables, and environmental sensors, offer significant benefits for managing AD by enabling continuous health monitoring and early detection of anomalies and providing valuable data for research. However, these benefits come with substantial security and privacy challenges. Different aspects of privacy and security concerns are summarized in Multimedia Appendix 5.

Data Types and Associated Risks

First, the nature of the data collected by these devices—ranging from personal details to medical records, geolocation, and behavioral patterns—makes them highly sensitive. Ensuring data security is paramount as it requires robust encryption methods to protect data both in transit and at rest. Encryption safeguards the data from unauthorized access during transmission and storage. Equally important is implementing strong access control measures, ensuring that only authorized individuals such as health care providers and approved caregivers can access the patient data. Regular security audits and vulnerability assessments are crucial to identifying and addressing potential security gaps, thereby maintaining the integrity of the monitoring systems.

Privacy concerns are equally critical, particularly given the stringent requirements set forth by regulations such as the General Data Protection Regulation (GDPR) in Europe and HIPPA (Health Insurance Portability and Accountability Act) in the United States [58-60]. These regulations mandate that

XSL•FO
RenderX

patients' personal health information be protected and that individuals have control over their data. Adherence to these regulations involves ensuring that patients (or their legal guardians) provide informed consent for data collection and use. This consent process must be clear and comprehensive, detailing what data will be collected, how they will be used, who will have access, and how long the data will be retained. Data minimization principles should be applied, collecting only the necessary data required for the intended purpose, thus reducing the risk of unnecessary exposure. Patients with AD may not fully understand the implications of data collection and their rights regarding data privacy. Therefore, it is imperative to involve caregivers or legal guardians in the consent process to ensure that the patients' rights and preferences are respected. Furthermore, patients with AD are particularly vulnerable to exploitation and abuse. Continuous monitoring can inadvertently infringe on their privacy and autonomy, raising ethical concerns. Balancing the need for constant monitoring to ensure safety with the patients' rights to privacy and dignity is a delicate task. Ethical considerations must guide the implementation of these technologies, ensuring that monitoring practices do not become overly intrusive.

Device and IoT Security

From a technical perspective, the integration of various monitoring devices such as smartphones, wearables, and environmental sensors introduces additional security challenges. Each device and sensor can be a potential entry point for cyberattacks. IoT devices are often targeted due to their typically lower security standards. To counteract these risks, it is crucial to ensure that all devices are secured against tampering and hacking. This involves regular firmware updates to patch vulnerabilities, the use of strong authentication mechanisms to prevent unauthorized device access, and the use of data anonymization techniques to protect patient identity in the event of a data breach [41,61].

Furthermore, the use of advanced technologies such as AI and machine learning in processing and analyzing the collected data presents its own set of risks. While these technologies can enhance the accuracy and efficiency of health monitoring, they also require access to large datasets, which can exacerbate privacy concerns if not handled properly. Ensuring that AI and machine learning models are trained on anonymized data and that they adhere to strict privacy standards is essential to mitigate these risks. Potential attackers pose significant threats to the security of remote monitoring systems. These attackers might attempt to intercept data transmission, gain unauthorized access to monitoring devices, or exploit vulnerabilities to extract sensitive information. Addressing these threats requires a multilayered security approach, including end-to-end encryption, secure communication protocols, and robust incident mechanisms to quickly address and mitigate breaches when they occur [47].

Balancing Safety and Patient Autonomy

While remote monitoring offers substantial benefits for managing AD, it also presents critical security and privacy challenges that must be carefully addressed. Ensuring robust data security through encryption, access controls, and regular audits; complying with privacy regulations; obtaining informed consent; and addressing the unique vulnerabilities of patients with AD are essential components of a secure and ethical monitoring system [62,63].

Balancing safety and patient autonomy in remote monitoring for AD care presents a complex challenge that requires a nuanced approach. While the technology offers substantial benefits, such as continuous health monitoring, early anomaly detection, and timely interventions, it must be implemented thoughtfully to prevent undue intrusion. The need for constant observation can infringe on the patients' sense of independence, raising ethical questions about dignity and consent. To address these concerns, remote monitoring systems should integrate customizable privacy settings that allow caregivers to modify the monitoring level based on the patients' current health status and preferences. Noncritical data anonymization and local data processing using edge computing can significantly reduce the sense of surveillance, ensuring that only essential information is transmitted to central servers [64].

Obtaining informed consent from patients with AD is another critical aspect due to cognitive impairments that often affect their decision-making capabilities. In such cases, involving caregivers or legal guardians becomes essential to uphold the patients' rights while ensuring that their best interests are protected [65,66]. Technologically secure consent management systems can be used to store and track consent records digitally. Leveraging blockchain technology for tamper-proof consent logs can enhance transparency and accountability, making it clear who has agreed to data collection and how the data will be used. This approach ensures compliance with regulations such as the GDPR [60] and HIPAA [58,67], reinforcing trust among patients, caregivers, and health care providers.

Adaptive monitoring systems that support variable levels of oversight can help maintain patient autonomy by shifting from comprehensive tracking to minimal oversight as needed. This adaptability allows caregivers to balance the safety needs with the patients' right to privacy. Moreover, implementing user-controlled privacy features empowers patients and their caregivers to configure data sharing preferences, such as choosing which types of data are collected and how frequently they are transmitted. This flexibility can significantly improve user satisfaction and trust in the monitoring system.

AI integration into monitoring systems should prioritize ethical standards, ensuring that data handling is minimal and transparent. For instance, using federated learning allows algorithms to be trained directly on local devices, avoiding the need to transfer raw data to external servers [68]. This method safeguards privacy while maintaining the system's ability to deliver valuable insights. In addition, explainable AI frameworks can provide justifications for alerts and decisions, ensuring that caregivers understand the reasoning behind the system's actions [69]. Such transparency bolsters confidence in the technology and facilitates informed decision-making.

To enhance safety protocols without compromising privacy, remote monitoring systems should use multilayered privacy protocols that combine encryption, robust user authentication, and differential privacy [70]. By injecting noise into datasets,

XSL•FO
RenderX

differential privacy makes it difficult to trace data back to individual patients while allowing for meaningful analysis. Regular security audits and ethical reviews should also be conducted to ensure that monitoring practices do not become overly invasive. Independent ethics committees can provide oversight, ensuring that the technology aligns with patient-centric values and preserves autonomy.

Maintaining patient dignity is paramount, especially when using advanced monitoring technology. Nonvisual sensors such as microradar [71] systems can detect falls and wandering without capturing images or videos, offering a less invasive alternative to camera-based surveillance. Moreover, patients and caregivers should have the option to customize notification settings, choosing to receive only high-priority alerts to minimize stress. This selective alert system can enhance quality of life by allowing patients to maintain a degree of normalcy and autonomy in their daily lives.

Potential Remote Monitoring Framework for Patients With AD (RQ 4)

Overview

Critical design considerations for a potential remote monitoring framework for patients with AD include ensuring privacy and data security to comply with HIPAA regulations, using nonintrusive monitoring technologies such as microradar sensors to maintain patient dignity, and incorporating AI-driven analytics for accurate detection of falls and wandering. The system should offer real-time alerts to caregivers, be user-friendly for both patients and caregivers, and integrate with existing health care platforms for seamless data sharing. In addition, providing customizable and scalable features can accommodate diverse patient needs and facilitate broader adoption.

The development of our proposed remote monitoring framework was directly informed by the insights gained from our literature review. Our findings revealed that existing monitoring solutions such as wearable sensors and camera-based systems present significant limitations in terms of user compliance, privacy concerns, and real-time responsiveness. Wearable devices require active patient participation, which may not always be feasible for individuals with ADRD, whereas camera-based systems raise ethical and privacy concerns. In response, our framework integrates microradar technology, which allows for the nonintrusive monitoring of falls and wandering behaviors without compromising user privacy. In addition, AI-powered analytics enhance the accuracy of detection by processing microradar data to identify anomalies in movement patterns. To support real-time caregiver intervention, our system uses Web Real-Time Communication (WebRTC) and peer-to-peer (P2P) communication for seamless alerts. By addressing these limitations and leveraging interdisciplinary advancements, our framework offers a comprehensive, privacy-conscious, and scalable solution for the monitoring of patients with ADRD. Different aspects pertaining to RQ 4 are discussed in the following subsections.

Ensuring Patient Safety

Patient safety is the cornerstone of any remote monitoring system for patients with AD. Wearable devices such as smartwatches and fitness trackers can monitor crucial health indicators such as gait, sleep patterns, physical activity, and falls. Environmental sensors placed in the patient's home can track temperature, motion, and other conditions to detect unusual behavior, such as wandering or falls, ensuring that immediate action can be taken [72].

Given the sensitivity of health data, robust security measures are essential. Data from wearable devices and sensors must be encrypted to prevent unauthorized access. Compliance with regulations such as the GDPR and HIPAA is necessary to protect patient privacy. Strong authentication methods and regular software updates can help maintain data integrity and security.

Ease of Use for Caregivers

The system must be user-friendly, as caregivers for individuals with ADRD may not be technologically savvy. A simple, intuitive interface is crucial, allowing for straightforward navigation and clear, actionable alerts to reduce ambiguity and stress. Customizable notifications and automated alerts ensure that caregivers are informed of critical issues without needing to constantly monitor the system while minimizing false alarms to avoid unnecessary stress. In addition, providing educational resources within the system helps caregivers understand how to best use the monitoring tools and improve their caregiving skills. This approach empowers caregivers, enhances efficiency, and improves patient safety by facilitating timely interventions [73,74].

Ethical Considerations

Ethical considerations for a potential remote monitoring framework for patients with AD include ensuring informed consent and respecting patient autonomy. Considering cognitive impairments, obtaining informed consent must involve patients as much as possible as well as legal representatives or caregivers when necessary to ensure understanding and voluntary participation. Privacy and confidentiality are paramount, requiring compliance with data protection regulations and secure data handling practices. The system should be minimally intrusive, using discreet technologies to balance safety with patient dignity and avoiding constant visible monitoring that could lead to a loss of independence. Transparency in communication about the system's operation and data use, as well as providing feedback mechanisms, builds trust and addresses concerns. Ethical design principles involving user-centered approaches and bias mitigation in AI, alongside regular ethical reviews and oversight by ethics committees, ensure that the system remains respectful and patient centered while enhancing safety [75,76].

Compatibility With Existing Systems

To be effective, the remote monitoring system for patients with AD must integrate seamlessly with existing health care infrastructure, such as electronic health records (EHRs) [77]. This integration ensures that all relevant health data from the monitoring system are accurately and efficiently captured within the broader health care framework. Using standardized protocols

XSL•FO
RenderX

and application programming interfaces is critical for facilitating data sharing and interoperability [78]. These standards, such as Health Level Seven Fast Healthcare Interoperability Resources, ensure that data from the monitoring system can be easily and securely exchanged with different EHR systems and other health IT platforms [79]. Seamless integration allows physicians, specialists, and other clinicians to access comprehensive, up-to-date information about patients, promoting coordinated care. For instance, if a remote monitoring system detects a fall or wandering incident, this information can be automatically logged in the patient's EHR, alerting HCPs in real time. This ensures that primary care physicians, specialists, and hospital staff have a unified view of the patient's health status and recent events, enabling more informed decision-making and timely interventions.

Moreover, compatibility with existing systems helps in reducing duplication of efforts and minimizes the risk of errors. HCPs can avoid manually entering data, which not only saves time but also reduces the likelihood of transcription errors. By ensuring that the remote monitoring data are automatically integrated into EHRs, the system enhances the accuracy and reliability of patient records. In addition to technical interoperability, organizational processes and workflows must be considered to ensure smooth integration. Training for physicians, nurses, and other HCPs on how to use the new system and interpret its data within the context of existing EHRs is essential. Developing clear protocols for how and when data from the monitoring system should be used in patient care can help ensure that the system is used effectively. By enabling seamless data sharing and integration, the system supports a more coordinated and comprehensive approach to health care, ultimately improving outcomes for patients with AD and reducing the burden on caregivers and medical professionals.

Scalability and Flexibility

A scalable system is essential for accommodating the increasing number of patients with AD who require care. As demand for remote monitoring solutions grows, the system must expand without compromising performance or reliability. This need involves designing the system architecture to support a larger user base and additional data processing and storage capabilities. Scalable infrastructure ensures that the system can handle a surge in the number of patients and caregivers, allowing health care providers to maintain high standards of care even as the user population increases. Moreover, incorporating standardized interfaces and protocols such as application programming interfaces and interoperability standards ensures that the system can easily connect with new devices and technologies, facilitating continuous improvement and innovation [80].

Flexibility in the remote monitoring system is equally important. The system should have a modular architecture that allows for the easy integration of new sensors and technologies as they become available. This modular approach ensures that the system can adapt to technological advancements and incorporate innovative solutions without requiring a complete redesign. It should also be configurable to meet the specific requirements of various health care settings, such as in-home care, assisted living facilities, or hospitals, providing tailored solutions that address the unique challenges of each environment. The system's scalability and flexibility ensure that it adapts to evolving health care needs, enhancing patient care and supporting the growth of remote monitoring technologies [80].

Technology Components

Wearable technology is essential for health monitoring, allowing for the real-time tracking of multiple health variables such as heart rate, physical activity, sleep habits, and location [56,72]. These devices help monitor the patient's daily routines and identify emergencies such as falls. Environmental sensors also enhance the patient's living environment by monitoring door activity, motion, and room temperature to detect unusual activity. For instance, if a patient leaves the house at an unusual time, an alert can be issued to the caregiver. Smartphone apps enable caregivers to easily monitor and receive notifications, enter observations, create personalized alerts, and access learning materials. Continuous connectivity ensures that caregivers are always informed about the patient's status.

Implementing AI and machine learning algorithms to analyze data collected from wearables and other devices allows for the identification of trends indicating potential health issues. These predictive analytics enable caregivers to anticipate problems and take preventive measures before they become critical, thus improving patient care and safety. Moreover, an IoT infrastructure is essential for ensuring seamless data transfer between all connected devices, with cloud computing offering scalable processing and storage capacity and edge computing minimizing delays by processing data locally. Effective data management is crucial for making sense of the vast amounts of data generated by the system, allowing for trend analysis and informed clinical decisions. Advanced analytics techniques can also provide personalized patient care. In addition, robust security technologies such as intrusion detection systems, firewalls, and encryption methods are vital to protect the system from threats, with frequent maintenance and updates ensuring ongoing security

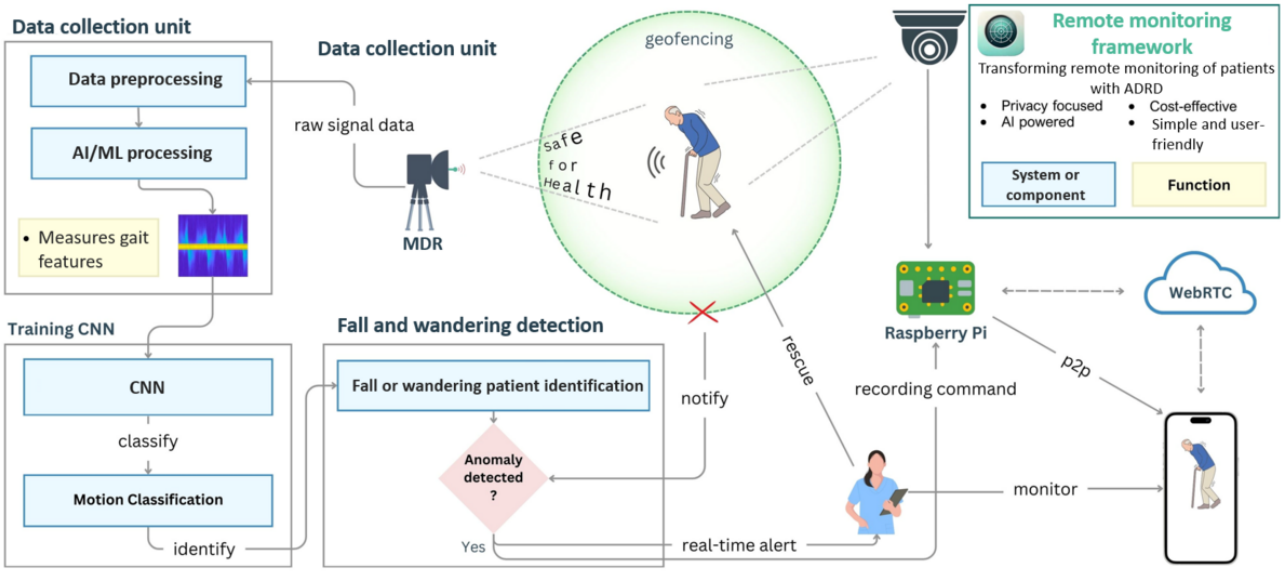
Comprehensive Framework

Overview

On the basis of the discussions in the previous subsections and the findings of previous studies, we propose a comprehensive remote monitoring framework, as shown in Figure 4, by incorporating the use of wearable technologies, sensors, AI-powered analytics, and a safe IoT infrastructure. The elements in the schematic framework are intricately connected and together create a comprehensive system for the remote monitoring of patients with AD. The framework is designed with a strong focus on health and safety. The use of microradar technology ensures that monitoring is nonintrusive and safe for patients, reducing the stress associated with traditional monitoring systems. This framework not only enhances patient safety but also alleviates the burden on caregivers by providing reliable, real-time support and intervention tools. The following is a detailed discussion of how each point contributes to the overall monitoring framework.

XSL•FO
RenderX

Figure 4. Schematic framework for remote monitoring of patients with Alzheimer disease. ADRD: Alzheimer disease and related dementias; AI: artificial intelligence; CNN: convolutional neural network; MDR: micro-Doppler radar; ML: machine learning; P2P: peer to peer; WebRTC: Web Real-Time Communication.



Data Collection

The data collection unit is the central component of the monitoring system and is responsible for gathering and processing the data required to track the patient's movements and health status effectively. At its core, the unit uses micro-Doppler radar technology to capture raw signal data by continuously monitoring the patient's movements in a nonintrusive manner [81]. This technology is crucial as it does not require the patient to wear any devices or interact with the system, thereby preserving the patient's privacy and minimizing any stress or discomfort that might arise from traditional monitoring systems. The raw data collected by the micro-Doppler radar undergoes a critical preprocessing stage in which essential gait features are extracted. This step is vital in filtering out noise and isolating the most relevant aspects of the patient's movements that could signal potential risks. Once preprocessed, the data are fed into advanced AI and machine learning algorithms, which analyze the extracted features to identify patterns in the patient's movements. This sophisticated processing enables the system to distinguish between normal daily activities and potentially hazardous situations such as falls or wandering, thereby enhancing the system's ability to ensure the patient's safety while maintaining their dignity and comfort.

Data Processing

The data processing component plays a critical role in enhancing the system's ability to accurately classify various patient movements, which is fundamental to ensuring effective monitoring and maintaining the safety of individuals with ADRD. The process begins with preprocessed data being fed into a convolutional neural network [82], a powerful AI model specifically designed to recognize and analyze patterns within the input data. This network is trained to classify different types of motions, enabling it to distinguish between normal daily activities such as walking or sitting and potentially hazardous incidents that may require immediate intervention, such as falls or wandering outside a designated safe zone. The convolutional neural network's ability to make these distinctions is vital to the system's overall accuracy and reliability as it allows for the early detection of unusual or dangerous behaviors. This, in turn, facilitates timely alerts and responses, ensuring that caregivers can intervene quickly and appropriately to protect the patient's well-being. Through this advanced data processing approach, the system not only enhances patient safety but also provides peace of mind to caregivers through the knowledge that the system is capable of accurately monitoring and responding to the specific needs of patients with ADRD.

Fall and Wandering Detection

The fall and wandering detection component is a critical part of the monitoring system, designed to ensure the continuous safety of patients by closely tracking their movements and identifying any signs of falling or wandering outside a predefined safe area. The system uses a geofencing approach in which a *defined* boundary or safe zone is established around the patient's environment. This safe zone is carefully configured based on the patient's needs and typical movement patterns. The system continuously monitors the patient's movements within this boundary using sophisticated pattern identification techniques. By analyzing these movement patterns in real time, the system can quickly detect any deviations from the norm. For example, if the patient crosses this virtual boundary or exhibits movements that suggest a fall, such as a sudden loss of balance or an abrupt change in posture, the system immediately flags this behavior as an anomaly. This continuous, vigilant monitoring is essential for identifying potential risks before they can lead to serious harm.

Once an anomaly is detected, the system's response should be both swift and automated. The anomaly detection process should be finely tuned to recognize specific triggers, such as the patient moving outside the safe zone or experiencing a fall. When such an event is identified, the system automatically triggers an alert, ensuring that caregivers are notified in real time. These alerts should be designed to reach caregivers through various channels,

XSL•FO
RenderX

such as mobile notifications or direct messages, providing them with immediate information about the patient's situation. This real-time notification system is crucial for enabling caregivers to respond promptly and appropriately to any incidents, whether it involves assisting the patient after a fall or guiding them back to safety after wandering. The combination of continuous monitoring and instantaneous alerting forms a robust safety net, significantly reducing the risks associated with falls and wandering in patients with ADRD.

Caregiver Notification and Support

The caregiver notification and support component is designed to empower caregivers with the tools and information they need to effectively monitor and respond to the needs of patients in real time. Central to this component are real-time alerts, which are instantly sent to caregivers' devices, ensuring that they are always aware of the patient's status. These immediate notifications are particularly crucial in emergency situations, such as when a fall or wandering event occurs. By receiving alerts the moment an anomaly is detected, caregivers can intervene promptly, providing necessary assistance or preventing further harm to the patient. This real-time communication is key to maintaining the safety and well-being of patients with ADRD.

In addition, the system is equipped with a recording command and rescue feature facilitated by a Raspberry Pi module. This module plays a vital role by capturing and recording the event as soon as an anomaly is detected. The recorded data are automatically sent to caregivers, offering them detailed insights into the incident. This documentation not only aids in immediate rescue actions but also serves as a valuable resource for later analysis. By reviewing recorded incidents, caregivers and medical professionals can better understand the circumstances leading up to the event, which is essential for medical evaluations and refining future care strategies. The ability to document and analyze these events ensures that care can be continuously improved, ultimately enhancing the quality of life of both patients and caregivers.

WebRTC and P2P Communication

The framework integrates advanced communication technologies to ensure seamless monitoring and interaction between caregivers and the system, providing a reliable and immediate connection regardless of location. A key component of this system is the use of WebRTC technology, which facilitates real-time P2P communication [83,84]. This technology allows caregivers to monitor the patient from their smartphones or other devices with ease and without delay, ensuring that they remain connected to the patient's status at all times. The use of WebRTC ensures that the communication is not only immediate but also direct, reducing latency and enhancing the responsiveness of the system.

This capability significantly enhances patient safety as it allows caregivers to maintain a continuous, real-time connection with the system, enabling them to respond quickly to any situation that may arise. Whether it is monitoring routine activities or responding to alerts triggered by the system, WebRTC ensures that caregivers can provide timely and effective interventions. The seamless monitoring facilitated by this technology is crucial for reinforcing the safety net around patients with ADRD, giving caregivers the confidence and tools they need to manage their loved ones' care effectively, no matter where they are.

Conclusions

This study highlights the revolutionary potential of remote monitoring technologies in AD care, emphasizing the critical balance among innovation, practicality, and empathy. The comprehensive analysis revealed that, while wearable devices and ambient sensors significantly enhance the ability to track and support patients with AD, they also introduce challenges related to privacy, data security, and user acceptance. The findings stress the necessity of involving patients and caregivers in the design process to ensure that these technologies meet their specific needs and regulatory requirements, which is crucial to the successful implementation of these technologies.

The proposed remote monitoring framework integrates cutting-edge microradar technology with AI-powered processing to create a nonintrusive, privacy-conscious solution for monitoring patients with AD and dementia. By focusing on real-time data processing, caregiver notification, and seamless communication, the system not only enhances patient safety but also provides critical support to caregivers, ultimately improving the quality of care for individuals with ADRD. Future research can focus on refining these systems to enhance their usability, accessibility, and security, thereby improving the overall caregiving experience and patient well-being. By advancing remote monitoring practices, we can create a supportive environment that alleviates caregiver burdens and promotes a higher quality of life for individuals living with AD.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[DOCX File , 31 KB - aging_v8i1e69175_app1.docx]

Multimedia Appendix 2
Remote monitoring technologies using wearable and portable device–based sensors.
[DOCX File , 45 KB - aging_v8i1e69175_app2.docx]

XSL•FO
RenderX

Multimedia Appendix 3
Remote monitoring technologies using nonwearable home sensors.
[DOCX File , 44 KB - aging_v8i1e69175_app3.docx]

Multimedia Appendix 4
Remote monitoring technologies for dementia care and their role in cognitive assessment.
[DOCX File , 49 KB - aging_v8i1e69175_app4.docx]

Multimedia Appendix 5
Different privacy and security considerations in remote monitoring.
[DOCX File , 14 KB - aging_v8i1e69175_app5.docx]

References

1. Cipriani G, Danti S, Picchi L, Nuti A, Fiorino MD. Daily functioning and dementia. Dement Neuropsychol 2020;14(2):93-102 [FREE Full text] [doi: 10.1590/1980-57642020dn14-020001] [Medline: 32595877]
2. Tatulian SA. Challenges and hopes for Alzheimer's disease. Drug Discov Today 2022 Apr;27(4):1027-1043 [FREE Full text] [doi: 10.1016/j.drudis.2022.01.016] [Medline: 35121174]
3. Breijyeh Z, Karaman R. Comprehensive review on Alzheimer's disease: causes and treatment. Molecules 2020 Dec 08;25(24):5789 [FREE Full text] [doi: 10.3390/molecules25245789] [Medline: 33302541]
4. Abeysinghe AA, Deshapriya RD, Udawatte C. Alzheimer's disease; a review of the pathophysiological basis and therapeutic interventions. Life Sci 2020 Sep 01;256:117996 [FREE Full text] [doi: 10.1016/j.lfs.2020.117996] [Medline: 32585249]
5. Lei P, Ayton S, Bush AI. The essential elements of Alzheimer’s disease. J Biol Chem 2021;296:100105 [FREE Full text] [doi: 10.1074/jbc.rev120.008207]
6. Castellani RJ, Rolston RK, Smith MA. Alzheimer disease. Dis Mon 2010 Sep;56(9):484-546 [FREE Full text] [doi: 10.1016/j.disamonth.2010.06.001] [Medline: 20831921]
7. Shin JY, Habermann B. Caregivers of adults living with Alzheimer's disease or dementia in 2020: a secondary analysis. J Gerontol Nurs 2022 Sep;48(9):15-25 [FREE Full text] [doi: 10.3928/00989134-20220805-02] [Medline: 36007215]
8. Reinhard SC, Given B, Petlick NH, Bemis A. Supporting family caregivers in providing care. In: Hughes RG, editor. Patient Safety and Quality: An Evidence-Based Handbook for Nurses. Rockville, MD: Agency for Healthcare Research and Quality (US); 2008.
9. Culberson JW, Kopel J, Sehar U, Reddy PH. Urgent needs of caregiving in ageing populations with Alzheimer's disease and other chronic conditions: support our loved ones. Ageing Res Rev 2023 Sep;90:102001 [FREE Full text] [doi: 10.1016/j.arr.2023.102001] [Medline: 37414157]
10. Koca E, Taşkapilioğlu Ö, Bakar M. Caregiver burden in different stages of Alzheimer's disease. Noro Psikiyatr Ars 2017 Mar;54(1):82-86 [FREE Full text] [doi: 10.5152/npa.2017.11304] [Medline: 28566965]
11. Manzini CS, do Vale FA. Emotional disorders evidenced by family caregivers of older people with Alzheimer's disease. Dement Neuropsychol 2020;14(1):56-61 [FREE Full text] [doi: 10.1590/1980-57642020dn14-010009] [Medline: 32206199]
12. Liao X, Huang Y, Zhang Z, Zhong S, Xie G, Wang L, et al. Factors associated with health-related quality of life among family caregivers of people with Alzheimer's disease. Psychogeriatrics 2020 Jul;20(4):398-405 [FREE Full text] [doi: 10.1111/psyg.12528] [Medline: 32027776]
13. Lott SA, Streel E, Bachman S, Bode K, Dyer J, Fitzer-Attas C, et al. Digital health technologies for Alzheimer's disease and related dementias: initial results from a landscape analysis and community collaborative effort. J Prev Alzheimers Dis 2024;11(5):1480-1489. [doi: 10.14283/jpad.2024.103] [Medline: 39350395]
14. Stavropoulos TG, Lazarou I, Diaz A, Gove D, Georges J, Manyakov NV, et al. Wearable devices for assessing function in Alzheimer's disease: a European public involvement activity about the features and preferences of patients and caregivers. Front Aging Neurosci 2021 Apr 12;13:643135 [FREE Full text] [doi: 10.3389/fnagi.2021.643135] [Medline: 33912025]
15. Dieffenderfer J, Brewer A, Noonan MA, Smith M, Eichenlaub E, Haley KL. A wearable system for continuous monitoring and assessment of speech, gait, and cognitive decline for early diagnosis of ADRD. In: Proceedings of the 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society. 2023 Presented at: EMBC 2023; July 24-27, 2023; Sydney, Australia. [doi: 10.1109/embc40787.2023.10339986]
16. Chandrasekaran R, Sadiq T M, Moustakas E. Usage trends and data sharing practices of healthcare wearable devices among US adults: cross-sectional study. J Med Internet Res 2025 Feb 21;27:e63879 [FREE Full text] [doi: 10.2196/63879] [Medline: 39982763]
17. Chung J, Pretzer-Aboff I, Parsons P, Falls K, Bulut E. Using a device-free wi-fi sensing system to assess daily activities and mobility in low-income older adults: protocol for a feasibility study. JMIR Res Protoc 2024 Nov 12;13:e53447 [FREE Full text] [doi: 10.2196/53447] [Medline: 39531268]

XSL•FO
RenderX

18. Tiersen F, Batey P, Harrison MJ, Naar L, Serban AI, Daniels SJ, et al. Smart home sensing and monitoring in households with dementia: user-centered design approach. JMIR Aging 2021 Aug 11;4(3):e27047 [FREE Full text] [doi: 10.2196/27047] [Medline: 34383672]
19. Hackett K, Xu S, McKniff M, Paglia L, Barnett I, Giovannetti T. Mobility-based smartphone digital phenotypes for unobtrusively capturing everyday cognition, mood, and community life-space in older adults: feasibility, acceptability, and preliminary validity study. JMIR Hum Factors 2024 Nov 22;11:e59974 [FREE Full text] [doi: 10.2196/59974] [Medline: 39576984]
20. Peterson CM, Mikal JP, McCarron HR, Finlay JM, Mitchell LL, Gaugler JE. The feasibility and utility of a personal health record for persons with dementia and their family caregivers for web-based care coordination: mixed methods study. JMIR Aging 2020 Jun 26;3(1):e17769 [FREE Full text] [doi: 10.2196/17769] [Medline: 32589158]
21. Alhammad N, Alajlani M, Abd-Alrazaq A, Epiphaniou G, Arvanitis T. Patients' perspectives on the data confidentiality, privacy, and security of mHealth apps: systematic review. J Med Internet Res 2024 May 31;26:e50715 [FREE Full text] [doi: 10.2196/50715] [Medline: 38820572]
22. Deng C, Shen N, Li G, Zhang K, Yang S. Digital isolation and dementia risk in older adults: longitudinal cohort study. J Med Internet Res 2025 Feb 19;27:e65379 [FREE Full text] [doi: 10.2196/65379] [Medline: 39969956]
23. Davis BH, Tamblin SP, Troutman-Jordan M. Expanding social interaction through online technology: using social media in vascular dementia. In: Huang L, Davis B, editors. Language, Aging and Society. Cham, Switzerland: Palgrave Macmillan; 2024.
24. Zaslavsky O, Kaneshiro J, Chu F, Teng A, Domoto-Reilly K, Chen AT. Virtual intervention for caregivers of persons with Lewy body dementia: pilot quasi-experimental single-arm study. JMIR Form Res 2022 Jul 29;6(7):e37108 [FREE Full text] [doi: 10.2196/37108] [Medline: 35904843]
25. Paolillo EW, Casaletto K, Clark A, Taylor J, Heuer H, Wise A, et al. Examining associations between smartphone use and clinical severity in frontotemporal dementia: proof-of-concept study. JMIR Aging 2024 Jun 26;7:e52831 [FREE Full text] [doi: 10.2196/52831] [Medline: 38922667]
26. Infographic: understanding different types of dementia. National Institutes of Health National Institute on Aging. URL: https://www.nia.nih.gov/health/alzheimers-and-dementia/understanding-different-types-dementia [accessed 2025-02-23]
27. Kalaria R. Similarities between Alzheimer's disease and vascular dementia. J Neurol Sci 2002 Nov 15;203-204:29-34 [FREE Full text] [doi: 10.1016/s0022-510x(02)00256-3] [Medline: 12417353]
28. Deng Z, Guo L, Chen X, Wu W. Smart wearable systems for health monitoring. Sensors (Basel) 2023 Feb 23;23(5):2479 [FREE Full text] [doi: 10.3390/s23052479] [Medline: 36904682]
29. Masoumian Hosseini M, Masoumian Hosseini ST, Qayumi K, Hosseinzadeh S, Sajadi Tabar SS. Smartwatches in healthcare medicine: assistance and monitoring; a scoping review. BMC Med Inform Decis Mak 2023 Nov 03;23(1):248 [FREE Full text] [doi: 10.1186/s12911-023-02350-w] [Medline: 37924029]
30. Haddaway NR, Page MJ, Pritchard CC, McGuinness LA. PRISMA2020: an R package and shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and open synthesis. Campbell Syst Rev 2022 Mar 27;18(2):e1230 [FREE Full text] [doi: 10.1002/cl2.1230] [Medline: 36911350]
31. Page MJ, McKenzie J, Bossuyt P, Boutron I, Hoffmann T, Mulrow C, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
32. Popp Z, Low S, Igwe A, Rahman MS, Kim M, Khan R, et al. Shifting from active to passive monitoring of Alzheimer disease: the state of the research. J Am Heart Assoc 2024 Jan 16;13(2):e031247 [FREE Full text] [doi: 10.1161/JAHA.123.031247] [Medline: 38226518]
33. Fares N, Sherratt RS, Elhajj IH. Directing and orienting ICT healthcare solutions to address the needs of the aging population. Healthcare (Basel) 2021 Feb 02;9(2):147 [FREE Full text] [doi: 10.3390/healthcare9020147] [Medline: 33540510]
34. Anghel I, Cioara T, Moldovan D, Antal M, Pop CD, Salomie I, et al. Smart environments and social robots for age-friendly integrated care services. Int J Environ Res Public Health 2020 May 27;17(11):3801 [FREE Full text] [doi: 10.3390/ijerph17113801] [Medline: 32471108]
35. Chouraki A, Tournant J, Arnal P, Pépin JL, Bailly S. Objective multi-night sleep monitoring at home: variability of sleep parameters between nights and implications for the reliability of sleep assessment in clinical trials. Sleep 2023 May 10;46(5):zsac319 [FREE Full text] [doi: 10.1093/sleep/zsac319] [Medline: 36583300]
36. Tong K, Attenborough K, Sharp D, Taherzadeh S, Deepak-Gopinath M, Vseteckova J. Acceptability of remote monitoring in assisted living/smart homes in the United Kingdom and associated use of sounds and vibrations—a systematic review. Appl Sci 2024 Jan 19;14(2):843 [FREE Full text] [doi: 10.3390/app14020843]
37. Read E, Woolsey C, Donelle L, Weeks L, Chinho N. Passive remote monitoring and aging in place: a scoping review. Can J Aging 2022 Aug 01;42(1):20-32 [FREE Full text] [doi: 10.1017/s0714980822000198]
38. Shiwani T, Relton S, Evans R, Kale A, Heaven A, Clegg A, Ageing Data Research Collaborative (Geridata) AI group, et al. New horizons in artificial intelligence in the healthcare of older people. Age Ageing 2023 Dec 01;52(12):afad219 [FREE Full text] [doi: 10.1093/ageing/afad219] [Medline: 38124256]

XSL•FO
RenderX

39. Purohit P, Khanpara P, Patel U, Kathiria P. IoT based ambient assisted living technologies for healthcare: concepts and design challenges. In: Proceedings of the 2022 Sixth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud). 2022 Presented at: I-SMAC 2022; November 10-12, 2022; Dharan, Nepal. [doi: 10.1109/i-smac55078.2022.9987375]
40. Giannios G, Mpaltadoros L, Alepopoulos V, Grammatikopoulou M, Stavropoulos TG, Nikolopoulos S, et al. A semantic framework to detect problems in activities of daily living monitored through smart home sensors. Sensors (Basel) 2024 Feb 08;24(4):1107 [FREE Full text] [doi: 10.3390/s24041107] [Medline: 38400265]
41. Machado SD, Tavares JE, Martins MG, Barbosa JL, González GV, Leithardt VR. Ambient intelligence based on IoT for assisting people with Alzheimer's disease through context histories. Electronics 2021 May 25;10(11):1260 [FREE Full text] [doi: 10.3390/electronics10111260]
42. Abdi S, de Witte L, Hawley M. Emerging technologies with potential care and support applications for older people: review of gray literature. JMIR Aging 2020 Aug 11;3(2):e17286 [FREE Full text] [doi: 10.2196/17286] [Medline: 32780020]
43. Patil S, Shobha T, Kumari JJ, Khanum R, Anjum R, Manjunath S. Medical assistance for Alzheimer's disease using smart specs. In: Proceedings of the 4th International Conference on Circuits, Control, Communication and Computing. 2022 Presented at: I4C 2022; December 21-23, 2022; Bangalore, India URL: https://doi.org/10.1109/I4C57141.2022.10057684 [doi: 10.1109/i4c57141.2022.10057684]
44. Amato F, Crovari P, Masciadri A, Bianchi S, Pasquarelli MG, Toldo M, et al. CLONE: a promising system for the remote monitoring of Alzheimer's patients: an experimentation with a wearable device in a village for Alzheimer's care. In: Proceedings of the 4th EAI International Conference on Smart Objects and Technologies for Social Good. 2018 Presented at: Goodtechs '18; November 28-30, 2018; Bologna, Italy. [doi: 10.1145/3284869.3284906]
45. Cho E, Kim S, Heo SJ, Shin J, Hwang S, Kwon E, et al. Machine learning-based predictive models for the occurrence of behavioral and psychological symptoms of dementia: model development and validation. Sci Rep 2023 May 18;13(1):8073 [FREE Full text] [doi: 10.1038/s41598-023-35194-5] [Medline: 37202454]
46. Sapci AH, Sapci HA. Innovative assisted living tools, remote monitoring technologies, artificial intelligence-driven solutions, and robotic systems for aging societies: systematic review. JMIR Aging 2019 Nov 29;2(2):e15429 [FREE Full text] [doi: 10.2196/15429] [Medline: 31782740]
47. Li C, Wang J, Wang S, Zhang Y. A review of IoT applications in healthcare. Neurocomputing 2024 Jan 14;565:127017 [FREE Full text] [doi: 10.1016/j.neucom.2023.127017]
48. Godkin FE, Turner E, Demnati Y, Vert A, Roberts A, Swartz RH, et al. Feasibility of a continuous, multi-sensor remote health monitoring approach in persons living with neurodegenerative disease. J Neurol 2022 May;269(5):2673-2686 [FREE Full text] [doi: 10.1007/s00415-021-10831-z] [Medline: 34705114]
49. Tian YJ, Felber NA, Pageau F, Schwab DR, Wangmo T. Benefits and barriers associated with the use of smart home health technologies in the care of older persons: a systematic review. BMC Geriatr 2024 Feb 14;24(1):152 [FREE Full text] [doi: 10.1186/s12877-024-04702-1] [Medline: 38355464]
50. Lazarou I, Stavropoulos TG, Mpaltadoros L, Nikolopoulos S, Koumanakos G, Tsolaki M, et al. Human factors and requirements of people with cognitive impairment, their caregivers, and healthcare professionals for mHealth apps including reminders, games, and geolocation tracking: a survey-questionnaire study. J Alzheimers Dis Rep 2021 Jun 11;5(1):497-513 [FREE Full text] [doi: 10.3233/ADR-201001] [Medline: 34368634]
51. Pou-Prom C, Raimondo S, Rudzicz F. A conversational robot for older adults with Alzheimer's disease. ACM Trans Hum Rob Interact 2020 May 31;9(3):1-25 [FREE Full text] [doi: 10.1145/3380785]
52. Mathkor DM, Mathkor N, Bassfar Z, Bantun F, Slama P, Ahmad F, et al. Multirole of the internet of medical things (IoMT) in biomedical systems for managing smart healthcare systems: an overview of current and future innovative trends. J Infect Public Health 2024 Apr;17(4):559-572 [FREE Full text] [doi: 10.1016/j.jiph.2024.01.013] [Medline: 38367570]
53. Sushames A, Edwards A, Thompson F, McDermott R, Gebel K. Validity and reliability of Fitbit flex for step count, moderate to vigorous physical activity and activity energy expenditure. PLoS One 2016 Sep 02;11(9):e0161224 [FREE Full text] [doi: 10.1371/journal.pone.0161224] [Medline: 27589592]
54. Mantua J, Gravel N, Spencer RM. Reliability of sleep measures from four personal health monitoring devices compared to research-based actigraphy and polysomnography. Sensors (Basel) 2016 May 05;16(5):646 [FREE Full text] [doi: 10.3390/s16050646] [Medline: 27164110]
55. Petrucci GJ, Freedson P, Masteller B, Cox M, Staudenmayer J, Sirard J. Sensitivity of the Misfit Shine™ to detect changes in laboratory-based and free-living physical activity. J Meas Phys Behav 2018 Mar;1(1):18-25 [FREE Full text] [doi: 10.1123/JMPB.2017-0006]
56. Murakami H, Kawakami R, Nakae S, Yamada Y, Nakata Y, Ohkawara K, et al. Accuracy of 12 wearable devices for estimating physical activity energy expenditure using a metabolic chamber and the doubly labeled water method: validation study. JMIR Mhealth Uhealth 2019 Aug 02;7(8):e13938 [FREE Full text] [doi: 10.2196/13938] [Medline: 31376273]
57. Vlami K, Pantelidi K, Dalamaga M, Karagianni F, Theodoropoulos K, Papiris S, et al. Psoriatic insomnia: a subjective and objective sleep evaluation. Acta Derm Venereol 2023 Feb 14;103:adv00864 [FREE Full text] [doi: 10.2340/actadv.v103.4507] [Medline: 36786739]

XSL•FO
RenderX

58. Oakley A. HIPAA, HIPPA, or HIPPO: what really is the heath insurance portability and accountability act? Biotechnol Law Rep 2023 Dec 13;42(6):306-318 [FREE Full text] [doi: 10.1089/blr.2023.29329.aso]
59. Bradford L, Aboy M, Liddell K. International transfers of health data between the EU and USA: a sector-specific approach for the USA to ensure an 'adequate' level of protection. J Law Biosci 2020 Oct 15;7(1):lsaa055 [FREE Full text] [doi: 10.1093/jlb/lsaa055] [Medline: 34221424]
60. Zaeem RN, Barber KS. The effect of the GDPR on privacy policies: recent progress and future promise. ACM Trans Manage Inf Syst 2020 Dec 08;12(1):1-20 [FREE Full text] [doi: 10.1145/3389685]
61. Gochoo M, Alnajjar F, Tan TH, Khalid S. Towards privacy-preserved aging in place: a systematic review. Sensors (Basel) 2021 Apr 28;21(9):3082 [FREE Full text] [doi: 10.3390/s21093082] [Medline: 33925161]
62. Uddin R, Koo I. Real-time remote patient monitoring: a review of biosensors integrated with multi-hop IoT systems via cloud connectivity. Appl Sci 2024 Feb 25;14(5):1876 [FREE Full text] [doi: 10.3390/app14051876]
63. Shojaei P, Vlahu-Gjorgievska E, Chow YW. Security and privacy of technologies in health information systems: a systematic literature review. Computers 2024 Jan 31;13(2):41 [FREE Full text] [doi: 10.3390/computers13020041]
64. Krishnamoorthy S, Dua A, Gupta S. Role of emerging technologies in future IoT-driven Healthcare 4.0 technologies: a survey, current challenges and future directions. J Ambient Intell Hum Comput 2021 May 12;14:361-407 [FREE Full text] [doi: 10.1007/s12652-021-03302-w]
65. Daly TP. Informing consent to antibodies in Alzheimer's disease. BMJ 2023 Oct 11;383:2350. [doi: 10.1136/bmj.p2350] [Medline: 37821123]
66. Gao SS, Chu CH, Young FY. Oral health and care for elderly people with Alzheimer's disease. Int J Environ Res Public Health 2020 Aug 07;17(16):5713 [FREE Full text] [doi: 10.3390/ijerph17165713] [Medline: 32784777]
67. Nass SJ, Levit LA, Gostin LO. HIPAA, the privacy rule, and its application to health research. In: Beyond the HIPAA Privacy Rule: Enhancing Privacy, Improving Health Through Research. Washington, DC: National Academies Press (US); 2009.
68. Li L, Fan Y, Tse M, Lin KY. A review of applications in federated learning. Comput Ind Eng 2020 Nov;149:106854. [doi: 10.1016/j.cie.2020.106854]
69. Dwivedi R, Dave D, Naik H, Singhal S, Omer R, Patel P, et al. Explainable AI (XAI): core ideas, techniques, and solutions. ACM Comput Surv 2023 Jan 16;55(9):1-33. [doi: 10.1145/3561048]
70. Mahmood Z, Jusas V. Blockchain-enabled: multi-layered security federated learning platform for preserving data privacy. Electronics 2022 May 19;11(10):1624. [doi: 10.3390/electronics11101624]
71. Feng X, Shan Z, Zhao Z, Xu Z, Zhang T, Zhou Z, et al. Millimeter-wave radar monitoring for elder’s fall based on multi-view parameter fusion estimation and recognition. Remote Sens 2023 Apr 16;15(8):2101. [doi: 10.3390/rs15082101]
72. Chakrabarti S, Biswas N, Jones LD, Kesari S, Ashili S. Smart consumer wearables as digital diagnostic tools: a review. Diagnostics (Basel) 2022 Aug 31;12(9):2110 [FREE Full text] [doi: 10.3390/diagnostics12092110] [Medline: 36140511]
73. Hussein T, Chauhan PK, Dalmer NK, Rudzicz F, Boger J. Exploring interface design to support caregivers' needs and feelings of trust in online content. J Rehabil Assist Technol Eng 2020 Dec 08;7:2055668320968482 [FREE Full text] [doi: 10.1177/2055668320968482] [Medline: 33343921]
74. Xie B, Tao C, Li J, Hilsabeck RC, Aguirre A. Artificial intelligence for caregivers of persons with Alzheimer's disease and related dementias: systematic literature review. JMIR Med Inform 2020 Aug 20;8(8):e18189 [FREE Full text] [doi: 10.2196/18189] [Medline: 32663146]
75. Nicholls SG, Carroll K, Nix HP, Li F, Hey SP, Mitchell SL, et al. Ethical considerations within pragmatic randomized controlled trials in dementia: results from a literature survey. Alzheimers Dement (N Y) 2022 May 02;8(1):e12287 [FREE Full text] [doi: 10.1002/trc2.12287] [Medline: 35509502]
76. Benson C, Friz A, Mullen S, Block L, Gilmore-Bykovskyi A. Ethical and methodological considerations for evaluating participant views on Alzheimer's and dementia research. J Empir Res Hum Res Ethics 2021;16(1-2):88-104 [FREE Full text] [doi: 10.1177/1556264620974898] [Medline: 33238781]
77. Evans RS. Electronic health records: then, now, and in the future. Yearb Med Inform 2016 May 20;Suppl 1(Suppl 1):S48-S61 [FREE Full text] [doi: 10.15265/IYS-2016-s006] [Medline: 27199197]
78. Gordon WJ, Patel V, Thornhill W, Bates DW, Landman A. Characteristics of patients using patient-facing application programming interface technology at a US health care system. JAMA Netw Open 2020 Oct 01;3(10):e2022408 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.22408] [Medline: 33107918]
79. Pais S, Parry D, Huang Y. Suitability of fast healthcare interoperability resources (FHIR) for wellness data. In: Proceedings of the 50th Hawaii International Conference on System Sciences. 2017 Presented at: HICSS 2017; January 4-7, 2017; Hilton Waikoloa Village, HI. [doi: 10.24251/hicss.2017.423]
80. Hampel H, Au R, Mattke S, van der Flier WM, Aisen P, Apostolova L, et al. Designing the next-generation clinical care pathway for Alzheimer's disease. Nat Aging 2022 Aug;2(8):692-703 [FREE Full text] [doi: 10.1038/s43587-022-00269-x] [Medline: 37118137]
81. Saho K, Uemura K, Sugano K, Matsumoto M. Using micro-doppler radar to measure gait features associated with cognitive functions in elderly adults. IEEE Access 2019 Feb 20;7:24122-24131. [doi: 10.1109/access.2019.2900303]

XSL•FO
RenderX

82. Sarvamangala DR, Kulkarni RV. Convolutional neural networks in medical image understanding: a survey. Evol Intell 2022;15(1):1-22 [FREE Full text] [doi: 10.1007/s12065-020-00540-3] [Medline: 33425040]
83. Lua EK, Crowcroft J, Pias M, Sharma R, Lim S. A survey and comparison of peer-to-peer overlay network schemes. IEEE Commun Surv Tutor 2005 Jun 30;7(2):72-93 [FREE Full text] [doi: 10.1109/comst.2005.1610546]
84. Suciu G, Stefanescu S, Beceanu C, Ceaparu M. WebRTC role in real-time communication and video conferencing. In: Proceedings of the Global Internet of Things Summit. 2020 Presented at: GIoTS 2020; June 3, 2020; Dublin, Ireland. [doi: 10.1109/giots49054.2020.9119656]

Abbreviations

AD: Alzheimer disease
ADRD: Alzheimer disease and related dementias
AI: artificial intelligence
BPSD: behavioral and psychological symptoms of dementia
EHR: electronic health record
GDPR: General Data Protection Regulation
HCP: health care professional
HIPAA: Health Insurance Portability and Accountability Act
IoMT: Internet of Medical Things
IoT: Internet of Things
MCI: mild cognitive impairment
mHealth: mobile health
P2P: peer to peer
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RQ: research question
WebRTC: Web Real-Time Communication

Edited by H LaMonica; submitted 23.11.24; peer-reviewed by P Wu, D Viderman, M OldeRikkert, S Groeneveld; comments to author 13.02.25; revised version received 24.02.25; accepted 25.02.25; published 14.05.25.

Please cite as:
Shaik MA, Anik FI, Hasan MM, Chakravarty S, Ramos MD, Rahman MA, Ahamed SI, Sakib N
Advancing Remote Monitoring for Patients With Alzheimer Disease and Related Dementias: Systematic Review
JMIR Aging 2025;8:e69175
URL: https://aging.jmir.org/2025/1/e69175
doi:10.2196/69175
PMID:

©Mohmmad Arif Shaik, Fahim Islam Anik, Md. Mehedi Hasan, Sumit Chakravarty, Mary Dioise Ramos, Mohammad Ashiqur Rahman, Sheikh Iqbal Ahamed, Nazmus Sakib. Originally published in JMIR Aging (https://aging.jmir.org), 14.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Enhancing Enrollment and Adherence in Long-Term Wearable Research on Dementia: Qualitative Systematic Review and Meta-Synthesis

Colleen M Peterson[1], PhD; Renée M St Louis[1], PhD; Carol Flannagan[1], PhD

University of Michigan Transportation Research Institute, Ann Arbor, MI, United States

Corresponding Author:
Colleen M Peterson, PhD
University of Michigan Transportation Research Institute
2901 Baxter Road
Ann Arbor, MI, 48109
United States
Phone: 1 734 489 1709
Email: cmpete@umich.edu

Abstract

Background: With the rapid expansion of wearable technologies, there is increased interest in their utility for passive data collection applications in research on aging. Wearables can be beneficial for research with people with dementia and their families, who have burdens that can make both study participation and reliable data collection more difficult, especially as dementia progresses, but their use also has challenges. Population-specific issues affecting the success of wearables for data collection can include remembering to wear a device, fluctuating acceptance of the device or study participation, and reliance on already strained caregivers.

Objective: This study aimed to systematically evaluate contemporary wearables research to describe persons with dementia's experiences with wearables, their desired qualities, and protocol needs to enhance participant buy-in and sustained wearing for better quality dementia research.

Methods: We used the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 checklist for systematic reviews and searched 3 scholarly databases using Medical Subject Headings (MeSH) terms for papers published since 2018 featuring the use or discussion of wearable devices for persons with dementia. We screened 1757 abstracts and retained 58 for full-text review.

Results: We present synthesized preferences, barriers, and facilitators to buy-in and adherence to wearables in dementia research. A total of 29 factors were categorized into 4 overarching categories aligned with study development: device selection, protocol considerations, enhancing recruitment, and promoting adherence.

Conclusions: These findings inform researcher guidelines for wearable device selection and protocol design to enhance the utility of wearables in future longitudinal research featuring persons with dementia and their caregivers.

(JMIR Aging 2025;8:e63768) doi:10.2196/63768

KEYWORDS

dementia; dementia care; Alzheimer disease; caregivers; cognitive impairment; wearable electronic devices; systematic review

Introduction

Background

The growing prevalence of dementia and its substantial economic and social impacts have led to increased attention and funding allocations for research supporting persons living with dementia and their caregivers as well as dementia prevention research [1]. Obtaining the perspective of persons living with dementia and their caregivers is critical to advancing our understanding of dementia to not only find effective treatments but also to address burdens often associated with providing care to this population [2,3]. Engaging persons living with dementia and their caregivers in research studies can be challenging for many reasons, including the fluctuating cognitive capacities of individuals with dementia, the logistical and emotional demands placed on caregivers, and the need for research protocols that accommodate the specific needs and

XSL•FO
RenderX

limitations of this population [3]. Innovative methods to increase meaningful engagement in research studies involving persons living with dementia and their caregivers are warranted.

Meeting Researcher Needs

Wearable technology is a noninvasive, passive way to collect a variety of data over long periods and offers the potential benefit of enhancing care and prevention efforts [4-6]. Remote data collection, such as that offered by wearable technology, has many potential benefits, including reducing recall bias and decreasing the burden placed on participants owing to frequent contact for research purposes. Identifying wearables and research protocols that meet both researcher and participant needs will help wearable technology be more effectively used in longitudinal dementia studies [7]. As the wearable technology market is expected to continue growing substantially [8], guidance for researchers to optimize the selection of devices to enhance buy-in and adherence in their studies is necessary. Wearable devices and research strategies tailored to persons living with dementia and their care providers are crucial for optimizing long-term research with this population [9].

Currently, longitudinal dementia research with wearables faces 3 overarching challenges. First, the challenges particular to persons with dementia and their caregivers and health care professionals must be considered to gain long-term buy-in and use, especially as dementia progresses and caregiving burdens increase [10,11]. Population-specific issues include remembering to wear or charge the device, ethical informed consent [12], and privacy concerns [13,14] that may differ from cognitively intact older adults [15,16]. Second, there is a multitude of wearables on the market with a variety of features and hardware forms at various costs that have not been evaluated in terms of this population's usability and adherence factors. Third, wearables must also provide consistent, accurate, and useful data that meet researcher needs (eg, activity monitors, GPS, and location information) and have limited or easily resolvable technical problems, user-related issues, or data access factors [17].

To date, there has not been a systematic evaluation of wearables for research that holistically examines whether wearables meet the needs of persons with dementia, their caregivers, and researchers to enhance participant buy-in and research utility. Moreover, with the rapid expansion and advancement of wearable technology in the last few years, new research on participant experiences and desires for wearables is necessary. Recent wearables research with this population has featured focus group discussions with limited representation of persons with dementia and their caregivers [10,18], as compared to the experiences of older adults without dementia [19]. Observational research that could inform wearable acceptance and adherence strategies has often been limited to very short-term evaluation periods (typically up to 7 days) with a handful of activity monitoring devices [20-22]. Identifying factors related to wearables that are acceptable to persons living with dementia and their care providers, as well as population-specific support systems, is necessary to strengthen the quality of research requiring long-term adherence [7].

Objectives

This study aimed to conduct a systematic review of dementia-related preferences, barriers, and facilitators to using wearables in research studies to inform future study development, including device selection and participant support protocols.

Methods

Data Sources

This systematic review followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 checklist (Multimedia Appendix 1) [23]. Original research studies were identified from the PubMed, MEDLINE, Scopus, and CINAHL databases (all last consulted in January 2023) and Google Scholar. In March 2024, approximately 500 results were reviewed to capture additional references since the 2023 search. The search strategy included a combination of Medical Subject Headings (MeSH) terms and keywords. Categorized according to the PICO (population, intervention, comparison, and outcomes) framework, search terms were used in OR statements in combination for the population ("dementia," "Alzheimer* Disease," "memory loss," and "cognitive impairment") and the intervention or indicator ("wearable*," "wearable sensor*," "tracker*," "biosensor*," "Fitbit," "wearable electronic*," "remote monitoring," and "pedometer"). Regarding comparison, the inclusion criteria did not require intervention trials, whereas for outcomes, the inclusion criteria did not specify the outcomes of the studies.

The full text of the relevant articles was retrieved from the Web of Science, Google Scholar, and library access at the University of Michigan. The articles were reviewed for additional related references.

Study Selection

The selection criteria for the literature search included (1) full-text data from primary peer-reviewed journal articles, commentaries, and conference proceedings; (2) use of remote monitoring via electronic devices that captured data of persons; (3) enrollment of, or reference to, populations living with Alzheimer disease, dementias, or chronic memory impairment or loss impairment; (4) publication in English from January 2018 to March 2024; and (5) documentation of any preferences, barriers, or facilitators to wearables device use. Notably, research featuring Huntington or Parkinson disease was included if participants were experiencing dementia-type impairment and otherwise met the inclusion criteria. Studies were excluded if they (1) did not meet the inclusion criteria, (2) used an implantable wearable or a device not intended for data collection (eg, hearing aids), (3) reported only a study protocol, or (4) included no person interaction (eg, device development without consulting with the population considered affected). Dissertations, books, book chapters, other reviews, and single-case studies were excluded.

CMP searched independently, and CMP and CF reviewed titles and abstracts for relevance and subsequently conducted the full review. CMP manually searched the reference sections of these full-text articles and key reviews for additional citations. RMSL

XSL•FO
RenderX

independently reviewed the search criteria and list of full-text articles to meet the eligibility criteria. The authors used EndNote (Clarivate), a bibliographic tool that allowed them to organize PDFs of articles by topic, along with other data about each article.

Data Extraction and Analysis

Data extracted (by CMP) from the full text of the eligible studies included, when available, publication details, study location, study design, study population, device type and location worn, and device-wearing duration. For the focus of this study, we extracted information related to the preferences, barriers, and facilitators of the use, adoption, maintenance, and adherence of wearable electronic devices by persons with dementia, including ideal device qualities and protocol needs for persons living with dementia and their caregivers. This information could come from the researchers (eg, procedures and processes attempted), persons with dementia (eg, reported preferences and difficulties inhibiting ongoing use), or persons caring for those living with dementia, whether health professionals or informal family caregivers (eg, encouraging persons to wear the device).

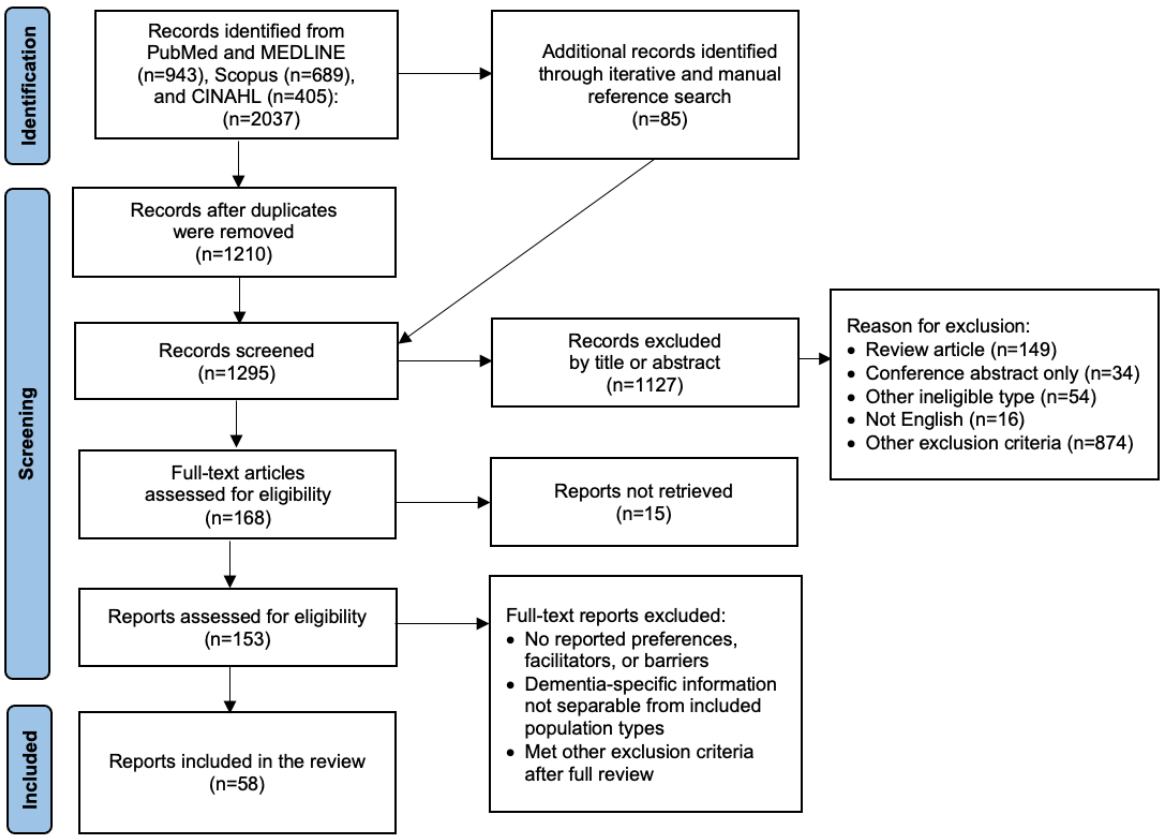
The extracted qualitative data on preferences, barriers, and facilitators were reviewed and initially synthesized for each study by CMP into a common factor coding structure using a grounded theory qualitative research approach [24]. This approach does not presuppose findings but relies on iterative "constant" comparison of collected data to develop codes and overarching themes. RMSL reviewed the summaries and factor synthesis, which CMP and RMSL then independently assigned to each study. The factor assignments were reviewed, and discrepancies were adjudicated together. We report the descriptive frequencies of the factors identified and noted in at least 5 studies and provide a qualitative synthesis of their implications for optimizing long-term dementia research.

Results

Overview

This systematic review identified 1757 studies through the databases as well as iterative and manual reference searches. After the eligibility criteria were applied, 3.3% (58/1757) of the studies were retained for data extraction analysis (Figure 1 [23]). Most studies (1127/1295, 87.03%) were excluded based on the abstract review because the device evaluated was not a wearable or the population had some cognitive impairment but did not specifically include dementia. The studies were primarily observational (eg, feasibility studies) and qualitative (eg, using interviews or focus groups), with only 19% (11/58) featuring some type of trial. Participants were most often persons living with dementia in the community (as opposed to residential care), their informal family care partners, and paid dementia care health professionals. The duration for which persons living with dementia were asked to wear the study device ranged from a few hours over several days to 12 months, though 7 days was a common wear period. Information pertaining to the individual studies, including the type of wearable, is summarized in Multimedia Appendix 2 [9,10,18,20,22,25-60].

Figure 1. PRISMA 2020 flow diagram for the systematic review of wearables in dementia research.



XSL•FO
RenderX

Barriers, facilitators, and preferences from the 58 studies were qualitatively synthesized into a total of 29 factors that fit into 4 overarching themes. These themes are categorized by issues dementia researchers should consider from study development to deployment: device selection, protocol considerations, enhancing recruitment, and promoting adherence. Although we categorized these factors by greatest relevance to a particular study section, it should be recognized that many of the factors are interconnected and related. The frequency and proportion of studies featuring each factor are detailed in Table 1. The factors present in each study and their cross-reference numbers are available in Multimedia Appendix 3 [9,10,18,20,22,25-77]. As we describe the factors in the results, we use the term caregivers inclusively to mean both informal (ie, unpaid, often family members) and paid caregivers (eg, health professionals and residential care staff) and provide relevant quotes that capture participant experiences. In the Discussion section, we offer a synthesis of recommendations for researchers planning to conduct longitudinal research with populations with dementia.

Table 1. Presence of synthesized factors across the studies (N=58).

Factor	Count, n (%)
Device selection	
Easy to use	21 (36)
Smaller size or weight	18 (31)
Comfort	15 (26)
Fits into routine	13 (22)
Unobtrusive	12 (21)
Tailorable	12 (21)
Materials	10 (17)
Aesthetics	8 (14)
Easy to wear	7 (12)
Water resistant	7 (12)
Protocol considerations	
Privacy concerns	15 (26)
Consent capacity	14 (24)
Battery	13 (22)
Adjustment period	10 (17)
Task requirements	9 (16)
Enhancing recruitment	
Technical anxiety	17 (29)
Multifunctional	16 (28)
Burden impact	16 (28)
Stigma	13 (22)
Caregiver buy-in	12 (21)
Connectivity	10 (17)
Promoting adherence	
Remote monitoring	19 (33)
Provide health insight	18 (31)
Technical support	17 (29)
Caregiver support	14 (24)
Safety	14 (24)
Remembering the device	12 (21)
Independence	10 (17)
Self-removal	7 (12)

XSL•FO
RenderX

Device Selection

Easy to Use

Given the care burden of family members or paid caregivers in residential settings, a wearable's ease of use was paramount when asking caregivers to be involved in a study. It was the most frequently identified factor in studies, with 36% (21/58) of studies mentioning it. The wearable device must be easy to use [25-27,61-65] with an intuitive interface and features [28-31]. The device should require little maintenance or technical support for everyday use, and any interaction with the device required by the study should be simple [10,28,32-35,66]. Some studies explicitly noted that perceived ease of use was associated with more positive attitudes toward wearing the device and less effort or sense of inconvenience in participation [27,36-39]. As a person living with dementia exclaimed in the study by Gris et al [38], "Older people need to use simple things!"

Smaller Size and Weight

Almost without exception, participants desired smaller and lighter wearables or expressed discontent about the size and weight of the device they were asked to wear in the study [34,35,40-42,63,64,66-68]. The influence of a wearable's size and weight was the most frequent physical form factor, with 31% (18/58) of the studies including comments about it, and it was the fourth most common factor identified overall. While rare, there were some exceptions to participants desiring smaller devices overall. Some suggested enlarging certain facets of a device, such as having a larger face or font to read [43] or larger buttons to interact with [27].

Comfort

The physical ease of wearing the device was important [28,31-33,38,41,44,67,69], with discomfort specifically cited as a reason why participants removed the device in several studies [34,45,66]. *Comfort* was sometimes explicitly linked with *unobtrusive*, as complementary qualities [20,29,61]. Nevertheless, prototype devices that were bulky or wearables with a more invasive application (eg, required taping to limbs or backs) were mostly tolerated in the short term. However, simply put, "If residents would show discomfort related to the wearable, the wearable would not be used" [61].

Fits Into Routine

Studies found that if a device type was already familiar or similar to what caregivers and persons living with dementia already used regularly, the wearable could be seen as convenient and more easily fit into a participant's lifestyle [10,30,32,33,46,67,70]. Wristwatches were the most frequently cited as meeting these criteria [20,29,35,38,41,47]. At the same time, devices might also be removed if accessory removal is also part of a participant's routine [20,32,41]: "To tackle such barriers [to using mHealth devices] will likewise require (informal) caregivers to change their attitudes towards using digital technologies and the integration of these technologies into routine care" [30].

Unobtrusive

A wearable that is not obvious or can be hidden was considered a distinct positive [25,29,35,38,70,71]. Having the ability to conceal a wearable alleviated some worry about a person living with dementia being resistant to wearing the device or removing it (refer to the Self-Removal subsection), which could be especially pertinent in studies involving individuals with more advanced stages of dementia [61,66,68]. Unobtrusive devices were also seen as reducing the potential of drawing attention to themselves or experiencing stigma associated with wearing a device [10,42,61] (refer to the Stigma subsection). Like the *fits into routine* factor, wristwatches were commonly cited as an unobtrusive solution [10,20], but something that could be slipped onto a belt or into a shoe was also noted as a viable option [10,29].

Tailorable

Having the option to use a wearable in a way that fits a person's particular style preferences was desired, from format (eg, neck pendant or watch) to style (eg, colors and design) [10,29,34,35,38,61,62,71], as some had concerns with the particular format that was provided in the study and desired a different way to wear the device [26,31,68]. Alternative device forms can include a spectrum of participants and their day-to-day presentations (eg, gender and culture) [10,18]. Offering a wearable in several different forms could help a device suit the *fits into routine* factor, but the *tailorable* factor highlights that singular forms may not fit into everyone's routine equally: "Several members in all four groups expressed a desire for more options in device styles...all groups asked if alternative styles could include a broach, belt clip, small bracelet, or ankle bracelet" [31].

Materials

Nonspecific material concerns or durability were cited as important considerations for wearables to be satisfactory to participants and researchers alike [29,41,63]. Specific material issues such as softness [34,63,66] and the need to make wearables safe for those with skin sensitivities or allergies were also mentioned [10,28,34,61]. Some participants stopped wearing the device altogether because it caused irritation [22] or even skin abrasions [48].

Aesthetics

The general appearance of the device design in its look and feel could be a barrier to everyday use and, thus, adherence [41,43,47,63,64]. Size, form, color, and device material all play a role in the aesthetic attractiveness of a wearable [31,34,35]. For example, participants in the study conducted by Jacklin et al [31] were dissatisfied with the prototype wearable they were testing, describing it as "big and clumsy looking" and "rubbery."

Easy to Wear

Diminishing fine motor skills in persons living with dementia made *easy to wear*, or how simply one can put on or take off the wearable, an issue for day-to-day use [10,62,65]. Wearables that were difficult for the wearer to put on or take off by themselves required caregiver support that could add stress to both the person living with dementia and their caregiver,

XSL•FO
RenderX

creating participation burden and a sense of lack of independence [34,48]. Some wearables were slightly more complicated and required specific sensor placement, which was either difficult to apply correctly or failed to remain securely attached [34,49,50].

The factor *easy to wear* is distinct from *easy to use* and may not play a role if a wearable is designed to be worn continuously. For example, a family caregiver reported that an accelerometer was "a little bit difficult to put on, so [the persons living with dementia] did do it but...it's not easy." The researchers echoed that "the wristband of the accelerometers could be difficult to fasten" but reported that afterward, "the accelerometers were not burdensome to wear" [65].

Water Resistant

Water-resistant or waterproof wearables were recommended [28,61,64] to minimize potential disruptions to data collection. Along with charging the device, a common point of noncompliance occurred when participants forgot to put the wearable back on after removing it for bathing (refer to the Remembering the Device subsection) [29,34]. By selecting a waterproof device and asking participants not to remove it, study protocols do not have to rely on persons living with dementia or their caregivers remembering to take off the device during bathing and, therefore, can reduce the number of study *task requirements* and related *burden impact* [44,63]. Cruz-Sandoval et al [29] suggested that minimizing the need to remove a device can reduce instances of "resistance" from persons living with dementia, and therefore, waterproof devices "can ease some of the burden on caregivers" as well.

Protocol Considerations

Privacy Concerns

Many studies noted that privacy was an issue, with participants voicing concerns about data security and who had access to the wearable's data [10,18,26,31,38,63,68,71]. For example, a caregiver's spouse worried about third-party breaches of her wearable data, explaining that "privacy and security are important to his wife because '...she understands just how easily those systems can be hacked'" [71].

At the same time, some persons living with dementia and their caregivers stated explicitly that privacy concerns were outweighed by the personal or societal benefits of their study participation [25,33], or they reported no privacy concerns [27]. Participants understood that data sharing was often necessary to meet the goals of the study, and privacy concerns were ameliorated when researchers established trust around data management: "Both informal caregivers and the nursing staff were not concerned about privacy issues regarding the wearable, as they consider biomedical variables as de-individualized values. The participants trusted the researchers and nursing home regarding good data handling, as they felt well-informed" [61].

Researchers were able to build trust and attenuate privacy concerns through transparency about data access and security measures during the informed consent processes and throughout the study [61,72]. Empowering participants to have control over privacy options (eg, allowing the collection of some but not all data) and offering the ability to turn off or select what data to monitor also helped [51,73].

Consent Capacity

The ability of persons living with dementia to understand and actively affirm their study participation is an ethical imperative but was specifically mentioned as a potential difficulty by several studies [38,71]. Especially in longer-term studies, seeking ongoing consent to participate was highlighted as necessary, with suggestions to check in regularly to assess the cognitive capacity of the persons living with dementia and their willingness to participate [20,51,74]: "Fluctuating cognition can be a significant barrier to the remote monitoring of isolated patients as forgetting about one's motivations to be monitored can lead to anxiety and agitation for being watched and, consequently, the disablement or destruction of equipment" [74].

At odds with fully informed consent was the general awareness of persons living with dementia regarding their participation or the device [32,43]. At this level of impairment, consent from a caregiver or another proxy for the persons living with dementia was sometimes sought instead, which could be an appropriate barrier to inclusion [22,69,70]. Another major challenge is reconciling what to do when a person living with dementia desires conflict with the caregiver's desire for them to engage in the study and wear the device [10,68]. Under consideration is the issue of dignity, respect for autonomy, and how ignoring these or not obtaining assent could lead to agitation: "In one case, [the study staff] visited the participant after he had refused to wear it on his wrist, saying his wife was tagging him 'like a dog'" [68]. In another study [10], a caregiver reported "[My husband] would have never given his consent even at the beginning—he liked the independence of making his own decisions and doing what he wanted. And I would have had to do it without his consent."

It was recommended that, when possible, researchers, caregivers of persons living with dementia, and their proxies should attend to their care recipient's wishes, whether from an explicit advanced directive [73] or knowledge of their comfort with similar technology [30].

Battery

The longevity of battery life and the ease of charging were noted as concerns for both researchers and participants [9,28,38,63,64,75]. Wearables with a battery life of at least several days were ideal [28,29,52,63], decreasing the need for participants to remember to charge the device [9,10,72] (ie, simplifying *task requirements*), which also minimized data collection interruptions for researchers. Low battery warnings from the device or study staff were seen as a useful tool to ensure an adequate charge of the battery [10,32,68]. Necessarily, the burden of charging was exacerbated when a caregiver was not on site to physically assist or when the charging mechanism was difficult to implement [35].

XSL•FO
RenderX

Adjustment Period

Participants may need time to acclimate to wearing the device (eg, establish wearing routines) and to resolve technical issues before consistent adherence and, therefore, consistent data quality, as adherence tended to improve after an initial period [9,29,32,33,41,53,61]. On the other hand, some studies found better compliance earlier in a study period [42,54,65], perhaps because of novelty. Adjustment periods featuring cognitive aids such as instructional pamphlets and study staff check-ins provided early troubleshooting and time to build new routines [9,29,33].

Task Requirements

Cognitive impairment and caregiving burden were barriers to protocol adherence when participants were asked to do more than simply wear the device [55,56,74,75]. When a study protocol required additional activity from either the persons living with dementia or a caregiver, such as completing bedtime and wake-up event markers, they were less likely to remember or be able to comply [9,50,57,72].

Enhancing Recruitment

Technical Anxiety

Although many interacted with the study wearables and related technology platforms successfully, participants' sense of a lack of technical proficiency and negative attitudes about technology were reported as a perceived barrier to enrollment [10,18,27,37,38,41,71] and adherence during the study [29,53,62]. Caregivers and persons living with dementia alike expressed concern about the complexity or technical know-how needed to operate the wearables [29,36,53,63,74]. Technical anxiety could present as worrying about damaging or losing the device [20,29,34]. The level of technical anxiety and attitudes toward the technology were related to age qualitatively and statistically [10,27,38,53,62] and previous experience with devices [36,46,76]. Participants in a study actually rated themselves as less technologically proficient at the end of the study because of difficulty with the wearable [42].

Multifunctional

Participants saw the potential for the device and platform to offer more information or other uses to the wearer as a benefit of being involved in the research [27,71]. Being able to track their health and activity, signal health events or emergencies (eg, wandering or fall alerts), provide directions, or offer location services, among others, were cited as desirable features [10,25,27,29-31,34,63,64,67,74]. Simpler functions such as information about time, date, and alarm capacity were also ideal wearable facets [10,32,35,41,52,63]: "Participants with dementia indicated that multiple purposes for the locator device, such as being able to measure heart rate, blood pressure and activity level, is advantageous as it would streamline the number of devices that they and their caregivers must contend with" [10]. Stavropoulos et al [63] reported, "Feedback suggested that some people with dementia may be more willing to accept technology that supports them in their daily functioning, in addition to assessing it."

Burden Impact

Caregivers especially liked the potential the wearables offered in terms of supporting their caregiving role and tasks, which, in turn, reduced related burden and stress [27,32,39,64,69]. Factors such as a wearable's capacity for *remote monitoring* and *providing health insights* could improve quality of life or provide peace of mind [33,39,63,67,72]. However, when the wearable was technically difficult to use (*easy to use)* or researchers required additional time or study activity (refer to the Task Requirements subsection), being involved in the study was perceived as a cause of additional burden and a deterrent to both enrollment and continued participation [29,32,33,58,59,61,70,74]. Notably, participants in several studies reported experiencing a mix of both reduced and heightened burden, which varied according to specific facets of the research or from using the wearable [32,33,65,69]: "The nursing staff indicated that they did not want to spend valuable care-time if the foreseen benefit of the wearable for the specific resident was not evident to them. Likewise, also informal caregivers found it important to know the potential individual benefits, before deciding whether to bother the resident with the wearable, as they anticipated some level of resistance to wearing the wristband by the residents" [61].

Stigma

Caregivers and persons living with dementia expressed concerns about the stigma associated with the wearable and the wearable identifying their care recipient (or themselves) as affected by the hidden disability [10,18,30,64]. There was concern about social acceptance and the visibility of the device, as well as the perceptions or comments that it might engender (eg, that they are under surveillance or need to be monitored) [33,34,61,72]. On the other hand, "social influence" was often a positive, with care providers or fellow persons living with dementia offering positive comments or encouraging the use of the wearable [29,37,39,76].

Families and professional care providers alike were sensitive to the negative connotation that some wearables might cause. An Alzheimer Society staff member reported that some families might avoid incorporating a MedicAlert bracelet in their care management plans because "It's a signifier or something that's perceived to bring stigma—[the persons living with dementia] don't want to be out there clearly marked in some way, that there's something wrong with them. Nor does the family" [18].

Caregiver Buy-In

Caregiver buy-in underlies many facets of wearable studies featuring persons living with dementia, as many studies relied on caregiver commitment and coordination with the researchers to do study activities (refer to the Caregiver Support subsection) [20,36,57,61,70]. Moreover, they played significant roles in facilitating understanding of the study by persons living with dementia [29,47] and in encouraging them to wear the device [39,41]. Some studies specifically cited the need for caregivers to have perceived usefulness or reasons to commit to reducing perceived potential burden or risks to enroll themselves and their care recipient in the study [61,72,76]. Gaining caregiver buy-in could be enhanced by reducing participation burdens

XSL•FO
RenderX

(eg, fewer *task requirements*, *easy-to-use* devices, and offering ongoing *technical support*) and experiencing personal benefit from the study (eg, *remote monitoring* and *providing health insights*).

Connectivity

Passive, remote monitoring is the defining advantage of wearable device studies. However, technical connectivity requirements can be a barrier to enrollment [10,31,61,69,74] and may create logistical problems for data transmission [18,31,56,77]. For example, requiring faster Wi-Fi or consistent Bluetooth linkage with a smartphone for more frequent data uploads [67,75,77] may introduce recruitment bias to individuals of higher socioeconomic status and those who are already technically inclined: "The lack of internet connection in patient homes and of funding for caregivers and family member to purchase assistive or communication devices are frequently coupled with skepticism or low abilities to engage with digital products" [74].

Promoting Adherence

Remote Monitoring

Remote monitoring was the most frequently cited desirable and useful feature (mentioned in 19/58, 33% of studies) when engaging in a wearable study among populations with dementia [10,25,27,34,59,61,69,76]. Geolocation-based remote monitoring enabled caregivers to track the movements of persons living with dementia when they were not present, including movement outside the home. Remote monitoring was also seen as a way to know about the well-being of persons living with dementia, for example, by the identification of agitated movement or interrupted sleep [32,40,77]. Remote monitoring enhanced the *independence* of the person living with dementia [27,38,63] while providing informal caregivers with peace of mind regarding the safety of persons living with dementia (eg, reduces *burden impact*) [46,63,64,67,72].

For some participants, the utility of the benefits of geolocation services conclusively outweighed potential *privacy concerns* [25,42]. However, for some, introducing the idea of being monitored could be seen as an infringement on autonomy and privacy [25,31], such as when "some participants feared that proud Anishinaabe older adults would be offended by the idea that their movements would be tracked" [31].

Provide Health Insight

Providing health insight to the participants was one of the top factors mentioned as a benefit to ongoing participation and use of a study's wearable. Paralleling the *multifunctional* factor as an aid to enrollment incentives, health insights included, for example, information on blood pressure, heart rate, sleep quality, and steps. Having ongoing access to this health data motivated participants' personal activity, helped caregivers monitor for health crises, and provided more general insight into the well-being of persons living with dementia [10,25,27,29,32,33,38,53,54,58,61,63,69,74,77]. In addition, they appreciated that this health information could be useful to residential staff and physicians in informing care decisions [25,27,32,34,61,64,69]. A caregiver in the study by Lazarou et al [25] said that with the information from the wearable, "we could monitor her daily activity and communicate any problems to the doctors. It helped us a lot to understand what is going on."

Technical Support

As *technical anxiety* was a barrier, providing initial technical training and ongoing technical support was identified as a complementary facilitator [20,26,37-39,50,74]. Staff availability to provide training at enrollment, support during an adjustment period, and assistance as needed throughout the study helped improve adherence [29,33,77]. Stand-alone instructional materials and cognitive aids were useful to participants who sought to troubleshoot issues on their own [9,10,47,66]. Along with data monitoring, regular proactive check-ins with participants made (or would have made) staff aware of technical problems and helped maintain better compliance [29,45,53,68,77].

For example, Gelonch et al [33] provided caregivers with initial training and ongoing technical assistance, and consequently, they supported their care recipient's use of a lifelogging camera: "The [persons living with dementia] said that the support and supervision provided by their caregiver in certain tasks or actions, such as connecting and loading the camera, were also key for them continuing with the experiment."

Even when apprehensive about using the technology (refer to the Technical Anxiety subsection), participants commonly showed a readiness to learn how to operate the wearable devices and follow study protocol. As recounted by Anderson et al [36], "Many were not necessarily technologically proficient caregivers, but they wanted to be useful. None of the caregivers implied that the research should not continue to be refined."

Caregiver Support

While the importance of *caregiver buy-in* was highlighted in recruitment, in this section, we underscore the importance of caregivers and their investment in the study for continued wearables adherence [36,39,71,77]. Many research protocols required caregiver activity to accomplish their goals by having them be the primary study participant or helping persons living with dementia complete study tasks [29,44] and, importantly, charging the wearable [35,77]. Many factors depend on the caregiver, such as helping persons living with dementia *remember the device* [20,41], assisting with putting it on (refer to the Easy to Wear subsection) [34,48], or understanding how to operate the wearable [33]. In addition, caregivers could be instrumental in explaining facets of the study and, in other ways, facilitating communication of research needs to their care recipient throughout the study [29,68].

Notably, relying on caregiver support can be a major ask (ie, increase *burden impact*): "The caregivers were able to be present almost all the time. They demonstrated personal investment in the care that the person living with dementia received. Their willingness to participate in, work with researchers, and use technology 24 hours a day for 30-60 days was a significant commitment" [36].

XSL•FO
RenderX

Safety

Broadly speaking, wearable technologies were seen as an instrument for enhancing the safety of persons living with dementia [45,46,71,72]. This potential positive was specifically identified as an ongoing benefit, often in tandem with or implied within *remote monitoring* and *providing health insight*. Wearables offer a sense of security to persons living with dementia and their care providers, aiding aging in place by alerting caregivers to life events or potential crises (eg, elopement or fall monitor) and detecting changes in health or behavioral patterns [28,31,34,59].

On rare occasions, the device itself was harmful, such as when an over-ear device caused skin abrasions [48] or when participants desired a different product or sensor design for safety reasons [69]. In contrast to other participants in their study, a caregiver in the study by Snyder et al [71] explained that they felt the wearable, an emergency response pendant, gave a false sense of security, saying, "My reluctance is that it doesn't seem...It can't stop something from happening to her. It just alerts you maybe if something does."

Remembering the Device

Adherence in wearable studies was affected by persons living with dementia's ongoing ability to remember what the device was for [22,29,42]. Not remembering the purpose of the device could also result in "confusion and agitation" [43]. More commonly, persons living with dementia simply forgot to wear the device [10,20,63,75], especially after removing it for bathing (if not *water resistant*) or sleep, often resulting in failure to put it back on [34,65]. Better adherence was reported earlier in the day [53,77], potentially because of sundowning: "Due to memory impairment participants had to be constantly reminded to wear the device and they could change their opinion from one day to the other. Activities that made wearing the device more salient, such as the need to take it off to charge it or bathing, proved challenging with some participants" [20].

Independence

Whether the device provided geolocating, physical activity monitoring, or health alerts, participants, especially caregivers, saw these devices and platform features as enhancing the independence and autonomy of persons living with dementia [10,27,39,63,67,71]. Participants' ongoing access to the data and ability to turn off data collection at their discretion contributed to a sense of agency within the study itself [33,38]. Similar to the potential harms of *remote monitoring*, Synder et al [71] framed this positive as part of the "conundrum between the benefits of having peace of mind, a sense of self-efficacy, and greater independence versus perceiving (remote monitoring technologies) as an invasion of privacy, a security risk, a false sense of security, and a distraction."

Self-Removal

Taking off the device when not instructed was an issue in several studies [44,60]. This could be temporary, such as during bathing, or routine, such as removing accessories before sleep [32,72]. In some cases, removal was longer term, often due to agitation (possibly related to sundowning) or discomfort [22,32,66]. Some researchers and care providers suggested or chose to put the wearable out of reach (eg, on their back) to prevent self-removal by persons living with dementia [60,61].

Discussion

Principal Findings

Overview

This review identified barriers, facilitators, and preferences that could improve the quality of research studies featuring wearable devices and persons living with dementia. The review included studies encompassing a variety of study designs and perspectives from researchers, persons living with dementia, and their caregivers (both unpaid family members and paid professionals), ensuring broad inclusion of research experiences. We offer a qualitative synthesis of these findings to enhance enrollment and adherence in long-term wearable research featuring persons living with dementia. Recommendations for future research fall within 3 main categories: device selection, ethical study design, and technical and data management plans.

Device Selection

Dementia Impairment Level Requires Different Devices

We identified the importance of selecting multifunctional devices with the capacity for participants to monitor their own or their care recipient's data, whether on the device itself or via a dedicated platform. However, what persons living with dementia and their caregivers preferred in terms of devices was sometimes different depending on the advancement of dementia in their target population. In earlier stages, persons living with dementia themselves were more engaged and desired more interactive, multifunctional devices that provided personal benefit to them. However, persons living with more severe dementia, from their perspective and their caregivers' perspectives, desired or interacted with only simpler wearable form factor and features (eg, just used the clock feature) [77]. Even less obtrusive devices or ones out of reach may also be warranted if they are more likely to remove the device themselves as dementia progresses.

The growing ubiquity of remote monitoring and wearable devices for everyday use likely makes stigma less of a concern as we move into the future because an older adult will not stand out when wearing something such as that, especially if it is unobtrusive or looks like an everyday wristwatch. The availability of the devices likely also drives expectations of what participants will personally gain from being in a research study, as people are more familiar with what type of information gathering is theoretically possible, and participants may expect to have access to that information too.

Wristwatches Fits Into Routine but Could Also Be Problematic

Wristwatch wearables were highly recommended as they are part of normal wear for many and offer different levels of multifunctionality. In addition, their unobtrusiveness can reduce stigma [10]. However, habitual device removal during sleep and bathing often conflicted with the research protocols requiring participants to wear the devices all the time [32,41]. Waterproof forms, along with clear instructions for the caregiver

XSL•FO
RenderX

and persons living with dementia to avoid removal in these situations, are useful. Researchers should understand that adjustment periods may be warranted with proactive reminders as participants acclimate to the newer routine and have fewer persistent issues [41].

Ethical Study Design

Attend to Fluctuating Consent Capacity in Persons Living With Dementia

Persons living with dementia feel actively engaged in the research and empowered over their health status when they can see their health data and decide when to engage in the study. However, progressive impairment can mean a person living with dementia may experience fluctuating ability to consent and understand the protocol for a long-term study. Researchers in longer-term studies should establish ongoing communication to monitor possible changes in device preferences or needs. Advanced cognitive decline from dementia can mean decision-making and consent ability are totally compromised for the person living with dementia. Reliance on a caregiver or another surrogate for consent is common, but studies emphasized that if the person living with dementia has given their thoughts on similar situations before, those preferences should be prioritized. Respecting participants' agency is paramount; recognizing and responding to the issues of ongoing consent in this population is key to conducting an ethical study. Actively checking in with participants on a regular basis in long-term studies to gain assent or consent for continued data collection is ethical and may also be required by institutional review boards, given expected cognitive changes.

Promote Inclusivity With Wearable and Protocol Choices

Researchers should be aware of how their study protocol and wearable device requirements may impact who can or would enroll in their study. Some studies, such as that by Cohen et al [53], suggested screening participants for technical knowledge and ability as a way of ensuring compliant participants are enrolled. However, this is not inclusive and could result in bias in recruitment. Instead, preexisting technology and connectivity needs (eg, stable Wi-Fi) should be minimized or eliminated to enhance distributive justice [73,78] and increase the representativeness of the study results. Researchers can reduce the burden and increase equity by not requiring constant wireless connectivity or by supplying the technology to participants. However, intermittent data collection may be at odds with proactive data monitoring for troubleshooting purposes. Staff phone calls or other lower-technology check-ins are recommended in this case.

If possible, different formats and styles (eg, watch vs pendant and watchband options) that are familiar and fit a participant's personal, cultural, and gender preferences should be offered [79]. However, it is important to balance these preferences with the type and quality of data that each format can reliably collect [34]. As evidenced by this review and other research efforts, user-centered codevelopment of wearables with persons living with dementia for research purposes is feasible [38,80] and should continue to be explored to enhance recruitment and adherence in longitudinal studies featuring persons living with dementia.

Data Management

Implement Data Management Plans

We categorized a total of 29 factors into 4 categories that aligned with study development (device selection, protocol considerations, enhancing recruitment, and promoting adherence). However, a fifth category—data management—was largely left uncommented on, as studies were primarily focused on the needs and preferences of the participants rather than the researchers. Participants may report ongoing adherence, but making sure a wearable is not only worn but also collecting data is paramount to conducting quality longitudinal research.

Researchers' data management plans should comprise data collection, storage, and processing for analysis [29,73]. For example, researchers conducting long-term studies should know how long their device batteries will last under their protocol and have procedures in place to limit both data loss and participant burden (eg, reminders to participants to charge their devices). By regularly checking data for completeness, researchers can identify technical problems and offer troubleshooting to resolve missing data issues. Researchers should also have a loss prevention plan for the wearable to reduce missing or inactive devices and the resulting data loss.

For example, Cruz-Sandoval et al [29] had a comprehensive data management plan for a wearable study in an assisted living setting. They combined weekly program visits with battery charging, data collection (ie, upload and synchronization with the study server), and addressed technical troubleshooting needs. They also reviewed the uploaded data for abnormal data patterns and double-checked with caregivers "as soon as possible" to understand if the abnormal data were valid or if they should flag it for deletion. As a loss prevention plan, the researchers also asked the staff to report on the location of the device at the end of each work shift. Particular data management plans could also influence device options; Muurling et al [81] recommended that real-time participant access to device data could both increase adherence and help alert researchers to data issues in close to real time.

Collective Data Management Efforts Could Enhance Support for Wearables Research

To reduce researcher data burden in individual studies, research groups could build infrastructure and codebases to share collectively in an open-access format (eg, on GitHub). Although researcher perspectives on the use of the devices and subsequent data management were generally missing in the studies, 1 study reported that their experience organizing, cleaning, and analyzing Fitbit data were "onerous and time consuming" and recommended that others use automated data management software [77]. Researchers should be aware that many commercial, proprietary wearables may not offer researchers access to the raw data, which may limit secondary utility and increase reliance on the original developers for measure validation. Other collective data enhancements include setting standard definitions or ideal compliance goals, which varied across studies or were not reported at all. Clear, agreed-upon

XSL•FO
RenderX

definitions for adherence and reporting standards are critical for improving long-term dementia research with wearable devices [9,47]. Indeed, a lack of specific, consistent adherence and attrition data from many studies made quantitative comparisons related to the factors impossible. When combined with real-time data monitoring, shared data expertise can help researchers determine whether seemingly incoherent data result from technical issues, allowing for timely resolution [81], or from changes in cognitive function and life circumstances as dementia progresses [82].

Informing Wearable Technology Adoption and Use Models

This systematic review evaluated preferences, barriers, and facilitators to wearables in a study setting, so everyday technology adoption measures may not be entirely applicable. Indeed, caregivers and persons living with dementia may endure adjustment periods, committing themselves to using difficult technologies if the knowledge gained will advance dementia prevention efforts or benefit other caregivers [82]. However, because of the lack of consistent quantitative data, we could not assess factor relationships quantitatively with wearable technology adoption and use models such as the unified theory of acceptance and use of technology model [83], its extended version—the unified theory of acceptance and use of technology 2 model [37], the wearable acceptance in health care model [84], or the technology acceptance model 3 [85]. These measures may also need to be updated as overall perceptions of technology change and its use becomes more ubiquitous, even among older adults. Surveys and other marketing research indicate that although fewer older adults use wearables compared to other age demographics, approximately 25% of US adults aged ≥65 were using wearables in 2023 [86], a rapid increase even in just the last few years [87]. Moreover, these measures relate to technology acceptance in general, while motivations underlying intentions and actual use of wearables may be different, especially as noted within research contexts. A research-specific technology acceptance questionnaire could help researchers understand perceptions directly related to their needs and identify additional barriers to recruitment and adherence.

There is also utility in a more systematic assessment of wearable device qualities themselves and how they relate to adherence. The use of measures such as the Quebec User Evaluation of Satisfaction With Assistive Technology (QUEST; version 2.0) can simplify device comparisons and inform future wearable choices or commercial development. Future reviews could also benefit from standard reporting of why particular devices were chosen in original studies.

Clinical Implications

We have reviewed and identified factors related to the use of wearables in long-term dementia research. However, wearables are also a potentially viable way to gather data for clinical decision-making because they can offer a way for clinicians to proactively monitor health data and circumvent some of the challenges around gathering a patient history from someone with dementia [64,88]. Supporting this potential, a large cohort study of cognitively healthy participants showed that changes in long-term 24-hour activity patterns were associated with elevated risk for the onset of Alzheimer disease, Parkinson disease, and cognitive decline [89]. Moreover, as wearable devices are being used more by the general population, clinicians are experiencing patients coming in with questions about what they have learned from their own wearable devices. Improving research in this area can help clinicians think about how they can interpret and use wearable data in a meaningful way (eg, personalized medicine). Clinicians could use the results of this paper to guide a conversation with their patients about selecting a wearable that is beneficial for meeting the needs of both the patient and their health care team, which could also enhance adherence.

Limitations and Strengths

Our findings are generally agnostic of cognitive impairment status or degree of dementia, reflecting the participant mix seen in the studies, which also often did not disambiguate by the severity or type of dementia (eg, Alzheimer disease and Lewy body disease), which may influence findings as well as the generalizability of findings. Because our analysis was not limited to specific device types or study goals (eg, physical activity or dementia prevention), not all factors may be applicable to all future research applications. We did not perform a formal risk of bias assessment because of the heterogeneity of the study types included, and we note that many studies may be considered at high risk of bias because of smaller sample sizes. However, considering the outcomes of interest for this study were qualitative in nature, the samples were considered sufficient.

Because our extraction captured all reported barriers and facilitators regardless of their origin, we identified the holistic perspective of researchers, persons living with dementia, and their care providers, including unpaid family members or paid health professionals. We systematically coded barriers, facilitators, and perspectives using a grounded, qualitative approach, which did not presuppose expected outcomes, allowing for new or unexpected insights. We offer a comprehensive synthesis of multiple types of studies featuring many wearables and outcomes of interest, identifying overarching themes core to long-term wearable research featuring populations with dementia.

Conclusions

We described factors critical to enhancing dementia study recruitment and wearable adherence, emphasizing how those factors should be considered collectively in an ethical context to meet both researcher needs and participant desires. Organized by study design order and literature prevalence, dementia researchers can use the findings of this systematic review to thoughtfully design quality, long-term wearable studies. The incorporation of wearable technology in dementia research presents a promising pathway to surmount traditional challenges associated with studying persons living with dementia and their caregivers. By leveraging these noninvasive tools, researchers can obtain continuous, objective data for use in advancing treatment, care, and prevention efforts. The results of this evaluation will facilitate quality improvement in ongoing research efforts and contribute to a more person-centered approach to developing effective interventions and prevention strategies.

XSL•FO
RenderX

Acknowledgments

This work was presented in part at the annual meeting of the Gerontological Society of America. This study was supported by the University of Michigan Network for Innovative Methods in Longitudinal Studies of Aging through the National Institute on Aging (grant #1 R24 AG077012-01). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institute on Aging.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA 2020 checklist.
[PDF File (Adobe PDF File), 110 KB - aging_v8i1e63768_app1.pdf]

Multimedia Appendix 2
Summary of included wearable research studies involving populations with dementia.
[DOCX File , 38 KB - aging_v8i1e63768_app2.docx]

Multimedia Appendix 3
Presence and frequency of factors across wearable research studies involving populations with dementia.
[DOCX File , 65 KB - aging_v8i1e63768_app3.docx]

References

1. GBD 2016 Dementia Collaborators. Global, regional, and national burden of Alzheimer's disease and other dementias, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol 2019 Jan;18(1):88-106 [FREE Full text] [doi: 10.1016/S1474-4422(18)30403-4] [Medline: 30497964]
2. Sallim AB, Sayampanathan AA, Cuttilan A, Ho R. Prevalence of mental health disorders among caregivers of patients with Alzheimer disease. J Am Med Dir Assoc 2015 Dec;16(12):1034-1041. [doi: 10.1016/j.jamda.2015.09.007] [Medline: 26593303]
3. Frank L, Jennings LA, Petersen RC, Majid T, Gilmore-Bykovskyi A, Schicker L, et al. Participation of persons with dementia and their caregivers in research. J Am Geriatr Soc 2021 Jul 10;69(7):1784-1792 [FREE Full text] [doi: 10.1111/jgs.17340] [Medline: 34245586]
4. Chen C, Ding S, Wang J. Digital health for aging populations. Nat Med 2023 Jul 18;29(7):1623-1630. [doi: 10.1038/s41591-023-02391-8]
5. Cullen A, Mazhar MK, Smith MD, Lithander FE, Ó Breasail M, Henderson EJ. Wearable and portable GPS solutions for monitoring mobility in dementia: a systematic review. Sensors (Basel) 2022 Apr 27;22(9):3336 [FREE Full text] [doi: 10.3390/s22093336] [Medline: 35591026]
6. Husebo BS, Heintz HL, Berge LI, Owoyemi P, Rahman AT, Vahia IV. Sensing technology to monitor behavioral and psychological symptoms and to assess treatment response in people with dementia. A systematic review. Front Pharmacol 2019 Feb 4;10:1699 [FREE Full text] [doi: 10.3389/fphar.2019.01699] [Medline: 32116687]
7. Daniels K, Bonnechère B. Harnessing digital health interventions to bridge the gap in prevention for older adults. Front Public Health 2023 Jan 8;11:1281923 [FREE Full text] [doi: 10.3389/fpubh.2023.1281923] [Medline: 38259780]
8. Hayward J, Chansin G, Zervos H. Wearable technology 2017-2027: markets, players, forecasts. IDTechEx. 2017. URL: https://www.idtechex.com/en/research-report/wearable-technology-2017-2027-markets-players-forecasts/536 [accessed 2025-03-17]
9. Guu TW, Muurling M, Khan Z, Kalafatis C, Aarsland D, ffytche D, et al. Wearable devices: underrepresentation in the ageing society. Lancet Digit Health 2023 Jun;5(6):e336-e337. [doi: 10.1016/s2589-7500(23)00069-9]
10. Neubauer N, Spenrath C, Philip S, Daum C, Liu L, Miguel-Cruz A. Identifying adoption and usability factors of locator devices for persons living with dementia. Dementia (London) 2022 Apr 29;21(3):862-881 [FREE Full text] [doi: 10.1177/14713012211065381] [Medline: 34964391]
11. Meiland F, Innes A, Mountain G, Robinson L, van der Roest H, García-Casal JA, et al. Technologies to support community-dwelling persons with dementia: a position paper on issues regarding development, usability, effectiveness and cost-effectiveness, deployment, and ethics. JMIR Rehabil Assist Technol 2017 Jan 16;4(1):e1 [FREE Full text] [doi: 10.2196/rehab.6376] [Medline: 28582262]

XSL•FO
RenderX

12. Zúñiga-Santamaría T, Vázquez-García CJ. Ethical issues in research in aging. In: García-Peña C, Pérez-Zepeda MU, Gutiérrez-Robledo LM, editors. Aging Research - Methodological Issues Second Edition. Cham, Switzerland: Springer; 2018:209-220.
13. Topfer LA. GPS locator devices for people with dementia. In: CADTH Issues in Emerging Health Technologies. Ottawa, ON: Canadian Agency for Drugs and Technologies in Health; 2016.
14. Bantry-White E. Supporting ethical use of electronic monitoring for people living with dementia: social work's role in assessment, decision-making, and review. J Gerontol Soc Work 2018 Apr 16;61(3):261-279. [doi: 10.1080/01634372.2018.1433738] [Medline: 29381128]
15. Paolillo EW, Lee SY, VandeBunte A, Djukic N, Fonseca C, Kramer JH, et al. Wearable use in an observational study among older adults: adherence, feasibility, and effects of clinicodemographic factors. Front Digit Health 2022 Jun 10;4:884208. [doi: 10.3389/fdgth.2022.884208] [Medline: 35754462]
16. Kononova A, Li L, Kamp K, Bowen M, Rikard RV, Cotten S, et al. The use of wearable activity trackers among older adults: focus group study of tracker perceptions, motivators, and barriers in the maintenance stage of behavior change. JMIR Mhealth Uhealth 2019 Apr 05;7(4):e9832 [FREE Full text] [doi: 10.2196/mhealth.9832] [Medline: 30950807]
17. Cho S, Ensari I, Weng C, Kahn MG, Natarajan K. Factors affecting the quality of person-generated wearable device data and associated challenges: rapid systematic review. JMIR Mhealth Uhealth 2021 Mar 19;9(3):e20738 [FREE Full text] [doi: 10.2196/20738] [Medline: 33739294]
18. Neubauer NA, Liu L. Influence of perspectives on user adoption of wander-management strategies. Dementia (London) 2021 Feb 12;20(2):734-758. [doi: 10.1177/1471301220911304] [Medline: 32164446]
19. Moore K, O'Shea E, Kenny L, Barton J, Tedesco S, Sica M, et al. Older adults' experiences with using wearable devices: qualitative systematic review and meta-synthesis. JMIR Mhealth Uhealth 2021 Jun 03;9(6):e23832 [FREE Full text] [doi: 10.2196/23832] [Medline: 34081020]
20. Favela J, Cruz-Sandoval D, Morales-Tellez A, Lopez-Nava IH. Monitoring behavioral symptoms of dementia using activity trackers. J Biomed Inform 2020 Sep;109:103520 [FREE Full text] [doi: 10.1016/j.jbi.2020.103520] [Medline: 32783922]
21. Kamil RJ, Bakar D, Ehrenburg M, Wei EX, Pletnikova A, Xiao G, et al. Detection of wandering behaviors using a body-worn inertial sensor in patients with cognitive impairment: a feasibility study. Front Neurol 2021;12:529661 [FREE Full text] [doi: 10.3389/fneur.2021.529661] [Medline: 33776875]
22. Buckley C, Cavadino A, Del Din S, Lord S, Taylor L, Rochester L, et al. Quantifying reliable walking activity with a wearable device in aged residential care: how many days are enough? Sensors (Basel) 2020 Nov 05;20(21):6314 [FREE Full text] [doi: 10.3390/s20216314] [Medline: 33167527]
23. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
24. Morse JM, Field PA. An overview of qualitative methods. In: Nursing Research. Boston, MA: Springer; 1996.
25. Lazarou I, Stavropoulos TG, Meditskos G, Andreadis S, Kompatsiaris I, Tsolaki M. Long-term impact of intelligent monitoring technology on people with cognitive impairment: an observational study. J Alzheimers Dis 2019 Aug 03;70(3):757-792. [doi: 10.3233/jad-190423]
26. Lee M, Mishra RK, Momin A, El-Refaei N, Bagheri AB, York MK, et al. Smart-home concept for remote monitoring of instrumental activities of daily living (IADL) in older adults with cognitive impairment: a proof of concept and feasibility study. Sensors (Basel) 2022 Sep 07;22(18):6745 [FREE Full text] [doi: 10.3390/s22186745] [Medline: 36146095]
27. Mishra RK, Park C, Momin AS, Rafaei NE, Kunik M, York MK, et al. Care4AD: a technology-driven platform for care coordination and management: acceptability study in dementia. Gerontology 2023 Sep 12;69(2):227-238 [FREE Full text] [doi: 10.1159/000526219] [Medline: 36096091]
28. Amato F, Bianchi S, Comai S, Crovari P, Pasquarelli MG, Imtiaz A, et al. CLONE: a promising system for the remote monitoring of Alzheimer's patients: an experimentation with a wearable device in a village for Alzheimer's care. In: Proceedings of the 4th EAI International Conference on Smart Objects and Technologies for Social Good. 2018 Presented at: Goodtechs '18; November 28-30, 2018; Bologna, Italy. [doi: 10.1145/3284869.3284906]
29. Cruz-Sandoval D, Favela J, Lopez-Nava IH, Morales A. Adoption of wearable devices by persons with dementia: lessons from a non-pharmacological intervention enabled by a social robot. In: Marques G, Bhoi AK, de Albuquerque VH, Hareesha KS, editors. IoT in Healthcare and Ambient Assisted Living. Singapore, Singapore: Springer; 2021:145-163.
30. Engelsma T, Yurt A, Dröes RM, Jaspers MW, Peute LW. Expert appraisal and prioritization of barriers to mHealth use for older adults living with Alzheimer's disease and related dementias: a Delphi study. Int J Med Inform 2022 Oct;166:104845 [FREE Full text] [doi: 10.1016/j.ijmedinf.2022.104845] [Medline: 35973365]
31. Jacklin K, Pitawanakwat K, Blind M, Lemieux AM, Sobol A, Warry W. Peace of mind: a community-industry-academic partnership to adapt dementia technology for Anishinaabe communities on Manitoulin Island. J Rehabil Assist Technol Eng 2020 Sep 16;7:2055668320958327 [FREE Full text] [doi: 10.1177/2055668320958327] [Medline: 32999732]
32. De Vito AN, Sawyer RJ2, LaRoche A, Arredondo B, Mizuki B, Knoop C. Acceptability and feasibility of a multicomponent telehealth care management program in older adults with advanced dementia in a residential memory care unit. Gerontol Geriatr Med 2020;6:2333721420924988 [FREE Full text] [doi: 10.1177/2333721420924988] [Medline: 32577434]

XSL•FO
RenderX

33. Gelonch O, Ribera M, Codern-Bové N, Ramos S, Quintana M, Chico G, et al. Acceptability of a lifelogging wearable camera in older adults with mild cognitive impairment: a mixed-method study. BMC Geriatr 2019 Apr 16;19(1):110 [FREE Full text] [doi: 10.1186/s12877-019-1132-0] [Medline: 30991948]
34. Godkin FE, Turner E, Demnati Y, Vert A, Roberts A, Swartz RH, et al. Feasibility of a continuous, multi-sensor remote health monitoring approach in persons living with neurodegenerative disease. J Neurol 2022 May 27;269(5):2673-2686 [FREE Full text] [doi: 10.1007/s00415-021-10831-z] [Medline: 34705114]
35. Hsu CK, Liu CC, Chang T, Liao JJ, Shu CM. Service design of a loss prevention device for older adults with dementia. Geriatrics (Basel) 2023 Sep 15;8(5):93 [FREE Full text] [doi: 10.3390/geriatrics8050093] [Medline: 37736893]
36. Anderson MS, Bankole A, Homdee N, Mitchell BA, Byfield GE, Lach J. Dementia caregiver experiences and recommendations for using the behavioral and environmental sensing and intervention system at home: usability and acceptability study. JMIR Aging 2021 Dec 06;4(4):e30353 [FREE Full text] [doi: 10.2196/30353] [Medline: 34874886]
37. Dai B, Larnyo E, Tetteh EA, Aboagye AK, Musah AI. Factors affecting caregivers' acceptance of the use of wearable devices by patients with dementia: an extension of the unified theory of acceptance and use of technology model. Am J Alzheimers Dis Other Demen 2020 Nov 03;35:1533317519883493 [FREE Full text] [doi: 10.1177/1533317519883493] [Medline: 31679390]
38. Gris F, D'Amen B, Lamura G, Paciaroni L, Socci M, Melchiorre MG. Personalized technological support for informal caregivers of older people with dementia: a co-design approach involving potential end users and healthcare professionals in three focus groups in Italy. Healthcare (Basel) 2023 Sep 28;11(19):2640 [FREE Full text] [doi: 10.3390/healthcare11192640] [Medline: 37830677]
39. Larnyo E, Dai B, Larnyo A, Nutakor JA, Ampon-Wireko S, Nkrumah EN, et al. Impact of actual use behavior of healthcare wearable devices on quality of life: a cross-sectional survey of people with dementia and their caregivers in Ghana. Healthcare (Basel) 2022 Jan 30;10(2):275 [FREE Full text] [doi: 10.3390/healthcare10020275] [Medline: 35206890]
40. Ahmed QA, Al-Neami AQ. A smart biomedical assisted system for Alzheimer patients. IOP Conf Ser Mater Sci Eng 2020 Jul 01;881:012110 [FREE Full text] [doi: 10.1088/1757-899X/881/1/012110]
41. Farina N, Sherlock G, Thomas S, Lowry RG, Banerjee S. Acceptability and feasibility of wearing activity monitors in community-dwelling older adults with dementia. Int J Geriatr Psychiatry 2019 Apr 20;34(4):617-624. [doi: 10.1002/gps.5064] [Medline: 30701592]
42. Megges H, Freiesleben SD, Rösch C, Knoll N, Wessel L, Peters O. User experience and clinical effectiveness with two wearable global positioning system devices in home dementia care. Alzheimers Dement (N Y) 2018;4:636-644 [FREE Full text] [doi: 10.1016/j.trci.2018.10.002] [Medline: 30519629]
43. McCarron HR, Zmora R, Gaugler JE. A web-based mobile app with a smartwatch to support social engagement in persons with memory loss: pilot randomized controlled trial. JMIR Aging 2019 Jun 18;2(1):e13378 [FREE Full text] [doi: 10.2196/13378] [Medline: 31518270]
44. Freytag J, Mishra RK, Street RLJ, Catic A, Dindo L, Kiefer L, et al. Using wearable sensors to measure goal achievement in older veterans with dementia. Sensors (Basel) 2022 Dec 16;22(24):9923 [FREE Full text] [doi: 10.3390/s22249923] [Medline: 36560290]
45. Mc Ardle R, Del Din S, Donaghy P, Galna B, Thomas AJ, Rochester L. The impact of environment on gait assessment: considerations from real-world gait analysis in dementia subtypes. Sensors (Basel) 2021 Jan 26;21(3):813 [FREE Full text] [doi: 10.3390/s21030813] [Medline: 33530508]
46. Hall A, Brown Wilson C, Stanmore E, Todd C. Moving beyond 'safety' versus 'autonomy': a qualitative exploration of the ethics of using monitoring technologies in long-term dementia care. BMC Geriatr 2019 May 24;19(1):145 [FREE Full text] [doi: 10.1186/s12877-019-1155-6] [Medline: 31126240]
47. Guu TW, Aarsland D, Ffytche DH. Feasibility of using wearable actigraphy for behavioural monitoring in care home residents living with moderate to severe Alzheimer's disease and agitation. Alzheimers Dement 2023 Dec 25;19(S19):e072283. [doi: 10.1002/alz.072283]
48. Musaeus CS, Waldemar G, Andersen BB, Høgh P, Kidmose P, Hemmsen MC, et al. Long-term EEG monitoring in patients with Alzheimer's disease using Ear-EEG: a feasibility study. J Alzheimers Dis 2022 Dec 06;90(4):1713-1723. [doi: 10.3233/jad-220491]
49. Dinesh K, Snyder CW, Xiong M, Tarolli CG, Sharma S, Dorsey ER, et al. A longitudinal wearable sensor study in Huntington's disease. J Huntingtons Dis 2020;9(1):69-81. [doi: 10.3233/JHD-190375] [Medline: 31868675]
50. Mc Ardle R, Morris R, Hickey A, Del Din S, Koychev I, Gunn RN, et al. Gait in mild Alzheimer's disease: feasibility of multi-center measurement in the clinic and home with body-worn sensors: a pilot study. J Alzheimers Dis 2018 Apr 10;63(1):331-341. [doi: 10.3233/jad-171116]
51. Berridge C, Zhou Y, Lazar A, Porwal A, Mattek N, Gothard S, et al. Control matters in elder care technology: evidence and direction for designing it in. DIS (Des Interact Syst Conf) 2022 Jun;2022:1831-1848 [FREE Full text] [doi: 10.1145/3532106.3533471] [Medline: 35969716]
52. Nickerson GY, Shade K. Effect of increasing physical activity on cognitive function in individuals with mild cognitive impairment: a knowledge translation to practice pilot project. J Gerontol Nurs 2021 May;47(5):14-18 [FREE Full text] [doi: 10.3928/00989134-20210401-01] [Medline: 34039095]

XSL•FO
RenderX

53. Cohen S, Waks Z, Elm JJ, Gordon MF, Grachev ID, Navon-Perry L, et al. Characterizing patient compliance over six months in remote digital trials of Parkinson's and Huntington disease. BMC Med Inform Decis Mak 2018 Dec 20;18:138. [doi: 10.1186/s12911-018-0714-7]
54. Eckert T, Bongartz M, Ullrich P, Abel B, Christian W, Kiss R, et al. Promoting physical activity in geriatric patients with cognitive impairment after discharge from ward-rehabilitation: a feasibility study. Eur J Ageing 2020 Sep 13;17(3):309-320 [FREE Full text] [doi: 10.1007/s10433-020-00555-w] [Medline: 32904844]
55. Chen JH, Lauderdale DS. Cognitive function, consent for participation, and compliance with wearable device protocols in older adults. J Gerontol A Biol Sci Med Sci 2019 Jan 16;74(2):269-273 [FREE Full text] [doi: 10.1093/gerona/gly032] [Medline: 29579176]
56. Dillon K, Prapavessis H. REducing SEDENTary behavior among mild to moderate cognitively impaired assisted living residents: a pilot randomized controlled trial (RESEDENT study). J Aging Phys Act 2021 Feb 01;29(1):27-35 [FREE Full text] [doi: 10.1123/japa.2019-0440] [Medline: 32580164]
57. Gibson RH, Gander PH. Monitoring the sleep patterns of people with dementia and their family carers in the community. Australas J Ageing 2019 Mar 07;38(1):47-51. [doi: 10.1111/ajag.12605] [Medline: 30525266]
58. Bowen ME, Gaynor B, Phillips LJ. Changes in physical and cognitive function predict sedentary behavior in older adults with mild cognitive impairment. Res Gerontol Nurs 2021 Nov;14(6):285-291. [doi: 10.3928/19404921-20211021-01] [Medline: 34807787]
59. Gedde MH, Husebo BS, Erdal A, Puaschitz NG, Vislapuu M, Angeles RC, et al. Access to and interest in assistive technology for home-dwelling people with dementia during the COVID-19 pandemic (PAN.DEM). Int Rev Psychiatry 2021 Jun 08;33(4):404-411 [FREE Full text] [doi: 10.1080/09540261.2020.1845620] [Medline: 33416012]
60. Kaenampornpan M, Khai ND, Kawattikul K. Wearable computing for dementia patients. In: Proceedings of the 16th International Conference on Computing and Information Technology. 2020 Presented at: IC2IT 2020; May 14-15, 2020; Pattaya, Thailand. [doi: 10.1007/978-3-030-44044-2_3]
61. Peeters MW, Schouten G, Wouters EJ. Wearables for residents of nursing homes with dementia and challenging behaviour: values, attitudes, and needs. Gerontechnology 2021 Jan 01;20(2):1-13. [doi: 10.4017/gt.2021.20.2.7.06]
62. Richeson NE, Croteau KA. A feasibility study examining use of the FitBit Zip™ vs. the Accusplit Eagle AC 120 XL pedometer to increase physical activity for persons with mild cognitive disorder. Activ Adapt Aging 2017 Oct 16;42(1):41-53. [doi: 10.1080/01924788.2017.1385367]
63. Stavropoulos TG, Lazarou I, Diaz A, Gove D, Georges J, Manyakov NV, et al. Wearable devices for assessing function in Alzheimer's disease: a European public involvement activity about the features and preferences of patients and caregivers. Front Aging Neurosci 2021 Apr 12;13:643135 [FREE Full text] [doi: 10.3389/fnagi.2021.643135] [Medline: 33912025]
64. Stavropoulos TG, Lazarou I, Strantsalis D, Nikolopoulos S, Kompatsiaris I, Koumanakos G. Human factors and requirements of people with mild cognitive impairment, their caregivers and healthcare professionals for eHealth systems with wearable trackers. In: Proceedings of the 2020 IEEE International Conference on Human-Machine Systems. 2020 Presented at: ICHMS 2020; September 7-9, 2020; Rome, Italy. [doi: 10.1109/ichms49158.2020.9209340]
65. van der Wardt V, Hancox JE, Burgon C, Bajwa R, Goldberg S, Harwood RH. Measuring physical activity levels in people with mild cognitive impairment or mild dementia. J Aging Phys Act 2021 Feb 01;29(1):10-16. [doi: 10.1123/japa.2019-0234] [Medline: 33049697]
66. Rose KM, Lach J, Perkhounkova Y, Gong J, Dandu SR, Dickerson R, et al. Use of body sensors to examine nocturnal agitation, sleep, and urinary incontinence in individuals with Alzheimer's disease. J Gerontol Nurs 2018 Aug 01;44(8):19-26. [doi: 10.3928/00989134-20180626-03] [Medline: 30059136]
67. Thorpe J, Forchhammer BH, Maier AM. Adapting mobile and wearable technology to provide support and monitoring in rehabilitation for dementia: feasibility case series. JMIR Form Res 2019 Oct 17;3(4):e12346 [FREE Full text] [doi: 10.2196/12346] [Medline: 31625951]
68. Wherton J, Greenhalgh T, Procter R, Shaw S, Shaw J. Wandering as a sociomaterial practice: extending the theorization of GPS tracking in cognitive impairment. Qual Health Res 2019 Feb 14;29(3):328-344 [FREE Full text] [doi: 10.1177/1049732318798358] [Medline: 30215572]
69. Raepsaet C, Serraes B, Verhaeghe S, Beeckman D. Integrating sensor technology in disposable body-worn absorbent products: a qualitative study to define user profile, (technical) criteria, conditions, and potential benefits. J Wound Ostomy Continence Nurs 2021;48(6):560-567. [doi: 10.1097/WON.0000000000000812] [Medline: 34781313]
70. Parry S, Chow M, Batchelor F, Fary RE. Physical activity and sedentary behaviour in a residential aged care facility. Australas J Ageing 2019 Mar 03;38(1):E12-E18. [doi: 10.1111/ajag.12589] [Medline: 30281184]
71. Snyder M, Dringus L, Maitland Schladen M, Chenail R, Oviawe E. Remote monitoring technologies in dementia care: an interpretative phenomenological analysis of family caregivers' experiences. Qual Report 2020 May 11;25(5):1233-1252. [doi: 10.46743/2160-3715/2020.4127]
72. Sharma N, Braakman-Jansen LM, Oinas-Kukkonen H, Croockewit JH, van Gemert-Pijnen J. Exploring the needs and requirements of informal caregivers of older adults with cognitive impairment from sensor-based care solutions: multimethod study. JMIR Aging 2023 Oct 25;6:e49319 [FREE Full text] [doi: 10.2196/49319] [Medline: 37878353]

XSL•FO
RenderX

73. Wangmo T, Lipps M, Kressig RW, Ienca M. Ethical concerns with the use of intelligent assistive technology: findings from a qualitative study with professional stakeholders. BMC Med Ethics 2019 Dec 19;20(1):98 [FREE Full text] [doi: 10.1186/s12910-019-0437-z] [Medline: 31856798]
74. Tiersen F, Batey P, Harrison MJ, Naar L, Serban AI, Daniels SJ, et al. Smart home sensing and monitoring in households with dementia: user-centered design approach. JMIR Aging 2021 Aug 11;4(3):e27047 [FREE Full text] [doi: 10.2196/27047] [Medline: 34383672]
75. Svetnik V, Wang TC, Ceesay P, Snyder E, Ceren O, Bliwise D, et al. Pilot evaluation of a consumer wearable device to assess sleep in a clinical polysomnography trial of suvorexant for treating insomnia in patients with Alzheimer's disease. J Sleep Res 2021 Dec 02;30(6):e13328. [doi: 10.1111/jsr.13328] [Medline: 34340251]
76. Sun Y, Kim HM, Xu Y, Wang Y, Kwong J, Kim SS, et al. GPS tracking in dementia caregiving: social norm, perceived usefulness, and behavioral intent to use technology. In: Proceedings of the 54th Hawaii International Conference on System Sciences. 2021 Presented at: HICSS 2021; January 5, 2021; Kauai, HI. [doi: 10.24251/hicss.2021.461]
77. O'Sullivan G, Whelan B, Gallagher N, Doyle P, Smyth S, Murphy K, et al. Challenges of using a Fitbit smart wearable among people with dementia. Int J Geriatr Psychiatry 2023 Mar 28;38(3):e5898. [doi: 10.1002/gps.5898] [Medline: 36814072]
78. Hansson SO. Technology and distributive justice. In: The Ethics of Technology: Methods and Approaches. Lanham, MD: Rowman & Littlefield Publishers; 2017:51-66.
79. Guu TW, Brem AK, Albertyn CP, Kandangwa P, Aarsland D, Ffytche D. Wrist-worn actigraphy in agitated late-stage dementia patients: a feasibility study on digital inclusion. Alzheimers Dement 2024 May 18;20(5):3211-3218 [FREE Full text] [doi: 10.1002/alz.13772] [Medline: 38497216]
80. Goerss D, Köhler S, Rong E, Temp AG, Kilimann I, Bieber G, et al. Smartwatch-based interventions for people with dementia: user-centered design approach. JMIR Aging 2024 Jun 07;7:e50107 [FREE Full text] [doi: 10.2196/50107] [Medline: 38848116]
81. Muurling M, de Boer C, Hinds C, Atreya A, Doherty A, Alepopoulos V, et al. Feasibility and usability of remote monitoring in Alzheimer's disease. Digit Health 2024 Apr 09;10:20552076241238133 [FREE Full text] [doi: 10.1177/20552076241238133] [Medline: 38601188]
82. Gaugler JE, Zmora R, Mitchell LL, Finlay J, Rosebush CE, Nkimbeng M, et al. Remote activity monitoring for family caregivers of persons living with dementia: a mixed methods, randomized controlled evaluation. BMC Geriatr 2021 Dec 18;21(1):715 [FREE Full text] [doi: 10.1186/s12877-021-02634-8] [Medline: 34922475]
83. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478. [doi: 10.2307/30036540]
84. Gao Y, Li H, Luo Y. An empirical study of wearable technology acceptance in healthcare. Ind Manag Data Syst 2015 Oct 19;115(9):1704-1723. [doi: 10.1108/IMDS-03-2015-0087]
85. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. Decis Sci 2008 May 09;39(2):273-315. [doi: 10.1111/j.1540-5915.2008.00192.x]
86. Use of wearable fitness or wellness technologies in the United States as of 2023, by age. Statista. URL: https://www.statista.com/statistics/1424304/use-of-fitness-wearables-in-the-us-by-age/ [accessed 2024-06-21]
87. Chandrasekaran R, Katthula V, Moustakas E. Patterns of use and key predictors for the use of wearable health care devices by US adults: insights from a national survey. J Med Internet Res 2020 Oct 16;22(10):e22443 [FREE Full text] [doi: 10.2196/22443] [Medline: 33064083]
88. Chinner A, Blane J, Lancaster C, Hinds C, Koychev I. Digital technologies for the assessment of cognition: a clinical review. Evid Based Ment Health 2018 May 20;21(2):67-71 [FREE Full text] [doi: 10.1136/eb-2018-102890] [Medline: 29678927]
89. Winer JR, Lok R, Weed L, He Z, Poston KL, Mormino EC, et al. Impaired 24-h activity patterns are associated with an increased risk of Alzheimer's disease, Parkinson's disease, and cognitive decline. Alzheimers Res Ther 2024 Feb 14;16(1):35 [FREE Full text] [doi: 10.1186/s13195-024-01411-0] [Medline: 38355598]

Abbreviations

MeSH: Medical Subject Headings
PICO: population, intervention, comparison, and outcomes
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QUEST: Quebec User Evaluation of Satisfaction With Assistive Technology

XSL•FO
RenderX

Edited by M Mardini; submitted 28.06.24; peer-reviewed by TW Guu, M Gagnon-Roy; comments to author 04.02.25; revised version received 25.03.25; accepted 27.04.25; published 31.07.25.

Please cite as:
Peterson CM, St Louis RM, Flannagan C
Enhancing Enrollment and Adherence in Long-Term Wearable Research on Dementia: Qualitative Systematic Review and Meta-Synthesis
JMIR Aging 2025;8:e63768
URL: https://aging.jmir.org/2025/1/e63768
doi:10.2196/63768
PMID:

©Colleen M Peterson, Renée M St Louis, Carol Flannagan. Originally published in JMIR Aging (https://aging.jmir.org), 31.07.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Barriers to and Facilitators of Digital Health Technology Adoption Among Older Adults With Chronic Diseases: Updated Systematic Review

Jennifer Hepburn[1*], BSc, MSc, MA; Lynn Williams[1*], BSc, MSc, PhD; Lisa McCann[2*], BSc, MSc, PhD

[1]Department of Psychological Sciences and Health, University of Strathclyde, Glasgow, United Kingdom

[2]Department of Computer and Information Sciences, University of Strathclyde, Glasgow, United Kingdom

[*]all authors contributed equally

Corresponding Author:
Jennifer Hepburn, BSc, MSc, MA
Department of Psychological Sciences and Health
University of Strathclyde
40 George Street
Glasgow, G1 1XQ
United Kingdom
Phone: 1 44 141 552 4400
Email: jennifer.hepburn@strath.ac.uk

Abstract

Background: Older adults with chronic diseases are key beneficiaries of digital health technologies, yet adoption remains inconsistent, particularly in rural areas and among certain demographic groups, such as older women.

Objective: This systematic review aimed to identify barriers to and facilitators of digital health adoption among older adults with chronic diseases, with particular attention to rural-urban differences, co-design, and equity-relevant factors.

Methods: This updated review built on a previously published review by extending the search to include PsycArticles, Scopus, Web of Science, and PubMed databases for studies published between April 2022 and September 2024. Gray literature from August 2021 onward was also included. Studies were eligible if they reported barriers to or facilitators of digital health adoption among adults aged ≥60 years with chronic diseases. Findings were mapped to the capability, opportunity, and motivation–behavior model and analyzed using the PROGRESS-Plus (place of residence; race, ethnicity, culture, and language; occupation; gender and sex; religion; education; socioeconomic status; and social capital–plus) equity framework. Quality was assessed using the Mixed Methods Appraisal Tool, and all results are reported in line with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines.

Results: In total, 12 studies from the original review were retained, with 17 new peer-reviewed studies added, yielding a total of 29 studies in addition to 30 documents identified in the gray literature search. Barriers included limited digital literacy and physical and cognitive challenges (capability); infrastructural deficits and usability challenges (opportunity); and privacy concerns, mistrust, and high satisfaction with existing care (motivation). Facilitators included tailored training and accessible design (capability), health care provider endorsement and hybrid care models (opportunity), and recognition of digital health benefits (motivation). Health care providers emerged as both facilitators and barriers, positively influencing adoption when engaged and trained but hindering it when lacking confidence or involvement. Comparative analysis of rural and urban contexts was limited by inconsistent reporting of equity-relevant variables. However, gray literature suggested that rural users face additional infrastructural challenges but express higher satisfaction with local care, potentially reducing motivation for digital uptake. Gender differences were observed in 5% (3/59) of the peer-reviewed studies and gray literature sources, with older women showing lower adoption and differing outcome priorities. Co-design enhanced adoption, especially when involving not just older adults but also health care providers and community stakeholders.

Conclusions: Digital health adoption among older adults is shaped by capability, opportunity, and motivation factors. Effective and equitable digital health strategies must address infrastructural and literacy barriers, engage health care providers through training and co-design, and ensure multistakeholder involvement. This review highlights that greater attention to standardized reporting of demographic variables, especially gender and rurality, is essential in digital health research to support inclusive implementation.

XSL•FO
RenderX

Trial Registration: PROSPERO International Prospective Register of Systematic Reviews CRD42024586893; https://www.crd.york.ac.uk/PROSPERO/view/CRD42024586893

International Registered Report Identifier (IRRID): RR2-https://doi.org/10.3399/bjgp25X742161

(JMIR Aging 2025;8:e80000) doi:10.2196/80000

KEYWORDS

older adults; chronic disease; digital health; technology adoption; rural health; co-design

Introduction

Background

The global population aged ≥60 years is projected to double between 2015 and 2050, placing increasing pressure on health care systems [1,2]. Population aging is widely recognized as one of the most pressing challenges facing health and social care worldwide [3]. Chronic disease is the leading cause of death worldwide, posing a major public health burden [4]. Ischemic heart disease, stroke, and chronic obstructive pulmonary disease (COPD) are among the top causes of mortality alongside Alzheimer disease, lung cancer, diabetes, chronic kidney disease, and cirrhosis [5]. These conditions typically require ongoing long-term management and monitoring, such as blood glucose tracking for diabetes and blood pressure tracking for cardiovascular disease. Multimorbidity (more than one chronic disease) is rising sharply, with an estimated two-thirds of adults aged ≥65 years expected to live with multiple chronic conditions by 2035 [6,7]. Rural and remote areas are disproportionately affected, where health systems often face additional challenges due to dispersed populations and resource constraints [8,9].

To address these pressures, innovative strategies are needed to enhance chronic disease management, particularly in settings where health care access is limited. Digital health technologies (DHTs), including telemedicine, mobile apps, remote monitoring devices, and electronic health records, offer promising tools to support health care delivery, improve health outcomes, and enhance system efficiency [10,11]. The global uptake of digital health accelerated during the COVID-19 pandemic, when remote monitoring and web-based consultations became critical [12-14]. When implemented effectively, digital health solutions have the potential to reduce clinician workload [15]; improve patient safety and care quality [16]; and decrease the need for in-person visits, which is an especially valuable benefit in rural areas, where patients often travel long distances to access health care [17]. Despite these potential benefits, there is evidence suggesting that older adults and those in rural areas are less likely to engage with DHTs compared to their urban counterparts [18-20]. Paradoxically, while these groups stand to benefit the most, they often remain digitally excluded.

The importance of digital health is underscored by the World Health Organization (WHO) *Global Strategy on Digital Health 2020-2025*, which promotes digital solutions to enhance health system resilience and equity [11]. Similarly, the regional digital health action plan for the WHO European Region 2023 to 2030 highlights the need to build digital capacity across the public and health care workforce [21]. These priorities have stimulated research into barriers to and facilitators of digital health adoption, with recurring themes such as infrastructural challenges, digital literacy gaps, and the critical role of health care providers in supporting uptake [22].

Although a growing body of literature explores digital health adoption, most studies focus on general populations or specific diseases, with limited attention to older adults with chronic conditions, a group at heightened risk of digital exclusion. For example, individuals aged ≥65 years are 18 times more likely never to have used a health app compared to those aged 18 to 24 years [23]. Despite increasing internet use during the pandemic, 34% of adults aged ≥75 years remain offline, and 13% of those aged ≥65 years do not own a mobile phone [24]. According to a report by Age UK, limited digital skills remain a significant barrier, with an estimated 4.7 million people aged ≥65 years in the United Kingdom lacking the basic digital skills required to use the internet effectively. In addition, the report found that 49% of people aged ≥50 years find it difficult to access their general practitioners (GPs) due to challenges with digital or telephone triage systems. These barriers contribute to feelings of frustration, disempowerment, and disengagement among older adults [25]. It is important to note that older adults are a diverse group whose digital engagement is shaped by the intersection of multiple factors, such as age, gender, socioeconomic status, educational level, place of residence, and health status, which can combine to create unique patterns of digital inclusion or exclusion. Understanding the common barriers to and facilitators of digital health adoption in this population is critical to inform more inclusive digital health strategies. To support this focus on equity, the PROGRESS-Plus (place of residence; race, ethnicity, culture, and language; occupation; gender and sex; religion; education; socioeconomic status; and social capital–plus) framework was selected to guide both the rationale for this review and the structured extraction of equity-related variables during analysis. PROGRESS-Plus is an equity-oriented framework that also considers the *Plus* domains of age, disability, and other contextual factors. This framework is increasingly used in systematic reviews to assess whether and how health interventions account for social determinants of health, ensuring that findings support equitable access and outcomes across diverse populations [26].

Gender is one such factor, with studies suggesting that older women, particularly those in low-income settings, are less likely to adopt DHTs than men due to lower digital confidence, greater privacy concerns, and different expectations regarding the purpose of DHTs [27-29]. These differences underscore the importance of including gender in evaluations of digital health access. Recent reviews have highlighted persistent inequities in digital health adoption; for example, a scoping review of the

XSL•FO
RenderX

WHO European Region found that digital health interventions were more often adopted by White, English-speaking, urban, and economically advantaged individuals, with limited evidence on underserved groups such as rural populations or ethnic minority groups [30]. Similarly, a review of digital health use during the COVID-19 pandemic found that rapid implementation sometimes exacerbated disparities, particularly among those with limited access to devices or low digital literacy [31]. These findings underscore the importance of applying equity frameworks such as PROGRESS-Plus when evaluating digital health uptake.

Previous Work

A previous scoping review that investigated barriers to and facilitators of older adults using eHealth identified common barriers such as lack of self-efficacy, insufficient knowledge and support, concerns over functionality, and lack of clear information about the benefits of digital health [32]. However, few reviews have focused specifically on older adults with chronic diseases, with the notable exception of a systematic review by Bertolazzi et al [33], which explored the barriers and facilitators influencing adoption of DHTs by older adults with chronic diseases. The review identified 5 domains influencing adoption: demographic and socioeconomic factors, health-related factors, dispositional attitudes, technology-related barriers, and social influences. The key facilitators were perceived usefulness, early introduction of technology, social support, and user-friendly design. Barriers included digital literacy, technical issues, a preference for existing ways of managing health, and lack of support.

Goal of This Study

This review significantly extended the work by Bertolazzi et al [33] in both scope and depth to capture a rapidly evolving digital health landscape. While the original review included studies up to April 2022, it missed a critical wave of literature emerging during and after the COVID-19 pandemic, a period marked by a surge in the deployment and evaluation of DHTs [34,35]. By identifying studies published between April 2022 and September 2024, this review broadened the evidence base to reflect more recent trends, challenges, and innovations in digital health adoption among older adults with chronic diseases.

In addition to updating the time frame, this review addresses key limitations of the earlier work by incorporating gray literature, an important yet often overlooked source in fast-moving fields such as digital health. Government reports, policy briefs, and non-peer-reviewed evaluations provide timely, real-world insights. This is particularly relevant to equity, as published literature on digital health often underrepresents older adults [36] and specifically older adults in rural communities [37]. As Landerdahl Stridsberg et al [38] note, “Gray literature is particularly relevant for digital health and HWT… Systematic reviews… should consider including gray literature based on a systematic approach.” This approach is consistent with PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidance [39], which recognises gray literature as a valid evidence source in systematic reviews where eligibility criteria are applied consistently and search processes are transparent.

This review also deepened the analysis by examining previously underexplored dimensions. While the original review noted the value of involving older adults in design, it did not assess the use or impact of co-design approaches. We addressed this by extracting data on co-design and its influence on adoption. Similarly, rurality, a key determinant of access and equity in digital health, was not systematically analyzed. Where reported, we extracted rural-urban comparisons to better understand how context shapes adoption. Equity-related reporting itself also varied widely in the previous review. To address this, we applied the PROGRESS-Plus framework to assess how studies accounted for equity-relevant factors such as gender, geography, and socioeconomic status, dimensions that are essential for informing inclusive digital health strategies.

Finally, we applied a theoretical lens absent from the original work. Using the capability, opportunity, and motivation–behavior (COM-B) model [40], we categorized barriers and facilitators to provide a more structured, behaviorally informed interpretation of the evidence. This theoretical grounding enhances the potential for actionable insights to inform intervention design and policy. This review is an update of the review by Bertolazzi et al [33] and applied the same inclusion criteria, with the search extended from April 2022 to September 2024. Limitations identified in the original review informed the scope and methods of this update, including the addition of gray literature and attention to rurality, co-design, and equity considerations.

Objectives

This review aimed to provide a more comprehensive understanding of the factors that influence digital health uptake among older adults with chronic diseases. The review was guided by the following questions:

1. What are the main factors hindering the adoption of DHTs by older adults with chronic diseases?
2. What are the main factors facilitating the adoption of these technologies?
3. How do these barriers and facilitators vary across equity-relevant factors such as place of residence (urban or rural) and gender?
4. In what ways does co-design support the adoption of DHTs for chronic disease management?

Methods

Protocol and Registration

This review followed a prespecified protocol, which was registered with PROSPERO (CRD42024586893).

Overview

This study is an updated systematic review that synthesizes research results on the barriers to and facilitators of the uptake of DHTs by older adults with chronic diseases following the PRISMA guidelines (see the PRISMA checklist in Multimedia Appendix 1) [39]. To address specific limitations identified in the original review, the update expanded the search to include gray literature, increased attention to rural contexts, and incorporated additional data extractions on co-design processes.

XSL•FO
RenderX

The full search strategies from Bertolazzi et al [33] were available and retained for reference. All studies included in the original review were rescreened using the same inclusion criteria as this update, with only eligible studies retained. In accordance with the PRISMA 2020 guidance for updated reviews [39], we searched only from April 2022 to September 2024 as the previous review had already conducted a comprehensive search from 2012 to April 2022.

Information Sources and Search Strategy

Four electronic databases were searched: PsycArticles, Scopus, Web of Science, and PubMed. Searches were conducted for publications dated April 2022 to September 2024 using a Boolean strategy combining keywords and MeSH (Medical Subject Headings) based on the population, intervention, comparison, and outcome framework [41] (Textbox 1; see Multimedia Appendix 2 for the search strategy). The date restriction was intentional as the period from 2012 to April 2022 had already been searched in the original review and rerunning it risked unnecessary duplication without yielding additional eligible studies.

Textbox 1. Search terms using the population, intervention, comparison, and outcome mnemonic.

Population

"elderly" OR "older adult" OR "ageing people"

Phenomena of interest

"technology" OR "gerontechnology"

Context

"chronic disease" OR "chronic illness" OR "long-term conditions" OR "chronic conditions"

Eligibility Criteria

The eligibility criteria used in the original review for the study selection process are presented in Textbox 2 [33].

The inclusion and exclusion criteria used in this review match those in Textbox 2 with the following amendment: studies were included if they were published between April 2022 and September 2024. Although the database searches only included empirical studies, the gray literature search included nonempirical publications. In line with the original review by Bertolazzi et al [33], *chronic disease* was defined as in their review: "health problems that are long lasting, generally progressive, and require ongoing management over a period of years or decades." This definition includes but is not limited to cardiovascular diseases (eg, hypertension and heart failure), respiratory diseases (eg, COPD and asthma), diabetes and other endocrine disorders, neurological conditions (eg, dementia and Parkinson disease), chronic kidney disease, and cancer (where ongoing management is required). For inclusion in this review, study participants had to be older adults (aged ≥60 years) with one or more diagnosed chronic diseases as defined previously.

Textbox 2. Inclusion and exclusion criteria used in the original review.

Inclusion criteria

- Population: participants aged ≥60 years and diagnosed with one or more chronic diseases
- Phenomenon of interest: focus on barriers to or facilitators of the adoption of digital health technologies for chronic disease management and technology targeting older adults
- Context: empirical studies (qualitative, quantitative, or mixed methods) and published in English between 2012 and 2022

Exclusion criteria

- Population: mixed-age samples (including those aged <60 years)
- Phenomenon of interest: studies evaluating a general health care service without focus on the specific enabling digital tools
- Context: nonempirical publications (eg, reviews, editorials, dissertations, conference abstracts, or narrative-only studies)

Searching for Gray Literature Studies

A targeted gray literature search was conducted between August 2024 and January 2025. This focused on identifying white papers, policy documents, and guidelines related to digital health implementation in rural settings for older adults with chronic conditions. To reflect developments since the publication of the WHO "Global strategy on digital health 2020-2025" [11], only documents published from August 2021 onward were included. The gray literature search included manual searches of international government websites and research organizations and targeted searches for dissertations and conference proceedings. A total of 30 gray literature documents were included. These were not formally quality appraised or counted among the included studies but were used to contextualize findings, particularly for rural settings.

XSL•FO
RenderX

Study Selection

Records were initially screened by title and abstract followed by full-text assessment for eligibility. Two reviewers independently screened titles and abstracts and then full texts for eligibility, with discrepancies resolved through consensus or a third reviewer. Duplicates were removed before screening. All studies from the original review were rescreened in full against the inclusion criteria for this update. Only those meeting the criteria were retained for synthesis alongside newly identified studies from the extended search period.

Data Extraction

Data were extracted using a data extraction tool within Covidence (Veritas Health Innovation) by 2 independent reviewers. Extracted data included (1) barriers to and facilitators of digital health adoption (outcomes); (2) age, gender, chronic conditions, urban or rural residence, and other equity variables relevant to the PROGRESS-Plus framework (participant characteristics); (3) type of DHTs (eg, telemedicine, mobile app, or wearable), setting, and whether co-design was reported (intervention characteristics); and (4) study design (qualitative, quantitative, or mixed methods), country, sample size, and publication year (study characteristics).

Where data were unclear or missing, this was noted in the extraction tables, and no assumptions were made.

Theoretical Framework for Data Interpretation and Equity Considerations

This review applied the COM-B framework [40] to interpret and map research findings on the barriers and facilitators influencing older adults with chronic diseases in their use of digital technology for self-management. As noted in the Introduction section, the PROGRESS-Plus framework was also applied to identify equity-relevant factors reported across the studies in line with Cochrane recommendations [26]. This approach enabled both behavioral and equity-focused interpretations of the findings.

Quality Assessment

The methodological quality of the included studies was assessed using the Mixed Methods Appraisal Tool (MMAT) [42]. Two reviewers independently conducted the quality assessment, with discrepancies resolved through discussion or consultation with a third reviewer.

Synthesis Methods

Due to heterogeneity in study designs, interventions, and reporting formats, a narrative synthesis was conducted. Extracted data were manually tabulated using Covidence and Microsoft Excel, and thematic analysis was applied to synthesize findings within the COM-B framework.

1. Eligibility for synthesis: all the included studies were eligible for narrative synthesis and were grouped by study design and type of digital health intervention.
2. Exploring heterogeneity: heterogeneity was explored narratively, with attention to differences across rural versus urban populations, intervention types, and study contexts.
3. Sensitivity analyses: no sensitivity analyses were conducted as no meta-analysis was performed and no pooled effect estimates were produced.
4. Risk-of-bias assessment: no formal assessment of publication bias was conducted as no meta-analysis was performed.
5. Certainty assessment: the certainty of evidence was not formally graded (eg, using the Grading of Recommendations Assessment, Development, and Evaluation) as the synthesis focused on qualitative and mixed methods findings.

Gray literature was presented separately from peer-reviewed studies to maintain transparency as these sources were not subject to the same methodological quality appraisal. This separation allows readers to interpret peer-reviewed evidence independently while still considering the contextual value of gray literature, which is integrated into the Discussion section.

Results

The findings presented in this section are from the updated review and include both newly identified studies and eligible studies retained from the original review following rescreening against the inclusion criteria. A brief comparison with the original review is provided in the Discussion section.

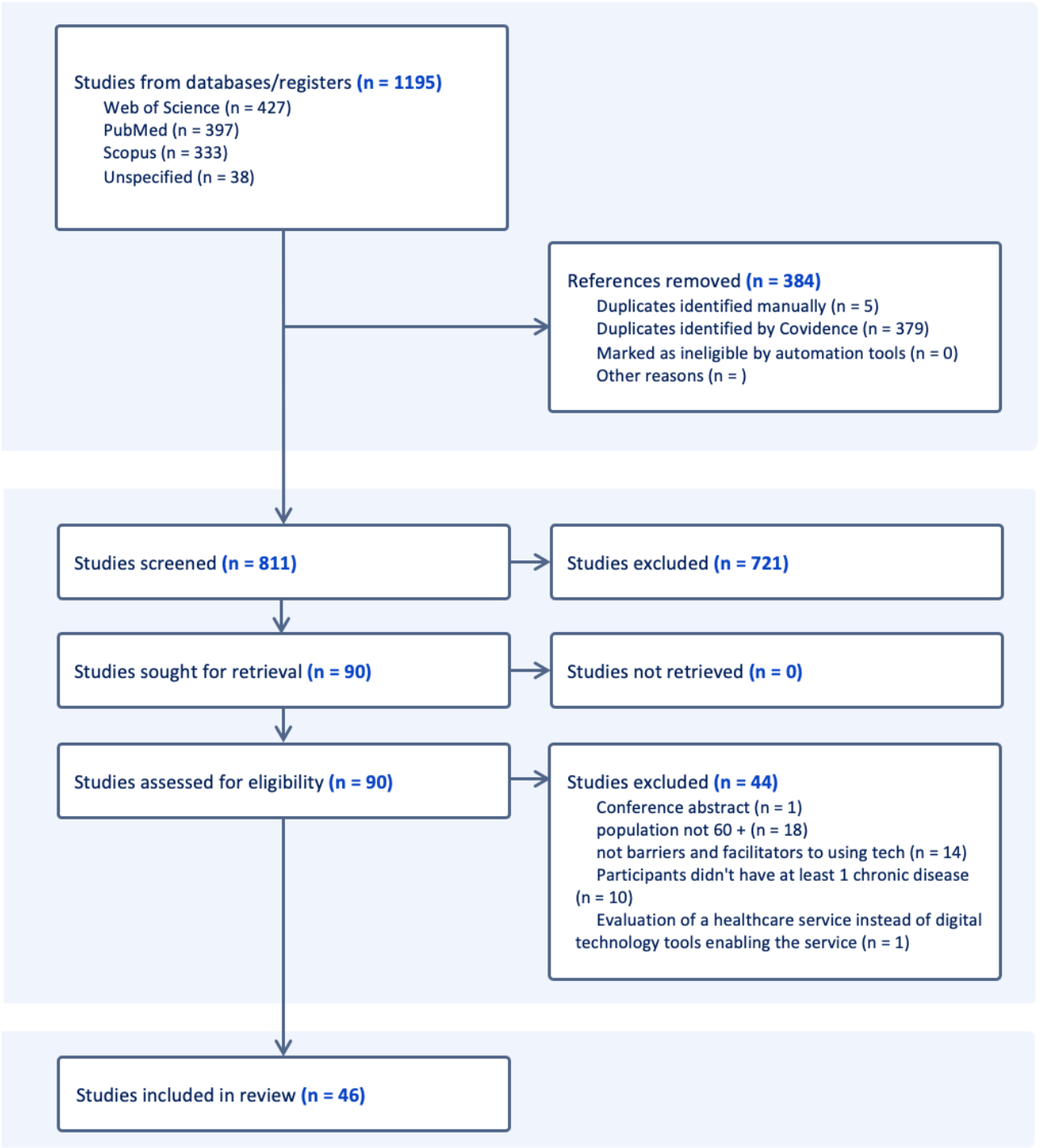
Identification of Studies

A total of 1963 records were retrieved across databases. After removing duplicates, of the 1963 records, 811 (41.31%) were screened. Two authors independently screened titles and abstracts followed by full-text assessment against the eligibility criteria, with discrepancies discussed with a third reviewer. As shown in the PRISMA flowchart (Figure 1), 46 studies were initially included, comprising 17 (37%) new studies published between April 2022 and September 2024. During the data extraction process, eligibility issues in several studies included in the previous review by Bertolazzi et al [33] were identified. Specifically, 59% (17/29) of the studies previously included did not meet the stated inclusion criteria upon re-evaluation and, therefore, were excluded from this review. This explains the difference between the 46 studies initially shown in the PRISMA diagram and the 29 peer-reviewed studies included in the final synthesis. Exclusion was based on one or more of the following factors: inclusion of mixed-age samples that did not isolate findings for adults aged ≥60 years, digital technologies not specifically targeting older adults, or the papers not including barriers to or facilitators of the adoption of DHTs for chronic disease management.

The 12 studies that did meet the criteria were retained and included alongside 17 newly identified studies. As a result, this updated review comprised a total of 29 peer-reviewed studies. Full details on the excluded studies and the reasons for their exclusion are provided in Multimedia Appendix 3 [43-60]. The PRISMA diagram reflects the peer-reviewed studies only. Gray literature was identified and screened separately, with results reported in a dedicated section of this manuscript.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart.



Characteristics of the Included Studies

A summary of study characteristics is provided in Multimedia Appendix 4 [61-89].

Among the 29 included studies, study designs comprised 1 (3%) randomized controlled trial [61] and 5 (17%) mixed methods studies [62-66]. Most studies were qualitative—38% (11/29) were interview-based studies [67-77], and 10% (3/29) used focus groups [78-80]. In total, 31% (9/29) of the studies used quantitative designs, including questionnaires and surveys [81-89].

The type of technologies used for managing chronic diseases varied widely, including wearable devices [67,72,74,77,89], mobile apps [61,68,70,73,85], digital health platforms [62,64,69,79], patient portals [80], telemonitoring [63,66,75], telemedicine [71,81,83], telehealth [76,87], information and communications technology services [78,84,86], assistive robots [65], and active video games [74].

XSL•FO
RenderX

Regarding the type and number of chronic diseases addressed, most studies (17/29, 59%) focused on multiple conditions [61-64,66,68,72,75,78,80-82,84,85,87-89]. A total of 24% (7/29) of the studies examined a single, specific chronic condition—cognitive impairment [65], COPD [67,70,74,76], or heart failure [73,77].

Barriers and Facilitators (COM-B Framework)

After summarizing the study characteristics, the barriers and facilitators were grouped under the COM-B domains: capability, opportunity, and motivation (see the list of barriers and facilitators in Multimedia Appendix 5 [61-118]).

Capability

Limited digital literacy [71,77,83-85,89] and lower educational attainment [74,76,85,86] were the most common barriers. Other barriers included unfamiliarity with smartphones [78] or the internet [69]; low health literacy [79]; early difficulties with the technology [75,77,79]; being acutely unwell at the time of installation [75]; and age-related visual, manual dexterity, or memory challenges [62,66,67,75,76,83].

Facilitators in this domain included the provision of clear instructions, structured training sessions, and technical support [63,64,68,70,71,75,77,79,80,82,85,87], as well as pretraining initiatives aimed at building users' technical skills before engaging with the technology [79]. Technology that was simple and easy to use [63,65,72,82] and early integration of the technology into care routines [65] further supported engagement.

Opportunity

Key barriers within the opportunity domain were primarily structural and contextual. This included limited access to digital devices [72,85], high technology costs [63,77-79], poor internet connectivity and technical issues [62,75,82,87], lack of goal personalization [68], single-condition focus [82], reduced social interaction [76,78], and reliance on family for support [69,70].

Practical and emotional support from family, friends, and caregivers promoted uptake [63,67,68,70,71,75,79,80,87], with involvement from adult children identified as particularly influential [67,76]. Co-design was reported in 24% (7/29) of the included studies [62-64,68,79,82,85], and person-centered approaches to co-design were specifically highlighted as facilitators in 29% (2/7) of these studies [82,85]. Related findings specific to health care providers are presented in the Health Care Providers section.

Motivation

Many older adults reported apprehension about using DHTs, often driven by concerns about data privacy [67,70,76] and mistrust of digital systems [85]. Emotional responses such as fear, anxiety, and embarrassment were commonly reported when navigating unfamiliar technologies [67,77,80,85] alongside a fear of making mistakes [79]. For some, the technology acted as a distressing reminder of illness or the possibility of recurrence [72,75,79], and negative emotions were triggered when health goals were not achieved or when data readings indicated concerning results [68,79]. Others viewed digital tools as an added burden [64], questioning their reliability [64,67,75,76] or doubting their value or benefit to health [68,69,73]. Feelings of ambivalence [76], fears of losing autonomy [78,79], and a general lack of interest in technology [85] further reduced motivation. In some cases, social perceptions also played a role, with concerns that using digital health tools might make individuals appear older or more dependent [78].

Key motivational facilitators included understanding the health benefits of the technology, gaining insights into condition management, setting goals, and feeling empowered to take an active role in care [63,64,67-69,72-74,76,81]. For some, the technology provided reassurance and a sense of safety, particularly when it enabled ongoing remote monitoring by health care professionals [66,75,78]. Positive experiences with the technology, including successful use [76] and receiving personalized feedback [79], further strengthened motivation to engage. Related findings specific to health care providers are presented in the Health Care Providers section.

Health Care Providers

This section aligns with the COM-B framework, with subthemes mapping primarily to the opportunity and motivation domains. It is presented separately due to the prominence and frequency of health care provider influences across the included studies.

Opportunity

In some instances, health care providers themselves acted as barriers, particularly when they appeared disinterested or lacked time to engage with patient-generated data [64,73,78]. Uptake was also less likely if the technology had not been recommended by a trusted health care professional [72]. Conversely, endorsement or active encouragement from trusted clinicians increased engagement [63,75-77,79,118]. Digital services that strengthened continuity of care or improved communication with clinicians were more likely to be accepted [64,69,70,80].

Motivation

Concerns that digital technologies might reduce or replace personal interaction with health care providers were common, especially among older adults who valued in-person care [71,72,75]. The perception that remote consultations were inferior to face-to-face interactions further contributed to reluctance [70]. However, clinician engagement with patient-generated data was also reported to enhance patient motivation [68,74], and collaborative involvement in care planning fostered a sense of empowerment [63]. Trust in health care providers remained a critical motivational driver for digital health uptake [71].

Rural and Urban Differences

A total of 14% (4/29) of the included studies reported on whether the study took place in a rural or urban area [66,72,75,88], and 75% (3/4) of these studies took place in rural areas. These studies identified the following: unreliable technology use in rural areas due to poor internet connectivity [75], no significant differences in telenursing needs between rural and urban participants regarding desired service features [88], and the fact that combining personalized telemonitoring with a case manager familiar with patients allowed for timely

XSL•FO
RenderX

responses to alerts without overburdening GPs or specialists [66].

Gray Literature: Rural-Specific Barriers and Facilitators

Gray literature findings are presented separately here for transparency but are synthesized with peer-reviewed evidence in the Discussion section. Additional barriers and facilitators specific to rural contexts were identified in the gray literature that were not reported in the included peer-reviewed studies.

Capability

Barriers to capability included the absence of locally delivered and culturally tailored training programs [90-95]. Among health care professionals, a lack of dedicated training, ongoing support, and clear guidance on telehealth use contributed to slower adoption and reduced confidence [93,94,97]. As previously discussed, limited digital literacy among older adults emerged as a key barrier in the peer-reviewed studies, and this was reinforced by the gray literature sources [90,94,95,98-109]. In addition, health care professionals' unfamiliarity with digital health tools and previous negative experiences further hindered engagement [96,97,107].

Facilitators included the provision of targeted local training and culturally relevant support materials tailored to rural communities [90,93,101]. These approaches were reported to enhance digital readiness and build confidence among both patients and clinicians [93].

Opportunity

Poor internet connectivity and fragmented infrastructure were consistently highlighted [90,91,96-102,106,107,109,111-113]. Additional challenges included limited promotion and awareness of available digital tools [102,113,114] alongside workforce constraints such as staff shortages, outdated or nonintegrated equipment, and inadequate resourcing [90,100,108,109,113].

Facilitators included investment in infrastructure, financial and logistical implementation support, and the use of low-technology or hybrid models such as telephone-based care to mitigate digital exclusion [100,115]. Involvement of trusted link workers, digital champions, community navigators, and clinicians was reported to support uptake and reduce burden on GPs [66,92,96,97].

Co-Design

Several sources also identified co-design as a facilitator to ensure that digital tools were relevant and usable in rural contexts. This included participatory approaches that involved older adults, families, caregivers, and community stakeholders to support adoption [90,97,104,107,110,114-117]. Barriers included the absence of participatory design [98], limited consideration of older adults' cognitive and physical needs [106,110], and insufficient community involvement during development [115]. Several sources noted that the lack of family or community input contributed to resistance or reduced the relevance of DHTs [110,115,116].

Motivation

Gray literature identified several motivational barriers affecting digital health uptake in rural areas. These included mistrust of digital platforms [90,95,99,108,110], fear that technology would replace human interaction [101], and a preference for face-to-face care [101-103,114]. Cultural values such as self-sufficiency [118], feelings of underrepresentation in national policy [106], and mistrust of the government [110] were also reported. A lack of face-to-face contact with health care professionals was associated with reduced trust and satisfaction [108]. Culture and trust were described as influencing comfort and willingness to engage with digital tools [110]. In addition, one source reported a positive association between GP availability and satisfaction with local health care, suggesting that perceptions of adequate in-person care provision may reduce motivation to adopt digital alternatives [116].

Equity and PROGRESS-Plus Findings

The PROGRESS-Plus analysis highlighted inconsistent reporting of equity-relevant factors. Characteristics such as race and ethnicity, socioeconomic status, and occupation were often underreported. Only 14% (4/29) of the studies provided data on place of residence (urban vs rural), and none offered a clear operational definition of rurality, limiting comparative analysis.

In total, 93% (27/29) of the studies reported on gender; however, only 11% (3/27) of these studies provided any comparative insights and identified gender as a factor influencing digital health adoption. Men were reported to have a stronger preference for staying up-to-date with technology [67], showed a preference for autonomy-enhancing services [78], and were less likely than women to report technology unreadiness [81]. Gray literature further highlighted that older women, particularly in low-income settings, were less likely to engage with DHTs [90,91,95,111].

This lack of data limited our ability to compare subgroups on many PROGRESS-Plus factors. Multimedia Appendix 6 [61-89] provides a list of these factors. These findings are consistent with broader evidence that reports that digital health adoption tends to favor younger, urban, and socioeconomically advantaged populations, whereas rural and minority groups remain underrepresented [31,32]. This highlights the need for more systematic equity-focused reporting in digital health research.

Quality Assessment

The MMAT was used to assess the methodological quality of all the included peer-reviewed studies. No studies were excluded based on quality in line with guidance for mixed methods systematic reviews [26]. Qualitative studies generally achieved high MMAT scores (4 to 5 stars), with strengths including clearly defined questions, appropriate data collection, and detailed analysis. Only 2/14, (14%) [67,76] explicitly reported reflexivity with the remaining studies (12/14, 86%) [68-75,77-80] focusing on other strategies to enhance rigour, limiting transparency about researcher influence. Quantitative studies were of moderate to high quality (3 to 4 stars), typically showing appropriate methods and outcome measures. However, several had methodological limitations with 3/9 (33%) [83,84,89] employing unclear or convenience sampling approaches, 3/9 (33%) [82,83,88] failing to report or partially reported response rates and 6/9 (67%) [42,82-85,89] relying at

XSL•FO
RenderX

least partially on unvalidated self-report measures. Mixed methods studies were generally of good quality. All 5/5 (100%) [62-66] integrated qualitative and quantitative strands effectively. The majority (4/5, 80%) [62-65] clearly justified their mixed methods design, and 1/5 (20%) [66] did not provide a rationale for combining methods. The full MMAT ratings for each study are provided in Multimedia Appendix 7 [61-89].

Discussion

Principal Findings

This systematic review synthesized evidence on the barriers to and facilitators of digital health adoption among older adults with chronic diseases using the COM-B model and PROGRESS-Plus framework. The main barriers to adoption were limited digital literacy and confidence (capability); poor internet access, high costs, and lack of integrated technical support (opportunity); and fears regarding privacy, mistrust, and satisfaction with existing services (motivation). Facilitators included user-friendly design and training (capability), trusted health care provider endorsement and hybrid care models (opportunity), and perceived benefits of digital health and positive user experiences (motivation).

Regarding rural and urban differences, a key insight from this review is the paradox that, although rural patients are physically distant from health care services and stand to benefit from digital health, they often report high satisfaction and trust in their existing local care. This may act as a motivational barrier, reducing perceived need to adopt digital tools. The gray literature identified a strong preference for face-to-face care [101-103,114]. This may be partly explained by the fact that rural regions often have better GP-to-patient ratios than urban areas, and strong patient-provider relationships may reinforce satisfaction with in-person care [116].

Gray literature sources identified additional challenges specific to rural settings. These included limited digital familiarity among health care professionals [96,107]; insufficient training and ongoing support [93,94]; and structural barriers such as workforce shortages, outdated equipment, and resource constraints [90,100,108,109,113]. One proposed solution to mitigate these issues is the introduction of case managers or community link workers who can engage older adults directly and help reduce the burden on overstretched health care providers [66,92,96,97].

This review also highlights gender as an underexplored factor not addressed in the original review [34]. Evidence from the gray literature indicated that older women, particularly in low-income settings, were less likely to engage with DHTs than men [90,91,95,111]. In addition, peer-reviewed studies suggested that men and women may differ in their preferences for how digital health services are delivered [67,78]. These findings underscore the importance of routinely reporting gender and other equity-related variables to prevent the design of one-size-fits-all digital health solutions. It is also important to acknowledge that sex and gender are frequently used interchangeably in the literature, including studies applying frameworks such as PROGRESS-Plus. This can potentially obscure important distinctions between biological sex and socially constructed gender roles and how they shape health behavior and digital access.

Regarding co-design, while the original review [33] acknowledged the value of involving older adults in design, this review extended that understanding by showing that involving health care providers and other stakeholders (eg, caregivers, local organizations, and community gatekeepers) in co-design processes is also essential in encouraging uptake [90,98,105,109,110,112,116,117,119]. This broader approach ensures that digital tools are not only user-friendly for patients but also practical for providers and sustainable within health care systems.

Limitations

Limitations include inconsistent reporting of equity-relevant variables across peer-reviewed studies. Because many primary studies did not report key demographic details, our ability to analyze subgroup differences was constrained. The exclusion of non–English-language studies is also a potential limitation.

Studies reporting successful adoption might be more likely to be published, potentially biasing the literature; however, our inclusion of gray literature helped mitigate this to some extent. Although considerable care was taken in conducting the gray literature search, the manual systematic search may have led to the omission of some relevant documents. The variability in how gray literature is indexed and the lack of consistent search tools make comprehensive retrieval challenging despite targeted efforts. Nonetheless, by synthesizing multiple data sources and using established frameworks, we believe that the findings identified are reliable and relevant.

Comparison With Previous Work

Our findings align with those of the original review [33] on general barriers such as digital literacy gaps, infrastructural barriers, and motivational challenges as key hurdles for older adults adopting digital health. This review added new insights on gender differences, rural satisfaction paradoxes, the dual role of health care providers, and the expanded role of multistakeholder co-design, broadening the evidence base.

Future Research Directions

This review highlights the need for future studies to improve demographic reporting, particularly regarding characteristics such as rurality and gender, which are associated with digital exclusion. To improve the clarity of such research, future studies should report sex and gender as separate variables in line with the Sex and Gender Equity Research Guidelines [120]. Applying frameworks such as the COM-B and PROGRESS-Plus during the design and reporting could improve consistency and enable a more meaningful interpretation of barriers to and facilitators of DHT engagement. In addition, there is a need to adopt an intersectional approach that recognizes how multiple factors such as age, gender, and rurality shape individual experiences. This is essential for understanding who is most at risk of digital exclusion and for designing interventions that address the complex and intersecting inequalities facing older adults. Finally, the development and application of equity indicators

XSL•FO
RenderX

for digital health interventions would ensure that equity considerations are systematically embedded in both research and implementation.

Conclusions

This review confirms that digital health adoption among older adults with chronic diseases is shaped by a complex mix of capability, opportunity, and motivation factors. While digital health offers potential to improve access and outcomes, high satisfaction with existing care (especially in rural areas) and provider-related barriers highlight the need for targeted strategies. Embedding multistakeholder co-design and paying close attention to equity considerations such as gender and rurality in reporting and design are likely to be critical to developing digital health solutions that are inclusive, effective, and sustainable.

Acknowledgments

This PhD research is possible due to a generous donation from Leslie Stretch.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[DOCX File , 268 KB - aging_v8i1e80000_app1.docx]

Multimedia Appendix 2
Search strategy.
[DOCX File , 20 KB - aging_v8i1e80000_app2.docx]

Multimedia Appendix 3
Excluded studies.
[DOCX File , 30 KB - aging_v8i1e80000_app3.docx]

Multimedia Appendix 4
Study characteristics.
[XLSX File (Microsoft Excel File), 23 KB - aging_v8i1e80000_app4.xlsx]

Multimedia Appendix 5
Barriers and facilitators.
[XLSX File (Microsoft Excel File), 38 KB - aging_v8i1e80000_app5.xlsx]

Multimedia Appendix 6
Mixed Methods Appraisal Tool results.
[XLS File (Microsoft Excel File), 51 KB - aging_v8i1e80000_app6.xls]

Multimedia Appendix 7
PROGRESS-Plus (place of residence; race, ethnicity, culture, and language; occupation; gender and sex; religion; education; socioeconomic status; and social capital–Plus) variable extraction.
[XLSX File (Microsoft Excel File), 12 KB - aging_v8i1e80000_app7.xlsx]

References

1. Ageing and health. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2025-05-29]
2. Cristea M, Noja GG, Dănăcică DE, Ştefea P. Population ageing, labour productivity and economic welfare in the European Union. Econ Res Ekon Istra 2020 May 02;33(1):1354-1376. [doi: 10.1080/1331677x.2020.1748507]
3. Rudnicka E, Napierała P, Podfigurna A, Męczekalski B, Smolarczyk R, Grymowicz M. The World Health Organization (WHO) approach to healthy ageing. Maturitas 2020 Sep;139:6-11 [FREE Full text] [doi: 10.1016/j.maturitas.2020.05.018] [Medline: 32747042]
4. About global NCDs. Centers for Disease Control and Prevention (CDC). URL: https://www.cdc.gov/globalhealth/healthprotection/ncd/global-ncd-overview.html [accessed 2025-05-29]

XSL•FO
RenderX

5. Aikins AD, Agyei-Mensah S, Agyemang C. Chronic Non-Communicable Diseases in Ghana. Multidisciplinary Perspectives. Accra, Ghana: Sub-Saharan Publishers; 2013.
6. Chobe M, Chobe S, Dayama S, Singh A, Metri K, Basa JR, et al. Prevalence of non-communicable diseases and its associated factors among urban elderly of six Indian states. Cureus 2022 Oct;14(10):e30123 [FREE Full text] [doi: 10.7759/cureus.30123] [Medline: 36381942]
7. Samal L, Fu HN, Camara DS, Wang J, Bierman AS, Dorr DA. Health information technology to improve care for people with multiple chronic conditions. Health Serv Res 2021 Oct;56 Suppl 1(Suppl 1):1006-1036 [FREE Full text] [doi: 10.1111/1475-6773.13860] [Medline: 34363220]
8. The future of NHS human resources and organisational development. NHS England. URL: https://www.england.nhs.uk/wp-content/uploads/2021/11/B0659_The-future-of-NHS-human-resources-and-organisational-development-report_22112021.pdf [accessed 2025-05-29]
9. Okobi E, Adigun AO, Ozobokeme O, Emmanuel O, Akinsanya PA, Okunromade O, et al. Examining disparities in ownership and use of digital health technology between rural and urban adults in the US: an analysis of the 2019 health information national trends survey. Cureus 2023 May;15(5):e38417. [doi: 10.7759/cureus.38417] [Medline: 37273368]
10. 20+ remote patient monitoring devices in healthcare. Binariks. URL: https://binariks.com/blog/remote-patient-monitoring-devices/ [accessed 2025-05-29]
11. Global strategy on digital health 2020-2025. World Health Organization. URL: https://www.who.int/publications/i/item/9789240020924 [accessed 2025-05-29]
12. Ndwabe H, Basu A, Mohammed J. Post pandemic analysis on comprehensive utilization of telehealth and telemedicine. Clinical eHealth 2024 Dec;7:5-14. [doi: 10.1016/j.ceh.2023.12.002]
13. Health at a glance 2023: OECD indicators. Organisation for Economic Co-operation and Development (OECD). URL: https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-2023_7a7afb35-en [accessed 2025-05-29]
14. Monaco A, Palmer K, Holm Ravn Faber N, Kohler I, Silva M, Vatland A, et al. Digital health tools for managing noncommunicable diseases during and after the COVID-19 pandemic: perspectives of patients and caregivers. J Med Internet Res 2021 Jan 29;23(1):e25652 [FREE Full text] [doi: 10.2196/25652] [Medline: 33464206]
15. A plan for digital health and social care. Department of Health and Social Care (UK). URL: https://www.gov.uk/government/publications/a-plan-for-digital-health-and-social-care/a-plan-for-digital-health-and-social-care [accessed 2025-05-29]
16. Digital clinical safety strategy. NHS England. URL: https://transform.england.nhs.uk/key-tools-and-info/digital-clinical-safety-strategy [accessed 2025-05-29]
17. Telemedicine: a key solution to rural healthcare challenges. Health Education England. URL: https://digital-transformation.hee.nhs.uk/blogs/telemedicine-a-key-solution-to-rural-healthcare-challenges [accessed 2025-05-29]
18. Alam K, Mahumud RA, Alam F, Keramat SA, Erdiaw-Kwasie MO, Sarker AR. Determinants of access to eHealth services in regional Australia. Int J Med Inform 2019 Nov;131:103960. [doi: 10.1016/j.ijmedinf.2019.103960]
19. Hall J, Skinner F, Tilley P, MacRury S. Service user preferences for diabetes education in remote and rural areas of the Highlands and Islands of Scotland. Rural Remote Health 2018 Mar;18(1):4326 [FREE Full text] [doi: 10.22605/RRH4326] [Medline: 29580062]
20. Currie M, Philip LJ, Roberts A. Attitudes towards the use and acceptance of eHealth technologies: a case study of older adults living with chronic pain and implications for rural healthcare. BMC Health Serv Res 2015 Apr 16;15:162 [FREE Full text] [doi: 10.1186/s12913-015-0825-0] [Medline: 25888988]
21. Regional digital health action plan for the WHO European region 2023-2030 (RC72). World Health Organization. URL: https://www.who.int/europe/publications/i/item/EUR-RC72-5 [accessed 2025-05-29]
22. Borges do Nascimento IJ, Abdulazeem H, Vasanthan LT, Martinez EZ, Zucoloto ML, Østengaard L, et al. Barriers and facilitators to utilizing digital health technologies by healthcare professionals. NPJ Digit Med 2023 Sep 18;6(1):161 [FREE Full text] [doi: 10.1038/s41746-023-00899-4] [Medline: 37723240]
23. Chowdhury SZ, Stevens S, Wu C, Woodward C, Andrews T, Ashall-Payne L, et al. An age-old problem or an old-age problem? A UK survey of attitudes, historical use and recommendations by healthcare professionals to use healthcare apps. BMC Geriatr 2023 Feb 24;23(1):172. [doi: 10.1186/s12877-023-03772-x]
24. The state of ageing 2023: summary report. Centre for Ageing Better. URL: https://ageing-better.org.uk/resources/summary-report-state-ageing-2023 [accessed 2025-05-29]
25. Age U. Offline and Overlooked: digital exclusion and its impact on older people. Age UK. URL: https://www.ageuk.org.uk/globalassets/age-uk/documents/reports-and-publications/reports-and-briefings/active-communities/offline-and-overlooked-report.pdf [accessed 2025-05-29]
26. O'Neill J, Tabish H, Welch V, Petticrew M, Pottie K, Clarke M, et al. Applying an equity lens to interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. J Clin Epidemiol 2014 Jan;67(1):56-64. [doi: 10.1016/j.jclinepi.2013.08.005] [Medline: 24189091]
27. Kim JK, Lee HY, Christensen MC, Merighi JR. Technology access and use, and their associations with social engagement among older adults: do women and men differ? J Gerontol B Psychol Sci Soc Sci 2017 Sep 01;72(5):836-845. [doi: 10.1093/geronb/gbw123] [Medline: 28073816]

XSL•FO
RenderX

28. Women and digital health: towards truly inclusive medicine. EDHEC Business School & Ipsos. URL: https://www.edhec.edu/en/research-and-faculty/edhec [accessed 2025-05-29]
29. Transforming lives: the impact of digital health innovations for older women. Consumer Technology Association (CTA). URL: https://www.cta.tech/articles/transforming-lives-the-impact-of-digital-health-innovations-for-older-women/[cited [accessed 2025-05-29]
30. Woolley KE, Bright D, Ayres T, Morgan F, Little K, Davies AR. Mapping inequities in digital health technology within the World Health Organization's European region using PROGRESS PLUS: scoping review. J Med Internet Res 2023 Apr 28;25:e44181 [FREE Full text] [doi: 10.2196/44181] [Medline: 37115613]
31. Yao R, Zhang W, Evans R, Cao G, Rui T, Shen L. Inequities in health care services caused by the adoption of digital health technologies: scoping review. J Med Internet Res 2022 Mar 21;24(3):e34144. [doi: 10.2196/34144] [Medline: 35311682]
32. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556. [doi: 10.1186/s12889-021-11623-w]
33. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: an integrative systematic review. BMC Public Health 2024 Feb 16;24(1):506 [FREE Full text] [doi: 10.1186/s12889-024-18036-5] [Medline: 38365698]
34. Wosik J, Fudim M, Cameron B, Gellad ZF, Cho A, Phinney D, et al. Telehealth transformation: COVID-19 and the rise of virtual care. J Am Med Inform Assoc 2020 Jun 01;27(6):957-962 [FREE Full text] [doi: 10.1093/jamia/ocaa067] [Medline: 32311034]
35. Monaghesh E, Hajizadeh A. The role of telehealth during COVID-19 outbreak: a systematic review based on current evidence. BMC Public Health 2020 Aug 01;20(1):1193 [FREE Full text] [doi: 10.1186/s12889-020-09301-4] [Medline: 32738884]
36. Poli A, Kelfve S, Klompstra L, Strömberg A, Jaarsma T, Motel-Klingebiel A. Prediction of (non)participation of older people in digital health research: exergame intervention study. J Med Internet Res 2020 Jun 05;22(6):e17884 [FREE Full text] [doi: 10.2196/17884] [Medline: 32501275]
37. Shiroma K, Miller J. Representation of rural older adults in AI health research: a systematic review. Innov Aging 2024;8(Suppl 1):258-5300 https://doi.org/10.1093/geroni/igae098.0833 [FREE Full text] [doi: 10.1093/geroni/igae098.0833]
38. Landerdahl Stridsberg S, Richardson MX, Redekop K, Ehn M, Wamala Andersson S. Gray literature in evaluating effectiveness in digital health and health and welfare technology: a source worth considering. J Med Internet Res 2022 Mar 23;24(3):e29307 [FREE Full text] [doi: 10.2196/29307] [Medline: 35319479]
39. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71]
40. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42]
41. Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. ACP J Club 1995;123(3):A12-A13. [Medline: 7582737]
42. Hong QN, Pluye P, Fàbregues S, Bartlett G, Boardman F, Cargo M, et al. Mixed Methods Appraisal Tool (MMAT), version 2018. Department of Family Medicine. URL: https://mixedmethodsappraisaltoolpublic.pbworks.com/w/file/fetch/127916259/MMAT_2018_criteria-manual_2018-08-01_ENG.pdf [accessed 2025-05-29]
43. AlMahadin G, Lotfi A, Zysk E, Siena FL, Carthy MM, Breedon P. Parkinson's disease: current assessment methods and wearable devices for evaluation of movement disorder motor symptoms - a patient and healthcare professional perspective. BMC Neurol 2020 Nov 18;20(1):419 [FREE Full text] [doi: 10.1186/s12883-020-01996-7] [Medline: 33208135]
44. Ancker JS, Witteman HO, Hafeez B, Provencher T, Van de Graaf M, Wei E. "You get reminded you're a sick person": personal data tracking and patients with multiple chronic conditions. J Med Internet Res 2015 Aug 19;17(8):e202 [FREE Full text] [doi: 10.2196/jmir.4209] [Medline: 26290186]
45. Banbury A, Parkinson L, Nancarrow S, Dart J, Gray L, Buckley J. Multi-site videoconferencing for home-based education of older people with chronic conditions: the Telehealth Literacy Project. J Telemed Telecare 2014 Oct;20(7):353-359. [doi: 10.1177/1357633X14552369] [Medline: 25399994]
46. Christiansen L, Sanmartin Berglund J, Anderberg P, Cellek S, Zhang J, Lemmens E, et al. Associations between mobile health technology use and self-rated quality of life: a cross-sectional study on older adults with cognitive impairment. Gerontol Geriatr Med 2021 May 25;7:23337214211018924 [FREE Full text] [doi: 10.1177/23337214211018924] [Medline: 34104685]
47. Duroseau N, Abramson T, Pergament K, Chan V, Govindavari JP, Ciraco C, et al. Acceptance of technology-based tools in a sample of Parkinson's patients. Chronic Illn 2017 Mar 08;13(1):3-13. [doi: 10.1177/1742395316653453] [Medline: 27269275]
48. Gellis ZD, Kenaley B, McGinty J, Bardelli E, Davitt J, Ten Have T. Outcomes of a telehealth intervention for homebound older adults with heart or chronic respiratory failure: a randomized controlled trial. Gerontologist 2012 Aug;52(4):541-552. [doi: 10.1093/geront/gnr134] [Medline: 22241810]
49. Rice F, Podolna O, Legere T, Ratcliffe J. From concept to care: health technology talent in Alberta. SSRN Journal 2024 Nov 05;6(2):195-211 [FREE Full text] [doi: 10.2139/ssrn.4831094]

XSL•FO
RenderX

50. Metting EI, Schrage AJ, Kocks JW, Sanderman R, van der Molen T. Assessing the needs and perspectives of patients with asthma and chronic obstructive pulmonary disease on patient web portals: focus group study. JMIR Form Res 2018 Nov 22;2(2):e22 [FREE Full text] [doi: 10.2196/formative.8822] [Medline: 30684436]
51. Nancarrow SA, Banbury A, Buckley J. Evaluation of a national broadband network-enabled telehealth trial for older people with chronic disease. Aust Health Rev 2016 Jan;40(6):641-648. [doi: 10.1071/AH15201] [Medline: 27028234]
52. Or CK, Tao D. A 3-month randomized controlled pilot trial of a patient-centered, computer-based self-monitoring system for the care of type 2 diabetes mellitus and hypertension. J Med Syst 2016 Apr;40(4):81. [doi: 10.1007/s10916-016-0437-1] [Medline: 26802011]
53. Price-Haywood EG, Harden-Barrios J, Ulep R, Luo Q. eHealth literacy: patient engagement in identifying strategies to encourage use of patient portals among older adults. Popul Health Manag 2017 Dec;20(6):486-494. [doi: 10.1089/pop.2016.0164] [Medline: 28384076]
54. Technology enhanced working: reports and conference contributions. Remote and Rural Healthcare Education Alliance. 2025 Sep 02. URL: https://learn.nes.nhs.scot/1527# [accessed 2025-09-02]
55. Robinson SA, Lee AA, Kennedy E, Harmon KJ. Factors associated with technology adoption among older adults with heart failure. J Card Fail 2020;26(7):649. [doi: 10.1016/j.cardfail.2020.02.007]
56. Scottish islands: data overview 2023. Scottish Government Riaghaltas na h-Alba. URL: https://www.gov.scot/publications/scottish-islands-data-overview-2023/pages/3/ [accessed 2025-09-02]
57. Smaerup M, Gronvall E, Larsen SB. Assistive technology for the elderly: effect on occupational performance and quality of life. J Assist Technol 2016;10(1):34. [doi: 10.1108/JAT-09-2015-0026]
58. Caldeira C, Naurain N, Connelly K. “I hope I never need one”: unpacking stigma in aging in place technology. In: ResearchGate. 2022 Apr 29 Presented at: CHI '22: CHI Conference on Human Factors in Computing Systems; April 29, 2022; New Orleans, LA p. 1-12 URL: https://www.researchgate.net/publication/360263400_I_hope_I_never_need_one_Unpacking_Stigma_in_Aging_in_Place_Technology/> [doi: 10.1145/3491102.3517586]
59. Zulfiqar A, Vaudelle O, Hajjam M, Geny B, Talha S, Letourneau D, et al. Results of the "GER-e-TEC" experiment involving the use of an automated platform to detect the exacerbation of geriatric syndromes. J Clin Med 2020 Nov 26;9(12):a [FREE Full text] [doi: 10.3390/jcm9123836] [Medline: 33256080]
60. Zulfiqar AA, Massimbo DN, Hajjam M, Gény B, Talha S, Hajjam J, et al. Glycemic disorder risk remote monitoring program in the COVID-19 very elderly patients: Preliminary results. Front Physiol 2021 Oct 27;12:749731 [FREE Full text] [doi: 10.3389/fphys.2021.749731] [Medline: 34777011]
61. Wong AK, Bayuo J, Wong FK, Chow KK, Wong SM, Lau AC. The synergistic effect of nurse proactive phone calls with an mHealth app program on sustaining app usage: 3-arm randomized controlled trial. J Med Internet Res 2023 May 01;25:e43678 [FREE Full text] [doi: 10.2196/43678] [Medline: 37126378]
62. Sheng Y, Bond R, Jaiswal R, Dinsmore J, Doyle J. Augmenting K-means clustering with qualitative data to discover the engagement patterns of older adults with multimorbidity when using digital health technologies: proof-of-concept trial. J Med Internet Res 2024 Mar 28;26:e46287 [FREE Full text] [doi: 10.2196/46287] [Medline: 38546724]
63. Ondiege B, Clarke M. Investigating user identification in remote patient monitoring devices. Bioengineering (Basel) 2017 Sep 13;4(3):76 [FREE Full text] [doi: 10.3390/bioengineering4030076] [Medline: 28952556]
64. Doyle J, Murphy E, Gavin S, Pascale A, Deparis S, Tommasi P, et al. A digital platform to support self-management of multiple chronic conditions (ProACT): findings in relation to engagement during a one-year proof-of-concept trial. J Med Internet Res 2021 Dec 15;23(12):e22672 [FREE Full text] [doi: 10.2196/22672] [Medline: 34914612]
65. Søraa RA, Nyvoll P, Tøndel G, Fosch-Villaronga E, Serrano JA. The social dimension of domesticating technology: interactions between older adults, caregivers, and robots in the home. Technol Forecast Soc Change 2021 Jun;167:120678. [doi: 10.1016/j.techfore.2021.120678]
66. Schoevaerdts D, Lerude MP, Tellier V, Pierard M, Voilmy D, Novella J. Home telemonitoring in smart rurality: results from the HIS2R interreg feasibility pilot study. Aging Clin Exp Res 2024 Mar 13;36(1):275-283. [doi: 10.1007/S40520-024-02709-7]
67. Arnaert A, Sumbly P, da Costa D, Liu Y, Debe Z, Charbonneau S. Acceptance of the Apple watch series 6 for telemonitoring of older adults with chronic obstructive pulmonary disease: qualitative descriptive study part 1. JMIR Aging 2023 Dec 26;6:e41549. [doi: 10.2196/41549]
68. Tahsin F, Austin T, McKinstry B, Mercer SW, Loganathan M, Thavorn K, et al. Examining use behavior of a goal-supporting mHealth app in primary care among patients with multiple chronic conditions: qualitative descriptive study. JMIR Hum Factors 2022 Nov 30;9(4):e37684 [FREE Full text] [doi: 10.2196/37684] [Medline: 36449335]
69. Mroz EL, Hernandez-Bigos K, Esterson J, Kiwak E, Naik A, Tinetti ME. Acceptability and use of an online health priorities self-identification tool for older adults: a qualitative investigation. PEC Innov 2023 Dec 15;3:100242 [FREE Full text] [doi: 10.1016/j.pecinn.2023.100242] [Medline: 38161685]
70. Zou X, Sun P, Chen M, Nan J, Gao J, Huang X, et al. Experience of older patients with COPD using disease management apps: a qualitative study. Healthcare (Basel) 2024 Apr 07;12(7):802 [FREE Full text] [doi: 10.3390/healthcare12070802] [Medline: 38610224]

XSL•FO
RenderX

71. Buawangpong N, Pinyopornpanish K, Pliannuom S, Nantsupawat N, Wiwatkunupakarn N, Angkurawaranon C, et al. Designing telemedicine for older adults with multimorbidity: content analysis study. JMIR Aging 2024 Jan 10;7:e52031 [FREE Full text] [doi: 10.2196/52031] [Medline: 38198201]
72. Kouri A, Gupta S, Straus SE, Sale JE. Exploring the perspectives and experiences of older adults with asthma and chronic obstructive pulmonary disease toward mobile health: qualitative study. J Med Internet Res 2023 Aug 22;25:e45955 [FREE Full text] [doi: 10.2196/45955] [Medline: 37606961]
73. Foster M, Xiong W, Quintiliani L, Hartmann CW, Gaehde S. Preferences of older adult veterans with heart failure for engaging with mobile health technology to support self-care: qualitative interview study among patients with heart failure and content analysis. JMIR Form Res 2022 Dec 20;6(12):e41317 [FREE Full text] [doi: 10.2196/41317] [Medline: 36538348]
74. Simmich J, Mandrusiak A, Russell T, Smith S, Hartley N. Perspectives of older adults with chronic disease on the use of wearable technology and video games for physical activity. Digit Health 2021 May 30;7:20552076211019900 [FREE Full text] [doi: 10.1177/20552076211019900] [Medline: 34104468]
75. Middlemass JB, Vos J, Siriwardena AN. Perceptions on use of home telemonitoring in patients with long term conditions - concordance with the Health Information Technology Acceptance Model: a qualitative collective case study. BMC Med Inform Decis Mak 2017 Jun 26;17(1):89 [FREE Full text] [doi: 10.1186/s12911-017-0486-5] [Medline: 28651588]
76. Jiang Y, Sun P, Chen Z, Guo J, Wang S, Liu F, et al. Patients' and healthcare providers' perceptions and experiences of telehealth use and online health information use in chronic disease management for older patients with chronic obstructive pulmonary disease: a qualitative study. BMC Geriatr 2022 Jan 03;22(1):9 [FREE Full text] [doi: 10.1186/s12877-021-02702-z] [Medline: 34979967]
77. Cajita MI, Hodgson NA, Lam KW, Yoo S, Han H. Facilitators of and barriers to mHealth adoption in older adults with heart failure. Comput Inform Nurs 2018 Aug;36(8):376-382 [FREE Full text] [doi: 10.1097/CIN.0000000000000442] [Medline: 29742549]
78. Jo HS, Hwang YS. Psychological factors that affect the acceptance and need for ICT services for older adults with chronic diseases. Gerontechnology 2021 Jan 01;20(2):1-11. [doi: 10.4017/gt.2021.20.2.411.01]
79. Malavasi M, Cesario L, Fiordelmondo V, Gherardini A, Hoogerwerf E, Lepore C, et al. Validation of an IoT-based home system for integrated care services: a qualitative investigation involving older adults with multiple chronic health conditions. J Ambient Intell Human Comput 2023 Aug 31;15(5):2779-2794. [doi: 10.1007/s12652-023-04677-8]
80. Portz JD, Bayliss EA, Bull S, Boxer RS, Bekelman DB, Gleason K, et al. Using the technology acceptance model to explore user experience, intent to use, and use behavior of a patient portal among older adults with multiple chronic conditions: descriptive qualitative study. J Med Internet Res 2019 Apr 08;21(4):e11604. [doi: 10.2196/11604]
81. Rodríguez-Fernández JM, Danies E, Hoertel N, Galanter W, Saner H, Franco OH. Telemedicine readiness across medical conditions in a US national representative sample of older adults. J Appl Gerontol 2021 Dec 02;41(4):982-992. [doi: 10.1177/07334648211056231]
82. Jimenez J, del Rio A, Berman AN, Grande M. Personalizing digital health: adapting health technology systems to meet the needs of different older populations. Preprints.org Preprint posted online on April 6, 2023. [doi: 10.20944/preprints202304.0078.v1]
83. Ufholz K, Sheon A, Bhargava D, Rao G. Telemedicine preparedness among older adults with chronic illness: survey of primary care patients. JMIR Form Res 2022 Jul 27;6(7):e35028. [doi: 10.2196/35028]
84. Ha J, Park HK. Factors affecting the acceptability of technology in health care among older Korean adults with multiple chronic conditions: a cross-sectional study adopting the senior technology acceptance model. Clin Interv Aging 2020;15:1873-1881 [FREE Full text] [doi: 10.2147/CIA.S268606] [Medline: 33061336]
85. Ventura MT, Giuliano AF, Buquicchio R, Bedbrook A, Czarlewski W, Laune D, et al. Implementation of the MASK-Air® app for rhinitis and asthma in older adults: MASK@Puglia pilot study. Int Arch Allergy Immunol 2022 Sep 22;183(1):45-50. [doi: 10.1159/000518032] [Medline: 34569536]
86. Chen Y, Yuan J, Shi L, Zhou J, Wang H, Li C, et al. Understanding the role of technology anxiety in the adoption of digital health technologies (DHTs) by older adults with chronic diseases in shanghai: an extension of the unified theory of acceptance and use of technology (UTAUT) model. Healthcare 2024 Jul 16;12(14):1421 [FREE Full text] [doi: 10.3390/healthcare12141421] [Medline: 39057564]
87. Tian XF, Wu RZ. Determinants of the mobile health continuance intention of elders with chronic diseases: an integrated framework of ECM-ISC and UTAUT. Int J Environ Res Public Health 2022 Aug 12;19(16):9980 [FREE Full text] [doi: 10.3390/ijerph19169980] [Medline: 36011615]
88. Yuan Y, Tao C, Yu P, Wang Y, Kitayama A, Takashi E, et al. Demand analysis of telenursing among empty-nest elderly individuals with chronic diseases based on the Kano model. Front Public Health 2022 Sep 28;10:990295 [FREE Full text] [doi: 10.3389/fpubh.2022.990295] [Medline: 36249233]
89. Herkert C, Graat-Verboom L, Gilsing-Fernhout J, Schols M, Kemps HM. Home-based exercise program for patients with combined advanced chronic cardiac and pulmonary diseases: exploratory study. JMIR Form Res 2021 Nov 09;5(11):e28634 [FREE Full text] [doi: 10.2196/28634] [Medline: 34751655]
90. The state of digital health report 2023: global digital health monitor. The State of Digital Health. URL: https://digitalhealthmonitor.org/stateofdigitalhealth23 [accessed 2025-05-29]

XSL•FO
RenderX

91. Innovation and leadership in healthy aging: global insights to inform policy and enhance the well-being of older adults. AARP, FP Analytics. URL: https://www.aarpinternational.org/File%20Library/Healthy%20Aging/InnovationInHealthyAgingReport.doi.10.26419-2fint.00050.001.pdf [accessed 2025-05-29]
92. Digital health and care strategy: enabling, connecting and empowering: care in the digital age. Scottish Government. URL: https://tinyurl.com/7pn4pmmp [accessed 2025-05-29]
93. Psychosocial applications of technology for health and wellness coaching of older adults with dementia and mild cognitive impairment and their carers in rural areas. Universidad de Salamanca. URL: https://gredos.usal.es/bitstream/handle/10366/156142/PDB_MolinariUlateM_Psychosocial.pdf?isAllowed=y&sequence=1 [accessed 2025-05-29]
94. Diabetes tech can't wait: closing the technology gap in Scotland. Scottish Parliament. URL: https://www.diabetes.org.uk/sites/default/files/2023-11/Diabetes%20Scotland%20-%20Diabetes%20Tech%20Can't%20Wait%20Report%202023_0.pdf?VersionId=z4OrebnDcjAXafeRIMrEp5uyAT9foqBL [accessed 2025-06-29]
95. Raja A. Task sharing management of non-communicable diseases using mHealth in rural India. Monash University. URL: https://bridges.monash.edu/articles/thesis/Task_sharing_management_of_non-communicable_diseases_using_mHealth_in_rural_India/22647550 [accessed 2025-06-29]
96. Terje AK, Munoz SA, Hall J. Evaluation of the mPower project 2017-2022: full report. University of the Highlands and Islands. URL: https://mpowerhealth.eu/wp-content/uploads/2022/07/Evaluation-of-the-mPower-Project-2017-2022-Full-Report.pdf [accessed 2025-06-29]
97. MBS Review Advisory Committee. Telehealth postimplementation review: final report. Department of Health and Aged Care. URL: https://www.health.gov.au/sites/default/files/2024-06/mbs-review-advisory-committee-telehealth-post-implementation-review-final-report.pdf [accessed 2025-06-29]
98. Raja M, Kymre IG, Bjerkan J, Galvin KT, Uhrenfeldt L. National digital strategies and innovative eHealth policies concerning older adults' dignity: a document analysis in three Scandinavian countries. BMC Health Serv Res 2023 Aug 10;23(1):848 [FREE Full text] [doi: 10.1186/s12913-023-09867-w] [Medline: 37563599]
99. Yusuf MO. Discourse, materiality, and the users of mobile health technologies: a Nigerian case study. Michigan Technological University. URL: https://digitalcommons.mtu.edu/etdr/1531/[cited [accessed 2025-06-29]
100. Igwama GT, Nwankwo EI, Emeihe EV, Ajegbile MD. AI-enhanced remote monitoring for chronic disease management in rural areas. Int J Appl Res Soc Sci 2024;6(8):1824-1847 [FREE Full text] [doi: 10.51594/ijarss.v6i8.1428]
101. Ufholz K, Sheon A, Bhargava D, Rao G. Telemedicine preparedness among older adults with chronic illness: survey of primary care patients. JMIR Form Res 2022 Jul 27;6(7):e35028 [FREE Full text] [doi: 10.2196/35028] [Medline: 35896013]
102. Khoda A, Njenga JK. mHealth enablers and inhibitors in a rural poor setting. In: Proceedings of the 2022 IST-Africa Conference. 2022 Presented at: IST-Africa '22; May 16-20, 2022; Dublin, Ireland p. 1-13 URL: https://ieeexplore.ieee.org/document/9845660 [doi: 10.23919/ist-africa56635.2022.9845660]
103. Greer DB, Abel WM. Exploring feasibility of mHealth to manage hypertension in rural black older adults: a convergent parallel mixed method study. Patient Prefer Adherence 2022 Aug;Volume 16:2135-2148 [FREE Full text] [doi: 10.2147/ppa.s361032]
104. Care in the digital age: delivery plan 2023-24. Digital Health and Care Scotland. URL: https://www.gov.scot/publications/care-digital-age-delivery-plan-2023-24/[cited [accessed 2025-06-29]
105. Donawa A, Powell C, Wang R, Chih M, AronoffSpencer E, Baker C. Note: assessing cancer patient usability of a mobile distress screening app. In: Proceedings of the 5th ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies. 2022 Presented at: COMPASS '22; June 29-July 1, 2022; Seattle, WA p. 643-649 URL: https://dl.acm.org/doi/10.1145/3530190.3534833 [doi: 10.1145/3530190.3534833]
106. National islands plan: annual report 2023. Scottish Government. URL: https://www.gov.scot/publications/national-islands-plan-annual-report-2023/[cited [accessed 2025-06-29]
107. Herkert C, Graat-Verboom L, Gilsing-Fernhout J, Schols M, Kemps HM. Home-based exercise program for patients with combined advanced chronic cardiac and pulmonary diseases: exploratory study. JMIR Form Res 2021 Nov 09;5(11):e28634 [FREE Full text] [doi: 10.2196/28634] [Medline: 34751655]
108. Berggren P, Pourazar E, Lundqvist A. Sweden: access to rural services by strengthening primary care with digital tools in remote areas of Sweden. World Health Organization. URL: https://www.who.int/europe/publications/m/item/sweden-access-to-rural-services-by-strengthening-primary-care-with-digital-tools-in-remote-areas-of-sweden-2022 [accessed 2025-06-29]
109. Legere T, Podolna O, Ratcliffe J, Rice F. From concept to care: health technology talent in Alberta. Information and Communications Technology Council. URL: https://ictc-ctic.ca/reports/concept-care-health-technology-talent-alberta [accessed 2025-06-29]
110. Rural health strategy. Ministry of Health. URL: https://www.health.govt.nz/system/files/2023-07/rural-health-strategy-oct23-v2.pdf [accessed 2025-06-12]
111. Digital health to support family planning providers. United States Agency for International Development. URL: https://www.fphighimpactpractices.org/wp-content/uploads/2020/04/DigitalHealth_HIP-Brief.pdf [accessed 2025-06-29]
112. The digital health blueprint and action plan 2023-2033. Australian Government Department of Health and Aged Care. URL: https://www.health.gov.au/resources/publications/the-digital-health-blueprint-and-action-plan-2023-2033?language=en [accessed 2025-06-29]

XSL•FO
RenderX

113. Bogulski CA, Pro G, Acharya M, Ali MM, Brown CC, Hayes CJ, et al. The association between rurality, dual Medicare/Medicaid eligibility and chronic conditions with telehealth utilization: An analysis of 2019–2020 national Medicare claims. J Telemed Telecare 2024 Feb 05;31(6):842-852. [doi: 10.1177/1357633x241226741]
114. A tale of two country sides: remote and rural health and medicine. University of Central Lancashire. URL: https://www.lancashire.ac.uk/assets/pdf/rural-medicine-and-health-report.pdf [accessed 2025-06-29]
115. Eberly LA, Tennison A, Mays D, Hsu CY, Yang CT, Benally E, et al. Telephone-based guideline-directed medical therapy optimization in Navajo nation: the Hózhó randomized clinical trial. JAMA Intern Med 2024 Jun 01;184(6):681-690 [FREE Full text] [doi: 10.1001/jamainternmed.2024.1523] [Medline: 38583185]
116. Scottish islands: data overview 2023. Scottish Government. URL: https://www.gov.scot/publications/scottish-islands-data-overview-2023 [accessed 2025-06-29]
117. Garcia JJ, Castilla C, Aguirre J, Martinez J, Liu W. Experiences in the design of localized eHealth tools for users facing inequality of access to healthcare. In: Proceedings of the 11th International Conference on Design, User Experience, and Usability: Design for Emotion, Well-being and Health, Learning, and Culture. 2022 Presented at: HCII '22; June 26-July 1, 2022; Virtual Event p. 130-148 URL: https://link.springer.com/chapter/10.1007/978-3-031-05900-1_8 [doi: 10.1007/978-3-031-05900-1_8]
118. Caldeira C, Nurain N, Connelly K. "I hope I never need one": unpacking stigma in aging in place technology. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 2022 Presented at: CHI '22; New Orleans, LA; April 29-May 5, 2022 p. 1-12 URL: https://dl.acm.org/doi/10.1145/3491102.3517586 [doi: 10.1145/3491102.3517586]
119. Rural Scotland data dashboard: overview. Scottish Government. URL: https://www.gov.scot/publications/scottish-islands-data-overview-2023/ [accessed 2025-06-29]
120. Heidari S, Babor TF, De Castro P, Tort S, Curno M. Sex and gender equity in research: rationale for the SAGER guidelines and recommended use. Res Integr Peer Rev 2016;1:2 [FREE Full text] [doi: 10.1186/s41073-016-0007-6] [Medline: 29451543]

Abbreviations

COM-B: capability, opportunity, and motivation–behavior
COPD: chronic obstructive pulmonary disease
DHT: digital health technology
GP: general practitioner
MeSH: Medical Subject Headings
MMAT: Mixed Methods Appraisal Tool
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROGRESS-Plus: place of residence; race, ethnicity, culture, and language; occupation; gender and sex; religion; education; socioeconomic status; and social capital–plus ("plus" includes additional factors such as age and disability)
WHO: World Health Organization

Edited by AF Abdulai; submitted 02.07.25; peer-reviewed by D Bright, F Hagenimana; comments to author 31.07.25; revised version received 04.08.25; accepted 22.08.25; published 11.09.25.

Please cite as:
Hepburn J, Williams L, McCann L
Barriers to and Facilitators of Digital Health Technology Adoption Among Older Adults With Chronic Diseases: Updated Systematic Review
JMIR Aging 2025;8:e80000
URL: https://aging.jmir.org/2025/1/e80000
doi:10.2196/80000
PMID:

©Jennifer Hepburn, Lynn Williams, Lisa McCann. Originally published in JMIR Aging (https://aging.jmir.org), 11.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital Measurement of Subjective Experiences in Alzheimer Disease and Related Dementias (AD/ADRD)

Colin Depp, PhD; Jason Holden, PhD; Eric Granholm, PhD

Department of Psychiatry, University of California, San Diego, 9500 Gilman Drive, MC 0737, La Jolla, CA, United States

Corresponding Author:
Eric Granholm, PhD
Department of Psychiatry, University of California, San Diego, 9500 Gilman Drive, MC 0737, La Jolla, CA, United States

Abstract

Symptoms such as loss of pleasure, agitation, and sadness are subjective experiences that contribute significantly to caregiver burden and health care costs in Alzheimer disease and related dementias (AD/ADRD). However, traditional self-report measures of subjective experiences are limited in AD/ADRD due to cognitive impairments and awareness. Passive sensing, which collects data without active participant input, has emerged as a promising approach to quantify aspects of subjective experiences. Smartphones, wearables, and in-home sensors can quantify mobility, physiology, speech, and social interaction markers of constructs relevant to AD/ADRD. Available research indicates potential but is largely at the proof-of-concept stage. In this Commentary, we discuss several roadblocks to future translation of passive sensing in measuring subjective experiences in AD/ADRD, including technical implementation, data harmonization, validation, ethical and privacy principles. Addressing these challenges could lead to transformative applications to care for AD/ADRD, enabling precise monitoring of behavioral symptoms and related treatment targets, ultimately improving quality of life for persons with AD/ADRD and their caregivers.

(JMIR Aging 2025;8:e71920) doi:10.2196/71920

KEYWORDS

Alzheimer's disease; measurement; passive sensing; subjective experience; mobile assessment

Introduction

Subjective quality of life is a primary concern of persons living with Alzheimer disease and Alzheimer disease-related dementias (AD/ADRD) and their caregivers. Subjective experiences, or an individual's internal and personal perception of events, emotions, or sensations, include important outcomes in AD/ADRD such as well-being, affect, pain, and loneliness. Additional subjective experiences, like loss of pleasure, agitation, and sadness, are among the behavioral symptoms of dementia that lead to the greatest caregiver burden and health care costs. Reliable and valid measurement approaches are thus critical to quantifying these subjective experiences as outcomes in treatment studies. AD/ADRD produce substantial roadblocks to the measurement of subjective experiences, such as cultural differences in how symptoms are expressed, variability in caregiver interpretation of the experiences of patients, or practical usability challenges associated with sensors in older adults. Self-report measures of subjective experiences can make substantial demands for cognitive abilities impaired in AD/ADRD (memory, verbal fluency, comprehension) and can be biased by cultural milieu, social situation, and memories and schemas of the subjective experience in question. In recent years, technological approaches to gathering data about subjective experiences have been developed, implemented, and tested. Passive sensing can be used to gather data proximal to a variety of experiences that would be difficult or impossible to gather without self-report. Moreover, a promise of passive sensors and related devices is their capacity to gather continuous information that could be used to understand within-person dynamics of subjective experiences, and ultimately better personalize interventions to unique timings and contexts. While there is much enthusiasm about the potential for such tools to support subjective measurement, there remain a number of challenges (and opportunities for innovation) to generate translatable and useful passive measurement tools in AD/ADRD. We briefly summarize some of these challenges and provide accompanying suggestions for future innovators in this area.

Passive Sensing Methods

Passive sensing occurs when data are acquired by a device without intentional input from the respondent. Sensing typically uses a variety of devices that can be broadly grouped into three clusters: a) smartphones, b) wearables, and c) in-home sensing devices [1,2]. Common smartphone sensor data streams evaluated in studies of subjective experiences include GPS, accelerometers, Bluetooth, and WiFi data streams. These tools can enable quantification of movement and nearby cell phones, eye movements and facial expressions, and ambient noise and speech features [3]. In addition to sensors, smartphones produce metadata such as calls, texts, app use, and other device interaction data [4]. Also available on mobile devices are more granular device interaction data that has been used to infer

XSL•FO
RenderX

cognitive and emotional processes, such as keyboard analytics like errors or typing speed [5]. Wearable sensors like smartwatches and rings produce data on mobility as well as other physiological parameters, such as heart rate variability, sleep patterns, galvanic skin response, temperature, or respiration rate. In-home sensors have a variety of form factors to include ones that are intended to be "invisible" and interactive ones like robots [6]. In-home sensors include infrared or pressure sensors designed to monitor mobility patterns or other interactions within physical spaces. Audio sensors have been used to gather socially-relevant data such as the quantification of linguistic or acoustic speech markers [7]. Video data includes collection from 3D cameras that can be used to generate facial or skeletal models that quantify deviations from normative trajectories [8]. Increasingly, researchers have attempted to integrate multiple sensor streams via computational approaches [9], and create feedback systems for caregivers leveraging passive sensing data [10]. As consumer-grade devices accumulate greater processing power and more novel and powerful sensors are created, it is easy to see why passive sensing research has burgeoned in recent years.

Challenges in Validating Passive Sensing of Subjective Experiences in ADRD

A first principle in considering and reporting on sensor-based subjective experience research is that subjective experiences are inferred rather than measured directly, as is the case with self-report (see [11] for an excellent summary). Therefore, terminology used in reports should be consistent with the notion that passive sensing is a proxy for but not a direct measure of the construct of interest. This does not mean that sensors cannot be useful for research on subjective constructs. A related point is that the current "gold standard" for validation of passive sensing measures of subjective experience is self-report, which is not free of biases or the impact of cognitive impairments in AD/ADRD. Patients may not recall recent experiences or fully comprehend language and questions. For example, global well-being (*what is your overall well-being?*) can diverge from aggregated immediate experience (*what is your well-being right now?*) measured via ecological momentary assessment (see [12]). It also should not be forgotten that self-report measures can be adapted to people with cognitive impairment or language production difficulties (eg, with visual analog scales [13]), and recent work indicates that active smartphone-based assessments and ecological momentary assessment are feasible and valid in older adults with mild cognitive impairment [14]. Finally, caregiver reports and clinician ratings of patient subjective experiences can also be used as "gold standards" to validate passive sensing measures, but these measures place an added burden on caregivers and providers and are not free of biases or the impact of the observer's capacity to infer patient experiences.

Central to validation of passive sensing measures in AD/ADRD research is this question: What if "gold standard" subjective experience measurement (eg, a questionnaire) is difficult or impossible to obtain due to the effects of cognitive impairment? As Kourtis et al [15] point out that in late-life depression and Alzheimer's disease, multiple passive sensors show promise in detecting multiple subjective experiences, like mood, loneliness, suicide risk, agitation, daily life functioning, and dementia onset and progress, but work is only beginning with these measures. For example, Au-Yeung et al [16] evaluated a home-based mobility sensor in adults with a dementia diagnosis, and documented mobility patterns in an effort to examine within-person dynamics of agitation and apathy. A pilot study by Galambos et al [17] tested an early dementia detection model in older adults, with home-based sensors detecting changes in the amount of time spent in the bedroom, in the living room, and mealtime activities that have congruence with health assessments. Other passive sensing methods use infrared sensors to assess time spent in the house versus out of the house [18] as well as the speed of ambulation through the house [19], which have correlated with cognitive function changes. Another study [20] found that wireless home-based sensors captured differences in activities of daily living patterns between dementia and healthy individuals. Smartphone-based Bluetooth detection of nearby cell phones, actigraphy, and GPS data have been used for predicting loneliness in college students [21] and between Bluetooth features and depression [3], but this work is only beginning in AD/ADRD.

As several reviews of the literature of sensor-based systems of subjective experiences have concluded [1,22] the great majority of research studies to date have been at the "proof of concept" phase. Few studies have been replicated, and few technologies have been evaluated in samples designed for rigorous validation. For research on passive sensing to advance the replicability of sensing, it could be useful to adapt systematic and coordinated methods being used to improve self-report. An enormous amount of effort has gone into developing researcher toolkits like PROMIS (Patient-Reported Outcomes Measurement Information System) and NeuroQoL (Quality of Life in Neurological Disorders) [23]. These National Institutes of Health–sponsored projects spearheaded development of common measures of subjective constructs by rigorous item development, reliability and replicability testing in diverse samples so as to mitigate bias, and application of advanced psychometric techniques toward item selection, short-form development to maximize scalability, along with establishment of convergent and divergent validity specifically in people with diminished cognitive ability (eg, use of visual analog scales). Perhaps some processes involved in these coordinated efforts could serve as models to enhance the utility of passive sensing.

Key roadblocks to creating such a uniform approach to enhancing the replicability of passive sensor data are the intricacies of feature extraction and computational processing of data streams. To move passive sensing toward a more standardized and sharable method, there are open science solutions to enhance transparency, including platforms to harmonize data standards, for example, the Collaborative Aging Research using Technology Initiative (CART) [24]. Repositories for research protocols and other collaborative approaches, such as checklists [25], are available to support best practices in data collection, processing, and validation. Application of Findable, Accessible, Interoperable, and Reusable (FAIR) [26] principles to data collection processes would enable the needed comparison

XSL•FO
RenderX

of the impact of different devices and software versions on outputs. These initiatives should support greater aggregation of results across studies to address fundamental questions about sensing relevant subjective experiences that persist: Which sensor or combination of sensors is most accurate and reliable in sensing aspects of which subjective experiences? How long and at what data density is required to obtain valid results? Which sensors offer the greatest balance of practicality in implementation, unobtrusiveness, cost and data processing demands, and validity and reliability? What are the best practices specific to people with ADRD and their caregivers in maximizing adherence?

While the coming years will undoubtedly bring a plurality of new, more-sensitive sensors, and new approaches to measuring aspects of subjective experience applicable to ADRD, it is essential that this research accompany robust consideration of the principles of the Belmont report in respect to respect for persons, beneficence, and justice. For the same reasons that self-report is hampered, so is informed consent in ADRD. Tools to simplify informed consent and assess and improve decisional capacity are available [24,27], but the technical and privacy considerations surrounding passive sensing (eg, comprehension) can present marked challenges even for people without cognitive impairment. Resources such as guidelines to identify best practices in obtaining proxy or individual informed consent and granular choices surrounding the type of data collected could maximize respect for the individual in the informed consent setting [28]. The industry has a major role in this, including the deployment of health research infrastructure and gatekeeping functions to ensure privacy standards are met and maintained during the research, alongside software updates.

Another piece of the puzzle is engagement in the community in regard to the setting of the sensor and also the results of the research. Few sensors or devices used in the collection of data in sensor studies were designed with older adults in mind. For example, wearable sensors may not be calibrated to age-associated factors such as the thinning of skin, and conclusions from those sensors may be erroneous [29]. Given that caregivers are often essential partners in the deployment and sustainment of data collection, their unique needs and challenges need to be considered [30]. Therefore, part of the researcher's task is to partner with older adults, caregivers, and other stakeholders through user-centered design before deployment [31]. As part of this process, researchers should query participants' desires for the return of information on sensing data, including at the individual level [32]. Community-engaged research approaches may be one way to expand diversity of inclusion in sensor-based research, which will be essential to understand biases that may be present.

Conclusion

Paradoxically, these many challenges to passive sensing in AD/ADRD stem from the very same reasons why such approaches offer promise, pressing unmet needs. The behavioral symptoms of dementia, including agitation or loss of pleasure, are among the single largest drivers of caregiver distress and institutionalization. Making the case for passive sensing in behavioral symptoms even more compelling are the potential risks, costs, and harms of current treatments for behavioral symptoms (eg, side effects, institutionalization). There are emerging systems that infer these subjective symptoms by use of passive sensing alongside contemporaneous dynamic data from potential contributors to the onset and sustainment of these symptoms that could be targets for intervention. Clinical translation opportunities could include more precise monitoring of change in subjective experiences that are targeted by interventions (eg, pharmacologic or nonpharmacologic treatments targeting loss of pleasure). Given the continuous nature of passive sensing, idiographic approaches are possible, wherein personalized interventions could be developed through user-centered design [33] in a just-in-time fashion based on data garnered from the individual. We note that similar to more active digital health solutions, validation is only one step and clinical implementation accompanies many additional challenges (eg, cost, urgent responses, provider burden) [34]. Thus, coordinated efforts at addressing roadblocks to the translation of passive sensing of subjective experience in ADRD could lead to transformative approaches that address critical unmet needs.

Conflicts of Interest

None declared.

References

1. Qirtas MM, Zafeiridi E, Pesch D, White EB. Loneliness and social isolation detection using passive sensing techniques: scoping review. JMIR Mhealth Uhealth 2022 Apr 12;10(4):e34638. [doi: 10.2196/34638] [Medline: 35412465]
2. Sheikh M, Qassem M, Kyriacou PA. Wearable, environmental, and smartphone-based passive sensing for mental health monitoring. Front Digit Health 2021;3:662811. [doi: 10.3389/fdgth.2021.662811] [Medline: 34713137]
3. Zhang Y, Folarin AA, Sun S, et al. Predicting depressive symptom severity through individuals' nearby bluetooth device count data collected by mobile phones: preliminary longitudinal study. JMIR Mhealth Uhealth 2021 Jul 30;9(7):e29840. [doi: 10.2196/29840]
4. Wang Y, Ren X, Liu X, Zhu T. Examining the correlation between depression and social behavior on smartphones through usage metadata: empirical study. JMIR Mhealth Uhealth 2021;9(1):e19046. [doi: 10.2196/19046]
5. Vesel C, Rashidisabet H, Zulueta J, et al. Effects of mood and aging on keystroke dynamics metadata and their diurnal patterns in a large open-science sample: a BiAffect iOS study. J Am Med Inform Assoc 2020 Jul 1;27(7):1007-1018. [doi: 10.1093/jamia/ocaa057] [Medline: 32467973]

XSL•FO
RenderX

6. Van Patten R, Keller AV, Maye JE, et al. Home-based cognitively assistive robots: maximizing cognitive functioning and maintaining independence in older adults without dementia. Clin Interv Aging 2020:1129. [doi: 10.2147/CIA.S253236]
7. Asci F, Costantini G, Di Leo P, et al. Machine-learning analysis of voice samples recorded through smartphones: the combined effect of ageing and gender. Sensors (Basel) 2020;20(18):5022. [doi: 10.3390/s20185022]
8. Du C, Graham S, Depp C, Nguyen T. Assessing physical rehabilitation exercises using graph convolutional network with self-supervised regularization. In: Du C, Graham S, Depp C, Nguyen T, editors. Presented at: 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). [doi: 10.1109/EMBC46164.2021.9629569]
9. Graham S, Depp C, Lee EE, et al. Artificial intelligence for mental health and mental illnesses: an overview. Curr Psychiatry Rep 2019 Nov;21(11):1-18. [doi: 10.1007/s11920-019-1094-0]
10. Thomas NW, Bhati DK, Ault L, et al. Smart Home Tech: an interface to provide feedback to caregivers of persons living with cognitive impairment. Alzheimers Dement 2022 Dec;18(S8):e067365. [doi: 10.1002/alz.067365]
11. LeDoux JE, Hofmann SG. The subjective experience of emotion: a fearful view. Curr Opin Behav Sci 2018 Feb;19:67-72. [doi: 10.1016/j.cobeha.2017.09.011] [Medline: 39355512]
12. Stone AA, Bachrach CA, Jobe JB, Kurtzman HS, Cain VS. The Science of Self-Report: Implications for Research and Practice: Psychology Press; 1999.
13. Horgas AL, Elliott AF, Marsiske M. Pain assessment in persons with dementia: relationship between self-report and behavioral observation. J Am Geriatr Soc 2009 Jan;57(1):126-132. [doi: 10.1111/j.1532-5415.2008.02071.x] [Medline: 19054191]
14. Moore RC, Ackerman RA, Russell MT, et al. Feasibility and validity of ecological momentary cognitive testing among older adults with mild cognitive impairment. Front Digit Health 2022;4. [doi: 10.3389/fdgth.2022.946685] [Medline: 35990012]
15. Kourtis LC, Regele OB, Wright JM, Jones GB. Digital biomarkers for Alzheimer's disease: the mobile/ wearable devices opportunity. NPJ Digit Med 2019;2:31119198. [doi: 10.1038/s41746-019-0084-2] [Medline: 31119198]
16. Au-Yeung WTM, Miller L, Beattie Z, et al. Monitoring behaviors of patients with late-stage dementia using passive environmental sensing approaches: a case series. Am J Geriatr Psychiatry 2022 Jan;30(1):1-11. [doi: 10.1016/j.jagp.2021.04.008] [Medline: 34039534]
17. Galambos C, Skubic M, Wang S, Rantz M. Management of dementia and depression utilizing in- home passive sensor data. Gerontechnology 2013 Jan 1;11(3):457-468. [doi: 10.4017/gt.2013.11.3.004.00] [Medline: 24049513]
18. Petersen J, Austin D, Mattek N, Kaye J. Time out-of-home and cognitive, physical, and emotional wellbeing of older adults: a longitudinal mixed effects model. PLoS One 2015;10(10):e0139643. [doi: 10.1371/journal.pone.0139643] [Medline: 26437228]
19. Kaye J, Mattek N, Dodge H, et al. One walk a year to 1000 within a year: continuous in-home unobtrusive gait assessment of older adults. Gait Posture 2012 Feb;35(2):197-202. [doi: 10.1016/j.gaitpost.2011.09.006] [Medline: 22047773]
20. Urwyler P, Stucki R, Rampa L, Müri R, Mosimann UP, Nef T. Cognitive impairment categorized in community-dwelling older adults with and without dementia using in-home sensors that recognise activities of daily living. Sci Rep 2017 Feb 8;7:42084. [doi: 10.1038/srep42084] [Medline: 28176828]
21. Wu C, Barczyk AN, Craddock RC, et al. Improving prediction of real-time loneliness and companionship type using geosocial features of personal smartphone data. Smart Health (2014) 2021 Apr;20:100180. [doi: 10.1016/j.smhl.2021.100180]
22. Husebo BS, Heintz HL, Berge LI, Owoyemi P, Rahman AT, Vahia IV. Sensing technology to monitor behavioral and psychological symptoms and to assess treatment response in people with dementia. a systematic review. Front Pharmacol 2019;10:1699. [doi: 10.3389/fphar.2019.01699] [Medline: 32116687]
23. Cella D, Nowinski C, Peterman A, et al. The neurology quality-of-life measurement initiative. Arch Phys Med Rehabil 2011 Oct;92(10 Suppl):S28-S36. [doi: 10.1016/j.apmr.2011.01.025] [Medline: 21958920]
24. Beattie Z, Miller LM, Almirola C, et al. The Collaborative Aging Research Using Technology Initiative: an open, sharable, technology-agnostic platform for the research community. Digit Biomark 2020;4(Suppl 1):100-118. [doi: 10.1159/000512208] [Medline: 33442584]
25. Nebeker C, Bartlett Ellis RJ, Torous J. Development of a decision-making checklist tool to support technology selection in digital health research. Transl Behav Med 2020 Oct 8;10(4):1004-1015. [doi: 10.1093/tbm/ibz074] [Medline: 31120511]
26. Wilkinson MD, Dumontier M, Aalbersberg IJJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 2016 Mar 15;3(1):160018. [doi: 10.1038/sdata.2016.18] [Medline: 26978244]
27. Parmigiani G, Del Casale A, Mandarelli G, et al. Decisional capacity to consent to treatment and research in patients affected by mild cognitive impairment. A systematic review and meta-analysis. Int Psychogeriatr 2022 Jun;34(6):529-542. [doi: 10.1017/S1041610220004056] [Medline: 33583459]
28. Bartlett Ellis R, Wright J, Miller LS, et al. Lessons learned: beta-testing the digital health checklist for researchers prompts a call to action by behavioral scientists. J Med Internet Res 2021 Dec 22;23(12):e25414. [doi: 10.2196/25414] [Medline: 34941548]
29. Graham SA, Jeste DV, Lee EE, et al. Associations between heart rate variability measured with a wrist-worn sensor and older adults' physical function: observational study. JMIR Mhealth Uhealth 2019 Oct 23;7(10):e13757. [doi: 10.2196/13757] [Medline: 31647469]

XSL•FO
RenderX

30. Zhai S, Chu F, Tan M, Chi NC, Ward T, Yuwen W. Digital health interventions to support family caregivers: an updated systematic review. Digit Health 2023;9:20552076231171967. [doi: 10.1177/20552076231171967] [Medline: 37223775]
31. Tiersen F, Batey P, Harrison MJC, et al. Smart home sensing and monitoring in households with dementia: user-centered design approach. JMIR Aging 2021;4(3):e27047. [doi: 10.2196/27047]
32. Shen FX, Baum ML, Martinez-Martin N, et al. Returning individual research results from digital phenotyping in psychiatry. Am J Bioeth 2024 Feb;24(2):69-90. [doi: 10.1080/15265161.2023.2180109]
33. Nahum-Shani I, Smith SN, Spring BJ, et al. Just-in-Time Adaptive Interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. Ann Behav Med 2018 May 18;52(6):446-462. [doi: 10.1007/s12160-016-9830-8]
34. Koh J, Tng GYQ, Hartanto A. Potential and pitfalls of mobile mental health apps in traditional treatment: an umbrella review. JPM 2022 Aug 25;12(9):1376. [doi: 10.3390/jpm12091376]

Abbreviations

AD/ADRD: Alzheimer disease and related dementias

Edited by F Puga; submitted 24.02.25; peer-reviewed by A Hartanto, WT Au-Yeung; revised version received 15.10.25; accepted 03.11.25; published 18.11.25.

Please cite as:
Depp C, Holden J, Granholm E
Digital Measurement of Subjective Experiences in Alzheimer Disease and Related Dementias (AD/ADRD)
JMIR Aging 2025;8:e71920
URL: https://aging.jmir.org/2025/1/e71920
doi:10.2196/71920

© Colin Depp, Jason Holden, Eric Granholm. Originally published in JMIR Aging (https://aging.jmir.org), 18.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

A Digitally Capable Aged Care Workforce: Demands and Directions for Workforce Education and Development

Kathleen Gray[1], PhD; Kerryn Butler-Henderson[2], PhD; Karen Day[3], PhD

[1]Centre for Digital Transformation of Health, University of Melbourne, Melbourne, Australia

[2]School of Nursing, Paramedicine, and Healthcare Sciences, Charles Sturt University, Bathurst, Australia

[3]Health Systems Group, University of Auckland, Auckland, New Zealand

Corresponding Author:
Kathleen Gray, PhD
Centre for Digital Transformation of Health
University of Melbourne
Parkville
Melbourne, 3010
Australia
Phone: 61 390355511
Email: kgray@unimelb.edu.au

Abstract

As the aged care sector undergoes digital transformation, greater attention is needed to development of digital health capability in its workforce. There are many gaps in our understanding of the current and future impacts of technology on those who perform paid and unpaid aged care work. Research is needed to understand how to make optimal use of both digital resources and human resources for better aged care. In this Viewpoint, we reflect on a workshop held during an international conference that identified shared concepts and concerns to shape further research into workforce capability. Digital technologies and digital data can increase quality of care in a system that operates through partnerships among service providers, service users, and community members. To realize this potential, digital health learning and development are needed in the aged care workforce. As digital dimensions of aged care services expand, the sector needs clearer direction to implement approaches to workforce learning and development. These must be appropriate to support the safe and ethical performance of care work and to increase the satisfaction of those who care and those for whom they care.

(JMIR Aging 2025;8:e54143) doi:10.2196/54143

KEYWORDS

aged care; digital health; digital literacy; education; older adults; professional development; digital transformation; digital resources; users; community; learning; support; safe; ethical; satisfaction

Why Focus on a Digitally Capable Aged Care Workforce?

The aged care sector is in the spotlight around the world due to the increasing health and well-being needs of aging populations. Examples are the Australian Royal Commission into Aged Care Quality and Safety [1], Japan's Social Welfare for the Elderly Act [2], and Germany's national report examining the human rights of older persons in long-term care [3]. Comparisons of aged care systems internationally focus on a variety of issues [4,5].

In such reports, the aged care workforce emerges as an essential factor. One part of this workforce is paid, structured, and regulated (eg, registered nurses) and formally counted in the health and care workforce. Another part of this workforce is informal—often unpaid and not structured or regulated (eg, relatives as unpaid caregivers) but contributing in incalculable ways [6]. Both parts operate under conditions that can be inequitable, inefficient, uncertain, and unsatisfactory. The overall situation of the aged care workforce demands widespread improvements in recruitment and retention, practice standards, skill levels, recognition, and reward [7].

Digital maturity in aged care service provider organizations lags behind mainstream health care services; there are differences between the two also in terms of information needs, technologies of interest, and administration. The medico-legal record for aged care services works differently from hospitals and primary care services, making data and information standardization problematic [8]. Robotics, sensors, and virtual reality have been

XSL•FO
RenderX

explored for over a decade in many countries for their potential to augment human care-giving, but with little opportunity for uptake due to cost and complexity [9]. Telehealth has been shown to have a promising place in aged care, especially since the COVID-19 pandemic, but funding and service supply remain problematic [10]. Nevertheless, digital transformation of the aged care sector is underway [11,12].

Despite this, digital skills of any kind rarely appear in aged care job descriptions or training programs. Further, national workforce strategies related to digital health capability scarcely address the digital skill shortages in the aged care sector. On one hand, there is hope that the sector will find the "right" staff for digitally enabled aged care and find ways to assure them of technology's quality and trustworthiness [13,14]. On the other hand, there is uneasiness that roles and responsibilities may become invisible, peripheral, or "care-adjacent" in the process of moving to technological efficiency and sustainability [15,16]. In short, the introduction of technology has still-unknown implications for the evolution of work in the aged care sector, with possible effects ranging from displacing workers to supporting productivity in current roles or to elevating responsibilities and recognition. However, much of the research to date into aged carers' digital experiences and expectations has had a short-term focus on technology implementations rather than consequences and outcomes [17,18].

Working with technology in particular ways can support specific well-being factors, including positive emotions, self-awareness, mindfulness, empathy, and compassion—so-called positive computing [19]—and there are reports that this approach can benefit aged care organizations and their workers [20]. But there is no guarantee that this will happen without a deliberate workforce strategy. This raises questions about how to proceed, which shape the aims of further research: How can digital transformation of the aged care sector be optimized through greater investment in the digital knowledge, skills, attributes, and attitudes of its workforce? How can such investment in turn improve working conditions and job satisfaction in the aged care workforce?

The authors explored these questions during a workshop that we facilitated at the July 2023 MedInfo conference in Sydney, Australia. There we presented our view that workforce research and development is needed to achieve optimal outcomes for workers in order to achieve these for service users as the aged care sector moves toward digital maturity. The ensuing discussions with workshop delegates contributed to refining the viewpoint we outline here.

What Knowledge and Skill Building Is Required for Digital Aged Care Work?

Digital aged care work is likely to deliver the greatest benefits when needs for fundamental, complementary digital knowledge and skills are met in both paid and unpaid workers. Table 1 summarizes and juxtaposes the capabilities that we propose.

Table 1. Digital capabilities needed in the aged care workforce.

For paid work	For unpaid work
Knowing about developments in the digital environment in which health and care operate	Knowing about the existence of digital systems and tools, and the ways they can support care users and workers
Knowing how to adopt digital tools to help with care coordination and monitoring	Knowing how to use tools to support care closer to home and enhance care provision
Knowing the importance for care quality of collecting and interpreting digital data collected during episodes of care	Knowing how to navigate a complex aged care system across multiple digital platforms
Ability with general digital proficiency and literacy, including information security	Ability with general digital proficiency and literacy
Ability to evaluate online information critically	Ability to access and interpret information from multiple online sources
Ability to consider usability and empathy when using digital tools with patients and clients	Ability to learn to engage with technology and be comfortable with technology
Ability to interpret biometric data and help the patient or client to understand it	Ability to be connected and share health data between care consumers and providers
Ability to judge when technology is not needed as well as when it is	Ability to access and use technology to overcome isolation

There are parallel aspects of capability, with different levels of sophistication, in the needs of paid and unpaid workers. Accordingly, paid workers must be cognizant of not only their own capabilities, but also the capabilities of the unpaid workers with whom they partner. Aside from the digital capabilities of direct consumers of aged care services, paid workers must consider and collaborate with the capabilities of volunteer carers among family, friends and community members. Thus, to build the digitally capable human resources that will support healthy and dignified aging for all, the sector certainly must increase levels of digital literacy in the paid workforce; digital literacy being defined as the ability to use technology to communicate, retrieve, and evaluate information to make decisions [21]. However, the sector also must increase levels of active engagement in digital citizenship among both paid and unpaid workers; in the concept of digital citizenship, people are increasingly expected to use a range of technologies to access services of all kinds and to participate on all levels in society [22].

Various approaches are possible to support digital knowledge building and skill building for the paid workforce. There seem to be gaps and requirements in terms of both formally accredited

XSL•FO
RenderX

training programs and less formal continuing professional development, as well as for both workplace-based learning and workplace-independent learning. The diverse and differentiated methods needed to uplift the digital capabilities of the unpaid workforce are an even broader challenge. Strategy and culture will have to change in the aged care sector to create an environment that encourages and supports all those who do the work to embrace learning and development for positive uses of technology. Among service providers, technology suppliers, industry and community associations, and government agencies, it is not yet clear where roles and responsibilities should sit for planning, implementing, and evaluating the changes that are needed to create and support a digitally capable aged care workforce.

Although these themes are similar to those arising in the mainstream health care workforce in some respects, they also recognize that the aged care ecosystem is distinctive and heterogeneous in terms of the array of partnerships among service providers, service users, and community members [23]. It may be useful to build a more detailed understanding of the digital capability gaps and needs in this sector by overlaying two other system-level concepts onto the idea of the aged care ecosystem—care ecosystems and workforce ecosystems. A care ecosystem is defined as an approach to configuring paid and unpaid workers, ambient assistive technologies, and digital data sharing as work-arounds for intractable workforce shortages in care, particularly in situations that do not call for advanced or intensive clinical services [24,25]. A workforce ecosystem is characterized by employees and contractors working interdependently with automated processes to achieve the goals of an organization and their personal goals; it is considered a guiding concept for the challenges of human resources management as artificial intelligence changes industries and societies [26,27]. A whole-system approach that acknowledges these concepts could provide a coherent framework to plan, implement, and evaluate the range of complementary measures that are needed for digitally enabled aged care work.

Conclusions and Next Steps

Digital technologies and digital data have the potential to improve many aspects of the experiences of the people who receive aged care and the people who provide it. The context in which we seek to develop the digital capabilities of the aged care workforce is challenging. There are strong sustainability pressures on service providers. Public policies, strategies, and sentiments favor innovative, distributed care. Social and technological services and solutions are not joined up. Workers across the sector are stretched thin, whether they are highly trained professionals (eg, registered nurses), minimally trained caregivers (eg, health and personal care assistants), or informal caregivers (eg families, friends, and neighbours).

Understanding how to build digital capability in this workforce is complex and multifaceted, with many unknowns. Inclusion is one: What learning and development methods are most efficient and effective to enable digital skills across the entirety of this workforce? Empowerment is another: How can aged care workers be enabled to advocate for and implement digital improvements that bring positive changes to care? Expertise is another: What career opportunities and leadership roles are appropriate to motivate digital aged care specialists and sustain their attention to workforce development?

Next steps for the work in this field include reviewing the literature to distill evidence from reports of small-scale efforts, workforce research, and development that engages real-world settings and stakeholders, as well as trials of scalable learning and professional and community development programs. To progress this research requires new collaborations among aged care workers, service provider organizations, aged care technology vendors, and digital health workforce researchers. We can improve the current standing of digital skills in the aged care workforce only if we first clarify what the workforce needs from education providers, employers, technology suppliers, and policy makers to enable and motivate them to use existing and emerging technologies safely and wisely. This work will provide important information to design learning and development that enhances digital capability, and it will establish expectations that allow us to evaluate changes in workforce status and performance as an integral part of the digital transformation of aged care.

Acknowledgments

We are grateful to the convenors of the MedInfo 2023 conference for the opportunity to conduct our workshop and to the delegates who attended and engaged in discussion. The 3 authors prepared the workshop together. To set up the workshop discussion, each author presented on a viewpoint of the topic, followed by KB-H and KD cochairing the discussion. All 3 authors contributed to writing and editing this Viewpoint and signed off the paper as ready for submission.

Conflicts of Interest

None declared.

References

1. Final report: care, dignity and response. Royal Commission into Aged Care Quality and Safety. 2021. URL: https://www.royalcommission.gov.au/aged-care/final-report [accessed 2025-03-28]

XSL•FO
RenderX

2. Nakamura S. Japan's welfare system for the elderly: past, present, and future. Asia Health and Wellbeing Initiative. URL: https://www.ahwin.org/wp-content/uploads/2020/06/201812-S-Nakamura-Japan%E2%80%99s-Welfare-System-for-the-Elderly-.pdf [accessed 2025-03-28]
3. Aronson P, Mahler C. Human rights of older persons in long-term care: German national report. European Network of National Human Rights Institutions. 2016. URL: https://ennhri.org/wp-content/uploads/2019/10/germany.pdf [accessed 2025-03-28]
4. Review of international systems for long-term care of older people. Research paper 2. Royal Commission into Aged Care Quality and Safety. 2020. URL: https://apo.org.au/node/274151 [accessed 2025-04-02]
5. Appendix D: international experience – inquiry report – caring for older Australians. Australian Government Productivity Division. 2011. URL: https://www.pc.gov.au/inquiries/completed/aged-care/report/26-aged-care-appendixd.pdf [accessed 2025-03-28]
6. Woolford M, Joyce C, Polacsek M. Strategies to optimize aged care volunteer recruitment and retention: a systematic review. J Gerontol Soc Work 2022 Jan;65(1):45-62. [doi: 10.1080/01634372.2021.1923605] [Medline: 33971797]
7. Who cares? Attracting and retaining care workers for the elderly. Organisation for Economic Co-operation and Development. 2020. URL: https://www.oecd.org/en/publications/who-cares-attracting-and-retaining-elderly-care-workers_92c0ef68-en.html [accessed 2025-03-28]
8. Bail K, Gibson D, Acharya P, Blackburn J, Kaak V, Kozlovskaia M, et al. Using health information technology in residential aged care homes: an integrative review to identify service and quality outcomes. Int J Med Inform 2022 Sep;165:104824 [FREE Full text] [doi: 10.1016/j.ijmedinf.2022.104824] [Medline: 35792376]
9. Penteridis L, D'Onofrio G, Sancarlo D, Giuliani F, Ricciardi F, Cavallo F, et al. Robotic and sensor technologies for mobility in older people. Rejuvenation Res 2017 Oct;20(5):401-410. [doi: 10.1089/rej.2017.1965] [Medline: 28482748]
10. Day K, Millner S, Johnson H. How nurses use telehealth to support health transitions of older adults. Stud Health Technol Inform 2016;231:23-30. [doi: 10.3233/978-1-61499-712-2-23] [Medline: 27782013]
11. Nordin S, Sturge J, Ayoub M, Jones A, McKee K, Dahlberg L, et al. The role of information and communication technology (ICT) for older adults' decision-making related to health, and health and social care services in daily life-a scoping review. Int J Environ Res Public Health 2021 Dec 23;19(1):151 [FREE Full text] [doi: 10.3390/ijerph19010151] [Medline: 35010408]
12. Fasoli A, Beretta G, Pravettoni G, Sanchini V. Mapping emerging technologies in aged care: results from an in-depth online research. BMC Health Serv Res 2023 May 23;23(1):528 [FREE Full text] [doi: 10.1186/s12913-023-09513-5] [Medline: 37221528]
13. Peters M, Marnie C. Human costs of aged care productivity: innovation versus staffing and skills mix. Collegian 2021 Jun;28(3):351-355 [FREE Full text] [doi: 10.1016/j.colegn.2020.11.002]
14. Wang Y, Feng Y, Liang C. How digital change and innovation in the workplace affect front-line employee retention: a cross-sectional study based on the aged care industry. In: Tu Y, Chi M, editors. E-Business. Digital Empowerment for an Intelligent Future. WHICEB 2023. Lecture Notes in Business Information Processing, vol 480. Cham, Switzerland: Springer; 2023:72-83.
15. Hamblin KA. Technology in care systems: displacing, reshaping, reinstating or degrading roles? New Technol Work Employ 2022 Mar;37(1):41-58 [FREE Full text] [doi: 10.1111/ntwe.12229] [Medline: 35911255]
16. Boeva Y, Berger A, Bischof A, Doggett O, Heuer H, Jarke J, et al. Behind the scenes of automation: ghostly care-work, maintenance, and interferences: exploring participatory practices and methods to uncover the ghostly presence of humans and human labor in automation. In: Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems. New York, NY: Association for Computing Machinery; 2023:1-5.
17. Rigby K, Redley B, Hutchinson AM. Change agent's role in facilitating use of technology in residential aged care: a systematic review. Int J Med Inform 2023 Nov;179:105216 [FREE Full text] [doi: 10.1016/j.ijmedinf.2023.105216] [Medline: 37734272]
18. Bail K, Merrick E, Gibson D, Hind A, Strickland K, Redley B. A co-designed health information system implementation into residential aged care: A mixed-method evaluation. J Nurs Scholarsh 2023 Mar;55(2):521-535. [doi: 10.1111/jnu.12840] [Medline: 36366792]
19. Calvo R, Peters D. Positive Computing: Technology for Wellbeing and Human Potential. Cambridge, MA: MIT Press; 2014.
20. Shearman P. The manifestation of positive computing in the aged care workforce. University of Technology Sydney Faculty of Business [PhD thesis]. 2021. URL: https://opus.lib.uts.edu.au/handle/10453/158673 [accessed 2025-03-28]
21. Nichols T, Stornaiuolo A. Assembling "digital literacies": contingent pasts, possible futures. Media Commun 2019;7(2):14 [FREE Full text] [doi: 10.17645/mac.v7i2.1946]
22. Holgersson J, Söderström E, Rose J. Digital inclusion for elderly citizens for a sustainable society. In: Proceedings of the 27th European Conference on Information Systems (ECIS), Stockholm & Uppsala, Sweden, June 8-14, 2019. Atlanta, GA: Association for Information Systems; 2019.
23. Digital transformation sector partners co-design meeting (2). Department of Health Australia. URL: https://www.health.gov.au/sites/default/files/documents/2022/07/digital-transformation-sector-partners-slide-presentation-6-june-2022.pdf [accessed 2025-03-28]

XSL•FO
RenderX

24. Baldissera TA, Camarinha-Matos LM, De Faveri C. Human factor in designing an elderly care ecosystem. In: Adewumi A, Misra S, editors. Handbook of Research on the Role of Human Factors in IT Project Management. Hershey, PA: IGI Global; 2020:106-131.
25. Tseng H, Huang H, Hsieh C. Active aging AI community care ecosystem design. In: Gao Q, Zhou J, editors. Human Aspects of IT for the Aged Population. Healthy and Active Aging. HCII 2020. Lecture Notes in Computer Science(), vol 12208. Cham, Switzerland: Springer; 2020.
26. Altman E, Kiron D, Schwartz J, Jones R. The future of work is through workforce ecosystems. MIT Sloan Management Review. 2021. URL: https://sloanreview.mit.edu/article/the-future-of-work-is-through-workforce-ecosystems/ [accessed 2025-03-28]
27. Gorski AT, Gligorea I, Gorski H, Oancea R. Workforce and workplace ecosystem – challenges and opportunities in the age of digital transformation and 4IR. Int Conf Knowl Based Org 2022 Jun;28(1):187-194 [FREE Full text] [doi: 10.2478/kbo-2022-0028]

Edited by J Wang; submitted 30.10.23; peer-reviewed by M Kleib, A Azizi; comments to author 23.01.24; revised version received 29.04.24; accepted 12.12.24; published 02.04.25.

Please cite as:
Gray K, Butler-Henderson K, Day K
A Digitally Capable Aged Care Workforce: Demands and Directions for Workforce Education and Development
JMIR Aging 2025;8:e54143
URL: https://aging.jmir.org/2025/1/e54143
doi:10.2196/54143
PMID:

©Kathleen Gray, Kerryn Butler-Henderson, Karen Day. Originally published in JMIR Aging (https://aging.jmir.org), 02.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

Advancing Emergency Care With Digital Twins

Haoran Li[1], BSc; Jingya Zhang[1], BSc; Ning Zhang[2], PhD; Bin Zhu[3], PhD

[1]School of Public Policy and Administration, Xi'an Jiaotong University, Xi'an, China

[2]Vanke School of Public Health, Tsinghua University, Beijing, China

[3]School of Public Health and Emergency Management, Southern University of Science and Technology, Shenzhen, China

Corresponding Author:
Bin Zhu, PhD
School of Public Health and Emergency Management
Southern University of Science and Technology
1088 Xueyuan Avenue
Nanshan District
Shenzhen, 518000
China
Phone: 86 13530405020
Email: zhub6@sustech.edu.cn

Abstract

Digital twins—dynamic and real-time simulations of systems or environments—represent a paradigm shift in emergency medicine. We explore their applications across prehospital care, in-hospital management, and recovery. By integrating real-time data, wearable technology, and predictive analytics, digital twins hold the promise of optimizing resource allocation, advancing precision medicine, and tailoring rehabilitation strategies. Moreover, we discuss the challenges associated with their implementation, including data resolution, biological heterogeneity, and ethical considerations, emphasizing the need for actionable frameworks that balance innovation with data governance and public trust.

(JMIR Aging 2025;8:e71777) doi:10.2196/71777

KEYWORDS

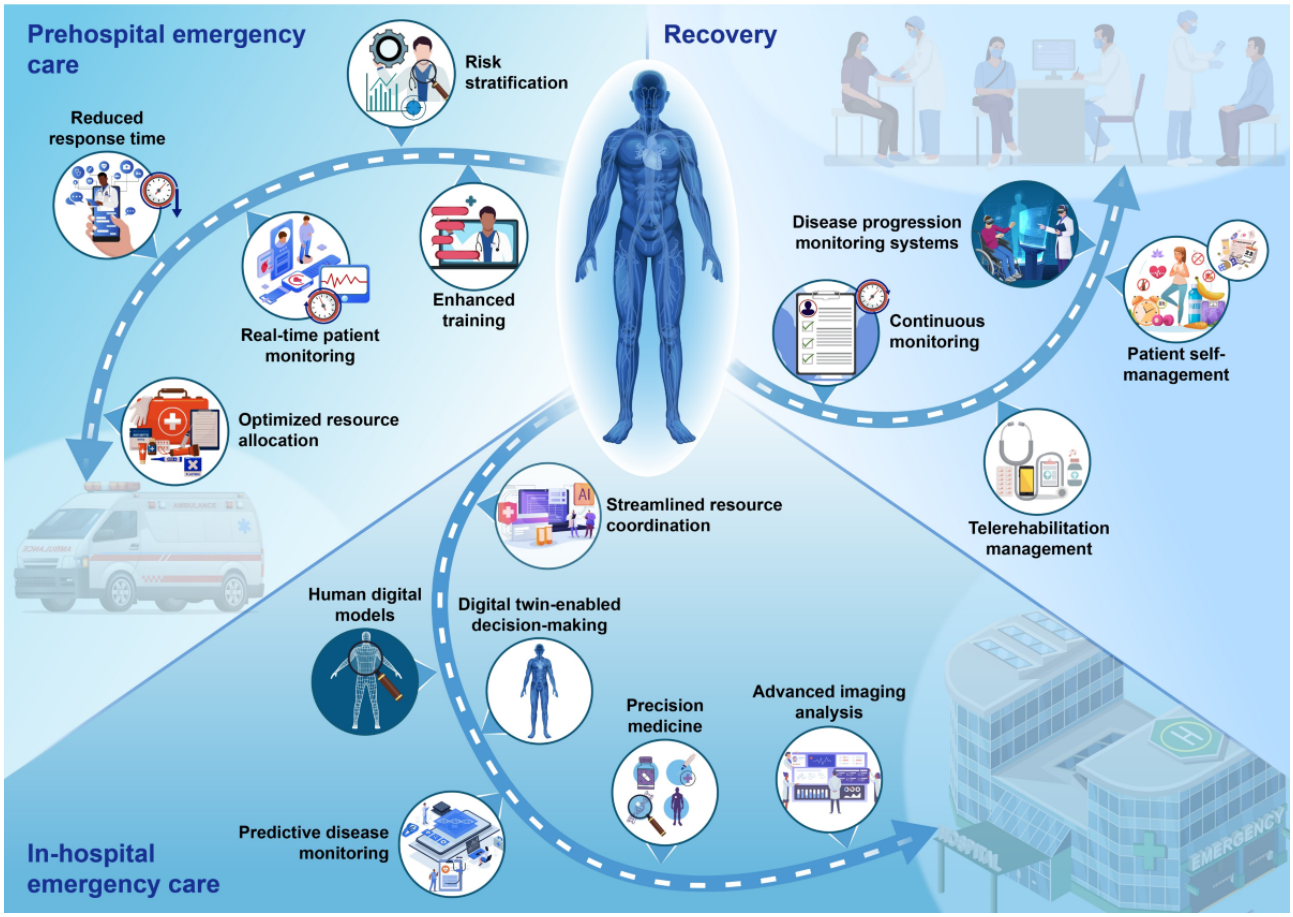
emergency care; digital twin; prehospital emergency care; in-hospital emergency care; recovery

Introduction

The concept of digital twins—dynamic, real-time simulations of processes, systems, or environments—has garnered increasing attention across various domains, from urban planning to medical monitoring [1]. Unlike predictive analytics and artificial intelligence (AI)–driven simulations, digital twins construct dynamic mirror models of physical entities through real-time data synchronization and simulation technologies, emphasizing virtual to physical mapping and closed-loop feedback to support full lifecycle management. In contrast, predictive analytics focus on trend forecasting based on historical data, whereas AI-driven simulations concentrate on multiscenario modeling of complex systems. The advantages of digital twins lie in their real-time capabilities, dynamic optimization, and potential for cross-system collaboration, enabling a closed-loop process from monitoring to decision-making, significantly enhancing efficiency and precision [2]. Emergency care, characterized by its demand for timeliness, precision, and adaptability, provides an ideal testing ground for this transformative technology. By bridging prehospital care, in-hospital management, and recovery, digital twins promise to reshape emergency medical systems for the future (Figure 1). Textbox 1 provides an explanation of the application of digital twins in prehospital first aid, in-hospital first aid, and rehabilitation.

XSL•FO
RenderX

Figure 1. The role of digital twins in transforming emergency medicine.



XSL•FO
RenderX

Textbox 1. The application of digital twins in prehospital first aid, in-hospital first aid, and rehabilitation.

Prehospital emergency care

- Optimized resource allocation: simulation of real-time first-aid scenarios to assess urban emergency needs and spatial distribution of resources, allowing for evidence-based planning and improved efficiency in resource use
- Reduced response time: minimization of procedural delays through real-time data sharing, enabling seamless communication and critical data updates across emergency personnel and systems
- Real-time patient monitoring: leveraging of wearable devices and sensors to create dynamic health profiles, facilitating immediate assessments and predictions of patient conditions (eg, vital signs such as heart rate and blood pressure)
- Risk stratification: aggregation and analysis of patient data to provide superior accuracy in predicting acute events (eg, cardiac arrests), outperforming traditional clinical risk assessments
- Enhanced training: use of simulation-based training environments to strengthen emergency response protocols, improve decision-making, and foster team coordination under high-pressure conditions

In-hospital emergency care

- Advanced imaging analysis: integration of multidimensional data to overcome sample limitations in medical imaging, enabling the development of robust diagnostic models
- Precision medicine: combination of genomic, environmental, and lifestyle data for tailored risk assessments and personalized treatment plans, empowering informed clinical decision-making
- Human digital models: use of digital replicas of anatomical structures to plan surgeries and minimize risks during procedures by simulating operations beforehand
- Digital twin–enabled decision-making: facilitation of real-time analysis and expert-guided interventions (eg, cardiopulmonary resuscitation and drug administration) via remote systems, reducing disparities in emergency care accessibility
- Predictive disease monitoring: use of Internet of Things devices and wearables to continuously monitor patient status, identifying early signs of complications and supporting timely interventions
- Streamlined resource coordination: ensuring efficient resource distribution and effective communication among health care teams through unified, up-to-date patient data sharing

Recovery

- Continuous monitoring: use of sensors and wearable devices to track recovery progress and dynamically adjust treatment plans to optimize outcomes
- Telerehabilitation management: enabling remote monitoring of patients through digital twins, providing clinicians with tools to assess and adjust rehabilitation remotely, particularly for mobility-impaired individuals
- Patient self-management: sharing digital twin insights with patients to encourage active participation in health management through lifestyle modifications and adherence to care plans
- Disease progression monitoring systems: integration of real-time monitoring data to predict the risk of disease exacerbations or complications, thereby intervening in time and triggering early warnings, allowing physicians to adjust treatment during critical windows and improve outcomes.

Prehospital Emergency Care

Current prehospital emergency care systems often struggle with suboptimal resource allocation due to their reliance on simplified decision-making models that fail to account for the complex, dynamic nature of urban environments [3]. Digital twins address these limitations by unifying diverse data streams into a single, adaptable model that enables real-time optimization of emergency resource allocation. These models incorporate multiple layers of data, from historical emergency call patterns to real-time traffic conditions, enabling more nuanced and effective resource deployment decisions.

Digital twins prove particularly valuable in prehospital emergency care, where time-sensitive decisions directly impact patient outcomes. By incorporating high-resolution information about building layouts, land use, population density, and traffic patterns, digital twins create actionable urban simulations that allow emergency response teams to plan and execute their operations with unprecedented efficiency. The integration of these various data sources enables more sophisticated approaches to emergency resource management than previously possible with traditional systems.

This transformation in emergency care delivery demonstrates the broader potential of digital twin technology to revolutionize urban health care systems through data-driven, predictive approaches to resource management. Moreover, wearables in ambulances provide dynamic health profiles, supporting real-time monitoring and accurate risk stratification for acute events. Simulation-based training environments can help improve emergency team readiness and coordination under high-pressure conditions.

XSL•FO
RenderX

In-Hospital Emergency Care

In hospitals, digital twins are redefining precision medicine by integrating diverse data sources to enable robust diagnostics, personalized treatments, and streamlined workflows [4]. By synthesizing real-time patient data with historical health records and predictive algorithms, digital twins facilitate more accurate disease modeling, enhancing clinical decision-making and tailoring interventions to individual patient profiles. This data-driven approach supports the early detection of complications, optimization of treatment pathways, and improvement of patient outcomes.

Preoperative simulations represent a critical application of digital twins in surgical care. By creating highly accurate digital replicas of patient anatomy, clinicians can simulate surgical procedures, anticipate potential challenges, and refine techniques before making the first incision. These virtual rehearsals significantly reduce surgical risks, improve procedural efficiency, and enhance patient safety. Complementing this, predictive monitoring through Internet of Things (IoT) devices enables the continuous assessment of patient vital signs, offering early warnings of complications such as sepsis or cardiac arrest. These alerts allow medical teams to respond swiftly, averting adverse outcomes and saving lives.

An emerging and transformative application is the development of medical device twins (MDTs), which provide virtual counterparts for critical hospital equipment, including ventilators, defibrillators, infusion pumps, and imaging devices. MDTs are created through digital twin technology and are consistent with the actual medical device in terms of function, performance, and status. They can be updated in real time to reflect the actual operation of the medical device. These digital twins continuously monitor device performance, tracking metrics such as wear and tear, calibration accuracy, and operational efficiency. By identifying potential failures before they occur, MDTs facilitate proactive maintenance, reducing downtime and ensuring the reliability of life-saving equipment during critical emergencies.

Beyond maintenance, MDTs contribute to resource optimization. By analyzing use patterns and forecasting demand surges, these systems can dynamically allocate equipment to areas of highest need, such as during mass casualty incidents or seasonal patient influxes. This predictive capability not only prevents bottlenecks in care delivery but also enhances hospital operational efficiency, ensuring that critical resources are always available when and where they are needed.

Recovery

In recovery, digital twins play a transformative role by enabling personalized and adaptive rehabilitation processes. These advanced systems integrate comprehensive patient data, including clinical records, biometric parameters, and lifestyle information, to develop dynamic treatment plans tailored to individual needs. Wearable devices continuously monitor patient progress in real time, providing crucial data to adjust therapies dynamically, ensuring that interventions remain effective and responsive to changes in a patient's condition [5,6]. This real-time feedback loop facilitates precision rehabilitation, minimizing setbacks and promoting steady recovery trajectories.

Telerehabilitation further amplifies these benefits by offering remote monitoring and therapy-adjustment capabilities. Patients can receive professional guidance and care modification without frequent in-person visits, reducing logistical burdens and enhancing access to rehabilitation services. Moreover, digital twins empower patients by offering insights into their recovery progress, encouraging active engagement in health management and fostering adherence to prescribed care plans.

A significant breakthrough in the recovery phase is disease relapse prediction and early warning, a feature that underscores the predictive power of digital twins. By leveraging longitudinal patient data and advanced predictive analytics, digital twin systems can detect subtle patterns indicative of potential disease recurrence. These insights enable the generation of real-time alerts for health care providers, prompting timely interventions that can mitigate complications and prevent hospital readmissions. For instance, changes in respiratory patterns or biomarkers could signal the early stages of a relapse, allowing for preemptive adjustments in treatment strategies.

Digital Twins in Real-World Applications

Currently, the application of digital twins in emergency care remains in its early developmental stages, with significant room for growth, yet they have already demonstrated potential in real-world scenarios. For instance, Pfizer and IBM's collaborative Project BlueSky focuses on developing remote patient monitoring systems that use wearable devices and environmental sensors to track disease symptoms in real time, capture dynamic data, and automatically identify abnormalities for timely intervention [7]. Similarly, Digital Orthopaedics and Dassault Systèmes jointly created a clinical decision support system capable of generating patient-specific computational models to reduce trial-and-error risks and achieve precision treatment [7]. During public health crises like COVID-19, digital twins have proven effective in enhancing emergency response capabilities by simulating real-time data collection and vaccination processes to identify optimal resource allocation strategies and refine contingency planning. These examples underscore the transformative role of digital twin technology in advancing emergency care efficiency and resilience [8].

In the cases mentioned above, digital twins have demonstrated significant advantages over current public health systems. Digital public health enables optimized resource allocation, reduces waste, and enhances emergency response efficiency, and it allows real-time monitoring of patient anomalies to accelerate emergency interventions. Centered on data-driven and intelligent capabilities, these advancements markedly improve the efficiency of resource allocation, response speed, and decision-making precision in emergency care.

In summary, digital twins represent a reliable tool in the era of digital intelligence, becoming increasingly feasible in the current data-rich context. The emergency response system serves as an excellent testing ground for digital twin applications, offering

XSL•FO
RenderX

significant potential across areas such as emergency resource planning, patient monitoring, and rehabilitation. However, the implementation of digital twins in the emergency domain is not a task to be completed overnight. It necessitates advancements in complementary fields, including urban modeling, wearable technology, and the IoT, to fully realize its potential. Despite their promise, the application of digital twins in emergency care faces substantial challenges. Currently, many medical digital twins remain confined to visualization, providing static representations rather than actionable tools for decision-making. For full realization, these systems must evolve beyond visualization to actionable frameworks. A second critical hurdle is data quality and resolution. Biological heterogeneity significantly influences disease progression and treatment outcomes, requiring vast, high-resolution datasets to build robust, accurate models. Lastly, ethical, privacy, and security concerns pose significant barriers to the adoption of digital twins in health care [9]. The integration of sensitive patient data with urban-level modeling raises important questions about data governance, requiring careful consideration to balance innovation with public trust. To address the aforementioned challenges, it is necessary to systematically overcome existing bottlenecks through technological optimization, privacy protection, ethical governance, and multiparty collaboration. On the privacy and security front, full lifecycle data protection must be strengthened by leveraging differential privacy and blockchain technology to anonymize sensitive information and enable traceable storage while implementing dynamic access controls to restrict unnecessary data access. Ethically and legally, clear boundaries and transparent guidelines for technology use must be established. Additionally, collaboration among governments, medical institutions, and technology companies is needed to develop unified technical standards and security evaluation systems alongside public education to build societal trust and ensure patient consent and data authorization. Only through parallel advancements in technology, regulatory frameworks, ethical constraints, and social collaboration can the safe implementation and sustainable development of digital twin technology in emergency scenarios be achieved.

Acknowledgments

This study was funded by the National Natural Science Foundation of China (72404117). The funder had no role in the design or conduct of the study; collection, management, analysis, or interpretation of the data; preparation, review, or approval of the manuscript; or the decision to submit the manuscript for publication.

Authors' Contributions

BZ and HL designed and supervised the research. HL drafted the manuscript. BZ revised the manuscript. JZ and NZ commented on the revision of the manuscript. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Chen J, Yi C, Okegbile SD, Cai J, Shen X. Networking architecture and key supporting technologies for human digital twin in personalized healthcare: a comprehensive survey. IEEE Commun Surv Tut 2024;26(1):706-746. [doi: 10.1109/comst.2023.3308717]
2. Li X, Feng M, Ran Y, Su Y, Liu F, Huang C, et al. Big data in earth system science and progress towards a digital twin. Nat Rev Earth Environ 2023 May 02;4(5):319-332. [doi: 10.1038/s43017-023-00409-w]
3. Fan C, Zhang C, Yahja A, Mostafavi A. Disaster city digital twin: a vision for integrating artificial and human intelligence for disaster management. Int J Inf Manage 2021 Feb;56:102049. [doi: 10.1016/j.ijinfomgt.2019.102049]
4. Laubenbacher R, Mehrad B, Shmulevich I, Trayanova N. Digital twins in medicine. Nat Comput Sci 2024 Mar;4(3):184-191. [doi: 10.1038/s43588-024-00607-6] [Medline: 38532133]
5. Sun T, He X, Li Z. Digital twin in healthcare: recent updates and challenges. Digit Health 2023;9:20552076221149651 [FREE Full text] [doi: 10.1177/20552076221149651] [Medline: 36636729]
6. Scott AK, Oyen ML. Virtual pregnancies: predicting and preventing pregnancy complications with digital twins. Lancet Digit Health 2024 Jul;6(7):e436-e437 [FREE Full text] [doi: 10.1016/S2589-7500(24)00086-4] [Medline: 38906606]
7. Mohapatra S, Bose S. An appraisal of literature for design and implementation of developing a framework for digital twin and validation through case studies. Health Technol 2020 Jul 04;10(5):1229-1237. [doi: 10.1007/s12553-020-00443-4]
8. De Benedictis A, Mazzocca N, Somma A, Strigaro C. Digital twins in healthcare: an architectural proposal and its application in a social distancing case study. IEEE J Biomed Health Inform 2023 Oct;27(10):5143-5154. [doi: 10.1109/JBHI.2022.3205506] [Medline: 36083955]
9. Willcox K, Segundo B. The role of computational science in digital twins. Nat Comput Sci 2024 Mar;4(3):147-149. [doi: 10.1038/s43588-024-00609-4] [Medline: 38532131]

XSL•FO
RenderX

Abbreviations

AI: artificial intelligence
IoT: Internet of Things
MDT: medical device twin

Edited by Y Luo; submitted 26.01.25; peer-reviewed by AK Jameil, J Wang; comments to author 10.03.25; revised version received 22.03.25; accepted 12.04.25; published 21.04.25.

Please cite as:
Li H, Zhang J, Zhang N, Zhu B
Advancing Emergency Care With Digital Twins
JMIR Aging 2025;8:e71777
URL: https://aging.jmir.org/2025/1/e71777
doi:10.2196/71777
PMID:

©Haoran Li, Jingya Zhang, Ning Zhang, Bin Zhu. Originally published in JMIR Aging (https://aging.jmir.org), 21.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluation of an AI-Based Clinical Decision Support System for Perioperative Care of Older Patients: Ethical Analysis of Focus Groups With Older Adults

Nina Parchmann[1], MA; Marcin Orzechowski[1], PD, Dr; Simone Brefka[2], Dr; Florian Steger[1], Prof Dr

[1]Institute of the History, Philosophy and Ethics of Medicine, Ulm University, Oberberghof 7, Ulm, Germany

[2]Institute for Geriatric Research at AGAPLESION Bethesda Ulm, Ulm University Medical Center, Ulm, Germany

Corresponding Author:
Nina Parchmann, MA
Institute of the History, Philosophy and Ethics of Medicine, Ulm University, Oberberghof 7, Ulm, Germany

Abstract

Background: The development and introduction of an artificial intelligence (AI)–based clinical decision support system (CDSS) in surgical departments as part of the "Supporting Surgery with Geriatric Co-management and AI" project addresses the challenges of an increasingly aging population. The system enables digital comanagement of older patients by providing evidence-based evaluations of their health status, along with corresponding medical recommendations, with the aim of improving their perioperative care.

Objective: The use of an AI-based CDSS in patient care raises ethical challenges. Gathering the opinions, expectations, and concerns of older adults (as potential patients) regarding the CDSS enables the identification of ethical opportunities, concerns, and limitations associated with implementing such a system in hospitals.

Methods: We conducted 5 focus groups with participants aged 65 years or older. The transcripts were evaluated using qualitative content analysis and ethically analyzed. Categories were inductively generated, followed by a thematic classification of participants' statements. We found that technical understanding did not influence the older adults' opinions.

Results: Ethical opportunities and concerns were identified. On the one hand, diagnosis and treatment could be accelerated, the patient-AI-physician interaction could enhance medical treatment, and the coordination of hospital processes could be improved. On the other hand, the quality of the CDSS depends on an adequate data foundation and robust cybersecurity. Potential risks included habituation effects, loss of a second medical opinion, and illness severity influencing patients' attitude toward medical recommendations. The risk of overdiagnosis and overtreatment was discussed controversially, and treatment options could be influenced by interests and finances. Additional concerns included challenges with time savings, potential declines in medical skills, and effects on the length of hospital stay.

Conclusions: To address the ethical challenges, we recommend allocating sufficient time for use of the CDSS and emphasizing individualized review of the CDSS results. Furthermore, we suggest limiting private financial sponsorship.

(JMIR Aging 2025;8:e71568) doi:10.2196/71568

KEYWORDS

medical ethics; older patient care; artificial intelligence; AI; clinical decision support system; focus groups

Introduction

Background

The global population is aging. A combination of increasing life expectancy [1] and enduring demographic shifts [2] has led to an increase in people aged 65 years or older [3]. Older patients face an increased risk of losing physical function and independence during hospitalization [4]. Comanagement of older patients, defined as a process of joint decision-making between surgical staff and a geriatrician, can improve outcomes such as mortality, length of hospital stay, and perioperative complications [5,6]. The rising prevalence of age-related conditions, including cognitive impairment and multimorbidity [7,8], further increases the demand for individualized, patient-centered treatment and resource-intensive medical care at a time when resources are already limited [2]. In parallel, the medical application of artificial intelligence (AI) is expanding, with the potential to enhance personalized medicine, patient care, prediction, and diagnostics while improving accuracy and efficiency and enabling earlier intervention [9].

A systematic review found the use of clinical decision support systems (CDSSs) to be effective with regard to both process- and patient-related outcomes [10].

XSL•FO
RenderX

The research project "Supporting Surgery with Geriatric Comanagement and AI" addresses the growing challenges in the health care system by developing and introducing an AI-based CDSS for perioperative care of older patients and subsequent care planning [11]. This system automatically processes and analyzes patient data to provide an evidence-based evaluation of the individual's health condition and corresponding medical recommendations. The CDSS aims to improve patient care and support medical professionals in surgical wards. However, the introduction of the CDSS also raises ethical challenges [12].

The benefits of medical AI systems are offset by concerns regarding accountability and transparency, biases in training data that may compromise safety, inequities in access, and data security risks [9]. Considering that older adults represent a patient group considered vulnerable and requiring particular protection, addressing AI-related ethical challenges to safeguard patient integrity is of major importance.

The application of large language models in hospitals as a means of improving patient care has been scientifically analyzed across various medical fields and purposes. However, to the best of our knowledge, an ethical evaluation of their use within a CDSS, particularly in perioperative care of older patients and incorporating the perspectives of potential patients, remains lacking.

Objectives

To explore the opinions, expectations, and concerns of potential older patients regarding the development and introduction of an AI-based CDSS in surgical wards caring for older patients, we conducted focus groups with older adults. We sought to (1) examine how the level of technical understanding influences their views and (2) identify the ethical opportunities, concerns, and limitations associated with implementing such a system in hospitals.

Methods

Ethical Considerations

Ethical approval was waived by the ethics committee of Ulm University, because no information on personal data, health-related data, or data on sexuality of the interviewees was collected (May 16, 2024). To ensure informed consent, all participants were sent a digital copy of the consent form before the discussion and were verbally briefed on the focus group procedures. The information covered data handling, privacy protection, voluntary participation, and the right to opt out, as well as details about the project and its funding. All participants provided written informed consent to participate, and no financial incentives were offered.

Overview

Conducting focus groups enabled us to explore the opinions, expectations, and concerns of older adults. This method allows participants to freely express their views, stimulate ideas through interaction, and provide a wide range of perspectives within each group [13,14].

Participant selection followed a deductive sampling approach. The inclusion criteria were age 65 years or older and sufficient German language skills to participate in a group discussion. Given that focus groups rely on oral conversation, we ensured that all participants met the language requirements through 3 steps. First, participants had to understand the written invitation and respond appropriately. Second, we conducted telephone conversations with participants to answer questions and schedule participation before each focus group. Third and last, before starting each focus group discussion, we verified German speaking and comprehension skills through brief conversational exchanges with all participants.

With the support of a gatekeeper, invitation letters were distributed through mailing lists targeting older adults interested in education, posted on an academic website, and further disseminated via snowball sampling. Participants were assigned to focus groups using a stratified randomization approach. First, interested individuals were divided into groups by gender. Next, an equal number of participants from each group were randomly selected and assigned to 3 different focus groups with unevaluated, mixed levels of technical understanding.

Individuals who had not yet been assigned to a group were asked to complete questionnaires assessing technical knowledge. The questionnaires were scored using a points system, and on the basis of these scores, participants were divided into 2 focus groups: one with higher technical understanding and one with lower technical understanding.

In total, 5 focus groups were conducted with 30 participants from the city of Ulm in southern Germany and surrounding areas in June and July 2024. The sessions were held face-to-face. The conversation, including technical explanations of all AI-related aspects, was held in simple language and at a slower speaking rate to accommodate age-related limitations and varying levels of technical knowledge among the participants.

The discussion guide was semistructured to allow flexible adaptation to the dynamics of each group and was organized into 5 thematic categories: general opportunities, risks and resources, accuracy and trust, data processing, patient-physician relationship, and limitations.

The discussions were recorded, transcribed, and anonymized. The quotations included in this study have been translated from German into English.

The transcripts were first analyzed using qualitative content analysis [15], followed by thematic analysis [16], while taking into account 6 analytic factors specific to focus groups [13]. Categories and subcategories were identified inductively, followed by a deductive analysis structured according to the questionnaire. Participant statements were coded and assigned to the established categories. First, 3 focus groups with mixed levels of technical understanding and the 2 groups with lower and higher technical understanding were analyzed separately, considering the 6 analytic factors [13], and then compared. As no major differences were detected between the groups with lower and higher technical understanding, all focus groups were subsequently combined for comparison and analysis.

XSL•FO
RenderX

Ethical Principles Considered in the Analysis

After this analysis, the categories and findings were reviewed by researchers with technical expertise in medicine, medical ethics, and communication and were ethically analyzed using the principle-based approach proposed by Beauchamp and Childress [17]. This biomedical approach is widely applied as a key method for reasoning about and analyzing ethical challenges in medicine. The ethical principles considered in the analysis reflect core medical values and comprise autonomy, beneficence, nonmaleficence, and justice. The principle of autonomy refers to decision-making based on the patient's individual preferences. Beneficence emphasizes maximizing patient benefit while balancing risks against potential benefits, and nonmaleficence aims to prevent harm to the patient. The principle of justice, also referred to as social justice, addresses the equitable and appropriate distribution of resources [17].

Results

Throughout the focus groups, older adults expressed their opinions, expectations, and concerns on an AI-based CDSS, which were organized into five categories: (1) data use, (2) patient-AI-physician interaction, (3) patient care, (4) resources, and (5) framework.

Data Use

Participants emphasized that the quality of the CDSS, including its technical functionality, depends on an adequate data foundation, while recognizing the patient as an uncertain data source. They pointed out the risk of blindly trusting memory-based information provided by older patients who might forget, cheat, or unintentionally provide inaccurate information in a hospital environment perceived as unusual or stressful.

Most discussants highlighted the importance of protecting health-related patient data and were critical of current data management practices. While participants stressed that data protection measures would be required, they anticipated that security gaps would likely persist:

> *[The data security] is full of holes [...] sometimes. […] Much is impeded but the reality is sometimes bitterly different.* [Participant]

Data manipulation and hacking attacks were classified as high risk. The discussants emphasized concerns about targeted data alteration and about physicians unknowingly referring to the falsified data if such manipulation remained undetected.

In addition, concerns were raised about data access by third parties. Most participants highlighted the risk of data misuse, while a few pointed out that the level of concern would depend on the individual's age and extent of misuse.

Patient-AI-Physician Interaction

The discussants noted that older patients' interactions with, and perspectives of, AI and the physician were influenced by individual personality traits. Three traits emerged from the discussions: (1) passive and trusting, (2) proactive, and (3) skeptical and experience-led (Table 1) .

Table . Representation of personality traits as influencing factors in the interactions and perspectives of older adults.

	Passive and trusting	Proactive	Skeptical and experience-led
Trust in physicians using AI[a]-based system	Complete trust in physician and health care system. Trust promotes acceptance of the use of any medical device.	Precondition: explanation of the system and involvement of the patient.	None: based on AI error rates and prior experiences of misdiagnosis.
Decision about using AI-based system	No involvement	Full involvement: right to demand use of AI-based system	Involvement: right to reject use of AI-based system
Participation (factor: cognitive capability)	None	Active participation of patient: entering their own data into the system	Gaining insights into visualizations as a basis for a discussion with the physician
Explainability of AI-based suggestions	Unnecessary	Highly important	Limitation: AI is not explainable (explainability is questionable when AI explains its own system)
Patient rights	Right not to know	Access right to the CDSS[b] to add missing data	Reviewing AI-based results

[a]AI: artificial intelligence.

[b]CDSS: clinical decision support system.

The relationship between the physician and the AI was expected to evolve into mutual supervision: the AI would oversee the physician's knowledge, while the physician would monitor the AI's accuracy.

The severity of illness was identified as an impact factor, with increasing severity making patients more receptive to treatment suggestions. In particular, in cases of severe illness, patients tended to place greater trust in potentially effective treatment options and to question the physician if their treatment suggestion seemed less effective:

> *If I suffer from a severe illness and AI says, "If this and that, it will turn out well." And my physician says "Um." Whom do I trust in this case?* [Participant]

According to participants, a habituation effect of the CDSS would develop. In the long term, patients would either trust the AI more than the physician or not notice its use. At the same

XSL•FO
RenderX

time, regular use of the CDSS would lead physicians to accept AI-based results as a given, place blind trust in the system, and thereby risk misdiagnosis.

While several discussants described the AI as providing a second opinion alongside the physician's assessment, others characterized the AI and physician together as forming a single first medical opinion. In this context, participants noted the risk that the option of obtaining a true second opinion could be lost if different physicians using the same CDSS arrived at similar conclusions.

Patient Care

Diagnosis and therapy provision were expected to accelerate, and the identification of rare and complex diseases was expected to become easier.

A few participants believed that the CDSS could help prevent overdiagnosis. However, others expressed concerns about the risk of overdiagnosis and overtreatment, which could lead to the prescription of more medication and prophylaxis programs, thereby increasing patients' sense of illness.

The discussants also emphasized the risk of treatment suggestions being influenced by calculations of cost of the lifetime gained, the patient's financial status and age, or third-party interests. The participants called for shared decision-making that accounts for the patient's preferences, individual circumstances, and stage of life:

> *The limitation is that the final decision is not made by the AI but by the human.* [Participant]

Resources

According to the discussants, language barriers could be reduced by providing medical professionals with guidance on patient-friendly communication and by offering staff from abroad information either in a widely spoken international language or in their native language.

Some discussants anticipated time savings through the reduction of complexity. However, others noted that such savings might depend on blind trust in the AI and reduced patient-physician communication.

Numerous participants also expressed concern about the potential loss of clinical intuition as a hallmark of good physicians and the inability to work without technical support:

> *I would have the concern that they push AI with the reason "I save employees, we save salaries, because we can just combine less qualified persons with the technology." I think that only works well with good physicians.* [Participant]

Framework

The hospital as an institution was expected to benefit from the CDSS through reduced patient length of stay and improved coordination of processes. However, discussants expressed concerns about a potential increase in patient numbers and a reduction in the time allocated for individual treatment. Furthermore, stress, lack of time, and convenience could lead physicians to blindly trust the CDSS:

> *The downside would perhaps be a transit operation. Each patient has five minutes, has only certain things to tell, to want and you press onto the computer and that is it.* [Participant]

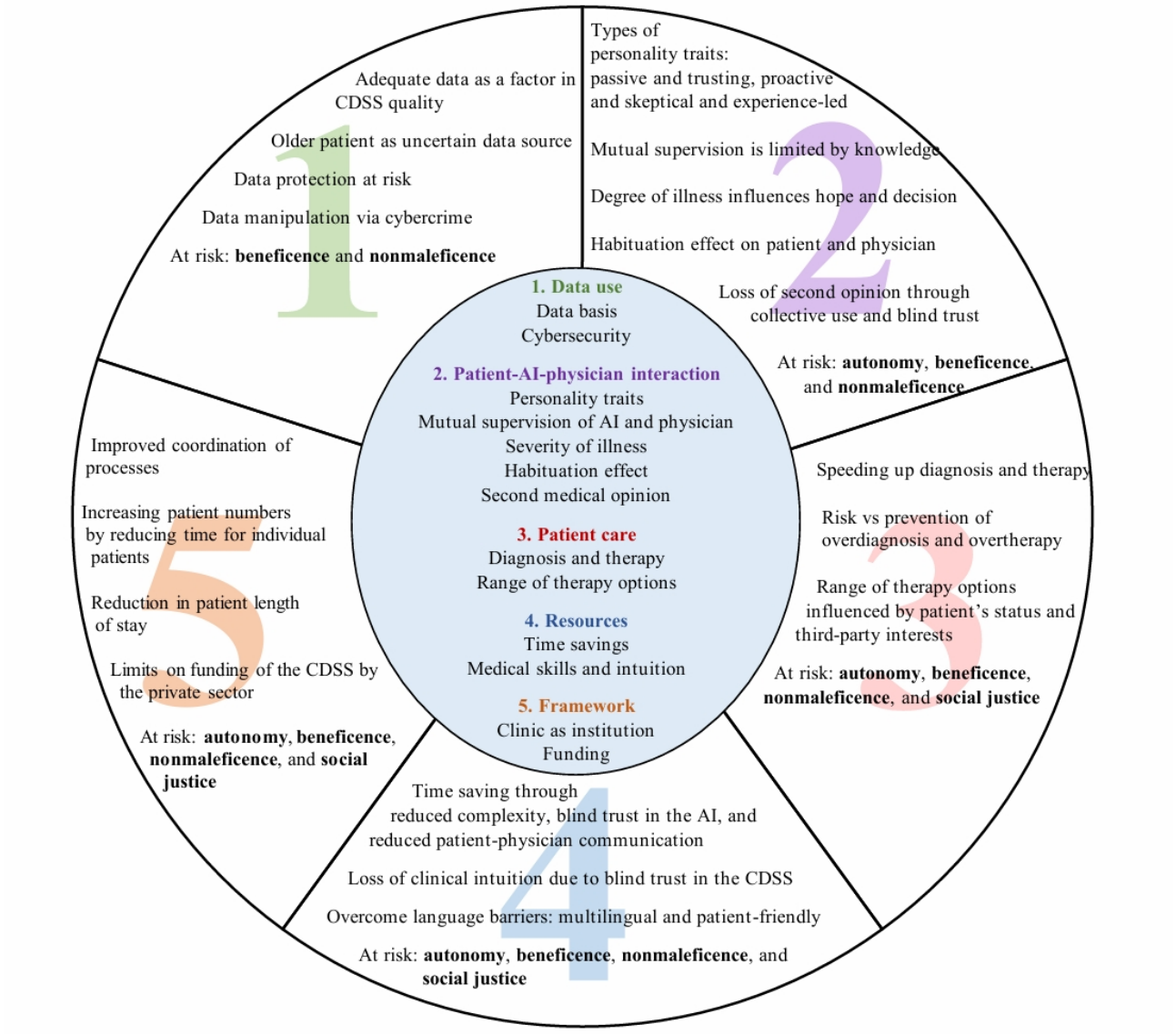
Most of the participants insisted that the AI-based system should not be financed by the private sector, as a precaution during its implementation. This would help prevent profit-oriented influences on the CDSS in the sensitive field of older patient care.

Discussion

Ethical opportunities, concerns, and limitations of an AI-based CDSS, as identified by older adults, are discussed with reference to the 4 principles proposed by Beauchamp and Childress [17]: autonomy, beneficence, nonmaleficence, and social justice (Figure 1).

XSL•FO
RenderX

Figure 1. Representation of ethical challenges raised by older adults regarding the introduction of an artificial intelligence (AI)–based clinical decision support system (CDSS) in older patient care.



Data Use

Participants described CDSS quality as dependent on an adequate data foundation and viewed the older patient as an uncertain data source. The completeness and accuracy of data were identified as the most relevant factors influencing system quality [18]. Low data quality, including insufficient or missing medical information, carries the risk of biased models being developed, which may lead to adverse results from the AI-based system [19]. However, self-reported data from older patients are considered less reliable [20], as they tend to forget or inaccurately recall information in hospital settings [20]. This uncertainty in the underlying data reduces system performance and quality, increasing the risk of incorrect medical suggestions and treatment. Consequently, patient harm from inappropriate treatment or diminished quality of medical care would jeopardize the principle of nonmaleficence.

Discussants expressed concern about cybersecurity. Cyberattacks are an increasing risk in the health care system [21]. Data that are altered, stolen, or manipulated [21] may go undetected [22] and could compromise CDSS recommendations. Medical decision-making based on false data puts patients at risk of receiving incorrect or inadequate treatment, thereby jeopardizing the principle of beneficence. In addition, the misuse of patients' personal data, violations of privacy, and the implementation of incorrect treatment resulting from altered data further threaten the principle of nonmaleficence.

Patient-AI-Physician Interaction

The personality traits of (1) passive and trusting, (2) proactive, and (3) skeptical and experience-led were identified as factors influencing older adults' perspectives. Personality traits shape individual preferences [23], affect older patients' perceptions of the CDSS [24], and influence expectations of the device.

The relationship between physician and AI was expected to develop into mutual supervision. Complementing each other's knowledge and detecting incorrect evaluations [25] through mutual supervision could prevent incorrect treatment and simultaneously improve medical care. This relationship could arise from physicians supervising the AI by interpreting its

XSL•FO
RenderX

results, while the AI minimizes the human errors of the physician [26]. The benefits that older patients experience from improvements in their medical care would support the principle of beneficence, whereas preventing patients from receiving incorrect medical treatment by minimizing human errors would strengthen the principle of nonmaleficence. However, mutual supervision remains imbalanced, given that the detection of AI-based errors is limited by the physician's knowledge [27] and the lack of explainability of black-box systems [28]. In such cases, rather than mutual supervision, a paternalistic relationship could develop between the AI and the physician [28].

Furthermore, illness severity was described as an impact factor. Life-threatening diseases tend to increase patients' hope for potential cures and create uncertainty when trusting a physician's less effective treatment options. Transparent communication of AI-based results is therefore necessary to enable patients to make autonomous, informed decisions and to demonstrate the range of available treatment options. However, physicians do not necessarily want to communicate AI-based results transparently to their patients; they would rather withhold information [29]. Withholding medical information hinders shared decision-making [29] and undermines a trustful patient-physician relationship. Uncertainty in trust may cause patients to experience worry and avoid decision-making [30]; it may also reduce their ability to make rational treatment choices. Consequently, the consideration of patients' individual preferences in medical decision-making, influenced by the severity of their illness, could be restricted, compromising the principle of autonomy.

Participants anticipated the development of a habituation effect. After becoming familiar with the application, patients might come to trust the AI more than the physician. Such distrust of the physician could lead patients to become skeptical of medical recommendations [31], thereby reinforcing uncertainty. Medical decision-making based on uncertainty and a general mistrust of physicians' recommendations undermines patients' independent decision-making, thereby jeopardizing the principle of autonomy. In addition, a general mistrust of medical recommendations could lead patients to decline beneficial treatment options, thus compromising the principle of beneficence. Physicians using the device regularly were also assumed to develop blind trust in the CDSS. Previous research has shown that habituation can lead users to take advantage of systems that permit unquestioned acceptance of suggestions, despite the high risks involved [32]. Blindly following CDSS suggestions that may not be accurate carries the risks of misdiagnosis and incorrect treatment, thereby jeopardizing the principle of nonmaleficence.

The discussants highlighted the risk of losing access to a genuine second medical opinion when different physicians using the same system reach similar conclusions. In this context, the combination of device and physician was described as a single first medical opinion. While a CDSS is designed to support physicians by providing a second medical evaluation, patients retain the option of consulting another physician for a second medical opinion [33]. However, if physicians collectively place blind trust in the CDSS and if similar systems generating similar results are used across hospitals, patients risk being deprived of a second medical opinion. Such circumstances would undermine patients' independent consideration of medical options and preference-based decision-making, thereby conflicting with the principle of autonomy.

Patient Care

The discussants emphasized that the CDSS would speed up diagnosis, particularly in rare and complex diseases, as well as facilitate the provision of treatment. The medical use of AI enables rapid analysis of patient data, leading to a reduction in the time required for diagnosis [34]. Early diagnosis in older patients provides a basis for timely initiation of treatment, lowering the risk of impairment and secondary complications [35]. As a result, overall medical care would improve. Prompt diagnosis may also enhance the healing process of older patients, thereby supporting the principle of beneficence. Furthermore, the prevention and decreased risk of medical complications through timely treatment provision aligns with the principle of nonmaleficence.

Several participants expressed concern that the CDSS could contribute to overdiagnosis and overtreatment, causing patients to feel increasingly ill, while others suggested that it could prevent these problems. Previous studies with AI-based screening devices have shown that AI can accurately detect cell changes that are not necessarily pathological but nevertheless result in overtreatment [36]. Overdiagnosis exposes patients to unnecessary medical treatments and the associated risk of harmful side effects [37], thereby conflicting with the principle of nonmaleficence. Overdiagnosis can also heighten patients' perception of illness, and when combined with treatment that offers no medical benefit, it further undermines the principle of beneficence. In addition, the preventable use of already limited resources fosters an imbalanced distribution of medical resources, conflicting with the principle of social justice. Conversely, if the CDSS helps prevent overdiagnosis and overtreatment, it could improve the provision of essential treatment, supporting the principle of beneficence. Preventing insufficient treatment in older patients and the resulting consequences would also reinforce the principle of nonmaleficence.

Participants expressed major concerns about treatment options being influenced by financial considerations and third-party interests. Security gaps and unauthorized access to data or algorithms could enable alterations to the AI-based system [38], thereby influencing CDSS outcomes. Restricting treatment options to serve profit-oriented interests of third parties could result in less effective treatment options for older patients. Such restrictions reduce older patients' capacity for informed decision-making and jeopardize the principle of autonomy. Limiting treatment options on the basis of third-party or financial interests carries the risk of excluding effective treatments, conflicting with the principle of beneficence. Moreover, narrowing treatment options in ways that compromise patient outcomes would undermine the principle of nonmaleficence.

Resources

Participants expected time savings through reduced complexity, blind trust in the CDSS, and avoidance of extensive

XSL•FO
RenderX

patient-physician communication. However, given the limited accuracy of a CDSS [39], physicians must review the AI-generated results individually to reduce the risk of incorrect treatment. Understanding the individual preferences and concerns of older patients requires time-intensive conversation [40] and is fundamental for decision-making. A lack of communication regarding the patient's preferences could lead to medical decision-making that does not sufficiently take these preferences into account, thereby jeopardizing the principle of autonomy. Relying blindly on the CDSS without examining the results for false implications could harm patient outcomes and compromise the principle of nonmaleficence. Together, blind trust in the CDSS and reduced patient-physician communication could undermine patient-centered medical care and treatment and, in turn, jeopardize the principle of beneficence.

The discussants expected the CDSS to help medical staff overcome language barriers. Effective communication between patients and medical professionals is necessary to identify individual concerns and to build trust [41]. Language barriers may cause patients to avoid follow-up appointments [42], leading to undetected postoperative complications. Poor medical care and worse outcomes resulting from communication difficulties [43] could be prevented by reducing language-related barriers. A multilingual, patient-centered use of the AI-based CDSS could therefore improve patient care and support the principle of beneficence. Improved communication about side effects, treatment risks, and patient concerns could support the principle of nonmaleficence. In addition, language support for medical staff from abroad can save time and human resources while laying the foundation for equitable medical treatment regardless of the language skills of patients or staff, thereby strengthening the principle of social justice.

Medical professionals who regularly use the CDSS were expected to lose clinical intuition and develop a dependency on technical support. Some researchers suggest that the CDSS could enhance medical skills by allowing physicians to learn from its results [44]. However, blind trust in the system may instead reduce medical skills and foster dependency [44]. As a result, the accuracy of medical treatment and thus the protection of patient integrity could become dependent on technology, jeopardizing the principle of beneficence. Moreover, the ability of physicians to verify suggested results and prevent incorrect treatment might no longer be ensured, thereby compromising the principle of nonmaleficence.

Framework

According to participants, hospitals would benefit from the CDSS through better coordination of processes, shorter patient stays, and reduced time spent on individual patient care, thereby enabling treatment of more patients. The shortest possible perioperative stay is financially the most profitable [45]. Minimizing length of stay while increasing patient volume enhances hospital revenue and saves resources. In addition, a higher number of patients could receive treatment, supporting the principle of social justice. However, a balance must be maintained between the needs of individual patients and those of the broader population. Older patients with multiple comorbidities, in particular, face a higher risk of mortality after premature discharge [46]. Profit-driven discharge of patients before medical necessity increases the risk of delayed complication detection and increased mortality in older patients, thereby jeopardizing the principle of nonmaleficence. Delayed treatment of postoperative complications could further expose patients to more severe consequences and subsequently compromise the principle of beneficence.

The discussants were emphatic about the need to restrict private-sector funding of the CDSS to prevent profit-oriented influence. As the algorithms and underlying data of a CDSS are adjustable [38], treatment options could be altered to favor sponsors. This would restrict the consideration of patients' individual preferences in medical decision-making and thereby jeopardize the principle of autonomy. Limiting treatment options for financial gain could also result in withholding effective treatments, undermining the principle of beneficence. Moreover, biased treatment suggestions could expose patients to incorrect medical care, compromising the principle of nonmaleficence.

Conclusions

This research indicates that older adults anticipate both opportunities and ethical challenges in data use, the patient-AI-physician interaction, patient care, resources, and framework when an AI-based CDSS is introduced in older patient care.

To address the ethical concerns raised by the limited number of potential patients in this qualitative research, we recommend allowing sufficient time for use of the CDSS to account for individual patient concerns and preferences. Furthermore, individualized review of AI-generated suggestions and avoidance of private sponsorship are recommended to reduce the risk of less effective medical treatments.

This research is limited by the fact that the participants lived in southern Germany and were privileged in having access to the internet and an academic network. Given the qualitative design, their subjective opinions do not represent average experiences and cannot be generalized to the global older population.

To further analyze the concerns of older adults, consideration of gender and differentiation by age ranges (eg, 65-74, 75-84, and ≥85 years) is recommended. On the basis of our small sample of participants from a single region in Germany, we recommend continuing the research with larger samples involving transregional and transnational participants to examine the scope and generalizability of our findings.

Acknowledgments

This work was supported by the German Federal Ministry of Research, Technology and Space

XSL•FO
RenderX

within the funding program “Strengthening Research in Geriatrics and Gerontology” (01GY2101). The “Supporting Surgery with Geriatric Comanagement and AI” study group comprises Karen Andersen-Ranberg, Sophia Andres, Sibylle Beck, Christian Bolenz, Simone Brefka, Raffael Cintean, Karen Clauss, Dhayana Dallmeier, Genia Decker, Miriam Deniz, Michael Denkinger, Vanessa Disque, Anna Lena Flagmeier, Marina Fotteler, Florian Gebhard, Julia Geiger, Janina Hahn, David Hansen, Thomas Hoffmann, Wolfgang Janni, Hans Kestler, Reinhold Kilian, Jochen Klenk, Thomas Derya Kocar, Aleyna Koese, Johann Kraus, Christoph Leinert, Elena Leinert, Fabia Mangold, Judith Mantz, Annabel Müller-Stierlin, Gabi Müller, Nadir Nasir, Graziano Onder, Desmond O’Neill, Marcin Orzechowski, Carlos Pankratz, Nina Parchmann, Nuh Rahbari, Johannes Schobel, Fabienne Schochter, Konrad Schütze, Tobias Skuban-Eiseler, Vytautas Stasiunaitis, Florian Steger, Walter Swoboda, Adriane Uihlein, Sebastian Voigt-Radloff, Martin Wehling, Pascal Wendel, Felix Wezel, Jana Willems, Jessica Wolf, Philip Wolf, Friedemann Zengerling, and Marco Zeptner. The authors thank Dr Markus Marquard and the Center for General Scientific Continuing Education of Ulm University for support in contacting and communicating with focus group participants.

Authors' Contributions

FS was responsible for funding acquisition. NP, MO, and FS were responsible for conceptualization and planning, conducting the focus groups, and investigation. NP wrote the original draft. NP, MO, SB, and FS reviewed and edited the manuscript. All authors have contributed to this paper and have approved the final version.

Conflicts of Interest

None declared.

References

1. Life expectancy at birth, total (years). World Bank Group. URL: https://data.worldbank.org/indicator/SP.DYN.LE00.IN [accessed 2024-09-14]
2. How university medical centers can best position themselves for changes across society. Charité – Universitätsmedizin Berlin. URL: https://www.charite.de/en/service/press_reports/artikel/detail/rethinking_european_healthcare_systems/ [accessed 2024-09-15]
3. Share of population that are aged 65 years and older in european countries in 2023. Statista. URL: https://www.statista.com/statistics/1105835/share-of-elderly-population-in-europe-by-country/ [accessed 2024-09-14]
4. Loyd C, Markland AD, Zhang Y, et al. Prevalence of hospital-associated disability in older adults: a meta-analysis. J Am Med Dir Assoc 2020 Apr;21(4):455-461. [doi: 10.1016/j.jamda.2019.09.015] [Medline: 31734122]
5. Van Grootven B, Flamaing J, Dierckx de Casterlé B, et al. Effectiveness of in-hospital geriatric co-management: a systematic review and meta-analysis. Age Ageing 2017 Nov 1;46(6):903-910. [doi: 10.1093/ageing/afx051] [Medline: 28444116]
6. Rapp K, Becker C, Todd C, et al. The association between orthogeriatric co-management and mortality following hip fracture. Dtsch Arztebl Int 2020 Jan 24;117(4):53-59. [doi: 10.3238/arztebl.2020.0053] [Medline: 32036854]
7. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. The Lancet 2012 Jul;380(9836):37-43. [doi: 10.1016/S0140-6736(12)60240-2]
8. Aubert CE, Streit S, Da Costa BR, et al. Polypharmacy and specific comorbidities in university primary care settings. Eur J Intern Med 2016 Nov;35:35-42. [doi: 10.1016/j.ejim.2016.05.022] [Medline: 27289492]
9. Aravazhi PS, Gunasekaran P, Benjamin NZ, et al. The integration of artificial intelligence into clinical medicine: trends, challenges, and future directions. Dis Mon 2025 Jun;71(6):101882. [doi: 10.1016/j.disamonth.2025.101882] [Medline: 40140300]
10. Damoiseaux-Volman BA, van der Velde N, Ruige SG, Romijn JA, Abu-Hanna A, Medlock S. Effect of interventions with a clinical decision support system for hospitalized older patients: systematic review mapping implementation and design factors. JMIR Med Inform 2021 Jul 16;9(7):e28023. [doi: 10.2196/28023] [Medline: 34269682]
11. Leinert C, Fotteler M, Kocar TD, et al. Supporting SURgery with GEriatric co-management and AI (SURGE-Ahead): a study protocol for the development of a digital geriatrician. PLoS ONE 2023;18(6):e0287230. [doi: 10.1371/journal.pone.0287230] [Medline: 37327245]
12. Skuban-Eiseler T, Orzechowski M, Denkinger M, Kocar TD, Leinert C, Steger F. Artificial intelligence-based clinical decision support systems in geriatrics: an ethical analysis. J Am Med Dir Assoc 2023 Sep;24(9):1271-1276. [doi: 10.1016/j.jamda.2023.06.008] [Medline: 37453451]
13. Krueger RA, Casey MA. Focus Groups- A Practical Guide for Applied Research, 5th edition: Sage Publications; 2015.
14. Hennink MM. Focus Group Discussions- Understanding Qualitative Research: Oxford University Press; 2014.
15. Mayring P. Qualitative Inhaltsanalyse- Grundlagen Und Techniken [Book in German]: Beltz; 2015.
16. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
17. Beauchamp T, Childress J. Principles of Biomedical Ethics, 8th edition: Oxford University Press; 2019.

XSL•FO
RenderX

18. Budach L, Feuerpfeil M, Ihde N, et al. The effects of data quality on machine learning performance. arXiv. Preprint posted online on Jul 29, 2022. [doi: 10.48550/arXiv.2207.14529]
19. Carini C, Seyhan AA. Tribulations and future opportunities for artificial intelligence in precision medicine. J Transl Med 2024 Apr 30;22(1):411. [doi: 10.1186/s12967-024-05067-0] [Medline: 38702711]
20. Savioli G, Ceresa IF, Giordano M, et al. The reliability of anamnestic data in the management of clostridium tetani infection in elderly. Front Med (Lausanne) 2021;8:684594. [doi: 10.3389/fmed.2021.684594] [Medline: 34778280]
21. Wasserman L, Wasserman Y. Hospital cybersecurity risks and gaps: review (for the non-cyber professional). Front Digit Health 2022;4:862221. [doi: 10.3389/fdgth.2022.862221] [Medline: 36033634]
22. Mirsky Y, Mahler T, Shelef I, et al. CT-GAN: malicious tampering of 3D medical imagery using deep learning. arXiv. Preprint posted online on Jan 11, 2019. [doi: 10.48550/arXiv.1901.03597]
23. Flynn KE, Smith MA. Personality and health care decision-making style. J Gerontol B Psychol Sci Soc Sci 2007 Sep;62(5):261-267. [doi: 10.1093/geronb/62.5.p261] [Medline: 17906167]
24. Redelmeier DA, Najeeb U, Etchells EE. Understanding patient personality in medical care: five-factor model. J Gen Intern Med 2021 Jul;36(7):2111-2114. [doi: 10.1007/s11606-021-06598-8] [Medline: 33506393]
25. Kempt H, Heilinger JC, Nagel SK. "I'm afraid I can't let you do that, Doctor": meaningful disagreements with AI in medical contexts. AI & Soc 2023 Aug;38(4):1407-1414. [doi: 10.1007/s00146-022-01418-x]
26. Göndöcs D, Dörfler V. AI in medical diagnosis: AI prediction & human judgment. Artif Intell Med 2024 Mar;149:102769. [doi: 10.1016/j.artmed.2024.102769] [Medline: 38462271]
27. Richard A, Mayag B, Talbot F, et al. What does it mean to provide decision support to a responsible and competent expert?: The case of diagnostic decision support systems. EURO J Decis Process 2020. [doi: 10.1007/s40070-020-00116-7]
28. Xu H, Shuttleworth KMJ. Medical artificial intelligence and the black box problem: a view based on the ethical principle of "do no harm". Intelligent Medicine 2024 Feb;4(1):52-57. [doi: 10.1016/j.imed.2023.08.001]
29. Lorenzini G, Arbelaez Ossa L, Shaw DM, Elger BS. Artificial intelligence and the doctor-patient relationship expanding the paradigm of shared decision making. Bioethics 2023 Jun;37(5):424-429. [doi: 10.1111/bioe.13158] [Medline: 36964989]
30. Hillen MA, Gutheil CM, Strout TD, Smets EMA, Han PKJ. Tolerance of uncertainty: conceptual analysis, integrative model, and implications for healthcare. Soc Sci Med 2017 May;180:62-75. [doi: 10.1016/j.socscimed.2017.03.024] [Medline: 28324792]
31. Williamson LD, Thompson KM, Ledford CJW. Trust takes two…. J Am Board Fam Med 2022 Dec 23;35(6):1179-1182. [doi: 10.3122/jabfm.2022.220126R1] [Medline: 36396414]
32. Hoesterey S, Onnasch L. The effect of risk on trust attitude and trust behavior in interaction with information and decision automation. Cogn Tech Work 2023 Feb;25(1):15-29. [doi: 10.1007/s10111-022-00718-y]
33. Isaacs D. The second opinion. J Paediatrics Child Health 2023 Mar;59(3):422-423. [doi: 10.1111/jpc.16246]
34. S Alshuhri M, Al-Musawi SG, Al-Alwany AA, et al. Artificial intelligence in cancer diagnosis: opportunities and challenges. Pathol Res Pract 2024 Jan;253:154996. [doi: 10.1016/j.prp.2023.154996] [Medline: 38118214]
35. Heppner HJ, Haitham H. Intensive care of geriatric patients-a thin line between under- and overtreatment. Wien Med Wochenschr 2022 Apr;172(5-6):102-108. [doi: 10.1007/s10354-021-00902-1] [Medline: 35006520]
36. Mori Y, Misawa M, Bernal J, et al. Artificial intelligence for disease diagnosis: the criterion standard challenge. Gastrointest Endosc 2022 Aug;96(2):370-372. [doi: 10.1016/j.gie.2022.04.057] [Medline: 35489396]
37. Gupta P, Gupta M, Koul N. Overdiagnosis and overtreatment; how to deal with too much medicine. J Family Med Prim Care 2020;9(8):3815. [doi: 10.4103/jfmpc.jfmpc_433_20]
38. Brown S, Davidovic J, Hasan A. The algorithm audit: scoring the algorithms that score us. Big Data Soc 2021 Jan;8(1). [doi: 10.1177/2053951720983865]
39. Gong Y, Min H, Jing X, Yu P. Challenges and opportunities of artificial intelligence in CDSS and patient safety. Stud Health Technol Inform 2024 Aug 22;316:1250-1254. [doi: 10.3233/SHTI240638] [Medline: 39176608]
40. Rostoft S, van den Bos F, Pedersen R, Hamaker ME. Shared decision-making in older patients with cancer - what does the patient want? J Geriatr Oncol 2021 Apr;12(3):339-342. [doi: 10.1016/j.jgo.2020.08.001] [Medline: 32839118]
41. Kwame A, Petrucka PM. A literature-based study of patient-centered care and communication in nurse-patient interactions: barriers, facilitators, and the way forward. BMC Nurs 2021 Sep 3;20(1):158. [doi: 10.1186/s12912-021-00684-2] [Medline: 34479560]
42. Chandrashekar P, Zhang R, Leung M, Jain SH. Impact of patient-physician language concordance on healthcare utilization. J Gen Intern Med 2022 Jun;37(8):2120-2122. [doi: 10.1007/s11606-021-06998-w] [Medline: 34342810]
43. Luan-Erfe BM, Erfe JM, DeCaria B, Okocha O. Limited English proficiency and perioperative patient-centered outcomes: a systematic review. Anesth Analg 2023 Jun 1;136(6):1096-1106. [doi: 10.1213/ANE.0000000000006159] [Medline: 36066429]
44. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digit Med 2020;3:17. [doi: 10.1038/s41746-020-0221-y] [Medline: 32047862]
45. Hashmi SA, Raja MHR, Arif A, Naseem Z, Pal KMB, Pal KMI. Reducing post-operative length of stay, is it worth the effort? World J Surg 2024 May;48(5):1096-1101. [doi: 10.1002/wjs.12115] [Medline: 38459712]

XSL•FO
RenderX

46. Aasbrenn M, Christiansen CF, Esen B, Suetta C, Nielsen FE. Mortality of older acutely admitted medical patients after early discharge from emergency departments: a nationwide cohort study. BMC Geriatr 2021 Jul 2;21(1):410. [doi: 10.1186/s12877-021-02355-y] [Medline: 34215192]

Abbreviations

AI: artificial intelligence
CDSS: clinical decision support system

Edited by M Antoniou; submitted 21.01.25; peer-reviewed by C Green, T Chang; revised version received 07.08.25; accepted 14.08.25; published 17.10.25.

Please cite as:
Parchmann N, Orzechowski M, Brefka S, Steger F
Evaluation of an AI-Based Clinical Decision Support System for Perioperative Care of Older Patients: Ethical Analysis of Focus Groups With Older Adults
JMIR Aging 2025;8:e71568
URL: https://aging.jmir.org/2025/1/e71568
doi:10.2196/71568

© Nina Parchmann, Marcin Orzechowski, Simone Brefka, Florian Steger. Originally published in JMIR Aging (https://aging.jmir.org), 17.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Comparative Diagnostic Accuracy of AI-Assisted Fluorine-18 Fluorodeoxyglucose Positron Emission Tomography Versus Structural Magnetic Resonance Imaging in Alzheimer Disease: Systematic Review and Meta-Analysis

Bingbing Wang[1*], PhD; Tailiang Zhao[1*], PhD; Rongrong Ma[2*], MSc; Xiaochuan Huo[1], PhD; Xiaoxiao Xiong[3], PhD; Minjie Wu[1], PhD; Yuran Wang[4], MSc; Liu Liu[3], PhD; Zhijiang Zhuang[3], PhD; Bin Wang[4,5], PhD; Jixin Shou[1*], PhD

[1]Department of Neurosurgery, Fifth Affiliated Hospital of Zhengzhou University, No. 3 Kangfu-qian Street, Zhengzhou, China

[2]Department of Anesthesiology and Perioperative Medicine, Xuchang Central Hospital, Henan University of Science and Technology, Xuchang, China

[3]Department of Neurology, Fifth Affiliated Hospital of Zhengzhou University, Zhengzhou, China

[4]The Fifth Clinical Medical College, Henan Medical College, Zhengzhou University, Zhengzhou, China

[5]Department of Cardiology, Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua University, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Tailiang Zhao, PhD
Department of Neurosurgery, Fifth Affiliated Hospital of Zhengzhou University, No. 3 Kangfu-qian Street, Zhengzhou, China

Abstract

Background: Neuroimaging is crucial in the diagnosis of Alzheimer disease (AD). In recent years, artificial intelligence (AI)–based neuroimaging technology has rapidly developed, providing new methods for accurate diagnosis of AD, but its performance differences still need to be systematically evaluated.

Objective: This study aims to conduct a systematic review and meta-analysis comparing the diagnostic performance of AI-assisted fluorine-18 fluorodeoxyglucose positron emission tomography (^{18}F-FDG PET) and structural magnetic resonance imaging (sMRI) for AD.

Methods: Databases including Web of Science, PubMed, and Embase were searched from inception to January 2025 to identify original studies that developed or validated AI models for AD diagnosis using ^{18}F-FDG PET or sMRI. Methodological quality was assessed using the TRIPOD-AI (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis–Artificial Intelligence) checklist. A bivariate mixed-effects model was employed to calculate pooled sensitivity, specificity, and summary receiver operating characteristic curve area (SROC-AUC).

Results: A total of 38 studies were included, with 28 moderate-to-high-quality studies analyzed. Pooled SROC-AUC values were 0.94 (95% CI 0.92 - 0.96) for sMRI and 0.96 (95% CI 0.94 - 0.98) for ^{18}F-FDG PET, demonstrating statistically significant intermodal differences (P=.02). Subgroup analyses revealed that for machine learning, pooled SROC-AUCs were 0.89 (95% CI 0.86 - 0.92) for sMRI and 0.95 (95% CI 0.92 - 0.96) for ^{18}F-FDG PET, while for deep learning, these values were 0.96 (95% CI 0.94 - 0.97) and 0.97 (95% CI 0.96 - 0.99), respectively. Meta-regression identified heterogeneity arising from study quality stratification, algorithm types, and validation strategies.

Conclusions: Both AI-assisted ^{18}F-FDG PET and sMRI exhibit high diagnostic accuracy in AD, with ^{18}F-FDG PET demonstrating superior overall diagnostic performance compared to sMRI.

(JMIR Aging 2025;8:e76981) doi:10.2196/76981

KEYWORDS

18F-FDG PET; Alzheimer disease; artificial intelligence; deep learning; machine learning; Preferred Reporting Items for Systematic Reviews and Meta-Analyses; PRISMA; sMRI; structural magnetic resonance imaging; systematic review

Introduction

Alzheimer disease (AD) is a progressive neurodegenerative disorder characterized by insidious onset, cognitive decline, and memory impairment. In the United States, approximately 6.7 million adults aged ≥65 years are affected by AD, while China faces an even greater burden, with over 13 million cases [1,2]. The prolonged disease course and high comorbidity rates

XSL•FO
RenderX

have established AD as one of the most fatal and economically burdensome conditions of the 21st century [3-5]. Accurate diagnosis of AD is critical for therapeutic decision-making and prognostic evaluation, particularly in the context of global population aging and increasing demands for precise patient stratification in clinical trials [6,7].

Neuroimaging modalities, including fluorine-18 fluorodeoxyglucose positron emission tomography (^{18}F-FDG PET) and structural magnetic resonance imaging (sMRI), have become cornerstone technologies in AD diagnostic frameworks (eg, National Institute on Aging-Alzheimer's Association criteria) because of their noninvasive nature and quantitative capabilities [8-10]. Recent advancements in artificial intelligence (AI) have revolutionized medical image analysis: machine learning (ML) enables data-driven predictive modeling beyond traditional rule-based programming, while deep learning (DL), an advanced ML paradigm, employs multilayer neural networks to extract high-level features from complex datasets, demonstrating transformative potential in neuroimaging [11-14]. Despite extensive research on AI-assisted sMRI or ^{18}F-FDG PET for AD diagnosis, two critical challenges hinder evaluation of model generalizability: (1) substantial heterogeneity in algorithm designs and validation frameworks, and (2) significant variability in the quality of individual studies [15]. Furthermore, no high-quality meta-analysis has comprehensively compared the diagnostic performance of these two imaging modalities.

This systematic review and meta-analysis addresses three critical objectives: (1) quantitative evaluation of diagnostic accuracy metrics for AI-assisted ^{18}F-FDG PET and sMRI based on moderate-to-high-quality evidence; (2) direct comparison of diagnostic performance between modalities; and (3) investigation of confounding factors, including study quality (moderate-to-high vs low), algorithm types (ML vs DL), and validation strategies (internal vs external), through meta-regression. The findings are anticipated to inform evidence-based optimization of AD diagnostic pathways.

Methods

Overview

This study adhered to the PRISMA-DTA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses of Diagnostic Test Accuracy) guidelines and was prospectively registered on PROSPERO (ID: CRD42023449927) to ensure transparency and minimize reporting bias [16,17]. Two reviewers independently conducted all stages of the review process, including title and abstract screening, full-text evaluation, data extraction, adherence assessment to reporting guidelines, and risk-of-bias evaluation. Discrepancies were resolved through group consensus.

Literature Search Strategy

Two investigators (TLZ and BW) systematically searched PubMed, Web of Science, and Embase from inception to January 2025 using a combination of Medical Subject Headings (MeSH) terms and free-text keywords. Additional searches were performed on clinical trial registries and OpenGrey to identify unpublished clinical trials and gray literature. Search terms encompassed four domains: (1) disease terminology (Alzheimer disease, AD, Alzheimer Syndrome); (2) imaging modalities (^{18}F-FDG PET, sMRI); (3) AI methodologies (ML, DL); and (4) diagnostic metrics (sensitivity, specificity, summary receiver operating characteristic curve area [SROC-AUC]). Reference lists of included studies were manually screened to identify additional relevant publications. The specific search strategies are provided in Table S1 in Multimedia Appendix 1.

Literature Inclusion and Exclusion Criteria

The inclusion criteria were as follows: (1) human studies developing or validating AI models using ^{18}F-FDG PET or sMRI to differentiate AD from normal controls; (2) AD diagnosis based on National Institute on Aging-Alzheimer's Association or International Working Group for New Research Criteria for Alzheimer’s Disease criteria; (3) availability of diagnostic performance metrics (eg, true positives, false positives, true negatives, false negatives) or explicit reporting of sensitivity and specificity; and (4) full-text availability in English. The exclusion criteria were as follows: (1) case reports, reviews, letters, or conference abstracts; (2) studies lacking sufficient diagnostic performance data; and (3) duplicate publications reporting on the same cohort without novel analyses.

Literature Screening and Data Extraction

Two reviewers (TZ and BW) independently performed title and abstract and full-text screening, followed by cross-verification to ensure accuracy. Two additional investigators (RM and XH) extracted data using predefined forms, including study characteristics (first author, publication year), model specifications (algorithm type, validation strategies), and diagnostic performance metrics (2×2 contingency tables, sensitivity, specificity). Extracted data were cross-checked for completeness and precision.

Quality Assessment of Included Studies

The methodological quality of included studies was evaluated using an adapted TRIPOD-AI (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis–Artificial Intelligence) checklist (Table S2 in Multimedia Appendix 1) by two independent reviewers (RM and XH) [18]. The instrument assessed four domains: (1) data quality (sample diversity, adequacy, preprocessing standardization), (2) model development (feature extraction rationale, algorithm selection, hyperparameter optimization), (3) validation methods (cross-validation rigor, external validation), and (4) clinical applicability (interpretability, net clinical benefit). Each of the nine items was scored 0‐2 (0=unsatisfied; 1=partially satisfied; 2=fully satisfied), with total scores categorized as high quality (≥16), moderate quality (10-15), or low quality (≤9). Disagreements were resolved through discussion.

Data Analysis

Diagnostic performance metrics were calculated using a bivariate mixed-effects model. SROC-AUC served as the primary outcome due to its threshold independence. Diagnostic

XSL•FO
RenderX

accuracy was classified per National Institutes of Health criteria as high (AUC≥0.90), moderate (0.70 - 0.89), or low (AUC<0.70) [19].

Heterogeneity Assessment and Publication Bias Evaluation

Heterogeneity was quantified using Cochran *Q* test (significance defined at the .05 level) and I^2 statistics (25%: low; 50%: moderate; 75%: high heterogeneity) [20]. Threshold effects were assessed using Spearman correlation between logit-transformed sensitivity and 1–specificity. Sensitivity analyses evaluated outlier influence by iteratively excluding individual studies. Univariable meta-regression analyses were conducted to assess the influence of potential confounding factors, including study quality (moderate to high vs low quality), algorithm type (ML vs DL), and validation strategy (internal vs external validation). Statistical significance of modifiers of the pooled effect size was defined at the .05 level after adjustment via the Holm-Bonferroni method. Publication bias was assessed via funnel plot asymmetry and Egger test, with significance defined at the .05 level. All analyses were conducted in Stata 16.0 using the MIDAS commands.

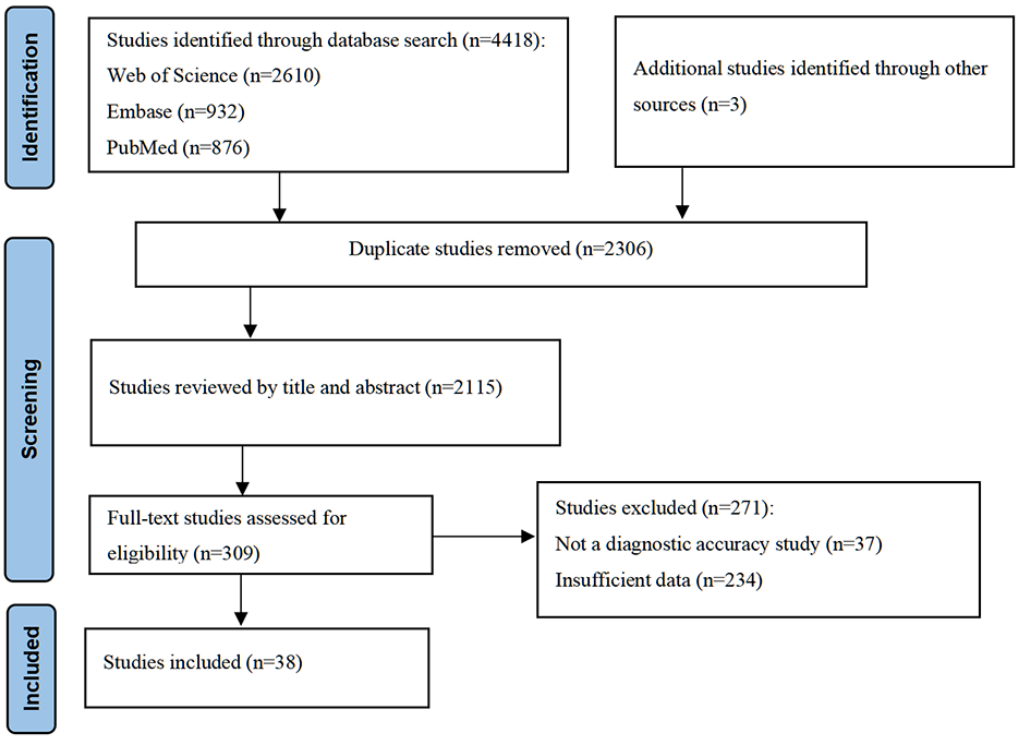
As this study focused on diagnostic accuracy, we employed the bivariate random-effects model via the MIDAS command in Stata to jointly synthesize sensitivity and specificity. This approach accounts for both within- and between-study variability, as well as the inherent correlation between sensitivity and specificity. In such models, conventional heterogeneity measures like I^2, although still reported for completeness, may not fully capture the joint variability of test performance measures. Therefore, to better understand heterogeneity, we additionally performed meta-regression and calculated joint likelihood ratio tests to explore potential effect modifiers. These methods align with recommended practices in diagnostic test accuracy meta-analyses [21,22].

Results

Literature Screening Process and Results

Initial searches identified 876 PubMed records (n=89 involving ^{18}F-FDG PET; n=787 sMRI), 932 Embase records (n=111 ^{18}F-FDG PET; n=821 sMRI), and 2610 Web of Science records (n=230 ^{18}F-FDG PET; n=2380 sMRI), supplemented by 3 additional studies. After deduplication in EndNote 20, 2 reviewers independently screened titles and abstracts using the Population, Intervention, Comparison, Outcome, Study Design (PICOS) criteria, followed by full-text evaluation against predefined inclusion/exclusion criteria. This rigorous process yielded 38 studies for systematic review and meta-analysis [23-60]. The literature screening process and results are shown in Figure 1.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) flow diagram.



Characteristics of Included Studies

Table 1 and Table S3 in Multimedia Appendix 1 detail the characteristics of the included studies. Imaging modalities comprised sMRI (n=15, 39%), ^{18}F-FDG PET (n=15, 39%), and combined sMRI/^{18}F-FDG PET (n=8, 22%). Data sources predominantly relied on open-access databases (n=29, 76%), with 28 (97%) from the Alzheimer's Disease Neuroimaging Initiative database and 1 (3%) from the Open Access Series of Imaging Studies database. Four studies (11%) used institutional data, while 5 (13%) combined Alzheimer's Disease Neuroimaging Initiative database with local datasets. Algorithmically, 26 studies (68%) employed ML, predominantly support vector machines (17/26, 65%), whereas 12 (32%)

XSL•FO
RenderX

utilized DL, primarily convolutional neural networks (9/12, 75%). Internal validation was implemented in 33 studies (87%), with 10-fold cross-validation in 16 (48%); only 5 studies (13%) incorporated external validation.

XSL•FO
RenderX

Table . Characteristics of all included studies (n=38).

Studies	Model development and internal validation							Model external validation	
	Definition	Modality	Data source	Algorithm used	Training/validation (ratio)	Testing	Type of internal validation	Data source	Definition
Zhang et al [23], 2011	AD[a]: 51; NC[b]: 52	sMRI[c]; PET[d]	ADNI[e] database	SVM[f]	NR[g]	NR	10-fold CV[h]	NR	NR
Yun et al [24], 2015	AD: 71; NC: 85	sMRI; PET	ADNI database	LDA[i]	NR	NR	LOOCV[j]	NR	NR
Westman et al [25], 2012	AD: 96; NC: 111	sMRI	ADNI database	OPLS[k]	NR	NR	7-fold CV	NR	NR
Vemuri et al [26], 2008	AD: 190; NC: 190	sMRI	Mayo Clinic	SVM	280	100	4-fold CV; hold-out	NR	NR
Suk et al [27], 2014	AD: 93; NC: 101	sMRI; PET	ADNI database	DBM[l]; SVM	NR	NR	10-fold CV	NR	NR
Sayeed et al [28], 2002	AD: 18; NC: 10	PET	Hammersmith Hospital	DFA[m]	NR	NR	LOOCV	NR	NR
Pan et al [29], 2019	AD: 247; NC: 246	PET	ADNI database	SVM	NR	NR	10-fold CV	NR	NR
Padilla et al [30], 2012	AD: 53; NC: 52	PET	ADNI database	SVM	NR	NR	LOOCV	NR	NR
Ni et al [31], 2021	AD: 638; NC: 629	PET	ADNI database	CNN[n]	1000	267	Hold-out	NR	NR
Magnin et al [32], 2009	AD: 16; NC: 22	sMRI	Pitié-Salpétriere Hospital	SVM	75%	25%	Bootstrap resampling	NR	NR
Lu et al [33], 2018	AD: 226; NC: 304	PET	ADNI database	MDNN[o]	NR	NR	10-fold CV	NR	NR
Liu et al [34], 2012	AD: 198; NC: 229	sMRI	ADNI database	SRC[p]	NR	NR	10-fold CV	NR	NR
Liu et al [35], 2018	AD: 93; NC: 100	PET	ADNI database	2D-CNN	NR	NR	10-fold CV	NR	NR
Li et al [36], 2015	AD: 44; NC: 45	PET	ADNI database; TUM database	GMM[q]; SVM	NR	NR	10-fold CV	NR	NR
Lerch et al [37], 2008	AD: 19; NC: 17	sMRI	Ludwig Maximilian University of Munich	QDA[r]	NR	NR	LOOCV	NR	NR
Kim et al [38], 2020	AD: 141; NC: 348	PET	ADNI database	CNN	NR	NR	NR	Severance dataset	AD: 80; NC: 72
Kim et al [39], 2020	AD: 139; NC: 347	PET	ADNI database	BEGAN[s]; SVM	NR	NR	NR	Severance dataset	AD: 73; NC: 68
Katako et al [40], 2018	AD: 94; NC: 111	PET	ADNI database	SVM	NR	NR	10-fold CV	NR	NR
Ismail et al [41], 2023	AD: 511; NC: 535	sMRI; PET	ADNI database	MultiAzNet; SVM	NR	NR	10-fold CV	NR	NR
Illán et al [42], 2011	AD: 95; NC: 97	PET	ADNI database	SVM	NR	NR	LOOCV	NR	NR
Hinrichs et al [43], 2009	AD: 89; NC: 94	sMRI; PET	ADNI database	LPboosting[t]	NR	NR	Leave-many-out CV	NR	NR

XSL•FO
RenderX

Studies	Model development and internal validation							Model external validation	
	Definition	Modality	Data source	Algorithm used	Training/validation (ratio)	Testing	Type of internal validation	Data source	Definition
Gray et al [44], 2012	AD: 50; NC: 54	PET	ADNI database	SVM	75%	25%	Monte Carlo CV	NR	NR
Gray et al [45], 2013	AD: 37; NC: 35	sMRI; PET	ADNI database	RF[u]	75%	25%	Monte Carlo CV	NR	NR
Feng et al [46], 2019	AD: 93; NC: 100	sMRI; PET	ADNI database	FSBi-LSTM[v]; 3D-CNN	NR	NR	10-fold CV	NR	NR
Cuingnet et al [47], 2011	AD: 137; NC: 162	sMRI	ADNI database	SVM	50%	50%	Hold-out	NR	NR
Chen et al [48], 2022	AD: 326; NC: 413	sMRI	ADNI-1 database; ADNI-2 database	2D-CNN; 3D-CNN	NR	NR	NR	ADNI-1 database; ADNI-2 database	AD: 326; NC: 413
Song et al [49], 2021	AD: 95; NC: 126	sMRI; PET	ADNI database	3D-CNN	NR	NR	10-fold CV	NR	NR
Li et al [50], 2019	AD: 130; NC: 162	PET	ADNI database	SVM	70%	30%	Monte Carlo CV	Huashan database	AD: 22; NC: 22
Ahila et al [51], 2022	AD: 220; NC: 635	PET	ADNI database	2D-CNN	90%	10%	Hold-out	NR	NR
Toussaint et al [52], 2012	AD: 80; NC: 80	PET	ADNI database	SVM	NR	NR	LOOCV	NR	NR
Tong et al [53], 2014	AD: 198; NC: 231	sMRI	ADNI database	SVM	NR	NR	LOOCV	NR	NR
Min et al [54], 2014	AD: 97; NC: 128	sMRI	ADNI database	SVM	NR	NR	10-fold CV	NR	NR
Jin et al [55], 2020	AD: 488; NC: 536	sMRI	In-house database; ADNI database	3D attention network	NR	NR	10-fold CV; leave center out CV	ADNI database; in-house database	AD: 488; NC: 536
Cho et al [56], 2012	AD: 128; NC: 160	sMRI	ADNI database	LDA	142	146	Hold-out	NR	NR
Chincarini et al [57], 2011	AD: 144; NC: 189	sMRI	ADNI database	RF; SVM	NR	NR	20-fold CV	NR	NR
Beheshti et al [58], 2017	AD: 92; NC: 94	sMRI	ADNI database	SVM	NR	NR	10-fold CV	NR	NR
Anandh et al [59], 2016	AD: 30; NC: 55	sMRI	OASIS[w] database	SVM	NR	NR	10-fold CV	NR	NR

XSL•FO
RenderX

Studies	Model development and internal validation							Model external validation	
	Definition	Modality	Data source	Algorithm used	Training/validation (ratio)	Testing	Type of internal validation	Data source	Definition
Amoroso et al [60], 2018	AD: 86; NC: 81	sMRI	ADNI database	RF	67	100	10-fold CV; hold-out	NR	NR

[a]AD: Alzheimer disease.
[b]NC: normal cognitive.
[c]sMRI: structural magnetic resonance imaging.
[d]PET: positron emission tomography.
[e]ADNI: Alzheimer's Disease Neuroimaging Initiative.
[f]SVM: support vector machine.
[g]NR: no report.
[h]CV: cross-validation.
[i]LDA: linear discriminant analysis.
[j]LOOCV: leave-one-out cross-validation.
[k]OPLS: orthogonal partial least squares.
[l]DBM: deep Boltzmann machine.
[m]DFA: discriminant function analysis.
[n]CNN: convolutional neural network.
[o]MDNN: multiscale deep neural network.
[p]SRC: sparse representation–based classifier.
[q]GMM: Gaussian mixture model.
[r]QDA: quadratic discriminant analysis.
[s]BEGAN: boundary equilibrium generative adversarial network.
[t]LP: linear program.
[u]RF: random forest.
[v]FSBi-LSTM: fully stacked bidirectional long short-term memory.
[w]OASIS: Open Access Series of Imaging Studies.

Quality Assessment Results of Included Studies

Nine (24%) studies were high quality, characterized by adequate sample sizes, rigorous validation (eg, multicenter external validation), and standardized reporting. Moderate-quality studies (n=19, 50%) met basic methodological standards (eg, cross-validation) but had limitations such as single-center data or insufficient clinical correlation analyses. Ten (26%) studies were low quality due to small sample sizes (<100 cases) and inadequate validation strategies, compromising external validity. The results of quality assessments for each study are detailed in Table S4 in Multimedia Appendix 1.

Data Analysis Results

Overall Diagnostic Performance

Analysis of 29 sMRI-related contingency tables demonstrated a pooled sensitivity of 88% (95% CI 86%-89%), a specificity of 92% (95% CI 90%-93%), and an SROC-AUC of 0.94 (95% CI 0.92-0.96). Subgroup analyses stratified by algorithm type revealed that ML (19 tables) achieved a sensitivity of 86% (95% CI 85%-88%), a specificity of 91% (95% CI 88%-93%), and an SROC-AUC of 0.88 (95% CI 0.85-0.90), while DL (10 tables) showed improved performance with a sensitivity of 88% (95% CI 86%-91%), a specificity of 92% (95% CI 90%-94%), and an SROC-AUC of 0.96 (95% CI 0.94-0.97). Stratification by study quality indicated that moderate-to-high-quality studies (24 tables) yielded a sensitivity of 87% (95% CI 85%-89%), a specificity of 91% (95% CI 89%-93%), and an SROC-AUC of 0.94 (95% CI 0.92-0.96), whereas low-quality studies (5 tables) reported elevated metrics with a sensitivity of 91% (95% CI 87%-94%), a specificity of 95% (95% CI 92%-97%), and an SROC-AUC of 0.98 (95% CI 0.96-0.99). Validation strategy–based analysis showed that internal validation (25 tables) achieved a sensitivity of 88% (95% CI 86%-90%), a specificity of 92% (95% CI 90%-94%), and an SROC-AUC of 0.95 (95% CI 0.92-0.96), while external validation (4 tables) demonstrated marginally lower performance with a sensitivity of 85% (95% CI 81%-89%), a specificity of 91% (95% CI 85%-94%), and an SROC-AUC of 0.93 (95% CI 0.90-0.95).

For ^{18}F-FDG PET (27 tables), pooled estimates demonstrated a sensitivity of 90% (95% CI 88%-92%), a specificity of 93% (95% CI 91%-94%), and an SROC-AUC of 0.96 (95% CI 0.94-0.98). Subgroup analyses stratified by algorithm type revealed that ML (15 tables) achieved a sensitivity of 89% (95% CI 86%-90%), a specificity of 91% (95% CI 88%-93%), and an SROC-AUC of 0.94 (95% CI 0.91-0.96), while DL (12 tables) exhibited superior performance with a sensitivity of 91% (95% CI 89%-93%), a specificity of 94% (95% CI 93%-96%), and an SROC-AUC of 0.98 (95% CI 0.96-0.99). Moderate-to-high-quality studies (21 tables) demonstrated a sensitivity of 90% (95% CI 88%-92%), a specificity of 93% (95% CI 91%-94%), and an SROC-AUC of 0.96 (95% CI

XSL•FO
RenderX

0.94-0.98), whereas low-quality studies (6 tables) showed comparable metrics with a sensitivity of 91% (95% CI 86%-94%), a specificity of 93% (95% CI 87%-96%), and an SROC-AUC of 0.96 (95% CI 0.94-0.98). Internal validation (24 tables) yielded a sensitivity of 90% (95% CI 89%-92%), a specificity of 93% (95% CI 91%-94%), and an SROC-AUC of 0.96 (95% CI 0.94%-0.97%), while external validation data (3 tables) were insufficient for SROC-AUC calculation but reported a sensitivity of 87% (95% CI 81%-93%) and a specificity of 95% (95% CI 91%-97%). The data analysis results are shown in Table 2.

XSL•FO
RenderX

Table . Summary estimates and meta-regression of pooled performance of all studies.

Main and subordinate directory	Tables, n	Sensitivity (%; 95% CI)	I^2	Specificity (%; 95% CI)	I^2	Joint *P* value	AUC[a] (95% CI)
sMRI[b]							
Overall	29	88 (86-89)	55.32	92 (90-93)	74.08	—[j]	0.94 (0.92-0.96)
Study quality						.02	
Medium-to-high quality	24	87 (85-89)	57.99	91 (89-93)	75.94		0.94 (0.92-0.96)
Low quality	5	91 (87-94)	10.2	95 (92-97)	0		0.98 (0.96-0.99)
Algorithm type						.40	
ML[d]	19	86 (85-88)	21.23	91 (88-93)	73.82		0.88 (0.85-0.90)
DL[e]	10	88 (86-91)	76.56	92 (90-94)	76.53		0.96 (0.94-0.97)
Validation strategy						.33	
Internal validation	25	88 (86-90)	46.38	92 (90-94)	71.65		0.95 (0.92-0.96)
External validation	4	85 (81-89)	72.26	91 (85-94)	83.65		0.93 (0.90-0.95)
SVM[f]	10	87 (85-90)	17.47	93 (90-95)	57.65		0.93 (0.90-0.95)
CNN[g]	9	88 (85-91)	78.19	92 (90-94)	77.55		0.96 (0.93-0.97)
^{18}F-FDG PET[k]							
Overall	27	90 (88-92)	42.04	93 (91-94)	61.49	—	0.96 (0.94-0.98)
Study quality						.96	
Medium-to-high quality	21	90 (88-92)	44.58	93 (91-94)	58.52		0.96 (0.94-0.98)
Low quality	6	91 (86-94)	42.73	93 (87-96)	72.3		0.96 (0.94-0.98)
Algorithm type						.01	
ML	15	89 (86-90)	1.61	91 (88-93)	45.64		0.94 (0.91-0.96)
DL	12	91 (89-93)	55.23	94 (93-96)	45.92		0.98 (0.96-0.99)
Validation strategy						.26	
Internal validation	24	90 (89-92)	34.61	93 (91-94)	64.9		0.96 (0.94-0.97)
External validation	3	87 (81-93)	—	95 (91-99)	—		—
SVM	11	90 (87-92)	0	92 (89-94)	37.5		0.94 (0.92-0.96)
CNN	8	91 (88-94)	69.76	93 (92-95)	34.63		0.97 (0.95-0.98)

[a]AUC: area under the receiver operating characteristic curve.

[b]sMRI: structural magnetic resonance imaging.

[c]Not applicable.

[d]ML: machine learning.

[e]DL: deep learning.

XSL•FO
RenderX

[f]SVM: support vector machine.

[g]CNN: convolutional neural network.

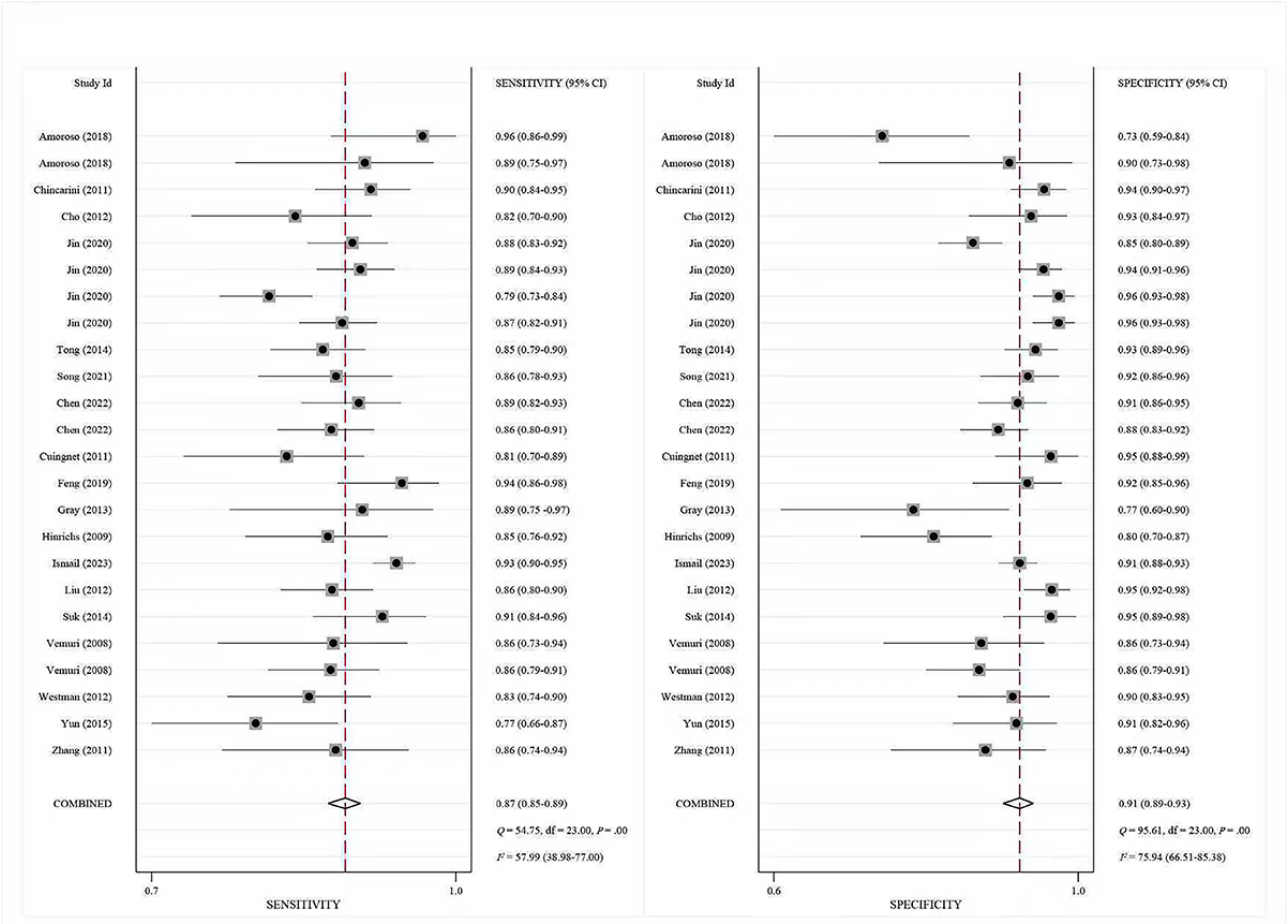
[h]^{18}F-FDG PET: fluorine-18 fluorodeoxyglucose positron emission tomograph.

Moderate-to-High-Quality Study Subgroup Analysis

For sMRI, pooled sensitivity, specificity, and SROC-AUC were 87% (95% CI 85% - 89%), 91% (95% CI 89% - 93%), and 0.94 (95% CI 0.92 - 0.96), respectively (see Figures 2 and 3). Stratification by algorithm type indicated that ML models achieved 86% (95% CI 84% - 88%) sensitivity, 89% (95% CI 86% - 92%) specificity, and an SROC-AUC of 0.89 (95% CI 0.86 - 0.92), while DL models demonstrated 88% (95% CI 86% - 91%) sensitivity, 92% (95% CI 90% - 94%) specificity, and an SROC-AUC of 0.96 (95% CI 0.94 - 0.97).

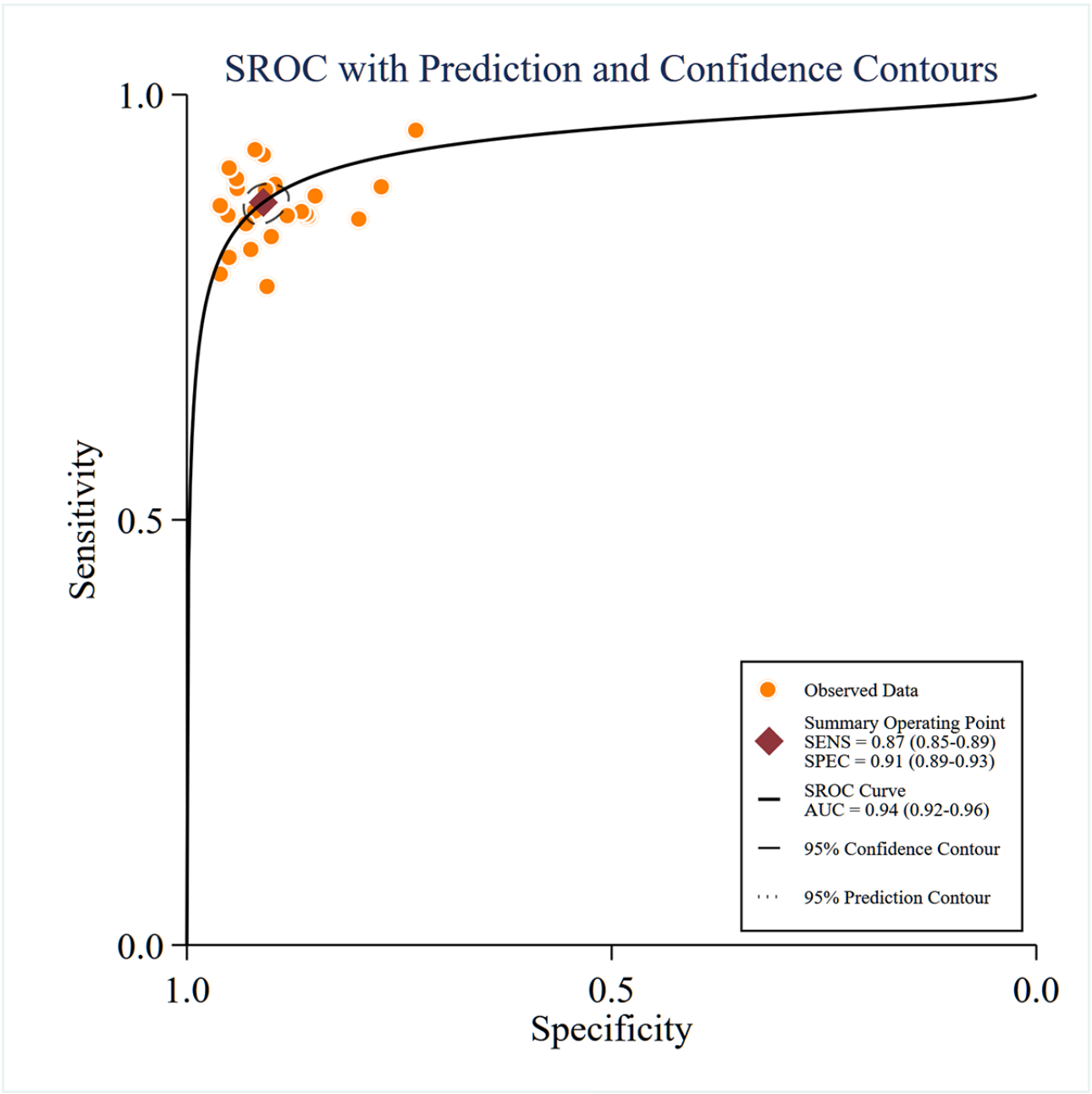
For ^{18}F-FDG PET, pooled sensitivity, specificity, and SROC-AUC were 90% (95% CI 88% - 92%), 93% (95% CI 91% - 94%), and 0.96 (95% CI 0.94 - 0.98), respectively (see Figures 4 and 5). Subgroup analyses revealed ML models achieved 89% (95% CI 86 - 91%) sensitivity, 91% (95% CI 87% - 93%) specificity, and an SROC-AUC of 0.95 (95% CI 0.92 - 0.96), whereas DL models outperformed with 91% (95% CI 89% - 93%) sensitivity, 94% (95% CI 93% - 96%) specificity, and SROC-AUC 0.97 (95% CI 0.96 - 0.99). The data analysis results are also shown in Table 3.

Figure 2. The pooled sensitivities and specificities based on moderate-to-high-quality studies for structural magnetic resonance imaging [23,24,25,26,27,34,41,43,45,46,47,48,49,53,55,57,60].



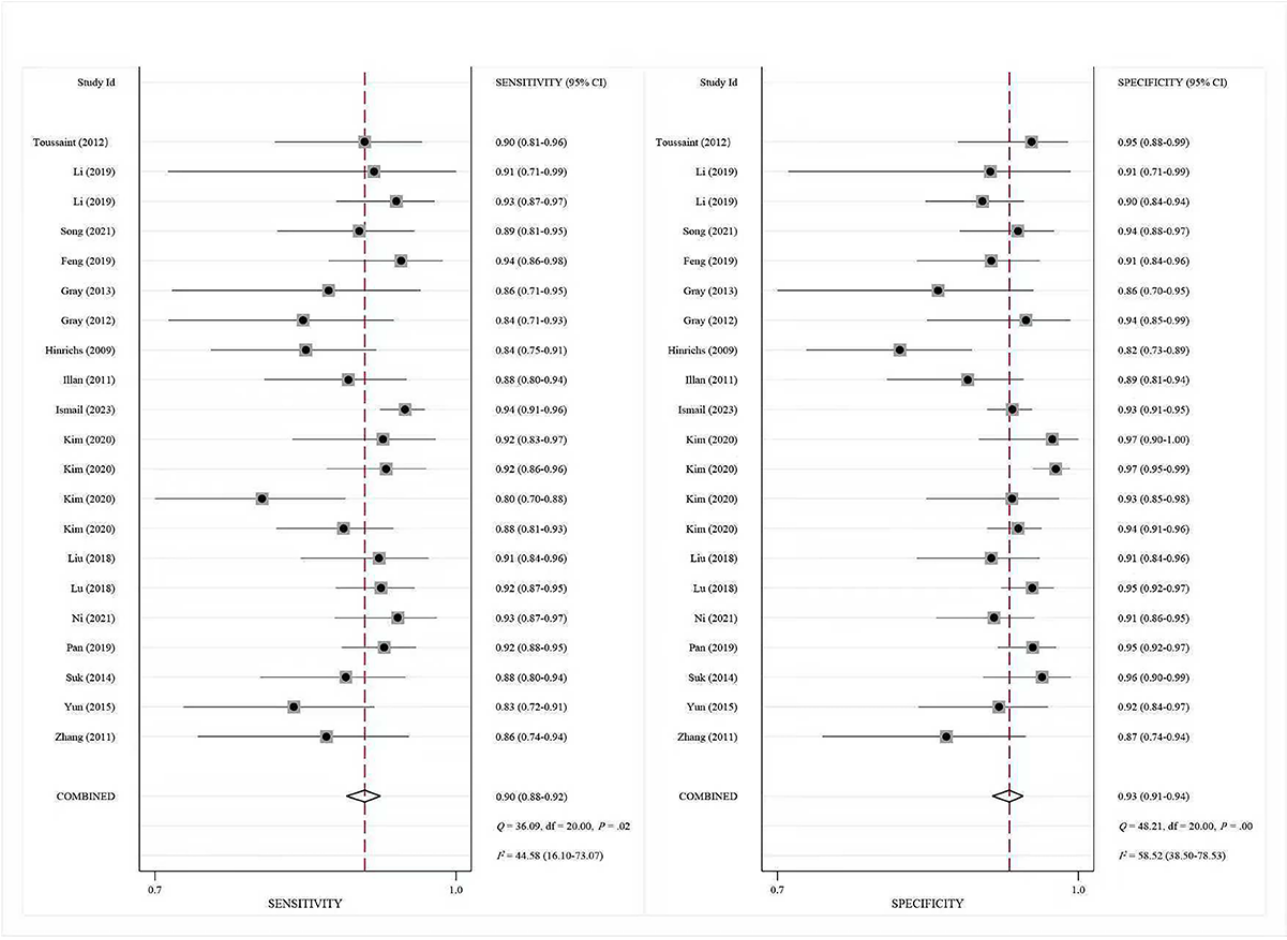
XSL•FO
RenderX

Figure 3. Summary receiver operating characteristic (SROC) based on moderate-to-high-quality studies curves for structural magnetic resonance imaging. AUC: area under the receiver operating characteristic curve area; SENS: sensitivity; SPEC: specificity.



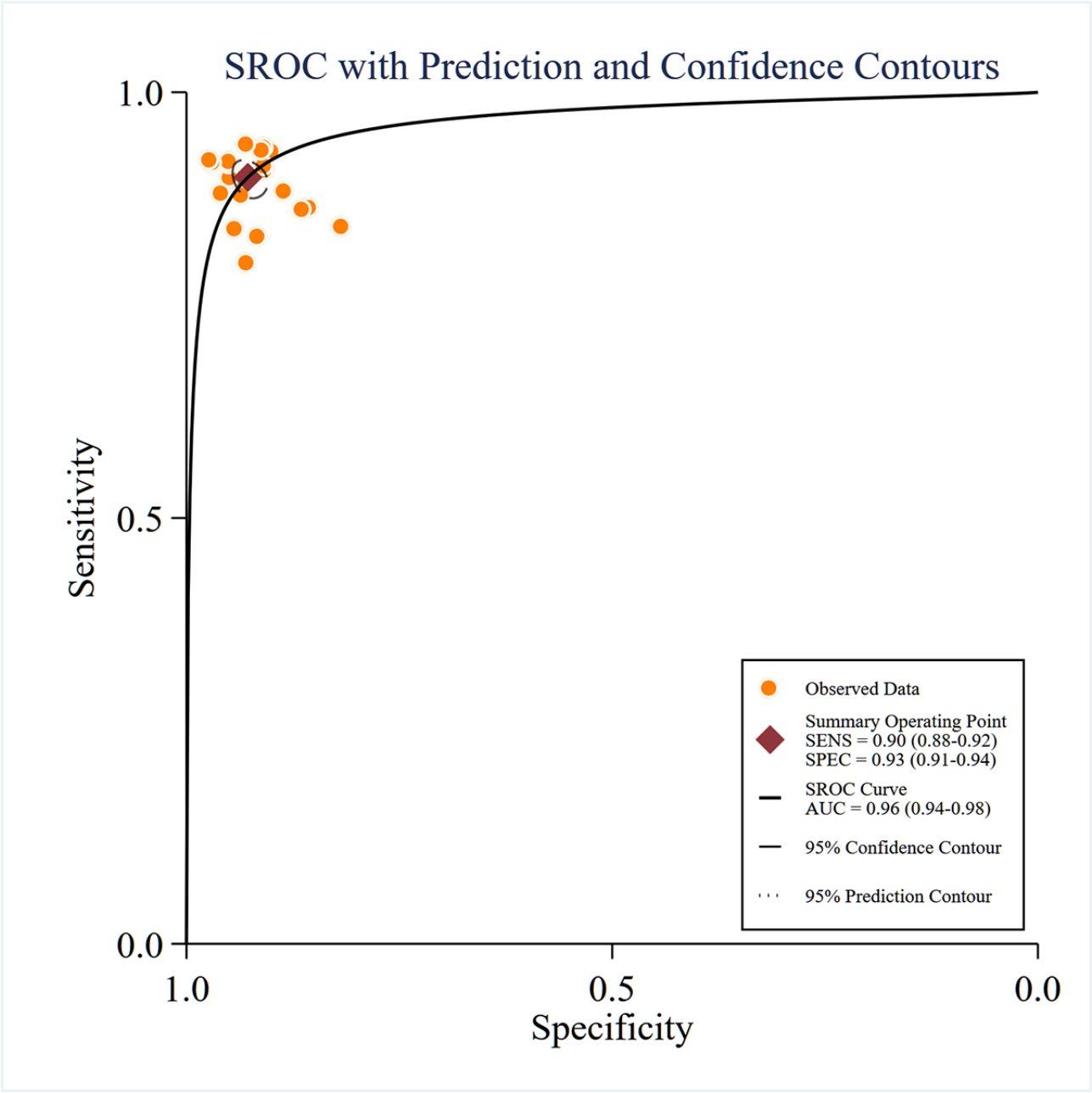
XSL•FO
RenderX

Figure 4. The pooled sensitivities and specificities based on moderate-to-high-quality studies for fluorine-18 fluorodeoxyglucose positron emission tomography (^{18}F-FDG PET) [23,24,27,29,31,33,35,38,39,41,42,43,44,45,45,46,49,50].



XSL•FO
RenderX

Figure 5. Summary receiver operating characteristic (SROC) based on moderate-to-high-quality studies curves for fluorine-18 fluorodeoxyglucose positron emission tomography (^{18}F-FDG PET). AUC: area under the receiver operating characteristic curve area; SENS: sensitivity; SPEC: specificity.



XSL•FO
RenderX

Table . Summary estimates and meta-regression of pooled performance of moderate-to-high-quality studies.

Main and subordinate directory	Tables, n	Sensitivity (%; 95% CI)	I^2	Specificity (%; 95% CI)	I^2	Joint *P* value	AUC[a] (95% CI)
Overall						.02	
sMRI[b]	24	87 (85-97)	57.99	91 (89-93)	75.94		0.94 (0.92-0.96)
^{18}F-FDG PET[c]	21	90 (88-92)	44.58	93 (91-94)	58.52		0.96 (0.94-0.98)
ML[d]							
sMRI	14	86 (84-88)	3.72	89 (86-92)	75.58		0.89 (0.86-0.92)
^{18}F-FDG PET	10	89 (86-91)	16.51	91 (87-93)	53.95		0.95 (0.92-0.96)
DL[e]							
sMRI	10	88 (86-91)	76.56	92 (90-94)	76.53		0.96 (0.94-0.97)
^{18}F-FDG PET	11	91 (89-93)	54.59	94 (93-95)	36.23		0.97 (0.96-0.99)

[a]AUC: area under the receiver operating characteristic curve.

[b]sMRI: structural magnetic resonance imaging.

[c]^{18}F-FDG PET: fluorine-18 fluorodeoxyglucose positron emission tomography.

[d]ML: machine learning.

[e]DL: deep learning.

Exploration of Statistical Heterogeneity Sources

Full Study Cohort

For sMRI contingency tables, sensitivity exhibited moderate heterogeneity (I^2=55.32%), while specificity showed high heterogeneity (I^2=74.08%). Threshold effect testing revealed no significant threshold effect (r=−0.101, P=.60). Subgroup analyses identified study quality stratification (P=.02) as a source of heterogeneity, with moderate-to-high-quality studies demonstrating lower sensitivity (87% vs 91%) and specificity (91% vs 95%) compared to low-quality studies. Algorithm type (P=.40) and validation strategy (P=.33) were not significant contributors.

For ^{18}F-FDG PET analyses, sensitivity and specificity displayed moderate (I^2=42.04%) and high heterogeneity (I^2=61.49%), respectively, with no threshold effect (r=−0.087, P=.67). Algorithm type (P=.01) significantly influenced heterogeneity, as ML models demonstrated lower sensitivity (89% vs 91%) and specificity (91% vs 94%) than DL. Study quality (P=.96) and validation strategy (P=.26) showed no significant impact. The data analysis results are also shown in Table 2.

Moderate-to-High-Quality Study Subgroups

For sMRI-based models, ML implementations demonstrated minimal sensitivity heterogeneity (I^2=3.72%) but high specificity heterogeneity (I^2=75.58%). Exclusion of studies using traditional ensemble algorithms (random forest, boosting) reduced specificity heterogeneity to moderate (I^2=45.47%), revealing that ensemble methods achieved higher sensitivity (89% vs 87%) but significantly lower specificity (79% vs 93%) compared to nonensemble ML models (P<.001). In contrast, DL-based sMRI models exhibited high heterogeneity for both sensitivity (I^2=76.56%) and specificity (I^2=76.53%). Removing external validation data mitigated heterogeneity to moderate (I^2=54.36%) and low levels (I^2=38.60%), respectively, with external validation studies showing significantly reduced sensitivity (85% vs 90%) and specificity (91% vs 93%) compared to internal validation (P<.001).

For ^{18}F-FDG PET-based models, ML implementations showed low sensitivity heterogeneity (I^2=16.51%) and moderate specificity heterogeneity (I^2=53.95%). DL models exhibited moderate sensitivity heterogeneity (I^2=54.59%) and low specificity heterogeneity (I^2=36.23%). The data analysis results also are shown in Table 3 and Table S5 in Multimedia Appendix 1.

Publication Bias Test

Funnel plot analysis of the full study cohort revealed no significant publication bias for sMRI (P=.69), whereas the ^{18}F-FDG PET subgroup exhibited significant bias (P<.001). In the moderate-to-high-quality cohort, both sMRI (P=.03) and ^{18}F-FDG PET (P=.01) demonstrated publication bias. However, subgroup analyses showed no significant bias for sMRI+ML (P=.06), sMRI+DL (P=.89), ^{18}F-FDG PET+ML (P=.08), or ^{18}F-FDG PET+DL (P=.28).

XSL•FO
RenderX

Discussion

Principal Findings

This meta-analysis confirms that AI-enhanced sMRI and ^{18}F-FDG PET achieve high diagnostic accuracy for AD, with pooled SROC-AUCs of 0.94 and 0.96, respectively, outperforming conventional visual assessments [61,62]. ^{18}F-FDG PET demonstrated superior overall performance (P=.02), likely attributable to its sensitivity to AD-specific metabolic abnormalities. Notably, ML amplified PET's diagnostic advantage (SROC-AUC: 0.95 vs 0.89), while DL narrowed the gap (SROC-AUC: 0.97 vs 0.96), highlighting DL's capacity to extract complex metabolic features and mitigate structural imaging limitations [63].

Subgroup analyses identified three key determinants of heterogeneity and performance variation. First, methodological quality significantly influenced sMRI models: moderate-to-high-quality studies reported lower sensitivity (87% vs 91%) and specificity (91% vs 95%) compared to low-quality studies (P=.02), suggesting inflated performance in the latter due to small sample sizes, single-center data, or nonstandardized preprocessing. Adherence to TRIPOD-AI guidelines and transparent reporting of preprocessing workflows are critical for future studies [18]. Second, algorithm type drove technical divergence in PET models: DL outperformed ML (sensitivity: 91% vs 89%; specificity: 94% vs 91%; P=.01) through end-to-end feature learning. However, ML's interpretability better aligns with clinical demands for transparency, whereas DL's reliance on large annotated datasets and computing resources limits its deployment in resource-constrained regions [64]. Third, validation strategies exposed generalizability limitations: external validation of sMRI+DL models showed reduced sensitivity (85% vs 90%) and specificity (91% vs 93%) compared to internal validation (P<.001).

The reduced diagnostic accuracy observed in externally validated models may stem from (1) data distribution shift—differences in feature distributions, class balance, and temporal trends between training and external datasets; (2) overfitting—models may have overfitted to training-specific noise or spurious patterns, limiting generalization; and (3) implementation and annotation inconsistencies—variations in data preprocessing, feature scaling, and labeling protocols across datasets [65].

Future Directions and Study Limitations

This study has two methodological constraints: potential selection bias from English-only inclusion and possible overestimation by prioritizing optimal contingency tables. Future research should focus on enhancing evidence robustness through multinational, multiethnic cohorts; improving transparency via open-source preprocessing codes, findable, accessible, interoperable, and reusable-compliant data sharing, and independent validation; and balancing performance and interpretability via explainable DL frameworks to meet clinical ethical standards [66].

We also acknowledge the risk of publication bias, as studies with positive results are more likely to be published, particularly in AI-related research where rapid progress and selective reporting may skew the literature. Additionally, language and geographic biases may exist since only English-language articles were included and most studies originated from a limited number of regions (eg, China, the United States, and Europe). This may limit the generalizability of our findings to underrepresented regions.

Despite ^{18}F-FDG PET's higher accuracy, its radiation exposure and cost hinder widespread screening [67]. Conversely, sMRI's cost-effectiveness and noninvasiveness position it as a first-line screening tool, with PET reserved for confirmatory testing in complex cases—a tiered diagnostic strategy. The 2023 Responsible AI for Social and Ethical Healthcare consensus implementation priorities, including transparent cost-benefit frameworks cross-modal standardization, and dynamic performance monitoring [68].

Comparison With Existing Reviews

Although previous reviews have explored the diagnostic accuracy of AI in AD, studies involving meta-analyses remain scarce [69-71]. Conducting such meta-analyses faces significant challenges, particularly due to substantial methodological heterogeneity across studies. This heterogeneity manifests in multiple dimensions, including variations in neuroimaging modalities, disparities in model validation strategies, and differences in algorithm types—all of which influence AD diagnostic performance and complicate the synthesis of evidence.

Borchert et al [70] conducted a comprehensive systematic review of 255 neuroimaging studies utilizing AI for dementia diagnosis and prognosis. Their findings demonstrated that discriminative models, particularly DL approaches, outperformed algorithmic classifiers in distinguishing AD patients from healthy controls. However, they emphasized critical methodological limitations, with conclusions primarily relying on qualitative synthesis rather than quantitative evidence.

In a systematic review and meta-analysis by Sun et al [71], the diagnostic accuracy of DL models based on ^{18}F-FDG PET for AD was investigated. While the study reported excellent diagnostic performance, notable heterogeneity was observed during meta-analysis, raising concerns about the reliability of the findings. Furthermore, the study focused exclusively on DL, overlooking the widespread application of traditional ML methods in current clinical research for AD diagnostic modeling.

In contrast, our meta-analysis incorporates a broader range of studies, rigorously controls methodological heterogeneity through stringent quality assessment and detailed subgroup analyses, and systematically evaluates the diagnostic accuracy of both ML and DL in AD. By emphasizing methodological rigor and the importance of external validation in AI-assisted neuroimaging for AD diagnosis, this study addresses critical gaps in the existing literature.

Although our subgroup comparisons between ML and DL models provide a useful overview of broad methodological trends, it must be noted that algorithm complexity, training data size, and model optimization procedures vary considerably

XSL•FO
RenderX

within each group. The observed performance differences may, therefore, reflect not only model class but also differences in dataset size, feature representation, and implementation quality. Future studies should aim to compare individual algorithms under standardized conditions.

Conclusions

In conclusion, AI can effectively support the diagnosis of AD using sMRI and ^{18}F-FDG PET imaging. Among these approaches, combining PET imaging with DL techniques yields the highest diagnostic accuracy. These findings suggest that a future direction lies in integrating precision neuroimaging with AI tools. To bring such systems into routine clinical use—helping doctors detect AD earlier, personalize treatment, and improve patient outcomes—future studies should focus on repeated validation with high-quality clinical datasets and the development of standardized implementation protocols.

Acknowledgments

This research was funded by the Henan Province Medical Science and Technology Public Relations Plan Joint Construction Project (LHGJ20230402).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Detailed database search process, characteristics of included studies, quality assessment results, and subgroup analysis findings.
[DOCX File, 59 KB - aging_v8i1e76981_app1.docx]

Checklist 1
PRISMA 2020 checklist.
[PDF File, 101 KB - aging_v8i1e76981_app2.pdf]

References

1. 2023 Alzheimer's disease facts and figures. Alzheimers Dement 2023 Apr;19(4):1598-1695. [doi: 10.1002/alz.13016]
2. Ren R, Qi J, Lin S, et al. The China Alzheimer Report 2022. Gen Psychiatr 2022;35(1):e100751. [doi: 10.1136/gpsych-2022-100751] [Medline: 35372787]
3. Jia J, Wei C, Chen S, et al. The cost of Alzheimer's disease in China and re-estimation of costs worldwide. Alzheimers Dement 2018 Apr;14(4):483-491. [doi: 10.1016/j.jalz.2017.12.006] [Medline: 29433981]
4. Scheltens P, Blennow K, Breteler MMB, et al. Alzheimer's disease. Lancet 2016 Jul 30;388(10043):505-517. [doi: 10.1016/S0140-6736(15)01124-1] [Medline: 26921134]
5. Pickett AC, Valdez D, White LA, et al. The CareVirtue digital journal for family and friend caregivers of people living with Alzheimer disease and related dementias: exploratory topic modeling and user engagement study. JMIR Aging 2024 Dec 24;7:e67992. [doi: 10.2196/67992] [Medline: 39719081]
6. Cahill S. WHO's global action plan on the public health response to dementia: some challenges and opportunities. Aging Ment Health 2020 Feb;24(2):197-199. [doi: 10.1080/13607863.2018.1544213] [Medline: 30600688]
7. Wu Y, Fu L, Li Q, et al. Recent advancements in the early diagnosis and treatment of Alzheimer's disease. Adv Therap 2023 Nov;6(11) [FREE Full text] [doi: 10.1002/adtp.202300181] [Medline: 36818419]
8. Jack CR, Bennett DA, Blennow K, et al. NIA‐AA Research Framework: toward a biological definition of Alzheimer's disease. Alzheimer's & Dementia 2018 Apr;14(4):535-562. [doi: 10.1016/j.jalz.2018.02.018]
9. Jack CR, Andrews JS, Beach TG, et al. Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup. Alzheimer's & Dementia 2024 Aug;20(8):5143-5169. [doi: 10.1002/alz.13859]
10. By S, Kahl A, Cogswell PM. Alzheimer's disease clinical trials: what have we learned from magnetic resonance imaging. J Magn Reson Imaging 2025 Feb;61(2):579-594. [doi: 10.1002/jmri.29462] [Medline: 39031716]
11. Kuo RYL, Harrison C, Curran TA, et al. Artificial intelligence in fracture detection: a systematic review and meta-analysis. Radiology 2022 Jul;304(1):50-62. [doi: 10.1148/radiol.211785] [Medline: 35348381]
12. Wang J, Xue L, Jiang J, et al. Diagnostic performance of artificial intelligence-assisted PET imaging for Parkinson's disease: a systematic review and meta-analysis. NPJ Digit Med 2024 Jan 22;7(1):17. [doi: 10.1038/s41746-024-01012-z] [Medline: 38253738]
13. Cheng Y, Malekar M, He Y, et al. High-throughput phenotyping of the symptoms of Alzheimer disease and related dementias using large language models: cross-sectional study. JMIR AI 2025 Jun 3;4:e66926. [doi: 10.2196/66926] [Medline: 40460418]

XSL•FO
RenderX

14. Bosco C, Shojaei F, Theisz AA, et al. Testing 3 modalities (voice assistant, chatbot, and mobile app) to assist older African American and Black adults in seeking information on Alzheimer disease and related dementias: wizard of Oz usability study. JMIR Form Res 2024 Dec 9;8:e60650. [doi: 10.2196/60650] [Medline: 39653372]
15. Rudroff T, Rainio O, Klén R. AI for the prediction of early stages of Alzheimer's disease from neuroimaging biomarkers—a narrative review of a growing field. Neurol Sci 2024 Nov;45(11):5117-5127. [doi: 10.1007/s10072-024-07649-8] [Medline: 38866971]
16. Jackson D, Turner R. Power analysis for random-effects meta-analysis. Res Synth Methods 2017 Sep;8(3):290-302. [doi: 10.1002/jrsm.1240] [Medline: 28378395]
17. McInnes MDF, Moher D, Thombs BD, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. JAMA 2018 Jan 23;319(4):388-396. [doi: 10.1001/jama.2017.19163] [Medline: 29362800]
18. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024 Apr:q902. [doi: 10.1136/bmj.q902]
19. Singh S, Chang SM, Matchar DB, Bass EB. Chapter 7: grading a body of evidence on diagnostic tests. J Gen Intern Med 2012 Jun;27 Suppl 1(Suppl 1):S47-S55. [doi: 10.1007/s11606-012-2021-9] [Medline: 22648675]
20. Ruppar T. Meta-analysis: how to quantify and explain heterogeneity? Eur J Cardiovasc Nurs 2020 Oct;19(7):646-652. [doi: 10.1177/1474515120944014] [Medline: 32757621]
21. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Stat Med 2001 Oct 15;20(19):2865-2884. [doi: 10.1002/sim.942] [Medline: 11568945]
22. Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 2005 Oct;58(10):982-990. [doi: 10.1016/j.jclinepi.2005.02.022] [Medline: 16168343]
23. Zhang D, Wang Y, Zhou L, Yuan H, Shen D. Multimodal classification of Alzheimer's disease and mild cognitive impairment. Neuroimage 2011 Apr 1;55(3):856-867. [doi: 10.1016/j.neuroimage.2011.01.008] [Medline: 21236349]
24. Yun HJ, Kwak K, Lee JM, Alzheimer's Disease Neuroimaging Initiative. Multimodal discrimination of Alzheimer's disease based on regional cortical atrophy and hypometabolism. PLoS ONE 2015;10(6):e0129250. [doi: 10.1371/journal.pone.0129250] [Medline: 26061669]
25. Westman E, Muehlboeck JS, Simmons A. Combining MRI and CSF measures for classification of Alzheimer's disease and prediction of mild cognitive impairment conversion. Neuroimage 2012 Aug 1;62(1):229-238. [doi: 10.1016/j.neuroimage.2012.04.056] [Medline: 22580170]
26. Vemuri P, Gunter JL, Senjem ML, et al. Alzheimer's disease diagnosis in individual subjects using structural MR images: validation studies. Neuroimage 2008 Feb 1;39(3):1186-1197. [doi: 10.1016/j.neuroimage.2007.09.073] [Medline: 18054253]
27. Suk HI, Lee SW, Shen D. Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. Neuroimage 2014 Nov 1;101:569-582. [doi: 10.1016/j.neuroimage.2014.06.077] [Medline: 25042445]
28. Sayeed A, Petrou M, Spyrou N, Kadyrov A, Spinks T. Diagnostic features of Alzheimer's disease extracted from PET sinograms. Phys Med Biol 2002 Jan 7;47(1):137-148. [doi: 10.1088/0031-9155/47/1/310] [Medline: 11814222]
29. Pan X, Adel M, Fossati C, et al. Multiscale spatial gradient features for 18F-FDG PET image-guided diagnosis of Alzheimer's disease. Comput Methods Programs Biomed 2019 Oct;180:105027. [doi: 10.1016/j.cmpb.2019.105027] [Medline: 31430595]
30. Padilla P, López M, Górriz JM, et al. NMF-SVM based CAD tool applied to functional brain images for the diagnosis of Alzheimer's disease. IEEE Trans Med Imaging 2012 Feb;31(2):207-216. [doi: 10.1109/TMI.2011.2167628] [Medline: 21914569]
31. Ni YC, Tseng FP, Pai MC, et al. Detection of Alzheimer's disease using ECD SPECT images by transfer learning from FDG PET. Ann Nucl Med 2021 Aug;35(8):889-899. [doi: 10.1007/s12149-021-01626-3] [Medline: 34076857]
32. Magnin B, Mesrob L, Kinkingnéhun S, et al. Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI. Neuroradiology 2009 Feb;51(2):73-83. [doi: 10.1007/s00234-008-0463-x] [Medline: 18846369]
33. Lu DH, Popuri K, Ding GW, Balachandar R, Beg MF. Multiscale deep neural network based analysis of FDG-PET images for the early diagnosis of Alzheimer's disease. Med Image Anal 2018 May;46:26-34. [doi: 10.1016/j.media.2018.02.002]
34. Liu MH, Zhang DQ, Shen DG. Ensemble sparse classification of Alzheimer's disease. Neuroimage 2012 Apr;60(2):1106-1116. [doi: 10.1016/j.neuroimage.2012.01.055]
35. Liu M, Cheng D, Yan W, Alzheimer's Disease Neuroimaging Initiative. Classification of Alzheimer's disease by combination of convolutional and recurrent neural networks using FDG-PET images. Front Neuroinform 2018;12:35. [doi: 10.3389/fninf.2018.00035] [Medline: 29970996]
36. Li R, Perneczky R, Yakushev I, et al. Gaussian mixture models and model selection for [18F] fluorodeoxyglucose positron emission tomography classification in Alzheimer's disease. PLoS ONE 2015 Apr;10(4):e0122731. [doi: 10.1371/journal.pone.0122731]
37. Lerch JP, Pruessner J, Zijdenbos AP, et al. Automated cortical thickness measurements from MRI can accurately separate Alzheimer's patients from normal elderly controls. Neurobiol Aging 2008 Jan;29(1):23-30. [doi: 10.1016/j.neurobiolaging.2006.09.013] [Medline: 17097767]

XSL•FO
RenderX

38. Kim HW, Lee HE, Oh K, Lee S, Yun M, Yoo SK. Multi-slice representational learning of convolutional neural network for Alzheimer's disease classification using positron emission tomography. Biomed Eng Online 2020 Sep 7;19(1):70. [doi: 10.1186/s12938-020-00813-z] [Medline: 32894137]
39. Kim HW, Lee HE, Lee S, Oh KT, Yun M, Yoo SK. Slice-selective learning for Alzheimer's disease classification using a generative adversarial network: a feasibility study of external validation. Eur J Nucl Med Mol Imaging 2020 Aug;47(9):2197-2206. [doi: 10.1007/s00259-019-04676-y] [Medline: 31980910]
40. Katako A, Shelton P, Goertzen AL, et al. Machine learning identified an Alzheimer's disease-related FDG-PET pattern which is also expressed in Lewy body dementia and Parkinson's disease dementia. Sci Rep 2018 Sep 5;8(1):13236. [doi: 10.1038/s41598-018-31653-6] [Medline: 30185806]
41. Ismail WN, P. P. FR, Ali MAS. A meta-heuristic multi-objective optimization method for Alzheimer's disease detection based on multi-modal data. Mathematics 2023 Feb;11(4):957. [doi: 10.3390/math11040957]
42. Illán IA, Górriz JM, Ramírez J, et al. 18F-FDG PET imaging analysis for computer aided Alzheimer's diagnosis. Inf Sci (Ny) 2011 Feb;181(4):903-916. [doi: 10.1016/j.ins.2010.10.027]
43. Hinrichs C, Singh V, Mukherjee L, et al. Spatially augmented LPboosting for AD classification with evaluations on the ADNI dataset. Neuroimage 2009 Oct 15;48(1):138-149. [doi: 10.1016/j.neuroimage.2009.05.056] [Medline: 19481161]
44. Gray KR, Wolz R, Heckemann RA, et al. Multi-region analysis of longitudinal FDG-PET for the classification of Alzheimer's disease. Neuroimage 2012 Mar;60(1):221-229. [doi: 10.1016/j.neuroimage.2011.12.071] [Medline: 22236449]
45. Gray KR, Aljabar P, Heckemann RA, Hammers A, Rueckert D. Random forest-based similarity measures for multi-modal classification of Alzheimer's disease. Neuroimage 2013 Jan 15;65:167-175. [doi: 10.1016/j.neuroimage.2012.09.065] [Medline: 23041336]
46. Feng C, Elazab A, Yang P, et al. Deep learning framework for Alzheimer's disease diagnosis via 3D-CNN and FSBi-LSTM. IEEE Access 2019;7:63605-63618. [doi: 10.1109/ACCESS.2019.2913847]
47. Cuingnet R, Gerardin E, Tessieras J, et al. Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. Neuroimage 2011 May 15;56(2):766-781. [doi: 10.1016/j.neuroimage.2010.06.013] [Medline: 20542124]
48. Chen L, Qiao HZ, Zhu F. Alzheimer's disease diagnosis with brain structural MRI using multiview-slice attention and 3D convolution neural network. Front Aging Neurosci 2022;14:871706. [doi: 10.3389/fnagi.2022.871706] [Medline: 35557839]
49. Song J, Zheng J, Li P, Lu X, Zhu G, Shen P. An effective multimodal image fusion method using MRI and PET for Alzheimer's disease diagnosis. Front Digit Health 2021;3:637386. [doi: 10.3389/fdgth.2021.637386] [Medline: 34713109]
50. Li Y, Jiang J, Lu J, Jiang J, Zhang H, Zuo C. Radiomics: a novel feature extraction method for brain neuron degeneration disease using 18F-FDG PET imaging and its implementation for Alzheimer's disease and mild cognitive impairment. Ther Adv Neurol Disord 2019;12:1756286419838682. [doi: 10.1177/1756286419838682] [Medline: 30956687]
51. A A, M P, Hamdi M, Bourouis S, Rastislav K, Mohmed F. Evaluation of neuro images for the diagnosis of Alzheimer's disease using deep learning neural network. Front Public Health 2022;10:834032. [doi: 10.3389/fpubh.2022.834032] [Medline: 35198526]
52. Toussaint PJ, Perlbarg V, Bellec P, et al. Resting state FDG-PET functional connectivity as an early biomarker of Alzheimer's disease using conjoint univariate and independent component analyses. Neuroimage 2012 Nov 1;63(2):936-946. [doi: 10.1016/j.neuroimage.2012.03.091] [Medline: 22510256]
53. Tong T, Wolz R, Gao Q, et al. Multiple instance learning for classification of dementia in brain MRI. Med Image Anal 2014 Jul;18(5):808-818. [doi: 10.1016/j.media.2014.04.006] [Medline: 24858570]
54. Min R, Wu G, Cheng J, Wang Q, Shen D. Multi-atlas based representations for Alzheimer's disease diagnosis. Hum Brain Mapp 2014 Oct;35(10):5052-5070. [doi: 10.1002/hbm.22531] [Medline: 24753060]
55. Jin D, Zhou B, Han Y, et al. Generalizable, reproducible, and neuroscientifically interpretable imaging biomarkers for Alzheimer's disease. Adv Sci (Weinh) 2020 Jul;7(14):2000675. [doi: 10.1002/advs.202000675] [Medline: 32714766]
56. Cho Y, Seong JK, Jeong Y, Shin SY. Individual subject classification for Alzheimer's disease based on incremental learning using a spatial frequency representation of cortical thickness data. Neuroimage 2012 Feb 1;59(3):2217-2230. [doi: 10.1016/j.neuroimage.2011.09.085] [Medline: 22008371]
57. Chincarini A, Bosco P, Calvini P, et al. Local MRI analysis approach in the diagnosis of early and prodromal Alzheimer's disease. Neuroimage 2011 Sep 15;58(2):469-480. [doi: 10.1016/j.neuroimage.2011.05.083] [Medline: 21718788]
58. Beheshti I, Demirel H, Matsuda H. Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm. Comput Biol Med 2017 Apr 1;83(109-19):109-119. [doi: 10.1016/j.compbiomed.2017.02.011] [Medline: 28260614]
59. Anandh KR, Sujatha CM, Ramakrishnan S. A method to differentiate mild cognitive impairment and Alzheimer in MR images using eigen value descriptors. J Med Syst 2016 Jan;40(1):26547845. [doi: 10.1007/s10916-015-0396-y] [Medline: 26547845]
60. Amoroso N, La Rocca M, Bruno S, et al. Multiplex networks for early diagnosis of Alzheimer's disease. Front Aging Neurosci 2018;10:365. [doi: 10.3389/fnagi.2018.00365] [Medline: 30487745]

XSL•FO
RenderX

61. Yamane T, Ikari Y, Nishio T, et al. Visual-statistical interpretation of (18)F-FDG-PET images for characteristic Alzheimer patterns in a multicenter study: inter-rater concordance and relationship to automated quantitative evaluation. AJNR Am J Neuroradiol 2014 Feb;35(2):244-249. [doi: 10.3174/ajnr.A3665] [Medline: 23907243]
62. Harper L, Fumagalli GG, Barkhof F, et al. MRI visual rating scales in the diagnosis of dementia: evaluation in 184 post-mortem confirmed cases. Brain (Bacau) 2016 Apr;139(Pt 4):1211-1225. [doi: 10.1093/brain/aww005] [Medline: 26936938]
63. Loddo A, Buttau S, Di Ruberto C. Deep learning based pipelines for Alzheimer's disease diagnosis: a comparative study and a novel deep-ensemble method. Comput Biol Med 2022 Feb;141:105032. [doi: 10.1016/j.compbiomed.2021.105032] [Medline: 34838263]
64. Pawar U, O'Shea D, Rea S, O'Reilly R. Explainable AI in healthcare. Presented at: 2020 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA); Jun 15-19, 2020; Dublin, Ireland. [doi: 10.1109/CyberSA49311.2020.9139655]
65. Birkenbihl C, Emon MA, Vrooman H, et al. Differences in cohort study data affect external validation of artificial intelligence models for predictive diagnostics of dementia - lessons for translation into clinical practice. EPMA J 2020 Sep;11(3):367-376. [doi: 10.1007/s13167-020-00216-z] [Medline: 32843907]
66. Barredo Arrieta A, Díaz-Rodríguez N, Del Ser J, et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion 2020 Jun;58:82-115. [doi: 10.1016/j.inffus.2019.12.012]
67. Mayerhoefer ME, Prosch H, Beer L, et al. PET/MRI versus PET/CT in oncology: a prospective single-center study of 330 examinations focusing on implications for patient management and cost considerations. Eur J Nucl Med Mol Imaging 2020 Jan;47(1):51-60. [doi: 10.1007/s00259-019-04452-y] [Medline: 31410538]
68. Goldberg CB, Adams L, Blumenthal D, et al. To do no harm - and the most good - with AI in health care. Nat Med 2024 Mar;30(3):623-627. [doi: 10.1038/s41591-024-02853-7] [Medline: 38388841]
69. Frizzell TO, Glashutter M, Liu CC, et al. Artificial intelligence in brain MRI analysis of Alzheimer's disease over the past 12 years: A systematic review. Ageing Res Rev 2022 May;77:101614. [doi: 10.1016/j.arr.2022.101614] [Medline: 35358720]
70. Borchert RJ, Azevedo T, Badhwar A, et al. Artificial intelligence for diagnostic and prognostic neuroimaging in dementia: A systematic review. Alzheimers Dement 2023 Dec;19(12):5885-5904. [doi: 10.1002/alz.13412] [Medline: 37563912]
71. Sun Y, Chen Y, Dong L, et al. Diagnostic performance of deep learning-assisted [18F]FDG PET imaging for Alzheimer's disease: a systematic review and meta-analysis. Eur J Nucl Med Mol Imaging 2025 Aug;52(10):3600-3612. [doi: 10.1007/s00259-025-07228-9] [Medline: 40159544]

Abbreviations

^{18}FDG-PET: fluorine-18 fluorodeoxyglucose positron emission tomography
AD: Alzheimer disease
AI: artificial intelligence
DL: deep learning
MeSH: Medical Subject Headings
ML: machine learning
PICOS: Population, Intervention, Comparison, Outcome, Study Design
PRISMA-DTA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses of Diagnostic Test Accuracy
sMRI: structural magnetic resonance imaging
SROC-AUC: summary receiver operating characteristic curve area
TRIPOD-AI: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis–Artificial Intelligence

Edited by M O'Connell; submitted 05.05.25; peer-reviewed by RT Potla, S Mohanadas, SA Mirshahvalad; revised version received 03.07.25; accepted 20.07.25; published 08.10.25.

Please cite as:
Wang B, Zhao T, Ma R, Huo X, Xiong X, Wu M, Wang Y, Liu L, Zhuang Z, Wang B, Shou J
Comparative Diagnostic Accuracy of AI-Assisted Fluorine-18 Fluorodeoxyglucose Positron Emission Tomography Versus Structural Magnetic Resonance Imaging in Alzheimer Disease: Systematic Review and Meta-Analysis
JMIR Aging 2025;8:e76981
URL: https://aging.jmir.org/2025/1/e76981
doi:10.2196/76981

XSL•FO
RenderX

© Bingbing Wang, Tailiang Zhao, Rongrong Ma, Xiaochuan Huo, Xiaoxiao Xiong, Minjie Wu, Yuran Wang, Liu Liu, Zhijiang Zhuang, Bin Wang, Jixin Shou. Originally published in JMIR Aging (https://aging.jmir.org), 8.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Toward Data-Informed Care in Long-Term Care: Qualitative Analysis

Suleyman Bouchmal[1], MSc; Katya YJ Sion[1], PhD; Jan PH Hamers[1,2], PhD, Prof Dr; Sil Aarts[1], PhD

[1]Department of Health Services Research, Limburg Living Lab in Ageing and Long-Term Care, Maastricht University, Duboisdomein 30, Maastricht, The Netherlands

[2]MeanderGroep Zuid-Limburg, Landgraaf, The Netherlands

Corresponding Author:
Suleyman Bouchmal, MSc
Department of Health Services Research, Limburg Living Lab in Ageing and Long-Term Care, Maastricht University, Duboisdomein 30, Maastricht, The Netherlands

Abstract

Background: In long-term care (LTC) for older adults, data on client, employee, and organization levels are collected in various ways, covering quality of care, life, and work. There is, however, a lack of understanding of how to introduce data-informed care in LTC and thus create value from data.

Objective: This study aims to investigate the experiences and perceptions of various stakeholders in LTC regarding data and data-informed care.

Methods: A qualitative study using the World Café cocreation technique was conducted with a diverse group of LTC stakeholders. Four questions were addressed: (1) What thoughts do you have when you hear the term “data” in relation to LTC? (2) What purposes do data have (in the future) in LTC? (3) What knowledge and skills are needed to enable data-informed care? (4) How can data contribute to and improve multidisciplinary learning? Stakeholders’ notes and the plenary summary were analyzed using conventional content analysis.

Results: Stakeholders included nurses, members of client councils, data specialists, researchers, and managers (N=20; mean age 50, SD 13 years). Five themes were identified: (1) despite uncertainty, the benefits of using data outweigh the associated risks; (2) the lack of accessibility and uniformity hinders integrating data-informed care; (3) human resources and finance departments pioneer data usage; however, potential lies in clinical decision-making; (4) data-informed care demands individual, collective, and organizational prerequisites; and (5) multidisciplinary collaboration enriches collective knowledge regarding data.

Conclusions: Introducing data-informed care requires enhancing data literacy of health care professionals, establishing clear communication about the role of data within the organization, and introducing new job positions, such as data scientists. Data-informed care was considered a multidisciplinary approach in which data have a supportive role to enhance collective understanding and are considered crucial for improving quality of care.

(JMIR Aging 2025;8:e69423) doi:10.2196/69423

KEYWORDS

long-term care; data; cocreation; data-informed; quality of care

Introduction

Data are crucial for making the right decisions in health care systems [1]. Information systems and electronic health records (EHRs) improve data management, offering potential benefits, such as enhancing the quality of care and increasing the knowledge levels of health care professionals [2,3]. For example, in clinical decisions, prompt data entry allows for timely access to information and better decision-making [4,5]. Additional examples include the facilitation of information-sharing, which enables collaborative approaches to decision-making [6,7].

In long-term care (LTC) for older adults, vast amounts of data are consistently collected on client, employee, and organizational levels [8,9]. Data can be collected in various ways, such as conversations, questionnaires, and the use of technologies like sensors or wearables. These data can, for example, provide new insights regarding quality of care, support the early detection of diseases, and enhance decision-making [10,11]. Despite the potential of data, integrating data into daily care practice is challenging due to uncertainties about interoperability regarding different data sources (ie, EHRs) and how to use this data for clinical decision-making [12-14]. In addition, health care professionals are often ambivalent about data and the use of information technology [15,16].

Clients in LTC often have complex and multifaceted needs that require, besides clinical data, additional considerations, such

XSL•FO
RenderX

as personal preferences [17,18]. Consequently, challenges arise in decision-making processes due to the context, available data, and characteristics of the decision-makers, such as experience and skills [19]. Previous research has argued that delivering care to individuals with complex care needs should be of high value, team-based, and data-informed [20,21]. Therefore, the informative role of data to support several processes, including decision-making and care delivery, is gaining increased attention in LTC [22-24]. While data-driven methods prioritize decisions based solely on information, data-informed care balances quantitative data with valuable insights from previous experiences and qualitative input from clients, relatives, and health care professionals [25,26]. As data-informed care has not yet been implemented in LTC, however, this may indicate that LTC clients are not receiving the most appropriate care [27].

To date, there is a lack of understanding of how data-informed care can be introduced in LTC settings. For example, the literature regarding the implications of data-informed care for different stakeholders (eg, caregivers, managers, nurses, and data specialists) is scarce [28,29]. Moreover, it remains uncertain which specific requirements are necessary and how LTC organizations and their stakeholders view the role of data in LTC [30]. Therefore, this study aims to investigate the experiences and perceptions of various stakeholders in LTC settings regarding data and data-informed care practices. The findings in this study provide insights into the status quo of data usage and highlight topics to transform toward data-informed care practices in LTC.

Methods

Design

This qualitative study used the cocreation method "The World Café," which allows an open conversation with a variety of stakeholders [31]. This creative method establishes informal discussions in small groups about central questions [32]. The chosen methodology aligns well with the exploratory aims of this study, making it a suitable approach for gaining initial insights [33]. In 3 constructive rounds, collaborative knowledge was created with the stakeholders and evaluated in a plenary session in which the findings were summarized and presented to the group for reflection and feedback [31].

Participants

This study was embedded within the interdisciplinary core group "Data Science of the Limburg Living Lab in Aging and Long-Term Care" (in Dutch: Academische Werkplaats Ouderenzorg Limburg [AWO-L]) [34]. This formal multidisciplinary network consists of a collaboration among Maastricht University, 9 large LTC organizations, and 3 educational institutions (secondary, vocational, bachelor's, or master's level). Together, the Limburg Living Lab delivers care to more than 50,000 clients daily with the support of 27,000 employees. The mission of the network is to conduct scientific research to contribute to the improvement of the quality of life for older adults and their relatives, quality of care, and quality of work for individuals active in LTC [34].

For this study, all organizations within the AWO-L Living Lab were approached and invited to participate based on purposive sampling. This study aimed to include a diverse group of stakeholders. In total, 3 LTC organizations and 2 educational institutes were registered. The principal investigator provided an informative presentation on location. Interested individuals were invited via an email in which the objectives of the study and practical information about the World Café sessions were provided, including the date, location, and time. For those who were interested in participating, a link was included in the email to register. Stakeholders were not required to have prior knowledge of data or data-informed care but were expected to be familiar with the LTC context—either through direct care provision for older adults or through involvement in research or education related to this setting. At the start of the World Café session, a presentation was delivered to introduce the concept of data-informed care. This allowed stakeholders to develop a common interpretation before engaging in group discussions.

Data Collection

Data collection took place in January 2024 using the seven design principles for conducting a World Café session: (1) set the context, (2) create a hospitable environment, (3) explore questions that matter, (4) encourage everyone's contribution, (5) connect diverse perspectives, (6) listen together for patterns and insights, and (7) share collective discoveries [35].

Starting the cocreation session, the stakeholders were welcomed with an introductory presentation to review the objectives of the World Café as well as an introduction into the concept of data-informed care. Time was allocated for questions from the stakeholders regarding the information provided by email or during the presentation. Thereafter, stakeholders received an informed consent form that detailed important aspects of the session. This included the type of information that was collected (ie, photos during the session, audio recordings of the plenary summary, and written answers given by the stakeholders), and how this information would be analyzed, processed, and stored. After signing the informed consent, the stakeholders were given a small toolkit containing Post-its, a notepad, markers, and pencils to provide written answers during the World Café session.

The World Café session was hosted in a public venue that simulated a restaurant's layout, with 4 round tables, as well as snacks and beverages. The World Café method is described as being conducted in a café-style layout of tables, which encourages enthusiasm, contribution, and participation [31,36]. Hence, the World Café facilitates an open conversation with a variety of stakeholders on several topics. The stakeholders were placed into 4 heterogeneous groups based on their profession and organization. Each group gathered at 1 table, and the group's composition remained unchanged during the cocreation session. In total, 4 tables (and thus 4 groups) were established [35]. All tables were hosted by a "moderator": a member of the research team with experience in qualitative data collection and in-depth knowledge about the research topic.

For each table, 1 central question and prompts were formulated during 3 meetings prior to the World Café with the research

XSL•FO
RenderX

team, based on their relevance to the aim of this study (Table 1) [37]. All moderators participated in a preparatory meeting prior to the World Café to collaboratively discuss the questions and prompts aligned with the study's aim. Moderators also received additional information on the session's structure, along with a written manual detailing the prompts assigned to each table. The prompts, which were subquestions regarding the central questions, were used to guide and enhance dialog to acquire a collective understanding.

Table . Central question per table.

Central questions	Examples prompts
T1 - What thoughts do you have when you hear the word "data" in relation to LTC[a]?	• Why do we collect data in LTC? • What information does a health care professional need to provide care? • What objection do you have to using data in the care process?
T2 - What purposes will data have (in the future) in LTC?	• Could data support improve the quality of care? • Why should a health care professional consider data usage in their work? • What role has data regarding decision-making?
T3 - What knowledge and skills are needed to enable data-informed care?	• How can data enhance care practices? • What competencies are required regarding data-informed care? • What systems or other resources does the health care professional need here?
T4 - How can data contribute to and improve multidisciplinary learning?'	• What is multidisciplinary working, and what role has learning and improving played in this? • How can multidisciplinary learning support care practices? • What challenges should be considered when integrating data into multidisciplinary learning?

[a]LTC: long-term care.

Each group discussed the corresponding central question for 30 minutes. In the first round, the stakeholders introduced themselves within their group. As the composition of the groups remained unchanged, introducing oneself once was sufficient [35]. After the moderators stated the central question, stakeholders started brainstorming individually and wrote their ideas down on Post-its before starting the conversation, ensuring that all stakeholders' perspectives were included [38]. Afterward, the dialog between the stakeholders started, and new ideas and answers were formulated on Post-its. The collaborative nature of the World Café methodology aligns with key principles of participatory research by fostering inclusive discussion and meaningful stakeholder engagement. In particular, it facilitates coconstruction, a higher level of participation as defined in adaptations of the participation ladder, where stakeholders actively contribute to shaping insights rather than functioning merely as data sources [39].

In the second and third rounds, the groups shifted to new tables. The moderators started by summarizing the insights of the previous rounds to the new group via the Post-its from the latter group. This allowed stakeholders to link with the previous group and build upon their work to discover new insights regarding the central question [40]. The moderators then stated the central question and let the stakeholders brainstorm individually again before starting the conversation. By expanding on each other's thoughts, the World Café sessions allowed "cross-pollination," meaning that sharing and interchanging information and thoughts enriches the group's capabilities and connects diverse perspectives, resulting in collective knowledge [31]. To conclude the 3 rounds, the 4 moderators provided a plenary summary of the overall findings from their tables. Additional comments or feedback shared during the plenary reporting of each table's insights were incorporated into the results. All stakeholders present during the world-café session were given the opportunity to share their opinion or provide comments based on the plenary reporting, in case something was missing or misinterpreted.

Data Analysis

The informed consent acquired additional information from each stakeholder, namely their age, gender, the name of their organization, their job title, and their years of work experience in their current profession, as described in the "Results" section. Descriptive analysis was performed to provide insights into the demographics of the stakeholders (eg, mean age, SDs, and range).

The notes on the Post-its and the written text on the sheet from the tables, as well as the audio record of the plenary summary, were transcribed verbatim [41]. The transcripts (ie, notes per table and the plenary summary) were subsequently uploaded to the qualitative analysis software MAXQDA (VERBI Software GmbH [42]). The first author familiarized himself with the data by reading all transcripts multiple times, which facilitated an in-depth understanding of the data [43]. Conventional content analysis was conducted to describe the experiences and perceptions of stakeholders' daily work with data in LTC [44,45]. Conventional content analysis avoids predetermined categories and inductively processes the data. Therefore, categories or themes, along with their names, emerge from the data when using a systematic approach [46,47]. This study

XSL•FO
RenderX

conducted the conventional content analysis by following up with systematic coding [46].

The first author conducted the coding process by performing a line-by-line analysis that identified meaningful units of information [48]. These text segments were "open coded" with overarching labels describing the condensed meaning. The descriptive labels of the open-coded segments were summarized by grouping similar codes to identify clusters of codes. By integrating the codes into clusters, overarching categories were formulated [49]. Establishing overarching categories by specifying subcategories is often referred to as axial coding [50]. This process was evaluated by all coauthors. The content and interpretation of the themes were discussed with the whole research team [51]. The translation process was checked for content and context independently within the research team, based on the central questions, used prompts, original notes written by the stakeholders, and translations conducted by the first author.

Ethical Considerations

The study protocol was approved by the ethics committee of Maastricht University (approval number: Faculty of Health, Medicine, and Life Sciences - FHML-REC/2024/055). All information about the aim of this study was communicated to the stakeholders in advance by email. Informed consent was provided, and participation was strictly voluntary. Withdrawal from the study was allowed at any moment for any or no reason. To guarantee the anonymity of the stakeholders, no names or organizations were documented.

Results

Overview

The descriptive characteristics of the stakeholders are reported in Table 2. In total, 20 stakeholders participated with a mean age of 50 (SD 13) years. The stakeholders had, on average, 16 (SD 9) years of working experience, and 55% (11/20) were female. This study included 7 different professions from LTC organizations and educational institutions.

Table . Descriptive characteristics of stakeholders (N=20).

Characteristics	Value
Age (years), mean (SD; range)	50 (13; 33-79)
Work experience (years), mean (SD; range)	16 (9; 3-40)
Sex (female), n (%)	11 (55)
Profession, n (%)	
Health care professionals[a]	6 (30)
Business Intelligence specialists[b]	3 (15)
Members client council	3 (15)
Education[c]	4 (20)
Managers	4 (20)

[a] health care professionals included 4 nurses, 1 specialist in involuntary treatment, and 1 specialist in client well-being.

[b]Business Intelligence specialists included 2 Business Intelligence architects and 1 specialist in data storage and exchange.

[c]Education included 2 teachers and 2 researchers.

Five themes were identified during the World Café session [52]: (1) despite uncertainty, the benefits of using data outweigh the associated risks; (2) the lack of accessibility and uniformity hinders integrating data-informed care; (3) human resources and finance departments pioneer data usage, yet potential lies in clinical decision-making; (4) data-informed care demands individual, collective, and organizational prerequisites; and (5) multidisciplinary collaboration enriches collective knowledge regarding data. The 5 themes, including their underlying axial codes and some examples of direct quotations in stakeholders' notes, are reported in Table 3.

XSL•FO
RenderX

Table . Overview of themes, axial codes, and stakeholders' notes.

Themes and axial codes	Stakeholders' notes (direct quotations)
Despite uncertainty, the benefits of using data outweigh the associated risks	
Knowledge about data	Data is just a word–data is not used (yet)–there is a lot of data in LTC.
Negative feelings	Data collection should not become the aim–fear of the unknown–risk that data steers a professional too much.
Current applications	AI in practice–eHealth (data that is not necessarily added or used by humans)–going from retrospective to real-time insights, to prospective improvements.
Ethical dilemmas	Uncertainty about reliability of data-Ethical dilemmas about data usage.
Data's potential for care	Data usage of quality of care–data supports tailored care to the client–identifying group selections.
The lack of accessibility and uniformity hinders integrating data-informed care	
Indicating variety of data	Demographical information – Clinical Measurements – kinds of data, including audio, video, qualitative and quantitative.
Uniformity in systems	Standardizing definitions in the systems—System often requires jargon, should switch to layman's' language—database structure is crucial for enhancing uniformity.
Simplification	Translating data to present understandable information for clients, informal caregivers and public–professionals will be supported if data becomes more comprehensive and easy to use–simplifying data improves multidisciplinary learning and improving.
Interoperability	Sharing data and associated solutions to increase data usage–LTC can continuously improve by connecting data between systems–Link external sources and strive for open data.
Electronic health record	Double checks in EHR–Data could support medication registration–Compliance cannot be tracked with data yet.
Information gain	Systems should be unambiguously and easy to use–Making data both findable and usable–With LTC, agreements are needed for accessibility and governance of data.
Accessibility	Creating overview of aging population (in region)–Matching eHealth and specific client(characteristics)–Insight into data of future clients.
Centralizing storage	Working towards one platform and format for different disciplines-Centralizing information–Reducing workload or pressure as data entry and retrieval are in one place.
HR and finance departments pioneer data usage, yet the potential lies in clinical decision-making	
Organizational opportunities	Determination of long-term vision, supported by financial numbers–Organize care differently (more care with less resources–Data can be used positively or negatively to finance care.
Data collection and analysis techniques	Collecting information on different levels, to gain insights what is important and for who–classifying systematic problem analysis–Identifying the different instruments for data collection will result in less administration costs.
Predictions	Data can support in findings patterns and trends–Estimating lifestyle and preferences of clients–Predicting care intensity.
Data-informed care demands individual, collective, and organizational prerequisites	
Education, training, and support	Education to understand the goal of data and information–Support needed for individuals with less digital literacy–Education or train employees systematically.
Teamwork and collective skills	Communicational skills within a team–Dare to innovate-mentality needed in a team–Every team should include an attention raiser or data specialist.
Individual contribution toward data-informed working	Feeling the urge to do own contribution independently or together with informal caregiver)–Taking control and responsibility–Understanding that data should support and not influence.

XSL•FO
RenderX

Themes and axial codes	Stakeholders' notes (direct quotations)
Requirements for learning and improving	Multidisciplinary collaboration with different layers (care, cure, management, IT)–Organizational or team objective should be formulated and communicated more clearly with everyone–Define responsibilities and roles.
Multidisciplinary collaboration enriches collective knowledge regarding data	
Central role for data	Data supports individuals in multidisciplinary approaches with shared and offering new ideas–Adhering to Evidence Based Practice working principles–Enhance dialogue and understanding between different professions.
Collaborative and shared decision-making	Integrating multiple sources and professions allows for better consideration and substantiation–Appropriately implementing SDM result in more, better dialogue with clients–making decisions as a team.
Striving for overarching goals	Teams should mitigate working with limited perspectives from other professions–by working multidisciplinary, the shared objective should be accepted by all–Crucial advantage is "multi-perspectivity" on the one subject.
Collective knowledge	Leaning-by-doing enhances individual digital literacy levels–Gaining knowledge about digital systems, finding information and using that information – Collective understanding acquired by individual contributions.

Despite Uncertainty, the Benefits of Using Data Outweigh the Associated Risks

When asking stakeholders what data means to them, the majority reported recognizing the presence of "data" in their organizations. They mentioned, however, that it is hard to describe data specifically. In addition, they expressed that data are not part of their daily tasks. As stakeholders discussed their role in collecting data, such as demographic information about their clients, they also indicated being aware of different types of data present in LTC (eg, images, audio, and quantitative).

> *Well, initially it was about numbers. Then you have to think of other things that we collect, for example data on wounds or falls or pain and stories in file - stories that employees write down in the files about clients.*

> *Information about personnel files but also sensor technology were added – this data was not collected by humans, but collected in a different way, such as video, audio.*

As the stakeholders considered data usage beneficial, they proposed that LTC should start shifting on the organizational level to increase the use of data. Stakeholders mentioned that such transitions should be done not only to enhance data usage but also to allow the evaluation of how decision-making is integrated in the organization. They emphasized the benefit of data use for the quality of care and discussed that LTC organizations should go from retrospective improvements to real-time insights and even to prospective improvements in the future. Several solutions were mentioned by the stakeholders to support such data transitions, for example, artificial intelligence. However, stakeholders noted that the use of data in practice entails both advantages and disadvantages due to underlying uncertainties, including ethical considerations and client privacy concerns.

> *First, we need to make sure that shared decision-making really happens. If that is not the case, how do you want to add the data part to introduce data-informed decision-making? Than we should think about the process, before adding data into it. Especially about how we want to offer it to that group.* [Client, relatives, and healthcare provider]

> *Implementation of data into solutions that can mimic ChatGPT in which insights from different disciplines can be linked together. Artificial intelligence may support care processes but could also bring more improvements over time. Everybody in the group [client, relatives, and healthcare provider] could experience then an added value.*

While summarizing the variety of data sources, stakeholders mentioned the negative aspects of using data. For example, a few stakeholders mentioned applications that could automatically collect (real-time) data (eg, sensors) but discussed their concern about the reliability of the data. Moreover, stakeholders indicated that using data for data-informed care includes the risk that health care professionals will be influenced too much by the data without assessing or recognizing their quality.

> *Again, not too critical, but critical enough to evaluate the data to see what do the data say, is it correct and does that match my gut feeling? Is that also how I see the client?*

The majority of the stakeholders indicated that using data could support gaining more insights into their clients. They wished for a supportive role for data to oversee their clients, including underlying data that might help them provide care. In addition, care processes on both an individual level (delivery-receiving care) and an organizational level (arranging and organizing care) can experience the benefits of using data, according to the stakeholders. They mentioned that data in LTC help deliver more personalized care and improve quality of care in the organization.

> *Because we want to improve the quality of care, we want to be able to get care tailored to the client by*

XSL•FO
RenderX

providing personalized care. Therefore, we will get a better picture of: OK, we now have this client in a group or subgroup with related data, for example, in a nursing home.

The Lack of Accessibility and Uniformity Hinders Integrating Data-Informed Care

A few stakeholders indicated that work processes were established for data entry and retrieval in their systems. Most stakeholders, however, mentioned that the systems in place did not complement their work processes and therefore negatively influenced their daily work. For example, stakeholders mentioned that they noticed differences in data entry between colleagues, mainly due to missing definitions or badly defined processes.

It is relevant to ensure uniformity by entering data in the same way in an organization. One person puts a "V" for woman (in Dutch: vrouw), and another tries the whole word "woman" or another enters an "f" for "female" in English.

Another example is weighing our clients. Do we need to weigh our clients with or without clothes: there's already a difference there.

The dialog of the stakeholders shifted to the shortcomings in the digital systems. They mentioned that data systems often require jargon, as decided by the developers. Stakeholders wished to switch to a more "layman's language" approach, as this would support professionals in translating information to make it more comprehensible to their clients, informal caregivers, and even the public.

The fact is that we are looking for clarity and uniformity, but at the same time you must be able to report in your own jargon. So ideally, we also use systems or even technology that can translate the jargon to everyone's needs.

Stakeholders noted that simplifying data presentation was only part of the solution. They also experienced a degree of ambiguity regarding the systems, with the main concern being the lack of central storage. Many recognized that entering data into different systems (eg, EHRs and personal files) increased their administrative burden. Centralizing information on 1 platform in a consistent format could, according to the stakeholders, reduce their workload and enhance data accessibility.

We would have a lot of new possibilities with developments in the organizations, such as gathering all information at one place–as this would allow you to plan much better, control better what has been said, and monitor more.

When it comes to potential, you can also think that more data together has a signaling function. Monitors, signals, help you to say at the right moments: Yes, now, we have to be careful [regarding the situation of this client] and we have to take action now.

Supposing that the uniformity and accessibility were deemed sufficient, stakeholders mentioned that connecting with other systems would be desirable. Stakeholders indicated that most of their clients were admitted from various organizations, such as hospitals or rehabilitation centers. As these organizations might use different systems, the connections needed to retrieve clients' data (eg, background and history) are often missing.

But in specific terms, this is the most important thing we find with this client, this is the most important data that comes with it. Also between other domains, such as primary care, general practitioners, and hospitals, everyone does things in their own way, and that does not connect well with each other.

For instance, this is how it works in practice now: You go to the dentist with a cavity, and afterwards you are being referred to a dental surgeon. You will hear probably that they should take another X-ray, because they cannot get the picture taken at the dentist practice, so twice the cost, twice the time, which becomes your concern.

Human Resources and Finance Departments Pioneer Data Usage, Yet Potential Lies in Clinical Decision-Making

Stakeholders noted that data are collected at all levels of LTC (eg, the client, department, and organization levels), each using different instruments and having various goals. They mentioned that, although data are not used for clinical decision-making, departments, such as human resources (HR) and finance, are already integrating data insights (eg, resource allocation and bed occupancy per department). Stakeholders observed, however, that data usage in LTC primarily focuses on HR and finance. These departments are supporting the organizations, and insights from these departments are crucial for determining the organization's long-term vision, sometimes supported by financial numbers. They pointed out that, by using systematic approaches (eg, problem analysis and running reports), these departments could be an example for other departments to show how information should be presented to a wider audience.

There are also different levels that you have to think about: you have the individual, but you can also learn from each other in the department, and you can learn from each other within the organization.

Data usage can be very effective and works for many people. You can show the performance of your staff, and you will consider more the allocation of resources, which is good for the sustainability of healthcare in the future.

For example, making an order, how to finance things, a lot can be automated based on data if we can properly see what data we already have.

Regarding data-collection and -analysis techniques, stakeholders mentioned the importance of gathering information at different levels to understand what is important and for whom. For example, stakeholders mentioned that working in a systematic order and using a limited number of different instruments for data collection could be used to reduce costs—both the financial cost and the burden experienced by health care providers.

XSL•FO
RenderX

> *We often see that we want to reduce the administrative tasks for an employee–which of course may lead to lower costs for the organization as a whole.*
>
> *Electronic health records of the person is a system in which a part of the wishes and needs is often missing and is not actually discussed when we look at how we can provide care for this person. Therefore, sometimes preparations in analyzing data could actually accelerate [the care] process.*

Stakeholders also discussed the predictive capabilities of data, noting that it can help find patterns and trends, estimate clients' lifestyles and preferences, and predict care intensities. They acknowledged that the data collected and analyzed could support care processes.

> *For example, if there is a client with behavioral issues, you hope to match them with a care professionals with the necessary expertise–and in a later stadium, you even want to predict to development of the behavioral issues among clients. Data servers than as a signaling function, where it monitors with you together and alerts you at the right moment.*

Data-Informed Care Demands Individual, Collective, and Organizational Prerequisites

To realize the benefits of using data for decision-making in LTC, stakeholders indicated that several transformations should be considered. First, stakeholders noted a need for training or education to understand data and its potential on an employee level. They indicated that the purpose of data collection is often unclear, as they experienced collecting data as a routine process. Stakeholders mentioned that the lack of clarity in work processes was leading to the underuse of data to provide tailored care or engage clients in dialog—both crucial for decision-making. Stakeholders mentioned that support would be needed for employees with less digital literacy.

> *Data literacy, learning the language, digital skills, are more often expected as basis for professionals, but it should not be considered as a basis [set of skills] and we should understand that some professionals really need support in this. Thinking about a digital coach.*
>
> *We also mentioned that such things should be included more in education. For future healthcare professionals, certain aspects will change, as the professionals we train now are definitely different from those we trained twenty years ago.*
>
> *To analyze data in groups, and perhaps, we should do this also during the education, to see whether we can let people get familiar with data earlier, learning what data are and what can be done with them.*

In addition, stakeholders mentioned the urgency of better top-down communication. According to them, management should clearly formulate objectives to improve data usage and clarify processes for data-informed care. Most stakeholders noted that collaboration across different layers and professions was required. Stakeholders also discussed the need for better-defined responsibilities and roles, as stakeholders described data-informed decision-making as a collaborative approach that includes multiple perspectives and experiences. Clear role definitions enhance clarity and focus on organizational objectives. Stakeholders suggested introducing roles, such as attention-raisers and data specialists, to maintain focus and improve data usage and understanding. The importance of these new roles within the LTC organizations lies in their specific skill sets, which contribute to both the team level (ie, fulfilling the collective need to understand and work with data) and the organizational level by facilitating and promoting the adoption of data-informed care. These roles allow for combining domain-specific knowledge with ICT (information and communication technology) and data expertise, enabling the bridging of the gap toward practical care applications.

> *If you are not clear about what your vision is as an organization, if you are not clear on what you want to focus on, how can a healthcare professional incorporate data into his or her work?*
>
> *Very important is the communication, collaboration and also a bit of trust.*
>
> *The question is if it is clear to everyone if we discuss a topic in a group setting? If some people say, that is not clear to me and then you immediately notice that everything that has been discussed may be correct, but that there are a number of concepts that were not communicated the right way.*

Stakeholders also noted the importance of teamwork and collective skills, including communication and a "dare to innovate" mentality. For individual contributions to data-informed work, stakeholders mentioned expecting pressure to contribute to group gatherings, including taking control and responsibility. First, stakeholders stressed the need for collaboration across various layers within an organization (ie, health care professionals, managers, and IT specialists).

> *It also requires a bit of openness, a bit of innovative capacity–especially from a diverse group of professionals. Suppose we all want to do something with data-informed working. Who could possibly do that? In fact, everyone should be aware of the importance of this, and we should work together with all layers.*

Second, clear communication of the organization's objects is important. It should be clear what vision the organization has about data usage.

> *You will also need a kind of outcome measure, a result and that addresses the needs of the client, this is our main objective. This should be formulated and communicated in a clear manner.*

Finally, responsibilities and roles should be better defined, as should the tasks of each individual.

> *You have to determine who is responsible for what, and what are we going to do–in which data can support to determine who is responsible for what. So*

XSL•FO
RenderX

you interact with the group get to say to everyone: yes, you had to do that, and you had to do that.

Multidisciplinary Collaboration Enriches Collective Knowledge Regarding Data

Stakeholders indicated that data-informed care needs a multidisciplinary approach in which data usage supports collaboration. Multidisciplinary collaboration involves health care professionals from LTC organizations and emphasizes the participation of clients and relatives. They mentioned that, by integrating data, experiences, and knowledge, new ideas could be generated. It was highlighted that acquiring general knowledge on data during multidisciplinary collaborations becomes more relevant when different professions gather. Moreover, they expected a collective understanding of data and client relevance. In addition, stakeholders mentioned that data should have a central role in a multidisciplinary approach or should even be a starting point.

> *Together, we mainly look at what the needs and wishes of the person are. What information do the relatives of a client already have? How can we achieve this in a combined intervention, a collaboration within our organization? That should be realized.*
>
> *We have to look more for common ground, where we can find each other, and see what we are doing the same, instead of just looking for the differences. We could go a lot further that way, towards a dream scenario–if this could be done with including data to improve multidisciplinary learning.*

Some stakeholders even mentioned that making decisions as a team allowed for comprehensive substantiation, promoting evidence-based practice and justification for care decisions. Stakeholders also noted that team decisions should not mitigate professional perspectives but rather take advantage of the diverse perspectives that occur in such collaborations.

> *When receiving the correct information to translate into data from which you have to choose again: which data are essential? Then you check this [information] multidisciplinary, in dialogue–before and during making a decision.*
>
> *As long as you have a shared goal and everybody is open to exchange [their] insights you can work together. And there are many stakeholders that all play a role to contribute and share their piece of knowledge.*

Stakeholders emphasized that multidisciplinary collaborations could improve individual competencies for all those involved (ie, professionals, clients, and informal caregivers). They noted that it was difficult to specify how and which competencies would be affected but that data perform a supportive role that enriches knowledge and experience—as data could simplify complex information by summarizing or illustrating it graphically. As for clients, some stakeholders mentioned that multidisciplinary approaches may result in better dialogs during consults and improved quality of care when data are available as a central component. Therefore, employees require suitable competencies to introduce data during dialogs with clients.

> *Not sure of how exactly that [specific set of competencies] would look like, but it will of course help a lot in the whole process to work with data and enhancing multidisciplinary collaborations.*
>
> *Therefore, we need the right match between good employees, good competencies and a specific client to work with. Then we could have more time to discuss what everybody in the group [staff, informal caregiver, and client] wants.*

Discussion

Principal Findings

This study aimed to provide insights into the experiences and perceptions of stakeholders from different LTC organizations (eg, health care professionals, client representatives, data and ICT specialists) regarding data and data-informed care. The findings indicate that these stakeholders are aware of the importance of data in LTC organizations but do not yet use data to their full potential. Stakeholders recognized its benefits, particularly in personalizing and improving care quality, and identified diverse types of data, such as quantitative data from wearables and qualitative data from EHRs or video materials. While engaging with data was not perceived as a regular part of their day-to-day professional lives, several stakeholders did recognize their role in collecting and using data (eg, health care professionals documenting information in EHRs and using these data when caring for clients). Data were primarily used for operational purposes, such as HR and finance departments. However, the integration of data into care decision-making remains limited. Stakeholders highlighted key conditions for improving data-informed care, including enhanced data literacy, clearer communication from leadership, and better-defined roles, responsibilities, and work processes. These conditions span different organizational levels—from individual competencies to system-wide coordination and support.

The findings align with previous literature, which emphasizes that health care professionals recognize the potential of data to enhance personalization and care quality, yet remain concerned about data accuracy and adequacy [53]. While data holds the potential to support a collective understanding of clients' needs and preferences, data is not yet systematically applied in everyday practice. Studies mentioned various factors that contribute to increased data use, such as clearly defined work processes [54], roles [55], and responsibilities [56,57]. In line with this, the uptake of data-informed care requires improvements in data literacy, clearer organizational communication, and clarifying data definitions [54-57]. Especially, data literacy plays a foundational role in enabling data-informed practices. It encompasses a set of competencies related to data interpretation, quality improvement, and communication within a multidisciplinary context [58,59]. Research indicates that several personal competencies may be crucial for adopting data-informed practices, such as problem-solving, reasoning, or critical thinking [60]. Understanding different levels of data literacy across professions

XSL•FO
RenderX

enables tailored training programs, such as eLearnings and digital coaches with job-specific cases, to support self-paced learning and improve adoption [23,61-63]. Furthermore, data-informed care is considered a multidisciplinary approach that creates new insights by integrating data, experiences, and knowledge from different disciplines [64,65]. Introducing new roles (eg, data scientist) might enhance data usage, as they require knowledge and skills to leverage data and generate predictions to uncover new insight related to care provision [56]. Especially when care needs are complex, the demand for the skills that support multidisciplinary approaches is growing, as they are often associated with an improved perceived quality of care [66-68]. In health care, research indicates that health care professionals in multidisciplinary teams continue to face challenges in using ICT systems and, consequently, data [28,69].

To date, there are multiple initiatives within LTC to contribute to data-informed care. For example, Business Intelligence (BI) solutions—consisting of systems and activities to process data into information—are often present in these organizations [70-72]. Such BI solutions allow LTC organizations to develop a platform to centralize data from different systems (eg, HR systems, EHRs, and supply chain systems) as they help streamline data entry and retrieval [73]. However, they have yet to be used to support data-informed care for clients. In this study, all participating organizations employed at least 1 BI specialist who supports management by providing data summarized in reports or dashboards. The use, however, is still mainly focused on support for the HR department, while its potential indicates a broader potential in LTC for clinical relevance and quality of care.

To our knowledge, this study is the first to provide an overview of the experiences and perceptions of a diverse group of stakeholders in LTC regarding the use of data. By conducting a World Café session, collective knowledge about data and data-informed care was collected. As all stakeholders had different backgrounds and years of working experience, multiple perspectives were included in this study [51]. To ensure the trustworthiness (ie, credibility and reliability) of this study, multiple actions were deployed. First, data-source triangulation was apparent, as stakeholders from three different LTC organizations and 2 educational institutes participated [74,75]. Second, to enhance trustworthiness in qualitative research, this study provided a detailed overview of the procedures in the data collection and analysis while having a constant dialog within the research team about the codes [76]. Third, to strengthen the confirmability, the main results of the World Café session were summarized and shared in a plenary, as a member check [76]. However, the results of this study should be viewed in light of some possible limitations. First, the authors adopted a reflexive and practice-informed position throughout the study, acknowledging their active role in shaping the research design, data collection, and interpretation of findings. Given the close collaboration with LTC organizations and the participatory nature of the research, the potential for bias was recognized. One factor might be confirmation bias, where researchers—consciously or unconsciously—may have interpreted the data in ways that align with their expectations or prior experiences within the LTC setting [77]. To mitigate this, multiple strategies were used, including triangulation of data sources and iterative coding with multiple researchers. These measures were intended to enhance the credibility of the findings while remaining sensitive to the context in which the research was embedded. Moreover, the plenary session with all stakeholders also served as a validation mechanism to enhance the credibility and interpretability of the findings and supports clarification of ambiguity. Second, the stakeholders who participated in this study might be more interested in data-informed care than the average stakeholder of LTC organizations [78]. Therefore, it might imply that the current result may not be applicable to all stakeholders in the LTC domain. Finally, to realize an impact within the participating organizations to transform toward data-informed care, more stakeholders should have been participating. It supports the dissemination of information and insights while simultaneously considering more experiences and perceptions.

This study identified several challenges that hinder integrating data into daily LTC practices. Addressing these challenges should be a key focus of future research. When prioritizing the adoption of data-informed care in LTC organizations, gaining insights into the status quo of their own data use might be the first step. For example, by conducting a data maturity assessment, several domains (eg, IT architecture, strategic management, data quality, and employee skills) can be evaluated regarding data usage and data-informed working and support for the transformation toward data-informed practices [79]. Gaining insight into these prerequisites for data-informed care within organizations might enhance the support for the development of strategies to embed data-informed care with primary processes. To realize the potential of data-informed care, LTC organizations must prioritize system interoperability, data quality, and a shared understanding of the data's purpose. Policies should support digital literacy at all levels and promote clear roles, responsibilities, and outcome-oriented data strategies. Building on recent national initiatives in the Netherlands, the Artificial Intelligence Navigator for LTC is well-positioned to integrate these insights for broader dissemination. Moreover, there is also a need to promote standardization and alignment of data entry processes across LTC organizations nationally. Variations in terminology, documentation practices, and digital systems currently limit the usability and reliability of data. Establishing uniform data entry guidelines and encouraging centralized, accessible information systems could reduce administrative burden, improve data quality, and prepare for real-time monitoring and personalized care planning nationwide.

Conclusion

Stakeholders in LTC recognize data as essential for personalized care; however, challenges like low standardization and fragmented systems hinder the transitions toward data-informed care. Enhancing data literacy, fostering organizational clarity, and implementing roles, such as data scientists, were considered necessities. Seen as a multidisciplinary approach, data-informed care aims to support collective understanding and inform decision-making, ultimately enriching quality of care by integrating diverse perspectives.

XSL•FO
RenderX

Data Availability

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

Conflicts of Interest

None declared.

References

1. Arowoogun JO, Babawarun O, Chidi R, Adeniyi AO, Okolo CA. A comprehensive review of data analytics in healthcare management: leveraging big data for decision-making. World J Adv Res Rev 2024;21(2):1810-1821 [FREE Full text] [doi: 10.30574/wjarr.2024.21.2.0590]
2. Kruse CS, Mileski M, Vijaykumar AG, Viswanathan SV, Suskandla U, Chidambaram Y. Impact of electronic health records on long-term care facilities: systematic review. JMIR Med Inform 2017 Sep 29;5(3):e35. [doi: 10.2196/medinform.7958] [Medline: 28963091]
3. Trzyna W, Cox B, Coustasse A. Impact of EMR/EHR and computer decision support systems on nursing homes and long-term care. Presented at: MBAA International; Mar 25, 2023; Chicago, IL, USA URL: https://mds.marshall.edu/mgmt_faculty/246/ [accessed 2025-08-26]
4. Zhang Y, Yu P, Shen J. The benefits of introducing electronic health records in residential aged care facilities: a multiple case study. Int J Med Inform 2012 Oct;81(10):690-704. [doi: 10.1016/j.ijmedinf.2012.05.013] [Medline: 22749424]
5. Meehan R. Electronic health records in long-term care: staff perspectives. J Appl Gerontol 2017 Oct;36(10):1175-1196. [doi: 10.1177/0733464815608493] [Medline: 26464335]
6. Elwyn G, Frosch D, Thomson R, et al. Shared decision making: a model for clinical practice. J Gen Intern Med 2012 Oct;27(10):1361-1367. [doi: 10.1007/s11606-012-2077-6] [Medline: 22618581]
7. Cranley LA, Yeung L, Tu W, McGillis Hall L. Healthcare aide involvement in team decision-making in long-term care: a narrative review of the literature. J Clin Nurs 2023 Aug;32(15-16):4217-4227. [doi: 10.1111/jocn.16573] [Medline: 36333860]
8. Kruse CS, Goswamy R, Raval YJ, Marawi S. Challenges and opportunities of big data in health care: a systematic review. JMIR Med Inform 2016 Nov 21;4(4):e38. [doi: 10.2196/medinform.5359] [Medline: 27872036]
9. Wasfy JH, Maddox TM. Worlds apart: the hype and reality of big data to improve health care. Circ Cardiovasc Qual Outcomes 2016 Sep;9(5):495-497. [doi: 10.1161/CIRCOUTCOMES.116.003150] [Medline: 27601457]
10. Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery. Lancet Oncol 2019 May;20(5):e262-e273. [doi: 10.1016/S1470-2045(19)30149-4] [Medline: 31044724]
11. Pirracchio R, Cohen MJ, Malenica I, et al. Big data and targeted machine learning in action to assist medical decision in the ICU. Anaesth Crit Care Pain Med 2019 Aug;38(4):377-384. [doi: 10.1016/j.accpm.2018.09.008]
12. Roski J, Bo-Linn GW, Andrews TA. Creating value in health care through big data: opportunities and policy implications. Health Aff (Millwood) 2014 Jul;33(7):1115-1122. [doi: 10.1377/hlthaff.2014.0147] [Medline: 25006136]
13. Galetsi P, Katsaliaki K, Kumar S. Values, challenges and future directions of big data analytics in healthcare: a systematic review. Soc Sci Med 2019 Nov;241:112533. [doi: 10.1016/j.socscimed.2019.112533] [Medline: 31585681]
14. Prainsack B. The value of healthcare data: to nudge, or not? Policy Studies 2020 Sep 2;41(5):547-562. [doi: 10.1080/01442872.2020.1723517]
15. Golz C, Aarts S, Hacking C, Hahn S, Zwakhalen SMG. Health professionals' sentiments towards implemented information technologies in psychiatric hospitals: a text-mining analysis. BMC Health Serv Res 2022 Nov 28;22(1):1426. [doi: 10.1186/s12913-022-08823-4] [Medline: 36443845]
16. Safi S, Thiessen T, Schmailzl KJ. Acceptance and resistance of new digital technologies in medicine: qualitative study. JMIR Res Protoc 2018 Dec 4;7(12):e11072. [doi: 10.2196/11072] [Medline: 30514693]
17. Cammer A, Morgan D, Stewart N, et al. The Hidden Complexity of Long-Term Care: how context mediates knowledge translation and use of best practices. Gerontologist 2014 Dec;54(6):1013-1023. [doi: 10.1093/geront/gnt068] [Medline: 23856027]
18. Ng R, Lane N, Tanuseputro P, et al. Increasing complexity of new nursing home residents in Ontario, Canada: a serial cross-sectional study. J Am Geriatr Soc 2020 Jun;68(6):1293-1300. [doi: 10.1111/jgs.16394] [Medline: 32119121]
19. Akyürek Ç, Sawalha R, Ide S. Factors affecting the decision making process in healthcare institutions. Acad Strateg Manag J 2015;14(10):430-439 [FREE Full text]
20. Chiolero A, Rodondi N, Santschi V. High-value, data-informed, and team-based care for multimorbidity. Lancet Public Health 2020 Feb;5(2):e84. [doi: 10.1016/S2468-2667(20)30007-4] [Medline: 32032562]
21. Goodwin N, Sonola L, Thiel V, Kodner D. Co-ordinated care for people with complex chronic conditions: key lessons and markers for success. : The King's Fund; 2013.

XSL•FO
RenderX

22. Aarts S. Ethical usage of data in long-term care: how do we proceed. In: Moral Design and Technology: Wageningen Academic; 2022:267-282. [doi: 10.3920/978-90-8686-922-0_15]
23. Doll J, Anzalone AJ, Clarke M, Cooper K, Polich A, Siedlik J. A call for a health data-informed workforce among clinicians. JMIR Med Educ 2024 Jun 17;10(1):e52290. [doi: 10.2196/52290] [Medline: 38889091]
24. Elliott J, Gordon A, Tong CE, Stolee P. "We've got the home care data, what do we do with it?": understanding data use in decision making and quality improvement. BMC Health Serv Res 2020 Mar 25;20(1):251. [doi: 10.1186/s12913-020-5018-9] [Medline: 32213180]
25. Fijten R, Wee L, Dekker A, Roumen C. Data-driven shared decision-making: a paradigm shift. J Radiat Oncol Inform 2021;11(2):4. [doi: 10.5166/jroi.11.2.1]
26. Haase CB, Brodersen JB. 88 data driven or data informed? How general practitioners use data to evaluate their own and colleagues' clinical work in clusters. 2023 Presented at: BMJ Evid Based Med; Aug 15, 2023 p. A43-A44. [doi: 10.1136/EBM-2023-POD.88]
27. Aarts S, Daniels R, Hamers J, Verbeek H. Data in long-term care for older adults [Article in Dutch]. Tijdschr Ouderengeneeskd 2020(5):92-96 [FREE Full text]
28. Sun Y, Ma X, Lindtner S, He L. Data work of frontline care workers: practices, problems, and opportunities in the context of data-driven long-term care. Proc ACM Hum-Comput Interact 2023 Apr 14;7(CSCW1):1-28. [doi: 10.1145/3579475]
29. Bossen C, Pine KH, Cabitza F, Ellingsen G, Piras EM. Data work in healthcare: an introduction. Health Informatics J 2019 Sep;25(3):465-474. [doi: 10.1177/1460458219864730] [Medline: 31405353]
30. Aarts S. Data-informed decision making in long-term care for older adults: what do we need. : International Psychohistorical Association; 2023 URL: https://assets.associationaccess.com/emails/ipa/ipa-bulletin-vol40no2-v2.pdf [accessed 2025-09-04]
31. Brown J. The World Café: Shaping Our Futures through Conversations That Matter: Berrett-Koehler; 2010. URL: https://books.google.nl/books?id=oymkkyQWmJAC&amp [accessed 2025-09-04]
32. Löhr K, Weinhardt M, Sieber S. The "World Café" as a participatory method for collecting qualitative data. Int J Qual Methods 2020 Jan 1;19:1609406920916976. [doi: 10.1177/1609406920916976]
33. Schiele H, Krummaker S, Hoffmann P, Kowalski R. The "research world café" as method of scientific enquiry: combining rigor with relevance and speed. J Bus Res 2022 Feb;140:280-296. [doi: 10.1016/j.jbusres.2021.10.075]
34. Verbeek H, Zwakhalen SMG, Schols J, Kempen G, Hamers JPH. The living lab in ageing and long-term care: a sustainable model for translational research improving quality of life, quality of care and quality of work. J Nutr Health Aging 2020;24(1):43-47. [doi: 10.1007/s12603-019-1288-5] [Medline: 31886807]
35. Schieffer A, Isaacs D, Gyllenpalm B. The World Café: part one. World Business Academy 2004;18(8):1-9 [FREE Full text]
36. MacFarlane A, Galvin R, O'Sullivan M, et al. Participatory methods for research prioritization in primary care: an analysis of the World Café approach in Ireland and the USA. Fam Pract 2017 Jun 1;34(3):278-284. [doi: 10.1093/fampra/cmw104] [Medline: 27677298]
37. Vogt EE, Brown J, Isaacs D. The Art of Powerful Questions: Catalyzing, Insight, Innovation, and Action: Whole Systems Associates: Whole Systems Associates; 2003. URL: https://www.sparc.bc.ca/wp-content/uploads/2020/11/the-art-of-powerful-questions.pdf [accessed 2025-08-26]
38. van Lierop MEA, Meijers JMM, van Rossum E, Rutten JER, Thoma-Lürken T, Zwakhalen SMG. How to establish workplace learning in long-term care: results from a World Café dialogue. BMC Nurs 2022 Aug 29;21(1):241. [doi: 10.1186/s12912-022-00999-8] [Medline: 36038912]
39. Arnstein SR. A ladder of citizen participation. J Am Inst Plann 1969 Jul;35(4):216-224. [doi: 10.1080/01944366908977225]
40. Fouché C, Light G. An invitation to dialogue: 'The World Café' in social work research. Qual Soc Work 2011;10(1):28-48. [doi: 10.1177/1473325010376016]
41. Poland BD. Transcription quality as an aspect of rigor in qualitative research. Qual Inq 1995 Sep;1(3):290-310. [doi: 10.1177/107780049500100302]
42. Kuckartz U, Rädiker S. Analyzing Qualitative Data with MAXQDA: Springer; 2019. [doi: 10.1007/978-3-030-15671-8]
43. Tesch R. Qualitative Research: Analysis Types and Software: Routledge; 2013. [doi: 10.4324/9781315067339]
44. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
45. Boeije H. Analyzing in qualitative research [Book in Dutch]. In: Denken En Doen: Uitgeverij Boom; 2005. URL: https://www.boom.nl/media/21/9789024425945_inkijkexemplaar_analyseren_in_kwalitatief_onderzoek.pdf [accessed 2025-08-26]
46. Glaser B, Strauss A. Discovery of Grounded Theory: Strategies for Qualitative Research: Routledge; 1967. [doi: 10.1097/00006199-196807000-00014]
47. Kondracki NL, Wellman NS, Amundson DR. Content analysis: review of methods and their applications in nutrition education. J Nutr Educ Behav 2002;34(4):224-230. [doi: 10.1016/s1499-4046(06)60097-3] [Medline: 12217266]
48. Bradley EH, Curry LA, Devers KJ. Qualitative data analysis for health services research: developing taxonomy, themes, and theory. Health Serv Res 2007 Aug;42(4):1758-1772. [doi: 10.1111/j.1475-6773.2006.00684.x] [Medline: 17286625]
49. Khandkar SH. Open Coding: University of Calgary; 2009, Vol. 23. URL: https://cspages.ucalgary.ca/~saul/wiki/uploads/CPSC681/opencoding.pdf [accessed 2025-08-26]

XSL•FO
RenderX

50. Corbin J, Strauss A. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory: Sage Publications; 2014. URL: https://methods.sagepub.com/book/mono/basics-of-qualitative-research/toc [accessed 2025-09-04]
51. Shenton AK. Strategies for ensuring trustworthiness in qualitative research projects. Edu Inf 2004;22(2):63-75. [doi: 10.3233/EFI-2004-22201]
52. O'Sullivan TA, Jefferson CG. A review of strategies for enhancing clarity and reader accessibility of qualitative research results. Am J Pharm Educ 2020 Jan;84(1):7124. [doi: 10.5688/ajpe7124] [Medline: 32292189]
53. Neves AL, Poovendran D, Freise L, et al. Health care professionals' perspectives on the secondary use of health records to improve quality and safety of care in England: qualitative study. J Med Internet Res 2019 Sep 26;21(9):e14135. [doi: 10.2196/14135] [Medline: 31573898]
54. Coleman C, Hudson S, Pederson B. Prioritized health literacy and clear communication practices for health care professionals. Health Lit Res Pract 2017 Jul;1(3):e91-e99. [doi: 10.3928/24748307-20170503-01] [Medline: 31294254]
55. Groysberg B, Slind M. Leadership is a conversation. Harv Bus Rev 2012 Jun;90(6):76-84. [Medline: 22741420]
56. Koltay T. Data governance, data literacy and the management of data quality. IFLA Journal 2016 Dec;42(4):303-312. [doi: 10.1177/0340035216672238]
57. Yoshida Y, Hirakawa Y, Hong YJ, et al. Factors influencing interprofessional collaboration in long-term care from a multidisciplinary perspective: a case study approach. Home Health Care Serv Q 2024;43(4):239-258. [doi: 10.1080/01621424.2024.2331452] [Medline: 38521999]
58. Dixon BE, Barboza K, Jensen AE, Bennett KJ, Sherman SE, Schwartz MD. Measuring practicing clinicians' information literacy. Appl Clin Inform 2017 Aug;26(1):149-161. [doi: 10.4338/ACI-2016-06-RA-0083]
59. Foadi N, Varghese J. Digital competence - a key competence for todays and future physicians. J Eur CME 2022;11(1):2015200. [doi: 10.1080/21614083.2021.2015200] [Medline: 34992949]
60. Schmidt DH, van Dierendonck D, Weber U. The data-driven leader: developing a big data analytics leadership competency framework. J Manag Dev 2023 Jul 19;42(4):297-326. [doi: 10.1108/JMD-12-2022-0306]
61. Farokhzadian J, Jouparinejad S, Fatehi F, Falahati-Marvast F. Improving nurses' readiness for evidence-based practice in critical care units: results of an information literacy training program. BMC Nurs 2021 May 18;20(1):79. [doi: 10.1186/s12912-021-00599-y] [Medline: 34001116]
62. Kateb S, Ruehle RC, Kroon DP, van Burg E, Huber M. Innovating under pressure: adopting digital technologies in social care organizations during the COVID-19 crisis. Technovation 2022 Jul;115:102536. [doi: 10.1016/j.technovation.2022.102536]
63. Aryee GFB, Amoadu M, Obeng P, et al. Effectiveness of eLearning programme for capacity building of healthcare professionals: a systematic review. Hum Resour Health 2024 Sep 2;22(1):60. [doi: 10.1186/s12960-024-00924-x] [Medline: 39223555]
64. Shaw M, Heyman B, Reynolds L, Davies J, Godin P. Multidisciplinary teamwork in a UK regional secure mental health unit: a matter for negotiation? Soc Theory Health 2007 Nov;5(4):356-377. [doi: 10.1057/palgrave.sth.8700103]
65. Roncaglia I. A practitioner's perspective of multidisciplinary teams: analysis of potential barriers and key factors for success. Psychol Thought 2016;9(1):15-23. [doi: 10.5964/psyct.v9i1.145]
66. Doornebosch AJ, Achterberg WP, Smaling HJA. Factors influencing interprofessional collaboration in general and during multidisciplinary team meetings in long-term care and geriatric rehabilitation: a qualitative study. BMC Med Educ 2024 Mar 15;24(1):285. [doi: 10.1186/s12909-024-05291-8] [Medline: 38486216]
67. Stühlinger M, Schmutz JB, Grote G. I hear you, but do I understand? The relationship of a shared professional language with quality of care and job satisfaction. Front Psychol 2019;10:1310. [doi: 10.3389/fpsyg.2019.01310] [Medline: 31214098]
68. Boorsma M, Frijters DHM, Knol DL, Ribbe ME, Nijpels G, van Hout HPJ. Effects of multidisciplinary integrated care on quality of care in residential care facilities for elderly people: a cluster randomized trial. CMAJ 2011 Aug 9;183(11):E724-E732. [doi: 10.1503/cmaj.101498] [Medline: 21708967]
69. Janssen A, Robinson T, Brunner M, Harnett P, Museth KE, Shaw T. Multidisciplinary teams and ICT: a qualitative study exploring the use of technology and its impact on multidisciplinary team meetings. BMC Health Serv Res 2018 Jun 13;18(1):444. [doi: 10.1186/s12913-018-3242-3] [Medline: 29898716]
70. Mettler T, Vimarlund V. Understanding business intelligence in the context of healthcare. Health Informatics J 2009 Sep;15(3):254-264. [doi: 10.1177/1460458209337446] [Medline: 19713399]
71. Böhringer M, Gluchowski P, Kurze C, Schieder C. A business intelligence perspective on the future internet. Presented at: Sustainable IT Collaboration Around the Globe 16th Americas Conference on Information Systems; Aug 12-15, 2010; Lima, Peru URL: https://aisel.aisnet.org/amcis2010/267/ [accessed 2025-08-26]
72. Spruit M, Vroon R, Batenburg R. Towards healthcare business intelligence in long-term care: an explorative case study in the Netherlands. Comput Human Behav 2014;30:698-707. [doi: 10.1016/j.chb.2013.07.038]
73. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Inf Sci Syst 2014;2:3. [doi: 10.1186/2047-2501-2-3] [Medline: 25825667]
74. Polit DF, Beck CT. Nursing Research: Generating and Assessing Evidence for Nursing Practice: Lippincott Williams & Wilkins; 2008.

XSL•FO
RenderX

75. Carter N, Bryant-Lukosius D, DiCenso A, Blythe J, Neville AJ. The use of triangulation in qualitative research. In: R NC, D BL, R AD, editors. Oncol Nurs Forum 2014 Sep;41(5):545-547. [doi: 10.1188/14.ONF.545-547] [Medline: 25158659]
76. Lincoln YS, Guba EG. Naturalistic Inquiry: Sage; 1985.
77. Berthet V, Teovanović P, de Gardelle V. A common factor underlying individual differences in confirmation bias. Sci Rep 2024 Nov 13;14(1):27795. [doi: 10.1038/s41598-024-78053-7] [Medline: 39537676]
78. Robinson OC. Sampling in interview-based qualitative research: a theoretical and practical guide. Qual Res Psychol 2014 Jan 2;11(1):25-41. [doi: 10.1080/14780887.2013.801543]
79. van den Berg B, Lukkien D, de Charro F, Snellen M, Beelen A, Bouwens M. How can the implementation of data-driven practices enhance efficiency and improve the quality of care in long-term care settings? [Article in Dutch]. TSG Tijdschr Gezondheidswet 2024;102(1):33-38. [doi: 10.1007/s12508-024-00425-1]

Abbreviations

AWO-L: Academische Werkplaats Ouderenzorg Limburg
BI: Business Intelligence
EHR: electronic health record
HR: human resources
ICT: information and communication technology
LTC: long-term care

Edited by PH Chau; submitted 29.11.24; peer-reviewed by C Recsky, HR Choi; revised version received 17.07.25; accepted 19.07.25; published 12.09.25.

Please cite as:
Bouchmal S, Sion KYJ, Hamers JPH, Aarts S
Toward Data-Informed Care in Long-Term Care: Qualitative Analysis
JMIR Aging 2025;8:e69423
URL: https://aging.jmir.org/2025/1/e69423
doi:10.2196/69423

© Suleyman Bouchmal, Katya YJ Sion, Jan PH Hamers, Sil Aarts. Originally published in JMIR Aging (https://aging.jmir.org), 12.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Machine Learning for Predicting Postoperative Functional Disability and Mortality Among Older Patients With Cancer: Retrospective Cohort Study

Yuki Hashimoto[1], PhD; Norihiko Inoue[1], MD, PhD; Takuaki Tani[2], PhD; Shinobu Imai[2], PhD

[1]Department of Clinical Data Management and Research, Clinical Research Center, National Hospital Organization Headquarters, 2-5-21 Higashigaoka, Meguroku, Japan

[2]Department of Pharmacoepidemiology, Showa University Graduate School of Pharmacy, Shinagawaku, Japan

Corresponding Author:
Yuki Hashimoto, PhD
Department of Clinical Data Management and Research, Clinical Research Center, National Hospital Organization Headquarters, 2-5-21 Higashigaoka, Meguroku, Japan

Abstract

Background: The global cancer burden is rapidly increasing, with 20 million new cases estimated in 2022. The world population aged ≥65 years is also increasing, projected to reach 15.9% by 2050, making cancer control for older patients urgent. Surgical resection is important for cancer treatment; however, predicting postoperative disability and mortality in older patients is crucial for surgical decision-making, considering the quality of life and care burden. Currently, no model directly predicts postoperative functional disability in this population.

Objective: We aimed to develop and validate machine-learning models to predict postoperative functional disability (≥5-point decrease in the Barthel Index) or in-hospital death in patients with cancer aged ≥ 65 years.

Methods: This retrospective cohort study included patients aged ≥65 years who underwent surgery for major cancers (lung, stomach, colorectal, liver, pancreatic, breast, or prostate cancer) between April 2016 and March 2023 in 70 Japanese hospitals across 6 regional groups. One group was randomly selected for external validation, while the remaining 5 groups were randomly divided into training (70%) and internal validation (30%) sets. Predictor variables were selected from 37 routinely available preoperative factors through electronic medical records (age, sex, income, comorbidities, laboratory values, and vital signs) using crude odds ratios (P<.1) and the least absolute shrinkage and selection operator method. We developed 6 machine-learning models, including category boosting (CatBoost), extreme gradient boosting (XGBoost), logistic regression, neural networks, random forest, and support vector machine. Model predictive performance was evaluated using the area under the receiver operating characteristic curve (AUC) with 95% CI. We used the Shapley additive explanations (SHAP) method to evaluate contribution to the predictive performance for each predictor variable.

Results: This study included 33,355 patients in the training, 14,294 in the internal validation, and 6711 in the external validation sets. In the training set, 1406/33,355 (4.2%) patients experienced worse discharge. A total of 24 predictor variables were selected for the final models. CatBoost and XGBoost achieved the largest AUCs among the 6 models: 0.81 (95% CI 0.80-0.82) and 0.81 (95% CI 0.80-0.82), respectively. In the top 15 influential factors based on the mean absolute SHAP value, both models shared the same 14 factors such as dementia, age ≥85 years, and gastrointestinal cancer. The CatBoost model showed the largest AUCs in both internal (0.77, 95% CI 0.75-0.79) and external validation (0.72, 95% CI 0.68-0.75).

Conclusions: The CatBoost model demonstrated good performance in predicting postoperative outcomes for older patients with cancer using routinely available preoperative factors. The robustness of these findings was supported by the identical top influential factors between the CatBoost and XGBoost models. This model could support surgical decision-making while considering postoperative quality of life and care burden, with potential for implementation through electronic health records.

(JMIR Aging 2025;8:e65898) doi:10.2196/65898

KEYWORDS

older patients with cancer; postoperative outcomes; functional disability; machine learning; decision-making

Introduction

The global cancer burden is rapidly increasing, with an estimated 20 million new cases and 9.7 million deaths in 2022 [1]. In Japan, the lifetime risk of being diagnosed with cancer is approximately 65.5% and 51.2% for men and women, respectively [2].

XSL•FO
RenderX

In addition, the global population is aging rapidly, with the proportion of those aged ≥65 years expected to increase from 9.1% in 2019 to 15.9% by 2050 [3]. Japan faces the most advanced stage of this demographic shift, with the population of older adults expected to increase from 28.8% in 2020 to 37.7% by 2050 [4]. Thus, cancer control for older patients has become an urgent issue worldwide, including in Japan.

Older patients with cancer often face challenges such as frailty [5], comorbidities [6-8], and socioeconomic status [6], which affect treatment outcomes and quality of life (QOL). Surgical resection is a key treatment, which requires careful consideration in older patients with cancer due to concerns about postoperative functional disability and its impact on long-term outcomes [9]. Some factors may influence surgical outcomes in older patients, including age [9], anemia [10], BMI [11,12], dementia [6,7], frailty [5], low household income [6], malnutrition [13], and smoking [14].

Considering the postoperative QOL and care burden on patients' families and society, it is important to predict not only postoperative mortality but also functional disability [15]. Hospital-associated disability, defined as functional disability following acute hospitalization, is recognized as a crucial outcome in older patients with significant impact on health care costs and long-term prognosis [16,17]. Some models can predict postoperative mortality [5,13]; however, few have addressed functional disability. A model after lower-extremity surgery [6] could predict the risk of in-hospital mortality and discharge to a nursing home, which is a surrogate for functional disability, in patients admitted from home. Currently, no model can directly predict functional disability after cancer surgery.

Therefore, this study aimed to develop a machine learning–based model for predicting postoperative functional disability and in-hospital mortality in patients with cancer aged ≥65 years using data from 70 hospitals across Japan. This approach will enable patients and their families to make informed decisions about undergoing cancer surgery, considering their postoperative QOL and care burden.

Methods

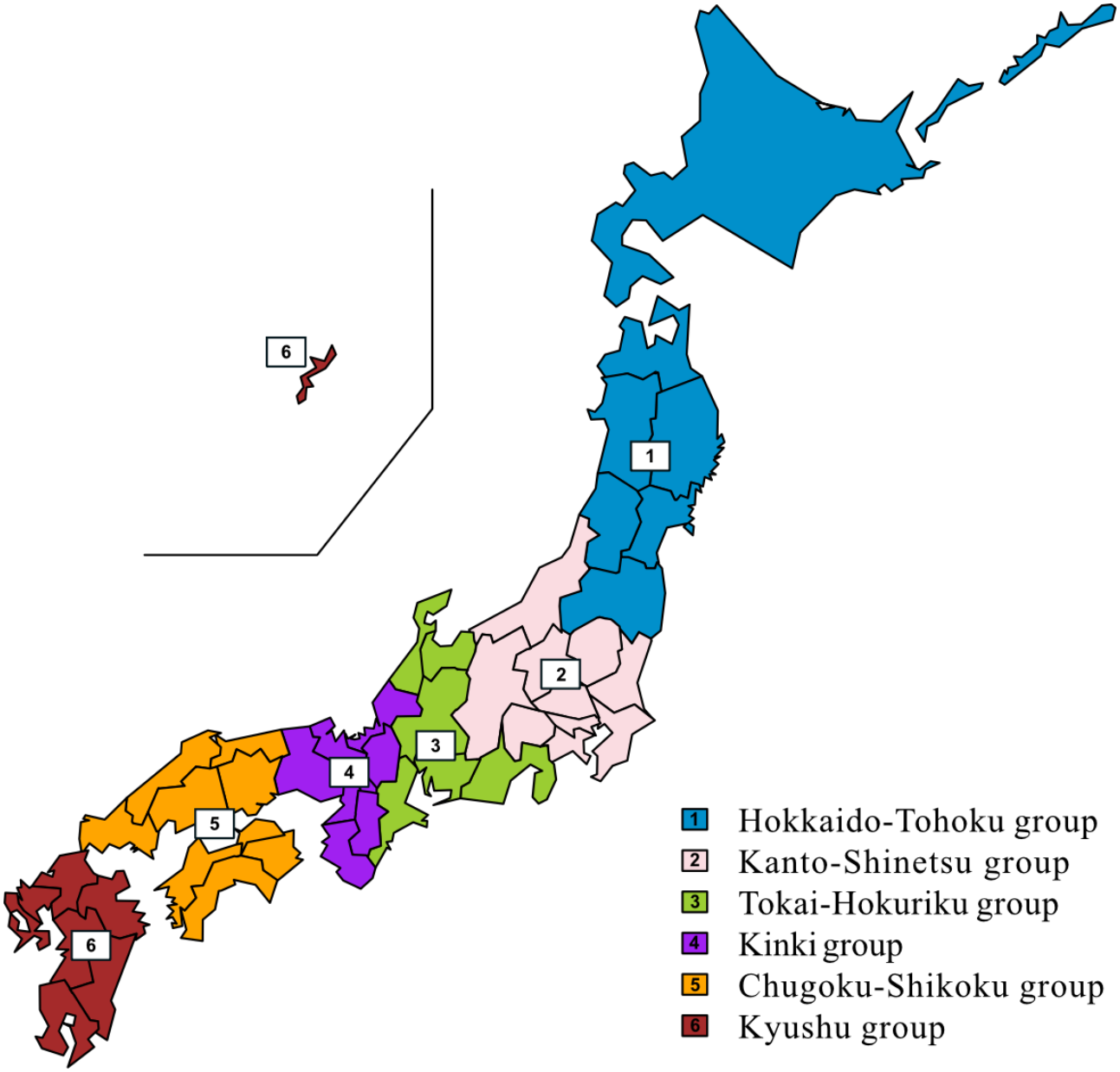
Study Design and Data Source

We conducted a retrospective cohort study to develop machine-learning models for predicting postoperative functional disability and in-hospital mortality in older patients with cancer. This study used data between April 2016 and March 2023 from 70 Japanese hospitals within the National Hospital Organization (NHO) database across 6 regional groups: Hokkaido-Tohoku group, Kanto-Shinetsu group, Tokai-Hokuriku group, Kinki group, Chugoku-Shikoku group, and Kyushu group (Figure 1) [18].

The NHO maintains 2 databases: (1) an administrative claims database based on the Diagnosis Procedure Combination–based Per-Diem Payment System [19] and a clinical information database based on the standardized structured medical record information exchange [20]. The administrative claims database contains patient information, such as age, sex, cost, comorbidities, complications, diagnosis, medical procedures, and medications. The clinical information database includes medical charts, laboratory data, and vital signs on a daily basis.

XSL•FO
RenderX

Figure 1. A total of 6 National Hospital Organization regional groups across Japan.



Participants

The study included patients aged ≥65 years who were admitted to NHO hospitals between April 2016 and March 2023 and underwent surgery for major cancers, including lung, stomach, colorectal, liver, pancreatic, breast, and prostate cancers. These cancer sites were selected because of their high incidence [2] and mortality rates [21] in Japanese and global cancer statistics [1]. The surgical procedures included both scopic and open surgeries under general anesthesia.

We excluded patients who had missing Barthel Index data at admission or discharge, were first included in the database at admission (no medical history available), underwent surgery more than 1 week after admission, or had the Barthel Index of 0 at admission. These exclusion criteria were applied because: (1) patients missing Barthel Index data could not be evaluated for outcome, (2) most of the predictor variables were missing if patients had no medical history, (3) we eliminated the effect of hospitalization on physical function from admission to surgery, and (4) the Barthel Index change from admission to discharge was an outcome variable, but patients with a minimum score of 0 at admission cannot show further decline [22].

Outcome Variables

The primary outcome was worse discharge, defined as either in-hospital death or postoperative functional disability. Postoperative functional disability was characterized as hospital-associated disability [16] (≥5-point decrease in the Barthel Index between admission and discharge) [17]. The Barthel Index consists of 10 items, including transfer, bathing, and stair climbing, used to evaluate activities of daily living on a scale of 0 - 100, with lower scores indicating a decline in physical function.

The secondary outcomes were health care costs and postoperative length of stay (LOS) in patients predicted to be at high risk and low risk for worse discharge. The optimal cut-off point for high-risk or low-risk classification was

XSL•FO
RenderX

determined using the Youden index [23] on the receiver operating characteristic (ROC) curve.

Predictor Variables

We identified the following 37 potential predictors of worse discharge based on previous studies: age (65 - 74, 75 - 84, and ≥85 years) [9], sex [24], underweight (BMI <18.5 kg/m^2) [12] and obesity (BMI ≥30 kg/m^2) [11], route of admission (home or nonhome), emergency admission, an estimated household income in the lowest tertile based on post-code (low income) [6], smoking (Brinkman Index ≥200) [14], functional dependence (Barthel Index ≤60) [22], surgical factors (open surgery, scopic surgery, and combined general and epidural anesthesia), the presence of gastrointestinal cancer (colorectal, liver, pancreas, and stomach), the cancer staging (0–II or III–IV) [25], the presence of recurrent cancer, Charlson Comorbidity Index (CCI) ≥3 [26,27], Hospital Frailty Risk Score (HFRS) ≥5 [28], comorbidities common in older patients (cerebrovascular disease, chronic pulmonary disease, congestive heart failure, dementia [8], diabetes, liver disease, myocardial infarction, peptic ulcer disease, peripheral vascular disease, and renal disease), medical history about cancer within 8 weeks before surgery that allowed the evaluation of preoperative chemotherapy [29] and radiotherapy [30] (chemotherapy, radiation, and surgery), vital signs (body temperature ≥38°C [31], systolic blood pressure >180 mm Hg [32]), and laboratory test values (albumin <3.5 g/dL [13], total bilirubin ≥2.0 mg/dL [33], creatinine ≥ 2.0 mg/dL [33], platelet <10^5/μL [33], and hemoglobin <11 g/dL [10]).

These predictors were measured as follows: age, sex, BMI, route of admission, emergency admission, income, cancer staging evaluated using the Tumor Nodes Metastasis (TNM) classification system [34], the presence of recurrent cancer, Brinkman Index, and Barthel Index were assessed at admission. Surgical factors were assessed during the surgery. The type of cancer was determined during the surgery using the *International Statistical Classification of Diseases and Related Health Problems, 10th Revision* (*ICD-10*) coding system [35]. Comorbidities were identified from *ICD-10* codes within 8 weeks before surgery (Table S1 in Multimedia Appendix 1). Notably, dementia was determined if it was either identified from *ICD-10* codes within 8 weeks before surgery or documented in the clinical summary at admission, as Japanese hospitals are required to include the dementia status of inpatients aged ≥65 years at admission [8]. Vital signs were measured closest to surgery after admission, and laboratory values were obtained using those measured closest to surgery within 8 weeks before surgery.

Statistical Analysis

We randomly selected one of the 6 NHO regional groups for the external validation set [36], while hospital data from the remaining 5 groups were randomly divided into the training (70%) and the internal validation (30%) sets [37]. The missing predictor variables were imputed using the "missRanger" [38], which is a random forest–based algorithm [39], assuming that the data are missing at random.

To summarize patient characteristics, continuous variables were expressed as mean (SD) or median (IQR), depending on the distribution of variables. The Wilcoxon rank-sum test or the Welch test was used to assess between-group differences. Categorical variables were expressed as proportions and compared using the χ^2 test.

In the training set, a double penalty was implemented by eliminating unnecessary variables to create a more practical model for clinical use [40]. The first penalty involved selecting predictor candidates with a crude odds ratio (OR) at P<.1 [41]. The second penalty to further narrow down the predictor candidates used the least absolute shrinkage and selection operator (Lasso) method, which allowed for the selection of clinically relevant variables with consistent relationships [36,42].

The selected factors were incorporated into 6 machine learning models: category boosting (CatBoost) [36,43], extreme gradient boosting (XGBoost) [36,42], logistic regression, neural networks [44], random forest [36], and support vector machine (SVM) [36]. Model performance was evaluated using the area under the ROC curve (AUC) with 95% CIs). Similarly, we calculated the accuracy, sensitivity, specificity, F_1-score, and the area under the precision-recall curve (PRAUC) to assess the performance of the models [36,37]. The precision-recall curve of the models was also shown.

We used the synthetic minority oversampling and random undersampling techniques to avoid overfitting owing to the imbalance between the positive and negative events [45]. The minority class was oversampled at 50%, 100%, and 200% of its original size, followed by random undersampling of the majority class to achieve equal numbers between classes. The models were trained using 10-fold cross-validation with grid search for hyperparameter optimization. Of these sampling ratios and hyperparameter combinations, those yielding the largest AUC were selected. We ranked the predictor variable using the Shapley additive explanations (SHAP) method [46] to assess the contribution of the predictors to the models. Moreover, considering higher cancer stages are associated with poorer postoperative outcomes [25], a multiple logistic regression was conducted to evaluate the interaction between cancer staging and other predictor variables using OR (95% CI). We examined interactions between cancer stage III–IV and the top 5 features based on the mean absolute SHAP value.

We analyzed the AUC difference between the models using the DeLong method [47], which was based on the model with the largest AUC in both internal and external validation. Sensitivity analyses were used to assess the impact of missing data and assessing 2 outcomes simultaneously (death or functional disability) on the models. For missing values, a complete case analysis was performed to confirm the robustness of the results obtained from the imputed data set. We also evaluated the AUC of the models when the models predicted only death in all patients or functional disability in patients with survival discharge. We conducted subgroup analyses by LOS, cancer type (breast, colorectal, liver, lung, pancreas, prostate, or stomach), and cancer staging (stage 0–I, II, III, or IV). For the LOS analysis, we calculated the 75th percentile of LOS for each cancer type separately and divided patients into LOS for each

XSL•FO
RenderX

cancer type <75th percentile (short-stay) and LOS ≥75th percentile (long-stay) groups [48].

For the analysis of secondary outcomes, we compared LOS and health care costs between patients at high and low risk based on the model with the highest AUC in both internal and external validation. Health care costs were assessed across various categories, including total, medical consultation, medication, medical procedure, surgical procedure, laboratory tests, hospital stay, and others. The currency conversion rate was 150 JPY to US $1.

All hypothesis tests had a 2-sided significance level of .05. All statistical analyses were performed using R version 4.3.1 (R Foundation for Statistical Computing) .

Ethical Considerations

Our study was approved by the Institutional Review Board of Showa University (approval number 2023 - 129-A). Individual consent was not required because this was an opt-out study. This study conforms to the principles outlined in the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) statement.

Results

In total, 54,360 patients from 70 hospitals were included in this study (Figure 2): 6711 in the external validation set from Kinki group, 33,355 in the training set, and 14,294 in the internal validation set from the remaining 5 regional groups. In the training set, 1406/33,355 (4.2%) patients experienced worse discharge (Table 1). These patients were older (24.5% [344/1406] vs 5.8% [1864/31,949] for age ≥85 years), more likely to be admitted from nonhome (8.1% [114/1406] vs 0.9% [281/31,949]), and had higher prevalence of dementia (28.7% [403/1406] vs 5.6% [1783/31,949]), gastrointestinal cancer (64.3% [904/1406] vs 41.7% [13,326/31,949]), and advanced cancer stage (32.6% [458/1406] vs 24.2% [7716/31,9] for stage III–IV).

In the training set, we selected 31 predictor variables for worse discharge using crude OR (Table S2 in Multimedia Appendix 1). Further selection using the Lasso method resulted in 24 factors: age (75 - 84 y and ≥85 y), male sex, BMI <18.5 kg/m^2, nonhome admission, emergency admission, low income, Barthel Index at admission ≤60, open surgery, gastrointestinal cancer, cancer stage III–IV, HFRS ≥5, cerebrovascular disease, congestive heart failure, dementia, diabetes, liver disease, myocardial infarction, medical history of chemotherapy, systolic blood pressure ≥180 mm Hg, albumin<3.5 g/dL, creatinine ≥ 2.0 mg/dL, platelet <10^5/μL, and hemoglobin <11 g/dL. All 6 models were developed using these 24 variables.

Furthermore, the AUCs were 0.81 (95% CI 0.80 - 0.82) for CatBoost, 0.81 (95% CI 0.80 - 0.82) for XGBoost, 0.79 (95% CI 0.78 - 0.80) for random forest, 0.79 (95% CI 0.78 - 0.80) for neural networks, 0.78 (95% CI 0.77 - 0.80) for SVM, and 0.78 (95% CI 0.77 - 0.80) for logistic regression (Figure 3). CatBoost and XGBoost were the 2 models with AUC ≥0.80 and showed similar values with relatively high accuracy (0.76), sensitivity (0.72), specificity (0.76), F_1-score (0.20), and area under the precision-recall curve (PRAUC) (0.22). The performance metrics and precision-recall curves for all models are shown in Table S3 and Figure S1 in Multimedia Appendix 1, respectively. In the top 15 influential factors based on the mean absolute SHAP value, the CatBoost and XGBoost models had the same combination for the 14 features: dementia, age ≥85 years, age 74 - 85 years, gastrointestinal cancer, albumin <3.5 g/dL, open surgery, male sex, hemoglobin <11 g/dL, low income, nonhome admission, Barthel Index at admission ≤60, BMI <18.5 kg/m^2, diabetes, and stage III–IV (Figure 4). There were no significant interactions between stage III–IV and the top 5 influential features that both models shared (Table S4) in Multimedia Appendix 1.

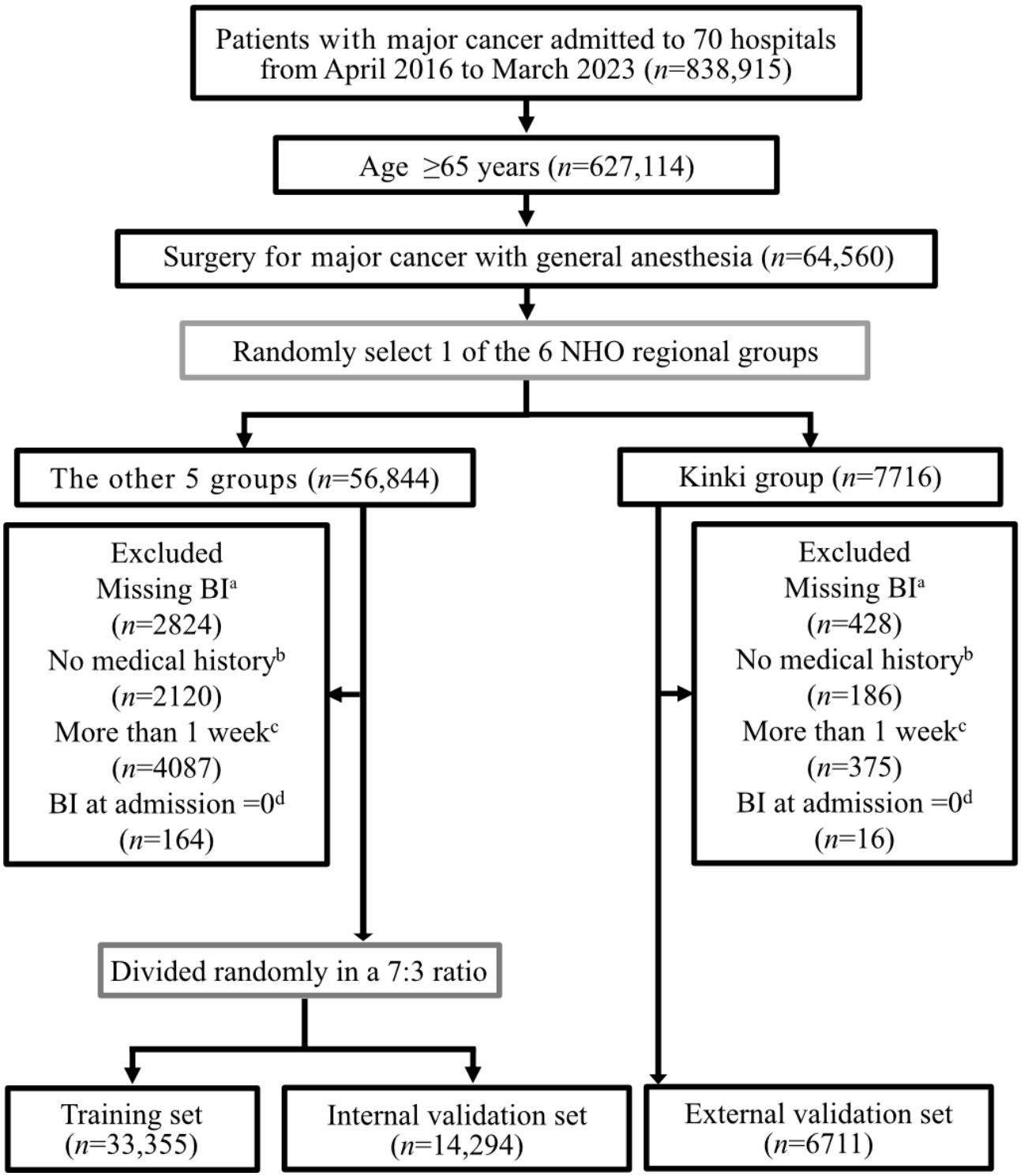
For both the internal and external validation set, the CatBoost model had the largest AUCs among the 6 machine- learning models: 0.77 (95% CI 0.75 - 0.79) and 0.72 (95% CI 0.68 - 0.75), respectively (Figure 5). In sensitivity analysis, all 6 models maintained comparable performance to the main analysis (Table S5 in Multimedia Appendix 1). The CatBoost model achieved relatively high AUCs and showed consistent performance in complete cases (0.78, 95% CI 0.76 - 0.80; 0.72, 0.68 - 0.76), death only (0.77, 0.71 - 0.82; 0.73, 0.65 - 0.81), and functional disability only (0.77, 0.75 - 0.79; 0.71, 0.68 - 0.75) for internal and external validation, respectively.

In subgroup analyses, the models maintained consistent performance for LOS and cancer staging (Table S5 in Multimedia Appendix 1). However, the performance varied based on cancer types: the CatBoost model achieved a larger AUC in patients with stomach cancer (internal: 0.80, 95% CI 0.76 - 0.84; external: 0.81, 95% CI 0.69 - 0.92) but a smaller AUC in patients with prostate cancer (0.53, 95% CI 0.40 - 0.66; 0.46, 95% CI 0.28 - 0.63) than the main analysis.

Based on the CatBoost model, patients at high risk had significantly longer LOS (internal: median 13, IQR 9 - 19 d vs median 9, IQR 7 - 13 d; external: median 13, IQR 10 - 19 d vs median 10.0, IQR 7.0 - 14.0 d) and higher total health care costs (internal: median US $11,048, IQR US $9191 - 13,106 d vs median US $10,092, IQR US $7894 - 11,893; external: median 11,069, IQR US $9401 - 13,499 vs median US $10,371, IQR US $8820 - 11,936) than patients at low risk (all P<.01). However, the high-risk group had slightly lower surgical procedure costs than the low-risk group in internal validation and was comparable to the low-risk group in external validation (Table 2).

XSL•FO
RenderX

Figure 2. Flow diagram of enrollment of study participants.(a) “Missing BI” means patients with missing Barthel Index at admission or discharge. (b) “No medical history” means patients included in the database within 8 weeks preceding surgery. (c) “More than 1 week” means patients who underwent surgery more than one week after admission. (d) “BI at admission=0” means patients with BI of 0 at admission. BI: Barthel Index; NHO: National Hospital Organization.



XSL•FO
RenderX

Table . Patient background with or without worse discharge in the training set.

Variable	No worse discharge (n=31,949)	Worse discharge (n=1406)	*P* value
Age (years), n (%)			<.01
65 - 74	18,010 (56.4)	405 (28.8)	
75 - 84	12,075 (37.8)	657 (46.7)	
≥85	1864 (5.8)	344 (24.5)	
Sex, n (%)			.05
Male	16,570 (51.9)	768 (54.6)	
Female	15,379 (48.1)	638 (45.4)	
BMI (kg/m^2), n (%)			
<18.5	2535 (7.9)	212 (15.1)	<.01
≥30	1170 (3.7)	57 (4.1)	.49
Route of admission, n (%)			
Home	31,668 (99.1)	1292 (91.9)	<.01
Nonhome	281 (0.9)	114 (8.1)	<.01
Nursing home	134 (0.4)	60 (4.3)	<.01
Other hospital	136 (0.4)	54 (3.8)	<.01
Others	11 (0)	0 (0)	1.00
Emergency admission, n (%)	252 (0.8)	42 (3)	<.01
Low income[a], n (%)	6020 (18.8)	352 (25)	<.01
Brinkman Index ≥200, n (%)	13,307 (41.7)	551 (39.2)	.07
Barthel Index[b]≤60, n (%)	422 (1.3)	139 (9.9)	<.01
Open surgery, n (%)	12,307 (38.5)	652 (46.4)	<.01
Scopic surgery, n (%)	19,642 (61.5)	754 (53.6)	.01
With epidural anesthesia[c], n (%)	15,742 (49.3)	769 (54.7)	<.01
Type of cancer, n (%)			<.01
Breast	6298 (19.7)	150 (10.7)	
Colorectal	7387 (23.1)	493 (35.1)	
Liver	1331 (4.2)	91 (6.5)	
Lung	9386 (29.4)	293 (20.8)	
Pancreas	786 (2.5)	62 (4.4)	
Prostate	2939 (9.2)	59 (4.2)	
Stomach	3822 (12)	258 (18.3)	
Gastrointestinal cancer	13,326 (41.7)	904 (64.3)	<.01
Cancer staging, n (%)			<.01
0-I	14,398 (45.1)	492 (35)	
II	9835 (30.8)	456 (32.4)	
III	4845 (15.2)	316 (22.5)	
IV	2871 (9)	142 (10.1)	
Stage III-IV	7716 (24.2)	458 (32.6)	<.01
Recurrent cancer, n (%)	2161 (6.8)	82 (5.8)	.19
Comorbidities, n (%)			

XSL•FO
RenderX

Variable	No worse discharge (n=31,949)	Worse discharge (n=1406)	*P* value
CCI[d] ≥3	4508 (14.1)	227 (16.1)	.04
HFRS[e] ≥5	270 (0.8)	55 (3.9)	<.01
Cerebrovascular disease	853 (2.7)	87 (6.2)	<.01
Chronic pulmonary disease	1832 (5.7)	95 (6.8)	.12
Congestive heart failure	823 (2.6)	72 (5.1)	<.01
Dementia	1783 (5.6)	403 (28.7)	<.01
Diabetes	4122 (12.9)	232 (16.5)	<.01
Liver disease	987 (3.1)	66 (4.7)	<.01
Myocardial infarction	223 (0.7)	17 (1.2)	.04
Peptic ulcer disease	1597 (5)	85 (6)	.09
Peripheral vascular disease	195 (0.6)	16 (1.1)	.02
Renal disease	341 (1.1)	27 (1.9)	<.01
Medical history within 8 weeks, n (%)			
Chemotherapy	2573 (8.1)	88 (6.3)	.02
Radiation	75 (0.2)	6 (0.4)	.25
Surgery	4220 (13.2)	215 (15.3)	.03
BT[f] ≥38˚C, n (%)	4212 (13.2)	190 (13.5)	.75
sBP[g] ≥180 mm Hg, n (%)	1898 (5.9)	126 (9)	<.01
Albumin <3.5 g/dL, n (%)	6669 (20.9)	543 (38.6)	<.01
T-Bil[h] ≥2.0 mg/dL, n (%)	252 (0.8)	14 (1)	.48
Creatinine ≥2.0 mg/dL, n (%)	639 (2)	65 (4.6)	<.01
Platelet $<10^5$/μL, n (%)	641 (2)	59 (4.2)	<.01
Hemoglobin <11 g/dL, n (%)	5896 (18.5)	523 (37.2)	<.01
Number of beds, n (%)			<.01
<300	1298 (4.1)	70 (5)	
300 - 499	17,922 (56.1)	707 (50.3)	
≥500	12,729 (39.8)	629 (44.7)	

[a]Low income" means an estimated household income in the lowest tertile based on ZIP code.

[b]Barthel Index ≤ 60" means the Barthel Index ≤ 60 at admission.

[c]With epidural anesthesia" means the combination of general and epidural anesthesia.

[d]CCI: Charlson Comorbidity Index.

[e]HFRS: Hospital Frailty Risk Score.

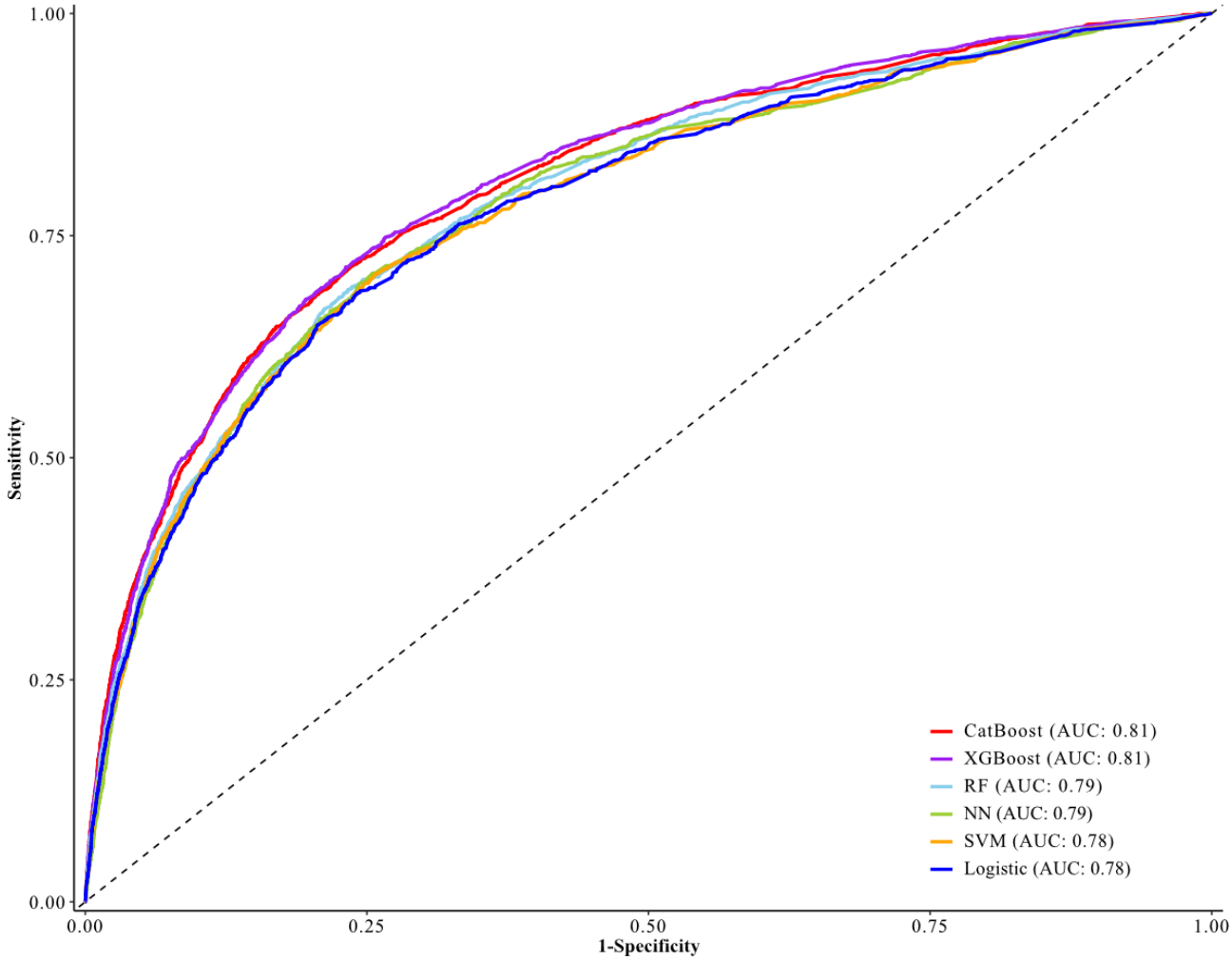
[f]BT: body temperature.

[g]sBP: systolic blood pressure.

[h]T-Bil: total bilirubin.

XSL•FO
RenderX

Figure 3. Receiver operating characteristic curves of 6 machine-learning models in the training set. AUC: area under the receiver operating characteristic curve; CatBoost: category boosting; Logistic: logistic regression; NN: neural networks; RF: random forest; SVM: support vector machine; XGBoost: extreme gradient boosting.



XSL•FO
RenderX

Figure 4. The top 15 features of predictor variables based on mean absolute Shapley additive explanations value. (**A**) The top 15 features selected by the category boosting model. (**B**) The top 15 features selected by the XGBoost: extreme gradient boosting model. (a) “Low income” means an estimated household income in the lowest tertile based on ZIP code. (b) “Barthel Index ≤60” means the Barthel Index ≤60 at admission. (c) “Chemotherapy” means that patients underwent chemotherapy within 8 weeks before surgery. (d) “Stage III-IV” means that patients had cancer staging III or IV using the TNM: Tumor Nodes Metastasis classification system.CatBoost: category boosting; SHAP: Shapley additive explanations; XGBoost: extreme gradient boosting.

XSL•FO
RenderX

Figure 5. Receiver operating characteristic curves of 6 machine learning models. (**A**) Receiver operating characteristic curves in the internal validation set. (**B**) Receiver operating characteristic curves in the external validation set. AUC: area under the receiver operating characteristic curve; CatBoost: category boosting; Logistic: logistic regression; NN: neural networks; RF: random forest; ROC: receiver operating characteristic; SVM: support vector machine; XGBoost: extreme gradient boosting.

XSL•FO
RenderX

Table . Adverse outcome between low-risk and high-risk group by category boosting model using the Youden index as cut-off value.

Variable	Internal validation set			External validation set		
	Low risk (n=10,788)	High risk (n=3506)	*P* value	Low risk (n=5459)	High risk (n=1252)	*P* value
Worse discharge, n (%)	221 (2)	386 (11)	<.01	116 (2.1)	113 (9)	<.01
Death, n (%)	25 (0.2)	33 (0.9)	<.01	16 (0.3)	18 (1.4)	<.01
Functional disability, n (%)[a]	196 (1.8)	353 (10.1)	<.01	100 (1.8)	95 (7.6)	<.01
LOS[b], median (IQR), days	9 (7-13)	13.0 (9-19)	<.01	10 (7-14)	13 (10-19)	<.01
Cost, median (IQR), US $						
Total	10,092 (7894-11,893)	11,048 (9191-13,106)	<.01	10,371 (8820-11,936)	11,069 (9401-13,499)	<.01
Medical consultation	77 (49-114)	97 (65-141)	<.01	75 (48-110)	97 (70-137)	<.01
Medication	68 (33, 139)	176 (87-362)	<.01	67 (34-139)	146 (76-333)	<.01
Medical procedure	26 (15-42)	40 (24-67)	<.01	27 (14-44)	28 (13-57)	<.01
Surgical procedure	6645 (5028-8253)	6620 (5513-8116)	.02	7203 (5825-8176)	6924 (5900-8212)	.46
Laboratory tests	449 (328-597)	516 (370-742)	<.01	443 (328-610)	526 (371-775)	<.01
Hospital stays	2404 (1873-3021)	3085 (2362-4073)	<.01	2484 (1940-3078)	3039 (2419-4178)	<.01
Others	0 (0-137)	100 (0-212)	<.01	0 (0-120)	0 (0-177)	<.01

[a]Functional disability" means a decrease in the Barthel Index by ≥5 points at discharge compared with admission.

[b]LOS: length of stay.

Discussion

Principal Findings

In this study, we developed and validated machine-learning models to predict postoperative functional disability and mortality in older patients with cancer. Our CatBoost model achieved good performance using routinely available preoperative factors from electronic health records, indicating the potential for clinical implementation. Although ethical training for hospital staff is essential to prevent unauthorized disclosure of prediction results, implementing this model within closed electronic health record systems could provide protection for patient privacy.

The previous model for lower-extremity surgery had an AUC of 0.72 in external validation [6], similar to our model; however, our model directly predicted functional disability using the Barthel Index rather than using nursing home discharge as a surrogate. The model performance remained consistent across sensitivity analyses for death and functional disability separately, and complete cases, indicating the robustness of our findings. Notably, the model demonstrated higher predictive performance in patients with stomach or colorectal cancer than the other cancers, making it especially valuable for surgical decision-making in these patients.

The CatBoost model identified patients at high risk who had longer LOS and higher health care costs; however, surgical procedure costs were comparable between patients at high and low risk. These findings suggested that the increased cost was based on the varying postoperative course. Our model can support decision-making for older patients with cancer and their families regarding cancer surgery by providing insights into potential postoperative QOL and care burden. Moreover, if patients at high risk choose to undergo cancer surgery, our model may enable health care providers to implement targeted interventions such as intensive postoperative rehabilitation. Early identification of patients at high risk, such as those aged ≥85 years with dementia, can help health care providers prepare support systems, including caregiver education, social work consultation for home health support, and coordination with multidisciplinary teams [15]. This proactive approach may help reduce caregiver burden and improve outcomes for both patients and their families.

Of the 6 machine-learning models, the CatBoost and XGBoost models, with AUC ≥0.80, had the same combination of 14 features in the top 15 influential factors. These factors include established risk factors for poor postoperative outcomes in older patients as identified in previous studies: dementia [7], older age [9] (≥85 y and 75 - 84 y), male sex [24], anemia [10] (hemoglobin <11 g/dL), low income [6], underweight [12] (BMI <18.5 kg/m²), diabetes [49], and cancer staging [25]. In addition,

XSL•FO
RenderX

several factors serve as proxies for known risk factors. For frailty [5], the factors include (1) open surgery, which generally results in a more pronounced postoperative functional disability compared with scopic surgeries; (2) nonhome admission, likely indicating that patients are too frail to live independently; and (3) Barthel Index ≤60 at admission, indicating severe dependence [22]. For malnutrition [13], the proxies are (1) albumin <3.5 g/dL, a marker of malnutrition, and (2) gastrointestinal cancer, which often involves a long time to restart food intake after surgery, increasing the risk of malnutrition compared with other cancer types. The consistency between the 2 models in identifying these factors further validates their importance in predicting postoperative outcomes. Although chemotherapy and creatinine ≥2.0 mg/dL were not common in both models, these factors, identified as influential factors in previous studies [33,50], might also be included as predictor variables in future models. Moreover, our models included several factors associated with social vulnerability, such as age ≥85 years, dementia, and low income. Without ensuring model transparency to health care providers, our model could unconsciously contribute to reduced surgical care access for vulnerable populations. Therefore, when implementing our model in clinical practice, health care providers should consider these model characteristics to ensure fair allocation of health care resources [51].

Limitations

Our study has some limitations. First, we only validated all models using Japanese data. While our CatBoost model showed moderate accuracy (AUC: 0.7 - 0.9) [52] in both internal and external validation, the AUC in the external validation was lower than that in the internal validation as observed in a previous study [6]. Further studies using data from other countries and ethnic groups are necessary to evaluate model robustness, including potential bias and overfitting [51], and confirm their applicability to different health care systems. However, our study used data from 70 hospitals across Japan, which may enhance generalizability within the country. In addition, considering the global trend of population aging, our models may prove valuable for other countries in the future, particularly when these countries reach levels of demographic aging similar to Japan's current situation.

Second, we did not have information on predictor variables such as marital status [6] because of the retrospective nature of the study. Despite this limitation, our models had good predictive performance in the validation sets. While our analysis showed no significant interactions between stage III–IV cancer and the top features, future studies incorporating additional variables may evaluate such interactions to enhance the predictive performance of models.

Finally, the long-term prognosis of patients classified as high-risk by our models remains unclear. Further research is required to determine the extent of functional recovery and mortality in these patients. At a minimum, postoperative functional disability in patients at high risk indicates an increased immediate post-discharge burden on family caregivers and health care resources.

Conclusions

Our CatBoost model achieved good performance for predicting postoperative functional disability and mortality in older patients with cancer. This model could support surgical decision-making for patients and families while guiding targeted interventions by health care providers. This model, which is based on routinely available preoperative factors, has the potential for implementation in clinical settings through electronic health records.

Acknowledgments

We thank Masato Koizumi of the Information Technology Department in the National Hospital Organization Headquarters for his advice on data handling. This research was partially funded by the Daiwa Securities Foundation, JSPS KAKENHI (grant number 23K19882), and Pfizer Health Research Foundation.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to being sensitive personal information. The analytic code is available from the corresponding author on request.

Authors' Contributions

YH, NI, TT, and SI contributed to conception and design and interpretation of data. YH assisted with drafting of the manuscript. YH, NI, TT, and SI handled critical review of the manuscript for important intellectual content. YH and NI contributed to statistical analysis. YH obtained funding. TT and SI assisted with supervision.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Table S1. ICD-10 codes selected as predictor variables. Table S2. Crude odds ratios of predictor variables for worse discharge. Table S3. Performance metrics of six machine learning models in training set. Table S4. Interaction between stage III–IV and top 5 features of predictor variables based on mean absolute SHAP value in training set. Table S5. AUCs of six machine learning

XSL•FO
RenderX

models for internal and external validation set in sensitivity and subgroup analyses. Figure S1. Precision-recall curve of six machine learning models in training set.
[DOCX File, 185 KB - aging_v8i1e65898_app1.docx]

References

1. Global cancer burden growing, amidst mounting need for services. World Health Organization. 2024. URL: https://www.who.int/news/item/01-02-2024-global-cancer-burden-growing--amidst-mounting-need-for-services [accessed 2025-05-06]
2. Cancer statistics. National Cancer Center, Center for Cancer Control and Information Services. 2019. URL: https://www.ncc.go.jp/en/cis/index.html [accessed 2025-05-06]
3. World population prospects. United Nations Department of Economic and Social Affairs Population Division. 2019. URL: https://population.un.org/wpp/assets/Files/WPP2019_Highlights.pdf [accessed 2025-05-06]
4. Cabinet Office Japan. Annual report on the ageing society [summary]. 2021 Jul URL: https://www8.cao.go.jp/kourei/english/annualreport/2021/pdf/2021#:~:text=population%20aged%2075%20years%20and,65%2D74%20years%20in%20size.&text=By%202065%2C%20one%20in%202.6,75%20years%20old%20and%20over [accessed 2025-05-06]
5. Shaw JF, Mulpuru S, Kendzerska T, et al. Association between frailty and patient outcomes after cancer surgery: a population-based cohort study. Br J Anaesth 2022 Mar;128(3):457-464. [doi: 10.1016/j.bja.2021.11.035] [Medline: 35034792]
6. Schaefer MS, Hammer M, Platzbecker K, et al. What factors predict adverse discharge disposition in patients older than 60 years undergoing lower-extremity surgery? The adverse discharge in older patients after lower-extremity surgery (ADELES) risk score. Clin Orthop Relat Res 2021 Mar 1;479(3):546-547. [doi: 10.1097/CORR.0000000000001532] [Medline: 33196587]
7. Masutani R, Pawar A, Lee H, Weissman JS, Kim DH. Outcomes of common major surgical procedures in older adults with and without dementia. JAMA Netw Open 2020 Jul 1;3(7):e2010395. [doi: 10.1001/jamanetworkopen.2020.10395] [Medline: 32662842]
8. Morishima T, Kuwabara Y, Saito MK, et al. Patterns of staging, treatment, and mortality in gastric, colorectal, and lung cancer among older adults with and without preexisting dementia: a Japanese multicentre cohort study. BMC Cancer 2023 Jan 19;23(1):67. [doi: 10.1186/s12885-022-10411-y] [Medline: 36658524]
9. Konishi T, Sasabuchi Y, Matsui H, Tanabe M, Seto Y, Yasunaga H. Long-term risk of being bedridden in elderly patients who underwent oncologic surgery: a retrospective study using a Japanese claims database. Ann Surg Oncol 2023 Aug;30(8):4604-4612. [doi: 10.1245/s10434-023-13566-5] [Medline: 37149549]
10. Boening A, Boedeker RH, Scheibelhut C, Rietzschel J, Roth P, Schönburg M. Anemia before coronary artery bypass surgery as additional risk factor increases the perioperative risk. Ann Thorac Surg 2011 Sep;92(3):805-810. [doi: 10.1016/j.athoracsur.2011.02.076] [Medline: 21592459]
11. Schaap LA, Koster A, Visser M. Adiposity, muscle mass, and muscle strength in relation to functional decline in older persons. Epidemiol Rev 2013;35(35):51-65. [doi: 10.1093/epirev/mxs006] [Medline: 23221972]
12. Roh L, Braun J, Chiolero A, et al. Mortality risk associated with underweight: a census-linked cohort of 31,578 individuals with up to 32 years of follow-up. BMC Public Health 2014 Apr 16;14(14):371. [doi: 10.1186/1471-2458-14-371] [Medline: 24739374]
13. Kim SW, Han HS, Jung HW, et al. Multidimensional frailty score for the prediction of postoperative mortality risk. JAMA Surg 2014 Jul;149(7):633-640. [doi: 10.1001/jamasurg.2014.241] [Medline: 24804971]
14. Minami T, Fujita K, Hashimoto M, et al. External beam radiotherapy combination is a risk factor for bladder cancer in patients with prostate cancer treated with brachytherapy. World J Urol 2023 May;41(5):1317-1321. [doi: 10.1007/s00345-023-04380-5] [Medline: 37024557]
15. Huang LW, Smith AK, Wong ML. Who will care for the caregivers? Increased needs when caring for frail older adults with cancer. J Am Geriatr Soc 2019 May;67(5):873-876. [doi: 10.1111/jgs.15863] [Medline: 30924526]
16. Loyd C, Markland AD, Zhang Y, et al. Prevalence of hospital-associated disability in older adults: a meta-analysis. J Am Med Dir Assoc 2020 Apr;21(4):455-461. [doi: 10.1016/j.jamda.2019.09.015] [Medline: 31734122]
17. Ogawa M, Yoshida N, Nakai M, et al. Hospital-associated disability and hospitalization costs for acute heart failure stratified by body mass index- insight from the JROAD/JROAD-DPC database. Int J Cardiol 2022 Nov 15;367(367):38-44. [doi: 10.1016/j.ijcard.2022.08.044] [Medline: 36029847]
18. Kanazawa N, Tani T, Imai S, Horiguchi H, Fushimi K, Inoue N. Existing data sources for clinical epidemiology: database of the National Hospital Organization in Japan. Clin Epidemiol 2022;14(14):689-698. [doi: 10.2147/CLEP.S359072] [Medline: 35615723]
19. Hayashida K, Murakami G, Matsuda S, Fushimi K. History and profile of diagnosis procedure combination (DPC): development of a real data collection system for acute inpatient care in Japan. J Epidemiol 2021 Jan 5;31(1):1-11. [doi: 10.2188/jea.JE20200288] [Medline: 33012777]
20. Yamana H, Moriwaki M, Horiguchi H, Kodan M, Fushimi K, Yasunaga H. Validity of diagnoses, procedures, and laboratory data in Japanese administrative data. J Epidemiol 2017 Oct;27(10):476-482. [doi: 10.1016/j.je.2016.09.009] [Medline: 28142051]

XSL•FO
RenderX

21. Cancer statistics, annual report. National Cancer Center Japan. 2022. URL: https://www.ncc.go.jp/en/publication_report/2022/ncphs/ncphs14.html [accessed 2025-05-06]
22. Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. J Clin Epidemiol 1989;42(8):703-709. [doi: 10.1016/0895-4356(89)90065-6] [Medline: 2760661]
23. Youden WJ. Index for rating diagnostic tests. Cancer 1950;3(1):32-35. [doi: 10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3] [Medline: 15405679]
24. Losurdo P, Mastronardi M, de Manzini N, Bortul M. Survival and long-term surgical outcomes after colorectal surgery: are there any gender-related differences? Updates Surg 2022 Aug;74(4):1337-1343. [doi: 10.1007/s13304-022-01323-4] [Medline: 35810269]
25. Bliton J, Parides M, Muscarella P, McAuliffe JC, Papalezova K, In H. Clinical stage of cancer affects perioperative mortality for gastrointestinal cancer surgeries. J Surg Res 2021 Apr;260(1-9):1-9. [doi: 10.1016/j.jss.2020.11.023] [Medline: 33310353]
26. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40(5):373-383. [doi: 10.1016/0021-9681(87)90171-8] [Medline: 3558716]
27. Birim O, Kappetein AP, Bogers AJJC. Charlson comorbidity index as a predictor of long-term outcome after surgery for nonsmall cell lung cancer. Eur J Cardiothorac Surg 2005 Nov;28(5):759-762. [doi: 10.1016/j.ejcts.2005.06.046] [Medline: 16157485]
28. Gilbert T, Neuburger J, Kraindler J, et al. Development and validation of a hospital frailty risk score focusing on older people in acute care settings using electronic hospital records: an observational study. Lancet 2018 May 5;391(10132):1775-1782. [doi: 10.1016/S0140-6736(18)30668-8] [Medline: 29706364]
29. Cats A, Jansen EPM, van Grieken NCT, et al. Chemotherapy versus chemoradiotherapy after surgery and preoperative chemotherapy for resectable gastric cancer (CRITICS): an international, open-label, randomised phase 3 trial. Lancet Oncol 2018 May;19(5):616-628. [doi: 10.1016/S1470-2045(18)30132-3] [Medline: 29650363]
30. Erlandsson J, Holm T, Pettersson D, et al. Optimal fractionation of preoperative radiotherapy and timing to surgery for rectal cancer (Stockholm III): a multicentre, randomised, non-blinded, phase 3, non-inferiority trial. Lancet Oncol 2017 Mar;18(3):336-346. [doi: 10.1016/S1470-2045(17)30086-4] [Medline: 28190762]
31. Kovacs C, Jarvis SW, Prytherch DR, et al. Comparison of the National Early Warning Score in non-elective medical and surgical patients. Br J Surg 2016 Sep;103(10):1385-1393. [doi: 10.1002/bjs.10267] [Medline: 27487317]
32. Hartle A, McCormack T, Carlisle J, et al. The measurement of adult blood pressure and management of hypertension before elective surgery: Joint Guidelines from the Association of Anaesthetists of Great Britain and Ireland and the British Hypertension Society. Anaesthesia 2016 Mar;71(3):326-337. [doi: 10.1111/anae.13348] [Medline: 26776052]
33. Park JH, Lee HJ, Oh SY, et al. Prediction of postoperative mortality in patients with organ failure after gastric cancer surgery. World J Surg 2020 May;44(5):1569-1577. [doi: 10.1007/s00268-020-05382-9] [Medline: 31993720]
34. Brierley JD, Gospodarowicz MK, Wittekind C. TNM Classification of Malignant Tumours: John Wiley & Sons; 2017. [doi: 10.1002/9780471420194]
35. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care 2005 Nov;43(11):1130-1139. [doi: 10.1097/01.mlr.0000182534.19832.83] [Medline: 16224307]
36. Liu C, Zhang K, Yang X, et al. Development and validation of an explainable machine learning model for predicting myocardial injury after noncardiac surgery in two centers in China: Retrospective Study. JMIR Aging 2024 Jul 26;7:e54872. [doi: 10.2196/54872] [Medline: 39087583]
37. Guan C, Ma F, Chang S, Zhang J. Interpretable machine learning models for predicting venous thromboembolism in the intensive care unit: an analysis based on data from 207 centers. Crit Care 2023 Oct 24;27(1):406. [doi: 10.1186/s13054-023-04683-4] [Medline: 37875995]
38. Mayer M. Package missranger: fast imputation of missing values. The Comprehensive R Archive Network. 2021 Mar 30. URL: https://cran.r-project.org/web/packages/missRanger/missRanger.pdf [accessed 2025-05-06]
39. Stekhoven DJ, Bühlmann P. MissForest--non-parametric missing value imputation for mixed-type data. Bioinformatics 2012 Jan 1;28(1):112-118. [doi: 10.1093/bioinformatics/btr597] [Medline: 22039212]
40. Li Y, Salmasian H, Vilar S, Chase H, Friedman C, Wei Y. A method for controlling complex confounding effects in the detection of adverse drug reactions using electronic health records. J Am Med Inform Assoc 2014;21(2):308-314. [doi: 10.1136/amiajnl-2013-001718] [Medline: 23907285]
41. Chen Q, Mao R, Zhao J, et al. Nomograms incorporating preoperative RDW level for the prediction of postoperative complications and survival in colorectal liver metastases after resection. Ann Palliat Med 2021 Apr;10(4):4143-4158. [doi: 10.21037/apm-20-2418] [Medline: 33832316]
42. Konishi T, Goto T, Fujiogi M, et al. New machine learning scoring system for predicting postoperative mortality in gastroduodenal ulcer perforation: A study using A Japanese nationwide inpatient database. Surgery 2022 Apr;171(4):1036-1042. [doi: 10.1016/j.surg.2021.08.031] [Medline: 34538648]
43. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A. CatBoost: unbiased boosting with categorical features. arXiv. Preprint posted online on Jun 28, 2017. [doi: 10.48550/arXiv.1706.09516]

XSL•FO
RenderX

44. Tu JV. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. J Clin Epidemiol 1996 Nov;49(11):1225-1231. [doi: 10.1016/s0895-4356(96)00002-9] [Medline: 8892489]
45. Oh T, Kim D, Lee S, et al. Machine learning-based diagnosis and risk factor analysis of cardiocerebrovascular disease based on KNHANES. Sci Rep 2022 Feb;12(1):35145205. [doi: 10.1038/s41598-022-06333-1]
46. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Presented at: NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems; Dec 4-9, 2017; Long Beach, CA. [doi: 10.48550/arXiv.1705.07874]
47. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. Biometrics 1988 Sep;44(3):837-845. [doi: 10.2307/2531595] [Medline: 3203132]
48. Bur AM, Brant JA, Newman JG, et al. Incidence and risk factors for prolonged hospitalization and readmission after transsphenoidal pituitary surgery. Otolaryngol Head Neck Surg 2016 Oct;155(4):688-694. [doi: 10.1177/0194599816652379] [Medline: 27221574]
49. Lopez LF, Reaven PD, Harman SM. Review: The relationship of hemoglobin A1c to postoperative surgical risk with an emphasis on joint replacement surgery. J Diabetes Complications 2017 Dec;31(12):1710-1718. [doi: 10.1016/j.jdiacomp.2017.08.016] [Medline: 29029935]
50. von Waldenfels G, Loibl S, Furlanetto J, et al. Outcome after neoadjuvant chemotherapy in elderly breast cancer patients - a pooled analysis of individual patient data from eight prospectively randomized controlled trials. Oncotarget 2018 Mar 16;9(20):15168-15179. [doi: 10.18632/oncotarget.24586] [Medline: 29632634]
51. Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring fairness in machine learning to advance health equity. Ann Intern Med 2018 Dec 18;169(12):866-872. [doi: 10.7326/M18-1990] [Medline: 30508424]
52. Fischer JE, Bachmann LM, Jaeschke R. A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. Intensive Care Med 2003 Jul;29(7):1043-1051. [doi: 10.1007/s00134-003-1761-8] [Medline: 12734652]

Abbreviations

AUC: area under the receiver operating characteristic curve
CatBoost: category boosting
CCI: Charlson Comorbidity Index
HFRS: Hospital Frailty Risk Score
***ICD-10*:** *International Statistical Classification of Diseases and Related Health Problems, 10th Revision*
Lasso: least absolute shrinkage and selection operator
LOS: length of stay
NHO: National Hospital Organization
OR: odds ratio
PRAUC: area under the precision-recall curve
QOL: quality of life
ROC: receiver operating characteristic
SHAP: Shapley additive explanations
TNM: Tumor Node Metastasis
TRIPOD: Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis
XGBoost: extreme gradient boosting

Edited by M Mardini; submitted 29.08.24; peer-reviewed by H Li, KE Cain, YG Luo; revised version received 18.03.25; accepted 18.03.25; published 14.05.25.

Please cite as:
Hashimoto Y, Inoue N, Tani T, Imai S
Machine Learning for Predicting Postoperative Functional Disability and Mortality Among Older Patients With Cancer: Retrospective Cohort Study
JMIR Aging 2025;8:e65898
URL: https://aging.jmir.org/2025/1/e65898
doi:10.2196/65898

© Yuki Hashimoto, Norihiko Inoue, Takuaki Tani, Shinobu Imai. Originally published in JMIR Aging (https://aging.jmir.org), 14.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium,

XSL•FO
RenderX

provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Unsupervised Deep Learning of Electronic Health Records to Characterize Heterogeneity Across Alzheimer Disease and Related Dementias: Cross-Sectional Study

Matthew West[1*], MPhys, SM; You Cheng[1*], PhD; Yingnan He[1*], MS; Yu Leng[1], SM; Colin Magdamo[1], BS; Bradley T Hyman[1], MD, PhD; John R Dickson[1], MD, PhD; Alberto Serrano-Pozo[1], MD, PhD; Deborah Blacker[1], MD, PhD; Sudeshna Das[1], PhD

Massachusetts General Hospital, Cambridge, MA, United States

[*]these authors contributed equally

Corresponding Author:
Sudeshna Das, PhD
Massachusetts General Hospital
65 Landsdowne Street
Cambridge, MA, 02139
United States
Phone: 1 617 768 8254
Email: sdas5@mgh.harvard.edu

Abstract

Background: Alzheimer disease and related dementias (ADRD) exhibit prominent heterogeneity. Identifying clinically meaningful ADRD subtypes is essential for tailoring treatments to specific patient phenotypes.

Objective: We aimed to use unsupervised learning techniques on electronic health records (EHRs) from memory clinic patients to identify ADRD subtypes.

Methods: We used pretrained embeddings of non-ADRD diagnosis codes (*International Classification of Diseases, Ninth Revision*) and large language model (LLM)–derived embeddings of clinical notes from patient EHRs. Hierarchical clustering of these embeddings was used to identify ADRD subtypes. Clusters were characterized regarding their demographic and clinical features.

Results: We analyzed a cohort of 3454 patients with ADRD from a memory clinic at Massachusetts General Hospital, each with a specialist diagnosis. Clustering pretrained embeddings of the non-ADRD diagnosis codes in patient EHRs revealed the following 3 patient subtypes: one with skin conditions, another with psychiatric disorders and an earlier age of onset, and a third with diabetes complications. Similarly, using LLM-derived embeddings of clinical notes, we identified 3 subtypes of patients as follows: one with psychiatric manifestations and higher prevalence of female participants (prevalence ratio: 1.59), another with cardiovascular and motor problems and higher prevalence of male participants (prevalence ratio: 1.75), and a third one with geriatric health disorders. Notably, we observed significant overlap between clusters from both data modalities (χ^2_4=89.4; P<.001).

Conclusions: By integrating *International Classification of Diseases, Ninth Revision* codes and LLM-derived embeddings, our analysis delineated 2 distinct ADRD subtypes with sex-specific comorbid and clinical presentations, offering insights for potential precision medicine approaches.

(JMIR Aging 2025;8:e65178) doi:10.2196/65178

KEYWORDS

Alzheimer disease and related dementias; electronic health records; large language models; clustering; unsupervised learning

XSL•FO
RenderX

Introduction

Background

Alzheimer disease (AD) is a neurodegenerative condition which affects more than 55 million people globally [1], and it is the seventh leading cause of death in the United States [2]. Despite its substantial public health burden, AD remains poorly understood, with limited treatment options available. AD and related dementias (ADRD) is an umbrella term that refers to multiple dementing illnesses, including AD, frontotemporal dementia (FTD), Lewy body dementia (LBD), and vascular dementia. AD is the most prevalent, accounting for around 60% to 80% of all dementia [2]. While these diseases have distinct clinical and neuropathological criteria, there is substantial overlap in both clinical presentation and autopsy findings at the individual patient level. For example, AD is clinically characterized by an amnestic-predominant dementia and neuropathologically defined by the build-up of amyloid beta (Aβ) plaques and neurofibrillary tangles formed by hyperphosphorylated tau protein [3]; however, these lesions are frequently accompanied by cerebrovascular disease (CVD) [4] or Lewy body pathology [5], which can influence clinical presentation. Likewise, LBD is defined by Lewy bodies but is also associated with plaques and tangles [6], which may accelerate the rate of cognitive decline [7]. This clinical and neuropathological heterogeneity limits our ability to target disease-modifying drugs to each specific neuropathological lesion. The so-called *amyloid hypothesis* has prevailed as the leading explanation of AD disease etiology, where it is held that Aβ toxicity leads to tau hyperphosphorylation, synaptic dysfunction, and neurodegeneration [8]. However, treatments targeting this hypothesis only show limited efficacy, which may stem in part from the clinical and neuropathological comorbidities [9], highlighting the need for a tailored approach to identify potential subtypes of disease and develop more effective targeted treatments.

Previous approaches to AD subtyping have focused on RNA expression, as well as brain imaging and cognitive assessments. Neff et al [10] identified 5 molecular subtypes of AD using RNA-sequencing signatures, characterized by different dysregulated pathways related to tau-mediated neurodegeneration, Aβ neuroinflammation, synaptic signaling, immune activity, mitochondria organization, and myelination. The Subtype and Stage Inference algorithm, applied to magnetic resonance imaging and positron emission tomography imaging data, identified distinct AD trajectories based on the rate and sequence of brain atrophy [11] and tau deposition [12]. Cognitive subtypes have also been identified based on memory, visuospatial and linguistic capabilities, and executive function [13-15]. These studies were limited to research cohorts with specific selection criteria, and it is unclear whether these subtypes can be extended to larger samples.

In contrast, real-world data, such as, electronic health records (EHRs), provide readily accessible large observational datasets and have been used for clustering AD or ADRD subtypes [16]. Unsupervised learning approaches on EHR datasets have revealed latent structure in conditions, such as autism [17,18] and Parkinson disease [19]. For AD subtyping, EHR-based approaches have used the *International Classification of Diseases* (*ICD*) or similar diagnostic codes, showing varying success depending on the methodology and population. Xu et al [20] used hierarchical clustering on EHR data from patients with AD, identifying subtypes related to CVD, mental illness, age of onset, and sensory problems. Alexander et al [21] found the following 5 patient subtypes: mental health, nontypical AD, typical AD, CVD, and men with cancer. They later identified a consistent subtype with early-onset AD, predominantly female participants, with a faster rate of progression using various machine learning methods [22]. Landi et al [23] used unsupervised deep learning to encode EHRs with temporal information, identifying early-onset AD, later-onset AD with mild comorbidities, and typical-onset AD with moderate symptoms. He et al [24] applied spectral clustering to EHRs of patients with AD, discerning 4 subtypes with significant demographic, mortality, and medication use differences. Tang et al [25] analyzed comorbidity patterns in EHRs of patients with AD, revealing sex-dependent variations. In another study, Tang et al [26] used EHRs with knowledge networks to predict AD onset and identify sex-specific genetic markers. These studies collectively highlight the varied methodologies and results in EHR-based research, emphasizing the complexity and potential of these approaches for a deeper understanding of AD.

However, none of these prior studies leveraged the richer representation of EHR data by embedding full sequences of clinical text. Transformers have emerged as state-of-the-art architecture for language modeling and are broadly characterized by the concept of attention [27]. Attention, named for its similarity to cognitive attention, enables the sharing of contextual information among word representations without directly encoding their sequence. The transformer used in this work is a version of the Bidirectional Encoder Representations from Transformers (BERT) architecture [28]. This architecture consists of an encoder which can be fine-tuned on downstream applications and domains. Specifically, we use Clinical BERT [29], which is pretrained on a large corpus of clinical notes from the critical care database Medical Information Mart for Intensive Care (MIMIC) [30].

Objectives

In this work, we used both pretrained embeddings of *ICD-9* code diagnostic data and transformer-derived embeddings of clinical notes. This dual approach addresses the limitations of previous studies by incorporating structured *ICD* codes, which allow us to study subtypes of patients with similar *ICD* codes (non-ADRD diagnosis in charts), and unstructured clinical notes, which capture detailed clinical history and manifestations provided by specialists. By combining these 2 modalities, we aimed to enhance the clustering of patient ADRD subtypes.

Methods

Cohort Selection Process

Patients were selected from the Massachusetts General Hospital (MGH) EHR database. The selection criteria included patients who had at least 2 MGH memory clinic visits (either an in-person office visit or a video telemedicine visit) from August

XSL•FO
RenderX

2015 to June 2022, were aged >50 years at their first visit, and had progress notes of substantial length (≥512 characters). These criteria were chosen due to the richness of the notes for the clustering analysis and the high quality of the ADRD diagnosis from specialists. From the identified patient cohort, 2 datasets were extracted as follows: one containing structured diagnostic *ICD* code data from the patients' entire medical history and another consisting of unstructured clinical notes authored by memory clinic specialists, limited to the most recent visit. We chose only the most recent visit note because it typically consolidates the patient's prior history, thereby reducing redundancy and providing a focused, up-to-date clinical snapshot. In addition, the dataset was filtered to exclude patients who did not have ADRD diagnoses (Multimedia Appendix 1), as well as those who lacked non-ADRD *ICD* codes (ie, patients who only had ADRD *ICD* codes were excluded).

Ethical Considerations

This study was approved by the Mass General Brigham Institutional Review Board (protocol 2015P001915), which granted a waiver of informed consent for secondary analysis of electronic health data. No participant compensation was provided. Electronic health data was queried from Epic and securely stored on servers within the Mass General Brigham firewall. Access was restricted to authorized study personnel, in full compliance with institutional privacy and data security policies.

Embedding Methodology

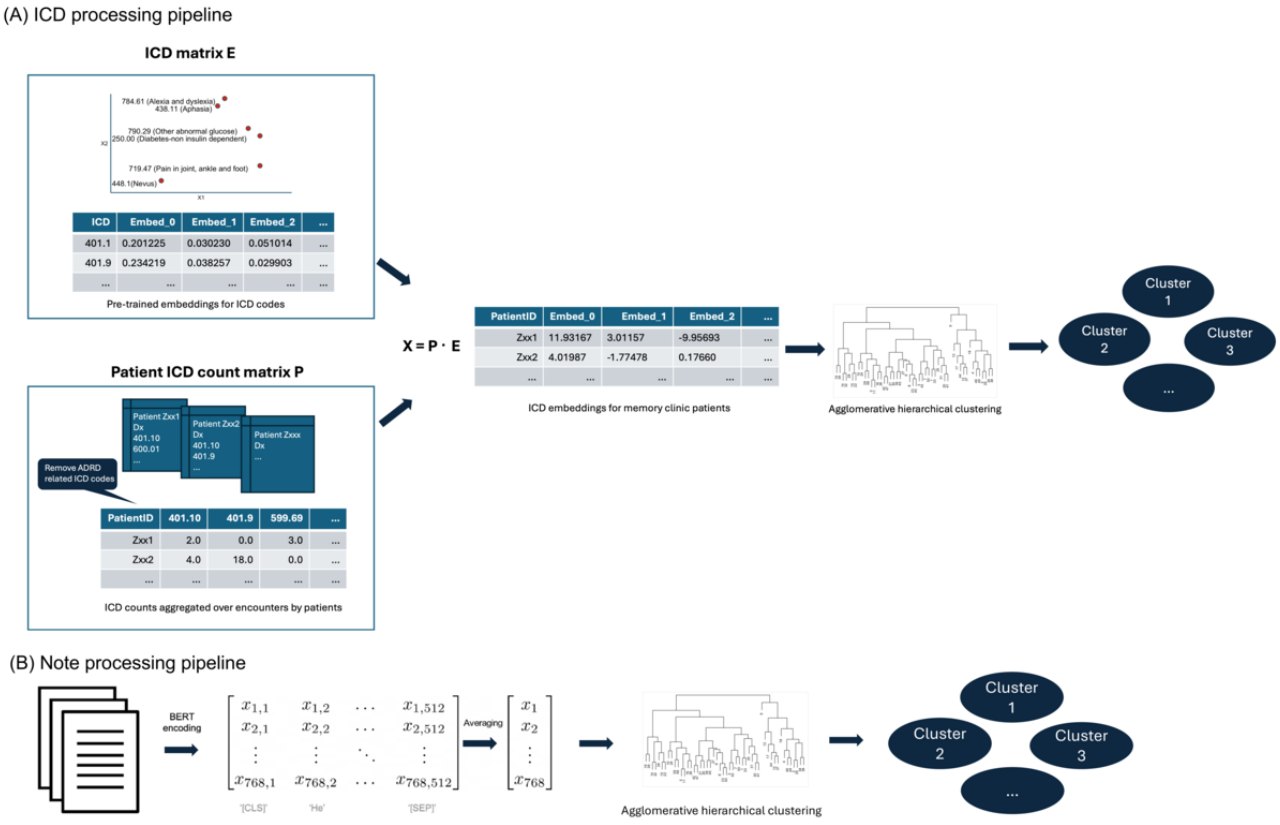
ICD Codes

Before clustering, it was necessary to derive a patient-level representation that encoded information relevant to phenotype in a single vector. While some prior work has relied on one-hot encoding (where categorical data are converted into binary vectors) of clinical data to represent patient phenotype, we leverage existing pretrained embeddings that capture relevant biomedical semantics in their latent representations of clinical concepts. In particular, we use a set of 300-dimensional embeddings for *ICD-9* codes, derived from prior work by Choi et al [31].

For a count-based encoded representation of *m ICD-9* codes across our cohort, $P \in R^{3454 \times m}$, and an embedding matrix, $E \in R^{m \times 300}$, our design matrix for clustering, $X \in R^{3454 \times 300}$, is given by the following matrix multiplication: $X = P \cdot E$.

This matrix multiplication sums the non-ADRD *ICD* embeddings across a patient record, and the resultant embedding is directly affected by the number of times each code appears in a patient's history. ADRD codes were dropped from the matrix *P* to not confound clustering based on structured ADRD phenotype. A schematic depicting the *ICD* representation pipeline is provided in Figure 1A.

Figure 1. Visualization of the clustering pipeline for (A) International Classification of Diseases (ICD) codes and (B) notes. For each subfigure, the workflow goes from left to right. BERT: Bidirectional Encoder Representations from Transformers.



XSL•FO
RenderX

Clinical Notes

Clinical notes were encoded using Clinical BERT before clustering. To derive patient-level representations of clinic notes, several preprocessing steps were undertaken. First, unwanted delimiter characters were stripped from patient notes, and notes were chunked into contiguous sections of up to 1024 characters. This resulted in a distribution of token numbers of ~200 to 300 per note following BERT's WordPiece encoding of input sequences. After passing through the transformer encoding, we took the final layer representation averaged over the 12 attention heads, such that each note was represented by a matrix of dimension (768, n), corresponding to each of the n input tokens having a 768-dimensional contextual vector representing it. Following this encoding, the representation was averaged over the token dimension to arrive at a single 768-dimensional vector for the whole note. This was explored using both simple averaging (arithmetic mean) over the token dimension, as well as attention-weighted averaging based on row-wise entropy of the final layer attention matrix. Attention-weighted averaging was used in patient representations due to the resultant lower inertia and increased silhouette score on average. A schematic depicting the note representation pipeline is provided in Figure 1B.

Attention-Weighted Averaging

For a given input sequence of length n, the final layer representation in a transformer model has an associated self-attention matrix, $A \in R^{n \times n}$. For BERT-based models, n is ≤512 due to the constraint on the length of the context window. To provide weights for averaging the embedding, $E \in R^{768 \times n}$, over the token dimension, we compute the row-wise differential entropy of the attention matrix, given in equation (1). Differential entropy is the continuous analog of Shannon entropy, which is usually only defined for discrete random variables [32]. In particular, the method described in Ebrahimi et al [33] is used to approximate the differential entropy, implemented in the Python (Python Software Foundation) library *SciPy version 1.7.3* [34], as the closed-form expression for the attention distribution for a given row $f(x)$ is not known analytically from the values of attention sampled. The differential entropy for a row i is given by



and the corresponding vector $h \in R^{n \times 1}$ corresponds to the entropy across every row. From the row-wise entropy, this vector is softmaxed to obtain the corresponding weights, $w \in R^{n \times 1}$, as follows:



The resultant embedding for a note sequence, $N \in R^{768 \times 1}$, is then given by the matrix multiplication as follows:



and the final patient-level representation is the simple average over all note fragments for a given patient, for their most recent encounter. A visualization of attention matrices with varying row-wise entropy and thus varying weighting per token is provided in Figures S1 and S2 in Multimedia Appendix 2.

Hierarchical Clustering

Clustering analysis was performed on *ICD-9* embeddings and clinical text representations to identify ADRD subtypes. We selected hierarchical agglomerative clustering with Ward linkage due to its ability to capture the hierarchical structure of clinical data, as seen in *ICD-9* codes (eg, metabolic disorders branching into type 1 and type 2 diabetes) and clinical notes (eg, cognitive impairment branching into memory loss and aphasia). Cluster quality was evaluated using elbow plots and silhouette scores, with implementation via *Scikit-learn v1.0.1* [35].

Optimal Transport

To address provider-specific effects in embeddings of clinical notes, we first applied Uniform Manifold Approximation and Projection (UMAP) to reduce the dimensionality of the data and then applied the earth mover distance transport approach using the Python *Optimal Transport* package [36]. *Optimal Transport* provides a mathematical framework to minimize the cost of transforming one distribution into another, which can address the problem of domain adaptation [37]. Domain adaptation involves adjusting data from different sources to make their data distributions more comparable, ensuring that models trained on these data perform well across various settings. In this context, the earth mover distance method was used to align the embeddings from various providers to a standard reference. This alignment ensured that the subsequent clustering analysis was less skewed by provider-related differences, allowing a more accurate interpretation of the underlying phenotypic variation.

Enrichment Analysis: ICD Clusters

ICD clusters were phenotypically characterized by testing for enrichment of *ICD-9* codes within each cluster. For each cluster, a 2×2 contingency table was generated for each *ICD-9* diagnosis code, comparing counts of patients with that code within cluster to counts of patients with that same code in other clusters. A chi-square test for enrichment was performed; the prevalence ratio (PR), calculated as the prevalence of each code in one cluster divided by its prevalence in the rest, was calculated to measure the strength of the association. A Bonferroni correction was applied to the resultant P values to correct for multiple comparisons, and *ICD* codes in each cluster were ranked by corrected P value to characterize the most significant enrichments. All P values in this investigation were two-sided, with a postcorrected α of 0.05 to determine significance. The top-10 significant diagnoses with the highest PR were extracted from each cluster for interpretation. If a cluster lacked significant diagnoses, the top diagnoses with the highest PR among the nonsignificant ones were selected. The enrichment analyses were conducted in Python version 3.8.15.

Topic Modeling: Note Clusters

We used BERTopic [38], using the Python package *BERTopic* v0.16.0, to identify representative topics and key terms within each note cluster. Embeddings obtained through optimal transport were directly used for clustering and topic assignment, bypassing the need for additional embedding and dimensionality

XSL•FO
RenderX

reduction steps. Before conducting cluster-based term frequency–inverse document frequency (TF-IDF) for topic assignment, the clinical text was preprocessed using a vectorizer to remove stop words and exclude common terms that appeared too frequently across most notes. Furthermore, to fine-tune and enhance the word representation of topics, we applied the KeyBERTInspired model, which extracts keywords by leveraging embeddings and cosine similarity to find the words with the closest semantic relationship to the note texts, thereby making them more representative of the topics. Following the extraction of representative terms within each cluster, we used GPT-4 (OpenAI) [39], a state-of-the-art large language model (LLM), to enhance interpretability. GPT-4 summarized the representative words provided by BERTopic into coherent themes with greater clinical significance, such as specific medical conditions, treatments, and medications. The PR, calculated as the prevalence of each word in one cluster divided by its prevalence in the rest, was calculated to measure the strength of the semantic relationship. The BERTopic modeling and analyses were conducted in Python version 3.9.6.

ADRD Diagnosis Categorization

The categorization of ADRD diagnoses was conducted using an extensive list of diagnostic names based on disease etiology. This list was meticulously reviewed for each unique diagnosis name recorded for the MGH memory clinic patients in the EHR system. The ADRD diagnoses categories included AD; dementia unspecified; FTD; LBD; vascular cognitive impairment (VCI); and others, such as posterior cortical atrophy (PCA), progressive supranuclear palsy, corticobasal degeneration, and primary progressive aphasia. An expert behavioral neurologist (JRD) provided critical input during this process, helping to develop a comprehensive mapping list that correlates specific diagnosis names with their corresponding ADRD categories. The application of this mapping to the data was performed using R version 4.3.2 (R Foundation for Statistical Computing). The full list of diagnosis names corresponding to ADRD diagnosis categories is provided in Multimedia Appendix 1.

Cluster Characterizations

To assess associations between clusters and sex, as well as ADRD diagnoses, we used the chi-square test. For each cluster, a 2×2 contingency table was generated for each variable, comparing the counts of patients with the characteristic within the cluster to those in other clusters. The PR, defined as the prevalence of a characteristic in one cluster divided by its prevalence in the remaining clusters, was calculated to measure the strength of the association: 1 indicates no difference in prevalence between the 2 groups, >1 indicates higher prevalence in the first group, and <1 indicates lower prevalence in the first group. In addition, to examine variations in the age of onset across clusters, we initially conducted a Kruskal-Wallis Rank Sum Test using the *stats* package from R. η^2 (calculated by subtracting the number of groups from the Kruskal-Wallis H statistic plus one, and then dividing this result by the total number of observations minus the number of groups) based on the H statistic was reported as the effect size: values closer to 0 indicate a smaller effect and values closer to 1 indicate a larger effect. Following significant findings, further post hoc analyses using the Dunn test were performed to delineate differences between groups. The *P* values were adjusted for multiple comparisons using the Benjamini-Hochberg method to control the false discovery rate (FDR). Age-of-onset data were rigorously annotated by human experts reviewing clinical notes; where notes did not specify an exact age of onset, the age at the first clinical visit within the memory clinic was used as an approximation. Finally, we conducted a chi-square test between *ICD* clusters and note clusters to test whether patient cluster assignment was consistent across note and *ICD* representations. If the contingency table was larger than 2×2, Cramér V (calculated as the square root of the chi-square statistic divided by the product of the sample size and the minimum dimension minus one) was reported as the effect size: 0 indicates no association and 1 indicates a strong association. Standardized residuals (standardized differences between the observed count and the expected count) were reported for each cell: values close to 0 indicate the observed count is close to the expected count, positive values indicate the observed count is higher than expected, and negative values indicate the observed count is lower than expected. All statistical analyses were conducted in R version 4.3.2.

Results

Study Population

Our final study population consisted of 3454 patients from the MGH tertiary care memory clinic with clinical notes and *ICD* codes in the EHR system. The average age of onset for patients was 72.1 (SD 9.5) years, with 1678 (48.58%) being female. The majority were White (n=3059, 88.56%), followed by Asian (n=90, 2.61%), Black or African American (n=77, 2.23%), American Indian or Alaska Native (n=4, 0.12%), and Native Hawaiian or Other Pacific Islander (n=1, 0.03%). In addition, 103 (2.98%) identified as belonging to other races, and race data were not available for 120 (3.47%) patients. Regarding ethnicity, 3020 (87.43%) identified as non-Hispanic, 106 (3.07%) as Hispanic, and ethnicity data were not available for 328 (9.5%) patients. AD was the most prevalent diagnosis, affecting 1317 (38.13%) patients, followed by dementia unspecified, which accounted for 1101 (31.88%) patients. Each encounter recorded only one diagnosis name, and only the most recent encounter was used. The patient selection details are illustrated in Figure 2. The demographic and ADRD diagnosis breakdowns are provided in Table 1 and ADRD categorization details are provided in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) diagram illustrating the selection of patients from the Massachusetts General Hospital (MGH) electronic health record (EHR) system. ADRD: Alzheimer disease and related dementias; Dx: diagnosis; ICD: International Classification of Diseases.

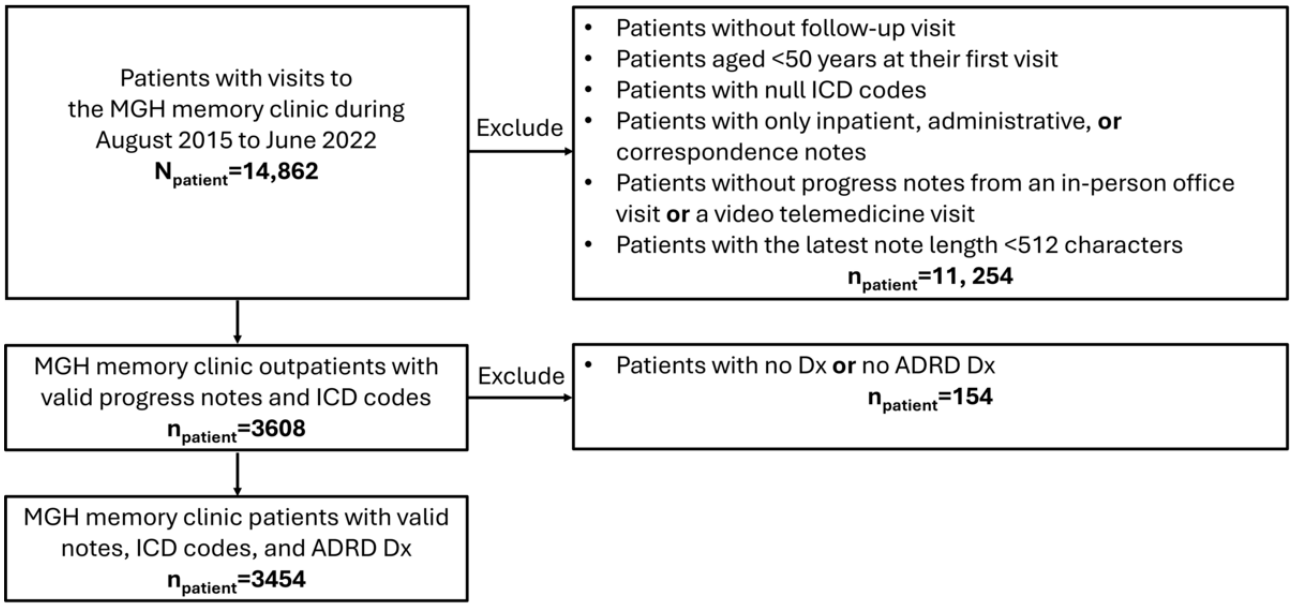


Table 1. Summary statistics of the final study population.

Characteristics	Total (N=3454)	AD[a] (n=1317)	Dementia unspecified (n=1101)	FTD[b] (n=190)	LBD[c] (n=261)	VCI[d] (n=96)	Other[e] (n=489)
Age of onset[f] (y), mean (SD)	72.1 (9.5)	74.6 (7.8)	73.3 (9.7)	63.8 (8.2)	71.4 (7.7)	76 (7.7)	65.1 (9.4)
Sex, n (%)							
Female	1678 (48.58)	717 (54.44)	538 (48.86)	78 (41.1)	64 (24.5)	41 (43)	240 (49.1)
Male	1776 (51.42)	600 (45.56)	563 (51.14)	112 (58.9)	197 (75.5)	55 (57)	249 (50.9)
Race, n (%)							
White	3059 (88.56)	1149 (87.24)	988 (89.74)	168 (88.4)	227 (87.0)	80 (83)	447 (91.4)
Black or African American	77 (2.23)	27 (2.05)	22 (2.00)	6 (3.2)	4 (1.5)	9 (9)	9 (1.8)
Asian	90 (2.61)	34 (2.58)	31 (2.82)	6 (3.2)	14 (5.4)	1 (1)	4 (0.8)
American Indian or Alaska Native	4 (0.12)	2 (0.15)	1 (0.09)	0 (0.0)	0 (0.0)	0 (0)	1 (0.2)
Native Hawaiian or Other Pacific Islander	1 (0.03)	1 (0.08)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Other	103 (2.98)	52 (3.95)	32 (2.91)	2 (1.1)	2 (0.8)	3 (3)	12 (2.5)
Unavailable	120 (3.47)	52 (3.95)	27 (2.45)	8 (4.2)	14 (5.4)	3 (3)	16 (3.3)
Ethnicity, n (%)							
Not Hispanic or Latino	3020 (87.43)	1133 (86.03)	987 (89.65)	161 (84.7)	228 (87.4)	83 (87)	428 (87.5)
Hispanic or Latino	106 (3.07)	53 (4.02)	37 (3.36)	2 (1.1)	3 (1.1)	3 (3)	8 (1.6)
Unavailable	328 (9.50)	131 (9.95)	77 (6.99)	27 (14.2)	30 (11.5)	10 (10)	53 (10.8)

[a]AD: Alzheimer disease.

[b]FTD: frontotemporal dementia.

[c]LBD: Lewy body dementia.

[d]VCI: vascular cognitive impairment.

[e]Includes posterior cortical atrophy, progressive supranuclear palsy, corticobasal degeneration, and primary progressive aphasia.

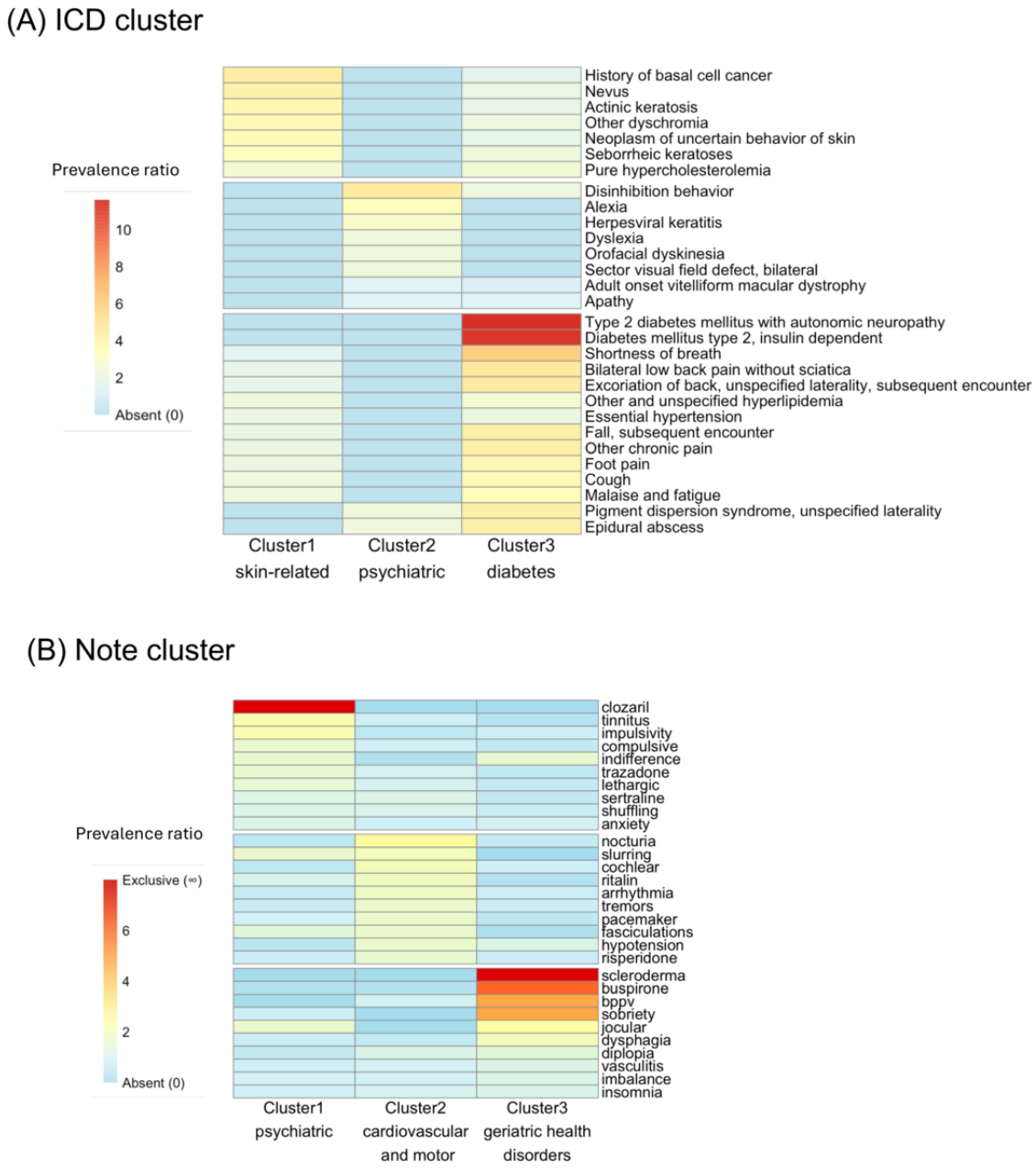
[f]Age of onset was manually annotated by experts viewing clinical notes; for notes without age of onset, we approximated with the age of first encounter.

XSL•FO
RenderX

ICD Clustering

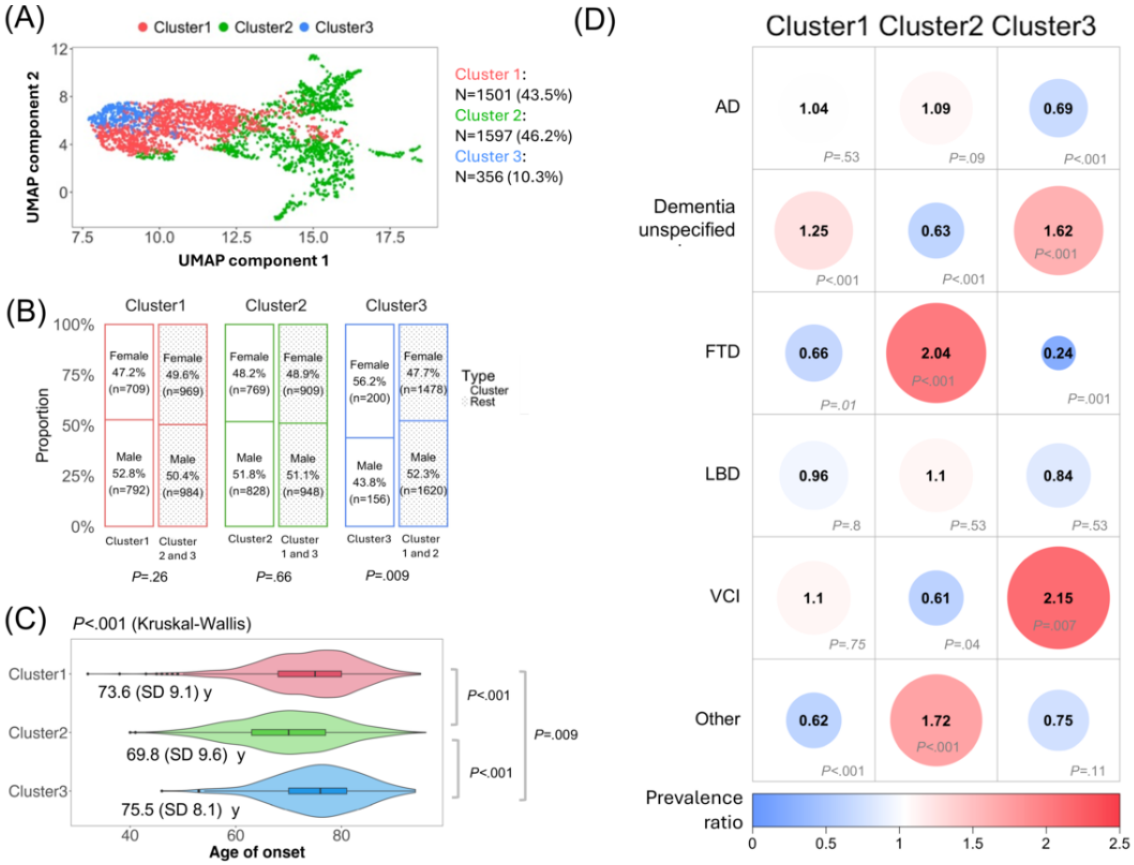
To investigate the clinical heterogeneity within ADRD, we clustered the embeddings of non-ADRD *ICD* codes assigned to this tertiary care sample from the MGH ADRD cohort. We hypothesized that this approach would reveal distinct clinical subtypes based on clinical comorbidities, which were associated with demographics (ie, age of onset and sex). The hierarchical agglomerative clustering method revealed 3 distinct clusters in the embeddings of non-ADRD *ICD* codes, as determined by the silhouette score. Figure 3A depicts a heatmap of enriched *ICD* codes across each cluster, and a 2D UMAP projection of these *ICD* code embeddings, colored by cluster, is displayed in Figure 4A. The distribution of patients across the clusters was as follows: cluster 1 included 1501 (43.46%) patients, cluster 2 included 1597 (46.24%) patients, and cluster 3 included 356 (10.31%) patients. Detailed summary statistics for these clusters are presented in Table 2.

Figure 3. Heatmap of enrichment in International Classification of Diseases (ICD) clusters and topics in note clusters. (A) This heatmap displays the enrichment of ICD-9 codes across ICD embedding clusters. Cluster 1 is primarily dominated by skin-related and certain cardiovascular conditions. Cluster 2 is marked by its exclusive and high prevalence ratios (PR) in psychiatric and behavioral conditions. Cluster 3 shows a diverse set of conditions with a significant prevalence of respiratory, pain-related, and complicated diabetic mellitus. (B) This heatmap displays representative words for each note cluster identified through topic modeling. Cluster 1 is primarily dominated by psychiatric manifestations and medications. Cluster 2 highlights cardiovascular, motor, and sensory issues. Cluster 3 covers a variety of symptoms and conditions, including autoimmune issues, behavioral and movement disorders, sleep disturbances, etc. In both (A) and (B), the color intensity of each code or word-cluster pairing reflects the PR (prevalence of code or word in observed group divided by prevalence in other groups) associated with that code or word. Words colored as exclusive were only present in one cluster.



XSL•FO
RenderX

Figure 4. Clustering of International Classification of Diseases (ICD) embeddings and their demographic and diagnostic associations. (A) Uniform Manifold Approximation and Projection (UMAP) visualization of ICD embeddings were characterized by 3 clusters: cluster 1 includes 1501 (43.5%) patients, cluster 2 comprises 1597 (46.2%) patients, and cluster 3 contains 356 (10.3%) patients. (B) Bar plot showing prevalence for sex by cluster, significance based on a chi-square test. Notably, cluster 3 has a significantly higher proportion of female participants compared to male participants (P_{FDR}=.009). (C) Violin plot illustrating the distribution of age of onset across clusters. Each violin plot shows the kernel density estimate of the data, with the center line representing the median age of onset. Box plot elements are overlaid, where the box limits indicate the upper and lower quartiles, and the whiskers extend to 1.5 times the IQR. Individual points are hidden for clarity. Significant differences are observed, with cluster 2 showing the earliest average age of onset at 69.8 (SD 9.6) years, and cluster 3 the latest at 75.5 (SD 8.1) years (P<.001). (D) Heatmap showing prevalence ratio for Alzheimer disease and related dementias (ADRD) diagnoses across clusters, significance derived from a chi-squared test. Clinical diagnoses include Alzheimer disease (AD); dementia unspecified; frontotemporal dementia (FTD); Lewy body dementia (LBD); vascular cognitive impairment (VCI); and others such as posterior cortical atrophy (PCA), progressive supranuclear palsy (PSP), corticobasal degeneration (CBD), and primary progressive aphasia (PPA). Significant distribution variations are evident across clusters. The circle size, color, and number indicate the magnitude of the prevalence ratio. Blue indicates underrepresentation while red indicates overrepresentation. *P* values were corrected with false discovery rate adjustments.



We examined differences in non-ADRD *ICD* code frequency across patient *ICD* embedding clusters (Figure 3A). In cluster 1, diagnoses, such as seborrheic keratoses (χ^2_1=243.15; P_{FDR}<.001; PR=3.32), actinic keratosis (χ^2_1=236.53; P_{FDR}<.001; PR=3.91), pure hypercholesterolemia (χ^2_1=199.13; P_{FDR}<.001; PR=2.71), history of basal cell cancer (χ^2_1=196.22; P_{FDR}<.001; PR=4.45), and nevus (χ^2_1=184.53; P_{FDR}<.001; PR=4.23), show notably high PRs. These diagnoses largely fall into skin-related disorders (such as various types of skin cancer and keratoses). Cluster 2 appears to be unique, with top PRs noted only for clinical signs, such as disinhibition behavior (χ^2_1=1.14; P_{FDR}=.99; PR=4.65), alexia (χ^2_1=0.426; P_{FDR}=.99; PR=3.49), and orofacial dyskinesia (χ^2_1=0.017; P_{FDR}=.99; PR=2.33), suggesting behavioral and psychiatric manifestations which are consistent with the FTD enrichment and earlier onset noted above. Notably, although these diagnoses did not reach significance in cluster 2, they were absent in the other clusters. Moreover, many diagnoses in cluster 2 were marked with a lack of PR, indicating a lower relevance of these diagnoses compared to clusters 1 or 3. Cluster 3 exhibited significant increases in diagnoses from a variety of categories, including respiratory issues (eg, cough: χ^2_1=339.47; P_{FDR}<.001; PR=3.77 and shortness of breath: χ^2_1=306.60; P_{FDR}<.001; PR=6.03), chronic pain (χ^2_1=460.55; P_{FDR}<.001; PR=4.80), musculoskeletal problems (eg, bilateral low back pain without sciatica: χ^2_1=456.91; P_{FDR}<.001; PR=4.80 and foot pain: χ^2_1=372.91; P_{FDR}<.001; PR=3.86), and complications of diabetes mellitus. Notably, diabetes mellitus (type 2 with autonomic neuropathy: χ^2_1=322.98; P_{FDR}<.001; PR=11.60 and insulin-dependent diabetes: χ^2_1=308.96; P_{FDR}<.001; PR=11.30) had exceptionally high PRs, suggesting a very strong association with these severe diabetes conditions in cluster 3.

XSL•FO
RenderX

Table 2. Summary statistics of International Classification of Diseases clusters.

Characteristics	Total (N=3454)	Cluster 1 (n=1501)	Cluster 2 (n=1597)	Cluster 3 (n=356)
Age of onset[a] (y), mean (SD)	72.1 (9.5)	73.6 (9.1)	69.8 (9.6)	75.5 (8.1)
Sex, n (%)				
Female	1678 (48.58)	709 (47.24)	769 (48.15)	200 (56.2)
Male	1776 (51.42)	792 (52.76)	828 (51.85)	156 (43.8)
Race, n (%)				
White	3059 (88.56)	1335 (88.94)	1421 (88.98)	303 (85.1)
Black or African American	77 (2.23)	37 (2.47)	25 (1.57)	15 (4.2)
Asian	90 (2.61)	40 (2.66)	40 (2.5)	10 (2.8)
American Indian or Alaska Native	4 (0.12)	0 (0)	2 (0.13)	2 (0.6)
Native Hawaiian or Other Pacific Islander	1 (0.03)	0 (0)	1 (0.06)	0 (0)
Other	103 (2.98)	51 (3.4)	34 (2.13)	18 (5.1)
Unavailable	120 (3.47)	38 (2.53)	74 (4.63)	8 (2.2)
Ethnicity, n (%)				
Not Hispanic or Latino	3020 (87.43)	1358 (90.47)	1326 (83.03)	336 (94.4)
Hispanic or Latino	106 (3.07)	45 (3)	42 (2.63)	19 (5.3)
Unavailable	328 (9.5)	98 (6.53)	229 (14.34)	1 (0.3)
ADRD Dx[b], n (%)				
AD[c]	1317 (38.13)	584 (38.91)	636 (39.82)	97 (27.2)
Dementia unspecified	1101 (31.88)	540 (35.98)	388 (24.3)	173 (48.6)
FTD[d]	190 (5.5)	64 (4.26)	121 (7.58)	5 (1.4)
LBD[e]	261 (7.56)	111 (7.4)	127 (7.95)	23 (6.5)
VCI[f]	96 (2.78)	44 (2.93)	33 (2.07)	19 (5.3)
Other[g]	489 (14.16)	158 (10.53)	292 (18.28)	39 (11)

[a]Age of onset was manually annotated by experts viewing clinical notes; for notes without age of onset, we approximated with the age of first encounter.

[b]Dx: diagnosis.

[c]AD: Alzheimer disease.

[d]FTD: frontotemporal dementia.

[e]LBD: Lewy body dementia.

[f]VCI: vascular cognitive impairment.

[g]Includes posterior cortical atrophy, progressive supranuclear palsy, corticobasal degeneration, and primary progressive aphasia.

Furthermore, statistical analyses revealed significant associations between ICD cluster membership, demographic variables, and diagnostic categories. Cluster 3 had an overrepresentation of female participants relative to clusters 1 and 2 (χ^2_1=8.8; P_{FDR}=.009; PR=1.178); however, clusters 1 and 2 showed no significant differences in sex distribution (cluster 1: χ^2_1=1.8; P_{FDR}=.26 and cluster 2: χ^2_1=0.2; P_{FDR}=.66; Figure 4B). In addition, the age of onset varied significantly among the ICD clusters (Kruskal-Wallis χ^2_2=182.6; P_{FDR}<.001, η^2=0.052). Cluster 2, with a mean age of onset of 69.8 (SD 9.6) years, had a significantly earlier age of onset compared with clusters 1 (Z=10.13; P_{FDR}<.001) and 3 (Z=−8.83; P_{FDR}<.001). Moreover, cluster 1 (mean 73.6, SD 9.1 years) had an earlier onset than cluster 3 (mean 75.5, SD 8.1 years; Z=−2.61; P_{FDR}=.009; Figure 4C). Cluster 1 was significantly enriched by dementia unspecified (χ^2_1=20.2; $_{FDR}$<.001; PR=1.252); cluster 2 was significantly enriched by FTD (χ^2_1=23.9; $_{FDR}$<.001; PR=2.039) and other rare ADRDs (χ^2_1=41; $_{FDR}$ <.001; PR=1.724); and cluster 3 was significantly enriched by VCI (χ^2_1=8.6; $_{FDR}$=.007; PR=2.147) and dementia unspecified (χ^2_1=50.2; $_{FDR}$<.001; PR=1.622). In contrast, no cluster was significantly enriched by AD, though AD was significantly underrepresented in cluster 3 (cluster 1: χ^2_1=0.6; $_{FDR}$=.53; cluster 2: χ^2_1=3.5; $_{FDR}$=.09; and cluster 3: χ^2_1=19.4; $_{FDR}$<.001; Figure 4D) or LBD (cluster 1: χ^2_1=0.1; $_{FDR}$=.80; cluster 2: χ^2_1=0.5; $_{FDR}$=.53; and cluster 3: χ^2_1=0.5; $_{FDR}$=.53; Figure 4D). Additional visualizations of

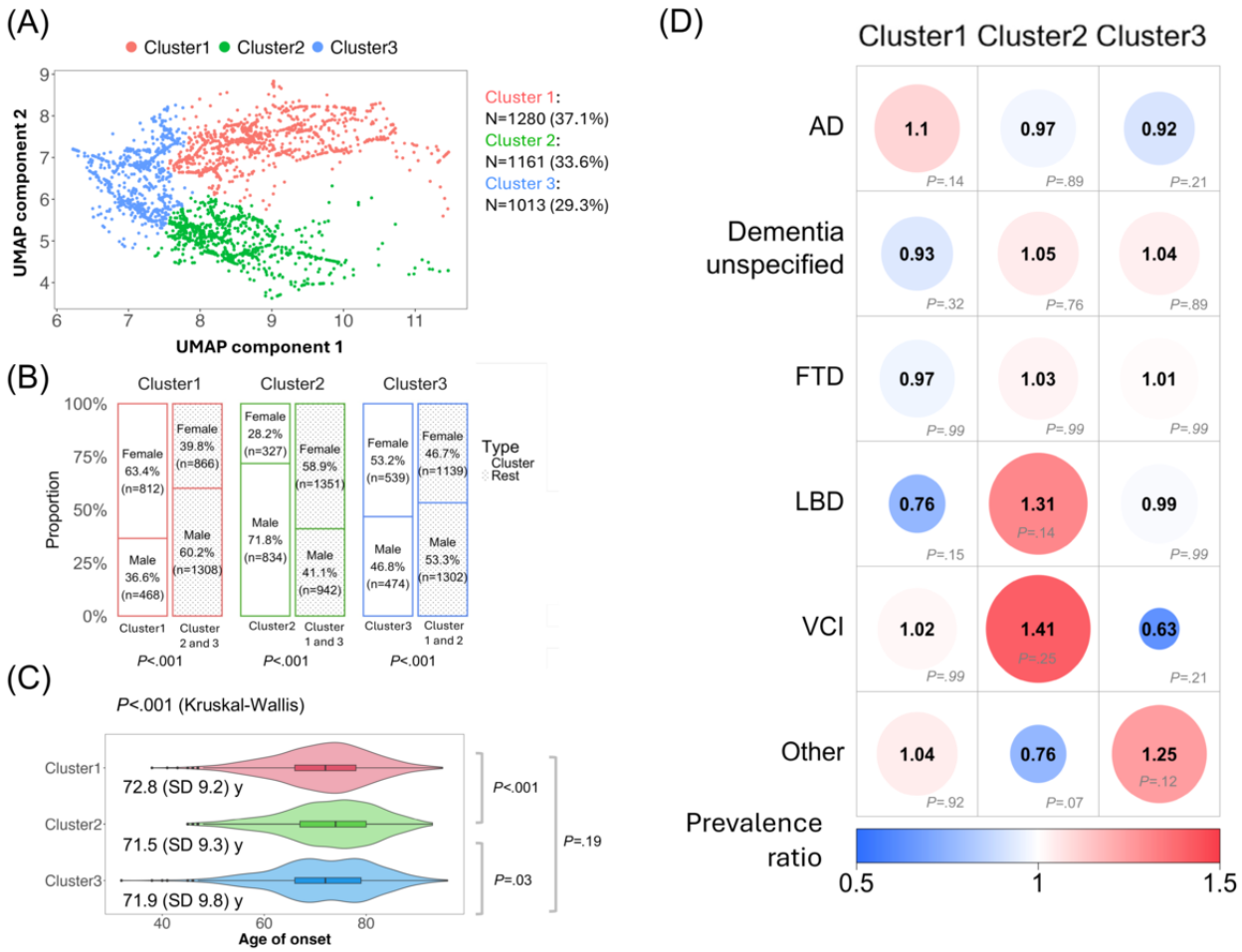
XSL•FO
RenderX

the UMAP projections colored by sex, age of onset, and ADRD diagnoses are available in Figures S3A, S3B, and S3C, respectively, in Multimedia Appendix 2.

Note Clustering

Initially, a provider effect was detected in the UMAP projection of note embeddings from the latest clinical notes of 3454 patients (Figure S4A in Multimedia Appendix 2). To address this, we used an optimal transport method, aligning the embeddings from all providers to the 2D embedding of a selected reference provider (Figure S4B in Multimedia Appendix 2). Following this alignment, hierarchical agglomerative clustering was applied to the adjusted note embeddings, revealing 3 distinct clusters, as determined by the silhouette score. The adjusted UMAP projection, color coded by cluster, is presented in Figure 5A. The patient distribution within these clusters was as follows: cluster 1 included 1280 (37.06%) patients, cluster 2 included 1161 (33.61%) patients, and cluster 3 included 1013 (29.33%) patients. Detailed summary statistics for each cluster are outlined in Table 3.

Figure 5. Clustering of note embeddings and their demographic and Alzheimer disease and related dementias (ADRD) diagnosis associations. (A) Uniform Manifold Approximation and Projection (UMAP) visualization of note embeddings characterized by 3 clusters: cluster 1 includes 1280 (37.1%) patients, cluster 2 includes 1161 (33.6%) patients, and cluster 3 includes 1013 (29.3%) patients. (B) Bar plot showing prevalence for sex by cluster, with significance based on chi-square tests. Notably, cluster 1 and cluster 3 were both enriched by female participants (P<.001) while cluster 2 was enriched by male participants (P<.001). (C) Violin plot illustrating the distribution of age of onset across clusters. Each violin plot shows the kernel density estimate of the data, with the center line representing the median age of onset. Box plot elements are overlaid, where the box limits indicate the upper and lower quartiles, and the whiskers extend to 1.5 times the IQR. Individual points are hidden for clarity. Significant differences are observed, with cluster 2 showing the latest average age of onset at 72.8 (SD 9.2) years, and cluster 1 the earliest at 71.5 (SD 9.3) years (P<.001). (D) Heatmap showing prevalence ratio for ADRD diagnoses across clusters, with significance derived from a chi-square test. Diagnoses include Alzheimer disease (AD); dementia unspecified; frontotemporal dementia (FTD); Lewy body dementia (LBD); vascular cognitive impairment (VCI); and others such as posterior cortical atrophy (PCA), progressive supranuclear palsy (PSP), corticobasal degeneration (CBD), and primary progressive aphasia (PPA). No significant distribution variations are observed across clusters (P>.05). The circle size, color, and number indicate the magnitude of the prevalence ratio. Blue indicates underrepresentation while red indicates overrepresentation. *P* values were corrected with false discovery rate (FDR) adjustments.



XSL•FO
RenderX

Table 3. Summary statistics of note clusters.

Characteristics	Total (N=3454)	Cluster 1 (n=1280)	Cluster 2 (n=1161)	Cluster 3 (n=1013)
Age of onset[a] (y), mean (SD)	72.1 (9.5)	71.5 (9.3)	72.8 (9.2)	71.9 (9.8)
Sex, n (%)				
Female	1678 (48.58)	812 (63.44)	327 (28.17)	539 (53.21)
Male	1776 (51.42)	468 (36.56)	834 (71.83)	474 (46.79)
Race, n (%)				
White	3059 (88.56)	1129 (88.20)	1032 (88.89)	898 (88.56)
Black or African American	77 (2.23)	30 (2.34)	26 (2.24)	21 (2.07)
Asian	90 (2.61)	36 (2.81)	25 (2.15)	29 (2.86)
American Indian or Alaska Native	4 (0.12)	1 (0.08)	1 (0.09)	2 (0.2)
Native Hawaiian or Other Pacific Islander	1 (0.03)	0 (0)	0 (0)	1 (0.1)
Other	103 (2.98)	34 (2.66)	44 (3.79)	25 (2.47)
Unavailable	120 (3.47)	50 (3.91)	33 (2.84)	37 (3.65)
Ethnicity, n (%)				
Not Hispanic or Latino	3020 (87.43)	1128 (88.13)	1004 (86.48)	888 (87.66)
Hispanic or Latino	106 (3.07)	36 (2.81)	39 (3.36)	31 (3.06)
Unavailable	328 (9.5)	116 (9.06)	118 (10.16)	94 (9.28)
ADRD Dx[b], n (%)				
AD[c]	1317 (38.13)	519 (40.55)	435 (37.47)	363 (35.83)
Dementia unspecified	1101 (31.88)	389 (30.39)	381 (32.82)	331 (32.68)
FTD[d]	190 (5.5)	69 (5.39)	65 (5.6)	56 (5.53)
LBD[e]	261 (7.56)	81 (6.33)	104 (8.96)	76 (7.5)
VCI[f]	96 (2.78)	36 (2.81)	40 (3.45)	20 (1.97)
Other[g]	489 (14.16)	186 (14.53)	136 (11.71)	167 (16.49)

[a]Age of onset was manually annotated by experts viewing clinical notes; for notes without age of onset, we approximated with the age of first encounter.

[b]Dx: diagnosis.

[c]AD: Alzheimer disease.

[d]FTD: frontotemporal dementia.

[e]LBD: Lewy body dementia.

[f]VCI: vascular cognitive impairment.

[g]Includes posterior.

We extracted common topics from each note cluster using topic modeling and examined the distribution of ADRD diagnoses across these clusters. In cluster 1, we found more terms related to psychiatric manifestations (eg, compulsive, indifference, and anxiety) and medications (eg, clozaril, trazodone, and sertraline), with a slight but nonsignificant enrichment in AD diagnosis (χ^2_1=4.9; P_{FDR}=.14; PR=1.10). Cluster 2 had more terms related to cardiovascular issues (eg, pacemaker, hypotension, and arrhythmia) and motor and sensory issues (eg, slurring, cochlear, and tremors), with a slight but nonsignificant enrichment in LBD (χ^2_1=4.6; P_{FDR}=.14; PR=1.31) and VCI (χ^2_1=2.5; P_{FDR}=.25; PR=1.41) diagnoses. Cluster 3 encompassed a wide variety of symptoms and conditions common in geriatric populations, including autoimmune (eg, scleroderma and vasculitis), behavioral changes and movement (eg, usual jocular behavior and imbalance, dysphagia), sleep (eg, insomnia), and sensory (eg, diplopia) problems, with a slight but not-significant enrichment in rare ADRD diagnoses (χ^2_1=3; P_{FDR}=.12; PR=1.25). Figure 3B depicts the list of representative words from each cluster and their PR, and Figure S5D in Multimedia Appendix 2 illustrates sentence examples from each cluster.

Furthermore, statistical analyses revealed significant associations of note cluster membership, with demographic variables, but not with ADRD diagnoses. First, both cluster 1 and 3 were significantly enriched by female participants (cluster 1: χ^2_1=178.7; P_{FDR}<.001; PR=1.593 and cluster 3: χ^2_1=12; P_{FDR}<.001; PR=1.14) while cluster 2 was enriched by male participants (χ^2_1=290.5; P_{FDR}<.001; PR=1.749; Figure 5B). In

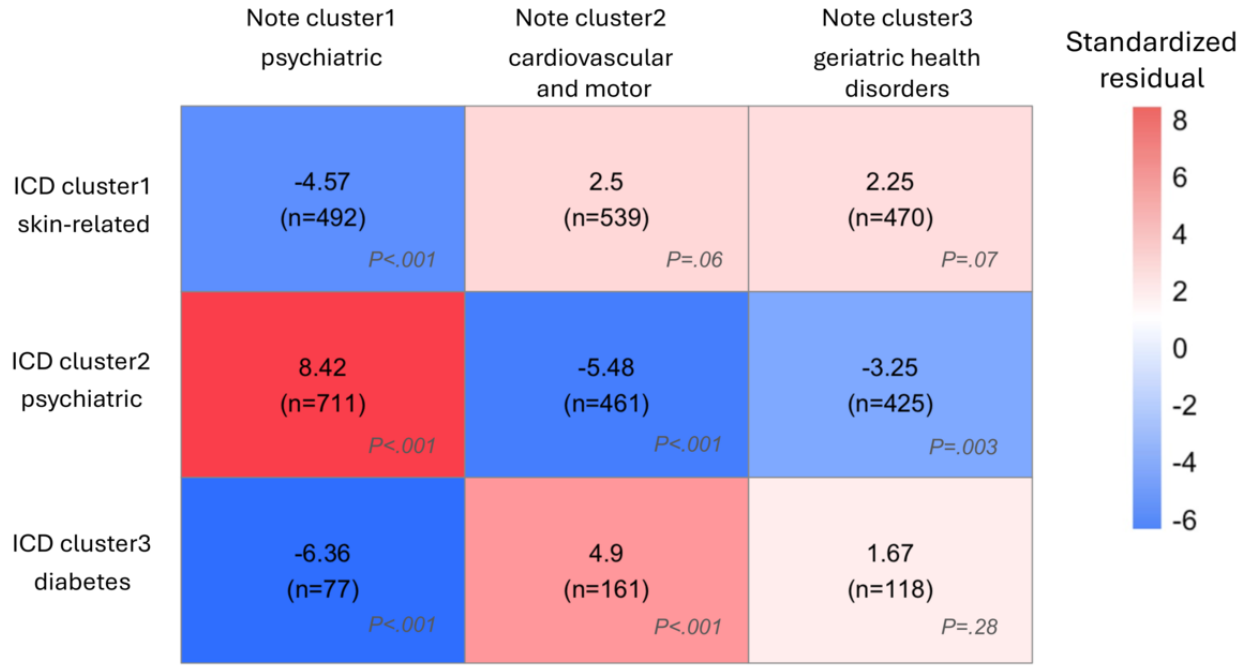
XSL•FO
RenderX

addition, age of onset varied significantly among the note clusters (Kruskal-Wallis χ^2_2=14.9; P<.001; η^2=0.004), with cluster 2 (mean 72.8, SD 9.2 years) having a significantly later age of onset compared to cluster 1 (Z=−3.82; P_{FDR}<.001) and cluster 3 (Z=2.31; P_{FDR}=.03), while cluster 1 (mean 71.5, SD 9.3 years) and cluster 3 (mean 71.9, SD 9.8 years) did not differ (Z=−1.33; P_{FDR}=.19; Figure 5C). However, no association was observed between note cluster membership and ADRD diagnoses (Cluster 1-AD: χ^2_1=4.9; P_{FDR}=.14; Cluster 1-dementia unspecified: χ^2_1=1.9; P_{FDR}=.32; Cluster 1-FTD: χ^2_1=0.02; P_{FDR}=.99; Cluster 1-LBD: χ^2_1=4.1; P_{FDR}=.15; Cluster 1-other: χ^2_1=0.19; P_{FDR}=.92; Cluster 1-VCI: χ^2_1<0.001, P_{FDR}=.99; Cluster 2-AD: χ^2_1=0.3; P_{FDR}=.89; Cluster 2-dementia unspecified: χ^2_1=0.6; P_{FDR}=.76; Cluster 2-FTD: χ^2_1=0.01; P_{FDR}=.99; Cluster 2-LBD: χ^2_1=4.6; P_{FDR}=.14; Cluster 2-other: χ^2_1=8.3; P_{FDR}=.07; Cluster 2-VCI: χ^2_1=2.5; P_{FDR}=.25; Cluster 3-AD: χ^2_1=3.1; P_{FDR}=.21; Cluster 3-dementia unspecified: χ^2_1=0.4; P_{FDR}=.89; Cluster 3-FTD: χ^2_1<0.001; P_{FDR}=.99; Cluster 3-LBD: χ^2_1<0.001; P_{FDR}=.99; Cluster 3-other: χ^2_1=6.1; P_{FDR}=.12; Cluster 3-VCI: χ^2_1=3; P_{FDR}=.21; Figure 5D). Additional visualizations of the UMAP projections colored by sex, age of onset, and ADRD diagnoses are provided in Figures S5A, S5B, and S5C, respectively, in Multimedia Appendix 2.

Comparison Between ICD Clusters and Note Clusters

Statistical analysis demonstrated significant associations between *ICD* and note clusters (χ^2_4=89.43; P<.001; Cramér V=0.114). Specifically, note cluster 1, characterized by more female participants (PR=1.593) and terms related to psychiatric manifestations and medications, significantly overlapped (standardized residual=8.42; P_{FDR}<.001) with *ICD* cluster 2, which is noted for the earliest onset of disease (mean 69.8, SD 9.6 years) and a higher prevalence of psychiatric disorders and higher proportion of patients with FTD (PR=2.039). In addition, note cluster 2, which had higher proportion of male participants (PR=1.749) and terms related to cardiovascular and motor issues, overlapped significantly (standardized residual=4.90; P_{FDR}<.001) with *ICD* cluster 3, which is marked by the oldest onset of disease (mean 75.5, SD 8.1 years), a higher occurrence of VCI (PR=2.147), and dementia unspecified (PR=1.622) and had high prevalence of diabetes. These findings suggest a meaningful pattern of cluster correspondence across modalities (Figure 6).

Figure 6. Heatmap of the association of International Classification of Diseases (ICD) clusters with note clusters. Heatmap of the association of ICD clusters with note clusters. ICD clusters were significantly associated with note clusters (P<.001). Post hoc analyses revealed that note cluster 1 was positively associated with ICD cluster 2 (standardized residual=8.42, P<.001) and negatively associated with ICD clusters 1 and 3 (cluster 1: standardized residual=−4.57, P<.001; cluster 3: standardized residual=−6.36, P<.001). Furthermore, note cluster 2 was positively associated with ICD cluster 3 (standardized residual=4.9, P<.001) and negatively associated with ICD cluster 2 (standardized residual=−5.48, P<.001). Finally, note cluster 3 was negatively associated with ICD cluster 2 (standardized residual=−3.25, P=.003). Each cell displays standardized residuals (standardized differences between the observed count and the expected count) along with the count of overlapping patients (n). Color bar represents the value of the standardized residual. P values were corrected with false discovery rate adjustments.



XSL•FO
RenderX

Discussion

Principal Findings

This study aimed to characterize the clinical heterogeneity across ADRD by applying representation learning techniques on patient EHRs in a tertiary care clinic. We used pretrained *ICD* code embeddings, a transformer architecture for encoding clinical notes, and unsupervised learning to identify distinct ADRD subtypes. This work represents the first example of clustering patients with ADRD using embeddings derived directly from an LLM, without prior rule-based extraction of relevant medical concepts from the clinical note. The *ICD* codes allowed us to investigate subtypes of patients with similar *ICD* codes (non-ADRD diagnosis in charts), while the clinical notes allowed us to capture clinical history and presentation recorded by memory specialists. Our results demonstrate distinct patterns of disease manifestation with significant overlap between the *ICD* and note clusters. The overlap of clusters between the 2 approaches suggests that the subtypes may reflect common underlying clinical heterogeneity, as distinct subtypes can be identified through different data modalities.

Our choice of hierarchical agglomerative clustering was guided by the hierarchical nature of our data, empirical evidence from prior dementia subtyping studies, and theoretical limitations of alternative algorithms. Alternative methods, such as K-means, Gaussian mixture models [40], and density-based spatial clustering of applications with noise (DBSCAN) [41], face theoretical limitations: K-means assumes spherical clusters that is difficult to satisfy in high-dimensional embeddings; DBSCAN relies on density thresholds that break down in such spaces; and Gaussian mixture models can be unstable with overlapping subtypes. In contrast, hierarchical clustering preserves these nested relationships, resulting in more homogeneous [42] and reproducible [43] clusters.

In our study, we identified 3 clusters from *ICD* embeddings, each characterized by distinct health conditions. Cluster 1 predominantly featured issues related to skin health, with a commonality of dementia unspecified diagnoses and an average age of onset in the early 70s. The association between skin health and dementia in this cluster may be attributed to age, as both conditions become more prevalent with advancing age. This aligns with findings from previous studies, which reported an increase in the prevalence of actinic keratosis [44] and seborrheic keratosis [45], as well as AD [46], with age. Cluster 2 is marked by early-onset, psychiatric and behavioral manifestations; enrichment in FTD and other less common forms of ADRD; and the earliest age of onset among our clusters, typically in the late 60s. This cluster extends the characterization found in previous studies that described a behavioral symptom subtype in patients [20,21] by demonstrating similar characteristics in a broader population of patients with ADRD. Cluster 3 encompasses a broad array of conditions like respiratory issues and severe diabetes, affecting older patients, more female participants than male participants, with a higher incidence of VCI and dementia unspecified, and an age of onset in the mid-70s. Reflecting the AD subtype identified by previous researchers, this group was characterized as being overall older and having more comorbidities [20]. Landi et al [23] further differentiated patients with AD by onset timing, which we also observed but across a more diverse set of ADRD diagnoses. Notably, our *ICD*-based clustering did not reveal clearly separated clusters in the 2D UMAP projections. While UMAP aims to preserve both local and global relationships when reducing high-dimensional data to a lower-dimensional space, some distortions may inevitably occur during dimensionality reduction. Alternatively, the overlapping clusters could reflect the complexity of comorbid conditions in ADRD, which may not form clearly distinguishable subgroups.

In addition, our analysis of clinical notes revealed 3 distinct subtypes. Cluster 1 featured terms related to psychiatric manifestations and medications, aligning with findings from previous studies [20,21]. Cluster 2 included terms related to cardiovascular and various motor and sensory issues, supported by previous studies that identified subtypes of CVD [20,21] and aligning with the predominant diagnoses of VCI and LBD within this cluster. Cluster 3 covered a wide array of health conditions, consistent with the higher occurrence of rare ADRD diagnoses, which tend to involve more heterogeneous health conditions. Notably, we observed significant overlap between *ICD* and note clusters, identifying 2 ADRD subtypes of interest that were concordant across the 2 data modalities: the first subtype, “psychiatric manifestations,” and the second subtype, “diabetes with cardiovascular or motor issues.” Thus, our analysis delineated 2 distinct ADRD subtypes with specific diagnostic and symptomatic profiles.

Our study identified sex differences across all clinical note clusters with substantial effects observed in note clusters 1 and 2. For example, note cluster 1 was significantly overrepresented in female participants and had a higher prevalence of AD (PR=1.1). It also overlapped significantly (P<.001) with *ICD* cluster 2, which was enriched for psychiatric and behavioral symptoms (eg, apathy). This aligns with Tang et al [25], who reported stronger psychiatric associations in female patients with AD, including greater links to depression. This pattern may be partially attributed to women’s greater likelihood of seeking mental health care [47,48]. In contrast, note cluster 2 was overrepresented in male participants, with higher prevalence of VCI (PR=1.41) and LBD (PR=1.31), consistent with Tang et al [25], who found vascular dementia was more common in male participants. The higher prevalence of VCI in male participants may be related to a greater burden of hypertension, particularly in early life [49]. In addition, the increased representation of male participants with LBD may reflect potential underdiagnosis in female participants [50,51]. Given the clinical impact of sex disparities in ADRD—particularly in AD and LBD [52]—future studies integrating longitudinal data and clinicopathological evidence will be crucial to disentangling biological influences from health care–seeking behaviors.

Another interesting observation relates to variations in age of onset, a key indicator of disease severity, across the identified subtypes in both clinical notes and *ICD*-based clusters, with notable differences in the *ICD*-derived subtypes. For instance, the early-onset *ICD* cluster 2 was enriched with psychiatric disorders and included diseases known with early-onset,

XSL•FO
RenderX

including FTD [53] and other rare ADRD categories, such as PCA [54]. In contrast, the late-onset *ICD* cluster 3, exhibited a higher prevalence of diverse health conditions, including respiratory issues (eg, cough and shortness of breath), chronic pain, musculoskeletal conditions, such as bilateral low back pain and foot pain, as well as diabetes mellitus. While chronic pain is not typically associated with ADRD, pain could indicate the general aging process [55], and relate to osteoporosis and osteoarthritis, which likely contribute to chronic pain in older adults. Other symptoms, such as foot pain and shortness of breath, may reflect comorbidities of diabetes.

Limitations

Our study has a few limitations. First, there are the challenges associated with using real-world EHR data. Differences in how health care providers document information, stemming from variations in training, personal documentation habits, and clinical judgment, may have contributed to inconsistencies. Furthermore, health care use patterns, such as visit regularity, may influence our clustering results. For example, variations in visit frequency could lead to overrepresentation of certain symptom clusters or skewed associations. Future studies adjusting for health care use patterns may help address this limitation. Second, the inclusion of long-term patient histories in clinical notes—where recent notes may capture both current and past symptoms—could introduce extraneous information, making it difficult to isolate content relevant to the latest diagnosis. This mixture of historical and recent data may have diluted the association between documented ADRD diagnoses and their actual clinical significance, leading to observed trends rather than clear associations. Furthermore, the repetition of relevant language across multiple encounters may have influenced the clustering process, potentially reflecting the frequency of patient visits to the memory clinic, rather than clinical characteristics. Third, our study is constrained by the absence of an independent validation cohort to confirm the identified clusters. While the relative overlap in clusters identified through *ICD* codes and note contents, along with the alignment with findings from previous research, offers some validation, the results could be strengthened by applying the same encoding and clustering techniques to an external validation cohort. To ensure external validity, these results need to be validated at other health care institutions. Fourth, another limitation of this study is that our subtyping characterization only focused on ADRD diagnoses based on etiology, but did not address the stage of disease, which clearly affects the neuropsychological profile. This calls for a focus on the heterogeneity of disease stage in future research. Moreover, there may be sex differences in who receives health care at different stages and ages, adding another layer of complexity to our findings. Finally, our study is limited by the lack of racial diversity in the cohort, with 88.6% of participants being White. Given known racial differences in AD incidence, comorbidities, and health care access [56-59]—and particularly the heightened impact of hypertension on AD risk in some minoritized groups [59,60]—our findings may not fully capture the spectrum of ADRD subtypes in these populations. Future studies with broader representation are necessary to improve the generalizability of our subtyping approach.

Future Directions

In future work, the preprocessing of clinical notes could be enhanced by implementing multiple methods, such as medspaCy [61], with a focus on targeting sections most relevant to diagnoses, such as medical history. To further improve the extraction and analysis of pertinent data, the use of emerging LLMs, such as GPT [39], should be explored. In addition, validating an independent dataset and enriching the patient population would help increase the robustness and reliability of the identified ADRD subtypes. To advance this work, we will use a dual-modality approach that leverages both structured and unstructured data sources, such as medications and imaging. A deep autoencoder that uses multiple modalities simultaneously could offer methodological improvements over our current practice of conducting parallel clustering analyses and relying on heuristic averaging of embeddings. Furthermore, explicitly using the temporal or graph properties of EHRs could yield more informative representations, enhancing unsupervised clustering capabilities, as has been shown in prior approaches in a supervised learning setting [62,63]. Ultimately, our goal is to develop machine learning models capable of predicting these ADRD subtypes from real-word health care systems. Such models may aid in more precise diagnostics, prognostics, and the formulation of targeted treatment strategies.

Acknowledgments

This study was funded by National Institute of Aging P30AG062421 (SD and BTH). The funder played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript. The authors acknowledge Yoonchul Shin and Qinglin Gou for annotating the age of onset from clinical notes.

Data Availability

The data used for this study are available from the Mass General Brigham Healthcare System, but restrictions apply to the availability of these data, which were used under permission for this study, and so are not publicly available. Code will be available on GitHub upon publication.

Authors' Contributions

MW was responsible for conceptualization, formal analysis, investigation, methodology, software, and writing original draft. YC was responsible for data curation, formal analysis, investigation, methodology, software, visualization, writing original draft,

XSL•FO
RenderX

review, and editing. YH was responsible for data curation and analysis, review, and editing. YL was responsible for formal analysis, methodology, and software. CM was responsible for conceptualization, review, and editing. BTH was responsible for review and editing. JRD was responsible for data curation, review, and editing. AS-P was responsible for review and editing. DB was responsible for review and editing. SD was responsible for conceptualization, funding acquisition, investigation, methodology, supervision, review, and editing.

Conflicts of Interest

JRD served on a scientific review board for I-Mab Biopharma. All other authors declare no conflicts of interest.

Multimedia Appendix 1
List of diagnosis names to Alzheimer disease and related dementias diagnosis categories. This document provides a list of diagnosis names from the electronic health records of patients at a memory clinic to various Alzheimer disease and related dementias diagnosis categories.
[PDF File (Adobe PDF File), 613 KB - aging_v8i1e65178_app1.pdf]

Multimedia Appendix 2
Visualizations of model attention and embedding representations. This appendix includes attention heatmaps from the final transformer layer across heads, as well as Uniform Manifold Approximation and Projection projections of International Classification of Diseases code, and clinical note embeddings, characterized by sex, age of onset, Alzheimer disease and related dementia diagnosis, and provider information.
[DOCX File , 2067 KB - aging_v8i1e65178_app2.docx]

References

1. Dementia fact sheet. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2024-04-29]
2. Alzheimer's disease fact sheet. National Institute on Aging. URL: https://www.nia.nih.gov/health/alzheimers-and-dementia/alzheimers-disease-fact-sheet [accessed 2024-04-29]
3. Bloom GS. Amyloid-β and tau: the trigger and bullet in Alzheimer disease pathogenesis. JAMA Neurol 2014 Apr 01;71(4):505-508. [doi: 10.1001/jamaneurol.2013.5847] [Medline: 24493463]
4. Stampfer MJ. Cardiovascular disease and Alzheimer's disease: common links. J Intern Med 2006 Sep 26;260(3):211-223 [FREE Full text] [doi: 10.1111/j.1365-2796.2006.01687.x] [Medline: 16918818]
5. Kotzbauer PT, Trojanowsk JQ, Lee VM. Lewy body pathology in Alzheimer's disease. J Mol Neurosci 2001 Oct;17(2):225-232. [doi: 10.1385/jmn:17:2:225] [Medline: 11816795]
6. Irwin DJ, Grossman M, Weintraub D, Hurtig HI, Duda JE, Xie SX, et al. Neuropathological and genetic correlates of survival and dementia onset in synucleinopathies: a retrospective analysis. Lancet Neurol 2017 Jan;16(1):55-65 [FREE Full text] [doi: 10.1016/S1474-4422(16)30291-5] [Medline: 27979356]
7. Ferreira D, Przybelski SA, Lesnick TG, Lemstra AW, Londos E, Blanc F, et al. β-Amyloid and tau biomarkers and clinical phenotype in dementia with Lewy bodies. Neurology 2020 Dec 15;95(24):e3257-e3268 [FREE Full text] [doi: 10.1212/WNL.0000000000010943] [Medline: 32989106]
8. Hardy J, Selkoe DJ. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. Science 2002 Jul 19;297(5580):353-356. [doi: 10.1126/science.1072994] [Medline: 12130773]
9. Mukhopadhyay S, Banerjee D. A primer on the evolution of aducanumab: the first antibody approved for treatment of Alzheimer's disease. J Alzheimers Dis 2021;83(4):1537-1552. [doi: 10.3233/JAD-215065] [Medline: 34366359]
10. Neff RA, Wang M, Vatansever S, Guo L, Ming C, Wang Q, et al. Molecular subtyping of Alzheimer's disease using RNA sequencing data reveals novel mechanisms and targets. Sci Adv 2021 Jan 08;7(2):eabb5398 [FREE Full text] [doi: 10.1126/sciadv.abb5398] [Medline: 33523961]
11. Young AL, Marinescu RV, Oxtoby NP, Bocchetta M, Yong K, Firth NC, Genetic FTD Initiative (GENFI), Alzheimer's Disease Neuroimaging Initiative (ADNI). Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with Subtype and Stage Inference. Nat Commun 2018 Oct 15;9(1):4273 [FREE Full text] [doi: 10.1038/s41467-018-05892-0] [Medline: 30323170]
12. Vogel JW, Young AL, Oxtoby NP, Smith R, Ossenkoppele R, Strandberg OT, Alzheimer's Disease Neuroimaging Initiative, et al. Four distinct trajectories of tau deposition identified in Alzheimer's disease. Nat Med 2021 May;27(5):871-881 [FREE Full text] [doi: 10.1038/s41591-021-01309-6] [Medline: 33927414]
13. Scheltens NM, Tijms BM, Koene T, Barkhof F, Teunissen CE, Wolfsgruber S, et al. Cognitive subtypes of probable Alzheimer's disease robustly identified in four cohorts. Alzheimers Dement 2017 Nov 17;13(11):1226-1236 [FREE Full text] [doi: 10.1016/j.jalz.2017.03.002] [Medline: 28427934]

XSL•FO
RenderX

14. Libon DJ, Drabick DA, Giovannetti T, Price CC, Bondi MW, Eppig J, et al. Neuropsychological syndromes associated with Alzheimer's/vascular dementia: a latent class analysis. J Alzheimers Dis 2014 Sep 16;42(3):999-1014. [doi: 10.3233/JAD-132147] [Medline: 25024329]
15. Scheltens NM, Galindo-Garre F, Pijnenburg YA, van der Vlies AE, Smits LL, Koene T, et al. The identification of cognitive subtypes in Alzheimer's disease dementia using latent class analysis. J Neurol Neurosurg Psychiatry 2016 Mar 17;87(3):235-243. [doi: 10.1136/jnnp-2014-309582] [Medline: 25783437]
16. Alashwal H, El Halaby M, Crouse JJ, Abdalla A, Moustafa AA. The application of unsupervised clustering methods to Alzheimer's disease. Front Comput Neurosci 2019 May 24;13:31 [FREE Full text] [doi: 10.3389/fncom.2019.00031] [Medline: 31178711]
17. Doshi-Velez F, Ge Y, Kohane I. Comorbidity clusters in autism spectrum disorders: an electronic health record time-series analysis. Pediatrics 2014 Jan;133(1):e54-e63 [FREE Full text] [doi: 10.1542/peds.2013-0819] [Medline: 24323995]
18. Luo Y, Eran A, Palmer N, Avillach P, Levy-Moonshine A, Szolovits P, et al. A multidimensional precision medicine approach identifies an autism subtype characterized by dyslipidemia. Nat Med 2020 Sep 10;26(9):1375-1379. [doi: 10.1038/s41591-020-1007-0] [Medline: 32778826]
19. Zhang X, Chou J, Liang J, Xiao C, Zhao Y, Sarva H, et al. Data-driven subtyping of Parkinson's disease using longitudinal clinical records: a cohort study. Sci Rep 2019 Jan 28;9(1):797 [FREE Full text] [doi: 10.1038/s41598-018-37545-z] [Medline: 30692568]
20. Xu J, Wang F, Xu Z, Adekkanattu P, Brandt P, Jiang G, et al. Data-driven discovery of probable Alzheimer's disease and related dementia subphenotypes using electronic health records. Learn Health Syst 2020 Oct 10;4(4):e10246 [FREE Full text] [doi: 10.1002/lrh2.10246] [Medline: 33083543]
21. Alexander N, Alexander DC, Barkhof F, Denaxas S. Using unsupervised learning to identify clinical subtypes of Alzheimer's disease in electronic health records. Stud Health Technol Inform 2020 Jun 16;270:499-503. [doi: 10.3233/SHTI200210] [Medline: 32570434]
22. Alexander N, Alexander DC, Barkhof F, Denaxas S. Identifying and evaluating clinical subtypes of Alzheimer's disease in care electronic health records using unsupervised machine learning. BMC Med Inform Decis Mak 2021 Dec 08;21(1):343 [FREE Full text] [doi: 10.1186/s12911-021-01693-6] [Medline: 34879829]
23. Landi I, Glicksberg BS, Lee HC, Cherng S, Landi G, Danieletto M, et al. Deep representation learning of electronic health records to unlock patient stratification at scale. NPJ Digit Med 2020;3:96 [FREE Full text] [doi: 10.1038/s41746-020-0301-z] [Medline: 32699826]
24. He Z, Tian S, Erdengasileng A, Charness N, Bian J. Temporal subtyping of Alzheimer's disease using medical conditions preceding Alzheimer's disease onset in electronic health records. AMIA Jt Summits Transl Sci Proc 2022;2022:226-235 [FREE Full text] [Medline: 35854753]
25. Tang AS, Oskotsky T, Havaldar S, Mantyh WG, Bicak M, Solsberg CW, et al. Deep phenotyping of Alzheimer's disease leveraging electronic medical records identifies sex-specific clinical associations. Nat Commun 2022 Feb 03;13(1):675 [FREE Full text] [doi: 10.1038/s41467-022-28273-0] [Medline: 35115528]
26. Tang AS, Rankin KP, Cerono G, Miramontes S, Mills H, Roger J, et al. Leveraging electronic health records and knowledge networks for Alzheimer's disease prediction and sex-specific biological insights. Nat Aging 2024 Mar 21;4(3):379-395 [FREE Full text] [doi: 10.1038/s43587-024-00573-8] [Medline: 38383858]
27. Vaswani A, Brain G, Shazeer N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. arXiv Preprint posted online on June 12, 2017 [FREE Full text]
28. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2019 Presented at: NAACL-HLT '19; June 2-7, 2019; Minneapolis, MN p. 4171-4186 URL: https://aclanthology.org/N19-1423.pdf
29. Alsentzer E, Murphy JR, Boag WH, Jindi D, Naumann T, McDermott M. Publicly available clinical BERT embeddings. In: Proceedings of the 2nd Clinical Natural Language Processing Workshop. 2019 Presented at: ClinicalNLP '19; June 7, 2019; Minneapolis, MN p. 72-78 URL: https://aclanthology.org/W19-1909.pdf [doi: 10.18653/v1/w19-1909]
30. Johnson AE, Pollard TJ, Shen L, Lehman LH, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. Sci Data 2016 May 24;3:160035 [FREE Full text] [doi: 10.1038/sdata.2016.35] [Medline: 27219127]
31. Choi Y, Chiu CY, Sontag D. Learning low-dimensional representations of medical concepts. AMIA Jt Summits Transl Sci Proc 2016;2016:41-50 [FREE Full text] [Medline: 27570647]
32. Shannon CE. A mathematical theory of communication. Bell Syst Tech J 1948;27:63. [doi: 10.1002/j.1538-7305.1948.tb00917.x]
33. Ebrahimi N, Pflughoeft K, Soofi ES. Two measures of sample entropy. Stat Probabil Lett 1994 Jun;20(3):225-234. [doi: 10.1016/0167-7152(94)90046-9]
34. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods 2020 Mar 3;17(3):261-272 [FREE Full text] [doi: 10.1038/s41592-019-0686-2] [Medline: 32015543]

XSL•FO
RenderX

35. Abraham A, Pedregosa F, Eickenberg M, Gervais P, Mueller A, Kossaifi J, et al. Machine learning for neuroimaging with scikit-learn. Front Neuroinform 2014;8:14 [FREE Full text] [doi: 10.3389/fninf.2014.00014] [Medline: 24600388]
36. Flamary R, Courty N. POT python optimal transport library. GitHub. URL: https://pythonot.github.io/ [accessed 2024-04-29]
37. Courty N, Flamary R, Tuia D, Rakotomamonjy A. Optimal Transport for Domain Adaptation. IEEE Trans Pattern Anal Mach Intell 2017 Sep 1;39(9):1853-1865. [doi: 10.1109/tpami.2016.2615921]
38. Grootendorst M. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv Preprint posted online on March 11, 2022 [FREE Full text]
39. Achiam OJ, Adler S, Agarwal S, Ahmad L, Akkaya I, Aleman FL, et al. GPT-4 technical report. arXiv Preprint posted online on March 15, 2023 [FREE Full text]
40. Reynolds D. Gaussian mixture models. In: Li SZ, Jain A, editors. Encyclopedia of Biometrics. Cham, Switzerland: Springer; 2009:659-663.
41. Ester M, Kriegel HP, Sander J, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining. 1996 Presented at: KDD '96; August 2-4, 1996; Portland, OR p. 226-231 URL: https://dl.acm.org/doi/10.5555/3001460.3001507
42. Ribino P, Napoli CD, Paragliola G, Serino L, Gasparini F, Chicco D. Exploratory analysis of longitudinal data of patients with dementia through unsupervised techniques. In: Proceedings of the 4th Italian Workshop on Artificial Intelligence for an Ageing Society. 2023 Presented at: AIxAS '23; November 6-9, 2023; Rome, Italy p. 1-21 URL: https://ceur-ws.org/Vol-3623/AIxAS_2023_paper_8.pdf
43. Meng W, Xu J, Huang Y, Wang C, Song Q, Ma A, et al. Autoencoder to identify sex-specific sub-phenotypes in Alzheimer's disease progression using longitudinal electronic health records. medRxiv Preprint posted online on July 11, 2024 [FREE Full text] [doi: 10.1101/2024.07.07.24310055] [Medline: 39040206]
44. Yaldiz M. Prevalence of actinic keratosis in patients attending the dermatology outpatient clinic. Medicine (Baltimore) 2019 Jul;98(28):e16465 [FREE Full text] [doi: 10.1097/MD.0000000000016465] [Medline: 31305480]
45. Sun MD, Halpern AC. advances in the etiology, detection, and clinical management of seborrheic keratoses. Dermatology 2022 Jul 26;238(2):205-217. [doi: 10.1159/000517070] [Medline: 34311463]
46. Alzheimer's disease facts and figures. Alzheimer's Association. URL: https://www.alz.org/alzheimers-dementia/facts-figures?gad_source=1 [accessed 2024-04-29]
47. Lim MT, Fong Lim YM, Tong SF, Sivasampu S. Age, sex and primary care setting differences in patients' perception of community healthcare seeking behaviour towards health services. PLoS One 2019 Oct 21;14(10):e0224260 [FREE Full text] [doi: 10.1371/journal.pone.0224260] [Medline: 31634373]
48. Susukida R, Mojtabai R, Mendelson T. Sex differences in help seeking for mood and anxiety disorders in the national comorbidity survey-replication. Depress Anxiety 2015 Nov 22;32(11):853-860. [doi: 10.1002/da.22366] [Medline: 25903117]
49. Connelly PJ, Currie G, Delles C. Sex Differences in the prevalence, outcomes and management of hypertension. Curr Hypertens Rep 2022 Jun 07;24(6):185-192 [FREE Full text] [doi: 10.1007/s11906-022-01183-8] [Medline: 35254589]
50. Bayram E, Coughlin DG, Rajmohan R, Litvan I. Sex differences for clinical correlates of substantia nigra neuron loss in people with Lewy body pathology. Biol Sex Differ 2024 Jan 19;15(1):8 [FREE Full text] [doi: 10.1186/s13293-024-00583-6] [Medline: 38243325]
51. Bayram E, Coughlin DG, Koga S, Ross OA, Litvan I, Dickson DW. Sex differences for regional pathology in people with a high likelihood of Lewy body dementia phenotype based on underlying pathology. Alzheimers Dement (Amst) 2025;17(1):e70083. [doi: 10.1002/dad2.70083] [Medline: 39886324]
52. Wei H, Masurkar AV, Razavian N. On gaps of clinical diagnosis of dementia subtypes: a study of Alzheimer's disease and Lewy body disease. Front Aging Neurosci 2023 Mar 21;15:1149036 [FREE Full text] [doi: 10.3389/fnagi.2023.1149036] [Medline: 37025965]
53. Lee SE. Frontotemporal dementia: clinical features and diagnosis. UpToDate. URL: https://www.uptodate.com/contents/frontotemporal-dementia-clinical-features-and-diagnosis [accessed 2024-04-29]
54. Chapleau M, La Joie R, Yong K, Agosta F, Allen IE, Apostolova L, et al. Demographic, clinical, biomarker, and neuropathological correlates of posterior cortical atrophy: an international cohort study and individual participant data meta-analysis. Lancet Neurol 2024 Feb;23(2):168-177 [FREE Full text] [doi: 10.1016/S1474-4422(23)00414-3] [Medline: 38267189]
55. Cheng Y, Ho E, Weintraub S, Rentz D, Gershon R, Das S, et al. Predicting brain amyloid status using the National Institute of Health Toolbox (NIHTB) for assessment of neurological and behavioral function. J Prev Alzheimers Dis 2024;11(4):943-957. [doi: 10.14283/jpad.2024.77] [Medline: 39044505]
56. Adkins-Jackson PB, Kraal AZ, Hill-Jarrett TG, George KM, Deters KD, Besser LM, et al. Riding the merry-go-round of racial disparities in ADRD research. Alzheimers Dement 2023 Oct;19(10):4735-4742. [doi: 10.1002/alz.13359] [Medline: 37394968]
57. Mielke MM, Aggarwal NT, Vila-Castelar C, Agarwal P, Arenaza-Urquijo EM, Brett B, et al. Consideration of sex and gender in Alzheimer's disease and related disorders from a global perspective. Alzheimer's & Dementia 2022 Dec 08;18(12):2707-2724 [FREE Full text] [doi: 10.1002/alz.12662] [Medline: 35394117]

XSL•FO
RenderX

58. Hernandez S, McClendon MJ, Zhou XH, Sachs M, Lerner AJ. Pharmacological treatment of Alzheimer's disease: effect of race and demographic variables. J Alzheimers Dis 2010 Jan 07;19(2):665-672 [FREE Full text] [doi: 10.3233/JAD-2010-1269] [Medline: 20110610]
59. Steenland K, Tan Y, Wingo T, Shi L, Xiao S, Wharton W. The effect of race and co-morbidities on Alzheimer's disease based on Medicare data. Alzheimers Dement 2023 May;19(5):1858-1864 [FREE Full text] [doi: 10.1002/alz.12838] [Medline: 36327171]
60. Akushevich I, Kolpakov S, Yashkin AP, Kravchenko J. Vulnerability to hypertension is a major determinant of racial disparities in Alzheimer's disease risk. Am J Hypertens 2022 Aug 01;35(8):745-751 [FREE Full text] [doi: 10.1093/ajh/hpac063] [Medline: 35581146]
61. Eyre H, Chapman AB, Peterson KS, Shi J, Alba PR, Jones MM, et al. Launching into clinical space with medspaCy: a new clinical text processing toolkit in Python. AMIA Annu Symp Proc 2021;2021:438-447 [FREE Full text] [Medline: 35308962]
62. Choi E, Xu Z, Li Y, Dusenberry M, Flores G, Xue E, et al. Learning the graphical structure of electronic health records with graph convolutional transformer. In: Proceedings of the 34th AAAI Conference on Artificial Intelligence. 2020 Apr 03 Presented at: AAAI '20; February 7-12, 2020; New York, NY p. 606-613. [doi: 10.1609/aaai.v34i01.5400]
63. Zhu W, Razavian N. Variationally regularized graph-based representation learning for electronic health records. In: Proceedings of the 2021 Conference on Health, Inference, and Learning. 2021 Presented at: CHIL '21; April 8-10, 2021; Virtual Event p. 1-13 URL: https://dl.acm.org/doi/10.1145/3450439.3451855 [doi: 10.1145/3450439.3451855]

Abbreviations

AD: Alzheimer disease
ADRD: Alzheimer disease and related dementia
Aβ: amyloid beta
BERT: Bidirectional Encoder Representations from Transformers
CVD: cerebrovascular disease
DBSCAN: density-based spatial clustering of applications with noise
EHR: electronic health record
FDR: false discovery rate
FTD: frontotemporal dementia
ICD: International Classification of Diseases
LBD: Lewy body dementia
LLM: large language model
MGH: Massachusetts General Hospital
MIMIC: Medical Information Mart for Intensive Care
PCA: posterior cortical atrophy
PR: prevalence ratio
TF-IDF: term frequency–inverse document frequency
UMAP: Uniform Manifold Approximation and Projection
VCI: vascular cognitive impairment

Edited by D Liu; submitted 07.08.24; peer-reviewed by R Zapata, D Chrimes, A Chaturvedi; comments to author 29.01.25; revised version received 19.02.25; accepted 10.03.25; published 31.03.25.

Please cite as:
West M, Cheng Y, He Y, Leng Y, Magdamo C, Hyman BT, Dickson JR, Serrano-Pozo A, Blacker D, Das S
Unsupervised Deep Learning of Electronic Health Records to Characterize Heterogeneity Across Alzheimer Disease and Related Dementias: Cross-Sectional Study
JMIR Aging 2025;8:e65178
URL: https://aging.jmir.org/2025/1/e65178
doi:10.2196/65178
PMID:

©Matthew West, You Cheng, Yingnan He, Yu Leng, Colin Magdamo, Bradley T Hyman, John R Dickson, Alberto Serrano-Pozo, Deborah Blacker, Sudeshna Das. Originally published in JMIR Aging (https://aging.jmir.org), 31.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR

XSL•FO
RenderX

Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Correction: Machine Learning Models for Frailty Classification of Older Adults in Northern Thailand: Model Development and Validation Study

Natthanaphop Isaradech[1,2*], MD; Wachiranun Sirikul[1,3,4*], MD, MSc; Nida Buawangpong[5], MD; Penprapa Siviroj[1], DrPH; Amornphat Kitro[1,2], MD, MCTM

[1]Department of Community Medicine, Faculty of Medicine, Chiang Mai University, 110, Intrawarorot Road, Meaung, Thailand

[2]Environmental and Occupational Medicine Excellence Center, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[3]Center of Data Analytics and Knowledge Synthesis for Health Care, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[4]Department of Biomedical Informatics and Clinical Epidemiology, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[5]Department of Family Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[*]these authors contributed equally

Corresponding Author:
Wachiranun Sirikul, MD, MSc
Department of Community Medicine, Faculty of Medicine, Chiang Mai University, 110, Intrawarorot Road, Meaung, Thailand

Related Article:

Correction of: https://aging.jmir.org/2025/1/e62942

(JMIR Aging 2025;8:e75690) doi:10.2196/75690

KEYWORDS

aged care; gerontology; geriatric; old; aging; clinical decision support; delivering health information and knowledge to the public; diagnostic systems; digital health; epidemiology; surveillance; diagnosis; frailty; machine learning; prediction; predictive; AI; artificial intelligence; Thailand; community dwelling; health care intervention; patient care

In “Machine Learning Models for Frailty Classification of Older Adults in Northern Thailand: Model Development and Validation Study” (JMIR Aging 2025;8:e62942) one error was noted.

Reference 44 was a duplicate of reference 36, which reads as follows:

Thinuan P, Siviroj P, Lerttrakarnnon P, Lorga T. Prevalence and potential predictors of frailty among community-dwelling older persons in northern Thailand: a cross-sectional study. Int J Environ Res Public Health. Jun 8, 2020;17(11):4077.

Reference 44 has therefore been removed, and all subsequent references have been reordered accordingly.

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 10.04.25; this is a non–peer-reviewed article;accepted 10.04.25; published 22.04.25.

Please cite as:
Isaradech N, Sirikul W, Buawangpong N, Siviroj P, Kitro A
Correction: Machine Learning Models for Frailty Classification of Older Adults in Northern Thailand: Model Development and Validation Study
JMIR Aging 2025;8:e75690
URL: https://aging.jmir.org/2025/1/e75690
doi:10.2196/75690

XSL•FO
RenderX

© Natthanaphop Isaradech, Wachiranun Sirikul, Nida Buawangpong, Penprapa Siviroj, Amornphat Kitro. Originally published in JMIR Aging (https://aging.jmir.org), 22.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Correction: Internet-Based Supportive Interventions for Family Caregivers of People With Dementia: Randomized Controlled Trial

Yanhong Xie[1*], MSc; Shanshan Shen[1*], MSc; Caixia Liu[2], BSc; Hong Hong[3], BSc; Huilan Guan[1], BSc; Jingmei Zhang[1], BSc; Wanqi Yu[4], BSc

[1]Geriatrics Department, Zhejiang Hospital, Hangzhou, China

[2]Nursing Department, Zhejiang Hospital, Hangzhou, China

[3]General Surgery, Zhejiang Hospital, Hangzhou, Zhejiang Province, China

[4]The Medical Record Statistics Department, Zhejiang Hospital, Hangzhou, China

[*]these authors contributed equally

Corresponding Author:
Hong Hong, BSc
General Surgery, Zhejiang Hospital, Hangzhou, Zhejiang Province, China

Related Article:

Correction of: https://aging.jmir.org/2024/1/e50847

(JMIR Aging 2025;8:e69493) doi:10.2196/69493

In "Internet-Based Supportive Interventions for Family Caregivers of People With Dementia: Randomized Controlled Trial" (JMIR Aging. 2024 Oct 4:7:e50847. doi: 10.2196/50847) the authors made three corrections.

The following funding information has been added to the end of the "Acknowledgements" section:

> *This work was supported by grants from the National Key Research and Development Program of China (program number: 2019YFE0113100), the Zhejiang Provincial Medicine and Health Technology Project (grant number: 2019KY003), and the Zhejiang Hospital 3060 Excellent Young Talents Training Project.*

Additionally, the authors would like to specify equal contribution between Yanhong Xie and Shanshan Shen as co-first authors, and Hong Hong and Caixia Liu as co-corresponding authors. The latter will be specified at the beginning of the "Acknowledgements" section as follows:

> *CL is the co-corresponding author on this work, and can be reached at the following email address: zjyyhlb2007@126.com.*

Finally, the following ORCID has been added to author Hong Hong:

> *0009-0001-1372-1670*

The corrections will appear in the online version of the paper on the JMIR Publications website on January 29, 2025, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 01.12.24; this is a non–peer-reviewed article;accepted 04.12.24; published 29.01.25.

Please cite as:
Xie Y, Shen S, Liu C, Hong H, Guan H, Zhang J, Yu W
Correction: Internet-Based Supportive Interventions for Family Caregivers of People With Dementia: Randomized Controlled Trial
JMIR Aging 2025;8:e69493
URL: https://aging.jmir.org/2025/1/e69493
doi:10.2196/69493

© Yanhong Xie, Shanshan Shen, Caixia Liu, Hong Hong, Huilan Guan, Jingmei Zhang, Wanqi Yu. Originally published in JMIR Aging (https://aging.jmir.org), 29.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction

XSL•FO
RenderX

in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Correction: Changes in Sarcopenia Status and Subsequent Cardiovascular Outcomes: Prospective Cohort Study

Yuanyue Zhu[1,2*], MD; Kan Wang[3,4*], PhD; Zuolin Lu[5], PhD; Feika Li[1,2], MD, PhD; Yu Xu[3,4], MD, PhD; Linhui Shen[1,2], MD, PhD; Yufang Bi[3,4], Prof Dr; Weiguo Hu[1,2], Prof Dr

[1]Department of Geriatrics, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[2]Medical Center on Aging, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[3]Department of Endocrine and Metabolic Diseases, Shanghai Institute of Endocrine and Metabolic Diseases, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[4]National Clinical Research Center for Metabolic Diseases (Shanghai), Key Laboratory for Endocrine and Metabolic Diseases of the National Health Commission, National Research Center for Translational Medicine, State Key Laboratory of Medical Genomics, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[5]School of Population Medicine and Public Health, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

*these authors contributed equally

Corresponding Author:
Weiguo Hu, Prof Dr
Department of Geriatrics
Ruijin Hospital, Shanghai Jiao Tong University School of Medicine
No. 197 Ruijin Second Road, Shanghai
Shanghai, 200025
China
Phone: 86 02164370045
Email: wghu@rjh.com.cn

Related Article:

Correction of: https://aging.jmir.org/2025/1/e69860

(JMIR Aging 2025;8:e84642) doi:10.2196/84642

In "Changes in Sarcopenia Status and Subsequent Cardiovascular Outcomes: Prospective Cohort Study" (JMIR Aging 2025;8:e69860) the authors noted one error.

In the originally published manuscript, the first author's name appeared as "Yuanyuan Zhu" and their ORCID ID appeared as "0009-0007-4341-9468".

The author's name has been corrected to read as "Yuanyue Zhu", and the ORCID ID was corrected to "0009-0001-6704-7036".

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 23.09.25; this is a non–peer-reviewed article;accepted 24.09.25; published 17.10.25.

Please cite as:
Zhu Y, Wang K, Lu Z, Li F, Xu Y, Shen L, Bi Y, Hu W
Correction: Changes in Sarcopenia Status and Subsequent Cardiovascular Outcomes: Prospective Cohort Study
JMIR Aging 2025;8:e84642
URL: https://aging.jmir.org/2025/1/e84642
doi:10.2196/84642
PMID:

XSL•FO
RenderX

©Yuanyue Zhu, Kan Wang, Zuolin Lu, Feika Li, Yu Xu, Linhui Shen, Yufang Bi, Weiguo Hu. Originally published in JMIR Aging (https://aging.jmir.org), 17.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Correction: Machine Learning Approach for Frailty Detection in Long-Term Care Using Accelerometer-Measured Gait and Daily Physical Activity: Model Development and Validation Study

Xiaoping Zheng[1*], PhD; Ziwei Zeng[1*], MSc; Kimberley S van Schooten[2,3], PhD; Yijian Yang[1,4], MD, PhD

[1]Department of Sports Science and Physical Education, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

[2]Neuroscience Research Australia, University of New South Wales, Sydney, Australia

[3]School of Population Health, University of New South Wales, Sydney, Australia

[4]CUHK Jockey Club Institute of Ageing, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

[*]these authors contributed equally

Corresponding Author:
Yijian Yang, MD, PhD
Department of Sports Science and Physical Education
The Chinese University of Hong Kong
G07 Kwok Sports Building. University Ave. The Chinese University of Hong Kong. N.T.
Hong Kong
China (Hong Kong)
Phone: 852 39434001
Fax: 852 26035781
Email: yyang@cuhk.edu.hk

Related Article:

Correction of: https://aging.jmir.org/2025/1/e77140

(JMIR Aging 2025;8:e85173) doi:10.2196/85173

In "Machine Learning Approach for Frailty Detection in Long-Term Care Using Accelerometer-Measured Gait and Daily Physical Activity: Model Development and Validation Study" (JMIR Aging 2025;1(8) e77140) the authors made a few corrections.

In the Results section of the Abstract, the sentence:

> *Explainable artificial intelligence analysis revealed that older adults with frailty exhibited more variable, complex, and asymmetric gait patterns, which were characterized by higher stride length variability, increased sample entropy, and a higher gait symmetry score.*

Has been revised to:

> *Explainable artificial intelligence analysis revealed that older adults with frailty exhibited more variable, complex, and asymmetric gait patterns, which were characterized by higher stride length variability, increased sample entropy, and a lower gait symmetry score.*

In the Results section, the sentence:

> *Higher gait symmetry score (indicating less symmetry; Multimedia Appendix 5) in the vertical direction contributes to the frail group, while lower values are associated with a nonfrailty classification.*

Has been revised to:

> *Higher gait symmetry score (indicating more symmetry; Multimedia Appendix 5) in the vertical direction contributes to the frail group, while lower values are associated with a nonfrailty classification.*

In the Results section, the sentence:

> *Compared with the nonfrail group, the frail group exhibited higher stride length variability, greater sample entropy in the vertical direction, and increased gait symmetry score in the vertical direction.*

Has been revised to:

> *Compared with the nonfrail group, the frail group exhibited higher stride length variability, greater sample entropy in the vertical direction, and decreased gait symmetry score in the vertical direction.*
>
> *In the Multimedia Appendix 5, the formula:*



Has been revised to:

XSL•FO
RenderX



The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to full-text repositories, the corrected article has also been resubmitted to those repositories.

Multimedia Appendix 1
Gait symmetry score calculation (Revised).
[DOCX File , 14 KB - aging_v8i1e85173_app5.docx]

Submitted 02.10.25; this is a non–peer-reviewed article;accepted 06.10.25; published 29.10.25.

Please cite as:
Zheng X, Zeng Z, van Schooten KS, Yang Y
Correction: Machine Learning Approach for Frailty Detection in Long-Term Care Using Accelerometer-Measured Gait and Daily Physical Activity: Model Development and Validation Study
JMIR Aging 2025;8:e85173
URL: https://aging.jmir.org/2025/1/e85173
doi:10.2196/85173
PMID:

©Xiaoping Zheng, Ziwei Zeng, Kimberley S van Schooten, Yijian Yang. Originally published in JMIR Aging (https://aging.jmir.org), 29.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Reconstruction of a Risk Analysis Tool That Uses Nursing Data for Frailty Assessment Among Older Adults: Derivation and Internal Evaluation Study

Brian G Celso[1], PhD; Sarah E Ser[2], BS, MS, PhD; Urszula A Snigurska[3], PhD; Laurence M Solberg[3,4], MD; Laurie J Duckworth[3], PhD; Elizabeth M Dunn[5], PhD; Mattia Prosperi[6], PhD; Robert J Lucero[7], MPH, PhD; Ragnhildur I Bjarnadottir[3], MPH, PhD

[1]Department of Surgery, UF Health Jacksonville, Jacksonville, FL, United States

[2]Department of Epidemiology, College of Public Health and Health Professions and College of Medicine, University of Florida, Gainesville, FL, United States

[3]College of Nursing, University of Florida, 1225 Center Drive, PO Box 100197, Gainesville, FL, United States

[4]Geriatrics Research, Education, and Clinical Center, North Florida/South Georgia Veterans Health System, Veterans Health Administration, Gainesville, FL, United States

[5]Clinical and Translational Science Institute, University of Florida, Gainesville, FL, United States

[6]Department of Epidemiology, College of Public Health and Health Professions, University of Florida, Gainesville, FL, United States

[7]School of Nursing, UCLA, Los Angeles, CA, United States

Corresponding Author:
Ragnhildur I Bjarnadottir, MPH, PhD
College of Nursing, University of Florida, 1225 Center Drive, PO Box 100197, Gainesville, FL, United States

Abstract

Background: Frailty screening for older adults is of particular importance for those with declining health and social risk factors. However, many existing frailty tools do not offer automated appraisal in clinical settings due to challenges in data collection and the complexity of current approaches. Although routine frailty screening is inconsistently implemented, elements of frailty are captured in the electronic health record (EHR) from hospital admissions data. Thus, further adjustments and adaptations are required to correctly identify frailty.

Objective: This study investigated existing frailty-related EHR data at two hospitals within a single health system to (1) identify key data elements and establish their availability in existing clinical workflows, and (2) use these data elements to reconstruct a validated and widely used frailty index.

Methods: This developmental study included encounters of older adult patients (aged ≥65 years) admitted to medical-surgical units at 2 academic hospitals in North Florida between January 2012 and May 2021. EHR data were used to reconstruct a frailty index modeled after the Risk Analysis Index. Optimal cut points for frailty classification were determined through receiver operating characteristic analysis. Multiple logistic regression models were compared to evaluate predictive performance for hospital mortality, and component importance was assessed by sequentially removing each frailty parameter from the comprehensive model. Age-stratified analyses were performed to evaluate the robustness of classifications across age groups and race/ethnicity. All regression models were estimated to address rare outcome events.

Results: A total of 10,863 hospital patients (45.5% male; mean age 75.4, SD 7.7 years) were included. Using optimal cut points, patients were classified as not frail (40.6%), prefrail (43.5%), or frail (14.9%), with corresponding mortality rates of 0.45%, 0.82%, and 3.24%. After adjustment for confounders, frail patients had significantly higher odds of mortality compared to not frail patients (odds ratio 7.14, 95% CI 4.30 - 11.8; P<.001). Component importance analysis identified shortness of breath (area under the receiver operating characteristic curve reduction: 0.051) and functional status (area under the receiver operating characteristic curve reduction: 0.047) as the strongest predictors. Age-stratified analyses confirmed the classification system's validity, with frail patients aged 65 to 75 years (mortality 3.97%) having higher mortality risk than not frail patients aged ≥86 years (mortality 0.48%).

Conclusions: This study developed and validated a 3-category frailty classification system with strong predictive validity for hospital mortality among older adults. The generated frailty index leverages existing EHR data to capture frailty without increasing provider workload or interfering with workflow. Component importance analysis identified respiratory dysfunction and functional limitations as key predictors. This automated approach to frailty assessment could improve risk stratification, providing critical

XSL•FO
RenderX

information to clinical teams with minimal burden and supporting clinical decision-making for hospitalized older adults. Future study of the overall reliability and validity of these derived frailty scores is warranted.

(JMIR Aging 2025;8:e67888) doi:10.2196/67888

KEYWORDS

frailty; risk assessment; nursing; clinical informatics; geriatrics

Introduction

Frailty, a geriatric syndrome characterized by increased vulnerability to adverse health outcomes, poses a significant challenge in the current health care system [1]. Recent guidelines from both the American Academy of Family Physicians and the International Conference of Frailty and Sarcopenia Research emphasize the importance of widespread screening for adults aged ≥65 years, particularly those with specific health conditions or social risk factors [2]. Routine frailty screening has been used in some preoperative settings [3]; however, it has been sporadic and inconsistently implemented in most acute care settings. Therefore, there is a need to improve and systematize the process of identifying frailty in these clinical settings.

Numerous tools have been developed to assess frailty. Some, such as the frailty phenotype [4], focus on diagnosing frailty, whereas others, such as the Frailty Index [5] and the Risk Analysis Index (RAI) [6,7], aim to identify risk factors for frailty and assess the risk of adverse events such as mortality, disability, or hospital readmissions. However, agreement among different frailty tools is only fair to moderate [8]. Despite the availability of validated tools, their implementation within acute care settings remains limited. This is largely due to challenges related to data collection and the complexity of existing approaches. The key challenge involves the reliance of many existing frailty assessment tools on primary assessments that are time-consuming and may not be feasible in acute care settings, especially with critically ill patients [9]. These tools often require data beyond routine documentation, which can be burdensome for already overworked health care professionals. Any future frailty assessment tool must easily integrate with the routine clinical workflow to facilitate the uptake of the tool by health care professionals.

In the context of modern digitalized collection of health data, frailty assessment in acute care settings can be facilitated by capturing the data that are already routinely collected as part of administrative operations and clinical workflows [10]. For example, in the United States, frailty is captured as a diagnosis under the Centers for Medicare and Medicaid Services. Kim et al [11,12] showed that a Medicare claims-based frailty index predicted disability, mobility impairment, recurrent falls, and number of skilled nursing facility days. Despite support for the validity of a claims-based frailty index, the utility of claims-based tools is limited to retrospective analyses given the delay in data availability. Among older adults in particular, an accurate and reliable way to assess frailty is essential to meet the projected demographic shifts in both the number and proportion of world population that are caused by increased longevity [13] and for the US context, the transition of baby boomers to retirement age [14]. This limitation is particularly problematic in view of the increasing emphasis on precision or personalized medicine and the need for rapid clinical decision-making in real or near real time among hospitalized patients.

Owing to these modern information technology demands and demographic pressures, a novel, digital tool is needed to assess the risk of frailty in hospitalized older adults. Therefore, any future frailty assessment tool must leverage the readily available data from electronic health records (EHRs) that capture patient data for 84% of non–federal acute care hospitals in the United States [15]. EHRs include a wide array of data about a patient and may be sufficient sources of all information that is necessary for a *hands-off* frailty assessment.

This study addresses the limitations of existing frailty assessment tools in acute care by offering a data-driven approach to frailty risk assessment. The long-term goal is to develop a prognostic index for frailty that automates the process of frailty risk assessment based on EHR data collected through routine acute care. The proposed tool is entitled “The Automated Frailty Index Tool” (A-FIT). A-FIT leverages concepts from existing, validated frailty assessment tools and operationalizes these concepts based on expertise from clinical practice. The primary aim of this research was to acquire accuracy and clinical decision support comparable to that of existing frailty index tools while substantially reducing assessment and documentation burdens. The findings will inform the development of a robust and clinically actionable frailty assessment and decision support tool for hospitalized older adults, contributing to the advancement of precision medicine within acute care.

Developing and validating a prediction model for a robust risk assessment tool is a multistep process [16-18]. For an EHR data-driven tool, this may involve substantial workflow assessments and multiple iterations of operationalization and calibration of predictors. This study serves as a critical first step in the development process. The objectives of this study were to investigate existing frailty-related EHR data at 2 hospitals within a single health system to identify key data elements and establish their availability in existing clinical workflows and to use these data elements to reconstruct a validated and widely used frailty index.

Methods

Study Design

This was a development study that used secondary data from a large retrospective observational study of EHR data. The study included encounters of adult patients (aged ≥65 years) admitted to a medical/surgical unit at 1 of 2 academic hospitals in North Florida between January 2012 and May 2021. The larger study

XSL•FO
RenderX

examined iatrogenic conditions among older adults in the acute care setting [19].

Ethical Considerations

This study was conducted under a University of Florida IRB protocol (IRB201900208) designated as exempt under Category 4(iii) for secondary research using identifiable health information. Under this determination, informed consent was waived in accordance with 45 CFR 46.104(d)(4)(iii) and HIPAA provisions for research (45 CFR 164.501 and 164.512[b]). All data were analyzed within a secure, HIPAA-compliant environment, with no direct patient contact and no identifiable information disclosed. Procedures adhered to institutional and national guidelines and the principles of the Declaration of Helsinki. Standard technical and administrative safeguards were in place to protect patient privacy and ensure compliance with institutional and federal data security requirements.

Conceptual Framework

The first step in the development process was informed by the Practice-Based Knowledge Discovery for Comparative Effectiveness Research framework proposed by Lucero and Bakken [20]. The framework, grounded in the Quality Health Outcomes Model [21] and the Knowledge Discovery in Databases [22] process, guides the extraction of actionable insights from electronic clinical data. This framework was chosen for its iterative approach to engaging practice-based experts in informatics research. A practice-based framework was crucial for developing a frailty index that not only demonstrated statistical and computational rigor but also possessed clear clinical utility. To this end, an interdisciplinary team of clinical experts was convened, including nurses, physicians, and physical therapists, as part of an iatrogenic conditions task force established to conduct the large retrospective observational study described earlier.

Analytic Sample

In alignment with existing frailty screening recommendations, the analytic sample for this study included only hospital patients aged ≥65 years. The original dataset consisted of 60,475 inpatients. Patients with missing data for any of the 11 frailty parameters, including patients without any documented assessments of 1 of the 4 ADLs or those for whom living status (independent or dependent) could not be clearly established from the recorded residence (eg,“country clinic,” “unable to assess,” and “homeless”) were excluded. Furthermore, 1 patient was excluded because their documented age was outside the plausible range for age (ie, ≥116 years). This can occur during admissions to the emergency room, for example, in the case of major traumatic injury, where information could not be collected directly from a patient or their family member. The resulting analytical sample consisted of 10,863 patients with complete data for all frailty parameters. Among these patients, 113 (1.04%) experienced in-hospital mortality, which served as the primary outcome for model development and validation.

Frailty Assessment Tool Construction

The second step in the development of the A-FIT involved a collaboration with the clinical expert task force to identify variables that were considered the most clinically meaningful representations of the RAI parameters and were expected to be widely and consistently documented within the medical/surgical setting. Through feedback from the team of clinical experts, important considerations were given to the stated goal that the tool enabled continuous assessment of frailty status across hospital admissions. To that end, 11 RAI frailty parameters were identified by the clinical task force.

For the next phase of development, the RAI was selected for comparison because the tool has an established validity across both surgical and nonsurgical settings. The RAI was developed and validated within the United States Department of Veterans Affairs (VA) health care system by adapting the Minimum Data Set Mortality Risk Index–Revised [23]. The RAI includes 11 assessment variables addressing various demographic, physiological, cognitive, and functional domains. As such, the RAI includes a broader range of factors beyond the 5 criteria of the frailty phenotype [4], but less than the Frailty Index [5], which requires consideration of at least 30 clinical variables. The RAI emphasizes effective distinction between frail and nonfrail patients in terms of increased risk of posttreatment complications and mortality. While the RAI was originally developed as a postoperative frailty assessment tool, it has been adapted and validated in nonoperative settings and across multiple hospitals.

To date, 3 versions of the index have been developed. The Clinical RAI (RAI-C) is a prospective screening tool in the form of a 14-item questionnaire addressing 11 frailty parameters [6]. The 11 frailty parameters are sex, age, cancer, renal failure, heart failure, cognitive decline, unintentional weight loss, poor appetite, shortness of breath at rest, activities of daily living (ADLs), and residence other than independent living. The Administrative RAI is a retrospective version that derives the same 11 frailty parameters from data available in surgical quality data sets [7]. The most recent version, called Risk Analysis Index-International Classification of Diseases (RAI-ICD), extends the application of RAI beyond the surgical population using the *International Classification of Diseases (ICD), Tenth Revision, Clinical Modification*, codes [24]. The RAI tools have numerous strengths, most notably their demonstrated validity and feasibility in assessing frailty in diverse and predominantly nonfrail populations. However, none of them fulfills the proposed criteria of automated assessment, thus requiring further adjustment and adaptation.

Our study extended the application of RAI by developing and validating empirically derived cut points for frailty risk categories. While previous applications of RAI primarily focused on binary classification (frail vs not frail), our approach sought to identify meaningful thresholds for 3 distinct risk categories (ie, not frail, prefrail, and frail) to enhance clinical utility and risk stratification.

Data Operationalization

To operationalize the 11 RAI frailty parameters identified by the clinical task force, the fourth step in the A-FIT development encompassed categorizing the 11 parameters. The first category comprised parameters assessed on admission that are unlikely to change during the hospitalization. The RAI parameters in this category were age, sex, and residence other than

XSL•FO
RenderX

independent living. The second category contained parameters assessed throughout hospitalization, for example, once a day or once a shift. These included shortness of breath and ADLs. Finally, there were parameters that are assessed on admission and may be updated as a patient's condition changes but not with regular frequency. These included unintentional weight loss and poor appetite, which are part of required nursing admission documentation but may not be assessed further unless directly relevant to the principal diagnosis or care plan. Similarly, current active medical diagnoses are documented on admission and only updated at discharge unless additional unexpected clinical problems emerge.

Informed by this practice-based data operationalization, the final step was to reconstruct a frailty assessment based on RAI parameters. Key data sources are listed in Textbox 1. While this study used retrospective data, an emphasis was placed on the use of data points available to clinicians at the point of care. Of the 11 frailty parameters, 7 had such accessible data points. Owing to limitations of available secondary data, the need to use administrative codes (ICD-9 and ICD-10) to represent cancer, heart failure, renal failure, and cognitive decline was necessary during this stage of development. Finally, hospital mortality was determined based on a discharge disposition of 'expired.'

Textbox 1. Automated Frailty Index Tool data operationalization.

Single time point

- Age on admission
- Sex
- Residence other than independent living (intake information)

Several time points

- Shortness of breath (at least one "dyspnea at rest" under Respiratory Pattern in Simple or Complex Assessment Flowsheet)
- Activities of daily living: eating, grooming, toileting, and walking (Activity Measure for Post-Acute Care assessment tool [25,26])

Continuous

- Unintentional weight loss
- Poor appetite (nutrition screening and admission documentation)
- Hospital problems or active diagnoses: cancer, heart failure, renal failure, cognitive decline (identified by *International Classification of Diseases, Ninth Revision*, or *International Classification of Diseases, Tenth Revision*, codes)

Statistical Analysis

Descriptive statistics

Descriptive statistics for each parameter, including frequency distribution (count and percent), central tendency (mean and median), and variability (SD), were generated following construction of the frailty assessment parameters as displayed in Textbox 1. A frailty score for each patient was calculated based on the scoring criteria proposed by Hall et al [6]. Differences between frailty categories were assessed using Kruskal-Wallis tests for continuous variables (age and ADL scores) and χ^2 tests for categorical variables. For significant Kruskal-Wallis results, post hoc pairwise comparisons were performed using Dunn test with Bonferroni correction. For the multicategory race/ethnicity variable, post hoc χ^2 tests were conducted with Bonferroni adjustment to identify specific between-group differences. In addition to statistical significance testing, standardized mean differences (SMDs) were calculated to assess the magnitude of differences between frailty categories, with SMD values >0.1, >0.3, and >0.5 considered to represent small, moderate, and large differences, respectively.

Frailty Classification

To determine optimal cut points for frailty classification, a data-driven approach was used. First, a binary cutpoint was identified using the Youden index (maximizing sensitivity+specificity–1) to distinguish frail from nonfrail patients. Subsequently, established data-driven cut points were established using receiver operating characteristic curve analysis and the Youden index to maximize discrimination. This approach yielded 3 clinically meaningful risk categories: not frail, prefrail, and frail. The robustness of these classifications was evaluated across age groups to ensure the index captured vulnerability beyond chronological age alone.

Model Development and Validation

To validate our frailty index and determine its incremental value beyond basic patient characteristics, several comparative models were constructed. The first model included only potential confounding factors (eg, hospital location, race/ethnicity, and smoking status). The second model added key clinical predictors (eg, age, sex, and functional status) to assess whether our composite frailty index provided additional predictive power beyond these established risk factors. The third and fourth models evaluated our frailty index in both continuous and categorical forms. Finally, a comprehensive model, including all individual frailty components, was constructed to determine whether our composite index adequately captured the predictive information contained in its constituent parts.

Multiple modeling approaches were compared to evaluate the predictive performance of the frailty index. Four models were constructed: (1) a confounders-only model, including hospital

XSL•FO
RenderX

location, race/ethnicity, and smoking status; (2) a key components model adding age, sex, and functional status (ADL) to the confounders; (3) a continuous frailty index model adding the index score to the confounders; and (4) a categorical frailty index model adding the 3-category frailty classification to the confounders. Additionally, these models were compared to a comprehensive model, including all individual frailty components.

Model performance was evaluated using discrimination (AUC) and calibration (Akaike Information Criterion [AIC]). Likelihood ratio tests were used to compare nested models. To assess the contribution of individual components to the frailty index's predictive ability, component importance analysis was conducted by sequentially removing each component and measuring the resulting change in AUC.

Additionally, age-stratified analyses were performed to evaluate the performance of the frailty classification system across different age groups (ie, 65 - 75, 76 - 85, and ≥86 years). For each age stratum, separate logistic regression models were fitted to estimate the odds of mortality associated with frailty categories, adjusting for potential confounders, including hospital location, race/ethnicity, and smoking status. All regression models were estimated using Firth's penalized likelihood approach to address potential issues related to rare events, as in-hospital mortality was an uncommon outcome in our dataset. Analyses were completed using R (version 4.3.1 [25]) with packages, including logistf, pROC, dplyr, tidyr, cutpointr, and dunn.test.

Results

Sample Characteristics

A total of 10,863 hospital patients aged ≥65 years were included. The average age was 75.4 (SD 7.7) years. Of the included participants, 4938 (45.5%) were male. With regard to race/ethnicity, 7724 participants (71.1%) were non-Hispanic White, 2406 participants (22.1%) were non-Hispanic Black, 314 participants (2.9%) were Hispanic, and the remaining 419 participants (3.9%) were other or unknown race/ethnicity. Hospital mortality occurred for 113 patients in the sample (1.04%). The frailty index scores ranged from 24 to 84. The skewness statistic was 0.39 (SE 0.02), indicating mild positive skew. The overall mean frailty score for the sample was 45.3 (SD 8.7).

Development of Frailty Categories

Using a data-driven approach, optimal cut points were identified to classify patients into 3 distinct risk categories. The first cutpoint at a frailty score of 54 was identified using the Youden index to maximize the sum of sensitivity (53.1%) and specificity (82.5%) for mortality prediction. A second cut point at a frailty score of 42 was established to distinguish between low-risk and intermediate-risk patients. On the basis of these thresholds, patients were classified as not frail (RAI <42), prefrail (RAI 42 - 53), and frail (RAI ≥54). The distribution of patients across these categories was as follows: not frail 40.8% (n=4434), prefrail 43.8% (n=4763), and frail 15.3% (n=1666). Hospital mortality rates showed a clear gradient across categories: 0.45% (95% CI 0.28%–0.70%) in not frail, 0.82% (95% CI 0.59%–1.12%) in prefrail, and 3.24% (95% CI 2.49%–4.23%) in frail patients.

Differences Between Frailty Categories

Significant differences were observed across frailty categories for both demographic and clinical characteristics and presented in Table 1. Post hoc pairwise comparisons using Dunn test with Bonferroni correction revealed significant age differences across all 3 frailty categories (all $P<.001$), with progressively higher median ages from not frail to prefrail to frail groups. For ADL scores, frail patients had significantly higher scores (indicating worse function) than both prefrail ($P<.001$) and not frail groups ($P=.046$), but no significant difference was found between not frail and prefrail groups ($P=.12$).

XSL•FO
RenderX

Table . Demographics based on frailty status.

	Overall	Not frail	Prefrail	Frail	SMD[a]	*P* value
n	10,863	4434	4763	1666		
Age, mean (SD), years	75.4 (7.7)	72.0 (68.0, 77.0)	76.0 (70.0, 83.0)	77.0 (71.0, 84.0)	0.454	<.001
ADL[b] score, median (interquartile range)	13.0 (12.0, 14.0)	13.0 (12.0, 14.0)	13.0 (11.0, 15.0)	13.0 (11.0, 14.0)	0.068	.002
Sex, n (%)					0.549	<.001
Female	5925 (54.5)	3084 (69.6)	2328 (48.9)	513 (30.8)		
Male	4938 (45.5)	1350 (30.4)	2435 (51.1)	1153 (69.2)		
Race/ethnicity, n (%)					0.065	.01
Hispanic	314 (2.9)	136 (3.1)	140 (2.9)	38 (2.3)		
Non-Hispanic Black	2406 (22.1)	917 (20.7)	1123 (23.6)	366 (22.0)		
Non-Hispanic White	7724 (71.1)	3193 (72.0)	3331 (69.9)	1200 (72.0)		
Other	419 (3.9)	188 (4.2)	169 (3.5)	62 (3.7)		
Hospital location, n (%)					0.08	<.001
Gainesville	8525 (78.5)	3497 (78.9)	3665 (76.9)	1363 (81.8)		
Jacksonville	2338 (21.5)	937 (21.1)	1098 (23.1)	303 (18.2)		
Dependent living, n (%)					0.427	<.001
No	10,101 (93.0)	4376 (98.7)	4383 (92.0)	1342 (80.6)		
Yes	762 (7.0)	58 (1.3)	380 (8.0)	324 (19.4)		
Shortness of breath, n (%)					0.505	<.001
No	10,184 (93.7)	4405 (99.3)	4488 (94.2)	1291 (77.5)		
Yes	679 (6.3)	29 (0.7)	275 (5.8)	375 (22.5)		
Weight loss, n (%)					0.523	<.001
No	9847 (90.6)	4328 (97.6)	4324 (90.8)	1195 (71.7)		
Yes	1016 (9.4)	106 (2.4)	439 (9.2)	471 (28.3)		
Poor appetite, n (%)					0.656	<.001
No	8694 (80.0)	4090 (92.2)	3733 (78.4)	871 (52.3)		
Yes	2169 (20.0)	344 (7.8)	1030 (21.6)	795 (47.7)		
Cancer diagnosis, n (%)					1.368	<.001
No	7849 (72.3)	4404 (99.3)	2998 (62.9)	447 (26.8)		
Yes	3014 (27.7)	30 (0.7)	1765 (37.1)	1219 (73.2)		
Heart failure, n (%)					0.49	<.001
No	8093 (74.5)	3855 (86.9)	3315 (69.6)	923 (55.4)		
Yes	2770 (25.5)	579 (13.1)	1448 (30.4)	743 (44.6)		
Renal failure, n (%)					0.788	<.001
No	7935 (73.0)	4035 (91.0)	3205 (67.3)	695 (41.7)		
Yes	2928 (27.0)	399 (9.0)	1558 (32.7)	971 (58.3)		

XSL•FO
RenderX

	Overall	Not frail	Prefrail	Frail	SMD[a]	*P* value
Cognitive decline, n (%)					0.272	<.001
No	9406 (86.6)	4139 (93.3)	3939 (82.7)	1328 (79.7)		
Yes	1457 (13.4)	295 (6.7)	824 (17.3)	338 (20.3)		
Mortality, n (%)					0.142	<.001
No	10,750 (99.0)	4414 (99.5)	4724 (99.2)	1612 (96.8)		
Yes	113 (1.04)	20 (0.5)	39 (0.8)	54 (3.2)		

[a]SMD: standardized mean difference.

[b]ADL: activities of daily living.

Analyses of race/ethnicity distribution using χ^2 tests with Bonferroni correction identified a significant difference between non-Hispanic Black and non-Hispanic White patients (*P*=.039). Non-Hispanic Black patients were less likely to be classified as not frail (38.1% vs 41.3%) and more likely to be classified as prefrail (46.7% vs 43.1%) compared to non-Hispanic White patients. The proportion classified as frail was similar across all racial/ethnic groups, ranging from 12.1% among Hispanic patients to 15.5% among non-Hispanic White patients.

Model Performance Comparison

Multiple models were compared to evaluate the predictive performance of our frailty index. A baseline model, including only potential confounders (eg, hospital location, race/ethnicity, and smoking status), showed poor discrimination for hospital mortality (AUC 0.592). A model incorporating key clinical predictors (eg, age, sex, and functional status) performed better (AUC 0.760). The continuous frailty index model demonstrated good discrimination (AUC 0.738), whereas the categorical frailty classification model achieved similar performance (AUC 0.717). A comprehensive model using all individual frailty components demonstrated the highest discrimination (AUC 0.887).

Likelihood ratio tests confirmed that both the continuous frailty index (χ^2_1=100.7; *P*<.001) and the 3-category classification (χ^2_2=75.4; *P*<.001) significantly improved prediction compared to the confounders-only model (Table 2). The AIC values showed a similar pattern, with the all-components model having the lowest AIC (935.47), followed by the continuous frailty index model (1142.02), the categorical model (1169.38), and the confounder-only model (1240.73).

XSL•FO
RenderX

Table . Frailty index model comparisons.

Model	AUC[a]	AIC[b]	AUC_CI	χ^2	*P* value
Model A (confounders only)[c]	0.592	1240.7	0.542 - 0.642	—[d]	—
Model B (key components)[e]	0.760	1252.3	0.713 - 0.807	—	—
Model C (continuous frailty index plus confounders)	0.738	1127.8	0.689 - 0.788	—	—
Model D[f] (categorical frailty index plus confounders)	0.717	1239.5	0.665 - 0.768	—	—
Model E (all components)	0.887	1142	0.853 - 0.921	—	—
Confounders only	0.592	1245.2	0.542 - 0.642	—	—
Key components	0.760	1169.4	0.713 - 0.807	—	—
Continuous frailty index	0.738	1250.4	0.689 - 0.788	100.7	<.001
Categorical frailty index	0.717	935.5	0.665 - 0.768	75.4	<.001
All components	0.887	1234.8	0.853 - 0.921	—	—

[a]AUC: area under receiver operating characteristic curve.

[b]AIC: calibration.

[c]Confounders (hospital location, race/ethnicity, and smoking status).

[d]Not applicable.

[e]Key components (age, sex, and activities of daily living).

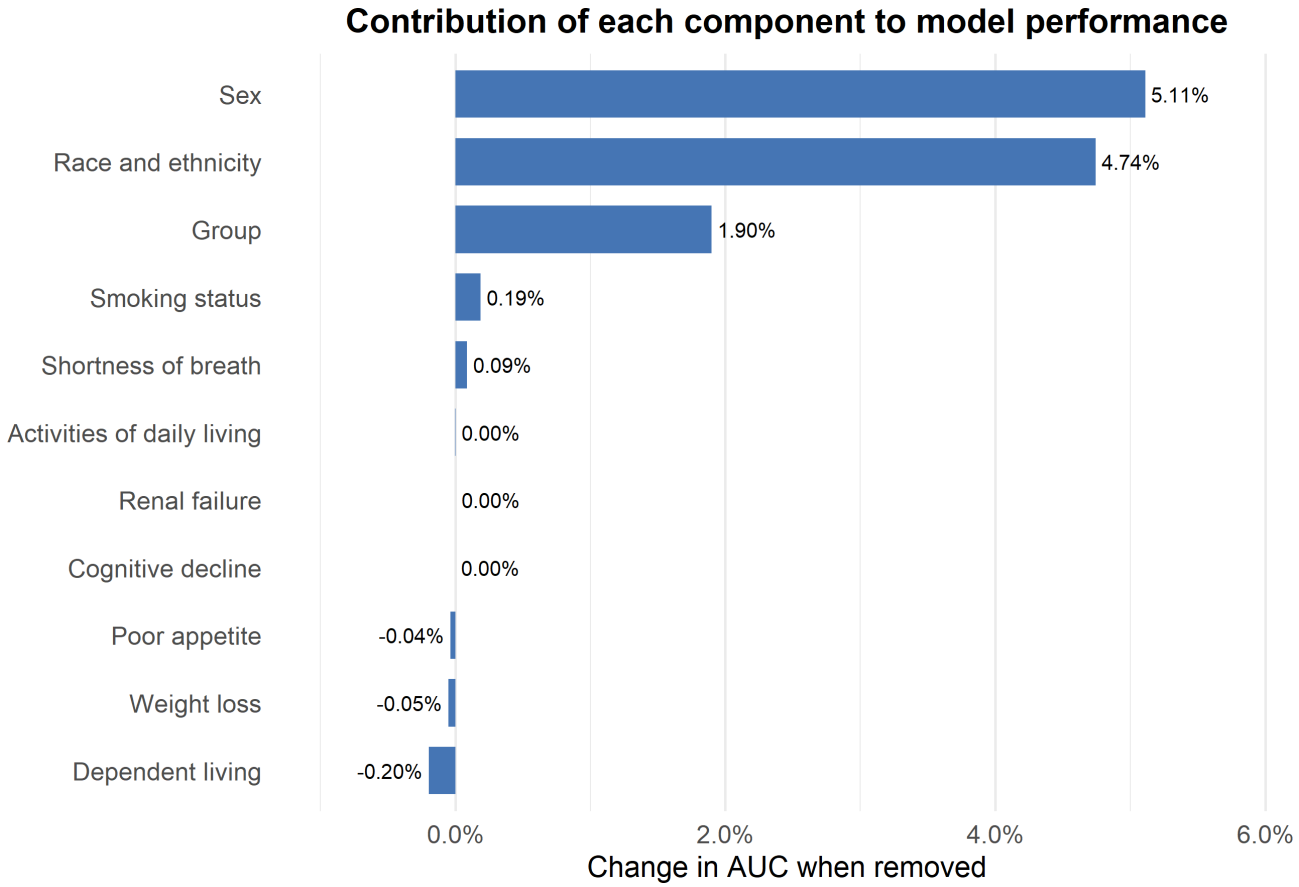
[f]Categorical age groups (ie, 65-75, 76-85, and ≥86 years).

Component Importance Analysis

To identify the key drivers of predictive performance, a component importance analysis was conducted by sequentially removing each frailty parameter from the comprehensive model shown in Figure 1. This analysis revealed that shortness of breath was the strongest predictor (AUC reduction: 0.051 when removed), followed by ADL limitations (AUC reduction: 0.047) and renal failure (AUC reduction: 0.019). Interestingly, age showed minimal independent contribution within this 65+ population (AUC reduction: –0.0005), suggesting that our frailty index captures vulnerability beyond chronological age.

XSL•FO
RenderX

Figure 1. Contribution of each component to model performance. ADL: activities of daily living; AUC: area under receiver operating characteristic curve.



Age-Stratified Analysis

As presented in Table 3, the proportion of patients classified as frail increased with age: 12% in the 65 to 75 years age group, 18% in the 76 to 85 years age group, and 26% in the ≥86 age group. In all age groups, the frail category identified patients at substantially higher risk, with mortality rates 6 to 8 times higher than patients classified as not frail. Notably, a frail patient in the 65 to 75 years age group (mortality 3.97%) had higher mortality risk than a not frail patient in the ≥86 years age group (mortality 0.48%).

Table . Age-stratified odds of mortality controlling for hospital location, race/ethnicity, and smoking status.

	Total (N)	Mortality, n (%)	Adjusted odds ratio	95% CI	*P* value
Age group (years)					
65 - 75					
Not frail	3066	15 (0.49)	*Reference*	—[a]	—
Prefrail	2287	17 (0.74)	1.56	0.78 - 3.13	.208
Frail	706	28 (3.97)	8.50	4.59 - 16.32	<.001
76 - 85					
Not frail	1159	4 (0.35)	*Reference*	—	—
Prefrail	1714	16 (0.93)	2.47	0.94 - 8.01	.069
Frail	611	16 (2.62)	6.42	2.41 - 21.02	<.001
≥86					
Not frail	209	1 (0.48)	*Reference*	—	—
Prefrail	762	6 (0.79)	1.15	0.23 - 11.12	.88
Frail	349	10 (2.87)	3.99	0.91 - 37.48	.07

[a]Not applicable.

XSL•FO
RenderX

After adjusting for hospital location, race/ethnicity, and smoking status, age-stratified logistic regression models showed varying strengths of association between frailty categories and mortality. In the 65 to 75 years age group, frail patients had significantly higher odds of mortality compared to not frail patients (odds ratio [OR] 8.50, 95% CI 4.59 - 16.32; *P*<.001), whereas prefrail status showed a positive but nonsignificant association (OR 1.56, 95% CI 0.78 - 3.13; *P*=.208).

In the 76 to 85 years age group, the association remained strong for frail patients (OR 6.42, 95% CI 2.41 - 21.02; *P*<.001), but prefrail status did not reach statistical significance (OR 2.47, 95% CI 0.94 - 8.01; *P*=.069). In the oldest age group (≥86 years), the point estimate for frail status remained elevated (OR 3.99, 95% CI 0.91 - 37.48) but did not reach statistical significance (*P*=.07).

Discussion

Principal Findings

The purpose of this research was to automate risk assessment and early frailty detection by using readily available EHR data, with the long-term goal of improving frailty detection while minimizing the documentation burden on health care providers. Our study not only developed a viable frailty index but also established empirically derived cut points for a 3-category frailty classification system with strong predictive validity for hospital mortality. This work demonstrates that routinely collected EHR data can be leveraged to create a clinically meaningful risk stratification tool for older adult medical/surgical patients.

Frailty Index Performance and Comparison With Previous Work

In this study, the average frailty index score was 45.3, considerably higher than the mean frailty scores reported in the development and validation of the original RAI instruments (11.8 for clinical RAI and 9.6 for RAI-administrative). This difference is largely explained by our focus on adults aged ≥65 years, whereas Hall et al [6] included all adults aged ≥18 years. Our 3-category classification system identified 14.9% of patients as frail (RAI ≥54), 43.5% as prefrail (RAI 42 - 53), and 40.6% as not frail (RAI <42). This distribution aligns with previous literature, suggesting that approximately 10% to 15% of community-dwelling older adults are frail, with a larger proportion in the prefrail category [27,28].

Our model comparison revealed that while individual frailty components provide optimal predictive power (AUC=0.887), the composite frailty index (AUC=0.738) and 3-category classification (AUC=0.717) offer substantial improvement over basic demographic factors alone (AUC=0.592). Interestingly, a simpler model using only age, sex, and functional status achieved slightly better discrimination (AUC=0.760) than our composite index, echoing previous findings about the powerful predictive value of functional status in geriatric populations [29-31]. However, our 3-category frailty classification offers conceptual advantages, clinical interpretability, and standardization that may outweigh the modest reduction in statistical performance.

Component Importance and Novel Insights

Our component importance analysis yielded several notable findings. First, respiratory dysfunction (shortness of breath) emerged as the strongest mortality predictor, with an AUC reduction of 0.051 when removed from the model. This aligns with growing recognition of respiratory parameters as important indicators of acute deterioration in hospitalized patients [32,33]. That shortness of breath was the most significant factor for the development of frailty, which is consistent with other research that revealed among older persons, respiratory impairment and frailty are strongly associated with one another [29,34]. Additionally, there is a substantial increase in vulnerability to adverse health outcomes, such as heart failure or chronic obstructive pulmonary disease, and risk of death when both are present. These conditions contribute to reduced physical activity, muscle weakness, and overall functional decline, thereby accelerating the progression toward frailty. Unfortunately, frailty is often underrepresented in respiratory management guidelines, which may contribute to limited recognition of its comorbidity in respiratory care [35].

Functional status (ADL) was the second most important component (AUC reduction: 0.047), consistent with extensive literature establishing functional measures as robust predictors of outcomes in older adults [36]. Notably, age contributed minimally to prediction beyond these clinical factors, confirming that our index captures frailty as a distinct clinical entity rather than merely reflecting chronological aging. This finding is particularly significant, given that our study focused exclusively on adults aged ≥65 years, suggesting that frailty assessment provides valuable risk stratification even within the older adult population.

Age-Stratified Analysis and Demographic Patterns

Our age-stratified analysis revealed that while the prevalence of frailty increases with age (12% in ages 65 - 75 years, 18% in ages 76 - 85 years, and 26% in ages ≥86 years), the predictive value of frailty categories remains robust across all age strata. The striking finding that frail patients aged 65 to 75 years had higher mortality risk than nonfrail patients aged ≥86 years underscores the clinical importance of frailty assessment beyond simple age-based approaches to risk stratification.

The gradient in ORs across age groups (OR decreasing from 8.50 to 6.42 to 3.99) suggests that while frailty remains an important predictor of mortality across all age groups, its relative impact may diminish somewhat in the oldest old. This pattern aligns with the concept of survival bias, where those who reach advanced age despite being classified as frail may represent a resilient subpopulation [37]. Nevertheless, the consistency of the association pattern across age strata supports the clinical utility of the frailty classification system throughout the geriatric age spectrum.

The declining statistical significance observed in the oldest age group likely reflects limited statistical power rather than absence of effect, as evidenced by the wide CIs. This highlights a common challenge in geriatric research where sample sizes often decrease with advancing age categories.

XSL•FO
RenderX

Our analysis of race/ethnicity distribution revealed significant differences between non-Hispanic Black and non-Hispanic White patients, with Black patients less likely to be classified as not frail (38.1% vs 41.3%) and more likely to be classified as prefrail (46.7% vs 43.1%). This finding adds to the growing body of evidence on health disparities in frailty prevalence [38,39] and highlights the importance of considering social determinants of health in frailty assessment and intervention planning.

The ORs for mortality demonstrate the strong discriminative ability of our classification system; after adjusting for confounders, prefrail patients had 1.8 times higher odds of mortality than not frail patients, whereas frail patients had 7.1 times higher odds. These magnitudes of association are clinically meaningful and comparable to established frailty measures in the literature [4-7].

Technical Implications

This study went beyond validation of an existing tool. Rather than using identical predictors and operational definitions from the original RAI development, the framework was adapted to work with routinely available EHR data, enhancing translatability across sites and settings. Our methodological approach to handling the rare outcome (hospital death) was rigorous, using Firth's penalized likelihood method to address potential bias in logistic regression with rare events.

Our 3-category classification system represents an important methodological advancement over binary approaches. By identifying a prefrail intermediate category, our system provides greater nuance for risk stratification and potentially enables earlier intervention before patients progress to more severe frailty. The use of SMDs in addition to statistical testing helped identify clinically meaningful differences between frailty categories, particularly for functional and respiratory parameters.

Clinical Implications

Frailty is recognized as an important predisposing factor for and predictor of adverse outcomes. Therefore, there is a necessity for frailty screening to identify potentially modifiable vulnerabilities among frail older adults. A practical application includes such procedures as the utilization of prehabilitation and optimized nutrition that have demonstrated beneficial effects on functional outcomes postoperatively [40-45]. This research extends these efforts beyond the surgical arena to also optimize hospitalized medical patients' condition and reduce the incidences of many possible complications that increase morbidity, prolong lengths of stay, and may lead to mortality.

The clinical implications of our findings are substantial. First, our 3-category frailty classification system provides clinicians with a practical framework for risk stratification that can be automatically generated from EHR data without additional documentation burden. Second, the identification of shortness of breath and functional status as the strongest predictors suggests potential targets for intervention and monitoring. Third, our finding that frailty predicts mortality independent of age challenges age-based stereotypes and encourages individualized risk assessment.

Our tool extends frailty assessment beyond the surgical arena to all hospitalized medical/surgical patients, addressing a significant gap in current practice. While much previous research on functional decline due to frailty has focused on surgical outcomes [46-48], our work demonstrates the relevance of frailty assessment across the broader inpatient population. Furthermore, addressing shortness of breath through appropriate medical intervention and pulmonary rehabilitation can potentially mitigate the risk of frailty development in older adults for improved clinical outcomes [49,50].

Frailty is also appreciated as a dynamic condition for older adults, who may experience changes in their frailty status over time. This may occur as either a gradual decline or sudden change in an acute care setting or possibly reversible with rehabilitation. Previous research has shown an association between older age and faster progression of frailty [51]. Additionally, patients who were postsurgery, especially emergency general surgeries, can induce or worsen frailty. Ng et al [52] showed that 33.3% of older patients undergoing emergency surgery experienced an increase in their frailty status 90 days postoperatively. Furthermore, 15.9% of patients were considered with a new frail condition impacting their discharge destination. Our automated approach creates opportunities for continuous monitoring of frailty status across hospital admissions, potentially allowing clinical teams to detect and respond to changes in patient vulnerability during hospitalization.

Strengths and Limitations

This study has several strengths, including the large sample size, rigorous methodological approach, and comprehensive validation of the frailty classification system. The use of data-derived cut points for our 3-category system provides a data-driven foundation for clinical implementation. Our component importance analysis identified key drivers of mortality prediction, offering valuable insights for focused assessment and intervention.

However, limitations must be acknowledged. First, our reliance on secondary data meant that some frailty parameters were derived from administrative codes rather than direct clinical assessment. Second, hospital mortality was a rare event (1.04%), necessitating specialized statistical approaches and potentially limiting statistical power for some analyses. Our use of Firth's penalized likelihood method helped address issues related to rare events, but validation in cohorts with more outcome events would strengthen confidence in our findings.

Additionally, our analytical sample excluded patients whose residential status could not be clearly established, including homeless individuals, potentially introducing selection bias. Future studies should investigate the inclusion of homelessness as an independent risk factor to improve generalizability. The observed differences in frailty distribution between racial/ethnic groups also warrant further investigation into potential disparities in frailty assessment and care.

While our study demonstrated the predictive validity of our frailty classification for in-hospital mortality, longer-term outcomes or other clinically relevant end points, such as

XSL•FO
RenderX

functional decline, discharge disposition, or readmission, were not assessed. Future research should explore these additional outcomes to further validate the clinical utility of our approach and include sequential measures to investigate how frailty changes during acute care so that clinicians can monitor and respond to changes in patient vulnerability during hospitalization.

Conclusions

This paper describes the development and validation of a 3-category frailty classification system based on readily available EHR data that demonstrate strong predictive validity for hospital mortality among older adults. Our approach provides meaningful risk stratification with mortality rates of 0.45%, 0.82%, and 3.24% across not frail, prefrail, and frail categories, respectively. Component importance analysis identified respiratory dysfunction and functional limitations as the strongest predictors, offering potential targets for monitoring and intervention.

This automated approach to frailty assessment could provide critical information to clinical teams with minimal additional documentation burden, potentially improving risk stratification and clinical decision-making for hospitalized older adults. Further research should explore how this frailty classification system performs for additional outcomes and whether implementation in clinical workflows improves patient care and outcomes.

Acknowledgments

This research was funded by the National Institute for Aging through the grant entitled Advancing Interdisciplinary Science of Aging through Identification of Iatrogenic Complications: The University of Florida Electronic Health Record Clinical Data Infrastructure for Enhanced Patient Safety among the Elderly (R33 AG062884-03).

Data Availability

The clinical EHR data analyzed in this study cannot be shared publicly due to HIPAA and institutional privacy restrictions. Deidentified analytic materials may be available from the corresponding author on reasonable request and with required institutional approvals. Study code is available upon reasonable request.

Conflicts of Interest

None declared.

References

1. Baxter CT, Offord NJ. Frailty on the acute medical take. Clin Med (Lond) 2022 Jul;22(4):298-301. [doi: 10.7861/clinmed.2022-0312] [Medline: 35882496]
2. 13th International Conference on Frailty & Sarcopenia Research (ICFSR). J Frailty Aging 2023 Mar;12(Suppl 1):2-32. [doi: 10.14283/jfa.2023.16]
3. Cappe M, Laterre PF, Dechamps M. Preoperative frailty screening, assessment and management. Curr Opin Anaesthesiol 2023 Feb 1;36(1):83-88. [doi: 10.1097/ACO.0000000000001221] [Medline: 36476726]
4. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
5. Peña FG, Theou O, Wallace L, et al. Comparison of alternate scoring of variables on the performance of the frailty index. BMC Geriatr 2014 Feb 24;14:25. [doi: 10.1186/1471-2318-14-25] [Medline: 24559204]
6. Hall DE, Arya S, Schmid KK, et al. Development and initial validation of the risk analysis index for measuring frailty in surgical populations. JAMA Surg 2017 Feb 1;152(2):175-182. [doi: 10.1001/jamasurg.2016.4202] [Medline: 27893030]
7. Arya S, Varley P, Youk A, et al. Recalibration and external validation of the Risk Analysis Index: a surgical frailty assessment tool. Ann Surg 2020 Dec;272(6):996-1005. [doi: 10.1097/SLA.0000000000003276] [Medline: 30907757]
8. McIsaac DI, Moloo H, Bryson GL, van Walraven C. The association of frailty with outcomes and resource use after emergency general surgery: a population-based cohort study. Anesth Analg 2017 May;124(5):1653-1661. [doi: 10.1213/ANE.0000000000001960] [Medline: 28431425]
9. Joseph B, Pandit V, Zangbar B, et al. Validating trauma-specific frailty index for geriatric trauma patients: a prospective analysis. J Am Coll Surg 2014 Jul;219(1):10-17. [doi: 10.1016/j.jamcollsurg.2014.03.020] [Medline: 24952434]
10. Shin J, Choi J, Kweon HJ, Han Y, Lee M. Hospital frailty risk score using electronic medical records and geriatric syndromes in an acute-care hospital. Geriatr Nurs 2025;62(Pt A):175-180. [doi: 10.1016/j.gerinurse.2025.01.016] [Medline: 39908784]
11. Kim DH, Schneeweiss S, Glynn RJ, Lipsitz LA, Rockwood K, Avorn J. Measuring frailty in Medicare data: development and validation of a claims-based frailty index. J Gerontol A Biol Sci Med Sci 2018 Jun 14;73(7):980-987. [doi: 10.1093/gerona/glx229] [Medline: 29244057]
12. Kim DH, Glynn RJ, Avorn J, et al. Validation of a claims-based frailty index against physical performance and adverse health outcomes in the health and retirement study. J Gerontol A Biol Sci Med Sci 2019 Jul 12;74(8):1271-1276. [doi: 10.1093/gerona/gly197] [Medline: 30165612]

XSL•FO
RenderX

13. United Nations, Department of Economic and Social Affairs, Population Division. World population prospects 2024: summary of results. : United Nations; 2024 URL: https://population.un.org/wpp/assets/Files/WPP2024_Summary-of-Results.pdf [accessed 2025-11-27]
14. Vespa JE, Armstrong DM, Medina L. Demographic turning points for the United States: population projections for 2020 to 2060. : US Department of Commerce, Economics and Statistics Administration, US Census Bureau; 2018.
15. Henry J, Pylypchuk Y, Searcy T, Patel V. Adoption of electronic health record systems among US non-federal acute care hospitals: 2008–2015. In: ASTP Health IT Data Brief: Washington (DC): Office of the Assistant Secretary for Technology Policy; 2016, Vol. 35:2008-2015 URL: https://www.ncbi.nlm.nih.gov/books/NBK612948/
16. Searle SD, Mitnitski A, Gahbauer EA, Gill TM, Rockwood K. A standard procedure for creating a frailty index. BMC Geriatr 2008 Sep 30;8:24. [doi: 10.1186/1471-2318-8-24] [Medline: 18826625]
17. Theou O, Haviva C, Wallace L, Searle SD, Rockwood K. How to construct a frailty index from an existing dataset in 10 steps. J Gerontol A Biol Sci Med Sci 2019;74(8):1271-1276. [doi: 10.1093/gerona/gly197]
18. Snigurska UA, Ser SE, Solberg LM, et al. Application of a practice-based approach in variable selection for a prediction model development study of hospital-induced delirium. BMC Med Inform Decis Mak 2023 Sep 13;23(1):181. [doi: 10.1186/s12911-023-02278-1] [Medline: 37704994]
19. Ser SE, Shear K, Snigurska UA, et al. Clinical prediction models for hospital-induced delirium using structured and unstructured electronic health record data: protocol for a development and validation study. JMIR Res Protoc 2023 Nov 9;12:e48521. [doi: 10.2196/48521] [Medline: 37943599]
20. Lucero RJ, Bakken S. Practice-based knowledge discovery for comparative effectiveness research: an organizing framework. Can J Nurs Res 2013 Mar;45(1):98-112. [doi: 10.1177/084456211304500109] [Medline: 25278645]
21. Mitchell PH, Ferketich S, Jennings BM. Quality health outcomes model. American Academy of Nursing Expert Panel on Quality Health Care. Image J Nurs Sch 1998;30(1):43-46. [doi: 10.1111/j.1547-5069.1998.tb01234.x] [Medline: 9549940]
22. Fayyad UM, Piatetsky-Shapiro G, Smyth P. From data mining to knowledge discovery: an overview. In: Fayyad US, Piatetsky-Shapiro G, Smyth P, Uthurusamy R, editors. Advances in Knowledge Discovery and Data Mining (2017).: AAAI Press/MIT Press; 1996 URL: https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/1230 [accessed 2025-11-27] [doi: 10.1609/aimag.v17i3.1230]
23. Niznik JD, Zhang S, Mor MK, et al. Adaptation and iInitial validation of Minimum Data Set (MDS) Mortality Risk Index to MDS Version 3.0. J Am Geriatr Soc 2018 Dec;66(12):2353-2359. [doi: 10.1111/jgs.15579] [Medline: 30335184]
24. Dicpinigaitis AJ, Khamzina Y, Hall DE, et al. Adaptation of the Risk Analysis Index for frailty assessment using diagnostic codes. JAMA Netw Open 2024 May 1;7(5):e2413166. [doi: 10.1001/jamanetworkopen.2024.13166] [Medline: 38787554]
25. R software. R: A language and environment for statistical computing. 2021. URL: https://www.R-project.org [accessed 2025-11-25]
26. Jette DU, Stilphen M, Ranganathan VK, Passek SD, Frost FS, Jette AM. Validity of the AM-PAC “6-Clicks” inpatient daily activity and basic mobility short forms. Phys Ther 2014 Mar;94(3):379-391. [doi: 10.2522/ptj.20130199] [Medline: 24231229]
27. Junek ML, Jones A, Heckman G, Demers C, Griffith LE, Costa AP. The predictive utility of functional status at discharge: a population-level cohort analysis. BMC Geriatr 2022 Jan 3;22(1):8. [doi: 10.1186/s12877-021-02652-6] [Medline: 34979946]
28. Inouye SK, Peduzzi PN, Robison JT, Hughes JS, Horwitz RI, Concato J. Importance of functional measures in predicting mortality among older hospitalized patients. JAMA 1998 Apr 15;279(15):1187-1193. [doi: 10.1001/jama.279.15.1187] [Medline: 9555758]
29. Collard RM, Boter H, Schoevers RA, Oude Voshaar RC. Prevalence of frailty in community-dwelling older persons: a systematic review. J Am Geriatr Soc 2012 Aug;60(8):1487-1492. [doi: 10.1111/j.1532-5415.2012.04054.x] [Medline: 22881367]
30. Hoogendijk EO, Afilalo J, Ensrud KE, Kowal P, Onder G, Fried LP. Frailty: implications for clinical practice and public health. Lancet 2019 Oct 12;394(10206):1365-1375. [doi: 10.1016/S0140-6736(19)31786-6] [Medline: 31609228]
31. Ofori-Asenso R, Chin KL, Mazidi M, et al. Global incidence of frailty and prefrailty among community-dwelling older adults: a systematic review and meta-analysis. JAMA Netw Open 2019 Aug 2;2(8):e198398. [doi: 10.1001/jamanetworkopen.2019.8398] [Medline: 31373653]
32. Kayser SA, Williamson R, Siefert G, Roberts D, Murray A. Respiratory rate monitoring and early detection of deterioration practices. Br J Nurs 2023 Jul 6;32(13):620-627. [doi: 10.12968/bjon.2023.32.13.620] [Medline: 37410682]
33. Mochizuki K, Shintani R, Mori K, et al. Importance of respiratory rate for the prediction of clinical deterioration after emergency department discharge: a single-center, case-control study. Acute Med Surg 2017 Apr;4(2):172-178. [doi: 10.1002/ams2.252] [Medline: 29123857]
34. Vaz Fragoso CA, Enright PL, McAvay G, Van Ness PH, Gill TM. Frailty and respiratory impairment in older persons. Am J Med 2012 Jan;125(1):79-86. [doi: 10.1016/j.amjmed.2011.06.024] [Medline: 22195532]
35. Verduri A, Clini E, Carter B, Hewitt J. Impact of frailty on symptom burden in chronic obstructive pulmonary disease. J Clin Med 2024 Feb 8;13(4):984. [doi: 10.3390/jcm13040984] [Medline: 38398298]

XSL•FO
RenderX

36. Osadnik CR, Brighton LJ, Burtin C, et al. European Respiratory Society statement on frailty in adults with chronic lung disease. Eur Respir J 2023 Aug;62(2):2300442. [doi: 10.1183/13993003.00442-2023] [Medline: 37414420]
37. Allen SK, Brown V, White D, et al. Multimodal prehabilitation during neoadjuvant therapy prior to esophagogastric cancer resection: effect on cardiopulmonary exercise test performance, muscle mass and quality of life-a pilot randomized clinical trial. Ann Surg Oncol 2022 Mar;29(3):1839-1850. [doi: 10.1245/s10434-021-11002-0] [Medline: 34725764]
38. Banack HR, Kaufman JS, Wactawski-Wende J, Troen BR, Stovitz SD. Investigating and remediating selection bias in geriatrics research: the selection bias toolkit. J Am Geriatr Soc 2019 Sep;67(9):1970-1976. [doi: 10.1111/jgs.16022] [Medline: 31211407]
39. Khan M, Nicholl BI, Hanlon P. Ethnicity and frailty: A systematic review of association with prevalence, incidence, trajectories and risks. Ageing Res Rev 2025 Jul;109:102759. [doi: 10.1016/j.arr.2025.102759] [Medline: 40306389]
40. Usher T, Buta B, Thorpe RJ, et al. Dissecting the racial/ethnic disparity in frailty in a nationally representative cohort study with respect to health, income, and measurement. J Gerontol A Biol Sci Med Sci 2021 Jan 1;76(1):69-76. [doi: 10.1093/gerona/glaa061] [Medline: 32147727]
41. Heil TC, Verdaasdonk EGG, Maas H, et al. Improved postoperative outcomes after prehabilitation for colorectal cancer surgery in older patients: an emulated target trial. Ann Surg Oncol 2023 Jan;30(1):244-254. [doi: 10.1245/s10434-022-12623-9] [Medline: 36197561]
42. Carli F, Bousquet-Dion G, Awasthi R, et al. Effect of multimodal prehabilitation vs postoperative rehabilitation on 30-day postoperative complications for frail patients undergoing resection of colorectal cancer: a randomized clinical trial. JAMA Surg 2020 Mar 1;155(3):233-242. [doi: 10.1001/jamasurg.2019.5474] [Medline: 31968063]
43. Molenaar CJL, Minnella EM, Coca-Martinez M, et al. Effect of multimodal prehabilitation on reducing postoperative complications and enhancing functional capacity following colorectal cancer surgery: the PREHAB randomized clinical trial. JAMA Surg 2023 Jun 1;158(6):572-581. [doi: 10.1001/jamasurg.2023.0198] [Medline: 36988937]
44. Berardi G, Vitiello A, Abu-Abeid A, et al. Micronutrients deficiencies in candidates of bariatric surgery: results from a single institution over a 1-year period. Obes Surg 2023 Jan;33(1):212-218. [doi: 10.1007/s11695-022-06355-8] [Medline: 36331725]
45. Aiolfi A, Griffiths EA, Sozzi A, et al. Effect of anastomotic leak on long-term survival after esophagectomy: multivariate meta-analysis and restricted mean survival times examination. Ann Surg Oncol 2023 Sep;30(9):5564-5572. [doi: 10.1245/s10434-023-13670-6] [Medline: 37210447]
46. Abdelfatah E, Ramos-Santillan V, Cherkassky L, Cianchetti K, Mann G. High risk, high reward: frailty in colorectal cancer surgery is associated with worse postoperative outcomes but equivalent long-term oncologic outcomes. Ann Surg Oncol 2023 Apr;30(4):2035-2045. [doi: 10.1245/s10434-022-12970-7] [Medline: 36648616]
47. Lee SH, Lim SW. Risk factors for postoperative delirium after colorectal surgery: a systematic review and meta-analysis. Int J Colorectal Dis 2020 Mar;35(3):433-444. [doi: 10.1007/s00384-019-03498-6] [Medline: 31897646]
48. Orouji Jokar T, Ibraheem K, Rhee P, et al. Emergency general surgery specific frailty index: a validation study. J Trauma Acute Care Surg 2016 Aug;81(2):254-260. [doi: 10.1097/TA.0000000000001120] [Medline: 27257694]
49. Maddocks M, Brighton LJ, Alison JA, et al. Rehabilitation for people with respiratory disease and frailty: an official American Thoracic Society workshop report. Ann Am Thorac Soc 2023 Jun;20(6):767-780. [doi: 10.1513/AnnalsATS.202302-129ST] [Medline: 37261787]
50. Tsai CL, Chang WP, Lin YK, Ho SC, Lin YH. Physical frailty related to cognitive impairment and COPD exacerbation: a cross-sectional study. Respir Med 2023 Mar;208:107129. [doi: 10.1016/j.rmed.2023.107129] [Medline: 36709919]
51. Mian H, Wildes TM, Vij R, Pianko MJ, Major A, Fiala MA. Dynamic frailty risk assessment among older adults with multiple myeloma: a population-based cohort study. Blood Cancer J 2023 May 10;13(1):76. [doi: 10.1038/s41408-023-00843-5] [Medline: 37164972]
52. Ng HJ, Rattray NJW, Quasim T, Moug SJ. Changes in frailty status and discharge destination post emergency laparotomy. World J Emerg Surg 2025 Apr 25;20(1):37. [doi: 10.1186/s13017-025-00612-8] [Medline: 40281633]

Abbreviations

A-FIT: Automated Frailty Index Tool
ADL: activities of daily living
AUC: area under the receiver operating characteristic curve
EHR: electronic health record
ICD: International Classification of Diseases
OR: odds ratio
RAI: Risk Analysis Index
RAI-C: Clinical Risk Analysis Index
SMD: standardized mean difference

XSL•FO
RenderX

Edited by A Quialheiro; submitted 23.10.24; peer-reviewed by A Lakdawala, C Bai, F Okeke; revised version received 08.08.25; accepted 28.09.25; published 02.12.25.

Please cite as:
Celso BG, Ser SE, Snigurska UA, Solberg LM, Duckworth LJ, Dunn EM, Prosperi M, Lucero RJ, Bjarnadottir RI
Reconstruction of a Risk Analysis Tool That Uses Nursing Data for Frailty Assessment Among Older Adults: Derivation and Internal Evaluation Study
JMIR Aging 2025;8:e67888
URL: https://aging.jmir.org/2025/1/e67888
doi:10.2196/67888

© Brian G Celso, Sarah E Ser, Urszula A Snigurska, Laurence M Solberg, Laurie J Duckworth, Elizabeth M Dunn, Mattia Prosperi, Robert J Lucero, Ragnhildur I Bjarnadottir. Originally published in JMIR Aging (https://aging.jmir.org), 2.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Insights Into Associations Between Oral and General Health Outcomes of Nursing Home Residents Based on InterRAI Data: Cross-Sectional Study

Emilie Schoebrechts[1], MSc; Johanna de Almeida Mello[1,2], MSc, PhD; Patricia Ann Ivonne Vandenbulcke[1], MSc; Hein Petrus Johannes van Hout[3,4], MSc, PhD; Jan De Lepeleire[5], MD, PhD; Anja Declercq[2,6], MSc, PhD; Dominique Declerck[1], DDS, PhD; Joke Duyck[1], DDS, PhD

[1]Population Studies in Oral Health, Department of Oral Health Sciences, KU Leuven, Kapucijnenvoer 7, Leuven, Belgium

[2]LUCAS, Center for Care Research and Consultancy, KU Leuven, Leuven, Belgium

[3]Departments of General Practice and Medicine for Older People, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

[4]Amsterdam Public Health Research Institute, Aging & Later Life Program, Amsterdam UMC, Amsterdam, The Netherlands

[5]Department of Public Health and Primary Care, Academic Center for General Practice, KU Leuven, Leuven, Belgium

[6]CeSO, Center for Sociological Research, KU Leuven, Leuven, Belgium

Corresponding Author:
Emilie Schoebrechts, MSc
Population Studies in Oral Health, Department of Oral Health Sciences, KU Leuven, Kapucijnenvoer 7, Leuven, Belgium

Abstract

Background: Oral health of nursing home residents is generally poor, largely due to age-related conditions that increase their vulnerability and make them more dependent on others for oral care. Poor oral health can significantly impact general health and well-being, highlighting the important role of caregivers in preventing, detecting, and addressing residents' oral health problems in time. The interRAI Suite of instruments, widely used for comprehensive health assessments, offers the opportunity for nondental caregivers to assess oral health as part of general health and well-being and facilitates the integration of oral health in general care planning.

Objective: Based on interRAI data, this study explored the associations between oral health and general health outcomes of nursing home residents in Flanders (Belgium) and the Netherlands.

Methods: This cross-sectional study included baseline interRAI assessments of 2362 Flemish and Dutch residents aged 65 years and older, collected by caregivers (eg, nurses, nurse aids) between October 2020 and February 2024. Validated outcome scales (eg, Activities of Daily Living Hierarchy Scale, Cognitive Performance Scale, and Depression Rating Scale), health conditions such as diabetes and low weight (BMI ≤19), and family support and participation in social activities, included in the interRAI instrument, provided information on residents' general health and well-being. Oral health was defined according to the 9 items of the optimized Oral Health Screener (OHS) for use in the interRAI instruments. Adjusted logistic regression models were used to explore the associations between oral and general health outcomes.

Results: Oral health problems were interconnected, with conditions in one oral structure impacting the health of others. When controlling for confounding variables, dependency on others for personal care and hygiene was significantly associated with poor oral hygiene (odds ratio [OR] 1.7, 95% CI 1.1 - 2.4), chewing difficulties (OR 2.2, 95% CI 1.3 - 3.5), compromised teeth (OR 1.7, 95% CI 1.1 - 2.5), and the need for dental referral (OR 1.8, 95% CI 1.4 - 2.3). Natural dentition and gender had a significant impact on oral health status. Low weight was associated with poor chewing function (OR 2.0, 95% CI 1.1 - 3.6) and dry mouth (OR 1.8, 95% CI 1.0 - 3.2), both associated with oral discomfort or pain (OR 6.8, 95% CI 3.3 - 14.0 and OR 3.5, 95% CI 1.8 - 6.9, respectively). Family support was identified as a facilitator in mitigating oral health problems.

Conclusions: The interRAI instrument, including the OHS, is a valuable tool for monitoring residents' health and identifying areas requiring additional support. The results demonstrated the importance of prioritizing oral health as an integral component of comprehensive care. Recognizing the interconnection between oral health and other health conditions, including indicators of well-being, can guide caregivers and other health care professionals in incorporating preventive and supportive oral care practices into daily care routines, contributing to improved resident health and well-being outcomes.

(JMIR Aging 2025;8:e72308) doi:10.2196/72308

KEYWORDS

advance care planning; long-term care; geriatric conditions; oral health; oral care; geriatric assessment

XSL•FO
RenderX

Introduction

Oral diseases are among the most widespread noncommunicable diseases, affecting nearly half of the population worldwide, with older adults among the most vulnerable persons [1]. Common oral diseases and conditions in older persons include dental caries, periodontitis, tooth loss, and dry mouth [2]. Although most oral diseases are largely preventable with adequate daily care and regular dental checkups, older adults face several challenges, making it difficult to maintain good oral health. These challenges include a reduced ability to self-care due to compromised general health, polypharmacy, difficult accessibility, and inadequate availability of dental care [3,4].

Suboptimal oral health can have major consequences for general health and quality of life and vice versa [5]. Research has shown a bidirectional relationship between periodontitis and systemic diseases such as diabetes and cardiovascular disease [6]. Individuals with cognitive impairment may have difficulty understanding the importance of oral health and maintaining consistent hygiene routines [3,7]. Conversely, tooth loss has been associated with cognitive decline, due to reduced chewing ability [8]. Wearing dentures at night can increase the risk of pneumonia due to accumulated plaque levels [3]. In addition, certain types of medications can reduce saliva production, increasing the risk of oral health problems [9]. Poor oral health can also lower a person's self-esteem, leading to social isolation and depressive symptoms [3,10]. Conversely, depression and other mental health conditions often result in neglected oral hygiene due to a lack of motivation or energy, which worsens oral health [3,10]. Given the complex interactions between oral and general health, it is important to address oral health problems in time to prevent further health deterioration.

Through regular oral health assessments, caregivers can help improve oral health of older adults by enabling prevention, early detection, and timely treatment of oral health problems [11,12]. In this regard, the interRAI Suite of instruments is useful due to its holistic approach and wide range of applications [13].

The interRAI Suite consists of comprehensive web-based assessment instruments for caregivers to collect information in a standardized way from care-dependent persons on key factors including physical and cognitive function, psychosocial factors, clinical situation, social support, and so on. These instruments have been introduced in more than 35 countries and are compatible across different care settings. The interRAI Suite provides reliable and high-quality data to improve continuity of care, promote a person-centered approach, and enhance the ability to assess clinical outcomes. The integration of collaborative action points (CAPs), which are trigger algorithms to identify care needs and the risk of adverse outcomes, facilitates collaborative decision-making and helps prevent health deterioration and improve quality of life. Since the interRAI instrument for use in long-term care facilities (LTCFs) (eg, nursing homes) includes oral health as an integral part of general health and well-being, it has the potential to integrate oral care into general care planning [13,14].

However, research has shown that the oral health section within version 9 of the interRAI instrument is incomplete and has limited validity [15]. Replacing the oral health section in the interRAI LTCF with an optimized oral health section may allow caregivers to more effectively identify oral health problems, while also considering the other non–oral health conditions indicated in the assessments that may serve as risk factors [16].

Therefore, an optimized oral health section, the ohr-interRAI, was developed and recently identified in a systematic review as the most appropriate oral health instrument for nondental caregivers to evaluate older adults' oral health [12,15,17]. Based on a Delphi study with 53 oral health experts from 34 countries, further improvements were made, resulting in the Oral Health Screener (OHS) for use in the interRAI instruments, which takes into account national and cultural differences that may exist regarding oral health, oral care, and oral behavior [18].

The OHS consists of 9 items: chewing function, discomfort or pain in the mouth, dry mouth, hygiene of removable dentures, oral hygiene, teeth, gums, tongue, and palate and inner surface of cheeks and lips. Based on the assessment of these items, the CAP *oral hygiene* and the CAP *referral to a dentist* will be activated when help with daily oral hygiene and a referral to a dentist is needed, respectively. To help caregivers identify oral health problems, photographs, guidelines on the general assessment, items and CAPs, and instruction videos are available [17,18].

Although not yet officially included in the global interRAI Suite of instruments, the OHS is already in use in nursing homes in Flanders (Belgium) and in the Netherlands. Since the interRAI LTCF instrument collects data on oral health and general health, including well-being, from residents, it allows for assessing the associations between these outcomes. This study is the first to explore such associations using the OHS. Given the limited validity of the current oral health section and the widespread use of the interRAI instruments, this study aims to support the implementation of the OHS in the interRAI instruments and to strengthen the interRAI's overarching goals. Understanding the links between oral health and other health problems can facilitate early intervention and comprehensive care planning, thereby improving residents' quality of life. In addition, recognizing these associations contributes to prevention, supporting cost-effective health care, and guides the development of policies and preventive programs targeting oral and general health in older populations.

Methods

Study Design and Inclusion Criteria

This cross-sectional study included interRAI LTCF baseline data from persons aged 65 years and older living in nursing homes in Flanders (Belgium) and in the Netherlands. The data were collected by professional caregivers (nurses, physiotherapists, occupational therapists, etc) using the interRAI LTCF instrument, including the OHS, to assess residents' oral and general health status. This study differs from a previous study by Schoebrechts et al [19], reporting the prevalence of oral health problems identified using the OHS, by including data from a larger number of residents and examining

XSL•FO
RenderX

associations with other health and well-being indicators. All data were collected between October 2020 and February 2024.

Ethical Considerations

This multicenter study was approved by the Belgian Privacy Commission and the Ethics Committee Research UZ/KU Leuven (B3222021000448). All participants or their family members gave informed consent for the use of their data for research purposes. To ensure privacy and confidentiality, all data were pseudonymized and data access was limited to the research team. No compensation was provided for participation.

Outcomes

Oral Health

Using the OHS, the assessment of chewing function, discomfort or pain in the mouth, and dry mouth requires a conversation or observation during meals, while the assessment of *hygiene of removable dentures*, *oral hygiene*, *teeth*, *gums*, *tongue*, and *palate and inner surface of cheeks and lips* requires a visual inspection in the mouth. These items are assessed on a scale that distinguishes between acceptable and unacceptable conditions, with photographs available for the 6 items requiring a visual inspection. In addition, caregivers have the option to indicate whether the assessment is not possible (eg, if persons resist) or not applicable (eg, no dentures). The CAP *oral hygiene* and the CAP *referral to a dentist* are activated when help with daily oral hygiene and a referral to a dentist are needed, respectively.

General Health

General health, including well-being, outcomes in the interRAI instrument were selected based on previous research indicating their relevance to oral health. The interRAI instrument includes internationally validated outcome scales to evaluate a person's clinical status, which are automatically calculated based on a number of individual items [20]. The following scales were included in the analyses. They were dichotomized using validated cutoff values, with higher scores indicating greater severity.

The Activities of Daily Living Hierarchy (ADLH) Scale, ranging from 0 to 6 (cutoff ≥3), refers to the degree of dependence on personal care, based on the items related to personal hygiene, toilet use, locomotion, and eating [21]. The Cognitive Performance Scale (CPS), ranging from 0 to 6 (cutoff ≥3), represents a person's cognitive status, including cognitive skills in daily decision-making, short-term memory, procedural memory, ability to make self-understood, and eating self-performance [22]. The Depression Rating Scale (DRS), ranging from 0 to 14 (cutoff ≥3), represents the presence of depressive symptoms based on several items such as making negative statements, expressions of unrealistic fears, and repetitive health complaints or crying [23]. The Changes in Health, End-Stage Disease, Signs, and Symptoms (CHESS) Scale, ranging from 0 to 5 (cutoff ≥3), demonstrates a person's health instability and is a strong predictor of mortality. The scale is based on several items such as changes in decision-making and activities of daily living status, dyspnea, dehydration, and weight loss [24]. The Communication Scale has a range from 0 to 8 (cutoff ≥4) and is calculated based on residents' ability to make themselves understood and to understand others [25]. The Revised Index of Social Engagement, ranging from 0 to 6 (cutoff ≥3), is calculated from items describing an individual's sense of initiative and social involvement in the nursing home. Unlike the other scales, a higher Revised Index of Social Engagement score indicates a positive outcome [26].

In addition, residents' general health was evaluated based on the presence of the following health diseases recorded in the interRAI instrument: dementia, diabetes mellitus, congestive heart failure, chronic obstructive pulmonary disease (COPD), Parkinson disease, anxiety, bipolar disorder, schizophrenia, and pneumonia. Other health conditions included aspiration, reflux, low weight (BMI ≤19), and polypharmacy (≥10 medications). Information on residents' mood or behavior status and psychosocial well-being was collected by including the interRAI items *resistance to care*, *reduced social interactions*, *strong and supportive relationships with family*, and the CAP *behavior*. This CAP is calculated based on residents' wandering behavior, verbal and physical abusive attitude, socially inappropriate or disruptive behavior, and resistance to care. In addition, information in the interRAI instrument such as *extensive to total dependency for personal hygiene*, *modified mode of nutritional intake* (due to not swallowing all types of food), and *severe vision impairment in adequate light* was included in the analyses as well as the item on having a dental checkup in the last year and information on a physician visit in the last 14 days. All items were dichotomized.

Covariates

The following variables were considered covariates: age (divided into 3 groups: 65 - 75, 76 - 85, and 86+ years), sex (female and male), smoking (smoker and nonsmoker), presence of teeth (based on the *not applicable* response option of item *teeth*), and presence of dentures (based on the *not applicable* response option of item *denture hygiene*).

Missingness Information

Analysis of missing values showed that they were random, which means that the probability of a missing value did not depend on the item itself or on certain characteristics of the assessors [19,27]. Possible causes for incomplete assessments with the interRAI instrument are discussed in detail by Vanneste et al [28].

Data Analyses

Descriptive statistics were used to illustrate the characteristics (eg, demographic information and oral and general health status) of the residents. Categorical variables were summarized as absolute numbers and percentages, and continuous variables were presented as means and accompanying standard deviations.

For bivariate analyses, the oral health items, originally assessed using 3 or 4 response options, were dichotomized considering only the options when an assessment was possible (ie, acceptable or unacceptable condition). Unadjusted logistic regression models were performed to explore associations between the oral health–related and the selected general health–related

XSL•FO
RenderX

variables. These analyses provided a preliminary understanding of the direct associations between the individual variables. Subsequently, multivariable logistic regression models were developed for each oral health item adjusted for oral and general health variables. Each model included a fixed set of explanatory variables (*ADLH/extensive to total dependency for personal hygiene*, *CPS*, *DRS*, *CHESS*, *Communication Scale*, *strong and supportive relationships with family*, and covariates), together with variables identified as significant in the bivariate analyses or supported by previous research. The variables considered as covariates were included in the regression models to control for potential confounding variables. *Denture hygiene*, *oral hygiene*, and *teeth* were excluded as explanatory variables in the models because these variables were assessed only in residents with dentures and natural teeth, respectively. Including these variables would have restricted the models to a limited subgroup of residents without dentures, without natural teeth, or without both. To address collinearity, variables measuring the same construct or part of another item (eg, *extensive to total dependency for personal hygiene* and *ADLH scale*, *resistance to care* and CAP *behavior*, and *presence of teeth* and *teeth*) were excluded from the models. *Extensive to total dependency for personal hygiene* variable was included instead of the ADLH Scale in the models for *denture hygiene*, *oral hygiene*, and CAP *oral hygiene*, as it may provide a better indication of these items' status. The associations between the variables were expressed using odds ratios (ORs) and 95% CIs. All statistical analyses were performed using IBM SPSS Statistics (version 28.0.1.1; IBM) and SAS Enterprise Guide (version 8.1; SAS Institute). A *P* value of <.05 was considered statistically significant.

Results

Baseline Characteristics

The data consisted of 2362 baseline assessments (ie, first assessments of the residents using the interRAI instrument, including the OHS), 789 from 12 Flemish nursing homes and 1573 from 26 Dutch nursing homes. The mean age of the residents was 82.17 (SD 7.53) years and most residents were female (1567/2362, 66.3%). About 8.0% (180/2346) of the residents were daily smokers (Table 1).

Table . Baseline characteristics of the participating nursing home residents.

Characteristics	Value (N=2362)
Age (years), mean (SD)	82.17 (7.53)
Age range (years), n (%)[a]	
65 - 75	633 (26.8)
76 - 85	1121 (47.5)
86+	608 (25.7)
Sex, n (%)[a]	
Female	1567 (66.3)
Male	795 (33.7)
Daily smoker, n (%)[a]	180 (7.7)

[a]Specific outcome per item.

Oral Health

Table 2 shows the distribution of the oral health conditions in the sample. Of the 3 self-reported items, chewing problems (290/2254, 12.9%) were the most reported condition by the nursing home residents. About 10.0% (226/2214) of the residents indicated having a dry mouth, and 4.0% (89/2219) reported discomfort or pain in the mouth. The condition of the teeth could be assessed in 49.8% (1147/2301) of the nursing home residents. This was found to be compromised in 25.6% (294/1147) of these residents. The second most commonly observed oral health problem was poor oral hygiene, identified in 20.6% (313/1519) of the residents with assessable oral hygiene (1519/2297, 66.1%). Hygiene of removable dentures could be assessed in 64.9% (1499/2310) of the participants with dentures and was found to be poor in 9.1% (136/1499). Gum problems were observed in 7.7% (155/2006) of the residents, problems with the tongue were observed in 4.0% (87/2057), and problems with the palate and inner surface of cheeks and lips were also observed in about 4.0% (76/1959) of the residents.

XSL•FO
RenderX

Table . The oral health status of the participating nursing home residents.

Oral health items	Residents, n/N[a] (%)
Chewing function (n=2323)	
Assessed	2254/2323 (97.0)
Good/acceptable	1964/2254 (87.1)
Poor/unacceptable	290/2254 (12.9)
Not assessed	69/2323 (3.0)
Cannot be assessed[b]	42/69 (60.9)
Not applicable (eg, blended food due to swallowing difficulties and tube feeding)	27/69 (39.1)
Missing oral health data	39/2362 (1.6)
Discomfort or pain in the mouth (n=2346)	
Can be assessed	2219/2346 (94.6)
No	2130/2219 (96.0)
Yes	89/2219 (4.0)
Not assessed	127/2346 (5.4)
Cannot be assessed	127/127 (100)
Missing oral health data	16/2362 (0.7)
Dry mouth (n=2346)	
Can be assessed	2214/2346 (94.4)
No	1988/2214 (89.8)
Yes	226/2214 (10.2)
Not assessed	132/2346 (5.6)
Cannot be assessed	132/132 (100)
Missing oral health data	16/2362 (0.7)
Hygiene of removable dentures (n=2310)	
Can be assessed	1499/2310 (64.9)
Good /acceptable	1363/1499 (90.9)
Poor/unacceptable	136/1499 (9.1)
Not assessed	811/2310 (35.1)
Cannot be assessed	183/811 (22.6)
Not applicable (no dentures)	628/811 (77.4)
Missing oral health data	52/2362 (2.2)
Oral hygiene (n=2297)	
Can be assessed	1519/2297 (66.1)
Good/acceptable	1206/1519 (79.4)
Poor/unacceptable	313/1519 (20.6)
Not assessed	778/2297 (33.9)
Cannot be assessed	217/778 (27.9)
Not applicable (no teeth or denture retainers)	561/778 (72.1)
Missing oral health data	65/2362 (2.7)
Teeth (n=2301)	
Can be assessed	1147/2301 (49.8)
Good/acceptable	853/1147 (74.4)

XSL•FO
RenderX

Oral health items	Residents, n/N[a] (%)
Poor/unacceptable	294/1147 (25.6)
Not assessed	1154/2301 (50.2)
Cannot be assessed	193/1154 (16.7)
Not applicable (no teeth and no root remnants)	961/1154 (83.3)
Missing oral health data	61/2362 (2.6)
Gums (n=2295)	
Can be assessed	2006/2295 (87.4)
Good/acceptable	1851/2006 (92.3)
Poor/unacceptable	155/2006 (7.7)
Not assessed	289/2295 (12.6)
Cannot be assessed	289/289 (100)
Missing oral health data	67/2362 (2.8)
Tongue (n=2308)	
Can be assessed	2057/2308 (89.1)
Good/acceptable	1970/2057 (95.8)
Poor/unacceptable	87/2057 (4.2)
Not assessed	251/2308 (10.9)
Cannot be assessed	251/251 (100)
Missing oral health data	54/2362 (2.3)
Palate and inner surface of cheeks and lips (n=2299)	
Can be assessed	1959/2299 (85.2)
Good/acceptable	1883/1959 (96.1)
Poor/unacceptable	76/1959 (3.9)
Not assessed	340/2299 (14.8)
Cannot be assessed	340/340 (100)
Missing oral health data	63/2362 (2.7)

[a]Data included in the model/total available oral health data.

[b]"Cannot be assessed" when the item could not be assessed adequately, for example, due to resistant behavior or cognitive impairment.

Based on the results of the assessments with the OHS, the CAP *oral hygiene* was activated in 16.8% (384/2289), indicating a need for help with daily oral care, and the CAP *referral to a dentist* was triggered in 27.5% (611/2221), indicating a need for dental treatment. In addition, caregivers indicated a dental checkup in the last year in 56.5% (1319/2333) of the nursing home residents.

General Health

Analyses of the outcome scales revealed that the majority of the residents were dependent on others for activities of daily living (1276/2351, 54.3%) and were cognitively impaired (1335/2341, 57.0%). Symptoms of depression were found in 36.7% (862/2346) and poor communication in 24.3% (5272/2351). Based on the CHESS Scale, health instability was present in 5.5% (126/2296) of the nursing home residents. Most residents (1902/2337, 81.4%) showed initiative and social involvement in the nursing home.

About 30.0% (706/2345) of the residents were experiencing dementia, 21.6% (506/2344) were experiencing diabetes, and 27.0% (634/2346) had behavior problems. The use of polypharmacy, indicated by taking 10 or more medications, was noted in 30.7% (721/2345). Strong and supportive relationships with family were reported by 88.2% (2070/2346) of the residents. More information on residents' general health and well-being is shown in Table 3.

XSL•FO
RenderX

Table . The general health status of the participating nursing home residents.

	Residents (N=2362), n/N[a] (%)
Scales	
Activities of Daily Living Hierarchy (0 - 6) ≥3	1276/2351 (54.3)
Cognitive Performance Scale (0 - 6) ≥3	1335/2341 (57.0)
Depression Rating Scale (0 - 14) ≥3	862/2346 (36.7)
Changes in Health, End-Stage Disease and Symptoms and Signs (0 - 5) ≥3	126/2296 (5.5)
Communication (0 - 8) ≥4	5272/2351 (24.3)
Revised Index of Social Engagement (0 - 6) ≥3	1902/2337 (81.4)
Diseases/disorders	
Anxiety	278/2350 (11.8)
Schizophrenia	28/2348 (1.2)
Bipolar	33/2349 (1.4)
Pneumonia	42/2348 (1.8)
Aspiration	187/2350 (8.0)
Reflux	74/2353 (3.1)
Dementia	706/2345 (30.1)
Diabetes mellitus	506/2344 (21.6)
Parkinson	157/2350 (6.7)
Chronic obstructive pulmonary disease	213/2344 (9.1)
Congestive heart failure	539/2342 (23.0)
Polypharmacy (≥10 medications)	721/2345 (30.7)
Low weight (BMI ≤19)	274/2257 (12.1)
Modified mode of nutritional intake	437/2356 (18.5)
Behavior problems	634/2346 (27.0)
Resistance to care	456/2349 (19.4)
Extensive to total dependency for personal hygiene	1139/2353 (48.4)
Severe vision impairment in adequate light	78/2348 (3.3)
Reduced social interactions	680/2347 (29.0)
Strong and supportive relationships with family	2070/2346 (88.2)
Physician visit in the last 14 days	901/2315 (38.9)
Dental checkup in last year	1319/2333 (56.5)

[a]Specific outcome per item/total available data per item.

Associations Between Oral Health and General Health Outcomes

Table S1 in Multimedia Appendix 1 and Tables 4 and 5 show the results of the bivariate analyses and the associated factors for each oral health item in the adjusted multivariable regression models, respectively.

XSL•FO
RenderX

Table . Factors associated with each oral health item (reference category: acceptable condition) in the adjusted logistic models.

	Chewing function (1564/2254)[a]		Discomfort or pain in the mouth (1609/2219)[a]		Dry mouth (1559/2214)[a]		Denture hygiene (1206/1499)[a]		Oral hygiene (1171/1519)[a]	
	OR[b]	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Age (years; 65-75: reference category)										
76 - 85	1.93	1.15 - 3.25[c]	0.88	0.42 - 1.84	0.77	0.49 - 1.21	1.11	0.63 - 1.96	0.95	0.63 - 1.45
>85	1.93	1.08 - 3.44[c]	1.04	0.45 - 2.39	1.14	0.69 - 1.89	0.94	(0.48 - 1.83	0.97	0.58 - 1.62
Sex (male: reference category)										
Female	0.65	0.42 - 1.02	0.73	(0.37 - 1.44	0.68	0.45 - 1.03	1.33	0.84 - 2.11	1.57	1.08 - 2.28[c]
Smoker	1.09	0.48 - 2.47	0.45	0.11 - 1.74	1.12	0.58 - 2.17	1.53	0.73 - 3.19	1.47	0.77 - 2.81
ADLH[d] Scale ≥3	2.16	1.34 - 3.49[e]	0.75	0.40 - 1.38	1.06	0.72 - 1.56	—[f]	—	—	—
Extensive to total dependency for personal hygiene	—	—	—	—	—	—	1.39	0.87 - 2.22	1.65	1.13 - 2.41[c]
CPS[g] ≥3	1.29	0.80 - 2.07	0.91	0.46 - 1.81	0.54	0.36 - 0.82[e]	1.05	0.63 - 1.75	0.98	[0.65 - 1.47
DRS[h] ≥3	0.85	0.56 - 1.28	1.86	0.99 - 3.50	1.18	0.80 - 1.73	0.71	0.42 - 1.19	1.28	0.87 - 1.87
CHESS[i] ≥3	0.71	0.32 - 1.58	0.99	0.32 - 3.09	2.14	1.09 - 4.22[c]	0.28	0.06 - 1.32	0.67	0.28 - 1.59
Communication Scale ≥4	1.61	0.99 - 2.62	0.99	0.42 - 2.35	0.79	0.44 - 1.43	1.07	0.56 - 2.06	0.95	0.59 - 1.53
RISE[j] ≥3	1.22	0.72 - 2.06	—	—	—	—	—	—	0.80	0.49 - 1.30
Family support	0.74	0.41 - 1.31	0.98	0.41 - 2.35	1.28	0.70 - 2.34	0.76	0.38 - 1.52	0.59	0.35 - 1.00[c]
Chewing problems	—	—	6.80	3.30 - 14.03[k]	2.77	1.60 - 4.78[k]	0.55	0.24 - 1.31	1.20	0.66 - 2.18
Discomfort or pain in the mouth	9.06	4.43 - 18.54[k]	—	—	3.53	1.80 - 6.90[k]	1.51	0.53 - 4.25	3.18	1.37 - 7.35[e]
Dry mouth	2.83	1.66 - 4.83[k]	3.47	1.82 - 6.65[k]	—	—	1.56	0.77 - 3.16	1.21	0.68 - 2.16
Gum problems	3.83	1.82 - 8.06[k]	2.38	0.86 - 6.58	0.71	0.30 - 1.70	5.31	2.26 - 12.44[k]	7.47	3.95 - 14.14[k]
Tongue problems	0.83	0.33 - 2.10	2.29	0.81 - 6.51	2.56	1.19 - 5.47[c]	6.06	2.64 - 13.94[k]	3.24	1.29 - 8.18[c]
Problems with palate and inner surface of cheeks and lips	1.07	0.41 - 2.79	1.73	0.52 - 5.79	2.51	1.01 - 6.25[c]	0.75	0.21 - 2.65	1.39	0.55 - 3.54
Presence of teeth	0.65	0.43 - 0.99[c]	0.66	0.35 - 1.25	0.66	0.45 - 0.98[c]	1.54	0.97 - 2.45	—	—
Presence of dentures	0.67	0.43 - 1.05	1.30	0.63 - 2.69	0.70	0.45 - 1.08	—	—	0.23	0.16 - 0.33[k]

XSL•FO
RenderX

	Chewing function (1564/2254)[a]		Discomfort or pain in the mouth (1609/2219)[a]		Dry mouth (1559/2214)[a]		Denture hygiene (1206/1499)[a]		Oral hygiene (1171/1519)[a]	
	OR[b]	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Modified mode of nutritional intake	8.77	5.60 - 13.73[k]	0.38	0.15 - 0.97[c]	0.78	0.43 - 1.40	—	—	1.09	0.61 - 1.96
Low weight (BMI 19)	2.02	1.12 - 3.64[c]	—	—	1.80	1.02 - 3.18[c]	—	—	1.48	0.84 - 2.63
Polypharmacy (≥10 medications)	—	—	1.82	0.98 - 3.36	1.47	1.00 - 2.16[c]	—	—	—	—
Resistance to care	1.13	0.67 - 1.92	—	—	0.49	0.26 - 0.94[c]	1.38	0.73 - 2.59	1.69	1.07 - 2.66[c]
CAP[l] behavior	—	—	1.39	0.69 - 2.82	—	—	—	—	—	—
Reduced social interactions	—	—	1.24	0.63 - 2.43	—	—	—	—	—	—
Severe vision impairment in adequate light	1.38	0.50 - 3.79	—	—	—	—	1.30	0.30 - 5.71	1.12	0.38 - 3.29
Congestive heart failure	0.57	0.35 - 0.93[c]	1.17	0.61 - 2.25	1.33	0.89 - 2.00	—	—	0.94	0.61 - 1.45
Aspiration	2.38	1.35 - 4.20[e]	—	—	2.28	1.21 - 4.28[c]	—	—	1.69	0.85 - 3.37
Pneumonia	2.07	0.52 - 8.19	0.71	0.09 - 5.28	1.63	0.51 - 5.17	0.88	0.17 - 4.45	1.41	0.41 - 4.81
Diabetes	—	—	—	—	—	—	—	—	—	—
Reflux	—	—	—	—	2.56	1.20 - 5.44[c]	—	—	—	—
COPD[m]	—	—	—	—	2.07	1.26 - 3.41[e]	—	—	—	—
Parkinson disease	—	—	—	—	1.36	0.68 - 2.71	—	—	—	—

[a]Data included in the model/total available oral health data.

[b]OR: odds ratio (significant OR >1: significant contributor; significant OR <1: significant protector).

[c]*P*<.05.

[d]ADLH: Activities of Daily Living Hierarchy.

[e]*P*<.01.

[f]Not included in the model.

[g]CPS: Cognitive Performance Scale.

[h]DRS: Depression Rating Scale.

[i]CHESS: Changes in Health, End-stage disease, Signs, and Symptoms.

[j]RISE: Revised Index of Social Engagement.

[k]*P*<.001.

[l]CAP: collaborative action point.

[m]COPD: chronic obstructive pulmonary disease.

XSL•FO
RenderX

Table . Factors associated with each oral health item (reference category: acceptable condition) in the adjusted logistic models.

	Teeth (895/1147)[a]		Gums (1657/2006)[a]		Tongue (1665/2057)[a]		Palate and inner surface of cheeks and lips (1605/1959)[a]		CAP[b] oral hygiene (1629/2289)[a]		CAP referral to a dentist (2081/2221)[a]	
	OR[c]	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Age (years; 65-75: reference category)												
76 - 85	1.72	1.10 - 2.68[d]	1.06	0.61 - 1.85	0.76	0.38 - 1.50	0.68	0.33 - 1.39	0.93	0.66 - 1.32	0.91	0.70 - 1.17
>85	1.51	0.87 - 2.61	1.13	0.58 - 2.19	1.11	0.52 - 2.36	0.82	0.36 - 1.85	0.83	0.55 - 1.26	0.88	0.65 - 1.20
Sex (male: reference category)												
Female	1.35	0.91 - 1.98	1.17	0.72 - 1.90	1.22	0.68 - 2.19	1.25	0.66 - 2.35	1.47	1.08 - 1.99[d]	0.97	0.77 - 1.22
Smoker	1.05	0.49 - 2.25	0.92	0.36 - 2.36	1.97	0.80 - 4.88	1.59	0.61 - 4.14	1.13	0.67 - 1.91	1.03	0.69 - 1.55
ADLH[e] Scale ≥3	1.66	1.12 - 2.47[d]	1.06	0.65 - 1.73	1.55	0.84 - 2.85	1.68	0.87 - 3.25	—[f]	—	1.81	1.43 - 2.30[g]
Extensive to total dependency for personal hygiene	—	—	—	—	—	—	—	—	1.63	1.20 - 2.22[h]	—	—
CPS[i] ≥3	0.82	0.54 - 1.25	1.05	0.62 - 1.77	1.12	0.59 - 2.12	0.53	0.26 - 1.11	1.07	0.77 - 1.50	1.09	0.85 - 1.41
DRS[j] ≥3	1.14	0.76 - 1.71	1.37	0.84 - 2.24	1.14	0.65 - 2.03	1.86	1.00 - 3.47[d]	1.11	0.81 - 1.53	1.27	1.02 - 1.59[d]
CHESS[k] ≥3	0.71	0.26 - 1.90	0.41	0.12 - 1.37	3.68	1.62 - 8.37[h]	2.16	0.84 - 5.54	0.77	0.37 - 1.61	1.11	0.72 - 1.72
Communication Scale ≥4	1.05	0.63 - 1.76	0.99	0.53 - 1.84	1.14	0.53 - 2.43	1.54	0.67 - 3.56	1.05	0.71 - 1.57	1.35	1.03 - 1.77[d]
RISE[l] ≥3	1.04	0.63 - 1.73	—	—	—	—	—	—	—	—	1.09	0.83 - 1.43
Family support	0.52	0.30 - 0.89[d]	0.59	0.32 - 1.10	0.91	0.42 - 2.00	0.56	0.26 - 1.21	0.73	0.48 - 1.12	0.68	0.50 - 0.92[d]
Chewing problems	2.20	1.19 - 4.07[d]	2.81	1.46 - 5.39[h]	0.69	0.29 - 1.67	1.62	0.69 - 3.79	0.99	0.59 - 1.66	—	—
Discomfort or pain in the mouth	2.02	0.78 - 5.22	2.94	1.28 - 6.72[d]	3.14	1.28 - 7.70[d]	4.84	2.03 - 11.54[g]	1.48	0.74 - 2.95	—	—
Dry mouth	1.56	0.85 - 2.86	1.16	0.58 - 2.34	3.17	1.64 - 6.14[g]	2.88	1.41 - 5.86[h]	1.30	0.81 - 2.08	—	—
Gum problems	7.23	3.92 - 13.30[g]	—	—	5.76	2.73 - 12.12[g]	—	—	8.08	4.66 - 14.00[g]	—	—
Tongue problems	3.32	1.25 - 8.80[d]	5.99	2.70 - 13.29[g]	—	—	—	—	3.68	1.92 - 7.06[g]	—	—
Problems with palate and inner surface of cheeks and lips	1.51	0.55 - 4.12	—	—	—	—	—	—	1.00	0.45 - 2.18	—	—
Presence of teeth	—	—	5.23	2.83 - 9.66[g]	0.63	0.34 - 1.16	1.81	0.92 - 3.56	—	—	—	—
Presence of dentures	0.48	0.33 - 0.70[g]	0.69	0.43 - 1.11	0.63	0.33 - 1.18	0.61	0.32 - 1.17	—	—	—	—

XSL•FO
RenderX

	Teeth (895/1147)[a]		Gums (1657/2006)[a]		Tongue (1665/2057)[a]		Palate and inner surface of cheeks and lips (1605/1959)[a]		CAP[b] oral hygiene (1629/2289)[a]		CAP referral to a dentist (2081/2221)[a]	
	OR[c]	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Modified mode of nutritional intake	0.97	0.53 - 1.75	1.12	0.56 - 2.26	1.16	0.52 - 2.58	0.77	0.31 - 1.90	0.71	0.44 - 1.14	3.47	2.64 - 4.57[g]
Low weight (BMI ≤19)	—	—	—	—	—	—	1.22	0.50 - 2.98	1.58	0.98 - 2.55	1.35	0.95 - 1.90
Polypharmacy (≥10 medications)	—	—	—	—	—	—	—	—	—	—	—	—
Resistance to care	1.83	1.13 - 2.98[d]	1.45	0.82 - 2.58	—	—	1.05	0.46 - 2.36	1.47	0.99 - 2.17	0.95	0.72 - 1.25
CAP behavior	—	—	—	—	—	—	—	—	—	—	—	—
Reduced social interactions	—	—	1.01	0.60 - 1.70	—	—	—	—	0.90	0.63 - 1.27	—	—
Severe vision impairment in adequate light	0.67	0.19 - 2.40	1.23	0.38 - 4.01	3.66	1.25-10.73[d]	—	—	1.28	0.52 - 3.13	1.29	0.75 - 2.24
Congestive heart failure	—	—	1.24	0.74 - 2.10	—	—	—	—	0.68	0.48 - 0.99[d]	0.93	0.72 - 1.20
Aspiration	—	—	1.06	0.44 - 2.56	0.83	0.28 - 2.48	—	—	—	—	1.50	1.04 - 2.17[d]
Pneumonia	—	—	1.79	0.49 - 6.56	1.56	0.29 - 8.23	1.44	0.27 - 7.67	1.30	0.45 - 3.77	2.11	1.00 - 4.44
Diabetes	—	—	0.79	0.44 - 1.43	2.46	1.38 - 4.36[h]	—	—	—	—	—	—
Reflux	—	—	—	—	—	—	—	—	—	—	—	—
COPD[m]	—	—	—	—	—	—	—	—	—	—	—	—
Parkinson disease	—	—	—	—	—	—	—	—	—	—	—	—

[a]Data included in the model/total available oral health data.

[b]CAP: collaborative action point.

[c]OR: odds ratio (significant OR >1: significant contributor; significant OR <1: significant protector).

[d]*P*<.05.

[e]ADLH: activities of daily living hierarchy.

[f]Not included in the model.

[g]*P*<.001.

[h]*P*<.01.

[i]CPS: Cognitive Performance Scale.

[j]DRS: Depression Rating Scale.

[k]CHESS: Changes in Health, End-stage disease, Signs, and Symptoms.

[l]RISE: Revised Index of Social Engagement.

[m]COPD: chronic obstructive pulmonary disease.

The following factors were significantly associated with *chewing problems*, after controlling for confounders: discomfort or pain in the mouth (OR 9.1, 95% CI 4.4 - 18.5), modified mode of nutritional intake (OR 8.8, 95% CI 5.6 - 13.7), gum problems (OR 3.8, 95% CI 1.8 - 8.1), dry mouth (OR 2.8, 95% CI 1.7 - 4.8), aspiration problems (OR 2.4, 95% CI 1.4 - 4.2), ADLH problems (OR 2.2, 95% CI 1.3 - 3.5), and low weight (OR 2.0, 95% CI 1.1 - 3.6). This means that older adults with these conditions were more likely to have chewing difficulties than older adults with no such conditions. In addition, older age (more than 76 years) was also associated with more chewing problems (OR approximately 2.0, 95% CI 1.1 - 3.4) than those aged 65 - 75 years. The presence of teeth (OR 0.7, 95% CI 0.4 - 1.0) contributed to a better chewing function. Having heart failure (OR 0.6, 95% CI 0.4 - 0.9) was associated with fewer chewing problems.

XSL•FO
RenderX

In the adjusted logistic model for *discomfort or pain in the mouth*, chewing function (OR 6.8, 95% CI 3.3 - 14.0) and dry mouth (OR 3.5, 95% CI 1.8 - 6.7) had significant ORs, which means that residents with chewing difficulties and dry mouth had a higher chance of discomfort or pain in the mouth than those with no such conditions. A modified mode of nutritional intake (OR 0.4, 95% CI 0.2 - 1.0) was associated with less discomfort or pain in the mouth.

Chewing difficulties (OR 2.8, 95% CI 1.6 - 4.8) and discomfort or pain in the mouth (OR 3.5, 95% CI 1.8 - 6.9) were also identified as associated factors for a *dry mouth*. In addition, reflux (OR 2.6, 95% CI 1.2 - 5.4), poor conditions of the tongue (OR 2.6, 95% CI 1.2 - 5.5) and palate and inner surface of cheeks and lips (OR 2.5, 95% CI 1.0 - 6.3), aspiration problems (OR 2.3, 95% CI 1.2 - 4.3), health instability derived from the CHESS Scale (OR 2.1, 95% CI 1.0 - 4.2), COPD (OR 2.1, 95% CI 1.3 - 3.4), and low weight (OR 1.8, 95% CI 1.0 - 3.2) contributed to experiencing a dry mouth. Residents taking 10 or more medications (OR 1.5, 95% CI 1.0 - 2.2) were also more likely to have a dry mouth. The presence of teeth (OR 0.7, 95% CI 0.5 - 1.0), cognitive impairment (OR 0.5, 95% CI 0.4 - 0.8), and resisting care (OR 0.5, 95% CI 0.3 - 0.9) showed protective effects, reducing the experience of dry mouth.

Gum (OR 5.3, 95% CI 2.3 - 12.4, and OR 7.5, 95% CI 4.0 - 14.1, respectively) and tongue problems (OR 6.1, 95% CI 2.6 - 13.9, and OR 3.2, 95% CI 1.3 - 8.2, respectively) are both contributing factors to poor *denture hygiene* and poor *oral hygiene*. In the oral hygiene model, residents with discomfort or pain in the mouth (OR 3.2, 95% CI 1.4 - 7.4), resistance to care (OR 1.7, 95% CI 1.1 - 2.7), or extensive to total dependence on personal hygiene (OR 1.7, 95% CI 1.1 - 2.4) were also more likely to have poor oral hygiene. It was found that males were more prone to poor oral hygiene than females (OR 1.6, 95% CI 1.1 - 2.3). Family support (OR 0.6, 95% CI 0.4 - 1.0) and the presence of dentures (OR 0.2, 95% CI 0.2 - 0.3) contributed to acceptable oral hygiene. Despite its lack of significance in the bivariate analyses, pneumonia status was included in the adjusted logistic models for denture hygiene and oral hygiene based on evidence from previous research. However, the models yielded no significant contribution to poor denture or oral hygiene.

The highest ORs in the adjusted logistic model for *teeth* were found for the condition of the gums (OR 7.2, 95% CI 3.9 - 13.3) and tongue (OR 3.3, 95% CI 1.3 - 8.8). This means that older adults presenting gum or tongue problems were more likely to have dental problems than those with no such conditions. Chewing difficulties yielded an OR of 2.2 (95% CI 1.2 - 4.1), indicating an association with poor condition of the teeth. When resistance to care (OR 1.8, 95% CI 1.1 - 3.0) or extensive dependence on personal care (ADLH; OR 1.7, 95% CI 1.1 - 2.5) was reported, the association with dental problems was significant. The same was true for adults aged 76 - 85 years (OR 1.7, 95% CI 1.1 - 2.7) compared with those aged 65 - 75 years. On the contrary, family support (OR 0.5, 95% CI 0.3 - 0.9) and the presence of dentures (OR 0.5, 95% CI 0.3 - 0.7) showed protective effects.

Gum problems were more frequently reported when residents had problems with the tongue (OR 6.0, 95% CI 2.7 - 13.3) or still had teeth (OR 5.2, 95% CI 2.8 - 9.7). In addition to both being associated with dry mouth, discomfort or pain in the mouth (OR 2.9, 95% CI 1.3 - 6.7) and chewing difficulties (OR 2.8, 95% CI 1.5 - 5.4) were also associated with gum problems. Based on evidence from previous research, diabetes and heart failure were included as variables in the adjusted logistic model. However, these conditions were not significantly associated with gum problems.

Gum problems (OR 5.8, 95% CI 2.7 - 12.1) also presented the highest OR in the adjusted logistic model for *tongue* problems, followed by health instability (OR 3.7, 95% CI 1.6 - 8.4). Severe vision impairment in adequate light (OR 3.7, 95% CI 1.3 - 10.7) and diabetes (OR 2.5, 95% CI 1.4 - 4.4) were also identified as contributors to tongue problems. Dry mouth (OR 3.2, 95% CI 1.6 - 6.1, and OR 2.9, 95% CI 1.4 - 5.9, respectively) and discomfort or pain in the mouth (OR 3.1, 95% CI 1.3 - 7.7, and OR 4.8, 95% CI 2.0 - 11.5, respectively) were both associated with tongue problems and *palate and inner surface of cheeks and lips*. In addition, depressive symptoms increased the chance of problems with the palate and inner surface of cheeks and lips (OR 1.9, 95% CI 1.0 - 3.5).

Gum problems (OR 8.1, 95% CI 4.7 - 14.0) were a major contributor to the activation of the CAP *oral hygiene*, as well as tongue problems (OR 3.7, 95% CI 1.9 - 7.1), highlighting the need for improved oral care. Extensive to total dependence on personal hygiene (OR 1.6, 95% CI 1.2 - 2.2) and being male compared with being female (OR 1.5, 95% CI 1.1 - 2.0) also increased the chance of activation of the CAP. The model showed that residents with heart failure (OR 0.7, 95% CI 0.5 - 1.0) were less likely to need help with daily oral care.

The CAP *referral to a dentist* was more likely to be activated when residents had a modified mode of nutritional intake (OR 3.5, 95% CI 2.6 - 4.6) or were dependent on personal care (ADLH; OR 1.8, 95% CI 1.4 - 2.3). Aspiration (OR 1.5, 95% CI 1.0 - 2.2), communication problems (OR 1.4, 95% CI 1.0 - 1.8), and depressive symptoms (OR 1.3, 95% CI 1.0 - 1.6) also increased the chance of referral to a dentist. Family support showed a protective effect in the activation of the CAP (OR 0.7, 95% CI 0.5 - 0.9).

Smoking was included in each adjusted logistic model but did not significantly contribute to the presence of oral health problems. Similarly, pneumonia was included in all models except the one for teeth (not significant in bivariate analysis) but showed no association when adjusted for the other variables. Reflux, COPD, and Parkinson disease were conditions included only in the model for dry mouth, with reflux and COPD showing significant associations with dry mouth, while Parkinson disease did not.

Discussion

Principal Findings

This study examined the associations between oral health and general health outcomes based on interRAI assessments of nursing home residents in Flanders and the Netherlands,

XSL•FO
RenderX

performed by their caregivers. This is the first study to explore these associations using the OHS, the optimized Oral Health Screener for use in the interRAI instruments.

Advancements in prevention, restorative techniques, and dental materials have enabled older adults to retain their natural dentition longer [29]. Despite changing patterns of tooth retention, tooth loss is still prevalent in the older population. Research indicates that most individuals aged 85 years and older in industrialized countries wear removable dentures, with the majority being completely edentulous [29]. In addition, nursing home residents tend to have fewer teeth than those living independently [29]. In this study, at least 42.0% (961/2301) of participating residents had lost all of their natural teeth and at least 65.0% (1499/2310) were wearing removable dentures. This finding aligns with the prevalence of edentulism (40.0% - 60.0%) and the use of removable dentures (40.0% - 85.0%) among nursing home residents in Europe, as reported in a systematic review by Janssens et al [30].

Natural teeth are much more complex and time-consuming to clean than an edentulous mouth with complete dentures [29]. Tooth loss is frequently the result of dental caries and gum inflammation, resulting from poor oral hygiene [31]. This study suggests that the combination of natural teeth and dentures is associated with better oral hygiene and condition of the teeth. However, the exact number of remaining teeth among the residents in this study was unknown. Poor brushing and inadequate cleaning of dentures allow the accumulation of plaque and food debris, increasing the risk of infections in the mouth [32]. The interrelated nature of the oral health problems identified in this study may be consistent with research showing that the microbial environment associated with poor hygiene affects not only teeth and gums but also the entire oral cavity including the tongue, palate, and inner surface of cheeks and lips [33,34]. These areas may be affected by conditions such as ulcers, fungal infections such as candidiasis, or mucosal irritation [32,35]. Wong et al [36] confirmed that poor oral hygiene among residents was probably due to inadequate brushing or cleaning of teeth and dentures, which increases the prevalence of dental problems, periodontal diseases, and other severe complications.

Age-related conditions, such as functional impairment, frailty, and chronic illnesses, reduce residents' ability to self-care and increase their reliance on caregivers for their daily care [3,5,37]. Such dependence on others for personal care was associated with a higher risk of chewing difficulties, poor condition of the teeth, and a higher need for dental referral. The increased retention of natural teeth, along with the growing need for care and support in old age, may explain why significantly more dental problems were observed in adults aged 76 - 85 years than in those a decade younger [29,38]. In addition, extensive to total dependence on personal hygiene was associated with poor oral hygiene and activation of the CAP oral hygiene, consistent with previous research on oral health among older institutionalized residents, identifying high care dependency as a predisposing factor for poor oral hygiene [36]. The association was particularly true for men compared with women and aligns with broader research indicating that women tend to have better oral health behaviors [39]. However, oral care is often neglected when assistance is needed for activities of daily living [29]. Research has shown that caregivers often lack knowledge of oral health, give lower priority to oral care over other care tasks, and consider the mouth as an intimate area of the body [15]. In addition, residents' resistance to care was often associated with oral health problems, as confirmed by this study [15]. Residents may also lack awareness of oral health problems due to health conditions such as cognitive impairment, sensory losses (eg, impaired vision or reduced sense of taste and smell), or the misconception that poor oral health is part of aging [3,37,40,41]. This may explain the association between severe vision impairment in adequate light and more tongue problems.

A higher need for dental referral was found when residents struggled to express themselves or understand others, which is consistent with previous research that suggests that nonverbal individuals may be more vulnerable to oral health problems [42]. Research suggests that older adults often visit the dentist only in the event of certain symptoms, such as pain, rather than regularly for preventive care [36]. As they are unable to communicate discomfort or pain, it can be challenging for caregivers to address these problems on time [42]. This highlights the need for targeted training and support for caregivers to address these challenges and ensure better oral health outcomes for residents [43]. Family support was identified as beneficial for oral health outcomes. Family members can assist with daily oral hygiene, or can ensure or demand consistent care, as well as bridge gaps in communication or resistance [44].

This study showed that the presence of teeth plays an important role in residents' ability to chew all types of food. The scoping review by Lahoud et al [8] demonstrated that tooth loss and musculoskeletal conditions, including frailty and sarcopenia—age-related loss of skeletal muscle mass and function, are the most common factors that negatively affect chewing function. Chewing difficulties are associated with discomfort or pain in the mouth and can negatively affect food selection and nutritional intake, increasing the risk of malnutrition [8]. This study confirmed that the type of food consumed may also contribute to chewing difficulties. Diets that primarily consist of softer foods can lead to a decline in chewing function due to reduced use of the masticatory muscles and teeth. Furthermore, insufficient intake of essential nutrients, such as folic acid and vitamin A, may exacerbate oral health problems, including chewing difficulties [8,45]. Consistent with the systematic review by Fan et al [46], this study identified discomfort or pain in the mouth, dry mouth, and periodontal diseases as contributing to chewing problems, underscoring the importance of maintaining good oral health to support proper chewing. Interestingly, individuals with heart failure experienced fewer chewing difficulties, which could result from lifestyle changes, including dietary adjustments or more frequent medical consultations. However, further research is needed to clarify these associations and explore underlying mechanisms.

Significant associations were found between chewing function, discomfort or pain in the mouth, and dry mouth. Chewing stimulates the salivary glands, enhancing the production of saliva, which facilitates the swallowing of food [47,48]. Lack of saliva can lead to dry mouth, which can cause discomfort or

XSL•FO
RenderX

pain in the mouth and make it difficult to chew food, often necessitating dietary adjustments [47]. The association between the presence of teeth and residents experiencing less dry mouth is consistent with research showing that individuals with natural teeth generally retain a more effective chewing function than those with complete dentures, as dentures alone do not provide the same mechanical stimulation to the salivary glands [49]. Saliva also contributes to the protection of oral mucosa from pathogens [48]. The latter could explain the association between dry mouth and poor condition of the tongue and palate and the inner surface of cheeks and lips. In addition, it is also important to note that in the earlier version of the OHS, the ohr-interRAI, the item Tongue received a negative (poor) score when a dry tongue was observed [17]. Since the adjustment of the item was made in 2023 and the assessments were collected between 2020 and 2024, it is possible that caregivers took this into account when assessing the mouth [18].

Dry mouth is a common side effect of medication, particularly among older adults, who often take multiple medications due to a range of chronic systemic diseases [50]. This may explain the association between health instability and dry mouth, in addition to dehydration—one of the items in the CHESS Scale indicating health instability and a known cause of dry mouth [51]. Reflux and COPD were also associated with dry mouth. Reflux may lead to dry mouth due to acid exposure, which irritates salivary glands, reducing their function [52]. Residents with COPD may experience dry mouth as a side effect of inhaled medications or due to chronic mouth breathing, which exacerbates dryness [53]. However, further research is necessary to clarify the underlying mechanisms. Interestingly, this study suggested that dry mouth is linked to better cognitive function, which may be explained by medications commonly prescribed for dementia that improve cognition while also causing dry mouth as a side effect.

This study yielded significant associations between depressive symptoms and poor conditions of the palate and inner surface of cheeks and lips, as well as an increased need for dental referrals. Antidepressants, commonly prescribed for the management of depression and other mental health problems, are known to increase the likelihood of dry mouth, increasing the prevalence of oral health problems, including mouth sores and cracked lips [50,54]. In addition, the associations may be explained by research indicating that individuals with depressive symptoms often experience a lack of motivation and energy, which may lead to neglect of oral hygiene practices, contributing to poor oral health outcomes [55,56]. Several studies confirm the link between depressive symptoms and poor oral health outcomes among older adults [57-59]. Although depression is not a normal part of aging, older adults—particularly nursing home residents compared with those living independently—are more likely to experience loneliness and social isolation, poor physical health, and restrictions in daily activities, which can increase the risk of depression [60]. Given the high prevalence of depressive symptoms and oral health problems in nursing home residents, timely identification and management of both conditions is important due to their bidirectional relationship [1,60,61].

Despite research showing evidence on the relationship between periodontitis and systemic diseases such as cardiovascular disease and diabetes, these diseases were not significantly linked to gum problems in this study [3]. This discrepancy may be due to methodological factors, such as the sample size, the participants included, or the assessments by the caregivers. Diabetes was, however, associated with tongue problems. Diabetes is common in nursing home residents, with reported prevalence rates ranging from 20.0% to 34.0%, aligning with the prevalence of 21.6% (506/2344) in this study [62]. This condition is well known for its association with oral health problems, including an increased risk of developing oral infections and alterations to the oral mucosa, such as a coated and fissured tongue or aphthous stomatitis [63].

To better understand the results of this study, some key considerations and limitations should be acknowledged. The generalizability of the results to the wider population is limited. All participants resided in nursing homes where a specific software system was used, integrating the OHS into the interRAI instrument. While using the interRAI instrument became mandatory in Flemish nursing homes since June 2023, this was not the case in the Netherlands. The Dutch participating nursing homes, all part of the Omring group, specifically requested to use the interRAI LTCF, including the OHS [64]. It is therefore possible that these homes were particularly motivated to work on improving residents' oral care and oral health, which may have impacted the results, in addition to possible differences in context and health practices between Flanders and the Netherlands. Schoebrechts et al [19] compared the oral health status of older adults living in Flemish and Dutch nursing homes assessed by nondental caregivers using the OHS and found more oral health problems among Flemish than Dutch residents. In addition, they reported some differences in general health status, including a higher level of dependence on others for activities of daily living and personal hygiene, as well as fewer depressive symptoms among Flemish residents compared with their Dutch counterparts [19]. Comparing associations between oral health and general health outcomes among nursing home residents in Flanders and the Netherlands was beyond the scope of this study but could be a subject for future research.

The quality and accuracy of the interRAI assessments also depend on the knowledge and qualifications of the caregivers completing them. Previous research has shown that nondental caregivers often underestimate oral health problems in nursing home residents compared with dental professionals [17]. This underestimation may be due to lack of knowledge, inadequate training, limited experience, or discomfort in addressing oral health [15,17]. In addition, underestimation of oral health problems may be attributed to some assessment criteria. For instance, oral and denture hygiene are assessed based on the extent of plaque accumulation, with the presence of plaque covering more than one-third of a tooth or denture surface. It should be noted that this may be a crude measure for the assessors. Another reason may be due to limitations of the observational assessment, particularly when residents have communication difficulties, cognitive impairment, or resist care.

In this study, the oral health assessments using the OHS were performed for the first time by various caregivers whose training

XSL•FO
RenderX

or education was unknown. Therefore, it can be assumed that oral health problems identified through the assessment may have been severe. This highlights the importance of regular oral health assessments by caregivers to identify oral health problems that may otherwise go unnoticed, as well as the need for adequate training and practice.

Some of the associations that are observed in this study can be deduced from logical connections or previous research. Other associations remain less well explained, as they appear to be nonlogical at first glance or contradict existing studies. Therefore, caution is warranted when interpreting the results. Further research may explore these associations more thoroughly, contributing to a better understanding of the relationship between oral and general health in nursing home populations. Given the mandatory use of the interRAI instrument in nursing homes in Flanders, analyses of larger datasets can also facilitate the determination of the validity of the results and how they sustain over time. However, it should be noted that a large sample size can lead to statistically significant findings that do not necessarily have clinical significance [65]. In addition, it is important to note that the interRAI instrument primarily records the occurrence of conditions. Further research is needed to identify causal factors and inform targeted interventions.

Strengths

Collecting data with the interRAI instrument can be considered as a strength of this study as it is a validated, comprehensive, holistic, and widely used tool for assessing residents' health and well-being. In addition, the integration of the OHS showed significant improvements over its current version, allowing for more accurate identification of oral health problems [66].

Conclusions

This study provided valuable insights into the associations between oral health and general health outcomes, underscoring the importance of a holistic approach that considers oral health as an integral part of general health and well-being. Recognizing these associations can enhance care planning and improve quality of care.

In practical terms, these findings support measures such as integrating regular oral health assessments into care protocols, ensuring timely access to oral health professionals, and providing caregivers with targeted training on oral care and identifying poor oral health. In addition, the results showed the importance of interdisciplinary collaboration to comprehensively address oral health. An integrated approach has the potential to improve both oral and general health outcomes, help reduce the risk of related complications, and ultimately enhance residents' quality of life.

Acknowledgments

The authors would like to thank all residents and caregivers in the participating nursing homes for their valuable contribution to this study. This work was supported by KU Leuven (C24M-20 - 063) and Research Foundation Flanders (FWO TBM T003220N).

Data Availability

The datasets generated and analyzed in this study are not publicly available due to data privacy policies; however, information on the official procedure for requesting the data can be obtained by contacting the corresponding author.

Authors' Contributions

ES contributed to conception and design, data acquisition, analysis and interpretation, and drafting and revising the manuscript. JdAM and JD contributed to conception and design, data interpretation, and substantively revising the manuscript. PAIV, HPJvH, JDL, AD, and DD contributed to conception and design and substantively revising the manuscript. All authors approved the submitted version and agreed to be personally accountable for all aspects of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Factors associated with each oral health item (reference category: acceptable condition) in the bivariate analyses.
[DOCX File, 46 KB - aging_v8i1e72308_app1.docx]

References

1. Global oral health status report: towards universal health coverage for oral health by 2030. World Health Organization. 2022. URL: https://www.who.int/publications/i/item/9789240061484 [accessed 2024-11-05]
2. Lipsky MS, Singh T, Zakeri G, Hung M. Oral health and older adults: a narrative Review. Dent J (Basel) 2024 Feb 1;12(2):30. [doi: 10.3390/dj12020030] [Medline: 38392234]
3. Hazara R. Oral health in older adults. Br J Community Nurs 2020 Aug 2;25(8):396-401. [doi: 10.12968/bjcn.2020.25.8.396] [Medline: 32757898]

XSL•FO
RenderX

4. Patel J, Wallace J, Doshi M, et al. Oral health for healthy ageing. Lancet Healthy Longev 2021 Aug;2(8):e521-e527. [doi: 10.1016/S2666-7568(21)00142-2] [Medline: 36098001]
5. Ástvaldsdóttir Á, Boström AM, Davidson T, et al. Oral health and dental care of older persons—a systematic map of systematic reviews. Gerodontology 2018 Dec;35(4):290-304. [doi: 10.1111/ger.12368] [Medline: 30129220]
6. Hopkins S, Gajagowni S, Qadeer Y, et al. Oral health and cardiovascular disease. Am J Med 2024 Apr;137(4):304-307. [doi: 10.1016/j.amjmed.2023.11.022] [Medline: 38141902]
7. Farsai PS. Cognitive impairment in older adults and oral health considerations: treatment and management. Clin Geriatr Med 2023 May;39(2):295-310. [doi: 10.1016/j.cger.2023.01.001] [Medline: 37045534]
8. Lahoud T, Yu AYD, King S. Masticatory dysfunction in older adults: a scoping review. J Oral Rehabil 2023 Aug;50(8):724-737. [doi: 10.1111/joor.13493] [Medline: 37183339]
9. Liu B, Dion MR, Jurasic MM, Gibson G, Jones JA. Xerostomia and salivary hypofunction in vulnerable elders: prevalence and etiology. Oral Surg Oral Med Oral Pathol Oral Radiol 2012 Jul;114(1):52-60. [doi: 10.1016/j.oooo.2011.11.014] [Medline: 22727092]
10. Doshi M. Medical and psychosocial considerations in older adults. In: BDJ Clinician's Guides: Springer International Publishing:59-78.
11. Chalmers JM, Pearson A. A systematic review of oral health assessment by nurses and carers for residents with dementia in residential care facilities. Spec Care Dentist 2005;25(5):227-233. [doi: 10.1111/j.1754-4505.2005.tb01654.x] [Medline: 16454098]
12. Rodrigues LG, Sampaio AA, da Cruz CAG, Vettore MV, Ferreira RC. A systematic review of measurement instruments for oral health assessment of older adults in long-term care facilities by nondental professionals. Gerodontology 2023 Jun;40(2):148-160. [doi: 10.1111/ger.12648] [Medline: 35908227]
13. Carpenter I, Hirdes JP. Chapter 3: using interRAI assessment systems to measure and maintain quality of long-term care. In: OECD Health Policy Studies: A Good Life in Old Age?: Monitoring and Improving Quality in Long-Term Care, 1st edition: OECD Publishing; 2013:93-139.
14. Vanneste D, Declercq A. The development of BelRAI, a web application for sharing assessment data on frail older people in home care, nursing homes, and hospitals. IGI Global 2018:457-483. [doi: 10.4018/978-1-5225-5490-5.ch025]
15. Krausch-Hofmann S, De Almeida Mello J, Declerck D, et al. The oral health-related section of the interRAI: evaluation of test content validity by expert rating and assessment of potential reasons for inaccurate assessments based on focus group discussions with caregivers. Gerodontology 2019 Dec;36(4):382-394. [doi: 10.1111/ger.12421] [Medline: 31274218]
16. Tran TD, Krausch-Hofmann S, Duyck J, et al. Association between oral health and general health indicators in older adults. Sci Rep 2018 Jun 11;8(1):8871. [doi: 10.1038/s41598-018-26789-4] [Medline: 29891862]
17. Krausch-Hofmann S, Tran TD, Janssens B, et al. Assessment of oral health in older adults by non-dental professional caregivers-development and validation of a photograph-supported oral health-related section for the interRAI suite of instruments. Clin Oral Investig 2021 Jun;25(6):3475-3486. [doi: 10.1007/s00784-020-03669-8] [Medline: 33196870]
18. Schoebrechts E, de Almeida Mello J, Vandenbulcke P, et al. International Delphi study to optimize the oral health section in interRAI. J Dent Res 2023 Jul;102(8):901-908. [doi: 10.1177/00220345231156162] [Medline: 36919900]
19. Schoebrechts E, de Almeida Mello J, Vandenbulcke PAI, et al. Oral health problems among Flemish and Dutch nursing home residents assessed by nondental caregivers using the novel oral health section for inclusion in interRAI. Innov Aging 2024;8(10):igae090. [doi: 10.1093/geroni/igae090] [Medline: 39464724]
20. Status and outcome scales. InterRAI. URL: https://interrai.org/applications/status-and-outcome-scales/ [accessed 2025-10-01]
21. Morris JN, Fries BE, Morris SA. Scaling ADLs within the MDS. J Gerontol A Biol Sci Med Sci 1999 Nov 1;54(11):M546-M553. [doi: 10.1093/gerona/54.11.M546]
22. Morris JN, Howard EP, Steel K, et al. Updating the Cognitive Performance Scale. J Geriatr Psychiatry Neurol 2016 Jan;29(1):47-55. [doi: 10.1177/0891988715598231]
23. Burrows AB, Morris JN, Simon SE, Hirdes JP, Phillips C. Development of a minimum data set-based depression rating scale for use in nursing homes. Age Ageing 2000 Mar 1;29(2):165-172. [doi: 10.1093/ageing/29.2.165]
24. Hirdes JP, Frijters DH, Teare GF. The MDS-CHESS scale: a new measure to predict mortality in institutionalized older people. J Am Geriatr Soc 2003 Jan;51(1):96-100. [doi: 10.1034/j.1601-5215.2002.51017.x] [Medline: 12534853]
25. Iduye S, Risling T, McKibbon S, Iduye D. Optimizing the InterRAI assessment tool in care planning processes for long-term residents: a scoping review. Clin Nurs Res 2022 Jan;31(1):5-19. [doi: 10.1177/10547738211020373] [Medline: 34056955]
26. Gerritsen DL, Steverink N, Frijters DHM, Hirdes JP, Ooms ME, Ribbe MW. A revised index for social engagement for long-term care. J Gerontol Nurs 2008 Apr;34(4):40-48. [doi: 10.3928/00989134-20080401-04]
27. Little RJA, Rubin DB. Statistical Analysis With Missing Data, 3rd edition: Wiley; 2020.
28. Vanneste D, De Almeida Mello J, Macq J, Van Audenhove C, Declercq A. Incomplete assessments: towards a better understanding of causes and solutions. The case of the interRAI home care instrument in Belgium. PLoS One 2015;10(4):e0123760. [doi: 10.1371/journal.pone.0123760] [Medline: 25875281]
29. Müller F, Shimazaki Y, Kahabuka F, Schimmel M. Oral health for an ageing population: the importance of a natural dentition in older adults. Int Dent J 2017 Sep;67(Suppl 2):7-13. [doi: 10.1111/idj.12329] [Medline: 29023743]

XSL•FO
RenderX

30. Janssens L, Petrauskiene E, Tsakos G, Janssens B. Clinical and subjective oral health status of care home residents in Europe: a systematic review. J Am Med Dir Assoc 2023 Jul;24(7):1013-1019. [doi: 10.1016/j.jamda.2023.03.021] [Medline: 37105236]
31. Oral health. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/oral-health [accessed 2024-11-06]
32. Coll PP, Lindsay A, Meng J, et al. The prevention of infections in older adults: oral health. J Am Geriatr Soc 2020 Feb;68(2):411-416. [doi: 10.1111/jgs.16154] [Medline: 31479533]
33. Petersen PE, Yamamoto T. Improving the oral health of older people: the approach of the WHO Global Oral Health Programme. Community Dent Oral Epidemiol 2005 Apr;33(2):81-92. [doi: 10.1111/j.1600-0528.2004.00219.x] [Medline: 15725170]
34. Villoria GEM, Fischer RG, Tinoco EMB, Meyle J, Loos BG. Periodontal disease: a systemic condition. Periodontol 2000;96(1):7-19. [doi: 10.1111/prd.12616]
35. Gendreau L, Loewy ZG. Epidemiology and etiology of denture stomatitis. J Prosthodont 2011 Jun;20(4):251-260. [doi: 10.1111/j.1532-849X.2011.00698.x] [Medline: 21463383]
36. Wong FMF, Ng YTY, Leung WK. Oral health and its associated factors among older institutionalized residents—a systematic review. Int J Environ Res Public Health 2019 Oct 26;16(21):4132. [doi: 10.3390/ijerph16214132] [Medline: 31717812]
37. Niesten D, van Mourik K, van der Sanden W. The impact of frailty on oral care behavior of older people: a qualitative study. BMC Oral Health 2013 Nov 1;13:61. [doi: 10.1186/1472-6831-13-61] [Medline: 24175989]
38. Khan HTA, Addo KM, Findlay H. Public health challenges and responses to the growing ageing populations. Public Health Chall 2024 Sep;3(3):e213. [doi: 10.1002/puh2.213] [Medline: 40496520]
39. Lipsky MS, Su S, Crespo CJ, Hung M. Men and oral health: a review of sex and gender differences. Am J Mens Health 2021;15(3):15579883211016361. [doi: 10.1177/15579883211016361] [Medline: 33993787]
40. Tenani CF, De Checchi MHR, Bado FMR, Ju X, Jamieson L, Mialhe FL. Influence of oral health literacy on dissatisfaction with oral health among older people. Gerodontology 2020 Mar;37(1):46-52. [doi: 10.1111/ger.12443] [Medline: 31746043]
41. Vandenbulcke PAI, de Almeida Mello J, Cornette V, et al. Nursing home residents' perspectives on oral health: an in-depth interview study. Int J Nurs Stud Adv 2024 Jun;6:100198. [doi: 10.1016/j.ijnsa.2024.100198] [Medline: 38746809]
42. Lauritano D, Moreo G, Della Vella F, et al. Oral health status and need for oral care in an aging population: a systematic review. Int J Environ Res Public Health 2019;16(22):4558. [doi: 10.3390/ijerph16224558]
43. Palmers E, Janssens L, Phlypo I, et al. Perceptions on oral care needs, barriers, and practices among managers and staff in long-term care settings for older people in Flanders, Belgium: a cross-sectional survey. Innov Aging 2022;6(5):igac046. [doi: 10.1093/geroni/igac046]
44. Hoben M, Clarke A, Huynh KT, et al. Barriers and facilitators in providing oral care to nursing home residents, from the perspective of care aides: a systematic review and meta-analysis. Int J Nurs Stud 2017 Aug;73:34-51. [doi: 10.1016/j.ijnurstu.2017.05.003] [Medline: 28531550]
45. Schuster AJ, Pastorino DA, Marcello-Machado RM, Faot F. Influence of age and time since edentulism on masticatory function and quality of life in implant-retained mandibular overdenture wearers: 1-year results from a paired clinical study. Int J Oral Maxillofac Implants 2019;34(6):1466-1474. [doi: 10.11607/jomi.7525] [Medline: 31711087]
46. Fan Y, Shu X, Leung KCM, Lo ECM. Association between masticatory performance and oral conditions in adults: a systematic review and meta-analysis. J Dent 2023 Feb;129:104395. [doi: 10.1016/j.jdent.2022.104395] [Medline: 36563840]
47. Chan AKY, Chu CH, Ogawa H, Lai EHH. Improving oral health of older adults for healthy ageing. J Dent Sci 2024 Jan;19(1):1-7. [doi: 10.1016/j.jds.2023.10.018] [Medline: 38303786]
48. Ogawa H, McKenna G, Kettratad-Pruksapong M. Prevention of oral functional decline. Int Dent J 2022 Aug;72(4S):S21-S26. [doi: 10.1016/j.identj.2022.05.008] [Medline: 36031322]
49. Walls AWG, Steele JG. The relationship between oral health and nutrition in older people. Mech Ageing Dev 2004 Dec;125(12):853-857. [doi: 10.1016/j.mad.2004.07.011] [Medline: 15563930]
50. Tan ECK, Lexomboon D, Sandborgh-Englund G, Haasum Y, Johnell K. Medications that cause dry mouth as an adverse effect in older people: a systematic review and metaanalysis. J Am Geriatr Soc 2018 Jan;66(1):76-84. [doi: 10.1111/jgs.15151] [Medline: 29071719]
51. 5 common questions about dry mouth. NIH MedlinePlus Magazine. URL: https://magazine.medlineplus.gov/article/5-common-questions-about-dry-mouth?utm_source=chatgpt.com [accessed 2025-01-31]
52. Ranjitkar S, Smales RJ, Kaidonis JA. Oral manifestations of gastroesophageal reflux disease. J Gastroenterol Hepatol 2012 Jan;27(1):21-27. [doi: 10.1111/j.1440-1746.2011.06945.x] [Medline: 22004279]
53. Devlin J. Patients with chronic obstructive pulmonary disease: management considerations for the dental team. Br Dent J 2014 Sep;217(5):235-237. [doi: 10.1038/sj.bdj.2014.756] [Medline: 25213520]
54. Global strategy and action plan on oral health 2023–2030. World Health Organization. 2024. URL: https://iris.who.int/bitstream/handle/10665/376623/9789240090538-eng.pdf?sequence=1 [accessed 2025-10-01]
55. Slack-Smith L, Hearn L, Scrine C, Durey A. Barriers and enablers for oral health care for people affected by mental health disorders. Aust Dent J 2017 Mar;62(1):6-13. [doi: 10.1111/adj.12429] [Medline: 27164018]

XSL•FO
RenderX

56. Kisely S, Sawyer E, Siskind D, Lalloo R. The oral health of people with anxiety and depressive disorders—a systematic review and meta-analysis. J Affect Disord 2016 Aug;200:119-132. [doi: 10.1016/j.jad.2016.04.040] [Medline: 27130961]
57. Hybels CF, Bennett JM, Landerman LR, Liang J, Plassman BL, Wu B. Trajectories of depressive symptoms and oral health outcomes in a community sample of older adults. Int J Geriatr Psychiatry 2016 Jan;31(1):83-91. [doi: 10.1002/gps.4292] [Medline: 25962827]
58. Friedlander AH, Friedlander IK, Gallas M, Velasco E. Late-life depression: its oral health significance. Int Dent J 2003 Feb;53(1):41-50. [doi: 10.1111/j.1875-595x.2003.tb00655.x] [Medline: 12653339]
59. Skośkiewicz-Malinowska K, Malicka B, Ziętek M, Kaczmarek U. Oral health condition and occurrence of depression in the elderly. Medicine (Baltimore) 2018 Oct;97(41):e12490. [doi: 10.1097/MD.0000000000012490] [Medline: 30313038]
60. Chau R, Kissane DW, Davison TE. Risk factors for depression in long-term care: a systematic review. Clin Gerontol 2019;42(3):224-237. [doi: 10.1080/07317115.2018.1490371] [Medline: 29920178]
61. Cross L. Is depression influenced by oral health? Evid Based Dent 2025 Mar;26(1):57-58. [doi: 10.1038/s41432-025-01118-7] [Medline: 39948157]
62. Idrees T, Castro-Revoredo IA, Migdal AL, Moreno EM, Umpierrez GE. Update on the management of diabetes in long-term care facilities. BMJ Open Diab Res Care 2022 Jul;10(4):e002705. [doi: 10.1136/bmjdrc-2021-002705]
63. Rohani B. Oral manifestations in patients with diabetes mellitus. World J Diabetes 2019 Sep 15;10(9):485-489. [doi: 10.4239/wjd.v10.i9.485]
64. BelRAI in geriatric care [Web page in Dutch]. Department of Care. URL: https://www.departementzorg.be/nl/belrai-binnen-de-ouderenzorg [accessed 2024-12-17]
65. Armstrong RA. Is there a large sample size problem? Ophthalmic Physiol Opt 2019 May;39(3):129-130. [doi: 10.1111/opo.12618] [Medline: 30994198]
66. Schoebrechts E, de Almeida Mello J, Vandenbulcke PAI, et al. Comparison of the oral health status of nursing home residents using the current and the newly developed interRAI oral health section (OHS-interRAI): a cross-sectional study. BMC Geriatr 2024 Nov 15;24(1):950. [doi: 10.1186/s12877-024-05494-0] [Medline: 39548377]

Abbreviations

ADLH: Activities of Daily Living Hierarchy
CAP: collaborative action point
CHESS: Changes in Health, End-Stage Disease, Signs, and Symptoms
COPD: chronic obstructive pulmonary disease
CPS: Cognitive Performance Scale
DRS: Depression Rating Scale
LTCF: long-term care facility
OR: odds ratio

Edited by R Bjarnadottir; submitted 07.02.25; peer-reviewed by LE Janssens, MKA van Selms; revised version received 04.07.25; accepted 28.08.25; published 17.10.25.

Please cite as:
Schoebrechts E, de Almeida Mello J, Vandenbulcke PAI, van Hout HPJ, De Lepeleire J, Declercq A, Declerck D, Duyck J
Insights Into Associations Between Oral and General Health Outcomes of Nursing Home Residents Based on InterRAI Data: Cross-Sectional Study
JMIR Aging 2025;8:e72308
URL: https://aging.jmir.org/2025/1/e72308
doi:10.2196/72308

© Emilie Schoebrechts, Johanna de Almeida Mello, Patricia Ann Ivonne Vandenbulcke, Hein Petrus Johannes van Hout, Jan De Lepeleire, Anja Declercq, Dominique Declerck, Joke Duyck. Originally published in JMIR Aging (https://aging.jmir.org), 17.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Urban Quality and Biochemical, Hematological, and Nutritional Markers in Older Adults: Cross-Sectional Geospatial Study

Carlos Mena[1*], PhD; Yony Ormazabal[1*], MSc; Nacim Molina[2], PhD; Eduardo Fuentes[2], PhD; Juan Carlos Cantillana[3], MSc; Victoria Villalobos[2], TM; Moises Sandoval[4], PhD; Ivan Palomo[2], PhD; Diego Arauna[2], PhD

[1]Longevity Center VITALIS, Faculty of Economics and Business, University of Talca, Avenida Lircay S/N, Talca, Chile, Talca, Chile

[2]Thrombosis Research Center, Department of Clinical Biochemistry and Immunohematology, Faculty of Health Sciences, Interuniversity Center of Healthy Aging (CIES), Longevity Center VITALIS, University of Talca, Avenida Lircay S/N, Talca, Chile

[3]Faculty of Administration and Economics, Universidad Tecnológica Metropolitana, Santiago, Chile

[4]Instituto de Nutrición y Tecnología de los Alimentos (INTA), University of Chile, Santiago, Chile

[*]these authors contributed equally

Corresponding Author:
Diego Arauna, PhD
Thrombosis Research Center, Department of Clinical Biochemistry and Immunohematology, Faculty of Health Sciences, Interuniversity Center of Healthy Aging (CIES), Longevity Center VITALIS, University of Talca, Avenida Lircay S/N, Talca, Chile

Abstract

Background: The urban environment is an important determinant of frailty, primarily through factors such as infrastructure that supports physical activity, availability of social and medical support, and access to nutritious food. Given the increasing aging population, understanding the link between urban quality, frailty, and metabolic health is crucial for effective public health and urban planning interventions.

Objective: This study aims to quantify the impact of distinct urban domains (built-environment characteristics, accessibility to essential services, availability of green and recreational spaces, and neighborhood socioeconomic context) on frailty status, nutritional profile, and hematological or biochemical biomarkers in community-dwelling older adults by integrating geospatial analysis.

Methods: A cohort of 251 older adults (aged older than 65 years) was studied. Frailty was assessed using the Frailty Trait Scale 5, and nutritional status was determined using the Controlling Nutritional Status score. Hematological and biochemical parameters were evaluated in a subset of 70 participants by MINDRAY automatic equipment. A spatial analysis of frailty was conducted by incorporating Geographic Information System layers that mapped the distribution of urban facilities, including fruit and vegetable shops, senior centers, pharmacies, emergency health centers, parks and squares, community centers, and exercise facilities. Statistical analyses included *t* tests, Mann-Whitney *U* test, ANOVA, and correlation analyses.

Results: The prevalence of frailty was 17.5%. Frail individuals exhibited significantly higher BMI (mean 31.5, SD 4.4 vs mean 28.5, SD 4.5 kg/m²; *P*=.0001). When comparing the upper (Q4) and lower (Q1) quartiles of urban quality, Q4 participants had higher Frailty Trait Scale 5 scores (mean 15.2, SD 7.4 vs mean 11.8, SD 6.4; *P*=.0334) and lower handgrip strength (mean 19.1, SD 4.4 vs mean 22.8, SD 7.3 kg; *P*=.006). Frail individuals resided significantly closer to emergency health centers (*P*=.0010), family health centers (*P*=.0412), and exercise facilities (*P*=.0322). In addition, bilirubin (Spearman ρ=0.33; *P*=.0049), serum iron (Spearman ρ=0.27; *P*=.0272), transferrin saturation (Spearman ρ=0.24; *P*=.0386), red blood cell count (Spearman ρ=0.26; *P*=.0303), and red blood cell distribution width (Spearman ρ=0.23; *P*=.0462) were positively correlated with urban quality. Frail participants also had higher Controlling Nutritional Status scores (*P*=.0323), which were positively correlated with urban quality (Spearman ρ=0.25; *P*=.0359).

Conclusions: Urban quality was significantly associated with hematological parameters, nutritional status, and frailty. Frail individuals in areas with better urban quality exhibit lower handgrip strength, higher frailty scores, and greater proximity to emergency rooms, community health centers, and exercise facilities. This spatial distribution may reflect higher accessibility to health care and recreational resources among frail participants. Urban planning and public health strategies should focus on creating age-friendly environments to prevent frailty and improve health outcomes.

(JMIR Aging 2025;8:e74313) doi:10.2196/74313

KEYWORDS

aging; frailty; geospatial analysis; urban quality; hematological parameters; nutritional status; biochemical parameters

XSL•FO
RenderX

Introduction

Global demographic changes have indicated an accelerated increase in the aging population, along with a rise in the prevalence of older adults [1]. This demographic transition has prompted the analysis of various key factors associated with the aging process, with frailty being one of the primary foci in geriatric medicine [2]. This multisystemic syndrome is characterized by a decline in physiological reserves and intrinsic capacity, which leads to an increased vulnerability to stressors and a higher risk of adverse outcomes such as falls, hospitalization, disability, and mortality [3,4]. The global prevalence of frailty varies between 12% and 24%, depending on the diagnostic tool used, and it is one of the key factors to prevent and reverse according to the World Health Organization recommendations [5]. The increase in the frailty prevalence among older adults poses significant challenges for health care systems, highlighting the need for comprehensive strategies to address its underlying determinants [2,3,6]. Recent studies have highlighted the significant role of the urban environment in frailty, influencing both its development and improvement, independent of individual factors [7]. The built environment, characterized by access to green spaces, health services, community resources, places for physical activity, diet, and social interactions, plays a key role in frailty [7-9]. This is reflected in various studies indicating that older adults living in areas with lower land-use diversity and more social incivilities may be at risk of developing mobility limitations. In contrast, those residing in urban areas with limited resources and higher pollution levels tend to exhibit higher mortality rates associated with frailty [10,11]. In this context, frailty and mobility limitations have been associated with the neighborhood physical characteristics, including land-use mix diversity, land-use mix access, street connectivity, walking infrastructure, esthetics, crime safety, and social incivilities [11,12]. Due to the relevance of frailty in older adults, the World Health Organization has recognized the need to create age-friendly environments to enhance healthy aging and reduce the risk of frailty [5]. In Latin America, urban inequalities are particularly pronounced, with limited access to critical services and stark disparities in urban quality [13]. Chile exemplifies this context, characterized by rapid urbanization, a pronounced rural-urban divide, and significant socioeconomic heterogeneity. While previous studies have linked environmental factors to frailty in high-income countries [7-9], evidence from Latin America remains limited, particularly concerning the interplay between urban quality, hematological or biochemical markers, and nutritional status in aging populations. Notably, no Chilean study to date has integrated geospatial analysis with clinical biomarkers to examine how specific urban features (eg, proximity to emergency centers or green spaces) correlate with frailty phenotypes. This gap is critical, as Chile's unique urban-rural dynamics and aging demographics may reveal patterns not observed in more developed settings [14]. In this study, we address this gap by combining Geographic Information System (GIS)–based urban metrics with multidimensional frailty assessments (Frailty Trait Scale 5 [FTS-5]), biochemical or hematological profiling, and nutritional status (Controlling Nutritional Status [CONUT] score) in a Chilean cohort, offering novel insights for regional public health strategies. By integrating GIS data with health and sociodemographic variables, we aim to analyze possible spatial clusters of frailty, assess the impact of urban quality, and highlight the potential of urban planning and policy interventions in mitigating frailty risks [8,9,15]. In addition to environmental factors, nutritional and hematological status is a key component associated with frailty. Urban environmental conditions have also been implicated in modulating blood-based health indicators [16,17]. Hematological abnormalities, such as variations in red blood cell (RBC) count and iron status, are linked to frailty, potentially through their effects on muscle function, oxygen transport, and fatigue. Moreover, several biochemical markers reflecting hepatic and renal function have been associated with frailty in older adults [18]. Understanding how urban quality influences these clinical parameters may provide novel perspectives for developing targeted frailty prevention strategies [17]. With this background, the aim of this study was to quantify the impact of distinct urban domains (built-environment characteristics, accessibility to essential services, availability of green and recreational spaces, and neighborhood socioeconomic context) on frailty status, nutritional profile, and hematological or biochemical biomarkers in community-dwelling older adults by integrating geospatial analysis.

Methods

Participants and Study Design

A representative sample of 251 older adults (aged 65 years and older; both men and women) from the older adult population of Talca, Chile (Maule Region), registered with Family Health Centers and community organizations, participated in this cross-sectional study. The inclusion criterion was adults aged 65 years and older. Participants with self-reported or medically documented cancer, Parkinson disease, or vascular accidents were excluded, as well as older individuals unable to walk or speak, and those currently on statin therapy [19]. The sample size calculation was based on an estimated frailty prevalence of 24.6% among older adults, with a 95% confidence level, 80% statistical power, and an anticipated 20% attrition rate. The gender distribution of the sample reflected that of the population aged 65 years and older, according to data from the National Socioeconomic Characterization Survey.

Frailty Diagnosis

Frailty was measured using the FTS-5, which assesses frailty across 5 key domains: nutrition, physical activity, nervous system, strength, and gait speed. Nutrition was evaluated using BMI [20]. Physical activity was assessed with the Physical Activity Scale for Elderly. Nervous system function was evaluated using the Progressive Romberg Test. Strength was determined by handgrip strength in the dominant hand. Gait speed was measured by time taken to complete a 3-meter walk at a usual pace. Each domain is scored on a scale from 0 to 10 points. Frailty was determined by summing the scores across all 5 domains, resulting in a total range of 0-50 points. A score above 25 was considered indicative of frailty. The FTS-5 Scale has demonstrated high diagnostic accuracy (AUC: 0.81 - 0.85 for frailty incidence) in urban-dwelling older adults, with strong

XSL•FO
RenderX

psychometric properties relative to Fried's phenotype (sensitivity 0.608; specificity 0.823) and the FRAIL Scale (sensitivity 0.642, specificity 0.814). Unlike Fried's criteria, which rely solely on physical characteristics, the FTS-5 incorporates neurophysical and functional dimensions, enhancing its sensitivity in heterogeneous urban populations.

Nutritional Status

Nutritional status was assessed using the CONUT score, a validated tool based on 3 biochemical parameters: serum albumin, total lymphocyte count, and total cholesterol levels [21]. Each parameter is assigned a score based on predefined thresholds, and the sum of these scores provides the overall CONUT score. Higher scores indicate a greater degree of malnutrition, with classifications ranging from normal (0 - 1) to severe malnutrition (≥9).

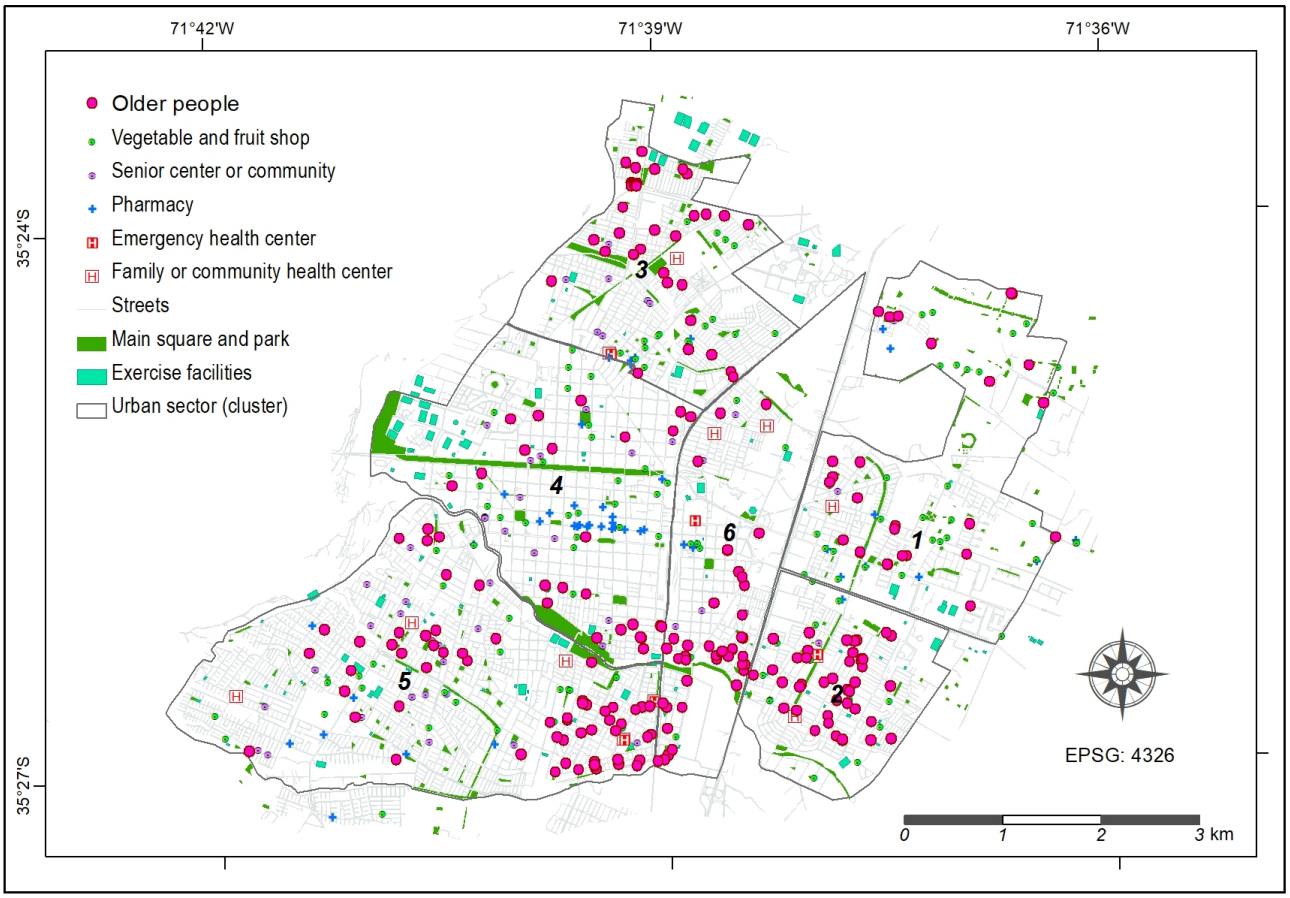
Hematological and Biochemical Profile

Seventy plasma and serum samples were randomly obtained from cohort participants. Blood samples were collected using ethylenediaminetetraacetic acid (4.5 mL) and without anticoagulant (5 mL) by phlebotomy. The ethylenediaminetetraacetic acid blood sample was used to determine hematological profile (BC-3600 MINDRAY) using automated equipment. The blood sample without anticoagulant was centrifuged at 240 g for 15 minutes to obtain serum, which was used to determine the biochemical profile (BC-300 MINDRAY) using automated equipment. The sample size was calculated using the GRANMO calculator with a 95% confidence level, 80% statistical power, and an expected loss rate of 20%.

Geospatial Database

A georeferenced database was created using GIS technology to represent the physical environment of the study area and the locations of the older adults participating in the study (Figure 1). Each participant was geographically located as a point object based on the residence data provided during the medical evaluation. All data were organized into a point feature layer, accompanied by a corresponding thematic table. The database included relevant factors related to urban physical infrastructure within the city, such as fruit and vegetable shops, senior centers or communities, pharmacies, emergency health centers, main squares and parks, family or community health centers, and exercise facilities. Each component was represented as a GIS layer, either in point or polygon form, at the neighborhood level, using geographical information from OSM (OpenStreetMap), Google Maps, and IDE Chile (Geospatial Infrastructure of Chile).

Figure 1. Locations of each older adult in the study within the city of Talca, alongside the available urban physical facilities.



The city of Talca (Maule Region, Chile) was further divided into 6 main clusters based on sociodemographic characteristics: cluster 1 represents the northeastern sector with a high socioeconomic status (SES), cluster 2 covers the southeastern sector with a low SES, cluster 3 includes the northern sector with a lower-middle socioeconomic class, cluster 4 corresponds to the historic center, cluster 5 encompasses the southern sector with a medium-high socioeconomic level, and cluster 6 refers

XSL•FO
RenderX

to the industrial center area. These clusters were defined using census data and municipal planning criteria, considering variables such as household income, education level, housing density, and land use. Classification was based on expert-informed manual grouping rather than statistical clustering algorithms.

Urban Quality Level

Each GIS layer was analyzed using the Euclidean distance method to assess the proximity of every location within the city to various urban facilities. The resulting distance layers were then classified into 3 zones: close, medium, and distant. Each zone was assigned a score on a scale from 1 to 3, with the closest proximity receiving a score of 3 and the most distant areas a score of 1. The distance ranges and corresponding scores for each urban facility were defined based on the local context, as shown in Table 1. These parameters had been previously applied in the same city [14] and are considered representative of 3 distinct levels of access efforts. The general criterion was that closer proximity to a facility indicated higher urban quality. A raster calculator was used to aggregate all layers into a summary index, with higher numerical values reflecting better urban quality in the corresponding area. The facilities included in this index were vegetable and fruit shops, senior centers or communities, pharmacies, emergency health centers, main squares and parks, family or community health centers, and exercise facilities, as detailed in Table 1. The summary index was computed as a continuous raster surface (50-m resolution) covering the entire urban area. Values were extracted at each participant's georeferenced location using a spatial join. While the index was later analyzed in relation to spatial clusters, its computation was conducted across the full urban extent rather than per cluster. All data management, processing, and analysis were performed using ArcGIS software (version 10; ESRI).

Table . Distance ranges and corresponding values assigned to each urban physical facility.

Facility	Distance ranges, m		
Vegetables and fruits shops	<300	300 - 600	>600
Senior centers or communities	<500	500 - 1000	>1000
Pharmacies	<500	500 - 1000	>1000
Emergency health centers	<1000	1000 - 2000	>2000
Main squares and parks	<200	200 - 400	>400
Family or community health centers	<700	700 - 1400	>1400
Exercise facilities (stadiums and green areas with urban assets designed for physical activity)	<400	400 - 600	>600
	Values		
	3	2	1

Statistical Analysis

Statistical analyses were conducted using GraphPad Prism 9. Continuous variables were expressed as mean (SD) or median with 95% CI. Categorical variables were expressed as percentages with 95% CI. Group differences were assessed using the Chi-square test with Yate's correction for proportions. For differences in means or medians, the *t* test, Mann-Whitney *U* test, or ANOVA was applied, depending on the context. Correlation analysis was performed using Spearman correlation coefficient (ρ). Statistical significance was considered at *P* values of <.0500.

Ethical Considerations

The institutional review board approval for this study was obtained from the Comité de Ética Científica of Universidad de Talca (reference number 06-2021). All procedures adhered to the ethical standards of the Comité de Ética Científica and the World Medical Association's Declaration of Helsinki. All participants provided written informed consent. All data were deidentified; stored on encrypted, password-protected institutional servers with access restricted to authorized personnel; and reported only in aggregate. Geospatial data were coded to administrative units to minimize reidentification. Participants participated voluntarily and received no compensation.

Results

Sociodemographic Characteristics and Spatial Cluster Distribution by Frailty Status

Table 2 shows the sociodemographic characteristics of the study cohort, as well as the geospatial distribution of nonfrail and frail individuals across the spatial clusters (1-6). The cohort consists of 17.5% (44/251) frail older adults and 82.5% (207/251) nonfrail older adults. No significant differences were observed in the prevalence of women between groups (*P*=.642). The frail group had a higher mean age (mean 75.6, SD 7.4 years) than the nonfrail group (mean 73.8, SD 5.2 years); however, this difference was not statistically significant (*P*=.0583). Conversely, the frail group exhibited a significantly higher mean BMI (mean 31.5, SD 4.4 kg/m²) than the nonfrail group (mean 28.5, SD 4.5 kg/m²) (*P*=.0001). Regarding geospatial distribution, no significant differences were observed between the groups.

XSL•FO
RenderX

Table . Sociodemographic description and geospatial distribution of the studied sample of older people according to frailty status.

	Nonfrail (n=207)	Frail (n=44)	*P* value
Variables			
Female % (95% CI)	73.9 (67.5 - 79.4)	77.2 (63.0 - 87.1)	.6424
Age (years), mean (SD)	73.8 (5.2)	75.6 (7.4)	.0583
BMI, mean (SD)	28.5 (4.5)	31.5 (4.4)	.0001
Spatial cluster			
Cluster 1 (95% CI)	12.6 (8.7 - 17.7)	6.8 (2.3 - 18.2)	.2792
Cluster 2 (95% CI)	17.4 (12.8 - 23.1)	18.2 (9.5 - 31.9)	.5918
Cluster 3 (95% CI)	17.9 (13.2 - 23.6)	22.7 (12.8 - 36.9)	.4536
Cluster 4 (95% CI)	11.6 (7.9 - 16.7)	6.8 (2.3 - 18.2)	.3531
Cluster 5 (95% CI)	23.1 (17.9 - 29.4)	27.3 (16.3 - 41.8)	.5640
Cluster 6 (95% CI)	17.4 (12.8 - 23.1)	18.2 (9.5 - 31.9)	.5918

Frailty Characterization According to Quartile Distribution of Summary Index

Figure 2 illustrates the different levels of urban quality based on a cumulative assessment (summatory index) of physical environmental factors considered in this study. These factors primarily encompass essential urban services and infrastructure relevant to the older adult cohort under investigation. Urban quality levels are closely linked to the accessibility of various facilities from each location within the city. In addition, Figure 2 displays the spatial distribution of individual participants, with each one identified by frailty status.

Figure 2. Frailty status of older adults across the summary index within the geospatial clusters defined for the city of Talca.

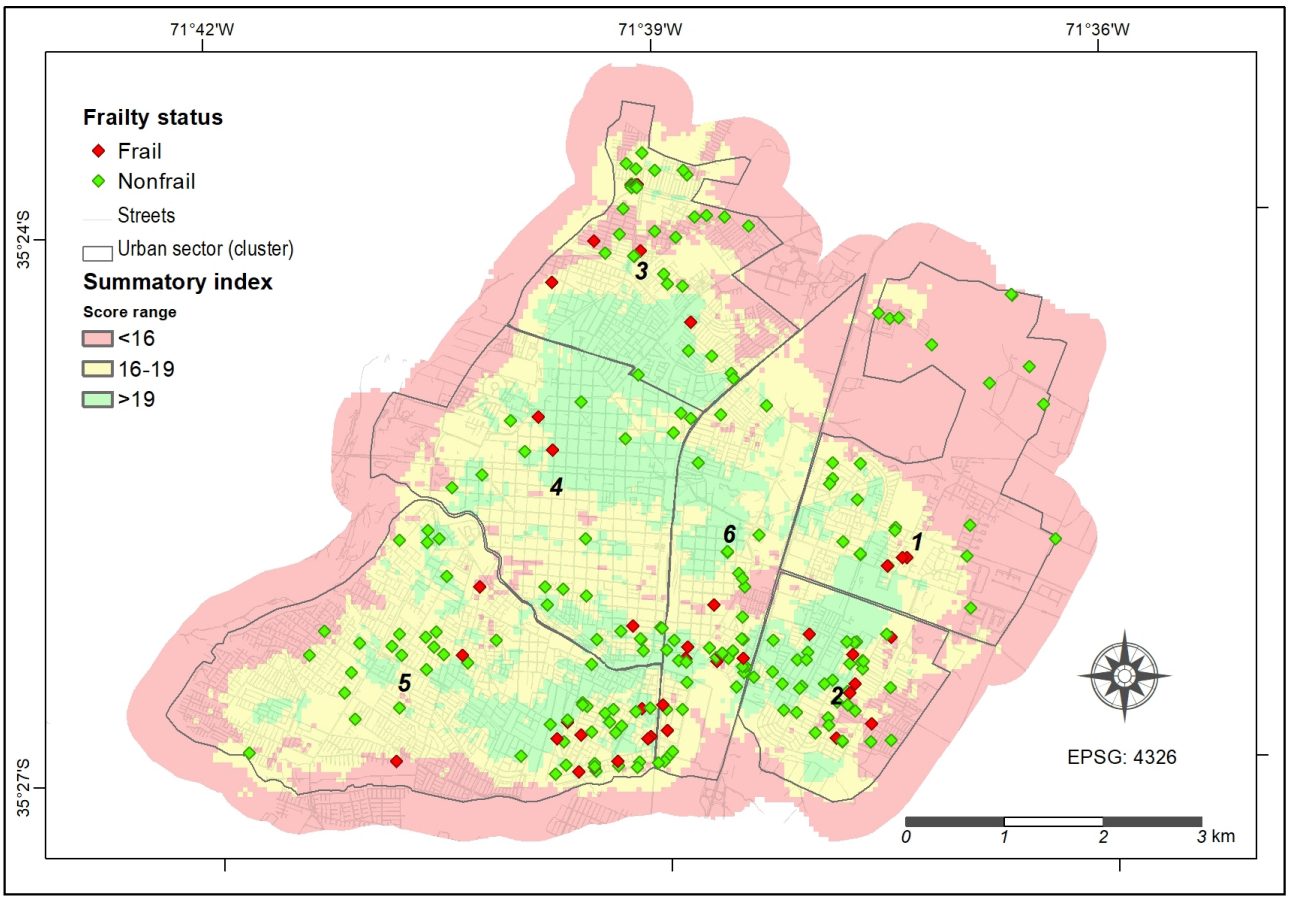


Table 3 shows mean values for frailty prevalence, frailty score, handgrip strength, gait speed, and BMI, grouped by quartiles of the urban quality summary index. While spatial clusters represent areas with homogeneous socioeconomic and urban characteristics (eg, cluster 1 corresponds to a high SES area in the northeast), the quartiles (Q1-Q4) classify areas based solely on their objective accessibility to services, regardless of SES. As shown in Figure 2, for instance, cluster 5 (classified as

XSL•FO
RenderX

medium-high SES) includes both Q1 (poorest access) and Q4 (best access) zones, highlighting that even socioeconomically advantaged areas may contain localized disparities in service provision. As previously described, lower values on the summary index indicate poorer urban quality, while higher values reflect better urban quality. Q1 represents the quartile with the lowest index values (poorest urban quality), and Q4 represents the highest values (best urban quality). No statistically significant differences were observed across quartiles for the analyzed variables concerning frailty prevalence (*P*>.0500). However, when comparing Q4 with Q1, participants in Q4 exhibited a significantly higher FTS-5 score (mean 15.2, SD 7.4 vs mean 11.8, SD 6.4; *P*=.0334) and a significantly lower handgrip strength (mean 19.1, SD 4.4 vs mean 22.8, SD 7.3; *P*=.0059). No significant differences were found between Q4 and Q1 for frailty prevalence, gait speed, or BMI (*P*>.0500).

Table . Frailty prevalence, handgrip strength, gait speed, and BMI according to quartile distribution of summary index.

Frail status	Q1 (12-17)	Q2 (>17 - 19)	Q3 (>19 - 20)	Q4 (>20 - 22)	*P* value
Frail % (95% CI)	15.3 (8.7 - 25.3)	17.8 (11.6 - 26.4)	17.3 (9.4 - 29.7)	23.1 (11.0 - 42.0)	.8462
FTS-5 Scale score, mean (SD)	11.8 (6.4)	13.6 (7.9)	13.7 (8.3)	15.2 (7.4)[a]	.2092
Handgrip strength (kg), mean (SD)	22.8 (7.3)	21.3 (7.3)	22.4 (7.2)	19.1 (4.4)[a]	.2620
Gait speed (m/s), mean (SD)	1.0 (0.3)	0.9 (0.2)	0.9 (0.2)	0.9 (0.2)	.6593
BMI (kg/m^2), mean (SD)	29.3 (4.3)	29.4 (4.1)	29.1 (4.4)	28.5 (4.2)	.7723

[a]*P* value<.0500 respect to Q1.

[b]FTS-5: Frailty Trait Scale 5.

Relationship Between Frailty and Distances to Urban Facilities and Urban Quality

Figure 3 shows a comparison of average distance to urban facilities by frailty status, as assessed using the FTS-5 clinical tool. No significant differences were observed in the distance between frail and nonfrail groups for key urban areas such as vegetable and fruit shops (Figure 3A), senior centers or communities (Figure 3B), pharmacies (Figure 3C), and main squares and parks (Figure 3E) (*P*>.0500). However, Figure 3A suggests a trend where frail individuals tend to live further from vegetable and fruit shops. In contrast, Figure 3D shows that the frail group is significantly closer to emergency health centers (*P*=.0010). Similarly, the frail group was significantly closer to family or community health centers (Figure 3F) (*P*=.0412) and stadiums and exercise facilities (Figure 3G) (*P*=.0322).

XSL•FO
RenderX

Figure 3. Comparison of mean distance for different frailty status diagnosed by Frailty Trait Scale 5 Scale with respect to relevant urban facilities: (A) vegetables and fruits shops, (B) senior centers or communities, (C) pharmacies, (D) emergency health centers, (E) main squares and parks, (F) family or community health centers, and (G) exercise facilities. The data presented are mean (SEM). Statistical analysis was performed using the *t* test or Mann-Whitney *U* test, as appropriate. *P<.0500. **P<.0001. NS: not significant.

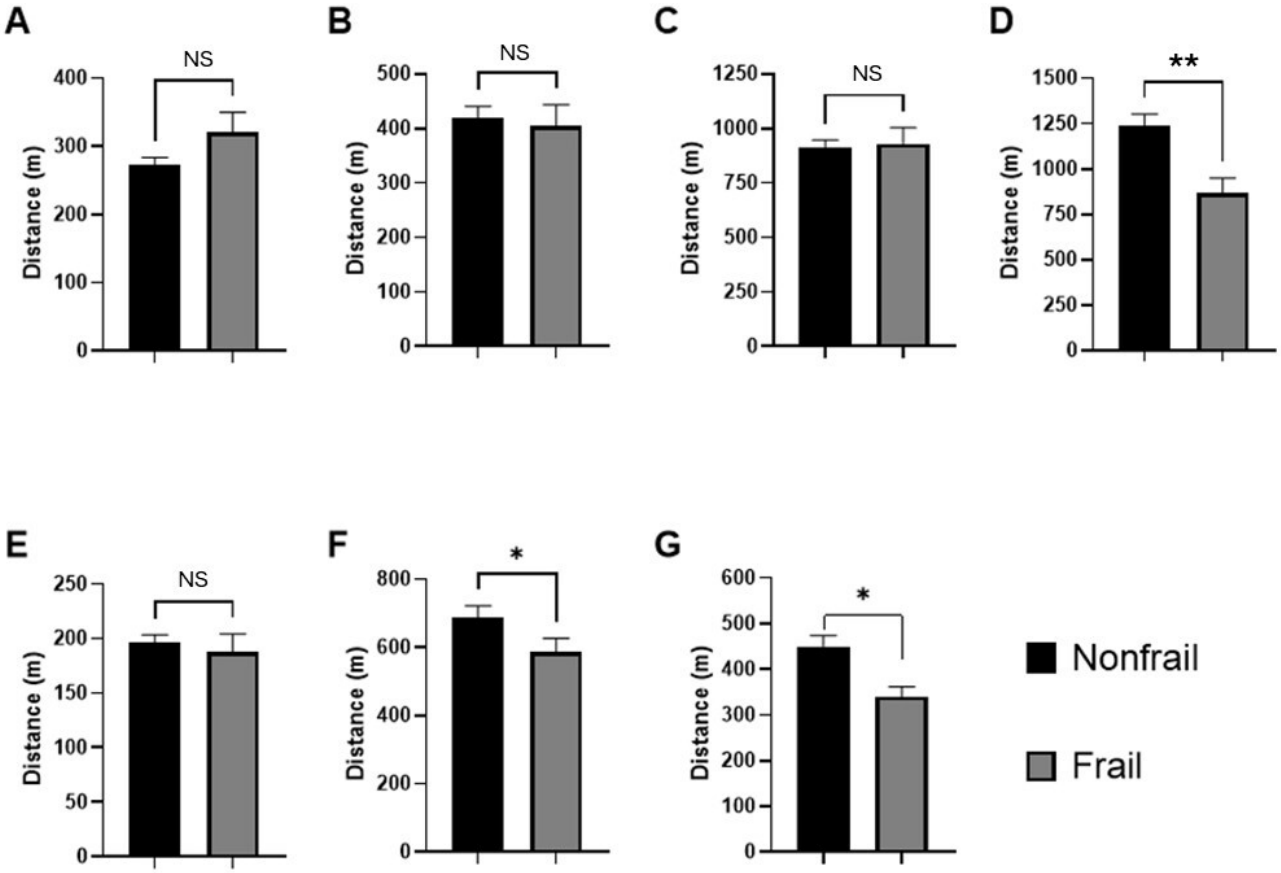
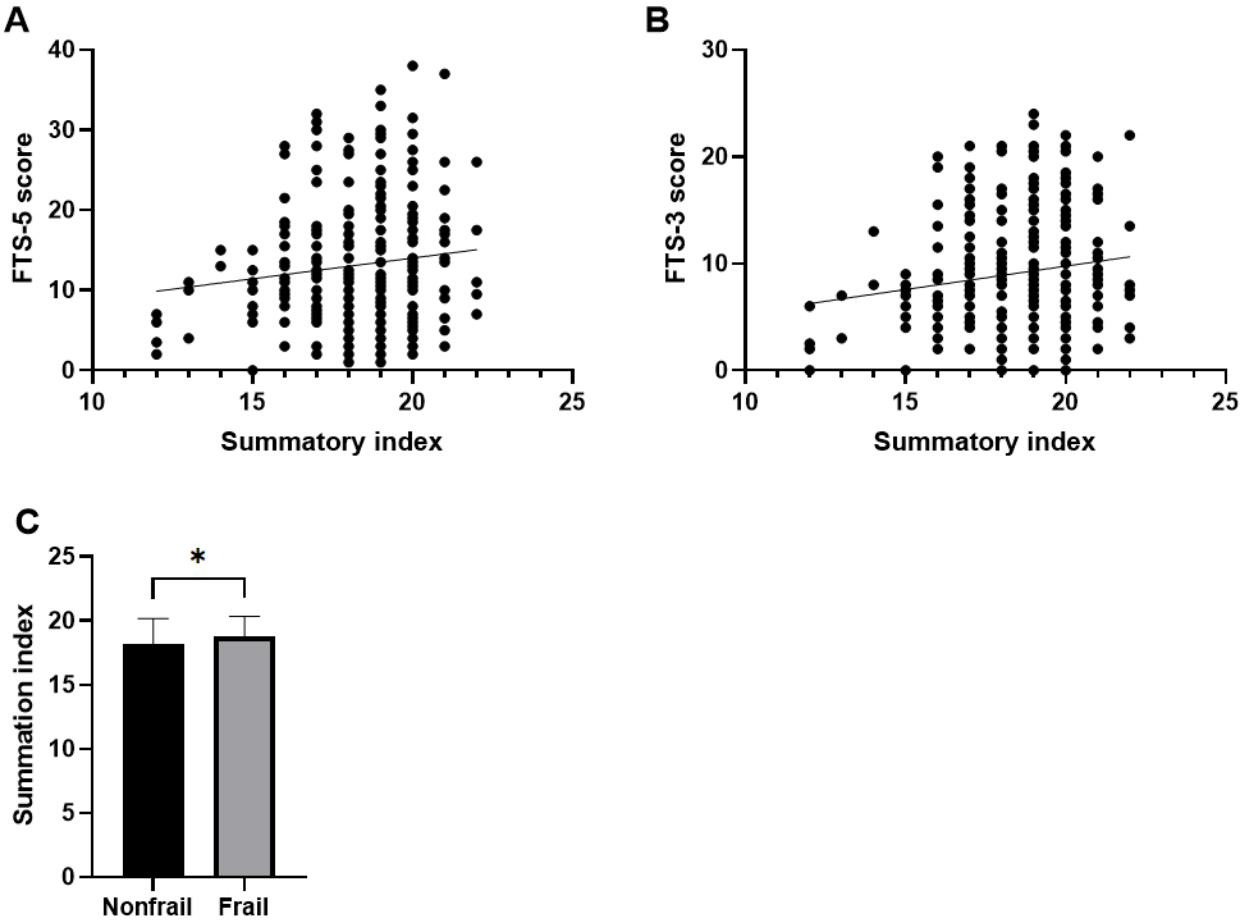


Figure 4 shows the correlation analysis between frailty scores and the summary index and compares this parameter by frailty status. In Figure 4A, the linear regression between the summary index and the FTS-5 score shows a positive correlation, where an increase in the summary index corresponds to a higher FTS-5 score. This correlation is statistically significant, with a Pearson *r* of 0.1319 (*P*=.0368). A similar trend is observed in Figure 4B, which illustrates the correlation between the summary index and the FTS-3 score (a shortened version of FTS-5), showing a Pearson *r* of 0.1562 (*P*=.0132). Figure 4C compares the summary index by frailty status, demonstrating a modest but significant increase in the frail group (*P*=.0310).

XSL•FO
RenderX

Figure 4. Linear regression between frailty scores and summary index, and comparison of the summary index by frailty status. (A) Linear regression between summary index and FTS-5 score, (B) linear regression between summary index and FTS-3 score, and (C) comparison of mean summary index by frailty status diagnosed by FTS-5. The data presented are mean (SEM). Statistical analysis was performed using the *t* test or Mann-Whitney *U* test, as appropriate. *P<.0500. FTS: Frailty Trait Scale.



Relationship Between Urban Quality and Biochemical and Hematological Parameters

Table 4 shows the results of the correlation analysis between the summary index and the biochemical and hematological parameters of the older adult participants. At the biochemical level, only bilirubin levels showed a significant correlation with the summary index (Spearman ρ=0.3325) (P=.0049). Figure 5A displays the linear regression for this parameter, illustrating a direct relationship where higher summary index values are associated with increased bilirubin levels in older adults. At the hematological level, several parameters showed a significant correlation with the summary index. Serum iron exhibited a positive correlation (Pearson r=0.2720; P=.0272), as illustrated in Figure 5E, where higher summary index values correspond to increased serum iron levels. Similarly, transferrin saturation (Spearman ρ=0.2478; P=.0386), RBC count (Pearson r=0.2629; P=.0303), and red blood cell distribution width (RDW) (Pearson r=0.2391; P=.0462) also displayed positive correlations with the summary index. The linear regression graphs in Figure 5B-E, corresponding to each parameter, respectively, further illustrate this direct relationship, where an increase in the summary index was associated with higher values of these hematological parameters.

XSL•FO
RenderX

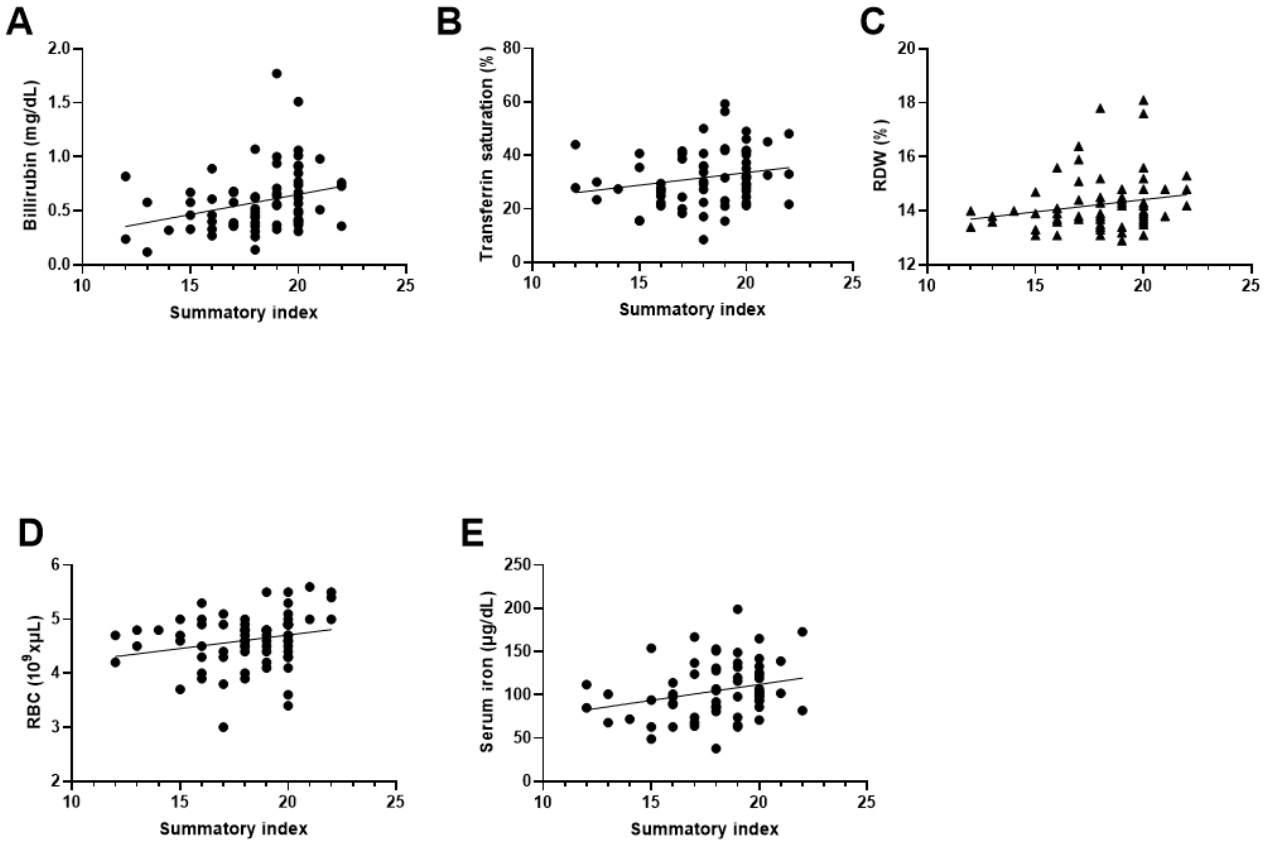
Table . Correlation analysis between summary index and biochemical and hematological parameters in older adults (n=70)[a].

Parameters	Spearman ρ	*P* value
Biochemicals		
GOT[b] (UI/L)	−0.08	.4647
Alkaline phosphatase (UI/L)	−0.09	.4103
LDH[c] (UI/L)	0.01	.9138
Bilirubin (mg/dL)	0.33	.0049
Glycemia (mg/dL)	0.32	−.1196
Uric acid (mg/dL)	0.09	.4329
Total cholesterol (mg/dL)	−0.13	.2537
BUN[d] (mg/dL)	−0.19	.0989
Calcium (mg/dL)	0.03	.7542
Phosphorus (mg/dL)	−0.10	.3741
Total proteins (mg/dL)	−0.11	.3265
Albumin (mg/dL)	0.10	.3683
CRP[e] (mg/dL)	−0.08	.4932
Hematological		
ESR[f] (mm/h)	−0.20	.0868
Serum iron (µg/dL)	0.27	.0272
TIBC[g] (µg/dL)	−0.02	.8634
Transferrin saturation (%)	0.24	.0386
Hemoglobin (g/dL)	0.19	.1122
Hematocrit (%)	0.21	.0758
RBC[h] count (x10^6 cells/µL)	0.26	.0303
MCHC[i] (%)	−0.07	.5511
MCV[j] (fL)	−0.11	.3537
RDW[k] (%)	0.23	.0462
WBC[l] count (cells/µL)	−0.06	.5770
Platelet count (cells/µL)	0.07	.5245
MPV[m] (fL)	0.07	.5626

[a]Nonparametric Spearman correlations were employed for all correlation analyses.
[b]GOT: glutamic-oxaloacetic transaminase.
[c]LDH: lactate dehydrogenase.
[d]BUN: blood urea nitrogen.
[e]CRP: C-reactive protein.
[f]ESR: erythrocyte sedimentation rate.
[g]TIBC: total iron-binding capacity.
[h]RBC: red blood cell.
[i]MCHC: mean corpuscular hemoglobin concentration.
[j]MCV: mean corpuscular volume.
[k]RDW: red blood cell distribution width.
[l]WBC: white blood cell.
[m]MPV: mean platelet volume.

XSL•FO
RenderX

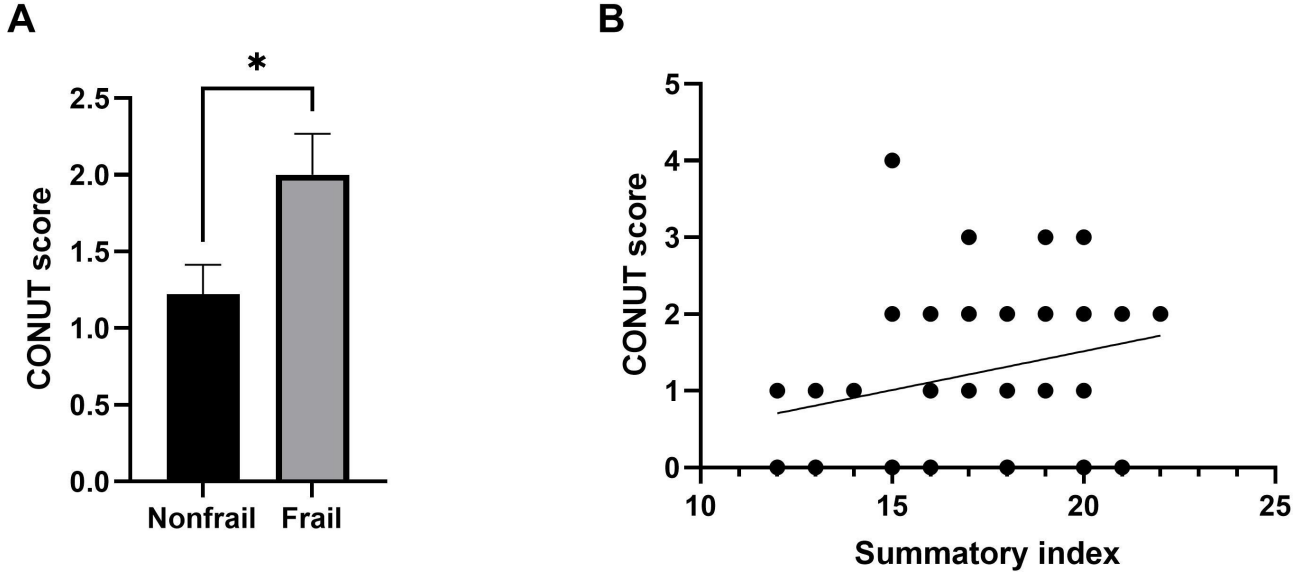
Figure 5. Linear regression analysis between summary index and hematological and biochemical parameters. (A) Bilirubin, (B) transferrin saturation, (C) RDW, (D) RBC, and (E) serum iron. Statistical analysis was performed using linear regression and Spearman correlation analysis. RBC: red blood cell count; RDW: red blood cell distribution width.



Nutritional Status Assessed by CONUT Score According to Frailty Status and Its Correlation With Urban Quality

Figure 6 shows a comparison of the CONUT score by frailty status and its correlation with urban quality, measured through the summary index. In Figure 6A, the frail group shows a significantly higher CONUT score (P=.0323), indicating poorer nutritional status. Meanwhile, Figure 6B displays the linear regression analysis between the CONUT score and the summary index, revealing a significant positive correlation between both variables (Spearman ρ=0.2513; P=.0359).

Figure 6. Relationship between nutritional status assessed using the CONUT score, frailty, and summary index in older adults. (A) Comparison of CONUT score based on frailty status according to the Frailty Trait Scale 5 Scale. (B) Linear regression analysis between CONUT score and summary index. Statistical analysis was performed by *t* test. *P<.05. CONUT: Controlling Nutritional Status.



XSL•FO
RenderX

Discussion

Principal Findings

This study examined the association between urban quality and frailty in older adults by integrating geospatial analysis with nutritional and hematological or biochemical parameters. Key findings revealed that higher urban quality was significantly associated with altered hematological markers (eg, elevated bilirubin, serum iron, and RBC count), poorer nutritional status (as indicated by higher CONUT score), and increased frailty prevalence. Frail individuals living in areas of higher urban quality also exhibited lower handgrip strength, higher frailty scores, and greater proximity to emergency health centers, community centers, and exercise facilities. In summary, older adults residing in neighborhoods with the greatest proximity to services, parks, and health facilities were more frequently classified as frail and exhibited weaker muscle strength, poorer nutritional status, and modest increases in iron-related and erythrocyte markers, while inflammatory and renal-hepatic indicators remained largely unaffected. While these spatial patterns suggest potential links between urban environment and health outcomes in aging populations, the cross-sectional design limits causal inference. Because all exposures and outcomes were assessed at a single point in time, the cross-sectional design precludes establishing temporal precedence; therefore, observed associations cannot be interpreted as causal. Reverse causation is plausible, and residual confounding from unmeasured factors may persist despite statistical adjustment. Several potential sources of bias should also be acknowledged. First, recruiting participants through primary care clinics may overrepresent older adults who are relatively health-seeking, potentially underestimating the true prevalence of frailty (selection bias). Second, minor geocoding errors and reliance on single time-point biomarker measurements may lead to exposure and outcome misclassification. Given these limitations, we emphasize the need for prospective studies to further investigate these associations and clarify their causal pathways.

The results indicate that frailty prevalence in the studied cohort was 17.5% (44/251), which aligns with previous research reporting frailty rates ranging from 12% to 27% in Chilean older adults [19,22]. The higher BMI observed in frail individuals is consistent with studies suggesting that obesity exacerbates physical limitations and inflammation, contributing to frailty [23-25]. Although the frail group had a higher mean age, the difference was not statistically significant. Age is a well-known factor for frailty, with prevalence increasing in advanced age, particularly in individuals aged 80 years and older [26,27]. However, it is important to consider that environmental and lifestyle factors may play a more substantial role than age alone [3]. The lack of significant differences in frailty prevalence across spatial clusters suggests a uniform distribution of risk factors in the evaluated area. This also highlights the need to expand the geospatial analysis to different urban and rural areas to capture the variations in frailty prevalence and associated risk [8].

Consistent with the previous findings, no significant difference in frailty prevalence was observed across quartiles based on the summary index. However, Q4 did show a higher prevalence of frailty than Q1, although this difference did not reach statistical significance, suggesting that a larger sample size may be needed to confirm this trend. In addition, Q4 exhibited a higher FTS-5 score, indicating a greater level of frailty. These findings suggest that frail older adults may be more likely to reside in areas with better urban quality, potentially reflecting adaptive behaviors related to their health status and quality of life [28,29]. For example, improved access to natural environments (an important component of urban quality) has been associated with reduced health inequality linked to income deprivation, indicating that older adults may choose to reside in better quality urban areas to support their health [30]. Likewise, increased greenery in urban neighborhoods has been linked to better physical and mental health, which may encourage frail individuals to settle in or remain within such environments to mitigate health declines [31].

On the other hand, Q4 also showed a significant decrease in handgrip strength, which correlates with a high FTS-5 score and greater frailty prevalence. Reduced handgrip strength is a marker of poorer muscle condition and is associated with increased risk of sarcopenia, chronic diseases, and mortality [32]. These findings suggest that addressing sarcopenia could be a valuable strategy in high-accessibility areas when designing interventions to prevent frailty and promote healthy aging in older adults [33].

This study also found that frail individuals tended to live significantly closer to exercise facilities, community health centers, and emergency medical services, which may reflect adaptive behaviors in response to health care accessibility. While proximity to these facilities suggests that frail older adults may prioritize access to spaces supporting physical activity and social interaction, closeness to emergency and community health centers likely provides critical support for managing both acute and chronic health conditions [31,34]. It is important to note that in this study, exercise facilities were operationally defined as stadiums and green areas equipped with infrastructure for physical activity. This definition represents a limitation when interpreting the associations observed. Moreover, the significant correlation between the summary index and frailty scores reinforces the hypothesis of a close relationship between frailty and urban environment quality. While our findings indicate that frail individuals live nearer to health and recreational facilities, actual use of these resources may be constrained by unmeasured sociodemographic barriers. These may include gender-related mobility limitations (particularly in older women), migration status (eg, language barriers), disability, and digital literacy (eg, limited access to digital health services) [35-37]. Future geospatial studies should incorporate individual-level data on these factors to better assess accessibility gaps in urban aging populations.

Limitations

In addition, some methodological limitations should be noted. The use of a summatory index, while straightforward and consistent with previous local studies [14], may oversimplify the relative importance of each facility type. More complex approaches (eg, weighted indices or principal component

XSL•FO
RenderX

analysis) may offer additional insights in future research. Moreover, spatial autocorrelation and the modifiable areal unit problem were not formally addressed. These are inherent challenges in geospatial analyses and should be explored in future studies using alternative spatial units and subjective environmental assessments. It should also be noted that the spatial clusters used in this study were not constructed based on direct SES indicators. Rather, they were defined using historical, structural, and industrial characteristics of the region. Therefore, any interpretation of these clusters as SES proxies should be made with caution. This limitation highlights the need for future studies to integrate direct SES variables when analyzing spatial health inequalities.

The results support the notion that environmental factors significantly influence the health status of older adults, revealing modest but statistically significant correlation between frailty, urban quality, and select biochemical and hematological parameters. Given the study's cross-sectional design, we emphasize the need for longitudinal research to confirm these associations and further explore causal mechanisms. The observed correlations offer valuable insights into the impact of urban environments on the nutritional status and metabolic health of older populations. Higher urban quality was associated with elevated bilirubin levels, which may reflect better liver function and more effective oxidative stress regulation. However, this correlation is modest and should be interpreted cautiously. Bilirubin, a well-established antioxidant, has been linked to reduced cardiovascular risk and lower systemic inflammation in prior studies [38]. Urban environments with better infrastructure are often associated with healthier lifestyles, including increased physical activity and dietary habits that support liver health [17]. Moreover, the positive correlation between urban quality and serum iron, RBC count, transferrin saturation, and RDW indicates a tentative association with certain hematological parameters [39]. These indicators are essential for evaluating anemia and overall oxygen transport capacity, both of which are critical determinants of functional status and frailty in older adults [18]. Conversely, poor urban quality, characterized by limited access to nutrient-dense food and inadequate health care services, has been linked to higher risks of anemia and micronutrient deficiencies [40]. Our findings are consistent with previous studies suggesting that the built environment influences dietary patterns, which may in turn affect hematological and biochemical health markers [16].

Our results also reveal an association between frailty and higher CONUT scores, underscoring the close interplay between nutritional status and frailty. Malnutrition is a well-established contributor to frailty, and the CONUT score has been validated as a reliable marker of nutritional risk in frail older adults [21,41]. Our findings indicate that frail individuals tend to have poorer nutritional status, which aligns with previous research linking malnutrition to sarcopenia, immune dysfunction, and increased morbidity [19,42,43]. The observed association between elevated CONUT scores and frailty also supports evidence connecting malnutrition to chronic inflammation and sarcopenia, key biological pathways underlying frailty [21,41]. While our study included C-reactive protein as an inflammatory marker (showing no significant correlation with urban quality), future studies should incorporate proinflammatory cytokines such as interleukin-6 and tumor necrosis factor α to more comprehensively assess inflammatory dysregulation [41]. It is important to note that the CONUT score's reliance on serum albumin (a negative acute-phase reactant) may reflect both malnutrition status and subclinical inflammation, complicating its interpretation. The cutoff values used (albumin <3.5 g/dL, lymphocytes <1500/mm^3, and total cholesterol <160 mg/dL) are based on the original CONUT validation [21]; however, their applicability and sensitivity to aging populations warrant further investigation. Interestingly, this positive correlation observed between urban quality and CONUT scores may indicate that better urban infrastructure supports improved nutritional status. Access to fresh food markets and community resources likely promotes better dietary intake, helping to reduce malnutrition and the burden of nutrition-related chronic diseases [44,45]. In contrast, lower urban quality may be associated with food insecurity and limited access to nutrient-dense foods, increasing the risk of frailty through inadequate protein and micronutrient intake [40,46].

It is important to emphasize that these findings highlight the need to address sociourban inequities in age-friendly urban planning. As demonstrated, targeted interventions, such as increasing opportunities for physical activity in low SES areas, can help mitigate health disparities. This suggests that improvements in urban quality should prioritize vulnerable subgroups to reduce frailty gradients across populations [47].

Notably, the paradoxical finding that frailer individuals resided closer to fresh food markets and yet exhibited poorer nutritional status (as indicated by higher CONUT scores) warrants further investigation. Although proximity to food sources is often assumed to enhance dietary intake, various barriers (such as mobility limitations, financial hardships, or lack of social support) may prevent frail older adults from effectively using these resources. This observation is consistent with prior studies indicating that physical access alone does not ensure food security or healthy eating behaviors, particularly among socially and physically vulnerable populations [40,45]. Similarly, the observed proximity of frail individuals to exercise facilities may reflect adaptive strategies to manage functional decline; however, actual engagement in physical activity may be constrained by comorbidities, pain, or reduced mobility [33]. These findings underscore the importance of complementing urban infrastructure with targeted, supportive interventions that bridge the gap between accessibility and actual utilization among frail older adults.

Conclusions

Our research demonstrates that among older adults in Chile, urban quality has a significant impact on their metabolic health, nutritional status, and frailty. Frail individuals residing in areas with higher urban quality exhibited lower handgrip strength and higher frailty scores. In addition, they tended to live closer to emergency rooms, community health centers, and exercise facilities. While this spatial pattern suggests potential accessibility advantages, its relationship with frailty requires further investigation. Moreover, higher CONUT scores in frail individuals and positive correlations between urban quality and

XSL•FO
RenderX

key biochemical or hematological parameters indicate that better urban environments may support healthier metabolic and nutritional profiles. Given the cross-sectional design and localized scope of this study, future research should incorporate longitudinal approaches and larger, more diverse cohorts to confirm these associations and explore the underlying mechanisms. Finally, urban planning strategies promoting age-friendly environments may help mitigate frailty risk; however, longitudinal studies are needed to clarify the roles of selective migration, aging in place, and disparities in urban development in shaping these outcomes.

Acknowledgments

This work has been funded by ANID/FONDECYT (grant 3230783), Interuniversity Center for Healthy Aging, Code RED211993, and ANID/FONDECYT (grant 1211136).

Authors' Contributions

DA, NM, YO, and JCC contributed to conceptualization. DA, YO, and VV participated in data curation and formal analysis. CM, IP, and EF participated in funding acquisition and project administration. DA, YO, JCC, and VV led the investigation. DA, NM, YO, JCC, and MS contributed to methodology. CM, IP, and MS led the resources. YO handled the software. CM and IP participated in supervision and validation. DA, YO, and VV contributed to writing—original draft and writing—review and editing.

Conflicts of Interest

None declared.

References

1. Grinin L, Grinin A, Korotayev A. Global aging: an integral problem of the future. how to turn a problem into a development driver. In: Sadovnichy V, Akaev A, Ilyin I, Malkov S, Grinin L, Korotayev A, editors. Reconsidering the Limits to Growth: A Report to the Russian Association of the Club of Rome: Springer International Publishing; 2023:117-135.
2. Morley JE, Vellas B, van Kan GA, et al. Frailty consensus: a call to action. J Am Med Dir Assoc 2013 Jun;14(6):392-397. [doi: 10.1016/j.jamda.2013.03.022] [Medline: 23764209]
3. Cesari M, Prince M, Thiyagarajan JA, et al. Frailty: an emerging public health priority. J Am Med Dir Assoc 2016 Mar;17(3):188-192. [doi: 10.1016/j.jamda.2015.12.016]
4. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
5. Integrated care for older people (ICOPE) implementation framework: guidance for systems and services. World Health Organization. 2019. URL: https://www.who.int/publications/i/item/9789241515993 [accessed 2025-10-23]
6. Dent E, Martin FC, Bergman H, Woo J, Romero-Ortuno R, Walston JD. Management of frailty: opportunities, challenges, and future directions. Lancet 2019 Oct 12;394(10206):1376-1386. [doi: 10.1016/S0140-6736(19)31785-4] [Medline: 31609229]
7. Yu R, Wang D, Leung J, Lau K, Kwok T, Woo J. Is neighborhood green space associated with less frailty? Evidence from the Mr. and Ms. Os (Hong Kong) Study. J Am Med Dir Assoc 2018 Jun;19(6):528-534. [doi: 10.1016/j.jamda.2017.12.015] [Medline: 29402649]
8. Taylor D, Barrie H, Lange J, Thompson MQ, Theou O, Visvanathan R. Geospatial modelling of the prevalence and changing distribution of frailty in Australia—2011 to 2027. Exp Gerontol 2019 Aug;123:57-65. [doi: 10.1016/j.exger.2019.05.010] [Medline: 31129145]
9. Abe T, Carver A, Sugiyama T. Associations of neighborhood built and social environments with frailty among mid‐to‐older aged Australian adults. Geriatrics Gerontology Int 2021 Oct;21(10):893-899. [doi: 10.1111/ggi.14253]
10. Meijer M, Mette Kejs A, Stock C, Bloomfield K, Ejstrud B, Schlattmann P. Population density, socioeconomic environment and all-cause mortality: a multilevel survival analysis of 2.7 million individuals in Denmark. Health Place 2012 Mar;18(2):391-399. [doi: 10.1016/j.healthplace.2011.12.001]
11. Sugiyama T, Sugiyama M, Mavoa S, Barnett A, Kamruzzaman M, Turrell G. Neighborhood environmental attributes and walking mobility decline: a longitudinal ecological study of mid-to-older aged Australian adults. PLoS One 2021;16(6):e0252017. [doi: 10.1371/journal.pone.0252017] [Medline: 34081707]
12. Martins BA, Visvanathan R, Barrie HR, et al. Built environment and frailty: neighborhood perceptions and associations with frailty, experience from the Nagoya Longitudinal Study. J Appl Gerontol 2021 Jun;40(6):609-619. [doi: 10.1177/0733464820912663] [Medline: 32368955]
13. Thumala D, Kennedy BK, Calvo E, et al. Aging and health policies in Chile: new agendas for research. Health Syst Reform 2017 Oct 2;3(4):253-260. [doi: 10.1080/23288604.2017.1353844] [Medline: 30359182]
14. Ormazábal Y, Arauna D, Cantillana JC, Palomo I, Fuentes E, Mena C. Unveiling the frailty spatial patterns among Chilean older persons by exploring sociodemographic and urbanistic influences based on geographic information systems: cross-sectional study. JMIR Aging 2024 Dec 15;8:e64254. [doi: 10.2196/64254]

XSL•FO
RenderX

15. Mena C, Ormazabal Y, Fuentes E, Palomo I. Impacts of physical environment perception on the frailty condition in older people. Geospat Health 2020 Jun 18;15(1). [doi: 10.4081/gh.2020.888] [Medline: 32575969]
16. Mena C, Fuentes E, Ormazábal Y, Palomo-Vélez G, Palomo I. Role of access to parks and markets with anthropometric measurements, biological markers, and a healthy lifestyle. Int J Environ Health Res 2015;25(4):373-383. [doi: 10.1080/09603123.2014.958134] [Medline: 25236622]
17. Di Ciaula A, Portincasa P. The environment as a determinant of successful aging or frailty. Mech Ageing Dev 2020 Jun;188:111244. [doi: 10.1016/j.mad.2020.111244] [Medline: 32335099]
18. Sepúlveda M, Arauna D, García F, Albala C, Palomo I, Fuentes E. Frailty in aging and the search for the optimal biomarker: a review. Biomedicines 2022 Jun 16;10(6). [doi: 10.3390/biomedicines10061426] [Medline: 35740447]
19. Palomo I, Giacaman RA, León S, et al. Analysis of the characteristics and components for the frailty syndrome in older adults from central Chile. The PIEI-ES study. Arch Gerontol Geriatr 2019;80(70-5):70-75. [doi: 10.1016/j.archger.2018.10.004] [Medline: 30390427]
20. García-García FJ, Carnicero JA, Losa-Reyna J, et al. Frailty Trait Scale-Short Form: a frailty instrument for clinical practice. J Am Med Dir Assoc 2020 Sep;21(9):1260-1266. [doi: 10.1016/j.jamda.2019.12.008] [Medline: 32005416]
21. Lo Buglio A, Bellanti F, Capurso C, Vendemiale G. Controlling Nutritional Status (CONUT) score as a predictive marker in hospitalized frail elderly patients. J Pers Med 2023 Jul 10;13(7):1119. [doi: 10.3390/jpm13071119] [Medline: 37511732]
22. Albala C, Lera L, Sanchez H, et al. Frequency of frailty and its association with cognitive status and survival in older Chileans. Clin Interv Aging 2017;12(995-1001):995-1001. [doi: 10.2147/CIA.S136906] [Medline: 28721027]
23. Yuan L, Chang M, Wang J. Abdominal obesity, body mass index and the risk of frailty in community-dwelling older adults: a systematic review and meta-analysis. Age Ageing 2021 Jun 28;50(4):1118-1128. [doi: 10.1093/ageing/afab039] [Medline: 33693472]
24. Arauna D, Cerda A, García-García JF, et al. Polypharmacy is associated with frailty, nutritional risk and chronic disease in Chilean older adults: remarks from PIEI-ES Study. Clin Interv Aging 2020;15:1013-1022. [doi: 10.2147/CIA.S247444] [Medline: 32636616]
25. Crow RS, Lohman MC, Titus AJ, et al. Association of obesity and frailty in older adults: NHANES 1999-2004. J Nutr Health Aging 2019;23(2):138-144. [doi: 10.1007/s12603-018-1138-x] [Medline: 30697622]
26. Verghese J, Ayers E, Sathyan S, et al. Trajectories of frailty in aging: prospective cohort study. PLoS One 2021;16(7):e0253976. [doi: 10.1371/journal.pone.0253976] [Medline: 34252094]
27. Hoogendijk EO, Afilalo J, Ensrud KE, Kowal P, Onder G, Fried LP. Frailty: implications for clinical practice and public health. Lancet 2019 Oct 12;394(10206):1365-1375. [doi: 10.1016/S0140-6736(19)31786-6] [Medline: 31609228]
28. Rose KC, Gitlin LN, Dennis MP. Readiness to use compensatory strategies among older adults with functional difficulties. Int Psychogeriatr 2010 Dec;22(8):1225-1239. [doi: 10.1017/S1041610210001584] [Medline: 20663239]
29. Diez Roux AV. Investigating neighborhood and area effects on health. Am J Public Health 2001 Nov;91(11):1783-1789. [doi: 10.2105/ajph.91.11.1783] [Medline: 11684601]
30. Mitchell R, Popham F. Effect of exposure to natural environment on health inequalities: an observational population study. Lancet 2008 Nov 8;372(9650):1655-1660. [doi: 10.1016/S0140-6736(08)61689-X] [Medline: 18994663]
31. Sugiyama T, Leslie E, Giles-Corti B, Owen N. Associations of neighbourhood greenness with physical and mental health: do walking, social coherence and local social interaction explain the relationships? J Epidemiol Community Health 2008 May;62(5):e9. [doi: 10.1136/jech.2007.064287] [Medline: 18431834]
32. Cheung CL, Nguyen USDT, Au E, Tan KCB, Kung AWC. Association of handgrip strength with chronic diseases and multimorbidity: a cross-sectional study. Age (Dordr) 2013 Jun;35(3):929-941. [doi: 10.1007/s11357-012-9385-y] [Medline: 22314403]
33. Seo Y, Kim M, Shin H, Won C. Perceived neighborhood environment associated with sarcopenia in urban-dwelling older adults: the Korean Frailty and Aging Cohort Study (KFACS). Int J Environ Res Public Health 2021 Jun 10;18(12):6292. [doi: 10.3390/ijerph18126292] [Medline: 34200703]
34. Giles-Corti B, Donovan RJ. The relative influence of individual, social and physical environment determinants of physical activity. Soc Sci Med 2002 Jun;54(12):1793-1812. [doi: 10.1016/s0277-9536(01)00150-2] [Medline: 12113436]
35. Mitra S, Yao M, Ritchie SG. Gender differences in elderly mobility in the United States. Transportation Res Part A Policy Pract 2021 Dec;154:203-226. [doi: 10.1016/j.tra.2021.10.015]
36. Whitaker KL, Krystallidou D, Williams ED, et al. Addressing language as a barrier to healthcare access and quality. Br J Gen Pract 2022 Jan;72(714):4-5. [doi: 10.3399/bjgp22X718013] [Medline: 34972791]
37. Seifert A. The digital exclusion of older adults during the COVID-19 pandemic. J Gerontol Soc Work 2020;63(6-7):674-676. [doi: 10.1080/01634372.2020.1764687] [Medline: 32401181]
38. Horsfall LJ, Nazareth I, Petersen I. Cardiovascular events as a function of serum bilirubin levels in a large, statin-treated cohort. Circulation 2012 Nov 27;126(22):2556-2564. [doi: 10.1161/CIRCULATIONAHA.112.114066] [Medline: 23110860]
39. Sales CH, Rogero MM, Sarti FM, Fisberg RM. Prevalence and factors associated with iron deficiency and anemia among residents of urban areas of São Paulo, Brazil. Nutrients 2021 May 31;13(6):1888. [doi: 10.3390/nu13061888] [Medline: 34072813]

XSL•FO
RenderX

40. Vilar-Compte M, Burrola-Méndez S, Lozano-Marrufo A, et al. Urban poverty and nutrition challenges associated with accessibility to a healthy diet: a global systematic literature review. Int J Equity Health 2021 Jan 20;20(1):40. [doi: 10.1186/s12939-020-01330-0] [Medline: 33472636]
41. Lo Buglio A, Bellanti F, Carmignano DFP, Serviddio G, Vendemiale G. Association between Controlling Nutritional Status (CONUT) score and body composition, inflammation and frailty in hospitalized elderly patients. Nutrients 2024 Feb 20;16(5):576. [doi: 10.3390/nu16050576] [Medline: 38474705]
42. Calcaterra L, Abellan van Kan G, Steinmeyer Z, Angioni D, Proietti M, Sourdet S. Sarcopenia and poor nutritional status in older adults. Clin Nutr 2024 Mar;43(3):701-707. [doi: 10.1016/j.clnu.2024.01.028] [Medline: 38320461]
43. Wei K, Nyunt MSZ, Gao Q, Wee SL, Yap KB, Ng TP. Association of frailty and malnutrition with long-term functional and mortality outcomes among community-dwelling older adults: results from the Singapore Longitudinal Aging Study 1. JAMA Netw Open 2018 Jul 6;1(3):e180650. [doi: 10.1001/jamanetworkopen.2018.0650] [Medline: 30646023]
44. Caspi CE, Sorensen G, Subramanian SV, Kawachi I. The local food environment and diet: a systematic review. Health Place 2012 Sep;18(5):1172-1187. [doi: 10.1016/j.healthplace.2012.05.006] [Medline: 22717379]
45. Larson NI, Story MT, Nelson MC. Neighborhood environments: disparities in access to healthy foods in the U.S. Am J Prev Med 2009 Jan;36(1):74-81. [doi: 10.1016/j.amepre.2008.09.025] [Medline: 18977112]
46. Bonnefoy M, Berrut G, Lesourd B, et al. Frailty and nutrition: searching for evidence. J Nutr Health Aging 2015 Mar;19(3):250-257. [doi: 10.1007/s12603-014-0568-3] [Medline: 25732208]
47. Rivera-Navarro J, Bonilla L, Gullón P, González-Salgado I, Franco M. Can we improve our neighbourhoods to be more physically active? Residents' perceptions from a qualitative urban health inequalities study. Health Place 2022 Sep;77:102658. [doi: 10.1016/j.healthplace.2021.102658]

Abbreviations

CONUT: Controlling Nutritional Status
FTS: Frailty Trait Scale
FTS-5: Frailty Trait Scale 5
GIS: Geographic Information System
RBC: red blood cell
RDW: red blood cell distribution width
SES: socioeconomic status

Edited by H LaMonica; submitted 21.03.25; peer-reviewed by B Olujobi, F Corica, I Ara, V Gonzalez; revised version received 01.08.25; accepted 02.08.25; published 27.10.25.

Please cite as:
Mena C, Ormazabal Y, Molina N, Fuentes E, Cantillana JC, Villalobos V, Sandoval M, Palomo I, Arauna D
Urban Quality and Biochemical, Hematological, and Nutritional Markers in Older Adults: Cross-Sectional Geospatial Study
JMIR Aging 2025;8:e74313
URL: https://aging.jmir.org/2025/1/e74313
doi:10.2196/74313

© Carlos Mena, Yony Ormazabal, Nacim Molina, Eduardo Fuentes, Juan Carlos Cantillana, Victoria Villalobos, Moises Sandoval, Ivan Palomo, Diego Arauna. Originally published in JMIR Aging (https://aging.jmir.org), 27.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Dynamic Associations of Social and Intellectual Activity With Frailty Trajectory in Middle-Aged and Older Adults in China: Nationwide Longitudinal Study

Fei Xu[1,2*], MPH; Xinlei Miao[1*], MD, PhD; Shuang Liu[1], MPH; Yangxuan He[1], MPH; Meng Li[1], MPH; Jiayi Deng[1], MPH; Yilin Zhu[1], MPH; Jingshan Jiang[1], MPH; Song Leng[1], MD, PhD

[1]Department of Health Management Center, The Second Affiliated Hospital of Dalian Medical University, No.467, Zhongshan Road, Shahekou District, Dalian, Liaoning, China

[2]School of Public Health, Dalian Medical University, Dalian, Liaoning, China

[*]these authors contributed equally

Corresponding Author:
Song Leng, MD, PhD
Department of Health Management Center, The Second Affiliated Hospital of Dalian Medical University, No.467, Zhongshan Road, Shahekou District, Dalian, Liaoning, China

Abstract

Background: Research on the effects of social and intellectual activities on frailty remains limited, particularly longitudinal studies examining how these activities influence the dynamic progression of frailty over time. We aimed to examine this association in a nationally longitudinal study.

Objective: This study aimed to investigate the relationship between social and intellectual and frailty trajectories in the middle-aged and older Chinese population. Specifically, it sought to identify distinct frailty trajectories and quantify the longitudinal protective effects of social and intellectual activities against adverse frailty progression, while exploring differential impacts across age, sex, and residence subgroups.

Methods: This study used data from the China Health and Retirement Longitudinal Study and group-based trajectory modeling identified frailty trajectories over seven years. Frailty was assessed using a 38-item frailty index. The frequency of social activities and intellectual activity were quantified separately. Multivariate ordinal logistic regression analyzed associations between activity frequencies and trajectory membership.

Results: Three different trajectories of frailty were identified in this study: “low progressive trajectories” (n=7208, 65.8%), “moderate progressive trajectories” (n=3061, 28.5%), and “high progressive trajectories” (n=609, 5.7%). Compared to nonparticipants, frequent social activity participation reduced the likelihood of transitioning to the “moderate progressive” trajectory (odds ratio [OR] 0.84, 95% CI 0.75 - 0.94; *P*=.004). Intellectual activity engagement lowered risks of both “moderate” (OR 0.77, 95% CI 0.63 - 0.94; *P*=.01) and “high progressive” trajectories (OR 0.63, 95% CI 0.40 - 0.99; *P*=.04). Subgroup analyses revealed differential effects by age, sex, and residence.

Conclusions: This study confirms the existence of heterogeneous long-term frailty trajectories among middle-aged and older Chinese adults, social and intellectual activities significantly mitigate frailty progression in Chinese middle-aged and older adults. Regular participation in structured social and intellectual activities can effectively delay frailty development in this population.

(JMIR Aging 2025;8:e80152) doi:10.2196/80152

KEYWORDS

frailty; social activity; intellectual activity; trajectory modelling; CHARLS; China Health and Retirement Longitudinal Study

Introduction

Frailty is a complex age-related clinical condition characterized by a decline in physiological capacity across multiple organ systems, leading to increased susceptibility to stressors [1]. As an emerging global health burden, frailty has significant implications for clinical practice and public health [2]. Globally, systematic assessments indicate that the prevalence of frailty among community-dwelling adults aged 65 years and older varies widely, ranging from 4.0% to 59.1% [3]. In a follow-up cohort study involving frail older adults across 28 countries, the incidence of comorbidities associated with frailty was reported to be 43.4 per 1000 person-years [4]. Frailty imposes a substantial burden on national health care utilization and expenditure costs. Given the increasing prevalence and growing harm caused by frailty, identifying it and implementing

XSL•FO
RenderX

interventions to slow its progression is crucial for the healthy [5,6].

As a multifactorial syndrome, frailty is influenced by a range of modifiable risk factors, among which, physical, social, and intellectual activities have garnered significant research interest for their potential protective roles. Numerous studies have confirmed that subjects with low social support or social engagement, measured by social frequency, are more likely to exhibit a rapid increase in frailty trajectory, often validated by professional tools such as the Fried phenotype or frailty indices [7-10]. Similarly, a prospective cohort study suggested that frequent intellectual activities—such as reading, playing games, or other cognitively stimulating pursuits—have been identified as a key modifiable factor for mitigating frailty [11]. While these studies underscore the potential protective role of social and intellectual activities, many focus primarily on frailty status or onset rather than its long-term dynamic progression. Furthermore, the specific correlation between the frequency of these activities and distinct longitudinal frailty trajectories remains underexplored in the context of an aging society, particularly within large national cohorts like CHARLS (China Health and Retirement Longitudinal Study).

In summary, this study aims to evaluate the association between social/intellectual activities and distinct frailty trajectories in middle-aged and older adults. The findings may inform strategies for preventing frailty onset and improving outcomes in individuals with frailty.

Methods

Study Population and Design

The analyses were performed based on the CHARLS. CHARLS is a longitudinal survey of inhabitants in mainland China aged 45 and older that provides information such as socioeconomic position and health status. It covers 17,708 participants from 150 counties of 28 provinces in China [12]. The data on socioeconomic status, lifestyles, medications, health status, and functioning assessments were collected. Details on the study design, sampling procedure, and details of the surveys are available in its report [12].

Our study used four waves of data (2011 - 2018). Participants were restricted to individuals aged ≥45 years who had attended the "health status and function" module at baseline, yielding an initial sample of 17,596. Exclusions comprised 3048 individuals: 497 were aged <45 years, 2436 lacked frailty composite indicator data, and 115 had missing social or intellectual activity data at baseline. During follow-up between 2011 and 2013, 2557 individuals were excluded due to attrition (including death, withdrawal, or loss of contact), while 878 newly eligible individuals were enrolled in 2013. A further 1991 participants were excluded between 2013 and 2018 due to loss to follow-up. A comparison of baseline characteristics between the final analytical cohort and participants lost to follow-up is provided in Multimedia Appendix 1.

Ethical Considerations

Ethical approval for the China Health and Retirement Longitudinal Study (CHARLS) was obtained from the Biomedical Ethics Review Committee of Peking University (Approval Number: IRB00001052-11015). Written informed consent was secured from all participants prior to enrollment. No direct compensation was provided to participants for this survey. For this secondary analysis, we used fully anonymized and deidentified data from CHARLS.

Assessment of Frailty

Frailty was assessed using a frailty index (FI), constructed following a standardized procedure [13]. The detailed method for the calculation of FI was reported in previously published studies [14,15]. Briefly, 38 deficit variables relevant to frailty were selected, including impairments in activities of daily living (ADL) and instrumental activities of daily living (IADL), covering 11 tasks such as personal hygiene, dressing, and financial management. The index also incorporated limitations in physical function (9 items), chronic diseases (9 items), psychological health markers (5 items), and subjective measures (eg, self-rated health).

Each variable was scored dichotomously (0=no deficit, 1=deficit present). The FI score was derived by summing all deficits and dividing by the total number of variables. Frailty was defined as FI ≥ 0.25. Full variable definitions and methodological details are provided in Multimedia Appendix 2.

Assessment of the Frequencies of Social and Intellectual Activities

In the CHARLS survey, participants reported their social and intellectual activities over the preceding month. Social activities, which included four items, were defined as: interacting with friends; attending sports/social clubs; participating in community organizations; and engaging in voluntary or charity work. Intellectual activities, comprising four items, were defined as: playing Ma-jong, playing chess, playing cards, or going to the community club; attending educational courses; investing in stocks; and browsing the internet. Each activity was assigned scores based on the number of participants (Multimedia Appendix 3). The frequency of participants in social activities and intellectual activities was counted separately, categorized into three groups: ≥3 (frequent participation), 1 - 2 (nonregular participation), and 0 (no participation) [11].

Covariates

Based on previous research, several sociodemographic and health-related covariates were included in the study. Sociodemographic covariates comprised age, sex, marital status, location of residence, educational attainment, and retirement status. Health-related covariates included current drinking, current smoking, physical activity levels, and inpatient care. Current drinking and smoking statuses were assessed via self-report using the following questions: "Do you currently smoke?" and "Do you currently drink alcohol?" Marital status was categorized as "married" or "unmarried." Location of residence was classified as "rural" or "urban," and educational attainment was grouped into three categories: "primary school or below," "middle school," and "high school or above." Physical activity levels were categorized based on participation frequency as "inactivity," "low-intensity activities," "moderate activities," or "vigorous activities." Inpatient care was evaluated

XSL•FO
RenderX

by the question: “Have you received inpatient care in the past year?” Retirement status was determined by the question: “Are you currently retired?”

Statistical Analysis

Continuous variables were assessed for normality and are presented as median (25th–75th percentiles) due to their non-normal distribution. Group comparisons were performed using the Kruskal-Wallis test, Pearson’s χ^2 test, Wilcoxon rank-sum test, and Fisher exact test, with false discovery rate correction for multiple comparisons. Categorical variables are summarized as frequencies and percentages, and group differences were assessed using χ^2 tests.

The Group-Based Trajectory Modeling (GBTM) approach was employed to investigate frailty trajectories in middle-aged and older adults. This method applies finite mixture modeling to identify observations with similar temporal developments and assign them to distinct trajectory groups. The optimal trajectory groups were selected based on the following criteria to the greatest extent possible [16]: (1) smaller absolute Bayesian Information Criterion (BIC) values; (2) larger BIC changes; (3) average posterior probability (APP)>0.7; (4) odds of correct classification (OCC) >5; (5) class proportions ≥5%; (6) relative entropy (Ek) is close to 1; (7) parsimony; and (8) interpretability of group patterns.

Subsequently, a multinomial logistic regression model was utilized to estimate the associations between social and intellectual activities and frailty trajectory patterns. Adjusted odds ratios (OR) and corresponding 95% CI are reported.

The multivariable-adjusted model incorporated the following covariates: age at baseline (continuous), sex (male vs female), education (primary school and below, middle school, high school and above), marital status (married vs others), smoking status (yes vs no), alcohol consumption (yes vs no), location of residence (rural vs. city), physical activity levels (inactivity, low-intensity activities, moderate activities, vigorous activities), inpatient care history (yes vs no), and retirement status (yes vs no). The benefits of social and intellectual activities may vary distinctly by subgroup. Sex-based differences in social roles [17], age-related transitions in health status [18], and China’s pronounced rural-urban disparities in infrastructure are all posited as key modifiers of these associations [19]. Thus, stratified analyses were additionally performed in separate models by age groups (<65 y and ≥65 y), sex (male vs female), and location of residence (rural vs city).

Two sensitivity analyses were conducted to assess the robustness of the associations between social/intellectual activities and frailty. First, the 878 newly enrolled participants in 2013 were excluded; second, frailty was stratified into three categories: nonfrailty, prefrailty, and frailty, to investigate its association with social and intellectual activities.

Results

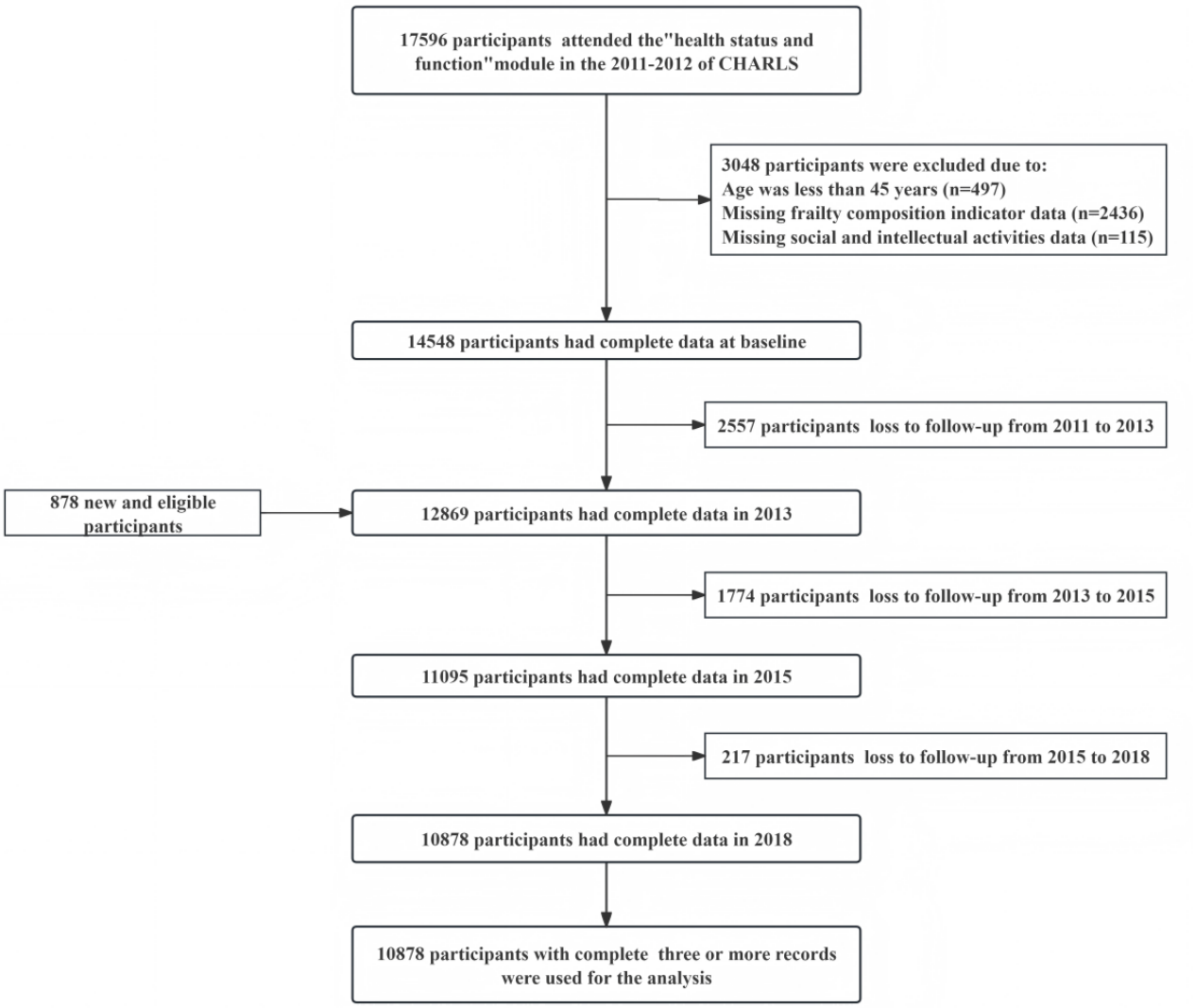
Baseline Characteristics

Of the initial 17,596 participants, a total of 10,878 individuals with three or more complete records comprised the final analytical cohort for trajectory modeling. The detailed flow of participant inclusion and exclusion is presented in Figure 1.

A total of 10,878 participants were included in the study, with 52.0% (n=5655) being female. Among them, 19.7% (n=2143) had a social activities score≥ 3, and 6.1% (n=662) had an intellectual activities score≥ 3. Most participants resided in rural areas, were married, had a primary school education or below, reported physical inactivity, and had retired (P<.001).

XSL•FO
RenderX

Figure 1. Schematic representation of the inclusion and exclusion process for 10,878 participants.



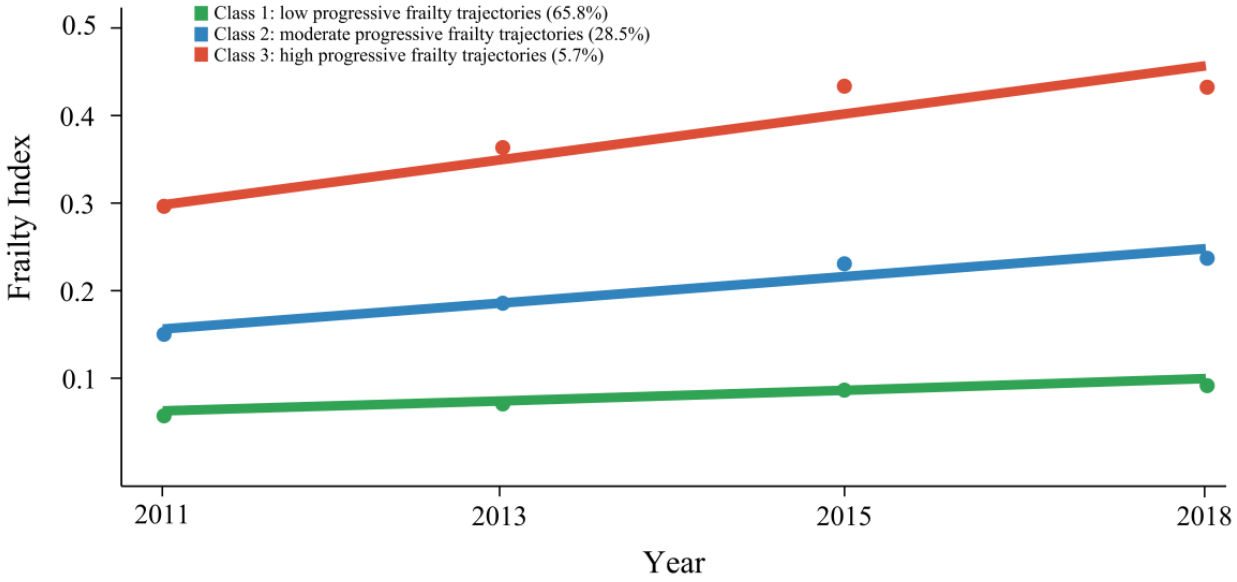
Estimated Frailty Trajectories

The trajectories model, consisting of three different trajectories, showed the best fit using the BIC, Akaike Information Criterion, APP, and interpretability in the trajectories analysis (Multimedia Appendix 4). Consequently, the GBTM with three trajectories was selected as the optimal model. Figure 2 illustrates the trajectories labeled as: (1) “low progressive” (n=7208, 65.8%), (2) “moderate progressive” (n=3,061, 28.5%), and (3) “high progressive” (n=609, 5.7%). The frailty function trajectories across the three groups are depicted in Figure 1. The “low progressive” group, comprising the majority of participants (65.8%), exhibited minimal FI values that remained consistently low throughout the observation period. The “moderate progressive” group (28.5%) showed gradual progression toward mild frailty over seven years, while the “high progressive” group (5.7%) displayed rapid FI escalation within approximately four years, followed by a plateau.

Comparative analysis of baseline characteristics across trajectory groups revealed that participants in the “high progressive” group were older, more likely to be female, married, rural residents, former smokers or alcohol consumers, and had higher educational attainment (high school or above). Additionally, this group exhibited higher rates of physical inactivity, hospitalization history, and reduced participation in social and intellectual activities compared to the “low progressive” group. Detailed distribution of baseline covariates and frailty trajectory groups was provided in Table 1.

XSL•FO
RenderX

Figure 2. Mean trajectories of frailty scores by increasing age among middle-aged and older adults.



XSL•FO
RenderX

Table . Baseline characteristics of the participants according to trajectories of frailty in middle-aged and older adults from CHARLS (China Health and Retirement Longitudinal Study).

Characteristics	Class 1, low progressive (n=7208)	Class 2, moderate progressive (n=3061)	Class 3, high progressive (n=609)	H/χ^2 test	*P* value
Age, years, median (IQR)	56.0 (50.0, 62.0)	60.0 (55.0, 66.0)	63.0 (57.0, 69.0)	366.9	<.001
Sex, n (%)				189.6	<.001
Female	3410 (47.3)	1855 (60.6)	390 (64.0)		
Male	3798 (52.7)	1206 (39.4)	219 (36.0)		
Current drinking, n (%)				175.4	<.001
Yes	2759 (38.3)	813 (26.6)	132 (21.7)		
No	4449 (61.7)	2248 (73.4)	477 (78.3)		
Current smoking, n (%)				91.7	<.001
Yes	2421 (33.6)	775 (25.3)	134 (22.0)		
No	4787 (66.4)	2286 (74.7)	475 (78.0)		
Married, n (%)				110.9	<.001
Yes	6637 (92.1)	2656 (86.8)	501 (82.3)		
No	571 (7.9)	405 (13.2)	108 (17.7)		
Residence, n (%)				59.6	<.001
Rural	4457 (61.8)	2109 (68.9)	433 (71.1)		
City	2751 (38.2)	952 (31.1)	176 (28.9)		
Education, n (%)				351.2	<.001
Primary school and below	4363 (60.5)	2364 (77.2)	507 (83.3)		
Middle school	1824 (25.3)	460 (15.0)	69 (11.3)		
High school and above	1021 (14.2)	237 (7.8)	33 (5.4)		
Physical activities, n (%)				34.8	<.001
Inactivity	4500 (62.4)	1865 (60.9)	399 (65.5)		
Low-intensity activities	617 (8.6)	306 (10.0)	79 (13.0)		
Moderate activities	903 (12.5)	424 (13.9)	67 (11.0)		
Vigorous activities	1188 (16.5)	466 (15.2)	64 (10.5)		
Inpatient care, n (%)				243.2	<.001
Yes	436 (6.0)	385 (12.6)	130 (21.3)		
No	6772 (94.0)	2676 (87.4)	479 (78.7)		
Retirement, n (%)				0.1	.94
Yes	807 (11.2)	349 (11.4)	67 (11.0)		
No	6401 (88.8)	2712 (88.6)	542 (89.0)		
Social activities, n (%)				7.4	.11
0	4375 (60.7)	1938 (63.3)	385 (63.2)		
1 - 2	1387 (19.2)	546 (17.8)	104 (17.1)		
≥3	1446 (20.1)	577 (18.9)	120 (19.7)		
Intellectual activities, n (%)				89.1	<.001
0	5576 (77.4)	2573 (84.1)	538 (88.3)		
1 - 2	1143 (15.9)	339 (11.0)	47 (7.7)		

XSL•FO
RenderX

Characteristics	Class 1, low progressive (n=7208)	Class 2, moderate progressive (n=3061)	Class 3, high progressive (n=609)	H/χ^2 test	*P* value
≥3	489 (6.7)	149 (4.9)	24 (4.0)		

Baseline Social, Intellectual Activities Scores and Frailty Trajectories

The multinomial regression analysis (Model 2) revealed distinct associations between social/intellectual activity scores and frailty trajectory membership. For social activities, compared to the reference group (score=0): participants with 1 - 2 scores showed no significant reduction in the odds of transitioning to the moderate progressive trajectory or the high progressive trajectory. Participants with ≥3 activity scores exhibited a 16% reduction in the odds of belonging to the moderate progressive trajectory (odds ratio [OR] 0.84; 95% CI 0.75 - 0.94; *P*=.004), while no significant association was observed for the high progressive trajectory.

For intellectual activities, compared to the reference group (score=0): participants with 1 - 2 scores demonstrated an 18% reduction in the odds of entering the moderate progressive trajectory (OR 0.82; 95% CI 0.72 - 0.94; *P*=.005) and a 40% reduction in the odds of transitioning to the high progressive trajectory (OR 0.60; 95% CI 0.43 - 0.83; *P*=.002). Participants with ≥3 activity scores showed even stronger protective effects: a 23% reduction in the odds for the moderate progressive trajectory (OR 0.77; 95% CI 0.63 - 0.94; *P*=.01) and a 37% reduction in the odds for the high progressive trajectory (OR 0.63; 95% CI 0.40 - 0.99; *P*=.04). Detailed results are presented in Table 2.

Table . Multinomial logistic regression analysis for the associations of social and intellectual activities with membership in the frailty trajectories group. Model 1 adjusted for age, sex, current drinking, current smoking; Model 2 plus additional adjustments for married, residence, education, physical activities, inpatient care, and retirement.

Variables	Model 1				Model 2			
	Moderate progressive (vs low progressive)		High progressive (vs low progressive)		Moderate progressive (vs low progressive)		High progressive (vs low progressive)	
	OR[a] (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Social activities scores								
0	1.00 (reference)	—[b]	1.00 (reference)	—	1.00 (reference)	—	1.00 (reference)	—
1-2	0.93 (0.83 - 1.05)	.26	0.94 (0.75 - 1.19)	.62	0.95 (0.85 - 1.07)	.43	0.92 (0.72,1.17)	.49
≥3	0.82 (0.73 - 0.92)	.001	0.80 (0.64 - 0.99)	.046	0.84 (0.75 - 0.94)	.004	0.87 (0.69,1.09)	.23
Intellectual activities scores								
0	1.00 (reference)	—	1.00 (reference)	—	1.00 (reference)	—	1.00 (reference)	—
1-2	0.78 (0.68 - 0.90)	<.001	0.54 (0.40 - 0.75)	<.001	0.82 (0.72 - 0.94)	.005	0.60 (0.43,0.83)	.002
≥3	0.67 (0.55 - 0.82)	<.001	0.51 (0.33 - 0.79)	.002	0.77 (0.63 - 0.94)	.01	0.63 (0.40,0.99)	.04

[a]OR: odds ratio.

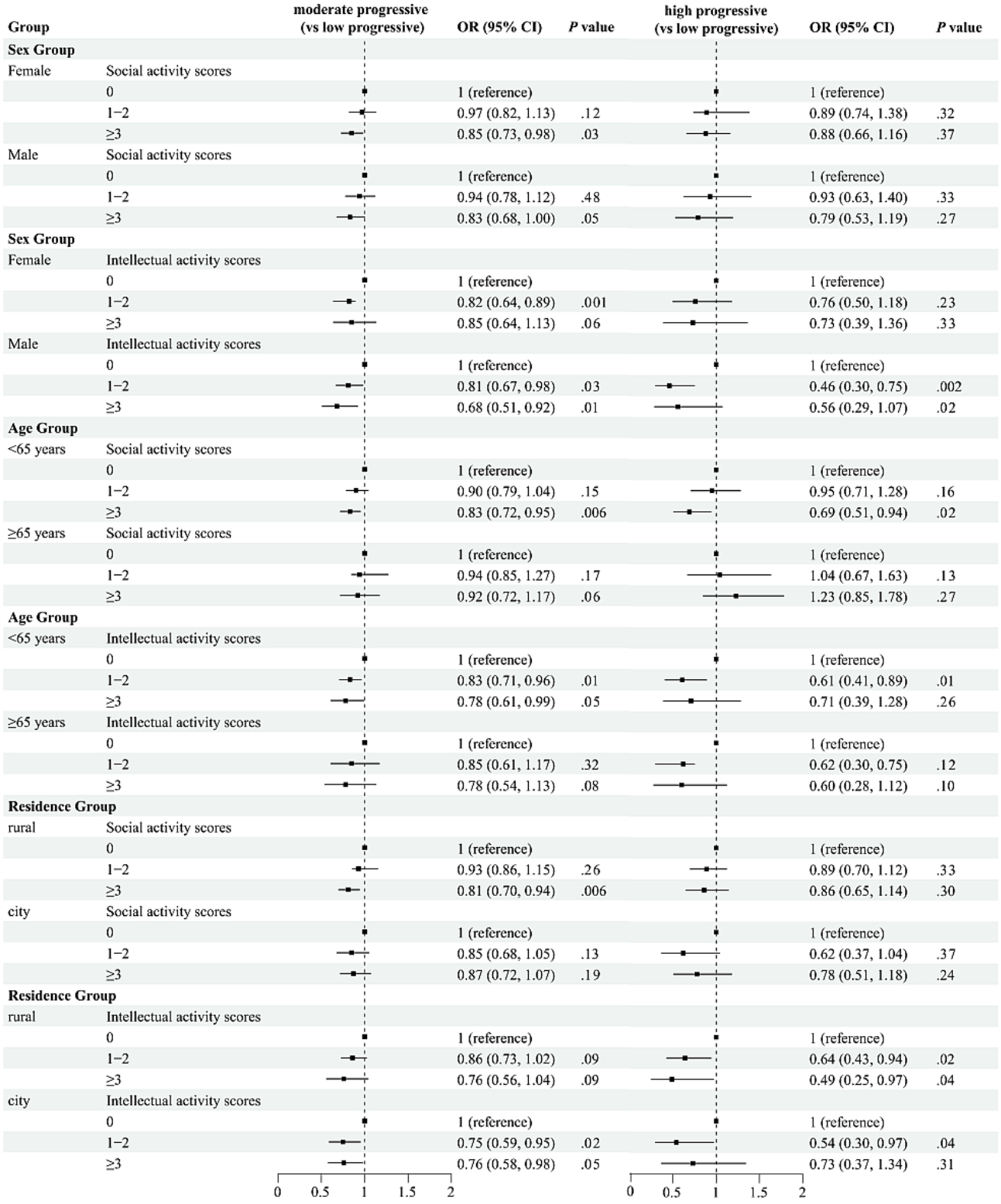
[b]Not applicable.

Stratified Analysis

The stratified analyses are shown in Figure 3. We found that significant associations between social activities scores and frailty trajectories seem to differ with sex and age. In men, higher social activity scores had a significant effect on the moderate progression of reducing frailty, (OR 0.85; 95% CI 0.73 - 0.98, *P*=.003). For women, intellectual activities demonstrated stronger protective effects against frailty progression. Women engaging in frequent intellectual activities had a 32% lower risk of entering the “moderate progressive” trajectory (OR 0.68; 95% CI 0.51 - 0.92, *P*=.001). At the same time, high levels of social activities have a positive role in preventing frailty in people aged <65 years. Frequent social activity reduced the risk of “moderate progressive” (OR 0.83; 95% CI 0.72 - 0.95, *P*=.05) and “high progressive” (OR 0.69; 95% CI 0.51 - 0.94, *P*=.02) trajectories. In terms of the geographical distribution, rural residents showed stronger associations between social activity and favorable vulnerable trajectories(OR 0.81; 95% CI 0.70 - 0.94, *P*=.006), while urban residents benefited more from structured intellectual engagement(OR 0.76; 95% CI 0.58 - 0.98, *P*=.05).

XSL•FO
RenderX

Figure 3. The association of intellectual activities and social activities, stratified by age group, sex group, and residence group. Odds ratio (OR), 95% CI, and *P* values test the significance of the association within each subgroup. Adjusted for age, sex, current drinking, current smoking, marital status, residence, education, physical activities, inpatient care, and retirement.



Sensitivity Analyses

In the sensitivity analyses, first, to account for potential bias introduced by participant selection, the cohort was reanalyzed after excluding 878 participants enrolled in 2013 (Table S1-S2 in Multimedia Appendix 5). Second, participants were categorized into three frailty states based on the frailty index to examine the association between distinct frailty states and engagement in social and intellectual activities. Across all sensitivity analyses, the results remained consistent, further supporting the robustness of the observed associations (Table S1-S2 in Multimedia Appendix 6).

XSL•FO
RenderX

Discussion

Principal Findings

In this longitudinal study, we identified three distinct frailty trajectories—low progressive, moderate progressive, and high progressive. Our findings demonstrated that frequent participation in social and intellectual activities is associated with a reduced risk of transitioning to adverse frailty trajectories. Unlike prior studies that primarily focused on frailty status or onset, our analysis captures the dynamic progression of frailty, thereby offering a longitudinal perspective on how modifiable psychosocial factors influence frailty development over time. Additionally, this study found that both social and intellectual activities can affect the frailty trajectory of different genders, different age groups, and different regions. These associations remained robust across sensitivity analyses, underscoring the potential protective role of engagement in social and intellectual activities against frailty progression. To our knowledge, this is one of the first studies to simultaneously examine the association of both social and intellectual activity frequencies with longitudinal frailty trajectories, providing a dynamic perspective on frailty progression and its modifiable psychosocial predictors.

Association Between Social Activity and Frailty Trajectories

Consistent with previous studies [20-22], our findings demonstrate that frequent participation in social activity is significantly associated with a reduced likelihood of progression along a moderate frailty trajectory. These associations can be explained by various factors: Firstly, social activity often involves physical exertion, which enhances muscle mass, strength, balance, and cardiopulmonary function [23], thereby directly counteracting established pathways linking skeletal muscle structure and function to frailty. However, whereas previous studies largely examined social support or isolation as static risk factors, our trajectory-based approach reveals how varying levels of social activity dynamically influence frailty progression over time. Previous studies have reported that loneliness and social isolation are established risk factors for frailty [24,25]. Social activity mitigates established risk factors for frailty, such as loneliness and social isolation, by providing emotional support and improving psychological well-being. Evidence also suggests that inflammation may play a regulatory role in the progression of frailty; inflammatory responses can be exacerbated by social isolation [26-29]. Thus, increased social activities could attenuate systemic inflammation and delay frailty progression.

Notably, we observed differential effects across subgroups. Our study found that men may derive greater benefits from social activities in slowing frailty progression, potentially due to their tendency to seek support through social networks to manage stress, whereas women may prioritize educational or cultural activities [17]. Notably, the protective effect was more pronounced in participants younger than 65 years of age in the present study, possibly because the middle-aged population socializes more frequently and is more affected. Additionally, geographic differences were evident. Social activities were particularly effective in reducing frailty among rural residents, possibly due to rural China's social structure and residential distribution, which leads to a strong reliance on kinship networks. These networks provide instrumental help and reduce loneliness through the incorporation of physical labor in agriculture-based social activity.

Association Between Intellectual Activity and Frailty Trajectories

Our findings corroborate earlier observational studies linking intellectual activity to a lower risk of frailty [11], but extend them by demonstrating a longitudinal and dose–response relationship between intellectual engagement and frailty trajectory. However, prior research has not adequately addressed the longitudinal effects of social and intellectual activities on frailty trajectories. Frailty often coexists with cognitive impairment [30]. Lack of intellectual activity increases the risks of cognitive impairment [31], which in turn creates a higher risk of frailty occurrences. In a cross-sectional study, Wang et al [32] found that engaging in intellectual activity in later life was associated with a lower prevalence of frailty, particularly among older adult women. This aligns with our observation in females that intellectual activity exerts a stronger protective effect against frailty progression in females. Furthermore, urban residents benefited more from intellectual activities; this may be due to urban areas with better-developed public infrastructure and information accessibility may facilitate cognitive stimulation through intellectual engagement.

Frailty has evolved from the initial concept to a complex, multidimensional bio-psycho-social syndrome [33]. Compared to previous interventions that have focused on physical activity and nutritional strategies, our findings highlight the importance of integrating social and psychological strategies. Based on the results of our study showing a significant association between participation in social activity and intellectual activity and reduced vulnerability development, this encourages the inclusion of frailty interventions in future social activity and intellectual activity intervention strategies in middle-aged and older adults. Specifically, tailored interventions should be implemented. For men, social activities should be actively promoted to delay the development process of frailty, while women should carry out a variety of learning skills from the perspective of intellectual activities. From the age group, we advocate intervention in the development of the frailty trajectory early in middle age, which can effectively prevent and reduce the development of frailty. In rural areas, enhancing public infrastructure could amplify the benefits of intellectual activities, whereas urban communities should prioritize social activity programs to address isolation risks. Furthermore, if social and intellectual activity are performed simultaneously and combined with physical exercise, it may have a positive impact on maintaining the normal operation of the physiological system, thus delaying the progression of cognitive fragility [34-36].

Strengths and Limitations of This Study

To our knowledge, this is the first study to investigate the relationship between social and intellectual activity and frailty trajectory. The strength of this study is that we used a relatively large nationally representative cohort of Chinese middle-aged and elderly people to evaluate the influence of social activity

XSL•FO
RenderX

and intellectual activity on frailty trajectory, and we found that frequent intellectual activity was associated with a lower frailty trajectory, suggesting that we provide a new direction for the management and prevention of frailty. Furthermore, state-of-the-art statistical models were used to fit frailty development trajectories. This approach helps to identify groups of individuals who have experienced similar levels and patterns over time, and more helps to identify people with similar frailty trajectories, thus rationing the intervention for the population.

However, limitations should be acknowledged. First, despite rigorous adjustment, residual confounding from unmeasured factors cannot be ruled out. Specifically, more granular data on socioeconomic status, nutritional status were not available. These factors may influence both activity participation and frailty risk. Second, self-reported activity frequencies are susceptible to recall bias, though nondifferential misclassification would likely attenuate observed associations. Third, the generalizability of our findings to other ethnic or cultural contexts may be limited. The CHARLS cohort, while nationally representative of middle-aged and older Chinese adults, has distinct sociodemographic characteristics that may differ from populations in Western countries.

Conclusion

This study provides novel evidence that social and intellectual activities are independently associated with favorable frailty trajectories in middle-aged and older adults. The sex- and age-specific effects highlight the need for tailored interventions—promoting social engagement in men and intellectual stimulation in women, particularly in younger populations. Furthermore, future research should investigate whether interventions that simultaneously combine social, intellectual, and physical activity could have synergistic effects in maintaining physiological system integrity and thus delaying frailty progression, building on existing evidence.

Acknowledgments

The authors thank all the participants.

Funding

This study was supported by the Dalian Science and Technology Talent Innovation Project (2024RQ059), Dalian Life and Health Sector Guidance Program (2024ZDJH01PT133), 1++ X Program for Large Cohort Study-Clinical Research Incubation Project (YXJS2024006), Basic Research Project of Higher Education Institutions of Liaoning Provincial Department of Education (LJ212510161043), Liaoning Provincial Science and Technology Joint Plan (2025JH2/101800182).

Data Availability

The datasets analyzed during this study are publicly available in the China Health and Retirement Longitudinal Study (CHARLS) repository [12,37].

Authors' Contributions

Conceptualization: FX, XM
Data curation: SL, YH
Formal analysis: JD, ML
Project administration: SL
Methodology: FX, XM
Supervision: FX, XM
Validation: JJ, YZ

Conflicts of Interest

None declared.

Multimedia Appendix 1
Comparison of baseline characteristics between the analytical sample and participants lost to follow-up.
[DOCX File, 19 KB - aging_v8i1e80152_app1.docx]

Multimedia Appendix 2
List of items included in the frailty index in this study.
[DOCX File, 19 KB - aging_v8i1e80152_app2.docx]

Multimedia Appendix 3
List of items included in the social and intellectual activity in this study.
[DOCX File, 14 KB - aging_v8i1e80152_app3.docx]

XSL•FO
RenderX

Multimedia Appendix 4
Fit statistics for global frailty group trajectories in middle-aged and older adults from CHARLS (China Health and Retirement Longitudinal Study).
[DOCX File, 15 KB - aging_v8i1e80152_app4.docx]

Multimedia Appendix 5
Sensitivity analysis 1.
[DOCX File, 23 KB - aging_v8i1e80152_app5.docx]

Multimedia Appendix 6
Sensitivity analysis 2.
[DOCX File, 23 KB - aging_v8i1e80152_app6.docx]

References

1. Dent E, Martin FC, Bergman H, Woo J, Romero-Ortuno R, Walston JD. Management of frailty: opportunities, challenges, and future directions. Lancet 2019 Oct 12;394(10206):1376-1386. [doi: 10.1016/S0140-6736(19)31785-4] [Medline: 31609229]
2. Hoogendijk EO, Afilalo J, Ensrud KE, Kowal P, Onder G, Fried LP. Frailty: implications for clinical practice and public health. Lancet 2019 Oct 12;394(10206):1365-1375. [doi: 10.1016/S0140-6736(19)31786-6] [Medline: 31609228]
3. Collard RM, Boter H, Schoevers RA, Oude Voshaar RC. Prevalence of frailty in community-dwelling older persons: a systematic review. J Am Geriatr Soc 2012 Aug;60(8):1487-1492. [doi: 10.1111/j.1532-5415.2012.04054.x] [Medline: 22881367]
4. Ofori-Asenso R, Chin KL, Mazidi M, et al. Global incidence of frailty and prefrailty among community-dwelling older adults: a systematic review and meta-analysis. JAMA Netw Open 2019 Aug 2;2(8):e198398. [doi: 10.1001/jamanetworkopen.2019.8398] [Medline: 31373653]
5. Walston J, Buta B, Xue QL. Frailty screening and interventions: considerations for clinical practice. Clin Geriatr Med 2018 Feb;34(1):25-38. [doi: 10.1016/j.cger.2017.09.004] [Medline: 29129215]
6. Woolford SJ, Sohan O, Dennison EM, Cooper C, Patel HP. Approaches to the diagnosis and prevention of frailty. Aging Clin Exp Res 2020 Sep;32(9):1629-1637. [doi: 10.1007/s40520-020-01559-3] [Medline: 32356135]
7. Ding YY, Kuha J, Murphy M. Multidimensional predictors of physical frailty in older people: identifying how and for whom they exert their effects. Biogerontology 2017 Apr;18(2):237-252. [doi: 10.1007/s10522-017-9677-9] [Medline: 28160113]
8. Fustinoni S, Santos-Eggimann B, Henchoz Y. Trajectories of phenotypical frailty over a decade in young-old community-dwelling adults: results from the Lc65+ study. J Epidemiol Community Health 2022 Mar;76(3):216-222. [doi: 10.1136/jech-2021-216633] [Medline: 34433618]
9. Jin Y, Yu R, Si H, et al. Effects of social support on frailty trajectory classes among community-dwelling older adults: The mediating role of depressive symptoms and physical activity. Geriatr Nurs 2022;45:39-46. [doi: 10.1016/j.gerinurse.2022.02.029] [Medline: 35303526]
10. Peek MK, Howrey BT, Ternent RS, Ray LA, Ottenbacher KJ. Social support, stressors, and frailty among older Mexican American adults. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences 2012 Nov 1;67(6):755-764. [doi: 10.1093/geronb/gbs081]
11. Huang Y, Guo X, Du J, Liu Y. Associations between intellectual and social activities with frailty among community-dwelling older adults in China: a prospective cohort study. Front Med 2021;8. [doi: 10.3389/fmed.2021.693818]
12. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). Int J Epidemiol 2014 Feb;43(1):61-68. [doi: 10.1093/ije/dys203] [Medline: 23243115]
13. Searle SD, Mitnitski A, Gahbauer EA, Gill TM, Rockwood K. A standard procedure for creating a frailty index. BMC Geriatr 2008 Sep 30;8:24. [doi: 10.1186/1471-2318-8-24] [Medline: 18826625]
14. Yin JH, Zeng YB, Zhou Z, Fang Y. Study on the status of frailty and related determinants among the elderly in China. Zhonghua Liu Xing Bing Xue Za Zhi 2018 Sep 10;39(9):1244-1248. [doi: 10.3760/cma.j.issn.0254-6450.2018.09.019] [Medline: 30293318]
15. Wang X, Chen Z, Li Z, et al. Association between frailty and risk of fall among diabetic patients. Endocr Connect 2020 Oct;9(10):1057-1064. [doi: 10.1530/EC-20-0405] [Medline: 33112808]
16. Nagin DS, Odgers CL. Group-based trajectory modeling in clinical research. Annu Rev Clin Psychol 2010 Mar 1;6(1):109-138. [doi: 10.1146/annurev.clinpsy.121208.131413]
17. Reid N, Young A, Shafiee Hanjani L, Hubbard RE, Gordon EH. Sex-specific interventions to prevent and manage frailty. Maturitas 2022 Oct;164:23-30. [doi: 10.1016/j.maturitas.2022.05.012] [Medline: 35780633]
18. Bellelli F, Consorti E, Hettiarachchige TMK, et al. Relationship among age, education and frailty in older persons. J Frailty Aging 2023;12(4):326-328. [doi: 10.14283/jfa.2023.39] [Medline: 38008985]

XSL•FO
RenderX

19. Xie X, Que J, Sun L, Sun T, Yang F. Association between urbanization levels and frailty among middle-aged and older adults in China: evidence from the CHARLS. BMC Med 2025 Mar 24;23(1):171. [doi: 10.1186/s12916-025-03961-y] [Medline: 40128743]
20. Makizako H, Shimada H, Doi T, et al. Social frailty leads to the development of physical frailty among physically non-frail adults: a four-year follow-up longitudinal cohort study. Int J Environ Res Public Health 2018 Mar 10;15(3):490. [doi: 10.3390/ijerph15030490] [Medline: 29534470]
21. Chon D, Lee Y, Kim J, Lee KE. The association between frequency of social contact and frailty in older people: Korean Frailty and Aging Cohort Study (KFACS). J Korean Med Sci 2018;33(51):e332. [doi: 10.3346/jkms.2018.33.e332]
22. Chong EY, Lim AHS, Mah FCY, Yeo LHW, Ng ST, Yi H. Assessing the psychosocial dimensions of frailty among older adults in Singapore: a community-based cross-sectional study. BMJ Open 2022 Feb;12(2):e047586. [doi: 10.1136/bmjopen-2020-047586]
23. Angulo J, El Assar M, Álvarez-Bustos A, Rodríguez-Mañas L. Physical activity and exercise: strategies to manage frailty. Redox Biol 2020 Aug;35:101513. [doi: 10.1016/j.redox.2020.101513] [Medline: 32234291]
24. Gale CR, Westbury L, Cooper C. Social isolation and loneliness as risk factors for the progression of frailty: the English Longitudinal Study of Ageing. Age Ageing 2018 May 1;47(3):392-397. [doi: 10.1093/ageing/afx188] [Medline: 29309502]
25. Kojima G, Taniguchi Y, Aoyama R, Tanabe M. Associations between loneliness and physical frailty in community-dwelling older adults: A systematic review and meta-analysis. Ageing Res Rev 2022 Nov;81:101705. [doi: 10.1016/j.arr.2022.101705] [Medline: 35932978]
26. Ferrucci L, Fabbri E. Inflammageing: chronic inflammation in ageing, cardiovascular disease, and frailty. Nat Rev Cardiol 2018 Sep;15(9):505-522. [doi: 10.1038/s41569-018-0064-2] [Medline: 30065258]
27. Soysal P, Stubbs B, Lucato P, et al. Inflammation and frailty in the elderly: a systematic review and meta-analysis. Ageing Res Rev 2016 Nov;31:1-8. [doi: 10.1016/j.arr.2016.08.006] [Medline: 27592340]
28. Smith KJ, Gavey S, RIddell NE, Kontari P, Victor C. The association between loneliness, social isolation and inflammation: A systematic review and meta-analysis. Neurosci Biobehav Rev 2020 May;112:519-541. [doi: 10.1016/j.neubiorev.2020.02.002] [Medline: 32092313]
29. Matthews T, Rasmussen LJH, Ambler A, et al. Social isolation, loneliness, and inflammation: a multi-cohort investigation in early and mid-adulthood. Brain Behav Immun 2024 Jan;115:727-736. [doi: 10.1016/j.bbi.2023.11.022] [Medline: 37992788]
30. Yuan Y, Peng C, Burr JA, Lapane KL. Frailty, cognitive impairment, and depressive symptoms in Chinese older adults: an eight-year multi-trajectory analysis. BMC Geriatr 2023 Dec 12;23(1):843. [doi: 10.1186/s12877-023-04554-1] [Medline: 38087208]
31. Lam LCW, Ong PA, Dikot Y, et al. Intellectual and physical activities, but not social activities, are associated with better global cognition: a multi-site evaluation of the cognition and lifestyle activity study for seniors in Asia (CLASSA). Age Ageing 2015 Sep;44(5):835-840. [doi: 10.1093/ageing/afv099] [Medline: 26271049]
32. Wang X, Lu Y, Li C, et al. Associations of lifestyle activities and a heathy diet with frailty in old age: a community-based study in Singapore. Aging (Milano) 2020 Jan 15;12(1):288-308. [doi: 10.18632/aging.102615]
33. Cohen CI, Benyaminov R, Rahman M, Ngu D, Reinhardt M. Frailty: a multidimensional biopsychosocial syndrome. Med Clin North Am 2023 Jan;107(1):183-197. [doi: 10.1016/j.mcna.2022.04.006] [Medline: 36402498]
34. Aartsen MJ, Smits CHM, van Tilburg T, Knipscheer K, Deeg DJH. Activity in older adults: cause or consequence of cognitive functioning? A longitudinal study on everyday activities and cognitive performance in older adults. J Gerontol B Psychol Sci Soc Sci 2002 Mar;57(2):153-162. [doi: 10.1093/geronb/57.2.p153] [Medline: 11867663]
35. Cheong CY, Nyunt MSZ, Gao Q, et al. Risk factors of progression to frailty: findings from the Singapore Longitudinal Ageing Study. J Nutr Health Aging 2020;24(1):98-106. [doi: 10.1007/s12603-019-1277-8] [Medline: 31886815]
36. Jia J, Zhao T, Liu Z, et al. Association between healthy lifestyle and memory decline in older adults: 10 year, population based, prospective cohort study. BMJ 2023 Jan 25;380:e072691. [doi: 10.1136/bmj-2022-072691] [Medline: 36696990]
37. China Health and Retirement Longitudinal Study (CHARLS). URL: https://charls.pku.edu.cn/ [accessed 2025-12-02]

Abbreviations

ADL: Activities of Daily Living
BIC: Bayesian Information Criterion
CHARLS: China Health and Retirement Longitudinal Study
FI: Frailty Index
GBTM: Group-Based Trajectory Modeling
IADL: Instrumental Activities of Daily Living
OR: odds ratio

XSL•FO
RenderX

Edited by A Garcia; submitted 05.07.25; peer-reviewed by L Lan, L Wen, W Zhou; accepted 14.11.25; published 15.12.25.

Please cite as:
Xu F, Miao X, Liu S, He Y, Li M, Deng J, Zhu Y, Jiang J, Leng S
The Dynamic Associations of Social and Intellectual Activity With Frailty Trajectory in Middle-Aged and Older Adults in China: Nationwide Longitudinal Study
JMIR Aging 2025;8:e80152
URL: https://aging.jmir.org/2025/1/e80152
doi:10.2196/80152

© Fei Xu, Xinlei Miao, Shuang Liu, Yangxuan He, Meng Li, Jiayi Deng, Yilin Zhu, Jingshan Jiang, Song Leng. Originally published in JMIR Aging (https://aging.jmir.org), 15.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Transitions in Sarcopenia Status and Cognitive Trajectories Among Middle-Aged and Older Adults in China: Longitudinal Cohort Study

Chun Luo[1], MM; Hao Wu[2], PhD; Xiaoying Shen[1], MM; Shuang Han[3], PhD; Lv Lin[4], MS; Bingyang Liu[5], MM

[1]Department of Endocrinology, The Affiliated Lihuili Hospital of Ningbo University, Ningbo, China

[2]Ningbo Institute of lnnovation for Combined Medicine and Engineering, The Affiliated Lihuili Hospital of Ningbo University, Ningbo, China

[3]Department of Geriatrics, Hangzhou First People's Hospital, Hangzhou, China

[4]Ningbo Municipal Center for Disease Control and Prevention, Ningbo, China

[5]Department of Geriatrics, The Affiliated Lihuili Hospital of Ningbo University, No. 57 Xingning Road, Ningbo, China

Corresponding Author:
Bingyang Liu, MM
Department of Geriatrics, The Affiliated Lihuili Hospital of Ningbo University, No. 57 Xingning Road, Ningbo, China

Abstract

Background: Baseline sarcopenia has been linked to cognitive decline in older adults; however, the impact of longitudinal changes in sarcopenia status on cognitive trajectories remains unclear.

Objective: This aims to examine the association between 2-year transitions in sarcopenia status and subsequent 5-year cognitive trajectories among middle-aged and older adults in China.

Methods: We analyzed data from 8189 participants (median age 58, IQR y; n=432952.9% female) in the China Health and Retirement Longitudinal Study. Sarcopenia status was determined in 2011 and 2013 according to the 2019 Asian Working Group for Sarcopenia criteria, and participants were classified into 7 transition groups based on status changes. Cognitive function was assessed from 2013 to 2018 using standardized *z* scores for executive function and episodic memory. Linear mixed-effects models were applied to evaluate associations between sarcopenia transitions and cognitive trajectories, adjusting for demographic, lifestyle, and health-related covariates.

Results: Progression from a nonsarcopenic state was associated with greater cognitive decline compared to stable nonsarcopenia (β=–0.016, 95% CI –0.026 to –0.007; *P*<.001), with greater decline observed among those progressing from possible sarcopenia to sarcopenia (β=–0.027, 95% CI –0.054 to –0.001; *P*=.04). In contrast, regression from possible sarcopenia was associated with improved cognitive performance (β=0.028, 95% CI 0.015-0.041; *P*<.001). No significant improvement was observed among individuals regressing from established sarcopenia. Subgroup analyses showed consistent benefits of regression from possible sarcopenia across sex, age, residence, and education groups, except among urban residents (*P*=.05).

Conclusions: Progression in sarcopenia status was independently associated with accelerated cognitive decline, whereas regression from possible sarcopenia predicted cognitive benefit. These findings highlight possible sarcopenia as a clinically actionable and potentially reversible stage and underscore the importance of early identification and intervention to preserve cognitive health in aging populations.

(JMIR Aging 2025;8:e78277) doi:10.2196/78277

KEYWORDS

aging; cognitive function; China Health and Retirement Longitudinal Study; longitudinal studies; sarcopenia

Introduction

Dementia currently affects more than 50 million people worldwide, and its prevalence is projected to reach 150 million by 2050 as the global population continues to age [1]. The societal and economic consequences are profound, with cumulative costs estimated to exceed US $14.5 trillion between 2020 and 2050 [2,3]. Despite intensive research efforts, current therapeutic options remain limited in effectiveness, reinforcing the urgent need to identify modifiable risk factors that may help delay or prevent cognitive decline [4].

Sarcopenia, a progressive skeletal muscle disorder characterized by decline in muscle mass, strength, and physical performance, is highly prevalent among older adults, affecting approximately 10%-27% of this population globally [5]. In addition to its established associations with frailty, falls, disability, and mortality [6-8], emerging evidence suggests a link between sarcopenia and neurocognitive impairment, including mild cognitive impairment and dementia [9-12]. Importantly, unlike many neurodegenerative conditions, sarcopenia is potentially modifiable through targeted interventions such as resistance exercise, adequate protein intake, and nutritional

XSL•FO
RenderX

supplementation, making it a promising intervention target for preserving cognitive health in aging populations [13-15].

However, most previous studies have focused on sarcopenia assessed at a single time point and have not accounted for its dynamic nature [9-12]. Sarcopenia may progress or regress over time due to changes in lifestyle or health status, and these transitions may differentially influence cognitive aging. Understanding the longitudinal impact of sarcopenia progression or recovery on cognitive trajectories may provide critical insights into optimal timing of interventions.

To address this gap, we analyzed data from the China Health and Retirement Longitudinal Study (CHARLS), a nationally representative cohort of middle-aged and older adults in China. We investigated whether 2-year transitions in sarcopenia status were associated with subsequent 5-year cognitive trajectories. To our knowledge, this is among the first large-scale longitudinal studies to evaluate the cognitive implications of dynamic changes in sarcopenia status, capturing both progression and regression patterns. We hypothesized that progression to possible sarcopenia or sarcopenia would be associated with accelerated cognitive decline, whereas regression from these states would be linked to a slower rate of cognitive deterioration. Additionally, we examined whether these associations differed across subgroups defined by sex, age, and residential status, and educational attainment, as these factors influence sarcopenia risk [16-18] and modify its relationship with cognitive function.

Methods

Study Population

This study used data from the CHARLS, a nationally representative cohort of Chinese adults aged ≥45 years [19]. The CHARLS adopted a multistage, stratified probability sampling strategy, with baseline interviews conducted in 2011 (wave 1) and follow-up assessments in 2013, 2015, and 2018 (waves 2 - 4). Sarcopenia status was evaluated in waves 1 and 2, while cognitive function was assessed from Wave 2 onward. Data from wave 5 (2020) were used only in sensitivity analyses due to the partial use of remote video interviews during the COVID-19 pandemic.

Of the 17,708 participants at baseline, we excluded those <45 years of age, those with missing sarcopenia or cognitive data, and those who reported psychiatric or memory-related disorders. Participants lacking sarcopenia assessments in wave 2 or without any cognitive follow-up data were further excluded [20,21].

Assessment of Sarcopenia

Sarcopenia was defined according to the 2019 criteria of the Asian Working Group for Sarcopenia [22]. Severe sarcopenia was classified as low muscle mass accompanied by both low muscle strength and poor physical performance. Sarcopenia was defined as low muscle mass combined with either low muscle strength or poor physical performance, whereas possible sarcopenia was defined as low muscle strength or reduced physical performance in the absence of low muscle mass.

Appendicular skeletal muscle mass was estimated using an anthropometric prediction equation validated against dual-energy X-ray absorptiometry (DXA), demonstrating high correlation (R^2=0.90) [23,24]. Sex-specific thresholds for low appendicular skeletal muscle mass or height squared were defined as the lowest 20% of the sex-specific distribution: <7.00 and <7.04 kg/m² for men and <5.25 and <5.32 kg/m² for women in waves 1 and 2, respectively [24,25]. Muscle strength was assessed by handgrip strength, measured twice for each hand using a handheld dynamometer; the maximum value was used for analysis. Low muscle strength was defined as <28 kg for men and <18 kg for women. Physical performance was evaluated using the 5-time chair stand test (≥12 s), walking speed (<1.0 m/s), or the Short Physical Performance Battery (≤9 points) [26].

Cognitive Function Assessment

Cognitive function was evaluated using structured face-to-face interviews covering 2 cognitive domains: executive function and episodic memory [27,28]. Executive function was assessed through 3 tasks: orientation to time (0 - 5 points), serial subtraction (0 - 5 points), and figure drawing (0 - 1 point), with a total score ranging from 0 to 11. Episodic memory was measured using immediate and delayed recall of a 10-word list, and the mean number of correctly recalled words across both trials was used as the final score (range: 0 - 10). Global cognitive function was calculated as the sum of executive function and episodic memory scores, yielding a total score from 0 to 21. All cognitive outcomes were standardized to *z* scores using the baseline mean and standard deviation, with higher values indicating better cognitive performance.

Covariates

Covariates were selected a priori based on previous research [29]. Demographic factors included age and sex. Socioeconomic variables encompassed education level (elementary or lower vs secondary or higher), marital status (married or partnered vs other), and place of residence (urban vs rural). Health behaviors included smoking status (ever vs never), alcohol use (ever vs never), and BMI (kg/m²). Chronic conditions included self-reported or medication-confirmed diagnoses of hypertension, diabetes, dyslipidemia, heart disease, and stroke.

Pain and depressive symptoms were also included as covariates due to their known associations with both sarcopenia and cognitive outcomes [30-33]. Pain was assessed dichotomously (yes or no). Depressive symptoms were measured using the 10-item Center for Epidemiological Studies Depression Scale (CES-D-10), a validated instrument in the CHARLS population that is widely used in aging research [34,35]. Scores on the CES-D-10 range from 0 to 30, with higher scores indicating more severe depressive symptoms; continuous CES-D-10 scores were included in all models.

Ethical Considerations

The study protocol of the CHARLS was reviewed and approved by the Biomedical Ethics Review Committee of Peking University (approval No. IRB00001052-11015). All study procedures involving human participants adhered to the ethical standards of Peking University and the Declaration of Helsinki

XSL•FO
RenderX

(1964) and its later amendments. Written informed consent was obtained from all participants prior to data collection. All data used in this secondary analysis were deidentified and publicly available; thus, no additional institutional review was required.

Statistical Analysis

Baseline Characteristics

Participants were categorized into groups based on changes in sarcopenia status from wave 1 to wave 2. Baseline characteristics were summarized using medians and IQRs for continuous variables and frequencies with percentages for categorical variables.

Sarcopenia Transitions and Cognitive Decline

The primary objective of this analysis was to estimate the rate of cognitive decline associated with different transition patterns in sarcopenia status. Linear mixed-effects models (LMMs) with random intercepts and random slopes were used to model individual cognitive trajectories over time. LMMs account for within-person correlation of repeated measures, accommodate participants with unequal numbers of observations, and are robust to unbalanced longitudinal data under the missing-at-random assumption. Cognitive decline was defined as the annual rate of change in standardized cognitive *z* scores over time, as estimated from the time effect in the LMMs. More negative regression coefficients (β) indicated a faster decline in cognitive function. This modeling approach has been widely applied in longitudinal studies of cognitive aging [36-39]. Given the robustness of LMMs to missing outcome data, no imputation procedure was implemented [40,41]. Model estimates were reported as β coefficients with 95% CIs.

Models were built sequentially. Model 1 adjusted for age, age squared (to capture nonlinearity), and sex. Model 2 further adjusted for education, marital status, residence, smoking, alcohol use, BMI, and BMI squared. Model 3 additionally controlled for chronic conditions.

Subgroup Analyses

Subgroup analyses were stratified by sex (male vs female), age group (<65 vs ≥65 y), residence (urban vs rural), and educational level (elementary or lower vs secondary or higher). These analyses were implemented using 3-way interaction terms between sarcopenia transition group, time, and each subgroup variable.

The sarcopenia transition variable included 3 categories to reflect changes in sarcopenia status between 2 survey waves: (1) progression from nonsarcopenia, (2) change in possible sarcopenia (including progression to sarcopenia and regression to nonsarcopenia), and (3) regression from sarcopenia.

Each subgroup variable was entered into the LMMs as a fixed effect together with its interaction terms with group and time. Effect modification was assessed using the overall *P* value for the 3-way interaction. Stratified associations within each subgroup were further reported to illustrate the pattern and magnitude of associations.

Sensitivity Analyses

Several sensitivity analyses were performed to assess the robustness of the findings. First, memory scores from wave 4 were harmonized using equipercentile equating to address potential inconsistencies arising from differences in test versions [42]. Second, a 9-category classification of sarcopenia transitions was applied to capture finer distinctions in status changes and assess whether more granular classification would yield similar associations with cognitive outcomes. Third, wave 5 cognitive data were included, despite partial implementation of remote video assessments during the COVID-19 pandemic, to incorporate the most recent follow-up and examine the stability of the results.

Additional Analyses

We further examined whether baseline sarcopenia status (wave 1) predicted subsequent cognitive decline during follow-up. In addition, to identify which diagnostic component contributed most to cognitive deterioration, we evaluated the unique and longitudinal effects of each sarcopenia component—low muscle strength, low muscle mass, and low physical performance—within the fully adjusted LMM (Model 3). For each component, we estimated the time × component interaction term, where a significant coefficient indicated that the component was associated with the rate of cognitive decline over time. We also calculated the semi-partial R^2 (part R^2) values and corresponding 95% CIs using 1000 bootstrap resamples, representing the proportion of variance in cognitive decline uniquely explained by each component after accounting for covariates.

All analyses were conducted using the R software (version 4.4.1; R Foundation for Statistical Computing), and part R^2 values were obtained with the *partR2* package (0.9.2). A 2-sided *P* value <.05 was considered statistically significant.

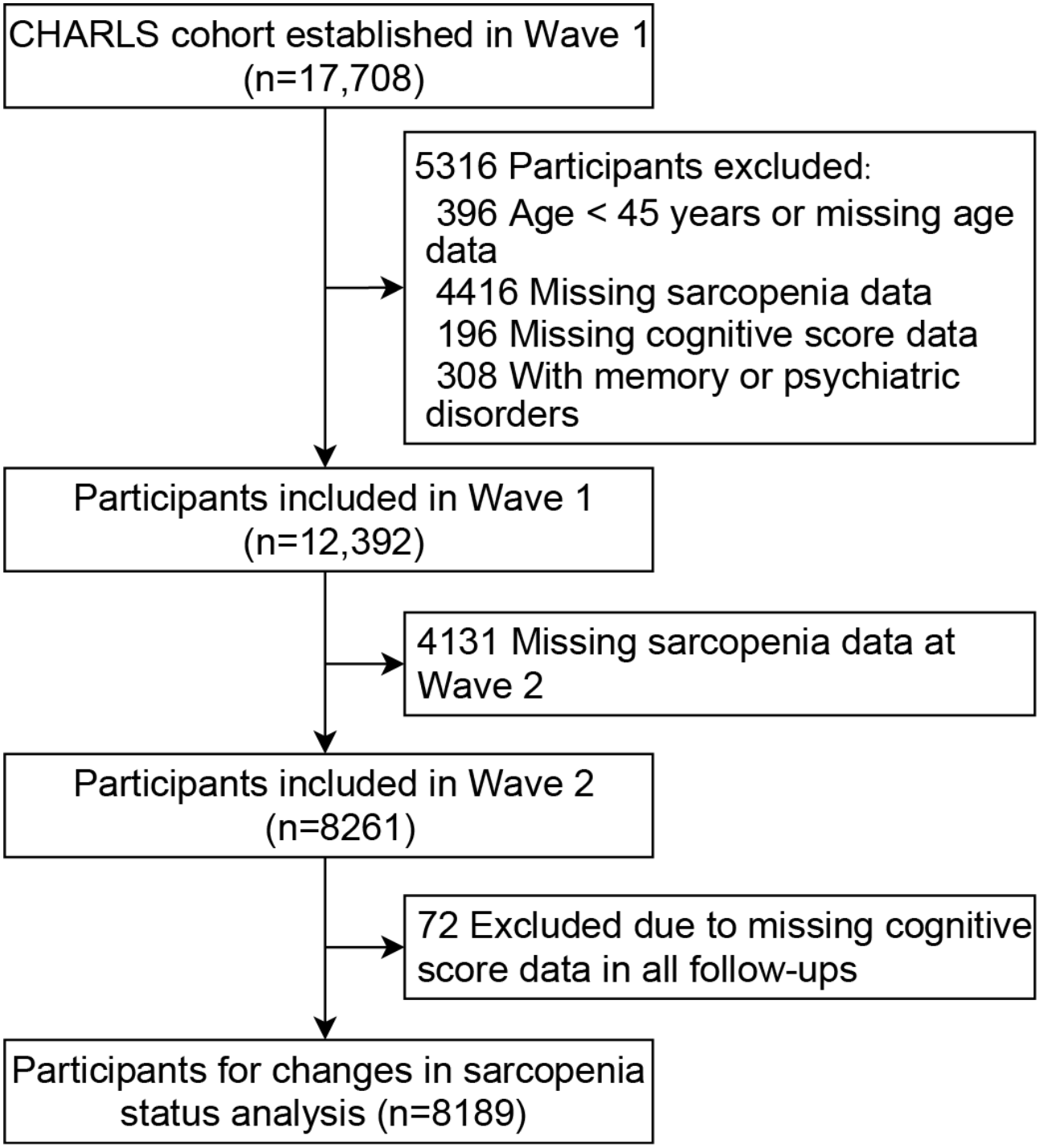
Results

Participant Grouping and Baseline Characteristics

A total of 8189 participants were included in the final analysis after applying the exclusion criteria (median age 58 y; n=4329, 52.9% female; Figure 1). Participants were initially classified into 9 groups based on sarcopenia status transitions between wave 1 and wave 2 (Table 1).

XSL•FO
RenderX

Figure 1. Flowchart illustrating participant selection and inclusion criteria. CHARLS: China Health and Retirement Longitudinal Study.



XSL•FO
RenderX

Table . Transition of sarcopenia status across wave 1 and wave 2[a].

Waves 1 and 2	Participants, n (% of total)	Trend of change
Nonsarcopenia (n=4870)		
Nonsarcopenia	3755 (45.9)	Stable
Possible sarcopenia	830 (10.1)	Progression
Sarcopenia	285 (3.5)	Progression
Possible sarcopenia (n=2010)		
Nonsarcopenia	1022 (12.5)	Regression
Possible sarcopenia	839 (10.2)	Stable
Sarcopenia	149 (1.8)	Progression
Sarcopenia (n=1309)		
Nonsarcopenia	161 (2.0)	Regression
Possible sarcopenia	96 (1.2)	Regression
Sarcopenia	1052 (12.8)	Stable

[a]The time interval between wave 1 and wave 2 survey was 2 years in the China Health and Retirement Longitudinal Study (CHARLS).

However, several transition groups—such as progression from nonsarcopenia to sarcopenia—comprised fewer than 3.5% (n=285) of the total sample, resulting in limited statistical power and unstable model estimates. To improve statistical robustness and model stability, transition groups with similar clinical patterns were consolidated. The final analysis included 7 groups: stable nonsarcopenia, progression from nonsarcopenia, stable possible sarcopenia, regression from possible sarcopenia, progression from possible sarcopenia, stable sarcopenia, and regression from sarcopenia. Baseline characteristics across the 7 transition groups are summarized in Table 2.

XSL•FO
RenderX

Table . Baseline characteristics of participants by changes in sarcopenia status.

Baseline	Nonsarcopenia		Possible sarcopenia			Sarcopenia	
Trend of change	Stable	Progression	Regression	Stable	Progression	Stable	Regression
Participants, n	3755	1115	1022	839	149	1052	257
Age (y), median (IQR)	55 (49-60)	59 (53-64)	57 (51-63)	62 (56-69)	66 (61-73)	69 (64-75)	64 (60-70)
Sex, n (%)							
Male	1908 (50.8)	527 (47.3)	403 (39.4)	319 (38.0)	67 (45.0)	495 (47.1)	141 (54.9)
Female	1847 (49.2)	588 (52.7)	619 (60.6)	520 (62.0)	82 (55.0)	557 (52.9)	116 (45.1)
Marital status, n (%)							
Married or partnered	3505 (93.3)	986 (88.4)	922 (90.2)	705 (84.0)	117 (78.5)	788 (74.9)	210 (81.7)
Other marital status	250 (6.7)	129 (11.6)	100 (9.8)	134 (16.0)	32 (21.5)	264 (25.1)	47 (18.3)
Education, n (%)[a]							
Elementary or below	2168 (57.7)	836 (75.0)	739 (72.3)	681 (81.2)	139 (93.3)	944 (89.8)	216 (84.0)
Secondary or above	1587 (42.3)	279 (25.0)	283 (27.7)	158 (18.8)	10 (6.7)	107 (10.2)	41 (16.0)
Residence, n (%)							
Rural	2289 (61.0)	767 (68.8)	694 (67.9)	562 (67.0)	102 (68.5)	822 (78.1)	206 (80.2)
Urban	1466 (39.0)	348 (31.2)	328 (32.1)	277 (33.0)	47 (31.5)	230 (21.9)	51 (19.8)
Smoking status, n (%)							
Never smokers	2252 (60.0)	683 (61.3)	669 (65.5)	560 (66.7)	94 (63.1)	597 (56.7)	133 (51.8)
Ever smokers	1503 (40.0)	432 (38.7)	353 (34.5)	279 (33.3)	55 (36.9)	455 (43.3)	124 (48.2)
Drinking status, n (%)							
Never drinkers	2330 (62.1)	764 (68.5)	726 (71.0)	655 (78.1)	116 (77.9)	739 (70.2)	169 (65.8)
Ever drinkers	1425 (37.9)	351 (31.5)	296 (29.0)	184 (21.9)	33 (22.1)	313 (29.8)	88 (34.2)
Hypertension, n (%)							
No	2407 (64.1)	665 (59.6)	608 (59.5)	350 (41.7)	88 (59.1)	648 (61.6)	164 (63.8)
Yes	1348 (35.9)	450 (40.4)	414 (40.5)	489 (58.3)	61 (40.9)	404 (38.4)	93 (36.2)
Dyslipidemia, n (%)[a]							
No	2480 (66.3)	790 (70.9)	668 (65.7)	522 (62.5)	104 (69.8)	838 (79.9)	189 (73.8)
Yes	1259 (33.7)	324 (29.1)	349 (34.3)	313 (37.5)	45 (30.2)	211 (20.1)	67 (26.2)
Diabetes, n (%)[a]							
No	3300 (88.0)	956 (85.8)	867 (84.9)	691 (82.5)	132 (88.6)	965 (91.9)	228 (89.1)
Yes	451 (12.0)	158 (14.2)	154 (15.1)	147 (17.5)	17 (11.4)	85 (8.1)	28 (10.9)
Heart disease, n (%)[a]							
No	3412 (91.3)	976 (87.9)	890 (87.7)	690 (82.7)	127 (85.8)	931 (89.1)	228 (89.1)
Yes	324 (8.7)	134 (12.1)	125 (12.3)	144 (17.3)	21 (14.2)	114 (10.9)	28 (10.9)
Stroke, n (%)[a]							
No	3706 (99.0)	1102 (98.9)	1006 (98.6)	798 (95.6)	146 (98.0)	1032 (98.4)	251 (98.0)
Yes	39 (1.0)	12 (1.1)	14 (1.4)	37 (4.4)	3 (2.0)	17 (1.6)	5 (2.0)
Pain, n (%)[a]							
No	2784 (74.2)	732 (65.7)	620 (60.7)	464 (55.4)	86 (57.7)	646 (61.5)	158 (61.5)

XSL•FO
RenderX

Baseline	Nonsarcopenia		Possible sarcopenia			Sarcopenia	
Trend of change	Stable	Progression	Regression	Stable	Progression	Stable	Regression
Yes	970 (25.8)	382 (34.3)	402 (39.3)	374 (44.6)	63 (42.3)	405 (38.5)	99 (38.5)
Depressive symptoms[a,b], median (IQR)	16 (13-20)	17 (13-22)	18 (14-23)	19 (15-25)	21 (16-25)	19 (15-25)	19 (15-24)
BMI (kg/m^2)[a], median (IQR)	23.8 (21.9-26.1)	22.9 (20.9-25.5)	24.1 (22.3-26.4)	24.8 (22.7-27.4)	21.8 (21.0-22.9)	19.1 (17.9-20.1)	19.6 (18.6-20.2)

[a]Missing data were observed for education (n=1), pain (n=4), diabetes (n=10), stroke (n=21), dyslipidemia (n=30), BMI (n=39), heart disease (n=45), and depressive symptoms (n=369).

[b]Depressive symptoms were assessed using the 10-item Center for Epidemiological Studies Depression Scale.

Sarcopenia Transitions and Cognitive Decline

In the fully adjusted model (Model 3), progression from nonsarcopenia was associated with a significantly faster decline in global cognitive function (β=–0.016, 95% CI –0.026 to –0.007; *P*<.001), as was progression from possible sarcopenia (β=–0.027, 95% CI –0.054 to –0.001; *P*=.04), compared with stable trajectories. In contrast, regression from possible sarcopenia was linked to improvement in global cognition (β=0.028, 95% CI 0.01-0.041; *P*<.001). Regression from sarcopenia did not confer a statistically significant cognitive benefit (Table 3).

Table . Associations between transitions in sarcopenia status and global cognitive decline[e].

Groups	Model 1[a]		Model 2[b]		Model 3[c]	
	β[d] (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value
Nonsarcopenia						
Stable	0 (reference)	—	0 (reference)	—	0 (reference)	—
Progression	–0.016 (–0.025 to –0.007)	<.001	–0.016 (–0.025 to –0.007)	<.001	–0.016 (–0.026 to –0.007)	<.001
Possible sarcopenia						
Stable	0 (reference)	—	0 (reference)	—	0 (reference)	—
Regression	0.027 (0.014 to 0.039)	<.001	0.026 (0.014 to 0.039)	<.001	0.028 (0.015 to 0.041)	<.001
Progression	–0.025 (–0.050 to 0)	.05	–0.028 (–0.053 to –0.003)	.03	–0.027 (–0.054 to –0.001)	.04
Sarcopenia						
Stable	0 (reference)	—	0 (reference)	—	0 (reference)	—
Regression	0.006 (–0.013 to 0.024)	.56	0.006 (–0.013 to 0.025)	.53	0.003 (–0.016 to 0.022)	.75

[a]Model 1 was adjusted for baseline age, age squared, and sex.

[b]Model 2 further adjusted for educational level, marital status, residence, smoking status, alcohol consumption, BMI, and BMI squared.

[c]Model 3 additionally included dyslipidemia, diabetes, hypertension, cardiovascular disease, stroke, depressive symptoms, and pain.

[d]β: regression coefficient.

[e]Data were analyzed using linear mixed-effects regression models. The β values (95% CI) represent the annual change in cognitive *z* scores (SD/year), relative to the reference group.

Domain-specific analyses yielded results consistent with the findings for global cognition. Progression from nonsarcopenia was associated with declines in executive function (β=–0.009, 95% CI –0.018 to –0.001; *P*=.03) and episodic memory (β=–0.021, 95% CI –0.034 to –0.009; *P*<.001). Regression from possible sarcopenia was associated with improvements in both executive function (β=0.014, 95% CI 0.001-0.026; *P*=.03) and episodic memory (β=0.039, 95% CI 0.022-0.056; *P*<.001). No significant improvements were found among those who regressed from sarcopenia (Tables S1–S2 in Multimedia Appendix 1).

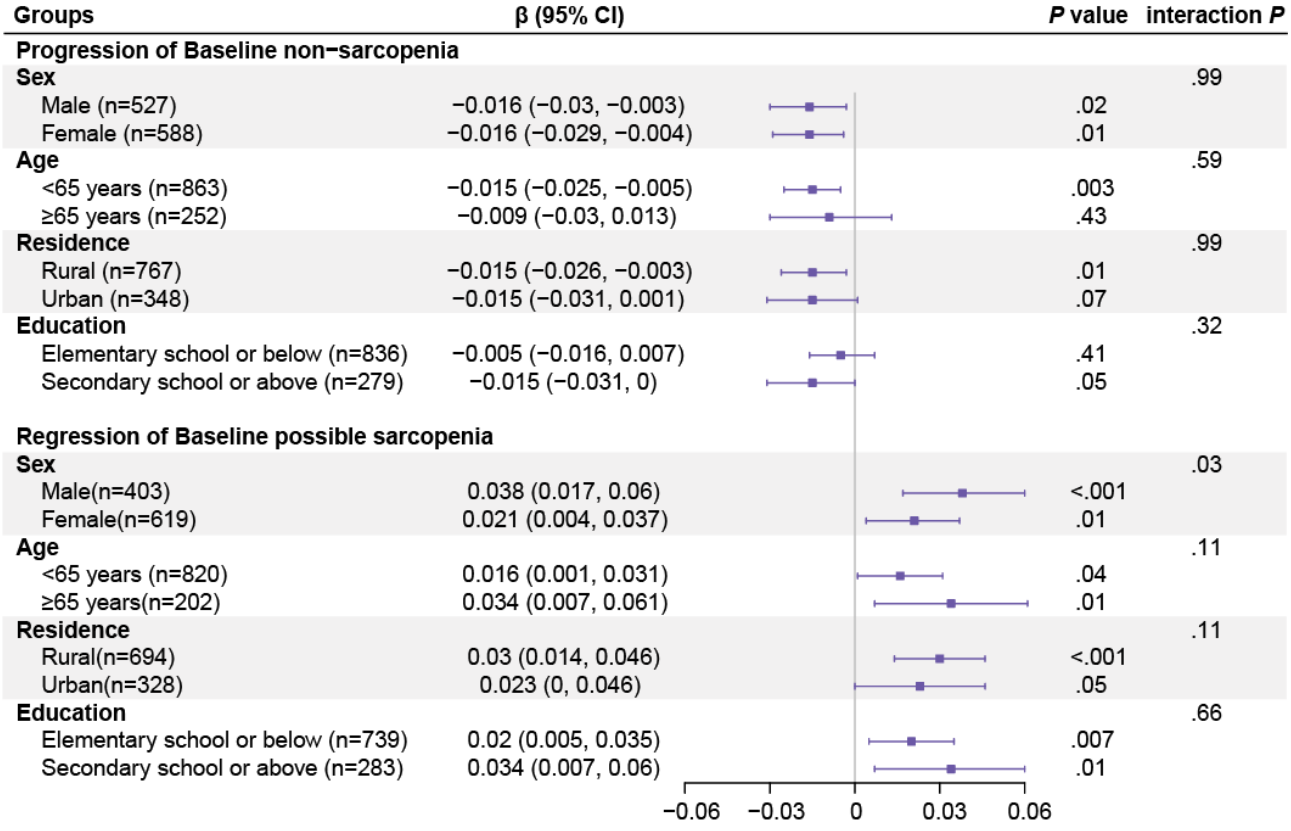
Subgroup Analyses

The results of subgroup analyses for global cognitive function are presented in Figure 2 and Figure S1 in Multimedia Appendix 1. Among participants without sarcopenia at baseline, those who progressed to possible sarcopenia or sarcopenia showed faster declines in cognition compared with those who remained nonsarcopenic. These associations were statistically significant

XSL•FO
RenderX

in both men and women, in adults aged <65 years, and among those living in rural areas. Marginal associations were observed among urban residents (*P*=.07) and participants with secondary school education or higher (*P*=.05). However, no significant effect modification was observed across sex, age, residence, or education (all *P* for interaction >.05).

Figure 2. Subgroup analyses of the association between sarcopenia status transitions and global cognitive decline. Results are presented for 2 key transitions: progression from nonsarcopenia and regression from possible sarcopenia. Associations were estimated using linear mixed-effects models, stratified by sex, age group (<65 vs ≥65 y), residential location (urban vs rural), and educational level. Models were adjusted for baseline age and age squared, sex (except in sex-stratified models), education (except in education-stratified models), marital status, residence (except in residence-stratified models), smoking status, alcohol use, BMI and BMI squared, dyslipidemia, diabetes, hypertension, cardiovascular disease, stroke, depressive symptoms, and pain. Error bars indicate 95% CIs. Because transitions from baseline possible sarcopenia include both progression to sarcopenia and regression to nonsarcopenia, a single omnibus interaction *P* value was calculated across these transitions; therefore, the same interaction *P* value is shown here and in Figure S1 in Multimedia Appendix 1.



Among participants with baseline possible sarcopenia, transitions demonstrated opposite associations with cognition. Regression to nonsarcopenia was consistently associated with slower cognitive decline across all subgroups, with significant associations observed in most subgroups (all *P*<.05), except for a marginal association among urban residents (*P*=.05). Conversely, progression from possible sarcopenia to sarcopenia was associated with faster cognitive decline, and a significant interaction with sex was observed (*P* for interaction=.03). Specifically, in the progression from possible sarcopenia to sarcopenia, cognitive decline was significant among men (β=–0.054, 95% CI –0.094 to –0.014; *P*=.008) but not among women (β=–0.004, 95% CI –0.039 to 0.031; *P*=.81).

Among participants with baseline sarcopenia, regression to possible sarcopenia or nonsarcopenia was not associated with significant cognitive benefit in any subgroup (all *P*>.05). Domain-specific subgroup analyses for executive function and episodic memory are presented in Tables S3-S4 in Multimedia Appendix 1.

Sensitivity Analyses

The findings were generally robust across all sensitivity analyses. When wave 4 memory scores were harmonized using equipercentile equating (Table S5 in Multimedia Appendix 1), the direction of associations remained unchanged; however, among participants who progressed from possible sarcopenia to sarcopenia, effect estimates were attenuated and no longer statistically significant. When analyses were repeated using the original 9-category classification of sarcopenia transitions (Tables S6-S7 in Multimedia Appendix 1), the results were broadly consistent with the primary analysis. The strongest associations were observed among participants progressing from nonsarcopenia to possible sarcopenia (β=–0.018, 95% CI –0.028 to –0.007; *P*<.001) and those regressing from possible sarcopenia to nonsarcopenia (β=0.028, 95% CI 0.015-0.041; *P*<.001). Incorporating wave 5 cognitive data (Table S8 in Multimedia Appendix 1) did not materially alter the associations, and effect estimates were similar in magnitude and direction to those observed in the main models. Overall, across all sensitivity analyses, progression from nonsarcopenia consistently predicted accelerated global cognitive decline, regression from possible

XSL•FO
RenderX

sarcopenia was associated with slower decline, and regression from established sarcopenia did not yield significant cognitive benefit.

Additional Analyses

We evaluated whether baseline sarcopenia status predicted subsequent cognitive decline. The participant selection flowchart is presented in Figure S2 in Multimedia Appendix 2, and baseline characteristics stratified by sarcopenia status are summarized in Table S9 in Multimedia Appendix 1. As shown in Table S10 in Multimedia Appendix 1, baseline sarcopenia was associated with faster decline in global cognition, executive function, and episodic memory.

We further investigated the relative contributions of individual diagnostic components of sarcopenia to cognitive decline. Longitudinal associations for each component are reported in Table S11 in Multimedia Appendix 1. Among the 3 components, low muscle strength was significantly associated with accelerated cognitive decline (β=–0.016, 95% CI –0.027 to –0.004; *P*=.006) and accounted for the largest proportion of explained variance (part R^2=0.002, 95% CI 0.000 - 0.022), as shown in Table S12 in Multimedia Appendix 1. In contrast, low muscle mass was not significantly associated with cognitive decline (β=–0.028; *P*=.23), and low physical performance demonstrated only a marginal association (β=–0.020; *P*=.06).

Discussion

This study used longitudinal data from a nationally representative cohort to examine the association between changes in sarcopenia status and cognitive trajectories among middle-aged and older adults. We found that progression from a nonsarcopenic state was significantly associated with accelerated cognitive decline, whereas regression from possible sarcopenia—but not from established sarcopenia—was linked to better cognitive performance. Importantly, the beneficial association of regression from possible sarcopenia with global cognition was consistently observed across multiple sensitivity and subgroup analyses. These findings highlight the importance of early detection and timely intervention during the possible sarcopenia stage, when the potential to preserve cognitive function and delay neurodegeneration may be greatest.

Previous CHARLS-based studies have reported associations between baseline sarcopenia and subsequent cognitive trajectories [11,35]. Our study extends this evidence by focusing on longitudinal changes in sarcopenia status (progression and regression), rather than baseline sarcopenia alone, using nationally representative data. In our cohort, 31.1% (n=2543) of the participants experienced a change in sarcopenia status over 2 years—15.4% (n=1264) progressed and 15.6% (n=1279) regressed—consistent with prior CHARLS findings [24]. Both progression from nonsarcopenia and progression from possible sarcopenia were consistently associated with steeper declines in global cognition, executive function, and episodic memory. In contrast, regression from possible sarcopenia was associated with improved cognitive outcomes. Notably, regression from established sarcopenia was not associated with meaningful cognitive gains, suggesting a limited therapeutic window once sarcopenia becomes entrenched.

In the subgroup analyses, participants who progressed to possible sarcopenia or sarcopenia exhibited faster cognitive decline than those who maintained a stable nonsarcopenic state. Significant associations were observed in certain subgroups (eg, men, women, individuals aged <65 y, and rural residents), whereas other subgroups showed attenuated or nonsignificant associations. None of the interaction terms for sex, age, residence, or education reached statistical significance. Consistent with guidance from Sun et al [43], the nonsignificant interactions observed in our study do not imply identical effects across subgroups but may reflect small true differences that our sample lacked adequate power to detect. Although the direction of associations between sarcopenia progression and cognitive decline was consistent across demographic strata, subtle population heterogeneity cannot be excluded. These findings underscore the need for larger, adequately powered prospective studies to validate potential subgroup-specific effects.

For instance, in the age-stratified analyses, the association between sarcopenia progression and accelerated cognitive decline was evident among participants aged <65 years but attenuated in those aged ≥65 years, although the interaction effect was not statistically significant. This finding contrasts with prior cross-sectional CHARLS research [44], which reported significant associations between sarcopenia and cognitive function in older adults. Although Du et al [44] examined cross-sectional relationships in participants aged ≥60 years, our longitudinal design captured within-person changes in cognition over time—an approach that better reflects temporal dynamics but is also influenced by attrition and survivorship effects. Selective attrition in longitudinal cohorts often leads to loss of participants with poorer cognitive function or declining health, leaving a more resilient survivor subset with reduced between-person variability [45]. Consequently, those at the greatest risk for cognitive decline may be disproportionately lost to follow-up, attenuating observed effects among older adults. Moreover, unaccounted competing risks such as mortality or severe morbidity may further underestimate associations; Austin and Fine [46] noted that treating participants who experience competing events as censored can bias estimates downward by excluding high-risk individuals. Finally, the smaller sample size in the ≥65-year group likely reduced statistical power and widened confidence intervals. Collectively, these factors likely contributed to the absence of statistically significant associations in older adults despite consistent effect directions across age strata.

Importantly, regression from possible sarcopenia was consistently associated with improved global cognitive function across all subgroups, with significant results (all *P*<.05) except among urban residents, where the association reached marginal significance (*P*=.05). These findings suggest that recovery from possible sarcopenia—representing an intermediate and potentially reversible stage—may provide a critical window for intervention to preserve cognitive health.

Our sensitivity analyses demonstrated that the main findings were robust. Incorporating wave 5 data, which reflect the most

XSL•FO
RenderX

recent cognitive assessments, yielded results nearly identical to the primary analyses, underscoring the stability and reproducibility of our findings. In contrast, when wave 4 memory scores were harmonized using equipercentile equating, associations in the progression from possible sarcopenia to sarcopenia group were attenuated and no longer statistically significant. This attenuation likely reflects the conservative nature of the equating method, which reduces variability across test versions and may compress score distributions, thereby lowering statistical power despite consistent effect directions [42].

Similarly, when the original 9-category classification of sarcopenia transitions was used, the association between progression from nonsarcopenia to sarcopenia and cognitive decline became weaker. This was primarily due to the very small number of participants in this transition group, which reduced statistical power and produced less stable estimates. Moreover, the strong association observed in the nonsarcopenia to possible sarcopenia group is biologically plausible. Possible sarcopenia represents an early transitional stage characterized predominantly by reduced muscle strength, a parameter that has been shown to be a more sensitive indicator of functional decline and adverse health outcomes than muscle mass alone. Consistent with our additional analyses, low muscle strength—but not low muscle mass—was independently associated with faster cognitive decline. These findings align with evidence from population-based cohorts and Mendelian randomization studies showing that handgrip strength is more strongly linked to cognitive performance than muscle mass [47,48]. Therefore, cognitive deterioration may become detectable during the possible sarcopenia stage, before the onset of advanced muscle wasting in established sarcopenia. Across all analyses, however, regression from possible sarcopenia consistently demonstrated cognitive benefits, reinforcing the notion that this transitional stage is particularly amenable to intervention [49].

Several biological mechanisms may explain these associations. The muscle-brain axis has been proposed as a key pathway linking musculoskeletal and cognitive health [50]. Skeletal muscle secretes myokines—signaling molecules that regulate systemic inflammation, metabolic homeostasis, and neuroplasticity [51,52]. Sarcopenia may reduce myokine production and disrupt neuroprotective signaling, whereas reversal through physical activity may restore these pathways [52-54]. In addition, sarcopenia progression is associated with chronic inflammation and oxidative stress—both implicated in cognitive decline and neurodegeneration [55-58]—which may be mitigated through improved muscle function [59,60]. Emerging evidence also suggests a role of gut microbiota dysbiosis in sarcopenia-related neuroinflammation, underscoring the potential importance of the gut-muscle-brain axis [61,62].

These findings have important implications for clinical practice and public health. Routine screening for possible sarcopenia should be incorporated into cognitive risk assessment protocols for aging populations. Such screening is feasible using simple measures—such as grip strength and chair-stand tests—that can be implemented in primary care or community settings without specialized equipment. Early detection may help identify individuals at elevated risk of cognitive decline who could benefit from timely, multidisciplinary interventions. In our cohort, more than half of participants with possible sarcopenia reverted to a nonsarcopenic state (Table 1), reinforcing the reversibility of early-stage functional decline and the importance of timely intervention. Growing evidence from intervention trials suggests that resistance training, alone or combined with nutritional supplementation, can improve muscle function and sarcopenia-related outcomes [13-15], supporting the potential reversibility of early-stage sarcopenia. Future studies should evaluate whether such interventions can also yield sustained cognitive benefits.

This study has several limitations. First, its observational nature precludes causal inference. While associations were consistent across sensitivity analyses, randomized controlled trials are needed to establish causality. Second, sarcopenia was assessed using a validated anthropometric prediction equation rather than dual-energy DXA, the reference standard. Although highly correlated with DXA (R^2=0.90) [23], some misclassification may have occurred. However, the 15% sarcopenia prevalence in our study aligns with previous meta-analyses using DXA or BIA, supporting its validity [5]. Third, modifications to the episodic memory test in 2018 may have introduced measurement inconsistencies; therefore, we applied equipercentile equating to harmonize memory scores across waves. Although this conservative adjustment attenuated associations in the progression from possible sarcopenia to sarcopenia group, the overall direction of effects remained consistent, supporting the robustness of the main findings [42]. Fourth, although physical exercise has been shown to be significantly associated with both cognitive function and sarcopenia [63,64], information on physical activity was missing in 61.0% (n=4995/8189) of participants in our cohort. Therefore, physical exercise was not included as a covariate in our analyses, and residual confounding related to physical activity cannot be excluded. Finally, residual confounding remains possible. Relevant factors such as hearing loss or social isolation—which may influence both sarcopenia and cognitive decline—were not fully captured and should be explored in future research [65].

Progression from a nonsarcopenic state—including progression from possible sarcopenia—is associated with accelerated cognitive decline, whereas regression from possible sarcopenia is linked to cognitive benefit. These findings identify possible sarcopenia as a critical and potentially reversible stage in the trajectory toward sarcopenia and cognitive impairment. They also support incorporating routine screening for possible sarcopenia into cognitive risk assessment protocols for aging populations. Given the simplicity, low cost, and feasibility of screening measures such as grip strength and chair-stand tests, implementation in primary care and community settings is both practical and scalable. Early identification and intervention at the possible sarcopenia stage may offer a key opportunity to preserve cognitive health and delay neurodegenerative processes. Future research should prioritize developing and evaluating targeted, evidence-based strategies to prevent sarcopenia progression and promote cognitive resilience in at-risk populations.

XSL•FO
RenderX

Acknowledgments

We sincerely thank the China Health and Retirement Longitudinal Study (CHARLS) team for making the data available for this study. We also appreciate the efforts of the field staff and the invaluable contributions of the participants. This study was supported by the Zhejiang Provincial Medical and Health Science and Technology Project (grant 2023KY1032) and the Natural Science Foundation of Ningbo (grant 2023J229). The funders had no role in the study design, data collection, analysis, interpretation, or manuscript preparation.

Data Availability

The data used in this study are from the CHARLS, which is available to registered users for academic research purposes. Access to the dataset requires user registration and approval through the CHARLS official website [66].

Authors' Contributions

CL and BL conceived and designed the study. CL, HW, XS, and SH conducted data curation, statistical analysis, and initial interpretation of the findings. CL and HW contributed to the literature review and drafting of the manuscript. LL and BL critically revised the manuscript for important intellectual content and supervised the overall project. All authors contributed to the final version of the article, approved it for submission, and agreed to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Associations between sarcopenia transitions and cognitive trajectories; subgroup analyses by demographic factors; three sensitivity analyses; and associations between baseline sarcopenia and cognition, including contributions of muscle strength, mass, and performance.
[XLSX File, 44 KB - aging_v8i1e78277_app1.xlsx]

Multimedia Appendix 2
Subgroup analyses of sarcopenia transition effects on global cognition and the participant flow for the baseline sarcopenia analysis.
[DOCX File, 228 KB - aging_v8i1e78277_app2.docx]

References

1. Nichols E, Steinmetz JD, Vollset SE, et al. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. Lancet Public Health 2022 Feb;7(2):e105-e125. [doi: 10.1016/S2468-2667(21)00249-8]
2. Su M, Wang T, Zou C, Cao K, Liu F. Global, regional, and national burdens of Alzheimer's disease and other forms of dementia in the elderly population from 1999 to 2019: a trend analysis based on the Global Burden of Disease Study 2019. Ibrain 2024;10(4):488-499. [doi: 10.1002/ibra.12181] [Medline: 39691425]
3. Chen S, Cao Z, Nandi A, et al. The global macroeconomic burden of Alzheimer's disease and other dementias: estimates and projections for 152 countries or territories. Lancet Glob Health 2024 Sep;12(9):e1534-e1543. [doi: 10.1016/S2214-109X(24)00264-X] [Medline: 39151988]
4. Reuben DB, Kremen S, Maust DT. Dementia prevention and treatment: a narrative review. JAMA Intern Med 2024 May 1;184(5):563-572. [doi: 10.1001/jamainternmed.2023.8522] [Medline: 38436963]
5. Petermann-Rocha F, Balntzi V, Gray SR, et al. Global prevalence of sarcopenia and severe sarcopenia: a systematic review and meta-analysis. J Cachexia Sarcopenia Muscle 2022 Feb;13(1):86-99. [doi: 10.1002/jcsm.12783] [Medline: 34816624]
6. Nunes-Pinto M, Bandeira de Mello RG, Pinto MN, et al. Sarcopenia and the biological determinants of aging: a narrative review from a geroscience perspective. Ageing Res Rev 2025 Jan;103(102587):102587. [doi: 10.1016/j.arr.2024.102587] [Medline: 39571617]
7. Kirk B, Cawthon PM, Arai H, et al. The conceptual definition of sarcopenia: Delphi consensus from the Global Leadership Initiative in Sarcopenia (GLIS). Age Ageing 2024 Mar 1;53(3):afae052. [doi: 10.1093/ageing/afae052] [Medline: 38520141]
8. Sayer AA, Cooper R, Arai H, et al. Sarcopenia. Nat Rev Dis Primers 2024 Sep 19;10(1):68. [doi: 10.1038/s41572-024-00550-w] [Medline: 39300120]
9. Hu Y, Peng W, Ren R, Wang Y, Wang G. Sarcopenia and mild cognitive impairment among elderly adults: the first longitudinal evidence from CHARLS. J Cachexia Sarcopenia Muscle 2022 Dec;13(6):2944-2952. [doi: 10.1002/jcsm.13081] [Medline: 36058563]

XSL•FO
RenderX

10. Beeri MS, Leugrans SE, Delbono O, Bennett DA, Buchman AS. Sarcopenia is associated with incident Alzheimer's dementia, mild cognitive impairment, and cognitive decline. J Am Geriatr Soc 2021 Jul;69(7):1826-1835. [doi: 10.1111/jgs.17206] [Medline: 33954985]
11. Xue M, Jia X, Shi X, et al. Association between sarcopenia and cognitive trajectories among middle-aged and older adults in China: a nationally representative cohort study. J Nutr Health Aging 2023;27(4):243-250. [doi: 10.1007/s12603-023-1906-1] [Medline: 37170430]
12. Amini N, Ibn Hach M, Lapauw L, et al. Meta-analysis on the interrelationship between sarcopenia and mild cognitive impairment, Alzheimer's disease and other forms of dementia. J Cachexia Sarcopenia Muscle 2024 Aug;15(4):1240-1253. [doi: 10.1002/jcsm.13485] [Medline: 38715252]
13. Guo H, Cao J, He S, et al. Quantifying the enhancement of sarcopenic skeletal muscle preservation through a hybrid exercise program: randomized controlled trial. JMIR Aging 2024 Nov 15;7:e58175. [doi: 10.2196/58175] [Medline: 39621937]
14. Dorhout BG, Wezenbeek N, de Groot L, Grootswagers P. Web-based exercise and nutrition intervention to improve leg muscle strength and physical functioning in older adults: pre-post pilot study. JMIR Form Res 2025 Jan 30;9:e54392. [doi: 10.2196/54392] [Medline: 39882993]
15. Cadellans Arroniz A, Romero Rodríguez D, Zárate V, et al. Effectiveness of iso-inertial resistance training on muscle power in middle-older adults: randomized controlled trial. JMIR Aging 2025 Aug 21;8:e66414. [doi: 10.2196/66414] [Medline: 40840466]
16. Lee HJ, Choi JY, Hong D, Kim D, Min JY, Min KB. Sex differences in the association between sarcopenia and mild cognitive impairment in the older Korean population. BMC Geriatr 2023 May 29;23(1):332. [doi: 10.1186/s12877-023-03911-4] [Medline: 37248457]
17. Lin YC, Chen ZJ, Tung HH, et al. Association between possible sarcopenia and domain-specific cognitive impairment in middle-aged and older adults: insights from the Gan-Dau Healthy Longevity Plan. Exp Gerontol 2024 Sep;194:112487. [doi: 10.1016/j.exger.2024.112487] [Medline: 38879092]
18. Gao L, Jiang J, Yang M, Hao Q, Luo L, Dong B. Prevalence of sarcopenia and associated factors in Chinese community-dwelling elderly: comparison between rural and urban areas. J Am Med Dir Assoc 2015 Nov 1;16(11):1003. [doi: 10.1016/j.jamda.2015.07.020] [Medline: 26385304]
19. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). Int J Epidemiol 2014 Feb;43(1):61-68. [doi: 10.1093/ije/dys203] [Medline: 23243115]
20. Shore S, Li H, Zhang M, et al. Trajectory of cognitive function after incident heart failure. Circ Heart Fail 2025 Mar;18(3):e011837. [doi: 10.1161/CIRCHEARTFAILURE.124.011837] [Medline: 39963777]
21. Xu Z, Zhang D, Yip BHK, et al. Combined mind-body physical exercise, cognitive training, and nurse-led risk factor modification to enhance cognition among older adults with mild cognitive impairment in primary care: a three-arm randomised controlled trial. Lancet Healthy Longev 2025 Apr;6(4):100706. [doi: 10.1016/j.lanhl.2025.100706] [Medline: 40294622]
22. Chen LK, Woo J, Assantachai P, et al. Asian Working Group for Sarcopenia: 2019 consensus update on sarcopenia diagnosis and treatment. J Am Med Dir Assoc 2020 Mar;21(3):300-307. [doi: 10.1016/j.jamda.2019.12.012]
23. Wen X, Wang M, Jiang CM, Zhang YM. Anthropometric equation for estimation of appendicular skeletal muscle mass in Chinese adults. Asia Pac J Clin Nutr 2011;20(4):551-556. [Medline: 22094840]
24. Zeng Q, Zhao L, Zhong Q, An Z, Li S. Changes in sarcopenia and incident cardiovascular disease in prospective cohorts. BMC Med 2024 Dec 31;22(1):607. [doi: 10.1186/s12916-024-03841-x] [Medline: 39736721]
25. Yang M, Hu X, Wang H, Zhang L, Hao Q, Dong B. Sarcopenia predicts readmission and mortality in elderly patients in acute care wards: a prospective study. J Cachexia Sarcopenia Muscle 2017 Apr;8(2):251-258. [doi: 10.1002/jcsm.12163] [Medline: 27896949]
26. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994 Mar;49(2):M85-M94. [doi: 10.1093/geronj/49.2.m85] [Medline: 8126356]
27. Luo C, Han S, Shen X, Wu H, Zhou J, Liu B. Sex-specific associations between estimated glucose disposal rate and cognitive decline in middle-aged and older adults in China: a longitudinal cohort study. Front Aging Neurosci 2025;17(1544352):1544352. [doi: 10.3389/fnagi.2025.1544352] [Medline: 39975601]
28. Lin L, Cao B, Chen W, Li J, Zhang Y, Guo VY. Association of adverse childhood experiences and social isolation with later-life cognitive function among adults in China. JAMA Netw Open 2022 Nov 1;5(11):e2241714. [doi: 10.1001/jamanetworkopen.2022.41714] [Medline: 36367722]
29. Jia L, Du Y, Chu L, et al. Prevalence, risk factors, and management of dementia and mild cognitive impairment in adults aged 60 years or older in China: a cross-sectional study. Lancet Public Health 2020 Dec;5(12):e661-e671. [doi: 10.1016/S2468-2667(20)30185-7] [Medline: 33271079]
30. Bell T, Franz CE, Kremen WS. Persistence of pain and cognitive impairment in older adults. J Am Geriatr Soc 2022 Feb;70(2):449-458. [doi: 10.1111/jgs.17542] [Medline: 34741304]
31. Chen J, Yan L, Chu J, Wang X, Xu Z. Pain characteristics and progression to sarcopenia in Chinese middle-aged and older adults: a 4-year longitudinal study. J Gerontol A Biol Sci Med Sci 2024 May 1;79(5):glae080. [doi: 10.1093/gerona/glae080] [Medline: 38469680]

XSL•FO
RenderX

32. Li Q, Cen W, Yang T, Tao S. Association between depressive symptoms and sarcopenia among middle-aged and elderly individuals in China: the mediation effect of activities of daily living (ADL) disability. BMC Psychiatry 2024 Jun 10;24(1):38858698. [doi: 10.1186/s12888-024-05885-y]
33. Yin J, John A, Cadar D. Bidirectional associations of depressive symptoms and cognitive function over time. JAMA Netw Open 2024 Jun 3;7(6):e2416305. [doi: 10.1001/jamanetworkopen.2024.16305] [Medline: 38861255]
34. Zhu F, Guo J, Zheng W. Associations between sarcopenia and depression in middle-aged and older adults: the moderating effect of smoking. Sci Rep 2024 Jul 2;14(1):15187. [doi: 10.1038/s41598-024-65343-3] [Medline: 38956420]
35. Jin S, Chao J, Jin Q, et al. Longitudinal trajectories of cognitive function among Chinese middle-aged and older adults: the role of sarcopenia and depressive symptoms. Brain Sci 2025 Apr 17;15(4):408. [doi: 10.3390/brainsci15040408] [Medline: 40309867]
36. Zhang H, Hao M, Hu Z, et al. Social frailty and its association with cognitive trajectories in older adults: a prospective cohort study. Alzheimers Res Ther 2025 Feb 19;17(1):50. [doi: 10.1186/s13195-025-01687-w] [Medline: 39972471]
37. Lo JW, Crawford JD, Lipnicki DM, et al. Trajectory of cognitive decline before and after stroke in 14 population cohorts. JAMA Netw Open 2024 Oct 1;7(10):e2437133. [doi: 10.1001/jamanetworkopen.2024.37133] [Medline: 39356504]
38. Wang J, Jiao D, Zhao X, et al. Childhood loneliness and cognitive decline and dementia risk in middle-aged and older adults. JAMA Netw Open 2025 Sep 2;8(9):e2531493. [doi: 10.1001/jamanetworkopen.2025.31493] [Medline: 40938598]
39. Bloomberg M, Brocklebank L, Hamer M, Steptoe A. Joint associations of physical activity and sleep duration with cognitive ageing: longitudinal analysis of an English cohort study. Lancet Healthy Longev 2023 Jul;4(7):e345-e353. [doi: 10.1016/S2666-7568(23)00083-1] [Medline: 37421962]
40. Li C, Zhu Y, Ma Y, Hua R, Zhong B, Xie W. Association of cumulative blood pressure with cognitive decline, dementia, and mortality. J Am Coll Cardiol 2022 Apr;79(14):1321-1335. [doi: 10.1016/j.jacc.2022.01.045]
41. Li C, Ma Y, Hua R, et al. Dose-response relationship between long-term blood pressure variability and cognitive decline. Stroke 2021 Oct;52(10):3249-3257. [doi: 10.1161/STROKEAHA.120.033697] [Medline: 34167328]
42. Wu Y, Zhang YS, Kobayashi LC, Mayeda ER, Gross AL. How to assess cognitive decline when test administration changes across study waves? Harmonizing cognitive scores across waves in the China Health and Retirement Longitudinal Study. J Alzheimers Dis Rep 2024;8(1):1661-1669. [doi: 10.1177/25424823241302759] [Medline: 40034362]
43. Sun X, Ioannidis JPA, Agoritsas T, Alba AC, Guyatt G. How to use a subgroup analysis: users' guide to the medical literature. JAMA 2014;311(4):405-411. [doi: 10.1001/jama.2013.285063] [Medline: 24449319]
44. Du H, Yu M, Xue H, Lu X, Chang Y, Li Z. Association between sarcopenia and cognitive function in older Chinese adults: evidence from the China health and retirement longitudinal study. Front Public Health 2022;10(1078304):1078304. [doi: 10.3389/fpubh.2022.1078304] [Medline: 36703834]
45. Hernandez R, Jin H, Lee PJ, et al. Attrition from longitudinal ageing studies and performance across domains of cognitive functioning: an individual participant data meta-analysis. BMJ Open 2024 Mar 7;14(3):e079241. [doi: 10.1136/bmjopen-2023-079241] [Medline: 38453191]
46. Austin PC, Ibrahim M, Putter H. Accounting for competing risks in clinical research. JAMA 2024 Jun 25;331(24):2125-2126. [doi: 10.1001/jama.2024.4970] [Medline: 38809526]
47. Sha T, Zhang Y, Wei J, et al. Sarcopenia and risk of cognitive impairment: cohort study and mendelian randomization analysis. JMIR Aging 2025 Jun 11;8:e66031. [doi: 10.2196/66031] [Medline: 40499039]
48. Tessier AJ, Wing SS, Rahme E, Morais JA, Chevalier S. Association of low muscle mass with cognitive function during a 3-year follow-up among adults aged 65 to 86 years in the Canadian longitudinal study on aging. JAMA Netw Open 2022 Jul 1;5(7):e2219926. [doi: 10.1001/jamanetworkopen.2022.19926] [Medline: 35796211]
49. Ji W, Lee D, Kim M, et al. Efficacy of a combined exercise and nutrition intervention study for outpatients with possible sarcopenia in community-based primary care clinics (ENdSarC): study protocol for a multicenter single-blinded randomized controlled trial. BMC Geriatr 2024 Oct 23;24(1):861. [doi: 10.1186/s12877-024-05434-y] [Medline: 39443873]
50. Arosio B, Calvani R, Ferri E, et al. Sarcopenia and cognitive decline in older adults: targeting the muscle-brain axis. Nutrients 2023 Apr 12;15(8):1853. [doi: 10.3390/nu15081853] [Medline: 37111070]
51. Severinsen MCK, Pedersen BK. Muscle-organ crosstalk: the emerging roles of myokines. Endocr Rev 2020 Aug 1;41(4):594-609. [doi: 10.1210/endrev/bnaa016] [Medline: 32393961]
52. Pedersen BK. Physical activity and muscle-brain crosstalk. Nat Rev Endocrinol 2019 Jul;15(7):383-392. [doi: 10.1038/s41574-019-0174-x] [Medline: 30837717]
53. Kim S, Choi JY, Moon S, Park DH, Kwak HB, Kang JH. Roles of myokines in exercise-induced improvement of neuropsychiatric function. Pflugers Arch - Eur J Physiol 2019 Mar;471(3):491-505. [doi: 10.1007/s00424-019-02253-8]
54. Sui SX, Williams LJ, Holloway-Kew KL, Hyde NK, Pasco JA. Skeletal muscle health and cognitive function: a narrative review. Int J Mol Sci 2020 Dec 29;22(1):255. [doi: 10.3390/ijms22010255] [Medline: 33383820]
55. Jimenez-Gutierrez GE, Martínez-Gómez LE, Martínez-Armenta C, Pineda C, Martínez-Nava GA, Lopez-Reyes A. Molecular mechanisms of inflammation in sarcopenia: diagnosis and therapeutic update. Cells 2022 Aug 1;11(15):2359. [doi: 10.3390/cells11152359] [Medline: 35954203]
56. Zhang H, Qi G, Wang K, et al. Oxidative stress: roles in skeletal muscle atrophy. Biochem Pharmacol 2023 Aug;214:115664. [doi: 10.1016/j.bcp.2023.115664] [Medline: 37331636]

XSL•FO
RenderX

57. Mekhora C, Lamport DJ, Spencer JPE. An overview of the relationship between inflammation and cognitive function in humans, molecular pathways and the impact of nutraceuticals. Neurochem Int 2024 Dec;181:105900. [doi: 10.1016/j.neuint.2024.105900] [Medline: 39522696]
58. Lyra E Silva NM, Gonçalves RA, Pascoal TA, et al. Pro-inflammatory interleukin-6 signaling links cognitive impairments and peripheral metabolic alterations in Alzheimer's disease. Transl Psychiatry 2021 Apr 28;11(1):251. [doi: 10.1038/s41398-021-01349-z] [Medline: 33911072]
59. Tseng PT, Zeng BS, Suen MW, et al. Efficacy and acceptability of anti-inflammatory eicosapentaenoic acid for cognitive function in Alzheimer's dementia: a network meta-analysis of randomized, placebo-controlled trials with omega-3 fatty acids and FDA-approved pharmacotherapy. Brain Behav Immun 2023 Jul;111:352-364. [doi: 10.1016/j.bbi.2023.04.017] [Medline: 37150266]
60. Song L, Li H, Fu X, Cen M, Wu J. Association of the oxidative balance score and cognitive function and the mediating role of oxidative stress: evidence from the National Health and Nutrition Examination Survey (NHANES) 2011-2014. J Nutr 2023 Jul;153(7):1974-1983. [doi: 10.1016/j.tjnut.2023.05.014] [Medline: 37187352]
61. Yan X, Li H, Xie R, et al. Relationships between sarcopenia, nutrient intake, and gut microbiota in Chinese community-dwelling older women. Arch Gerontol Geriatr 2023 Oct;113(105063):105063. [doi: 10.1016/j.archger.2023.105063] [Medline: 37216814]
62. Chen C, Liao J, Xia Y, et al. Gut microbiota regulate Alzheimer's disease pathologies and cognitive disorders via PUFA-associated neuroinflammation. Gut 2022 Nov;71(11):2233-2252. [doi: 10.1136/gutjnl-2021-326269] [Medline: 35017199]
63. Shen Y, Shi Q, Nong K, et al. Exercise for sarcopenia in older people: a systematic review and network meta-analysis. J Cachexia Sarcopenia Muscle 2023 Jun;14(3):1199-1211. [doi: 10.1002/jcsm.13225] [Medline: 37057640]
64. Yoon DH, Lee JY, Song W. Effects of resistance exercise training on cognitive function and physical performance in cognitive frailty: a randomized controlled trial. J Nutr Health Aging 2018;22(8):944-951. [doi: 10.1007/s12603-018-1090-9] [Medline: 30272098]
65. Livingston G, Huntley J, Liu KY, et al. Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. Lancet 2024 Aug 10;404(10452):572-628. [doi: 10.1016/S0140-6736(24)01296-0] [Medline: 39096926]
66. China Health and Retirement Longitudinal Study. URL: http://charls.pku.edu.cn [accessed 2025-11-03]

Abbreviations

BMI: body mass index
CES-D-10: 10-item Center for Epidemiological Studies Depression Scale
CHARLS: China Health and Retirement Longitudinal Study
DXA: X-ray absorptiometry
LMM: linear mixed-effects model

Edited by PH Chau; submitted 29.05.25; peer-reviewed by Y Dong, Z Xu; revised version received 12.10.25; accepted 26.10.25; published 16.12.25.

Please cite as:
Luo C, Wu H, Shen X, Han S, Lin L, Liu B
Transitions in Sarcopenia Status and Cognitive Trajectories Among Middle-Aged and Older Adults in China: Longitudinal Cohort Study
JMIR Aging 2025;8:e78277
URL: https://aging.jmir.org/2025/1/e78277
doi:10.2196/78277

© Chun Luo, Hao Wu, Xiaoying Shen, Shuang Han, Lv Lin, Bingyang Liu. Originally published in JMIR Aging (https://aging.jmir.org), 16.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Machine Learning Models for Frailty Classification of Older Adults in Northern Thailand: Model Development and Validation Study

Natthanaphop Isaradech[1,2*], MD; Wachiranun Sirikul[1,3,4*], MD, MSc; Nida Buawangpong[5], MD; Penprapa Siviroj[1], DrPH; Amornphat Kitro[1,2], MD, MCTM

[1]Department of Community Medicine, Faculty of Medicine, Chiang Mai University, 110, Intrawarorot Road, Meaung, Thailand

[2]Environmental and Occupational Medicine Excellence Center, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[3]Center of Data Analytics and Knowledge Synthesis for Health Care, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[4]Department of Biomedical Informatics and Clinical Epidemiology, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[5]Department of Family Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[*]these authors contributed equally

Corresponding Author:
Wachiranun Sirikul, MD, MSc
Department of Community Medicine, Faculty of Medicine, Chiang Mai University, 110, Intrawarorot Road, Meaung, Thailand

Related Article:

This is a corrected version. See correction statement: https://aging.jmir.org/2025/1/e75690

Abstract

Background: Frailty is defined as a clinical state of increased vulnerability due to the age-associated decline of an individual's physical function resulting in increased morbidity and mortality when exposed to acute stressors. Early identification and management can reverse individuals with frailty to being robust once more. However, we found no integration of machine learning (ML) tools and frailty screening and surveillance studies in Thailand despite the abundance of evidence of frailty assessment using ML globally and in Asia.

Objective: We propose an approach for early diagnosis of frailty in community-dwelling older individuals in Thailand using an ML model generated from individual characteristics and anthropometric data.

Methods: Datasets including 2692 community-dwelling Thai older adults in Lampang from 2016 and 2017 were used for model development and internal validation. The derived models were externally validated with a dataset of community-dwelling older adults in Chiang Mai from 2021. The ML algorithms implemented in this study include the k-nearest neighbors algorithm, random forest ML algorithms, multilayer perceptron artificial neural network, logistic regression models, gradient boosting classifier, and linear support vector machine classifier.

Results: Logistic regression showed the best overall discrimination performance with a mean area under the receiver operating characteristic curve of 0.81 (95% CI 0.75 - 0.86) in the internal validation dataset and 0.75 (95% CI 0.71 - 0.78) in the external validation dataset. The model was also well-calibrated to the expected probability of the external validation dataset.

Conclusions: Our findings showed that our models have the potential to be utilized as a screening tool using simple, accessible demographic and explainable clinical variables in Thai community-dwelling older persons to identify individuals with frailty who require early intervention to become physically robust.

(JMIR Aging 2025;8:e62942) doi:10.2196/62942

KEYWORDS

aged care; gerontology; geriatric; old; aging; clinical decision support; delivering health information and knowledge to the public; diagnostic systems; digital health; epidemiology; surveillance; diagnosis; frailty; machine learning; prediction; predictive; AI; artificial intelligence; Thailand; community dwelling; health care intervention; patient care

Introduction

The world population is moving toward an aging society. As health care technology improves, people are expected to live longer and healthier [1]. According to the World Health Organization, the population aged ≥60 years will increase from 1 billion in 2020 to 2.1 billion in 2050 and the number of people aged ≥80 years will reach 426 million in 2050 [2]. Researchers predicted that the proportion of people in Thailand aged ≥60

XSL•FO
RenderX

years would be more than 20% of the population in 2025 and more than 30% in 2031 [3,4].

The prevalence of frailty is high among older adults aged ≥60 years [5]. Global frailty prevalence ranges from approximately 10% to 12% [6-11]. The percentage varies by age, gender, and frailty classification tool. In Thailand, frailty prevalence was 22.1%, which is twice the global frailty prevalence, according to the Thai National Health Examination Survey cohort in 2018. Specifically, Thailand's northern region frailty prevalence was found to be 15% [12,13]. This creates concerns about the increasing aging population in Thailand.

Frailty is defined as a clinical state of increased vulnerability due to the age-associated decline of an individual's body resulting in increased morbidity and mortality when exposed to everyday or acute stressors [14,15]. This clinical syndrome is associated with decreased quality of life [7], slow gait speed [16], more depressive symptoms, higher BMI, reduced cognitive function [17], decreased strength [12], and increased risk of fall, hospital re-admission, and all-cause mortality in the older adult population [13,18-20]. Frailty has become a crucial research topic because this clinical syndrome can be reversed. Studies have shown that early detection and intervention can revert individuals from a frail to a fit state [21-23].

However, incorporating frailty evaluation into clinical practice in a primary care context is challenging due to increased administrative tasks, time limitations, and a lack of diagnostic effectiveness [15,24,25]. In the age of technology, health informatics has become an important role in health care research [26,27]. Several studies have applied information technology to frailty detection in primary care settings using machine learning models and artificial intelligence [28-30]. A study from Canada showcased an efficient frailty identification tool using the XGBoost machine learning model. Features used for the model were medication, medical billing codes, and other primary care clinical data [31]. Another example showed the development of a predictive machine learning model for frailty conditions based on a database of demographic data and clinical characteristics [32]. In China, a study simplified the Frailty Index assessment for older individuals using machine learning techniques and showed that logistic regression was the best performing and most interpretable model, with a mean area under the receiver operating characteristic curve (AUC) of 0.974 in the internal validation dataset [33]. Another study in China also developed and validated models using data from 6997 older adult participants to predict frailty risk, with random forest (RF) and logistic regression (LR) achieving AUC values of 0.77 and 0.76, respectively [34].

Despite the abundance of evidence of conducting frailty assessments using machine learning globally and in Asia, we have found no integration of data science tools and frailty screening and surveillance studies in Thailand. Most studies focused on the risk factors and their association with frailty syndrome but failed to show application in real-world settings [12,13,35,36]. This leads to our research question, "Can machine learning models predict frailty in community-dwelling older adults in Northern Thailand?" Therefore, we propose an approach to frailty detection using a machine learning model in the community-dwelling population from Lampang and Chiang Mai, Thailand, to effectively screen frailty status among Thai community-dwelling older adults and to help decrease clinicians' burden of work.

Methods

Source of Data

Development and Internal Validation Datasets

The datasets were derived from a cross-sectional study carried out in Lampang in 2016 and 2017; it is a northern Thai province with one of the highest aging indexes [37]. This study included older adults aged ≥60 years. Those with dementia (as determined by the Thai Mental State Examination), blindness, deafness, bedridden status, disabilities, or severe acute diseases were excluded.

To represent urban (8 villages), semiurban (8 villages), and rural (8 villages) communities, 24 villages in 3 districts were chosen. The records of the primary care unit were used to compile lists of community-dwelling older adults. The details of the data collection in this study are published elsewhere [36]. A total of 2228 older adults from this study were derived for model development and internal validation.

External Validation Dataset

The datasets used for external validation were derived from a cross-sectional study in Kuamung (suburban), Sankampang, Chiang Mai, Thailand, in 2021. The participants were included and excluded from the study with the same criteria as described in the internal validation datasets. A total of 464 older adults from this study were used for external validation of the derived models.

Predictors

Characteristics and Demographics

Assessments were conducted through a questionnaire at the participant's residence. The questionnaire included questions about sociodemographic information (age, household living arrangement, gender, and education level), self-reported medical diagnoses (such as hypertension, diabetes mellitus, and heart disease), level of physical activity per week, and exhaustion.

Anthropometric Variables

Anthropometric measures included BMI, waist circumference, and calf circumference (CC). A handheld dynamometer (Takei TKK5001) was used to assess handgrip strength. Height and calf circumference were measured with standard tape (Tajima brand, PIT-20BL model), and weight was measured with a calibrated weighing scale (Shaper Disney). The CC was measured over the unclothed area at the maximum diameter on the left leg. The tape was wrapped snugly around the calf and measured to the nearest 0.1 cm. All measurements were administered by 10 qualified field investigators, and the measurements were standardized by the principal investigator.

Outcome Variable

Frailty was evaluated based on Fried's phenotype [38], which includes the following five criteria:

XSL•FO
RenderX

1. Unintentional weight loss: The participant will be asked "In the last ten years, have you lost more than 10 pounds intentionally (not due to diet or exercise)?" If yes, then frail for weight loss criterion. At follow-up, weight loss was calculated as: (weight in previous year – current measured weight)/(weight in previous year)=K. If K≥0.05 and the subject did not report that he/she was trying to lose weight (ie, unintentional weight loss of at least 5% of previous year's body weight), then frail for weight loss=yes.
2. Exhaustion, assessed with self-report using Fried's method of assessment. The participant was first asked to self-assess whether she/he felt exhausted. If yes, she/he would be asked to rate the severity of the exhaustion. Ratings of 2-4 suggested a positive assessment.
3. Physical activity, based on the short version of the Minnesota Leisure time Activity questionnaire, which asked about walking, chores (moderately strenuous), mowing the lawn, raking, gardening, hiking, jogging, biking, exercise cycling, golf, single tennis, doubles tennis, racquetball, and calisthenics. For men, those with <383 kcals of physical activity per week were frail. For women, those with <270 kcals of physical activity per week were frail.
4. Walk time, stratified by gender and height (gender-specific cutoff for medium height).
    - Height ≤173 cm and walk time ≥7 seconds for men.
    - Height >173 cm and walk time ≥6 seconds for men.
    - Height ≤159 cm and walk time ≥7 seconds for women.
    - Height >159 cm and walk time ≥6 seconds for women.
5. Grip strength, stratified by gender and BMI quartiles.
    a. For men:
        - BMI ≤24 and grip strength ≤29 for men
        - BMI 24.1 - 26 and grip strength ≤30
        - BMI 26.1 - 28 and grip strength ≤30
        - BMI >28 and grip strength ≤32
    b. For women:
        - BMI ≤23 and grip strength ≤17
        - BMI 23.1-26 and grip strength ≤17.3
        - BMI 26.1-29 and grip strength ≤18
        - BMI <29 and grip strength ≤21

Older adults with 3 or more phenotypes were considered to have physical frailty, and those who had 1 or 2 phenotypes were classified as prefrailty. According to the study design of the derived datasets, the outcome was concurrently measured with the predictors.

Sample Size Calculation

The sample size for model development was calculated using the *pmsampsize* package via STATA (version 16; StataCorp). A minimum sample size required 16.41 events per predictor, which was 1519 patients with 263 frailty cases based on a maximum candidate predictor of 16, frailty prevalence of 17.29% in the dataset for model development, small overfitting defined as expected shrinkage of predictor effects by 10% or less, a small difference in the developed model's apparent and optimism-adjusted values of 0.15 ($R^2_{Nagelkerke}$), as suggested by Riley et al [39]. For the external validation, the sample size was calculated using the *pmvalsampsize* package via STATA. From the available data of 464 patients with 192 frailty cases, it was able to precisely estimate confidence interval widths of observed/expected statistic of 0.22, calibration slope of 0.60, and concordance statistic of 0.10 [40].

Statistical Analysis Methods and Synthesis of the Results

Data exploration was performed using descriptive statistics to determine the data quantity, quality, and distribution. A frequency and a percentage were used to describe categorical variables. For continuous variables, a mean with SD was used for parametrically distributed data, and a median with an IQR was used for nonparametrically distributed data. The comparison of characteristics between the frail and nonfrail populations was performed using the chi-square test for categorical variables, the independent *t* test for parametric variables, and the rank-sum test for nonparametric variables. The methods for feature selection, model development, and validation are described in the section "Model Development and Validation". Estimates of model discrimination and optimism are reported as the mean AUC with 95% CI across all repetitions of cross-validation for the internal validation and the AUC with 95% CI from 1000 bootstrapping samples for the external validation. To further explain model performance, we also created model calibration plots and calculated secondary metrics of prediction models, including the confusion matrix and specificity, sensitivity, and predictive values.

Missing Data and Imputation

There was only missing data in participants' age in the internal validation dataset (4/2228, 0.18%); therefore, a complete case analysis was performed on the dataset. The external validation dataset contained missing data, including age (2/464, 0.43%), BMI (11/464, 2.37%), waist and calf circumference (3/464, 0.64%), handgrip strength (2/464, 0.43%), and frailty status (4/464, 0.86%). We performed multiple imputations of missing data, except for frailty status (the outcome of interest), using the predictive mean matching imputation with 5 nearest neighbors via the *KNNImputer* from the *Scikit-Learn library 1.1.2*. Four participant records that did not have frailty status were removed (list-wise deletion).

Model Development and Validation

Feature Selection

The classification models were developed using the variables from the derived datasets. The variables for model development were selected using both a data-driven method and domain expertise. For a data-driven selection, a multivariable logistic regression with stepwise backward elimination was performed to determine statistically significant variables (*P* value less than .20). For the *P* value threshold of .20, we used a higher threshold to give priority to clinical reasoning in selecting variables by domain expertise and associated factors from the previous studies, along with statistical significance, which would allow more important variables to be entered into the model.

XSL•FO
RenderX

Rebalancing Data Strategy

We resolved unbalanced classification for the dataset's minority class by using the Synthetic Minority Oversampling Technique (SMOTE) to enhance model decision boundaries via the *imblearn.over_sampling.SMOTE* package [41,42].

Model Development and Internal Validation

The model development and validation were performed using Python (version 3.9; Python Software Foundation). The machine learning algorithms implemented in this study include the k-nearest neighbors (KNN) algorithm, RF machine learning algorithms, multilayer perceptron artificial neural network (MLP), gradient boosting classifier (GBC), linear support vector machine classifier (SVM), and LR models via the *Scikit-Learn library 1.1.2*. The hyperparameters were determined by using a grid search via the *GridSearch CV* package with 10-fold cross-validation on the derived dataset to determine the parameters of each model that led to the best discriminative performance. For 10-fold cross-validation, the derived dataset was divided into 10 folds of data and repeated 10 times to perform model training and testing. For each iteration, 9 folds of data were used to train the model and then it was tested with the remaining fold to ensure that almost all the derived data were used to train and test the models. The discriminative performance of the derived models was assessed by computing a confusion matrix and sensitivity, specificity, and predictive values, as well as AUC with a 95% CI.

The model calibration was evaluated using the calibration plot, which indicated the congruence between the observed proportion of the actual probability of outcome and the mean predicted probability (MPP) from the derived models.

External Validation

The derived models were validated again with the external validation dataset to determine the model optimism and calibration. The discriminative performance and the model optimism were re-evaluated and presented by discriminative performance matrices and a 95% CI of AUC from the 1000 bootstrapping samples, respectively. The model calibration using the external validation dataset was re-evaluated using the calibration plot.

Ethical Considerations

This study complies with the research with exemption category and has been certified by the Research Ethics Committee of the Faculty of Medicine, Chiang Mai University (study code: COM-2565 - 09159, number EXEMPTION 9159/2022). We requested a waiver of informed consent because this study is a retrospective analysis that exclusively utilizes anonymized secondary data from our research database, without collecting any additional information from medical records or other sources. All participants have previously provided informed consent for the primary data collection as described elsewhere [35,36,43]. All personal patient data were anonymized by removing citizen ID numbers, hospital numbers, addresses, and contact information from the dataset. The investigator cannot trace or identify individuals. The study results were reported in accordance with the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis plus Artificial Intelligence (TRIPOD+AI) statement [44].

Results

Baseline Characteristics of Participants in Each Dataset

Baseline characteristics of participants in this study are shown in Table 1. The participants in the internal validation dataset had a mean age of 71.0 years. Most of the participants were male and had finished primary school. Among 2228 old adults, 2160 lived with either their spouse, relative, or children while the others were living alone. The average BMI of the participants was 32.6 (SD 7.4) kg/m^2. The prevalence of hypertension, dyslipidemia, type 2 diabetes mellitus, and heart disease were 45.26%, 19.67%, 16.21%, and 4.31%, respectively. The averages of waist circumference, calf circumference, handgrip strength, and walk time were 83.40 (SD 11.23) cm, 32.40 (SD 4.35) cm, 32.40 (SD 6.68) kg, and 6.42 (SD 2.01) minutes, respectively. Overall, 9.79% of the participants were exhausted and 16.3% had physical activity higher than 150 minutes per week. The prevalence of frailty was 17.3% (n=385).

XSL•FO
RenderX

Table . The characteristics of participants in the development and internal validation datasets and the external validation dataset.

Characteristics	Development and internal validation datasets (Lampang, 2016 - 2017; N=2228)	External validation datasets (Chiang Mai, 2021; N=464)	*P* value
Age (years), mean (SD)	70.96 (7.49)	70.68 (5.58)	.45
Gender, n (%)			
Male	1569 (70.45)	193 (41.59)	<.001
Female	658 (29.55)	271 (58.41)	
Household living arrangement, n (%)			
Living alone	160 (7.18)	42 (9.05)	<.001
Living with spouse	1177 (52.85)	9 (1.94)	
Living with children	823 (36.96)	283 (60.99)	
Living with relatives or others	67 (3.01)	130 (28.02)	
BMI (kg/m^2), mean (SD)	32.64 (7.40)	22.73 (3.89)	.001
Education, n (%)			
No education	192 (8.62)	12 (2.61)	<.001
Primary school	1745 (78.36)	398 (86.52)	
Secondary school or higher	290 (13.02)	50 (10.87)	
Underlying diseases, n (%)			
Hypertension	1008 (45.26)	237 (51.08)	.02
Dyslipidemia	438 (19.67)	79 (17.03)	.19
Type 2 diabetes mellitus	361 (16.21)	78 (16.81)	.75
Heart diseases	96 (4.31)	18 (3.88)	.68
Anthropometric variables, mean (SD)			
Waist circumference (cm)	83.40 (11.23)	81.16 (10.80)	<.001
Calf circumference (cm)	32.40 (4.35)	32.85 (4.55)	.04
Walk time (min), mean (SD)	6.42 (2.01)	8.50 (5.38)	<.001
Exhaustion, n (%)	218 (9.79)	189 (40.74)	<.001
Adequate level of physical activity as defined by the World Health Organization, n (%)	363 (16.30)	104 (22.42)	<.001
Grip strength (kg), mean (SD)	32.40 (6.68)	19.87 (7.29)	<.001
Frailty, n (%)	385 (17.29)	192 (41.74)	<.001

In the external validation dataset, the participants had a mean age of 70.68 years. Most of the participants were male and had finished primary school. Among 464 persons, 422 lived with either their spouse, relative, or children, while the others were living alone. The average BMI of the participants was 22.73 (SD 3.89) kg/m^2. The prevalence of hypertension, dyslipidemia, type 2 diabetes mellitus, and heart disease were 51.08%, 17.03%, 16.81%, and 3.88%, respectively. The averages of waist circumference, calf circumference, grip strength, and walk time were 81.16 (SD 10.80) cm, 32.85 (SD 4.55) cm, 19.87 (SD 7.29) kg, and 8.50 (SD 5.38) minutes, respectively. The prevalence of participants with exhausted state was 40.74%, and 22.42% of the participants participated in physical activity more than 150 minutes per week.

Comparing the variables between the 2 datasets, we found that gender, household living arrangement, BMI, education, waist circumference, walk time, exhaustion, grip strength, and level of physical activity were significantly different.

Model Development

We used a dataset from Lampang (2016 - 2017; N=2228) for model development. The association between candidate predictors and frailty by univariate analysis is reported in Table 2. Feature selection for model development was selected by a backward elimination approach via a multivariable logistic regression and expert judgment. We chose the following features as model predictors: age, gender, status, underlying diseases (hypertension and dyslipidemia), BMI, waist and calf circumference, and level of exhaustion. Finally, the derived data included 385 participants with frailty and 1842 participants

XSL•FO
RenderX

without frailty, and all candidate predictors, as shown in Table 2, were used in model development.

Table . The comparison of participants' characteristics and their associations with frailty in the development and internal validation datasets.

Characteristics	Adjusted odds ratio – full model (95% CI)	*P* value	Adjusted odds ratio – reduced model (95% CI)	*P* value
Age (years)	1.04 (1.02 - 1.07)	<.001	1.09 (1.06 - 1.11)	<.001
Gender				
Male	Reference		Reference	
Female	1.93 (1.26 - 2.94)	.002	0.90 (0.63-1.27)	.55
Status				
Living alone	Reference			
Living with others	0.92 (0.46 - 1.86)	.82	0.87 (0.45 - 1.68)	.68
BMI (kg/m^2)	0.92 (0.90 - 0.95)	<.001	0.91 (0.88 - 0.93)	<.001
Education				
No education	Reference			
Primary school	1.00 (0.59 - 1.70)	.99		
Secondary school or higher	1.34 (0.65 - 2.77)	.42		
Underlying diseases				
Hypertension	1.47 (1.03 - 2.10)	.03	1.27 (0.92 - 1.75)	.15
Dyslipidemia	1.22 (0.78 - 1.91)	.38	1.45 (0.97 - 2.17)	.07
Type 2 diabetes mellitus	0.84 (0.54 - 1.31)	.44		
Heart disease	1.03 (0.45 - 2.34)	.95		
Anthropometric variables				
Waist circumference (cm)	1.02 (1.00 - 1.04)	.046	1.02 (1.00 - 1.04)	.02
Calf circumference (cm)	1.05 (1.01 - 1.09)	.01	1.04 (1.01 - 1.08)	.02
Walk time (min)	1.89 (1.71 - 2.10)	<.001	1.95 (1.78 - 2.13)	<.001
Exhaustion	20.39 (12.61 - 32.94)	<.001	28.23 (18.22 - 43.73)	<.001
Adequate level of physical activity as defined by the World Health Organization	4.49 (3.06 - 6.59)	<.001		
Grip strength	0.83 (0.79 - 0.85)	<.001		

However, the model appeared to be poor in frailty prediction performance and imbalanced as the classifiers intended to classify only the majority class (accuracy paradox). Therefore, a rebalancing strategy by SMOTE was applied to counter this problem. The oversampling data were generated and rebalanced the minority group in a 1:1 ratio. The details of the final model's hyperparameters using *GridSearch CV* are presented in Table S1 in Multimedia Appendix 1.

Adjusted odds ratios (aORs) of the full model were obtained from a multivariable logistic regression with all features. aORs of the reduced model were obtained from a multivariable logistic regression with stepwise backward elimination (P<.10) and feature selection based on the domain expertise.

Discrimination Performance of Internal Validated Models

We evaluated the model's performance by 10-fold cross-validation. The discrimination performances of the models are presented in Table 3 and Figure 1. The overall discrimination performance was presented by mean AUC 10-fold cross-validation. The KNN model achieved the highest overall performance with a mean AUC of 0.85 (95% CI 0.82 - 0.88), followed by MLP (AUC 0.81, 95% CI 0.72 - 0.89), LR (AUC 0.81, 95% CI 0.75 - 0.86), and SVM (AUC 0.75, 95% CI 0.75 - 0.86), respectively. In addition, the KNN model had the highest sensitivity and specificity (89% and 76%, respectively). RF had the lowest discrimination performance with almost all metrics. Other metrics that were not affected by data rebalancing, like positive predictive value (PPV) and negative predictive value (NPV), were also used to express the performance of the models. The best performances for both

XSL•FO
RenderX

PPV and NPV were found in the KNN model (0.79 and 0.87, respectively).

Table . Discrimination and optimism of internal validated models with 95% CIs.

Models	AUC[a], mean (95% CI)	Predictive values		Sensitivity	Specificity
		Positive	Negative		
LR[b]	0.81 (0.75 - 0.86)	0.74 (0.68 - 0.80)	0.74 (0.69 - 0.78)	0.74 (0.69 - 0.79)	0.73 (0.65 - 0.81)
KNN[c]	0.85 (0.82 - 0.88)	0.79 (0.76 - 0.82)	0.88 (0.86 - 0.90)	0.89 (0.87 - 0.91)	0.76 (0.71 - 0.81)
RF[d]	0.70 (0.60 - 0.79)	0.65 (0.58 - 0.73)	0.70 (0.63 - 0.76)	0.79 (0.74 - 0.83)	0.55 (0.39 - 0.71)
MLP[e]	0.81 (0.72 - 0.89)	0.73 (0.64 - 0.81)	0.75 (0.68 - 0.81)	0.78 (0.71 - 0.84)	0.69 (0.57 - 0.81)
GBC[f]	0.74 (0.72 - 0.89)	0.66 (0.59 - 0.74)	0.70 (0.63 - 0.78)	0.78 (0.73 - 0.83)	0.58 (0.43 - 0.73)
SVM[g]	0.75 (0.75 - 0.86)	0.74 (0.68 - 0.80)	0.73 (0.69 - 0.78)	0.74 (0.69 - 0.79)	0.73 (0.65 - 0.81)

[a]AUC: area under the receiver operating characteristic curve.

[b]LR: logistic regression.

[c]KNN: k-nearest neighbors.

[d]RF: random forest.

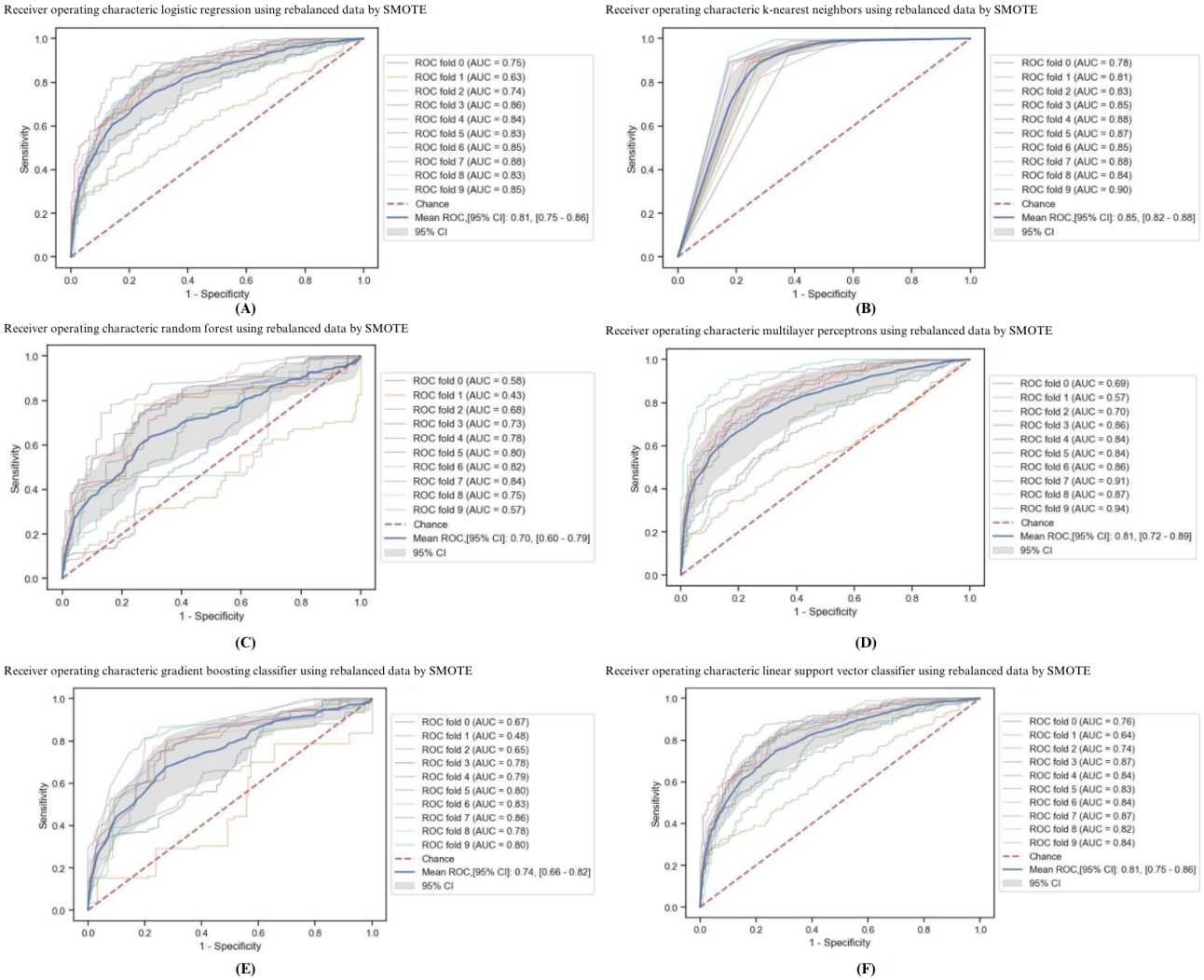
[e]MLP: multilayer perceptron artificial neural network.

[f]GBC: gradient boosting classifier.

[g]SVM: linear support vector machine classifier.

XSL•FO
RenderX

Figure 1. Receiver operating characteristic curves from 10-fold cross-validation of the rebalanced learning classifiers by SMOTE: (A) logistic regression model: mean AUC 0.81 (95% CI 0.75 - 0.86); (B) k-nearest neighbors model: mean AUC 0.85 (95% CI 0.82 - 0.88); (C) random forest model: mean AUC 0.70 (95% CI 0.60 - 0.79); (D) multilayer perceptron model: mean AUC 0.81 (95% CI 0.72 - 0.89); (E) gradient boosting classifier model: mean AUC 0.74 (95% CI 0.66 - 0.82); (F) linear support vector machine classifier model: mean AUC 0.75 (95% CI 0.75 - 0.86). ROC: receiver operating characteristic curve; SMOTE: synthetic minority oversampling technique.



Discrimination Performance of External Validated Models

We validated our trained model with an external validation dataset to evaluate the bias and variance of our trained models. SMOTE was not applied to the dataset since it was already well-balanced between participants with and without frailty. The performance of machine learning models validated by the external validation dataset is shown in Table 4.

XSL•FO
RenderX

Table . Discrimination and optimism of external validated models.

Models and model prediction		True label (frailty/non-frailty)	AUC[a], mean (95% CI)	Predictive values		Sensitivity	Specificity
				Positive	Negative		
Logistic regression							
	Frailty	140/65	0.75 (0.71 - 0.78)	0.68	0.80	0.73	0.76
	Nonfrailty	52/207					
K-nearest neighbors							
	Frailty	41/35	0.54 (0.51 - 0.57)	0.54	0.61	0.21	0.87
	Nonfrailty	151/237					
Random forest							
	Frailty	137/63	0.75 (0.71 - 0.78)	0.69	0.79	0.71	0.77
	Nonfrailty	55/209					
Multilayer perceptron artificial neural network							
	Frailty	98/38	0.68 (0.65 - 0.72)	0.72	0.71	0.51	0.86
	Nonfrailty	94/234					
Gradient boosting classifier							
	Frailty	89/31	0.69 (0.65 - 0.72)	0.74	0.70	0.46	0.89
	Nonfrailty	103/241					
Linear support vector machine classifier							
	Frailty	131/60	0.73 (0.70 - 0.77)	0.69	0.78	0.68	0.78
	Nonfrailty	61/212					

[a]AUC: area under the receiver operating characteristic curve.

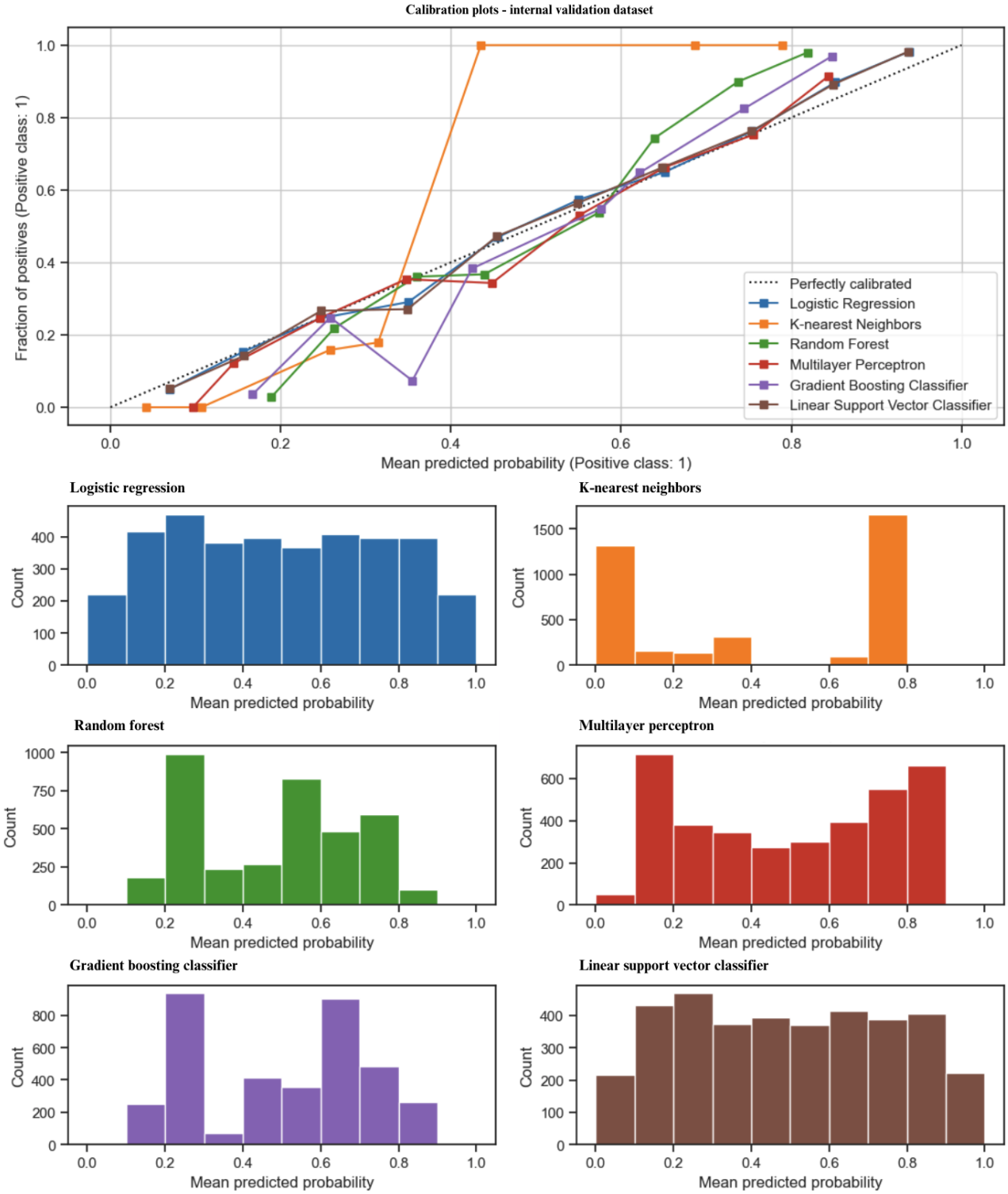
The overall discrimination performance was presented by mean AUC and 95% CI via 1000-fold bootstrapping. The LR and RF model achieved the highest overall performance with mean AUC 0.75 (95% CI 0.71 - 0.78), followed by SVM (mean AUC 0.73, 95% CI 0.70 - 0.77), and GBC (mean AUC 0.69, 95% CI 0.65 - 0.72), respectively. The LR model had the highest sensitivity (73%) and the MLP model had the highest specificity (86%). KNN had the lowest discrimination performance with all metrics. Other metrics that were not affected by data rebalancing, like PPV and NPV, were also used to express the performance of the models. The best performance for PPV was found in the GBC model (0.74); for NPV, it was found in the LR model (0.80).

Model Calibration

The model calibration was visualized with the calibration plot, which compared the expected probability of frailty, and the mean 10-fold cross-validation predicted the probability of each model. In the internal validation dataset (Figure 2), the LR, SVM, and MLP were well-calibrated with the expected probability of the data. The distribution of predicted probabilities of the models was well-balanced between 0.00 and 1.00. The KNN classifier aligned well with the MPP between 0.00 and 0.40 but it overestimated the high predicted probability. In contrast, the GBC underestimated MPP between 0.20 and 0.40 but aligned well with the rest of the expected probability. These results suggest that LR, SVM, and MLP are the most reliable for balanced predictions, while KNN, GBC, and RF require careful consideration depending on the probability range of interest.

XSL•FO
RenderX

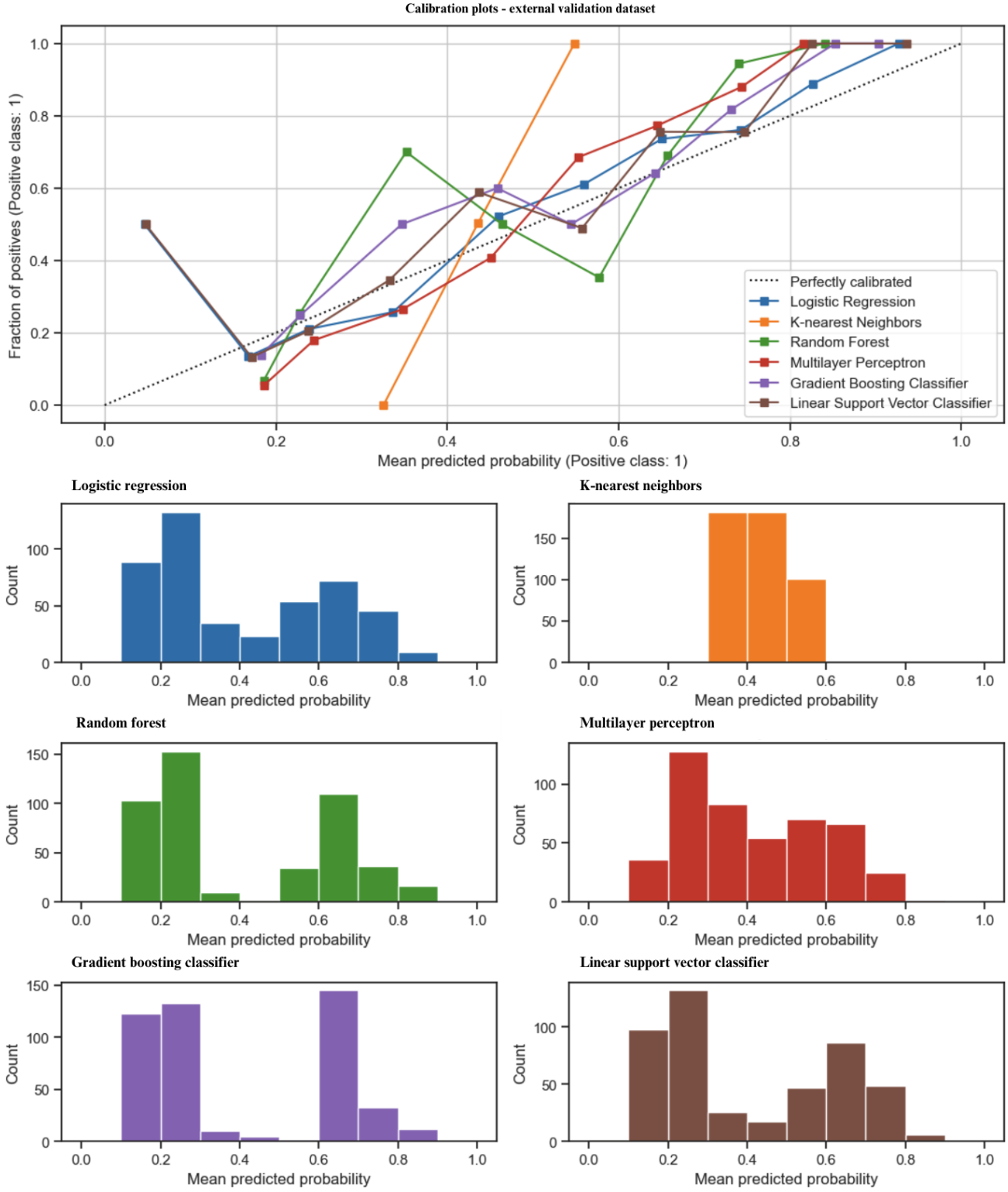
Figure 2. Model calibration of the developed models with internal validation.



In the external validation dataset (Figure 3), the LR, SVM, and MLP were almost perfectly calibrated to the expected probability of the data with a slight overestimation in the 0.40 - 0.80 MPP for MLP and an overestimation in the 0.00 - 0.40 MPP for LR and SVM. These 3 models have a similarly balanced distribution of MPP. RF aligned poorly with an overestimation in the low and high MPP, and it underestimated in the range between 0.40 and 0.70. The GBC aligned well with the calibration plot after 0.60 but overestimated before 0.40. The KNN classifier was poorly calibrated to the plot, giving only MPP in the range of 0.30 - 0.60.

XSL•FO
RenderX

Figure 3. Model calibration of the developed models with external validation.



Discussion

Model Performance

The internal validated models all showed remarkable discrimination performance, with the KNN model having the highest overall discrimination performance. However, the KNN model performed poorly in external validation datasets, indicating overfitting and the inability to generalize its results to other populations. In terms of model optimism, the LR, SVM, and MLP models had better discrimination capabilities, despite a slight decline in the models' performance during external validation. The mean AUCs of these models remained in the range of fair to good performance. When considering all aspects of model performance, the LR model was the most preferable because it provided good discrimination (AUC), an optimal range of predicted probabilities, and good calibration in both the internal and external validation datasets. On the other hand, other models had predicted probabilities that were not extreme enough, resulting in poor to fair calibrations, and they lacked

XSL•FO
RenderX

consistency between the internal and external validation datasets. In comparing the performance and validation aspects of previous frailty prediction models, most studies showed better performance than our best model (LR with a mean AUC of 0.81 in the internal validation set) [45]. The ELSA cohort study (2023) [46] reported the best performance from an RF model with an internal AUC of 0.92. Similarly, Bu et al (2023) [47] presented a multivariate LR model for predicting frailty in diabetic patients that had an AUC of 0.88 in the internal validation set. Additionally, a study on predictive modeling for frailty in older people using various machine learning methods found that artificial neural networks and SVM had the best performance in predicting mortality, with accuracies around 0.78 - 0.79 [32]. Although these studies reported high internal performance metrics, the absence of external validation raises concerns about their generalizability, especially for our focus on the community-dwelling older adult population in general. Our LR showed a mean AUC of 0.75 for the external validation dataset, which dropped slightly from the internal validated results, demonstrating a better edge in the community setting for our focused population. Furthermore, the predictors included in the previously reported model played an important role in prediction performance. Particularly, the utilization of handgrip strength in the study conducted by Bu et al [47] and the assessment of balance and chair stand in the ELSA cohort study [46] could potentially explain the high reported performance in both studies. These predictors either formed part of the frailty phenotype or served as surrogates for physical performance. However, incorporating these predictors into the models necessitated a trade-off between their value added to model performance and the difficulties of model utilization due to the time and skill required for the assessments. In our study, we proposed more parsimonious models using simpler predictions, which still achieved satisfactory performance for frailty screening.

Feature Selection and Findings Explanation

This study's findings suggest that machine learning models can be effective in classifying frailty status among community-dwelling older adults in Northern Thailand using simple predictors including, age, gender, household living arrangement, underlying diseases (hypertension and dyslipidemia), BMI, waist and calf circumference, and level of exhaustion. Age, BMI, waist circumference, and calf circumference are all potent risk factors for frailty in older adults. When the human body deviates from the normal physiologic process of aging, our levels of estrogen and testosterone gradually decline. These hormones play a vital role in maintaining muscle and bone mass, enhancing strength, and promoting optimal nervous system function [48,49]. As a result, the aging process can accelerate the decline of muscle, bone, and the nervous system, transforming an individual from fit to frail. Additionally, the female population is at higher risk of frailty because the normal bone turnover cycle is disrupted by estrogen deficiency during menopause, increasing bone resorption over deposition, and resulting in net bone loss in women [50]. Household living arrangement was also added to model features as we found that there were studies that showed an association of social adversity and support with frailty status [51,52]. Lastly, we selected level of exhaustion as a predictor in our models, as it had the highest crude aOR among all other features, and we found that it is highly feasible and time-efficient to acquire this data in real clinical settings, compared to other Fried's phenotypes, which involve multiple anthropometric and physical performance tests. These predictors also showed a significant contribution to the frailty prediction in the previous studies [46,47]. Most of our features used in the model are easily collectible, which makes it highly feasible and time-efficient to acquire all data in real clinical settings without depending on a high-level professional, making our model highly applicable for early frailty screening.

Limitations

The most important limitation of this study's models is the generalizability of the models. Our internal and external validation datasets were collected from a community-dwelling population in Lampang and Chiang Mai, respectively, which we confidently believe means that our models have high generalizability to the general population in Thailand. From our perspective, we assume that the models could also be implemented in other Asian countries, as studies showed similar frailty prevalence and population characteristics such as anthropometric measurements, age, household living arrangement, and underlying diseases of hypertension and diabetes [53-58].

However, we encourage conducting validation studies for other regions and populations before clinical application as frailty risk factors do vary across countries as well as in regions with different socioeconomic and health care contexts [59-61]. Another limitation of our study is spectrum bias, as our models were only able to distinguish between frail and robust older adults, despite Fried's criteria having 3 stages of frailty: robust, prefrail, and frail [62]. Nonetheless, we do not believe this bias will significantly affect our study's primary objective, which is the early detection of frailty in older adults to effectively prescribe nutrition and exercise interventions.

Implications

Identifying frailty early helps patient gain access to interventions like nutritional support and exercise programs faster, improving outcomes for older adults by preventing frailty progression, reducing falls, hospitalizations, and mortality, and enhancing quality of life [63,64]. We see 2 possibilities for the practical application of our developed and validated models. One option is to develop a web application that serves as a frailty screening tool that could be self-assessed by individuals or be used in outpatient clinic settings to screen patients. This would help health care providers efficiently identify patients with frailty who require closer monitoring or interventions. We have implemented the validated models to run on our web application, which can be accessed at the link in the Data Availability section.

Another option is to incorporate our machine learning models into electronic medical record or health surveillance systems. The machine learning models could be integrated with the electronic medical record system to provide automated frailty probability scores for each patient. This would enable health

XSL•FO
RenderX

care professionals to identify patients who require closer monitoring or interventions and could help optimize treatment plans for the community-dwelling population at risk of frailty. Furthermore, this could help future researchers retrieve and analyze frailty data from the hospital easier, leading to a better understanding of the factors that contribute to frailty and the development of more effective interventions, which also promotes more efficient use of resources within the health care system.

Conclusion

Machine learning models were fairly good at classifying frailty status among Thai community-dwelling older adults using age, gender, household living arrangement, underlying diseases (hypertension and dyslipidemia), BMI, waist and calf circumference, and level of exhaustion as predictors. The LR and RF models demonstrated the best discrimination performance and model calibration in both the internal and external validation datasets.

There are 2 potential practical applications for utilizing the study findings. These include creating a web application for self-screening or individual screening and incorporating machine learning models into electronic medical record or surveillance systems to provide automated frailty probability scores for individual patients. We advocate for further research on model external validation and temporal recalibration to enhance the model's practicality and applicability to the specific context in which it is used.

Acknowledgments

This work was supported by Chiang Mai University.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. The validated model's prediction and application interface can be accessed through the web application [65].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary materials
[DOCX File, 15 KB - aging_v8i1e62942_app1.docx]

References

1. World population ageing, 2019: highlights. United Nations Digital Library. URL: http://digitallibrary.un.org/record/3846855 [accessed 2025-03-23]
2. Ageing and health. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2022-08-02]
3. Population ageing in Thailand. ERIA. URL: http://www.eria.org/publications/population-ageing-in-thailand/ [accessed 2022-08-02]
4. Tantirat P, Suphanchaimat R, Rattanathumsakul T, Noree T. Projection of the number of elderly in different health states in Thailand in the next ten years, 2020-2030. Int J Environ Res Public Health 2020 Nov 23;17(22):8703. [doi: 10.3390/ijerph17228703] [Medline: 33238588]
5. O'Caoimh R, Sezgin D, O'Donovan MR, et al. Prevalence of frailty in 62 countries across the world: a systematic review and meta-analysis of population-level studies. Age Ageing 2021 Jan 8;50(1):96-104. [doi: 10.1093/ageing/afaa219] [Medline: 33068107]
6. Siriwardhana DD, Hardoon S, Rait G, Weerasinghe MC, Walters KR. Prevalence of frailty and prefrailty among community-dwelling older adults in low-income and middle-income countries: a systematic review and meta-analysis. BMJ Open 2018 Mar 1;8(3):e018195. [doi: 10.1136/bmjopen-2017-018195] [Medline: 29496895]
7. Nguyen AT, Nguyen LH, Nguyen TX, et al. Frailty prevalence and association with health-related quality of life impairment among rural community-dwelling older adults in Vietnam. Int J Environ Res Public Health 2019 Oct 12;16(20):3869. [doi: 10.3390/ijerph16203869] [Medline: 31614836]
8. Siriwardhana DD, Weerasinghe MC, Rait G, Falcaro M, Scholes S, Walters KR. Prevalence of frailty in rural community-dwelling older adults in Kegalle district of Sri Lanka: a population-based cross-sectional study. BMJ Open 2019 Jan 25;9(1):e026314. [doi: 10.1136/bmjopen-2018-026314] [Medline: 30782757]
9. Wu C, Smit E, Xue QL, Odden MC. Prevalence and correlates of frailty among community-dwelling Chinese older adults: the China Health and Retirement Longitudinal Study. J Gerontol A Biol Sci Med Sci 2017 Dec 12;73(1):102-108. [doi: 10.1093/gerona/glx098] [Medline: 28525586]

XSL•FO
RenderX

10. Collard RM, Boter H, Schoevers RA, Oude Voshaar RC. Prevalence of frailty in community-dwelling older persons: a systematic review. J Am Geriatr Soc 2012 Aug;60(8):1487-1492. [doi: 10.1111/j.1532-5415.2012.04054.x] [Medline: 22881367]
11. Vaingankar JA, Chong SA, Abdin E, et al. Prevalence of frailty and its association with sociodemographic and clinical characteristics, and resource utilization in a population of Singaporean older adults. Geriatr Gerontol Int 2017 Oct;17(10):1444-1454. [doi: 10.1111/ggi.12891] [Medline: 27576598]
12. Semmarath W, Seesen M, Yodkeeree S, et al. The association between frailty indicators and blood-based biomarkers in early-old community dwellers of Thailand. Int J Environ Res Public Health 2019 Sep 17;16(18):3457. [doi: 10.3390/ijerph16183457] [Medline: 31533354]
13. Srinonprasert V, Chalermsri C, Aekplakorn W. Frailty index to predict all-cause mortality in Thai community-dwelling older population: a result from a National Health Examination Survey cohort. Arch Gerontol Geriatr 2018;77:124-128. [doi: 10.1016/j.archger.2018.05.002] [Medline: 29751290]
14. Xue QL. The frailty syndrome: definition and natural history. Clin Geriatr Med 2011 Feb;27(1):1-15. [doi: 10.1016/j.cger.2010.08.009] [Medline: 21093718]
15. Dent E, Morley JE, Cruz-Jentoft AJ, et al. Physical frailty: ICFSR International Clinical Practice Guidelines for Identification and Management. J Nutr Health Aging 2019;23(9):771-787. [doi: 10.1007/s12603-019-1273-z] [Medline: 31641726]
16. Setiati S, Laksmi PW, Aryana I, et al. Frailty state among Indonesian elderly: prevalence, associated factors, and frailty state transition. BMC Geriatr 2019 Jul 3;19(1):182. [doi: 10.1186/s12877-019-1198-8] [Medline: 31269921]
17. Robertson DA, Savva GM, Kenny RA. Frailty and cognitive impairment--a review of the evidence and causal mechanisms. Ageing Res Rev 2013 Sep;12(4):840-851. [doi: 10.1016/j.arr.2013.06.004] [Medline: 23831959]
18. End of life care in frailty: falls. British Geriatrics Society. URL: https://www.bgs.org.uk/resources/end-of-life-care-in-frailty-falls [accessed 2022-08-02]
19. Cheng MH, Chang SF. Frailty as a risk factor for falls among community dwelling people: evidence from a meta-analysis. J Nurs Scholarsh 2017 Sep;49(5):529-536. [doi: 10.1111/jnu.12322] [Medline: 28755453]
20. Davis-Ajami ML, Chang PS, Wu J. Hospital readmission and mortality associations to frailty in hospitalized patients with coronary heart disease. Aging and Health Research 2021 Dec;1(4):100042. [doi: 10.1016/j.ahr.2021.100042]
21. Travers J, Romero-Ortuno R, Bailey J, Cooney MT. Delaying and reversing frailty: a systematic review of primary care interventions. Br J Gen Pract 2019 Jan;69(678):e61-e69. [doi: 10.3399/bjgp18X700241] [Medline: 30510094]
22. Bray NW, Smart RR, Jakobi JM, Jones GR. Exercise prescription to reverse frailty. Appl Physiol Nutr Metab 2016 Oct;41(10):1112-1116. [doi: 10.1139/apnm-2016-0226] [Medline: 27649859]
23. Marcucci M, Damanti S, Germini F, et al. Interventions to prevent, delay or reverse frailty in older people: a journey towards clinical guidelines. BMC Med 2019 Oct 29;17(1):193. [doi: 10.1186/s12916-019-1434-2] [Medline: 31660959]
24. Turner G, Clegg A, British Geriatrics Society, Age UK, Royal College of General Practioners. Best practice guidelines for the management of frailty: a British Geriatrics Society, Age UK and Royal College of General Practitioners report. Age Ageing 2014 Nov;43(6):744-747. [doi: 10.1093/ageing/afu138] [Medline: 25336440]
25. Bergman H, Ferrucci L, Guralnik J, et al. Frailty: an emerging research and clinical paradigm--issues and controversies. J Gerontol A Biol Sci Med Sci 2007 Jul;62(7):731-737. [doi: 10.1093/gerona/62.7.731] [Medline: 17634320]
26. Savel TG, Foldy S, Centers for Disease Control and Prevention. The role of public health informatics in enhancing public health surveillance. MMWR Suppl 2012 Jul 27;61(3):20-24. [Medline: 22832993]
27. Gamache R, Kharrazi H, Weiner JP. Public and population health informatics: the bridging of big data to benefit communities. Yearb Med Inform 2018 Aug;27(1):199-206. [doi: 10.1055/s-0038-1667081] [Medline: 30157524]
28. Yang H, Chang J, He W, Wee CF, Yit JST, Feng M. Frailty modeling using machine learning methodologies: a systematic review with discussions on outstanding questions. IEEE J Biomed Health Inform 2024;29(1):631-642. [doi: 10.1109/JBHI.2024.3430226]
29. Bian C, Ye B, Chu CH, McGilton KS, Mihailidis A. Technology for home-based frailty assessment and prediction: a systematic review. Gerontechnology 2020 Nov 2;19(3):1-13. [doi: 10.4017/gt.2020.19.003.06]
30. Mohanty SD, Lekan D, McCoy TP, Jenkins M, Manda P. Machine learning for predicting readmission risk among the frail: explainable AI for healthcare. Patterns (N Y) 2022 Jan 14;3(1):100395. [doi: 10.1016/j.patter.2021.100395] [Medline: 35079714]
31. Aponte-Hao S, Wong ST, Thandi M, et al. Machine learning for identification of frailty in Canadian primary care practices. Int J Popul Data Sci 2021;6(1):1650. [doi: 10.23889/ijpds.v6i1.1650] [Medline: 34541337]
32. Tarekegn A, Ricceri F, Costa G, Ferracin E, Giacobini M. Predictive modeling for frailty conditions in elderly people: machine learning approaches. JMIR Med Inform 2020 Jun 4;8(6):e16678. [doi: 10.2196/16678] [Medline: 32442149]
33. Huang L, Chen H, Liang Z. Enhancing the convenience of frailty index assessment for elderly Chinese people with machine learning methods. Sci Rep 2024 Oct 5;14(1):23227. [doi: 10.1038/s41598-024-74194-x] [Medline: 39369089]
34. Zhang W, Wang J, Xie F, et al. Development and validation of machine learning models to predict frailty risk for elderly. J Adv Nurs 2024 Dec;80(12):5064-5075. [doi: 10.1111/jan.16192] [Medline: 38605460]
35. Seesen M, Sirikul W, Ruangsuriya J, Griffiths J, Siviroj P. Cognitive frailty in Thai community-dwelling elderly: prevalence and its association with malnutrition. Nutrients 2021 Nov 25;13(12):4239. [doi: 10.3390/nu13124239] [Medline: 34959791]

XSL•FO
RenderX

36. Thinuan P, Siviroj P, Lerttrakarnnon P, Lorga T. Prevalence and potential predictors of frailty among community-dwelling older persons in Northern Thailand: a cross-sectional study. Int J Environ Res Public Health 2020 Jun 8;17(11):4077. [doi: 10.3390/ijerph17114077] [Medline: 32521642]
37. Haque MN. Active ageing level of older persons: regional comparison in Thailand. J Aging Res 2016;2016:9093018. [doi: 10.1155/2016/9093018] [Medline: 27375903]
38. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
39. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. BMJ 2020 Mar 18;368:m441. [doi: 10.1136/bmj.m441] [Medline: 32188600]
40. Riley RD, Snell KIE, Archer L, et al. Evaluation of clinical prediction models (part 3): calculating the sample size required for an external validation study. BMJ 2024 Jan 22;384:e074821. [doi: 10.1136/bmj-2023-074821] [Medline: 38253388]
41. Lemaître G, Nogueira F, Aridas CK. Imbalanced-learn: a Python toolbox to tackle the curse of imbalanced datasets in machine learning. J Mach Learn Res 2017;18(17):1-5.
42. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling Technique. jair 2002 Jun 1;16:321-357. [doi: 10.1613/jair.953]
43. Chittrakul J, Siviroj P, Sungkarat S, Sapbamrer R. Physical frailty and fall risk in community-dwelling older adults: a cross-sectional study. J Aging Res 2020;2020:3964973. [doi: 10.1155/2020/3964973] [Medline: 32714612]
44. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024 Apr 16;385:e078378. [doi: 10.1136/bmj-2023-078378] [Medline: 38626948]
45. Velazquez-Diaz D, Arco JE, Ortiz A, et al. Use of artificial intelligence in the identification and diagnosis of frailty syndrome in older adults: scoping review. J Med Internet Res 2023 Oct 20;25(1):e47346. [doi: 10.2196/47346] [Medline: 37862082]
46. Leme DDC, de Oliveira C. Machine learning models to predict future frailty in community-dwelling middle-aged and older adults: the ELSA cohort study. J Gerontol A Biol Sci Med Sci 2023 Oct 28;78(11):2176-2184. [doi: 10.1093/gerona/glad127] [Medline: 37209408]
47. Bu F, Deng XH, Zhan NN, et al. Development and validation of a risk prediction model for frailty in patients with diabetes. BMC Geriatr 2023 Mar 27;23(1):172. [doi: 10.1186/s12877-023-03823-3] [Medline: 36973658]
48. Dobbs MB, Buckwalter J, Saltzman C. Osteoporosis: the increasing role of the orthopaedist. Iowa Orthop J 1999;19:43-52. [Medline: 10847516]
49. Cruz-Jentoft AJ, Bahat G, Bauer J, et al. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing 2019 Jul 1;48(4):601-601. [doi: 10.1093/ageing/afz046] [Medline: 31081853]
50. Truumees E. Chapter 22 - Osteoporosis. In: Errico TJ, Lonner BS, Moulton AW, editors. Surg Manag Spinal Deform: W.B. Saunders; 2009:351-372. [doi: 10.1016/B978-141603372-1.50025-1]
51. Gomes CDS, Guerra RO, Wu YY, et al. Social and economic predictors of worse frailty status occurrence across selected countries in North and South America and Europe. Innov Aging 2018 Sep;2(3):igy037. [doi: 10.1093/geroni/igy037] [Medline: 30569024]
52. Lee Y, Chon D, Kim J, Ki S, Yun J. The predictive value of social frailty on adverse outcomes in older adults living in the community. J Am Med Dir Assoc 2020 Oct;21(10):1464-1469. [doi: 10.1016/j.jamda.2020.03.010] [Medline: 32362535]
53. Yu R, Wu WC, Leung J, Hu SC, Woo J. Frailty and its contributory factors in older adults: a comparison of two Asian regions (Hong Kong and Taiwan). Int J Environ Res Public Health 2017 Sep 21;14(10):1096. [doi: 10.3390/ijerph14101096] [Medline: 28934150]
54. Liu ZY, Wei YZ, Wei LQ, et al. Frailty transitions and types of death in Chinese older adults: a population-based cohort study. Clin Interv Aging 2018;13:947-956. [doi: 10.2147/CIA.S157089] [Medline: 29805253]
55. To TL, Doan TN, Ho WC, Liao WC. Prevalence of frailty among community-dwelling older adults in Asian countries: a systematic review and meta-analysis. Healthcare (Basel) 2022 May 12;10(5):895. [doi: 10.3390/healthcare10050895] [Medline: 35628034]
56. Huynh TQH, Pham TLA, Vo VT, Than HNT, Nguyen TV. Frailty and associated factors among the elderly in Vietnam: a cross-sectional study. Geriatrics (Basel) 2022 Aug 20;7(4):85. [doi: 10.3390/geriatrics7040085] [Medline: 36005261]
57. Norazman CW, Adznam SN, Jamaluddin R. Physical frailty among urban-living community-dwelling older adults in Malaysia. Int J Environ Res Public Health 2020 Sep 9;17(18):6549. [doi: 10.3390/ijerph17186549] [Medline: 32916809]
58. Murayama H, Kobayashi E, Okamoto S, et al. National prevalence of frailty in the older Japanese population: findings from a nationally representative survey. Arch Gerontol Geriatr 2020;91:104220. [doi: 10.1016/j.archger.2020.104220] [Medline: 32810734]
59. Morkphrom E, Srinonprasert V, Sura-amonrattana U, et al. Severity of frailty using modified Thai frailty index, social factors, and prediction of mortality among community-dwelling older adults in a middle-income country. Front Med (Lausanne) 2022;9:1060990. [doi: 10.3389/fmed.2022.1060990] [Medline: 36569139]
60. Qin Y, Hao X, Lv M, Zhao X, Wu S, Li K. A global perspective on risk factors for frailty in community-dwelling older adults: a systematic review and meta-analysis. Arch Gerontol Geriatr 2023 Feb;105:104844. [doi: 10.1016/j.archger.2022.104844] [Medline: 36335672]

XSL•FO
RenderX

61. Zeng XZ, Meng LB, Li YY, et al. Prevalence and factors associated with frailty and pre-frailty in the older adults in China: a national cross-sectional study. Front Public Health 2023;11:1110648. [doi: 10.3389/fpubh.2023.1110648] [Medline: 37554734]
62. Willis BH. Spectrum bias--why clinicians need to be cautious when applying diagnostic test studies. Fam Pract 2008 Oct;25(5):390-396. [doi: 10.1093/fampra/cmn051] [Medline: 18765409]
63. Rockwood K, Song X, Mitnitski A. Changes in relative fitness and frailty across the adult lifespan: evidence from the Canadian National Population Health Survey. CMAJ 2011 May 17;183(8):E487-E494. [doi: 10.1503/cmaj.101271] [Medline: 21540166]
64. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet 2013 Mar 2;381(9868):752-762. [doi: 10.1016/S0140-6736(12)62167-9] [Medline: 23395245]
65. Frailty Classificiation Using Machine Learning Model. URL: https://frailty-commed-cmu.streamlit.app [accessed 2025-03-23]

Abbreviations

aOR: adjusted odds ratio
AUC: area under the receiver operating characteristic curve
CC: calf circumference
GBC: gradient boosting classifier
KNN: k-nearest neighbors
LR: logistic regression
MLP: multilayer perceptron artificial neural network
MPP: mean predicted probability
NPV: negative predictive value
PPV: positive predictive value
RF: random forest
SMOTE: synthetic minority oversampling technique
SVM: linear support vector machine classifier
TRIPOD+AI: Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis plus Artificial Intelligence

Edited by J Wang; submitted 05.06.24; peer-reviewed by Y Huang, Y Huusi; revised version received 26.11.24; accepted 28.02.25; published 02.04.25.

Please cite as:
Isaradech N, Sirikul W, Buawangpong N, Siviroj P, Kitro A
Machine Learning Models for Frailty Classification of Older Adults in Northern Thailand: Model Development and Validation Study
JMIR Aging 2025;8:e62942
URL: https://aging.jmir.org/2025/1/e62942
doi:10.2196/62942

© Natthanaphop Isaradech, Wachiranun Sirikul, Nida Buawangpong, Penprapa Siviroj, Amornphat Kitro. Originally published in JMIR Aging (https://aging.jmir.org), 2.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Association Between Sleep Duration and Cognitive Frailty in Older Chinese Adults: Prospective Cohort Study

Ruixue Cai[1,2], PhD; Jianqian Chao[1], PhD; Chenlu Gao[2,3], PhD; Lei Gao[2,3,4,5], MBBS; Kun Hu[2,3,4,5], PhD; Peng Li[2,3,4,5], PhD

[1]Department of Epidemiology and Health Statistics, School of Public Health, Southeast University, Nanjing, China

[2]Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital, 149 13th Street, Boston, MA, United States

[3]Division of Sleep Medicine, Harvard Medical School, Boston, MA, United States

[4]Division of Sleep and Circadian Disorders, Departments of Neurology and Medicine, Brigham and Women's Hospital, Boston, MA, United States

[5]Broad Institute of MIT and Harvard, Cambridge, MA, United States

Corresponding Author:
Peng Li, PhD
Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital, 149 13th Street, Boston, MA, United States

Abstract

Background: Disturbed sleep patterns are common among older adults and may contribute to cognitive and physical declines. However, evidence for the relationship between sleep duration and cognitive frailty, a concept combining physical frailty and cognitive impairment in older adults, is lacking.

Objective: This study aimed to examine the associations of sleep duration and its changes with cognitive frailty.

Methods: We analyzed data from the 2008 - 2018 waves of the Chinese Longitudinal Healthy Longevity Survey. Cognitive frailty was rendered based on the modified Fried frailty phenotype and Mini-Mental State Examination. Sleep duration was categorized as short (<6 h), moderate (6 - 9 h), and long (>9 h). We examined the association of sleep duration with cognitive frailty status at baseline using logistic regressions and with the future incidence of cognitive frailty using Cox proportional hazards models. Restricted cubic splines were used to explore potential nonlinear associations.

Results: Among 11,303 participants, 1298 (11.5%) had cognitive frailty at baseline. Compared to participants who had moderate sleep duration, the odds of having cognitive frailty were higher in those with long sleep duration (odds ratio 1.71, 95% CI 1.48 - 1.97; *P*<.001). A J-shaped association between sleep duration and cognitive frailty was also observed (*P*<.001). Additionally, during a mean follow-up of 6.7 (SD 2.6) years among 5201 participants who were not cognitively frail at baseline, 521 (10%) participants developed cognitive frailty. A higher risk of cognitive frailty was observed in participants with long sleep duration (hazard ratio 1.32, 95% CI 1.07 - 1.62; *P*=.008).

Conclusions: Long sleep duration was associated with cognitive frailly in older Chinese adults. These findings provide insights into the relationship between sleep duration and cognitive frailty, with potential implications for public health policies and clinical practice.

(JMIR Aging 2025;8:e65183) doi:10.2196/65183

KEYWORDS

aging; frailty; cognition; cohort study; sleep duration; sleep quality; longitudinal study

Introduction

Physical frailty and cognitive impairment are prevalent among older adults and have individually been associated with adverse health outcomes [1,2]. They often coincide with aging and can be bidirectionally linked to each other [3,4], prompting the introduction of the concept of cognitive frailty—the coexistence of both physical frailty and cognitive impairment [5]. The necessity is further justified by the findings that cognitive frailty poses an even greater risk of adverse outcomes compared to the isolated effects of the 2 conditions [6,7]. Identifying factors associated with or contributing to cognitive frailty is important and will provide a better understanding of the underlying mechanisms.

Disturbed sleep patterns, such as an increased number of awakenings, abnormal nighttime sleep duration, and poorer sleep quality, are common among older adults [8]. Numerous epidemiological studies identified sleep disturbances as risk factors for physical frailty and cognitive impairment, such as shorter or longer sleep duration, insomnia, and sleep apnea [9,10]. In addition, recent studies have linked long sleep duration to cognitive frailty, but they are from cross-sectional analyses

XSL•FO
RenderX

[11,12]. A longitudinal study in Mexican older adults reported associations between increasing sleep duration trajectory with greater odds for mild cognitive impairment and frailty separately [13]. Whether sleep patterns are associated with the development of cognitive frailty as a unified concept remains to be elucidated.

To address this knowledge gap, this study aimed to assess the associations of sleep duration with cognitive frailty in a nationally representative, large cohort of older Chinese adults. By focusing on an older Chinese adult population, our findings may offer insights into the relationship between sleep duration and cognitive frailty within this demographic, with potential implications for public health policies and clinical practice.

Methods

Study Design and Participants

We used data from the Chinese Longitudinal Healthy Longevity Survey (CLHLS). As detailed previously, CLHLS is an ongoing, nationwide representative longitudinal cohort study of Chinese adults aged 65 years and older from 23 out of 31 provinces or municipalities or autonomous regions in mainland China [14,15]. The CLHLS started in 1998 with surveys on sociodemographic characteristics, lifestyle, cognitive function, psychological status, and physical function conducted every 2 - 3 years.

In this study, we used data from 4 waves (2008, 2011, 2014, and 2018) of the CLHLS. To maximize statistical power and leverage the large sample size, we first examined cross-sectional associations between sleep duration and sleep quality, with cognitive frailty at baseline (ie, the 2008 wave). For this set of analyses, we included 11,303 participants who had cognition and physical frailty assessments at baseline. We then examined whether baseline sleep duration and sleep quality, as well as change in sleep duration, were associated with the future development of cognitive frailty. For these longitudinal analyses, we excluded 1298 participants who had cognitive frailty at baseline, 4780 participants who had no follow-up cognitive frailty assessment, and 24 participants who had missing value in sleep duration in the 2011 wave. As a result, 5201 participants entered the subsequent longitudinal analysis.

Assessment of Cognitive Frailty

Cognitive frailty was identified as the simultaneous presence of both physical frailty and cognitive impairment [5].

Physical frailty was assessed by a modified Fried frailty phenotype, based on 5 components: shrinking, weakness, low mobility, exhaustion, and inactivity [16,17]. Shrinking was defined as having a BMI of 18.5 kg/m^2 or less. Weakness was determined if participants self-reported that they were unable to lift a bag weighing 5 kg or above. Low mobility was identified if participants self-reported that they were unable to walk 1 km or longer in a row. Exhaustion was identified if participants answered “always,” “often,” or “sometimes” to the question “Do you feel the older you get, the more useless you are?” Inactivity was defined as participants doing the following activities 1 time per week or less: housework, outside activity, garden work, raising domestic animals or pets, playing cards or mah-jongg, and social activity (organized). Participants were considered physically frail when meeting ≥3 of the 5 criteria.

Cognitive impairment was assessed by a validated Chinese version of the Mini-Mental State Examination (C-MMSE) [15,18]. The C-MMSE includes 24 items regarding orientation, memory, attention, calculation, and language, ranging from 0 to 30. Based on a prior study, participants were considered cognitive impairment if the C-MMSE score was ≤22 [15].

Assessment of Sleep Variables

Sleep duration was estimated based on the answers to the question “How many hours do you sleep normally?” Based on their answers, participants were further categorized into short (<6 h), moderate (6 - 9 h), and long (>9 h) sleep durations. This categorization aligned with previous research indicating that both short and long sleep durations are associated with poorer health outcomes [19]. The rate of change in sleep duration was calculated by taking the difference between sleep durations reported in the 2008 and 2011 waves (ie, sleep duration in the 2011 wave minus sleep duration in the 2008 wave) and dividing it by the individual follow-up interval in years. Sleep quality was assessed based on answers to the question “How about the quality of your sleep?” A score of “poor” was assigned if the answer was “so-so,” “bad,” or “very bad,” and a score of “good” was assigned otherwise.

Covariates

We considered the following covariates that have been identified to be associated with physical or cognitive outcomes in previous studies [20,21]: age, sex, education, marital status, residence, economic status, loneliness, smoking status, drinking status, and multimorbidity. Education was categorized as “Not Educated” if the participant had not attended any schools and “With Formal Education” if the participant had attended schools more than 1 year. Marital status was defined as currently married and living with a spouse, and others (eg, separated, divorced, widowed, and never married). Residence was categorized into living in urban (eg, city or town) or rural areas. Economic status was categorized into economic dependence or independence based on their income source. Loneliness was identified if participants answered “always,” “often,” or “sometimes” to the question “Do you feel lonely and isolated?” Smoking status was categorized as “Never smoked,” and “Former or current smoker.” Drinking status was categorized into “Never drank,” and “Former or current drinker.” Multimorbidity was defined from self-reports and categorized as “Yes” if having 2 or more of the following chronic diseases: hypertension, diabetes, heart disease, stroke or cerebrovascular disease, respiratory disease, cancer, and Parkinson disease, or “No” if otherwise.

Statistical Analysis

Listwise deletion was applied to both cross-sectional and longitudinal analyses to handle missing data on covariates. The *t* tests for continuous variables and chi-square tests for categorical variables were used to compare the differences in the baseline characteristics.

Multivariate logistic regression models were performed to assess the associations of sleep duration and sleep quality with

XSL•FO
RenderX

cognitive frailty at baseline (ie, the 2008 wave). We performed 3 models: model A was unadjusted; model B was adjusted for age, sex, and education; model C was further adjusted for marital status, residence, economic status, loneliness, smoking status, drinking status, and multimorbidity.

Cox proportional hazard (PH) models were used to test the associations between sleep duration and sleep quality with incident cognitive frailty during the follow up. Death was treated as a right-censored event at the time of occurrence. The follow-up duration was defined as the time interval in years from baseline to the date of the first occurrence of cognitive frailty, death, or the latest available data, whichever occurred first. We rounded the time to the nearest integer year to account for group-tied events. Similarly, we performed 3 models: model A was unadjusted; model B was adjusted for age, sex, and education; model C was further adjusted for marital status, residence, economic status, loneliness, smoking status, drinking status, and multimorbidity. The PH assumption of the Cox models was assessed using Schoenfeld residuals. No major violations of the PH assumption were detected for exposure variables.

Restricted cubic spline (RCS) curves were plotted to explore the nonlinear cross-sectional and longitudinal associations between sleep duration and cognitive frailty. Additionally, several sensitivity analyses were conducted to verify the robustness of our findings. First, we adjusted for each chronic disease individually and for all chronic diseases simultaneously. Second, we performed stratified analyses by sex and age. Furthermore, we performed similar sets of Cox PH models to investigate the associations of changes in sleep duration and incident cognitive frailty.

All statistical analyses were performed in R (version 4.1.2; R Core Team). Statistical significance was considered at a 2-tailed α level of .05.

Ethical Considerations

The CLHLS was approved by the Duke University Institutional Review Board (Pro00062871) and the Peking University Biomedical Ethics Committee (IRB00001052–13074). All participants provided written informed consent. The data used in this study were deidentified to protect participant privacy and confidentiality. No compensation was provided to participants.

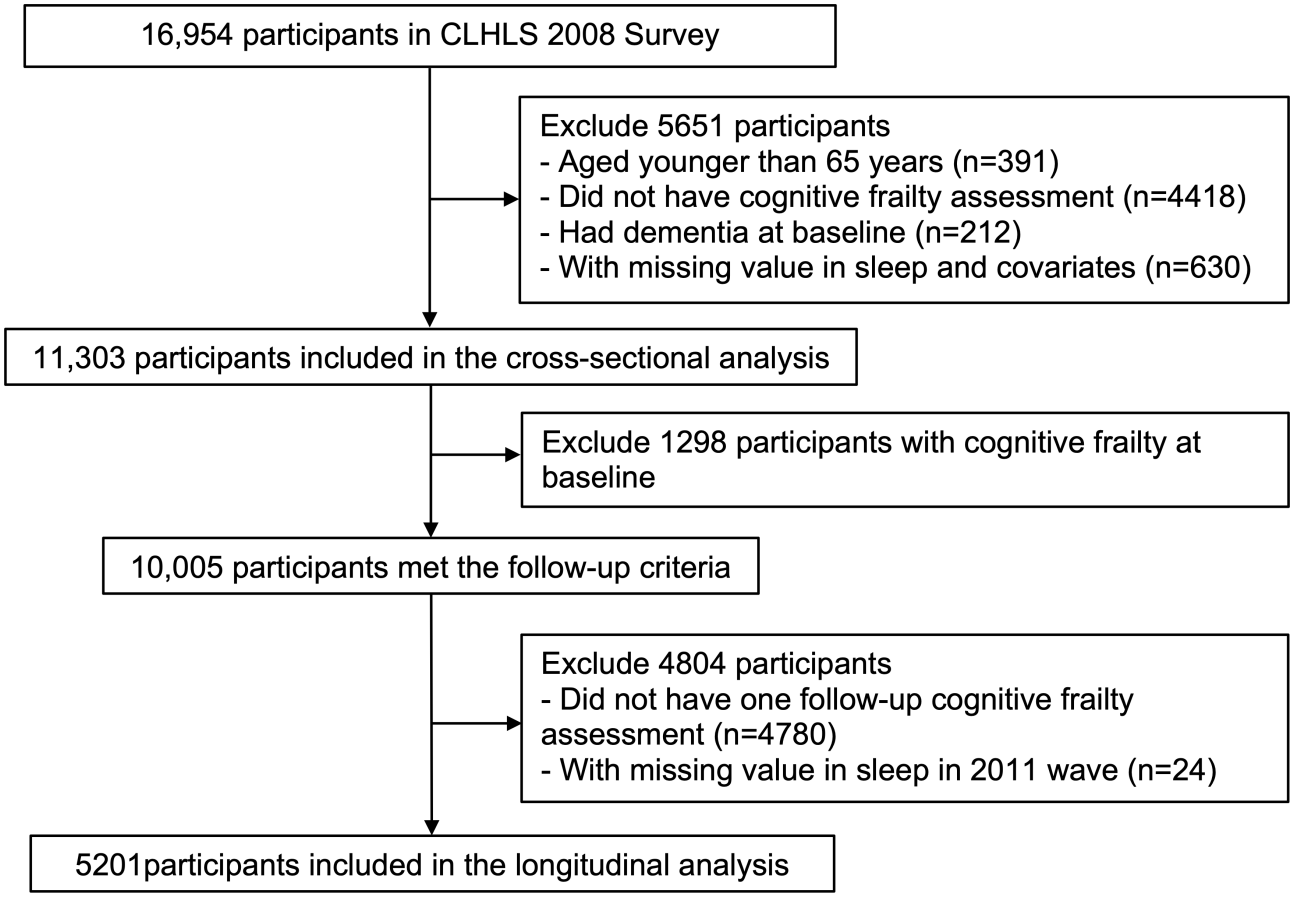
Results

Participant Characteristics

The participant flowchart is presented in Figure 1. Of the 11,303 participants included in the cross-sectional analysis, 6037 (53.4%) participants were female. The mean age at baseline was 84.7 (SD 11.0) years old. Approximately 11.5% (n=1298) of all participants had cognitive frailty at baseline. The numbers of participants who reported short, moderate, or long sleep duration were 1327 (11.7%), 7207 (63.8%), and 2769 (24.5%), respectively. Poor sleep quality was found in 3796 (33.6%) participants. Baseline characteristics stratified by cognitive frailty status at baseline are summarized in Table 1. Additionally, baseline characteristics categorized by sleep duration are presented in Multimedia Appendix 1. Baseline characteristics of 5201 participants included in the longitudinal analysis are summarized in Multimedia Appendix 2.

XSL•FO
RenderX

Figure 1. Flowchart of participants through the study. CLHLS: Chinese Longitudinal Healthy Longevity Survey.



XSL•FO
RenderX

Table . Baseline characteristics of participants by cognitive frailty status at baseline.

	Overall (n=11,303)	Non-CF[a] (n=10,005)	CF (n=1298)	*P* value
Age (years), mean (SD)	84.7 (11.0)	83.3 (10.6)	95.3 (7.2)	<.001
Sex, n (%)				<.001
Male	5266 (46.6)	4966 (49.6)	300 (23.1)	
Female	6037 (53.4)	5039 (50.4)	998 (76.9)	
Education, n (%)				<.001
Not educated	6548 (57.9)	5457 (54.5)	1091 (84.1)	
With formal education	4755 (42.1)	4548 (45.5)	207 (15.9)	
Marital status, n (%)				<.001
Married and living with spouse	4098 (36.3)	3979 (39.8)	119 (9.2)	
Others	7205 (63.7)	6026 (60.2)	1179 (90.8)	
Current residence, n (%)				.12
Urban	4647 (41.1)	4140 (41.4)	507 (39.1)	
Rural	6656 (58.9)	5865 (58.6)	791 (60.9)	
Economic status, n (%)				<.001
Dependence	7929 (70.2)	6716 (67.1)	1213 (93.5)	
Independence	3374 (29.9)	3289 (32.9)	85 (6.5)	
Loneliness, n (%)				<.001
Yes	3581 (31.7)	2943 (29.4)	638 (49.2)	
No	7722 (68.3)	7062 (70.6)	660 (50.8)	
Smoking status, n (%)				<.001
Never smoked	7271 (64.3)	6230 (62.3)	1041 (80.2)	
Former or current smoker	4032 (35.7)	3775 (37.7)	257 (19.8)	
Drinking status, n (%)				<.001
Never drank	7670 (67.9)	6665 (66.6)	1005 (77.4)	
Former or current drinker	3633 (32.1)	3340 (33.4)	293 (22.6)	
Multimorbidity, n (%)				.02
Yes	1068 (9.4)	969 (9.7)	99 (7.6)	
No	10,235 (90.6)	9036 (90.3)	1199 (92.4)	
Sleep quality, n (%)				<.001
Good	7507 (66.4)	6706 (67.0)	801 (61.7)	
Poor	3796 (33.6)	3299 (33.0)	497 (38.3)	
Sleep duration, n (%)				<.001
Short (<6 h)	1327 (11.7)	1179 (11.8)	148 (11.4)	
Moderate (6 - 9 h)	7207 (63.8)	6557 (65.5)	650 (50.1)	
Long (>9 h)	2769 (24.5)	2269 (22.7)	500 (38.5)	

[a]CF: cognitive frailty.

Association Between Sleep Duration and Cognitive Frailty at Baseline

Compared to participants who had moderate sleep duration, the odds of having cognitive frailty were higher in those who had long sleep duration (odds ratio [OR] 1.71, 95% CI 1.48 - 1.97; *P*<.001) in the initial model adjusted for age, sex, and education Table 2. The association remained statistically significant with a similar OR in the augmented model after further adjustments of marital status, residence, economic status, loneliness, smoking status, drinking status, and multimorbidity (OR 1.69, 95% CI 1.46 - 1.95; *P*<.001). There was no statistically significant

XSL•FO
RenderX

difference in the odds of having cognitive frailty between participants with short and modest sleep durations in the initial model adjusted for age, sex, and education. Compared with those who had a good sleep quality, participants who had a poor sleep quality were at a higher risk for cognitive frailty in both initial and the augmented models with further adjustments (initial model: OR 1.63, 95% CI 1.40 - 1.89; *P*<.001; augmented model: OR 1.41, 95% CI 1.21 - 1.64; *P*<.001; Table 2). In addition, RCS analysis revealed a J-shaped association between sleep duration and cognitive frailty at baseline (*P* for nonlinear<.001; Figure 2).

The results of sensitivity analyses are summarized in Multimedia Appendices 3 and 4. The associations between long sleep duration and baseline cognitive frailty remained consistent when adjusting for each chronic disease individually and when accounting for all chronic diseases simultaneously (Multimedia Appendix 3). For individuals younger than 80 years old, the association of long sleep duration with cognitive frailty became not statistically significant, whereas for individuals older than 80 years old, the association persisted. The association remained significant in sex-stratified models (Multimedia Appendix 4).

Table . Association between sleep duration and cognitive frailty at baseline.

	Model A[a]		Model B[b]		Model C[c]	
	OR[d] (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Age	__[e]	—	1.12 (1.12 - 1.13)	<.001	1.12 (1.11 - 1.13)	<.001
Sex (female)	—	—	1.80 (1.54 - 2.11)	<.001	1.56 (1.30 - 1.87)	<.001
Not educated	—	—	1.78 (1.49 - 2.13)	<.001	1.57 (1.31 - 1.90)	<.001
Married and living with spouse	—	—	—	—	1.21 (0.97 - 1.52)	.09
Rural residence	—	—	—	—	0.86 (0.75 - 0.99)	.04
Economic dependence	—	—	—	—	2.24 (1.75 - 2.89)	<.001
Loneliness	—	—	—	—	1.82 (1.60 - 2.09)	<.001
Smoker	—	—	—	—	0.91 (0.76 - 1.10)	.33
Drinker	—	—	—	—	1.03 (0.87 - 1.22)	.73
Multimorbidity	—	—	—	—	1.47 (1.14 - 1.87)	.002
Poor sleep quality	1.62 (1.41 - 1.85)	<.001	1.63 (1.40 - 1.89)	<.001	1.41 (1.21 - 1.64)	<.001
Sleep duration (hours)						
Moderate (6 - 9)	reference	reference	reference	reference	reference	reference
Short (<6)	1.01 (0.82 - 1.23)	.95	1.04 (0.83 - 1.29)	.75	1.06 (0.85 - 1.32)	.61
Long (>9)	2.47 (2.17-2.82)	<.001	1.71 (1.48 - 1.97)	<.001	1.69 (1.46 - 1.95)	<.001

[a]Model was unadjusted.

[b]Model was adjusted for age, sex, and education at baseline.

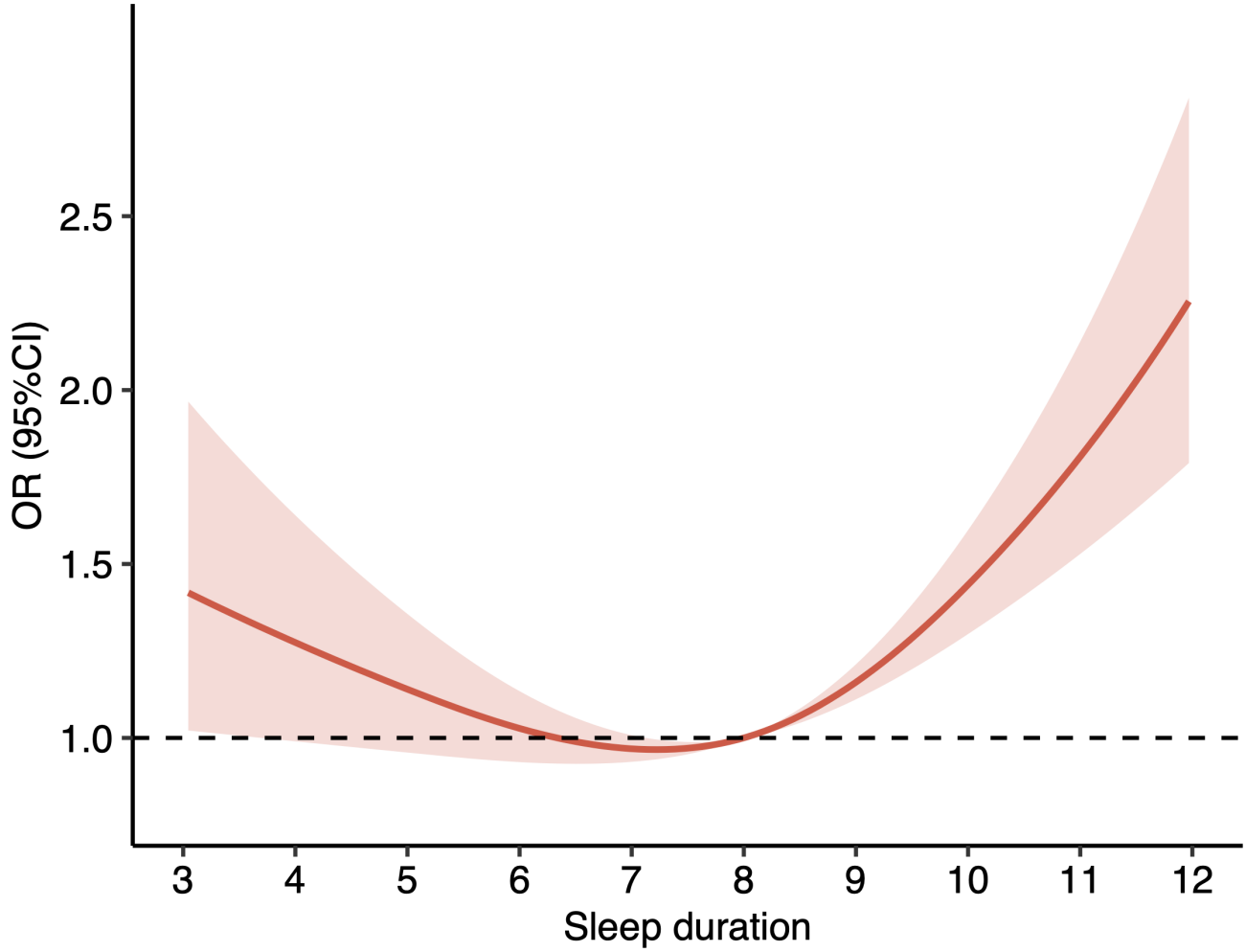
[c]Model was adjusted for age, sex, education, marital status, residence, economic status, loneliness, smoking status, drinking status, and multimorbidity at baseline.

[d]OR: odds ratio.

[e]Not applicable.

XSL•FO
RenderX

Figure 2. Nonlinear association between sleep duration and cognitive frailty at baseline using restricted cubic spline. The logistic regression model was used to fit the data, with solid lines representing the estimated OR and shaded areas representing the corresponding 95% CIs. OR: odds ratio.



Association Between Sleep Duration and Incident Cognitive Frailty During Follow-Up

Over a mean of 6.7 (SD 2.6) years of follow up, 521 (10%) participants out of the 5201 participants developed cognitive frailty. In the initial model adjusted for age, sex, and education, as compared with participants who had moderate sleep duration at baseline, a higher risk of cognitive frailty was observed in participants who had long sleep duration at baseline (hazard ratio [HR] 1.32, 95% CI 1.07 - 1.62; *P*=.008; Table 3). The association persisted after further adjusting for marital status, residence, economic status, loneliness, smoking status, drinking status, and multimorbidity (HR 1.30, 95% CI 1.06 - 1.59; *P*=.01). There was no statistically significant difference in the hazard of incident cognitive frailty between participants with short sleep duration and moderate sleep duration in the initial model. Besides, there was no significant difference in the risk of incident cognitive frailty between participants with good and poor sleep quality. Additionally, RCS analysis showed no significant nonlinear association between sleep duration and cognitive frailty during follow up (*P* for nonlinear=.29; Figure 3).

XSL•FO
RenderX

Table . Association between sleep duration and incident cognitive frailty during follow up.

	Model A[a]		Model B[b]		Model C[c]	
	HR[d] (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Age	__[e]	—	1.09 (1.08 - 1.10)	<.001	1.09 (1.08 - 1.10)	<.001
Sex (female)	—	—	1.98 (1.60 - 2.45)	<.001	1.61 (1.25 - 2.08)	<.001
Not educated	—	—	1.84 (1.47 - 2.31)	<.001	1.72 (1.36 - 2.17)	<.001
Married and living with spouse	—	—	—	—	1.27 (1.01 - 1.59)	.04
Rural residence	—	—	—	—	1.03 (0.86 - 1.24)	.74
Economic dependence	—	—	—	—	1.47 (1.14 - 1.91)	.003
Loneliness	—	—	—	—	1.22 (1.01 - 1.47)	.04
Smoker	—	—	—	—	1.02 (0.80 - 1.30)	.85
Drinker	—	—	—	—	0.77 (0.61 - 0.97)	.03
Multimorbidity	—	—	—	—	1.33 (0.96 - 1.83)	.08
Poor sleep quality	1.32 (1.08 - 1.60)	.006	1.22 (1.00 - 1.49)	.05	1.14 (0.93 - 1.39)	.22
Sleep duration (hours)						
Moderate (6 - 9)	reference	reference	reference	reference	reference	reference
Short (<6)	0.96 (0.72 - 1.27)	.76	0.90 (0.68 - 1.21)	.49	0.94 (0.70 - 1.25)	.66
Long (>9)	1.74 (1.42 - 2.14)	<.001	1.32 (1.07 - 1.62)	.008	1.30 (1.06 - 1.59)	.01

[a]Model was unadjusted.

[b]Model was adjusted for age, sex, and education at baseline.

[c]Model was adjusted for age, sex, education, marital status, residence, economic status, loneliness, smoking status, drinking status and multimorbidity at baseline.

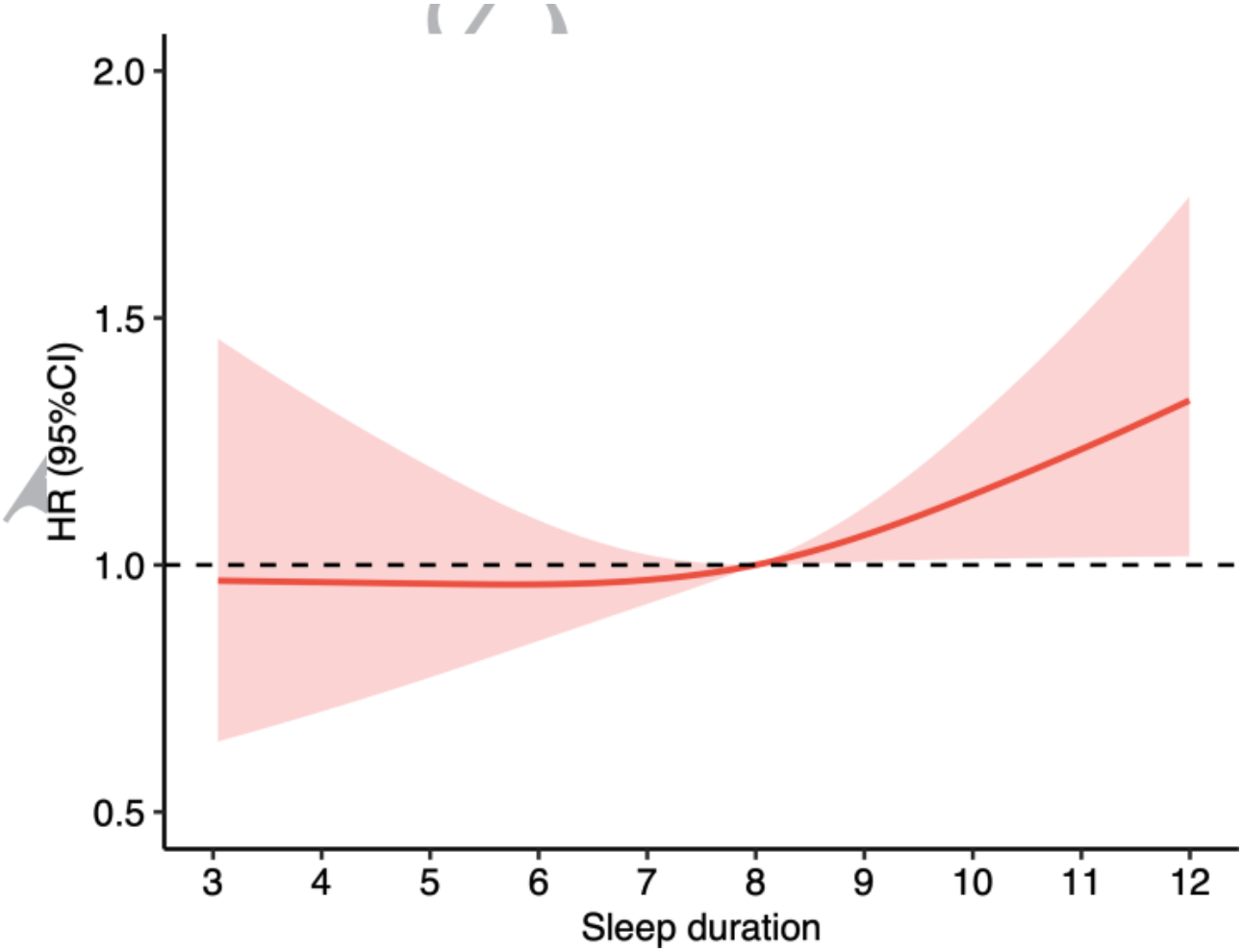
[d]HR: hazard ratio.

[e]Not applicable.

The results of sensitivity analyses are summarized in Multimedia Appendices 5 and 6. The associations between long sleep duration and baseline cognitive frailty status were consistent when adjusting for each chronic disease individually, or when adjusting for all chronic diseases simultaneously (Multimedia Appendix 5). For individuals younger than 80 years old, the association of long sleep duration with cognitive frailty became not statistically significant, whereas for individuals older than 80 years old, the association persisted. The association was not significant in male participants but remained statistically significant in female participants (Multimedia Appendix 6).

XSL•FO
RenderX

Figure 3. Nonlinear association between sleep duration and cognitive frailty during follow up using restricted cubic spline. The Cox proportional model was used to fit the data, with solid lines representing the estimated HR and shaded areas representing the corresponding 95% CIs. HR: hazard ratio.



Rate of Change in Sleep Duration and Cognitive Frailty

The mean annual rate of change in sleep duration was –0.02 (SD 0.87) hours/year. In the model adjusted for age, sex, education, sleep quality, and sleep duration at baseline, a faster increase in sleep duration per year was associated with a higher risk for cognitive frailty: for 1 hour/year increase in sleep duration, the HR was 1.14 (95% CI 1.03 - 1.27; *P*=.01). The association remained significant after further adjusting for marital status, residence, economic status, loneliness, smoking status, drinking status, and multimorbidity (HR 1.15, 95% CI 1.03 - 1.27; *P*=.01; Multimedia Appendix 7).

Discussion

Principal Findings

In this large community-based prospective study of over 10,000 Chinese older adults, we found that long sleep duration was consistently associated with cognitive frailty status at baseline, as well as an increased risk of cognitive frailty incident during follow up. We also found that a faster annual increase in sleep duration was associated with a higher risk of cognitive frailty during follow up.

The associations of sleep duration with physical frailty or cognitive decline are complex with mixed findings in prior research. A study using data from the National Health and Nutrition Examination Survey demonstrated that prolonged sleep duration (ie, ≥10 h) was associated with physical frailty in older adults, while a prospective study of 309 older Mexican adults showed that participants with either short sleep duration (ie, ≤5 h) or long sleep duration (ie, ≥9 h) had a higher risk of physical frailty [22,23]. Likewise, a Japanese cohort study of 623 older adults reported that long sleepers (>8 h) had a higher risk of cognitive impairment, while a cohort study found that self-reported short sleepers (<7 h) had an increased risk of cognitive decline [24,25]. However, few studies explored the association between sleep duration and cognitive frailty, that is, the coexistence of physical frailty and cognitive impairment. Two recent cross-sectional studies reported an association of long sleep duration with cognitive frailty among community-dwelling older adults and in older adults with heart failure [11,12]. In this study, we confirmed such a cross-sectional association between long sleep duration and cognitive frailty. More intriguingly, we showed consistent evidence from longitudinal analyses that participants with long sleep duration at baseline had an increased risk of developing cognitive frailty in the future. Of note, we did not find a significant association between short sleep duration and cognitive frailty in this study.

XSL•FO
RenderX

We also used the annual rate of change in sleep duration to assess the relationship between sleep duration changes and cognitive frailty. The results showed that a faster annual increase in sleep duration was associated with a higher risk of cognitive frailty, independent of the baseline sleep duration category. Notably, this study accounted for variations in follow-up years by using the annual rate of sleep duration change instead of the absolute change in sleep duration, minimizing classification bias and providing a more accurate representation of the relationship between sleep duration trajectories and cognitive frailty.

The mechanisms underlying the association of sleep duration with incident cognitive frailty remain unclear. However, several potential pathways can be suggested. First, inappropriate sleep duration (either shorter or longer than ideal) has been linked to cardiovascular health [26], which has been reported as a risk factor for both physical frailty and cognitive decline [27,28]. Second, increased sleep duration has also been linked with elevations in C-reactive protein, a marker for systemic inflammation, and interleukin-6, a proinflammatory biomarker [29,30]. Prior studies have shown that even low-level inflammation may pose an increased risk for the development of physical frailty and cognitive impairment [31,32]. Finally, long sleep duration may also be an indicator of or linked to disturbances in circadian control, the internal biological clock that prepares bodily responses to environmental light-dark changes. Circadian disturbances have been previously reported to be associated with both physical frailty and cognitive impairment or decline [33-36].

Several previous studies showed that sleep quality is closely related to impairments in both physical and cognitive domains [37,38]. In addition, 2 recent cross-sectional studies reported that poor sleep quality was associated with cognitive frailty in older Chinese adults living in nursing homes, and among Thai community-dwelling older adults during the COVID-19 restrictions [39,40]. Our observed cross-sectional association between sleep quality and cognitive frailty aligns with these previous findings. However, we did not find evidence of an association between sleep quality and the development of cognitive frailty in this study. It is important to note that the assessment of sleep quality in this study relied on a single question, which may be overly simplistic and may not fully capture the complexity of sleep behaviors. Future studies should use more comprehensive approaches to assess sleep health, considering its multidimensional nature, which includes factors such as duration, timing, efficiency, satisfaction, and alertness [41].

Strengths and Limitations

To the best of our knowledge, this study was the first to assess the association between sleep duration and incident cognitive frailty in a nationally representative sample of older Chinese adults. Strengths of this study include its longitudinal design and a large sample size. This study also has several limitations. First, the sleep duration and sleep quality were from self-reports and may be subjected to recall biases, especially in older adults and individuals with cognitive impairment. Future studies, ideally with an objective sleep monitoring approach (such as combining actigraphy and sleep diary or polysomnography), are warranted. Second, a growing body of studies have linked daytime napping behaviors with physical frailty, cognitive decline, Alzheimer dementia, or cognitive frailty [40,42,43]. Similarly, sleep disorders, such as sleep apnea, have been consistently linked with physical frailty, cognitive impairment, and dementia [10,44]. However, these potential sleep-related influences were not collected in the CLHLS. Future studies may consider incorporating these multidimensional structs when assessing sleep health and its relationships with physical and cognitive outcomes. Finally, the participants in this study were, on average, older than 80 years at baseline. In subgroup analyses, the association between long sleep duration and cognitive frailty was not statistically significant in participants younger than 80 years. This may be due to the limited number of cognitive frailty events in this subgroup, which could have resulted in insufficient statistical power to detect a significant association. Therefore, caution should be taken when translating our findings into younger populations.

Conclusions

In this large cohort study of over 10,000 older Chinese adults, long sleep duration was consistently associated with an increased risk of cognitive frailty, both cross-sectionally and longitudinally. These findings underscore the importance of sleep monitoring in aging populations. Future research using objective sleep assessments and investigating the underlying biological mechanisms is warranted to inform more effective prevention strategies.

Acknowledgments

This work is supported by the Key Project of Philosophy and Social Science Research in Colleges and Universities in Jiangsu Province (2022SJZD141 to JC) and the BrightFocus Foundation Alzheimer's Research Program (A2020886S to PL). PL is additionally supported by the Funds to Sustain Research Excellence from the Brigham Research Institute and a start-up fund from the Department of Anesthesia, Critical Care and Pain Medicine at Massachusetts General Hospital. CG is additionally supported by Alzheimer's Association (AARFD-22 - 928372). KH is additionally supported by the National Institute on Aging (RF1AG064312; R01AG083799) and a start-up fund from the Department of Anesthesia, Critical Care and Pain Medicine at Massachusetts General Hospital. LG is additionally supported by the Alzheimer's Association (AACSF-23 - 1148490).

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Baseline characteristics of participants by sleep duration groups
[DOCX File, 20 KB - aging_v8i1e65183_app1.docx]

Multimedia Appendix 2
Baseline characteristics of participants by incident cognitive frailty status during follow-up.
[DOCX File, 20 KB - aging_v8i1e65183_app2.docx]

Multimedia Appendix 3
Association between sleep and cognitive frailty at baseline with individual chronic disease adjustments.
[DOCX File, 19 KB - aging_v8i1e65183_app3.docx]

Multimedia Appendix 4
Association between sleep and baseline cognitive frailty stratified by age and sex.
[DOCX File, 19 KB - aging_v8i1e65183_app4.docx]

Multimedia Appendix 5
Association between sleep and incident cognitive frailty with individual chronic disease adjustments.
[DOCX File, 19 KB - aging_v8i1e65183_app5.docx]

Multimedia Appendix 6
Association between sleep and incident cognitive frailty stratified by age and sex.
[DOCX File, 19 KB - aging_v8i1e65183_app6.docx]

Multimedia Appendix 7
Rate of change in sleep duration and cognitive frailty.
[DOCX File, 18 KB - aging_v8i1e65183_app7.docx]

References

1. Bae JB, Han JW, Kwak KP, et al. Impact of mild cognitive impairment on mortality and cause of death in the elderly. J Alzheimers Dis 2018;64(2):607-616. [doi: 10.3233/JAD-171182] [Medline: 29914024]
2. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet 2013 Mar 2;381(9868):752-762. [doi: 10.1016/S0140-6736(12)62167-9] [Medline: 23395245]
3. Boyle PA, Buchman AS, Wilson RS, Leurgans SE, Bennett DA. Physical frailty is associated with incident mild cognitive impairment in community-based older persons. J Am Geriatr Soc 2010 Feb;58(2):248-255. [doi: 10.1111/j.1532-5415.2009.02671.x] [Medline: 20070417]
4. Song X, Mitnitski A, Rockwood K. Nontraditional risk factors combine to predict Alzheimer disease and dementia. Neurology (ECronicon) 2011 Jul 19;77(3):227-234. [doi: 10.1212/WNL.0b013e318225c6bc] [Medline: 21753161]
5. Kelaiditi E, Cesari M, Canevelli M, et al. Cognitive frailty: rational and definition from an (I.A.N.A./I.A.G.G.) International Consensus Group. J Nutr Health Aging 2013 Nov;17(9):726-734. [doi: 10.1007/s12603-013-0367-2]
6. Feng L, Nyunt MSZ, Gao Q, et al. Physical frailty, cognitive impairment, and the risk of neurocognitive disorder in the Singapore longitudinal ageing studies. GERONA 2016 Mar 24:glw050. [doi: 10.1093/gerona/glw050]
7. Solfrizzi V, Scafato E, Lozupone M, et al. Additive role of a potentially reversible cognitive frailty model and inflammatory state on the risk of disability: the Italian longitudinal study on aging. Am J Geriatr Psychiatry 2017 Nov;25(11):1236-1248. [doi: 10.1016/j.jagp.2017.05.018] [Medline: 28689645]
8. Li J, Vitiello MV, Gooneratne NS. Sleep in normal aging. Sleep Med Clin 2018 Mar;13(1):1-11. [doi: 10.1016/j.jsmc.2017.09.001] [Medline: 29412976]
9. Bubu OM, Brannick M, Mortimer J, et al. Sleep, cognitive impairment, and Alzheimer's disease: a systematic review and meta-analysis. Sleep 2017 Jan 1;40(1). [doi: 10.1093/sleep/zsw032] [Medline: 28364458]
10. Sun R, Xie Y, Jiang W, Wang E, Li X. Effects of different sleep disorders on frailty in the elderly: a systematic review and meta-analysis of observational studies. Sleep Breath 2023 Mar;27(1):91-101. [doi: 10.1007/s11325-022-02610-5] [Medline: 35416613]
11. Zhao Y, Lu Y, Zhao W, et al. Long sleep duration is associated with cognitive frailty among older community-dwelling adults: results from West China Health and Aging Trend study. BMC Geriatr 2021 Oct 27;21(1):608. [doi: 10.1186/s12877-021-02455-9] [Medline: 34706663]
12. Seo EJ, Son YJ. The prevalence of cognitive frailty and its association with sleep duration and depression among older adults with heart failure. Clin Gerontol 2024;47(3):416-425. [doi: 10.1080/07317115.2022.2125858] [Medline: 36117469]

XSL•FO
RenderX

13. Salinas-Rodríguez A, Manrique-Espinoza B, Moreno-Tamayo K, Guerrero-Zúñiga S. Trajectories of sleep duration and quality and their association with mild cognitive impairment, frailty, and all-cause mortality. Sleep Health 2024 Apr;10(2):240-248. [doi: 10.1016/j.sleh.2023.12.002] [Medline: 38238122]
14. Zeng Y. Chinese longitudinal healthy longevity survey and some research findings. Geriatrics Gerontol Int 2004 Sep;4(s1). [doi: 10.1111/j.1447-0594.2004.00147.x]
15. Zeng Y, Feng Q, Hesketh T, Christensen K, Vaupel JW. Survival, disabilities in activities of daily living, and physical and cognitive functioning among the oldest-old in China: a cohort study. Lancet 2017 Apr;389(10079):1619-1629. [doi: 10.1016/S0140-6736(17)30548-2]
16. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
17. Wang HY, Lv X, Du J, Kong G, Zhang L. Age- and gender-specific prevalence of frailty and its outcomes in the longevous population: the Chinese longitudinal healthy longevity study. Front Med (Lausanne) 2021;8:719806. [doi: 10.3389/fmed.2021.719806] [Medline: 34409056]
18. Yi Z, Vaupel JW. Functional capacity and self–evaluation of health and life of oldest old in China. J Soc Issues 2002 Jan;58(4):733-748. [doi: 10.1111/1540-4560.00287]
19. Li J, Cao D, Huang Y, et al. Sleep duration and health outcomes: an umbrella review. Sleep Breath 2022 Sep;26(3):1479-1501. [doi: 10.1007/s11325-021-02458-1] [Medline: 34435311]
20. Vatanabe IP, Pedroso RV, Teles RHG, et al. A systematic review and meta-analysis on cognitive frailty in community-dwelling older adults: risk and associated factors. Aging Ment Health 2022 Mar 4;26(3):464-476. [doi: 10.1080/13607863.2021.1884844]
21. Yuan C, Zhang Y. Modifiable risk factors of cognitive frailty in community-dwelling older adults: a systematic review and meta-analysis. Int J Nurs Knowl 2024 Apr;35(2):177-185. [doi: 10.1111/2047-3095.12437] [Medline: 37378541]
22. Baniak LM, Yang K, Choi J, Chasens ER. Long sleep duration is associated with increased frailty risk in older community-dwelling adults. J Aging Health 2020;32(1):42-51. [doi: 10.1177/0898264318803470] [Medline: 30270714]
23. Moreno-Tamayo K, Manrique-Espinoza B, Morales-Carmona E, Salinas-Rodríguez A. Sleep duration and incident frailty: the rural frailty study. BMC Geriatr 2021 Jun 16;21(1):368. [doi: 10.1186/s12877-021-02272-0] [Medline: 34134643]
24. Keil SA, Schindler AG, Wang MX, et al. Longitudinal sleep patterns and cognitive impairment in older adults. JAMA Netw Open 2023 Dec 1;6(12):e2346006. [doi: 10.1001/jamanetworkopen.2023.46006] [Medline: 38048131]
25. Kinoshita K, Otsuka R, Takada M, et al. Dietary amino acid intake and sleep duration are additively involved in future cognitive decline in Japanese adults aged 60 years or over: a community-based longitudinal study. BMC Geriatr 2023 Oct 11;23(1):653. [doi: 10.1186/s12877-023-04359-2] [Medline: 37821805]
26. Lloyd-Jones DM, Allen NB, Anderson CAM, et al. Life's essential 8: updating and enhancing the American Heart Association's construct of cardiovascular health: a presidential advisory from the American Heart Association. Circulation 2022 Aug 2;146(5):e18-e43. [doi: 10.1161/CIR.0000000000001078] [Medline: 35766027]
27. Palta P, Griswold M, Ranadive R, et al. Midlife cardiovascular health and robust versus frail late-life status: the atherosclerosis risk in communities study. J Gerontol A Biol Sci Med Sci 2022 Jun 1;77(6):1222-1229. [doi: 10.1093/gerona/glab310] [Medline: 34661638]
28. Cho S, Yang PS, Kim D, et al. Association of cardiovascular health with the risk of dementia in older adults. Sci Rep 2022 Sep 19;12(1):15673. [doi: 10.1038/s41598-022-20072-3]
29. Irwin MR, Olmstead R, Carroll JE. Sleep disturbance, sleep duration, and inflammation: a systematic review and meta-analysis of cohort studies and experimental sleep deprivation. Biol Psychiatry 2016 Jul 1;80(1):40-52. [doi: 10.1016/j.biopsych.2015.05.014] [Medline: 26140821]
30. Patel SR, Zhu X, Storfer-Isser A, et al. Sleep duration and biomarkers of inflammation. Sleep 2009 Feb;32(2):200-204. [doi: 10.1093/sleep/32.2.200]
31. Singh-Manoux A, Dugravot A, Brunner E, et al. Interleukin-6 and C-reactive protein as predictors of cognitive decline in late midlife. Neurology (ECronicon) 2014 Aug 5;83(6):486-493. [doi: 10.1212/WNL.0000000000000665] [Medline: 24991031]
32. Wang J, Maxwell CA, Yu F. Biological processes and biomarkers related to frailty in older adults: a state-of-the-science literature review. Biol Res Nurs 2019 Jan;21(1):80-106. [doi: 10.1177/1099800418798047] [Medline: 30198309]
33. Cai R, Gao L, Gao C, et al. Circadian disturbances and frailty risk in older adults. Nat Commun 2023 Nov 16;14(1):7219. [doi: 10.1038/s41467-023-42727-z] [Medline: 37973796]
34. Gao L, Li P, Gaykova N, et al. Circadian rest-activity rhythms, delirium risk, and progression to dementia. Ann Neurol 2023 Jun;93(6):1145-1157. [doi: 10.1002/ana.26617] [Medline: 36808743]
35. Li P, Gao L, Gaba A, et al. Circadian disturbances in Alzheimer's disease progression: a prospective observational cohort study of community-based older adults. Lancet Healthy Longevity 2020 Dec;1(3):e96-e105. [doi: 10.1016/s2666-7568(20)30015-5] [Medline: 34179863]
36. Li P, Hu K. Circadian disturbances and age-related disorders. Adv Biol (Weinh) 2023 Nov;7(11):e2300405. [doi: 10.1002/adbi.202300405] [Medline: 37986670]

XSL•FO
RenderX

37. Gildner TE, Liebert MA, Kowal P, Chatterji S, Snodgrass JJ. Associations between sleep duration, sleep quality, and cognitive test performance among older adults from six middle income countries: results from the Study on Global Ageing and Adult Health (SAGE). J Clin Sleep Med 2014 Jun 15;10(6):613-621. [doi: 10.5664/jcsm.3782] [Medline: 24932140]
38. Wai JT, Yu DF. The relationship between sleep-wake disturbances and frailty among older adults: a systematic review. J Adv Nurs 2020 Jan;76(1):96-108. [doi: 10.1111/jan.14231] [Medline: 31588595]
39. Griffiths J, Seesen M, Sirikul W, Siviroj P. Malnutrition, depression, poor sleep quality, and difficulty falling asleep at night are associated with a higher risk of cognitive frailty in older adults during the COVID-19 restrictions. Nutrients 2023 Jun 23;15(13):2849. [doi: 10.3390/nu15132849] [Medline: 37447178]
40. Liu S, Hu Z, Guo Y, Zhou F, Li S, Xu H. Association of sleep quality and nap duration with cognitive frailty among older adults living in nursing homes. Front Public Health 2022 Aug 25;10:963105. [doi: 10.3389/fpubh.2022.963105]
41. Buysse DJ. Sleep health: can we define it? Does it matter? Sleep 2014 Jan 1;37(1):9-17. [doi: 10.5665/sleep.3298]
42. Li P, Gao L, Yu L, et al. Daytime napping and Alzheimer's dementia: a potential bidirectional relationship. Alzheimers Dement 2023 Jan;19(1):158-168. [doi: 10.1002/alz.12636] [Medline: 35297533]
43. Zhang Y, Zhou L, Ge M, Lin X, Dong B. Association between daytime nap duration and risks of frailty: findings from the China Health and Retirement Longitudinal Study. Front Public Health 2022;10:1098609. [doi: 10.3389/fpubh.2022.1098609] [Medline: 36777767]
44. Bubu OM, Andrade AG, Umasabor-Bubu OQ, et al. Obstructive sleep apnea, cognition and Alzheimer's disease: a systematic review integrating three decades of multidisciplinary research. Sleep Med Rev 2020 Apr;50:101250. [doi: 10.1016/j.smrv.2019.101250] [Medline: 31881487]

Abbreviations

C-MMSE: Chinese version of the Mini-Mental State Examination
CLHLS: Chinese Longitudinal Healthy Longevity Survey
HR: hazard ratio
OR: odds ratio
PH: proportional hazard
RCS: restricted cubic spline

Edited by F Puga; submitted 13.08.24; peer-reviewed by IC Wu, WQ Song; revised version received 10.03.25; accepted 20.03.25; published 23.04.25.

Please cite as:
Cai R, Chao J, Gao C, Gao L, Hu K, Li P
Association Between Sleep Duration and Cognitive Frailty in Older Chinese Adults: Prospective Cohort Study
JMIR Aging 2025;8:e65183
URL: https://aging.jmir.org/2025/1/e65183
doi:10.2196/65183

© Ruixue Cai, Jianqian Chao, Chenlu Gao, Lei Gao, Kun Hu, Peng Li. Originally published in JMIR Aging (https://aging.jmir.org), 23.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Identification of Target Body Composition Parameters by Dual-Energy X-Ray Absorptiometry, Bioelectrical Impedance, and Ultrasonography to Detect Older Adults With Frailty and Prefrailty Status Using a Mobile App in Primary Care Services: Descriptive Cross-Sectional Study

Beatriz Ortiz-Navarro[1*], MD; José Losa-Reyna[2,3*], PhD; Veronica Mihaiescu-Ion[4,5*], PT, MSc; Jerónimo Garcia-Romero[6*], MD, Prof Dr; Margarita Carrillo de Albornoz-Gil[6*], MD, PhD; Alejandro Galán-Mercant[4,5*], PhD, PT, Prof Dr

[1]Public Andalusian Health System, Málaga, Spain

[2]CIBER of Frailty and Healthy Aging, Madrid, Spain

[3]Sports Science Research Centre, King Juan Carlos University, Madrid, Spain

[4]Institute of Biomedicine of Cádiz, Cádiz, Spain

[5]MOVE-IT Research Group, Department of Nursing and Physiotherapy, Faculty of Health Sciences, University of Cádiz, Cádiz, Spain

[6]Área de Educación Física y Deportiva, Facultad de Medicina, Universidad de Málaga, Boulevard Louis Pasteur, Málaga, Spain

[*]all authors contributed equally

Corresponding Author:
Margarita Carrillo de Albornoz-Gil, MD, PhD
Área de Educación Física y Deportiva, Facultad de Medicina, Universidad de Málaga, Boulevard Louis Pasteur, Málaga, Spain

Abstract

Background: Frailty syndrome in older adults represents a significant public health concern, characterized by a reduction in physiological reserves and an increased susceptibility to stressors. This can result in adverse health outcomes, including falls, hospitalization, disability, and mortality. The early identification and management of frailty are essential for improving quality of life and reducing health care costs. Conventional assessment techniques, including dual-energy X-ray absorptiometry (DXA), bioelectrical impedance analysis (BIA), and muscle ultrasound (US), are efficacious but frequently constrained in primary care settings by financial and accessibility limitations.

Objective: The aim of this study is to analyze the differences in anthropometric characteristics, physical function, nutritional status, cognitive status, and body composition among older adults identified as frail, prefrail, or robust in primary care services using the PowerFrail mobile app. Furthermore, the study assesses the predictive capacity of body composition variables (whole-body phase angle [WBPhA] via BIA, US-measured rectus femoris muscle thickness, and DXA-derived lean mass) in identifying frailty and evaluates their feasibility for implementation in primary care.

Methods: A descriptive cross-sectional study was conducted with 94 older adult participants aged between 70 and 80 years, recruited through the Andalusian Health Service in Spain. Frailty status was classified using the PowerFrail App, which integrates muscle power assessment and provides personalized physical activity recommendations. Body composition was measured using WBPhA (BIA), muscle US, and DXA. Statistical analyses included 1-way ANOVA for group comparisons, logistic regression to investigate associations, and receiver operating characteristic curve analysis to evaluate the predictive accuracy of the body composition measures.

Results: Participants were categorized into frail (n=28), prefrail (n=33), and robust (n=33) groups. All body composition measures exhibited high specificity in detecting frailty, with varying sensitivity. Unadjusted US showed the highest specificity but low sensitivity (10.7%). WBPhA and right leg lean mass (LeanM RL) demonstrated significant predictive capabilities, especially when adjusted for age and sex, with area under the curve values ranging from 0.678 to 0.762. The adjusted LeanM RL model showed a good balance between sensitivity (35.7%) and specificity (93.9%; *P*=.045), indicating its potential as a reliable frailty predictor. These findings are consistent with previous research emphasizing the importance of muscle mass and cellular health in frailty assessment.

Conclusions: Body composition variables, particularly WBPhA, LeanM RL, and US, are effective predictors of frailty in older adults. The PowerFrail mobile app, combined with advanced body composition analysis, offers a practical and noninvasive method

XSL•FO
RenderX

for early frailty detection in primary care settings. Integrating such technological tools can enhance the early identification and management of frailty, thereby improving health outcomes in the aging population.

(JMIR Aging 2025;8:e67982) doi:10.2196/67982

KEYWORDS

frailty syndrome; older adults; body composition; bioelectrical impedance analysis; muscle ultrasound, dual-energy X-Ray absorptiometry; mobile health apps; primary care; PowerFrail App

Introduction

Frailty syndrome in older adults is a multifaceted clinical condition characterized by decreased physiological reserves and increased vulnerability to stressors, which elevates the risk of adverse health outcomes such as falls, hospitalization, disability, and mortality [1,2]. As the global population ages, the prevalence of frailty is projected to rise, posing significant challenges to health care systems worldwide [3]. Early identification and management of frailty are crucial for enhancing quality of life, maintaining functional independence, and reducing health care costs associated with frailty-related complications [4,5].

Primary care facilities are uniquely positioned to play a pivotal role in the early detection and management of frailty due to their accessibility and continuous engagement with the aging population [6]. Implementing effective screening tools within primary care can facilitate timely interventions, thereby mitigating the progression of frailty and its associated adverse outcomes [7]. However, the integration of comprehensive frailty assessments into routine primary care practice remains limited, often due to time constraints, lack of standardized tools, and insufficient training among primary care providers [8].

Body composition variables, particularly those related to muscle mass and tissue quality, are essential in identifying and predicting frailty syndrome [9]. Techniques such as dual-energy X-ray absorptiometry (DXA), bioelectrical impedance analysis (BIA), and muscle ultrasound (US) have been extensively used to assess body composition and muscle status in older adults [10]. The whole-body phase angle (WBPhA), obtained through BIA, serves as an indicator of cellular health and nutritional status, with lower values associated with increased frailty and poorer clinical outcomes [11,12]. Similarly, muscle US offers a noninvasive way to assess muscle thickness and quality, aiding in the detection of sarcopenia and frailty [13]. Although DXA is considered the gold standard for measuring bone and muscle mass, its high cost and limited accessibility in primary care facilities make it necessary to explore alternative assessment methods [14].

In recent years, the advent of mobile health (mHealth) apps has introduced innovative solutions for the assessment and monitoring of complex geriatric syndromes, including frailty. The PowerFrail app represents a pioneering effort in this domain, being the first clinical and scientifically validated app designed to assess muscle power and frailty in older adults in a user-friendly manner [15,16]. This app not only facilitates the screening process but also provides individualized recommendations for improvement and tailored physical activity regimens, thereby supporting personalized intervention strategies. The utilization of validated mobile apps and new trends around artificial intelligence in primary care hold significant promise for enhancing the early detection and management of frailty, offering advantages such as accessibility, ease of use, and the ability to provide real-time feedback and recommendations [17,18].

Moreover, integrating mHealth tools into primary care can bridge gaps in health care delivery by enabling continuous monitoring and follow-up, which are critical for managing chronic conditions and preventing the escalation of frailty [18]. Previous studies have highlighted the effectiveness of mobile apps in improving health outcomes among older adults by facilitating timely interventions and enhancing patient engagement [19]. These findings underscore the potential of mHealth solutions to complement traditional assessment methods, providing a comprehensive approach to frailty management in primary care settings.

From above, the main objective of this study is to analyze the differences in variables related to anthropometric characteristics, physical function, nutritional status, cognitive status, and body composition in phenotypes of frail, prefrail, and robust older adults identified in primary care services. Frailty levels will be classified using the PowerFrail app. Additionally, the second objective is to assess the predictive capacity of body composition variables (US, bioimpedance, and DXA) in identifying older adults with frailty and to evaluate their implementation in primary care services. We used the technological tools WBPhA, US for rectus femoris muscle thickness, and DXA for bone mineral density and lean mass assessment. Our results could demonstrate that these parameters are useful predictors for the identification of frailty, in line with previous findings, thus supporting the potential integration of these tools into primary care practices for the early detection and management of frailty, integrating for the first time mHealth technologies with advanced body composition analysis systems.

Methods

Ethical Considerations

This was a descriptive cross-sectional study that evaluated clinical, physiological, body composition, and psychometric variables in a sample of older adult participants. Recruitment and data collection took place between March and June of 2023. Participants were recruited through advertisements from the public Andalusian Health Service system, Spain. Participants were recruited from primary care centers through referrals by health care professionals. Recruitment was conducted in collaboration with general practitioners, who identified potential

XSL•FO
RenderX

participants meeting the inclusion criteria. The study adhered to the ethical principles of the Declaration of Helsinki for medical research involving human subjects. All participants provided written informed consent. This study was approved by the Costa del Sol Institutional Ethics Committee, under protocol number BON22, on December 23, 2022. All participants provided written informed consent prior to their inclusion in the study. Participants' privacy and confidentiality were ensured throughout the study: all data were anonymized prior to analysis, and identifying information was stored securely and separately from the study data. No identifiable images or information of participants are included in this publication. Participants did not receive compensation for their participation..

Participants

Inclusion criteria for the study were as follows:

- Individuals aged between 70 and 80 years, inclusive.
- Absence or presence of mild cognitive impairment as determined by a Barthel index score [20] greater than 95.
- Sit-to-stand test score [21] between 2.5 and 3.6 for men or between 1.9 and 3 for women.
- Signed informed consent.

Exclusion criteria were the following:

- Individuals younger than 70 years or older than 80 years.
- Individuals with moderate or severe cognitive impairment.
- Sit-to-stand test score [21] less than 2.5 or greater than 3.6 for men or less than 1.9 or greater than 3 for women.
- Individuals living in institutions.
- Individuals with pacemakers or metal prostheses.

Variables and Procedure

The aim of this study was to identify differences in measured parameters to classify individuals into different states of frailty within primary care. To identify different states of frailty, the sit-to-stand test was used, and relative power was calculated using the equations validated by Losa-Reyna et al [16]. Using the cutoff points determined by Losa-Reyna et al [22], participants were classified into 3 groups: frail, prefrail, and robust. This test was performed in the health center's office.

After identifying the study subjects, a detailed evaluation of the remaining variables was carried out in the Functional Testing Laboratory of the Physical Education and Sports Area of the University of Málaga. Participants came to the laboratory fasting (a minimum of 3 hours) and without having done any previous exercise, wearing comfortable and light clothing without metal objects. The tests were carried out between 9:30 AM and 2:00 PM. Once in the laboratory, 2 qualified researchers administered the different tests and assessments; body composition determinations were performed fasting and in the early morning, and the rest of the scheduled evaluations were carried out after breakfast.

Clinical and Demographic Variables

We collected the following clinical and demographic variables:

- Age and sex: Recorded during the initial interview (age in years; sex as male or female).
- Body mass index: Calculated by dividing weight (measured in kilograms using a Digital Scale Extra Large Seca Robusta 813) by height (measured in meters using a stadiometer) squared (m^2).
- Number of medications (drugs): Extracted from the clinical history of the Diraya system (Andalusian Health Service database), with the participants' consent.
- Education level: Recorded during the initial interview.

Body Composition Variables

We collected the following body composition variables:

- Muscle architecture: Muscle thickness and pennation angle of the rectus femoris muscle of the dominant leg were measured using US (Logiq Book XP Ultrasound System and 8L linear transducer). The rectus femoris was examined with the participant in a supine position, with the US operator standing on the ipsilateral side of the participant. The evaluation protocol previously described by Mateos-Angulo et al [23] was used.
- DXA: Lean mass of arms and legs, total lean body mass, skeletal muscle index, and bone mineral density were determined using a Hologic Horizon A DXA scanner (Hologic Inc). Each subject was examined by a certified technician. The distinction between bone and soft tissue, the detection of edges, and regional demarcations were performed using computer algorithms with APEX Corporation Software (version 5.6.0.7). For each scan, patients were asked to remove all materials that could attenuate the X-ray beam, including jewelry. Due to the sensitivity of the soft tissue analysis, the patient should only wear a paper gown for the scan. There should be no pillow on the scan, as the material would affect the soft tissue measurement. The densitometer calibration was checked daily with the standard calibration block supplied by the manufacturer.
- BIA: Total body water, intracellular, extracellular water, and phase angle (WBPhA) variables were determined by multifrequency bioimpedance using the Inbody 770 model. Multifrequency segmental data were obtained that accurately determined total body water, intracellular and extracellular water, impedance (Xc and R), and phase angle (Z) in the 5 body segments (right arm, left arm, trunk, right leg, and left leg).

Cognitive and Nutritional Status

To assess cognitive and nutritional status, the following questionnaires were administered digitally using the Google Forms application.

- Cognitive capacity: Cognitive status was evaluated [24] using the General Practitioner Assessment of Cognition, score 0 to 8, a rapid, reliable, and specific test for the detection of dementia in primary care. This instrument is considered an efficient alternative to others, such as the Mini-Mental State Examination, due to its rapid administration and lack of bias related to gender, education level, or mental health.
- Nutritional status screening: Nutritional status was assessed using the Mini Nutritional Assessment (MNA) [25], score

XSL•FO
RenderX

0 to 14, a widely used scale in the geriatric population. The first part of this test serves as a screening tool to detect the risk of malnutrition, with a cutoff value of 10 points or less. It is a widely used method in older adults and has been validated in different clinical contexts.

Physical Function Evaluation

To evaluate the physical function of the participants, the Short Physical Performance Battery (SPPB) was used, score 0 to 12 [26]. This battery includes three tests:

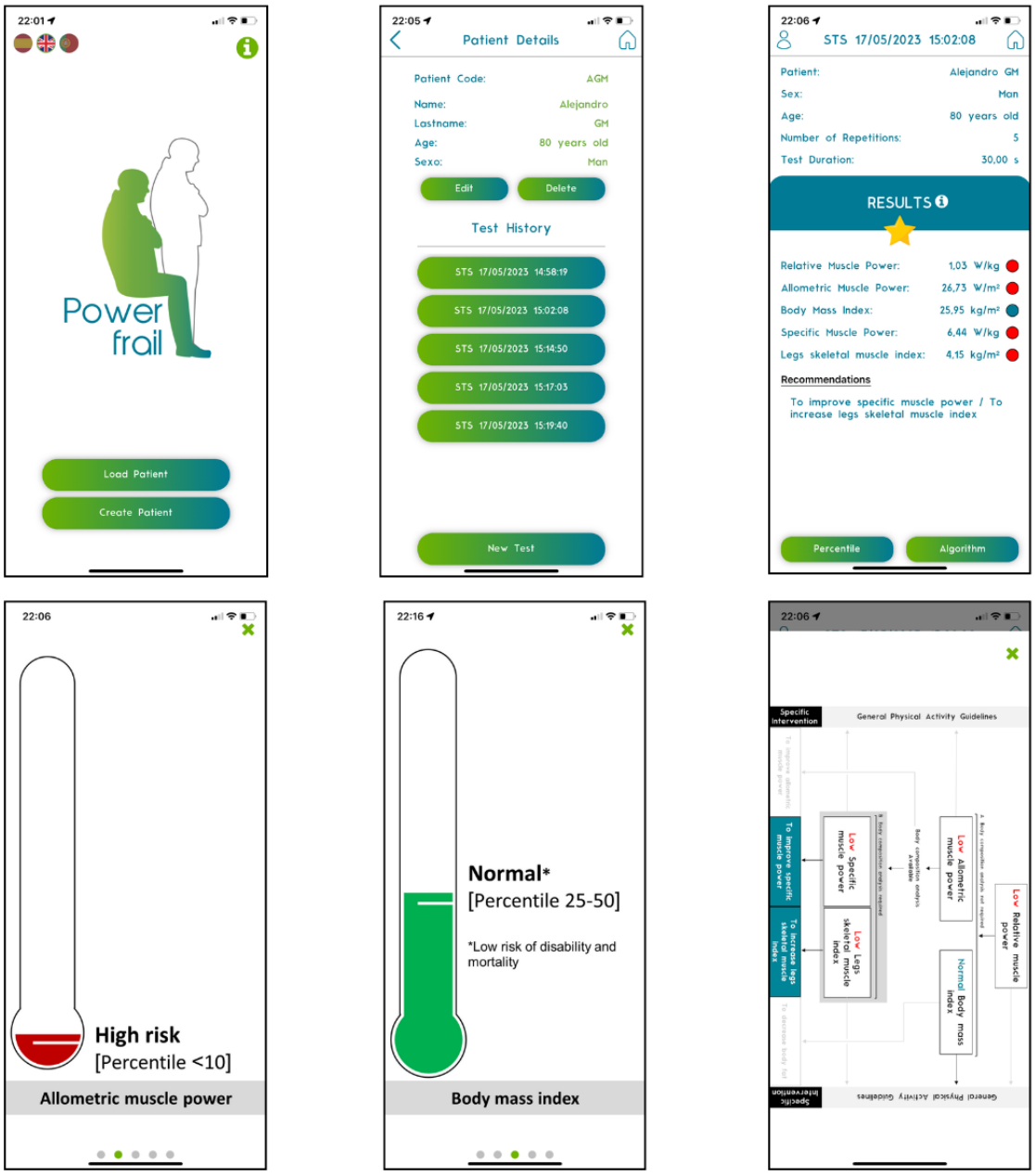
- Walking speed: Measured over a 4-meter distance, expressed in meters per second. Given that gait speed measurement is used as a tool for detecting frailty in older adults from the general population due to its high sensitivity, simplicity, and feasibility, we also used the cutoff points for frailty and prefrailty related to gait speed and sarcopenia in accordance with the consensus document developed by Cruz-Jentoft et al [27]. In this framework, values below 0.6 m/s are identified as frail, between 0.6 and 1 m/s as prefrail, and above 1 m/s as robust [27].
- Static balance: Assessed in 3 different positions.
- Sit-to-stand test: Evaluation of the time it takes for participants to stand up and sit down from a chair 5 times.

In addition to the tests included in the SPPB, two tests were performed:

- Mobile lower limb relative muscle power (RPOW): To measure RPOW, participants performed 5 repetitions of standing up from and sitting down onto a chair with a height of 0.46 m, following the protocol validated by Alcazar et al [15]. The test was performed using the PowerFrail app (Figure 1), developed and validated by Losa-Reyna et al [16], installed on a stable smartphone.
- Handgrip strength: Isometric handgrip strength, expressed in kilograms, was measured using a Takei Physical Fitness Test adjustable dynamometer, following a standardized protocol, as indicated by Roberts et al [28]. Participants performed the test in an erect standing position, with shoulders adducted and arms extended parallel to the body, without touching their torso. Two attempts were made for each extremity, and the maximum value was considered, regardless of hand dominance.

XSL•FO
RenderX

Figure 1. PowerFrail mobile app screenshots.



Sample Size Calculation

The sample size was estimated using an ANOVA model for 3 independent groups (robust, prefrail, and frail). The calculation was performed using the G*Power software. For the sample size calculation, a power of 90% (1–β=.90), a significance level of 5% (α=.05), and an effect size of 0.4 were used. The result was a total sample size of 84 participants. This size was obtained with a noncentrality parameter (λ) of 13.44 and a critical *F* of 3.11, with 2 degrees of freedom in the numerator and 81 in the denominator. To avoid losses in the initial calculation, enough participants were recruited to maintain the robustness of the study. Finally, 94 subjects participated in the study, distributed across 3 groups: robust, prefrail, and frail.

Statistical Analysis

To compare the differences on all variables between robust participants and those with frailty and prefrailty, a 1-way ANOVA was performed on the total sample. Logistic regression analysis was performed to investigate the relationship between body composition and frailty. WBPhA, thickness US, and DXA lean muscle from the lower limbs were entered into the regression model as independent variables, as they were found to have significant differences between the studied groups. Considering that age and sex may influence the relationship between body composition and frailty, these factors were introduced as a confounding variable into an adjusted regression model, treating age as a continuous covariate and sex as a categorical factor. The dependent variable was a binary indicator of frailty, coded as 1 for participants with frailty and 0 for robust participants and those with prefrailty.

XSL•FO
RenderX

To extract values for body composition to identify the presence of prefrailty or frailty, we conducted an analysis using the receiver operating characteristic (ROC) curve. This analysis was applied only to body composition that was significant in the logistic regression analysis. In the ROC analysis, the outcome variable was the presence or absence of frailty. The test variable was the body composition that was significantly associated with frailty. Youden index [29] was calculated with the following formula: Youden index = sensitivity + specificity – 1. The area under the curve (AUC), sensitivity, and specificity were calculated to evaluate the accuracy of the identified predictive models. The AUC could distinguish between nonpredictive (AUC<0.5), less predictive (0.5<AUC<0.7), moderately predictive (0.7<AUC<0.9), and highly predictive (0.9<AUC<1) values, as well as perfect prediction (AUC=1) [30].

Results

The background information of the participants is shown in Table 1. Of the 94 participants in this study, 28 were frail, 33 were prefrail, and 33 had a robust profile. The average age for each frailty category was 76.5, 75.3, and 74.0 years, respectively. The percentages of men and women with frailty and prefrailty were 32% and 52% for men and 68% and 48% for women, respectively. All investigated variables were significantly different among the frailty categories, except BMI, MNA score, skeletal mass index, left arm lean mass, right arm lean mass, and total body lean mass.

XSL•FO
RenderX

Table . Descriptive characteristics of the sample (n=94).

	Frailty[a]	Prefrailty	Robust	*P* value[b]	Post hoc[a]
Particpants, n (%)	28 (30)	33 (35)	33 (35)		
Sex (female; male), n (%)	19 (68); 9 (32)	16 (48); 17 (52)	17 (52); 16 (48)		
Age (years), mean (SD)	76.46 (2.92)	75.34 (2.86)	73.99 (2.59)	.004	*F*>R
BMI (kg/m^2), mean (SD)	28.35 (5.34)	27.87 (4.73)	28.65 (5.27)	.82	
RPOW[c] (W/kg), mean (SD)	1.69 (0.54)	2.60 (0.32)	3.32 (0.46)	<.001	*F*<P<R
MNA[d] (score, 0 - 14), mean (SD)	13.00 (1.28)	12.94 (1.71)	13.30 (0.98)	.52	
GPCOG[e] (score, 0 - 8), mean (SD)	5.46 (1.73)	6.60 (1.32)	7.15 (1.00)	<.001	*F*<P,R
SPPB[f] (score, 0 - 12), mean (SD)	8.07 (2.26)	10.79 (1.24)	11.75 (0.43)	<.001	*F*<P<R
5STS[g] (s), mean (SD)	18.47 (6.44)	12.01 (1.40)	9.27 (1.31)	<.001	*F*<P<R
WALK4m[h] (s), mean (SD)	5.92 (2.25)	4.07 (0.75)	3.92 (0.61)	<.001	*F*<P<R
GS4m[i] (m/s), mean (SD)	0.74 (0.21)	1.01 (0.16)	1.04 (0.15)	<.001	*F*<P,R
Handgrip max (kg), mean (SD)	19.24 (5.39)	25.48 (6.57)	26.25 (7.50)	<.001	*F*<P,R
Number of drugs[j], mean (SD)	6.89 (2.81)	4.73 (3.13)	3.09 (2.20)	<.001	*F*>P>R
US[k] (cm), mean (SD)	0.92 (0.32)	1.02 (0.25)	1.16 (0.26)	.004	*F*<R
WBPhA[l] (°), mean (SD)	4.36 (0.55)	4.63 (0.61)	4.95 (0.43)	<.001	*F*<R
SMI[m], mean (SD)	6.51 (1.04)	7.09 (1.08)	7.20 (1.33)	.07	
Left arm lean mass (g), mean (SD)	2088.24 (624.67)	2485.68 (720.65)	2408.92 (772.28)	.08	
Right arm lean mass (g), mean (SD)	2274.00 (597.82)	2624.04 (787.75)	2661.02 (915.64)	.12	
Left leg lean mass (g), mean (SD)	5975.11 (1494.27)	7147.72 (1794.43)	7158.07 (2192.25)	.02	*F*<P,R
Right leg lean mass (g), mean (SD)	5982.38 (1573.34)	7368.42 (1823.40)	7314.78 (2040.34)	.006	*F*<P,R
Total body lean mass (g), mean (SD)	42,130.41 (8174.96)	48,123.28 (11,063.12)	47,858.20 (12,130.27)	.06	

[a]F: frailty; P: prefrailty; R: robust.

[b]*P* value adjusted for comparing a family of 3.

[c]RPOW: relative muscle power.

[d]MNA: Mini Nutritional Assessment.

[e]GPCOG: General Practitioner Assessment of Cognition.

[f]SPPB: Short Physical Performance Battery.

[g]5STS: 5 sit-to-stand time.

[h]WALK4m: time to walk 4 meters.

[i]GS4m: gait speed over 4 meters at a normal pace.

XSL•FO
RenderX

[j]Drugs: number of daily medications.

[k]US: ultrasound on rectus femoris.

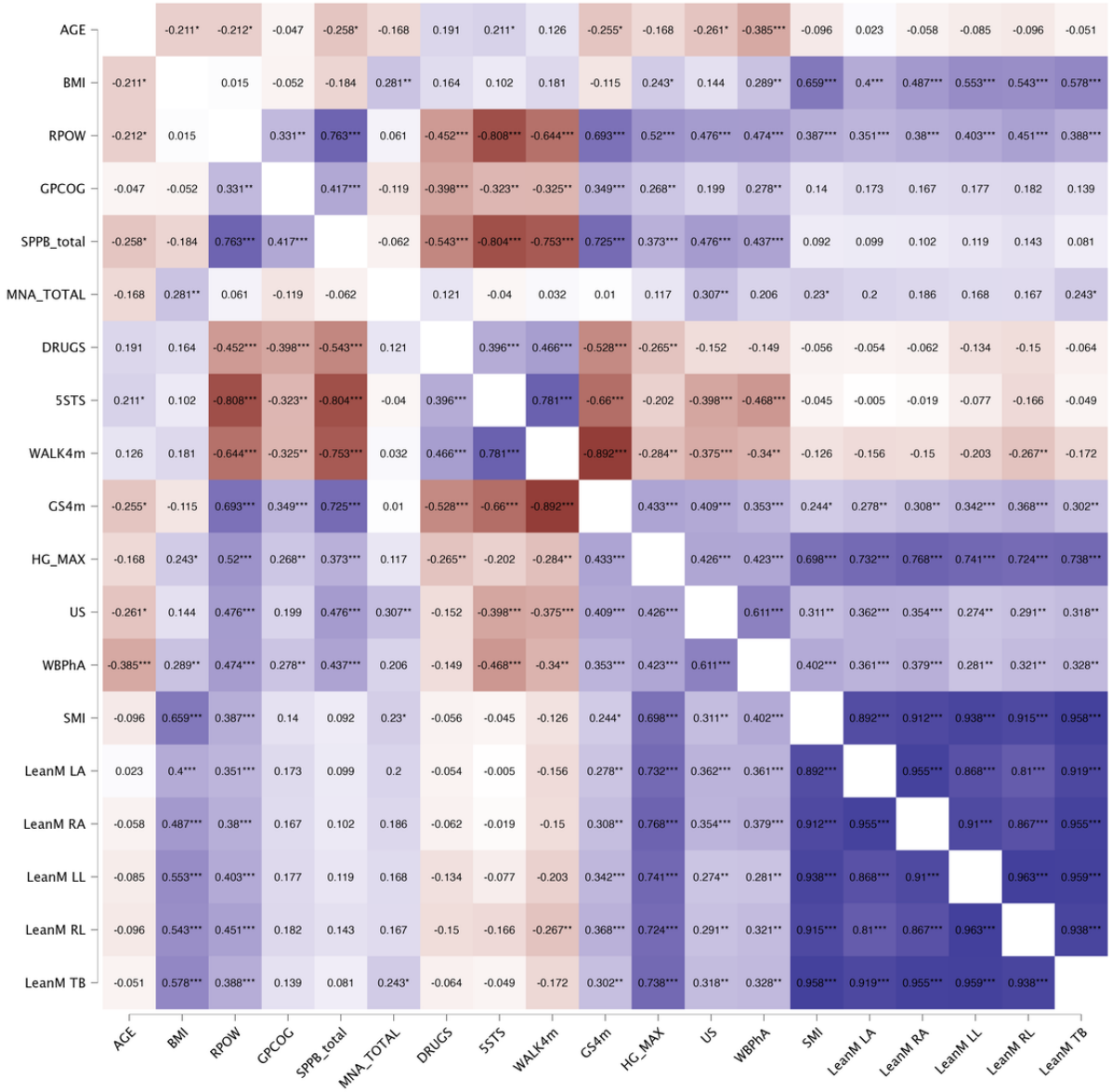
[l]WBPhA: whole-body phase angle.

[m]SMI: skeletal muscle index.

The correlation analysis is presented through a heat map in Figures 2-4. Starting with the primary variable identified as a frailty criterion (RPOW), we initially analyzed the behavior of this variable in relation to the other studied variables. Additional correlation analyses were also explored, such as examining the behavior and relationship of handgrip strength. In this regard, we conducted a first analysis on the entire sample (Figure 2), then one each for women (Figure 3) and men (Figure 4). In the entire sample (Figure 1), we found strong positive and negative correlations between the RPOW and all the variables, except MNA score, however, these relationships changed for the women-only and men-only samples. In the complete sample (Figure 2), the results identified stronger correlations between SPPB and RPOW, 5 sit-to-stand time and RPOW, drugs and RPOW, GS4m and RPOW, and handgrip and RPOW. Regarding the analysis of the correlations separately for women and men (Figures 3 and 4), the results were as expected. Specifically, as observed in the cited figures, there were relationships involving body composition variables (BIA, DXA, and US) and expressions of strength, such as lower limb power and, more markedly, handgrip strength. The observed differences in correlations between sexes are consistent with expectations, given the complex biological, hormonal, and social interactions that affect males and females differently during the aging process. It is important to consider these factors when interpreting the results and when designing interventions that address frailty and functional decline in older adults effectively and equitably.

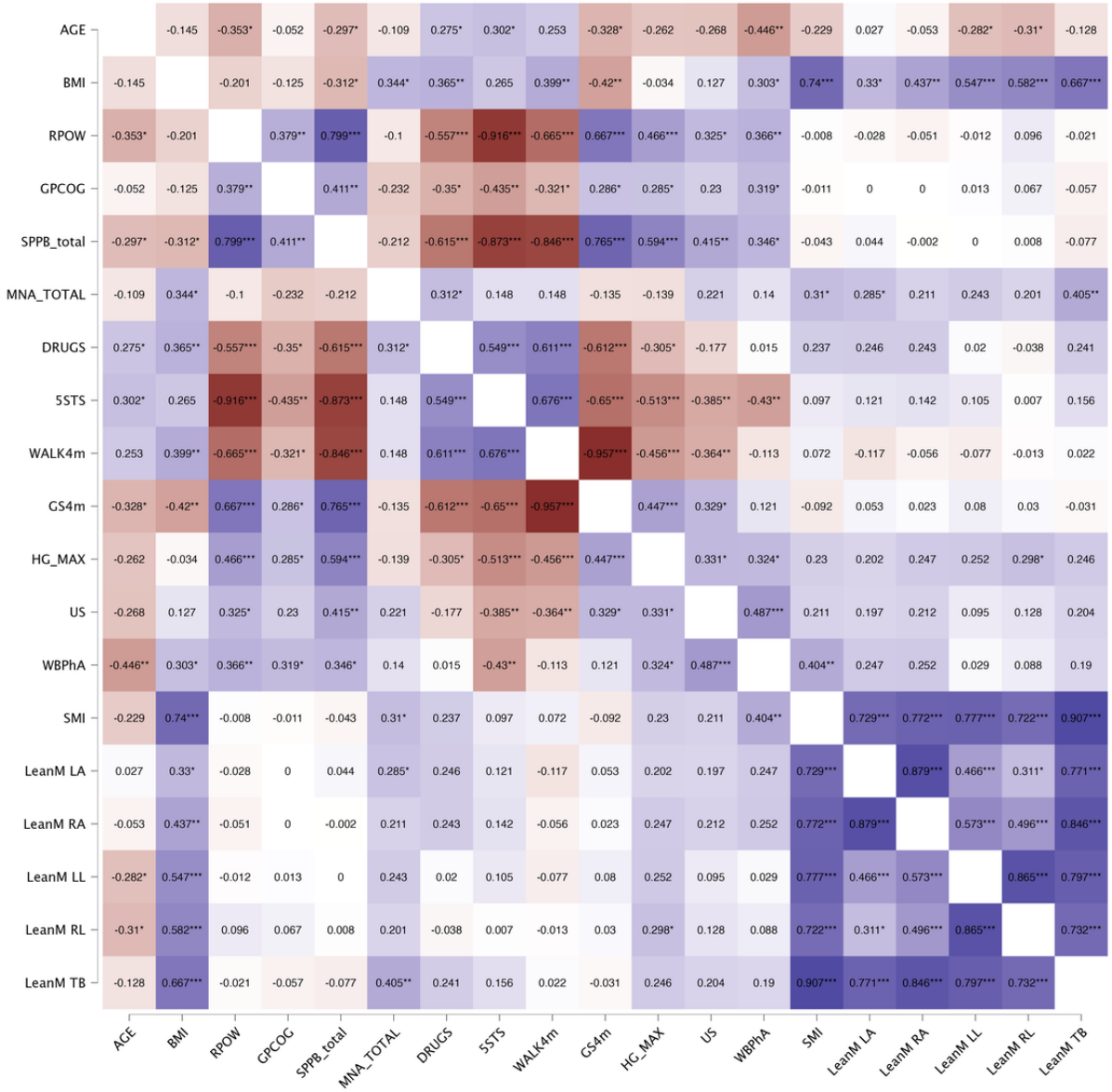
XSL•FO
RenderX

Figure 2. Simple correlation analysis results for the whole sample (n=94). Pearson correlations of all studied variable were carried out on the total sample. 5STS: 5 sit-to-stand time; DRUGS: number of daily medications; GPCOG: General Practitioner Assessment of Cognition score; GS4m: gait speed over 4 meters at a normal pace; HG: handgrip; LeanM LA: left arm lean mass; LeanM LL: left leg lean mass; LeanM RA: right arm lean mass; LeanM RL: right leg lean mass; LeanM TB: total body lean mass; MNA: Mini Nutritional Assessment score; RPOW: relative muscle power; SMI: skeletal muscle index; SPPB: Short Physical Performance Battery score; US: ultrasound on rectus femoris; WALK4m: time to walk 4 meters; WBPhA: whole-body phase angle. *P<.05, **P<.01, ***P<.001.



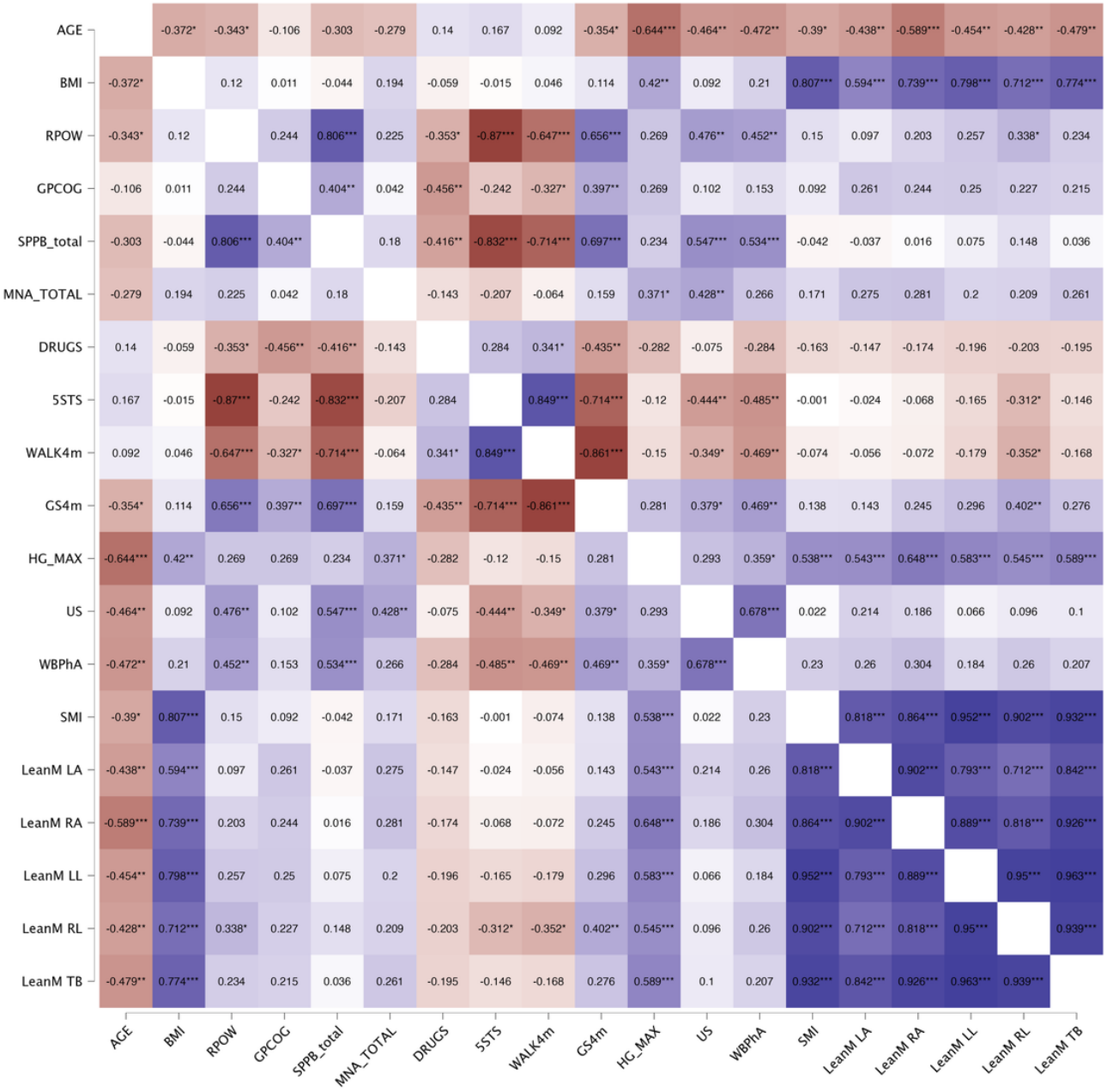
XSL•FO
RenderX

Figure 3. Simple correlation analysis results for women (n=52). Pearson correlations were conducted for all studied variables on the sample of women. 5STS: 5 sit-to-stand time; DRUGS: number of daily medications; GPCOG: General Practitioner Assessment of Cognition score; GS4m: gait speed over 4 meters at a normal pace; HG: handgrip; LeanM LA: left arm lean mass; LeanM LL: left leg lean mass; LeanM RA: right arm lean mass; LeanM RL: right leg lean mass; LeanM TB: total body lean mass; MNA: Mini Nutritional Assessment score; RPOW: relative muscle power; SMI: skeletal muscle index; SPPB: Short Physical Performance Battery score; US: ultrasound on rectus femoris; WALK4m: time to walk 4 meters; WBPhA: whole-body phase angle. *P<.05, **P<.01, ***P<.001.



XSL•FO
RenderX

Figure 4. Simple correlation analysis results for men (n=42). Pearson correlations were conducted for all studied variables on the sample of men (n=42). 5STS: 5 sit-to-stand time; DRUGS: number of daily medications; GPCOG: General Practitioner Assessment of Cognition score; GS4m: gait speed over 4 meters at a normal pace; HG: handgrip; LeanM LA: left arm lean mass; LeanM LL: left leg lean mass; LeanM RA: right arm lean mass; LeanM RL: right leg lean mass; LeanM TB: total body lean mass; MNA: Mini Nutritional Assessment score; RPOW: relative muscle power; SMI: skeletal muscle index; SPPB: Short Physical Performance Battery score; US: ultrasound on rectus femoris; WALK4m: time to walk 4 meters; WBPhA: whole-body phase angle. *P<.05, **P<.01, ***P<.001.



The logistic regression analyses presented in Table 2 demonstrate the predictive capacity of various diagnostic factors for frailty in older adults. The Nagelkerke R^2 values indicate the proportion of variance in frailty status explained by each model, both before and after adjusting for age and sex. The initial model for US shows that US measurements of the rectus femoris explained 10.1% of the variance in frailty (Nagelkerke R^2=0.101; P=.009). After adjusting for age and sex, the explained variance increases to 20.6% (adjusted Nagelkerke R^2=0.206; P=.002), indicating a stronger association when these factors are considered. WBPhA exhibited a significant association with frailty, explaining 16.4% of the variance in the unadjusted model (Nagelkerke R^2=0.164; P=.001). The adjusted model further enhanced the explained variance to 24.5% (adjusted Nagelkerke R^2=0.245; P<.001), underscoring its robustness as a predictive factor. Both left leg lean mass and right leg lean mass (LeanM RL) were significant predictors of frailty. The unadjusted models explained 12.4% (P=.003) and 16.4% (P<.001) of the variance, respectively. After adjustment, the explained variance increased to 21.5% for the left leg (P=.001) and 24.6% for the right leg (P<.001).

XSL•FO
RenderX

Table . Results of the logistic regression analyses (n=94).

	Nagelkerke R^2	*P* value	Adjusted Nagelkerke R^{2a}	*P* value
Ultrasound on rectus femoris	0.101	.009	0.206	.002
WBPhA[b]	0.164	.001	0.245	<.001
Lean mass, left leg	0.124	.003	0.215	.001
Lean mass, right leg	0.164	<.001	0.246	<.001

[a]Adjusted Nagelkerke R^2 included age as a covariable and sex as a factor.

[b]WBPhA: whole-body phase angle.

The ROC analyses indicated that all evaluated measures demonstrated a high degree of specificity in detecting frailty in older adults. It is notable that unadjusted muscle US demonstrated a high level of specificity, although sensitivity was relatively low. The WBPhA and LeanM RL also exhibited significant predictive capabilities, with AUC and *P* values indicating good discrimination. When adjusting for sex and age, there was a notable improvement in sensitivity, particularly in the case of LeanM RL, which maintained statistical significance (*P*=.045). These findings suggest that the lean mass of the right leg, adjusted for demographic factors, may be a promising indicator for predicting frailty in this population. On the other hand, it is worth noting that the adjusted WBPhA presented the highest Youden index value (0.417), which suggests that it may be the most effective measure in terms of balancing sensitivity and specificity for predicting frailty. Similarly, the adjusted LeanM RL also had a high Youden index (0.296), which establishes it as another promising measure.

Discussion

Principal Findings

This study provides significant evidence on the predictive capacity of body composition variables and their relationship with frailty syndrome in older adults. Technology tools such as BIA, WBPhA, US of the anterior rectus quadriceps muscle thickness, and DXA, were used in this study. The older adult frailty status was calculated using the PowerFrail mobile app [15] (Figure 1). This app is the first scientifically based app that allows the assessment of the muscle power and frailty of older adults in a simple way. Our results showed that these parameters are useful predictors for identifying frailty, in line with previous findings.

This study found several significant correlations between variables related to body composition, physical performance, and cognitive status in female participants, but these were not the same as the values observed in their male counterparts. This may be because women experience more functional limitations than men during the aging process [31]. Women tend to suffer a greater decline in physical function with age, which has been attributed to hormonal changes, particularly the decrease in estrogen levels during menopause [32]. Menopause is associated with changes in body composition characterized by an increase in body fat and a progressive decrease in muscle mass and strength [33]. These alterations can lead to a higher prevalence of sarcopenia and frailty in older women compared to men.

Based on these postmenopausal changes, muscle and fat composition might be more closely related to physical performance in women, which we observed in our correlation results from Figures 2-4. Specifically, the associations between body composition measures (evaluated through BIA, DXA, and US) and expressions of strength, such as lower limb power from the mobile app and handgrip strength, were more pronounced in women. This suggests that changes in muscle quality and quantity may have a greater impact on the physical function of older women.

Although previous research has indicated that muscle quality contributes to physical capacity in older adults, the effect of sex differences on this association has not been thoroughly investigated [34]. Some studies have suggested that muscle quality independently predicts physical function in older men but not in women [35]. However, our findings differ, as we observed significant correlations in women. Unlike previous studies, we included analyses of physical function status, which may explain the different results. The underlying mechanisms of these sex-specific differences still need to be investigated and clarified.

Contributing factors may include differences in muscle structure composition, hormonal influences, and neuromuscular activation patterns between men and women [36]. Additionally, the accuracy and sensitivity of measurement techniques, such as BIA and US, may vary between sexes due to differences in body fat distribution and hydration status [37]. Further research is needed to explore these factors and understand how they influence the relationship between body composition and physical function in older adults.

The connection between body composition and frailty indicators has been the focus of extensive research. However, to our knowledge, no studies have simultaneously compared DXA, BIA, and US, specifically for the identification of frailty syndrome in a primary care setting. Our study addresses this gap by evaluating the predictive capacities of these 3 modalities, offering a comprehensive analysis that can inform clinical practice regarding the most suitable and practical tools for frailty assessment in older adults. Although previous studies have compared two modalities, our inclusion of a third (US) provides a broader perspective on body composition assessment tools available for primary care settings. For instance, research has demonstrated that muscle US is an emerging tool for diagnosing sarcopenia, with studies summarizing its diagnostic accuracy [38]. Additionally, studies have evaluated the reliability and

XSL•FO
RenderX

validity of sarcopenia diagnosis using BIA compared with the gold standard, DXA, assessing the predictive accuracy of BIA for diagnosis [39,40]. However, these studies did not include US in their comparisons. Rossini-Venturini et al [41] highlighted that the anthropometric prediction equations developed in their study provide a reliable, practical, and low-cost instrument to assess the components that change the most during the aging process, corroborating our findings. This perspective emphasizes the significance of considering diverse elements of body composition in the evaluation of health among older adults. Although some studies have recognized muscle mass, assessed via DXA or BIA, as a critical element in forecasting frailty [42,43], our analysis demonstrated that lower extremity lean mass (left left lean mass and LeanM RL) did reveal a direct relationship in the models. The present results are in line with findings that indicate muscle mass reduction alone does not suffice to predict frailty without factoring in physical performance [44]. In fact, physical performance, which can be measured through functional tests such as gait speed or grip strength, is vital for evaluating frailty status among older adults [45].

There has been a growing acknowledgment in recent literature regarding the significance of assessing both the quantity and quality of muscle. Xu et al [46] demonstrated that body composition, encompassing both muscle mass and quality, is significantly associated with frailty in older adult inpatients. The current literature suggests that it is not enough to analyze the quantity of muscle mass in absolute terms; quality is equally important. Quality can be assessed through methods such as US and BIA, which provide insights into muscle integrity and performance in terms of its functional capacity. In this context, analyzing muscle quality through US could effectively complement DXA, which has traditionally been considered the gold standard in the assessment of body composition. Although DXA provides valuable information on lean mass and fat mass, it does not provide details on muscle distribution and quality, critical for understanding frailty in older adults [47].

Moreover, our findings indicated that muscle strength and gait speed are important indicators of frailty, corroborating the work of Tsukasaki et al [48], who found a strong association between muscle strength, gait speed, and cross-sectional muscle area determined by midthigh computed tomography. These findings reinforce the idea that comprehensive evaluations of muscle function should be integrated into frailty assessments. Therefore, muscle composition emerges as a potential public health assessment by enabling the clinical quantification of muscle mass and an estimation of physical function in the older adult population [49].

Our findings demonstrated that all measures exhibited high specificity but varying sensitivity in detecting frailty (Table 3). Our ROC analysis showed moderate predictive ability for WBPhA, US thickness, and lean mass of legs from DXA, with an AUC between 0.678 and 0.749. This aligns with the results from previous studies [50,51], which emphasized that models combining body composition measurements with physical performance tests can improve the predictive ability of frailty. Unadjusted US showed perfect specificity but low sensitivity (10.7%), indicating it is highly effective at correctly identifying nonfrail individuals but less capable of detecting those who are frail. This aligns with previous studies suggesting that muscle US, while precise in measuring muscle thickness, may have limitations in sensitivity due to operator dependency and variability [52].

Table . Results of receiver operating characteristic analysis (n=94)

	Sensitivity	Specificity	Youden index[a]	AUC[b]	*P* value
US[c]	0.107	1.000	0.107	0.678	.02
US (adjusted)[d]	0.321	0.939	0.260	0.732	.24
WBPhA[e]	0.240	0.952	0.192	0.704	.003
WBPhA (adjusted)[d]	0.480	0.937	0.417	0.762	.09
LeanM LL[f]	0.107	0.985	0.092	0.683	.009
LeanM LL (adjusted)[d]	0.250	0.924	0.174	0.741	.17
LeanM RL[g]	0.286	0.939	0.225	0.703	.003
LeanM RL (adjusted)[d]	0.357	0.939	0.296	0.749	.045

[a]Youden index = sensitivity + specificity – 1.

[b]AUC: area under the curve.

[c]US: ultrasound on rectus femoris.

[d]Model adjusted by sex and age.

[e]WBPhA: whole-body phase angle.

[f]LeanM LL: left leg lean mass.

[g]LeanM RL: right leg lean mass.

Adjusting the US measure for age and sex increased sensitivity to 32.1% but reduced specificity to 93.9%, and the adjusted model did not reach statistical significance (*P*=.24). This suggests that while adjustments improve sensitivity, they may

XSL•FO
RenderX

not be sufficient to make US a standalone diagnostic tool for frailty in primary care settings. The WBPhA showed an unadjusted sensitivity of 24% and specificity of 95.2%, with an AUC of 0.704 (*P*=.003). After adjusting for age and sex, sensitivity improved to 48%, specificity slightly decreased to 93.7%, and AUC increased to 0.762, although the *P* value was nonsignificant (*P*=.09). These results suggest that WBPhA, particularly when adjusted for sex and age factors, has potential as a screening tool for frailty. Previous research has shown that lower phase angle values are associated with decreased muscle function, poor nutritional status, and higher frailty risk [12,53]. The noninvasive nature and ease of use of BIA make WBPhA a practical option for primary care, although standardization of measurement protocols is necessary.

LeanM RL emerged as a significant predictor when adjusted for age and sex (*P*=.045), with sensitivity and specificity of 35.7% and 93.9%, respectively, and an AUC of 0.749. This indicates that the adjusted LeanM RL model had a good balance between sensitivity and specificity and may be valuable in predicting frailty. The importance of lower limb muscle mass in frailty assessment is well-documented. For instance, [54] reported that decreased appendicular lean mass is associated with physical disability and increased risk of adverse outcomes in older adults. However, the use of DXA in primary care is limited due to cost, accessibility, and exposure to low-dose radiation.

In comparison, the unadjusted left leg lean mass showed similar specificity (98.5%) but low sensitivity (10.7%), and the adjusted model did not achieve statistical significance. This suggests that while lean mass measurements are informative, the right leg may provide more predictive value than the left in this context, possibly due to dominance or functional differences, although further research is needed to confirm this observation. Overall, our results indicated that LeanM RL, adjusted for age and sex, may be the most effective measure among those studied for predicting frailty. This is significant because identifying reliable, accessible markers for frailty is crucial for early intervention. Given the limitations of DXA in primary care, exploring alternative methods to estimate lean mass, such as predictive equations or portable devices, could enhance feasibility. The high specificity observed across all measures suggests they are effective in ruling out frailty in individuals without frailty. However, the variable sensitivity underscores the need for multicomponent assessment tools. The Comprehensive Geriatric Assessment remains the gold standard for frailty evaluation but is resource-intensive [55]. Incorporating measures like WBPhA and simplified lean mass assessments could enhance screening efficiency in primary care.

Limitations, Clinical Implications, and Future Directions

Limitations of the study include the sample size and homogeneity in some variables, such as BMI and arm lean mass (left arm and right arm), which could have affected the accuracy of the predictive results. Heterogeneity in these measurements has been reported in previous research [56], suggesting that variability in body composition could influence frailty prediction. There is potential for the integration of advanced body composition analysis tools and mobile technology into primary care services for the early identification and management of frailty syndrome in older adults. Methods such as WBPhA, US of the rectus femoris muscle, and DXA analysis proved to be effective predictors of frailty, offering high specificity in detection. The implementation of the PowerFrail mobile app enables rapid and personalized assessments, optimizing preventive interventions. These tools not only improve the accuracy of clinical evaluations but also reduce costs associated with frailty-related complications, promoting healthier aging and alleviating the burden on public health care systems. In addition, while this study incorporated a range of assessment tools, it is essential for future research to concentrate on validating these measures in larger and more diverse cohorts to ensure their wider applicability. The cross-sectional design of this study further restricts the ability to infer causal relationships between body composition variables and frailty. Longitudinal investigations could provide valuable information on how variations in body composition affect the progression of frailty over time. Understanding these dynamics could facilitate developing more effective and personalized interventions for the older adult population. Finally, future studies should explore the integration of these mHealth technologies with advanced body composition analysis systems to optimize early detection and management of frailty.

Conclusions

The findings of this study reinforce the utility of various body composition evaluations, such as WBPhA, LeanM RL measured by DXA, and quadriceps thickness assessed by US, as effective indicators for predicting frailty in older adults, aligning with previous research. However, our results highlight the necessity of not relying exclusively on muscle mass as a predictor of frailty. It is essential to incorporate assessments of muscle function and physical performance into clinical evaluations to enhance the accuracy of identifying individuals susceptible to frailty. The combination of tools such as WBPhA, bioimpedance, and US, along with the PowerFrail app, could provide a more complete and accurate assessment of the health status of older adults, allowing effective preventive interventions to be implemented in primary care services.

Acknowledgments

This study received funding from the Call for Biomedical Research and Innovation Projects from the Biomedical Research Institute of Cádiz, Andalusian Health System Grant (PP11-004-2023); PUENTE Project Call for the Enhancement of Proposals from the State Plan of the University (PR2022-018); and Plan Andaluz de Investigación, Desarrollo e Innovación (PAIDI 2020; DOC_01412) from the Andalusian Government.

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Bandeen-Roche K, Xue QL, Ferrucci L, et al. Phenotype of frailty: characterization in the women's health and aging studies. J Gerontol A Biol Sci Med Sci 2006 Mar;61(3):262-266. [doi: 10.1093/gerona/61.3.262] [Medline: 16567375]
2. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
3. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet 2013 Mar 2;381(9868):752-762. [doi: 10.1016/S0140-6736(12)62167-9] [Medline: 23395245]
4. Rockwood K, Mitnitski A. Frailty defined by deficit accumulation and geriatric medicine defined by frailty. Clin Geriatr Med 2011 Feb;27(1):17-26. [doi: 10.1016/j.cger.2010.08.008] [Medline: 21093719]
5. Galán-Mercant A, Cuesta-Vargas AI. Detección precoz de la fragilidad, tecnología aplicada al movimiento humano para la prevención de la discapacidad [Article in Spanish]. Fisioterapia 2017 May;39(3):135-136. [doi: 10.1016/j.ft.2016.10.002]
6. Puts MTE, Toubasi S, Andrew MK, et al. Interventions to prevent or reduce the level of frailty in community-dwelling older adults: a scoping review of the literature and international policies. Age Ageing 2017 May 1;46(3):383-392. [doi: 10.1093/ageing/afw247] [Medline: 28064173]
7. Dent E, Morley JE, Cruz-Jentoft AJ, et al. International Clinical Practice Guidelines for Sarcopenia (ICFSR): screening, diagnosis and management. J Nutr Health Aging 2018;22(10):1148-1161. [doi: 10.1007/s12603-018-1139-9] [Medline: 30498820]
8. Rolland Y, Czerwinski S, Abellan Van Kan G, et al. Sarcopenia: its assessment, etiology, pathogenesis, consequences and future perspectives. J Nutr Health Aging 2008;12(7):433-450. [doi: 10.1007/BF02982704] [Medline: 18615225]
9. Cruz-Jentoft AJ, Baeyens JP, Bauer JM, et al. Sarcopenia: European consensus on definition and diagnosis. Age Ageing 2010 Jul 1;39(4):412-423. [doi: 10.1093/ageing/afq034]
10. Beaudart C, McCloskey E, Bruyère O, et al. Sarcopenia in daily practice: assessment and management. BMC Geriatr 2016 Oct 5;16(1):170. [doi: 10.1186/s12877-016-0349-4] [Medline: 27716195]
11. Norman K, Pichard C, Lochs H, Pirlich M. Prognostic impact of disease-related malnutrition. Clin Nutr 2008 Feb;27(1):5-15. [doi: 10.1016/j.clnu.2007.10.007] [Medline: 18061312]
12. Matias CN, Nunes CL, Francisco S, et al. Phase angle predicts physical function in older adults. Arch Gerontol Geriatr 2020;90:104151. [doi: 10.1016/j.archger.2020.104151] [Medline: 32563736]
13. Gupta M, Lehl SS, Lamba AS. Ultrasonography for assessment of sarcopenia: a primer. J Midlife Health 2022;13(4):269-277. [doi: 10.4103/jmh.jmh_234_22] [Medline: 37324795]
14. Buckinx F, Landi F, Cesari M, et al. Pitfalls in the measurement of muscle mass: a need for a reference standard. J Cachexia Sarcopenia Muscle 2018 Apr;9(2):269-278. [doi: 10.1002/jcsm.12268] [Medline: 29349935]
15. Alcazar J, Losa-Reyna J, Rodriguez-Lopez C, et al. The sit-to-stand muscle power test: an easy, inexpensive and portable procedure to assess muscle power in older people. Exp Gerontol 2018 Oct 2;112:38-43. [doi: 10.1016/j.exger.2018.08.006] [Medline: 30179662]
16. Losa-Reyna J, Alcazar J, Rodríguez-Gómez I, et al. Low relative mechanical power in older adults: an operational definition and algorithm for its application in the clinical setting. Exp Gerontol 2020 Dec;142:111141. [doi: 10.1016/j.exger.2020.111141] [Medline: 33127413]
17. Velazquez-Diaz D, Arco JE, Ortiz A, et al. Use of artificial intelligence in the identification and diagnosis of frailty syndrome in older adults: scoping review. J Med Internet Res 2023 Oct 20;25:e47346. [doi: 10.2196/47346] [Medline: 37862082]
18. Liu N, Yin J, Tan SSL, Ngiam KY, Teo HH. Mobile health applications for older adults: a systematic review of interface and persuasive feature design. J Am Med Inform Assoc 2021 Oct 12;28(11):2483-2501. [doi: 10.1093/jamia/ocab151] [Medline: 34472601]
19. Soto-Bagaria L, Eis S, Pérez LM, et al. Mobile applications to prescribe physical exercise in frail older adults: review of the available tools in app stores. Age Ageing 2023 Dec 1;52(12):afad227. [doi: 10.1093/ageing/afad227] [Medline: 38157286]
20. Mahoney FI, Barthel DW. Functional evaluation: the Barthel index. Md State Med J 1965 Feb;14:61-65. [Medline: 14258950]
21. Bohannon RW. Sit-to-stand test for measuring performance of lower extremity muscles. Percept Mot Skills 1995 Feb;80(1):163-166. [doi: 10.2466/pms.1995.80.1.163] [Medline: 7624188]
22. Losa-Reyna J, Baltasar-Fernandez I, Alcazar J, et al. Effect of a short multicomponent exercise intervention focused on muscle power in frail and pre frail elderly: a pilot trial. Exp Gerontol 2019 Jan;115:114-121. [doi: 10.1016/j.exger.2018.11.022] [Medline: 30528641]
23. Mateos-Angulo A, Galán-Mercant A, Cuesta-Vargas AI. Muscle thickness contribution to sit-to-stand ability in institutionalized older adults. Aging Clin Exp Res 2020 Aug;32(8):1477-1483. [doi: 10.1007/s40520-019-01328-x] [Medline: 31463929]
24. Villarejo A, Puertas-Martín V. Utilidad de los test breves en el cribado de demencia [Article in Spanish]. Neurología 2011 Sep;26(7):425-433. [doi: 10.1016/j.nrl.2010.12.002]

XSL•FO
RenderX

25. Muñoz Díaz B, Molina-Recio G, Romero-Saldaña M, et al. Validation (in Spanish) of the Mini Nutritional Assessment survey to assess the nutritional status of patients over 65 years of age. Fam Pract 2019 Mar 20;36(2):172-178. [doi: 10.1093/fampra/cmy051] [Medline: 29873713]
26. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994 Mar;49(2):M85-M94. [doi: 10.1093/geronj/49.2.m85] [Medline: 8126356]
27. Cruz-Jentoft AJ, Bahat G, Bauer J, et al. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing 2019 Jan 1;48(1):16-31. [doi: 10.1093/ageing/afy169] [Medline: 30312372]
28. Roberts HC, Denison HJ, Martin HJ, et al. A review of the measurement of grip strength in clinical and epidemiological studies: towards a standardised approach. Age Ageing 2011 Jul;40(4):423-429. [doi: 10.1093/ageing/afr051] [Medline: 21624928]
29. Martínez-Camblor P, Pardo-Fernández JC. The Youden index in the generalized receiver operating characteristic curve context. Int J Biostat 2019 Apr 3;15(1). [doi: 10.1515/ijb-2018-0060] [Medline: 30943172]
30. Swets JA. Measuring the accuracy of diagnostic systems. Science 1988 Jun 3;240(4857):1285-1293. [doi: 10.1126/science.3287615] [Medline: 3287615]
31. Iannuzzi-Sucich M, Prestwood KM, Kenny AM. Prevalence of sarcopenia and predictors of skeletal muscle mass in healthy, older men and women. J Gerontol A Biol Sci Med Sci 2002 Dec;57(12):M772-M777. [doi: 10.1093/gerona/57.12.m772] [Medline: 12456735]
32. Maltais ML, Desroches J, Dionne IJ. Changes in muscle mass and strength after menopause. J Musculoskelet Neuronal Interact 2009;9(4):186-197. [Medline: 19949277]
33. Sipilä S, Narici M, Kjaer M, et al. Sex hormones and skeletal muscle weakness. Biogerontology 2013 Jun;14(3):231-245. [doi: 10.1007/s10522-013-9425-8] [Medline: 23636830]
34. Barbat-Artigas S, Rolland Y, Vellas B, Aubertin-Leheudre M. Muscle quantity is not synonymous with muscle quality. J Am Med Dir Assoc 2013 Nov;14(11):852. [doi: 10.1016/j.jamda.2013.06.003] [Medline: 23896368]
35. Goodpaster BH, Park SW, Harris TB, et al. The loss of skeletal muscle strength, mass, and quality in older adults: the health, aging and body composition study. J Gerontol A Biol Sci Med Sci 2006 Oct;61(10):1059-1064. [doi: 10.1093/gerona/61.10.1059] [Medline: 17077199]
36. Clark BC, Manini TM. Sarcopenia =/= dynapenia. J Gerontol A Biol Sci Med Sci 2008 Aug;63(8):829-834. [doi: 10.1093/gerona/63.8.829] [Medline: 18772470]
37. Kyle UG, Genton L, Slosman DO, Pichard C. Fat-free and fat mass percentiles in 5225 healthy subjects aged 15 to 98 years. Nutrition 2001;17(7-8):534-541. [doi: 10.1016/s0899-9007(01)00555-x] [Medline: 11448570]
38. Fu H, Wang L, Zhang W, Lu J, Yang M. Diagnostic test accuracy of ultrasound for sarcopenia diagnosis: a systematic review and meta-analysis. J Cachexia Sarcopenia Muscle 2023 Feb;14(1):57-70. [doi: 10.1002/jcsm.13149] [Medline: 36513380]
39. Lee JH, Kim HJ, Han S, Park SJ, Sim M, Lee KH. Reliability and agreement assessment of sarcopenia diagnosis through comparison of bioelectrical impedance analysis and dual-energy X-ray absorptiometry. Diagnostics (Basel) 2024 Apr 25;14(9):899. [doi: 10.3390/diagnostics14090899] [Medline: 38732314]
40. Lee SY, Ahn S, Kim YJ, et al. Comparison between dual-energy X-ray absorptiometry and bioelectrical impedance analyses for accuracy in measuring whole body muscle mass and appendicular skeletal muscle mass. Nutrients 2018 Jun 7;10(6):738. [doi: 10.3390/nu10060738] [Medline: 29880741]
41. Rossini-Venturini AC, Veras L, Abdalla PP, et al. Multicompartment body composition analysis in older adults: a cross-sectional study. BMC Geriatr 2023 Feb 9;23(1):87. [doi: 10.1186/s12877-023-03752-1] [Medline: 36759773]
42. Kołodziej M, Sebastjan A, Ignasiak Z. Appendicular skeletal muscle mass and quality estimated by bioelectrical impedance analysis in the assessment of frailty syndrome risk in older individuals. Aging Clin Exp Res 2022 Sep;34(9):2081-2088. [doi: 10.1007/s40520-021-01879-y] [Medline: 34118025]
43. Kołodziej M, Kozieł S, Ignasiak Z. The use of the bioelectrical impedance phase angle to assess the risk of sarcopenia in people aged 50 and above in Poland. Int J Environ Res Public Health 2022 Apr 13;19(8):4687. [doi: 10.3390/ijerph19084687] [Medline: 35457554]
44. Taylor JA, Greenhaff PL, Bartlett DB, Jackson TA, Duggal NA, Lord JM. Multisystem physiological perspective of human frailty and its modulation by physical activity. Physiol Rev 2023 Apr 1;103(2):1137-1191. [doi: 10.1152/physrev.00037.2021] [Medline: 36239451]
45. Jiménez-García JD, Ortega-Gómez S, Martínez-Amat A, Álvarez-Salvago F. Associations of balance, strength, and gait speed with cognitive function in older individuals over 60 years: a cross-sectional study. Appl Sci (Basel) 2024;14(4):1500. [doi: 10.3390/app14041500]
46. Xu L, Zhang J, Shen S, et al. Association between body composition and frailty in elder inpatients. Clin Interv Aging 2020;15:313-320. [doi: 10.2147/CIA.S243211] [Medline: 32184580]
47. Oba H, Matsui Y, Arai H, et al. Evaluation of muscle quality and quantity for the assessment of sarcopenia using mid-thigh computed tomography: a cohort study. BMC Geriatr 2021 Apr 13;21(1):239. [doi: 10.1186/s12877-021-02187-w] [Medline: 33849469]

XSL•FO
RenderX

48. Tsukasaki K, Matsui Y, Arai H, et al. Association of muscle strength and gait speed with cross-sectional muscle area determined by mid-thigh computed tomography - a comparison with skeletal muscle mass measured by dual-energy X-ray absorptiometry. J Frailty Aging 2020;9(2):82-89. [doi: 10.14283/jfa.2020.16] [Medline: 32259181]
49. Alarcón-Rivera M, Cornejo-Mella C, Cáceres-Aravena C, Concha-Cisternas Y, Fernández-Valero P, Guzmán-Muñoz E. Relationship between appendicular muscular mass index and physical function in older people. AIMS Public Health 2024;11(1):130-140. [doi: 10.3934/publichealth.2024006] [Medline: 38617413]
50. Masanés F, Rojano I Luque X, Salvà A, et al. Cut-off points for muscle mass - not grip strength or gait speed - determine variations in sarcopenia prevalence. J Nutr Health Aging 2017;21(7):825-829. [doi: 10.1007/s12603-016-0844-5] [Medline: 28717813]
51. Bahat G, Tufan A, Tufan F, et al. Cut-off points to identify sarcopenia according to European Working Group on Sarcopenia in Older People (EWGSOP) definition. Clin Nutr 2016 Dec;35(6):1557-1563. [doi: 10.1016/j.clnu.2016.02.002] [Medline: 26922142]
52. Pillen S, Tak RO, Zwarts MJ, et al. Skeletal muscle ultrasound: correlation between fibrous tissue and echo intensity. Ultrasound Med Biol 2009 Mar;35(3):443-446. [doi: 10.1016/j.ultrasmedbio.2008.09.016] [Medline: 19081667]
53. Norman K, Stobäus N, Pirlich M, Bosy-Westphal A. Bioelectrical phase angle and impedance vector analysis--clinical relevance and applicability of impedance parameters. Clin Nutr 2012 Dec;31(6):854-861. [doi: 10.1016/j.clnu.2012.05.008] [Medline: 22698802]
54. Studenski SA, Peters KW, Alley DE, et al. The FNIH sarcopenia project: rationale, study description, conference recommendations, and final estimates. J Gerontol A Biol Sci Med Sci 2014 May;69(5):547-558. [doi: 10.1093/gerona/glu010] [Medline: 24737557]
55. Ellis G, Gardner M, Tsiachristas A, et al. Comprehensive geriatric assessment for older adults admitted to hospital. Cochrane Database Syst Rev 2017 Sep 12;9(9). [doi: 10.1002/14651858.CD006211.pub3] [Medline: 28898390]
56. Cesari M, Gambassi G, van Kan GA, Vellas B. The frailty phenotype and the frailty index: different instruments for different purposes. Age Ageing 2014 Jan;43(1):10-12. [doi: 10.1093/ageing/aft160] [Medline: 24132852]

Abbreviations

AUC: area under the curve
BIA: bioelectrical impedance analysis
DXA: dual-energy X-ray absorptiometry
GS4m: gait speed over 4 meters at a normal pace
LeanM RL: right leg lean mass
mHealth: mobile health
MNA: Mini Nutritional Assessment
ROC: receiver operating characteristic
RPOW: relative muscle power
SPPB: Short Physical Performance Battery
US: ultrasound
WBPhA: whole-body phase angle

Edited by M Antoniou; submitted 29.10.24; peer-reviewed by C Garcia-Munoz, EO GaitÁn, MN Shalaby; revised version received 08.01.25; accepted 10.02.25; published 15.05.25.

Please cite as:
Ortiz-Navarro B, Losa-Reyna J, Mihaiescu-Ion V, Garcia-Romero J, Carrillo de Albornoz-Gil M, Galán-Mercant A
Identification of Target Body Composition Parameters by Dual-Energy X-Ray Absorptiometry, Bioelectrical Impedance, and Ultrasonography to Detect Older Adults With Frailty and Prefrailty Status Using a Mobile App in Primary Care Services: Descriptive Cross-Sectional Study
JMIR Aging 2025;8:e67982
URL: https://aging.jmir.org/2025/1/e67982
doi:10.2196/67982

© Beatriz Ortiz-Navarro, José Losa-Reyna, Veronica Mihaiescu-Ion, Jerónimo Garcia-Romero, Margarita Carrillo de Albornoz-Gil, Alejandro Galán-Mercant. Originally published in JMIR Aging (https://aging.jmir.org), 15.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging,

XSL•FO
RenderX

is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Association Between the Frailty Index Based on Laboratory Tests and All-Cause Mortality in Hospitalized Older Adults: Retrospective Cohort Study

Eyal Pasternak[1,2], MD; Tamar Freud[1,3], PhD; Yan Press[1,2,3,4,5,6], MD

[1]Division of Community Health, Faculty of Health Sciences, Ben-Gurion University of the Negev, Beer- Sheva, Israel

[2]Soroka Medical Center, Beer-Sheva, Israel

[3]Siaal Research Center for Family Medicine and Primary Care, Faculty of Health Sciences, Ben-Gurion University of the Negev, Beer- Sheva, Israel

[4]Department of Geriatrics, Soroka Medical Center, Ben-Gurion University of the Negev, Reger Boulevard, POB 151, Beer-Sheva, Israel

[5]Unit for Community Geriatrics, Division of Health in the Community, Ben-Gurion University of the Negev, Beer-Sheva, Israel

[6]Center for Multidisciplinary Research in Aging, Ben-Gurion University of the Negev, Beer-Sheva, Israel

Corresponding Author:
Yan Press, MD
Division of Community Health, Faculty of Health Sciences, Ben-Gurion University of the Negev, Beer- Sheva, Israel

Abstract

Background: Frailty is a common issue among hospitalized older adult patients and is associated with numerous adverse health outcomes. Assessing frailty facilitates better decision-making for treatment plans, patient placement, and discharge planning. Approximately a decade ago, the frailty index based on laboratory tests (FI-Lab) metric was introduced. Although this index has been shown in numerous studies to predict adverse medical outcomes, including mortality, it has not been extensively evaluated among patients hospitalized in internal medicine departments for diverse indications.

Objective: The aim of the study was to investigate the relationship between FI-Lab at admission and all-cause mortality during hospitalization and after discharge in patients aged 65 years and older admitted for diverse clinical indications to internal medicine departments.

Methods: This retrospective cohort study included patients aged 65 years and older hospitalized in the internal medicine departments of a large tertiary hospital. Data included demographic variables, comorbidity, and all-cause mortality. The FI-Lab was calculated based on 16 available blood tests, as well as blood pressure and heart rate measurements. We used Cox proportional hazards regression models to evaluate associations with mortality. Model performance was assessed using the C-index and time-dependent receiver operating characteristic (ROC) curves. Hospitalization data were collected from December 25, 2016, to January 7, 2023.

Results: During the study period, 31,443 patients were hospitalized in internal medicine departments, and FI-Lab was calculable for 31,398 of them. The mean age of the patients was 77.6 (SD 8.2) years, and 52.1% (16,346/31,443) were women. The mean FI-Lab score was 0.38 (SD 0.15). Based on FI-Lab scores, patients were categorized into 4 groups: robust, mildly prefrail, moderately prefrail, and frail. After adjusting for age, sex, and comorbidities, frail and prefrail patients exhibited higher mortality rates than robust patients. For each 0.01 increase in the FI-Lab score (as a continuous variable), adjusted analyses revealed a 5.5% increase in in-hospital mortality, a 2.9% increase in mortality within the first year after hospitalization, and a 1.9% increase in mortality beyond the first year.

Conclusions: The FI-Lab is a readily available and informative metric of frailty in older hospitalized patients. Calculating this index can assist physicians with identifying patients at high risk of mortality and provide meaningful information to support clinical decision-making.

(JMIR Aging 2025;8:e70204) doi:10.2196/70204

KEYWORDS

FI-Lab; all-cause mortality; hospitalized; older adults; risk factor

Introduction

Frailty is a condition of increased vulnerability to impaired recovery of homeostasis following stress [1]. Frailty is highly prevalent among adults. In a nationally representative sample of Medicare enrollees aged 65 years and older, the prevalence of prefrail individuals was 45.5%, while the prevalence of frail individuals was 15.3% [2]. The rate of frailty among hospitalized

XSL•FO
RenderX

patients is even higher. In a study of more than 22,000 individuals aged 65 years and older admitted to internal medicine departments, 43% were classified as prefrail, and 20% were classified as frail [3]. A higher prevalence was found in a meta-analysis by Ligthart-Melis et al [4], with 47% classified as prefrail and 36% classified as frail.

Among older hospitalized patients, frailty is associated with longer hospital stays [3,5-7], a higher likelihood of admission to an intensive care unit (ICU) and extended time in it [3], falls during hospitalization [6], delirium [6], pressure ulcers [6], an increased likelihood of institutionalization after discharge [3,6], a greater risk of readmission [3,8], higher hospitalization costs [9], and even increased mortality [3,5,6,10,11].

In 2001, two primary models of frailty were introduced: the cumulative deficit model [12] and the phenotype model [13]. Based on these two models, more than 65 tests have been developed to assess frailty [14]. Identifying frailty among hospitalized patients is essential as it enables, on one hand, risk stratification [7,15] and, on the other hand, the development of an intervention plan tailored for frail patients. Such interventions have been shown to shorten hospital stay [16], reduce the need for medical services after discharge, lower the likelihood of readmission, increase patient satisfaction, and even decrease mortality [17]. Despite this, frailty assessment among hospitalized patients is limited because it requires completion of a long checklist, a time-consuming process that clinicians in busy departments often cannot handle [18]. Some frailty assessment tools rely on self-reported data (eg, weight loss in the absence of objective data, information on physical activity, or feelings of exhaustion) or require patient cooperation during assessment (eg, walking speed and handgrip strength measurements) [19]. It is quite likely that such assessments would be difficult to conduct during acute hospitalization. To address these challenges, a frailty index based on laboratory tests (FI-Lab) was introduced for the first time in 2014 [20]. Since then, numerous studies have been conducted using various versions of FI-Lab. These have demonstrated its association with adverse health outcomes, including longer hospital stays [21-25]; an increased likelihood of ICU admission or prolonged ICU stays [22,23]; a greater risk for readmission [21,22,26,27]; institutionalization after discharge [24,27]; greater postdischarge care needs [21]; and increased in-hospital mortality [24,25,27-32], mortality within the first year after discharge [18,27,28,33-37], and even beyond 1 year after discharge [19,21,32,36,38-41].

As demonstrated in our literature review (see Multimedia Appendix 1), the majority of studies have focused on populations with specific medical conditions (eg, heart failure, pneumonia, sepsis). Only a limited number of studies (eg, [38,40]) have examined populations admitted to internal medicine departments for diverse indications. Moreover, to our knowledge, no studies to date have explored the relationship between FI-Lab and all-cause mortality among hospitalized patients in Israel.

The aim of this study was to evaluate the association between FI-Lab at the time of admission and mortality, both during hospitalization and following discharge, among patients aged 65 years and older admitted for diverse indications to internal medicine departments in a large tertiary care hospital in Israel.

Methods

Study Setting and Population

This was a retrospective cohort study using data from electronic medical records (EMRs) for patients aged 65 years and older who were hospitalized for any reason in the Internal Medicine Division at Soroka University Medical Center (SUMC) in southern Israel. SUMC is the third-largest tertiary medical center in Israel and the largest in southern Israel, with a capacity of 1200 beds. The Division of Internal Medicine consists of 7 internal medicine wards.

The study population included men and women 65 years of age and older who were hospitalized in one of the study wards for any reason during the 6-year period from December 25, 2016, to January 1, 2023. The follow-up period for mortality ended on November 16, 2024. It should be noted that, if a patient had multiple hospitalizations during the follow-up period, the first hospitalization was considered the index event.

Data Collection

The following data were collected from the EMR: sociodemographic, clinical, laboratory, and mortality.

Sociodemographic data included age, sex, and ethnicity.

Clinical data included the primary reason for hospital admission (according to the emergency department diagnosis) and comorbid diseases. The Charlson Comorbidity Index (CCI) [42] was automatically calculated. The first measurement of systolic and diastolic blood pressures and heart rate at triage in the emergency department was recorded.

Laboratory data included venous blood tests performed during evaluations in the emergency department or within the first 48 hours of hospitalization. We focused on 16 blood tests, at least 14 of which are routinely performed in the emergency department: hemoglobin, hematocrit, mean corpuscular volume, mean corpuscular hemoglobin, white blood cells, lymphocyte number, platelets, mean platelet volume, glucose, creatinine, urea, sodium, potassium, chloride, protein, and albumin. The latter two were usually performed during hospitalization.

All-cause mortality was defined as in-hospital mortality, mortality within 1 year after discharge, and mortality during the entire study period. Data on all-cause mortality were obtained from the SUMC computerized system, which is regularly updated with information from the Ministry of Interior's database.

F1-Lab Construction

A total of 19 items were used to construct the FI-Lab, including 16 blood tests from venous blood samples and 3 vital signs recorded during triage in the emergency department: systolic blood pressure, diastolic blood pressure, and heart rate. Each item was dichotomized based on the normal reference intervals defined by the SUMC laboratory norms. Values outside the reference range were assigned a score of 1, while values within the reference range were assigned a score of 0. The FI-Lab score

XSL•FO
RenderX

was calculated by summing the number of deficits present and dividing the total by the number of items included. The FI-Lab score ranged from 0 to 1, with higher scores indicating greater frailty. Only patients who had more than 85% of the necessary items (n>16) were included. The FI-Lab scores were categorized by quartiles and as a continuous variable (0.01-point increase).

Power Analysis

To estimate statistical power, we assumed exponentially distributed survival times and relied on outcome rates from a previously published cohort study with a military veteran population [27] in which 27.5% of the high-risk group and 8.2% of the low-risk group died within 1 year. Based on these mortality rates and an expected research population of 30,000 patients, with approximately 12% classified as high risk, we estimated the hazard ratio (HR) to be 3.76.

Power analysis for time-to-event outcomes was conducted using the method for proportional hazards regression by Schoenfeld [43].

This calculation was implemented in R, using standard functions to compute the normal quantiles and logarithmic transformation of the HR. The resulting estimate showed that only 116 events were required to achieve 99% power. Given the expected >2000 events in our study, we concluded that the study was well-powered to detect a significant difference in survival outcomes.

Statistical Analysis

Statistical analyses were performed using the SPSS version 29 (IBM Corp) and R software (version 4.4.3). A *P* value <.05 was considered statistically significant. All tests were 2-sided. Categorical variables are described as frequencies and percentiles. Continuous variables, such as age, are described as mean (SD). The comparison of specific characteristics between frailty status groups, as assessed using FI-Lab scores, was carried out using the χ^2 test for categorical variables and Kruskal-Wallis test for continuous variables, such as length of hospital stay, which did not meet the assumption of normal distribution.

Cox proportional hazards assumptions were evaluated using the Schoenfeld residuals test (cox.zph function in R) and by visual inspection of log(-log[survival]) plots. No significant violations of the proportional hazards assumption were detected for the variables included in the models (global test *P*>.05).

Cox regression models were built for all-cause mortality according to frailty status, as assessed using FI-Lab scores, at 4 different points in time: in-hospital mortality, mortality during the first year after discharge, mortality beyond the first year after discharge, and mortality at any time. These models are presented as adjusted survival curves for all-cause mortality at different levels of frailty status, as assessed using FI-Lab scores, and as adjusted HRs for FI-Lab scores, age, gender, CCI scores, and the 4 main indications for hospitalization (cardiovascular, musculoskeletal, respiratory, and fever).

To evaluate the predictive performance of FI-Lab, we used the Harrell C-index (concordance index) and time-dependent receiver operating characteristic (ROC) curves based on Cox proportional hazards models. Adjusted survival curves were generated using the *adjustedCurves* package (version 0.6.2) in R (version 4.4.3) [44,45]. The analysis was based on Cox proportional hazards models fitted with the coxph() function from the *survival* package [46], and adjusted estimates were visualized using *ggplot2* [47] and the *survminer* package (version 0.4.9) [48].

Each plot illustrates the ROC curves for the multivariable model with and without the FI-Lab variable. The comparison between models with and without the FI-Lab variable was performed using time-dependent areas under the curve (AUCs), and statistical differences were tested using a bootstrap method with 2000 resamples. The AUC values are provided for both models, along with the *P* values from the statistical comparison of AUCs.

Ethical Considerations

The study was approved by the Helsinki Committee of the Soroka Medical Center (SOR-0297 - 23). The study was based on retrospective analysis of anonymized EMR data. Informed consent was not required according to national regulations and institutional review board policy. All data were handled in compliance with institutional data privacy and confidentiality standards.

Results

Study Population Characteristics

During the study period, 31,443 patients aged 65 years and older were hospitalized in the Internal Medicine Division at SUMC. Of these, 31,398 had sufficient data to calculate FI-Lab.

The key characteristics of the 31,443 patients for whom FI-Lab was calculated are presented in Table 1. The average age was 77.6 (SD 8.2) years, 16,346 (52.1%) were women, and 27,824 (88.7%) were Jewish. The 3 primary reasons for hospitalization, as recorded in the emergency department, were cardiovascular problems (6491/31,443, 20.7%), musculoskeletal issues (5460/31,443, 17.4%), and respiratory problems (3141/31,443, 10%). The average CCI score was 3.7 (SD 2.8).

XSL•FO
RenderX

Table . Comparison of the study population (N=31,443): survivors versus nonsurvivors.

	Survivors (n=16,260)	Nonsurvivors (n=15,138)	*P* value
Age (years), mean (SD)	74.8 (7.1)	80.7 (8.4)	<.001
Age (years), median (range)	74 (65-111)	81 (65-111)	<.001
Gender, n (%)			.003
Male	7662 (47.1)	7390 (48.8)	
Female	8598 (52.9)	7746 (51.2)	
Missing	0 (0)	2 (0.01)	
Ethnicity, n (%)			<.001
Jewish	14,261 (87.9)	13,563 (89.6)	
Arab	1963 (12.1)	1574 (10.4)	
Missing	36 (0.2)	1 (0.01)	
Primary reason for hospital admission, n (%)			<.001
Cardiovascular	3647 (22.4)	2844 (18.8)	
Musculoskeletal	3771 (23.2)	1689 (11.2)	
Respiratory	1341 (8.2)	1800 (11.9)	
Fever	1358 (8.4)	1777 (11.7)	
Infection	1057 (6.5)	1212 (8)	
Gastrointestinal	859 (5.3)	887 (5.9)	
Other	4227 (26)	4929 (32.6)	
Total CCI[a] score, mean (SD)	2.9 (2.3)	4.5 (2.9)	<.001
Total CCI score, median (range)	2 (0 - 18)	4 (0 - 22)	<.001
FI-Lab score, mean (SD)	0.33 (0.14)	0.43 (0.15)	<.001
FI-Lab score, median (range)	0.31 (0 - 0.89)	0.42 (0 - 0.94)	<.001
Frailty status assessed using FI-Lab[b] scores, n (%)			<.001
Robust (<0.2632)	6083 (37.4)	2545 (16.8)	
Mildly prefrail (0.2633 - 0.0.3684)	3772 (23.2)	2563 (16.9)	
Moderately prefrail (0.3685 - 0.4737)	4240 (26.1)	5090 (33.6)	
Frail (>0.4738)	2165 (13.3)	4940 (32.6)	
Hospitalization duration (days), mean (SD)	2.9 (6.4)	4.3 (7.6)	<.001
Hospitalization duration (days), median (range)	1 (0 - 246)	2 (0 - 139)	<.001
Timing of mortality, n (%)			—[c]
Died while hospitalized	—	1658 (11)	
Died in the first year	—	5174 (34.2)	
Died after 1 year	—	8306 (54.9)	

[a]CCI: Charlson Comorbidity Index.

XSL•FO
RenderX

[b]FI-Lab: frailty index based on laboratory tests.

[c]Not applicable.

F1-Lab Scores

The mean FI-Lab score was 0.38 (SD 0.15), with a median of 0.37 (range 0.0-0.94).

Since the FI-Lab scores were not normally distributed, each quartile contained an unequal number of patients. The first quartile consisted of 8628 patients (8628/31,443, 27.5%) who had an FI-Lab score <0.2632 and were classified as "robust." The second quartile consisted of 6335 patients (6335/31,443, 20.2%) who had an FI-Lab score between 0.2633 and 0.3684 and were classified as "mildly prefrail." The third quartile consisted of 9930 patients (9930/31,443, 29.7%) who had an FI-Lab score between 0.3685 and 0.4737 and were classified as "moderately prefrail." The fourth quartile consisted of 7105 patients (7105/31,443, 22.6%) who had an FI-Lab score >0.4737 and were classified as "frail."

The minimum follow-up time was 1 day, and the maximum was 7.8 years, with an average of 3.4 (SD 2.4) years and a median of 3.2 years. During the follow-up period, 15,138 patients passed away. Of the entire sample of 31,398 patients, 1658 (5.3%) died during hospitalization, 5174 (16.5%) died within the first year after discharge, and 8306 (26.5%) died after the first year and up to the end of the follow-up period.

There were no statistically significant differences between the 45 patients excluded from the sample and the 31,443 patients included in the sample, in terms of age (P=.83), gender (P=.35), ethnicity (P=.80), CCI score (P=.46), or length of hospital stay (P=.25). In-hospital mortality rates were higher among the 45 patients excluded from the sample (15.6% vs 5.3%, P=.02). Differences in mortality rates during the first year after discharge (13.2% vs 17.4%, P=.66) and beyond the first year after discharge (39.4% vs 33.8%, P=.61) were not statistically significant.

The distribution of patient characteristics according to frailty status groups assessed using FI-Lab scores is presented in Table 2. As the FI-Lab score increased, patients in the group tended to be older (P<.001) and have a higher burden of comorbidities, as measured using the CCI (P<.001). Male representation was lower in both extreme groups (robust and frail) compared with the other 2 groups (P<.001). No differences in FI-Lab scores were found between Jewish and Arab patients.

The length of hospital stay was higher in the robust group than in the mildly prefrail group, but as the severity of frailty increased, a longer hospital stay was observed (P<.001).

XSL•FO
RenderX

Table . Comparison of patient characteristics by frailty status group, assessed using the frailty index based on laboratory tests (FI-Lab) score.

Characteristic	Frailty status group				*P* value
	Robust (n=8628)	Mildly prefrail (n=6335)	Moderately prefrail (n=9330)	Frail (n=7105)	
Gender, n (%)					<.001
Male	3975 (46.1)	3113 (49.1)	4640 (49.7)	3324 (46.8)	
Female	4653 (53.9)	3222 (50.9)	4690 (50.3)	3781 (53.2)	
Age (years), mean (SD)	75.4 (7.6)	77.2 (8.0)	78.4 (8.3)	79.8 (8.5)	<.001
Age (years), median (range)	74 (65-107)	77 (65-108)	78 (65-111)	80 (65-111)	<.001
Ethnicity, n (%)					.65
Jewish	7615 (88.4)	5625 (88.9)	8281 (88.8)	6303 (88.8)	
Arab	1004 (11.6)	701 (11.1)	1041 (11.2)	791 (11.2)	
Missing	9 (0.1)	9 (0.1)	8 (0.1)	11 (0.2)	
Total CCI[a] score, mean (SD)	2.9 (2.4)	3.4 (2.6)	3.9 (2.8)	4.5 (3.1)	<.001
Total CCI score, median (range)	2 (0 - 18)	3 (0 - 19)	3 (0 - 22)	4 (0 - 20)	<.001
Mortality, n (%)					<.001
Alive	6083 (70.5)	3772 (59.5)	4240 (45.4)	2165 (30.5)	
Died while hospitalized	105 (1.2)	128 (2)	515 (5.5)	910 (12.8)	
Died in the first year	710 (8.2)	746 (11.8)	1789 (19.2)	1929 (27.1)	
Died after a year	1730 (20.1)	1689 (26.7)	2786 (29.9)	2101 (29.6)	
Hospitalization duration (days), mean (SD)	2.76 (6.1)	2.45 (4.5)	4.0 (7.4)	5.1 (8.8)	<.001

[a]CCI: Charlson Comorbidity Index.

Mortality

In-hospital mortality rates within the first year after discharge and beyond that increased with higher FI-Lab scores. For example, in the robust group, in-hospital mortality was 1.2% (105/8628), compared with 2% (128/6335) in the mildly prefrail group, 5.5% (515/9330) in the moderately frail group, and 12.8% (910/7105) in the frail group. Overall, only 2165 patients (2165/7105, 30.5%) in the frail group were alive at the end of the follow-up period, whereas 6083 (70.5%) of the 8628 patients in the robust group survived until the end of the follow-up period (*P*<.001).

Table 3 presents the results of the Cox proportional hazard models for FI-Lab as a predictor of all-cause mortality. In the unadjusted model (Model 1); after adjustment for age and gender (Model 2); after adjustment for age, gender, and CCI (Model 3); and after adjustment for age, gender, CCI, and primary indication for hospitalization (Model 4), FI-Lab, categorized dichotomously into frailty groups (model type a) or used as a continuous variable (model type b), was consistently associated with an increased risk for mortality.

XSL•FO
RenderX

Table . Cox proportional hazard models of frailty index based on laboratory tests (FI-Lab) for prediction of all-cause mortality.

Model and variables	In-hospital mortality		Mortality during the first year after discharge		Mortality beyond the first year after discharge	
	Hazard ratio (95% CI)	*P* value	Hazard ratio (95% CI)	*P* value	Hazard ratio (95% CI)	*P* value
1a						
FI-Lab score						
Robust	1	—[a]	1	—	1	—
Mildly prefrail	1.67 (1.29-2.16)	<.001	1.47 (1.33-1.63)	<.001	1.51 (1.41-1.61)	<.001
Moderately prefrail	4.65 (3.77-5.74)	<.001	2.62 (2.4-2.85)	<.001	2.07 (1.95-2.2)	<.001
Frail	11.34 (9.27-13.88)	<.001	4.34 (3.98-4.73)	<.001	2.91 (2.73-3.1)	<.001
1b						
FI-Lab score (continuous)	1.06 (1.06-1.06)	<.001	1.04 (1.04-1.04)	<.001	1.03 (1.03-1.03)	<.001
2a						
F1-Lab score						
Robust	1	—	1	—	1	—
Mildly prefrail	1.53 (1.18-1.98)	<.001	1.34 (1.21-1.48)	<.001	1.36 (1.27-1.45)	<.001
Moderately prefrail	4 (3.24-4.94)	<.001	2.23 (2.04-2.43)	<.001	1.76 (1.65-1.87)	<.001
Frail	9.21 (7.51-11.29)	<.001	3.52 (3.23-3.84)	<.001	2.36 (2.22-2.52)	<.001
2b						
F1-Lab score (continuous)	1.06 (1.05-1.06)	<.001	1.03 (1.03-1.03)	<.001	1.02 (1.02-1.02)	<.001
3a						
F1-Lab score						
Robust	1	—	1	—	1	—
Mildly prefrail	1.48 (1.15-1.92)	<.001	1.27 (1.15-1.41)	<.001	1.29 (1.21-1.38)	<.001
Moderately prefrail	3.76 (3.04-4.64)	<.001	2 (1.83-2.18)	<.001	1.6 (1.5-1.7)	<.001
Frail	8.38 (6.83-10.29)	<.001	2.97 (2.72-3.25)	<.001	2.07 (1.94-2.21)	<.001
3b						

XSL•FO
RenderX

Model and variables	In-hospital mortality		Mortality during the first year after discharge		Mortality beyond the first year after discharge	
	Hazard ratio (95% CI)	*P* value	Hazard ratio (95% CI)	*P* value	Hazard ratio (95% CI)	*P* value
F1-Lab score (continuous)	1.06 (1.05-1.06)	<.001	1.03 (1.03-1.03)	<.001	1.02 (1.02-1.02)	<.001
4a						
F1-Lab score						
Robust	1	—	1	—	1	—
Mildly prefrail	1.44 (1.11-1.86)	.01	1.22 (1.1-1.35)	<.001	1.25 (1.17-1.34)	<.001
Moderately prefrail	3.54 (2.86-4.38)	<.001	1.85 (1.7-2.02)	<.001	1.52 (1.43-1.62)	<.001
Frail	7.63 (6.2-9.39)	<.001	2.65 (2.43-2.9)	<.001	1.91 (1.79-2.04)	<.001
4b						
FI-Lab score (continuous)	1.05 (1.05-1.06)	<.001	1.03 (1.02-1.03)	<.001	1.02 (1.02-1.02)	<.001

[a]Not applicable.

It can be seen that, in all models, the crude model (Model 1); the model adjusted for age and sex (Model 2); the model adjusted for age, sex, and CCI (Model 3); and the model adjusted for age, sex, CCI, and the primary indication for hospitalization (Model 4), an increase in frailty level, as measured using the FI-Lab, was associated with an increase in all-cause mortality during hospitalization, within the first year postdischarge, and throughout the entire follow-up period.

This association was evident when FI-Lab was treated as a categorical variable (robust, mildly prefrail, moderately prefrail, and frail) as well as when it was treated as a continuous variable (per 0.01-point increase in FI-Lab).

For example, in Model 4a (adjusted for age, gender, CCI, and the primary indication for hospitalization), the odds of a patient in the frail group not surviving hospitalization were more than 7 times higher than for those in the robust group (*P*<.001). Similarly, the odds of a patient in the moderately prefrail group dying during hospitalization were more than 3 times higher than for robust patients (*P*<.001). When analyzed as a continuous variable, in Model 4b, every 0.01 increase in the FI-Lab score was associated with a 5.3% increase in in-hospital mortality (*P*<.001), a 2.6% increase in mortality during the first year after discharge (*P*<.001), and a 1.7% increase in mortality by the end of the follow-up period (*P*<.001).

The full multivariable Cox regression results models, including all core confounders (age, gender, CCI, and main reason for hospitalization), are presented in Multimedia Appendix 1 for reference and comprehensive interpretation.

Figure 1 presents the adjusted survival curves for age, gender, CCI, and primary indication for hospitalization of the study population during hospitalization (Figure 1A), during the first year after discharge (Figure 1B), beyond the first year (Figure 1C), and any time (Figure 1D), stratified by FI-Lab–based frailty status. In all frail groups, patients had the lowest survival rates in all the examined time frames (all *P*<.001).

According to the ROC analysis (Figure 2), the inclusion of FI-Lab significantly improved the model performance for (Figure 2A) mortality during index hospitalization (AUC=0.783 vs 0.704, *P*<.001). All the other models were statistically significant but with a lower impact of the FI-Lab: (Figure 2B) mortality within the first year after hospitalization (AUC=0.729 vs 0.712, *P*<.001), (Figure 2C) mortality beyond the first year after hospitalization (AUC=0.655 vs 0.655, *P*=.02), and (Figure 2D) mortality at any time (AUC=0.789 vs 0.773, *P*<.001). This reflects the method's high sensitivity in large datasets and should be interpreted in light of the limited clinical impact of small AUC gains.

XSL•FO
RenderX

Figure 1. Confounder-adjusted survival curves for all-cause mortality according to frailty status assessed using frailty index based on laboratory tests (FI-Lab) scores (A) during the index hospitalization, (B) during the first year after the index hospitalization, (C) beyond the first year after discharge, and (D) at any time.

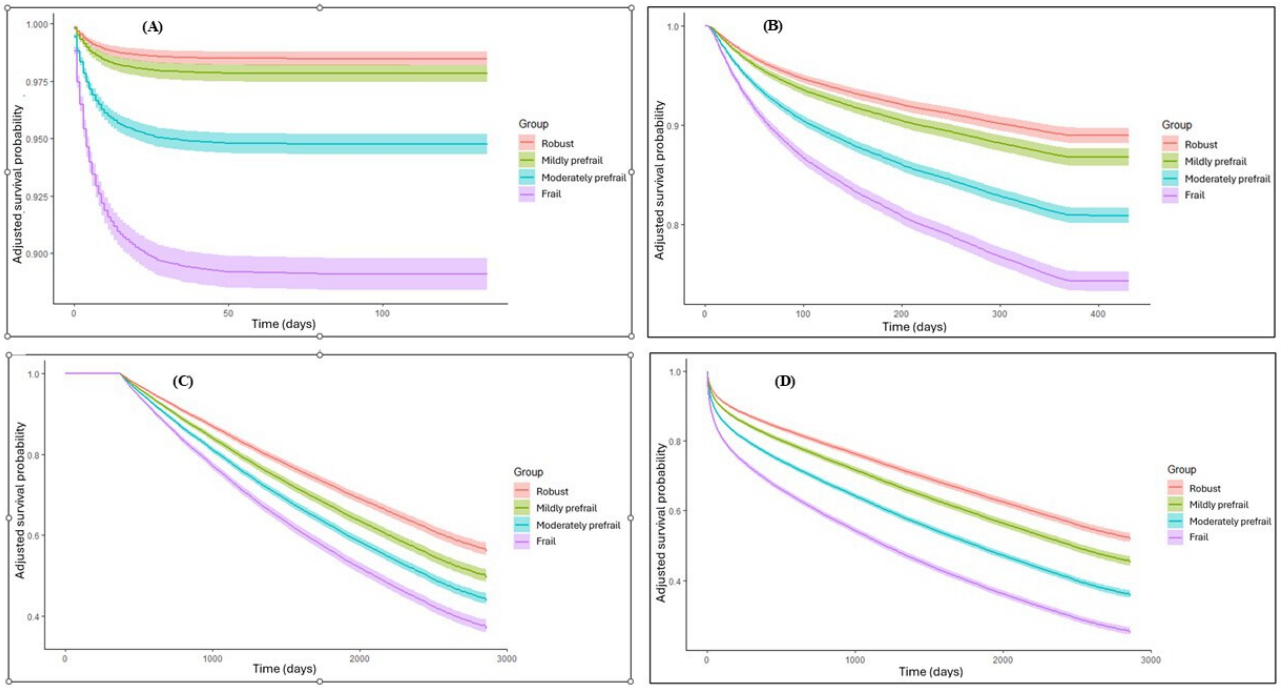
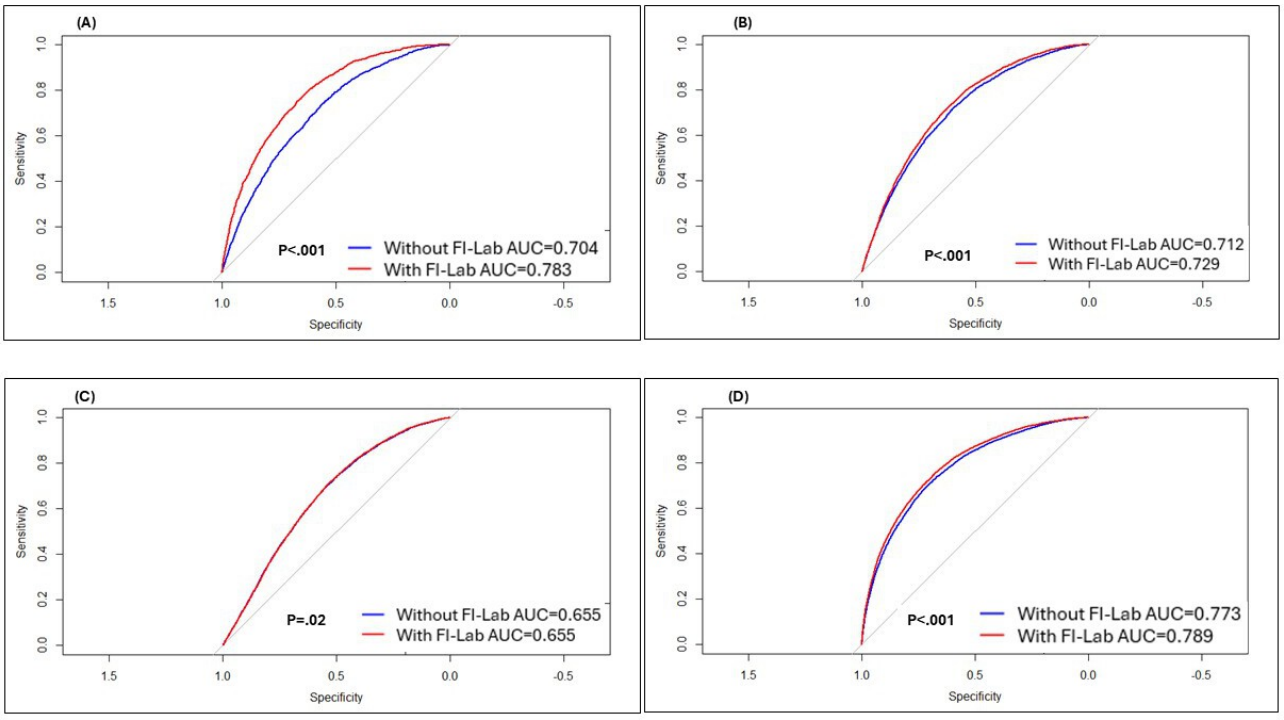


Figure 2. Comparison of receiver operating characteristic (ROC) curves for mortality (A) during the index hospitalization, (B) during the first year after the index hospitalization, (C) beyond the first year after the index hospitalization, and (D) at any time. AUC: area under the curve; FI-Lab: frailty index based on laboratory tests.



Discussion

In this study, involving a sample of more than 33,000 patients aged 65 years and older who were hospitalized in the internal medicine departments of a tertiary medical center, FI-Lab was found to be an independent prognostic factor for increased mortality, both in the short term and long term, whether analyzed as a dichotomous or continuous variable. These results are consistent with prior research with older hospitalized populations. Multimedia Appendix 2 summarizes 25 studies, all of which demonstrated that FI-Lab is a reliable predictor of mortality in this patient group.

Variations can be seen among the studies in terms of mortality rates, even during comparable periods. For example, in the study by Ding et al [29], each 0.01 increase in FI-Lab was associated

XSL•FO
RenderX

with a 3% increase in in-hospital mortality, whereas in the study by Nakashima et al [25], an identical increase in the FI-Lab resulted in a 9.8% increase in in-hospital mortality. This could depend, of course, on the type of study (the first study was retrospective, while the second was prospective); different in-hospital mortality rates (21.6% vs 7.4%), which probably were associated with many variables; the type of patient (patients with sepsis in the first study vs acutely admitted to the geriatric wards in the second study); the number of FI-Lab variables (33 vs 40); and other factors.

Although the association between FI-Lab and all-cause mortality in this study was statistically significant across all time points (during hospitalization, 1-year postdischarge, and throughout the entire follow-up period; see Table 3), the strength of this association declined over time. This is also reflected in the decrease in AUC in the ROC analysis (Figure 2), which was higher during hospitalization than in later time periods.

Additionally, the lower difference in AUC with and without the inclusion of FI-Lab for predicting mortality beyond the first year and for mortality at any time further suggests that the association between FI-Lab and mortality weakens over time. A similar trend has been observed in several other studies in the field (eg, [25,28]). This phenomenon is likely due to the emergence of new variables over time, such as changes in comorbidities, frailty status, and functional status, which may influence mortality and, in turn, alter the relationship between FI-Lab measured at the time of hospitalization and subsequent mortality. A future prospective study that accounts for these evolving variables throughout the follow-up period could help confirm this hypothesis.

The use of FI-Lab in hospitalization offers several notable advantages. As aforementioned, frailty assessment during hospitalization using “classical” models may involve completing a lengthy checklist of data [18]. In some cases, these assessments rely on subjective data (such as a patient’s self-report of physical activity and weight loss) or require the patient’s cooperation (eg, grip strength or walking speed tests) [19]. Therefore, performing frailty assessments using these models can be challenging in a time- and resource-constrained clinical setting.

In contrast, FI-Lab is typically based on laboratory tests that are routinely performed upon admission and, in some cases, supplemented by basic clinical data such as blood pressure and heart rate measurements. As a result, obtaining the data required for FI-Lab calculations generally does not involve additional effort on the part of clinicians [22,24,31]. In cases of acute illness, laboratory results may remain within normal ranges if there is sufficient reserve capacity in the affected organs. A higher FI-Lab score reflects the severity of the damage, number of affected organs and systems, and individual’s level of frailty [40]. FI-Lab likely captures subclinical changes at the organ and system level and may indicate acute or chronic conditions that have the potential to lead to adverse medical outcomes [18]. However, since FI-Lab assesses the current level of acute or chronic accumulated organ dysfunction, it does not necessarily reflect the patient’s comorbidity status prior to the acute illness [25]. In fact, FI-Lab may assess the current state of biological reserves and predict mortality better than comorbidity indices [34]. For these reasons, FI-Lab can serve as a readily available and effective tool for assessing frailty in older hospitalized patients.

However, it is also important to recognize the tool’s limitations. The results of a study by Jäger et al [35] showed that FI-Lab values may vary between the time of hospital admission and subsequent measurements, and these changes significantly impact the prediction of mortality within 6 months and 1 year. Therefore, in cases of prolonged hospitalization, it may be necessary to reassess FI-Lab periodically.

Another significant limitation of the FI-Lab is that it does not assess functional status and cannot replace clinical frailty assessments or comprehensive geriatric evaluations. Relying solely on FI-Lab without a complementary geriatric assessment may lead to significant errors in prognosis prediction and the development of intervention plans [49]. Over-reliance on FI-Lab can, on one hand, prevent patients with high scores from receiving intensive treatment [23] and, on the other, deprive patients with low scores of the support required for frail patients, such as prevention of delirium and functional decline. Therefore, although this method may serve as an initial tool, it *should not be* the sole approach for frailty assessment during hospitalization.

This study has several advantages. First, to our knowledge, it is the largest study conducted among hospitalized patients aged 65 years and older. The applicability of the tool was highlighted by the finding that, among more than 33,000 patients admitted to SUMC during the study period, FI-Lab could not be calculated for only 45 patients. The study included patients treated for a broad spectrum of medical problems across all internal medicine departments of the medical center, minimizing the influence of specific departmental policies on treatment and prognosis. The research is based on data from SUMC’s EMR, which was established nearly 3 decades ago and is considered highly reliable. Additionally, mortality data were obtained through linkage with the Ministry of Interior, ensuring the reliability of these records. Finally, the extended follow-up period in our study represents yet another significant advantage.

This study has several limitations. It was conducted in a single center using a retrospective design, which inherently introduces risks of selection bias and limits the availability of key clinical parameters such as functional, cognitive, and social status, as well as other frailty measures and environmental factors. These data, which are critical for a comprehensive assessment of a patient’s prognosis, are typically not recorded in the medical records of internal medicine departments and could not be retrospectively obtained. Additionally, although the inclusion of a heterogeneous population of hospitalized patients may enhance the generalizability of our findings, it also limits applicability to specific clinical subgroups (eg, patients with pneumonia, sepsis, or myocardial infarction). The primary outcome examined was all-cause mortality, without exploring other clinically relevant outcomes such as delirium, in-hospital falls, institutionalization, readmissions, or length of stay. FI-Lab was assessed only once using laboratory data obtained at admission and during the first days of hospitalization, which may not reflect a consistent time point due to potential changes

XSL•FO
RenderX

in lab values over time. Furthermore, the FI-Lab used in this study was based on 19 laboratory variables; some potentially important tests (eg, liver function) were excluded due to inconsistent availability. As demonstrated in prior studies (see Multimedia Appendix 1), there is no consensus on the optimal number or type of variables to include in FI-Lab, with considerable variability across cohorts. Although our primary aim was to evaluate the predictive utility of FI-Lab, we were unable to incorporate other validated frailty measures (eg, such as the 7-point Clinical Frailty Scale [50] or the Hospital Frailty Risk Score [51]) due to data limitations. It is important to note that previous research in the field has reported inconsistent associations between FI-Lab and other frailty indices [18,20,21,25-27,33,34,38,49,52-58].

Finally, although many associations reached statistical significance—likely due to the large sample size—some effect sizes were small and may not be clinically meaningful at the individual level. Nonetheless, such modest associations may have important implications at the population level, reinforcing the potential of FI-Lab as a scalable tool for risk stratification in hospital settings.

Although many of the associations observed in this study reached statistical significance (eg, $P<.001$), this is not unexpected given the large sample size (>31,000 patients). It is important to distinguish between statistical and clinical significance. For example, the HR of 1.017 per unit increase in FI-Lab reflects a modest effect size that may not be clinically meaningful at the individual level. However, at the population level, even small shifts in risk associated with frailty—as captured by laboratory values—may lead to substantial differences in outcomes across large health care systems. These findings support the utility of FI-Lab as a scalable risk stratification tool while also emphasizing the need for cautious interpretation of small effect sizes in clinical decision-making.

In conclusion, this study, conducted with a large cohort of patients aged 65 years and older who were hospitalized for diverse indications in the internal medicine departments of a large tertiary hospital, demonstrated that FI-Lab is an independent predictor of all-cause mortality during and after hospitalization, even after adjusting for multiple covariates, including the primary indication for admission.

FI-Lab may serve as an effective and easily accessible tool, though not a standalone method, for assessing frailty in older hospitalized patients. Automating FI-Lab calculation at the time of hospital admission could provide emergency department clinicians with additional, meaningful information to support clinical decision-making.

Data Availability

The datasets generated and analyzed during this study are not publicly available due to legal and data security reasons but are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Full Cox proportional hazard models of the frailty index based on laboratory tests (FI-Lab) for predicting all-cause mortality.
[DOCX File, 36 KB - aging_v8i1e70204_app1.docx]

Multimedia Appendix 2
Summary of 25 studies that examined the association between the frailty index based on laboratory tests (FI-Lab) and mortality.
[DOCX File, 78 KB - aging_v8i1e70204_app2.docx]

References

1. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet 2013 Mar 2;381(9868):752-762. [doi: 10.1016/S0140-6736(12)62167-9] [Medline: 23395245]
2. Bandeen-Roche K, Seplaki CL, Huang J, et al. Frailty in older adults: a nationally representative profile in the United States. J Gerontol A Biol Sci Med Sci 2015 Nov;70(11):1427-1434. [doi: 10.1093/gerona/glv133] [Medline: 26297656]
3. Bonjour T, Waeber G, Marques-Vidal P. Trends in prevalence and outcomes of frailty in a Swiss university hospital: a retrospective observational study. Age Ageing 2021 Jun 28;50(4):1306-1313. [doi: 10.1093/ageing/afaa278] [Medline: 33453112]
4. Ligthart-Melis GC, Luiking YC, Kakourou A, Cederholm T, Maier AB, de van der Schueren MAE. Frailty, sarcopenia, and malnutrition frequently (co-)occur in hospitalized older adults: a systematic review and meta-analysis. J Am Med Dir Assoc 2020 Sep;21(9):1216-1228. [doi: 10.1016/j.jamda.2020.03.006] [Medline: 32327302]
5. Chong E, Ho E, Baldevarona-Llego J, Chan M, Wu L, Tay L. Frailty and risk of adverse outcomes in hospitalized older adults: a comparison of different frailty measures. J Am Med Dir Assoc 2017 Jul 1;18(7):638. [doi: 10.1016/j.jamda.2017.04.011] [Medline: 28587850]

XSL•FO
RenderX

6. Hubbard RE, Peel NM, Samanta M, Gray LC, Mitnitski A, Rockwood K. Frailty status at admission to hospital predicts multiple adverse outcomes. Age Ageing 2017 Sep 1;46(5):801-806. [doi: 10.1093/ageing/afx081] [Medline: 28531254]
7. Romero-Ortuno R, Wallis S, Biram R, Keevil V. Clinical frailty adds to acute illness severity in predicting mortality in hospitalized older adults: an observational study. Eur J Intern Med 2016 Nov;35:24-34. [doi: 10.1016/j.ejim.2016.08.033] [Medline: 27596721]
8. Comans TA, Peel NM, Hubbard RE, Mulligan AD, Gray LC, Scuffham PA. The increase in healthcare costs associated with frailty in older people discharged to a post-acute transition care program. Age Ageing 2016 Mar;45(2):317-320. [doi: 10.1093/ageing/afv196] [Medline: 26769469]
9. García-Nogueras I, Aranda-Reneo I, Peña-Longobardo LM, Oliva-Moreno J, Abizanda P. Use of health resources and healthcare costs associated with frailty: the FRADEA study. J Nutr Health Aging 2017;21(2):207-214. [doi: 10.1007/s12603-016-0727-9] [Medline: 28112778]
10. Joosten E, Demuynck M, Detroyer E, Milisen K. Prevalence of frailty and its ability to predict in hospital delirium, falls, and 6-month mortality in hospitalized older patients. BMC Geriatr 2014 Jan 6;14:1. [doi: 10.1186/1471-2318-14-1] [Medline: 24393272]
11. Wallis SJ, Wall J, Biram RWS, Romero-Ortuno R. Association of the clinical frailty scale with hospital outcomes. QJM 2015 Dec;108(12):943-949. [doi: 10.1093/qjmed/hcv066] [Medline: 25778109]
12. Mitnitski AB, Mogilner AJ, Rockwood K. Accumulation of deficits as a proxy measure of aging. ScientificWorldJournal 2001 Aug 8;1:323-336. [doi: 10.1100/tsw.2001.58] [Medline: 12806071]
13. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
14. Buta BJ, Walston JD, Godino JG, et al. Frailty assessment instruments: systematic characterization of the uses and contexts of highly-cited instruments. Ageing Res Rev 2016 Mar;26:53-61. [doi: 10.1016/j.arr.2015.12.003] [Medline: 26674984]
15. Basic D, Shanley C. Frailty in an older inpatient population: using the clinical frailty scale to predict patient outcomes. J Aging Health 2015 Jun;27(4):670-685. [doi: 10.1177/0898264314558202] [Medline: 25414168]
16. Leite HT, Manhães AC, Antunes LA, Chan T, Hajj-Boutros G, Morais JA. The implementation of a geriatrics co-management model of care reduces hospital length of stay. Healthcare (Basel) 2022 Oct 29;10(11):2160. [doi: 10.3390/healthcare10112160] [Medline: 36360501]
17. Rezaei-Shahsavarloo Z, Atashzadeh-Shoorideh F, Gobbens RJJ, Ebadi A, Ghaedamini Harouni G. The impact of interventions on management of frailty in hospitalized frail older adults: a systematic review and meta-analysis. BMC Geriatr 2020 Dec 3;20(1):526. [doi: 10.1186/s12877-020-01935-8] [Medline: 33272208]
18. Ritt M, Jäger J, Ritt JI, Sieber CC, Gaßmann KG. Operationalizing a frailty index using routine blood and urine tests. Clin Interv Aging 2017;12:1029-1040. [doi: 10.2147/CIA.S131987] [Medline: 28721031]
19. Jin X, Ren Y, Shao L, et al. Prevalence of frailty and prediction of mortality in Chinese cancer patients using a frailty index-based clinical algorithm-a multicentre study. Cancer Med 2021 Sep;10(18):6207-6217. [doi: 10.1002/cam4.4155] [Medline: 34318626]
20. Howlett SE, Rockwood MRH, Mitnitski A, Rockwood K. Standard laboratory tests to identify older adults at increased risk of death. BMC Med 2014 Oct 7;12:171. [doi: 10.1186/s12916-014-0171-9] [Medline: 25288274]
21. Ellis HL, Wan B, Yeung M, et al. Complementing chronic frailty assessment at hospital admission with an electronic frailty index (FI-Laboratory) comprising routine blood test results. CMAJ 2020 Jan 6;192(1):E3-E8. [doi: 10.1503/cmaj.190952] [Medline: 31907228]
22. Kim Y, Song K, Kang CM, Lee H. Impact of preoperative laboratory frailty index on mortality and clinical outcomes in older surgical patients with cancer. Sci Rep 2022;12(1):9200. [doi: 10.1038/s41598-022-13426-4]
23. Lim A, Choi M, Jang Y, Lee H. Preoperative frailty based on laboratory data and postoperative health outcomes in patients undergoing coronary artery bypass graft surgery. Heart Lung 2022;56:1-7. [doi: 10.1016/j.hrtlng.2022.05.005] [Medline: 35598421]
24. Nagae M, Umegaki H, Nakashima H, Nishiuchi T. FI-Lab in the emergency department and adverse outcomes among acutely hospitalized older adults. Arch Gerontol Geriatr 2025 Feb;129:105649. [doi: 10.1016/j.archger.2024.105649] [Medline: 39368270]
25. Nakashima H, Nagae M, Komiya H, et al. Combined use of the Clinical Frailty Scale and laboratory tests in acutely hospitalized older patients. Aging Clin Exp Res 2023 Sep;35(9):1927-1935. [doi: 10.1007/s40520-023-02477-w] [Medline: 37378813]
26. Li M, She Q, Tu J, et al. Association of frailty with clinical outcomes in chronic obstructive pulmonary disease: a retrospective longitudinal cohort study. Heliyon 2023 May;9(5):e15764. [doi: 10.1016/j.heliyon.2023.e15764]
27. Ysea-Hill O, Gomez CJ, Mansour N, et al. The association of a frailty index from laboratory tests and vital signs with clinical outcomes in hospitalized older adults. J Am Geriatr Soc 2022 Nov;70(11):3163-3175. [doi: 10.1111/jgs.17977] [Medline: 35932256]
28. Bai W, Hao B, Xu L, Qin J, Xu W, Qin L. Frailty index based on laboratory tests improves prediction of short-and long-term mortality in patients with critical acute myocardial infarction. Front Med (Lausanne) 2022 Dec 6;9:1070951. [doi: 10.3389/fmed.2022.1070951] [Medline: 36561712]

XSL•FO
RenderX

29. Ding H, Li X, Zhang X, Li J, Li Q. The association of a frailty index derived from laboratory tests and vital signs with clinical outcomes in critical care patients with septic shock: a retrospective study based on the MIMIC-IV database. BMC Infect Dis 2024 Jun 10;24(1):573. [doi: 10.1186/s12879-024-09430-w] [Medline: 38853273]
30. Gu JJ, Liu Q, Zheng LJ. A frailty assessment tool to predict in-hospital mortality in patients with acute exacerbations of chronic obstructive pulmonary disease. Int J Chron Obstruct Pulmon Dis 2021;16:1093-1100. [doi: 10.2147/COPD.S300980] [Medline: 33907395]
31. Veronese N, Briganò V, Ciriminna S, et al. Prognostic value of a laboratory index of frailty in older patients hospitalized for COVID-19: the COMEPA study. J Frailty Aging 2024;13(3):213-217. [doi: 10.14283/jfa.2024.29] [Medline: 39082764]
32. Zhao H, Tu J, She Q, et al. Prognostic significance of frailty in hospitalized elderly patients with community-acquired pneumonia: a retrospective cohort study. BMC Geriatr 2023 May 17;23(1):308. [doi: 10.1186/s12877-023-04029-3] [Medline: 37198576]
33. Guan L, Soh CH, Reijnierse EM, Lim WK, Maier AB. Association of a modified laboratory frailty index with adverse outcomes in geriatric rehabilitation inpatients: RESORT. Mech Ageing Dev 2022 Apr;203:111648. [doi: 10.1016/j.mad.2022.111648] [Medline: 35219637]
34. Hao B, Xu W, Gao W, et al. Association between frailty assessed using two electronic medical record-based frailty assessment tools and long-term adverse prognosis in older critically ill survivors. J Nutr Health Aging 2023;27(8):649-655. [doi: 10.1007/s12603-023-1961-6] [Medline: 37702338]
35. Jäger J, Sieber CC, Gaßmann KG, Ritt M. Changes of a frailty index based on common blood and urine tests during a hospital stay on geriatric wards predict 6-month and 1-year mortality in older people. Clin Interv Aging 2019;14:473-484. [doi: 10.2147/CIA.S191117] [Medline: 30880928]
36. Sohn B, Choi JW, Hwang HY, Jang MJ, Kim KH, Kim KB. Frailty index is associated with adverse outcomes after aortic valve replacement in elderly patients. J Korean Med Sci 2019 Aug 12;34(31):e205. [doi: 10.3346/jkms.2019.34.e205] [Medline: 31392853]
37. Wang S, Wang L, Wang Y, et al. Association between frailty index based on laboratory tests and all-cause mortality in critically ill patients with heart failure. ESC Heart Fail 2024 Dec;11(6):3662-3673. [doi: 10.1002/ehf2.14948] [Medline: 38979803]
38. Engvig A, Wyller TB, Skovlund E, et al. Association between clinical frailty, illness severity and post-discharge survival: a prospective cohort study of older medical inpatients in Norway. Eur Geriatr Med 2022 Apr;13(2):453-461. [doi: 10.1007/s41999-021-00555-8] [Medline: 34417977]
39. Kim CH, Kang Y, Kim JS, Sohn SH, Hwang HY. Association between the frailty index and clinical outcomes after coronary artery bypass grafting. J Chest Surg 2022 Jun 5;55(3):189-196. [doi: 10.5090/jcs.21.147] [Medline: 35440518]
40. Klausen HH, Petersen J, Bandholm T, et al. Association between routine laboratory tests and long-term mortality among acutely admitted older medical patients: a cohort study. BMC Geriatr 2017 Mar 1;17(1):62. [doi: 10.1186/s12877-017-0434-3] [Medline: 28249621]
41. Wang Y, Zhang R, Shen Y, Su L, Dong B, Hao Q. Prediction of chemotherapy adverse reactions and mortality in older patients with primary lung cancer through frailty index based on routine laboratory data. Clin Interv Aging 2019;14:1187-1197. [doi: 10.2147/CIA.S201873] [Medline: 31308642]
42. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40(5):373-383. [doi: 10.1016/0021-9681(87)90171-8] [Medline: 3558716]
43. Schoenfeld DA. Sample-size formula for the proportional-hazards regression model. Biometrics 1983 Jun;39(2):499-503. [Medline: 6354290]
44. R Core Team. R: a language and environment for statistical computing version 443. R Foundation for Statistical Computing. 2024. URL: https://www.R-project.org/ [accessed 2025-08-14]
45. Denz R. AdjustedCurves: confounder-adjusted survival curves and cumulative incidence functions. The Comprehensive R Archive Network. URL: https://CRAN.R-project.org/package=adjustedCurves [accessed 2025-08-15]
46. Therneau TM, Lumley T, Atkinson E, Crowson C. Survival: survival analysis. The Comprehensive R Archive Network. URL: https://CRAN.R-project.org/package=survival [accessed 2025-08-15]
47. Wickham H. Ggplot2: Elegant Graphics for Data Analysis: Springer-Verlag; 2016.
48. Kassambara A, Kosinski M, Biecek P, Fabian S. Survminer: drawing survival curves using “ggplot2”. The Comprehensive R Archive Network. URL: https://CRAN.R-project.org/package=survminer [accessed 2025-08-18]
49. Nakashima H, Watanabe K, Komiya H, et al. Frailty index based on common laboratory tests for patients starting home-based medical care. J Am Med Dir Assoc 2024 Sep;25(9):105114. [doi: 10.1016/j.jamda.2024.105114] [Medline: 38950586]
50. Rockwood K, Song X, MacKnight C, et al. A global clinical measure of fitness and frailty in elderly people. CMAJ 2005 Aug 30;173(5):489-495. [doi: 10.1503/cmaj.050051] [Medline: 16129869]
51. Gilbert T, Neuburger J, Kraindler J, et al. Development and validation of a Hospital Frailty Risk Score focusing on older people in acute care settings using electronic hospital records: an observational study. Lancet 2018 May 5;391(10132):1775-1782. [doi: 10.1016/S0140-6736(18)30668-8] [Medline: 29706364]

XSL•FO
RenderX

52. Blodgett JM, Pérez-Zepeda MU, Godin J, et al. Frailty indices based on self-report, blood-based biomarkers and examination-based data in the Canadian Longitudinal Study on Aging. Age Ageing 2022 May 1;51(5):afac075. [doi: 10.1093/ageing/afac075] [Medline: 35524747]
53. Blodgett JM, Theou O, Howlett SE, Rockwood K. A frailty index from common clinical and laboratory tests predicts increased risk of death across the life course. Geroscience 2017 Aug;39(4):447-455. [doi: 10.1007/s11357-017-9993-7] [Medline: 28866737]
54. Blodgett JM, Theou O, Howlett SE, Wu FCW, Rockwood K. A frailty index based on laboratory deficits in community-dwelling men predicted their risk of adverse health outcomes. Age Ageing 2016 Jul;45(4):463-468. [doi: 10.1093/ageing/afw054] [Medline: 27076524]
55. Chao CT, Huang JW, Chiang CK, Hung KY, COhort of GEriatric Nephrology in NTUH (COGENT) Study Group. Applicability of laboratory deficit-based frailty index in predominantly older patients with end-stage renal disease under chronic dialysis: a pilot test of its correlation with survival and self-reported instruments. Nephrology (Carlton) 2020 Jan;25(1):73-81. [doi: 10.1111/nep.13583] [Medline: 30834584]
56. Mitnitski A, Collerton J, Martin-Ruiz C, et al. Age-related frailty and its association with biological markers of ageing. BMC Med 2015 Jul 13;13:161. [doi: 10.1186/s12916-015-0400-x] [Medline: 26166298]
57. Yang M, Zhuo Y, Hu X, Xie L. Predictive validity of two frailty tools for mortality in Chinese nursing home residents: frailty index based on common laboratory tests (FI-Lab) versus FRAIL-NH. Aging Clin Exp Res 2018 Dec;30(12):1445-1452. [doi: 10.1007/s40520-018-1041-7] [Medline: 30259498]
58. Rockwood K, McMillan M, Mitnitski A, Howlett SE. A frailty index based on common laboratory tests in comparison with a clinical frailty index for older adults in long-term care facilities. J Am Med Dir Assoc 2015 Oct 1;16(10):842-847. [doi: 10.1016/j.jamda.2015.03.027] [Medline: 25952475]

Abbreviations

AUC: area under the curve
CCI: Charlson Comorbidity Index
EMR: electronic medical record
FI-Lab: frailty index based on laboratory tests
HR: hazard ratio
ICU: intensive care unit
ROC: receiver operating characteristic
SUMC: Soroka University Medical Center

Edited by F Puga; submitted 17.12.24; peer-reviewed by HS Zhang, M Nagae; revised version received 25.06.25; accepted 01.07.25; published 10.09.25.

Please cite as:
Pasternak E, Freud T, Press Y
Association Between the Frailty Index Based on Laboratory Tests and All-Cause Mortality in Hospitalized Older Adults: Retrospective Cohort Study
JMIR Aging 2025;8:e70204
URL: https://aging.jmir.org/2025/1/e70204
doi:10.2196/70204

© Eyal Pasternak, Tamar Freud, Yan Press. Originally published in JMIR Aging (https://aging.jmir.org), 10.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Handwriting in Mild Cognitive Impairment: Reliability Assessment and Machine Learning–Based Screening

Simone Toffoli[1], MSc; Carlo Abbate[2], PhD; Francesca Lunardini[3], PhD; Edoardo Corno[1], MSc; Nicholas Diani[2], MSc; Alessia Gallucci[2], PhD; Emanuele Tomasini[2]; Pietro Davide Trimarchi[2], PhD; Simona Ferrante[1,4], PhD

[1]Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan, Italy

[2]IRCCS Fondazione Don Carlo Gnocchi ONLUS, Milan, Italy

[3]Center for Clinical Neuroscience, Hospital Los Madroños, Madrid, Spain

[4]LEARNLab, Joint Research Platform, IRCCS Istituto Neurologico Carlo Besta, Milan, Italy

Corresponding Author:
Simone Toffoli, MSc
Department of Electronics, Information and Bioengineering
Politecnico di Milano
Via Ponzio 34/5
Milan, 20133
Italy
Phone: 39 0223999004
Email: simone.toffoli@polimi.it

Abstract

Background: Mild cognitive impairment (MCI) is a precursor of dementia. Therefore, MCI identification and monitoring are crucial to delaying dementia onset. Given the limits of existing clinical tests, objective support tools are needed.

Objective: This work investigates quantitative handwriting analysis, tailored to enable domestic monitoring, as a noninvasive approach for MCI screening and assessment.

Methods: A sensorized ink pen, used on paper and equipped with sensors, memory, and a communication unit, was used for data acquisition. The tasks included writing a grocery list and free text to mimic daily life handwriting, and a clinical dictation test (parole-non-parole [PnP] test), featuring regular, irregular, and made-up words, aimed at assessing MCI dysgraphia. From the recorded data, 106 indicators describing the performance in terms of time, fluency, exerted force, and pen inclination were computed. A total of 57 patients with MCI were recruited, of whom 45 performed a test-retest protocol. The indicators were examined to assess their test-retest reliability. The indicators from the test repetition were used to assess their relationship with the scores of clinical tests via correlation analysis. For the PnP test, differences in the indicators among the 3 types of words were statistically investigated. These analyses were conducted separately for the cursive (2/3 of the sample) and block letters (1/3 of the sample) allographs, with the level of significance set at 5%. Data from healthy older adults were available for the grocery list (34 participants) and free text (45 participants) tasks. These were exploited to build machine learning classification models for the distinction between patients with MCI and healthy controls.

Results: When dealing with reliability, 93% and 44% of the indicators were characterized by a significant reliability of at least moderate intensity for cursive and block letters respectively. As for the correlation analysis, patients with preserved cognitive status and daily life functionality were associated with significantly better temporal performances, both in free writing and PnP. The analysis of PnP highlighted the presence of surface dysgraphia in the recruited sample, as irregular words showed significantly worse temporal indicators with respect to regular and made-up ones. The classification models' built-in free writing data achieved accuracies ranging from 0.80 to 0.93 and F_1-scores from 0.81 to 0.92 according to the input dataset.

Conclusions: The presented results suggest the suitability of ecological handwriting analysis for the all-around monitoring of MCI, from early screening to disease progression evaluation.

(JMIR Aging 2025;8:e73074) doi:10.2196/73074

KEYWORDS

mild cognitive impairment; handwriting; sensorized ink pen; machine learning; parole-non-parole test

XSL•FO
RenderX

Introduction

In the last decades, the world has witnessed progressive aging. This trend brings about various challenges at the social, economic, and health care levels, as it directly affects citizens' health, compromising several aspects of their daily lives [1]. First, older individuals experience greater difficulty in responding to external stimuli, a condition that leads to isolation and inertia. Second, a gradual, albeit sometimes slow, decline is observed, causing a state referred to as "frailty" [2]. In this state, the individual becomes more susceptible to motor-related accidents, resulting in disability and an increase in hospitalization. Furthermore, aging is linked to a higher prevalence of neurodegenerative diseases, for which treatments that slow down the progression of the disease are the only available interventions.

Among these disorders, dementia is one of the most prevalent, being characterized by a decline in cognitive function, memory loss, and changes in behavior [3]. Dementia leads to a gradual loss in the ability to perform daily tasks independently, causing patients to often require round-the-clock care and supervision. This can translate into emotional and physical toll on family members and caregivers, who may experience stress, burnout, and financial strain. The health care systems are not immune to the negative consequences of the disorder, since it is associated with extensive long-term care, medications, and support services [4].

Clinically, a transitional phase between the healthy state and dementia has been identified and denoted as mild cognitive impairment (MCI) [1]. MCI can manifest heterogeneously. While patients with amnestic MCI experience memory-related cognitive loss, participants diagnosed with nonamnestic MCI face impairments in the domains of attention, language, and visuospatial ability [5]. From an epidemiological standpoint, the aggregate prevalence of MCI among older people ranges from 15% to 20% [6-9]. The problem lies in the relevant yearly percentage (up to 15%) of cases that are diagnosed with dementia after having previously been in the MCI state. This underscores the need for MCI early screening, aimed at preventing the progression to dementia [10]. Unfortunately, the broad spectrum of symptoms and the typically nonsignificant impact on the participants' independent life make the detection of MCI quite challenging.

Cognitive tests are typically used in clinical practice for this purpose since they are quicker and cheaper than imaging techniques [11,12]. The existing tests, all conducted under the clinician's supervision, vary in the assessed cognitive domains (eg, memory, language, attention) and in the tasks proposed to the individual, but are in general capable of providing a snapshot of one's cognitive status. The most common are the Mini-Mental State Examination (MMSE) [12], the Montreal Cognitive Assessment (MoCA) [13], and the clock drawing test (CDT) [14]. Despite being largely used, such clinical tests experience a few drawbacks. As they depend on the clinician administering the test, interrater reliability and subjectivity in the assessment are a concern. Moreover, to assess all the relevant cognitive domains, a combination of different tests should be considered, thus leading to increased administration time. Finally, the scales are usually characterized by low granularity, thus making it difficult to establish universally accepted cutoff scores. Thus, there is a need for a cheap, noninvasive, objective solution to complement the standard clinical procedure for MCI screening [6].

In this sense, the recent literature claimed the importance of handwriting: given its complexity, it can serve as a biomarker for cognitive decline [15]. This is because handwriting is the result of the interaction between various cognitive processes: attention, language, short and long-term memory, motor planning, visual-spatial skills, in-hand manipulation, fine motor control, and sensory awareness of the fingers [16]. With the goal of characterizing the handwriting gesture of patients with MCI or with dementia with respect to healthy participants, several studies have been conducted. In Werner et al [17], a group of 41 controls, 31 patients with MCI, and 22 patients with mild dementia were recruited and performed copying tasks of increasing complexity on a digitizer. Significant differences emerged, with both groups of patients showing increased execution times and reduced pressure during the motor task. Conversely, the Livescribe Echo Pen, used on the Livescribe paper sheet, was used in Kawa et al [16] to compare the handwriting of 37 controls and 37 patients diagnosed with MCI. Again, writing slowness emerged as a peculiar characteristic of the latter group. This was coupled with significantly higher and wider written traces. The extracted features were then used to build a linear discriminant classifier, achieving an accuracy of 70%. The combination of a specific digital pen and paper was also adopted in [18], investigating Chinese handwriting in patients affected by Alzheimer disease, for which MCI is a common precursor. The results confirmed the existing literature, as worse temporal measures, both on paper and in the air, were associated with the disease. A higher degree of pressure applied to the writing surface and a poorer control of the same were revealed as well. A classification study, aimed at distinguishing patients in the early stage of dementia from healthy participants, was conducted in [19]. Cursive loops were acquired with a WACOM digitizer, to then apply unsupervised clustering followed by a Bayesian-based classifier. The approach reached an accuracy of 74%. Execution speed, kinematics complexity, acceleration, and pressure were computed from the data recorded by a WACOM digital graphic tablet in [20]. Handwriting tasks, ranging from drawings to spontaneous sentence production, were performed on a paper sheet fixed onto the tablet surface by a group of 17 healthy participants (healthy control [HC]), 12 patients with MCI, and 23 patients with dementia (AD). Performance results ranged from 69.2% (HC vs MCI vs AD) to 96.6% (HC vs MCI). Pressure features were the most influential in the model output. In Chai et al [21], simple graphic tasks, going from repeatedly writing the letter "T" to drawing a pentagram, were performed by 39 HCs and 40 patients with MCI on a commercial tablet using a digital pen. After extracting 8 features, a binary support vector classifier achieved an F_1-score of 83.1%.

Given the evidence that emerged in the literature, this work aims to evaluate handwriting in patients with MCI by changing the paradigm in terms of the acquisition tool and the proposed

XSL•FO
RenderX

handwriting tasks, with an eye on the feasibility of remote, ecological monitoring. Indeed, while able to provide relevant, objective data on the handwriting performance, limitations can be pointed out for the approaches found in the literature. Digitizers can still represent a barrier for older adults [22], who are typically the ones affected by MCI. Moreover, the writing execution on the digitizer surface, characterized by a reduced friction with respect to paper, is inevitably altered [23]. On the other hand, the methods proposed in studies by Kawa et al [16] and Qi et al [18] rely on the combination of a digital device and writing surface, potentially hampering their everyday employment. For these reasons, in this study, handwriting data were collected with a sensorized ink pen [24] that writes on normal paper. As for the tasks, the "Parole-non-Parole" test (PnP), aimed at assessing the origin of dysgraphic manifestations associated with MCI, was administered together with 2 unconstrained exercises, aimed at mimicking the natural handwriting one could perform in the domestic scenario. From the collected data, a series of indicators was extracted, and 3 specific objectives were investigated. First, the test-retest reliability of the indicators was assessed, both for the clinical PnP test and the 2 ecological tasks, to evaluate whether the proposed approach could be suitable for the longitudinal monitoring of handwriting in patients with MCI. Second, the indicators were examined in terms of their quantitative support for the clinical evaluation. This was done by assessing the indicators' correlation with the scores of clinical tests. Moreover, the PnP was specifically analyzed to understand whether the indicators can be used to support the identification of lexical and phonological dysgraphia. Last, the unconstrained handwriting tasks were considered to build classification models to distinguish between patients with MCI and healthy participants, with the aim of enabling the ecological, domestic monitoring of handwriting as a screening tool for the first signs of cognitive decline.

Methods

Ethical Considerations

The protocol received the approval (ID 07_20/05/2021) from the Ethical Committee of the section IRCCS Fondazione Don Carlo Gnocchi of the Ethical Committee IRCCS Regione Lombardia. An information sheet for study participation was provided to all eligible participants, and informed consent was obtained from each of them.

Participants and Protocol

Participants were recruited by IRCCS Fondazione Don Carlo Gnocchi Milano. The inclusion criteria were (1) confirmed diagnosis of MCI according to the recommendations by Albert et al [25]; (2) being aged older than 65 years; (3) having completed elementary school; (4) having a Clinical Dementia Rating (CDR) scale [26] score less than or equal to 1; (5) obtaining negative neurological examination results for both lateral and extrapyramidal signs; (6) having trail making test A (TMT-A) results within the normal range for age and education according to the Italian calibration by Giovagnoli et al [27]; and (7) not having motor limitations that hinder task performance. Candidates were excluded in case of the presence of cognitive impairment of other origin (cerebral stroke, multiple sclerosis, and Parkinson disease), psychiatric disorders in either recent or remote anamnesis, severe sensory or intellectual deficits, and history of substance abuse or alcohol misuse. The recruited patients underwent clinical assessment, including MMSE, CDR, CDT, TMT-A, basic activity of daily living (B-ADL) [28], and instrumental activity of daily living (I-ADL) [29].

The sample size was determined by taking the assessment of the extracted indicators' test-retest reliability as the main aim, since the scientific literature lacks similar works on patients with MCI. The only comparable study pertains to a systematic review that reports the psychometric properties of motor assessments (mobility, walking, balance, and functional performance) for patients with dementia, yielding intraclass correlation coefficients (ICC) between 0.42 and 0.99 [30]. On the other hand, the reliability of the proposed sensorized pen was demonstrated, the results showing $0.68<ICC<0.99$, only in older adults without any diagnoses [24]. Given the scarcity of specific evidence in the population with MCI, a conservative approach was selected. Considering the minimum ICC value of 0.42 as a target, a statistical power of 80% and a dropout rate of 10%, the sample size was set to 40.

The protocol was administered under the clinician's supervision in a test-retest fashion on the same day, with at least a 30-minute interval between the 2 repetitions (test and retest) of the protocol. The time interval was chosen to ensure that the participants were under the same conditions between test and retest. In this sense, potential learning effects were not a concern given the nature of the protocol (no memory-related tasks were administered). Each patient was provided with blank sheets of paper to perform three writing tasks in random order, 1 minute apart: (1) writing a grocery list of at least 6 items (List); (2) composing a content-free text of at most 7 lines (Text); and (3) executing the PnP test [31], where the participant wrote a series of 15 words under dictation. These included 7 regular words (simple and easily understandable Italian words: "tavolo," "cartone," "attore," "sorella," "postini," "cuoceva," and "democratiche"), 3 irregular words (requiring proper semantic and lexical comprehension of the Italian language to be written correctly: "cieco," "nacquero," and "conoscenza"), and 5 made-up (MU) words (words that do not exist in the Italian dictionary: "sabomi," "fule," "descia," "sterpanzi," and "getrunna"). The order in which words were presented was fixed across participants and protocol repetitions. Orthographic errors were then evaluated to assess the eventual presence of lexical dysgraphia, which is associated with errors in irregular words, or phonological dysgraphia, linked with errors in MU words.

Sensorized Ink Pen

The handwriting exercises were carried out using a sensorized ink pen [24], able to record linear acceleration and angular velocity on 3 orthogonal axes, together with the force applied on the pen tip with a sampling frequency of 50 Hz. The device was successfully used for the quantitative analysis of drawing tasks in patients with Parkinson disease [32] and demonstrated to be capable of recognizing healthy, older adults of different ages from unconstrained handwriting samples [33]. Data acquisition was controlled by the operator administering the

XSL•FO
RenderX

protocol through a custom iOS app. On top of managing the start and stop of the recordings, the app allows users to store 4 different types of labels upon pressing buttons displayed on the user interface. Labels were used to tag the 3 types of words of the PnP test, so that they could be successively analyzed.

Data Analysis

The data recorded by the sensorized ink pen were elaborated in MATLAB R2022b (MathWorks, Inc) to extract 106 indicators measuring heterogeneous characteristics of the handwriting gesture, namely the temporal performance, the force patterns applied on the writing surface, the movement fluency, the pen inclination with respect to the vertical axis (tilt), the high frequency oscillations both in the domains of time and frequency. The complete description of the indicators can be found in [32-35]. The indicators presented in the Results section are described in Textbox 1 for each domain. The details on their computation can be found in the Multimedia Appendix 1. The ones with (S) after the name were selected a priori for the analysis of reliability, relationships with clinical scores, and MCI dysgraphia characterization. If all the indicators within a domain were selected a priori, (S) is found at the end of the name of the domain.

XSL•FO
RenderX

Textbox 1. Description of the indicators extracted from the sensorized ink pen data, divided by domain.

Temporal domain (S):

- Execution time: the time in seconds required to complete the execution.
- Rel stroke num: the number of strokes (tracts on paper) generated in the time unit.
- Mean on sheet: the average time in seconds required to produce a stroke.
- On sheet CV: the coefficient of variation (CV) of the time required to produce a stroke.
- On sheet ratio: the percentage of time spent with the tip in contact with the paper with respect to the total time required to complete the execution.
- Mean in air: the average time in seconds spent with the pen in the air.
- In air CV: the CV of the time spent with the pen in the air.
- Mean pause: the average duration in seconds of pauses (in air moments longer than 2 seconds).
- Pause num: the number of pauses in the execution.
- Air sheet ratio: the ratio between the mean in air and the mean on the sheet.

Fluency domain:

- LDLJ A (S): the logarithmic dimensionless jerk of the acceleration signal.
- LDLJ G (S): the logarithmic dimensionless jerk of the angular velocity signal.
- SPARC (S): the spectral arc length of the angular velocity signal.
- G NC (S): the number of changes (NC) in the time unit in the angular velocity signal.
- A NC: the number of changes (NC) in the time unit in the acceleration signal.
- Cons peak diff G (S): the absolute average difference in deg/s between consecutive extrema in the angular velocity signal.

Force domain (S):

- Mean force: the mean force in arbitrary units exerted while the tip is on paper.
- Force overshoot: the difference between the maximum and median exerted force, in arbitrary units.
- Force NC: the NC in the time unit in the force signal.
- Cons peak diff F: the absolute average difference in arbitrary units between consecutive extrema in the force signal.

Pen inclination domain:

- Mean tilt (S): the average pen inclination with respect to gravity in degrees.
- Tilt CV (S): the CV of the pen inclination with respect to gravity.
- Tilt median bandwidth: the width of the frequency interval that contains 50% of the tilt power spectrum total power.

High frequency oscillations (time domain):

- A ApEn: approximate entropy of the acceleration signal.

High frequency oscillations (frequency domain):

- G peak power: the power at the angular velocity power spectrum peak.
- G RPW 8: the median relative power of the angular velocity power spectrum in a 3 Hz interval centered around 8 Hz.
- G RPW 11: the median relative power of the angular velocity power spectrum in a 3 Hz interval centered around 11 Hz.
- G Dom RPW 11: the relative power of the dominant angular velocity power spectrum in a 3 Hz interval centered around 11 Hz
- A peak frequency: the frequency at which the peak in the acceleration power spectrum is found.
- A peak power: the power at the acceleration power spectrum peak.
- A median bandwidth: the width of the frequency interval that contains 50% of the acceleration power spectrum total power.
- A Dom 68% Peak Bandwidth: the width of the frequency interval, centered around the spectral peak, that contains 68% of the acceleration power spectrum total power.

XSL•FO
RenderX

Statistical Analysis

Overview

The same software was used to run the statistical analyses, with the significance level set at 5%. Given the differences in the indicators in cursive and block letters (BL) handwriting—the choice of the allograph was left free to the participant—the analyses were conducted separately for the 2 allographs. The pipeline was common for each goal: after testing for the indicator normality with the Lilliefors test, the proper statistical test was applied.

Reliability

The indicators test-retest reliability in the MCI group for the List, Text, and PnP tasks, as well as for PnP regular, irregular, and MU words separately, was the first aspect to be evaluated. Before computing the reliability, a paired sample test was conducted, namely the *t* test for normally distributed indicators, the Wilcoxon test otherwise. If the paired test rejected the null hypothesis (ie, the mean/median of the difference between test and retest value is equal to 0), the indicator was deemed not reliable. Otherwise, the ICC (absolute agreement) or the Kendall W was computed if the indicator was normal or nonnormal, respectively. Both parameters range between 0 and 1. The indicator reliability was interpreted as follows: nonreliable if ICC<0.5 or W<0.2; fairly reliable if 0.2≤W<0.4; moderately reliable if 0.5≤ICC< 0.75 or 0.4≤W<0.6; good reliability if 0.75≤ICC<0.9 or 0.6≤W<0.8; optimal reliability if ICC≥0.9 or W≥0.8. Notably, the test-retest reliability analysis did not involve the indicators measuring variability (namely, all the indicators measuring variance or coefficient of variation [CV]).

The successive statistical analyses were solely based on the indicators extracted from the first repetition of the task, to avoid including eventual learning phenomenon in the data.

Relationship With Clinical Scores

For the second aim (ie, support to the clinical evaluation), the correlation between the indicators and the clinical parameters (MMSE, CDT, B-ADL, I-ADL, TMT-A, and number of errors in PnP test, either globally and divided by word type) was computed to measure to which extent the former can be used to approximate the results of cognitive/functional tests. Either the Pearson (*r*, normal) or Spearman (ρ, nonnormal) correlation coefficients were considered with the following interpretation: weak correlation, $0.2 \leq |r, \rho| < 0.4$; moderate correlation, $0.4 \leq |r, \rho| < 0.6$; strong correlation, $0.6 \leq |r, \rho| < 0.8$; very strong correlation, $|r, \rho| \geq 0.8$. In this case, the PnP task was considered as a whole (ie, no differentiation among word types).

MCI Dysgraphia Characterization

To provide quantitative information for the detection of lexical and phonetical dysgraphia, handwriting indicators were statistically compared among the different types of words in the PnP task. After averaging the indicators across words of the same type, the differences among the 3 types were studied using statistical methods for paired samples: repeated measures analysis of variance for normal data and Friedman test for the nonnormal case. Following such tests, in case the null hypothesis was rejected, a post hoc comparison was conducted to determine which word groups exhibited significant differences using the Bonferroni method.

Classification of Unconstrained Handwriting Tasks

Last, binary classification models for the distinction between patients diagnosed with MCI and HC participants were developed in Python (version 3.10; Python Software Foundation). The models were based only on the unconstrained writing tasks presented in this study (List and Text) to collect evidence on the suitability of ecological handwriting for MCI detection. The data for HC participants were collected during other acquisition campaigns, using the same device. These participants signed an informed consent for their voluntary participation. It is worth noting that only participants with age ≥65 years and MMSE ≥27 were included in the HC group, to avoid considering participants with possibly relevant cognitive deterioration [36]. The age threshold of 65 years was based on the age of the youngest participant in the MCI group. Importantly, HCs were excluded if diagnosed with any neurological, musculoskeletal, or cardiovascular conditions that could impair handwriting.

The starting datasets were as follows: (1) List in cursive allograph, L (C); (2) Text in cursive allograph, T (C); (3) List and Text in cursive allograph, L+T (C); (4) List in BL allograph, L (BL); (5) Text in BL allograph, T (BL); (6) List in all allographs (L); and (7) Text in all allographs (T). Independent of the considered dataset, the model's input consisted of handwriting indicators only, excluding the ones showing a correlation coefficient greater than 0.9 in absolute value with at least another indicator [37]. A 10-fold stratified repeated cross-validation approach was adopted in the training phase, setting the F_1-score as the metric to be optimized, given that, in general, the datasets were not balanced. At each iteration, the seed was changed and the whole dataset was divided into train, validation, and test sets, adopting 0.70, 0.15, and 0.15 ratios, respectively. The validation set was exploited to optimize the algorithm hyperparameters through a randomized search approach running 30 iterations. The test set was used to evaluate the model's performance on data not seen by the algorithm in the previous phases. The global performances were then assessed by averaging the classification metrics (accuracy, recall, precision, F_1-score, and area under the precision recall curve) over the 10 folds. Within this framework, several strategies were adopted. Given the general imbalance of the target classes, both undersampling and oversampling techniques were implemented in the training and validation sets through the *imbalanced-learn* library [38]. The former included random undersampling, the NearMiss method, and the Estimated Nearest Neighbors method. As for the latter, Synthetic Minority Over-sampling Technique (SMOTE), SMOTE-support vector machine (SVMSMOTE), Borderline SMOTE, and ADASYN were considered. Last, the selectKBest feature selection method [39] was used to reduce the dataset dimensionality, with the number of features K to be tried ranging from 10 to 50 with discrete steps of 10. Support vector classifier, random forest, and a pool of boosting algorithms—namely AdaBoost, gradient boosting classifier (GBC), XGBoost (XGB), LightGBM (LGBM), and CatBoost—were tried. The details on the space

XSL•FO
RenderX

of explored hyperparameters for each model are provided in the Multimedia Appendix 1. After the identification of the best performing models on the test set, the importance of indicators in their predictions was evaluated through Shapley Additive Explanations (SHAP) analysis [40]. The results of the SHAP analysis were used to qualitatively explore the patterns of misclassified samples: for each model, the trend of the 5 most relevant indicators for the predictions in misclassified samples was compared with the median trend of the same indicators in the MCI and HC groups through polar plots. To do so, the indicators were normalized between 0 and 1 for visualization purposes.

Results

Participants and Dataset

The characteristics of the recruited participants are reported in Table 1, separately for the whole sample of patients with MCI, which was exploited for the classification analysis, and for the sample that performed test and retest acquisitions, considered for the statistical analyses. Among the 57 patients with MCI enrolled, 40 were amnestic and 17 were nonamnestic. In more detail, the sample included: 3 pure amnestic MCIs (with isolated memory disturbance); 37 multiple-domain amnestic MCIs (with memory disturbance associated with disturbance of other cognitive faculties); 8 multiple-domain nonamnestic MCIs (with disturbance of other cognitive faculties, eg, attention, executive, spatial, and language, without memory disturbance); 9 single-domain nonamnestic MCIs (with disturbance of a single cognitive faculty, mostly dysexecutive).

As for the HC sample, information is displayed separately for the List and Text tasks. For this group, only the MMSE score was available. When comparing the MMSE score of participants with MCI and HC, the score of the MCI whole group was significantly lower than that of the HC sample who performed the List (P=5.4×10^{-05}) and the HC sample who performed the Text (P=6.4×10^{-07}).

The time required for completing the PnP test, given in seconds as median (IQR), was 126.42 (102.56-155.27) seconds at test and 122.56 (95.16-141.47) seconds at retest.

The allographs adopted by patients who underwent the test-retest protocol are reported hereafter:

- List, 29 cursive, 12 BL, 4 mixed (ie, BL in one repetition and cursive in the other).
- Text, 33 cursive, 11 BL, 1 mixed.
- PnP, 31 cursive, 11 BL, 3 mixed.

The allograph distribution of the data used in the classification problem is shown in Table 2. Here, the information refers to the available handwriting data, since both the test and retest acquisitions of the MCI group were considered, while HC participants performed the tasks multiple times.

XSL•FO
RenderX

Table 1. Sample characteristics.

Characteristics	MCI[a] (whole; N=57)	MCI (test-retest; n=45)	HC[b] (list; n=34)	HC (text; n=45)
Sex				
Female	32	30	18	23
Male	16	15	16	22
NA[c]	9	—[d]	—	—
Age (years), mean (SD)	78.92 (5.59)	78.22 (5.15)	73.35 (7.79)	73.44 (7.50)
Education, median (IQR)	10 (8-13)	12 (8-13)	8 (5-13.5)	8 (5-15)
MMSE[e], median (IQR)	27 (24-28)	27 (24-28)	29 (28-30)	29 (28-30)
CDR[f], median (IQR)	0.5 (0.5-0.5)	0.5 (0.5-0.5)	—	—
CDT[g], median (IQR)	3 (3-5)	3 (3-5)	—	—
B-ADL[h], median (IQR)	6 (5-6)	6 (5-6)	—	—
I-ADL[i], median (IQR)	6 (4-8)	6 (4-8)	—	—
TMT-A[j], median (IQR)	64 (48-89)	67 (51-90)	—	—
PnP[k] errors, median (IQR)	4 (2-5)	4 (2-5)	—	—
PnP errors R[l], median (IQR)	1 (0-1)	1 (0-1)	—	—
PnP errors I[m], median (IQR)	1 (0-1)	1 (0-1)	—	—
PnP errors MU[n], median (IQR)	2 (1-3)	2 (1-3)	—	—

[a]MCI: mild cognitive impairment.
[b]HC: healthy control.
[c]NA: not available.
[d]Not applicable.
[e]MMSE: Mini-Mental State Examination.
[f]CDR: clinical dementia rating.
[g]CDT: clock drawing test.
[h]B-ADL: basic activity of daily living.
[i]I-ADL: instrumental activity of daily living.
[j]TMT-A: trail making test–A.
[k]PnP: parole-non-parole test.
[l]R: regular words.
[m]I: irregular words.
[n]MU: made-up words.

XSL•FO
RenderX

Table 2. Allograph distribution in the datasets used for the binary classification problem.

Task	Cursive	Block letters	NA[a]
MCI[b]			
List	62	28	20
Text	68	21	3
List and text	61	21	0
HC[c]			
List	26	21	5
Text	106	34	5
List and text	17	6	0

[a]NA: not available.

[b]MCI: mild cognitive impairment.

[c]HC: healthy control.

Reliability

The analysis regarded the data collected from participants with MCI, denoted as "MCI Test-Retest" in Table 1. Participants who adopted different allographs across the test-retest protocol were excluded. For the sake of brevity, the results are reported for a subset of temporal, force, fluency, and tilt indicators, since these domains are the most commonly investigated in the literature.

Table 3 displays the results obtained on the 2 unconstrained handwriting tasks, separately for cursive ("C" columns) and block letters ("BL" columns) allograph.

In the List executed in cursive, optimal, good, and moderate reliability was revealed for 6 of 15, 8 of 15 (mean pause not reaching significance), and 1 of 15 indicators, respectively. The trend was confirmed in the Text task, with 7, 5, and 2 indicators showing optimal, good, and moderate reliability. Only a single unreliable indicator emerged, namely "Mean Force." When dealing with BL, the results were comparable for the Text, with only 2 indicators being not reliable. Among the reliable ones, statistical significance was not reached in 4 of 13 cases (mean in air, mean pause, LDLJ A, mean tilt). The situation was instead slightly worse for the List task: reliability was mostly of moderate intensity (10/15 indicators), coupled with 8 cases with a *P* value greater than the critical threshold of .05.

The indicators' reliability related to the PnP task performed in cursive (Table 4) was optimal, as the analysis yielded 14 out of 15 significant results of at least good reliability. Significance was not revealed only for the mean in air. The breakdown of the 3 types of words (Table 4) revealed that the positive results obtained on the whole task were due to the contribution of regular (Force NC being the only nonreliable indicator) and irregular words (besides cons peak diff G, all the indicators were significantly reliable). MU words were indeed characterized by 5 out of 15 unreliable indicators and 2 nonsignificant results among the reliable ones. Furthermore, in this case, the BL allograph was associated with worse outcomes in the whole PnP (Table 5: 5 unreliable indicators and 6 at least moderately reliable but not significant ones). Here, the poorest results happened with irregular words (6 reliable indicators reaching significance, 6 not reliable), while MU words revealed better reliability values than the corresponding in cursive, although only 4 showed a *P* value lower than .05.

Figure 1 displays a heatmap of the overall reliability intensity across indicators—the same reported in Tables 3-5—and tasks.

XSL•FO
RenderX

Table 3. Reliability results for List ("L" columns) and Text ("T" columns). Refer to Textbox 1 for descriptions of the indicators.

Domain and indicator	L (C)[a]			T (C)			L (BL)[b]			T (BL)		
	N[c]	Rel[d]	*P* value	N	Rel	*P* value	N	Rel	*P* value	N	Rel	*P* value
Temporal domain												
Rel stroke num (number/s)	0	0.83[e]	1.8×10^{-5}	0	0.89[e]	8.2×10^{-9}	0	0.67[f]	.046	0	0.89[e]	.001
Mean on sheet (s)	0	0.65[f]	.005	1	0.83[g]	.01	0	0.69[f]	.03	0	0.00[h]	.74
On sheet ratio	0	0.76[e]	3.0×10^{-4}	0	0.90[g]	8.5×10^{-10}	0	0.68[f]	.04	0	0.95[g]	1.7×10^{-5}
Mean in air (s)	1	0.83[g]	.02	0	0.81[e]	7.5×10^{-6}	1	0.63[f]	.24	1	0.84[g]	.08
Mean pause (s)	1	0.70[e]	.09	0	0.55[f]	.02	1	0.55[f]	.36	1	0.86[g]	.07
Air sheet ratio	1	0.78[e]	.03	0	0.83[e]	2.3×10^{-6}	0	0.55[f]	.12	1	0.92[g]	.05
Fluency domain												
LDLJ A	0	0.89[e]	8.3×10^{-8}	1	0.91[g]	.003	0	0.55[f]	.11	1	0.77[e]	.12
LDLJ G	1	0.90[g]	.007	0	0.89[e]	6.4×10^{-9}	1	0.89[g]	.05	0	0.77[e]	.01
SPARC	0	0.88[e]	2.9×10^{-7}	0	0.93[g]	1.3×10^{-11}	0	0.67[f]	.04	1	0.92[g]	.05
Cons peak diff G (deg/s)	1	0.86[e]	.01	1	0.90[g]	.004	0	0.61[f]	.08	0	0.75[e]	.02
Force domain												
Mean force (arbitrary)	0	0.93[g]	2.4×10^{-9}	0	NA[h,i]	NA	0	0.94[g]	2.2×10^{-5}	0	0.91[g]	.001
Force OVS (arbitrary)	1	0.94[g]	.005	0	0.97[g]	1.1×10^{-16}	0	0.94[g]	3.0×10^{-5}	0	0.89[e]	.001
Force NC (number/s)	0	0.8[e]	4.4×10^{-6}	0	0.89[e]	4.4×10^{-9}	0	0.82[e]	.003	0	0.90[g]	3.3×10^{-4}
Cons peak diff F (arbitrary)	1	0.87[g]	.01	0	0.97[g]	4.4×10^{-16}	0	0.65[f]	.06	0	0.41[h]	.20
Tilt domain												
Mean tilt (deg)	1	0.81[g]	.02	0	0.69[f]	8.7×10^{-4}	1	0.83[g]	.08	1	0.90[g]	.05

[a]C: cursive.

[b]BL: block letters.

[c]N: normality (0 normal, 1 nonnormal).

[d]Rel: reliability.

[e]Good reliability.

[f]Moderate reliability.

[g]Excellent reliability.

[h]Not reliable.

[i]Indicator failed the paired samples test.

XSL•FO
RenderX

Table 4. Reliability results for parole-non-parole test (PnP) test and regular (R), irregular (I), and made-up (MU) words in cursive. Refer to Textbox 1 for descriptions of the indicators.

Domain and indicator	PnP (C)[a]			R			I			MU		
	N[b]	Rel[c]	*P* value	N	Rel	*P* value	N	Rel	*P* value	N	Rel	*P* value
Temporal domain												
Rel stroke num (number/s)	1	0.93[d]	.003	1	0.94[d]	.003	0	0.69[e]	.001	1	0.87[d]	.008
Mean on sheet (s)	1	0.84[d]	.01	1	0.97[d]	.002	1	0.93[d]	.003	1	NA[f,g]	NA
On sheet ratio	0	0.84[h]	2.5×10^{-5}	0	0.89[d]	2.8×10^{-8}	1	0.74[h]	.049	1	0.59[e]	.24
Mean in air (s)	1	0.72[h]	.06	0	0.86[h]	2.8×10^{-7}	1	0.75[h]	.04	1	0.64[h]	.14
Mean pause (s)	1	0.81[d]	.02	1	0.82[d]	.02	1	0.77[h]	.03	1	0.75[d]	.04
Air sheet ratio	1	0.80[d]	.02	0	0.86[h]	3.2×10^{-7}	0	0.65[e]	.004	0	0.28[g]	.20
Fluency domain												
LDLJ A	1	0.90[d]	.005	0	0.83[h]	3.6×10^{-6}	0	0.87[h]	2.2×10^{-7}	1	NA[g]	NA
LDLJ G	1	0.93[d]	.003	1	0.86[d]	.01	0	0.90[d]	6.3×10^{-9}	1	NA[g]	NA
SPARC	1	0.91[d]	.004	1	0.95[d]	.002	1	0.83[d]	.014	1	0.92[d]	.004
Cons peak diff G (deg/s)	1	0.91[d]	.004	1	0.92[d]	.004	1	0.70[h]	.077	1	NA[g]	NA
Force domain												
Mean force (arbitrary)	0	0.92[d]	8.1×10^{-10}	0	0.93[d]	1×10^{-10}	0	0.91[d]	3.6×10^{-9}	0	0.93[d]	2.9×10^{-10}
Force OVS (arbitrary)	0	0.89[h]	2.8×10^{-8}	0	0.94[d]	4.2×10^{-11}	1	0.92[d]	.004	1	0.91[d]	.004
Force NC (number/s)	0	0.91[d]	4.9×10^{-9}	0	NA[g]	NA	0	0.78[h]	4.2×10^{-5}	0	0.69[e]	.001
Cons peak diff F (arbitrary)	0	0.94[d]	4.3×10^{-12}	0	0.90[d]	6.4×10^{-9}	0	0.85[h]	8.6×10^{-7}	1	0.89[d]	.006
Tilt domain												
Mean tilt (deg)	0	0.88[h]	4.4×10^{-8}	0	0.90[d]	8.3×10^{-19}	0	0.89[h]	6.2×10^{-8}	0	0.89[h]	4.8×10^{-8}

[a]C: cursive.

[b]N: normality (0 normal, 1 nonnormal).

[c]Rel: reliability.

[d]Excellent reliability.

[e]Moderate reliability.

[f]Indicator failed the paired samples test.

[g]Not reliable.

[h]Good reliability.

XSL•FO
RenderX

Table 5. Reliability results for parole-non-parole test (PnP) test and regular (R), irregular (I), and made-up (MU) words in block letters. Refer to Textbox 1 for descriptions of the indicators.

Domain and indicator	PnP (BL)[a]			R			I			MU		
	N[b]	Rel[c]	*P* value	N	Rel	*P* value	N	Rel	*P* value	N	Rel	*P* value
Temporal domain												
Rel stroke num (number/s)	0	0.77[d]	.01	1	0.72[d]	.15	0	0.79[d]	.007	0	0.57[e]	.10
Mean on sheet (s)	0	0.41[f]	.17	0	0.84[d]	.003	0	0.86[d]	.002	0	0.77[d]	.01
On sheet ratio	0	0.93[g]	3.4×10^{-4}	1	0.83[g]	.09	0	NA[f,h]	NA	1	0.63[d]	.25
Mean in air (s)	1	0.79[d]	.11	0	0.79[d]	.01	1	NA[f]	NA	0	0.66[e]	.06
Mean pause (s)	0	NA[f]	NA	1	0.83[g]	.09	1	NA[f]	NA	1	0.74[d]	.14
Air sheet ratio	0	NA[f]	NA	0	0.79[d]	.009	1	NA[f]	NA	0	0.63[e]	.07
Fluency domain												
LDLJ A	0	0.37[f]	.24	0	0.04[f]	.47	0	0.68[d]	.04	0	0.10[f]	.44
LDLJ G	1	0.77[d]	.12	0	0.66[e]	.06	0	0.67[d]	.03	0	0.76[d]	.02
SPARC	0	0.50[e]	.15	0	0.75[d]	.02	1	0.83[g]	.08	1	0.91[g]	.05
Cons peak diff G (deg/s)	0	0.51[e]	.15	0	0.45[d]	.20	0	0.48[f]	.16	0	0.51[e]	.15
Force domain												
Mean force (arbitrary)	1	0.90[g]	.05	1	0.90[g]	.05	1	0.93[g]	.05	1	0.88[g]	.06
Force OVS (arbitrary)	1	0.89[g]	.06	0	0.88[d]	.001	0	0.82[d]	.004	1	0.87[g]	.07
Force NC (number/s)	0	0.87[d]	.001	0	0.67[e]	.06	1	0.58[e]	.32	0	0.51[e]	.12
Cons peak diff F (arbitrary)	0	0.87[d]	.001	0	0.85	.003	0	0.38[f]	.25	0	0.70[e]	.04
Tilt domain												
Mean tilt (deg)	0	0.41[f]	.17	0	0.55[e]	.08	1	0.45[i]	.52	0	0.65[e]	.04

[a]BL: block letters.

[b]N: normality (0 normal, 1 nonnormal).

[c]Rel: reliability.

[d]Good reliability.

[e]Moderate reliability.

[f]Not reliable.

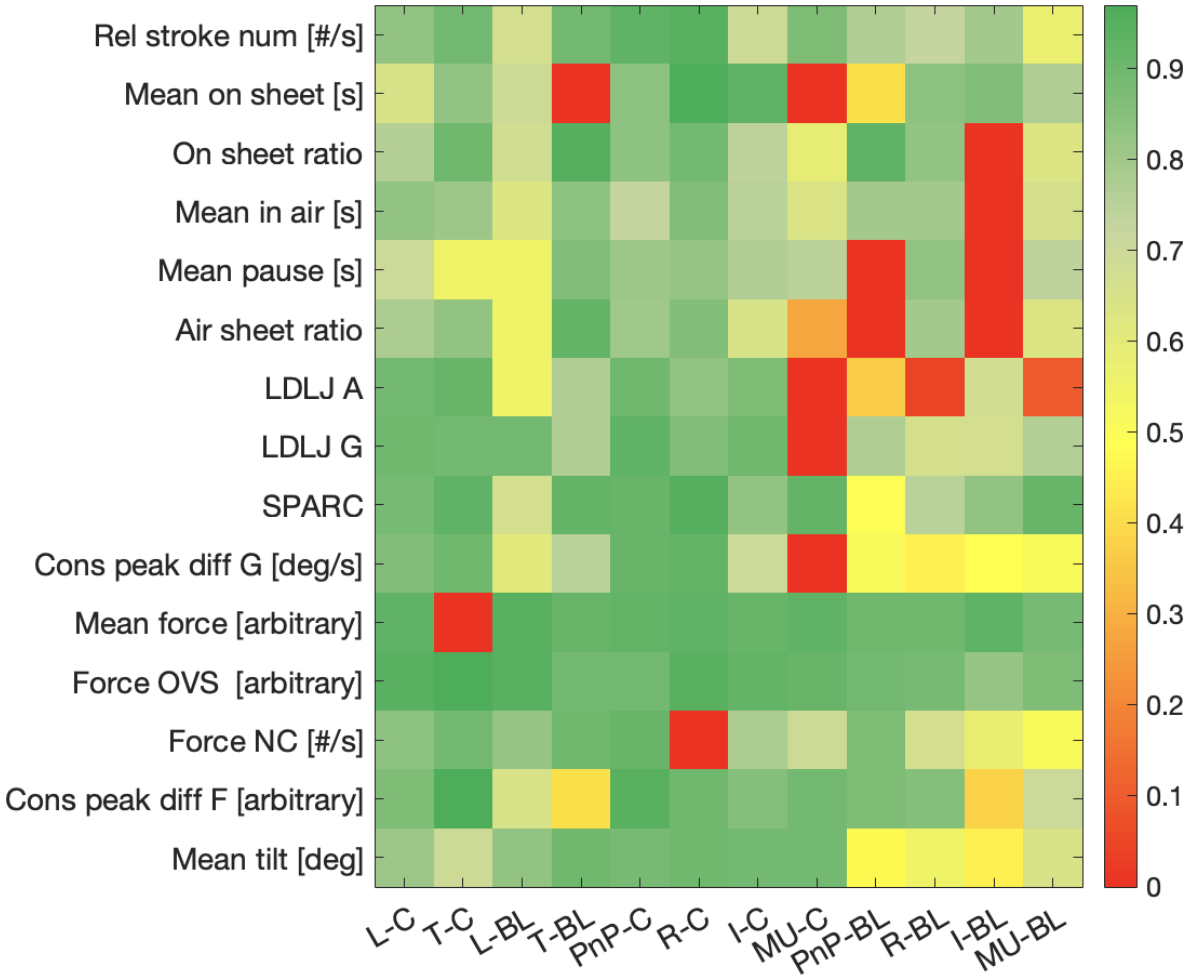
[g]Excellent reliability.

[h]Indicator failed the paired samples test.

[i]Fair reliability.

XSL•FO
RenderX

Figure 1. Reliability heatmap. Each row represents an indicator extracted from the sensorized ink pen with its measurement unit (if absent, the indicator is dimensionless), while each column represents a task. The color bar on the right encodes the reliability intensity. For the sake of visualization, here there is no distinct interpretation between ICC and Kendall W for the color coding: only the reliability value is considered. BL: block letters; C: cursive; I: irregular words in PnP; L: list; LDLJ A: logarithmic dimensionless jerk of the acceleration signal; LDLJ G: logarithmic dimensionless jerk of the angular velocity signal; MU: made-up words in PnP; NC: number of changes; OVS: overshoot; PnP: parole-non-parole; R: regular words in PnP; SPARC: spectral arc length of the angular velocity signal; T: text.



Relationship With Clinical Scores

The statistically significant correlations of handwriting indicators with clinical parameters are reported in Table 6. Overall, significant correlations were identified mainly for the unconstrained tasks in BL allograph. The I-ADL score was also significantly correlated with the indicators extracted from the Text in cursive allograph, while the cursive List exhibited no significant results at all. The PnP test, on the other hand, revealed relationships with the computed indicators mainly when considering the number of errors in the irregular words. The correlation absolute values ranged from moderate to strong, and their direction followed the expected trends.

XSL•FO
RenderX

Table 6. Correlation results. The correlation coefficient value and its corresponding *P* value are reported for all the tasks, separately for cursive (C) and block letters (BL) allographs. The clinical parameter "PnP errors - I" refers to the number of errors with irregular words in the parole-non-parole test (PnP) test. The correlations, in this case, were computed only with the indicators extracted from irregular words. Refer to Textbox 1 for descriptions of the indicators.

Clinical parameter and indicator	L[a] (C)[b]		T[c] (C)		L (BL)[d]		T (BL)		PnP (C)		PnP (BL)	
	corr[e]	*P* value	corr	*P* value	corr	*P* value	Corr	*P* value	corr	*P* value	corr	*P* value
MMSE[f]												
Rel stroke num (number/s)	__[g]	—	—	—	0.52	.04	0.66	.00	—	—	—	—
On sheet ratio	—	—	0.41	.03	0.62	.01	0.57	.02	—	—	0.51	.04
Mean in air (s)	—	—	—	—	–0.56	.02	–0.66	.00	—	—	—	—
Mean pause (s)	—	—	—	—	—	—	–0.63	.01	—	—	–0.51	.05
Pause num (number)	—	—	–0.42	.02	–0.68	.004	—	—	—	—	–0.74	.001
Air sheet ratio	—	—	—	—	–0.58	.02	–0.57	.02	—	—	—	—
CDR[h]												
On sheet CV	—	—	—	—	—	—	–0.80	.047	–0.80	.047	—	—
Mean in air (s)	—	—	0.80	.047	—	—	0.80	.047	—	—	—	—
Air sheet ratio	—	—	—	—	—	—	0.80	.047	—	—	—	—
CDT[i]												
Mean force (arbitrary)	—	—	—	—	0.50	.05	—	—	—	—	0.53	.03
Force OVS (arbitrary)	—	—	—	—	0.59	.02	0.58	.02	—	—	—	—
Force NC (number/s)	—	—	—	—	0.63	.01	—	—	—	—	—	—
Cons peak diff F (arbitrary)	—	—	—	—	0.57	.02	—	—	—	—	0.74	.001
I-ADL[j]												
Rel stroke num (number/s)	—	—	0.42	.02	0.57	.02	0.65	.01	—	—	—	—
Mean on sheet (s)	—	—	–0.42	.02	—	—	—	—	—	—	—	—
On sheet ratio	—	—	—	—	0.66	.01	0.65	.01	—	—	—	—
Mean in air (s)	—	—	–0.49	.01	–0.63	.01	–0.83	.00	—	—	—	—
Mean pause (s)	—	—	–0.41	.03	—	—	–0.52	.04	—	—	—	—
Pause num (number)	—	—	–0.48	.01	–0.57	.02	—	—	—	—	–0.69	.003
Air sheet ratio	—	—	—	—	–0.61	.01	–0.73	.001	—	—	—	—
TMT-A[k]												
Mean in air (s)	—	—	0.47	.01	—	—	0.55	.03	—	—	—	—
Mean force (arbitrary)	—	—	—	—	—	—	–0.51	.04	—	—	—	—
Force NC (number/s)	—	—	–0.39	.04	–0.51	.045	—	—	–0.41	.03	—	—
Cons peak diff F [arbitrary]	—	—	—	—	–0.58	.02	–0.59	.02	—	—	—	—
PnP errors – I												

XSL•FO
RenderX

Clinical parameter and indicator	L[a] (C)[b]		T[c] (C)		L (BL)[d]		T (BL)		PnP (C)		PnP (BL)	
	corr[e]	*P* value	corr	*P* value	corr	*P* value	Corr	*P* value	corr	*P* value	corr	*P* value
Mean on sheet (s)	—	—	—	—	—	—	—	—	—	—	–0.65	.01
On sheet ratio	—	—	—	—	—	—	—	—	—	—	–0.69	.007
Air sheet ratio	—	—	—	—	—	—	—	—	—	—	0.70	.005
G NC (number/s)	—	—	—	—	—	—	—	—	–0.54	.002	–0.61	.02

[a]L: list.

[b]C: cursive.

[c]T: text.

[d]BL: block letters.

[e]corr: correlation.

[f]MMSE: Mini-Mental State Examination.

[g]Not applicable.

[h]CDR: clinical dementia rating.

[i]CDT: clock drawing test.

[j]I-ADL: instrumental activity of daily living.

[k]TMT-A: trail making test-A.

MCI Dysgraphia Characterization

The statistically significant results of the post hoc comparison of indicators (ie, the indicators for which the repeated measures ANOVA or Friedman test rejected the null hypothesis) among regular, irregular, and MU words are reported in Table 7 for both cursive ("C" columns) and block letters ("BL" columns). Most significances regarded irregular words.

XSL•FO
RenderX

Table 7. Results of the parole-non-parole test post-hoc comparison among different word types. "Type 1" and "Type 2" indicate the type of words being compared. The columns "LB," "Diff," and "UB" represent, respectively, the lower bound, the estimate, and the upper bound of the difference in the indicator between the type 1 and type 2 words. Refer to Textbox 1 for descriptions of the indicators.

Indicator	Type 1	Type 2	C[a]					BL[b]				
			P value[c]	LB[d]	Diff	UB[e]	*P* value (PH)	*P* value	LB	Diff	UB	*P* value (PH)[f]
Execution time (s)												
	R[g]	I[h]	1.7×10^{-6}	–1.52	–0.90	–0.28	.001	3.3×10^{-4}	–2.05	–1.14	–0.24	.007
	MU[i]	I	1.7×10^{-6}	–1.92	–1.30	–0.68	1.4×10^{-6}	3.3×10^{-4}	–2.33	–1.43	–0.52	4.7×10^{-4}
Rel stroke num (number/s)												
	R	I	4.4×10^{-5}	0.42	1.03	1.65	1.9×10^{-4}	.025	0.07	0.35	0.63	.015
	MU	I	4.4×10^{-5}	0.35	0.97	1.58	5.4×10^{-4}	.025	0.07	0.36	0.66	.015
On sheet ratio												
	R	I	.88	—[j]	—	—	—	1.6×10^{-4}	0.03	0.10	0.17	.005
	MU	I	.88	—	—	—	—	1.6×10^{-4}	0.03	0.12	0.22	.007
Mean in air (s)												
	R	I	5.9×10^{-5}	–1.28	–0.67	–0.05	.029	3.4×10^{-4}	–2.33	–1.43	–0.52	4.7×10^{-4}
	MU	I	5.9×10^{-5}	–1.75	–1.13	–0.52	3.4×10^{-5}	3.4×10^{-4}	–2.05	–1.14	0.24	.007
In air CV												
	R	MU	.001	0.06	0.22	0.38	.004	.008				
	MU	I	.001	–0.33	–0.17	0.00	.044	.008	0.05	–0.32	0.59	.019
Mean pause (s)												
	R	I	1.5×10^{-5}	–1.39	–0.82	–0.24	.002	.002	–1.84	–1.04	–0.23	.006
	MU	I	1.5×10^{-5}	–1.66	–1.08	–0.51	1.8×10^{-5}	.002	0.20	–1.00	1.80	.009
Pause num (number)												
	MU	I	.002	–0.92	–0.55	–0.18	.001	.018	–1.20	–0.64	–0.01	.016
Tilt CV												
	R	I	.003	0.25	0.87	1.48	.002	6.3×10^{-4}	0.52	1.43	2.33	4.7×10^{-4}
	MU	I	.003	—	—	—	—	6.3×10^{-4}	0.02	0.93	1.83	.042
LDLJ G												
	R	MU	.032	–1.28	–0.67	–0.05	.029	.001	–0.39	–0.22	–0.04	.013
	MU	I	.032	—	—	—	—	.001	0.11	0.34	0.57	.004

[a]C: cursive.

[b]BL: block letters.

[c]*P* value of the statistical test investigating the differences among the 3 types of words (ANOVA or Friedman test).

[d]UB: upper bound.

[e]LB: lower bound.

[f]PH: *P* value of the pairwise post hoc comparison.

[g]R: regular.

[h]I: irregular.

[i]MU: made-up.

[j]Not applicable.

XSL•FO
RenderX

Classification of Unconstrained Handwriting Tasks

The best classification performances for each dataset are displayed in Table 8, together with the adopted training strategy in terms of under- or oversampling and feature selection. Classification metrics on the test set are given as average (SD) over the 10 folds of stratified repeated cross-validation.

Table 8. Binary classification performances on the test set. For details, see the Methods section (Data Analysis subsection).

Dataset	Model	Sampling[a]	FS[b]	Accuracy	Recall	Precision	F_1-score	PRC[c] AUC[d]
L[e] (C)[f]	RF[g]	SMOTE[h]	1	0.84 (0.05)	0.93 (0.07)	0.83 (0.08)	0.87 (0.04)	0.82 (0.07)
T[i] (C)	CB[j]	SMOTE	0	0.86 (0.06)	0.87 (0.08)	0.81 (0.14)	0.83 (0.07)	0.76 (0.12)
L+T (C)	AB[k]	No	0	0.84 (0.09)	0.96 (0.08)	0.85 (0.09)	0.90 (0.07)	0.84 (0.09)
L (BL)[l]	CB	SVMSMOTE[m]	1	0.80 (0.14)	0.90 (0.17)	0.78 (0.18)	0.81 (0.13)	0.77 (0.17)
T (BL)	SVC[n]	RUS[o]	0	0.93 (0.07)	1.00 (0.00)	0.87 (0.14)	0.92 (0.09)	0.87 (0.14)
L	XGB[p]	SMOTE	0	0.84 (0.07)	0.88 (0.07)	0.88 (0.09)	0.87 (0.06)	0.85 (0.09)
T	SVC	RUS	0	0.86 (0.08)	0.85 (0.12)	0.79 (0.12)	0.81 (0.10)	0.73 (0.13)

[a]The adopted sampling strategy is reported.

[b]FS is 1 if feature selection was applied, 0 otherwise.

[c]PRC: precision recall curve.

[d]AUC: area under the curve.

[e]L: list.

[f]C: cursive.

[g]RF: random forest.

[h]SMOTE: Synthetic Minority Oversampling Technique.

[i]T: text.

[j]CB: CatBoost.

[k]AB: AdaBoost.

[l]BL: block letters.

[m]SVMSMOTE: Synthetic Minority Oversampling Technique-Support Vector Machine.

[n]SVC: support vector classifier.

[o]RUS: random undersampling.

[p]XGB: XGBoost.

The top 5 most important indicators for each model, according to the SHAP analysis, are presented in Table 9.

A total of 21 indicators resulted among the possible 35 (5 top relevant indicators for 7 datasets), the most relevant ones. The most prevalent was Tilt CV, being present in 5 out of 7 datasets (one occurrence of Tilt CV_T in the L+T [C] dataset), followed by G peak power (4/7). The indicators A peak frequency and force NC were found in 3/7 cases, while cons peak diff G, A ApEn, and execution time resulted in 2/7 datasets. The remaining indicators occurred only once. Importantly, the indicators occurring multiple times were always characterized by the same direction in the MCI group. A high degree of consistency was revealed for the 3 models trained on cursive samples. Indeed, Tilt CV was the most important indicator in all 3 cases, while cons peak diff G, A ApEn, and G peak power were found in the 2 models trained with List cursive samples (L [C]; L+T [C]). For the 2 models trained on BL samples, the opposite trend occurred, with no superpositions among their 5 most relevant indicators.

Given these outcomes, the analysis of misclassified samples was carried out for cursive-samples models only.

In the L (C) model, G peak power showed trends compatible with those of patients with MCI in 50% (5/10) of false positive (FP) samples. The other recurrent altered indicators in FPs were cons peak diff G and air sheet ratio (4/10, 40%). Tilt CV and G peak power were found to be in line with that of HC participants in 83% (5/6) and 67% (4/6) of the false negatives (FNs), respectively.

In T (C), FPs were mainly associated with problems in temporal indicators (execution time: 11/16, 69%; mean pause: 10/16, 63%) and in a peak frequency (11/16, 69%). This last indicator was the main contributor to the presence of FNs (12/16, 75%), together with Tilt CV and mean pause (10/16, 63%).

When combining the information of the 2 tasks (L+T [C]), Tilt CV_T was found to deviate from the expected values in most misclassified samples: 67% in both FPs (4/6) and FNs (2/3). The other main contributor to FPs was A NC (3/6, 50%).

XSL•FO
RenderX

Table 9. For each indicator, the indicator name, the corresponding absolute Shapley Additive Explanations value, and the direction in patients with MCI are presented for each dataset. The results are related to the best-performing model in each dataset, presented in Refer to Textbox 1 for descriptions of the indicators.

Dataset	#1 Indicator	#2 Indicator	#3 Indicator	#4 Indicator	#5 Indicator
L[a] (C)[b]	Tilt CV 0.093 (+)[c]	Cons peak diff G 0.057 (+)	A ApEn 0.044 (+)	Air sheet ratio 0.019 (+)	G peak power 0.019 (+)
T[d] (C)	Tilt CV 0.129 (+)	Force NC 0.090 (–)	Execution time 0.083 (+)	A_Peak frequency 0.058 (–)	Mean pause 0.042 (+)
L+T (C)	Tilt CV_T[e] 0.090 (+)	A ApEn 0.059 (+)	A NC 0.056 (+)	Cons peak diff G 0.036 (+)	G peak power 0.010 (+)
L (BL)[f]	LDLJ A 0.108 (–)[g]	A_Peak frequency 0.105 (–)	G peak power 0.066 (+)	Force overshoot 0.041 (+)	On sheet CV 0.022 (+)
T (BL)	Force NC 0.042 (–)	G RPW 11 0.040 (+)	A Dom 68% peak bandwidth 0.039 (+)	G RPW 8 0.035 (+)	Tilt median bandwidth 0.032 (–)
L	Tilt CV 0.077 (+)	Execution time 0.073 (+)	A peak frequency 0.041 (–)	G peak power 0.027 (+)	G Dom RPW 11 0.024 (+)
T	Force NC 0.063 (–)	A peak power 0.043 (+)	A median bandwidth 0.043 (+)	G NC 0.039 (+)	Tilt CV 0.030 (+)

[a]L: list.

[b]C: cursive.

[c]+ denotes increased indicator values.

[d]T: text.

[e]Tilt CV_T in the L+T (C) dataset refers to the Tilt CV indicator in the text task. The other indicators in the row refer to the list task.

[f]BL: block letters.

[g]– denotes decreased indicator values.

Interestingly, 3 HCs and 2 MCIs were misclassified across all 3 cursive datasets. As for the HC group:

- PEN024 was consistently misclassified because of trends in A ApEn and temporal indicators (execution time, mean pause, and air sheet ratio) compatible with those of patients with MCI.
- PEN025 instead showed alterations in the kinematics of the handwriting movement, with high values of Tilt CV in the T (C) task and great angular velocity oscillations power in both tasks (G peak power).
- PENT002 exhibited a heterogeneous pattern, with difficulties in the force and pen inclination domains in the T (C) task. Altered values of cons peak diff G were among the reasons for the wrong prediction both in L (C) and L+T (C).

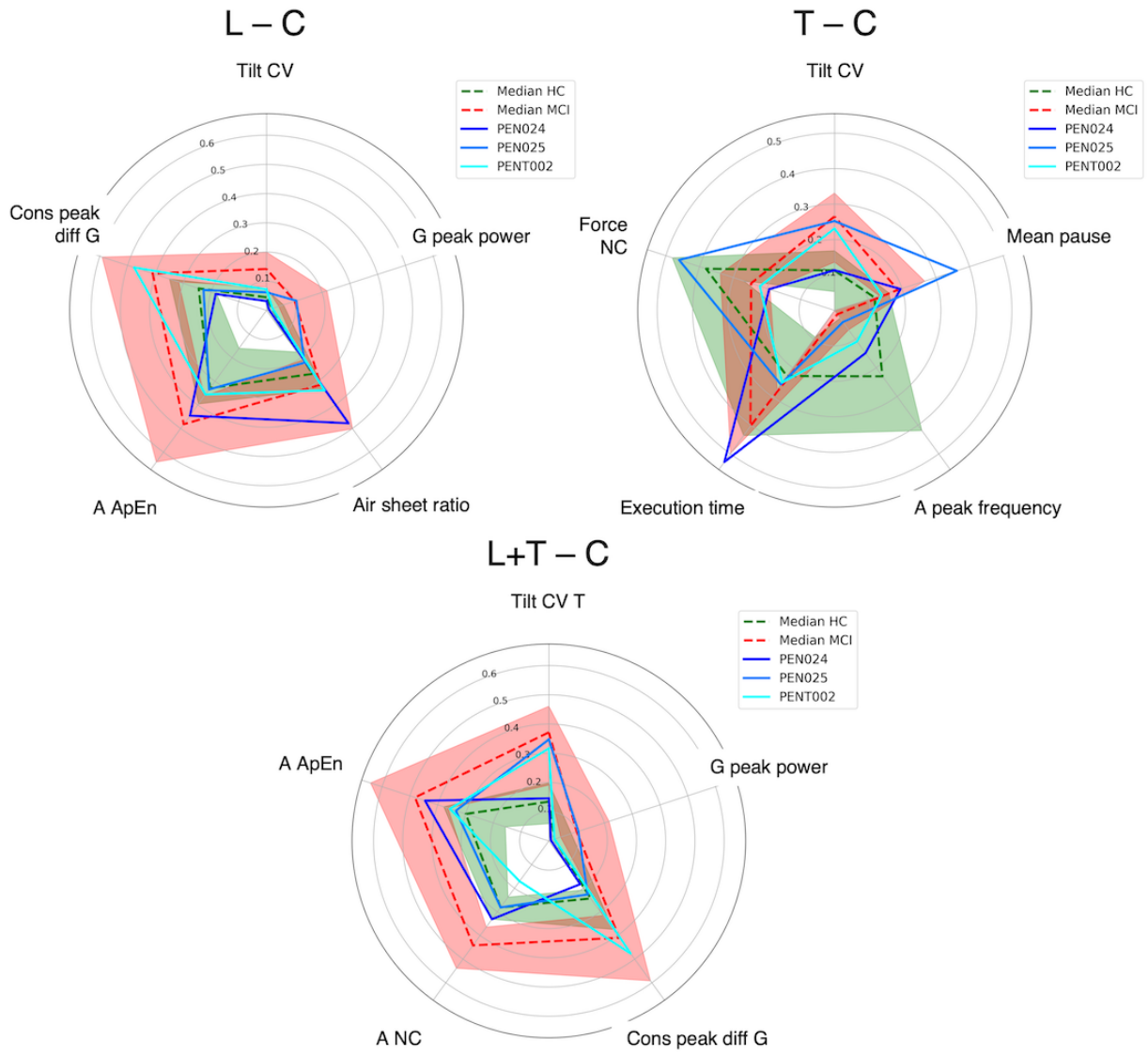
Moving to the patients with MCI:

- DGSMN07 revealed a high coherence across tasks. The modulation of rotational kinematics (cons peak diff G and G peak power) was similar to that of HCs, in turn causing the same behavior in Tilt CV.
- DGMSN17 was misclassified in all datasets due to the trend of Tilt CV. The effect of temporal indicators was identified in the T (C) task.

The indicators' trend for these samples is presented in the following radar plots for each classification model (in order, L [C], T [C], and L+T [C]). For each model, the most relevant indicator is displayed at the top of the radar. Then, indicators are displayed counterclockwise in decreasing order of relevance. All the values are normalized between 0 and 1 for visualization purposes. The dashed green line represents the indicators' median value, while the green area represents the interval between the 25th and the 75th percentile in the HC group. The red elements represent the same for the group of patients with MCI. The solid lines represent the indicator's value for the sample under examination. Blue hues are used for HC samples in Figure 2, while orange hues are used for MCI samples in Figure 3.

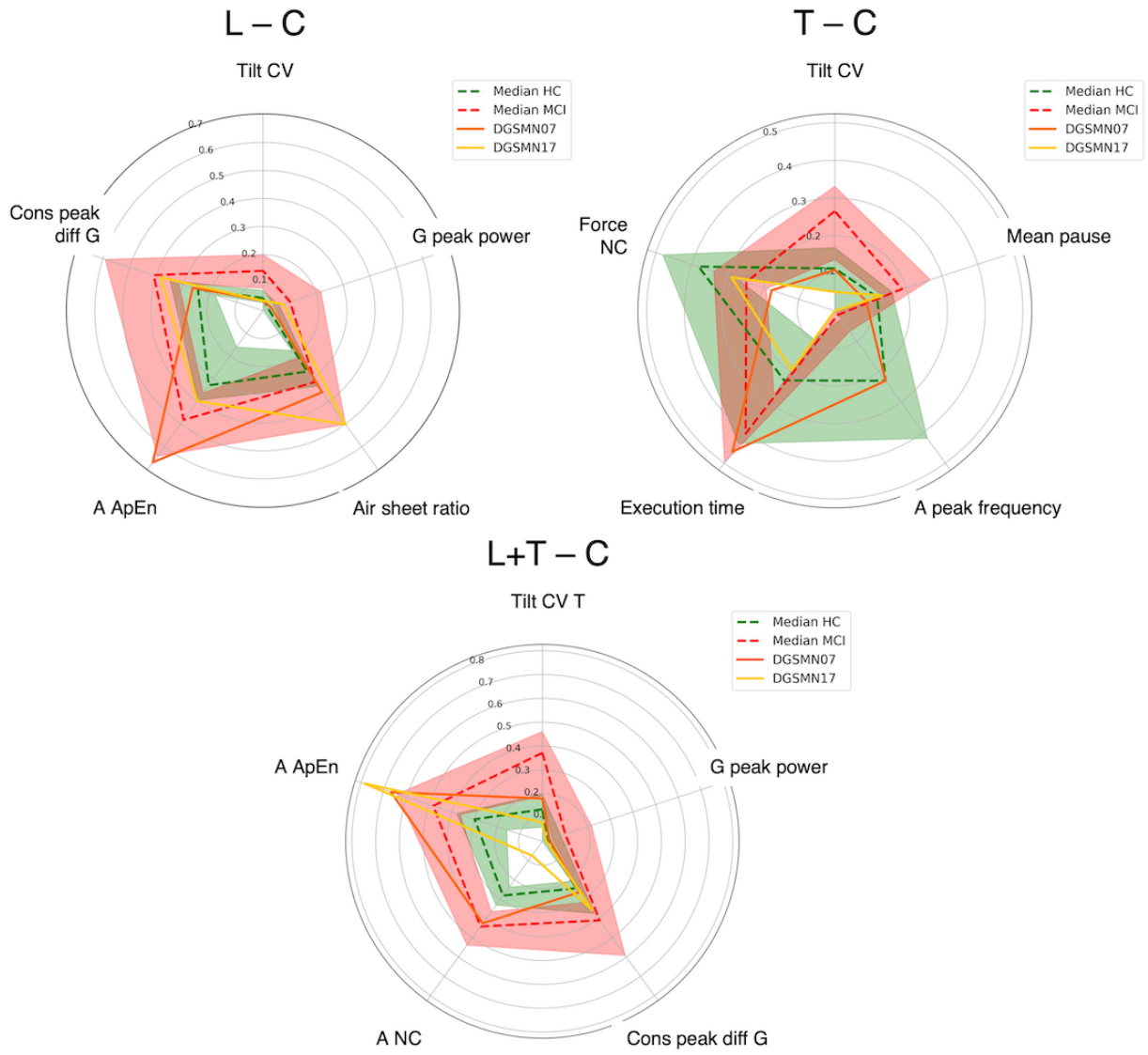
XSL•FO
RenderX

Figure 2. Radar plots for HC samples from participants PEN024, PEN025, and PENT002. A ApEn: approximate entropy of the acceleration signal; C: cursive; CV: coefficient of variation; HC: healthy control; L: list; MCI: mild cognitive impairment; T: text; Tilt CV T: tilt CV indicator in the text task.



XSL•FO
RenderX

Figure 3. Radar plots for MCI samples from participants DGSMN07 and DGSMN17. A ApEn: approximate entropy of the acceleration signal; C: cursive; CV: coefficient of variation; HC: healthy control; L: list; MCI: mild cognitive impairment; T: text; Tilt CV T: tilt CV indicator in the text task.



Discussion

Principal Findings

This work presented the quantitative analysis of handwriting in a population of patients diagnosed with MCI. With respect to the existing body of literature on the topic, an ecological approach for data acquisition was proposed. Handwriting tasks were executed on traditional paper using a sensorized ink pen able to record kinematic and dynamic signals. Moreover, content-free handwriting composition tasks were proposed to the participants to mirror the writing that can be carried out in everyday life. These aspects were deemed necessary for the development of a solution that potentially unlocks the monitoring of handwriting in the home environment. However, such features alone are not sufficient for the purpose.

That is why the first goal of the work was the assessment of the test-retest reliability of the information extracted from the acquisition protocol. High reliability is an essential requirement for any measurement instrument, particularly for solutions meant to detect anomalies in unsupervised longitudinal frameworks, in order to avoid high FP rates and alert fatigue. Importantly, the reliability analysis was conducted by separating different writing styles, namely cursive and BL, given that they differ in both temporal and kinematic parameters. The results were excellent for the cursive allograph, considering the 3 tasks, 42 out of the 45 presented indicators resulted significantly reliable, with reliability values mostly rated as good or optimal. This percentage dropped for BL: this was caused mainly by *P* values not reaching statistical significance, rather than by low reliability values. The outcome could be explained by the reduced number of participants who adopted the BL allograph in the recruited sample, which could also be the reason behind the overall better results found for cursive. Interestingly, the best results were achieved with the unconstrained tasks (List and Text), regardless of the adopted allograph. This could imply that one's peculiar handwriting process characteristics are preserved independently of the content itself, thus making the analysis of unsupervised

XSL•FO
RenderX

handwriting activity a promising approach for the remote tracking of MCI progression. On the other hand, the reliability of the indicators from the standardized PnP test in BL was not always verified. The unreliability of the mean pause indicator, measuring the average duration of moments longer than 2 seconds spent with the pen in the air, could, in theory, be due to differences in the dictation timing of the operator between test and retest acquisitions.

Regarding the support for the clinical evaluation, the correlation analysis revealed a scarce presence of significant results (around 10% of the total correlation values). Such an outcome should not be interpreted negatively for 2 reasons. First, the administered clinical tests are not meant to assess the handwriting performance specifically. Even the CDT and the TMT-A, which require the patients to draw and trace lines for their completion, are not evaluated according to the graphomotor performance itself, but rather on their accuracy. The same holds for the PnP test: the scoring is based on the orthographic correctness of the written words, while the proposed indicators are aimed at characterizing the handwriting process. Thus, a direct correspondence between the clinical scores and the handwriting indicators does not exist. Moreover, the CDR and B-ADL scores were characterized by low granularity (the former presented with only 2 distinct values, namely 0.5 and 1), resulting in poor significance. Second, and most importantly, the obtained significant correlations were of at least moderate intensity and in line with the existing literature [17]. Diving into detail, the highest number of correlations was found between temporal handwriting indicators and the MMSE and the I-ADL scores. As the MMSE and I-ADL increased, rel stroke num and on on-sheet ratio increased as well, while the mean in air, the mean pause, pause num, and air sheet ratio decreased. Therefore, in the recruited sample, participants who are more cognitively intact and functional in daily life tend to write faster thanks to a reduced degree of hesitation. Another relevant finding concerns the 2 tests involving the use of a pen or pencil. Those who performed well in the CDT and TMT-A applied a greater force on the writing surface, coupled with a higher number of changes in the force signal direction (the opposite signs of the correlation values are due to the opposite interpretation of the tests score). Such a pattern is associated in the literature with the healthy status [19]. Last, significant correlations between indicators and the number of errors in the PnP test emerged only for the errors in Irregular words. Here, when the BL allograph was used, a higher number of errors was associated with a faster performance on the sheet: the mean time spent with the pen on paper (mean on sheet) and the percentage of time spent on paper when writing a word (on sheet ratio) both showed a negative correlation. This could suggest that errors in irregular words occurred because the participants paid reduced attention (hence the increased speed on the sheet) to the words to be written, despite requiring orthographic knowledge to be written correctly. Linking the observed correlation patterns with specific domains of cognitive function is not trivial. The clinical tests used in the correlational study are useful in providing general measures of one's cognitive functioning or some of its high-level variables, such as processing speed, or measures of the patient's level of autonomy. In contrast, they are not thought of, nor have they shown particular accuracy and validity in detecting dysfunctions of individual cognitive functions such as memory, attention, executive functions, critical thinking, language, spatial abilities, visual perception, and so on. Any attempt to explain the obtained correlations in terms of individual cognitive domains would therefore be highly speculative and unsupportable with the available data. However, to a first approximation, the correlations between temporal indicators and MMSE and I-ADL scores seem to closely resemble what clinicians call cognitive hesitancy, an aspect of general insecurity in test responses that can be framed in the context of executive function dysfunctionality.

Dealing with the quantitative information derived from the PnP test, another neat trend emerged for irregular words. Independent of the handwriting style and orthographic errors, irregular words posed the greatest challenge to the recruited sample. With respect to the other types of words, irregular ones were consistently and significantly associated with worse in-air temporal performances (ie, longer times spent with the pen in the air), as revealed by mean in air and mean pause. This translated into a significantly slower rate of grapheme production (Rel stroke num) and, consequently, into the longest times required to write the words (execution time). In the BL allograph, alterations were also found in the on-sheet execution, as irregular words exhibited around 10% less time spent with the pen on the writing surface (on-sheet ratio). Last, irregular words were linked to a reduced variability in the pen inclination (Tilt CV), possibly suggesting an impaired manipulation of the writing instrument when dealing with this type of word. According to the classical cognitive models of handwriting, the marked alteration of the handwriting process found in the current work for irregular words, which require semantic and lexical knowledge to be written correctly compared with regular and MU words, could underline alterations in the lexical pathways as a result of dysfunctions in the orthographic retrieval process.

Overall, the implemented quantitative indicators seem able to capture relevant aspects of the handwriting gesture that are linked to the participant's general conditions. On top of that, the indicators can also identify altered patterns that can provide a deeper understanding of the nature of the patient's handwriting impairments. This stresses once again the suitability of handwriting monitoring in the population with MCI, as impairments in the gesture could be associated with a cognitive or functional decline.

The last evaluated aspect was the ability of the indicators to distinguish unconstrained handwriting samples generated by patients with MCI from those written by older adults without any diagnoses and MMSE ≥ 27. The choice of focusing on the List and Text tasks was motivated by their potential application in the domestic scenario. In such a condition, the information could be gathered in ecological modality without the need for the operator's supervision. The classification performances were good, with F_1-scores selected as the evaluation criterion given the datasets' imbalance, ranging from 0.81 to 0.92. Importantly, the results do not appear to be task or allograph-dependent. These results highlight the power of the proposed analysis: the extracted handwriting process indicators do provide relevant

XSL•FO
RenderX

information on the participants' cognitive status, regardless of the writing content and the allograph adopted to realize it. When comparing with the existing literature, the current results are placed between the accuracies of 74% in [19] and 96.6% in [20]. It is, however, worth pointing out that the latter classification performance, related to the execution of a spontaneous sentence, was obtained in different conditions. A digitizer was used, thus providing information also on the written trace, not available in this study, on a narrow sample size (12 patients with MCI and 17 controls). Moreover, the authors reported the participants' age, which was not balanced between groups, as one of the most discriminative features for the model. Such a variable was not included in the models presented here.

The exploration of the relevance of indicators in the predictions of models provided great insight into the aspects that mostly differentiate the execution of patients with MCI and HC participants. This was true particularly for the cursive allograph, since heterogeneous results were found in the models based on BL samples, likely because of the reduced sample size. In general, the outcome of the SHAP analysis was not the expected one. Despite anticipating temporal indicators to carry the highest discriminative power, reflecting hesitations and greater processing times for patients with MCI, their occurrence among the most relevant was somewhat limited: only 5 out of 35 temporal indicators emerged, never placed in first or second position. The biggest determinant in the classification models was the variability of the pen inclination with respect to the vertical axis (Tilt CV), as it was found among the 5 most relevant indicators in 6 out of 7 models, being in the first place in 4 cases. In the classification models built on the cursive List samples (L [C] and L+T [C]), the relevance of Tilt CV, with high values always steering the classification towards the MCI group, was coherently associated with the presence of other kinematic indicators, both in the fluency and in the frequency domains. Such an outcome could open new perspectives on the evaluation of the handwriting performance of patients with MCI, traditionally focused on execution time, product quality, and orthographic accuracy. Researchers and clinicians should also consider the handwriting process kinematics that underlie the generation of the written trace. The results of the SHAP analysis were the basis for the inspection of misclassification patterns in the cursive-based models. These were homogeneous in FNs, for whom Tilt CV was the main reason why the models predicted them as belonging to the HC group. When this was not the case, the cause was still found in the handwriting kinematics. The situation was variable for FPs. In L (C), three archetypes emerged: (1) the participants who experience alterations in the variability of pen inclination; (2) the participants with excessive amplitude in rotational movements; (3) a heterogeneous group, with difficulties both in kinematics and in-air durations. The same number of patterns showed up in T (C). The first was again characterized by abnormal values of Tilt CV, while the second exhibited the opposite behavior, with no impairments in Tilt CV. The last archetype was the most anticipated one, where increased times were the sole reason for misclassifications. When combining List and Text in cursive, 2 main patterns were highlighted. Once again, the most prevalent had Tilt CV_T as the cause of FPs occurrence, while the other one was the entropy of the acceleration signal. Overall, the analysis revealed the multifaceted nature of the handwriting process in older adults, highlighting the importance of the participant-specific exploration of the model predictions. This would allow one to unveil one's peculiar pattern of handwriting impairments, thus helping the clinician in identifying the motor and cognitive aspects that need a thorough assessment.

It is important to highlight the limitations found in this work. Regarding the recruitment, depression and anxiety disorders, relevant for various aspects of cognitive motor processes, were not formally measured during patients' assessment. However, patients with a history of anxiety disorders or depression were excluded from the sample. So, caution should be placed when trying to generalize the presented findings to the population with MCI. As for the inclusion of control participants, their selection was solely based on their MMSE score. While they did not report any neurological diagnoses, they did not undergo any thorough clinical examination aimed at assessing their cognitive status. Thus, the presence of cognitively impaired participants in the control group cannot be excluded. In the statistical analysis of the available data, the percentage of recruited patients who adopted the BL allograph (around 1/3 of the total) was limited with respect to cursive. This translated into a reduced number of statistical significances in the reliability analysis, while the binary classification models trained on BL samples were characterized by greater SDs of the classification metrics across the 10 folds of the stratified repeated cross-validation, with respect to their cursive counterparts. Despite the imbalance in the adopted allograph, we do believe that the performance difference between BL-trained and cursive-trained classification models does not imply a reduced discriminative power of the former writing style. Although additional BL samples are needed to prove this hypothesis, this allograph should be better suited for the binary classification problem based on the proposed tool and indicators. Indeed, being the allograph characterized by a highly standardized motor pattern of stroke generation, the main source of variability between MCI and HC participants should be found in the indicators measuring the in-air performance, which should reflect longer processing times and hesitations in the former group. When dealing with cursive, the participants' personal handwriting style comes into play, making the classification task more challenging. Just as an example, greater times on-sheet could indicate either slowness of movement or a writing style characterized by longer, interconnected strokes. In general, the choice of the allograph was left free to the participant to guarantee an ecological execution of the task and to align with the PnP test instructions, which do not specify how the participant should write, but the presented findings should be confirmed by increasing the BL sample size. The free choice turned out to be a liability in a limited number of cases, where patients adopted different allographs in the test and retest acquisitions, thus making them unsuitable for the reliability analysis. Given the impossibility of characterizing specific cognitive domains with the battery of clinical tests used in this work, the inclusion of more specific clinical tests for different cognitive domains is planned as a next step to refine the

XSL•FO
RenderX

correlation analysis. In this regard, the positive results obtained with the words and nonwords writing test, which is the only instrument investigating a specific cognitive function in the study, indicate that such a research objective seems promising and feasible. With respect to the digitizers, the used sensorized ink pen cannot directly provide information on the trace made on paper. Future work should focus on proper methods for the extraction of product-related handwriting information to reduce this gap, particularly for the cursive allograph. Other promising directions can be found in multimodal approaches for the classification of patients with MCI. For example, the authors in [21] demonstrated that the fusion of handwriting and EEG parameters can improve the classification performance with respect to considering a single data source alone. Of course, the feasibility and user friendliness of the protocol should be carefully designed in such a case. The results also found the basis for additional research: a longitudinal study could be conducted, possibly in an unsupervised scenario. Last, the clinical community would benefit from comparative investigations on the discriminative efficacy of various handwriting-based screening methods. They should also consider their implementation complexity to provide an all-around informed documentation regarding the utility and feasibility of their integration into clinical practice.

Conclusions

The relevance of handwriting in the population with MCI was confirmed in this work, which described a novel approach for its quantitative characterization. The analyses revealed 2 potential complementary scenarios for its fruitful application. On one side, information related to the participant's clinical conditions can be quickly and noninvasively extracted, both through the PnP test and unconstrained handwriting. On the other hand, the handwriting tool emerges as a viable solution to track old adults' conditions in a transparent and ecological way, fostering MCI remote monitoring in a home setting.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Details on indicators computation and hyperparameter tuning.
[DOCX File , 32 KB - aging_v8i1e73074_app1.docx]

References

1. Anderson ND. State of the science on mild cognitive impairment (MCI). CNS Spectr 2019;24(1):78-87. [doi: 10.1017/S1092852918001347] [Medline: 30651152]
2. Gray SL, Anderson ML, Hubbard RA, LaCroix A, Crane PK, McCormick W, et al. Frailty and incident dementia. J Gerontol A Biol Sci Med Sci 2013;68(9):1083-1090 [FREE Full text] [doi: 10.1093/gerona/glt013] [Medline: 23419778]
3. Prince M. Epidemiology of dementia. Psychiatry 2007;6(12):488-490. [doi: 10.1016/j.mppsy.2007.10.001]
4. Bremer P, Cabrera E, Leino-Kilpi H, Lethin C, Saks K, Sutcliffe C, RightTimePlaceCare Consortium. Informal dementia care: consequences for caregivers' health and health care use in 8 European countries. Health Policy 2015;119(11):1459-1471. [doi: 10.1016/j.healthpol.2015.09.014] [Medline: 26518906]
5. Roberts R, Knopman DS. Classification and epidemiology of MCI. Clin Geriatr Med 2013;29(4):753-772 [FREE Full text] [doi: 10.1016/j.cger.2013.07.003] [Medline: 24094295]
6. Pessoa R, Bomfim A, Ferreira B, Chagas M. Diagnostic criteria and prevalence of mild cognitive impairment in older adults living in the community: a systematic review and meta-analysis. Arc Clin Psychiatr 2019;46(3):72-79. [doi: 10.1590/0101-60830000000197]
7. Chen P, Cai H, Bai W, Su Z, Tang Y, Ungvari GS, et al. Correction to: global prevalence of mild cognitive impairment among older adults living in nursing homes: a meta-analysis and systematic review of epidemiological surveys. Transl Psychiatry 2023;13(1):104 [FREE Full text] [doi: 10.1038/s41398-023-02408-3] [Medline: 36997507]
8. Ritchie K. Mild cognitive impairment: an epidemiological perspective. Dialogues Clin Neurosci 2004;6(4):401-408 [FREE Full text] [doi: 10.31887/DCNS.2004.6.4/kritchie] [Medline: 22034212]
9. Gillis C, Mirzaei F, Potashman M, Ikram MA, Maserejian N. The incidence of mild cognitive impairment: a systematic review and data synthesis. Alzheimers Dement (Amst) 2019;11:248-256 [FREE Full text] [doi: 10.1016/j.dadm.2019.01.004] [Medline: 30911599]
10. Vessio G. Dynamic handwriting analysis for neurodegenerative disease assessment: a literary review. Appl Sci 2019;9(21):4666. [doi: 10.3390/app9214666]
11. O'Caoimh R, Timmons S, Molloy DW. Screening for mild cognitive impairment: comparison of "MCI Specific" screening instruments. J Alzheimers Dis 2016;51(2):619-629 [FREE Full text] [doi: 10.3233/JAD-150881] [Medline: 26890758]
12. Zhuang L, Yang Y, Gao J. Cognitive assessment tools for mild cognitive impairment screening. J Neurol 2021;268(5):1615-1622. [doi: 10.1007/s00415-019-09506-7] [Medline: 31414193]
13. Freitas S, Simões MR, Alves L, Santana I. Montreal cognitive assessment: validation study for mild cognitive impairment and alzheimer disease. Alzheimer Dis Assoc Disord 2013;27(1):37-43. [doi: 10.1097/WAD.0b013e3182420bfe] [Medline: 22193353]

XSL•FO
RenderX

14. Breton A, Casey D, Arnaoutoglou NA. Cognitive tests for the detection of mild cognitive impairment (MCI), the prodromal stage of dementia: meta-analysis of diagnostic accuracy studies. Int J Geriatr Psychiatry 2019;34(2):233-242. [doi: 10.1002/gps.5016] [Medline: 30370616]
15. De Stefano C, Fontanella F, Impedovo D, Pirlo G, Scotto di Freca A. Handwriting analysis to support neurodegenerative diseases diagnosis: a review. Pattern Recognit Lett 2019;121:37-45. [doi: 10.1016/j.patrec.2018.05.013]
16. Kawa J, Bednorz A, Stępień P, Derejczyk J, Bugdol M. Spatial and dynamical handwriting analysis in mild cognitive impairment. Comput Biol Med 2017;82:21-28. [doi: 10.1016/j.compbiomed.2017.01.004] [Medline: 28126631]
17. Werner P, Rosenblum S, Bar-On G, Heinik J, Korczyn A. Handwriting process variables discriminating mild alzheimer's disease and mild cognitive impairment. J Gerontol B Psychol Sci Soc Sci 2006;61(4):P228-P236. [doi: 10.1093/geronb/61.4.p228] [Medline: 16855035]
18. Qi H, Zhang R, Wei Z, Zhang C, Wang L, Lang Q, et al. A study of auxiliary screening for alzheimer's disease based on handwriting characteristics. Front Aging Neurosci 2023;15:1117250 [FREE Full text] [doi: 10.3389/fnagi.2023.1117250] [Medline: 37009455]
19. Kahindo C, El-Yacoubi MA, Garcia-Salicetti S, Rigaud A, Cristancho-Lacroix V. Characterizing early-stage alzheimer through spatiotemporal dynamics of handwriting. IEEE Signal Process. Lett 2018;25(8):1136-1140. [doi: 10.1109/lsp.2018.2794500]
20. Garre-Olmo J, Faúndez-Zanuy M, López-de-Ipiña K, Calvó-Perxas L, Turró-Garriga O. Kinematic and pressure features of handwriting and drawing: preliminary results between patients with mild cognitive impairment, alzheimer disease and healthy controls. Curr Alzheimer Res 2017;14(9):960-968 [FREE Full text] [doi: 10.2174/1567205014666170309120708] [Medline: 28290244]
21. Chai J, Wu R, Li A, Xue C, Qiang Y, Zhao J, et al. Classification of mild cognitive impairment based on handwriting dynamics and qEEG. Comput Biol Med 2023;152:106418. [doi: 10.1016/j.compbiomed.2022.106418] [Medline: 36566627]
22. Fischer SH, David D, Crotty BH, Dierks M, Safran C. Acceptance and use of health information technology by community-dwelling elders. Int J Med Inform 2014;83(9):624-635 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.06.005] [Medline: 24996581]
23. Gerth S, Klassert A, Dolk T, Fliesser M, Fischer MH, Nottbusch G, et al. Is handwriting performance affected by the writing surface? comparing preschoolers', second graders', and adults' writing performance on a tablet vs. paper. Front Psychol 2016;7:1308 [FREE Full text] [doi: 10.3389/fpsyg.2016.01308] [Medline: 27672372]
24. Lunardini F, Di Febbo D, Malavolti M, Cid M, Serra M, Piccini L, et al. A smart ink pen for the ecological assessment of age-related changes in writing and tremor features. IEEE Trans Instrum Meas 2021;70:1-13. [doi: 10.1109/tim.2020.3045838]
25. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to alzheimer's disease: recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease. Alzheimers Dement 2011;7(3):270-279. [doi: 10.1016/j.jalz.2011.03.008] [Medline: 21514249]
26. Morris JC. Clinical dementia rating: a reliable and valid diagnostic and staging measure for dementia of the alzheimer type. Int Psychogeriatr 1997;9 Suppl 1:173-6; discussion 177. [doi: 10.1017/s1041610297004870] [Medline: 9447441]
27. Giovagnoli AR, Del Pesce M, Mascheroni S, Simoncelli M, Laiacona M, Capitani E. Trail making test: normative values from 287 normal adult controls. Ital J Neurol Sci 1996;17(4):305-309. [doi: 10.1007/BF01997792] [Medline: 8915764]
28. Jefferson AL, Byerly LK, Vanderhill S, Lambe S, Wong S, Ozonoff A, et al. Characterization of activities of daily living in individuals with mild cognitive impairment. Am J Geriatr Psychiatry 2008;16(5):375-383 [FREE Full text] [doi: 10.1097/JGP.0b013e318162f197] [Medline: 18332397]
29. Gold DA. An examination of instrumental activities of daily living assessment in older adults and mild cognitive impairment. J Clin Exp Neuropsychol 2012;34(1):11-34. [doi: 10.1080/13803395.2011.614598] [Medline: 22053873]
30. Trautwein S, Maurus P, Barisch-Fritz B, Hadzic A, Woll A. Recommended motor assessments based on psychometric properties in individuals with dementia: a systematic review. Eur Rev Aging Phys Act 2019;16:20 [FREE Full text] [doi: 10.1186/s11556-019-0228-z] [Medline: 31700552]
31. Capasso R, Miceli G. Esame Neuropsicologico per l'Afasia: E.N.P.A. Berlin and Heidelberg, Germany: Springer Science & Business Media; 2001.
32. Toffoli S, Lunardini F, Parati M, Gallotta M, De Maria B, Longoni L, et al. Spiral drawing analysis with a smart ink pen to identify Parkinson's disease fine motor deficits. Front Neurol 2023;14:1093690 [FREE Full text] [doi: 10.3389/fneur.2023.1093690] [Medline: 36846115]
33. Lomurno E, Toffoli S, Febbo DD, Matteucci M, Lunardini F, Ferrante S. Age group discrimination via free handwriting indicators. IEEE J Biomed Health Inform 2025;29(1):56-67. [doi: 10.1109/JBHI.2024.3444497] [Medline: 39146174]
34. Toffoli S, Lunardini F, de Isla CG, Ferrante S. AI-based ecological monitoring of handwriting to early detect cognitive decline. : IEEE; 2023 Presented at: IEEE EMBS International Conference on Biomedical and Health Informatics (BHI); 2023 October 15-18; Pittsburgh, Pennsylvania, USA p. 1-4. [doi: 10.1109/bhi58575.2023.10313370]
35. Toffoli S, Lunardini F, Parati M, Gallotta M, Muletti M. Classification of patients with Parkinson's disease using free handwriting features collected through a smart ink pen. Universidad de Las Palmas de Gran Canaria. 2021. URL: https://accedacris.ulpgc.es/handle/10553/116097 [accessed 2025-08-12]

XSL•FO
RenderX

36. Salis F, Costaggiu D, Mandas A. Mini-mental state examination: optimal cut-off levels for mild and severe cognitive impairment. Geriatrics (Basel) 2023;8(1):12. [doi: 10.3390/geriatrics8010012] [Medline: 36648917]
37. Ulivi M, Meroni V, Orlandini L, Prandoni L, Rossi N, Peretti GM, et al. Opportunities to improve feasibility, effectiveness and costs associated with a total joint replacements high-volume hospital registry. Comput Biol Med 2020;121:103775. [doi: 10.1016/j.compbiomed.2020.103775] [Medline: 32568670]
38. Imbalanced-Learn Documentation. 2024. URL: https://imbalanced-learn.org/stable/ [accessed 2025-08-12]
39. Machine learning in python. scikit-learn. 2024. URL: https://scikit-learn.org/stable/ [accessed 2025-08-12]
40. Lundberg S, Lee S. A unified approach to interpreting model predictions. In: NIPS'17: Proceedings of Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017 Presented at: 31st International Conference on Neural Information Processing Systems; December 4 - 9, 2017; Long Beach, California, USA p. 4766-4775 URL: https://arxiv.org/pdf/1705.07874v1 [doi: 10.48550/arXiv.1705.0787]

Abbreviations

B-ADL: basic activity of daily living
BL: block letters
C: cursive
CDR: clinical dementia rating
CDT: clock drawing test
CV: coefficient of variation
FN: false negative
FP: false positive
HC: healthy control
I-ADL: instrumental activity of daily living
ICC: intraclass correlation coefficient
L: list
MCI: mild cognitive impairment
MMSE: Mini-Mental State Examination
MU: made-up
NC: number of changes
PnP: parole-non-parole test
RF: random forest
SMOTE: Synthetic Minority Oversampling Technique
SHAP: Shapley Additive Explanations
SVC: support vector classifier
SVMSMOTE: Synthetic Minority Oversampling Technique-support vector machine
T: text
TMT-A: trail making test-A
XGB: XGBoost

Edited by M Gray; submitted 26.02.25; peer-reviewed by C Udensi, EV Cores; comments to author 19.05.25; revised version received 30.06.25; accepted 10.07.25; published 23.09.25.

Please cite as:
Toffoli S, Abbate C, Lunardini F, Corno E, Diani N, Gallucci A, Tomasini E, Trimarchi PD, Ferrante S
Handwriting in Mild Cognitive Impairment: Reliability Assessment and Machine Learning–Based Screening
JMIR Aging 2025;8:e73074
URL: https://aging.jmir.org/2025/1/e73074
doi:10.2196/73074
PMID:

©Simone Toffoli, Carlo Abbate, Francesca Lunardini, Edoardo Corno, Nicholas Diani, Alessia Gallucci, Emanuele Tomasini, Pietro Davide Trimarchi, Simona Ferrante. Originally published in JMIR Aging (https://aging.jmir.org), 23.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Ethical and Quality of Care–Related Challenges of Digital Health Twins in Care Settings for Older Adults: Scoping Review

Md Shafiqur Rahman Jabin[1,2], MSc, PhD; Ayesha Mirza[3], BE, PGCE; Adaobi Ilodibe[4], BSc, MSc; Tillal Eldabi[4], MSc, PhD; Emilia Vann Yaroson[5], MSc, PhD

[1]Department of Medicine and Optometry, Linnaeus University, Universitetsplatsen 1, Kalmar, Sweden

[2]Faculty of Health Studies, University of Bradford, Bradford, United Kingdom

[3]Faculty of Engineering and Digital Technologies, University of Bradford, Bradford, United Kingdom

[4]School of Management, University of Bradford, Bradford, United Kingdom

[5]Management School, Sheffield University, Sheffield, United Kingdom

Corresponding Author:
Md Shafiqur Rahman Jabin, MSc, PhD
Department of Medicine and Optometry, Linnaeus University, Universitetsplatsen 1, Kalmar, Sweden

Abstract

Background: Digital health twins (DHTs) have been evolving with their diverse applications in medicine, particularly in care settings for older adults, in response to the increasing demands of older adults. Despite its numerous benefits, the optimal implementation of DHTs has faced several challenges, particularly in terms of ethics and quality of care. Given the continuous rise in the need for such care and the evident potential for DHTs to meet these needs, this review seeks to identify and address the gaps in research knowledge to enhance DHT implementation.

Objective: The review aims to compile and synthesize the best available evidence regarding the issues associated with quality of care, the ethical implications of DHTs, and the strategies undertaken to overcome those challenges in care settings for older adults.

Methods: The review followed the Joanna Briggs Institute (JBI) methodology as a guide. The published studies were searched through CINAHL, MEDLINE, JBI, and Web of Science. The unpublished studies were searched through Mednar, Trove, OCLC WorldCat, and Dissertations and Theses. Studies published in English from 2016 were considered. This review included studies of older individuals (aged 60 years and older) undergoing care delivery associated with DHTs and respective care providers. The concept involved the application of technology, and the context included studies based on care settings for older adults. A broad scope of evidence, including quantitative, qualitative, text, and opinion studies, was considered. In addition, 2 independent reviewers screened the titles and abstracts and reviewed the full text.

Results: The results will be presented in a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram. A total of 2 draft charting tables were developed and presented. A summary of the characteristics of the included studies was then described in terms of location, study sites, timing, participants, and outcomes measured or phenomena of interest. A result-based convergent (integrated) synthesis design was used to identify 5 key challenges. Those challenges included (1) data security and privacy concerns, (2) equity and accessibility of health care, (3) effectiveness concerning context, timing, and location, (4) ethical implications regarding autonomy, consent, and overdiagnosis, and (5) the impact of DHTs on health care workflows and provider workload.

Conclusions: The studies reviewed reveal several critical characteristics regarding the implementation of DHT technologies and their associated ethical considerations, particularly in terms of safety, equity, timing, location, participant characteristics, and workflow impact. The implications of these challenges emphasize the necessity for more practical ethical guidelines and policy frameworks to mitigate the potential risks associated with DHT application in older care. Further research should be conducted to examine other dimensions of the quality of care, such as access, timeliness, acceptability, and appropriateness.

International Registered Report Identifier (IRRID): RR2-10.2196/51153

(JMIR Aging 2025;8:e73925) doi:10.2196/73925

KEYWORDS

patient safety; health equity; effectiveness; equality; accessibility; social care; data security; right to privacy; patient consent; overdiagnosis; PRISMA; Preferred Reporting Items for Systematic Reviews and Meta-Analyses

XSL•FO
RenderX

Introduction

Digital health, with its tools for telemedicine, remote monitoring, medication management, and social connection, is a reassuring presence in the care of older adults. It is not just about enhancing their health, independence, and quality of life, but also about promoting healthy aging. This is achieved through the promotion of preventive care, disease self-management, and the personalized interventions that address the specific needs of older adults [1,2]. Digital health interventions can be tailored to individual needs and abilities (personalized care), addressing specific health concerns like falls, sarcopenia, and mental health challenges to support independence and well-being [3,4].

The advent of digital health twins (DHTs) marks a pivotal evolution in health care and social care, particularly in the management of care for older adults. These sophisticated virtual representations simulate individual health profiles by integrating extensive datasets, promising personalized health and social care interventions that can significantly enhance patient outcomes [5,6]. Despite the considerable advantages that DHTs present—improvements in effectiveness, safety, and accessibility of care—their implementation raises critical ethical dilemmas and challenges related to the quality of care that necessitate rigorous examination [7,8]. DHTs can be more than just a digital replica or a virtual model of patients (physical twins); they can be a sophisticated representation designed to faithfully mirror the real-world system in real time, analyze their behavior, and provide predictive insights using advanced simulation, machine learning, and reasoning to inform decision-making. The analytical and predictive capability of a digital twin makes it distinct from a dummy replica of the physical system [9,10]. DHTs are generated from multimodal patient data, population data, and real-time updates on patient and environmental variables [11]. For the convenience of the reader, the term “digital health twins” (DHTs) has been consistently maintained throughout the article. However, several other related terms, such as “digital twins” and “personal digital twins,” have also been used to depict the findings, thereby maintaining the originality and credibility of the included studies, primarily in the results section and tables.

DHTs are considered the pinnacle of personalized health care due to their ability to create a virtual replica of a patient’s health information, created using various data sources—from genetic information to real-time biometric data [6,12]. One of the primary concerns associated with DHTs is their potential to redefine established health norms. Creating personalized health models could inadvertently establish new benchmarks affecting the safety and effectiveness of health and social care interventions across diverse patient populations [13]. This technology has the potential to vastly improve the accuracy and efficiency of health care delivery, leading to better outcomes for patients and providers alike [12].

However, deploying such innovations must be approached with caution to mitigate risks related to data privacy and security, particularly among older adults who are often more susceptible to data breaches [14]. DHTs present heightened privacy and security risks compared to other technologies due to their reliance on sensitive, real-time patient data and their complex, interconnected nature. The continuous and urgent influx of personal health information, combined with the potential for remote access and manipulation, makes DHTs a prime target for cyberattacks and privacy breaches, highlighting the need for collaborative efforts among professionals to address this issue [15]. As digital health solutions become increasingly pervasive, the ethical ramifications surrounding data ownership, consent, and potential misuse warrant careful consideration [16,17].

DHTs can revolutionize personalized care planning. Numerous studies have demonstrated that these models can enhance the customization of interventions, leading to improved health outcomes [16]. However, the inherent personalization also introduces ethical considerations regarding data representation and equitable access to care. Bahrami et al [18] emphasized concerns regarding equitable access to tailored drug therapies, which may inadvertently exacerbate existing disparities in health and social care.

The potential of DHTs to facilitate early detection of health issues, such as dementia, is widely acknowledged [19]. These systems can enhance care safety and effectiveness through timely interventions and provoke significant privacy and autonomy concerns [20]. The ethical implications of using technology for health monitoring necessitate the development of robust frameworks that prioritize patient autonomy while promoting engagement with digital health solutions [21]. This highlights the importance of considering ethical aspects of DHT in health and social care.

Furthermore, the intersection of DHTs with traditional health and social care practices raises profound ethical questions concerning the equilibrium between potential benefits and harms. For instance, Lin et al [22] scrutinized the risk of overdiagnosis in colorectal cancer screening programs, highlighting the likely exacerbation of this issue by DHTs. The unintended consequences of digital health interventions include patient anxiety and unnecessary medical procedures, underscoring the necessity for thorough ethical scrutiny [23]. As health and social care transitions toward a data-driven paradigm, integrating emerging technologies must be guided by ethical principles that prioritize patient welfare.

Incorporating innovative technologies, including virtual reality (VR), within DHT frameworks opens new avenues for training and treatment [24]. While VR simulations can enhance health and social care training and improve treatment outcomes, they also present ethical challenges regarding informed consent and data use. Ensuring these technologies are used ethically and responsibly is paramount to maintaining public trust in digital health innovations.

This research investigates the ethical and quality of care-related challenges associated with DHTs in older adult care settings. A scoping review has been chosen because it is designed to map key concepts and examine the literature in a research area, providing an overview of the current evidence available. By synthesizing existing literature and identifying knowledge gaps, particularly as articulated in the protocol by Jabin et al [8], this study aims to provide a comprehensive understanding of the

XSL•FO
RenderX

implications of DHTs. In doing so, it will underscore the necessity for frameworks that ensure technological advancements in health and social care do not compromise the fundamental ethical principles of equity, autonomy, and patient-centered care. This reinforces the need for ethical guidelines in the digital health field, where integrating novel technologies must be closely aligned with ethical considerations without compromising the quality of health and social care. Specifically, the review questions are:

- What problems are faced by older individuals (their family and relatives), and health and social care providers, associated with the application of DHTs in care settings for older adults?
- What are the documented issues related to the quality of care for older adults, such as safety, equity, effectiveness, and accessibility, concerning DHTs?
- What are the ethical challenges concerning the application of DHTs used in care settings for older adults?
- What strategies have been evaluated and implemented in care settings for older adults that address the challenges associated with DHTs?

Methods

The proposed scoping review was conducted under the guidance provided by the Joanna Briggs Institute (JBI) methodology for scoping reviews [25]. The JBI methodology for scoping reviews offers a systematic and structured approach for mapping and reviewing vast areas of study.

Eligibility Criteria

Overview

This scoping review included studies by following the population, concept, and context (PCC) mnemonics. These mnemonics were used as a guide (not a policy); therefore, the inclusion criteria of this scoping review included a detailed description of the types of participants, concepts, and context, as well as search strategies, data extraction, charting, analysis, and presentation of the results. The eligibility criteria are listed and described in detail (see Table 1 and Table S1 in Multimedia Appendix 1) of the original protocol published by Jabin et al [8].

XSL•FO
RenderX

Table . Quality of care, ethical challenges, and strategies of included studies.

Study reference (year)	Quality of care	Ethical challenges	Strategies
Bruynseels et al [12] (2018)	Redefines health norms, affecting safety and effectiveness	Inequality and discrimination risks	Governance for data transparency and privacy
Liu et al [26] (2019)	Enhances monitoring and prediction, affecting effectiveness and accessibility	—[a]	Cloud-based framework to improve care delivery
Chakshu et al [16] (2019)	Noninvasive health monitoring enhances safety and effectiveness	Not directly addressed	Semiactive digital twin model for health monitoring
Calderita et al [27] (2020)	CPS[b] integration improves caregiving, potentially affecting safety and accessibility	Not directly addressed	CPS-AAL[c] systems for quality-of-life enhancement
Vidal et al [28] (2020)	Care coordination through integrated living models (ILMs) and advancing noninvasive sensors for accurate health monitoring	Transparent and unbiased use of diverse health data and algorithms	Multidisciplinary collaboration, extending research across all age groups, and enhancing caregiver coordination through ILM
Barbiero et al [11] (2021)	Digital twin framework could improve personalized care plans	Personalization implicates ethical considerations	Framework as a strategy for preventative medicine
Jovanovic et al [29] (2021)	Affects health intervention effectiveness with vaccination strategies	Impact of information dissemination on ethics	Vaccination strategy simulation model
Kobayashi et al [19] (2021)	System for early dementia sign detection enhances care safety and effectiveness	Privacy and autonomy concerns	Early detection system for dementia care
Khan et al [30] (2022)	Unobtrusive sensors for data collection are relevant to safety and effectiveness	Privacy concerns due to data collection	Microwave sensing for nonintrusive data collection
Bahrami et al [7] (2022)	Tailored drug therapy digital twin addresses safety and effectiveness	Raises questions about equitable access	Physics-based digital twin for personalized medicine
Wickramasinghe et al [31] (2022)	Advocates for digital twins in dementia care for effectiveness and personalization	Not directly addressed	Clinical decision support model for dementia care
Sahal et al [32](2022)	PDTs (patient-delivered therapies) improve health care effectiveness.	Blockchain hints at addressing privacy	Reference framework for personalized health care
Zhou et al [33] (2022)	Metaverse in cognitive decline intervention could affect accessibility and effectiveness	Privacy and autonomy issues with meta-hospitals	Metaverse as a strategy for nonpharmacological interventions
Alves et al [5] (2022)	VR[d] simulator for robot training impacts safety and effectiveness	Consent in data used for simulation training	VR simulation for training without patient data
Bahrami et al [18] (2023)	Effects of physiological features on pain relief, impacting effectiveness	Equity concerns with tailored therapy	Digital twin control of therapy for pain management
Bui [21] (2023)	Insights into elderly perceptions of DHT[e] affect accessibility and acceptance	User acceptance implicates autonomy and privacy issues	Behavioral intention understanding for DHT adoption
Lin et al [22] (2023)	Assesses overdiagnosis in CRC[f] screening, affecting safety and effectiveness	Raises ethical concerns about harm versus benefit	Digital twin approach to mitigate overdiagnosis
Zhao et al [20] (2023)	Assisted experience of seniors with technology affects accessibility and effectiveness	Ethical considerations of autonomy and equity	Digital twin remote collaboration enhancement
Thamotharan et al [23] (2023)	HDT[g] for diabetes management impacts safety and effectiveness	Personalized treatment raises privacy and equity issues	HDT framework for individualized diabetes treatment
Cai et al [34] (2023)	STRIDE[h] for ADRD[i] risk evaluation improves effectiveness and safety in gait analysis	Privacy concerns with gait data collection	STRIDE with MDR[j] and DL[k] for noninvasive risk evaluation

XSL•FO
RenderX

[a]Not available.

[b]CPS: Cyber-Physical System.

[c]CPS-AAL: Cyber-Physical System for Ambient Assisted Living.

[d]VR: virtual reality.

[e]DHT: digital health twin.

[f]CRC: colorectal cancer.

[g]HDT: health digital twin.

[h]STRIDE: Specific context needed, often refers to "Strategies for Risk Identification and Development" or other frameworks.

[i]ADRD: Alzheimer disease and related dementias.

[j]MDR: multidrug resistance.

[k]DL: deep learning.

Participants

This review included studies of older individuals (aged 60 years and older) undergoing older care associated with DHTs, irrespective of gender and diversity, including age, ethnicity, socioeconomic status, disorders, and disability. Studies on caregivers (family or friends—paid or unpaid) and care providers (licensed or unlicensed) involved in older care concerning DHT were also included.

Concept

The key concept in this scoping review is the process and application of DHT. Studies evaluated the application of DHT involving older care providers, older individuals, or family, friends, and relatives.

Context

The scoping review considered studies in care settings for older adults associated with DHT, such as geriatric wards of primary health care, hospitals or clinics, old-age homes, nursing homes, care homes, and home care facilities for older individuals.

Search Strategy

Databases were searched for both published and unpublished studies. The approach to searching for studies for a scoping review followed the standard 3-step method [8]. The first step involved an initial limited search of relevant databases, followed by an analysis of the text words in the title and abstract, as well as the index terms used to describe the article. The search for published studies included a 2-way search strategy. One way was to search the journal and reference databases, such as CINAHL, MEDLINE, JBI, and Web of Science. Another way was to search article-based (journal) databases, such as ACM Digital Library, IEEE Xplore, and BMJ Journals. The search for unpublished studies included Mednar, Trove, OCLC WorldCat, and Dissertations and Theses. A second search was undertaken across all included databases using all identified keywords and index terms. Additional search strategies, including citation searches for specific researchers or articles (eg, gold-standard articles) and chain searches (reviewing the reference lists of systematically selected articles), were included to complement the search for published and unpublished papers. Studies, such as reviews (systematic, scoping, and umbrella) and letters to editors, were excluded. Studies published in English were considered. While the precise origin of the specific phrase “digital health twin” is hard to pinpoint, the foundational concepts of digital twins began gaining significant traction in healthcare around 2016, driven by advancements in related fields like the internet of medical things and artificial intelligence [24].

Types of Studies

This scoping review considered experimental and quasi-experimental study designs, including randomized controlled trials, nonrandomized controlled trials, before-and-after studies, and interrupted time-series studies. In addition, analytical observational studies, including prospective and retrospective cohort studies, case-control studies, and analytical cross-sectional studies, were also considered for inclusion. This review considered descriptive observational study designs, including case series, individual case reports, and descriptive cross-sectional studies for inclusion.

Qualitative studies were also considered, focusing on qualitative data, including, but not limited to, designs such as phenomenology, grounded theory, ethnography, qualitative description, and action research. Text and opinion papers regarding the benefits, challenges, and strategies to overcome the challenges posed by DHT were also considered, as the scoping review includes a broad scope of evidence.

Source of Evidence Selection

Following the search, all identified citations were collated and uploaded into EndNote 20 (Clarivate Analytics), and duplicates were removed. Following a pilot test, titles and abstracts were then screened by 2 or more independent reviewers for assessment against the inclusion criteria for the review. Potentially relevant sources were retrieved in full, and their citation details were imported into a Microsoft Excel sheet instead of the JBI System for the Unified Management, Assessment, and Review of Information [35] due to a lack of resources. In addition, 2 independent reviewers assessed the full text of selected citations in detail against the inclusion criteria. Reasons for excluding sources of evidence in full text that did not meet the inclusion criteria were recorded and reported in the scoping review. Any disagreements that arose between the reviewers at each stage of the selection process were resolved through discussion or with the assistance of an additional reviewer (as an arbitrator). The search results and the study inclusion process were reported in full in the final scoping review and presented in the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) flow diagram [36,37].

XSL•FO
RenderX

Data Extraction

A total of 2 independent reviewers used a data extraction tool they had developed to extract data from the papers included in the scoping review (see Table 1). The data extracted included specific details about the participants, concept, context, study methods, and key findings relevant to the review questions.

A draft charting table was developed as a data extraction tool. While extracting data from each included evidence source, the charting table was modified and revised as necessary. The modifications were detailed in the scoping review. Any reviewer disagreements were resolved either by the reviewer or with the assistance of an additional reviewer (as an arbitrator).

We extracted variables to ensure alignment with our research questions, while some variables may emerge during full-text screening; therefore, it was essential to predefine the key data extraction elements (as the protocol is structured). Clearly outlining these variables ensured consistency in data collection, enhanced transparency in the scoping review methodology, and strengthened the study's replicability. We also followed the JBI Manual for Evidence Synthesis as a guide, which provided a basic data extraction table template to ensure completeness, methodological consistency, and adherence to best practices in data extraction [38].

Results

Overview

The results are presented as a "map" of the data extracted from the included papers, in a tabular form (as necessary) and in a descriptive format that aligns with the review's objective and scope. A clear explanation for each category is provided, accompanied by a narrative summary that describes how the results relate to the review's objectives and questions. This scoping review synthesizes key findings from the literature on the application of DHTs in older adult care settings.

A result-based convergent synthesis design was used in which all types of studies, such as quantitative and qualitative, were analyzed and presented [39]. The findings were then integrated using convergent synthesis, which identified five key challenges. Those challenges included (1) data security and privacy concerns, (2) equity and accessibility of health care, (3) effectiveness concerning context, timing, and location, (4) ethical implications regarding autonomy, consent, and overdiagnosis, and (5) the impact of DHTs on health care workflows and provider workload.

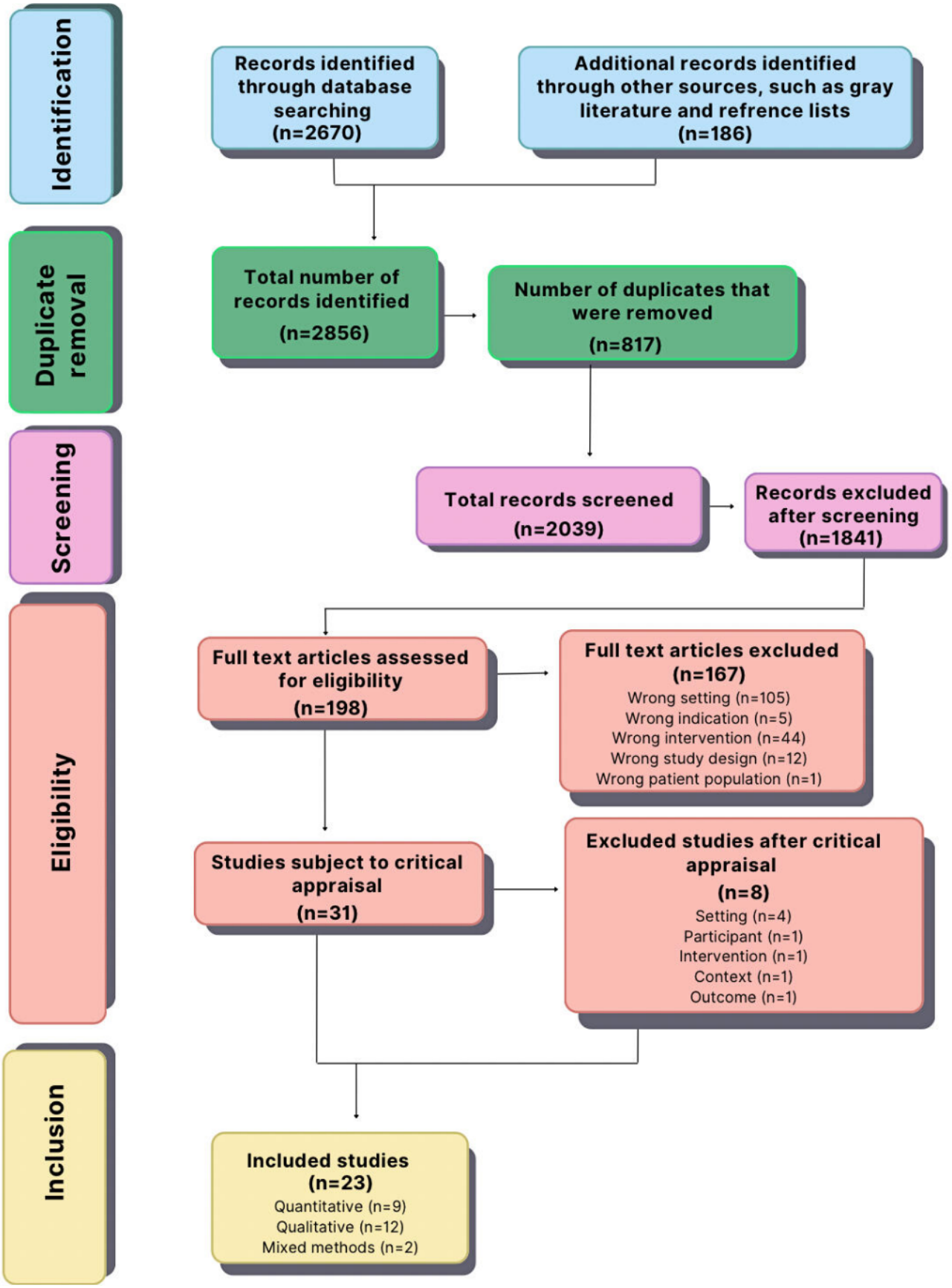
Search Strategies and Screening

Figure 1 shows the results of the search strategy and inclusion of studies as a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram. The total number of studies identified was 856, which included searching bibliographic databases (n=786 studies) and searching for gray literature and other sources, such as reference lists of included studies (n=70 studies). After removing 344 duplicates, a total of 512 records were screened. The titles and abstracts of each of the 512 studies were screened independently by 2 reviewers (the principal author as the first reviewer of all studies and the contributing authors as second reviewers on a similar number of studies each); 354 studies were excluded after screening because they were not relevant to the review question. A full-text review assessing the eligibility of the remaining 158 studies resulted in 20 studies being advanced to the final list of included studies, while 138 studies were excluded (see Figure 1).

A search strategy using Boolean search strings in databases is provided in Table S2 in Multimedia Appendix 2, and the inclusion and exclusion criteria used during the review process are outlined in Table S3 in Multimedia Appendix 2.

XSL•FO
RenderX

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram of the studies in the scoping review.



XSL•FO
RenderX

Summary of the Characteristics of the Included Studies

The studies reviewed emphasize the integration of DHTs in health care, offering a wide range of applications, including virtual simulations, disease monitoring, and health care delivery.

Location, Study Sites, and Timing

The studies vary in terms of location and timing (see Table S1 in Multimedia Appendix 1). For example, Bruynseels et al [12] explore the potential of DHTs in health care settings, but do not specify particular sites or study duration. On the other hand, Khan et al [30] use microwave sensors in a static care-home model, which likely involves specific care-home locations, showcasing a practical implementation of DHTs in elderly care. The studies span different periods, with some focusing on contemporary advancements [32] and others taking a longer-term view of digital health transformation [26]. All included studies encompass a range of countries (n=14) globally, with the majority originating from the United Kingdom and Europe. The publication of these 20 studies, spanning from 2018 to 2023, suggests that DHT has gained attention only recently.

Participants

Participant involvement varies across the studies (see Table S1 in Multimedia Appendix 1). For example, Jovanovic et al [29] focus on vaccination strategy simulations, meaning their study does not involve direct patient participation but relies on data for modeling purposes. In contrast, the model by Bahrami et al [7] drug diffusion in an online patient, suggesting that their study could involve patient-based data for the creation of DHTs. Studies like those delve into ethical issues related to using DHTs in care settings for older adults, which would likely involve patient consent and ethical considerations for vulnerable groups [12,22,28].

Outcomes Measured/Phenomenon of Interest

The outcomes measured in the studies often focus on the effectiveness of DHT technology in improving health care outcomes (see Table S1 in Multimedia Appendix 1). For example, Liu et al [26] assess the use of cloud-based DHTs for elderly health care services, measuring outcomes such as patient monitoring and health service optimization. In contrast, studies like the one by Cai et al [34] focus on the early detection of Alzheimer disease, aiming to measure how DHTs could assist in identifying disease risk early, ultimately improving diagnosis and treatment outcomes.

Claims and Findings

The reviewed literature consistently highlighted that DHTs and associated technologies offer transformative potential for improving health care delivery, particularly through personalization, predictive analytics, and real-time monitoring (see Multimedia Appendix 1). However, shared concerns about ethics, equity, and data privacy were repeatedly emphasized across the studies. For example, Bruynseels et al [12] argued that DHTs raise significant ethical, legal, and societal issues, particularly regarding equality and potential discrimination, which resonated with Chakshu et al [16], who, although focusing on noninvasive detection of carotid stenosis, implicitly align with Bruynseels and colleagues' [12] ethical stance by advocating for procedural alignment and ethical validation in clinical application. Both stressed the need for patient autonomy and safeguarding rights in digital health innovation [12,16].

Similarly, the concerns around data privacy and security raised by Khan et al [30], who examined unobtrusive microwave sensors in care homes, were mirrored in Kobayashi et al [19], who discussed privacy and autonomy concerns in their dementia detection system. Both highlighted that the collection of personal health data—whether through sensors or DHTs—demands robust ethical and legal frameworks to protect vulnerable populations, particularly older adults [19,30].

In terms of personalized care and equity, Bahrami et al [7] proposed a physics-based digital twin for tailored drug therapy, focusing on personalized medicine, which aligned closely with Barbiero et al [11], who developed a general digital twin framework emphasizing modularity and scalability in representing biomedical data. Both studies demonstrated how, although helpful for precision treatment, personalized health care models could worsen disparities if access is not sufficiently addressed [7,11]. Wickramasinghe et al [31], who emphasized the necessity of inclusive clinical decision support in dementia care, share this issue.

Further illustrating a shared commitment to enhancing health care efficiency and safety, while acknowledging ethical concerns such as informed consent and data governance, is the integration of cutting-edge technologies, including VR simulators by Alves et al [5] and personal digital twins (PDTs) by Sahal et al [32]. Alves and colleagues' [5] focus on online training environments and Sahal and colleagues' [32] blockchain-enabled PDT framework were both operating in tandem, suggesting technological steps to ensure openness and confidence [5,32]. Sahal and colleagues' [32] blockchain-enabled PDT framework and Alves' emphasis on online training environments complemented each other by recommending technological measures to guarantee transparency and trust [5,32].

A notable similarity existed between Jovanovic et al [29] and Lin et al [22] in their use of simulation and modeling to guide public health interventions—Jovanovic et al through vaccination strategy simulations and Lin et al via overdiagnosis assessments in colorectal cancer screening [22,29]. Both emphasized the importance of data-driven decision-making to optimize health outcomes and minimize unintended consequences, such as overdiagnosis or inefficient resource allocation [22,29].

Finally, the user perception and adoption challenges explored by Bui [21], regarding older adults' attitudes toward DHTs, were closely reflected in Zhao et al [20], who examined older users' expectations of intelligent technologies. Both studies concluded that user acceptance was influenced by privacy concerns, ease of use, and perceived benefits, underlining the critical need for human-centered design in DHT solutions [20,21].

Integrated Synthesis

This review identified a range of challenges based on the key characteristics of the included studies (see Table 1), which are briefly described hereinafter:

XSL•FO
RenderX

Safety Concerns and Data Privacy

Safety remains a primary concern in the deployment of DHTs, particularly regarding patient data security and privacy.

Risks to Patient Privacy

DHTs, as virtual representations of patients, relied on integrating extensive multimodal datasets [12], including health records, environmental data, and real-time monitoring information [5]. While these technologies offered the potential to improve safety through personalized health monitoring and predictive capabilities, they also raised significant risks to patient privacy.

Unauthorized Data Access and Misuse

Older adults are often more vulnerable to data breaches due to cognitive impairments or a lack of digital literacy [21], and require additional safeguards [34]. Studies by Liu et al [26] and Khan et al emphasized the potential for unauthorized data access and misuse [26], which could undermine trust in DHT technologies. This concern is amplified in older populations [30] who may lack familiarity with digital systems and are more susceptible to exploitation.

Protecting patient data through secure and transparent privacy protocols is essential to ensure the safe adoption of these technologies in care settings for older adults.

Equity and Access to Care

Equity in access to DHT technologies is another significant concern, especially in the context of older adult care. Although DHTs hold promise for personalized care, disparities in access to technology may exacerbate existing health inequities [16].

Technological Barriers Affect Older Care

Older adults, particularly those from disadvantaged socioeconomic backgrounds, may have limited access to the required digital infrastructure or struggle with digital literacy. Furthermore, some older individuals may face technological barriers, such as a lack of internet access or difficulty operating digital devices, which limit their ability to benefit from DHT-based care solutions.

Disparities in Health Outcomes

Research studies by Bahrami et al [7,18] highlighted how inequitable access to tailored drug therapies and health interventions could worsen disparities in health outcomes if these technologies were not broadly accessible or affordable. Barbiero et al [11] similarly argued that without equitable distribution and access to DHT-based care models, exacerbating the social determinants of health and creating new forms of inequality in health care delivery were at risk.

To achieve true personalization of care, DHTs must be deployed with an emphasis on overcoming these access barriers, ensuring that all older adults benefit equally from these advancements.

Effectiveness of DHT Implementation in Relation to Location, Timing, and Context

The location and timing of DHT implementation significantly influence its effectiveness in care settings for older adults. Liu et al [26] and Bruynseels [12] showed that the efficacy of DHTs could vary considerably across different care environments, including home care, nursing homes, and hospital settings.

Limited Resources Cause Challenges in DHT Implementation

The technological infrastructure of these settings played a crucial role in determining the feasibility and success of DHT integration. For example, in-home care environments, where resources might be limited, implementing DHTs could be more challenging compared to well-equipped health care facilities [29].

Rapid Deployment Leads to Insufficient Validation of DHT Implementation

Furthermore, the timing of DHT deployment can influence its impact. This then emphasized the integration of DHTs during a health crisis (such as the COVID-19 pandemic), accelerated adoption, but also raised challenges related to rapid deployment [21], including insufficient validation and potential overreliance on untested technologies. The timing of DHT use in disease progression was also critical—early-stage interventions, particularly in conditions such as dementia, are most beneficial, as they could significantly improve patient outcomes [34].

Therefore, the contextual environment in which DHTs are deployed must be carefully considered to maximize the benefits of these technologies.

Ethical Concerns: Autonomy, Consent, and Overdiagnosis

Ethical concerns are central to implementing DHTs, particularly in terms of patient autonomy, informed consent, and the risk of overdiagnosis.

Informed Consent Becomes a Complex Ethical Issue

Many older adults, especially those with cognitive impairments, might face challenges in providing informed consent for digital health interventions [30]. Informed consent in these contexts became a complex ethical issue, as health care providers must balance the benefits of personalized care with the need to protect vulnerable patients' rights to autonomy and decision-making [22,26,28].

The Issue of Overdiagnosis Leads to Harmful Side Effects and Unnecessary Treatments

In addition to autonomy and consent, the issue of overdiagnosis was another significant ethical concern. Lin et al pointed out that the predictive capabilities of DHTs might lead to overdiagnosis, particularly in screening programs, resulting in unnecessary treatments and interventions. This concern was especially relevant for older populations [22], where the risks associated with overdiagnosis—such as the potential for harmful side effects or unnecessary treatments—are more pronounced due to comorbidities and frailty [26].

While DHTs offer the potential for early detection, they must be used judiciously to avoid these unintended consequences and ensure that interventions are genuinely beneficial rather than causing harm.

XSL•FO
RenderX

Impact on Care Workflow and Provider Workload

Introducing DHTs into care settings for older adults may significantly alter the care workflow, with both positive and negative implications.

The Need for Enhancing Care Efficiency

Studies by Bahrami et al [7] and Bui [21] highlighted that DHTs could enhance care efficiency by automating specific tasks [21], such as continuous health monitoring, data collection, and decision-making support [7]. These improvements could reduce the burden on health care providers and streamline care coordination. However, integrating these technologies also required adaptation to new systems and workflows, which might lead to initial disruptions and increased workload.

The Need for an Increase in Training Requirements for Existing Practices

Furthermore, as health care providers become familiar with DHT technologies, there may be an increase in training requirements and adjustments to existing practices, which could potentially lead to a higher workload in the short term. While the long-term benefits of DHTs, such as reduced manual tasks and improved diagnostic accuracy, are clear, the transition period must be carefully managed to avoid burnout or resistance among health care staff.

Therefore, successfully integrating DHTs into care settings for older adults requires thoughtful planning to balance the expected workflow improvements with the challenges of training and adapting health care staff to new technologies.

Strategies to Mitigate DHT Challenges

Several studies propose strategies to mitigate these ethical challenges that prioritize inclusivity, transparency, and patient autonomy (see Multimedia Appendix 1). Governance frameworks were aimed at ensuring the data had openness and privacy [12]. On the other hand, Sahal et al [32] and Zhou et al [33] advocated for integrating blockchain technologies and decentralized models to enhance data security and user control. In addition, Vidal et al [28] and Wickramasinghe et al [31] highlighted the importance of multidisciplinary collaboration and integrated care models, such as integrated living models, which promoted coordinated care and supported the ethical implementation of DHTs in practice [28,31].

Collectively, these studies illustrated a shared understanding of both the opportunities and ethical responsibilities associated with the deployment of DHTs in older adult care settings. There has been a consistent call for robust ethical oversight, patient-centered governance, and policies that ensure equitable access to these technologies [12,21,23].

Further Suggestions

Several studies offer suggestions for improving DHT integration into health care. Most studies highlighted ethical and quality of care challenges in care settings for older adults and recommend developing clear guidelines for using DHT (see Multimedia Appendix 1). For instance, Venkatesh et al [9] suggested considering computation, implementation, and regulation challenges to ensure the proper functioning of health digital twins for precision medicine. These suggestions highlighted the importance of addressing technical and ethical hurdles for DHTs to integrate into health and care systems fully.

Discussion

Principal Findings

The implementation of DHTs in health and social care for older adults presents both significant opportunities and inherent risks. In this scoping review, 20 studies were included from a total of 856 identified through a systematic literature search. This review aims to gain insights into the best available evidence on the implementation of DHTs for older adults in health and social care. This discussion critically examines five key challenges that emerge from the literature. These challenges include (1) data security and privacy concerns; (2) equity and accessibility of health care; (3) effectiveness in relation to context, timing, and location; (4) ethical implications regarding autonomy, consent, and overdiagnosis; and (5) the impact of DHTs on health care workflows and provider workload. Overcoming such issues will ensure that DHTs contribute to the quality of care of older individuals and do not take away from it.

The increasing reliance on DHTs requires robust data protection, particularly for vulnerable populations, such as older individuals. DHTs are based on large multimodal datasets, including biometric data, health records, and real-time environmental sensing, which raise significant ethical and privacy concerns [26,30]. Several studies highlight the risks associated with unauthorized data access and potential misuse, particularly for older individuals who may lack familiarity with digital systems and, consequently, are more vulnerable to exploitation [5,12]. Although DHTs provide substantial benefits, such as continuous health monitoring [40] and predictive health care analytics [7], they also introduce risks, including data breaches and unwarranted surveillance [19]. The establishment of ethical guidelines that mandate transparency, accountability, and patient consent is essential in curbing these threats [20]. Blockchain technology, as proposed by Sahal et al [32], could potentially provide a safe manner of data management; nonetheless, empirical investigations are required to ascertain its usefulness in real-life health care settings.

While DHTs offer potential advancements in personalized health care, concerns remain regarding equitable access. The digital divide, particularly amongst older cohorts within lower socioeconomic segments, stands to exacerbate health care disparities [21]. Furthermore, digital twin-controlled drug therapies and individualized health interventions may inadvertently become privileged health care solutions, accessible predominantly to wealthy populations, thereby further aggravating existing disparities [11]. Equity must take precedence while deploying DHTs, with policy measures to curb accessibility barriers through government policies, ring-fenced finances, and digital literacy programs. Otherwise, integrating DHTs will likely widen health and care inequalities rather than bridge them.

The performance of any digital health technology dramatically depends on the environment in which they are implemented

XSL•FO
RenderX

[41,42]. Studies show that DHTs perform optimally in well-structured health care settings, such as hospitals and specialist clinics, with strong technological infrastructures that support real-time patient location and accurate data integration [26]. The usefulness of such technologies in home-care environments where technology and care capabilities are limited has not been adequately established [27]. On the other hand, the timing of DHT implementation is also a critical factor influencing its success. The rapid adoption of DHTs during the COVID-19 pandemic raised concerns regarding premature reliance on untested digital health solutions [20]. Similarly, DHT interventions conducted early in the course of dementia care are more beneficial than interventions done in later stages [19], and improve pain management and reduce patient outcome variability with the effect of physiological features on the achieved pain relief [18]. Therefore, strategies for implementing DHT must be customized according to individual health status, disease stages, and diverse care settings to optimize their impact.

The application of DHTs generates complex ethical questions, primarily about patient autonomy and informed consent. Older people with impaired cognition may be unable to fully grasp the implications of consenting to continuous digital health monitoring [29]. In addition, coercion, the compulsion that persons may feel in adopting DHTs by virtue of health care policy or social culture, is of interest to voluntary involvement [21]. The second ethical concern is overdiagnosis [43,44]. DHT predictive algorithms can lead to an overestimation of disease risk and, consequently, unnecessary medical procedures that increase patient anxiety and place additional pressures on health care systems (Lin et al [22]). While early diagnosis is one of the significant advantages of DHTs, careful calibration is necessary to distinguish between necessary interventions and overmedicalization. Ethical guidelines governing the use of DHT must prioritize informed patient choice over algorithmic assumptions, so that clinical decisions are made appropriately based on clinical need [28,45].

The presence of DHTs in geriatric care settings has conflicting effects on health care workers. On the one hand, DHTs can streamline care processes, reducing administrative workload through automated data collection and real-time patient monitoring, thereby enhancing clinical decision-making [7,20]. On the other hand, the transition to DHT-based care models necessitates significant changes to clinical workflows, potentially increasing short-term workload and requiring extensive training [31]. This necessitates that health care professionals be adequately trained to interpret and apply DHT-generated insights to patient care plans. Research by Alves et al [5] suggests that virtual reality (VR)-based training simulations facilitate the integration of digital twin technology before full-scale implementation. However, unless carefully managed, there is a risk of overreliance on automated decision-support systems, potentially undermining the clinical judgment of health care professionals [23]. Consequently, the integration of DHTs must be approached with caution, ensuring a balance between technological efficiency and the human expertise essential to high-quality patient care.

Strengths and Limitations of the Study

The study provides a strong ethical analysis, but it would benefit from more involvement of health care professionals and older adults, which could have enriched the findings with practical insights. The study highlights the importance of balancing technological innovation with ethical considerations to ensure DHTs enhance, rather than compromise, the quality of care for older adults. One of the key advantages is that this provides valuable insight into the benefits. A key strength is a systematic approach; the guidelines by JBI's methodology for scoping reviews ensure methodological rigor. This study is comprehensive, as it consolidates evidence from various studies on the safety, equity, practicality, ethics, and impact on care workflow of these digital clinical points of care. This illustrates the potential of DHTs in early disease detection and dementia, particularly in highlighting their potential contribution to preventive health care [19]. However, concerns over overdiagnosed patients and unintended consequences, including heightened anxiety and unnecessary procedures, highlight the lack of empirical validation [22].

A significant limitation is the use of secondary data drawn mostly from the published literature, which may exhibit bias or an evidence gap, particularly in the context of mask implementation for DHTs in practice. Another drawback of the study is that the latest search was conducted in October 2023; no further search could have been performed due to lack of resources, including funding and the research team's time commitment. For the same reason, that is, the lack of resources, we also did not follow the JBI Methodology strictly, but instead used it as a guide. Although the search identified 786 studies, some may have been missed. Due to the strict adherence to quality assurance in the gray literature search, including the credibility, objectivity, and accuracy of non–peer-reviewed sources, no unpublished articles could have been included in this review, except the study by Bui [21] (thesis) in 2023. Another limitation is that only studies published in English were included. However, the findings of 20 included studies relate to a range of 14 countries worldwide, with most studies emanating from the United Kingdom and Europe.

Ethical Frameworks and Policy Development

The following key considerations should guide future practice and policy:

- The implications of this study emphasize the necessity for solid ethical guidelines and policy frameworks to mitigate the potential risks associated with DHT application in older adult care and to ensure that the patient remains the ultimate decision-maker [30,46,47].
- Patients' autonomy and data privacy should be protected while fair access to technology that has the potential to enhance patient care is maintained [20,21].

Health Care System: Adaptation

To support the effective integration of DHTs into care for older adults, several key considerations for system adaptation can be identified:

XSL•FO
RenderX

- This study clearly highlights the pressing need for health and social care providers and institutions to integrate DHTs thoughtfully [48].
- It is imperative to strike a balance between the benefits offered by these technologies, particularly their predictive and monitoring capabilities. Concerns may then be related to overdiagnosis, increased workloads, and the potential for disparities in care delivery [7,22].
- Limitations associated with relying solely on secondary data could indicate that future research should prioritize empirical validation through direct engagement with health care professionals, caregivers, and older adults [49,50].
- Engagement is essential for refining the practical applications of DHTs and effectively assessing their real-world impact [9,19,29].

Future Research and Stakeholder Engagement

The following key directions are proposed to guide future research and strengthen stakeholder engagement:

- Given the limitations of relying on secondary data, future studies should prioritize empirical validation through co-design with stakeholders, such as health care professionals, caregivers, and older adults.
- All to refine the practical applications of DHTs and assess their real-world impact [19,29].
- To address stakeholder engagement and practical relevance, the findings can be effectively translated into practice through the development of guidelines and frameworks.
- Further research can examine other dimensions of the quality of care, such as access, timeliness, acceptability, and appropriateness.

Conclusion

This scoping review was conducted to assess the available evidence on implementing DHTs for the care of older adults, highlighting both significant opportunities and inherent risks. A systematic search of the literature yielded 20 studies from an initial 856 records. The studies revealed several critical characteristics regarding the implementation and ethical considerations of DHT technologies, particularly in terms of safety, equity, timing, location, participant characteristics, and their impact on workflow. A total of 5 key challenges were identified through the integrated synthesis, which is crucial for understanding the potential and limitations of DHTs in enhancing care for older adults. The challenges included (1) data security and privacy concerns; (2) equity and accessibility of health care; (3) effectiveness concerning context, timing, and location; (4) ethical implications regarding autonomy, consent, and overdiagnosis; and (5) the impact of DHTs on health care workflows and provider workload. The implications of these challenges underscore the need for more practical ethical guidelines and policy frameworks to mitigate the potential risks associated with DHT application in older care, ensuring that patients remain the ultimate decision-makers. Further research should be conducted to examine other dimensions of the quality of care, such as timeliness, acceptability, and appropriateness.

Acknowledgments

The authors would like to thank Evalill Nilsson and the librarians from Linnaeus University and the University of Bradford for their guidance and assistance in this research. A small grant, specifically the Brad-Attain Research Placement, was received from the University of Bradford for a student working on this project. A publishing grant has been obtained from Linnaeus University as a part of the University Library's research support.

Data Availability

All data generated or analyzed during this study are included in this published article.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Key characteristics of the included studies.
[DOCX File, 43 KB - aging_v8i1e73925_app1.docx]

Multimedia Appendix 2
Additional tables on the search strategy on databases and the inclusion and exclusion criteria.
[DOCX File, 21 KB - aging_v8i1e73925_app2.docx]

Checklist 1
PRISMA-ScR checklist.
[PDF File, 660 KB - aging_v8i1e73925_app3.pdf]

References

XSL•FO
RenderX

1. Alruwaili MM, Shaban M, Elsayed Ramadan OM. Digital health interventions for promoting healthy aging: a systematic review of adoption patterns, efficacy, and user experience. Sustainability 2023;15(23):16503. [doi: 10.3390/su152316503]
2. Jabin MSR, Pan D, Nilsson E. Characterizing patient details-related challenges from health information technology-related incident reports from Swedish healthcare. Front Digit Health 2024;6:1260521. [doi: 10.3389/fdgth.2024.1260521] [Medline: 38380372]
3. Daniels K, Bonnechère B. Harnessing digital health interventions to bridge the gap in prevention for older adults. Front Public Health 2023;11:1281923. [doi: 10.3389/fpubh.2023.1281923] [Medline: 38259780]
4. Jabin MSR, Pan D, Nilsson E. Characterizing healthcare incidents in Sweden related to health information technology affecting care management of multiple patients. Health Informatics J 2022;28(2):14604582221105440. [doi: 10.1177/14604582221105440] [Medline: 35762538]
5. Alves SFR, Uribe-Quevedo A, Chen D, Morris J, Radmard S. Developing a VR simulator for robotics navigation and human robot interactions employing digital twins. Presented at: 2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW); Mar 12-16, 2022; Christchurch, New Zealand. [doi: 10.1109/VRW55335.2022.00036]
6. Prakosa A, Arevalo HJ, Deng D, et al. Personalized virtual-heart technology for guiding the ablation of infarct-related ventricular tachycardia. Nat Biomed Eng 2018 Oct;2(10):732-740. [doi: 10.1038/s41551-018-0282-2] [Medline: 30847259]
7. Bahrami F, Rossi RM, Defraeye T. Predicting transdermal fentanyl delivery using physics-based simulations for tailored therapy based on the age. Drug Deliv 2022 Dec;29(1):950-969. [doi: 10.1080/10717544.2022.2050846] [Medline: 35319323]
8. Jabin MSR, Yaroson EV, Ilodibe A, Eldabi T. Ethical and quality of care-related challenges of digital health twins in older care settings: protocol for a scoping review. JMIR Res Protoc 2024 Feb 23;13:e51153. [doi: 10.2196/51153] [Medline: 38393771]
9. Venkatesh KP, Raza MM, Kvedar JC. Health digital twins as tools for precision medicine: considerations for computation, implementation, and regulation. NPJ Digit Med 2022 Sep 22;5(1):150. [doi: 10.1038/s41746-022-00694-7] [Medline: 36138125]
10. Björnsson B, Borrebaeck C, Elander N, et al. Digital twins to personalize medicine. Genome Med 2019 Dec 31;12(1):4. [doi: 10.1186/s13073-019-0701-3] [Medline: 31892363]
11. Barbiero P, Viñas Torné R, Lió P. Graph representation forecasting of patient's medical conditions: toward a digital twin. Front Genet 2021;12:652907. [doi: 10.3389/fgene.2021.652907] [Medline: 34603366]
12. Bruynseels K, Santoni de Sio F, van den Hoven J. Digital twins in health care: ethical implications of an emerging engineering paradigm. Front Genet 2018;9:31. [doi: 10.3389/fgene.2018.00031] [Medline: 29487613]
13. Taylor BJ. Developing an integrated assessment tool for the health and social care of older people. Br J Soc Work 2012 Oct 1;42(7):1293-1314. [doi: 10.1093/bjsw/bcr133]
14. Rodríguez-Aguilar R, Marmolejo-Saucedo JA. Conceptual framework of Digital Health Public Emergency System: digital twins and multiparadigm simulation. EAI Endorsed Trans Perv Health Tech 2020;6(21):e3. [doi: 10.4108/eai.13-7-2018.164261]
15. Yi H. Improving cloud storage and privacy security for digital twin based medical records. J Cloud Comp 2023;12(1):151. [doi: 10.1186/s13677-023-00523-6]
16. Chakshu NK, Sazonov I, Nithiarasu P. Towards enabling a cardiovascular digital twin for human systemic circulation using inverse analysis. Biomech Model Mechanobiol 2021 Apr;20(2):449-465. [doi: 10.1007/s10237-020-01393-6] [Medline: 33064221]
17. Jabin MSR. Identifying and characterising problems arising from interactions between medical imaging and health information technology as a basis for improvements in practice, in Australian Institute for Precision Health. : University of South Australia, Adelaide; 2019.
18. Bahrami F, Rossi RM, De Nys K, Defraeye T. An individualized digital twin of a patient for transdermal fentanyl therapy for chronic pain management. Drug Deliv Transl Res 2023 Sep;13(9):2272-2285. [doi: 10.1007/s13346-023-01305-y] [Medline: 36897525]
19. Kobayashi T, Fukae K, Imai T, Arai K. Dementia sign detection system using digital twin. Presented at: 2021 Ninth International Symposium on Computing and Networking (CANDAR); Nov 23-26, 2021; Matsue, Japan. [doi: 10.1109/CANDAR53791.2021.00025]
20. Zhao Y, Mao M, Wang Y. Research on the elderly user assisted experience based on digital twin remote collaboration system. In: Design Studies and Intelligence Engineering: IOS Press; 2023. [doi: 10.3233/FAIA220717]
21. Bui K. Factors influence the behavioural intention to use digital twin healthcare by the elderly: a cross-sectional survey. Tilburg University. 2023. URL: https://arno.uvt.nl/show.cgi?fid=161529 [accessed 2025-10-25]
22. Lin TY, Chiu SYH, Liao LC, Chen SLS, Chiu HM, Chen THH. Assessing overdiagnosis of fecal immunological test screening for colorectal cancer with a digital twin approach. NPJ Digit Med 2023 Feb 10;6(1):24. [doi: 10.1038/s41746-023-00763-5] [Medline: 36765093]
23. Thamotharan P, Srinivasan S, Kesavadev J, et al. Human digital twin for personalized elderly type 2 diabetes management. J Clin Med 2023 Mar 7;12(6):2094. [doi: 10.3390/jcm12062094] [Medline: 36983097]

XSL•FO
RenderX

24. Katsoulakis E, Wang Q, Wu H, et al. Digital twins for health: a scoping review. NPJ Digit Med 2024 Mar 22;7(1):77. [doi: 10.1038/s41746-024-01073-0] [Medline: 38519626]
25. Peters MDJ, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. Int J Evid Based Healthc 2015 Sep;13(3):141-146. [doi: 10.1097/XEB.0000000000000050] [Medline: 26134548]
26. Liu Y, Zhang L, Yang Y, et al. A novel cloud-based framework for the elderly healthcare services using digital twin. IEEE Access 2019;7:49088-49101. [doi: 10.1109/ACCESS.2019.2909828]
27. Calderita LV, Vega A, Barroso-Ramírez S, Bustos P, Núñez P. Designing a cyber-physical system for ambient assisted living: a use-case analysis for social robot navigation in caregiving centers. Sensors (Basel) 2020 Jul 18;20(14):4005. [doi: 10.3390/s20144005] [Medline: 32708496]
28. Vidal PP, Vienne-Jumeau A, Moreau A, et al. An opinion paper on the maintenance of robustness: towards a multimodal and intergenerational approach using digital twins. Aging Med (Milton) 2020 Sep;3(3):188-194. [doi: 10.1002/agm2.12115] [Medline: 33103039]
29. Jovanović R, Davidović M, Lazović I, Jovanović M, Jovašević-Stojanović M. Modelling voluntary general population vaccination strategies during COVID-19 outbreak: influence of disease prevalence. Int J Environ Res Public Health 2021 Jun 8;18(12):6217. [doi: 10.3390/ijerph18126217] [Medline: 34201285]
30. Khan S, Saied IM, Ratnarajah T, Arslan T. Evaluation of unobtrusive microwave sensors in healthcare 4.0-toward the creation of digital-twin model. Sensors (Basel) 2022 Nov 5;22(21):8519. [doi: 10.3390/s22218519] [Medline: 36366218]
31. Wickramasinghe N, Ulapane N, Andargoli A, et al. Digital twins to enable better precision and personalized dementia care. JAMIA Open 2022 Oct;5(3):ooac072. [doi: 10.1093/jamiaopen/ooac072] [Medline: 35992534]
32. Sahal R, Alsamhi SH, Brown KN. Personal digital twin: a close look into the present and a step towards the future of personalised healthcare industry. Sensors (Basel) 2022 Aug 8;22(15):5918. [doi: 10.3390/s22155918] [Medline: 35957477]
33. Zhou H, Gao JY, Chen Y. The paradigm and future value of the metaverse for the intervention of cognitive decline. Front Public Health 2022;10:1016680. [doi: 10.3389/fpubh.2022.1016680] [Medline: 36339131]
34. Cai F, Patharkar A, Wu T, Lure FYM, Chen H, Chen VC. STRIDE: Systematic radar intelligence analysis for ADRD risk evaluation with gait signature simulation and deep learning. IEEE Sens J 2023 May 15;23(10):10998-11006. [doi: 10.1109/jsen.2023.3263071] [Medline: 37547101]
35. Munn Z, Aromataris E, Tufanaru C, et al. The development of software to support multiple systematic review types: the Joanna Briggs Institute System for the Unified Management, Assessment and Review of Information (JBI SUMARI). Int J Evid Based Healthc 2019 Mar;17(1):36-43. [doi: 10.1097/XEB.0000000000000152] [Medline: 30239357]
36. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
37. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
38. Aromataris E. JBI Manual for Evidence Synthesis 2024. [doi: 10.46658/JBIMES-24-01]
39. Hong QN, Pluye P, Bujold M, Wassef M. Convergent and sequential synthesis designs: implications for conducting and reporting systematic reviews of qualitative and quantitative evidence. Syst Rev 2017 Mar 23;6(1):61. [doi: 10.1186/s13643-017-0454-2] [Medline: 28335799]
40. Jabin MSR. Operational disruption in healthcare associated with software functionality issue due to software security patching: a case report. Front Digit Health 2024;6:1367431. [doi: 10.3389/fdgth.2024.1367431] [Medline: 38550716]
41. Jabin MSR, Nilsson E, Nilsson AL, Bergman P, Jokela P. Digital Health Testbeds in Sweden: an exploratory study. Digit HEALTH 2022;8:20552076221075194. [doi: 10.1177/20552076221075194] [Medline: 35186314]
42. Pan D, Jabin MSR. Software-related challenges in Swedish healthcare through the lens of incident reports: a desktop study. Digit HEALTH 2023;9:20552076231203600. [doi: 10.1177/20552076231203600] [Medline: 37744748]
43. Jabin MSR, Schultz T, Mandel C, et al. A mixed-methods systematic review of the effectiveness and experiences of quality improvement interventions in radiology. J Patient Saf 2022 Jan 1;18(1):e97-e107. [doi: 10.1097/PTS.0000000000000709] [Medline: 32433438]
44. Jabin MSR, Schultz T, Hibbert P, Mandel C, Runciman W. Effectiveness of quality improvement interventions for patient safety in radiology: a systematic review protocol. JBI Database System Rev Implement Rep 2016 Sep;14(9):65-78. [doi: 10.11124/JBISRIR-2016-003078] [Medline: 27755318]
45. Steen M, Wepa D, Bergman P, Jabin MSR. Assessing the healthcare quality issues for digital incident reporting in Sweden: incident reports analysis. Digit HEALTH 2023;9:20552076231174307. [doi: 10.1177/20552076231174307] [Medline: 37188073]
46. Marasinghe KM, Chaurasia A, Adil M, Liu QY, Nur TI, Oremus M. The impact of assistive devices on community-dwelling older adults and their informal caregivers: a systematic review. BMC Geriatr 2022 Nov 24;22(1):897. [doi: 10.1186/s12877-022-03557-8] [Medline: 36424533]
47. Jabin MSR, Samuriwo R, Chilaka M, Yaroson EV. Effectiveness and experiences of quality improvement interventions in older adult care: protocol for a mixed methods systematic review. JMIR Res Protoc 2024 Apr 18;13:e56346. [doi: 10.2196/56346] [Medline: 38635311]

XSL•FO
RenderX

48. Chase JG, Zhou C, Knopp JL, et al. Digital twins in critical care: what, when, how, where, why? IFAC-PapersOnLine 2021;54(15):310-315. [doi: 10.1016/j.ifacol.2021.10.274]
49. Croatti A, Gabellini M, Montagna S, Ricci A. On the integration of agents and digital twins in healthcare. J Med Syst 2020 Aug 4;44(9):161. [doi: 10.1007/s10916-020-01623-5] [Medline: 32748066]
50. Kamel Boulos MN, Zhang P. Digital twins: from personalised medicine to precision public health. J Pers Med 2021 Jul 29;11(8):745. [doi: 10.3390/jpm11080745] [Medline: 34442389]

Abbreviations

DHT: digital health twin
JBI: Joanna Briggs Institute
PCC: population, concept, and context
PDT: personal digital twin
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
VR: virtual reality

Edited by A Quialheiro; submitted 13.03.25; peer-reviewed by K Blackmore, MK Milic; revised version received 07.08.25; accepted 29.09.25; published 28.10.25.

Please cite as:
Jabin MSR, Mirza A, Ilodibe A, Eldabi T, Vann Yaroson E
Ethical and Quality of Care–Related Challenges of Digital Health Twins in Care Settings for Older Adults: Scoping Review
JMIR Aging 2025;8:e73925
URL: https://aging.jmir.org/2025/1/e73925
doi:10.2196/73925

© Md Shafiqur Rahman Jabin, Ayesha Mirza, Adaobi Ilodibe, Tillal Eldabi, Emilia Vann Yaroson. Originally published in JMIR Aging (https://aging.jmir.org), 28.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using Wearable Sensors to Measure and Predict Personal Circadian Lighting Exposure in Nursing Home Residents: Model Development and Validation

Shevvaa Beiglary[1], MS; Yanxiao Feng[2], PhD; Nan Wang[3], PhD; Neda Ghaeili[1], MS; Ying-Ling Jao[4], PhD; Yo-Jen Liao[4], PhD; Yuxin Li[1], BAE; Julian Wang[1], EngD, PhD

[1]Department of Architectural Engineering, Pennsylvania State University, 556 White Course Drive, University Park, PA, United States

[2]School of Applied Engineering and Technology, New Jersey Institute of Technology, Newark, NJ, United States

[3]Department of Civil and Environmental Engineering, Northwestern University, Evanston, IL, United States

[4]Ross and Carol Nese College of Nursing, Pennsylvania State University, University Park, PA, United States

Corresponding Author:
Julian Wang, EngD, PhD
Department of Architectural Engineering, Pennsylvania State University, 556 White Course Drive, University Park, PA, United States

Abstract

Background: Lighting, especially circadian lighting, significantly affects people with dementia, influencing sleep patterns, daytime alertness, and behavioral symptoms such as agitation. Since individuals experience and respond to light differently, measuring personal lighting exposure is essential for understanding its impact on health. Without individual data, the connection between lighting and health outcomes remains unclear. Wearable sensors provide a practical way to track personal light exposure, helping researchers better assess its effects on circadian rhythms and overall well-being.

Objective: This study aims to develop and validate both calibration and predictive models using wearable lighting sensors to assess individual circadian lighting exposure accurately. By leveraging machine learning techniques and empirical data, we seek to establish a reliable method for health care researchers and practitioners to investigate and optimize lighting conditions for improved circadian health in nursing homes, especially for residents with dementia.

Methods: A combination of controlled laboratory experiments and on-site data collection was conducted using professional spectrophotometer measurements as ground truth. Calibration models were developed for photopic lux and correlated color temperature, while predictive models estimated circadian metrics such as circadian stimulus. The sensors and the developed models were implemented in a real-world health care research project about bright light therapy intervention at 2 assisted-living facilities.

Results: The calibration models for photopic lux and correlated color temperature demonstrated strong accuracy, with an adjusted R^2 of 0.858 and 0.982, respectively, ensuring reliable sensor measurements. Predictive models for circadian stimulus were developed using both simple regression and machine learning techniques. The random forest model outperformed linear regression, achieving an adjusted R^2 of 0.915 and a cross-validation R^2 of 0.857, demonstrating high generalization capability. Upon the implementation of these models, significant individual variations in circadian light exposure were found in the study, highlighting the significance of customized lighting evaluations. These results confirm the effectiveness of wearable sensors, combined with the developed calibration and predictive modeling, in accurately assessing personal circadian light exposure and supporting lighting-related health care research.

Conclusions: This study introduces an effective and scalable approach to circadian light assessment using wearable sensors and predictive modeling. By replacing labor-intensive and costly spectrometer measurements, the proposed methodology enables continuous, cost-effective monitoring in health care environments. However, challenges related to sensor wearability, durability, and user compliance were identified, underscoring the need for further sensor design refinements. Future research should focus on refining sensor integration, expanding case studies, and developing adaptive lighting interventions to enhance circadian health in vulnerable populations.

(JMIR Aging 2025;8:e72338) doi:10.2196/72338

KEYWORDS

circadian; lighting; wearable sensors; long-term care facilities; predictive modeling; personalized lighting assessment; dementia; Alzheimer disease

XSL•FO
RenderX

Introduction

Background of Circadian Lighting and Health

The human circadian system is a complex biological mechanism that regulates various physiological and behavioral processes, including sleep-wake cycles, hormone secretion, and cognitive function [1]. This system is primarily entrained by light-dark cycles, with exposure to light during the day and darkness at night being crucial for maintaining a healthy circadian rhythm [2,3]. Disruption of the circadian system, commonly known as circadian rhythm disruption, has been associated with a wide range of adverse health outcomes, including sleep disorders, metabolic disorders, cardiovascular diseases, and even certain types of cancer [1,4].

Lighting plays a crucial role in the entrainment of the circadian system. Exposure to light, particularly blue-enriched short-wavelength light, during the biological night can suppress the production of melatonin, a key hormone involved in the regulation of the sleep-wake cycle [5,6]. This suppression can lead to disruptions in the circadian rhythm and associated health problems. Conversely, exposure to appropriate lighting during the day can help maintain a healthy circadian rhythm and promote various physiological and cognitive benefits [7].

In the context of nursing homes, residents often experience disruptions in their circadian rhythms due to a variety of factors, including reduced exposure to natural light, irregular sleep-wake patterns, and the use of artificial lighting that may not be optimized for circadian entrainment [8,9]. These disruptions can have significant negative impacts on the health and well-being of nursing home residents, leading to an increased risk of sleep disorders, cognitive impairment, and other chronic health conditions [9]. Recent studies have further highlighted the importance of circadian-effective lighting interventions in improving sleep quality, circadian rhythms, and depression in older adults with dementia [10-12].

Moreover, researchers have emphasized the need for tailored lighting solutions that consider the specific needs and characteristics of the nursing home population. For example, Figueiro et al [13,14] have discussed the importance of circadian lighting for older adults with dementia, as well as explored the role of light and circadian rhythms in this population.

Personal Circadian Lighting Exposure

Assessing individual circadian lighting exposure has become increasingly vital in the health care sector, especially in long-term care facilities, where individuals' circadian rhythms can be profoundly affected by their lighting conditions. In recent years, there has been an increasing focus on the advancement of wearable sensor technologies capable of precisely and conveniently quantifying an individual's light exposure throughout the day. This is crucial for comprehending the effects of lighting on circadian rhythms, hormonal balance, and overall health and wellness.

The significance of precisely quantifying individual circadian lighting exposure in nursing homes is paramount. Multiple studies have underscored the significant impact of lighting on the physical and mental health of older adults, especially those with dementia. A study by Figueiro et al [15] demonstrated that customized lighting interventions in long-term care facilities resulted in notable enhancements in sleep, depression, and agitation among individuals with dementia. This highlights the necessity for precise and individualized lighting exposure data to guide the design and execution of these interventions.

Nonetheless, the existing spectrometer-based methods for quantifying individual circadian lighting exposure are frequently unfeasible for extensive application in nursing homes and other health care environments. These devices are generally cumbersome, expensive, and labor-intensive, rendering them inappropriate for prolonged monitoring and ongoing data collecting. Markvart et al [16] underscored the difficulties in calibrating and rectifying the output from wearable light exposure devices, stressing the necessity for more efficient but effective measuring technology.

Proposed Work

In this study, we developed predictive models to assess personal circadian lighting exposure using wearable lighting sensors. These sensors measured key lighting parameters, including correlated color temperature (CCT), photopic lux, and irradiance levels at the RGB channels. The calibration of these measurements was meticulously conducted through controlled laboratory experiments under various electrical lighting conditions, supplemented by on-site data collection using professional instruments to ensure accuracy. The predictive models, built upon these calibrated lighting parameters, were designed to predict circadian melanopic lux, melanopic equivalent daylight illuminance, and circadian stimulus (CS) values. To demonstrate the practical application of these wearable sensors and the developed models, we employed them in a bright light therapy project within 2 nursing homes. Here, we analyzed variations in circadian lighting exposure across different individuals, showcasing the models' effectiveness in capturing personal circadian light (CL) environments and their potential impact on health and well-being.

Scientific Contributions

This work makes significant scientific contributions by introducing calibration and predictive models built upon machine learning algorithms and rigorous datasets, which effectively replace the costly and complex processes of spectrometer measurement and circadian lighting calculations. These models enable accurate prediction of circadian melanopic lux and CS values via basic lighting measures, making long-term monitoring of personal circadian lighting exposure feasible with wearable sensors.

Furthermore, by showcasing the application of these models in real-world settings, the study highlights the significant variability in CL exposure among individuals, underscoring the importance of personalized lighting assessments.

Finally, by facilitating wearable technology for CL assessment, this study paves the way for more accessible and practical approaches to understanding and optimizing light environments for individual circadian health and well-being, particularly in contexts such as nursing homes where continuous monitoring is crucial.

XSL•FO
RenderX

Literature Review: Current Circadian Lighting Measurement Methods

Accurate measurement of CL exposure is essential for understanding the impact of lighting on the human circadian system and for developing effective lighting interventions. One of the primary instruments used for measuring CL exposure is the spectroradiometer, which is capable of providing detailed spectral information about the light environment [17,18].

Spectroradiometers are sophisticated instruments that can measure the spectral power distribution (SPD) of light, which is the relative energy of the different wavelengths of light present in a given environment. This information can be used to calculate various metrics related to CL exposure, such as the CL adjusted (CL_A) and the CS [19,20].

In addition to spectroradiometers, other instruments such as wearable light sensors and calibrated digital cameras have also been used to measure CL exposure [17,18,21]. These methods offer the advantage of being able to capture light exposure data over an extended period, providing a more comprehensive understanding of an individual's CL exposure patterns.

Predictive modeling approaches have also been developed to estimate CL exposure based on various environmental and personal factors. For example, models have been proposed that use information about the location, time of day, and weather conditions to predict the spectral composition and intensity of light in a given environment [7,22]. These models can be particularly useful in situations where direct measurement of light exposure is not feasible, such as in large-scale studies or in environments where spectroradiometers are not available.

In the context of nursing homes, the use of wearable sensors and predictive modeling techniques can be particularly valuable for assessing the CL exposure of residents. By combining these methods with the data obtained from spectroradiometers, researchers and health care professionals can gain a more comprehensive understanding of the lighting conditions in nursing homes and develop targeted interventions to optimize circadian entrainment and improve the health and well-being of residents [17,23].

Wearable Sensors for Circadian Lighting Measurements

Recent years have witnessed initiatives to create more compact, cost-effective, and user-friendly wearable sensor devices for tracking personal circadian lighting exposure. These devices seek to address the shortcomings of conventional spectrometer-based solutions, offering a practical and accessible method for evaluating an individual's light exposure throughout the day.

A primary difficulty in the development of wearable sensor systems is the compromise between measurement precision and user comfort. Figueiro et al [17] evaluated the efficacy of 3 practical field devices for measuring individual light exposures and activity levels, underscoring the necessity for more adaptable and precise wearable sensor technologies. The contemporary generation of spectrometer-based systems frequently compromises wearability because of their dimensions and measurement principles, in addition to being expensive and labor-intensive to operate.

Researchers are investigating alternate methods for measuring human circadian illumination exposure to tackle these problems. A recent work by Wang et al [24] introduced an innovative wearable sensor device that integrates a tiny light sensor with a microcontroller and wireless communication features, facilitating continuous and discreet monitoring of an individual's light exposure. This integrated sensor system has the potential to address the shortcomings of conventional spectrometer-based solutions, facilitating broader implementation in nursing homes and other health care environments.

Kubicki et al [25] created a wearable sensor platform that integrates numerous light sensors and a machine learning algorithm to assess circadian-effective light exposure, offering a more precise and user-friendly option than current systems. This advanced sensor technology, together with predictive modeling and data analysis, enables health care providers to comprehend the correlation between lighting and the health and well-being of nursing home residents more effectively.

Emerging sensor technologies may provide exciting opportunities to assist people with dementia and their carers in the context of an aging population. Lu et al [26] provided an in-depth examination of how wearable health devices can noninvasively measure vital signs, activity, and sleep in people with vulnerabilities, such as cognitive impairment. Such technologies offer the advantage of being passive (ie, no active engagement necessary on the part of the person with dementia) and are capable of continuous data collection. On this basis, Tiersen et al [27] used user-centered design to design smart home sensing solutions for dementia households, highlighting the need for personalized monitoring solutions that address the issues both patients and caregivers face in their day-to-day life.

Furthermore, Rose et al [28] addressed the emotional and psychological consequences of the introduction of such technologies and reported family caregivers' overall acceptability of using digital technology to record stress, but significant concerns were expressed about privacy invasion and technical complexity. Taking these concerns head-on, Berridge et al [29] created a new self-administered decision support tool to facilitate early advance planning for technology use in dementia care, allowing people in the early stages of cognitive impairment to continue to communicate their preferences about future monitoring and assistance technologies before declining to more advanced impairment. Together, these studies provide evidence that, albeit advantages of sensor technologies for monitoring and providing services to individuals with dementia, implementation should take into account user preferences, safeguarding the patient's privacy, and integrating sensors in existing care practices.

Methods

Ethical Considerations

The study was approved by the institutional review board at The Pennsylvania State University (hereinafter referred to as "Penn State") (STUDY00020216). Resident participants were

XSL•FO
RenderX

assessed for their capacity to provide informed consent using the evaluation to sign consent form. Informed consent was obtained from all residents who were deemed capable. For all residents, regardless of their ability to consent, consent was also obtained from their legally authorized representatives. Resident participants did not receive compensation. However, nursing home staff who contributed information related to residents' mood and behavioral assessments received a US $25 gift card per visit. All hard copy documents containing identifiable information were stored in locked, secure cabinets accessible only to institutional review board–approved research team members. Electronic research data were deidentified and stored on a secure research drive at Penn State, accessible only to authorized members of the research team.

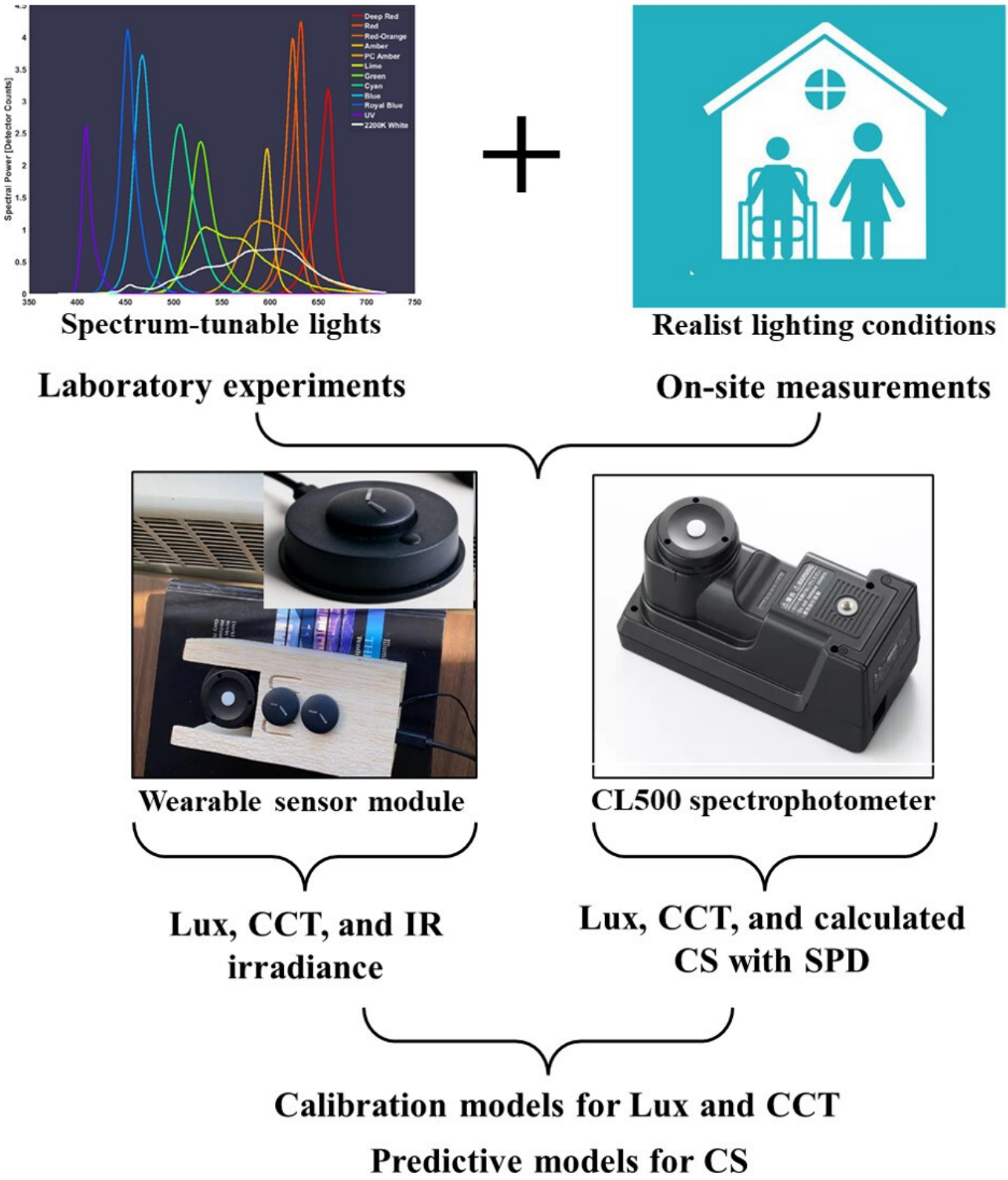
Experiments and Data Collection Procedure

To obtain comprehensive and sufficient datasets (Multimedia Appendix 1) for subsequent calibration and modeling, we conducted both laboratory experiments and field data collection in 2 assisted-living facilities. The laboratory experiments and data collection were completed at the Penn State Architectural Engineering Lighting Laboratory.

Using a professional spectrophotometer, we obtained ground truth data by placing 3 randomly selected wearable sensors and the spectrophotometer in the same position in the laboratory under various electrical lighting conditions. Moreover, to ensure the accuracy and reliability of using such models in assisted living facilities, we also conducted a similar procedure and setup in 2 selected nursing homes. In the laboratory, we simulated relatively more controllable lighting scenarios compared with the field, where we conducted several site visits on different days to capture variations related to weather, room conditions, and interior setup. The same set of instruments, wearable sensors, methods, and experimental setup was used for both laboratory and field measurements. Figure 1 shows the diagram of the major instruments used in this study. The spectrophotometer (CL500) output for photopic lux and CCT served as the reference for calibrating the wearable sensors through regression models. Additionally, the spectrophotometer's SPD data were used to calculate CL_A and CS, which then acted as ground truth data for building predictive models based on the wearable sensor data. Detailed information on the instruments, experimental setup, lighting conditions, and metrics is provided in the subsequent sections. It is of importance to note that both laboratory lighting scenarios and real-world lighting scenarios included daylighting, due to its direct impact on human biology as one of the main themes for our experiment.

XSL•FO
RenderX

Figure 1. Diagram for experiments and data collection. CCT: correlated color temperature; CS: circadian stimulus; IR: infrared; SPD: spectral power distribution.



Wearable Light Sensor

In this study, we adopted the LYS Button as one of our data collection tools. The LYS Button is incorporated with the LYS app, enabling users to collect data on individual daily CL exposure. The LYS wearable button is a compact and lightweight device and is ideal for individuals such as our human participants and assisted-living residents impacted by mild to moderate dementia, who tend to not respond well to large objects hanging off their clothing. The wearable sensor is an ideal tool when optimizing lighting environments. The LYS Button can track various parameters including photopic illuminance, color temperature, irradiance at R (465 nm), G (525 nm), B (615 nm), and infrared (IR) (1100 nm) channels, and motion. The key features of this wearable sensor are shown in Table 1. One of the drawbacks of the LYS Button is that it does not directly output circadian lighting levels, such as melanopic lux and CS. The availability of light data through the R, G, and B channels using the LYS device, which can be converted to melanopic

XSL•FO
RenderX

data, was of importance and value, hence why it was selected for this study.

Table . LYS Button: key features.

Feature	Specification
Illuminance measurement	0.1-100,000 lux
Color temperature measurement	1800-10,000 K
Continuous monitoring and data logging	Enabled
Connectivity	Bluetooth to mobile app
Battery life	Up to 7 days per charge
Design	Compact, discreet, and wearable
Companion mobile app	Data visualization and analysis

Spectrophotometer

CL500A Illuminance Spectrophotometer (Konica Minolta) was used In this study, as the tool to obtain the ground truth data for developing calibration and predictive models (Multimedia Appendix 2) from the wearable sensor data. It can measure illuminance, CCT, color-rendering index, and many other color or light parameters. Also, the instrument uses a silicon photodiode array and a diffraction grating to provide full SPD data, which can be then computed to obtain CL_A and CS. The other key features of CL500A are listed in Table 2.

Table . CL500A Spectroradiometer: key features.

Feature	Specification
Spectral measurement range	360-780 nm, 1-nm resolution
Illuminance measurement range	0.01-100,000 lx
Calculated parameters	Chromaticity coordinates, CCT[a], and CRI[b]
Display	Large, easy-to-read color liquid crystal display
Data transfer	USB interface to a computer
Design	Compact and lightweight, portable
Calibration	Traceable to national standards

[a]CCT: correlated color temperature.

[b]CRI: color-rendering index.

Laboratory Data Collection

Overview

We conducted our experiment at the Penn State University Architectural Engineering Lighting Laboratory, adhering to specified standards to ensure the precision and dependability of our results. Our methods conform to relevant international standards, including ISO 17025 for laboratory quality management. The laboratory is equipped with NIST-traceable luminaires, which allowed us to create a variety of electrical lighting conditions. Photopic lux levels ranged from 10 to 10,000, and CCT values spanned a broad range between 1500 to 10,500 K. The CCT distribution was achieved using 4 primary indoor light sources with CCT of 1500 K, 3000 K, 4200 K, and 6500 K. These sources were used individually or in combination, and—along with natural daylight entering through windows—they enabled coverage of the broad CCT range. Illuminance levels were similarly generated using these 4 dimmable light sources, allowing fine control over light intensity. Higher data density was intentionally designed and collected around 2 key illuminance ranges: one representing typical indoor artificial lighting and another corresponding to daylight-accessible conditions. As a result, we generated 100 different lighting scenarios by combining various lighting intensities and CCT conditions in the lighting laboratory. For each scenario, we waited approximately 10 minutes to ensure stability before logging 5 data points. The LYS Button wearable sensors automatically documented and uploaded the measured data to a central cloud, while the CL500A's data were transmitted directly to a computer via cable connection.

National Institute of Standards and Technology

NIST-traceable luminaires, which allowed us to create a variety of electrical lighting conditions. Photopic lux levels ranged from 10 to 10,000, and CCT values spanned a broad range between 1500 to 10,500 K. The CCT distribution was achieved using 4 primary indoor light sources with CCT of 1500 K, 3000 K, 4200 K, and 6500 K. These sources were used individually or in combination, and—along with natural daylight entering through windows—they enabled coverage of the broad CCT range. Illuminance levels were similarly generated using these 4 dimmable light sources, allowing fine control over light intensity. Higher data density was intentionally designed and collected around 2 key illuminance ranges: one representing typical indoor artificial lighting and another corresponding to

XSL•FO
RenderX

daylight-accessible conditions. As a result, we generated 100 different lighting scenarios by combining various lighting intensities and CCT conditions in the lighting laboratory. For each scenario, we waited approximately 10 minutes to ensure stability before logging 5 data points. The LYS Button wearable sensors automatically documented and uploaded the measured data to a central cloud, while the CL500A's data were transmitted directly to a computer via cable connection.

We affixed 3 randomly selected wearable sensors to the CL500A spectrophotometer at the same position within the laboratory, enabling the study team to obtain concurrent readings from both instruments. This method facilitated a direct comparison of the measurements and the detection of any inconsistencies between the 2 sensors (Figure 2). The CL500A's output for photopic lux and CCT served as reference points to calibrate the wearable sensors through regression models. Additionally, the CL500A's SPD data were used to calculate CS and melanopic lux, which then acted as ground truth data for building predictive models based on the wearable sensor data.

Figure 2. Data collection at Penn State University lighting laboratory.



On-Site Data Collection

The on-site data collection with the CL500A and LYS Buttons was executed in different residential rooms of 2 assisted-living facilities located in Pennsylvania. We intentionally visited the nursing homes under varying conditions, such as different weather, times of day, weekdays and weekends, and room types (resident rooms, hallway, dining hall, and activity rooms). Furthermore, in each room, the measurements were conducted at various distances from the windows, spanning from 0 to 9 feet, to evaluate the devices' reactions to varying amounts of daylight. The electrical lights were also modulated to form different daylight-electrical light combination scenarios. The window blinds were positioned at 4 distinct slat angles, 135°, 45°, and 90°, as well as the fully retracted position, to evaluate the effect of window shade on illumination conditions. The measurement stations were arranged at multiple viewing angles, ranging from 0°, directly facing the window, to 270° to assess their effectiveness across diverse orientations. The primary light sources in the 2 nursing homes consisted of 3 different CCT types: 1800 K, 2700 K, and 4000 K. Through controlled switching of these sources and adjustment of blinds to modulate daylight access, a broad range of illuminance and CCT variations was achieved. Similar to the laboratory settings, illuminance levels were primarily concentrated in lower ranges typical of dim indoor environments without daylight and in moderate ranges when daylight was present. CCT values were predominantly in the lower spectrum due to the warm indoor lighting but extended to higher levels when daylight contributed to the scene. All these steps and setups aimed to capture sufficient variations in lighting conditions from the window zone (Figure 3A) and the main bed area (Figure 3B) for subsequent calibration and prediction as displayed.

XSL•FO
RenderX

Figure 3. On-site data collection in the window area (A) and the living area (B) by authors at the assisted-living facility.



Circadian Metrics Used for Data Processing

CS, closely associated with CL_A, is a metric that directly quantifies the extent of circadian system activation resulting from specific light exposure [20]. Because neither spectrophotometers (CL500A) nor wearable sensors (LYS Button) can output CS directly, we had to extract the raw spectral data from CL500A and then input them into the CS calculator developed by Rensselaer Polytechnic Institute [19,30-32] as seen in equation 1. CS [33], shown in equation 2, is derived from the CL_A value.

.

(2)CS=0.7∗[1−11+(CLA355.7)1.1026]

Equation values and parameters are shown in Textbox 1.

Textbox 1. Equation values and parameters.

CL_A: circadian light adjusted.

E_λ: light source spectral irradiance.

Mc_λ: melanopsin sensitivity (corrected for crystalline lens transmittance) [34].

k=0.2616

S_λ: S-cone fundamental.

a_{b-y}=0.7

mp_λ: macular pigment transmittance.

a_{rod} = 3.3

V_λ: photopic luminous efficiency function [35].

RodSat = 6.5 Wm^{-2}

V_λ': scotopic luminous efficiency function [35].

XSL•FO
RenderX

Calibration and Predictive Modeling Methods

Calibration Modeling of Photopic Lux and CCT

The calibration process was implemented to ensure the measurement accuracy for the parameters of photopic lux and CCT from LYS. Simple regression calibration was applied first using the data obtained from the reference devices. The original data can be found in Multimedia Appendix 2. The candidate parameters from the reference devices included red, green, and blue (RGB) value, IR light, illuminance, and color temperature.

The photopic lux was calibrated based on the illuminance measurements from CL500A and the measurements from LYS about illuminance, RGB, IR, and CCT. The measurement accuracy of LYS was calibrated for its performance under different lighting conditions. Similarly, CCT from LYS was also calibrated based on the CCT measurements from CL500A. The data screening process was first conducted to check the negative values that were invalid and to enhance the data quality. The feature selection of these candidate parameters for the simple regression model was based on their significant test results. The models were developed and evaluated using a train-test split for the assessment of modeling metrics. In the train-test split method, 70% of the total data were used as a training set to train the model, and 30% were used as test data to value the model's performance. Evaluation metrics including R^2, mean squared error (MSE), and mean absolute error (MAE) were used to indicate the goodness of fit of the model. Equations 3-5 show the mathematical formula used for each of the implemented evaluation metrics.

$$(3)\ MAE=1N\sum i=1N|yi-y\hat{}|$$

$$(4)\ MSE=1N\sum i=1N(yi-y\hat{})2$$

$$(5)\ R2=1-\sum(yi-y\hat{})2\sum(yi-\bar{y})2$$

where $y\hat{}$ is the predicted value of y, and $\bar{y}$ is the mean.

Predictive Modeling of CS

The prediction models were developed based on the available parameters' information for the CS from CL500A. The original data can be found in Multimedia Appendix 2. The calculation of CS has been studied in diverse ways for quick and efficient computation. The LRC Circadian Light Combined Calculator was applied in the study in section "Circadian Metrics Used for Data Processing." The calculator was programmed to enable the output of CS by the input of SPD for a lighting source. The spectral lighting with the wavelength from 360 to 780 nm could be obtained from the CL500A and thus the values of CS were obtained using the CS calculator. The spectral lighting data with a 1-nm interval were first integrated into a 5-nm increment for an efficient computation in the calculator.

After the calculation of the CS, a CL prediction model was developed based on the measured values from LYS including RGB value, IR light, illuminance, and color temperature. The simple regression model was applied first based on the train-test split to check the prediction accuracy using evaluation metrics of R2, MAE, and MSE. To better evaluate the reliability of a prediction model, especially for the data collected under different lighting environments, the 5-fold cross-validation shown in equation 6 was further used to identify the optimal simple regression model.

$$(6)\ CV(k)=1k\sum i=1nerrori$$

where the error can refer to evaluation metrics such as MAE, MSE, and R^2.

To achieve high and satisfactory prediction accuracy, a random forest model was used for predictive analysis if the simple regression model did not meet the desired performance criteria. A random forest model enabled a multitude of decision trees during training and output the mean prediction for regression of the individual trees. This bagging technique, where different subsets of the dataset are used to train each decision tree, made the overall model more robust and less prone to overfitting. Compared with a linear regression, it could provide a comparison of the feature importance which can contribute to the understanding of the influence of the input variables on the output.

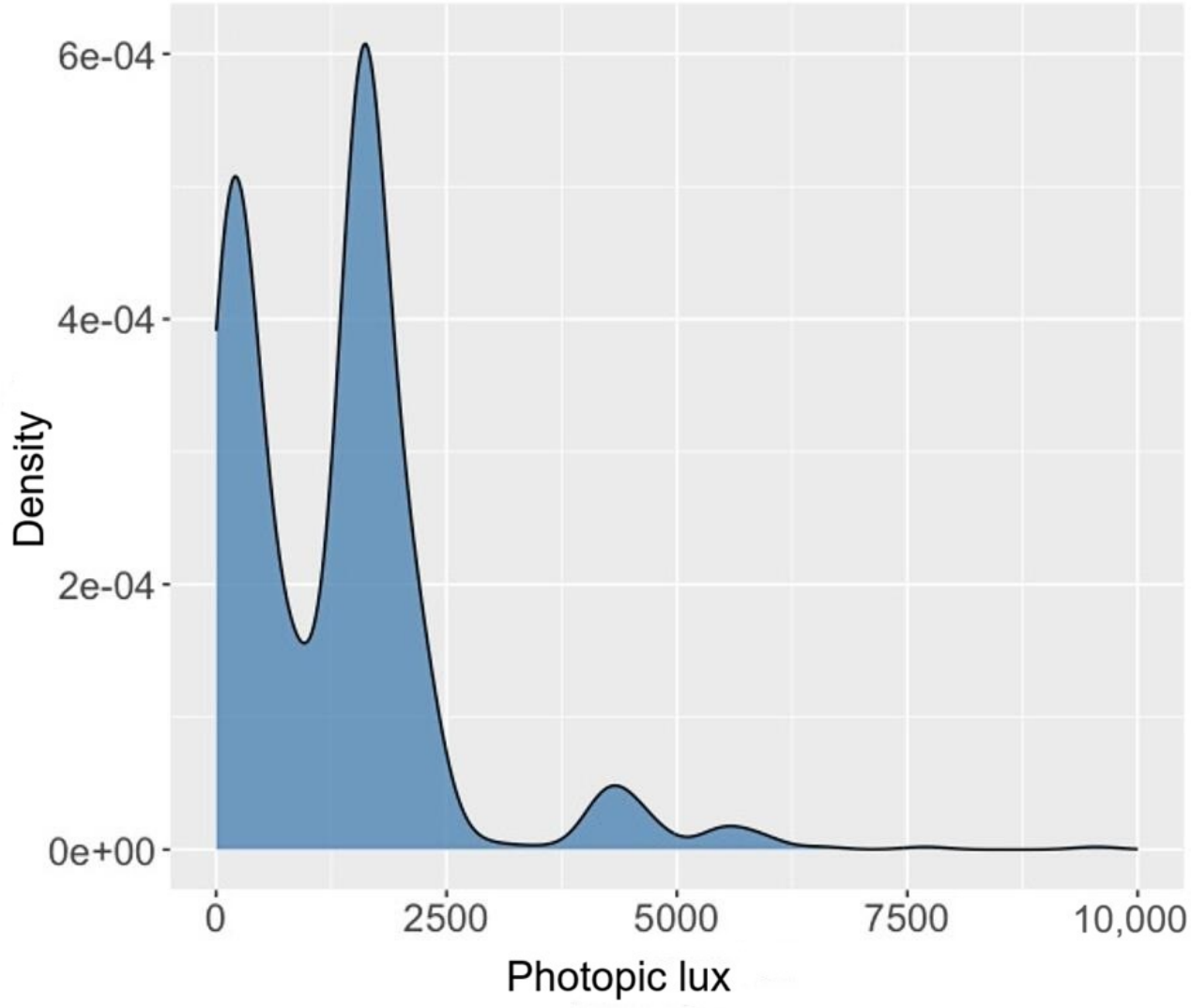
Results

Calibration Model for Photopic Lux

Figure 4 shows the distribution of photopic lux measurements collected from our experiments at both the laboratory and on-site. The data covered a range from ultra-low to high lighting levels, corresponding to 2 main locations. The first peak (at lower lux values) likely represented interior spaces illuminated by artificial lighting, while the second peak (at higher lux values) corresponded to areas near windows with natural daylight exposure.

XSL•FO
RenderX

Figure 4. Photopic lux distribution during data measurement.



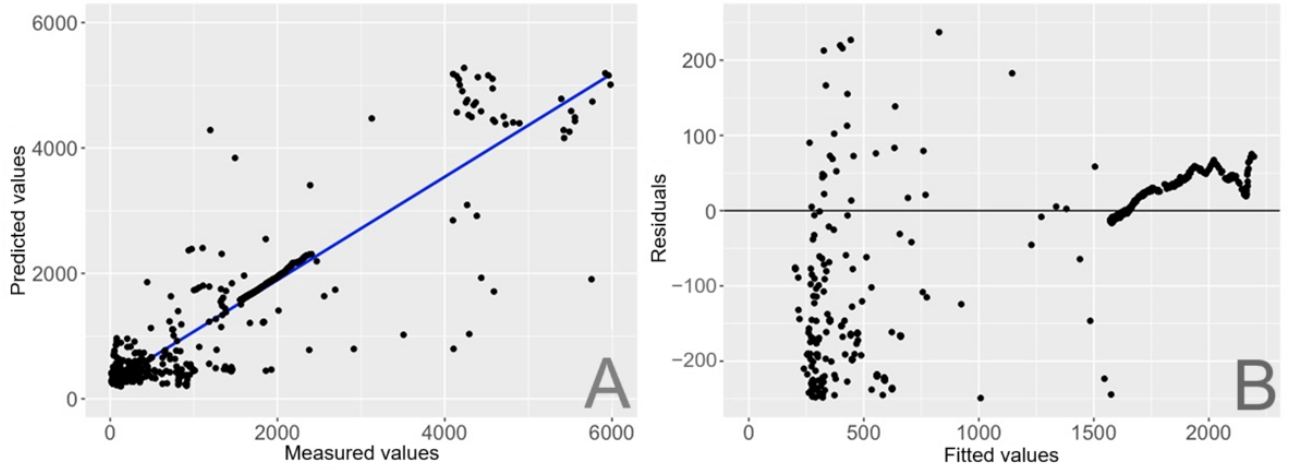
For the photopic lux calibration, a simple regression model with acceptable accuracy was obtained. The input parameters, including the LYS Button's output, which include CCT (Kelvin), IR channel irradiance, and illuminance (lux), were selected during the modeling analysis. The calibration function is shown in equation 7. This calibration function adjusts the original measurements from the LYS Button (CCT', IR, and lux') to more accurately reflect true photopic lux (Y). The coefficients (0.0596, 1.6064, and 0·8432) represent the contribution of each parameter to the final calibrated value.

$$(7)\ Y(\text{Photopiclux}) = 0.0596 * CCT' + 1.6064 * IR + 0.8432 * \text{lux}$$

Where CCT', IR, and lux' refer to the original output from the LYS Button—CCT (Kelvin), irradiance of near-IR wavebands at the peak of 1100 nm (W/m²), and illuminance level (lux), respectively—the comma differentiates the final calibrated parameters, CCT, and lux. This may be due to small changes in solar light or daylight, which can significantly boost the lux measurement, as near-IR sources are primarily from daylight. In contrast, the original lux' measurement has a smaller coefficient, suggesting that it was initially less accurate and thus required a substantial adjustment. The CCT' shows a minor but necessary adjustment. Together, these coefficients correct the original sensor outputs to better match the true photopic lux values. The overall adjusted R^2 and MAE for this calibration model are 0.858 and 355, respectively. The goodness-of-fit plot is shown in Figure 5 (A), indicating that the predicted values generally correspond to the measured values. Plot (B) in Figure 5 shows the residuals against the fitted values, in which the residuals are widely dispersed especially at lower fitted values. However, at the higher fitted value range, there is a clear pattern, possibly suggesting model misspecification on the presence of nonlinear relationships. However, for general indoor lighting conditions, the calibration model shows a general trend of aligning with the measured values.

XSL•FO
RenderX

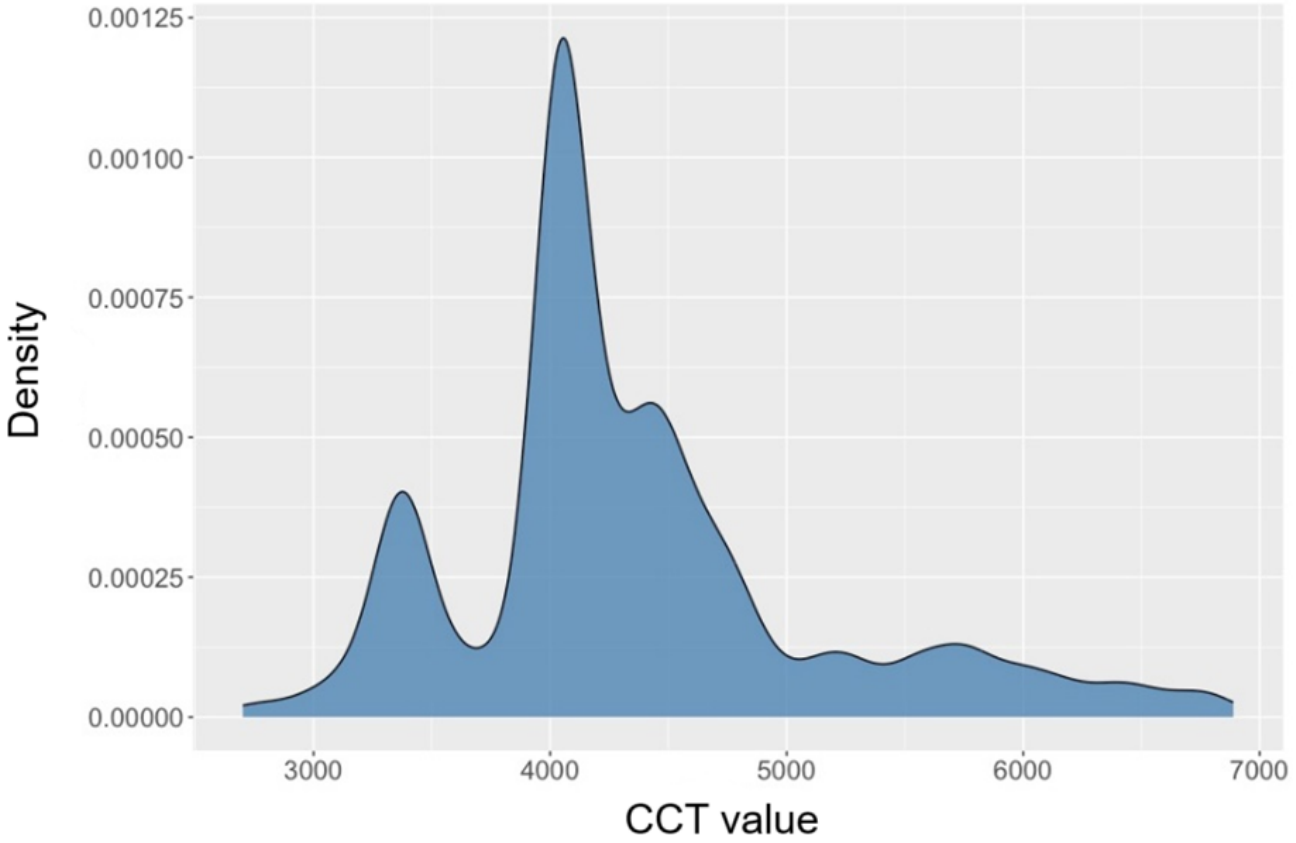
Figure 5. Calibration results for photopic lux: (A) Predicted versus measured values. (B) Residuals versus fitted values.



Calibration Model for CCT

Figure 6 illustrates the distribution of CCT values recorded during the data collection procedure. Similar to the distribution of photopic lux, it represents 2 key environments. The initial peak (around 3000 - 4000 K) likely corresponds to artificial interior lighting with a warmer color temperature, while the subsequent distribution (around 4000 - 5000 K) may represent areas near windows with daylight, featuring a cooler color temperature.

Figure 6. Correlated color temperature distribution during data measurement. CCT: correlated color temperature.



Similar to the calibration model of photopic lux, for the CCT calibration, a simple regression model with satisfactory accuracy was also obtained. The input parameters, including the LYS Button's output—CCT (Kelvin), IR channel irradiance, and illuminance (Lux)—were selected again during the CCT calibration model development. The calibration equation is shown in equation 8:

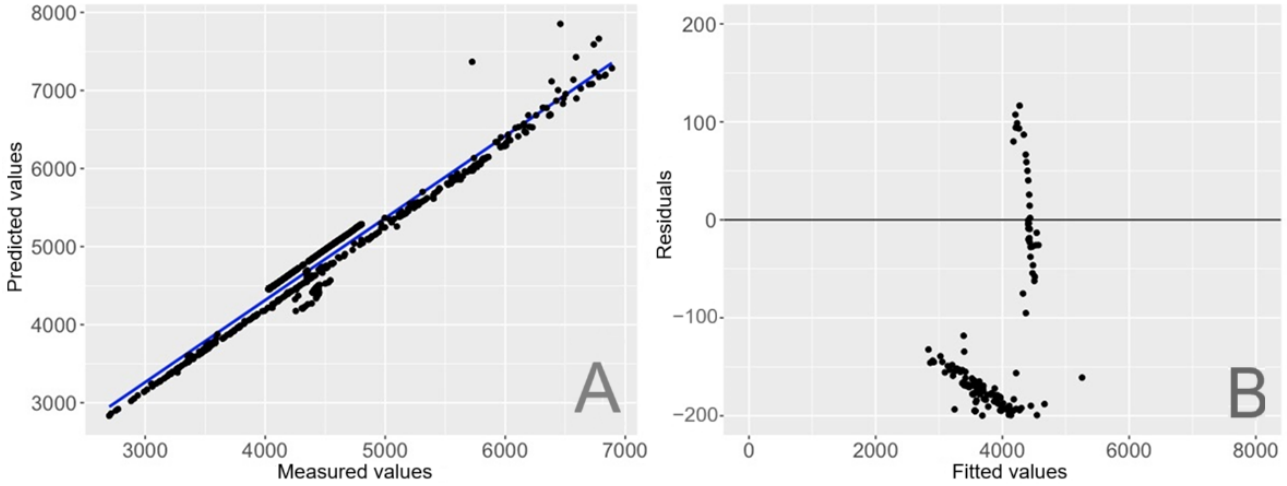
$$(8)\quad Y(CCT)=1.048\times CCT'-0.5427\times IR+0.1761\times lux'$$

In the calibration model, the Kelvin component, with a coefficient of 1.025, shows a strong correlation with the true CCT, indicating that the original CCT values are quite close but require a slight adjustment. The IR component has a relatively strong negative influence on the CCT, highlighting the significant impact of solar IR on the calibration process, likely due to its effect on perceived color temperature, particularly from daylight sources. Together, these coefficients adjust the original sensor outputs to better match the true CCT

XSL•FO
RenderX

values. The overall adjusted R^2 and MAE for this calibration model are 0.982 and 342, respectively. The goodness-of-fit plot is shown in Figure 7A, and the model bias and variance patterns in Figure 7B.

Figure 7. Calibration for correlated color temperature. (A) Predicted versus measured values. (B) Residuals versus fitted values.

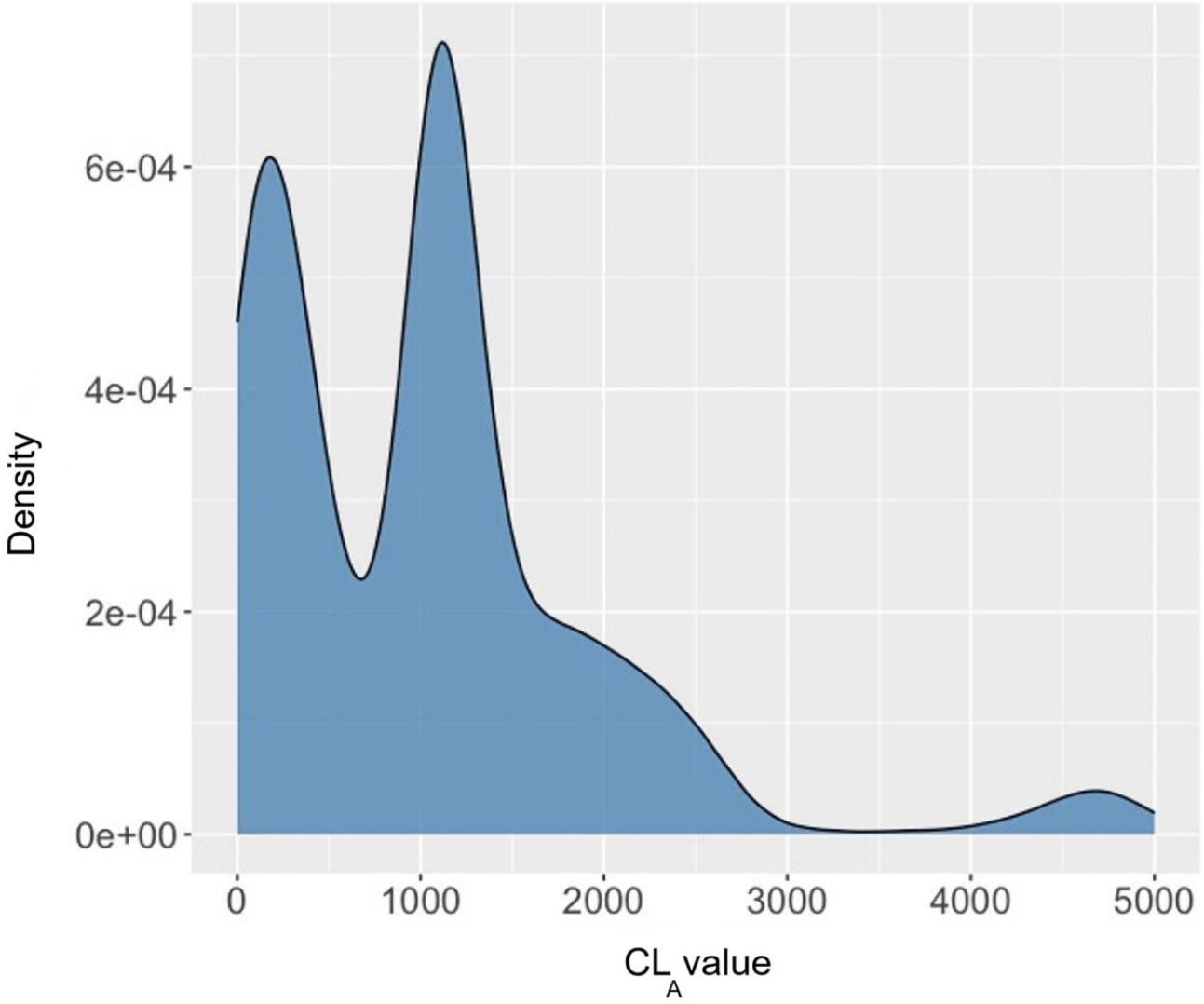


Evaluation of Predictive Models for CL_A

Figure 8 represents the distribution of CL_A values collected, which also reflects 2 major lighting conditions. The first peak (at lower CL_A values) likely corresponds to areas with artificial lighting, providing minimal circadian stimulation. The second peak (at higher CL_A values) likely represents zones near windows with natural daylight, which contribute more significantly to CL_A exposure. The remaining distribution tail illustrates sporadic occurrences of very high CL_A values, likely in brightly lit areas with substantial daylight, which is also aligned with the photopic lux data distribution.

XSL•FO
RenderX

Figure 8. Circadian light adjusted (CLA) distribution during data measurement.



Based on the modeling procedure illustrated in section "Laboratory Data Collection," we obtained both simple regression models and machine learning–based models. As shown in Table 3, the machine learning–based model shows much more accurate performance relative to the linear regression model, with an adjusted R^2 of 0.86.

Regarding the machine learning models, we also conducted the importance feature ranking analysis. Figure 9 shows the variable importance ranking results using 2 metrics %IncMSE (percentage increase in MSE) and IncNodePurity (increase in Node Purity). %IncMSE measures the increase in the model's prediction error when a particular variable is randomly permuted while all others are kept constant. A higher value indicates greater importance, while IncNodePurity is based on the total decrease in node impurities contributed by a variable across all the trees in the forest. As shown in this context, the IR value has the highest in both metrics consistently, which is similar to the calibration model results in CCT and photopic Lux calibration—the potential daylight effects on indoor IR levels drive the variations of circadian lighting level changes. The sensor's original CCT value had a strong direct impact on the model's predictive accuracy, while it might not be crucial for creating pure or homogenous nodes, meaning it did not contribute much to the structure of the decision trees.

Table . Model comparison.

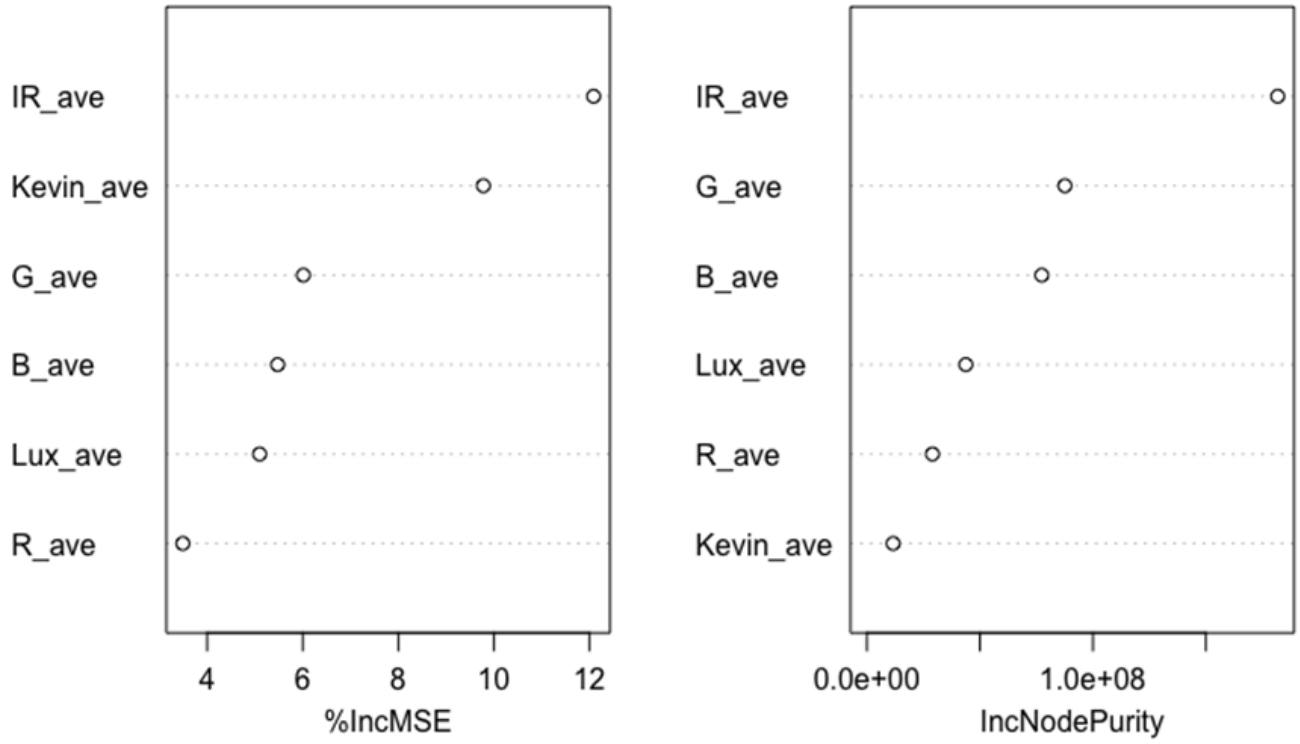
Modeling techniques	Performance	MAE[a]	RMSE[b]
Random forest	0.8567 (cross-validation R^2)	180.53	645
Linear regression	0.645 (adjusted R^2)	491.5	1050

[a]MAE: mean absolute error.

[b]RMSE: root-mean-square error.

XSL•FO
RenderX

Figure 9. Variable importance.



Applications in Circadian Lighting Health Research in Nursing Homes

To understand the functionality and usage of the wearable sensors with calibrated and predicted models, we applied the sensors to 29 older adults with dementia in 2 assisted-living facilities as part of a National Institutes of Health–supported project. A post hoc power analysis using GPower 3.1 [36] indicated that with a total sample size of 29 and α=.05, the study had 87 % power [37] to detect a medium effect size (Cohen *d*=0.53), as reported in the previous studies [14,15]. During the personalized data collection from residents, the LYS device was pendant-worn or clipped on the upper body area (eg, upper chest, shirt collar, and shoulder) to more closely approximate eye light exposure. Staff and nurses have been guided to make sure that the attachment of devices is in secure positions as well as constant monitoring of the residents to ensure that the device has not been accidentally removed from the human participant. Researchers also visited sites twice a week to ensure the sensor attachment. Additionally, the LYS device supports real-time remote monitoring. We implement automated algorithms on the cloud to detect potential nonwear or malfunction events (eg, no change in light and motion data readings for more than 2 hours during the daytime and 4 hours during the nighttime). When such events were identified, alerts would be sent to the research team who would follow up with on-site verification and resolution as needed.

The sensor data for each individual was automatically uploaded to a central computer (Mac Mini) located in the dining hall of the nursing homes to ensure Bluetooth connectivity to all participant residents and then retrieved through a remote connection. By using the calibrated and predicted sensor measurement methods developed in this work, we were able to analyze each resident's circadian lighting exposure, enabling more rigorous and effective intervention analysis.

We then selected specific residents with the widest range of variabilities and differences to demonstrate the applicability of our developed methods within the following sections. It is important to note that any analysis shown in this section is of exploratory value; the relatively small focus of the study in this section was to illustrate the measurement methodology for potential use in future studies on circadian lighting health.

Intervariability of Residents' Lighting Exposures

Understanding the intervariability of residents' lighting exposures is crucial for accurately analyzing the effects of interventions, as it accounts for the differences in physical environments and personal factors among individuals. These 2 residents were particularly selected due to their differences in the dementia stages, the room's window orientation direction, and daily routine and schedules. To do this, we particularly extracted 2 residents' hourly sensor data and then calculated weekly averages of photopic lux, CCT, and CS. Resident 1's room faced south, and resident 2's faced east. Resident 1 had mild dementia, and resident 2 had moderate dementia. It was also observed that resident 1 benefited from the facility's activities such as spending several hours during the day in the sun-lit dining area and having scheduled outdoor visits more frequently than resident 2. Table 4 denotes key differences between the 2 residents selected for detailed analysis in this study, providing context for the observed variations in CS, lux, and CCT measurements in the subsequent sections.

Figures 10-12 are the detailed comparisons in terms of their variations in CS, Lux, and CCT based on our calibrated and predicted measurement models, using a total of 268 data points collected across 5 project weeks and processed using the machine learning and regression models in section "Calibration

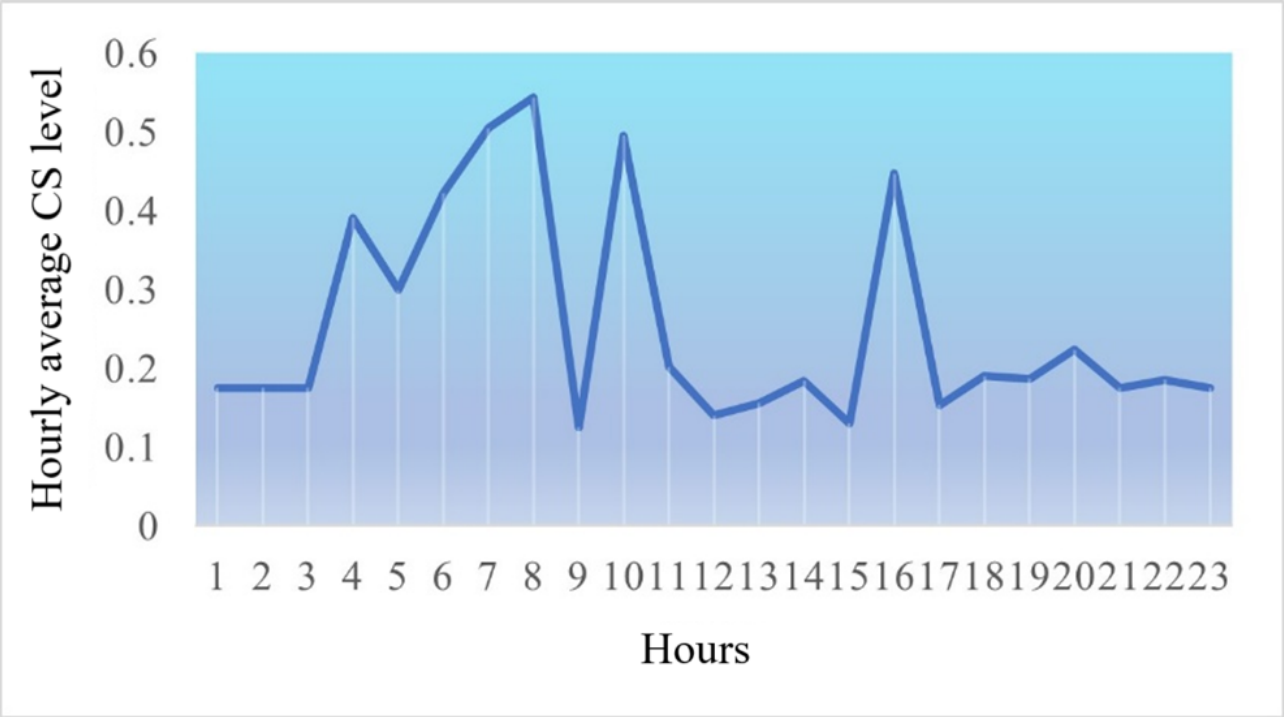
XSL•FO
RenderX

and Predictive Modeling Methods" in R Studio [Posit] and visualized using Python features (the main coding information used in this work can be found in Multimedia Appendix 1).

Table . Comparative characteristics of selected study residents.

Characteristic	Resident 1	Resident 2
Dementia stage	Mild	Moderate
Room orientation	South-facing	East-facing
Activity level	High; regular participation	Limited; infrequent participation
Outdoor access	Frequent scheduled visits	Minimal exposure
Daily routine	Structured; time in sun-lit areas	Less structured; more time in room
Sleep pattern	Generally stable	Erratic; disrupted
Light exposure	Consistent throughout day	Morning-dominant
Daytime behavior	Alert; minimal sundowning	Variable alertness; evening agitation

Figure 10. Hourly circadian stimulus intravariability. CS: circadian stimulus.



XSL•FO
RenderX

Figure 11. Hourly illuminance intravariability.

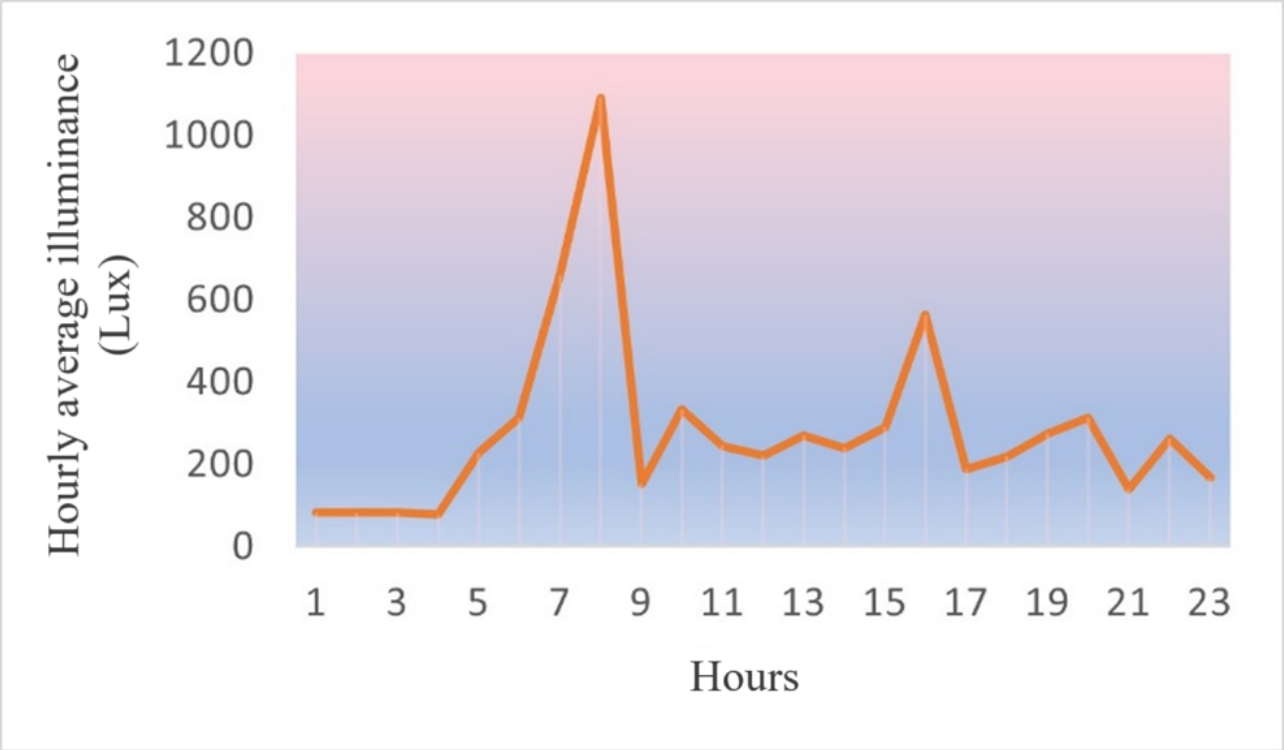
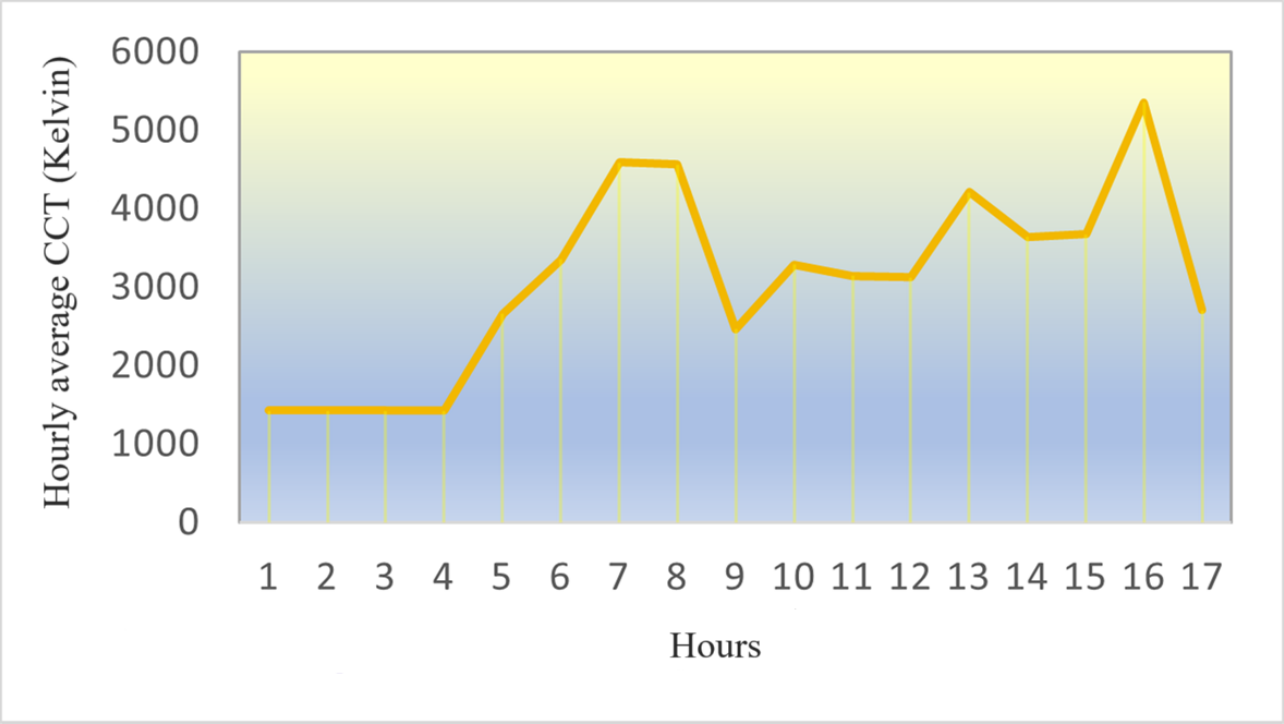


Figure 12. Hourly daytime correlated color temperature variations. CCT: correlated color temperature.



CS Intervariability

Multimedia Appendix 3 visualizes the weekly daytime and nighttime CS value variations of the selected 2 residents across the project. When comparing nighttime CS levels between resident 1 (mean 0.196, variance 0.00223) and resident 2 (mean 0.228, variance 0.00167), the paired *t* test analysis revealed no statistically significant difference between the 2 residents (2-tailed t^4=−1.01; *P*=.368). Although resident 2 showed a slightly higher mean nighttime CS value, this difference was not substantial enough to reach statistical significance at the α level of .05. Analysis of daytime CS levels revealed that resident 1 (mean 0.414, variance 0.00418) demonstrated marginally higher values than resident 2 (mean 0.326, variance 0.00113). This difference approached statistical significance in the paired *t* test (2-tailed t^4=2.43; *P*=.072). While not meeting the conventional *P*<.05 threshold for statistical significance, this trend suggested that resident 1 potentially experienced greater

XSL•FO
RenderX

circadian-effective light exposure during daytime hours than resident 2.

Both residents have higher CS levels during the daytime compared with nighttime throughout the period. Comparatively, across those weeks, resident 1 with moderate dementia generally had higher daytime CS levels and lower nighttime CS levels than resident 2 in most weeks, with the highest average daytime CS of 0.5 for resident 1 during week 5 and 0.38 for resident 2 during week 2, and the highest average nighttime CS of 0.29 for resident 2 during week 5 and 0.24 for resident 1 during week 4. The differences in both daytime and nighttime CS levels between the 2 residents can be attributed to several factors, in particular their different dementia stages, different daytime activity participation levels, and bedroom orientations as pointed out in section "Intervariability of Residents' Lighting Exposures."

Resident 2 with moderate dementia (reported erratic nighttime behaviors and activities) living in an east-facing room and lower daytime CS levels had less participation in the facility's activities and as a result had reduced daylight exposure. As expected, resident 2 displayed higher nighttime CS due to disrupted sleep. Their east-facing orientation provided less consistent and intense daylight than the south-facing room of resident 1, impacting the overall daytime and nighttime CS-level variations between the two.

Lux Intervariability

With regard to weekly daytime and nighttime illuminance level comparisons, there were notable differences between the 2 residents. For nighttime illumination levels, statistical analysis revealed a significant difference between resident 1 (mean 141.58 lux, variance 2558.22) and resident 2 (mean 233.58 lux, variance 4042.12), with the paired *t* test showing statistical significance (2-tailed t^4=−2.78; *P*=.050). This finding indicates that resident 2 experienced significantly higher illumination levels during nighttime hours than resident 1, which aligns with observations of resident 2's reported sleep disturbances and increased nighttime activity. For daytime illumination levels, data analysis comparing resident 1 (mean 422.24 lux, variance 4438.70) and resident 2 (mean 376.36 lux, variance 493.39) showed no statistically significant difference (2-tailed t^4=1.32; *P*=.258). While resident 1 experienced somewhat higher average daytime illumination levels, the considerable variance in resident 1's measurements and the relatively small sample size likely contributed to the nonsignificant finding. This suggests that both residents received comparable amounts of light exposure during daytime hours despite their different room orientations and activity patterns.

As seen in Multimedia Appendix 3, resident 1, who had a milder dementia stage than resident 2 and resided in a south-facing bedroom, was exposed to a sufficient amount of daytime light exposure. This abundance of natural light not only supported resident 1's circadian rhythms, as discussed in section "Intervariability of Residents' Lighting Exposures," but also had the potential to positively impact their mood, alertness, and overall well-being according to the observations. In contrast, resident 2, who had a moderate dementia stage and an east-facing bedroom, experienced slightly lower daytime lux levels than resident 1. However, it is important to note that resident 2's room would still receive natural light exposure, particularly during the early morning hours when the direct sunrise exposure provides a high level of illumination, as both residents ultimately had sufficient average amounts of lux close to 400 per recommendations and standards.

It is worth mentioning that the differences in daytime lux levels between the 2 residents can also be attributed to their varying levels of participation in daily activities and programs. Resident 1, with a lower dementia stage, was likely more engaged in general activities scheduled by the facility, which made a further contribution to their increased exposure to daytime light, due to the activity area being sun-lit from morning to afternoon. In contrast, resident 2, with a moderate dementia stage, may have had reduced participation in activities and programs, leading to slightly lower daytime lux levels. Resident 2 also had significantly higher nighttime illuminance exposure due to erratic sleep patterns.

CCT Intervariability: Accuracy Challenges of CCT Measurements in Dim Conditions

A primary obstacle in quantifying CCT in low-light conditions is the intrinsic difficulty in precisely measuring the SPD of the light source. Rea and Freyssinier [38] emphasize that under low-light situations, the spectral sensitivity of the human eye alters, resulting in possible disparities between the measured CCT and the perceived CCT. This may lead to mistakes when using conventional colorimetric systems, which are calibrated for elevated light levels. Chen et al [39] examined the efficacy of different CCT measurement methods in low-light conditions, showcasing that these challenges are especially evident in low-light conditions, where the signal-to-noise ratio of the measurement may be considerably diminished. Li et al [40] assert that conventional CCT calculation methods may be inadequate for these light sources, as they can display intricate SPDs that may result in erroneous CCT estimations. Owing to these constraints, we exclusively examined the daytime CCT to align more closely with reality.

Analysis of nighttime CCT levels revealed a notable difference between resident 1 (mean 2098.10 K, variance 515642.97) and resident 2 (mean 3733.34 K, variance 1760014.09), with results approaching statistical significance (2-tailed t^4=−2.26; *P*=.086). This trend suggests that resident 2 was exposed to lighting with a substantially cooler color temperature during nighttime hours compared with resident 1, which may help explain some of the observed differences in nighttime behavior and sleep quality, as cooler (bluer) light is known to potentially disrupt circadian rhythms and melatonin production during evening hours. Examination of daytime CCT values revealed no statistically significant difference between resident 1 (mean 4219.44 K, variance 226319.65) and resident 2 (mean 3880.88 K, variance 520702.89) according to the paired *t* test analysis (2-tailed t^4=0.81; *P*=.465). While resident 1 experienced slightly higher color temperature exposure during daytime hours, suggesting some exposure to cooler, more bluish light that typically enhances alertness, the substantial variance in measurements and small sample size resulted in a nonsignificant finding. This

XSL•FO
RenderX

indicates that both residents generally experienced comparable light spectral qualities during daytime, despite differences in room orientation and activity patterns.

With regard to the CCT variations seen in Multimedia Appendix 3, resident 1 tended to experience a slightly higher daytime CCT than resident 2. This can be primarily attributed to the consistent natural light exposures that resident 1 receives for a longer duration throughout the day, due to the south-facing orientation of their room. The increased exposure to this cooler, bluer-toned light can have beneficial effects on the resident's circadian rhythms, cognitive function, and overall well-being.

In contrast, resident 2's east-facing room receives less direct sunlight exposure throughout the day, leading to a relatively lower daytime weekly average CCT. While resident 2's daytime CCT levels are lower than those of resident 1, it is important to note that the early morning sunlight coming from the east, even with a lower CCT, can still help entrain the resident's sleep-wake cycle and promote better synchronization with the natural light-dark cycle. It is seen that resident 1's higher average levels of CCT are also partly due to their participation in outdoor activities, while resident 2 did not participate as frequently.

In sum, the analysis of lighting exposure intervariability among residents, using the regression and machine learning models developed in this study, may provide help to identify moderating factors that may influence the effectiveness of lighting interventions across different residents.

Intravariability of Residents' Lighting Exposures

Examining the intravariability of each resident's light exposure provides insights into the daily fluctuations and consistency of their lighting environments. Especially for circadian lighting intervention studies, this understanding allows researchers to understand individual characteristics and adjust lighting interventions to ensure that target conditions are maintained. From a long-term application perspective, this can also enable lighting system adjustments and personalization to maintain optimal circadian lighting exposure. As such, we particularly select 1 resident who exhibits a mild stage of dementia, with intermittent reports of midnight roaming behaviors.

The hourly data shown in Figures 10-12 were obtained from this resident on 1 weekend day in October with partly cloudy sky conditions. Their room was oriented toward the east, which allows for direct sunshine exposure from sunrise until approximately 9 AM. During the afternoon hours, this resident tended to spend more time in their bedroom, where they received minimal sunshine exposure. In particular, this specific room orientation and the resident's daily routine help explain the observed surge in CS, lux, and CCT measurements during the morning hours.

This resident's CS levels can be explained by their room orientation (east-facing window) as well as their daily schedule on the specific day of data collection. During the lunch and early dinner hours (before 11 AM and around 4 PM), the resident was seated in the dining room, which has large windows and receives abundant natural light throughout the day (Figure 10). Throughout the night, they had low CCT levels, which points to their uninterrupted sleep that evening.

As the sun rises, it shines directly through this resident's east-facing room, contributing to a high level of illumination. Later in the day, around 4 PM, the resident was exposed to the direct, lower-angle rays of the setting sun on the western side of the building (Figure 11). During the evening, illuminance remained at a low level and it was in accordance with the observation of their normal sleep.

The daytime CCT for this resident seems to plummet between 4 PM and 5 PM. This occurrence was likely due to the resident stepping outside briefly scheduled by their daily routine and then going back to the interior where there were much lower CCT conditions (Figure 12). Nighttime CCT data were not processed due to the constraints mentioned in section "CCT Intervariability."

The detailed analysis of this resident's lighting exposure and daily activities highlights the significant role of factors such as room orientation, daily routines, and engagement in facility programs in shaping their overall circadian rhythms and well-being. The data collected showcased the potential benefits of optimizing lighting and daily schedules for residents with mild dementia to promote better engagement and awareness, and the data processing with the models we developed in this study provided insights into the relationship between their circadian lighting exposure and engagement in daily activities. These results can help health care providers, nurses, and families of older adults with dementia to ensure that they are making the necessary lighting and routine adjustments for these residents.

Summary of Application Findings

The application here demonstrated the effectiveness of wearable sensors, combined with calibration and predictive models, in assessing individual circadian lighting exposure. By capturing personal light exposure data, these models enabled a detailed analysis of how factors such as room orientation, daily routines, and activity engagement influence circadian rhythms. This approach highlights the significant variability in lighting experiences among individuals, which traditional group-level assessments may overlook.

The small focus of the study was to validate the measurement methodology for use in future studies of health impacts. We recognize that additional validated research will be needed to perform robust effect analyses in larger and more diverse populations and to determine the broader impact on health. Such efforts can greatly enhance circadian lighting research, enabling more targeted interventions and improving data-driven decision-making in health care environments.

Discussion

Principal Findings

Despite its contributions, this study encountered several challenges related to sensor wearability, durability, and data collection. In the application projects, some dementia residents found the sensors uncomfortable, reducing adherence, while

XSL•FO
RenderX

others unintentionally removed or misplaced them, leading to a 40% loss or damage rate in that project. This highlights the need for more discreet, lightweight, and tamper-resistant sensor designs that seamlessly integrate into daily wear. Data collection also faced limitations, as site-specific environmental conditions and seasonal variations may have influenced results. The study's reliance on a single type of sensor raises concerns about generalizability, particularly in RGB and IR channel sensitivity. Future research should include multisite, long-term studies across diverse lighting conditions and validate models using multiple sensor types to improve accuracy and applicability.

To recapitulate, although the use of LYS Button wearable sensors enabled us to collect continuous personal CL exposure data, factors such as sensor placement variability, individual characteristics, and environmental complexities inherent in nursing homes might affect the generalizability of the developed measurement method. We stress that these issues are common when pioneering new methodologies in real-world settings. By advancing sensor design, expanding case studies, and integrating adaptive lighting solutions, future research can enhance the scalability and effectiveness of personalized circadian lighting assessments. These improvements will support more precise data-driven health care interventions, ultimately contributing to better sleep, cognitive function, and overall well-being in long-term care residents.

Conclusions and Future Work

Using wearable lighting sensors, we developed and validated calibration and predictive models to evaluate individual lighting exposure in terms of photopic lux, CCT, and circadian lighting levels. Controlled laboratory studies and on-site data collection were conducted to calibrate sensor values, as follows: To adjust the raw data from the LYS Button and improve its accuracy, we developed a simple regression calibration using key sensor outputs—CCT (CCT in Kelvin), IR channel irradiance, and raw illuminance (lux). The regression analysis generated coefficients that reflect each parameter's contribution to estimating true photopic lux, with the IR measurement showing a particularly strong influence likely due to its sensitivity to daylight variations. Overall, the calibration model, with an adjusted R^2 of 0.858 and with reasonable MAE, shows that although some deviations remain (especially under low light conditions), it reliably aligns the sensor's outputs with the measured lux values in general indoor lighting scenarios.

As with photopic lux calibration, a simple regression model was developed for CCT using the raw outputs (CCT, IR, and lux) from the LYS Button. The model adjusted the raw CCT values by attributing a strong positive weight to the original CCT output and incorporating the effects of IR and lux, which helped account for variations due to daylight influence. With an adjusted R^2 of 0.982 and an MAE of 342, the calibration effectively corrects the sensor outputs to match the true CCT observed in various lighting environments.

Predictive models were then developed to forecast CS levels based on the calibrated lighting factors, with the random forest model outperforming linear regression, achieving an adjusted R^2 (test) of 0.915 and a cross-validation R^2 of 0.857. The IR value was identified as the most influential factor in predicting circadian lighting levels, further reinforcing the reliability of the model. These results demonstrate the robustness of data-driven modeling for CL assessment, providing a highly accurate and scalable method for personal circadian exposure evaluation.

We then implemented wearable sensors and developed models in a bright light therapy project in 2 assisted-living facilities to demonstrate how these models can be used practically. These field applications in 2 assisted-living facilities revealed notable interindividual variations in CL exposure, with room orientation, daily routines, and engagement in facility activities playing key roles. For instance, residents with higher daytime light exposure, particularly from natural sources, demonstrated more stable circadian patterns, whereas those with limited exposure exhibited signs of disrupted rhythms. Similarly, the residents who have increased engagement in everyday routines, and activities, and availability to steady daylight in the south-facing orientation also had relatively high and stable circadian lighting exposure. When assessed individually only, a specific resident with mild dementia, an east-facing room, and reported nighttime wandering were selected for study on 1 weekend day for 24 hours, and their lighting exposure data were very much in accordance with their behavior and activity patterns within that period. These findings underscore the importance of the lighting assessments at the individual resident level in investigating and optimizing environmental conditions for residents with dementia and related health conditions.

Acknowledgments

Research reported in this work was supported by the National Institutes of Health/National Institute of Aging (number 1R21AG078544-01: The Effect of Smart Ambient Bright Light for Nursing Home Residents with Alzheimer's Disease and Related Dementias).

Authors' Contributions

SB contributed to conceptualization, writing, original draft, formal analysis, review and editing, data collection, analysis, and visualization; YF contributed to writing, data modeling, and visualization; NW and NG contributed to data collection and data processing; YLJ participated in writing, review, and editing, and worked as co-corresponding author (yuj15@psu.edu); YJL contributed to data collection, review, and editing; YL participated in data collection; and JW participated in project administration, methodology, investigation, and writing.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Datasets used.
[XLSX File, 1076 KB - aging_v8i1e72338_app1.xlsx]

Multimedia Appendix 2
Model codes.
[DOCX File, 13 KB - aging_v8i1e72338_app2.docx]

Multimedia Appendix 3
Weekly daytime and nighttime lighting condition variations.
[DOCX File, 701 KB - aging_v8i1e72338_app3.docx]

References

1. Czeisler CA. Perspective: casting light on sleep deficiency. Nature New Biol 2013 May 23;497(7450):S13-S13. [doi: 10.1038/497S13a]
2. Duffy JF, Czeisler CA. Effect of light on human circadian physiology. Sleep Med Clin 2009 Jun;4(2):165-177. [doi: 10.1016/j.jsmc.2009.01.004] [Medline: 20161220]
3. Berson D. Strange vision: ganglion cells as circadian photoreceptors. Trends Neurosci 2003 Jun;26(6):314-320. [doi: 10.1016/S0166-2236(03)00130-9]
4. Boudreau P, Dumont GA, Boivin DB. Circadian adaptation to night shift work influences sleep, performance, mood and the autonomic modulation of the heart. PLoS One 2013;8(7):e70813. [doi: 10.1371/journal.pone.0070813] [Medline: 23923024]
5. Turley K, Rafferty J, Bond R, Mulvenna M, Ryan A, Crawford L. Evaluating the Impact of a daylight-simulating luminaire on mood, agitation, rest-activity patterns, and social well-being parameters in a care home for people with dementia: cohort study. JMIR Mhealth Uhealth 2024 Nov 29;12:e56951. [doi: 10.2196/56951] [Medline: 39611803]
6. Brainard GC, Hanifin JP, Greeson JM, et al. Action spectrum for melatonin regulation in humans: evidence for a novel circadian photoreceptor. J Neurosci 2001 Aug 15;21(16):6405-6412. [doi: 10.1523/JNEUROSCI.21-16-06405.2001] [Medline: 11487664]
7. Figueiro MG, Steverson B, Heerwagen J, et al. The impact of daytime light exposures on sleep and mood in office workers. Sleep Health 2017 Jun;3(3):204-215. [doi: 10.1016/j.sleh.2017.03.005]
8. Shochat T, Cohen-Zion M, Tzischinsky O. Functional consequences of inadequate sleep in adolescents: a systematic review. Sleep Med Rev 2014 Feb;18(1):75-87. [doi: 10.1016/j.smrv.2013.03.005]
9. Ancoli-Israel S, Gehrman P, Martin JL, et al. Increased light exposure consolidates sleep and strengthens circadian rhythms in severe Alzheimer's disease patients. Behav Sleep Med 2003 Feb;1(1):22-36. [doi: 10.1207/S15402010BSM0101_4]
10. Figueiro MG. Light, sleep and circadian rhythms in older adults with Alzheimer's disease and related dementias. Neurodegener Dis Manag 2017 Apr;7(2):119-145. [doi: 10.2217/nmt-2016-0060] [Medline: 28534696]
11. Figueiro MG, Plitnick B, Roohan C, Sahin L, Kalsher M, Rea MS. Effects of a tailored lighting intervention on sleep quality, rest-activity, mood, and behavior in older adults with Alzheimer disease and related dementias: a randomized clinical trial. J Clin Sleep Med 2019 Dec 15;15(12):1757-1767. [doi: 10.5664/jcsm.8078] [Medline: 31855161]
12. Jao YL, Wang J, Liao YJ, et al. Effect of Ambient Bright Light on Behavioral and Psychological Symptoms in People With Dementia: A Systematic Review. Innov Aging 2022 May 6;6(3):igac018. [doi: 10.1093/geroni/igac018]
13. Figueiro MG, Pedler D, Plitnick B, Zecena E, Leahy S. Tailored lighting intervention (TLI) for improving sleep-wake cycles in older adults living with dementia. Front Physiol 2023;14:1290678. [doi: 10.3389/fphys.2023.1290678] [Medline: 38169838]
14. Figueiro MG, Sahin L, Kalsher M, Plitnick B, Rea MS. Long-term, all-day exposure to circadian-effective light improves sleep, mood, and behavior in persons with dementia. J Alzheimers Dis Rep 2020 Aug 4;4(1):297-312. [doi: 10.3233/ADR-200212] [Medline: 33024938]
15. Figueiro MG, Plitnick BA, Lok A, et al. Tailored lighting intervention improves measures of sleep, depression, and agitation in persons with Alzheimer's disease and related dementia living in long-term care facilities. Clin Interv Aging 2014;9:1527-1537. [doi: 10.2147/CIA.S68557] [Medline: 25246779]
16. Markvart J, Hansen Å, Christoffersen J. Comparison and correction of the light sensor output from 48 wearable light exposure devices by using a side-by-side field calibration method. LEUKOS 2015 Jul 3;11(3):155-171. [doi: 10.1080/15502724.2015.1020948]

XSL•FO
RenderX

17. Figueiro MG, Hamner R, Bierman A, Rea MS. Comparisons of three practical field devices used to measure personal light exposures and activity levels. Light Res Technol 2013 Aug;45(4):421-434. [doi: 10.1177/1477153512450453] [Medline: 24443644]
18. Bierman A, Klein TR, Rea MS. The Daysimeter: a device for measuring optical radiation as a stimulus for the human circadian system. Meas Sci Technol 2005 Nov 1;16(11):2292-2299. [doi: 10.1088/0957-0233/16/11/023]
19. Lewis P, Christoforou R, Ha PP, Wild U, Schweiker M, Erren TC. Architecture, light, and circadian biology: A scoping review. Science of The Total Environment 2024 Dec;955:177212. [doi: 10.1016/j.scitotenv.2024.177212]
20. Rea MS, Figueiro MG, Bierman A, Bullough JD. Circadian light. J Circadian Rhythms 2010 Feb 13;8(1):2. [doi: 10.1186/1740-3391-8-2] [Medline: 20377841]
21. Khademagha P, Aries MBC, Rosemann ALP, van Loenen EJ. Implementing non-image-forming effects of light in the built environment: a review on what we need. Build Environ 2016 Nov;108:263-272. [doi: 10.1016/j.buildenv.2016.08.035]
22. Figueiro MG, Rea MS. A working threshold for acute nocturnal melatonin suppression from “white” light sources used in architectural applications. J Carcinog Mutagen 2013;4(3). [doi: 10.4172/2157-2518.1000150]
23. Shochat T, Barker DH, Sharkey KM, Van Reen E, Roane BM, Carskadon MA. An approach to understanding sleep and depressed mood in adolescents: person-centred sleep classification. J Sleep Res 2017 Dec;26(6):709-717. [doi: 10.1111/jsr.12550] [Medline: 28573658]
24. Wang Y, Zhang X, Chen H. From needs to control: a review of indicators and sensing technologies for occupant-centric smart lighting systems. Energy Build 2025 Jul;339:115740. [doi: 10.1016/j.enbuild.2025.115740]
25. Kubicki M, Ryan JT, Shaw JM, Depner CM. Overview: wearable devices for monitoring sleep and circadian physiology. In: Kushida CA, editor. Encyclopedia of Sleep and Circadian Rhythms, 2nd edition: Academic Press; 2023:1-5. [doi: 10.1016/b978-0-12-822963-7.00198-5]
26. Lu L, Zhang J, Xie Y, et al. Wearable health devices in health care: narrative systematic review. JMIR Mhealth Uhealth 2020 Nov 9;8(11):e18907. [doi: 10.2196/18907] [Medline: 33164904]
27. Tiersen F, Batey P, Harrison MJC, et al. Smart home sensing and monitoring in households with dementia: user-centered design approach. JMIR Aging 2021 Aug 11;4(3):e27047. [doi: 10.2196/27047] [Medline: 34383672]
28. Rose L, Saha S, Flowers E, et al. Family caregiver perspectives on digital methods to measure stress: qualitative descriptive study. J Med Internet Res 2025 Apr 24;27:e66034. [doi: 10.2196/66034] [Medline: 40273447]
29. Berridge C, Turner NR, Liu L, et al. Advance planning for technology use in dementia care: development, design, and feasibility of a novel self-administered decision-making tool. JMIR Aging 2022 Jul 27;5(3):e39335. [doi: 10.2196/39335] [Medline: 35896014]
30. CS calculator. Lighting Research Center. URL: https://www.lrc.rpi.edu/cscalculator [accessed 2025-07-18]
31. Rea MS, Figueiro MG, Bierman A, Hamner R. Modelling the spectral sensitivity of the human circadian system. Light Res Technol 2012 Dec;44(4):386-396. [doi: 10.1177/1477153511430474]
32. Rea MS, Figueiro MG, Bierman A, Hamner R. Corrigendum. Light Res Technol 2012 Dec;44(4):516. [doi: 10.1177/1477153512467607]
33. Rea MS, Nagare R, Figueiro MG. Modeling circadian phototransduction: quantitative predictions of psychophysical data. Front Neurosci 2021;15:615322. [doi: 10.3389/fnins.2021.615322] [Medline: 33613181]
34. Wyszecki G, Stiles WS. Color Science: Concepts and Methods, Quantitative Data and Formulae: John Wiley & Sons; 1982.
35. Light as a true visual quantity: principles of measurement. Report number CIE 041 - 1978. : Commission Internationale de l’Éclairage; 1994.
36. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods 2009 Nov;41(4):1149-1160. [doi: 10.3758/BRM.41.4.1149]
37. Cohen J. Statistical Power Analysis for the Behavioral Sciences, 2nd edition: Lawrence Erlbaum Associates; 1988.
38. Rea MS, Freyssinier JP. Color rendering: beyond pride and prejudice. Color Research & Application 2010 Dec;35(6):401-409. [doi: 10.1002/col.20562]
39. Chen Z, Yu X, Wang C, Khanh TQ, Liu Q. Limitations of correlated colour temperature in visual perception and its implications for evaluating the colour rendition quality of indoor lighting. Build Environ 2025 Feb;269:112436. [doi: 10.1016/j.buildenv.2024.112436]
40. Li C, Cui G, Melgosa M, et al. Accurate method for computing correlated color temperature. Opt Express 2016;24(13):14066. [doi: 10.1364/OE.24.014066]

Abbreviations

CCT: correlated color temperature
CL: circadian light
CL_A: circadian light adjusted
CS: circadian stimulus
IR: infrared
MAE: mean absolute error

XSL•FO
RenderX

MSE: mean squared error
RGB: red, green, and blue
SPD: spectral power distribution

Edited by H LaMonica; submitted 07.02.25; peer-reviewed by A Ghosh, X Zhang; revised version received 11.05.25; accepted 12.06.25; published 11.09.25.

Please cite as:
Beiglary S, Feng Y, Wang N, Ghaeili N, Jao YL, Liao YJ, Li Y, Wang J
Using Wearable Sensors to Measure and Predict Personal Circadian Lighting Exposure in Nursing Home Residents: Model Development and Validation
JMIR Aging 2025;8:e72338
URL: https://aging.jmir.org/2025/1/e72338
doi:10.2196/72338

© Shevvaa Beiglary, Yanxiao Feng, Nan Wang, Neda Ghaeili, Ying-Ling Jao, Yo-Jen Liao, Yuxin Li, Julian Wang. Originally published in JMIR Aging (https://aging.jmir.org), 11.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using Deep Learning to Perform Automatic Quantitative Measurement of Masseter and Tongue Muscles in Persons With Dementia: Cross-Sectional Study

Mahdi Imani[1], PhD, Biomed Eng; Miguel G Borda[2,3], MD, PhD; Sara Vogrin[1], MD, MSc; Erik Meijering[4], PhD; Dag Aarsland[2,5], MD, PhD; Gustavo Duque[6], MD, PhD

[1]Department of Medicine, Melbourne Medical School, University of Melbourne, St. Albans, Australia

[2]Centre for Age-Related Medicine (SESAM), Stavanger University Hospital, Stavanger, Norway

[3]Department of Neurology, Clínica Universidad de Navarra, Pamplona, Spain

[4]School of Computer Science and Engineering, UNSW Sydney, Sydney, Australia

[5]Department of Old Age Psychiatry, Institute of Psychiatry, Psychology, and Neuroscience, King's College London, London, United Kingdom

[6]Bone, Muscle & Geroscience Group, Research Institute of the McGill University Health Centre, McGill University, 1001 Decarie Blvd, Room EM1.3226, Montreal, QC, Canada

Corresponding Author:
Gustavo Duque, MD, PhD
Bone, Muscle & Geroscience Group, Research Institute of the McGill University Health Centre, McGill University, 1001 Decarie Blvd, Room EM1.3226, Montreal, QC, Canada

Abstract

Background: Sarcopenia (loss of muscle mass and strength) increases adverse outcomes risk and contributes to cognitive decline in older adults. Accurate methods to quantify muscle mass and predict adverse outcomes, particularly in older persons with dementia, are still lacking.

Objective: This study's main objective was to assess the feasibility of using deep learning techniques for segmentation and quantification of musculoskeletal tissues in magnetic resonance imaging (MRI) scans of the head in patients with neurocognitive disorders. This study aimed to pave the way for using automated techniques for opportunistic detection of sarcopenia in patients with neurocognitive disorder.

Methods: In a cross-sectional analysis of 53 participants, we used 7 U-Net-like deep learning models to segment 5 different tissues in head MRI images and used the Dice similarity coefficient and average symmetric surface distance as main assessment techniques to compare results. We also analyzed the relationship between BMI and muscle and fat volumes.

Results: Our framework accurately quantified masseter and subcutaneous fat on the left and right sides of the head and tongue muscle (mean Dice similarity coefficient 92.4%). A significant correlation exists between the area and volume of tongue muscle, left masseter muscle, and BMI.

Conclusions: Our study demonstrates the successful application of a deep learning model to quantify muscle volumes in head MRI in patients with neurocognitive disorders. This is a promising first step toward clinically applicable artificial intelligence and deep learning methods for estimating masseter and tongue muscle and predicting adverse outcomes in this population.

(JMIR Aging 2025;8:e63686) doi:10.2196/63686

KEYWORDS

artificial intelligence; machine learning; sarcopenia; dementia; masseter muscle; tongue muscle; deep learning; head; tongue; face; magnetic resonance imaging; MRI; image; imaging; muscle; muscles; neural network; aging; gerontology; older adults; geriatrics; older adult health

Introduction

Age-related muscle wasting and neurodegeneration, clinically presented as sarcopenia and dementia, respectively, are the major drivers of frailty, falls, and disability in older adults worldwide [1]. Sarcopenia is characterized by loss of muscle mass, strength, and function in older adults. Aging is the leading risk factor, but conditions such as chronic diseases, inflammation, sedentarism, and malnutrition promote sarcopenia onset and progression [2]. Sarcopenia has a 10% overall prevalence globally in older persons, 29% in the community, 14%-33% in long-term care settings, and up to 50% in the very old (>80) [3,4]. Despite being a common and relevant health-related condition, it is unseen and underdiagnosed,

XSL•FO
RenderX

particularly in older persons with cognitive disorders. To diagnose sarcopenia, measurement of muscle mass, muscle performance, and strength is necessary [2]. Estimating muscle performance and strength is accessible and cheap with traditional methods, such as gait speed and grip strength, respectively [5]. However, techniques such as dual X-ray absorptiometry or body magnetic resonance imaging (MRI) are necessary to accurately assess lean or muscle mass. These methods can increase costs and time and are impractical in settings such as dementia clinics [6].

Dementia patients are highly affected by sarcopenia, with a prevalence of around 60% - 70% [7,8]. People living with neurodegenerative diseases are more prone to experience difficulties due to malnutrition, being sedentary, and falling; therefore, having sarcopenia increases the risk of adverse outcomes. Sarcopenia is not only a risk factor for adverse outcomes for those with dementia but also promotes cognitive loss in healthy older adults [9]. Therefore, diagnosing sarcopenia in people with neurodegenerative diseases is relevant and necessary.

Head MRI is a widely used diagnostic method for assessing dementia and Alzheimer disease (AD), as it offers intricate representations of the brain's anatomy and physiology. In clinical practice, MRI is often combined with other imaging techniques and cognitive assessments to support the diagnosis of these conditions. The utilization of MRI has seen an upward trend in recent times, as it has become an instrumental tool for the early detection and tracking of the evolution of dementia and AD. According to estimates, 60% - 80% of patients diagnosed with dementia or AD undergo MRI as part of their diagnostic evaluation [6].

Mastication and deglutition muscles such as the masseter and tongue are visible in brain MRI scans [10,11]. These muscles can reflect not only age-associated general muscle decline but also systemic processes due to highly complex interactions with the immune system and the inflammatory response, the nervous system, and the crossroads of several components of the frailty syndrome [12,13]. Indeed, in a previous publication, we have reported that manually segmented masseter predicts mortality and clinical short-term and long-term outcomes in several contexts [14]. Moreover, head muscles such as the tongue and masseter could be cost-effective alternatives to estimate muscle mass in dementia and other common conditions such as head cancer, stroke, or cranioencephalic trauma [14,15]. However, manual and semi-automatic techniques are labor-intensive and time-consuming, making the image processing task for large studies difficult, expensive, and, most importantly, impractical to apply in a clinical setting. Therefore, in the present study, we aimed to use MRI scans of the head opportunistically to develop an automated deep learning method to evaluate sarcopenia.

Methods

Population and Data Source

The "Dementia Study of Western Norway" (DemVest) is a long-term study between 2005 and 2013, with ongoing follow-up assessments. Participants were referred from dementia clinics in Hordaland and Rogaland and were insured by the same national insurance scheme. The study's methodology is described elsewhere [15]. Those with moderate or severe dementia, delirium, past bipolar or psychotic conditions, terminal illness, or newly diagnosed somatic diseases impacting cognition, function, or participation were excluded.

For this study, subjects with dementia with Lewy bodies (DLB) or mild AD who had baseline MRI scans were included. Out of 111 participants (85 AD and 26 DLB), 33 AD and 20 DLB participants with MRI images for brain and muscle measurement were selected based on the quality of the images and clear delineation of the regions of interest. The diagnosis of dementia was made in accordance with the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) criteria, and patients were classified as AD or DLB [16]. A mini-mental state examination score of ≥20 or a clinical dementia rating global score of 1 was chosen as the definition for mild dementia. The diagnosis was based on inclusion but could be modified with clinical evolution, consensus, and autopsy [15]. Participants were evaluated through structured assessments, and medical records were used to gather complete medical history and comorbidity data. In total, 56 participants had pathological diagnoses with 80% accuracy compared to clinical criteria, which reflects a reliable initial clinical diagnosis [17].

Ethical Considerations

This study was approved by the regional ethics committee (approval code: 2010/633) and the Norwegian authorities for the collection of medical data. All data were handled and kept following national data privacy protocols. All participants signed an informed consent form before inclusion in the study.

Imaging

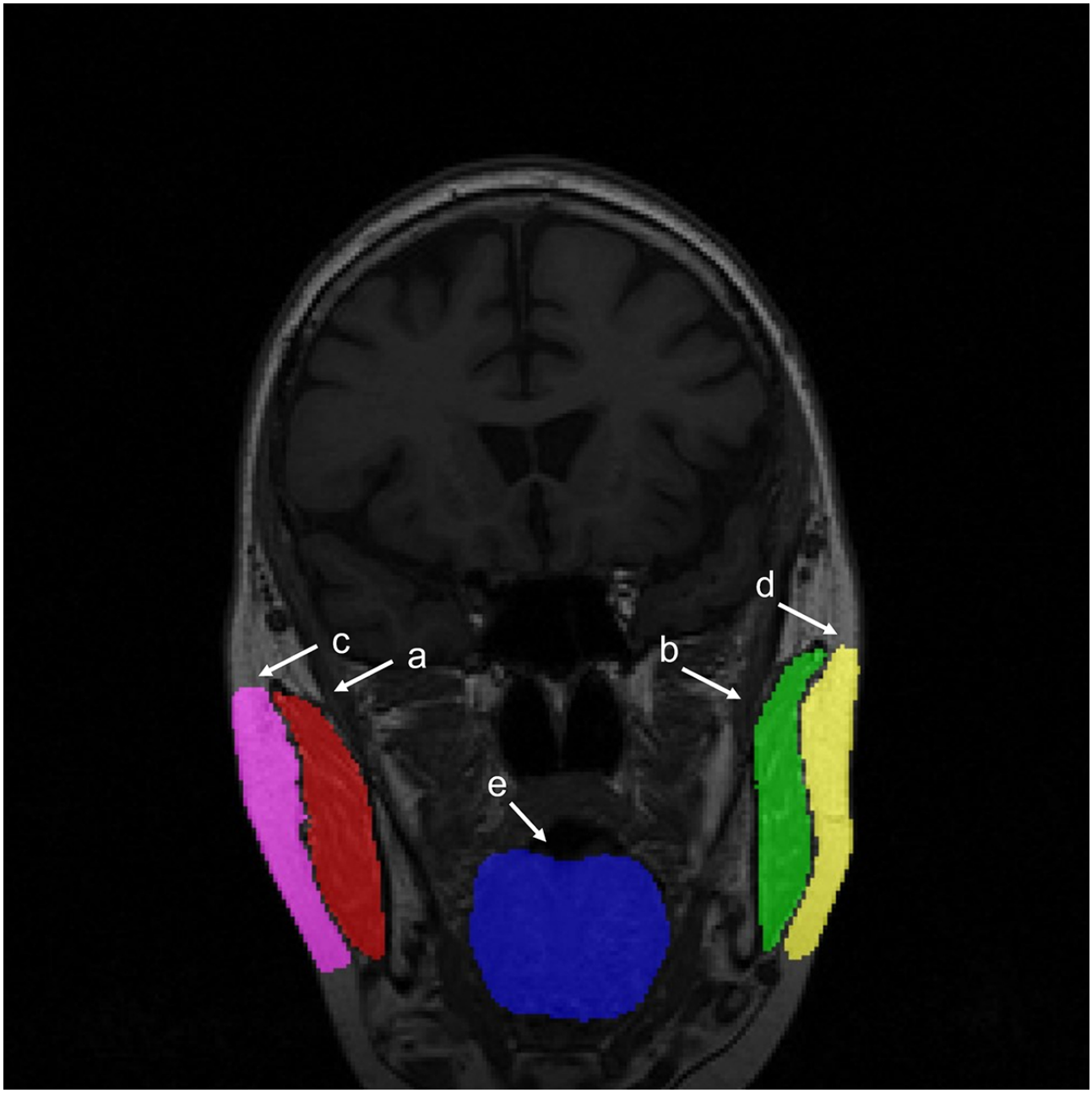
All images were acquired at baseline. A 1.5-T Philips Intera-scanner was used to obtain MRI images. The acquisition protocol for 3D T1-weighted sequence was as follows: flip angle of 30°, repetition time/echo time of 10.0/4.6 ms, number of excitations of 2, 2-mm slice thickness with 1-mm spacing between the slices (1-mm slices with no gap), matrix of 256×256 pixels, and field of view of 26 cm. Those with movement artifacts and inadequate image quality were removed from the data using visual quality checks. A standardized preprocessing method for harmonizing multiple collections of MRIs was applied, which consisted of movement correction and intensity normalization following previously validated techniques [15].

Ground Truth Image Segmentation

Ground truth (GT) images were segmented using interactive pixel techniques made available by SliceOmatic software (TomoVision) following a manual method previously reported [18]. The size of the masseter muscle was used as a reference for the selection criteria of the slices. In total, 5 slices were selected from the ones with both right and left masseter muscles at their largest. For each slice, 5 tissues were segmented: left and right masseter muscles, left and right subcutaneous fat, and tongue muscle (Figure 1). The masseter muscle on each side was used as a reference to segment subcutaneous fat.

XSL•FO
RenderX

Figure 1. Example of segmented tissues overlaid on the original MRI. (**A**) Right masseter muscle, (**B**) left masseter muscle, (**C**) right subcutaneous fat, (**D**) left subcutaneous fat, and (**E**) tongue muscle.



Network Architecture

We studied and compared 6 different U-Net-like architectures. The original U-Net architecture [19] was first designed to segment medical images, and many other researchers have tried to improve its performance by integrating additional techniques into its architecture [20]. U-Net consists of a contracting path (encoder) and an expansive path (decoder) with skip connections between these 2 paths. The network learns features from the provided image and the mask at increasingly higher spatial scales by gradually down-sampling to lower resolutions through the encoding path. The expansive path then gradually increases the resolution of the output from the encoding path to the original image size, resulting in a probability map as an output, indicating the chance of each pixel belonging to a specific tissue. One important feature of U-Net is its skip connections, which concatenate feature maps from the encoding path to the corresponding block in the expansive path, making it possible to maintain small details crucial in medical image segmentation.

We have included 5 variants of U-Net in this study, including Attention U-Net, Dense U-Net, Residual U-Net, Inception U-Net, and U-Net++. Attention U-Net is desirable since it allows the model to focus on specific objects and ignore unnecessary areas [21]. In Dense U-Net, the traditional U-Net blocks are replaced with a dense block, enabling the model to segment objects with greater distinction. This feature is important in medical practice since tissues are often very close and sometimes overlap [22]. Residual U-Net architecture tries to tackle the vanishing gradient issue, a common problem in designing deep neural networks [23]. In most cases, the same organ's size can vary between patients, which can cause limitations on the segmentation capability of the model. By using filters with

XSL•FO
RenderX

different sizes, Inception U-Net attempts to overcome this problem [24]. Lastly, U-Net++ aids the classic U-Net model to more accurately segment images by providing semantic information from a dense network of skip connections as an intermediary grid between the encoding and decoding paths [25]. We also included a Wide U-Net architecture to eliminate the effect of increased trainable parameters. This model has the same architecture as U-Net but with more feature maps per layer (30, 60, 120, 240, and 480) compared to the original U-Net (16, 32, 64, 128, and 256). Hence, it will serve as a control to compare the models with larger numbers of trainable parameters with the base U-Net (Table 1).

Table . Number of trainable parameters for each model.

Model	Trainable parameters
U-Net	2,164,390
Attention U-Net	2,233,270
U-Net++	2,555,702
Inception U-Net	5,529,526
Residual U-Net	6,877,110
Dense U-Net	7,666,320
Wide U-Net	7,596,306

Training Procedure

All models were trained with the mini-batch stochastic gradient descent algorithm using the Adam optimizer. A batch size of 8 was selected. The learning rate was 0.0001. The training was done for 200 epochs. The values of hyperparameters were empirically tuned for best performance. Categorical cross-entropy (CSE) was selected as the loss function for this study [26]. The CSE loss function minimizes the distance between 2 distributions (the predicted labels and the GT labels). CSE is one of the most popular loss functions for image segmentation and has shown excellent performance in muscle segmentation [27,28]. All experiments were implemented with open-source software: Python (version 3.7.13), TensorFlow (version 2.8.2), and Keras (version 2.8.0).

Model Evaluation

The results of the experiments were evaluated using 2 main measures: Dice similarity coefficient (DSC) and average symmetric surface distance (ASSD). The DSC represents the agreement between the GT labels and predicted labels that models generate:

$$DSC_{P,G} = 2(P \cap G)(P+G)$$

where $\cap$ is the intersection and P and G are the 2 labels. DSC ranges between 0 and 1, where 0 indicates no agreement and 1 indicates perfect agreement. In our study, DSC is presented as a percentage.

The ASSD measures the average distance from pixels on the boundary of predicted labels to corresponding pixels on the boundary of the GT labels:

where BP and BG are the boundaries of predicted labels and corresponding reference labels, respectively. d(v,B) is the shortest Euclidean distance between voxel v and boundary B. An ASSD of 0 indicates a perfect match between predicted and reference labels. The ASSD was measured in mm.

We used k-fold cross-validation in evaluating the models. This technique splits the data set into k subsets (folds). The deep learning models are trained on all but one of the subsets (k–1), and then the models are evaluated on the subset that was not used for training. This process is repeated k times, and the average of the results is reported. We used k=10.

Additional metrics, including Jaccard coefficient, precision, recall, sensitivity, specificity, and F_1-score, are presented in Multimedia Appendix 1. The results are shown as mean and standard deviation and median and interquartile range across k-fold (k=10) cross-validation.

Statistical Methods

We explored the association between the BMI and the muscle and fat areas and volumes in the MRI images using individual linear regression models adjusted by sex and age. No adjustments for multiple testing were made. All assumptions were checked. All *P*-values were evaluated at a 5% level. The analysis and graphs were carried out using R (version 4.2.2).

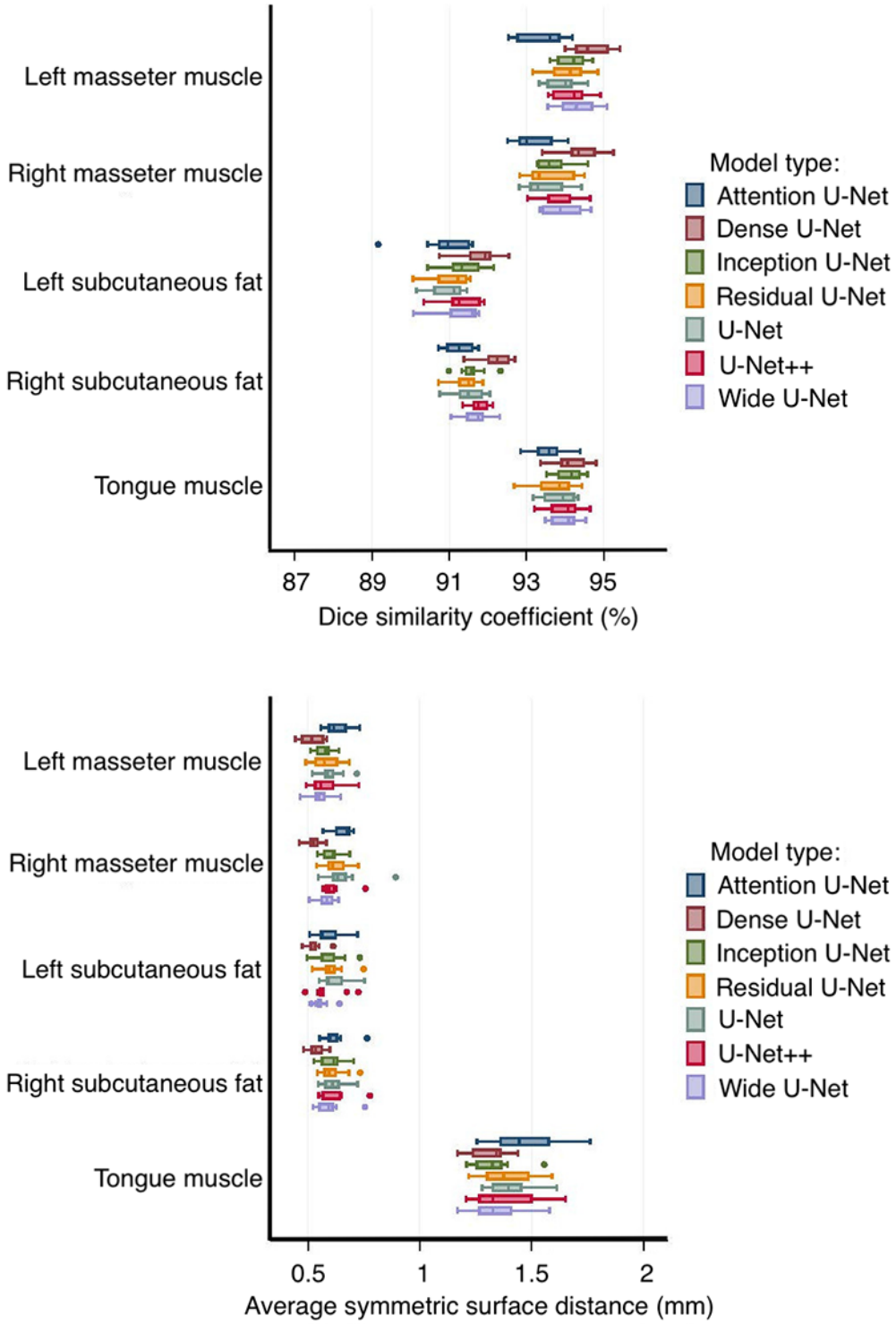
Results

Segmentation

DSC and ASSD were used to quantitatively analyze the segmentation results (Figure 2 and Table 2). The Dense U-Net model has a higher average DSC and lower ASSD than other models. In contrast, Attention U-Net has the lowest average DSC score compared to other models. This result is confirmed by the higher ASSD for Attention U-Net for all tissues. Other models have shown almost similar results for all areas. Additional metrics have been presented in Multimedia Appendix 1. The results from these metrics confirm the findings of the study mentioned in this section; hence, they were omitted from the main manuscript.

XSL•FO
RenderX

Figure 2. Box plot of k-fold (k=10) cross-validation results for attention U-Net, dense U-Net, inception U-Net, residual U-Net, U-Net, U-Net++, and wide U-Net. Top: DSC in percentage. Bottom: ASSD in mm.



XSL•FO
RenderX

Table . Mean DSC and ASSD for test and validation results for k-fold (k=10) cross-validation. Standard deviation for the measurements in this table is presented in Multimedia Appendix 1.

	Model	Test set					Validation set				
		T[a]	LM[b]	RM[c]	LSF[d]	RSF[e]	T	LM	RM	LSF	RSF
DSC[f] (%)	Attention U-Net	93.57	93.4	93.19	90.9	91.26	93.43	93.38	93.23	91.05	91.43
	Dense U-Net	94.12	94.66	94.4	91.76	92.22	94.07	94.7	94.49	91.89	92.42
	Inception U-Net	94.07	94.12	93.69	91.36	91.59	93.85	94.27	93.73	91.45	91.81
	Residual U-Net	93.71	94.03	93.53	91.05	91.46	93.74	94.1	93.67	91.35	91.68
	U-Net	93.83	93.94	93.46	90.97	91.49	93.88	93.93	93.52	91.18	91.67
	Wide U-Net	94	94.29	93.93	91.35	91.68	94.03	94.3	93.99	91.59	92.01
ASSD[g] (mm)	Attention U-Net	1.47	0.63	0.66	0.59	0.62	1.53	0.67	0.65	0.6	0.59
	Dense U-Net	1.3	0.52	0.53	0.53	0.54	1.37	0.51	0.53	0.52	0.54
	Inception U-Net	1.33	0.57	0.61	0.59	0.61	1.44	0.56	0.63	0.6	0.59
	Residual U-Net	1.39	0.58	0.62	0.61	0.61	1.38	0.58	0.63	0.58	0.6
	U-Net	1.41	0.6	0.65	0.62	0.61	1.45	0.6	0.64	0.61	0.59
	Wide U-Net	1.35	0.55	0.58	0.56	0.6	1.34	0.55	0.58	0.56	0.57

[a]T: tongue muscle.

[b]LM: left masseter muscle.

[c]RM: right masseter muscle.

[d]LSF: left subcutaneous fat.

[e]RSF: right subcutaneous fat.

[f]DSC: Dice similarity coefficient.

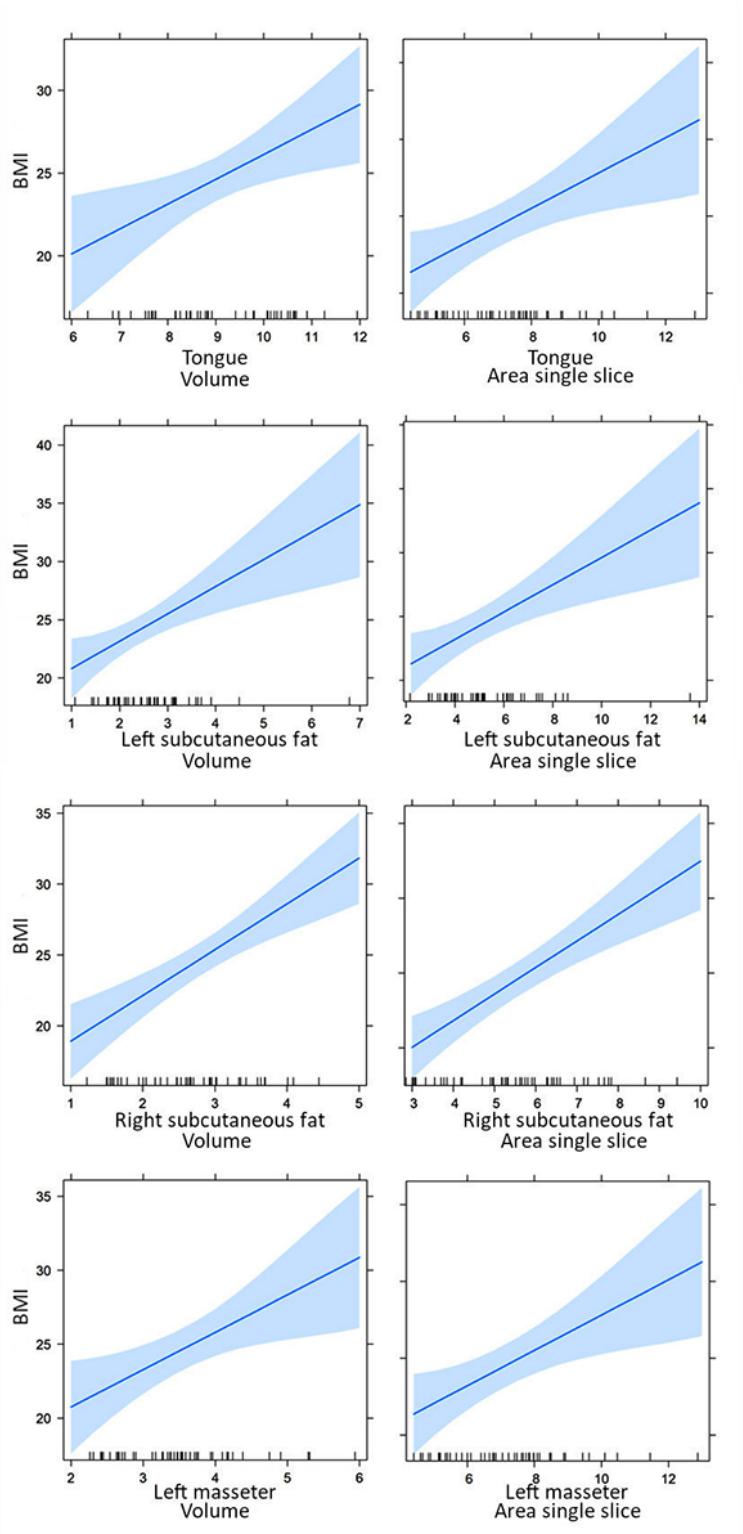
[g]ASSD: average symmetric surface distance.

Clinical Validation

To prove the clinical validity of the measurements, we evaluated the association between the segmented muscles and subcutaneous fat and BMI. We found a significant positive association between tongue muscle, left masseter muscle, and left and right subcutaneous fat and BMI (Figure 3). The area of a single slice as well as the volume of 5 slices per patient was calculated for this experiment. The results were adjusted by age and sex with $P<.05$ (Multimedia Appendix 2).

XSL•FO
RenderX

Figure 3. Individual linear regression repressing the relationship between quantitative results (area of a single slice and volume of 5 slices) from the Dense U-Net model and BMI.



Discussion

In this study, we evaluated the performance of six deep learning models for segmentation of the masseter muscles, subcutaneous fat, and tongue muscle in MRI images of the head. Several variations of the U-Net architecture were trained and tested using k-fold cross-validation. The use of deep neural networks for segmenting musculoskeletal tissues in patients with AD and LBD is a novel experimental contribution to the deep learning-based segmentation literature as well as the clinical literature.

Our study demonstrated that the Dense U-Net model performed superiorly to the other models in all evaluated regions by achieving an overall DSC of 93.43% and ASSD of 0.68 mm on the test data. The remaining models demonstrated comparable outcomes, except the Attention U-Net, which achieved slightly

XSL•FO
RenderX

less accurate results in all regions with an overall DSC of 92.46% and ASSD of 0.79 mm on the test data. Notably, despite having a comparable number of trainable parameters, the Dense U-Net model demonstrated a higher DSC and lower ASSD than the Wide U-Net model. The Dense U-Net model achieves superior performance with a similar number of parameters, suggesting that its architectural efficiency, rather than parameter quantity alone, drives this improvement.

The validation set produced similar results to the training set, indicating that the trained models did not suffer from overfitting. Although we applied data augmentation to the training data set (results not shown), it did not significantly improve the accuracy of the segmentation models.

Furthermore, we observed a significant correlation between the results and BMI, a well-established measure of nutrition and body composition. This underscores the validity and clinical relevance of this method. If these localized muscle measurements correlate with BMI, it suggests that they may reflect broader nutritional and body composition states.

Sarcopenia is a relatively newly recognized condition for which neuromuscular degeneration, central nervous system alpha motor unit loss, and fat infiltration into muscle are the most distinctive proven and observed pathogenic features, leading to loss of muscle mass and strength and predisposing to physical frailty [2]. We have proposed that non-invasive assessment of intermuscular adipose tissue and muscle mass by image analysis could constitute a viable method to diagnose sarcopenia and predict its associated outcomes, the clinical impact of which is also under study by our group [29].

The validity and clinical implications of measuring the masseter muscle have been shown in previous studies [30,31]. In recent work by our group, we compared the diagnostic capacity for sarcopenia between the gold-standard dual X-ray absorptiometry and our measurements of head muscles. The results showed that both methods had equivalent accuracy [32]. In older adults with glioblastoma, a decreased masseter diameter on preoperative imaging was associated with shorter overall survival and 90-day mortality after surgical resection [33]. In addition, low masseter muscle was significantly associated with worse overall survival in patients aged 65 years or older, diagnosed with squamous cell carcinoma of the head and neck and treated with curative intent [34]. Another study evaluated post-operative results after carotid endarterectomy; low masseter mass was associated with a prolonged hospital stay and recurrent stroke within 5 years [35]. In another study, preoperative masseter mass was a predictor of postoperative pneumonia in patients with esophageal cancer [36]. Additionally, other studies have shown that the masseter muscle can be used as a nutritional biomarker. The masseter muscle, analyzed via computed tomography (CT) anthropometry, showed a statistically significant association with systemic nutritional biomarkers [37]. On the other hand, the tongue has shown to be a good marker of prognosis, as tongue strength has shown to be helpful in diagnosing sarcopenia [38]. Previous studies from our group also report that tongue muscle volume is correlated with malnutrition and even brain structures in patients with dementia [18,29].

Therefore, the approach we present in this paper can be opportunistically used to quantify muscle volumes and investigate the implications of having low muscle mass in the masseter or the tongue in people with brain, head, and neck diseases. Thus, it is an important first step toward developing a more efficient method to estimate masseter and tongue muscle with a better capacity to be implemented in clinical practice. Manual and semi-automatic techniques have been employed in several studies for masseter muscle segmentation in MRI [39]. A recent study used shape determination to segment the masseter muscle in MRI images [40]. In this approach, a manual contour for 8 slices must be defined by the user, and the model then determines the shape for the remaining scans, reducing the time and labor required for segmentation. However, this technique still requires manual segmentation, which can be time-consuming and prone to user error compared to our approach.

Model-based techniques have also been explored for the segmentation of the mastication muscle [41,42]. Although these techniques have shown high accuracy (>90%), there must be a distinct boundary between the masseter muscle and surrounding tissues to ensure accuracy. This distinct boundary refers to visible differences in intensity, texture, or anatomical structure on imaging, which enable the models to accurately separate the muscle from adjacent tissues, such as fat or bone.

Machine and deep learning approaches have been widely used to segment muscles in various body parts. CT scans and cone beam CT scans have primarily been investigated for measuring the head organs, including masseter muscle using these techniques [43-45]. In a previous study, a basic U-Net model was applied to segment the masseter muscle in head CT scans to investigate hemifacial microsomia [46]. The mean DSC reported in that study was 79.4% for the experimental group and 82.4% for the group with mandible deformities, which are lower compared to the results obtained in our study. In a study using CT scan of the head for segmentation of masticatory muscles, deep learning techniques were superior to atlas-based techniques, achieving a mean DSC of 83% [47]. To the best of our knowledge, the techniques considered in our study have not yet been used to segment musculoskeletal tissues in MRI images of the head.

Our study had some limitations. The Attention U-Net model demonstrated the lowest DSC values and the highest ASSD values among the evaluated models, indicating suboptimal segmentation accuracy across tissues. This underperformance may be attributed to the Attention mechanism's inability to effectively focus on the target tissue, leading to dispersed attention, particularly in smaller structures or regions with indistinct boundaries. To address this limitation, future work could involve refining the Attention mechanism to enhance its specificity and focus on regions of interest. Alternatively, exploring models that prioritize multi-scale feature extraction and detail preservation may provide improved segmentation performance, particularly for small or complex anatomical regions.

The DemVest study had some limitations that may have impacted the results. Selection bias might have been present if

XSL•FO
RenderX

patients with more complex health conditions were included, as primary care referrals were used. The study was not specifically designed for the paper's purpose, which could limit the data analysis and interaction control. For example, the absence of a healthy control group prevents us from determining whether the observed muscle volume characteristics are specific to individuals with AD or neurocognitive disorders, or if they represent normal variations associated with aging. However, the primary objective of this study was to demonstrate the feasibility and accuracy of our deep learning model in quantifying muscle volumes using head MRI, rather than to establish definitive differences between diseased and healthy populations. In addition, the sample size is relatively small, and there is a risk that our results may not generalize to larger populations. Additionally, the MRI scans were obtained from a single center using a single MRI machine, which may affect the model's generalizability. Longitudinal analyses were not conducted because imaging was only performed at baseline, in line with the initial primary objectives of the DemVest study and the resources allocated for image acquisition.

These limitations should be considered when interpreting this study's results and addressed in future studies seeking broader application of the proposed approach. Whether masseter and tongue volumes in these muscles correlate with lean body mass and inflammaging and could be predictors of neurocognitive conditions and their associated outcomes remains unknown.

On the other hand, this study has several strengths, including well-characterized data, detailed and exhaustive diagnosis to correctly identify people with dementia, and an automated quality control and analysis pipeline. Additionally, the results fill a gap in the literature and provide insights into a possible method to efficiently diagnose sarcopenia in context when a head MRI is already available.

In summary, to our knowledge, this is the first study that validates deep learning methods that could be easily implemented in clinical practice to measure masseter and tongue muscle volumes with a solid potential to become biomarkers with strong predictive value for adverse outcomes in older persons with dementia. Since imaging is widely used in memory clinics worldwide, this opportunistic approach to image analyses could become standard practice in those settings. However, further large longitudinal studies are still required.

Acknowledgments

We want to thank the participants, researchers, and technical staff that made the DemVest study possible, as well as the staff and facilities provided by (SESAM) Centre for age-related medicine at Stavanger, Norway. This work was supported by the Norwegian government, through hospital owner Helse Vest (Western Norway Regional Health Authority) number 911973. It is also funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust, King's College London, and the Australian Institute for Musculoskeletal Science (AIMSS). Mahdi Imani is supported by an Australian Government Research Training Program (TRP) Scholarship.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Table presenting extended accuracy evaluation metrics, including Dice Similarity Coefficient (DSC), Jaccard Coefficient (JC), Average Symmetric Surface Distance (ASSD), Precision, Recall, Sensitivity, Specificity, and F1-Score. Results are presented as mean (standard deviation) and median (interquartile range) across 10-fold cross-validation.
[XLSX File, 20 KB - aging_v8i1e63686_app1.xlsx]

Multimedia Appendix 2
Table presenting association between area and volumes from MRI scans and BMI in each region of study.
[DOCX File, 16 KB - aging_v8i1e63686_app2.docx]

References

1. WHO G. WHO Methods and Data Sources for Global Burden of Disease Estimates 2000– 2011. : Department of Health Statistics and Information Systems; 2020.
2. Cruz-Jentoft AJ, Bahat G, Bauer J, et al. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing 2019 Jan 1;48(1):16-31. [doi: 10.1093/ageing/afy169] [Medline: 30312372]
3. Shafiee G, Keshtkar A, Soltani A, Ahadi Z, Larijani B, Heshmat R. Prevalence of sarcopenia in the world: a systematic review and meta- analysis of general population studies. J Diabetes Metab Disord 2017;16:21. [doi: 10.1186/s40200-017-0302-x] [Medline: 28523252]
4. Cruz-Jentoft AJ, Landi F, Schneider SM, et al. Prevalence of and interventions for sarcopenia in ageing adults: a systematic review. Report of the International Sarcopenia Initiative (EWGSOP and IWGS). Age Ageing 2014 Nov;43(6):748-759. [doi: 10.1093/ageing/afu115] [Medline: 25241753]

XSL•FO
RenderX

5. Morley JE. Sarcopenia: diagnosis and treatment. J Nutr Health Aging 2008;12(7):452-456. [doi: 10.1007/BF02982705] [Medline: 18615226]
6. Tan LF, Lim ZY, Choe R, Seetharaman S, Merchant R. Screening for frailty and sarcopenia among older persons in medical outpatient clinics and its associations with healthcare burden. J Am Med Dir Assoc 2017 Jul 1;18(7):583-587. [doi: 10.1016/j.jamda.2017.01.004] [Medline: 28242192]
7. Cobo CS, Pérez V, Hermosilla C, Nuñez M, Lorena PD. Prevalence of sarcopenia in elderly with dementia institutionalized: a multicenter study. J Aging Res 2014;3:178-181. [doi: 10.14283/jarcp.2014.31]
8. Sperlich E, Fleiner T, Zijlstra W, Haussermann P, Morat T. Sarcopenia in geriatric psychiatry: feasibility of the diagnostic process and estimation of prevalence within a hospital context. J Cachexia Sarcopenia Muscle 2021 Oct;12(5):1153-1160. [doi: 10.1002/jcsm.12748] [Medline: 34151538]
9. Peng TC, Chen WL, Wu LW, Chang YW, Kao TW. Sarcopenia and cognitive impairment: a systematic review and meta-analysis. Clin Nutr 2020 Sep;39(9):2695-2701. [doi: 10.1016/j.clnu.2019.12.014] [Medline: 31917049]
10. van Spronsen PH, Weijs WA, Valk J, Prahl-Andersen B, van Ginkel FC. Comparison of jaw-muscle bite-force cross-sections obtained by means of magnetic resonance imaging and high-resolution CT scanning. J Dent Res 1989 Dec;68(12):1765-1770. [doi: 10.1177/00220345890680120901] [Medline: 2600258]
11. Lee YH, Lee KM, Auh QS. MRI-based assessment of masticatory muscle changes in TMD patients after Whiplash injury. J Clin Med 2021 Apr 1;10(7):1404. [doi: 10.3390/jcm10071404] [Medline: 33915742]
12. Tay L, Ding YY, Leung BP, et al. Sex-specific differences in risk factors for sarcopenia amongst community-dwelling older adults. Age (Dordr) 2015 Dec;37(6):121. [doi: 10.1007/s11357-015-9860-3] [Medline: 26607157]
13. Fung FY, Koh YLE, Malhotra R, et al. Prevalence of and factors associated with sarcopenia among multi-ethnic ambulatory older Asians with type 2 diabetes mellitus in a primary care setting. BMC Geriatr 2019 Apr 29;19(1):122. [doi: 10.1186/s12877-019-1137-8] [Medline: 31035928]
14. van Heusden HC, Chargi N, Dankbaar JW, Smid EJ, de Bree R. Masseter muscle parameters can function as an alternative for skeletal muscle mass assessments on cross-sectional imaging at lumbar or cervical vertebral levels. Quant Imaging Med Surg 2022 Jan;12(1):15-27. [doi: 10.21037/qims-21-43] [Medline: 34993057]
15. Aarsland D, Rongve A, Nore SP, et al. Frequency and case identification of dementia with Lewy bodies using the revised consensus criteria. Dement Geriatr Cogn Disord 2008;26(5):445-452. [doi: 10.1159/000165917] [Medline: 18974647]
16. Bell CC. DSM-IV: Diagnostic and Statistical Manual of Mental Disorders. J Am Med Assoc 1994 Sep 14;272(10):828. [doi: 10.1001/jama.1994.03520100096046] [Medline: 7933395]
17. Skogseth R, Hortobágyi T, Soennesyn H, et al. Accuracy of clinical diagnosis of dementia with Lewy bodies versus neuropathology. J Alzheimers Dis 2017;59(4):1139-1152. [doi: 10.3233/JAD-170274] [Medline: 28731443]
18. Borda MG, Bani Hassan E, Weon JH, et al. Muscle volume and intramuscular fat of the tongue evaluated with MRI predict malnutrition in people living with dementia: a 5-year follow-up study. J Gerontol A Biol Sci Med Sci 2022 Feb 3;77(2):228-234. [doi: 10.1093/gerona/glab224] [Medline: 34338751]
19. Falk T, Mai D, Bensch R, et al. U-Net: deep learning for cell counting, detection, and morphometry. Nat Methods 2019 Jan;16(1):67-70. [doi: 10.1038/s41592-018-0261-2] [Medline: 30559429]
20. Ronneberger O, Fischer P, Brox T, editors. U-net: convolutional networks for biomedical image segmentation. Presented at: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference; Oct 5-9, 2015; Munich, Germany.
21. Oktay O, Schlemper J, Folgoc LL, Lee M, Heinrich M, Misawa K, et al. Attention U-Net: learning where to look for the pancreas. arXiv. Preprint posted online on 2018arXiv:180403999.
22. Cai S, Tian Y, Lui H, Zeng H, Wu Y, Chen G. Dense-UNet: a novel multiphoton in vivo cellular image segmentation model based on a convolutional neural network. Quant Imaging Med Surg 2020 Jun;10(6):1275-1285. [doi: 10.21037/qims-19-1090] [Medline: 32550136]
23. Zhang ZX, Liu QJ, Wang YH. Road extraction by deep residual U-Net. IEEE Geosci Remote Sensing Lett 2018 May;15(5):749-753. [doi: 10.1109/LGRS.2018.2802944]
24. Punn NS, Agarwal S. Inception U-Net architecture for semantic segmentation to identify nuclei in microscopy cell images. ACM Trans Multimedia Comput Commun Appl 2020 Feb 29;16(1):1-15. [doi: 10.1145/3376922]
25. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. UNet++: redesigning skip connections to exploit multiscale features in image segmentation. IEEE Trans Med Imaging 2020 Jun;39(6):1856-1867. [doi: 10.1109/TMI.2019.2959609] [Medline: 31841402]
26. Zhang ZL, Sabuncu MR. Generalized cross entropy loss for training deep neural networks with noisy labels. In: Adv Neural Inf Process Syst 2018, Vol. 31. [Medline: 000461852003034]
27. Garcia-Garcia A, Orts-Escolano S, Oprea S, Villena-Martinez V, Martinez-Gonzalez P, Garcia-Rodriguez J. A survey on deep learning techniques for image and video semantic segmentation. Appl Soft Comput 2018 Sep;70:41-65. [doi: 10.1016/j.asoc.2018.05.018]
28. Aldoj N, Biavati F, Michallek F, Stober S, Dewey M. Automatic prostate and prostate zones segmentation of magnetic resonance images using DenseNet-like U-net. Sci Rep 2020 Aug 31;10(1):14315. [doi: 10.1038/s41598-020-71080-0] [Medline: 32868836]

XSL•FO
RenderX

29. Borda MG, Castellanos-Perilla N, Tovar-Rios DA, Ferreira D, Duque G, Aarsland D. Tongue muscle mass is associated with total grey matter and hippocampal volumes in Dementia with Lewy Bodies. Arch Gerontol Geriatr 2022;100:104647. [doi: 10.1016/j.archger.2022.104647] [Medline: 35134612]
30. Uhlich R, Hu P. Sarcopenia diagnosed using masseter muscle area predictive of early mortality following severe traumatic brain injury. Neural Regen Res 2018 Dec;13(12):2089-2090. [doi: 10.4103/1673-5374.241451] [Medline: 30323130]
31. Wallace JD, Calvo RY, Lewis PR, et al. Sarcopenia as a predictor of mortality in elderly blunt trauma patients: comparing the masseter to the psoas using computed tomography. J Trauma Acute Care Surg 2017 Jan;82(1):65-72. [doi: 10.1097/TA.0000000000001297] [Medline: 27820555]
32. Borda MG, Duque G, Pérez-Zepeda MU, et al. Using magnetic resonance imaging to measure head muscles: an innovative method to opportunistically determine muscle mass and detect sarcopenia. J Cachexia Sarcopenia Muscle 2024 Feb;15(1):189-197. [doi: 10.1002/jcsm.13362] [Medline: 38050325]
33. Morshed RA, Young JS, Casey M, et al. Sarcopenia diagnosed using masseter muscle diameter as a survival correlate in elderly patients with glioblastoma. World Neurosurg 2022 May;161:e448-e463. [doi: 10.1016/j.wneu.2022.02.038] [Medline: 35181534]
34. McGoldrick D, Alsabbagh AY, Pettit L, Shaikh M, Bhatia S. Masseter muscle defined sarcopenia and survival in head and neck cancer patients at a single cancer centre. Br J Oral Maxillofac Surg 2020 Dec;58(10):e157-e158. [doi: 10.1016/j.bjoms.2020.10.073]
35. Hogenbirk RNM, Banning LBD, Visser A, et al. Association between masseter muscle area and thickness and outcome after carotid endarterectomy: a retrospective cohort study. J Clin Med 2022 May 30;11(11):3087. [doi: 10.3390/jcm11113087] [Medline: 35683474]
36. Kamada T, Ohdaira H, Ito E, et al. Association between masseter muscle sarcopenia and postoperative pneumonia in patients with esophageal cancer. Sci Rep 2022 Sep 30;12(1):16374. [doi: 10.1038/s41598-022-20967-1] [Medline: 36180776]
37. Hwang Y, Lee YH, Cho DH, Kim M, Lee DS, Cho HJ. Applicability of the masseter muscle as a nutritional biomarker. Medicine (Baltimore) 2020 Feb;99(6):e19069. [doi: 10.1097/MD.0000000000019069] [Medline: 32028430]
38. Chen KC, Lee TM, Wu WT, Wang TG, Han DS, Chang KV. Assessment of tongue strength in sarcopenia and sarcopenic dysphagia: a systematic review and meta-analysis. Front Nutr 2021;8:684840. [doi: 10.3389/fnut.2021.684840] [Medline: 34249993]
39. Volk GF, Karamyan I, Klingner CM, Reichenbach JR, Guntinas-Lichius O. Quantitative magnetic resonance imaging volumetry of facial muscles in healthy patients with facial palsy. Plast Reconstr Surg Glob Open 2014 Jun;2(6):e173. [doi: 10.1097/GOX.0000000000000128] [Medline: 25289366]
40. Ng HP, Foong KWC, Ong SH, et al. Quantitative analysis of human masticatory muscles using magnetic resonance imaging. Dentomaxillofac Radiol 2009 May;38(4):224-231. [doi: 10.1259/dmfr/75198413] [Medline: 19372110]
41. Ng HP, Ong SH, Hu Q, Foong KWC, Goh PS, Nowinski WL. Muscles of mastication model-based MR image segmentation. Int J CARS 2006 Nov 23;1(3):137-148. [doi: 10.1007/s11548-006-0046-4]
42. Ng HP, Ong SH, Liu J, et al. 3D segmentation and quantification of a masticatory muscle from MR data using patient-specific models and matching distributions. J Digit Imaging 2009 Oct;22(5):449-462. [doi: 10.1007/s10278-008-9132-1] [Medline: 18516642]
43. Jiang Y, Shang F, Peng J, et al. Automatic masseter muscle accurate segmentation from CBCT using deep learning-based model. J Clin Med 2022 Dec 21;12(1):55. [doi: 10.3390/jcm12010055] [Medline: 36614860]
44. Zhang Y, Pei Y, Qin H, et al. Masseter muscle segmentation from cone-beam ct images using generative adversarial network. In: Zha H, editor. Presented at: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI); Apr 8-11, 2019; Venice, Italy. [doi: 10.1109/ISBI.2019.8759426]
45. Zhu W, Huang Y, Zeng L, et al. AnatomyNet: deep learning for fast and fully automated whole-volume segmentation of head and neck anatomy. Med Phys 2019 Feb;46(2):576-589. [doi: 10.1002/mp.13300] [Medline: 30480818]
46. Han W, Xia W, Zhang Z, et al. Radiomics and artificial intelligence study of masseter muscle segmentation in patients with hemifacial microsomia. J Craniofac Surg 2023;34(2):809-812. [doi: 10.1097/SCS.0000000000009105] [Medline: 36728424]
47. Chen W, Li Y, Dyer BA, et al. Deep learning vs. atlas-based models for fast auto-segmentation of the masticatory muscles on head and neck CT images. Radiat Oncol 2020 Jul 20;15(1):176. [doi: 10.1186/s13014-020-01617-0] [Medline: 32690103]

Abbreviations

AD: Alzheimer disease
ASSD: average symmetric surface distance
CSE: categorical cross-entropy
CT: computed tomography
DLB: dementia with Lewy bodies
DSC: Dice similarity coefficient
DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition
GT: ground truth

XSL•FO
RenderX

MRI: magnetic resonance imaging

Edited by M Gray; submitted 26.06.24; peer-reviewed by Q Gang, Z Zhao; revised version received 17.12.24; accepted 02.01.25; published 19.03.25.

Please cite as:
Imani M, Borda MG, Vogrin S, Meijering E, Aarsland D, Duque G
Using Deep Learning to Perform Automatic Quantitative Measurement of Masseter and Tongue Muscles in Persons With Dementia: Cross-Sectional Study
JMIR Aging 2025;8:e63686
URL: https://aging.jmir.org/2025/1/e63686
doi:10.2196/63686

© Mahdi Imani, Miguel G Borda, Sara Vogrin, Erik Meijering, Dag Aarsland, Gustavo Duque. Originally published in JMIR Aging (https://aging.jmir.org), 19.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Older Adults' Perspectives and Acceptance of AI-Driven Health Technologies: Qualitative Study

Arkers Kwan Ching Wong[1], PhD; Jessica Hiu Toon Lee[1], BSc; Yue Zhao[2], PhD; Qi Lu[2], PhD; Shulan Yang[3], MSc; Vivian Chi Ching Hui[1], PhD

[1]School of Nursing, The Hong Kong Polytechnic University, 1 Yuk Choi Rd, Hung Hom, Kowloon, Hong Kong, China (Hong Kong)

[2]School of Nursing, Tianjin Medical University, Tianjin, China

[3]Nursing Department, Zhejiang Hospital, Hangzhou, China

Corresponding Author:
Arkers Kwan Ching Wong, PhD
School of Nursing, The Hong Kong Polytechnic University, 1 Yuk Choi Rd, Hung Hom, Kowloon, Hong Kong, China (Hong Kong)

Abstract

Background: Artificial intelligence (AI) is increasingly being applied in various health care services due to its enhanced efficiency and accuracy. As the population ages, AI-based health technologies could be a potent tool in older adults' health care to address growing, complex, and challenging health needs. This study aimed to investigate perspectives on and acceptability of the use of AI-led health technologies among older adults and the potential challenges that they face in adopting them. The findings from this inquiry could inform the designing of more acceptable and user-friendly AI-based health technologies.

Objective: The objectives of the study were (1) to investigate the attitudes and perceptions of older adults toward the use of AI-based health technologies; (2) to identify potential facilitators, barriers, and challenges influencing older adults' preferences toward AI-based health technologies; and (3) to inform strategies that can promote and facilitate the use of AI-based health technologies among older adults.

Methods: This study adopted a qualitative descriptive design. A total of 27 community-dwelling older adults were recruited from a local community center. Three sessions of semistructured interviews were conducted, each lasting 1 hour. The sessions covered five key areas: (1) general impressions of AI-based health technologies; (2) previous experiences with AI-based health technologies; (3) perceptions and attitudes toward AI-based health technologies; (4) anticipated difficulties in using AI-based health technologies and underlying reasons; and (5) willingness, preferences, and motivations for accepting AI-based health technologies. Thematic analysis was applied for data analysis. The Theoretical Domains Framework and the Capability, Opportunity, Motivation, and Behavior (COM-B) model behavior change wheel were integrated into the analysis. Identified theoretical domains were mapped directly to the COM-B model to determine corresponding strategies for enhancing the acceptability of AI-based health technologies among older adults.

Results: The analysis identified 9 of the 14 Theoretical Domains Framework domains—knowledge, skills, social influences, environmental context and resources, beliefs about capabilities, beliefs about consequences, intentions, goals, and emotion. These domains were mapped to 6 components of the COM-B model. While most participants acknowledged the potential benefits of AI-based health technologies, they emphasized the irreplaceable role of human expertise and interaction. Participants expressed concerns about the usability of AI technologies, highlighting the need for user-friendly and tailored AI solutions. Privacy concerns and the importance of robust security measures were also emphasized as critical factors affecting their willingness to adopt AI-based health technologies.

Conclusions: Integrating AI as a supportive tool alongside health care providers, rather than regarding it as a replacement, was highlighted as a key strategy for promoting acceptance. Government support and clear guidelines are needed to promote ethical AI implementation in health care. These measures can improve health outcomes in the older adult population by encouraging the adoption of AI-driven health technologies.

(JMIR Aging 2025;8:e66778) doi:10.2196/66778

KEYWORDS

artificial intelligence–based health technologies; health technology; AI-based health technology; machine learning; ML; artificial intelligence; AI; algorithm; model; analytics; perceptions; acceptability; gerontology; geriatrics; older adult; elderly; older person; older people; aging; mobile phone

XSL•FO
RenderX

Introduction

Background

Artificial intelligence (AI) refers to a computerized system that is capable of executing a wide range of tasks that typically involve human intelligence—ranging from physical tasks, cognitive functions, and problem-solving to decision-making—and can be performed without explicit instructions from humans [1]. AI can be classified into 4 main subsets—machine learning, natural language processing (NLP), physical robots, and robotic process automation. The field of AI is rapidly advancing, and AI technologies have been widely applied in different aspects of health care services, including diagnostic assistance, health screening, and imaging interpretations [2]. It has been described as a "second set of eyes" for medical practitioners [3]. Furthermore, over the past decade, the application of AI systems has increasingly centered on empowering patients to actively manage their health and boost their participation in the shared decision-making process, giving rise to diverse AI-driven products such as robots, smart assistants, virtual or augmented reality, wearable devices, and mobile apps [4]. Particularly notable is the emergence of AI-powered chatbot services, which permit patients to seek medical advice and receive triage for their conditions in a prompt and cost-effective manner [5]. With the debut of ChatGPT by OpenAI in November 2022 and the popularity it has gained, it is envisioned that AI-based conversational large language models with NLP abilities could conceivably revolutionize health care practice and education, contributing to substantial transformative shifts [6,7].

The world is currently encountering a significant expansion in the aging population, with the number of people aged 60 years or above anticipated to rise to 1.4 billion by 2030 and 2.1 billion by 2050 [8]. Conventional older adult health care has relied heavily on in-person monitoring; however, the further intensification of the shortage of health care workers could in the long run present a global challenge to the sustainability of delivering quality medical care [9]. In recent years, digital technologies, such as smart older adult health care products, which involve the deep integration of AI, have been used in older adult health care sectors [10]. Evidence has shown that AI-based technologies can play a constructive role in improving the physical and psychological well-being, quality of life, and independent living of older adults [11]. However, there is a great digital divide between the older and younger generations, with the former being the age group that is the least likely to have access to computers and the internet, due to physical obstacles such as physical disabilities and to psychological factors such as a lack confidence in using technology as well as ethical concerns [12,13].

While AI-based health technologies have rapidly evolved, and some studies have comprehensively discussed the advantages of applying them in the field of geriatrics, insufficient academic attention has been paid to acquiring a clear understanding of the attitudes, perceptions, and experiences of seniors toward these technologies [10]. This gap may implicitly influence the acceptance of, and motivation to use, these technologies in the future. In addition, while studies have primarily focused on the perceptions of medical practitioners and patients regarding AI health technologies, the perspectives of older adults—a key user group—have received limited attention [14-16]. For instance, although some research has explored older adults' views on general AI-powered technologies, their relevance to AI health technologies remains uncertain [17]. Concerns about accuracy, reliability, and trust in AI tools compared with in-person medical advice further underscore the need for targeted investigation. This study addresses this gap by exploring older adults' perceptions of and acceptability toward AI-based health technologies, with a specific focus on practical strategies to enhance adoption. By identifying barriers, facilitators, and preferences, the findings have clear clinical implications, offering actionable insights for health care providers, policy makers, and AI developers. These insights can inform the design of tailored, user-friendly AI tools and guide their ethical implementation in health care, ultimately improving older adults' health outcomes.

Theoretical Framework

The interview guide for this study was developed using the Theoretical Domains Framework (TDF) and the Capability, Opportunity, Motivation, and Behavior (COM-B) model behavior change wheel to comprehensively explore factors influencing the acceptability of AI-based health technologies among older adults. The TDF consists of 14 domains that integrate behavioral determinants derived from over 30 psychological theories [18]. To ensure comprehensive coverage of these domains, the interview guide focused on five key areas: (1) general impressions of AI, (2) previous experiences with AI-based technologies, (3) perceptions and attitudes toward AI-based health technologies, (4) expected difficulties and underlying reasons, and (5) willingness. These focus areas were aligned with the study objectives and systematically mapped to the relevant TDF domains, as summarized in Table 1. This approach ensured that all 14 domains were addressed, either directly or indirectly, during the interviews.

XSL•FO
RenderX

Table . The interview questions guided by the Theoretical Domains Framework (TDF).

Focus area	Aligned TDF domains	Example question
General impressions of AI[a]	Knowledge, emotion, and optimism	"What comes to mind when you think about artificial intelligence and its role in health technologies?"
Previous experiences with AI-based technologies	Memory, attention, and decision processes; knowledge; and skills	"Have you used any AI-based technologies before? Can you describe your experience?"
Perceptions and attitudes toward AI-based health technologies	Social or professional role and identity; beliefs about capabilities; beliefs about consequences; emotion; and optimism	"What are your thoughts about using AI-based technologies for managing your health?"
Expected difficulties and underlying reasons	Environmental context and resources; skills; social influences; behavioral regulation; and memory, attention, and decision processes	"What challenges do you think you might face when using AI-based health technologies? Why do you think these occur?"
Willingness	Intentions; goals; beliefs about capabilities; and reinforcement	"Would you be willing to use AI-based health technologies? Why or why not?"

[a]AI: artificial intelligence.

To further enhance the applicability of findings, the identified TDF domains were mapped to the COM-B behavior change wheel (Table 2). The COM-B model summarizes 6 sources of behavior—social or physical opportunity, automatic or reflective motivation, and physical or psychological capability. This model provided a practical framework for identifying specific behavior change factors and strategies to improve the acceptance and adoption of AI-based health technologies among older adults [19].

Table . Capability, Opportunity, Motivation, and Behavior (COM-B) model components and its relation to Theoretical Domains Framework (TDF) domains.

COM-B components		TDF domains
Capability		
	Psychological capability	• Knowledge • Skills • Memory, attention, and decision processes
	Physical capability	• Behavioral regulation • Skills
Opportunity		
	Social opportunity	• Social influences
	Physical opportunity	• Environmental context and resources
Motivation		
	Reflective motivation	• Social or professional role and identity • Beliefs about capabilities • Optimism • Beliefs about consequences • Intentions • Goals
	Automatic motivation	• Social or professional role and identity • Optimism • Reinforcement • Emotion

The interview guide included open-ended questions designed to elicit rich, detailed responses.

This structured mapping ensured that the interview guide covered all 14 TDF domains comprehensively, while tailoring the questions to elicit insights into factors influencing the acceptability of AI-based health technologies among older adults. While alternative models, such as the technology acceptance model and the unified theory of acceptance and use of technology are commonly used in technology adoption studies, these models primarily focus on individual perceptions, such as perceived usefulness and ease of use (technology acceptance model) or performance expectancy and effort expectancy (unified theory of acceptance and use of technology).

XSL•FO
RenderX

These constructs, while valuable, may not fully capture the complex interplay of factors influencing older adults' adoption of AI technologies.

The systematic alignment between the interview guide and the TDF framework facilitated the identification of theoretical mediators and behavioral determinants. This mapping subsequently informed the integration of the findings with the COM-B behavior change wheel, allowing for the development of tailored strategies to promote the use of AI-based health technologies in this population.

Objectives

The objectives of the study were (1) to investigate the attitudes and perceptions of older adults toward the use of AI-based health technologies; (2) to identify potential facilitators, barriers, and challenges influencing older adults' preferences toward AI-based health technologies; and (3) to inform strategies that can promote and facilitate the use of AI-based health technologies among older adults.

Methods

Study Design

A qualitative descriptive design was adopted for this study [20]. This approach helps in the effort to uncover and understand the experiences, attitudes, and perceptions of people, making it appropriate for this study. The findings from a qualitative descriptive study are able to inform strategies that promote and facilitate the use of AI-based health technologies, which makes it particularly useful for this research. To explore perceptions on and acceptability of the use of AI-based technologies in health maintenance among older adults, and to provide insights into their subjective views, in-depth semistructured interviews guided by an interview guide were conducted.

Participants and Recruitment

Community-dwelling older adults aged 60 years or above were recruited from a local community center using a convenience sampling method. Eligible participants included those with or without previous exposure to AI tools who were willing to participate and share their experiences. Exclusion criteria included underlying physical, psychological, or neurodevelopmental problems that impaired interaction, mental instability, cognitive impairments such as traumatic brain injury, substance abuse, dementia, severe psychiatric disorders, intellectual disability, or being bedridden.

Sampling and Sample Size

Convenience sampling approach was employed, guided by the principle of data saturation. The target sample size was 25 participants, which is commonly regarded as sufficient for qualitative descriptive studies to achieve thematic depth and richness [21]. Ultimately, 27 older adults were recruited, providing adequate data to explore the study's focus areas.

Data Collection

Data collection involved 3 sessions of semistructured interviews with each participant, each lasting 1 hour. To minimize ambiguity, AI was classified into 4 categories: machine learning, NLP, physical robots, and robotic process automation. Interviews covered five focus areas: (1) general impressions of AI; (2) previous experiences with AI-based technologies; (3) perceptions and attitudes toward AI-based health technologies; (4) anticipated difficulties in using AI-based health technologies and their underlying reasons; and (5) willingness, preferences, and motivations in accepting AI-based health technologies.

In the first focus area, examples of AI tools were not initially provided to elicit spontaneous responses. However, recognizing that participants might not be fully aware of everyday interactions with AI, the second focus area included real-life examples to clarify potential misconceptions. Examples of machine learning technologies included fitness tracking devices, such as Google Fitbit and Apple Watch, which monitor physical activity and provide insights into health trends. For NLP, the interviewer referenced virtual assistants like Siri (Apple Inc) and Alexa (Amazon Inc), as well as health care chatbots that answer patient inquiries. Computer vision was illustrated with examples of AI systems that analyze medical images, such as detecting abnormalities in x-rays or CT scans. This approach enabled a nuanced exploration of participants' experiences with general AI technologies versus their attitudes toward AI for health purposes. The data were securely stored in Cloud storage, accessible only to the research team.

Data Analysis

Thematic analysis was used in this study [22]. The interview recordings were first translated into English, transcribed, read, and reread to gain a sense of the participants' experiences. Initial ideas were noted down. Next, the data were coded to extract key insights from the transcript. The codes were reviewed in a weekly discussion with the supervisor to obtain agreement on the interpretations. The underlying subthemes that emerged from the codes were identified and the subthemes were further clustered into themes. The themes, subthemes, and representative quotes were used to present the findings. Data were mapped to TDF to identify underlying theoretical mediators that influenced the perceptions and acceptability to the older adults of the use of AI-based health technologies. Once these theoretical domains were identified, they were mapped to components of the COM-B behavior change wheel that matched theoretical mediators with corresponding key strategies.

Ethical Considerations

The study protocol and procedures were approved by the Ethical Committee of the Hong Kong Polytechnic University (approval HSEARS20230810006) on August 29, 2023, and adhered to the ethical standards of the Declaration of Helsinki. Before starting the study, all participants provided informed consent during a face-to-face interview with the researcher. All the participants were fully informed about the nature and purpose of the study. They were guaranteed the right to withdraw from the study at any time without adverse consequences. The collected data were encrypted and stored in a password-protected database. Although the researchers did not foresee any significant risks associated with the proposed study, the researchers recognized that older adults might experience discomfort or fatigue from sitting through the interviews. To address this, the researchers ensured that participants were given

XSL•FO
RenderX

scheduled breaks throughout the interview. In addition, participants were encouraged to request breaks as needed to ensure that they were comfortable during the entire interview process. Participants who completed the interviews received an HK $100 (US $12.80) supermarket gift voucher to compensate for transportation costs.

Results

Participant Characteristics

The study included 27 participants, with 18 females (66.7%) and 9 males (33.3%). The mean age was 69.44 (SD 6.7) years. Most participants were married (19/27, 70.4%) and had a high school education (14/27, 51.9%). A majority lived with family (19/27, 70.4%) and rated their health as good (14/27, 58.3%). In addition, 66.7% (18/27) reported having chronic diseases. Regarding technology, 51.9% (14/27) felt somewhat comfortable using smartphones or tablets, and 59.3% (16/27) were somewhat familiar with tech products. Over 80% (23/27, 85.2%) of participants had heard of AI health technology, though their familiarity was limited. AI product usage varied, with approximately half of the participants reporting they used AI products often or occasionally, while the other half reported they used them seldom or were unfamiliar with them. Most participants (21/27, 77.8%) believed AI was to some extent helpful in health management, and 66.7% (18/27) had a positive overall impression of AI. Further details about the characteristics of the participants are provided in Table 3.

XSL•FO
RenderX

Table . Participant characteristics.

Characteristic		Total (N=27)
Gender, n (%)		
	Men	9 (33.3)
	Women	18 (66.7)
Age (y)		
	Mean (SD)	69.44 (6.17)
	Median (IQR)	70 (66-74)
Marital status, n (%)		
	Single	1 (3.7)
	Married	19 (70.4)
	Divorced	1 (3.7)
	Widowed	6 (22.2)
Education level, n (%)		
	Primary school and below	4 (14.8)
	Junior high school	7 (25.9)
	High school	14 (51.9)
	College or university	2 (7.4)
	Postgraduate degree	0 (0)
Living status, n (%)		
	Live with family	19 (70.4)
	Live alone	7 (25.9)
	Did not complete	1 (3.7)
Self-rated health status, n (%)		
	Poor	1 (3.7)
	Fair	12 (44.4)
	Good	14 (58.3)
Any chronic diseases, n (%)		
	Yes	18 (66.7)
	No	9 (33.3)
Comfort level with using smartphone or tablet, n (%)		
	Very comfortable	10 (37.0)
	Somewhat comfortable	14 (51.9)
	Not very comfortable	2 (7.4)
	Completely uncomfortable	0 (0)
	Did not complete	1 (3.7)
Familiarity with technology products, n (%)		
	Very familiar	0 (0)
	Somewhat familiar	16 (59.3)
	Not very familiar	8 (29.6)
	Completely unfamiliar	2 (7.4)
	Did not complete	1 (3.7)
Have heard of AI [a] health technology, n (%)		
	Yes, very familiar	0 (0)

XSL•FO
RenderX

Characteristic		Total (N=27)
	Yes, somewhat familiar	11 (40.7)
	Yes, not very familiar	12 (44.4)
	No, completely unfamiliar	3 (11.1)
	Did not complete	1 (3.7)
Frequency of AI product usage, n (%)		
	Often used	6 (22.2)
	Occasionally used	6 (22.2)
	Seldom used	5 (18.5)
	Never used	7 (25.9)
	Did not complete	3 (11.1)
Attitude toward AI's helpfulness in health management, n (%)		
	Very helpful	2 (7.4)
	To some extent helpful	21 (77.8)
	Not helpful	0 (0)
	Uncertain	2 (7.4)
	Did not complete	2 (7.4)
Overall impression about AI technology, n (%)		
	Positive	18 (66.7)
	Neutral	6 (22.2)
	Negative	0 (0)
	Uncertain	1 (3.7)
	Did not complete	2 (7.4)

[a]AI: artificial intelligence.

Objective 1: Investigate Older Adults' Attitudes and Perceptions Related to the Use of AI-Based Health Technologies

General Impressions of AI-Based Health Technologies

Most participants had positive views of AI-based health technologies, recognizing their potential to improve health outcomes by enhancing health monitoring, providing personalized recommendations, and assisting in decision-making. However, there was also limited awareness and understanding of specific AI-based health technologies, with some participants confusing them with general technologies such as smartphones. In addition, concerns were raised about fraud and scams associated with AI-based health technologies.

Attitudes and Perceptions Toward AI-Based Health Technologies

Privacy and data security were recurring themes in participants' discussions of AI-based health technologies. Many participants expressed concerns about the potential misuse of personal data and highlighted the importance of robust security measures. These concerns were particularly pronounced among participants with limited experience using technology or those influenced by media reports of data breaches. However, a subset of participants viewed privacy as a societal issue rather than a specific risk associated with AI-based technologies. For instance, one participant remarked:

> *Yes, it's acceptable.…Uh, what privacy do you have? There is no privacy in the whole world. Everyone's mobile phone is being monitored.* [Participant 3]

Another noted:

> *Not to mention privacy nowadays, from where we sit now there is completely no privacy…even if you listen to this song now, then this song will automatically appear (on your phone). When you are viewing clothes, then you will see the clothes later on (on your phone).* [Participant 5]

These contrasting perspectives highlight the varying levels of concern about privacy among older adults and underscore the need for transparent communication about data security in the development and deployment of AI-based health technologies.

Trustworthiness and Accuracy

Trust in AI-based health technologies varied among participants, influenced by previous experiences, information sources, and perceptions of accuracy. While some acknowledged the utility of AI in self-monitoring, others expressed reservations about its accuracy and trustworthiness, stressing the irreplaceable value of human judgment and expertise in health care.

XSL•FO
RenderX

Objective 2: Identify Potential Facilitators, Barriers, and Challenges That Influence Older Adults' Preferences Toward AI-Based Health Technologies

Technological Barriers

Participants identified multiple technological challenges that hindered their ability to use AI-based health technologies effectively. These challenges included operational difficulties, limited digital literacy, and lack of access to required resources, such as hardware or stable internet connectivity. These barriers were particularly evident among participants with physical or cognitive impairments due to age-related degeneration.

For instance, 1 participant described how physical limitations affected their interaction with technology:

> *I have degeneration, too. I could remember a lot of things before. But now I can't. My body movements are slow now, so it's difficult for me to use AI-based health technologies.* [Participant 3]

Another significant barrier was self-perceived incompetence in using digital tools, often leading to frustration or feelings of inadequacy:

> *"I feel like I am stupid after learning something...."* [Participant 7]

In addition, resource-related challenges, such as a lack of access to necessary hardware (eg, smartphones) and stable internet connections, were raised, further limiting participants' ability to engage with AI technologies.

Emotional and Psychological Barriers

Emotional and psychological factors emerged as significant barriers to the adoption of AI-based health technologies. These included fear of technology, skepticism regarding its reliability, and concerns about losing human connection in health care interactions. The prevalence of scams and fraudulent activities in the digital age further compounded participants' fears, making them hesitant to trust digital tools.

For example, 1 participant expressed a deep apprehension about interacting with technology:

> *I don't feel comfortable with these technologies because they seem complicated, and I worry about making mistakes.* [Participant 6]

Concerns about scams were also prevalent, with another participant noting:

> *Nowadays, you hear so much about fraud. It makes me scared to trust anything online, even if it looks helpful.* [Participant 4]

These findings illustrate that emotional and psychological barriers are not only about individual fears but also reflect broader societal concerns regarding trust and safety in the digital world.

Perceived Usefulness and Relevance

Participants perceived AI-based health technologies as useful, particularly those that addressed specific health needs such as the monitoring of blood pressure, blood sugar, and cholesterol. They also appreciated technologies that could assist with health monitoring, reminders, and personalized care.

> *I hope that these three things could be more mature, help to reduce blood pressure, blood sugar, or blood cholesterol, whatever. At least I know whether my blood sugar is good or not; if it is not good, I will eat something and exercise to increase my physical capacity.* [Participant 3]

> *Yeah, like an alarm. As a reminder to alert you. Well, high blood pressure, high blood sugar, high cholesterol level, those we get when are older. If AI can help with these three things, it is definitely good.* [Participant 2]

Acknowledgment by Authorities

The endorsement of AI-based health technologies by official authorities or regulatory bodies significantly facilitated their acceptance and adoption. Participants felt more confident and trusting when these technologies were validated by reputable sources, such as health care agencies or government bodies.

> *If it is something organized by the Hospital Authority or the government, it will be more trustworthy. Because sometimes, if it is made by pharmaceutical companies, it may not necessarily be that clear.* [Participant 2]

> *Well, if you talk about it being official, like the doctors and government, I will assume that this AI is reliable, you will basically feel at ease.* [Participant 1]

> *I have to see if the AI comes from a large or a small company. For example, Hong Kong Polytechnic University, Baptist University — they are more reliable then.* [Participant 4]

Objective 3: Inform Strategies That Promote and Facilitate the Use of AI-Based Health Technologies Among Older Adults

Integration of Human Expertise

While recognizing the potential benefits of AI, participants emphasized the need for AI technologies to complement rather than replace human expertise in health care. They suggested that AI could be useful as an auxiliary tool for making preliminary medical suggestions, but that the final decision-making should involve consultations with a human being.

> *I just ask the AI, but I can ignore its answers or I will think about it again. It is just a reference and I won't believe it completely. Well, when you see a doctor, you won't just see one doctor, you will see many many right?...AI is just the first contact point.* [Participant 2]

User-Friendly Design

To enhance the acceptability of AI-based health technologies, it is crucial to develop solutions that are user-friendly and tailored to the specific needs of older adults. A user-friendly design should include simplified and intuitive interfaces that

XSL•FO
RenderX

minimize cognitive load, with features such as large, high-contrast text and icons for better visibility and ease of use. Voice command functionality can further improve accessibility for users with limited dexterity or vision impairments. In addition, providing step-by-step tutorials, user manuals in multiple formats (eg, videos and print), and accessible customer support can address challenges related to digital literacy. These features collectively ensure that the technology is not only easy to use but also aligns with the physical and cognitive capabilities of older adults, thereby fostering greater adoption and satisfaction.

Addressing Privacy and Security Concerns

Given the concerns around privacy and data security, the development of AI-based health technologies should include robust security measures to protect the personal information of users. Addressing these concerns is essential to building trust and encouraging adoption among older adults.

Government and Regulatory Support

Participants highlighted the importance of government and regulatory support in promoting the ethical implementation of AI in health care. Clear guidelines and endorsements from authoritative bodies can foster trust and enhance the acceptability of AI-based health technologies among older adults.

Table 4 provides a detailed description of how the identified barriers and facilitators toward the use of AI-based health technologies were mapped to the TDF domains and corresponding COM-B components. These mappings were derived from participants' responses during the interviews. The barriers and facilitators were categorized based on their alignment with the TDF domains and subsequently mapped to the COM-B components to identify actionable strategies for behavior change.

Table . Identified Theoretical Domains Framework (TDF) domains and Capability, Opportunity, Motivation, and Behavior (COM-B) components related to the adoption of artificial intelligence (AI)–based health technologies.

Barrier or facilitator	COM-B components	TDF domains	Illustrative participant quote
Barrier: limited technological skills	Capability	Knowledge and skills	"I find it difficult to use these technologies because my body movements are slow, and I don't know how to start." [P3]
Facilitator: perceived usefulness and relevance	Motivation	Beliefs about consequences, intentions, and goals	"If AI can help monitor my blood pressure or remind me, it would be very helpful." [P2]
Barrier: lack of official recognition or endorsement	Opportunity	Social influences	"If it's recognized by the government or hospitals, I would trust it more." [P1]
Facilitator: confidence in trustworthiness and accuracy	Motivation	Beliefs about capabilities	"If the results are accurate, I can rely on it for some guidance." [P4]
Barrier: privacy and data security concerns	Motivation	Emotion	"I worry about my personal information being stolen if I use this technology." [P5]
Facilitator: AI as an auxiliary tool for initial suggestions	Motivation	Beliefs about consequences and intentions	"I can use AI for some advice, but I still prefer to consult a doctor for the final decision." [P2]

Discussion

Principal Findings

The aim of the study was to provide insights into the perceptions and acceptability of AI-based health technologies among older adults, using the TDF and the COM-B behavior change wheel as theoretical frameworks. These frameworks were instrumental in identifying key behavioral determinants and mapping them to actionable strategies for enhancing AI adoption. In total, 9 of the 14 TDF domains (knowledge, skills, social influences, environmental context and resources, beliefs about capabilities, beliefs about consequences, intentions, goals, and emotion) were identified and mapped to 6 COM-B components (psychological capability, physical capability, social opportunity, physical opportunity, reflective motivation, and automatic motivation). By applying these frameworks, the study provided a structured approach to understanding the factors influencing older adults' behavior toward AI adoption. The responses of the participants shed light on various aspects of the adoption and utilization of AI-based health technologies among older adults. First, the participants expressed a range of attitudes and perceptions toward AI-based health technologies, including curiosity, skepticism, and enthusiasm toward the use of AI in the management and maintenance of health. These perceptions align with the TDF domains of knowledge and beliefs about consequences, which underscore the need to enhance understanding of AI and clearly communicate its benefits to older adults. The interviews provided a deeper understanding of the factors influencing these attitudes and the potential benefits and concerns associated with AI adoption in the context of the participants' health. They shared their perceptions of AI technologies, addressing aspects such as reliability, trustworthiness, and their perception of the impacts on health care outcomes. These insights informed the reflective motivation component of the COM-B framework, highlighting the

XSL•FO
RenderX

importance of fostering trust and reliability in AI-based technologies to enhance adoption.

In addition, the study revealed an interesting finding regarding older adults' preferences for using AI in asking for advice and self-health maintenance. Contrary to expectations, participants expressed a higher level of trust and reliability in real persons, particularly physicians, when making decisions about pharmacological interventions such as changing drug dosages. Participants believed that AI could play a valuable role in providing information and acting as a reference tool for self-health maintenance. However, they expressed reservations about relying solely on AI for decisions related to medication management. This finding is linked to the TDF domain of social influences and the social opportunity component of the COM-B framework, as participants emphasized the irreplaceable role of human expertise and interpersonal trust in health care decision-making. They viewed AI as a useful source of information but considered the expertise and experience of physicians and other health care professionals to be more reliable and crucial in making decisions regarding medical interventions and treatment. The emotion domain of TDF also played a role, as participants' skepticism and emotional barriers highlighted the importance of addressing concerns through education and reassurance. The findings were also in line with a mixed methods study indicating that the lack of empathy and a professional human approach made AI chatbots less acceptable to some users [15]. This insight has significant implications for the development of AI-based health technologies. It might indicate that older adults desire a collaborative approach that combines the benefits of AI with the guidance and expertise of real-person health care providers [23-25]. Developers could focus on creating AI systems that could provide accurate and evidence-based information, acting as a reference tool to support older adults' self-health maintenance efforts. In this context, AI could assist older adults in accessing reliable and up-to-date information about medications, potential side effects, and alternative treatment options. AI-based systems could also offer personalized recommendations for lifestyle modifications, nonpharmacological interventions, and self-care strategies that align with the preferences and needs of older adults.

While privacy and confidentiality are often considered to be potential concerns when using AI in health care, surprisingly, many expressed a relaxed and relatively pessimistic attitude toward privacy. One of the reasons for this was that the perceived benefits outweighed the concerns. Most expressed the view that the potential benefits of AI-based health technologies have overshadowed any concerns about privacy or confidentiality. They may have viewed the potential improvement in their health outcomes or access to personalized care as more significant than the potential risks to their privacy. Another underlying reason may be that the participants may have had other concerns that they considered more important than privacy and confidentiality in the context of AI-based health technologies. For example, they may have been more focused on the usability, effectiveness, or reliability of the technology. Regardless, the absence of expressed concerns does not necessarily indicate the lack of importance of privacy and confidentiality. Rather, it highlights the need for researchers and developers to proactively address these concerns and ensure that robust privacy and security measures are in place when designing and implementing AI-based health technologies. It is crucial for future research and development efforts to emphasize the importance of privacy and confidentiality in AI-based health technologies, which involves implementing strong data protection measures, obtaining informed consent from users, and transparently communicating how their data will be used and safeguarded.

To enhance older adults' acceptance of AI-based health technologies, strategies should focus on improving technological skills through tailored training programs, securing official recognition and endorsement from health care authorities, addressing concerns over privacy and data security with robust protocols, emphasizing the trustworthiness and accuracy of AI tools, positioning AI as auxiliary aids for medical suggestions, highlighting the usefulness and relevance of AI applications, and providing emotional support to mitigate psychological barriers. These strategies are directly informed by the COM-B framework, targeting capability, opportunity, and motivation as the key drivers of behavior change. By implementing these targeted strategies, older adults might be encouraged to embrace AI technologies, leading to improved health care outcomes and increased engagement with innovative health care solutions. Meanwhile, government endorsements could also contribute to the establishment of guidelines, standards, and regulations that ensure the ethical use of AI in health care, addressing concerns related to privacy, data security, and quality assurance.

Limitations

The limitations of this study were the sample size and representativeness of participants. While we recognize that this sample size may not fully capture the diversity of the population, the study aimed to offer initial, in-depth insights into older adults' perceptions of AI-based health technologies. Future studies could address subgroup comparisons by incorporating stratified sampling to examine variations across different demographic or socioeconomic groups. In addition, while quantitative comparisons were not a focus of this study, future research could include mixed method approaches to statistically assess participant perceptions and examine their relationships with demographic and contextual factors by using validated survey tools. Last but not the least, many participants had limited before knowledge or experience with AI-driven health technologies. Some participants conflated AI technologies with general digital tools, such as mobile apps, due to a lack of familiarity with the distinction. To address this, the interviewer provided concise explanations and real-life examples of AI subsets, including machine learning, NLP, computer vision, and expert systems. While this approach helped participants better understand the discussion, their responses were still shaped by their interpretations and familiarity with technology. Consequently, the findings reflect older adults' perceptions and acceptance of health technologies they associate with AI, rather than definitive perspectives on AI-driven health technologies.

Conclusion

While the majority of participants recognized the potential advantages of AI-based health technologies, they underscored

XSL•FO
RenderX

the indispensable role of human expertise and interaction. Vital strategies for improving acceptability include crafting user-friendly and customized AI solutions, addressing privacy concerns, ensuring robust security measures, and integrating AI as a complementary tool alongside health care providers. Government backing and guidelines can play a pivotal role in advancing ethical AI integration in health care, fostering trust, and enhancing personalized care for older adults, ultimately leading to improved health outcomes in this population.

Conflicts of Interest

None declared.

References

1. Esmaeilzadeh P, Mirzaei T, Dharanikota S. Patients' perceptions toward human-artificial intelligence interaction in health care: experimental study. J Med Internet Res 2021 Nov 25;23(11):e25856. [doi: 10.2196/25856] [Medline: 34842535]
2. Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol 2017 Jun 21;2(4):230-243. [doi: 10.1136/svn-2017-000101] [Medline: 29507784]
3. Cai CJ, Winter S, Steiner D, Wilcox L, Terry M. "Hello AI": uncovering the onboarding needs of medical practitioners for human-AI collaborative decision-making. Proc ACM Hum-Comput Interact 2019 Nov 7;3(CSCW):1-24. [doi: 10.1145/3359206]
4. Zhang Z, Genc Y, Wang D, Ahsen ME, Fan X. Effect of AI explanations on human perceptions of patient-facing AI-powered healthcare systems. J Med Syst 2021 May 4;45(6):64. [doi: 10.1007/s10916-021-01743-6] [Medline: 33948743]
5. Fan X, Chao D, Zhang Z, Wang D, Li X, Tian F. Utilization of self-diagnosis health chatbots in real-world settings: case study. J Med Internet Res 2021 Jan 6;23(1):e19928. [doi: 10.2196/19928] [Medline: 33404508]
6. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. Healthcare (Basel) 2023 Mar 19;11(6):887. [doi: 10.3390/healthcare11060887] [Medline: 36981544]
7. Vaishya R, Misra A, Vaish A. ChatGPT: Is this version good for healthcare and research? Diabetes & Metabolic Syndrome: Clinical Research & Reviews 2023 Apr;17(4):102744. [doi: 10.1016/j.dsx.2023.102744]
8. Ageing. World Health Organization. URL: https://www.who.int/health-topics/ageing#tab=tab_1 [accessed 2025-01-31]
9. Ho A. Are we ready for artificial intelligence health monitoring in elder care? BMC Geriatr 2020 Sep 21;20(1):358. [doi: 10.1186/s12877-020-01764-9] [Medline: 32957946]
10. Lee CH, Wang C, Fan X, Li F, Chen CH. Artificial intelligence-enabled digital transformation in elderly healthcare field: scoping review. Advanced Engineering Informatics 2023 Jan;55:101874. [doi: 10.1016/j.aei.2023.101874]
11. Czaja SJ, Ceruso M. The promise of artificial intelligence in supporting an aging population. J Cogn Eng Decis Mak 2022 Dec;16(4):182-193. [doi: 10.1177/15553434221129914]
12. Ball C, Francis J, Huang KT, Kadylak T, Cotten SR, Rikard RV. The physical-digital divide: exploring the social gap between digital natives and physical natives. J Appl Gerontol 2019 Aug;38(8):1167-1184. [doi: 10.1177/0733464817732518] [Medline: 29165009]
13. Chu CH, Nyrup R, Leslie K, et al. Digital ageism: challenges and opportunities in artificial intelligence for older adults. Gerontologist 2022 Aug 12;62(7):947-955. [doi: 10.1093/geront/gnab167] [Medline: 35048111]
14. Fritsch SJ, Blankenheim A, Wahl A, et al. Attitudes and perception of artificial intelligence in healthcare: a cross-sectional survey among patients. Digit Health 2022;8:20552076221116772. [doi: 10.1177/20552076221116772] [Medline: 35983102]
15. Nadarzynski T, Miles O, Cowie A, Ridge D. Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: a mixed-methods study. Digit Health 2019;5:2055207619871808. [doi: 10.1177/2055207619871808] [Medline: 31467682]
16. Shinners L, Aggar C, Grace S, Smith S. Exploring healthcare professionals' understanding and experiences of artificial intelligence technology use in the delivery of healthcare: an integrative review. Health Informatics J 2020 Jun;26(2):1225-1236. [doi: 10.1177/1460458219874641] [Medline: 31566454]
17. Shandilya E, Fan M. Understanding older adults' perceptions and challenges in using AI-enabled everyday technologies. Presented at: Chinese CHI 2022: The Tenth International Symposium of Chinese CHI; Oct 22-23, 2022; Guangzhou, China p. 105-116. [doi: 10.1145/3565698.3565774]
18. Richardson M, Khouja CL, Sutcliffe K, Thomas J. Using the theoretical domains framework and the behavioural change wheel in an overarching synthesis of systematic reviews. BMJ Open 2019 Jun;9(6):e024950. [doi: 10.1136/bmjopen-2018-024950]
19. Buchanan H, Newton JT, Baker SR, Asimakopoulou K. Adopting the COM-B model and TDF framework in oral and dental research: a narrative review. Community Dent Oral Epidemiol 2021 Oct;49(5):385-393. [doi: 10.1111/cdoe.12677] [Medline: 34270110]
20. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [Medline: 10940958]
21. Boddy CR. Sample size for qualitative research. QMR 2016 Sep 12;19(4):426-432. [doi: 10.1108/QMR-06-2016-0053]

XSL•FO
RenderX

22. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
23. Wong AKC, Wong FKY. The psychological impact of a nurse-led proactive self-care program on independent, non-frail community-dwelling older adults: a randomized controlled trial. Int J Nurs Stud 2020 Oct;110:103724. [doi: 10.1016/j.ijnurstu.2020.103724] [Medline: 32777605]
24. Wong AKC, Wong FKY, Chow KKS, Wong SM, Bayuo J, Ho AKY. Effect of a mobile health application with nurse support on quality of life among community-dwelling older adults in Hong Kong: a randomized clinical trial. JAMA Netw Open 2022 Nov 1;5(11):e2241137. [doi: 10.1001/jamanetworkopen.2022.41137] [Medline: 36350651]
25. Wong AKC, Bayuo J, Su JJ, et al. Exploring the experiences of community-dwelling older adults on using wearable monitoring devices with regular support from community health workers, nurses, and social workers: qualitative descriptive study. J Med Internet Res 2024 Aug 7;26:e49403. [doi: 10.2196/49403] [Medline: 39110493]

Abbreviations

AI: artificial intelligence
COM-B: Capability, Opportunity, Motivation, and Behavior
NLP: natural language processing
TDF: Theoretical Domains Framework

Edited by F Puga; submitted 23.09.24; peer-reviewed by HJ Kim, L Zhou, A Nakagomi; revised version received 23.12.24; accepted 07.01.25; published 12.02.25.

Please cite as:
Wong AKC, Lee JHT, Zhao Y, Lu Q, Yang S, Hui VCC
Exploring Older Adults' Perspectives and Acceptance of AI-Driven Health Technologies: Qualitative Study
JMIR Aging 2025;8:e66778
URL: https://aging.jmir.org/2025/1/e66778
doi:10.2196/66778

© Arkers Kwan Ching Wong, Jessica Hiu Toon Lee, Yue Zhao, Qi Lu, Shulan Yang, Vivian Chi Ching Hui. Originally published in JMIR Aging (https://aging.jmir.org), 12.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Development and Feasibility Study of HOPE Model for Prediction of Depression Among Older Adults Using Wi-Fi-based Motion Sensor Data: Machine Learning Study

Shayan Nejadshamsi[1,2,3], PhD; Vania Karami[1,2,3], PhD; Negar Ghourchian[4], PhD; Narges Armanfard[1,5], PhD; Howard Bergman[2], MD, FRCP; Roland Grad[2], MD, CM, MSc; Machelle Wilchesky[2,3,6], PhD; Vladimir Khanassov[2], MD; Isabelle Vedel[2], MD, PhD; Samira Abbasgholizadeh Rahimi[1,2,3,7], BEng, PhD

[1]Mila—Quebec Artificial Intelligence Institute, Montreal, QC, Canada

[2]Family Medicine Department, Faculty of Medicine and Health Sciences, McGill University, Montreal, QC, Canada

[3]Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, QC, Canada

[4]Aerial Technologies, Montreal, QC, Canada

[5]Department of Electrical and Computer Engineering, Faculty of Engineering, McGill University, Montreal, QC, Canada

[6]Donald Berman Maimonides Centre for Research in Aging, Montreal, QC, Canada

[7]Faculty of Dental Medicine and Oral Health Sciences, McGill University, Montreal, Canada

Corresponding Author:
Samira Abbasgholizadeh Rahimi, BEng, PhD
Family Medicine Department
Faculty of Medicine and Health Sciences
McGill University
5858 Côte des Negies
Montreal, QC, H3S 1Z1
Canada
Phone: 1 5143987375
Email: samira.rahimi@mcgill.ca

Abstract

Background: Depression, characterized by persistent sadness and loss of interest in daily activities, greatly reduces quality of life. Early detection is vital for effective treatment and intervention. While many studies use wearable devices to classify depression based on physical activity, these often rely on intrusive methods. Additionally, most depression classification studies involve large participant groups and use single-stage classifiers without explainability.

Objective: This study aims to assess the feasibility of classifying depression using nonintrusive Wi-Fi–based motion sensor data using a novel machine learning model on a limited number of participants. We also conduct an explainability analysis to interpret the model's predictions and identify key features associated with depression classification.

Methods: In this study, we recruited adults aged 65 years and older through web-based and in-person methods, supported by a McGill University health care facility directory. Participants provided consent, and we collected 6 months of activity and sleep data via nonintrusive Wi-Fi–based sensors, along with Edmonton Frailty Scale and Geriatric Depression Scale data. For depression classification, we proposed a HOPE (Home-Based Older Adults' Depression Prediction) machine learning model with feature selection, dimensionality reduction, and classification stages, evaluating various model combinations using accuracy, sensitivity, precision, and F_1-score. Shapely addictive explanations and local interpretable model-agnostic explanations were used to explain the model's predictions.

Results: A total of 6 participants were enrolled in this study; however, 2 participants withdrew later due to internet connectivity issues. Among the 4 remaining participants, 3 participants were classified as not having depression, while 1 participant was identified as having depression. The most accurate classification model, which combined sequential forward selection for feature selection, principal component analysis for dimensionality reduction, and a decision tree for classification, achieved an accuracy of 87.5%, sensitivity of 90%, and precision of 88.3%, effectively distinguishing individuals with and those without depression. The explainability analysis revealed that the most influential features in depression classification, in order of importance, were

XSL•FO
RenderX

"average sleep duration," "total number of sleep interruptions," "percentage of nights with sleep interruptions," "average duration of sleep interruptions," and "Edmonton Frailty Scale."

Conclusions: The findings from this preliminary study demonstrate the feasibility of using Wi-Fi–based motion sensors for depression classification and highlight the effectiveness of our proposed HOPE machine learning model, even with a small sample size. These results suggest the potential for further research with a larger cohort for more comprehensive validation. Additionally, the nonintrusive data collection method and model architecture proposed in this study offer promising applications in remote health monitoring, particularly for older adults who may face challenges in using wearable devices. Furthermore, the importance of sleep patterns identified in our explainability analysis aligns with findings from previous research, emphasizing the need for more in-depth studies on the role of sleep in mental health, as suggested in the explainable machine learning study.

(JMIR Aging 2025;8:e67715) doi:10.2196/67715

KEYWORDS

depression; classification; machine learning; artificial intelligence; older adults

Introduction

Depression is a prevalent mental health disorder characterized by emotional dysregulation, leading to persistent sadness, loss of interest, and anhedonia [1-3]. The rising incidence of depression among older adults has become a significant public health issue [4-6]. Early detection of depression and corresponding intervention are vital for improving mental health outcomes and reducing the overall burden on individuals and health care systems [7-9]. Traditional methods for assessing depression include various approaches that typically require in-person evaluations, specialized training in comprehensive geriatric assessments, and reliance on clinical judgment and questionnaires, which can be challenging and resource-intensive [10-12]. These methods require older adults to visit clinical settings frequently, increasing strain on health care facilities and reducing data collection opportunities. Additionally, many older adults prefer to remain in their homes and be remotely monitored in that environment, highlighting the need for remote care solutions in this demographic [13,14].

Physical activity and mobility are among the important factors in evaluating depression, with strong correlations established between these parameters and depression assessments [1,15]. The advent of the Internet of Things has enabled continuous and remote monitoring of physical activity. Several studies have used statistical methods to analyze the relationship between physical activity, as measured by wearable devices, and depression [16-21]. As the field of artificial intelligence (AI) advances, machine learning models have emerged as promising tools for depression classification using physical activity data [22]. For instance, Adamczyk and Malawski [23] used data from wearable actigraph watches in 3 classification models: logistic regression (LR), support vector machine (SVM), and random forest (RF) comparing automatic and manual feature engineering for depression classification. Bai et al [24] used phone use, sleep data, and step counts from 334 participants, using 2 feature selection methods (L1-based feature selection) and 6 machine learning models (decision tree [DT], k-nearest neighbors, naive Bayes, LR, SVM, and RF) for mood classification. Chikersal et al [25] analyzed data from smartphones and fitness trackers of 138 college students to identify those experiencing depressive symptoms, using nested randomized LR for feature selection and AdaBoost with gradient boosting classifier. Dai et al [26] used heart rate, energy expenditure, sleep, and other activity data from wearable Fitbit devices for depression remission detection in 106 participants within 2 intervention and control groups, using a multitask learning algorithm comprising 2 dense layers with shared parameters. Similarly, Griffiths et al [27] classified depression using activity and sleep data from Fitbit devices of 24 participants through an RF model. Espino-Salinas et al [28] used wrist-worn accelerometers to measure physical activity in 55 participants, applying a 2D-convolutional neural network (CNN) and a deep neural network for depression classification. Jakobsen et al [29] used RF, deep neural network, and CNN algorithms for depression classification with wrist-worn actigraph data from 55 participants. Jung et al [30] used gait accelerometry data and a bidirectional long short-term memory network–based classifier to assess depression in 45 older adults. Other studies have also explored the use of wearable device data combined with various classification methods, such as 1D-CNN [31], deep convolutional neuro-fuzzy [32], Ensemble models [33], and extreme gradient boosting [34].

These recent studies highlight the integration of wearable devices with machine learning algorithms as a promising approach for continuous, remote, home-based monitoring and early detection of depression. However, wearable device-based approaches face challenges due to their intrusive nature [35]. Participants are required to wear sensors or devices, which may lead to issues with compliance, comfort, and data accuracy, especially over extended periods [36], particularly among older adults [37]. These challenges can result in inaccurate data collection [38]. Nonwearable methods, such as remote monitoring through ambient sensors, offer potential solutions to these issues [37,39,40]. These approaches can alleviate concerns related to device adherence and physical discomfort, providing a more seamless integration into daily life. Additionally, our literature review indicates that most studies on depression classification use a relatively large number of participants and primarily use single-stage classifiers. Many studies also focus on detailed aspects of physical activity data (eg, body displacement, acceleration) using intrusive wearable sensors, which, while effective, present challenges related to user comfort and compliance. This reliance on wearable technology underscores the need for exploring alternative, less intrusive methods.

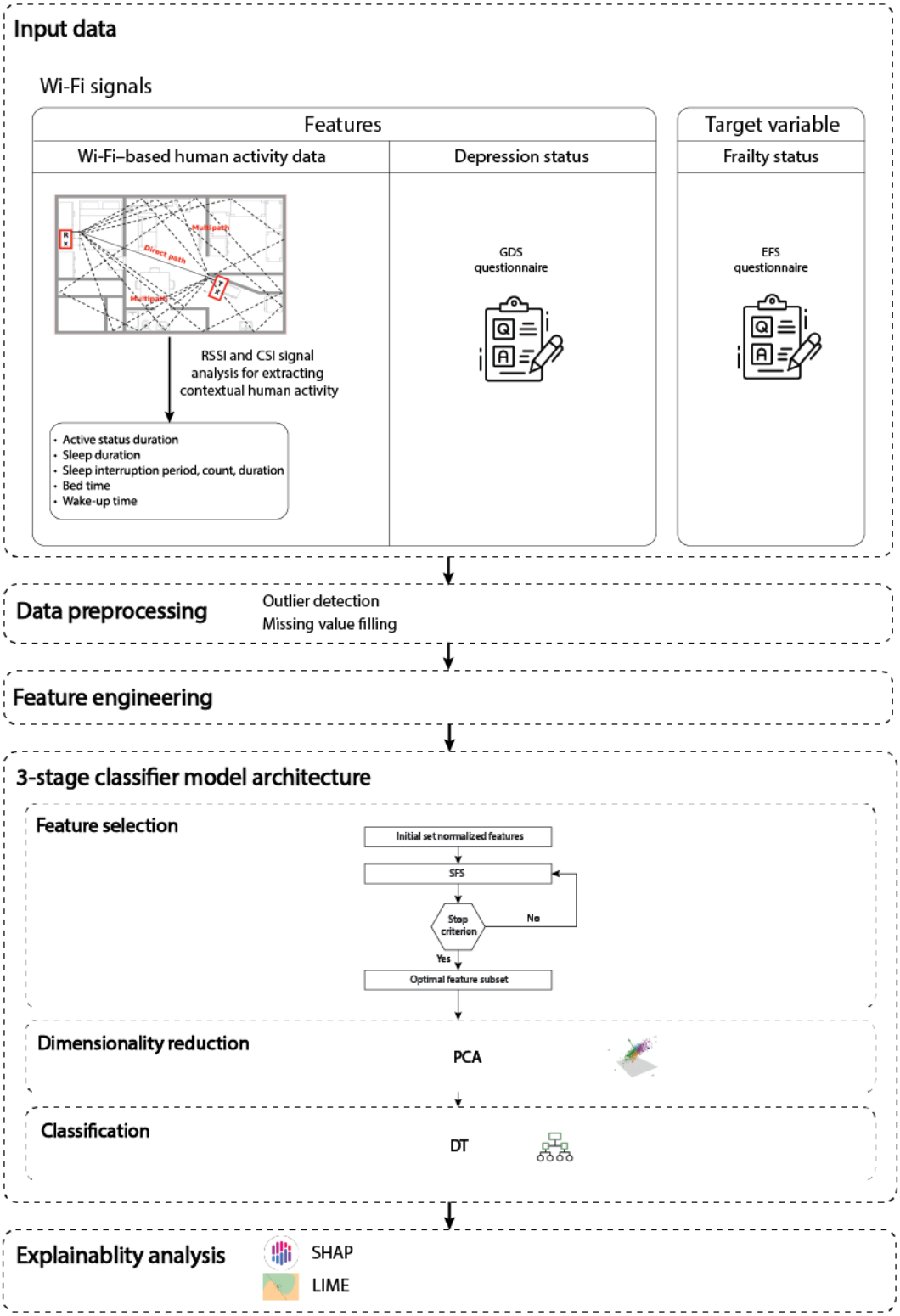
XSL•FO
RenderX

Recently, Wi-Fi–based sensing in smart homes has emerged as an alternative method for detecting and monitoring contextual human activity and movement [41]. Wi-Fi–based technologies are increasingly adopted due to their existing infrastructure in homes and minimal additional setup costs [42]. These technologies use signal metrics such as received signal strength indicator (RSSI) and channel state information (CSI) to analyze Wi-Fi signal characteristics, offering human activity identification compared to invasive wearable sensors, image analysis, or video-based systems [43]. Leveraging Wi-Fi–based activity data with machine learning presents a viable approach for different diseases, specifically depression classification, which is the focus of this study.

Technological and AI-based methods for depression classification predominantly rely on wearable sensors and are often conducted on large participant groups and typically use statistical techniques or single-stage machine learning classifiers. Furthermore, these approaches often overlook the explainability analysis of the models, which causes a lack of understanding of the underlying decision-making processes and the contribution of individual features to the model's predictions. To address these challenges, this feasibility study introduces a novel 3-stage machine learning model that incorporates feature selection, dimensionality reduction, and classification for depression classification. This model is specifically designed for a limited number of participants using low cost, easily installable Wi-Fi data, which provides continuous insights into human indoor activities. In addition to creating a model that functions effectively with small sample sizes, this study integrates explainable AI techniques to enhance the interpretability of the model's predictions. This approach ensures that the insights derived from the model are transparent and comprehensible, providing clarity on how specific features contribute to the classification outcomes. Figure 1 illustrates a schematic overview of our proposed framework.

XSL•FO
RenderX

Figure 1. Structure of the automatic Wi-Fi–based depression classification framework. CSI: channel state information; DT: decision tree; EFS: Edmonton Frailty Scale; GDS: Geriatric Depression Scale; LIME: local interpretable model-agnostic explanations; PCA: principal component analysis; RSSI: received signal strength indicator; SFS: sequential forward selection; SHAP: Shapely addictive explanations.



Methods

Data Acquisition

We begin by outlining the data acquisition process for the participants, followed by an in-depth explanation of our proposed HOPE model.

Study Cohort

Our recruitment approach used both digital and in-person strategies. Digital outreach was conducted through email campaigns, social media platforms, and digital posters. Prospective participants were provided with detailed information about the study, and those who expressed interest received a consent form. A member of the research team then coordinated

XSL•FO
RenderX

the setup of the monitoring equipment. Participants were compensated for their time and involvement with an e-gift card.

Participants in this study were required to be aged 65 years or older, capable of communicating in English or French, and have access to an internet connection at home. Exclusion criteria were as follows: (1) individuals with mental or physical conditions that would impede their ability to participate in the study and its 6-month follow-up, such as gait or balance disorders, active mental health issues, or the use of mobility aids such as canes; and (2) individuals with current substance use disorder, including alcohol or drugs, due to their potential impact on physical mobility. However, individuals with a history of substance use disorder who were no longer consuming were considered eligible.

Experimental Protocol

This study used a nonintrusive Wi-Fi–based motion sensor system to facilitate remote monitoring of human activity. Initially, we collected demographic information (ie, age and gender) from the participants. Subsequently, Wi-Fi data were acquired through our remote monitoring technology [44]. This device, installed at the network access point, detects Wi-Fi signals in environments conducive to passive sensing, including private residences and public spaces where Wi-Fi is prevalent [45]. In such indoor settings, Wi-Fi signals exhibit stability in the absence of individuals but fluctuate significantly with the presence and movement of people [45]. These signal variations correspond to distinct patterns associated with human movements and activities, thus providing valuable data for activity monitoring [46]. The collected data were subsequently transferred to secure cloud storage via an internet connection. Our team has developed advanced signal processing and AI-based algorithms to process raw Wi-Fi RSSI and CSI measurements. These algorithms standardize signal variations and translate RSSI and CSI fluctuations into a detailed set of contextual information related to human activity [45]. This information encompasses daily activity duration, bedtime, wake-up time, total sleep duration, and sleep interruption information. In addition to using Wi-Fi–based activity data, frailty and depression statuses were also assessed using validated assessment tools. The Edmonton Frailty Scale (EFS), which measures multiple dimensions of frailty [47-49], and the 15-item Geriatric Depression Scale (GDS) [50] were administered at the end of the experiment.

Analytical Framework

The methodical steps of our proposed model are detailed in the following subsections. During the development, implementation, and reporting, we adhered to the Minimum Information About Clinical Artificial Intelligence Modeling (MI-CLAIM) guidelines [51], following best practices designed to promote transparency and reproducibility of our AI model.

Data Preparation and Feature Extraction

To prepare the contextual human activity data as input for depression classification, obtained from our Wi-Fi signal analysis software, we designed a preprocessing stage. This process involves handling missing values and outliers to ensure the integrity of the data [52,53]. To enhance analytical depth and improve model performance, we implemented feature engineering on contextual human activity data. This process involves extracting a variety of new features, such as the mean and SD of bedtime and wake-up times, mean and SD of sleep duration (in hours), total count and mean of sleep interruptions, total duration and mean duration of sleep interruptions (in hours), longest continuous sleep duration (in hours), percentage of nights with sleep disturbances, and metrics related to daily activity, including the mean and SD of total daily activity and hourly activity durations, as well as peak activity hour. These derived features, combined with EFS data, were incorporated into our 3-stage machine learning classification model. The depression status, determined using the GDS, was used to label the samples for classification purposes.

HOPE Model Development

The proposed HOPE Model was designed for depression classification in older adults using nonintrusive Wi-Fi–based motion sensor data. Due to the limited number of participants, it is necessary to provide an efficient pipeline for preprocessing, feature extraction, and classification. The limited number of participants and high dimensionality of features required a tailored multistage machine learning pipeline to maximize classification accuracy. Furthermore, ensuring that our depression classification model is explainable to clinicians is important, as highlighted in our previous works [54,55]. To address this, we incorporated explainable machine learning techniques such as Shapley additive explanations (SHAP) and local interpretable model-agnostic explanations (LIME). To achieve these goals, our proposed HOPE model was structured into 3 stages of machine learning architecture: feature selection, dimensionality reduction, and classification followed by post hoc explainability analysis using SHAP and LIME. Each stage plays a critical role in refining the data and ensuring that the final classification is both accurate and interpretable. Table 1 provides details on the various techniques used and evaluated at each stage.

XSL•FO
RenderX

Table 1. Methods used at each phase of our 3-stage architecture.

Feature selection	Dimensionality reduction	Classification	Explainability analysis
CFS[a] [56]	PCA[b] [57]	NB[c] [58]	SHAP[d] [59]
SFS[e] [60]	FA[f] [61]	LR[g] [62]	LIME[h] [63]
MI[i] [64]	LDA[j] [65]	kNN[k] [66]	—[l]
SelectKBest [67]	kPCA[m] [68]	SVM[n] [69]	—
RFE[o] [70]	—	Decision tree [71]	—
—	—	RF[p] [72]	—
—	—	GBM[q] [73]	—
—	—	XGboost [74]	—
—	—	LightGBM [75]	—
—	—	Voting classifier [76]	—
—	—	Bagging classifier [77]	—
—	—	AdaBoost [78]	—

[a]CFS: correlation-based selection.
[b]PCA: principal component analysis.
[c]NB: naive Bayes.
[d]SHAP: Shapley addictive explanations.
[e]SFS: sequential forward selection.
[f]FA: factor analysis.
[g]LR: logistic regression.
[h]LIME: local interpretable model-agnostic explanations.
[i]MI: mutual information.
[j]LDA: linear discriminant analysis.
[k]kNN: k-nearest neighbor.
[l]Not applicable.
[m]kPCA: kernel principal component analysis.
[n]SVM: support vector machine.
[o]RFE: recursive feature elimination.
[p]RF: random forest.
[q]GBM: gradient boosting machine.

Feature selection is performed to reduce the dimensionality of the dataset by identifying the most relevant features for depression classification, enhancing both the speed and accuracy of the classification model [79]. The reduced subset of features serves as input for the subsequent dimensionality reduction stage. The validity of the chosen selected features is investigated using correlation analysis in the Results section. Dimensionality reduction techniques are applied to further refine the feature set compared to the initial feature selection stage [80] and to minimize overfitting. The dimensionally reduced features from the second stage were then processed in the third stage, which focused on classification. In this stage, the classification model processes the features derived from the earlier stages to categorize samples into 2 target classes: "participants with depression" and "participants without depression." The classification algorithm leverages the patterns identified in the features, such as sleep duration and interruptions, to make predictions. The classification task involved assigning a probability score to each sample, determining the likelihood of belonging to either class based on the relationships in the feature set. A decision boundary was then established to assign the final class label for each sample. The classification process was systematically evaluated to ensure robustness and reliability, focusing on separating the 2 groups effectively even with the small dataset. The machine learning classification pipeline was designed to minimize the risk of overfitting by using techniques such as feature selection and dimensionality reduction, ensuring that the most informative and relevant features were used for prediction. To conclude, we used explainable AI techniques to interpret the model's predictions, focusing on identifying the most influential features and their impact on classification outcomes.

Following the training and evaluation of all potential combinations for each stage, the architecture using sequential forward selection (SFS) for feature selection, principal component analysis (PCA) for dimensionality reduction, and DT for classification emerged as the most effective configuration. This SFS-PCA-DT framework (Figure 1)

XSL•FO
RenderX

demonstrated superior performance compared to other combinations.

Our proposed model is supported by different considerations. The SFS algorithm incrementally selects features to improve classification performance and is particularly suited for datasets with a smaller number of participants [81]. Unlike filter methods, SFS acts as a wrapper technique that pairs with a machine learning classification algorithm, providing greater stability in performance [81]. The method starts with no selected features and progressively adds them based on their ability to enhance cross-validation outcomes. In the second stage, PCA serves as a highly effective tool for reducing the dimensionality of data in an unsupervised manner [82]. It converts the initial set of features into a reduced number of uncorrelated components, maintaining the bulk of the data's variance. This reduction is important for preventing overfitting, especially with limited sample sizes. As a classification algorithm, DT algorithm, known for its strength in binary classification tasks, effectively uses the streamlined feature set generated in earlier stages. It models data by learning simple decision rules inferred from the input features, creating a tree-like structure. Each internal node in the tree represents a decision based on a feature, each branch represents an outcome of the decision, and each leaf node represents a class label [83]. The integration of SFS, PCA, and DT results in an efficient model that aligns with established methodologies and theoretical principles in the field. In this study, comprising only 4 participants, the training and validation procedure were carefully designed to minimize overfitting and to achieve reliable model generalization. To this end, we used a 4-fold cross-validation strategy. Each fold consisted of 3 participants for training and 1 participant for testing, ensuring that every participant contributed to both training and testing in separate iterations. This approach was repeated 10 times with different random seeds to account for variations in the training process, further enhancing the robustness of the performance metrics. During the training phase, a range of hyperparameter optimization techniques, including random search [84], Bayesian optimization [85], and Hyperband [86] was performed for each component of the 3-stage pipeline. For example, the DT classifier's maximum depth, minimum sample split, and criterion parameters are tuned using Bayesian optimization within predefined search spaces. Similarly, for PCA, the optimal number of components is optimized to maximize variance retention while preventing overfitting. The SFS algorithm was guided by internal cross-validation within the training set to identify the most predictive subset of features.

Evaluation Metrics

To validate the effectiveness of our depression classification method, we used 4 evaluation metrics: accuracy [87], sensitivity [87], precision [87], and F_1-score. Accuracy provides a comprehensive measure of the model's overall performance. Sensitivity helps ensure that the model accurately identifies as many true cases of depression as possible, minimizing the risk of missing individuals who actually have the condition [88]. To further validate the stability of the model, we present the training and test accuracies against the hyperparameter variations, demonstrating the model convergence.

Ethical Considerations

The study received approval from McGill University's Institutional Ethics Committee (A06-B18-21A), allowing data collection and analysis for this project. Written informed consent was collected from all participants prior to their involvement in the study. All collected data were anonymized immediately after collection, with no personally identifiable information retained to ensure participant confidentiality. Participants received a $20 e-gift card as compensation for their time. At every stage of the research, we adhered to the ethical principles outlined in the Declaration of Helsinki [89] and the Tri-Council Policy Statement [90].

Results

Clinical Study Insights

Six community-dwelling older adults residing in Montreal, Canada, were recruited for this study between May 2022 and September 2022. However, 2 participants withdrew due to internet connectivity issues, resulting in the use of data from the remaining 4 participants for the analysis. The EFS results indicated that 2 participants exhibited moderate frailty (scores ranging from 6 to 11), while the other 2 participants were classified as nonfrail (scores of 5 or below). GDS results suggested that 1 participant had depressive symptoms (score=10), while the other 3 participants did not (score range=0-4).

Some of the participants' demographic and clinical characteristics are detailed in Table 2. Over a 6-month period, the activities of each participant were continuously monitored at 15-minute intervals using Wi-Fi motion sensors. Following the identification of potential input features derived from Wi-Fi signals and questionnaire data, we designed and developed a 3-stage architecture as outlined in the methodology section. To support future research and ensure the reproducibility of our findings, the code for our model is openly accessible on our lab's GitHub repository [91].

Table 2. Demographic and clinical characteristics of the participants of this study.

	Participant without depression (n=3)	Participant with depression (n=1)
Sex (female/male), n/n	1/2	1/0
Age (years), mean (SD)	67.05 (3.70)	65.50 (0.00)
Edmonton frailty scale, mean (SD)	3.34 (3.21)	7.00 (0.00)
Geriatric depression scale, mean (SD)	4.00 (1.00)	10.00 (0.00)

XSL•FO
RenderX

We assessed a total of 240 model configurations, which resulted from the combination of 5 feature selection methods, 4 dimensionality reduction techniques, and 12 classifiers. Each configuration was trained 10 times to account for variations in performance metrics, ensuring the robustness of our findings. To mitigate the overfitting risk, we used k-fold cross-validation [92]. We experimented with different initial feature sets to optimize model performance, and among the various hyperparameter tuning methods, Bayesian optimization consistently yielded superior results.

Model Performance and Validation

This section presents the outcomes of our classification model evaluation. Table 3 summarizes the top-performing configurations among the 240 model variations that we tested. The SFS-PCA-DT model, which integrates SFS, PCA, and DT, emerged as the leading performer across multiple metrics. Its relatively high accuracy indicates the model's ability to distinguish between individuals with and without depression. The model's high sensitivity ensures that individuals with depression are correctly identified. This is critical in clinical settings, where depression, particularly among older adults, often goes unrecognized despite its severe impact on cognitive function [93-96], quality of life [97,98], and mortality risk [99,100]. Ensuring high sensitivity reduces the likelihood of missed diagnoses, which is important for timely and effective treatment. However, the relatively high standard deviation in both accuracy and sensitivity suggests that the model's performance may vary, indicating occasional instances of less reliable predictions.

Table 3. Average classification performance of the top 5 architectures.

Model	Accuracy (%)	Sensitivity (%)	Precision (%)	F_1-score (%)
SFS[a] – PCA[b] – DT[c]	87.50 (12.50)	90.00 (20.00)	88.34 (18.34)	86.00 (14.74)
SFS + FA[d] + DT	85.00 (16.58)	83.34 (25.82)	90.00 (20.00)	81.67 (18.93)
SFS + PCA + LR[e]	82.50 (19.53)	85.00 (22.91)	83.34 (25.82)	80.00 (20.82)
SFS + PCA + SVM[f]	82.50 (19.53)	90.00 (20.00)	75.00 (22.42)	80.00 (20.82)
MI[g] + PCA + LR	80.00 (10.00)	85.00 (22.91)	78.34 (22.42)	76.00 (13.06)

[a]SFS: sequential forward selection.

[b]PCA: principal component analysis.

[c]DT: decision tree.

[d]FA: factor analysis.

[e]LR: logistic regression.

[f]SVM: support vector machine.

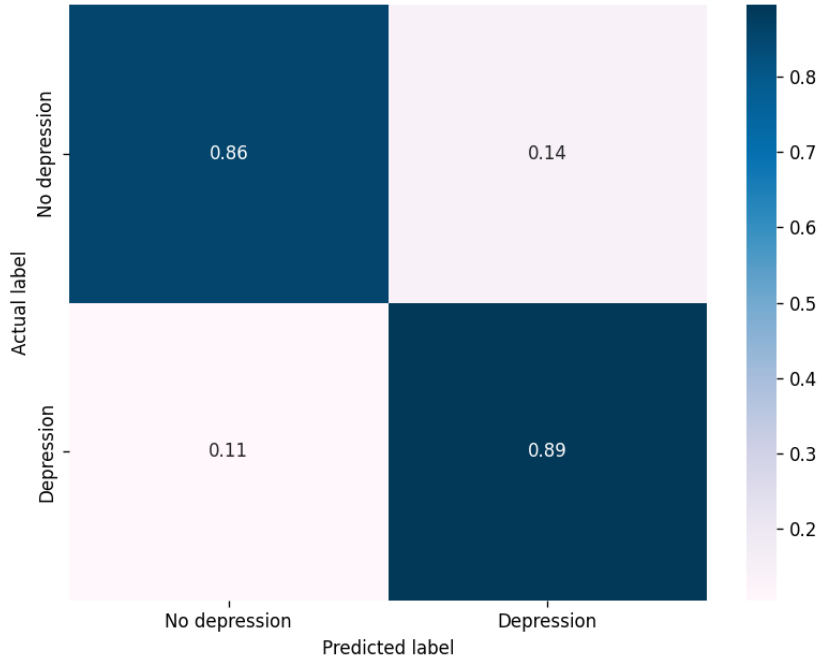
[g]MI: mutual information.

Table 3 illustrates that SFS is frequently featured among the highest-performing models, including the top model with 87.50% accuracy, and 3 other strong contenders. Mutual information also proves effective, appearing in 1 model with 80.00% accuracy, indicating that both SFS and mutual information are potent feature selection techniques for depression classification with limited samples. PCA seems to be the preferred method for dimensionality reduction, being used in 4 of the top 5 models. Among classifiers, DT stands out, featuring in the top 2 models with 87.50% and 85.00% accuracy. Other classifiers, such as LR and SVM, also perform well, each appearing in models with an accuracy exceeding 80%. The results highlight interesting tradeoffs, such as the SFS + PCA + SVM model, which, while slightly lower in accuracy (82.50%), maintains a high sensitivity (90.00%). This supports the practice of evaluating models using multiple metrics, especially in situations where the application involves clinical diagnoses, where accurately identifying true positives is crucial.

Figure 2 displays the averaged confusion matrix for the top-performing SFS – PCA – DT model used to classify depression status. Due to the constraint of having only 4 samples, we used a 4-fold cross-validation strategy, with each fold being tested 10 times to ensure a thorough evaluation. The model showed promising results, accurately classifying 86% of individuals without depression as not having depression and 89% of individuals with depression as having depression. The average false positive rate, where individuals without depression were incorrectly classified as having it, was 14%, while the average false negative rate, where individuals with depression were incorrectly classified as not having it, was 11%.

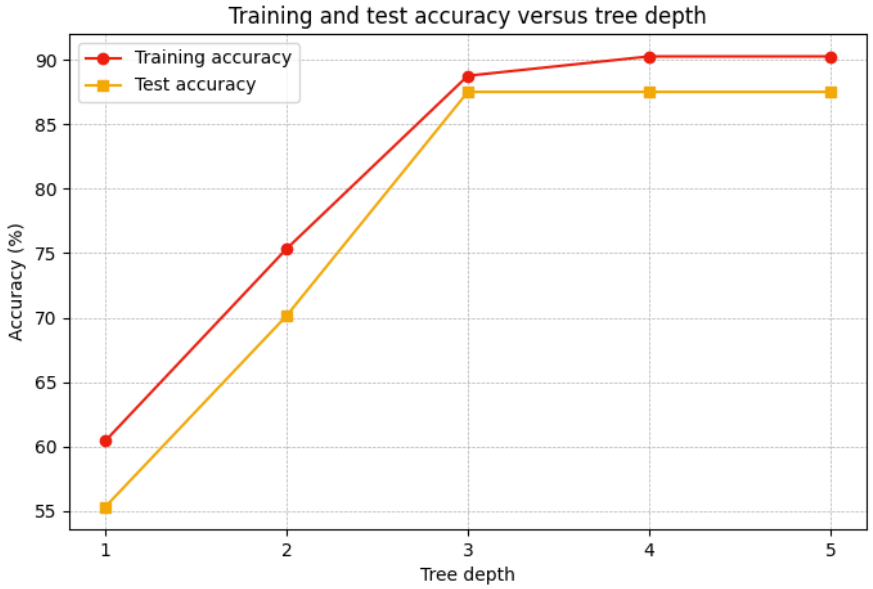
XSL•FO
RenderX

Figure 2. Confusion matrix for the top-performing model (SFS – PCA – DT). DT: decision tree; PCA: principal component analysis; SFS: sequential forward selection.



To validate the convergence of the proposed algorithm, we analyzed the relationship between tree depth and accuracy on both the training and test datasets. Figure 3 demonstrates the training and test accuracy of the best performing model (SFS – PCA – DT) as a function of tree depth. The training accuracy increases consistently with tree depth, stabilizing at its maximum, reflecting that the model can fully capture the training data as depth increases. The test accuracy improves initially with increasing tree depth but stabilizes beyond a depth of 3. These observations confirm that the proposed algorithm achieves convergence in terms of performance tradeoffs between model complexity and generalization.

Figure 3. Training and test accuracy as a function of tree depth, demonstrating convergence of our proposed model.



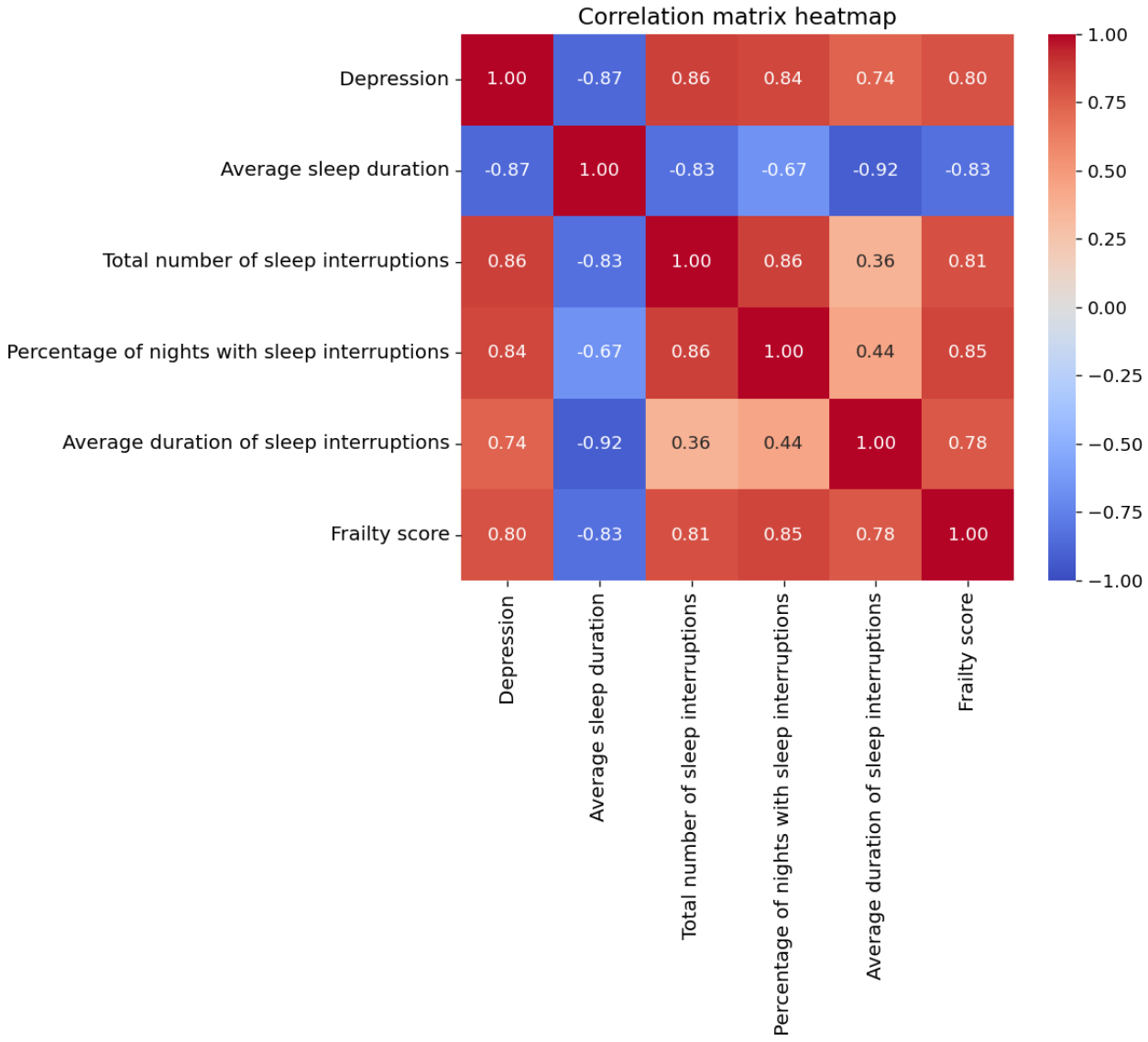
Feature Selection and Analysis

In our 3-stage classification model, we implement a combination of feature selection and dimensionality reduction techniques to improve the efficacy of our machine learning approach [80]. The features selected by SFS included “average sleep duration,” “total number of sleep interruptions,” “percentage of nights with sleep interruptions,” “average duration of sleep interruptions,” and “EFS.” Correlation analysis of these features revealed notable associations with depression status. Such analysis helps identify how variations in these features might be related to changes in depression, providing valuable insights for clinicians and researchers to develop more effective diagnostic tools and treatments.

XSL•FO
RenderX

Figure 4 illustrates a strong negative correlation between depression and average sleep duration. Conversely, depression was positively correlated with the total number of sleep interruptions, percentage of nights with sleep interruptions, average duration of sleep interruptions, and EFS. Blue cells indicate negative correlation values, while red cells represent positive correlations. Darker colors signify stronger correlations. The high correlation values highlight the significance of these factors in understanding and potentially classifying older adults with depression, aligning with findings from previous studies [101-105]. For example, Vallance et al [106] demonstrated that engaging in daily activities can alleviate the adverse effects of depression among older adults. Furthermore, several studies have highlighted a connection between depression and frailty [107,108]. Vaughan et al [109] showed that the prevalence of both depression and frailty among individuals aged 55 years and older exceeds 10%. These findings confirm the association between depression and the features incorporated in our model.

Figure 4. Correlation matrix heatmap between depression and selected features by sequential forward selection.



Comparative Analysis With Baseline Models

In this section, we evaluate the performance of our proposed 3-stage architecture for depression classification, which leverages Wi-Fi–based contextual human activity data, against baseline models previously outlined in the introduction. Although direct comparisons are inherently difficult due to differences in data acquisition methods (wearable devices), feature sets, and sample sizes, this analysis serves to contextualize the effectiveness of our approach for our case study with a limited number of participants.

For the baseline models, we selected the most current classification architectures used in the literature for depression classification. To ensure a fair comparison, each baseline model was trained and tested using the same feature set applied in our experiment. The resulting performance metrics for each baseline model are presented in Table 4. Despite the challenges associated with our smaller sample size, the comparison offers valuable insights into the relative efficacy of our method.

XSL•FO
RenderX

Table 4. Average performance across different baseline machine learning models.

Model architecture	Accuracy (%)	Sensitivity (%)	Precision (%)	F_1-score (%)
RF[a] [23,27]	12.50	5.00	5.00	N/A[b]
SVM[c] [23]	15.00	10.00	10.00	N/A
LR[d] [23]	22.50	15.00	13.34	N/A
XGBoost [34]	25.00	10.00	10.00	N/A
L1-based feature selection + DT[e] [24]	32.50	35.00	18.34	N/A
L1-based feature selection + RF [24]	22.50	15.00	13.34	N/A
L1-based feature selection + kNN[f] [24]	22.50	15.00	13.34	N/A
L1-based feature selection + NB[g] [24]	30.00	25.00	15.00	N/A
L1-based feature selection + LR [24]	37.50	45.00	25.00	N/A
L1-based feature selection + SVM [24]	25.00	10.00	10.00	N/A
Randomized LR + AdaBoost [25]	55.00	73.33	55.00	N/A
HOPE model[h]	87.50	90.00	88.34	86.00

[a]RF: random forest.

[b]N/A: data not applicable.

[c]SVM: support vector machine.

[d]LR: logistic regression.

[e]DT: decision tree.

[f]kNN: k-nearest neighbor.

[g]NB: naive Bayes.

[h]Best performed proposed model.

As shown in Table 4, traditional single-stage machine learning classifiers such as RF, LR, and SVM demonstrate relatively lower performance, with accuracy ranging from 12.50% to 22.50%. Among these, LR achieves relatively higher accuracy. XGBoost exhibits better performance than the traditional models. Incorporating feature selection techniques further improves the performance of these models. Specifically, combining L1-based feature selection with various classifiers results in modest performance gains, while the randomized LR combined with AdaBoost achieves a significant improvement, reaching 55.00% accuracy and 73.33% sensitivity. Our proposed 3-stage architectures significantly surpass all other baseline models across all metrics. While most baseline models struggle with sensitivity and precision, often scoring below 15%, our best proposed model demonstrates substantial enhancements in these metrics with 90.00% sensitivity and 88.34% precision, indicating a superior capability to correctly identify positive cases and reduce false positives.

Model Explainability

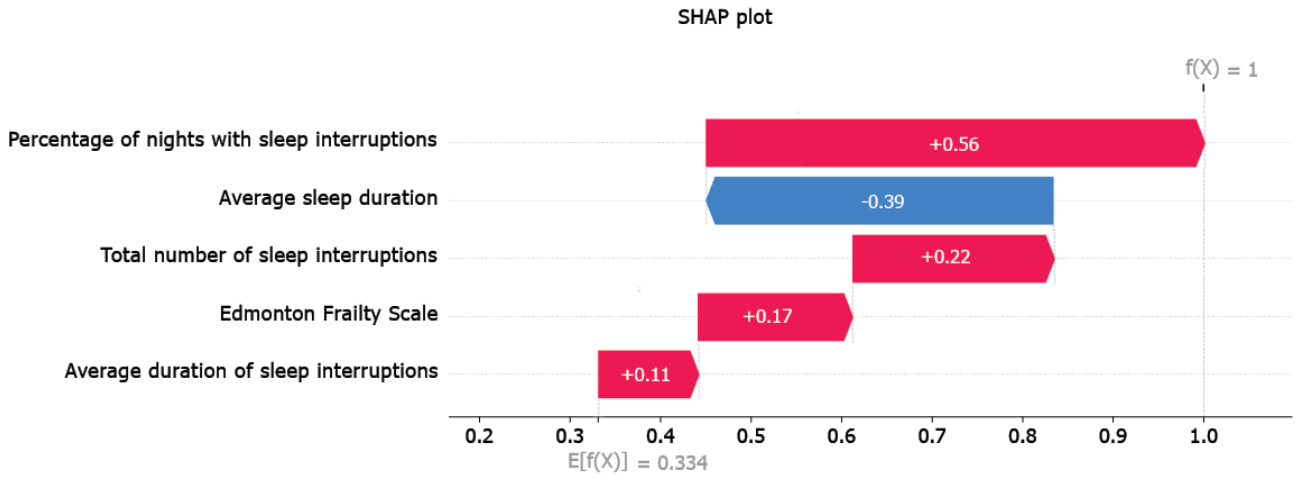
To enhance our understanding of the decision-making processes within our proposed model, we used SHAP [50] and LIME [54] for model interpretability analysis. These model-agnostic methods can be applied across various machine learning models, providing valuable insights into our model's predictive behavior. By integrating these interpretability techniques, we aim to improve the transparency and potential clinical relevance of our depression classification framework. These methods help us identify which features most significantly influence the model's predictions, particularly in the context of depression classification.

SHAP Analysis

The SHAP waterfall plot (Figure 5) illustrates the relative importance of features for depression classification, with red and blue colors representing positive and negative contributions, respectively. Among the features, "the percentage of nights with sleep interruptions" is the most impactful, positively correlating with depression risk, indicating that frequent sleep disturbances are a strong predictor of depression. Conversely, the average sleep duration exhibits a substantial negative impact on depression prediction, suggesting that longer sleep durations are associated with a reduced likelihood of depression. Sleep-related variables continue to play a pivotal role in the model's predictions; both the total number of sleep interruptions and the average duration of these interruptions contribute positively to depression risk, further underscoring the importance of uninterrupted sleep in depression diagnosis. Although the frailty scale is included in the model, its influence is relatively minor compared to sleep-related features.

XSL•FO
RenderX

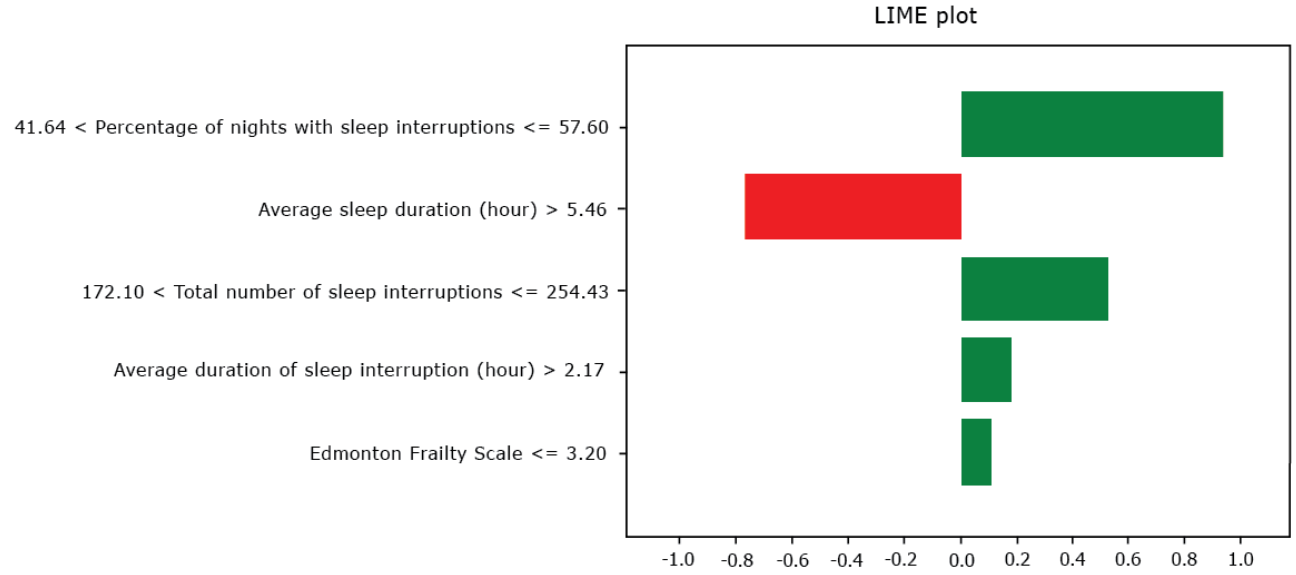
Figure 5. SHAP analysis. SHAP: Shapley additive explanations.



LIME Analysis

The LIME plot (Figure 6) provides a complementary view of feature importance, with green and red colors indicating positive and negative influences, respectively. Consistent with the SHAP results, LIME identifies "the percentage of nights with sleep interruptions" as the most critical feature in the classification of depression of our proposed model. Similarly, the average sleep duration is shown to have a significant negative impact on depression classification, in line with SHAP findings. The total number of sleep interruptions also ranks highly with a positive influence on depression risk, again aligning with SHAP results. A notable difference between the 2 methods is the relatively lower impact of the frailty scale in the LIME analysis, which requires further investigations.

Figure 6. LIME analysis. LIME: local interpretable model-agnostic explanations.



Discussion

Our research uses Wi-Fi–based motion sensors to extract daily activities, which are then used in our proposed machine learning method for depression classification.

Our study findings confirm the feasibility of using Wi-Fi–based motion sensors for depression classification among older adults. Our proposed HOPE (Home-Based Older Adults' Depression Prediction) model achieved an accuracy of 87.5%, sensitivity of 90%, and precision of 88.3%. The most influential features identified were sleep-related metrics, such as average sleep duration and sleep interruptions, highlighting the importance of sleep patterns in depression classification. These findings suggest that Wi-Fi–based monitoring offers a nonintrusive and effective alternative to conventional wearable technologies for depression assessment. These conventional methods, while effective, often present challenges in terms of participant compliance, particularly among older adults, due to their burdensome and sometimes uncomfortable nature. In contrast, our Wi-Fi–based approach is nonintrusive and allows for continuous monitoring without requiring participants to wear or interact with any devices. This can significantly enhance participant compliance and the integrity of the data collected over extended periods. Compared to other nonintrusive monitoring technologies, such as camera-based methods, our Wi-Fi–based approach has distinct advantages. Wi-Fi infrastructure is prevalent in most homes and does not pose

XSL•FO
RenderX

privacy risks, making it a cost-effective and scalable solution for continuous health monitoring. Furthermore, unlike previous studies that rely on microlevel body displacement and accelerometer data, our study emphasizes macrolevel physical activity features like sleep patterns and overall activity levels shift is crucial as it highlights the potential of using broader, more easily obtainable metrics to assess depression status. Our findings demonstrate that these macrolevel features are not only feasible but also effective measures for depression classification, broadening the scope of nonintrusive monitoring technologies in mental health research. The next steps can be extracting more detailed types of human activity using nonintrusive Wi-Fi data and expanding more on using this type of data acquisition for depression classification.

Additionally, our proposed model demonstrates relatively high performance compared to other classification models presented in existing depression classification studies, even with a limited sample size. Many studies using physical activity data from wearable devices often benefit from larger datasets and frequently use single-stage classifiers or deep neural networks. These models generally show strong performance with abundant data; however, their effectiveness diminishes when applied to smaller datasets, such as the one in this feasibility study. To address the limitations imposed by our smaller sample size, we designed a 3-stage machine learning classification architecture, which combines feature selection, dimensionality reduction, and classification into a multistep process. This approach allows for the extraction of the most relevant features while minimizing noise, thereby improving classification performance. Despite the small sample size, our model consistently outperformed conventional single-stage classifiers, highlighting the strength of both the machine learning architecture and the selected human activity features—particularly sleep patterns and activity levels—used for depression classification. This also underscores the adaptability of our model to different data scales, making it a more versatile option for future research where data availability might be limited. While this model shows promising results, however, caution is needed in interpreting these results. Future work should aim to enhance its robustness and generalizability by expanding the dataset. Collecting Wi-Fi–based physical activity data from a larger and more diverse sample would not only improve the model's statistical power but also allow for a more comprehensive evaluation of its performance across different population groups, such as varying age ranges and health conditions. This would be particularly valuable in developing a scalable solution for real-world applications. Additionally, the integration of advanced machine learning techniques, such as deep neural networks or hybrid models combining traditional classifiers with deep learning components, could further enhance classification accuracy.

Our study is distinctive not only in its methodological approach but also in its emphasis on model explainability, a crucial aspect often overlooked in prior research on depression classification. Explainability is essential in health care applications, where understanding the factors driving a model's decision is critical for clinical adoption and trust. By using SHAP and LIME, we were able to dissect the decision-making process of our model and pinpoint the most influential features for classifying depression. Both explainability analyses converge on the identification of sleep interruption features as key predictors in the depression classification of our proposed model. Among these, the "percentage of nights with sleep interruptions," "average sleep duration," and "total number of sleep interruptions" emerged as the primary driving factors. These findings align with existing literature that highlights the strong correlation between sleep disturbances and depression. However, our approach goes a step further by quantifying the impact of these features on the classification outcomes, providing a more nuanced understanding of their role. These findings suggest that future tools for depression assessment may benefit from a stronger focus on sleep quality and patterns and further investigations are required in this regard.

The integration of sensors and AI is transforming health care, yet the application of these technologies in depression classification remains underdeveloped and lacks extensive investigation. This study aimed to create an automated machine learning system for the continuous, remote monitoring and assessment of daily physical activity among older adults in a home setting, with the goal of distinguishing between individuals with and without depression.

In summary, although there were some challenges, our results suggest that using Wi-Fi–based data to capture contextual human activities is a promising and efficient method for classifying depression. The model we developed, leveraging data from Wi-Fi motion sensors, showed strong potential in accurately identifying early signs of depression and paving the way for more advanced and accessible mental health monitoring technologies among community dwelling older adults..

Acknowledgments

This study was funded by the Ministry of Innovation of Quebec, McGill University, Aerial Tech, and Mila—Quebec AI Institute. The authors acknowledge support from these institutions. SAR is Canada Research Chair (Tier II) in Artificial Intelligence and Advanced Digital Primary Health Care, received salary support from a Research Scholar Junior 1 Career Development Award from the Fonds de Recherche du Québec-Santé during part of this project, and her research program is supported by the Natural Sciences Research Council Discovery (grant 2020-05246).

Data Availability

The data analyzed during this study are not publicly available due to restrictions placed by the McGill University Institutional Ethics Committee, but the deidentified data are available from the corresponding author on reasonable request.

XSL•FO
RenderX

Authors' Contributions

Conceptualization was led by SAR, who established the study's goals, design, and research questions and obtained the funding for the project. The methodology was developed by SAR, and SN. The data collection was done by SAR and VK. Data curation was managed by SAR, SN, and VK. Formal analysis was conducted by SAR and SN. The original draft was written by SAR, and SN. Reviewing and editing were a collaborative effort with all authors. Supervision and overall project leadership were provided by SAR.

Conflicts of Interest

None declared.

References

1. Abd-Alrazaq A, AlSaad R, Shuweihdi F, Ahmed A, Aziz S, Sheikh J. Systematic review and meta-analysis of performance of wearable artificial intelligence in detecting and predicting depression. NPJ Digit Med 2023;6(1):84 [FREE Full text] [doi: 10.1038/s41746-023-00828-5] [Medline: 37147384]
2. Su YA, Si T. Progress and challenges in research of the mechanisms of anhedonia in major depressive disorder. Gen Psychiatr 2022;35(1):e100724 [FREE Full text] [doi: 10.1136/gpsych-2021-100724] [Medline: 35309242]
3. Parra F, Benezeth Y, Yang F. Automatic assessment of emotion dysregulation in American, French, and Tunisian adults and new developments in deep multimodal fusion: cross-sectional study. JMIR Ment Health 2022;9(1):e34333 [FREE Full text] [doi: 10.2196/34333] [Medline: 35072643]
4. Choi J, Lee S, Kim S, Kim D, Kim H. Depressed mood prediction of elderly people with a wearable band. Sensors 2022;22(11):4174 [FREE Full text] [doi: 10.3390/s22114174] [Medline: 35684797]
5. Gao Y, Jia Z, Zhao L, Han S. The effect of activity participation in middle-aged and older people on the trajectory of depression in later life: national cohort study. JMIR Public Health Surveill 2023;9:e44682 [FREE Full text] [doi: 10.2196/44682] [Medline: 36951932]
6. Song S, Seo Y, Hwang S, Kim HY, Kim J. Digital phenotyping of geriatric depression using a community-based digital mental health monitoring platform for socially vulnerable older adults and their community caregivers: 6-week living lab single-arm pilot study. JMIR Mhealth Uhealth 2024;12:e55842 [FREE Full text] [doi: 10.2196/55842] [Medline: 38885033]
7. Colizzi M, Lasalvia A, Ruggeri M. Prevention and early intervention in youth mental health: Is it time for a multidisciplinary and trans-diagnostic model for care? Int J Ment Health Syst 2020;14(1):23 [FREE Full text] [doi: 10.1186/s13033-020-00356-9] [Medline: 32226481]
8. Owen D, Antypas D, Hassoulas A, Pardiñas AF, Espinosa-Anke L, Collados JC. Enabling early health care intervention by detecting depression in users of web-based forums using language models: longitudinal analysis and evaluation. JMIR AI 2023;2:e41205 [FREE Full text] [doi: 10.2196/41205] [Medline: 37525646]
9. Huang S, Wang Y, Li G, Hall BJ, Nyman TJ. Digital mental health interventions for alleviating depression and anxiety during psychotherapy waiting lists: systematic review. JMIR Ment Health 2024;11:e56650 [FREE Full text] [doi: 10.2196/56650] [Medline: 39255015]
10. Smith KM, Renshaw PF, Bilello J. The diagnosis of depression: current and emerging methods. Compr Psychiatry 2013;54(1):1-6 [FREE Full text] [doi: 10.1016/j.comppsych.2012.06.006] [Medline: 22901834]
11. Minaeva O, Riese H, Lamers F, Antypa N, Wichers M, Booij SH. Screening for depression in daily life: development and external validation of a prediction model based on actigraphy and experience sampling method. J Med Internet Res 2020;22(12):e22634 [FREE Full text] [doi: 10.2196/22634] [Medline: 33258783]
12. Glavin D, Grua EM, Nakamura CA, Scazufca M, Ribeiro Dos Santos E, Wong GHY, et al. Patient health questionnaire-9 item pairing predictiveness for prescreening depressive symptomatology: machine learning analysis. JMIR Ment Health 2023;10:e48444 [FREE Full text] [doi: 10.2196/48444] [Medline: 37856186]
13. Camp N, Johnston J, Lewis MGC, Zecca M, Di Nuovo A, Hunter K, et al. Perceptions of in-home monitoring technology for activities of daily living: semistructured interview study with community-dwelling older adults. JMIR Aging 2022;5(2):e33714 [FREE Full text] [doi: 10.2196/33714] [Medline: 35511248]
14. Chan A, Cai J, Qian L, Coutts B, Phan S, Gregson G, et al. In-home positioning for remote home health monitoring in older adults: systematic review. JMIR Aging 2024;7:e57320 [FREE Full text] [doi: 10.2196/57320] [Medline: 39622026]
15. Gianfredi V, Blandi L, Cacitti S, Minelli M, Signorelli C, Amerio A, et al. Depression and objectively measured physical activity: a systematic review and meta-analysis. Int J Environ Res Public Health 2020;17(10):3738 [FREE Full text] [doi: 10.3390/ijerph17103738] [Medline: 32466242]
16. Arrieta H, Rezola-Pardo C, Echeverria I, Iturburu M, Gil SM, Yanguas JJ, et al. Physical activity and fitness are associated with verbal memory, quality of life and depression among nursing home residents: preliminary data of a randomized controlled trial. BMC Geriatr 2018;18(1):80. [doi: 10.1186/s12877-018-0770-y] [Medline: 29580209]
17. Dillon CB, McMahon E, O'Regan G, Perry IJ. Associations between physical behaviour patterns and levels of depressive symptoms, anxiety and well-being in middle-aged adults: a cross-sectional study using isotemporal substitution models. BMJ Open 2018;8(1):e018978 [FREE Full text] [doi: 10.1136/bmjopen-2017-018978] [Medline: 29358436]

XSL•FO
RenderX

18. Jung S, Lee S, Lee S, Bae S, Imaoka M, Harada K, et al. Relationship between physical activity levels and depressive symptoms in community-dwelling older Japanese adults. Geriatr Gerontol Int 2018;18(3):421-427. [doi: 10.1111/ggi.13195] [Medline: 29052928]
19. Ludwig VM, Bayley A, Cook DG, Stahl D, Treasure JL, Asthworth M, et al. Association between depressive symptoms and objectively measured daily step count in individuals at high risk of cardiovascular disease in South London, UK: a cross-sectional study. BMJ Open 2018;8(4):e020942 [FREE Full text] [doi: 10.1136/bmjopen-2017-020942] [Medline: 29654044]
20. Ku PW, Steptoe A, Liao Y, Sun WJ, Chen LJ. Prospective relationship between objectively measured light physical activity and depressive symptoms in later life. Int J Geriatr Psychiatry 2018;33(1):58-65. [doi: 10.1002/gps.4672] [Medline: 28181713]
21. Raudsepp L, Riso EM. Longitudinal association between objectively measured walking and depressive symptoms among Estonian older adults. J Aging Phys Act 2017;25(4):639-645. [doi: 10.1123/japa.2016-0303] [Medline: 28338399]
22. Kumar P, Garg S, Garg A. Assessment of anxiety, depression and stress using machine learning models. Procedia Comput Sci 2020;171:1989-1998. [doi: 10.1016/j.procs.2020.04.213]
23. Adamczyk J, Malawski F. Comparison of manual and automated feature engineering for daily activity classification in mental disorder diagnosis. Comput Inf 2021;40(4):850-879. [doi: 10.31577/cai_2021_4_850]
24. Bai R, Xiao L, Guo Y, Zhu X, Li N, Wang Y, et al. Tracking and monitoring mood stability of patients with major depressive disorder by machine learning models using passive digital data: prospective naturalistic multicenter study. JMIR Mhealth Uhealth 2021;9(3):e24365 [FREE Full text] [doi: 10.2196/24365] [Medline: 33683207]
25. Chikersal P, Doryab A, Tumminia M, Villalba DK, Dutcher JM, Liu X, et al. Detecting depression and predicting its onset using longitudinal symptoms captured by passive sensing. ACM Trans Comput Hum Interact 2021;28(1):1-41. [doi: 10.1145/3422821]
26. Dai R, Kannampallil T, Zhang J, Lv N, Ma J, Lu C. Multi-task learning for randomized controlled trials. Proc ACM Interact Mob Wearable Ubiquitous Technol 2022;6(2):1-23. [doi: 10.1145/3534591]
27. Griffiths C, da Silva KM, Leathlean C, Jiang H, Ang CS, Searle R. Investigation of physical activity, sleep, and mental health recovery in treatment resistant depression (TRD) patients receiving repetitive transcranial magnetic stimulation (rTMS) treatment. J Affect Disord Rep 2022;8:100337 [FREE Full text] [doi: 10.1016/j.jadr.2022.100337] [Medline: 35619990]
28. Espino-Salinas CH, Galván-Tejada CE, Luna-García H, Gamboa-Rosales H, Celaya-Padilla JM, Zanella-Calzada LA, et al. Two-dimensional convolutional neural network for depression episodes detection in real time using motor activity time series of depresjon dataset. Bioengineering 2022;9(9):458 [FREE Full text] [doi: 10.3390/bioengineering9090458] [Medline: 36135004]
29. Jakobsen P, Garcia-Ceja E, Riegler M, Stabell LA, Nordgreen T, Torresen J, et al. Applying machine learning in motor activity time series of depressed bipolar and unipolar patients compared to healthy controls. PLoS One 2020;15(8):e0231995 [FREE Full text] [doi: 10.1371/journal.pone.0231995] [Medline: 32833958]
30. Jung D, Kim J, Mun KR. Identifying depression in the elderly using gait accelerometry. Annu Int Conf IEEE Eng Med Biol Soc 2022;2022:4946-4949. [doi: 10.1109/EMBC48229.2022.9871877] [Medline: 36086152]
31. Kulam S. Time-series classification with uni-dimensional convolutional neural networks: An experimental comparison with long short-term memory networks [dissertation].: University of Oslo; 2019. URL: https://www.duo.uio.no/handle/10852/73250 [accessed 2024-12-07]
32. Kumar A, Sangwan SR, Arora A, Menon VG. Depress-DCNF: A deep convolutional neuro-fuzzy model for detection of depression episodes using IoMT. Appl Soft Comput 2022;122:108863. [doi: 10.1016/j.asoc.2022.108863]
33. Nguyen DK, Chan CL, Adams Li AH, Phan DV. Deep stacked generalization ensemble learning models in early diagnosis of depression illness from wearable devices data. 2021 Presented at: ICMHI 2021: 2021 5th International Conference on Medical and Health Informatics; May 14-16, 2021; Kyoto, Japan. [doi: 10.1145/3472813.3472815]
34. Opoku Asare K, Moshe I, Terhorst Y, Vega J, Hosio S, Baumeister H, et al. Mood ratings and digital biomarkers from smartphone and wearable data differentiates and predicts depression status: a longitudinal data analysis. Pervasive Mobile Comput 2022;83:101621. [doi: 10.1016/j.pmcj.2022.101621]
35. Park C, Mishra R, Golledge J, Najafi B. Digital biomarkers of physical frailty and frailty phenotypes using sensor-based physical activity and machine learning. Sensors 2021;21(16):5289 [FREE Full text] [doi: 10.3390/s21165289] [Medline: 34450734]
36. Akbari G, Nikkhoo M, Wang L, Chen CPC, Han DS, Lin YH, et al. Frailty level classification of the community elderly using Microsoft Kinect–based skeleton pose: a machine learning approach. Sensors (Basel) 2021;21(12):4017 [FREE Full text] [doi: 10.3390/s21124017] [Medline: 34200838]
37. Utsumi T, Hashimoto M. A development of an early detection system of pre-frailty in senior citizens living inside. 2023 Presented at: 17th International Conference on Ubiquitous Information Management and Communication (IMCOM); Seoul, Korea; January 03-05, 2023. [doi: 10.1109/imcom56909.2023.10035612]

XSL•FO
RenderX

38. Piau A, Wild K, Mattek N, Kaye J. Current state of digital biomarker technologies for real-life, home-based monitoring of cognitive function for mild cognitive impairment to mild Alzheimer disease and implications for clinical care: systematic review. J Med Internet Res 2019;21(8):e12785 [FREE Full text] [doi: 10.2196/12785] [Medline: 31471958]
39. Tegou T, Kalamaras I, Tsipouras M, Giannakeas N, Votis K, Tzovaras D. A low-cost indoor activity monitoring system for detecting frailty in older adults. Sensors 2019;19(3):452 [FREE Full text] [doi: 10.3390/s19030452] [Medline: 30678280]
40. Goonawardene N, Tan HP, Tan LB. Unobtrusive detection of frailty in older adults. In: Zhou J, Salvendy G, editors. Human Aspects of IT for the Aged Population Applications in Health, Assistance, and Entertainment. Cham, Switzerland: Springer International Publishing; 2018.
41. Mohtadifar M, Cheffena M, Pourafzal A. Acoustic- and radio-frequency–based human activity recognition. Sensors 2022;22(9):3125 [FREE Full text] [doi: 10.3390/s22093125] [Medline: 35590815]
42. Wang S, Zhou G. A review on radio based activity recognition. Digit Commun Netw 2015;1(1):20-29. [doi: 10.1016/j.dcan.2015.02.006]
43. Showmik IA, Sanam TF, Imtiaz H. Human activity recognition from Wi-Fi CSI data using principal component-based wavelet CNN. Digit Signal Process 2023;138:104056. [doi: 10.1016/j.dsp.2023.104056]
44. Pulle V, Rozon M, Zandifar A, Yadavalli SH, Asghari VR, Minayi Jalil A, et al. Systems and methods for remote monitoring care using wireless sensing technologies. Justia. URL: https://patents.justia.com/patent/20240027564 [accessed 2024-10-29]
45. Allegue M, Ghourchian N, Rozon M, Carignan E, Hopper M, Gonzalez JF, et al. WiFi motion intelligence: The fundamentals. Aerial Technologies Inc. URL: https://aerial.ai/files/common/The_fundamentals_Aerial_Telefonica_PCP-final_aligned.pdf [accessed 2025-01-14]
46. Vitkovic J, Le C, Lee SL, Clark RA. The contribution of hearing and hearing loss to balance control. Audiol Neurootol 2016;21(4):195-202. [doi: 10.1159/000445100] [Medline: 27251708]
47. Dent E, Kowal P, Hoogendijk EO. Frailty measurement in research and clinical practice: a review. Eur J Intern Med 2016;31:3-10. [doi: 10.1016/j.ejim.2016.03.007] [Medline: 27039014]
48. Rolfson DB, Majumdar SR, Tsuyuki RT, Tahir A, Rockwood K. Validity and reliability of the Edmonton Frail Scale. Age Ageing 2006;35(5):526-529 [FREE Full text] [doi: 10.1093/ageing/afl041] [Medline: 16757522]
49. Perna S, Francis MD, Bologna C, Moncaglieri F, Riva A, Morazzoni P, et al. Performance of Edmonton Frail Scale on frailty assessment: its association with multi-dimensional geriatric conditions assessed with specific screening tools. BMC Geriatr 2017;17(1):2 [FREE Full text] [doi: 10.1186/s12877-016-0382-3] [Medline: 28049443]
50. Yesavage JA. The use of self-rating depression scales in the elderly. In: Poon LW, Crook T, Davis KL, Eisdorfer C, Gurland BJ, Kaszniak AW, et al, editors. Handbook for Clinical Memory Assessment of Older Adults. Washington, DC: American Psychological Association; 1986.
51. Norgeot B, Quer G, Beaulieu-Jones BK, Torkamani A, Dias R, Gianfrancesco M, et al. Minimum information about clinical artificial intelligence modeling: the MI-CLAIM checklist. Nat Med 2020;26(9):1320-1324 [FREE Full text] [doi: 10.1038/s41591-020-1041-y] [Medline: 32908275]
52. Heo YJ. Population aging and house prices: Who are we calling old? J Econ Ageing 2022;23:100417 [FREE Full text] [doi: 10.1016/j.jeoa.2022.100417]
53. Sarawagi S, Bhamidipaty A. Interactive deduplication using active learning. 2002 Presented at: KDD02: The Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; July 23-26, 2002; Edmonton, Canada. [doi: 10.1145/775047.775087]
54. Ghadiri P, Yaffe MJ, Adams AM, Abbasgholizadeh-Rahimi S. Primary care physicians' perceptions of artificial intelligence systems in the care of adolescents' mental health. BMC Prim Care 2024;25(1):215 [FREE Full text] [doi: 10.1186/s12875-024-02417-1] [Medline: 38872128]
55. Rahimi SA, Légaré F, Sharma G, Archambault P, Zomahoun HTV, Chandavong S, et al. Application of artificial intelligence in community-based primary health care: systematic scoping review and critical appraisal. J Med Internet Res 2021;23(9):e29839 [FREE Full text] [doi: 10.2196/29839] [Medline: 34477556]
56. Hall MA. Correlation-Based Feature Selection for Machine Learning [dissertation]. Hamilton, New Zealand: University of Waikato; 1999. URL: https://www.lri.fr/~pierres/donn%E9es/save/these/articles/lpr-queue/hall99correlationbased.pdf [accessed 2025-01-14]
57. Abdi H, Williams LJ. Principal component analysis. WIREs Computational Stats 2010;2(4):433-459. [doi: 10.1002/wics.101]
58. Rish I. An empirical study of the naive Bayes classifier. 2001 Presented at: IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence; August 4, 2001; Seattle, WA URL: https://www.semanticscholar.org/paper/An-empirical-study-of-the-naive-Bayes-classifier-Watson/2825733f97124013e8841b3f8a0f5bd4ee4af88a
59. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. 2017 Presented at: NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems; December 4-9, 2017; Long Beach, CA URL: https://papers.nips.cc/paper_files/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html
60. Marcano-Cedeño A, Quintanilla-Domínguez J, Cortina-Januchs MG, Andina D. Feature selection using sequential forward selection and classification applying artificial metaplasticity neural network. 2010 Presented at: IECON 2010—36th Annual Conference on IEEE Industrial Electronics Society; November 7-10, 2010; Glendale, AZ. [doi: 10.1109/iecon.2010.5675075]

XSL•FO
RenderX

61. Fodor IK. A survey of dimension reduction techniques. Office of Scientific and Technical Information. 2002. URL: https://www.osti.gov/biblio/15002155 [accessed 2022-05-09]
62. Jr DWH, Lemeshow S. Applied Logistic Regression. Hoboken, NJ: John Wiley & Sons; 2004.
63. Peltola T. Local interpretable model-agnostic explanations of Bayesian predictive models via Kullback-Leibler projections. ArXiv. Preprint posted online on October 5, 2018 2018. [doi: 10.48550/arXiv.1810.02678]
64. Vergara JR, Estévez PA. A review of feature selection methods based on mutual information. Neural Comput Appl 2013;24(1):175-186. [doi: 10.1007/s00521-013-1368-0]
65. Balakrishnama S, Ganapathiraju A. Linear Discriminant Analysis—A Brief Tutorial [dissertation].: Mississippi State University; 1995. URL: https://www.semanticscholar.org/paper/LINEAR-DISCRIMINANT-ANALYSIS-A-BRIEF-TUTORIAL-Balakrishnama-Ganapathiraju/163bb761cd7c5ca8c786bf4816a10308f1396ffb [accessed 2024-10-29]
66. Peterson LE. K-nearest neighbor. Scholarpedia 2009;10(2):1883. [doi: 10.4249/scholarpedia.1883]
67. Tariq MA. A study on comparative analysis of feature selection algorithms for students grades prediction. J Inf Organ Sci 2024;48(1):133-147. [doi: 10.31341/jios.48.1.7]
68. Cao LJ, Chua KS, Chong WK, Lee HP, Gu QM. A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine. Neurocomputing 2003;55(1-2):321-336. [doi: 10.1016/s0925-2312(03)00433-8]
69. Hearst MA, Dumais ST, Osuna E, Platt J, Scholkopf B. Support vector machines. IEEE Intell Syst Appl 1998;13(4):18-28. [doi: 10.1016/j.aca.2011.07.027] [Medline: 21889629]
70. Granitto PM, Furlanello C, Biasioli F, Gasperi F. Recursive feature elimination with random forest for PTR-MS analysis of agroindustrial products. Chemometrics Intell Lab Syst 2006;83(2):83-90. [doi: 10.1016/j.chemolab.2006.01.007]
71. Song Y, Lu Y. Decision tree methods: applications for classification and prediction. Shanghai Arch Psychiatry 2015;27(2):130-135 [FREE Full text] [doi: 10.11919/j.issn.1002-0829.215044] [Medline: 26120265]
72. Breiman L. Random forests. Machine Learning 2001;45(1):5-32. [doi: 10.1023/A:1010933404324]
73. Friedman JH. Greedy function approximation: A gradient boosting machine. Ann. Statist 2001;29(5):1189-1232. [doi: 10.1214/aos/1013203451]
74. Chen T, He T, Benesty M, Khotilovich V, Tang Y, Cho H, et al. xgboost: Extreme Gradient Boosting. R Foundation for Statistical Computing. 2014. URL: https://cran.r-project.org/web/packages/xgboost/index.html [accessed 2024-07-24]
75. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. LightGBM: A highly efficient gradient boosting decision tree. 2017 Presented at: NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems; December 4-9, 2017; Long Beach, CA.
76. Ruta D, Gabrys B. Classifier selection for majority voting. Inf Fusion 2005;6(1):63-81. [doi: 10.1016/j.inffus.2004.04.008]
77. Zareapoor M, Shamsolmoali P. Application of credit card fraud detection: based on bagging ensemble classifier. Procedia Comput Sci 2015;48:679-685. [doi: 10.1016/j.procs.2015.04.201]
78. Hastie T, Rosset S, Zhu J, Zou H. Multi-class AdaBoost. Stat Interface 2009;2(3):349-360. [doi: 10.4310/sii.2009.v2.n3.a8]
79. Khalid S, Khalil T, Nasreen S. A survey of feature selection and feature extraction techniques in machine learning. Sci Inf Conf 2014:024-032 [FREE Full text] [doi: 10.14308/ite000494]
80. Tsai CF, Sung YT. Ensemble feature selection in high dimension, low sample size datasets: parallel and serial combination approaches. Knowl Based Syst 2020;203:106097. [doi: 10.1016/j.knosys.2020.106097]
81. Soares I, Dias J, Rocha H, Do Carmo Lopes M, Ferreira B. Feature selection in small databases: a medical-case study. 2016 Presented at: XIV Mediterranean Conference on Medical and Biological Engineering and Computing 2016; March 31-April 2, 2016; Paphos, Cyprus. [doi: 10.1007/978-3-319-32703-7_158]
82. Lever J, Krzywinski M, Altman N. Principal component analysis. Nature Methods 2017;14(7):641-642. [doi: 10.1038/nmeth.4346]
83. Saed AAA, Jaharadak AA. Implementation with performance evaluation of decision tree classifier for uncertain data. Lit Rev 2022;5(5):125-132. [doi: 10.1017/cbo9780511810114.020]
84. Bergstra J, Bengio Y. Random search for hyper-parameter optimization. J Mach Learn Res 2012;13(1):281-305 [FREE Full text]
85. Frazier P. A tutorial on Bayesian optimization. ArXiv. Preprint posted online July 8, 2018 2018 [FREE Full text] [doi: 10.1287/educ.2018.0188]
86. Li L, Jamieson KG, DeSalvo G, Rostamizadeh A, Talwalkar A. Hyperband: A novel bandit-based approach to hyperparameter optimization. J Mach Learn Res 2016:6765-6816 [FREE Full text]
87. Shreffler J, Huecker MR. Diagnostic Testing Accuracy: Sensitivity, Specificity, Predictive Values and Likelihood Ratios. Treasure Island, FL: StatPearls Publishing; 2024.
88. Liu W, Jiang Y, Zhang J, Ma Z. RFARN: Retinal vessel segmentation based on reverse fusion attention residual network. PLoS One 2021;16(12):e0257256 [FREE Full text] [doi: 10.1371/journal.pone.0257256] [Medline: 34860847]
89. Goodyear MDE, Krleza-Jeric K, Lemmens T. The Declaration of Helsinki. BMJ 2007;335(7621):624-625 [FREE Full text] [doi: 10.1136/bmj.39339.610000.BE] [Medline: 17901471]
90. Tri-Council policy statement: ethical conduct for research involving humans—TCPS 2 (2022). Government of Canada. URL: https://ethics.gc.ca/eng/policy-politique_tcps2-eptc2_2022.html [accessed 2024-10-29]

XSL•FO
RenderX

91. Shayan Nejadshamsi. AID project - Depression Github Repository. URL: https://github.com/rahimi-s-lab/HOPE-model [accessed 2025-02-07]
92. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in python. J Mach Learn Res 2015;12(85):2825-2830. [doi: 10.1145/2786984.2786995]
93. Akyol Y, Durmus D, Dogan C, Bek Y, Canturk F. Quality of life and level of depressive symptoms in the geriatric population. Turk J Rheumatol 2010;25(4):165-173. [doi: 10.5152/tjr.2010.23]
94. de Jonge P, Ormel J, Slaets JPJ, Kempen GIJM, Ranchor AV, van Jaarsveld CHM, et al. Depressive symptoms in elderly patients predict poor adjustment after somatic events. Am J Geriatr Psychiatry 2004;12(1):57-64. [Medline: 14729560]
95. Helvik AS, Engedal K, Krokstad S, Stordal E, Selbæk G. A comparison of depressive symptoms in elderly medical inpatients and the elderly in a population-based health study (the Nord-Trøndelag Health Study 3). Nord J Psychiatry 2012;66(3):189-197. [doi: 10.3109/08039488.2011.614958] [Medline: 21958367]
96. Ordu Gokkaya NK, Gokce-Kutsal Y, Borman P, Ceceli E, Dogan A, Eyigor S, et al. Pain and quality of life (QoL) in elderly: the Turkish experience. Arch Gerontol Geriatr 2012;55(2):357-362. [doi: 10.1016/j.archger.2011.10.019] [Medline: 22104759]
97. Hsu HC. Physical function trajectories, depressive symptoms, and life satisfaction among the elderly in Taiwan. Aging Ment Health 2009;13(2):202-212. [doi: 10.1080/13607860802342201] [Medline: 19347687]
98. Lue BH, Chen LJ, Wu SC. Health, financial stresses, and life satisfaction affecting late-life depression among older adults: a nationwide, longitudinal survey in taiwan. Arch Gerontol Geriatr 2010;50:S34-S38. [doi: 10.1016/S0167-4943(10)70010-8] [Medline: 20171454]
99. Feng L, Yap KB, Ng TP. Depressive symptoms in older adults with chronic kidney disease: mortality, quality of life outcomes, and correlates. Am J Geriatr Psychiatry 2013;21(6):570-579. [doi: 10.1016/j.jagp.2012.12.020] [Medline: 23567405]
100. Van der Weele GM, Gussekloo J, De Waal MWM, De Craen AJM, Van der Mast RC. Co-occurrence of depression and anxiety in elderly subjects aged 90 years and its relationship with functional status, quality of life and mortality. Int J Geriatr Psychiatry 2009;24(6):595-601. [doi: 10.1002/gps.2162] [Medline: 19031476]
101. Ben-Zeev D, Scherer EA, Wang R, Xie H, Campbell AT. Next-generation psychiatric assessment: using smartphone sensors to monitor behavior and mental health. Psychiatr Rehabil J 2015;38(3):218-226 [FREE Full text] [doi: 10.1037/prj0000130] [Medline: 25844912]
102. Byrne JEM, Bullock B, Brydon A, Murray G. A psychometric investigation of the sleep, circadian rhythms, and mood (SCRAM) questionnaire. Chronobiol Int 2019;36(2):265-275. [doi: 10.1080/07420528.2018.1533850] [Medline: 30395721]
103. Difrancesco S, Lamers F, Riese H, Merikangas KR, Beekman ATF, van Hemert AM, et al. Sleep, circadian rhythm, and physical activity patterns in depressive and anxiety disorders: a 2-week ambulatory assessment study. Depress Anxiety 2019;36(10):975-986 [FREE Full text] [doi: 10.1002/da.22949] [Medline: 31348850]
104. Ghandeharioun A, Fedor S, Sangermano L, Ionescu D, Alpert J, Dale C, et al. Objective assessment of depressive symptoms with machine learning and wearable sensors data. 2017 Presented at: 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII); October 23-26, 2017; San Antonio, TX. [doi: 10.1109/acii.2017.8273620]
105. Jacobson NC, Weingarden H, Wilhelm S. Using digital phenotyping to accurately detect depression severity. J Nerv Ment Dis 2019;207(10):893-896. [doi: 10.1097/NMD.0000000000001042] [Medline: 31596769]
106. Vallance JK, Eurich D, Lavallee C, Johnson ST. Daily pedometer steps among older men: associations with health-related quality of life and psychosocial health. Am J Health Promot 2013;27(5):294-298. [doi: 10.4278/ajhp.120316-QUAN-145] [Medline: 23402230]
107. Aprahamian I, Borges MK, Hanssen DJC, Jeuring HW, Oude Voshaar RC. The frail depressed patient: a narrative review on treatment challenges. Clin Interv Aging 2022;17:979-990 [FREE Full text] [doi: 10.2147/CIA.S328432] [Medline: 35770239]
108. Lohman MC, Mezuk B. Frailty and depression: comorbidity in the context of imperfect measurement. J Am Geriatr Soc 2013;61(3):474 [FREE Full text] [doi: 10.1111/jgs.12138] [Medline: 23496193]
109. Vaughan L, Corbin AL, Goveas JS. Depression and frailty in later life: a systematic review. Clin Interv Aging 2015;10:1947-1958 [FREE Full text] [doi: 10.2147/CIA.S69632] [Medline: 26719681]

Abbreviations

AI: artificial intelligence
CNN: convolutional neural network
CSI: channel state information
DT: decision tree
EFS: Edmonton Frailty Scale
GDS: Geriatric Depression Scale
LIME: local interpretable model-agnostic explanations
LR: logistic regression

XSL•FO
RenderX

MI-CLAIM: Minimum Information About Clinical Artificial Intelligence Modeling
PCA: principal component analysis
RF: random forest
RSSI: received signal strength indicator
SFS: sequential forward selection
SHAP: Shapely addictive explanations
SVM: support vector machine

Edited by Y Du; submitted 18.10.24; peer-reviewed by PH Liao, C Silva; comments to author 05.12.24; revised version received 12.12.24; accepted 19.12.24; published 03.03.25.

Please cite as:
Nejadshamsi S, Karami V, Ghourchian N, Armanfard N, Bergman H, Grad R, Wilchesky M, Khanassov V, Vedel I, Abbasgholizadeh Rahimi S
Development and Feasibility Study of HOPE Model for Prediction of Depression Among Older Adults Using Wi-Fi-based Motion Sensor Data: Machine Learning Study
JMIR Aging 2025;8:e67715
URL: https://aging.jmir.org/2025/1/e67715
doi:10.2196/67715
PMID:40053734

©Shayan Nejadshamsi, Vania Karami, Negar Ghourchian, Narges Armanfard, Howard Bergman, Roland Grad, Machelle Wilchesky, Vladimir Khanassov, Isabelle Vedel, Samira Abbasgholizadeh Rahimi. Originally published in JMIR Aging (https://aging.jmir.org), 03.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Model-Based Feature Extraction and Classification for Parkinson Disease Screening Using Gait Analysis: Development and Validation Study

Ming De Lim[1], BCS; Tee Connie[1], PhD; Michael Kah Ong Goh[1], PhD; Nor 'Izzati Saedon[2], MBBS

[1]Faculty of Information Science and Technology, Multimedia University, Melaka, Malaysia

[2]Department of Medicine, Faculty of Medicine, Universiti Malaya, Kuala Lumpur, Malaysia

Corresponding Author:
Tee Connie, PhD
Faculty of Information Science and Technology
Multimedia University
Jalan Ayer Keroh Lama
Melaka, 75450
Malaysia
Phone: 60 62523592
Fax: 60 62318840
Email: tee.connie@mmu.edu.my

Abstract

Background: Parkinson disease (PD) is a progressive neurodegenerative disorder that affects motor coordination, leading to gait abnormalities. Early detection of PD is crucial for effective management and treatment. Traditional diagnostic methods often require invasive procedures or are performed when the disease has significantly progressed. Therefore, there is a need for noninvasive techniques that can identify early motor symptoms, particularly those related to gait.

Objective: The study aimed to develop a noninvasive approach for the early detection of PD by analyzing model-based gait features. The primary focus is on identifying subtle gait abnormalities associated with PD using kinematic characteristics.

Methods: Data were collected through controlled video recordings of participants performing the timed up and go (TUG) assessment, with particular emphasis on the turning phase. The kinematic features analyzed include shoulder distance, step length, stride length, knee and hip angles, leg and arm symmetry, and trunk angles. These features were processed using advanced filtering techniques and analyzed through machine learning methods to distinguish between normal and PD-affected gait patterns.

Results: The analysis of kinematic features during the turning phase of the TUG assessment revealed that individuals with PD exhibited subtle gait abnormalities, such as freezing of gait, reduced step length, and asymmetrical movements. The model-based features proved effective in differentiating between normal and PD-affected gait, demonstrating the potential of this approach in early detection.

Conclusions: This study presents a promising noninvasive method for the early detection of PD by analyzing specific gait features during the turning phase of the TUG assessment. The findings suggest that this approach could serve as a sensitive and accurate tool for diagnosing and monitoring PD, potentially leading to earlier intervention and improved patient outcomes.

(JMIR Aging 2025;8:e65629) doi:10.2196/65629

KEYWORDS

model-based features; gait analysis; Parkinson disease; computer vision; support vector machine

Introduction

Background

Parkinson disease (PD) is a common neurological disease that affects millions of individuals worldwide. This disorder gradually impairs a person's ability to move their body, resulting in a variety of crippling symptoms, including tremors, stiff muscles, and sluggish motions. To effectively manage PD and provide timely measures, accurate and early diagnosis is essential.

Traditionally, clinical evaluations conducted by medical professionals, which can be arbitrary and inconsistent, have

XSL•FO
RenderX

been used to diagnose PD. In distant or underdeveloped areas, access to specialized PD care may be restricted, delaying diagnosis and therapy. The prevalence of PD is predicted to rise as the world's population ages, placing additional demand on health care resources and highlighting the need for easily available and effective diagnostic methods.

In this regard, current technology has shown promise in addressing the difficulties associated with PD diagnosis, especially in the areas of machine learning and deep learning. The limits of conventional clinical procedures can be overcome with the promise of early and objective identification provided by machine learning techniques.

This study investigated the use of machine learning for PD identification with a focus on gait characteristics. The goal was to develop a simple and noninvasive method for screening PD symptoms early through analyzing gait patterns.

In this study, kinematic features during the turning phase of the timed up and go (TUG) assessment were extracted and analyzed. Gait variabilities occurring during body turning are more easily identified, as turning involves complex motor coordination and balance adjustments, making it a challenging movement. This increased complexity can accentuate subtle abnormalities in gait patterns that might not be as apparent during straight walking. Turning is particularly difficult for individuals with PD, as it requires precise control and stability, often revealing difficulties such as freezing of gait (FoG), reduced step length, and asymmetrical movements. Therefore, analyzing the turning phase allows for a more sensitive and accurate detection of PD symptoms, providing a reliable indicator of whether a participant exhibits PD-related gait abnormalities. This focus enhances the effectiveness of the diagnostic tool, which contributes to the facilitation of early and accurate identification of PD.

Conventional Methods

Challa et al [1] proposed an advanced predictive model for PD by using machine learning algorithms, such as multilayer perceptron, boosted logistic regression, random forest, and BayesNet. Their investigation applied the Parkinson Progression Markers Initiative dataset, an extensive dataset containing data from patients with PD and healthy participants. The experimental results showcased remarkable performance improvements over existing methods, boasting accuracy rates of 96.09% for training and 95.45% for testing in the multilayer perceptron algorithm, 96.5854% for training and 96.02% for testing in the BayesNet algorithm, 95.45% for training and 94.87% for testing in the random forest algorithm, and the highest accuracy achieved by the boosted logistic regression algorithm with 97.159% for training and 96.97% for testing. The area under the curve (AUC) of the receiver operating characteristic curve reached an impressive 98.9% for the boosted logistic regression algorithm, emphasizing its robust predictive capabilities. This research represents a significant stride in health care, providing a reliable model for early PD prediction, crucial for timely diagnosis and intervention in addressing this global public health challenge.

In 2019, Polat [2] investigated the recognition of FoG cases in individuals with PD, using a logistic regression classifier trained and tested on a dataset comprising 16 samples. The study meticulously assessed the classifier's performance using a comprehensive set of 10 performance measures, such as accuracy, miss rate, false discovery rate, false positive rate, false omission rate, sensitivity, specificity, precision, and negative predictive value. Impressively, the logistic regression classifier exhibited a noteworthy accuracy of 81.3% in accurately classifying FoG cases. The research further compared the performance of linear regression with 4 alternative models: linear support vector machine (SVM), quadratic SVM, cubic SVM, and k-nearest neighbors (KNN), revealing that the proposed logistic regression model surpassed its counterparts with the highest accuracy of 81.3% in classifying FoG datasets for individuals with PD. This outcome underscores the superior performance of the logistic regression model in the context of FoG classification.

Vidya and Sasikumar [3] conducted a comprehensive study on the application of multiclass SVM in using gait analysis to identify and grade the severity of PD. The researchers used a publicly accessible dataset containing Vertical Ground Reaction Force Sensors and implemented kinematic analysis to extract spatiotemporal features crucial for the diagnostic process. Their suggested framework included a multiregression strategy to normalize gait time series data and a correlation-based feature selection method. A total of 4 distinct SVM kernel functions such as linear, Gaussian, quadratic, and cubic were rigorously evaluated across 3 different walking tests to gauge their performance. Impressively, the quadratic SVM classifier emerged as the most effective, achieving an outstanding average accuracy of 98.65%. This result surpassed existing state-of-the-art methods, showcasing the robustness and efficacy of the proposed SVM-based approach for PD diagnosis and severity rating.

Moreover, Fang [4] performed a study focusing on predicting PD through the application of machine learning techniques. The research extensively compared the accuracy and recall of 3 distinct algorithms: KNN, random forest, and naive Bayesian. Addressing the inherent limitations of equal weighting in traditional KNN, the study introduced an entropy weight method to enhance KNN's performance, specifically mitigating equal-weighting issues. In addition, the research delved into a voice-based Unified Parkinson Disease Rating Scale (UPDRS) prediction scheme, leveraging algorithms to effectively predict UPDRS scores from voice data of patients with PD. Notably, the study used the University of California Irvine dataset, showcasing that the refined KNN algorithm surpassed its traditional counterpart, achieving a notable accuracy rate increase from 91.8% to 93.8%. This research contributes significantly to the advancement of PD prediction methodologies, particularly emphasizing the pivotal role of improved weighting mechanisms in enhancing algorithmic accuracy.

In addition, Gundala et al [5] conducted a comprehensive study using the random forest algorithm for the recognition of PD using the spiral handwritten dataset. The machine learning technique, widely recognized for its efficacy in processing handwritten designs, was applied by partitioning the dataset into subgroups on the basis of features and constructing decision

XSL•FO
RenderX

trees for each feature. The dataset, sourced from the Kaggle website, provided the foundation for training the random forest model. Notably, the algorithm's strength lies in its ability to combine outputs from multiple decision trees through majority voting, resulting in a remarkable accuracy rate of 91%. The study emphasizes the robustness of the random forest algorithm in accurately identifying PD on the basis of the features extracted from handwritten drawings. The results underscore the algorithm's effectiveness in leveraging the diversity of decision trees for enhanced predictive accuracy in the context of PD diagnosis.

Moreover, Govindu and Palwe [6] analyzed Multidimensional Voice Program audio data collected from both patients with PD and healthy individuals. Among the machine learning models evaluated, the random forest classifier emerged as the most effective, achieving a detection accuracy of 91.83% and a sensitivity of 0.95. This model outperformed other techniques, such as SVM, KNN and logistic regression, which were also assessed for their classification capabilities. The superior performance of the random forest classifier highlights its robustness and reliability for detecting PD in its early stages, demonstrating the promise of machine learning in enhancing diagnostic accuracy.

Deep Learning Methods

Pereira et al [7] proposed an innovative method for early identification of PD using a convolutional neural network (CNN) trained on handwritten dynamics data obtained by a smart pen. The CNN effectively learned relevant features from the signals generated during individual exams, enabling discrimination between individuals with and without PD on the basis of these learned features. Experimental results showcased the superiority of the CNN over raw data, with the ImageNet architecture, using 128 times 128 images and a 75% training dataset split, yielding the best overall accuracy of 83.77%. Despite these promising results, the study acknowledged challenges in achieving consistent recognition rates over control individuals, emphasizing the need for further refinement in the proposed approach.

Later, Grover et al [8] introduced a deep learning methodology using a deep neural network (DNN) constructed with TensorFlow and Keras, containing 3 hidden layers with 10, 20, and 10 neurons each and an input layer with 16 units, concluding with an output layer representing the classes *severe* and *nonsevere*. The DNN was trained on the Parkinson telemonitoring voice dataset from the University of California Irvine machine learning repository, comprising biomedical voice measurements from 42 participants. The experiments focused on predicting the severity of PD based on total UPDRS and motor UPDRS scores. The DNN exhibited substantial accuracy improvements over previous research, achieving a classification accuracy of 94.4422% for total UPDRS and 83.367% for motor UPDRS on the training dataset. While test dataset results were comparatively lower, with 62.7335% accuracy for total UPDRS and 81.6667% for motor UPDRS, the proposed DNN classifier showcased enhanced performance compared to previous studies, emphasizing its potential for accurate severity prediction in PD.

Aal et al [9] conducted an optimized approach for early PD detection, using speech features extracted from 2 datasets, dataset 1 and dataset 2, containing recordings from both healthy individuals and patients with PD. Using Mel-frequency cepstral coefficients and delta Mel-frequency cepstral coefficients, the authors proposed a deep learning model combining a recurrent neural network (RNN) with a long short-term memory (LSTM) layer. Comparative evaluation of alternative ML techniques, including SVM, KNN and RNN with stochastic gradient descent, revealed superior performance of the proposed RNN-LSTM model, optimized with the adaptive moment estimation optimizer. The model exhibited remarkable testing accuracy rates of 95.8% on dataset 1 and 90.24% on DS2, accompanied by high recall, precision, and F_1-score on both datasets. Aal et al [9] further demonstrated the model's superiority over existing methods for PD detection, solidifying its potential as an effective tool for early diagnosis.

Apart from that, Ouhmida et al [10] delved into the early recognition of voice-based PD through the application of CNN and artificial neural network (ANN). Using 2 distinct datasets, the study showcased the superior accuracy of CNN over ANN in their experiments. Dataset 1 encompassed 195 voice recordings from 31 individuals, while dataset 2 featured 240 recordings from 80 participants, with a balanced distribution of 40 patients with PD and 40 healthy individuals. The deep learning models, specifically CNN and ANN, underwent training and testing on both datasets, revealing CNN's accuracy rates of 93.10% and 88.89% for datasets 1 and 2, respectively. In contrast, the ANN model achieved slightly lower accuracies, with 82.76% for dataset 1 and 72.22% for dataset 2. Ouhmida et al [10] also highlighted the intricate layers comprising each model, with the ANN model featuring 2 hidden layers and the CNN model incorporating convolution, normalization, activation, softmax, and classification layers. The paper expressed an intent to extend the research by exploring additional deep learning methods and implementing a hybrid system integrating diverse techniques and datasets.

Biswas et al [11] proposed an approach for the early recognition of PD by proposing 2 distinct deep learning models tailored for hand-drawn graphics. The first model, a 2D CNN, processed preprocessed images of spirals, circles, and meanders as input, achieving notable accuracies of 83.6% on circles, 61.5% on spirals, and 67.8% on meanders. Another model, an innovative LSTM model, operated on timeline-series signals and demonstrated an overall accuracy of 0.78. Leveraging the NewHandPD dataset for training and testing, the authors conducted 4 comprehensive experiments, presenting the results. The study posited that early PD detection through these advanced models could potentially enhance treatment outcomes and elevate the overall quality of life for affected individuals.

Apart from that, Khaskhoussy and Ayed [12] evaluate the effectiveness of SVM CNN for classifying data obtained from speech tasks. Total 2 types of input data were analyzed: the raw speech signal values and i-vector features of dimensions 100, 200, and 300. The classification performance was assessed using 5 evaluation metrics: accuracy, precision, recall or sensitivity, specificity, and F_1-score. For a test dataset of 28 participants,

XSL•FO
RenderX

the approach achieved outstanding results, including 100% accuracy, a precision of 0.99, a recall of 0.98, a specificity of 0.96, and an F_1-score of 0.98. A table summarizing the state-of-the-art methods is presented in Table 1.

Table 1. A summary of the state-of-the-art methods.

Study	Methods	Input type	Dataset	Performance
Challa et al [1]	MLP[a], BayesNet, RF[b], and boosted logistic regression	Nonmotor symptoms	PPMI[c] dataset	Best overall accuracy: 96.97%
Polat [2]	Logistic regression	FoG[d]	FoG dataset (Parkinson disease)	81.3%
Vidya and Sasikumar [3]	SVM[e]	Gait features	Gait analysis dataset	98.65%
Fang [4]	KNN[f]	Voice records	UCI[g] dataset	93.8%
Gundala et al [5]	RF	Handwritten drawings	Kaggle handwritten drawings dataset	91%
Pereira et al [7]	CNN[h]	Handwritten dynamics	Public dataset of handwritten dynamics extracted by a smart pen	Best overall accuracy: 83.77%
Grover et al [8]	DNN[i]	Voice dataset	Parkinson telemonitoring voice dataset	Total UPDRS[j]: train 94.44%, test 62.73% and motor UPDRS: train 83.37%, test 81.67%
Aal et al [9]	RNN[k]-LSTM[l]	Speech features	Dataset 1 and dataset 1[2]	Dataset 1: 95.8% and dataset 12: 90.24%
Ouhmida et al [10]	CNN ANN[m]	Voice	Dataset 1, dataset 2	Dataset 1: 93.10% and dataset 2: 88.89%
Biswas et al [11]	LSTM	Hand drawings	NewHandPD	78.7%

[a]MLP: multilayer perceptron.
[b]RF: random forest.
[c]PPMI: Parkinson Progression Markers Initiative.
[d]FoG: freezing of gait.
[e]SVM: support vector machine.
[f]KNN: k-nearest neighbors.
[g]UCI: University of California Irvine
[h]CNN: convolutional neural network.
[i]DNN: deep neural network.
[j]UPDRS: Unified Parkinson Disease Rating Scale.
[k]RNN-: recurrent neural network.
[l]LSTM: long short-term memory.
[m]ANN: artificial neural network.

Methods

Overview

The study used a methodical approach that started with data collection from young adults, older adults, and patients with PD, followed by obtaining consent, particularly from those diagnosed with the disease. The TUG assessment was a key part of the data collection process. Video enhancements and preprocessing were performed to enhance the quality of the videos. After that, the key points on the human body were obtained using a human pose estimation technique. Next, features such as shoulder distance, step and stride lengths, cadence, and speed were extracted to analyze the gait patterns. The Butterworth filter was applied to refine the data, and peaks were identified to calculate steps and turning durations. Finally, SVM [2] was used to distinguish between the different groups based on the extracted features, aiming to improve the prediction and analysis of PD symptoms.

Dataset

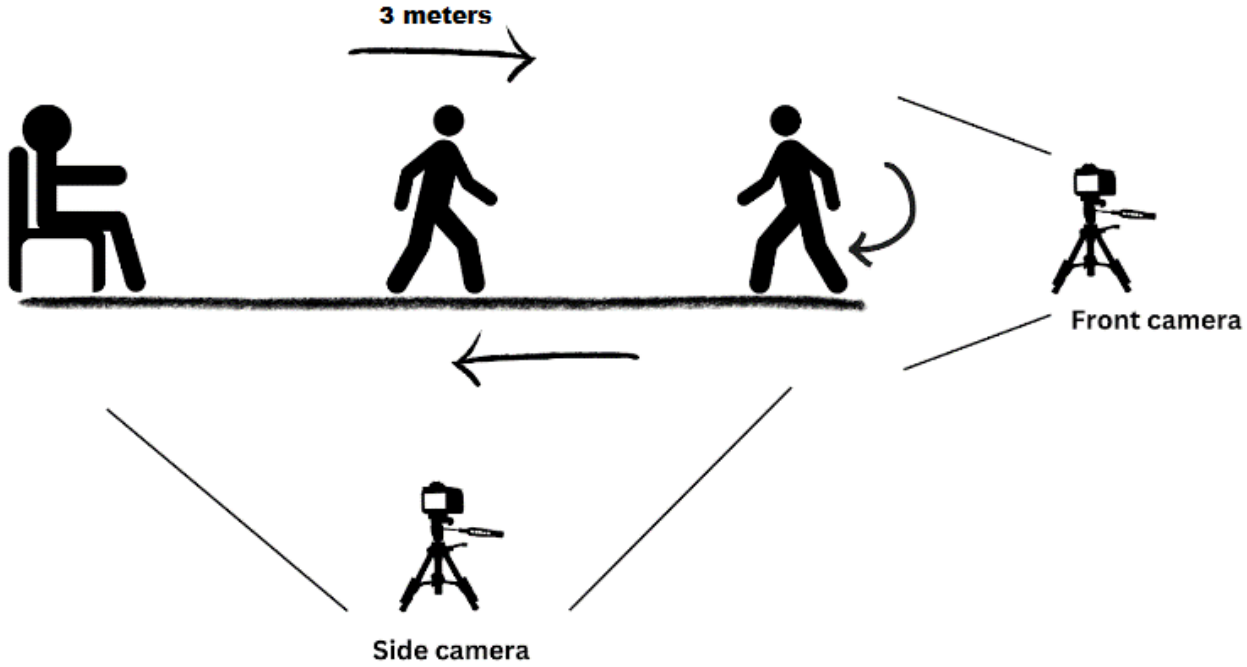
The self-collected dataset comprised video recordings of 28 individuals performing the TUG assessment, a standard test used to evaluate mobility and balance. The dataset comprised 3 distinct cohorts: young adults aged between 21 and 33 years, older individuals aged >60 years, and older individuals diagnosed with PD. Consents were obtained from the participants before data collection. The videos captured various gait patterns, providing a comprehensive set of data for analyzing kinematic features, such as step length, stride length, and joint angles. This dataset allowed for controlled conditions and detailed annotations, ensuring high-quality data for feature extraction and analysis.

XSL•FO
RenderX

TUG Assessment

Figure 1 shows a TUG workflow, outlining the sequential steps involved in the activity. The process began with the individual seated at the designated starting point. Upon initiating the recording, the participant stood up and proceeded to walk a distance of 3 meters in a typical manner. Upon reaching the 3-meter mark, the individual executed a turn, walking back to the starting point, where they concluded the TUG process by returning to a seated position. This comprehensive description encapsulated the entire TUG procedure, providing a clear understanding of the task's progression. To enhance data capture, 2 phones were used, one on the front and the other on the side.

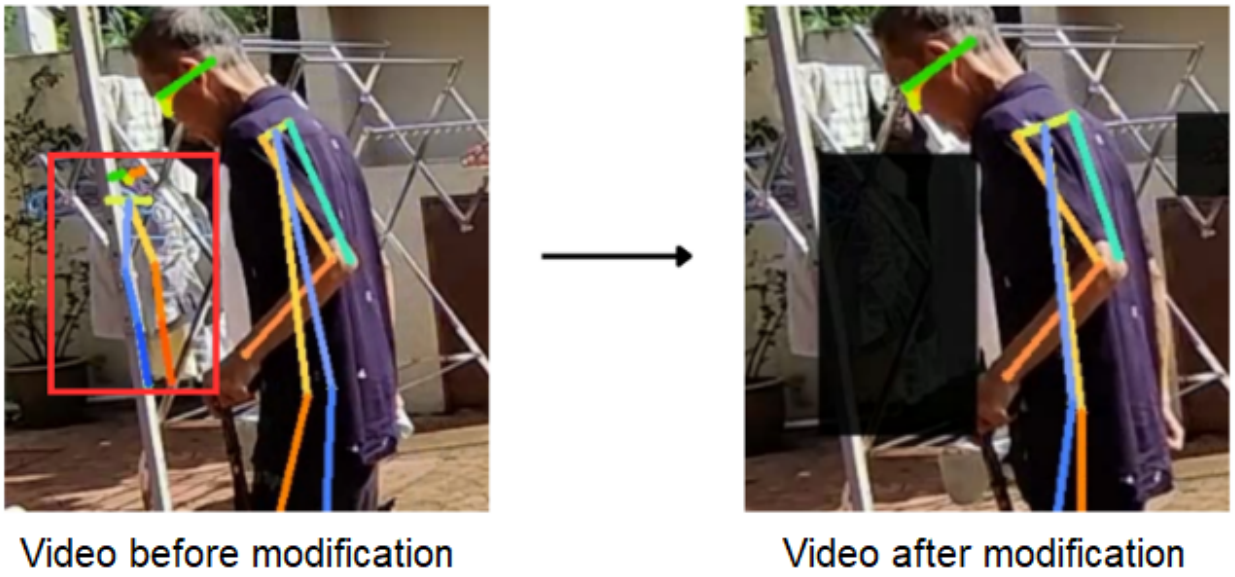
Figure 1. Timed up and go (TUG) assessment shooting site configuration plan.



Video Enhancement

The presence of background clutter in the original video introduces noise, which significantly affects the accuracy of letter pose extraction. Applying human pose estimation technique directly to unprocessed video may result in inaccurate keypoint coordinates, affecting the reliability of analysis results. To alleviate this problem, a preprocessing step was crucial. Specifically, the video was subjected to a customized processing method in which a black block was used to mask certain areas susceptible to noise. This strategic approach ensured that no noise interferes with the subsequent process, resulting in more precise and reliable coordination extraction. Figure 2 shows an example of enhancement performed on the video images.

Figure 2. Video enhancement.

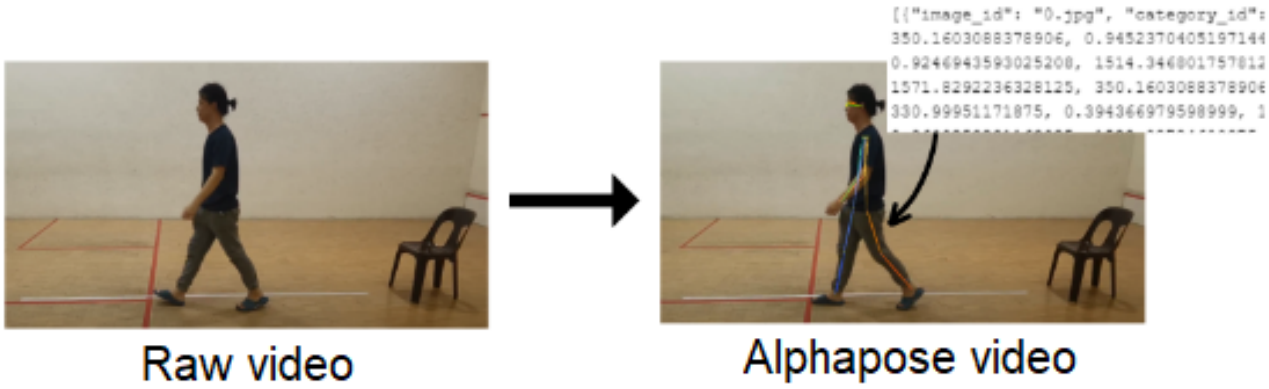


XSL•FO
RenderX

Human Pose Estimation

In this study, AlphaPose was used to identify 17 key points from the human body, each corresponding to a specific anatomical location. The process yielded a JSON file containing coefficients for 17 key points. It is important to note that the JSON file, at this stage, has not undergone data processing. The raw file was populated with various elements, including image_id, category_id, key points, and scores within each set. The image_id represented the frame of the video, while the category_id served to identify the object (set as 1 for a person). The key points section contains coordinates for body part locations and corresponding detection confidence, formatted as x1, y1, c1, x2, y2, c2, and so forth, with "c" denoting the confidence score (Figure 3).

Figure 3. Gait analysis using AlphaPose.



Data Extraction

In the data preprocessing phase, the first step involved loading the JSON file and carefully filtering key points and image IDs. In addition, confidence scores were eliminated. The next focus was to mark the 17 identified body key points, including nose, left eye (LEye), right eye (REye), left ear (LEar), right ear (REar), left shoulder (LShoulder), right shoulder (RShoulder), left elbow (LElbow), right elbow (RElbow), left wrist (LWrist), right wrist (RWrist), left hip (LHip), right hip (RHip), left knee (LKnee), right knee (Rknee), left ankle (LAnkle), and right ankle (RAnkle). Next, according to the specific video file name, key point coordinates for each image ID (frame) were marked, ranging from Nose_x and Nose_y to RAnkle_x and RAnkle_y.

The results from the data preprocessing stage boasted a comprehensive structure, encompassing 36 columns that capture key information. Each column was curated to provide a detailed representation of the dataset. Starting with the file_name, which specified the associated video file, and the image_id indicating the frame or image ID corresponding to each set of key points, the subsequent 34 columns focus on the x and y coordinates of 17 distinct body key points.

Frame Segmentation

Frame segmentation is a process to assign a time interval to each frame in the video which helps to determine the occurrence of a turning event. The total duration of the video is divided by the total number of frames (rows). This calculation yields the time interval per frame, assuming a constant frame rate throughout the video. With this interval calculated, a new column named "time_duration" is added to the DataFrame. Each row in the "time_duration" column is populated with the cumulative time, calculated as the frame's index in the DataFrame multiplied by the time interval. This method provided a time stamp for each frame, which is essential for synchronizing the data with the video and analyzing the timing of the detected body turning events. The pseudocode presented in Figure 4 describes the step-by-step procedure for frame segmentation.

XSL•FO
RenderX

Figure 4. Pseudocode describing the step-by-step procedure for frame segmentation.

Algorithm 1: Pseudocode for Frame Segmentation

Data: *df*
Result: *turning_events*

```
function frame_segmentation(DataFrame df)
  for each row in df:
    shoulder_distance = sqrt((RShoulder_x − LShoulder_x)^2 + (RShoulder_y − LShoulder_y)^2)
    set df[shoulder_distance] to shoulder_distance
  end for

  for each row in df:
    if df[shoulder_distance] > threshold then
      set df[turning_event] to TRUE
    else
    set df[turning_event] to FALSE
  end if
 end for

 Return turning_events
End function

# Main Process
DataFrame df = Load data into a DataFrame
Float distance_threshold = Set the distance threshold

DataFrame turning_events = detect_turning_events(df)

Integer total_rows = Length of DataFrame df  # Total number of rows
Float total_duration_seconds = Set the total duration of the video in seconds

Float time_interval = total_duration_seconds / total_rows

For each row in DataFrame turning_events:
  Calculate time_duration = index_of_row * time_interval
  Add time_duration to DataFrame turning_events
```

Noise Filtering

The Butterworth filter was used as the noise filtering tool in this study. The Butterworth filter is a type of signal processing filter that plays a pivotal role in enhancing the clarity of movement coordinate data extracted from videos. By applying the Butterworth filter to the raw coordinates, high-frequency noise is effectively attenuated, resulting in a smoother trajectory on the graph. This smoothed representation provides a clearer visualization of the participant's movements, reducing interference from irrelevant fluctuations.

Combined with the Butterworth filter, a peak detection algorithm was used to identify important points on the graph. These peaks correspond to key events in motion, such as steps taken by the participant. The peak identification process helped in extracting basic features such as the number of steps throughout the video. This kind of feature extraction is particularly valuable for specific PD aspects of analysis and quantification. The Butterworth filter is expressed using equation 1:



where $y(t)$ is the filtered signal; $x(t)$ is the input signal; fc is the cutoff frequency; f is the frequency of the signal; and n is the order of the filter.

In this study, an order of 4 was used for the Butterworth filters because it provides a smooth yet sufficiently steep roll-off in the transition band. While higher-order filters provide sharper roll-offs, they also require more computational resources. Therefore, a fourth-order filter appeared to be a reasonable compromise, being computationally efficient while providing adequate filtering. On the other hand, the peak detection equation is given by equation 2:

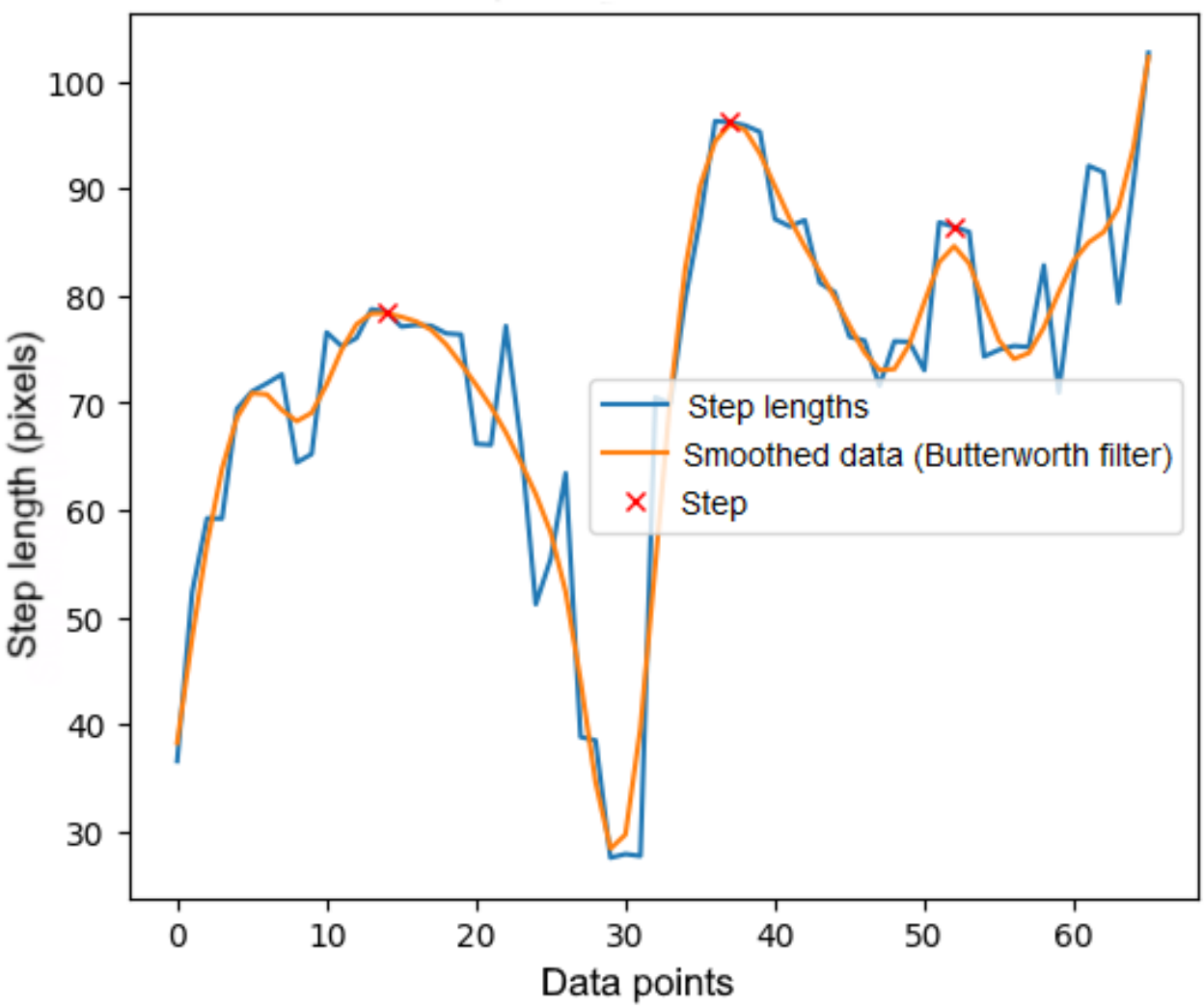
XSL•FO
RenderX



where $y(t)$ is the filtered signal; T is the set of all time points in the signal $y(t)$; t_i represents the time point where a peak occurs; and *height* is the threshold height for peak detection.

Only peaks that have an amplitude greater than or equal to the threshold, h, are detected and included in the output. In this study, the value for the height parameter was set to 100. An example of the graph for noise filtering with peaks is illustrated in Figure 5. This example demonstrates that the application of the Butterworth filter smooths the signal and yields an accurate step count during body turning.

Figure 5. An example of a graph processed with a Butterworth filter, showing peaks.



PD Recognition Using Model-Based Features

Certain gait features can help identify abnormalities or changes in walking patterns, which are particularly useful for diagnosing and monitoring conditions such as PD. By focusing on the body turning period, we aimed to capture the most challenging part of the gait cycle, where gait variabilities for PD are more easily observed. Turning involves complex motor coordination and balance adjustments, making it a difficult movement. This complexity can accentuate subtle abnormalities in gait patterns that might not be as apparent during straight walking. Therefore, analyzing gait features during turns provides a more sensitive and accurate assessment of PD-related gait abnormalities, enhancing the effectiveness of early diagnosis and monitoring. The gait features used in this study are summarized in Table 2.

XSL•FO
RenderX

Table 2. Gait analysis feature definitions and formulas.

Feature	Definition	Formula
Shoulder distance	Average horizontal distance between the left and right shoulders	mean(Shoulder_Distance($LShoulder_x$,$LShoulder_y$,$RShoulder_x$,$LShoulder_y$))
Step length	Average distance between the left ankle and right ankle	mean(Step_Length($LAnkle_x$,$LAnkle_y$,$RAnkle_x$,$RAnkle_y$))
Stride length	Average distance covered in 1 full stride, which consists of 2 steps (1 by each foot)	mean(Stride_Length($RAnkle_x$,$RAnkle_y$))
Angle of both knees	Degrees at the knee joint by considering the vectors formed by the knee to hip and knee to ankle points	Θ_{Knee}(Hip_x,Hip_y,$Knee_x$,$Knee_y$,$Ankle_x$,$Ankle_y$)
Angle of both hips	Degrees at the hip joint by considering the vectors formed by the hip to knee and hip to shoulder midpoint	Θ_{Hip} ($Shoulder_{mid}$,$Shoulder_{midy}$, Hip_x. Hip_y, $Knee_x$, $Knee_y$)
Symmetrical leg	Degrees for each ankle by considering the vectors formed by the knee to hip midpoint and knee midpoint to hip midpoint	Θ_{Leg}($Knee_{mid}$,$Knee_{midy}$, Hip_{mid} . Hip_{midy}, $Knee_x$, $Knee_y$)
Symmetrical arm	Degrees for each arm by considering the vectors formed by the arm to shoulder midpoint and arm midpoint to shoulder midpoint	Θ_{Arm}($Shoulder_{mid}$,$Shoulder_{midy}$, $Elbow_{mid}$. $Elbow_{midy}$)
Trunk Angle 1 (vertical)	Degrees by considering the vertical reference vector and the vector from the hip midpoint to the nose	Θ_{Trunk1}(Hip_{mid} ,Hip_{midy}, $Nose_x$. $Nose_y$)
Trunk angle 2 (horizontal)	Degrees by considering the vector from the hip midpoint to the shoulder midpoint and a horizontal reference vector	Θ_{Trunk2}($Shoulder_{mid}$,$Shoulder_{midy}$, Hip_{mid} ,Hip_{midy})
Shank angle	Degrees at the hip joint by considering the vectors formed by the left ankle to knee midpoint and right ankle to knee midpoint	Θ_{Shank}($Ankle_x$,$Ankle_y$,$Shoulder_{mid}$,$Shoulder_{midy}$)

Body Turning Duration Calculation

Figure 6 presents 2 graphs related to the analysis of shoulder distance during a body turn. The left graph (Figure 6A) displays the original, unprocessed data, illustrating the raw measurements of shoulder distance over time. The fluctuations observed may correspond to the natural movement variations during a turn. In contrast, the right graph (Figure 6B) exhibits data that have been smoothed using a Butterworth filter, a technique that mitigates short-term fluctuations and reveals the underlying pattern of movement more clearly.

The smoothed data enables effective determination of the peaks of the signal, which identifies the significant turning points. These peaks are highlighted by the red "X" marks in the figure and represent moments where the shoulder distance reaches its maximum, indicating a complete turn or a change in direction. By focusing on the 2 highest peaks, the graph underscores the most substantial turning events, thereby minimizing the potential for misinterpreting minor variations as significant movements.

In the context of shoulder distance, *LShoulder* and *RShoulder* would correspond to the left and right shoulders, with their respective *x* and *y* coordinates. The shoulder distance, *Shoulder_Distance*, which is a measure of how far apart the shoulders are, is calculated as follows (equation 3):

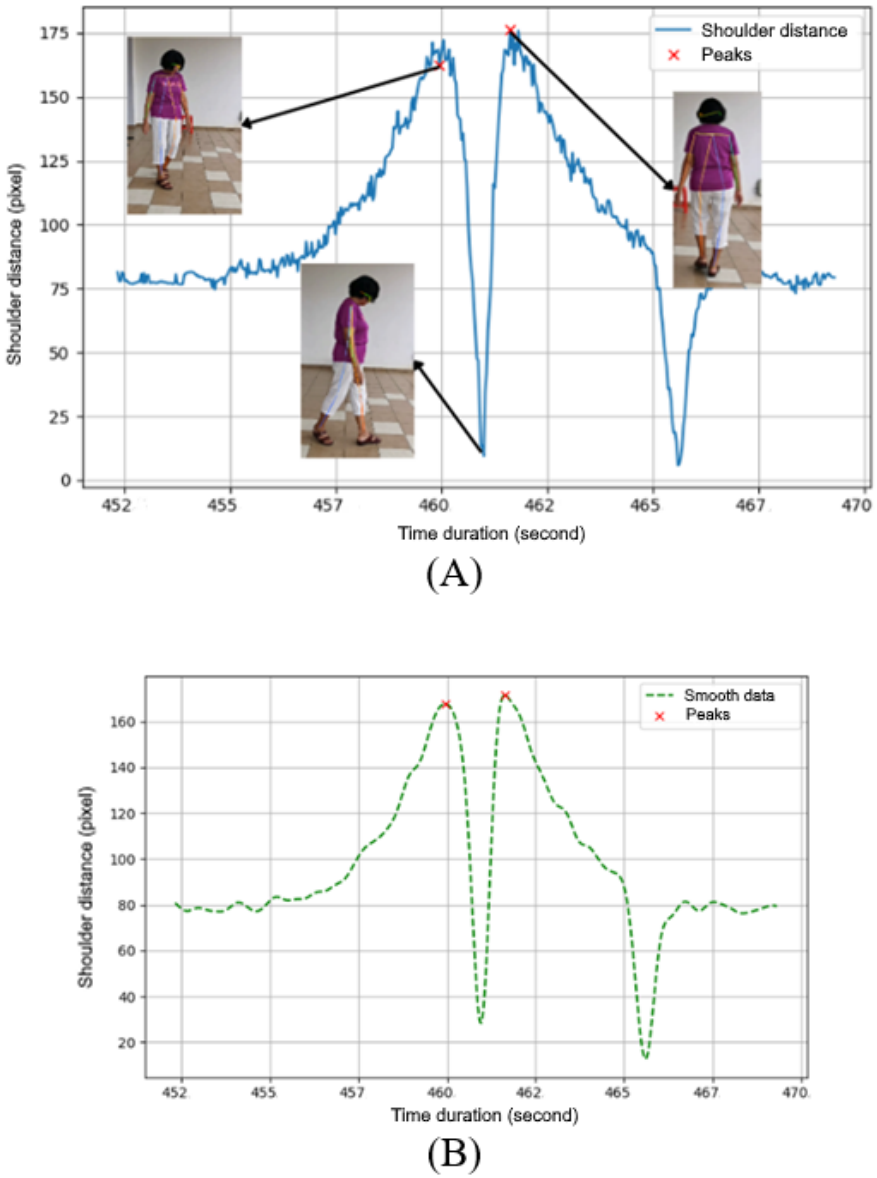


The indexes between these 2 peaks are then taken to determine the exact time the turn occurred. The difference between these peak times is calculated to find the duration between consecutive turning points. Let $T_{peak,i}$ be the time of the *i*th peak and *n* is the total number of peaks detected. Equation 4 calculates the total time duration between the first and the last peak by adding up the durations between each pair of consecutive peaks,



XSL•FO
RenderX

Figure 6. Shoulder distance analysis with peak detection: (A) graph before smoothing and (B) graph after smoothing.



Ethical Considerations

Ethical approval was granted by the Research Ethics Committee Multimedia University (approval number EA0422022). Informed consent was obtained from all the participants. A statement about this is given in the Methods section. Identifiable features of research participants are not visible in the manuscript. No compensation was provided to the participants.

Results

Overview

This section presents the experiment results and discussion for the gait features extracted from the self-collected dataset, focusing on the effectiveness of these features in distinguishing between normal and PD-affected gait patterns using a SVM classifier.

For the experiments, principal component analysis was applied to reduce the dimension of the dataset. The number of components for principal component analysis was set to 8, corresponding to the number of participants with PD. This step helps in capturing the most significant features while reducing computational complexity.

The dataset was additionally divided into a 70:30 ratio. This indicates that the model was trained using 70% of the data, with the remaining 30% set aside for testing and performance evaluation. By splitting the data this way, a significant portion of the data were used to train the model while maintaining enough for a thorough assessment. In total, 10 trials of the experiments were performed, and the average results were recorded.

Participant Details

There were 3 groups of participants in the self-collected dataset, namely, patients with PD, older adults, and adolescents. A significant portion of the participants (8/28, 29%) were diagnosed with PD, while the remaining participants were healthy (20/28, 71%). Most of the participants were Chinese (24/28, 85%), with a smaller representation of Indian (3/28, 11%) and Malay (1/28, 4%) individuals. There was a higher

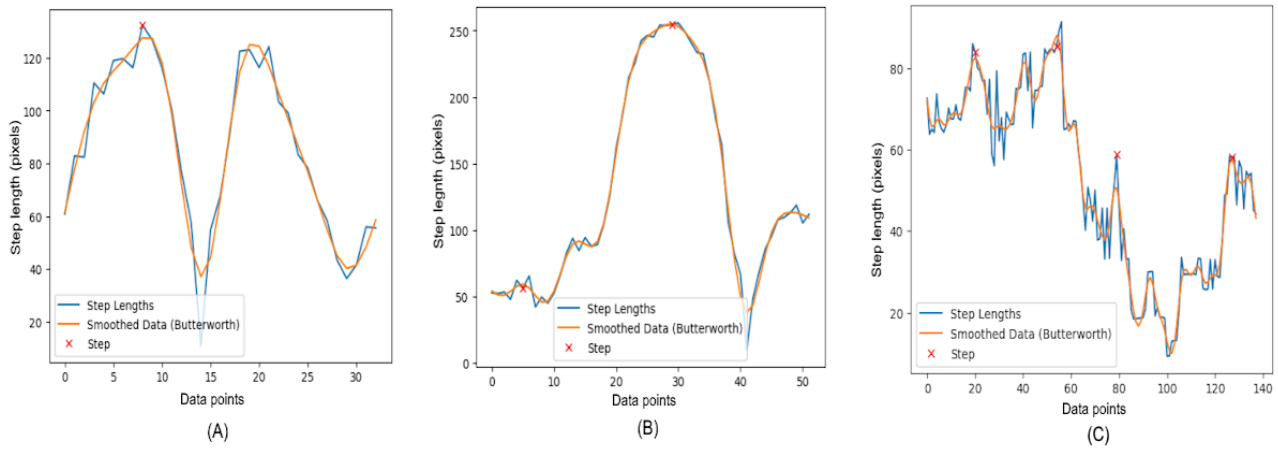
XSL•FO
RenderX

proportion of male participants (21/28, 75%) compared to female participants (7/28, 25%). Finally, most participants fell within the age range of 60 to 69 years (14/28, 50%), followed by the age range of 20 to 29 years (9/28, 32%), and smaller proportions in other age ranges. These results collectively illustrate the diversity and focus of the study population.

Total Steps During Body Turning

Figure 7 presents 3-line graphs that correspond to the step lengths for different demographic groups: younger individuals, older individuals, and individuals with PD. The step lengths are measured during a turning sequence.

Figure 7. A comparative step analysis by step length between different groups: (A) younger individuals, (B) older individuals, and (C) individuals with Parkinson disease.



The graph of younger individuals (Figure 7A) demonstrates a more variable stepping pattern with pronounced peaks, suggesting agility and a higher range of motion (ROM) during turns. The steps are uneven in length, but the peaks return to baseline quickly, indicating swift changes in direction and a potentially more dynamic gait.

In contrast, the graph of older individuals (Figure 7B) has fewer, more rounded peaks, indicative of a more cautious or steady turning strategy. The step lengths are generally more uniform, with smoother transitions between steps. This could reflect a more deliberate and potentially less stable gait, as is often seen with aging.

The graph of individuals with PD (Figure 7C) demonstrates a significantly different pattern, with smaller step lengths and a higher frequency of steps, which may point to the short, shuffling steps often associated with PD. The peaks are less pronounced and more erratic, highlighting the challenges individuals with PD may face in maintaining a regular stepping pattern.

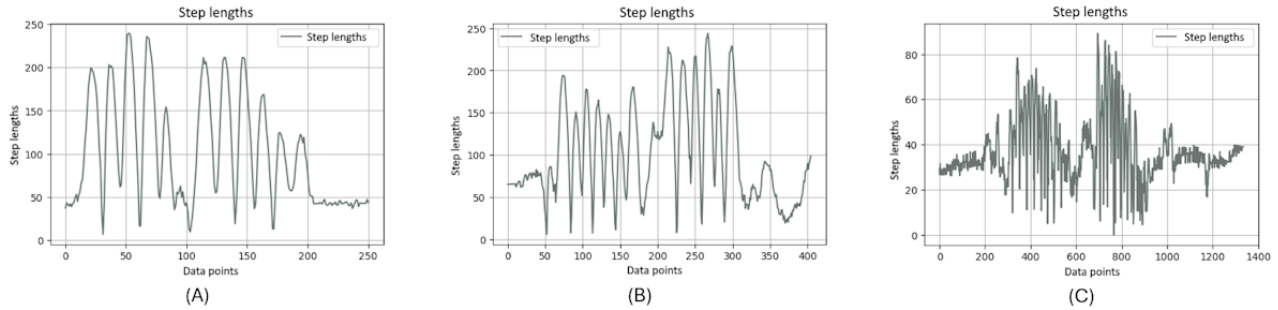
Step Length

Step length refers to the distance between the left and right ankle for each step. It is determined by the coordinates of the left ankle ($LAnkle_x$, $LAnkle_y$) and the right ankle ($RAnkle_x$, $RAnkle_y$) at a particular instance in time. The formula to calculate step length is provided in equation 5. This formula gives the straight-line distance between 2 points in a 2D space, such as a video frame or a motion capture system's coordinate plane.



In Figure 8, the graphs illustrate the differences in gait patterns across 3 demographic groups during a TUG test. Each graph plots the variation in step lengths, which reflect the distance between the left and right ankle during individual steps.

Figure 8. A comparative step length analysis between different groups: (A) younger individuals, (B) older individuals, and (C) individuals with Parkinson disease.



The graph of younger individuals (Figure 8A) is characterized by high peaks and deep valleys, indicating significant variability in step lengths. This could suggest a dynamic and robust gait, with the ability to take both long and short steps, perhaps adjusting speed or direction more frequently.

In contrast, the graph of older individuals (Figure 8B) shows a moderate level of variability with somewhat rounded peaks.

XSL•FO
RenderX

The reduced height of the peaks compared to the young group suggests shorter steps on average, which may be a sign of a more cautious approach to movement, possibly due to decreased mobility or balance concerns that come with age.

The graph of individuals with PD (Figure 8C) differs markedly from the other 2. It has a much higher frequency of smaller fluctuations, and the overall range of step lengths is noticeably lower. This pattern indicates the short, shuffling steps that are often observed in individuals with PD, reflecting the challenges they face with gait initiation and continuation.

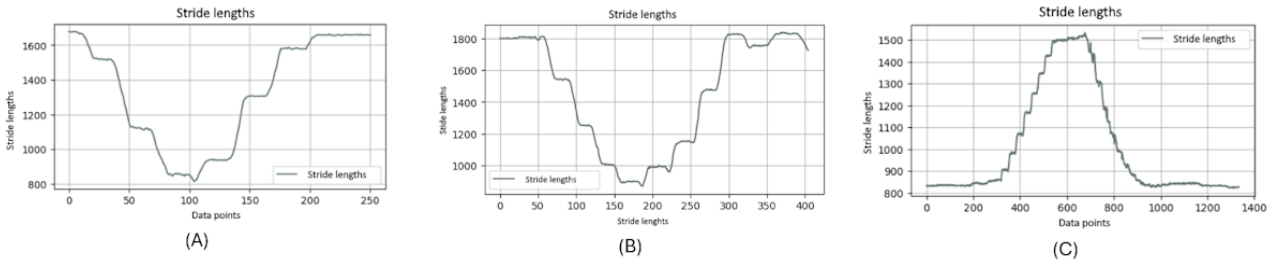
Stride Length

Stride length is typically defined as the distance covered in one full stride, consisting of 2 steps (one by each foot). The equation to calculate the stride length for each frame is presented in equation 6. It calculates the stride length as the Euclidean distance from a point (possibly the origin) to the coordinates of the right ankle for each frame.



Figure 9 shows the stride length of the right ankle throughout a single TUG video. The graph exhibits the dynamic nature of stride lengths, with the vertical axis representing the stride length and the horizontal axis corresponding to data points that likely represent sequential frames of the video.

Figure 9. A comparative stride length analysis between different groups: (A) younger individuals, (B) older individuals, and (C) individuals with Parkinson disease.



It is noteworthy that this graph clearly distinguishes between 2 types of motion: the rising and falling edges indicate the leg in motion, not in contact with the ground, and the flat regions signify the stance phase where the leg is stationary on the floor, awaiting the next step. The consistency in the flat regions indicates moments when the leg is at rest between strides.

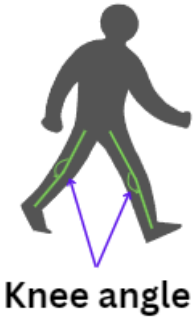
It is evident that the graphs of younger and older individuals (Figures 9A and 9B) display similar frequencies in stride patterns, whereas the graph of the individuals with PD (Figure 9C) significantly diverges. It illustrates a strikingly different frequency in stride lengths, standing out from the more consistent rhythmic patterns observed in the younger- and older-individual groups.

Moreover, from these charts, it is obvious that when younger and older people complete a stride, their feet leave the ground to a greater extent, because a normal stride will take a relatively large range. On the contrary, when patients with PD complete a stride, their feet leave the ground to a relatively smaller extent. The amplitude of the ground is relatively short, which can be considered a symptom of PD.

Angle of Both Knees

Analyzing the positions of the hip, knee, and ankle for each leg to determine the angles at both knees can assist in recognizing PD (Figure 10).

Figure 10. An illustration of the knee angle position.



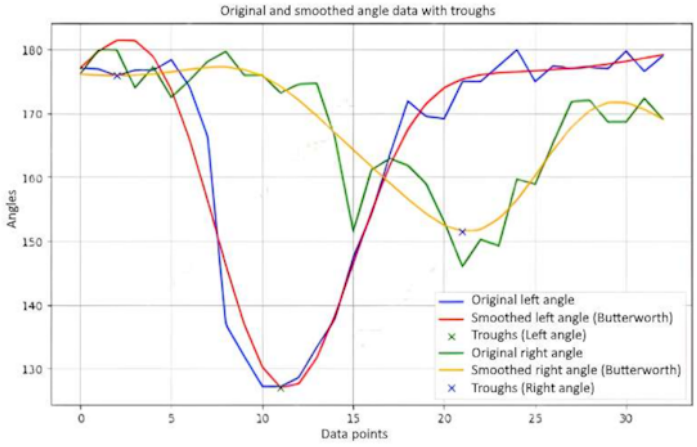
The knee angles provide crucial insights into the degree of bending or extension at the knee joints, which are essential for understanding gait patterns. PD often affects movement and posture, leading to distinctive gait abnormalities. By quantifying these knee angles, the calculation helps in detecting and assessing these gait characteristics, aiding in the identification and monitoring of PD symptoms.

Given the knee, hip, and ankle joint positions, the knee angle and the Θ_{knee} can be computed as follows (equation 7):

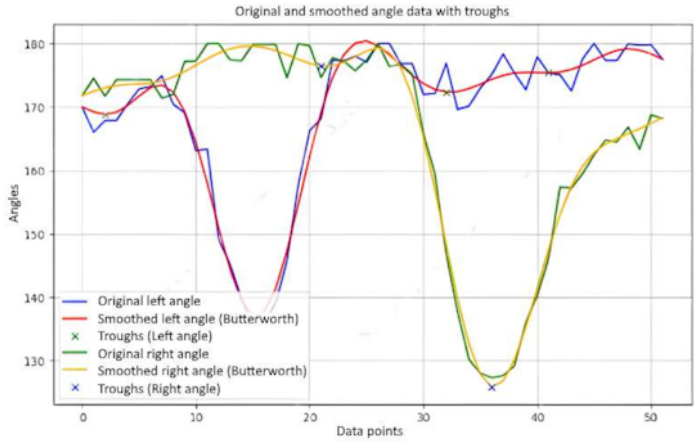
XSL•FO
RenderX

Figure 11 presents 3-line graphs comparing the knee joint angles of the 3 groups: younger individuals, older individuals, and individuals with PD. The graph of the younger individuals group (Figure 11A) illustrates smooth and regular oscillations in knee joint angles for both the left and right legs. As they turn their body, the angles show significant flexion and extension, reflecting the youthful ability to execute the turn in a single, fluid step. The peaks and troughs are well-defined, indicating robust and agile movements typical of healthy, young individuals. This pattern highlights their efficient and coordinated gait.

Figure 11. A comparative analysis of the knee angle between different groups: (A) younger individuals, (B) older individuals, and (C) individuals with Parkinson disease.



(A)



(B)



(C)

XSL•FO
RenderX

In the older-individual group graph (Figure 11B), the knee joint angles are more subdued, with 2 distinct curves for each turn. This indicates that older individuals take 2 steps to complete the body turn, reflecting a more cautious and segmented approach. The movement patterns are slower and less dynamic compared to the younger individual group, with a reduced ROM.

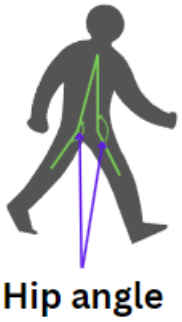
The graph for the group of individuals with PD (Figure 11C) shows highly irregular and erratic knee joint angles. The curves exhibit frequent sharp peaks and drops, indicating inconsistent and disrupted movement as individuals with PD attempt to turn their bodies. These fluctuations reflect the challenges faced by patients with PD, including tremors, rigidity, and difficulty maintaining smooth and coordinated movements. The graph captures their struggle to control knee flexion and extension, leading to a more disordered and interrupted gait pattern during the turning motion.

These graphs highlight clear differences in turning movements: young individuals exhibit fluid, single step turns; older individuals use a careful 2-step approach; and patients with PD show irregular, disrupted turning patterns.

Angle of Both Hips

The ROM angle while turning the body is used to determine the angle of both hips (Figure 12). The positions of the hip, knee, and shoulder are used to find the angle formed at the hip joint. This measurement is crucial for assessing how the hip moves during turns, providing insights into the flexibility and coordination of the lower body. Understanding the hip ROM is especially valuable for identifying movement patterns in different populations, such as detecting mobility issues or assessing gait in individuals with PD.

Figure 12. An illustration of the hip angle position.



The hip angle is calculated as follows (equation 8):

The graphs showing the cosine of the hip angle are shown in Figure 12. The figure shows 3-line graphs comparing hip angles during body turns for the 3 groups. Each graph represents the hip angles for both the left and right sides, with lines depicting both the original and smoothed data.

The graph of the younger individuals group (Figure 12A) indicates relatively stable and moderate fluctuations in hip angles during body turns. Both hips show consistent movement with shallow peaks and troughs, reflecting balanced and controlled motion. Younger individuals maintain a steady range of hip angles, suggesting efficient and coordinated hip movement when turning.

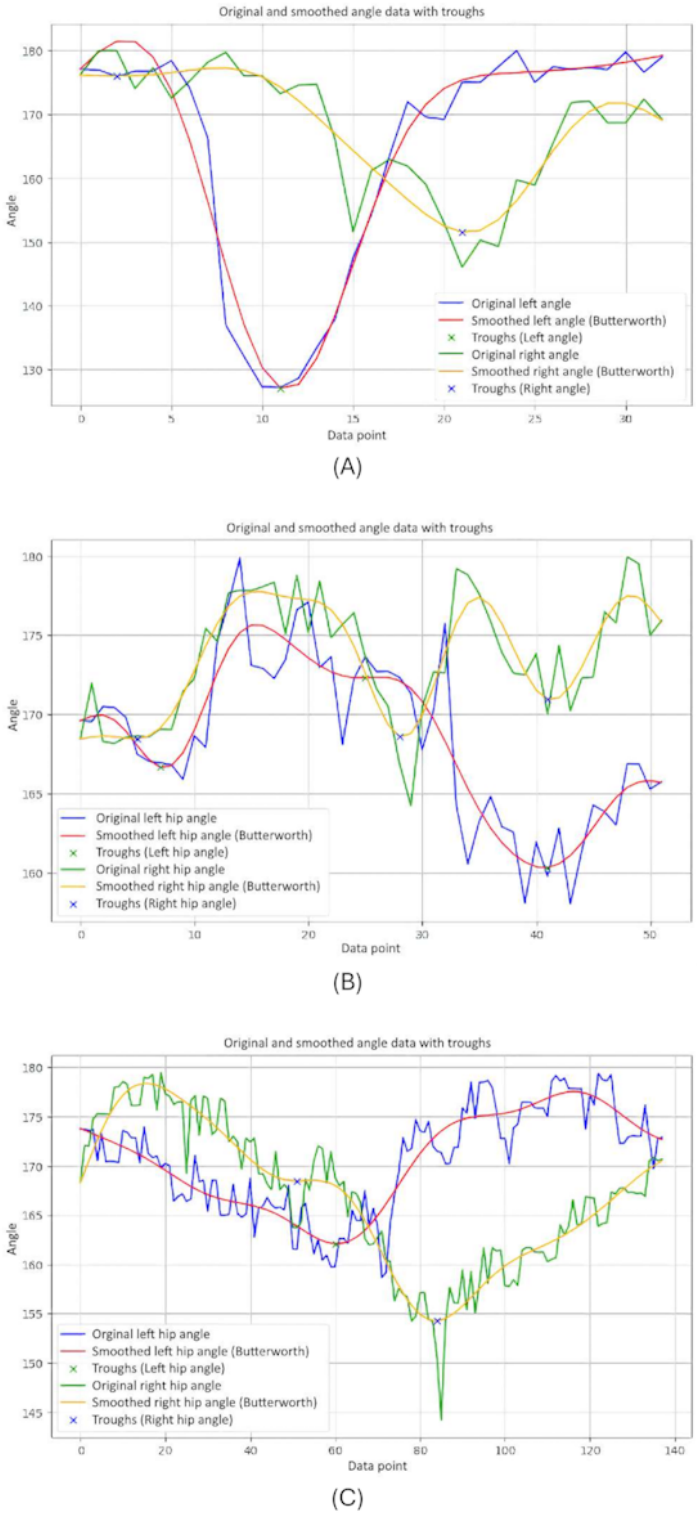
In the older-individual group graph (Figure 12B), the hip angles display more pronounced and variable fluctuations compared to the young group. The curves reveal that older individuals experience larger and more variable hip movements during turns. This increased variability suggests that older people adjust their hip movements significantly to maintain balance and stability, resulting in less smooth and more fluctuating hip angles.

The group of individuals with PD (Figure 12C) shows significant irregularity and instability in hip angles. The curves exhibit erratic and sharp changes, highlighting the difficulty individuals with PD face in maintaining consistent hip movement during turns. The frequent and abrupt fluctuations in hip angles are characteristic of the disease, where symptoms such as rigidity and tremors disrupt smooth and coordinated movements.

Figure 13 illustrates clear differences in hip angle dynamics during turns among the 3 groups. Younger individuals (Figure 13A) have controlled and consistent hip movements, older individuals (Figure 13B) show more variability and larger fluctuations in hip angles, and those with PD (Figure 13C) experience erratic and unstable hip movements. These patterns reflect how age and neurological conditions affect the coordination and efficiency of hip movement during body turns.

XSL•FO
RenderX

Figure 13. A comparative analysis of the hip angle between different groups: (A) younger individuals, (B) older individuals, and (C) individuals with Parkinson disease.

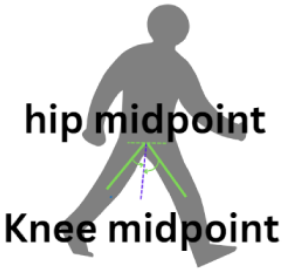


Leg Symmetry

The symmetry of leg movements is calculated using the angles at the knee joints (Figure 14). The goal was to measure and evaluate the balance and symmetry in the leg movements, which is crucial for understanding gait and stability, especially during turning motions. This information is particularly useful for analyzing the movement patterns in different populations, such as detecting gait irregularities in individuals with conditions such as PD.

XSL•FO
RenderX

Figure 14. An illustration of the symmetrical leg position.



The leg angle is computed as follows (equations 9-11):

Next, the angles of the left and right legs are compared:

The leg symmetry is then assessed as:

Figure 15 displays 3-line graphs comparing the leg symmetry during body turns for the 3 participant groups. Each graph plots the symmetrical leg movements, showing the original and smoothed angles of the left and right legs as they turn.

The graph of the younger-individual group (Figure 15A) shows closely aligned curves for the left and right leg angles, indicating high symmetry in their leg movements during turns. Both legs exhibit similar patterns with minimal deviation between them. This close alignment reflects the balanced and coordinated movements typical of young, healthy individuals, suggesting that they maintain a consistent and symmetrical gait.

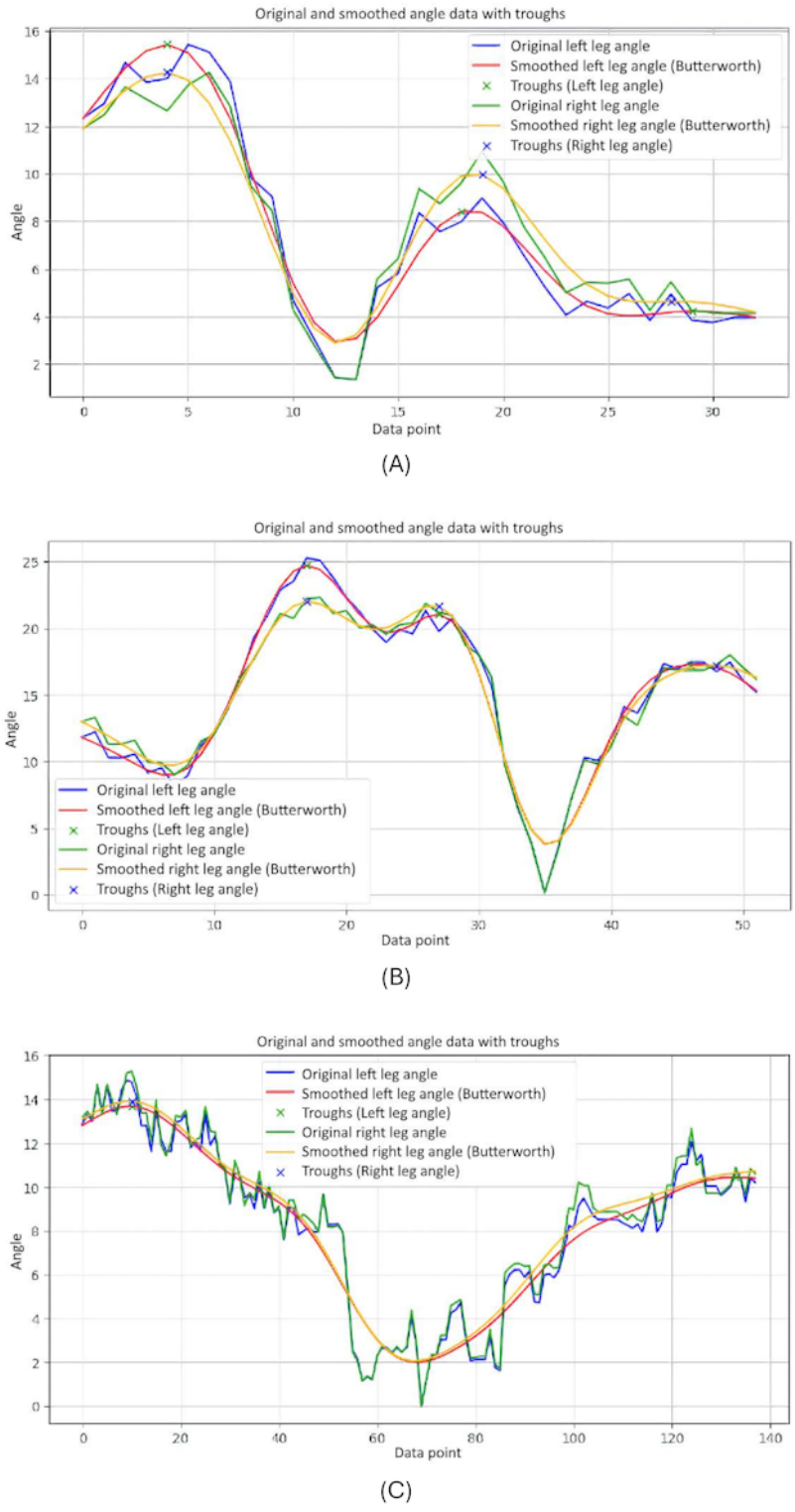
The older-individual group graph (Figure 15B) reveals more noticeable fluctuations between the left and right leg angles. While the overall patterns still follow a similar trajectory.

On the other hand, the graph of individuals with PD (Figure 15C) shows significant irregularities and less alignment between the left and right leg angles. The curves are erratic and display frequent sharp deviations, reflecting low symmetry in their leg movements. This lack of alignment is characteristic of PD, where symptoms such as tremors and rigidity cause disrupted and uncoordinated movements.

This comparison highlights how age, neurological conditions, and health status impact the ability to maintain balanced and symmetrical gait during turning motions.

XSL•FO
RenderX

Figure 15. A comparative analysis of the symmetrical leg between different groups: (A) younger individuals, (B) older individuals, and (C) individuals with Parkinson disease.

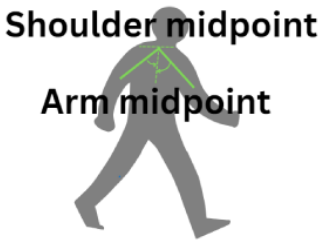


Arm Symmetry

The symmetry of arm movements was analyzed using the angles at the elbows to the shoulders (Figure 16). The purpose was to measure how balanced and coordinated the arm movements are, especially during activities such as turning the body. This assessment was important for understanding the ROM and the symmetry in arm movements, which can provide insights into overall body coordination and detect possible imbalances or movement disorders.

XSL•FO
RenderX

Figure 16. An illustration of the symmetrical arm position.



The calculation for arm symmetry was similar to that of finding the symmetry in legs (equation 12):



Figure 17 presents 3-line graphs comparing the arm symmetry during body movements for 3 groups. Each graph plots the angles of the left and right arms over time, showing both the original and smoothed data.

The graph of the younger-individual group (Figure 17A) displays closely aligned curves for the left and right arm angles. Both arms move in a highly synchronized manner, with minimal differences between the left and right angles throughout the motion.

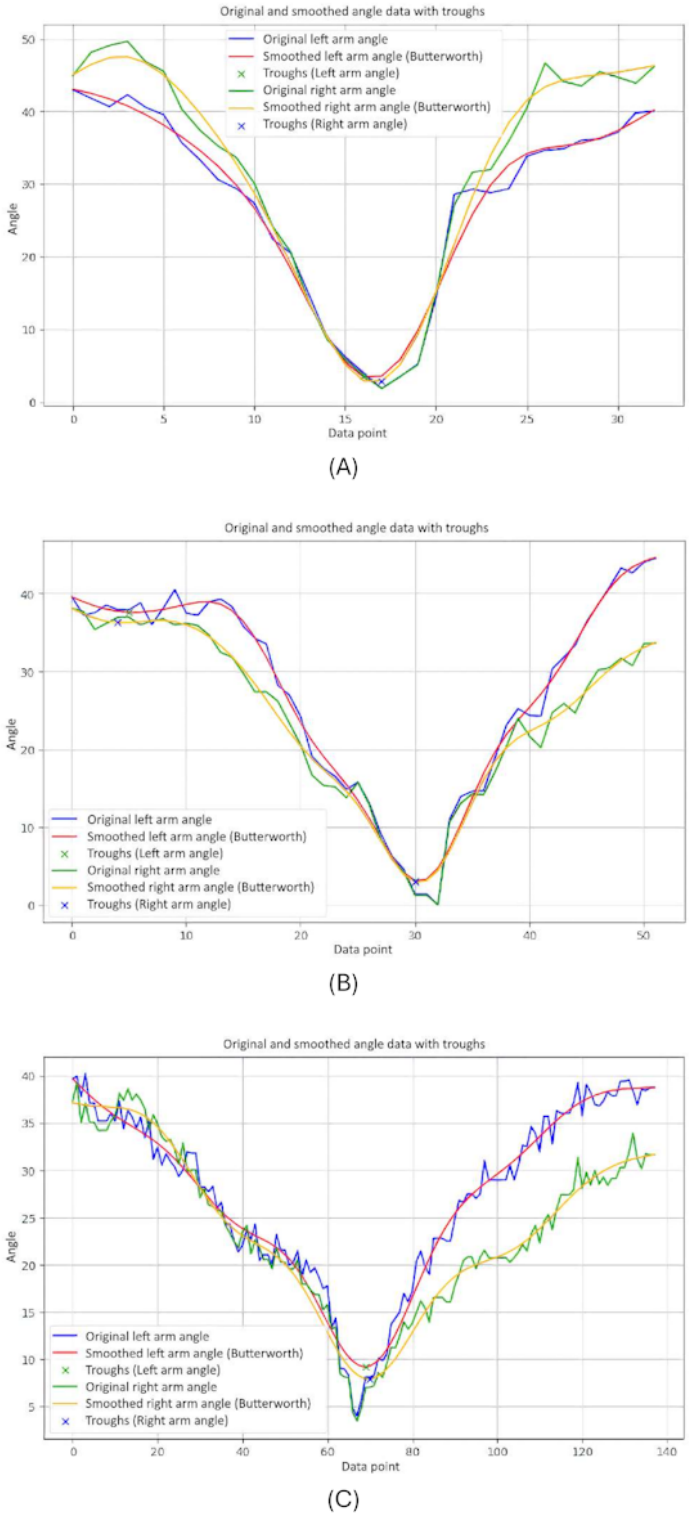
The older-individual group graph (Figure 17B) shows a slightly more varied pattern compared with the young group. While the overall movements of the left and right arms remain relatively synchronized, there are noticeable differences in the magnitude and timing of the angles.

The graph of individuals with PD (Figure 17C) exhibits significant irregularities and discrepancies between the left and right arm angles. The curves are erratic and often diverge sharply, indicating substantial asymmetry in arm movements.

We observed that younger individuals demonstrate highly synchronized and symmetrical arm movements, older individuals show moderate symmetry with some variability, and those with PD experience significant asymmetry and irregularity in their arm movements. These patterns underscore how age and neurological conditions impact the coordination and balance of arm movements during body turns.

XSL•FO
RenderX

Figure 17. A comparative analysis of the symmetrical arms between different groups: (A) younger individuals, (B) older individuals, and (C) individuals with Parkinson disease.

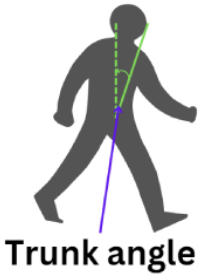


Trunk Angle 1 (Vertical)

The relative positions of the hips and nose are used to calculate the trunk angle (Figure 18). The trunk angle helps to understand how much the upper body tilts or bends relative to the lower body, especially during activities such as turning. This angle helps in assessing posture and balance, providing insights into the coordination and alignment of the trunk with the hips during movement. It is particularly useful in evaluating movement patterns and detecting postural deviations in various populations, including those with movement disorders such as PD.

XSL•FO
RenderX

Figure 18. An illustration of the trunk angle 1 vertical position.



The calculation of the trunk angle 1 is given as follows (equation 13):



Figure 19 displays 3 graphs comparing trunk angles relative to the vertical axis for 3 groups. Each graph shows the trunk angle over a series of data points during body movements, with both original and smoothed data represented. The trunk angle is used to observe the severity of hunchback, providing insights into the degree of forward bending or curvature of the upper body.

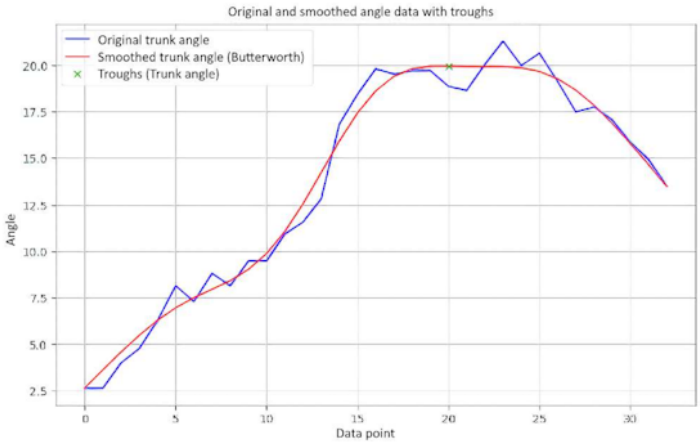
The graph of the younger-individual group (Figure 19A) exhibits a consistent and smooth progression of trunk angles. The trunk angle gradually increases and decreases within a narrow range, reflecting a well-coordinated and balanced posture with minimal forward bending. Younger individuals maintain a steady alignment with minimal deviations, indicating a lack of significant hunchback severity.

The older-individual group graph (Figure 19B) shows more variability in trunk angles compared to the young group. The curves display noticeable fluctuations and less consistency, with several abrupt changes in angle. These variations suggest that older individuals experience more difficulty in maintaining a stable trunk posture, leading to less smooth and more erratic movements. The increased trunk angles and variability may indicate a greater tendency toward forward bending, suggesting a moderate severity of hunchback as they struggle to maintain an upright posture.

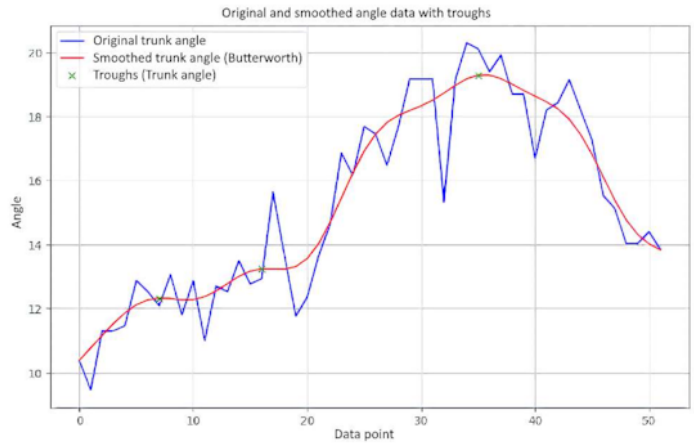
The graph of individuals with PD (Figure 19C) reveals significant irregularities and instability in trunk angles. The curves are highly erratic, with frequent sharp changes and a wide range of deviations from the vertical alignment. These abrupt shifts indicate substantial difficulties in maintaining consistent trunk posture. Individuals with PD struggle with controlling their trunk movements, leading to frequent tilting and misalignment relative to the vertical axis. The pronounced forward bending and high variability in trunk angles reflect severe hunchback, characteristic of the disease's impact on posture and movement control.

XSL•FO
RenderX

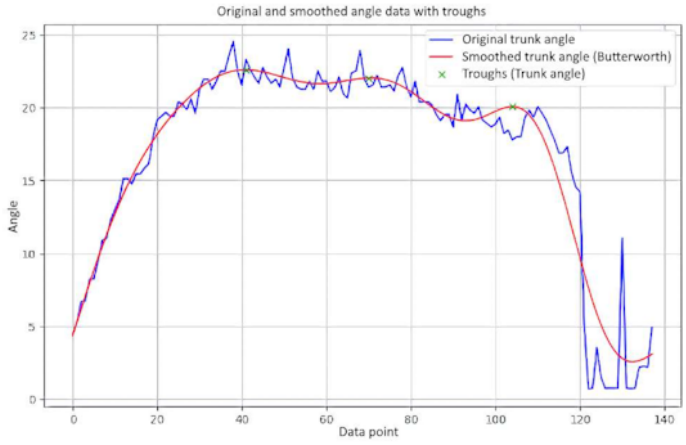
Figure 19. A comparative analysis of the trunk angle 1 between different groups: (A) younger individuals, (B) older individuals, and (C) individuals with Parkinson disease.



(A)



(B)



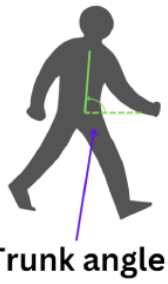
(C)

Trunk Angle 2 (Horizontal)

Trunk angle 2 analyzes the positions of the shoulders and hips (Figure 20). This angle is crucial for understanding the alignment of the trunk relative to the lower body, particularly during movements such as turning. By evaluating the trunk angle, we can assess the severity of hunchback or forward bending, which provides valuable insights into posture and stability.

XSL•FO
RenderX

Figure 20. An illustration of the trunk angle 2 horizontal position.



The calculation for trunk angle 2 is similar to that for trunk angle 1. The only difference is that as trunk angle 2 is horizontal, the default value must be set at the *x* coordinate instead of the keypoint. The formula to calculate trunk angle 2 is given as follows (equation 14):

Figure 21 shows graphs comparing trunk angles relative to the horizontal axis for 3 groups. These angles help assess the severity of forward bending or misalignment.

The graph of the younger-individual group (Figure 21A) exhibits stable and controlled trunk angles with minimal deviations, indicating well-balanced posture and low severity of hunchback.

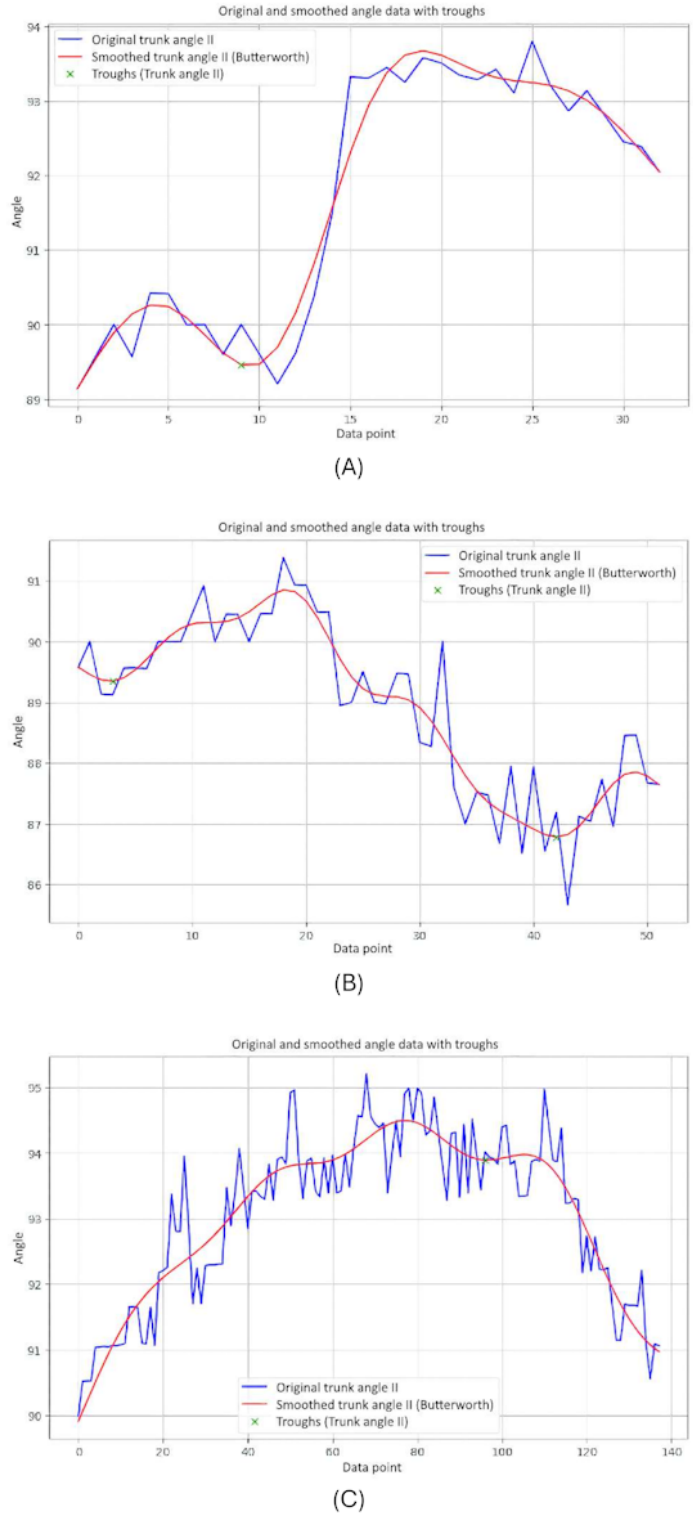
The graph of the older individuals' group (Figure 21B) showed more variability and sudden changes in trunk angle, reflecting moderate forward lean and greater difficulty maintaining stable horizontal alignment, indicating moderate kyphosis severity.

The graph of the individuals with PD (Figure 21C) shows significant instability and frequent sharp changes in trunk angles, indicating severe difficulties in maintaining consistent trunk posture and severe hunchback severity due to the disease's impact on movement control.

Overall, younger individuals maintain a low severity of hunchback, older individuals exhibit moderate severity, and those with PD show severe hunchback, reflecting their challenges in maintaining horizontal alignment.

XSL•FO
RenderX

Figure 21. A comparative analysis of the trunk angle 2 horizontal between different groups: (A) younger individuals, (B) older individuals, and (C) individuals with Parkinson disease.

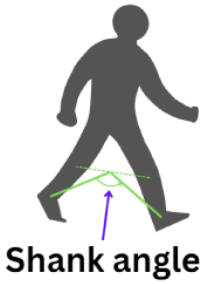


Shank Angle

The shank angle computes the positions of the knees and ankles (Figure 22). This angle is useful to understand the alignment and movement of the lower leg, especially during activities such as turning. By evaluating the shank angle, we can assess the coordination and stability of the lower leg movements, providing valuable insights into gait patterns and detecting any abnormalities in movement, which is particularly useful for analyzing conditions such as PD.

XSL•FO
RenderX

Figure 22. An illustration of the shank angle position.



The formula to calculate shank angle is provided in equation 15:

Figure 23 presents the graphs comparing shank angles during body movements for 3 groups. Each graph shows the shank angle over a series of data points, with both original and smoothed data represented.

The graph of the younger-individual group (Figure 23A) demonstrates relatively stable and consistent shank angles. Both the original and smoothed curves align closely, indicating smooth and coordinated movements. The angles show moderate fluctuations within a narrow range, reflecting balanced and controlled lower leg movements typical of healthy young individuals.

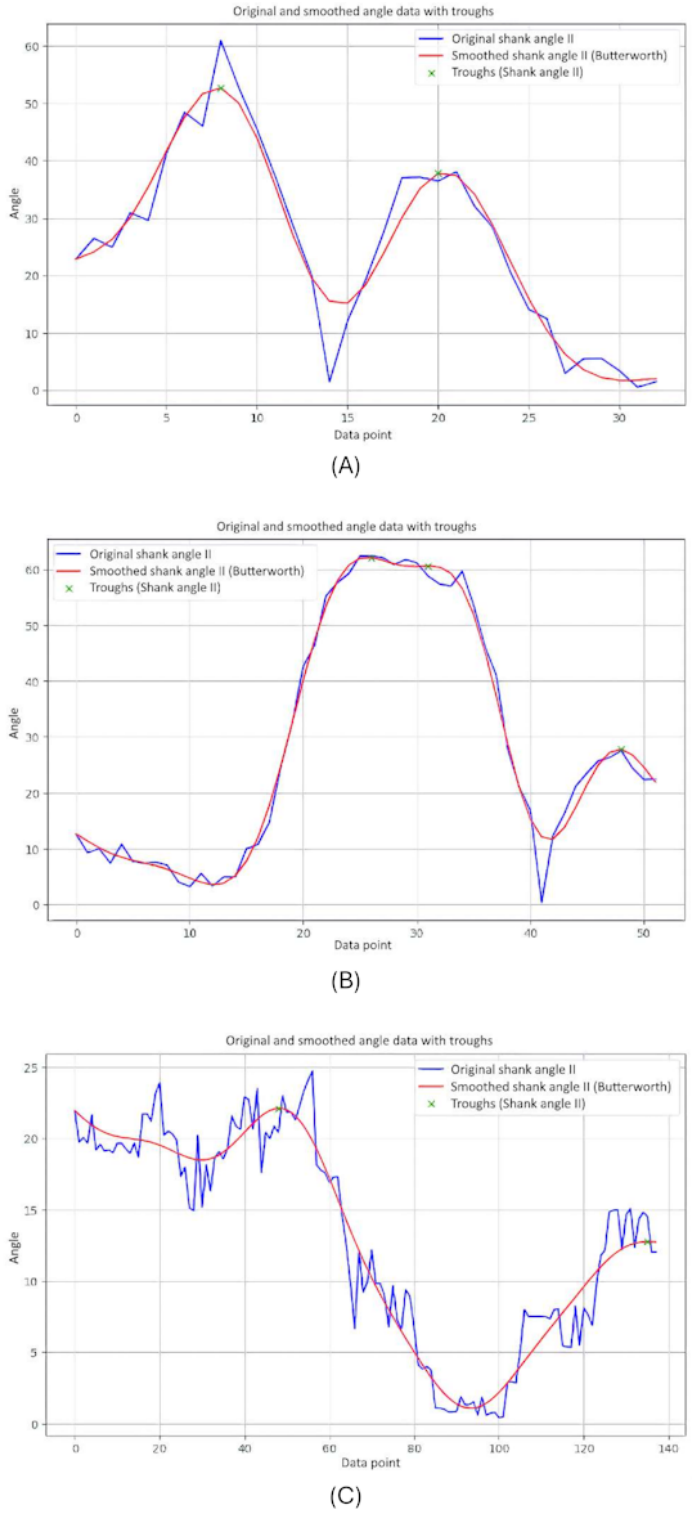
The older-individual group graph (Figure 23B) shows shank angles that are stable and consistent, such as the younger individuals group. Both the original and smoothed curves align closely, indicating smooth and coordinated lower leg movements. The angles show moderate fluctuations within a narrow range, reflecting balanced and controlled movements. This suggests that older individuals perform nicely, maintaining stable shank angles comparable to the younger individuals group.

The graph of individuals with PD (Figure 23C) reveals significant irregularities and instability in shank angles. The curves are highly erratic, with frequent sharp changes and a wide range of deviations. These abrupt shifts indicate substantial difficulties in maintaining consistent lower leg movements. Individuals with PD struggle with controlling their shank angles, leading to frequent misalignment and instability. This pattern is characteristic of the disease, where symptoms such as rigidity and tremors disrupt smooth and coordinated movements.

Figure 23 highlights the differences in shank angle behavior among the 3 groups. Younger and older individuals exhibit stable and coordinated shank movements with minimal variability, indicating balanced and controlled lower leg movements. In contrast, those with PD display significant instability and erratic shank movements, underscoring the impact of the condition on lower leg control and stability.

XSL•FO
RenderX

Figure 23. A comparative analysis of the shank angle between different groups: (A) younger individuals, (B) older individuals, and (C) individuals with Parkinson disease.



Classification Report

A comprehensive evaluation of the SVM model is provided by the classification report for the test set, as shown in Table 3. The report includes the overall accuracy of the model, as well as each class's recall, precision, and F_1-score. Although the training accuracy was 0.95, the test accuracy was 0.89, indicating that the model fit the training data well with minimal overfitting. The classification report shows that class 0 (normal class) had high precision and recall, while class 1 (PD class) has lower recall. This disparity in recall may be due to the smaller sample size of class 1.

XSL•FO
RenderX

Table 3. Classification report for distinguishing normal and PD[a] classes using gait features (test accuracy: 0.89 and train accuracy: 0.95).

Classes or metrics	Precision	Recall	F_1-score	Support
0 (normal)	0.88	1.00	0.93	7.00
1 (PD)	1.00	0.50	0.67	2.00
Accuracy	—[b]	—	0.89	9.00
Macroaverage	0.94	0.75	0.80	9.00
Weighted average	0.90	0.89	0.87	9.00

[a]PD: Parkinson disease.

[b]Not available.

AUC and Loss Curve

Figure 24 provides specifics regarding the model's capacity to distinguish between the 2 classes. Moderate discrimination capability is indicated by the test set's AUC of 0.6429. An indicator of how well the model performs across various training data subsets is provided by the cross-validation scores. With a mean cross-validation score of 0.9, the model's cross-validation scores are (0.75, 1.00, 1.00, 0.75, 1.00). This shows that the model continues to work well when the data are folded in different ways.

Figure 24. The receiver operating characteristic (ROC) curve.

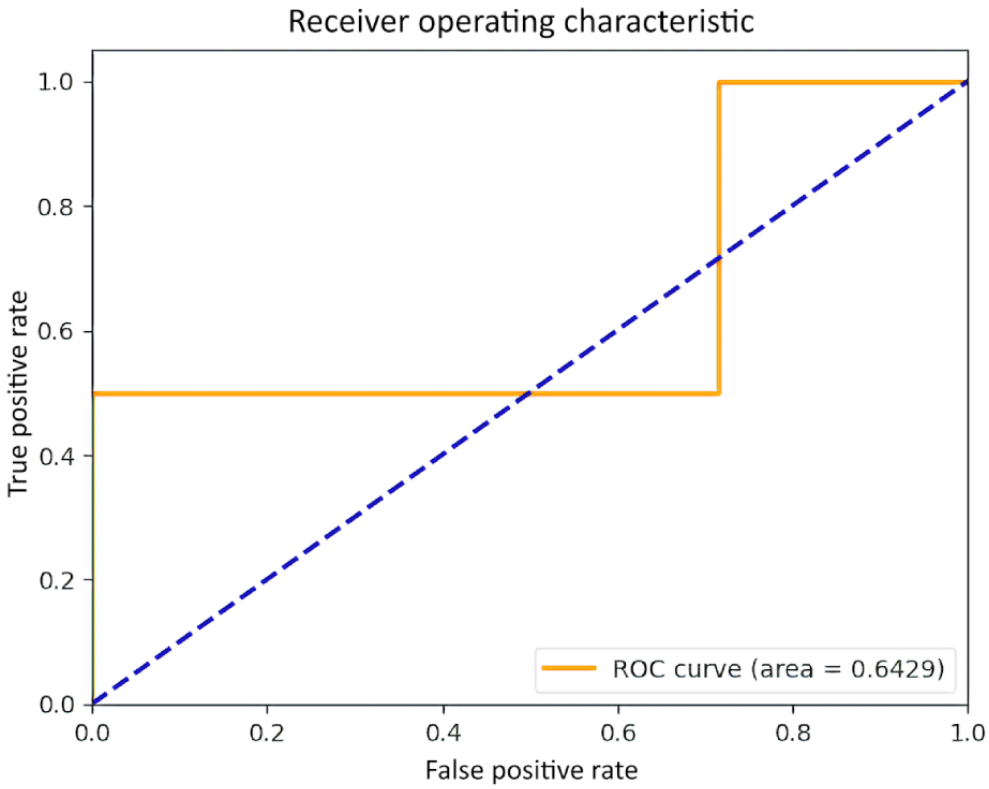
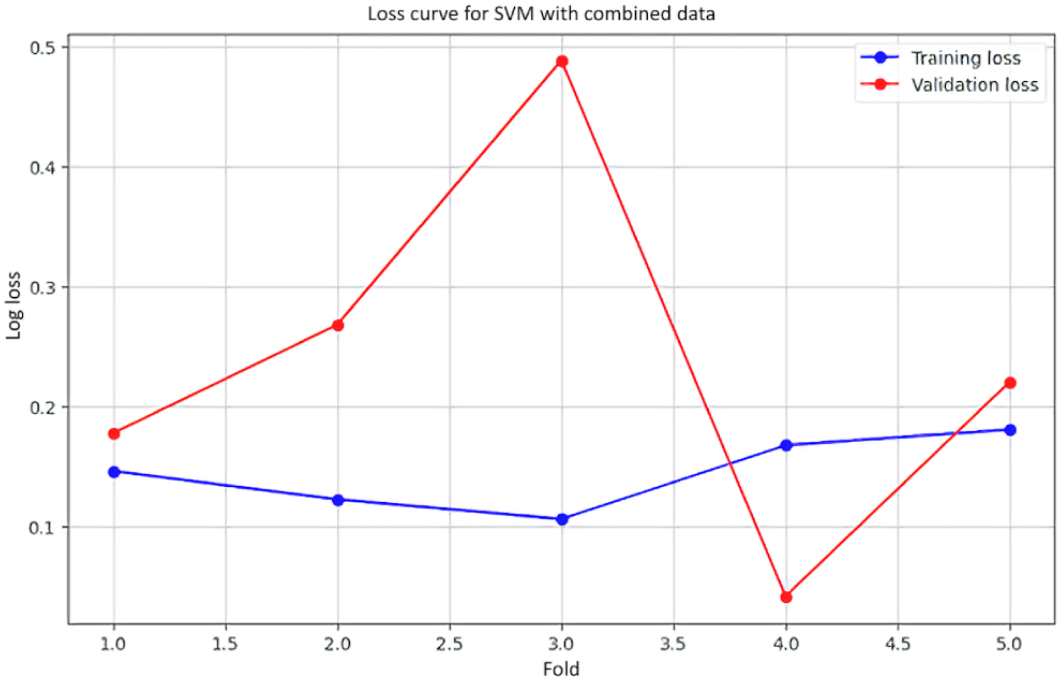


Figure 25 shows the training and validation loss curve, illustrating the log loss for both training and validation sets across various folds of cross-validation. The blue curve indicates the training loss, which stays relatively stable with minor fluctuations. In contrast, the red curve, representing the validation loss, exhibits significant fluctuations, peaking notably at fold 3. This implies that while the model fits the training data well, it experiences considerable variability in performance on unseen data, suggesting possible overfitting or sensitivity to specific data subsets. The difference between the training and validation loss also underscores the model's difficulty in generalizing from the training set to the validation set, highlighting the need for potential adjustments in model complexity or training strategy to achieve more consistent performance.

XSL•FO
RenderX

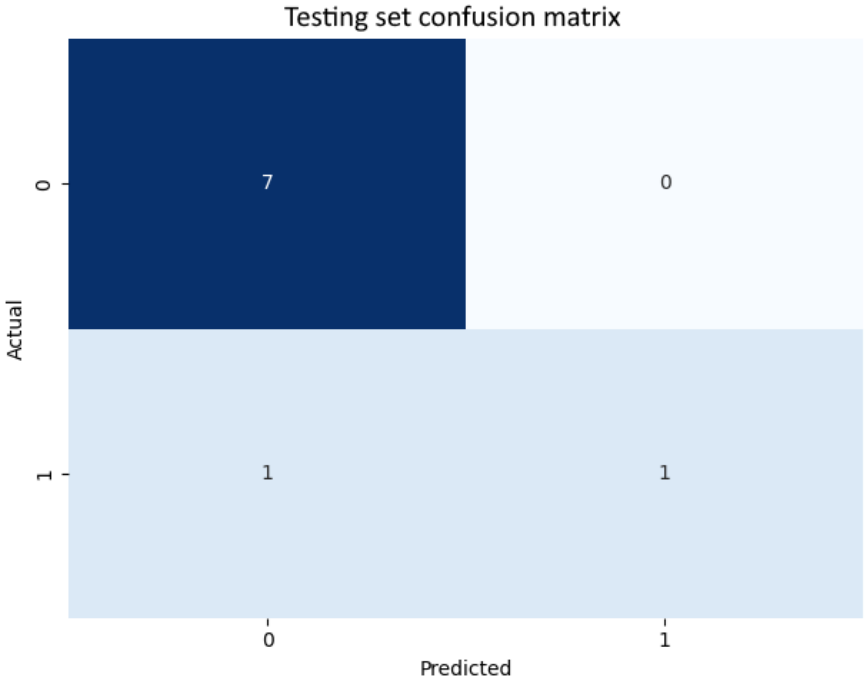
Figure 25. The training and validation loss curve for support vector machine (SVM) with combined data.



Confusion Matrix

The confusion matrix depicted in Figure 26 provides a thorough analysis of the predictions of the model on the test set. It reveals that while the model correctly predicted 7 out of 7 occurrences of class 0 (non-PD), it incorrectly classified 1 out of 2 instances of class 1 (PD). This suggests that although the model is highly accurate in predicting class 0, it faces challenges with class 1, likely due to class imbalance in the dataset.

Figure 26. The testing set confusion matrix for model-based features.



Discussion

This study presents a noninvasive approach for early detection of PD through the analysis of model-based gait features. Using kinematic characteristics, such as shoulder distance, step length, stride length, knee and hip angles, leg and arm symmetry, and trunk angles, we aimed to identify subtle gait abnormalities associated with PD. Data were collected through controlled video recordings of the TUG assessment, and the extracted features were processed using advanced filtering techniques and analyzed using SVM classifier.

XSL•FO
RenderX

The results demonstrate that the model-based features were highly effective in distinguishing between normal and PD-affected gait patterns, achieving high accuracy, precision, recall, and F_1-score. These findings support the potential of model-based gait analysis as a noninvasive, accessible tool for early diagnosis and monitoring of PD. The study also highlights the importance of addressing class imbalance and refining the feature extraction process to further enhance the performance of the classifier.

Future research should focus on validating these findings with larger datasets, exploring other machine learning models, and integrating additional features to improve the robustness and accuracy of PD detection systems. This approach could significantly benefit clinical settings, providing a reliable, noninvasive method for early diagnosis and improved patient outcomes.

Acknowledgments

This research is supported by the Fundamental Research Grant Scheme (FRGS/1/2020/ICT02/MMU/02/5) and Malaysian-Jordan Matching Fund.

Conflicts of Interest

None declared.

References

1. Challa KN, Pagolu VS, Panda G, Majhi B. An improved approach for prediction of Parkinson's disease using machine learning techniques. In: Proceedings of the International Conference on Signal Processing, Communication, Power and Embedded System. 2016 Presented at: SCOPES 2016; October 3-5, 2016; Paralakhemundi, India. [doi: 10.1109/scopes.2016.7955679]
2. Polat K. Freezing of gait (FoG) detection using logistic regression in Parkinson's disease from acceleration signals. In: Proceedings of the Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science. 2019 Presented at: EBBT 2019; April 24-26, 2019; Istanbul, Turkey. [doi: 10.1109/ebbt.2019.8742042]
3. Vidya B, Sasikumar P. Gait based Parkinson’s disease diagnosis and severity rating using multi-class support vector machine. Appl Soft Comput 2021 Dec;113:107939. [doi: 10.1016/j.asoc.2021.107939]
4. Fang Z. Improved KNN algorithm with information entropy for the diagnosis of Parkinson's disease. In: Proceedings of the International Conference on Machine Learning and Knowledge Engineering. 2022 Presented at: MLKE 2022; February 25-27, 2022; Guilin, China. [doi: 10.1109/mlke55170.2022.00024]
5. Gundala S, Harichandana C, Jaasmitha A, Brundha PL. Parkinson’s disease identification by voice and handwritten drawings using Xgboost and Random Forest Algorithms. In: Proceedings of the 3rd International Conference on Computing and Information Technology. 2023 Presented at: ICCIT 2023; September 13-14, 2023; Tabuk, Saudi Arabia. [doi: 10.1109/iccit58132.2023.10273905]
6. Govindu A, Palwe S. Early detection of Parkinson's disease using machine learning. Procedia Comput Sci 2023;218:249-261. [doi: 10.1016/j.procs.2023.01.007]
7. Pereira CR, Weber SA, Hook C, Rosa GH, Papa JP. Deep learning-aided Parkinson's disease diagnosis from handwritten dynamics. In: Proceedings of the 29th SIBGRAPI Conference on Graphics, Patterns and Images. 2016 Presented at: SIBGRAPI 2016; October 4-7, 2016; Sao Paulo, Brazil. [doi: 10.1109/sibgrapi.2016.054]
8. Grover S, Bhartia S, Akshama, Yadav A, KR S. Predicting severity of Parkinson’s disease using deep learning. Procedia Comput Sci 2018;132:1788-1794. [doi: 10.1016/j.procs.2018.05.154]
9. Aal HA, Taie S, El-Bendary N. An optimized RNN-LSTM approach for Parkinson’s disease early detection using speech features. Bull Electr Eng Inform 2021 Oct;10(5):2503-2512. [doi: 10.11591/eei.v10i5.3128]
10. Ouhmida A, Terrada O, Raihani A, Cherradi B, Hamida S. Voice-based deep learning medical diagnosis system for parkinson's disease prediction. In: Proceedings of the International Congress of Advanced Technology and Engineering. 2021 Presented at: ICOTEN 2021; July 4-5, 2021; Taiz, Yemen. [doi: 10.1109/icoten52080.2021.9493456]
11. Biswas S, Kaur N, KR S. Early detection of Parkinson’s disease from hand drawings using CNN and LSTM. In: Proceedings of the 4th International Conference on Artificial Intelligence and Speech Technology. 2022 Presented at: AIST 2022; December 9-10, 2022; Delhi, India. [doi: 10.1109/aist55798.2022.10065159]
12. Khaskhoussy R, Ayed YB. Improving Parkinson’s disease recognition through voice analysis using deep learning. Pattern Recognit Lett 2023 Apr;168:64-70. [doi: 10.1016/j.patrec.2023.03.011]

Abbreviations

ANN: artificial neural network
AUC: area under the curve
CNN: convolutional neural network

XSL•FO
RenderX

DNN: deep neural network
FoG: freezing of gait
KNN: k-nearest neighbors
LSTM: long short-term memory
PD: Parkinson disease
ROM: range of motion
SVM: support vector machine
TUG: timed up and go
UPDRS: Unified Parkinson Disease Rating Scale

Edited by G Seçkin; submitted 21.08.24; peer-reviewed by A Samraj, PS Teh, TT Verlekar; comments to author 18.11.24; revised version received 26.11.24; accepted 20.01.25; published 08.04.25.

Please cite as:
Lim MD, Connie T, Goh MKO, Saedon N'
Model-Based Feature Extraction and Classification for Parkinson Disease Screening Using Gait Analysis: Development and Validation Study
JMIR Aging 2025;8:e65629
URL: https://aging.jmir.org/2025/1/e65629
doi:10.2196/65629
PMID:

©Ming De Lim, Tee Connie, Michael Kah Ong Goh, Nor 'Izzati Saedon. Originally published in JMIR Aging (https://aging.jmir.org), 08.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Identifying Deprescribing Opportunities With Large Language Models in Older Adults: Retrospective Cohort Study

Vimig Socrates[1,2*], MS; Donald S Wright[3,4*], MD, MHS; Thomas Huang[3], BS; Soraya Fereydooni[3], BS; Christine Dien[2], BS; Ling Chi[3], BS; Jesse Albano[5], PharmD, BCPS, BCEMP; Brian Patterson[6], MD, MPH; Naga Sasidhar Kanaparthy[3,4], MD, MPH; Catherine X Wright[7], MD; Andrew Loza[1,3], MD, PhD; David Chartash[1,8], PhD; Mark Iscoe[1,3], MD, MHS; Richard Andrew Taylor[1,3,9], MD, MHS

[1]Department of Biomedical Informatics and Data Science, School of Medicine, Yale University, New Haven, CT, United States

[2]Program of Computational Biology and Bioinformatics, Yale University, New Haven, CT, United States

[3]Department of Emergency Medicine, School of Medicine, Yale University, New Haven, CT, United States

[4]VA Connecticut Healthcare System, US Department of Veterans Affairs, West Haven, CT, United States

[5]Department of Pharmacy, Yale New Haven Hospital, New Haven, CT, United States

[6]BerbeeWalsh Department of Emergency Medicine, University of Wisconsin–Madison, Madison, WI, United States

[7]Section of Cardiovascular Medicine, Department of Internal Medicine, Yale School of Medicine, New Haven, CT, United States

[8]School of Medicine, University College Dublin, Dublin, Ireland

[9]Department of Biostatistics, School of Public Health, Yale University, New Haven, CT, United States

[*]these authors contributed equally

Corresponding Author:
Richard Andrew Taylor, MD, MHS
Department of Biomedical Informatics and Data Science
School of Medicine
Yale University
464 Congress Avenue
Suite 260
New Haven, CT, 06510
United States
Phone: 1 2037854058
Email: richard.taylor@yale.edu

Abstract

Background: Polypharmacy, the concurrent use of multiple medications, is prevalent among older adults and associated with increased risks for adverse drug events including falls. Deprescribing, the systematic process of discontinuing potentially inappropriate medications, aims to mitigate these risks. However, the practical application of deprescribing criteria in emergency settings remains limited due to time constraints and criteria complexity.

Objective: This study aims to evaluate the performance of a large language model (LLM)–based pipeline in identifying deprescribing opportunities for older emergency department (ED) patients with polypharmacy, using 3 different sets of criteria: Beers, Screening Tool of Older People's Prescriptions, and Geriatric Emergency Medication Safety Recommendations. The study further evaluates LLM confidence calibration and its ability to improve recommendation performance.

Methods: We conducted a retrospective cohort study of older adults presenting to an ED in a large academic medical center in the Northeast United States from January 2022 to March 2022. A random sample of 100 patients (712 total oral medications) was selected for detailed analysis. The LLM pipeline consisted of two steps: (1) filtering high-yield deprescribing criteria based on patients' medication lists, and (2) applying these criteria using both structured and unstructured patient data to recommend deprescribing. Model performance was assessed by comparing model recommendations to those of trained medical students, with discrepancies adjudicated by board-certified ED physicians. Selective prediction, a method that allows a model to abstain from low-confidence predictions to improve overall reliability, was applied to assess the model's confidence and decision-making thresholds.

XSL•FO
RenderX

Results: The LLM was significantly more effective in identifying deprescribing criteria (positive predictive value: 0.83; negative predictive value: 0.93; McNemar test for paired proportions: χ^2_1=5.985; *P*=.02) relative to medical students, but showed limitations in making specific deprescribing recommendations (positive predictive value=0.47; negative predictive value=0.93). Adjudication revealed that while the model excelled at identifying when there was a deprescribing criterion related to one of the patient's medications, it often struggled with determining whether that criterion applied to the specific case due to complex inclusion and exclusion criteria (54.5% of errors) and ambiguous clinical contexts (eg, missing information; 39.3% of errors). Selective prediction only marginally improved LLM performance due to poorly calibrated confidence estimates.

Conclusions: This study highlights the potential of LLMs to support deprescribing decisions in the ED by effectively filtering relevant criteria. However, challenges remain in applying these criteria to complex clinical scenarios, as the LLM demonstrated poor performance on more intricate decision-making tasks, with its reported confidence often failing to align with its actual success in these cases. The findings underscore the need for clearer deprescribing guidelines, improved LLM calibration for real-world use, and better integration of human–artificial intelligence workflows to balance artificial intelligence recommendations with clinician judgment.

(JMIR Aging 2025;8:e69504) doi:10.2196/69504

KEYWORDS

deprescribing; large language models; geriatrics; potentially inappropriate medication list; emergency medicine; natural language processing; calibration

Introduction

Polypharmacy, widely defined as the regular use of at least 5 medications, is common in older adults and at-risk populations [1]. In fact, approximately 30% of patients aged 65 years or older have polypharmacy [2], and nearly half of older emergency department (ED) patients are discharged with one or more new medications [3]. Although necessary and beneficial for some patients, polypharmacy can increase the risk of negative consequences for patients, including ED visits, adverse drug events, falls, disability, and inappropriate medication use [1]. While definitions differ, deprescribing is generally defined as a structured process by which potentially inappropriate medications (PIMs) are identified and withdrawn under the supervision of a health care provider. In some definitions, the process is described as evaluating the risk-benefit tradeoff, focusing on situations where the potential or actual harms of a medication outweigh its benefits, considering the patient's individual care goals and quality of life [2,4,5].

Deprescribing tools, such as the Screening Tool of Older People's Prescriptions (STOPP) [6] and Beers criteria [7], have been developed to help providers assess and identify PIMs based on a patient's medication list [7-9]. These explicit assessments are criterion-based with clear standards but are often impractical to implement in time-constrained clinical settings, such as ED, due to the need to evaluate multiple clinical indications and specialist-prescribed medications [10]. Attempts to digitize these criteria into electronic clinical decision support (CDS) have raised difficulties, typically requiring a labor-intensive coding process and unstructured information such as free text from patient records to contextualize certain criteria [11,12].

Large language models (LLMs) have been shown to interpret complex clinical situations and offer recommendations, from differential diagnoses to care management, leading to growing interest in their application in the medical field [13-16]. Moreover, they have been shown to extract medication-related data such as medication name, dosage, and frequency, necessary for the application of deprescribing criteria [17]. Finally, LLMs are excellent in-context learners, requiring very little labeled data to make predictions [18]. This reduces the annotation burden for time-constrained ED physicians while improving the use of unstructured patient records to contextualize patient medication lists. However, the majority of clinical reasoning evaluations on LLMs have been conducted using standardized exams (the United States Medical Licensing Examination) or digital case reports [14,19]. Their ability to perform clinical reasoning and calibrate responses over physician-generated text remains understudied.

Here, we propose to evaluate the performance of an LLM-based data pipeline in recommending deprescribing options for older adult ED patients at discharge based on 2 leading deprescribing criteria, Beers and STOPP. We have also included a recently developed list of criteria, Geriatric Emergency Medication Safety Recommendations (GEMS-Rx), intended to prevent the initiation of inappropriate medications in the acute care setting, as similar deprescribing lists specific to this care environment are not available [3]. Through this work, we hope to evaluate whether an LLM-based CDS system can effectively triage medications eligible for electronic deprescribing in older adults. Successful implementation of such a system would help address gaps in electronic deprescribing by using an LLM to contextualize recommendations within individual patient records and reduce manual development in CDS tools.

Methods

Ethical Considerations

This study was conducted with approval from the Institutional Review Board (IRB) at Yale University, under protocol number 2000035077. The IRB determined that this research qualifies for exemption as it involves secondary analysis of existing electronic health record (EHR) data, with no additional patient contact or data collection. The original data were collected with patient consent, and the current analysis adheres to the conditions of that consent and IRB approval, permitting

XSL•FO
RenderX

secondary analysis without the need for additional consent. All data were de-identified prior to analysis to ensure patient confidentiality. This study complies with ethical standards and guidelines for research involving human subjects.

Patient Cohort

All older adults (aged 65 years and older) with polypharmacy (5 or more active outpatient medications) presenting to an ED in a large academic medical center in the Northeast United States between January 2022 and March 2022 were identified. Due to budgetary constraints and a lack of prior evidence regarding the performance of LLMs in this task to guide a power calculation, we selected a random sample of 100 unique patients for evaluation.

Identification of Consensus-Based High-Yield Criteria

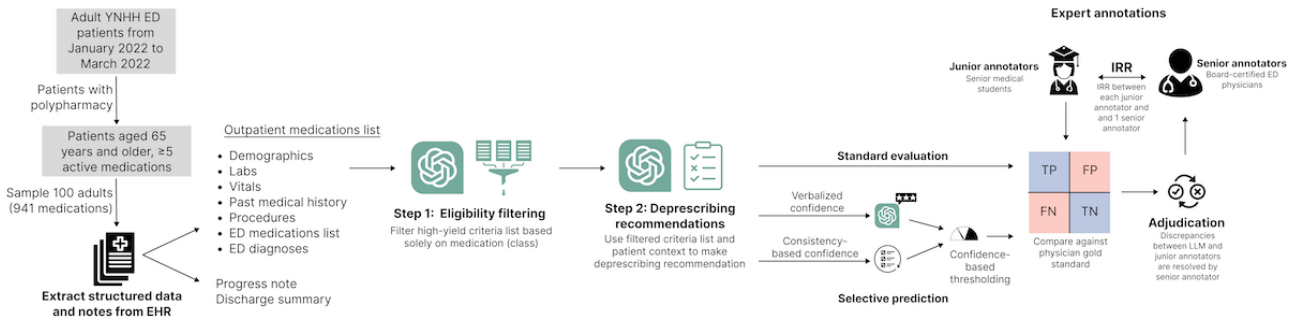
We conducted a consensus-based evaluation to filter three preexisting deprescribing lists (ie, STOPP [6], Beers [7], and GEMS-Rx [3]) into a focused set of high-yield deprescribing criteria for the LLM to use in its recommendations. High-yield criteria were defined as those posing a significant clinical risk to the patient and being identifiable within the electronic health records (EHRs). To identify these criteria, we evaluated 180 recommendations across 2 key dimensions: clinical risk and EHR computability. The consensus panel consisted of 6 board-certified physicians (in Emergency Medicine, Internal Medicine, and Clinical Informatics) and 1 ED pharmacist. Each member of the group individually reviewed each of the criteria and rated them on a 5-point Likert scale. We selected the top 50% of criteria with an average score greater than 3 on both dimensions, calculated across all experts, as high-yield criteria. We further elaborate on this consensus process and the final set of criteria in the Results section.

The need to filter the criteria before proceeding with the study was identified in our preliminary research [20], which revealed that one of the main causes of discrepancies between physicians and LLMs arose from ambiguous inclusion or exclusion conditions in deprescribing criteria. For example, criteria like "Statins for primary cardiovascular prevention in persons aged ≥85 with established frailty with expected life expectancy likely less than 3 years" include elements—such as "established frailty" and "expected life expectancy"—that are challenging to quantify and therefore difficult to implement computationally. The dimensions used to filter criteria were chosen to ensure that an LLM-enabled CDS tool prioritizes meaningful recommendations from high-quality EHR data, enabling accurate, actionable deprescribing recommendations and reduction of alert fatigue. We present the final set of high-yield criteria based on the average results of the consensus study.

Deprescribing Recommendations by GPT-4o

The study was approved under an exemption by the Yale University institutional review board prior to commencement (HIC# 2000035077). All patient-level data were deidentified prior to use with the LLM. We leveraged Microsoft's Azure OpenAI GPT-4o (GPT-4o model version: 2024-08-06 and OpenAI API version: 2024-02-15-preview) to produce deprescribing recommendations through a 2-stage process, as shown in Figure 1. In stage 1, GPT-4o was prompted to filter the full list of high-yield criteria solely based on the patient's medication list, ignoring inclusion or exclusion conditions. In stage 2, GPT-4o was prompted to use its previously filtered criteria list, along with structured (eg, demographics, lab values, vitals, and past medical history) and unstructured (most recent progress note and discharge summary) information, to determine if the patient satisfied any deprescribing criteria and the medication should be recommended for deprescribing. Each medication was evaluated individually to prevent errors from simultaneous processing, such as misattribution of criteria or medication omissions. To ensure optimal performance, we engineered prompts in an iterative fashion [21], using 1 patient at a time from a set of patients (up to 10% of the cohort) not used in the subsequent evaluation. After each evaluation, prompts were adjusted to correct any systematic errors (eg, instances where no relevant criteria in step 1 led to noncriteria-based deprescribing recommendations in step 2) by the LLM. After our third prompt yielded an output without any identifiable errors, we stopped the iterative prompt development process. Consequently, the final 2 patients initially reserved for this purpose were included in the final cohort evaluation (n=92 patients, 626 medications). Aside from the consistency-based method described later, all LLM calls were performed with a fixed temperature (temperature=0; low randomness in generated responses) and seed to ensure reproducibility and deterministic outputs.

Figure 1. Overview of the evaluation pipeline, consisting of a 2-step GPT-4o process, performance comparison with junior annotators (medical students), and final adjudication by senior annotators (board-certified physicians). ED: emergency department; EHR: electronic health record; FN: false negative; FP: false positive; IRR: interrater reliability; TN: true negative; TP: true positive; YNHH: Yale New Haven Hospital.



This 2-stage process was developed to correct errors identified when both stages were accomplished at once. Our initial testing revealed that providing the LLM with the full set of criteria and medications led to simple reasoning errors. The large number

XSL•FO
RenderX

of criteria, combined with the simultaneous processing of the complete patient medication list, resulted in inaccuracies in applying individual criteria to specific medications. Separating the process of criteria filtering and application both reduces confusion due to large input context sizes [22] and ensures that extraneous context does not distract the LLM [23]. The full prompts for step 1 and step 2 are included in Figures S5 and S6 in Multimedia Appendix 1.

Selective Prediction Methods

In addition to evaluating an LLM's ability to make deprescribing recommendations, we assessed whether its confidence estimates were well-calibrated and examined their impact on predictive performance. To do so, we collected GPT-4o's decision confidence for both steps using 2 validated confidence elicitation methods: chain-of-thought verbalized confidence and self-random sampling with average-confidence aggregation, referred to as consistency-based confidence [24]. In verbalized confidence, we asked the LLM to explicitly estimate its confidence for each step following its decision. For the consistency-based approach, we followed the best practices established in prior work [24] by sampling the LLM multiple times (number of samples=5) with high temperature (T=0.8; high randomness in generated responses) and used a majority vote weighted by the confidence of each response to determine the final deprescribing recommendation.

In a human-in-the-loop decision-making system, the LLM's confidence would be used to determine if the model should abstain due to low certainty regarding its own decision. In practice, this case would be considered too difficult for the LLM and forwarded to an expert reviewer. This human-in-the-loop decision-making pipeline is known as selective prediction and has been commonly found to improve performance in non–text-based applications [25,26]. We evaluated both selective prediction methods using risk-coverage curves [27,28], substituting risk for the F_1-score (a measure of the predictive performance of a model balancing precision and recall) to capture the full range of predictive performance. Coverage was also expressed inversely as the deferring fraction, representing the proportion of instances where the LLM abstained from making a decision. We conducted a more in-depth analysis of the method that proved to be more effective.

Comparison and Adjudication With Clinical Experts

In this study, we used a rigorous human review and adjudication process to assess model performance. Two trained senior medical students (M4) classified all medications in the test cohort using a 2-stage pipeline, after first computing interrater reliability (IRR) on an adjudication set of 75 medications from 5 patients. Similar to the LLM pipeline, for each medication, a medical student determined (1) if there exists a relevant high-yield criteria based on the medication list, and (2) whether the medication should be recommended for deprescribing. Discrepancies between the students (junior annotators) and the LLM across both stages were adjudicated by 2 board-certified ED physicians (senior annotators). Similarly to the junior annotators, we measured the IRR between the 2 senior ED physicians, prior to adjudication on the full set of discrepancies. Finally, we classified the errors leading to incorrect recommendations by the LLM, leveraging a prior evaluation framework [29]. We classified each error as 1 of 4 error types: incorrect reading comprehension, incorrect recall of knowledge, incorrect reasoning step, and not enough information, as described in Table 1.

Table 1. Definitions of GPT-4o error types inspired by framework from Lièven et al [29] relevant to deprescribing recommendations.

GPT-4o error types	Definition	Example
Incorrect reading comprehension	Includes misunderstanding of order of text, such as when a medication is dependent on another medication in a specific arrangement. Also includes ignoring information provided in the input text, such as missing a relevant category explicitly stated in the recommendations.	GPT failed to recognize acetaminophen by name from the list of STOPP[a] criteria in a patient at risk for malnutrition or liver disease.
Incorrect recall of knowledge	Includes failure to recognize classes of medications or other medical facts necessary to perform the task.	GPT correctly recognized amlodipine was a calcium channel blocker but failed to recognize it was more broadly an antihypertensive.
Incorrect reasoning step	Faulty reasoning, such as inappropriate assumptions or leads of logic unsupported by the clinical data.	GPT recommended discontinuing warfarin in a patient with a therapeutically elevated INR[b] after assuming that this elevated INR was due to a bleeding disorder.
Not enough information	Inappropriate application of missing data leading to potentially unreliable conclusions, such as assuming abnormality of a missing laboratory study.	GPT recommended discontinuing a QT[c] prolonging antidepressant based on the possibility of QT prolongation without any ECG[d] data or history of abnormal QT interval.

[a]STOPP: Screening Tool of Older People's Prescriptions.

[b]INR: International normalized ratio.

[c]QT: QT interval.

[d]ECG: electrocardiogram.

XSL•FO
RenderX

Data Analysis

We evaluated whether the LLM or the medical student was correct, using a gold standard derived from senior annotator (board-certified ED physicians) adjudication of discrepancies. To compare their proportions of correct responses, we applied the McNemar test, a statistical method commonly used to analyze paired nominal data, such as diagnostic accuracy from different assessments applied to the same cases [30]. All analysis was performed using Python (version 3.9; Python Software Foundation). Statistical testing was carried out using *statsmodels* (version 0.14.4) [31] and all visualizations were generated using *seaborn* (version 0.13.2) [32] and *matplotlib* (version 3.8.2) [33].

Results

Patient Cohort

In total, we identified 10,977 unique patients across 15,161 emergency department encounters from January 2022 to March 2022 meeting our selection criteria, from which 100 patients were randomly selected (Table 2). As our criteria only pertain to oral medication, nonoral medications were subsequently filtered out, resulting in 712 total oral medications across the cohort and a median of 6 oral medications per patient (Figure 2). From our initial study cohort of 100 patients, 10 patients were set aside for both prompt engineering and calculation of IRR between junior annotators. Fewer iterations were needed to refine the prompt than initially anticipated, so the remaining 2 patients were included in the final study cohort. This resulted in a final evaluation cohort of 92 patients, encompassing a total of 626 medications. Based on the mechanism of action, statins were the most common medication class (atorvastatin and rosuvastatin; 6.8% combined), followed by proton pump inhibitors (pantoprazole, esomeprazole, and omeprazole; 4.6% combined). When classified by therapeutic effect, antihypertensive agents were most prevalent (amlodipine, lisinopril, losartan, and valsartan; 10.6% combined).

Table 2. Demographic overview of the 100 patients included in the evaluation (N=100).

Characteristics	Values
Age, mean (SD)	75.8 (7.6)
Sex, n (%)	
Female	63 (63)
Male	37 (37)
Race, n (%)	
Asian	1 (1)
Black or African American	16 (16)
White or Caucasian	69 (69)
Other or not listed	13 (13)
None	1 (1)
Ethnicity, n (%)	
Hispanic or Latino	10 (10)
Non-Hispanic	90 (90)
Smoking status, n (%)	
Former smoker	47 (47)
Never smoker	43 (43)
Current every day smoker	8 (8)
Passive smoke exposure—never smoker	1 (1)
Light tobacco smoker	1 (1)
Number of medications, median (IQR)	6.0 (5.0-8.2)

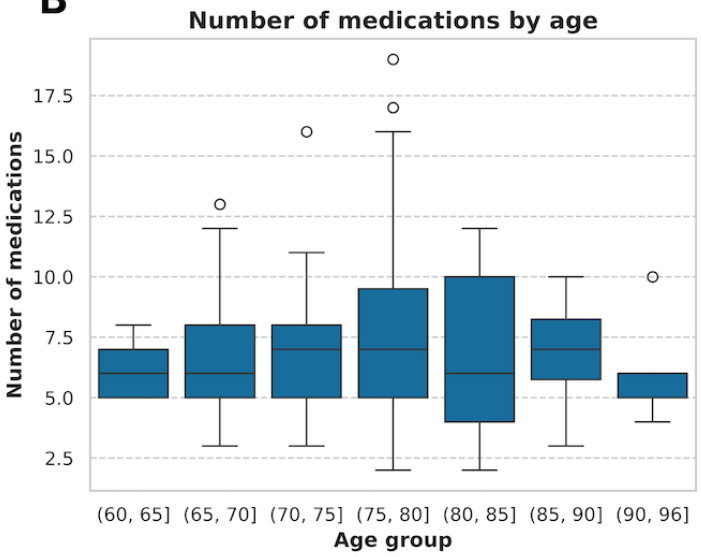
XSL•FO
RenderX

Figure 2. Medication information about baseline cohort. (A) The top 20 most commonly prescribed medications represented as a percentage of the total medication set. (B) Distribution of medications across different age groups.

A

Medication name	Total medications (%)
Amlodipine	4.6
Atorvastatin	4.2
Metoprolol succinate	3.3
Aspirin	3.3
Lisinopril	2.6
Rosuvastatin	2.6
Hydrochlorothiazide	2.4
Losartan	2.1
Gabapentin	2.1
Famotidine	1.8
Prednisone	1.8
Escitalopram	1.8
Sennosides	1.6
Lorazepam	1.6
Pantoprazole	1.6
Esomeprazole	1.5
Apixaban	1.5
Acetaminophen	1.5
Omeprazole	1.5
Valsartan	1.3

B



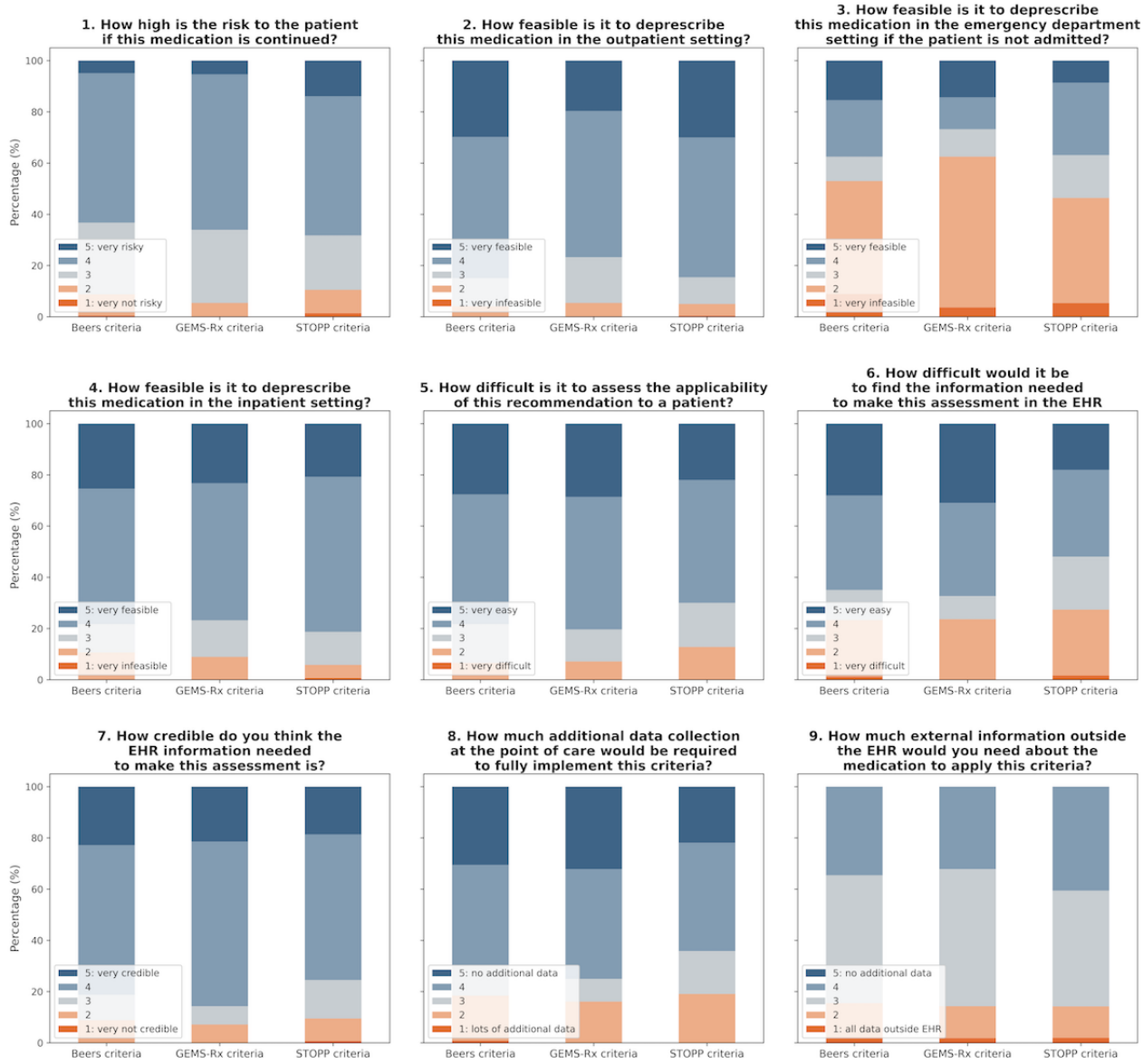
Evaluation of Consensus-Based High-Yield Criteria

To streamline evaluation by the LLM, we filtered criteria using the average scores from an expert consensus panel. All criteria (n=180) were evaluated based on their scores for clinical risk (Q1) and EHR computability (Q2-Q5), as assessed by an expert panel; average scores by deprescribing list are shown in Figure 3. Any criteria scoring less than 3 on both clinical applicability and EHR computability were excluded resulting in 161 criteria. From the remaining set, the top 50% were selected, resulting in 81 high-yield criteria across all 3 deprescribing lists.

On average, STOPP criteria had the lowest clinical risk and EHR computability ratings, while the Beers criteria had the highest, contributing to their respective adoption rates of 45.9% and 62.5% among high-yield criteria. Reduced inclusion of STOPP criteria was primarily attributed to panelists' concerns that the information required was not readily accessible within the EHR and would necessitate additional data at the point of care. We also present the results of feasibility in various clinical settings by criteria list in Figure S1 in Multimedia Appendix 1, showing more likelihood for deprescribing in both outpatient and inpatient contexts, compared to the ED, across all 3 criteria lists. A scatter plot of risk versus EHR computability of high-yield criteria is presented in Figure S2 in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 3. Average distribution of results used to filter high-yield criteria on a 5-point Likert scale from a consensus study by an expert panel (n=7) split by 3 criteria lists: Beers, GEMS-Rx, and STOPP. EHR: electronic health record; GEMS-Rx: Geriatric Emergency Medication Safety Recommendations; STOPP: Screening Tool of Older People's Prescriptions.



Deprescribing Recommendations by GPT-4o

We next evaluated the LLM's deprescribing recommendations by comparing them to those of medical students, resolving any discrepancies through adjudication by board-certified EM physicians. As shown in Figure 4, 315 medications (50.3% of the total) lacked relevant high-yield deprescribing criteria. The LLM effectively identified these cases, achieving an F_1-score of 0.86 (precision=0.83, recall=0.90). Among those medications with relevant criteria, 64 (10.2% of the total) were cases where either GPT-4o or the medical student recommended deprescribing. In this second step, which involved applying relevant criteria to make a recommendation, the LLM performed less effectively, with an F_1-score of 0.58 (precision=0.47, recall=0.76).

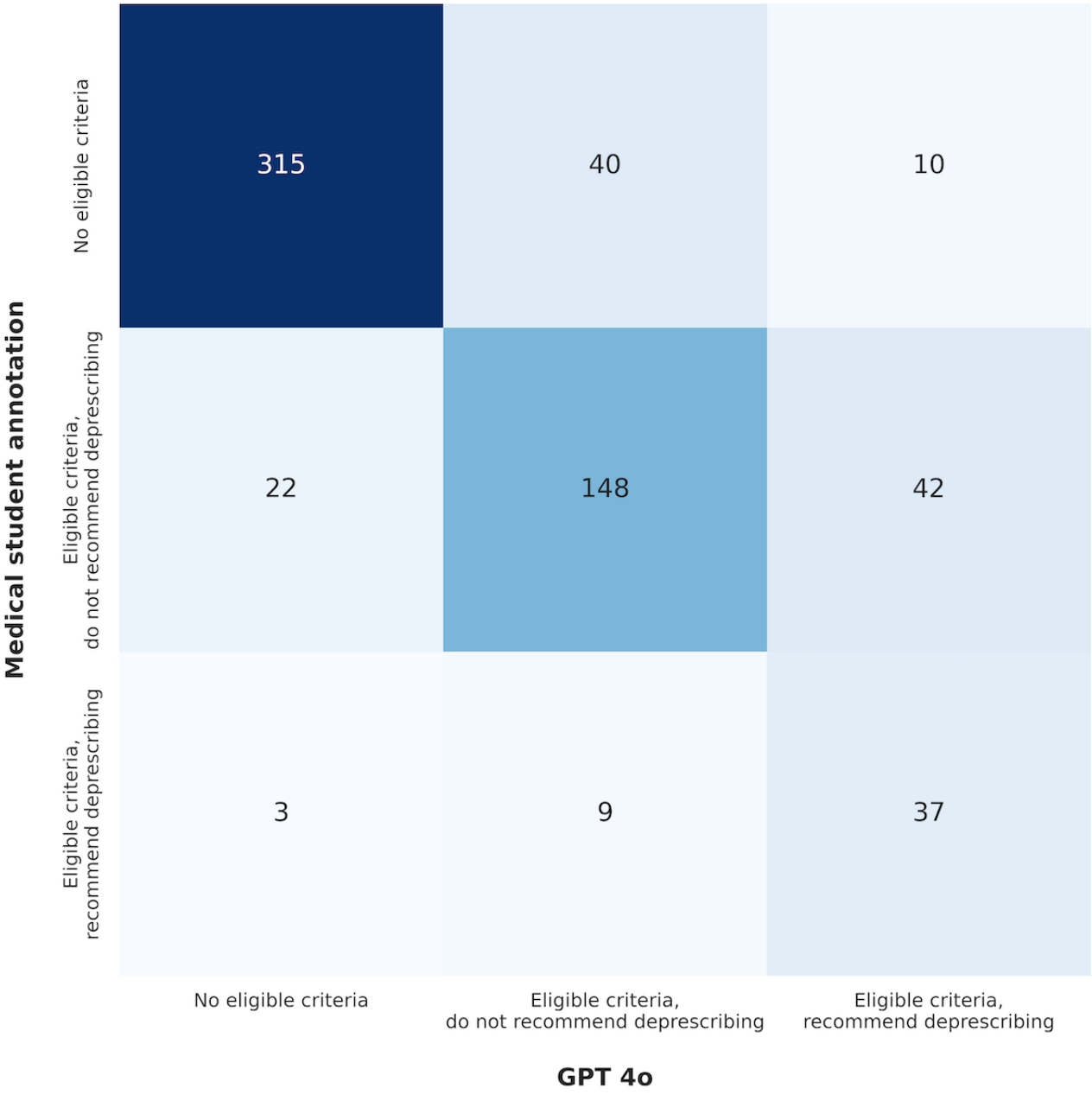
For cases where the LLM and the medical students differed, 2 senior annotators (board-certified Emergency Medicine physicians) adjudicated 126 discrepancies after standardizing the codebook and verifying IRR (Cohen *k*: eligibility=0.795, deprescribing=0.745). Notably, the confusion matrix (Figure 4) revealed that a major source of discrepancy was the significantly higher likelihood of the LLM to recommend deprescribing (11.6%) compared to the medical students (1.91%). The confusion matrix describes all possible outcomes when comparing the LLM with the medical students (eg, the medical student recommended deprescribing with eligible criteria and the LLM found no eligible criteria). The adjudication yielded similar results to those observed in comparison with medical students (Figure 5A). Across all discrepancies, GPT-4o was significantly more effective in criteria filtering compared to the medical student (McNemar test for paired proportions [30]: χ^2_1=5.985; *P*=.015). However, in applying relevant criteria, GPT-4o performed worse than the medical students, though the difference was not statistically significant (McNemar test: χ^2_1=1.818; *P*=.178). Adjudication was chosen as the gold standard for resolving discrepancies because preliminary

XSL•FO
RenderX

research showed low IRR among medical students for both steps of the process (Cohen *k*: step 1=0.68, step 2=0.33) across 75 medications.

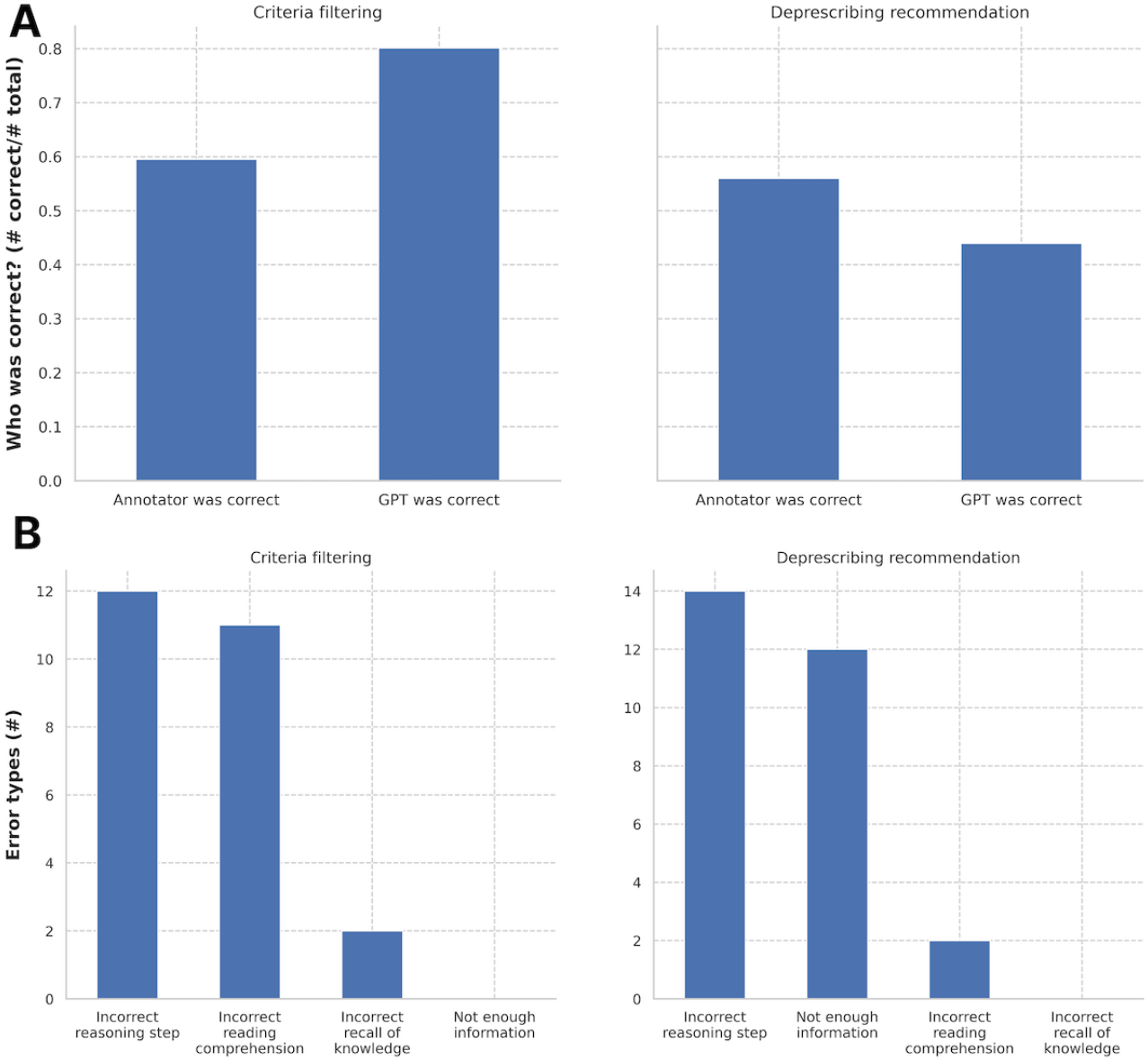
In cases where GPT-4o was incorrect, error analysis highlighted key patterns across the 2 steps. For criteria filtering, senior annotators observed that errors often stemmed from incorrect reasoning or reading comprehension issues (Figure 5B). Similarly, in making recommendations, the most common error was incorrect reasoning, followed by cases where the LLM lacked sufficient information to make an accurate determination and subsequently made inappropriate assumptions.

Figure 4. Confusion matrix (n=626 medications). The joint confusion matrix across both steps showing alignment and discrepancies between the GPT-4o model and medical students.



XSL•FO
RenderX

Figure 5. (A) Expert adjudication (n=126). Adjudication by senior clinical expert comparing junior annotators and GPT-4o in both criteria filtering and deprescribing recommendation tasks. (B) GPT-4o error modes. Types of errors by GPT-4o in the adjudication set (n=126) for both criteria filtering and deprescribing recommendation tasks.



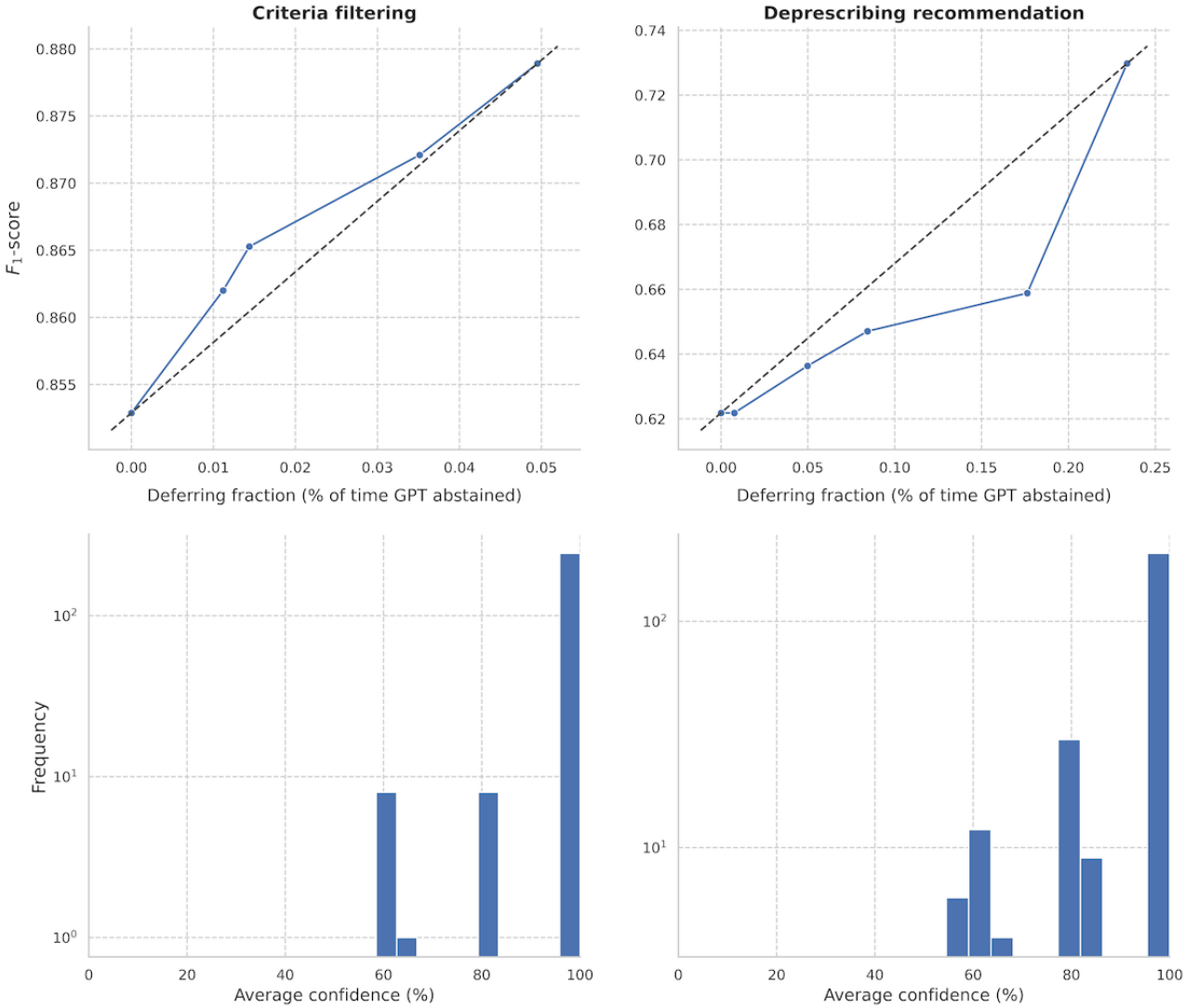
Selective Prediction Methods

Finally, we investigated the impact of incorporating confidence estimates from the LLM to guide selective prediction, allowing the model to abstain from making predictions in cases of low confidence. We compared 2 approaches: verbalized confidence and consistency-based confidence. Verbalized confidence demonstrated a narrower range of F_1-scores overall, with step 1 (eligibility) ranging from 0.860 to 0.863 and step 2 (deprescribing) from 0.58 to 0.69, as shown in Figure S3 in Multimedia Appendix 1. In contrast, consistency-based confidence exhibited broader and higher F_1-score ranges, with step 1 spanning 0.85 to 0.88 and step 2 ranging from 0.62 to 0.73 (Figure 6A). These results suggest that consistency-based confidence provides more flexibility and improved performance compared to verbalized confidence across both steps.

In particular, consistency-based selective prediction demonstrates a positive linear relationship between accuracy in deprescribing recommendations and deferring fractions. However, despite some minor improvements, we find that the LLM is poorly calibrated, as shown in Figure 6B. Despite consistency-based weighting, the confidence distribution is severely left-skewed with the minimum confidence being 58.5% in eligibility filtering and 54.5% in deprescribing recommendations.

XSL•FO
RenderX

Figure 6. Consistency-based selective prediction. (A) Range of F1-scores resulting from the application of consistency-based selective prediction in both steps of the deprescribing pipeline. The dotted line shows ideal performance as a function of deferring fraction. (B) Distribution of confidences for deprescribing recommendations from GPT-4o on a log scale.



Discussion

Principal Findings

In this retrospective cohort study evaluating deprescribing opportunities for PIMs among older adults with polypharmacy in the ED, we found that LLMs effectively identify relevant criteria from verified lists but are less adept at applying these criteria to individual patient cases. GPT-4o's performance was compared to that of medical students in a 2-step pipeline: filtering for criteria-eligible medications and making specific deprescribing recommendations. Adjudication by senior clinicians was used to resolve discrepancies, and selective prediction methods were tested to improve the model's reliability. The results offer insights into both the capabilities and limitations of LLMs in a real-world clinical context, highlighting key areas for improvement in both LLM frameworks and deprescribing guidelines.

Effectiveness of the 2-Step LLM Pipeline

The LLM demonstrated strengths in the initial filtering step, accurately identifying a high proportion of medications that matched deprescribing criteria, thus offering the potential to support clinicians in rapidly screening complex medication lists. In fact, the LLM outperformed medical students by a significant margin (80.1% vs 59.5% correct, McNemar test: P=.02). The adjudication, combined with strong overall performance (maximum F_1-score: 87.8%) using selective prediction methods, suggests that the LLM can effectively minimize the number of criteria requiring final review for deprescribing recommendations. Cases of misclassification were relatively uncommon and primarily related to nonstandard drug class names or overly broad groups, which could be improved with refined deprescribing criteria. However, in the second step—making specific deprescribing recommendations—the LLM encountered considerable difficulty, particularly when dealing with ambiguous criteria, missing information, and nuanced clinical scenarios. For example, thiazide diuretics are recommended for deprescribing in cases of significant hypokalemia, hyponatremia, hypercalcemia, or a history of gout. However, GPT-4o recommended deprescribing without access to current electrolyte values, instead basing its suggestion on a history of chronic kidney disease, a condition associated with

XSL•FO
RenderX

potential electrolyte imbalances but not meeting the relevant inclusion criteria. If implemented in clinical decision support, these inaccuracies might contribute to increased alert fatigue and extend the time required to interpret LLM-generated recommendations, potentially offsetting the intended efficiency gains in identifying deprescribing opportunities.

Our findings on the LLM's performance in identifying medications with relevant deprescribing criteria based on eligibility guidelines align with evidence from clinical trial matching literature [34], where LLMs have shown performance comparable to physicians in applying such criteria to identify eligible patients. One study used a similar "filter-and-apply" pipeline, in which trials were first filtered and then matched to patients, showcasing the effectiveness of this approach [35]. However, despite successes in eligibility filtering, challenges remain when applying complex criteria to specific patient cases. Similar errors to those observed in our work have been reported, such as incorrectly identifying patients who meet partial criteria, or assuming a patient with breast cancer does not have lung cancer simply because it is not explicitly mentioned [36,37]. Overall, while LLMs hold promise for reducing the time burden in determining deprescribing eligibility, their application requires careful consideration, particularly in tasks involving complex clinical reasoning.

Role of Selective Prediction in Clinical Decision-Making

To address the model's limitations in clinical decision-making, we implemented selective prediction methods, which allowed the LLM to "abstain" from making a recommendation in cases of low confidence. Selective prediction marginally improved the LLM's filtering accuracy by enabling it to defer uncertain cases to human reviewers. However, the effectiveness of this approach was limited by the poorly calibrated confidence levels assigned by the LLM to its decisions. Specifically, the LLM displayed a minimum confidence level of 54%, even in cases where its recommendations were incorrect. This indicates a tendency toward overconfidence, particularly in its deprescribing recommendations. While verbalized confidence is known to be overconfident in clinical question answering [38], our results contradict recent work that suggests that consistency-based methods alleviate some of these concerns [39]. This discrepancy underscores the importance of task-specific confidence thresholds and suggests that selective prediction, while useful, is not a one-size-fits-all solution in complex clinical applications.

Improved uncertainty calibration in LLMs could enhance selective prediction methods, optimizing physician-artificial intelligence (AI) workflows in clinical settings. Future applications of a well-calibrated deprescribing CDS tool could flag cases where critical information is missing (eg, antipsychotics at unchanged doses for more than 3 months without a documented medication review). This approach could streamline medication filtering while preserving human oversight, allowing clinicians to focus on complex cases where LLM reliability is limited.

Need for Clearer Deprescribing Guidelines

A notable finding from this study is the need for clearer and more consistent deprescribing guidelines. Ambiguities in criteria definitions, such as those related to medication administration routes and drug classes, present substantial barriers to automation and contribute to discrepancies between human and model interpretations. Additionally, the model often recommended deprescribing medications that, while potentially inappropriate, required specific contextual qualifiers (eg, patient's life expectancy, nutritional status, frailty status) to justify deprescribing—criteria that the LLM misapplied due to lack of explicit context or ambiguous language in the guidelines. It is important to note that current deprescribing criteria were not originally designed for direct implementation in CDS systems, but rather as general recommendations for prescribing physicians. Reorganizing these criteria into a structured, explicit framework tailored for CDS use could reduce ambiguity, improve the model's performance, and support more consistent application in clinical practice. In general, streamlining deprescribing criteria to ensure consistent application across clinical contexts could improve model reliability and help standardize deprescribing practices.

Implications for LLM Use in Clinical Practice and Future Directions

The results of this study underscore the promise of LLMs in enhancing deprescribing workflows by providing rapid filtering of PIMs, which could alleviate some of the burdens on health care providers. However, this work also highlights the limitations of current LLMs in complex, context-sensitive clinical decision-making tasks. The LLM's frequent tendency to overrecommend deprescribing, as compared to medical students, indicates that clear boundaries for medication eligibility and exclusion are critical for reducing false positives in automated recommendations. Cases with the potential for human harm were observed, such as suggesting deprescribing anticoagulation in a patient with recent thromboembolism. Additionally, cases were seen in which the LLM recommended deprescribing without citing a specific criterion. These behaviors suggest that strong guardrails on the LLM are needed to ensure safe, high-quality recommendations. Enhancing guideline specificity, particularly for complex inclusion or exclusion criteria, could reduce both human and model error rates and may foster greater acceptance of AI-assisted deprescribing tools among clinicians. Our findings highlight the potential value of human-AI collaboration frameworks. For example, a human-in-the-loop framework (a model in which humans review difficult cases the LLM cannot resolve) could involve LLMs assisting in the identification of deprescribing opportunities while deferring final recommendations to clinicians [40]. Alternatively, the LLM could focus on the initial filtering of relevant deprescribing criteria for specific medications, leaving the recommendation task entirely to the clinician, thereby leveraging the LLM's strength in mapping medications or medication classes to appropriate criteria efficiently [41]. These approaches not only leverage the model's efficiency in data processing but also mitigate risks associated with erroneous recommendations, particularly in this high-risk clinical context. Future research should focus on refining LLM architectures to better handle the nuances of clinical reasoning and context interpretation, perhaps by incorporating more advanced natural language processing techniques and domain-specific training.

XSL•FO
RenderX

Additionally, efforts to standardize deprescribing guidelines would greatly benefit the development of automated tools in this area, making them more reliable and broadly applicable.

Limitations

This study has several limitations that warrant consideration. First, the retrospective nature of our analysis, relying on historical data from EHRs, may not fully capture the complexity of real-time clinical decision-making in emergency settings. The study's focus on a single large academic medical center limits the generalizability of our findings to other settings with different patient populations, documentation patterns, and health care practices. Second, the selective prediction methods, while providing insights into the LLM's confidence, were not universally effective, particularly in the nuanced task of deprescribing recommendations. The model's performance in these recommendations highlights the challenge of translating structured criteria into actionable clinical decisions, especially when faced with ambiguous inclusion or exclusion conditions. Additionally, the model's reliance on textual prompts and structured EHR data may not fully account for nuanced clinical contexts that influence deprescribing decisions. Third, the small sample size for detailed analysis (100 patients) limits the statistical power and may not reflect broader patterns of medication use and deprescribing needs. Cost may be a barrier to larger sample sizes in the future, as the total application programming interface utilization fees were approximately US $300-$400 over these 100 patients and the cost to both evaluate and implement the system in the real world would scale linearly with the study population. Additionally, the study relied on medical students for initial annotation. While these annotations were reviewed and adjudicated by board-certified physicians, this process may introduce variability, potentially affecting the reliability of their use as a gold standard in selective prediction methods. Our process for selecting these criteria also did not explicitly include any geriatricians, though did include a range of individuals who regularly care for older adults. Finally, the criteria used (STOPP, Beers, and GEMS-Rx) were selected based on their perceived clinical risk and EHR computability, which may not encompass all relevant deprescribing scenarios. The lack of standardized guidelines for implementing deprescribing criteria in LLMs also poses a challenge to consistency and accuracy.

Conclusions

This study demonstrates the potential of LLMs to augment clinical decision support by effectively filtering deprescribing criteria for older adults with polypharmacy in ED. While the LLM showed promise in identifying medications eligible for deprescribing, it faced challenges in making nuanced deprescribing recommendations, underscoring the need for human oversight in AI-driven processes. Future research should prioritize refining the model by addressing ambiguities in deprescribing criteria and integrating broader clinical context, such as longitudinal data from prior progress notes and discharge summaries, to enable the detection of relevant clinical trends. Expanding the dataset and exploring more effective strategies for integrating human judgment with AI capabilities will help overcome limitations in generalizability, helping optimize patient care. The findings underscore the potential of LLMs in AI-enabled automated CDS tools for deprescribing while emphasizing the need to refine deprescribing criteria and establish clearer guidelines to support the integration of AI into clinical practice.

Acknowledgments

VS has received funding from the Yale University-Mayo Clinic Center of Excellence in Regulatory Science and Innovation for this work. MI receives funding through CTSA (grant UL1 TR001863) from the National Center for Advancing Translational Science, a component of the National Institutes of Health. This publication's contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

Authors' Contributions

VS, MI, and RAT conceived of the study with help from DSW in the design of its analysis. VS, CD, and LC provided data engineering and analyzed the data. VS prepared figures. DSW, JA, BP, NSK, CXW, AL, and RAT participated in the consensus study. TH, SF, DSW, and MI annotated notes for model training and testing. VS, DSW, LC, and RAT drafted the manuscript, and all authors contributed substantially to its revision. RAT takes responsibility for the paper as a whole.

Conflicts of Interest

RAT receives unrelated support from grants from the National Institutes of Health, Gordon and Betty Moore Foundation, the Federal Drug Administration, the Agency for Healthcare Research & Quality, and Beckman Coulter, Inc as well as options from Vera Health for serving as an advisor. DC is the associate editor for *JMIR Medical Education*.

Multimedia Appendix 1
Additional figures and tables.
[DOCX File , 1455 KB - aging_v8i1e69504_app1.docx]

References

XSL•FO
RenderX

1. Halli-Tierney AD, Scarbrough C, Carroll D. Polypharmacy: evaluating risks and deprescribing. Am Fam Physician 2019;100(1):32-38 [FREE Full text] [Medline: 31259501]
2. Scott IA, Hilmer SN, Reeve E, Potter K, Le Couteur D, Rigby D, et al. Reducing inappropriate polypharmacy: the process of deprescribing. JAMA Intern Med 2015;175(5):827-834. [doi: 10.1001/jamainternmed.2015.0324] [Medline: 25798731]
3. Skains RM, Koehl JL, Aldeen A, Carpenter CR, Gettel CJ, Goldberg EM, et al. Geriatric emergency medication safety recommendations (GEMS-Rx): modified Delphi development of a high-risk prescription list for older emergency department patients. Ann Emerg Med 2024;84(3):274-284. [doi: 10.1016/j.annemergmed.2024.01.033] [Medline: 38483427]
4. Reeve E, Gnjidic D, Long J, Hilmer S. A systematic review of the emerging definition of 'deprescribing' with network analysis: implications for future research and clinical practice. Br J Clin Pharmacol 2015;80(6):1254-1268 [FREE Full text] [doi: 10.1111/bcp.12732] [Medline: 27006985]
5. Hohl CM, Dankoff J, Colacone A, Afilalo M. Polypharmacy, adverse drug-related events, and potential adverse drug interactions in elderly patients presenting to an emergency department. Ann Emerg Med 2001;38(6):666-671. [doi: 10.1067/mem.2001.119456] [Medline: 11719747]
6. O'Mahony D, Cherubini A, Guiteras AR, Denkinger M, Beuscart JB, Onder G, et al. STOPP/START criteria for potentially inappropriate prescribing in older people: version 3. Eur Geriatr Med 2023;14(4):625-632 [FREE Full text] [doi: 10.1007/s41999-023-00777-y] [Medline: 37256475]
7. American Geriatrics Society Beers Criteria® Update Expert Panel. American Geriatrics Society 2023 updated AGS Beers Criteria® for potentially inappropriate medication use in older adults. J Am Geriatr Soc 2023;71(7):2052-2081. [doi: 10.1111/jgs.18372] [Medline: 37139824]
8. Candeias C, Gama J, Rodrigues M, Falcão A, Alves G. Potentially inappropriate medications and potential prescribing omissions in elderly patients receiving post-acute and long-term care: application of screening tool of older people's prescriptions/screening tool to alert to right treatment criteria. Front Pharmacol 2021;12:747523 [FREE Full text] [doi: 10.3389/fphar.2021.747523] [Medline: 34737705]
9. Kaufmann CP, Tremp R, Hersberger KE, Lampert ML. Inappropriate prescribing: a systematic overview of published assessment tools. Eur J Clin Pharmacol 2014;70(1):1-11. [doi: 10.1007/s00228-013-1575-8] [Medline: 24019054]
10. Lee S, Bobb Swanson M, Fillman A, Carnahan RM, Seaman AT, Reisinger HS. Challenges and opportunities in creating a deprescribing program in the emergency department: a qualitative study. J Am Geriatr Soc 2023;71(1):62-76 [FREE Full text] [doi: 10.1111/jgs.18047] [Medline: 36258309]
11. Scott IA, Pillans PI, Barras M, Morris C. Using EMR-enabled computerized decision support systems to reduce prescribing of potentially inappropriate medications: a narrative review. Ther Adv Drug Saf 2018;9(9):559-573. [doi: 10.1177/2042098618784809] [Medline: 30181862]
12. Anrys P, Boland B, Degryse JM, De Lepeleire J, Petrovic M, Marien S, et al. STOPP/START version 2-development of software applications: Easier said than done? Age Ageing 2016;45(5):589-592. [doi: 10.1093/ageing/afw114] [Medline: 27496930]
13. Clusmann J, Kolbinger FR, Muti HS, Carrero ZI, Eckardt JN, Laleh NG, et al. The future landscape of large language models in medicine. Commun Med 2023;3(1):141 [FREE Full text] [doi: 10.1038/s43856-023-00370-1] [Medline: 37816837]
14. Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, et al. How does ChatGPT perform on the United States Medical Licensing Examination (USMLE)? The implications of large language models for medical education and knowledge assessment. JMIR Med Educ 2023;9:e45312 [FREE Full text] [doi: 10.2196/45312] [Medline: 36753318]
15. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. PLOS Digit Health 2023;2(2):e0000198. [doi: 10.1371/journal.pdig.0000198] [Medline: 36812645]
16. Savage T, Nayak A, Gallo R, Rangan E, Chen JH. Diagnostic reasoning prompts reveal the potential for large language model interpretability in medicine. NPJ Digit Med 2024;7(1):20 [FREE Full text] [doi: 10.1038/s41746-024-01010-1] [Medline: 38267608]
17. Goel A, Gueta A, Gilon O, Liu C, Erell S, Nguyen LH, et al. LLMs accelerate annotation for medical information extraction. Proc Mach Learn Res 2023:82-100 [FREE Full text]
18. Agrawal M, Hegselmann S, Lang H, Kim Y, Sontag D. Large language models are few-shot clinical information extractors. In: Goldberg Y, Kozareva Z, Zhang Y, editors. Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing. Abu Dhabi, United Arab Emirates: Association for Computational Linguistics; 2022:1998-2002.
19. Kanjee Z, Crowe B, Rodman A. Accuracy of a generative artificial intelligence model in a complex diagnostic challenge. JAMA 2023;330(1):78-80 [FREE Full text] [doi: 10.1001/jama.2023.8288] [Medline: 37318797]
20. SAEM24 Abstracts. Academic emergency medicine. Journal Article 2024;31(S1):8-401. [doi: 10.1111/acem.14906]
21. Safranek CW, Huang T, Wright DS, Wright CX, Socrates V, Sangal RB, et al. Automated HEART score determination via ChatGPT: honing a framework for iterative prompt development. J Am Coll Emerg Physicians Open 2024;5(2):e13133 [FREE Full text] [doi: 10.1002/emp2.13133] [Medline: 38481520]
22. Liu N, Lin K, Hewitt J, Paranjape A, Bevilacqua M, Petroni F, et al. Lost in the middle: how language models use long contexts. Trans Assoc Comput Linguist 2024:157-173. [doi: 10.1162/tacl_a_00638]

XSL•FO
RenderX

23. Shi F, Chen X, Misra K, Scales N, Dohan D, Chi E, et al. Large language models can be easily distracted by irrelevant context. 2023 Presented at: ICML'23: Proceedings of the 40th International Conference on Machine Learning; July 23, 2023; Honolulu, HI.
24. Xiong M, Hu Z, Lu X, Li Y, Fu J, He J, et al. Can LLMs Express Their Uncertainty? An Empirical Evaluation of Confidence Elicitation in LLMs. In: Can LLMs express their uncertainty? An empirical evaluation of confidence elicitation in LLMs. 2024 Presented at: The Twelfth International Conference on Learning Representations; May 7, 2024; Vienna, Austria URL: https://openreview.net/forum?id=gjeQKFxFpZ
25. Leibig C, Allken V, Ayhan MS, Berens P, Wahl S. Leveraging uncertainty information from deep neural networks for disease detection. Sci Rep 2017;7(1):17816 [FREE Full text] [doi: 10.1038/s41598-017-17876-z] [Medline: 29259224]
26. Chua M, Kim D, Choi J, Lee NG, Deshpande V, Schwab J, et al. Tackling prediction uncertainty in machine learning for healthcare. Nat Biomed Eng 2023;7(6):711-718. [doi: 10.1038/s41551-022-00988-x] [Medline: 36581695]
27. Geifman Y, El-Yaniv R. Selective classification for deep neural networks. 2017 Presented at: NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems; December 4-9, 2017; Long Beach, CA p. 4885-4894.
28. El-Yaniv R, Wiener Y. On the foundations of noise-free selective classification. J Mach Learn Res 2010;11(53):1605-1641. [doi: 10.5555/1756006.1859904]
29. Liévin V, Hother CE, Motzfeldt AG, Winther O. Can large language models reason about medical questions? Patterns 2024;5(3):100943 [FREE Full text] [doi: 10.1016/j.patter.2024.100943] [Medline: 38487804]
30. Kim S, Lee W. Does McNemar's test compare the sensitivities and specificities of two diagnostic tests? Stat Methods Med Res 2017;26(1):142-154. [doi: 10.1177/0962280214541852] [Medline: 24996898]
31. Seabold S, Perktold J. statsmodels: Econometric and statistical modeling with python. SciPy Proc 2010. [doi: 10.25080/majora-92bf1922-011]
32. Waskom ML. seaborn: Statistical data visualization. J Open Source Softw 2021;6(60):3021. [doi: 10.21105/joss.03021]
33. Hunter JD. Matplotlib: A 2D graphics environment. Comput Sci Eng 2007;9(3):90-95. [doi: 10.1109/mcse.2007.55]
34. Gupta S, Basu A, Nievas M, Thomas J, Wolfrath N, Ramamurthi A, et al. PRISM: patient records interpretation for semantic clinical trial matching system using large language models. NPJ Digit Med 2024;7(1):305 [FREE Full text] [doi: 10.1038/s41746-024-01274-7] [Medline: 39468259]
35. Ferber D, Hilgers L, Wiest IC, Leßmann ME, Clusmann J, Neidlinger P, et al. End-to-end clinical trial matching with large language models. ArXiv Preprint posted online on July 18, 2024. [doi: 10.48550/arXiv.2407.13463]
36. Beattie J, Neufeld S, Yang D, Chukwuma C, Gul A, Desai N, et al. Utilizing large language models for enhanced clinical trial matching: a study on automation in patient screening. Cureus 2024;16(5):e60044 [FREE Full text] [doi: 10.7759/cureus.60044] [Medline: 38854210]
37. Nievas M, Basu A, Wang Y, Singh H. Distilling large language models for matching patients to clinical trials. J Am Med Inform Assoc 2024;31(9):1953-1963. [doi: 10.1093/jamia/ocae073] [Medline: 38641416]
38. Omar M, Agbareia R, Glicksberg BS, Nadkarni GN, Klang E. Benchmarking the confidence of large language models in clinical questions. MedRxiv Preprint posted online on September 10, 2024. [doi: 10.1101/2024.08.11.24311810]
39. Savage T, Wang J, Gallo R, Boukil A, Patel V, Safavi-Naini SAA, et al. Large language model uncertainty proxies: discrimination and calibration for medical diagnosis and treatment. J Am Med Inform Assoc 2024:e254. [doi: 10.1101/2024.06.06.24308399]
40. Mosqueira-Rey E, Hernández-Pereira E, Alonso-Ríos D, Bobes-Bascarán J, Fernández-Leal A. Human-in-the-loop machine learning: a state of the art. Artif Intell Rev 2023;56(4):3005-3054. [doi: 10.1007/s10462-022-10246-w]
41. Naikar N, Brady A, Moy G, Kwok HW. Designing human-AI systems for complex settings: ideas from distributed, joint, and self-organising perspectives of sociotechnical systems and cognitive work analysis. Ergonomics 2023;66(11):1669-1694 [FREE Full text] [doi: 10.1080/00140139.2023.2281898] [Medline: 38018437]

Abbreviations

AI: artificial intelligence
CDS: clinical decision support
ED: emergency department
EHR: electronic health record
GEMS-Rx: Geriatric Emergency Medication Safety Recommendations
IRR: interrater reliability
LLM: large language model
PIM: potentially inappropriate medication
STOPP: Screening Tool of Older People's Prescriptions

XSL•FO
RenderX

Edited by M Mardini; submitted 02.12.24; peer-reviewed by K Lee, I Scott; comments to author 02.01.25; revised version received 16.01.25; accepted 21.02.25; published 11.04.25.

Please cite as:
Socrates V, Wright DS, Huang T, Fereydooni S, Dien C, Chi L, Albano J, Patterson B, Sasidhar Kanaparthy N, Wright CX, Loza A, Chartash D, Iscoe M, Taylor RA
Identifying Deprescribing Opportunities With Large Language Models in Older Adults: Retrospective Cohort Study
JMIR Aging 2025;8:e69504
URL: https://aging.jmir.org/2025/1/e69504
doi:10.2196/69504
PMID:

©Vimig Socrates, Donald S Wright, Thomas Huang, Soraya Fereydooni, Christine Dien, Ling Chi, Jesse Albano, Brian Patterson, Naga Sasidhar Kanaparthy, Catherine X Wright, Andrew Loza, David Chartash, Mark Iscoe, Richard Andrew Taylor. Originally published in JMIR Aging (https://aging.jmir.org), 11.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Machine Learning Approach for Frailty Detection in Long-Term Care Using Accelerometer-Measured Gait and Daily Physical Activity: Model Development and Validation Study

Xiaoping Zheng[1*], PhD; Ziwei Zeng[1*], MSc; Kimberley S van Schooten[2,3], PhD; Yijian Yang[1,4], MD, PhD

[1]Department of Sports Science and Physical Education, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

[2]Neuroscience Research Australia, University of New South Wales, Sydney, Australia

[3]School of Population Health, University of New South Wales, Sydney, Australia

[4]CUHK Jockey Club Institute of Ageing, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

[*]these authors contributed equally

Corresponding Author:
Yijian Yang, MD, PhD
Department of Sports Science and Physical Education
The Chinese University of Hong Kong
G07 Kwok Sports Building. University Ave. The Chinese University of Hong Kong. N.T.
Hong Kong
China (Hong Kong)
Phone: 852 39434001
Fax: 852 26035781
Email: yyang@cuhk.edu.hk

Related Article:

This is a corrected version. See correction statement: https://aging.jmir.org/2025/1/e85173

Abstract

Background: Frailty affects over 50% of older adults in long-term care (LTC), and early detection is critical due to its potential reversibility. Wearable sensors enable continuous monitoring of gait and physical activity, and machine learning has shown promise in detecting frailty among community-dwelling older adults. However, its applicability in LTC remains underexplored. Furthermore, dynamic gait outcomes (eg, gait stability and symmetry) may offer more sensitive frailty indicators than traditional measures like gait speed, yet their potential remains largely untapped.

Objective: This study aimed to evaluate whether frailty in LTC facilities could be effectively identified using machine learning models trained on gait and daily physical activity data derived from a single accelerometer.

Methods: This study is a cross-sectional secondary analysis of baseline data from a 2-arm cluster randomized controlled trial. Of the 164 individuals initially enrolled, 51 participants (age: mean 85.0, SD 9.0 years; female: n=24, 47.1%) met the inclusion criteria of completing all assessments required for this study and were included in the final analysis. Frailty status was assessed using the fatigue, resistance, ambulation, incontinence, loss of weight, nutritional approach, and help with dressing (FRAIL-NH) scale. Participants completed a 5-meter walking task while wearing a 3D accelerometer. Following this task, the accelerometer was used to record daily physical activity over approximately 1 week. A total of 34 dynamic and spatial-temporal gait outcomes, 3 physical activity variables, and 6 demographic characteristics were extracted. Five conventional machine learning models were trained to classify frailty status using a leave-one-out cross-validation approach. Model performance was evaluated based on accuracy and the area under the receiver operating characteristic curve. To enhance model interpretability, explainable artificial intelligence techniques were used to identify the most influential predictive outcomes.

Results: The extreme gradient boosting model demonstrated the optimal performance with an accuracy of 86.3% and an area under the curve of 0.92. Explainable artificial intelligence analysis revealed that older adults with frailty exhibited more variable, complex, and asymmetric gait patterns, which were characterized by higher stride length variability, increased sample entropy, and a lower gait symmetry score.

XSL•FO
RenderX

Conclusions: Our findings suggest that dynamic gait outcomes may serve as more sensitive indicators of frailty than spatial-temporal gait outcomes (eg, gait speed) in LTC settings, offering valuable insights for enhancing frailty detection and management.

(JMIR Aging 2025;8:e77140) doi:10.2196/77140

KEYWORDS

frailty; long-term care; gait; physical activity; machine learning

Introduction

Frailty is associated with adverse health outcomes and increased dependency, severely increasing the risk of early mortality in older adults [1]. In long-term care (LTC) facilities, approximately 52.3% of older adults are affected by frailty [2], posing a significant threat to residents' well-being and sustainability in health care. Effective and efficient frailty screening methods are urgently needed to mitigate these negative outcomes and the progression of frailty among older adults in LTC facilities.

Currently, there is no internationally recognized gold standard for screening or diagnosing frailty for older adults, especially LTC residents. Over 60 different frailty assessment tools have been developed [3]. Fried's frailty phenotype (FFP) is widely used among community-dwelling older adults [4], while the fatigue, resistance, ambulation, incontinence, loss of weight, nutritional approach, and help with dressing (FRAIL-NH) scale is specifically designed for nursing home settings [5]. However, these assessments are time-consuming, require trained health care professionals, and are often conducted at a single time point, normally after a concern for declined intrinsic capacity has been raised [6]. Given that the transition from healthy aging to frailty is a dynamic and potentially reversible process [7], continuous monitoring is critical for early interventions, which may effectively prevent the frailty progression [8].

Wearable sensor technology and machine learning have emerged as promising alternatives to traditional frailty assessments [9-12]. Accelerometers can continuously monitor older adults' gait and daily physical activity, while machine learning has been successfully applied to analyze accelerometer data and detect frailty status in community-dwelling older adults [9-12]. Specifically, self-selected gait speed and moderate-to-vigorous physical activity (MVPA) time derived from accelerometer data are significantly correlated with frailty status in community-dwelling older adults [9-12]. However, it remains unclear whether these findings apply to the LTC population, as these individuals often have various physical and cognitive conditions that may affect their mobility [13]. Additionally, LTC residents tend to be more sedentary due to the constrained nature of their living environments [14]. Furthermore, beyond spatial-temporal gait outcomes, the role of dynamic gait outcomes, such as gait regularity and stability derived from accelerometer signals [13], remains underexplored in identifying older adults with frailty. These dynamic outcomes have demonstrated greater sensitivity in distinguishing between different patient groups, such as individuals with chronic low back pain [15-17], and may also be valuable for frailty prediction among older adults.

Apart from this, the emergence of explainable artificial intelligence techniques has addressed the critical challenge of interpreting complex machine learning models in health care applications. These techniques also support compliance with legal and regulatory requirements and help foster trust between clinicians and patients. Among them, Shapley additive explanations (SHAP) is a widely adopted approach grounded in cooperative game theory [16,18]. SHAP enables researchers and clinicians to understand not only which but also how outcomes such as gait variability contribute to the recognition of frailty. Incorporating SHAP enhances model transparency and strengthens clinical relevance, particularly in high-stakes settings like LTC facilities.

Therefore, this study aimed to evaluate whether gait outcomes and daily physical activity data derived from a single accelerometer could identify older adults with frailty in LTC facilities. We hypothesized that machine learning could accurately identify older adults with frailty. By incorporating SHAP, we further hypothesized that accelerometer-derived dynamic gait outcomes would be more sensitive than spatial-temporal gait outcomes in distinguishing between older adults with or without frailty in LTC facilities.

Methods

Study Design

This study is a cross-sectional, secondary analysis of baseline data from a 2-arm cluster randomized controlled trial [19]. Baseline data were collected from 20 LTC facilities in Hong Kong. The inclusion criteria were as follows: (1) aged 65 years or older; (2) able to rise from a chair and stand for at least 20 seconds; (3) confirmed by a physician as medically stable to participate in the study; and (4) able to wear an accelerometer continuously for 5 days to measure daily physical activity. Older adults were excluded if they (1) were unable to comprehend instructions, (2) could not complete the exercise program due to medical conditions, or (3) had legal blindness [19].

Ethical Considerations

The study adhered to the principles of the World Medical Association Declaration of Helsinki and Good Clinical Practice. All participants provided written informed consent, either personally or through a family member, prior to study enrollment. Ethics approval has been acquired from the Research Ethics Board at the Chinese University of Hong Kong and the Joint CUHK-NTEC Clinical Research Ethics Committee (SBRE-21-0413 and CREC-2022-459) [19].

XSL•FO
RenderX

Data Collection and Processing

Frailty was assessed using the FRAIL-NH scale, with a total score ranging from 0 to 14 [5]. Participants were categorized into 2 groups based on their total score following the cut-off values suggested in a previous study [20]: nonfrail (0-1) and frail (2-14).

During the study, a single accelerometer (ActiGraph GT9X Link IMU) was used, with a sampling frequency of 100 Hz and a range of ±16 gravitational acceleration units (g). Participants were instructed to walk at a comfortable speed over a 5-meter distance twice and the accelerometer was affixed near the L3 vertebra. From the accelerometer data, 34 gait outcomes were extracted from the data (details are provided in Multimedia Appendix 1). These outcomes represent the pace (gait speed, stride length, stride time, stride frequency, and acceleration root mean square), regularity (stride regularity and gait symmetry score), smoothness (index of harmonicity and harmonic ratio), predictability (sample entropy), and stability (maximal Lyapunov exponent and maximal Lyapunov exponent normalized per stride by time) of gait and have been described previously [15].

Regarding daily physical activity, participants were instructed to wear the accelerometer on the right side of their waist (at the anterior superior iliac spine) for 5 consecutive days, removing it only during bathing or before bed [21]. A valid day required at least 8 hours of wear time, and data from at least 3 valid days were included in the analysis using ActiLife software (version 6.13.4, ActiGraph LLC) [22]. Physical activity intensity was classified using established cut-off points for similar populations: sedentary behavior (0-50 counts/min), light physical activity (LPA; 51-759 counts/min), and MVPA (≥760 counts/min) [22]. Three outcomes, including the percentage of time spent in sedentary behavior, LPA, and MVPA, were extracted from the data.

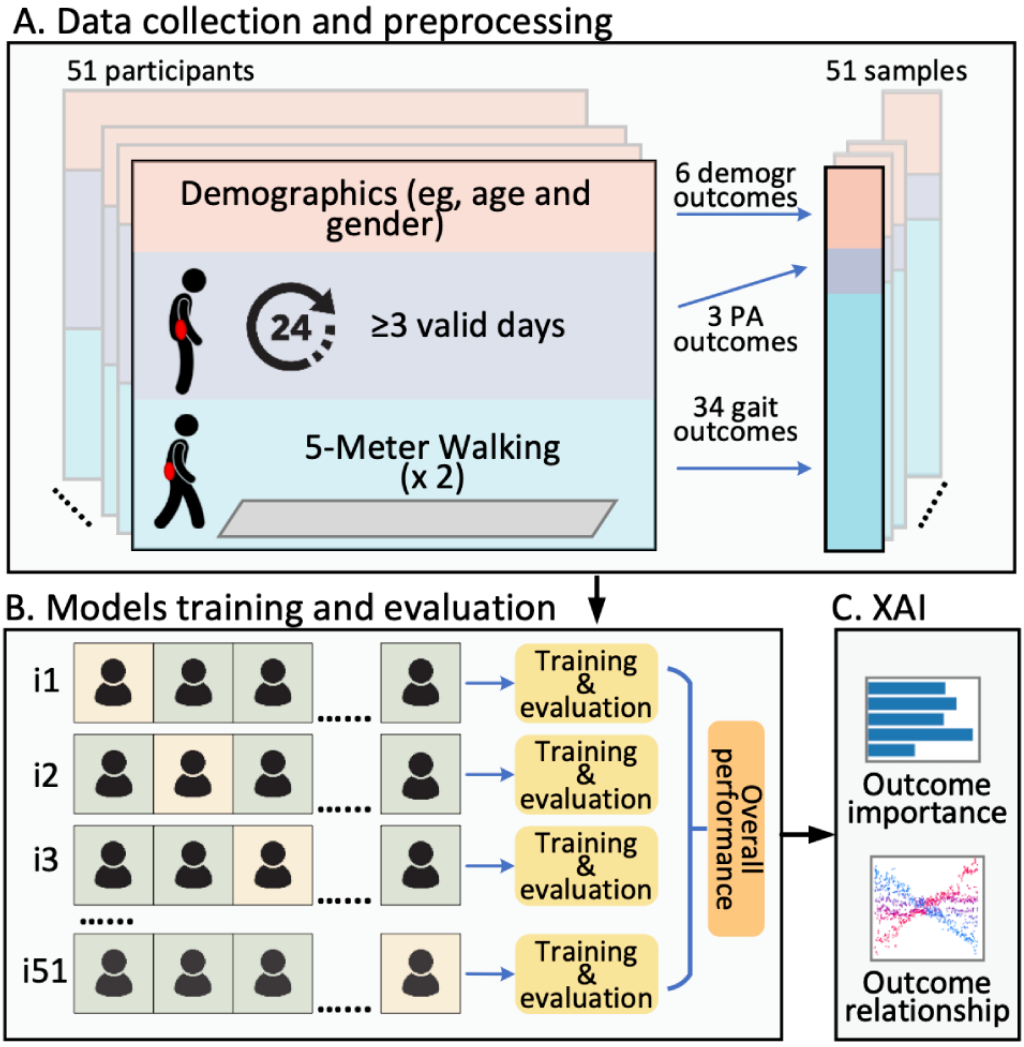
Six demographic characteristics, specifically age, gender, height, body weight, BMI, and use of mobility aids, were collected. The use of mobility aids was categorized into 3 levels: independent, use of cane, and use of walker.

Model Development and Interpretation

The gait outcomes were averaged across 2 trials, and the daily physical activity outcomes were averaged across valid days. As a result, a total of 43 outcomes (Figure 1A) were used as input to train 5 machine learning classifiers, including extreme gradient boosting (XGBoost), random forest, support vector machine, naive Bayes, and K-nearest neighbors, with labels determined by the FRAIL-NH scale. Detailed descriptions of the corresponding data preprocessing of each model are provided in Multimedia Appendix 2.

XSL•FO
RenderX

Figure 1. Data processing and analysis. (A) Data collection and preprocessing for demographic, daily physical activity (PA), and gait data. (B) Models training and evaluation based on leave-one-out cross-validation. (C) Explainable artificial intelligence (XAI) for estimating the importance of each outcome and the relationships between different outcomes.



Model performance for each classifier was evaluated using leave-one-out cross-validation (LOO-CV), where the data from 1 participant was left out for testing, and the remaining participants were used for training (Figure 1B). This process was repeated for each participant to ensure an unbiased assessment of model performance. For each iteration, a Bayesian optimization approach [23] with 5-fold cross-validation was used to identify the best hyperparameters for the classifier. This process followed a nested cross-validation design [24]. The details of the parameter space and settings of each machine learning model are provided in Multimedia Appendix 3. Model performance was evaluated by aggregating results from LOO-CV to compute the following metrics: accuracy, sensitivity, specificity, area under the receiver operating characteristic curve (AUC), and F_1-score.

To explore the association between gait outcomes, daily physical activity, and frailty, SHAP was used to interpret the results of the best classifier. SHAP estimates the contribution of each outcome to the classification process, as well as the interactions between selected outcomes using principles from game theory (Figure 1C).

In this study, SHAP values were computed for each testing participant during each iteration of the LOO-CV, using the corresponding model trained without that participant. The SHAP values from all iterations were then aggregated, leveraging the additivity property of SHAP [16,18], to generate a global interpretation of feature importance across the entire dataset. This approach ensures faithful, subject-specific explanations while maintaining the validity of the cross-validation framework.

To allow for the verification and reproduction of results, the project repository has been made publicly available [25]. It includes the source code and log files of all experiments.

Statistical Analysis

Statistical analyses were conducted using SPSS version 27.0 (IBM Corporation) with a significance level set at $P<.05$. Continuous variables were presented as mean (SD) for normally distributed data or median (IQR) for non-normally distributed data. Categorical variables were expressed as frequency and percentage. The normality of continuous variables was assessed using the Shapiro-Wilk test. To compare demographic characteristics between the nonfrail and frail groups, independent samples *t* tests were used for normally distributed

XSL•FO
RenderX

variables, while the Mann-Whitney U test was applied for non-normally distributed variables. For categorical variables, the Chi-square test (or Fisher exact test, where appropriate) was used to compare proportions between the 2 groups.

Results

Participant Characteristics

Among the 164 participants recruited from 20 LTC facilities, 113 were excluded for the following reasons: aged younger than 65 years, deceased, or withdrew from participation (n=15); unwilling to wear an accelerometer or insufficient accelerometer data for daily activity (n=82); and missing accelerometer data for gait (n=16). This left 51 participants for data analysis, with their demographic characteristics summarized in Table 1. Statistically significant differences between the nonfrail and frail groups were observed in body weight (nonfrail: mean 57.3, SD 10 kg; frail: mean 50.7, SD 9.4 kg; *P*=.02) and the use of walking aids (*P*=.008).

Table 1. Demographic characteristics of participants, mean (SD).

Demographic	Nonfrail (n=26)	Frail (n=25)	*P* value
Age (years), mean (SD)	83.2 (10.6)	87.0 (7.0)	.14
Gender, n			.48
Female	11	13	
Male	15	12	
Height (cm), mean (SD)	159.0 (9.7)	157.2 (10.9)	.52
Body weight (kg), mean (SD)	57.3 (10.0)	50.7 (9.4)	.02
BMI (kg/m^2), mean (SD)	22.6 (3.5)	20.7 (4.0)	.07
Walking aids, n			.008
Independent	16	5	
Use of cane	9	16	
Use of walker	1	4	

Model Performance

Table 2 presents all performance metrics for each model and Figure 2 shows their confusion matrices to offer a detailed view of individual model performance. XGBoost demonstrated the best classification performance compared to other machine learning models, achieving an AUC of 0.92 (Figure 3), with an accuracy of 86.3%, sensitivity of 0.85, specificity of 0.88, and an F_1-score of 0.86.

Table 2. Performance metric comparison between machine learning models.

Model	Accuracy (%)	AUC[a]	Sensitivity	Specificity	F_1-score
NB[b]	68.6	0.73	0.65	0.72	0.68
KNN[c]	62.7	0.70	0.62	0.64	0.63
SVM[d]	70.6	0.78	0.69	0.72	0.71
RF[e]	66.7	0.79	0.69	0.64	0.68
XGBoost[f]	86.3	0.92	0.85	0.88	0.86

[a]AUC: area under the receiver operating characteristic curve.

[b]NB: naive Bayes.

[c]KNN: K-nearest neighbors.

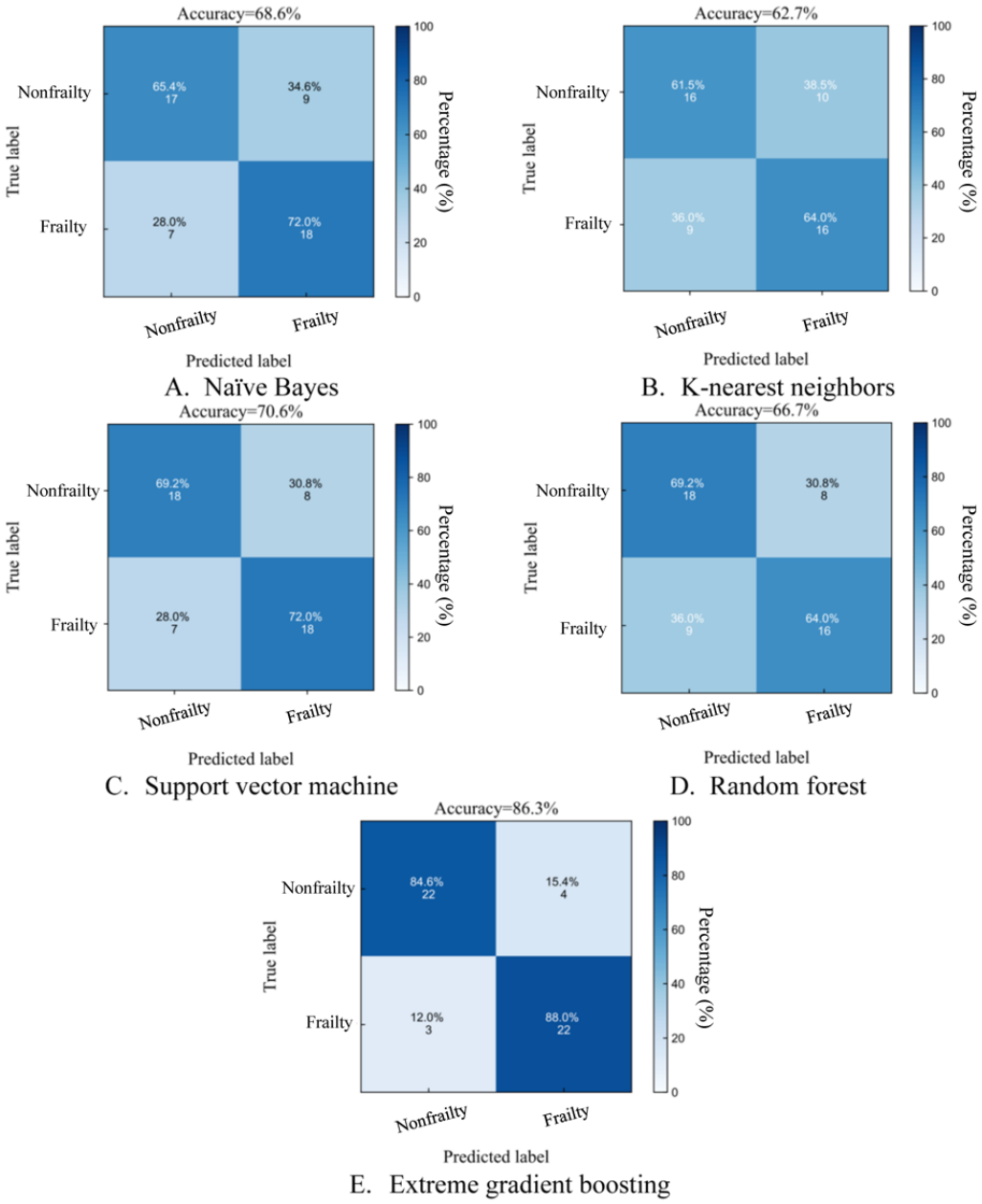
[d]SVM: support vector machine.

[e]RF: random forest.

[f]XGBoost: extreme gradient boosting.

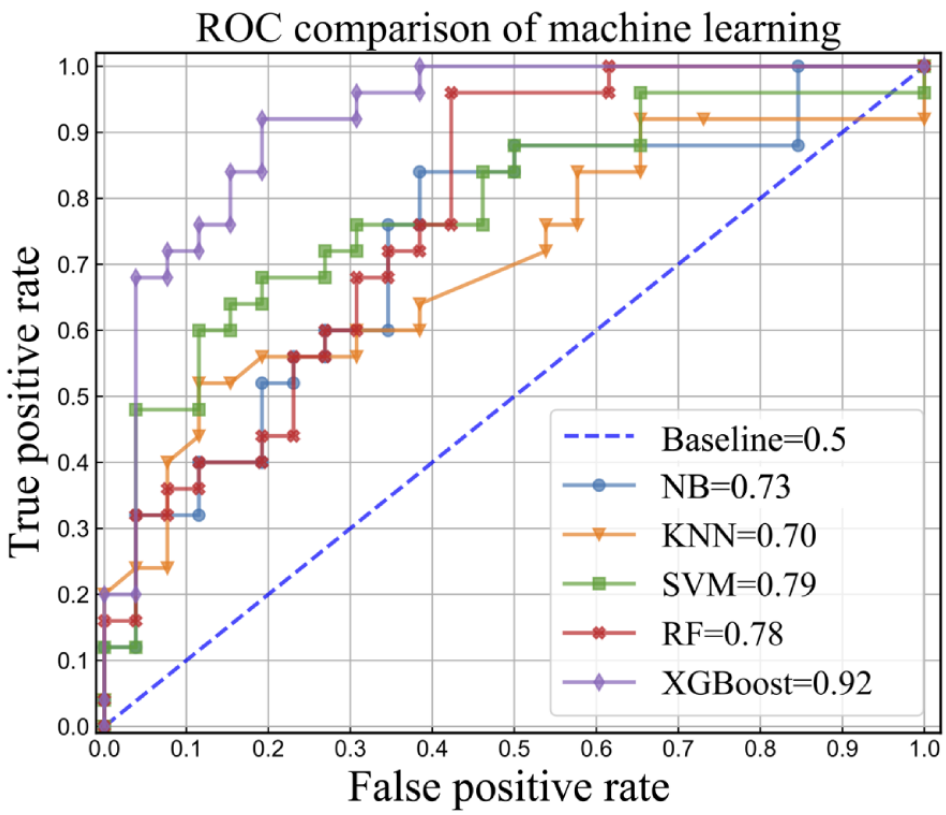
XSL•FO
RenderX

Figure 2. Confusion matrix comparisons between machine learning classifiers.



XSL•FO
RenderX

Figure 3. Comparison of the area under the curve between machine learning models. ROC: receiver operating characteristic curve; NB: naive Bayes; KNN: K-nearest neighbors; SVM: support vector machine. RF: random forest; XGBoost: extreme gradient boosting.



SHAP was used to interpret the classification process of XGBoost, as it demonstrated the best performance. The top 10 outcomes based on absolute SHAP values are presented in Multimedia Appendix 4. The elbow plot identified an "elbow" at 5, suggesting that the top 5 outcomes were the most influential in the classification. These key outcomes include stride length variability, sample entropy in the vertical direction, body weight, BMI, and gait symmetry score in the vertical direction.

Figure 4 illustrates the relationships between these outcomes and their corresponding SHAP values. Negative SHAP values indicate a contribution toward the nonfrail group, while positive values indicate a contribution toward the frail group. Larger absolute SHAP values represent a greater influence on classification. The figure shows that predicted older adults without frailty tend to have lower stride length variability, whereas participants with high sample entropy (greater gait complexity) are more likely to be classified as frail. Similarly, higher body weight and BMI values are associated with the nonfrail group. Higher gait symmetry score (indicating more symmetry; Multimedia Appendix 5) in the vertical direction contributes to the frail group, while lower values are associated with a nonfrailty classification.

Figure 4. Relationship between Shapley additive explanations (SHAP) value and key outcomes.

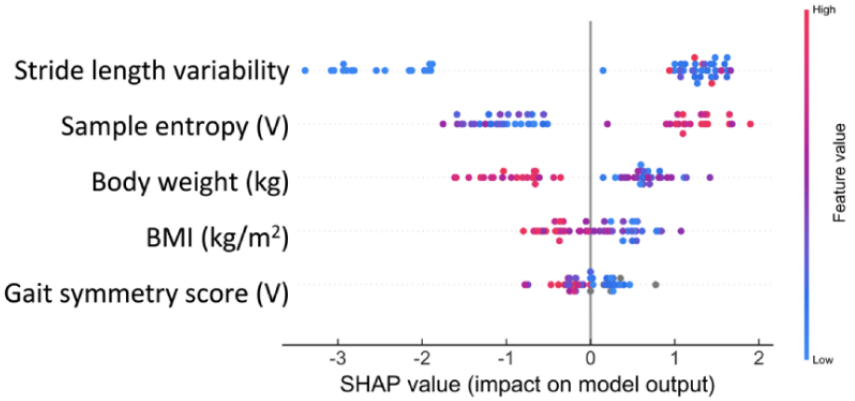
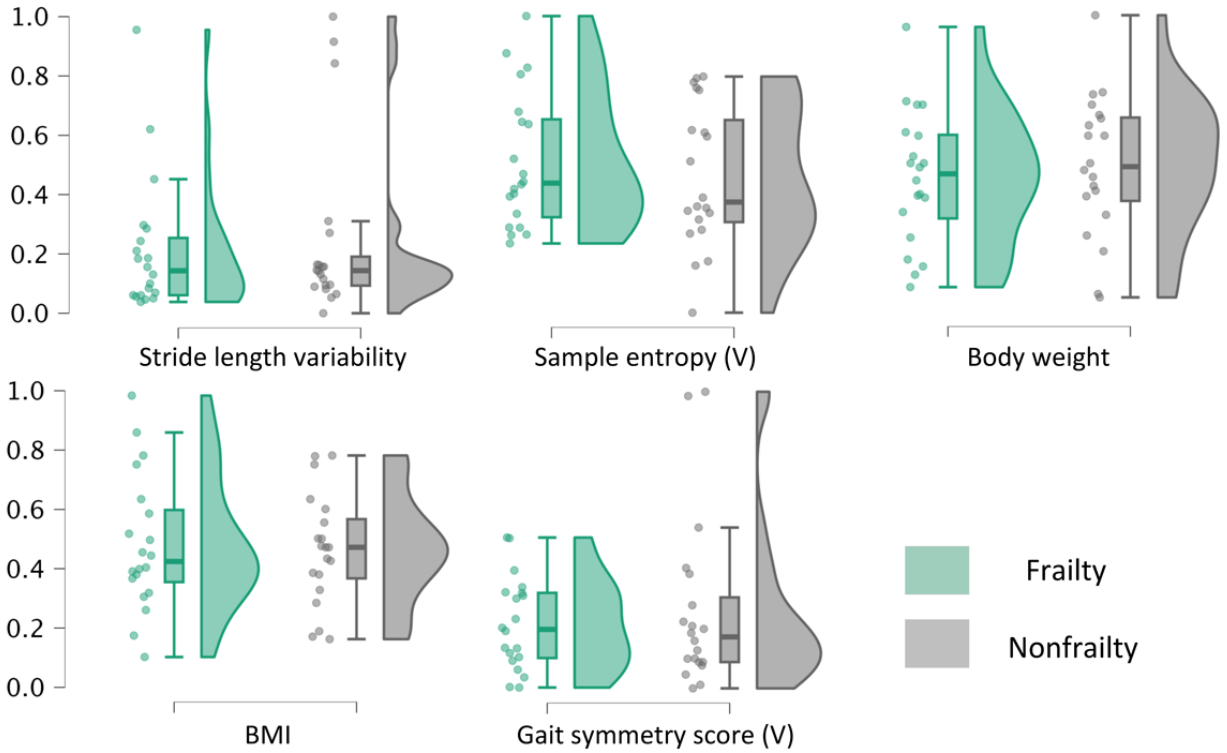


Figure 5 presents raincloud plots of the 5 key outcomes. The raincloud plot combines a kernel density plot with a box plot, with dots representing individual data points. All values were normalized to a 0-1 scale for comparability. Compared with the nonfrail group, the frail group exhibited higher stride length variability, greater sample entropy in the vertical direction, and decreased gait symmetry score in the vertical direction.

XSL•FO
RenderX

Figure 5. Raincloud plot for the top 5 key outcomes (normalized to a 0-1 scale). The dots represent individual data points; the thick line within each box plot represents the median, while the box boundaries show the interquartile range; the density plot reflects a kernel density estimate, illustrating the distribution of data across different values—wider sections indicate higher data concentration. V: vertical direction.



Discussion

Principal Findings

To the best of our knowledge, this study is among the first investigations that used machine learning in conjunction with accelerometer-derived gait and daily physical activity data to detect frailty among older adults in LTC facilities. Our results confirm that machine learning models, particularly XGBoost, demonstrated accurate classification performance in a LOO-CV (AUC=0.92, accuracy=86.3%). Consistent with our hypothesis, dynamic gait outcomes, including stride length variability, sample entropy (vertical direction), and gait symmetry (vertical direction), are more sensitive in distinguishing frailty status among LTC-dwelling older adults compared with traditional spatial-temporal parameters (eg, gait speed). These biomechanical markers, combined with anthropometric measures (body weight and BMI), emerged as the most influential predictors.

Stride length variability contributed the most to the classification process, with higher variability observed in the frail group. Previous research has reported similar findings, showing that older adults with frailty exhibit greater stride length variability in community settings [26,27]. Larger stride-to-stride fluctuations in stride length may reflect reduced accuracy and reliability of the neuromuscular systems that regulate gait rhythm [28]. Given that increased fall risk [29], age-related physical decline [30], and cognitive impairments are linked to frailty [31], it is not surprising that older adults with frailty exhibit higher gait variability [32-34]. Sample entropy was used to quantify gait predictability and complexity [35]. Higher entropy values (less predictability) in the frail group may suggest disrupted locomotor control, likely tied to neurological or neuromuscular dysfunction [36-38]. This aligns with findings that gait irregularity discriminates older adults with prefrailty from those without frailty [39] in community-dwelling settings and correlates with pathologies such as Parkinson disease and high fall risk [38]. Gait symmetry, which provides insight into the consistency between 2 consecutive steps, was lower in older adults with frailty, indicating greater degrees of asymmetry. Left-right gait coordination is primarily governed by localized neuronal networks, or central pattern generators, as observed in animal studies [40]. In healthy individuals, this coordination occurs automatically with limited cognitive input [41]. However, with aging, sarcopenia, or cognitive impairment, gait becomes less automatic and left-right coordination may increasingly depend on attentional demands, particularly in individuals with frailty [41,42]. Additionally, studies have shown no strong association between gait variability and asymmetry, suggesting that asymmetry represents an independent measure of gait impairment linked to distinct pathological causes [41]. Collectively, these findings suggest that older adults with frailty in LTC facilities exhibit more variable, irregular, and asymmetric gait patterns compared to those without frailty.

The frail group had significantly lower body weight compared with the nonfrail group, which aligns with sarcopenia's role in frailty pathogenesis, where muscle mass loss contributes to both weight reduction and functional decline [43]. Additionally, the lower BMI may reflect nutritional deficits commonly seen in older adults residing in LTC facilities, further exacerbating frailty progression [44]. These findings highlight the value of

XSL•FO
RenderX

integrating anthropometric data with dynamic gait metrics for comprehensive frailty screening in LTC settings.

Notably, dynamic gait outcomes contributed more to the classification process than gait speed. In community-dwelling older adults, self-selected gait speed has been widely recognized as a key indicator for distinguishing frailty from nonfrailty [11,12] and has been incorporated into screening tools like the FFP [45]. However, in this study, the low SHAP value of gait speed indicates its limited contribution to the classification process among older adults in LTC facilities. One possible explanation is that self-selected gait speed is influenced by multiple factors beyond walking capacity. For instance, in this study, 21.6% of older adults had experienced at least one fall in the past year. Due to fear of falling or other safety concerns, they may voluntarily reduce their gait speed despite being physically capable of walking faster [46]. Another possible explanation is the environmental constraints of LTC facilities in Hong Kong. The care homes in Hong Kong typically have narrow spaces, and gait measurements were conducted in hallways, which may have influenced gait speed. As a result, gait speed alone may not be a sufficiently sensitive measure for frailty classification in this population and settings.

Previous studies have reported that older adults with frailty tend to have lower levels of daily physical activity [47], and higher physical activity levels have been shown to protect against frailty [48]. In this study, the frail group exhibited a lower percentage of time spent in sedentary behavior (P=.001) and a higher percentage of time in LPA (P=.001), but no significant difference was observed in MVPA (Multimedia Appendix 6). However, SHAP values indicated that physical activity outcomes contributed minimally to model predictions. This suggests that in LTC settings, the current activity measures, based on time spent percentage, may not strongly differentiate frailty status (Multimedia Appendix 4). One possible reason is that the current physical activity metrics are relatively heterogeneous, covering a broad range of percentage values with considerable variability and outliers. To enhance frailty classification, more advanced measures such as postural transitions (eg, sit-to-stand [49]) and physical activity patterns [50] should be considered to develop a more robust model capable of accurately distinguishing between frail and nonfrail states among this vulnerable population.

Compared with traditional frailty assessment tools, such as the FFP and the short physical performance battery [51,52], the proposed method leverages wearable sensor-derived gait and physical activity outcomes analyzed through machine learning. This approach enables a more objective, continuous, and data-driven assessment of frailty. It also holds promise for the early detection of prefrailty and frailty in a scalable, unobtrusive, and real-time manner across diverse care settings, ultimately supporting more timely and targeted interventions.

Limitations

The walking task in this study was limited to a 5-meter distance due to space constraints in Hong Kong LTC facilities. While 2.4-, 4.6-, or 6-meter walking tests are commonly used for clinical gait analysis [53] and longer walks likely would have led to fatigue, a longer distance could have provide more reliable estimates of gait characteristics. For instance, sample entropy typically requires at least 10 seconds of signal data for reliable computation. In this study, the average walking speed was 0.46 m/s (SD 0.15 m/s), meaning that the recorded acceleration data exceeded 10 seconds, meeting the requirement for valid sample entropy calculations. Additionally, we tested the stability of sample entropy (Multimedia Appendix 7), which showed that using 70% of the current data length was sufficient to obtain stable results. Therefore, we believe that the sample entropy findings in this study are reliable. Given the limited number of participants in this study, a LOO-CV approach was used to improve the generalizability of the machine learning results and reduce the risk of overfitting. Lastly, the cross-sectional design precludes causal inferences, and the FRAIL-NH scale, though validated for LTC settings [5,20], may not fully capture frailty's dynamic progression. Longitudinal studies with repeated measures are warranted to validate these findings.

Conclusions

This study demonstrated that machine learning models, particularly XGBoost, effectively identify frailty in LTC older adults using accelerometer-derived gait and daily physical activity data. Our findings underscore the superiority of dynamic gait outcomes, including stride length variability, sample entropy (vertical direction), and gait symmetry (vertical direction), over traditional spatial-temporal outcomes in distinguishing frailty status. These outcomes reflect nuanced neuromuscular insights into frailty pathophysiology. The complementary role of anthropometric measures (body weight and BMI) further hints toward an interplay between sarcopenia, nutritional status, and functional decline in this population.

Acknowledgments

This study was supported by the Research Grants Council of Hong Kong (grant CUHK 24618722).

Conflicts of Interest

None declared.

Multimedia Appendix 1
The included gait outcomes.
[DOCX File , 16 KB - aging_v8i1e77140_app1.docx]

XSL•FO
RenderX

Multimedia Appendix 2
Details of data preprocessing for each model.
[DOCX File , 14 KB - aging_v8i1e77140_app2.docx]

Multimedia Appendix 3
Hyperparameter space.
[DOCX File , 15 KB - aging_v8i1e77140_app3.docx]

Multimedia Appendix 4
Absolute Shapley additive explanations values for the top 10 outcomes and elbow plot.
[DOCX File , 94 KB - aging_v8i1e77140_app4.docx]

Multimedia Appendix 5
Gait symmetry score calculation.
[DOCX File , 14 KB - aging_v8i1e77140_app5.docx]

Multimedia Appendix 6
Raincloud plots for the daily physical activity outcomes.
[DOCX File , 101 KB - aging_v8i1e77140_app6.docx]

Multimedia Appendix 7
The averaged sample entropy values in the vertical direction over all participants based on different length of acceleration signals.
[DOCX File , 143 KB - aging_v8i1e77140_app7.docx]

References

1. Zhang X, Dou Q, Zhang W, Wang C, Xie X, Yang Y, et al. Frailty as a predictor of all-cause mortality among older nursing home residents: a systematic review and meta-analysis. J Am Med Dir Assoc 2019 Jun;20(6):657-663.e4. [doi: 10.1016/j.jamda.2018.11.018] [Medline: 30639171]
2. Kojima G. Prevalence of frailty in nursing homes: a systematic review and meta-analysis. J Am Med Dir Assoc 2015 Nov 01;16(11):940-945. [doi: 10.1016/j.jamda.2015.06.025] [Medline: 26255709]
3. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. The Lancet 2013 Mar;381(9868):752-762. [doi: 10.1016/s0140-6736(12)62167-9]
4. Buta BJ, Walston JD, Godino JG, Park M, Kalyani RR, Xue Q, et al. Frailty assessment instruments: systematic characterization of the uses and contexts of highly-cited instruments. Ageing Res Rev 2016 Mar;26:53-61 [FREE Full text] [doi: 10.1016/j.arr.2015.12.003] [Medline: 26674984]
5. Kaehr E, Visvanathan R, Malmstrom TK, Morley JE. Frailty in nursing homes: the FRAIL-NH Scale. J Am Med Dir Assoc 2015 Feb;16(2):87-89. [doi: 10.1016/j.jamda.2014.12.002] [Medline: 25556303]
6. Cobo A, Villalba-Mora E, Pérez-Rodríguez R, Ferre X, Rodríguez-Mañas L. Unobtrusive sensors for the assessment of older adult's frailty: a scoping review. Sensors (Basel) 2021 Apr 23;21(9):2983 [FREE Full text] [doi: 10.3390/s21092983] [Medline: 33922852]
7. Wleklik M, Uchmanowicz I, Jankowska EA, Vitale C, Lisiak M, Drozd M, et al. Multidimensional approach to frailty. Front Psychol 2020;11:564 [FREE Full text] [doi: 10.3389/fpsyg.2020.00564] [Medline: 32273868]
8. Cameron ID, Fairhall N, Langron C, Lockwood K, Monaghan N, Aggar C, et al. A multifactorial interdisciplinary intervention reduces frailty in older people: randomized trial. BMC Med 2013 Mar 11;11:65 [FREE Full text] [doi: 10.1186/1741-7015-11-65] [Medline: 23497404]
9. Osuka Y, Chan LL, Brodie MA, Okubo Y, Lord SR. A wrist-worn wearable device can identify frailty in middle-aged and older adults: the UK Biobank study. J Am Med Dir Assoc 2024 Oct;25(10):105196. [doi: 10.1016/j.jamda.2024.105196] [Medline: 39128825]
10. Leghissa M, Carrera, Iglesias CA. Machine learning approaches for frailty detection, prediction and classification in elderly people: a systematic review. Int J Med Inform 2023 Oct;178:105172. [doi: 10.1016/j.ijmedinf.2023.105172] [Medline: 37586309]
11. Vavasour G, Giggins OM, Doyle J, Kelly D. How wearable sensors have been utilised to evaluate frailty in older adults: a systematic review. J Neuroeng Rehabil 2021 Jul 08;18(1):112 [FREE Full text] [doi: 10.1186/s12984-021-00909-0] [Medline: 34238323]
12. Ruiz-Ruiz L, Jimenez AR, Garcia-Villamil G, Seco F. Detecting fall risk and frailty in elders with inertial motion sensors: a survey of significant gait parameters. Sensors (Basel) 2021 Oct 19;21(20):6918 [FREE Full text] [doi: 10.3390/s21206918] [Medline: 34696131]

XSL•FO
RenderX

13. Canavan PK, Cahalin LP, Lowe S, Fitzpatrick D, Harris M, Plummer-D'Amato P. Managing gait disorders in older persons residing in nursing homes: a review of literature. J Am Med Dir Assoc 2009 May;10(4):230-237. [doi: 10.1016/j.jamda.2009.02.008] [Medline: 19426938]
14. Douma JG, Volkers KM, Engels G, Sonneveld MH, Goossens RH, Scherder EJ. Setting-related influences on physical inactivity of older adults in residential care settings: a review. BMC Geriatr 2017 Apr 28;17(1):97 [FREE Full text] [doi: 10.1186/s12877-017-0487-3] [Medline: 28454563]
15. Zheng X, Reneman MF, Echeita JA, Schiphorst Preuper RH, Kruitbosch H, Otten E, et al. Association between central sensitization and gait in chronic low back pain: insights from a machine learning approach. Comput Biol Med 2022 May;144:105329 [FREE Full text] [doi: 10.1016/j.compbiomed.2022.105329] [Medline: 35231802]
16. Zheng X, Otten E, Reneman MF, Lamoth CJ. Explaining deep learning models for age-related gait classification based on acceleration time series. Comput Biol Med 2025 Jan;184:109338. [doi: 10.1016/j.compbiomed.2024.109338] [Medline: 39536383]
17. Zhou Y, Romijnders R, Hansen C, Campen JV, Maetzler W, Hortobágyi T, et al. The detection of age groups by dynamic gait outcomes using machine learning approaches. Sci Rep 2020 Mar 10;10(1):4426 [FREE Full text] [doi: 10.1038/s41598-020-61423-2] [Medline: 32157168]
18. Shapley LS. A value for n-person games. In: Kuhn H, Tucker A, editors. Contributions to the Theory of Games II. Princeton, NJ: Princeton University Press; 1953.
19. Yang Y, Zeng Z, van Schooten KS, Sum RK, Shen J, Ho C, et al. Effects of a multicomponent physical activity programme, Mobility-Fit, compared with a standard care lower limb strengthening programme, to promote safe mobility among older adults in care facilities: protocol for a cluster randomised controlled trial. BMJ Open 2024 Jan 24;14(1):e082403 [FREE Full text] [doi: 10.1136/bmjopen-2023-082403] [Medline: 38267249]
20. Theou O, Tan EC, Bell JS, Emery T, Robson L, Morley JE, et al. Frailty levels in residential aged care facilities measured using the frailty index and FRAIL-NH scale. J Am Geriatr Soc 2016 Nov;64(11):e207-e212. [doi: 10.1111/jgs.14490] [Medline: 27783396]
21. Yatsugi H, Chen T, Chen S, Narazaki K, Nagayoshi S, Kumagai S, et al. Normative data of objectively measured physical activity and sedentary time in community-dwelling older Japanese. Int J Environ Res Public Health 2021 Mar 30;18(7):3577 [FREE Full text] [doi: 10.3390/ijerph18073577] [Medline: 33808283]
22. Zeng Z, Hsu CL, van Schooten KS, Yang Y. Sex differences in the associations of accelerometer-determined physical activity with physical and cognitive function in older adults living in long-term care. Front Public Health 2024;12:1446286 [FREE Full text] [doi: 10.3389/fpubh.2024.1446286] [Medline: 39588160]
23. Wu J, Chen XY, Zhang H, Xiong LD, Lei H, Deng SH. Hyperparameter optimization for machine learning models based on Bayesian optimization. J Electron Sci Technol 2019;17(1):26-40. [doi: 10.1007/978-1-4842-6579-6_4]
24. Lewis M, Spiliopoulou A, Goldmann K, Pitzalis C, McKeigue P, Barnes M. nestedcv: an R package for fast implementation of nested cross-validation with embedded feature selection designed for transcriptomics and high-dimensional data. Bioinform Adv 2023;3(1):vbad048 [FREE Full text] [doi: 10.1093/bioadv/vbad048] [Medline: 37113250]
25. Zheng X. Identifying frail older adults in long-term care facilities using machine learning on gait and daily physical activity data from a single accelerometer. GitHub. URL: https://github.com/xzheng93/Frailty_Identification [accessed 2025-04-30]
26. Schwenk M, Howe C, Saleh A, Mohler J, Grewal G, Armstrong D, et al. Frailty and technology: a systematic review of gait analysis in those with frailty. Gerontology 2014;60(1):79-89 [FREE Full text] [doi: 10.1159/000354211] [Medline: 23949441]
27. Pradeep Kumar D, Toosizadeh N, Mohler J, Ehsani H, Mannier C, Laksari K. Sensor-based characterization of daily walking: a new paradigm in pre-frailty/frailty assessment. BMC Geriatr 2020 May 06;20(1):164 [FREE Full text] [doi: 10.1186/s12877-020-01572-1] [Medline: 32375700]
28. Hausdorff JM. Gait variability: methods, modeling and meaning. J Neuroeng Rehabil 2005 Jul 20;2:19 [FREE Full text] [doi: 10.1186/1743-0003-2-19] [Medline: 16033650]
29. Chittrakul J, Siviroj P, Sungkarat S, Sapbamrer R. Physical frailty and fall risk in community-wwelling older adults: a cross-sectional study. J Aging Res 2020:3964973 [FREE Full text] [doi: 10.1155/2020/3964973] [Medline: 32714612]
30. Ghoraani B, Boettcher LN, Hssayeni MD, Rosenfeld A, Tolea MI, Galvin JE. Detection of mild cognitive impairment and Alzheimer's disease using dual-task gait assessments and machine learning. Biomed Signal Process Control 2021 Feb;64:102249 [FREE Full text] [doi: 10.1016/j.bspc.2020.102249] [Medline: 33123214]
31. Kelaiditi E, Cesari M, Canevelli M, van Kan GA, Ousset P, Gillette-Guyonnet S, IANA/IAGG. Cognitive frailty: rational and definition from an (I.A.N.A./I.A.G.G.) international consensus group. J Nutr Health Aging 2013 Sep;17(9):726-734 [FREE Full text] [doi: 10.1007/s12603-013-0367-2] [Medline: 24154642]
32. Montero-Odasso M, Almeida Q, Bherer L, Burhan A, Camicioli R, Doyon J, Canadian GaitCognition Network. Consensus on shared measures of mobility and cognition: from the Canadian Consortium on Neurodegeneration in Aging (CCNA). J Gerontol A Biol Sci Med Sci 2019 May 16;74(6):897-909 [FREE Full text] [doi: 10.1093/gerona/gly148] [Medline: 30101279]
33. Verghese J, Holtzer R, Lipton RB, Wang C. Quantitative gait markers and incident fall risk in older adults. J Gerontol A Biol Sci Med Sci 2009 Aug;64(8):896-901 [FREE Full text] [doi: 10.1093/gerona/glp033] [Medline: 19349593]

XSL•FO
RenderX

34. Kang HG, Dingwell JB. Separating the effects of age and walking speed on gait variability. Gait Posture 2008 May;27(4):572-577. [doi: 10.1016/j.gaitpost.2007.07.009] [Medline: 17768055]
35. Al Abiad N, van Schooten KS, Renaudin V, Delbaere K, Robert T. Association of prospective falls in older people with ubiquitous step-based fall risk parameters calculated from ambulatory inertial signals: secondary data analysis. JMIR Aging 2023 Nov 24;6:e49587 [FREE Full text] [doi: 10.2196/49587] [Medline: 38010904]
36. Smith SA, Micklewright D, Winter SL, Mauger AR. Muscle pain from an intramuscular injection of hypertonic saline increases variability in knee extensor torque reproduction. J Appl Physiol (1985) 2021 Jan 01;130(1):57-68 [FREE Full text] [doi: 10.1152/japplphysiol.00139.2020] [Medline: 33151778]
37. Borg FG, Laxåback G. Entropy of balance--some recent results. J Neuroeng Rehabil 2010 Jul 30;7:38 [FREE Full text] [doi: 10.1186/1743-0003-7-38] [Medline: 20670457]
38. Castiglia SF, Trabassi D, Conte C, Ranavolo A, Coppola G, Sebastianelli G, et al. Multiscale entropy algorithms to analyze complexity and variability of trunk accelerations time series in subjects with Parkinson's disease. Sensors (Basel) 2023 May 22;23(10):4983 [FREE Full text] [doi: 10.3390/s23104983] [Medline: 37430896]
39. Zhong R, Rau PP, Yan X. Application of smart bracelet to monitor frailty-related gait parameters of older Chinese adults: a preliminary study. Geriatr Gerontol Int 2018 Sep;18(9):1366-1371. [doi: 10.1111/ggi.13492] [Medline: 30105810]
40. Marder E, Calabrese RL. Principles of rhythmic motor pattern generation. Physiol Rev 1996 Jul;76(3):687-717. [doi: 10.1152/physrev.1996.76.3.687] [Medline: 8757786]
41. Yogev G, Plotnik M, Peretz C, Giladi N, Hausdorff JM. Gait asymmetry in patients with Parkinson's disease and elderly fallers: when does the bilateral coordination of gait require attention? Exp Brain Res 2007 Mar;177(3):336-346. [doi: 10.1007/s00221-006-0676-3] [Medline: 16972073]
42. Schwenk M, Mohler J, Wendel C, D'Huyvetter K, Fain M, Taylor-Piliae R, et al. Wearable sensor-based in-home assessment of gait, balance, and physical activity for discrimination of frailty status: baseline results of the Arizona frailty cohort study. Gerontology 2015;61(3):258-267 [FREE Full text] [doi: 10.1159/000369095] [Medline: 25547185]
43. Cruz-Jentoft A, Bahat G, Bauer J, Boirie Y, Bruyère O, Cederholm T, Writing Group for the European Working Group on Sarcopenia in Older People 2 (EWGSOP2),the Extended Group for EWGSOP2. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing 2019 Jul 01;48(4):601 [FREE Full text] [doi: 10.1093/ageing/afz046] [Medline: 31081853]
44. Norman K, Haß U, Pirlich M. Malnutrition in older adults-recent advances and remaining challenges. Nutrients 2021 Aug 12;13(8):2764 [FREE Full text] [doi: 10.3390/nu13082764] [Medline: 34444924]
45. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, Cardiovascular Health Study Collaborative Research Group. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
46. Suri A, Hubbard ZL, VanSwearingen J, Torres-Oviedo G, Brach JS, Redfern MS, et al. Fear of falling in community-dwelling older adults: What their gait acceleration pattern reveals. Comput Methods Programs Biomed 2024 Feb;244:108001. [doi: 10.1016/j.cmpb.2023.108001] [Medline: 38199138]
47. Zhao W, Hu P, Sun W, Wu W, Zhang J, Deng H, et al. Effect of physical activity on the risk of frailty: a systematic review and meta-analysis. PLoS One 2022;17(12):e0278226 [FREE Full text] [doi: 10.1371/journal.pone.0278226] [Medline: 36454790]
48. Huisingh-Scheetz M, Wroblewski K, Kocherginsky M, Huang E, Dale W, Waite L, et al. The relationship between physical activity and frailty among U.S. older adults based on hourly accelerometry data. J Gerontol A Biol Sci Med Sci 2018 Apr 17;73(5):622-629 [FREE Full text] [doi: 10.1093/gerona/glx208] [Medline: 29106478]
49. Park C, Mishra R, Golledge J, Najafi B. Digital biomarkers of physical frailty and frailty phenotypes using sensor-based physical activity and machine learning. Sensors (Basel) 2021 Aug 05;21(16):5289 [FREE Full text] [doi: 10.3390/s21165289] [Medline: 34450734]
50. Zheng X, Reneman MF, Preuper RHS, Otten E, Lamoth CJ. Relationship between physical activity and central sensitization in chronic low back pain: insights from machine learning. Comput Methods Programs Biomed 2023 Apr;232:107432 [FREE Full text] [doi: 10.1016/j.cmpb.2023.107432] [Medline: 36868164]
51. da Câmara SMA, Alvarado BE, Guralnik JM, Guerra RO, Maciel Á. Using the short physical performance battery to screen for frailty in young-old adults with distinct socioeconomic conditions. Geriatr Gerontol Int 2013 Apr;13(2):421-428. [doi: 10.1111/j.1447-0594.2012.00920.x] [Medline: 22882512]
52. Ramírez-Vélez R, López Sáez de Asteasu M, Morley J, Cano-Gutierrez C, Izquierdo M. Performance of the short physical performance battery in identifying the frailty phenotype and predicting geriatric syndromes in community-dwelling elderly. J Nutr Health Aging 2021;25(2):209-217 [FREE Full text] [doi: 10.1007/s12603-020-1484-3] [Medline: 33491036]
53. Bortone I, Sardone R, Lampignano L, Castellana F, Zupo R, Lozupone M, et al. How gait influences frailty models and health-related outcomes in clinical-based and population-based studies: a systematic review. J Cachexia Sarcopenia Muscle 2021 Apr;12(2):274-297 [FREE Full text] [doi: 10.1002/jcsm.12667] [Medline: 33590975]

XSL•FO
RenderX

Abbreviations

AUC: area under the receiver operating characteristic curve
FRAIL-NH: fatigue, resistance, ambulation, incontinence, loss of weight, nutritional approach, and help with dressing scale
FFP: Fried's frailty phenotype
LOO-CV: leave-one-out cross-validation
LPA: light physical activity
LTC: long-term care
MVPA: moderate-to-vigorous physical activity
SHAP: Shapley additive explanations
XGBoost: extreme gradient boosting

Edited by M Gray; submitted 08.05.25; peer-reviewed by ME Hernandez, S Garcia-de-Villa; comments to author 24.06.25; revised version received 17.07.25; accepted 08.08.25; published 15.09.25.

Please cite as:
Zheng X, Zeng Z, S van Schooten K, Yang Y
Machine Learning Approach for Frailty Detection in Long-Term Care Using Accelerometer-Measured Gait and Daily Physical Activity: Model Development and Validation Study
JMIR Aging 2025;8:e77140
URL: https://aging.jmir.org/2025/1/e77140
doi:10.2196/77140
PMID:

©Xiaoping Zheng, Ziwei Zeng, Kimberley S van Schooten, Yijian Yang. Originally published in JMIR Aging (https://aging.jmir.org), 15.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Representations of Older Adults' Digital Literacy in Canadian News Media: Critical Discourse Analysis Using Unified Theory of Acceptance and Use of Technology 2

Alixe Ménard[1,2], BHSc, MSc; Alexandra Maharaj[3], BHSc; Sophie Harb[1], BHSc; Sarah Fraser[1,2], PhD; Tracey O'Sullivan[1,2], PhD

[1]Interdisciplinary School of Health Sciences, Faculty of Health Sciences, University of Ottawa, 200 Lees Ave, Ottawa, ON, Canada

[2]LIFE Research Institute, University of Ottawa, Ottawa, ON, Canada

[3]Telfer School of Management, University of Ottawa, Ottawa, ON, Canada

Corresponding Author:
Alixe Ménard, BHSc, MSc
Interdisciplinary School of Health Sciences, Faculty of Health Sciences, University of Ottawa, 200 Lees Ave, Ottawa, ON, Canada

Abstract

Background: The transition of social and health services to digital platforms has significant implications for the inclusion and well-being of older adults. Digital literacy is a key determinant of health and equity, particularly as services shift to digital platforms. For older adults, this is crucial for accessing essential services and information, aiding in emergency resource navigation and information access.

Objective: This study examines how Canadian news media portray older adults' digital literacy, using the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) framework to guide the analysis.

Methods: News media articles from Canada, published from May 4, 2023, to July 4, 2024, were analyzed. Articles were identified through database searches and manual reviews, with inclusion criteria focusing on mentions of older adults and digital literacy. Critical discourse analysis was conducted using UTAUT2 factors to interpret the media narratives.

Results: Of the 103 articles reviewed, 54 met the inclusion criteria. Four key themes emerged: (1) Performance and effort expectancy: media frequently depicted older adults as needing significant help with digital technologies, indicating potential digital ageism. (2) Social influence: fearmongering narratives suggest insufficient technological skills could result in social exclusion and limited access to essential services. (3) Facilitating conditions: articles highlighted older adults' susceptibility to digital scams and fraud, reinforcing digital ageism. (4) Hedonic motivation: media portrayals emphasized active aging, illustrating how digital tools, such as home security systems, can enhance independence and quality of life.

Conclusions: Many media narratives on older adults' digital literacy perpetuate digital ageism. The study highlights how negative portrayals of older adults' digital skills and their exclusion from digital spaces underscore the need for more inclusive media representations. Findings suggest that media outlets could play a crucial role in shifting to more supportive portrayals of older adults' engagement with technology. Improving media portrayals can enhance digital engagement and preparedness among aging populations, contributing to better overall quality of life.

(JMIR Aging 2025;8:e69373) doi:10.2196/69373

KEYWORDS

digital divide; media representation; older adults; disaster planning; digital technology

Introduction

Background

Digital literacy refers to the diverse skill levels individuals possess in using digital technologies and navigating their associated risks with discernment and confidence [1]. As a determinant of health, the ability to navigate digital environments shapes access to health resources and broader health equity outcomes [2-4]. Digital literacy involves the combination of cognitive abilities and technical skills, allowing individuals to effectively use digital technology for information retrieval, critical evaluation, content creation, and communication [4]. The rapid shift of social and health services to digital platforms has profound implications for the inclusion and well-being of older adults [5].

The intersection of an aging population with rapid technological advancements has fostered the misconception that older adults are technologically incompetent [6]. This belief exemplifies digital ageism, which encompasses age-related biases in how digital technology is developed, applied, and perceived by

XSL•FO
RenderX

society [7]. Digital ageism refers to the systemic exclusion or marginalization of older adults in digital environments and is embedded in both the design and dissemination of technology [8]. It has been categorized as occurring at both individual and organizational levels. At the individual level, digital ageism involves the failure to consider the specific needs, preferences, and abilities of older adults in the design and use of technology [9]. At the organizational level, it is perpetuated by a lack of training and awareness among technology developers, including website designers and service providers, leading to digital platforms that are not accessible or inclusive of older adult users [9]. These issues are further compounded by how emerging technologies, such as artificial intelligence, are developed and deployed often without adequate representation of older adults in training data or user testing processes [9]. Digital ageism fosters stereotypes that assume older adults lack digital skills [8]. In actuality, older adults possess diverse digital competencies and engage with information and communication technology in various ways [7].

Digital inclusion is essential not only for everyday activities, such as staying socially connected, but also for navigating high-stakes situations. The consequences of digital exclusion become particularly pronounced during emergencies, when timely access to digital information and services is critical. As digital ageism limits older adults' engagement with technology, it can reduce their ability to respond effectively in times of crisis. This was especially evident during the COVID-19 pandemic, which saw a rise in the use of telehealth [10], digital banking channels [11], and web-based grocery services [12] to mitigate contact exposure [13]. More broadly, digital platforms offer invaluable resources for accessing emergency services, support networks, and relief supplies during disasters [14]. The swift dissemination of accurate and timely information, often facilitated through digital channels, is thus necessary for informed decision-making and prompt action during disasters [15].

Digital literacy equips individuals with the skills to navigate digital platforms efficiently, enabling them to access emergency alerts, evacuation routes, and safety guidelines without delay [14]. Consequently, it is essential to promote digital literacy among older adults to enhance their ability to navigate digital resources effectively, stay informed about disaster risks, and access timely assistance during a crisis [16]. Empowering older adults with digital literacy skills can foster greater resilience and preparedness within aging populations, ultimately contributing to more effective disaster response and recovery efforts [17]. Of note, digital literacy includes a broad set of competencies such as the technical skills required to use digital devices, the ability to access, evaluate and manage information, critically assess media content, and navigate digital systems such as web browsers [18]. Given this broad definition, it is reductive to assume that low digital literacy is inherent to older adults solely because they did not come of age during the digital era [18]. This stereotype promotes the falsehood that older adults are less capable of digital engagement and that technology is a privilege reserved for younger users [18]. Such assumptions diminish older adults' self-efficacy and interest in being digitally active [19]. Therefore, it is essential to critically assess the language used when discussing older adults in digital contexts to avoid reinforcing the digital divide as a socially constructed and preventable inequity [19].

Theoretical Framework

The Unified Theory of Acceptance and Use of Technology (UTAUT) model has previously been used to explain behavior and intentions related to the adoption of digital technology [20,21]. Technology-related behaviors are evaluated based on 5 factors that influence acceptance and use: (1) performance expectancy; (2) effort expectancy; (3) social influence; (4) facilitating conditions; and (5) age, gender, and experience [22]. In 2012, the original UTAUT model was expanded into UTAUT2, incorporating additional factors such as hedonic motivation, price value, and habit [23].

Performance expectancy refers to the perceived usefulness of digital technology [21]. Effort expectancy assesses the ease of use of the technology. Social influence evaluates the extent to which an individual's social network values the use of digital technology. Facilitating conditions encompass the resources and support available to an individual for using the technology [21]. UTAUT suggests that behavioral intention is influenced by factors such as age, gender, and experience [21]. In UTAUT2, hedonic motivation relates to the enjoyment experienced when using digital technology [23]. Price value refers to the cognitive assessment of the balance between the perceived benefits and the monetary costs. Finally, habit describes the extent to which digital technology use is automatic (based on previous usage) [23].

In this study, we leverage UTAUT2 to evaluate news media representations of older adults' digital literacy within the context of disaster risk management. Specifically, we investigate how newspapers depict older adults as agents with varying degrees of digital proficiency using UTAUT2 factors. Our guiding question was: What prevailing narratives surround digital literacy and older adults, and how are they depicted in news media narratives related to disaster risk management? While this focus shaped our search strategy and inclusion criteria described below, we ultimately found few disaster media articles that addressed older adults' digital inclusion or exclusion. This absence became a significant finding in itself, prompting a broader examination of how digital literacy and older adults are framed in mainstream media more generally.

Methods

Positionality

As researchers, we acknowledge that our backgrounds and experiences shape our perspectives and approaches to this study. Our research team comprises 5 women, 2 of whom are aged 50 years and older, all of whom are digitally literate. These factors influence our understanding of digital literacy and our interpretation of media representations of older adults. Our collective experience as women, including those who have navigated digital landscapes for decades, informs our sensitivity to issues of age and gender bias in media portrayals. Our digital literacy enables us to critically assess the technical aspects of digital engagement, while our personal and professional

XSL•FO
RenderX

experiences with media analysis and expertise in ageism research provide a nuanced lens through which we view the data. We have been conscientious in our approach to the data and remained reflexive in our interpretations by holding weekly team meetings where we discussed our relational awareness to ensure that "while the data are interpreted through the eyes and cultural standards of the researcher, the effects of the latter are monitored and mitigated" [24].

Data Collection

We investigated digital Canadian news media articles related to older adults, digital literacy, and disaster risk management, shared from May 4, 2023, to July 4, 2024. These dates were selected to reflect the World Health Organization's downgrade of the COVID-19 pandemic [25], allowing us to evaluate postpandemic discourse [25]. The news databases ProQuest, Factiva, and Eureka were systematically searched with the help of a Research Librarian (TR) to uncover relevant articles. The keywords related to digital literacy and disasters were informed by government documents, including the Government of Canada, the United Nations Educational, Scientific, and Cultural Organization, and the Intergovernmental Panel on Climate Change. The search strategies for each database are presented in Multimedia Appendix 1. Furthermore, the Canadian Broadcasting Corporation, the Canadian Television Network, and Radio-Canada were manually searched using the terms "seniors," "elderly," "older," or "aged" in English, and "aîné," "aînée," "aînés," or "aînées" in French.

Included articles met the following criteria: (1) mention of older adults; (2) mention of digital literacy; and (3) written in English or French. At the outset, we established "mention of disasters" as an inclusion criterion. However, due to the scarcity of articles meeting this criterion during the initial data collection phase, we made the decision to remove it. This adjustment is reflected as an important finding in the results section. All articles were imported into Zotero (Corporation for Digital Scholarship) and initially checked by the primary author (A Ménard). The senior author then reviewed 50% of the articles (SF), while the remaining 50% were divided among the other research team members (A Maharaj, SH, and TO), each reviewing one-third of this portion for a second check.

Data Analysis

UTAUT2 was used to organize the data [23], and critical discourse analysis (CDA) was subsequently used for a deeper analysis of the data [26]. The UTAUT2 factors served as a priori constructs through which we organized the data deductively to represent the underlying discursive concept of digital literacy through technology acceptance in older adults [27]. In attempting to organize data according to UTAUT2 factors, we discovered that price value and habit were not discussed in media narratives, which led to their removal from our codebook. We also combined performance expectancy and effort expectancy because of their overlap, whereby articles described behavioral intention as being informed by ease of use, directly impacting perceived usefulness of technology. CDA was used to examine and interpret the construction and perpetuation of discourses [28] concerning digital literacy among older adults, particularly in the context of disaster risk management depicted in Canadian news media. CDA is based on the premise that the use of certain language is intentional, regardless of whether individuals are aware of their discursive choices [29]. Throughout our analysis, using CDA, news media language was understood as constructing social reality, not reflecting it [30]. Through this lens, we sought to understand theories of reality perpetuated in news media and relations of power encoded in the data [31] as they pertain to digital literacy among older adults. Further, CDA allowed us to examine how discourse reflects and perpetuates power structures and inequalities [29].

All authors first familiarized themselves with the data. Following this, the first author (A Ménard) conducted an inductive analysis of the data using open coding [32]. Subsequently, the coauthors (A Maharaj, SH, SF, and TO) independently coded a portion of the included articles. The research team met weekly to ensure consistency in coding, identify emerging patterns, and reach consensus on article inclusion. All codes were then grouped to produce a preliminary codebook, which was refined by the first author (A Ménard). The data were further analyzed by aligning the codes closely with the research question and organizing them deductively according to UTAUT2. Initial coding by all authors was conducted using Microsoft Excel, while the second round of coding by the first author was conducted using NVivo 12 (QSR International).

Ethical Considerations

This study involved a scan of publicly available digital newspaper articles and did not involve human participants, personal data, or an intervention. As such, it did not require ethics approval in accordance with the Tri-Council Policy Statement (TCPS 2) and the policies of the University of Ottawa.

Results

Overview

We identified a total of 103 articles after the removal of duplicate studies. Despite a thorough search, no articles emerged that addressed digital literacy in the context of disaster risk management for older adults. This notable absence of relevant disaster-related narratives necessitated a shift in focus, leading to an analysis centered solely on the portrayal of digital literacy among older adults as depicted in the available news media. Following this pivot, a total of 54 articles (Multimedia Appendix 2) [33-86] met the inclusion criteria and were retained. The included articles underwent re-evaluation by 3 members of the research team (A Ménard, SF, and TO) to further confirm their eligibility for inclusion. Of the included articles, 17 were from Ontario [33-49], 16 from Quebec [50-65], 9 from British Columbia [66-74], 3 from Nova Scotia [75-77], 1 from Alberta [78], 1 from New Brunswick [79], 2 from Manitoba [80,81], and 5 were Canada-wide [82-86]. A total of 30 articles [33-40,42-45,47,48,67-78,80,81] were written in English, while the remaining 24 [41,46,50-65,79,82-86] were written in French.

Here, we present the themes emerging from media discourses, according to a combination of 5 UTAUT2 factors: (1) performance and effort expectancy (combined) (ie, how the media portrays the benefits and ease of use of digital technologies for older adults); (2) social influence (ie, the

XSL•FO
RenderX

media's depiction of the impact of social circles, such as family, friends, and societal norms, on older adults' decisions to adopt and use digital technologies); (3) facilitating conditions (ie, how the media presents the availability of resources and support systems that enable older adults to use digital technologies), and (4) hedonic motivation (ie, how the media portrays the enjoyable aspects of digital technology use and how age-related characteristics impact older adults' motivation to engage with digital tools).

Theme 1: Performance and Effort Expectancy

The media discourse on digital technology and older adults revealed how narratives construct perceptions of performance and effort expectancy related to technology use within this population. The media depicted older adults as a homogeneous group with limited digital literacy, framing them with low performance expectancy and high effort expectancy. This portrayal reinforced a narrative that older adults require substantial assistance from younger generations to learn and navigate digital technologies. The media also emphasized the importance of older adults gaining digital literacy to ensure continuous or lifelong learning, shaping discourse around skill-building and access to digital technology. These narratives framed digital literacy as critical to the inclusion of older adults in society, reflecting a broader societal discourse that highlights both the anticipated benefits (performance expectancy) and perceived effort (effort expectancy) required for older adults to engage with digital technologies. Table 1 illustrates how these themes were represented in media narratives, reflecting the critical discourse on performance and effort expectancy.

XSL•FO
RenderX

Table . Performance and effort expectancy.

UTAUT2[a] factors	Contents	Items
Performance expectancy (PE)	• PE1: "[T]echnology is constantly oriented to younger and younger children, and less user-friendly for elders. Many seniors do not have computers and cell phones, but to navigate the challenges of our "modern" society, one can no longer rely on businesses or government offices having a human being at the end of a phone line, ready and available for help. (…) Seniors wait for hours on a phone call promising "customer help." They show up at 6 AM to get lab tests (with the lab opening at 7 AM), perhaps because they are not comfortable with being, or able to be, in a "text line." (Comox Valley Record, British Columbia) • PE2: "Other types of customers are also affected [by bank closures], those who are less comfortable with technology, such as seniors, who generally prefer face-to-face contact." (Le Droit, Ontario) • PE3: "I'm a senior and do so resent all the generalities that we 65-year-olds are not very adept with the technologies. I frequently adjust my digital watch, use the cellular mobile telephone to send voice calls, drive a fuel-injected horseless carriage, program the remote on my TV, change the taped message on my answering machine, and, hard to believe, back up the family files onto 51/4" floppy disks once per month." (Canadian Broadcasting Corporation, Ontario)	Limited performance and lifelong learning
Effort expectancy (EE)	• EE1: "Cyber Seniors Technology Tutoring is a free, one-on-one tutoring program aimed at seniors who have little or no experience with technology. The program partners each participant with a community volunteer tutor. The two work together on developing basic computer, tablet or cellphone skills targeted to the learner's specific device and goals." (Eagle Valley News, British Columbia) • EE2: "She helped me make notes, and then had me practice, so that she could see that the message got through to me." (The Brandon Sun, Manitoba) • EE3: "[T]hey have limited knowledge of or access to digital banking services." (The Brandon Sun, Manitoba)	Personalized instruction and tailored learning goals

[a]UTAUT2: Unified Theory of Acceptance and Use of Technology 2.

Theme 2: Social Influence

Media narratives surrounding digital literacy for older adults frequently used fearmongering tactics, suggesting a lack of technological skills could lead to social exclusion, difficulties in maintaining connections with grandchildren, and restricted access to essential services like telehealth. These portrayals framed digital literacy as crucial for bridging generational gaps and reinforcing intergenerational relationships, while also emphasizing its growing importance for accessing digital banking and web-based grocery services. Many articles depicted older adults as "vulnerable," portraying them as being left behind by technological advancements. Consequently, these narratives advocated for community centers and libraries to develop programs aimed at enhancing digital literacy, helping older adults remain socially active and integrated members of society. Textbox 1 illustrates the social influence affecting older adults' digital literacy, as depicted in these media narratives, highlighting factors such as societal expectations and community support.

XSL•FO
RenderX

Textbox 1. Social influence (Unified Theory of Acceptance and Use of Technology 2 factor).

Social influence (SI) captures how individuals' behavior is shaped by the perceptions, support, and encouragement of others in their community.

- SI1. "Community centers or libraries are often places seniors would need to visit in order to get their hands on technology. However, teaching a new skill to others costs nothing, according to Garcia. 'If you don't teach people, they will never know how to do it, so I think it is very beneficial for the seniors community to have people guide them on ways to keep them connected." (Richmond News, British Columbia)
- SI2. "It's all about keeping all levels of our community connected." (Saltwire, Nova Scotia)
- SI3. "Telehealth is becoming more accessible, many of the elderly, my grandparents included, lack computer skills to book online appointments." (Vancouver Province, British Columbia)

Themes: community support, staying connected, and intergenerational contact.

Theme 3: Facilitating Conditions

Media portrayals constructed older adults as needing protection from fraud or scams within the digital sphere, reflecting broader discourses of vulnerability and dependence. These portrayals underscore the necessity for digital safety measures, increased awareness of scams, and fraud prevention training specifically targeted toward this demographic. In addition, media narratives depicted older adults as "vulnerable" to loneliness and social isolation due to perceived deficiencies in digital literacy, framing social participation and connectivity as key solutions. The discourse highlighted the importance of facilitating conditions, such as access to technology, technical support, and training programs through local libraries, as crucial resources enabling older adults to engage with digital technologies. By focusing on these supportive measures, media narratives constructed a need for comprehensive systems to enhance digital literacy, reinforcing a view of older adults as needing assistance to navigate the digital landscape effectively. However, this framing of older adults as targets of digital scams may simultaneously discourage technology use and perpetuate negative stereotypes about their technological competencies compared with younger adults. Textbox 2 illustrates the facilitating conditions affecting older adults' digital literacy as depicted in these media narratives.

Textbox 2. Facilitating conditions (Unified Theory of Acceptance and Use of Technology 2 factor).

Facilitating conditions (FC) refer to the resources and support available to help older adults use digital technologies. These include training programs, access to devices, and institutional support to ensure safety and reduce isolation.

- FC1: "The purpose of the workshop was to equip seniors with the knowledge and tools to protect themselves from common scams in an increasingly digital world." (Wiarton Echo, Ontario)
- FC2: "Vancouver police are handing free mobile phones to low-income seniors to ensure they can call 911 in an emergency." (Vancouver Sun, British Columbia)
- FC3: "Learning computer skills helps seniors overcome feelings of isolation and loneliness." (The Chronicle Herald, Nova Scotia)

Themes: safety, scams, and loneliness.

Theme 4: Hedonic Motivation

In this context, hedonic motivation is defined as the enjoyment and pleasure derived from digital technology. Media depictions of older adults using digital technologies often construct a narrative of enjoyment and pleasure, reflecting the concept of active aging and its role in enhancing quality of life. The media highlighted the positive aspects of digital technology, such as its potential to support aging in place and promote independence and self-efficacy through tools like home security systems and voice-activated assistants. However, these portrayals often framed enjoyment not in terms of amusement or leisure, but through the lens of safety, control, and reassurance, reflecting a more utilitarian interpretation of hedonic motivation. These reinterpretations had age-attributed nuances, revealing how media narratives framed older adults' interactions with digital tools based on their specific needs and challenges rather than for amusement or pleasure. For instance, difficulties in navigating digital customer service, accessing government support, and managing telehealth were common barriers depicted. This suggests a shift in how "reward" is represented for older adults in digital contexts, highlighting satisfaction derived from security and autonomy, rather than entertainment as is normally described in younger adults. These representations illustrated how media discourse on hedonic motivation intersected with practical considerations, shaping perceived motivation and engagement of older adults with digital tools. Such framings may also reflect underlying ageist assumptions that older adults engage with technology primarily for functional rather than pleasurable purposes. Textbox 3 illustrates how hedonic motivation influences older adults' digital literacy as depicted in these media narratives.

XSL•FO
RenderX

Textbox 3. Hedonic motivation (Unified Theory of Acceptance and Use of Technology 2 factor).

Hedonic motivation (HM) refers to the enjoyment, satisfaction, or pleasure derived from using technology. For older adults, this motivation is often linked to enhanced quality of life, emotional comfort, and cognitive benefits.

- HM1: "Helping seniors get connected to important services and family members is essential in supporting their full participation in society and improving their quality of life." (Vernon Morning Star, British Columbia)
- HM2: "Just having a system welcome you home by turning the lights on can be a comfort, not to mention a safety precaution." (The Chronicle Herald, Nova Scotia)
- HM3: "By pursuing computer use, social activities, and games in both midlife (age 50 - 65) and late life, benefits were noted, and a reduced risk of mild cognitive impairment was found." (The Canadian Press, Canada-wide)

Themes: satisfaction, safety, and cognition.

Discussion

Principal Findings

Our exploration of media narratives on digital literacy and older adults highlighted several key themes. Performance and effort expectancy were often portrayed negatively, with older adults seen as having limited digital skills and facing significant challenges in learning new technologies, thus necessitating extensive support. Social influence was emphasized through fearmongering, depicting older adults as "vulnerable" to social exclusion and advocating for community-based digital literacy programs to address societal expectations and support. Facilitating conditions were portrayed as crucial, with media focusing on the need for access to digital technology, technical support, and fraud prevention, though this focus on vulnerabilities may inadvertently discourage digital technology adoption by older adults. Finally, hedonic motivation was reflected in media portrayals that emphasized the enjoyment and benefits of digital technology use for enhancing independence and quality of life, while also acknowledging age-attributed challenges. Overall, media narratives both highlighted and complicated the factors influencing older adults' digital literacy, intertwining practical and motivational aspects with underlying digital ageist undertones.

Older adults were predominantly depicted as deficient in digital literacy, reinforcing digital ageism by suggesting they will be left behind unless they adapt to new technologies [9]. These narratives often used fear tactics, emphasizing the potential loss of critical services such as banking, groceries, and health care if older adults do not embrace digital technology. In addition, there was recurring portrayal of older adults as particularly susceptible to scams and fraud, with reliance on younger adults to bridge the digital literacy gap. However, recent research has found no evidence that older adults are more likely than individuals of other ages to be victims of cyber fraud [87]. Research on susceptibility to scams indicates that older adults who are more susceptible typically have lower levels of cognition [88]. However, this finding does not apply to all older adults and may indicate ageist generalizations perpetuated by news media.

Comparison With Previous Work

The media portrayals described above not only reinforce stereotypes about older adults' digital competence but also highlight a critical issue: their exclusion from essential areas such as disaster risk preparedness [89]. The exclusion of older adults in the context of digital literacy and disaster risk leads to their systematic marginalization in disaster preparedness efforts [14]. In disaster scenarios, digital technology is crucial for supporting communication between individuals, their families, and health authorities; social media is often used to disseminate health information to the masses; and telehealth serves as the primary conduit for health care during public health emergencies and subsequent lockdowns [14]. As society increasingly relies on digital tools for disseminating crucial information and coordinating emergency responses, individuals with limited digital proficiency face heightened vulnerability [90]. This exclusion exacerbates their risk during disasters, underscoring the urgent need for inclusive strategies to enhance digital literacy and ensure equitable access to disaster preparedness resources [90]. As such, it is important to consider how the narrative of older adults lacking digital literacy and thus being "vulnerable" in disasters perpetuates digital ageism and deters them from meaningfully using digital technology [91].

Performance expectancy positively predicts behavioral intention, suggesting that individuals who perceive significant benefits from digital literacy are more likely to intend to use digital technology [92]. Similarly, effort expectancy is also a positive predictor, indicating that individuals who find digital technology easy to use are more likely to intend to adopt it, thereby enhancing their digital literacy [92]. That said, it is important to distinguish between ageist stereotypes and legitimate barriers that may affect older adults' engagement with digital technology. While media portrayals often attribute digital disengagement solely to age-related decline or resistance to change, they frequently overlook key structural and design-related challenges [93]. For example, small screen sizes, complex interfaces, limited accessibility features, sensory or cognitive impairments, and the cost of digital devices or internet access can all contribute to lower digital engagement among older adults [94]. These barriers are not inherent to aging but rather reflect gaps in inclusive design, digital equity, and the need to address concerns related to privacy and safety [95]. Recognizing these factors helps reframe the issue from individual "inability" to systemic inaccessibility. Moreover, some media narratives have begun to highlight initiatives that promote inclusive digital literacy, such as community-based training programs, intergenerational tech mentoring, and age-friendly design innovations, which represent important counter-narratives to digital ageism [96,97].

XSL•FO
RenderX

However, media portrayals of older adults as lonely and socially isolated, combined with depictions of their difficulties with digital technology, can negatively impact these predictors [98]. When media frames digital technology as a crucial tool for alleviating loneliness and overcoming isolation, it highlights the perceived benefits (performance expectancy) but simultaneously underscores the challenges (effort expectancy). This dual portrayal may diminish behavioral intention by emphasizing both the urgency and complexity of adopting digital technology, potentially discouraging older adults from engaging with digital technology [99]. Social influence positively predicts behavioral intention, meaning that individuals who perceive support from important people in their social networks are more likely to intend to use digital technology [100]. Conversely, those who do not feel such support are less likely to intend to adopt digital technology [100].

Media representations of older adults as "vulnerable" to social exclusion and reliant on community centers and libraries for digital literacy workshops highlight the role of social support in digital technology adoption [101]. Facilitating conditions positively predict behavioral intention, suggesting individuals who believe in the availability of resources to aid their digital technology use are more likely to intend to use it [92]. In contrast, those who perceive a lack of such resources are less inclined to adopt digital technology [92]. Media portrayals that depict older adults as "vulnerable" to scams and frauds exacerbate concerns about internet safety and complicate their digital technology use, reinforcing the need for supportive conditions [102].

Hedonic motivation, or the enjoyment derived from using digital technology, is another positive predictor of behavioral intention [92]. Individuals who find technology enjoyable are more likely to intend to use it [92]. Our findings further align with existing literature, which suggests social perceptions of older age act as a barrier to digital technology use among this population, reinforcing stereotypes about older adults and their use of technology [91,103,104]. Notably, literature suggests that ageism influences the use and acceptance of digital technology through the inferiorization, patronage, deprioritization, and exclusion of older adults in the digital age [7,9,105,106]. Our findings contribute to the literature by highlighting how the media perpetuates digital ageism and may discourage older adults from meaningfully engaging with digital technology. Information disseminated through news media has been recognized as a potential source of negative beliefs and biased behaviors toward older adults [107]. This underscores how news media are uniquely positioned to transform narratives surrounding older adults and digital literacy. Rather than reinforcing ageist stereotypes, media outlets can foster inclusion of older adults in the digital age, promote positive representations, and support their engagement with technology, which can enhance their quality of life and participation in society.

The findings from this study also highlight how older adults are forgotten within narratives around disaster risk preparedness using digital literacy. This exclusion reflects a broader pattern of ignoring the needs of older adults in critical areas such as disaster preparedness, further perpetuating ageism and limiting their access to essential information resources. Future research should explore strategies to better integrate older adults into disaster risk management narratives and assess the impact of inclusive digital literacy initiatives on their preparedness and resilience.

Limitations

This paper offers valuable insights into how Canadian news media portrayals of older adults and digital literacy reinforce digital ageism and marginalization of older adults in disaster risk preparedness. While our analysis highlights how certain media portrayals of older adults in digital contexts may reinforce ageist narratives, we acknowledge the risk of overinterpreting all negative depictions as inherently ageist. Some representations, such as those addressing digital scams or usability challenges, may reflect legitimate concerns rather than discriminatory assumptions. As ageism researchers, we approached the data critically but remained mindful of this interpretive tension, which we recognize as a limitation and a reflection of our analytical lens. This study benefits from a robust methodological approach, including a systematic search of multiple databases and detailed CDA using UTAUT2 [23,26]. The use of diverse media sources and languages adds depth to the findings [108]. However, limitations include the decision to constrain the focus of the study on Canadian media, which may not capture global perspectives on digital literacy for older adults in disaster risk preparedness. Future research could address these limitations by expanding the scope to include international media and by developing strategies to integrate digital literacy in older adults into disaster preparedness narratives more effectively. Furthermore, we did not include social media in our search and focused strictly on newspapers and broadcast news websites (eg, Canadian Broadcasting Corporation). Yet, we recognize that social media (eg, Instagram and Facebook) has become a key medium for news sharing [109], and has changed the production flow and delivery of news for many broadcasting corporations [110]. Future research should investigate perceptions of digital literacy in older adults as discussed on social media and explore user sentiment in the comments under posts from key news organizations.

Conclusion

While this study was initially designed to examine media narratives at the intersection of digital literacy and disaster resilience among older adults, the limited presence of such discussions in news media led to a necessary broadening of scope. This absence, however, is a significant finding in itself, highlighting the exclusion of older adults from critical public discourse on digital preparedness during crises. This paper underscores the critical impact of media narratives on the digital literacy of older adults, revealing how these portrayals perpetuate digital ageism and contribute to the marginalization of older adults in essential areas such as disaster risk preparedness. By highlighting the negative portrayal of older adults' digital competencies and their exclusion from disaster preparedness discussions, the study demonstrates the need for a shift in media representation. The findings suggest media outlets can fulfill a pivotal role in transforming the narrative, promoting more inclusive and supportive portrayals of older

XSL•FO
RenderX

adults' engagement with digital technology. Future research should seek to broaden the scope to include international perspectives and explore strategies for more effective integration of older adults into digital literacy and disaster preparedness narratives.

Acknowledgments

This work was supported by a grant from the Social Sciences and Humanities Research Council of Canada (435-2022-0718 to TO). The funders had no role in the study's design; collection, analyses, or interpretation of data; writing of the manuscript; or decision to publish the results. The authors thank Téa Rokolj, a research librarian (University of Ottawa), for devising the search strategies for newspaper databases and advising on searching the websites of the Canadian Broadcasting Corporation (CBC), Canadian Television Network (CTV), and Radio-Canada.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy for older adults and digital literacy.
[DOCX File, 17 KB - aging_v8i1e69373_app1.docx]

Multimedia Appendix 2
References for newspaper articles.
[DOCX File, 20 KB - aging_v8i1e69373_app2.docx]

References

1. Arias López MP, Ong BA, Borrat Frigola X, et al. Digital literacy as a new determinant of health: a scoping review. PLOS Digit Health 2023 Oct;2(10):e0000279. [doi: 10.1371/journal.pdig.0000279] [Medline: 37556425]
2. Crawford A, Serhal E. Digital health equity and COVID-19: the innovation curve cannot reinforce the social gradient of health. J Med Internet Res 2020 Jun 2;22(6):e19361. [doi: 10.2196/19361] [Medline: 32452816]
3. Gillie M, Ali D, Vadlamuri D, Carstarphen KJ. Telehealth literacy as a social determinant of health: a novel screening tool to support vulnerable patient equity. J Alzheimers Dis Rep 2022;6(1):67-72. [doi: 10.3233/ADR-210024] [Medline: 35360275]
4. van Kessel R, Wong BLH, Clemens T, Brand H. Digital health literacy as a super determinant of health: more than simply the sum of its parts. Internet Interv 2022 Mar;27:100500. [doi: 10.1016/j.invent.2022.100500] [Medline: 35242586]
5. DeWalt DA, Broucksou KA, Hawk V, et al. Developing and testing the health literacy universal precautions toolkit. Nurs Outlook 2011;59(2):85-94. [doi: 10.1016/j.outlook.2010.12.002] [Medline: 21402204]
6. Mubarak F, Suomi R. Elderly forgotten? Digital exclusion in the information age and the rising grey digital divide. Inquiry 2022;59:469580221096272. [doi: 10.1177/00469580221096272] [Medline: 35471138]
7. Chu CH, Nyrup R, Leslie K, et al. Digital ageism: challenges and opportunities in artificial intelligence for older adults. Gerontologist 2022 Aug 12;62(7):947-955. [doi: 10.1093/geront/gnab167] [Medline: 35048111]
8. Chu C. Digital ageism and its implications for older adult inclusion. Innovation in Aging 2023 Dec 21;7(Supplement_1):312-312. [doi: 10.1093/geroni/igad104.1036]
9. Manor S, Herscovici A. Digital ageism: a new kind of discrimination. Human Behav and Emerg Tech 2021 Dec;3(5):1084-1093. [doi: 10.1002/hbe2.299]
10. Shaver J. The state of telehealth before and after the COVID-19 pandemic. Prim Care 2022 Dec;49(4):517-530. [doi: 10.1016/j.pop.2022.04.002] [Medline: 36357058]
11. Kaur SJ, Ali L, Hassan MK, Al-Emran M. Adoption of digital banking channels in an emerging economy: exploring the role of in-branch efforts. J Financ Serv Mark 2021 Jun;26(2):107-121. [doi: 10.1057/s41264-020-00082-w]
12. Trude ACB, Lowery CM, Ali SH, Vedovato GM. An equity-oriented systematic review of online grocery shopping among low-income populations: implications for policy and research. Nutr Rev 2022 Apr 8;80(5):1294-1310. [doi: 10.1093/nutrit/nuab122] [Medline: 35076065]
13. Shen H, Namdarpour F, Lin J. Investigation of online grocery shopping and delivery preference before, during, and after COVID-19. Transp Res Interdiscip Perspect 2022 Jun;14:100580. [doi: 10.1016/j.trip.2022.100580] [Medline: 35291491]
14. Lokmic-Tomkins Z, Bhandari D, Bain C, Borda A, Kariotis TC, Reser D. Lessons learned from natural disasters around digital health technologies and delivering quality healthcare. Int J Environ Res Public Health 2023 Mar 3;20(5):4542. [doi: 10.3390/ijerph20054542] [Medline: 36901559]
15. Simon T, Goldberg A, Adini B. Socializing in emergencies—a review of the use of social media in emergency situations. Int J Inf Manage 2015 Oct;35(5):609-619. [doi: 10.1016/j.ijinfomgt.2015.07.001]

XSL•FO
RenderX

16. Martínez-Alcalá CI, Rosales-Lagarde A, Pérez-Pérez YM, Lopez-Noguerola JS, Bautista-Díaz ML, Agis-Juarez RA. The effects of Covid-19 on the digital literacy of the elderly: norms for digital inclusion. Front Educ 2021 Jul;6. [doi: 10.3389/feduc.2021.716025]
17. Wilson-Menzfeld G, Gates JR, Moreland M, Raw H, Johnson A. Learning digital skills online: empowering older adults through one-to-one, online digital training provision. Front Psychol 2023;14:1122277. [doi: 10.3389/fpsyg.2023.1122277] [Medline: 37213359]
18. Vercruyssen A, Schirmer W, Geerts N, Mortelmans D. How "basic" is basic digital literacy for older adults? Insights from digital skills instructors. Front Educ 2023 Sep;8. [doi: 10.3389/feduc.2023.1231701]
19. Lu SY, Yoon S, Yee WQ, Heng Wen Ngiam N, Ng KYY, Low LL. Experiences of a community-based digital intervention among older people living in a low-income neighborhood: qualitative study. JMIR Aging 2024 Apr 25;7:e52292. [doi: 10.2196/52292] [Medline: 38662423]
20. Liu CH, Chen YT, Kittikowit S, Hongsuchon T, Chen YJ. Using unified theory of acceptance and use of technology to evaluate the impact of a mobile payment app on the shopping intention and usage behavior of middle-aged customers. Front Psychol 2022;13:830842. [doi: 10.3389/fpsyg.2022.830842] [Medline: 35310288]
21. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
22. Ayaz A, Yanartaş M. An analysis on the unified theory of acceptance and use of technology theory (UTAUT): acceptance of electronic document management system (EDMS). Computers in Human Behavior Reports 2020 Aug;2:100032. [doi: 10.1016/j.chbr.2020.100032]
23. Venkatesh V, Thong J, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 2012;36(1):157. [doi: 10.2307/41410412]
24. Yip SY. Positionality and reflexivity: negotiating insider-outsider positions within and across cultures. International Journal of Research & Method in Education 2024 May 26;47(3):222-232. [doi: 10.1080/1743727X.2023.2266375]
25. Cheng K, Wu C, Gu S, Lu Y, Wu H, Li C. WHO declares the end of the COVID-19 global health emergency: lessons and recommendations from the perspective of ChatGPT/GPT-4. Int J Surg 2023 Sep 1;109(9):2859-2862. [doi: 10.1097/JS9.0000000000000521] [Medline: 37246993]
26. Wodak R, Meyer M. Methods of Critical Discourse Analysis: SAGE Publications, Ltd; 2001. [doi: 10.4135/9780857028020]
27. Willey-Sthapit C, Jen S, Storer HL, Benson OG. Discursive decisions: signposts to guide the use of critical discourse analysis in social work. Qual Soc Work 2022 Jan;21(1):129-146. [doi: 10.1177/1473325020979050]
28. Fairclough N. Critical discourse analysis. In: The Routledge Handbook of Discourse Analysis: Routledge; 2013. [doi: 10.4324/9781315834368]
29. Mullet DR. A general critical discourse analysis framework for educational research. J Adv Acad 2018 May;29(2):116-142. [doi: 10.1177/1932202X18758260]
30. Coyle A. Discourse analysis. In: Research Methods in Psychology, 3rd edition: Thousand Oaks, CA, US: Sage Publications, Inc; 2006:366-387.
31. Yazdannik A, Yousefy A, Mohammadi S. Discourse analysis: a useful methodology for health-care system researches. J Educ Health Promot 2017;6:111. [doi: 10.4103/jehp.jehp_124_15] [Medline: 29296612]
32. Coates WC, Jordan J, Clarke SO. A practical guide for conducting qualitative research in medical education: part 2-coding and thematic analysis. AEM Educ Train 2021 Aug;5(4):e10645. [doi: 10.1002/aet2.10645] [Medline: 34585038]
33. Carol Baldwin L. Why governments should support seniors clubs. : The Canadian Press; 2023 Jul 28.
34. Leon C. $255K in federal funds support London projects for seniors. Strathroy Age Dispatch. 2023 Jul 6.
35. Become a lifeguard at any age with new Clarington program. Clarington This Week. 2023 Jul 4. URL: https://www.durhamregion.com/news/become-a-lifeguard-at-any-age-with-new-clarington-program/article_2e4885d4-5cd8-5f1a-8ef3-2b9c2718c330.html [accessed 2025-08-23]
36. Kerr P. Bruce County Public Library holds anti-scam workshop. Wiarton Echo. 2023 Dec 12.
37. City launches Healthy Aging Passport program. Brantford Expositor. 2023 Jun 13. URL: https://www.brantfordexpositor.ca/news/local-news/city-launches-healthy-aging-passport-program [accessed 2025-08-25]
38. Edwards J. Collingwood's Breaking Down Barriers sees massive growth thanks to pandemic pivoting: "I think it propelled us"; more than 3,000 people accessed services in 2022. : Collingwood Connection; 2023 Sep 22. URL: https://www.simcoe.com/news/collingwood-s-breaking-down-barriers-sees-massive-growth-thanks-to-pandemic-pivoting-i-think-it/article_debdc60b-26ce-54b6-b95c-78f2151c032c.html [accessed 2025-08-25]
39. El Elaimy L. Cybercriminals target Burnaby residents. The Canadian Press. 2024 Mar 5.
40. Federal grant for Cambridge older adult home will target resident programs. Cambridge Times. 2023 Jul 13. URL: https://www.cambridgetimes.ca/news/federal-grant-for-cambridge-older-adult-home-will-target-resident-programs/article_86504804-5580-5690-b8cd-2598b8ad01de.html [accessed 2025-08-25]
41. Gril E. Fermetures de services et de guichets bancaires: des clientèles vulnérables. Le Droit Online. 2024 Feb 23. URL: https://www.ledroit.com/partenaires/protegez-vous/2024/02/23/fermetures-de-services-et-de-guichets-bancaires-des-clienteles-vulnerables-RBDPTG27S5CULHKCTKNMCRC74A/ [accessed 2025-08-25]

XSL•FO
RenderX

42. Guelph programs for seniors get $165K from feds. Guelph Mercury-Tribune. 2023 Jun 2. URL: https://www.guelphmercury.com/news/guelph-programs-for-seniors-get-165k-from-feds/article_e85177cf-3e5d-5c18-8fc4-1865bd5c7d6e.html [accessed 2025-08-25]
43. Brierley A. Helping older adults feel less isolated and lonely this holiday season; National Institute on Ageing report finds as many as 41% of Canadians aged 50 years and older are at risk of social isolation. Barrie Advance. 2023 Dec 12.
44. Morga A. How to talk to older people in your life about scams. National Post. 2023 Nov 10. URL: https://nationalpost.com/pmn/life-pmn/how-to-talk-to-older-people-in-your-life-about-scams [accessed 2025-08-25]
45. Port Colborne program helps seniors age well. Port Colborne The Leader. 2023 Jun 28. URL: https://www.niagarathisweek.com/news/port-colborne-program-helps-seniors-age-well/article_6ac1e45e-4071-53e7-b32e-6db39942009e.html [accessed 2025-08-25]
46. Dubuc C. Un projet pour aider les retraités à réintégrer le marché du travail dans l'est ontarien. ICI Radio-Canada - Ottawa-Gatineau. 2023 Jul 3. URL: https://ici.radio-canada.ca/nouvelle/1992619/aines-retraite-ontario-employeur-retour-travail [accessed 2025-08-25]
47. Seniors, caregivers and entrepreneurs turn to tech as Canada's population gets older than ever. CBC News. 2023 Oct 27. URL: https://www.cbc.ca/news/canada/toronto/agetech-innovation-week-brings-seniors-caregivers-entrepreneurs-together-1.7007985 [accessed 2025-08-25]
48. Seniors community grants total $90,000 for programs in Lanark-Frontenac-Kingston. Perth Courier. 2023 Jun 2. URL: https://www.insideottawavalley.com/news/seniors-community-grants-total-90-000-for-programs-in-lanark-frontenac-kingston/article_0b2b5fb4-ccf8-5e98-b18c-7d4757cb9ad7.html [accessed 2025-08-25]
49. Canadore College is leading the way to healthier aging. North Bay Nugget. 2024 Mar 2. URL: https://www.nugget.ca/news/local-news/canadore-college-is-leading-the-way-to-healthier-aging [accessed 2025-08-25]
50. 100 000$ pour les ainés de Beauport-Limoilou. Mon Limoilou. 2023 May 26. URL: https://monlimoilou.com/2025/100-000-pour-les-aines-de-beauport-limoilou/ [accessed 2025-08-25]
51. Des cours pour améliorer la santé cognitive des aînés à Baie-Comeau. Radio-Canada. 2023 Sep 12. URL: https://ici.radio-canada.ca/nouvelle/2009711/centre-education-adultes-estuaire-formation-aine [accessed 2025-08-25]
52. Gillou, 78 ans, tiktokeur. La Presse. 2024 Jan 21. URL: https://www.lapresse.ca/societe/chroniques/2024-01-21/culture-web/gillou-78-ans-tiktokeur.php [accessed 2025-08-25]
53. L'AQDR Jonquière veille aux intérêts des aînés. Le Quotidien. 2023 Dec 16. URL: https://www.lequotidien.com/la-vitrine/2023/12/16/laqdr-jonquiere-veille-aux-interets-des-aines-DGLYCUPSPVDTLOSUVCKQO7TIWQ/ [accessed 2025-08-25]
54. La fracture numérique: un enjeu crucial pour les personnes aînées. Journal de rue de l'Estrie. 2023 Dec 1. URL: https://www.jdrestrie.ca/2023/12/05/la-fracture-numerique-un-enjeu-crucial-pour-les-personnes-ainees/ [accessed 2025-08-25]
55. Le guichet automatique, ce nouveau «Serpuarien». Le Nouvelliste. 2024 Feb 5. URL: https://www.lenouvelliste.ca/opinions/parole-au-lecteur/2024/02/05/le-guichet-automatique-ce-nouveau-serpuarien-OQ5H7VR26RFHJIAKJXRTM6A2MY/ [accessed 2025-08-25]
56. Descôteaux D. Les aînés seront plus durement affectés par ce changement. Journal de Québec. 2023 Nov 4.
57. LGBTQ passé 50 ans: de l'angoisse aux solutions. Le Quotidien. 2023 May 13. URL: https://www.lequotidien.com/la-vitrine/vie-active-50/2023/05/13/lgbtq-passe-50-ans-de-langoisse-aux-solutions-D2SMB6TEFJBE3FQMBCCQRZLBQ4/ [accessed 2025-08-25]
58. Littéracie numérique: nouveau projet pour prévenir les inégalités. L'Écho La Tuque - Haut St-Maurice Online. 2024 Feb 18. URL: https://www.lechodelatuque.com/infolettre/litteracie-numerique-nouveau-projet-pour-prevenir-les-inegalites/ [accessed 2025-08-25]
59. Longueuil: penser la cité pour les aînés. Le Courrier du Sud. 2024 Jan 18. URL: https://www.lecourrierdusud.ca/longueuil-penser-la-cite-pour-les-aines/ [accessed 2025-08-25]
60. Faucher E. Nouvelle vague de fraudes ciblant les aînés à Trois-Rivières. Le Nouvelliste. 2023 Dec 6. URL: https://www.lenouvelliste.ca/actualites/justice-et-faits-divers/2023/12/06/nouvelle-vague-de-fraudes-ciblant-les-aines-a-trois-rivieres-44PFZZ4ISVB43LE54YKRAFI2YE/ [accessed 2025-08-25]
61. Bérubé G. Responsable d'être escroqué. Le Devoir. 2024 Feb 17.
62. Archambault H. Santé: des aînés de plus en plus déments. Journal de Montréal. 2023 Oct 20.
63. Brunet L. Techno-maltraitance Notre vie quotidienne s'est déplacée vers le monde des transactions « en ligne». Le Devoir. 2023 Jun 30. URL: https://www.ledevoir.com/opinion/idees/793838/idees-techno-maltraitance [accessed 2025-08-25]
64. Une entreprise granbyenne veut combattre l'exclusion numérique. Granby Express. 2023 Dec 27. URL: https://www.granbyexpress.com/infolettre/une-entreprise-granbyenne-veut-combattre-lexclusion-numerique/ [accessed 2024-03-21]
65. Les aînés qui fréquentent internet réduiraient leur risque de démence. Radio-Canada. 2023 May 13. URL: https://ici.radio-canada.ca/nouvelle/1979416/stimulation-cerveau-connaissances-sante-mentale-cognitive?depuisRecherche=true [accessed 2025-08-25]
66. Cyber seniors tutoring in Sicamous. Eagle Valley News. 2023 Sep 21.
67. Leung V. Digital connection: keeping seniors up-to-date with online news. Richmond News. 2023 Dec 18. URL: https://www.richmond-news.com/in-the-community/digital-connection-keeping-seniors-up-to-date-with-online-news-7992108 [accessed 2025-08-25]

XSL•FO
RenderX

68. Pass J. For Shona: honour and help our elders – benefit is it just might change your life. Comox Valley Record. 2023 Oct 19. URL: https://www.comoxvalleyrecord.com/community/for-shona-honour-and-help-our-elders-it-just-might-change-your-life-5888770 [accessed 2025-08-25]
69. Smith J. Is it your year to volunteer? Vernon Morning Star. 2024 Jan 18.
70. McCrae D. My family has been left by the health-care wayside. The Province. 2023 Dec 6. URL: https://theprovince.com/opinion/op-ed/devin-mccrae-my-family-has-been-left-by-the-health-care-wayside [accessed 2025-08-25]
71. Seyd J. Ottawa grants $732K to help seniors in West Vancouver. North Shore News. 2024 Mar 13.
72. Scammers target seniors. Here's how to protect them. Burns Lake Lakes District News. 2024 Jan 31.
73. McManus T. Today is World Elder Abuse Awareness Day in New Westminster. New Westminster Record. 2023 Jun 16. URL: https://www.newwestrecord.ca/local-news/today-is-world-elder-abuse-awareness-day-in-new-westminster-7152771 [accessed 2025-08-25]
74. Crawford T. VPD hands out free 911-only mobile phones for low income seniors. Vancouver Sun. 2023 Sep 20. URL: https://vancouversun.com/news/local-news/vancouver-police-free-911-phones-for-seniors [accessed 2025-08-25]
75. Campbell F. "We paid our dues": senior calls for government help to keep aging residents at home. PNI Atlantic News. 2023 Dec 8. URL: https://www.saltwire.com/atlantic-canada/we-paid-our-dues-senior-calls-for-government-help-to-keep-aging-residents-at-home-100918762 [accessed 2024-04-02]
76. Connors C. Ding-dongs and blue cheese sauce: Tech Tea with Teens helps Cape Breton seniors navigate digital world. PNI Atlantic News. 2024 Mar 3. URL: https://www.saltwire.com/atlantic-canada/ding-dongs-and-blue-cheese-sauce-tech-tea-with-teens-helps-cape-breton-seniors-navigate-digital-world-100944244 [accessed 2025-08-25]
77. Field K, LeBlanc L. Senior living: technology helps many seniors overcome isolation. The Chronicle Herald. 2023 Sep 26.
78. Sethi C. Comment: coming together, literally, can solve our housing crisis. Western Investor. 2023 Jun 24. URL: https://www.westerninvestor.com/alberta/comment-coming-together-literally-can-solve-our-housing-crisis-7190487 [accessed 2025-08-25]
79. Restigouche: un site web pour le rayonnement des aînés. Acadie Nouvelle. 2023 Jul 1.
80. McDougall M. Students help seniors with tech skills. The Brandon Sun. 2023 Jun 10. URL: https://www.brandonsun.com/local/2023/06/10/students-help-seniors-with-tech-skills [accessed 2025-08-25]
81. Iantorno M. Who winds up left behind when retail is automated? The Brandon Sun. 2024 Feb 16.
82. AGE-WELL, en collaboration avec le Réseau canadien des soins aux personnes fragilisées, annonce 17 projets de recherche dans le cadre du Programme de financement Catalyst pour le vieillissement en santé. AGE-WELL. 2023 Jun 8. URL: https://agewell-nce.ca/fr/archives/17223 [accessed 2025-08-25]
83. Le gouvernement du Canada finance 112 projets au Manitoba pour assurer l'autonomie des aînés au sein de leur communauté. Canada NewsWire. 2023 May 23. URL: https://www.newswire.ca/fr/news-releases/le-gouvernement-du-canada-finance-112-projets-au-manitoba-pour-assurer-l-autonomie-des-aines-au-sein-de-leur-communaute-842040510.html [accessed 2025-08-25]
84. Les banques répondent aux besoins financiers uniques des aînés. Canada NewsWire. 2023 Jun 5. URL: https://www.newswire.ca/fr/news-releases/les-banques-repondent-aux-besoins-financiers-uniques-des-aines-860847659.html
85. Montréal-Nord se mobilise pour améliorer la qualité de vie de ses aîné.e.s. Canada NewsWire. 2023 Jun 8. URL: https://www.newswire.ca/fr/news-releases/montreal-nord-se-mobilise-pour-ameliorer-la-qualite-de-vie-de-ses-aine-e-s-852164152.html [accessed 2025-08-25]
86. Ève Ménard i de journalisme local. Vieillir, ailleurs dans le monde. La Presse Canadienne - Le fil radio. 2023 Sep 28.
87. Shang Y, Wu Z, Du X, Jiang Y, Ma B, Chi M. The psychology of the internet fraud victimization of older adults: a systematic review. Front Psychol 2022;13:912242. [doi: 10.3389/fpsyg.2022.912242] [Medline: 36132192]
88. James BD, Boyle PA, Bennett DA. Correlates of susceptibility to scams in older adults without dementia. J Elder Abuse Negl 2014;26(2):107-122. [doi: 10.1080/08946566.2013.821809] [Medline: 24499279]
89. Bell SA, Singer D, Solway E, Kirch M, Kullgren J, Malani P. Predictors of emergency preparedness among older adults in the United States. Disaster Med Public Health Prep 2021 Oct;15(5):624-630. [doi: 10.1017/dmp.2020.80] [Medline: 32475374]
90. Tilahun B, Gashu KD, Mekonnen ZA, Endehabtu BF, Angaw DA. Mapping the role of digital health technologies in prevention and control of COVID-19 pandemic: review of the literature. Yearb Med Inform 2021 Aug;30(1):26-37. [doi: 10.1055/s-0041-1726505] [Medline: 34479378]
91. Mannheim I, Varlamova M, van Zaalen Y, Wouters EJM. The role of ageism in the acceptance and use of digital technology. J Appl Gerontol 2023 Jun;42(6):1283-1294. [doi: 10.1177/07334648231163426] [Medline: 36917039]
92. Nordhoff S, Louw T, Innamaa S, et al. Using the UTAUT2 model to explain public acceptance of conditionally automated (L3) cars: a questionnaire study among 9,118 car drivers from eight European countries. Transportation Research Part F: Traffic Psychology and Behaviour 2020 Oct;74:280-297. [doi: 10.1016/j.trf.2020.07.015]
93. Chen Y, Yin Q, Zhou T, et al. Digital inclusion pathways and influencing factors among older adults in outpatient settings: a grounded theory study. Clin Interv Aging 2025;20:701-716. [doi: 10.2147/CIA.S518045] [Medline: 40454304]

XSL•FO
RenderX

94. Fakhimi MM, Hughes A, Gustavson AM. Evaluating smart home usability and accessibility in early detection and intervention of mental health challenges among older adults: a narrative review and framework. JAL 2025 Mar;5(1):3. [doi: 10.3390/jal5010003]
95. Xu H, Xie B, Chang CH. Empowering the care of older adults through the use of technology. Work Aging Retire 2024 Jan;10(1):1-5. [doi: 10.1093/workar/waad030] [Medline: 38196828]
96. Mulligan K. Digital inclusion, online participation and health promotion: promising practices from community-led participatory journalism. Glob Health Promot 2023 Jun;30(2):35-39. [doi: 10.1177/17579759221126150] [Medline: 36287081]
97. Juris JJ, Bouldin ED, Uva K, Cardwell CD, Schulhoff A, Hiegl N. Virtual intergenerational reverse-mentoring program reduces loneliness among older adults: results from a pilot evaluation. Int J Environ Res Public Health 2022 Jun 10;19(12):7121. [doi: 10.3390/ijerph19127121] [Medline: 35742370]
98. Donovan NJ, Blazer D. Social isolation and loneliness in older adults: review and commentary of a national academies report. Am J Geriatr Psychiatry 2020 Dec;28(12):1233-1244. [doi: 10.1016/j.jagp.2020.08.005] [Medline: 32919873]
99. Kebede AS, Ozolins LL, Holst H, Galvin K. Digital engagement of older adults: scoping review. J Med Internet Res 2022 Dec 7;24(12):e40192. [doi: 10.2196/40192] [Medline: 36477006]
100. Kelly S, Kaye SA, Oviedo-Trespalacios O. What factors contribute to the acceptance of artificial intelligence? A systematic review. Telematics and Informatics 2023 Feb;77:101925. [doi: 10.1016/j.tele.2022.101925]
101. Kamin ST, Beyer A, Lang FR. Social support is associated with technology use in old age. Z Gerontol Geriat 2020 May;53(3):256-262. [doi: 10.1007/s00391-019-01529-z]
102. Havers B, Tripathi K, Burton A, Martin W, Cooper C. A qualitative study exploring factors preventing older adults from reporting cybercrime and seeking help. CrimRxiv . [doi: 10.21428/cb6ab371.8c4e3181]
103. McDonough CC. The effect of ageism on the digital divide among older adults. GGM 2016 Jun;2(1):1-7. [doi: 10.24966/GGM-8662/100008]
104. Quan-Haase A, Williams C, Kicevski M, Elueze I, Wellman B. Dividing the grey divide: deconstructing myths about older adults' online activities, skills, and attitudes. American Behavioral Scientist 2018 Aug;62(9):1207-1228. [doi: 10.1177/0002764218777572]
105. Köttl H, Allen LD, Mannheim I, Ayalon L. Associations between everyday ICT usage and (self-)ageism: a systematic literature review. Gerontologist 2023 Aug 24;63(7):1172-1187. [doi: 10.1093/geront/gnac075] [Medline: 35670751]
106. Rosales A, Fernández-Ardèvol M. Ageism in the era of digital platforms. Convergence (Lond) 2020 Dec;26(5-6):1074-1087. [doi: 10.1177/1354856520930905] [Medline: 33239961]
107. Kroon AC, Trilling D, Van Selm M, Vliegenthart R. Biased media? How news content influences age discrimination claims. Eur J Ageing 2019 Mar;16(1):109-119. [doi: 10.1007/s10433-018-0465-4] [Medline: 30886565]
108. Canestrino R, Magliocca P, Li Y. The impact of language diversity on knowledge sharing within international university research teams: evidence from TED project. Front Psychol 2022;13:879154. [doi: 10.3389/fpsyg.2022.879154] [Medline: 35529548]
109. Knutson B, Hsu TW, Ko M, Tsai JL. News source bias and sentiment on social media. PLoS One 2024;19(10):e0305148. [doi: 10.1371/journal.pone.0305148] [Medline: 39441772]
110. Hendrickx J, Vázquez-Herrero J. Dissecting social media journalism: a comparative study across platforms, outlets and countries. Journal Stud 2024 Jul 3;25(9):1053-1075. [doi: 10.1080/1461670X.2024.2324318]

Abbreviations

CDA: critical discourse analysis
UTAUT: Unified Theory of Acceptance and Use of Technology
UTAUT2: Unified Theory of Acceptance and Use of Technology 2

Edited by G Seçkin; submitted 28.11.24; peer-reviewed by A Nüssler, E Maranesi, M Hess; revised version received 26.06.25; accepted 14.07.25; published 29.08.25.

Please cite as:
Ménard A, Maharaj A, Harb S, Fraser S, O'Sullivan T
Representations of Older Adults' Digital Literacy in Canadian News Media: Critical Discourse Analysis Using Unified Theory of Acceptance and Use of Technology 2
JMIR Aging 2025;8:e69373
URL: https://aging.jmir.org/2025/1/e69373
doi:10.2196/69373

XSL•FO
RenderX

© Alixe Ménard, Alexandra Maharaj, Sophie Harb, Sarah Fraser, Tracey O'Sullivan. Originally published in JMIR Aging (https://aging.jmir.org), 29.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Older Adult's Views of the Age-Inclusivity of Physical Activity Websites Using the Think Aloud Method: Qualitative Analysis

Veda Clemson[1], BSc, MSc; Elisabeth Grey[2,3], BSc, MSc, PhD; Julie Barnett[4], BSc, PhD; Ella Burfitt[1], BSc; Fiona Gillison[1], BSc, MSc, PhD

[1]Department for Health, University of Bath, Claverton Down, Bath, United Kingdom

[2]The National Institute for Health and Care Research Applied Research Collaboration West (NIHR ARC West), University Hospitals Bristol and Weston NHS Foundation Trust, Bristol, United Kingdom

[3]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom

[4]Department of Psychology, University of Bath, Bath, United Kingdom

Corresponding Author:
Fiona Gillison, BSc, MSc, PhD
Department for Health, University of Bath, Claverton Down, Bath, United Kingdom

Abstract

Background: Older adults are the least active in our society and may face additional barriers to taking part in physical activity compared with those experienced by younger people because of factors such as lower digital literacy and negative stereotypes of aging.

Objective: This study aimed to explore how older adults navigate websites that provide access to physical activity opportunities and facilities and make judgments about their suitability.

Methods: Semistructured interviews were embedded within a think-aloud approach. Participants were shown a series of websites and asked to navigate through the websites as if they were going to take up what was on offer, articulating their thoughts and comments out loud as they progressed. Participants viewed up to 4 websites, rotated from a pool of 8, including leisure centers, exercise products, gyms, or community organizations. Additional questions were asked about perceptions of the inclusivity of the websites at the end of the interview. Digital recordings were made and transcribed verbatim, and analyzed using thematic analysis.

Results: Nineteen participants (6 male and 13 female) aged between 65 and 84 years were recruited from southern England; one-third reported having poor digital ability prior to taking part. Three overarching themes relating to the research question were identified as follows: (1) signals of age-inclusivity, (2) limiting beliefs, and (3) confidence in making judgments. Older adults inferred a lot of information about how welcome they would be in physical activity settings from the images and language used on websites. They showed a preference for imagery that was inclusive of age, body shape, and physical ability, not only for those depicting older adults themselves. Some adults reported firm views about the type and intensity of physical activity that is appropriate for older adults, and many expressed a specific dislike of gyms, based on both the (young) age of most users and perceived emphasis towards aesthetic rather than health-related exercise. While most participants could navigate websites successfully, they preferred to visit venues and speak to staff to gain greater confidence that they would feel welcome and that the activities would be at a suitable level.

Conclusions: Websites providing access to physical activity could be more inclusive of older adults by using more diverse imagery, providing clearer descriptions of the activities on offer, along with details of the level of fitness or ability needed to take part, and providing alternatives to web-based booking. Additional societal-level approaches to reducing age-limiting self-stereotyping may also be useful in expanding the opportunities for older adults to access mainstream provisions for physical activity.

(JMIR Aging 2025;8:e68951) doi:10.2196/68951

KEYWORDS

ageism; inclusivity; digital literacy; older adult; physical activity

XSL•FO
RenderX

Introduction

Older adults are the least active in our society [1], yet they could derive significant benefits from physical activity in terms of preventing and managing chronic disease, promoting strength and function [2], and reducing morbidity [3]. The UK government recommends 150 minutes of moderate-intensity activity per week for older adults, as well as breaking up sedentary periods and including exercise that improves muscle strength, balance, and flexibility [4,5]. Data for 2022 to 2023 from Sport England estimate that around 53% of people older than 65 years achieve this, with 36% achieving less than 30 minutes of activity per week [6]. Activity in later life reduces the risk of disability, the development of noncommunicable diseases, and the decline in functional capacity associated with aging [7]. There may also be benefits to cognitive health [8].

Older adults experience many of the same barriers reported by younger adults to physical activity, including time, cost, and enjoyment, and lack of confidence or fitness [9,10]. Additional barriers that are more commonly seen in older adulthood relate to physical function, such as ill health or disability, safety concerns, and fear of falling [10]. However, there are also indirect barriers that may influence engagement with facilities and activities in older age, including poorer digital literacy, digital exclusion [11], and negative stereotypes around aging [12]. Considering digital literacy, while many older adults do own smartphones or other digital devices, digital acceptability and access decline with age [13]. The physical activity industry mirrors the rest of society in having become more digitized, for example, through moving to web-based bookings, payments, and advertising, and the development of increasingly technical exercise equipment. These changes have an unintended negative impact on access for older adults.

Older adults face several age stereotypes relevant to the physical activity domain, often being portrayed as frail, easily fatigued, and physically inactive [14]. Some of this stereotyping can be observed in how older adults are treated by others, for example, fitness professionals commonly have (and communicate to greater or lesser degrees) low expectations of older adults' capabilities, which in turn discourages older adults from taking part [15]. Providing positive examples of the capabilities of older adults that present a very high level of ability (ie, exceptional fitness) can also be problematic to some older adults who feel they may not be able to match up to expectations, putting them off taking part [16]. Age stereotypes for physical activity are not only held by younger adults, older adults themselves often share these negative perceptions, for example, feeling that their aging body means they are less able to exercise, and that people their age are out of place in physical activity settings, such as gyms and health clubs [17].

The theory of stereotype embodiment sets out one way in which self-stereotyping may develop, proposing that people internalize negative stereotypes throughout their lives, developing negative attitudes toward older adults, which are directed toward themselves as they age [18]. This leads to unconscious effects on behavior, particularly in settings where age stereotypes are primed (eg, depictions of older adults with limited ability), impacting health behaviors [19]. Little research from this perspective has been undertaken specifically in relation to physical activity, and research exploring the impact of age stereotyping on physical activity from other theoretical perspectives has resulted in some conflicting results. For example, in response to age-based stereotype threats (ie, pressure to avoid confirming or exacerbating stereotypes when taking part in an activity; [20]), some studies show poorer performance in physical activities, such as walking speed [21], while others show improved performance, for example in fitness training [22]. Thus, research to provide greater insight into older adults' responses to how they are portrayed during their first approach to physical activity settings is warranted and could help us to understand the variation in effects.

The aim of this study was to explore how websites providing information about physical activity opportunities may influence the barriers or facilitators of physical activity for older adults. Using a "think aloud" approach, our objectives were to (1) explore the social signals older adults identified of inclusivity, fit, or welcome, and (2) capture their views about the digital capability required to find relevant information. Together, this would provide information on how older adults experience web-based gateways to access physical activity opportunities and what information they draw on in assessing whether the activity on offer was suitable for them.

Methods

Ethical Considerations

Ethics approval was granted by the Research Ethics Committee for Health at the University of Bath (EP 20/21 105). All participants provided informed written consent, and any identifying information was anonymized prior to data analysis. Interviews were conducted between January 2022 and August 2023.

Design

This study was situated within a constructivist research paradigm, selecting a "think aloud" approach in order to explore and understand people's subjective experiences and focus on meaning-making in real time [23]. The think aloud approach has shown previous utility in exploring human-computer interactions and decision-making [24] and the usability of websites in health settings [25].

The approach encourages participants to verbalize their thoughts or feelings while navigating web-based resources. In our study, this was embedded within semistructured interviews to ensure interviewees had been asked comparable questions when interpreting the findings, and to allow participants to express their thoughts and experiences at a more general level [26]. The interviews were all conducted in person by slim, White, young women (aged 18 - 23 years); we include this detail as it was commented on by some participants, so it may have influenced some of their responses.

Participants

For this study, we considered adults to be aged 65 years or older, in line with the UK Office for National Statistics [27]. The

XSL•FO
RenderX

inclusion criteria were individuals aged 65 years and older, who can read and speak fluent English, self-assessed as physically able to, and interested in taking part in physical activity. Participants were recruited through advertisements posted in local community centers, on local social media platforms (such as "Nextdoor"), and through the authors' existing networks. We aimed to recruit 20 older adults to provide a range of experiences, focusing on advertising in lower socioeconomic areas. Participants received travel expenses and refreshments but no payment for their time.

Procedure

Volunteers were sent a participant information sheet and consent form and invited to set up a time for their interview; these arrangements were typically made via email, but could be by phone or letter if preferred. Interviews were conducted in public spaces of the participants' choice (to ensure they would feel comfortable) using a laptop provided by the researcher. Participants were asked to nominate a café local to them or could choose to come to the University. At the start of the interview, the researcher talked through the process and obtained written consent. The researcher then presented between 1 and 4 example websites in turn, working through the think aloud process with each until the participant wished to stop or the 1 hour had passed; participants were offered the choice whether to continue at the end of each website viewing. The order in which websites were presented was rotated to ensure that the findings related to a variety of website designs and topics.

Participants were encouraged to provide an honest opinion on the websites, add any thoughts which came to mind [28], and were prompted to reflect on what they were doing as they browsed (eg, "why did you click on that?"; Multimedia Appendix 1 contains the full topic guide). In line with a relaxed think aloud approach, the interviewer only interacted with the participant when asking prompts or in response to a participant's request for help with navigating the websites [28]. At the end of browsing websites, participants were asked to evaluate the content overall. One example of the questions asked for this was "were there any aspects of the websites that made you feel included or excluded?".

The pool of example websites was chosen to reflect different aspects of the physical activity industry following public engagement activity with 7 older adults as part of a related study (ie, what they consider as part of the industry), and discussion with an advisor from a national representative body. Our aims were to include both commercial and noncommercial services, sites that advertise equipment as well as facilities, and ensure we featured a range of activities that are accessible to someone starting out (eg, activities that do not require extensive technical skills, and can be started at low intensity). The research team reviewed the final set to ensure we included variations in website designs (Multimedia Appendix 1 lists the specific websites chosen). Two examples were available from each of the following 5 categories: local authority providers of physical activity services; fitness industry providers (eg, generalist gyms that focused on building strength and endurance); fitness industry product providers (eg, personal fitness equipment and clothing); community organizations that encourage physical activity (eg, walking groups); and UK physical activity campaigns. The order in which the websites were presented was rotated.

After the interview, participants provided information on their demographics, usual physical activity level, and self-rated health.

Data Analysis

Interviews were audio-recorded using a digital voice recorder, transcribed verbatim by VC, and anonymized. Analysis was conducted using a 6-step thematic analysis [29], including familiarization, generating initial codes (including a code book), organizing codes into clusters of shared meaning (themes and subthemes), and refining and naming the themes with the wider research team. Two transcripts were coded by VC and FG to refine the codebook, and the remainder were then coded by VC and the themes were developed iteratively.

Website tracking software was used to generate a recording of participants' progress in order to verify which specific pages participants were looking at as they spoke. This information has been added in square brackets within quotes where appropriate.

Results

Sample

Participants (6 male and 13 female) were aged between 65 and 84 years and based in Southern England. Fourteen (74%) participants had postschool qualifications, and the majority of the participants considered themselves to be healthy (95% scored >7 on a 10-point scale), with moderate fitness (most reported a steady or brisk walking pace) (Table 1).

Around a third of participants considered that they had poor digital ability, but the majority were able to navigate the websites without difficulty, and the challenges experienced most came from using unfamiliar hardware (ie, using the researcher's laptop, rather than a usual PC at home) rather than difficulty using common website functions (scrolling down, search function, clicking through different pages). At the extremes, 2 participants were not confident in navigating websites unaided, so they asked the researcher to do this for them (P11 and P18), while one participant was a website designer prior to retiring (P4).

Three overarching themes relating to the research question were identified as follows: (1) signals of age-inclusivity, reflecting what participants read into the images and language used on websites; (2) limiting beliefs, reflecting how participants related beliefs about their own physical capability, and age stereotypes to website content; and (3) trust, reflecting older adults' reservations about what could be understood from web-based information alone.

XSL•FO
RenderX

Table . Participant characteristics.

	Age group (years)	Gender	Ethnic group	Highest educational attainment	Usual walking pace	Health rating[a]
1	65 - 69	Man	White	Bachelor degree	Steady	8
2	80 - 84	Woman	White	Bachelor degree	Brisk	7
3	75 - 79	Woman	White	Master or postgraduate qualification	Brisk	7
4	65 - 69	Woman	Black or African	GCSE/O level[b]	Slow	3
5	75 - 79	Woman	White	Master or postgraduate qualification	Steady	8
6	70 - 74	Man	White	GCSE/O level	Brisk	8
7	65 - 69	Man	White	Bachelor degree	Steady	7
8	70 - 74	Man	White	Master or postgraduate qualification	Steady	8
9	75 - 79	Man	White	Master or postgraduate qualification	Steady	7
10	65 - 69	Woman	White	Master or postgraduate qualification	Steady	7
11	65 - 69	Woman	White	Bachelor degree	Brisk	8
12	65 - 69	Woman	White	Bachelor degree	Steady	9
13	75 - 79	Woman	White	Master or postgraduate qualification	Slow	10
14	65 - 69	Woman	White	Master or postgraduate qualification	Steady	9
15	65 - 69	Woman	White	No formal qualifications	Slow	8
16	65 - 69	Woman	White	Master or postgraduate qualification	Brisk	8
17	80 - 84	Woman	White	A Level or equivalent	Steady	7
18	75 - 79	Woman	Asian or Asian British	No formal qualifications	Brisk	7
19	65 - 69	Man	Black or African	Master or postgraduate qualification	Slow	8

[a]Self-rated on a scale of 1 - 10 where 1 is very bad health and 10 is excellent health.

[b]GCSE/O level: General Certificate of Secondary Education Ordinary Level.

Theme 1: Signals of Age-Inclusivity

Images

Participants drew on images, language, and the type of physical activity or service offered in judging whether websites were relevant to them. Images were often the first source of information that participants considered ("I noticed almost immediately there was a mix you know, that in photos there were older people" [P9]) and were often taken very literally as an indication of who uses a particular facility. Participants extrapolated from who was depicted as to whether they would be welcome as a participant:

> *This is the gym… oh… group exercise... oh let's see if there's any for old people. [inspects the photograph]. There's a young lady… that's not for me! Absolutely not for me. If they had a photograph of a group with a couple of older people involved in the same group yep, I don't mind being with young people at all, except I wouldn't like to be the only one.* [P2]

> *It's modelled by people, young people without wobbly bits.* [P15]

While seeing older adults depicted was interpreted as a clear indicator that a website was offering something intended for them, most participants used a broader benchmark and judged websites as age-inclusive when they saw images of people of different ages, genders, races, and abilities, as well as people with disabilities. For example, on articulating why certain websites looked age-inclusive, participants stated:

> *[Because] they're all normal people, as in, of an age that I can relate to. So it's not, it's not big punchy*

XSL•FO
RenderX

> *"do this to build your muscle" type thing it's all about… it's all about keeping active and keeping healthy through, through activity, which is good.* [P6]

> *That one is, I think that's very, the seems quite welcoming that, I'd have a look at it, I would if I saw that picture... Because it's a family there, and there's a child there, it is not as though someone's just exercising and looks really fit (laughs). I mean, I'm not saying they don't look fit but I mean it's not a real over-fit.* [P17]

In addition to images, participants responded positively to clear statements that all ages are welcome:

> *Um… they [community rowing club] particularly stress the headline [ages] 13 to 100 so, it definitely was not focused just on the… you know the normally you associate the, I suppose I associate rowing with the boat race, young fit men or women from um from the Universities etc. That's my first reaction so um this seemed quite inclusive in that sense, quite friendly, seems quite informative.* [P9]

Similarly, images of people of all body sizes were considered positive and inclusive, regardless of their age. As such, body size appeared to serve as a proxy for participants believing an activity would be achievable for them:

> *But I mean, but you see that's still good to get people going isn't it because they're all shapes and sizes [in the image] and I'd rather do classes with people like that then just really old people so that's quite nice* [P17]

> *They are… women and they are older women, kind of thing that they are not… slim ones so they really appeal to the likes of me that it can be me, joining them, you see.* [P7]

While wanting to see older adults depicted, participants could be very particular in how old, or how active the older adults were. Some interviewees distanced themselves from websites where older adults were depicted as being too old or inactive:

> *They are very old people and they don't appeal to me at all. I couldn't interact with anybody like that. I mean, okay, they might get satisfaction out of it but nah. It's not me, it's just not me.* [P15]

Very few older adults expressed an interest in exercise sessions that are solely available to older adults, preferring to see indicators of inclusivity to the whole community, with older and larger people depicted as part of this. This may reflect (as suggested by the quote from P15) the degree to which participants have actively embraced an older identity.

Language

In some cases, the language used on websites signaled to participants that these services or facilities were not for them. This was particularly evident in websites relating to gyms and exercise classes, which included new "branded" terms such as classes called "SHRED," "Trax," and "Rush". Participants viewed this terminology as an immediate turn-off, and an indicator that such activities were not for them:

> *It's delving into a sort of more modern language that I don't necessarily understand, but that's probably more my problem than anybody else's!* [P6]

> *Oh… what are glute exercises? Oh… glute, did you know about glute muscles? I didn't.* [P13]

> *Shred (laughing)… see I, well I'm just looking I'm thinking right, so I immediately focused in on classes and I'm not going to go and do a class that says shred because um I just know that that is just, you know, I wouldn't do it.* [P12]

Some participants were explicit in describing the lack of explanation of such classes as discriminatory:

> *I don't think for my particular age group that they are informative enough and I find that they are a bit discriminative and they're… ageist. That's it.* [P15]

Theme 2: Limiting Beliefs

Perceived Physical Capability

Most participants recognized and described a deterioration in their physical ability as they had aged and expressed a desire to avoid injuries by attempting to do too much. While some participants had long-term conditions such as arthritis that limited their activity, for many, this reflected a fear of injury based on the perceived limitations of age alone. This caution influenced their responses to websites based on the type of exercise on offer:

> *With the equipment, I'd end up… tearing something or ending up injuring something of my body. So, I've come to the conclusion that you know, probably it's best not to be going into gyms.* [P4]

This concern was also reflected in participants' desire for more information about the intensity of activities than was provided on most websites. This is often related to expected levels of fitness and capability to start engaging in the exercise on offer, reflecting the fact that although they might have been experienced in some activities (ie, not a beginner), they knew they were much less fit than when they last took part. For example, 2 participants commented on the website of a local walking group that was marketed towards older adults, and depicted older adults in their imagery:

> *It would be helpful if there was a bit more information, to give you… the severity of the walk, is it rough terrain or not. Because obviously as you get older, the flatter the walk the more… probability that you'll actually go onto it…* [P6]

> *Older people do tend to walk a little bit slower so you want to know is this going to take me a couple of hours, or is this going to take all day?* [P6]

> *My concern would be walking out for the first time … I'd be worried about lagging behind and upsetting people who are experienced walkers. If they had a 'break me in walk' and then oh you are broken in a bit now so you can do the next one, and the next one.* [P15]

Similar concerns were offered about fitness classes:

XSL•FO
RenderX

> *It [the website] didn't exclude me, it's just I, my body and everything else that comes with it, makes me feel… a bit worried to try to go and lift weights or do some strenuous thing that I know my body can't take… Do I need to be a certain fitness level before? That's something I need to know.* [P4]

Others expressed concerns about potentially embarrassing themselves in mixed settings or pushing themselves too far. For example, on looking at a strength training class video, one participant commented:

> *If you're not fit I don't think you would ever look [to go] there because you would feel such a fool. You would, you would think I'm a wimp, I couldn't cope with that, so I wouldn't go.* [P17]

Some older adults also expressed concern that younger instructors or trainers would not be able to adapt to what was needed for older adults, and thus that what was demanded of them would be too great. These participants tended to use the age of instructors depicted on websites as a signal of whether a service or facility was appropriate for them:

> *I don't know whether I've just got more respect in general for older people because they've got more experience, maybe that, I just don't quite believe that someone – sorry about this [referring to the young age of the interviewer] – does a three-year course in whatever, fitness, can necessarily have all the information they need to help older people.* [P11]

> *[I'd need] way more support, but who's going to be supporting me, I mean, some squeaky chicky ain't going to come over and say 'Alright, old love let's sort you out!'* [P15]

Preference for older instructors was not a universally held view, and those who had more experience exercising in mixed-age groups provided examples of younger instructors who adapted classes well. They also flagged the need to differentiate people by ability rather than age:

> *My Pilates class is basically for over 50's but they want to stop it too because they say people don't like being defined by their age… so it's like how do you find another way of saying over 55 without patronising somebody?... But at the same time sometimes you don't know what you're looking for. There's some people who are over 65 and they don't need to look at an over 65 sort of over 50s class because they're really fit and they can still do stuff in an ordinary class.* [P12]

Stereotyping

While participants were generally pragmatic and objective in acknowledging age-related changes in physical capability, there was also some evidence of stereotypical thinking. Some of this appeared to relate to self-stereotyping of what types of physical activity older adults should be doing. The example below relates to a comment made while reading through a list of exercise classes at a municipal leisure center.

> *... I'd go for body balance but yoga, no, spin, sounds too fast, those aren't right for old people… If its high intensity, we [older adults] shouldn't be doing things high intensity.* [P2]

However, many of the interview excerpts coded into this theme related to gyms. It was often hard to disentangle participants' pre-existing attitudes towards gyms as a type of business (and their experience of them being predominantly used by young people), from the specific aspects of the website presented ("I'm actually a little bit against gyms anyway, because, it's a modern thing" [P7]). These views typically reflected participants' attitudes towards gyms rather than what was presented on a website, as well as their experience of gyms.

> *I didn't like going in the actual gym bit because it was all people a lot younger than me and they were all trying to outdo each other.* [P15]

> *I see these [gyms] and I only tend to see young people. You know, cool looking sort of young people going into (particular gym), and that doesn't, that doesn't do anything - this website - to disabuse me of that viewpoint. It comes across as a younger, commercial first [place].* [P9]

> *I mean I wouldn't go anywhere near that place [gym]. You know I, not being rude, but I mean cause to me that is just full of… people who are like nearly old enough to be my grandchild so I'm like just wouldn't (laughing), so I wouldn't go in there no.* [P12]

In some cases, the stereotypical thinking reflected assumptions that gyms are a place people go to do image—rather than health—enhancing activity. Often, the negative statements about young people are conflated with assumptions about why people use gyms, given the greater prevalence of young people using them.

> *It's all a bit showy-off-y to me [going to the gym], rather than, I suspect… intended just to get you fit and keep you fit. It's a bit, it's all a bit show-y.* [P7]

> *When I have been into the odd hotel gym there's generally some very, very narcissistic p- you know people there looking at themselves more in the mirror than they are [exercising], you know what I mean, it er, to a younger person you know a, beautiful blonde there or really fit young bloke you know, you know doing whatever I'm thinking, yep, you don't need that.* [P9]

Some participants were less critical of the reasons why people attend gyms, but in believing gyms are for building muscle, and that building muscle is something that only younger people do, they assumed that gyms were not for them. This was reinforced by images of young fit adults lifting weights:

> *When you get over 65, you're more interested probably in er… you're not going to build muscle, you're not going to, you know, attain any major goals. You want to, keep mobile keep flexible, and if there were more things to build upon flexibility, um… that would be more appealing to people of an older age*

XSL•FO
RenderX

group, yes… There doesn't seem to be any of that on this website. [P6]

They [images] all look quite frightening to me, because they all look so much younger and fitter! … I mean look at the, these guys are all pretty stacked, so. I'll never look like that. Not that I ever did! [P6]

Theme 3: Confidence in Making Judgments

Participants preferred websites where information was presented clearly on the first page, where the font was simple, without too many distractions from moving images. The majority considered information about the proximity of facilities ("See, straight away I think what I want to do is find out, where's the nearest one for me?" [P6]), costs, and links to a person they could contact to be of primary importance. They wanted to be able to see this information without needing to click through multiple screens.

I want to know what the membership is and it's giving me "hear from our members." I want to know where it is, how much it is, swimming pay as you go [P15]

So this is, they're not going to give me any information about courses until… until I open the portal and I'm not interested in that. [P5]

Even among those comfortable with using websites and digital tools, many participants expressed a preference to establish some personal contact before committing; they looked for a phone number of someone to talk to, or stated that they would not book via web, but would use the site to gather information and then visit the venue to book. For some, the need to verify information in person reflected a lack of trust in what is presented via web, rather than poor digital literacy.

It's all a marketing ploy, not really giving me the information that I really want. [P15]

I… can't tell, what they would be like from the website… but it's all part of an advertise-advertising their product um so, so yo-you… that's how I would expect it to be I'd expect it to be advertising a product I wouldn't expect to see any truth there [P7]

This hesitancy may also reflect participants' interest in factors that cannot be ascertained from a website, like the atmosphere of the venue or how they are treated by staff.

If I was going to join the gym, I'd go down there and physically look at it. So, um, but only based on having… drawn up a-a-a-a list of 3 say [from their websites]… and that's what I would do and then go and look at them, and see what I thought and judge it on, on um… the people that are um… are um running it so forth, do you get a welcome er a good welcome there, do they seem to know what they're talking about. [P7]

Hesitancy about using websites to make decisions on the appropriateness of physical activity opportunities appeared to differ from general comfort with digital devices or web-based shopping. For example, P7, who would not trust the "truth" of a leisure center website, was comfortable with buying sports equipment via web ("I buy my sports stuff online, or off mates"). Others reported being able to use a device, but not being comfortable with website bookings as a specific function:

Well I'm not very good at websites to be honest, and actually although I'm on it on my iPad with the gym, my son has to book me online all the time. I can't do it, it's sad init? [P18]

Discussion

Principal Findings

This study explored older adults' reactions and views of the age-inclusivity of websites providing access to physical activity opportunities. Together, the themes showed how older adults drew on the imagery and language of websites to inform their decisions, as well as some existing schemas of what types and intensities of activity are appropriate for older adults (ie, implicit ageism). Participants were encouraged by websites depicting people of mixed ages, body shapes, and abilities using their service, and were not looking for services offered purely for older adults. Many older adults judged whether an opportunity was appropriate for them from factors that signaled the intensity and challenge of the physical activity on offer, motivated by a wish to avoid injury or embarrassment, and felt the ability to keep up with others. Information on the level of ability of fitness required to take part was rarely presented directly, which may have caused participants to rely on images and language to make these inferences. Participants showed a greater preference for and trust in websites that provided clear and up-front information on basic considerations such as cost, location, and contact details. Nonetheless, they reported still being reluctant to believe an opportunity was for them without having visited a venue and spoken to staff.

The think aloud approach has been used in previous research to explore how people use and make sense of websites, and what they base their judgments on [28]. For example, users have been shown to use short case examples to understand the relevance of the site by drawing constant comparisons between the examples presented, their own experience [28]. While in this study few websites that we used included case studies (although these were popular where they did), we did observe a similar effect in how participants were extrapolating information from the individuals shown in the images based on their experience and existing attitudes; for example, some participants inferred knowledge about a "typical user's" motivation and fitness level just from a photo. This process was most closely observed in older adults' responses to images of young people, using this information to form judgments that a facility or service was not for them. In part, this reflected participants' explicit narratives of how their capability and motives for physical activity had changed over time, such that they judged what is suitable for young adults would not be suitable for them. While some settings (eg, gyms [30] and surfing [15]) are already considered a young adults' domain by default, our study showed that in other settings too, seeing images predominantly of young adults was used as a heuristic for assessing an activity as irrelevant or inappropriate. This response was diminished when images included mixed age groups, ethnicities, and body shapes alongside young adults.

XSL•FO
RenderX

Past work has described how older adults who take part in physical activity need to navigate the tension between experiencing the benefits of physical activity, noticing their own physical decline, and dealing with social norms about what type of physical activity is acceptable in older age [30]. This tension appeared to be visible in our study, acting as a lens through which participants were evaluating whether the activities that the website was promoting were relevant and welcoming to them. In line with other work [10], we found that older adults' beliefs about their capabilities were a common barrier to physical activity, and evidence of an apparent acceptance of the inevitability of physical decline [17]. Current health messaging, as well as marketing of physical activity facilities and opportunities, may assume that older adults are interested because they believe that there is scope for improvement, so exposing this alternative perspective—that older adults largely do not feel that improvement in strength and fitness is feasible—could be important to inform future approaches.

While participants in this study showed a good ability to navigate websites and good digital literacy, they still exhibited some lack of confidence in applying this. Two participants either asked for help from the Researcher in navigating the websites or reported that they normally ask a younger family member to make web-based bookings or purchases for them. This aligns with previous research demonstrating that even among those with high levels of competency in using websites, older adults may have lower perceptions of their usefulness [31] and acceptability [11]. In this study, this lower perceived usefulness may relate to the fact that information older adults consider crucial was not provided; websites rarely provided explicit statements about welcoming all ages, contact details, clarity about the level of starting fitness required, or information on how demands can be tailored. However, alternatives to web-based booking, whether by phone or in person, will likely remain an important part of providing an age-inclusive offer, to allow older adults to test out the atmosphere of a venue and sample interactions with staff.

Strengths and Limitations

The use of think aloud interviews worked well in this setting by encouraging participants to verbalize their thoughts and experiences while interacting with websites in "real time." It is argued that this often results in more honest accounts of user experience [32]. A limitation of this approach was that participants were navigating websites that they had not chosen for themselves, on a device that was not familiar to them. Participants may have been less distracted by the mechanics of navigating a website if using a familiar device. Our findings may also have been influenced by the specific websites we chose. While we attempted to mitigate this by identifying a range of different websites and rotating these through interviews, providing choice to participants (eg, allowing them to nominate the types of physical activity there were interested in) may have helped to provide a greater test of participants' views on the inclusivity of the website, rather than the type of activity it offered. In particular, it was hard to disentangle participants' comments on the desirability of gyms and the imagery that accompanied them, reported in theme 2, from their preexisting attitudes.

We attempted to recruit a diverse sample by advertising in areas populated by people of lower income and with greater ethnic diversity, and by seeking support from gatekeepers (eg, local Black and minority ethnic support groups for older adults). However, the majority of those willing to take part were White, highly educated, and living in urban areas, which is a limitation of this study. Other studies conducted with older adults report greater challenges with website navigation [25], which we may have found if we had recruited a sample with a more diverse socioeconomic background. Our participants were also relatively healthy; this is less a limitation and more a reflection of our design, as in order to address our research question, we recruited individuals for whom the websites were relevant, that is, those who are capable of engaging in physical activity. However, we may have discovered different barriers had we included more people with limited physical ability. Related to this, the age of our participants spanned around 2 decades from 65 to 84 years, from those still working to those who have been retired for 20 years. People at either end of this age group, as well as those experiencing greater or lesser physical decline or limitations, may have had different perspectives on who the websites depicted and what was offered. Within this small qualitative study, we were not able to tease out these differences.

Implications and Future Directions

The findings of this study suggest that older adults are more likely to consider that services and facilities are relevant and attractive to them when websites depict a diverse set of users, including an explicit statement indicating people of all ages are welcome, and present clear information to users to judge the level of fitness and ability likely to be required to take part. Remaining physically active for longer is key to reducing age-related disability and disease [7] and promoting good cognitive health [8]. In addition, including older adults in intergenerational settings can foster the mental and physical health benefits associated with social inclusion and connection [33,34]. If increased physical activity participation in the growing population of older adults is to be achieved, their needs need to be accommodated affordably within mainstream facilities and services.

A common reason why older adults wanted more information on the likely demands of the physical activity on offer was to reassure them that they would not be pushed too far, feel embarrassed, and that the activity could be adapted safely for someone of their age and level of ability. Classifications referring to the level of experience (beginner, intermediate, etc) may not be meaningful when someone has historical experience but has not taken part in many years, and a lack of explanation alongside activities that have nontraditional names can lead older adults to infer that this is not intended for them. Evidence of venues using older adults as instructors and coaches, and information to confirm that activities can be adapted to be appropriate for all ability levels and ages (where this is the case) would also increase older adults' confidence in trying out physical activity opportunities. While many older adults can navigate websites competently, they expressed a preference for simple information and were put off by too many moving images and the need to click through multiple links to find information.

XSL•FO
RenderX

Conclusions

This study has provided insight into what older adults notice and look for in websites when exploring opportunities to increase their physical activity. It flagged how their pre-existing beliefs and implicit assumptions may influence whether they consider new options for activity, taking into account perceptions of their own capability, social norms, and some stereotypes around age and motivation for physical activity. These existing beliefs may interact with web content to exacerbate or potentially challenge stereotypes, to encourage older adults to feel more welcome to use mainstream facilities and services that support physical activity.

Acknowledgments

We would like to thank Ms Natanya Cassidy and Ms Abigail Russell for their contribution to the interviews, as well as all the participants who gave up their time to take part. EG's time is supported by the National Institute for Health and Care Research Applied Research Collaboration West (NIHR ARC West)

Authors' Contributions

FG, JB, EG, and EB contributed to the conceptualization of the study. VC, EB, and AR were responsible for data collection. VC and FG conducted the analysis. FG and VC drafted the initial version of the manuscript. All authors contributed to further writing, editing, and refinement of the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary materials containing the interview topic guide and a list of websites used in the study.
[DOCX File, 19 KB - aging_v8i1e68951_app1.docx]

References

1. Lindsay T, Westgate K, Wijndaele K, et al. Descriptive epidemiology of physical activity energy expenditure in UK adults (The Fenland study). Int J Behav Nutr Phys Act 2019 Dec 9;16(1):126. [doi: 10.1186/s12966-019-0882-6] [Medline: 31818302]
2. Dipietro L, Campbell WW, Buchner DM, et al. Physical activity, injurious falls, and physical function in aging: an umbrella review. Med Sci Sports Exerc 2019 Jun;51(6):1303-1313. [doi: 10.1249/MSS.0000000000001942] [Medline: 31095087]
3. Chudasama YV, Khunti KK, Zaccardi F, et al. Physical activity, multimorbidity, and life expectancy: a UK Biobank longitudinal study. BMC Med 2019 Jun 12;17(1):108. [doi: 10.1186/s12916-019-1339-0] [Medline: 31186007]
4. Sparling PB, Howard BJ, Dunstan DW, Owen N. Recommendations for physical activity in older adults. BMJ 2015 Jan 21;350:h100. [doi: 10.1136/bmj.h100] [Medline: 25608694]
5. Physical activity guidelines: UK chief medical officers' report. : Department of Health and Social Care; 2019.
6. Active lives survey. : Sport England; 2024.
7. Cunningham C, O' Sullivan R, Caserotti P, Tully MA. Consequences of physical inactivity in older adults: a systematic review of reviews and meta‐analyses. Scandinavian Med Sci Sports 2020 May;30(5):816-827. [doi: 10.1111/sms.13616]
8. Bherer L, Erickson KI, Liu-Ambrose T. A review of the effects of physical activity and exercise on cognitive and brain functions in older adults. J Aging Res 2013;2013(1):657508. [doi: 10.1155/2013/657508] [Medline: 24102028]
9. Bethancourt HJ, Rosenberg DE, Beatty T, Arterburn DE. Barriers to and facilitators of physical activity program use among older adults. Clin Med Res 2014 Sep;12(1-2):10-20. [doi: 10.3121/cmr.2013.1171] [Medline: 24415748]
10. Spiteri K, Broom D, Bekhet AH, de Caro JX, Laventure B, Grafton K. Barriers and motivators of physical activity participation in middle-aged and older-adults - a systematic review. J Aging Phys Act 2019 Sep 1;27(4):929-944. [doi: 10.1123/japa.2018-0343] [Medline: 31141447]
11. Gallistl V, Rohner R, Hengl L, Kolland F. Doing digital exclusion - technology practices of older internet non-users. J Aging Stud 2021 Dec;59:100973. [doi: 10.1016/j.jaging.2021.100973] [Medline: 34794717]
12. Marquet M, Chasteen AL, Plaks JE, Balasubramaniam L. Understanding the mechanisms underlying the effects of negative age stereotypes and perceived age discrimination on older adults' well-being. Aging Ment Health 2019 Dec 2;23(12):1666-1673. [doi: 10.1080/13607863.2018.1514487]
13. Choudrie J, Pheeraphuttranghkoon S, Davari S. The digital divide and older adult population adoption, use and diffusion of mobile phones: a quantitative study. Inf Syst Front 2020 Jun;22(3):673-695. [doi: 10.1007/s10796-018-9875-2]
14. Ng R. Societal age stereotypes in the U.S. and U.K. from a media database of 1.1 billion words. Int J Environ Res Public Health 2021 Aug 21;18(16):8822. [doi: 10.3390/ijerph18168822] [Medline: 34444578]
15. Jin B, Harvey IS. Ageism in the fitness and health industry: a review of the literature. J Aging Phys Act 2021 Feb 1;29(1):99-115. [doi: 10.1123/japa.2019-0230] [Medline: 32759410]

XSL•FO
RenderX

16. Wangler J, Jansky M. Media portrayal of old age and its effects on attitudes in older people: findings from a series of studies. Humanit Soc Sci Commun 2023;10(1):1-9. [doi: 10.1057/s41599-023-01671-5]
17. Evans AB, Sleap M. "You feel like people are looking at you and laughing": older adults' perceptions of aquatic physical activity. J Aging Stud 2012 Dec;26(4):515-526. [doi: 10.1016/j.jaging.2012.07.004] [Medline: 22939548]
18. Levy B. Stereotype embodiment: a psychosocial approach to aging. Curr Dir Psychol Sci 2009 Dec 1;18(6):332-336. [doi: 10.1111/j.1467-8721.2009.01662.x] [Medline: 20802838]
19. Levy BR, Myers LM. Preventive health behaviors influenced by self-perceptions of aging. Prev Med 2004 Sep;39(3):625-629. [doi: 10.1016/j.ypmed.2004.02.029] [Medline: 15313104]
20. Lamont RA, Swift HJ, Abrams D. A review and meta-analysis of age-based stereotype threat: negative stereotypes, not facts, do the damage. Psychol Aging 2015 Mar;30(1):180-193. [doi: 10.1037/a0038586] [Medline: 25621742]
21. Barber SJ, Hamel K, Ketcham C, Lui K, Taylor-Ketcham N. The effects of stereotype threat on older adults' walking performance as a function of task difficulty and resource evaluations. Psychol Aging 2020 Mar;35(2):250-266. [doi: 10.1037/pag0000440] [Medline: 31971413]
22. Deshayes M, Palermo A, Korchi K, Philippe AG. Influence of negative stereotype on physical activity level among older adults during a training session. Front Sports Act Living 2022;4:998724. [doi: 10.3389/fspor.2022.998724] [Medline: 36506719]
23. Koro-Ljungberg M, Douglas EP, Therriault D, Malcolm Z, McNeill N. Reconceptualizing and decentering think-aloud methodology in qualitative research. Qual Res 2013 Dec;13(6):735-753. [doi: 10.1177/1468794112455040]
24. Cotton DRE, Gresty KA. The rhetoric and reality of e‐learning: using the think‐aloud method to evaluate an online resource. Assess Eval High Educ 2007 Oct;32(5):583-600. [doi: 10.1080/02602930601116920]
25. Bolle S, Romijn G, Smets EMA, Loos EF, Kunneman M, van Weert JCM. Older cancer patients' user experiences with web-based health information tools: a think-aloud study. J Med Internet Res 2016 Jul 25;18(7):e208. [doi: 10.2196/jmir.5618] [Medline: 27457709]
26. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a knowing researcher. Int J Transgend Health 2023;24(1):1-6. [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
27. Profile of the older population living in england and wales in 2021 and changes since 2011. Office for National Statistics. 2025. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/ageing/articles/profileoftheolderpopulationlivinginenglandandwalesin2021andchangessince2011/2023-04-03 [accessed 2025-03-17]
28. Breuning M, Schäfer-Fauth L, Lucius-Hoene G, Holmberg C. Connecting one's own illness story to the illness experiences of others on a website-an evaluation study using the think aloud method. Patient Educ Couns 2020 Jan;103(1):199-207. [doi: 10.1016/j.pec.2019.08.014] [Medline: 31466883]
29. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
30. Tulle E, Dorrer N. Back from the brink: ageing, exercise and health in a small gym. Ageing Soc 2012 Oct;32(7):1106-1127. [doi: 10.1017/S0144686X11000742]
31. Campelo AM, Katz L. Older adults' perceptions of the usefulness of technologies for engaging in physical activity: using focus groups to explore physical literacy. Int J Environ Res Public Health 2020 Feb 11;17(4):1144. [doi: 10.3390/ijerph17041144] [Medline: 32053937]
32. Alhadreti O, Mayhew P. Rethinking thinking aloud: a comparison of three think-aloud protocols. Presented at: CHI '18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems; Apr 21-26, 2018; Montreal, Canada. [doi: 10.1145/3173574.3173618]
33. Krzeczkowska A, Spalding DM, McGeown WJ, Gow AJ, Carlson MC, Nicholls LAB. A systematic review of the impacts of intergenerational engagement on older adults' cognitive, social, and health outcomes. Ageing Res Rev 2021 Nov;71:101400. [doi: 10.1016/j.arr.2021.101400] [Medline: 34237435]
34. Zhou F, Zhang H, Wang HY, Liu LF, Zhang XG. Barriers and facilitators to older adult participation in intergenerational physical activity program: a systematic review. Aging Clin Exp Res 2024 Feb 12;36(1):39. [doi: 10.1007/s40520-023-02652-z] [Medline: 38345651]

Edited by Y Jiang; submitted 27.11.24; peer-reviewed by A Chasteen, D Sullivan, M Griffin; revised version received 17.03.25; accepted 01.05.25; published 19.06.25.

Please cite as:
Clemson V, Grey E, Barnett J, Burfitt E, Gillison F
Exploring Older Adult's Views of the Age-Inclusivity of Physical Activity Websites Using the Think Aloud Method: Qualitative Analysis
JMIR Aging 2025;8:e68951
URL: https://aging.jmir.org/2025/1/e68951
doi:10.2196/68951

XSL•FO
RenderX

© Veda Clemson, Elisabeth Grey, Julie Barnett, Ella Burfitt, Fiona Gillison. Originally published in JMIR Aging (https://aging.jmir.org), 19.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Factors Influencing Older Adults' Perception of the Age-Friendliness of Their Environment and the Impact of Loneliness, Technology Use, and Mobility: Quantitative Analysis

Eric Balki[1], BSc, MA, MSc, PhD; Niall Hayes[2], BSc, MSc, PhD; Carol Holland[1], BSc, PhD

[1]Department of Health Research, Faculty of Health and Medicine, Lancaster University, Lancaster, United Kingdom

[2]Keyworth Building, Leeds University, Leeds, United Kingdom

Corresponding Author:
Eric Balki, BSc, MA, MSc, PhD
Department of Health Research
Faculty of Health and Medicine
Lancaster University
Health Innovation One, Sir John Fisher Dr
Bailrigg
Lancaster, LA1 4AT
United Kingdom
Phone: 44 1524 65201
Email: e.balki@lancaster.ac.uk

Abstract

Background: The World Health Organization's (WHO) publication on age-friendly environments (AFEs) imagines future cities to become more age-friendly to harness the latent potential of older adults, especially those who have restricted mobility. AFE has important implications for older adults in maintaining social connections, independence, and successful aging-in-place. However, technology is notably absent in the 8 intersecting domains of AFEs that the WHO imagines improve older adult well-being, and we investigated whether technology should form a ninth domain. While mobility was severely restricted, the COVID-19 pandemic provided an opportunity to test how older adults' perceptions of their AFE changed and what role technology was playing.

Objective: This study examined how life-space mobility (LSM), a concept for assessing patterns of functional mobility over time, and loneliness impacted perceived AFEs and the moderating effect of technology. It also explores whether technology should play a greater role as the ninth domain of the WHO's imagination of the AFE of the future.

Methods: In this cross-sectional quantitative observation study, data from 92 older adults aged 65-89 years were collected in England from March 2020 to June 2021 during the COVID-19 pandemic. The Life-space Questionnaire, Technology Experience Questionnaire, UCLA (University of California, Los Angeles) Loneliness Scale, and age-friendly environment assessment tool were used. Correlation and moderation analyses were used to investigate relationships between variables.

Results: Most participants (86/92, 93%) had not left their immediate town in the previous 4 weeks before the interview. Restricted LSM was positively correlated to the age-friendly environment assessment tool, that is, rising physical isolation was linked to a better perception of AFEs; however, we discovered this result was due to the moderating impact of increased use of technology, and that restricted LSM actually had a negative effect on AFEs. Loneliness was correlated negatively with the perception of AFEs, but technology use was found to moderate the impact of loneliness.

Conclusions: Pandemic-related LSM restrictions impacted perceived AFEs and loneliness negatively, but technology played a moderating role. The findings demonstrate that technology could be considered as a ninth domain in the WHO's assessment of AFEs for older adults and that there is a need for its explicit acknowledgment.

(JMIR Aging 2025;8:e67242) doi:10.2196/67242

KEYWORDS

COVID-19; age-friendliness of environments; physical isolation; digital communication technologies; loneliness; cross-sectional; WHO; World Health Organization; older adults; reduced mobility; age friendliness of environments; adult well-being; social

XSL•FO
RenderX

connections; aging in place; life-space mobility; LSE; functional mobility; UCLA loneliness scale; age-friendly environment assessment tool; AFEAT

Introduction

Background

Enabling the development of age-friendly environments (AFEs), defined as a physical and social setting that helps people age well and participate in their communities to promote "aging-in-place" for the mental and physical well-being of older adults, has become an increasingly important policy issue. This is in part a response to an aging population dynamic, urbanization, intensification of housing concerns, and community preferences, causing an increase in the deterioration of mental wellness, anxiety, stress, and depression among other disorders [1]. Older adults' perception of AFEs directly affects their quality of life (QOL) and known predictors of depression, including loneliness [2,3]. In urban environments, older adults tend to spend much of their time in their local neighborhoods and are therefore sensitive to change [4,5]. Environmental degradation, such as lack of resources, restricted access to health care, or crime, brings additional challenges.

Previous studies have demonstrated that the perception of AFEs affects the action space of older people, affecting their social participation [6]. It is known to positively moderate the relationship between frailty, mental well-being, and depression, particularly in older adults with poor health, limited mobility, and cognitive decline [1,7,8]. Moreover, a lack of social opportunities can result in loneliness, social isolation, and worsened depression, all of which exacerbate cognitive decline, QOL, and frailty [8-14]. Aboderin et al [15] and Ng et al [16] also noted a reduction in older adults' psychological resilience due to sustained stress and anxiety resulting from crowding, lack of space, and other issues that revolve around a poor age-centered environment. Other research has demonstrated a correlation between loneliness and AFEs [17-21], and identified factors linking the psychological health of economically disadvantaged older adults with their perception of AFE. Rantakokko et al [22] described the impact of person-environment interaction on mental well-being, demonstrating a mechanism by which older adults who experienced loneliness perceived obstacles with their environment. Stafford [23] demonstrated that older age insularity and the accompanying withdrawal from social interaction can result in deteriorating relationships, loneliness, poorer mental health, and a diminished perception of AFEs. If the environment is not conducive to aging-in-place, it causes difficulties in accessing services [24] and an increased risk of physical and mental health care needs [25]. As such, the perception of AFEs must be considered carefully to improve understanding of older adults aging-in-place and mental well-being.

Life-space mobility (LSM) describes the physical environment a person inhabits daily, structured into various zones (called life-space zones), centered on an anchor (eg, bedroom), and expanding outward into the rest of the house, house perimeter, local community, neighborhood, or town [26,27]. This concept corresponds to individuals' functional mobility enabling meaningful participation in community activities. It reflects how older adults move across life-space zones over a given period while incorporating frequency and independence.

According to the United Nations Development Program report "The Sustainable Development Goals and COVID-19" [28], there are unprecedented ongoing burdens on mental health for older adults. Social distancing orders during the COVID-19 pandemic were conceived as advantageous to protect potentially vulnerable populations, such as older adults [29], from disease transmission; however, it has led to long-lasting effects on mental well-being.

Restricted LSM impedes older adults' access to their [30-32] choice of environments and is associated with potential adverse mental health outcomes [27]. In addition, previous research has demonstrated that LSM restrictions reduce participation in out-of-home [33] activities and negatively affect QOL [9,34], which is in turn associated with loneliness [35].

Although pandemics create restricted mobility [36], older adults with knowledge of technology can capitalize on a digitally enabled AFE to avoid loneliness and LSM restrictions that create a strain on mental well-being. It enables older adults to engage in internet-based activities they may have enjoyed in person and gain access to health care, civic services, therapy, counseling [37], and resources that help with mental well-being [38,39]. We therefore expected both variables (LSM and loneliness) to impact AFEs negatively and for technology to moderate their impact.

Feldman and Oberlink [40] pioneered the concept of AFEs, identifying important components such as social engagement, enhancing independence, and optimizing mental functioning and well-being. Age-friendliness is derived from an ecological model of aging, in which a person's mental well-being results from the interaction between their functional and cognitive capacity (competencies) and the environmental characteristics that exert pressure on these competencies (environmental stress) [41].

The age-friendly environment assessment tool (AFEAT) was designed to assess older adults' perception of their environment and focuses on individual-oriented age-friendliness and individual-environment interaction, providing a more holistic picture [11,42,43]. Self-perceived AFEs are associated with improved QOL and mental health regardless of older adult frailty or abilities [8].

Although there is currently no universally accepted definition of an AFE, the World Health Organization (WHO) tentatively defines such communities as those where "policies, services, settings and structures support and enable people to age actively." The WHO [1] publication imagines future cities to become more age-friendly to harness the latent potential of older adults through 8 intersecting domains addressing obstacles to older adults' mental well-being: respect or social inclusion, outdoor spaces and buildings, housing, social participation, transportation, communication and information, civic

XSL•FO
RenderX

participation and employment, and community support and health services. Explicit mention of technology is notably absent.

The WHO framework has been criticized for the exiguous technology element in the domains, prompting a reexamination [44]. Incorporating technology into AFEs in its broadest sense has become an increasingly important area for aging independence and mental well-being [45] in recent years.

Marston and van Hoof [44] discussed the incorporation and use of technology within the AFE assessment agenda, and a recent systematic umbrella review elaborated on the advantages of technology interventions for social connectedness [46]. Their call for the inclusion of digital technology as a domain in AFE evaluations is beginning to resonate.

For example, Pedell et al [47,48] advocated for digital elements to encompass all aspects of environmental age-friendliness, in addition to mental and physical aspects, to realize benefits for older adults. Moreover, Reuter et al [49] acknowledged the WHO's age-friendly city initiative, considering an aging population amid increasing urbanization. However, they determined that such initiatives overlooked technology as a critical component of global digitalization.

Research Objectives

Our study aimed to examine the potentially complex relationship between internal (loneliness) and external (LSM) factors that influence older people's perception of AFEs and to determine whether technology moderates their impact. We aimed to answer a question with an important implication: should technology be included as a ninth domain in the WHO's global age-friendly cities guide in the assessment of AFEs in cities and communities for older adult mental well-being?

We set out to test the following hypotheses:

- Hypothesis (H1): LSM restriction is negatively correlated with perception of AFEs.
- H2: Increased feelings of loneliness are correlated with poor perception of AFEs.
- H3: Technology use moderates the impact of LSM restrictions on the perception of AFEs.
- H4: Technology use moderates the impact of loneliness on the perception of AFEs after considering the LSM restriction effect.

Methods

Study Design

This was a cross-sectional quantitative observational design. This report follows the STROBE (Strengthening Reporting of Observational Studies in Epidemiology) guidelines [50] (STROBE checklist mentioned in Multimedia Appendix 1).

Setting

Participants were recruited from the United Kingdom. Data were collected from January 16, 2020, to June 21, 2021, a period when social distancing mandates were enforced and social engagement outside the home was restricted.

Participants

Eligible participants were required to be living in their own home, be proficient in English, and be 65 years or older. Older people who lived in nursing or care homes or those with cognitive decline or mental health issues were excluded. Volunteers were recruited via advertisements posted in resource centers for older adults, housing associations, third-sector organizations, social activity clubs, local senior groups, direct human interaction, and word-of-mouth recommendations. Volunteers were instructed to either call and leave a voicemail or send an email indicating their willingness to participate; a return call confirmed their eligibility.

G*Power software (Version 3.1; Erdfelder, Faul, and Buchner) was used to calculate the minimum sample size required for the empirical validation of the tested moderation model. Multiple regression was used, with effect size f^2 of 0.15, power of 0.80, and 3 predictors. The recommended sample size was determined to be 87. A total of 110 participants enrolled; however, 18 did not complete all questionnaires and were excluded. The sample achieved included 92 people between the ages of 65 and 92 years (mean age 74.6, SD 7.23 years). All participants identified as either male or female, with more females (55/92, 60%) than males. More than 89% of the participants were White, with less than 11% of minority ethnicities (7 British Asian, 3 British Black). Having collated various demographic information such as age, gender, ethnicity, and education level, we were able to ascertain that participants emanated from diverse sociodemographic backgrounds.

Variables and Measures

All participants filled out a health history questionnaire based on the Scientific Advisory Group for Emergencies (SAGE) Encyclopedia of Communication Research Methods [51].

Loneliness was measured using the 20-item UCLA (University of California, Los Angeles) Loneliness Scale [52], with scores ranging from 20 to 80. Higher scores reflected higher loneliness (Cronbach α=0.88).

Utilization of technology was evaluated using the Technology Experience Questionnaire [53]. The participants were given a list of technologies (communication, computer, daily, health, recreational, and transportation technology) and asked to rate their familiarity with and use of each on a 5-point scale. Scores ranged from 0 to 180, with higher scores indicating greater use and familiarity with technology (Cronbach α=0.84).

The perception of the AFE was assessed by the AFEAT. This is a 10-item measure that uses a 5-point Likert scale, scoring items from (1) strongly disagree to (5) strongly agree and gauging participants' perceptions of their home, their local communities, the resources within the environment, and their appropriateness for meeting their daily needs. The scores ranged from 0 to 50, with higher scores representing a more positive perception of the age-friendliness of the environment (Cronbach α=0.75).

LSM was measured using the Life-Space Questionnaire [26]. Participants were asked yes or no questions about specific places they visited in the last 4 weeks, starting with another room in

XSL•FO
RenderX

their current residence and increasing the distance to a location outside England. The scores ranged from 9 to 18, with higher scores demonstrating greater restriction of LSM (Cronbach α=0.90).

Procedure

Telephone surveys collected information on loneliness, technology use, LSM, and perceptions of AFEs, in addition to basic demographic information (eg, age, education, gender, and ethnicity). Google Analytics was used to collect and tabulate the data. Participants completed the assessments across 14 months.

Statistical Analysis

Analyses were performed using IBM SPSS version 28, using a minimum significance level of 95% probability. There were no missing data, and participants completed all questions. The variables of AFEAT, loneliness, technology use, and LSM were inspected for kurtosis and skewness to assess their distribution deviations from normality via a histogram with simulated overlapping normal curves. Moreover, the homoscedasticity of the residuals was checked using a standardized residual versus a standardized predicted plot. Using the Mahalanobis (P<.001) and Cook distances, we determined whether high leverage points, significant outliers, or highly influential points exist by examining a scatterplot matrix of the dependent and continuous independent variables. A linear regression was performed to check the included variance caused by the data point and if it needed to be removed from the dataset. The criterion for discarding observations was the inability to meet 2 of the distance measures' 3 gauges. However, no outliers were found that would significantly impact the findings, and thus, none were removed. The confirmation of the independence of the observations and the assumption of no autocorrelation in the residuals was checked using the Durbin-Watson d-statistic.

The initial descriptive analyses contained means, frequencies, and SD. Pearson product-moment correlation coefficients (r) were calculated to determine whether associations exist between the variables. The same correlational analysis was used to determine whether the perception of AFE is correlated with LSM during the pandemic (H1) and whether loneliness is correlated with the perception of AFEs (H2).

Hayes' [54] PROCESS macro for SPSS with model 1 was applied to investigate the moderating effects of technology use on the relationship between LSM and AFE (H3) and technology use on the relationship between loneliness and AFEs. If the standardized coefficients of the interaction terms were significant (P<.05) or marginally significant (P<.09), we conducted a simple slope test to examine the interaction effect at different levels to reveal the nature of significant interactions to further explain the moderating effect.

Ethical Considerations

Participants accessed an information sheet either via email or read on the phone and were allowed to ask questions before giving their consent. Before completing the questionnaires, each participant gave informed consent to volunteer without compensation and participate in the study. All participants were fully anonymized.

All participants were informed of their rights to withdraw at any point in the research and informed about anonymity. The ethical procedures were aligned with the guidelines of the British Psychological Society, and the study received ethical approval from the Lancaster University Faculty of Health and Research Ethics Committee (reference number FHMREC19121). Data captured via telephone first confirmed the participant's identity and was recorded in spreadsheets and anonymized thereafter.

Results

Overview

Table 1 shows mean, SD, kurtosis, and skewness values and Shapiro Wilk test results. Kurtosis and skewness values had a relatively small range of ±1; we determined that the normal distribution deviation was insignificant. The distributions of the variables of loneliness, AFE, technology experience, and LSM were close to normal.

Table 1. Descriptive statistics for loneliness, life-space, age-friendly environment assessment tool, and technology experience (N=92).

	Scores					
	Minimum	Maximum	Mean (SD)	Skewness	Kurtosis	Shapiro Wilk test
UCLA[a] loneliness score	21	80	47.49 (17.814)	0.204	−1.630	0.192[b]
Life-space mobility	10	18	13.70 (1.595)	0.515	0.035	0.201[b]
Age-friendliness of environment	0	35	19.51 (9.687)	0.019	−1.399	0.168[b]
Technology experience	48	175	116.87 (40.951)	−0.260	−1.624	0.204[b]

[a]UCLA: University of California, Los Angeles.

[b]P<.001 under moderate.

Participants demonstrated high levels of loneliness, with 44% of older adults demonstrating loneliness scores greater than 50, with scores above 40-50 considered moderate loneliness and scores greater than 50 considered high [55]. LSM scores were high, with more than 93% of the participants scoring >11, showing that they had not been outside their immediate town. Previous prepandemic studies with similar sample sizes and methodology reported almost half that score [26]. The

XSL•FO
RenderX

perceptions of AFEs were mixed, with a mean score of 19.51 (SD 9.69), demonstrating both positive and negative perceptions.

Prepandemic data from Garner and Holland's [8,56] studies provided a mean of 42.2, taken from 132 participants based in England, indicating a more positive perception of AFEs before the pandemic. Most participants scored above 125 (56%) for technology, demonstrating frequent use and familiarity with technology in general [53].

Next, we calculated the Pearson correlation coefficients to establish the relationships between loneliness, technology, LSM, and AFE perception to test H1 and H2. A correlation matrix of the variables was examined and is presented in Table 2.

Table 2. Correlational analysis between variables (N=92).

Variables	UCLA[a] Loneliness Scale	Life-space mobility	Age-friendliness of environment	Technology experience
UCLA Loneliness Scale				
r	1	−0.483[b]	−0.698[b]	−0.631[b]
P value	—	.002	.006	.003
Life-space mobility				
r	−0.483[b]	1	0.461[b]	0.430[b]
P value	.003	—	.003	.004
Age-friendliness of environment				
r	−0.698[b]	0.461[b]	1	0.667[b]
P value	.003	.004	—	.003
Technology experience				
r	−0.631[b]	0.430[b]	0.667[b]	1
P value	.003	.004	.002	—

[a]UCLA: University of California Los Angeles.

[b]*P*<.01.

[c]Not applicable.

LSM is Negatively Correlated With the Perception of AFE During the Pandemic (H1)

The correlation between LSM and AFE perception was statistically significant (r=0.461, P<.001) but positively correlated, which was contrary to the hypothesis (Table 2). This meant that higher LSM scores associated with restricted mobility were correlated with a greater positive perception of AFEs. Although this rejected H1, it was a notable result.

Loneliness Is Negatively Correlated to a Perception of AFEs (H2)

The correlation between loneliness and AFE perception was statistically significant (r=−0.698, P<.001) and negatively correlated. This meant that greater loneliness was correlated with more negative perceptions of AFEs, thus confirming H2.

Technology Use Moderates the Impact of LSM Restriction on the Perception of AFEs (H3)

Model 1 was used in the PROCESS 4.0 macro for SPSS to examine the moderation effect proposed in H3 [54], as shown in Figure 1, which shows the moderation role played by technology use in the relationship between LSM and AFE perception.

Here, all continuous variables were converted to z scores for use in the model as suggested by Frazier et al [57] and Hayes [54] (ie, via z scoring, expressed as the deviation from their sample means in SD units). The unconditional interaction of LSM and technology use was insignificant (β=0.1921, t_1=1.963; P=.06; Table 3).

All variables in the model are standardized and brought into the regression equation.

Using the Aiken and West [58] method, a simple slope test was used to analyze the conditional effect of technology use between LSM and AFEs (ie, whether technology use moderates the relationship between LSM and AFEs). As illustrated in Figure 2, when technology experience was high, LSM and AFEs were significantly positively correlated (β simple: mean 1, SD 2.2037, t=3.2216; P=.003), indicating that older adults' perception of their environment was more positive when they used technology more.

In contrast, the relationship between LSM and AFEs was not obvious when the technology experience was low (β simple: mean −1, SD −0.3429, t=−0.3510; P=.73). Thus, there appears to be a positive relationship between LSM restriction and perception of AFEs when technology use was high but not when it was low (explaining the unexpected direction of H1). Furthermore, note the slight downward slope indicating the negative impact of LSM restrictions on the perception of AFEs when technology experience was low. This hints that in the absence of technology use, LSM restrictions had a detrimental impact on the perception of AFEs. This confirms H3 and

XSL•FO
RenderX

explains the initial rejection of the unconditional (overall) moderation impact of technology experience on LSM by a small margin (P=.06).

Figure 1. Moderating role of technology experience on the relationship between life-space mobility and age-friendliness of environments.

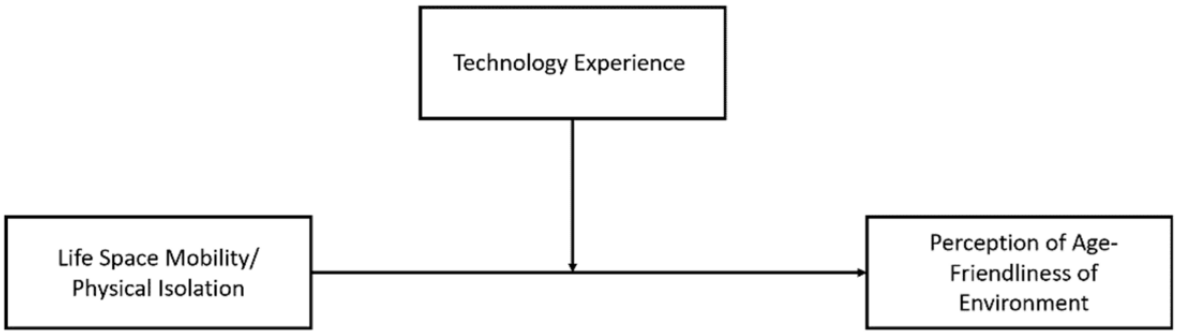
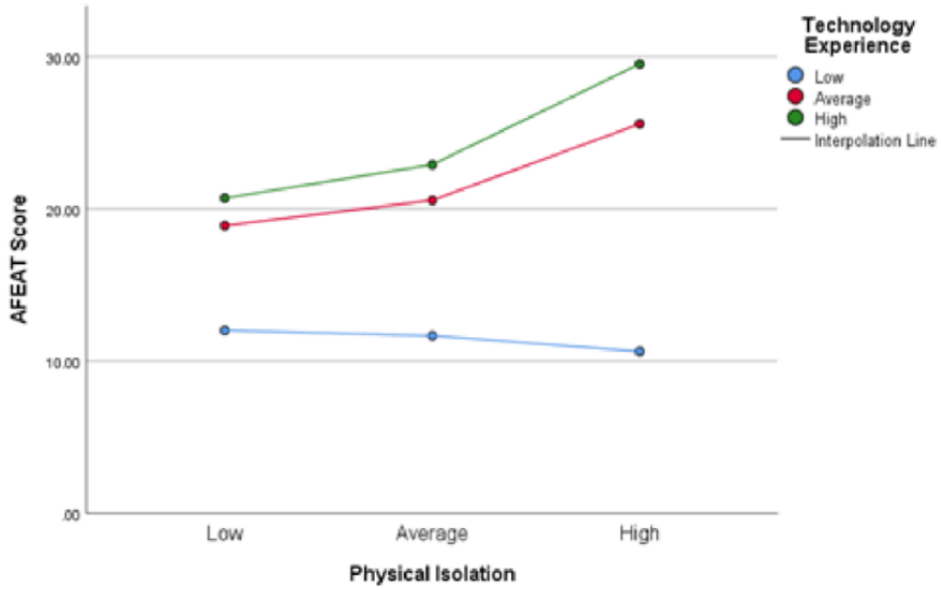


Table 3. Interaction between life-space mobility and technology experience.

	Overall fit indicators			Significance of standardized coefficient	
	R	R^2	*F* test (*df*)	β	*t* test (*df*)
Variables	0.71	0.50	29.7962 (1,90)		
Life-space mobility (ZLS1)				0.1787	2.1002[a] (90)
Technology experience (ZTE1)				0.6155	7.1839[a] (90)
ZLS1*ZTE1				0.1921	1.9628 (90)

[a]P<.001.

Figure 2. Moderating role of technology experience on the relationship between life-space mobility and age-friendly environment (simple slope test). AFEAT: age-friendly environment assessment tool.



Technology Use Moderates the Impact of Loneliness on the Perception of AFEs After Considering the Impact of LSM (H4)

Model 1 was used in the PROCESS 4.0 macro for SPSS to examine the moderation effect of technology use on the relationship between loneliness and perception of AFEs, as proposed in H4 [54] and shown in Figure 3, which shows the moderating effect of technology experience on the relationship between loneliness and the perception of AFEs while controlling for LSM.

All continuous variables were converted to *z* scores. As shown in Table 4, the relationship between loneliness and technology experience was significant (β=−0.3829, t_{90}=−5.1518; P<.001), but the impact of LSM when added to the regression model was not significant (β=0.7151, t_{90}=1.665; P=.07), showing that

XSL•FO
RenderX

although technology use had a moderating impact on the relationship between loneliness and AFE perception, the contribution of LSM was not significant on the model once these other variables had been taken into account.

All variables in the model are standardized and brought into the regression equation.

We then used the simple slope test to analyze the conditional effect of technology on the impact of loneliness on the perception of AFEs to further understand the impact. As evidenced in Figure 4, the link between loneliness and AFEs was not as obvious when technology was low (β simple: mean −1, SD −0.0819, t_1=−1.5394; *P*=.13). Conversely, when technology was high, the impact of moderation was apparent more clearly (β simple: mean 1, SD −0.4983, t_{90}=−7.2636; *P*<.001). Thus, we can conclude that technology moderates the impact of loneliness on the perception of AFEs when the technology experience is high, confirming H4.

Figure 3. Moderating effect of technology experience in the relationship between loneliness and age-friendliness of environments after controlling for life-space mobility.

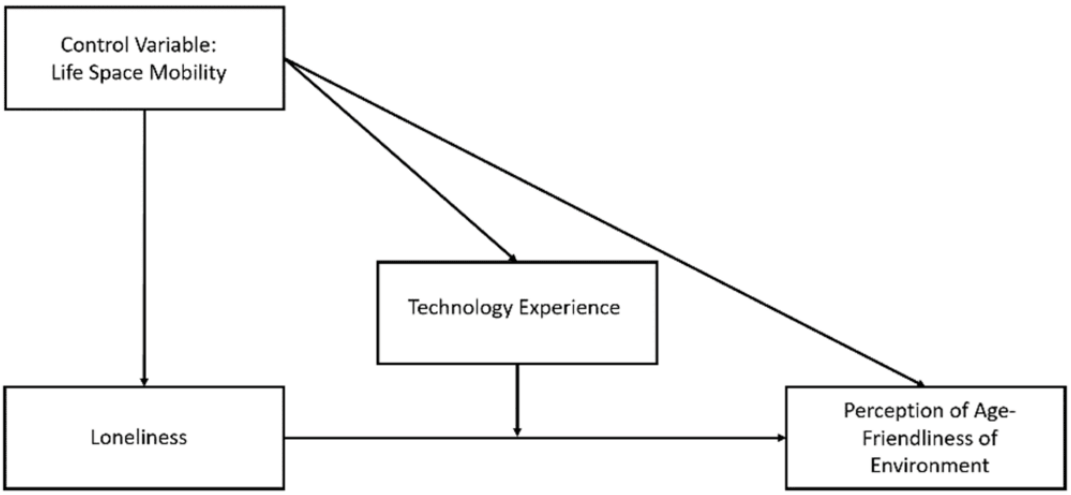


Table 4. The relationship between loneliness and technology experience and the impact of life-space mobility.

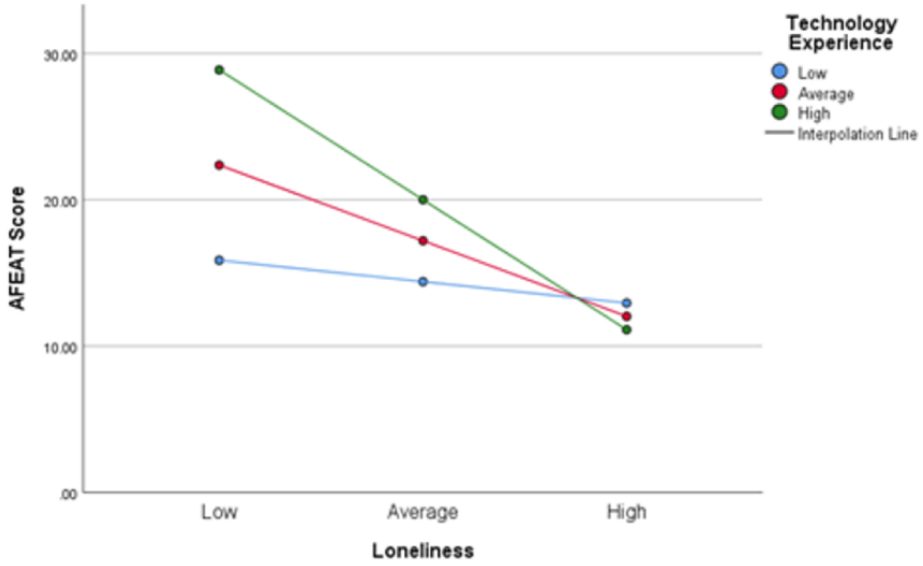
	Overall fit indicators			Significance of standardized coefficient	
	R	R^2	*F* test (*df*)	β	*t* test
Variables	0.82	0.68	45.8722 (1,90)		
Loneliness (ZL1)	—[a]	—	—	−0.5335	−6.2794[b]
Technology experience (ZTE1)	—	—	34.78 (1,90)	0.2888	3.5746[b]
ZL1*ZTE1	—	—	—	−0.3829	−5.1518[b]
Life-space mobility (ZLS1)	—	—	—	0.1177	1.665

[a]Not applicable.

[b]*P*<.001.

XSL•FO
RenderX

Figure 4. Moderating effect of technology experience in the relationship between loneliness and age-friendliness of environments after controlling for life-space mobility (simple slope test).



Discussion

Principal Findings

This study found higher levels of loneliness during the height of the COVID-19 pandemic among the older adult sample compared with the commonly available prepandemic data. For instance, Victor and Bowling [59] reported an average loneliness level of 30%, whereas Hawkley and Cacioppo [60] found the prevalence of loneliness in older adults to be approximately 25%, compared with the 44% found in this study. Numerous studies have found an increase in loneliness during the pandemic [61-63].

Higher loneliness was found to be correlated with a more negative perception of AFE, confirming H2 and in line with what has been theorized by the prevailing literature [18-20,22]. Shortfalls in emotional and social fulfillment have been highlighted as a predictor of loneliness, and the lack of opportunities for older adults to socialize, access important key services, and ability to engage with the community may have worsened perceptions. When older adults perceive that their desired quantity or quality of social engagements is met, they are less likely to experience loneliness and may also have favorable attitudes toward AFEs. Previous research has suggested that this relationship is likely to be complex [19] and bidirectional. AFE perception can be an indicator of a mechanism for aging-in-place, preserving older adults' physical setting of choice, providing a sense of community attachment, and allowing them to engage with a developed social network, with associated familiarity and better mental well-being.

This study found that greater use of technology clearly moderated the relationship between loneliness and the perception of AFEs. It also moderated the LSM-AFE relationship. Furthermore, the ability to use technology successfully to adapt to challenging experiences during lockdown emerged as a potential buffer against the impact of loneliness on AFE perception. Other studies that examined technology use during the pandemic revealed an increase, indicating adaptation to LSM restrictions via alternative pathways [37,64]. An increase in LSM restriction was linked to a positive perception of AFEs, rejecting H1, which at first seemed counterintuitive. However, the moderation effect of technology, such that this relationship existed only when technology use was high, explained the unexpected direction of the relationship. Older adults were overcoming physical restrictions barriers, where the ability to replace previously in-person activities with those online may have impacted perceptions. Other studies that examined AFEs also noted an increase in positive perception of AFEs as the pandemic progressed, but these could also have been linked to the easing of restrictions [56]. Overall, technology use may have improved the negative impacts of loneliness on the perception of AFEs helping with older adult mental health.

In testing H4, we determined that when technology experience was high, it had a moderation effect on the loneliness-AFE relationship. Ng et al [16] attempted to explain the relationship between internet use and loneliness through a moderation-mediation mechanism between internet use, perception of AFE, and loneliness, and depression. Their findings were consistent with those of Park et al [24] and Domènech-Abella et al [65], who confirmed the moderating effect of the internet on the age-friendliness-loneliness-depression mechanism, which may also explain our results. For example, Booth et al [66] discovered a partial mediation effect of feelings of helplessness, social isolation, depression, and distrust between psychological distress and perceived AFE (especially concerning security). Social isolation is a well-established predictor of loneliness and depression, and loneliness has also been linked in the literature. Taken together, these hint at a causal route between loneliness and perception of AFEs. Increased anxiety can lead to increased loneliness and mental health issues, potentially leading to a reduced fit between the older adult and the environment [67,68].

XSL•FO
RenderX

Despite studies highlighting the potential moderating role of technology between loneliness and AFE perception, previous results were always unclear about a direct link between these 3 variables. Our study found a clear moderating role for technology in the relationship between loneliness and AFEs, which is a notable finding.

When using the simple slope test, where technology use was low, we found that LSM negatively affected perceived AFEs when technology use was low. This explained our initial counterintuitive result and supported earlier findings [24]. To substitute or overcome confinement, older adults may have developed alternative routes of access to AFE domains, such as social, civic engagement, and access to services through technology.

Access to the internet may have helped reduce older people's boredom, and there have been examples from recent studies where older adults were able to access informal help networks, reading, or online game groups, as well as participate in community-based activities like attending virtual church gatherings [69,70]. Alternative social and emotional outlets to combat loneliness through access via the internet, to previously in-person services (eg, primary care, counseling) may have contributed to network socialization, allowing older adults to continue feeling like they are a part of their environment and reducing loneliness.

The implications of our study can be applied to situations outside of the pandemic context. Studies have highlighted that older adults prefer to live in their own homes and interact with their local community, where they have developed relationships over time and do so as long as possible [71,72]. A negative perception of AFEs can be viewed as a barometer indicating poor person-environment fit [72], associated with poorer mental health outcomes [25,71]. Therefore, technology could be a solution for those who ordinarily have limited LSM, are at risk for social isolation, and have a negative perception of AFEs.

Our study confirmed the link between LSM, loneliness, and perception of AFEs, as well as the moderating role technology played during the COVID-19 pandemic, advancing the findings of previous studies in this rapidly evolving body of literature. We also strengthened the argument that the WHO's Global Age-Friendly Cities framework would benefit older adults more by including technology use as an additional ninth domain.

Study Limitations

This study had several limitations. For example, the sampling could have been predisposed to participants literate in digital resources and more socially connected. Generally, such participants may have experienced less loneliness [73]. Furthermore, a cross-sectional design cannot establish causality [74]. Although the sample size was small, the statistical power, effect size, precision, type, complexity of analysis, study population variability, and homogeneity overcame this shortcoming. The results are an important contribution to the discourse on the role being played by AFEs in the mental well-being of older adults and the role of technology.

Although it is unlikely that participants will experience the same level of LSM restrictions after the pandemic, this allowed us to examine the studied measures in a normally inaccessible environment. However, we cannot conclude with certainty that the pandemic caused the observations because we did not have a prepandemic assessment for the same participants. Nevertheless, other prepandemic studies supported our hypothesis of mobility restrictions' impact on loneliness.

Conclusion

Despite the limitations, the results suggest that the vulnerability of older adults during the pandemic and their exposure to loneliness and negative perceptions of AFEs increased, and technology played an important role in moderating these influences. Given the advancements in technology, the WHO's 8 domains of AFEs may be obsolete with their narrow implicit recognition of technology. Our study demonstrates the significance of an explicit recognition of technology in the evaluation of AFEs as an integral component of all aspects of older adults' daily lives. Future studies may also wish to look further into the impact of demographics and differences between genders.

Community and mental health service access could be improved by providing online access. Older adults need cheap access to internet infrastructure, and health community centers should provide technology training, attend online meetings, or use mental health applications.

Researchers in the field of loneliness in older adults are encouraged to use our results to inform initiatives to reduce the mental health risks for older adults in vulnerable crises, such as the pandemic and civil insurrections [75,76]. Appropriate consideration of these factors will aid decision makers in developing robust and effective strategies during times of crisis, as well as in assisting an aging population with aging-in-place.

Conflicts of Interest

None declared.

Multimedia Appendix 1
STROBE Checklist.
[DOCX File , 35 KB - aging_v8i1e67242_app1.docx]

References

XSL•FO
RenderX

1. Global age-friendly cities: a guide. World Health Organization. 2007. URL: https://www.who.int/publications/i/item/9789241547307 [accessed 2007-10-05]
2. Phillipson C, Baars J. Social theory and social ageing. Ageing Soc 2007;22:68-84. [doi: 10.4135/9781446278918.n4]
3. Phillipson C. Developing age-friendly urban communities: critical issues for public policy. Public Policy Aging Rep 2014 Dec 27;25(1):4-8. [doi: 10.1093/ppar/pru052]
4. Buffel T, Phillipson C, scharf T. Ageing in urban environments: developing 'age-friendly' cities. Crit Soc Policy 2012 May 22;32(4):597-617. [doi: 10.1177/0261018311430457]
5. Phillipson C. Developing Age-Friendly Communities: New Approaches to Growing Old in Urban Environments. Berlin, Heidelberg: Springer; 2010.
6. Cerin E, Nathan A, van Cauwenberg J, Barnett DW, Barnett A, Council on EnvironmentPhysical Activity (CEPA) – Older Adults working group. The neighbourhood physical environment and active travel in older adults: A systematic review and meta-analysis. Int J Behav Nutr Phys Act 2017;14(1):15 [FREE Full text] [doi: 10.1186/s12966-017-0471-5] [Medline: 28166790]
7. Young AF, Russell A, Powers JR. The sense of belonging to a neighbourhood: can it be measured and is it related to health and well being in older women? Soc Sci Med 2004;59(12):2627-2637. [doi: 10.1016/j.socscimed.2004.05.001] [Medline: 15474215]
8. Garner IW, Holland CA. Age-friendliness of living environments from the older person's viewpoint: development of the age-friendly environment assessment tool. Age Ageing 2020;49(2):193-198. [doi: 10.1093/ageing/afz146] [Medline: 31790132]
9. Balki E, Hayes N, Holland C. The impact of social isolation, loneliness, and technology use during the COVID-19 pandemic on health-related quality of life: observational cross-sectional study. J Med Internet Res 2022;24(10):e41536 [FREE Full text] [doi: 10.2196/41536] [Medline: 36260401]
10. Mahmood A, Seetharaman K, Jenkins HT, Chaudhury H. Contextualizing innovative housing models and services within the age-friendly communities framework. Gerontologist 2022;62(1):66-74. [doi: 10.1093/geront/gnab115] [Medline: 34355769]
11. Hsu HC. Associations of city-level active aging and age friendliness with well-being among older adults aged 55 and over in Taiwan. Int J Environ Res Public Health 2020;17(12):4304 [FREE Full text] [doi: 10.3390/ijerph17124304] [Medline: 32560170]
12. Berg-Weger M, Morley JE. Editorial: Loneliness and social isolation in older adults during the COVID-19 pandemic: implications for gerontological social work. J Nutr Health Aging 2020;24(5):456-458 [FREE Full text] [doi: 10.1007/s12603-020-1366-8] [Medline: 32346678]
13. Hwang TJ, Rabheru K, Peisah C, Reichman W, Ikeda M. Loneliness and social isolation during the COVID-19 pandemic. International Psychogeriatrics 2020;32(10):1217-1220. [doi: 10.1017/s1041610220000988]
14. Au A, Lai DWL, Yip H, Chan S, Lai S, Chaudhury H, et al. Sense of community mediating between age-friendly characteristics and life satisfaction of community-dwelling older adults. Front Psychol 2020;11:86 [FREE Full text] [doi: 10.3389/fpsyg.2020.00086] [Medline: 32194465]
15. Aboderin I, Kano M, Owii HA. Toward "Age-Friendly Slums"? Health challenges of older slum dwellers in nairobi and the applicability of the age-friendly city approach. Int J Environ Res Public Health 2017;14(10):1259 [FREE Full text] [doi: 10.3390/ijerph14101259] [Medline: 29053576]
16. Ng SI, Lim XJ, Hsu HC, Chou CC. Age-friendliness of city, loneliness and depression moderated by internet use during the COVID-19 pandemic. Health Promot Int 2023;38(3) [FREE Full text] [doi: 10.1093/heapro/daac040] [Medline: 35437585]
17. Perlman D, Peplau LA. In: Gilmour R, Duck S, editors. Toward a social psychology of loneliness. In Personal Relationships: 3. Relationships in Disorder. Cambridge, Massachusetts, USA: Academic Press; 1981:56.
18. Dinkins CS. Seeing oneself in the face of the other: the value and challenge of human connectedness for older adults. J Psychosoc Nurs Ment Health Serv 2017;55(7):13-17. [doi: 10.3928/02793695-20170619-03] [Medline: 28671237]
19. Kearns A, Whitley E, Tannahill C, Ellaway A. 'Lonesome Town'? Is loneliness associated with the residential environment, including housing and neighbourhood factors? J Community Psychol 2015;43(7):849-867 [FREE Full text] [doi: 10.1002/jcop.21711] [Medline: 26740728]
20. Lee MS, Park J. Television shopping at home to alleviate loneliness among older consumers. amj 2017;18(4):139. [doi: 10.15830/amj.2017.18.4.139]
21. Gibney S, Zhang M, Brennan C. Age-friendly environments and psychosocial wellbeing: a study of older urban residents in Ireland. Aging Ment Health 2020;24(12):2022-2033. [doi: 10.1080/13607863.2019.1652246] [Medline: 31402677]
22. Rantakokko M, Iwarsson S, Vahaluoto S, Portegijs E, Viljanen A, Rantanen T. Perceived environmental barriers to outdoor mobility and feelings of loneliness among community-dwelling older people. J Gerontol A Biol Sci Med Sci 2014;69(12):1562-1568. [doi: 10.1093/gerona/glu069] [Medline: 24864307]
23. Stafford PB. The Global Age-Friendly Community Movement: A Critical Appraisal, 5. New York: Berghahn Books; 2018.
24. Park S, Kim M, Chung S. Age-friendly environments and depressive symptoms among Korean adults: the mediating effects of loneliness. Aging Ment Health 2021;25(6):1060-1070. [doi: 10.1080/13607863.2020.1755827] [Medline: 32321293]

XSL•FO
RenderX

25. Gaugler JE, Duval S, Anderson KA, Kane RL. Predicting nursing home admission in the U.S: a meta-analysis. BMC Geriatr 2007;7(1):13 [FREE Full text] [doi: 10.1186/1471-2318-7-13] [Medline: 17578574]
26. Stalvey BT, Owsley C, Sloane ME, Ball K. The life space questionnaire: a measure of the extent of mobility of older adults. J Appl Gerontol 1999;18(4):460-478. [doi: 10.1177/073346489901800404]
27. Xue QL, Fried LP, Glass TA, Laffan A, Chaves PHM. Life-space constriction, development of frailty, and the competing risk of mortality: the Women's Health And Aging Study I. Am J Epidemiol 2008;167(2):240-248. [doi: 10.1093/aje/kwm270] [Medline: 17906296]
28. Sachs J. Sustainable Development Report 2020: the Sustainable Development Goals and Covid-19 Includes the SDG Index and Dashboards. Cambridge: Cambridge University Press; 2021.
29. Kotwal AA, Holt-Lunstad J, Newmark RL, Cenzer I, Smith AK, Covinsky KE, et al. Social isolation and loneliness among san francisco bay area older adults during the COVID-19 shelter-in-place orders. J Am Geriatr Soc 2021;69(1):20-29 [FREE Full text] [doi: 10.1111/jgs.16865] [Medline: 32965024]
30. Emlet CA, Moceri JT. The importance of social connectedness in building age-friendly communities. J Aging Res 2012;2012:1-9 [FREE Full text] [doi: 10.1155/2012/173247] [Medline: 22162807]
31. Lui CW, Everingham JA, Warburton J, Cuthill M, Bartlett H. What makes a community age-friendly: A review of international literature. Australas J Ageing 2009;28(3):116-121. [doi: 10.1111/j.1741-6612.2009.00355.x] [Medline: 19845650]
32. Peel C, Sawyer Baker P, Roth DL, Brown CJ, Brodner EV, Allman RM. Assessing mobility in older adults: the UAB study of aging life-space assessment. Phys Ther 2005;85(10):1008-1119. [Medline: 16180950]
33. Theis KA, Furner SE. Shut-In? Impact of chronic conditions on community participation restriction among older adults. J Aging Res 2011;2011 [FREE Full text] [doi: 10.4061/2011/759158] [Medline: 21837277]
34. Rantakokko M, Iwarsson S, Kauppinen M, Leinonen R, Heikkinen E, Rantanen T. Quality of life and barriers in the urban outdoor environment in old age. J Am Geriatr Soc 2010;58(11):2154-2159. [doi: 10.1111/j.1532-5415.2010.03143.x] [Medline: 21054297]
35. Coyle CE, Dugan E. Social isolation, loneliness and health among older adults. J Aging Health 2012;24(8):1346-1363. [doi: 10.1177/0898264312460275]
36. Groarke JM, Berry E, Graham-Wisener L, McKenna-Plumley PE, McGlinchey E, Armour C. Loneliness in the UK during the COVID-19 pandemic: Cross-sectional results from the COVID-19 psychological wellbeing study. PLoS One 2020;15(9):e0239698 [FREE Full text] [doi: 10.1371/journal.pone.0239698] [Medline: 32970764]
37. Moore RC, Hancock JT. Older adults, social technologies, and the coronavirus pandemic: challenges, strengths, and strategies for support. Soc Media Soc 2020 Aug 11;6(3). [doi: 10.1177/2056305120948162]
38. Shah SGS, Nogueras D, van Woerden HC, Kiparoglou V. Evaluation of the effectiveness of digital technology interventions to reduce loneliness in older adults: systematic review and meta-analysis. J Med Internet Res 2021;23(6):e24712 [FREE Full text] [doi: 10.2196/24712] [Medline: 34085942]
39. Dowson B, Atkinson R, Barnes J, Barone C, Cutts N, Donnebaum E, et al. Digital approaches to music-making for people with dementia in response to the COVID-19 pandemic: current practice and recommendations. Front Psychol 2021;12:625258 [FREE Full text] [doi: 10.3389/fpsyg.2021.625258] [Medline: 33967893]
40. Feldman PH, Oberlink MR. Developing community indicators to promote the health and well-being of older people. Fam Community Health 2003;26(4):268-274. [doi: 10.1097/00003727-200310000-00004] [Medline: 14528133]
41. Lawton MP. Environment and other determinants of well-being in older people. Gerontologist 1983;23(4):349-357. [doi: 10.1093/geront/23.4.349] [Medline: 6352420]
42. Gerst-Emerson K, Jayawardhana J. Loneliness as a public health issue: the impact of loneliness on health care utilization among older adults. Am J Public Health 2015;105(5):1013-1019. [doi: 10.2105/ajph.2014.302427]
43. Luciano A, Pascale F, Polverino F, Pooley A. Measuring age-friendly housing: a framework. Sustainability 2020;12(3):848. [doi: 10.3390/su12030848]
44. Marston HR, van Hoof J. "Who Doesn't think about technology when designing urban environments for Older People?" a case study approach to a proposed extension of the WHO's age-friendly cities model. Int J Environ Res Public Health 2019;16(19):3525 [FREE Full text] [doi: 10.3390/ijerph16193525] [Medline: 31547204]
45. Sánchez-González D, Rojo-Pérez F, Rodríguez-Rodríguez V, Fernández-Mayoralas G. Environmental and psychosocial interventions in age-friendly communities and active ageing: A systematic review. Int J Environ Res Public Health 2020;17(22):8305 [FREE Full text] [doi: 10.3390/ijerph17228305] [Medline: 33182710]
46. Balki E, Hayes N, Holland C. Effectiveness of technology interventions in addressing social isolation, connectedness, and loneliness in older adults: Systematic umbrella review. JMIR Aging 2022;5(4):e40125 [FREE Full text] [doi: 10.2196/40125] [Medline: 36279155]
47. Belam G. Yoga as an intervention for older peoples mental health: a literature review. WWOP 2020;24(3):159-169. [doi: 10.1108/wwop-05-2020-0017]
48. Pedell S, Borda A, Keirnan A, Aimers N. Combining the Digital, Social and Physical Layer to Create Age-Friendly Cities and Communities. Int J Environ Res Public Health 2021;18(1):325 [FREE Full text] [doi: 10.3390/ijerph18010325] [Medline: 33466259]

XSL•FO
RenderX

49. Reuter A, Liddle J, Scharf T. Digitalising the age-friendly city: Insights from participatory action research. Int J Environ Res Public Health 2020;17(21):8281 [FREE Full text] [doi: 10.3390/ijerph17218281] [Medline: 33182453]
50. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med 2007;4(10):e296 [FREE Full text] [doi: 10.1371/journal.pmed.0040296] [Medline: 17941714]
51. Allen M. The SAGE Encyclopedia of Communication Research Methods. Washington DC: SAGE; 2017.
52. Russell DW. UCLA Loneliness Scale (Version 3): reliability, validity, and factor structure. J Pers Assess 1996;66(1):20-40. [doi: 10.1207/s15327752jpa6601_2] [Medline: 8576833]
53. Czaja SJ, Charness N, Fisk AD, Hertzog C, Nair SN, Rogers WA, et al. Factors predicting the use of technology: findings from the center for research and education on aging and technology enhancement (CREATE). Psychol Aging 2006;21(2):333-352 [FREE Full text] [doi: 10.1037/0882-7974.21.2.333] [Medline: 16768579]
54. Hayes AF. Partial, conditional, and moderated mediation: quantification, inference, and interpretation. Commun Monogr 2018;85:4-40.
55. Mullins LC. Encyclopedia of Gerontology, 2nd ed. Berlin, Heidelberg,: Springer; 2007.
56. Garner IW, Varey S, Navarro-Pardo E, Marr C, Holland CA. An observational cohort study of longitudinal impacts on frailty and well-being of COVID-19 lockdowns in older adults in England and spain. Health Soc Care Community 2022;30(5):e2905-e2916. [doi: 10.1111/hsc.13735] [Medline: 35089638]
57. Frazier PA, Tix AP, Barron KE. Testing moderator and mediator effects in counseling psychology research. J Counsel Psychol 2004 Jan;51(1):115-134. [doi: 10.1037/0022-0167.51.1.115]
58. Aiken LS, West SG. Multiple Regression: Testing and Interpreting Interactions. Washington DC: SAGE; 1991.
59. Victor CR, Bowling A. A longitudinal analysis of loneliness among older people in Great Britain. J Psychol 2012;146(3):313-331. [doi: 10.1080/00223980.2011.609572] [Medline: 22574423]
60. Hawkley LC, Cacioppo JT. Loneliness matters: a theoretical and empirical review of consequences and mechanisms. Ann Behav Med 2010;40(2):218-227 [FREE Full text] [doi: 10.1007/s12160-010-9210-8] [Medline: 20652462]
61. Bu F, Steptoe A, Fancourt D. Loneliness during a strict lockdown: Trajectories and predictors during the COVID-19 pandemic in 38,217 United Kingdom adults. Soc Sci Med 2020;265:113521 [FREE Full text] [doi: 10.1016/j.socscimed.2020.113521] [Medline: 33257177]
62. Elran-Barak R, Mozeikov M. One month into the reinforcement of social distancing due to the COVID-19 outbreak: subjective health, health Behaviors, and Loneliness among people with chronic medical conditions. Int J Environ Res Public Health 2020;17(15):5403 [FREE Full text] [doi: 10.3390/ijerph17155403] [Medline: 32727103]
63. Krendl AC, Perry BL. The impact of sheltering in place during the COVID-19 pandemic on older adults' social and mental well-being. J Gerontol B Psychol Sci Soc Sci 2021;76(2):e53-e58 [FREE Full text] [doi: 10.1093/geronb/gbaa110] [Medline: 32778899]
64. Balki E, Hayes N, Holland C. Loneliness and older adults: psychological resilience and technology use during the COVID-19 pandemic-a cross sectional study. Front Aging 2023;4:1184386 [FREE Full text] [doi: 10.3389/fragi.2023.1184386] [Medline: 37434741]
65. Domènech-Abella J, Switsers L, Mundó J, Dierckx E, Dury S, De Donder L. The association between perceived social and physical environment and mental health among older adults: mediating effects of loneliness. Aging Ment Health 2021;25(5):962-968. [doi: 10.1080/13607863.2020.1727853] [Medline: 32067471]
66. Booth J, Ayers SL, Marsiglia FF. Perceived neighborhood safety and psychological distress: exploring protective factors. The Journal of Sociology & Social Welfare 2012;39(4):137. [doi: 10.15453/0191-5096.3703]
67. Balki E, Holland C, Hayes N. A mixed methods exploration of older adult use and acceptance of digital communication technology for social connectedness during the Covid-19 pandemic,. J Med Internet Res 2023;25:e41535 [FREE Full text] [doi: 10.2196/41535] [Medline: 37531187]
68. Balki E, Hayes N, Holland C. The indirect impact of educational attainment as a distal resource for older adults on loneliness, social isolation, psychological resilience, and technology use during the COVID-19 pandemic: cross-sectional quantitative study. JMIR Aging 2023;6:e47729 [FREE Full text] [doi: 10.2196/47729] [Medline: 37999938]
69. Cohen-Mansfield J, Muff A, Meschiany G, Lev-Ari S. Adequacy of web-based activities as a substitute for in-person activities for older persons during the COVID-19 pandemic: survey study. J Med Internet Res 2021 Jan 22;23(1):e25848 [FREE Full text] [doi: 10.2196/25848] [Medline: 33439851]
70. Gibney S, Moore T, Shannon S. Loneliness in later life: a cross-sectional survey analysis of place-based factors in Ireland. QAOA 2019;20(2):80-96. [doi: 10.1108/qaoa-04-2018-0015]
71. Greenfield EA. Support from neighbors and aging in place: can NORC programs make a difference? Gerontologist 2016;56(4):651-659. [doi: 10.1093/geront/gnu162] [Medline: 26035886]
72. Park H, Park S, Shephard RJ, Aoyagi Y. Yearlong physical activity and sarcopenia in older adults: the nakanojo study. Eur J Appl Physiol 2010;109(5):953-961. [doi: 10.1007/s00421-010-1424-8] [Medline: 20336310]
73. Chronic PSO. Interventions to combat loneliness for older adults. A review of reviews 2018;964((Suppl. 1)). [doi: 10.1093/geroni/igy031.3571]

XSL•FO
RenderX

74. Sygit-Kowalkowska E, Szrajda J, Weber-Rajek M, Porażyński K, Ziółkowski M. Resilience as a predicator of mental health of incarcerated women. Psychiatr Pol 2016;51(3):549-560. [doi: 10.12740/pp/onlinefirst/62617]
75. Abdulkerim S, Albari M, MacKenzie A, Thurston A. Title: report on a survey of mobility disability among Syrian refugees living in Turkey. Int J Educational Res Open 2021;2:100059. [doi: 10.1016/j.ijedro.2021.100059]
76. Fund KS. Access to Health Care and Minority Ethnic Groups. (King?s Fund). 2006. URL: https://assets.kingsfund.org.uk/f/256914/x/9e3f53c8f1/access_health_care_minority_ethnic_groups_2006.pdf [accessed 2025-04-03]

Abbreviations

AFE: age-friendly environment
AFEAT: Age-Friendly Environment Assessment Tool
H: hypothesis
LSM: life-space mobility
QOL: quality of life
SAGE: Scientific Advisory Group for Emergencies
STROBE: Strengthening Reporting of Observational Studies in Epidemiology
WHO: World Health Organization

Edited by A Quialheiro; submitted 06.10.24; peer-reviewed by C Yang, C Franceschetti; comments to author 29.01.25; revised version received 25.02.25; accepted 02.04.25; published 06.05.25.

Please cite as:
Balki E, Hayes N, Holland C
Factors Influencing Older Adults' Perception of the Age-Friendliness of Their Environment and the Impact of Loneliness, Technology Use, and Mobility: Quantitative Analysis
JMIR Aging 2025;8:e67242
URL: https://aging.jmir.org/2025/1/e67242
doi:10.2196/67242
PMID:

©Eric Balki, Niall Hayes, Carol Holland. Originally published in JMIR Aging (https://aging.jmir.org), 06.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Buffering Effects of Internet Use on Caregiving-Related Health Impacts and Loneliness Among Older Informal Caregivers in California: Cross-Sectional Study

Xiang Qi[1], BSN, PhD; Ruotong Liu[1], PhD; Eunjung Ko[1], PhD; Yaolin Pei[2], PhD; Bei Wu[1], PhD

[1]Rory Meyers College of Nursing, New York University, 433 1st Ave, New York, NY, United States

[2]School of Nursing, The University of Texas at Austin, Austin, TX, United States

Corresponding Author:
Xiang Qi, BSN, PhD
Rory Meyers College of Nursing, New York University, 433 1st Ave, New York, NY, United States

Abstract

Background: Loneliness has emerged as a global public health issue, with recent data indicating that 27.6% of adults aged 65 to 80 report feelings of loneliness despite the postpandemic resumption of social activities. Older caregivers face unique challenges that may exacerbate feelings of loneliness due to the demanding nature of caregiving responsibilities. While internet use has been suggested as a potential intervention to reduce loneliness, its moderating effect on the relationship between caregiving-related health effects and loneliness remains understudied.

Objective: This study aims to investigate: (1) the association between caregiving-related health effects and loneliness among older informal caregivers; (2) the relationship between internet use frequency and loneliness; and (3) whether internet use moderates the association between caregiving-related health effects and loneliness.

Methods: We analyzed cross-sectional data from the 2019‐2020 California Health Interview Survey, focusing on 3957 informal caregivers aged 65 and older. Loneliness was measured using a modified 3-item UCLA Loneliness Scale. Health effects of caregiving were assessed by self-reported physical or mental health problems due to caregiving responsibilities. Internet use frequency was measured on a 4-point scale. Multivariable linear regressions were used to test the study aims, adjusting for sociodemographic factors, health status, and caregiving-context characteristics.

Results: Among participants, 475 (12.0%) reported experiencing physical or mental health problems due to caregiving responsibilities. After adjusting for covariates, caregivers who experienced health problems related to caregiving reported higher levels of loneliness compared to those who did not (β=0.76, SE .07, *P*<.001). More frequent internet use was associated with a lower level of loneliness (β=−0.11, SE 0.03, *P*<.001). Additionally, internet use significantly moderated the relationship between caregiving-related health effects and loneliness (β=−.16, SE 0.07, *P*=.02), suggesting that the negative impact of caregiving-related health effects on loneliness was attenuated among caregivers who used the internet more frequently.

Conclusions: Caregiving-related health effects are associated with increased loneliness among older informal caregivers, but more frequent internet use may both directly reduce loneliness and buffer against the adverse impact of caregiving on loneliness. These findings align with recent research highlighting the potential of technology-based interventions to combat social disconnection among older adults. Health care providers and policy makers should consider implementing programs that enhance internet access among older caregivers as part of comprehensive strategies to address loneliness in this vulnerable population.

(JMIR Aging 2025;8:e74209) doi:10.2196/74209

KEYWORDS

social isolation; social disconnection; family caregiving; information and communication technology; digital technology; stress-process model

Introduction

Loneliness in older adults has emerged as a growing global public health concern, with recent data indicating that 27.6% of adults aged 65 to 80 continue to experience feelings of loneliness despite the postpandemic resumption of social activities [1]. Loneliness is associated with multiple adverse health outcomes, including depression, heart disease, cognitive decline, sleep disturbance, and premature mortality [2-5]. The US Surgeon General's 2023 Advisory on social isolation and loneliness equated the health impact of lacking social connection to smoking 15 cigarettes daily, underscoring the severity of this issue [6].

The significance of this issue is magnified for older informal caregivers, who face unique challenges and heightened risks of

XSL•FO
RenderX

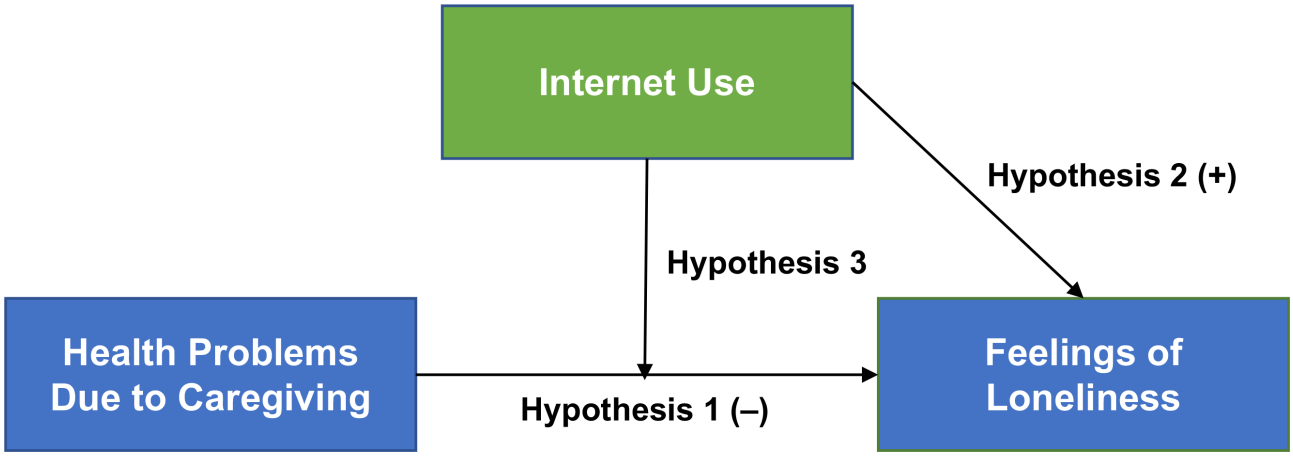
loneliness due to the demands and stress associated with caregiving responsibilities [7,8]. A recent analysis revealed that 15% of older caregivers experienced loneliness, with caregivers of persons with dementia having 1.62 times higher odds of experiencing loneliness compared to other caregivers [9]. For older informal caregivers, their advanced age can intensify their vulnerability to adverse health effects stemming from the burden of caregiving. In addition, older informal caregivers, who provide unpaid care to their family members or friends, often experience isolation due to the restrictive nature of caregiving duties that limit social interactions [7,8]. This isolation can exacerbate those adverse health effects associated with caregiving, creating a cycle of declining mental and physical health [10]. Given these challenges, identifying accessible interventions to alleviate loneliness in older caregivers is crucial. Increasingly, internet use has been considered a potential means to maintain social connections for those with limited in-person interaction opportunities.

While research on the general population has shown mixed or even negative associations between the widespread adoption of the internet and digital communication technologies and loneliness [11-13], the unique circumstances of informal caregivers may create a different context for digital engagement [14,15]. Caregivers face specific mobility constraints and social limitations that may make internet-based connections particularly valuable. For older informal caregivers, the internet may offer dual benefits: serving as a source of caregiving information and emotional support while facilitating social connections that reduce feelings of loneliness [16]. Recent evidence suggests that internet use may buffer against the negative health impacts of loneliness for older informal caregivers [17], enhance life satisfaction through improved interpersonal connections [18], and provide access to specialized support communities [15]. A recent scoping review indicated that technology-based interventions have shown preliminary evidence for reducing social isolation and loneliness among informal caregivers of persons with dementia [15]. However, despite the potential of internet use to serve as a resource for this population, research examining the intersection of caregiving, internet use, and loneliness within this population remains limited. This study aims to fill these gaps by examining the associations between caregiving-related health effects, internet use, and feelings of loneliness among older informal caregivers in California. Furthermore, it investigates whether internet use moderates the relationship between caregiving-related health effects and loneliness.

Using the Stress Process Model as a theoretical framework [19], this study considers how internet use may function as a coping resource that buffers against the adverse impact of caregiving stressors on feelings of loneliness. The Stress Process Model posits that caregiving experiences involve primary stressors (direct pressures of caregiving) and secondary stressors (indirect pressures such as financial strain), which can contribute to negative health outcomes, including loneliness [19]. In this study, we interpret caregiving-related health problems as a secondary stressor within the Stress Process Model framework, representing the negative consequences of primary caregiving stressors on caregivers' well-being. Within this framework, internet use may serve as a coping resource that provides both informational and emotional support, potentially mitigating the impact of caregiving-related stress on loneliness.

This study tests three hypotheses: (1) experiencing adverse health effects from caregiving is associated with a higher level of loneliness; (2) frequent internet use is associated with a lower level of loneliness; and (3) internet use protects against the adverse caregiving-related health effects on feelings of loneliness among older informal caregivers. The conceptual framework with three hypotheses is shown in Figure 1.

Figure 1. Conceptual framework.



Methods

Data Source and Sample Selection

This study used cross-sectional data from the 2019 - 2020 California Health Interview Survey (CHIS), which is the largest statewide health survey in the United States [20]. CHIS uses a complex, multistage sampling design to collect comprehensive health-related information from California's community-dwelling population [20,21]. The survey is conducted either online or by telephone and collects data on sociodemographic characteristics, health-related behaviors, health status, health care access, and utilization patterns. The 2019 - 2020 CHIS surveyed 44,109 adults with a screened response rate of 72.0%, producing a sample that is representative of California's diverse population. Notably, CHIS is conducted

XSL•FO
RenderX

in multiple languages and uses sampling strategies to reach various racial and ethnic and socioeconomic groups, ensuring broad representativeness of the state's population [20].

From the full CHIS sample, we restricted our analytic sample to respondents who met the following criteria: (1) aged 65 years or older; (2) reported providing unpaid help to family members or friends with illnesses or disabilities in the past year; and (3) had complete information on the key study variables (internet use, health effects of caregiving, and loneliness). After applying these criteria, our final analytic sample consisted of 3957 older informal caregivers.

Measurement

Dependent Variable: Loneliness

Loneliness was assessed using the modified version of the validated 3-item UCLA Loneliness Scale [22], demonstrating good reliability and validity in older adult populations [4,23]. Participants responded to three questions: "How much of the time do you feel (1) lack of companionship, (2) left out, and (3) isolated from others?" Each item was rated on a 3-point scale: "1=hardly ever or never," "2=some of the time," and "3=often." [22] Responses were summed to create a continuous loneliness score ranging from 3 to 9, with higher scores indicating greater levels of loneliness. The internal consistency of the 3-item UCLA Loneliness Scale in our sample was excellent (Cronbach α=0.974).

Independent Variables

Health effects of caregiving: The health impact of caregiving was measured by asking respondents "whether you had experienced physical or mental health problems as a result of their caregiving responsibilities in the past 12 months." Responses were dichotomized (yes/no), with "yes" indicating the presence of caregiving-related health problems.

Frequency of internet use: Internet use was measured by asking participants the question: "how often you used the internet?" Responses were coded on a 4-point scale: "1=less than a few times a day," "2=few times a day," "3=many times a day," and "4=almost constantly." Following previous research [24], we treated internet use frequency as a continuous variable, with higher scores indicating more frequent use of the internet.

Covariates

We included several potential covariates based on established associations with caregiving, internet, and mental health among older adults [25-28]. Sociodemographic covariates included age (continuous, in years), sex (0=male; 1=female), race/ethnicity (1=non-Hispanic White; 2=non-Hispanic Black; 3=Hispanic or Latino; 4=non-Hispanic Asian; and 5=others), U.S. born (0=not born in the U.S.; 1=born in the US), education level (0=less than high school; 1=high school or above), economic status measured by family income relative to the federal poverty level (0=<200%; 1=≥200%), and marital status (0=married/living with partner; 1=widowed/separated/divorced/never married). We also controlled for health status, including self-rated health (measured on a 5-point Likert scale from 1=poor to 5=excellent) and the number of self-reported chronic conditions (ie, hypertension, asthma, diabetes, heart disease, vision/hearing impairment). Additionally, we included caregiving context variables that might influence the relationship between our primary variables of interest: experiencing financial stress due to caregiving (measured on a 4-point Likert scale from 1=extremely stressful to 4=not at all stressful), caregiving hours per week (ranged 1 to 125), the care recipient's living arrangement (1=living alone; 2=living with caregiver or other family members; and 3=living in a nursing home/assisted-living facility/other living situations), relationship to the care recipient (1=spouse/partner; 2=parent/parent-in-law; 3=other relatives; and 4=nonrelative), and whether the caregiver supported the care recipient with Alzheimer disease and related dementia (0=nondementia caregiver; 1=dementia caregiver).

Statistical Analysis

We first conducted descriptive analyses of all study variables, presenting continuous variables as means (SD) and categorical variables as frequencies and percentages (%). To visually examine the relationship between internet use frequency and loneliness, we created violin plots displaying the distribution of UCLA Loneliness Scale scores across the four internet use frequency categories. We also performed one-way ANOVA tests with post hoc Tukey Honestly Significant Difference (HSD) comparisons to assess differences in mean loneliness scores across internet use frequency groups.

To test our hypotheses, we used a series of multivariable linear regression models with loneliness as the dependent variable. After adjustment for all covariates, Model 1 examined the association between caregiving-related health problems and loneliness. Model 2 examined the association between internet use and loneliness. Model 3 examined the associations of both caregiving-related health problems and internet use with loneliness. Model 4 incorporated an interaction term between caregiving-related health problems and internet use to assess whether internet use moderated the association between caregiving-related health problems and loneliness, while controlling for the same set of covariates.

To facilitate the interpretation of significant interaction effects, we conducted post hoc analyses and created visual representations showing the relationship between caregiving-related health problems and loneliness at different levels of internet use. We also calculated adjusted means of loneliness scores for different combinations of caregiving-related health problems and internet use frequency to illustrate the interaction effect.

All analyses incorporated the complex survey design features of CHIS, including stratification, clustering, and sampling weights, to ensure that results were representative of older informal caregivers in California. We addressed missing data using multiple imputations for covariates with missing values less than 5%; cases with missing values on key study variables were excluded as per our sample selection criteria. We conducted all statistical analyses using Stata MP (version 17.0; StataCorp, with a significance level set at P<.05 for all tests. All analyses were conducted by a multidisciplinary team with expertise in gerontology, nursing, and biostatistics, providing confidence in the study's methodological rigor.

XSL•FO
RenderX

Ethical Considerations

This study was exempt from the Institutional Review Board review at New York University as CHIS is a publicly available dataset.

Results

Participants' Characteristics

Table 1 displays the sociodemographic characteristics, health status, and caregiving context of the 3957 older informal caregivers in California who participated in the study. The participants had a mean age of 72.46 years (SD 5.88), with 58.6% (n=2,319) being female. The sample was predominantly non-Hispanic White (78.6%, n=3,109), with 11.7% (n=462) born outside the United States. Regarding educational attainment, most participants (86.7%, n=3,432) had at least some college education, and the majority (83.0%, n=3,283) reported family income at or above 200% of the federal poverty level.

XSL•FO
RenderX

Table . Characteristics of participants from the California Health Interview Survey 2019 - 2020.

Variables	Participants, (N=3957), n (%)
Sociodemographic variables	
Age (years old), mean (SD, range)	72.46 (5.88; 65-85)
Sex (female), n (%)	2319 (58.6)
Race/ethnicity, n (%)	
Non-Hispanic White	3109 (78.6)
Non-Hispanic Black	130 (3.3)
Hispanic/Latinx	357 (9.0)
Non-Hispanic Asian	263 (6.6)
Other	98 (2.5)
Born outside of the United States, n (%)	462 (11.7)
Education, n (%)	
<High school	525 (13.3)
≥ High school	3432 (86.7)
% Federal poverty level, n (%)	
<200%	674 (17.0)
≥200%	3283 (83.0)
Marital status, n (%)	
Married/living with partner	1780 (45.0)
Widowed/separated/divorced/never married	2177 (55.0)
Health status	
Smoking, n (%)	
Currently smoker	198 (5.0)
Quit smoking	1495 (37.8)
Never smoking	2264 (57.2)
Body mass index, mean (SD, range)	26.8 (6.11; 12.65 - 69.35)
Self-rated health, mean (SD, range)	2.44 (0.96; 1-5)
Number of chronic conditions[a], mean (SD, range)	1.19 (1.00; 0-5)
Caregiving context, n (%)	
Experience financial stress due to caregiving, continuous	3.47 (0.81; 1-4)
Caregiving hours per week	15.39 (21.24; 1-125)
Care recipient's living condition, n (%)	
Living alone	1010 (25.5)
Living with caregivers or other family members	2069 (52.3)
Nursing home/assisted living facility/other situations	878 (22.2)
Relationship to the care recipient, n (%)	
Spouse/partner	1099 (27.8)
Parent/parent-in-law	971 (24.5)
Other relatives	925 (23.4)
Nonrelative	962 (24.3)
Dementia caregiver, n (%)	
Yes	911 (23.0)
No	3046 (77.0)

XSL•FO
RenderX

Variables	Participants, (N=3957), n (%)
Internet use frequency, mean (SD; range)	2.71 (0.84; 1-4)
Have health problems due to caregiving, n (%)	475 (12.0)
Loneliness, mean (SD; range)	3.98 (1.41; 3 - 9)

[a]Chronic conditions: hypertension, asthma, diabetes, heart disease, vision/hearing impairment.

In terms of caregiving context, participants reported low to moderate levels of financial stress due to caregiving, with a mean score of 3.47 (SD 0.81) on a 4-point scale, where higher scores indicate less stress. The mean caregiving hours per week were 15.39 (SD 21.24) on a scale ranging from 1 to 125. Approximately half (52.3%, n=2069) of the care recipients lived with the caregivers or other family members, while 1010 (25.5%) lived alone and 878 (22.2%) resided in nursing homes, assisted living facilities, or other living arrangements. The sample exhibited a fairly even distribution across caregiver relationship types: 27.8% (n=1099) were caring for spouses or partners, 24.5% (n=971) for parents or parents-in-law, 23.4% (n=925) for other relatives, and 24.3% (n=962) for nonrelatives such as friends or neighbors. In addition, a total of 911 (23.0%) care recipients had Alzheimer disease and related dementia. Notably, 12.0% (n=475) of the caregivers reported experiencing physical or mental health problems due to their caregiving responsibilities.

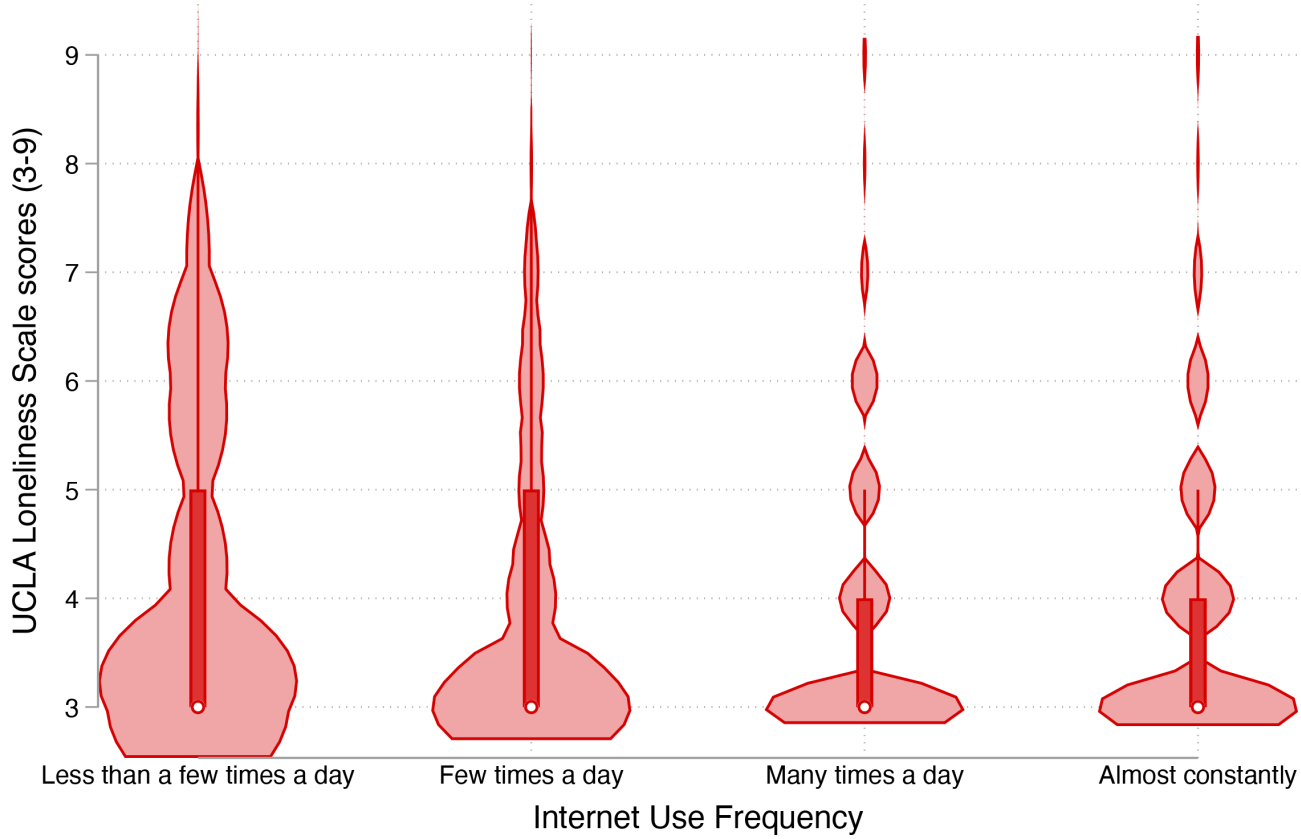
Regarding technology use, participants reported moderate levels of internet use, with a mean score of 2.71 (SD 0.84) on a 4-point scale, where a higher score indicates more frequent use. The mean UCLA Loneliness Scale score was 3.98 (SD 1.41) on a scale ranging from 3 to 9, indicating relatively low levels of loneliness in the overall sample.

Associations Between Health Effects of Caregiving, Internet Use, and Loneliness

Figure 2 presents the violin plot depicting the distribution of loneliness scores across the four internet use frequency categories. The distribution shape changes notably across categories, with the 'less than a few times a day' group showing the highest mean score of loneliness (4.14, SD 1.49) and the widest distribution, indicating greater variability in loneliness experience among infrequent internet users. In contrast, the 'almost constantly' group exhibited the lowest median score (3.85, SD 1.27), with a narrower distribution concentrated toward lower loneliness values. One-way ANOVA confirmed significant differences in loneliness across the four categories ($F_{3,\ 3953}$=7.05, P<.001). Post hoc tests revealed significant differences between all pairs of internet use frequency groups (all Ps<.05) except between the "many times a day" and "almost constantly" groups.

XSL•FO
RenderX

Figure 2. Violin plots of loneliness scores by internet use frequency categories among older informal caregivers (n=3957). The width of each violin represents the probability density of data at different values. White dots indicate medians, red boxes show interquartile ranges, and red lines represent the range excluding outliers. ANOVA testing confirmed significant differences between groups (*P*<.001). Mean and sample sizes: less than a few times a day (mean 4.14, SD1.49; n=188), few times a day (mean, SD 3.99, 1.47; n=1568), many times a day (mean 3.85, SD 1.33; n=1398), and almost constantly (mean 3.84, SD 1.27; n=803).



Regression Analysis of Caregiving-Related Health Effects, Internet Use, and Their Interaction

In Model 1 (Table 2), after adjusting for sociodemographic characteristics, health status, and caregiving context variables, caregivers who reported experiencing physical or mental health problems due to caregiving had significantly higher loneliness scores compared to those who did not report such problems (β=0.76, SE 0.07, *P*<.001). The regression analysis in Model 2 also revealed a significant inverse relationship between internet use frequency and loneliness (β=−0.11, SE 0.03; *P*<.001). Specifically, more frequent Internet use was associated with a lower level of loneliness after controlling for socio-demographic factors, health status, and caregiving context variables (Model 3). Additionally, we conducted a sensitivity analysis excluding self-rated health and number of chronic conditions from the regression models; the association between caregiving-related health problems and loneliness remained significant and of similar magnitude, suggesting that our findings were not driven by adjusting for overall health status.

XSL•FO
RenderX

Table . Multivariable linear regression of internet use and health problems of caregiving with loneliness.

Variables	Model 1	Model 2	Model 3	Model 4
	β coefficient (SE)			
Health problems of caregiving	0.76 (0.07)[a]	--	0.75 (0.07)[a]	1.19 (0.21)[a]
Internet use frequency[b]	--	−0.11 (0.03)[a]	−0.10 (0.02)[a]	−0.08 (0.02)[a]
Health problems of caregiving × internet use	--	--	--	−0.16 (0.07)[c]
Age, continuous	−0.01(0.00)[c]	−0.01(0.00)[c]	−0.01(0.00)[c]	−0.01(0.00)[c]
Female (ref. Male)	0.07 (0.04)	0.12 (0.05)[d]	0.08 (0.04)	0.08 (0.04)
Race/ethnicity (ref. Non-Hispanic White)				
Non-Hispanic Black	−0.32 (0.09)[a]	−0.34 (0.09)[a]	−0.30 (0.09)[a]	−0.30 (0.09)[a]
Hispanic/Latino	−0.10 (0.06)	−0.08 (0.06)	−0.08 (0.06)	−0.09 (0.06)
Non-Hispanic Asian	−0.33 (0.09)[a]	−0.34 (0.09)[a]	−0.36 (0.09)[a]	−0.36 (0.09)[a]
Other	−0.16 (0.11)	−0.19 (0.11)	−0.14 (0.11)	−0.15 (0.11)
Born outside of the United States (ref. Born in the United States)	0.39 (0.07)[a]	0.36 (0.07)[a]	0.39 (0.07)[a]	0.39 (0.07)[a]
Education (ref.<High school)				
≥ High school	0.04 (0.05)	0.07 (0.05)	0.02 (0.05)	0.02 (0.05)
% Federal poverty level (ref.<200%)				
≥200%	−0.04 (0.05)	−0.03 (0.05)	−0.05 (0.05)	−0.05 (0.05)
Marital status (ref. Married/living with partner)				
Widowed/separated/divorced/never married	−0.57 (0.05)[a]	−0.60 (0.05)[a]	-0.56 (0.05)[a]	−0.57 (0.05)[a]
Smoking (ref. Currently smokes)				
Quit smoking	0.07 (0.10)	0.06 (0.10)	0.05 (0.10)	0.05 (0.10)
Never smoking	−0.06 (0.10)	−0.08 (0.10)	−0.09 (0.10)	−0.09 (0.10)
Body mass index, continuous	−0.01 (0.00)[c]	−0.01 (0.00)[d]	−0.01 (0.00)[c]	−0.01 (0.00)[c]
Self-rated health, continuous[e]	−0.16 (0.03)[a]	−0.21 (0.03)[a]	−0.17 (0.03)[a]	−0.17 (0.03)[a]
Number of chronic conditions[f], continuous	0.02 (0.02)	0.01 (0.02)	0.02 (0.02)	0.02 (0.02)
Caregiving hours per week, continuous	−0.01 (0.00)[a]	−0.01 (0.00)[a]	−0.01 (0.00)[a]	−0.01 (0.00)[a]
Experience financial stress due to caregiving, continuous	0.23 (0.03)[a]	0.30 (0.03)[a]	0.23 (0.03)[a]	0.23 (0.03)[a]
Care recipient's living condition (ref. Living alone)				
Living with caregivers or other family members	0.12 (0.06)[c]	0.11 (0.06)	0.11 (0.06)	0.11 (0.06)
Nursing home/assisted living facility/other situations	0.08 (0.07)	0.08 (0.07)	0.07 (0.07)	0.07 (0.07)
Relationship to the care recipient (ref. Spouse/partner)				
Parent/parent-in-law	−0.02 (0.07)	−0.05 (0.07)	−0.03 (0.07)	−0.02 (0.07)
Other relatives	0.00 (0.06)	−0.05 (0.06)	−0.01 (0.06)	−0.00 (0.06)

XSL•FO
RenderX

Variables	Model 1	Model 2	Model 3	Model 4
	β coefficient (SE)			
Dementia caregiver (ref. No)				
Yes	0.67 (0.15)[a]	0.59 (0.17)[a]	0.69 (0.16)[a]	0.72 (0.17)[a]

[a]*P*<.001.

[b]A higher score indicates more frequent Internet use.

[c]*P*<.05.

[d]*P*<.01.

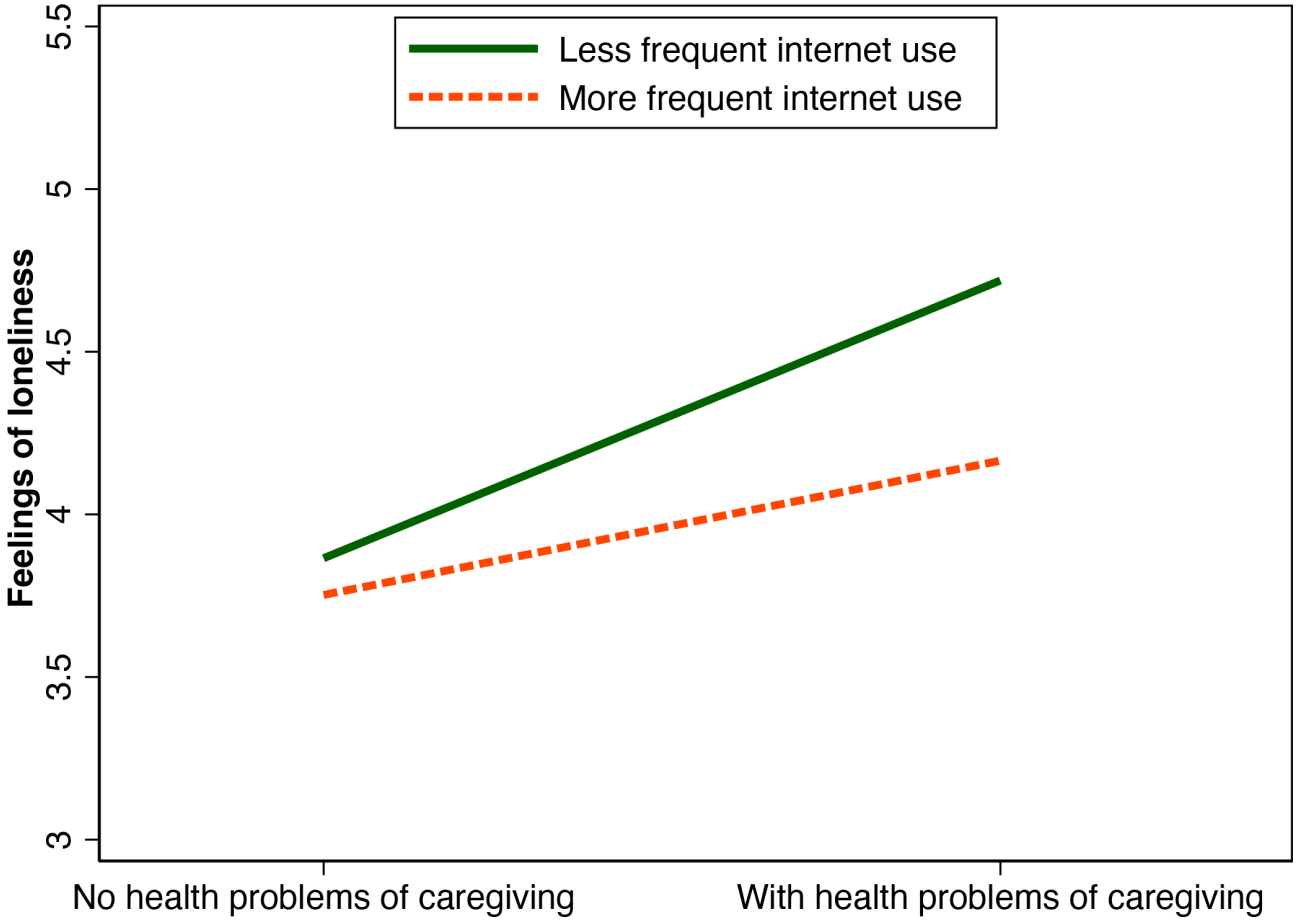
[e]A higher score indicates better self-rated health (5-point Linkert Scale).

[f]Chronic conditions: hypertension, asthma, diabetes, heart disease, vision/hearing impairment.

To examine the moderating role of internet use on the association between caregiving-related health problems and loneliness, we included an interaction term between internet use and caregiving-related health problems in the multivariable regression model (Model 4 in Table 2). The interaction term was statistically significant (β=−.16, SE 0.07, *P*=.02), indicating that internet use significantly attenuated the negative impact of caregiving-related health problems on loneliness.

Figure 3 illustrates this interaction effect, showing that the positive association between caregiving-related health problems and loneliness was stronger among caregivers with low internet use compared to those with high internet use. For caregivers with low internet use, experiencing health problems due to caregiving was associated with a substantial increase in loneliness scores. In contrast, for caregivers with high internet use, the difference in loneliness scores between those with and without caregiving-related health problems was less pronounced.

Figure 3. The moderating effect of internet use on the association between caregiving-related health problems and loneliness (California Health Interview Survey 2019 - 2020). The model was controlled for sociodemographics, health status, and caregiving context characteristics.



Discussion

Principal Findings

We hypothesized that older caregivers who experienced caregiving-related health problems would report higher loneliness, that more frequent internet use would be associated with lower loneliness, and that internet use would buffer the impact of caregiving-related health problems on loneliness. Consistent with these hypotheses, our study revealed three key findings regarding loneliness among older informal caregivers.

XSL•FO
RenderX

First, caregivers who reported experiencing physical or mental health problems due to caregiving had significantly higher levels of loneliness compared to those who did not report such problems, supporting our first hypothesis. Second, more frequent internet use was associated with lower levels of loneliness, confirming our second hypothesis. Third, internet use significantly moderated the relationship between caregiving-related health problems and loneliness, supporting our third hypothesis.

The observed association between caregiving-related health problems and increased loneliness aligns with previous research on the mental health consequences of caregiving burden [29-31]. Loneliness among older caregivers can stem from various sources, including emotional loneliness within the relationship with their care recipients, especially if the care recipient has dementia [9]. The nature of caregiving often necessitates significant time commitment and can lead to social isolation, as caregivers frequently sacrifice their social lives and personal needs to provide care [7,31]. This isolation can be particularly pronounced for spousal caregivers, who typically bear the most intensive care duties and often experience diminished ability to participate in social activities. The continuous concern for their care recipients can result in a sacrifice of personal freedom and a decline in their sense of identity within societal roles [30].

Our finding that more frequent internet use is associated with lower loneliness scores among older caregivers contributes to the growing literature on the potential benefits of digital technology for this population [15,25]. This relationship may reflect the ability of internet-based communication to facilitate social connections despite the time and mobility constraints that often accompany caregiving responsibilities [32]. The internet can serve as a bridge that enhances the intergenerational connections, mitigates feelings of loneliness among older informal caregivers, and boosts their overall quality of life [15,25]. By mitigating space and time constraints, the internet provides a channel through which caregivers can boost interaction within their interpersonal networks, alleviate perceived isolation, and maintain good mental status [33]. These benefits may be particularly valuable for older caregivers who face challenges in maintaining in-person social connections due to their caregiving duties. Research has shown that technology-enabled support interventions can increase feelings of social bond, social support, and shared experiences [25,33,34], all of which can alleviate loneliness and its negative effects on mental health among older adults. Digital platforms can also foster social and coping skills and help change maladaptive perceptions and thoughts, thereby decreasing loneliness (eg, evidence from internet-based caregiver support interventions shows improved coping and reduced isolation) [33,34]. Our findings regarding the association between internet use and a lower level of loneliness among older caregivers align with these mechanisms.

Additionally, our study found that internet use moderates the association between caregiving-related health problems and loneliness. This suggests that internet use may serve as a protective factor against loneliness for caregivers experiencing health problems due to their caregiving responsibilities. Consistent with the Stress Process Model's framework—which suggests that coping resources buffer the negative effects of caregiving stressors on outcomes like loneliness – our findings indicate that frequent internet use likely served as a coping resource that attenuated the impact of caregiving-related health problems on loneliness [19]. The internet may offer dual benefits for older caregivers experiencing health problems, as it can serve both as a source of health information and a means of maintaining social connections despite physical limitations or time constraints [25,26,33,34]. Digital technology can enable caregivers to access health resources, connect with health care providers, and join support groups relevant to their specific health concerns, while simultaneously facilitating social interaction with friends and family.

Our findings have several implications for interventions targeting loneliness among older caregivers. First, they suggest that promoting internet use among older caregivers may be a promising strategy for reducing loneliness, particularly for those experiencing health problems due to caregiving. Improving digital health literacy is crucial for enabling older caregivers to effectively utilize online resources. Recent evidence shows that tailored interventions can significantly enhance eHealth literacy in older populations, thereby empowering caregivers to find, evaluate, and apply online health information [35,36]. Health care providers should consider assessing both caregiving-related health problems and internet use patterns when evaluating the risk of loneliness among older caregivers. Those with caregiving-related health problems and low internet use may be particularly vulnerable to loneliness and could benefit from targeted interventions. Moreover, policy makers should consider investing in infrastructure and programs that increase internet accessibility and digital literacy among older adults, particularly those with caregiving responsibilities. This could include subsidized internet access, device distribution programs, and community-based digital literacy training tailored to older adults. Government, policy makers, and society should prioritize their efforts to promote internet literacy among older informal caregivers, advance the development of internet and digital infrastructure, and bridge the urban-rural divides.

This study has several limitations that need to be acknowledged. First, CHIS's cross-sectional design precludes causal inferences regarding the relationships between caregiving-related health problems, internet use, and loneliness. Longitudinal studies are needed to examine how changes in caregiving-related health problems and internet use over time influence trajectories of loneliness among older caregivers. Second, our measure of internet use focused solely on frequency rather than specific types of internet activities or duration of online interactions. Future research should examine how different patterns and durations of internet use (eg, social media, entertainment, health information seeking) may differentially affect loneliness among older caregivers. Understanding internet use patterns and device adoption across older caregivers and care recipients is increasingly recognized in enhancing geriatric care [37]. Third, health problems due to caregiving measures are self-reported and may not exclusively reflect problems caused by caregiving alone. It is challenging to definitively attribute health issues solely to caregiving responsibilities with a cross-sectional self-report item, and we caution that this variable should be

XSL•FO
RenderX

interpreted as a perceived impact of caregiving on health. Finally, although our sample was representative of older informal caregivers in California, the generalizability of our findings to other geographic contexts may be limited. Future studies should examine these relationships in diverse populations and settings, including rural areas with limited internet infrastructure and communities with varying levels of technological integration. Future research should explore the development of tailored digital interventions specifically designed to address loneliness among older caregivers. These interventions could integrate components addressing both caregiving-related stressors and social connection, potentially leveraging artificial intelligence to personalize content based on individual caregiving contexts and preferences. In addition, emerging technologies such as social robots have shown promise in supporting older people with dementia and their informal caregivers [38] and could be explored as complementary tools to reduce loneliness in caregiving contexts.

Conclusion

Our study provides evidence that internet use is associated with lower levels of loneliness among older informal caregivers and could buffer against the negative impact of caregiving-related health problems on loneliness. This study highlights the potential of digital technology as a tool for addressing loneliness in this vulnerable population. As the population ages and the number of older caregivers increases, developing effective strategies to support their psychosocial well-being becomes increasingly important. Internet-based interventions represent a promising approach, particularly for caregivers facing health challenges related to caregiving responsibilities. Future research and policy efforts should focus on enhancing internet accessibility, promoting digital literacy, and developing age-friendly digital platforms that meet the specific needs of older caregivers.

Acknowledgments

The authors thank all California Health Interview Survey participants and staff for their contribution to this study.

Funding

This study is partially supported by National Institutes of Health (K99AG076871, P30AG083257, P50MD017356). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability

The California Health Interview Survey (CHIS) datasets are publicly available at the UCLA Center for Health Policy Research. Researchers may obtain the datasets after sending a data user agreement to the CHIS team [39].

Conflicts of Interest

None declared.

References

1. Malani PN, Solway E, Kirch M, Singer DC, Roberts JS, Kullgren JT. Loneliness and social isolation among US older adults. JAMA 2025 Jan 21;333(3):254-257. [doi: 10.1001/jama.2024.23213] [Medline: 39652320]
2. Domènech-Abella J, Mundó J, Switsers L, van Tilburg T, Fernández D, Aznar-Lou I. Social network size, loneliness, physical functioning and depressive symptoms among older adults: examining reciprocal associations in four waves of the Longitudinal Aging Study Amsterdam (LASA). Int J Geriatr Psychiatry 2021 Oct;36(10):1541-1549. [doi: 10.1002/gps.5560] [Medline: 33908639]
3. Holt-Lunstad J, Smith TB, Baker M, Harris T, Stephenson D. Loneliness and social isolation as risk factors for mortality: a meta-analytic review. Perspect Psychol Sci 2015 Mar;10(2):227-237. [doi: 10.1177/1745691614568352] [Medline: 25910392]
4. Qi X, Malone SK, Pei Y, Zhu Z, Wu B. Associations of social isolation and loneliness with the onset of insomnia symptoms among middle-aged and older adults in the United States: a population-based cohort study. Psychiatry Res 2023 Jul;325:115266. [doi: 10.1016/j.psychres.2023.115266] [Medline: 37245484]
5. Shen C, Rolls ET, Cheng W, et al. Associations of social isolation and loneliness with later dementia. Neurology (ECronicon) 2022 Jul 11;99(2):e164-e175. [doi: 10.1212/WNL.0000000000200583] [Medline: 35676089]
6. US Surgeon General. Our epidemic of lonelinesss and isolation. : United States Department of Health and Human Services; 2024 URL: https://www.hhs.gov/sites/default/files/surgeon-general-social-connection-advisory.pdf [accessed 2025-11-25]
7. Hajek A, Kretzler B, König HH. Informal caregiving, loneliness and social isolation: a systematic review. Int J Environ Res Public Health 2021 Nov 18;18(22):12101. [doi: 10.3390/ijerph182212101] [Medline: 34831857]
8. Yanguas J, Pinazo-Henandis S, Tarazona-Santabalbina FJ. The complexity of loneliness. Acta Biomed 2018 Jun 7;89(2):302-314. [doi: 10.23750/abm.v89i2.7404] [Medline: 29957768]

XSL•FO
RenderX

9. Chen XS, Lydon EA, Mathias VF, Rogers WA, Mudar RA, Raj M. Loneliness among older caregivers: an analysis of the 2020 California Health Interview Survey. J Appl Gerontol 2025 Jul;44(7):1193-1204. [doi: 10.1177/07334648241301484] [Medline: 39577446]
10. Cacioppo JT, Cacioppo S. Loneliness in the modern age: an evolutionary theory of loneliness (ETL). In: Advances in Experimental Social Psychology: Elsevier; 2018, Vol. 58:127-197. [doi: 10.1016/bs.aesp.2018.03.003]
11. Bonsaksen T, Ruffolo M, Price D, et al. Associations between social media use and loneliness in a cross-national population: do motives for social media use matter? Health Psychol Behav Med 2023;11(1):2158089. [doi: 10.1080/21642850.2022.2158089] [Medline: 36618890]
12. Mozafar Saadati H, Mirzaei H, Okhovat B, Khodamoradi F. Association between internet addiction and loneliness across the world: a meta-analysis and systematic review. SSM Popul Health 2021 Dec;16:100948. [doi: 10.1016/j.ssmph.2021.100948] [Medline: 34754896]
13. Pei Y, Qi X, Li G, et al. Unequal effects of the lockdown on mental health in Shanghai: the moderating and mediating role of neighborhood environment and online social connections. J Community Psychol 2025 Jan;53(1):e23177. [doi: 10.1002/jcop.23177] [Medline: 39789845]
14. Choi M, Kong S, Jung D. Computer and internet interventions for loneliness and depression in older adults: a meta-analysis. Healthc Inform Res 2012 Sep;18(3):191-198. [doi: 10.4258/hir.2012.18.3.191] [Medline: 23115742]
15. Mao W, Qi X, Chi I, Wichinsky L, Wu B. Technology-based interventions to address social isolation and loneliness among informal dementia caregivers: a scoping review. J Am Med Dir Assoc 2023 Nov;24(11):1700-1707. [doi: 10.1016/j.jamda.2023.08.005] [Medline: 37678415]
16. Blusi M, Kristiansen L, Jong M. Exploring the influence of internet‐based caregiver support on experiences of isolation for older spouse caregivers in rural areas: a qualitative interview study. Int J Older People Nursing 2015 Sep;10(3):211-220. [doi: 10.1111/opn.12074]
17. Hua Z. Internet usage buffers the effect of loneliness on subjective health among informal caregivers of older adults. Psychol Health Med 2025 Feb 7;30(2):234-251. [doi: 10.1080/13548506.2024.2417313]
18. Li J, Liang N, Wu Y, Ma X, Liao Z. Effect of internet usage on the life satisfaction of the elderly spousal caregivers of disabled elderly. Sci Rep 2024;14(1):22294. [doi: 10.1038/s41598-024-73298-8]
19. Pearlin LI, Mullan JT, Semple SJ, Skaff MM. Caregiving and the stress process: an overview of concepts and their measures. Gerontologist 1990 Oct;30(5):583-594. [doi: 10.1093/geront/30.5.583] [Medline: 2276631]
20. Ponce NA, Lavarreda SA, Yen W, Brown ER, DiSogra C, Satter DE. The California Health Interview Survey 2001: translation of a major survey for California's multiethnic population. Public Health Rep 2004;119(4):388-395. [doi: 10.1016/j.phr.2004.05.002] [Medline: 15219795]
21. Sherr S, Goyle A, Langdale K, Best J. C. CHIS 2019–2020 Methodology Report Series. Report 5: weighting and variance. : UCLA Center for Health Policy Research; 2021 URL: https://healthpolicy.ucla.edu/sites/default/files/2023-05/chis_2019_2020_methodologyreport5_weightingandvarianceestimation_final-ucla-10sep2021.pdf [accessed 2025-11-25]
22. Hughes ME, Waite LJ, Hawkley LC, Cacioppo JT. A short scale for measuring loneliness in large surveys: results from two population-based studies. Res Aging 2004;26(6):655-672. [doi: 10.1177/0164027504268574] [Medline: 18504506]
23. Qi X, Belsky DW, Yang YC, Wu B. Association between types of loneliness and risks of functional disability in older men and women: a prospective analysis. Am J Geriatr Psychiatry 2023 Aug;31(8):621-632. [doi: 10.1016/j.jagp.2023.02.046] [Medline: 36935279]
24. Gonzalez M, Sanders-Jackson A, Emory J. Online health information-seeking behavior and confidence in filling out online forms among Latinos: a cross-sectional analysis of the California Health Interview Survey, 2011-2012. J Med Internet Res 2016 Jul 4;18(7):e184. [doi: 10.2196/jmir.5065] [Medline: 27377466]
25. Sherifali D, Ali MU, Ploeg J, et al. Impact of internet-based interventions on caregiver mental health: systematic review and meta-analysis. J Med Internet Res 2018 Jul 3;20(7):e10668. [doi: 10.2196/10668] [Medline: 29970358]
26. Xie Y, Shen S, Liu C, et al. Internet-based supportive interventions for family caregivers of people with dementia: randomized controlled trial. JMIR Aging 2024 Oct 4;7:e50847. [doi: 10.2196/50847] [Medline: 39365914]
27. Qi X, Wang K, Pei Y, Đoàn LN, Yi SS, Wu B. Psychological distress in Asian American informal caregivers: an analysis by disaggregated ethnic groups. Ethn Health 2025 Feb 17;30(2):232-253. [doi: 10.1080/13557858.2024.2430287]
28. Qi X, Pei Y, Li G, Tang W, Hall BJ, Wu B. Different sources of social capital and loneliness among adults in Shanghai, China: exploring sex differences. Community & Applied Soc Psy 2025 Mar;35(2):e70089. [doi: 10.1002/casp.70089]
29. Calderón-Larrañaga A, Kirvalidze M, Dahlberg L, Sacco LB, Morin L. Effectiveness of interventions to address the negative health outcomes of informal caregiving to older adults: protocol for an umbrella review. BMJ Open 2021 Nov 9;11(11):e053117. [doi: 10.1136/bmjopen-2021-053117] [Medline: 34753765]
30. Li L, Wister AV, Mitchell B. Social isolation among spousal and adult-child caregivers: findings from the Canadian Longitudinal Study on Aging. J Gerontol B Psychol Sci Soc Sci 2021 Aug 13;76(7):1415-1429. [doi: 10.1093/geronb/gbaa197] [Medline: 33170276]
31. Ross A, Perez A, Wehrlen L, et al. Factors influencing loneliness in cancer caregivers: a longitudinal study. Psychooncology 2020 Nov;29(11):1794-1801. [doi: 10.1002/pon.5477] [Medline: 32672866]

XSL•FO
RenderX

32. Nowland R, Necka EA, Cacioppo JT. Loneliness and social internet use: pathways to reconnection in a digital world? Perspect Psychol Sci 2018 Jan;13(1):70-87. [doi: 10.1177/1745691617713052] [Medline: 28937910]
33. Magid K, Sagui-Henson SJ, Sweet CC, Smith BJ, Chamberlain CEW, Levens SM. The impact of digital mental health services on loneliness and mental health: results from a prospective, observational study. IntJ Behav Med 2024 Jun;31(3):468-478. [doi: 10.1007/s12529-023-10204-y]
34. Shah SGS, Nogueras D, van Woerden HC, Kiparoglou V. Evaluation of the effectiveness of digital technology interventions to reduce loneliness in older adults: systematic review and meta-analysis. J Med Internet Res 2021 Jun 4;23(6):e24712. [doi: 10.2196/24712] [Medline: 34085942]
35. Bevilacqua R, Strano S, Di Rosa M, et al. EHealth literacy: from theory to clinical application for digital health improvement: results from the ACCESS Training Experience. Int J Environ Res Public Health 2021 Nov 10;18(22):11800. [doi: 10.3390/ijerph182211800] [Medline: 34831555]
36. Barbati C, Maranesi E, Giammarchi C, et al. Effectiveness of eHealth literacy interventions: a systematic review and meta-analysis of experimental studies. BMC Public Health 2025 Jan 23;25(1):288. [doi: 10.1186/s12889-025-21354-x] [Medline: 39849354]
37. Shaffer KM, Chow PI, Cohn WF, Ingersoll KS, Ritterband LM. Informal caregivers' use of internet-based health resources: an analysis of the Health Information National Trends Survey. JMIR Aging 2018 Dec 18;1(2):e11051. [doi: 10.2196/11051] [Medline: 31518244]
38. Amabili G, Cucchieri G, Margaritini A, et al. Social robotics and dementia: results from the eWare Project in Supporting Older People and Their Informal Caregivers. Int J Environ Res Public Health 2022 Oct 16;19(20):13334. [doi: 10.3390/ijerph192013334] [Medline: 36293915]
39. California Health Interview Survey (CHIS). Center for Health Policy Research. URL: https://healthpolicy.ucla.edu/our-work/california-health-interview-survey-chis [accessed 2025-11-25]

Abbreviations

CHIS: California Health Interview Survey
HSD: Honestly Significant Difference

Edited by G Seçkin; submitted 19.03.25; peer-reviewed by G Amabili, W Tang; revised version received 27.09.25; accepted 14.11.25; published 27.11.25.

Please cite as:
Qi X, Liu R, Ko E, Pei Y, Wu B
Buffering Effects of Internet Use on Caregiving-Related Health Impacts and Loneliness Among Older Informal Caregivers in California: Cross-Sectional Study
JMIR Aging 2025;8:e74209
URL: https://aging.jmir.org/2025/1/e74209
doi:10.2196/74209

© Xiang Qi, Ruotong Liu, Eunjung Ko, Yaolin Pei, Bei Wu. Originally published in JMIR Aging (https://aging.jmir.org), 27.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Comparative Analysis of 16 Aging Concepts and Their Influence on Aging Narratives: Bibliometric and Content Analysis

Na Xiao[1], MA; Bo Xia[1], PhD; Laurie Buys[2], PhD; Connie Susilawati[3], PhD; Martin Larbi[1], PhD; Qing Chen[2], PhD

[1]School of Architecture and Built Environment, Faculty of Engineering, Queensland University of Technology, D Block Level 4 401, Gardens Point, Brisbane, Australia

[2]Faculty of Health Sciences, Australian Catholic University, Brisbane, Australia

[3]School of Economics and Finance, Faculty of Business and Law, Queensland University of Technology, Brisbane, Australia

Corresponding Author:
Na Xiao, MA
School of Architecture and Built Environment, Faculty of Engineering, Queensland University of Technology, D Block Level 4 401, Gardens Point, Brisbane, Australia

Abstract

Background: Globally, various aging concepts (such as healthy aging, successful aging, and active aging) have emerged to promote the goal of "aging well" and have gained widespread attention in academia, policy, and practice to change the negative narrative on aging. However, whether and how these aging concepts have contributed to changing the negative narratives remains unclear. Moreover, they are not clearly defined nor widely agreed upon, often creating ambiguity and confusion.

Objective: This paper aims to provide a comprehensive review and comparative analysis of 16 aging concepts, with a particular focus on how their evolution in research has contributed to shifting the narrative surrounding aging.

Methods: This study used the bibliometric software VosViewer (Center for Science and Technology Studies) to visualize international collaboration among countries and cocitation networks among journals. This helped identify which countries and journals play central roles in research on aging concepts and revealed how academic contributions are distributed globally. Additionally, content analysis supported by the corpus linguistics software AntConc (Waseda University) was conducted to examine and compare the main focuses, applications, challenges, and future research directions of these concepts.

Results: The findings indicate that while all 16 aging concepts share the common goal of improving the quality of life for older adults, they offer different perspectives, encompassing health management, social participation, mental health, and technological innovation. Key challenges to achieving the goal of each aging concept were identified, including unequal access to health care resources, barriers to social participation, and difficulties in adopting technology.

Conclusions: The overall impact of these aging concepts on reshaping negative aging narratives remains relatively limited. Future efforts should focus on advancing technology, optimizing policies, enhancing social support systems, and fostering global collaboration to provide innovative and sustainable solutions that promote the overall well-being of older adults.

(JMIR Aging 2025;8:e72011) doi:10.2196/72011

KEYWORDS

aging concepts; health aging; successful aging; active aging; bibliometric analysis

Introduction

The global trend of an aging population is accelerating, with a steadily increasing proportion of people older than 65 years of age worldwide [1,2]. Many governments view this demographic change as an increasingly pressing challenge, particularly in the areas of health and social security [3]. Policy makers commonly associate aging with higher incidences of chronic diseases and increased reliance on long-term care services, which places significant burdens on public health budgets and resource allocation [4-7]. Similarly, many researchers view aging as a growing challenge, emphasizing how the prevalence of neurodegenerative diseases, diabetes, and other chronic conditions in older populations significantly increases medical demands, thereby intensifying pressure on health care systems [8-11].

Governmental focus on fiscal sustainability, along with academia's emphasis on the challenges of the aging process, has fostered a prevailing negative narrative of aging as an inevitable social and economic burden. Policy documents frequently underscore the financial pressure aging places on pension systems and long-term care insurance, driving related legislation and measures [12]. This "cost-centric" perspective is also prevalent and echoed in academic studies, which highlight the strain aging poses to the labor market and its potential threat to national economic productivity [13-15].

XSL•FO
RenderX

In response to the prevalent focus on the negative impacts of aging, some researchers and international organizations (eg, World Health Organization [WHO]) have sought to reshape this narrative by introducing aging concepts like "healthy aging," "successful aging," and "active aging" [16]. These concepts seek to shift the focus from the diseases and functional decline associated with aging to the positive aspects of fostering holistic well-being and active social participation among older adults through preventive and supportive measures [17-21]. For example, the concept of active aging highlights the importance of continued participation in social, economic, cultural, and physical activities for older adults, and it emphasizes staying active in all areas of life to improve both the quality of life and overall health [22]. These concepts collectively aim to counteract the dominant "cost-centric" and deficit-based narratives of aging, promoting a more holistic view of older adults as active, valuable members of society.

However, despite the initial intentions behind these aging concepts, whether or not they have succeeded in counteracting the dominant negative narrative around aging remains largely unknown. In fact, some researchers even argue that these concepts may inadvertently reinforce certain aspects of the negative view of aging [23-25]. For example, Waddell et al [26] mentioned that achieving "successful aging" often emphasizes individual responsibility for maintaining health and well-being, overlooking broader social and environmental factors, setting unrealistic standards that marginalize those with health challenges, thus reinforcing a deficit-based view, and framing aging as a problem to be managed rather than a natural process. Additionally, the commercialization of "active aging" and "healthy aging" (eg, the "silver market" and "silver economy") reinforces the notion that aging requires continuous self-improvement, perpetuating ageist views that devalue individuals who do not meet ideals of productivity or health [27,28].

Moreover, the definitions and applications of these aging concepts often vary significantly, leading to confusion and misinterpretation in both academic and practical contexts. For instance, "successful aging" from biomedical research is commonly defined as the absence of major diseases and the maintenance of good physical function in older people [29]. In contrast, successful aging in sociological and psychological studies emphasizes the ability of older adults to maintain social connections and mental well-being [30]. These definitional inconsistencies hinder cross-disciplinary understanding and cooperation and introduce uncertainties into the development and implementation of public health strategies and policy implementation.

Therefore, this study aims to examine whether, and in what ways, the application of aging concepts in academic research has contributed to reframing the dominant negative narrative surrounding aging. To achieve this, a systematic review of aging-related literature was conducted, leading to the identification of 16 key aging concepts. These concepts were then used as keywords to retrieve relevant papers from the Scopus database. Scopus's built-in analysis tools, along with the bibliometric software VosViewer (Center for Science and Technology Studies), were used to map the distribution of publications across disciplines and journals, and to visualize international research collaborations and relationships between frequently cocited journals.

Furthermore, the linguistic software AntConc (Waseda University) supported a comprehensive content analysis, examining core themes, practical applications, persistent challenges, and emerging trends associated with these aging concepts. Based on these analyses, the study not only charts how aging concepts are defined, applied, and interpreted within academic discourse but also evaluates whether their application signals a shift from deficit-based to empowerment-oriented narratives.

To support this interpretive approach, the study draws on narrative theory (eg, [31]), which emphasizes how language and discourse shape the meanings ascribed to aging. Complementing this, the social construction of aging framework (eg, [32]) informs our understanding of aging concepts as dynamic constructs that are embedded in broader cultural, institutional, and policy contexts. Together, these theoretical models guide our analysis of how aging concepts contribute to either deficit-based or empowerment-oriented representations of later life.

Methods

Data Collection

This study adopts a systematic literature review approach, aiming to ensure the comprehensiveness and credibility of data through extensive searching, iterative reading, filtering, and cross-validation. The process began with the use of Google Scholar to identify aging concepts through close reading of academic journal papers, books, and conference proceedings. This initial stage was further enhanced by including authoritative documents from international organizations such as the WHO, the European Commission, and relevant government reports. This multilayered approach enabled the identification of 16 aging concepts as follows: (1) healthy aging, (2) successful aging, (3) active aging, (4) positive aging, (5) productive aging, (6) optimal aging, (7) social aging, (8) creative aging, (9) smart aging, (10) healthy-active aging, (11) adaptive aging, (12) robust aging, (13) resilient aging, (14) conscious aging, (15) intelligent aging, and (16) vital aging.

These 16 aging concepts were then used as keywords to search in the Scopus database. Scopus was selected for its extensive multidisciplinary coverage, inclusion of high-quality peer-reviewed journals across both biomedical and social sciences, and compatibility with bibliometric tools such as VosViewer. Its structured metadata and citation indexing made it especially suitable for large-scale quantitative and linguistic analysis. While Scopus provides a robust foundation for this study, future research may benefit from incorporating other databases (eg, Web of Science and PubMed) for triangulation and broader validation.

Initial searches occasionally returned irrelevant results, such as papers focused on animal aging or the degradation of materials and mechanical systems. To refine the dataset and focus on

XSL•FO
RenderX

human aging, population-specific terms were added. The final search string was:

TITLE-ABS-KEY(“healthy aging” OR “active aging” OR “healthy-active aging” OR “smart aging” OR “intelligent aging” OR “successful aging” OR “creative aging” OR “social aging” OR “conscious aging” OR “resilient aging” OR “adaptive aging” OR “positive aging” OR “productive aging” OR “robust aging” OR “optimal aging” OR “vital aging”)AND(“elderly” OR “elderly people” OR “elders” OR “the aged” OR “the elderly population” OR “older adults” OR “older individuals” OR “older persons” OR “older generation” OR “senior citizens” OR “senior population” OR “seniors” OR “golden-agers” OR “retirees” OR “pensioners” OR “aging population” OR “aging adults” OR “geriatric population”)

Both American (“aging”) and British (“ageing”) spellings were used to ensure comprehensive retrieval. For compound terms such as “healthy-active aging,” both hyphenated and nonhyphenated forms were included. As of October 21, 2024, the final search yielded 10,642 papers in English-language peer-reviewed journals.

Data Analysis

A combination of bibliometric, corpus linguistic, and qualitative content analysis methods was used to examine how aging concepts are applied and framed in academic literature. The built-in analytical tools of the Scopus database were used to perform an initial descriptive analysis of the dataset, including (1) the distribution of publications across various journals, (2) the trends in the number of publications over the years, (3) the spread of research across various academic disciplines, (4) and the emergence time of 16 aging concepts.

Following this, the bibliometric software VosViewer was used to visualize the co-occurrence of (1) geographic collaboration among countries, and (2) cocited journal references. The goal was to identify key regions contributing significantly to aging research and highlight highly influential journals. VosViewer is widely used for bibliometric analysis and scientific knowledge mapping, known for its exceptional capacity to manage large datasets, such as the 10,642 papers retrieved in this study. These network analyses offer insight into the structural dynamics of academic discourse, including disciplinary concentration and geographic dominance, which serve as contextual indicators for understanding the breadth and inclusivity of narrative development around aging concepts. In both cocitation and collaboration networks, total link strength (TLS) was used as a metric to quantify the strength of connections between nodes (eg, journals or countries). TLS is automatically calculated by VosViewer based on the frequency and intensity of co-occurrence—cocitation in journal networks, or coauthorship in country networks. A higher TLS value indicates stronger linkage and greater structural centrality within the network.

For the comparative analysis, 2 datasets were extracted:

- Dataset 1 consisted of a full set of 10,642 papers based on the 16 aging concepts, with extracted text from titles, abstracts, and keywords.
- Dataset 2 consisted of a focused set of 1402 papers retrieved using the keywords “age well,” “aging well,” and “ageing well,” also limited to titles, abstracts, and keywords.

The first dataset was concept-driven, focusing on how predefined aging concepts are applied in scholarly research, while the second dataset was goal-oriented, capturing how the broader notion of “aging well” naturally emerges in academic discourse. The second dataset was provided as a comparative lens, enabling an assessment of whether and how the selected aging concepts were embedded within mainstream narratives surrounding positive aging outcomes.

Then, linguistic analysis was performed using the software AntConc. Word frequency counts identified how often each aging concept appeared in both datasets, providing a basis for comparing their relative prominence. AntConc’s Keywords in Context function supported in-depth qualitative content analysis by extracting 30-word windows surrounding each concept for further analysis. Specifically, the initial step involved analyzing text from Dataset 1, which formed the foundation for subsequent analysis. The Keywords in Context function was then used to locate occurrences of the aging concepts within the text, extracting 30 words before and after each keyword to form target sentences. In the first round of content analysis, these target sentences were categorized into several main themes. The primary themes included:

- Core content: describes the main ideas of each aging concept.
- Application areas: refer to the domains where these concepts have been applied, such as public health or social policy.
- Challenges: highlight the difficulties encountered in achieving the goals of these aging concepts.
- Future trends: focus on the development directions and potential applications of these concepts.

After completing the initial analysis, a second round of content analysis was conducted. In this phase, differences between the concepts were compared to provide a more summarized description—a “big picture,” so to speak. The process followed a clear and logical flow: beginning with large-scale text extraction, identifying keyword contexts, conducting 2 rounds of thematic content analysis, and concluding with a comparative synthesis. The analytical procedures are illustrated in Figure 1.

To evaluate whether aging concepts have contributed to shifting the negative narrative on aging, the content analysis examined the presence and emphasis of themes reflecting either deficit-based or empowerment-based narratives. Concepts framed primarily around disease, burden, and decline were interpreted as maintaining or even reinforcing a traditional negative discourse, while those emphasizing well-being, participation, adaptability, and holistic development were viewed as indicative of a shifting, more positive narrative change. The distribution of disciplines and publication trends also served as supporting indicators: a concentration of research within medical domains tended to reinforce traditional narratives, while an increasing presence in social, policy, or interdisciplinary domains suggested progress toward a more balanced narrative. These interpretive criteria guided the

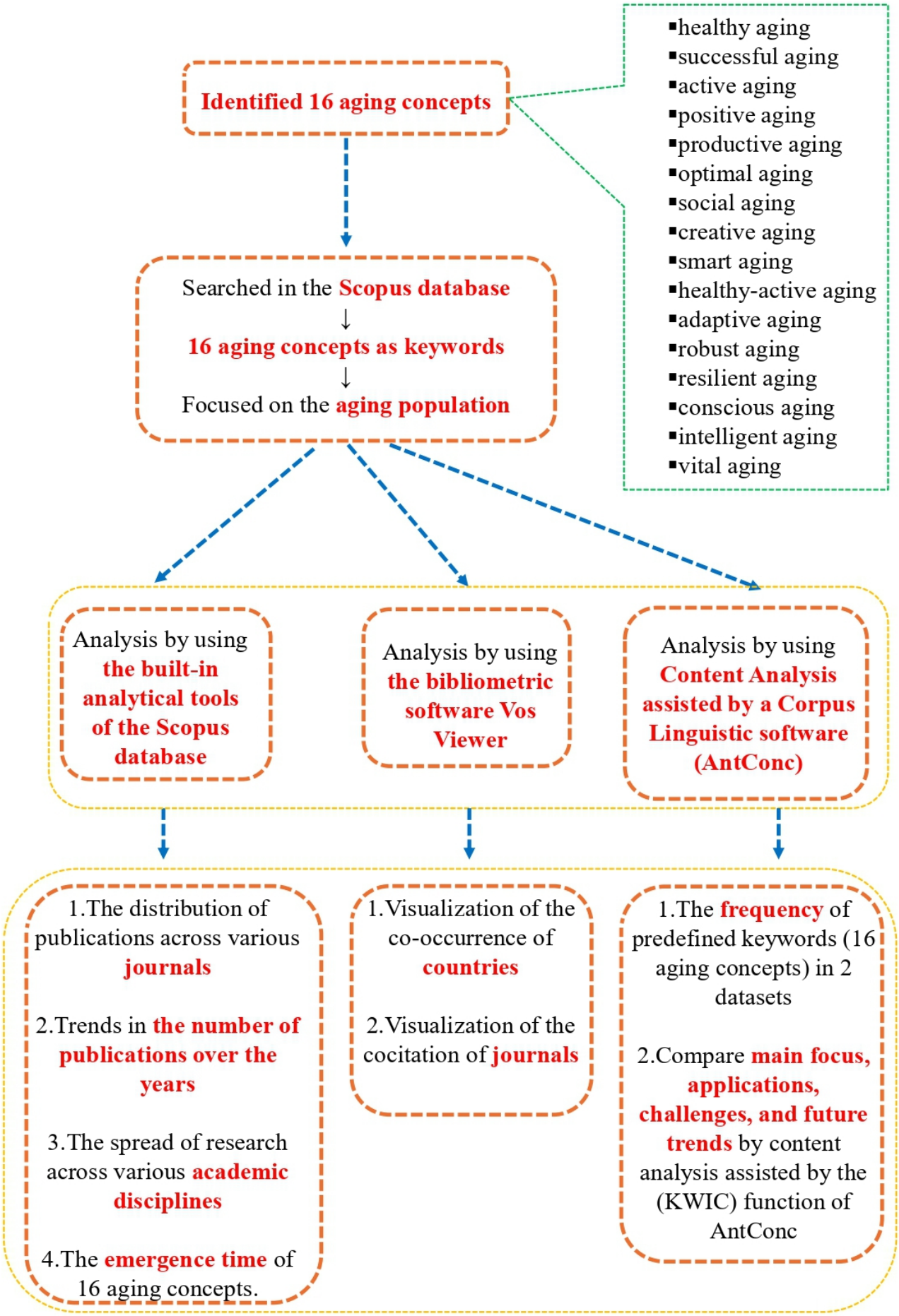
XSL•FO
RenderX

qualitative assessment of whether and how narrative transformation is occurring in the academic literature.

XSL•FO
RenderX

Figure 1. Research procedure. KWIC: Keywords in Context.

XSL•FO
RenderX



XSL•FO
RenderX

Results

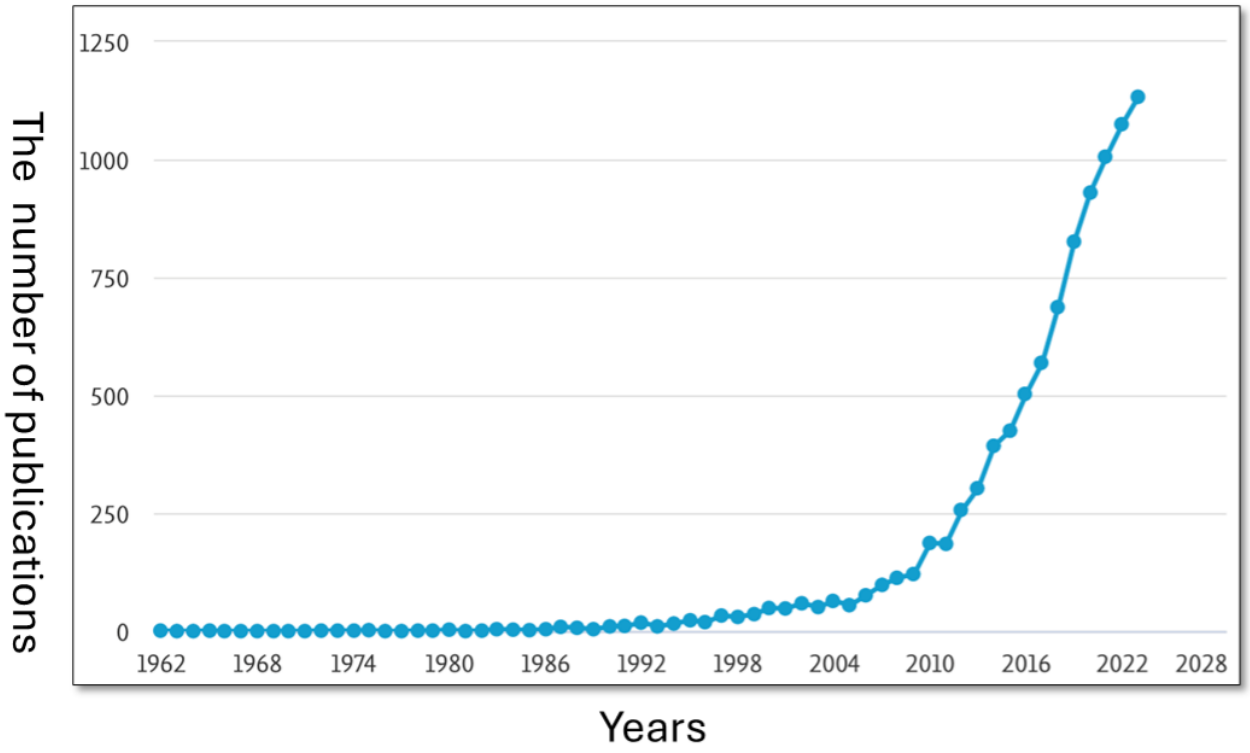
Publication Trends and Disciplinary Distribution

Distributions of Publication Volume in Years

The line chart (Figure 2) illustrates the annual publication volume trend over the 62-year period from 1962 to 2023, based on 9465 papers out of the 10,642 retrieved, as data for 2024 represents only part of the year, which affects the trend line.

It is clearly shown that there is an exponential increase in the number of papers published, particularly after 2011, which indicates a growing interest and emphasis on research mentioning aging concepts. In the early 50 years (from 1962 to 2011), the number of publications remained steady, but slow growth, generally below 250 per year. Then, in the near-decade (from 2012 to 2020), the number of publications experienced a sharp increase, with the count reaching between 250 and 1000 papers. This upward trajectory continued between 2021 and 2023, when the number of publications exceeded 1000 papers.

Figure 2. Trends in publications mentioning aging concepts (2016-2023).



Distributions Cross Disciplines

The disciplinary distribution of publications is presented in Multimedia Appendix 1. Medicine takes the lead with 34.2%, clearly positioning it as the central discipline in research applying aging concepts. Social Sciences (12.4%) and Biochemistry and Genetics (9.8%) follow at a distance, indicating that the research disciplines extend beyond medicine to encompass the social sciences and biosciences. In contrast, Health Professions account for 3%, Environmental Science for 2.8%, and Computer Sciences for only 1.5%, indicating that these fields contribute fewer aging-related publications in absolute terms.

This distribution highlights an important structural feature of aging research: while it remains heavily concentrated in medical domains—often reflecting deficit-based narratives centered on disease, decline, and dependency—there is a gradual increase in contributions from disciplines that approach aging from more social, psychological, or interdisciplinary perspectives. This expansion suggests a modest shift toward a more multidimensional, empowerment-oriented discourse. However, it is worth noting that this analysis does not account for the total number of publications within each discipline. Therefore, it cannot measure the relative importance of aging within each field. Future studies could address this limitation by normalizing the data against overall publication volume within each discipline.

Distribution of Journals

Table 1 presents the distribution of the top 35 journals and the number of papers published in each (all with over 50 publications). This distribution reflects the broad scope and complexity across areas of research applying aging concepts, spanning fields such as public health, geriatrics, neuroscience, neurodegenerative diseases, aging psychology and mental health, sociology, social sciences, nutrition, and multidisciplinary aging studies. The top 5 journals in this list are the *International Journal of Environmental Research and Public Health*, with 322 papers; *BMC Geriatrics*, with 204 papers; *PLOS One*, with 148 papers; *Aging and Society*, with 144 papers; and *Frontiers in Aging Neuroscience*, with 142 papers.

XSL•FO
RenderX

Table . Top 35 journals ranking by the number of papers (as of October 21, 2024).

No.	Journal name	Papers, n
1	*International Journal of Environmental Research and Public Health*	322
2	*BMC Geriatrics*	204
3	*PLOS One*	148
4	*Aging and Society*	144
5	*Frontiers in Aging Neuroscience*	142
6	*Journals of Gerontology Series A: Biological Sciences and Medical Sciences*	140
7	*Neurobiology of Aging*	133
8	*Journals of Gerontology Series B: Psychological Sciences and Social Sciences*	128
9	*The Gerontologist*	103
10	*Frontiers in Public Health*	99
11	*Educational Gerontology*	95
12	*Experimental Gerontology*	91
13	*Nutrients*	87
14	*Aging and Mental Health*	86
15	*Archives of Gerontology and Geriatrics*	86
16	*BMC Public Health*	86
17	*Psychology and Aging*	86
18	*NeuroImage*	85
19	*Aging Clinical and Experimental Research*	77
20	*Journal of the American Geriatrics Society*	73
21	*Journal of Nutrition, Health and Aging*	72
22	*Scientific Reports*	71
23	*Journal of Aging and Social Policy*	62
24	*Journal of Alzheimer's Disease*	62
25	*Social Science and Medicine*	62
26	*Age and Aging*	61
27	*Journal of Aging and Health*	61
28	*BMJ Open*	60
29	*Frontiers in Psychology*	57
30	*Journal of Applied Gerontology*	56
31	*Gerontology*	53
32	*Aging International*	52
33	*Journal of Aging Studies*	52
34	*Activities, Adaptation and Aging*	51
35	*Aging*	50

Co-Occurrence of Cocited Journals

Figure 3 presents a cocitation map of journals in research applying aging concepts generated using VosViewer software. Table 2 lists the top 30 journals ranked by the number of citations. Together, they show key journals' academic influence and interconnections.

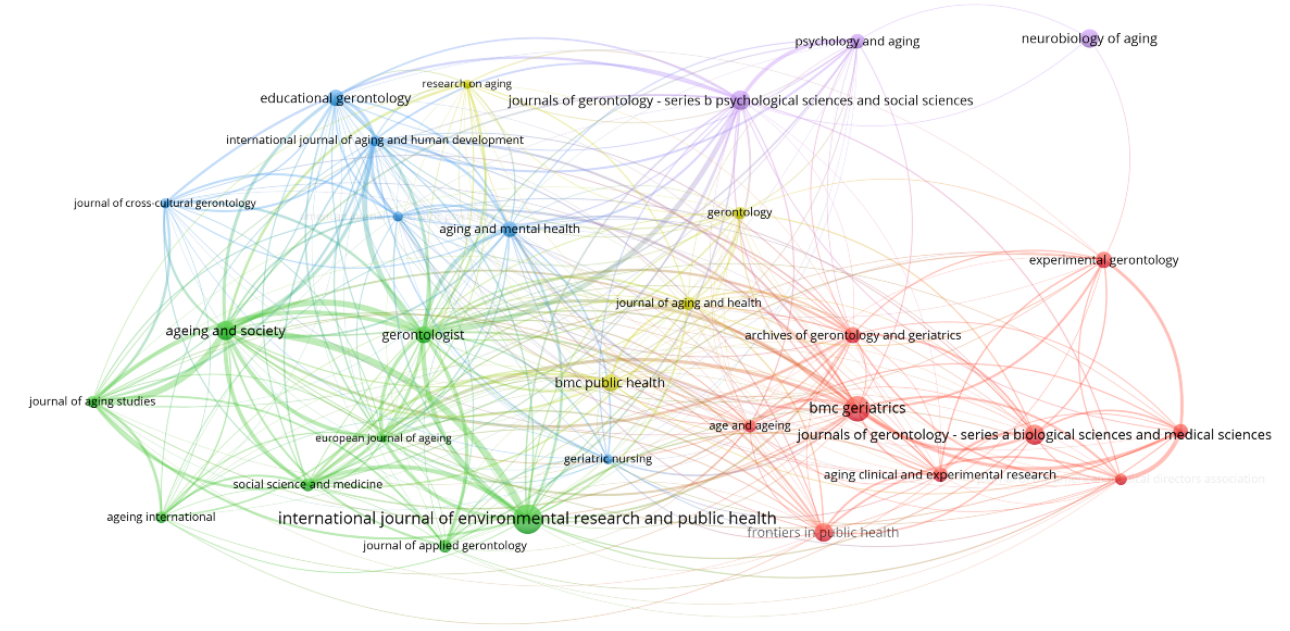
Specifically, larger nodes, such as *Journals of Gerontology—Series A: Biological Sciences and Medical Sciences* (6981 citations), *Gerontologist* (6774 citations),

XSL•FO
RenderX

Journals of Gerontology—Series B: Psychological Sciences and Social Sciences (5016 citations), and *Aging and Society* (4784 citations), indicate that these journals hold considerable academic influence in research applying aging concepts. In addition to the node size, different colors in the network distinguish various research domains, such as social sciences and public health (green), clinical and biological aging (red), cognitive aging and neuroscience (purple), educational and developmental aspects of aging (blue), and a multidisciplinary approach combining medical and social sciences (yellow). Moreover, the thickness of the connecting lines between journals reflects the strength of their citation relationships, quantified by TLS, which indicates how frequently the journals are cocited.

The top 3 journals in terms of TLS are *Gerontologist* (TLS*:* 345), *Aging and Society* (TLS: 317), *and International Journal of Environmental Research and Public Health* (TLS: 230), highlighting strong citation linkages within the research network. As a cumulative metric automatically calculated by VosViewer, TLS represents the total strength of cocitation links a journal has with others. A higher TLS means the journal is frequently and widely cocited, suggesting its centrality in the knowledge network. This clustering of cocitation relationships, dominated by biomedical and gerontology-focused journals, reflects the structural concentration of aging discourse, indicating a limited narrative shift as defined in our methodological framework.

Figure 3. The map of the cocitation network of journals generated by VosViewer.



XSL•FO
RenderX

Table . Cocitations of top 30 journals: ranked by TLS[a].

No.	Journals	Citations, n	TLS
1	*Gerontologist*	6774	345
2	*Aging and Society*	4784	317
3	*International Journal of Environmental Research and Public Health*	3646	230
4	*International Journal of Aging and Human Development*	1544	193
5	*BMC Geriatrics*	3397	183
6	*Social Science and Medicine*	3063	173
7	*Aging and Mental Health*	2526	171
8	*Journals of Gerontology—Series B: Psychological Sciences and Social Sciences*	5016	161
9	*Journal of Aging Studies*	2323	153
10	*European Journal of Aging*	793	145
11	*Archives of Gerontology and Geriatrics*	1763	128
12	*American Journal of Geriatric Psychiatry*	3017	123
13	*Educational Gerontology*	1377	123
14	*Journal of Aging and Health*	1861	112
15	*Journals of Gerontology—Series A: Biological Sciences and Medical Sciences*	6981	108
16	*Journal of Nutrition, Health and Aging*	1858	99
17	*Age and Aging*	1958	93
18	*Journal of Applied Gerontology*	749	91
19	*Journal of Cross-Cultural Gerontology*	725	91
20	*Psychology and Aging*	4752	90
21	*Frontiers in Public Health*	654	87
22	*Journal of the American Medical Directors Association*	2411	82
23	*BMC Public Health*	1974	81
24	*Research on Aging*	1052	80
25	*Aging Clinical and Experimental Research*	1273	79
26	*Geriatric Nursing*	882	79
27	*Aging International*	720	72
28	*Experimental Gerontology*	2890	70
29	*Geriatrics and Gerontology International*	461	64
30	*Gerontology*	1675	64

[a]TLS: total link strength.

XSL•FO
RenderX

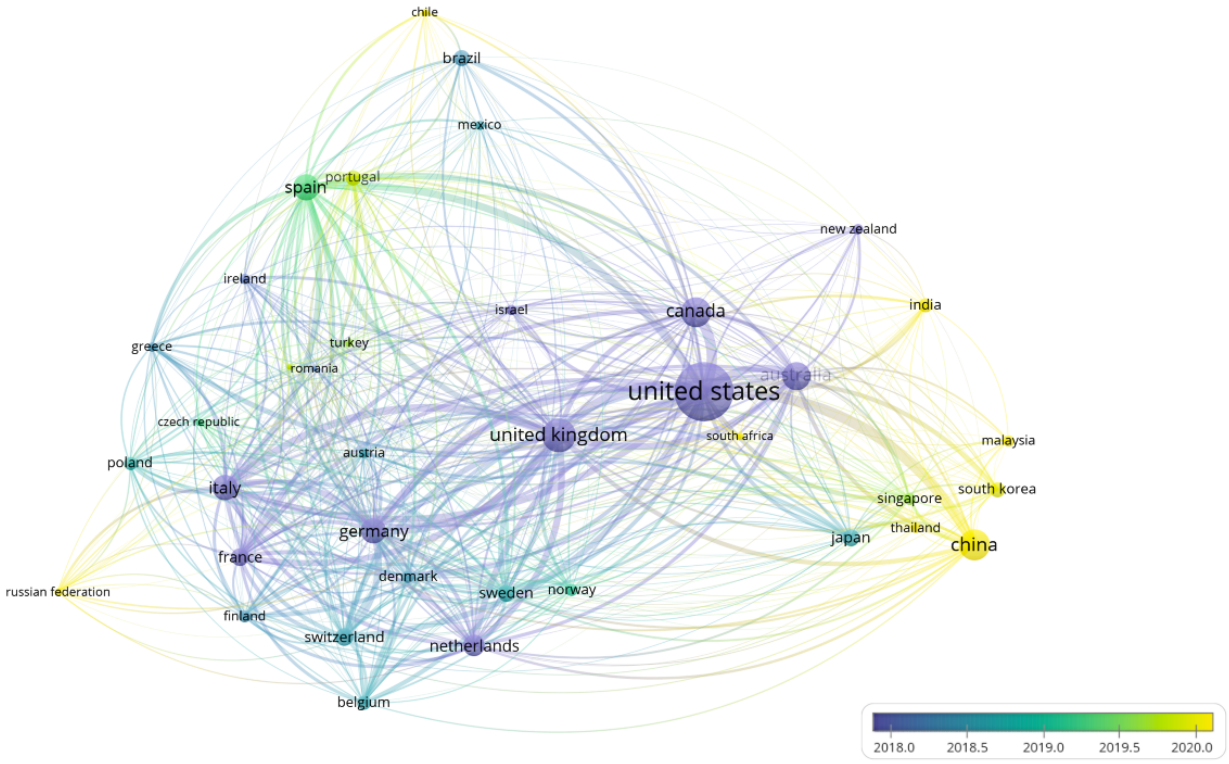
Distributions in Geographical Areas

Figure 4 presents a collaboration map of countries generated by VOSviewer, while the number of papers ranked and the TLS for each country are listed in Table 3. These 2 tables provide a comprehensive overview of global collaboration in research applying aging concepts.

The countries such as the United States, China, and the United Kingdom play leading roles in the global collaboration network. Specifically, the size of the nodes reflects each country's importance in the network, with larger nodes indicating a higher number of published papers. The United States leads with 3262 papers, followed by China with 1447 papers, and the United Kingdom with 1094 papers (Table 3). The lines connecting the nodes represent the collaboration between countries, where thicker lines signify stronger collaborative relationships. This strength is quantified by the TLS, as shown in Table 3. In this regard, the United States, the United Kingdom, and China rank the highest in TLS as well, with values of 1792, 1425, and 779, respectively. In this case, TLS reflects the cumulative strength of coauthorship ties a country maintains with others, helping identify central players in the global landscape of aging research collaboration.

Additionally, in earlier years (eg, around 2017), collaborations were predominantly centered in Western countries such as the United States, Italy, Germany, France, Switzerland, and the Netherlands, as indicated by the blue and purple lines in Figure 4. In contrast, more recent collaborations (eg, around May 2019) are represented by green and yellow lines, showing a growing focus on research applying aging concepts in Asian countries, with China leading the way. This shift has also seen increased collaborations involving countries like South Korea, Thailand, Malaysia, and Singapore, reflecting a broadening of the global research network. Nevertheless, the persistent dominance of high-income Western countries in collaborative centrality suggests an imbalanced global discourse, potentially constraining the diversification of narrative perspectives in aging research.

Figure 4. Global collaboration network map generated by VosViewer.



XSL•FO
RenderX

Table . Top 30 countries: ranked by the number of papers.

No.	Country	Papers, n	Total link strength
1	United States	3262	1792
2	China	1447	779
3	United Kingdom	1094	1425
4	Canada	871	595
5	Australia	776	726
6	Spain	664	800
7	Italy	658	944
8	Germany	635	871
9	Netherlands	464	661
10	France	338	624
11	Switzerland	325	638
12	Sweden	298	532
13	Japan	282	279
14	Brazil	259	175
15	Portugal	242	313
16	South Korea	220	151
17	Belgium	206	375
18	India	200	141
19	Poland	178	261
20	Singapore	168	216
21	Denmark	161	300
22	Norway	160	251
23	Finland	151	324
24	Malaysia	132	82
25	Israel	124	175
26	Ireland	118	185
27	Austria	117	302
28	New Zealand	106	86
29	Greece	103	314
30	Thailand	101	87

Analysis of Aging Concepts: Frequencies, Emergence, and Categorization

Frequencies of 16 Aging Concepts in Aging-Related Research

Multimedia Appendix 2 presents the frequency of aging concepts identified through the corpus linguistic software AntConc, analyzing titles, abstracts, and keywords across 10,642 papers (dataset 1). These frequency data reflect the academic attention given to different aging concepts, and higher frequencies suggest greater scholarly interest, while lower frequencies may indicate emerging or more niche research areas.

The most frequently used aging concept is “healthy aging,” which appears 9710 times and indicates that health is a major focus in aging-related research. Other commonly used aging concepts include “successful aging” (5237 times) and “active aging” (3210 times), reflecting attention on maintaining function and engagement in older age. On the other hand, less frequently mentioned concepts are “robust aging” (8 times) and “conscious aging” (17 times), suggesting that these concepts are less explored or emerging in aging-related research.

Emergence Time of 16 Aging Concepts

Figure 5 presents the years when aging concepts were first explicitly applied to older adult populations in academic literature, revealing their increasing diversity over time. Notably, to explore the differences in the initial emergence time and application of these concepts, a comparison was made between the search results of 16 aging concepts when used as standalone

XSL•FO
RenderX

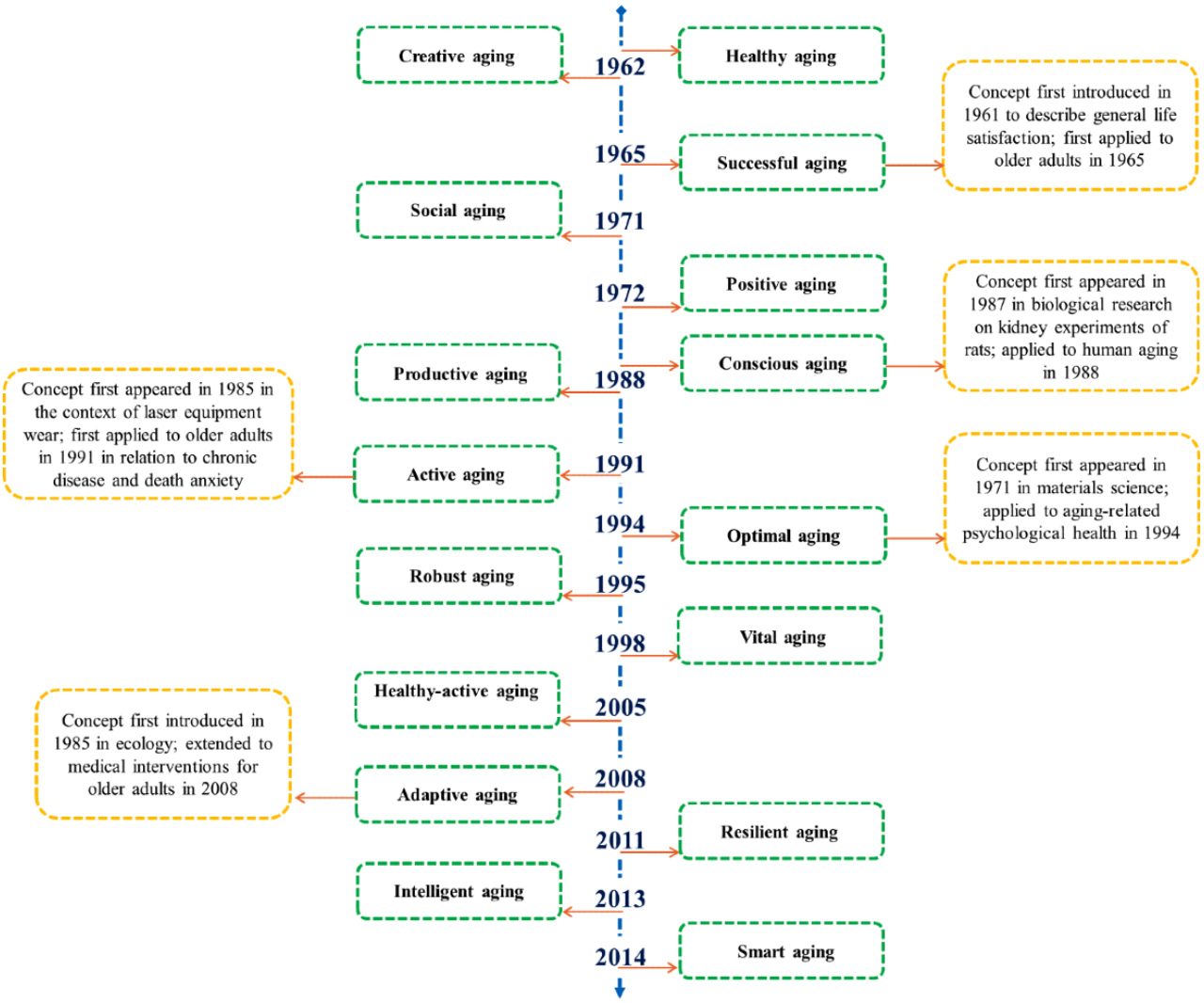
keywords and when combined with keywords related to the aging population. In Figure 5, the emergence time of each aging concept reflects the first documented use in academic literature where the term is explicitly applied to human aging or older adults, rather than its first-ever mention in any unrelated scientific field.

In terms of the time dimension, the earliest aging concepts are "healthy aging" (1962), "creative aging" (1962), and "successful aging" (1965). These concepts were initially introduced from nutrition, gerontology, and medicine perspectives. The first concept was "healthy aging," which originated from a nutritional perspective, focusing on how nutrition can slow down physical decline and prevent diseases to improve the quality of life for older adults [33]. "Creative aging" initially encouraged reflection on the reasons behind the pursuit of longevity, emphasizing that beyond disease management, attention should be given to personal growth and the richness of life experiences [34,35]. The concept of "successful aging" first appeared in 1961 in general psychological literature, but its first explicit use in aging-related research occurred in 1965 [36]. Research in 1965, focusing on the facts of impacting aging for older adults, highlighted the need to distinguish between aspects of aging due to the natural aging process and those more closely related to individual lifestyles or situational factors unrelated to aging [37].

From the 1970s to the 2000s, perspectives on aging concepts gradually expanded, introducing a range of ideas, including social aging in 1971, optimal aging in 1994, positive aging in 1972, adaptive aging in 2008, active aging in 1991, conscious aging in 1988, productive aging in 1988, robust aging in 1995, and vital aging in 1998. Specifically, in 1971, the concept of "social aging" was introduced for the first time, emphasizing that the irreversible loss of an individual's ability to choose what they can do is the most important indicator of aging, whether social or physical aging [38]. That same year, the concept of "optimal aging" also emerged, initially referring to the aging process of metals [39]. It was not until 1994 that the term was applied to older adults, proposing that a positive mental outlook could help them cope constructively with the challenges of aging [40]. In 1972, the term "positive aging" appeared for the first time in the literature, mainly in relation to controlling Hodgkin's disease [41]. In 1985, "adaptive aging" was first mentioned in the field of ecology [42], and by 2008, the concept had been applied in medical research related to older adults [43]. The term "active aging" initially appeared in the literature related to laser equipment wear [44], but was first applied to human aging in 1991, when discussions on chronic disease and death anxiety associated the term with older adults [45]. In 1987, the term "conscious aging" first appeared in biological research related to kidney experiments on rats [46]. The year 1988 is used in the timeline to reflect the earliest application of the term to human aging. In 1988, the concept of "productive aging" emerged, which highlighted how aging amplified economic, political, and moral issues, shifting the focus toward social policy [47]. "Robust aging" was first introduced in 1995, with criteria including greater social interaction, better health and vision, and fewer significant life events for older individuals [48]. Finally, in 1998, the concept of "vital aging" was introduced in psychological health programs, focusing on the development of mental health skills and the effectiveness of changing attitudes, beliefs, and behaviors, providing the motivation and support needed for individuals to make important lifestyle changes and create a more vibrant and productive later life [49].

XSL•FO
RenderX

Figure 5. Timeline of aging concepts first applied to older adult populations in academic research. The main timeline indicates the year when each concept was first applied in academic research, focusing specifically on human aging or older adults. Additional dates in red indicate when the concept first appeared in academic literature, including in nonaging or nonhuman contexts such as materials science, ecology, or animal research.



In the 21st century, concepts like "healthy-active aging" (2005) and "resilient aging" (2011) were introduced, and with the rapid advancement of technology, newer concepts such as "intelligent aging" (2013) and "smart aging" (2014) also emerged. In particular, "healthy-active aging" was first introduced in 2005, suggesting that free health checks and financial advice can be effective for older adults, while knowledge acquisition and social activities enhance their sense of control, especially in impoverished areas [50]. The concept of "resilient aging" was introduced in 2011, offering a useful framework for understanding the mechanisms of aging in the context of disease-related pathologies [51]. "Intelligent aging" emerged in 2013 in response to the growing aging population and shrinking tax base, increasing demand for health care and housing services, while IAServ, a cloud-based platform, aims to provide personalized, scalable, and cost-effective home care through modern technology [52]. In 2014, the concept of "smart aging" was first introduced, emphasizing a positive acceptance of later life and framing aging as a developmental stage toward intellectual maturity [53].

In summary, the years presented in Figure 5 are selected based on a consistent criterion: only the first documented applications explicitly related to human aging or older adult populations are included. Aging concepts have evolved from a narrow focus on health and success to encompass a broader, multidimensional view that includes psychological, social, and technological aspects. Over time, society's understanding of aging has deepened, and future research in this field will likely emphasize interdisciplinary approaches to address the complex challenges of aging. This shift in perspective not only reflects changing societal needs but also provides more comprehensive strategies for dealing with the aging population.

Top 3 Aging Concepts: Publication Growth and Trends Comparison

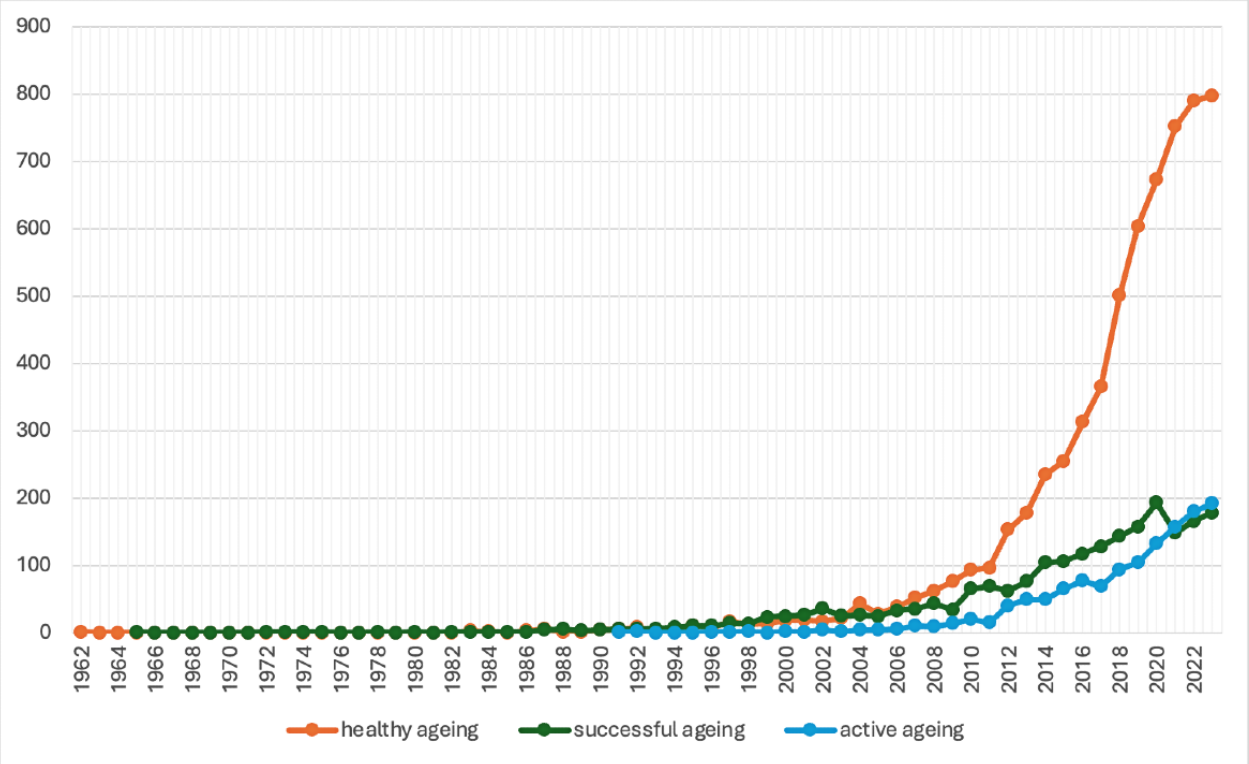
Figure 6 shows the top 3 popular aging concepts among the 16 identified: "healthy aging," "successful aging," and "active aging," reflecting significant growth from 1962 to 2023, but at varying rates. "Healthy aging" saw a sharp rise in attention after 2010, with a particularly rapid increase after 2020, becoming the dominant aging concept mentioned in research. On the other hand, "successful aging" and "active aging" gained more steady attention starting from the 1990s. "Active aging" experienced slightly faster growth after 2000, and especially after 2019,

XSL•FO
RenderX

surpassing "successful aging" in growth rate, reflecting an increased emphasis on older adults' participation in social activities and maintaining an active lifestyle.

This trend suggests that while "healthy aging" has become the dominant concept, the steady growth of "active aging," particularly its acceleration after 2019, highlights an increasing emphasis on the importance of older adults actively participating in society and maintaining an engaged lifestyle. Compared to other more established concepts, there is likely to be a greater future focus on leveraging "active aging" to improve the quality of life for older adults, encouraging their active social integration and promoting their overall well-being.

Figure 6. Successful aging, healthy aging, and active aging concepts: publication trends over time (1961-2023). Frequency of 16 aging concepts in achieving "age well."



Multimedia Appendix 3 shows the frequency of 16 aging concepts in the text of titles, abstracts, and keywords across 1402 papers (dataset 2: searched in the Scopus database using the keywords "age well" OR "aging well" OR "ageing well" limited to English-language journal papers), identified through the corpus linguistic software AntConc. This comparison evaluates how effectively the identified aging concepts from dataset 1 are reflected in the more naturalistic academic discourse surrounding "aging well," thereby indicating their narrative integration and relevance across different contexts. The comparison of the use frequency of different aging concepts reveals that "successful aging (417 times)" is the popular aging concept in discussions of how to achieve "age well," followed closely by "healthy aging (247 times)" and "active aging (180 times)."

Additionally, less frequently mentioned concepts, such as "positive aging" and "productive aging," are mentioned 70 and 14 times, indicating that while less prominent, there is still interest in aging's psychological and social contribution aspects. Aging concepts such as "adaptive aging" and "resilient aging" appear very rarely, with just 3 and 1 occurrences, suggesting that these concepts, though relevant to coping with aging challenges, have not yet gained significant traction in the literature. Likewise, concepts like "intelligent aging" and "smart aging" were not found, reflecting that while there may be much research into how technology supports older adults in adapting to their environments, they have yet to become mainstream but may hold potential for future exploration.

Categorization and Core Focus of 16 Aging Concepts

Table 4 summarizes 16 aging concepts, detailing their core focus and application areas and categorizing them into 4 groups, including (1) health and well-being, (2) social engagement, (3) technology and innovation, and (4) resilience and adaptation. Analyzing these concepts reveals how the academic community prioritizes different aspects of aging and applies them in practice. Although all the concepts share the common goal of improving the quality of life for older adults, each contributes a distinct perspective to the field of aging-related research.

The concepts of aging, such as healthy aging, successful aging, optimal aging, positive aging, and vital aging, can be categorized as part of the Health and Well-being Aspect. These concepts aim to promote health and overall well-being, yet they approach this goal from varying angles. Healthy aging primarily promotes overall health through public health initiatives and disease prevention. Successful aging highlights the importance of maintaining both physical and mental health, focusing on how psychological well-being and social support enhance quality of

XSL•FO
RenderX

life. Optimal aging takes a broader approach, addressing physical, emotional, and cognitive health, aiming to extend a healthy lifespan through community well-being and health promotion programs. Positive aging focuses more on mental well-being, such as maintaining a positive attitude and emotional resilience. Finally, vital aging stresses lifelong adaptation for their well-being, encouraging older adults to remain independent and active through lifelong learning and adaptability.

XSL•FO
RenderX

Table . Comparisons of the core focus of 16 aging-related concepts.

Categorizations and aging concepts	Core focus	Application areas
Health and well-being aspect		
Healthy aging	Focuses on achieving a high quality of life through promoting overall health and preventing disease during the aging process.	• Public health • Older people care • Disease prevention
Successful aging	Focuses on maintaining physical and mental health in later life and enhancing quality of life.	• Older people health management • Mental health • Social policy
Optimal aging	Focuses on promoting health, longevity, and overall well-being by addressing physical, emotional, and cognitive aspects of aging, ensuring individuals remain independent and active as they age.	• Health promotion • Community well-being • Emotional and sensory wellness • Public health programs • Evidence-based information • Dissemination
Positive aging	Focuses on the importance of maintaining a positive attitude, psychological resilience, and emotional well-being.	• Mental health • Emotional well-being • Psychological therapy
Vital aging	Emphasizes lifelong adaptation to maximize health, independence, and cognitive and emotional well-being, promoting active and successful aging through various programs and interventions.	• Lifelong adaptation • Mental fitness • Lifelong learning • Environmental policies • Cultural context • Community-based programs
Creative aging	Focuses on enhancing the psychological and social well-being of older adults through artistic and cultural activities.	• Art or culture engagement • Psychological therapy
Social engagement aspect		
Active aging	Focuses on the active participation of older adults in social activities and maintaining physical vitality.	• Social participation • Community activities • Sports and recreation
Social aging	Focuses on the roles and interactions of older adults in society, examining how social, economic, and cultural factors influence the aging process.	• Sociological research • Social policy • Analysis of older adults' social roles
Productive aging	Focuses on the active participation of older adults in socially productive activities, such as volunteering, caregiving, and work, contributing to society while enhancing individual well-being.	• Community volunteering • Social support • Health literacy • Cultural differences • Government policies • Workforce participation • Lifelong learning
Conscious aging	Emphasizes deep reflection on the aging process, spiritual growth, and contemplation of life.	• Spiritual requirements • Lifelong learning • Personal growth
Health and well-being and social engagement aspect		
Healthy-active aging	Combines elements of healthy aging and active aging to maintain vitality and health in older adults.	• Health promotion • Community health • Older people care
Technology and innovation aspect		

XSL•FO
RenderX

Categorizations and aging concepts	Core focus	Application areas
Smart aging	Focuses on supporting the quality of life of older adults through smart technologies such as virtual reality, data privacy protection, and smart cities.	• Smart homes • Health monitoring • Smart city development
Intelligent aging	Mainly involves the application of smart technology in health monitoring and management.	• Smart health monitoring • Personalized health care • AI[a]-assisted living
Resilience and adaptation aspect		
Adaptive aging	Emphasizes optimizing design and infrastructure to accommodate the physical and cognitive changes associated with aging.	• Public infrastructure design • Product design • Biomedical research
Resilient aging	Emphasizes the adaptability and psychological resilience of older adults in the face of challenges.	• Mental health • Coping strategies • Crisis management
Robust aging	Focuses on maintaining vitality across physical, emotional, cognitive, and social domains, with an emphasis on resilience and adaptability throughout the aging process.	• Productive involvement • Affective status • Functional status • Cognitive status • Physical activity • Social engagement • Cellular biomarkers • Cross-generational aging

[a]AI: artificial intelligence.

In a related area, aging concepts of active aging, social aging, productive aging, creative aging, and conscious aging all fall under the aspect of social engagement, but they each emphasize different areas. Active aging encourages older adults to stay involved in social and physical activities, promoting vitality. Social aging examines the social roles of older adults and the impact of social, economic, and cultural factors. Productive aging highlights older adults' contributions to society through activities like volunteering and work, enhancing well-being. Creative aging improves emotional health through engagement in art and culture, while conscious aging emphasizes personal and spiritual growth through reflection on life. Notably, healthy-active aging belongs to both healthy-active aging and social engagement aspects because it combines elements of both healthy and active lifestyles, supporting older adults in maintaining vitality and well-being by staying physically and mentally active.

The technology and innovation aspect encompasses aging concepts like smart aging and intelligent aging, both of which address the role of technology in supporting older adults. Although both concepts highlight the importance of smart technologies, they diverge in their specific applications. Smart aging focuses on enhancing daily life through technologies such as smart homes and health monitoring systems, while intelligent aging integrates artificial intelligence (AI) and data science to provide personalized health care solutions, illustrating the increasing role of AI in aging research. While both are grounded in technology-based theories, intelligent aging emphasizes more prominently data-driven approaches and AI.

Within the resilience and adaptation aspect, concepts like resilient aging, adaptive aging, and robust aging focus on older adults' capacity to adjust to the changes accompanying aging. Resilient aging highlights psychological resilience, emphasizing older adults' ability to cope with challenges. On the other hand, adaptive aging explores how design and infrastructure can be optimized to accommodate physical and cognitive changes in aging. Robust aging takes a more comprehensive approach, incorporating resilience across physical, cognitive, emotional, and social domains and integrating biological markers to assess longevity and health outcomes.

Challenges and Future Trends of Achieving Aging Concepts

The 16 aging concepts reflect the academic community's vision of promoting "aging well" from various perspectives. Through content analysis, Table 5 outlines the challenges in achieving the goal of each aging concept and future trends in applying each aging concept. Although the challenges—such as unequal access to resources, technological barriers, and cultural differences—are significant, there are promising opportunities for innovative and sustainable solutions. Advances in technology and a growing global focus on aging will create opportunities for policy makers, researchers, and communities to collaborate in addressing these challenges. By doing so, older adults worldwide can anticipate a safer, healthier, and more fulfilling aging process, regardless of their social, cultural, or economic backgrounds.

XSL•FO
RenderX

Table . Challenges and future work of 16 aging-related concepts.

Aging concepts	Challenges in achieving the goal of each aging concept	Future trends in applying each aging concept
Healthy aging	• Lack of interdisciplinary collaboration and long-term investment • Unequal access to medical resources globally	• Greater integration of healthy aging initiatives in public health policies • Growing focus on preventive health management due to accelerated global population aging
Successful aging	• Challenges in defining "success" across diverse cultural backgrounds • Difficulty in setting a universal standard due to significant variation in individual health status	• Increasing emphasis on social policies aimed at supporting older adults to maintain a high quality of life in their later years • Expanded community support, health care, and mental health services for older adults
Active aging	• Challenges of physical and environmental barriers for sustained participation • Difficulty for older adults in economically and infrastructurally disadvantaged areas	• Advances in technology enabling digital activities and remote participation as new means for older adults to engage socially • Increased participation of older adults in social activities through web-based communities and digital courses
Positive Aging	• Challenges due to negative stereotypes of aging hindering broad acceptance • Limited resources to address mental health issues in older adults across sectors	• Increasing awareness of mental health leading to greater focus on the mental health and emotional well-being of older adults • More mental health services and support programs emerging to address the needs of older adults
Smart aging	• Barriers due to high costs of technology application and unresolved data privacy issues • Challenges in older adults' acceptance of technology	• Further development and widespread adoption of smart homes, wearable devices, and health monitoring technologies • Enhanced safety and independence for older adults living at home through advanced technologies
Social aging	• Challenges arising from differences in social structure and culture for older adults • Difficulty in developing universal solutions due to varying social challenges across countries and communities	• Increased focus on aging issues leading to more policies and research on role transitions of older adults in society • Emphasis on cultural and social support to enhance social participation of older adults
Creative aging	• Insufficient resources for creative activities in certain areas • Limited opportunities for older adults in remote or resource-scarce communities	• Remote art programs and digital creative workshops providing opportunities for older adults to engage in creative activities when in-person participation is not possible • Growing importance of digital platforms postpandemic facilitating creative engagement for older adults
Adaptive aging	• Challenges due to the diversity and complexity of older adults' needs for product design • Significant time and financial resources required to promote and apply adaptive design in public infrastructure	• Increasing focus on adaptive and accessible design in the construction of future cities and public infrastructure • Ensuring safe and comfortable environments for older adults through improved urban and infrastructure design
Resilient aging	• Challenges due to considerable variability in individual psychological resilience • Imbalances in social support systems affecting older adults' ability to build resilience across countries and communities	• Fostering psychological resilience in older adults becoming a priority amid increasing global social and environmental crises • Further development of psychological support programs and community support networks for older adults

XSL•FO
RenderX

Aging concepts	Challenges in achieving the goal of each aging concept	Future trends in applying each aging concept
Healthy-active aging	• Challenges in coordinating resources across multiple fields for multidimensional health • Difficulty in achieving health goals in economically underdeveloped or resource-limited regions	• Expansion of community health programs, particularly through digital health platforms • Promotion of healthy behaviors among older adults via digital health initiatives
Intelligent aging	• Challenges due to high cost and complexity of technology limiting its application range • Difficulty in acceptance of smart technologies by older adults	• Increasing role of AI[a] and big data in personalized health management • Emergence of more AI-based health solutions for older adults
Conscious aging	• Challenges due to varying cultural importance placed on the spiritual life of older adults • Limited societal attention to spiritual health compared to physical health hindering development in this area	• Increasing global attention to mental and spiritual health leading to greater promotion of spiritual support programs for older adults • More widespread adoption of practices such as meditation and yoga among older adults
Productive aging	• Challenges due to disparities in access to social support and education • Varying cultural perspectives on aging as a barrier • Unequal access to workforce participation opportunities for older adults	• Increasing promotion of productive aging through policy initiatives, focusing on the integration of older adults in the community and workforce • Key trends of lifelong learning and flexible work opportunities for older adults
Robust aging	• Challenges due to variability of aging experiences across generations • Disparities in access to health care and social support systems • Differences in biomarker availability and interpretation	• Advancements in biotechnology and understanding of cellular aging driving robust aging strategies toward personalized health interventions • Importance of cross-generational health programs and adaptive aging practices in future public health initiatives
Optimal aging	• Challenges due to gender disparities in aging experiences • Unequal access to health care and wellness programs • Need for broader dissemination of reliable health information	• Growing emphasis on evidence-based health interventions and gender-specific aging strategies • Increasing focus on optimal aging through personalized health plans, community support systems, and stress management techniques
Vital aging	• Challenges due to disparities in educational opportunities • Environmental inequalities impacting aging experiences • Cultural differences in aging experiences across regions • Accessibility issues of programs in different regions	• Focus on expanding vital aging programs globally, emphasizing integration of environmental policies • Attention to mental fitness initiatives to enhance aging experiences across different cultural contexts

[a]AI: artificial intelligence.

One of the primary challenges in achieving the goal of these aging concepts is the global disparity in medical and health care resources, particularly in economically disadvantaged regions where access to medical services is limited. This inequality hampers older adults' ability to maintain both physical and mental health and benefit from preventive health care strategies. In the future, public health efforts should focus on integrating preventive health care with personalized interventions, ensuring that older adults have access to the necessary resources to support their well-being, regardless of location.

In addition to health care challenges, promoting social participation among older adults presents another major obstacle. Factors such as social structures, economic development, and cultural differences often create barriers to maintaining social connections and engaging with communities. Many societies prioritize physical health while neglecting emotional and spiritual needs. Cultural differences also influence the emphasis placed on spiritual life, making it difficult to implement universally applicable support systems. However, with growing global awareness of mental health, more programs are likely to address the emotional and spiritual well-being of older adults.

XSL•FO
RenderX

Practices such as meditation, yoga, and other spiritual activities are expected to become more widespread, fostering a sense of fulfillment and psychological balance. Additionally, in many regions, older adults struggle to fully participate in social life due to a lack of flexible employment options and inadequate lifelong learning opportunities, further limiting their engagement. Balancing physical health with mental and spiritual well-being remains crucial. To improve the quality of life for older adults, future policies should focus on strengthening social support networks and providing meaningful opportunities for social participation, which in turn will enhance their mental health and sense of belonging.

Adopting technology to improve older adults' lives also presents unique challenges. While technological solutions like smart homes, wearable devices, and health monitoring systems hold great potential, their high costs and complexity often limit widespread adoption, particularly in resource-constrained regions. Privacy concerns further complicate the integration of these technologies into everyday life. Future technological developments must focus on reducing costs, improving accessibility, and ensuring data security to make these tools more inclusive. Additionally, creative activities remain unevenly distributed, especially in remote areas, which are key to promoting older adults' emotional and social well-being. Expanding web-based platforms will provide older adults with greater access to creative opportunities, helping bridge the gap between regions with varying resource availability.

Finally, fostering older adults' adaptability to changing environments and promoting psychological resilience is crucial for their overall well-being. Developing these qualities requires significant time and financial investment. Public infrastructure and services must be optimized to meet the diverse needs of older adults, ensuring they can live in safe and comfortable environments. Building community support systems and mental health programs will also be critical in fostering psychological resilience, especially as global social and environmental crises increasingly affect vulnerable populations. Future urban planning efforts should prioritize adaptive design to create environments that support older adults' physical, mental, and emotional health.

These challenges and future trends were derived from the content analysis process and thematically coded based on recurring contextual patterns. They also provide insight into the limitations of existing aging narratives and help identify conceptual gaps that could inform a more inclusive and multidimensional shift in aging discourse.

Discussion

Principal Findings

This study found that although a wide range of aging concepts have been introduced in academic literature, their overall impact on reframing negative aging narratives remains limited. The analysis revealed that most applications of these concepts continue to be concentrated in biomedical fields, with relatively less integration of social and economic dimensions. While some concepts promote multidimensional perspectives, their practical influence on shifting aging discourse is still fragmented and uneven. The demographic shift toward an aging population has led to rising concerns about "ageism," which reinforces stereotypes of older adults as cognitively and physically declining, resistant to change, or burdensome to society [54,55]. These views contribute to social exclusion and negatively affect older adults' mental and physical well-being. Academic research plays a key role in challenging such narratives by adopting a more comprehensive view of aging that encompasses physical, cognitive, emotional, social, and economic dimensions [56,57]. Recognizing older adults' contributions—through work, caregiving, community participation, and cultural knowledge—can reshape how society perceives aging and promote age-inclusive policies [58-60].

The primary goal of this study is to explore whether and how research applying aging concepts contributes to shifting the negative narrative on aging in academia. The overall results show that while great efforts have been made, progress has been limited. As outlined in the Methods section, the dominance of biomedical framing and the lack of social and economic integration indicated a limited narrative shift. While the bibliometric analysis reveals a significant increase in aging concept research across disciplines, the concentration in medical fields and minimal attention to social and economic aspects still reflect an uneven narrative shift. Although many of these studies aim to promote "aging well," they often take a narrow approach, focusing mainly on medical aspects such as disease control, health maintenance, and longevity. This limited perspective, while contributing to a partial change in the perception of aging, can inadvertently reinforce stereotypes (thinking: equating aging with poor health), prejudice (feeling: viewing older adults as less capable), and discrimination (acting: excluding older adults from work and daily life) [61]. Aging is a natural part of life and requires medical support across the human lifespan, but a holistic perspective on aging should extend beyond medical aspects to include social, economic, and other dimensions, generating a more balanced and comprehensive understanding of aging [62-64].

While the emergence of diverse aging concepts reflects an aspirational effort to counter negative narratives, our findings suggest that their actual impact has been uneven and often limited. Several factors may help explain this discrepancy. Many aging concepts lack consistent definitions, making them difficult to operationalize across disciplines and policy settings. The continued dominance of a biomedical focus further restricts their narrative potential. Furthermore, some concepts—such as "successful" or "active" aging—have been critiqued for reinforcing idealized standards or placing the responsibility of healthy aging on individuals, which may unintentionally perpetuate ageist assumptions. These limitations highlight the gap between conceptual intention and actual discursive impact.

These findings can be further interpreted through the lens of narrative theory and the social construction of aging, which suggests that aging concepts do not merely describe reality but actively shape how aging is perceived and valued in society. This theoretical perspective helps explain the persistence of deficit-based frames, and the limited extent of narrative change remains despite growing conceptual diversification.

XSL•FO
RenderX

At the same time, however, research involving these aging concepts has also acknowledged the multidimensional nature of aging and highlighted the need to address it holistically in both research and policy. The content analysis of this study categorized all 16 aging concepts into four groups based on their main focus and application, contributing to a more comprehensive understanding: (1) health and well-being (eg, healthy aging, successful aging, optimal aging, positive aging, vital aging, and creative aging); (2) social engagement (eg, active aging, social aging, productive aging, and conscious aging); (3) technology and innovation (eg, smart aging and intelligent aging); and (4) resilience and adaptation (eg, adaptive aging, resilient aging, and robust aging). Although all these concepts are built upon the foundation of health, their differing emphases reflect the multidimensional nature of aging research in academia. For instance, concepts centered on health and well-being focus on enhancing physical health and quality of life, while social engagement concepts underscore the social roles and contributions of older adults.

Nevertheless, despite these expanded areas of focus, practical applications often remain fragmented, limiting their broader impact on societal perceptions of aging. Specifically, the content analysis in this study, which also explored challenges and future trends, revealed significant obstacles to achieving the goals of aging concepts, such as unequal access to health care, technological barriers, and cultural differences. For example, “smart aging” initiatives leveraging digital technology, though beneficial for enhancing independence, face barriers such as limited access and high costs, restricting their reach [65]. Similarly, resilience-oriented concepts that promote coping strategies for psychological and physical challenges often lack widespread implementation and do not significantly influence public policy or community practices [66]. These findings underscore the complexity of effectively implementing aging concepts and point to potential pathways for progress.

To drive a meaningful shift in the negative narrative on aging, interdisciplinary strategies are essential, bridging current gaps to create a cohesive and impactful transformation [67]. Several successful approaches from other domains offer promising models for adaptation in aging research. For example, universal design in urban planning can inform inclusive spaces and housing that support physical and cognitive aging. Lifelong learning platforms from the education sector can promote digital literacy and social engagement. Resilience-building frameworks, developed in clinical psychology and disaster management, offer coping strategies that can be tailored to later-life transitions. Integrating such practices into aging-focused interventions can help bring empowerment-based narratives into real-world application.

These strategies must ensure that technological advancements and global awareness translate into tangible benefits that are accessible, affordable, and inclusive for all [68,69]. Future policies should adopt a nuanced understanding of aging by fostering integrated research that unites medical, social, and technological aspects, thus enabling a holistic perspective that can reshape both academic and societal narratives on aging. For instance, practical strategies like expanding social networks, offering flexible employment, and fostering lifelong learning opportunities are crucial, as they not only enhance older adults’ sense of belonging and mental well-being but also drive a transformative shift in societal perceptions and approaches to aging [70-72]. Without such efforts, there is a risk of reinforcing existing inequalities rather than resolving them. As indicated by the cocitation and collaboration analyses, current aging research remains structurally concentrated within biomedical networks and Western regions, suggesting that future narrative shifts require more interdisciplinary and globally distributed scholarly engagement.

Limitations and Future Directions

While this study offers valuable insights for both researchers and policy makers, it has certain limitations. The exclusive reliance on the Scopus database may have led to the omission of relevant studies from other sources. Additionally, although multiple analytical tools such as VosViewer and AntConc were used, differences in their functionalities may introduce interpretation biases. Furthermore, this study did not fully explore these aging concepts’ roles in policy-making and cross-cultural contexts, leaving room for future research in these important areas.

Future research could address several areas identified in this study. First, expanding the data sources to include additional databases (eg, Web of Science or PubMed) may broaden disciplinary representation. Second, exploring how aging concepts are implemented and interpreted in policy settings would help assess their practical influence. Third, cross-cultural comparisons could reveal differences in how aging narratives are constructed globally. Finally, mixed method approaches combining bibliometric, linguistic, and qualitative data could provide a more comprehensive understanding of both academic and societal impacts of aging discourse.

Conclusions

This study comprehensively analyzes 16 aging concepts, elucidating their main focuses, applications, and the challenges and future trends associated with their implementation. The analysis reveals an overall picture of the academic landscape, including the distribution of journals and disciplines, as well as cocitation networks and regional collaborations. “Healthy aging” emerged as the most dominant concept according to its frequency in the 2 datasets. A content analysis further categorized the 16 concepts into four primary groups based on focus and application, enhancing our understanding of their roles: (1) health and well-being (eg, healthy aging, successful aging, optimal aging, positive aging, vital aging, and creative aging); (2) social engagement (eg, active aging, social aging, productive aging, and conscious aging); (3) technology and innovation (eg, smart aging and intelligent aging); and (4) resilience and adaptation (eg, adaptive aging, resilient aging, and robust aging). Notably, “healthy aging” and “creative aging” were the earliest concepts, first appearing in 1962, marking the beginning of an evolving discourse on aging. While these concepts collectively aim to enhance the quality of life for older adults and reshape the negative narrative on aging, progress has been uneven, with many efforts remaining siloed and largely focused on medical aspects. Although there has been a shift toward more holistic perspectives, including social engagement,

XSL•FO
RenderX

resilience, and technological adaptation, these approaches have yet to achieve widespread and integrated impact.

This classification was developed through systematic content analysis, grouping concepts by their primary thematic focus and area of application. It serves not only as a descriptive framework but also offers practical relevance. For researchers, it highlights patterns across disciplines and identifies gaps for future study. For policymakers, it enables more targeted approaches—for instance, by aligning interventions with specific conceptual clusters such as promoting social participation or enhancing technological access for older adults. This structured understanding of the aging discourse can guide more coordinated, interdisciplinary strategies in both research and policy development.

This study contributes uniquely to the existing literature in several ways. First, it provides a comprehensive cross-conceptual comparison of 16 aging frameworks, offering a panoramic view that is rarely achieved in prior studies focused on single concepts. Second, by integrating bibliometric mapping with corpus linguistic analysis, it introduces a multimethod approach that captures both the structural distribution and the narrative functions of aging discourse. Third, the study applies a theoretically grounded lens—drawing on narrative theory and the social construction of aging—to assess how conceptual framing influences the persistence or transformation of age-related narratives. Together, these contributions advance a more integrated and reflexive understanding of how academic discourse shapes societal perspectives on aging.

Significant challenges persist, including unequal access to health care, technological barriers, and cultural differences, which complicate the effective implementation of these aging concepts. Addressing these challenges requires coordinated strategies that ensure accessibility, affordability, and inclusivity, translating technological advancements and global awareness into tangible benefits for all older adults. Future policies must be informed by a nuanced understanding of aging, emphasizing preventive health care, personalized interventions, and social and psychological support. Practical measures like expanding social networks, offering flexible employment, and fostering lifelong learning opportunities are essential for promoting a sense of belonging and enhancing mental well-being among older adults.

Data Availability

The datasets analyzed during this study are available through the Scopus database (Elsevier, Amsterdam, Netherlands). The data are publicly accessible to researchers with institutional access to Scopus. The studies reported in the manuscript were not preregistered.

Authors' Contributions

Conceptualization: NX and BX.
Methodology: NX, BX, and CS.
Investigation: NX, BX, and ML.
Formal analysis: NX and ML.
Supervision: BX and LB.
Writing—original draft: NX.
Writing—review and editing: BX, LB, CS, ML, QC, and NX.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Disciplinary distribution of publications.
[PNG File, 184 KB - aging_v8i1e72011_app1.png]

Multimedia Appendix 2
Frequency of 16 aging-related concepts by using AntConc.
[DOCX File, 15 KB - aging_v8i1e72011_app2.docx]

Multimedia Appendix 3
Comparisons of the frequency of 16 aging concepts in achieving “Age Well.”
[DOCX File, 14 KB - aging_v8i1e72011_app3.docx]

References

1. Lutz W, Sanderson W, Scherbov S. The coming acceleration of global population ageing. Nature New Biol 2008 Feb 7;451(7179):716-719. [doi: 10.1038/nature06516] [Medline: 18204438]

XSL•FO
RenderX

2. World population prospects 2024: summary of results. : United Nations; 2024 URL: https://www.ijhpm.com/article_3254.html [accessed 2025-09-19]
3. Maj-Waśniowska K, Jedynak T. The issues and challenges of local government units in the era of population ageing. Adm Sci 2020 Jun 18;10(2):36. [doi: 10.3390/admsci10020036]
4. Bloom DE, Chatterji S, Kowal P, et al. Macroeconomic implications of population ageing and selected policy responses. Lancet 2015 Feb;385(9968):649-657. [doi: 10.1016/S0140-6736(14)61464-1]
5. Gahan P, Harbridge R, Healy J, Williams R. The ageing workforce: policy dilemmas and choices. Aust J Public Adm 2017 Dec;76(4):511-523. [doi: 10.1111/1467-8500.12232]
6. Mosca I, van der Wees PJ, Mot ES, Wammes JJG, Jeurissen PPT. Sustainability of long-term care: puzzling tasks ahead for policy-makers. Int J Health Policy Manage 2016 Aug 17;6(4):195-205. [doi: 10.15171/ijhpm.2016.109]
7. Sanz I, Velázquez FJ. The role of ageing in the growth of government and social welfare spending in the OECD. Eur J Polit Econ 2007 Dec;23(4):917-931. [doi: 10.1016/j.ejpoleco.2007.01.003]
8. Dillin A, Gottschling DE, Nyström T. The good and the bad of being connected: the integrons of aging. Curr Opin Cell Biol 2014 Feb;26:107-112. [doi: 10.1016/j.ceb.2013.12.003] [Medline: 24529252]
9. Maresova P, Javanmardi E, Barakovic S, et al. Consequences of chronic diseases and other limitations associated with old age—a scoping review. BMC Public Health 2019 Nov 1;19(1):1431. [doi: 10.1186/s12889-019-7762-5] [Medline: 31675997]
10. Pawelec G, Goldeck D, Derhovanessian E. Inflammation, ageing and chronic disease. Curr Opin Immunol 2014 Aug;29:23-28. [doi: 10.1016/j.coi.2014.03.007] [Medline: 24762450]
11. Ward SA, Parikh S, Workman B. Health perspectives: international epidemiology of ageing. Best Pract Res Clin Anaesthesiol 2011 Sep;25(3):305-317. [doi: 10.1016/j.bpa.2011.05.002] [Medline: 21925398]
12. Harper S. Economic and social implications of aging societies. Science 2014 Oct 31;346(6209):587-591. [doi: 10.1126/science.1254405] [Medline: 25359967]
13. Kudo S, Mutisya E, Nagao M. Population aging: an emerging research agenda for sustainable development. Soc Sci (Basel) 2015 Dec;4(4):940-966. [doi: 10.3390/socsci4040940]
14. Maestas N, Mullen KJ, Powell D. The effect of population aging on economic growth, the labor force, and productivity. Am Econ J Macroecon 2023 Apr 1;15(2):306-332. [doi: 10.1257/mac.20190196]
15. Powell JL. The power of global aging. Ageing Int 2010 Mar;35(1):1-14. [doi: 10.1007/s12126-010-9051-6]
16. Foster L, Walker A. Active and successful aging: a European policy perspective. Gerontologist 2015 Feb;55(1):83-90. [doi: 10.1093/geront/gnu028] [Medline: 24846882]
17. Bauman A, Merom D, Bull FC, Buchner DM, Singh MAF. Updating the evidence for physical activity: summative reviews of the epidemiological evidence, prevalence, and interventions to promote “Active Aging”. Gerontologist 2016 Apr;56(Suppl 2):S268-S280. [doi: 10.1093/geront/gnw031]
18. Cosco TD, Prina AM, Perales J, Stephan BCM, Brayne C. Operational definitions of successful aging: a systematic review. Int Psychogeriatr 2014 Mar;26(3):373-381. [doi: 10.1017/S1041610213002287] [Medline: 24308764]
19. Estebsari F, Dastoorpoor M, Khalifehkandi ZR, et al. The concept of successful aging: a review article. Curr Aging Sci 2020;13(1):4-10. [doi: 10.2174/1874609812666191023130117] [Medline: 31657693]
20. Lu W, Pikhart H, Sacker A. Domains and measurements of healthy aging in epidemiological studies: a review. Gerontologist 2019 Jul 16;59(4):e294-e310. [doi: 10.1093/geront/gny029] [Medline: 29897451]
21. Tavares RE, Jesus MD, Machado DR, Braga VAS, Tocantins FR, Merighi MAB. Healthy aging from the perspective of the elderly: an integrative review. Rev Bras Geriatr Gerontol 2017;20(6):878-889. [doi: 10.1590/1981-22562017020.170091]
22. Bowling A. Enhancing later life: how older people perceive active ageing? Aging Ment Health 2008 May;12(3):293-301. [doi: 10.1080/13607860802120979] [Medline: 18728941]
23. Holstein MB, Minkler M. Self, society, and the “new gerontology”. Gerontologist 2003 Dec;43(6):787-796. [doi: 10.1093/geront/43.6.787] [Medline: 14704376]
24. Langmann E, Weßel M. Leaving no one behind: successful ageing at the intersection of ageism and ableism. Philos Ethics Humanit Med 2023 Nov 24;18(1):22. [doi: 10.1186/s13010-023-00150-8] [Medline: 38001533]
25. Levy SR, Apriceno M, Department of Psychology, Stony Brook University, USA. Ageing: the role of ageism. OBM geriatr 2019 Jun 20;3(4):1-1. [doi: 10.21926/obm.geriatr.1904083]
26. Waddell C, Van Doorn G, Power G, Statham D. From successful ageing to ageing well: a narrative review. Gerontologist 2024 Dec 13;65(1):gnae109. [doi: 10.1093/geront/gnae109] [Medline: 39139052]
27. Krzyminiewska G. Ethical dilemmas of the silver economy. EiP 2020 Mar 31;19(1):61. [doi: 10.12775/EiP.2020.005]
28. Lipp B, Peine A. Ageing as a driver of progressive politics? What the European Silver Economy teaches us about the co-constitution of ageing and innovation. Ageing Soc 2024 Jul;44(7):1481-1493. [doi: 10.1017/S0144686X22000903]
29. Montross LP, Depp C, Daly J, et al. Correlates of self-rated successful aging among community-dwelling older adults. Am J Geriatr Psychiatry 2006 Jan;14(1):43-51. [doi: 10.1097/01.JGP.0000192489.43179.31] [Medline: 16407581]
30. Depp CA, Glatt SJ, Jeste DV. Recent advances in research on successful or healthy aging. Curr Psychiatry Rep 2007 Feb;9(1):7-13. [doi: 10.1007/s11920-007-0003-0] [Medline: 17257507]
31. Riessman CK. Narrative Analysis: Sage Publications; 1993.

XSL•FO
RenderX

32. Calasanti TM, Gender SKF. Social Inequalities, and Aging: Rowman Altamira; 2001. URL: https://books.google.com.au/books?id=y45SCthkXE8C [accessed 2025-05-02]
33. Davidson CS, Livermore J, Andersen P, Kaufman S. The nutrition of a group of apparently healthy aging persons. Am J Clin Nutr 1962 Mar;10(3):181-199. [doi: 10.1093/ajcn/10.3.181] [Medline: 13883781]
34. Bortz EL. Creative aging. J Am Geriatr Soc 1962 Jul;10(7):553-557. [doi: 10.1111/j.1532-5415.1962.tb00150.x] [Medline: 13871506]
35. Bortz EL. Cooperative, creative aging—a maturing challenge. We ripen or we rot*. J Am Geriatr Soc 1962 Nov;10(11):916-918. [doi: 10.1111/j.1532-5415.1962.tb00498.x]
36. Havighurst RJ. Successful aging. Bull Natl Assoc Second Sch Princ 1961 Feb;49(298):61-66. [doi: 10.1177/019263656504929805]
37. Maddox GL. Fact and artifact: evidence bearing on disengagement theory from the Duke geriatrics project. Hum Dev 1965;8(2–3):117-130. [doi: 10.1159/000270296] [Medline: 14338370]
38. Goldman S. Social aging, disorganization, and loss of choice. Gerontologist 1971;11(2):158-162. [doi: 10.1093/geront/11.2_part_1.158] [Medline: 5098322]
39. Spektorova SI, Utochkin LE. Aging of alloy AL9 without preliminary quenching. Met Sci Heat Treat 1971 Oct;13(10):884-886. [doi: 10.1007/BF00713833]
40. Marcoen A. Spirituality and personal well-being in old age. Ageing Soc 1994 Dec;14(4):521-536. [doi: 10.1017/S0144686X00001896]
41. Campos JL. Continuous positive ageing in Hodgkin's disease. Br J Radiol 1972 Dec;45(540):917-922. [doi: 10.1259/0007-1285-45-540-917] [Medline: 4647142]
42. Borchert R. Calcium-induced patterns of calcium-oxalate crystals in isolated leaflets of Gleditsia triacanthos L. and Albizia julibrissin Durazz. Planta 1985 Aug;165(3):301-310. [doi: 10.1007/BF00392226] [Medline: 24241134]
43. Goldsmith TC. Aging, evolvability, and the individual benefit requirement; medical implications of aging theory controversies. J Theor Biol 2008 Jun 21;252(4):764-768. [doi: 10.1016/j.jtbi.2008.02.035] [Medline: 18396295]
44. Nash FR, Sundburg WJ, Hartman RL, et al. Implementation of the proposed reliability assurance strategy for an InGaAsP/InP, planar mesa, buried heterostructure laser operating at 1.3 µm for use in a submarine cable. AT&T Tech J 1985 Mar;64(3):809-860. [doi: 10.1002/j.1538-7305.1985.tb00449.x]
45. Schorr JA, Farnham RC, Ervin SM. Health patterns in aging women as expanding consciousness. ANS Adv Nurs Sci 1991 Jun;13(4):52-63. [doi: 10.1097/00012272-199106000-00007] [Medline: 2059005]
46. Corman B, Michel JB. Glomerular filtration, renal blood flow, and solute excretion in conscious aging rats. Am J Physiol Regul Integr Comp Physiol 1987 Oct 1;253(4):R555-R560. [doi: 10.1152/ajpregu.1987.253.4.R555]
47. Somers AR. Aging in the 21st century: projections, personal preferences, public policies—a consumer view. Health Policy 1988;9(1):49-58. [doi: 10.1016/0168-8510(88)90117-0] [Medline: 10286666]
48. Garfein AJ, Herzog AR. Robust aging among the young-old, old-old, and oldest-old. J Gerontol B Psychol Sci Soc Sci 1995 Mar;50(2):S77-S87. [doi: 10.1093/geronb/50b.2.s77] [Medline: 7757843]
49. Cusack SA, Thompson WJA. Mental fitness: developing a vital aging society: Mens sana in corpore sano. Int J Lifelong Educ 1998 Sep;17(5):307-317. [doi: 10.1080/0260137980170503]
50. Secker J, Bowers H, Webb D, Llanes M. Theories of change: what works in improving health in mid-life? Health Educ Res 2005 Aug;20(4):392-401. [doi: 10.1093/her/cyg133] [Medline: 15507489]
51. Negash S, Bennett DA, Wilson RS, Schneider JA, Arnold SE. Cognition and neuropathology in aging: multidimensional perspectives from the Rush Religious Orders Study and Rush Memory And Aging Project. Curr Alzheimer Res 2011 Jun;8(4):336-340. [doi: 10.2174/156720511795745302] [Medline: 21222592]
52. Su CJ, Chiang CY. IAServ: an intelligent home care web services platform in a cloud for aging-in-place. Int J Environ Res Public Health 2013 Nov 12;10(11):6106-6130. [doi: 10.3390/ijerph10116106] [Medline: 24225647]
53. Nouchi R, Kawashima R. Benefits of "smart ageing" interventions using cognitive training, brain training games, exercise, and nutrition intake for aged memory functions in healthy elderly people. In: Tsukiura T, Umeda S, editors. Memory in a Social Context: Springer; 2017:219-232.
54. Global Report on Ageism, 1st edition: World Health Organization; 2021.
55. The risks of ageism model: how ageism and negative attitudes toward age can be a barrier to active aging. Soc Issues Policy Rev 2017 Jan;11(1):195-231. [doi: 10.1111/sipr.12031]
56. Dannefer D, Perlmutter M. Development as a multidimensional process: individual and social constituents. Hum Dev 1990;33(2-3):108-137. [doi: 10.1159/000276506]
57. Teater B, Chonody JM. What attributes of successful aging are important to older adults? The development of a multidimensional definition of successful aging. Soc Work Health Care 2020 Mar;59(3):161-179. [doi: 10.1080/00981389.2020.1731049] [Medline: 32100639]
58. Gonzales E, Matz-Costa C, Morrow-Howell N. Increasing opportunities for the productive engagement of older adults: a response to population aging. Gerontologist 2015 Apr;55(2):252-261. [doi: 10.1093/geront/gnu176] [Medline: 26035601]
59. Martinez IL, Crooks D, Kim KS, Tanner E. Invisible civic engagement among older adults: valuing the contributions of informal volunteering. J Cross Cult Gerontol 2011 Mar;26(1):23-37. [doi: 10.1007/s10823-011-9137-y] [Medline: 21243418]

XSL•FO
RenderX

60. Ranzijn R. The potential of older adults to enhance community quality of life: links between positive psychology and productive aging. Ageing Int 2002 Mar;27(2):30-55. [doi: 10.1007/s12126-002-1001-5]
61. Iversen TN, Larsen L, Solem PE. A conceptual analysis of Ageism. Nord Psychol 2009 Jan;61(3):4-22. [doi: 10.1027/1901-2276.61.3.4]
62. Chen YRR, Schulz PJ. The effect of information communication technology interventions on reducing social isolation in the elderly: a systematic review. J Med Internet Res 2016 Jan 28;18(1):e18. [doi: 10.2196/jmir.4596] [Medline: 26822073]
63. Krawczyk K, Chelkowski T, Laydon DJ, et al. Quantifying online news media coverage of the COVID-19 pandemic: text mining study and resource. J Med Internet Res 2021 Jun 2;23(6):e28253. [doi: 10.2196/28253] [Medline: 33900934]
64. Wong AKC, Lee JHT, Zhao Y, Lu Q, Yang S, Hui VCC. Exploring older adults' perspectives and acceptance of AI-driven health technologies: qualitative study. JMIR Aging 2025 Feb 12;8:e66778. [doi: 10.2196/66778] [Medline: 39937162]
65. Kibria MG. Knowledge creation model in woo enabled smart ageing iot service platform. 2016 Presented at: 2016 Eighth International Conference on Ubiquitous and Future Networks (ICUFN); Vienna, Austria p. 526-531. [doi: 10.1109/ICUFN.2016.7537088]
66. Hicks MM, Conner NE. Resilient ageing: a concept analysis. J Adv Nurs 2014 Apr;70(4):744-755. [doi: 10.1111/jan.12226] [Medline: 23919385]
67. Hennessy CH, Walker A. Promoting multi-disciplinary and inter-disciplinary ageing research in the United Kingdom. Ageing Soc 2011 Jan;31(1):52-69. [doi: 10.1017/S0144686X1000067X]
68. Chrysikou E, Biddulph JP, Loizides F, et al. Creating resilient smart homes with a heart: sustainable, technologically advanced housing across the lifespan and frailty through inclusive design for people and their robots. Sustainability 2024;16(14):5837. [doi: 10.3390/su16145837]
69. Kulkarni M. Digital accessibility: challenges and opportunities. IIMB Manag Rev 2019 Mar;31(1):91-98. [doi: 10.1016/j.iimb.2018.05.009]
70. An exploration of older worker flexible working arrangements in smaller firms. Human Res Manag J 2016 Jan;26(1):12-28. [doi: 10.1111/1748-8583.12074]
71. Braun MT. Obstacles to social networking website use among older adults. Comput Human Behav 2013 May;29(3):673-680. [doi: 10.1016/j.chb.2012.12.004]
72. Dauenhauer J, Steitz DW, Cochran LJ. Fostering a new model of multigenerational learning: older adult perspectives, community partners, and higher education. Educ Gerontol 2016 Jul 2;42(7):483-496. [doi: 10.1080/03601277.2016.1157419]

Abbreviations

AI: artificial intelligence
TLS: total link strength
WHO: World Health Organization

Edited by A Quialheiro; submitted 01.02.25; peer-reviewed by JA Gonzalez-Oreja, WF Khaw; revised version received 14.05.25; accepted 23.07.25; published 23.09.25.

Please cite as:
Xiao N, Xia B, Buys L, Susilawati C, Larbi M, Chen Q
Comparative Analysis of 16 Aging Concepts and Their Influence on Aging Narratives: Bibliometric and Content Analysis
JMIR Aging 2025;8:e72011
URL: https://aging.jmir.org/2025/1/e72011
doi:10.2196/72011

© Na Xiao, Bo Xia, Laurie Buys, Connie Susilawati, Martin Larbi, Qing Chen. Originally published in JMIR Aging (https://aging.jmir.org), 23.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Experience and Impact of Digital Technologies on Indigenous Populations in New Zealand During the COVID-19 Pandemic and Cyclone Gabrielle: The Kaupapa M ori Methodology

Dianne Wepa[1], PhD; Shiji Thomas[2,3], MSc; Md Shafiqur Rahman Jabin[2,4], PhD

[1]School of Community and Public Health, Auckland University of Technology, Auckland, New Zealand

[2]Faculty of Health Studies, University of Bradford, Bradford, United Kingdom

[3]School of Health and Society, University of Salford, Salford, United Kingdom

[4]Department of Medicine and Optometry, Linnaeus University, Universitetsplatsen 1, Kalmar, Sweden

Corresponding Author:
Md Shafiqur Rahman Jabin, PhD
Faculty of Health Studies, University of Bradford, Bradford, United Kingdom

Abstract

Background: Pandemics, such as COVID-19, and climate change–related catastrophic weather events are increasing, impacting social connectedness within communities by disrupting social cohesion, increasing loneliness, and affecting mental health and social well-being. Digital technology, in addition to being used for communication, education, and business transactions, also plays a vital role in maintaining a country's health and well-being, as well as sustaining economic growth.

Objective: This study aimed to explore the experiences of Māori kaumātua in using digital technology to meet their health needs within Ngāti Kahungunu, North Island, New Zealand, during the COVID-19 pandemic and Cyclone Gabrielle.

Methods: This qualitative study employed the Kaupapa Māori methodology to understand the challenges, resilience, and approaches used by Māori to maintain connectedness and access essential services. An inductive approach to thematic analysis, as recommended by Braun and Clarke, was used to ensure a thorough and robust data analysis. The user characteristic was assessed on a semantic level using the information provided in the narrative text.

Results: The findings highlight the role of digital technology in disaster management and underscore the urgent need to address digital disparities in support of vulnerable populations. In this study, 14 individuals were interviewed, comprising 71% (n=10) women and 29% (n=4) men. These participants fell into different age groups, with 9 participants being 65 years or older (older adults). Of the total participants, 43% (n=6) were limited users, 43% (n=6) comprised confident users, and the rest (n=2; 14%) were normal users. A total of 6 themes emerged from the interview data: social connectedness and resilience, digital literacy and access to information, barriers to telecommunications and digital technology, cultural appropriateness and psychological barriers, perceived threats of feeling insecure, and impact on mental health and emotional well-being.

Conclusions: Vulnerable situations such as pandemics and extreme weather events can have tremendous effects on the lives of Indigenous people who live remotely. The study also focused on the actions that should be taken to mitigate these challenges and overcome difficult circumstances, such as the pandemic and the cyclone. The recommendations include a better health care system and improved coordination among care providers, user-friendly digital solutions, ensuring local funding and community services, establishing training processes for basic digital skills, and fostering leadership and partnerships with Indigenous New Zealanders.

(JMIR Aging 2025;8:e73974) doi:10.2196/73974

KEYWORDS

community service; COVID-19; Cyclone Gabrielle; digital exclusion; digital solutions; disaster management; health inequalities; Māori kaumātua; Ngāti Kahungunu; training and education

Introduction

Background

Pandemics, such as COVID-19, and climate change–related catastrophic weather events are increasing, impacting social connectedness within communities by disrupting social cohesion, increasing loneliness, and affecting mental health and social well-being [1,2]. Displaced populations and resource scarcity can reduce social functioning. However, the breakdown of local support systems is one of the key factors that maintain social functioning and can have a significant impact on communities. Psychological impacts such as anxiety and fear of future events diminish social engagement, particularly for

XSL•FO
RenderX

vulnerable groups such as older people [1,2]. Indigenous older people are generally more vulnerable (or at risk) in each of those domains, such as COVID-19, natural disasters such as Cyclone Gabrielle, and digital technologies [3-5].

The Context of the COVID-19 Pandemic in New Zealand

New Zealand, a small island nation with a population of approximately 5 million, is a diverse community, with significant representation from the Māori communities (16%), Asia (15%), and the Pacific Islands (7%), as well as Europe and other countries (62%) [6]. Māori are the Indigenous people of New Zealand. They migrated from Polynesia to New Zealand more than 800 years ago. Their tradition and culture are deeply rooted in respect for nature, community, and their kaumatua (elders) [7]. The New Zealand government introduced new strategies in response to the COVID-19 pandemic. One of the strategies includes a 4-tiered alert system that indicates the severity of the COVID-19 situation and corresponding measures [8]. This system, which ranges from Alert Level 1 (prepare) to Alert Level 4 (lockdown), was not only highly regarded worldwide but also proved effective in managing the crisis [9].

Most of the workforce, excluding essential workers such as staff from supermarkets, hospitals, and transport, was advised to work remotely during the lockdown period [9]. Daily media briefings, led by government officials like the Prime Minister and Director-General of Health, were broadcast on the radio to keep the population updated about the impact of the disease in the country and the plans to eliminate it [10]. Websites and the national emergency broadcast system were used to advise the public regarding safety measures [11]. Important information from the "Unite Against COVID-19 website" and the number for free helplines were also shared on social media platforms, such as Facebook, Twitter, Instagram, LinkedIn, and WhatsApp, by government agencies [4]. These ways of communication and working meant that people had to rely on digital technology during the crisis period.

The Context of Cyclone Gabrielle in New Zealand

Cataclysmic weather episodes are increasing globally, adversely contributing to the rise in health-related risks [12], food security, livelihoods, clean water supply, and human security [13]. Climate change–related health effects are variable for different populations as they depend on geographical positioning, health system capacity and capability, demographics, and the disease burden of communicable and non-communicable diseases [14]. The world faces a distinctive challenge with the rapidly changing climate and the rising frequency of extreme weather events. In the recent decade, the prevalence of extreme weather events has escalated due to climate change and global warming [12,14]. Coronese and colleagues established a compelling relationship between climate change and the heightened frequency and severity of extreme weather events in numerous regions of the globe [15]. The Intergovernmental Panel on Climate Change [16] delineates the implications of climate change and climate-induced extreme weather events on socioeconomic systems, including crop damage, heightened wildfire risk, more frequent power outages, increased susceptibility to food or water shortages, and political system vulnerability [5].

Within the context of New Zealand, severe weather and climate events are frequently experienced, primarily associated with flooding that disconnects different communities and groups of people from accessing support in the form of health care services, medical and food supplies, safe drinking water supply, and safe areas for rehabilitation during disaster occurrences [17]. The year 2023 began with a series of emergencies declared in Aotearoa, New Zealand [18]. Regions, such as Coromandel/Thames and Te Tairawhiti Bay of Plenty, were hit by Cyclone Hale [19,20], not only causing considerable damage to the livelihoods of the residents but also uprooting the entire coastal ecosystem from its essence [17]. Disastrously, this was followed by Cyclone Gabrielle, which caused significant flooding in many different parts of Tamaki Makaurau and caused the government to declare a national emergency [21,22]. The devastating impact of Cyclone Gabrielle was particularly evident in regions of Hawke's Bay and Gisborne [21]. The floodwater left thousands stranded and trapped in high-risk landslide areas [23].

The Use of Digital Technologies by Indigenous New Zealanders

Over the last 2 decades, significant use of technology has been seen, mainly for communication, education, and business purposes [24]. However, in 2019, the Department of Internal Affairs reported that 1 in 5 of the population, mainly people from Māori and Pacific Islands, people living in social housing or rural areas, older persons, and people self-identifying as the unemployed and the disabled, did not have access to or the skills to use digital technology [25]. Digital technology in this context refers to electronic devices, such as computers and mobile phones, and their infrastructure (connection modalities) used for communication and sharing information. The report also highlighted that digital exclusion affected their access to education and communication and significantly impacted a person's well-being, underscoring the urgent need to address this issue, and different recommendations were proposed to overcome digital exclusion [25].

Similarly, many studies have reported health inequalities with Māori, resulting in high morbidity and mortality rates [26,27]. The incidence rate of complications and deaths has been high in the Māori communities during previous outbreaks of influenza and other communicable diseases [28]. However, there is potential for introducing positive changes by understanding the lived experiences of Māori people in using digital technology to meet their health care needs.

Study Aims

The primary aim of this exploratory qualitative study was to understand the lived experience of Māori communities in using digital technologies to meet their health care needs during the simultaneous occurrence of the COVID-19 pandemic and Cyclone Gabrielle. This study focused on the experience of Māori communities, particularly the older generation (known as kaumātua), during the COVID-19 pandemic and Cyclone Gabrielle, emphasizing the importance of digital technology in maintaining social connectedness and accessing vital services. The study explored the following research questions:

XSL•FO
RenderX

1. What was the experience of day-to-day life using digital technologies for the Indigenous New Zealanders during the COVID-19 pandemic and Cyclone Gabrielle?
2. How did the pandemic and the natural disaster affect the use of digital technologies and access to information?
3. What recommendations should be considered for using digital technologies and information access during a pandemic and natural disaster?

Methods

Overview

This was an exploratory qualitative study based on the Kaupapa Māori methodology [29]. This methodology involves research developed, conducted, and reported by Māori to promote and support their community. This research approach is grounded in Māori worldviews, values, and practices. It is a decolonizing methodology developed by and for Māori, seeking to challenge Western-dominated ways of knowing and conducting research. The Kaupapa Māori methodology prioritizes Māori aspirations, language, culture, and autonomy throughout the research process [29-31]. Kaupapa Māori research is often participatory and collaborative, focusing on creating knowledge in partnership with Māori communities. In this sense, it aligns strongly with the interpretivist approach, as it emphasizes the importance of understanding people's subjective experiences, values, and cultural contexts [32].

Data Collection

In this study, 14 Māori families from various communities within the tribe or iwi, known as Ngāti Kahungunu, participated in semistructured interviews conducted between December 2022 and February 2023. A purposive sampling method and the interpretivist approach were employed to select participants from Māori communities, as this Phase 2 study builds upon a previous research conducted by Wepa and colleagues [30], which involved interviewing a small population group. The population group represented the senior members and older adults of Māori families since they are generally more vulnerable (or at risk) in each of those domains, such as COVID-19, natural disasters such as Cyclone Gabrielle, and digital technologies. Participants who were at increased risk for adverse events were excluded. Participants ranged in age from 45 to 90 years old. A participant information sheet was provided to all participants.

These families were interviewed as a group as per the Kaupapa Māori methodology. Under this methodology, research was developed, conducted, and reported by Māori to promote and support their community [29-31]. A research assistant connected to each community supported the primary investigator in conducting the semistructured interviews from December 10, 2022, to February 16, 2023. The interviews lasted an average of 40 minutes, with the longest interview being approximately 90 minutes. The interviews were targeted at the family's senior member, such as older adults; in some cases, other family members, such as daughters, sons, and grandchildren, participated in a support role. This is because some older adults experienced difficulties with hearing or vision, making it challenging for them to participate in an interview. In such cases, family members helped by clarifying questions and reading materials aloud. Semistructured interviews were conducted by JTH and RTH (research assistants) using the interview guide (Textbox 1). The research assistants were chosen since they were familiar with the values and cultural background of the Indigenous community in New Zealand. However, the interview was not only restricted to the interview guide but also included additional social demographic information, such as age, gender, location, living conditions, which were also collected. The interviews were audio-recorded and supplemented by written field notes.

Textbox 1. Interview guide for semistructured interviews.

Interview guide

1. How did you use digital technology (during COVID-19 and Cyclone Gabrielle)?
2. Given your experiences, do you have any suggestions for what improvements could be made?
3. What process do you think providers of services could consider?
4. What process do you think (whanau, hapu, iwi) people could consider?
5. Is there something else you would like to share?
6. Is there anything else you would like to ask me?

Data Analysis

The recordings were transcribed to free-text narratives by RS (research assistant), as the researcher's background, training, and familiarity with the Indigenous community aligned with the subject matter. The free-text narratives were analyzed by MSRJ and DW and reviewed by ST to minimize bias and enhance the validity of the results, since these researchers' expertise lies within the realm of qualitative thematic analysis. An inductive approach to thematic analysis, as recommended by Braun and Clarke, was used to ensure a thorough and robust data analysis. This comprehensive approach was chosen to ensure the validity and reliability of the results, instilling confidence in the research findings [33], which involved identifying keywords and phrases from the transcripts to indicate potential relevant concepts. The extracted concepts were then grouped into a number of themes of similar nature. Discussions were held between MSRJ, ST, and DW to achieve a consensus when there was a difference in opinion, ensuring a collaborative and inclusive research process [34,35]. This collaborative approach was chosen to ensure that all perspectives were

XSL•FO
RenderX

considered, making the research process more inclusive and comprehensive [34].

Ethical Considerations

Written informed consent was obtained from them before the interviews. The process of obtaining consent involved explaining the study, its potential risks and benefits, and the participants' rights. Participation in the study was voluntary, and the participants were given the option to attend interviews at their homes or a meeting place of their choice. Transportation was arranged for participants who attended interviews at a communal place. The research grant covered participant travel costs. Ethics approval was granted by the Auckland University of Technology Ethics Committee on 10 November 2022 (Ref No. 22/30). Any personal or sensitive information collected was deidentified, and the free-text narratives were aggregated for analysis.

Results

Participants Characteristics

For this study, 14 individuals were interviewed, comprising 71% (n=10) women and 29% (n=4) men. These participants fell into different age groups, with 9 participants being 65 years or older (older adults; Table 1).

Of 14 participants, 6 lived in a rural community, with 1 residing outside of it; 2 close to it; 2 in central Hastings, Hawkes Bay; and 1 at regional Ngati Kahungunu area. Among the remaining participants, 3 resided in Havelock North, 1 in Matahiwi Marae Base, 1 in regional Ngati Kahungunu area, 1 at the central location, and 1 outside Hastings, Hawkes Bay (Table 1). These are the places where Māori populations were mainly represented.

The participants older than 65 years and living with disabilities or preexisting health conditions were particularly vulnerable. The health conditions ranged from mobility, vision, hearing, cardiovascular or respiratory diseases, reliance on medication or health services for their daily needs and overall well-being.

All participants were isolated at home during the COVID-19 pandemic, with 64% (n=9) living with family members or dependents, such as sons, daughters, and grandchildren, and the rest (n=5; 36%) isolated alone with no family members. Of all, 36% (n=5) of the total participants, who were isolated alone, were dependent on others, such as friends, family (not living with them), and community members, for support (Table 1). Participants reported strong kinship networks as a substantial social capital central to the communities, enabling them to practice strong social connectedness or cohesion.

All participants except 1 received help from the local community (eg, Whānau, Marae) members, family members, or local health providers (eg, Māori; Table 1). They received food, water, sanitizing products, and other health needs. Some individuals were provided with services, such as pick-ups and drop-offs for health appointments and vaccinations, by local health providers. They could rely on community members and health providers for home visits, funding, access to COVID-19 information, and other medical needs.

Some participants felt comfortable and confident offering support and advice to other community members affected by COVID-19, as well as distributing food packages, sanitizing products, and medical kits. Other participants received little help during the COVID-19 pandemic; instead, they relied on non-family members for support and struggled to find information independently (Table 1). The person who did not receive support during the pandemic experienced worsening unmet needs, reduced access to care, increased distress, and heightened loneliness.

Of the total participants, 50% (n=7) had a good or excellent understanding of COVID-19 information and requirements (eg, isolation, staying home; Table 1). Of these 7 participants, 5 were either health workers, former health workers, or affiliated with health organizations within the community. This group of participants was proactive in offering health advice, including guidance on accessing information and selecting suitable digital media. Their support extended beyond the domain of information and technology, including vaccinations through local health providers and transportation for elderly individuals who were ill. The remaining 50% (n=7) lacked a good understanding (fair or basic) of COVID-19, needing help and heavily relying on other family members for information and health support (Table 1).

XSL•FO
RenderX

Table . Participant characteristics.

P[a]	Age (years)	G[b]	Living area	Living with FM[c] or dependents or alone	Help received	Understanding of COVID-19 (health care workers)	DT[d] usage	Internet access
P1	55+	M[e]	Mathiwi Marae Base	FM or dependents	Yes	Good (Ministry of Health)	Confident	No issues
P2	55+	F[f]	Paki Paki, Hastings Hawkes Bay	FM or dependents	Yes	Fair	Normal	No issues
P3	85+	M	Close to central Hastings Hawkes Bay	Alone	Yes	Basic	Limited	Had issues
P4	55+	F	Havelock North	FM or dependents	Yes	Excellent (midwife)	Confident	No issues
P5	60+	F	Paki Paki, outside Hastings Hawkes Bay	Alone	Yes	Good (volunteer)	Confident	Had issues
P6	65+	M	Flaxmere area	Alone	Yes	Basic	Limited	Had issues
P7	40+	F	Paki Paki, close to Hastings Bay	FM or dependents	Yes	Good (caretaker)	Confident	Had issues
P8	65+	F	Havelock North	Alone	No	Fair	Limited	Had issues
P9	65+	F	Central Hawkes Bay	FM or dependents	Yes	Good (ex-health care worker)	Confident	Had issues
P10	85+	F	Regional Ngati Kahungunu area	FM or dependents	Yes	Fair	Limited	No issues
P11	65+	F	Outside Hastings	FM or dependents	Yes	Good	Normal	Had issues
P12	85+	M	Havelock North	Alone	Yes	Good	Confident	Had issues
P13	75+	F	Central Hawkes Bay	FM or dependents	Yes	Fair	Limited	No issues
P14	65+	F	regional Ngati Kahungunu area	FM or dependents	Yes	Basic	Limited	Had issues

[a]P: participants.

[b]G: gender.

[c]FM: family member.

[d]DT: digital technology.

[e]M: male.

[f]F: female.

Sources of Information

A number of sources of information were identified: the Ministry of Health website, COVID number (eg, 0800), TV and radio news or updates, COVID app (eg, COVID Sign in), tracking app with QR code, and social media (eg, Facebook). To access information and health requirements, the participants used their own devices, including cellular phones, smartphones, telephones, TVs, computers, and laptops.

A total of 5 (36%) of the participants reported having access to the internet with no issues, while the rest (64%) described their internet access as inconsistent and expensive (Table 1).

Digital Technology User Characteristics

Of the total participants, 43% (n=6) were limited users, 43% (n=6) comprised confident users, and the rest (n=2; 14%) were normal users (Table 1). The user characteristic was assessed by using the information provided in the narrative text on the semantic level (without making any assumptions).

XSL•FO
RenderX

Before the pandemic and the cyclone, the limited users did not know how to use such technologies. Therefore, during the pandemic and the cyclone, these users had to depend on family members, such as daughters and grandchildren, to access government information or regular updates from the TV. The grandchildren indicated that the use of those technologies remained limited or did not improve among the senior family members after the pandemic and the cyclone.

> *I had to rely on my granddaughter, who could access information through Facebook.* [P13]

Confident users could navigate and access various information sources and the online world and regulate social media (eg, Facebook), frequently using different apps before the pandemic and the cyclone. The frequent use of digital technologies enabled these participants to quickly search for COVID-19 information during the pandemic. They could use a QR-scanning app to access information on health requirements and track the latest emergency updates from the government (eg, the Ministry of Health website). The increased use of technology during the pandemic and the cyclone also enabled the participants to connect with family members through social media. After the pandemic and the cyclone, confident users continued to use technologies more efficiently and regularly. They upgraded their knowledge and use of various other apps, such as Microsoft Teams (virtual meetings), and helped other family members with shared information.

> *Navigating information through various means made me feel confident and act on necessary actions.* [P12]

Before COVID-19 and the cyclone, most people used digital technology mainly for social media, such as Facebook. Their regular use of social media helped them keep in touch with their family, and apps such as "COVID Sign-in" supported them in accessing updated information from the government about the ongoing pandemic and the cyclone. Although this user group continued to use various social media and apps to a significant extent, they relied on their family members for extended usage of apps, such as QR scanning, after the pandemic.

> *I use Facebook regularly. This helped me accessing updated information from the government.* [P2]

Themes That Emerged as a Result of the Study

In total, 6 themes emerged from the interview data: social connectedness and resilience, digital literacy and access to information, barriers to telecommunications and digital technology; cultural appropriateness and psychological barriers, perceived threats of feeling insecure, and impact on mental health and emotional well-being.

Theme 1: Social Connectedness and Resilience

Kaumātua, with their resilience and adaptability, have been the cornerstone of fostering community resilience. Their effective use of digital platforms to stay connected with whānau and access health services has been a testament to their resourcefulness. Their collective action, community solidarity, and reliance on traditional support networks have been instrumental in overcoming the challenges posed by isolation during COVID-19 and Cyclone Gabrielle [3]. Most participants felt safe and fortunate to have family and friends to rely on during the COVID-19 pandemic. The importance of love and a sense of belonging within family and community was expressed by many participants, while experiencing social isolation during the pandemic and cyclone. They felt confident in the information provided by family members, such as grandchildren, and assessed their health needs, including vaccination for COVID-19. They could also rely on the community for help, as they felt safer and more confident within it. They felt the need to connect with relatives living outside New Zealand. This further enabled them to rely on digital technology (eg, smartphones) to stay updated and informed, not only for themselves but also for their family, the community, and the outside world.

> *A lot of good medicine is basic love, how do we support each other? Through aroha (love).* [P4]

> *No department knows me better than my moko (grandchild) does.* [P10]

The participants considered connectedness as an essential factor in maintaining good health and well-being. As such, social connectedness can help neutralize vulnerabilities within a socioeconomic context through neighborhood networks, community cohesion, and trustworthiness, which can empower community members to work toward shared goals and objectives during disasters.

> *You can only look to your community for help, nobody else.* [P10]

Theme 2: Digital Literacy and Access to Information

The digital divide was evident, with many participants relying on the tech-savvy younger generation for technology use and access to information. This intergenerational support has been crucial in bridging the gap. However, some community individuals lacked the basic digital skills needed to navigate the internet effectively, and others were required to develop advanced digital skills to enhance their confidence and motivation. The impact of limited digital literacy on independent access to health-related information is significant, and coupled with inconsistent internet connectivity and high costs, it hinders their ability to access health-related information independently [3].

> *First it was COVID-19 and now it's the cyclone, no coordinated approach by the government–where will it end?* [P6]

The participants relied heavily on social media to stay in touch with their family members. However, their lack of knowledge and access to digital devices hindered their ability to maintain communication during periods of isolation and the cyclone. Difficulty in using digital devices affected their ability to access online health facilities. They suggested introducing more user-friendly digital solutions to overcome these challenges. Moreover, local care providers demonstrated an apparent lack of coordination among themselves, which could have been further improved by exchanging and updating information promptly and delivering one-on-one services as needed.

XSL•FO
RenderX

> *We used to have a health clinic, now you're told to go online, nobody around to help do that if you don't have someone in the family to help.* [P3]

Theme 3: Barriers to Telecommunications and Digital Technology

During Cyclone Gabrielle, widespread power outages and telecommunications failures severely affected the community's ability to communicate and access emergency services [3]. Participants emphasized the need for improved infrastructure, better internet connections, and localized communication channels to address these challenges.

> *A rich man moved out to our beach and complained about the internet, then a new tower got built soon after.* [P11]
>
> *People are building empires, need to cut the red tape and put that money back into the people.*
>
> *[P9]*

It was a challenging situation for those who were not digitally savvy, as they had to rely heavily on others for health needs and information. The government should have played a more significant role in addressing these gaps; for example, providing appropriate "match support" could have improved the situation.

> *Good phone and internet coverage was only available with one company.* [P7]

Theme 4: Cultural Appropriateness and Psychological Barriers

A preference for face-to-face interactions over digital communication was evident for the Māori communities. Various concerns about online safety and privacy, as well as the potential fear of losing the value of their own culture and language, contributed to the reluctance to engage with digital technologies. This necessitated digital inclusion initiatives and online services to be designed through the lens of the local community, incorporating their cultural values, learning styles, and needs. Fostering local community-led digital projects and ensuring online environments are safe and respectful is required. Ensuring online safety is crucial to building trust and confidence in the digital world. It is equally important to provide access to culturally relevant content, making users feel understood and catered to.

> *If we are supposed to use QR codes, where is the privacy around that?* [P1]

Theme 5: Perceived Threats of Feeling Insecure

The frequent visits of health providers were seen as "unwanted people" turning up at home, posing a significant risk to security during the time of isolation. Social media would have been a better choice than meeting in person, as it reduced the risk of contracting COVID-19 associated with in-person visits. For example, one of the participants felt intense fear and uncertainty when returning to her own community, as she was unsure about how it might affect her if she contacted her community members.

> *She didn't like people or unwanted people turning up at her home (during Covid-19). She felt unsafe and insecure.* [P5]

Theme 6: Impact on Mental Health and Emotional Well-Being

Following the cyclone, the area experienced constant power cuts and load shedding, which was particularly worrisome since the internet and phone lines were down for most of the time. However, in the face of these challenges, the Māori community came together in a display of unity and support. The disruption to community services and communication, combined with the displacement from homes and separation from family and community, particularly elders and children, loss of property, and disruption to cultural practices, had a significant impact on the mental health and well-being of Māori communities. But the community's solidarity was a powerful force that helped overcome these difficulties.

> *At some point, I felt mentally disturbed and did not know what to do.* [P6]

Discussion

Principal Findings

This study aimed to explore the experiences of people from Māori communities in using digital technology to meet their personal health needs during the COVID-19 pandemic and Cyclone Gabrielle. It examined their ability to access digital devices and their skills and capabilities in using these devices to access vital information. Our study's findings provide valuable insights into the role of digital technology in disaster management and underscore the urgent need to address digital disparities in support of vulnerable populations, such as older adults. In total, 6 themes emerged from the interview data: social connectedness and resilience, digital literacy and access to information, barriers to telecommunications and digital technology, cultural appropriateness and psychological barriers, perceived threats of feeling insecure, and impact on mental health and emotional well-being. This section presents the interpretation of the results and their significance by relating them to existing research, particularly in terms of the impact of digital technologies during the pandemic and the natural disaster, followed by a set of recommendations, and the strengths and limitations of the study.

Maintaining social connectedness during disasters and pandemic situations is crucial for several reasons. Social connections provide emotional support, reducing feelings of isolation, stress, and anxiety that often accompany such crises [36,37]. During disasters, strong social networks can facilitate the sharing of vital information, resources, and aid, enhancing community resilience and individual well-being. Additionally, connectedness ensures that vulnerable populations, such as the elderly or those with disabilities, receive the necessary care and support [37,38]. In the context of pandemics, where physical distancing measures are implemented, maintaining social connections through virtual means helps uphold mental health and prevents the adverse effects of loneliness [39]. Social connectedness also fosters a sense of community and solidarity,

XSL•FO
RenderX

which is crucial for collective action and adherence to public health measures. Overall, maintaining social bonds during crises is vital for both psychological health and effective disaster response and recovery [25,40].

Impact of Digital Technologies During the COVID-19 Pandemic

Digital health technologies are applicable in various areas, such as chronic disease self-management, as well as storing, sharing, and transferring confidential clinical data. These apps have enhanced the adherence to therapeutic regimes and provided a reliable platform for communication with health care professionals, enabling timely interventions [41-43]. Previous research on telemedicine has demonstrated its ability to save costs and travel time, particularly benefiting older populations in underserved areas. It can also reduce waiting times for patients for medical appointments, thereby reducing the time to diagnosis [43,44]. Additionally, especially during times of pandemics, it can help to reduce the risk of contagion and infections [45].

Most older individuals in the community relied on local health care workers and family members for food, water, and sanitary supplies, exposing themselves to unwanted health and safety risks. During this period, information about support services and preventive measures was disseminated to the public through digital technologies, including television, radio, health websites, and social media. However, the results of this study revealed that nearly half of the community people had limited knowledge of using this technology and were dependent on family members to attain important information. Furthermore, two-thirds of respondents reported challenges associated with internet connections, including high costs, limited accessibility, and difficulties using the internet. These issues are compounded by a gap in communication and information sharing, suggesting that digital technology is not equally accessible, resulting in a disparity in its use. These findings align with the report published by the Department of Internal Affairs in 2019, which stated that Māori people had limited access to internet connections and digital technologies, and inadequate funding for basic digital training [25]. These challenges highlighted the need for training and education to enhance the foundational digital skills for Māori communities [46].

There was a clear need for improved communication and a more effective system to reach COVID-19 victims. Information and information-sharing hubs could be better utilized to facilitate community links and support client work. Although the use of a QR scanning app enabled easy access to and management of information, a more effective approach would have been to maintain an up-to-date database with each community registered. The database would have helped people understand COVID-19 requirements and contributed to building the community [47,48]. Moreover, the inconsistency in the information shared by local health providers clearly indicated a lack of coordination among service providers [49,50]. This could have been further improved by exchanging and updating information in a timely manner and delivering one-on-one service as required [50,51].

Impact of Digital Technologies During Cyclone Gabrielle

The frequent occurrences of extreme weather events have a significant impact on social connectedness within communities. Disasters can disrupt these social networks, displace individuals from their homes, and create feelings of isolation and loneliness. This is particularly important for elderly populations, who may rely heavily on their social connections for support and companionship. Recognizing and preserving social connectedness are crucial for effective disaster management and community resilience [52]. The findings of this study indicate that digital technology plays a key role in maintaining health and well-being during crises, as vital information can be shared through social connectedness. The lack of digital literacy and reliable telecommunications infrastructure disproportionately affected Māori communities' ability to access timely information on safe practices during the COVID-19 pandemic and Cyclone Gabrielle, highlighting the urgent need for targeted interventions. Strengthening community-driven initiatives and integrating indigenous knowledge into disaster management plans can enhance resilience and support [53].

According to a recent report, the displacement triggered by Cyclone Gabrielle has destroyed numerous houses and forced a significant number of evacuations, affecting approximately 10,500 individuals. In the Opotiki district alone, heavy rainfall led to the evacuation of 1335 individuals, accounting for a substantial 15% of the total population. Despite these challenges, the affected communities have shown remarkable resilience [54]. Our study, a detailed analysis of the impact of Cyclone Gabrielle on digital public health, employed a qualitative research method. Like most studies of extreme weather events, we have reported that Cyclone Gabrielle had severe and long-lasting effects on public health. The individuals who faced this cyclone, coupled with power cuts and loss of internet connection, have experienced a significant impact on their psychological well-being and quality of life. This led to an increased level of anxiety, depression, and post-traumatic stress, which are likely to persist in the long term [55]. Within the affected region, the long-term devastation of primary industries, coupled with widespread damage such as heavy flooding, landslides, and infrastructure destruction, has necessitated a national state of emergency. Immediate action is crucial to mitigate the long-term effects of this disaster [56].

During such extreme weather events, poor access to reliable telecommunication services (like phones and internet connections) is exacerbated within populations that already face challenges such as remoteness [57]. The reliance on care support is a significant factor influencing an individual's awareness of risks and ability to control them [58].

Recommendations

It is essential to remember that Indigenous communities are incredibly diverse, each with its unique challenges and opportunities for digital inclusion. Addressing these challenges necessitates a multifaceted, collaborative, and, most importantly, a long-term commitment from governments, technology companies, nonprofits, and Indigenous communities. Each stakeholder plays a crucial role in developing sustainable and

XSL•FO
RenderX

culturally appropriate solutions [3,59]. Therefore, based on the study findings and the considerations emerging from various literature and public reports, we suggest the following recommendations to help Indigenous communities overcome challenging situations, such as the pandemic and extreme weather events.

The need for a user-friendly and culturally relevant digital solution/technology was evident, as older individuals found it difficult to learn how to use digital technology. The frustration caused by the long waiting time for the national COVID 0800 number was accepted by those affected by the outbreak. A local COVID number could have replaced this to facilitate the needs of the local community and its people [50,60]. The need for a more user-friendly, culturally relevant, and simplified version of the digital solution is paramount, particularly for older individuals, to make it "fit for purpose" [25,61]. It is crucial that these programs are community-led, culturally safe, and accessible to people of all ages [3,62], especially to reduce digital disparity among Māori communities.

Ensure local funding—people directly linked to the local community were expected to have more access to food, water, and medical supplies [63]. There is sufficient evidence that the complex administrative requirements and logistical challenges associated with accessing government funding and programs can pose significant threats to Indigenous communities with limited resources [64]. More local funding needs to be provided for low-income families coping with isolation, food scarcity, and health conditions [60].

The need for additional community services—specifically, health care for older people—was needed. Further support (eg, doctor visits, regular phone calls to check up on health) for older individuals who would also have been considered [52], particularly helping them access the information and the health requirements [65], for example, how to download and use a mobile app [3]. The government needs to apply the lessons learned to those in most need and work with the local communities to provide additional services [50,60].

Establish a training program for the local community—focusing on digital apps such as online banking, social media, and email, as well as accessing information on the internet. Many community members, especially older adults, lack proficiency in these areas. By providing training and education, older adults will develop fundamental skills, enabling them to utilize technology for their future needs [56,66,67]. Government support is crucial in ensuring equitable access to digital technologies and resources, particularly for underserved communities and those lacking digital literacy. With the help of the relevant authority, necessary training and resources can be provided to help individuals develop the digital skills needed to participate more fully in the digital economy [68,69].

Fostering leadership and partnerships—governments and mainstream service providers may overlook the unique needs and priorities of diverse Indigenous communities, failing to tailor policies and programs to their specific contexts. The absence of Indigenous-led initiatives is particularly concerning, as they are not only crucial but also the most effective way to develop culturally relevant and effective digital inclusion strategies. Recognizing the expertise of Indigenous communities in understanding their own needs is essential for respectful and impactful policy-making [25,70]. The implementation of these strategies necessitates fostering collaboration between local health care providers, regional and government funding bodies, community health partners, and the national health service in New Zealand (Te Whatu Ora) [31].

Strengths and Limitations of the Study

The Kaupapa Māori was chosen for several reasons: first, the primary investigator identifies as Māori and is connected to the community through whakapapa or genealogical ties and developed a body of knowledge to improve Māori health status over her academic career; second, the elders from her community approached her to investigate their concerns about the lack of access to health services within rural and remote communities in their tribal region; and third, this publication is a continuation of a 3-phase study that utilized Kaupapa Māori methodology since its commencement [48]. This study also followed Standards for Reporting Qualitative Research (SRQR), which improved the clarity, completeness, transparency, and credibility of the study [71] (Checklist 1).

To ensure cultural safety protocols and kaupapa Māori methodology were adhered to, 2 female kaumatua or elders (RS and JT) assisted with access to the various communities and remained present throughout the interviews. This approach ensured that the elders provided a moderating influence on the research process, and Māori cultural protocols were adhered to. Unexpectedly, Cyclone Gabrielle occurred partway through the interviews, prompting the research team to include accounts from participants about this event during the data collection phase. However, a caution has been raised regarding the structured interviews to ensure they cover both the COVID-19 pandemic and Cyclone Gabrielle equally. Moreover, limitations include the small number of participants and reliance on family members to assist with interviews, which may affect data accuracy.

Another limitation was that the exact age of the participants was not collected or documented. The precise ages of individuals—especially older people—sometimes could not be determined. The exact age could only be determined if the birth was recorded. Moreover, in traditional Māori and Pacific societies prior to European contact, written birth records were not kept. People often tracked age in relation to events (such as wars, voyages, famines, and natural disasters) rather than exact dates.

Conclusions

Vulnerable situations such as pandemics and extreme weather events can have tremendous effects on the lives of Indigenous people who live remotely. This study provided insights into how the COVID-19 pandemic and Cyclone Gabrielle posed significant challenges for Māori communities, particularly for vulnerable older individuals, highlighting the importance of digital technology and telecommunication services in maintaining social connectedness and accessing health care services. The study also focused on the actions that should be taken to mitigate these challenges and overcome difficult

XSL•FO
RenderX

circumstances, such as the pandemic and Cyclone Gabrielle. The recommendations include a better health care system and improved coordination among care providers, user-friendly digital solutions, ensuring local funding and community services, establishing training processes for basic digital skills, and fostering leadership and partnerships with Indigenous New Zealanders.

Acknowledgments

The authors would like to thank Jean Te Huia, Rose Smith, and Rina Te Huia for their overall support, guidance, and assistance in this research. This research was funded by the Health Research Council's Māori Health Research Development Grant (No. 22/835). A small grant, that is, the Brad-Attain Research Placement for a student working in this project, was received from the University of Bradford, the Brad-Attain team. Moreover, a publishing grant was obtained from Linnaeus University as part of the University Library's research support.

Data Availability

The manuscript reported all supporting data.

Authors' Contributions

Conceptualization: DW
Data analysis: DW
Editing-original draft manuscript: DW
Methodology: DW
Writing: DW
Data analysis: ST
Editing: ST
Literature review: ST
Review Concept: ST
Writing: ST
Review Concept: MSRJ
Data analysis: MSRJ
Editing: MSRJ
Manuscript structure refining: MSRJ
Methodology: MSRJ
Overall coordination: MSRJ
Supervision: MSRJ
Writing: MSRJ

Conflicts of Interest

None declared.

Checklist 1
SRQR checklist. This study followed Standards for Reporting Qualitative Research (SRQR) [71].
[PDF File, 964 KB - aging_v8i1e73974_app1.pdf]

References

1. Nitschke JP, Forbes PAG, Ali N, et al. Resilience during uncertainty? Greater social connectedness during COVID-19 lockdown is associated with reduced distress and fatigue. Br J Health Psychol 2021 May;26(2):553-569. [doi: 10.1111/bjhp.12485] [Medline: 33099800]
2. Kanakis K, McShane C. Preparing for disaster: preparedness in a flood and cyclone prone community. Aust J Emerg Manag 2016;31(2):18-24 [FREE Full text]
3. Wepa D. Digital disparities: tech solutions for indigenous communities. Open Access Government 2025;46(1):26-27. [doi: 10.56367/OAG-046-11905]
4. Lips A, Eppel E. Effects of covid-19 on digital public services. 2021 URL: https://openaccess.wgtn.ac.nz/articles/report/Effects_of_Covid-19_on_digital_public_services/24502513?file=43046464 [accessed 2025-10-16]
5. The Intergovernmental Panel on Climate Change. AR5 synthesis report: climate change 2014. 2014. URL: https://www.ipcc.ch/report/ar5/syr [accessed 2025-02-19]

XSL•FO
RenderX

6. Jefferies S, French N, Gilkison C, et al. COVID-19 in New Zealand and the impact of the national response: a descriptive epidemiological study. Lancet Public Health 2020 Nov;5(11):e612-e623. [doi: 10.1016/S2468-2667(20)30225-5] [Medline: 33065023]
7. New Zealand Government. Māori culture and heritage. 2021 URL: https://www.govt.nz/browse/history-culture-and-heritage/maori-language-culture-and-heritage/maori-culture-and-heritage/?utm_source=chatgpt.com [accessed 2025-10-16]
8. Cumming J. Going hard and early: Aotearoa New Zealand's response to Covid-19. Health Econ Policy Law 2022 Jan;17(1):107-119. [doi: 10.1017/S174413312100013X] [Medline: 33663626]
9. Gauld R. A review of public policies on Covid-19: the New Zealand experience. PAP 2023 May 17;26(1):10-20. [doi: 10.1108/PAP-04-2022-0028]
10. Beattie A, Priestley R. Fighting COVID-19 with the team of 5 million: Aotearoa New Zealand government communication during the 2020 lockdown. Soc Sci Humanit Open 2021;4(1):100209. [doi: 10.1016/j.ssaho.2021.100209]
11. Kohlenbach T. Five process lessons from New Zealand's response to the Covid-19 pandemic. 2020. URL: https://www.processexcellencenetwork.com/organizational-change/articles/a-team-of-five-million-five-process-lessons-from-new-zealands-response-to-the-covid-19-pandemic [accessed 2025-01-21]
12. Rocque RJ, Beaudoin C, Ndjaboue R, et al. Health effects of climate change: an overview of systematic reviews. BMJ Open 2021 Jun 9;11(6):e046333. [doi: 10.1136/bmjopen-2020-046333] [Medline: 34108165]
13. Clarke B, Otto F, Stuart-Smith R, Harrington L. Extreme weather impacts of climate change: an attribution perspective. Environ Res: Climate 2022 Sep 1;1(1):012001. [doi: 10.1088/2752-5295/ac6e7d]
14. Makharia GK, Sadeghi A, Leddin D, Costello A. Impact of climate change on vulnerable populations. Gut 2023 Dec;72(12):2201-2204. [doi: 10.1136/gutjnl-2023-331195]
15. Coronese M, Lamperti F, Keller K, Chiaromonte F, Roventini A. Evidence for sharp increase in the economic damages of extreme natural disasters. Proc Natl Acad Sci USA 2019 Oct 22;116(43):21450-21455. [doi: 10.1073/pnas.1907826116]
16. IPCC. Climate change 2014: mitigation of climate change contribution of working group III to the fifth assessment report of the intergovernmental panel on climate change. 2014 URL: https://www.dia.govt.nz/Government-Inquiry-into-the-Response-to-the-North-Island-Severe-Weather-Events Department of Internal Affairs [accessed 2025-10-16]
17. Aldridge J, Bell R. Climate change impacts to extreme weather events associated with insured losses in New Zealand: a review. Environ Res: Climate 2025 Mar 31;4(1):012001. [doi: 10.1088/2752-5295/ada1f0]
18. Department of Internal Affairs. Report of the government inquiry into the response to the north island severe weather events. 2024 URL: https://knowledgeauckland.org.nz/media/lvbnuvqf/report-of-the-government-inquiry-response-north-island-severe-weather-events-april-2024.pdf [accessed 2025-09-16]
19. Dimitrof S. Cyclone Hale leaves its mark, not finished with Te Tai Rāwhiti. 2023 URL: https://www.teaonews.co.nz/2023/01/11/cyclone-hale-leaves-its-mark-not-finished-with-te-tai-rawhiti/ [accessed 2025-10-16]
20. Thames Coromandel District Council. Thames Coromandel District Council Cyclone Hale leaves long tail of aftermath in our district. URL: https://www.tcdc.govt.nz/Our-Council/News-Media-and-Public-Notices/Latest-News/Cyclone-Hale-leaves-long-tail-of-aftermath-in-our-district [accessed 2025-01-21]
21. McCabe K. Cyclone Gabrielle causes national state of emergency in New Zealand. 2023. URL: https://www.rmets.org/metmatters/cyclone-gabrielle-causes-national-state-emergency-new-zealand [accessed 2025-01-21]
22. Stats NZ. Impacts of cyclones and flooding on economic statistics in Aotearoa New Zealand: conceptual and measurement methods. 2023. URL: https://www.stats.govt.nz/reports/impacts-of-cyclones-and-flooding-on-economic-statistics-in-aotearoa-new-zealand-conceptual-and-measurement-methods-2023 [accessed 2025-01-21]
23. McClure T. Cyclone Gabrielle worst storm to hit New Zealand this century, says PM. . 2023 URL: https://www.theguardian.com/world/2023/feb/13/cyclone-gabrielle-new-zealand-declares-national-state-of-emergency#:~:text=%E2%80%9CCyclone%20Gabrielle%20is%20the%20most,Chris%20Hipkins%2C%20said%20on%20Tuesday [accessed 2024-10-16]
24. United Nations. The impact of digital technologies. URL: https://www.un.org/en/un75/impact-digital-technologies [accessed 2025-01-21]
25. Grimes A, White D. Digital inclusion and wellbeing in New Zealand. Motu Economic and Public Policy Research. : Motu Economic and Public Policy Research; 2019 URL: https://www.motu.nz/our-research/wellbeing-and-macroeconomics/well-being-and-sustainability-measures/digital-inclusion-and-wellbeing-in-new-zealand [accessed 2025-09-17]
26. Crengle S, Davie G, Whitehead J, de Graaf B, Lawrenson R, Nixon G. Mortality outcomes and inequities experienced by rural Māori in Aotearoa New Zealand. Lancet Reg Health West Pac 2022 Nov;28:100570. [doi: 10.1016/j.lanwpc.2022.100570] [Medline: 36042896]
27. Gurney J, Stanley J, Sarfati D. The inequity of morbidity: disparities in the prevalence of morbidity between ethnic groups in New Zealand. J Comorb 2020;10:2235042X20971168. [doi: 10.1177/2235042X20971168] [Medline: 33224894]
28. Wilson N, Barnard LT, Summers JA, Shanks GD, Baker MG. Differential mortality rates by ethnicity in 3 influenza pandemics over a century, New Zealand. Emerg Infect Dis 2012;18(1):71-77. [doi: 10.3201/eid1801.110035]

XSL•FO
RenderX

29. Walker S, Eketone A, Gibbs A. An exploration of kaupapa Maori research, its principles, processes and applications. Int J Soc Res Methodol 2006 Oct;9(4):331-344. [doi: 10.1080/13645570600916049]
30. Smith GH, Hoskins T, Jones A. Kaupapa māori theory: indigenous transforming of education. In: Hoskins TK, Jones A, editors. Critical Conversations in Kaupapa Maori: Hui Publishers; 2017:70-81 URL: https://natlib.govt.nz/records/38199629?utm_source=chatgpt.com [accessed 2025-10-16]
31. Wepa D, Wilson D. Struggling to be involved: an interprofessional approach to examine Maori whanau engagement with healthcare services. J Nurs Res Pract 2019;03(3):1-5. [doi: 10.37532/jnrp.2019.3(3).1-5]
32. Cram F, Adcock A. Kaupapa māori research. In: Liamputtong P, editor. Handbook of Qualitative Cross-Cultural Research Methods: Edward Elgar Publishing; 2022:56-84. [doi: 10.4337/9781800376625.00013]
33. Braun V, Clarke V. What can "thematic analysis" offer health and wellbeing researchers? Int J Qual Stud Health Well-being 2014;9:26152-26152. [doi: 10.3402/qhw.v9.26152] [Medline: 25326092]
34. Jabin MSR. The need for a refined classification system and national incident reporting system for health information technology-related incidents. Front Digit Health 2024;6:1422396. [doi: 10.3389/fdgth.2024.1422396] [Medline: 39131183]
35. Embregts P, Taminiau EF, Heerkens L, Schippers AP, van Hove G. Collaboration in inclusive research: competencies considered important for people with and without intellectual disabilities. Policy Practice Intel Disabi 2018 Sep;15(3):193-201. [doi: 10.1111/jppi.12248]
36. Burnett C, Purkey E, Davison CM, et al. Spirituality, community belonging, and mental health outcomes of indigenous peoples during the COVID-19 pandemic. Int J Environ Res Public Health 2022 Feb 21;19(4):2472. [doi: 10.3390/ijerph19042472] [Medline: 35206662]
37. Bryant H. How māori support whānau to promote digital inclusion. : Massey Research Online; 2024 URL: https://mro.massey.ac.nz/bitstreams/2999ef9c-5958-44e9-995c-2f82b4e8635a/download?utm_source=chatgpt.com [accessed 2025-09-17]
38. Ottoni CA, Winters M, Sims-Gould J. "We see each other from a distance": neighbourhood social relationships during the COVID-19 pandemic matter for older adults' social connectedness. Health Place 2022 Jul;76:102844. [doi: 10.1016/j.healthplace.2022.102844]
39. Wickramaratne PJ, Yangchen T, Lepow L, et al. Social connectedness as a determinant of mental health: a scoping review. PLoS ONE 2022;17(10):e0275004. [doi: 10.1371/journal.pone.0275004] [Medline: 36228007]
40. Oetzel JG, Ruru S, Zhang Y, et al. Enhancing well-being and social connectedness for Māori elders through a peer Education (Tuakana-Teina) programme: a cross-sectional baseline study. Front Public Health 2021;9:775545. [doi: 10.3389/fpubh.2021.775545] [Medline: 34957027]
41. Power T, Wilson D, Best O, et al. COVID-19 and Indigenous Peoples: an imperative for action. J Clin Nurs 2020 Aug;29(15-16):2737-2741. [doi: 10.1111/jocn.15320] [Medline: 32412150]
42. Yashadhana A, Pollard-Wharton N, Zwi AB, Biles B. Indigenous Australians at increased risk of COVID-19 due to existing health and socioeconomic inequities. Lancet Reg Health West Pac 2020 Aug;1:100007. [doi: 10.1016/j.lanwpc.2020.100007] [Medline: 34173592]
43. Grodecki K, Goulding A, Suraweera N. Building strong digital foundations: digital fluency-focused partnerships among early childhood education service providers and public libraries in Aotearoa New Zealand. J Aust Libr Inf Assoc 2024 Oct;73(4):462-484. [doi: 10.1080/24750158.2024.2386643]
44. Russo JE, McCool RR, Davies L. VA telemedicine: an analysis of cost and time savings. Telemed J E Health 2016 Mar;22(3):209-215. [doi: 10.1089/tmj.2015.0055] [Medline: 26305666]
45. Coombes CE, Gregory ME. The current and future use of telemedicine in infectious diseases practice. Curr Infect Dis Rep 2019 Oct 19;21(11):41. [doi: 10.1007/s11908-019-0697-2] [Medline: 31630276]
46. World Health Organization. Global strategy on digital health 2020-2025. 2021 URL: https://www.who.int/publications/i/item/9789240020924?utm_source=chatgpt.com [accessed 2025-08-23]
47. Jabin MSR, Pan D. Software-related challenges in Swedish healthcare through the lens of incident reports: a desktop study. Digit Health 2023;9:20552076231203600. [doi: 10.1177/20552076231203600] [Medline: 37744748]
48. Wepa D, Smith R, Gemmell L. Reconnecting Māori in a post-COVID-19 world: a blessing in disguise. AlterNative Int J Indig Peoples 2023 Dec;19(4):873-881. [doi: 10.1177/11771801231198131]
49. Pan D, Nilsson E, Jabin MSR. A review of incidents related to health information technology in Swedish healthcare to characterise system issues as a basis for improvement in clinical practice. Health Informatics J 2024;30(3):14604582241270742. [doi: 10.1177/14604582241270742] [Medline: 39116887]
50. Reweti A. Understanding how whānau-centred initiatives can improve Māori health in Aotearoa New Zealand. Health Promot Int 2023 Aug 1;38(4):daad070. [doi: 10.1093/heapro/daad070] [Medline: 37493232]
51. Saeri AK, Cruwys T, Barlow FK, Stronge S, Sibley CG. Social connectedness improves public mental health: investigating bidirectional relationships in the New Zealand attitudes and values survey. Aust N Z J Psychiatry 2018 Apr;52(4):365-374. [doi: 10.1177/0004867417723990] [Medline: 28803484]
52. Elers P, Dutta MJ, Elers S. Culturally centring digital inclusion and marginality: a case study in Aotearoa New Zealand. New Media & Society 2022 Feb;24(2):311-327. [doi: 10.1177/14614448211063180]
53. Motsumi MM, Nemakonde LD. A framework to integrate indigenous knowledge into disaster risk reduction to build disaster resilience: insights from rural South Africa. DPM 2024 Dec 16;33(6):73-85. [doi: 10.1108/DPM-08-2024-0220]

XSL•FO
RenderX

54. Radio New Zealand. Cyclone Gabrielle. 2023. URL: https://www.odc.govt.nz/our-council/news/?item=id:21344qj1e17q9svo7iq8 [accessed 2024-12-05]
55. Mason K, Lindberg K, Haenfling C, et al. Social vulnerability indicators for flooding in Aotearoa New Zealand. Int J Environ Res Public Health 2021 Apr 9;18(8):3952. [doi: 10.3390/ijerph18083952] [Medline: 33918651]
56. Akram S, Mushtaq S. Environmental change and floods: the long-ignored effects of displacement on mental health. Front Public Health 2024;12:1434123. [doi: 10.3389/fpubh.2024.1434123] [Medline: 39744366]
57. Marshall A, Wilson CA, Dale A. Telecommunications and natural disasters in rural Australia: the role of digital capability in building disaster resilience. J Rural Stud 2023 May;100:102996. [doi: 10.1016/j.jrurstud.2023.03.004]
58. Fjordside S, Morville A. Factors influencing older people's experiences of participation in autonomous decisions concerning their daily care in their own homes: a review of the literature. Int J Older People Nurs 2016 Dec;11(4):284-297. [doi: 10.1111/opn.12116] [Medline: 27019374]
59. Warschauer M. Technology and Social Inclusion: Rethinking the Digital Divide: MIT Press; 2004. [doi: 10.7551/mitpress/6699.001.0001]
60. Tenbensel T, Cumming J, Willing E. The 2022 restructure of Aotearoa New Zealand's health system: will it succeed in advancing equity where others have failed? Health Policy 2023 Aug;134:104828. [doi: 10.1016/j.healthpol.2023.104828] [Medline: 37263868]
61. Jabin MSR, Nilsson E, Nilsson AL, Bergman P, Jokela P. Digital health testbeds in Sweden: an exploratory study. Digit HEALTH 2022;8:20552076221075194. [doi: 10.1177/20552076221075194] [Medline: 35186314]
62. Hartnett M, Fields A. Digital inclusion in New Zealand. JOFDL 2020;23(2):1-4. [doi: 10.61468/jofdl.v23i2.391]
63. Wills O, Moore PM. Addressing the digital divide: the economic case for increasing digital inclusion: NZIER report to the Digital Council for Aotearoa. : Digital Council for Aotearoa; 2022 URL: https://www.digital.govt.nz/assets/Uploads/Addressing-the-digital-divide-the-economic-case-for-increasing-digital-inclusion.pdf [accessed 2025-08-17]
64. Garneau M. Barriers to economic development in indigenous communities. report of the standing committee on indigenous and northern affairs. : House of Commons; 2022 URL: https://www.ourcommons.ca/Content/Committee/441/INAN/Reports/RP11714230/inanrp02/inanrp02-e.pdf [accessed 2025-09-17]
65. Jabin MSR, Pan D, Nilsson E. Characterizing patient details-related challenges from health information technology-related incident reports from Swedish healthcare. Front Digit Health 2024;6:1260521. [doi: 10.3389/fdgth.2024.1260521] [Medline: 38380372]
66. Jabin MSR, Steen M, Wepa D, Bergman P. Assessing the healthcare quality issues for digital incident reporting in Sweden: incident reports analysis. Digit HEALTH 2023;9:20552076231174307. [doi: 10.1177/20552076231174307] [Medline: 37188073]
67. Jabin MSR. Operational disruption in healthcare associated with software functionality issue due to software security patching: a case report. Front Digit Health 2024;6:1367431. [doi: 10.3389/fdgth.2024.1367431] [Medline: 38550716]
68. Department of Internal Affairs Te Tari Taiwhenua, New Zealand Government. The Digital Inclusion Blueprint—Te Mahere Mō Te Whakaurunga Matihiko 2019. URL: https://www.digital.govt.nz/assets/Documents/113Digital-Inclusion-BlueprintTe-Mahere-mo-te-Whakaurunga-Matihiko.pdf [accessed 2025-09-17]
69. Digital inclusion user insights—māori. Ngā tirohanga kaiwhakamahi whakaurunga matihiko. URL: https://www.digital.govt.nz/dmsdocument/177~report-digital-inclusion-user-insights-maori/html [accessed 2015-08-13]
70. Aziz MGA. Digital inclusion for indigenous communities: bridging the digital divide. 2024 URL: https://www.linkedin.com/pulse/digital-inclusion-indigenous-communities-bridging-abdul-aziz-odrsc [accessed 2025-09-17]
71. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for Reporting Qualitative Research. Acad Med 2014 Sep;89(9):1245-1251. [doi: 10.1097/ACM.0000000000000388]

Edited by A Garcia; submitted 14.03.25; peer-reviewed by FL Uribe, R Samuriwo; revised version received 02.10.25; accepted 07.10.25; published 28.10.25.

Please cite as:
Wepa D, Thomas S, Jabin MSR
The Experience and Impact of Digital Technologies on Indigenous Populations in New Zealand During the COVID-19 Pandemic and Cyclone Gabrielle: The Kaupapa Māori Methodology
JMIR Aging 2025;8:e73974
URL: https://aging.jmir.org/2025/1/e73974
doi:10.2196/73974

© Dianne Wepa, Shiji Thomas, Md Shafiqur Rahman Jabin. Originally published in JMIR Aging (https://aging.jmir.org), 28.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium,

XSL•FO
RenderX

provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital Literacy and Its Association With Subjective Health Status and Healthy Lifestyle Behaviors Among Korean Older Adults: Cross-Sectional Study

Soon Young Lee[1*], MPH; Yejin Kim[2*], MPH; Bomgyeol Kim[3], PhD; Sang Gyu Lee[4], PhD; Suk-Yong Jang[5], PhD; Tae Hyun Kim[6], PhD

[1]Department of Strategic Planning Part, Kangbuk Samsung Hospital, Seoul, Republic of Korea

[2]Department of Public Health, Graduate School, Yonsei University, Seoul, Republic of Korea

[3]Mo-Im Kim Nursing Research Institute, College of Nursing, Yonsei University, Seoul, Republic of Korea

[4]Department of Preventive Medicine, College of Medicine, Yonsei University, Seoul, Republic of Korea

[5]Department of Biohealth Policy Analysis, Graduate School of Transdisciplinary Health Sciences, Yonsei University, Seoul, Republic of Korea

[6]Department of Biohealth Industry Management, Graduate School of Transdisciplinary Health Sciences, Yonsei University, 50-1 Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Tae Hyun Kim, PhD
Department of Biohealth Industry Management, Graduate School of Transdisciplinary Health Sciences, Yonsei University, 50-1 Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea

Abstract

Background: With an aging population driven by advances in medical technology, digital literacy has become essential for improving the quality of life of older adults, enhancing access to health information, and promoting healthy lifestyles. Furthermore, the COVID-19 pandemic may have influenced the subjective health perceptions and healthy lifestyle behaviors of older adults. However, there is limited research exploring the relationship between digital literacy, subjective health perceptions, and healthy lifestyle behaviors in Korea.

Objective: This study aimed to investigate digital literacy's impact on Korean older adults' subjective health status and healthy lifestyle behaviors.

Methods: Data of 8664 respondents (aged 65 years and older) from the 2020 National Survey of the Older Koreans were analyzed. Digital literacy was measured based on the use of IT devices (ITDs), difficulty using online information, and inconvenience of ITDs. Statistical analyses, such as the Rao-Scott chi-square test, Wilcoxon rank sum test, and multiple regression analysis, were conducted.

Results: Respondents with above-average ITD use (adjusted odds ratio [aOR] 1.73, 95% CI 1.50 - 1.99) and less difficulty using online information (aOR 1.41, 95% CI 1.24 - 1.61) had higher odds of perceiving themselves as healthy. Conversely, high difficulty using ITDs was associated with lower odds of respondents perceiving themselves as healthy (aOR 0.84, 95% CI 0.82 - 0.87). Furthermore, high ITD use predicted engagement in healthy lifestyle behaviors (aOR 1.51, 95% CI 1.33 - 1.72), whereas high difficulty using ITDs predicted lower odds of engagement (aOR 0.94, 95% CI 0.92 - 0.97). In contrast, there was no difference in the odds of engaging in healthy lifestyle behaviors regardless of difficulty using online information (aOR 1.03, 95% CI 0.92 - 1.15).

Conclusions: This study underscores the significant association between digital literacy and improved health outcomes among older adults. Promotion of digital literacy and relevant policies is essential to help older adults effectively obtain health information online, thereby improving their quality of life and overall health.

(JMIR Aging 2025;8:e64974) doi:10.2196/64974

KEYWORDS

digital literacy; healthy lifestyle behaviors; older adults; subjective health status; quality of life

XSL•FO
RenderX

Introduction

Advances in medical technology have increased life expectancy, thereby contributing to the growth of an aging population worldwide [1]. South Korea became an aging society in 2017, and it is predicted to transition to a super-aged society by 2025 [2]. Old age is characterized by a general decline in physical and mental functioning, leading to a range of health issues, including senile and chronic degenerative diseases and impaired daily functioning [3]. Prevention, treatment, and management of diseases are necessary to improve of older adults' health care and quality of life. In addition, health information and proactive responses for health care are important [4]. Particularly, the first generation of Korean "baby boomers" (born between 1955 and 1972), specifically those born in 1955, has been incorporated into the older adult population since 2020. Thus, this is a crucial time for discussion on older adults' health care.

Generally, older adults obtain health information from health care providers, friends and relatives, television, newspapers, radio, and other mass media to maintain their health [5]. However, rapid technological developments have increased the opportunities for acquiring health information and news on the internet, leading to digital literacy's increased importance [6-8]. Digital technology initially spread among young people, men, and highly educated people and is gradually spreading to older adults, women, and less educated people. The aging trend and the rapid development of new media technology suggest that older adults will have to use digital technology more in the future [9]. Digital literacy comprises the ability to understand the information generated on a digital device and implement it accordingly [10]. In an information society, digital literacy's importance is expanding to all areas of life as a key factor in improving quality of life [11,12].

The manner in which older adults perceive their health informs how they relate to society. In addition, their subjective well-being is intimately linked to their perceived health status [13]. Meanwhile, health concerns can be a primary motivation for searching health information online [14]. Specifically, the use of the internet promotes the health of older adults by improving access to health information, encouraging healthy lifestyles, and enhancing social interaction [15]. While older adults recognize the effectiveness and usefulness of using health information through digital media such as the internet for health care, they experience challenges when actually attempting to obtain health information online [16,17]. According to a 2022 report on the digital divide, vulnerable groups such as people with disabilities, low-income individuals, farmers, and older adults have markedly lower digital use capabilities, with the latter experiencing the highest level of information vulnerability [18].

The growing importance of digital literacy for older adults has prompted a variety of studies. However, most Korean research is survey-based or restricted to specific geographic regions or disease populations, focusing primarily on health knowledge, mental health aspects such as depression and self-efficacy, or quality of life [19-22]. While subjective health perceptions and healthy lifestyle behaviors of older adults may have been influenced by COVID-19 pandemic, these studies were conducted before the pandemic [23-25]. Furthermore, the rapid progress of aging and the advancement of IT, along with gaps in the acquisition of information, unchecked spread of health information, and dissemination of unverified information, have created growing concerns among the older adult population vis-à-vis the lack of reliable information [26-28].

This study seeks to examine 2 key hypotheses: (1) older adults with higher levels of digital literacy are more likely to perceive themselves as healthy and (2) higher levels of digital literacy among older adults are associated with a greater likelihood of engaging in health-promoting behaviors. Using a nationally representative sample, we aim to investigate the association between different aspects of digital literacy and older adults' subjective health status and healthy lifestyle behaviors.

Methods

Data Source and Study Population

This cross-sectional study used data from the 2020 National Survey of the Older Koreans (NSOK), conducted by the Ministry of Health and Welfare and the Korea Institute for Health and Social Affairs (KIHASA). The NSOK is a survey conducted every 3 years since 2008, in accordance with Article 5 of the Welfare of Senior Citizens Act. It systematically collects data from more than 10,000 older adults aged 65 years and older who are living in ordinary residences across 17 cities and provinces nationwide. Trained surveyors carry out the survey in 969 districts selected based on regional representation. They survey all households with older residents, considering factors such as area of residence, age, gender, education, and marital status [29]. This survey aimed to identify the living conditions, characteristics, and needs of older adults to develop welfare policies to improve their quality of life. In addition, it sought to develop policies in response to Korea's aging society by examining changes in the characteristics of older adults through the accumulation of time-series data [29].

NSOK calculates weights in 3 steps: using design weights, nonresponse-adjusted weights, and poststratification weights. First, the design weights are directly derived from the sampling design, using a 2-stage cluster sampling method. This method involves probability proportional to size (PPS) sampling, where sample areas are selected based on the number of households in each area, followed by the selection of households within each sampled area. Nonresponse adjustment is conducted by upweighting responding households using the reciprocal of the response rate. Final weights are calculated by applying a raking ratio method, and respondent weights, adjusted for design and nonresponse weights, are poststratified to match the distribution of older adults in the population, thereby enhancing the accuracy of estimates. The 2020 NSOK used data from the 2019 Census, the most recent available at the time of survey completion, to calculate final weights by applying population data for older adults by gender, age group (65 - 74 years or 75 years and older), and household size (1-person household, 2+ persons) across the 17 metropolitan areas [21].

XSL•FO
RenderX

The study excluded 1256 of the 10,097 respondents who answered "don't know" or "not applicable" to the digital literacy survey. In addition, if the respondent has a physical or mental illness, the cohabitant or a close representative of the study subject responds on their behalf. In this case, only the objective information of the older adults was surveyed, and the main focus of this study was not surveyed [28]. Therefore, the results were generalized to older Koreans by excluding the 177 respondents who responded on behalf of their cohabitant. These exclusions ensured the objectivity of the health indicator responses, as the survey focused on the subjective judgment of the original respondents [29]. Finally, the study included 8664 participants.

Ethical Considerations

The Institutional Review Board (IRB) of the Yonsei Medical Center approved a consent exemption for this study (No. 4-2023-0361). The 2020 Elderly Survey used in this study was anonymized by the data provider, KIHASA, and the tools and processes used in the survey were approved by the IRB of KIHASA (No. 2020 - 36).

Measurements

Subjective Health Status and Healthy Lifestyle Behaviors

The dependent variables of this study included subjective health status and healthy lifestyle behaviors. Subjective health status has been widely used in various studies as a valid proxy for actual health status, given its reliable predictive value for health outcomes [30-32]. It was categorized into good (very healthy and healthy) and poor (fair, poor, and very poor). Healthy lifestyle behaviors were defined as smoking cessation, decreased alcohol consumption, and exercise [33]. Smoking cessation was defined as having quit smoking, decreased alcohol consumption was defined as not drinking at all in the previous year or drinking less than twice a week, and exercise was defined as answering "yes" to the question, "Do you usually exercise for at least 10 minutes on a consistent basis?"

Digital Literacy

This study's main variable was digital literacy, which was divided into 3 subcategories. It was scored based on the degree of use of IT devices (ITDs), difficulty using information provided online, and inconvenience of ITD usage.

For the item on the degree of ITD usage, 1 point was assigned to each of the activities in Textbox 1, which involved using a PC, mobile phone, or tablet. The sample was divided into 2 groups according to the respondent's score, namely above average (3 points or more) and below average (3 points or less).

For the item on the difficulty of using information provided online, the degree of difficulty in using information or services required for daily life via the internet was categorized into "difficult" and "not difficult." The inconvenience of ITD usage was assessed considering tasks such as making online reservations, conducting purchases using machines such as kiosks and automated teller machines, and shopping at stores that only accept cards. For each task, the level of inconvenience was scored on a scale from 0 ("not at all inconvenient") to 2 ("inconvenient"). Finally, the overall level of inconvenience of using ITDs was rated on a scale from 0 ("not at all inconvenient") to 8 ("inconvenient"). A Cronbach α of 0.88 was identified for ITD usage, and 0.80 for inconvenience of using ITDs.

Textbox 1. Activities that involved using a PC, mobile phone, or tablet.

- Receiving messages.
- Sending messages.
- Searching for and viewing information.
- Taking photos and videos.
- Listening to music.
- Playing games.
- Watching videos.
- Using social media.
- Using e-commerce platforms.
- Using online financial transactions.
- Searching for and installing applications.
- Other activities.

Covariates

The covariates included sociodemographic characteristics and health-related factors. Sociodemographic characteristics included gender, age, residence, educational level, household types, and household income. Age was classified into five 5-year intervals (65 - 69, 70 - 74, 75 - 79, 80 - 84, and ≥85 years old). Residence was classified into "rural (eup/myeon)" and "urban (dong)." Educational status was classified into "elementary school or less," "middle and high school," and "university and above." Household type was classified into "single," "married couple," and "cohabitant (nonspouse)." Household income encompassed all household members' earned, business, property, and transfer incomes. This comprised the annualized gross household income for the most recent year (July 1, 2019, to

XSL•FO
RenderX

June 30, 2020), converted to a monthly basis and divided into quartiles.

Health-related factors included the number of chronic diseases respondents had. Chronic diseases included conditions that lasted for more than 3 months and had been diagnosed by a doctor. These were categorized as "none," "1‐2 chronic diseases," and "3 or more chronic diseases," encompassing 32 conditions.

Statistical Analysis

A frequency analysis was conducted to identify the participants' general characteristics. The Rao-Scott chi-square test and Wilcoxon rank sum test were conducted to evaluate and compare these characteristics. Multiple logistic regression analysis examined the associations of digital literacy with subjective health status and healthy lifestyle behaviors. In addition, a subgroup analysis was conducted to determine whether gender led to differences in the association of digital literacy with subjective health status and healthy lifestyle behaviors. All analyses were conducted using SAS software (version 9.4; SAS Institute), while statistical significance was set to $P<.05$. All analyses were weighted by NSOK to account for differences in survey completion rates across NSOK's survey regions.

Results

The participants' demographic and socioeconomic characteristics are summarized in Table 1. Of the 8664 study participants, all weighted by region, 55.9% (4649) had above-average scores in ITD usage, while 72.3% (6264) found it difficult to use information available online. The level of inconvenience of ITD usage was 4.2 out of 8.

Participants were mostly female (n=5077, 55.7%), aged 65‐69 years (n=3282, 35.8%), urban residents (n=6256, 76.1%), educated up to middle and high school (n=4571, 54.4%), married (n=4502, 59.7%), in household income level 4 (n=2169, 29.1%), and had 1‐2 chronic diseases (n=4936, 56.6%).

XSL•FO
RenderX

Table . Distribution of demographic and socioeconomic characteristics.

Variable	Total, n (%)[a]	Subjective health status			Healthy lifestyle behaviors		
		Healthy, n (%)[a]	Nonhealthy, n (%)[a]	*P* value[b]	Engagement, n (%)[a]	Nonengagement, n (%)[a]	*P* value[b]
Total	8664 (100)	4471 (51.1)	4193 (48.9)		3945 (46.9)	4719 (53.1)	
Degree of ITD[c] usage				<.001			<.001
Above average	4649 (55.9)	3038 (64.8)	1611 (35.2)		2329 (51.8)	2320 (48.2)	
Below average	4015 (44.1)	1433 (33.8)	2582 (66.2)		1616 (40.7)	2399 (59.3)	
Using information provided online				<.001			.05
Difficult	6264 (72.3)	2898 (45.6)	3366 (0.5)		2773 (46.1)	3491 (53.9)	
Nondifficult	2400 (27.7)	1573 (65.6)	827 (0.3)		1172 (48.9)	1228 (51.1)	
Inconvenience of ITD usage (score), mean (SD)	4.2 (2)	3.7 (1.9)	4.8 (2)	<.001	4.1 (2)	4.4 (2)	<.001
Gender				<.001			<.001
Male	3587 (44.3)	2097 (57)	1490 (43)		1427 (41.4)	2160 (58.6)	
Female	5077 (55.7)	2374 (46.5)	2703 (53.5)		2518 (51.3)	2559 (48.7)	
Age (years)				<.001			<.001
65 - 69	3282 (35.8)	2235 (69.3)	1047 (30.7)		1567 (48.2)	1715 (51.8)	
70 - 74	2193 (23.9)	1147 (53.9)	1046 (46.1)		1049 (48.6)	1144 (51.4)	
75 - 79	1668 (22.4)	640 (38.3)	1028 (61.7)		769 (48.8)	899 (51.2)	
80 - 85	1085 (12.9)	322 (27.2)	763 (72.8)		414 (40.9)	671 (59.1)	
≥85	436 (5)	127 (27)	309 (73)		146 (37.2)	290 (62.8)	
Residence				<.001			<.001
Urban (dong)	6256 (76.1)	3361 (52.8)	2895 (47.2)		3055 (50.4)	3201 (49.6)	
Rural (eup/myeon)	2408 (23.9)	1110 (45.9)	1298 (54.1)		890 (35.8)	1518 (64.2)	
Educational status				<.001			<.001
Elementary school or less	3609 (39.1)	1320 (34.3)	2289 (63.4)		1522 (42.5)	2087 (57.5)	
Middle and high school	4571 (54.4)	2790 (60.6)	1781 (39)		2123 (48.2)	2448 (51.8)	
University and above	484 (6.5)	361 (73.3)	123 (25.4)		300 (63)	184 (37)	

XSL•FO
RenderX

Variable	Total, n (%)[a]	Subjective health status			Healthy lifestyle behaviors		
		Healthy, n (%)[a]	Nonhealthy, n (%)[a]	*P* value[b]	Engagement, n (%)[a]	Nonengagement, n (%)[a]	*P* value[b]
Household type				<.001			.01
Single	2693 (19.8)	1117 (39.9)	1576 (60.1)		1160 (43)	1533 (57)	
Married couple	4502 (59.7)	2625 (57)	1877 (43)		2110 (47.8)	2392 (52.2)	
Cohabitant (nonspouse)	1469 (20.5)	729 (44.7)	740 (55.3)		675 (48.3)	794 (51.7)	
Household income				<.001			<.001
Level 1 (lowest)	2158 (20.7)	843 (38.4)	1315 (61.6)		799 (37.3)	1359 (62.7)	
Level 2	2171 (24.2)	965 (42.3)	1206 (57.7)		1044 (49.2)	1127 (50.8)	
Level 3	2166 (26)	1286 (58.2)	880 (41.8)		1020 (46.6)	1146 (53.4)	
Level 4 (highest)	2169 (29.1)	1377 (61.2)	792 (38.8)		1082 (52.1)	1087 (47.9)	
Number of chronic diseases				<.001			<.001
None	1533 (16.9)	1262 (82.1)	271 17.9		730 (48.2)	803 (51.8)	
1 - 2	4936 (56.6)	2696 (54.1)	2240 (45.9)		2160 (44.6)	2776 (55.4)	
3 or more	2195 (26.5)	513 (25)	1682 (75)		1055 (51.1)	1140 (48.9)	

[a]Weighted percentage.

[b]Rao-Scott chi-squared probability for comparison between (1) healthy and nonhealthy groups, and (2) groups engaging in healthy lifestyle behaviors and those not engaging in healthy lifestyle behaviors.

[c]ITD: IT device.

Of participants with above-average ITD usage, 64.8% (n=3038) perceived themselves as healthy and 51.8% (n=2329) engaged in healthy lifestyle behaviors. Of participants who reported difficulty using online information, 45.6% (n=2898) perceived themselves as healthy and 46.1% (n=2773) engaged in healthy lifestyle behaviors. Participants who perceived themselves to be healthy exhibited an ITD inconvenience score of 3.7 and those who engaged in healthy lifestyle behaviors exhibited a score of 4.1.

The association of digital literacy with subjective health status and healthy lifestyle behaviors is shown in Table 2. Compared with participants with below-average ITD usage, the odds of participants with above-average ITD usage were 1.73-times more likely to perceive themselves as healthy (adjusted odds ratio [aOR] 1.73, 95% CI 1.50 - 1.99, *P*<.001) and 1.51-times more likely to engage in healthy lifestyle behaviors (aOR 1.51, 95% CI 1.33 - 1.72, *P*<.001). Participants who did not find it difficult to use information available online had 1.41-times (aOR 1.41, 95% CI 1.24 - 1.61, *P*<.001) higher odds of perceiving themselves as healthy, compared with those who found it difficult. Higher ITD difficulty scores were also associated with 0.84-times lower odds of respondents perceiving themselves as healthy (aOR 0.84, 95% CI 0.82 - 0.87, *P*<.001) and 0.94-times lower odds of engaging in healthy lifestyle behaviors (aOR 0.94, 95% CI 0.92 - 0.97, *P*<.001).

XSL•FO
RenderX

Table . Digital literacy and its association with subjective health status and healthy lifestyle behaviors.

Variable	Perceived as healthy		Engaging in healthy lifestyle behaviors	
	aOR[a,b] (95% CI)	*P* value	aOR[a] (95% CI)	*P* value
Degree of ITD[c] usage				
Below average	Ref.[d]	—[e]	Ref.	—
Above average	1.73 (1.50-1.99)	<.001	1.51 (1.33-1.72)	<.001
Using information provided online				
Difficult	Ref.	—	Ref.	—
Nondifficult	1.41 (1.24-1.61)	<.001	1.03 (0.92-1.16)	.59
Inconvenience of ITD usage (score)	0.84 (0.82-0.87)	<.001	0.94 (0.92-0.97)	<.001

[a]Adjusted for covariates including gender, age, residence, educational status, household type, household income, and number of chronic diseases.

[b]aOR: adjusted odds ratio.

[c]ITD: IT device.

[d]Ref.: reference.

[e]Not applicable.

The association of digital literacy with subjective health status and healthy lifestyle behaviors by gender is shown in Tables 3 and 4. Both men and women with a higher than average degree of ITD usage were more likely to perceive themselves as healthy than those with an average or lower degree of ITD usage, and women in particular were nearly twice as likely (male aOR 1.64, 95% CI 1.34 - 1.99, *P*<.001; female aOR 1.91, 95% CI 1.60 - 2.27, *P*<.001). Similarly, people who did not experience any difficulties in using information provided online were more likely to perceive themselves as healthy than those who did experience difficulties, but this was more likely to be the case for men (male aOR 1.60, 95% CI 1.32 - 1.94, *P*<.001; female aOR 1.27, 95% CI 1.06 - 1.51, *P*=.01). Women with a higher degree of ITD usage were 1.54 times more likely to engage in healthy lifestyle behaviors than women with a lower degree of ITD usage, and this was slightly higher than for men (male aOR 1.47, 95% CI 1.20 - 1.81 *P*<.001; female aOR 1.54, 95% CI 1.30 - 1.82, *P*<.001). The less comfortable they were with using ITD, the less likely they were to perceive themselves as healthy (male aOR 0.84, 95% CI 0.80 - 0.88, *P*<.001; female aOR 0.85, 95% CI 0.81 - 0.88, *P*<.001) and to practice healthy lifestyles (male aOR 0.95, 95% CI 0.90 - 0.99, *P*=.02; female aOR 0.94, 95% CI 0.90 - 0.97, *P*<.001), with similar results for men and women.

Table . Digital literacy and its association with subjective health status by gender.

Variable	Perceived as healthy			
	Male		Female	
	aOR[a,b] (95% CI)	*P* value	aOR[a] (95% CI)	*P* value
Degree of ITD[c] usage				
Below average	Ref.[d]	—[e]	Ref.	—
Above average	1.64 (1.34-1.99)	<.001	1.91 (1.60-2.27)	<.001
Using information provided online				
Difficult	Ref.	—	Ref.	—
Nondifficult	1.60 (1.32-1.94)	<.001	1.27 (1.06-1.51)	.01
Inconvenience of ITD usage (score)	0.84 (0.80-0.88)	<.001	0.85 (0.81-0.88)	<.001

[a]Adjusted for covariates including age, residence, educational status, household type, household income, and number of chronic diseases.

[b]aOR: adjusted odds ratio.

[c]ITD: IT device.

[d]Ref.: reference.

[e]Not applicable.

XSL•FO
RenderX

Table . Digital literacy and its association with healthy lifestyle behaviors by gender.

Variable	Engaging in healthy lifestyle behaviors			
	Male		Female	
	aOR[a,b] (95% CI)	*P* value	aOR[a] (95% CI)	*P* value
Degree of ITD[c] usage				
Below average	Ref.[d]	—[e]	Ref.	—
Above average	1.47 (1.20-1.81)	<.001	1.54 (1.30-1.82)	<.001
Using information provided online				
Difficult	Ref.	—	Ref.	—
Nondifficult	1.08 (0.91-1.29)	.38	1.00 (0.85-1.18)	.96
Inconvenience of ITD usage (score)	0.95 (0.90-0.99)	.02	0.94 (0.90-0.97)	<.001

[a]Adjusted for covariates including age, residence, educational status, household type, household income, and number of chronic diseases.

[b]aOR: adjusted odds ratio.

[c]ITD: IT device.

[d]Ref: reference.

[e]Not applicable.

Discussion

Principal Findings and Comparison With Previous Work

This study examined the association between older people's digital literacy and their subjective health status and healthy lifestyle behaviors. Research has shown that older people with a high level of digital literacy, those who actively use information devices, and those who have easy access to online information are more likely to feel healthy and maintain healthy lifestyle behaviors. Conversely, those who experience difficulty accessing information online and are uncomfortable using information devices are less likely to engage in healthy lifestyle behaviors.

This study's findings are consistent with those of previous studies showing that both the use and amount of information affect subjective health status. In addition, whether older people own digital devices, how often they use them, and the extent of their internet activity are likely to affect their digital literacy and health [34-36]. According to a study in the United States, searching for and mastering relevant medical knowledge on the internet is beneficial for older adults with high blood pressure and heart disease, improving their health care and reducing disease rates [37]. In addition, using the internet could improve the health of older individuals by increasing their ability to exercise; for example, internet intervention programs helped them select exercise activities in the areas of endurance, flexibility, strengthening, and balance enhancement, which helped older adults increase their exercise frequency, thereby improving their physical health [38,39]. In China, older people with a high frequency of digital device usage experience a positive impact on their lifestyle, including improvements in diet, sleep, and exercise [15,40].

Empowering older people with a wide range of digital skills, from simple computer use to usage of social networks, emails, and online meetings, could enrich their lives, enhance their cognitive capabilities, and make them more productive and efficient members of society [41]. This could significantly contribute to the improvement of their perceived health status [42]. People who understand health information available online can use it effectively in their daily life to acquire valuable insights regarding their illnesses, increase their involvement as patient in the treatment process, and positively assess their subjective health status [43]. The internet provides a wealth of information, creating and increasing thematic content, especially for groups intimately associated with older adults, which helps maintain intergenerational relationships and promotes connected social capital [44]. With this social capital, older adults with a high level of digital literacy could fulfill their desire for belonging and love, which in turn could enable them to evaluate their health more positively [45].

However, unlike previous studies, this study found no difference in healthy lifestyle behaviors between those with and without difficulty using online information. Older people with high digital literacy have easier access to health information and use it effectively to positively impact their health care. Furthermore, they regularly access the internet for health information [27,35]. A previous study reported that adults aged 50 - 60 years who are not in good health access the internet for health care purposes because they want to be actively involved in health care decisions [46]. A few older individuals use emails and social networking sites to promote cancer screening to colleagues [47]. Online mediation, such as effectively managing hemoglobin A_{1c} levels in patients with diabetes through social networking site–based interventions, improved knowledge about the disease and had a positive impact on health behaviors [48,49].

Also in South Korea, Min [50] reported that the use of internet information decreased with age, regardless of physical and environmental accessibility, but this measure of use was comparable with this study's degree of ITD usage. Previous

XSL•FO
RenderX

studies have also highlighted the positive effects of online health information seeking psychological and physical domains, such as increased self-efficacy and improved health behaviors [51,52]. However, these studies were conducted before the COVID-19 pandemic, leaving a gap in understanding health behaviors during and after the pandemic.

The 2020 NSOK survey was conducted during the COVID-19 pandemic—a period marked by reduced outings among older adults due to social distancing policies and concerns about weakened immune systems. This context likely contributed to an overall decline in healthy lifestyle behaviors, particularly exercise, regardless of difficulties in using online information [23-25]. The reliance of Korean older adults on interpersonal sources for health information is deeply rooted in cultural values such as Confucianism and collectivism [53,54]. For health-related information, older adults in South Korea often rely on various sources, including health care professionals, mass media, local health centers, and family members. Particularly, adult children act as intermediaries in accessing and interpreting digital health information. In addition, a strong cultural trust in health care professionals reinforces a preference for direct consultations rather than independent online searches [55-57]. These sources play an important role in supporting health behaviors, even without the use of the internet, and help create an environment where information remains accessible despite low levels of online information use. In addition, older adults may have relied on household members or health center staff to mediate and explain online information, rather than accessing it directly [58-60]. This indirect approach could have resulted in an underestimation of the actual challenges older adults face when navigating online information.

Limitations

The present study had a few limitations. First, this is a part-sectional study. As we measured the variables at the time of the survey, we could only derive meaning in relation to the association between older adults' digital literacy and the causal relationship between their subjective health status and health practices. Second, as life expectancy increases, it is possible to divide the older adult population aged 85 years or older into different age groups. However, the NSOK is limited to detailed age-specific analysis. Nevertheless, this study used the latest data from national units, which ensured the representation of South Korea's older adults. In this study, digital literacy was divided into specific categories such as the degree of ITD usage, the difficulty of accessing online information, and the inconvenience of ITDs. This highlights the importance of linking older people's digital literacy to their subjective health status and health practices.

Conclusions

This study expands the existing knowledge on the subject by using a national-scale survey of South Korea's older adult population. To improve the subjective health and health practices of older adults, digital literacy education and related policies must be promoted. Furthermore, it is important to provide information education, so that older individuals can identify the required information, even among commercial and vague content. Health care and public institutions should provide information, booking systems, guidance, and other services to help older adults make use of these resources. We hope that this study will serve to improve older adults' quality of life and contribute to their health and welfare.

Data Availability

The data can be accessed from the Ministry of Health and Welfare and the Korea Institute for Health and Social Affairs (KIHASA) homepage. However, restrictions apply to data availability. The data were used with permission for the current study and, therefore, are not publicly available. The data are available from the corresponding author upon reasonable request and with prior permission from the KIHASA. Applications to use the National Survey of the Older Koreans (NSOK) data will be reviewed by the inquiry committee of research support, and once approved, raw data will be provided to the applicant for a fee.

Authors' Contributions

SYL was involved in study conception, statistical design, and manuscript drafting. YK and BK were involved in study conception, statistical design, data analysis, and manuscript drafting. SGL and SYJ oversaw the manuscript's development and helped revise it. THK was involved in study conception, statistical design, manuscript development, and manuscript revision. All the authors critically revised the manuscript for important intellectual content and approved its final version.

Conflicts of Interest

None declared.

References

1. Lee H, Lim JA, Nam HK. Effect of a digital literacy program on older adults' digital social behavior: a quasi-experimental study. Int J Environ Res Public Health 2022 Sep 29;19(19):12404. [doi: 10.3390/ijerph191912404] [Medline: 36231707]
2. 2023 population projections. Daejeon: Statistics Korea. 2023. URL: https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1BPA003&conn_path=I2 [accessed 2024-06-13]
3. Murman DL. The impact of age on cognition. Semin Hear 2015 Aug;36(3):111-121. [doi: 10.1055/s-0035-1555115] [Medline: 27516712]

XSL•FO
RenderX

4. Checton MG, Greene K. Elderly patients' heart-related conditions: disclosing health information differs by target. Psychol Health Med 2015;20(5):594-604. [doi: 10.1080/13548506.2014.986141] [Medline: 25434731]
5. Chaudhuri S, Le T, White C, Thompson H, Demiris G. Examining health information-seeking behaviors of older adults. Comput Inform Nurs 2013 Nov;31(11):547-553. [doi: 10.1097/01.NCN.0000432131.92020.42] [Medline: 23974574]
6. An S, Lim Y, Chung S. Psychological wellbeing through health information behavior: effects of the elderly's digitigal literacy. Korea Gerontol Soc 2020 Oct 1;40(5):833-854. [doi: 10.31888/JKGS.2020.40.5.833]
7. Global strategy on digital health 2020–2025. : World Health Organization; 2021 URL: https://www.who.int/docs/default-source/documents/gs4dhdaa2a9f352b0445bafbc79ca799dce4d.pdf [accessed 2025-05-30]
8. Arias López MDP, Ong BA, Borrat Frigola X, et al. Digital literacy as a new determinant of health: a scoping review. PLOS Digit Health 2023 Oct;2(10):e0000279. [doi: 10.1371/journal.pdig.0000279] [Medline: 37824584]
9. Schäffer B. The digital literacy of seniors. Research in Comparative and International Education 2007 Mar;2(1):29-42. [doi: 10.2304/rcie.2007.2.1.29]
10. Gilster P. Digital Literacy: Wiley Computer Publishing; 1997.
11. Leahy D, Dolan D. Digital literacy: a vital competence for 2010. Presented at: IFIP International Conference on Key Competencies in the Knowledge Society; Sep 20-23, 2010; Brisbane, Australia. [doi: 10.1007/978-3-642-15378-5_21]
12. Park H, Jin M, Park J, Lim K. The analysis of national and international curriculum for digital literacy education. J Educ Cult 2021;27(5):75-101. [doi: 10.24159/joec.2021.27.5.75]
13. Cockerham WC, Sharp K, Wilcox JA. Aging and perceived health status. J Gerontol 1983 May;38(3):349-355. [doi: 10.1093/geronj/38.3.349] [Medline: 6841932]
14. McMullan RD, Berle D, Arnáez S, Starcevic V. The relationships between health anxiety, online health information seeking, and cyberchondria: systematic review and meta-analysis. J Affect Disord 2019 Feb 15;245:270-278. [doi: 10.1016/j.jad.2018.11.037] [Medline: 30419526]
15. Zhou Y, Bai Y, Wang J. The impact of internet use on health among older adults in China: a nationally representative study. BMC Public Health 2024;24(1):1065. [doi: 10.1186/s12889-024-18269-4]
16. Lee JY. A study on the use of information and social computing service by the elderly. Journal of the Korean Society for information Management 2012 Mar 30;29(1):375-393. [doi: 10.3743/KOSIM.2012.29.1.375]
17. Miller LMS, Bell RA. Online health information seeking: the influence of age, information trustworthiness, and search challenges. J Aging Health 2012 Apr;24(3):525-541. [doi: 10.1177/0898264311428167] [Medline: 22187092]
18. The report on the digital divide [Article in Korean]. : Ministry of Science and ICT, National Information Society Agency; 2022 URL: https://www.nia.or.kr/site/nia_kor/ex/bbs/View.do?cbIdx=81623&bcIdx=25353&parentSeq=25353 [accessed 2025-06-04]
19. Ahn JI, Seo YK. An analysis of sub-factors of digital media literacy divide: focusing on the effects of generation and economic status. Journal of Digital Convergence 2014 Feb 28;12(2):69-78. [doi: 10.14400/JDC.2014.12.2.69]
20. Choi HI, Song I. The mediating effect of self-efficacy between the elderly's digital information literacy and life satisfaction. J Korea Acad Ind Coop Soc 2020;21(6):246-255. [doi: 10.5762/KAIS.2020.21.6.246]
21. Kim M, Kim Y, Um S. A study about the types of digital information utilization and life satisfaction of middle-aged and the older people. J Korean Assoc Reg Inf Soc 2020;21:51-74. [doi: 10.22896/karis.2020.23.1.003]
22. Park S, Kim B. Predictors of internet use among older adults with diabetes in South Korea: survey study. JMIR Med Inform 2020 Dec 23;8(12):e19061. [doi: 10.2196/19061] [Medline: 33277232]
23. Machado CLF, Pinto RS, Brusco CM, Cadore EL, Radaelli R. COVID-19 pandemic is an urgent time for older people to practice resistance exercise at home. Exp Gerontol 2020 Nov;141:111101. [doi: 10.1016/j.exger.2020.111101] [Medline: 33049354]
24. Roschel H, Artioli GG, Gualano B. Risk of increased physical inactivity during COVID-19 outbreak in older people: a call for actions. J Am Geriatr Soc 2020 Jun;68(6):1126-1128. [doi: 10.1111/jgs.16550] [Medline: 32392620]
25. Kim J, Kim Y, Ha J. Changes in daily life during the COVID-19 pandemic among South Korean older adults with chronic diseases: a qualitative study. Int J Environ Res Public Health 2021 Jun 24;18(13):6781. [doi: 10.3390/ijerph18136781] [Medline: 34202534]
26. Kim S, Choi H. Still hungry for information: information seeking behavior of senior citizens in South Korea. Presented at: iConference 2014 Proceedings; Mar 4-7, 2014; Berlin, Germany p. 889-894. [doi: 10.9776/14300]
27. Levy H, Janke AT, Langa KM. Health literacy and the digital divide among older Americans. J Gen Intern Med 2015 Mar;30(3):284-289. [doi: 10.1007/s11606-014-3069-5] [Medline: 25387437]
28. Mubarak F, Suomi R. Elderly forgotten? Digital exclusion in the information age and the rising grey digital divide. INQUIRY 2022;59:469580221096272. [doi: 10.1177/00469580221096272] [Medline: 35471138]
29. Lee YK, Kim S, Hwang N, Lim J, Joo B, Namgung E, et al. 2020 National Survey of Older Koreans. : Korea Institute for Health and Social Affairs; 2020.
30. Choi Y. Is subjective health reliable as a proxy variable for true health? A comparison of self-rated health and self-assessed change in health among middle-aged and older South Koreans. Health and Social Welfare Review 2016 Dec;36(4):431-459. [doi: 10.15709/hswr.2016.36.4.431]

XSL•FO
RenderX

31. Benyamini Y, Idler EL, Leventhal H, Leventhal EA. Positive affect and function as influences on self-assessments of health: expanding our view beyond illness and disability. J Gerontol B Psychol Sci Soc Sci 2000 Mar;55(2):107-116. [doi: 10.1093/geronb/55.2.p107] [Medline: 10794189]
32. Heller DA, Ahern FM, Pringle KE, Brown TV. Among older adults, the responsiveness of self-rated health to changes in Charlson comorbidity was moderated by age and baseline comorbidity. J Clin Epidemiol 2009 Feb;62(2):177-187. [doi: 10.1016/j.jclinepi.2008.05.009] [Medline: 18722090]
33. Peel NM, McClure RJ, Bartlett HP. Behavioral determinants of healthy aging. Am J Prev Med 2005 Apr;28(3):298-304. [doi: 10.1016/j.amepre.2004.12.002] [Medline: 15766620]
34. Cherid C, Baghdadli A, Wall M, et al. Current level of technology use, health and eHealth literacy in older Canadians with a recent fracture-a survey in orthopedic clinics. Osteoporos Int 2020 Jul;31(7):1333-1340. [doi: 10.1007/s00198-020-05359-3] [Medline: 32112118]
35. Tennant B, Stellefson M, Dodd V, et al. eHealth literacy and Web 2.0 health information seeking behaviors among baby boomers and older adults. J Med Internet Res 2015 Mar 17;17(3):e70. [doi: 10.2196/jmir.3992] [Medline: 25783036]
36. Papp-Zipernovszky O, Horváth MD, Schulz PJ, Csabai M. Generation gaps in digital health literacy and their impact on health information seeking behavior and health empowerment in Hungary. Front Public Health 2021;9:635943. [doi: 10.3389/fpubh.2021.635943] [Medline: 34055714]
37. Meischke H, Eisenberg M, Rowe S, Cagle A. Do older adults use the Internet for information on heart attacks? Results from a survey of seniors in King County, Washington. Heart Lung 2005;34(1):3-12. [doi: 10.1016/j.hrtlng.2004.06.006] [Medline: 15647729]
38. Irvine AB, Gelatt VA, Seeley JR, Macfarlane P, Gau JM. Web-based Intervention to promote physical activity by sedentary older adults: randomized controlled trial. J Med Internet Res 2013;15(2):e19. [doi: 10.2196/jmir.2158]
39. Guo E, Li J, Luo L, Gao Y, Wang Z. The effect and mechanism of internet use on the physical health of the older people-empirical analysis based on CFPS. Front Public Health 2022;10:952858. [doi: 10.3389/fpubh.2022.952858] [Medline: 36339140]
40. Cui K, Zou W, Ji X, Zhang X. Does digital technology make people healthier: the impact of digital use on the lifestyle of Chinese older adults. BMC Geriatr 2024 Jan 22;24(1):85. [doi: 10.1186/s12877-023-04651-1] [Medline: 38254001]
41. Martínez-Alcalá CI, Rosales-Lagarde A, Pérez-Pérez YM, Lopez-Noguerola JS, Bautista-Díaz ML, Agis-Juarez RA. The effects of Covid-19 on the digital literacy of the elderly: norms for digital inclusion. Front Educ 2021;6. [doi: 10.3389/feduc.2021.716025]
42. Prieto-Flores ME, Moreno-Jiménez A, Fernandez-Mayoralas G, Rojo-Perez F, Forjaz MJ. The relative contribution of health status and quality of life domains in subjective health in old age. Soc Indic Res 2012 Mar;106(1):27-39. [doi: 10.1007/s11205-011-9791-z]
43. Song JH, Shin SJ. The effects of e-Health literacy and subjective health status on health-seeking behaviors of elderly using the internet in the community. J Digit Converg 2020;18(1):321-332. [doi: 10.14400/JDC.2020.18.1.321]
44. Lu H, Kandilov IT. Does mobile internet use affect the subjective well-being of older Chinese adults? An instrumental variable quantile analysis. J Happiness Stud 2021 Oct;22(7):3137-3156. [doi: 10.1007/s10902-021-00365-6]
45. Yang Y, Zeng D, Yang F. Internet use and subjective well-being of the elderly: an analysis of the mediating effect based on social capital. Int J Environ Res Public Health 2022 Sep 24;19(19):12087. [doi: 10.3390/ijerph191912087] [Medline: 36231387]
46. Lee K, Hoti K, Hughes JD, Emmerton L. Dr Google and the consumer: a qualitative study exploring the navigational needs and online health information-seeking behaviors of consumers with chronic health conditions. J Med Internet Res 2014 Dec 2;16(12):e262. [doi: 10.2196/jmir.3706] [Medline: 25470306]
47. Cutrona SL, Roblin DW, Wagner JL, et al. Adult willingness to use email and social media for peer-to-peer cancer screening communication: quantitative interview study. JMIR Res Protoc 2013 Nov 28;2(2):e52. [doi: 10.2196/resprot.2886] [Medline: 24287495]
48. Yaagoob E, Hunter S, Chan S. The effectiveness of social media intervention in people with diabetes: an integrative review. J Clin Nurs 2023 Jun;32(11-12):2419-2432. [doi: 10.1111/jocn.16354] [Medline: 35545822]
49. Kim JN, Lee S. Communication and cybercoping: coping with chronic illness through communicative action in online support networks. J Health Commun 2014;19(7):775-794. [doi: 10.1080/10810730.2013.864724] [Medline: 24559492]
50. Min Y. The digital divide among internet users: an analysis of digital access, literacy, and participation. J Commun Res 2011 Feb;48(1):150-187. [doi: 10.22174/jcr.2011.48.1.150]
51. Kim S, Love F, Quistberg DA, Shea JA. Association of health literacy with self-management behavior in patients with diabetes. Diabetes Care 2004 Dec;27(12):2980-2982. [doi: 10.2337/diacare.27.12.2980] [Medline: 15562219]
52. Hall AK, Bernhardt JM, Dodd V, Vollrath MW. The digital health divide: evaluating online health information access and use among older adults. Health Educ Behav 2015 Apr;42(2):202-209. [doi: 10.1177/1090198114547815] [Medline: 25156311]
53. Koh EK, Koh CK. Caring for older adults: the parables in Confucian texts. Nurs Sci Q 2008 Oct;21(4):365-368. [doi: 10.1177/0894318408324320] [Medline: 18953016]

XSL•FO
RenderX

54. Palley HA, Usui C. Social policies for the elderly in the Republic of Korea and Japan: a comparative perspective*. Soc Policy Adm 1995 Sep;29(3):241-257. [doi: 10.1111/j.1467-9515.1995.tb00467.x]
55. Koo S, Cho J, Ahn E, et al. Use of health information among older adults: trust and source of health information by gender and region. Health and Social Welfare Review 2016 Jun;36(2):454-479. [doi: 10.15709/hswr.2016.36.2.454]
56. Jung WS, Kang HG, Han S, Kim E. Recognition and use of health information for preliminary elderly and elderly people. Journal of Digital Convergence 2021;19(6):419-427. [doi: 10.14400/JDC.2021.19.6.419]
57. Chang SJ, Lee KJ, Kim IS, Lee WH. Older Korean people's desire to participate in health care decision making. Nurs Ethics 2008 Jan;15(1):73-86. [doi: 10.1177/0969733007083936] [Medline: 18096583]
58. Kim DJ, Lee YS, Jeon ER, Kim KJ. Present and future of AI-IoT-based healthcare services for senior citizens in local communities: a review of a South Korean Government digital healthcare initiatives. Healthcare (Basel) 2024;12(2):281. [doi: 10.3390/healthcare12020281]
59. Jang Y, Jung J, Park NS, Kim MT, Chung S. COVID-19 health information sources and their associations with preventive behaviors: a typological study with older residents in Seoul, South Korea. PLoS One 2023;18(8):e0290824. [doi: 10.1371/journal.pone.0290824] [Medline: 37651388]
60. Weber W, Reinhardt A, Rossmann C. Lifestyle segmentation to explain the online health information-seeking behavior of older adults: representative telephone survey. J Med Internet Res 2020 Jun 12;22(6):e15099. [doi: 10.2196/15099] [Medline: 32530433]

Abbreviations

aOR: adjusted odds ratio
IRB: Institutional Review Board
ITD: IT device
KIHASA: Korea Institute for Health and Social Affairs
NSOK: National Survey of the Older Koreans

Edited by J Wang; submitted 01.08.24; peer-reviewed by J Kim, M Tomita; revised version received 02.03.25; accepted 11.04.25; published 05.06.25.

Please cite as:
Lee SY, Kim Y, Kim B, Lee SG, Jang SY, Kim TH
Digital Literacy and Its Association With Subjective Health Status and Healthy Lifestyle Behaviors Among Korean Older Adults: Cross-Sectional Study
JMIR Aging 2025;8:e64974
URL: https://aging.jmir.org/2025/1/e64974
doi:10.2196/64974

© Soon Young Lee, Yejin Kim, Bomgyeol Kim, Sang Gyu Lee, Suk-Yong Jang, Tae Hyun Kim. Originally published in JMIR Aging (https://aging.jmir.org), 5.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Selection, Optimization, and Compensation Strategies Used by Older Adults to Live Well With Technology: Qualitative Study

Wei Qi Koh[1], PhD; Kristiana Ludlow[2,3], PhD; Jacki Liddle[1,4], PhD; Nancy A Pachana[5], PhD

[1]School of Health and Rehabilitation Sciences, The University of Queensland, 84A Services Road, St Lucia, Brisbane, Australia

[2]Centre for Health Services Research, Faculty of Health, Medicine and Behavioural Sciences, The University of Queensland, Brisbane, Australia

[3]Australian Frailty Network, The University of Queensland, Brisbane, Australia

[4]Occupational Therapy Department, Princess Alexandra Hospital, Brisbane, Australia

[5]School of Psychology, The University of Queensland, Brisbane, Australia

Corresponding Author:
Wei Qi Koh, PhD
School of Health and Rehabilitation Sciences, The University of Queensland, 84A Services Road, St Lucia, Brisbane, Australia

Abstract

Background: With rapid digitalization, technologies are increasingly integrated as part of our everyday lives and are becoming increasingly essential for individuals to participate in society. Technology presents opportunities to support healthy aging. Examples include digital health monitoring and opportunities to maintain social connectedness through online platforms. However, the processes in which older adults select and integrate technologies into their daily lives have not been well examined.

Objectives: This study uses the Selection, Optimization, and Compensation (SOC) model to understand how older adults integrate technology into their everyday lives to live well. The two key research questions are as follows: (1) How do older adults describe their technology use and their choices, analyzed with respect to SOC processes? (2) How do older adults perceive that technology is a part of living well?

Methods: A descriptive qualitative study was conducted. Purposive sampling was used to recruit older adults who were aged 55 years and older, were living in the community, spoke and understood English, and resided in Australia. Five in-person focus groups comprising 20 older adults were conducted. Data were analyzed using hybrid inductive and deductive reflexive thematic analyses, based on the SOC model.

Results: All participants resided in Brisbane, Queensland. Older adults selected technology intentionally to enhance different aspects of their daily lives. Perceived "involuntary" selection of technology could lead to feelings of resentment or frustration. Optimization strategies included self-monitoring, integrating technology into daily routines, digital literacy and proficiency, and problem-solving skills. Compensatory strategies included choosing alternative technology that suited participants' abilities or skills and seeking support through informal or formal avenues.

Conclusions: These findings provide important considerations for technology developers to design technology in collaboration with older adults to ensure that they align with needs and preferences. Digital literacy is not sufficient to empower older adults to use technology; when empowering older adults to use technology, service providers should also consider facilitating other intrinsic and extrinsic resources and skills.

(JMIR Aging 2025;8:e75019) doi:10.2196/75019

KEYWORDS

technology; healthy aging; selective optimization and compensation; aging well; older adults

Introduction

Background

According to the World Health Organization, healthy aging entails creating the environments and opportunities that enable people to be and do what they do throughout their lives [1]. As populations are aging globally, healthy aging is a priority in many countries, and interventions and programs have been developed to support older adults to avoid negative impacts associated with disability or to maintain or improve their functional ability. Interventions have also focused on supporting older people to sustain or enhance participation in social engagement and other meaningful activities [2]. Examples include education, problem-solving and goal setting, leisure exploration [2], and activities to improve physical capacities [3]. Synthesized evidence suggests that physical activity programs can improve healthy aging by reducing physical limitations and improving psychosocial outcomes [4]. Similarly, non–exercise-based interventions, such as leisure-based interventions or programs to increase older adults'

XSL•FO
RenderX

self-awareness or self-acceptance, have led to improvements in older adults' quality of life [2].

With rapid digitalization, technologies are increasingly integrated as part of our everyday lives and are becoming increasingly essential for individuals to participate in society. For instance, payments for services and goods have become digitized through mainstream telephone-based services, app-based services, and automated teller machines (ATMs) [5]. Correspondingly, technology presents new opportunities to digitize interventions aimed at supporting healthy aging [6]. Examples of technology that have been developed and leveraged to support healthy aging include web platforms, mobile health apps, and wearable technology, which have been designed to support older adults in self-monitoring and decision-making, measuring physiological, cognitive, or physical metrics, or delivering interventions [7]. While these technologies are intended to simplify and enhance older adults' daily lives, they can also make everyday lives more complex [5]. To understand how technologies can be best used to support healthy aging, it is necessary to understand how they have been assimilated in older adults' everyday lives. Several studies have been conducted to understand older adults' experiences with technologies. For instance, Reyes and colleagues conducted a study in Australia to understand older adults' choices of, and motivations to use technologies for health-related self-management [8]. They also sought to understand barriers and facilitators to technology use. The researchers found that older people leveraged technology for health monitoring and to remain active and connected. Facilitators include social and organizational influences, accessibility, and cost, whereas some barriers include perceived information overload, ease of use, and concerns about privacy and accuracy. These findings were echoed in other international studies conducted in Canada [9] and the United Kingdom [10]. While existing research has provided valuable collective insights into how older adults experience technology and barriers and facilitators to their use, the processes in which older adults select and integrate technologies into their daily lives have not been well examined.

Selective Optimization With Compensation

The Selective Optimization with Compensation (SOC) model [11] is a meta-theoretical model focusing on the "how" of aging well; it describes processes that older adults undertake to reach their goals, particularly considering that their resources may become more limited over time. According to the SOC, when personal and environmental resources become limited, individuals make decisions to allocate these resources to pursue and to achieve their goals. This involves (1) selection, (2) optimization, and (3) compensation. Selection entails setting or prioritizing goals through loss-based selection or elective selection. For instance, older adults might choose to cease participation in certain activities that might become more effortful and instead select other activities deemed to be more meaningful. Optimization entails acquiring and refining resources and skills. Examples include learning new skills or knowledge needed to engage in chosen activities. Finally, compensation involves acquiring or substituting resources to achieve their goals. For example, older adults might leverage technology or support services to help them participate in activities. According to Freund and Baltes, the orchestrated use of these 3 action strategies effectively manages the loss of resources across individuals' lifespans and conserves resources [12]. Given this, the SOC model could provide important theoretical guidance on bridging an established theory on successful aging with digital engagement. There is limited research on the application of the SOC model to the understanding of how technology is leveraged as part of, or integrated into, older adults' process of selection, optimization, and compensation. This study aims to understand how technology is integrated into older adults' everyday lives to live well using the SOC model. Specifically, the key research questions are as follows:

1. How do older adults describe their technology use and their choices, analyzed with respect to SOC processes?
2. How do older adults perceive that technology is a part of living well?

Methods

A qualitative descriptive approach [13,14] underpinned our qualitative inquiry, where we stayed close to the "surface of the data and events" to gather rich descriptions of the views of participants and describe the phenomena from the viewpoints of participants.

Sampling, Recruitment, and Context

Purposive sampling was used to identify and recruit participants who met the following criteria: (1) aged 55 years and older, (2) living independently in the community, (3) able to speak and understand English, and (4) residing in Australia. Different avenues were contacted to advertise this study to prospective participants. These included national and local senior networks, age-friendly university networks, and the University of Queensland's (UQ) Healthy Living, an interprofessional clinic promoting healthy aging and well-being. Individuals who were interested in participating in this study were invited to share their preferred contact information with the research team. The lead researcher (WQK) contacted potential participants by email and/or a phone call to provide more information about the study and to arrange a focus group session at their preferred time and date. Online and in-person focus group options were planned, and recruitment was not limited to specific states in Australia. However, all participants who volunteered to join this study lived in Brisbane, and all focus group sessions were conducted in-person. No prospective participants declined participation. Brisbane is the capital city in the state of Queensland in Australia, with older adults constituting approximately 17% of the population. The local government, Brisbane City Council (BCC), has several initiatives to support seniors in Brisbane, and these include centers with organized activities, exercise programs, and transport concessions [15]. Brisbane has good digital infrastructure, and efforts have been made to improve digital accessibility and close the digital divide [16].

Data Collection

Data collection took place between April and June 2024. Five face-to-face focus groups, comprising 20 participants in total, were conducted at a community health service for older adults.

XSL•FO
RenderX

Each focus group comprised 3–5 participants and lasted approximately an hour. The topic guide (Multimedia Appendix 1) was developed and discussed with a consumer advisor. Each focus group was facilitated by a moderator (WQK). At the outset of each session, the moderator reiterated the research purpose, explained the session structure, and answered any questions. All participants were offered a AU $20 (US $14) gift voucher as a token of appreciation for their time and participation. The average duration of focus groups was 67 minutes. All sessions were audio-recorded and transcribed verbatim.

Data Analysis

Data were imported into NVivo14 for data management. Data were analyzed using reflexive thematic analysis [17] both inductively and deductively based on SOC to identify strategies used by participants. First, WQK immersed herself in the data by reading and re-reading the transcripts and listening to the audio recording. During this process, data were coded descriptively and inductively through open coding. Next, codes were organized based on the SOC model and grouped into themes and subthemes. This was an iterative and reflective process, where the researcher moved back and forth between preliminary themes. Two meetings were arranged with the research team to discuss and explore the wider team's perspectives (JL, KL, and NP) on the data and codes. For example, after organizing codes based on the SOC model, team discussions provided a sounding board as to whether data fitted into the preliminary themes. From discussions, subthemes like "passive selection of technology" were created to avoid "force-fitting" data into the "selection" component of the existing SOC model, acknowledging that while participants "select" technology, this could be based solely on circumstantial reasons. This process provided valuable leverage of multidisciplinary expertise to achieve a richer and more holistic interpretation of the data and the interrelationships between themes.

Researcher Reflexivity

Considering the reflexive nature of our analytical approach, it is imperative to reflect on how the authors' background and experiences shaped data analysis and interpretation [17,18]. The interviewer (WQK) is a lecturer in occupational therapy and an experienced qualitative interviewer, with more than 7 years of experience working clinically with older adults. This allowed WQK to understand and embrace older adults' perspectives and drive to remain independent. As an international researcher from Singapore, WQK could miss out on participants' nuances relating to local contexts in Australia. Field notes were taken and discussed with the research team (JL, KL, and NP) who contributed their professional and research expertise during data analysis to situate our findings in both local and international contexts. They also contributed their expertise in situating findings with theoretical underpinning.

Trustworthiness

This study followed a methodologically rigorous process, based on Lincoln and Guba's criteria for trustworthiness [19]. First, to ensure transparency and confirmability, data management and analysis of the reflexive thematic analysis process took place on the NVivo software. Next, findings were discussed with the research team during the analysis process to obtain a more holistic, multidisciplinary perspective of the data collected, providing credibility. Another strength is in its use of an established theoretical framework (SOC) to guide data analysis. In addition, the criterion of transferability was aided through having clear descriptions of participants' characteristics and direct quotations. The Standards for Reporting Qualitative Research (SRQR) checklist was to report the conduct of this study (Checklist 1) [20].

Ethical Considerations

Ethical approval was obtained from the University of Queensland (Ref:2023/HE002389). Written informed consent was obtained from all participants, who were also provided with the opportunity to clarify questions prior to the commencement of data collection. Participants' data were stored securely and confidentially in alignment with approved institutional ethical guidelines. Each participant was provided a AU $20 (US $14) gift card as a token of appreciation for their time for participation. To ensure anonymity, participants were given pseudonyms for data analysis and reporting.

Results

Demographic Characteristics

Participants' demographic information is summarized in Table 1. Participants' mean age was 75.4 years, and there was a similar proportion of men and women. More than half of the sample (n=12, 60%) had a Bachelor's degree or higher, and most participants were unemployed or retired. All participants (n=20, 100%) were regular users of smartphones, computers, or laptops. Most described proficiency and daily use of these technologies. Approximately two out of three have had experiences using pedometers or smartwatches; one third are regular users (n=6, 30%) and others have seen or used them previously (n=7, 35%). None of the participants had used virtual reality technologies. Strategies relating to technology use will be described using the SOC model with 3 overarching themes: (1) elective selection of technology— which are further divided into 2 subthemes: "elective, active selection" and "passive selection"; (2) optimization strategies, and (3) compensation strategies.

XSL•FO
RenderX

Table . Participants' characteristics.

Characteristics	Participants
Gender, n (%)	
Men	11 (55)
Women	9 (45)
Mean (SD) age, years (range)	75.4 (68-82)
Educational attainment, n (%)	
Bachelor's degree	12 (60)
Diploma/advanced diploma	7 (35)
Certificate III/IV (vocational qualification)	1 (5)
Employment status, n (%)	
Part-time or casual employment	2 (10)
Retired	18 (90)

Theme 1: Elective Selection of Technology

This theme described ways in which participants intentionally and unintentionally chose to use technology to facilitate different aspects of participation in daily living and well-being. The subthemes include (1) elective, "active" selection of technology and (2) passive selection of technology, where participants chose technology but perceived the choice as being "forced" or unavoidable.

Subtheme 1—Elective, "Active" Selection of Technology

All participants described regular use of their mobile phones, laptops, or personal computers, and a few used other technologies, such as iPads, tablets, and smartwatches. Participants conveyed that technologies and applications were purposively selected to help with daily life organization. For example, some described keeping digital records of their accounts or finances, organizing photographs or photobooks electronically, or moving from using physical diaries or calendars to digital versions for more effective management of schedules and appointments.

> *I never seem to have the calendar ... besides, I share it with my husband, so he can see what it's like. When you get older, as so many… medical, specialists… we're trying to keep track of it all.* [Olivia, 74]

Some participants described using technology as a means for leisure and social participation. For instance, a few shared their experiences engaging in mobile games such as quizzes or crosswords, which were described as fun or as means to pass time. Technology was said to provide opportunities for mental stimulation and lifelong learning through information seeking and learning through the internet. Examples included tracing family history and increasing knowledge in various topics of interest, such as medications, artificial intelligence (AI), or other educational topics. Participants expressed that technology provided opportunities to feel connected with the world:

> *Yeah. I'm a great believer in current affairs. If I have a few moments to scroll, I'm on Twitter … the commentary on things that are going on is amazing. You just feel so in touch with the world, you know?.* [Liam, 69]

A few participants described using activity tracking applications to monitor daily activities, such as physical activity levels (eg, step count) and physiological features (eg, heart rate and sleep). One participant expressed that this served as supplementary information that was provided to his general practitioner (GP) during medical consults:

> *It's handy to have, particularly in a medical interaction, like I go to my GP, and I've been recently regularly, two times a month perhaps, for a variety of different reasons, and I like to be able to say that this is what I'm doing, this is how I'm going, and she can monitor my overall activity rates and things like that … I find that particularly useful.* [John, 72]

In addition, participants reported using technology to stay socially connected with family and friends, particularly those living overseas; to reconnect with old friends; and to connect with new friends and likeminded people:

> *I use the cliche apps like Facebook to keep up with friends back in Hungary, family in Australia … Fair few discussion groups on various things like cooking.* [Liam, 69]

> *… maintain old friendships. Yeah. And renew old friendships … I didn't know that [name of friend] would age so well, oh that's great, I'm going to the hairdressers (laughs).* [Hazel, 69]

Subtheme 2—Passive Selection of Technology

Many participants described technology uptake as passive or "forced," where technology selection was seen as essential or unavoidable, regardless of their willingness to use them. This was described as resulting from rapid digitization of services such as banking and electronic commerce. Participants conveyed that this could affect or threaten older adults' participation in meaningful activities and daily routines:

> *We're forced into technology. The thing is you haven't got any choice anymore because essentially when I had to get the blue card for PCYC [Police Citizens*

XSL•FO
RenderX

> *Youth Club – a network of non-profit organisations], the only way you could do it was online… the other thing is that I just got my license renewed and I thought, well, I'll see if I can get this new digital license. The only way you can get your digital license is you've got to download an app.* [Jack, 76]

While most participants identified themselves as being relatively proficient in navigating technology, feelings of being forced into using technology in their day-to-day lives were expressed and described as frustrating and upsetting. They explained that this affected how they engage in activities:

> *I'm annoyed and angry with technology… for instance, now… you're running a book sale, but you've got to register it, and if you want to register it, you've got to have an account, and you've got to have a code name, a code word, and you've got to nominate a time. I've got friends up at [regional book club], people who love books, but they refuse to sign up because they can't nominate a time to get there from [regional area] by public transport ….* [Samuel, 80]

Theme 2: Optimization Strategies

This theme described strategies that were identified by participants to optimize the use of their selected technology to ensure meaningful life participation. It also described how participants optimized their life using technology. These included one's ability to self-monitor, to successfully adjust to and/or integrate technology to fit into daily lives, and to leverage one's digital literacy, positive attitude, and problem-solving skills.

Many described the importance of monitoring the relevance of technology, keeping in mind their intended use of technology for well-being. In other words, they described active deliberation about when technology can be, or has been, beneficial in helping them live well. Some participants expressed the importance of using technology purposefully to avoid time wastage or harmful effects:

> *I choose what I like to see, who I want to talk, and also connecting with family. WhatsApp, I call my sister once a month or once a week. We talk about things. It's something that I really like … not the silly thing of scrolling and scrolling and scrolling [social media], and spending hours doing … that young people do? No. I decide what to do.* [Evelyn, 76]

Participants explained that ongoing deliberation about the utility of technology encouraged them to continue to engage with their chosen technological mediums in some instances and disengage in other instances. For example, while some described leveraging mobile applications to track their daily activities (eg, step count) for self-monitoring, they reported eventually disengaging from these apps after gaining a sense of their activity levels to benefit from the activity without the drawback of the technology:

> *I found that helpful in the beginning, but I just don't use it anymore. I have an idea of how long my walks are … I'd rather enjoy the walk then look … if you go out and, you know, bush or national parks or something, you want to look at your surroundings. You don't want to be embedded in this thing [mobile phone].* [Amelia, 80]

Another optimization strategy was the investment in time and resources to keep updated with technology and be aware of potential risks, such as misinformation or financial scams. Some participants identified digital proficiency, previous professional experience, or personal experiences as assets that enable them to problem-solve through challenges. For example, to keep updated with technology updates, one participant actively acquires knowledge through searching for YouTube videos. A few participants also described the importance of having a positive attitude in seeking resources or in equipping oneself with skills to overcome current or prospective challenges, such as risk management or active problem-solving:

> *We'll have problems, but we need the confidence to think, 'Oh yeah, look, a little bit of sorting out.' I mean, you know, we'll be able to understand it, ask. Whereas some people are just so confused, it's too much for them. But I think if you have a healthy attitude towards aging, you think, 'Well look, I'll make sense eventually, I'll seek out help, I'll look at the YouTube videos, I'll work it out.* [Luke, 79]

Theme 3: Navigating Technology Use and Compensation Strategies

This theme described challenges that participants encountered during technology use and strategies that were discussed as ways to adapt to challenges or limitations; these included choosing alternative technologies that are better aligned with their knowledge and abilities and seeking formal and/or informal support. These strategies enabled them to continue engaging with technologies in ways that support their lifestyles and routines without perceived detractions.

Some participants described functional changes or disabilities that limited their abilities to use technology effectively and safely. One participant provided the example of visual fatigue when looking at smaller fonts, particularly through her mobile phone. Correspondingly, she described choosing an alternative technology that was more suited to her physical abilities (eg, laptop with a larger font size). Some participants also described experiencing challenges or feeling stuck when they experienced technical issues or when they were not able to navigate technology use:

> *I was doing a post to Facebook a couple of days ago, and I wanted to reorder the photographs that I was taking, and they were in an event, and I wanted them to be sequential, and I thought, 'Damn it, I don't know how to do it. I want to get this post out; I don't have time now.* [Luke, 79]

A few participants described seeking informal support from their friends or family members, particularly the younger generation whom they described as being more technology-savvy:

> *I must admit I call him [husband], as when I get stuck, he's very good with technology. And I just call him*

XSL•FO
RenderX

> *... if he fails, we call [a service] or whatever it is* [Emily, 71]
>
> *Or a 15-year-old grandchild* [Hazel, 69]

One participant expressed concerns that relying on family members to assist with technology could lead to burden:

> *My daughter occasionally drops in and tries to sort out things and explain things, but I don't want to burden her anymore.* [Samuel, 80]

A few participants described seeking external support from other sources, such as a local nonprofit organization that assists people who are aged older than 50 years to become competent in computer and internet use:

> *I am going to shortly join a group that's like a technical men's shed [non-profit organization to promote socialization and shared activities among men] … I'm hoping my technological problems like reordering the photographs on Facebook and that sort of thing will be handy … to figure out these little perks that I don't quite know at the moment yet.* [Luke, 79]

Discussion

Principal Findings

This research aimed to understand ways in which older adults perceive they use technology to live well. Specifically, this work utilized the SOC model to understand the "hows"; how older people select or opt out of using technology (selection), how they optimize their resources and skills to use technology to live well (optimization), and how they compensate for a lack of resources and skills (compensation) to maintain meaningful participation in the face of challenges. Corresponding with the "elective selection" ontology of the SOC model, findings revealed that older adults selected technology intentionally to enhance different aspects of their daily lives. Perceived "involuntary" selection of technology could lead to feelings of resentment or frustration. In some instances, this led to participation in activities or routines that did not involve technologies. Corresponding with the "optimization" component of the SOC model, optimization strategies where older adults used technology to improve daily living (and vice versa) included self-monitoring, technology integration into their daily routines, digital literacy and proficiency, and problem-solving skills. Corresponding with the "compensation" aspect of the SOC model, participants' compensatory strategies included choosing alternative technology that suited participants' abilities or skills and seeking support through informal or formal avenues. Our participant cohort was living in Brisbane, Queensland, with good digital infrastructure across the city; the cohort identified themselves as being relatively digitally savvy and well-educated. Given this, findings may be more resonant with this population.

Technology Selection—The "Whats" and "Whys"

Participants in our study reported voluntarily and purposefully selecting technologies that facilitated improvements in various aspects of their daily lives, such as improving the effectiveness and efficiency of self-management (eg, activities within daily routines, health management) and participating in leisure, social, cognitive, and physical activities. These are congruent with the literature on how older adults perceive successful aging [21]. The most frequently used technologies were smartphones, laptops, and/or desktop computers, and the most commonly discussed technological applications were social media and messaging apps, internet searches, digital organizers, and calendars. Purposeful and meaningful selection of technology appeared to encourage uptake and continued use. Our work bolsters findings from several studies showing that the perceived benefits of technology and their meaningfulness facilitated uptake [22], whereas low interest or a lack of understanding about the purpose of technology impeded older adults' use or continued use [23-25]. A possible explanation is that purposeful, autonomous selection of technologies empowers older adults with intrinsic motivation, which could in turn encourage sustainable use of technology beyond the novelty period [26]. On the contrary, some technologies were selected nonautonomously or, in some participants' words, being "forced" upon them with circumstantial changes. These might have been further affected by the COVID-19 pandemic, which has led to a particular acceleration of service digitalization [27,28]. When technology use is perceived as involuntary, this could lead to reduced satisfaction, a sense of oppression, or frustration [29]. These findings highlight the importance of ensuring autonomy and providing older adults with a choice in selecting technologies that are seen as meaningful in improving their daily living. This is particularly important given the continual interest and development of technology to support older adults to live well.

Technology Use—The 'Hows'

Participants in this study described leveraging different cognitive skills, including metacognition and problem-solving skills, to optimize the use of their selected technology. Metacognition entails actively monitoring and evaluating information to ascertain whether one is moving toward a goal and making decisions on how to make changes to achieve the goal [30]. In our study, older adults described metacognitive processes used to ensure that they use their selected technology in ways that aligned with their personal preferences, lifestyles, and routines. Such monitoring and self-regulation skills have been found to improve health outcomes, such as preventing illness and maintaining health [31].

Older adults also described ways in which they overcame problems that were encountered, such as technical issues or working around software or hardware updates. Such problem-solving skills have been viewed as an important aspect of successful aging [32]. According to Chen and colleagues, problem-solving entails having "fluid resources," which encompasses abstraction and performing mental operations, and crystallized resources, which entail gaining resources through cumulative experience and knowledge [32]. Given that fluid resources are based on reasoning and processing abilities, they can decline with age, particularly in later adulthood [33,34]. On the contrary, crystallized resources can be maintained or increased over time up to late adulthood [34]. Our participant cohort was generally highly educated. Most reported high digital literacy skills as part of their previous professional employment.

XSL•FO
RenderX

This might have contributed to the development of crystallized resources, which allowed them to work through issues and to consider other avenues, such as YouTube tutorials, as a means to equip themselves with the skills and knowledge to overcome challenges. In addition, technology is rapidly developing alongside potential benefits and risks such as misinformation and online scams or fraud. According to National Seniors Australia, more than 600,000 scam reports were made in 2023 with older adults being most impacted [35]. Likewise, the National Council of Aging revealed that worldwide, older adults aged 60 years and older reported US$3.4 billion total fraud loss in 2023 [36]. Accordingly, problem-solving skills appear to be requisite to afford older adults the confidence and abilities to use technologies in calculated ways. This includes ensuring that the benefits of technology outweigh their risks and for technology use to be meaningfully aligned with older adults' personal well-being goals.

Some participants described a positive outlook and a "can do" mentality as being an important strategy for technology use. Positive attitudes have been found to be a psychological resource in healthy aging [37,38] as studies have found that older people with a more positive attitude were more likely to have better self-reported health and well-being [39], participate in community and physical activities [40], and activate health-promoting behaviors [41]. In our study, such behaviors include trying to find solutions to overcome potential problems encountered to continue using technology in a meaningful way. This suggests that when designing interventions to support older adults to engage in technology, considerations should also be made to inculcate positive attitudes and consider measuring attitudinal changes as an outcome measure.

Some older adults may lack the confidence or ability to use technology in ways that they perceive as meaningful. For instance, some participants in this study expressed a lack of confidence to overcome issues encountered and may lack "crystallized resources" such as digital proficiency to navigate issues. In such instances, support appears crucial, including informal support from family members and formal support from organizations. While informal support is valuable, some study participants also expressed concerns about burdening their family, particularly if they sought help more frequently than they hoped. Such self-perceived burden can affect older adults' self-concept [42]. Considering that technology is continuously evolving, it is more than likely that any users of technology, including older adults, may require support to troubleshoot potential issues. While family members might be able to bridge obstacles [43], it is important to consider sustainable ways, such as community training hubs, which older adults can receive technological support as required. While such services might be available, many participants in this study did not seem aware of such services; this resonates with findings from other studies [43,44].

Implications for Practice

There is increasing interest in using technology to support healthy aging, as seen through continued development and research in this field, alongside trends of rapid development of digital infrastructure in developed countries worldwide. Despite this, a lack of technology uptake and sustained use remains a pertinent issue [25], and efforts continue to be made to improve digital uptake particularly among older adults nationally [45] and internationally [46]. Our findings provide practical implications to support aged care organizations, centers, and care providers to support older adults to use and integrate digital technologies into their daily lives. Regarding technology selection, our study suggests that meaningful selection of technology by older adults provides them with a sense of autonomy and motivation for sustained use. Selected technologies must serve a purpose, such as meaningfully supporting aspects of their daily lives. Moving forward, aged care organizations, centers for seniors, and health care professionals working to prescribe and equip older adults to use technologies should consider empowering them with the autonomy to reflect on their needs and abilities and to choose technologies that can be personalized to their routines and preferences [47]. Forced selection can be counterproductive in supporting older adults' well-being, and it must be acknowledged that older adults may opt out of using technologies. If forced selection occurs, additional supports are likely necessary to support older adults in using technologies and to mitigate the risk of excluding participation.

In terms of optimization strategies to use technology purposefully, cognitive processes, including metacognition and problem-solving, appear necessary to empower older adults with the confidence to navigate the use of technology in ways that align with their intended purpose. This is particularly true of technologies that are dynamic, which can affect ways of using or operating them [48]. Likewise, a positive attitude toward using technology and its challenges appears to be an important, facilitatory resource. Therefore, service developers or health care professionals working to equip older adults with technological skills should also consider the cognitive aspects of using technology and empower older adults to gain skills and knowledge to use them in reasoned ways. Examples might include empowering older adults with problem-solving skills [49], building self-efficacy and confidence in managing prospective challenges, or positive thinking training [50]. Regarding compensatory strategies, being supported to use technologies and navigate challenges is important [51]. Such resources may include family support and external support services like community hubs. Given that not all older adults may have family support, it is important to ensure that external support services are readily accessible to promote digital inclusion [52] and to minimize the negative repercussions of digital exclusion, such as poorer health outcomes and social isolation [53].

In terms of leveraging the SOC model as a theoretical framework—our findings underlined its utility and success in providing an established framework to guide our analysis and interpretation of the "how" of individual older adults could integrate technologies into their daily lives. However, individuals' goals and strategies are often interrelated with contextual demands. For example, our subtheme "passive selection of technology" was largely influenced by participants' perceptions and experiences of a rapidly developing digital infrastructure that influenced their uptake of technology.

XSL•FO
RenderX

However, the SOC model provides limited consideration of contextual and societal factors that can influence the aging process. Accordingly, researchers leveraging the SOC model to understand digital engagement should take into consideration its limitations and reflect how contextual demands influence individuals' goals and strategies.

Limitations

The limitations of this study must be acknowledged. While several recruitment avenues were used to recruit participants, all participants were recruited from a healthy aging clinic and were relatively physically active. There was no option in our demographic survey to capture data of frequency of technology use in a week, and this should be acknowledged as a limitation given that this knowledge could provide more information on participants' digital proficiency. However, most identified and described themselves as being technologically savvy and relatively well-educated, with all reporting weekly use of computers/laptops and smartphones. Participants' digital proficiency could lead to response bias, and findings from this research may be more resonant with this population and may be less representative considering other demographic groups, particularly older adults who are less active and technologically savvy. Finally, this study was limited to participants who could speak English, and as such, our findings may have excluded non-English speaking older adults' perspectives and experiences. Future studies might consider further exploring understanding technology use and non-use in other population groups.

Conclusions

Framed with respect to the SOC model, this study contributes to a theoretical understanding of how older adults use technology to live well. Older adults purposefully select technologies that align with their priorities in living well, and this appears to be a facilitator for technology uptake and sustained use. Older people leverage different resources and skills to use their selected technologies in meaningful ways. These include metacognitive skills, problem-solving skills, positive attitudes, and internal and external support. These findings provide important considerations for technology developers to design technology in collaboration with older adults to ensure that they align with needs and preferences. It is also evident that digital literacy is not sufficient to empower older adults to use technology; when empowering older adults to use technology, service providers should also consider facilitating other resources and skills.

Acknowledgments

We would like to thank all participants who contributed to this study and to UQ Healthy Living for supporting our recruitment and data collection process. This research was funded by the University of Queensland's School of Health and Rehabilitation Sciences New Staff Start-Up Grant.

Data Availability

The study was not preregistered. The data generated and analyzed in this study are not publicly available to maintain participant privacy and confidentiality. However, deidentified parts of the interview transcripts may be obtained from the corresponding author upon reasonable request.

Authors' Contributions

WQK planned the study, sought ethics approval, recruited participants, conducted focus groups and analyzed the data, and wrote the paper. JL, NAP, and KL contributed to the study design, interpretation of findings, and critical revision of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Topic Guide.
[DOCX File, 16 KB - aging_v8i1e75019_app1.docx]

Checklist 1
Standards for Reporting Qualitative Research (SRQR) Checklist.
[DOCX File, 23 KB - aging_v8i1e75019_app2.docx]

References

1. World Health Organization. Healthy ageing and functional ability. 2020 Jul 14. URL: https://www.who.int/news-room/questions-and-answers/item/healthy-ageing-and-functional-ability [accessed 2025-02-15]
2. Koh WQ, Wora N, Liong NWL, Ludlow K, Pachana NA, Liddle J. Non-exercise-based interventions to support healthy aging in older adults: a systematic review and meta-analysis of randomized controlled trials. Gerontologist 2024 Dec 17;65(2):gnae156. [doi: 10.1093/geront/gnae156] [Medline: 39478349]

XSL•FO
RenderX

3. Moreno-Agostino D, Daskalopoulou C, Wu YT, et al. The impact of physical activity on healthy ageing trajectories: evidence from eight cohort studies. Int J Behav Nutr Phys Act 2020 Jul 16;17(1):92. [doi: 10.1186/s12966-020-00995-8] [Medline: 32677960]
4. Pinheiro MB, Oliveira JS, Baldwin JN, et al. Impact of physical activity programs and services for older adults: a rapid review. Int J Behav Nutr Phys Act 2022 Jul 14;19(1):87. [doi: 10.1186/s12966-022-01318-9] [Medline: 35836187]
5. Wallcook S, Malinowsky C, Charlesworth G, Ryd C, Nygård L. Everyday technology's interplay in the lives of people with dementia: A multiple case study in the rural North of England. J Rural Stud 2024 Feb;106:103203. [doi: 10.1016/j.jrurstud.2024.103203]
6. World Health Organization. Promoting healthy ageing in a digital world. 2024 Jul 14. URL: https://www.who.int/europe/news-room/events/item/2024/09/17/default-calendar/promoting-healthy-ageing-in-a-digital-world [accessed 2025-09-16]
7. Ienca M, Schneble C, Kressig RW, Wangmo T. Digital health interventions for healthy ageing: a qualitative user evaluation and ethical assessment. BMC Geriatr 2021 Jul 2;21(1):412. [doi: 10.1186/s12877-021-02338-z] [Medline: 34215209]
8. Garcia Reyes EP, Kelly R, Buchanan G, Waycott J. Understanding older adults' experiences with technologies for health self-management: interview study. JMIR Aging 2023;6:e43197. [doi: 10.2196/43197]
9. Kabacińska K, Vu K, Tam M, Edwards O, Miller WC, Robillard JM. "Functioning better is doing better": older adults' priorities for the evaluation of assistive technology. Assist Technol 2023 Jul 4;35(4):367-373. [doi: 10.1080/10400435.2022.2113180] [Medline: 35972791]
10. Vaportzis E, Clausen MG, Gow AJ. Older adults perceptions of technology and barriers to interacting with tablet computers: a focus group study. Front Psychol 2017 Oct 4;8:1687. [doi: 10.3389/fpsyg.2017.01687] [Medline: 29071004]
11. Baltes PB, Baltes MM, Eds. Psychological perspectives on successful aging: the model of selective optimization with compensation. In: Successful Aging: Perspectives from Behavioural Sciences: Cambridge University Press; 1990:1-34 URL: https://tinyurl.com/395ktst8 [accessed 2025-09-16]
12. Freund B, Baltes P. The Orchestration of Selection, Optimization and Compensation: An Action-Theoretical Conceptualization of a Theory of Developmental Regulation Control of Human Behavior, Mental Processes, and Consciousness: Essays in Honour of the 60th Birthday of August Flammer: Lawrence Erlbaum Associates Publishers; 2000:35-58 URL: https://tinyurl.com/494w2rc4 [accessed 2025-09-16]
13. Sandelowski M. Whatever happened to qualitative description. Res Nurs Health 2000;23:334-340. [doi: 10.1002/1098-240X(200008)23:43.0.CO;2-G]
14. Sandelowski M. What's in a name? Qualitative description revisited. Res Nurs Health 2010 Feb;33(1):77-84. [doi: 10.1002/nur.20362] [Medline: 20014004]
15. Brisbane city council. Seniors in Brisbane. 2025 Jul. URL: https://www.brisbane.qld.gov.au/community-support-and-safety/support-for-community/seniors-in-brisbane [accessed 2025-02-25]
16. Queensland government. Digital Economy Strategy | Strategies and initiatives corporateName=The State of Queensland; jurisdiction=Queensland. URL: https://www.qld.gov.au/about/how-government-works/strategies-and-initiatives/digital-economy-strategy [accessed 2025-02-19]
17. Braun V, Clarke V. Thematic Analysis: A Practical Guide: Sage Publishing; 2021. URL: https://uk.sagepub.com/en-gb/eur/thematic-analysis/book248481 [accessed 2025-09-16]
18. Jamieson MK, Govaart GH, Pownall M. Reflexivity in quantitative research: a rationale and beginner's guide. Social & Personality Psych 2023 Apr;17(4):e12735. [doi: 10.1111/spc3.12735]
19. Lincoln YS, Guba EG. Naturalistic Inquiry: Sage Publishing; 1985. URL: https://tinyurl.com/mtpyuasp [accessed 2025-09-16]
20. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
21. Burton E, Teater B, Chonody J, Alford S. What does it mean to successfully age?: multinational study of older adults' perceptions. Gerontologist 2024 Oct 1;64(10):gnae102. [doi: 10.1093/geront/gnae102] [Medline: 39126227]
22. Liddle J, Stuart A, Worthy P, et al. "Building the threads of connection that we already have": the nature of connections via technology for older people. Clin Gerontol 2021 Aug 8;44(4):406-417. [doi: 10.1080/07317115.2020.1852638]
23. Chung J, Brakey HR, Reeder B, Myers O, Demiris G. Community-dwelling older adults' acceptance of smartwatches for health and location tracking. Int J Older People Nurs 2023 Jan;18(1):e12490. [doi: 10.1111/opn.12490] [Medline: 35818900]
24. Dequanter S, Fobelets M, Steenhout I, et al. Determinants of technology adoption and continued use among cognitively impaired older adults: a qualitative study. BMC Geriatr 2022 Apr 28;22(1):376. [doi: 10.1186/s12877-022-03048-w] [Medline: 35484488]
25. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556. [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
26. Meng L, Ma Q. Live as we choose: the role of autonomy support in facilitating intrinsic motivation. Int J Psychophysiol 2015 Dec;98(3 Pt 1):441-447. [doi: 10.1016/j.ijpsycho.2015.08.009] [Medline: 26344180]
27. Siette J, Ludlow K, Dodds L, Strutt P, Wuthrich V. "Provided a window on the world and lessened my feeling of isolation": older adults' perceived COVID-19 impact and technology use in Australia during recurrent lockdowns. BMC Geriatr 2024 Feb 28;24(1):206. [doi: 10.1186/s12877-024-04807-7] [Medline: 38419001]

XSL•FO
RenderX

28. Organisation for Economic Co-operation and Development. Spurring growth and closing gaps through digitalisation in a post-COVID world: policies to LIFT all boats. : Organisation for Economic Co-Operation and Development (OECD); 2021 Nov URL: https://www.oecd.org/en/publications/spurring-growth-and-closing-gaps-through-digitalisation-in-a-post-covid-world-policies-to-lift-all-boats_b9622a7a-en.html [accessed 2025-09-17] [doi: 10.1787/b9622a7a-en]
29. Fang H, Wan X, Zheng S, Meng L. The spillover effect of autonomy frustration on human motivation and its electrophysiological representation. Front Hum Neurosci 2020;14:134. [doi: 10.3389/fnhum.2020.00134] [Medline: 32390813]
30. Hertzog C, Dunlosky J. Metacognition in later adulthood: spared monitoring can benefit older adults' self-regulation. Curr Dir Psychol Sci 2011 Jun;20(3):167-173. [doi: 10.1177/0963721411409026] [Medline: 24478539]
31. Li Y, Tang J, Ma X, Zhang X, Xue Y, Zhao X. Impact of metacognition on health-related behavior: a mediation model study. J Environ Public Health 2023;2023:6547804. [doi: 10.1155/2023/6547804] [Medline: 36761241]
32. Chen X, Hertzog C, Park DC. Cognitive predictors of everyday problem solving across the lifespan. Gerontology 2017;63(4):372-384. [doi: 10.1159/000459622] [Medline: 28273664]
33. Craik FI, memory J. The Handbook of Aging and Cognition: Lawrence Eribaum Associates; 1992:51-110. [doi: 10.4324/9780203837665]
34. Park DC, Lautenschlager G, Hedden T, Davidson NS, Smith AD, Smith PK. Models of visuospatial and verbal memory across the adult life span. Psychol Aging 2002 Jun;17(2):299-320. [doi: 10.1037/0882-7974.17.2.299] [Medline: 12061414]
35. National seniors Australia. Seniors top scammers' hit list. 2024 Jul 14. URL: https://nationalseniors.com.au/news/latest-news/seniors-top-scammers-hit-list [accessed 2025-02-20]
36. National council on aging. Top 5 Scams Targeting Seniors. 2025 Jul 14. URL: https://www.ncoa.org/article/top-5-financial-scams-targeting-older-adults/ [accessed 2025-02-20]
37. Bryant C, Bei B, Gilson K, Komiti A, Jackson H, Judd F. The relationship between attitudes to aging and physical and mental health in older adults. Int Psychogeriatr 2012 Oct;24(10):1674-1683. [doi: 10.1017/S1041610212000774] [Medline: 22647646]
38. Mock SE, Eibach RP. Aging attitudes moderate the effect of subjective age on psychological well-being: evidence from a 10-year longitudinal study. Psychol Aging 2011 Dec;26(4):979-986. [doi: 10.1037/a0023877] [Medline: 21728444]
39. Bryant C, Bei B, Gilson KM, Komiti A, Jackson H, Judd F. Antecedents of attitudes to aging: a study of the roles of personality and well-being. Gerontologist 2016 Apr;56(2):256-265. [doi: 10.1093/geront/gnu041] [Medline: 24793646]
40. Sánchez Palacios C, Trianes Torres MV, Blanca Mena MJ. Negative aging stereotypes and their relation with psychosocial variables in the elderly population. Arch Gerontol Geriatr 2009;48(3):385-390. [doi: 10.1016/j.archger.2008.03.007] [Medline: 18448179]
41. Levy BR, Myers LM. Preventive health behaviors influenced by self-perceptions of aging. Prev Med 2004 Sep;39(3):625-629. [doi: 10.1016/j.ypmed.2004.02.029] [Medline: 15313104]
42. Pedroso-Chaparro MDS, Márquez-González M, Vara-García C, et al. Guilt for perceiving oneself as a burden in healthy older adults. Associated factors. Aging Ment Health 2021 Dec;25(12):2330-2336. [doi: 10.1080/13607863.2020.1822291] [Medline: 32964745]
43. Mizrachi Y, Shahrabani S, Nachmani M, Hornik A. Obstacles to using online health services among adults age 50 and up and the role of family support in overcoming them. Isr J Health Policy Res 2020 Aug 21;9(1):42. [doi: 10.1186/s13584-020-00398-x] [Medline: 32825840]
44. Takase A, Matoba Y, Taga T, Ito K, Okamura T. Middle-aged and older people with urgent, unaware, and unmet mental health care needs: practitioners' viewpoints from outside the formal mental health care system. BMC Health Serv Res 2022 Nov 23;22(1):1400. [doi: 10.1186/s12913-022-08838-x] [Medline: 36419047]
45. Australian Government Department of Health D and A. Aged care data and digital strategy for older people, their families and carers. Australian Government Department of Health, Disability and Ageing. URL: https://www.health.gov.au/our-work/aged-care-data-and-digital-strategy/for-older-people [accessed 2025-02-14]
46. Meylan P, Alam M. Digital inclusion for all: ensuring access for older adults in the digital age. FP Analytics. 2023. URL: https://www.aarpinternational.org/file%20library/resources/2023-a-fpa-aarp-digital-inclusion-final.pdf [accessed 2025-01-11]
47. Baraković S, Baraković Husić J, van Hoof J, et al. Quality of life framework for personalised ageing: a systematic review of ICT solutions. Int J Environ Res Public Health 2020 Apr 24;17(8):2940. [doi: 10.3390/ijerph17082940] [Medline: 32344521]
48. Peek STM, Wouters EJM, van Hoof J, Luijkx KG, Boeije HR, Vrijhoef HJM. Factors influencing acceptance of technology for aging in place: a systematic review. Int J Med Inform 2014 Apr;83(4):235-248. [doi: 10.1016/j.ijmedinf.2014.01.004]
49. Nielsen TL, Holst-Stensborg HW, Nielsen LM. Strengthening problem-solving skills through occupational therapy to improve older adults' occupational performance - a systematic review. Scand J Occup Ther 2023 Jan;30(1):1-13. [doi: 10.1080/11038128.2022.2112281] [Medline: 35995214]
50. Taherkhani Z, Kaveh MH, Mani A, Ghahremani L, Khademi K. The effect of positive thinking on resilience and life satisfaction of older adults: a randomized controlled trial. Sci Rep 2023 Mar 1;13(1):3478. [doi: 10.1038/s41598-023-30684-y] [Medline: 36859479]

XSL•FO
RenderX

51. Harris MT, Blocker KA, Rogers WA. Older adults and smart technology: facilitators and barriers to use. Front Comput Sci 2022;4. [doi: 10.3389/fcomp.2022.835927]
52. Tirado-Morueta R, Rodríguez-Martín A, Álvarez-Arregui E, Ortíz-Sobrino M, Aguaded-Gómez JI. The digital inclusion of older people in Spain: technological support services for seniors as predictor. Ageing Soc 2023 Jun;43(6):1409-1435. [doi: 10.1017/S0144686X21001173]
53. Wang Y, Wu Z, Duan L, et al. Digital exclusion and cognitive impairment in older people: findings from five longitudinal studies. BMC Geriatr 2024;24(1):406. [doi: 10.1186/s12877-024-05026-w]

Abbreviations

AI: artificial intelligence
BCC: Brisbane City Council
SOC: selective optimization with compensation

Edited by F Puga; submitted 26.03.25; peer-reviewed by A Adekola, AD Eaton, SB Guo; revised version received 13.07.25; accepted 15.08.25; published 19.09.25.

Please cite as:
Koh WQ, Ludlow K, Liddle J, Pachana NA
Selection, Optimization, and Compensation Strategies Used by Older Adults to Live Well With Technology: Qualitative Study
JMIR Aging 2025;8:e75019
URL: https://aging.jmir.org/2025/1/e75019
doi:10.2196/75019

© Wei Qi Koh, Kristiana Ludlow, Jacki Liddle, Nancy A Pachana. Originally published in JMIR Aging (https://aging.jmir.org), 19.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Free-Text Responses in a Nationally Representative Experimental Survey about End-of-Life Care Choices: ChatGPT-4o-Assisted Qualitative Analytical Study

Elizabeth M Goldberg[1], MD, ScM; Mario Macis[2], PhD; Megan Bounds[1], MPH; Jonathan Gomez Picazo[1], BS; Lauren Hersch Nicholas[3,4], PhD, MPP

[1]Department of Emergency Medicine, CU Anschutz, 12631 East 17th Avenue 2nd Floor, Room 2519, Aurora, CO, United States

[2]Department of Economics and Management, Carey Business School, Johns Hopkins University, Baltimore, MD, United States

[3]Division of Geriatrics, Department of Medicine, CU Anschutz, Aurora, CO, United States

[4]Center for Bioethics and Humanities, CU Anschutz, Aurora, CO, United States

Corresponding Author:
Elizabeth M Goldberg, MD, ScM
Department of Emergency Medicine, CU Anschutz, 12631 East 17th Avenue 2nd Floor, Room 2519, Aurora, CO, United States

Abstract

Background: Little is known about how surrogates make end-of-life care choices for patients who lack the ability to make decisions for themselves.

Objective: The study aims (1) to identify key themes that emerged from participants' free-text responses to a large nationally representative vignette survey about surrogate decision-making in end-of-life care and (2) to determine if an advanced artificial intelligence (AI) chatbot could assist us in accurately and efficiently performing qualitative analyses.

Methods: Our dataset included 3931 free-text responses from a nationally representative survey of 6109 individuals. In this qualitative study, we first familiarized ourselves with the free-text responses and hand-coded the first 200 responses until we reached saturation. We then created a codebook, initial themes, subthemes, and illustrative quotes. Subsequently, we prompted ChatGPT-4o to analyze the entire dataset of 3931 responses and identify frequent keywords and generate themes and quotable quotes. We validated responses by comparing the AI's keyword counts to qualitative software (NVivo, Lumivero) counts and cross-validating AI-generated quotes with the original transcripts.

Results: We identified several key themes: surrogates more often chose comfort care for care recipients with dementia, particularly at advanced stages. They also strongly weighed the patients' perceived quality of life and functional status. Many reported making surrogate decisions based on their own lived experiences or values, rather than making decisions aligned with the patients' previously stated wishes. There was no significant difference between the AI and qualitative software's keyword counts. The most frequent keywords included "life" (2051/81,713, 2.51%), "quality" (903/81,713, 1.11%), and dementia (507/81,713, 0.62%). Overall, AI-generated themes closely aligned with aforementioned human-generated themes. Manual coding of the first 200 free-text responses required 4 hours, including codebook development. In contrast, ChatGPT-4o generated themes in <10 seconds using the predefined codebook. However, dataset preparation, output verification, iterative prompting, debugging, and validation required several weeks.

Conclusions: Surrogates often base end-of-life decisions on dementia stage, perceived quality of life, and their own lived experiences, rather than patient preferences. Using an AI chatbot to perform qualitative analysis on free-text responses may help extend the work of qualitatively trained investigators, especially for large datasets such as free-text responses to large surveys.

(JMIR Aging 2025;8:e76335) doi:10.2196/76335

KEYWORDS

Chatbot; geriatrics; comfort care; life-extending care; qualitative research; survey

Introduction

Background

Many chronically ill older adults recognize the importance of discussing their end-of-life desires with their surrogates, family, friends, and clinicians, in the event they are unable to communicate during an emergency. Having advanced care directives reduces surrogate stress by avoiding the urgent decision-making process during emergencies and potential undesired treatment. It also may increase the likelihood that patients receive the care they desire, particularly in end-of-life situations where they may be unable to express their wishes [1].

XSL•FO
RenderX

Despite the increased use of advanced care directives, there are still incongruent care preferences between patients and their surrogates [2].

Gaps and Preliminary Studies

This discordance in end-of-life decision-making suggests that better facilitation and understanding of end-of-life decision-making are needed [3]. Previous studies have focused on the effectiveness of advanced care directives, caregiver communication, or physician influence only [1,4-6]. However, the end-of-life decision-making process is complex, and individuals must consider multiple factors when deciding. Our team created a vignette-based cross-randomized survey to better understand surrogate factors in making end-of-life decisions [7]. Quantitative results suggest that surrogates were less likely to recommend life-sustaining treatment for persons living with dementia, that surrogates were more likely to ignore patient preferences for life-extending treatments when the person had dementia, and that surrogates were more likely to choose treatments that matched their own preferences when patients' wishes were unclear.

Objectives

In this study, our *objective* was to analyze free-text responses to the survey question "What were the main considerations that led you to give the answers you provided in the module above (which captured your end-of-life care choice for your care recipient)?" Given this survey included 6109 individuals, we decided to leverage artificial intelligence (AI) to assist our team in performing a qualitative analysis. Alternative existing approaches to analyzing such a large volume of data, including qualitative thematic analysis, have drawbacks including the need to select and analyze a smaller subset of responses to avoid the extensive resource use needed to perform traditional qualitative analysis [8]. Here, we share the findings of our qualitative analysis of free-text responses relating to end-of-life care decision making and detail how we used an AI-generative chatbot, ChatGPT-4o, to facilitate the analysis.

Methods

Data Sources

Overview

We analyzed free-text responses from a nationally representative, web-based experimental survey that included 6109 adults aged ≥18 years [7]. Thirty-five percent of survey respondents did not provide a free-text response to the question soliciting a rationale for the end-of-life care choice made in the vignette and thus were not included in this analysis. This survey was developed by co-authors LHN and MM. Both are PhD health economists and professors with experience studying the interaction between health care use and economic outcomes through combining survey, administrative, and clinical data. No relationship was established between the research team and participants before study commencement. A professional survey company, Ipsos Knowledge Network survey platform, recruited participants between December 8, 2022, and December 19, 2022, and ensured complete anonymity to the researchers, reducing social-desirability bias and other experimenter-demand concerns [9]. Surveys were available in English and Spanish, and participants were provided with devices to complete the survey if they lacked home computers or internet. This vignette-style survey was intended to systematically investigate surrogate decision-making in the context of dementia-related cognitive impairment and end-of-life care for older adults [7]. Further details on the type of data collected, timeline of data collection, and procedures are detailed in our previously published paper [7].

The objective of this paper was to analyze all the free-text responses qualitatively that were provided to the question "What were the main considerations that led you to give the answers you provided (Comfort Care or Life-extending Care) in the module above (which captured your end-of-life care choice for your care recipient)?" We followed the Standards for Reporting Qualitative Research (SRQR) guidelines (Checklist 1) in reporting this work [10]. Our dataset did not include any personal health information as this was an anonymous survey to a nationally representative group. Thus, we did not provide any personal health information to ChatGPT. While ChatGPT was used as a tool to perform efficient coding and theme generation, it was not used for writing assistance.

Summary

We performed several steps to conduct this qualitative analysis of free-text responses including (1) following traditional thematic analysis procedures to code and generate themes on a subset of free-text responses [11], (2) prompting ChatGPT-4o to use our codebook to generate its own themes and quotable quotes, and (3) debugging and validating the AI responses. We first familiarized ourselves with the data by reading and rereading the first 200 responses to the free-text question. Then, we hand-coded the first 200 responses, generating a preliminary code book and iteratively improving it through team discussion. Next, we created a list of preliminary themes and illustrative quotes following traditional thematic analysis procedures.

After providing the data and our codebook to ChatGPT-4o (GPT-4 variant, May 13, 2024; used via Google Chrome browser), we prompted it to generate relevant keywords, keyword counts, themes, and illustrative quotes. We performed several steps to validate the AI-generated results, including comparing AI-generated themes and quotes with themes and quotes generated by qualitatively trained human professionals and comparing keyword counts with counts generated by a commonly used qualitative analytic software, NVivo (Lumivero) [12]. Finally, we put the findings into the context of the current literature.

Phase 1: Generating a Valid Approach

Prior to initiating the analysis, we conducted a literature review to identify how AI had been used in previous qualitative research. In particular, our methods were shaped by studies by Nashwan and Abukhadijah [13] and Van Veen et al [14], who provided prompt examples. Only limited studies have used AI approaches for analyzing free-text responses, so we also compared ideas on how to validate our AI findings with other academic teams doing similar work. These discussions yielded several suggestions, including comparing themes side by side

XSL•FO
RenderX

between AI and human-generated themes and doing a keyword count comparison with commonly used qualitative software.

Phase 2: Familiarization With the Data and Hand Coding

First, a physician scientist with experience caring for patients at the end of life and graduate-level qualitative training (EMG) and two other qualitatively trained research staff members (MB and JGP) read the first 200 participant responses. Then, 2 team members (EMG and MB) inductively coded the responses to the first 200 responses together. At this point, we reached thematic saturation, whereby no new significant themes continued to emerge, and concluded manual coding [15,16]. During this session, the team members created a preliminary coding schema in which we identified 6 major codes and several subcodes relating to end-of-life decision-making (Table S1 in Multimedia Appendix 1). The codebook was then iteratively refined through group discussion. Then, a third team member independently coded the same 200 responses (JGP). Finally, the principal investigator of the study (LHN), who designed the survey and had performed a preliminary human qualitative analysis after reading a sample of the responses, provided additional suggestions on the codes from her analysis, which we incorporated into the final codebook after team discussion.

Phase 3: Human-Identified Themes, Subthemes, and Illustrative Quotes

Once we felt confident that our codebook reflected the topics and scope of the data from the first 200 respondents, we generated themes and quotes that were illustrative of the major and minor themes. Investigator triangulation (multiple investigators with multiple areas of expertise) was used to establish the trustworthiness of our findings [17]. We compiled a table to display themes and quotes.

Phase 4: AI Analyses

Next, we used ChatGPT-4o to extend our human-generated analysis by providing it with the entire dataset of 3931 responders to the free-text question. We also piloted other generative AI tools (Microsoft Copilot and ChatGPT-3) but found that ChatGPT-4o provided the most usable responses and allowed us to upload larger amounts of data. Before entering the dataset in ChatGPT, we cleaned the data by removing all nonresponses. Free text containing responses of “none,” “n/a,” “na,” “no,” “nothing,” “no answer,” or “…,” were removed from the dataset submitted to ChatGPT to limit error. These responses were instead coded as “0” to indicate no free text was available. Data were uploaded via a Microsoft Word document containing a table with headings labeled “Free Text Responses,” “Age,” “Race/Ethnicity,” and “Gender.” ChatGPT processed the entire document at one time and required no chunking of information submission. No application programming interface or interface was used.

Our initial prompt to ChatGPT provided important context on the task, including that we planned to perform a qualitative analysis of free-text responses and that the AI should act as a qualitative analyst using our codebook and data to identify themes and quotable quotes. Although we considered having ChatGPT-4o inductively code the responses without human input, we anticipated that this would require much more investment by our team on the back end to iteratively prompt and refine themes. We also wanted to avoid the risk of irrelevant or erroneous themes or excessive AI “creativity.” We also prompted it to find the most common keywords in the text and generate keyword counts.

To illustrate how we completed this analysis, so it can be replicated, we share several sample prompts and how we validated responses:

1. The first prompt set up the research study context, its methods, and the AI’s role. An example prompt is: “This is a large survey study focused on end-of-life decisions, and you are a qualitative researcher analyzing the free-text responses to these questions.”
2. Next, we asked ChatGPT-4o to code the data using our codebook. A sample prompt includes: “This is a preliminary coding schema for our research. Using this codebook, code the entire sample. How did you do this?”
3. We asked ChatGPT-4o to create themes and subthemes and identify quotable quotes or key phrases that “make an impression on people from the text.” We then manually searched the dataset in NVivo to ensure the quote was verbatim and not fabricated.

Then, we asked the AI to create keywords identified in the entire dataset. We first quantified keywords by querying NVivo using the “word frequency/query” function. NVivo created a list of the most frequently used words in the dataset with the exception of proper names. NVivo provided the keyword, its frequency in the dataset, and the weighted percentage out of all words in the dataset. Then, we provided each keyword to ChatGPT-4o in addition to supplying it with a data file which included the full dataset of responses. We then queried ChatGPT-4o for exact keyword counts and weighted percentages, for example, “How often is the word ‘life’ found in this dataset, and what is the weighted percentage of ‘life’ among all words in the dataset?” by asking it to find how many times the specific keyword could be identified in the dataset.

Ethical Considerations

The study was found to be exempt by the Johns Hopkins Homewood institutional review board (HIRB00012400), and we followed SRQR guidelines in reporting this work. Our dataset did not include any personal health information as this was an anonymous survey of a nationally representative group. Thus, we did not provide any personal health information to ChatGPT. The survey participants received compensation for their participation; the amount is unknown and proprietary due to a lump sum getting paid to Ipsos, the survey company.

Results

Overview

Free-text responses were created by 3931 (65%) surveyed participants. Free-text responders were significantly older, had higher educational attainment, were less likely to be working full time, and were more likely to be a caregiver for another adult than nonresponders. They were also significantly more likely to be married, have higher income, and carry a diagnosis

XSL•FO
RenderX

of a chronic illness. Over half of free-text responders were people with chronic illnesses. A total of 1438 (37%) of them identified as caregivers, of which 601 (42%) stated that they provided ≥20 hours of care each week (Table 1). No significant differences were noted for sex, race, or region of the country.

XSL•FO
RenderX

Table . Participant demographics by free-text responder status.

Characteristics	Answered free-text question	
	No answers (n=2088)	Answer provided (n=3931)
Age (y), mean (SD)	51.19 (16.92)	55.08 (16.78)
Sex, n (%)		
Male	1016 (48.70)	1898 (48.28)
Female	1072 (51.34)	2033 (51.72)
Race, n (%)		
Black, non-Hispanic	229 (11.0)	400 (10.2)
Hispanic	103 (4.9)	148 (3.8)
Other, non-Hispanic	238 (11.4)	479 (12.2)
White, non-Hispanic	1449 (69.40)	2787 (70.90)
2+races, non-Hispanic	69 (3)	117 (3)
Education, n (%)		
No high school diploma or GED[a]	149 (7.1)	183 (4.7)
High school graduate (high school diploma)	652 (31.2)	963 (24.5)
Some college or associate degree	585 (28)	1099 (27.96)
Bachelor’s degree	392 (18.8)	882 (22.4)
Master’s degree or higher	310 (14.8)	804 (20.5)
Employment status, n (%)		
Working full time	980 (46.9)	1698 (43.20)
Working part time	277 (13.3)	553 (14.1)
Not working	831 (3.8)	1680 (42.74)
Marital status, n (%)		
Now married	1240 (59.39)	2418 (61.51)
Widowed	93 (4)	198 (5)
Divorced	207 (9.9)	424 (10.8)
Separated	31 (1.5)	63 (1.6)
Never married	517 (24.8)	828 (21.1)
Household income (US $), n (%)		
<10,000	98 (4.7)	135 (3.4)
10,000 to 24,999	205 (9.8)	319 (8.1)
25,000 to 49,999	371 (17.8)	645 (16.4)
50,000 to 74,999	345 (16.5)	657 (16.7)
75,000 to 99,999	247 (11.8)	501 (12.7)
100,000 to 149,999	384 (18.4)	702 (17.9)
≥150,000	438 (21)	972 (24.7)
Region of residence, n (%)		
Northeast	363 (17.4)	697 (17.7)
Midwest	486 (23.3)	873 (22.2)
South	787 (37.7)	1493 (37.98)
West	452 (21.6)	868 (22.1)
Caregiver for adult family member or friend, n (%)		
No	1504 (72.03)	2493 (63.4)

XSL•FO
RenderX

Characteristics	Answered free-text question	
	No answers (n=2088)	Answer provided (n=3931)
Yes	579 (27.8)	1438 (36.58)
Weekly hours spent caring for another adult, n (%)		
0-5	123 (21.8)	300 (20.9)
5-9	111 (19.6)	240 (16.8)
10-14	72 (13)	203 (14.2)
15-19	28 (5)	88 (6)
≥20	231 (40.9)	601 (15.3)
Chronic illness		
No	1066 (51.05)	1751 (44.54)
Yes	1022 (48.95)	2180 (55.46)

[a]GED: General Educational Development.

Human coding identified six major themes: (1) rationales provided for making end-of-life decisions, (2) the role of physical function in decision-making surrounding end-of-life care, (3) the role of dementia in decision-making surrounding end-of-life care, (4) impact on caregivers, (5) quality of life, and (6) basing decision on personal preferences or "I would" statements. Subthemes identified by human coding included: expected trajectory, drawing on individual experiences to decide, the caregiver feels a duty to choose to extend life, patient autonomy or preferences, ethics, emotional strain on family, limited quality of life with dependence on others, and living a life with value (Table 2).

XSL•FO
RenderX

Table . Human-derived major themes, subthemes, and quotable quotes.

Major theme and subtheme	Quotable quotes
Rationales provided for making end-of-life decisions	
Expected trajectory	• *If there is no brain activity found then and no family to contest then make patient comfortable. Many things to be considered in all cases based on doctor advice and family concerns. All cases and circumstances are not the same!* [White non-Hispanic male with chronic disease, 72 y] • *Mr. Jones is not able to eat, speak and move around. He would be in bed for the rest of his life, and that's not how he should live his life. It's very hard to be living under his conditions.* [Hispanic female adult, 54 y]
Drawing on personal experiences to decide	• *My own mother has middle to late stage Alzheimer's and I know her quality of life is already so impaired that it's no way to live. She looks at a newspaper and asks if it contains the "numbers and directions for her"—she has trouble recognizing what a fork is. Alzheimer's is heartbreaking and undignified for her. She is no longer who she was and does not know her own family. This is why I would choose for not invasive methods.* [White non-Hispanic female adult, 57 y]
A caregiver feels a duty to choose to extend life	• *Even if the person can be kept alive, it's not really living if the person is kept alive only to ease the guilt of a caregiver or significant other. I think the caregiver feels guilty if they do not try everything possible to extend life.* [White non-Hispanic female adult, 72 y]
Deciding based on patient autonomy and preferences	• *She had decided before she became ill what she wanted so family wouldn't have to decide that for her. I truly believe we should follow the patient's wishes if written.* [Black non-Hispanic male caregiver, 73 y] • *If a patient has a do not resuscitate order it should be followed, if not measures to sustain life [should] be used.* [White non-Hispanic male with chronic disease, 72 y] • *Having had both parents in similar situations when in their 80s. Neither had an advanced directive, but me and sibs were clearly told for years from each that they did not want to live longer if bedridden.* [White, non-Hispanic male with chronic disease, 67 y]
Role of physical function in decision making surrounding end-of-life care	• *Life has value, Mr. Smith has limited mobility but otherwise is fully aware of his own existence. I believe Mr. Smith wants to live!* [White non-Hispanic male adult, 61 y]

XSL•FO
RenderX

Major theme and subtheme	Quotable quotes
Role of dementia in decision making surrounding end-of-life care	• *Mrs. Jones will never get better with the dementia. She will slowly lose her ability to talk and eat and the dementia will turn into Alzheimer's disease, and she will pass away a horrible and painful death.* [White non-Hispanic female with chronic disease, 65 y] • *She has dementia that will never improve and there is not much life after that diagnosis. In short there is very little quality of life and it will only get worse.* [White non-Hispanic male adult, 56 y] • *He has dementia already and extending his life would not make his life any better!!* [White non-Hispanic female with chronic disease, 70 y]
Impact on caregivers	
Emotional strain on family	• *I choose not to put the family through the pain of prolonging life that may not be beneficial and could hurt them more in the long run.* [Black non-Hispanic female adult, 55 y] • *I know of this happening to a friend's parent and it was an emotion[al] ride for both. Neither person had a good life until the parents passing.* [White non-Hispanic female adult, 67 y]
Quality of life	
Limited quality of life with dependence on others	• *Quality of life is extremely important to me. Being dependent on others is not a good quality of life.* [White non-Hispanic female with chronic disease, 75 y] • *If he is going to be treated with dignity, and respect, maybe he would have a half-way decent quality of life. Even if he isn't aware of it. I would certainly hope so.* [White non-Hispanic female caregiver, 70 y]
Live a life with value	• *If you can't have quality of life your better off dying.* [White non-Hispanic female adult, 68 y]
Basing decision on personal preferences or "I would" statements	• *If I was confined to bed for the rest of my life, I would choose comfort care if the opportunity arose.* [White non-Hispanic male with chronic disease, 31 y] • *I put myself in those situations, primarily. Since I would not want unnecessary life extension, especially at 85+, I would not want this for anyone else. Interestingly, although I can fluctuate on my religious beliefs, the fact that I can consider an afterlife is comforting, and a reason to not artificially extend a few years here at the end.* [White non-Hispanic male caregiver, 68 y]

ChatGPT-4o identified 7 major themes from the free-text responses and also identified subthemes and representative quotes (Table 3). We further asked ChatGPT-4o to indicate the demographics of the responder (Table S2 in Multimedia Appendix 2) and indicate if the proxy was choosing comfort care, life-extending care, or both, which it performed accurately

XSL•FO
RenderX

(Table S3 in Multimedia Appendix 3). In Table S3 in Multimedia Appendix 3, we also categorized errors made by the AI, including fabrication of quotes and substantial modifications of quotes. To aid in the replication of this work, we summarized the additional steps taken to ensure rigor in our analytic approach in Multimedia Appendix 4.

Table . ChatGPT-4o-derived themes using the codebook generated by humans.

Major theme and subtheme	Quotable quotes
Respect for advanced directives	
Honoring patient preferences	*The caregivers should follow the advanced directives.* [White non-Hispanic female, 51 y]
Clear directives prior to illness	*She said prior she did not want to have the invasive treatments. It is honoring her wishes.* [White non-Hispanic female, 67 y]
Quality of life	
Independence versus dependence	*If no quality of life, why take drastic measures to extend it?* [White non-Hispanic male, 88 y]
Burden of dementia	
Enjoyment of life and meaningful existence	*Life in bed is not a good life. Waste of resources that could be used on others.* [White non-Hispanic male, 75 y]
Loss of memory and cognitive decline	*With dementia, he lives in a state of confusion and fear. This will continue to get worse and quality of life will worsen even more.* [White non-Hispanic female, 54 y]
Emotional and financial toll on caregivers	*Having a caregiver 24-7 would cause an undue burden on family. Some family members would not be able to handle watching a family member deteriorate in front of their eyes.* [Black non-Hispanic male, 60 y]
Caregiver burden	
Emotional hardship for families	*The dementia has taken over, there is no quality of life. How much time are you buying to keep him in a state of no memory?* [White non-Hispanic female, 66 y]
Financial strain of prolonged care	*My father and my mother-in-law both had dementia, and it was very difficult and expensive to keep them alive.* [White non-Hispanic male, 62 y]
Ethical and religious considerations	
Sanctity of life	*Life, to all extents, is precious, given by a power greater than the limitations of medical care.* [White non-Hispanic male, 81 y]
Beliefs about life and death	*We are not the ones who should take away life.* [White non-Hispanic male, 87 y]
Patient dignity	
Maintaining respect during end-of-life care	*You should be able to die with dignity. Laying in bed is not living, it is existing.* [White non-Hispanic male, 72 y]
Avoiding perceived indignities	*End it so your son can get on with his own life.* [White non-Hispanic female, 75 y]
Decision-making challenges	
Lack of clear information	*It's a tough call to make when all the information is just a few paragraphs.* [White non-Hispanic male, 56 y]
Emotional difficulty in making choices	*I am unable to respond because I don't believe I should decide on another person's life.* [White non-Hispanic female, 83 y]

XSL•FO
RenderX

Validation and Lessons Learned Using Generative AI

In Table 4, we depict a comparison of human- and ChatGPT-4o-generated major themes and subthemes. Several themes overlapped between human (black text) and AI (purple text) thematic analysis. While the AI did not always name the theme the same as our human analysts, the meaning was similar. Overall, there were several differences observed between human-generated and AI-generated tables, including that the AI closely relied on the codebook to generate themes, representative quotes were shorter, and more themes were generated by the AI. In addition, AI-generated themes were more descriptive—summarizing directly what participants said—rather than interpretative (eg, offering latent meaning or suggesting why participants may have said what was stated). ChatGPT created themes such as "quality of life" without extrapolating underlying motivations or cultural values, such as human-derived "drawing on individual experiences to decide."

Table . Comparison of human-derived and artificial intelligence–derived thematic analysis.

Themes	Human-generated analysis	ChatGPT-4o analysis
Quality of life	✓	✓
Rationales provided for making end-of-life choices	✓	
Expected trajectory	✓	
Drawing on individual experiences to decide	✓	
Caregiver feels duty to choose to extend life	✓	
Deciding based on patient autonomy and preferences (respect for advanced directives)	✓	✓
Role of physical function in decision making surrounding end-of-life care	✓	
Role of dementia in decision making surrounding end-of-life care (burden of dementia)	✓	✓
Impact on caregivers (caregiver burden)	✓	✓
Basing decision on personal preferences or "I would" statements	✓	
Decision-making challenges		✓
Ethical and religious considerations		✓
Patient dignity		✓

We note that the AI generated 3 themes not included in our human-generated analysis: decision-making challenges, ethical and religious considerations, and patient dignity. While codes relevant to these topics existed in the human codebook, they were not prioritized in the final theme set due to their relatively lower frequency in the subset of responses reviewed manually. It is possible that these themes appeared more frequently across the full dataset analyzed by ChatGPT, which highlights the potential for AI methods to surface lower prevalence but potentially meaningful perspectives that might be missed in human review.

During the validation of ChatGPT-4o-generated quotable quotes by human analysts, we identified 3 major errors. First, ChatGPT-4o would often combine statements from multiple participants into one quote. Second, it sometimes ignored the prompt to create verbatim quotes and instead modified the quote to be more concise or grammatically correct. Third, it sometimes fabricated quotes entirely (hallucination). For the first error, combining statements, we addressed the challenge by reformatting the data into a table rather than using Microsoft Word, so the AI could more easily identify when one quote ended and another started. This improved ChatGPT-4o's quotable quote reliability substantially. Our previous data format separated participant responses with spaces only. For the second error, paraphrasing, we reprompted ChatGPT-4o that exact quotes were needed, and they could not be modified. With reprompting, ChatGPT-4o was able to list verbatim quotable quotes. Below we display an example of the second error. While the meaning is the same, the content is paraphrased. For the third error, complete fabrication of quotes, we removed these from the tables and reprompted the AI to choose new verbatim quotes followed by verifying their existence in the dataset. We show more examples of all 3 types of errors in Table S3 in Multimedia Appendix 3.

Original quote by responder:

> *Life is precious and we should not try to end it of our own accord. It is God's role to extend or end a life.*

ChatGPT-4o-generated quote:

> *Life is precious and should be preserved as long as possible, according to God's will.*

In Table 5, we listed the 11 most common keywords identified by NVivo within the free-text responses and provided the counts and weighted percentages for keywords in NVivo and ChatGPT-4o. We chose the 11 most common keywords because starting at the 12th keyword, the frequency of the count was

XSL•FO
RenderX

much lower. We identified that NVivo and ChatGPT-4o were using different methodologies to provide keyword counts and weighted percentages. Initial prompts to ChatGPT-4o generated counts and weighted percentages that were lower than NVivo, suggesting undercounting. We show how we corrected the discrepancy with the following example. To correct the count for the keyword "life," we told ChatGPT-4o that NVivo had come up with a count of 2051 for the keyword "life" and asked it what it had done incorrectly, as its initial count for life was 1984. ChatGPT-4o responded that the NVivo count was correct, and what it was doing wrong was "The initial discrepancy was due to not properly accounting for cases where 'life' was followed by punctuation or appeared at the end of a sentence or line." We then asked ChatGPT-4o to repeat the experiment using the correct method. Counts were still discrepant after this correction, though the ChatGPT-4o counts were closer overall to the NVivo counts after this reprompting. We then discovered that NVivo excluded "stop words," which were considered insignificant. When we changed the NVivo settings to include the stop words, we then produced considerably closer weighted percentages, which are shared in Table 5.

Table . Keywords generated by NVivo compared to ChatGPT-4o-generated keywords and their frequencies.

Keyword	NVivo keyword count (weighted percentage), n (%)	ChatGPT-4o keyword count (weighted percentage), n (%)
Life	2051 (2.51)	2051 (2.48)
Quality	903 (1.11)	904 (1.09)
Care	551 (0.68)	551 (0.67)
Dementia	507 (0.62)	509 (0.61)
Want	481 (0.59)	481 (0.58)
Able	366 (0.45)	366 (0.44)
Live	366 (0.45)	366 (0.44)
Wishes	351 (0.43)	351 (0.42)
Family	346 (0.42)	346 (0.42)
Still	321 (0.39)	321 (0.39)
Think	255 (0.31)	255 (0.31)

Time Considerations and Replication

Human coding was time-intensive; our team spent 4 hours coding the first 200 responses to the free-text question, recoding the first 200 responses, and creating and revising the initial codebook. ChatGPT-4o took <10 seconds to generate themes using our codebook. However, preparing the dataset for ChatGPT-4o, verifying the content, subsequent reprompting and verification, identifying methods for debugging, and validating took several weeks.

Due to ChatGPT-4o being a large language model (LLM), replication of similar outputs using the same prompt and dataset was challenging. We often received varying answers to the same command using ChatGPT-4o, although the overall major themes identified within the dataset remained relatively consistent. Once we reformatted responses into a table from the initial word file, ChatGPT-4o had more consistent response outputs to the same prompt.

Discussion

Principal Findings

In this ChatGPT-4o-assisted qualitative analysis, we identified several themes related to how surrogates make end-of-life care choices for patients. Proxies' decisions are informed by prior caregivers' prior lived experiences, expectations of quality of life, how advanced the stage of dementia is in the patient, and also personal values, including religious beliefs. In performing this analysis, we also learned several lessons about using AI tools to extend the capacity of our qualitative team. We learned that ChatGPT-4o is capable of using a human-created codebook to search free-text responses to create themes, representative quotes, and keywords. In addition, we learned that ChatGPT-4o-generated themes are primarily descriptive and closely match codes identified by humans. We also found that validation by humans is necessary, and ChatGPT-4o can sometimes provide insights on errors and correct them when prompted to do so.

This unique cross-randomized experiment allowed our team to provide new insights into what role dementia plays in surrogate decision-making. We found that many surrogates consider the stage of dementia to make end-of-life decisions. Caregivers may correctly identify that dementia is incurable and that extending life when dementia is advanced may not allow the individual to live the way they once did, free of the support of others. Others identified that quality of life in dementia may be tied to the ability to recognize and socialize with others, and when recognition diminishes, extending life may not be desired. Prior research also found that some surrogates equate dementia with loss of the "self" and view recognition as a key marker of dignity and quality of life [18]. Some felt extending life was warranted when the patient previously requested it or because of their religious beliefs.

Another common theme was drawing on individual experiences or preferences to make decisions. A key finding in the

XSL•FO
RenderX

quantitative survey and the AI qualitative analysis was that when preferences were not clear, surrogates made choices that aligned with their own wishes. Personal beliefs about how one would feel if faced with these life choices were weighed in making the decision. However, ethically, proxies should make choices in line with patient preferences and not their own [19]. In some cases, proxies chose comfort care after projecting what *their* future self may feel like when aged or if *they* had the functional status of the patient described in the vignette. These survey and qualitative findings could inform clinical practice in several important ways: (1) clinicians should feel empowered to inform proxies that decisions should be made that are in line with patient preferences and (2) education should be provided to proxies about advanced dementia and its impact on quality of life, as well as the impact life-extending care may have on patient comfort and well-being.

Other literature on end-of-life choices and surrogate choices have found that while 90% of caregivers acknowledge that dementia is incurable, only 40% see it as life-limiting [20]. One study indicates that caregivers of people living with dementia who have chronic illnesses and who discuss goals of care with clinicians are more likely to have accurate prognostic understanding on the end-of-life care for persons living with dementia [21]. In addition, caregivers with better prognostic understanding were more likely to state a preference for comfort-focused care, and their care recipients were less likely to receive burdensome interventions and have greater comfort during the dying process. Studies have also recognized racial differences in advanced care planning, with African American persons reporting lower rates of advanced care planning completion (89% vs 73%) and lower preference for comfort care [22]. Several studies indicate that more education of proxies about the stages of dementia and end-of-life care options would be helpful to improve the quality of proxy decision-making [21,22].

The use of natural language processing AI-augmented qualitative analysis has been described as the "Way of the Future" [23]; however, few investigators have used it to perform qualitative analysis [24]. One team used structured topic modeling coupled with thematic analysis to analyze free-text data from nearly 38,000 English patrons with the aim to improve the user experience of public-facing services in the United Kingdom [25]. This team tested models with 5 to 40 topics and differing covariates and then had 2 human coders conduct thematic analysis to interpret the topics. In contrast with this approach, our approach did not require team members to have requisite skills in machine-learning methods and thus may be more accessible for qualitatively trained teams. Other AI approaches to analyzing large quantities of free text have included using natural language processing and machine learning models to develop robust classification models [26], but this approach also requires data science or related technical expertise.

Other teams have used ChatGPT-4o and more accessible LLMs to analyze qualitative data [23,27,28]. Experiences of these investigators can be summarized as overall positive in that these LLMs were found to be capable of producing satisfactory analyses that come close to human-generated analyses, but the performance of different LLMs varied, was not predictable, and serial prompting was necessary. Many have stated that themes are more descriptive than interpretive, yet it was easy to use and required less effort than manual coding. We also found that AI themes often did not go beyond what was explicitly stated by the participant and failed to explore latent meaning or social context. This could pose a barrier to using these tools in research where interpretative themes are essential. In addition, researchers found that even when using the same prompts and dataset, slightly different output was generated as AI generative chatbots "learn" over time. While this could improve performance, it also introduces problems with reproducibility. Thus, our team found it important to carefully log all prompts and responses in a log. In contrast to our approach, Carvalho et al [27] used a "series of chain-of-thought prompts to dig deeper into the provided response." This approach had 2 advantages. It allowed the investigators to guide the chatbot to provide them with more positive themes when negative ones were initially suggested, and it surfaced less commonly encountered themes, allowing for more diversity of perspectives. Carvalho et al [27] also found that the paid version of ChatGPT-4o, which our team also used, generated the most accurate responses compared to the unpaid version and Google Gemini. This was particularly true when it was well-prompted and after multiple repetitions were performed in the same analysis.

Because there is currently no established gold standard for using advanced chatbots for qualitative analysis, our team made several deliberate methodological decisions. These included providing the AI a human-generated codebook rather than allowing it to inductively code responses and generate themes "independently." Other teams may choose to allow the AI full discretion in its coding, and this approach could be more efficient and lead to more novel themes. We made this choice because our team had ample qualitative expertise, which allowed for us to code a subset of responses and ensure the codebook reflected key domains of interest. However, this approach may have reduced the AI's ability to discover novel themes and constrained "creativity."

Barriers to the use of AI in qualitative research include concerns over data privacy, a need for upskilling in AI among trained qualitative investigators, and a lack of established approaches to use the tool [24]. Information such as transcripts uploaded to AI platforms is used to develop the underlying algorithms that allow the computer to learn, which some academics feel is problematic from a data privacy standpoint. It is important for academics to follow institutional best practices such as obtaining approval from their ethical review boards prior to engaging in this work. Some institutions may advise using enterprise AI solutions or other tools that allow investigators to select that inputted data will not be externally shared or used to improve the AI's algorithm. Another barrier has been that senior investigators trained in qualitative methods may not possess the skills to use AI for qualitative analysis or teach others how to use it. Similarly, a roadmap is not currently available that outlines how rigorous qualitative analysis can be done with AI assistance. We believe that AI will be increasingly used to analyze large qualitative datasets, and as its use and popularity grow, investigators must learn through trial and error how to best use it and debug it.

XSL•FO
RenderX

Limitations

Free-text responses are an important source of data in surveys that collect quantitative measures because they can provide context to answer choices; however, unlike semistructured interviews, there is no opportunity to prompt the interviewee to elaborate on points or explain their statement when it is not clear. We found sentence fragments or short replies that made it difficult to understand what the surrogate's perspective was and could be interpreted differently depending on the person reading the text. Although our research team has a multidisciplinary background, we recognize that we bring our own bias to our work, and this may have affected our interpretation of our findings. We made efforts to increase the validity of ChatGPT-4o's analysis by asking it to show its work, comparing its output with existing qualitative software, and completing our own a priori human analysis on a subsample of the qualitative data, but the use of this tool for qualitative analysis is still new, and thus, there may be errors that we did not recognize.

Conclusions

Participants in a vignette-style nationally representative survey asked to act as proxies provided several different rationales for why they chose to extend care or focus on comfort care in their hypothetical care recipients. Choices were often made based on whether the care recipient had dementia and how advanced it was, perceived quality of life, or the caregivers' lived experiences or values, rather than making decisions aligned with the care recipients' previously stated wishes. This suggests that it is in an individual's best interest to share their care preferences with their intended proxy and engage in advance care planning prior to losing health care decision-making capacity.

Using an AI tool to perform qualitative analysis on free-text responses may help extend the work of qualitatively trained investigators, especially for large datasets such as free-text responses to large surveys. Our advice to investigators who aim to use ChatGPT-4o or other AI generative chatbots is to first complete a human-generated analysis using traditional qualitative coding, provide a codebook to the AI, anticipate the need for multiple prompts, try different LLMs, plan validation approaches and time for debugging, and keep a prompt and response log.

Acknowledgments

During the preparation of this work, the authors tested ChatGPT-4o's ability to qualitatively analyze the free-text responses for end-of-life choices. After using this tool or service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication. This project received funding from the National Institute on Aging (R01AG059205).

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

EMG and LHN conceived of the paper. LHN and MM obtained funding, designed**,** curated, and quantitatively analyzed the survey data. EMG, MB, and JGP qualitatively analyzed the free-text responses and wrote the original manuscript. EMG, MB, and JGP executed the methods and validated the analysis. EMG supervised the investigation and administered the project. All authors had full access to all of the data and contributed to the methodological design, reviewed, and edited the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Preliminary coding scheme.
[DOCX File, 15 KB - aging_v8i1e76335_app1.docx]

Multimedia Appendix 2
The themes, quotes, and demographic information produced by ChatGPT-4o after prompting demonstrate its ability to identify and match demographic information.
[DOCX File, 16 KB - aging_v8i1e76335_app2.docx]

Multimedia Appendix 3
Initial artificial intelligence (AI)-generated themes by type of care (comfort care vs life-extending vs both) and human verification of AI-generated representative quotes with error type.
[DOCX File, 19 KB - aging_v8i1e76335_app3.docx]

XSL•FO
RenderX

Multimedia Appendix 4
Additional method material to aid in replication of the study.
[DOCX File, 25 KB - aging_v8i1e76335_app4.docx]

Checklist 1
SRQR checklist.
[DOCX File, 17 KB - aging_v8i1e76335_app5.docx]

References

1. Dixon J, Karagiannidou M, Knapp M. The effectiveness of advance care planning in improving end-of-life outcomes for people with dementia and their carers: a systematic review and critical discussion. J Pain Symptom Manag 2018 Jan;55(1):132-150. [doi: 10.1016/j.jpainsymman.2017.04.009] [Medline: 28827062]
2. Fried TR, Zenoni M, Iannone L, O'Leary JR. Assessment of surrogates' knowledge of patients' treatment goals and confidence in their ability to make surrogate treatment decisions. JAMA Intern Med 2019 Feb 1;179(2):267-268. [doi: 10.1001/jamainternmed.2018.5299] [Medline: 30477019]
3. Ramplin C. Establishing a structured plan to provide high-quality end-of-life care in community settings. Br J Community Nurs 2019 Mar 2;24(3):120-127. [doi: 10.12968/bjcn.2019.24.3.120] [Medline: 30817209]
4. Chi HL, Cataldo J, Ho EY, Rehm RS. Please ask gently: using culturally targeted communication strategies to initiate end-of-life care discussions with older Chinese Americans. Am J Hosp Palliat Care 2018 Oct;35(10):1265-1272. [doi: 10.1177/1049909118760310] [Medline: 29514489]
5. Pocock LV, Wye L, French LRM, Purdy S. Barriers to GPs identifying patients at the end-of-life and discussions about their care: a qualitative study. Fam Pract 2019 Oct 8;36(5):639-643. [doi: 10.1093/fampra/cmy135] [Medline: 30649266]
6. Ramsaroop SD, Reid MC, Adelman RD. Completing an advance directive in the primary care setting: what do we need for success? J Am Geriatr Soc 2007 Feb;55(2):277-283. [doi: 10.1111/j.1532-5415.2007.01065.x] [Medline: 17302667]
7. Nicholas LH, Langa KM, Halpern SD, Macis M. How do surrogates make treatment decisions for patients with dementia: an experimental survey study. Health Econ 2024 Jun;33(6):1211-1228. [doi: 10.1002/hec.4810] [Medline: 38358920]
8. Lennon R, Calo W, Miller E, Zgierska A, Van Scoy L, Fraleigh R. Using artificial intelligence to support rapid, mixed-methods analysis: developing an automated qualitative assistant (AQUA). Ann Fam Med 2022 Apr 1;20(20 Suppl 1):20. [doi: 10.1370/afm.20.s1.2893] [Medline: 36944057]
9. Kuziemko I, Norton MI, Saez E, Stantcheva S. How elastic are preferences for redistribution? Evidence from randomized survey experiments. Am Econ Rev 2015 Apr 1;105(4):1478-1508. [doi: 10.1257/aer.20130360]
10. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251. [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
11. Braun V, Clarke V. Thematic analysis. In: Cooper HM, Camic PM, Long DL, Panter AT, Rindskopf D, Sher KJ, editors. APA Handbook of Research Methods in Psychology, Vol 2: Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological: American Psychological Association; 2012:57-71. [doi: 10.1037/13620-004]
12. NVivo. Lumivero. 2018 Oct 10. URL: https://lumivero.com/products/nvivo/ [accessed 2025-10-23]
13. Nashwan AJ, Abukhadijah H. Harnessing artificial intelligence for qualitative and mixed methods in nursing research. Cureus 2023;15(11):e48570. [doi: 10.7759/cureus.48570]
14. Van Veen D, Van Uden C, Blankemeier L, et al. Clinical text summarization: adapting large language models can outperform human experts. Res Sq 2023 Oct 30:rs.3.rs-3483777. [doi: 10.21203/rs.3.rs-3483777/v1] [Medline: 37961377]
15. Tran VT, Porcher R, Tran VC, Ravaud P. Predicting data saturation in qualitative surveys with mathematical models from ecological research. J Clin Epidemiol 2017 Feb;82:71-78. [doi: 10.1016/j.jclinepi.2016.10.001] [Medline: 27789316]
16. Guest G, Bunce A, Johnson L. How many interviews are enough?: An experiment with data saturation and variability. Field methods 2006;18(1):59-82. [doi: 10.1177/1525822X05279903]
17. Denzin NK. The Research Act: A Theoretical Introduction to Sociological Methods: Routledge; 2009. URL: https://www.routledge.com/The-Research-Act-A-Theoretical-Introduction-to-Sociological-Methods/Denzin/p/book/9780202362489 [accessed 2025-10-23]
18. Cruise CE, Lashewicz BM. Dementia and dignity of identity: a qualitative evidence synthesis. Dementia (London) 2022 May;21(4):1233-1249. [doi: 10.1177/14713012211072929] [Medline: 35232314]
19. Schuklenk U. New frontiers in end-of-life ethics (and policy): scope, advance directives and conscientious objection. Bioethics 2017 Jul;31(6):422-423. [doi: 10.1111/bioe.12372] [Medline: 28608972]
20. Gabbard J, Johnson D, Russell G, et al. Prognostic awareness, disease and palliative understanding among caregivers of patients with dementia. Am J Hosp Palliat Care 2020 Sep;37(9):683-691. [doi: 10.1177/1049909119895497] [Medline: 31854201]
21. Balasubramanian I, Andres EB, Poco LC, Malhotra C. Prognostic understanding among caregivers of persons with dementia: a scoping review. J Am Geriatr Soc 2025 Mar;73(3):947-955. [doi: 10.1111/jgs.19245] [Medline: 39446015]

XSL•FO
RenderX

22. Pettigrew C, Brichko R, Black B, et al. Attitudes toward advance care planning among persons with dementia and their caregivers. Int Psychogeriatr 2020 May;32(5):585-599. [doi: 10.1017/S1041610219000784] [Medline: 31309906]
23. Hitch D. Artificial intelligence augmented qualitative analysis: the way of the future? Qual Health Res 2024 Jun;34(7):595-606. [doi: 10.1177/10497323231217392] [Medline: 38064244]
24. Morgan DL. Exploring the use of artificial intelligence for qualitative data analysis: the case of ChatGPT. Int J Qual Methods 2023 Oct;22. [doi: 10.1177/16094069231211248]
25. Bondaronek P, Papakonstantinou T, Stefanidou C, Chadborn T. User feedback on the NHS test & trace service during COVID-19: the use of machine learning to analyse free-text data from 37,914 England adults. Public Health Pract (Oxf) 2023 Dec;6:100401. [doi: 10.1016/j.puhip.2023.100401] [Medline: 38099087]
26. Alexander G, Bahja M, Butt GF. Automating large-scale health care service feedback analysis: sentiment analysis and topic modeling study. JMIR Med Inform 2022 Apr 11;10(4):e29385. [doi: 10.2196/29385] [Medline: 35404254]
27. Carvalho T, Negm H, El-Geneidy A. A comparison of the results from artificial intelligence–based and human-based transport-related thematic analysis. Findings 2024. [doi: 10.32866/001c.94401]
28. Lee VV, van der Lubbe SCC, Goh LH, Valderas JM. Harnessing ChatGPT for thematic analysis: are we ready? J Med Internet Res 2024 May 31;26:e54974. [doi: 10.2196/54974] [Medline: 38819896]

Abbreviations

AI: artificial intelligence
LLM: large language model
SRQR: Standards for Reporting Qualitative Research

Edited by Y Jiang; submitted 21.04.25; peer-reviewed by A Uchenna, T Amusa; revised version received 10.09.25; accepted 22.09.25; published 29.10.25.

Please cite as:
Goldberg EM, Macis M, Bounds M, Picazo JG, Nicholas LH
Free-Text Responses in a Nationally Representative Experimental Survey about End-of-Life Care Choices: ChatGPT-4o-Assisted Qualitative Analytical Study
JMIR Aging 2025;8:e76335
URL: https://aging.jmir.org/2025/1/e76335
doi:10.2196/76335

© Elizabeth M Goldberg, Mario Macis, Megan Bounds, Jonathan Gomez Picazo, Lauren Hersch Nicholas. Originally published in JMIR Aging (https://aging.jmir.org), 29.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX